BUYER CASE STUDY

# Los Alamos National Laboratory: Seven Decades of Computing Leadership

Steve Conway
Robert Sorensen

Earl C. Joseph, Ph.D.

## IDC OPINION

Los Alamos National Laboratory (LANL) is one of three U.S. Department of Energy/National Nuclear Security Administration (DOE/NNSA) labs whose core mission is to ensure confidence in the safety, security, performance, and reliability of America's nuclear stockpile. In pursuit of this important mission over seven decades, LANL has developed unrivaled experience in evaluating new, large-scale supercomputer systems and providing important feedback about the systems and their constituent technologies to their private sector OEM developers. IDC believes that this experience positions LANL well to assume an important leadership role in assessing today's vibrant global high-performance computing (HPC) technology development scene, including non-U.S. initiatives and questions about the future of technologies for processors, accelerators, storage and networking, power and cooling, and more. Further:

- LANL has been leveraging its expertise and resources to support a broad spectrum of scientific disciplines and a growing number of important collaborations with U.S. industry. LANL paved the way for the supercomputer era and has remained at the forefront of computing leadership.
- IDC expects LANL's experience with the coming "Trinity" supercomputer to make important contributions to the body of knowledge needed for success in the exascale supercomputing era.

## IN THIS BUYER CASE STUDY

This IDC Buyer Case Study benefits from recent IDC discussions with officials of LANL. LANL hosts a major U.S. government-funded supercomputing center that has twice housed the world's most powerful supercomputer and that will soon operate Trinity, a system that represents an important milestone en route to exascale computing capability. Since its founding in 1943, LANL has operated 100 leading-edge scientific computers, including early iconic systems such as the LANL-developed Mathematical Analyzer, Numerical Integrator, and Computer (MANIAC). LANL is widely recognized as a leader in exploiting high-performance computing in the interests of national security and advanced scientific and engineering research. The lab has been expanding its support for industrial collaborations and novel research on technologies for large-scale supercomputer systems.

January 2015, IDC #253567

## SITUATION OVERVIEW

# Organization Overview

## *LANL: A History of Computing Firsts*

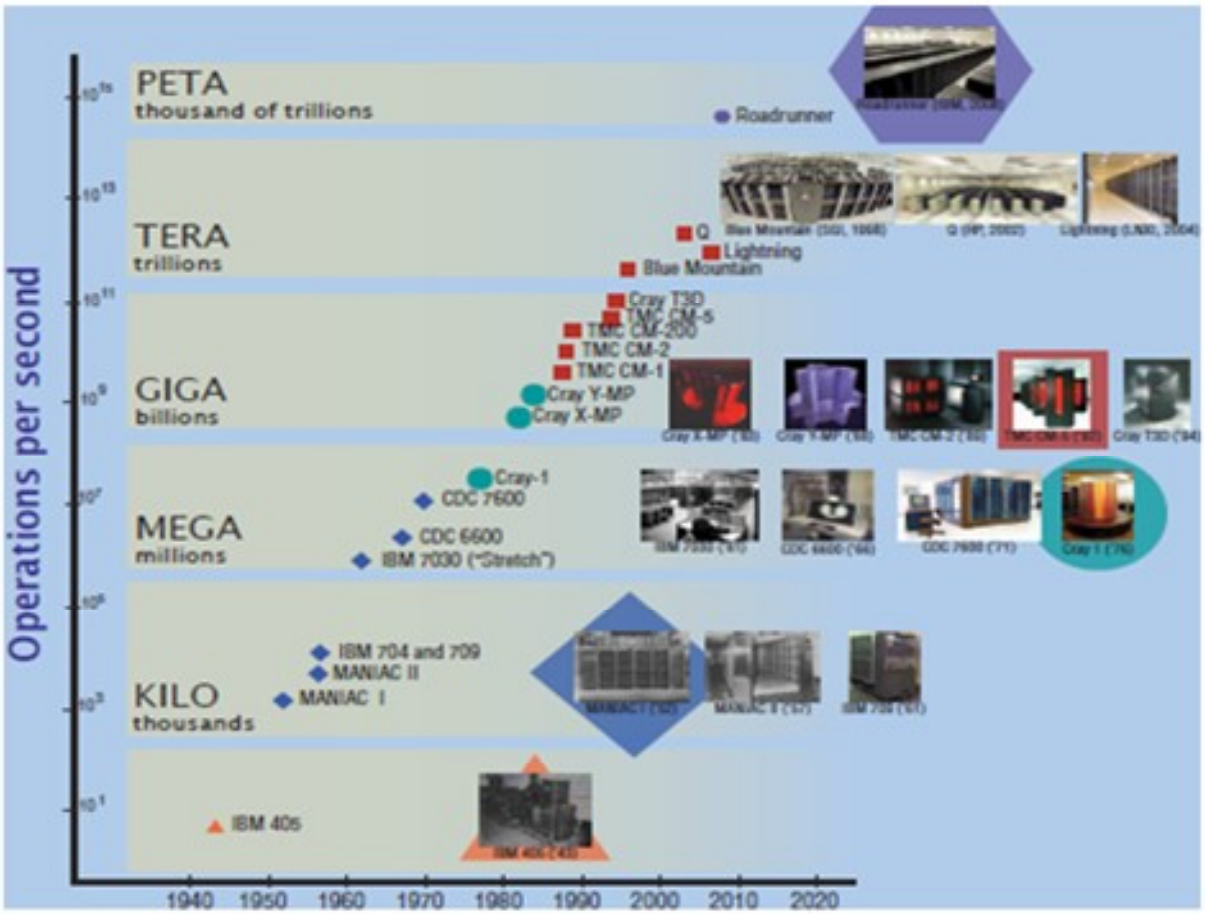
As an introduction to the history of supercomputing at LANL, consider this impressive short list of first-ever (and very early) achievements:

