



Management von Linux Installationen am DESY

Uwe Ensslin

24. 2. 2005

Das Umfeld

- 2 Standorte, Hamburg (HH) und Zeuthen (Zn)
- Beschleunigerbetrieb (HH)
- 4 Hochenergie Experimente (HH)
- Neue Projekte: XFEL, Petra III
- Synchrotronstrahlungslabor HASYLAB (HH)
- PITZ Photoinjektor (Zn)
- AMANDA/ICECUBE Experiment (Zn)



Linux bei DESY



- **Bereitstellung von Infrastruktur zur Speicherung und Archivierung von Experimentdaten**
 - ★ Rate während der Datennahme ca. 500 GB/Tag
 - ★ Archivierung in 4 STK Silos (ca. 100 TB pro Jahr)
- **Unterstützung des Forschungsbetriebes**
 - ★ Gewährleistung des performanten Zugriffs auf Experimentdaten
- **Betrieb von Computing Ressourcen für die Experimente**
 - ★ Troubleshooting von > 1000 Rechnern
 - ★ Betreuung von Diensten



- **Wir erwarten eine Zunahme der Anzahl der zu betreuenden Rechner durch**
 - ★ Den Bau von XFEL (http://xfel.desy.de/content/e169/index_eng.html)
 - ★ Den Ausbau von PETRA (<http://petra3.desy.de/>)
 - ★ Den Aufbau einer GRID Infrastruktur (<http://grid.desy.de/>)
 - ★ Ein neues Teilchenphysikprogramm nach HERA
- **Wir erwarten keinen Ausbau der Personalkapazität**

Die Herausforderung



- Die schiere Zahl der zu betreuenden Systeme O(1500)
- Große Bandbreite an zu unterstützenden HW Komponenten
- Das Alter der betriebenen Systeme
 - ★ Ausfälle, Leistungsfähigkeit --> Migrationsträgheit
- Support neuer Hardware
- Unterstützung von 64 Bit Systemen
- 554 installierte SuSE Pakete, 467 DESY Anwendungssoftware Pakete
 - ★ Aktualisierung der installierten Pakete (Bugfixes, Security)
- Legacy-Software

- **Hardware**
 - ★ Standardisierung von Desktops durch eine Arbeitsgruppe
 - ★ IT Empfehlungen für Server und Notebooks
- **Software**
 - ★ Ablösung SuSE Professional durch Scientific Linux (<http://www.scientificlinux.org>)
- **Tools, die ..**
 - ★ .. ein zentrales Management
 - ★ und verteilte, automatische Administration erlauben



- **Die DESY Frameworks sind Programmsuites zur**
 - ★ **Verwaltung maschinenbezogener Daten**
 - ★ **Speicherung von Regeln und Verfahren**
 - ★ **Erzeugung von technischen Infrastrukturdaten**
 - ★ **Durchführung von (Post-) Installation und Konfiguration**
 - ★ **Pflege installierter Maschinen**
- **Aus historischen Gründen wurden zwei Frameworks parallel entwickelt und auf die Standorte optimiert**

- **Der Weg zu einem gebrauchsfähigen Rechner**
 - ★ **Zentrale Registrierung des Systems (Manuell)**
 - **IP-Adresse, Routing, Interface, Partitionierung, SCSI vs. IDE, Gruppen- und Clusterzugehörigkeit, 'Service Flags', Ansprechpartner, Standort, Inventarnummer**
 - ★ **Die Bereitstellung der technischen Infrastruktur (Tool)**
 - **NIS, DNS, DHCP/QIP, PXE**
 - **Erzeugung von YaST(2)-, Kickstart- und Jumpstart Profilen**
 - ★ **Anstoß der Installation**
 - **Boot/Installationsmedien, PXE-Boot**
 - **Installationsscript erlaubt Installation im Batch Modus**
 - ★ **Warten auf die Login Maske**



- **Die Basisinstallation**

- ★ **Installiert das OS, initiale Netzwerkkonnektivität und Tools**
- ★ **Die Boot/Installationsmedien enthalten**
 - **Einen Zeiger auf den Installationsserver und das Maschinenprofil (YaST 2 XML Profil oder 'kickstart' Datei)**
 - **Den Installationskernel, der alle im Einsatz befindlichen Maschinentypen unterstützt**
- ★ **Der Installationskernel bootet und startet das Installationsprogramm der Distribution (YaST2 oder Anaconda)**
- ★ **Das Installationsprogramm führt, gesteuert durch das Maschinenprofil, eine automatische Basisinstallation durch ..**
- ★ **.. und übergibt im Anschluß die Kontrolle an die Postinstallation**



