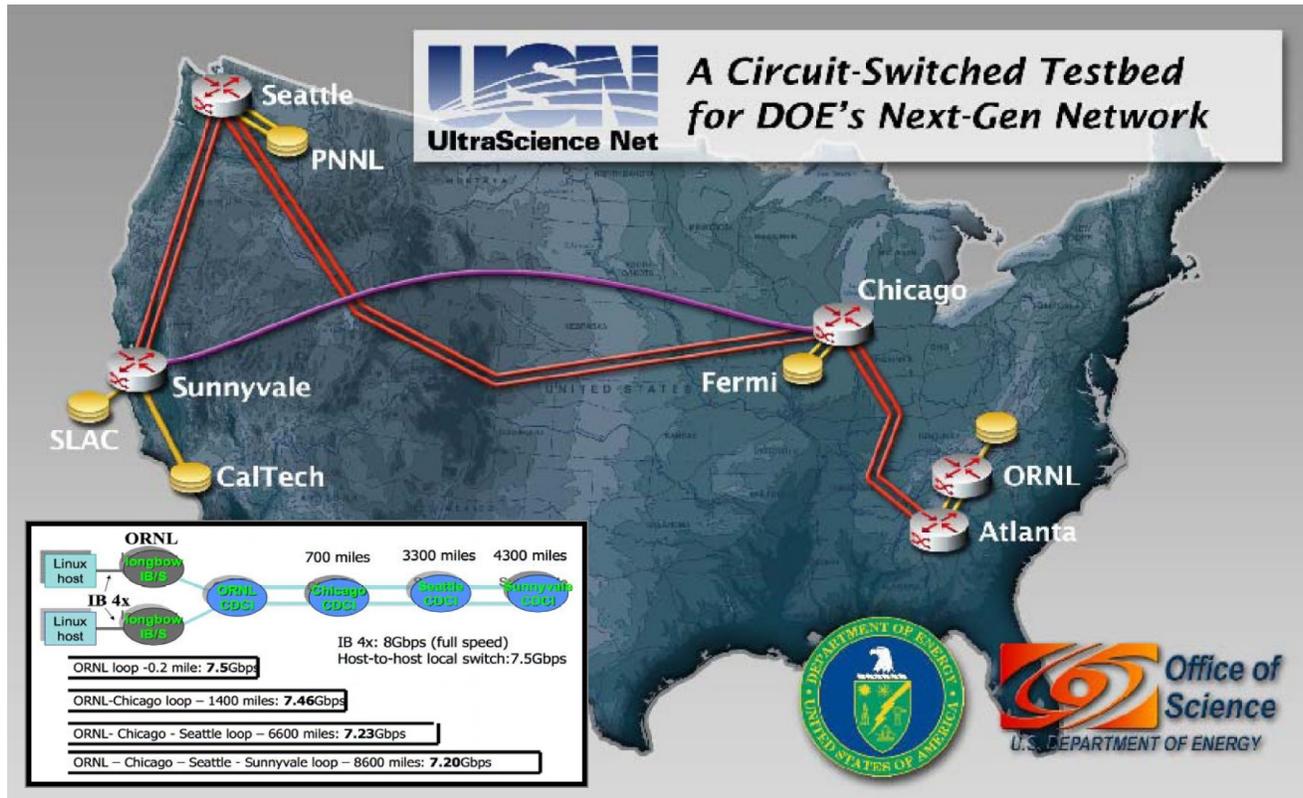# InfiniBand RDMA spans UltraScience Net

US Department of Energy's Oak Ridge National Laboratory demonstrates very high efficiency RDMA transport from Obsidian's Longbow XR InfiniBand range extenders over the entire length of UltraScience Net - 8,600 miles

Background

The DOE operates many of the world's largest "capability" supercomputers, producing and consuming data files that are growing exponentially in size as ever larger machines are deployed - increased data capacities enable larger, higher-fidelity physics simulations. However, this growth trend presents a critical problem to the DOE - when supercomputer users need to move super-sized files, shared-infrastructure TCP/IP-based networks are unable to handle the bandwidth overload. Clustering techniques for processors and storage have allowed the computing machinery to scale, but the conventional approaches to networking them together have not kept pace.

The InfiniBand protocol can support reliable connection modes as an alternative to TCP/IP, and uses a lossless flow-control mechanism better suited to highly efficient transport, but is normally limited to very short connection distances within or between equipment racks.

Obsidian Strategics developed a family of InfiniBand range extension products that allow native InfiniBand to ride 10 Gigabit optical WANs (10 Gigabit Ethernet, OC-192 SONET (Packet Over SONET/ ATM AAL5/ GFP-F) across global distances.

ORNL has been evaluating Obsidian Longbow XR devices since September 2006, running a variety of tests over USN, which is well suited as a next-generation networks testbed. Comprising a pair of OC-192 SONET channels between five locations spanning the country, it can be configured to provide 10Gbits/s network loops of varying lengths (and therefore optical delay).

The Demonstration

Using InfiniBand's low CPU overhead Remote Direct Memory Access (RDMA), the demonstration in ORNL's booth streams data across all 8,600 miles of USN with an essentially flat bandwidth-distance curve.

By dynamically reconfiguring UltraScience Net's loopback during the live demonstrations, actual RDMA bandwidth numbers from the Longbow XR pair are presented on the screen as the loopback point is adjusted from Oak Ridge to Atlanta, Chicago, Seattle and finally Sunnyvale, California.

This kind of plug-and-play performance - with no in-situ tuning - and independent of distance, is in striking contrast to normal network performance.

Applications

Bulk transport of Petascale data sets – as files continue to grow in size, it becomes increasingly infeasible to share them across networks unless transport efficiency can be raised to close to wire speed. The local-side protocol must be convenient and native to clustered storage architectures, and must also have a bandwidth roadmap at least as aggressive as wide area optical transport roadmaps. *Range-extended InfiniBand is the only approach meeting these requirements today.*