Sponsored by: **SGI and Intel**

Authors:
Peter Rutten
Matthew Marden

March 2016

The Value Proposition of Scale-Up x86 Servers: Cost-Efficient and Powerful Performance for Critical Business Insights

EXECUTIVE SUMMARY

It is becoming increasingly clear that scale-up server solutions not only represent a comparable alternative to scale-out or distributed environments but also, for a variety of reasons, can have a demonstrably beneficial impact on an organization's ability to gain insights and compete more successfully. More and more organizations — from small and medium-sized businesses (SMBs) to the largest enterprises — are deploying compute-intensive and data-intensive applications to analyze huge data sets so as to obtain new knowledge and insights and build innovative products or marketing strategies around them. Quite a few of these applications have traditionally resided on HPC clusters, and others have been deployed in distributed environments. But the limitations of clusters to perform analytics on very large data sets in-memory and scale without severe latency implications caused by the cluster networking have led to a revived interest in scale-up platforms.

This IDC study on the business value of scale up shows that scale up can yield improved performance, including increased resource utilization, improved application performance, less unplanned downtime, and extended datacenter life. Furthermore, we found that scale-up consolidation can result in combined savings of 33% thanks to operational cost reductions as a result of reduced server management staffing requirements and lower costs associated with power and cooling, software licensing, and IT infrastructure. This study also looked at the SGI UV 3000 and SGI UV 300 — two powerful scale-up systems that have the scalability and memory capabilities to execute some of the most demanding compute- and data-intensive workloads.

Situation Overview

Scale Up Versus Scale Out Defined

To meet their growing needs for workloads, IT organizations can expand the capability of their x86 server infrastructure by either scaling up by adding fewer, more capable servers or scaling out by adding multiple relatively less capable servers:

- **Scale up.** Scale up is achieved by putting the workload on a bigger, more powerful server (e.g., migrating from a two-socket server to a four- or eight-socket x86 server in a rack-based or blade form factor). This is a common way to scale databases and a number of other workloads. It has the advantage of allowing organizations to avoid making significant changes to the workload; IT managers can just install the workload on a bigger box and keep running it the way they always have.
- **Scale out.** Scale out refers to expanding to multiple servers rather than a single bigger server. The use of availability and clustering software (ACS) and its server node management, which enables IT managers to move workloads from one server to another or to combine them into a single computing resource, represents a prime example of scale out. Scale out usually offers some initial hardware cost advantages (e.g., four 2-socket servers may cost less than a single 16-socket server). (IDC notes that each socket may have more than one processor.) In many cases, the redundancy offered by a scale-out solution is also useful from an availability perspective. However, IDC research has also shown that scale-out solutions can drive up opex to undesirable levels. At the same time, large data volumes and required processing capabilities are taxing scale-out systems.

Scale out refers to expanding to multiple servers rather than a single bigger server.

As organizations choose to entrust an increasing number of demanding, mission-critical workloads to scale-up servers, they must make sure that they are not introducing single points of failure at the server level for these critical applications. This means that customers should choose servers that incorporate advanced reliability, availability, and serviceability (RAS) features. RAS servers incorporate features that improve resilience by allowing processing to continue in the case of hardware failure, ensure the attainment of agreed-upon uptime service levels, and provide improved diagnostic and error detection tools to enable them to be maintained more quickly and easily.

Scenarios for Scale-Up Server Deployments

The ability to perform advanced business and technical computing has become critical for most businesses. Compute-intensive applications, such as computer-aided engineering (CAE), next-generation sequencing, scientific simulations, and advanced analytics, as well as data-intensive applications, such as SAP HANA, Oracle Database In-Memory, data analytics, visualizations, and real-time streaming, are becoming core competitive competencies for organizations across all industries. Given the complex computations these applications perform and the data volumes they need to manage, they require many processors and large amounts of memory.

Compute-intensive applications have been running on high-performance computing (HPC) clusters, scaling out to however many nodes are needed to perform massive sequencing, engineering, or simulation tasks.

