insideHPC

The insideHPC Guide to

Flexible HPC

BROUGHT TO YOU BY TYAN

Executive Summary

Organizations that implement high-performance computing (HPC) technologies have a wide range of requirements. From small manufacturing suppliers to national research institutions, using significant computing technologies is critical to creating innovative products and leading-edge research. No two HPC installations are the same. For maximum return, budget, software requirements, performance and customization all must be considered before installing and operating a successful environment.

Challenges in HPC Environments

While all users of HPC technology want the fastest performance available, price and power consumption always seem to come into play, whether in the initial planning or at a later time. Standard performance measures exist that may or may not relate to an end user's application mix, but it is important to understand the various benchmark results that go into determining the performance of a CPU, a server or an overall cluster.

The acknowledged leader in providing enterprise-grade, high-performance CPUs is Intel®. Although other companies do provide instruction-compatible CPUs for data center computing, the leader in both market share and technology is Intel. When looking at CPU performance, a number of factors must be considered. A measure of CPU performance, although not the most important one, is the clock rate. Basically, the faster the clock rate of the chip, the faster an application will run. The increase in clock rates on a specific line of CPUs marched forward at a predictable and amazing rate until about 2005, when clock rates leveled out at about 3GHz. As the clock rates on popular processors began to level off, the increase in the number of independent cores on a single chip began to increase. At first, two cores per chip appeared, quickly followed by quad-core designs. The technology has advanced at such a rapid rate that today's most modern CPUs contain more than 20 cores per chip. The absolute performance of a system must account for the clock rate, instructions executed per clock tick, the number cores and the number of sockets in a single server.

Contents

Requirements in HPC Environments

Organizations that require HPC technologies will have a wide range of requirements, depending on their workloads. While some applications can run hundreds or even thousands of cores, other applications cannot take advantage of more than one core. Various programming techniques can be used for applications to take advantage of more than one core. Within a server that is running a single OS, OpenMP directives can be used to distribute different threads to different cores. However, when an application requires more cores than are available in a single server, the Message Passing Interface (MPI) API is used, and can distribute and keep track of various parts of the workloads on different servers. Then, within the server, either OpenMP or the MPI API can be used.

Single-socket systems can be used for applications that are not easily able to run and take advantage of multi-core systems. Dual-socket systems are typically the most popular because the core count is reasonable for many applications (up to 44 cores per system), but there still aren't too many. Four-socket systems are ideal for applications that can scale significantly yet still require a single OS installation. Eight-socket systems are available for the most demanding applications, but do come with a cost overhead. One-, two-, four- and eight-socket systems basically perform the same tasks, yet in a more parallel environment the amount of memory available using similar memory dual in-line memory modules (DIMMS) can push the decision to a larger system. With a larger number of sockets per CPU board, more memory slots will be available, thus increasing the size of the data that an application can address.

With a larger number of sockets per CPU board, more memory slots will be available, thus increasing the size of the data that an application can address.

Many HPC installations will need a variety of servers. With intelligent resource management software, a combination of one-, two-, four- and eight-socket systems can be put to maximum use. In a large environment where multiple jobs are running simultaneously, some low in parallelism and some high in parallelism will be best-suited to installing a variety of server capabilities.

Recently, a new class of computing hardware has become available that can greatly accelerate certain applications. While main CPUs now top out (as of May 2016) at 22 cores per socket, accelerators can contain more than 100 cores per system. The Intel ®Xeon Phi ™coprocessor is an example of this new type of hardware, where applications that are highly parallel and contain algorithms that can be spread over hundreds of threads can take advantage of. While not typical, some benchmarks show more than 100 times improvement in the performance of an application when using an accelerator. For very compute-intensive applications, it may be critical to house more than one of these accelerators in a single system. This would give an application access to between 200 and 300 computing elements in a single enclosure and a single operating system. A tight coupling of the CPUs with the coprocessor can enable significant acceleration of certain classes of applications.

