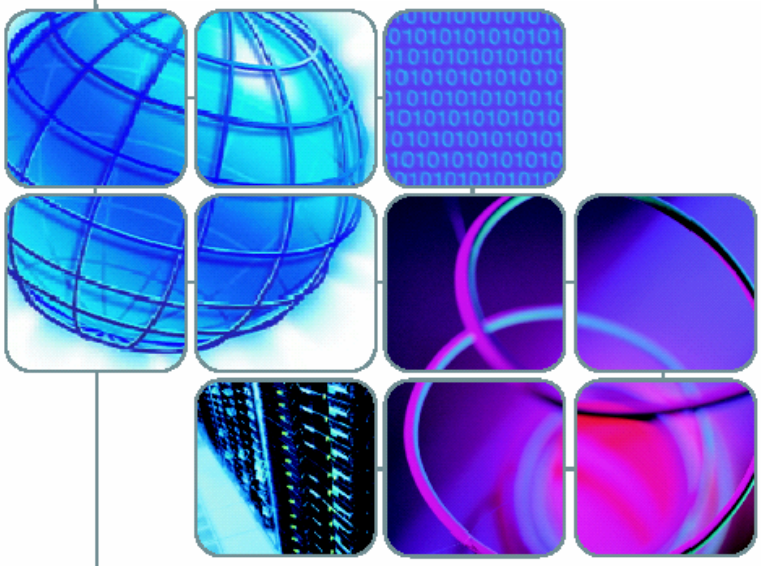
AdvancedTCA Modular Communications Platform Design Guide*

Guidelines for Modular Communications Platform Building-Block and System Interoperability

Version 2.0 Release 1.0 October 29, 2004



hp invent
intel
Avocent OSATechnologies
RadiSys
Marconi
continuous COMPUTING CORPORATION
ELMA Your Solution Partner
Mellanox Technologies
Schroff
IP Fabrics
PERFORMANCE TECHNOLOGIES
kontron
SIEMENS
Littelfuse
ADLINK TECHNOLOGY INC.
MOTOROLA
SBS Technologies
Diversified Technology An Ergon Co.
ADVANTECH Network Computing
CARLO GAVAZZI
Pigeon Point Systems

Contents

Figures

Tables

1 Introduction

Modular Communications Platforms (MCP) are industry standards-based communications infrastructure platforms and building blocks that enable efficiencies through the entire value chain including solution flexibility, faster time-to-market, vendor choice, and cost benefits. Through a growing ecosystem of standards-based suppliers, Modular Communications Platforms provide Network Equipment Providers (NEPs) with reusable development and deployment platforms, and an avenue to revive innovative product design in a climate of severe resource restraint.

Advanced Telecommunications Computing Architecture or AdvancedTCA® is a new series of industry-standard specifications for the next generation of carrier grade communications equipment. These specifications, driven by over 100 companies within the PCI Industrial Computers Manufacturing Group (PICMG*), incorporate the latest trends in high speed interconnect technologies, next generation processors, and improved reliability, manageability and serviceability, resulting in a new blade (board) and chassis (shelf) form factor optimized for communications. The principal objective of AdvancedTCA has been to provide standardized platform architecture for carrier-grade telecommunication applications, with support for carrier-grade features.

The primary benefits of this platform are:

- Improved price-to-performance ratio with significant cost savings
- Open architecture and standard interfaces that provide a choice of vendors at various levels of integration (in contrast to being captive to a single vendor solution)
- Robust hardware and software components that incorporate redundancy, failover, hardware failure prediction, and that prevent problems for end-users
- Customized system that is optimized to run vertical stacks of applications

This chapter introduces the audience and purpose of the AdvancedTCA Modular Communications Platform Design Guide. Additionally, this chapter summarizes the content of the guide and defines guideline compliance.

1.1 Audience for this Guide

This guide is for OEM hardware and software designers, architects and system integrators to develop, implement, and advantageously use systems built around the AdvancedTCA Modular Communications Platform for the telecommunications equipment market. This guide assumes that the reader understands the AdvancedTCA family of specifications.

1.2 Purpose of this Guide

The AdvancedTCA MCP Design Guide is intended as a companion to the AdvancedTCA specifications and serves as a tool to aid designers in developing a common, interoperable communications platform. This document contains a set of preferred implementation guidelines for promoting building block interoperability on the base AdvancedTCA specification. This Design Guide targets communications network elements.

The AdvancedTCA family of specifications has a number of options that enable the support of a large variety of applications. However, this could increase the likelihood that AdvancedTCA building blocks may not interoperate. Key success factors for wide scale deployments of a technology include; standardization and interoperability among various implementations. This document offers guidance to AdvancedTCA equipment manufacturers on elements/building blocks that are compliant with the specifications and also interoperable with components from other vendors.

The AdvancedTCA family of specifications does not define how to implement some features that are essential for carrier grade network elements, such as, network clocks, automatic protection, and shelf management. This guide provides guidelines on implementing some of these gaps in the specification. By

following the guidelines in this guide, vendors can ensure that their components will interoperate with other complying systems.

Telecommunications systems are currently dominated by proprietary hardware systems that also have a huge investment in software. For NEPs to take advantage of cheaper Commercial Off The Shelf (COTS) hardware, it is important that the underlying features and interfaces of the modular hardware facilitate an easy migration of the NEP's software onto modular hardware. This guide offers guidance to the NEPs on how to build systems adhering to the Design Guide.

1.3 Sections of this Guide

This second edition of the AdvancedTCA MCP Design Guide has been reorganized and expanded to provide clearer guidelines on how to build AdvancedTCA components and systems that are driven by the need for strong interoperability.

Chapter	Description
Section 1: Overview	
1 Introduction	Read this chapter for an introduction and an overview of the document.
2 Modular Communications Platform Architecture	Read this chapter for a high level architectural overview of the AdvancedTCA Modular Communications Platform (MCP). This is a general model that will be used as the framework to derive specific architectures for different Network Elements (NE).
Section 2: Building Blocks	
3 AdvancedTCA Overview	Read this chapter for an overview of the AdvancedTCA specifications.
4 MCP Backplane Fabric Topologies and Technologies	The AdvancedTCA family of specifications has a rich set of options and configurations for backplane fabric. Read this chapter to understand the fabric options that are recommended for building MCP systems.
5 The AdvancedTCA Shelf	Read this chapter to learn the essential requirements of an AdvancedTCA Shelf that allows it to be used by modules from multiple vendors. It is expected that the shelf will normally be sourced from a single vendor. Therefore it must be able to interoperate with multiple vendors' AdvancedTCA boards and Rear Transition Modules (RTMs).
6 AdvancedTCA Boards	Read this chapter for guidelines on building AdvancedTCA boards and the features that should be present in various board types.
7 Shelf Management	Read this chapter to understand the architecture of the shelf manager and the standard interface to the system manager. The interface between the shelf manager and the system manager is not covered in great detail in the AdvancedTCA specifications.
8 AdvancedTCA Stratum Clock Distribution	Read this chapter to understand the usage of the AdvancedTCA synchronization clock interfaces to properly support and distribute a Stratum 3/3E external reference within and between AdvancedTCA shelves.

Chapter	Description
9 AdvancedTCA Automatic Protection Switching	Read this chapter to understand the support for Automatic Protection Switching (APS) in SONET end systems such as Routers, Media Gateways, and Multi Service Switches.
10 Carrier Grade OS	Read this chapter for the requirements of a carrier grade operating system used in MCP systems.
Section 3: Certification and Design Practices	
11 Certification and Practices	Read this chapter to understand the certifications and practices that should be followed when developing AdvancedTCA MCPs.
12 Design Practices	Read this chapter to understand the common design practices that should be followed when building AdvancedTCA building blocks.
Appendices	
Appendix A: Requirement Naming Convention	Read this appendix to understand the two-letter naming convention used in the creation of the Requirement Ids in this document.
Appendix B: Carrier Grade OS	Read this appendix to understand the CG OS features.
Appendix C: Glossary	Read this appendix for the abbreviations and unique terms used in this document.

1.4 Guideline Compliance

To ensure they receive the benefits of standardization and interoperability, equipment purchasers will look for Modular Communications Platform Design Guide compliance as an attribute of modular communications hardware. The following sub-sections describe the levels of guideline compliance.

Individual vendors are responsible for ensuring their products meet the requirements of the design guide before identifying a product as Modular Platform Design Guide Version 2.0 Compliant. Modular platform adopters shall be provided with tools, such as test suites, that will aid in determining adherence to the requirements, as well as opportunities to test interoperability between products from multiple vendors.

To see the latest version of the AdvancedTCA MCP Design Guide, go to:

http://www.intel.com/go/atcadg

1.4.1 Required Compliance

If a feature is ***required,*** the feature must be implemented to comply with the guidelines. Other architectural layers assume the presence of a required feature.

Note: Some required features might be nested under parent features marked as recommended or optional. In this situation, the nested ***required*** feature must be implemented for compliance only if implementers include the recommended or optional feature. These nested features are marked as ***required if*** or ***recommended if***.

When a feature is defined as ***required***, the term ***shall*** is used in the wording of the requirement.

1.4.2 Recommended Compliance

If a feature is ***recommended*****,** the feature does not need to be implemented. Implementation is, however, encouraged. The presence of a recommended feature is not to be assumed nor precluded by other architectural levels.

An implementation, which does not include a ***recommended*** feature, must be prepared to operate with another implementation, which does include the feature, though perhaps with reduced functionality. Similarly, an implementation, which does include the feature, must be prepared to operate with another implementation, which does not include the feature (except, of course, for the functionality the feature provides).

When a feature is ***recommended,*** the term ***should*** is used in the wording of the requirement.

1.4.3 Optional Compliance

If a feature is ***optional***, the feature does not need to be implemented. There should be no dependencies on optional features by other architectural levels. Omitting an optional feature should have no impact on other required features or proper operation of the system.

Operating systems are not required to support optional features, but must operate in the presence or absence of those features.

When a feature is ***optional***, the term ***may*** is used in the wording of the requirement.

1.4.4 Tables of Requirements

This document lists requirements in a tabular format. Each table contains the following columns:

- ID: This column refers to an alphanumeric identifier. Each requirement has a unique identifier. The identifier comprises a 2-letter feature name followed by a sequential numbering system starting with 001. Refer to Appendix A: Requirement Naming Convention for a list of the 2-letter feature names.
- Feature Name: This column lists the feature to which this requirement corresponds.
- Requirements: This column contains the requirements to be met for required, recommended or optional compliance. Whether a requirement is required, recommended or optional can be determined by use of terminology in the requirement.

For example:

ID	Feature Name	Requirements
TC-001	Table Conventions	All requirements ***shall*** be listed in a table with associated ID, Feature Name and Requirements columns.

1.5 Terminology use

This document uses the terms *blades* and *boards* interchangeably. Use of the terms *MCP system* or *MCP component* indicate compliance to the requirements of the Design Guide.

2 Modular Communications Platform Architecture

This chapter provides a high-level architectural overview of the AdvancedTCA Modular Communications Platform (MCP). This is a general model that will be used as the framework to derive specific architectures for different Network Elements (NEs) focusing on converged communications and compute systems in carrier networks.

2.1 Architecture Overview

A Network Element (NE) is an entity that is configured to provide support for an application such as a multi-service switch, voice switch, edge router, Radio Network Controller (RNC) and so on in a carrier's communications network.

An NE is typically composed of one or more chassis, each of which in turn is composed of multiple blades for specific functionalities. The architecture is described in terms of such modular building blocks.

At the highest level, an MCP NE consists of the following elements:

- One or more Servers
- One or more Switches
- System (Server-Switch) Interconnects
- Network Management Server (external to the NE)

Figure 2-1 depicts a high-level block diagram of such a Network Element. Note that in all further discussions, a reference to a Server or a Switch indicates the logical function represented by each rather than to a physical entity.

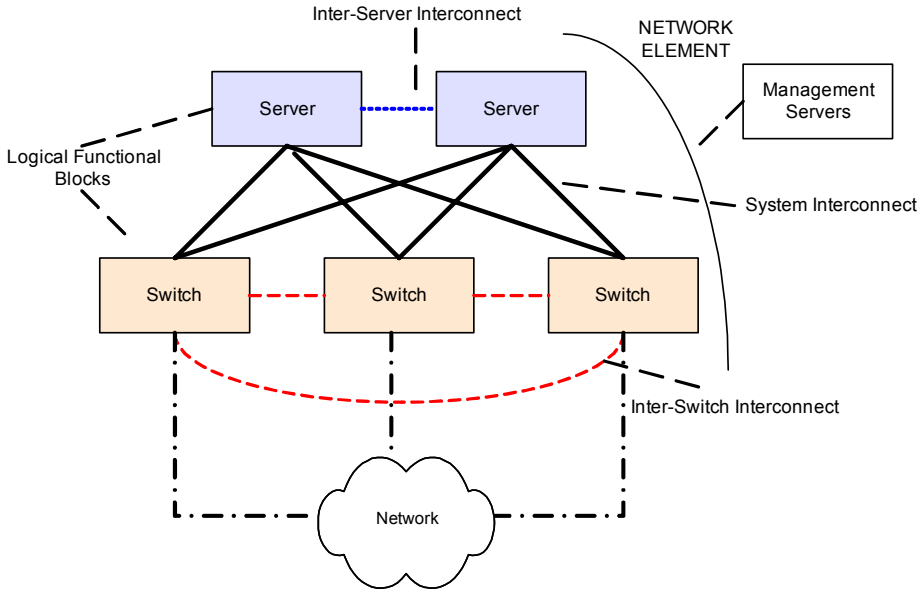


Figure 2-1 Network Element Structure

The following sub-sections discuss the functionality associated with each of the high-level building blocks and variations of their configurations within a Network Element.

There is substantial commonality across board types between network elements allowing large amounts of reuse. Additionally, customizing the boards using standard mezzanines based on the Advanced

Mezzanine Card form factor further reduces the number of total cards needed and allows for scalability over a wide performance range.

2.1.1 Server

A server (or Telecommunications Server) is the component of an NE that is primarily responsible for application-level software. The application-level software depends on the type of NE. It may support a softswitch for VOIP networks, Home Location Register (HLR), the Mobile Switching Center server including the Visiting Location Register (VLR), billing, provisioning, SLA management and so on. It generally interfaces with one or more switches and provides resource management and control capabilities across the components of the Network Element.

A server provides services in the following three planes:

Plane	Services
Control plane	The Control Plane is responsible for the real-time control of the data plane. This includes signaling protocols that determine the structure of the user information data flow.
Management plane	The Management Plane supports the management of faults, configuration, accounting, performance and security (FCAPS). Configuration may include application software parameters, configuration of addresses (IP, SS7 Point Codes etc.), signaling and routing protocols, physical and service interfaces, redundancy etc. It also supports the human interface needed to manage the server.
Data plane	The (user) Data Plane is the functionality associated with the transfer of user or protocol data from an ingress port (or internal data source) to an egress port (or internal data sink). In a server, Protocol Processing Blades (compute cards) support application processing for the Data Plane functionality. The Data Plane traffic in this case is associated with the protocol stacks and application software. The Data Path requirements in terms of Quality of Service are generally not as stringent for a Server as in the case of a Switch.

Servers may also have storage resources. The various blades that constitute the server are connected by the control plane, data plane, and management plane interconnects.

2.1.2 Switch

A switch is the component of an NE that is primarily responsible for data (bearer) path processing. As in the case of a server, a switch also provides service in three planes:

Plane	Services
Control plane	The Control Plane is responsible for the real-time control of the data plane. This includes signaling protocols that determine the structure of the user information data flow.
Management plane	The Management Plane supports the management of faults, configuration, accounting, performance and security (FCAPS). Configuration may include application software parameters, configuration of addresses (IP, SS7 Point Codes etc.), signaling and routing protocols, physical and service interfaces, redundancy etc. It also supports the human interface needed to manage the switch.
Data plane	The Data Plane in a switch is also referred to as the User or Forwarding Plane. It differs from the Data Plane in a Server in that it typically supports

Plane	Services
	the transport of bearer/user data as opposed to the transport of protocol or application data. Line and Service Blades support the Data Plane functionality. The data path in a Switch comprises the physical interfaces, framers and network processors, DSPs etc. involved in packet and media processing including providing support for data path interworking.

Additionally, a Switch might have a separate Storage Card or may have some storage capabilities on the Control blade itself.

Depending on the physical configuration of the Server and the Switch, the three planes of functionality may be shared among the servers and switches or they may be independent. Refer to Section 2.1.4 for the possible physical configurations.

2.1.3 Interconnect

Name	Description
System Interconnect	The interconnect between the servers and the switches depends on their physical configuration. If the server and the switch are in distinct chassis, they may be connected via Ethernet with an application-defined protocol for communication. For example, in a media-gateway communicating with a server (softswitch), MGCP is typically used as the protocol for communication. If the server and switch are integrated within a chassis, the Control Plane Interconnect (and optionally the Data Plane Interconnect) functions as the interconnect between the two.
Inter-Server Interconnect	The interconnect between servers depends on their physical configuration. For example, if multiple servers are in distinct chassis, then they may be connected via Ethernet. If, on the other hand, they are integrated into a single chassis (that is, functionality is distributed over multiple server blades in the chassis), the Control Plane Interconnect and optionally the Data Plane Interconnect functions as the interface between them.
Inter-Switch Interconnect	The interconnect between switches is similar to that between servers. In addition, switches may be interconnected using user/subscriber interfaces. For example, an ATM over SONET trunk interface on a Line Card on each switch may be used to connect two switches, thus providing increased switch capacity and user ports across the two switches.

2.1.4 Physical Configuration

Multiple options exist for configuring Server and Switch Building Blocks comprising a NE. Although the physical configurations might vary, the functionality associated with a server and a switch remains unchanged.

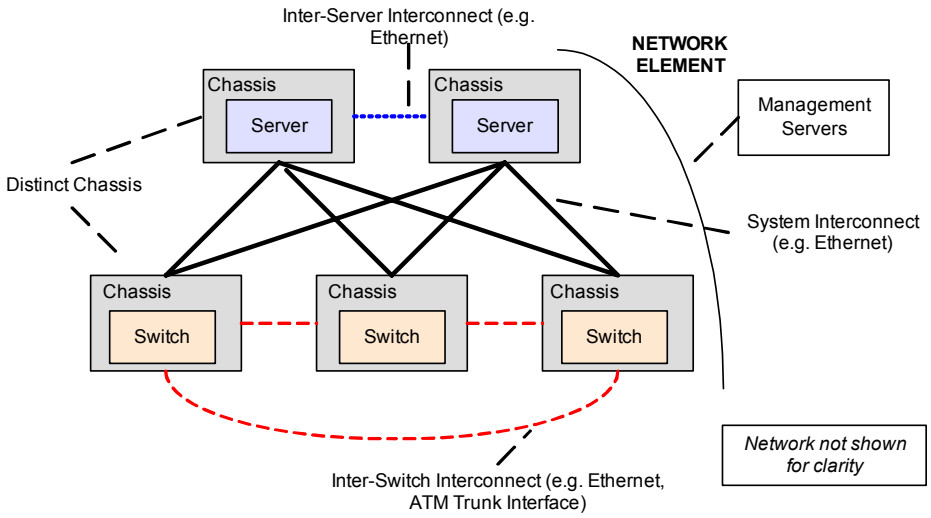


Figure 2-2 Network Element Configurations: Distinct Server and Switch Chassis

Figure 2-2 shows Server and Switch functionality in distinct chassis interconnected via the System Interconnect.

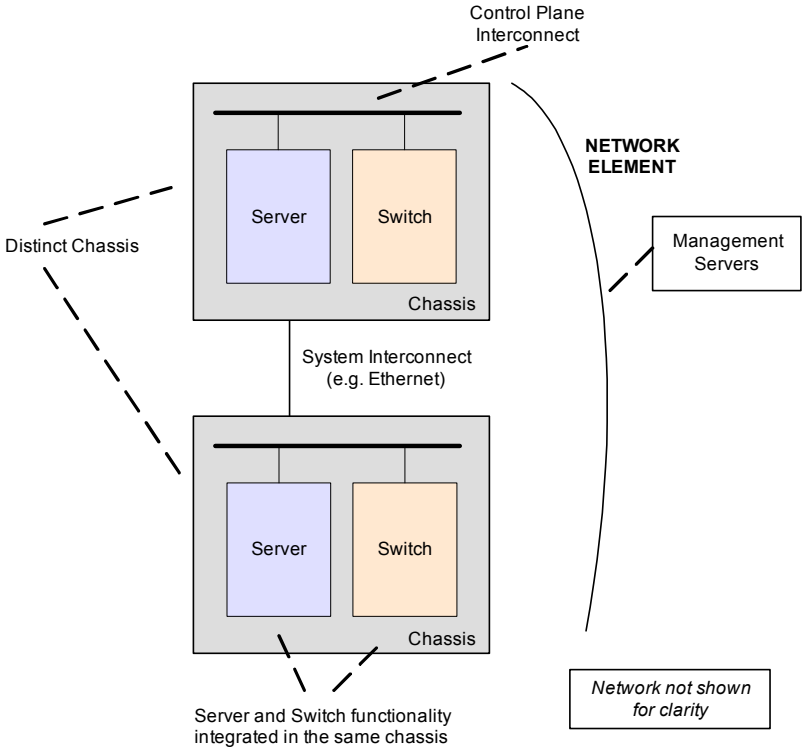


Figure 2-3 Network Element Configurations: Server and Switch functionality integrated into a Chassis

Figure 2-3 depicts Server and Switch functionality integrated into the same chassis, interconnected via the Control Plane Interconnect and optionally via the Data Plane Interconnect.

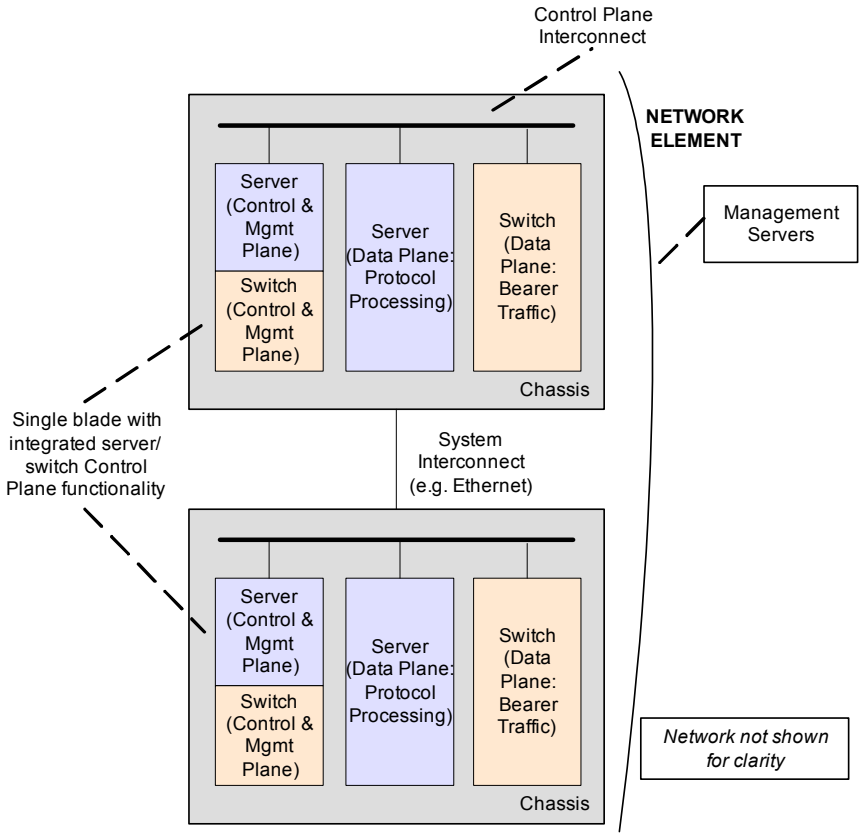


Figure 2-4 Network Element Configurations: Server and Switch Control Plane functionality integrated onto a single Control Blade in chassis

Figure 2-4 depicts the Control and Management Plane functionality of the Server and Switch integrated into one or more Control Blades in the same chassis.

2.2 Server Architecture

This section provides an overview of the architecture of a server. Figure 2-5 depicts the key components of a server.

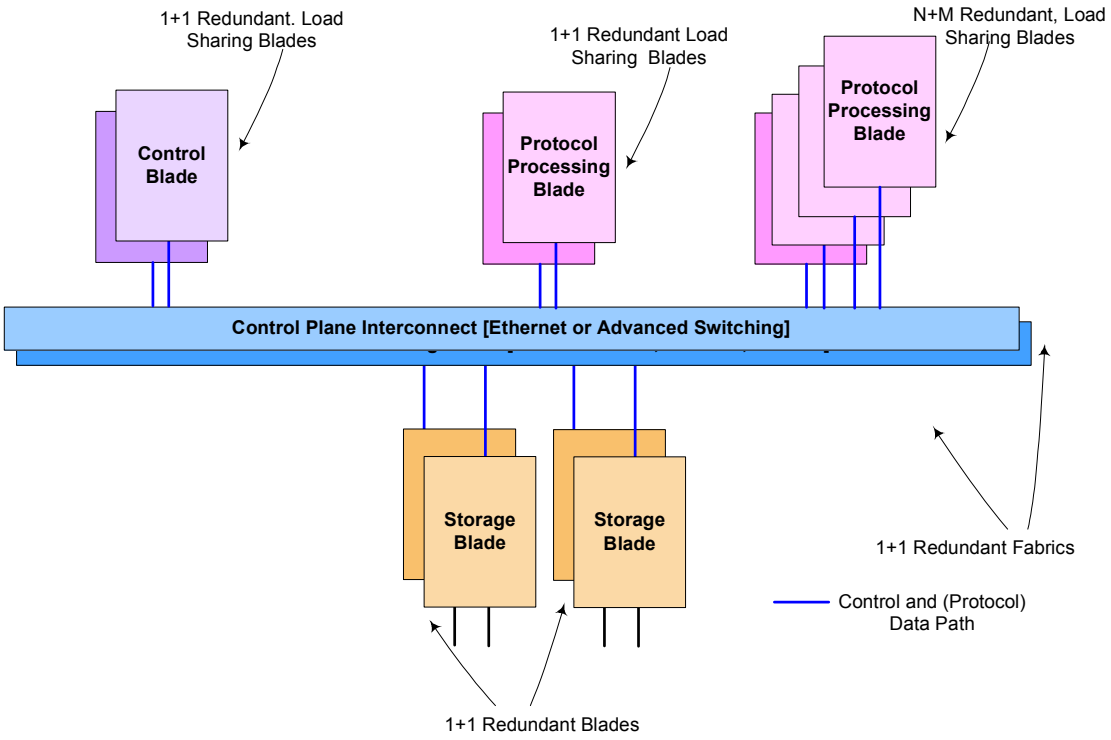


Figure 2-5 Server Architecture

The high-level hardware building blocks included in a Server are as follows:

Building Block	Description
Control blade	The Control Blade executes middleware software such as management of blades, event and alarm management, fault management, resource management etc. It also provides redundant access for external management, typically via Ethernet and additionally a serial port.
Protocol Processing blade	Protocol Processing Blade functionality is application dependent. In a softswitch, the blades may perform SS7 call processing, provisioning, billing etc. In an RNC, the blades implement the Transport Network Layer components, providing SCCP support using Broadband SS7 or SIGTRAN.
Storage blade	A server may comprise Storage Blades. If separate ones are not present, the Control Blade could include storage capabilities. The Protocol Processing Blades may also include some storage. For example, a Protocol Processing Blade involved in billing may store billing-related data.
Control plane interconnect	All of the blades are connected via the Control Plane Interconnect which may be Ethernet or Advanced Switching. Details of this interconnect and specific technologies and topologies supported by MCP platforms are in Chapter 4. Note that there is no direct correlation between the control

Building Block	Description
	plane interconnect and the AdvancedTCA base and fabric interfaces. The control plane interconnect may use either the AdvancedTCA base or fabric interfaces depending on the needs of the particular NE being considered.

Other than the components shown in Figure 2-5, all servers have power supplies, fans/ fan trays and management access. A server may optionally possess a Data Plane Interconnect.

The following sections expand the Server Building Blocks from a hardware and software perspective and identify the next level of modular building blocks.

2.2.1 Control Blade

A Control Blade, as depicted in Figure 2-6, consists of primarily one or more Control CPUs which support middleware functionalities such as resource management, alarm/event management etc. It may optionally have some storage such as a flash disk to store configuration information, logs, statistics etc. The Control Blade is connected to the Control Plane Interconnect via the Base or Backplane Fabric Interface. It may additionally interface with the Data Plane Interconnect also, if one exists in the server.

Additionally, in an AdvancedTCA chassis, there is a bus, referred to as the Intelligent Platform Management Bus (IPMB) over which the Intelligent Platform Management Interface (IPMI) protocol runs. IPMI provides support for autonomous monitoring, logging and recovery control functions (status of blades, ability to power-on/off blades etc.) independent on the main processors.

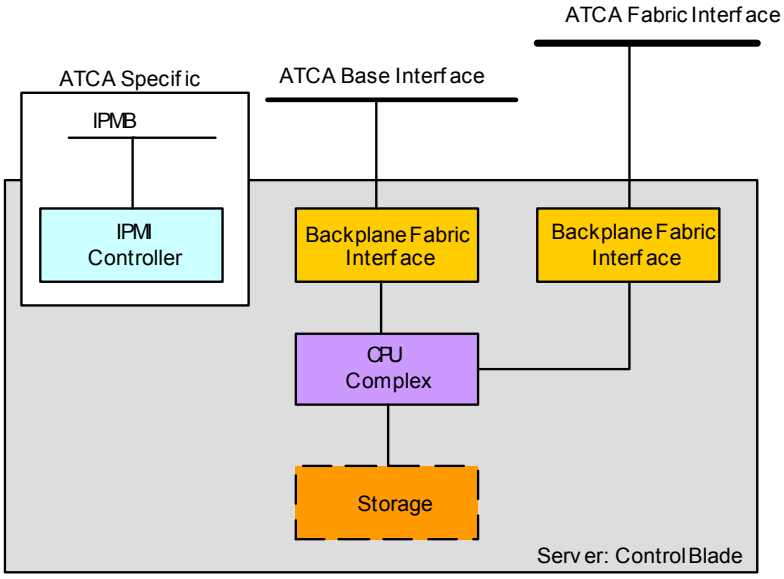


Figure 2-6 Control Blade

2.2.2 Protocol Processing Blade

A Protocol Processing Blade, as depicted in Figure 2-7 is similar to a Control Blade. The main difference lies in the software that executes on the Protocol Processing Blades. In most cases, the same hardware can be used for Control Blades and Protocol Processing Blades.

Protocol Processing Blades may also have access to line interfaces. For example, they may have T1/E1 interfaces if they are involved in SS7 processing. To retain a lot of commonality with the Control Blade, these interfaces may be on a mezzanine card.

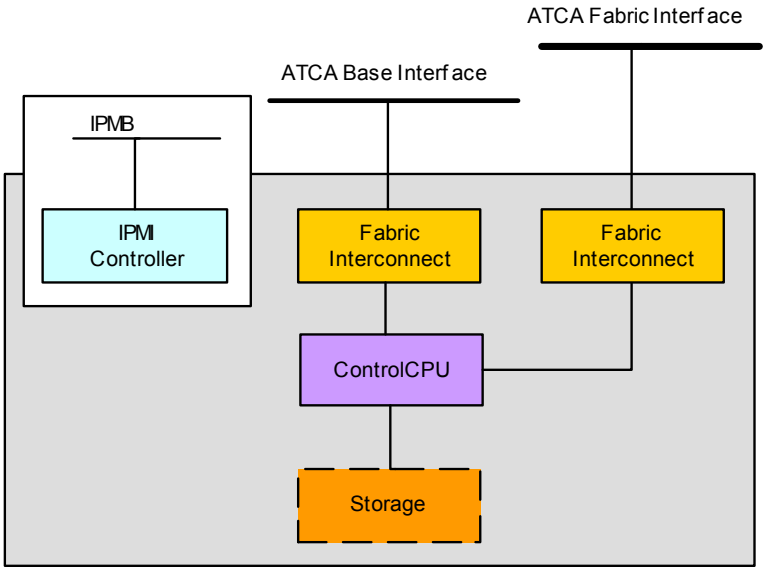


Figure 2-7 Protocol Processing Blade

2.2.3 Storage Blade

There are several options for Storage Blades:

- Internal Storage: Storage Controller and Storage are integrated on the same blade
- Internal Storage: Storage Controller and Storage are on distinct blades
- External Storage: Storage Controller Blade in chassis interfaces with external storage such as JBOD or RAID. This blade may also be used to control storage on a Storage Blade within the chassis.

In all cases, the components include one or more CPUs, a TOE/iSCSI device, a disk controller, and storage disks. The examples depicted here show the use of iSCSI. If the fabric interface is based on PICMG 3.1, Fibre Channel connections on the backplane are also allowed and will often be a simpler solution.

Figure 2-8 depicts a storage blade in which the storage controller and storage are integrated.

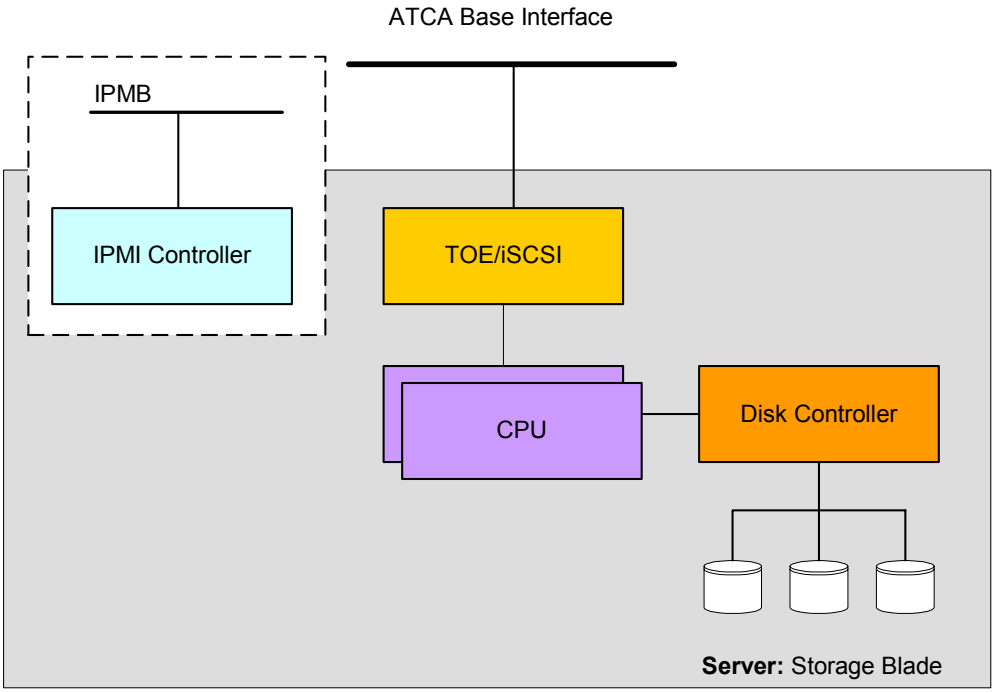


Figure 2-8 Internal Storage -- Storage Controller and Storage on same Blade

Figure 2-9 shows a configuration where the storage controller and storage elements are different blades.

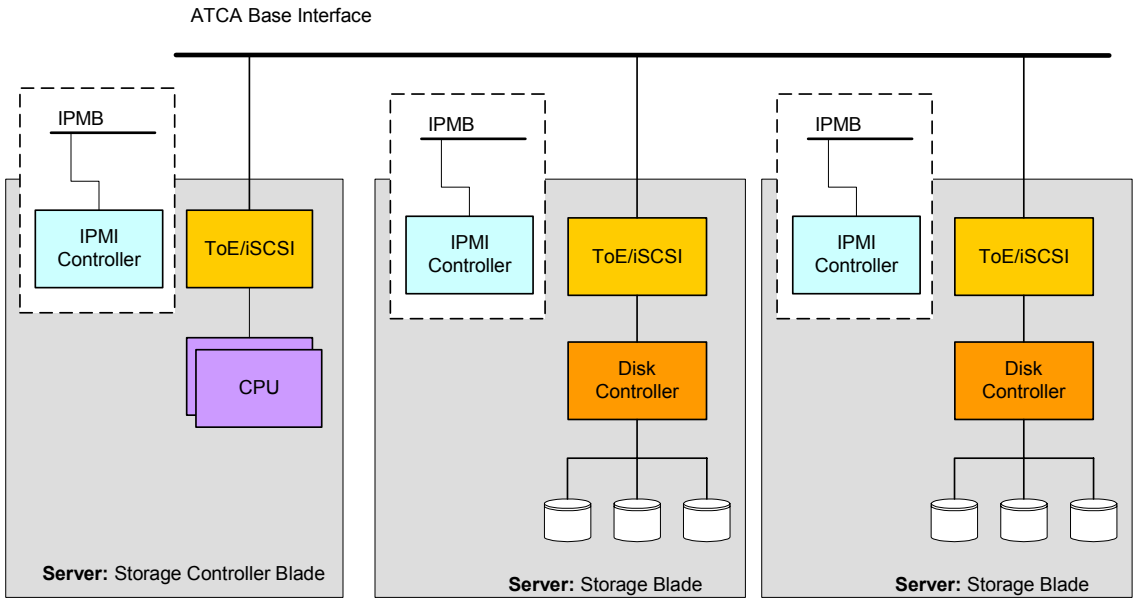


Figure 2-9 Internal Storage -- Storage Controller Blade and Storage Blades

Figure 2-10 depicts a configuration with external storage elements.

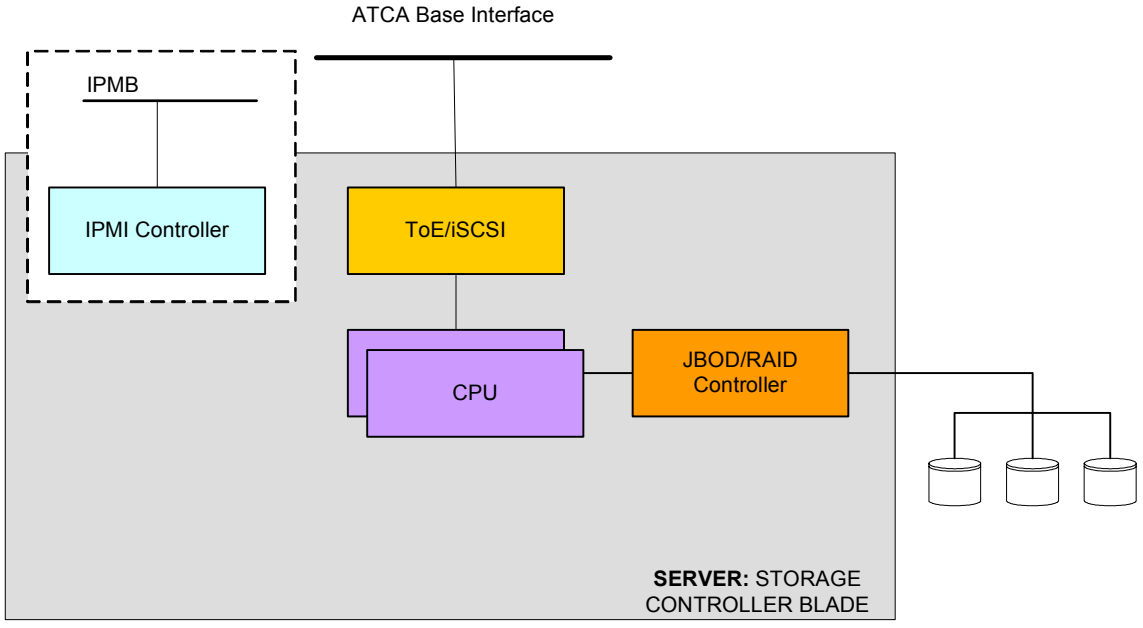


Figure 2-10 External Storage

2.3 Switch Architecture

This section provides an overview of the architecture of a switch. Figure 2-11 depicts the key components of a switch.

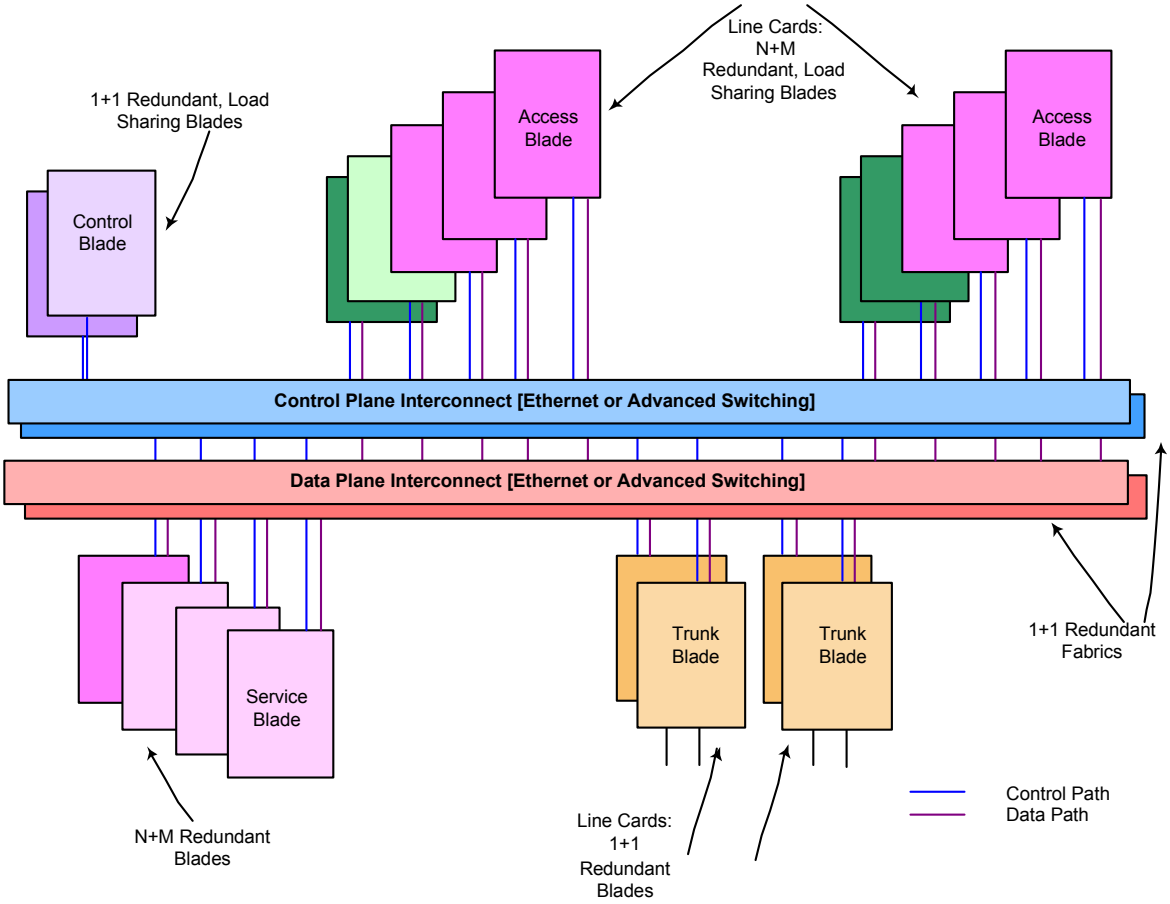


Figure 2-11 Switch Architecture

The high-level hardware building blocks included in a Switch are the following:

Building Block	Description
Control blade	The Control Blade in a switch is similar to the control blade in a server. It executes middleware software such as management of blades, event and alarm management, fault management, resource management etc. It also provides back-up access for external management, typically via Ethernet and a serial port. Additionally, it is likely to have some storage capabilities and a network clock source.
Line card	Line Cards provide data path support. They perform packet processing and interworking functions (that is, protocol conversion such as converting frame relay to ATM using FRF.5 or converting TDM streams to ATM Circuit Emulation Services) on data entering/ leaving the switch. They comprise physical interfaces on which data is received or transmitted, framers to process the physical layer format of the data, and one or more

Building Block	Description
	network processing units to perform packet processing functions such as classification, policing, shaping etc. There are two classes of Line Cards; they differ mainly in the interface capacity and to some extent the packet processing functionality. However, the hardware building blocks are, similar. • Access Blade: Interfaces on these blades typically connect to subscriber devices such as IADs, PBXs, and so on • Trunk Blade: High-capacity interfaces on these blades typically connect to core networks such as the PSTN, IP, or ATM backbone networks. Network elements at the edge of the network will have both access and trunk line cards. Network elements in the core of the network will typically only have what we term here as trunk line cards.
Service blade	Service Blades support the data path by providing additional resources that are computation-intensive. For example, a Service Blade might support tone detection, echo cancellation, voice compression/ decompression etc. Data requiring additional processing is routed from an Access Blade to the Service Blade over the Data Plane Interconnect. The Service Blade then performs the necessary function such as data compression and sends the data to the Trunk Blade over the Data Plane Interconnect.
Control plane interconnect	All the blades are connected via the Control Plane Interconnect which is Ethernet or Advanced Switching.
Data plane interconnect	All the Access, Trunk, and Service Blades are additionally connected via the Data Plane Interconnect for the transport of bearer traffic. The control blade may also be connected to the data plane interconnect. For details of fabric technology and topology refer to Chapter 4. Note that there is no direct correlation between the control plane or data interconnect and the AdvancedTCA base and fabric interfaces. The control plane or data plane interconnect may use either the AdvancedTCA base or fabric interfaces depending on the needs of the particular NE being considered.

2.3.1 Control Blade

The Control Blade in a switch is similar to the control blade in a server. However it may have additional capabilities, such as being able to connect to a network clock source and having a mechanism to distribute the clock to the Line Cards. Network synchronization is critical to ensure that transmission quality is at its highest level. Without proper synchronization, timing slips can occur which can degrade a network's functionality. Building Integrated Timing Source (BITS) clocks are used to overcome this condition. The BITS clock extracts timing E1, T1, or 64 kbps signal information from a dedicated timing source, and this timing signal becomes the master timing reference for a given network element. Alternatively the clock can be derived from a line interface such as a T1/E1 or SONET/SDH interface. Note that the clock module can also be located on other blades such as the switch blades that can be expected to be used in all configurations of the network element.

If a single type of Control Blade is being developed for use in a Server and a Switch, it should encompass the functionality required by both. Figure 2-12 illustrates the structure of a Control Blade.

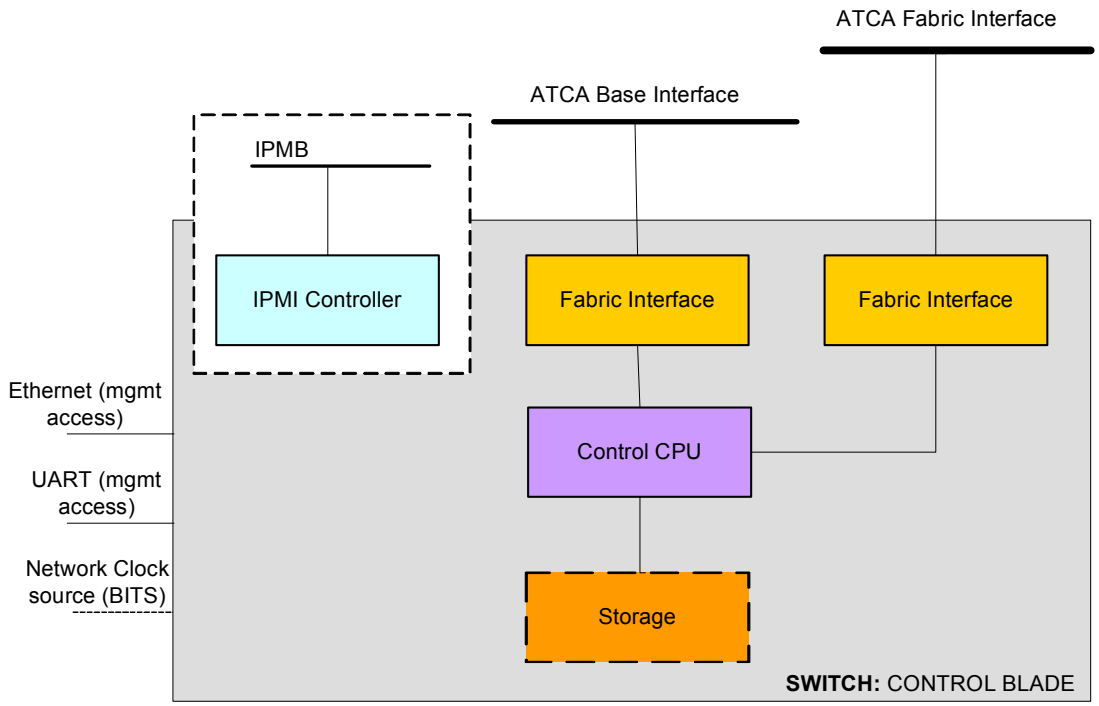


Figure 2-12 Control Blade

2.3.2 Line Card (Blade)

Line Cards can be of two types: Access Blades and Trunk Blades.

Regardless of the type of line card, the set of building blocks required are similar. Each line card has physical interfaces connected to MACs or Framers, which feed the data from network interfaces to one or more Network Processing Units (NPUs) that may have varying degrees of programmability. The Physical Interfaces on a Trunk Blade are of higher capacity than on the Access Blades. The packet processing stages within the NPUs vary depending on the protocol with which the data is associated, the services provided by the blade, and whether the blade is an Access or Trunk Blade. The data eventually makes its way to the Data Plane Interconnect (via the Backplane Fabric Interface) to be sent out on a different network interface. In some cases, the line card may perform local routing to local ports.

The Control CPU is optional in a Line Card. If present, it provides support for configuration of the line card and communication with the Control Blade. It may also be part of a distributed or hierarchical control processing system, handling some portion of the overall data plane control and protocol processing load. It handles “exceptions” locally, if it can, and forwards them to the control blade, if it cannot. If absent, an integrated controller within the NPUs may provide this functionality. Line Cards are connected to the Control Plane Interconnect and the Data Plane Interconnect. They communicate with the Control Blade over the Control Interconnect and transport bearer traffic using the Data Plane Interconnect. Figure 2-13 depicts the structure of a Line Card.

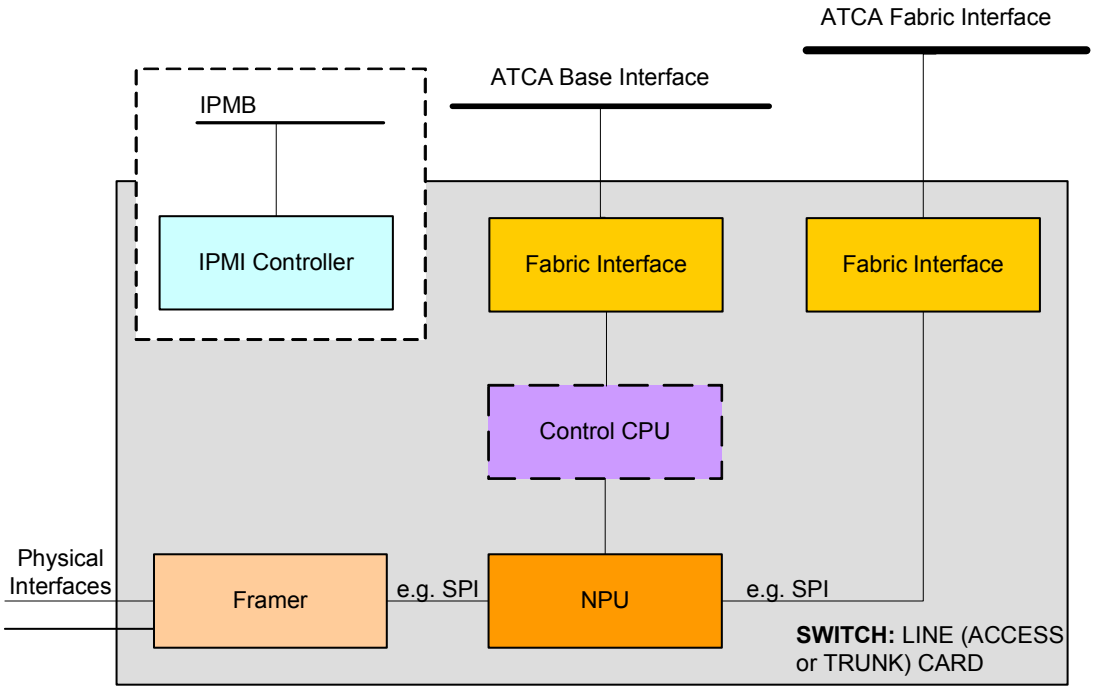


Figure 2-13 Line Card (Blade)

2.3.3 Service Blade

Service Blades support the data path by providing additional resources that are compute-intensive. They typically comprise multiple NPUs and/or DSPs for data path processing. Optionally, they may also include a Control CPU which provides support for communication with the Control Blade over the Control Plane Interconnect. As with the Line Card, if the Control CPU is absent, an integrated controller within the NPUs/DSPs is responsible for communication with the Control Blade for configuration, status, exception processing, protocol offload, etc.

