FOR: Infrastructure
& Operations
Professionals

Server And Data Center Predictions For 2013

by Richard Fichera, March 8, 2013

KEY TAKEAWAYS

Newer Workloads Will Expand Faster

While some workloads report higher growth than others, Forrester has not seen sign of flat or negative growth in legacy processing requirements. Newer workloads like analytics, big data, and online web services such as search, social/digital media, and gaming are driving requirements for computing, storage, and networking at an accelerating pace.

There Is An Ongoing Need To Maintain Operational Efficiency

Operational and financial pressures continue along with an increasingly complex regulatory environment. I&O groups are bracing themselves for the implications of the new Dodd-Frank Wall Street Reform and Consumer Protection Act, which will trigger new requirements with its more than 200 new rules regulating financial services firms.

Continually Execute Infrastructure Transformation Initiatives

The pace of infrastructure transformation as part of the process of providing new capabilities continues, although it has slowed slightly from the previous year. These efforts continue alongside and as part of efforts to virtualize and standardize the data center environment and embrace cloud technologies.

Fulfill The Need For Long-Term Capacity And Visibility

As processing requirements escalate, many organizations are discovering that they need to pay attention to long-term capacity planning and make better use of their existing facilities. Many IT organizations will deem it more important to optimize data center capacity and resource consumption to avoid costly new build/buy projects.

Server And Data Center Predictions For 2013

by Richard Fichera
with Christopher Voce and Eric Chi

WHY READ THIS REPORT

This report outlines Forrester's solution for infrastructure and operations (I&O) executives working on data center infrastructure. This report is designed to help you build a road map for 2013. As we transition from 2012, the landscape is a combination of the familiar and the new. The drumbeat of new servers and new x86 CPU offerings will continue, with announcements from all of the major players. Forrester also expects to see the early commercialization of ARM CPUs, which will offer additional options for energy-efficient processing. On the data center front, expect that the pressures on power-efficient operations will continue unabated, driving adoption of emerging data center infrastructure management (DCIM) software. Modular data center options will proliferate as entry prices decline and additional options appear. This report focuses on eight key predictions that I&O leaders should consider when planning for their server and data center infrastructure moves.

Table Of Contents

Notes & Resources

Forrester gathered information across many client inquiries and industry briefings.

Related Research Documents

Updated Q3 2011: Power And Cooling Heat Up The Data Center
September 21, 2011

Updated Q2 2011: Are Converged Infrastructures Good For IT?
June 22, 2011

Put ARM-Based Servers On Your Server Planning Horizon
March 15, 2011

SERVER AND DATA CENTER PREDICTIONS THAT WILL SHAPE THE MARKET IN 2013

It's time polish off our crystal ball again and look at the remainder of the year. In many ways, our 2013 predictions are a continuation of those in 2012. The data center infrastructure picture is one that must, by its nature, change in an evolutionary rather than revolutionary fashion. Here is what we expect over the next year — and what I&O pros should do about it.

1. The Next Major x86 Server Refresh Cycle Will Come Late In The Year

The underlying heartbeat of server technology — semiconductor improvements — will continue unabated in 2013, with an aggressive refactoring of Intel's Xeon CPUs and Advanced Micro Devices' (AMD's) attempt to counter with its own innovations. The most influential event — the volume introduction of the next-generation CPUs and servers based on Intel's Ivy Bridge architecture — will likely occur late in the year. These CPUs will represent a unification of Intel's server line around the Ivy Bridge architecture with a 22 nm implementation. The initial releases will overlap the positioning of the current E5 CPUs, and subsequently the high E7 CPUs will make a major jump to both a new architecture and a 22 nm process. Our expectation is that core counts will rise modestly and some of the additional real estate will be used for embedded accelerators such as cryptography, virtual machine (VM) acceleration, and additional security features already introduced in other Ivy Bridge chips.[1] Energy efficiency will get an additional boost from the superior gate design of Intel's 22 nm process.[2]

2013 will also see volume shipments of servers based on Intel's recently announced Centerton Atom-S CPUs, low-power CPUs targeted at low-power dense servers intended for hyperscale data centers. These 22 nm CPUs will be the first real SOC products, containing all required elements for a complete entry server except for memory and network, and should improve Intel's competitive position relative to emerging ARM products.

Expect major systems vendors to have systems ready to ship at product introduction, for the most part minor revisions of the most recent cycle of server products in the 1S through 4S segments. Larger systems may see more aggressive innovation, as they were largely bypassed in the last major round of server architecture innovation in 2012.