- LANL played a major role in developing the Monte Carlo simulation method.
- LANL hosted serial number 1 of what were arguably the world's first supercomputers, the IBM 7030 "Stretch" system and the Cray-1 supercomputer. (This means LANL also operated the first serial supercomputer and the first vector supercomputer.)
- LANL's work morphed into the Cray-1 operating system and the earliest predecessor to the Berkeley Fast File System that sired all modern Unix/Linux file systems.
- LANL used its 1,024-core Thinking Machines CM-5 massively parallel processing (MPP) supercomputer to advance the state of the art in exploiting SIMD data parallelism. It can be argued that today's GPGPU is analogous to a CM-5 system on a card.
- The IBM "Roadrunner" supercomputer at LANL was the first to break the Linpack petaflop barrier; Roadrunner, with its mix of AMD Opteron and IBM Cell processors, was also the first hybrid supercomputer in which CPUs connected via a PCI bus to accelerators sped up data parallel work. In that sense, Roadrunner arguably was the earliest ancestor of today's and tomorrow's hybrid systems, including the DOE "Summit" and "Sierra" supercomputers due out in 2018.
- LANL has been a pioneer in reusing evaporated water to cool supercomputers, saving 47 million gallons of water in 2013 alone.

Figure 1 shows the major supercomputers LANL has operated, with the rise in peak performance over the years.

## FIGURE 1

**Major Supercomputers Operated by LANL, 1940–2020**




Source: LANL, 2014

## *The Pioneering Years (1943–1960): LANL Paves the Way for Supercomputing*

Los Alamos Scientific Laboratory (LASL) (as LANL was previously called) was founded in 1943 as a primary R&D site within the Manhattan Project whose aim was to create the world's first nuclear weapons. To speed that urgent work, in the same year LASL scientists turned to early computing technology such as the IBM 601 multiplier and the IBM 405, an electric punched card accounting machine that could print alphanumerical summary reports.

With these devices, "a single problem took about three months to complete" (see Roger Lazarus, *et al., Computing at LASL in the 1940s and 1950s,* DOE, 1978). To cut solution times, the scientists quickly graduated to early electronic computers as soon as those became available. Notable examples included:

- **ENIAC.** In an arrangement brokered by John von Neumann, LASL scientists used this iconic electronic computer housed at the University of Pennsylvania in the late 1940s to carry out the first important thermonuclear calculation.
- **UNIVAC I and II.** LASL scientists used time on these systems located in New York and Philadelphia.
- **MANIAC.** This electronic computer developed at LASL was used from 1952 to 1958 to simulate the thermonuclear process. MANIAC ran a broad range of scientific applications, along with chess. It also played a major role in the development of the Monte Carlo method.
- **IBM 701.** The IBM 701, delivered to LASL in 1953, was the first truly commercial electronic computer used at the laboratory. Other early IBM computers at LASL included the follow-on IBM 704 and 709 series.

