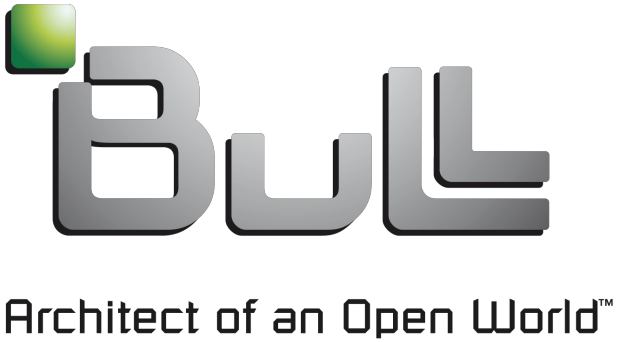


Etat des lieux des outils d'installations de cluster

Formation CNRS, Autrans 10/2009

Plan

- Introduction sur les outils de déploiement
- Les distributions standards et leurs méthodes de déploiement intégrées
- Les différents outils de déploiement
- Les distributions spécialisées
- Configuration par outils automatisés
- Perspectives

SERVI WARE
A Bull Group Company

BULL

Plan

- **Introduction sur les outils de déploiement**
- Les distributions standards et leurs méthodes de déploiement intégrées
- Les différents outils de déploiement
- Les distributions spécialisées
- Configuration par outils automatisés
- Perspectives

SERVI WARE
A Bull Group Company

BULL

Introduction sur les outils de déploiement

- Objectifs

  - Installer / réinstaller un ensemble de serveurs identiques ou non, quelque soit leur état

  - Rapidité

  - Flexibilité (en fonction du matériel)

  - Convivialité

  →L'importance donnée à chaque critères va aiguiller le choix des outils

SERVI WARE
A Bull Group Company

BULL

Introduction sur les outils de déploiement

- Moyens
  - Techniques d'installations par le réseau
  - Parallélisation des installations
  - Contrôle distant des serveurs utilisés (carte de management / KVM)
  - Interface (graphique, WEB, ligne de commande)
  - Outils spécifiques pour clusters

Introduction sur les outils de déploiement

Cluster

- Relative homogénéité du matériel
- Homogénéité de l'environnement logiciel
- Peu de changement par rapport à l'installation originale (ou réinstallation)
- Environnement « ami »

Parc informatique

- Hétérogénéité du matériel
- Hétérogénéité de l'environnement logiciel
- Évolution de l'environnement logicielle au court du temps
- Problématique de sécurité + ou - forte

SERVI WARE
A Bull Group Company

BULL

Etapes de l'installation

« Avec les mains »

- Allumer le serveur
- Insérer le DVD
- Choix dans le bios pour booter sur DVD
- Clic, clic, clic (réponses aux questions diverses)
- Reboot
- …
- Configuration « cluster »

« Via le réseau »

- Boot du serveur
- Choix bios pour boot réseau
- Fournir un mini OS via le réseau
- Installation automatique (réponses enregistrés)
- Reboot
- …
- Configuration « cluster »

SERVI WARE
A Bull Group Company

BULL

Contrôle à distance

- Objectifs

  - Remplacer les interventions humaines
    - Remplacer l'appui sur le bouton on/off de chacun des serveurs
    - Remplacer le branchement d'un clavier et d'un écran sur chacun des serveurs

  - Rendre possible l'hébergement du cluster dans un centre distant

Contrôle à distance

- Cartes de management
  - « Mini ordinateur » embarqué dans chaque serveur
  - Port Ethernet dédié ou partagé
  - Possède sa propre IP
  - Relié à la carte mère du serveur
  - Attention: différences entre chaque modèles (même pour un même constructeur)

SERVIWARE
A Bull Group Company

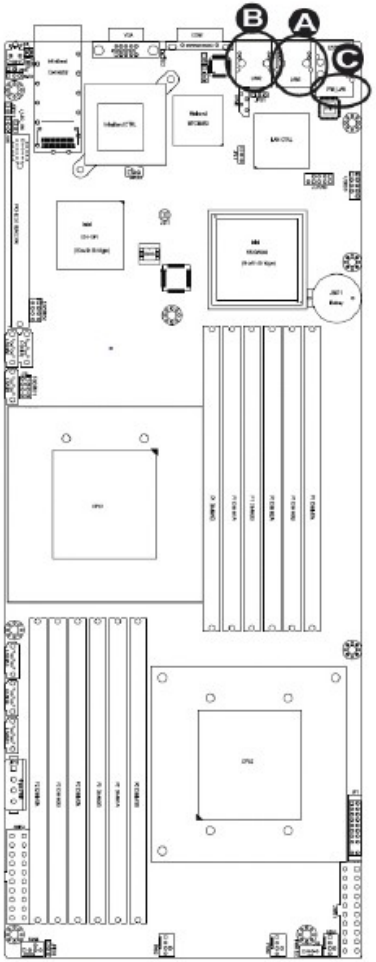
BULL

Contrôle à distance

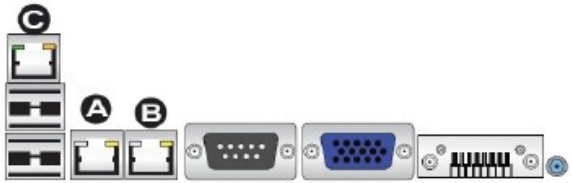
- Cartes de management
  - Fonctionnalités (suivant les modèles)
    - Interface Telnet, SSH, ...
    - Interface WEB
    - KVM (Keyboard Video Mouse) par applet JAVA
    - Authentification par login/password
    - on/off, reset
    - Clignotement d'une LED
    - Affichage des log hardware
    - ...

