The Computerworld Honors Program

Honoring those who use Information Technology to benefit society

Final Copy of Case Study

LOCATION:
Markham, Ontario, Canada

YEAR:
2011

STATUS:
Laureate

CATEGORY:
Collaboration

ORGANIZATION:
The European Organization for Nuclear Research (CERN)

ORGANIZATION URL:
http://www.cern.ch

PROJECT NAME:
Large Hadron Collider (LHC) Cluster Deployment with Platform Computing

PROJECT OVERVIEW
The European Organization for Nuclear Research (CERN) has a long history of driving both scientific and technical innovation since 1954, including Tim Berners-Lee's invention of HTTP and the World Wide Web to, more recently, the creation of the Large Hadron Collider (LHC), the largest and highest-energy particle accelerator in existence. Computing power is essential to ensuring that CERN's 12,000+ scientists and researchers across 68 countries can collaborate to help solve the mysteries of the universe. To deal with large datasets, massive throughput requirements, a globally dispersed network of scientists, finite budgets, and an emphasis on standards and openness, CERN needs a shared computing infrastructure that can support any collection of hardware, virtual machine hypervisors and operating systems in a cost-effective manner. When the organization started building a new cluster for the Large Hadron Collider in 2008 it sought an open, cost-effective platform for its expansive network of scientists that could manage both virtual and physical servers. The resulting technology deployment has created one the world's largest cloud computing environments for scientific collaboration. The LHC project presented CERN with some new challenges. Overall utilization of the new cluster was not optimal, and administering the heterogeneous cluster was labor-intensive. To capture, simulate, reconstruct and analyze the trillions of computing events generated by the LHC experiments, utilization of available machines needed to improve, and system administration needed to be simplified. The team at CERN responsible for delivering computing resources to scientists wanted not only to reduce the cost of managing heterogeneous virtual and physical machines, but also to delegate some administrative tasks to users by providing self-service options. The first step to improve utilization was to virtualize the machines. The LHC cluster uses many different hypervisors including HyperV, KVM, and Xen. This allows experiment teams greater flexibility to use the best platform for their applications. However, managing multiple virtualization technologies adds a level of complexity to the entire environment. A dynamic IT solution for enterprises to build and run private clouds was sought. CERN selected two products, Platform Computing's ISF and an open-source solution for an

in-depth evaluation as candidate tools to manage both the virtual and physical environment as a whole. Platform ISF gives CERN more control and provides administrative efficiency by automating myriad tasks that cannot be done by hypervisors alone. Forming the core of one of CERN's private clouds, it manages the resources and application environment offerings and provides a contract interface to enable scientists to reserve and use resources on demand. Most importantly, it unites the diverse computing platforms into a single dynamically shared infrastructure, dramatically improving CPU utilization with fewer resources.

SOCIETAL BENEFITS

In deploying one the world's largest cloud computing environments for scientific collaboration, CERN is enabling scientists around the globe to achieve maximum efficiency in transforming massive amounts of raw research data into knowledge that advances our understanding of fundamental physics and the origins of the universe.

PREVIOUS PROJECT UPDATED/EXPANDED?

No. CERN is taking a phased approach by starting small, and by evaluating two candidate solutions. Based on the experience they are gathering and the way the system reacts, they will add more and more machines to it. While the project is currently in production, it has not yet reached its full planned capacity.

PROJECT IMPLEMENTATION COMPLETE?

No

PROJECT BENEFIT EXAMPLE

Improves CPU utilization and the delivery of computing resources to CERN scientists working around the world. — Allows scientists to choose their application environments and manage their processing workloads themselves, eliminating labor- and cost-intensive administration processes for IT department. — Allows for increased scalability of the resources provided to scientists within constant head counts. — "We are limited in the amount of power and cooling available," says Tony Cass, group leader of the former Fabric Infrastructure and Operations group at CERN. "We want to wring every last drop out of the resources we have to do physics. Even 10 percent more capacity means that much more toward improving the physics." — reported in Scientific American, February 8, 2010 ("CERN Gears Up Its Computers for More Atom Smashing," by Larry Greenemeier)

IS THIS PROJECT AN INNOVATION, BEST PRACTICE? Yes

ADDITIONAL PROJECT INFORMATION

"CERN Gears Up Its Computers for More Atom Smashing" Scientific American, February 8, 2010:
http://www.scientificamerican.com/article.cfm?id=cern-lhc-cloud-computing