White Paper

Next-Generation Server Technology: The Key to Speed, Productivity and Reliability

NGST-WP101-x

This white paper reveals the little-known fact that in a server, it's the core logic and not the CPU that does most of the work and establishes the system's capabilities. Armed with this inside track on how to exploit advances in server technology to achieve a next-generation network, you can deploy systems that run faster, respond quicker, work in denser environments, consume less power and generate less heat than the systems you are using today.

April 15, 2005

Introduction

Although central processors get all the glory, the core logic in servers does most of the work and plays a key role in shaping a system's more visible attributes. Contemporary servers need the ability to absorb and process data that streams through broadband networks at gigabit speeds. Workloads like this demand the utmost in memory bandwidth and input/output capacity. Unlike contemporary desktop systems that loaf during the day and sleep at night, servers must operate on a 24/7 schedule, with no timeouts for rest and relaxation. It goes without saying that the suppliers of these systems must do all they can to ensure data integrity and continuous, trouble-free operation. The bigger the server, and the more critical its mission, the more it needs these RAS (Reliability, Availability and Serviceability) features. Since 1997, System I/O™ core logic from Broadcom's ServerWorks™ unit has played a key role in helping system suppliers meet these demanding requirements.

A server's core logic determines the number of processors the system can accommodate. It defines the bandwidth, capacity and reliability of the memory subsystem, along with the number and speed of the input/output cards in the system and the aggregate I/O bandwidth of the system. The core logic's architecture plays a key role in determining the overall performance of the server. Poorly designed core logic can throttle system performance.

Architecture 101 for Server Core Logic

To satisfy these demanding requirements, core logic chipsets typically contain two distinct types of components. The first, commonly referred to as the North Bridge, links the central processor(s) to the main memory system. Since Intel®-based servers allow up to four processors to share a common bus, the ***North Bridge*** must control this shared bus so the traffic flows in a fair and efficient manner. The North Bridge also serves as the system's DRAM controller and exerts a great influence on overall system operation, given the impact main memory latency plays in determining processor performance.

The second type of component, commonly referred to as the ***South Bridge*** or ***I/O Bridge***, links the North Bridge to the I/O buses in the system. South Bridges often contain support for a variety of low-performance and/or legacy I/O devices as well. In order to support high aggregate rates of I/O, system designers often want to include several I/O Bridges in a single system. To support this feature, the North Bridge must provide a separate link for each of several South or I/O bridges. This approach facilitates the implementation of specialized I/O Bridges that incorporate networking or storage interfaces, and thus improves overall performance. The manner in which these bridges communicate with one another can have a profound effect on system expandability and throughput.

Figure 1 illustrates the architecture of the Broadcom Grand Champion™ GC-LE System I/O™ core logic chipset, one of the most widely used chipsets in the Intel®-based server market. Let's take a quick tour of this configuration to understand how the core logic accomplishes its tasks. The large chip in the middle, labeled "CMIC-LE," serves as the North Bridge. The arrow entering the CMIC-LE at the top represents an eight-byte-wide processor front-side bus (FSB) that operates at speeds of 400 MHz or 533 MHz, and transfers data at a peak rate of 4.2 Gbps. In order to feed the processors at this peak rate, the memory system must deliver up to 4.2 Gbps of data each second. Fortunately (but it's no accident), the dual banks of DDR266 memory (266 MHz Double Data Rate DRAM, sometimes referred to as "PC2100") can deliver an aggregate rate of 4.2 Gbps. If you had only one bank of memory, or if it ran at a lower speed (like 200 MHz), the processor would be constrained by memory system bandwidth, and system performance would suffer.

To support this dual-bank mode of operation, memory modules must be installed in pairs; inexpensive servers might used dual 256 MB DIMMs for a total of 512 MB, while higher-end boxes might use up to eight 2 GByte DIMMs for a total of 16 GBytes, more than most Intel® Xeon™ processors want or need today.