Service Blades do not have external interfaces since they are not responsible for the ingress or egress of network data. They are only responsible for processing the bearer path traffic as it transits through the switch.

Figure 2-14 shows the structure of a Service Blade.

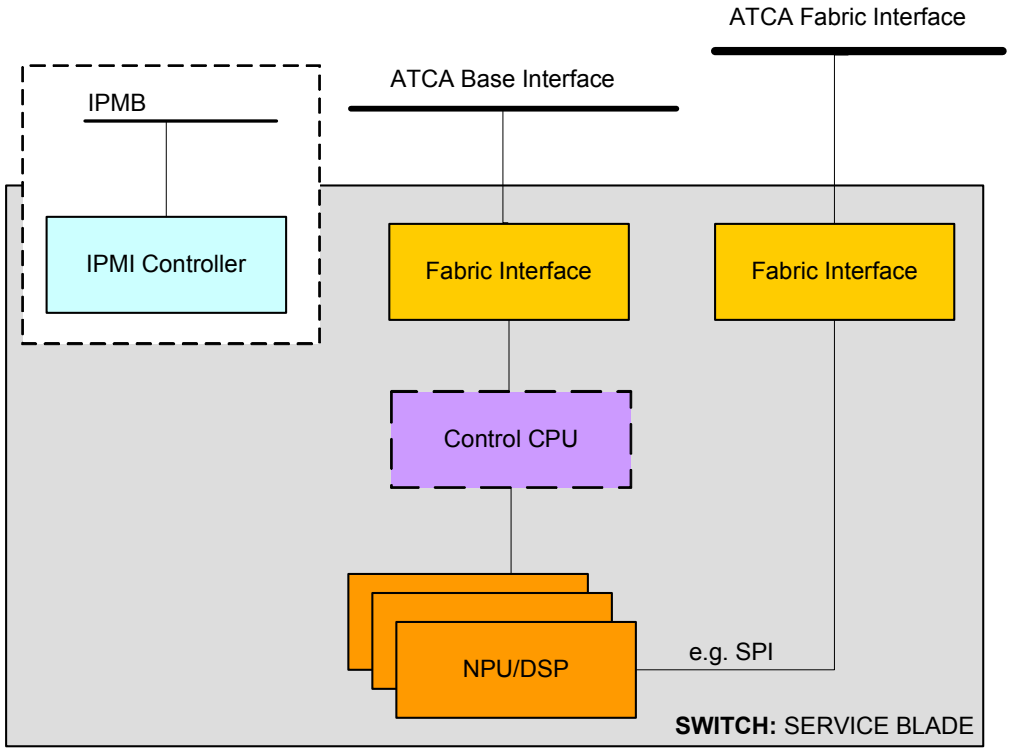


Figure 2-14 Service Blade: Hardware Building Blocks

2.4 Server and Switch Software Architecture

As with Hardware Building Blocks, Software Building Blocks support Control, Data, and Management Plane functionality.

The key difference between the software in a server and the software in a switch is in the software components of the control blade. A server consists of a control blade that performs various system management functions, as well as connection management and a set of protocol processing blades dedicated for protocol processing. Both the control blade and the protocol processing blades in a server may be configured in a redundant, load sharing mode. The software architecture for these boards would be very similar to that shown in Figure 2-18.

A control blade in a switch also implements system management, connection management, and protocol processing functions. A key additional function that a switch control blade implements is data plane management. Examples of data plane management functions initiated by a switch control blade are setting up forwarding table entries in a router or setting up MGCP connections in a media gateway. The software components on a switch typically initiate a data plane management function in response to a communication from the server. In a media switch for example, an MGCP connection request from the server results in the switch control blade performing a set of data plane management functions to service the request.

2.4.1 Control Plane to Data Plane Interaction

MCP network elements ideally use standard APIs from the Network Processing Forum (NPF) for control plane and data plane communications. The NPF defines a software framework and a suite of APIs that allow for the management of data plane components and control plane services in a manner that abstracts vendor-specific implementation details, allowing control components to interact with the data plane in a consistent manner. This enables the development of network applications that function without change in a variety of physical implementations.

The control plane software functions can be either fully encapsulated in a control blade, or they can be distributed across a control blade and a control CPU on the line card. The exact distribution of control and data plane functions varies from system to system and depends on the requirements of the overall system design. The communication between the control plane and data plane entities when they are on separate physical components is always based on the Forwarding and Control Element Separation (ForCES) protocol that is being standardized in the IETF. A key part of the ForCES protocol is the definition of a standard model of the data plane allowing capabilities to be discovered dynamically permitting new blades and services to be added to the system at runtime.

The following figures provide a simple illustration of three sample variations on the split between Control and Data Plane functionalities. The hardware architectures in the three cases offer a system designer varying degrees of scalability to match the specific system requirements.

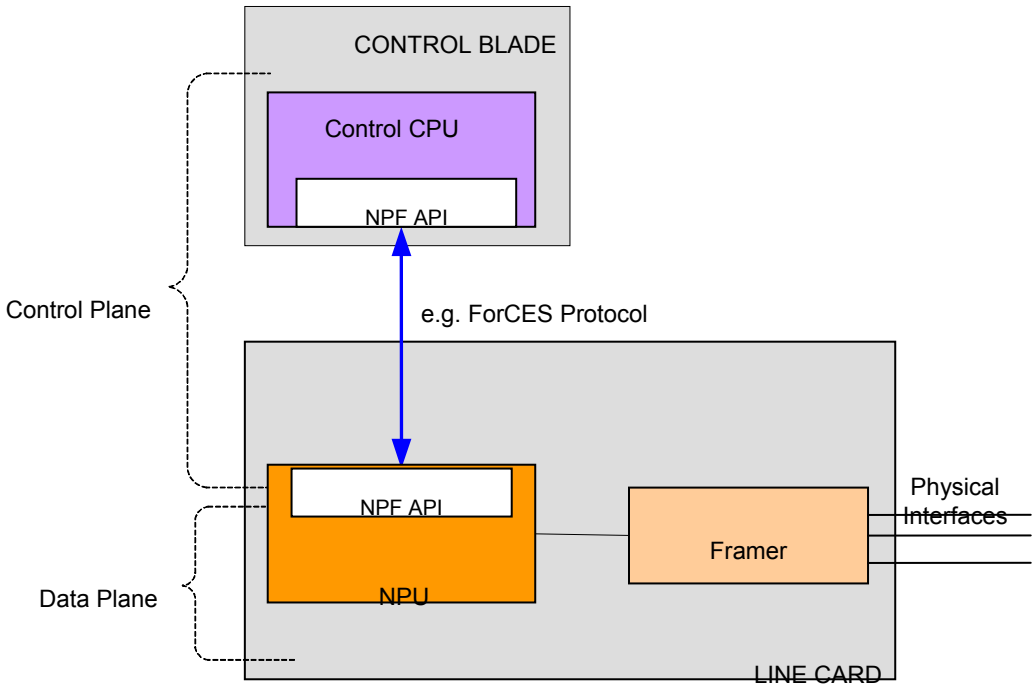


Figure 2-15 Centralized versus Distributed Control Plane: Case I

In the first case shown in Figure 2-15, the Line Card has no distinct Control CPU. So, some components on the Line Card are responsible for handling the transport of control and management data from and to the Control Blade. The Control Plane functionality in this case extends from the Control Blade to the NPF API functions of the NPU. Data Plane functionality is supported by the rest of the NPU that is involved in packet processing.

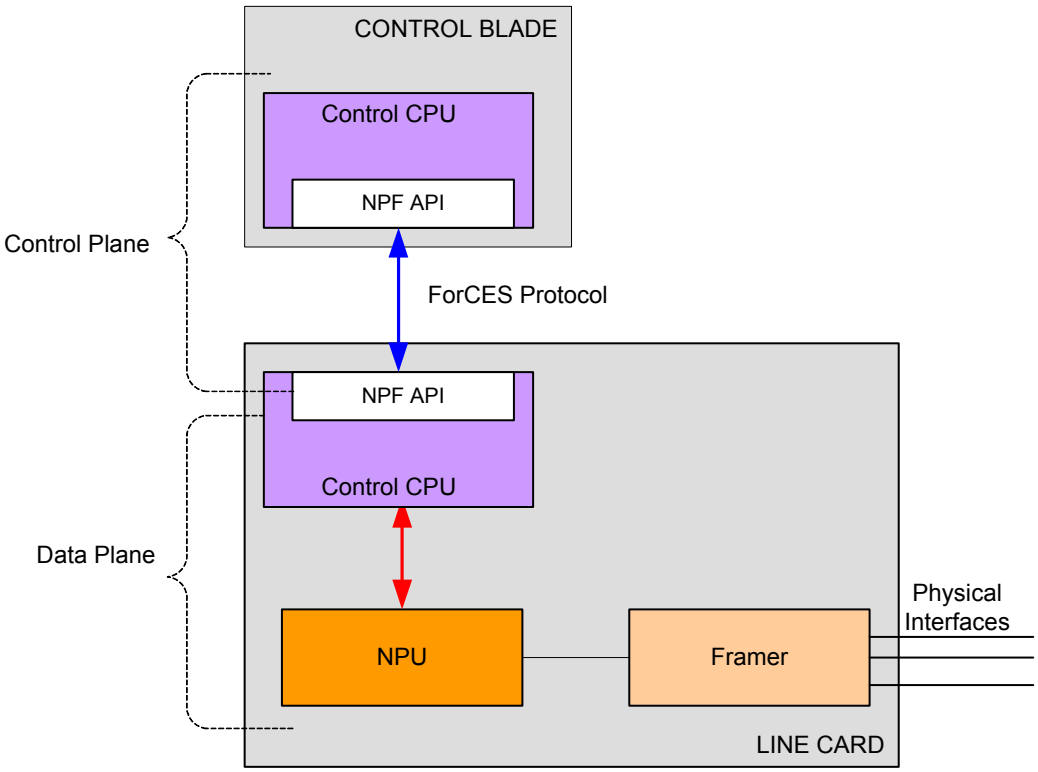


Figure 2-16 Centralized versus Distributed Control Plane: Case II

In the second case shown in Figure 2-16, the Line Card has a distinct Control CPU. This Control CPU on the Line Card may participate in the data plane functionality and supports handling the transport of control and management data from and to the Control Blade. The Control Plane functionality in this case extends from the Control Blade to the NPF API component of the Control CPU on the Line Card. Despite its name, the Control CPU on the Line Card does not execute Control Plane code. The control processor and the NPU support the required data plane functions.

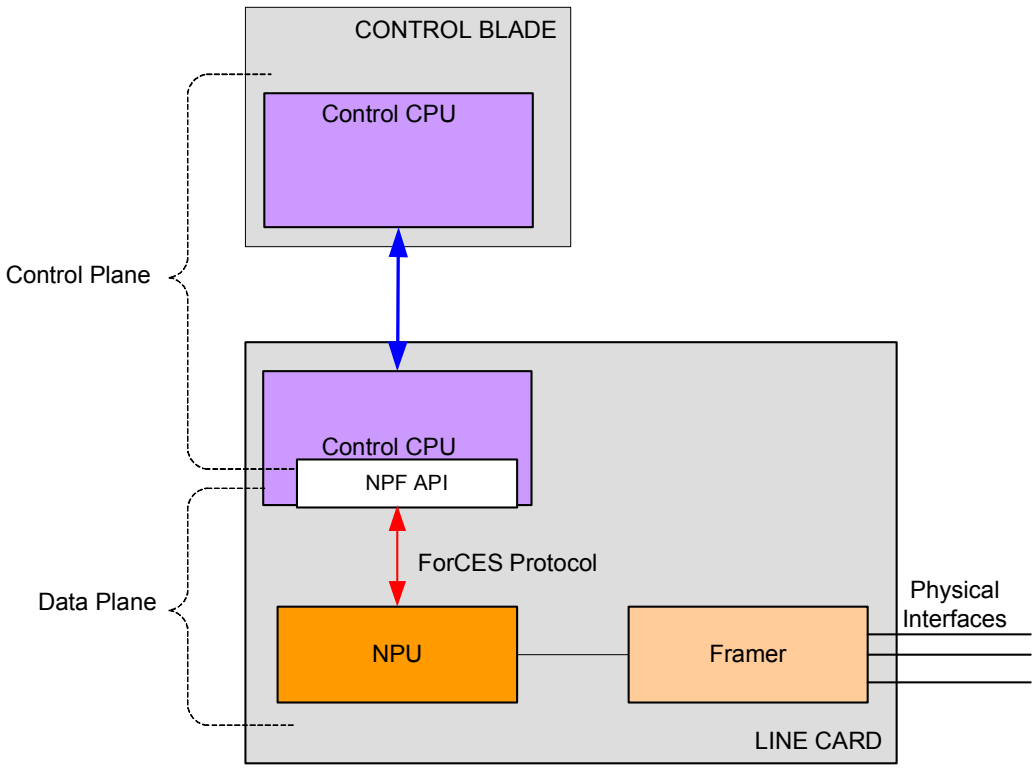


Figure 2-17 Centralized versus Distributed Plane: Case III

In the third case also, the Line Card has a distinct Control CPU as shown in Figure 2-17. But the Control CPU on the line card hosts control functions and plays no part in data plane processing. The control plane function is distributed across a Control Blade and all the Line Cards in the system.

2.4.2 Control Plane

Figure 2-18 illustrates the control plane software building blocks which reside primarily on the Control Blade.

At a high-level, the software building blocks comprise the following main components enumerated bottom up:

- Carrier Grade OS
- Messaging/Transport Plug-in
- Middleware
- Application Services
- Management Object Model

From a functional perspective, the components may be classified into two main classes:

- Components that provide infrastructure (represented by the blocks on the left-hand side of the figure)
- Control Plane protocols and stacks providing services

Some of these components such as the NPF Functional API (FAPI) implementation are specific to a switch control blade, while other components such as NPF Service API (SAPI) may be implemented either on a server control/protocol processing blade or on a switch control blade.

The following sections discuss the functionality associated with each of these building blocks.

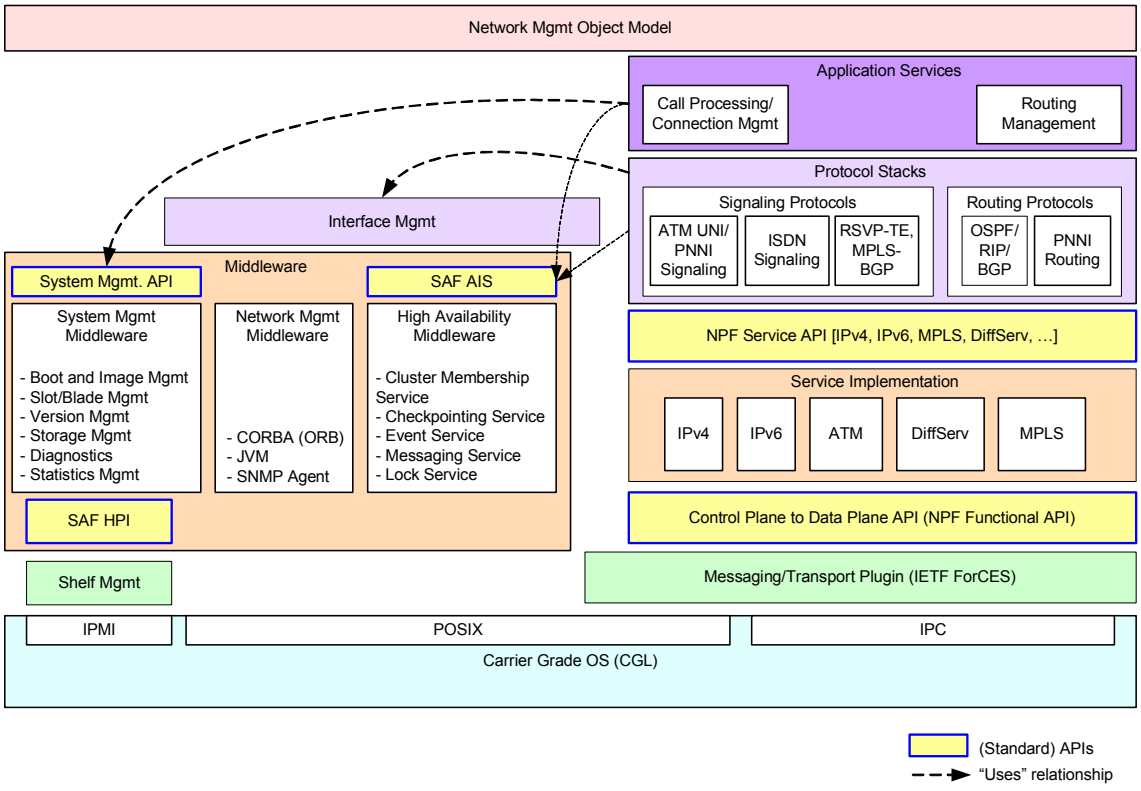


Figure 2-18 Control Plane: Example of software building blocks

2.4.2.1 Carrier Grade OS

The Carrier Grade Operating System (often Linux) provides an abstraction of the hardware and supports real-time scheduling, hardened drivers for devices, logging, events etc. In addition, it provides support for Inter-Process Communication (IPC), POSIX standards, and the IPMI driver to be used for Shelf Management.

For details of Carrier Grade Linux, refer to Chapter 10.

2.4.2.2 Shelf Management

The Shelf Manager is the entity in an AdvancedTCA chassis that uses IPMI to retrieve information on the blades, their types, sensor readings, events etc., and passes it to the system manager when required.

The interfaces between the shelf manager and the system manager (CIM and SNMP) and between the blade middleware and IPMI services (HPI) are described in Chapter 7.

2.4.2.3 Messaging/ Transport Plug-in

The messaging/ transport plug-in provides a protocol for the transport of control messages between the control plane and data plane. It provides a level of abstraction over data plane resources. This allows the control plane to configure and manage the data plane while making only minimal assumptions about the actual physical configuration of the data plane.

2.4.2.4 Middleware

Middleware supports three main classes of functionality:

Functionality	Description
System management middleware	The System Management Middleware Building Block supports a wide range of infrastructure services. Some of the core services include: • *Boot and Image Management*: Manages the installation or download of the appropriate software image for the blade(s) in question. • *Slot/Blade Management*: Manages the blades in the chassis – their power up/boot up, controlled shutdown, support for different administrative modes, configuration of the blades after boot-up etc. It also provides support for high availability via configuring Protection Groups. Two Trunk Blades may be configured as a 1+1 Protection Group; similarly, a few Access Blades may be configured as members of an N+M Protection Group. • *Version Management*: Supports management of different hardware versions and software image versions, and maintains a compatibility matrix across the mix of hardware and software. This ensures that the software running on a piece of hardware is compatible with the hardware, or different software building blocks are compatible with each other. • *Storage Management*: Provides access to storage of data such as configuration of the switch, statistics, logs etc. The storage may reside on the Control Blade, on a separate blade or specifically, a Storage Blade in the switch. It may also be external to the server or switch. • *Statistics Management*: Manages the collection of statistics from the blades, especially from the Line Cards, and possibly, support for interval and total statistics, which may be used for diagnostics and billing/usage measurement purposes. • *Diagnostics*: Supports execution of diagnostic tests and relays the information collected to management or in the case of faults, to the Availability Management Middleware Building Block. • The Network Management Middleware Building Block provides support for protocols or agents to support management access into the server or switch and the capability to process management-initiated requests or internal notifications (such as SNMP traps) that need to be delivered to external management servers. This entity provides the underlying support for the management of user-sessions for access to the switch. The service may be provided by an ORB, SNMP agent, JVM etc.
Availability management middleware	The Availability Management Middleware Building Block encompasses services that an NE needs to provide support for Fault Management and High Availability. This entity provides support to detect faults, classify them and take appropriate action to recover from them. In some cases, the action may be to autonomously switch over to an available backup resource. In other cases, there may be no recovery possible except for

Functionality	Description
	notification to a Network Element Management system (discussed in Chapter 2.5). Diagnostics included within the System Management Middleware Building Block also have a role to play in fault detection. Availability Management supports the following services: • Cluster Membership Management: Provides support for applications to retrieve information of the current membership of nodes in the cluster, and notification of membership changes. • Checkpointing Service: Provides support for applications to checkpoint their data incrementally for protection against failures. • Event Service: Provides support for applications to register for notification on occurrence of certain events. Publishers publish events, and subscribers are notified of the event. In essence, provides a mechanism for publishers and subscribers to communicate asynchronously. Events may be triggered due to external conditions (such as an alarm condition on a line) or may be triggered due to internal conditions in the switch such as a faulted blade. • Message Service: Supports inter-process communication within a node or across nodes in a cluster via message queues and message queue groups. • Lock Service: Supports distributed lock services for synchronizing access to shared resources.

Refer to the SA Forum Application Interface Specification for details on the functionality supported by each.

2.4.2.5 Interface Management

Interface management manages the physical ports and service interfaces in the server or switch. These include the management access ports as well as the bearer traffic ports. Examples of services provided by interface management include setting the attributes of a physical port such as whether the port is to be channelized, setting an optical port to SONET or SDH mode, and establishing and managing Inverse Multiplexing over ATM (IMA) groups. Interface Management also supports high availability via the configuration of Protection Groups for interfaces. For example, if two Trunk Blades are configured to provide 1+1 redundancy, the interfaces/ports associated with the blades also support 1+1 redundancy. In the case of SONET ports, they would support 1+1 SONET APS. So, in effect, there are two levels of redundancy:

- Blade-level redundancy
- Port-level redundancy

It is the responsibility of Interface Management to hide interface level redundancy from the higher-layers of application software so that switchovers (or failovers) from one port to another, do not impact the functions of the higher-layer services. This may be accomplished by exposing a single service interface to application software. This service interface should always be associated with the currently active port. A switchover should result in a change in the mapping between the service interface and the physical port it is associated with it, but should not impact higher level software (such as connections set up on the service interface, signaling protocols using the interface for signaling etc.) that is dependent on the service interface.

Current work in the NP Forum is expected to provide standard interfaces for many of the services referred to above.

2.4.2.6 Application Services

The functionality at the Application Layer is dependent on the type of NE. The Application Services component comprises several layers of software:

- Applications: The range of applications supported includes call processing, route management etc.
- Protocol Stacks: The specific protocol stacks are dependent on the type of NE. Examples are:
    - OSPF, RIP, BGP for IP Routing
    - PNNI for ATM Routing
    - ATM UNI/PNNI Signaling for ATM call setup/teardown
    - RSVP-TE, MPLS-BGP for signaling in an IP network
    - MGCP for communication with a softswitch
    - TCP/IP for management connections and applications such as proxies and load balancers

2.4.3 Data Plane

Figure 2-19 illustrates the software building blocks on the Line Card of a switch. Although the Data Plane provides support for the transport of user traffic, there is a component of the Data Plane that is responsible for configuring the data path with the services to be supported.

From a high-level (bottom-up) perspective, the Software Building Blocks comprise the following main components:

- Carrier Grade OS
- Messaging/Transport Plug-in
- Functional APIs and implementation

From a functional perspective, the components may be classified into two main groups:

- Components on the CPU of a Line Card, providing support to configure and manage the services to be supported (represented by the blocks on the left and center of the figure)
- Functional Blocks on the micro engines, for packet processing in the data path (represented by the blocks on the right-hand side of the figure)

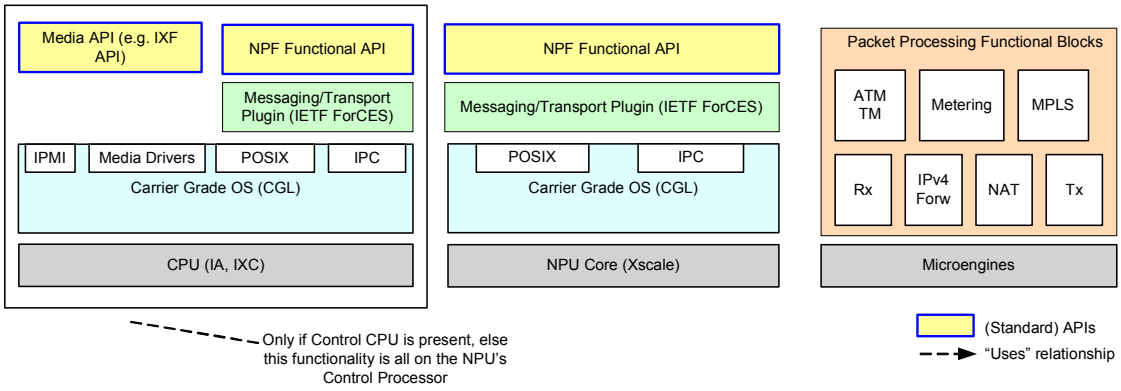


Figure 2-19 Data Plane: Software building blocks

The next few sections discuss the functionality associated with each of these Building Blocks.

Many of the data plane components described here are available as part of the Intel® Internet Exchange Architecture (Intel® IXA) Software Development Kit and can be enhanced to meet specific requirements. These components are developed in the context of the Intel® IXA Portability Framework, which is a software infrastructure for writing modular, portable software for network applications. The framework encompasses the following:

- The Intel® IXA SDK for developing packet processing functional blocks. Includes support for NPF Functional APIs
- The CP-PDK which provides an interface for the control plane to manage functional blocks in the data plane. The CP-PDK supports the IETF ForCES protocol as well as the NPF Service APIs
- An array of tools, such as simulators, compilers, debuggers, and libraries, such as OS abstraction layers and hardware abstraction layers, that can assist in developing data plane software

Building block	Description
Carrier grade OS	As before, the Carrier Grade Operating System (often Linux* or VxWorks*) provides an abstraction of the hardware; supports real-time scheduling, hardened drivers for devices, support for logging, events etc. In addition, it provides support for Inter-Process Communication (IPC), POSIX standards and IPMI driver. It also provides support for media devices such as framers, MACs and LIUs.
Media API (IXF API)	The IXF API provides an Intel standard interface for configuring and managing Intel IXF devices (framers and MACs). The IXF API provides an interface which abstracts the underlying hardware device so that driver code may be ported (with minimal modifications) to other devices similar in functionality (e.g. porting existing software for controlling an Intel OC-3 framer/ mapper to also be able to control an Intel OC-48 framer/ mapper).
Messaging/ transport Plug-in	The messaging/ transport plug-in provides a protocol for the transport of control messages between the control plane and data plane. It provides a level of abstraction over the data plane resources. This allows the control plane to configure and manage the data plane while making only minimal assumptions about the actual components residing in the data plane (e.g. a data plane consisting of multiple blades, a single blade, multiple processors on a single blade, etc).
NPF Functional APIs	The NPF Functional APIs provide support for configuration of the data path. These industry-specified APIs expose a standard interface for managing logical functional blocks (LFBs) in the data plane (e.g. Forwarding block, classification block, etc.).
Packet Processing Functional Blocks	Packet processing functional blocks perform the actual processing of packets received on network interfaces. The functionality supported depends on the protocols supported in the Data Plane. Functional blocks might include IP forwarding, packet classification, metering, policing, shaping, data compression, etc.

2.5 Network Element Management

A Network Element needs to be managed and is typically managed as a single entity even though the Server and the Switch functionality might be distributed across multiple chassis. A Network Element may be managed using several protocols: CORBA, SNMP, TL1, CIM, telnet to an embedded CLI etc. The Network Element exposes an object model to an external Network Element Management application e.g. TIVOLI, HP, OPENVIEW. The network element administrator uses the object model to perform various OAM operations such as system configuration, statistics monitoring and status monitoring. Network equipments typically expose multiple management interfaces, a programmable interface such as CORBA or CIM over SOAP/XML as well as a command line or graphical user interface. The command line interface is typically used when physical access is required to manage the equipment either due to security reasons or due to equipment faults.

The monitoring and management of all resources within the Network Element including the following:

- Discovery, provisioning, and inventory of physical system resources, such as blades, slots, power supplies and power entry modules, fans and fan trays, sensors, interlocks, etc
- Provisioning and inventory of all software applications running on the server such as network management, OAMP, value-add user apps (billing QoS, etc), and management applications

- Maintenance of all inventoried hardware and software in a system model or schema stored in a persistent storage medium such as disk or flash
- Fault detection and preventative diagnostics through monitoring the health of all hardware and software resources inventoried through heart-beating and process monitoring
- Fault isolation, diagnosis, failover, repair, and restoration

Management support is typically provided via network (typically Ethernet) access from the Control Blade(s) of the Network Element. Redundant paths are normally provided for access to the NE. Also, serial port connectivity is typically provided to help with major troubleshooting and diagnostics.

Fault management, Configuration, Accounting, Performance, and Security (FCAPS) as defined by ITU-T and standardized by the International Standards Organization (ISO) defines a model for the objectives of network element management. The five aspects of management in this model are:

Aspect	Description
Fault management	Involves the detection of faults in an NE and the actions taken to recover from them. For example, the NE would need to support the generation, logging, and notification of an alarm event on an interface, followed by the execution of alarm correlation to perform root-cause analysis; this would help determine the corrective action to be taken.
Configuration	Involves support for the monitoring of, and provisioning the NE and its sub-systems. The types of services to be provisioned depend on the type of NE. For example, in the case of a multi-service switch, it would need to support the provisioning of services such as IP, MPLS, and ATM at the interfaces, and the configuration of connections for carrying user traffic. Additionally, support also needs to be provided for the persistent storage of configuration data.
Accounting	Involves support for the fair distribution of NE resources among subscribers. For example, service usage would need to be monitored for billing purposes.
Performance	Involves management of the overall performance of the network in an attempt to maximize throughput and avoid bottlenecks. As an example, statistics would need to be collected for various services being provided, and the data analyzed. Based on the analysis, corrective action might need to be taken.
Security	Involves protection of the network against unauthorized users and the configuration of access rights for authorized users. For example, user accounts need to support passwords, automatic password ageing, access permissions etc. In addition, an audit trail might be needed to log operator actions.

This page intentionally left blank.

3 AdvancedTCA Overview

PICMG 3.x is a family of specifications that defines a new generation architecture for building high end, “carrier grade” equipment. The PICMG 3.x specifications are oriented around switch fabric technology instead of a conventional parallel bus. This is an evolution from PICMG 2.16 based systems where the Ethernet based fabric backplane was first introduced.

The specifications provide enough information to allow board, backplane, and chassis vendors to independently develop products that will be interoperable when integrated. Details include board dimensions, equipment practice, connectors, power distribution, and a robust system management architecture that can be used independent of the switch fabric link technology.

PICMG 3.x is a term sometimes used to describe the family of specifications. It is used interchangeably with PICMG 3. The specifications include:

Specification	Description
PICMG 3.0	This is the overall general specification that defines the mechanics, board dimensions, power distribution, power and data connectors, and system management. This is a ‘fabric agnostic’ specification that is intended to serve the needs of a variety of fabrics currently on the market and allows the flexibility to support new fabric interfaces.
PICMG 3.1	This specification defines an Ethernet switch fabric over the generic backplane fabric interconnect. It provides for data rates up to 10 Gbit/sec. per link. It also defines the use of Fibre Channel.
PICMG 3.2	This specification defines how InfiniBand systems are built within the architecture and specifies link physical layers, protocols, and protocol mappings.
PICMG 3.3	This specification defines a StarFabric implementation over the backplane providing TDM, cell, control, and packet connectivity over the same fabric.
PICMG 3.4	This specification defines the use of both PCI Express Base and Advanced Switching based on PCI Express within an AdvancedTCA platform.
PICMG 3.5	This specification defines the Advanced Fabric Interconnect/ S-RapidIO.

Figure 3-1 shows the complete PICMG 3 family of specifications. However, the Design Guide intends to narrow the options for greater interoperability and focuses on PICMG 3.0 and fabric implementations based on PICMG 3.1 and PICMG 3.4 as detailed in Chapter 4.

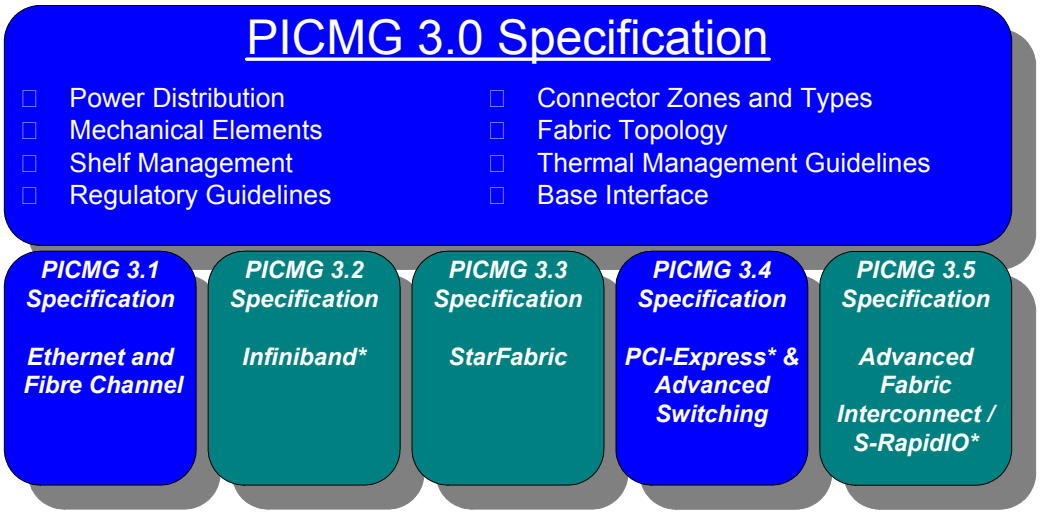


Figure 3-1 AdvancedTCA Family of Specifications

The base specification, PICMG 3.0 was voted and approved in December 2002. Since then, sub-specifications 3.1-3.4 have also been approved. In 2003, the committee worked on and completed PICMG 3.0 ECR (Engineering Change Record) release (approved in January 2004).

Since the approval of the specification, PICMG has conducted several AdvancedTCA Interoperability Workshops (AIWs) with the participation of various component manufacturers. These AIW sessions provided a medium for the companies to test their products with each other for interoperability and compliance. Intel has also started to put together a set of open sourced compliance tools (ACTS: AdvancedTCA Compliance Test Suite) to automate many of the AdvancedTCA compliance testing to facilitate greater interoperability.

A PICMG Module subcommittee focuses on the development of Advanced Mezzanine Card (AMC) specifications. The committee plans to finalize the specification in Q3 of 2004. Once completed, AMC has the potential to augment the next generation of modularity in the system.

The subsequent chapters of this Design Guide examine key requirements and design issues in the AdvancedTCA Building blocks (Fabric, Shelf, Board and Shelf Management) as well as discuss the complementary software building blocks such as Carrier Grade OS which are essential to develop a modular communication system.

4 MCP Backplane Fabric Topologies and Technologies

This chapter specifies the communications and storage backplane fabric technologies and topologies mandated for AdvancedTCA MCP network elements. To ensure a high degree of interoperability, only a subset of the many options permitted in the AdvancedTCA specification is allowed. This chapter also specifies how TDM voice traffic is to be carried over the backplane fabric in an interoperable fashion.

4.1 Communications Fabric Technologies for MCP

AdvancedTCA MCP network elements use Ethernet or Advanced Switching for their backplane fabrics in accordance with PICMG 3.0, PICMG 3.1, and PICMG 3.4.

These fabric technologies allow a wide range of network elements ranging from blade servers to communications systems such as media gateways, wireless infrastructure systems, and edge routers to be built. These fabric technologies allow line cards of approximately OC48c to 4x1GbE capacity to be supported. Specialized faster trunk or uplink interfaces (OC192c or 10GbE) can be supported on the switch cards which serve as an aggregation point for traffic.

Ethernet, on the AdvancedTCA Base Interface, is expected to be the backplane fabric technology used in the first generation of AdvancedTCA systems. These systems are mostly telco-grade blade servers or communications systems with relatively low throughput such as the network elements used in wireless telephony systems. Bandwidth beyond 1Gbps can be provided by using Ethernet on the AdvancedTCA Fabric Interface where up to four GbE links may be implemented with the option also of using 10GbE.

As the communications aspect of AdvancedTCA systems increases in the future, as for example when edge routers are built on AdvancedTCA, Advanced Switching provides the fabric features required such as quality of service, traffic isolation, and congestion management. For these systems, Advanced Switching provides an interoperable fabric technology with the features typically associated with today's proprietary communications fabrics.

Note that the use of Fibre Channel according to PICMG 3.1 is also permitted for storage applications and is described in Chapter 4.6.

ID	Feature Name	Requirements
FT-1	Fabric Technology	MCP network elements ***shall*** only use Ethernet (PICMG 3.0 and PICMG 3.1) and Advanced Switching (PICMG 3.4) as backplane fabric technologies.

4.2 Fabric Topologies for MCP

MCP network elements most commonly use a star topology with the switch blade replicated if needed for redundancy. Note that the switch blade does not have to be a dedicated blade and may also be able to perform other centralized control or data path functions such as providing a trunk interface. Some systems may benefit from a mesh topology and so this is allowed in specialized cases (low slot count systems, etc). These requirements apply only to the AdvancedTCA Fabric Interface. As required by PICMG 3.0 the Base Interface always uses a star configuration.

ID	Feature Name	Requirements
FT-2	Fabric Topology	MCP network elements ***should*** only use star configurations for their fabrics (via the AdvancedTCA Fabric Interface).
FT-3	Fabric Topology	For high availability, a redundant switch ***may*** be used in either a load shared or active/standby mode.
FT-4	Fabric Topology	MCP network elements ***may*** use mesh configurations for their fabrics (via the AdvancedTCA Fabric Interface).

4.3 Base Interface (PICMG 3.0)

All blades in MCP network elements must provide support for the base interface in accordance with PICMG 3.0. This includes blades that also support PICMG 3.1 or PICMG 3.4 on their AdvancedTCA Fabric Interfaces.

ID	Feature Name	Requirements
FT-5	Base Interface	All blades in MCP network elements ***shall*** support the AdvancedTCA Base Interface.
FT-6	Base Interface	Blades in MCP network elements ***should*** support the maximum 1Gbps speed on the AdvancedTCA Base Interface.

4.4 Ethernet (PICMG 3.1)

If multiple Ethernet links from a blade to a switch are to be aggregated or “trunked”, this aggregation must be performed utilizing IEEE* 802.3ad Section 43.

ID	Feature Name	Requirements
FT-7	PICMG 3.1 Fabrics	Blades that support Ethernet (PICMG 3.1) as a fabric technology ***may*** use both the Base and Fabric interfaces simultaneously (using independent Ethernet interfaces on each).
FT-8	PICMG 3.1 Fabrics	If multiple Ethernet links are aggregated, the aggregation ***shall*** be per IEEE* 802.3ad Section 43.
FT-9	PICMG 3.1 Fabrics	Ethernet Switch blades ***shall*** provide Layer 2 including IEEE* 802.1p and IEEE* 802.1q switching services.
FT-10	PICMG 3.1 Fabrics	Ethernet Switch blades ***may*** provide Layer 3 and higher layer switching services.
FT-11	PICMG 3.1 Fabrics	The IEEE* 802.1p priorities (or IEEE* 802.1d class of service) ***should*** be used according to the guidelines in IEEE* 802.1d.

4.5 Advanced Switching (PICMG 3.4)

Advanced Switching is used on the Fabric Interface as specified in PICMG 3.4. Note that while PICMG 3.4 permits the use of PCI Express, MCP network elements do not use PCI Express as a backplane interconnect.

Since only a straightforward (possibly redundant) star configuration is required, in most cases AS fabric management can be kept simple and possibly limited to out-of-band programming of the switch configuration by a local processor on the switch blade.

ID	Feature Name	Requirements
FT-12	Advanced Switching	PICMG 3.4 based AdvancedTCA blades ***shall*** support Advanced Switching as a backplane technology.
FT-13	Advanced Switching	Blades that support Advanced Switching ***shall*** support one of the three following sets of lane widths: (a) x1 only (b) x1 and x2 (c) x1, x2 and x4.
FT-14	Advanced Switching	Blades that support Advanced Switching ***shall*** support at least four (4) Ordered VCs for unicast bearer traffic. This provides the minimum traffic isolation capabilities required for many communications systems especially when a combined control/data interconnect is used.

ID	Feature Name	Requirements
FT-15	Advanced Switching	Blades that support Advanced Switching ***shall*** support at least one (1) Bypassable VC for fabric management traffic.
FT-16	Advanced Switching	Blades that support Advanced Switching ***shall*** support at least one (1) multicast VC.
FT-17	Advanced Switching	Blades that support Advanced Switching ***shall*** support segmentation and re-assembly using PI-2 for packet protocol interfaces (PIs) using the in-order variant of PI-2.
FT-18	Advanced Switching	Blades that support Advanced Switching ***may*** support segmentation and re-assembly using PI-2 for packet PIs using the out-of-order variant of PI-2.
FT-19	Advanced Switching	The native segmentation and re-assembly mechanism specified for SQ, SLS and SDT ***shall*** be used for those specific PIs.
FT-20	Advanced Switching	Blades that support Advanced Switching ***shall*** support Status Based Flow Control.
FT-21	Advanced Switching	Blades that support Advanced Switching ***shall*** support virtual output queues to at least sixteen (16) destinations with at least four (4) traffic classes.
FT-22	Advanced Switching	Blades that support Advanced Switching ***shall*** support Source Injection Rate Control on all virtual output queues (also referred to as Connection Queues).
FT-23	Advanced Switching	Blades that support Advanced Switching ***shall*** support the minimum bandwidth egress scheduler on all egress ports of switches.
FT-24	Advanced Switching	Blades that support Advanced Switching ***shall*** support the “perishable bit”.
FT-25	Advanced Switching	Blades that support Advanced Switching ***shall*** support the expected communications PIs from the ASI SIG which are expected to be compatible with Network Processing Forum (NPF) Messaging.
FT-26	Advanced Switching	Blades that support Advanced Switching ***may*** support the PCI Express to Advanced Switching Bridging Architecture (PI-8).
FT-27	Advanced Switching	Blades that support Advanced Switching ***may*** support the Simple Load Store (SLS) PI.
FT-28	Advanced Switching	Blades that support Advanced Switching ***may*** support the Simple Queueing (SQ) PI.
FT-29	Advanced Switching	Blades that support Advanced Switching ***may*** support the Socket Data Transfer (SDT) PI.
FT-30	Advanced Switching	Blades that support PI-8 or SLS/SQ ***should*** support more than the required single Bypassable VC.

4.6 Storage Fabrics

The only permitted dedicated storage fabric in AdvancedTCA systems is the Fibre Channel as defined in PICMG 3.1. Therefore in systems where PICMG 3.1 is employed on the AdvancedTCA fabric interface, the Fibre Channel may be used on one or two channels.

Using the Fibre Channel reduces the number of channels available for Ethernet, and system design must accommodate this. One common configuration that is likely for telco-grade blade servers built on AdvancedTCA is to use the Base interface for communications and use the Fibre Channel (and no Ethernet) on the fabric interface.

Note that PICMG 3.4 and PICMG 3.1 when used with 10Gb Ethernet make no allowance for the Fibre Channel and so systems that use these fabrics cannot also use Fibre Channel as a storage backplane fabric. These systems will need to use the emerging IP-based block storage protocols such as Internet Small Computer Systems Interface (iSCSI) and Fibre Channel over TCP/IP (FCIP) over their base or PICMG 3.1 or PICMG 3.4 fabrics.

ID	Feature Name	Requirements
FT-31	Storage Fabric	MCP systems that require block storage services over the backplane ***should*** use iSCSI or FCIP block storage protocols if they use either PICMG 3.4 Advanced Switching or PICMG 3.1 10Gb Ethernet for the backplane fabric.

4.7 TDM Operation

Certain network elements such as media servers need a mechanism for transmitting 64kbps telephony TDM streams between blades. Traditionally this has been provided by specialized TDM buses which have generally been proprietary but more recently have been based on such standards as H.100 and H.110 from the Enterprise Computer Telephony Forum (ECTF).

AdvancedTCA has no provision for a H.110 like dedicated bus for 64kbps telephony TDM streams. Therefore the equivalent function must be provided over the base and fabric interfaces.

PICMG is standardizing a Layer 2.5 TDM encapsulation mechanism which consists of two protocols, SFP – System Fabric Plane encapsulation shim for Ethernet, and I-TDM – Internal Time Division Multiplex encapsulated protocol optimized for PICMG based switched packet fabrics. I-TDM defines two modes of operation for different application, a 1ms mode for media server and media gateway like applications, and 125us mode for PSTN switching type applications. SFP + ITDM will provide interoperability between boards that send and receive telephony TDM streams in MCP systems that use Ethernet for the backplane fabric.

SFP and I-TDM are designed to allow them to be implemented with a software interface very similar to what is currently used for H.110 based systems to ease migration to AdvancedTCA.

For systems using Advanced Switching as the backplane fabric, the Advanced Switching Interconnect Special Interest Group (ASI SIG) plans defining a Generic Tunneling PI to perform the encapsulation / shim functions for Advanced Switching as SFP does for Ethernet for the encapsulation of I-TDM frames. The Generic Tunneling PI plus PICMG I-TDM will provide interoperability for systems that encapsulate TDM at Layer 2.5.

Many systems will encapsulate telephony TDM streams at a higher layer using the Real-time Protocol (RTP) over UDP and IP. Blades in these systems will be interoperable by conforming to the various payload format specifications for TDM encapsulation in RTP published by the IETF.

ID	Feature Name	Requirements
FT-32	TDM Flows	MCP systems that use Ethernet as the backplane fabric and exchange TDM flows at Layer 2.5 ***shall*** encapsulate the TDM flows according to the

ID	Feature Name	Requirements
		ITDM and SFP specifications expected from PICMG later in 2004.
FT-33	TDM Flows	MCP systems that use Advanced Switching as the backplane fabric and exchange TDM flows at Layer 2 ***shall*** encapsulate the TDM flows according to the Generic Tunneling PI plus I-TDM specifications expected from the ASI SIG later in 2004.
FT-34	TDM Flows	MCP systems that exchange TDM flows at a layer higher than Layer 2.5 ***shall*** do so using a standard RTP payload format for TDM multiplexing over RTP, UDP, and IP.

This page intentionally left blank.

5 The AdvancedTCA Shelf

The AdvancedTCA* Shelf was conceived to provide a standards-based modular hardware platform that can support equipment developed for the next generation network infrastructure. Shelf configurations include but are not limited to for use in 600 mm ETSI racks and cabinets, and 19 inch and 23 inch width EIA racks and cabinets.

An AdvancedTCA MCP shelf would typically encompass the following functional components:

- A Subrack which provides the mechanical interface to PICMG 3.0 Boards and consists of Guide Rails, ESD discharge clips, alignment/keying, Handle interface, Face Plate mounting hardware, EMC gasketing, and Backplane interface
- A PICMG 3.0 compliant backplane providing interconnect for module power, IPMB management interconnect to modules and boards, Telco clock, update ports, as well as base interconnect and fabric interconnects
- Power subsystem that consists of Power Entry Modules (PEMs) supporting redundant -48V nominal (or -60V) power feeds to the shelf
- Cooling subsystem that consists of air filtration and Fan Modules supporting the maximum cooling requirements of the AdvancedTCA shelf
- Shelf Management
- Shelf FRU Information
- Telco Alarm module

The MCP shelf is normally provided by a single vendor as a fully integrated building block that is ready to be used by the Network Equipment Providers to build a system using AdvancedTCA Front boards and RTMs from multiple vendors. The MCP shelf supports all mandatory PICMG 3.0 requirements unless otherwise noted in this chapter. This chapter provides additional functional requirements that should be satisfied by all MCP shelves to ensure interoperability with AdvancedTCA compliant Front boards and RTMs. This chapter also provides recommended implementation guidelines for some features.

Figure 5-1 shows an example of a 14 slot AdvancedTCA Shelf fully populated with boards.

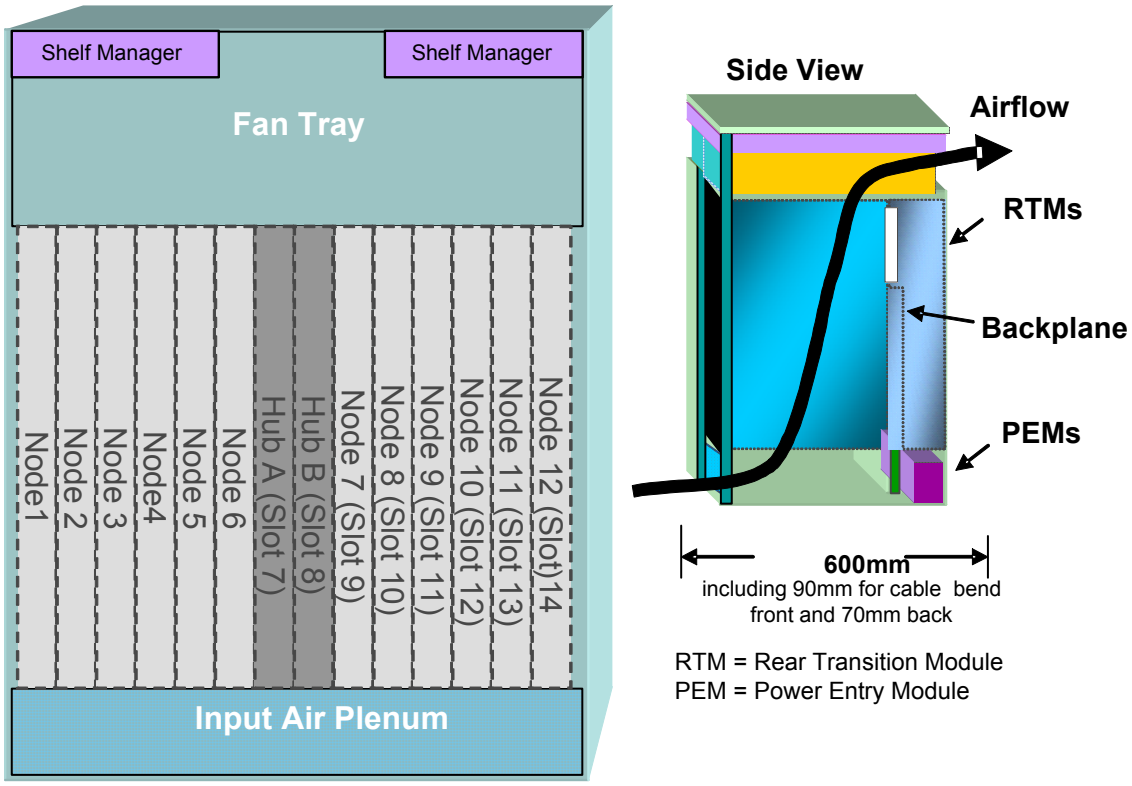


Figure 5-1 Example of a 14 slot AdvancedTCA Shelf and components

5.1 Shelf Design

This section covers the guidelines and requirements for building an AdvancedTCA MCP shelf. In general, the MCP shelf should be designed to meet the high availability required by the carrier environment. The shelf design should build in the necessary redundancy that is needed to achieve this level of availability. Since the shelf is the central building block for any system, it is important that the shelf is able to work with a wide variety of AdvancedTCA boards from multiple vendors. A fully configured shelf should be designed to be capable of providing 99.999% of availability, and should also be able to function normally without any backup modules installed. This allows the end user to use the same shelf in an environment where lower cost is more important than system availability.

