

Unique NEBS Telecom Chassis Enables a Scalable HA System Strategy

In today's embedded communications systems, rack space in the Central Office (CO), or the Data Center is scarce and valuable. With five nines availability goals often required, and 1U or better mounting pitch desired for server equipment becoming commonplace, a CompactPCI platform supporting these features is necessary in the marketplace. In addition, many users are looking for scalable system components that can be used with various processor architectures to build telecom application platforms. In order to provide such a building block, one must provide a system chassis that addresses the needs of the telecom community. The first step in the design process is to define the necessary features for a Carrier Grade system platform.

General Purpose Carrier or Telecom Grade Chassis

In order to develop such a platform, let's consider how the telecom community has addressed high availability system architecture, and collect the requirements for a Carrier Grade platform.

Requirements of a Carrier Grade CompactPCI Platform

Carrier Class hardware is a term used to describe a chassis designed for use in telecommunications environments, typically where NEBS Level 3 and ETSI requirements must be met. Typically, a Carrier Class chassis must meet many of the following requirements, including:

- A redundancy scheme to support high availability, five nines or greater
- Modular architecture, supporting removable active components
- Front to back airflow
- Component insertion and removal from the front of the chassis
- All I/O from the rear
- Hot plug, hot swap modules
- Alarm capability - major, minor, critical, and rack
- Redundant power supplies
- -48 VDC input and/or 60 VDC (ETSI)

In addition, an approach that supports scaling the system is often desirable.

Approaches to High Availability

In today's marketplace, there are many schemes being considered to implement high availability systems. Much discussion has taken place regarding Active/Active and Active/Standby schemes. System chassis such as Motorola's 8216 and Ziatech's 5083 series have been developed to allow switching of the CompactPCI domains to take place in order to provide redundancy.

Reviewing approaches being used in the industry, one finds that many use a loosely coupled approach of distributed processors to accomplish the goal. This approach can be referred to as clustering. In this approach, part or all of the processing computer is duplicated, either on a one for one basis, or in an N+1 fashion, depending on the needs of the application. In discussing these approaches with telecom users, there are a few schools of thought. Some want the particular node, or chassis itself to be very reliable, with modular components, where the



chassis will maintain or degrade its performance only slightly as components fail. With this strategy, users are looking to replace boards or power supplies while the system is running. Another school of thought is to define the box as the field replaceable unit (FRU), therefore making the FRU as cost effective as possible, and achieving the availability goal by duplicating the FRU.

Both of these approaches are being used by carrier companies to build communication systems. Since they are both being used by the market, one cannot discount either strategy. Our goal as suppliers is to build systems that will be consumed and meet the market needs. Therefore, a middle of the road strategy would also seem to be desirable, where one considers building low cost nodes or processing elements that may be duplicated and easily replaced, but fit into existing, loosely coupled system strategies currently being deployed. These strategies typically have machines networked together using Ethernet, with a heartbeat monitor running to determine if the other machines on the network are functioning.

In a clustering approach, typically copies of the application along with the operating system will be duplicated across N machines. In some applications, this is two or three; in others where the data set and tasks can be divided or partitioned, this can be across N processors. Such a scalable architecture is applicable to systems where growth is a consideration, such as in HLR or VLR systems used in the wireless marketplace.

Considering the issues associated with the user's HA system strategy, a design that addresses low cost, loosely coupled systems should be considered, starting with the chassis.

The CompactPCI specifications provide a standard to design to, relative to implementation

Chassis Considerations

In order to support system architectures that will approach or exceed five nines availability, one needs to provide a reliable platform within which the logic cards can be contained. The primary goal of the chassis design is hardware fault tolerance.

Issues with building reliable chassis include:

- Redundant cooling systems
- Redundant power supplies both AC and DC
- Split power inlet
- Chassis environmental alarms

Chassis Cooling

Adequate cooling is *fundamental* to building a NEBS compliant chassis. Here, a push-pull scheme can be used to provide adequate airflow through the card payload, and provide for redundant air movers. Further, multiple fans can be used in each module so that a single fan failure does not cause a total loss of function in either module.

Power System

The CompactPCI Power Interface Specification defines methods for implementing a N+1 Power Supply scheme. With CPCI, one no longer needs to use a wire harness to connect the power supplies to the backplane. Power supplies are now available in both 3U and 6U form factors that can be plugged directly into the mid-plane, without cables. This increases the reliability of the system. Using an N+1 scheme, a power supply can fail, while the second supply provides the necessary power to operate the system. Using plug-in supplies allows the defective power supply to be removed and replaced, without powering down the system.

Alarm Functions

Typically, the system chassis should be capable of monitoring temperature, fan status, power supply status, and supply voltages. The system chassis should be capable of reporting Major, Minor, Rack and Critical Alarms.

