



Epigenomics Case Study

The Situation

Epigenomics is a molecular diagnostics company focusing on the development and commercialization of in vitro diagnostic tests for cancer screening and cancer specialty applications such as monitoring and classification. As such, they have become an innovative leader in cancer screening and diagnostic testing,

Epigenomics is currently developing screening techniques for colorectal, lung, and prostate cancers. Epigenomics offers diagnostic testing that utilizes DNA methylation, a technology that catches cancer in its earliest stages. Additionally, the company identifies potential biomarkers for use in drug development by its customers.

Time to market is critical for a company's success and even more so in the competitive medical test and classification landscape. For Epigenomics, developing the newest cancer screening techniques required concurrent and timely access to critical research data from amongst their researchers, both in Seattle, Oregon and from their sister site in Berlin, Germany.

Prior to contacting Cutting Edge Networked Storage, Epigenomics' storage infrastructure consisted of mainly multiple DAS (Direct Attached Storage) storage devices attached to multiple database and application servers, which in turn were connected through an internal Gigabit Ethernet network. The storage infrastructure in Berlin was similar to that in Seattle. The non-centralized storage network had to manage Epigenomics' ever increasing data proliferation - both structured (such as databases) and unstructured (such as images, video/audio, word documents, emails and spreadsheets).

Further, the infrastructure lacked the seamless ability to scale processing power and capacity as requirements grew since it was difficult to ascertain the capacity growth requirements from disparate storage devices vs. a centralized and managed common storage pool. In the interim, Berlin's datacenter implemented a Network Appliance NAS storage solution, which proved adequate, but was prohibitively expensive for the smaller Seattle site.

The Challenge

The goal of the company was to find a cost effective, flexible, highly available, scalable and feature rich storage solution to help it achieve both economic and operational benefits by consolidating disparate storage resources, reducing backup and recovery windows, complying with government regulations on data retention management, and reducing storage TCO with a tiered-storage architecture that could be replicated in both sites. Further, the new solution would have to help solve key problems such as ineffective storage utilization, the costs of managing storage, and the ability to manage storage growth surrounding data replication, disaster recovery / business continuance and backup.

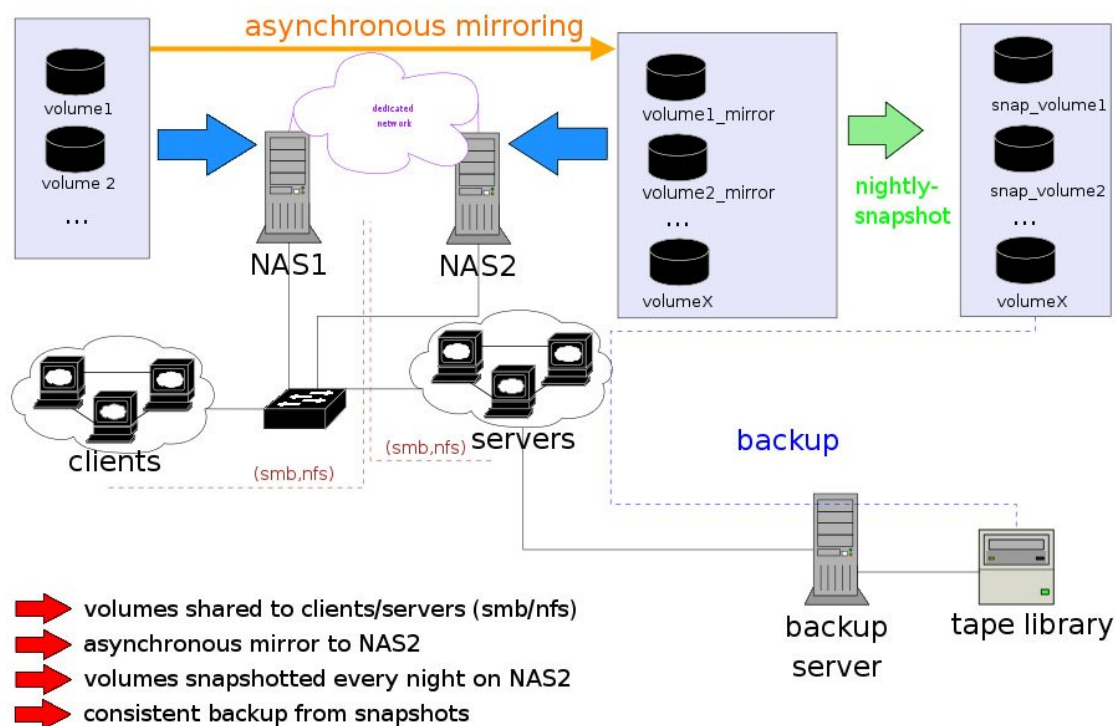
It was determined early on that the company's near term needs did not require the performance and complexity associated with Fibre Channel. Fibre Channel is widely used in enterprise data centers because of its end-to-end intelligence, consistently high performance, low latency, support of multi-path fabrics and its high availability. On the downside, Fibre Channel does not have wide industry support at present to use it for a server cluster interconnect. Also, as the number of server connections increases and the price of individual servers decline, the high cost of Fibre Channel and of Host Bus Adapters in particular becomes a concern. Further, Fibre Channel still exhibits poor interoperability amongst different vendor products after almost a decade in deployment.

