



Aus Einhundert Terabyte Rohdaten destillieren die **DESY-Rechner** jährlich zehn Terabyte Analysedaten und, Open-E NAS-Module in Storage-Servern sorgen dafür, dass sie blitzschnell den Wissenschaftlern online zur Verfügung stehen.



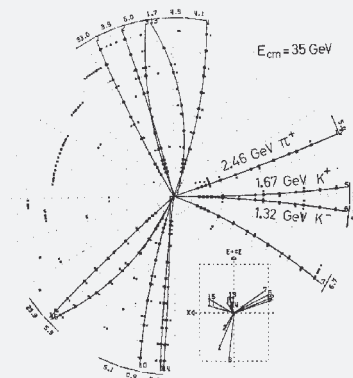
Einhundert Terabyte an Rohdaten

Das **Deutsche Elektronen SYNchrotron (DESY)** in Hamburg zählt zu den national wie international bedeutendsten Einrichtungen der physikalischen Grundlagenforschung. In seinen beiden Forschungsbereichen Teilchenphysik an dem Beschleuniger HERA und Forschung mit Photonen im HASYLAB fallen pro Jahr rund Einhundert Terabyte an Rohdaten an, wobei der größte Teil aus Messungen stammt, ein nicht unerheblicher Teil aber aus Monte-Carlo-Simulationen besteht, was die Forscher zu messen erwarten.

Damit sich die Teilchenphysiker ein Bild des Geschehens machen können, müssen diese aus physikalischen Detektoren stammenden Messwerte von Lichtmengen, Hit-Patterns, Ladungen oder Strömen ebenso wie die in Simulationen generierten Rohdaten spezielle Filter- und Kombinationsprozesse durchlaufen. So trennen die Wissenschaftler zum einen die Spreu vom Weizen – oder in der Sprache der Physik – die Information vom Rauschen. Zum anderen fügt dieser rechenintensive Vorgang die Signale aus den verschiedenen Teilen der Detektoren zu physikalischen Aussagen

zusammen, die als Spurenbilder oder Diagramme dargestellt werden, weswegen ihn die Physiker Rekonstruktion nennen.

Mit diesem rechenzeitaufwändigen Vorgang sind am **DESY** weitgehend automatisiert arbeitende Rechnerfarmen beaufschlagt. Sie werden mit den Rohdaten gefüttert, die aus den Experimenten in eine Tape-Library fließen. Im Rechenzentrum des **DESY** sind mit dieser Aufgabe für jedes der vier HERA-Experimente 100 Doppel-CPU-PCs besetzt. Sie verdichten die **100 TByte Rohdaten** auf rund **10 TByte Analysedaten** und legen sie zur Weiterverarbeitung durch die Wissenschaftler auf vernetzten File-Servern ab.



Aus 100 Terabyte Rohdaten jährlich "rekonstruieren" die Wissenschaftler Elementarteilchenbahnen wie der hier gezeigten Gloun-Entdeckung.



Das Rohdatenlager – die Tape-Library – ist die Datenschatzkammer des DESY. Trotz Robotik ist ein Bandspeicher viel zu langsam, weshalb die Daten auf File-Servern bereitgestellt werden müssen.

IBM – SGI – LINUX ...

- Die Geschichte der massenhaften Speicherung von Daten des Großbeschleunigers HERA und deren Vorhaltung beginnt beim DESY im Jahr 1994. „Im Rechenzentrum arbeiteten seinerzeit Mainframes des Typs IBM 370/168“, erinnert sich Dr. Knut Woller, der seine Position – selbst ganz Physiker – scherzhaft mit IT-„Halbleiter“ beschreibt. Mit der weltweit einmaligen Hadron-Elektron-Ring-Anlage (HERA) hielten erstmals UNIX-Workstations Einzug ins DESY-Rechenzentrum: Die mit 32 RISC-CPU's bestückten SGI Challenge XL waren Universal-Maschinen mit eigenem Speicher-Subsystem.
- „Die SGI-Lösung konnte zwar pro Rechner hunderte von Forscherinnen und Forschern bedienen und war überaus erfolgreich“, erinnert sich Dr. Knut Woller. „Diese Systeme mit der schnell wachsenden Datenmenge zu skalieren erwies sich jedoch als fast unmöglich“. So hätte DESY nur die Wahl gehabt, die gleiche Anlage erneut anzuschaffen, „was weder das Budget noch die Vernunft bergab.“

So beschränkt die HERMES-Kollaboration am DESY bereits Mitte 1996 den Weg zur weltweit ersten LINUX-PC-Farm der Teilchenphysik, einem vernetzten Rechnerverbund, bestehend aus zehn Doppel-Processor-PCs, bestückt mit Pentium Pro 200 von Intel. Linus Thorwald – Vater des Open-Source-UNIX – musste mehrfach konsultiert werden, um in Kooperation mit den DESY-Physikern den LINUX-Kernel im Multiprozessorbetrieb stabil zu machen, „denn Symmetrisches MultiProcessing, kurz SMP, unter hoher Last konnte das LINUX von damals noch nicht gut“.