ID	Feature Name	Requirements
SH-1	Shelf Design	A fully configured shelf ***shall*** have an availability of 99.999% or better based on both hardware and software.
SH-2	Shelf Design	A fully configured shelf ***shall*** continue to function without interruption with the loss of any single shelf element (including a FRU cooling unit).
SH-3	Shelf Design	All blades in a shelf ***shall*** continue to operate even if a shelf manager is unavailable with the understanding that the shelf management functions will be unavailable until the shelf management functionality is restored.

5.1.1 Connectors

ID	Feature Name	Requirements
SH-4	Shelf Design	Connectors ***shall*** meet all requirements specified in PICMG 3.0.
SH-5	Shelf Design	All Backplane connectors ***should*** be Press fit connectors.
SH-6	Shelf Design	Shelf Modules (PEMs, Fan Tray, and Filter Tray) ***should*** use blind mate connectors.
SH-7	Shelf Design	Cable connectors, if used, ***shall*** be polarized connectors with a latching mechanism.

The latching mechanism prevents the connector getting loose during vibration.

5.1.2 Accessibility

Avoid side interference by designing to eliminate side-facing tool movements. Design tool movements to be either front-facing (accessible from the front) or rear-facing (accessible from the rear). Depending on the Rack design and whether it is placed next to another Rack, after a Shelf is placed on a Rack, side restrictions may make it time-consuming and difficult for operators to service a shelf.

ID	Feature Name	Requirements
SH-8	Shelf Design	A Shelf ***shall*** not place any serviceable components on the side of the shelf.

5.1.3 Rack Mounting

ID	Feature Name	Requirements
SH-9	Shelf Design	A Shelf ***shall*** support mounting for both English (per EIA-310, IEC 60297-1) and Metric (per IEC 60917-2).
SH-10	Shelf Design	A Shelf ***shall*** support front and center mounting within a rack.

5.1.4 Backplane Support

ID	Feature Name	Requirements
SH-11	Shelf Design	A Shelf that has a backplane in Zone 3 ***shall*** support boards that do not have any Zone 3 connectivity.
SH-12	Shelf Design	Connectors on the extended backplane ***should*** be the same as those in Zone 2.
SH-13	Shelf Design	A Shelf ***shall*** be able to support backplanes with a thickness of 8mm.

Having connectors on the extended backplane that are the same as those in Zone 2 enables OEMs to implement their existing proprietary bus on the backplane in the Zone 3 area as a first step towards a transition to full AdvancedTCA compliance. With the backplane in the Zone 3 area, RTMs will not be possible in the corresponding RTM slots.

5.1.5 Fasteners

ID	Feature Name	Requirements
SH-14	Shelf Design	Metric fasteners ***shall*** be used for mounting shelves and modules within the shelf.

5.1.6 Handling

ID	Feature Name	Requirements
SH-15	Shelf Design	A Shelf ***shall*** be able to rest on its front or rear surfaces without toppling over.

5.1.7 Surface Finish

Avoid corrosion issues by the use of painted surfaces, post-plated metal such as steel that is nickel-plated after machining (cutting and bending can damage metal surfaces that include cut edges and tight-radius bends) or use stainless steel or similar materials that do not rust or corrode.

Design with cosmetic considerations of the client in mind. Avoid troublesome fingerprint problems on board backing plates, shelf surfaces, and filler plates. Recommend painting some surfaces, using post-plated or stainless materials, and careful fabrication.

ID	Feature Name	Requirements
SH-16	Shelf Design	Surface Reflectance and Color ***shall*** be as per GR-63-CORE, O4-76 unless otherwise specified by the customer.
SH-17	Shelf Design	A Shelf ***shall*** be of quality design without any sharp edges.

5.1.8 EMI/EMC Design Guidelines

Refer to Chapter 11 Certification and Practices for the EMC design guidelines.

5.1.9 Power Connections

Torque requirements should be placed on the shelf next to the power connection and ground lugs. Dual lugs are better. Design to eliminate cable rotation. For example, use a fitting into which the cable terminal seats.

Note that the cable conductor size is often determined by the allowable voltage drop. As a result the conductor size is often larger than that required by UL. 1V drop is normally allowed each way from the power plant to the shelf.

ID	Feature Name	Requirements
SH-18	Shelf Design	A Shelf ***should*** provide dual hole lugs for Power feed cables as per UL 486A.
SH-19	Shelf Design	A Shelf ***shall*** provide strain relief mechanisms that are capable of withstanding a force of 90 lbs.
SH-20	Shelf Design	The voltage drop across the power feed cables between power plant and shelf ***shall*** be 1V or less.

ID	Feature Name	Requirements
SH-21	Shelf Design	A Shelf ***shall*** provide a mechanism to prevent cable rotation under vibrations and high current.

5.1.10 Internal Wiring

Although cables can be used to interconnect the parts within a shelf, they should be avoided as much as possible for shelves since they are prone to failure. A better approach is to use PCBs to interconnect the components without cables. Cables can fail due to pinching, scraping, or cutting when boards and other removable units are inserted or removed.

ID	Feature Name	Requirements
SH-22	Shelf Design	Internal wiring of the shelf ***should*** be done using PCBs instead of cables wherever it is practical. If cables are used, then all cables shall be fixed to the shelf so that it does not obstruct FRU insertion and removal.

5.1.11 Cable Management

The shelf should provide appropriate cable management so that cables are not dropping down and blocking the airflow, either in the front or in the rear of the shelf, and it is much easier to service the shelf.

ID	Feature Name	Requirements
SH-23	Shelf Design	FRU replacement and door operation ***shall*** not be impeded by cables or their management.
SH-24	Shelf Design	Cables and Fibres ***should*** be down draped by 45 degrees to lessen mechanical stress.

5.1.12 Subrack

The mechanical requirements for the Subrack are described in Section 2.6 of PCMG 3.0. It is important that the Subrack be designed to meet all requirements in the specification since the Subrack features are mandatory to guarantee interoperability of boards in the shelf. Subrack designers should pay particular attention to meeting all tolerances.

5.2 Shelf FRU Device

ID	Feature Name	Requirements
SH-25	Shelf FRU	A Shelf FRU Device ***shall*** be implemented in a hot swappable module.

5.3 Telco Alarm Module

ID	Feature Name	Requirements
SH-26	TAM	If a dedicated TAM is implemented, it ***shall*** reside on IPM controller and ***shall*** be connected to IPMB-0.

Refer to Chapter 12 for telco alarm connector pin assignments and voltage levels.

5.4 Backplane

The Backplane distributes power, Metallic Test Bus, ring generator bus, and low level shelf management signals through the Zone 1 connectors. The Base Interface, Fabric Interface, Update Channel Interface, and Synchronization Clock Interface signals are distributed through the Zone 2 connectors. This section covers the mechanical and electrical requirements for building a compliant backplane. The usage models for the Zone 2 interfaces are described in other parts of this document. The physical view of dual-dual star backplane is shown in Figure 5-2.

The mechanical specifications of the backplane are described in Section 2.5 of PICMG 3.0.

It is important that a backplane never fails and one way to improve reliability is to ensure that the backplane is entirely passive. The backplane should only contain traces, connectors, and a minimal amount of resistors and capacitors. By not putting any active components, even simple integrated circuits, it is possible to substantially improve the reliability of a backplane.

ID	Feature Name	Requirements
SH-27	Backplane	Backplanes ***shall*** not contain any active components.
SH-28	Backplane	The mean time to replace a failed backplane ***shall*** not be more than 3 hours. Note that the backplane repair typically will require the power and all the Front Boards and RTMs to be removed from the shelf.

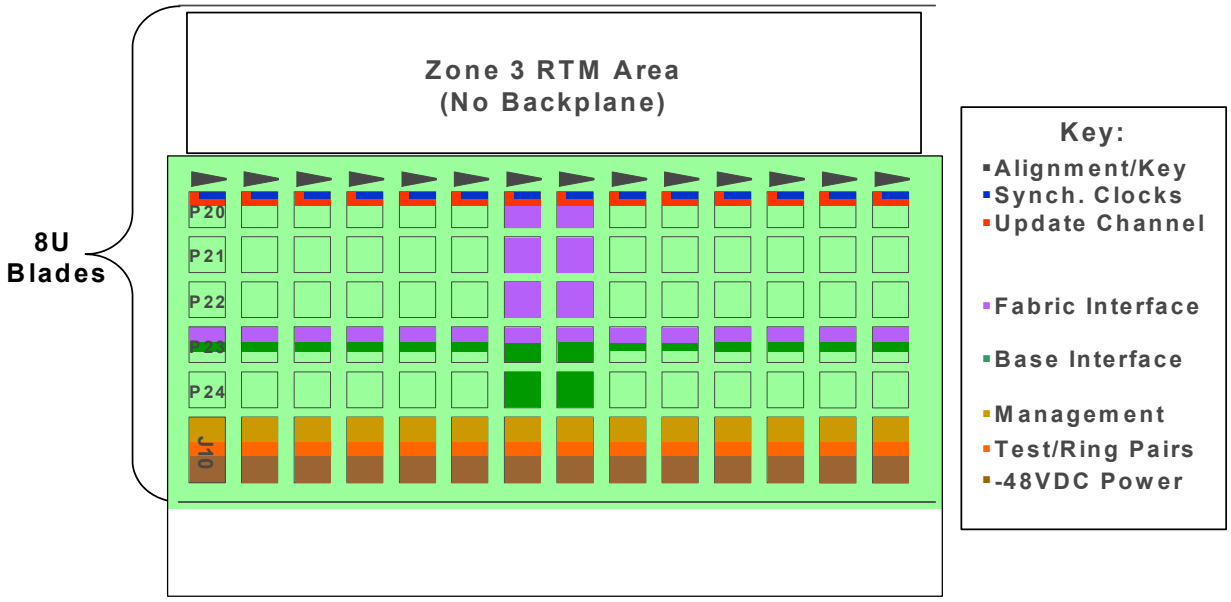


Figure 5-2 AdvancedTCA Backplane

5.4.1 Logical versus Physical slot mapping

The mapping between logical and physical slots in an Advanced TCA Backplane is implementation dependent.

ID	Feature Name	Requirements
SH-29	Backplane	Logical and physical Slot mapping ***should*** be optimized to reduce trace length over the backplane and thereby reduce cost for a backplane that supports only star topology.

ID	Feature Name	Requirements
SH-30	Back Plane	In support of PICMG 3.0 all shelves ***shall*** have slot numbering easily legible for the operator on the front subrack and the rear subrack.

By locating the Fabric Slots (lowest numbered logical slot) in the center slots of the backplane, the trace length is reduced by half on the backplane. This makes it possible to implement topology supporting up to 3.125 Gbaud signaling rates on a low cost printed circuit material. The mapping between logical and physical slots on this implementation is shown in Figure 5-3.

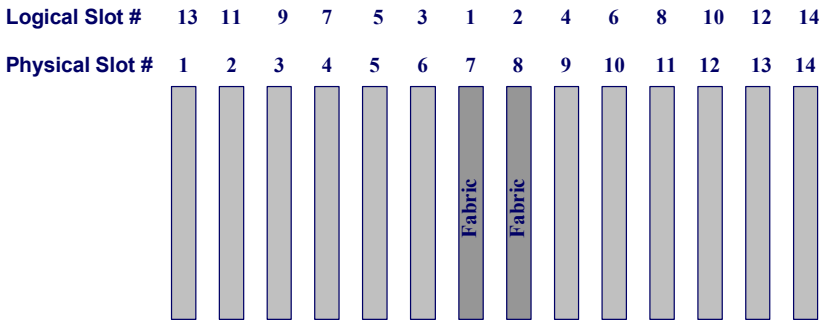


Figure 5-3 Logical and Physical Slot Mapping

5.4.2 Update Ports

Update Channels are private links between boards in any two slots. All PICMG 3.0 backplanes are required to support Update Channels (P20 connector).

ID	Feature Name	Requirements
SH-31	Backplane	A Backplane ***shall*** meet the requirements in PICMG 3.0 section 6.2.2 Zone 2 interface support requirements.
SH-32	Backplane	A Backplane ***shall*** connect the Update channels of redundant Hub slots together. In a backplane with two Hub slots, Logical slots 1 and 2 are connected.
SH-33	Backplane	A Backplane ***should*** route the Update Channel between non-adjacent slots to better support wide boards.
SH-34	Backplane	Update Channel interconnect diagram ***shall*** be provided on the shelf.
SH-35	Backplane	Update Channel interconnect diagram ***shall*** be visible with all boards installed.

5.4.3 Synchronization Clocks

ID	Feature Name	Requirements
SH-36	Backplane	Backplane ***shall*** support all the three redundant synchronization clocks as per PICMG 3.0 section 6.2.2 Zone 2 interface support requirements.

5.4.4 Backplane Routing Requirements

5.4.4.1 Zone 2 Routing Guidelines

This section defines the electrical requirements and design guidelines for Zone 2 interconnect on the backplane.

ID	Feature Name	Requirements
SH-37	Backplane	Backplane ***shall*** comply with the following PICMG 3.0 requirements: 1) Section 6.2.3.1 Base and Fabric Interface requirements 2) Section 6.2.3.2 Synchronization Clock Interface Requirements 3) Section 6.2.3.3 Update Channel Interface requirements
SH-38	Backplane	Fabric and Update Channel Interface ***shall*** be capable of supporting 3.125GBaud signaling rate.
SH-39	Backplane	Backplanes ***should*** use the following PICMG 3.0 backplane design guidelines: 1) Section 8.2.4 Base and Fabric Interface Backplane design guidelines 2) Section 8.2.5 Synchronization Clock Interface Backplane design guidelines 3) Termination resistors ***may*** be placed directly on the backplane at either end of the clock traces as shown in Figure 5-4

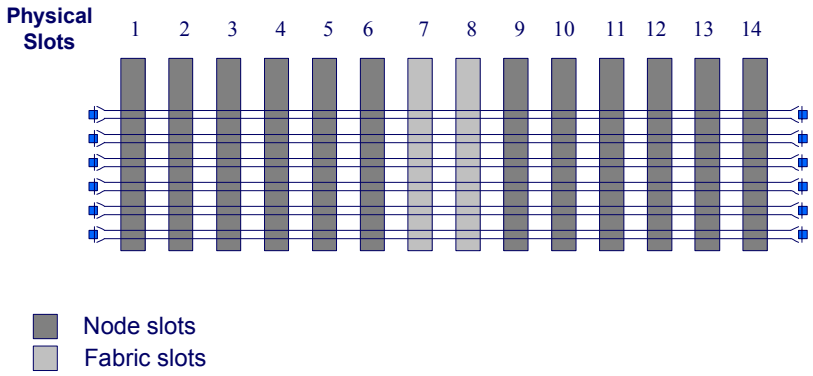
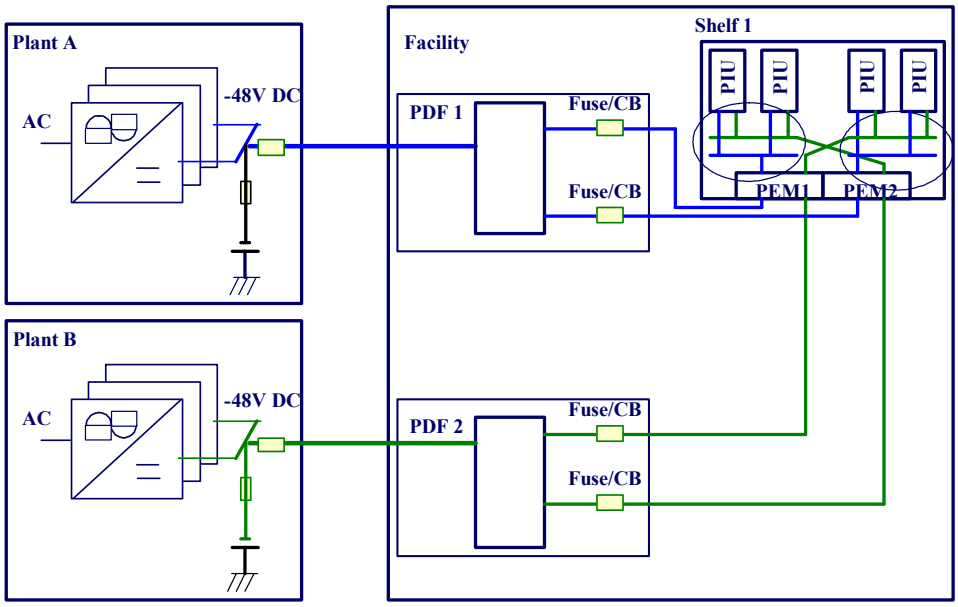


Figure 5-4 Synchronization Clock Terminations

5.5 Shelf Power Subsystem

Figure 5-5 shows power distribution to an AdvancedTCA shelf in a typical facility having two power plants. A shelf with two power domains is shown in Figure 5-6. To reduce feed cable size and impact of short circuit to the rest of the equipment, two feeds from each power plant powers the shelf as shown in Figure 5-6.This implementation provides plant, PEM, and feed level redundancy.



PDF: Power Distribution Frame

Figure 5-5 AdvancedTCA Shelf Power Distribution

Figure 5-6 illustrates a backplane power distribution with two power domains.

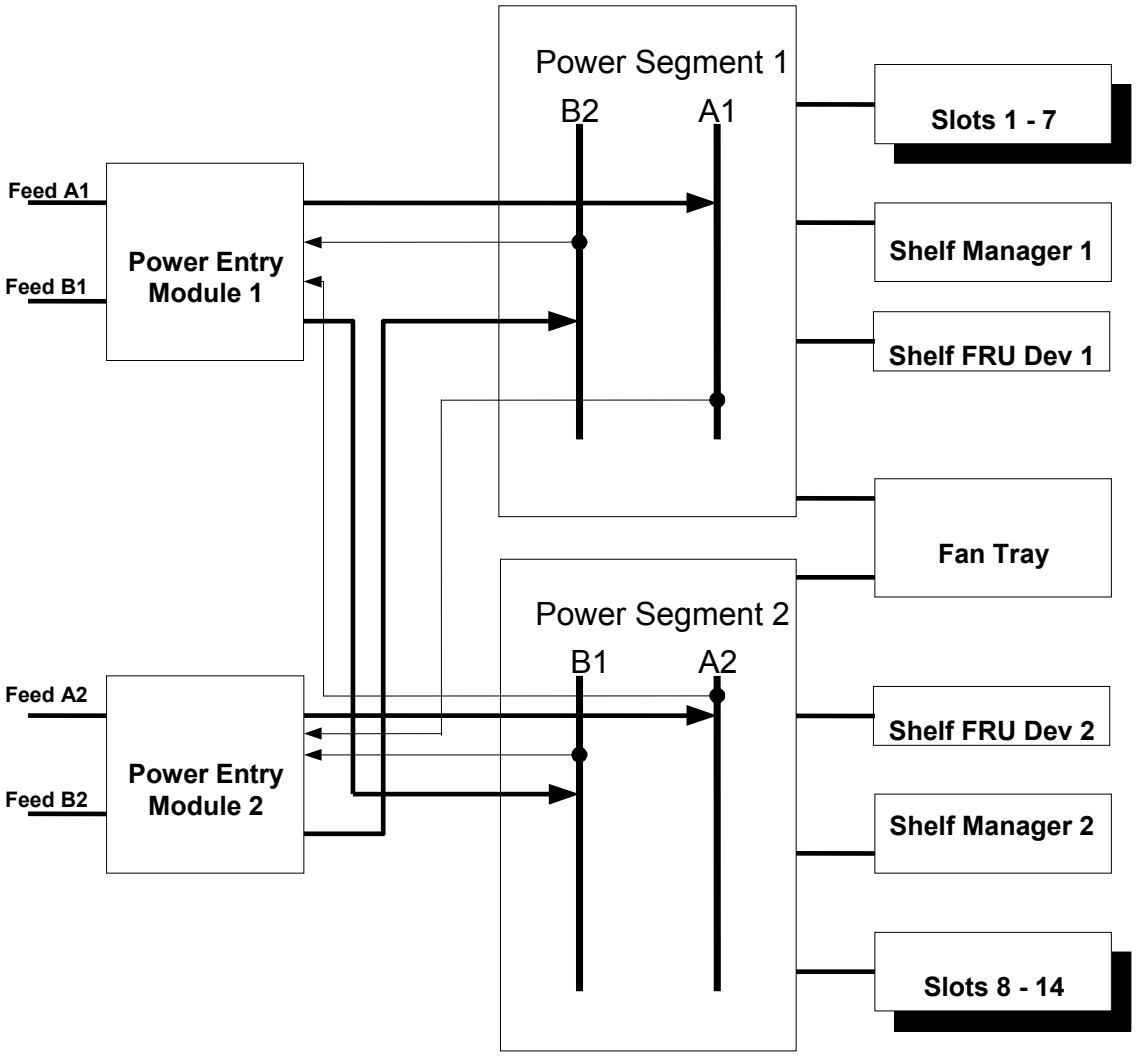


Figure 5-6 Backplane power distribution with two power domains for large shelf with more than 7 slots

ID	Feature Name	Requirements
SH-40	Shelf Power	A high capacity (greater than 7 slots) shelf ***should*** implement two power domains. This is required to keep the total current per feed below 60 amps so that the shelf can be easily deployed in the field.

The power distribution architecture supports redundancy of PEMs, power feeds, and power plants. The backplane is divided into two power segments of approximately equal power capacity. 1+1 redundant power planes power each segment and each power plane receives power from a separate power feed.

Segmented power planes have several advantages, especially in larger shelves. Since both the segments are fully isolated all the way to the power plants A and B, malfunction in one segment affects only the boards and modules connected to that segment. Segmented power planes also halves the feed capacity in a two-power segment implementation making cabling easier. It also results in lower short circuit current surges.

In this implementation, redundant modules such as Hubs, Shelf FRU Devices, and Shelf Managers are powered from separate power segments to ensure malfunction in one segment does not bring down the entire shelf.

A feed from each power plant is connected to each PEM. PEM1 receives power from feeds A1 and B1, while PEM2 receives power from feeds A2 and B2. This enables labeling of PEM feed terminals close to the terminals on the PEMs themselves using a common face plate.

5.5.1 Operating Voltage Range

<table>
<tr><th>ID</th><th>Feature Name</th><th>Requirements</th></tr>
<tr><td>SH-41</td><td>Shelf Power</td><td>A Shelf shall meet the following operating and non-operating voltage range:
<table>
<tr><th>Type of Operation</th><th>PICMG3.0 Specification</th><th>Design Guide Requirement</th></tr>
<tr><td>Operating</td><td>-44VDC to -72VDC</td><td>-39VDC to –72VDC</td></tr>
<tr><td>Non-Operating</td><td>0VDC to -40.5VDC & -72VDC to -75VDC</td><td>0VDC to < –39VDC & > -72VDC to –75VDC</td></tr>
</table></td></tr>
<tr><td>SH-42</td><td>Shelf Power</td><td>Application of non-operating voltage range shall not cause any damage to the equipment.</td></tr>
</table>

The Design Guide requires operation up to -39VDC at the lower limit compared to 40.5VDC specified in PICMG 3.0. This accommodates the requirements of certain carriers. During degraded operating voltage range, shelf operation can be impacted by 0V transients at the shelf input.

5.5.2 Sample Feed Capacity Calculation

Consider a 14 slot AdvancedTCA shelf with the following power budget that implements a dual power domain described earlier.

Board/Module	Power/Unit	Number of Units	Total Power
AdvancedTCA Slots	200W	14	2800W
Fan Tray	200W	1	200W
Shelf Managers	30W	2	60W
PEMs, Shelf FRU Device	15W	2 each	30W
Total Power			3090W

Total feed current at –39V = 3085W / 39V = 80A approximately

Since either power segment could power Fan Tray, the following calculation arrives at the maximum current per feed.

Total power without the Fan Tray load = 2885W

Total feed current without the Fan Tray load = 2885W / 39V = 74A approximately

Maximum current per feed without the Fan Tray load = 37A

Maximum Fan Tray current = 200W / 39V = 5A approximately

Since the Fan Tray could be powered by either power segment depending on its voltage level,

Maximum total current per feed = 37A + 5A = 42A

Apart from fans, power segment load imbalance could occur when a double wide module spans both the power segments, but derives all its power from a single power segment. Shelf Manager handles such cases by monitoring loading on a power segment basis. In the sample shelf shown in this guide, this is not an issue, since power segment transition occurs between the two Hub slots.

ID	Feature Name	Requirements
SH-43	Shelf Power	The shelf power distribution path ***shall*** be designed to support load current at -32VDC feed voltage.
SH-44	Shelf Power	IR drop from the Power Entry Module (PEM) input terminals to the farthest Front Board slot ***shall*** be 1V or less in either direction.
SH-45	Shelf Power	A Shelf ***should*** implement built-in power conditioning. It could be implemented in a Power Entry Module (PEM).
SH-46	Shelf Power	A Shelf ***shall*** support redundant feeds and PEMs. The PEMs provide the connections to the facility's power distribution and the shelf backplane. For high capacity shelves with power requirements up to 3kW per shelf use of traditional PDUs (Power Distribution Units) is not efficient. PDUs also take up additional rack space.
SH-47	Shelf Power	If a shelf implements dual power domains, then redundant boards (e.g. Hub, Shelf Manager) ***shall*** be powered from different segments.
SH-48	Shelf Power	FRU **should** generally not draw more power than was allocated during power negotiation. Startup currents **may** momentarily require more current, but the power draw **should** not exceed the allocated power budget by more than 20% during startup. This startup condition **shall** not last more than 25 ms and **shall** not occur more than once per FRU activation

5.5.3 PEM Block Diagram

The block diagram of a typical PEM implementation is shown in Figure 5-7. This PEM supports two feeds, one each from power plant A and B.

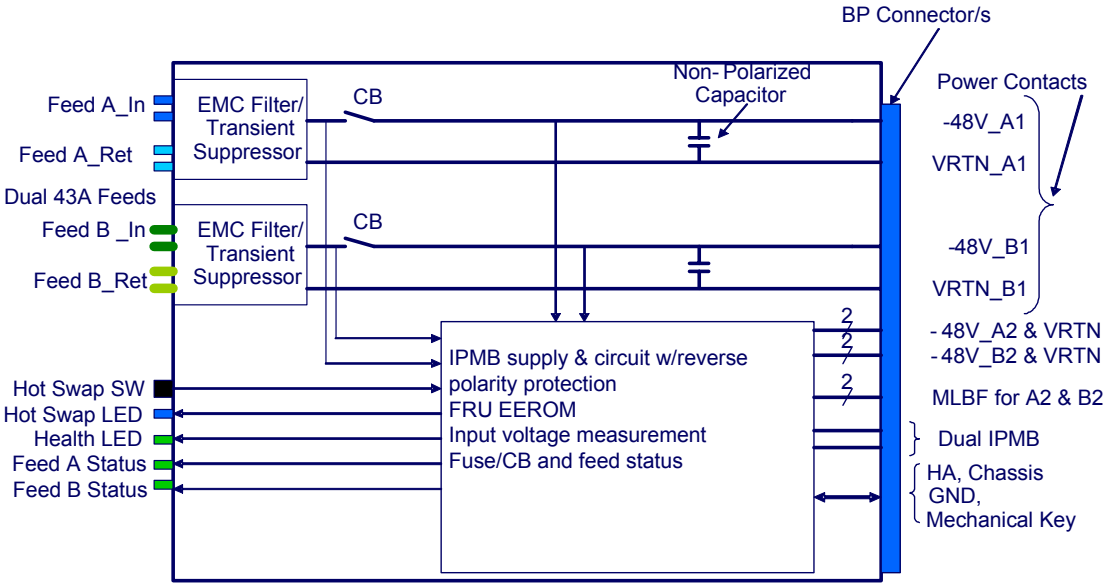


Figure 5-7 Typical PEM Implementation

Figure above shows the layout of a managed PEM

ID	Feature Name	Requirements
SH-49	Shelf Power	A PEM ***shall*** be swappable without interrupting service when redundant PEMS are installed in the shelf.
SH-50	Shelf Power	A PEM ***shall*** be managed by the Shelf Manager.
SH-51	Shelf Power	A PEM ***shall*** support over current protection.
SH-52	Shelf Power	A PEM ***shall*** support the detection of feed failure.
SH-53	Shelf Power	A PEM ***shall*** support the detection of the activation of the over current circuit.
SH-54	Shelf Power	A PEM ***shall*** support measurement of feed voltage to monitor input voltage.
SH-55	Shelf Power	A PEM ***shall*** provide support for a Health LED.
SH-56	Shelf Power	A PEM ***shall*** support conducted emission filtering, as described in Section 4.1.5.2 of PICMG 3.0.
SH-57	Shelf Power	A PEM control and status circuitry ***shall*** also derive its power from the redundant PEM. This is so that the input feed failures can be reported via management system.

ID	Feature Name	Requirements
SH-58	Shelf Power	A PEM ***shall*** support capacitive load to balance feed cable inductance to avoid ringing on power up.
SH-59	Shelf Power	A PEM ***shall*** support conducted emission filtering, as described in Section 4.1.5.2 of PICMG 3.0.

Reverse polarity protection of feed voltages is not recommended within the PEMs to reduce power dissipation and voltage drops.

5.5.3.1 Over Current Protection

Shelf level over current protection prevents catastrophic shorts from damaging the facility. Shorts on individual boards are protected via the board's fuses on the power and return lines that are mandated by the PICMG* 3.0 specification.

Circuit breakers also provide a convenient, reliable method of shutting off current flow to the shelf during installation or in lab environments.

ID	Feature Name	Requirements
SH-60	Shelf Power	A PEM ***should*** use a single pole circuit breaker (CB) for over current protection.
SH-61	Shelf Power	Circuit breaker trip point ***shall*** be rated at approximately 125% of the average maximum current rating.

Circuit breakers avoid the need for keeping an inventory of different types of fuses used in various AdvancedTCA and non-AdvancedTCA shelves.

5.5.3.2 Feed Cable Inductance Compensation

Mismatch between shelf impedance and power distribution system impedance can result in oscillations on application of power, short circuit or installation of a board/module. To avoid this, capacitors are provided within the PEMs. Total capacitance should be more than 0.3uF per watt of equipment load. For example, for a feed that supplies up to 1.5kW of power, the capacitance works out to be 500uF per feed.

ID	Feature Name	Requirements
SH-62	Shelf Power	The total capacitance provided in the PEMs to compensate for the mismatch between shelf impedance and power distribution system impedance ***should*** be more than 0.3uF per watt of equipment load.

The power supply system source impedance is a function of its rated capacity. The larger the capacity, the lower is the source impedance. This is because the power cable is sized to limit voltage drop in the supply path.

Figure 5-8 shows the Power-Supply system source impedance.

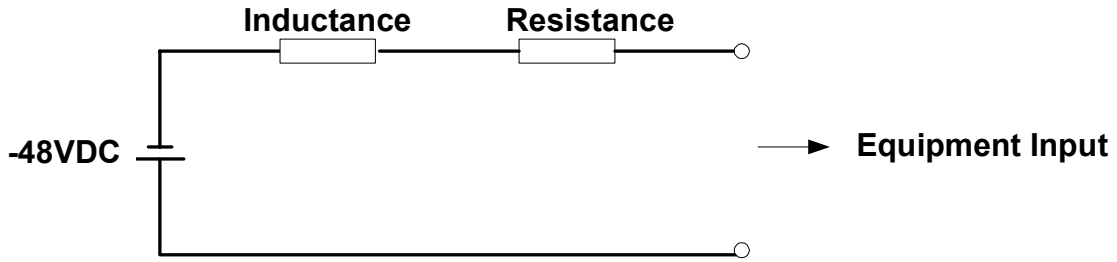


Figure **5-8** Power Supply System Source Impedance

R, L	Feed Power	
	0W – 750W	750W – 1500W
Resistance	30milliOhms	15.4milliOhms
Inductance	18.2uH	9.2uH

The values are based on a typical power distribution system and could vary from facility to facility.

5.5.3.3 Contact Sequencing

BP interface signals are shown in Figure 5-7.

ID	Feature Name	Requirements
SH-63	Shelf Power	The contacts ***should*** mate in the following sequence on insertion of PEM. On removal, the sequence is reversed: **Note:** The power plant feeds to the PEM being inserted or removed ***shall*** be turned off and the circuit breakers on the PEM ***shall*** be in the off position. 1. ESD Contact 2. Mechanical key 3. Shelf Ground 4. –48V_A, VRTN_A, -48V_B, VRTN_B to BP and from BP 5. IPMB, HA, Logic Ground 6. VRTN_MLBF_A & VRTN_MLBF_B

Sequencing of Plant A and Plant B –48V inputs from the BP (from the second PEM) to the PEM being inserted has not been implemented since the current draw is very minimal and a soft power-on device can be implemented. Also, sequencing of 48V input and return signals from the same feed is not implemented to reduce connector cost. This does not affect DC/DC converter operation. PICMG 3.0 compliant 8U boards require both of the above sequencing.

5.5.3.4 Decoupling of backplane power planes

Refer to Chapter 12 for Decoupling guidelines.

5.6 Hardware Management Interface

The hardware management interface is implemented on J10 connector on the backplane along with power and Ring Voltage signals. The connector, signals, and pin assignments are given in the PICMG 3.0 specification.

5.6.1 Hardware Addresses for Slots other than Front Board Slots

Hardware addresses for front board slots and Shelf FRU Devices are specified in PICMG 3.0. For all other modules including Shelf Managers, the hardware address is shelf implementation specific and is not specified by AdvancedTCA. Shelf FRU Information maps a hardware address to logical and physical addresses. Shelf Managers, Fan Tray, and PEMs were also assigned 8 bits (including parity) of hardware address for consistency with AdvancedTCA address implementation. This also enables PEMs and Fan Trays to be on the Dual IPMB bus in AdvancedTCA slots in other implementations. These address allocations are vendor specific.

ID	Feature Name	Requirements
SH-64	Hardware Management	Dedicated Shelf FRU slots ***shall*** implement the Hardware Addresses specified in Table 5-1.
SH-65	Hardware Management	Shelf infrastructure modules should implement the addresses listed in Table 5-1.This requirement assumes all platform infrastructure modules are on separate IPMB nodes.
SH-66	Hardware Management	The Shelf Manager or other module slots ***shall*** be mapped to 8h and 9h if they front the Shelf FRU Information to meet PICMG 3.0 requirement.

Table 5-1 Typical Hardware Addresses Slots other than Front Boards

Board/Module	HA[6:0]	Comments
Shelf FRU slot 1	8h	Seven bits plus parity
Shelf FRU slot 2	9h	
Fan tray slot 1	21h	
Fan tray slot 2	22h	
…		
PEM slot 1	31h	Seven bits plus parity
PEM slot 2	32h	
…		

5.7 Shelf Management

The Shelf Management subsystem consists of the shelf manager and the Intelligent Platform Management Bus (IPMB) which is used to connect the shelf managers to the intelligent boards and modules in the shelf.

5.7.1 Intelligent Platform Management Bus (IPMB)

There are also two implementation topologies for the IPMB: radial or bused topology. The radial implementation has a separate dual IPMB from the shelf manager to each intelligent FRU. In the bused implementation, there is a single dual IPMB that interconnects all intelligent FRUs to the shelf managers.

The advantages of the radial IPMB are:

- A private IPMB for each front board which provides better isolation and security
- Dedicated IPMB bandwidth to each device

The disadvantages of the radial IPMB are:

- More complex to implement

- Can only be implemented in a dedicated shelf manager

The advantages of the bused IPMB are:

- Simpler implementation
- Can be used by dedicated and non-dedicated shelf managers

The disadvantages of the bused IPMB are:

- Limit to the number of devices that can be placed on the bus
- All devices share the same IPMB bandwidth

An alternate implementation that takes advantage of both the radial and bus implementations is the hybrid topology, shown in Figure 5-9. This implementation has a Dual Star IPMB interconnect for Front Boards and dual bus IPMB interconnect for platform infrastructure modules such as PEMs, Fan Trays, and Shelf FRUs. Dual star hardware within a Shelf Manager is implementation specific. There could be separate IPMI controllers behind each IPMB node within a Shelf Manager. In an alternate implementation, a single IPMI controller could support all the IPMB nodes on each star using separate IPMB drivers in star configuration.

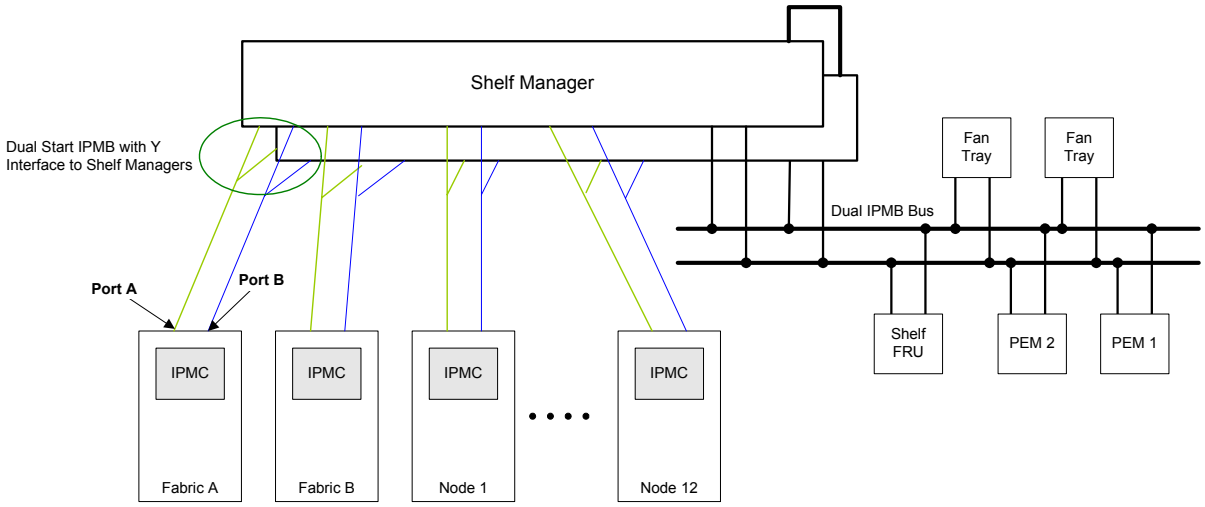


Figure 5-9 Hybrid IPMB Topology using Dedicated Shelf Managers

5.7.2 Shelf Manager

The Shelf Manager manages the shelf hardware. There are two Shelf Manager implementation models – Dedicated Shelf Manager and Non-Dedicated Shelf Manager. In the case of the Dedicated Shelf Manager, all shelf management tasks and applications are autonomously hosted on dedicated compute hardware. In the case of the Non-Dedicated Shelf Manager, all shelf management tasks and applications are hosted on compute HW that may be shared with either System Management software or non-management application software.

5.7.2.1 Dedicated Shelf Manager

It has the following advantages:

1. Does not require ShMC capabilities on any of the front boards, and hence provides greater flexibility in the selection of front boards
2. It limits fault domain. For example, if a fabric board also implements the Shelf Manager fails, it does not result in automatic fail-over of the Shelf Manager.

3. Supports both Hybrid and Dual Bus IPMB topologies
4. Telco alarms could be either integrated with the Shelf Manager or implemented in a separate module

It has the following disadvantages:

1. Requires dedicated Shelf Manager slots which takes up valuable shelf space

5.7.2.2 Non-Dedicated Shelf Manager

This implementation supports dual bus IPMB topology only. Each IPMB bus is connected to all the front boards as well as the platform infrastructure modules as shown in Figure 5-10. The platform infrastructure modules can be intelligent or managed by another module.

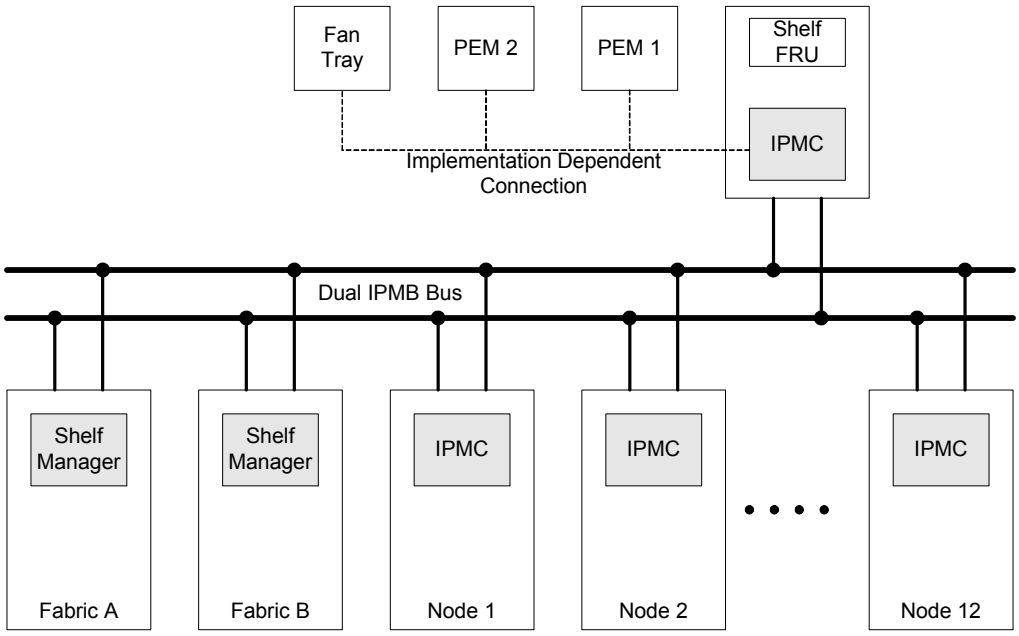


Figure 5-10 Non-Dedicated Shelf Manager

The non-dedicated shelf manager has the following advantages:

1. Neither dedicated Shelf Manager slots, nor dedicated Shelf Manager FRU modules are required
2. Low cost
3. Increased reliability through simple, homogeneous design of management subsystem. Simpler backplane IPMB routing
4. Reduction in number of FRUs within system both increases reliability and reduces inventory thus lowering system cost of ownership

In a typical implementation, Shelf Manager will be implemented as a mezzanine card on a Fabric Board. The Telco alarm module may also be implemented on, or in conjunction with, this fabric board.

The non-dedicated shelf manager has the following disadvantages:

1. The disadvantages associated with the tight coupling of Shelf Manager with a Front Board.

ID	Feature Name	Requirements
SH-67	Shelf Manager	A Shelf with less than or equal to 7 slots ***may*** support one Shelf Manager.
SH-68	Shelf Manager	A Shelf with more than 7 slots ***shall*** support two Shelf Managers. When a smaller shelf is deployed, redundancy is often provided at the shelf level.
SH-69	Shelf Manager	A Shelf with a dedicated Shelf Manager ***may*** support either Hybrid or Dual Bus IPMB Shelf Managers and topologies.
SH-70	Shelf Manager	A Shelf ***should*** support dedicated Shelf Manager.
SH-71	Shelf Manager	Shelf Managers ***should*** be accessible from the front.

Note: Pin assignments, mechanical specifications and other implementation details of dedicated and non-dedicated Shelf Managers may be added in a later draft of this Design Guide.

5.8 Shelf Cooling Subsystem

This section provides requirements and guidelines for meeting the cooling requirements of the AdvancedTCA MCP shelf. The shelf will use forced air convection cooling with air drawn from the front of the shelf, filtered to meet GR-063-CORE requirements, and forced over the front boards and RTMs before exiting through the rear. Multiple fans or blowers shall be used to provide the required volume and reliability. The shelf shall provide complete seals on all joints and faceplate edges to allow for more efficient cooling.

5.8.1 Shelf Airflow Management

According to PICMG 3.0, the maximum power dissipation for a single slot is 200W and is 400W for a double wide slot. This amount of heat has to be removed from hot components by the airflow through the slot propelled by air movers without causing the average airflow temperature at the chassis exit to be more than 65~70^{o}C. For the short term catastrophe failure scenario, chassis inlet airflow temperature is 55^{o}C, that means, the maximum airflow temperature rise across the chassis between inlet and exit shall not exceed 10~15^{o}C to meet with thermal environment requirements (PICMG 3.0 Table 5-1). PICMG 3.0 Fig 5-3 has characterized slot power dissipation rate vs. volumetric rate of cooling air. It shows the minimum airflow requirement is 37CFM to cool 200W per slot at the sea level. This volumetric airflow rate CFM (cubic feet per minute) varies upon the airflow density at different sea elevation. The total airflow will be distributed to cool front board and RTM. The PEM usually does not need dedicated airflow from the slot. The airflow partition for the Front board and RTM is governed by the following equations.

Volumetric airflow rate for Front board is

$$\dot{V}_{FB} = \left(\frac{P_{FB}}{P_{FB} + P_{RTM}}\right) \times CFM_{slot}$$

Volumetric airflow rate for the RTM is

$$\dot{V}_{RTM} = \left(\frac{P_{RTM}}{P_{FB} + P_{RTM}}\right) \times CFM_{slot} \text{ or: } \dot{V}_{RTM} = CFM_{slot} - \dot{V}_{FB}$$

Where:

P_{FB} and P_{RTM} are the total TDP (thermal dissipation rate) on front board and RTM in the same AdvancedTCA slot,

CFM_{slot} is the total airflow introduced into the slot.

Airflow partition may be controlled by dedicated venting, ducting or blocking mechanism at chassis air inlet plenum, fan tray, subrack or backplane.

For example, in an AdvancedTCA slot, there are 185W TDP heat generation on front board and 15W heat on RTM. The slot airflow requirements to cool 200W at sea level within 10^{o}C air temperature rise envelop are: 2.78CFM for the RTM and 34.22CFM for the front board which add up to 37CFM shown in Figure 5-11.

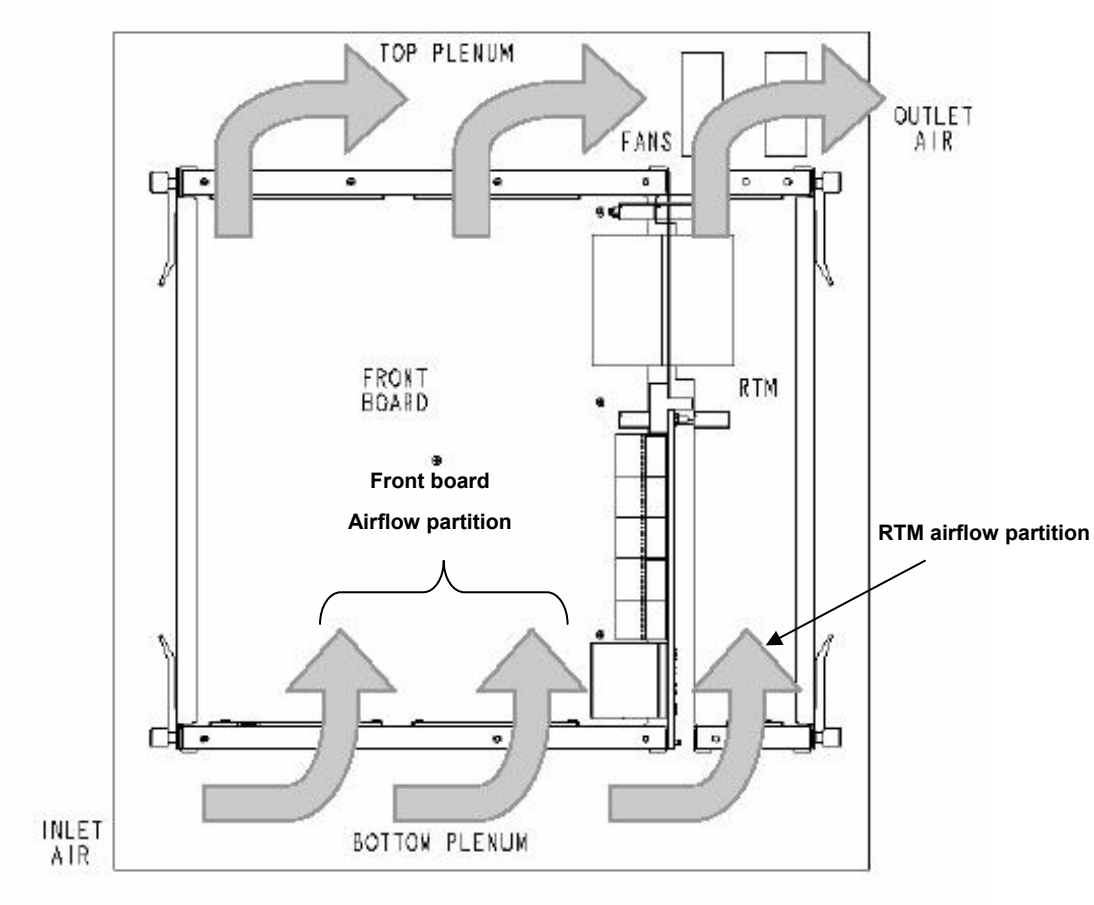


Figure 5-11 Airflow partition in the Shelf

For the most stringent chassis design, a slot requires 200W cooling capacity for front board with and without the presence of the RTM. Dedicated air flow for RTM cooling is highly recommended for this scenario.

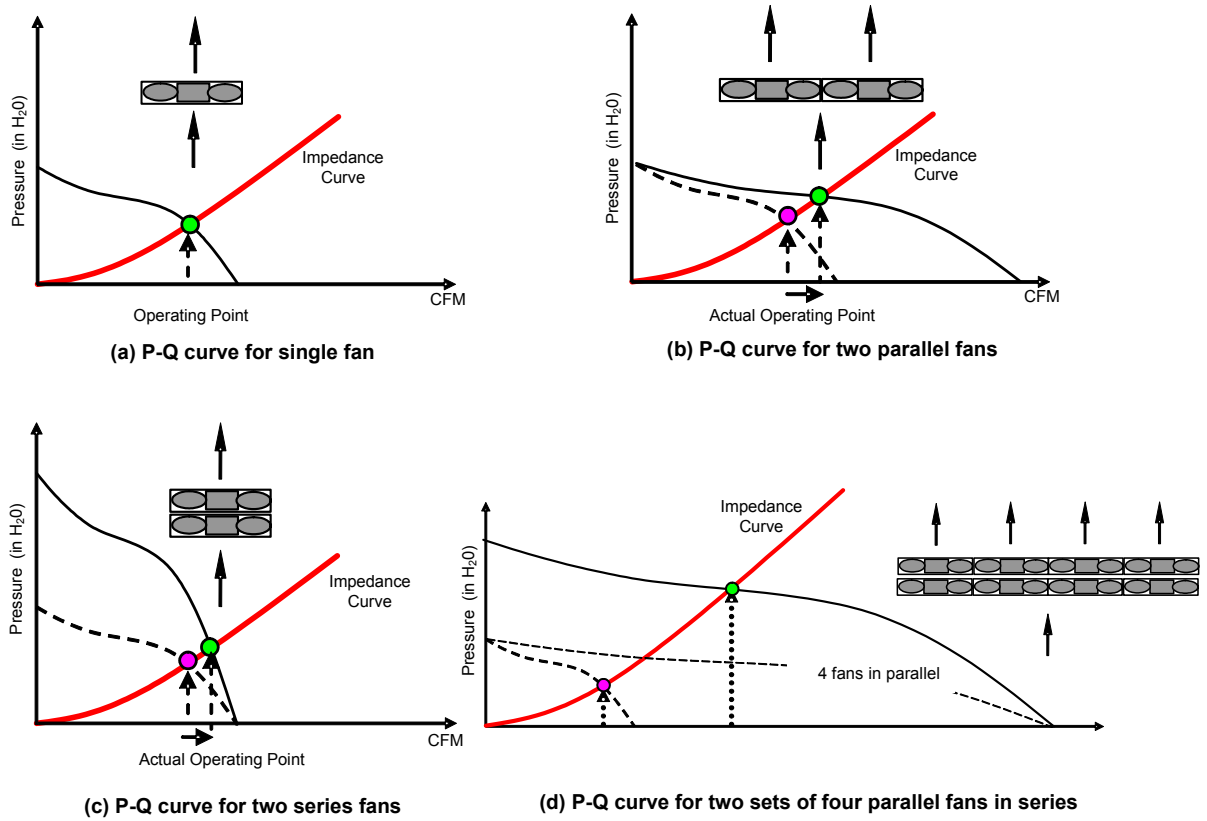
5.8.1.1 Fan design

Shelf cooling capacity is a dependent of air mass flow rate through chassis slots. Airflow rate is a function of system impedance, airflow density associated with sea level and airflow mover P-Q characterization. Chassis air cooling capacity is also determined by the maximum air temperature rise allowance

5.8.1.1.1 Fan set configurations and system performance

Figure 5-12(a) displays a single fan P-Q curve and system impedance curve. P-Q curve characterizes volumetric airflow (CFM) at different air pressure through a fan. A fan may have different fan curve at different fan speed (RPM). The general fan design guideline is to choose a fan with large static CFM and pressure. A fan may perform differently upon total chassis airflow resistance represented by chassis

impedance. A chassis impedance is characterized by airflow resistance generated at chassis air entry and exhaust plenum, filters, blade/RTM and its components/modules, cables, filter, EMI shielding grid, PEM, shelf management modules, etc. Actual airflow through a fan to chassis is dominated by system operating point, which is the cross point of P-Q curve and system impedance curve. Measures to reduce chassis impedance increases actual airflow rate from the fan through the chassis. To enhance the airflow performance in a chassis, a group of fans are designed for a multi-slot chassis. There are two generic fan combinations: fan in series (Figure 5-12(b)) and fan in parallel (Figure 5-12(c)). Fan in parallel can double the maximum fan airflow capacity but not the actual airflow rate due to non-linearity of the impedance a P-Q curves. Maximum backpressure remains unchanged. Fan in series set may double the maximum fan pressure, but it cannot change maximum airflow capacity. With the enforced air pressure, the actual airflow may be improved (Figure 5-12(c)). Figure 5-12(d) is an array of eight fans with two rows of four parallel fans in series. The actual airflow greatly increased.