Additionally, it was determined that a pure Windows based system would not support Epigenomics' growth due to its availability and flexibility limitations. Upon further analysis, they required a system that allowed seamless file exchange via a VPN with their Berlin office, strong flexibility, remote monitoring, and the ability to export multiple protocols.

The Solution

Epigenomics researched multiple vendors before deciding on Cutting Edge Networked Storage Solutions and its EdgeWare™ empowered ISS (Integrated Storage Solutions) NAS and iSCSI offerings. Networks based on Ethernet technology greatly simplify network operations, leverage a large existing pool of trained personnel, and allow the use of widely available management and monitoring applications. OpenE was considered but its operating system lacked the flexibility and data protection features of Cutting Edge's EdgeWare OS. The ISS's rich feature set, scalability, flexibility, and its user friendly and intuitive management GUI (graphical user interface) were key priorities in Epigenomics' decision making process. According to Dennis Bieling, Systems Administrator for Epigenomics, "The Cutting Edge solution allows us to seamlessly exchange files via our dedicated VPN with Berlin and to mirroring our high priority file trees. Furthermore, EdgeWare's flexibility allows easy export of multiple protocols like CIFS, HTTP, NFS and FTP, as well as integrating SNMP for remote monitoring and diagnostics. Coupled with ISS's affordability and failover capability made Cutting Edge the clear choice."

Epigenomics chose twin Cutting Edge ISS 9.6 TB integrated NAS servers as a starting point for their Seattle data center. The ISS combines enterprise level features with the ability incorporate high performance SAS drives for compute intensive applications or cost effective SATA drives where massive capacity and low TCO are priorities. Additionally, the ISS offers enterprise level features like IP Failover, high fault tolerance, snapshot data preservation, remote mirroring, and remote replication.



The present integration of the CuttingEdge NAS servers into the data center infrastructure as shown in the above diagram in Seattle includes 24 Dell Power Edge servers to store database and application software. All but 2 servers are Linux servers, while the remaining two servers run on Windows. All desktop clients run on Windows XP via a Gigabit Ethernet network. The Cutting Edge ISS NAS servers, acting as online data repositories, support all the clients and servers within the data center. The clients and servers are connected through a 3Com layer 3 master switch with traffic prioritization, which in turn, ports to all other switches with 3 lines per switch. The ISS solution also deployed NIC bonding which allows the bandwidth of multiple NIC cards to be aggregated and converted into one, increasing the data transfer capacity and speed.

The Cutting Edge ISSs run asynchronously on the dedicated Gigabit Ethernet network. Both ISS storage servers share all volumes running CIFS and NFS, with their own directories and file trees to both the clients and servers. The ISS stores home directories, emails, applications, profile data, and all documents.