Dem HERMES-Ansatz folgten im Jahr 1997 die Experimente H1, ZEUS und das HASyLab. Und schließlich erhielt Dr. Woller im Jahr 1998 vom Rechenzentrumsleiter den Auftrag, beim Aufbau eines DESY-weiten LINUX-Supports mitzuwirken.

So beauftragte DESY die Firma SuSE, das Installationswerkzeug YaST ihrer LINUX-Distribution, zur Vereinfachung der Bereitstellung und der Wartung um eine Reihe von Autoinstallationsroutinen zu erweitern, denn die Zahl der Linux-Server im Rechenzentrum war schnell auf einige hundert angewachsen.

Im Jahr 2000 machte sich das DESY-Rechenzentrum an den radikalen Umbau der bestehenden Speichersysteme von SCSI- und FibreChannel (FC) auf die wesentlich preiswertere IDE-RAID-Technologie – und stieß dabei fast zwangsläufig auf Open-E.



Technik von gestern: Dr. Knut Woller vor den ausgemusterten SGI-Challenge-Systemen.

...Open-E

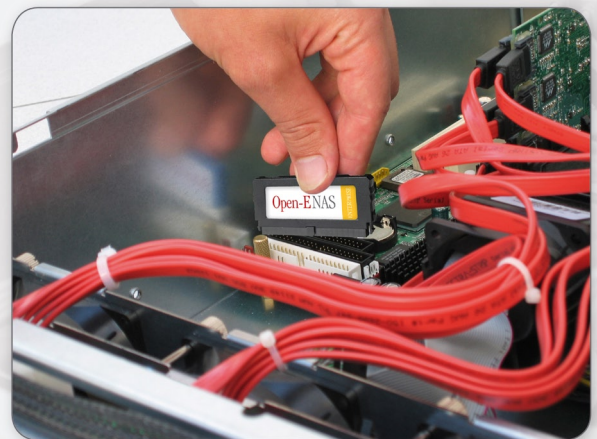


Technik von heute: Das Open-E NAS beim DESY fasst netto 2,4 TByte. Ein Open-E NAS-Modul steuert zwei 3Ware-Controller, an denen jeweils acht 200-GByte-Harddisks hängen, von denen eine als RAID-5-Parity und eine als Hot-Spare-Platte dient.

- Zwei Gründe hatten diese Entwicklung begünstigt: Zum einen die Verfügbarkeit von IDE-Festplatten großer Speicherkapazität – IBM hatte im Jahr 2000 Festplatten mit einer Kapazität von 75 GByte auf den Markt gebracht. Zum anderen die Entwicklung der IDE-RAID Controller von 3Ware, deren herausragende Eigenschaft ihre Multi-Controller-Fähigkeit ist.
- Mit diesen Komponenten ließen sich in einer Standard-x86-Maschine für 40.000 DM 1,5 TByte Speicherkapazität bereitstellen – einschließlich Server.
- Zwar gab es anfangs einige Probleme mit der Ausfallsicherheit der IBM-Platten. Daher wechselte das DESY zunächst zu Western Digital um schließlich bei Seagate als Hersteller der 200 GB Harddisks zu enden, die heute im Rechenzentrum unter Linux laufen und mit 2,4 TByte „pro Büchse“ ein Gesamtspeichervolumen von 150 bis 200 TByte bereitstellen.
- Doch es ging noch besser. Denn obwohl ein Open-Source-Betriebssystem erheblich preiswerter ausfällt als beispielsweise ein Storage Server 2003 von Microsoft – die Aufwände für die Wartung blieben eine große Herausforderung für ein Rechenzentrum; ganz zu schweigen davon, dass Wissenschaftler zu ausgeprägtem Individualismus neigen, was zu einer Vielfalt an Betriebssystemen auf deren Rechnern und Notebooks führt. „In Jahr 2003 trat erstmals Open-E an uns heran, die wir bis dato als kompetenten Distributor für unsere RAID Controller kannten“, erinnert sich Dr. Woller. Johan van den Boogaart stellte dem DESY einen kompakten, pflegeleichten NAS-Server in Aussicht, der die wichtigsten Dateisysteme allesamt unterstütze – Samba, NFS und NTFS – und in einem IDE-Stecker Platz fand.

“Open-E macht einen Storage-Server zur Plug-and-Play-Appliance.”