- **What I&O needs to know:** These processors will feature substantially improved performance per watt and improvement in absolute performance at the same clock speed. They will also feature more sophisticated power management capabilities that are accessible from the vendor's management consoles.

- **What I&O needs to do about it:** Once again, Forrester expects the power efficiency gains from the upcoming introductions to be significant enough that I&O groups may want to evaluate earlier upgrades of in-place servers to gain these efficiencies — especially those who've hit or will soon run into the power limits available to their racks but need more capacity. Forrester strongly recommends that I&O groups perform this analysis on their workloads as soon as is practical.

2. ARM-Based Servers Will Be Slower To Appear From Major Vendors Than Predicted

Our previous prediction was that ARM-based servers would generate high volumes in 2012, but the development of the market and ecosystem has been even slower than we expected due to two major developments: 1) the slower-than-expected progress on a full 64-bit ARM server CPU, which pushes off mainstream server adoption of ARM until full 64-bit systems can be evaluated, and 2) the focus of major vendors such as HP and Dell on their near-term efforts on servers based on Intel's new Centerton Atom SOC offerings, which offer 100% x86 code compatibility combined with what Forrester expects to be significant improvements in power per core and workload per watt compared to previous x86 CPUs.

The ARM providers countered with the Cortex A53/A57 architecture, a full-featured 64-bit server CPU. Unfortunately, it will not be available until late 2013 at the earliest, with systems probably not appearing until 2014. At the same time, the ARM ecosystem has received a major boost from AMD, which announced that it has taken out an ARM architecture license and also purchased SeaMicro, an innovative fabric-based server.

- **What I&O needs to know:** ARM servers are no longer science fiction, and if you want a system for investigative and development purposes, you will be able to buy one from one or more of the smaller server vendors or possibly latch onto one of a limited supply of Redstone systems from HP's developer's program. Real production systems will probably have to wait until 2014 when full 64-bit systems are available.

  Forrester's take is that if the 64-bit ARM architecture appears on schedule and offers verifiable benefits over Intel in the form of better than two to three times performance per watt, then it still has a viable future. Otherwise, ARM technology will continue to be relegated to its current strongholds in embedded and mobile devices — where "relegated" is a pejorative term for being incredibly successful.

- **What I&O needs to do about it:** Initially, the burden is on the application development function to evaluate the suitability of these new systems, but as these investigations mature, I&O will have to get involved to understand the implications. In terms of timing, there is no need to focus significant I&O resources on this until mid-to-late 2013.

3. AMD: A Multi-Pronged Approach — But Odds Are Against It

AMD put some price/performance pressure on Intel with its recent Magny-Cours CPUs, which pack up to 16 cores. However, this offering has been an overall disappointment in terms of raw performance for mainstream server applications due to deficiencies inherent in the Bulldozer core. AMD has announced its anticipated Piledriver core as the successor to the Bulldozer. With Intel's clear march to new 22 nm server chips in 2013, unless Piledriver greatly outperforms Bulldozer, AMD is going to find it hard to compete for mainstream server business — and may be relegated to niches where high core-count at low power and price make it competitive.

March 8, 2013

The two unknowns in making accurate predictions about AMD's future as a server vendor are its integrated CPU/GPU technology used on its desktop and laptop products and its execution of its ARM strategy. Overall, we think the most positive scenario will be that AMD maintains its position relative to Intel, and that significant share gains are unlikely in 2013.

- **What I&O needs to know:** AMD will survive as the No. 2 server CPU vendor. Aside from performance differences driven by the differing number of cores and performance of the cores, AMD products will remain completely software-compatible with Intel products.

- **What I&O needs to do about it:** Like we suggested last year, reassess your Intel versus AMD server investments in light of the new announcements, and for large procurements, don't be afraid to play one vendor off the other. In the final analysis, there are few workloads that can't be handled adequately by either vendor.

4. Hybrid Scalar/GPU Computing Will Become Mainstream

In 2010, we saw the introduction of general purpose graphics processing unit (GPGPU) technology as an adjunct to conventional computing architectures. GPUs are highly effective for some computing tasks, once the high cost of programming them is overcome. By the end of 2011, all the necessary elements were in place for an accelerated acceptance into mainstream computing: ubiquitous hardware support by all the major systems vendors, mature software development environments (despite some vendor-driven splits between OpenCL and Cuda), and at least two major suppliers in Nvidia and AMD. The penultimate step in the mainstreaming of GPUs was the successful introduction of AMD's Fusion desktop processors, which combine an x86 scalar core with an integrated GPU. This served as a catalyst for further GPU development tool activity and raised the awareness of GPUs as a potential solution across a wide range of developers.