- **Die Postinstallation**

- ★ spielt die Software der Distribution ein, die nicht Teil der Basisinstallation ist
- ★ konfiguriert das System
 - ➔ Eintrag statischer IP-Konfiguration, Mail-Server, Clusterspezifisches Root-Passwort, Nutzerzugang, AFS-Konfiguration, DESY Printing System, ssh keys, ..
- ★ initiiert die Installation der DESY-Anwendungssoftware
- ★ erzeugt die nötigen Hooks für die Wartung von Konfiguration und Software
- ★ Führt einen Reboot durch, nach welchem die Maschine benutzungsfähig ist.



- **Laufende Aktualisierung installierter Systeme durch das Framework**
 - ★ Konfigurationsänderungen werden zentral konfiguriert und dann auf Maschinen(gruppen) eingepflegt
 - ★ Einspielen von aktualisierten RPMs von SuSE und DESY
 - ★ Neue RPMs als Erweiterung des Funktionsumfangs werden zentral konfiguriert und eingepflegt
 - ★ Technisch: Aktivierung des Frameworks via 'cron'
 - **Nachts:** Routine, größere Änderungen
 - **Tags:** Routine, kleinere Änderungen
 - **Stündlich:** Nach Bedarf, etwa Security Updates

Sonderfall: Kernel Upgrades (HH)



- Erfolgen nach Absprache mit den Nutzern
- Es gibt eine automatische Vorbereitung durch das Framework
- Die Aktivierung geschieht durch kontrollierten Reboot
- Auch dies ließe sich automatisieren ...

DESY

Erfahrungen in Hamburg




- Das Framework ist seit 12 Jahren im Einsatz und wird kontinuierlich entwickelt und angepaßt
- Der Einsatz erfolgte in der Vergangenheit plattformübergreifend auf AIX, IRIX, HP-UX, Solaris und Linux
- Die Einarbeitung neuer Mitarbeiter und die Nutzung durch 'Gelegenheits-Anwender' ist problematisch
- Eine aktive Einbindung der Admins der Experimente konnte nicht realisiert werden
- SQL-artige Query Möglichkeiten fehlen
- Es gibt inzwischen neue Tools, die auf das kommende Umfeld (LCG) zugeschnitten sind