### The Lab's Crucial Checkout-and-Feedback Role

Today it is difficult to fully appreciate the experimental nature of these early electronic computers when they first arrived at LANL (we'll use the lab's current name from now on). Beyond the Turing-von Neumann baseline schema for the architecture, very little about these machines was standard or proven effective in real-world use on arrival at the lab. The indispensable job of testing these novel computing machines in practice and providing rapid, continual feedback fell heavily to LANL users. This established the pattern for the symbiotic relationships LANL still has with makers of leading-edge supercomputers today.

To a very important extent, the ability of LANL users to give early electronic computers rigorous workouts, to provide feedback along the way to the computers' developers and, in many cases, to create needed technologies that were missing (e.g., software) paved the way for the supercomputer era that followed. LANL played a major role in hardening early supercomputers for use in production computing facilities. Through trial-and-error experimentation, the lab also helped define the hardware, software, and networking technologies that would become standard supercomputing components.

## *The Early Supercomputer Era (1960–1980): LANL Usage Helps Define the Future*

The dividing line between supercomputers and their precursors is arbitrary, but historians typically base their choice of the first "supercomputer" on one of the first scientific computers that incorporated transistors instead of vacuum tubes. LANL was an early user of several of the first transistorized supercomputers. In detail:

- The first specimen of the IBM 7030 supercomputer, nicknamed Stretch, was delivered to LANL in 1961 and used until 1971. The transistorized Stretch design contributed substantially to the notably successful IBM System/360 (S/360) mainframe computer introduced in 1964. The S/360 was a hit with business and scientific organizations alike. IBM and LANL codeveloped the Stretch assembly program and operating system, so it is fair to say that LANL has made important contributions to business computing as well as scientific supercomputing.

- When it reached LANL in 1966, the CDC 6600 – the first computer designed by Seymour Cray – was 10 times faster than any other computer, including IBM Stretch, the record holder till then. The considerable speed advantage of CDC 6000 was the immediate impetus for coining the term *supercomputer* to differentiate the Control Data machine.
- In 1976, LANL received serial number 1 of the Cray-1 supercomputer, Seymour Cray's first creation at Cray Research. The new company's fate hung on the machine's performance during a six-month trial period. Thankfully, the fledgling supercomputer met LANL's stringent requirements, and the try and buy turned into a revenue-producing sale for start-up Cray Research. This episode once again illustrates LANL's important role in testing and validating new supercomputers. (The Cray-1 installation also began LANL's long experience operating supercomputers at high altitudes. LANL is 1.4mi above sea level, high enough for concerns about cosmic events.)

## *The Age of Heterogeneous Parallelism (1980–Today): LANL Passes the Giga, Tera, and Peta Milestones*

The history of parallel computing dates back at least to the 1960s, and arguably before then, depending on how you define the term. Vector supercomputers began going modestly parallel in the early 1980s, at a time when they dominated the HPC industry. The Cray X-MP series (1982) featured two vector processors, and the successor Cray Y-MP product line (1988) could hold up to eight processors. LANL acquired and put both of these commercial supercomputers through their paces early in their lifetimes.

But the rise of parallelism didn't begin in earnest until the advent of microprocessor-based, massively parallel processing supercomputers, especially in the second half of the 1980s and the 1990s. Emblematic of the MPP movement was the Thinking Machines CM-5 launched in 1991. The CM-5 system at LANL, outfitted with the maximum configuration of 1,024 processor cores, qualified as the world's fastest supercomputer on the June 1993 inaugural TOP500 supercomputers list. The system had peak performance of 131 gigaflops and sustained performance of 59.7 gigaflops on the Linpack test.

LANL broke the teraflops barrier in 2002 with its homegrown, 288-core "Space Simulator" machine that had peak performance of more than 1.4 teraflops.

Fast-forward to 2008, and a LANL supercomputer once again heads the TOP500 list. This time it's Roadrunner, an IBM system that achieved multiple firsts. It was the first supercomputer to pass the petaflop milestone (1.026 petaflops) on the Linpack test used to rank the TOP500 systems. As noted previously, IBM Roadrunner was also the first leading supercomputer to exploit a hybrid mix of base processors (AMD Opteron) and accelerators (IBM PowerXCell 8i) connected via a PCI bus.

