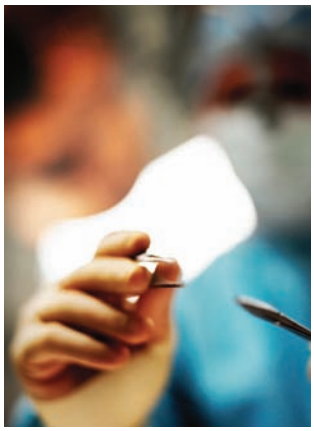
Altair | PBS Works™

PBS Professional at the National Institutes of Health (NIH): Scaling Up with PBS Professional to Fill Increasing Demands





Key Highlights

Industry
Life Sciences

Challenge
Open source solution could not provide adequate scalability, and required a system with open APIs to enable close integration with NIH workflows.

Altair Solution
When a node goes down, PBS Professional allows the job to be requeued without bringing the system back up.

Benefits
- Ability to schedule parallel jobs with awareness of the network topology within the clusters
- Automatically ensures each job runs completely on nodes connected to one switch or the other

Growing Scientific HPC Resources at NIH

Twenty-seven institutes and centers make up the National Institutes of Health. Many of their acronyms, listed on the NIH website (nih.gov), can be as mysterious as hieroglyphics: NHGRI, NIGMS, NIAMS. Others, such as NCI (National Cancer Institute) and NIMH (National Institute of Mental Health) are familiar to most of us. By any name or acronym, this cluster of organizations in Bethesda, Maryland is a major force in health sciences research.

NIH is widely known as a granting agency that funds research at scores of universities and institutions across the U.S. But research within its own institutes covers a mind-boggling array of disciplines and fields of study, much of it at the molecular level, and most of it extremely compute-intensive. NIH is well equipped with processing power to serve the institutes' scientists, and PBS Professional® is used for batch job management on its Linux cluster.

Growing an HPC Infrastructure for NIH Researchers

Helix Systems is the informal name for the HPC resources within the Division of Computer System Services, which is part of NIH's Center for Information Technology (CIT). Over twenty years ago, scientists were submitting jobs to a DEC System 10, the group's original large system. Since then, the group has used IBM, Convex, and SGI compute technology. It entered the cluster era in 1999 with an 80-CPU system known as Biowulf that has grown to more than 2,500 processors to become the primary resource at Helix Systems.

Keywords: Life Sciences, NIH, PBS Professional

NIH Success Story

NIH is widely known as a granting agency that funds research at scores of universities and institutions across the U.S. The legacy Origin systems, which run IRIX, operate as standalone units, and their jobs are submitted separately.

Helix Systems serves as a central computing facility for the intramural NIH program, and Biowulf is its primary computational resource. The cluster is used only by on-campus NIH scientists, and only for science. It is not homogeneous; it consists of more than 1,250 dual-processor systems, most of which have Opteron processors, although it includes several hundred Xeons from tearlier upgrades. In fact, the cluster has moved through generations of processors, starting with Pentiums, then Athlons, and on to Xeons and Opterons. This heterogeneity is due to the NIH budgeting system, which enables CIT to add several hundred nodes a year rather than make a major purchase every three or four years.

