IDC

Analyze the Future

WHITE PAPER

The Strategy of the Market Share Leader in High-Performance Computing

Sponsored by: Hewlett-Packard

Steve Conway
Earl Joseph, Ph.D.
Richard Walsh
Jie Wu

February 2009

Global Headquarters: 5 Speen Street Framingham, MA 01701 USA P.508.872.8200 F.508.935.4015 www.idc.com

IDC OPINION

The recent growth of the high-performance computing (HPC) market, powered almost exclusively by the adoption of Linux clusters, has been dramatic. But the burgeoning complexity of technical clusters — the dominant species of HPC system — could have a dampening effect if it is not adequately addressed. A major cause of cluster complexity is rampant hardware parallelism: systems *averaging* thousands of processors, each of them a multi-core chip whose core count doubles every 18–24 months. Add to this hardware parallelism trend the additional issues of third-party software costs, weak interconnect performance, the difficulty of scaling many applications beyond a single node, storage and data management, power, cooling, and facility space, and cluster complexity quickly begins to skyrocket.

According to IDC research, Hewlett-Packard (HP) has moved into the HPC market share (revenue) leadership position principally by working to alleviate cluster complexity through a coordinated strategy of investment and innovation, HPC-centric product planning and design, external partnerships, application expertise and focus, system integration, sales, and support — focusing on alleviating this complexity for datacenter administrators and end users. HP's broad product portfolio for HPC also leverages the company's innovations for the mainstream enterprise IT market and advances from HP Labs. HP has also taken a customer-centric, technology-agnostic approach that offers buyers a wide range of technology and product choices. And the company has amassed the in-house domain expertise needed to act as a trusted advisor to HPC users.

Although the revenue leadership position could change hands in the future, IDC believes that HP's approach to the HPC market positions the company well for sustaining its strong presence and exploiting near-term growth trends and opportunities in this market. HP's strategy is closely aligned with fast-growing HPC market segments (sub-$250,000 price points), and HP has recently scored some notable wins in the high-mindshare, high-end HPC market.

IN THIS WHITE PAPER

This white paper explores the reasons behind HP's rise to market share (revenue) leadership in the fast-growing global market for HPC servers, also known as technical servers. The paper begins with an overview of the HPC market's historical and projected growth, requirements and trends, and customer pain points. The overview is followed by a close examination of HP's differentiated strategy for addressing this market — the company's approaches to R&D, products, marketing, sales, and support. Finally, IDC discusses the opportunities and challenges confronting HP in this market, offers guidance for HPC buyers and vendors, and provides a future outlook.

SITUATION OVERVIEW

The HPC market has undergone another paradigm shift over the past decade, and especially since 2002. In the late 1990s, the high end of the HPC market was still dominated by proprietary purpose-built supercomputers and the rest of the market by RISC-Unix SMP servers. The Intel-based "ASCI Red" MPP supercomputer, arguably the first sustained teraflops system, provided 1.6 peak teraflops for $55 million, or about $30 million per peak teraflop.

By 2003, the dominant high-end systems were clusters of RISC-Unix commercial off-the-shelf (COTS) systems because their standard technologies could scale to large system sizes while providing better price/performance than special-architecture supercomputers. RISC-Unix SMP servers dominated the rest of the market, but something new was beginning to appear — Linux-based "Beowulf" clusters based on x86-processor servers, mostly built by users, often out of obsolete personal computers, and managed catch-as-catch-can with open source or user-developed software.

Fast-forward to 2008, and two-thirds of HPC server revenue belonged to industry-standard Linux clusters. Clusters today not only dominate the workgroup, departmental, and enterprise HPC market segments for systems priced below $500,000 but also have made strong inroads into the priciest "supercomputer" segment that includes the world's largest, most powerful HPC systems.

Every one of the systems listed on the November 2008 TOP500 list of the world's most powerful supercomputers had peak performance of more than one teraflop, and more than 80% were clusters. This revolution in architecture — from scale-up to scale-out and from proprietary to industry standard — has been accompanied by dramatic reductions in the cost of HPC cycles and a corresponding surge in market growth. A teraflops system now sells for a street price well under $100,000, enabling many smaller organizations and business units, some in new-to-HPC vertical markets, to exploit the proven benefits of HPC: faster time to market, superior product quality, and novel insights that can create lasting competitive advantage.

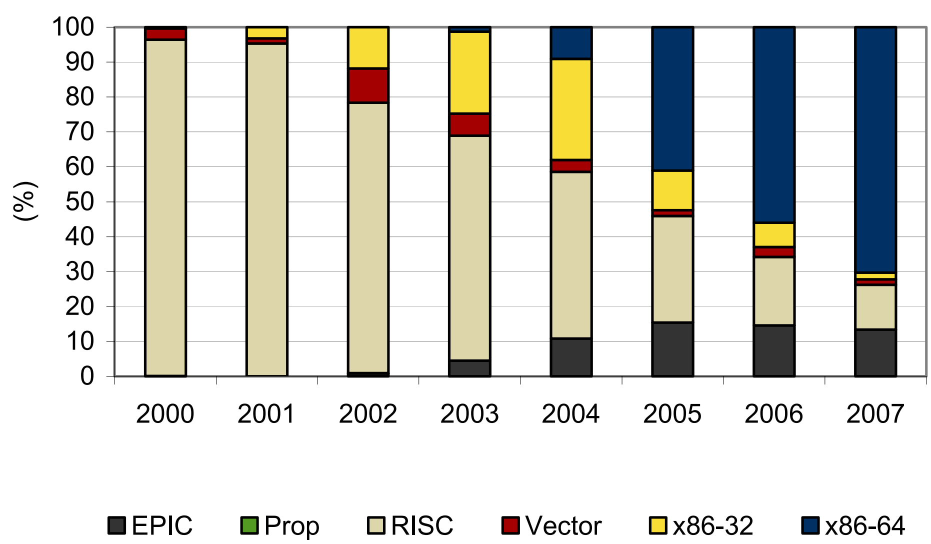
During this period of dramatic change, HP clearly "got" the underlying technology trends driving change. Rather than circling the wagons to protect its RISC-Unix business, HP expanded its strategic focus to Linux clusters and encouraged the move toward lower-cost industry-standard solutions. Starting in 2003, HP emerged as the HPC revenue leader, widening its lead in 2007 and 2008. The following sections provide a closer look at HP's strategy during this period, its impact on the market, and prospects for the next few years.

