



The **DESY** computers distill one hundred terabytes of raw data into ten terabytes of analysis data annually. Open-E NAS modules in the storage servers ensure that this data is made available online for the scientists to access at lightning speeds.



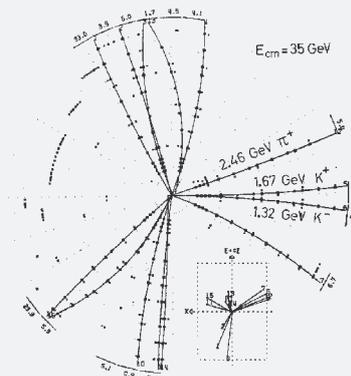
One Hundred Terabytes of Raw Data

The **Deutsches Elektronen SYNchrotron (DESY, German Electron Synchrotron)** in Hamburg, Germany, is one of the most important national and international centers for pure physics research. The center has two areas of activity: particle physics research using the HERA accelerator and research with photons in the HASYLAB. These activities produce around one hundred terabytes of raw data each year. The majority of this data is made up of measurements and, in no small part, Monte Carlo simulations, the results of which the researchers expect to measure.

In order for the particle physicists to create a picture about what is going on, data about quantities of light, hit patterns, charges and currents captured using physical detectors as well as raw data generated in the simulations must be subjected to special filter and combination processes. On the one hand, this is how the scientists separate the wheat from the chaff - or, to put it in physics jargon, separate the information from the noise. On the other, this compute-intensive procedure collates signals from various parts of the detectors into physical expressions, which are displayed as trace

images or diagrams, which is why the physicists refer to this process as reconstruction.

Automated computer farms are predominantly involved in handling this run-time intensive process at **DESY**. The computer farms are fed with raw data, which flows from the experiments into a tape library. Each of the four HERA experiments requires 100 dual CPU PCs to compute this task in the **DESY** computer center. They compress **100 TB of raw data** into around **10 TB of analysis data** and store this information on networked file servers so that the scientists can process the data further.



The scientists use 100 TB of raw data each year to "reconstruct" elementary particle paths, which helped them to discover the gluon, which is shown here.

IBM – SGI – LINUX ...



The raw data store - the tape library - is DESY's data treasure chamber. Despite the use of robotics, a tape storage system is far too slow for DESY's requirements, which is why the data has to be made available on file servers.

- The story of the mass storage and maintenance required for the data produced by the large-scale HERA accelerator at DESY all starts in 1994. IT head Dr. Knut Woller, himself a physicist, recalls that *"at that time, the computer center was running IBM 370/168 mainframes."* With the launch of the Hadron-Elektron-Ring-Anlage (HERA, Hadron Electron Ring Accelerator), which is the only one of its kind in the world, UNIX workstations started to find their way into the DESY computer center for the first time. The SGI Challenge XLs equipped with 32 RISC CPUs were general-purpose machines with their own storage subsystems.
- *"The SGI solution was able to serve hundreds of researchers per computer and was extremely successful,"* says Dr. Woller. *"But scaling these systems to keep up with the fast-growing volumes of data proved almost impossible."* And according to Dr. Woller, this meant DESY would have had to purchase the same system again, *"which neither the budget nor reason allowed."*

This is why the HERMES collaboration at DESY decided to introduce the first Linux PC farm for particle physics worldwide as early as mid-1996. This is a networked computer farm comprising ten dual processor PCs equipped with Intel Pentium Pro 200s. The DESY physicists worked in conjunction with Linus Thorwald - the father of open source UNIX. He had to be consulted repeatedly so that they could stabilize the Linux kernel for multiprocessing, *"as the Linux available back then didn't cope that well with symmetrical multiprocessing, or SMP for short, at high loads."*

The H1, ZEUS and HASYLAB experiments followed the HERMES approach in 1997. And the computer center manager then gave Dr. Woller the task of participating in the development of a DESY-wide Linux support system in 1998.

To simplify the maintenance and delivery of data, DESY commissioned SuSE to add a series of auto-installation routines to the YaST installation tool in their Linux distribution, as the number of Linux servers in the computer center had quickly grown to several hundred.

The DESY computer center began the radical task of switching the existing storage systems from SCSI and Fibre Channel (FC) technology to the substantially cheaper IDE RAID technology in 2000. This led DESY almost inevitably to Open-E.



Yesterday's technology: Dr. Knut Woller in front of the retired SGI Challenge systems.

...Open-E



Today's technology: The Open-E NAS at DESY holds a net 2.4 TB. One Open-E NAS module manages two 3Ware controllers. Eight 200 GB hard disks are connected to each of the controllers. One of the controllers is used for RAID-5 parity and the other serves as a hot spare.

