Using CompactPCI to build carrier-grade systems

By Fred Rehhausser

Companies of every sort are rushing to “.com” their businesses, creating new opportunities and challenges for service providers. As a result, these providers need ways to quickly differentiate themselves in the marketplace and to provide a highly reliable, scalable, and available level of service. In this article, Fred describes new CompactPCI solutions that enable Telecom OEMs and Network Equipment Providers create High-Availability systems.

The arrival of the “dot-com” age has magnified the importance of service-level availability. Internet users now expect their network services to deliver the same level of uptime and reliability that their telephony services have always provided. For telecom OEMs and network equipment providers (NEPs), this creates an opportunity for competitive differentiation based on system availability levels.

Carriers and NEPs are now looking for commercial off-the-shelf (COTS) platforms that incorporate High-Availability (HA) “hooks”. In this way, OEMs and NEPs can build their own value-added HA systems with the desired features and services, reducing time-to-market and engineering development costs.

The emergence of today’s dynamic, Internet-driven market

In the not too distant past, AT&T dominated the telecom market, providing end-to-end solutions that worked together remarkably well. Innovation proceeded at a deliberate pace, with new switch designs being deployed every 20 years or so. Today, however, the market moves to Internet time. Consequently, change has become the constant in telecom, and successful players need to be able to upgrade continuously to stay competitive. This new dynamic has, of course, been fueled by the intense competition resulting from the rise of the Internet and the World Wide Web.

At the same time, overwhelming fragmentation in software and hardware solutions has led to major interoperability issues, especially as the telecom market wrestles with convergence that combines voice, data, and video onto a single line. The emergence of specifications for converged information and their quick adoption provides significant increases in network traffic of 50-75 percent over traditional networks. Today’s networks were built for voice calls that last an average of three to four minutes. In contrast, the typical Internet connection is 30 minutes long, creating a tremendous strain on existing networks. This situation is creating an opportunity for new competitors, many with much cheaper infrastructure costs, to enter the market with converged wireline, wireless, switching, or access networks.

As embedded telecom developers rush to fill the transmission gap created by the Internet and convergence, they must address the exacting demands of high availability. With the increasing pervasiveness of the Internet, consumers are demanding the same concept of high availability inherent in the telecom world. In other words, people want non-stop availability inherent in the telecom world. With the increasing pervasiveness of the Internet, consumers are demanding the same concept of high availability inherent in the telecom world. In other words, people want non-stop availability inherent in the telecom world. In other words, people want non-stop availability inherent in the telecom world. In other words, people want non-stop availability inherent in the telecom world.

Consumers expect this same level of reliability to be extended to all their Internet and telecom services, especially if they are converged into a single entity.

Availability defined

Every technology vendor seems to have its own definition of availability, and over the precise meaning of terms such as “high availability”, “continuous availability”, “permanent availability”, and others. Only one definition has any meaning to the people who actually use the technology; application or service-level availability. If end users cannot access the applications or network services on which they depend, then the system is down. And only one level of availability is acceptable to end users: 100 percent.

Consequently, the overarching design goal in developing a HA-aware, COTs architecture and product line is eliminating every source of downtime, planned and unplanned, in both hardware and software. The uptime target is to have no periods of service outage, even though underlying components may fail. The achievable availability objective in the short term is 99.999% (the “five nines” availability), which translates to five minutes of downtime per year.

The goal of a High-Availability system is maximum uptime, but extremely high accuracy is not always a critical design requirement. For example, a video server can occasionally sustain the loss of a frame and is not a problem for users as long as the service itself remains available. The focus must be on building systems that have increased uptime, and that uptime is a critical parameter for telecom systems. The system can suffer some types of faults and recover, and the recovery process may have some impact on the user. A preferable situation is a system that is resistant to software errors, making it more suitable as a platform for delivery of new services.

...only one level of availability is acceptable to end users

100 percent.

Why CompactPCI for Telco

In the new era of multi-service networks and proliferating hardware devices, proprietary technology has become a major obstacle to competitiveness. Proprietary technology decreases agility, limits flexibility in service delivery, and locks telecom companies into limited sources of innovation at relatively high prices. That is why adherence to industry standards and open technology is fast becoming a strategic imperative.

Because CompactPCI is a standard based on many existing standards and has broad support by many popular operating systems, NEPs are able to significantly accelerate their time to market for new and upgraded versions of their systems. Being an industrial-strength standard, CompactPCI is well suited for a wide
In particular, it is ideal for embedded processing applications like system controllers, DSP cards, and custom I/O cards, and can be used in central office applications to deploy additional front-end services to augment an OEM's legacy switch offering.