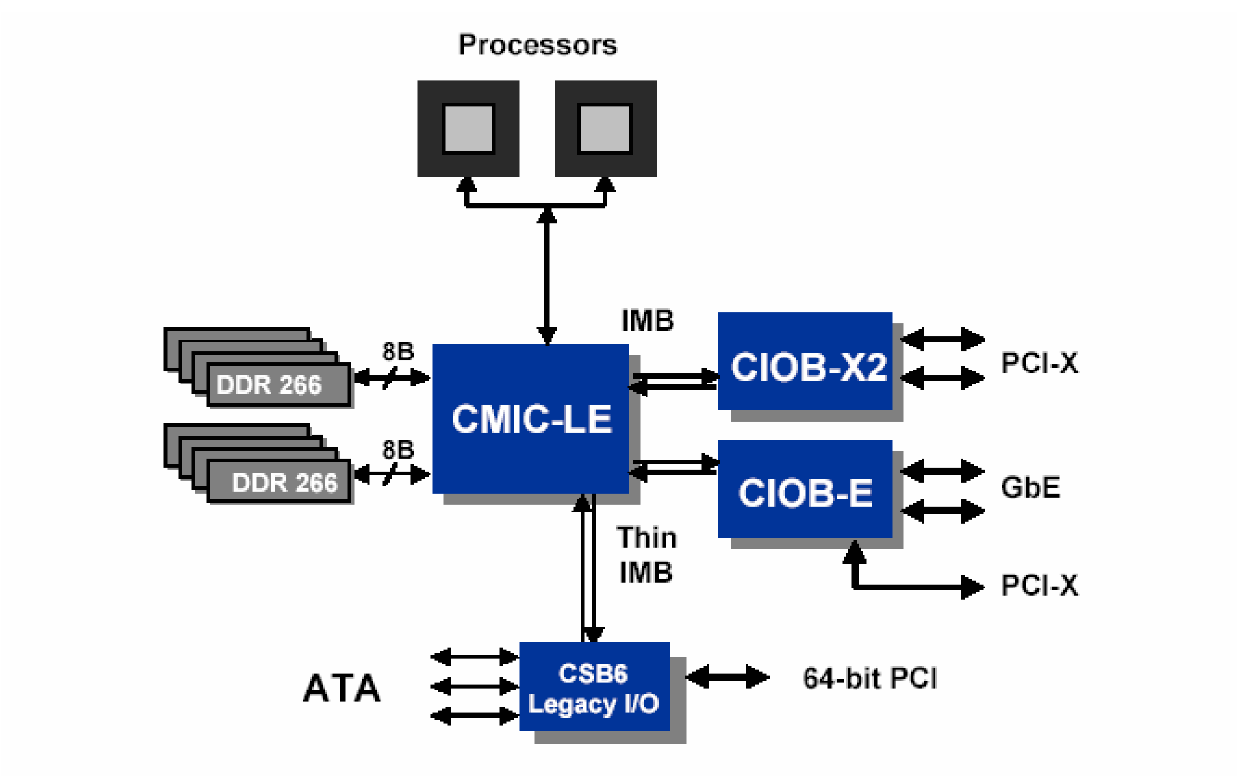

Figure 1: Broadcom Grand Champion™ LE System I/O™ = Core Logic

The memory controller in the CMIC-LE supports standard functions like ECC (Error Correcting Code) that detect and correct minor errors in the physical memory. The CMIC-LE also supports more advanced features like memory scrubbing, which corrects small errors before they become big ones, and even more exotic features like “Chipkill™,” which allows the system to recover if an entire memory chip selfdestructs, and “memory sparing,” which allows the system to switch out an entire bank of memory and replace it with one held in reserve for just such an emergency. Features like these previously existed only in mainframes, but now can be found in systems priced under $2,000.