- DESY gab bei Comptronic Storage-Server der neuen Generation in Auftrag: „Comptronic baut die Server nach der DESY-Spezifikation“, erklärt Dr. Knut Woller. In ein 4U-Gehäuse packt der DESY-Handelspartner von den Hamburger Landungsbrücken je zwei 3Ware-Controller mit jeweils acht SATA-Festplatten – und ein NAS-Modul von Open-E.
- „Damit erreichen wir eine Speicherdichte von gut zwei Terabyte pro Server“, freut sich Dr. Woller über den Coup und bestätigt: „DESY kauft für die Linux-Server durchgängig Escalade-Controller von 3Ware“. Mit den 6000er-ATA-Controllern hatte der DESY-Storage-Server zwei bis drei MByte pro Sekunde geschafft, mit den 7000ern waren es 30 MByte pro Sekunde. Die erste SATA-Generation hatte das Rechenzentrum sicherheitshalber übersprungen. „Heute betreiben wir Escalade 9500 Controller im RAID-5-Verbund und erreichen damit einen Durchsatz von 100 MByte pro Sekunde. Damit sind sie für viele Anwendungen konkurrenzfähig zu wesentlich teureren SCSI- oder FibreChannel-RAID-Systemen“.
- Kein Problem für den DESY-Backbone – der transportiert ein Gigabyte pro Sekunde...



Passt auf dem IDE-Port, enthält Betriebssystem und NAS-Server und steuert mehrere 3Ware-SATA-RAID Controller – das NAS-Modul von Open-E.



Zufrieden lehnt sich Dr. Knut Woller zurück. Das jüngste Update der Enterprise-Version des Open-E NAS-Moduls gibt ihm den Active Directory Support für die Notebooks der Wissenschaftler.

Kostenersparnis und Sicherheitsgewinn

- Aber beeindruckende Geschwindigkeit ist längst nicht alles. Kosten und Sicherheit sind zwei weitere Argumente, die für die Open-E-Lösung sprechen. „Das Open-E NAS kennt keinen lokalen User“, erläutert Dr. Woller. „So kann sich auch kein ungebetener einloggen – ein echtes Plus an Sicherheit!“.
- Und auf der Kostenseite rechnet Dr. Woller vor: „Für zwei Systemfestplatten, ein Floppy- oder CD-ROM-Bootlaufwerk und einen Windows Storage Server 2003 kann man leicht ein Open-E Enterprise-Modul kaufen“. Aber das sind nur die Anschaffungskosten. Zusätzliches Sparpotenzial sieht er bei Manpower und Downtime: „Ein selbst bei LINUX immer wieder fälliges Upgrade des Betriebssystems dauert länger als ein Flash-Upgrade des Open-E NAS-Moduls“.
- Das jüngste macht ihn richtig glücklich: „Seit dem Upgrade vom 3. Dezember 2004 haben wir Active Directory Support, den wir vor allem für unsere Windows-Notebook-User so gut brauchen können, die ihre Festplatten ins NAS sichern“. Denn so braucht das DESY-Rechenzentrum nicht mehrere Zugangskontrollsysteme zu pflegen. Ferner sehr begehrt: Das neue Multi-Snapshot mit Scheduling-Funktion und die Aufteilung administrativer Aufgaben in verschiedenen Admin-Accounts.

Fazit: Günstiger geht es nicht. Open-E liefert ideale NAS-Lösungen für den riesigen Speicherbedarf von Großunternehmen

- Denn mit dem Speicherbedarf eines Großunternehmens kann DESY locker mithalten. Die laufenden Forschungsprojekte am DESY werden für weiteres Wachstum sorgen. Und wenn – was durchaus im Rahmen des Möglichen ist – auch noch der neue ca. 35 km lange Internationale Linear Collider (ILC) kommen sollte, wird der Storage-Bedarf noch einmal explodieren. Dieses Zukunftsprojekt schießt Elektronen derart wuchtig auf ihre Antiteilchen, die Positronen, dass man tiefer und genauer in die Struktur und Entstehung der Materie und des Universums blicken kann als je zuvor. Der ILC soll entweder in den USA, in Japan oder in Europa gebaut werden. Kommt er nach Europa, wäre der Standort mit großer Wahrscheinlichkeit bei DESY in Hamburg.
- Der zu erwartende Speicherbedarf lässt Dr. Woller kalt: „Mit dem Open-E NAS-Enterprise-Modul sind Datenmengen im Multi-Terabyte-Bereich locker handhabbar.“ iSCSI sei unaufhaltsam im Kommen: „iSCSI wird sich in Rechenzentren durchsetzen“, sinniert er. „Das hätte ich vor einem Jahr noch nicht geglaubt. Aber während sich bei Fibre Channel von einer Generation zur nächsten der Durchsatz verdoppelt, wächst die Bandbreite von Ethernet noch immer in Zehnerpotenzen“.
- Da trifft es sich, dass der iSCSI-Support zu den CeBIT-Neuheiten 2005 von Open-E zählt.

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>