(a) P-Q curve for single fan

(b) P-Q curve for two parallel fans

(c) P-Q curve for two series fans

(d) P-Q curve for two sets of four parallel fans in series

Figure 5-12 P-Q Curves for Fan Configurations

When two coupled fans are integrated in series into one unit, the fan group performance is very sensitive to the distance between the two fans due to the airflow interference at front fan exit and back fan entry. The key parameters to impact the optimal distance are fan hub diameter, fan geometry and RPM.

5.8.1.1.2 Redundancy requirement

In order to satisfy system reliability requirement, an AdvancedTCA chassis design must be able to provide adequate airflow to dissipate 200W per slot without making the exit airflow temperature exceed 70°C system even when any single fan fails. Each fan is coupled with a back up fan to meet this requirement. During the normal state, every fan operates at low speed. When a fan failure occurs, the chassis air temperature sensor will detect a rise in chassis temperature and the shelf manager will notify the fan controller to increase the speed of the fan. There are two scenarios for system cooling:

1. Operating short term failure at -5°C to 55°C ambient in less than 96 consecutive hours, less than 15 days/year and less than 15 occurrences in one year.
2. Single cooling unit fails under the normal operating conditions with 5°C to 40°C ambient temperature.

For the 1st worst case scenario, every fan operates at full speed. For the 2nd worst case scenario, at least the coupled fan for the failed fan operates at full speed to provide the airflow and the cooling must remain unchanged.

5.8.1.1.3 Fan configurations

There are a variety of fan design configurations. But there are three basic types for redundant fan design:

- Fans at chassis inlet air plenum ("push") as Figure 5-14.
- Fans at chassis outlet or exhaust air plenum ("pull") as Figure 5-13.
- Fans at both inlet and outlet plenums ("push and pull") as Figure 5-15. Fans are placed either horizontally or vertically outside or inside chassis plenum.

Comparisons between push fan design and pull fan design:

1. Push fan has less dusting issues because it builds positive air pressure inside chassis to prevent dust from unsealed leakage around chassis.
2. Push fan introduces air impingement or better cooling at blade front edge where critical components such as CPU are located.
3. The disadvantages of push fan are; the fan motor heat generations are added into the chassis which may degrade the cooling capacity. It may also need additional industrial design for the protruded front fan setting.
4. With a pull fan, the filter and fan are at the opposite ends of the chassis which gives much more freedom and simplicity in the design versus a push fan design where the filter and fan must compete for the same space which increases complexity in both design and service.

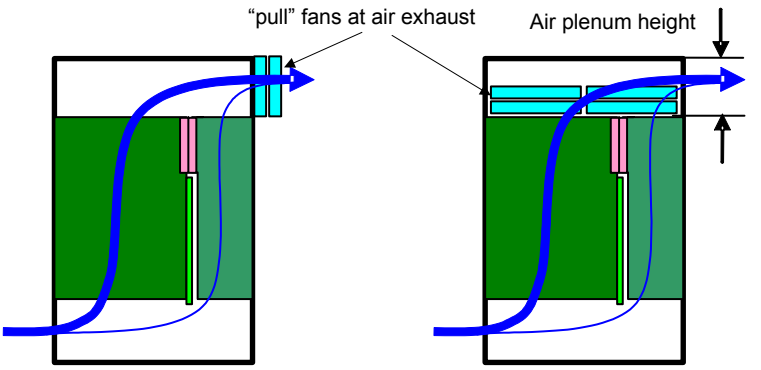
Comparisons between push or pull and push-pull fan design:

Push-pull fan design has least airflow interference issue for coupled fans like two fans integrated together. But it needs two sets of fan trays and fan control modules and connectors that would increase manufacturing cost and reduces system reliability. It may also have an issue on airflow pattern inconsistencies when a fan is in failure mode that would result in insufficient cooling for the components (U1) at the regions impacted by airflow deviation (Figure 5-16).

Comparison between horizontal and vertical fan designs:

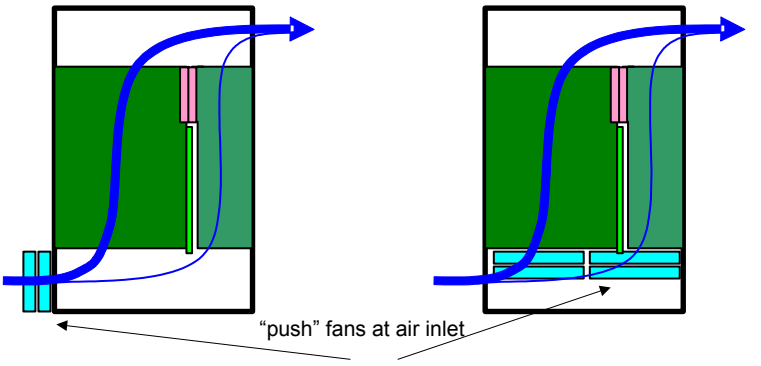
Horizontal coupled fan design as shown in Figure 5-13(b) and Figure 5-14(b) is good if there is enough air plenum height to fit thickness of two fan, distance between two coupled fans and airflow gap between fan group to chassis top/bottom. This is very difficult to realize especially for the small form factor chassis platforms.

Vertical coupled fans have less airflow resistance due to low back pressure with the requirement being small diameter fan. And usually when the fan diameter is small, fan performance is low. This can be compensated by using multiple arrays of fans that would provide more uniform airflow to each slot compared with few large fans.



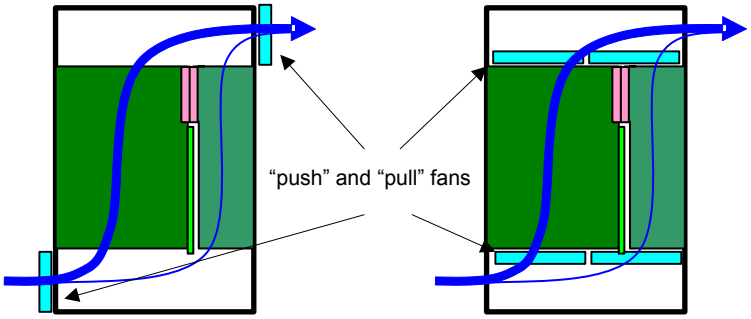
(a) Vertical fan sets (b) Horizontal fan sets

Figure 5-13 Fan at air exhaust plenum ("pull")



(a) Vertical fan sets (b) Horizontal fan sets

Figure 5-14 Fan at air inlet plenum ("push")



(a) Vertical fan sets (b) Horizontal fan sets

Figure 5-15 Fan at both air inlet/exhaust plenums ("push and pull")

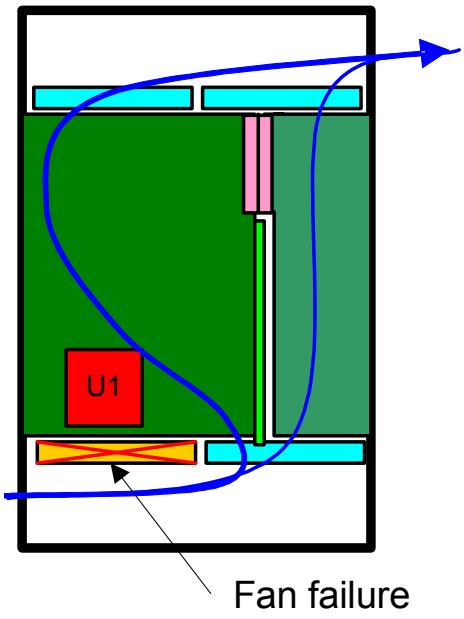


Figure 5-16 Airflow path change leaves airflow dead zone

5.8.2 Cooling Requirements

The previous section should be used to design the cooling subsystem to meet the requirements outlined in this section. The hottest ambient temperature the service personnel are expected to replace components in the chassis is 40 degree C. To meet the NEBS requirements, the boards need to survive 55 degree C during the fan swap operation. The boards should have a 15 degree margin.

ID	Feature Name	Requirements
SH-72	Shelf Cooling	Aisle facing surface temperature ***shall*** not exceed 38C @ 26C ambient, GR-63-CORE, O4-15.
SH-73	Shelf Cooling	Sound level ***shall*** not exceed 65dBA @ as per GR-63-CORE, R4-72 at 26C ambient. This limit can be exceeded under single fan failure or higher ambient temperature conditions.
SH-74	Shelf Cooling	Sound level ***shall*** be less than 7.2 bels sound power as per ETS 300 753:1997.
SH-75	Shelf Cooling	Blank panels/fillers ***shall*** be provided to ensure desired air flow in a partially populated shelf. This includes absence of PEM, blades, shelf manager module, or other components that could impact air flow within the shelf.
SH-76	Shelf Cooling	Any single failure in the shelf's cooling system ***shall*** not affect the operation of any blade in the chassis at an incoming ambient temperature of 40 degree C. In order to account for HVAC failures, a shelf with all fans operating at 55 degree C shall not cause any blade to fail.

ID	Feature Name	Requirements
SH-77	Shelf Cooling	A Shelf ***shall*** be capable of cooling 200W of maximum average power in each front slot, in a typical configuration over the operating temperature and altitude range.
SH-78	Shelf Cooling	A Shelf ***should*** be capable of cooling 20W of maximum average power in each RTM slot, in a typical configuration over the operating temperature and altitude range.
SH-79	Shelf Cooling	Shelf vendor ***shall*** provide Slot Impedance Curve as per Section 5.4.2.1 and Slot Fan Flow Curve as per Section 5.4.2.2 of PICMG 3.0 specification. It shall be generated for the worst case conditions described above.
SH-80	Shelf Cooling	The shelf cooling system ***should*** be designed so that the fans typically run at somewhere between 50% and 75% of their maximum speed during normal operation. This results in lower noise, higher average fan life, and thermal headroom so that it can increase the fans to full speed if the facility's air conditioning fails or if a fan fails. For acoustics, the typical operating noise ***should*** be under 65 decibels, according to NEBS.
SH-81	Shelf Cooling	Fan Trays ***shall*** be front accessible.
SH-82	Shelf Cooling	The mean time to replace a failed fan unit without impacting service ***shall*** be less than 2 minutes while the components are running.
SH-83	Shelf Cooling	A Shelf ***shall*** support an air filter as per NEBS GR-63-CORE requirements R4-64 through R4-69.
SH-84	Shelf Cooling	A Shelf Manager ***shall*** be able to detect the presence of an air filter.
SH-85	Shelf Cooling	A Shelf ***shall*** allow the air filter to be easily inspected by a technician and replaced without interrupting the operation of the shelf.
SH-86	Shelf Cooling	The mean time to replace the filter ***should*** be 15 minutes.
SH-87	Shelf Cooling	Fan trays **shall** run at 100% if communication to the shelf manager is lost.

Note that NEBS is in the process of redefining the acoustic noise specification in term of sound power. System Integrator usually measures acoustic noise on a fully configured shelf.

In order to maximize surface area, reduce airflow resistance and to enable filter replacement without introducing accumulated dust from the filter into the shelf, the air filter could be placed in a tray which is pulled out for filter replacement.

5.9 Regulatory, Environmental, Reliability and Compliance Requirements

ID	Feature Name	Requirements
SH-88	Regulatory, Environmental, Reliability and Compliance Requirements	A Shelf ***shall*** comply with NEBS Level 3 and ETSI Specifications. Refer to Chapter 11 Certification and Practices for details.

6 AdvancedTCA Boards

Two board form factors are described in detail in PICMG 3.0 specifications:

- Front Boards containing the desired electronic functions and connectors required to interface with these functions.
- Rear Transition Modules (RTMs) providing user defined input and output connectivity to the companion Front Board from the rear.

All AdvancedTCA Front Boards must have the following common elements: front panel, ejector handles, hot-swap interface (switch and blue LEDs), IPMC, power interface, temperature sensors, payload and Zone 1 connector to the backplane. In most cases, the board will also have a Zone 2 connector to the backplane to connect to the base and fabric interfaces, update channels and the clock signals. Payload is the general term used to include the hardware on the board except that associated with IPMI management. Chapter 2 described the architecture of the payload types used in network elements, namely Control, Protocol Processing, Storage, Hub, Service and Line cards. The payload may also include standard modules such as PMC, PrPMC and AMC to implement the required functions.

This chapter specifies additional requirements for implementing the common front board elements to promote a high degree of interoperability and standardization, as well as requirements for the various payload options. The essential components and interfaces of the Front Board assembly are shown in Figure 6-1.

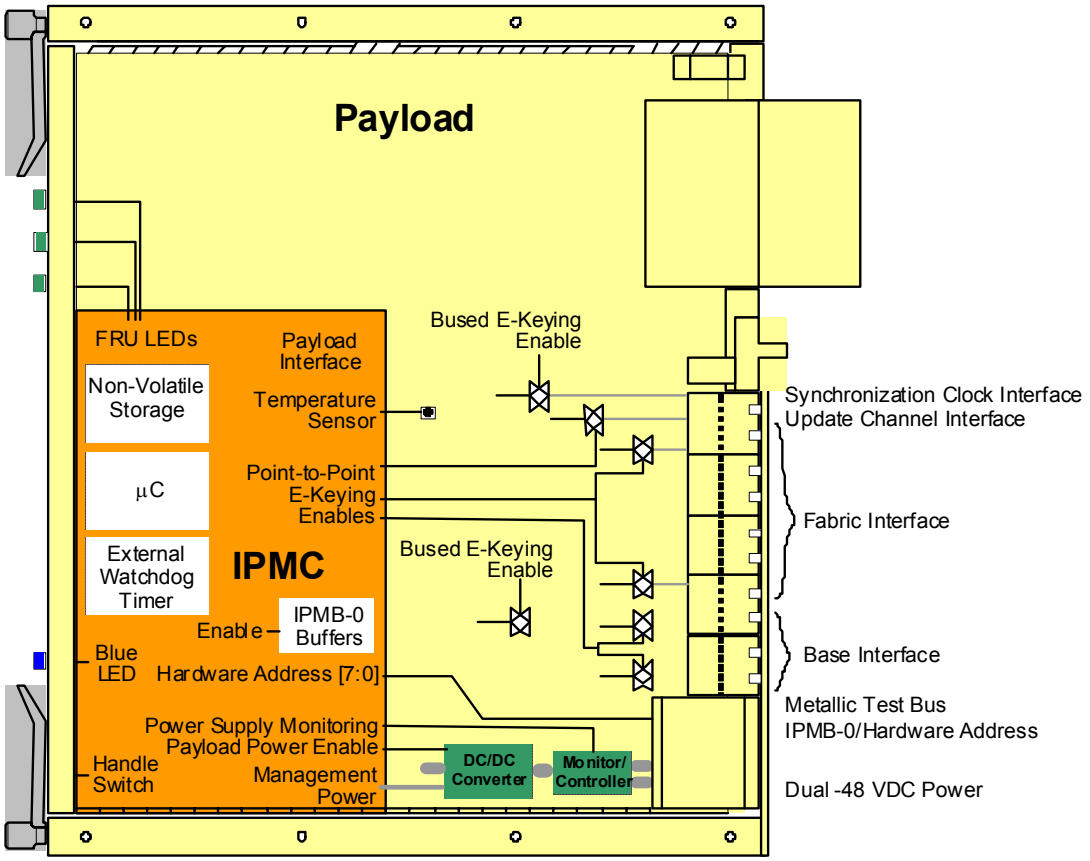


Figure 6-1 AdvancedTCA Front Board Assembly (not to scale)

6.1 Front Board Assembly Common Requirements

The primary goal of this section is to provide additional requirements and guidelines to ensure that the Front Board is designed to operate in a wide range of environments. The key to interoperability is to ensure that the mechanical, thermal, power, and management interfaces of the front board are specified with few options. In addition, the standardization of the interfaces around some of the common components such as the management circuitry will promote greater reuse across boards and hence drive down the overall cost of the boards as well as improve reliability. A secondary goal of this section is to define interface requirements for these internal building blocks.

6.1.1 Mechanical

The mechanical specification of Front Boards is covered in Chapter 2 of PICMG 3.0. There should be a high degree of interoperability if all the mandatory requirements are satisfied.

6.1.1.1 EMC Gasket

Fabric covered gaskets are recommended over beryllium copper gaskets.

ID	Feature Name	Requirements
BD-1	Front Board Mechanical	Front boards ***shall*** use fabric gaskets to meet the Face Plate EMC gasket requirements defined in PICMG 3.0.
BD-2	Front Board Mechanical	Front Board EMC gaskets ***shall*** withstand 500 insertion- / extraction-cycles without degradation. The slot on the left of the slot being tested shall be populated during the test.

6.1.1.2 Handles

Face plate handles are specified in Section 2.2.7.1 of PICMG 3.0. MCP boards shall also satisfy the requirements below.

ID	Feature Name	Requirements
BD-3	Front Board Mechanical	The handles on Front boards ***shall*** have a latching device.

6.1.2 Look and Feel

Since most network elements will likely be made up of boards from more than one supplier, it is important that the physical appearance of the boards maintain the same look and feel among the suppliers. This section provides additional guidelines that should be followed by board designers to ensure that the system integrator is able to easily customize their boards to maintain a uniform look and feel across their system.

ID	Feature Name	Requirements
BD-4	Front Board Mechanical	The board vendor ***may*** provide support for attaching faceplates provided by the system integrator.

6.1.2.1 Face Plate Labels

The PICMG 3.0 specification provides a list of recommended guidelines for the placement of commonly used labels to provide a more consistent look among vendors Front Boards and RTMs. These guidelines ***shall*** be mandatory for MCP Front Boards and RTMs.

ID	Feature Name	Requirements
BD-5	Front Board	If symbols are used to label LEDs and connectors on the face plate, then

ID	Feature Name	Requirements
	Mechanical	industry recognized symbols ***shall*** be used.
BD-6	Front Board Mechanical	The space on the Front Panel ***shall*** be reserved for the barcodes of the system integrator or network equipment provider.
BD-7	Front Board Mechanical	The equipment manufacturer ***may*** place any barcodes on the Front Panel for the system integrator before shipping the board.

6.1.2.2 LEDs

ID	Feature Name	Requirements
BD-8	Front Board Mechanical	All LEDs on MCP based boards ***shall*** comply with the mandatory and recommended requirements specified in chapter 2 and 3 of PICMG 3.0.

6.1.3 Thermal Guidelines

The PICMG 3.0 specification allows up to 200W of power to be delivered to the Front Board. This power will be dissipated into microprocessors, chipsets, modules, and other electrical components on the board and generate heat. The generated heat will cause silicon, PCB, and other electrical component temperature to rise. The components need to be adequately cooled so that they continue to operate within their thermal specifications.

The components on the boards are cooled by forced convection with the assistance of fans and blowers on the shelf. In an ideal situation, there would be even air flow across the slots and across the boards. Since air takes the path of least resistance, it is important to place components so that the air flow is uniform across all slots and across the board. Since the cool air enters the shelf from the bottom, components such as processors that generate the most heat should be placed at the bottom.

ID	Feature Name	Requirements
BD-9	Front Board Thermal	Front Board components that generate the most heat such as processors ***should*** be placed in areas of the board that receive highest airflows.
BD-10	Front Board Thermal	Front Board components ***should*** be placed to ensure that the airflow is uniform across the board.
BD-11	Front Board Thermal	Front Board or RTM vendor ***shall*** clearly document performance degradation, if any, under a single failure condition in the cooling system.

6.1.4 Power Interface

All Front Boards receive their power from dual -48V DC feeds on the Zone 1 connection to the backplane. The specification of the power interface is described in chapter 4 of PICMG 3.0. This section provides additional requirements and guidelines for the power interface for all MCP Front boards.

Table 6-1 Board Operating Voltage Range

Type of Operation	PICMG3.0 Specification	Design Guide Requirement
Normal	-39.5VDC to –72VDC	-38VDC to –72VDC
Non-Operating	0VDC to –39.5VDC & -72VDC to –75VDC	0VDC to < –38VDC & > -72VDC to –75VDC

Some carriers require operation below –40.5VDC and require the Shelf Manager to gracefully shut down boards when the shelf input voltage remains less than –39VDC for more than 100ms. Hence the lower limit recommended for board input voltage is –38VDC, which gives 2V margin for drops across backplane connector, OR-ing diodes, fuses and filter inductors. Most DC/DC converters are specified to work at –36VDC or below based on the loading.

AdvancedTCA boards shall use hot swap controllers to limit in rush current to bulk capacitors on power up and insertion. They shall not use EARLY power signals.

Hot swap controller provides the following functions:

1) Limits in-rush current to PICMG 3.0 specified limits
2) Implements over and under voltage shut down
3) Enables DC/DC converter only after the board is fully seated

ID	Feature Name	Requirements
BD-12	Front Board Power Interface	A Shelf ***shall*** meet the operating and non-operating voltage range shown in Table 6-1.
BD-13	Front Board Power Interface	Boards that use more than 40W ***should*** use hot swap controllers to limit in rush current to bulk capacitors on power up and insertion instead of using the EARLY power signals.
BD-14	Front Board Power Interface	Boards ***shall*** be able to detect a blown fuse in the power circuitry and report this to management.

6.1.5 Management Circuitry

All Front Boards are required to have management circuitry which is responsible for the low-level hardware management of the board. The management circuitry consists of an Intelligent Platform Management Controller (IPMC), non-volatile storage, bus isolator, and an external watchdog. The requirements of the IPMC and its interaction with the Front board and Shelf Manager are described in chapter 3 of the PICMG 3.0 specification. This section provides additional requirements and guidelines to ensure that the board management circuitry is able to interoperate with Shelf Manager and the payload.

Figure 6-1 shows the components of the management circuitry and the main interfaces. At the heart of the management circuitry is a microcontroller sub system which provides the core services required by the IPMC. The management circuitry should also provide non-volatile storage that is used to store FRU information, SEL, and SDR. The non-volatile storage should use a SEEPROM which normally has a longer life than Flash devices.

6.1.5.1 System Event Log (SEL)

ID	Feature Name	Requirements
BD-15	Front Board IPMC	The IPMC ***shall*** implement a local SEL of at least 4KB to keep local copies of events generated by the board if payload processor and interface to IPMC is available.
BD-16	Front Board IPMC	The local SEL ***shall*** be manageable by the payload processor if payload processor and interface to IPMC is available.

6.1.5.2 Payload Interface

The payload interface enables communications between the IPMC and its payload processor(s). For maximum implementation flexibility, AdvancedTCA leaves the details of the Payload Interface entirely up to the implementer. This has a negative impact on interoperability. In order to improve the interoperability across this interface, all compliant boards shall provide a common API in the board support package to communicate over the payload interface. This will allow the system integrator to easily port the board application to a different board.

The list of commands that shall be supported by the IPMC on the Payload Interface is shown in Table 6-2. Each command is grouped by general function, with an indication of whether the command is mandatory (M), optional (O), or Not Required (No Req) for AdvancedTCA IPM Controllers and Payload Interface.

Table 6-2 Payload Interface IPMI Commands

	IPMI Spec. section	NetFn	CMD	IPM Controller req.	Payload Interface Requirement	Additional Notes and Requirements
IPM Device "Global" Commands				**M**	**M**	
Get Device ID	17.1	App	01h	M	M	
Cold Reset	17.2	App	02h	O	O	If IPM controller implements this command, it should allow access of this command via payload interface. This command should not affect the state of the payload
Warm Reset	17.3	App	03h	O	O	If IPM controller implements this command, it should allow access of this command via payload interface. This command should not affect the state of the payload
Get Self Test Results	17.4	App	04h	M	M	
Manufacturing Test On	17.5	App	05h	O	O	
Set ACPI Power State	17.6	App	06h	O	O	
Get ACPI Power State	17.7	App	07h	O	O	
Get Device GUID	17.8	App	08h	O	O	

	IPMI Spec. section	NetFn	CMD	IPM Controller req.	Payload Interface Requirement	Additional Notes and Requirements
Broadcast "Get Device ID	17.9	App	01h	M	No Req	Should only be accessible via IPMB
BMC Watchdog Timer Commands				**M**	M	
Reset Watchdog Timer	21.5	App	22h	M	M	
Set Watchdog Timer	21.6	App	24h	M	M	
Get Watchdog Timer	21.7	App	25h	M	M	
BMC Device and Messaging Commands				**O**	M	O/M: If any in this group is implemented then O/M is mandatory.
Set BMC Global Enables	18.1	App	2Eh	O/M	M	
Get BMC Global Enables	18.2	App	2Fh	O/M	M	
Clear Message Flags	18.3	App	30h	O/M	M	
Get Message Flags	18.4	App	31h	O/M	M	
Enable Message Channel Receive	18.5	App	32h	O	O	
Get Message	18.6	App	33h	O/M	M	
Send Message	18.7	App	34h	O/M	M	
Read Event Message Buffer	18.8	App	35h	O	M	
Get BT Interface Capabilities	18.9	App	36h	O/M	O/M	Mandatory if BT implemented
Master Write-Read	18.10	App	52h	O/M	M	MasterWriteRead should disallow any writes to IPMB
Get System GUID	18.13	App	37h	O	O	
Get Channel Authentication Capabilities	18.12	App	38h	O	O	
Get Session Challenge	18.14	App	39h	O	O	

	IPMI Spec. section	NetFn	CMD	IPM Controller req.	Payload Interface Requirement	Additional Notes and Requirements
Activate Session	18.15	App	3Ah	O	O	
Set Session Privilege Level	18.16	App	3Bh	O	O	
Close Session	18.17	App	3Ch	O	O	
Get Session Info	18.18	App	3Dh	O	O	
Get AuthCode	18.19	App	3Fh	O	O	
Set Channel Access	18.20	App	40h	O	O	
Get Channel Access	18.21	App	41h	O	O	
Get Channel Info	18.22	App	42h	O	O	
Set User Access	18.23	App	43h	O	O	
Get User Access	18.24	App	44h	O	O	
Set User Name	18.25	App	45h	O	O	
Get User Name	18.26	App	46h	O	O	
Set User Password	18.27	App	47h	O	O	
Chassis Device Commands				**O**	**O**	All Chassis control commands when sent to IPMC on board are applicable to the board only.
Get Chassis Capabilities	22.1	Chassis	00h	O	O	
Get Chassis Status	22.2	Chassis	01h	O	O	
Chassis Control	22.3	Chassis	02h	O	O	
Chassis Reset	22.4	Chassis	03h	O	O	
Chassis Identify	22.5	Chassis	04h	O	O	
Set Chassis Capabilities	22.6	Chassis	05h	O	O	

	IPMI Spec. section	NetFn	CMD	IPM Controller req.	Payload Interface Requirement	Additional Notes and Requirements
Set Power Restore Policy	22.7	Chassis	06h	O	O	
Get System Restart Cause	22.9	Chassis	07h	O	O	
Set System Boot Options	22.10	Chassis	08h	O	O	
Get System Boot Options	22.11	Chassis	09h	O	O	
Get POH Counter	22.12	Chassis	0Fh	O	O	
Event Commands				**M**	M	
Set Event Receiver	23.1	S/E	00h	M	M	The Event Received shall always be IPMB address 20h and LUN 0
Get Event Receiver	23.2	S/E	01h	M	M	
Platform Event (a.k.a. “Event Message”)	23.3	S/E	02h	M	M	
PEF and Alerting Commands				**O**	**O**	
Get PEF Capabilities	24.1	S/E	10h	M	M	
Arm PEF Postpone Timer	24.2	S/E	11h	M	M	
Set PEF Configuration Parameters	24.3	S/E	12h	M	M	
Get PEF Configuration Parameters	24.4	S/E	13h	M	M	
Set Last Processed Event ID	24.5	S/E	14h	M	M	
Get Last Processed Event ID	24.6	S/E	15h	M	M	
Alert Immediate	24.7	S/E	16h	O	O	
PET Acknowledge	24.8	S/E	17h	O	O	
Sensor Device Commands				**M**	**M**	

	IPMI Spec. section	NetFn	CMD	IPM Controller req.	Payload Interface Requirement	Additional Notes and Requirements
Get Device SDR Info	29.2	S/E	20h	M	M	
Get Device SDR	29.3	S/E	21h	M	M	
Reserve Device SDR Repository	29.4	S/E	22h	M	M	
Get Sensor Reading Factors	29.5	S/E	23h	O	O	
Set Sensor Hysteresis	29.6	S/E	24h	O	O	
Get Sensor Hysteresis	29.7	S/E	25h	O	O	
Set Sensor Threshold	29.8	S/E	26h	O	O	
Get Sensor Threshold	29.9	S/E	27h	O	O	
Set Sensor Event Enable	29.10	S/E	28h	O	O	
Get Sensor Event Enable	29.11	S/E	29h	O	O	
Re-arm Sensor Events	29.12	S/E	2Ah	O	O	
Get Sensor Event Status	29.13	S/E	2Bh	O	O	
Get Sensor Reading	29.14	S/E	2Dh	M	M	
Set Sensor Type	29.15	S/E	2Eh	O	O	
Get Sensor Type	29.16	S/E	2Fh	O	O	
FRU Device Commands				**M**	**M**	
Get FRU Inventory Area Info	28.1	Storage	10h	M	M	
Read FRU Data	28.2	Storage	11h	M	M	
Write FRU Data	28.3	Storage	12h	M	M	
SDR Device Commands				**O**	**O**	

	IPMI Spec. section	NetFn	CMD	IPM Controller req.	Payload Interface Requirement	Additional Notes and Requirements
Get SDR Repository Info	27.9	Storage	20h	M	M	
Get SDR Repository Allocation Info	27.10	Storage	21h	O	O	
Reserve SDR Repository	27.11	Storage	22h	M	M	
Get SDR	27.12	Storage	23h	M	M	
Add SDR	27.13	Storage	24h	O/M	M	
Partial Add SDR	27.14	Storage	25h	O/M	M	
Delete SDR	27.15	Storage	26h	O/M	M	
Clear SDR Repository	27.16	Storage	27h	O/M	M	
Get SDR Repository Time	27.17	Storage	28h	O/M	M	
Set SDR Repository Time	27.18	Storage	29h	O/M	/M	
Enter SDR Repository Update Mode	27.19	Storage	2Ah	O	M	
Exit SDR Repository Update Mode	27.20	Storage	2Bh	M	M	
Run Initialization Agent	27.21	Storage	2Ch	O	O	
SEL Device Commands				**O**	**M**	
Get SEL Info	25.2	Storage	40h	M	M	
Get SEL Allocation Info	25.3	Storage	41h	O	M	
Reserve SEL	25.4	Storage	42h	O	M	
Get SEL Entry	25.5	Storage	43h	M	M	
Add SEL Entry	25.6	Storage	44h	M	M	
Partial Add SEL Entry	25.7	Storage	45h	M	M	

	IPMI Spec. section	NetFn	CMD	IPM Controller req.	Payload Interface Requirement	Additional Notes and Requirements
Delete SEL Entry	25.8	Storage	46h	O	O	
Clear SEL	25.9	Storage	47h	M	M	
Get SEL Time	25.10	Storage	48h	M	M	
Set SEL Time	25.11	Storage	49h	M	M	
Get Auxiliary Log Status	25.12	Storage	5Ah	O	O	
Set Auxiliary Log Status	25.13	Storage	5Bh	O	O	
LAN Device Commands				**O/M**	**O/M**	
Set LAN Configuration Parameters	19.1	Transport	01h	O/M	O	
Get LAN Configuration Parameters	19.2	Transport	02h	O/M	O	
Suspend BMC ARPs	19.3	Transport	03h	O/M	O	
Get IP/UDP/RMCP Statistics	19.4	Transport	04h	O	O	
Serial/Modem Device Commands				**O**	**O**	
Set Serial/Modem Configuration	20.1	Transport	10h	O/M	O	
Get Serial/Modem Configuration	20.2	Transport	11h	O/M	O	
Set Serial/Modem Mux	20.3	Transport	12h	O	O	
Get TAP Response Codes	20.4	Transport	13h	O	O	
Set PPP UDP Proxy Transmit Data	20.5	Transport	14h	O	O	
Get PPP UDP Proxy Transmit Data	20.6	Transport	15h	O	O	
Send PPP UDP Proxy Packet	20.7	Transport	16h	O	O	
Get PPP UDP Proxy Receive Data	20.8	Transport	17h	O	O	

	IPMI Spec. section	NetFn	CMD	IPM Controller req.	Payload Interface Requirement	Additional Notes and Requirements
Serial/Modem Connection Active	20.9	Transport	18h	O	O	
Callback	20.10	Transport	19h	O	O	
Set User Callback Options	20.11	Transport	1Ah	O	O	
Get User Callback Options	20.12	Transport	1Bh	O	O	
Bridge Management Commands (ICMB)				**O**	**O**	
Get Bridge State	[ICMB]	Bridge	00h	O	O	
Set Bridge State	[ICMB]	Bridge	01h	O	O	
Get ICMB Address	[ICMB]	Bridge	02h	O	O	
Set ICMB Address	[ICMB]	Bridge	03h	O	O	
Set Bridge Proxy Address	[ICMB]	Bridge	04h	O	O	
Get Bridge Statistics	[ICMB]	Bridge	05h	O	O	
Get ICMB Capabilities	[ICMB]	Bridge	06h	O	O	
Clear Bridge Statistics	[ICMB]	Bridge	08h	O	O	
Get Bridge Proxy Address	[ICMB]	Bridge	09h	O	O	
Get ICMB Connector Info	[ICMB]	Bridge	0Ah	O	O	
Get ICMB Connection ID	[ICMB]	Bridge	0Bh	O	O	
Send ICMB Connection ID	[ICMB]	Bridge	0Ch	O	O	
Discovery Commands (ICMB)				**O**	**O**	
Prepare For Discovery	[ICMB]	Bridge	10h	O	O	
Get Addresses	[ICMB]	Bridge	11h	O	O	

	IPMI Spec. section	NetFn	CMD	IPM Controller req.	Payload Interface Requirement	Additional Notes and Requirements
Set Discovered	[ICMB]	Bridge	12h	O	O	
Get Chassis Device ID	[ICMB]	Bridge	13h	O	O	
Set Chassis Device ID	[ICMB]	Bridge	14h	O	O	
Bridging Commands (ICMB)				**O**	**O**	
Bridge Request	[ICMB]	Bridge	20h	O	O	
Bridge Message	[ICMB]	Bridge	21h	O	O	
Event Commands (ICMB)				**O**	**O**	
Get Event Count	[ICMB]	Bridge	30h	O	O	
Set Event Destination	[ICMB]	Bridge	31h	O	O	
Set Event Reception State	[ICMB]	Bridge	32h	O	O	
Send ICMB Event Message	[ICMB]	Bridge	33h	O	O	
Get Event Destination	[ICMB]	Bridge	34h	O	O	
Get Event Reception State	[ICMB]	Bridge	35h	O	O	
OEM Commands for Bridge NetFn				**O**	**O**	
OEM Commands	[ICMB]	Bridge	C0h-FEh	O	O	
Other Bridge Commands				**O**	**O**	
Error Report	[ICMB]	Bridge	FFh	O	O	
AdvancedTCA®	PICMG® 3.0 Table			**M**	**M**	
Get PICMG Properties	3-9	PICMG	00h	M	M	
Get Address Info	3-8	PICMG	01h	M	M	

	IPMI Spec. section	NetFn	CMD	IPM Controller req.	Payload Interface Requirement	Additional Notes and Requirements
Get Shelf Address Info	3.13	PICMG	02h	O	No Req	
Set Shelf Address Info	3-14	PICMG	03h	O	No Req	
FRU Control	3-22	PICMG	04h	M	M	
Get FRU LED Properties	3-24	PICMG	05h	M	M	
Get LED Color Capabilities	3-25	PICMG	06h	M	M	
Set FRU LED State	3-26	PICMG	07h	M	M	
Get FRU LED State	3-27	PICMG	08h	M	M	
Set IPMB State	3-51	PICMG	09h	M	M	
Set FRU Activation Policy	3-17	PICMG	0Ah	M	M	
Get FRU Activation Policy	3-18	PICMG	0Bh	M	M	
Set FRU Activation	3-16	PICMG	0Ch	M	M	
Get Device Locator Record ID	3-29	PICMG	0Dh	M	M	
Set Port State	3-41	PICMG	0Eh	O/M	No Req	Should not be allowed through payload interface
Get Port State	3-42	PICMG	0Fh	O/M	M	
Compute Power Properties	3-60	PICMG	10h	M	M	
Set Power Level	3-62	PICMG	11h	M	No Req	Should not be allowed through payload interface
Get Power Level	3-61	PICMG	12h	M	M	
Renegotiate Power	3-66	PICMG	13h	O	O	
Get Fan Speed Properties	3-63	PICMG	14h	M	No Req	

	IPMI Spec. section	NetFn	CMD	IPM Controller req.	Payload Interface Requirement	Additional Notes and Requirements
Set Fan Level	3-65	PICMG	15h	O/M	No Req	
Get Fan Level	3-64	PICMG	16h	O/M	No Req	
Bused Resource	3-44	PICMG	17h	O/M	M	
Get IPMB Link Info	3-49	PICMG	18h	O/M	No Req	

ID	Feature Name	Requirements
BD-17	Front Board IPMC	If the board implements a Payload Interface, the board vendor ***may*** provide an Open IPMI driver for the board if LINUX* is supported. The latest version of Open IPMI source and documentation is at http://sourceforge.net/projects/openipmi/.
BD-18	Front Board IPMC	If the board implements a Payload Interface, the board vendor ***may*** provide an Open HPI framework for the board. The latest version of Open HPI source and documentation is at http://sourceforge.net/projects/openhpi/.
BD-19	Front Board IPMC	If a board supports LINUX* and implements a Payload Interface, the IPMC ***shall*** support over its Payload Interface all IPMI commands indicated as mandatory in the “Payload Interface Requirement” column of Table 6-2.
BD-20	IPMC Retries	IPMCs ***shall*** retry all IPMI commands 5 times under all error conditions, except if the IPMB bus has been determined to be hung. This requirement converts the timing specification C1 in Table 4-1 of the IPMI 1.5 specification from recommended to mandatory.
BD-21	IPMC Timeout	IPMCs ***shall*** provide retries based on a maximum timeout of 250 ms as defined by the IPMB specification or timeouts increasing linearly; starting at 60ms and increasing to 250 ms for the fifth retry.

All FRUs convey their IPMI-related events to the Shelf Manager over the IPMB bus, which has bandwidth of 100 Kbps. A single IPMB packet may take in excess of 2.5 milliseconds to transmit on the IPMB, not counting retries or processing time at either end.

ID	Feature Name	Requirements
BD-22	IPMB latencies	Applications or blades dependent on sub-second communication (i.e., 50ms failover timing) ***shall not*** depend on IPMI-based communication paths. Base or Fabric Interfaces ***should*** be used for these timing-critical functions.

6.1.5.3 IPMI Sensors

IPMI sensors can be used to monitor a variety of hardware and software capabilities in a modular system. These sensors can be discrete sensors (with specifically defined states) and threshold-based sensors

(typically derived from analog-to-digital [A2D] converters). Three of the most common threshold-based sensors are as follows:

- Voltage
- Temperature
- Fan Speed

Each of these threshold-based sensors needs the following alarm limits to be set:

- Upper and lower limits for critical alarms
- Upper and lower limits for non-critical alarms

The above alarms generated by boards and modules will be mapped into minor, major, and critical Telco alarms by the CMM.

The information below is to provide the designer with a base point to start from with regards to the margins and percentages to be used for setting the "upper critical", "upper non-critical", "lower critical" and "lower non-critical" voltage threshold settings. Note that after setting the thresholds it is imperative that they be "tuned" based on information and results gathered from physically monitoring each of the voltages with an oscilloscope specifically at the inputs of the voltage monitoring device.

6.1.5.3.1 Terminology

IPMI defines three levels of thresholds: non-critical, critical, and non-recoverable. The PICMG* 3.0 specification defines the three levels as minor, major, and critical. The word "critical" can indicate a second-level or third-level error, depending on which nomenclature is used.

6.1.5.3.2 False Errors vs. Missed Errors

Since the devices used for measuring the analog signals are subject to errors, a choice sometimes needs to be made between generating false error notifications vs. potentially missing valid errors. In such cases, if a product malfunction is likely if the error is missed, the decision should be made to ensure errors are reported, even if it means that false errors will be generated.

The above notwithstanding, all efforts should be made to eliminate the generation of false errors if at all possible. This may mean that non-critical thresholds will need to be eliminated to prevent the generation of false error messages.

ID	Feature Name	Requirements
BD-23	Front Board IPMC	Critical thresholds ***shall*** be defined for any sensor that indicates a condition where component failure or unreliable hardware operation is imminent. Documentation given to customers (i.e., Technical Product Specifications) ***shall*** indicate that critical thresholds are the thresholds that indicate potential failures and non-critical thresholds indicate conditions outside normal operating conditions but within the safe operating range of the affected components.
BD-24	Front Board IPMC	Non-critical thresholds ***may*** be defined for sensors where there is sufficient room between normal operating ranges and the critical thresholds; if there is not sufficient room, non-critical thresholds ***shall not*** be defined for those sensors. Non-critical thresholds ***may*** be used for predictive failure analysis, if they are defined.
BD-25	Front Board IPMC	Non-recoverable thresholds ***shall not*** be defined except where required by governing specifications (i.e., AdvancedTCA upper non-recoverable temperature limits).

6.1.5.3.3 Raw Values, Errors and Hysteresis

This includes the following:

- Raw values
- Errors
- Hysteresis
- Documentation Requirements

Raw Values

To understand how Sensor Data Record (SDR) thresholds are set, it is important to understand how they are used. Most threshold based sensors are interpreted as follows:

Real value = Raw Value x Multiplier + Offset

The Raw Value is always a single-byte value. This means that sensors do not have infinite precision, they're limited to 256 values.

The IPMI controllers (IPMCs) operate off the raw values. They never do the conversion to the real values as indicated in the formula above. Thresholds are set by raw values, and an IPMI controller compares the raw values read to the raw threshold values to determine if an event should be generated. The conversion to real values only occurs when a sensor or event value needs to be presented to a higher-level interface (but this is typically done by the Shelf Manager or higher-level software, not the IPMI controller).

Errors

Each sensor device is subject to some sort of error, and the technical specifications for that component should detail the exact error budget for that component. Some of the more common A2D converters with I^2C interfaces have a total unadjusted error of 2.5%.

A common misconception is that a variance in the input (reference) voltage to the component is an additive element to the overall error budget. Most components will account for any variance in their input voltage in the estimate for their measurement error.

Hysteresis

One potential with a threshold-based sensor is that the analog input for that sensor could hover near the threshold value. If a sensor's upper critical threshold had a raw value of 87 and the actual value fluctuated between 86 and 87 (right at the line between them), the sensor could generate multiple events from what is essentially a steady-state condition. Every time the sensor read 87, the IPMC would generate an *assertion* event to indicate the error occurred, and every time the sensor read 86, the IPMC would generate a *de-assertion* event to indicate that the sensor was below the threshold. This could easily generate hundreds or thousands of events when there was little to no change in the system. This could fill up the System Event Log (SEL) and prevent other events from being recorded.

To eliminate this issue, each SDR defines hysteresis values for threshold-based sensors. The purpose of the hysteresis value is to "debounce" the sensor reading and only generate new events when sensor values change by a significant enough amount.

In the example above, a hysteresis value of 0 (no hysteresis) was essentially used. If a hysteresis value of 1 was used, there would be a raw value buffer of 1 between the assertion and de-assertion values. Thus, an event would still be generated when the sensor went to 87, but a de-assertion event would not be generated until the sensor went to 85. A hysteresis value of 2 would prevent de-assertion until the sensor went down to 84, etc.

In the example above, if the hysteresis value is set to 1, the sensor would always be de-asserted at 85 and always asserted at 87. If the raw value of the sensor is 86, it may be either asserted or de-asserted, depending on whether the assertion threshold or hysteresis boundary was most recently crossed.

Technically, there are two hysteresis values for each sensor; one applies to upper thresholds and one to lower thresholds. However, in practice, the upper and lower hysteresis values for a given sensor are typically the same value.

In some cases, a larger hysteresis value may be used, but larger hysteresis values run the risk of never de-asserting a momentary crossing of a threshold. For example, if the nominal raw value of the sensor in the example above is 84, the event may never be de-asserted with a hysteresis value of 2 or more, even though a raw value of 85 may be well within the normal operating range of the affected components. In general, the hysteresis value for threshold-based SDRs should be 1.

ID	Feature Name	Requirements
BD-26	Front Board IPMC	All threshold-based SDRs ***shall*** define a hysteresis value of at least 1.

Documentation Requirements

ID	Feature Name	Requirements
BD-27	Front Board IPMC	Customer documentation for all threshold-based sensors ***shall*** include the following thresholds: • Lower Non-Recoverable (LNR) • Lower Critical (LC) • Lower Non-Critical (LNC) • Nominal value (not technically a threshold, but included here for completeness) • Upper Non-Critical (UNC) • Upper Critical (UC) • Upper Non-Recoverable For each threshold defined above, the value ***shall*** be defined in real-world limits (i.e., 1.223V). Both the assertion and de-assertion values ***shall*** be provided for each threshold. Additionally, the raw sensor threshold values from the monitoring device (i.e., 87) ***may*** also be provided.

6.1.5.3.4 Voltage Limits

Voltage limits within a shelf could be divided into two groups:

1. Limits specified by the standards.
   Example: Backplane voltage and power supply limits specified by PICMG* 2.0 and PICMG2.11 respectively. Also applies to voltages such as DC input voltages to chassis.
2. Board/module voltage limits. Specified by the board vendor based on operating voltage and absolute voltage limits of parts on a board/module as well as onboard supply tolerances.

Backplane limits are typically specified in the relevant standard and are applicable at slot inputs. Voltages could vary from slot to slot due to IR drop and voltage sensor location. For this discussion, it is assumed that the power plane designs are good and variation from slot to slot is negligible or within the limits of the applicable specifications. (For example, PICMG 3.0 allows up to a 1V drop from the shelf's DC input to the blade's power connector.)

Setting Limits for Internal Voltage Rails

The measurement of voltages is typically not precise down to the last milliVolt. By the time the error budgets and hysteresis values are taken into account, it is not always possible to set both non-critical and critical thresholds without the risk of generating substantial quantities of false alarms. To prevent this situation, most non-critical voltage thresholds will be eliminated.

Determining the Critical Working Voltage Range

Each voltage rail will have one or more components using that rail. For each component on that rail, determine the minimum and maximum operating voltages for each component. Then select the *minimum maximum* and the *maximum minimum* voltage for each rail. For example, consider the following components:

Component	Minimum Voltage	Maximum Voltage
A	**1.40**	1.65
B	1.33	1.78
C	1.21	**1.61**
Critical Working Voltage	1.40	1.61

The Critical Working Voltage (CWV) is 1.40 to 1.61 Volts, based on the most restrictive requirements of the components on that voltage rail. Note that the minimum and maximum voltages may be from different components or the same component.

The CWV provides a starting point for setting the lower and upper critical thresholds for the selected rail.

Note that there may be a voltage drop between the voltage supply and the component using it. Similarly, there may be a voltage drop between the supply and the device measuring the voltage. Such offsets need be taken into account (determined through measurement with scopes or logic analyzers) when determining the CWV. This includes the following:

- Setting Thresholds
- Validating the Supply Voltage
- Negative Voltages

Setting Thresholds

If the measurement device has a measurement error of 2.5%, whatever threshold is set will have about a 5% (±2.5%) variance centered around the selected threshold value. The threshold is set to match the CWV. This is a straightforward and easy to explain method. Under most circumstances, the voltage rail should be more than 2.5% tighter than the CWV, so the risk of false errors is mitigated. Voltage margining of ±5% is typically done on the voltage rails for design validation; this is 2x the error budget, so the chance of missing critical errors is substantially reduced.

ID	Feature Name	Requirements
BD-28	Front Board IPMC	The measurement device ***should*** be precise enough so that its error budget is less than half the margined voltage level for each voltage. This can be adjusted by either selecting a more precise A2D converter or increasing the voltage margining for each voltage in the system.

Validating the Supply Voltage

In an ideal world, the voltage supply to each rail would be constant and would not vary. However, there must be some allowance for some variation in the supply voltage. In the following

requirements, *N%* refers to the error budget of the voltage monitoring device (2.5% in the examples above).

ID	Feature Name	Requirements
BD-29	Front Board IPMC	The design engineer ***shall*** define the expected voltage range of each internally generated voltage rail. The design engineer ***shall*** ensure that these limits are tested and met during the design validation stage.
BD-30	Front Board IPMC	The power rail ***should*** be designed such that the voltage rail is at least N% tighter than the CWV. IF the voltage rail is within N% (total measurement error) of the CWV, the LC and/or UC thresholds ***shall*** be adjusted outwards accordingly to maintain at least an N% margin between the operating range of the voltage rail and the thresholds in order to avoid nuisance events. In such cases, the voltage margining ***shall*** be adjusted accordingly to keep at least 2*N% (2x of the measurement error) buffer between the threshold and margin limit.

Negative Voltages

Negative Voltages are to be treated as absolute value disregarding the negative sign. In this case increased value of the A2D raw value represents the increased absolute value of the voltages. For example, a voltage level of -72V should be considered "higher" than -48V. Upper Critical Limits for negative voltages will have a higher absolute voltage limit than Lower Critical Limits for that same sensor.

6.1.5.3.5 Temperature Limits

Under normal operation, the ambient air to a shelf will be in the range of 5°C to 40°C. In the case of an HVAC (Heating, Ventilation, and Air Conditioning) plant failure, the range of temperatures can be -5°C to 55°C. (Refer to NEBS GR-63, R4-7—note the footnote in the table). Depending on the placement of the temperature sensor relative to the ambient air coming into the shelf, the temperature could rise between 0 and 15°C from the ambient temperature.

ID	Feature Name	Requirements
BD-31	Front Board IPMC	Sensors not receiving pre-heated air ***shall*** assume an ambient temperature rise (T_r) of 0°C. Temperature sensors receiving pre-heated air ***shall*** assume a T_r of 15°C, unless a more precise limit for that sensor's T_r can be established.
BD-32	Front Board IPMC	Some equipment types (such as CMMs and fan control boards) may be designed to operate in areas of pre-heated air. For these components, the temperature sensor ***shall*** assume a T_r of 15°C. Using the T_r value as described above, the temperature thresholds ***shall*** be set according to Table 6-3**.**

Table 6-3 Temperature Threshold Settings

Limit Type	Description	Comment
Upper Non-Recoverable	Upper Critical + 10°C	With the assumption that the component still can operate at least 10°C above its published limit
Upper Critical Limit	55°C + Tr.	Hard drive and battery may or may not meet this limit.