2012 saw further maturation of the ecosystem and the introduction of the Intel Xeon Phi, an x86 compatible many-core accelerator that is intended to compete with Nvidia and AMD Fusion. Intel's offering lags behind its competition in raw performance but has the advantage of being compatible with current x86 codes. As a tangible token of GPU implementation becoming more mainstream, the No. 1 computer on the top 500 list is now a hybrid system with x86 and GPU processors.

- **What I&O needs to know:** All major vendors offer servers with integrated GPUs, in both rack and blade or dense rack configurations. There are also better development tools and skilled programmers to take advantage of GPUs. The initial burden is on the internal app dev groups and the independent software vendor (ISV) community, but for any application that has inherent parallelism, GPUs offer major potential performance gains.

- **What I&O needs to do about it:** I&O should work with the major application stakeholders to determine which applications can be rewritten to take advantage of GPU capabilities. Many ISVs, including Matlab, Mathematica, MSC Nastran, Ansys, Abaqus, and Autodesk, currently

support GPUs, and I&O groups should work with major application stakeholders to identify potential candidates. The introduction of the Xeon Phi offers a path to rapid migration at the sacrifice of peak potential performance.

5. Converged Infrastructures Will Continue To Evolve

All current vendors have updated their core portfolios of blade and converged infrastructures (CIs) with new offerings incorporating the latest Intel technology. Last year, we predicted that HP and IBM would make major architectural refreshes of their products, and we were half-correct, as IBM announced its new PureSystems products, a major architectural overhaul of its blade products line. HP's recent announcements in 2013 complete the tier one system vendor changes to its blade product lineup this year in preparation for Intel's next major server CPU deliveries.

In addition to changes in the core blade server offerings, we expect an acceleration of recent momentum behind storage-centric converged infrastructure, particularly solutions optimized for virtualized environments, particularly as the base technology for private cloud deployments. Currently the purest exemplars of this trend are Nutanix and SimpliVity, both of which have announced innovative new modular architectures that dramatically simplify the provisioning and scaling of VMware clusters by means of innovative storage management technology and simplified physical VM host server cluster scalability. Other storage-centric solutions have also appeared, including an impressive offering from HDS. Taken as a class, these systems will offer:

- **More granular abstraction of physical components.** Cisco one-upped all incumbent vendors with its UCS architecture, which offered a richer presentation of physical components as software objects, and IBM's new Flex offerings have the best storage integration story of the major vendors.

- **Increased use of integrated fabrics for system scaling.** We can see clues about future server implementation by looking at the HP Redstone server, unveiled as part of its ARM-based server announcement. Expect the passive backplane architecture of current blades to give way to a more scalable high-speed fabric, with more flexible topologies for aggregation and management. Intel's public statements outlining its fabric technology strategy add to the likelihood that fabric technology will come down in price and implementation complexity.

  The systems from Nutanix and SimpliVity also reflect this trend toward integrated fabrics, with the systems being hooked together via a 10 Gb Ethernet fabric that handles all storage traffic between nodes.

- **Even more focus on power and cooling.** Power and cooling efficiency was one of the primary values offered by the early generations of blade servers, and Forrester expects that this focus will continue. Blade enclosures offer more opportunities to optimize power and cooling than a conventional rack server because of their integrated design.

- **Better systems management.** Expect vendors to improve the current blade server management tools to accommodate the more complex software environments of future systems. While all of the vendor implementations will differ, Forrester expects that one common thread running through them will be federation of multiple system enclosures into a seamless management presentation. IBM has made major strides with its pure systems, and products from Nutanix and SimpliVity set a new bar for simplicity of scaling clusters and for storage management in a VMware environment.

- **Continued focus on opex improvements.** A major theme in 2012 was improved opex via simplification of the provisioning, scaling, updating, and management of the CI. Forrester expects this trend to continue, with major investments in these areas by all players.

Above and beyond the underlying hardware and system management tools, the vendors will compete to extend the CI concept to include composite application services, presented as abstractions including servers, network, storage, and security policies. HP, with its CloudSystem portfolio, is a prime example of this trend and will continue to enhance this already strong offering. IBM's PureApplication systems, which leverage application, OS, and infrastructure patterns, now offer a strong alternative to HP's CloudSystem.

- **What I&O needs to know:** The major challenge for I&O groups will be understanding the complexities of the available options and coordinating with all the affected stakeholders in the organization. The complexity of CI, along with its inherent lack of standardization above the basic VM layer, tends to lead to vendor lock-in. This isn't necessarily bad as long as the tradeoffs for lack of vendor choice are mitigated by tangible benefits from the solution.