Off to the right of the CMIC-LE sit two I/O bridges, labeled “CIOB-X2” and “CIOBE.” The CIOB-X2 (Champion™ I/O Bridge for PCI-X™) controls two PCI-X bus controllers, each operating at speeds up to 1 Gbps. (A “new improved” version, based on the recently approved PCI-X 2.0 specification, operates at speeds for 2 or 4 Gbps, and is coming soon to a server near you.) When the devices on these PCI-X buses operate at full speed, the CIOB-X2 must move data through the North Bridge at rates up to 2 Gbps. Fortunately (but again, not by accident), the Inter-Module Bus (IMB) that links the CIOB-X2 to the CMIC-LE can transfer up to 3.2 Gbps, so there’s bandwidth to spare. (But be careful; some competitive chipsets can’t match the IMB’s bandwidth, and find their performance constrained by the path between the North Bridge and the I/O bridges.) The CIOB-E (Champion™ I/O Ethernet Bridge) includes support for two integrated Gigabit Ethernet (GbE) controllers, along with another PCI-X bus. It may be the only device in the world that combines all these features in a single chip. Integrating GbE support into the core logic reduces system cost and saves power (we’ll have more to say about that later). The CIOB-E’s Ethernet controllers contain features that optimize system performance and manageability. Each Ethernet Media Access Control (MAC) unit includes two high-speed RISC processor cores that can be utilized for advanced packet filtering and CPU offloading. The CIOB-E’s MAC follows the same programming model as the discrete Broadcom® BCM5704 GbE controller, and thus software stacks developed for the discrete controller can be utilized without modification with the CIOB-E, and vice versa. The CIOB-E contains two integrated PHYs for standard 10/100Mb Ethernet cabling, but system designers can also use external SERDES (Serializer/Deserializer) chips to run the Ethernet signals over the backplanes often used in blade-type configurations. This is really cool stuff!

The CSB6 (Champion™ South Bridge, version 6) hangs off the bottom of the CMIC-LE. Its primary task is to handle so-called legacy functions like booting the system and handling multi-processor I/O interrupts, but it also provides support for a 64-bit PCI bus, four USB ports and three ATA disk drive interfaces.

Broadcom offers system suppliers a variety of North Bridges and I/O Bridges that address a wide range of price/performance and capacity alternatives. The system suppliers, in turn, arrange these components in a TinkerToy® manner, to create small, medium and large Intel®-based servers. Broadcom’s “secret sauce,” the key to this modularity, is its Inter-Module Bus, or IMB. The IMB supports unidirectional data transfers at speeds up to 1.6 Gbps and bidirectional transfers at speeds up to 3.2 Gbps. Broadcom’s scalable IMB technology allows the company to evolve its products in an orderly manner, and protects the investments system suppliers and end users make in advanced server technology.

Industry Standards Drive the Intel®-Based Server Market

Over the past five years, the Intel®-based portion of the server market has grown far more rapidly than segments based on alternative processor architectures. The superior price/performance of Intel®-based server solutions, compared with alternative platforms, has driven much of this growth. Industry standards play a key role in enabling these price/performance advantages. Industry standards allow component suppliers to compete for the most effective solutions, and permit system suppliers and end users to select best-of-breed components with a high degree of confidence that the elements they select will operate properly in the system. Component suppliers, platform suppliers and end users must all remain in sync with one another in order to maintain this momentum. Over the past five years, Broadcom's ServerWorks™ unit has demonstrated its ability to lead the industry with regard to the evolution of memory technology standards (from 66 MHz and 133 MHz SDRAM to 200 MHz and 266 MHz DDR) and I/O bus technology standards (from PCI to PCI-X to PCI-X 2.0). The industry has continually validated Broadcom's technology selections. This record suggests that end users can be assured that systems based on Broadcom technology will have long and useful lives, and are unlikely to encounter technological dead-ends.

Who is Using Broadcom's Core Logic?

It's unlikely that an IT professional like you would ever pick up the phone and purchase core logic chipsets directly from Broadcom, unless you wanted a few to display on your desk. System suppliers design Broadcom core logic into a wide range of end-user products, but rarely tout the ServerWorks™ brand in their promotional activities. (You can usually find this information buried deep in their data sheets if you look hard enough.) To make it easier for you to find systems that use Broadcom technology, we've included in Appendix A of this paper a partial list of currently-available systems that use Broadcom's core logic. Not surprisingly, the list is quite long. After all, ServerWorks™ has shipped more than 10 million core logic chipsets over the past five years. An impressive statistic, especially given that the Intel®-based server market just recently achieved a shipment rate of five million systems per year.