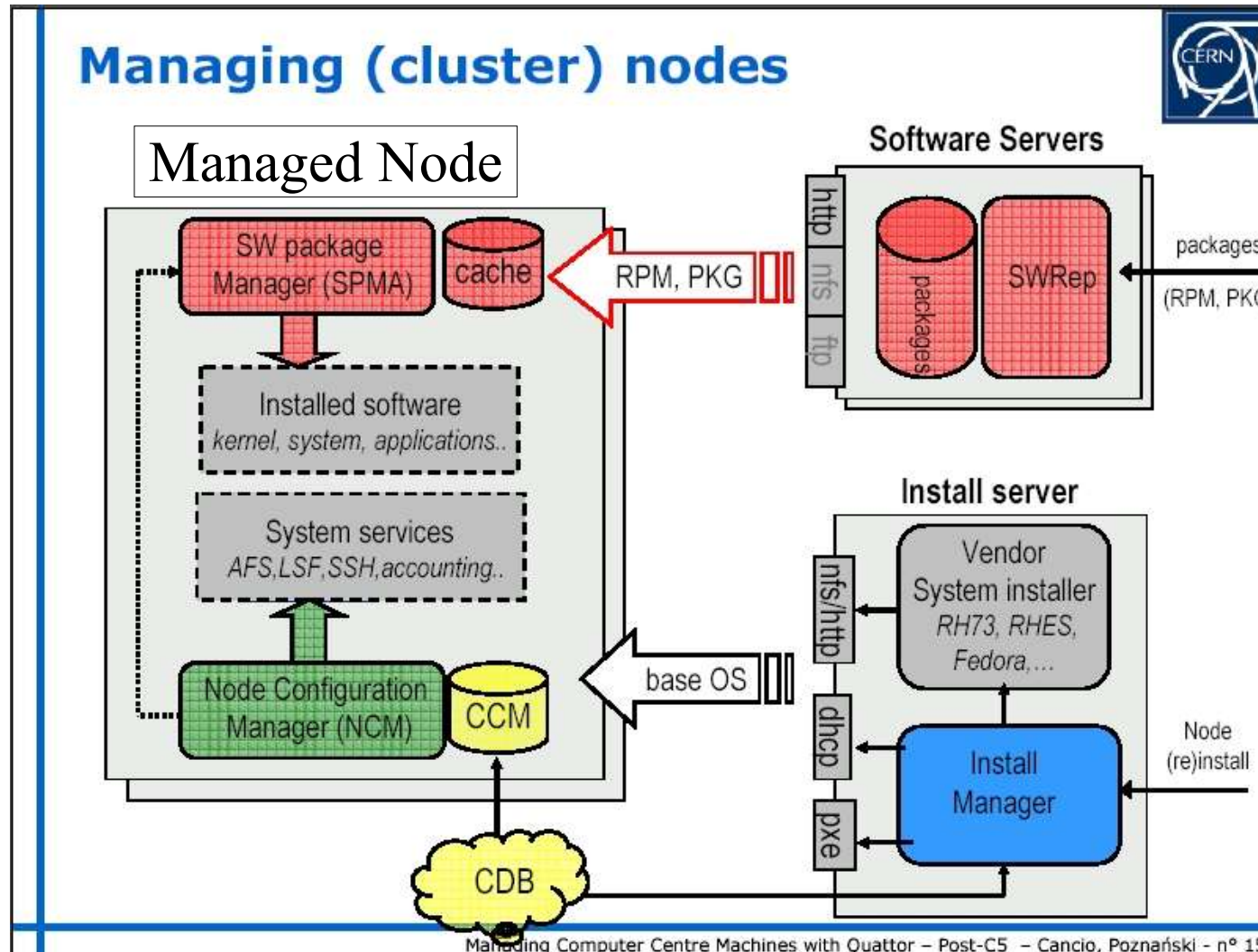
Andere Welten ..



- **Es gibt eine große Bandbreite von Tools**
 - ★ **Kommerziell: HP Openview, Tivoli (IBM), SUN N1, ..**
 - ★ **Linux Distributionen: YaST 2 + YOU, Anaconda + up2date**
 - ★ **Generische Tools erleichtern Installation und Pflege:**
 - **apt, YUM**
 - **cfengine**
 - ★ **Site-spezifische Lösungen**
 - **Beispiel DESY Hamburg: wboom/salad**
 - **Beispiel DESY Zeuthen : VAMOS/cfengine**
 - ★ **Neuere Tools in der Hochenergiephysik (bzw. im (LCG-)GRID Umfeld):**
 - **ROCKS (USA) (<http://rocks.npaci.edu>)**
 - **Quattor (Europa, CERN) (<http://www.quattor.org>)**

- Quattor ist eine Toolsuite zur Systeminstallation und -Konfiguration (<http://www.quattor.org>)
- Entwickelt im Rahmen des European Data GRID (EDG) Projektes am CERN
- Weitergeführt durch CERNs IT Department und Partner, insbesondere Universidad Autónoma de Madrid (UAM) im Rahmen des ELFms (Extremely Large Fabric management system) Projektes (<http://elfms.web.cern.ch/elfms/>) 
- Einsatz am CERN auf > 2100 Maschinen (“CPU worker nodes, tape and disk servers, infrastructure servers”)
- Einsatz an der UAM auf Desktops

Quattor: Managed Nodes



Managing Computer Centre Machines with Quattor - Post-C5 - Cancio, Poznański - n° 15

Quattor: Komponenten



- **Configuration Database (CDB)**
 - ★ Beschreibung von Maschineneigenschaften in 'PAN Templates' (s.u.)
 - ★ Speicherung in der CDB mit SQL Interface zur Abfrage
- **Software Repository (SWRep)**
 - ★ Mehrere Architekturen und Releases
 - ★ Software als 'Packages' (RPM oder PKG)
- **Automated Installation Infrastructure (AII)**
 - ★ Ist ein DHCP Server, der das Bootimage vorgibt (s.u.)
 - ★ Generiert Kickstart Profile
 - ★ Propagiert Konfigurationsänderungen bei Neuinstallationen