Compute-intensive applications have been running on high-performance computing (HPC) clusters, scaling out to however many nodes are needed to perform massive sequencing, engineering, or simulation tasks. Data-intensive applications, like large databases, may be running on legacy scale-up platforms or on distributed environments. In both cases, performance is increasingly impeded as data sets grow — either because of the limitations of clusters and distributed environments, which require performance-draining networking, or because of the latencies that result from reading data from and writing to disks and/or flash.

Both types of applications can get a tremendous performance boost from recent technologies that allow an entire data set to be held in memory, instead of on a disk, while performing computations on the data. This is especially true for a scale-up system in which large amounts of memory are placed close to the processors, which are connected via advanced connective topologies.

Some of these applications therefore lend themselves particularly well to scale-up rather than scale-out deployments. IDC has identified three groups of workloads that we believe can benefit significantly from scale-up use-case scenarios.

Compute-Intensive Applications

- **Computer-aided engineering.** The various processes involved in CAE are extremely compute intensive with increasingly larger data sets and more and more variables that must be addressed in the engineering design, as well as ever more complex analyses performed and a growing demand for better visualizations of the results.
- **Next-generation sequencing.** Next-generation sequencing is essentially defined by throughput, scalability, and speed, as opposed to traditional DNA sequencing, which was slow and costly. The compute intensity of modern complex genome research, such as sequencing an entire genome or discovering new RNA, requires the ability to place huge amounts of data in-memory and process the data with a large number of processors that are closely interconnected in a single system.

- **Scientific simulations.** The sheer breadth of scientific simulation solutions that scientists and businesses use to resolve academic and product development problems is indicative of how pervasive simulation techniques have become. Their sophistication and their ability to simulate the most complex real-world systems, whether economic, biological, geographical, or of some other nature, in all their facets are rapidly increasing. And the demands to perform these simulations faster and faster are persistent.

- **Advanced analytics.** Advanced analytics focuses on predicting behaviors, trends, and phenomena, thereby enabling businesses to adjust their business strategies accordingly. Advanced analytics is used widely by marketing departments, for example, and it is being adopted by businesses in all industries. The statistical complexities of performing predictive analytics or risk management on huge data sets in near real time require a platform that is not bogged down by processing power or scaling limitations.

Data-Intensive Applications

Increasingly, sophisticated data analytics on large data sets must be performed on real-time or near-real-time transactional data rather than on older data lifted out of a database.

- **Data analytics.** Increasingly, sophisticated data analytics on large data sets must be performed on real-time or near-real-time transactional data rather than on older data lifted out of a database. The ability to keep transactional data in memory and execute computations on that data in a scale-up environment provides significant advantages.

- **SAP and SAP HANA.** SAP HANA is an in-memory database that is designed to perform transactions and analytics on the same platform in real time. SAP HANA requires an in-memory platform, but the database can run on scale-up and scale-out systems (e.g., for Business Warehouse). However, running Business Suite on HANA — or running the next-generation SAP S/4HANA applications — is best done on a scale-up system.

- **Oracle Database In-Memory.** Oracle Database software products support both scale-out and scale-up computing. Deploying Oracle Database 12c on a scale-up in-memory server provides noteworthy performance benefits.

- **Visualization.** Data visualization is a burgeoning component of big data analytics across all disciplines, whether purely academic or within the context of industrial R&D. Visualizations have taken on a critical role in the human understanding of large sets of complex data, and generating visualizations quickly, for very large data sets, can require a platform with extreme processing and scaling capabilities.

» **Real-time streaming — sensors, satellites, self-driving cars.** The onslaught of data streams that the Internet of Things will spawn and the opportunities for insights and capabilities that lie buried within them are enormous. The ability to capture data streams, analyze them, and detect unique moments or anomalies within them is rapidly becoming a critical competency for academic and commercial organizations. The largest raw data streams may be managed in a cloud, but the subsets on which real-time analyses are performed will be processed more efficiently in an in-memory, scale-up environment.

Custom Software Applications

Custom applications, many of which are legacy applications running on aging hardware, still drive many mission-critical workloads in enterprises today.