- **What I&O needs to do about it:** Aggressively pursue your current vendor for updates and visibility to futures — the sheer volume of new product technology in the CI space will be difficult to track. Where there is the opportunity for a major architectural change to infrastructure as well as applications, there is also an opportunity to introduce a new cycle of competition into your environment. Additionally, new offerings will give you more choices for lowering opex.[3]

6. DCIM Solutions Will Proliferate, Ushering In An Age Of Confusion

In 2011 and years prior, we saw the emergence of the first generation of DCIM solutions. In 2012, the major data center infrastructure vendors Schneider and Emerson rolled out their integrated DCIM solutions, and Forrester expects 2013 to see the rapid shift in market share toward these major players at the expense of smaller vendors.

DCIM solutions offer data center operators the ability to take a holistic look at their entire physical data center environment and to not only understand what is happening in real time but also make intelligent decisions about future changes and investments. Forrester defines a rough taxonomy of functions for DCIM software as follows:

- **Inventory and discovery.** The solution collects real-time physical infrastructure data, including physical rack and equipment location, circuit capacities, CRAC/cooling data, and thermal data. To perform these functions, the software vendor needs to collect data from the equipment of multiple and, in many cases, competing vendors, along with data from other sources, including other CMDBs. Prospective purchasers need to understand in detail how the initial facility discovery and ongoing updates will be handled.

- **Continuous data collection.** Once the DCIM software is installed, it needs to collect data from the environment on a continuous basis. The granularity of the data collection, the type of database in which it's stored, and the extent to which it can be used for other analyses are characteristics that differentiate the vendors.

- **Consolidated display.** DCIM solutions offer a display and reporting of real-time data, trends, and exceptions. One of the earliest functions of DCIM software was as a unified dashboard for displaying a single-source view of power, physical, and thermal information. Bridging the worlds of facilities and IT operations was one of the initial selling propositions of the first-generation DCIM applications and continues to be a focus of many DCIM products.

- **Trend analysis.** The next major step in functionality is the ability to display trends for power and cooling. This will offer data center operators the ability to better predict and diagnose conditions that have exceeded preset limits.

- **Model solutions for implementation.** This is still a work in progress for most vendors, but the ability to model potential solutions to power and workload problems is one of the areas where the systems vendors and infrastructure equipment suppliers intersect. To effectively propose an optimal set of allocations, the software needs to understand the behavior of the system at a granular level, the potential workloads, and the details of the power and cooling environment. High-end IT systems management ISVs, such as HP, CA, and BMC Software, have traditionally offered this, but it's appearing in many of the DCIM offerings as well. The degree to which the potential solution simulations are aware of the underlying workloads and the accuracy and granularity of the simulations are all differentiators.

The universe of DCIM vendors is evolving and expanding — more than 60 vendors currently identify themselves as supplying DCIM solutions, with an average of two or three new ones cropping up each month. If DCIM follows the same trajectory as other emerging solution spaces, we can expect considerable dropout of smaller players combined with the mergers and acquisitions (M&A) activity that is already occurring. The integration of DCIM with conventional management tools is a work in progress, and the competitive battle among DCIM vendors will likely be defined by the degree to which this integration can be made seamless and useful for I&O professionals. The challenge for I&O professionals will be separating myth from reality in a rapidly changing vendor landscape:

■ **What I&O needs to know:** The need for DCIM is not new — problems with managing the instantaneous operational aspects of power and cooling and planning for future changes have been around for years. What is new is the availability of single solutions that now begin to merge the world of facilities, I&O, and even application service delivery. At a minimum, I&O groups need to forge new working relationships with facilities to begin to realize the power of these solutions.

■ **What I&O needs to do about it:** I&O groups should be prepared to invest time in evaluating and selecting a DCIM offering. Forrester believes that the benefits are real — such efforts will be rewarded with incremental power and cooling efficiencies ranging from 5% to 30%, depending on the initial state of the environment.

7. Modular Data Center Options Will Continue To Evolve

Even after upgrading systems and optimizing power and cooling, eventually there comes a time when new data center space is needed. Either as a result of major corporate changes such as M&A, growth of current business requirements, or new lines of business, many enterprises can be expected to outgrow a current facility in the next three to five years. The story has been the same for the past several decades. What has changed are the options available for new data center capacity. To the old menu of hosting providers and colocation we now have two new major options — cloud and modular data centers.

