

PNNL, SGI to team on storage technology development for data-intensive computing

June 24 2004

Project aims to accelerate scientific research by shifting computation of large data to storage devices

The Department of Energy's Pacific Northwest National Laboratory (PNNL) today announced a research alliance aimed at enabling a new generation of fast and efficient storage technology for data-intensive computing. Part of a long-term collaboration between PNNL and Silicon Graphics, the alliance includes options for more than 2.5 petabytes of storage over the next two years.

PNNL will conduct research into "active storage," a groundbreaking effort to shift computation and transformation of data from client computers to storage devices. According to Dr. Eng Lim Goh, SGI senior vice president and chief technology officer, the effort "holds the promise of dramatic productivity breakthroughs for a broad range of computing disciplines saddled by large data."

The effort combines the expertise of SGI and PNNL in advanced storage technologies with the laboratory's mission to address national priorities in the chemical, physical and biological sciences. As the first phase of the alliance, SGI Professional Services will deliver a single 380 terabyte file system this summer to the William R. Wiley Environmental Molecular Sciences Laboratory located at PNNL.

PNNL scientists will be able to take raw data sets stored on the file server and conduct computations to identify data signatures and patterns

before the data is transferred to client systems.

"By developing methods to perform computing inside the file system, we will be able to reduce the amount of redundant data transfers, which routinely undermines productivity and lengthens the time to solution," said Scott Studham, PNNL associate director for advanced computing. "This vastly more efficient approach to data-intensive storage promises to significantly speed scientific discoveries in life sciences, national security, and even film and video production."

The new file system is expected to sustain write rates in excess of 8GB/sec and demonstrate single client write rates of more than 600MB/sec. To achieve this performance, the new file system will leverage Lustre, an open source, object-oriented file system with development lead by Cluster File System Inc., with funding from DOE. Lustre currently is used on four of the top five supercomputers, including the PNNL cluster based on 1,900 Intel® Itanium® 2 processors.

"The research alliance taps SGI's expertise as a leading provider of storage solutions designed specifically for data-intensive environments, with robust and combinable solutions for intelligent consolidation, data lifecycle management and data protection," Goh said. SGI also plans to evaluate how the research effort may contribute to the evolution of the company's existing SGI® InfiniteStorage CXFSTM shared file systems.

"In this alliance with PNNL, we are committed to developing and delivering innovative storage technologies that solve problems unique to data-intensive environments," Goh said, noting that scientific and engineering dataset sizes are growing, generated by increasingly comprehensive simulations, or collected from increasingly sensitive and multi-modal sensors.

To increase the value of these datasets, SGI anticipates that data-intensive computing methods may emerge as another branch in computational science. "We've built systems with large, monolithic, globally addressable memories to contain these datasets in their entirety, which is one approach of solving the problem," Goh said. "The alliance with PNNL will work on another approach: in-storage analysis. We look forward to the possibility of incorporating results from this research into future versions of CXFS."

PNNL (www.pnl.gov) is a DOE Office of Science laboratory that solves complex problems in energy, national security, the environment and life sciences by advancing the understanding of physics, chemistry, biology and computation. PNNL employs 3,800 staff, has a \$600 million annual budget, and has been managed by Ohio-based Battelle since the lab's inception in 1965.

The original press release can be found [here](#).

Citation: PNNL, SGI to team on storage technology development for data-intensive computing (2004, June 24) retrieved 25 April 2024 from <https://phys.org/news/2004-06-pnnl-sgi-team-storage-technology.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.