

Human Genome Sequencing Center Pushes Boundaries of Disease Research with Force10 Networks

Customer PROFILE

Customer
Human Genome
Sequencing Center



Industry
Medical Research

Application
Data Center Core

Highlights
HGSC installs Force10 E1200 switch/routers to reliably and cost effectively transport data between researchers, the supercomputer and storage facilities with line-rate, non-blocking performance.



Baylor's Human Genome Sequencing Center performs in one day the same amount of DNA sequencing that would have taken a month to complete in 1998.

In 1996, Baylor College of Medicine (Houston, Texas) was one of six institutions chosen to do a pilot program supporting the final phase of the Human Genome Project. In 1999, it was one of just three chosen to complete that final phase.

The human genome, comprised of over three billion base pairs, was expected to take 15 years to accurately sequence. But new technologies that increase sequencing productivity over 30-fold shortened the term by over two years.

For its part, Baylor's Human Genome Sequencing Center (HGSC) runs 77 sequencing machines all day, every day, performing in one day the same amount of DNA sequencing that would have taken a month to complete in 1998. The daily data stream into and out of HGSC is huge, and all of it is important to meeting the goal of 99.99 percent sequence accuracy. To support data transport between researchers, the supercomputer and the storage facilities with line-rate, non-blocking performance, reliability and cost effectiveness, HGSC installed the Force10 Networks TeraScale E1200 switch/routers.

Not Interested in Fancy Networking

There is a lot of esoteric technology involved in genome sequencing, but the packets produced are not isochronous, like those for streaming video or voice. "What we needed, frankly, was resilient, high throughput networking," said David Parker, HGSC's manager of systems. "The goal was to enable a limited number of highly skilled researchers to meet the stringent time targets," he explained, "and the budgets for doing so were far from unlimited."

To meet those objectives, HGSC required a network that could transport data at non-blocking, full line-rate speeds. It had to be scalable to allow for program expansion. And, it had to have impressive price/performance ratios, to boot. "That old image of government deep pockets is an old image," Parker smiled. "We got to be one of the three final-phase centers because we could show both competence and lower costs. Our network infrastructure played a big part in that," he added.

With the Force10 E1200, HGSC had the flexibility, scalability, and robustness it required. "Our expanded capabilities called for a line-rate,



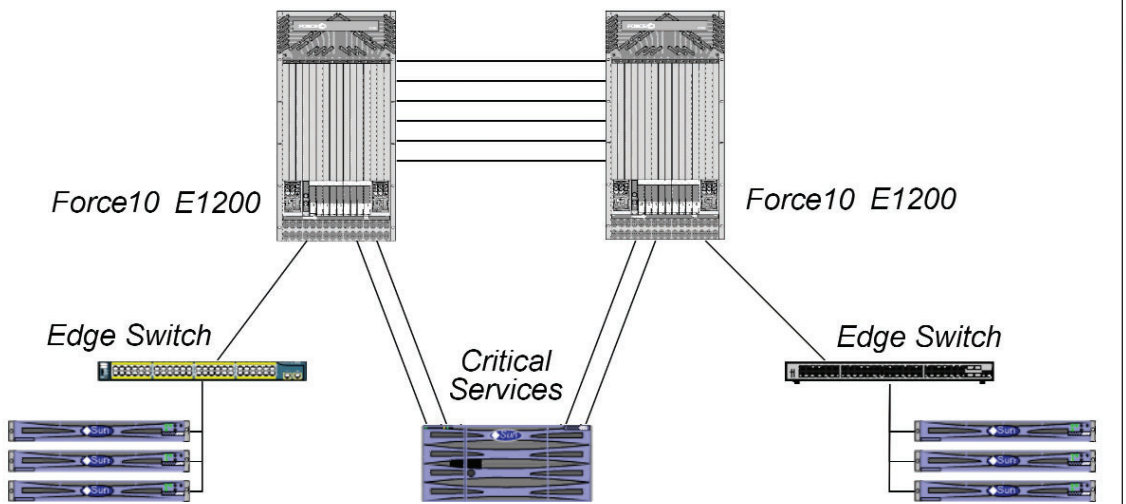
Human Genome Sequencing Center Pushes Boundaries of Disease Research with Force10 Networks

Customer PROFILE

“Our expanded capabilities called for a line-rate, non-blocking core switching infrastructure that could move large data sets on and off our computing infrastructure quickly... A pair of E1200s gave us that without strain.”

David Parker

HGSC's manager of systems,
Baylor College of Medicine



Human Genome Sequencing Center core-access architecture features 3.36 Tb/sec switching capacity at two billion packets/sec.

non-blocking core switching infrastructure that could move large data sets on and off our computing infrastructure quickly," Parker explained. "A pair of E1200s gave us that without strain." It also provided the flexibility of mixing 10 GbE and GbE line cards to meet whatever load balancing needs might emerge over time.

Additionally, the Force10 E-Series provided the scalability that would allow the center to expand its network from 100 nodes to 300, and possibly even 1,000 nodes in the future.

Not Just for Genome Work

With sequencing of human chromosome 3 complete in 2006, the Human Genome Sequencing Center finished the final stage of its multi-year project to sequence the human genome. Researchers are now using the genome information and the powerful supercomputer maintained by the center to discover the genetic basis for human disease.

The kind of network performance and value HGSC is getting from Force10 could be applied to other aspects of research, such as pharmaceuticals and complex-organ research.

For example, a key aspect of genome research is identifying those segments that are genes. Genes direct the production of proteins, which determine our hereditary qualities. "And proteins, themselves, are a huge research opportunity," says Parker.

"Protein folding is a key to designing drugs, theoretically, without having to deal with huge amounts of experimentation." Protein folding refers to the three-dimensional shape that proteins quickly exhibit after being created, and which enables them to perform their proper functions.

Another area of research that lends itself to generating lots of data and data sharing is neural modeling. "In trying to understand how our brains work, we need to better understand how neurons work, and which areas of the brain are associated with which functions," Parker explained. "In some ways modeling all the possible synaptic combinations makes sequencing three billion base pairs look easy."

But, he added, both protein-folding and neural-modeling research are highly collaborative, data-intensive efforts. "They produce a ton of binary data that needs to be moved around between databases and researchers with great reliability and consistency. A non-blocking, line-rate network would be a good fit there, too," he concluded.

Looking to the Future

With the final phase of decoding work complete, HGSC is leveraging that research to develop a deeper understanding of diseases and possible cures. The Force10 E-Series gives HGSC the high performance network it requires to continue expanding the boundaries of scientific research.



Force10 Networks, Inc.
350 Holger Way
San Jose, CA 95134 USA
www.force10networks.com

408-571-3500 PHONE
408-571-3550 FACSIMILE

© 2007 Force10 Networks, Inc. All rights reserved. Force10 and the Force10 logo are registered trademarks, and EtherScale, FTOS, SFTOS, and TeraScale are trademarks of Force10 Networks, Inc. All other brand and product names are trademarks or registered trademarks of their respective holders. Information in this document is subject to change without notice. Certain features may not yet be generally available. Force10 Networks, Inc. assumes no responsibility for any errors that may appear in this document.

CP33 607 v1.3

[PAGE 2 OF 2]