Telecom OEMs and service providers have embraced CompactPCI technology because it merges the performance, scalability, and reliability of proprietary bus specifications such as VME with the cost-efficiency and flexibility of open standards like PCI. Sun chose the CompactPCI specification as the hardware foundation for HA board products because it is gaining rapid acceptance in the industry and is flexible enough to accommodate a wide range of telecom applications.

Ever at the forefront of innovative computer technology, Sun has invested a significant amount of effort in developing a new HA architecture around CompactPCI technology. Called the CP2000 Program, it is both a High-Availability architecture for CompactPCI and a new generation of CompactPCI board-level products. It is a long-term project focused on putting the power of Sun's highly reliable, highly available server and computer architecture into a CompactPCI form.

With this program, telecom OEMs, NEPs, and enterprise customers can directly leverage Sun’s massive investments and expertise in High-Availability technology at a component level. It makes high-end, data-center-class reliability, accessibility, and scalability (RAS) features available on cost-effective, industry-standard CompactPCI boards. The result is that OEMs and NEPs can substantially reduce the cost and development time of creating specialized, High-Availability systems. By leveraging this off-the-shelf solution, they can focus on providing value-added technologies and services rather than developing basic hardware and software platforms.

**Extending the CompactPCI for High-Availability needs**

For all its strengths, CompactPCI still has room for improvement when it comes to supporting the stringent High-Availability requirements demanded by the Telco industry. That is why Sun has deliberately enhanced the CompactPCI specification to allow for High-Availability implementations.

The current PICMG specification for CompactPCI focuses on only the hardware aspects of achieving High-Availability. The specification allows CompactPCI products to support hot-swap capabilities, but provides no application programming interface (API) or framework for device drivers, middleware, or application software to consistently take advantage of hot swap. The result has been a series of point solutions in which the card vendor writes a high-level device driver or middleware component to handle failover and hot swap, creating a spider’s web of paradigms and interfaces to maintain.

As a leading hardware and operating system vendor, Sun found itself in a unique position to help solve this problem. Sun’s CP2000 program focuses on both the hardware and software aspects of maintaining High Availability in CompactPCI-based systems. Sun developed High-Availability features for CompactPCI and intends to submit the best attributes of its CP2000 architecture for public review and comment through the PICMG.

For example, take the physical connection. This program uses superior materials and extensive testing that goes beyond PICMG specifications. Sun’s CompactPCI HA architecture uses the 220-pin, 2mm “hard metric” connector, ensuring adequate shielding and grounding for low ground bounce and reliable operation in noisy environments. For high-service-ability, these CompactPCI boards use rear-panel connectors, minimizing service time during board replacement. The backplane of this CompactPCI system is entirely passive with no active components. As a result, failed boards can easily be replaced with no impact to service or application availability. In addition, this architecture provides excellent shock and vibration characteristics and delivers low power consumption (~8mA), and support for vertical card orientation provides excellent cooling.

Another weakness in the current hot swap CompactPCI specification is that cards are either on the PCI bus or they are completely disconnected. This makes it difficult to perform diagnostic work without jeopardizing the system integrity. To eliminate this major drawback, the CP2000 includes an intelligent peripheral management interface (IPMI) bus that supports an out-of-band communication link between the system controller and the satellite card. The auxiliary bus functions as a comprehensive debug port without affecting the main PCI bus, significantly improving the overall reliability of the main PCI bus.

Leveraging the capabilities of the IPMI bus even further, the CP2000 architecture adds a second system controller to the PCI bus. This circumvents the problem of having two PCI bus masters by using the IPMI bus to allow the two system controllers to agree on who is mastering the bus. In fact, this activity is so important that the CP2000 architecture introduces a second IPMI bus to perform this operation exclusively. The result is that the CP2000 architecture removes a key single point of failure, the system controller itself. The only other single point of failure is the CompactPCI bus itself. The CP2000 architecture protects this by disabling (tristating) any problem cards. The bus itself is completely passive to promote high MTBF numbers.