For very compute-intensive applications, it may be critical to house more than one of these accelerators in a single system. This would give an application access to between 200 and 300 computing elements in a single enclosure and a single operating system.

Industries

A number of industries require HPC technologies, to either bring products to market faster or model and simulate physical phenomena.

1 BIO-INFORMATICS, LIFE SCIENCES and COMPUTATIONAL CHEMISTRY

For these industries, the advancement in computational power has enabled companies to offer new drugs and medicines by simulating how certain components interact with other components. The computational power available in a single server containing multiple CPUs and accelerators brings personalized medicine closer than ever before. This is definitely an industry that can take advantage of the wide range of hardware that is available today.

2 ELECTRONIC DESIGN AUTOMATION

Designing the latest computer processors and other microelectronic devices requires significant computing power. While not an overly parallel application, millions of simulations must be run in order to verify the inner workings of a new CPU or related device. Large clusters of low core-count systems would be ideal for this industry so that runs can all be done in parallel, thus reducing the chance for expensive hardware bugs. Bringing a new product to market faster can increase profits and result in a market-leading product family.

3 FINANCIAL SERVICES

The ability to make faster decisions with more data can directly increase the profits of a trading organization. With the latest systems, more simulations based on previous data and trends can be run, increasing the confidence in the outcomes. Trading decisions can be made milliseconds faster than a competitor whose systems (clock speed) may be older, leading to a competitive advantage.

While some critical applications that a large manufacturing company will run are not highly parallel, a large cluster with varying types of systems allows for multiple design teams to share computer resources. This results in a more optimized product and faster time to market.

4 MANUFACTURING

Large-scale computer simulations have been driving the competitive edge for manufacturing companies for many years. For accurate simulations that involve both structural and computational fluid dynamics, they are ideally run in large clusters with varying numbers of cores. Recent advances in commercial computational fluid dynamics (CFD) codes can take advantage of hundreds of cores to produce more accurate and more complete simulations. While some critical applications that a large manufacturing company will run are not highly parallel, a large cluster with varying types of systems allows for multiple design teams to share computer resources. This results in a more optimized product and faster time to market.

5 DATA MINING, ANALYTICS and DATABASES

Although not traditional HPC, the use of significant computing power for industries that must sift through terabytes of data is required. High-performing CPUs are needed, as is a lot of high-speed main memory to hold the massive amounts of data. Some of these tasks can be run on a cluster using a variety of computing hardware.

6 WEATHER, CLIMATE MODELING and ATMOSPHERIC RESEARCH

HPC systems are used every day to simulate near-term weather events or long-range climate forecasts. Finer meshes with more physics simulated can lead to more accurate day-to-day or even hour-to-hour weather forecasts. Long-range climate modeling also benefits from a large amount of computing power. In addition, weather and climate simulation codes can take advantage of thousands of cores and are highly scalable. Many of these applications can also use the power of accelerators to increase performance significantly.

7 OIL, GAS and PETROLEUM EXPLORATION

Making correct and accurate decisions when deciding where to drill for oil and gas can save millions of dollars. HPC technologies can aid in determining where to drill and can simulate the best methods for oil and gas extraction. Simulating these processes can take advantage of all the computing power that an organization can budget for. Large clusters of a wide range of hardware can be effectively used to look for valuable deposits, which reduces the expense at the front end of the exploration pipeline.

8 SCIENTIFIC RESEARCH

Many disciplines that are being researched at universities, national labs and in the research departments of leading companies rely on highly optimized applications and a wide range of computer systems. Together, large clusters can be created that allow researchers to simulate many processes or search for new answers. Many of these applications are highly parallel and are custom-designed. These applications can take advantage of the latest in CPU architectures and heterogeneous environments. Tweaking the system to perform at its maximum performance for a specific application requires extreme knowledge of the application, as well as a deep understanding of the CPU's strengths and weaknesses.