Custom applications, many of which are legacy applications running on aging hardware, still drive many mission-critical workloads in enterprises today. As demands on these applications grow in terms of data volumes, performance requirements, and RAS features, the underlying hardware cannot keep pace. As custom applications run up against system limitations, IT organizations are faced with a critical decision: whether to scale up or scale out. For several years, IT elected to migrate to a scale-out environment, only to be faced with significant increases in opex while being unable to sustain the performance and RAS that legacy systems had provided. Some of these organizations are now switching back to scale up, while others are leapfrogging the scale-out scenario and going directly from legacy scale-up systems to modern scale-up platforms.

Measuring the Business Value of Scale-Up X86 Servers

IDC's research demonstrates that organizations can achieve cost savings and staff time efficiencies when they run the right workloads on scale-up servers. A comparative analysis of the average costs of using fewer larger servers instead of more distributed server environments shows that organizations can operate such workloads at a 33% average lower cost.

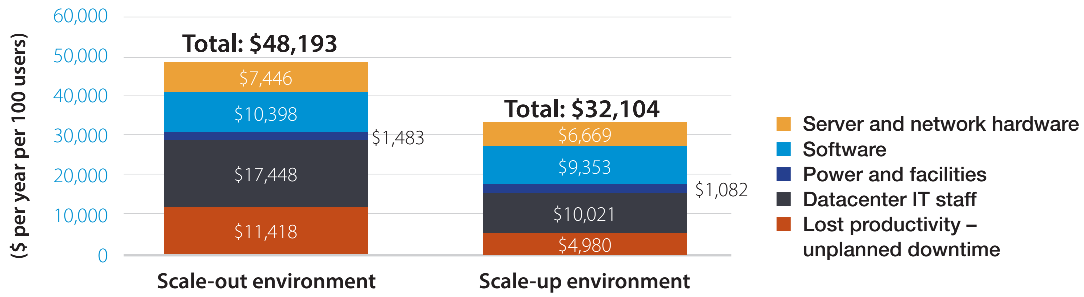
For this study, IDC considered research conducted from 2012 to 2015 with organizations that decided to run certain workloads on larger scale-up servers rather than smaller, more distributed servers. These organizations either compared the costs to run workloads on scale-up and more distributed servers or provided insight on their migrations of workloads from distributed to larger servers. Interviews conducted for this research covered a variety of qualitative and quantitative topics designed to understand the impact on these organizations of their choice of scale-up servers in terms of infrastructure and datacenter costs, IT staff time requirements, impact of unplanned downtime, and ability of scale-up servers to support business operations.

IDC's Business Value research found that these organizations have reduced costs and achieved value in all of these areas with scale-up servers, despite spending more for scale-up servers on a per-server basis. By enabling these organizations to run certain workloads in a consolidated server environment, these larger servers have provided cost and staff time efficiencies. Further, their reliability and high availability have contributed to making users and business operations more effective by minimizing the frequency and duration of outages to systems and business applications.

Figure 1 presents IDC's findings with regard to efficiencies for these organizations of running certain workloads in a scale-up environment compared with a more distributed environment on a per-100-user basis. It demonstrates that the organizations considered for this study are supporting these workloads with scale-up servers at an average cost of $32,104 per year per 100 users, which is 33% less than the cost of running the same workloads in a comparable distributed server environment ($48,193 per year per 100 users).

FIGURE 1

Annual Cost of Scale-Up Servers Versus Scale-Out Servers per 100 Users in Selected Environments



	Scale-Out Environment	Scale-Up Environment	% Difference
Server and network hardware	$7,446	$6,669	10%
Software	$10,398	$9,353	10%
Power and facilities	$1,483	$1,082	27%
Datacenter IT staff	$17,448	$10,021	43%
Lost productivity — unplanned downtime	$11,418	$4,980	56%
Total	$48,193	$32,104	33%

Source: IDC Business Value Research, 2012-2015

Notes:

- **Server and network hardware** includes the annualized cost of server and network hardware taking into account initial costs and replacement cycles, as well as ongoing maintenance fees.
- **Software includes fees** for software licensing and provider maintenance and support.
- **Power and facilities** includes the cost for power and cooling, datacenter space, and so forth for the server environments being considered.
- **Datacenter IT staff** includes the cost of IT staff time to manage and implement all aspects of datacenter operations for the workloads running on the server environments being considered.
- **Lost productivity — unplanned downtime** includes the cost of lost productive time for users caused by system/application outages and periods of unavailability.