## *Heading Toward Exascale: The Next-Generation Trinity Supercomputer*

The next major performance milestone being targeted by the worldwide HPC community is an exaflop, a quintillion ($10^{18}$) floating-point operations per second. IDC expects the United States, China, Japan, and one or more European nations to reach this milestone during 2020–2024. This achievement will require overcoming daunting challenges in scalability, density, reliability/resiliency, power efficiency, and other areas.

An important advance on the road toward exascale computing is the Trinity supercomputer that is scheduled for a multiphase deployment at LANL in late 2015 and 2016. Trinity is funded by the DOE's National Nuclear Security Administration and will serve users at all three NNSA national laboratories: LANL, Lawrence Livermore National Laboratory, and Sandia National Laboratories. IDC expects the peak performance for this Cray XC heterogeneous supercomputer to be more than 30 petaflops, with contributions to that total coming from Intel's Xeon (formerly Haswell) CPUs tightly integrated with Intel's Xeon Phi (Knights Landing) accelerators. The storage system will have a capacity of 82PB.

## *Major Collaborations with Industry*

Los Alamos has an active industry-partnering program and has worked over the past five years with more than 250 large and small companies to address national technology challenges. A recent addition to LANL's industry-partnering capabilities is the Richard P. Feynman Center for Innovation (FCI). The center acts as an innovation catalyst to help enable and accelerate the path from scientific creativity to the deployment of solutions.

### Chevron-LANL Collaboration

Chevron Corp. and LANL are collaborating in a joint research project to improve the recovery of hydrocarbons trapped in oil shales and slow-flowing oil formations. The goal of the Chevron-Los Alamos collaboration is to develop an environmentally responsible and commercially viable process to recover crude oil and natural gas from western U.S. oil shales. The joint research and development focuses on oil shale formations in the Piceance Basin in Colorado. The work will include reservoir simulation and modeling, as well as experimental validation of new recovery techniques, including a form of *in-situ* (inground) processing that has the potential to mitigate greenhouse gas emissions. Chevron has applied to participate in the Bureau of Land Management's research, development, and demonstration leasing program in the Piceance Basin. Chevron plans to use the 160-acre lease to evaluate the technologies developed through its alliance with Los Alamos, subject to both approval from the bureau and the success of the research program. The U.S. Geological Survey estimates the United States holds 2 trillion barrels of oil shale resources, with about 1.5 trillion barrels of those resources located in the Western United States, primarily in Wyoming, Colorado, and Utah.

### GE Oil & Gas

Chevron reported that the alliance with Los Alamos has already led to several breakthroughs in oil and gas technology, including the reduction of ultrahigh casing pressures in deepwater wells and improved well performance. The Chevron-LANL alliance also incubated the innovative swept-frequency acoustic interferometry metering technology, which GE Oil & Gas plans to incorporate into GE flow meter products.

### EMC-LANL CRADA Collaboration

In 2011, LANL signed an Umbrella Cooperative Research and Development Agreement (CRADA) with EMC Corp. Together, LANL and EMC agreed to enhance, design, build, test, and deploy new cutting-edge technologies in an effort to meet some of the nation's most difficult information technology challenges. The CRADA involves six general categories of technology development in which LANL and EMC plan to collaborate over a five-year period, including high-performance computing, data storage, cybersecurity, data sharing and mobility, cloud computing, large-scale analytics, and materials science.

This first project task statement (PTS) under the Umbrella CRADA is focused on support for the U.S. Department of Energy's exascale initiative and other data-intensive programs. The LANL–EMC collaboration for the exascale initiative is aimed at boosting HPC levels to the exaflops – 1,000 times faster than current petascale capabilities. The project involves design and development of an open source, extremely scalable data management middleware library called the parallel log-structured file system (PLFS), which will be used on a range of computing platforms from small clusters to the largest supercomputers in the world. (A side note is that burst buffers using solid state, nonvolatile storage were first prototyped in this project.)