Quattor: Komponenten

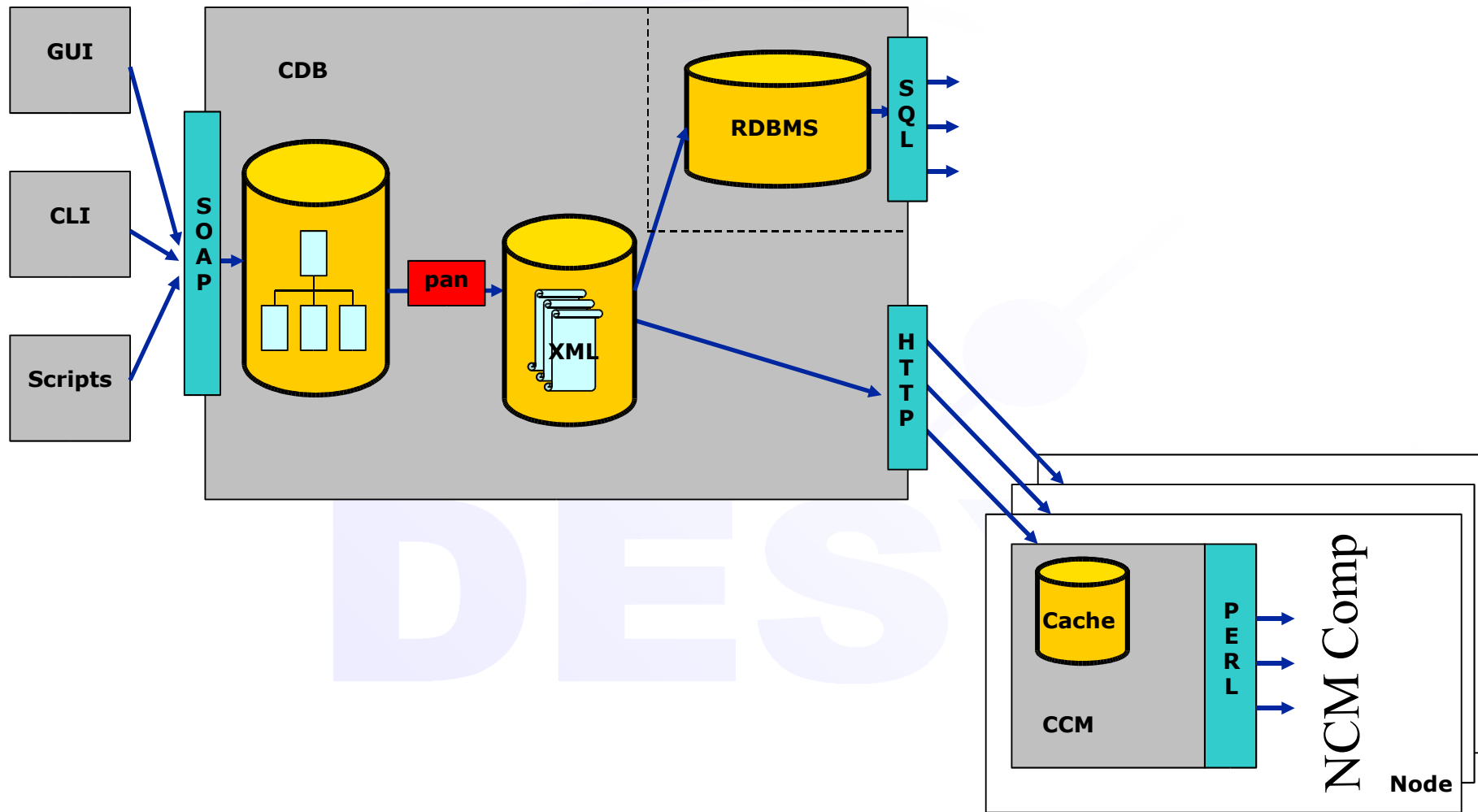
- **Der Configuration Cache Manager (CCM)**
 - ★ Synchronisiert eine lokale Kopie der Maschinenkonfiguration
 - ★ Aktiviert das NCM
- **Node Configuration Management (NCM)**
 - ★ Überführt die *vorhandene* Konfiguration der Maschine in die (in der CDB) *definierte*
 - ★ Ruft eine Reihe von 'NCM Komponenten'
- **NCM Komponenten**
 - ★ Perl Module, die Einzelaspekte der Konfigurationsdefinition umsetzen
 - ➔ Domainname, /etc/fstab, ...