- Two factors favored this development: The first was the availability of IDE hard disks with large storage capacities - IBM having launched hard disks with a capacity of 75 GB on the market in 2000. The second was 3Ware's production of IDE RAID controllers, whose outstanding feature was their multi-controller capability.
- Using these components, it was possible to put 1.5 TB of storage capacity - and the server - on a standard x86 machine for DM 40,000.
- However, there were problems with the reliability of the IBM disks at the start. Because of this, DESY first switched over to using products from Western Digital and ultimately turned to Seagate as its manufacturer of choice for the 200 GB hard disks. These hard disks are running under Linux in the computer center today and, packing 2.4 TB 'per socket', they deliver a total memory capacity of between 150 TB and 200 TB.
- But there was still room for improvement. Even though open source operating systems are considerably lower in price than, for example, a Microsoft Storage Server 2003, the costs of maintenance remained a major issue for the computer center. Added to this is the distinctive individualism of the scientists, which means that they all run different operating systems on their computers and laptops. "Open-E first approached us in 2003. Up until then, we had known them as a professional distributor for our RAID controllers," recalls Dr. Woller. Johan van den Boogaart offered DESY a compact, easy-to-maintain NAS server, which supported all the important file systems - such as Samba, NFS and NTFS - and which fitted into an IDE connector.

"Open-E Turns a Storage Server into a Plug-and-Play Appliance."

- DESY ordered the latest generation of storage servers from Comptronic. "Comptronic builds the servers according to DESY's specifications," explains Dr. Woller. DESY's commercial partner, based on the Hamburg jetties, packs two 3Ware controllers - with eight SATA hard disks each - and one NAS module from Open-E into one 4U housing.
- "This is how we achieve a storage density of a good two terabytes per server," says Dr. Woller, pleased at his coup. He goes on to confirm that: "DESY only buys Escalade controllers from 3Ware for the Linux servers." With the 6000 series ATA controllers, the DESY storage server had managed about two to three megabytes a second; with the 7000 series, the server achieved 30 megabytes per second. As a precaution, the computer center skipped the first SATA generation. "Today we run Escalade 9500 controllers in a RAID 5 farm and achieve a throughput of 100 MB per second. This means that in many applications, they can compete with the considerably more expensive SCSI or Fiber Channel RAID systems."
- No problem for the DESY backbone - it transports one gigabyte per second...



It fits onto the IDE port, contains both the operating system and the NAS server, and controls several 3Ware SATA RAID controllers - the NAS module from Open-E.



Dr. Woller sits back with satisfaction. The latest update to the Enterprise version of the Open-E NAS module provides him with the Active Directory Support he needs for the scientists' laptops.

Cost Savings and Increased Security

- But impressive speeds are not the whole story. Costs and security were two more reasons why **DESY** chose the Open-E solution. *"The Open-E NAS solution does not recognize any local users,"* explains Dr. Woller. *"This means no one can log in without authorization - a real plus in terms of security!"*.
- And Dr. Woller also underlines the issue of cost: *"For the price of two system disks, one floppy or CD-ROM boot drive and a Windows Storage Server 2003, you can easily afford one Open-E Enterprise Module."* Yet this is merely the cost of acquisition. There is extra potential for cutting costs in terms of manpower and downtime: *"Even the Linux operating system needs updating regularly, and that takes longer than the flash upgrade for the Open-E NAS module."*
- Dr. Woller is very pleased with the latest upgrade: *"We now have Active Directory Support since we installed the upgrade from December 3, 2004. Above all, we really need this for those Windows laptop users who back up their hard disks on the NAS."* This is a benefit because, thanks to Active Directory Support, the **DESY** computer system no longer needs to maintain several access control systems. The new multi-snapshot feature, which offers a scheduling function and distributes administrative tasks over several admin accounts, is also a welcome addition.

Conclusion: It Doesn't Get Any Better. Open-E Provides Ideal NAS Solutions to Suit Large Companies' Massive Storage Requirements

This is important because **DESY** can easily match any large company in terms of its data storage requirements. The current research projects at **DESY** also mean this demand for storage is set to grow. And if the new 22 mile long International Linear Collider (ILC) is built at the **DESY** site - which could well happen - then the storage requirements will explode once more. This future project would fire electrons at positrons, their antimatter counterparts, with such force that it would provide a tool for examining the structure and origins of matter and the universe at greater depth and with more accuracy than ever before. The ILC may be built in either the USA, Japan or Europe. If Europe is selected, the most likely place to site the ILC would be at the **DESY** in Hamburg.

- Dr. Woller remains unperturbed by the potential amounts of storage this would require: *"Data volumes in the multiple terabyte range are easy to manage thanks to the Open-E NAS Enterprise Module."* According to Dr. Woller, the arrival of the iSCSI standard is unstoppable: *"iSCSI will establish itself as the standard in computer centers,"* reflects Dr. Woller. *"I wouldn't have believed that only a year ago. But while the throughput in Fibre Channel technology doubles from one generation to the next, Ethernet bandwidth continues to grow in orders of magnitude."*
- So it is appropriate that iSCSI support will be one of the new developments Open-E will present at CeBIT 2005.

open-e

Open-E GmbH
Lindberghstraße 5
D-82178 Puchheim, Germany

Tel.: +49 (89) 800777-0
Fax: +49 (89) 800777-17
<http://www.open-e.com>