Incremental innovation: How much more can be designed into CompactPCI?

By Bob Ehlers

Kaizen is the Japanese term that means continuous improvement and is taken from the words kai *meaning continuous and* zen *meaning improvement. Some translate* kai *to mean change and* zen *to mean good, or for the better. In 1995, the PCI Industrial Computer Manufacturers Group (PICMG) introduced CompactPCI, or PICMG 2.x. At the time, it represented a sea change for the embedded computing industry. The form factor became widely adopted, particularly in the telecommunications industry. CompactPCI anticipated and accommodated the leading embedded computing technologies including high-end processors, high-speed bus backplanes, and state-of-the-art power delivery and cooling.*

Over the last nine years, CompactPCI has undergone a continuing advance of point specifications*, which have incrementally improved the architecture, adding features such as hot-swap, redundancy, telephony bus, packet switching on the backplane, and support for telephony mezzanine modules. However, the core specification has remained basically unchanged. The form factor has served well as a foundation for consistent system design and interoperability.*

Unfortunately, the technologies that CompactPCI has leveraged and was originally designed to accommodate have continued to evolve, to the point where the form factor has reached its maximum carrying capacity, as specified. CompactPCI is challenged by processor power and cooling requirements, board density (space), storage, and I/O throughput.

Therein lies the challenge for the industry: Has CompactPCI been taken as far as it can go through incremental improvement? Can CompactPCI be sustained through incremental innovation? Or should CompactPCI be phased out in favor of other form factors that are designed from the beginning to accommodate the technology infrastructure as it exists today or is anticipated to require in the future?

What does the industry want?

Nine years is a long time for any technology. During this time, many companies have adopted their infrastructure, facilities, and staffing to accommodate the CompactPCI based systems they have purchased:

- Central Offices, data centers, switching centers, and industrial automation control rooms are all sized to support the typical 19-inch rack with 6-12U chassis. The space is allocated.
- Power distribution systems are sized to support about 1kW per chassis or 4kW per rack.
- Cooling systems are designed to dissipate the same amount of power.
- Spare inventories exist to keep systems operating.
- Management architectures have been developed.
- Technicians have been trained in system maintenance.
- Special purpose boards have been developed for customer applications.

There is critical mass for CompactPCI, or a derivative thereof that supports backward compatibility and similar physical footprint. CompactPCI enjoys a large installed base, and those who have made these investments would like to see them protected. No one is saying that they would prefer to move to a larger footprint, if the current footprint could be made to support contemporary performance.

Offering a solution that will allow for contemporary levels of performance within the existing system footprint is what the industry really wants, not a forklift upgrade of the entire architecture. The goal is to preserve existing investments while providing a future for performance and scalability. Large cost savings can be realized by supporting incremental innovation to the current architecture to mitigate power, thermal, board real estate, I/O, and storage limitations.

Any architecture changes made to support the smaller form factors for embedded systems like CompactPCI can always be ported up to larger form factors. Porting back down to smaller form factors from large form factors will never be as easy or even possible in many cases. Let us apply *kaizen* to embedded system design.

The challenges

Power input

Developers must address a number of issues relating to CompactPCI thermal and power. The architecture is currently air-cooled. The slot pitch, as well as the board depth and height, are fixed. The backplane connectors are defined with fixed pin counts, wire gauge, and pin mappings.

Manufacturers can employ strategies to increase the power input issues within CompactPCI systems. First, spreading power out across more pins can allow for additional power. Increasing power on higher voltage pins, such as 12 VDC, and then converting this power back to usable voltages on the board is another way to accomplish power delivery.

Making power delivery systems redundant and managed is another issue. Power supplies deliver more power in smaller footprints. Supporting highly available operations is critical, and doing so without doubling the number of supplies required in the system is essential. The power supplies must also be more intelligent about how they behave in a group, as opposed to individual performance. Load sharing, active management of voltage and current levels, notification of impending failure, and cost all need to be considered in designing a power delivery system.

This leaves the issue of how to cool the additional heat generated in the system.