The Tipping Point: 2001–2003

As recently as 2001, systems based on x86 processors in industry-standard servers accounted for less than 5% of all HPC server revenue and clusters employing any kind of processor barely exceeded 10% of the HPC market (see Figure 1). The leading HPC system vendors were IBM, HP, Sun, and SGI, all focused on proprietary RISC-processor, Unix SMP servers. However, there were indications of changes to come. The U.S. Department of Energy's Accelerated Strategic Computing Initiative (ASCI) program was giving a major boost to scale-out architectures by demonstrating that large clusters of RISC-Unix servers — the architecture behind ASCI Blue, White, and Q — could deliver scalable performance on an increasing number of applications at a lower cost than previous purpose-built supercomputers. Intel and AMD processors were challenging RISC processors in floating-point performance at a fraction of the cost, and Linux was maturing in HPC contexts.

FIGURE 1

Total HPC Revenue Share by Processor Type, 2000–2007

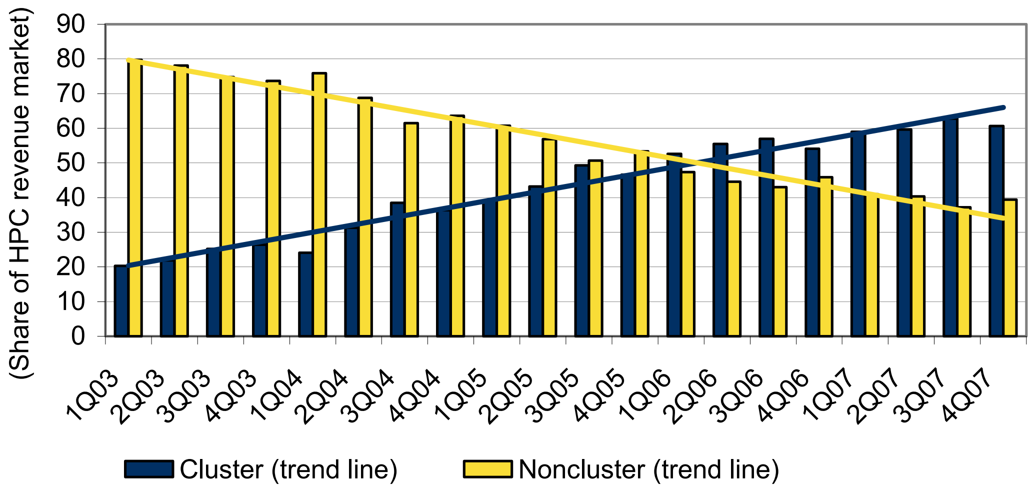


Source: IDC, 2008

Just two years later, in 2003, year-over-year cluster growth skyrocketed to 160% (see Figure 2) and x86-based systems captured about 25% of all HPC revenue. This was also the tipping point in HP's strategy. HP was in the midst of moving from its own PA-RISC processor to Intel's new Itanium. The company had just acquired Compaq and already faced the daunting task of moving the AlphaServer HPC installed base to HP products as Alpha was also phased out. HP had a successful Unix business model for HPC and good reasons to stay focused there.

FIGURE 2

Growth in HPC Clusters, 2003–2007



Source: IDC, 2008

Despite the rapid growth of industry-standard clusters, major roadblocks still were holding up their market adoption. Although clusters offered the prospect of much improved price/performance, they were complex, difficult to administer, and unreliable and had not garnered the support of the HPC ISVs necessary to drive volume. No major vendor offered industry-standard clusters as a product, and no major HPC customer had applied them in other than experimental, largely academic settings.

Against this backdrop, HP decided to invest in Linux-based industry-standard clusters. HP's HPC Division recognized that the price/performance advantage of the scale-out model would motivate suppliers to overcome other roadblocks and eventually surpass RISC-Unix in the market. HP chose to accelerate market adoption of Linux clusters by alleviating those roadblocks, even if that meant cannibalizing its own RISC-Unix HPC business.

The five key elements of HP's strategy were:

- Offer flexible industry-standard choices for customers
- Make clusters easy to configure and purchase
- Optimize and pretest configurations for performance and reliability
- Deliver a full solution stack through collaborative development
- Drive emerging HPC technologies into fully supported HP products

In short, HP's goal was to help drive the transformation to industry-standard clusters by systematically alleviating their drawbacks and enabling users to realize their price/performance benefits. The company hoped to stimulate overall market growth as well as to increase its share.