Processing Logic Considerations

In order to implement a HA system, a strategy must be established for CPU fault tolerance and I/O Card fault tolerance.

Hardware fault tolerance issues to be considered include:

- Use of redundant processors
- Redundant I/O cards
- Domain switching
- Multiple fault domains
- N+1 consideration
- Critical component identification
- Single points-of-failure

Hybrid Approach: Fault Tolerant Hardware vs. System Duplication

Providing hardware fault tolerance at the chassis level, while providing processing fault tolerance at the CPU level through a clustering approach, is a way to implement a clustering architecture at lower cost.

Each processor becomes a redundant element. None of the processors is dependent on each other from a system resource point-of-view.

A Cost Effective HA Solution

Developing a strategy that employs clustering at the board level can be a cost effective means to implement a high availability system. Use of single slot, single board computers with multiple Ethernet ports provides a mechanism to loosely couple N processors, allowing for multiple fault domains. It is possible for all communication between the nodes to take place over Ethernet using an IP solution. In this manner, no inter-CPU communication is required over the CPCI bus.

Many system designers today have satisfied the requirement of redundancy by replicating the entire Processing Node, including I/O and mass storage. This can be costly, although less complex than active/passive multiple domain systems, where CPUs are switched on and off the bus. In these systems, a passive CPU will be running, waiting to take over for a failed CPU, allowing it to drive a payload of I/O cards in the failed domain.

To date, the software to support warm domain fail takeover has been complex, and is now being used to implement such systems.

Using a strategy where the Node is a Board and it represents a FRU, is a middle of the road approach. In this case, application programs are typically loaded over Ethernet and are executed from RAM. A local hard drive maybe mounted on the CPU for a boot image, or for application storage. Other storage may be mounted remotely over NFS, to a reliable storage subsystem, such as a RAID device.

In such a scheme, graceful degradation of performance is possible. Failure of a single FRU may cause no loss or partial loss of capability.

Advantages of the Single Slot FRU

Combining I/O with the CPU function is now possible using the PMC standard and available hardware. PMCs (PCI Mezzanine Cards) with PIMs, (PMC Interface Modules) allow single slot CPU and I/O combinations to be formed, where the FRU still requires only a single slot. PIMs allow the I/O to be mapped to specific connector schemes on the rear transition module for the CPU.

PMCs support E1/T1 interfaces, SCSI, Serial, Ethernet, and other types of I/O.

Segmented vs. Standard CompactPCI Backplane

If a system is comprised of N CPU nodes, connected using TCP/IP over Ethernet, then an interconnecting backplane bus is not required. This approach allows the processors to operate independently from each other. In this case, limiting the CPUs to a single slot, with each CPU supplying its own clock eliminates dependence on a single system controller.

Operating in a cluster scheme for redundancy allows one to operate without a backplane bus. This scheme further indicates that in future systems, in certain system applications, the Ethernet fabric may be of more interest than the actual CPCI bus. The CPCI bus may then be used for maintenance or configuration, rather than actually passing data. Other switched

require a single slot allow the spacing to reach six CPUs per 4U chassis, thus supporting a pitch of 6/4U, or .66U per CPU, where the densities include space for modular removable power supplies.

This scheme supports a better than 1U spacing factor within a chassis or equipment rack. This can become a decision factor considering the cost of Central Office or Data Center space for telecommunications equipment.

NEBS Considerations

Often the system engineer is faced with the implications of meeting NEBS, and asks the question what does it take to pass NEBS Level 3? While many are familiar with Safety and EMC regulatory issues, which on their own are quite involved, the

