



Platform LSF helps CERN reconstruct the BIG BANG!

Customer
CERN

Solution
Platform LSF

Industry
Government Research

Results

Challenges

- CERN needed to manage huge amounts of data collected from high-energy accelerators, and the application workload generated by 6,500 scientists worldwide
- CERN needed to harness optimize the performance of compute- and data-intensive applications, while reducing support requirements and creating a reliably scaling computing environment to meet drastic growth in requirements

- The computing cluster is achieving new levels of application performance, and results are being delivered to users when they need them
- Minimal support or manual intervention is required to keep workload management system running smoothly, thus reducing staff requirements
- Complex data management software workload is also being managed, enabling different users to access many data sets as needed
- Platform LSF workload management solution is successfully scaling up as computing infrastructure and workload grow

“We’ve had a very fruitful collaboration with Platform. We’ve both benefited from this in terms of improvements in areas such as queuing techniques and failover, and in the interesting new ways of using the Platform LSF product, such as for grid-level scheduling.”

Dr. Tony Cass
Leader of CERN’s Computing Fabric,
Infrastructure & Operations Group

Success of particle accelerator is in the grid

The European Organization for Nuclear Research (CERN) has just finished building the world’s largest particle accelerator, the Large Hadron Collider (LHC), where some 6,500 high-energy particle physicists from labs and universities around the world will conduct experiments to understand the nature of matter as it existed immediately following the Big Bang.

Smashing beams of particles into each other generates staggering amounts of data that is captured and stored in CERN’s custom-built data storage system. Using special software, physicists at CERN, or in locations around the world, analyze the data for information that will help them prove or disprove their theories on particles and their interactions. This analysis, and the many simulations that typically accompany it, requires that large numbers of compute- and data-intensive jobs be submitted to CERN’s computing centre.

Delivering timely results

With an average of 70,000 jobs scheduled each day, and daily peaks reaching 100,000 jobs, “Platform LSF ensures that our computing cluster delivers results to our physicist users when they want them,” says Dr. Tony Cass, Leader of the Computing Fabric, Infrastructure & Operations Group at CERN.

Platform’s High Performance Computing (HPC) software played a key role in making the project possible. The scheduling and workload prioritization capabilities of Platform LSF allow CERN to distribute application workload from to the 40,000+ CPUs in LHC’s computing grid efficiently enough to perform the massive amounts of calculations and processes required. Most of the post-experimental analysis uses custom software written at CERN but commercial applications in use include: Ansys, AutoCAD, Cadence, LabVIEW, Matlab, Mathematica, Opera, PVSS, CATIA, StarCD, and Saber—among many others.





Allocating the workload across the entire Linux cluster as required, Platform LSF matches the jobs being submitted to the computing resources available and keeps the resources running at close to maximum capacity. Platform LSF also balances job priorities and handles the different user requests—some want short jobs to run quickly, some are wanted over night, while simulations may need to run for weeks.

When CERN started building the new cluster for LHC, replacing Platform LSF was not even up for discussion. "I've been responsible for our use of Platform LSF since 1997 and always thought that it was the one part of our facility we would not need to change for LHC. We believe it can handle the challenge by simply scaling up to manage even more jobs, faster," says Cass.

Staggering amounts of data

With experiments getting more complex and the detection equipment getting more sophisticated all the time, however, CERN needed to harness the full potential of its computing resources in order to optimize the performance of the physicists' compute- and data-intensive applications. To handle the exponential growth in data, Platform LSF is used for the data management software workload within CERN's special storage system, giving physicists quick and easy access to the experimental data.

Platform LSF tackles unique storage management challenge

To handle its unique data storage requirements—capturing vast amounts of data very quickly from the particle detector equipment—CERN developed its own hierarchical storage management system called CASTOR. StorageTek tape robots provide the platform for the CASTOR system, which also includes a big spooling disk based on a dedicated array of clustered Linux file servers.

When a physicist submits a job requiring access to data, special storage management software running on each of the Linux file servers in CASTOR finds and retrieves the data, which could be distributed across thousands of storage units, spools it to the file server, and then points the user's application to the data. With only limited file server resources available, however, a storage management software scheduling requirement arises when many different users request access to many different data sets.

"Platform LSF helps us address our storage management challenge by scheduling user requests for data sets to the least-loaded disk server, the same way that it schedules job processing to the least-loaded application server," explains CERN's Cass. "It chooses the file servers based on a variety of resource utilization factors such as CPU usage, I/O load, available disk space, and so on."

The System Just Keeps On Ticking

When asked to elaborate on the role of Platform LSF at CERN overall, Dr. Cass explained that it wasn't a simple case of needing it to help CERN make better use of idle or underutilized resources, as it is for many other organizations. Dr. Cass says they've had no problems leaving their system running on its own for extended periods of time without needing any support, even when a key support position remained unfilled for six months. Platform LSF kept their system running smoothly with no breakdowns. "We find Platform LSF to be reliable, bullet-proof software that just keeps running and handling the workload," enthuses Cass. "We've been able to reduce the amount of staff we have to put into managing this area," he adds.

Platform™

Platform Computing is the leader in grid and cloud computing software that dynamically connects IT resources to workload demand according to business policies. Over 2,000 of the world's largest organizations rely on our solutions to improve IT productivity and reduce data center costs. Platform has strategic relationships with Cray, Dell, HP, IBM, Intel, Microsoft, Red Hat, and SAS. Building on 16 years of market leadership, Platform continues to help data centers be more efficient, responsive and dynamic. Visit www.platform.com.

World Headquarters
Platform Computing Inc.
3760 14th Avenue
Markham, Ontario
Canada L3R 3T7
Tel: +1 905 948 8448
Fax: +1 905 948 9975
Toll-free tel: 1 877 528 3676
info@platform.com

Sales - Headquarters
Toll-free tel: 1 877 710 4477
Tel: +1 905 948 8448

North America
New York: +1 646 290 5070
San Jose: +1 408 392 4900
Detroit: +1 248 359 7820

Europe
Basingstoke: +44 (0) 1256 883756
London: +44 (0) 20 7977 1480
Paris: +33 (0) 1 41 10 09 20
Düsseldorf: +49 2102 61039 0
Munich: +49 89 517397 52
Oslo: +44 1256 883756
info-europe@platform.com

Asia-Pacific
Beijing: +86 10 82276000
Xi'an: +86 029 87607400
asia@platform.com
Tokyo: +81(0)3-6302-2901
info-japan@platform.com
Singapore: +65 6307 6590
lliew@platform.com