HP and the HPC Market, 2003–2009 and Beyond

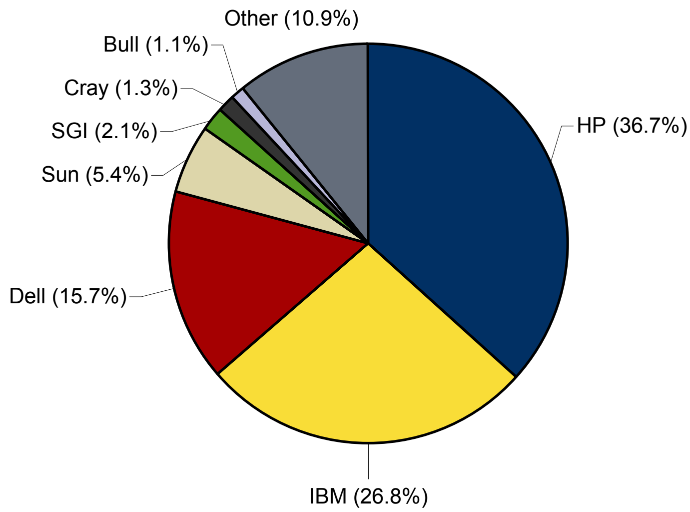
Propelled by cluster revenue, the HPC market saw stellar growth in the years after 2003, including three years of double-digit growth from 2003 to 2005, followed by a still-impressive 9% year-over-year growth between 2005 and 2006, and 15% growth in 2007. In the second half of 2008, the global economic recession throttled down the rapid growth of the HPC server market. IDC estimates that full-year 2008 HPC server revenue will come in at about $9.6 billion, down 4.2% from 2007. IDC's base case forecast shows HPC server revenue declining 5.4% in 2009, resuming modest growth in 2010, and rebuilding to robust 9.6% growth and $11.7 billion revenue in 2012. If our base case assumptions hold true, HPC will remain one of the bright spots in the IT space over the next few years.

The majority of the growth will continue to come from Linux clusters priced below $500,000 — and especially those priced under $250,000. The Technical Workgroup segment (systems selling for under $100,000) is being expanded by the emergence of a potentially important new category of entry-level supercomputer products, such as HP's Cluster Express, and by a potential surge in the migration of desktop technical computing users to technical servers.

HP was one of the pioneers among major HPC vendors when it introduced cluster products in 2003 and the Unified Cluster Portfolio (UCP) in 2004. All major HPC players have benefited from the market's growth in recent years, but HP outpaced the overall market by achieving five consecutive years of double-digit growth from 2002 to 2007. HP has been the HPC revenue leader or coleader since 2003, garnering 37% of HPC server market revenue in the second quarter of 2008 (see Figure 3).

FIGURE 3

Vendor HPC Market Shares in 2Q08



Source: IDC, 2008

HP's broad portfolio of scale-out clusters made up more than 50% of the company's overall technical server sales in 2007 and 2008, a higher percentage than for other vendors with mixed server portfolios. This means that HP also sells a substantial number of nonclustered technical computing systems, including scale-up systems as large as the 128-core Superdome.

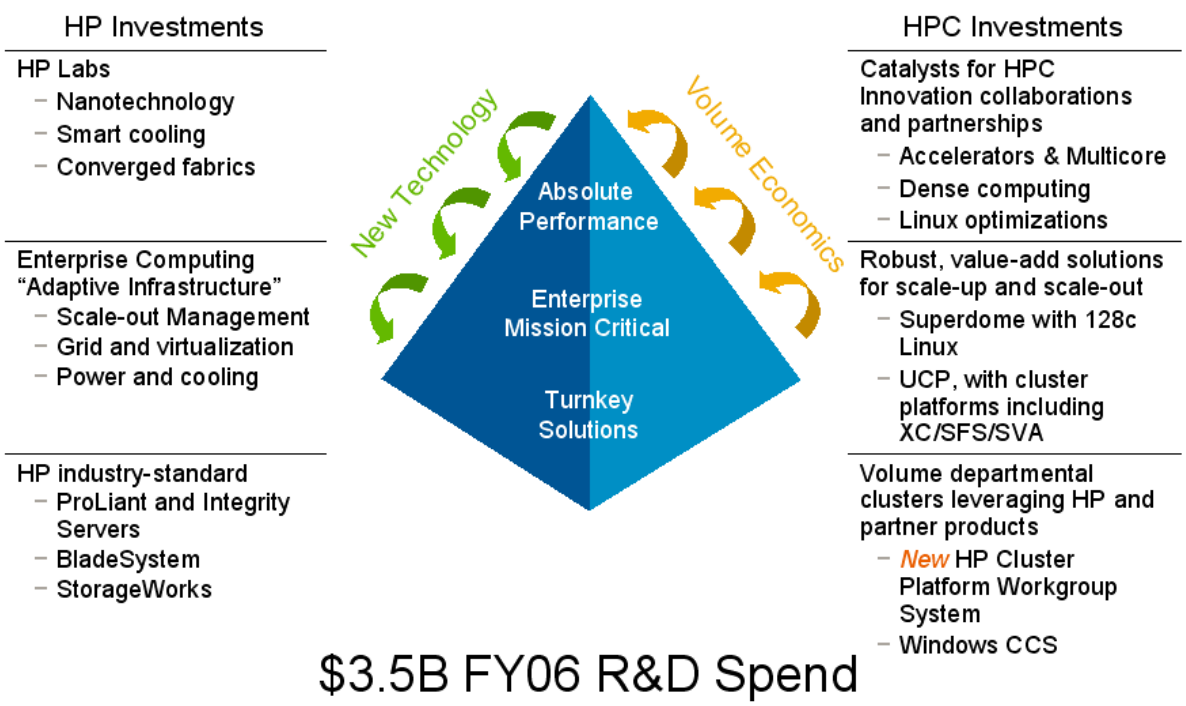
Offering clusters that heavily leverage standard technologies allows HP to roll out new products quickly and at prices that are more affordable than those of HPC systems with a larger portion of custom content. HP adds differentiated value to its branded clusters not only via HPC-specific innovation but also by exploiting non-HPC-specific innovations achieved elsewhere in the company.