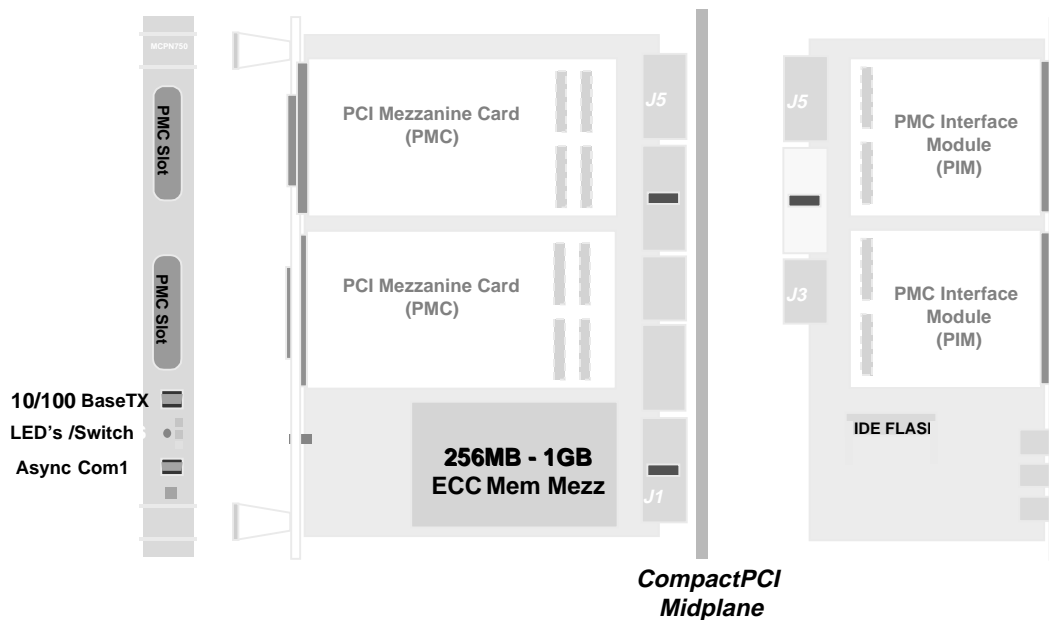


Figure 3.0

fabric architectures, such as Rapid I/O or Stargen may become more useful for passing local high bandwidth data than using the CPCI bus.

Physical Constraints

System Space Considerations

Often, it is the goal of the designer to fit as much processing hardware into the smallest possible space. Depending on the redundancy scheme, this may require packing in as many CPUs or processing elements as possible. Further, one needs to review how much of the hardware must be replicated. With an approach where all of the critical hardware per node fits in a single slot, the desire is to fit as many processors as possible into a given rack space. Lately, a goal has been to build 1U high self-contained chassis, where the chassis itself becomes the FRU.

With CPCI cards requiring .8 inch spacing (4HP), nodes that

NEBS specifications add many additional physical and environmental requirements to achieve compliance. A partial list of tests is included in Table 1. For the uninitiated, the NEBS qualification process is long, time consuming, and often risky and costly.

Compliance vs. Certification

Today compliance is often defined to mean **designed to meet** by the industry, where a chassis and payload may have been taken through its paces at the testing laboratory, but the configuration that you will use has not been certified. Here, the user should examine their specific needs, and ask questions of the supplier to see if the system solution has been certified and listed, or if it has been designed to meet the requirements of a certain operating environment.

Table 1.0

NEBS GR-63 Physical and Environmental Test Matrix

Transportation and Storage
Low Temperature Exposure & Thermal Shock
High Temperature Exposure & Thermal Shock
High Relative Humidity Exposure
Temperature, Humidity & Altitude
Operating Temperature & Relative Humidity
Altitude
Heat Dissipation
Fire Test
Equipment Assembly Fire Spread
Needle Flame
Handling Drop Test
Earthquake Environment
Office Vibration
Transportation Vibration
Airborne Contaminants
Acoustic Noise
Illumination

Design Considerations relative to NEBS

Operating Temperature Extremes

To comply with the Bell Core Standards, the system hardware must operate indefinitely at 40 degrees C, with short-term upper operating temperature excursion to 55 degrees C for 96 hours.

With the inlet temperature at 55 degrees C, the exhaust temperature will be higher, usually 8 to 15 degrees C, meaning the air temperature used to cool the components on the board will probably reach 65 to 70 degrees C along the way.

For power supplies, these operating temperatures translate into reduced output power, where many supplies are de-rated after they reach a 40 degree C operating point. Here, the power system must have adequate margin for the payload to meet the operating point at the upper temperature extreme.

System qualification implications include:

- Design issues with adequate power supply margin
- Difficulty in meeting the requirements with devices such as hard drives.

Flammability Testing

In this case, the chassis payload and cables must not act well as fuel, in case of a fire in the Central Office or Data Center. The Spread Flame and Needle Flame tests subject the chassis internals to an open flame. Restrictions exist relative to flame spread, and flame permeation beyond the chassis. Chassis layout and airflow considerations are critical to passing this test. Fire resistant materials must be used throughout the design, including the material used for the components within circuit card assemblies and power supplies. Plastic card guides will melt, smolder, and smoke. Don't expect to get anything back in working order from this test.

Shock

With respect to structural integrity, the system chassis must be designed to pass severe shock testing. Earth Quake Zone Four is comparable to a catastrophic earthquake. Adequate mounting schemes must be provided. This is an operational test.

Gaseous Contaminants Testing

Another test required for certification is gaseous contaminants. This test involves a 14-day chamber soaking of the hardware, running at 30 degrees C and 70% humidity, while subjected to an atmosphere of particulates, volatile organic compounds, and reactive gases. After this test is complete, the system is then fully powered and functionally tested. While not necessarily practical relative to today's operating environment in the CO, it is still required to pass NEBS.

