

PBS Professional at the Translational Genomics Research Institute (TGen): Cutting Time to Discovery



Key Highlights

Industry

Life Sciences

Challenge

A vendor-supported solution that would meet the need to provide flexible job scheduling on Saguaro, TGen's 512-node production cluster.

Altair Solution

PBS Professional has the flexibility to fit TGen's needs, and was easy to install, running with little to no support needed.

Benefits

- Provides flexibility to run large jobs multiple nodes while running thousands of small serial jobs on a single node

Customer Profile

In recent years, corporations and research institutions around the world have applied massive computational resources to defining the makeup of the human genome. One of the greatest challenges is to translate that knowledge into therapeutics and diagnostics — which is the mission of the Translational Genomics Research Institute (TGen), a remarkable non-profit organization founded by a joint effort between the State of Arizona, Arizona State Municipal Governments, Indian Tribal Community, educational institutions, private foundations and corporate entities. TGen's work is not only to make genetic discoveries, but also to translate discoveries into benefits for human health in the form of new diagnostic tests and therapies.

With support from Arizona State University (ASU), TGen established its High Performance Biocomputing Center (HPBC) on the ASU Tempe Campus to give its scientists the powerful computational resources they need to discover how genetic changes contribute to disease progression and resistance to therapy. In the words of the HPBC's mission statement, these resources “help empower researchers' ability to rapidly translate genomic discoveries into diagnosis and treatment.”

Setting Up TGen's Computational Heartbeat

TGen was established in Phoenix in 2002 with an initial staff of 23 scientists. In March 2003, ASU procured high-performance

TGen Success Story

TGen established its HPBC to give its scientists the powerful computational resources they need. *“One of the ringing endorsements I can give PBS Professional... It just works.”*

computing machines, including a 512-node IBM eServer Cluster 1350, to support TGen’s translational genomics research program.

In April 2003, IBM began installation of the eServer Cluster, running Red Hat Linux on 1024 Intel Xeon processors. By July, the HPBC was running production-type tests of gene sequencing and other processes on the cluster system, which is known as Saguaro. By late September it was in full production.

Altair’s PBS Professional® workload management software was an integral element of Saguaro from the start.

“We had installed OpenPBS on our 16-node development cluster, a scaled-down version of Saguaro that came on line quite a bit sooner than Saguaro,” says James Lowey, Manager of High Performance Computing Systems for TGen. “But it didn’t provide many of the things that PBS Professional does. One of the critical factors PBS Professional gives me is accounting — the ability to look at the

number of jobs that are run and the amount of time each job takes.

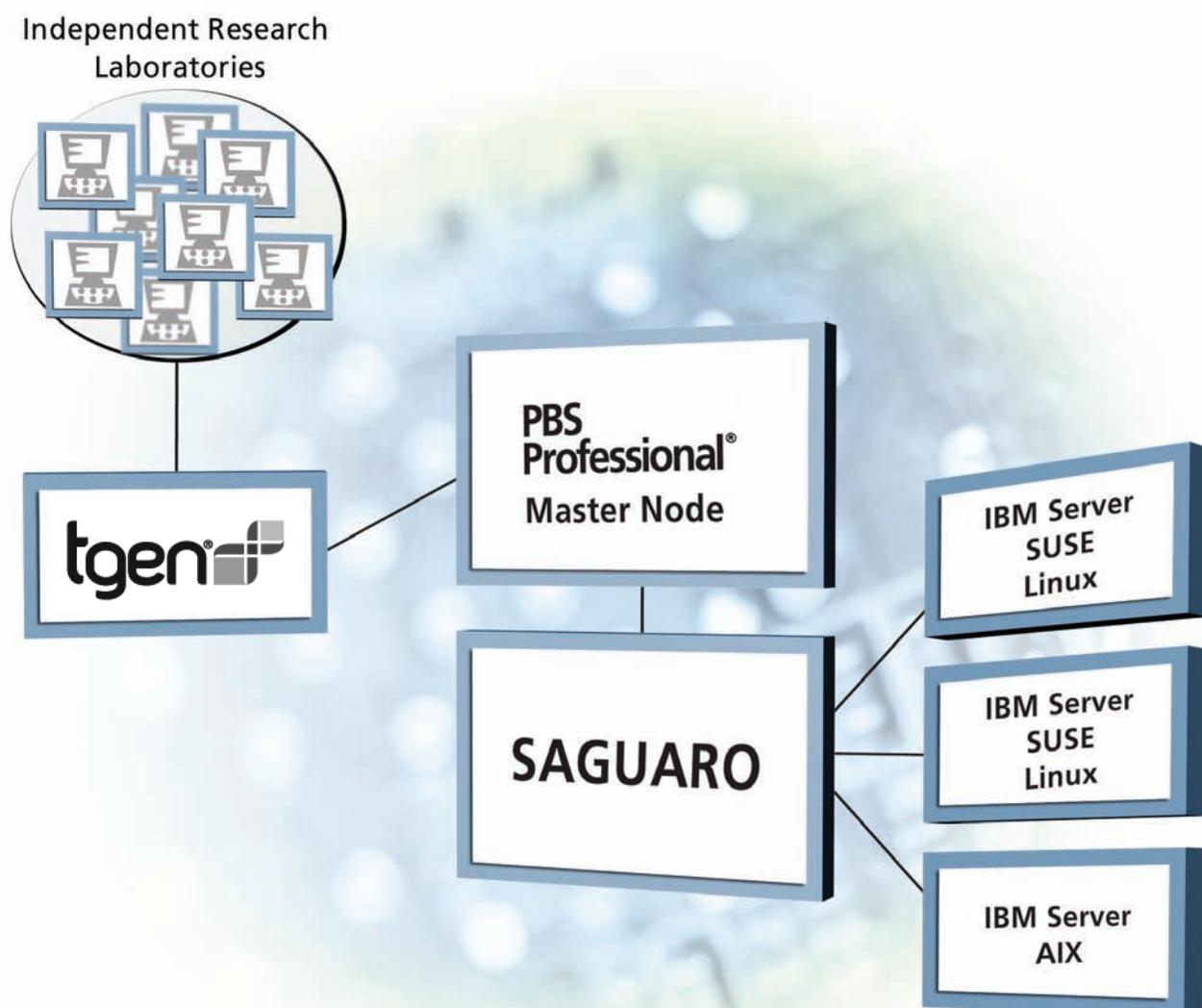
“Ultimately, TGen’s CIO, Dr. Edward Suh, chose PBS Professional because we needed a vendor-supported product that would meet our need to provide flexible job scheduling on Saguaro, the 512-node production cluster. We’ve been very pleased with the product’s capabilities.”