If you scan the table in the appendix, you will see a wide range of two-way and four-way systems from established server suppliers like Dell, HP and IBM. You might be surprised to see that leading storage suppliers like EMC and Network Appliance have also adopted Broadcom technology. SAN and NAS systems consume tremendous amounts of I/O bandwidth as they move data back and forth between high-performance storage devices and high-performance network interfaces. Broadcom's core logic delivers the high I/O performance these storage system suppliers need to make their products fly.

Many of the systems in the appendix are packaged as dual-processor systems in 1U enclosures or blade configurations, a “scale out” approach often referred to as “dense computing” that is especially well suited to edge-of-network applications. There is a reason why suppliers prefer Broadcom technology for these configurations. Broadcom’s platforms with integrated Gigabit Ethernet controllers typically consume 20 watts less power than systems based on alternate technologies. That may not sound like much, but when you pack 42 of these 1U boxes into a rack, that’s almost a kilowatt of power you don’t need to supply and heat you don’t need to remove from your data center.

ServerWorks – A Subsidiary of Broadcom Corporation

Many observers were puzzled a few years ago when Broadcom, a company that specializes in semiconductors for high-bandwidth communications, acquired ServerWorks, a supplier to the computational server market. Since that time, the reasons for this acquisition have become clear, as computing and communications technology have continued to converge. At a recent conference Paul Otellini, Intel’s president, observed, “Soon every communication device will compute. And every computing device will communicate.” In order to implement this convergence, organizations need access to intellectual property in both computing and communication spheres. The acquisition gave ServerWorks the access it needed to Broadcom’s industry-leading Ethernet technology, and gave Broadcom access to ServerWorks’ high-performance input/output architecture, now known as System I/O™. The Ethernet capability incorporated in the previously discussed CIOB-E marks the first demonstration of the advantages of this new organizational arrangement.

What’s an IT Professional to Do?

As an IT professional involved in server selection and deployment, you face a number of challenges. Odds are, the traffic entering your data centers is growing by leaps and bounds, and the processing to handle the transactions embedded in this data is growing due to application complexity and security concerns. Unfortunately, your budget for equipment to handle these new requirements hasn’t increased as fast as the workload, if it’s increased at all. You want to take advantage of the enhanced performance of the latest Intel® server technology, but you can’t afford to throw out everything in your shop and start all over again. Fortunately, there is a solution to these problems.

Contemporary servers based on Intel® Xeon™ technology and Broadcom’s System I/O™ core logic offer dramatically higher levels of performance than systems you acquired only a few years ago, but they maintain compatibility with the software systems and I/O devices you added back then. These traditions of backward compatibility and planned evolution play a central role in directing the product strategies of Broadcom. They allow end users of equipment based on Broadcom technology to incrementally expand and upgrade their installed base of equipment without fear of obsolescence.

This applies whether they need to upgrade from 10/100Mb Ethernet to GbE, or from Pentium II® Xeon™ to the latest Xeon™ technology. All in all, they enable end users to exploit advances in server technology in order to deploy systems that run faster, respond more quickly, operate more reliably, are easier to service, work in denser environments, consume less power and generate less heat than the systems they are using today.