Monitoring the overall system's health is easy with EdgeWare's SNMP based system monitoring and alert feature. Epigenomics chose to use Argus on top of EdgeWare, allowing both ISS NAS servers to be a part of the monitoring loop and to centralize monitoring of all servers. EdgeWare's advanced monitoring system notifies network administrators on statuses such as health disk status, system temperature, RAID status, and spare drives on the ISS servers. The EdgeWare OS also generate logs that that record monitoring activities on a timely basis for later analysis.

Additionally, because Epigenomics' data center is an integrated security environment, the ISSs act in concert with a NIS server and an LDAP server. To ensure that the data is safe and to avert a worst case disaster recovery scenario, Epigenomics also employs a multi-layered protection scheme to optimize business continuance. All network servers run independently of one another.

An asynchronous mirror is run every 24 hours on a typical 1.3 TB volume. An IP snapshot in turn, is taken off those mirrors to a backup server and Snapshots of all volumes are taken every 6 hours and used to write backups to a near line tape library. For this purpose a backup client is installed on the ISSs and data is pulled by the backup server, even though a network agent running on both Cutting Edge NAS servers would allow direct backups from the servers to a direct attached tape library. Epigenomics decided to integrate both CuttingEdge NAS servers into the already existing backup infrastructure instead.

Incremental backups are run daily A full backup on all servers are run every other week, usually on weekends where peak capacities approach 3 TB. For natural or man-made disasters, Epigenomics stores tape based backups at an offsite location.

Future Implementation of Cutting Edge ISS solutions

With the imminent release of failover protection from Cutting Edge, Epigenomics is planning to add up to 12 TB per ISS server and clustering them in the near future. Added Bieling, "at the moment, there are only asynchronous mirrors of some of the volumes. We plan to add the extra capacity while incorporating failover to allow for full synchronous operation."

Cutting Edge's ISS storage servers allowed convenient centralized management, unparalleled reliability, streamlined data protection via snapshot technology, and the capability to manage Linux and Windows based environments seamlessly. The ISS solution accommodated the client agent for Epigenomics' backbone backup server, making it simple to integrate Cutting Edge's product into the company's existing system. In comparing Cutting Edge's ISS solutions to their prior storage solution in Berlin, Bieling concluded, "Cutting Edge's NAS solutions match the Network Appliance solution in performance and is roughly 50% of the price."

The Cutting Edge Advantage

The Integrated Storage Solution (ISS) Series from Cutting Edge Networked Storage Solutions is ideal for small to medium size businesses that require affordable, dedicated network attached storage (NAS), IP SAN (IP storage area network), block level iSCSI storage solutions or support for DAS exchange consolidation that offer extremely easy setup, flexible file management, effortless in-chassis capacity expansion, sustained high-speed operation and utmost reliability. ISS makes available the option of mixing and matching low cost SATA drives and high performance 10k/15k RPM SAS drives to achieve the perfect balance between performance, capacity and cost. Built for mission critical applications like RAID protected centralized multi-protocol storage, remote replication, mirroring and data repositories, ISS speeds data access and integrates networking and file system operations in a single unit, lessening the burden of storage consolidation and file sharing in heterogeneous networks. The ISS product family provides the best cost performance and capacity storage solutions available today.

The ISS Series is powered by Cutting Edge's 3rd generation field proven EdgeWare, a fully integrated IP 64-bit storage operating system that provides a storage centric, easy-to-use and easy-to-support, heterogeneous network file system. EdgeWare offers complete NAS and iSCSI functionality and same time use in a single operating system, coupled with best-in-class performance, versatility and security for SAN, LAN and WAN applications. EdgeWare's rich feature set includes built-in remote replication for disk-to-disk backup, mirroring, IP Failover, multiple snapshot capability, iSCSI target, web-based management, RAID 6 support and built-in security features that make it a strong alternative to expensive Fibre Channel connectivity, with comparable features and performance. EdgeWare affords almost limitless volume sizes, extremely high performance of up to 200 MB/s, increased data reliability, greater security and increased scalability. iSCSI target efficiency is also improved by supporting multiple iSCSI initiators on different volumes without sacrificing NAS, CIFS and NFS performance.