Integration concentrates heat

One phenomenon that system and board designers are faced with is chip-level integration. Greater and greater performance is possible by moving from discrete components to monolithic integrated components, and space requirements are often smaller for a single integrated component than for many discrete components. However, integration concentrates heat in one portion of the board. Concentration of heat creates the challenge of how to move enough air across the component to dissipate its heat, and requires greater levels of airflow.

In embedded system design, there is a point of diminishing returns for chip integration. This point occurs when the concentration of functions, particularly functions such as video graphics and USB, which are appropriate for desktop and enterprise server applications but may be optional in embedded designs, results in a single, monolithic chip. Therefore, the embedded designer has little choice as to board layout for thermal efficiency. Chip manufacturers need to realize that embedded systems may not necessarily benefit from integration of the same feature set as what the enterprise market requires. Pragmatically, it is doubtful that chip manufacturers will act on this, as it is generally the embedded market that leverages the enterprise market for parts to achieve economies of scale. Certain chip solutions are better than others. These will be discussed a little later.

Specialization of function: Compartmentalization

Some functions, routinely considered part of the integrated feature set of an embedded CPU board, can be spread more evenly across board and system designs. For instance, certain system management architectures allow chassis-level hardware management to occur in small, specialized intelligent shelf management modules. System designers have great flexibility when placing these modules in a system. Moving the system management function from a large size single board computer to a smaller footprint module frees single board computers to be dedicated to application processing. Specialization of hardware can yield more even thermal distribution in a system.

Spot cooling

Even following compartmentalization, with heat intensive activities spread more evenly throughout the board and chassis, we are still left with large, integrated, power hungry chips. The semiconductor industry paradigm is to shrink dies, add transistors (increase density), increase package size, and increase clock frequencies to boost CPU performance. The CPU market targets volume market requirements, which are largely driven by the desktop and enterprise server market. Desktop or server space is not at much of a premium. Chassis can easily support delivering and cooling 450 or more watts, which is more than sufficient for most high-end processors and supporting chipsets.

Additionally, as clock frequencies for onboard signaling increase, part placement must be closer together, further concentrating heat in one location of the board. If the embedded market wants to keep up with the performance curve of the desktop and server market, new cooling paradigms are needed. Parts must be placed closer together in order to reduce packet/data/electron transit times. Note that this actually drives system designers towards tighter board layouts, which goes against the conventional thinking that we should be increasing board size and spreading components out in order to get better cooling. Solving the cooling problem in the *Eurocard* form factor used by CompactPCI is actually a better solution than moving to a larger form factor because of timing limitations.

The solution to thermal issues in embedded systems is to look beyond air cooling as the only way to dissipate heat. Cooling can be delivered in a small footprint to meet the needs of current and future embedded system components.

Board space

The CompactPCI specification uses the Eurocard form factor, which calls out a 6U high board that is 160 mm deep and has a 0.8-inch slot pitch. At the time of its inception, this board space was sufficient for the CPU and supporting chipsets for the processors of the day and even looking forward. But over the last nine years, more and more functionality has crept into designs. Most jumps in performance have brought with them incremental space requirement increases. Increases from 10 Mbps to 100 Mbps to 1 Gbps Ethernet have driven larger MAC and PHY components. The jump from Intel Pentium II to Pentium III to Pentium 4 drove larger processor packages. The jump from 32/33 to 64/66 to PCI-X has driven larger bridge chips. Increases in video processor memory and processor capacity have driven larger chips.

Conversely, at the same time, chipsets have consolidated from many discrete components to single, monolithic components. In many cases this has reduced total space requirements. These integrations also result in an overall component cost reduction.

When designing an embedded architecture, each individual feature needs to be examined for placement. Is the feature best implemented as part of an integrated chipset solution, or is it best moved to another location in the architecture? What is the cost of moving the feature to a different location? Can the transport required to support the feature be extended to the new location (bandwidth and timing)?

There are several key functions and features that can be moved off the base CPU PCB, but doing so requires significant planning in architecture design.