HP's Strategy: A Closer Look

HP formulated its HPC strategic vision shortly after the Compaq acquisition and has implemented it consistently for the past five years. While grounded firmly in the volume economics of industry-standard computing, it also leverages the vibrant concepts and technologies spun off by the market's constant quest for more scalable performance. Figure 4 is a depiction of this vision.

FIGURE 4

Translating Market Leadership into Customer Value



Source: HP, 2008

The left side of Figure 4 focuses on innovation. HP forms close collaborations with HPC customers driven to innovate by their need for performance and scale. Through research and advanced development, often in partnership with other leading industry-standard technology providers, emerging technologies either demonstrate their viability for the broader market or become niche offerings of specialized companies.

For those technologies with volume potential (see the right side of Figure 4), HP's product development groups commercialize and launch high-quality, fully supported versions, enabling volume economics to reduce costs and broaden market adoption. In this process, HP embraces components from partners and concentrates its own energies on ensuring that the customer solution — in this case HPC clusters — is well integrated, reliable, and fully supported.

The strategy has two important benefits for HPC customers. First, it accelerates the transition of new technologies from their low-volume, high-cost, early-adopter phase to their high-volume, low-cost, mature phase. Customers of all sizes benefit from the reduced costs. Second, because HP collaborates rather than competes with other technology providers, the strategy stimulates innovation throughout the ecosystem and gives HPC customers a wider choice of new products. HP shares its market success with partners, but benefits as the solution integrator and service provider.

Beneath the umbrella of this strategic vision, HP focused on the five initiatives (listed earlier) designed to drive the growth of industry-standard clusters. In the following sections, we take a closer look at each initiative and its impact on the market.

Offer Flexible Industry-Standard Choices for Customers

A significant driver of HP's HPC market share growth has been the variety and configurability of the cluster products the company offers. HP has capitalized on the component-construction mindset of HPC cluster buyers, which allows buyers to assemble systems that meet their specific requirements by standardizing and modularizing almost every aspect of the system configuration process. This begins with a variety of standard-form-factor blades or rackmounted nodes built around a choice of processor technologies — those from the Xeon, Itanium, and Opteron families — along with dual- and quad-core options and dual- and quad-socket boards.

c-Class Blades: Optimized for HPC

IDC research has shown that although rackmounted clusters remain popular, buyer preferences have shifted decidedly in favor of bladed servers and clusters. Driving this shift in preference is the recent significant improvements to blade technologies for HPC. HP's c-Class BladeSystem is an example of a blade architecture optimized for HPC — and the impact that can have on market uptake of a new technology.

HP was one of the first major HPC vendors to begin offering blade servers (early 2004). Like all blades of that generation, however, HP's p-Class had several deficiencies that ruled it out for most HPC users:

- Top-bin processor speeds were not available in blades (too hot for the required density).
- High-performance interconnects were not supported.
- Densities and power consumption were only marginally better than rack servers.
- Blade prices for similar configurations were significantly higher than those for rackmounted configurations.

In developing its c-Class BladeSystem architecture, HP systematically addressed these issues. The result was a design optimized for HPC. A few examples are as follows:

- **HP Active Cool fans.** HP's custom Active Cool fan is designed specifically to efficiently cool any of the cluster offerings. While each enclosure uses a unique power supply, all are actively managed as an array to respond to the systems' power draw by increasing or decreasing draw and delivered power. This design provides ample cooling even for top-bin processors and offers significant power consumption savings as well.
- **Integrated interconnects in the blade enclosure.** The c7000 blade enclosure holds 8 full-height or 16 half-height server or storage blades vertically and has 8 interconnect bays, up to 6 integrated 2250W power supplies, and up to 10 HP Active Cool fans. DDR InfiniBand is supported directly on the midplane.

- **Engineered for RAS.** HP has engineered its c-Class BladeSystem for reliability, availability, and serviceability (RAS) operation. Power supplies can be fitted in an N+N or N+1 configuration and are globally managed to reduce power drawn if any blade in the enclosure is idle. Interconnects can also be fitted in high-bandwidth, dual-rail, redundant configurations. The Onboard Administrator can also be dual-configured. All components are hot pluggable for rapid return to fully redundant operation in the event of a failure. Many features are included to minimize the impact of heat-related failures: a ducted design of the enclosure to ensure critical components get maximum draw off; custom heat sinks for the processors; thermal control through enclosure-based, software-controlled power budgeting; and a patented fan design that raises reliability, draws less power, and cools more efficiently.