Limit Type	Description	Comment
Upper non Critical Limit	40°C + Tr.	
Nominal	25°C	Not truly a threshold, but defined for documentation purposes as a frame of reference.
Lower non Critical Limit	5°C	
Lower Critical Limit	-5°C	Does not have to start operating at -5°C, but must be able to allow ambient to drop this low while in operation.
Lower Non-Recoverable	Undefined	

6.1.5.3.6 Fan Speed

The primary purpose for setting fan speed thresholds is to alert the administrator when a fan is failing. Since most fan failures occur by a gradual slowing of one or more fans, a properly set threshold can alert the administrator *before* the fan actually fails. This allows a replacement part to be ordered and the fan to be fixed during a normal service call rather than forcing a truck roll for an emergency fan replacement.

Properly set fan thresholds can save network operators hard dollars in terms of operational expenses, and this can be a crucial factor in making product decisions.

This includes:

- Fan Thresholds

  A fan speed variation of as little as 10% needs to be detected. However, administrators are allowed to manually change the fan speed by 20% or more. Any static fan threshold will be virtually useless if it is set outside this range.

- Detecting Fan Speed Differences

  Controllers for fans can set the minimum and maximum fan speeds expected for those fans. The addition of a differential fan speed sensor will also provide dynamic fan speed detection to solve the problems above.

ID	Feature Name	Requirements
BD-33	Front Board IPMC	The IPMC monitoring the fans ***should*** define a differential fan speed sensor with LNC and UNC thresholds set at an RPM reading just beyond the expected variation between fans in the system. The LC and UC thresholds ***should*** be twice the RPM value of the LNC thresholds. Each time the differential fan speed sensor is checked, it ***should*** do the following: Select the next fan in the series (do a circular queue of all the fans). Get the speed of that fan. Call this the *fan0* speed. Go through fans 1-n. Compare the fanx speed to fan0 speed. If the result is less than x RPM or more than y RPM, generate an LC, LNC, UNC, or UC event as appropriate. As this cycles through each of the fans as a reference fan, any discrepancy between the fans becomes apparent if the difference is large enough. Note that the offset, multiplier, etc. for the fan0 sensor can be the same as the ones for the real fans. This will allow the difference to be reported to the user in RPM as easily as an individual fan's speed.

6.1.6 Boot Firmware

There are many boot firmware implementations available with varying degrees of functionality. At this point, this guide does not recommend any one implementation, but the goal of this section is to specify a set of functional requirements that should be provided by the boot firmware on all MCP-compliant boards. A common set of functions for the boot firmware will simplify the system integrators' task of upgrading and deploying software on systems built with boards from many vendors.

The main task of the boot firmware is to prepare the board to boot operating systems or embedded applications, provide services to the OS (particularly early in the boot process), and provide manageability data to the system. The boot firmware is normally stored in Flash on the board and the OS or embedded applications are stored on devices that are either accessible locally or remotely by the board. The boot firmware must provide the capability of supporting the different set of devices that can be used for storing the OS.

The boot firmware typically consists of two main parts:

- Boot Initialization
- Boot Monitor

6.1.6.1 Boot Initialization

This is the first piece of code that the processor executes from Flash after it is released from the reset state. This code performs the following steps:

- Minimum initialization of CPU core and memory controllers so that the boot monitor can be loaded into RAM
- Read a jumper or switch on the board to determine which Boot Monitor to load into RAM
- RAM test – limited fragment of RAM, used by Boot Monitor, is tested with pattern tests of each memory location within tested area
- Validate the boot monitor binary image
- The Boot Monitor code is copied from flash into RAM and then executed

The boot initialization code occupies the first sector (or sectors) of flash memory associated with a particular processor – the boot sector. The boot sector must be locked (write-protected) to avoid the possibility that a user accidentally overwrites the boot sector with an incorrect image, which would render the board unusable without first repairing it with hardware tools such as a JTAG flash programmer.

6.1.6.2 Boot Monitor

The Boot Monitor should provide a wide set of tools for downloading and executing programs on the board, as well as tools for manipulating the board's environment. It should be used for both product development (debug support) and for end product deployment (flash and network booting).

The Boot Monitor should provide the following capabilities:

- Power On Self Tests (POST)
- Boot Manager
- Console for configuration and management
- Local and network boot support
- Communication to the IPMC
- Error logging
- Security

There should be provision for a backup copy of the Boot Monitor which should be invoked by the Boot Initialization code if the active copy of the Boot Monitor is corrupted.

ID	Feature Name	Requirements
BD-34	Front Board Boot Firmware	Boards ***shall*** have a backup copy of the boot monitor which shall be invoked if the Boot Initialization code detects that the active copy of the boot monitor is corrupt.
BD-35	Front Board Boot Firmware	Boards ***shall*** support the remote update of all Firmware on the board and the restoration of configuration parameters. This applies to all IPMI firmware and BIOS.

6.1.6.2.1 Power-On Self Test (POST)

The Power-On Self Test (POST) performs initialization and basic tests of the CPU core and main baseboard components on which the CPU depends. When POST is completed, results are stored in Boot Status Record. The components tested by POST are board dependent but it is expected that all peripherals on the board will be tested. All POST errors should be logged in the IPMC SEL.

ID	Feature Name	Requirements
BD-36	Front Board Boot Firmware	Boards ***should*** log all POST errors in the IPMC SEL.

6.1.6.2.2 Boot Manager

The Boot Manager is the main part of Boot Monitor and controls the whole booting process. First it runs POST and then reads the configuration stored in flash to determine next steps.

Generally, Boot Manager enters the console prompt and waits for user input. At this moment, the user can run utilities built in the Boot Monitor or execute operating software (diagnostics, or operating system or OS-loader). The Boot Manager is responsible for running proper utilities once the console command is recognized.

Optionally, the user can define a startup script that is stored in flash together with other configuration parameters. A startup script is a set of console commands to be executed on startup. When a startup script includes a command executing operating software, then operating software will be started automatically to take over control of the processor (it does not return to Boot Monitor in this case). Otherwise, when all commands from the startup script are completed, Boot Monitor displays a console prompt.

6.1.6.2.3 Boot Monitor Console

The Boot Monitor uses the console as an output port from the early phases of POST. This is normally a front panel serial port. It is used to report the status (progress and errors) of the POST phases that follow.

The console can also be used as an interactive user interface (Command Line Interface) that allows the user to manage the boot process or run utilities built in to Boot Monitor.

The Command Line Interface is activated (Boot Monitor breaks normal execution and enters console prompt) in the following cases:

- After completion of POST when the startup script is disabled (default case).
- When the user presses Ctrl-C on the console during POST execution (after UART ports have been initiated and tested). In this case, the current phase of POST is completed first and then Boot Monitor enters the console prompt, waiting for user input.
- When startup script is completed and there was no command executing operating software.

- When Ctrl-C was pressed during the execution of the startup script (currently executed command is finished first).

The examples of features provided by Boot Monitor console include:

- **Execution of download utilities:** The user can download a file into a selected memory area using X/Y modem (via serial port) or TFTP (via Ethernet), and can manually define IP addresses necessary for TFTP transfer or use BOOTP client to get them from the host.
- **Execution of flash tools:** The user can write a selected area of memory into flash. Usually it will be a file previously downloaded into RAM. The image can also be loaded from flash to RAM or can be deleted from flash.
- **Definition of configuration:** The user can define configuration parameters which include local IP addresses, default server's IP address or content of startup script. The configuration is stored in a dedicated sector of flash.
- **Code execution:** The user can execute the code located in selected address of RAM or flash.
- **Diagnostics:** It should be possible to run single tests or complete diagnostics of the entire board. Some tests can be executed with parameters that allow the user to modify test execution. Some tests can be repeated a specified number of times or run until Ctrl-C is pressed.

6.1.6.2.4 Security

The Boot Monitor should provide two levels of security: user and supervisor.

6.1.7 Debug Interfaces

All front boards should provide a serial interface to each CPU that can be used to output debugging information. There should also be a debugging interface such as JTAG or ITP present for all CPUs on the board. The debugging interface is connected to a header for use during manufacturing and board level testing.

6.2 Common Front Board Types

The high level architecture of the blades used to build MCP Systems is described in Chapter 2. This section describes the detailed requirements of each of the blade types: Control, Protocol Processing, Storage, Hub, Service, and Line cards. The functional requirements for these common board types are described in more detail in the Network Element profiles.

7. Shelf Management

Shelf Management refers to the Shelf Manager (ShMgr) acting as the single point of management for all Managed Field Replaceable Units (FRUs) in the shelf for out of band (OOB) management. Out of Band (OOB) management in this section refers to the actions performed by communication between the Shelf Manager and all managed FRUs through IPMB-0 (Out Of Band Management Bus). The Shelf Manager interacts with Managed FRUs in the chassis to perform Shelf Management using an Intelligent Platform Management Interface (IPMI). Refer to Section 3.1.1 and Section 3.3 of the AdvancedTCA® specification for details.

In a hierarchical model, a System Management entity is a level above the Shelf Manager. It communicates with shelf managers on one or more shelves to get/set information. It can also directly communicate with boards with managed FRUs within the shelves to get/set information that may not be available from the Shelf Manager. The scope of this chapter is limited to Shelf Manager-related functionality.

Section 3 of the AdvancedTCA specification covers the architecture and detailed requirements for shelf management and its relationship and dependencies with IPMI. The Design Guide narrows down options available within the specification by documenting recommendations or requirements for shelf management solutions. This chapter also documents Shelf Manager requirements for communication with system management software.

Section 7.1 summarizes the capabilities of shelf management as detailed in the AdvancedTCA specification. Further sub-sections list additional requirements for functionalities that ensure higher levels of reliability and availability, such as redundancy, failover, and recovery.

Section 7.11 delves into the external interfaces required by the Shelf Management solution to interface with system management software or operators.

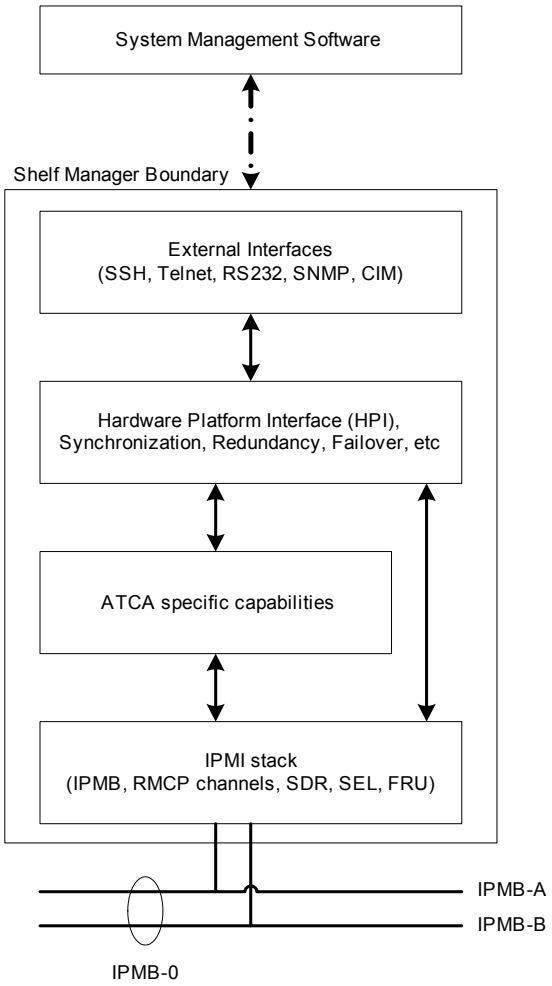
7.1 Overview

The Shelf Manager is a single point of out of band (OOB) management for the AdvancedTCA shelf by managing the shelf and all managed resources in the shelf. While the list below is not exhaustive, it illustrates key functionalities and provides guidance for additional capabilities required on a Shelf Manager implementation.

- Hot swap management (Section 3.2.4 of AdvancedTCA specification)
- Power Management (Section 3.9 of AdvancedTCA specification)
- Electronic Keying Management (Section 3.7 of AdvancedTCA specification)
- Cooling Management (Section 3.9.3.3 of AdvancedTCA specification)
- Out of Band access for control of managed devices (power on, reset, power cycle, power off)
- Status and health monitoring of all managed devices
- Fault management and recovery of back plane IPMB-0 errors that require action when potential or actual fault conditions are detected. (Section 3.8.7 of AdvancedTCA specification)
- Notification of states, status, and faults to management sub systems outside of the Shelf Management (System Management)

The AdvancedTCA shelf management architecture utilizes significant aspects of the Intelligent Platform Management Interface (IPMI) infrastructure. The Shelf Manager consists of the following:

1. **Shelf Management Controller (ShMC).** This is the communications and control hardware used by the Shelf Manager software/ firmware to communicate with and control the other elements in the shelf management system. ShMC provides the Shelf Manager with the interface to the back plane IPMB-0 for sending and receiving IPMI messages between managed FRUs.
2. The **logical entity** responsible for managing power, cooling and interconnects (with Electronic Keying) in AdvancedTCA shelf. The shelf manager also routes messages from the system manager to IPMB-0, provides interface to system repositories, and responds to event messages. It provides a single point of management for the shelf.



** - Shelf Management may also include other interfaces like FTP, SFTP, SLP, Telnet, SSH as required*

Figure 7-1 Shelf Manager Architecture

Figure 7-1 illustrates a generic block diagram of a shelf manager that includes the IPMI stack that provides the capabilities defined by the IPMI v1.5 specification. AdvancedTCA specific capabilities which

provide power management, hot swap management, E-Keying and cooling management capabilities can be integrated with the IPMI stack. Service Availability Forum (SAF) has defined a Hardware Platform Interface (HPI). This interface provides the application programming interface (API) for management applications to interact with underlying hardware subsystem. HPI implementations collect incoming event information, identify error categories, and provide collective health information. The current version (B.01.01) is recommended for adoption. Lastly, the external interface layer provides programmatic and human interface into the shelf manager.

7.2 Management Bus (IPMB)

Regardless of the bus topology (radial/bus), the following timeouts and retry mechanisms are required to be implemented.

ID	Feature Name	Requirements
SM-1	Management Bus	Shelf managers ***shall*** retry all IPMI commands 5 times under all error conditions, except if the IPMB bus has been determined to be hung. This requirement converts the timing specification C1 in Table 4-1 of the IPMB specification from recommended to mandatory.
SM-2	Management Bus	Shelf managers ***shall*** provide retries based on a maximum timeout of 250 ms as defined by the IPMB specification or timeouts increasing linearly; starting at 60ms and increasing to 250ms by the last retry.

All FRUs convey their events to the Shelf Manager over the IPMB bus, which has maximum bandwidth of 100kbps, as defined in PICMG 3.0 specification. Given the nature of this bus, and that the event information further may need to propagate over 10/100 Ethernet interface from shelf manager to system manager, the following requirement allows management applications to take potential actions on events with minimum latencies.

ID	Feature Name	Requirements
SM-3	IPMB latencies	It is recommended that applications/blades that need 50 ms failover timings do not rely on event information from shelf managers to make decisions for failover conditions. Blades ***may*** have local hardware management agents that communicate with IPMI controllers on the blade, which read in events and make failover decisions.

7.3 Hot Swap State transition handling

The PICMG 3.0 specification defines the state machine for managing hot swap transitions of all managed FRUs in the shelf. This section provides additional requirement to manage FRUs more efficiently such that the presence of a FRU in the shelf can be determined to the fullest extent possible.

ID	Feature Name	Requirements
SM-4	Hot Swap	Shelf managers ***should*** continue to ping the Managed FRU in the M1 state (when coming into the M1 state from either the M2 or M6 state) until it stops responding. At that time, it would be safe to conclude that the FRU has entered the M0 state and log a M1 → M0 transition event on behalf of the FRU.

7.4 Self Test Support

The IPMI stack on the shelf manager can support the ability to report out self test errors through IPMI commands and eventually reflect it through all of its supported interfaces like SNMP, CIM, and CLI. This capability enables root causing basic errors on shelf manager, due to which intended service may not be available.

ID	Feature Name	Requirements
SM-5	Self Test	Shelf managers ***should*** support reporting out self test errors *(refer to Get Self Test Results command)* as defined in IPMI v1.5 specification and reflect the same information through interfaces used for communication with system management.

7.5 Sensor Support

The shelf manager shall support sensor types defined by the IPMI v1.5 specification. If the shelf manager implementation supports IPMI v2.0, additional sensor types defined should be supported and integrated into the overall health of managed entities, while ensuring backward compatibility with the IPMI v1.5 specification.

IPMI specification defines a broad range of sensor types that can be used to map sensors; sometimes these predefined sensors may not be adequate, requiring definition of OEM specific sensors. Events are generated to the Shelf Manager for OEM sensors as well; hence, it is required that shelf management designers implement a framework such that integrating OEM sensors into the overall health status information is possible. When OEM sensors are implemented on the Shelf Manager, those OEM sensors should also be integrated into the overall health status as described in Section 7.7.

ID	Feature Name	Requirements
SM-6	Sensor Support	Shelf manager implementations ***shall*** provide detailed documentation on any OEM sensors defined on the shelf manager, to enable ease of integration with system management.

Shelf Managers provide Ethernet connectivity for communication with external system management entities and human interfaces. Since this connectivity is important for overall shelf manager availability, the following sensor support is required by each Shelf Manager for every Ethernet link to which it is connected.

ID	Feature Name	Requirements
SM-7	Sensor Support	Shelf manager implementations ***shall*** monitor Ethernet connectivity and log the loss of connectivity into the system event log (SEL). Sensor type 27h as defined by the IPMI specification shall be used for this purpose.

7.6 System Event Log (SEL)

In the AdvancedTCA environment, platform event messages notify the shelf manager of all events, including error conditions, state changes, and hot swap state changes on the managed FRUs. Since storage space on shelf managers is expected to be finite, it is possible that the system event log (SEL) fills up and cannot accept new entries. Under these conditions, new entries will not be written to the SEL (as per IPMI specification) and this could affect the overall operation of the shelf. The following requirements allow shelf managers to process events under conditions when the SEL is full or gets corrupted.

ID	Feature Name	Requirements
SM-8	System Event Log	Shelf manager implementations ***shall*** send the event logging disabled (sensor type 10h) event through either the SNMP trap or PET, which allows proactive notification of SEL full indications to system management entities. IPMI v1.5 specification defines the sensor type 10h for Event Logging Disabled indications (section 36.2).
SM-9	System Event Log	Shelf manager implementations ***shall*** provide at least 10 Kbytes of persistent storage space for the IPMI System Event Log (SEL).
SM-10	System Event Log	Shelf managers ***shall*** process all events when the SEL is full. Though events will not be logged into the SEL, the Shelf Manager shall process events and take necessary actions (e.g. SNMP traps, PET, cooling management).
SM-11	System Event Log	Shelf managers ***shall*** respond with “SEL erase in progress” response code (81h) to Platform Event Message requests when the SEL is being erased.

7.7 Hardware Platform Interface (HPI)

HPI is the programming interface for interface providers and management applications to get access to underlying hardware management subsystem. HPI provides consistent interface to interact with all resources in the shelf, including blades, shelf FRU, Fan Trays and shelf management entities.

ID	Feature Name	Requirements
SM-12	HPI Support	Shelf manager implementations ***should*** provide Hardware Platform Interface (HPI) implementation (SAI-HPI-B.01.01 or later) that SNMP agents, CIM providers, and other interfaces can leverage off.
SM-13	AdvancedTCA to HPI mapping	Shelf manager implementations ***should*** comply with the AdvancedTCA to HPI mapping specification being developed by the Service Availability Forum. Since this is a work in progress, readers of this document are requested to track the progress of this effort. Eventually, the mapping effort will be reviewed by PICMG for adoption into its specifications.

7.8 Health Status Reporting

Since the Shelf Manager is the single point of OOB management of the shelf, the shelf manager shall track all prevailing OOB error conditions in the shelf and provide collective information on the health of the managed entities at a managed FRU and shelf level.

This mechanism allows operators and system management entities to query and determine the health of individual AdvancedTCA managed FRUs without having to read system event logs, query current readings and then determine the state of each sensor and collective health. This functionality is considered the responsibility of the shelf manager.

The Shelf Manager provides the collective health status of managed FRUs based on the sensors available from managed FRUs through OOB access (IPMI).

ID	Feature Name	Requirements
SM-14	Health Status Reporting	The shelf manager ***shall*** monitor the health of itself, redundant shelf managers, and all managed FRUs.
SM-15	Health Status Reporting	The shelf manager ***shall*** be able to report known health status by providing a severity level (minor, major or critical) and the list of associated

ID	Feature Name	Requirements
		events that contribute to each severity. This information shall be reported through all interfaces (e.g. SNMP, CLI, CIM) except RMCP.
SM-16	Health Status Reporting	Shelf managers ***shall*** be able to report health status on a per managed FRU level. This information shall be reported through all interfaces (e.g. SNMP, CLI, CIM) except RMCP.
SM-17	Health Status Reporting	Shelf managers ***shall*** be able to report collective health status of the entire shelf. This information shall be reported through all interfaces (e.g. SNMP, CLI, CIM) except RMCP.

7.9 Redundancy

Since MCP platforms are designed for high levels of reliability and availability, redundant shelf managers must be present in the shelf. Redundant Shelf Managers function in active-standby configuration, such that the active shelf manager communicates with managed FRUs and performs shelf management. The active and standby shelf managers typically communicate with each other to synchronize information and keep track of the operational state of the redundant shelf managers. A failover from active to standby occurs in case of the active shelf manager is unable to manage the shelf due to failure or deterioration of the health of the active shelf manager.

ID	Feature Name	Requirements
SM-18	Redundancy	Redundancy ***shall*** be implemented in accordance with requirements SH-67 and SH-68 Redundant shelf managers ***shall*** function in an active-standby configuration.
SM-19	Redundancy	Shelf managers ***shall*** be able to failover to the redundant shelf manager when the active is either unable to maintain its level of functionality or so instructed through programmatic or human action.
SM-20	Redundancy	Shelf managers **shall** synchronize critical data between the active and standby shelf managers. Examples of critical data include SDR repository, health information, SEL, configuration information and any other specific data required for the reliable operation of the shelf and shelf manager functionality.
SM-21	Redundancy	A Shelf manager ***shall*** be able to report redundancy status information and integrate this information into the overall health information of the shelf manager.

7.10 Failsafe Upgrades

Shelf managers are intended to be field upgradeable, hence shelf management solutions need to provide reliable and fail safe mechanisms for software updates/upgrades.

Normal updates/upgrades are the cases where shelf management images/binaries are changed under normal operating conditions, without errors. Recovery mode refers to cases where either the corruption of code or an earlier update/upgrade failure results in reduced functionality, which may be limited to accepting updates only (TFTP).

When shelf manager designers are considering options to provide a redundant boot loader for recovery mode purposes, the following should be considered.

1. Ability of the boot loader to accept updates over the network, using either FTP or TFTP.

2. If the shelf manager software resides on flash and the hardware has more than one flash part, both copies of the boot loader shall reside on different flash parts, such that failure of one flash part does not cause failure of the entire board.
3. On initialization of the board, the boot loader shall check the integrity of the redundant copy of the boot loader, perform self tests on critical hardware components like the NIC, any other PCI devices, memory and log errors, which should then be reflected into the IPMI based system event log (SEL).

ID	Feature Name	Requirements
SM-22	Failsafe Upgrades	Shelf manager software/firmware ***shall*** be able to be updated over the network for normal upgrades/updates and recovery from upgrade/update failures. Network connectivity for recovery-based upgrades ***may*** be limited to a single subnet (e.g. TFTP).
SM-23	Failsafe Upgrades	Shelf manager software/firmware ***shall*** be able to be reliably upgraded in the field by providing redundant copies of boot loader software, which can be used to update FW/SW using either FTP or TFTP to recover from upgrade failures. It ***shall*** be possible to configure one copy of boot loader software to be write protected at all times.
SM-24	Failsafe Upgrades	Shelf manager software/firmware ***shall*** be able to re-enumerate the operational state of entire shelf if it loses track of the current state of FRUs in the shelf. For example, A Shelf with 2 shelf managers is running with boards powered on. Both shelf managers are then physically extracted from the shelf and reinserted. The shelf manager ***should*** be able to communicate with all boards and re-sync operational states of the boards.

7.11 External Interfaces

MCP shelf managers need to provide additional interfaces for communication outside of the shelf. The network protocols listed need to be supported on shelf managers for secure communication and shelf discovery by system management entities.

Service Location Protocol (SLP) is an IETF standards protocol that provides a framework to allow networking applications to discover the existence, location, and configuration of services.

ID	Feature Name	Requirements
SM-25	External Interfaces	MCP shelf managers ***shall*** provide at least 1 IEEE 802.3 Ethernet interface. This interface shall be implemented to the base Ethernet interface.
SM-26	External Interfaces	An additional Ethernet interface on the front panel ***may*** be implemented.
SM-27	External Interfaces	Shelf managers ***should*** support Secure Shell (SSH), Telnet, Secure FTP (SFTP), and File Transfer Protocol (FTP).
SM-28	External Interfaces	Shelf managers ***shall*** provide the capability to disable Telnet and FTP for security reasons.
SM-29	External Interfaces	Shelf managers ***should*** provide Service Location Protocol (SLP) for automated discovery of shelves in an operational environment. SLP v2 should be implemented as per RFC2608.
SM-30	External Interfaces	If SLP is supported, the Shelf Manager ***shall*** provide a SLP Service Agent (SA) for network discovery.
SM-31	External	If SLP is supported, the SLP discovery string ***shall*** be customizable by

ID	Feature Name	Requirements
	Interfaces	users.
SM-32	External Interfaces	If SLP is supported, the SLP discovery string ***shall*** be stored in the SLP configuration file.
SM-33	External Interfaces	If SLP is supported, the SLP service agent on the shelf manager ***shall*** respond to multicast and unicast service requests.
SM-34	External Interfaces	If SLP is supported, the SLP service agent ***shall*** attempt to register itself with directory agents at initialization only.

The following sections list additional external interfaces and their requirements for shelf manager implementations.

7.11.1 Front Panel Serial Port (RS232)

The RS232 port is used for physical access into the shelf manager for initial configuration and deployment. This interface is used on the shelf manager for initial configuration of the shelf parameters, when network connectivity may not be available.

ID	Feature Name	Requirements
SM-35	Serial Port	Shelf managers ***should*** provide an RS232 serial port.
SM-36	Serial Port	When shelf managers support the serial port, the connector layout ***shall*** be as shown in Section 12.3.

7.11.2 System Management Interface

The shelf manager shall support the following interfaces for integration with system management.

ID	Feature Name	Requirements
SM-37	System Interface	Shelf managers ***may*** support command line interface.
SM-38	System Interface	If shelf managers support SNMP, a standardized SNMP MIB ***should*** be supported as discussed in the next section.
SM-39	System Interface	If shelf managers support the CIM interface, a standardized CIM schema and SOAP as wire protocol ***shall*** be supported. Standardization of the CIM schema for AdvancedTCA is in progress.
SM-40	System Interface	Shelf managers ***should*** support either SNMP or CIM based interface.
SM-41	System Interface	It is suggested that system management applications do not use raw RMCP (IPMI over LAN) to integrate with system management. RMCP ***may*** be used for debugging, IPMI based FW updates to IPMI controllers on IPMB-0, and similar purposes.
SM-42	System Interface	Shelf Management solutions may already provide proprietary interfaces (e.g.: RPC, CLI, SNMP), as standardized interfaces for communication between shelf and system management are not available. The Design Guide provides pointers for standardization efforts of shelf to system management interfaces in standards bodies (PICMG, DMTF & SAF), which need to be supported on all shelf management solutions. Hence, it

ID	Feature Name	Requirements
		is required that all shelf management solutions ***should*** provide a transition path to customers to migrate from proprietary interface to standardized interfaces as described in this document.

7.11.2.1 Command Line Interface (CLI)

The shelf manager should support command line interfaces into the shelf manager for operator access and possibly scripting.

ID	Feature Name	Requirements
SM-43	CLI	Shelf managers ***may*** provide command line interface (CLI) through serial, ssh, and telnet interfaces.
SM-44	CLI	The Server Management Working Group (SMWG) within DMTF is working to define a standardized CLI format. At the writing of this version of the Design Guide, SMWG effort is not complete. It is recommended that shelf manager designers follow this effort and implement the format as per the SMWG specifications. When implementing shelf manager CLI before the release of the SMWG specification, shelf manager designers ***should*** align their implementations with SMWG effort as closely as possible, so that further updates to conform to SMWG requirements can be made smoothly.

7.11.2.2 Simple Network Management Protocol (SNMP)

The Simple Network Management Protocol (SNMP) is used to interact with system management software. SNMP is the primary interface into the shelf manager for monitoring and configuring the shelf and its parameters.

ID	Feature Name	Requirements
SM-45	SNMP	If SNMP is implemented, SNMP v2 and SNMP v3 ***should*** be supported.
SM-46	SNMP	If SNMP is implemented, Shelf managers ***should*** provide the ability to configure between SNMP v2 and v3.
SM-47	SNMP	If SNMP is implemented, Shelf managers ***should*** provide the ability to disable set operations through SNMP v2 (only get operations will be executed).
SM-48	SNMP	If SNMP is implemented, Shelf managers ***should*** support standardized SNMP MIB being developed by Systems Management Working Group (SMWG) in Service Availability Forum (SAF) that is intended to support AdvancedTCA. This MIB leverages AdvancedTCA to HPI mapping specification developed by Service Availability Forum, as described in Chapter 7.7. Eventually, the SNMP MIB supporting AdvancedTCA will be reviewed by PICMG for adoption into its specifications.
SM-49	SNMP	Since shelf management solutions may already support proprietary MIBs, a transition path ***shall*** be provided to customers, wherein support for both SNMP MIBs (proprietary and standardized) are provided so as to enable customers to migrate to standardized MIBs over period of time.
SM-50	SNMP	The root of the MIB in the tree ***shall*** be configurable such that the root MIB can be made to reside under any OEM's node.

SNMP traps should be sent to system management entity for events detected by the shelf manager.

ID	Feature Name	Requirements
SM-51	SNMP	Shelf managers ***shall*** send SNMP traps with severity level matching the severity of the event itself.
SM-52	SNMP	Shelf managers ***shall*** provide ability to send SNMP traps to at least 5 destinations. This allows for reliable transfer of event information when system management entity may be experiencing errors.
SM-53	SNMP	Shelf managers ***shall*** provide ability to configure SNMP trap destinations. (IP Address).
SM-54	SNMP	Shelf managers ***shall*** provide ability to configure SNMP Trap port.
SM-55	SNMP	Shelf managers ***shall*** provide ability to configure SNMP community name.

7.11.2.2.1 Platform Event Trap (PET) support

Platform Event Trap is used to send IPMI platform events to external network entities. PET data includes the actual data received in the Platform Event Request; along with ability to send normal event information, this capability also enables the shelf manager to forward events to the system manager that cannot be interpreted in standardized manner (e.g. Platform Events for OEM sensors).

ID	Feature Name	Requirements
SM-56	PET support	If standardized SNMP is not supported, Shelf managers ***shall*** support platform event traps (PET) as per PET v1.0 specification.
SM-57	PET support	If PET is supported, Shelf managers ***shall*** send platform events to SNMP trap address destinations as described in above section. This is required over and above any trap destinations set by PEF configurations.

7.11.2.3 Remote Management and Configuration Protocol (RMCP)

RMCP provides IPMI over LAN communication with the shelf manager and the boards in the shelf. RMCP is not suggested for use to integrate shelf to system managers. The intent of RMCP is for debugging and specification compliance only.

ID	Feature Name	Requirements
SM-58	RMCP	Shelf managers ***shall*** support MD2 and MD5 authentication.
SM-59	RMCP	Shelf managers ***shall not*** permit clear password and no authentication options on their RMCP interfaces for security purposes. (Refer response byte 3 of Get Channel Authentication Capabilities command in IPMI v1.5 specification.)
SM-60	RMCP	Shelf managers ***shall*** implement message bridging as described in clarification E315, Addenda, Errata, and Clarifications document revision 5 of the IPMI v1.5 rev 1.1 specification. http://www.intel.com/design/servers/ipmi/spec.htm

RMCP-based message bridging functionality allows system manager software to directly send messages to AdvancedTCA blades. While this functionality allows capabilities like IPMB-based firmware updates or some other specific feature, this capability could potentially be used by system manager software to directly send messages to boards to enable/disable E-Keys or some other capability that should be controlled by the shelf manager.

ID	Feature Name	Requirements
SM-61	RMCP	Since it is expected that the shelf manager is responsible for reliable operation of the shelf, shelf manager implementations ***should*** snoop commands being bridged from RMCP channel into the back plane IPMB-0 and incase commands fall into category that shelf manager ***should*** be in control of, the request ***should not*** be bridged to the board and an in sufficient privilege return code be returned to the requestor. This guideline is over and above authentication and privilege checks for commands as per IPMI specification.

7.11.2.4 Common Information Model (CIM)

The Common Information Model (CIM) is a common data model of an implementation-neutral schema for describing overall management information in a network/enterprise environment. CIM is comprised of a Specification and a Schema. The Specification defines the details for integration with other management models (i.e. SNMP's MIBs or the CMIP's GDMO) while the schema provides the actual model descriptions.

7.11.2.4.1 CIM Object Manager

The Object Manager is the key component of a CIM-compliant manageability framework. It is aware of the supported CIM schema, supports multiple independent providers to contribute concrete implementations and instantiations based on the CIM schema and responds to CIM operations requested by various entities (providers, management applications and consoles). The Object Manager is made schema-aware based on reading the CIM schema in the appropriate format: (a) MOF (Managed Object Format) via a MOF compiler, or (b) CIM/XML through the CIM/XML Handler. The Object Manager may subsequently store the schema in some internal representation format.

While DMTF defines CIM schema covering all possible objects that need modeling, AdvancedTCA Shelf Management requires a subset of the entire schema. The following section describes the requirement for supporting the subset of the schema for AdvancedTCA.

7.11.2.4.2 CIM Schema/ Profile

The Systems Management Working Groups (SMWG) within the Service Availability Forum is chartered with the task of mapping HPI to CIM, which would be the basis for the CIM schema and the profiles supporting AdvancedTCA and would also align with CIM schema extensions for modular servers being developed within DMTF. Eventually, the CIM schema and profiles will be reviewed by PICMG for adoption into its specifications. This model enables maximum reuse of schema definitions between Enterprise and Telco based CIM implementations.

ID	Feature Name	Requirements
SM-62	AdvancedTCA CIM Schema/Profile	Shelf managers ***should*** support CIM schema/profile being developed by the Systems Management Working Group (SMWG*) in the Service Availability Forum (SAF*).

7.12 Standards Interaction/Linkage

Given the standards based building block approach in Modular Communication Platforms (MCP), Figure 7-2 illustrates the features and interfaces that may be implemented on shelf management solutions along with information on which standards body develops these interfaces, allowing for maximized interoperability.

Note: Figure 7-2 is a conceptual illustration of shelf management architecture.

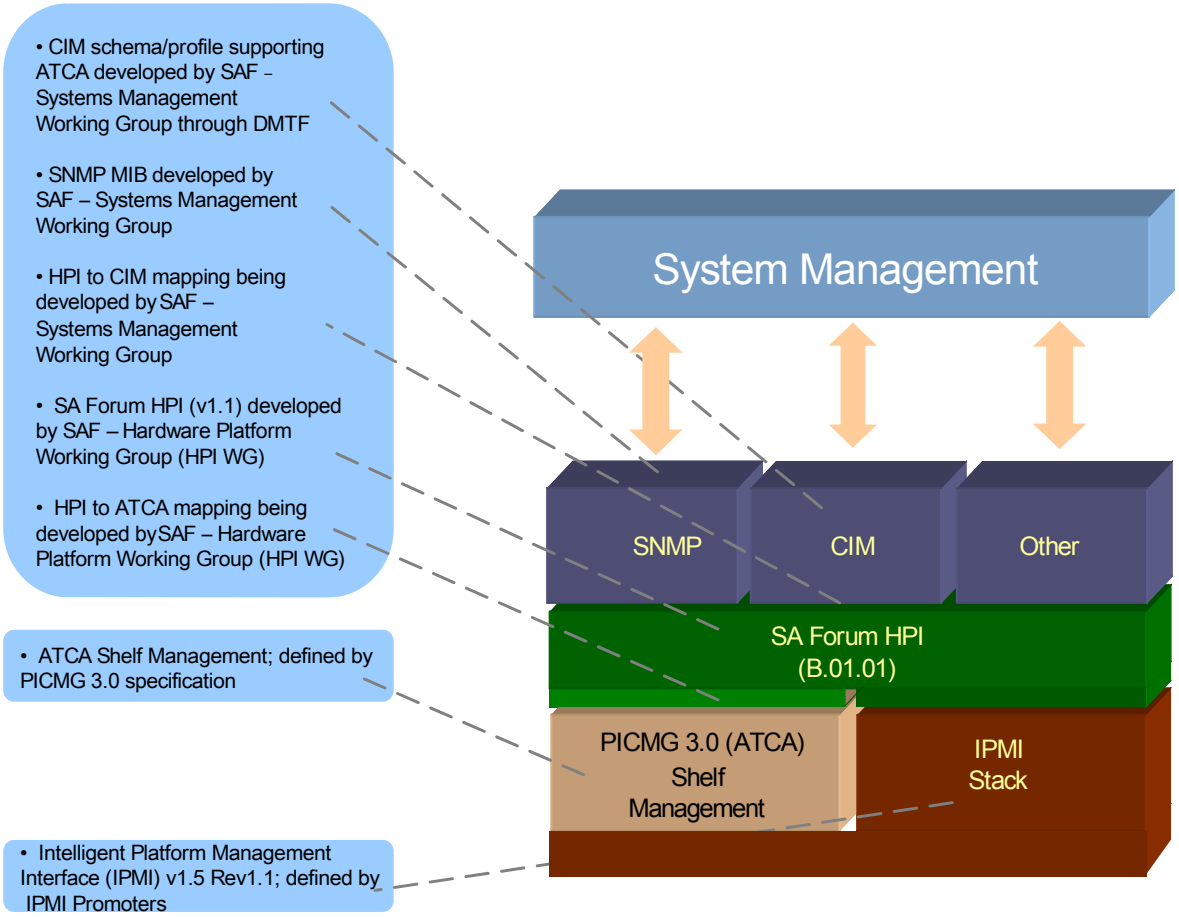


Figure 7-2 Shelf Management Architecture

The specifications being developed by the SAF* to support AdvancedTCA (AdvancedTCA to HPI mapping, SNMP MIB, CIM schema/profile) will be reviewed by PICMG for adoption/linkage into its specifications.

Given that some of the standardized interface definitions are in progress, Shelf management solutions should provide a transition path to customers so that migration to standard interfaces over period of time is achieved with small side effects.

8. AdvancedTCA Stratum Clock Distribution

8.1 System Issues

Many telecommunications applications using the AdvancedTCA architecture must connect to external networks, SONET/SDH networks for example, that require strict timing relationships between multiple line card interfaces and the external network timing reference. The AdvancedTCA specification defines dedicated clock interfaces across the backplane to facilitate synchronization of all network boards installed within a shelf. This section describes the use of AdvancedTCA synchronization clock interfaces to properly support and distribute a Stratum 3/3E (or 4/4E) external reference source within and between AdvancedTCA shelves. The system must properly capture the external reference timing source and distribute a synchronized timing signal (clock) to all boards within the shelf and possibly to additional shelves.

The AdvancedTCA specification (PICMG 3.0) defines backplane routing and termination requirements to facilitate reliable distribution of synchronization clocks to all boards within the shelf. However, additional guidance is warranted to ensure consistent, interoperable use of these interfaces. This section describes a modular design approach using mezzanine cards as a synchronization clock building block that can easily be integrated into an AdvancedTCA system. Design recommendations are provided to design the System Master Clock Generator (SMCG) mezzanine card which sources the system synchronization clocks along with baseboard design guidelines needed to support the SMCG or other mezzanine functions that make use of the backplane synchronization clocks. The SMCG circuitry can also be integrated onto specific AdvancedTCA boards rather than using a mezzanine card in which case the functionality is expected to match the described design recommendations.

8.2 Network Reference Clock

The digital SONET and SDH networks connect all system agents to a global reference timing source that adheres to strict accuracy requirements defined as Stratum levels. The Primary Reference Source (PRS) or Stratum 1 clocks source all network elements via a master-slave synchronized network comprising hierarchical clock sources that adhere to differing Stratum level quality standards. Proper connection to the network reference clock source is critical to maintaining system interoperability throughout the network and to minimizing dropped frames and effects of jitter that can impact voice quality.

8.3 Stratum Clock Network

With the SDH and SONET networks, all the clocks are traceable to a master clock (PRC), the requirements for which are defined in G.811 (PRC) by the International Telecommunication Union (ITU). These are equivalent to the Stratum 1 clocks or PRS, which are defined by North American Standards.

The secondary levels of clock regenerators deployed in the network are called the node clocks or Synchronization Supply Units, and are defined by the G.812 from ITU-T. They are equivalent to stratum 2 clocks from the North American standard. These are also called SETS and have the same level as the Building Integrated Timing Supply (BITS) clocks, which provide clocks to an entire system or building and the performance of which is very close to the stratum 1 clock (PRC). As part of SDH equipment, the Synchronization Equipment Timing Supply (SETS) Unit is defined in G.783, which describes the building blocks of SDH equipment. The clock equipment function is commonly called the Slave Equipment Clock (SEC); equivalent to the SONET Minimum Clock (SMC) standards defined by the North American standards (Stratum 3E). Figure 8-1 shows the Master-slave synchronization hierarchy structure and relevant ITU and ANSI quality standards.

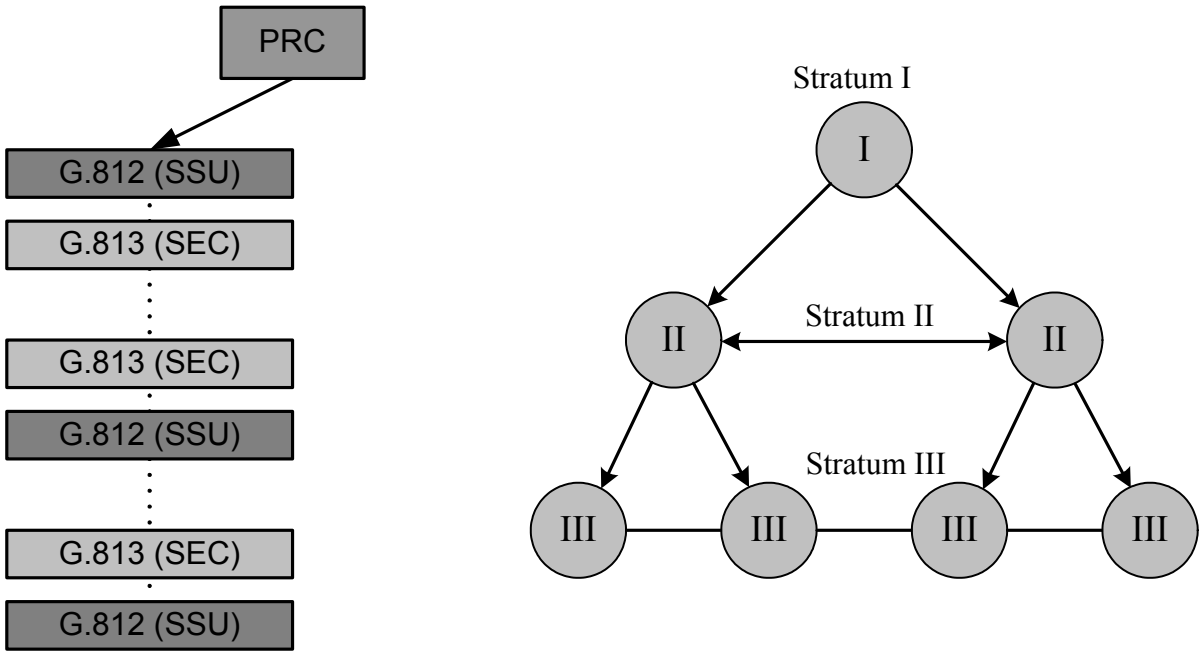


Figure 8-1 ITU and North American Synchronization Structure

8.3.1 Applicable GR and ETSI Standards

1. PRC according to ITU-T G.811
2. SSU/BITS according to ITU-T G.812
3. SEC according to ITU-T G.813
4. SETS according to ITU-T G.783
5. Stratum 3/3E according to Telcordia GR-1244-CORE
6. Stratum 4/4E according to Telcordia GR-1244-CORE

8.4 External Timing Reference Connection Models

SONET and SDH network equipment may be configured to support different timing modes.

Timing Mode	Description
External Timing mode	This is the preferred mode of synchronizing SONET network elements. In this mode a higher quality clock (G.812 or BITS wherever applicable) traceable to the PRS/ PRC is used to synchronize the system. Non-traffic carrying primary rate signals is input to the System Master Clock Generation Unit to generate the synchronized timing clocks (system clocks) that can be used by all system agents.
Line Timing mode	In this mode the Network Element (NE) has the capability to directly derive clock timing from an incoming OC-N signal and time all outgoing OC-N, STS-1, and STS-3 signals from this clock. The primary application for this timing mode is in access rings where no BITS clock is available for external timing.

Timing Mode	Description
Internal Timing mode	In this mode the Network Element uses its internal clock to time the outgoing OC-N/STS-1 signals. This timing mode is used in point-to-point configurations when no BITS clock is available in the office or the node is not capable of generating Hold-over clocks.
Hold-over timing mode	In this mode the clock module is able maintain frequency and phase when it loses the reference, using a Phased Lock Loop (PLL) to generate the frequency locked timing for a period of time while awaiting the return of a valid reference. Hold-over capability and design requirements may vary and this mode is intended only for temporary conditions.

Several connection models to reference synchronization signals are discussed in this section.

8.4.1 External Timing Mode - A Building Integrated Timing Supply (BITS)

A BITS connection scheme requires a simply redundant connection to each platform as shown Figure 8-2. In this External Timing configuration the shelf is physically connected to a Primary Reference Clock (PRC). The external reference is connected directly to the Primary and Secondary system Master Clock Generators which capture the Reference Clock and distributes synchronized clocks to other boards within the shelf.

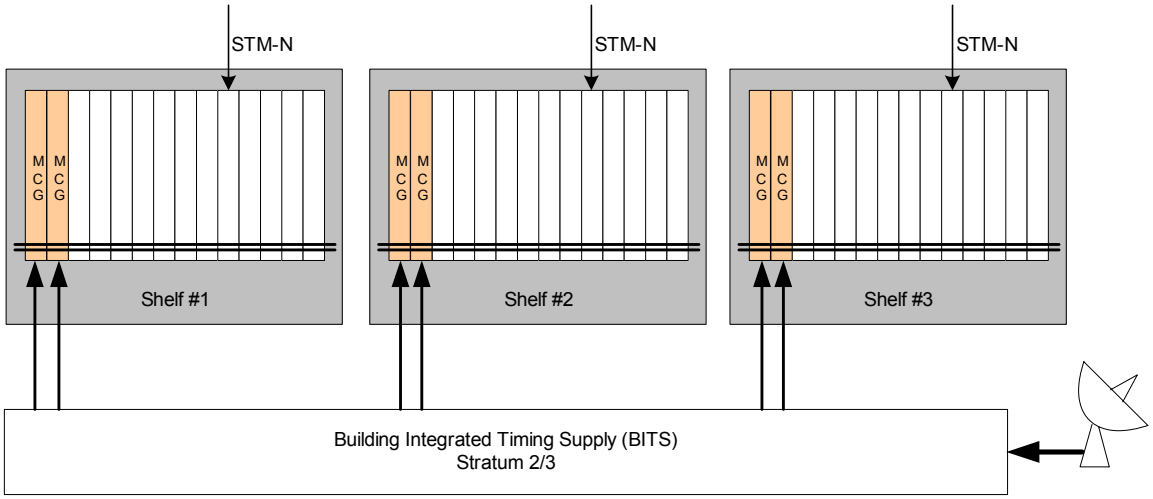


Figure 8-2 BITS Signal Distribution

8.4.2 External Timing Mode - Synchronization Supply Unit (SSU)

SSU distribution requires the connected platforms to pass a line captured reference signal to the SSU (Dual outputs) and the SSU returns a phase synchronized signal back to the platform. The system Master Clock Generator supports the external connections to the SSU and distributes synchronized clocks to other boards within the shelf.

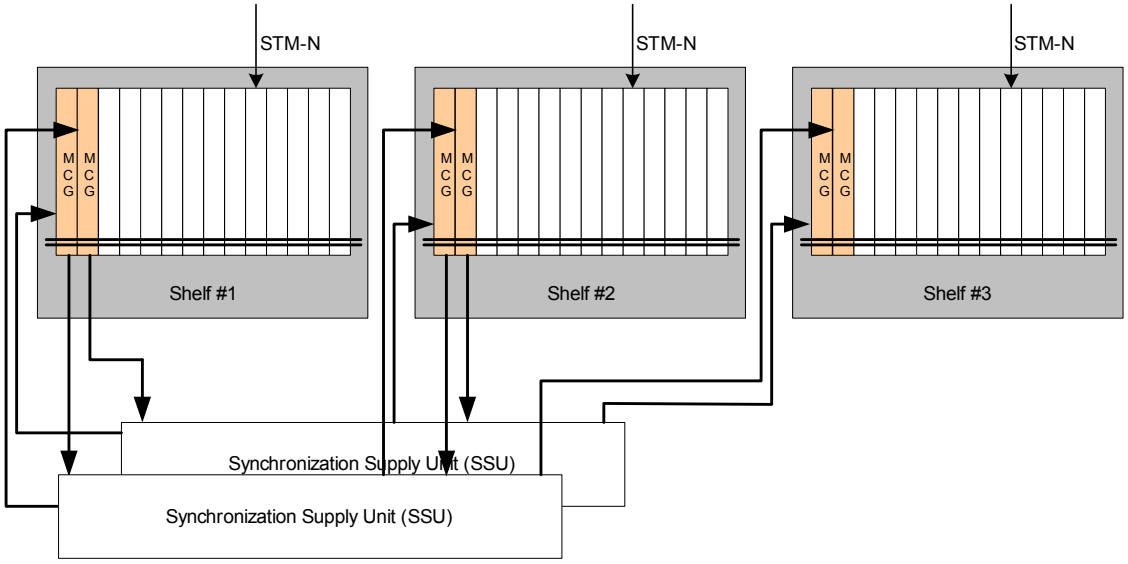


Figure 8-3 SSU Reference Signal Distribution

8.4.3 Line Timing Mode

In this connection model the network reference timing source is derived directly from an incoming OC-N signal on a typical line card rather than a dedicated reference connected directly to the system MCG. The line card forwards the reference timing source to the system Master Clock Generator via the AdvancedTCA backplane (CLK3A/3B). The system Master Clock Generator can then supply the synchronized clocks (CLK1 and CLK2) needed for other boards within the system and supply reference timing externally to additional shelves (children) as shown in Figure 8-4.