I/O throughput

Bandwidth is on a continual march towards greater heights. Ethernet has moved from 2 Mbps to 10 Mbps to 100 Mbps to 1 Gbps to 10 Gbps. PCI has moved from 32/66 to 64/66 to 64/133. Hublink, Hypertransport, RIO, ATA, and others have all experienced similar performance gains. These performance gains have exponentially improved device-to-device communications and simultaneously have increased the demands made of source and target devices in processing data transmitted over these various architectures. As devices receive and send data at higher rates of speed, they are burdened with processing the protocols used to transmit this data. Error checking, addressing, data typing, encryption, and other processes all make demands of the processor. All of this occurs before the CPU can even process the data.

System designers need strategies for reducing the burden placed on the CPU in utilizing these higher data rates. These strategies, when implemented as part of an architecture that considers such factors as space, heat, and power delivery, can allow exponential performance improvements by reducing CPU burden and dispersing processing more evenly in the chassis.

CPU selection

Many attributes make a good CPU for any particular application. What works for a desktop or enterprise server may not necessarily be best for embedded systems. However, as convergence occurs between computing and telecommunications and as applications become embedded into the communications network, a base level of similarity in features among desktop, enterprise, data center, and embedded telecommunications computing environments must exist. On a typical PICMG 2.16 board IP packets move from the backplane through the bridges to the CPU for stack processing (see Figure 1).

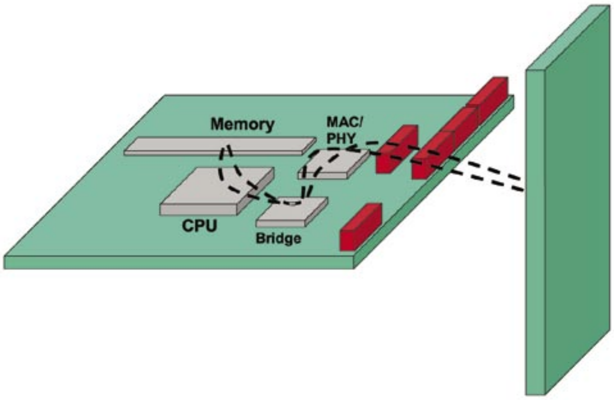

Figure 1

Integrated application development environments move applications from the developer's desk to a lab test environment to a production environment. Recognition of this when designing an embedded computing architecture, which is generally the target production environment, can ultimately:

- Reduce the cost of application development.
- Improve determinism.
- Enhance reliability.
- Simplify management.

Selecting a CPU that can benefit from *compartmentalization* is also important. CPUs that have a limited range of supporting chipset feature options or that are overly integrated can result in a large, immovable mass. Sometimes this is the best solution, sometimes it is not. If the features sought for a specific application can be closely mapped to a highly integrated CPU and support chipset, then this may be the best direction. If application requirements are broad, then this approach generally does not work well. The trick is to look for applications that can benefit from integration, moving to a system on a chip where appropriate.

One key distinction between desktop or development environments and production environments is the need for user interface. Chips like the Intel P4 are designed to support integrated graphics, and such user peripheral devices as USB, printer ports, keyboards, and mouse. Many of these features have become integrated into the CPU or support chipsets. Marketing managers for CPU product lines have driven the creation of benchmarks, marketing messaging, and collateral, which stress the performance of various CPUs for user interface functions (video refresh rates, vector draw). Many customers ask for *performance* without really understanding what it is that they are seeking. Many CPUs with *lower performance,* as portrayed though these marketing message filters, actually have better performance characteristics for embedded applications than their *higher performance* desktop or enterprise counterparts.

CPU onboard bus architectures and future scalability are also key selection criteria. Creating a nonblocking I/O path from the processor to support chips and processes to the network is essential. Balancing bandwidth between the CPU, bridges, memory, mass storage devices, peripheral devices, and the network is more of an art than a science, with balance being the operative concept. Processor capacity must be balanced with the sum of I/O pipes feeding the processor.