- **Reductions in power, space, and complexity.** HP reports that compared with rackmounted clusters, c-Class enhancements require:

  - Up to 40% less power consumption

  - Up to 50% less floor space

  - Up to 75% less cabling

- **Inclusion of InfiniBand interconnects.** HP's BladeSystem DDR, 4X InfiniBand mezzanine card is a good example of the company's focus on c-Class support for HPC. This card is based on Mellanox's low-latency ConnectX InfiniBand chip and software technology. ConnectX was introduced in late 1Q07 by Mellanox. HP's blade-ready DDR mezzanine card was made available only three months later in June 2007. This card was the first DDR, InfiniBand HCA in a blade-compatible form factor to be delivered to the marketplace. HP introduced a companion ConnectX-based DDR, InfiniBand c-Class switch module at the same time, as well as external switches from partners Voltaire and Cisco.

Market acceptance of c-Class blades — led by HPC — has been nothing short of phenomenal. In the first two years following their introduction, they captured 176 entries in, or 35% of, the TOP500 of the world's most powerful supercomputers, more than any other system architecture and more blade installations than all other vendors combined.

HP offers both blade and rackmounted HPC clusters. HP's blade offerings provide higher density, lower TCO, and higher reliability in a more fully integrated design. HP's rackmounted offerings provide more flexibility in memory choices, interconnects, and other components. In both solutions, HP provides a broad set of options with processors, operating systems, interconnects, software middleware stacks, and other elements to directly match the needs of varying customer workloads.

Interconnect, Storage, and Software Choices

HP provides a modularized and varied set of interconnects that include a mix of Ethernet and InfiniBand switches as well as single- and dual-port NICs, some from HP and some from other name-brand vendors. The company has long-standing relationships with Myricom and Quadrics.

HP's storage offerings are similarly designed in modular and enclosure-friendly form factors. For instance, the company's half-height, All-in-One Storage Blade fits into both of HP's rack enclosures to serve all associated compute blades, while a sibling storage blade is designed to augment the storage capacity of single server blades. HP also offers a half-height in-enclosure tape blade, as well as a full range of rackmounted storage subsystems that can be optionally configured into clusters.

Above these hardware permutations, HP offers a wide range of its own software solutions (see the software discussion below) as well as partner-developed tools. These include operating systems (RHEL, SuSE, Windows HPC, HP-UX), math and message passing libraries (HP MLIB and MPI), cluster management tools (ICE Linux, CMU), cluster file systems (HP Scalable File Share [HP SFS], PolyServe), and cluster visualization tools (HP Scalable Visualization Array [HP SVA]).

Make Clusters Easy: HP's Unified Cluster Portfolio

While so many options make it possible for every HPC market segment to optimize its choice of cluster configuration, they also force enormous complexity into the process of configuring, ordering, building, and supporting the resulting clusters. To address this complexity, HP has adopted a standard, build-to-order approach for clusters.

The HP UCP is in essence a selection of reference architectures whose hardware and software elements have been carefully selected to ensure that they will work well together at user sites. HP converts the architectures into a set of menus that reflect the many permutations of components that can be integrated successfully. From these menus, the company creates a set of tools that simplify and automate the process of configuring a UCP cluster along with the related manufacturing documentation that enables HP factories to build clusters reliably in volume.

HP uses these configuration tools on its direct-order Web sites and distributes them to its sales force and channel partners. The configurators take as inputs a few key parameters describing the desired cluster (e.g., server type, processor model and speed, memory, storage, interconnect, and software and service options). They output completely designed clusters with a detailed bill of materials and a list price. Salespeople and customers can easily evaluate alternative configurations, including their cost differences. Customers can have the cluster built to order by HP or a channel partner, or they can order the components and integrate the cluster themselves.

Optimize and Pretest for Performance and Reliability

While UCP simplifies clusters, its chief goal is to provide buyers with a choice of preconfigured, pretested cluster solutions that they can successfully deploy very rapidly, with minimal technical knowledge and without experiencing glitches or harmful delays to their own core mission. UCP helps shorten the overall cycle time to a smoothly working HPC solution.

Users choose from a wide range of tested options — processors, interconnects, operating systems, middleware, and application software — from both the open source community and the commercial community. UCP also includes scalable storage for HPC clusters, HP SFS, and HP SVA. These elements have passed stringent tests to gain HP's stamp of approval. This approach may seem straightforward on paper, but in practice it requires a significant investment by HP in hardware, software, vendor relationships, and personnel to set up and run smoothly; hence IDC believes it is beyond the reach of smaller vendors. IDC expects HP to continue strengthening and broadening UCP as new technologies mature.

When built to order by HP, UCP clusters are configured exactly as specified by customers, integrated and tested in HP factories, and shipped in racks to simplify final assembly at the customer site. Customer benchmarks and final tests are also performed in the company's manufacturing integration centers. These centers, where hardware and software components as well as experienced personnel are readily available, represent a real differentiator for HP from an operational perspective. While competitors claim to have similar procedures and processes in place, HP has been able to leverage the capabilities of its world-class manufacturing, business, and technical teams to continuously improve cluster reliability and total customer experience.

Deliver a Full Solution Stack Through Collaborative Development

When HP shifted its focus to industry-standard clusters in 2003, the company had a strong portfolio of HPC software running on Unix, but none at all on Linux. Indeed, there were few if any Linux software solutions that demonstrated the quality and reliability to satisfy mainstream, production-oriented HPC shops. Recognizing this lack of proven software as an important barrier to market acceptance, HP committed to delivering a complete solution stack on Linux. Unlike its major competitors, however, HP chose to collaborate with other software development groups, integrating the best open source and proprietary components and developing its own IP only where it could add substantial value.