TYAN® Solutions

TYAN is a leading provider of Intel® Xeon®-based systems for a wide range of industries. When it comes to providing the component computing solutions that are required for a range of customers, TYAN's long-time expertise in this area is critical for a successful implementation. End users can become more productive faster by working with suppliers that offer a wide range of servers and associated hardware.

TYAN offers a wide range of systems with a variety of CPU sockets and a range of storage and network options. This is especially important, as every end-user environment has different requirements that must be met.

All of the TYAN Intel Xeon server solutions are optimized to run applications extremely fast. Using Intel Xeon CPUs designed onto a board that offers a significant number of memory slots provides users with the performance they expect. From one-socket to four-socket solutions, TYAN offers systems that incorporate the Intel Xeon CPUs as well as the Intel Xeon Phi coprocessors. TYAN offers a wide range of systems with a variety of CPU sockets and a range of storage and network options. This is especially important, as every end-user environment has different requirements that must be met.

For the most demanding and cost-sensitive environments, the Intel® Xeon® E5-2600 v4 series of processors emerges as the leader. The latest processors from Intel, nicknamed "Broadwell-EP," are included in all varieties in TYAN servers. The most popular are those systems based on the E5-2600 v4 processors. The Intel Xeon E5-2600 v4 builds upon leading-edge performance that users have come to expect. Numerous world records have been achieved from business-oriented benchmarks to computationally intensive ones, at all configurations.

The Intel Xeon E5-2600 v4 processors can be used in two-socket systems, with a maximum addressability of 3TB memory (by deploying 128GB 3DS LRDIMMS). As the typical workhorse of an HPC data center, up to 44 cores are available for computation (two sockets of 22 cores each). TYAN has designed these systems so that a two-socket server can address up to 24 DIMMs in a server. By populating 64GB DIMMs in all slots, 1.5TB memory can be addressed in a compact form factor.

A four-socket system is ideal for those that require maximum performance, and whose applications must address a significant amount of memory.

For the most demanding HPC users whose applications require a single operating system for parallelization, TYAN can deliver:

- a four-socket server that contains the Intel® Xeon® processor E7-8800 v4, or
- Intel® Xeon® processor E7- 4800 v4 CPUs

A four-socket system can address up to 6TB of memory, together with using 64 (Intel Xeon 4800 v4) to 96 (Intel Xeon E7-8800 v4) cores. This system is ideal for those that require maximum performance, and whose applications must address a significant amount of memory.

The most demanding application requires high core counts, maximum amounts of memory, and fast and persistent storage. As solid state disks (SSDs) become more widely used in computing systems, it is important for vendors to partner with leading suppliers to offer a wide range of SSDs. Since SSDs have no moving parts and are extremely high-performance as compared with hard disk drives (HDDs), modern HPC servers must be designed to incorporate these devices, in similar form factors as HDDs. The Intel ®SSD product line delivers the highest performance available, while maintaining enterprise-grade reliability. In addition, various Intel SSD DC products can scale in performance as computing and storage needs grow.

All of the aforementioned systems can incorporate from one to four Intel Xeon Phi coprocessors. The combination of high core count, high memory configurations and number of coprocessors gives both developers and users a supercomputer in a rack-mount enclosure. With more than one Intel Xeon Phi coprocessor installed in a server, applications that are able to take advantage of this amount of computing power will truly shine. However, it is important to make sure that the application has a significant amount of parallelism built in, and is designed to take advantage of the single instruction, multiple data architecture.

At the very high end of the server offerings from TYAN is a true supercomputer in 4U. The FT76-B7922 system can:

- hold up to four of the Intel Xeon CPUs
- has room for up to four of the Intel Xeon Phi coprocessors

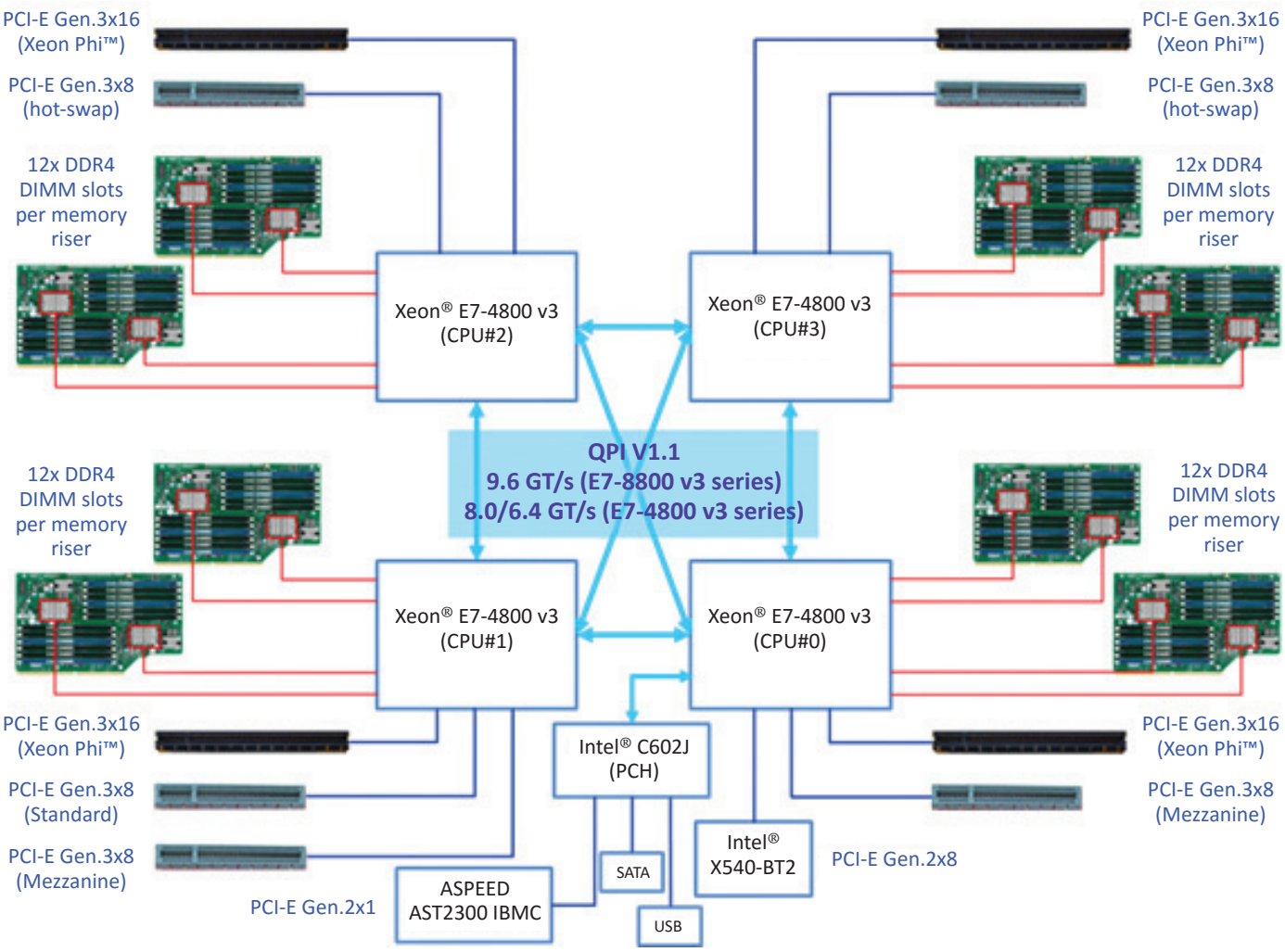
With all of this power in a single server, application tuning is critical, or the machine can be set up to run multiple applications concurrently. TYAN's FT76-B7922 provides for highly parallelized application deployment in various HPC segments, such as graphic virtualization, big data analysis and cloud computing.

The FT76-B7922 is the first multi-purpose server platform that simultaneously supports scale-up (fat node) and scale-out (many-core CPU node).

The FT76-B7922 is the first multi-purpose server platform that simultaneously supports scale-up (fat node) and scale-out (many-core CPU node). The platform accommodates four Intel® Xeon® processor E7-8800/4800 v4 CPUs and four Intel Xeon Phi 7120P (5110P/3120P) cards, or four GPUs within the enclosure for typical HPC applications. With 1:1 CPU to Intel Xeon Phi coprocessor ratio or 1:1 CPU to GPU ratio, TYAN's FT76-B7922 provides a high price/performance and performance/watt for HPC community that needs CPU- and GPU-intensive computing workloads.

The quad-socket Intel Xeon E7-8800/4800 v4 platform allows TYAN to offer a very large memory footprint. It gives customers a maximum memory pool that can reach 6TB of RAM when 96x 64GB DDR4 DIMMs are deployed. Such a large amount of memory allows the use of several high-performance applications, including HPC visualization and in-memory computing. Entire databases can be loaded into RAM, allowing very rapid access and manipulation of very large datasets. The following figure is a block diagram of the FT76-B7922.

TYAN has a complete product line that addresses any corporation or research organization's requirement. The combination of various Intel Xeon CPUs, Intel Xeon Phi coprocessors, and wide variety of memory choices and storage options make TYAN's one of the most complete and flexible lineups in the industry. Depending on the requirements and the budget of the end-user organization, a wide array of servers can be obtained, and when included as part of a cluster, can deliver the highest performance.



Successful Implementations

A successful implementation at customer sites requires that the supplier does more than deliver a defined set of hardware. Suppliers must be able to understand the needs of the customer, even when these needs are unique.

RANGE OF PRODUCTS

As discussed previously, customers may require a range of servers and storage solutions. TYAN has a wide range of products that can fit anyone's needs, whether the environment requires single-, dual- or quad-socket servers.

LATEST TECHNOLOGY

CPU, memory and storage suppliers continuously improve their products and offerings. It is critical to customers that their server vendor keeps pace with the latest offerings and incorporates these new and faster products quickly into their product line. TYAN accomplishes this by having superb relationships with its suppliers.

CUSTOMIZATION

For some customers, the ability for the supplier to customize offerings for specific environments or workloads can be critical to a successful implementation. Technically TYAN can customize BIOS, management firmware, enclosure look and feel, even board-level functions per requirement. However, the final scope of customization depends on project schedule and budget. TYAN specializes in working with customers on all aspects of the implementation phase to ensure that the systems and the clusters are working optimally.

CLOUD SERVICE PROVIDER

In addition to supplying the latest hardware systems to customers, TYAN is a leader in developing optimized servers for organizations that are creating a cloud infrastructure. These systems are designed for very high performance with low power requirements and can act as a dedicated server for a specific application or as a larger system to run different applications simultaneously.

Summary

TYAN has developed a reputation over the past 25 years for delivering state-of-the-art servers and board-level products that deliver exceptional performance. Using the Intel Xeon processors and the Intel Xeon Phi coprocessors, TYAN has been able to create a wide range of products that can be customized and delivered directly to the customer. With many years of experience and the best technology that is available today, TYAN responds to the most demanding requirements for HPC customers. For organizations that need the latest in technology and a strong and able supplier, TYAN works with customers to accomplish this.

Contact Information

USA

3288 Laurelview Court, Fremont, CA 94538
Tel: +1-510-651-8868
Pre-Sales: +1-510-651-8868 x5120
Fax: +1-510-651-7688
Email: Sales_america@tyan.com

TAIWAN

5F, Builiding B, No. 209, Sec. 1, Nan Gang Rd.,
Nan Gang Dist.,Taipei 11568, Taiwan
Tel: +886-2-2652-5888
Fax: +886-2-2652-5805
Email: Tsung.yi@mic.com.tw
Sales_Taiwan@mic.com.tw