Leaving aside cloud for the moment, which in many ways implies a more fundamental architectural change and is not mutually exclusive with new data center requirements, modular data centers represent a significant shift in provisioning new data center capacity. Forrester expects modular data centers to continue their rapid growth in 2013, with the end-of-year percentage of new data center spend in the range of 10% to 15%. A continuation of 2012 trends, offerings will continue to diversify, with more choice in "room within room" offerings, and smaller modular entry points for self-contained pod offerings, along with enhanced management software as differentiators:

■ **What I&O needs to know:** Building new data center capacity is a major undertaking for most enterprises and involves almost the entire chain of corporate stakeholders, including business units, finance, facilities, I&O, and executive management. For I&O, the most commonly missing element is accurate visibility into future requirements.

■ **What I&O needs to do about it:** I&O needs to build long-range capacity planning processes into IT and encourage business units to participate. Modular data centers' two stellar benefits will remain intact in 2013: They are somewhat cheaper than conventional data centers and much, much faster to acquire. A new pod can be ordered and delivered from all of the major suppliers in anywhere from six to 12 weeks with infrastructure installed in 18 to 24 months for conventional data center space.

8. The SDDC Will Gain Momentum As An Organizing Metaphor For Virtual Infrastructure

Forrester believes that the concept of the software-defined data center (SDDC), originally popularized by VMware, will gather momentum as an organizing concept for complex virtualized infrastructures. The SDDC, as discussed in Forrester's I&O IT infrastructure playbook, will provide a completely software-based abstraction of the data center, including servers, storage, and networks.[4]

- **What I&O needs to know:** The SDDC promises to offer an integrated architecture that allows the merger of legacy architectures, cloud computing, and workload-centric architectures into a single, manageable architecture. I&O professionals need to continue to stay up to date on the evolution of virtual infrastructure.

- **What I&O needs to do about:** This is a long-term evolution, and initially all that will change are some of the names of products from familiar vendors. Longer term, major players such as VMware, Microsoft, and others will deliver more technology that plugs into the concept of a software-defined data center. I&O groups need to track the competitive position of the major players, particularly VMware and Microsoft, and within the next 24 months should have a plan for how their organizations plan to implement major components of software-defined data center for their own operations.

ENDNOTES

[1] Intel has discussed a feature called APICv that will improve the efficiency of hypervisors by reducing the number of times the hypervisor code must exit to process an interruption, resulting in lower CPU overhead as well as reduced latency, both of which translate into higher performance for the VM workloads.

[2] For more information on energy efficiency, see the September 21, 2011, "Updated Q3 2011: Power And Cooling Heat Up The Data Center" report.

[3] For more information on lowering opex, see the June 22, 2011, "Updated Q2 2011: Are Converged Infrastructures Good For IT?" report.

[4] Over the last two decades, IT infrastructure and operations (I&O) teams have been in a race with complexity. As business requirements have mounted, generations of technology have been implemented to try and stem the tide. The result: data centers characterized by complex, static physical configurations, incomplete virtualization efforts, and multiple silos of management. But the emergence of better virtualization management tools, software-defined networks, and converged infrastructure stacks has opened the potential to what Forrester calls the software-defined data center (SDDC) — a comprehensive abstraction of a complete data center. For more information, see the November 12, 2012, "The Software-Defined Data Center Is The Future Of Infrastructure Architecture" report.

About Forrester

A global research and advisory firm, Forrester inspires leaders, informs better decisions, and helps the world's top companies turn the complexity of change into business advantage. Our research-based insight and objective advice enable IT professionals to lead more successfully within IT and extend their impact beyond the traditional IT organization. Tailored to your individual role, our resources allow you to focus on important business issues — margin, speed, growth — first, technology second.

FOR MORE INFORMATION

To find out how Forrester Research can help you be successful every day, please contact the office nearest you, or visit us at www.forrester.com. For a complete list of worldwide locations, visit www.forrester.com/about.

CLIENT SUPPORT

For information on hard-copy or electronic reprints, please contact Client Support at +1 866.367.7378, +1 617.613.5730, or clientsupport@forrester.com. We offer quantity discounts and special pricing for academic and nonprofit institutions.

Forrester Focuses On
Infrastructure & Operations Professionals

You are responsible for identifying — and justifying — which technologies and process changes will help you transform and industrialize your company's infrastructure and create a more productive, resilient, and effective IT organization. Forrester's subject-matter expertise and deep understanding of your role will help you create forward-thinking strategies; weigh opportunity against risk; justify decisions; and optimize your individual, team, and corporate performance.

« IAN OLIVER, client persona representing Infrastructure & Operations Professionals