Three HP software products exemplified this approach: XC cluster manager, SFS data manager, and SVA. Taken as an integrated package, the three products addressed what HP viewed as the cornerstones of HPC productivity: scalable computation, data management, and visualization. Yet HP delivered XC and SFS on a completely new OS platform with robust functionality less than 18 months after starting the initiative.

To accelerate time to market, XC was designed to leverage existing components. It contains many elements, but one of the three key choices was the Simple Linux Utility for Resource Management (SLURM), an open source project of Lawrence Livermore National Labs. It allocates compute nodes to users; provides a framework for starting, executing, and monitoring parallel jobs; and manages a queue of pending work. To actually do the monitoring, XC employs Ganglia, a scalable distributed monitoring system developed as an open source project at U.C. Berkeley. It is based on a hierarchical design targeted at federations of clusters and achieves very low per-node overheads and high concurrency. For job submission and scheduling, HP chose Platform Computing's LSF because its powerful and flexible policy-based scheduler had established it as the category leader.

While HP maintained a traditional royalty-based relationship with LSF, it forged new ground with the SLURM, Ganglia, and other open source communities. HP became an active member, returned its own code integrating these components to the open source pool, made significant new functional contributions (particularly to SLURM), and in general attempted to help improve both the functionality and the quality for all community members, whether they were XC customers or not.

SFS incorporated a totally new cluster file system, Lustre, and was brought to market through collaboration between the U.S. Department of Energy's ASCI PathForward program, Cluster Resources Inc., and HP. The collaboration aimed to deliver an object-based distributed file system that would scale to tens of thousands of compute nodes, petabytes of data, and terabytes per second of aggregate I/O bandwidth. Cluster Resources developed the file system and HP tested it, added its own IP in the form of management tools for the file system and the underlying storage hardware, and released the combination as HP SFS. Lustre is now widely used in the largest public sector clusters, while HP SFS has a significant base of public and private sector customers around the world. Interestingly, HP's support for the Lustre-based SFS continues even though Sun acquired CFS more than one year ago.

SVA also started as an ASCI PathForward project. The program originally included both special-purpose hardware and a library of software tools to enable cluster-based image compositing. Its goal was to scale visualization performance to handle massive data sets, extremely high resolution, and real-time streaming and manipulation of composited images. Halfway through the project, both HP and the Department of Energy realized that, given the rapid advancement in GPU technologies and in PCI Express bus speeds, the hardware component was unnecessary. But HP continued with software development and released SVA as an extension of XC, as well as its Parallel Compositing Library, which contains a proposed standard API for parallel compositing along with an open source reference implementation. SVA customers, HP, and its partners now form the community for this library. They continue to invest in its evolution and hope it will form the basis of an industry-standard approach to cluster-based compositing, a standard that — like the Message Passing Interface (MPI) — could encourage ISV support for cluster visualization.

HP MPI: Making It Easier for ISVs

In 2003, Linux clusters were missing not only management middleware but also virtually all the key ISV applications needed to drive volume acceptance. While ISVs were moving to support individual servers under Linux, they were reluctant to take on the complexities of cluster support, which at the time required them to test and certify a different MPI library with every interconnect and to build an additional executable image for each such pair. They were also reluctant to support multiple Linux distributions and multiple x86 processor architectures because of the dramatic impact on their support matrix.

Because of its long-standing close partnerships with HPC ISVs, HP was keenly aware of these difficulties. MPI is the key enabling technology to scale application performance across servers in a cluster, and lack of application support would have had a major dampening effect on market adoption of Linux clusters. Working closely with its partners, HP developed a plan to alleviate these difficulties by extending the capabilities of its own MPI library, HP MPI. HP's key concept was to eliminate the need for ISVs to create and test multiple executables, thus making Linux cluster support relatively simple and inexpensive. The company accomplished this by:

- Supporting all the major Linux distributions (multiple releases of Red Hat and SuSE)

- Supporting all the major industry-standard processor architectures (Xeon, Opteron, Itanium)

- Supporting all the major HPC interconnects (InfiniBand, Gigabit Ethernet, 10 Gigabit Ethernet, Quadrics, Myrinet, and others)

- Licensing ISVs to package the HP MPI runtime with their application executable at no charge and supporting their customers regardless of the server they use

ISVs were quick to respond. HP announced the program in early 2004. In June, ANSYS (then a leading independent supplier of structural analysis applications, now part of Fluent) signed on as the first official program participant. Today, nearly all leading HPC ISVs (some 25 according to HP) support Linux MPI, most using HP MPI.

Drive Emerging HPC Technologies into Fully Supported HP Products

The final key element of HP's strategy is its Catalysts for Innovation program. HP clearly recognizes not only the key role of innovation in sustaining its market leadership but also that new technologies must be compatible with the industry-standard ecosystem to achieve volume and drive down costs. HP therefore views itself as a catalyst in recognizing problem areas and evaluating new technologies, but it emphasizes a collaborative approach with industry participants and interested customers. Catalysts for Innovation is a vehicle to drive emerging technologies toward market adoption while sharing risks and rewards. It is in sharp contrast to the inclination of some HP competitors to perform research largely on a proprietary basis.

The conceptual framework for HP's Catalysts for Innovation program was formed in 2003 when HP won its ASCI PathForward contracts for file systems and visualization. Its success in delivering products, along with similar collaborations in the commercial launch of InfiniBand clusters (with Mellanox, Voltaire, and others), led HP to formalize the Catalysts program in 2006.