Putting Saguaro to Work

Today, HPBC typically serves about 65 accounts on Saguaro, most of them within TGen. Scientific collaborators at ASU and other research institutions are also active users of the resource. They use BLAST, AMBER, Gaussian, and other commercial and in-house-developed applications to run thousands of jobs on Saguaro. PBS Professional provides the flexibility to run large jobs across, say, 128 nodes, while running thousands of small serial jobs on a single node.

Saguaro, a 16-node development cluster, and three IBM SMP compute servers — two running SUSE Linux and one running AIX — all run on a high-performance SAN that connects to Saguaro over three Cisco 4006 switches. Users can watch their jobs interactively using a 1TB IBM GPFS Parallel File System that is accessible to every node on the cluster.

One characteristic of PBS Professional that has helped HPBC cope with the demand for Saguaro’s resources is hands-off dependability and simplicity of maintenance. “One of the ringing endorsements I can give PBS Professional is that once we got it set up and working, I have not had to do anything to it at all,” says Lowey. “I went through last year and upgraded my entire cluster to Red Hat EL3.0. Part of that process was reinstalling PBS Professional. I followed the instructions in the manual and it took about 20 minutes. It was quite simple.”



Looking Ahead: Upgrades and a Web-Based Interface

One of HPBC's goals is to move TGen to a web-based job submission model, and an internal web-based data analysis website is already operating. Another goal is flexible queues tied together with a switch architecture, which will enable HPBC to run, say, 32-node jobs on a single switch blade, or 128-node jobs on a single switch,

removing the latency of switch-to-switch communication. These and other advances will involve PBS Professional.

TGen's successful experience with PBS Professional will soon lead to an upgraded version, which HPBC is currently evaluating. Of particular interests are the job array, redundancy and failover features of the current release. Saguaro has received heavy

utilization since it came into full production in late 2003, and an increase in failures is inevitable. PBS Professional's Automatic Job Recovery will automatically redo any interrupted job upon detecting that the nodes have gone down.

"I'm excited that upgrading my production cluster and looking at other uses of PBS Professional in our HPC environment are among my goals this year," said Lowey.

Visit the PBS Works library of
Success Stories
at www.pbsworks.com

About Altair

Altair empowers client innovation and decision-making through technology that optimizes the analysis, management and visualization of business and engineering information. Privately held with more than 1,800 employees, Altair has offices throughout North America, South America, Europe and Asia/Pacific. With a 27-year-plus track record for high-end software and consulting services for engineering, computing and enterprise analytics, Altair consistently delivers a competitive advantage to customers in a broad range of industries. Altair has more than 3,000 corporate clients representing the automotive, aerospace, government and defense, and consumer products verticals. Altair also has a growing client presence in the electronics, architecture engineering and construction, and energy markets.

About PBS Works

PBS Works™, Altair's suite of on-demand cloud computing technologies, allows enterprises to maximize ROI on existing infrastructure assets. PBS Works is the most widely implemented software environment for managing grid, cloud, and cluster computing resources worldwide. The suite's flagship product, PBS Professional®, allows enterprises to easily share distributed computing resources across geographic boundaries. With additional tools for portal-based submission, analytics, and data management, the PBS Works suite is a comprehensive solution for optimizing HPC environments. Leveraging a revolutionary "pay-for-use" unit-based business model, PBS Works delivers increased value and flexibility over conventional software-licensing models.

www.pbsworks.com



Altair Engineering, Inc., World Headquarters: 1820 E. Big Beaver Rd., Troy, MI 48063-2031 USA
Phone: +1.248.614.2400 • Fax: +1.248.614.2411 • www.altair.com • info@altair.com