Workload Management

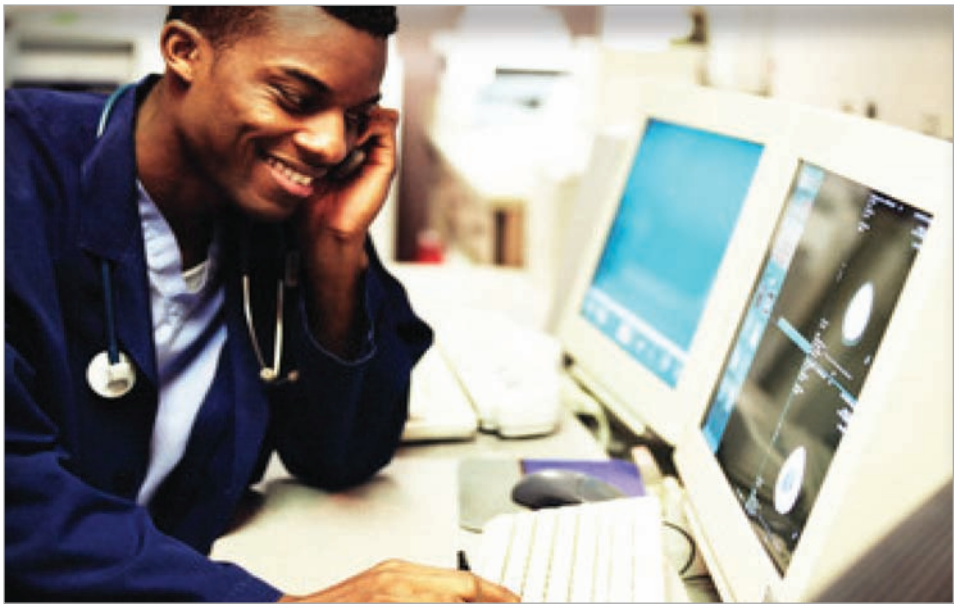
CIT used OpenPBS as its workload management system when it first brought its 80-processor cluster on line in 1999. It stayed with OpenPBS as the cluster grew. But when the cluster broke the 300-CPU level, OpenPBS could no longer provide adequate scalability, and CIT decided to move a more robust queuing system. In the past, when nodes went down with OpenPBS, CIT had to manually repair the system before requeuing jobs. When a node goes down now, PBS Professional® allows the job to be requeued without bringing the system back up.

CIT also needed a system with open APIs to enable close integration with NIH workflows. An example is Swarm, which CIT developers wrote to provide simultaneous submissions of large numbers of jobs with slightly different parameters. The next upgrade will need to provide technology that fulfills this need because individual researchers have submitted Swarms of 6,000 jobs each, which has placed a substantial load on the cluster. Helix Systems administrators need a more efficient way of handling such large submissions.

Another critical feature that CIT depends on, and will depend on even more in the future, is the ability to schedule parallel jobs with an awareness of the network topology within the clusters. In the past, users would have to manually ensure that their jobs were submitted to one of two Myrinet switches. PBS Professional automatically ensures that each job runs completely on nodes connected to one switch or the other.

Workload Management at NIH

In the past, when nodes went down with OpenPBS, NIH's Center for Information Technology had to manually repair the system before requeuing jobs. When a node goes down now, PBS Professional allows the job to be requeued without bringing the system back up.



Matching Applications to Platforms

In addition to Biowulf, Helix Systems hardware currently includes three SGI servers for computational work: an 8-CPU Origin 2400, a 32-CPU Origin 3400, and a 32-processor Altix. Four Network Appliance filers serve files from eight terabytes of high-performance online storage and 20 TB of nearline storage. The legacy Origin systems, which run IRIX, operate as standalone units, and their jobs are submitted separately. The Altix server, on the other hand, runs Linux, so CIT is using PBS Professional to integrate it into the cluster as a fat node. In total, CIT has deployed PBS Professional to manage the computational workload on 2,726 processors.

The idea is to match the application to the platform. Most applications that are run on Helix Systems compute resources, such as BLAST and Allegro, run well using the cluster model, either as parallel jobs or as what Helix Systems administrators call a Swarm of single-threaded jobs. Helix Systems maintains its SGI systems for jobs that don't fit those models. Parallel jobs in applications such as AMBER and Gaussian, which have such critical latency requirements that even high-performance networks like Myrinet or Infiniband aren't fast enough, are run on the Altix, and so are SMP jobs that need more than two processors or a great deal of memory.

Helix Systems has to maintain a general-purpose system because it runs all popular bio applications, some of which are parallel and some of which are single-threaded. NIH scientists run more than 40 applications on Helix Systems for Sequence Analysis, Phylogenetic/Linkage Analysis, Computational Chemistry/Molecular Modeling, Proteomics/Mass Spectrometry, Mathematical Analysis/Statistics, and Structural Biology (biowulf.nih.gov/apps/index.html). The greatest overall consumers of compute cycles are bioinformatics, statistical studies and molecular dynamics.

Hundreds of NIH scientists have accounts for the use of these resources, accessing the system through SSH. Active users, defined by CIT as accumulating at least 30 node-hours in a month, numbered 122 in a recent month. Some institutes have their own HPC resources and rarely call on CIT resources; others rely on Helix Systems entirely.

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

Altair Engineering, Inc., World Headquarters: 1820 E. Big Beaver Rd., Troy, MI 48083-2031 USA
Phone: +1.248.614.2400 • Fax: +1.248.614.2411 • www.altair.com • info@altair.com