Finally, the function of a CPU board in embedded design has transitioned from being primarily a control processor to being an application processor. Control is now more of a distributed function residing either on intelligent peripheral boards or in special purpose devices such as intelligent system managers. Application processor characteristics significantly differ from those of a control processor. Applications are more demanding from a memory and compute throughput perspective. Floating-point performance becomes more critical, while integer performance continues to scale up linearly. In the architecture shown in Figure 2, IP packets move from the backplane to the mezzanine, where they are fully processed and delivered directly to the CPU as data.

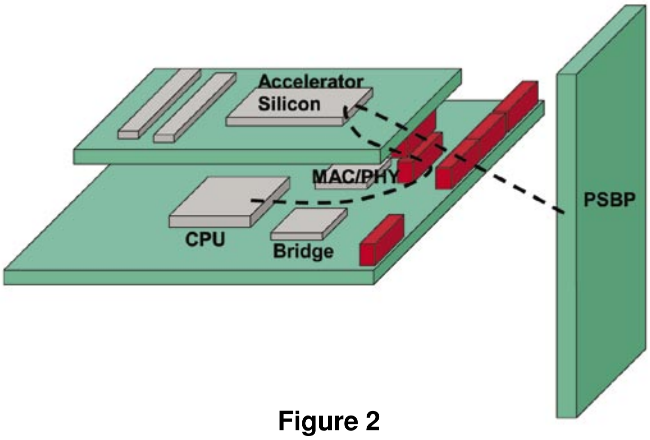

Figure 2

Storage

The slot pitch, board size, and thermal limitations of embedded systems have significantly limited the amount of mass storage that can be incorporated into a system. Generally speaking, disk drives have been limited to 2.5-inch drives, taken from the laptop/mobile component market. These devices are low density, slow, and not very reliable when compared to other larger drive form factors.

When a company wanted to implement mass storage in an embedded system, it generally has meant either incorporating external storage through a PMC host bus adapter attached to an external disk array, which can be very expensive and wastes space. Or, the customer accepts the attributes of 2.5-inch drives and builds in redundancy to offset reliability and capacity limitations. In either case, a lot of money is spent.

Additionally, using onboard storage on a CPU board adds to space, thermal, and power design requirements. Moving storage off the CPU board can have great benefits in processing density, costs, and reliability. Once again, an architecture that is thoughtfully implemented can address storage requirements through compartmentalization.

An iSCSI host bus adapter functionality allows for mass storage abstraction. Once storage is abstracted from the application processor, it can be managed through the intelligent shelf manager, acting as a load balancer.

An example:

1. A CPU board is inserted into a vacant system slot.
2. The intelligent shelf manager identifies the type of board it is and examines overall system loading and functional requirements.
3. Based on a system management policy stored in the intelligent shelf manager, the CPU board is given an iSCSI target boot partition.
4. The CPU board then boots up based on the application configuration on the iSCSI target, joins the load balancing group, and is ready to run.

This is just one example of the functionality that can be achieved through incremental innovation in system management.

Conclusion

Innovation can come in many forms, from sea change to incremental improvements. Often the most influential innovations are derived from continuous incremental improvements, rather than monumental changes. Innovations such as the implementation of iSCSI, TCP/IP, IPsec offload, Intelligent Platform Management Interface management, and closed system cooling solutions all incrementally contribute to a vastly superior system architecture with plenty of headroom for today and tomorrow.

***Bob Ehlers** is Performance Technologies' director of business development. He has worked for the company for three years in product management and marketing. Previously, Bob was vice president of operations for HipBone, Inc., an Internet collaboration hosted service provider, and vice president of business services at Inktomi Corporation. He is a graduate of California Polytechnic State University with a BS in Business Administration and also holds a professional certificate in Telecommunications Management from UC Berkeley.*

For further information, contact Bob at:

Performance Technologies
1050 Southwood Drive
San Luis Obispo, CA 93401
Tel: 805-783-6153
Fax: 805-541-5088
E-mail: bob.ehlers@pt.com
Website: www.pt.com