The Catalysts for Innovation program defines a set of emerging technology areas that HP, with inputs from its partners and customers, wishes to investigate collaboratively. HP assigns a small number of engineers to "kick the tires" through lab evaluations, pilot projects, and so forth, with help from interested partners and customers. The collaborators report results at least twice yearly (at HP's Consortium for Advanced Scientific and Technical [CAST] users' group meetings). Findings often form the basis for papers and presentations at technical conferences as well.

So far, HP has identified 10 emerging technology areas for Catalysts investigations:

- Accelerators

- Multi-core optimization

- Dense computing
- Grid computing
- Converged fabrics
- Advanced power and cooling
- WAN file system caching
- Parallel compositing
- Linux optimizations
- Unified Parallel C

The first two projects exemplify HP's philosophy.

Early in 2007, HP launched its HPC Multi-core Optimization Program, which is aimed at reducing the difficulties HP's HPC customers face in getting acceptable, sustained performance from today's multi-core microprocessor architectures. HP has already implemented numerous multi-core optimization features in its own products (HP MPI, XC) and made significant technology contributions to the open-sourced SLURM. In addition, HP hopes to leverage the best practices and capabilities of vendor partners (e.g., Intel, AMD, Allinea, PGI, Visual Numerics, Stanford University) and customers through this program. Areas of collaboration include products and software tools designed to facilitate parallel application development in multi-core environments.

HP's Multi-core Optimization Program Toolkit is an initial result of this effort. HP will roll out new products and features through the program over time. The project, products, test results, and associated white papers are available at **http://www.hp.com/techservers/hpccn/hpccollaboration/ADCatalyst/Multi-core_Optimization_Program.html**.

With a similar purpose in mind (bringing higher sustained parallel performance to its customers), HP started its HPC Accelerator Program to investigate the current HPC acceleration technologies. Through partnerships with hardware and software vendors (e.g., NVIDIA, ClearSpeed, AMD, Celoxica, RapidMind, Mitrionics), HP is qualifying acceleration options on its servers and will selectively sell and support accelerators as a part of its UCP program.

Readers can find more details about HP's Catalysts for Innovation program at **http://www.hp.com/techservers/hpccn/hpccollaboration/ADCatalyst/**.

The Role of HP Labs

The Catalysts for Innovation program focuses on advanced development, that is, the near-term commercialization of promising technologies. For longer-term research, however, the company has its own innovation powerhouse, HP Labs, which explores and develops future-oriented technologies. HP Labs Director Prith Banerjee has a strong background in HPC. Formerly dean of the College of Engineering at the

University of Illinois at Chicago, Dr. Banerjee's research interests include parallel and distributed computing, compilers, and computer-aided design.

Earlier this year, Dr. Banerjee reorganized HP Labs into a smaller number of individual labs working on 20–30 large, collaborative projects. His goal is to leverage collaboration with universities, government researchers, venture capital firms, and start-up companies to amplify innovation in HP Labs and accelerate the transfer of scientific discoveries into commercial products.

One of the new labs, the Exascale Computing Lab (ECL), under the direction of Norm Jouppi, has selected four focus areas:

- Dramatically improve the manageability, power consumption, and availability of future computing systems
- Develop effective parallel programming techniques and software tools for the manycore era
- Create new capabilities through virtualization and novel hardware/software platforms
- Dramatically improve performance/cost of systems by leveraging future technologies (e.g., nanophotonics)

In IDC's view, creation of the ECL reinforces HP's long-term commitment to high-performance computing and improves its opportunity for continued market leadership. ECL's programs take dead aim at the most pressing long-term barriers to exaflops clusters. HP's collaborative approach will stimulate an entire ecosystem of researchers and partners to innovate. While the current Catalysts projects bring solutions to market for the next several years, ECL will be inventing the next generation of new technologies that will find their way into commercial solutions in five to 10 years.

HPC and Cloud Computing

While the scale-out architectural concept started with the HPC market and the Beowulf project in 1993, it has more recently found commercial traction with the largest of Internet service providers and their experiments with "cloud computing." HP reports that HPC and the Web 2.0 market share a remarkable convergence of technology requirements, market structure, and "computing as a service" concepts. These shared features include:

- Scale-out solution architecture, in which extra capacity or performance is achieved by adding homogeneous computing or storage elements
- Extreme focus on cost-reducing designs, both in commodity components and in power consumption, cooling requirements, and floor space/density
- Rejection of traditional hardware features for high availability in favor of lower-cost software solutions

- Concentrated buying power among industry leaders that requires radically different vendor processes to succeed

- Innovative delivery of computing services using the cloud model, that is, over the Internet, with virtualized but scalable resources, with billing by consumption

The parallels between cloud and HPC are important because they provide clues about how vendors can successfully meet the needs of the fast-growing broader scale-out market. Computing vendors that succeeded in HPC as the market shifted from scale-up to scale-out will have a significant advantage in cloud computing as well.

IDC believes that HP's HPC leadership positions the company very well to succeed in the broader scale-out market and cloud computing. In particular, HP has developed:

- A centralized, highly knowledgeable scale-out design team

- Dedicated concierge service team

- Supply chain optimizations for delivering large clusters

- Deployment support for large clusters — HP's solution architects and design engineers provide firsthand support in deploying its large clusters, while onsite services teams lay the local groundwork but are assisted by remote experts in troubleshooting and acceptance/performance testing.

HP has honed these capabilities and skills over many years in a series of successful large cluster deployments. Recent successes include Pacific Northwest National Laboratories (2,300 ProLiant DL185 G5 servers), Tata's Computational Research Laboratories (1,800 BL460c c-Class blades), and the Swedish federal government (2,128 BL460c c-Class blades).