Designing a system to pass the environmental, EMC and Safety requirements and objectives of NEBS is a challenging effort and a balancing process for the engineering team. For example, increasing the airflow in a system to pass the short-term operating temperature of 55 degrees C may result in failing a test such as the flame spread test, or the acoustic noise test. Changing one variable in a system may affect the results of other tests.

Polaris-HA: General Purpose Carrier Grade Solution

The Polaris-HA is a system chassis that implements the features described above.

The chassis has been designed with all modules accessible from the front. All I/O is available in the rear. The chassis uses N+1, 3U pluggable CPCI Power Supplies.

All modules including the alarm, intake fan, exhaust fan, are removable. The alarm, intake fan, exhaust fan, and power supplies are hot swappable.

The chassis is 4U high by 19" wide by 13.5" deep. It operates from either -48VDC, or from a 110/220 VAC source.

The Polaris-HA allows up to five CPUs to be contained, and operate independently from each other within a CompactPCI chassis. In the 4U high chassis, this equates to .66U per CPU, allowing the Polaris-HA to be space efficient.

Front and rear views of the chassis are found in Figure 1 and Figure 2 respectively.



Figure 1.0 Front View of Carrier Grade chassis

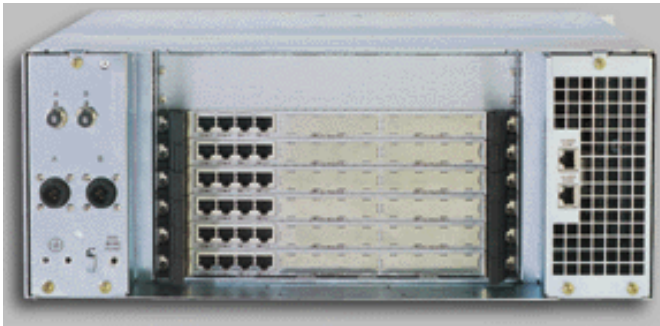


Figure 2.0 Rear View of Polaris chassis

The chassis has been certified to meet NEBS Level 3 and ETSI, using both SPARC and PowerPC CPUs. It has been designed to meet and is pending certification for Pentium CPCI processors from Motorola.

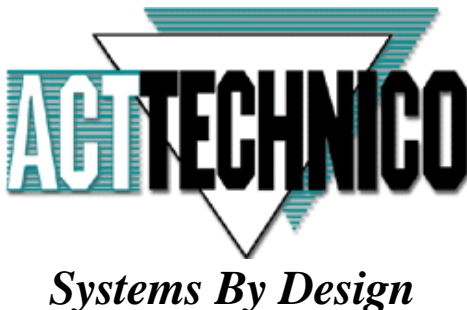
To achieve rugged construction in order to pass shock and vibration tests, the chassis was constructed of steel. The chassis must provide substantial structural integrity to pass the shock and vibration tests.

Environmental monitoring is accomplished through the Alarm Module. The Alarm Module reports via a dual RS232 serial ports using a simple binary protocol, internal temperature, power status, voltage levels, fan status, and the presence of active boards by monitoring the slot Healthy signal. The Alarm Module generates Major, Minor and Critical Alarms.

System High Availability Features

Detection

The Alarm Module is capable of monitoring the Healthy signal to determine if a board is present and powered on. After insertion and power up, the board can go through the system boot process where it is typically fed an application program from an application server over Ethernet.



Recovery Feature

The Alarm Module, via software control is capable of resetting any of the six slots within the chassis. This feature can be used to restart a CPU if an application becomes corrupt or non-responsive, or if a CPU has become non-responsive.

2.16 IP Fabric System Extensions

Currently, a PICMG 2.16 version of the chassis is being planned, where the chassis would contain five CPU slots wired for IP fabric connection, and two 3U switch slots. Two one-gigabit Ethernet ports would provide pipes into the fabric. In a segmented scheme, this would further extend the usefulness of this architecture in an IP environment.

Summary

The Polaris-HA represents a general purpose carrier grade system chassis suitable for implementation of loosely coupled high availability systems. The chassis allows users to achieve efficient use of rack real estate by allowing five CPUs to be used per 4U shelf. Versions of the chassis have been qualified with mixed CPU architectures, including SPARC, and PowerPC. This system chassis is ideal where users desire to run Solaris, Linux, or VxWorks hosted applications within the same platform. The platform has been certified to meet NEBS Level 3 and ETSI standards, and certain configurations have been certified.

Ken Grob, Vice President of ACT/Technico, is the author of this paper. Ken earned his degree in Engineering from Drexel University (BSEE). His experience includes working for Kulicke & Soffa and IBM before he founded Advanced Control Technology and Technico with cofounder Chris Kuebler.

One Ivybrook Blvd., Suite 180
Ivyland, PA 18974
Tel: (215) 957-9102
(800) 445-6194
Fax: (215) 957-9074
Web: www.acttechnico.com