Organizations reported that they are achieving these cost savings and time efficiencies because scale-up servers meet the demands of high-performance workloads better than more distributed server environments. Specifically, IDC's research shows that organizations can leverage scale-up servers to optimize their hardware, software, and ongoing datacenter costs; reduce the amount of IT staff time needed to support and manage these environments; and minimize the impact of unplanned downtime on users and business operations.

IT Infrastructure Cost Efficiencies

Organizations interviewed by IDC reported that they are able to reduce their datacenter infrastructure and operations costs with scale-up servers despite spending more on purchasing servers on a per-server basis.

Organizations interviewed by IDC reported that they are able to reduce their datacenter infrastructure and operations costs with scale-up servers despite spending more on purchasing servers on a per-server basis. They achieve savings because scale-up servers provide more compute and better virtualization capabilities than the distributed servers they considered, meaning that their scale-up environments support workloads substantially more efficiently on a per-server basis. Specifically, organizations indicated that the strong performance and virtualization capabilities of larger servers are helping them support selected workloads at a cost that is lower than that of scale-out solutions. IDC's research shows that these organizations spend an average of 10% less per year as a result of benefits in the following areas:

- **Physical servers.** Scale-up servers are more powerful, must be replaced less frequently, and support higher virtualization density on average than more distributed servers. As a result, organizations need fewer servers to support equivalent workload environments. IDC's research shows that organizations considered for this study run the same workloads on 30% fewer scale-up servers than they would have needed with other servers. This smaller server footprint carries through for these organizations to further savings in other datacenter hardware, including network hardware. In total, IDC found that organizations spend an average of 10% less per year on server and network hardware with their larger servers.
- **Software.** Consolidating server environments on larger servers also helps organizations optimize software licensing costs. Licenses are sometimes priced on a per-CPU basis, so more distributed server environments can increase overall licensing costs and reduce licensing use efficiency. IDC calculates that organizations considered for this study spend an average of 10% less per year on software licenses than they would with a more distributed server footprint.

- **Datacenter operational costs.** Larger servers can also help organizations reduce their power and datacenter space costs. As discussed, having a consolidated server footprint can mean that organizations use less power and space overall, and many larger servers are designed to be as power and space efficient as possible for running demanding workloads.

IT Staff Time Efficiencies

IDC's research shows that the IT staff time needed to deploy, support, and deliver business applications is often at least as costly as hardware and software.

IDC's research shows that the IT staff time needed to deploy, support, and deliver business applications is often at least as costly as hardware and software. As a result, maximizing the efficiency of IT staff time is critical to maintaining the lowest possible cost of operations and to freeing up staff to focus on supporting the business or driving IT innovation. IDC's research shows that organizations can leverage scale-up servers to limit the amount of IT staff time needed to support business applications compared with more distributed server environments. IDC has found that organizations require 43% less staff time on average per 100 users to run selected workloads with scale-up servers than with more distributed servers.

Organizations attributed IT staff efficiencies with larger servers to a number of factors, including:

- **Fewer servers.** Running workloads on scale-up servers reduces the total number of processors and servers that an organization must manage and support. Because there are tasks, including patching, tuning, and maintenance, that IT teams generally must perform for each physical server within their datacenter environments, running the same workloads on fewer larger servers frees up staff time that can be reinvested in supporting business operations and other higher-value tasks.
- **Optimized workloads.** Scale-up servers are often designed for high performance for specific, demanding workloads, which can mean that they experience fewer problems. In addition, larger servers face fewer capacity limitations than less powerful servers, reducing the staff time needed to deploy additional compute resources and reallocate workloads.
- **Virtualization density.** Scale-up servers are often able to accommodate higher virtualization densities, which further reduces the physical infrastructure environment that IT teams must manage. In addition, because larger servers are often designed to support virtualized workloads, organizations reported to IDC that the process of extending and managing virtualization is less time consuming on a per-workload basis in their scale-up environments than on other servers.

IDC Analyze the Future

Using scale-up servers for the right workloads can also reduce the impact of unplanned downtime and other outages on employees and business operations.

