

## HEPNT

---

### Hauptthemen

- Migration zu Windows 2000
- Automatisierte W2000 Installation
- Authentisierung und Passwortsynchronisation mit UNIX
- Webservice mit Windows

## Migration zu W2000

---

Pläne in fast allen Labs

meist noch im Testbetrieb

Ausarbeitung von Strategien für

- Installationsservice
- Strukturierung und Einsatz Active Directory
- Authorisierung (Group Policies)

## Automatisierte W2000

### Installation

---

#### Mehrere Varianten in Diskussion

- native W2000 Tools: Remote Install Service (RIS) und Group Policies (GPO)
- Herstellertools, z.B. Drive Image Professional und WinInstall LE
- eigene Lösungen (z.B. bei DAPNIA)

Grundtenor: W2000 Tools sehr komplex, trotzdem teilweise nicht flexibel genug (GPO), noch kein Patentrezept in Sicht

## Authentisierung

---

JLAB hat Tool zur Passwortsynchronisierung zwischen UNIX und NT

Vortrag über Probleme mit Authentisierungsprotokollen

## Webservices mit Windows

---

### Vorschlag vom CERN:

- Abbildung vieler Services auf http
- Webinterface zum transparenten Zugriff
- Bis zu Zugriff auf Filesystem (WEBDAV)
- Verstecken von Servern hinter einheitlichem Namensschema ([www.cern.ch/service](http://www.cern.ch/service))

Probleme aus meiner Sicht: “Tunnelung” aller Services über http, damit Aushebelung der Sicherheitsmechanismen

## UNIX

---

### Hauptthemen:

- Aufbau und Automatisierung von Farmen
- Tools zur Systemüberwachung
- Batchsysteme
- LHC Computing
- AFS

## Aufbau und Automatisierung von Farmen

---

Viele Labs bauen Linux Farmen auf

- Basis PC oder Alpha Plattform
- Dickes Netzwerk, wenn für Parallelprogrammierung vorgesehen
- Aufbau meist Formfaktor 1U und 2U in Racks
- Fileserver mit IDE Platten bei DESY, FNAL

Automatisierte Installation mit Redhat, SuSE, eigenen Tools (Tailor SLAC)

## Linuxfarmen bei DESY

---

Automatisierte Installation als gemeinsames Projekt zwischen DESY und SuSE

Konzept wurde Anfang 2000 entwickelt

Erste Testinstallation im Dez. bei DESY

Aufbau eines Installationservers

Software und Dokumentation wird Bestandteil der Standard SuSE Distribution

## Tools zur Systemüberwachung

---

### Mindestens 5 Eigenentwicklungen

- FNAL: NGOP (in Entwicklung)
- CERN: PEM (in Entwicklung)
- IN2P3: GMS (in Benutzung, JAVA basiert)
- DESY (Zeuthen): Scout, weiterentwickeltes SLAC Tool
- SLAC: Ranger (Neuentwicklung, Erprobung)

keine Koordinierung gewollt!?

## Batchsysteme

---

LSF nach wie vor am weitesten verbreitet

Kundenbindung durch Kampfpreise

Einige Sites weichen auf PBS aus

- Erfahrungen dazu klingen nicht gut

Condor auch im Einsatz, andere Zielrichtung

- Beschäftigung leerstehender Computer
- Aufwand/Nutzen in Ordnung?

## LHC Computing

---

Basis: PC Farmen (1000..10 000 Systeme)

daher neue Konzepte für

- HW Configuration Management
- SW Installation Mgmt (OS und Applications)
- System monitoring und Problem Management

Umsetzung des Grid-Konzeptes

- Veröffentlichung der angebotenen Dienste
- Authentisierung/Authorisierung, Policies
- Scheduling, Vorreservierung für wichtige Aufgaben
- garantierte Quality of Service, Verfügbarkeit von Ressourcen

## AFS (nicht im Web)

---

AFS Support in Zukunft über 1st level IBM Support

2nd level Support dann wie bisher

Preise für Support werden neu ausgehandelt

- keine Site Lizenzen mehr?
- Andsere Gebührenmodelle
- in Europa teilweise seit 2 Jahren keine Vertragsverhandlungen mehr

Ende des Services für Version 3.6 12/2002

keine Pläne für neue AFS Version

kein kommerzielles Interesse an AFS





## Grid Anwendungen

---

### Multidisziplinäre Simulationen

Probleme mit großem Parameterraum, der auf viele Einzelruns aufgeteilt wird

### Koordinierte Analysen großer Datenmengen

- Datensätze (Petabytes) in HEP Kollaborationen

### Benutzung weit verteilter Datenarchive

(Metereologie usw.)