Contrôle à distance

- Cartes de management (exemple d'un Bull R422E2)



A: LAN 1
B: LAN 2
C: IPMI Dedicated LAN



SERVI WARE
A Bull Group Company

BULL

Contrôle à distance

- Cartes de management
  - Chaque constructeur a sa carte de management

    - Ilo (HP)
    - Ilom, alom, elom (SUN)
    - IMM / Carte RSA (IBM)
    - AMM bladecenter (IBM)
    - SIMSO+ (SuperMicro)
    - ...

    → la plupart des cartes exportent une interface IPMI

SERVI WARE
A Bull Group Company

BULL

Contrôle à distance

- IPMI
  - « Intelligent Platform Management Interface »
  - Protocole standard
  - De plus en plus répandu
  - Non spécifique à un constructeur
  - Accès de façon unique aux différentes cartes de management

SERVI WARE
A Bull Group Company

BULL

Contrôle à distance

- IPMI, fonctionnalités
  - On/off, reset
  - Accès aux valeurs de sondes (température, voltage, ...)
  - Log hardware
  - SoL (Serial Over Lan)
  - Interroger à distance (réseau Ethernet) ou depuis l'OS (module Linux)
  - ...

Contrôle à distance

- IPMI, outils permettant de dialoguer avec ce protocole

  - ipmitool
  - OpenIPMI
  - ipmiutil
  - GNU Freeipmi
  - ...

SERVI WARE
A Bull Group Company

BULL

Contrôle à distance

- IPMI, exemple

```
ipmitool -U root -P changeme -H node08-mgt chassis status
System Power         : on
Power Overload       : false
Power Interlock      : inactive
Main Power Fault     : false
Power Control Fault  : false
Power Restore Policy : unknown
Last Power Event     : command
Chassis Intrusion    : inactive
Front-Panel Lockout  : inactive
Drive Fault          : false
Cooling/Fan Fault    : false
Front Panel Control  : none
```


SERVI WARE
A Bull Group Company

BULL

Contrôle à distance

- PDU manageables
  - Prises électriques avec possibilité arrêt/marche à distance (attention: indépendance de chaque prise)
  - Switchs de ports série
    - câblage de chaque port série des serveurs sur ces switchs
    - Interfaces: WEB, telnet, SSH, ...

Séquence de boot réseau

- BIOS (Basic Input Output System)
  - Premier code exécuté par la machine
  - Possibilité de configurer la carte de management
  - Activer le boot réseau
  - Passe la main au firmware de la carte réseau

Séquence de boot réseau

- PXE (Preboot Execution Environment)
  - Méthode de boot réseau
  - Intégré au firmware de la carte réseau
  - Requête DHCP avec la MAC adresse
  - Récupération, via TFTP, d'un code à exécuter
  - Met à disposition une pile réseau utilisable par le code téléchargé
  - Protocole poussé par Intel et repris par tous les constructeurs de carte réseau(sur les machines x86_64)

Séquence de boot réseau

- DHCP (Dynamic host configuration protocol)
  - Serveur sur la frontale du cluster
  - Configuration automatique des paramètres IP
  - Adresse IP, Netmask, DNS, gateway, ...
  - Réseau local
  - Donne le nom du fichier PXE à récupérer
  - Données envoyées aux clients suivant leurs MAC adresses

Séquence de boot réseau

- DHCP (exemple d'un fichier de configuration)

```
ddns-update-style interim;
ignore client-updates;

subnet 192.168.42.0 netmask 255.255.255.0 {

        option routers                  192.168.42.1;
        option subnet-mask              255.255.255.0;
        range dynamic-bootp 192.168.42.50 192.168.42.200;
        default-lease-time 21600;
        filename "pxelinux.0";
        next-server 192.168.42.1;
        max-lease-time 43200;
}
```


SERVI WARE
A Bull Group Company

BULL

Séquence de boot réseau

- Adresse MAC (Media Access Control)
  - Identifiant physique stockée dans la carte réseau
  - Ces identifiant sont utilisés notamment dans les réseaux
    - Ethernet
    - Réseaux sans fil Bluetooth
    - Réseaux sans fil Wi-Fi
    - ...
  - 6 octets: FF:FF:FF:FF:FF:FF
  - 2^48 adresses (281474 976 710 656)
  - 24 bits par constructeur (16 777 216)