Risk Mitigation — Availability

Using scale-up servers for the right workloads can also reduce the impact of unplanned downtime and other outages on employees and business operations. On average, organizations considered for this study reported that unplanned downtime exerts a 56% lower cost in terms of lost productive time with scale-up servers than with more distributed servers. In addition, scale-up servers have helped these organizations limit the potential loss of revenue that can occur during interruptions to business processes and services. IDC's research shows that organizations benefit from scale-up servers in the following ways:

» **Fewer points of failure.** Scale-up server environments run the same workloads on fewer servers, limiting the number of potential failure points and translating to fewer outages. However, it should be noted that this can increase the potential impact of server-related downtime when it does occur in scale-up server environments by impacting more users.

» **Designed for high availability.** Scale-up servers are often designed to run workloads that require the highest possible levels of availability and uptime. As a result, they are often designed to meet stringent availability standards and service-level agreements (SLAs) with failover capabilities and other fault-tolerant features.

SGI UV In-Memory Scale-Up Servers

The SGI UV platform, built on Intel Xeon processors, is a powerful single system image server. It is particularly efficient for compute- and data-intensive applications such as computational biology, genome reconstruction, high content analysis, and systems biology. The system's overall performance and its ability to efficiently run very demanding workloads, whether they are ongoing tasks or new tasks designed to obtain innovative insights, can result in an advantageous ROI for customers. With directly or SAN attached RAID, MAID, active archive, or JBOD storage solutions from SGI and other leading vendors, SGI UV enables immediate access to any number of CPUs on the system, delivering maximum computational dynamic range thanks to an all-to-all NUMAlink — all processors interconnect over direct links — or enhanced hypercube topology, which results in much lower and more consistent memory access latencies (see Figure 2).

The SGI UV 3000 was designed for the kind of compute-intensive applications mentioned previously, such as computer-aided engineering, genetic sequencing, and scientific modeling.

SGI UV 3000

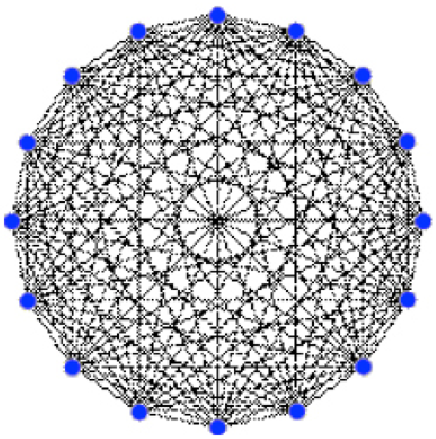
The SGI UV 3000 was designed for the kind of compute-intensive applications mentioned previously, such as computer-aided engineering, genetic sequencing, and scientific modeling. The system has extreme scaling capabilities and can run the most demanding technical applications. At the heart of the system lies the SGI NUMAlink 6 ASIC interconnect, one per blade (there are eight blades), which links a large number of processors via high-bandwidth connections. The system runs on Intel Xeon E5-4600 v3 processors and scales from 4 sockets to 256 sockets as well as up to 64TB of cache-coherent shared memory, meaning that updates to data are propagated to all caches holding copies of that data and at high speed, which is an important aspect of its powerful scaling capacity.

SGI UV 300

The SGI UV 300 has been designed for data-intensive, in-memory, I/O demanding applications such as high-performance data analytics (HPDA), complex data visualization, and real-time streaming. The system runs on Intel Xeon E7-8800 v3 processors and can scale from 4 sockets to 64 sockets as well as up to 64TB of cache-coherent shared memory. The SGI UV 300 interconnect is the latest-generation NUMAlink 7 ASIC, which has a range of capabilities that business-critical applications increasingly require, most notably in-memory, real-time analytics; preventive fraud detection; and data streaming.

FIGURE 2

SGI UV 300 All-to-All Topology: Up to 32 Sockets



Source: SGI, 2016

Examples of SGI UV Deployments

Inova

Inova is a not-for-profit healthcare provider in Washington, D.C., with five hospitals, a Level 1 Trauma Center, and a Level 4 Neonatal Intensive Care unit. Inova is known for the Inova Heart and Vascular Institute (IHVI), the Inova Translational Medicine Institute (ITMI) on genomics, the Inova Neuroscience Institute, and the Inova Children's Hospital.