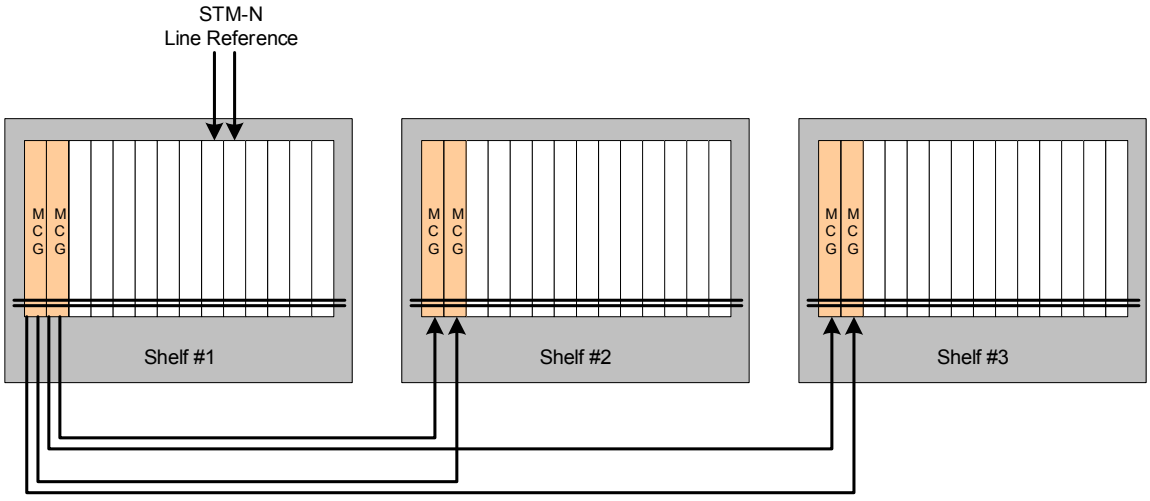


Figure 8-4 Parent/child Reference Clock Distribution

8.4.4 Redundant Connections

Each system connected to a network (network element) must have redundant access to the external reference timing source. Redundant access means each element (shelf) has at least two connections to the network timing source; one acts as the Primary connection and the other as the Secondary (Back-up) connection. To maximize fault tolerance, the Primary and Secondary connections should reside on separate boards within the shelf. In the event the primary source becomes unstable, the system must be capable of switching-over to the secondary source within strictly defined Stratum level requirements. The AdvancedTCA specification defines redundant connections for each clock interface to facilitate Primary/Secondary switch-over. This allows each board in the shelf to have direct connection to the Primary and Secondary reference timing source. Each board that receives the reference timing source must monitor the incoming clocks and switch-over if the Primary line becomes unstable.

Primary and Secondary timing agents may utilize a hardware connection (cable) between them coordinate switch-over events and/or to synchronize phase and frequency of their outputs. Boards that source timing clocks to the system may utilize Hold-over mode to maintain the internal synchronous clock output until the external reference is recovered or switched over to the Secondary (back-up) input. Use of Hold-over mode is only acceptable for limited duration and cannot replace a Secondary timing reference connection.

8.5 AdvancedTCA Synchronization Clock Interface

The AdvancedTCA specification defines three sets of bused lines across the backplane (the Synchronization Clock Interface) to facilitate distribution of system clocks to any installed boards within the shelf. All AdvancedTCA backplanes are required to support these bused lines, while boards may support a connection to any, all, or none of the clock buses depending upon the needs of the application.

The Synchronization Clock Interface supports differential signal connections bused to all system slots in the shelf and are intended for MLVDS signaling. The lines are terminated on the backplane. The three separate clock interfaces (labeled as CLK1A & 1B, CLK2A & 2B, and CLK3A & 3B) are defined as redundant connections for primary-to-secondary fail-over in the event the primary clock source is interrupted or fails. The CLK1A/1B lines are designated to carry 8 kHz signals and CLK2A/2B is required to carry 19.44 MHz signals for system synchronization purposes. The CLK3A/3B lines are user definable and, thus, the signal frequency may vary from one system to another. The AdvancedTCA specification does not dictate how these clock sources are derived.

ID	Feature Name	Requirements
SI-1	Synchronization Clock Interface	For systems incorporating Line Timing mode, CLK3A/3B ***should*** be used to distribute the captured external reference source (typically 2.048 MHz or 1.544 MHz). Only boards directly connected to the line interface may drive the CLK3 bus. Only boards supporting the Master Clock Generator (MCG) function shall receive the CLK3A/B signal.
SI-2	Synchronization Clock Interface	The CLK1A/1B and CLK2A/2B buses ***shall*** be sourced only by boards that support the system Master Clock Generator (MCG) function. The MCG ***shall*** provide 8 kHz and 19.44 MHz clocks that are frequency and phase locked (synchronized) to the external reference timing source.
SI-3	Synchronization Clock Interface	All boards needing connection to the system Synchronization Interface ***shall*** ensure proper electrical signaling for the Synchronization Interface as described in the PICMG 3.0 specification.

8.5.1 Clock Signaling

This section provides a summary of electrical requirements for AdvancedTCA backplanes and boards for each of the Synchronization Interface buses. Refer to PICMG 3.0 Section 6.7.3 for signaling requirements.

The clock lines on the AdvancedTCA backplane are intended for M-LVDS differential signaling configured as a multi-drop bus with a differential impedance of 130 Ohms +/- 10%. Each differential signal pair is length matched to better than 8.5 ps and is terminated on the backplane.

8.5.2 CLK 1

ID	Feature Name	Requirements
SI-4	Synchronization Clock Interface	All CLK 1A or CLK 1B sources ***shall*** have a nominal frequency of 8 kHz and be specified to meet or exceed Stratum 4E clock requirements (Telcordia GR-1244-CORE). The rising edge of the 8 kHz clock ***shall*** occur at regular 125uS intervals. Rising edge skew between the Primary and Secondary clock ***shall*** be less than 10 nS.

8.5.3 CLK 2

ID	Feature Name	Requirements
SI-5	Synchronization Clock Interface	All CLK 2A or CLK 2B sources ***shall*** have a nominal frequency of 19.44 MHz and be specified have a duty cycle of 50% +/-10%. CLK 2A/2B signals ***shall*** meet jitter attenuated clock requirements of a Stratum 3 or 3E clock (Telcordia GR-1244-CORE) or SDH Equipment Clock Option I or II (per ITU-T G.813). Rising edge skew between the Primary and Secondary clock ***shall*** be less than 10 nS.

8.5.4 CLK 3

The quality requirements for CLK 3 and CLK 3A buses are not specified in the AdvancedTCA specification with the following exceptions: Clock signals for CLK 3A/3B must have a frequency of less than 100 MHz. Also, when the CLK 3A/3B clock signal is derived from an external network timing reference that is a multiple of 8 kHz, it must have a frequency of either 8 kHz, 1.544 MHz, 2.048 MHz or 19.44 MHz and the rising edge shall occur on regular intervals with a high time of at least 45 ns.

8.5.5 Clock Distribution Models

The AdvancedTCA specification describes two models, centralized and distributed, for use of synchronization clocks between backplane connected boards. In the centralized model, synchronized clock sources are driven by dedicated boards within the shelf. Source redundancy is handled by having a Primary and Secondary agent for sourcing the system synchronization clocks. In the distributed model, each board in the shelf is capable of sourcing any of the synchronization clocks to the system. While the distributed model offers a higher level of redundancy and flexibility, it is more difficult to manage. For this reason, this section focuses on the centralized distribution model in which the central clock distribution agent is a mezzanine based Master Clock Generation function.

8.5.6 Redundant Clock Usage

The A/B redundant buses can be thought of as Primary and Secondary lines for each clock. This allows redundant clock sources to be available to all system boards simultaneously. For high availability, Primary

and Secondary clock sources (system Master Clock Generators) are typically supplied by different boards within the shelf. The Primary clock source drives the A or B bus line and the Secondary clock source drives the other line. Assignment of Primary and Secondary agents and which corresponding bus line each board drives, requires some system management coordination (refer to Section 8.5.1). Primary and Secondary clock sources are also required to minimize nominal phase differences (rising edge skew) between the two clock outputs. In the event of separate boards supplying the Primary and Secondary system clocks, the boards may utilize an external connection to maintain phase synchronization. The backplane Update Channel interface can be used for this purpose or an external (cabled) connection via the Zone 3 Rear Transition Module (RTM) interface or the front panel. Because of the tight synchronization requirements between the Primary and Secondary agents, it is expected these boards (mezzanine cards) come from a common supplier. Multi-vendor interoperability between Primary and Secondary agents is not expected.

Boards receiving synchronization clocks are required to monitor the status of both A and B lines and be capable of dynamically switching between Primary and Secondary (A and B lines) sources in the event of a fault on the Primary. This A/B switch-over mechanism must occur without impacting any external interfaces and, thus, must be managed in hardware for maximum performance.

ID	Feature Name	Requirements
SI-6	Synchronization Clock Interface	For high availability, Primary and Secondary clock sources (system Master Clock Generators) ***shall*** be supplied by different boards within the shelf. The Primary source agent should drive CLKA while the Secondary agent drives CLKB buses.
SI-7	Synchronization Clock Interface	Primary and Secondary clock sources ***shall*** minimize nominal phase difference (rising edge skew) between the A/B clock outputs. The boards ***may*** utilize an external connection to maintain phase synchronization. The backplane Update Channel interface or an external (cabled) connection via the Zone 3 Rear Transition Module (RTM) interface or the front panel ***may*** be used for this purpose.
SI-8	Synchronization Clock Interface	Boards receiving synchronization clock inputs ***shall*** monitor the status of both A and B lines and be capable of dynamically switching between Primary and Secondary (A and B lines) sources in the event of a fault on the Primary. This A/B switch-over mechanism shall occur without impacting any external interfaces and, thus, must be managed in hardware for maximum performance

8.6 AdvancedTCA Bused Electronic Keying Mechanism

Boards wishing to drive signals onto the backplane Synchronization Clock Interface must request permission of the Shelf Manager before enabling clock driver outputs. Specific IPMI commands defined in PICMG 3.0 allow agents to request control of a bused interface. Commands allow the Shelf Manager to have a board relinquish control of the bus and allow the Shelf Manager to query the board's status to determine who owns the bus at any point in time. These commands allow the system to configure and manage the bused interfaces and dynamically respond to bus requests.

The AdvancedTCA system management bused electronic keying mechanism has provisions for identifying requests from boards wanting to drive the Primary clock source to the system. Agents requesting to drive the system clock must provide a high impedance to any/all Synchronization Interface buses until it receives acknowledgement from the Shelf Manager. Once designated by the Shelf Manager to be the Primary clock source, that agent can designate a Secondary source agent through an out-of-band interface. The Shelf Manager does not designate the secondary (back-up) clock agent.

Synchronization of the rising edge skew between the Primary and Secondary agent can also be managed via an out-of-band interface.

In the event of a Primary clock failure, a secondary agent must be available to replace the failed clock. In most cases, however, this fail-over event is handled by a hardware mechanism between previously assigned Primary and Secondary clock agents. Regardless of the switch-over method employed, whenever the Primary timing connection fails the Shelf Manager must be notified.

8.7 Modular Clock Synchronization

The goal of the Modular Communications Platform is to identify system building blocks and ensure they are easy to integrate into systems. The modular approach proposed in this section describes the design requirements for an AdvancedTCA Board to properly interface to the AdvancedTCA synchronization clock backplane buses whether receiving or driving synchronized system clocks or network reference clocks to/ from the backplane. The modular approach is based on the following system assumptions:

- The centralized system clock source agent will have a System Master Clock Generation mezzanine card (or integrated circuit of similar function) installed on the baseboard. System Master Clock enabled boards will drive phase synchronized signals onto the CLK 1 (8 kHz) and CLK 2 (19.44 MHz) buses.
- Systems will support redundant connections to external network reference signal and support automatic switch-over if the Primary input signal becomes unstable.
- All boards needing connection to the system synchronization clocks shall ensure proper electrical signaling for the Synchronization Interface as described in the PICMG 3.0 specification. Boards are not required to connect to all three buses of the Synchronization Interface.
- All boards receiving synchronization signals from the backplane connect to both A and B buses (Primary and Secondary) and automatically switch-over to the Secondary input in the event the Primary input becomes unstable. Due to response time requirements, switch–over is typically handled by receiver hardware circuitry. Notification of the switch-over event may be sent to the board's IPMC.
- All boards will properly communicate Synchronization Interface connections to the AdvancedTCA Shelf Manager as defined in PICMG 3.0.

These functions can be facilitated via three hardware building block components:

1. The system Master Clock Generator (MCG) can be implemented as a mezzanine add-in card
2. The external clock capture function can be implemented as a mezzanine add-in card; either integrated into the system MCG (External Line Mode) or into the line interface card (Line Timing Mode).
3. The backplane Synchronization Interface support circuit (Clock Distribution Circuit) resides on all AdvancedTCA boards and directs clock signals to/ from the system backplane and the MCG and/ or line interface mezzanine cards.

8.7.1 System Master Clock Generator Design Requirements

Implementing the system Master Clock Generator (MCG) function as a mezzanine card allows it to be added to any system as modular building block component. There are many high quality PLL devices available today that could be designed onto a small form factor mezzanine add-in card to provide Stratum 3/3E or 4/4E quality clock sourcing to an AdvancedTCA platform.

The function of the MCG unit is to lock onto a reference clock signal derived from line interfaces or external signals, monitor its quality, and generate a high stability, low jitter clock signal to be distributed to all units in the shelf and to make it available outside of the shelf as a synchronization means for other equipment. This centralized distribution model works in primary/secondary arrangement in which each MCG unit is installed onto a different AdvancedTCA board, one sourcing the CLKA buses and the other

sourcing the CLKB buses. In this manner, the Primary and Secondary MCGs provide non-interrupted, transient-free service, even in case of failure of one of the two units. Agents receiving the synchronized clocks from the Primary and Secondary MCG sources must monitor both the A/B inputs and switch-over inputs if the Primary becomes unstable.

Figure 8-5 and Figure 8-6 depict a centralized distribution model with dedicated Primary and Secondary system MCG sources for External Timing Mode and Line Timing Mode respectively.

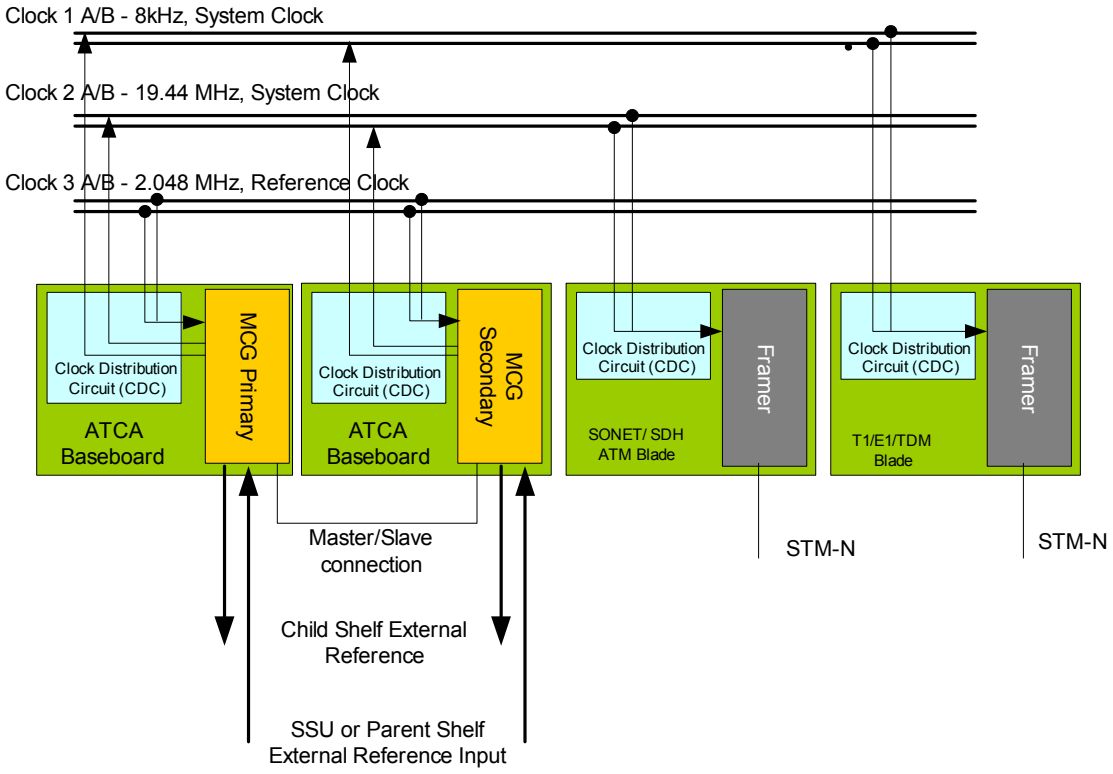


Figure 8-5 Centralized System Clock Distribution with External Line Reference

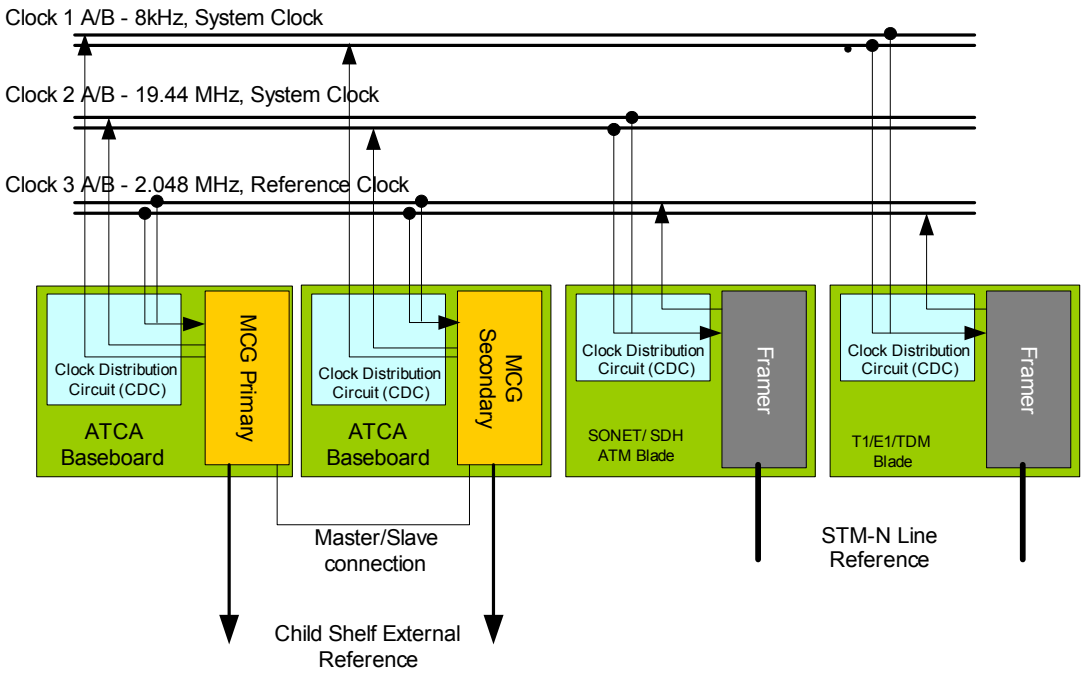


Figure 8-6 Centralized System Clock Distribution with Line Reference

8.7.1.1 Advanced Mezzanine Card Clock Support

The Advanced Mezzanine Card (AMC) specification defines three clock lines to match the three buses of the AdvancedTCA Synchronization Interface, making it a good vehicle for the system Master Clock Generator function. Other mezzanine definitions such as PMC may also be used. Using an AMC module to supply the system Master Clock Generation (MCG) function is an easy and flexible way for a system integrator to add the this building block service to any platform. The MCG mezzanine card can be installed onto any two boards in the system including the switch fabric boards (Hub Boards) in the system. This section describes an AMC mezzanine based Master Clock Generation module that supports the various requirements for sourcing system clocks within the shelf and among multiple shelves.

8.7.1.2 Master Clock Generator Features

The system Master Clock Generator (MCG) shall support the following features:

ID	Feature Name	Requirements
SI-9	Synchronization Clock Interface	MCG ***shall*** source system synchronization clock outputs for CLK 1 (8 kHz) and CLK 2 (19.44 MHz) AdvancedTCA clock synchronization interfaces. • MCG card outputs ***shall*** be driven onto either CLK 1A/2A buses or onto CLK 1B/2B depending on whether it is the Primary or Secondary system clock source. Complies with ITU-T G.812 (Type I) clock node requirements for wander and/or jitter quality and ITU-T G.813, G.823, G.824 and Telcordia GR-1244-CORE, GR-253-CORE and/or GR-499-CORE for PDH interfaces. • Internal switch-over between clock units ***shall*** have no impact on external interfaces according to standards ITU-

ID	Feature Name	Requirements
		T G.812, G.813 and Telcordia GR-253-CORE, GR-1244-CORE.
SI-10	Synchronization Clock Interface	MCG ***shall*** support dual external inter-shelf reference outputs (2.048 MHz/1.544 MHz synchronized to main PLL) • The output clock interface at 2.048 MHz (called T4 according to ETSI EN 300 462-2-1 V1.2.1) at 75/120 Ohms (G.703). • Derived DS1/T1 interface according to GR-253-CORE at 100 Ohms.
SI-11	Synchronization Clock Interface	The MCGs ***shall*** support a synchronization signal between two physically connected (cabled) MCG mezzanine cards. Switch-over between Primary/Secondary MCGs ***shall*** have no impact on external interfaces according to G.812, G.813, GR-253-CORE, GR-1244-CORE.
SI-12	Synchronization Clock Interface	MCG ***shall*** support dual (Redundant) inputs for external reference signal (DS1/E1/2.048 MHz/1.544 MHz) • The MCG card can directly receive external timing via redundant 2.048 MHz (called DS1 interfaces according to ETSI EN 300 462-2-1 V1.2.1) with 75/120 Ohm according to ITU-T G.703. • The MCG can derive its synchronization input via redundant DS1/T1 interfaces according to Telcordia GR-253-CORE.
SI-13	Synchronization Clock Interface	The MCG ***shall*** be capable of receiving a reference input from CLK 3 from which to generate synchronization clock outputs to CLK1 and/or CLK 2.
SI-14	Synchronization Clock Interface	The MCG ***shall*** support Primary/Secondary select input. (Direct connect between AMC clock generation modules), (8kHz MLVDS). This is a synchronization signal between two physically connected (cabled) MCG mezzanine cards.
SI-15	Synchronization Clock Interface	Mezzanine based MCG ***shall*** support a Control Interface (PCI Express to carrier board) to facilitate processor control to configure mezzanine functions.

Figure 8-7 shows the various connections between the MCG and the AdvancedTCA base board.

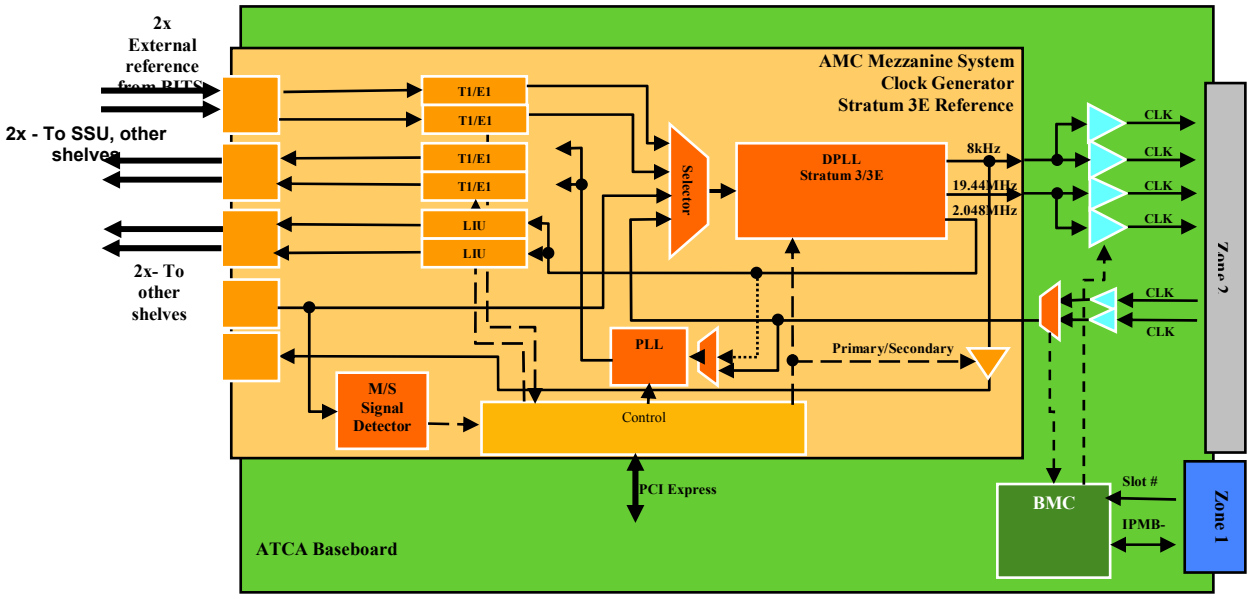


Figure 8-7 Mezzanine based System Clock Generator

8.7.1.3 Primary / Secondary Select and Synchronization

A direct connection between the Primary and Secondary MCG is a simple means to ensure phase and frequency coordination between the two sources. The Primary agent can drive a signal to the Secondary to indicate that it is actively sourcing the system clock resources (on CLK 1A and/or CLK2A). If the connection provides a synchronized clock output, the Secondary agent can also use this signal to phase synchronize its outputs that are driven onto CLK1B and/or CLK2B. Restart of the original Primary resource may be managed via AdvancedTCA IPMI interface or by application software. This cabled connection can be facilitated via the front panel, rear transition module (zone 3) interface, or the Update Channel interface (zone 2).

8.7.2 Base Board Requirements

All boards needing network timing synchronization are required to support a local Clock Distribution Circuit (CDC) for properly interfacing to the backplane synchronization interface band routing clocks to local resources and mezzanines.

8.7.2.1 Baseboard Clock Interface Design Requirements

The baseboard must be able to detect the presence of the MCG mezzanine and communicate the capability to drive system synchronized clocks and/or reference sources onto the backplane. (Define IPMI detection sequence) The baseboard Clock Distribution Circuit (CDC) must connect the mezzanine clock signals to appropriate CLK lines. Clock outputs to the backplane must be capable of switching between either A or B lines, but need not drive both simultaneously. Backplane inputs must monitor and select A or B lines to route the signal to the mezzanine cards and other board resources as needed.

To support the centralized system clock distribution model, all boards needing to synchronize to the network reference signal will operate in one of several modes:

- Boards that only receive the system synchronized clocks shall monitor both A and B bus inputs for CLK 1 (8 kHz) and/or CLK 2 (19.44 MHz).
- Boards that drive synchronized system clock signals (MCG mezzanine installed) shall source either the A or B lines for CLK 1 and CLK 2. These boards will also be capable of receiving and monitoring CLK 3A and 3B.

- Boards that capture the external synchronization signal from the network line connection (Line Timing Mode) shall drive a reference signal onto CLK 3A or CLK 3B.

All boards that need to synchronize to network reference timing signals must support a local clock distribution circuit (CDC) that provides the following functions:

- Provides an electrically compliant interface to the AdvancedTCA backplane clock buses.
- Should be configurable to support various modes of operation. For example, boards that need only receive synchronization timing from the backplane may only receive signals from the CLK 1A/1B buses and will configure the CLK 1 and CLK 2 connections as inputs. If, however, the board has a MCG mezzanine card installed, it may need to source the CLK 1A/1B or CLK 2A/2B buses and will configure these connections as outputs.

ID	Feature Name	Requirements
SI-16	Synchronization Clock Interface	All AdvancedTCA boards needing to interface to the backplane Synchronization Interface for either on-board or mezzanine based resources ***shall*** support a Clock Distribution Circuit (CDC) with the following capabilities: • Backplane outputs ***shall*** be switchable to drive either A or B lines • Boards that only receive the system synchronized clocks (backplane inputs) ***shall*** monitor both A and B bus inputs for CLK 1 (8 kHz) and/or CLK 2 (19.44 MHz) • Boards that drive synchronized system clock signals (MCG mezzanine installed) ***shall*** source either the A or B lines for CLK 1 and CLK 2. These boards ***shall*** also be capable of receiving and monitoring CLK 3A and 3B. • Boards that capture the external synchronization signal from the network line connection (Line Timing Mode) ***shall*** be capable of driving a reference signal onto CLK 3A or CLK 3B

8.7.2.2 Control of Clock Interfaces

The AdvancedTCA specification defines a Bused Resource Control Command to facilitate the negotiation of agents wishing to source the Synchronization Interface buses. This section describes the usage scenarios of bused resource control commands to facilitate the negotiation of clock synchronization buses (CLK1A/1B, CLK2A/2B).

- IPMI Bused Resource Control Command shall be used to request Primary system clock ownership. The centralized distribution model means this process is performed at system initialization only (not dynamic). The agent requesting to source a clock bus issues this command to the Shelf Manager. If granted permission this agent becomes the Primary agent and may source both clock A/B lines on the backplane (e.g. CLK1A/1B). If a redundant source is desired to simultaneously drive one of the bused lines, the Primary agent must negotiate and assign the Secondary agent permission to do so. The IPMI Bused Resource Control Command is not used to determine the Secondary agent. The agent requesting control of CLK1 and/or CLK2 buses shall have system MCG capabilities.
- The hardware connection between MCG cards determines the Primary/Secondary relationship with each other. The “secondary” card must drive the alternative bus to the Primary MCG. This configuration need not change even if the Primary/Secondary status reverses due to a glitch on

the Primary. This connection should also be used to provide phase synchronization between Primary and Secondary MCGs.

- Determination of the Primary source of the STM line captured reference onto CLK 3A/B buses shall be managed via Bused Resource Control Command. Boards with STM line interfaces request control of CLK 3A/B buses and assigns permission to Secondary agents as needed.

There is a need for defining the common software protocols and interfaces which will resolve the following issues:

- Error notification from source agents to Shelf Manager
- Control functions between carrier and mezzanine MCG card

8.8 Summary

Distribution of synchronization clocks referenced to a network timing source is a fundamental requirement of all systems connected to digital networks such as SDH/SONET. AdvancedTCA has been architected to enable the distribution of clocks within the system and proper use of the AdvancedTCA backplane Synchronization Clock Interface can facilitate interoperability of multi-vendor boards within a common shelf. In addition to proper connection to the backplane, a clock distribution circuit on AdvancedTCA boards must also facilitate clock interfaces to mezzanine sites to enable the integration of Master Clock Generation and Line Interface functions that source synchronization clocks and capture external reference clocks for use by the entire system. The AMC mezzanine specification is particularly well suited to ensure easy integration of these important clock functions due to defined interfaces for clocking functions that match the AdvancedTCA interfaces. Definition of common software protocols and interfaces must be further developed to complete the modular design approach and maximize multi-vendor interoperability of the clocking hardware building blocks.

9. AdvancedTCA Automatic Protection Switching

This chapter defines the support for Automatic Protection Switching (APS) in SONET end systems such as Routers, Media Gateways and Multi Service Switches. More general SONET transport network support (required for ring-based Add/Drop Multiplexers, Digital Cross Connects etc.) is not considered.

Note: *At the writing of this version of the Design Guide, the information provides example of the support currently provided. The requirements to be complied with will be added in a later draft of this document.*

9.1 Reference Documents

1. Telcordia GR-253-CORE: SONET Transport Systems
2. ITU-T G.841: Types and Characteristics of SDH network protection architectures
3. ITU-T G.808.1: Generic Protection Switching – Linear Trail and Subnetwork Protection network
4. OIF Implementation Agreement OIF-SFI4-01.0 – Proposal for a common electrical interface between SONET framer and serializer/ deserializer parts for OC192 interfaces
5. OIF Implementation Agreement OIF-TFI5-0.1.0 – TFI-5 TDM Fabric to Framer Interface

9.2 SONET and APS

The SONET Section, Line and Path are shown in Figure 9-1 with their associated overhead in the STS-N frame. Note: SDH has the same structure but uses the terms Regenerator Section, Multiplex Section, and Path.

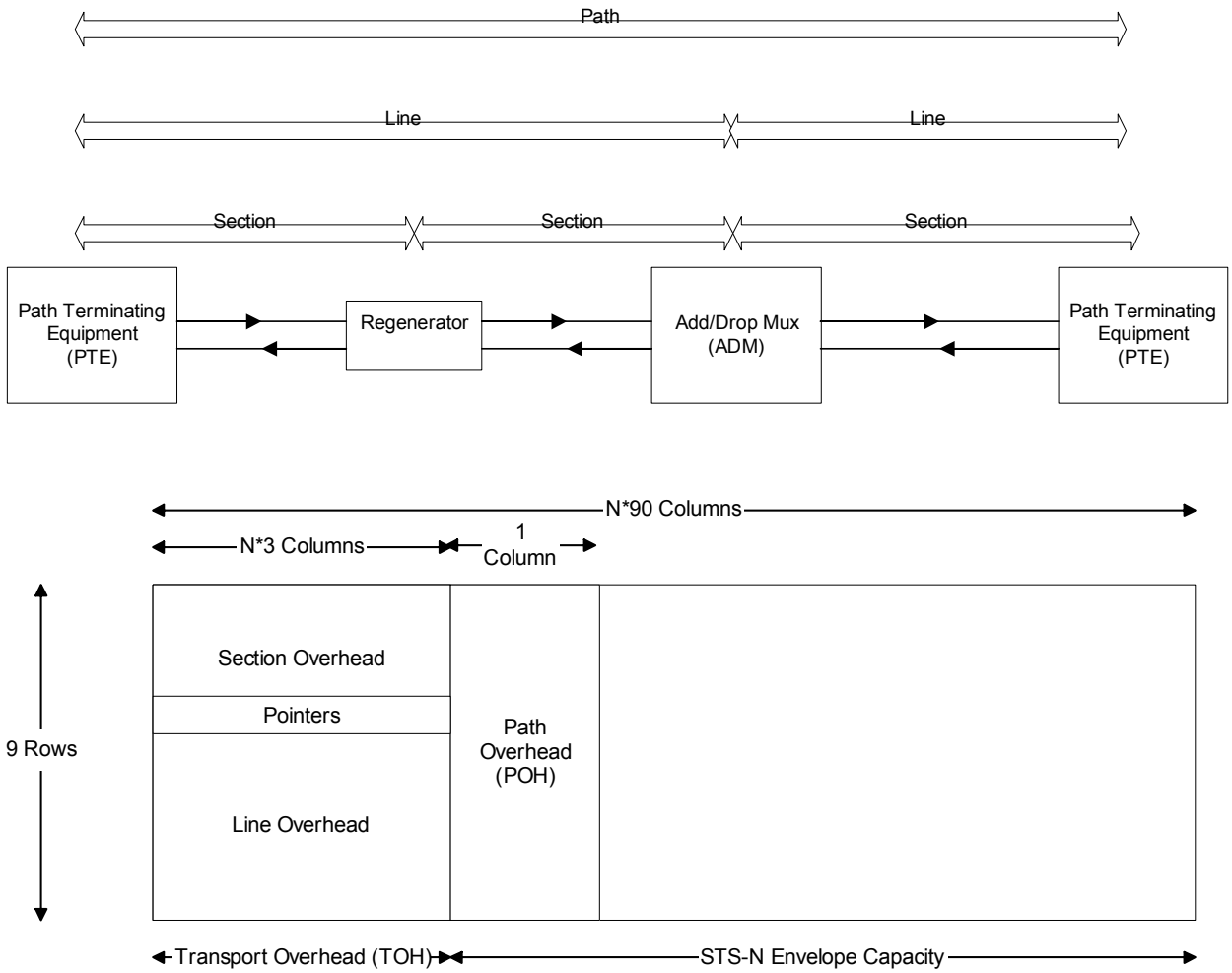


Figure 9-1 SONET Section, Line, and Path

Several Automatic Protection Switching (APS) schemes have been defined for SONET. This section addresses only Linear APS operating between the Path Terminating Equipment (for example, a Router, Media Gateway, ATM switch) and the adjacent Add/Drop Multiplexer. SONET protection schemes used (by Add/Drop Multiplexers) in SONET rings such as Bidirectional Line Switched Rings (BLSR) and Unidirectional Path Switched Rings (UPSR) are not covered. Linear APS operates at the SONET Line layer between two nodes connected by one or more working links and a single protection link (the SDH equivalent is called Multiplex Section Protection or MSP). There are two general APS architectures: 1+1 and 1:n (of which 1:1 is a special case).

In the 1+1 architecture, the head-end signal is permanently bridged to the working and protection lines so that the same payload is transmitted on both lines. The tail-end receiving equipment selects either the

working or protection signal based on local switching criteria (e.g. signal failure, signal degrade etc.). Unidirectional switching is the default 1+1 mode and only requires the failed line to be switched, that is, the receiver simply selects the protection line instead of the working line. Bidirectional mode can be configured and requires that both directions are switched on a single failure. This co-coordinated switching is effected by using an in band control protocol carried in the K1, K2 bytes of the Line overhead on the protection line. Note: Although coordinated switching is not required in unidirectional mode the K1, K2 bytes are still used to indicate local actions and mode of operation (uni/bidirectional).

By default, 1+1 switching uses nonrevertive switching where the protection line is still used after the working line has been repaired until a manual command to switch back has been invoked; revertive mode (where the traffic is automatically switched back to the working line on repair) may also be provided as a configurable option. In the 1:n architecture the head-end signal from any of the n (1 to 14) working lines can be bridged to the single protection line. Head-end to tail-end signaling uses the K1, K2 bytes and because the head-end is switchable, the protection line can be used to carry extra (low priority) traffic when not in use. All switching is revertive and both unidirectional and bidirectional modes are applicable (with bidirectional as the default).

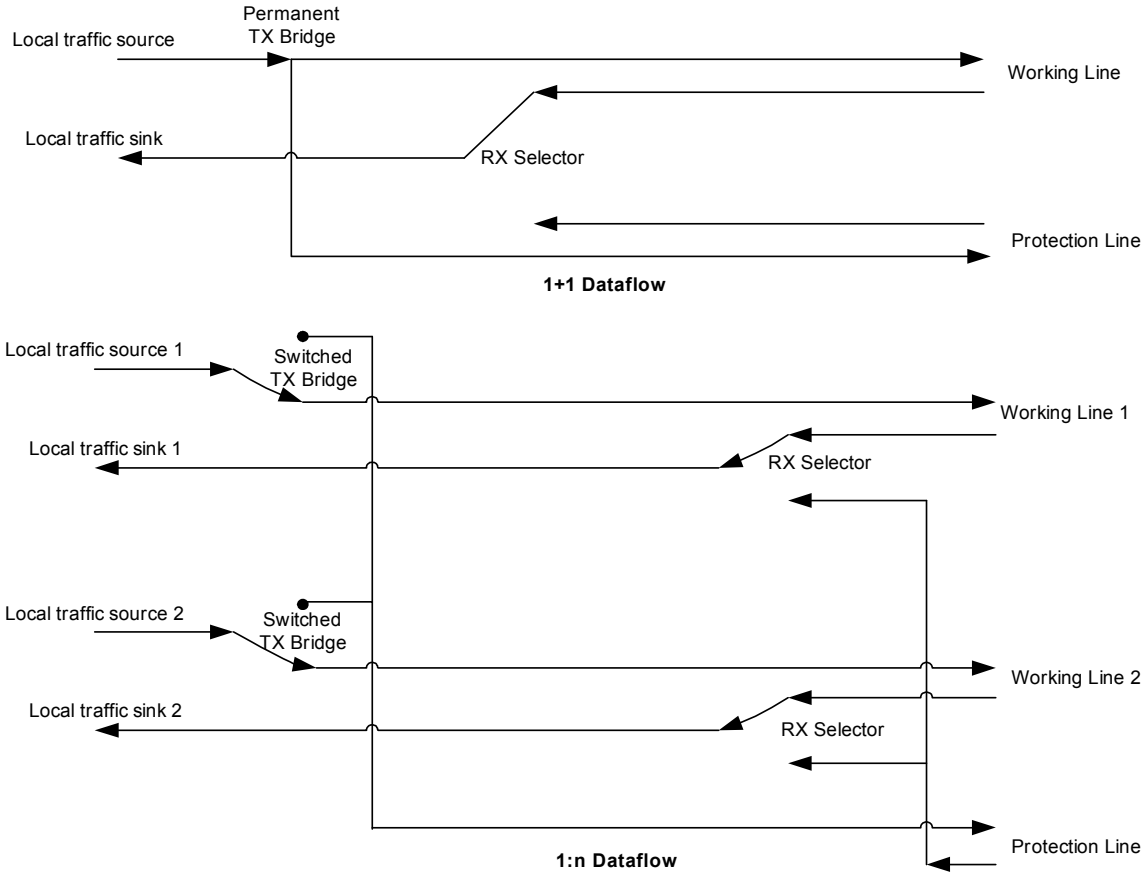


Figure 9-2 APS Dataflow (protection not switched in)

9.3 The Target Application

This chapter addresses **only** SONET end system equipment such as Routers, Media Gateways and Multi Service Switches. These are packet-based systems that typically forward traffic among a mix of non-SONET interfaces (e.g. Gigabit Ethernet) and SONET interfaces. They gain access to the SONET transport network via a directly connected SONET ADM. By contrast, SONET transport equipment deploy

more sophisticated configurations (such as protected SONET rings), support more extensive transport network functions and typically forward SONET frames between interfaces (requiring, in AdvancedTCA systems, the transmission of SONET frames over the AdvancedTCA fabric); this type of system is **explicitly outside the scope of this design**.

SONET end systems typically support application packet traffic with SONET I/O used simply for transporting this packet data. They are required to terminate each of the SONET layers. The SONET I/O card will always terminate the two lower layers (line and section) but there is an open question on where the SONET path is terminated. In traditional TDM systems, it is believed that the path is often terminated on a different card on the other side of the (TDM) bus. In the packet-based systems being addressed here, there is a distinct benefit in localizing all SONET processing (mapping, all overhead processing, framing etc.) to the dedicated SONET I/O boards; for example:

- The application boards are independent of the transmission network
- There is no requirement to transport SONET over packet-based fabrics

A sample architecture is shown in Figure 9-3.

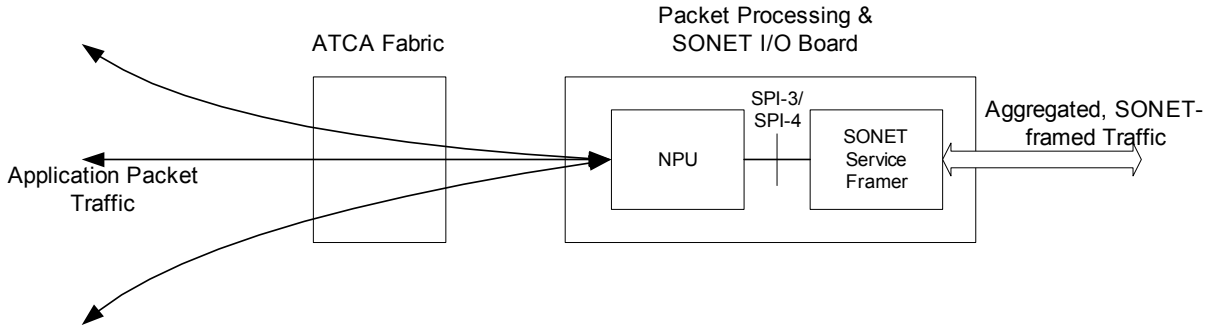


Figure 9-3 SONET End System Architecture

This generic example could support OC192/quad OC48 applications requiring a high end NPU and a SPI-4 packet interface; the same principles also apply to lower bandwidth applications where the NPU and Service framer would use the SPI-3 interface (with a gasket if necessary).

9.4 1+1 Linear APS Equipment Protection

A SONET end system addressed by this design connects directly to a SONET ADM. The required SONET protection functions include:

- **Linear APS:** This operates at the SONET line layer and protects the point-to-point link between the SONET end system and the ADM.
- **1+1 protection:** This is the most commonly deployed APS modes and in some ways is simpler (e.g. fixed bridging). Identical Path Overhead should be transmitted on both working and protection lines. Consequently, the permanent bridge must be located after the generation of the Path Overhead which effectively rules out the possibility of using fabric multicasting for the bridge.
- **Equipment protection:** Here the working and protection SONET interfaces are terminated on different boards consequently protecting against both line (external) and equipment (internal) failure. Note: Facility protection where both interfaces are terminated on the same board is not explicitly addressed because the required bridging is local to the board (and hence not a significant problem).

An example of the data path in 1+1 APS equipment protection is shown in Figure 9-4.

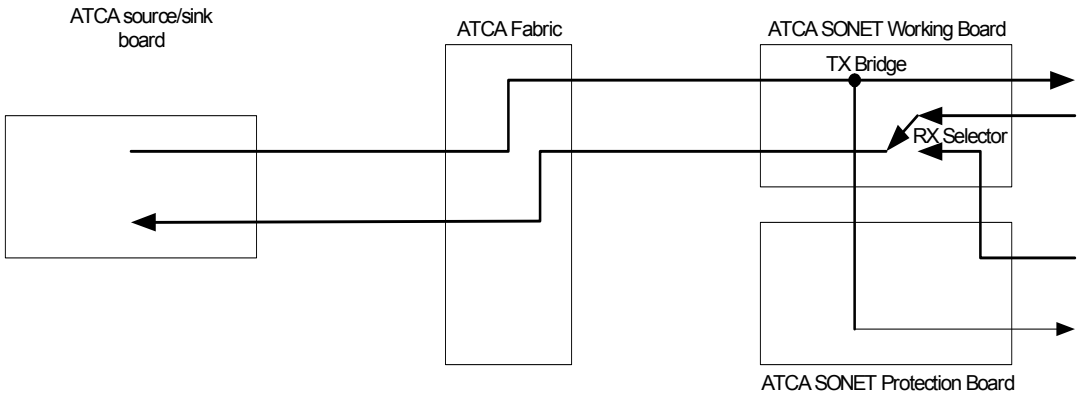


Figure 9-4 AdvancedTCA 1+1 Equipment Protection

The major logical functions required to support linear 1+1 APS are:

- **Permanent TX Bridge** continuously bridges the traffic for transmission to both the working and protection interfaces such that the Path Overhead and payload are transmitted identically on both lines. It is performed on one of the AdvancedTCA SONET Boards (normally working).
- **RX Selector** is provided by one of the AdvancedTCA SONET Boards selecting one of the two flows (based on the established local criteria) for forwarding through the fabric.

9.5 Design Scope

The high level design considers **only** the SONET datapath for 1+1 Linear APS equipment protection in SONET end systems (as described in previous sections); in particular it focuses on the required permanent bridging (i.e. data traffic replication) between AdvancedTCA boards. The following topics are outside the scope of the design:

- Control mechanisms used to support APS, such as state synchronization, board failure detection, ordered switches and so on
- Error indications to software applications, such as protection line failure, working board failure and so on
- Transport of application packet data (or SONET frames) through the AdvancedTCA fabric and delivery to the correct board
- 1:n APS since this has no special data path requirements (there is no bridging/stream replication) and switching (state synchronization etc.) is based on control mechanisms
- Advanced Mezzanine Cards will be addressed in a future phase

9.6 APS Dataflow, Interfaces and Clocking

Figure 9-5 shows example data flows and typical interfaces for 1+1 Linear APS equipment protection.

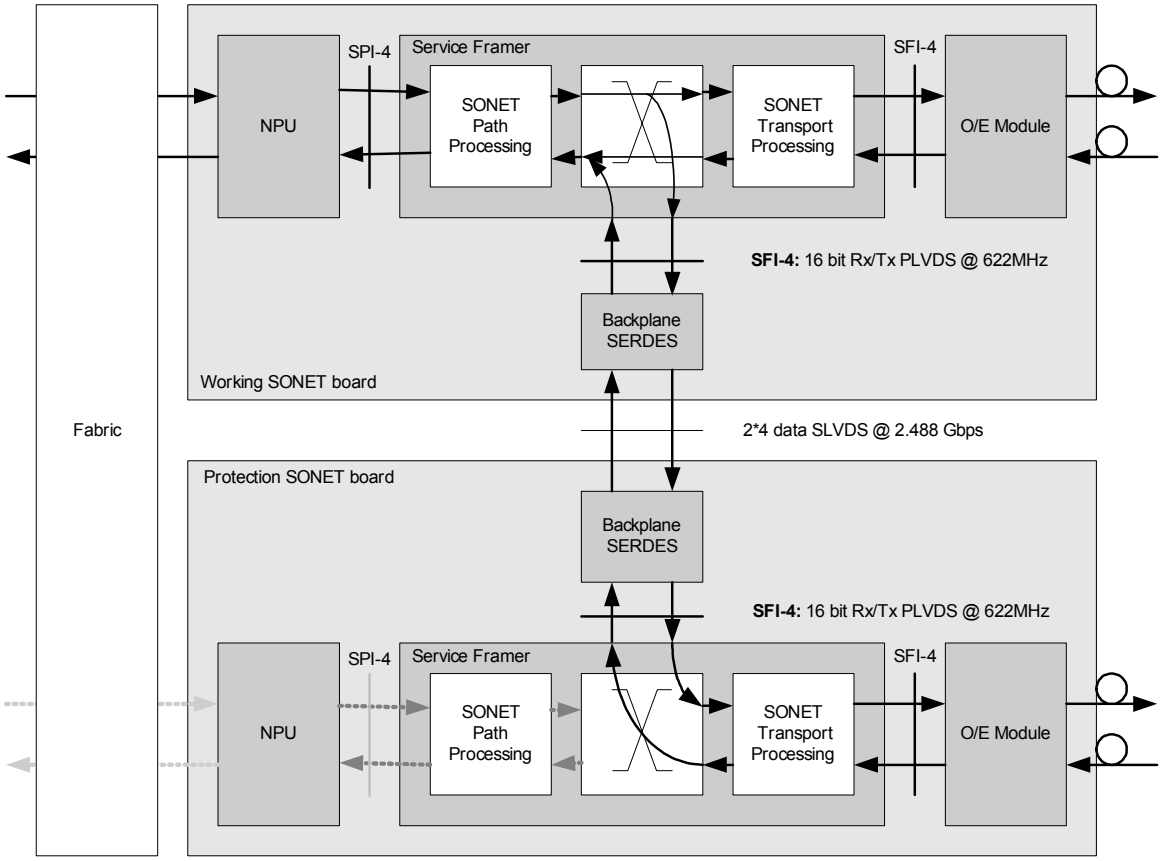


Figure 9-5 APS Dataflow

9.6.1 Data Flow

The required SONET data flows are:

- **Packet to SONET TX:**
    - In equipment protection, a mechanism must deliver the source traffic (over the fabric) to the correct board. Normally this would be the working board but must be switched to the protection board on failure/removal of the working board. It is an implementation option to use fabric multicasting for delivering the source traffic to both boards (thereby making the SONET board failures transparent to the source module) where it is ignored by the protection board. An alternate option is for APS software to control the dynamic switching of source traffic to the correct SONET AdvancedTCA board (e.g. switching to the protection SONET board on failure of the working SONET board).
    - The SONET Path Overhead and payload are generated on the working board and then transmitted towards the SONET Transport processing block on **both** the working and protection boards.
    - There is always an inter-board flow from ***TX working to TX protection***. This can be used to convey the K1, K2 APS signaling bytes and optionally used to carry control/status messages encoded in spare fields in the SONET Transport overhead. Note: This proprietary inter-board signaling could be controlled by a local processor via the framer's registers or possibly via a dedicated FPGA connected to the framer's TOH insertion port.
- **SONET RX to Packet:**
    - SONET traffic received from the working line undergoes Transport (line/section) processing and is then made available to the RX selector function.
    - SONET traffic received from the protection line undergoes Transport (line/section) processing and is then passed to the working board. One of the two incoming flows is selected (based on local criteria) for Path processing before being forwarded to the local NPU (and on to the fabric).
    - There is always an inter-board flow from ***RX protection to RX working*** allowing the working board to monitor the K1, K2 bytes carried on the protection line and optionally receive control/status messages encoded in the SONET Transport overhead.

9.6.2 Interfaces

TDM Fabric to Framer Interface (TFI-5) is the industry standard for 10G SONET signals across a backplane and many modern service framers support this standard. It was intended as a framer to switch interface but can be used as a framer to framer interface. The TFI-5 interface is physically carried over multiple 2.488 Gbps links with each link corresponding to the standard SONET STS-48 rate (or SDH STM-16). TFI-5 is a frame-based protocol (where the frame is a simplification of the SONET frame) and has three layers:

1. The Link layer provides framing, scrambling, line deskew, and line error monitoring
2. The Connection layer provides for supervision of errors and connectivity for each STS-1 timeslot. Note: The Connection later is not actually required in this application.
3. The Mapping layer coordinates the transport of data with a higher bandwidth than is supported by a single TFI-5 link (analogous to Inverse Multiplexing). An example is carrying OC192 over 4 separate 2.488 Gbps links.