Appendix A: Some Systems that Use Broadcom Core Logic

Supplier	Model	Broadcom Core Logic	Intel® Processor	Notes
Dell	PowerEdge 1650™	HE-SL	PentiumIII (DP)®	
	PowerEdge 1655MC™	III-LE	PentiumIII (DP)®	Blade
	PowerEdge 1750™	GC-LE	Xeon DP™	
	PowerEdge 2650™	GC-LE	Xeon DP™	
	PowerEdge 4600™	GC-HE	Xeon DP™	
	PowerEdge 600SC™	GC-SL	Pentium 4®	
	PowerEdge 6600™	GC-HE	Xeon MP™	
	PowerEdge 6650™	GC-HE	Xeon MP™	
	PowerEdge 1600SC™	GC-SL	Xeon DP™	
	PowerEdge 650™	GC-SL	Pentium 4®	
	PowerVault 725N™	GC-SL	Pentium 4®	Storage
	PowerVault 775N™	GC-LE	Xeon DP™	Storage
EMC	CLARiiON	GC-LE	Xeon DP™	Storage
Fujitsu-Siemens	PRIMERGY BX300	III-LE	PentiumIII (DP)®	Blade
	PRIMERGY BX600	GC-LE	Xeon MP™	Blade
	PRIMERGY RX100	GC-SL	Pentium 4®	
	PRIMERGY RX300	GC-LE	Xeon DP™	
	PRIMERGY RX600	GC-LE	Xeon MP™	
	PRIMERGY TX150	GC-SL	Pentium 4®	
	PRIMERGY TX200	GC-LE	Xeon DP™	
	PRIMERGY TX300	GC-LE	Xeon DP™	
	PRIMERGY TX300	GC-LE	Xeon MP™	
HP	ProLiant™ BL10e	III-LE	PentiumIII®	Blade
	ProLiant™ BL10e G2	GC-SL	Pentium-M®	Blade
	ProLiant™ BL20p	HE-SL	PentiumIII (DP)®	Blade

Supplier	Model	Broadcom Core Logic	Intel® Processor	Notes
HP (contd)	ProLiant™ BL20p G2	GC-LE	Xeon DP™	Blade
	ProLiant™ BL40p	GC-LE	Xeon MP™	Blade
	TC 2120	GC-SL	Pentium 4®	
	ProLiant™ DL320 G2	GC-SL	Pentium 4®	
	ProLiant™ DL360 G3	GC-LE	Xeon DP™	
	ProLiant™ DL380 G3	GC-LE	Xeon DP™	
	ProLiant™ DL560	GC-LE	Xeon MP™	
	ProLiant™ DL580 G2	GC-HE	Xeon MP™	
	ProLiant™ DL310	GC-SL	Pentium 4®	
	ProLiant™ DL330 G3	GC-SL	Xeon DP™	
	ProLiant™ DL350 G3	GC-LE	Xeon DP™	
	ProLiant™ DL370 G3	GC-LE	Xeon DP™	
	ProLiant™ DL530 G2	GC-HE	Xeon DP™	
	ProLiant™ DL570 G2	GC-HE	Xeon DP™	
IBM	eServer® X225	GC-LE	Xeon DP™	
	eServer® X235	GC-LE	Xeon DP™	
	eServer® X255	GC-HE	Xeon MP™	
	eServer® X305	GC-SL	Pentium 4®	
	eServer® X335	GC-LE	Xeon DP™	
	eServer® X343	HE-SL	PentiumIII (DP)®	
	eServer® X345	GC-LE	Xeon DP™	
	xSeries BladeCenter HS20	GC-LE	Xeon DP™	Blade

Supplier	Model	Broadcom Core Logic	Intel® Processor	Notes
NEC	Express 5800 Model 120Ef	GC-SL	Xeon DP™	
	Express 5800 Model 120Mf	GC-LE	Xeon DP™	
	Express 5800 Model 140Rc-4	GC-HE	Xeon MP™	
	Express 5800 Model 320Lb	GC-LE	Xeon DP™	
	Express 5800 Model 340Ha	GC-LE	Xeon MP™	
Network Appliance	F800	III-HE	Xeon MP™	Storage
	FAS900	GC-HE	Xeon MP™	Storage
Stratus	Stratus® ftServer® 3300	GC-LE	Xeon DP™	
	Stratus® ftServer® 5600	GC-LE	Xeon DP™	
	Stratus® ftServer® 6600	GC-LE	Xeon MP™	

Phone: 949-450-8700
Fax: 949-450-8710
E-mail: info@broadcom.com
Web: www.broadcom.com

BROADCOM CORPORATION
16215 Alton Parkway, P.O. Box 57013
Irvine, California 92619-7013

NGST-WP101-R 04/15/05