In 2015, ITMI chose an SGI UV 2000 system, the predecessor of the previously discussed SGI UV 3000, to enable its researchers to obtain greater insights from its genomic databases, among the largest in the world, to deliver more accurate diagnoses faster.

In 2015, ITMI chose an SGI UV 2000 system, the predecessor of the previously discussed SGI UV 3000, to enable its researchers to obtain greater insights from its genomic databases, among the largest in the world, to deliver more accurate diagnoses faster. The overall goal was to move from a reactive to a predictive system, which required analyzing enormous amounts of data in its genome database.

The SGI UV 2000 consists of 32 blades on 4 rack units built around Intel Xeon E5-4650L processors. The storage solution is SGI InfiniteStorage 5600, which is very I/O intensive at 50GBps, and provides 1PB of storage capacity in a single file system. The system's memory-to-processor ratio is very high at 32GB per real core, enabling extremely fast file reading. ITMI has started several studies on the UV 2000 to date — one to help predict and prevent pre-term births; one to find correlations between disease, genomics, patient demographics, and environmental factors; and one using whole genome sequencing to diagnose sick infants.

Sigma-Aldrich

Sigma-Aldrich is a life science and high-tech company that manufactures and distributes more than 250,000 chemicals, biochemicals, and other products to more than 1.4 million research, industrial, and commercial customers around the world. To perform the real-time analysis needed to manage its business processes and growing data demands, Sigma-Aldrich decided to switch its scale-out HANA appliance for a scale-up HANA solution with a road map beyond the current eight-socket limitation.

Sigma-Aldrich deployed 24 appliances of SGI UV for SAP HANA that are built around Intel Xeon E7-8890 v2 15-core processors, providing 50TB of in-memory compute power to run all its SAP applications as well as SAP Business Warehouse. The system will allow Sigma-Aldrich to merge its entire SAP environment into a single Business Suite on HANA instance on a scale-up, single-node system.

Sigma-Aldrich has stated to SGI that "SGI has offered us a unique scale-up technology for SAP HANA beyond the normal limits. This allows us to realize our vision of transforming our entire enterprise to a 'real-time enterprise,' without worrying about our ability to support the exponential growth of big data."

FedCentric

A team of researchers from Georgetown University, FedCentric Technologies LLC, and Frederick National Laboratory for Cancer Research designed an experiment to demonstrate that a graph database, running on an in-memory system, can be a "very powerful and useful tool" for cancer research. More specifically, they aimed to determine whether graph database structures are applicable for mining variants from individuals and populations in a scalable manner and understanding their impact by integrating with known annotations.

To execute the experiment, which required a powerful, scale-up, in-memory system, they chose an SGI UV 300 provided by FedCentric Labs, which scales up to 1,152 cores and 64TB of memory. After the experiment, the team concluded that an in-memory graph database allows researchers to run known queries and enables them to develop algorithms to explore complex correlations.

They also found that "the performance using the graph database for finding details on specific nodes or a list of nodes is better [than] or equal to [the performance of] a well-architected relational database." In other words, the SGI UV 300 was able to handle the complexity and data intensity of a massive graph database without a degradation of performance.

Challenges and Opportunities

For Organizations

As workloads become increasingly compute and data intensive, many organizations are finding that their infrastructure is unable to process the huge data sets that are available to them or quickly execute the computations that would yield strategically important knowledge.

As workloads become increasingly compute and data intensive, many organizations are finding that their infrastructure is unable to process the huge data sets that are available to them or quickly execute the computations that would yield strategically important knowledge. Costly legacy scale-up platforms often do not have the memory capabilities, the I/O speeds, and the processing power for these modern workloads. Low-capex distributed environments or HPC clusters are equally impeded but have additional networking latencies to overcome as well as surging opex implications. These organizations have an opportunity to dramatically improve their ability to run compute- and data-intensive applications, at lower cost, on modern scale-up systems.

For SGI

SGI is a smaller vendor in the server market, traditionally focused on very powerful, high-RAS systems for a customer base that requires them. This market is expanding beyond the scientific and research community and entering the datacenter as enterprises want more and more capabilities to better analyze the huge volumes of data they process every day. For SGI, the opportunity is the challenge, as the market is essentially coming to it. The firm's sales capabilities and brand-name recognition might not be those of an HP, an IBM, or an Oracle, but we believe that businesses would do well to investigate SGI's scale-up offerings before making a decision.