Quattor: Pan Templates



- 'Pan' ist eine Sprache
 - ★ Möglichkeit der vollständigen, hierarchischen Beschreibung eines Clusters ausgehend von der Hardware (Mainboard, Platte, NIC) über Konfigurationsdetails (Partitionierung) bis zur Liste der Maschinen der Clusters.
 - ★ Die Beschreibung erlaubt Vererbung:
 - “Die Maschine ist ein DESY Standard-PC der Spezifikation 03/11.”
 - “Die Maschine erhält die Konfiguration des 'Dirac'-Clusters”
- Ein Template beschreibt ein Objekt in Pan
- Ein 'Compiler'
 - ★ validiert die Template-Hierarchie
 - ★ erzeugt eine XML Version der Maschinenbeschreibung

Quattor: Configuration Database



Quattor: Installationsprozeß



- **Registrierung/Maschinendaten**
 - ★ Erzeugung eines Templates für die Maschine
 - Maschinenspezifika wie IP Name/Adresse
 - 'Includen' allgemeiner Templates
 - ★ Eintragen des Templates in die Configuration Database
 - ★ Ein kompletter Satz von Beispieltemplates für eine Site wird mitgeliefert
- **Anstoß der Installation**
 - ★ Maschinen werden auf 'Booten via PXE' eingestellt (BIOS)
 - ★ Im AII auf 'Booten des Installationsimage' umgestellt
 - ★ Nach der Installation benachrichtigt die Maschine das AII und dieses stellt auf 'Booten des Produktionssystems' zurück

Quattor: Installationsprozeß



- **Ablauf der Installation**

- ★ Das Installationsimage bootet und startet das ..
- ★ .. Installationswerkzeug der Distribution (anaconda)
- ★ Die vom AII generierte Software Selektion umfaßt auch die Quattor RPMs
- ★ Die Postinstallation im Kickstart-File startet den CCM ..
- ★ .. der die Maschine in einen gebrauchsfähigen Zustand überführt

Quattor: Aktualisierung



- **Fallstudie: Aktualisierung von 'pdksh'**
 - ★ Das neue pdksh RPM wird in das Repository eingefügt
 - ★ Das Template mit der Liste aller RPMs im Repository wird aktualisiert
 - ★ Das Template mit der Liste der auf einem (auf den) Cluster(n) zu installierenden RPMs wird angepaßt
 - ★ Das Einfügen der geänderten Templates aktualisiert die Configuration Database
 - ★ Der Configuration Cache Manager auf allen betroffenen Maschinen erkennt die geänderte Konfiguration und veranlaßt den Software Package Management Agent (SPMA) das neue RPM zu installieren

Quattor: Status



- Eine Zusammenfassung der Nutzung innerhalb der Hochenergiephysik findet sich unter <http://www.rhic.bnl.gov/hepix/talks/041020am/leiva.pdf>
- Stand am DESY (HH)
 - ★ Die Einführung von Quattor ist erklärtes Ziel von DESY IT
 - ★ Eine Produktionseinführung steht bevor,
 - Start auf GRID-Maschinen und Compute-Clustern unter SL 3
 - ★ Die Einarbeitung ist aufwendig, wird unterstützt durch Kurse am/vom CERN
 - ★ Einschätzung: Ab O(20) gleichartigen Maschinen lohnt der Aufwand
 - ★ Vorteil: Ein integriertes System, eine abfragbare Datenbasis

Quattor: Status



- **Stand am DESY (HH)**
 - ★ **Stand: Node Configuration Management Komponenten**
 - **Standart-Installation wird durch mitgelieferte Komponenten abgedeckt ..**
 - **.. die auf dem mitgelieferten Konfigurartionsbaum basieren**
 - ★ **Nachteil: Noch kein durchgängiges Delegationsmodell**
 - **Ist von den Entwicklern als Problem erkannt und in Arbeit**

DESY