TFI-5 is a serial protocol with three signals:

1. TFIDATA carries data between the two framers; the same signal is applicable in both directions.

2. TFIREFCK provides the timing reference for all the TFI-5 TFIDATA signals
3. TFI8KREF provides a reference to frame boundaries for all TFI-5 devices

Some service framers carry the inter-framer APS flows on the parallel, SFI-4 (SERDES Framer Interface) physical interface with 16 RX/TX data lines (and clocks) clocked at 622 MHz to support bidirectional OC192. This parallel interface (PLVDS) is connected to an external backplane SERDES to perform the necessary serialization (SLVDS) for transmission over the backplane TFI-5 interface. In some cases the TFI-5 layers are supported by the service framer itself so the backplane SERDES can be a very simple (non-SONET aware) device**.** Service framers that do not support the TFI-5 protocol require a more complex external backplane SERDES, such as to perform the TFI-5 link layer functions such as line deskew etc. Appropriate devices are available from Vitesse, AMCC and others.

9.6.3 SONET Clocking

There are various options for configuring SONET clocking in the AdvancedTCA system. One option is to use loop timing where the receiver recovers clock from the received signal; in quad OC48 operation, this requires SONET pointer processing to rate adapt between the independent lines clocks and the TFI system clock which may cause pointer movement within the SONET frame. This can be avoided by using a separate Stratum 3 clock source routed to both working and protection boards. Clock distribution using an external clock source (Stratum 3) is outlined in Figure 9-6. This shows that the external clock is supplied to both the line and backplane SERDES as the SFI-4 REFCLK which is then distributed throughout the two boards. The TFI clock signal (TFIREFCLK) and frame reference (TFI8KREF) are generated on the working board and supplied to the protection board.

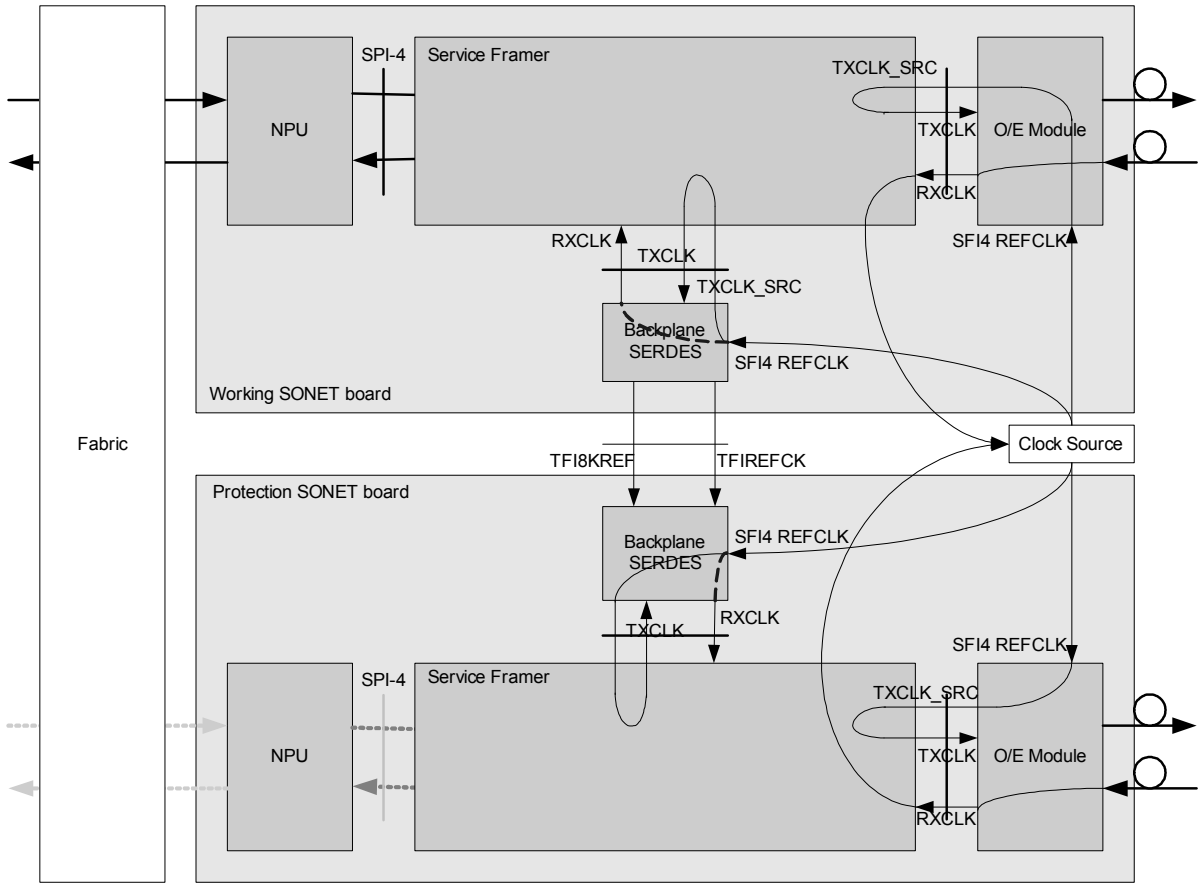


Figure 9-6 Clock Distribution

In a configuration where the clock source is duplicated on each board, the working board would be the master with the protection board being the slave; both clocks would be distributed between boards.

In all cases, clocking can be distributed between boards using the AdvancedTCA Synchronization Clock Interface.

9.7 AdvancedTCA APS Systems

9.7.1 The Update Channel Interface

In systems where the SONET I/O boards are in fixed, dedicated slots, the serialized, multi-link TFI-5 interface can be transported between the boards using the AdvancedTCA Update channel - a point-to-point interface formed from 10 differential signaling pairs normally routed over the system backplane between adjacent logical AdvancedTCA slots. (In dedicated slots, it is believed that systems with fixed working/protection SONET slots are normally acceptable.) Each signaling pair can support unidirectional 2.488 Gbps (OC48) so the 10 differential signaling pairs in the Update channel can be used as follows:

- Four of the pairs can be allocated to carry data from the working to protection boards supporting the ***TX working to TX protection*** flow at up to quad OC48 or OC192
- Four of the pairs can be allocated to carry data from the protection to working boards supporting the ***RX protection to RX working*** flow at up to quad OC48 or OC192
- One pair can be allocated for the TFIREFCK signal from the working to the protection board
- One pair can be allocated for the TFI8KREF signal from the working to the protection board (as an alternative, this can be recovered as part of frame delineation on one of the TFI-5 lanes)

The proposal to use the OC48 Update signaling pairs is independent of the APS granularity supported within the Service framers (e.g. OC192, OC48, OC12 etc.). However, some framers do not support the 16 * OC12 operation on both the APS and line interfaces.

9.7.2 Scenarios

Each scenario focuses on APS actions related to the data path. Consequently, although the application software is notified of each detected failure (typically this involves intervention by a local control processor on each board), and the corresponding alarm is raised as part of normal OAM, these actions are not explicitly noted.

9.7.3 Initialization

The Working and Protection cards are configured in slots connected by the AdvancedTCA Update channel. The system is initialized such that all SONET TX traffic is unicast to the Working card. The APS Permanent Bridging function is always located on the Working card with a corresponding permanent Working TX to Protection TX flow over the AdvancedTCA Update channel. The APS RX selector function is always located on the Working card with a corresponding permanent Protection RX to Working RX flow over the AdvancedTCA Update channel. The steady state data flows after initialization are shown in Figure 9-7.

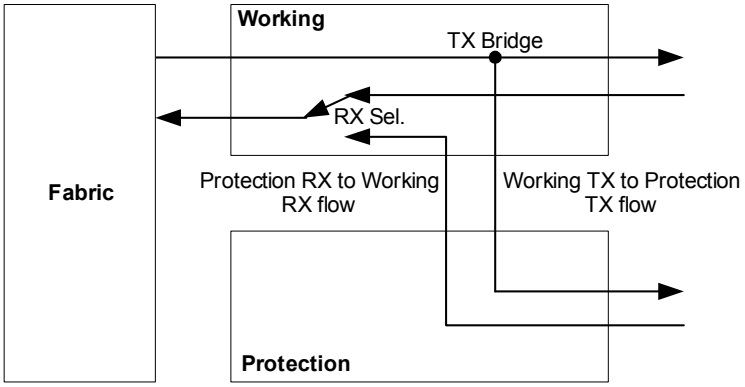


Figure 9-7 Steady State Data flows

For clarity, the figures show a single RX Selector appropriate for a single OC192 line. Service framers can typically switch at a much finer granularity such that a failure on one line has no effect on other lines.

9.7.4 Working Line Failure

Failure of the transmit line is detected by the remote peer which effects an APS switch. Failure of the receive line is detected by the local working board which effects an APS switch causing it to select the protection line. The K1, K2 bytes are then used to signal to the remote peer (and carried between the working and protection boards via the TX working to TX protection flow). The data flows after the local APS switch are shown in Figure 9-8.

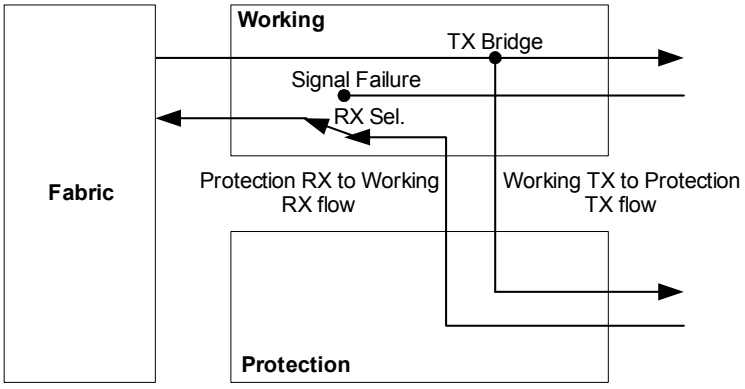


Figure 9-8 Working RX Line Failure

Under the default non-revertive case, the data flows remain unchanged after the working line is repaired but the Signal Failure is no longer detected on the Working card. An explicit switch request is required to revert to the flows shown in Figure 9-7.

9.7.5 Update Channel Failure

A failure occurring anywhere on the Working TX to Protection TX path such as, the Backplane SERDES, AdvancedTCA connectors, backplane trace, is detected by the local protection board as a Signal Failure

as shown in Figure 9-9. No data is transmitted on the protection line and that is detected by the remote peer which signals this failure via the K1/K2 bytes. There is no local APS action. The APS feature is no longer supported in the TX direction.

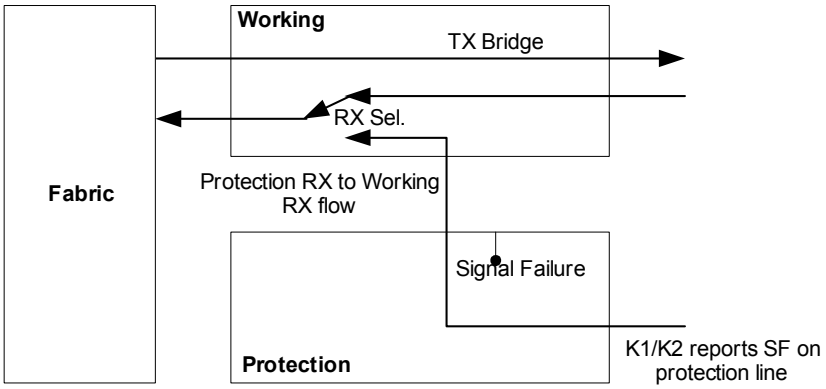


Figure 9-9 Update Channel Failure (TX Direction)

A failure occurring anywhere on the Protection RX to Working RX path, such as the Backplane SERDES, AdvancedTCA connector, backplane trace, is detected by the local working board as shown in Figure 9-10. The APS feature is no longer supported in the RX direction; subsequent failure of the Working RX line will result in loss of service.

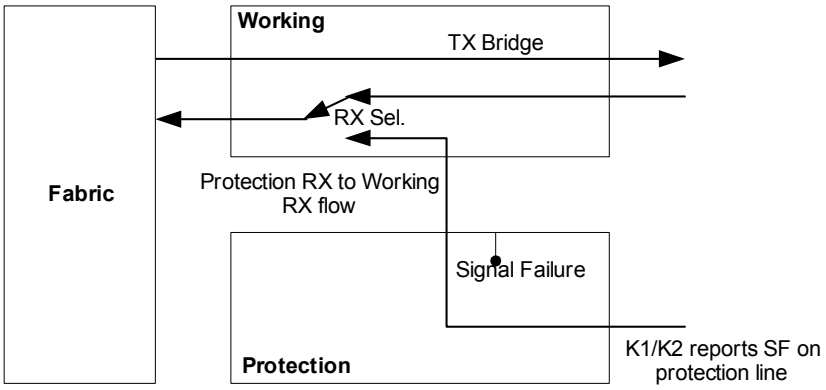


Figure 9-10 Update Channel Failure (RX Direction)

9.7.6 Protection Line Failure

Failure of the transmit line is detected by the remote peer but no remote APS switch occurs. Failure of the receive line is detected by the local protection board (and indirectly by the local working board) but no local APS switch occurs.

9.7.7 Card Failure

Failure of either card results in the loss of the APS service; the remaining board is ordered to enter the non-APS (straight through) mode. Figure 9-11 depicts an example of a working card failing.

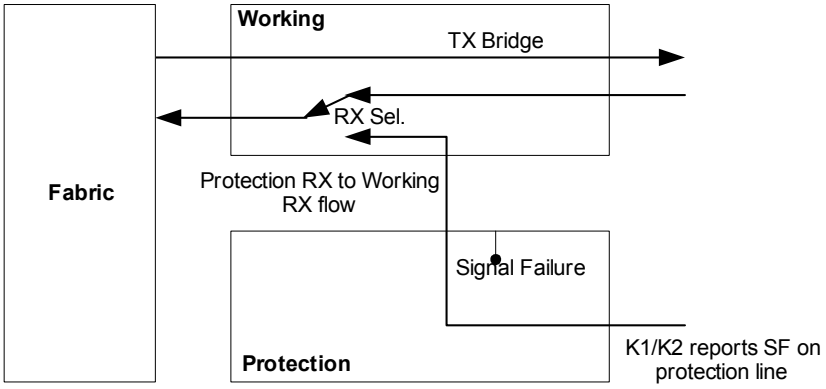


Figure 9-11 Working Card Failure

Depending on the exact failure mode, typical indications include:

- Signal failure is detected by the protection card on the TX Update Channel
- No data is transmitted on either line causing the remote peer to report Signal Failure of both lines via the K1/K2 bytes
- No data enters the fabric from the working board

The exact criteria for detecting board failure are outside the scope of this guide. On detecting failure of the Working card, the control plane:

- Orders the Protection board to enter the non-APS mode (receiving TX data from the fabric, transmitting RX data to the fabric)
- Instructs the source cards to redirect their TX data to the Protection card
- Optionally disables Fabric access by the Working card

The resulting flows are shown in Figure 9-12.

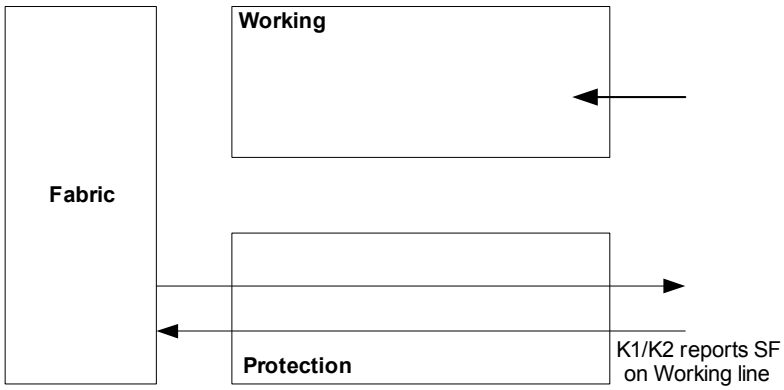


Figure 9-12 APS Switch to Protection Card

On repair of the Working card, both cards are instructed to enter APS mode and the source cards ordered to switch TX traffic to the Working card. In a non-revertive mode, the Working card will still select the Protection line as shown in Figure 9-13.

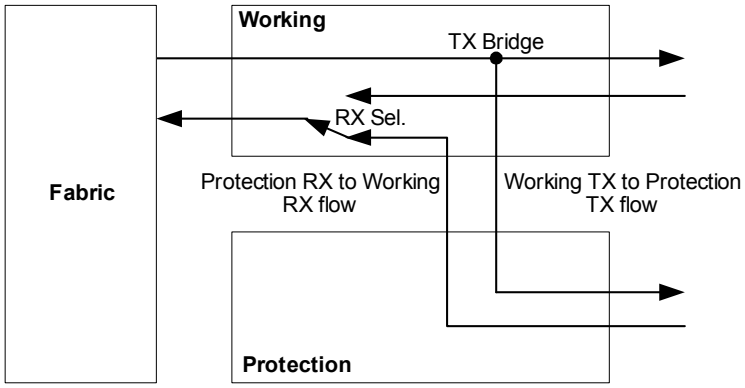


Figure 9-13 Working Card Repair

9.8 Open Questions

The scenarios described above illustrate one proposed design and raise several questions for further consideration:

1. Is it acceptable to have a non-redundant Update channel? This approach results in the loss of the APS service (see Section 9.7.5) on a single failure on the backplane. Sufficient Update Channel lanes exist to support 1+1 redundancy for OC48 applications but not OC192. To address OC192 redundancy, an alternative could be developed as follows:
    a. Distribute all clocking signals via the AdvancedTCA Synchronization Clock interface thereby freeing up 2 of the 10 Update Channel lanes

b. Group the Update Channel interface into 2 groups of 5 lanes (one group for each direction)

c. Support 1:4 redundancy within each OC192 group

2. Is it optimal for the source cards to unicast their TX traffic and therefore be compelled to switch destinations on RX card failure? Alternative approaches that could be considered are:

    a. Multicast source TX traffic to both Working and Protection cards where it is discarded by the protection card unless the working card has failed. This means that SONET card failures are transparent to the source cards although there may be QoS implications associated with the multicast mechanism.

    b. Dual unicast TX traffic to both Working and Protection cards where it is discarded by the protection card unless the working card has failed. This provides APS transparency (as above) at the expense of fabric usage.

3. Is it preferable, in some designs, to route the *TX Working to TX Protection* and *RX Protection to RX Working* flows via the fabric? This eliminates use of the Update Channel at the expense of fabric usage.

This page is intentionally left blank.

10. Carrier Grade OS

A Carrier Grade Operating System (CG OS) is a key component of a highly available MCP system architecture. In addition to providing a robust software platform for high availability middleware and applications, a CG OS includes specific features that enhance the availability and serviceability of the MCP system.

ID	Feature Name	Requirements
OS- 1	MCP based Carrier Grade Servers ***should*** host a Carrier Grade Operating System.	A detailed list of CG OS features and their description found in **Appendix B**.

This page is intentionally left blank.

11. Certification and Practices

The AdvancedTCA Platform is intended to be designed, at a minimum, to meet the following certifications. Depending on the equipment type and customer expectations, there may be additional required certifications required.

CAUTION: As noted in Chapter 1 of the PICMG* 3.0 specification, AdvancedTCA* systems are designed to be installed and serviced by trained service personnel only, not by equipment operators. The primary reason for this is that voltages over 60VDC that can be present in AdvancedTCA systems.

11.1 System Power

ID	Feature Name	Requirements
CP-1	System Power	When the DC power source (plant power) is -48V, the system power ***shall*** meet the requirements of the PICMG 3.0 specification for a -48V two- or three wire system. In addition it ***shall*** meet the requirements of the following: • ETS 300 132-2 • ANSI T1.315
CP-2	System Power	When the DC power source (plant power) is -60V, the system power ***shall*** meet the requirements of the PICMG 3.3 specification for a -60 V two- or three wire system. In addition, it ***shall*** meet the requirements of the following: ETS 300 132-2

11.2 System Grounding

ID	Feature Name	Requirements
CP-3	System grounding	System grounding ***shall*** be of a multi-point type and meet the requirements of the following: • ETS 300 253 • IEC60950 • UL60950 • GR-1089-CORE

11.3 Climatic Requirements

Storage Temperature and Humidity

ID	Feature Name	Requirements
CP-4	Climatic Requirements	Hardware ***shall*** meet the requirements of the following: • ETSI EN300 019-2-1 class 1.2 • NEBS GR-63-CORE (2002) Sections 4.1and 5.1

ID	Feature Name	Requirements
		• IEC 60068-2-1, IEC 60068-2-2, IEC 60068-2-30 and IEC 60068-2-56

11.3.2 Transport Temperature and Humidity

ID	Feature Name	Requirements
CP-5	Climatic Requirements	Hardware ***shall*** meet the requirements of the following: • ETSI EN300 019-2-2 class 2.3 • NEBS GR-63-CORE (2002) sections 4.1 and 5.1 • IEC 60068-2-1, IEC 60068-2-2, IEC 60068-2-14, IEC 60068-2-30 and IEC 60068-2-56

11.3.3 Operation Temperature and Humidity

ID	Feature Name	Requirements
CP-6	Climatic Requirements	Hardware ***shall*** meet the requirements of the following: • ETSI EN300 019-2-2 class 3.1E • NEBS GR-63-CORE (2002) section 4.1 and 5.1 • IEC 60068-2-1, IEC 60068-2-2, IEC 60068-2-14 and IEC 60068-2-56 NEBS GR-3028 Section 5.1

11.3 Dynamics Requirements

11.4.1 Transport

ID	Feature Name	Requirements
CP-7	Dynamics	Hardware ***shall*** meet the requirements of the following: • NEBS GR-63-CORE (2002) Section 4.4.4 and 5.4.3 • IEC 60068-2-29, IEC 60068-2-32, IEC 60068-2-64 • ETSI EN300 019-2-2 class T2.3

11.4.2 Storage

ID	Feature Name	Requirements
CP-8	Dynamics	Hardware ***shall*** meet the requirements of the following: • NEBS GR-63-CORE (2002) Sections 4.3.1 and 5.4.3 • IEC 60068-2-6, IEC 60068-2-64 • ETSI EN300 019-2-1 class T1.2

11.4.3 Operation

ID	Feature Name	Requirements
CP-9	Dynamics	Hardware ***shall*** meet earthquake standards per: • NEBS GR-063-CORE (2002) Zone 4, sections 4.4.1 and 5.4.1 • IEC 60068-2-6 and IEC 60068-2-64 • ETSI EN300 019-2-3 class 3.1
CP-10	Dynamics	Hardware ***shall*** meet the requirements of the following: • NEBS GR-63-CORE (2002) sections 4.4.3 and 5.4.2 • IEC 60068-2-27 • ETSI EN300 019-2-3, class 3.1

11.4 Electromagnetic Interference (Emissions)

ID	Feature Name	Requirements
CP-11	Electromagnetic Interference	Equipment ***shall*** meet the requirements of the following: • FCC Part 15 Class A • GR-1089-CORE • EN 300 386 Class A • CISPR 22 Class A • EN55022 Class A • ICES 003 Class A
CP-12	Electromagnetic Interference	The equipment ***should*** also meet the requirements of the following: • FCC Part 15 Class B • EN 300 386 Class B • CISPR 22 Class B • EN55022 Class B • ICES 003 Class B

11.5 Electromagnetic Compatibility (Immunity)

ID	Feature Name	Requirements
CP-13	Electromagnetic Compatibility	Electromagnetic shielding ***shall*** meet the requirements of the following: • NEBS GR-1089-CORE • EN55024 • CISPR24 • EN 300 386

Note: Certain factors aid in determining noise susceptibility, such as if the hardware is installed in a rack, if it has doors, if the doors are closed, etc.

11.6 Acoustics

ID	Feature Name	Requirements
CP-14	Acoustics	Equipment ***shall*** meet the requirements of the following: • NEBS GR-1089-CORE sections 4.6 and 5.6 • ETS 300 753

11.7 Transients and Electrical Noise

ID	Feature Name	Requirements
CP-15	Transients and Electrical Noise	Equipment ***shall*** meet the requirements of the following: • NEBS GR-1089-CORE (2002) section 7.5.3 • ETSI EN300 132-2 • EN 300 386

11.8 Safety

ID	Feature Name	Requirements
CP-16	Safety	Hardware ***shall*** meet the safety requirements specified in the following: • EN 60950 • UL 60950 • NEBS GR-1089-CORE • IEC 60950

11.9 Flammability

ID	Feature Name	Requirements
CP-17	Flammability	Hardware ***shall*** be fire resistant according to the following: • GR-63-CORE (2002), sections 4.2 and 5.2 • UL60950 • EN60950 • IEC60950

11.10 Reliability

ID	Feature Name	Requirements
CP-18	Reliability	The reliability of the hardware ***shall*** be specified in MTBF per one of the following international standards, and the method used ***shall*** be specified: • TR-332 • MIL-HDBK-217 • HRD5

ID	Feature Name	Requirements
		• RFD93 CNET • SN29500

11.11 Altitude

ID	Feature Name	Requirements
CP-19	Altitude	Hardware ***shall*** meet the altitude requirements of the following: • NEBS GR-63-CORE (2002) sections 4.1.3 and 5.1.3

11.12 Airborne Contaminants

ID	Feature Name	Requirements
CP-20	Airborne contaminants	Hardware ***shall*** support uninterrupted service in the presence of airborne contaminants according to the requirements of the following: • NEBS GR-63-CORE (2002) sections 4.5 and 5.5 • IEC 60068-2-60 • IEC 60068-2-68

This page is intentionally left blank.

12 Design Practices

This chapter describes the common design practices to be followed when building AdvancedTCA components.

12.1 Over Voltage Transients

Voltage transients are short duration surges of electrical energy. These surges (transients) are random in nature and can cause damage to the electrical circuits or even personal injury if not handled properly. Industry estimates report that up to 70% of field failures of electronic equipment are due to electrical overstress conditions. The AdvancedTCA platform needs to address these transients to ensure that these threats do not disrupt the operation of any of the I/O boards in the AdvancedTCA platform.

Examples of voltage transients are lightning strikes, Electro Static Discharge (ESD), inductive load switching, or a power line crossing. Numerous electronics standards groups have analyzed transient voltage occurrences using accepted monitoring and testing methods. The characteristics of several over voltage transients are shown in Table 12-1. Most countries require telecommunications circuits to be tested and certified to meet specific standards such as Telcordia GR-1089, TIA/EIA-18-968 (was FCC Part 68) and ITU K.20, .21 and .45.

Table 12-1 Example of Transients and Magnitudes

	Voltage	Current	Rise Time	Duration
Lightning	25kV	20kA	10us	1ms
Switching	600V	500A	50us	500ms
EMP	1kV	10A	20ns	1ms
ESD	15kV	30A	1-5ns	100ns

12.1.1 Power Cross Transients

Power cross transients occur when an AC power line comes in contact with or is located parallel to a telecom phone line. The induced AC voltage will couple to the telecom line causing over voltage transients. For example, the GR - 1089 first level AC Fault Test requires that the equipment continue to operate after the AC transient. AdvancedTCA platform needs to address these transients to assure that high voltage and current surges do not disrupt the operation of any of the AdvancedTCA boards in the AdvancedTCA platform.

12.1.1.1 Transient Protection Devices

Devices used to protect against transients are Gas Discharge Tubes (GDTs), Metallic Oxide Varistors (MOVs), Ceramic and Polymer based Varistors, Transient Voltage Suppressors (TVS) diodes, Polymer ESD suppressors, and Thyristors. These devices are typically applied between the Vsignal and Vground configurations. Figure 12-1 shows a block diagram for typical protection circuit configurations.

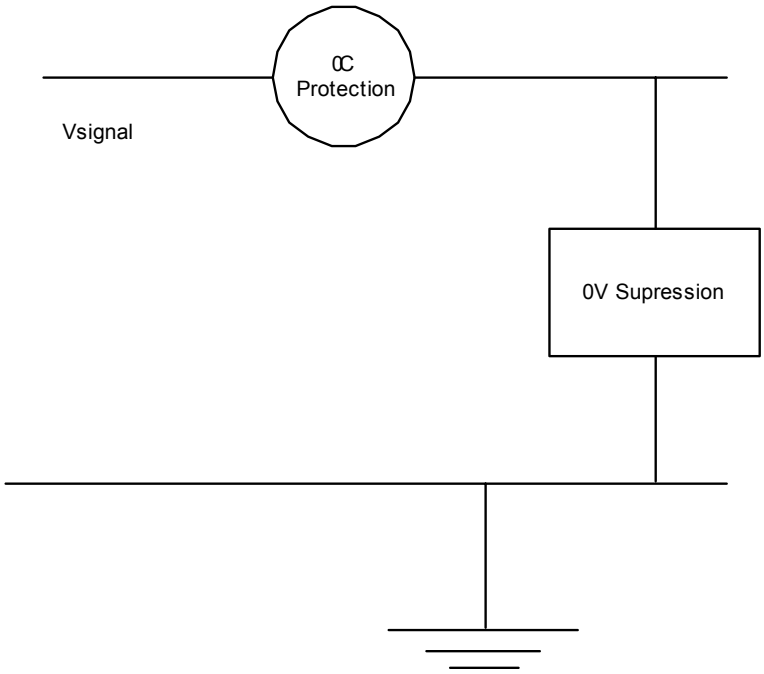


Figure 12-1 Typical Circuit Protection Configurations

Tip and ring pairs must be protected from both lightning and power cross transients. These circuits are used in a variety of applications and require different circuit configurations for each application. Figure 12-2 shows a typical telecom circuit protection configuration using several transient protection devices.

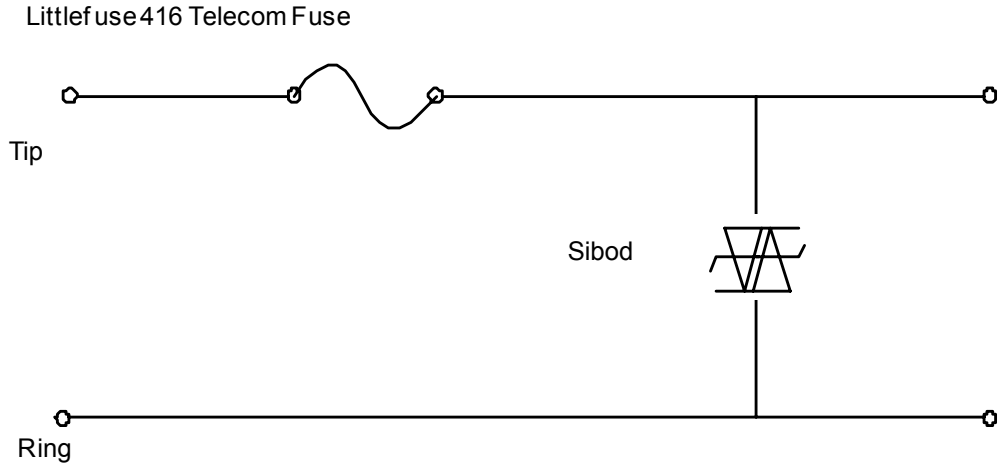


Figure 12-2 Typical Protection for Telecom Circuits

12.1.2 Electro Static Discharge (ESD) Design Considerations

ESD transients are caused when two objects with different voltage potentials come in contact with or close proximity to each other. Very fast rise times and very high peak voltages and currents characterize ESD transients. Due to the hot swap functionality and high speed I/O, the AdvancedTCA platform needs to address ESD transients on all I/O lines to assure reliable operation of any of the AdvancedTCA boards in the TCA platform. Most countries use a version of the IEC 61000-4-2 standard for testing electronic

equipment for ESD susceptibility up to level 4, which is 8 kV direct contact and 15 kV air discharge. Figure 12-3 shows the IEC 61000-4 Human Body Model (HBM) pulse, used for ESD testing.

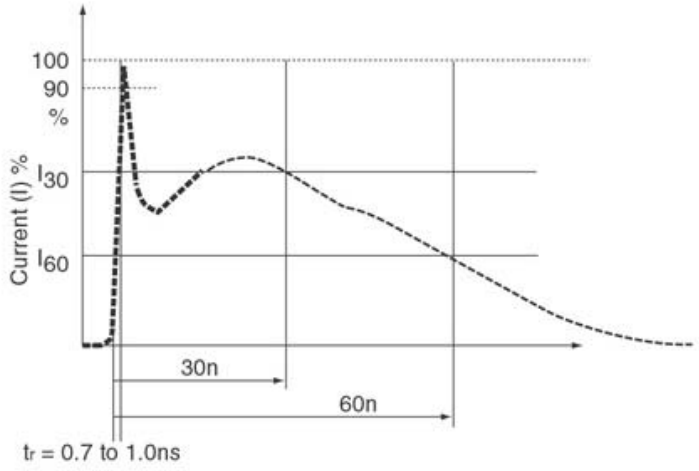


Figure 12-3 IEC 61000 - 4 Waveform

ID	Feature Name	Requirements
DP-1	Over Voltage Transients	Designers ***should*** include ESD suppressors on all I/O signal lines that could be exposed to an air discharge or direct contact ESD transient. In particular, USB 2.0, Ethernet, input/ output devices and serial communications lines ***should*** be protected.

12.1.2.1 Copper Ethernet Port

This port is common to 10/100/1000 Base T connections and uses the RJ 45 connector as a standard interface. Circuit protection is required to meet specific industry standards for over voltage protection and over current protection.

ID	Feature Name	Requirements
DP-2	Over Voltage Transients	RJ 45 connectors with common mode chokes (parallel and series for both TX and RX lines) and with high voltage blocking capacitors ***should*** be used. For example, the SI- 50132 from InNet technologies™ has been found to meet these requirements adequately.
DP-3	Over Voltage Transients	Low capacitance Thyristors on each transmit and receive pair ***may*** be required to meet industry standards such as Telcordia GR-1089, TIA/EIA-18-968 (was FCC Part 68), and ITU K.20, .21 and .45. Telecom rated fuses ***may*** also be required to meet specific surge tests for certain countries.
DP-4	Over Voltage Transients	When surge tests are not required, the transmit and receive pairs ***should*** be protected by ESD protection devices and high voltage blocking capacitors as shown in Figure 12-4.

Low-capcitance capcitors, Littelfuse™ GDTs, Sibod Thyristors, and telecom rated fuses (461) are designed for inter building applications. Figure 12-4 shows a basic telecom protection circuit using

Littelfuse GDTs, Sibod Thyristors, and telecom 461 fuses.

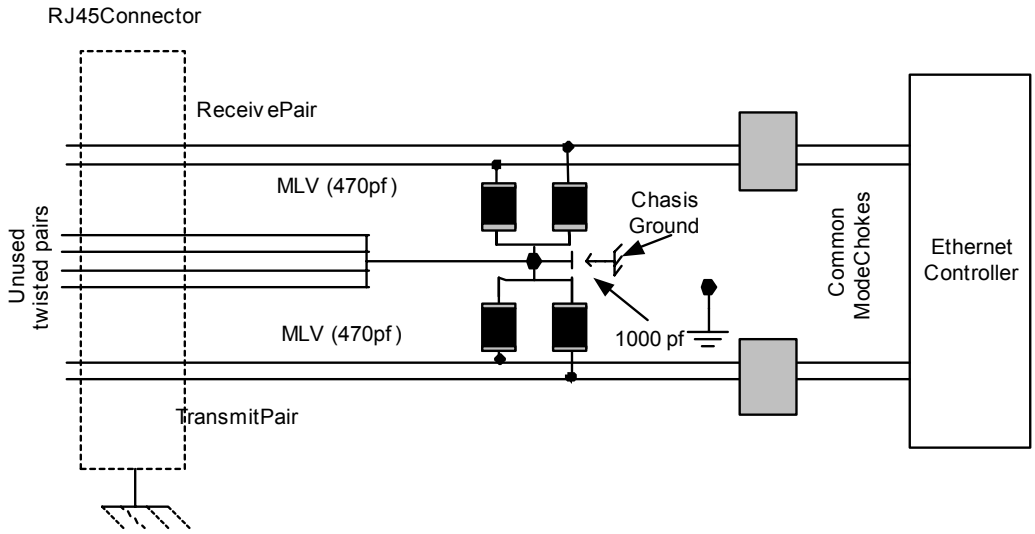


Figure 12-4 Typical Intrabuilding Ethernet Circuit Protection Configuration

ID	Feature Name	Requirements
DP-5	Over Voltage Transients	The shields of the Ethernet connectors ***should*** be connected to chassis ground.
DP-6	Over Voltage Transients	Differential impedance ***should*** be maintained in a 100 ohm differential application.
DP-7	Over Voltage Transients	When routing differential pairs, identical trace lengths ***should*** be maintained. Differences in trace lengths translate directly to signal skew.
DP-8	Over Voltage Transients	Differential pair traces ***should*** be maintained on the same side of the PCB to minimize discontinuities.
DP-9	Over Voltage Transients	A separation of 40 mils ***should*** be maintained between RX and TX pairs to minimize crosstalk.
DP-10	Over Voltage Transients	When a board or system has large numbers of Ethernet cables, then shielded cables ***should*** be used to pass class B limits.

12.1.3 Air Discharge

Air discharges can occur in both lightning and ESD transients. Specific circuit protection devices handle both types of transients.

Lightning transients occur to wiring that is outside a building. The most susceptible are power and telephone lines wiring, either overhead or buried.

All wiring that enters a building must utilize circuit protection devices to prevent damage to electronic equipment connected to these wires. The most common circuit protection devices for this application are

Gas Discharge Tubes (GDT), Metallic Oxide Varistors (MOV), and high voltage Transient Suppression Diodes (TVS). Fuses are also used in conjunction with these devices to limit current transients.

Air discharge ESD transients can cause significant damage to electronic equipment due to the high voltage, high current characteristics. The IEC 61000-4-2 standard addresses this transient.

ID	Feature Name	Requirements
DP-11	Over Voltage Transients	Designers ***should*** include ESD suppressors on all signal lines that could be exposed to an air discharge ESD transient. In particular, USB 2.0, Ethernet, and serial communications lines ***should*** be protected.

12.1.4 Contact Discharge

Contact discharges are the major source for ESD transients. The human body model (HBM) is utilized in the IEC 61000-4-2 specification for testing electronic equipment against direct contact ESD transients. The hot swap feature of the AdvancedTCA Chassis provides an opportunity for ESD transients.

ID	Feature Name	Requirements
DP-12	Over Voltage Transients	Due to the high voltage, high current characteristics, designers ***should*** include ESD suppressors on all signal lines that could be exposed to a human body.
DP-13	Over Voltage Transients	Back plane I/O, USB 2.0, Ethernet and serial communications lines ***should*** be protected.

Note: *The ESD discharge system of AdvancedTCA was not designed to handle the full body potential for hot-inserting AdvancedTCA boards and that properly maintained static discharge wrist straps are mandatory for integration and service personnel. The residual ESD on the person handling the board must be limited to 1 kV or the discharge resistors will arc-over and potentially upset the system.*

12.2 Typical Telecom Protection Circuit for Primary and Secondary Protection

The section describes EMC grounding guidelines for plug-in boards/modules and the backplanes. Safety grounding requirements are covered in another section.

12.2.1 Grounding Guidelines for Boards/Modules

12.2.1.1 PCB Grounding Guidelines

ID	Feature Name	Requirements
DP-14	Telecom Protection	Designs ***should*** incorporate power and ground planes of at least 1 oz. copper.
DP-15	Telecom Protection	In designs that require a cut power plane to isolate different voltages, there ***should*** be the following: • Additional adjacent GND layers with 4 mils separation to each other, for example, use the power GND planes as an effective power supply bypass capacitor. This capacitance is more effective than a capacitor of equivalent value because the planes have no inductance or Equivalent Series Resistance (ESR). OR • Stitching high frequency capacitors between isolated parts of power plane.

ID	Feature Name	Requirements
DP-16	Telecom Protection	Copper layers connected to the chassis ground ***shall*** be placed on the top and bottom layers at the front of the boards. Both layers ***shall*** be connected together by vias and also connected to every shield of front panel connectors (compute boards - video, serial, Ethernet, USB, mouse / KB connector) and board mounting holes.
DP-17	Telecom Protection	Logic Ground (Ground plane) ***shall*** be separated from the Chassis Ground but must have the ability to be connected together using capacitors located close to every connector. The value of the capacitors will be determined during the design phase. The caps should be 603 size and 100VDC rated. The decision on whether to place the capacitors will depend on EMC and ESD test results.
DP-18	Telecom Protection	Provision ***shall*** be made to optionally connect Logic Ground to Chassis Ground using a jumper. This jumper is for installation by an OEM or NEP. Such a connection is helpful for an inter-Shelf interface of single ended signals. It is expected that such a jumper will not be required to support single ended signals if the proper high frequency bypass is implemented on boards as described earlier. In case such a jumper is required, a decision on whether to ship a board with the jumper already in place will be taken at a later date based on customer feedback. There is no plan to make this option software programmable as it introduces more complexity and more parts.
DP-19	Telecom Protection	The shield of the connectors on the back of the boards (backplane connectors) ***should*** be connected to Logic Ground.
DP-20	Telecom Protection	For telecom or other metallic I/O cards, ground and voltage planes ***should*** be eliminated from under or near the I/O connectors to minimize the coupling of EMI onto the I/O signals. If DC ground connections are needed on the connector, they should be of minimum trace width and at least 1" long before connecting to the ground plane.
DP-21	Telecom Protection	Filter capacitors ***shall*** be connected between the Chassis Ground and -48V planes immediately after the onboard fuses and close to the backplane power connector. The 48V to chassis ground isolation HiPot requirement is 1414VDC. Capacitors **shall** be rated to withstand a HiPot test at 1414VDC.

Note: Do not cut up the power or ground planes in an effort to steer current path. This usually produces more noise, not less.

12.2.1.2 Clock Generator Design Guidelines

In designing the clock generator, the following are guidelines to be considered:

- Should have the ability to use Spread Spectrum modulation for EMI reduction. This technique is quite useful for achieving class B certification.
- Shall isolate the chip's power supply from the rest of the circuit through the use of a ferrite bead.
- Shall place on-chip series damping resistor at clock output drivers for low noise and EMI reduction.
- For high speed traces, reduce if not eliminate vias to minimize impedance discontinuities.
- Keep traces as short as possible and use rounded or 45-degree corners.

- Eliminate/reduce stub lengths.
- Bypass capacitors should be located on the same side of the PCB as the clock generator and must be located as close as possible to power and ground pins of the chip. Furthermore, it is recommended that the capacitors be located between the pins and the vias to the planes. The value of the bypass capacitors will be determined during the design phase.

12.2.1.3 USB Port (1.1 and 2.0)

ID	Feature Name	Requirements
DP-22	Telecom Protection	USB interfaces are high-speed serial interfaces and ***should*** be protected against ESD transients. The plug and play nature of the USB interface provides the opportunity for inducing ESD transients into the AdvancedTCA chassis.
DP-23	Telecom Protection	USB 2.0 is very sensitive to ESD and board designers ***should*** include ESD protection on the data lines and the Vbus line ***should*** be protected against over current events from a down steam USB peripheral. A Positive Thermal Coefficient (PTC) ***may*** be used.
DP-24	Telecom Protection	The shield of the USB connector ***should*** be connected to the chassis ground.
DP-25	Telecom Protection	USB 1.1 lines ***should*** be protected against EMI and ESD. The SPUSB1 series ESD suppressor from Littelfuse is suitable for this purpose and provides ESD protection, EMI filtering, and data line termination. This device ***should*** be placed as close to the USB connector as possible. The Littelfuse SPUSB1 device is suitable for this purpose by providing ESD, EMI, and termination in one device. Figure 12-5 shows the SPUSB1 device in the USB circuit.

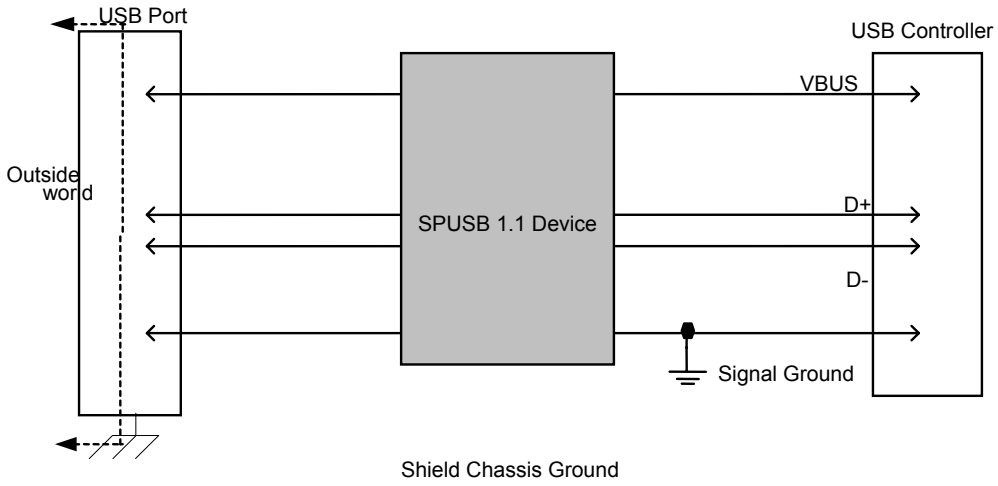


Figure 12-5 SPUSB1 Protection Circuit for USB 1.1 Interfaces

Figure 12-6 shows the alternate circuit protection configuration for USB1.1 circuits.

ID	Feature Name	Requirements
DP-26	Telecom Protection	Alternately, the USB 1.1 lines port ***should*** be protected against ESD and EMI by placing a 120 Ohm ferrite bead in series and low capacitance (47pf) ESD suppressor connected to logic ground. Littelfuse Multi-layer Varistors (MLVs) are suitable for this purpose.

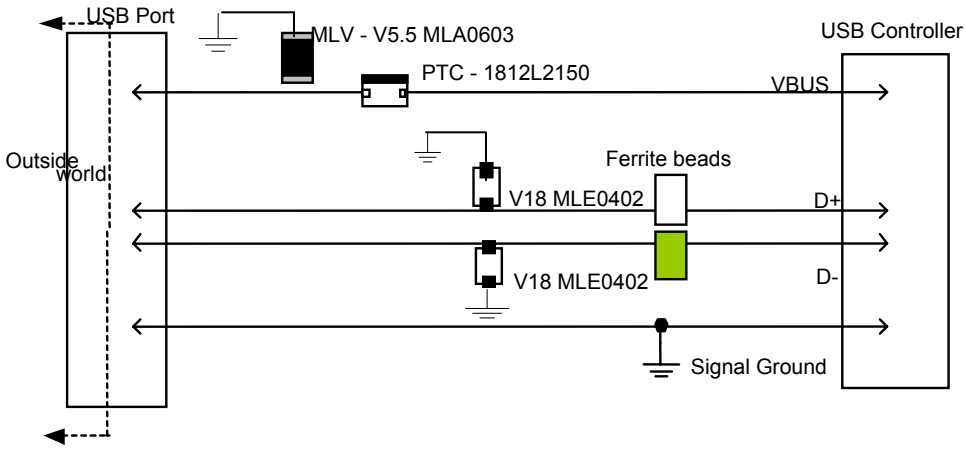


Figure 12-6 Alternate Protection Circuit for USB 1.1 Interfaces

The Pulseguard series ESD suppressors from Littelfuse are suitable for this purpose as outlined in the Intel High Speed USB Platform Design Guide, revision 1.0, section 5. Figure 12-7 shows a proper USB 2.0 circuit protection configuration. MLV devices can be added as shown to improve performance against ESD transients on the USB connector shield.

ID	Feature Name	Requirements
DP-27	Telecom Protection	USB 2.0 lines ***should*** be protected against ESD. Due to the high-speed nature of the USB 2.0 interface, high-speed ESD suppressors are required to meet signal integrity requirements.

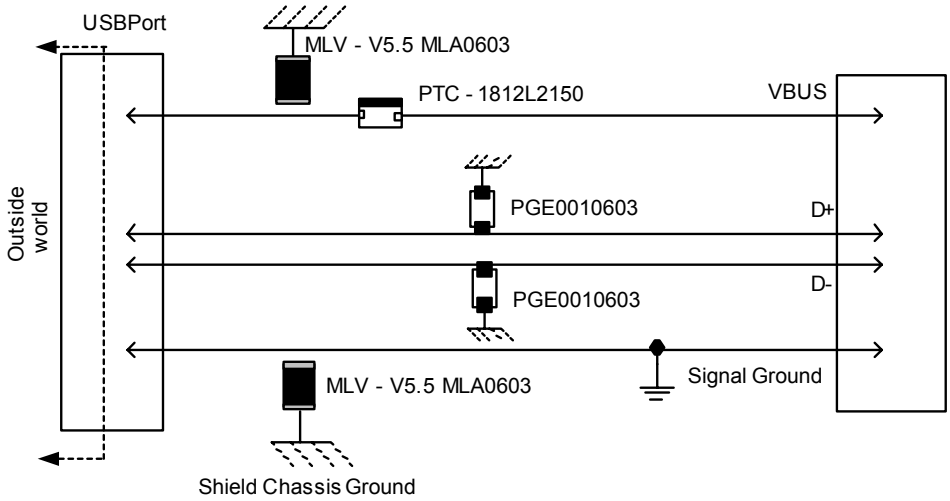


Figure 12-7 USB 2.0 Circuit Protection Configuration

12.2.1.4 Keyboard/Mouse Port

Telecom equipment operators and service personnel utilize these interfaces to connect to the telecom equipment and the opportunity exists for inducing ESD transients into the AdvancedTCA chassis through the keyboard or mouse.

ID	Feature Name	Requirements
DP-28	Telecom Protection	The shield of the keyboard or mouse connector ***should*** be connected to the chassis ground.
DP-29	Telecom Protection	The signal lines ***should*** be protected against ESD and EMI by placing a 120 Ohm ferrite bead in series and low capacitance (470pf) ESD suppressor connected to logic ground.

Littelfuse Multi-layer Varistors (MLVs) are suitable for this purpose as shown in Figure 12-8.

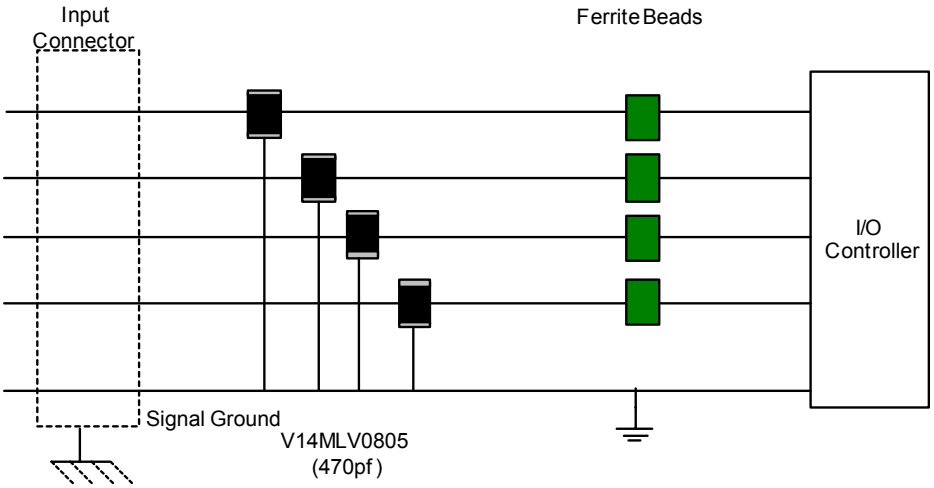


Figure 12-8 EMI/ESD Circuit for USB 1.1 and Keyboard/Mouse Circuits

12.2.1.5 RJ 45 Serial Port

This port is common for basic serial connections and subject to ESD transients. This interface provides the opportunity for inducing ESD transients into the AdvancedTCA chassis.

ID	Feature Name	Requirements
DP-30	Telecom Protection	Data lines ***should*** be filtered and protected with low capacitance Multi-Layer Varistors (470pf) connected to the logic ground as shown in Figure 12-8. Littelfuse MLV devices are suitable for this purpose.
DP-31	Telecom Protection	The shield of the serial port connector ***should*** be connected to chassis ground.
DP-32	Telecom Protection	A shielded cable is required to meet Class B limits. Interbuilding RJ-45 circuits ***may*** have circuit protection on each tip/ring pair. The basic circuit for this type of protection is shown in Figure 12-9.