Summary and Conclusion

As organizations, small and large, are increasingly relying on compute- and data-intensive applications to analyze massive amounts of data from which they hope to obtain critical new insights to provide them with a competitive edge, scale-up server solutions have been found to have the capabilities organizations need with an advantageous cost-benefit ratio.

As organizations, small and large, are increasingly relying on compute- and data-intensive applications to analyze massive amounts of data from which they hope to obtain critical new insights to provide them with a competitive edge, scale-up server solutions have been found to have the capabilities organizations need with an advantageous cost-benefit ratio. This is especially the case in situations where the data sets are becoming excessively large, as massive data sets compute much faster and more efficiently on an in-memory scale-up system with a very large memory footprint than they would in clustered solutions, which are hampered by network latencies.

IDC has found (using the methodology described in the Appendix) that scale-up systems yield a number of benefits when used for the right workloads: improved performance, better resource utilization, improved application performance, less unplanned downtime, and extended datacenter life. Our research also determined that scale-up consolidation can yield savings of 33% due to lower opex. Contributing factors are reduced server management staffing requirements as well as lower costs associated with power and cooling, software licensing, and IT infrastructure.

This study also looks at the SGI UV platform and the SGI UV 3000 and SGI UV 300 — two scale-up systems that are typically used for very demanding compute- and data-intensive tasks in fields such as computational biology, genome reconstruction, business analytics, and systems biology. Given the increasing data volumes that enterprises must manage for insights, the near-real-time speed with which enterprises wish to obtain such insights, and the increasing complexity levels with which their data scientists are operating, The SGI UV platform can serve as a powerful platform for enterprises and life science centers alike.

Appendix

IDC leveraged data and information from the IDC Business Value database to inform the results presented in this study. IDC conducts hundreds of interviews per year with end-user organizations with regard to their IT environments and their use of IT solutions. For this project, IDC considered the experiences of 21 organizations interviewed from 2012 to 2015 with regard to their decision to run and support certain workloads on fewer servers rather than on more distributed server environments. IDC asked these organizations about costs associated with their use of these larger servers compared with smaller, more distributed servers based on their analyses and experiences when migrating from more distributed environments to larger servers.

IDC Analyze the Future

In calculating the values described in this study, IDC used a number of assumptions, which are summarized as follows:

- Time values are multiplied by burdened salary (salary + 28% for benefits and overhead) to quantify efficiency and manager productivity savings. For this study, IDC used an average fully burdened salary of $100,000 per year for IT staff members and $70,000 per year for other users of IT services.
- Downtime values are a product of the number of hours of downtime multiplied by the number of users affected.
- The impact of unplanned downtime is quantified in terms of impaired end-user productivity.
- Lost productivity is a product of downtime multiplied by burdened salary.

Because every hour of downtime does not equate to a lost hour of productivity, IDC attributed only a fraction of the result to savings. As part of our assessment, we asked each company what fraction of downtime hours to use in calculating productivity savings and the reduction in lost revenue. IDC then taxed the revenue at that rate.

Further, because IT solutions require a deployment period, the full benefits of the solution are not available during deployment. To capture this reality, IDC prorated the benefits on a monthly basis and then subtracted the deployment time from the first-year savings.

Note: All numbers in this document may not be exact due to rounding.

About IDC

International Data Corporation (IDC) is the premier global provider of market intelligence, advisory services, and events for the information technology, telecommunications and consumer technology markets. IDC helps IT professionals, business executives, and the investment community make fact-based decisions on technology purchases and business strategy. More than 1,100 IDC analysts provide global, regional, and local expertise on technology and industry opportunities and trends in over 110 countries worldwide. For 50 years, IDC has provided strategic insights to help our clients achieve their key business objectives. IDC is a subsidiary of IDG, the world's leading technology media, research, and events company.

IDC Global Headquarters

5 Speen Street
Framingham, MA 01701
USA
508.872.8200
Twitter: @IDC
idc-insights-community.com
www.idc.com

IDC Analyze the Future