## PRObE Center Aims to Advance Research About Large-Scale Supercomputers

In 2012, the National Science Foundation (NFS), with support from LANL, the New Mexico Consortium, and Carnegie Mellon University, opened the Parallel Reconfigurable Observational Environment (PRObE) Center, a one-of-a-kind computer systems research center located in Los Alamos. The center hosts the world's first supercomputing system dedicated to research on and about large-scale systems – an important topic for the march toward exascale capability.

Using funding from NSF and employing more than 2,000 recently retired computers from LANL, the PRObE Center is the world's first facility where computer systems researchers from across the nation have onsite and remote access to a dedicated large-scale supercomputer for the testing of supercomputing and big data systems software. Researchers can load their software and "own" their assigned portion of the supercomputer during their development projects.

The New Mexico Consortium is using PRObE to engage the community's young people. PRObE provides unique opportunities to dozens of local undergraduates and high-school students who come to Los Alamos to learn about supercomputing.

## Challenges and Solution

### *Challenges*

IDC sees LANL's main challenges as follows:

- LANL faces mission-related challenges associated with science-based nuclear stockpile stewardship, especially accurately predicting the effects of aging on the safety, performance, and reliability of the stockpile.
- LANL and other leading HPC sites in the United States and around the world face a common challenge in securing the funding needed to advance to exascale computing capability and beyond.
- Assuming adequate funding is available, LANL and other HPC leaders confront an array of daunting (and exciting) technical challenges that must be mastered en route to exascale performance.

### *Solution*

IDC is confident that LANL's supercomputing history positions the lab to master the challenges it faces as well as any HPC site in the world.

### *Opportunities*

IDC sees LANL's main opportunities as follows:

- LANL has exceptional opportunities to continue leveraging DOE investments in the lab for the benefit of advanced open scientific research of national importance as well as to help boost industrial innovation and competitiveness in the United States.
- Especially (but not exclusively) through PRObE Center activities, LANL has varied opportunities to extend its impressive achievements in advancing the design and use of large-scale supercomputer systems. (LANL researchers compete with university and other researchers for time on the PRObE Center supercomputer.)
- IDC believes that LANL is poised to help meet the national security, scientific, and industrial challenges of the 21st century.

## Results

LANL has amassed enormous experience related to the design, functioning, and usage of large-scale supercomputers systems in pursuit of the lab's primary mission to ensure confidence in the safety, security, performance, and reliability of America's nuclear stockpile. Today, LANL has a well-deserved reputation as one of the world's leading high-performance computing sites. The lab has leveraged its experience to play a major role in advancing the state of the art in supercomputing. LANL has also leveraged government/taxpayer investments by making large-scale supercomputer systems available to industry for advanced, precompetitive R&D initiatives.

## ESSENTIAL GUIDANCE

- **Stay on top of technology developments in the increasingly varied, competitive global HPC industry.** Gone is the era when HPC technology leadership was a two-nation race between the United States and Japan. As the HPC community approaches the exascale era, the competitive field has become more crowded, with China, leading European nations, and potentially India as new entrants. In each of these geographies, increased recognition of HPC's strategic value for scientific and economic competitiveness has produced a push for developing indigenous HPC technologies. In China alone, this has already led to half a dozen processor development initiatives. Along with this global competition, HPC buyers/users face a growing array of questions about the future roles of technologies for processors, accelerators, storage and networking, power and cooling, data-intensive and cloud computing, and more. LANL has an admirable history as an early, effective evaluator of emerging technologies and is well positioned to assume a leadership role in assessing today's vibrant, global technology development scene.
- **Document and promote ROI successes.** IDC believes that government officials will increasingly demand ROI evidence to help justify substantial new investments in leadership-

class HPC systems and related resources, especially in today's still-challenging economy. To the extent that this is feasible, LANL should quantify, document, and promote ROI achieved by the lab's scientists, in the form of innovations, and by LANL's industrial partners in collaborations with LANL. IDC is collecting examples of HPC-related ROI for a three-year DOE study grant and would welcome as many examples as possible from LANL and its partners.