**Software considerations and extensions**

Sun’s HA architecture for I/O card hot swap also focuses on extending the software capabilities. Software is the area least addressed by the PICMG, but is the most important in maintaining high availability. In the case of I/O card failures, it is relatively simple to implement a traditional hardware-based hot-swap technique. PICMG has sub-specifications in place defining how to do this. However, implementing a mechanism for handling the failure of an I/O controller or system controller is more problematic. Sun’s solution provides a way to make applications hot-swap aware. Failures are handled seamlessly in software, and failed hardware can be “swapped out”, all while the system remains online.

Should failure occur, the system relies on alternate pathing to automatically redirect disk and network operations to a predefined alternate path and I/O cards can be serviced without systems disruption. Here’s how it works. Each I/O device connects to two I/O controllers with two separate electrical pathways to the I/O device. If one system controller fails, the application is informed and automatically switches to the alternate controller.

Alternate pathing also enables the system to be dynamically reconfigured, while the system is up and running, without a system reboot. For example, an administrator can remove a board from the system, service it and re-insert the board, all without halting the operating system or terminat-
Using CompactPCI to build carrier-grade systems

ing any user applications. Now administrators can perform online repair and reconfiguration of servers, directly increasing application or service-level availability.

Software utilities tightly integrate the system controller with the satellite boards. For example, satellite cards communicate with the system controller over the backplane using a special driver. This driver will enable a variety of programming models to be implemented. One common approach is to plug this low-level driver into the IP stack and treat each satellite card as an IP host. One can think of each satellite card as a separate workstation on a LAN where the LAN is the high-bandwidth, low-latency PCI bus. This architecture also has a mechanism to allow booting across the backplane, such as satellite cards boot from the system controller, and also a comprehensive set of management utilities that are mainly derived from the IPMI bus interfaces.

HA CompactPCI architecture configuration

In a typical CompactPCI system for a Telco based on the CP2000 architecture, the CompactPCI backplane or shelf defines the compute node. It usually has between 1-16 slots. At one slot there is a standalone CPU. At eight slots there is one system controller and seven spare slots for additional CompactPCI cards. Seven slots are the maximum a single system controller can drive. A key feature of CP2000-class cards is that they can be installed in either system controller or satellite slots. Thus it is possible to construct distributed or loosely coupled multi-computers.

If failure occurs, for example in one particular line card such as a T1 card or an Ethernet card, an alternative card located on the same shelf is rapidly switched in and used as its backup. The failover can be very fast since it is within the existing operating system and application image, thereby avoiding application restart. This approach is designed to address line card failures specifically. And that makes a lot of sense when you consider the typical configuration of many of these systems. Quite often the ratio of line cards to system controller cards is between 4:1 and 12:1. This means that any fault is much more likely to occur in a line card than a system controller, especially if one considers down-stream networking issues. With this ratio it makes far more sense to have some number of line-card spares within the shelf, or failover capabilities within a shelf.

The implementation is fairly obvious for I/O cards such as T1/E1 and Ethernet. However, satellite CPU cards (as in a distributed MP machine) create a more complex issue. Sun’s concept is to have specific support within the operating system (often an RTOS) to allow failover between satellite cards. The ChorusOS operating environment, for example, does have support for this. In other cases one might consider special “reliable” programming paradigms or distributed databases. The CP2000 program will provide some functionality for this; however, other operating systems and even the applications themselves may provide specific support.

Telco Ready

Starting with the solid foundation offered by the existing CompactPCI standard, it is possible to create a Telco-ready framework that can meet the High-Availability needs of the demanding telecommunications industry. Using a standards-based approach will enable OEMs and NEPs to quickly construct a highly responsive, flexible infrastructure without compromising on High-Availability requirements. The beauty of this approach is that carrier-grade systems are easily tailored and modified to support new Telco requirements as they emerge in the fast-paced communications arena.

Fred Rehhausser is considered one of the foremost experts in the embedded computer market. Industry analysts credit him with helping to launch Sun Microsystems’ Microelectronics high availability CompactPCI family, positioning Chorus Systems as the Telecom OS, establishing Motorola as the dominant VME boards vendor, building the FORCE COMPUTERS SPARC-based VME boards, and repositioning AMPRO’s line of embedded computers as a macro-component. Rehhausser’s experience includes more than 20 years in senior or executive management positions at Sun Microsystems’ Microelectronics, Chorus Systems, FORCE COMPUTERS, and Motorola’s Computer Group. Rehhausser earned his BSEE from the Massachusetts Institute of Technology (MIT), and a MSE from the University of Pennsylvania. He is the author of numerous papers on bus architecture and bus technology, and holds several patents. For more information, you can contact Fred at: