IT@Intel White Paper
Intel IT
High-Performance Computing
May 2013

High-Performance Computing For Silicon Design

- Four generations of HPC have successfully enabled Intel® silicon tapeout, reducing tapeout time from 25 to less than 10 days.
- Intel has taped out several silicon products with HPC-1 alone, delivering a return on investment of USD 44.72 million.[1]

Shesha Krishnapura
Senior Principal Engineer, Intel IT

Ty Tang
Senior Principal Engineer, Intel IT

Vipul Lal
Senior Principal Engineer, Intel IT

Raju Nallapa
Principal Engineer, Intel IT

Doug Austin
Storage and Backup Architect, Intel IT

Shaji Achuthan
Senior Staff Engineer, Intel IT

Executive Overview

Designing Intel microprocessors is extremely compute intensive. Tapeout is a final step in silicon design and its computation demand is growing continuously for each generation of silicon process technology. Intel IT adopted high-performance computing (HPC) to address this very large computational scale and realized significant improvements in computing performance, reliability, and cost.

We treated the HPC environment as a holistic computing capability—ensuring all key components were well designed, integrated, and operationally balanced with no bottlenecks. We designed our HPC model to scale to meet future needs, with HPC generations aligned with successive generations of Intel® process technology.

The first-generation HPC environment (HPC-1), supporting 45nm processor tapeout, included innovative approaches and technologies to increase scalability, such as the following:

- A parallel storage system providing 10x scalability compared with our previous system based on traditional file servers, together with high-speed backup.
- Large-memory compute servers based on a unique modular non-uniform memory access (NUMA) design, offering significant cost advantages.
- Batch compute servers based on multi-core Intel® Xeon® processors, offering substantial performance increases.
- Optimization of our license server and job scheduler to handle thousands of simultaneous design jobs.

HPC-1 successfully enabled 45nm processor tapeout, delivering net present value (NPV) of USD 44.72 million to Intel. We subsequently developed three new generations of HPC environment (HPC-2, HPC-3, and HPC-4), with further scalability increases to support the tapeout of 32nm, 22nm, and 14nm processors, respectively.

Since deployment, our HPC environment has supported a 29.87x increase in compute demand, with a 20x increase in stability. In addition, tapeout time was reduced from 25 days for the first 65nm process technology-based microprocessor in a non-HPC compute environment to 10 days for the first 45nm process technology-based microprocessor in an HPC-enabled environment. The success of the HPC environment was due to factors such as careful alignment of technology with business needs, informed risk taking, and disciplined execution. We are continuing to develop the future HPC generations to enable tapeout of successive generations of Intel® processors.

Contents

IT@INTEL

The IT@Intel program connects IT professionals around the world with their peers inside our organization – sharing lessons learned, methods and strategies. Our goal is simple: Share Intel IT best practices that create business value and make IT a competitive advantage. Visit us today at www.intel.com/IT or contact your local Intel representative if you'd like to learn more.

BUSINESS CHALLENGE

Microprocessor design is extraordinarily complex—and as a result, requires huge amounts of computing capacity. About 50,000 of the servers in Intel's worldwide environment are dedicated to silicon design.

Each new generation of process technology—such as the transition from 65nm to 45nm processors—brings a substantial increase in complexity, requiring a major increase in design compute performance.

Though increased performance is needed across the entire design process, the requirement is particularly acute at the highly compute-intensive tapeout stage.

Tapeout is a process where Intel chip design meets manufacturing. As shown in Figure 1, it is the last major step in the chain of processes leading to the manufacture of the masks used to make microprocessors.

During tapein, the stage that immediately precedes tapeout, Intel chip design teams create multi-gigabyte hierarchical layout databases specifying the design to be manufactured. During tapeout, these layout databases are processed using electronic design automation (EDA) tools. These tools apply extremely compute-intensive resolution enhancement techniques (RET) to update layout data for mask manufacturability and verify the data for compliance to mask manufacturing rules.

A key EDA application within the tapeout stage is optical proximity correction (OPC), which makes it possible to create circuitry that contains components far smaller than the wavelength of light directed at the mask. OPC is a complex, compute-bound process. To accelerate the process, OPC applications take advantage of distributed parallel processing; tasks are divided into thousands of smaller jobs that run on large server clusters.

It is critical to complete tapeout as fast as possible—and to minimize errors—because delays at this stage can mean slipped project deadlines and even a missed market window.

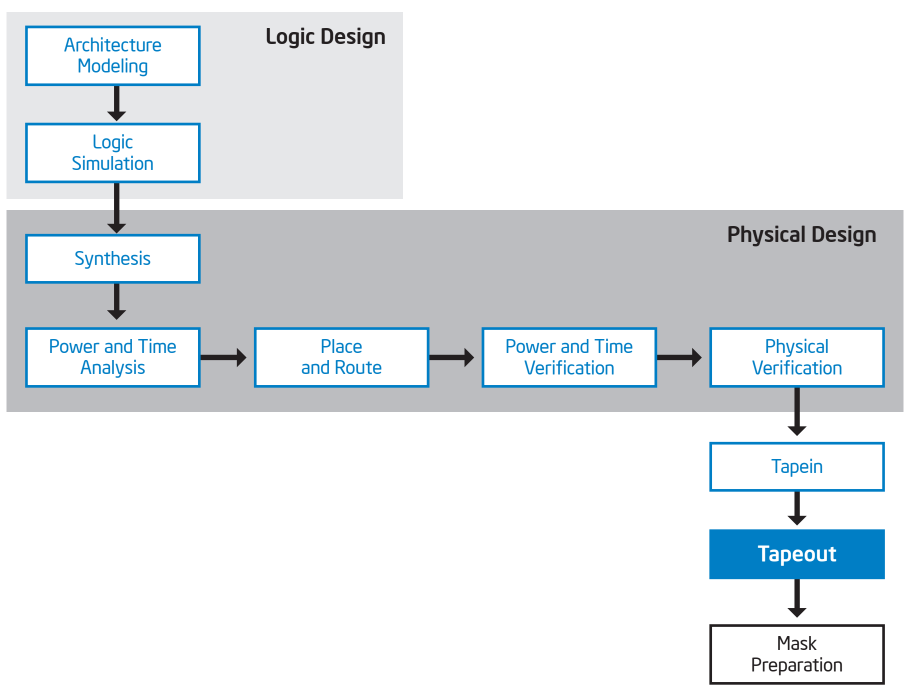


Figure 1. The phases of silicon design.

Tapeout Challenges

Up to and including the 65nm process technology generation, tapeout computing was managed as an extension of our general-purpose design computing environment.

However, as we prepared for the transition to the first Intel® 45nm processors, it became apparent that we needed a new approach to create a cost-effective, reliable, and predictable environment capable of supporting the increased demands of 45nm processor tapeout.

Overall, we anticipated that we would need a 10-fold increase in scalability. Key challenges included:

- Storage. We anticipated a requirement for a 10x increase in storage system throughput. However, our existing production network-attached storage (NAS) file servers were already experiencing I/O bottlenecks even before the transition to 45nm technology.
- Compute servers. The compute servers used to run the largest tapeout jobs could not support the anticipated 4x increase in physical memory requirements.
- Stability. Our existing production environment was not designed to support very large-scale tapeout computing. Because of this, it was less reliable than desired, leading to more than 20 tapeout delays per quarter.
- Cost. We needed to solve these technical challenges while meeting the requirement to reduce capital expenditure by USD 20 million.

We expected this growth trend to continue in future process generations. This meant we needed an approach that could both support 45nm tapeout and subsequently scale to meet future needs.

To solve these challenges, we set out to develop a high-performance computing (HPC) environment optimized for tapeout processing, using large compute server clusters and disruptive technologies to deliver substantial increases in scalability and performance.

SOLUTION: HIGH-PERFORMANCE COMPUTING STRATEGY

In 2005, we created an HPC strategic program to develop a highly scalable and reliable tapeout compute environment that is capable of delivering optimal results. Developing our HPC environment presented significant challenges because this was the first time HPC was attempted for semiconductor design.

Strategic objectives included the following:

- Leverage industry and internal expertise to architect a leading-edge HPC environment
- Design a solution that is highly customized for tapeout
- Use open standards to develop an agile environment
- Regularly benchmark and adopt best-in-class HPC technology

Our immediate goal was to enable the tapeout of the first Intel 45nm processors to meet our committed deadline to Intel product groups.

Our longer-term objective was to develop an HPC generational model that could meet future needs, aligned in lockstep with successive generations of Intel® process technology, as shown in Figure 2. Each HPC generation would provide a major increase in capacity to support the demands of the corresponding new processor generation.

For the first generation of the HPC environment (HPC-1), our goal was to achieve an overall 10x increase in scalability.

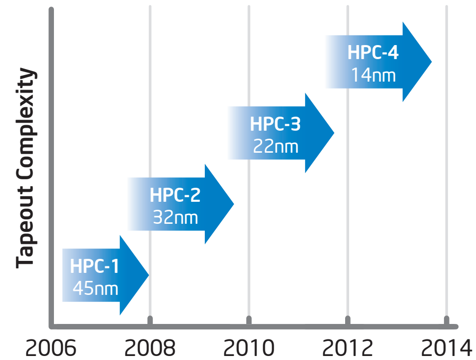


Figure 2. We aligned our high-performance computing (HPC) environment with process technology generations.

Our approach was to treat the HPC environment as a holistic computing capability—ensuring that critical components were well-designed, integrated, and operationally balanced with no single bottleneck. These components were the following:

- Storage and backup
- Compute servers
- Network
- Batch clustering and job scheduling
- Application license servers
- Enterprise Linux* OS
- Application and platform tuning

The solution stack that delivers our HPC environment is shown in Figure 3.

We assessed the performance of each component using real tapeout workloads. We identified bottlenecks and the improvements needed in each area. Then members of the HPC program and Intel's manufacturing group jointly defined the HPC-1 specifications.

We have continued this approach with subsequent HPC generations to achieve the increases in scalability required for successive processor generations.

Beginning in 2007 we designed and implemented the second, third, and fourth generations of our HPC environment to provide the increased compute resources required to support tapeout of 32nm, 22nm, and 14nm processors.

We have made substantial improvements in the key components, outlined in the following sections.

Storage and Backup

We identified storage performance and scalability as significant bottlenecks. We implemented a parallel storage system to deliver the anticipated 10x increase in required scalability in HPC-1; by HPC-4, scalability has increased 27x. We combined this with a faster backup solution capable of handling the required throughput and much larger disk volumes.

PARALLEL STORAGE

For the 65nm processor generation, we had been using traditional NAS file servers, which were able to serve only 400 distributed clients and had a 400 GB volume size limit.

For the 45nm generation, we needed to support up to at least 4,000 clients—a 10x increase—and volume sizes up to 3 TB. To achieve this with the existing solution would have required at least 10 additional storage server racks. This was not an option because of the resulting increases in our data center footprint as well as power and cooling costs. An additional problem was that the need to replicate large design datasets across multiple storage servers to work around scalability limitations affected the productivity of our design engineers.

We therefore decided to research parallel storage solutions that would not only satisfy our current storage needs but also easily scale to future demands. The storage solution needed to deliver higher performance with a significantly lower total cost of ownership (TCO).

We considered more than 20 possible solutions and selected one after an extensive evaluation, including onsite testing with real tapeout workloads that consumed more than 1 million CPU hours.

The deployment of our parallel storage solution was a milestone; it was a pioneering use of parallel storage in an IT organization in the semiconductor industry.

Parallel storage specifications

Our parallel storage system is based on an array of blade servers, each powered by an Intel® CPU and including one or two hard drives, RAM, and a UNIX*-like OS kernel. Most blades are used to store data; storage capacity can be increased or decreased by adding or removing a blade. The system also includes blades that provide metadata services.

For HPC-1, our system consisted of 110 blades—100 storage blades and 10 metadata blades—interfacing through gigabit Ethernet (GbE) with a total uplink bandwidth of 40 gigabits per second (Gb/s).

In successive HPC generations we upgraded the parallel storage system to provide even greater performance and

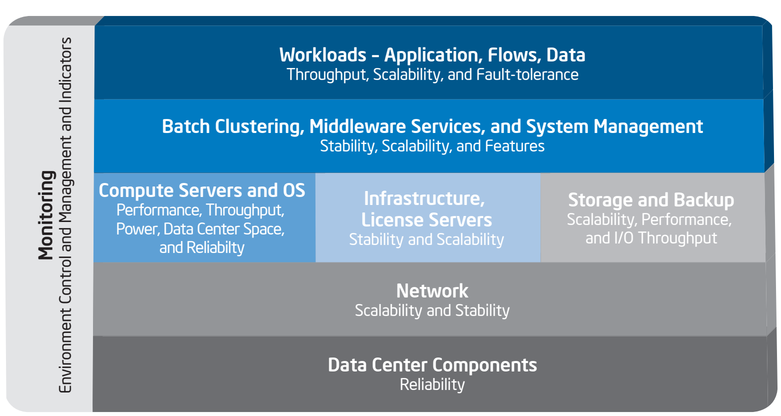


Figure 3. The high-performance computing (HPC) solution stack is well-designed, integrated, and operationally balanced.

scalability. In the upgraded system, each blade included a faster Intel® processor as well as more RAM and the latest hard drive storage technology.

Specifications of the HPC-1, HPC-2, HPC-3, and HPC-4 parallel storage blades are summarized in Table 1.

Parallel storage advantages

The parallel storage system has delivered major advantages over our previous file servers.

- **Scalability.** In HPC-1 we were able to substitute one parallel server for every 10 conventional storage servers, as shown in Figure 4. This 10:1 consolidation ratio translated into huge cost savings because of reduced space requirements and energy consumption.
- **Performance.** For specific portions of the workflow, we saw up to a 300-percent performance improvement in HPC-1 compared to the previous storage solution.
- **Volume size.** The maximum volume size increased by a factor of 16, from 400 GB to 6.4 TB, easily supporting our requirement for 3 TB-plus volumes in HPC-1.

For the 14nm processor generation, the HPC-4 parallel storage environment has scaled to support up to 40,000 distributed client CPU cores and up to 10 TB volume sizes.

Table 1. Parallel Storage System Specifications for HPC-1 through HPC-4

Component	HPC-1	HPC-2	HPC-3	HPC-4
Storage Blade CPU Specification	Intel® Celeron® processor 1.2 GHz, 256 KB L2 cache	Intel® Celeron® processor M370 1.5 GHz, 1 MB L2 cache	Intel® Celeron® processor M370 1.5 GHz, 1 MB L2 cache	Intel® Xeon® processor LC3518 1.73 GHz, 256 KB L2, 2 MB L3 cache
Chipset	Intel® 440GX Chipset	Intel® 3100 Chipset	Intel® 3100 Chipset	Intel® BD3420 PCH
Bus	100 MHz FSB	400 MHz FSB	400 MHz FSB	DMI 2.5 GT/s
RAM	512 MB	2 GB	4 GB	8 GB
RAM Type	PC 100 SDRAM	ECC DDR2-400	ECC DDR2-400	ECC DDR3-800
Hard Drives	2x SATA 3.0 Gb/s, 400 GB, 7200 RPM, 8 MB cache	2x SATA 3.0 Gb/s, 500 GB, 7200 RPM, 16 MB cache	1x SATA 3.0 Gb/s, 1 TB, 7200 RPM, 32 MB cache	2x SATA 3.0 Gb/s, 2 TB, 7200 RPM, 64 MB cache
Raw Storage System Capacity	80 TB	100 TB	93 TB	180 TB

HPC-1 First-generation high-performance computing environment; **HPC-2** Second-generation high-performance computing environment; **HPC-3** Third-generation high-performance computing environment; **HPC-4** Fourth-generation high-performance computing environment

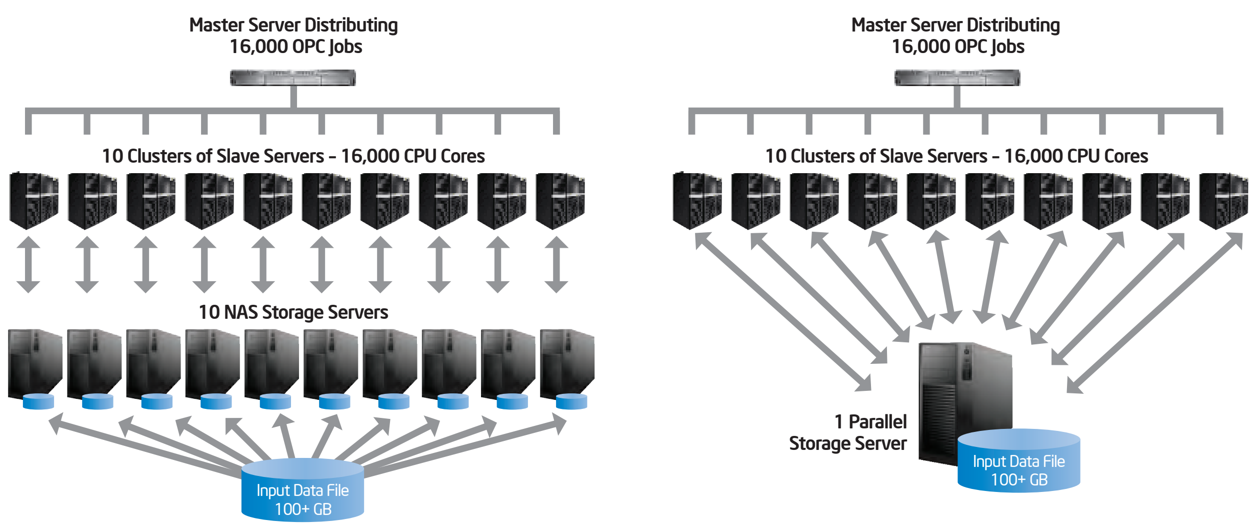


Figure 4. Consolidation with the high-performance computing (HPC) parallel storage environment.

BACKUP

The HPC-1 requirements greatly exceeded the capabilities of our previous backup infrastructure. HPC-1 included disk volumes larger than 3 TB; to meet our service-level agreement, we needed to complete backup of these volumes within 24 hours. This required a single-stream throughput of at least 35 MB/s.

At the time, this requirement was challenging because few available tape drives offered this level of performance. However, we identified an upcoming product that offered 120 MB/s raw performance per tape drive. After verifying performance, we coupled two of these drives with media servers running Linux*, which enabled us to more easily use them with the parallel storage system.

When combined with the parallel storage system, this setup delivered aggregate read throughput of more than 200 MB/s. As a result, we were able to support 3 TB volumes without compromising our backup, archive, and restore service levels.

We have been able to continue scaling our HPC backup solution to support up to 10 TB volume sizes today.

Compute Servers

Our tapeout environment includes thousands of servers that support highly compute-intensive applications. The increased demands of 45nm tapeout and beyond presented significant challenges in the following areas:

LARGE-MEMORY COMPUTE SYSTEMS

The largest tapeout jobs, such as design rule check (DRC) workloads, require servers with a very large RAM capacity. We also use these large-memory servers as master servers for distributed OPC applications.

The maximum certified memory capacity of servers in our pre-HPC tapeout environment was 128 GB. However, we knew that the increased complexity of 45nm processors would result in tapeout jobs that required up to 4x this memory capacity.

Moving to a higher-end system based on our existing architecture to support large memory capacity would have increased costs significantly. We therefore set a goal of implementing a system based on a modular architecture that could scale to meet future needs while meeting our aggressive cost objectives.

We identified a unique modular system based on non-uniform memory access (NUMA) architecture, capable of accommodating up to 32 Intel® Xeon® processors and 512 GB of RAM.

While this system provided the scalability we needed, the situation also created new challenges. There wasn't a Linux OS optimized for NUMA platforms, and neither the server nor the EDA applications were qualified for use in our environment.

We took a two-step approach: We first focused on deploying a 256 GB configuration to enable tapeout of the first 45nm processor, followed by a larger 512 GB system for tapeout of subsequent high-volume 45nm processors.

256 GB SOLUTION

Our initial objective was to create a system based on four nodes, each with four processors and 64 GB of RAM, and compare performance with the previous solution. The architecture is shown in Figure 5.

This required close collaboration with the suppliers of the server hardware and the OS. We formed a joint system enablement team and worked intensively with a pre-release

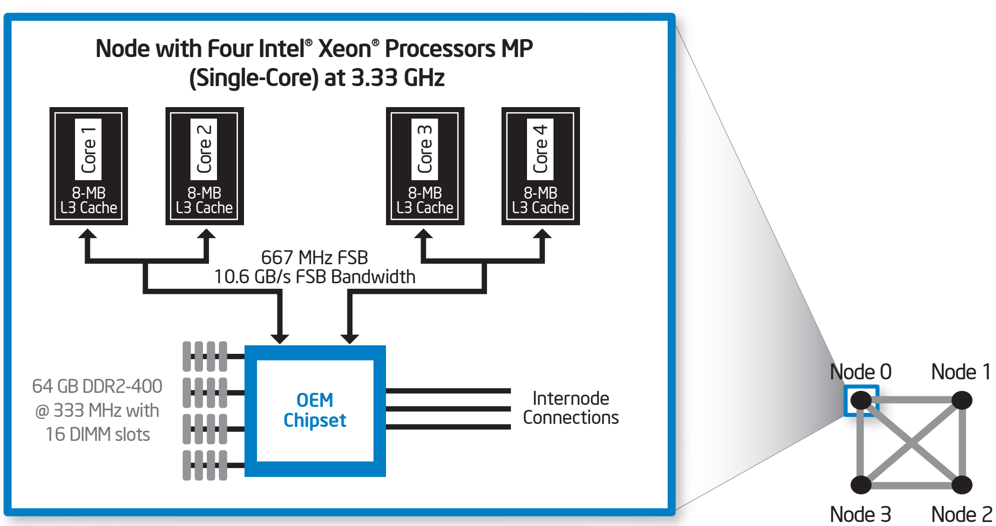


Figure 5. First-generation high-performance computing environment (HPC-1) large-memory system. Left: One node based on four single-core processors with 64 GB of RAM. Right: How four nodes interconnect to create a 256 GB system.

version of the OS to help ensure that it ran effectively on the system. We also worked with the OS supplier to conduct numerous performance and reliability tests.

As a next step, we worked closely with the EDA supplier to certify its memory-intensive DRC application on the new platform. Our efforts to resolve critical functionality, reliability, and performance issues achieved a remarkable result: We deployed the production system on the same day that the OS release was officially launched.

The new system successfully delivered substantial performance improvements and the ability to run bigger workloads. Large workloads ran 79 percent faster, compared with the previous server architecture.

512 GB SOLUTION

Our objective was to enable an eight-node system with 32 CPUs and up to 512 GB of RAM, analyze the scalability and stability, and qualify the system in time to support tapeout of high-volume 45nm processors. We connected eight of the nodes illustrated in Figure 5; the interconnectivity is shown in Figure 6.

We evaluated this system when running DRC workloads consuming up to 512 GB of RAM. We tested multiple workloads in a variety of configurations, including single and multiple concurrent workloads using local and network file systems. We found that the system was able to scale to run these workloads with no performance degradation.

HPC-1 LARGE-MEMORY COMPUTE SERVER REFRESH

When Intel® Xeon® processor 7100 series was released, with two cores per processor, we adopted these processors as standard. The overall system architecture remained the same, but each individual node now was equipped with additional cores and a larger L3 cache. An individual node is shown in Figure 7.

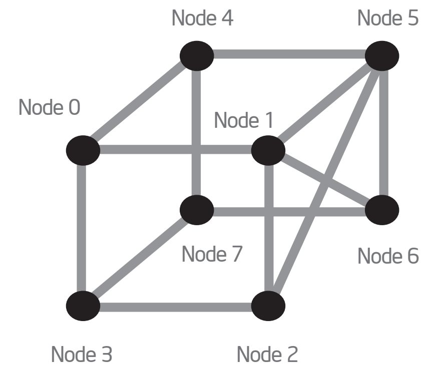


Figure 6. Interconnectivity for first-generation high-performance computing environment (HPC-1) large-memory compute server with eight nodes and 512 GB of RAM.

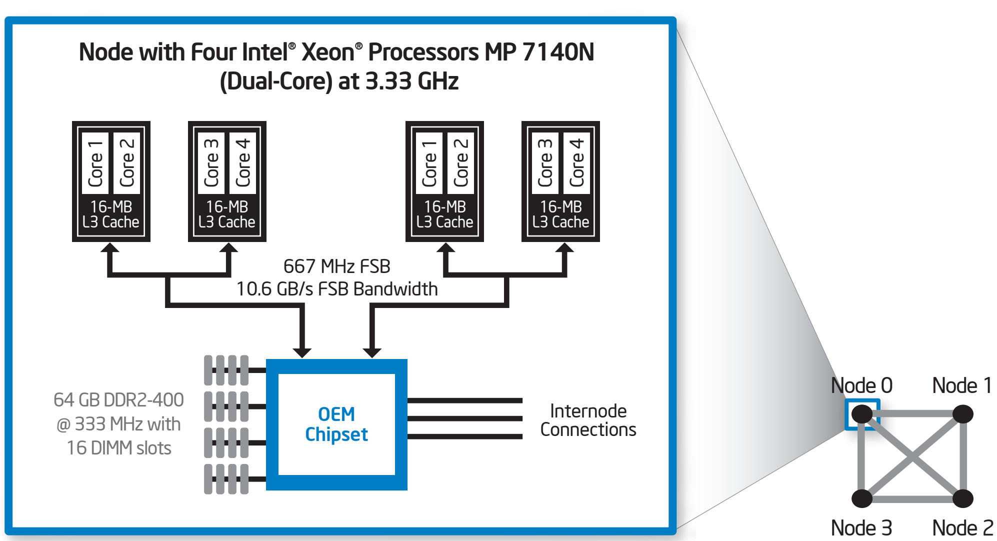


Figure 7. First-generation high-performance computing environment (HPC-1) large-memory refresh server. Left: One node based on four dual-core processors with 64 GB of RAM. Right: How four nodes interconnect to create a 256 GB system.

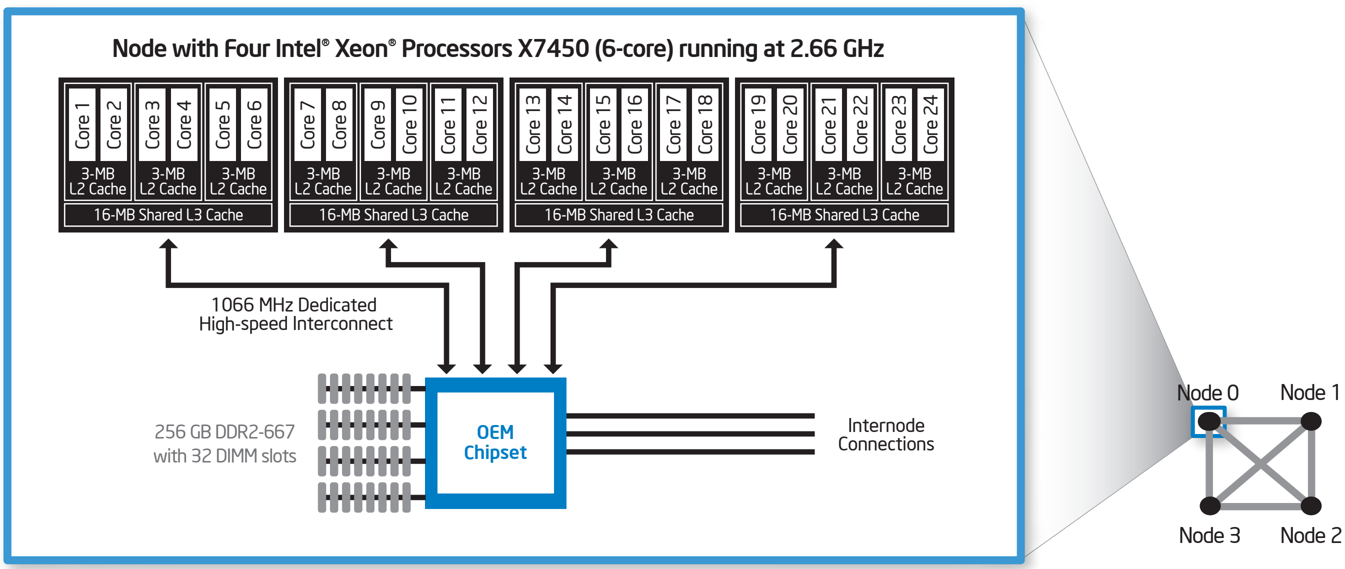


Figure 8. Node in a second-generation high-performance computing environment (HPC-2) large-memory compute server. Left: One node based on four 6-core processors with 256 GB of RAM. Right: How four nodes interconnect to create a 1 TB system.

HPC-2 LARGE-MEMORY COMPUTE SERVER WITH 1 TB OF RAM

For HPC-2, we took advantage of the introduction of the 45nm Intel® Xeon® processor 7400 series, with six cores per processor, to create a 96-core system with 1 TB of RAM. This consists of a four-node cluster in which each node has 256 GB of RAM and 24 processor cores. The architecture is shown in Figure 8.

HPC-3 AND HPC-4 LARGE-MEMORY COMPUTE SERVER WITH 2 TB OF RAM

For HPC-3, we started by connecting two nodes of the Intel® Xeon® processor 7500 series with 8 cores per processor and 1 TB of RAM per node to create 64 cores per system with 2 TB of RAM. Later, we took advantage of the newer Intel® processor® E7-8800 product family 10-core processor platforms to create 80 cores per system with 2 TB of RAM for high volume HPC-3 and HPC-4. For HPC-4, we adopted Intel® Solid-State Drives to enable fast swap on the Intel processor E7-8800 product family large-memory compute servers. The two architectures are shown in Figures 9 and 10, respectively.

All four generations of the large-memory compute servers are compared in Table 2.

BATCH COMPUTE SERVERS

Compute-intensive tapeout jobs such as OPC are handled by large clusters of batch compute servers operating in parallel in a master-slave configuration. To illustrate the scale of the challenge, there may be as many as 40,000 OPC jobs executing concurrently on thousands of servers.

We achieved major performance improvements by taking advantage of multi-core Intel Xeon processors as they became available. Our pre-HPC environment relied on single-core processors, but we subsequently moved to dual-core and then quad-core, six-core, and eight-core processors.

Our tapeout workload results provided real-world proof of a key theoretical advantage of multi-core processors: that performance scales with the number of cores within an HPC cluster.

Servers based on Intel Xeon processors with four cores showed a consistent ability to run twice as many jobs as servers with prior generation dual-core processors and delivered faster runtimes with a relative throughput of 4.8x compared to older generation single-core processors. The relative throughput scaling has increased to 19.4x by HPC-4 with eight-core processors.

The performance benefits achieved with faster Intel Xeon processor-based batch compute servers in HPC-1 translated directly into a reduction in data center space and energy requirements.

As new Intel server processors are released, we have continued to incorporate servers based on these processors into our environment. This delivers continuing increases in performance for key applications such as OPC and simulation, as shown in Figure 11.

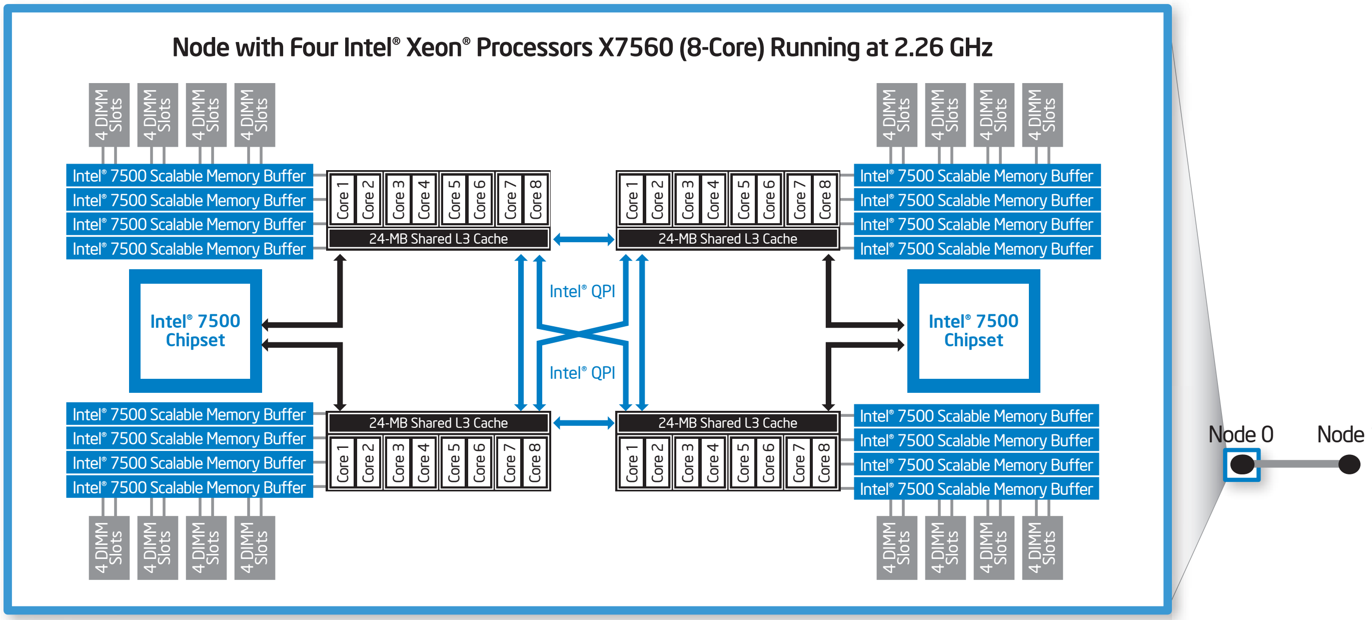


Figure 9. Node in a third-generation high-performance computing environment (HPC-3) large-memory compute server. Two such nodes are connected to form a single 2 TB system. Left: One node based on four 8-core processors with 1 TB of RAM. Right: Two nodes interconnect to create a 2 TB system.

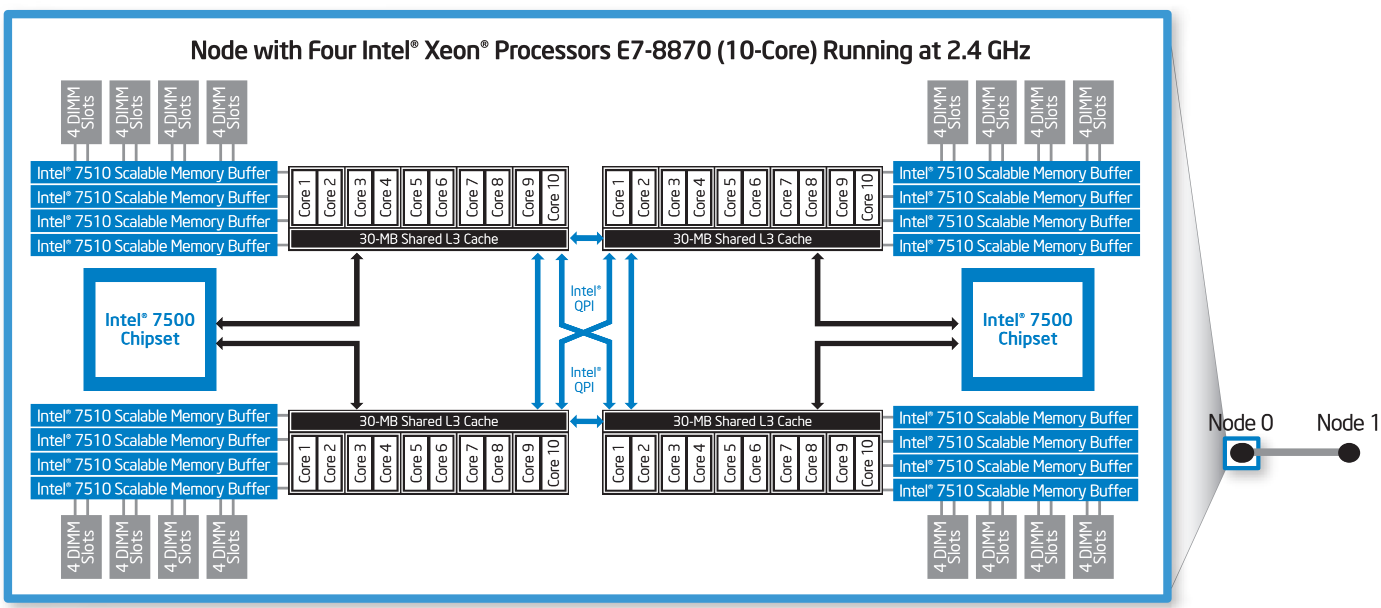
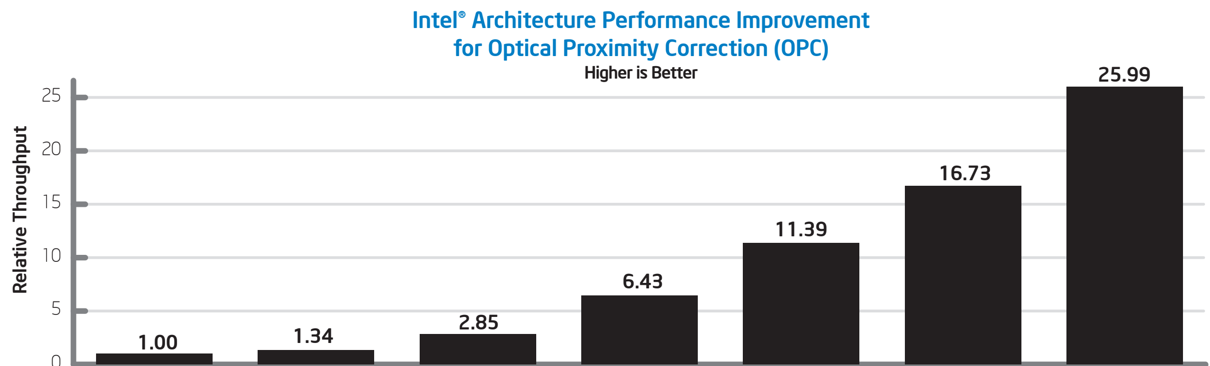


Figure 10. Node in a third-generation and fourth-generation high-performance computing environment (HPC-3 and HPC-4) large-memory compute server. Two such nodes are connected to form a single 2 TB system. Left: One node based on four 10-core processors with 1 TB of RAM. Right: Two nodes interconnect to create a 2 TB system.

Table 2. Comparison of HPC-1, HPC-2, HPC-3, and HPC-4 Large-Memory Compute Servers

	HPC-1	HPC-2	HPC-3	HPC-4
Total CPU Cores	32 or 64	96	64 or 80	80
Memory Capacity	512 GB	1 TB	2 TB	2 TB plus Intel® Solid-State Drive fast swap
Data Center Rack Space Needed	24 rack units	16 rack units	8 rack units	8 rack units
Power Consumed	7.3 kW	3.6 kW	2.52 kW or 2.34 kW	2.34 kW

HPC-1 First-generation high-performance computing environment; HPC-2 Second-generation high-performance computing environment; HPC-3 Third-generation high-performance computing environment; HPC-4 Fourth-generation high-performance computing environment



	64-bit Intel® Xeon® Processor with 1 MB L2 Cache (3.6 GHz)	64-bit Intel® Xeon® Processor with 2 MB L2 Cache (3.8 GHz)	Intel® Xeon® Processor 5150 (2.66 GHz)	Intel® Xeon® Processor E5450 (3.0 GHz)	Intel® Xeon® Processor X5570 (2.93 GHz)	Intel® Xeon® Processor X5675 (3.06 GHz)	Intel® Xeon® Processor E5-2680 (2.7 GHz)
OPC Jobs (One OPC job per core)	2	2	4	8	8	12	16
Runtime (HH:MM:SS)	10:40:12	07:58:31	03:44:20	01:39:32	00:56:11	00:38:16	00:24:38
Relative Throughput	**1.00**	**1.34**	**2.85**	**6.43**	**11.39**	**16.73**	**25.99**

Figure 11. Servers based on successive generations of multicore Intel® Xeon® processors continue to deliver improvements in batch computing performance.

Network

By carefully characterizing data transfer requirements, we determined the need to increase bandwidth and provide high availability across the tapeout environment. We made the following upgrades:

- For HPC-1: all master and large-memory compute servers to at least 2x 1 GB/s network connection with switch-level failover capabilities; all slave servers to at least 100 Mb/s.
- By HPC-4: all master and large-memory compute servers to 2x 10 GB/s; all slave servers to 1 Gb/s.

We provided 6x 1 GB/s uplinks per rack in HPC-1 and 2x 10 GB/s uplinks per rack by HPC-4. Also, we configured the two uplinks to connect to two different switches and virtual LANs (VLANs) for redundancy in case of link or upstream switch failure.

Batch Clustering: Job Scheduler Improvements

Tapeout involves scheduling thousands of simultaneous OPC batch jobs as efficiently as possible. Heavy job loading exposed quality issues in the batch job scheduler, resulting in a higher level of job failures and lower server utilization.

We devised a systematic test method based on synthetic jobs that did not generate load on the CPU. This enabled us to analyze and stress test the job scheduler code—scheduling up to 11,000 production machines while the machines were still being used for regular production work. As a result, we were able to execute several million test jobs per day.

This method was key to developing an improved scheduler as well as to detect and fix bugs, because it allowed us to rapidly test combinations of hardware and OS scheduler configurations.

Our improved scheduler cut in half the time required for job submission and scheduling. It also supported three independent job queues, resulting in a 13.5x increase in the total number of jobs supported by our tapeout resources.

EDA Application License Servers

EDA application license server performance was a factor constraining the growth of our tapeout environment. Random job failures occurred when the license servers were heavily loaded, resulting in an inability to check out more licenses.

As when optimizing the job scheduler, testability was a key challenge. It was impractical to extensively test the license servers using the actual EDA application, because this would have required the dedicated use of more than 5,000 production server CPUs over several days.

We overcame this obstacle by working with suppliers to develop a methodology for testing simultaneous license checkout of 1,000 keys per second from a single machine—while running regular production jobs. This enabled us to stress-test the license servers and validate new software and configuration combinations.

This approach led to the discovery of a fundamental bug in the license server application that limited scalability and enabled suppliers to fix it before it impacted our growing production environment.

We used the same method to demonstrate to our EDA application supplier that license servers based on Intel® architecture were stable and more scalable than the RISC-based servers used in our pre-HPC production environment. The move to Intel® architecture-based license servers meant that our design and tapeout computing environment was completely based on Intel architecture.

Enterprise Linux* OS

To improve the stability of batch computing, we standardized on the same enterprise Linux OS on all our HPC large-memory and batch computing servers. As we took advantage of new hardware platforms, we worked with the OS supplier to enhance and optimize the OS to support new hardware features. We also worked with the OS supplier to resolve bugs and to help ensure interoperability between new and existing platforms.

Application and Platform Tuning

To take full advantage of multicore platforms, we have optimized BIOS settings for processors, memory, and hard drive operation modes to achieve a further 20-percent performance improvement. We also periodically performed internal stress tests to help ensure that the efficiency of our HPC cluster is comparable with top-ranked GbE supercomputing clusters in the Top500*.

HPC BENEFITS

The use of HPC-1 to enable tapeout of Intel's breakthrough 45nm processors delivered significant value to Intel. Financial analysis showed that HPC-1 alone delivered net present value (NPV) of USD 44.72 million, of which USD 22.68 million was directly attributable to the first generation of the parallel storage solution and USD 16.64 million to the large-memory compute servers. Batch compute server improvements reduced requirements for data center space, power, and cooling, resulting in USD 5.4 million NPV.

HPC-2, HPC-3, and HPC-4 have continued to deliver substantial increases in scalability and performance, as shown in Table 3.

In addition to providing the major increases in compute capacity required for new processor generations, HPC has dramatically improved the stability of our tapeout environment. The number of issues impacting tapeout declined sharply after the implementation of HPC-1, and this improvement has been sustained even as the environment has supported continuous growth in demand. As shown in Figure 12, since deployment, HPC has supported more than a 29.87x increase in demand, with a 20x increase in stability.

Intel Tapeout Computing Metrics

Pre-HPC versus HPC Environment Comparison

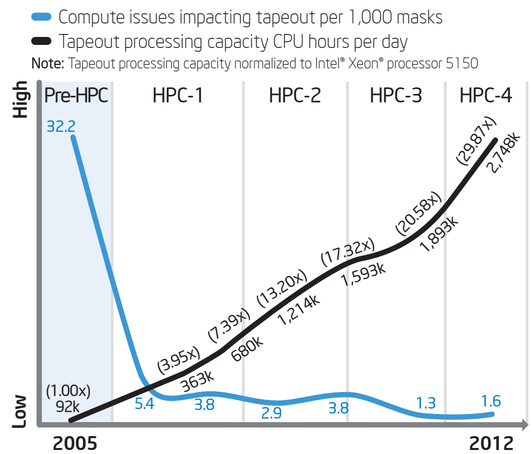


Figure 12. High-performance computing (HPC) has shown increased stability even as demand has increased.

Table 3. Summary of HPC Generational Performance Improvements Note: Generation scaling improvements shown in parenthesis.

2006 → 2007 → 2008 → 2009 → 2010 → 2011 → 2012 → 2013

HPC Technology Areas	**HPC-1** Optimize for 45nm Support >=65nm	**HPC-2** Optimize for 32nm Support >=45nm	**HPC-3** Optimize for 22nm Support >=32nm	**HPC-4** Optimize for 14nm Support >=22nm
BATCH CLUSTERING – Stability, Scalability, Features				
Systems per Pool	8,500 (1.3x)	11,000 (1.3x)	11,000 (1x)	11,000 (1x)
Jobs per Pool	20,000+ (1.5x)	30,000+ (1.5x)	120,000+ (4x)	180,000+ (1.5x)
STORAGE AND BACKUP – Scalability, Performance, I/O Throughput				
I/O Spec Throughput	5,120 (10x)	5,120 (1x)	14,080 (2.75x)	14,080 (1x)
High-volume Manufacturing I/O Throughput	3,200+ MB/s	3,500+ MB/s (1.1x)	5,300+ MB/s (1.5x)	5,300+ MB/s (1x)
Volume Size	3.2 TB (8x)	6.4 TB (2x)	6.4 TB (1x)	11 TB (1.72x)
Single-Stream Performance	70 MB/s (1x)	160 MB/s (2.3x)	240 MB/s (1.5x)	240 MB/s (1x)
Hardware and Software Capacity	Parallel Storage-Generation 1 100 TB (8x)	Parallel Storage-Generation 2 100 TB (1x)	Parallel Storage-Generation 3 100 TB (1x, SSD)	Parallel Storage-Generation 4 360 TB (3.6x)
NETWORK – Scalability, Stability				
Storage	40 Gb/s (10x)	40 Gb/s (1x)	100 Gb/s (2.5x)	300 Gb/s (3x)
Master	1 Gb/s (10x)	2x 1 Gb/s (1x, Redundancy)	2x 10 Gb/s (10x, Redundancy)	2x 10 Gb/s (1x, Redundancy)
Slave	100 Mb/s (1x)	100 Mb/s (1x)	100 Mb/s - 1 Gb/s (1x - 10x)	1 Gb/s (1x)
COMPUTE SERVERS – Optimized for Performance, Throughput, Capacity, Power, and Data Center Space				
Large RAM Server Performance Throughput	512 GB (4x) 1.6x to 5x	1 TB (2x) 1.7x	2 TB (2x) 2.24x	2 TB (1x, SSD) 2.24x
Batch Node Performance Throughput	2-Socket/Dual-Core/16 GB 2.1x	2-Socket/Quad-Core/32 GB 2.3x	2-Socket/4-8 Core/72-128 GB 1.53x	2-Socket/ 8+ Cores/128-256 GB 1.4x
OS – New Hardware Feature Support, Scalability, Stability, Performance				
Enterprise Feature	Stable, Intersystem NUMA Support	Multi-Core Optimized	Power, Performance Optimized	Power, Performance Optimized

KEY LEARNINGS AND FUTURE PLANS

The success of HPC was based on several key factors.

- **Alignment of technology with business requirements.** In specifying the HPC solution, we carefully aligned business and technical requirements, resulting in a system that delivered the scalability required to support 45nm and beyond processor tapeout. We are continuing to use this model to align successive HPC and process technology generations.
- **Informed risk-taking.** To optimize solutions for HPC, we needed to take risks, such as our pioneering decisions to use our parallel storage system and the modular large-memory compute servers. Implementing these solutions required significant ecosystem development. Our team understood that there was a significant risk, with concerns about supplier maturity and the viability of the solution in production use, yet we strongly believed the system would deliver great rewards to Intel. The fact that these solutions worked and enabled 45nm processor—and newer generations, up to 14nm processor—tapeout demonstrated that the risk level was appropriate.
- **Governance.** We adopted a holistic view of HPC capabilities and created a clear computing roadmap. Disciplined governance then helped ensure that we executed according to this roadmap. Intel IT and business groups acted as a single team with collective responsibility; a joint manufacturing and IT committee reviewed and approved computing recommendations.

We are currently developing the fifth HPC generation (HPC-5) to support the tapeout of 10nm processors. As with previous generations, we expect to optimize the throughput of 10nm tapeout applications with significant, balanced improvements across all HPC components. This includes major performance improvements in the areas of storage, compute servers, batch clustering, and network bandwidth. These concepts and learnings are also being applied to other phases of silicon design.

CONCLUSION

Our pioneering HPC approach to silicon design enabled tapeout of the industry's first 45nm processors and numerous follow-on products.

Delivering this solution required replacing our old computing model with an innovative approach aligned with the requirements of Intel process technology generations. Intel's manufacturing group recognized two components of our environment—the parallel storage solution and large-memory Intel Xeon processor-based NUMA systems—as pillars supporting the successful completion of the first 45nm processors. Intel has taped out several silicon products with HPC-1 alone, delivering ROI of USD 44.72 million and reducing tapeout time from 25 to less than 10 days.[1] We are continuing to develop new HPC generations as Intel process technology advances to meet scalability requirements, while keeping tapeout time to less than 10 days..

For more information on Intel IT best practices, visit www.intel.com/it.

ACRONYMS

DDR2	Double Data Rate 2
DDR3	Double Data Rate 3
DMI	direct media interface
DRC	design rule check
ECC	error correction code
EDA	electronic design automation
FSB	front-side bus
GB	gigabyte
GbE	gigabit ethernet
Gb/s	gigabits per second
GHz	gigahertz
GT/s	giga-transfers per second
HPC	high-performance computing
HPC-1	first-generation HPC environment
HPC-2	second-generation HPC environment
HPC-3	third-generation HPC environment
HPC-4	fourth-generation HPC environment
KB	kilobyte
kW	kilowatt
MB	megabyte
Mb/s	megabits per second
MB/s	megabytes per second
NAS	network-attached storage
NPV	net present value
NUMA	non-uniform memory access
OPC	optical proximity correction
RET	resolution enhancement techniques
RPM	revolutions per minute
SATA	Serial Advanced Technology Attachment
SSD	solid-state drive
TB	terabyte
TCO	total cost of ownership
VLAN	virtual LAN

[1] Tapeout time was reduced from 25 days for the first 65nm process technology-based microprocessor in a non-HPC compute environment to less than 10 days through an HPC-enabled environment. Financial analysis showed that HPC-1 alone delivered a net present value (NPV) of USD 44.72 million.

 Printed in USA ♻ Please Recycle 0513/WWES/KC/PDF 328926-002US