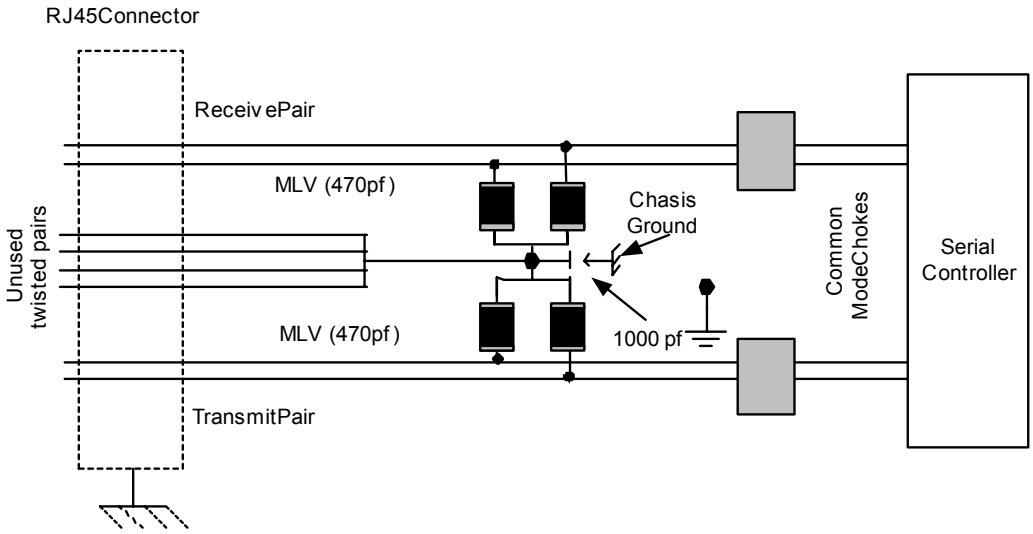


Figure 12-9 Intrabuilding RJ-45 Protection Circuit

12.2.1.6 Phone/Modem Circuits

Over voltage and over current circuit protection is required when interfaces are connected to telephone circuits, such as modems. Tip and ring circuits are subject to lightning and power cross transients and protection is required to prevent these transients from entering the AdvancedTCA chassis.

For CPE rated interfaces, each transmit and receive pair is required to pass all telecom equipment and must be tested and certified to meet industry standards such as Telcordia GR-1089, TIA/EIA-18-968 (was FCC Part 68), and ITU K.20,.21 and.45. Telecom rated fuses are also required to meet specific surge tests for certain countries. Littelfuse GDTs, Sibod Thyristors, and telecom rated fuses (461) are designed for this purpose. Figure 12-10 shows a typical telecom protection circuit for phone line/modem ports.

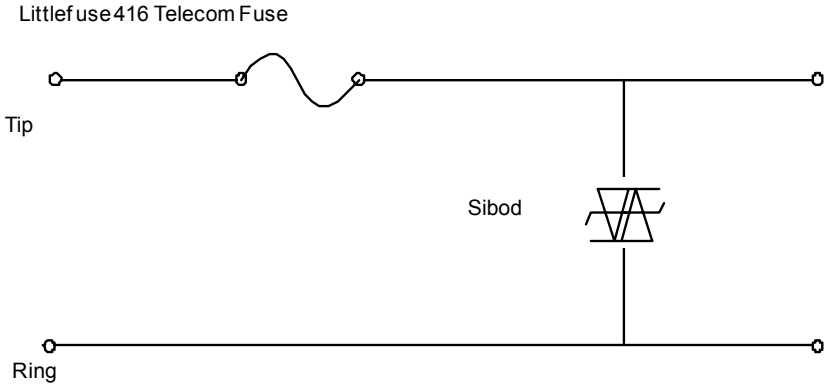


Figure 12-10 Phone Line/Modem Port Protection Circuit

12.2.1.7 Primary and Secondary Protection Circuit

Telecom circuits that interface the outside world require both primary and secondary protection. This requires a combination circuit with over voltage and over current devices that work together to suppress damaging transients. **Figure 12-11** shows a circuit that provides both primary and secondary protection for ungrounded tip-ring pairs. This circuit can be used for interface circuits that connect to telecom lines that are susceptible to lightning and power cross transients.

ID	Feature Name	Requirements
DP-33	Telecom Protection	Circuit protection devices ***should*** be tested together to assure proper operation under current industry standards.

Littelfuse circuit protection devices meet this requirement. All telecom equipment must be tested and certified to meet industry standards such as Telcordia* GR-1089, TIA/EIA-18-968 (was FCC Part 68), and ITU K.20, .21 and.45.

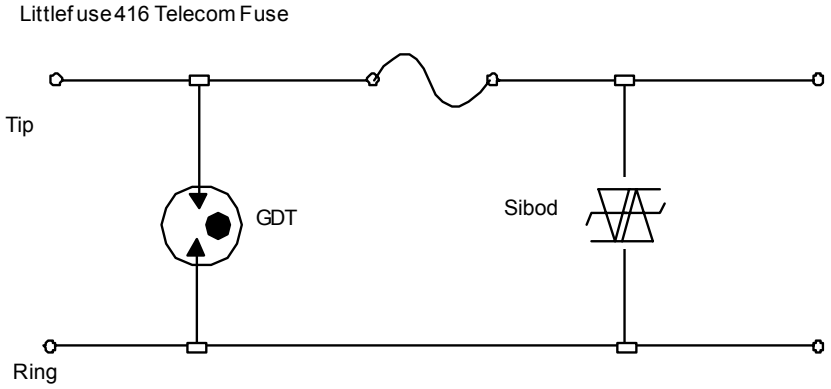


Figure 12-11 Primary and Secondary Protection Telecom Circuits

12.2.1.7.1 Longitudinal and Metallic Protection Circuit

Grounded systems require specific application circuits to provide complete protection for primary and secondary transients. The main difference in these circuits is the reference to ground from the tip and ring signal lines. In this case, each line must have over voltage and over current protection connected to ground. **Figure 12-12** shows the proper circuit protection configuration for these circuits. Tip and ring circuits are subject to lightning and power cross transients and protection is required to prevent these transients from entering the AdvancedTCA chassis.

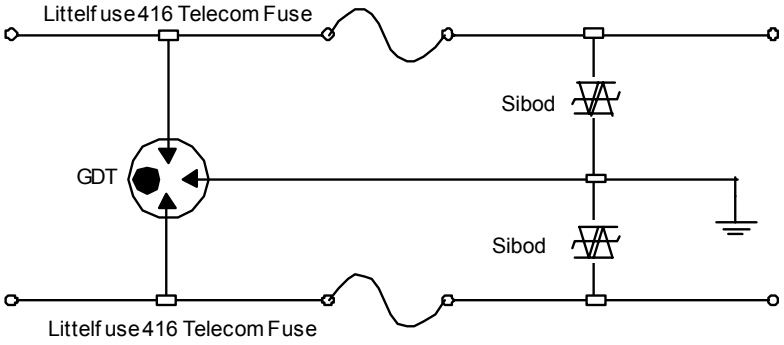


Figure 12-12 Telecom Protection for Grounded Systems

12.2.2 Grounding Guidelines for the BP

ID	Feature Name	Requirements
DP-34	Telecom Protection	Backplanes ***shall*** have high frequency decoupling capacitors between Logic Ground and Chassis Ground located close to every backplane-mounting hole. The value of the capacitors will be determined during the

ID	Feature Name	Requirements
		design phase.
DP-35	Telecom Protection	The capacitors ***should*** be 603 size and 50V rated. Decision on whether to place the capacitors will depend on the EMC test results.
DP-36	Telecom Protection	Filter capacitors ***shall*** be connected between the Chassis Ground and -48V planes immediately after the onboard fuses and close to the backplane power connector. 0.1uF, 250VDC capacitors in 1812 package size are desired for EMC performance. For example, VJ9427/X7R from Vishay Vitramon* is suitable for this usage.

12.3 Front Panel Connectors and Pin Assignments

12.3.1 Serial Port

Serial ports are implemented on compute boards and fabric boards for OOB management. The connectors are 8-pin RJ-45, when the modem interface is not supported and 10-pin RJ-45 when the modem interface is supported.

The commonly available adapter is an 8-pin RJ 45 to DE-9. This adapter could also be used with a 10-pin RJ45 connector. However, the additional 2-signals on the 10-pin connector will not be available when this adapter is used.

Table 12-2 8-pin Serial RJ-45 Pin Assignments

Pin Number	Signal
1	RTS
2	DTR
3	TXD
4	GND
5	GND
6	RXD
7	DSR
8	CTS

The following port supports an external modem.

Table 12-3 10 Pin, Serial RJ-45 Pin Assignments

Pin Number	Signal
1	RI
2	RTS
3	DTR
4	TXD

Pin Number	Signal
5	GND
6	GND
7	RXD
8	DSR
9	CTS
10	CD

12.3.2 USB Port

ID	Feature Name	Requirements
DP-37	Telecom Protection	At least one USB port ***should*** exist on the front panels of all compute boards, both embedded and non-embedded. This port is useful for interfacing to floppy, CD-R, CD-RW, or DVD media boards as well as keyboard and mouse.
DP-38	Telecom Protection	Wherever possible, the port ***should*** be USB2.0 compliant; otherwise it ***shall*** be USB 1.0 compliant.

12.3.3 Ethernet Ports

12.3.3.1 Gigabit Ethernet

Pin assignments for the cable interface are given in Table 12-4. Pin assignments for the printed board interface might change depending on the vendor.

Table 12-4 Gb Ethernet in Assignments

Pin Number	Signal
1	Differential pair 1 positive signal
2	Differential pair 1 negative signal
3	Differential pair 0 positive signal
4	Differential pair 2 positive signal
5	Differential pair 2 negative signal
6	Differential pair 0 negative signal
7	Differential pair 3 positive signal
8	Differential pair 3 negative signal

12.3.3.2 Fast Ethernet

Pin assignments for the cable interface are in Table 12-5. Pin assignments for the printed board interface might change depending on the vendor.

Table 12-5 Pin Assignments for the Cable Interface

Pin Number	Signal
1	Differential pair 1 positive signal
2	Differential pair 1 negative signal
3	Differential pair 0 positive signal
4	Differential pair 2 positive signal
5	Differential pair 2 negative signal
6	Differential pair 0 negative signal
7	Differential pair 3 positive signal
8	Differential pair 3 negative signal

12.3.4 Telco Alarm Connector

Telco alarm connectors have not been defined for this type of equipment. The wiring arrangement in Table 12-6 is compatible with Sun Microsystems* Netra Alarm connector. Intel recommended signal names are shown as well.

Table 12-6 Telco Alarm Connector Pin Assignments

Pin Number	Intel Signal Name	Other Popular Signal Name
1	MinorReset0+	Reset0+
2	MinorReset0-	Reset0-
3	MajorReset1+	Reset1+
4	MajorReset1-	Reset1-
5	CriticalAlarm-NO	System-NO
6	CriticalAlarm-NC	System-NC
7	CriticalAlarm-COM	System-COM
8	MinorAlarm-NO	Alarm0-NO
9	MinorAlarm-NC	Alarm0-NC
10	MinorAlarm-COM	Alarm0-COM
11	MajorAlarm-NO	Alarm1-NO
12	MajorAlarm-NC	Alarm1-NC
13	MajorAlarm-COM	Alarm1-COM
14	PwrAlarm-NO	Alarm2-NO
15	PwrAlarm-COM	Alarm2-COM

The alarm signals are connected to a set of dry contact relays that must be capable of carrying up to 100Vdc or 1 A with a maximum rating of 30 W / 30 VA. The reset inputs are timed pulse inputs that are used to clear the minor and major alarm states. (There is no reset for the critical state.) These reset inputs

must be optically isolated from the rest of the system. Reset is accomplished by asserting a voltage differential from 3.3V to 48V for between 200 and 300 ms. The acceptable voltage range is from 0 to 48 VDC continuous and up to 60 VDC at a 50% duty cycle. The current dawn by this input may be as high as 12 mA with no damage resulting from a reversal in polarity.

12.3.4.1 Connector Locations

ID	Feature Name	Requirements
DP-39	Telecom Protection	Connector locations ***shall*** be as per the PIU Front Panel Layout Guidelines.

12.4 Telecom Protection Circuit for Grounded Systems

12.4.1 Creepage and Clearance

Boards and modules ***shall*** comply with the spacing guidelines found in the standards listed. Ensure that spacings are maintained on the boards (especially top and bottom surfaces of the board) and as part of the board install into a chassis.

ID	Feature Name	Requirements
DP-40	Creepage and Clearance	Creepage and Clearance ***shall*** meet the requirements of the following standards: • UL60950 • IEC60950 • EN60950

The following creepage and clearance information can be used as an ***example*** only. Note that interpretation of creepage and clearance guidelines can vary between certifying agencies. Certain geographies may have more stringent requirements than those listed below. This information should always be carefully reviewed with your safety certifier to ensure accuracy. The following is one example from one certification agency:

Table 12-7 Creepage and Clearance Guidelines

Circuit Type	Insulation Type	Creepage	Clearance	Working Voltage
-48V, VRTN to –48V, VRTN	Functional	1.4mm	0.6mm	72VDC
-48V, VRTN to Shelf Ground	Basic	1.4mm	0.7mm	72VDC
-48V, VRTN to Logic Ground, IPMB, Zone 2, 3.3V	Basic	1.4mm	0.7mm	72VDC
-48V, VRTN to MT	Basic	2.0mm	0.9mm	200VDC
-48V, VRTN to RG	Basic	1.5mm	0.9mm	153pVAC/108VRMS
Shelf Ground to IPMB, Zone 2, 3.3V	Functional	1.2mm	0.4mm	< 15VDC
Shelf Ground to Logic Ground	Functional	0.4mm	0.4mm	0V to couple of volts
Shelf Ground to MT	Basic	2.0mm	0.9mm	200VDC

Circuit Type	Insulation Type	Creepage	Clearance	Working Voltage
Shelf Ground to RG	Basic	2.0mm	0.9mm	153pVAC/108VRMS
Among IPMB, 3V, Logic Ground, Zone 2	Functional	N/A	N/A	<15VDC
Logic Ground, IPMB, Zone 2, 3.3V to MT	Basic	2.0mm	0.9mm	200VDC
Logic Ground, IPMB, Zone 2, 3.3V to RG	Basic	2.0mm	0.9mm	153pVAC/108VRMS
MT to RG	Functional	2.0mm	0.6mm	200VDC
MT to MT	Functional	2.0mm	0.6mm	200VDC
RG to RG	Functional	1.5mm	0.6mm	153pVAC/108VRMS

Note: The *-48V signal also includes EARLY and ENABLE signals on the Zone 1 connector.*

The above guidelines are based on the following assumptions:

1. Pollution Degree 2
2. Material Group is IIIb
3. Transients are less than 800V on the Ring Generator and Metallic Test signals
4. –48V and VRTN are considered to be TNV2 circuits.

If the peak working voltage does not exceed 71V, there is no requirement for distance through insulation.

ID	Feature Name	Requirements
DP-41	Creepage and Clearance	If the peak working voltage exceeds 71V, the following rules apply: a. For functional and basic insulations, there is no requirement at any peak working voltage for distance through insulation. b. Supplementary or Reinforced insulations ***shall*** have a minimum distance through insulation of 0.4 mm.
DP-42	Creepage and Clearance	For functional insulation, clearances, and creepage, distances ***shall*** satisfy one of the following requirements: a, b, or c. Clearances and creepage distances ***shall*** also satisfy one of these requirements for insulation between a secondary circuit and an inaccessible conductive part that is earthed for functional reasons. a. They meet clearance and creepage distance requirements for functional insulation in 2.10 or b. They withstand the electric strength tests for functional insulation in 5.2.2 or c. They are short circuited where a short circuit could cause: i. Overheating of any material creating a risk of fire, unless the material that could be overheated is flammability Class V-1 or better or ii. Thermal damage to basic, supplementary, or reinforced

ID	Feature Name	Requirements
		insulation, thereby creating a risk of electric shock.

12.5 Symbols

A library of graphical symbols suitable for use in actual product graphics is available at http://www.intel.com/go/atcadg.

Appendix A: Requirement Naming Convention

This appendix describes the two-letter naming convention used in the creation of the Requirement Ids in each chapter of this document:

Chapter	ID	Feature Name
4	FT	Fabric Topology
5	SH	Shelf
6	BD	Boards
7	SM	Shelf Management
8	SI	Synchronization Interface
9	AP	Automatic Protection Switching
10	OS	Carrier Grade OS
11	CP	Certification and Practices
12	DP	Design Practices

This page is intentionally left blank.

Appendix B: Carrier Grade OS

B.1 CG OS Features

The features of a Carrier Grade OS can be categorized as shown in Figure B-1 Carrier Grade OS Features.

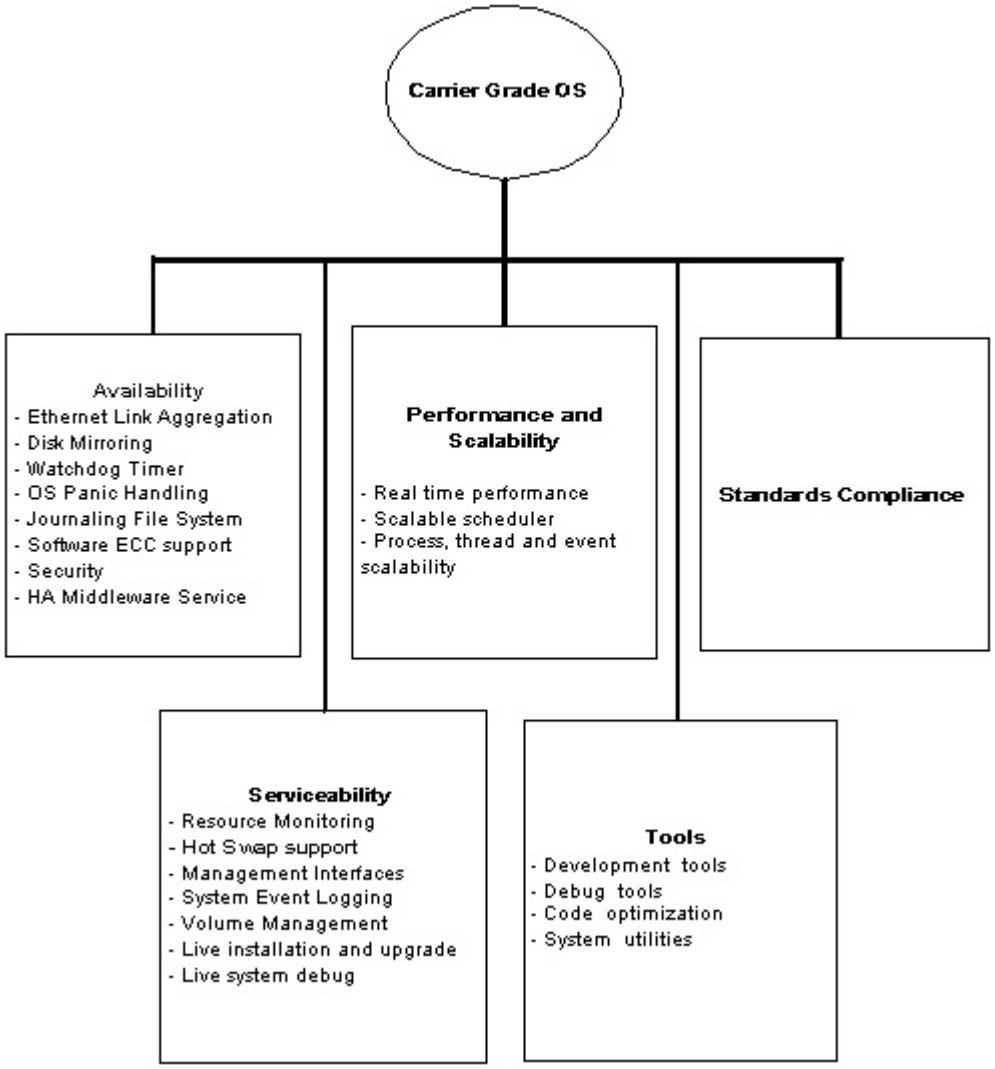


Figure B-1 Carrier Grade OS Features

B.1.1 Availability features

The Availability features of a CG OS fall into three main functional categories:

Feature	Description
Platform availability software	This monitors the redundant sets of hardware, detects failures, and fails over to the standby hardware when the active fails. Ethernet Bonding, Disk Mirroring, Watchdog Timer support are features in this category.
Enhancements	Specific software enhancements build robustness into the OS and make it resilient to device failures, and security attacks. Panic error handling, Journaling File System support, software ECC support, security features and application heartbeat monitoring are features in this category.
Essential services and APIs	A third category of CG OS HA features includes software that provides essential services and APIs for HA middleware and applications so that system level HA mechanisms like clustering can be built on top of the CG OS.

The following discussion describes these features in more detail.

Feature	Description
Ethernet Bonding and Link Aggregation	To ensure Highly Available Ethernet connectivity between nodes and to an external network in an AdvancedTCA system, the node boards will typically include redundant Ethernet controllers (NICs). The Ethernet Bonding driver in the CG OS makes the redundant NICs appear as a single virtual Ethernet interface, and supports automatic failover from one NIC to another. In addition as mentioned in Section 4.4 when multiple Ethernet links are “trunked” together for aggregating bandwidth, the bonding driver should support Dynamic Link Aggregation compliant with the IEEE 802.3ad standard.
Software disk mirroring support	AdvancedTCA systems need to ensure that disk data corruption or disk failures do not lead to system downtime. Employing a Redundant Array of Inexpensive Disks (RAID) system is one way to meet this requirement. The Software RAID 1 feature in the CG OS mirrors disk data on multiple disks, detects disk failures, and ensures continued access to data in the event of disk failures. In addition, the feature supports transparent re-synching of spare drives that are added to the disk array to replace defective ones.
Watchdog timer support	Watchdog timers are often used to detect application hang conditions. The CG OS should provide the driver that interfaces with the system watchdog timer.
Handling of irrecoverable OS errors (panics)	When an irrecoverable OS kernel error occurs, the CG OS should provide a set of configurable functions including logging the “panic” in the system event log, options to reboot, power down, or power cycle. The Shelf Management Module can send an SNMP trap in response to the “panic” event that can result in automatic recovery procedures initiated from a management console.
Support for resilient	A CG OS should provide support for resilient file systems that permit rapid

Feature	Description
file systems	file system recovery from system crashes. They also improve file system robustness. A Journaling File System (JFS) is an example of a resilient file system. Using database journaling techniques, a JFS can restore a file system to a consistent state in a short time. A non-resilient file system would require a complete file system check in the event of an error, and take much longer to restore the file system.
Software ECC support	A CG OS should provide support for detection and logging of parity and ECC memory errors. This facilitates early detection of memory failure and appropriate corrective action before the system fails due to memory failure.
Security features	Security attacks can compromise a system operation and result in downtime. A CG OS should possess host and network security features to prevent that from happening. Secure user authentication schemes, access control mechanisms, features that prevent and contain effects of “root” compromise, cryptographic software and APIs, IPSec support, secure key exchange mechanisms like IKE, are some of the key features a CG OS is required to have.
OS services and APIs for HA middleware and applications	AdvancedTCA High Availability systems often complement the component level redundancy features with application and node level redundancy mechanisms such as clustering. Applications and HA middleware that implement HA clusters rely on CG OS services to access and control platform events like hot swap. In addition, the CG OS may provide core clustering services like system heartbeats, cluster communication, and data check pointing. To build interoperable systems, it is important that these services are exposed to middleware and applications through standard APIs.

B.1.2 Serviceability Features

These are CG OS features that support servicing and managing hardware and software on carrier grade systems.

Feature	Description
Resource monitoring	A capability to monitor the resource utilization of the software drivers and applications is critical to the early detection of hardware and software faults. Some of the critical resources that need to be monitored are disk space usage, CPU and memory utilization, and network usage. In addition to providing resource monitoring subsystems, a CG OS should provide a resource monitor framework that makes it easy to “plug-in” additional resource monitors, and an API to access the monitoring subsystems.
Hot swap support	The CG OS must be able to support the hot swapping of Field Replaceable Units (FRUs) without bringing the system down. For example hot swapping AdvancedTCA storage blades can result in new devices entering the system or existing devices leaving the system. In the removal of a storage blade, devices that are in use by applications may be removed. The CG OS must support features like “Forced Unmount” to recover from such conditions. As FRUs in an AdvancedTCA chassis undergo hot swap, logical device names assigned by the OS based on the

Feature	Description
	location of the device in the enumeration sequence, are likely to change. This might result in the same device having different logical names at different times during the operation of the system. This lack of persistence of device names can give rise to system configuration issues, such as configuration files that embed logical device names will have to change. A CG OS should support persistent naming of devices.
Management protocols and interfaces	To manage AdvancedTCA shelves and server blades remotely, a CG OS should support the industry standard systems and platform management protocols like SNMP and RMCP. SNMP support should include support for SNMP traps that enable asynchronous event notification from the SNMP agent to the SNMP manager. SNMP protocol support should be provided for IPv4 and IPv6 environments, and should include full support for MIB – II including a rich set of baseline MIBs, such as HOST-RESOURCES-MIB, UCD-SNMP-MIB, NET-SNMP-MIB, NET-SNMP-AGENT-MIB, SNMP-NOTIFICATION-MIB, SNMP-TARGET-MIB, etc. Remote Management Control Protocol (RMCP) provides the capability for IPMI messaging over the LAN. The CG OS should support RMCP to enable communication between the system manager and the shelf manager. Section 7.9 discusses the management of AdvancedTCA shelves using CIM. When the system manager of a system of AdvancedTCA shelves is a CIM client, an AdvancedTCA compute blade or shelf manager that is based on a CG OS should provide the software framework to support third party Common Information Model (CIM) Provider and Object Manager applications. The CIMOM will use CIM providers to map CIM objects to underlying platform-specific interfaces such as HPI, SNMP, etc.
System event logging	A CG OS should include a system wide event logging framework with features such as: • Support for text and binary log data • Event records tagged with fixed structure attributes • Access to logged events through standard APIs • Ability to read events from the log that match a user-specified filter • Ability to be notified when specific events occur • Secure remote access to logged events • Flexible framework that allows configuration of log destinations, size, etc.
Volume management support	Volume Managers allow hard disks to be managed without incurring downtime, by grouping physical disks into logical volumes. Logical volumes can be enlarged in size by adding storage space during runtime. If the file system on the logical volume supports it, they can be enlarged without interrupting the applications using them. AdvancedTCA storage applications require the CG OS to support a volume manager.
Live installation and upgrade of software	The ability to upgrade and install software on a live system considerably enhances the system's serviceability. An example of a situation where this may be required is when security vulnerability in a carrier grade system is detected and a software patch is to be applied without taking down the system. A CG OS should support live software upgrades locally and

Feature	Description
	remotely.
Live system debug and dump support	CG OS kernel and application dumps and tools to analyze the dumped data are critical in debugging certain runtime failures. The dumps can be taken after a crash of an application or even while an application is running. Features like multiple dump targets (disk, network, and memory), configurable dump sizes, and data further enhance this feature.

B.1.3 Scalability and performance features

Scalability features of a CG OS permit acceptable increase in performance and capacity of the carrier grade systems with addition of hardware resources. Performance features in the CG OS enable applications run on the carrier grade systems to achieve performance levels that meet their requirements. It is important to recognize that performance requirements vary widely between applications and hence the choice of a particular CG OS depends largely on the application requirement.

Feature	Description
Real time performance support	Event response times dictate the real time performance of a CG OS. Interrupt, scheduling, and context switch latencies in the system determine response time. When application tasks have strict response time deadlines that need to be met under all workloads, the requirement is called a hard real time requirement. Soft real time applications have less strict requirements. CG OS features like pre-emptible OS kernel, low latency schedulers, flexible scheduling policies, and schedulers that can force affinity between processes interrupt handlers and CPUs in a multiprocessor system, and priority inheritance mechanisms that prevent priority inversion, significantly enhance the real time performance of the CG OS.
Scalability features	A CG OS is likely to be deployed across different types of AdvancedTCA platforms, ranging from embedded platforms with limited computing and memory resources to medium and large scale servers with multiple physical and logical (hyper threaded) CPUs, and large memory and I/O capacities. This calls for software that scales well across these platforms.
Scalable scheduler support	When the number of CPUs and software tasks in a system increases, the OS scheduler must schedule the tasks on the different CPUs taking into account among other things: o The load balancing between the CPUs o Maintenance of process affinity to CPUs so that a process is not scheduled on a different CPU every time, causing cache thrashing o Being aware of processor features like hyper threading, so that processes are scheduled on available physical CPUs first before they are scheduled on logical CPUs o Ensuring that the scheduling algorithm takes roughly the same time (O(1) complexity) to schedule the processes when the number of processes and CPUs in the system increases
Process, thread and event scaling	A CG OS should accommodate a large number of processes and threads while delivering the required performance. Process or thread creation times should stay within acceptable limits with increasing workload. Other

Feature	Description
	OS features like fast mutual exclusion primitives (mutexes) that reduce thread and process synchronization time, and efficient mechanisms to deliver a large number of low level asynchronous events to user space applications help enhance the scalability of the OS.
Network and I/O scalability	Features in a CG OS that enhance network and I/O scalability include: o Features that reduce interrupt overload due to heavy network traffic like, support for polled mode network drivers o Support for copy avoidance techniques that minimize buffer copies between user and kernel space memory o Asynchronous I/O capability

B.1.4 Standards Compliance

One of the key goals of AdvancedTCA modular communication platforms is interoperability between different vendor hardware. A CG OS can extend this interoperability to application software, by supporting key industry standard interfaces. Applications written to such standard APIs can also be easily ported to run on other Operating Systems. IEEE* POSIX, SA Forum's Hardware Platform Interface (HPI) and Application Interface Specification (AIS), the LINUX* Standards Base (LSB) are some of the standards that a CG OS should implement to support the development of interoperable application software.

B.1.5 Development, debug, and optimization tools

The availability of easy to use software development, debug, and optimization tools can speed up the software development cycle and help highlight application design issues early in the cycle. A CG OS should provide a rich set of tools in the following main categories:

Tools	Description
Development tools	These include compilers, Integrated Development Environments (IDE), source control and versioning tools, performance libraries, software development kits (SDK) for special runtime environments like Java. Intel compilers and performance libraries help generate optimized code for Intel processors including advanced support for Hyper-Threading Technology and Streaming SIMD Extensions 3 (SSE3). Microcode programming for the IXP network processors can present a greater challenge than programming general purpose processors. Special compilers that can simplify the complex task of writing IXP microcode by providing a more "C" like language and providing a certain amount of automation with respect to distributing functionality across micro-engines, optimizations, etc. can speed up the development time considerably.
Debug tools	These include user space and kernel debuggers, thread enabled debuggers and core dumps, trace tool kits, fault injection framework and are extremely valuable during the code development and debug phase. Other tools such as dynamic probes that allow insertion of software instrumentation into a running system help in debugging random software issues that are difficult to characterize.
Code optimization	This is a critical part of the software development process. The CG OS

Tools	Description
	should support a rich set of profiling and optimization tools. Such tools can help locate “hotspots” in the code that can be optimized to extract maximum performance. The Intel VTune™ Performance Analyzer is a powerful tool in this category, which can identify sections of code that cause significant micro-architectural events like cache misses, mis-prediction of branches, pipeline stalls, etc and helps focus the optimization effort on such code.
System utilities	The CG OS should support a set of system utilities that can be used from the command line as well as applications to get valuable run-time information on the system parameters like CPU utilization, process information like time spent waiting for I/O, virtual memory statistics, network statistics, etc.

B.2 Carrier Grade OS Industry Efforts

B.2.1 Carrier Grade Linux*

Keeping in mind the special requirements that telecom applications place on the CG OS (as discussed above), there is an industry effort to define a version of the open source Linux* operating system to meet those requirements. Open Source Development Lab (OSDL) has created the Carrier Grade Linux* Working Group (CGLWG*) to develop an architecture and requirement definition for a Carrier Grade Linux* (CGL). Based on open enrollment, the CGLWG* comprises open source software developers and equipment suppliers. The CGLWG* is chartered with gathering requirements, identifying open source projects that meet those requirements, promoting development of open source projects to meet requirements, and publishing a Requirements Definition document that will serve as a guide to vendors of Linux distributions who want to release Carrier Grade versions of Linux. Interoperability between different vendor implementations of CGL is achieved through CGL requirements that specify standard interfaces like LSB, POSIX, etc. More information on the CGLWG, their charter and other details can be found at http://www.osdl.org/lab_activities/carrier_grade_linux/documents.html.

B.2.2 CG Linux Requirements

The current version of the CGL requirements document is 2.0 and can be found at http://www.osdl.org/docs/carrier_grade_linux_requirements_definition___version_20.pdf. The requirements are assigned a category to indicate the core nature of the requirement. The table below provides a sampling of **some of the priority 1** requirements for CG Linux. The POC Reference column refers to open source projects that implement the requirement. The POC projects have a Maturity field (not shown in table below) that rates their stability (Production or Experimental).

Table B-1 Requirements for Carrier Grade Linux

Name	Priority	Category	POC reference
Linux Standard Base Compliance	1	Standards	http://www.linuxbase.org/
IPv6	1	Standards	http://linux-ipv6.org/
IPSECv6	1	Standards	"
MIPv6	1	Standards	"
POSIX Threads	1	Standards	Native POSIX Thread Library (NPTL): http://people.redhat.com/drepper/nptl/
SNMP for IPv4 & IPv6	1	Standards	http://www.netsnmp.org
Hot Plug Insertion	1	Platform	http://sourceforge.net/projects/linux-hotplug http://www.sourceforge.net/projects/atca-hotswap
Hot Plug Deletion	1	Platform	"
Ethernet Link Aggregation	1	Platform	http://www.sourceforge.net/projects/bonding http://sourceforge.net/projects/e1000/
Ethernet Link Failover	1	Platform	"
RAID 1 Support	1	Platform	Linux Kernel http://scsirastools.sf.net
Resilient Filesystem Support	1	Platform	Linux Kernel
Disk and Volume Management	1	Platform	LVM http://sistina.com/products_lvm.htm
Watchdog Time Interface Requirements	1	Availability	http://sourceforge.net/projects/openipmi
Application Heartbeat Monitor	1	Availability	http://linux-ha.org
Software Remote Update and Installation	1	Availability	System Installation Suite (SIS): http://sisuite.sourceforge.net/
Kernel Message Structuring	1	Serviceability	http://evlog.sf.net
Dynamic Debug / Probe Insertion	1	Serviceability	Kernel Dynamic Probes (dprobes) (http://oss.software.ibm.com/developerworks/opensource/linux/projects/dprobes/)
Platform Signal Handler	1	Serviceability	Machine Check Architecture (http://sourceforge.net/projects/ia32mcahandler)

Name	Priority	Category	POC reference
User-Level (gdb) Debug Support for Threads	1	Tools	Thread-GDB (http://sourceforge.net/projects/thread-gdb)
Kernel Debugger	1	Tools	KDB (http://oss.sgi.com/projects/kdb/) KGDB (http://sourceforge.net/projects/kgdb/)
Efficient Low-level Asynchronous Events	1	Scalability	Asynchronous Event Mechanism (AEM): http://sourceforge.net/projects/aem/
Soft Real Time Support	1	Performance	Linux kernel
Cluster Node Membership with Failure Detection	1	Clustering – Cluster Fault Handling	
Cluster Communication Service – Logical Addressing	1	Clustering – Cluster Communication and Messaging	
Cluster Communication Service – Fault Handling	1	Clustering – Cluster Communication and Messaging	
Support for Generic Kernel Security Modules	1	Security – Access Control	Linux Security Module: http://lsm.immunix.org/
Support for IKE	1	Security – Confidentiality and Integrity	FreeS/WAN: http://www.freeswan.org/

This page is intentionally left blank.

Appendix C: Glossary

The abbreviations and unique terms used in this document are listed below:

Abbreviation	Description
ADM	Add/Drop Multiplexor
AIS	Application Interface Specification
ANSI	American National Standards Institute
AdvancedTCA*	Advanced Telecommunications Compute Architecture
APS	Automatic Protection Switching
ATM	Asynchronous Transfer Mode
Backplane	A passive circuit capable of connecting one or more link ports from each Node Slot with one or more Nodes or Fabric Slots. The backplane also provides for power distribution and management bus connectivity for all Node and Fabric Boards.
BaseCX (1000BaseCX)	IEEE 802.3 physical layer for 1Gb/s Ethernet over a shielded copper medium, star topology, 25m maximum segment length. Most closely resembling a backplane star topology copper interconnect.
BGP	Border Gateway Protocol
BITS	Building Integrated Timing Source
BLSR	Bidirectional Line Switched Rings
Board (or Module)	An assembled PCB card that plugs into a shelf slot.
CGL	Carrier Grade Linux
CGLWG	Carrier Grade Linux Working Group
CGOS	Carrier Grade Operating System
CIM	Common Information Model
CIMOM	CIM Object Manager
CLI	Command Line Interface
CMM	Chassis Management Module - In the Telco world a chassis is know as a Shelf, hence a CMM is referred to as a SMM (Shelf Management Module)
COTS	Commercial Off The Shelf – versus custom
CPU	Central Processing Unit. A microprocessor.
DSP	Digital Signal Processor
Dual Star Topology	A fabric topology in which two switch resources provide redundant connections to all end points within the network. The Base Interface is defined as a Dual Star fabric topology for all PICMG 3.0 compliant backplanes. For the Fabric Interface, Dual Star provides the minimum redundant fabric environment required for compliant backplane configuration. Up to 14 Node Board/Slots utilize two Fabric Channels to support a connection to each of two Fabric

Abbreviation	Description
	Boards/Slots. Fabric Boards/Slots support a connection to all Node Boards/Slots within a shelf and to the other Fabric Board/Slot. (Also Dual Star Backplane.)
ECTF	Enterprise Computer Telephony Forum
ESD	Electro Static Discharge
ESR	Equivalent Series Resistance
ETSI	Environmental Testing for Telecommunication Equipment based on requirements of European Telecommunications Standards Institute
Fabric Board	A board intended for use in a star topology backplane that provides connectivity to a number of Node Boards within the backplane. Fabric Boards may support either or both the Base Interface and Fabric Interface. Boards utilizing the Fabric Interface will typically provide switching resources to all 15 available Fabric Channels. Compliant Fabric Interface support is defined within PICMG 3.X subsidiary specifications. Fabric Boards supporting the Base Interface are installed into Logical slots 1 & 2 and utilize all 16 Base Channels to provide IEEE* 802.3ab compatible Ethernet switching resources to up to 14 Node Boards and the other Fabric Board.
Fabric Channel	A physical connection within the Fabric Interface composed of up to 8 differential signal pairs (2 adjacent rows) along the Zone 2 ZD connector. Fabric Channels are mapped to ZD connectors J20/P20 through J23/P23. Any system slot in a backplane may support between 2 to 15 Channels. Each Fabric Channel is the endpoint of a slot-to-slot connection such that a system slot/board with 2 Channels supports connections to 2 other systems slots/boards. Fabric Channels are numbered 1 through 15 and slots/boards always support them in sequential order starting with Fabric Channel 1. Fabric Channels are sub-divided into four 2-pair Ports and may be Single Port (2-pair), Double Port (4-pair) or Full Channel (8-pair) implementations.
Fabric Interface	A Zone 2 interface that provides connections comprised of up to 8 differential signal pairs (Channel) between end-points. Compliant backplanes may support the Fabric Interface in a variety of configurations including Full Mesh and Dual Star topologies. Boards that support the Fabric Interface may be configured as Node Boards, Fabric Boards or Mesh Enabled Boards. Compliant board implementations of the Fabric Interface are defined by the PICMG 3.x subsidiary specifications.
Fabric Slot	A slot in a backplane that is capable of supporting Fabric Boards, Mesh Enabled Boards or Node Boards. In a Star topology backplane, Fabric Slots must reside in Logical Slots 1 & 2, and possibly Logical Slots 3 & 4. All slots in a Full Mesh backplane are considered Fabric Slots. Fabric Slots support both the Base Interface and Fabric Interface. Typically Fabric Slots located in Logical Slots 1 & 2 are capable of supporting both Base Interface and Fabric Interface Fabric Boards. Logical Slots 1 & 2 are always Fabric Slots regardless of the Fabric Topology. These slots support 16 Base Channels and 15 Fabric Channels each. Fabric Slots located in Logical Slots 3-16 support 2 Base Channels and up to 15 Fabric Channels.
Frame	A physical or logical entity that may contain one or more shelves. Also called a rack, or, if enclosed, a cabinet.
ForCES	Forwarding and Control Element Separation
Front Board	Front Board includes a PCB, a Panel and connects with the Zone 1 and Zone 2 Backplane connectors and optionally may connect with a Zone 3 Midplane

Abbreviation	Description
	Connector or directly to a RTM connector. A board that conforms to PICMG 3.0 mechanicals (8U x 280mm) and is installed into the front portion of a shelf or frame. (that is, Front Plug-in Board or Front Board.)
FRU	Field Replaceable Unit. Logically represents an entity that can be queried for sensor data, and perhaps an entity that has a satellite management controller. Example FRUs can include an entire SBC or a power supply.
Full Channel	A Fabric Channel connection that uses all 8 differential signal pairs between end-points.
Full Mesh Topology	A fabric topology in which all network end-points have a direct connection to all other end-points. Full Mesh configurations may be supported within the Fabric Interface to provide a highly redundant fabric environment capable of very large aggregate bandwidth capacity across the shelf. Full Mesh configured backplanes are capable of supporting Mesh Enabled Boards or Fabric and Node Boards installed in a dual star arrangement. (That is, Full Mesh Backplane.)
Gb/s	Gigabits per second
GB/s	Gigabytes per second
GDT	Gas Discharge Tube
HA	High Availability
HLD	High-Level Design
HLR	Home Location Register
Hot Swap	A specific specification from the PICMG 3.X Specification.
HP	IEC 60297-xxx Subrack aperture width for Panels (1 Horizontal Pitch (HP)=5.08mm; max 84HP for 19inch Subracks).
HWM	Hardware Management. The combination of hardware, firmware and software used to monitor and control the hardware resources in a system. Also called Platform Management.
I^2C	Inter-integrated Circuit bus. A multi-master, 2-wire serial bus used as the basis for current IPMBs.
IA-32	32-bit Intel Architecture. Intel's family of 32-bit processors.
IA-64	64-bit Intel Architecture. Intel's 64-bit processors.
IAD	Integrated Access Device
IBIS	I/O Buffer Information Specification
ICMB	Intelligent Chassis Management Bus. A character-level transport for inter-shelf communications between intelligent chassis.
IDE	Integrated Development Environment
IEC	International Electro technical Commission
IEEE	Institute of Electrical and Electronics Engineers
IETF	Internet Engineering Task Force
I/O or IO	Input / Output
IP	Internet Protocol
IPMB	Intelligent Platform Management Bus. Name for the architecture, protocol, and

Abbreviation	Description
	implementation of an out-of-band bus that provides a communication path for platform management information.
IPMI	Intelligent Platform Management Interface. A specification and mechanism for providing inventory management, monitoring, logging, and recovery control for elements of a computer system that is independent of the main processors, BIOS, and operating system.
Intel® IXA SDK	Intel® Internet Exchange Architecture (Intel® IXA) Software Development Kit
iSCSI	Internet Small Computer Systems Interface
Isolation	A method of reducing the inductive and/or capacitive coupling (i.e., crosstalk) between an aggressor and victim.
J2EE	Java* 2 Enterprise Edition
JBOD	"Just A Bunch Of Disks"
JFS	Journaling File System
JVM	Java Virtual Machine
LAN	Local Area Network
LFB	Logical Functional Blocks
LIU	Line Interface Unit
Logical Slot	A System Slot within a shelf defined by the Zone 1 geographic address. Every system slot has a unique Logical Slot # (maximum of 16 Logical slots per shelf). Logical Slot numbers are used to determine Channel mapping between system slots. PICMG 3.0 defines a direct correlation between Channel numbers and Logical Slot numbers. For example, Channel 1 (Base and/or Fabric Channels) of every system slot establishes a direct connection to Logical Slot 1. Logical Slots 1 & 2 are always designated as Fabric Slots.
LP	Link Port. A physical port composed of 8 pin pairs.
LSB	Linux Standards Base
MAC	Media Access Controller
LVDS	Low Voltage Differential Signaling, a modern alternative to PECL for backplane point-to-point applications
MCP	Modular Communications Platform
Mesh	A backplane topology in which there is a dedicated link port connection between each slot and every other slot. In this topology there is no distinction between Node and Fabric Slots.
MIB	SNMP - Management Information Base
MIB-II	SNMP – MIB for managing networks of TCP/IP based internets
MGCP	Media Gateway Control Protocol
MLV	Multi Layer Varistor
Module (or Board)	An assembled PCB card that plugs into a shelf slot. These terms are used interchangeably.
MOV	Metallic Oxide Varistor
MPLS	Multi Protocol Label Switching
NEBS	Network Equipment Building System
NE	Network Element

Abbreviation	Description
NEP	Network Equipment Provider
Node Board	A board intended for use in a star topology backplane that requires connectivity to a central Fabric Board within the backplane. Node Boards may support either or both the Base Interface and Fabric Interface. Boards supporting the Fabric Interface utilize Fabric Channels 1 & 2 (and possibly 3 & 4). Compliant Fabric Interface support is defined within PICMG 3.x subsidiary specifications. Boards supporting the Base Interface utilize Base Channels 1 & 2 only to support 10/100 or 1000BASE-T Ethernet as defined within PICMG 3.0.
Node Slot	A slot in a backplane that supports Node Boards only. A Node Slot is not capable of supporting a Fabric Board, thus never occupy Logical Slots 1 or 2. Node Slots apply only to backplanes designed to support Fabric Star topologies. Node Slots support both the Base Interface and Fabric Interface. Typically, a Node Slot supports two or four Fabric Channels and two Base Channels. Each two Channel Node Slot establishes connections to Logical Slots 1 & 2 respectively and four Channel Node Slots establish connections to Logical Slots 1, 2, 3 & 4 respectively.
NPTL	Native POSIX Thread Library
NPU	Network Processor Unit
OAM&P	Operations, Administration, Management, and Provisioning
OOB	Out Of Band
ORB	Object Request Broker
OSDL	Open Source Development Labs
OSPF	Open Shortest Path First
OSS	Operational Support System
PCB	The Printed Circuit Board (PCB) definition used for both, the Front Board and the optional RTM.
Panel	A Panel is mounted to the front end of a Front Board or RTM providing for EMC protection and I/O interface. The panel includes injector/extractor handles, alignment pins and keying features.
PBX	Private Branch Exchange
PECL	Positive/Pseudo Emitter Coupled Logic – A low power alternative to CMOS/TTL
PEM	Power Entry Module
Physical Port	A port that physically exists. It is supported by one of many physical interface (PHY) type components.
Physical Slot	A system slot within a shelf defined by its physical location within a shelf. Physical Slot 1 is always designated to the left most justified slot and increment sequentially to the right (or bottom justified to top). A system slot may have different Physical slot and Logical Slot number designations. The Shelf FRU ROM provides a mapping of Physical Slot designations to Logical Slot designations for every slot.
PICMG*	PCI Industrial Computer Manufacturers Group
PMC	Peripheral Management Controller. Synonymous with the IPMI-defined satellite management controller.

Abbreviation	Description
PNNI	Private-Network-Network Interface
POSIX	Portable Operating System Interface
PSB	Packet Switching Backplane
PSTN	Public Switched Telephone Network
PTC	Positive Thermal Coefficient
RADIUS	RADIUS protocol used for authentication (usually) of dial-up user accounts
RAID	Redundant Array of Independent Disks
RAS/RASM	Reliability Availability Serviceability and Management
RDBMS	Relational Database Management System
Rear Board	A board that is installed into the rear portion of a shelf or frame. Typically used for I/ O cabling. These boards conform to PICMG 3.0 and IEEE 1101.11 mechanicals (8U x 70mm). (Also Rear Panel Board, Rear Transition Board, Rear Transition Module, Line Interface Unit and Rear Blade.)
RIP	Routing Information Protocol
RMCP	Remote Management Control Protocol (IPMI over LAN)
RNC	Radio Network Controller
RSVP-TE	Resource Reservation Protocol
RTM	Rear Transition Module, reference 8Ux6HPx70 may plug directly into a Front Board or may plug into a specified Zone 3 Midplane connector.
SA Forum	Service Availability Forum
SCCP	Signaling Connection Protocol Part
SCTP	Stream Control Transmission Protocol
SDH	Synchronous Digital Hierarchy
SEL	Sensor Event Log. An IPMI defined term. Maintained by the BMC
SFI-4	SERDES Framer Interface
ShFRU	Shelf Field Replaceable Unit (Stores Shelf Unique ID)
ShMC	Shelf Management Controller. Provides control of and connection to IPMB bus for communicating (using IPMI) with intelligently managed shelf devices, modules, and boards.
Shelf	The shelf consists of the sub-rack, backplane, fabric/node boards, cooling devices, front boards, RTMs, power supplies, and so on. Also known as a chassis.
Shelf Manager	A logical concept of intelligence (software) to perform specific functionality (with associated hardware) to provide management of a shelf (or group of shelves).
SIGTRAN	Signaling Transport
Single Port	A Fabric Channel connection that only uses 2 differential signal pairs (of a possible 8-pair) between end-points.
SLA	Service Level Agreement – A level of service agreed upon between a service provider and an their customer
Slots	A Slot defines the position of one Front Board and or one RTM. Front Boards and RTMs are inline.

Abbreviation	Description
SMBus	A two-wire serial bus. Slightly different electrical and timing characteristics than I^2C. This technology is owned by Intel Corporation.
SMC	Satellite Management Controller. Defined by IPMI. Also known as a peripheral controller or peripheral management controller.
SMM	Shelf Management Module. A concept of hardware and software responsible for managing the shelf modules and board plugged into the shelf backplane. The exact combination of components that make up a SMM may be form factor and implementation specific, however certain functionality is deemed required. Typically it must provide a hardware controller to connect to devices that are to be managed (intelligently managed), a physical interconnect (e.g. IPMB buses), support a communications protocol (e.g. IPMI), and have an external interface(s). If a Shelf Manager is one of the included components it has additional specific requirements (e.g. support for IPMI commands) and support for specific external interfaces and software (e.g. LAN connectivity and RMCP encapsulation of IPMI over IP).
SNMP	Simple Network Management Protocol
SOAP	Simple Object Access Protocol
SONET	Synchronous Optical NETwork
SS7	Signaling System 7
SSE3	Streaming SIMD Extensions 3
Star	A backplane topology in which there is one or more dedicated link port connections between each Node Slot and the Fabric Slot(s).
System	A managed entity that could include one or more Node and Fabric Boards, and/or one or more shelves, and/or one or more frames.
System Slot	Any mechanically compliant slot in PICMG 3.0 shelf.
Sub-rack	The sub-rack provides the interface to PICMG 3.X boards and consists of the card guides, ESD discharge, alignment/keying, injector/ejector interface, and front panel mounting hardware, EMI gasketing, and backplane interface. The sub-rack is a subset of the shelf.
Switch Slot	See Fabric Slot
T1E1.8	Standards Committee T1-Telecommunications
TCP/IP	Transmission Control Protocol/ Internet Protocol
TDM	Time Division Multiplexing
Telephony Clock Interface	A Zone 2 interface that provides synchronization clocks between all system slots. There are 3 distinct, redundant clock pairs that are bussed between all system slots. Backplanes must support the telephony clock interface; boards may support any or all of the clock interfaces.
TFI-5	TDM Framer to Framer Interface
TOE	TCP/IP Offload Engine
TVS	Transient Voltage Supressors
U	IEC 60297-xxx rack, shelf and subrack height increments (1U=44.45mm, 1U = 1.75 inches).
UDP	User Datagram Protocol – A connectionless transport used for streaming data such as audio and video over low bandwidth links

Abbreviation	Description
UNI	User Network Interface
VDC	Volts Direct Current
Victim	A signal inductively and/or capacitively coupled to an aggressor signal and considered to be a receptor of noise that distorts the signal being transferred between transmitter and receiver.
VLR	Visiting Location Register
VPN	Virtual Private Network
XAUI	10Gb/s Attachment Unit Interface
XML	eXtensible Markup Language