## LEARN MORE

## Related Research

Additional research from IDC in the Technical Computing hardware program includes the following documents:

- High-Performance Computing in Practice: HPC User Forum, September 15–17, 2014, Seattle, Washington (IDC #251975, October 2014)
- New Cray Urika-XA System Targets HPC-Big Data Convergence (IDC #lcUS25203814, October 2014)
- Emergent Use Cases in High-Performance Data Analysis: HPC User Forum, September 15-17, 2014, Seattle, Washington (IDC #251976, October 2014)
- Experiences with Accelerators and Coprocessors in High-Performance Computing: HPC User Forum, September 15–17, 2014, Seattle, Washington (IDC #251973, October 2014)
- Major Global High-Performance Computing Initiatives: HPC User Forum, September 15–17, 2014, Seattle, Washington (IDC #251971, October 2014)
- Lenovo Completes Acquisition of IBM's x86 Server Business (IDC #lcUS25176214, September 2014)
- USPS Touts Benefits of HPC for Big Data (IDC #lcUS25123014, September 2014)
- Worldwide Broader HPC 2014–2018 Forecast: Servers, Storage, Software, Middleware, and Services (IDC #248835, June 2014)
- When Massive Data Never Becomes Big Data (IDC #lcUS24922014, June 2014)
- Worldwide Technical Computing Server 2014–2018 Forecast (IDC #248779, May 2014)
- Perspectives on High-Performance Data Analysis: The Life Sciences (IDC #248348, May 2014)
- Global HPC Market Dynamics in 2013 (IDC #248137, April 2014)
- Industrial Partnership Programs and High-Performance Computing: HPC User Forum, April 7-9, 2014, Santa Fe, New Mexico (IDC #248113, April 2014)
- International Perspectives on Industrial High-Performance Computing Partnerships: HPC User Forum, April 7–9, 2014, Santa Fe, New Mexico (IDC #248122, April 2014)
- Worldwide HPC Public Cloud Computing 2014–2017 Forecast (IDC #247846, April 2014)
- IDC's Worldwide High-Performance Computing Predictions 2014 (IDC #WC20140211, February 2014)

- Micron Demonstrates Technologies to Address Emerging Challenges in Big Data Applications (IDC #244843, December 2013)
- Market Analysis Perspective: Worldwide HPC, 2013 — Directions, Trends, and Customer Requirements (IDC #244742, December 2013)
- HPDA Pulse Results: 2013 Hardware and Storage Market Analysis (IDC #244493, November 2013)
- Catalyst Supercomputer Heralds Shift to More Balanced Architectures (IDC #IcUS24437513, November 2013)
- China Eyes 10,000-Fold Data Reduction for Internet of Things (IDC #IcUS24392513, October 2013)
- HPC User Forum, October 2013, Seoul, Korea (IDC #243786, October 2013)
- Tools and Techniques for Technical Computing in Life Sciences: HPC User Forum, September 2013, Boston, Massachusetts (IDC #243778, October 2013)
- Perspectives on Quantum Computing: HPC User Forum, September 2013, Boston, Massachusetts (IDC #243777, October 2013)
- High-Performance Data Analysis in the Life Sciences: HPC User Forum, September 2013, Boston, Massachusetts (IDC #243774, October 2013)
- Chinese Research in Processor Designs for High-Performance Computing and Other Uses (IDC #243502, October 2013)
- The Broader HPC Market 2012-2017 Forecast: Servers, Storage, Software, Middleware, and Services (IDC #242742, August 2013)
- IDC's Worldwide Technical Server Taxonomy, 2013 (IDC #242725, August 2013)
- 10 Things CIOs Should Know About High-Performance Computing (IDC #241565, June 2013)
- Worldwide High-Performance Data Analysis 2013–2017 Forecast (IDC #241315, June 2013)
- Top Issues for HPC Sites: HPC User Forum, April 29–May 1, 2013, Tucson, Arizona (IDC #241463, June 2013)
- Advanced Visualization: HPC User Forum, April 29–May 1, 2013, Tucson, Arizona (IDC #241446, June 2013)
- Livermore Lab Expands Industry Partnerships: Economic Security Is Vital for National Security (IDC #240232, March 2013)
- Advanced Research at the South Ural State University Supercomputing Center (IDC #238225, December 2012)
- China Confirms Plans for 100 PLFOPS Supercomputer by 2015 (IDC #IcUS23797112, November 2012)
- HPC End-User Site Update: RIKEN Advanced Institute for Computational Science (IDC #233690, March 2012)

## 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.

## Global Headquarters

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

---