### Bündelung von Supercomputern zur Bearbeitung von Problemen in Echtzeit

## Grid Services

---

### Wichtige Komponenten

- Ressource discovery
- Execution Queues Management
- Management verteilter Anwendungen
- Monitoring und Error Recovery
- Checkpoint/Restart
- Access Control
- Global File System
- Grid Kommunikationsbibliotheken

## Grid Implementierungen

---

Einige Projekte mit laufenden  
Gridkomponenten (z.B. NASA's  
Information Power Grid, IPG )

Implementierungen basieren meist auf Globus

Wichtige Globus Services

- Auth per Zertifikationsserver (wie DESY CA)
- Optimierter Filetransfer (Globus-ftp)
- Kommunikation mit Globus MPI

## Infos zu Grid

---

Vorträge zum Thema Grid bei DESY

- Vorträge aus CERN, FNAL, KFA geplant

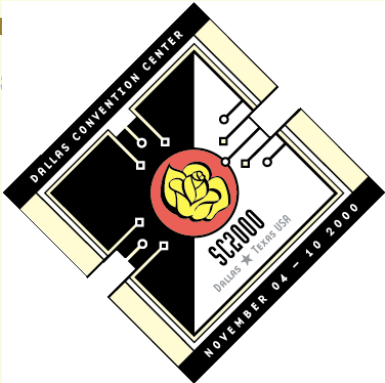
<http://www-fp.globus.org/>

Supercomputing 2000

---

**COTS (Commodity-Off-The-Shelf) Cluster System for High Performance Computing**

- Cluster components
- Cplant, Bunyip, KLAT2
- Special architectures
- Grape-6



Supercomputing 2000, COTS (Commodity-Off-The-Shelf) Cluster System for HPC

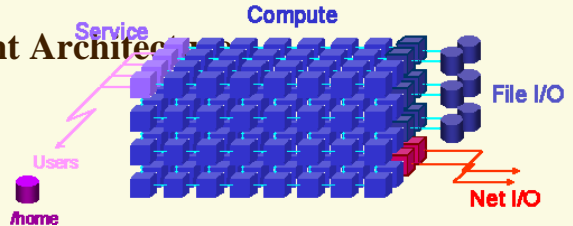
---

**Cluster components:**

- Standard PCs, workstations, rack mounts with PentiumIII, Alpha processors
- Network: Myrinet2000 (Myricom, 500 Mbytes/sec), QNet (QSW, 340 Mbytes/sec), (Fast Ethernet)
- Software Modell:
  - MPI - Message passing interface, MPI-2, MPICH, ...

Supercomputing 2000, COTS Cluster System for HPCcont...

- **CPlant at Sandia Lab (large cluster installation):**
- **Up to 20000 Alpha EV7 in 2004/5**
- **CPlant Architecture**



The diagram illustrates the CPlant architecture. It features a large grid of blue nodes representing compute nodes. A purple node labeled 'Users' is connected to a 'Service' node. The main grid is labeled 'Compute'. To the right, there are 'File I/O' and 'Net I/O' components.

Supercomputing 2000, COTS Cluster System for HPC, CPlant

- **Service nodes:**
- **Users interaction, launch parallel programs, provide input, receive output, debug, and monitor performance**
- **Typically configured with the features of standard workstation plus debuggers and performance monitors**
- **Compute nodes:**
- **Compute partition provides the compute cycles to the applications**
- **File I/O:**
- **High-performance parallel filesystem,**

Supercomputing 2000, COTS Cluster System for HPCcont...

**Cluster applications:**

- Bunyip, Australian National University:**
  - 92cent/Mflops, Ultra-Large-Scale Neural-Network Training on a PIII Cluster
  - Gordan Bell award for best price/performance,
  - 98 PIII 550 MHz dual CPU PCs, Fast Ethernet switching,
  - use of 32-bit PIII SIMD streaming instructions, 163 GFlops sustained
- KLAT2, University of Kentucky:**
  - High-Cost CFD on a low cost Cluster
  - 64 Athlon 700 MHz single CPU PCs, 4 fast Ethernet interfaces/each

Supercomputing 2000, Special architecture  
GRAPE-6 (GRAvity Pipe)

1.349 Tflops (768k particles) simulation of black holes in a galactic center,  
University of Tokyo

**Architecture:**

```
graph LR; Host[Host] --- PCI[PCI]; PCI --- HI1[Host interface]; PCI --- HI2[Host interface]; HI1 --- NB1[Network Boards]; HI2 --- NB2[Network Boards]; NB1 --- PB1[Processing Boards]; NB1 --- PB2[Processing Boards]; NB1 --- PB3[Processing Boards]; NB2 --- PB4[Processing Boards]; NB2 --- PB5[Processing Boards]; NB2 --- PB6[Processing Boards];
```

Supercomputing 2000, Special architecture  
GRAPE-6

1.349 Tflops (768k particles) simulation of black holes  
in a galactic center,  
University of Tokyo

Processor chip:

SSRAM      Memory interface      Predictor Pipeline      Force Pipeline

Network Interface

Supercomputing 2000, Future Supercomputing, SRC-6

**SRC-6 SMP**

SRC founded 1996  
by Seymour Cray

Multi Adaptive-Processor

Up to 512 Intel CPUs

256 memory ports

Segment 1, Board 1

Board 2

Board 16

µProc    I/O    Bridge    Private Memory    MAP

Read Crossbar

Write Crossbar

Common Memory