HP has also balanced its efforts to win the largest HPC procurements with investments to win business in the large number of smaller HPC customers. Through its high-end engagements, HP gained critical experience in designing, building, and deploying large scale-out systems. The company applied this learning to create standard products and volume processes for the broader market.

As the cloud market matures, IDC expects HP to apply the same strategies. While the largest accounts in both HPC and Web 2.0 will drive technology evolution and cost reductions, this top tier will be supplemented over time by a larger number of smaller customers who want to operate their own clouds but want an IT partner to supply the infrastructure. We also expect that over time, some of the top-tier companies will stop designing hardware and developing middleware simply because it is not their core strength. HP's balanced approach and demonstrated scale-out knowledge and experience should serve it well in the cloud computing space.

Organizing for the Future: The Scalable Computing and Infrastructure Group

In May, HP merged its High Performance Computing and Web 2.0 business groups into the new Scalable Computing and Infrastructure (SCI) group. IDC believes that HP is anticipating the next set of major market shifts and organizing for the future.

Although HPC and Web 2.0 customers have radically different views of performance scaling requirements (fine-grained parallel and floating-point intensive versus embarrassingly parallel and integer/transaction oriented), their common needs can still be met effectively with a single model for sales, solution design, delivery, and support.

HP's new SCI organization expressly aims at creating synergy in scale-out strategy, hardware, and solutions. HP is taking steps to extend its HPC leadership into Web 2.0 and cloud computing.

CHALLENGES AND OPPORTUNITIES

Opportunities

- **Stay the course that made HP the market leader.** By maintaining the fundamental strategy underpinning its success to date, HP should be able to fully benefit from the rising tide of technical server sales.

- **Zero in even harder on "ease of everything."** HP is already a leader in making technical clusters easier to acquire, deploy, and operate. The company reports that its online SMB store, with its approachable system configurator, which does not require follow-up from a salesperson, is a growing success story. But technical server, storage, and networking technologies will be increasingly complex to integrate and manage; hence, HP will need to redouble its efforts to hide these underlying intricacies from buyers, programmers, administrators, and end users. This will be important for organic growth based on selling more within HP's installed base, as well as for attracting the important stream of new HPC users now entering the technical server market from the world of technical desktop computing.

- **Trumpet HP's market leadership, including innovation.** It is an axiomatic marketing/sales principle that people prefer to buy from a leader. Now that HPC growth has made this market large enough for a firm of HP's size to pay serious attention to and HP is the revenue leader, the company should become far more vocal about its HPC commitment and standing than it has in the past. That includes the important role played by HPC-related investment and innovation at HP. In sum, greater attention to visibility should help boost revenue.

Challenges

- **Continue to approach high-end, high-mindshare procurements with caution.** A cornerstone of HP's success in the HPC market has been the company's policy of avoiding procurements that provide insufficient margin. Although many high-end procurements make good business sense, inadequately funded procurements are also common at the high end of the market — and are mixed blessings for the vendors that win them. In its relatively new revenue leadership position, HP may be more tempted to pursue low margin/no margin procurements in return for the high mindshare they confer. The challenge for HP will be to strike the right balance between financial prudence and high-end opportunities for advanced innovation and mindshare.

- **Better understand the needs of low-end users.** IDC expects substantial growth to continue at the low end of the HPC server market for systems priced below $250,000 — and especially those priced under $100,000. The latter Technical Workgroup segment is being further expanded by the emergence of a potentially important new category of entry-level supercomputer products, including the HP Cluster Express, and a potential surge in the migration of desktop technical computing users to technical servers. The challenge here is that not enough is known yet about the preferences and requirements of low-end HPC users and desktop technical computing users. A related challenge is simply finding the identities of more of these organizations for marketing and sales purposes. To fully exploit low-end growth opportunities, HP (and other vendors) will need to learn considerably more about them than is known today.

- **Broaden the definition of HPC leadership beyond the TOP500 list.** HP addresses the entire HPC market, from the low end to the high end, but has moved into the market leadership position based primarily on business success and revenue — not based on the firm's presence at the top of the TOP500 list of the world's most powerful supercomputers. And although HP systems have penetrated the top 10 positions on this list, a challenge for HP and the whole HPC industry will still be to broaden the definition of HPC leadership and achievement beyond only the performance on the useful but one-dimensional Linpack test that is the basis for the TOP500 rankings.

CONCLUSION

Growth in the HPC server market has been driven primarily by Linux clusters, which accounted for two-thirds of 2008 revenue and should remain the dominant species of HPC system for the foreseeable future. As clusters of any type (or OS) become larger and increasingly complex, HPC administrators and end users will seek relief in the form of vendor approaches that alleviate this complexity. Vendors that effectively address this complexity, while also meeting customer requirements for performance, price/performance, and environmental factors (power, cooling, facility space), will have greater chances to exploit HPC market growth.

Because many customers in HPC and other markets prefer to buy from a leader, HP's market leadership position will give the company an added selling advantage. HP's rise to revenue leadership and the strategy that powered this rise can also serve as proof points for the growing importance of the low end and midrange of the HPC market — and for the "ease-of-everything" requirements of these customers. And HP's policy of pursuing high-end opportunities aggressively, but selectively, can serve as an object lesson for some less cautious vendors. Now that HP has reestablished its presence in the high-end HPC market, IDC expects the company to continue to win new high-end business.