→Distinction de chaque noeud de calcul d'un cluster

SERVI WARE
A Bull Group Company

BULL

Séquence de boot réseau

- TFTP (Trivial File Transfer Protocol)
  - Transfert de fichiers
  - Protocole client/serveur
  - UDP
  - Ne gère pas le listage
  - Pas d'authentification / pas de chiffrement
  - Performances limitées

Séquence de boot réseau

- PXELINUX (syslinux)

  - Boot un noyau Linux par le réseau
  - Se chaîne à la ROM PXE des cartes réseau
  - « pxelinux.0 »
  - Cherche des fichiers via TFTP
  - Récupère le noyau Linux et l'initrd et donne la main

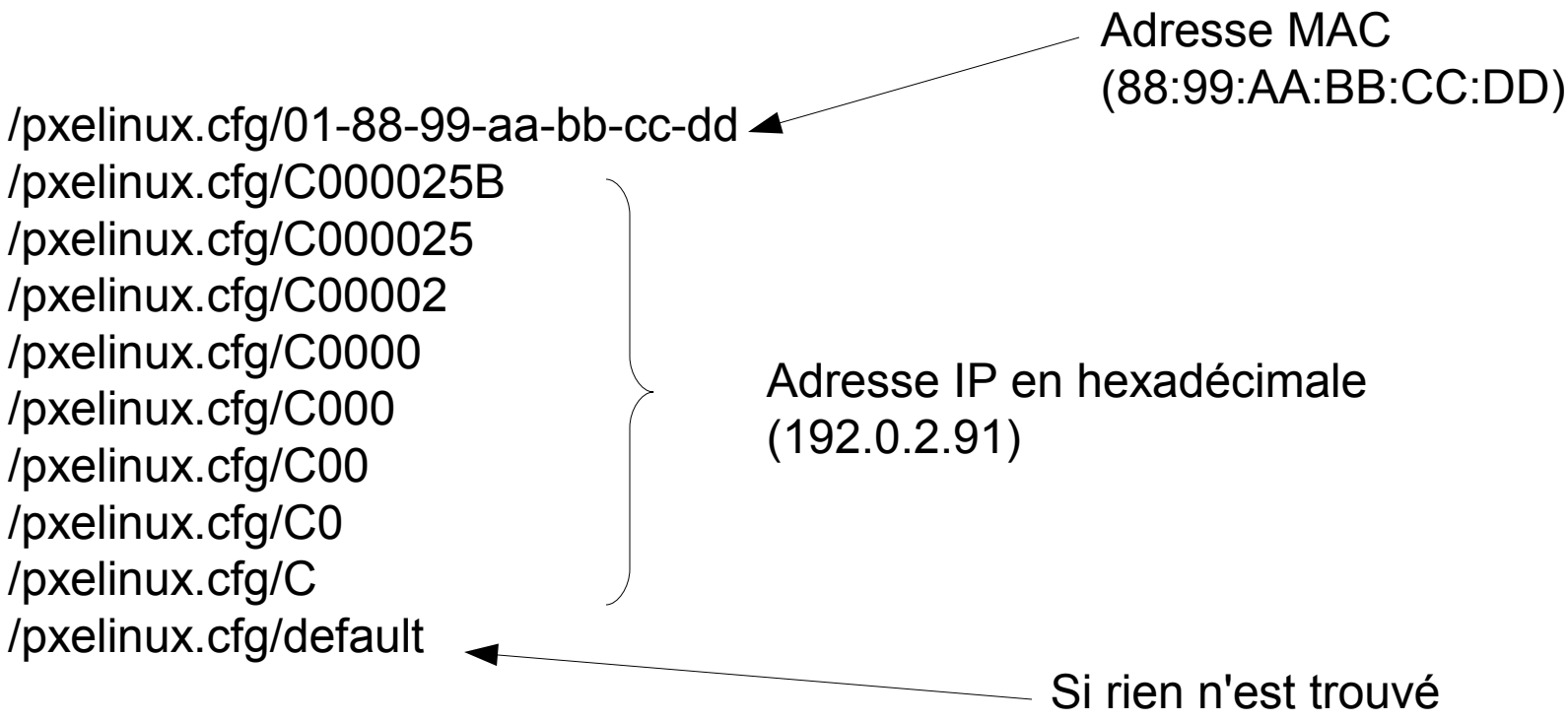
    - Initrd: image contenant un système de fichier qui sera chargé en mémoire vive. Il servira de système minimal pour la suite de l'installation.

SERVI WARE
A Bull Group Company

BULL

Séquence de boot réseau

- PXELINUX (syslinux): recherches successives



SERVIWARE
A Bull Group Company

BULL

Séquence de boot réseau

- PXELINUX (syslinux): exemple de fichier de configuration (sur le serveur TFTP)

```
DEFAULT hd

LABEL hd
   LOCALBOOT 0
```

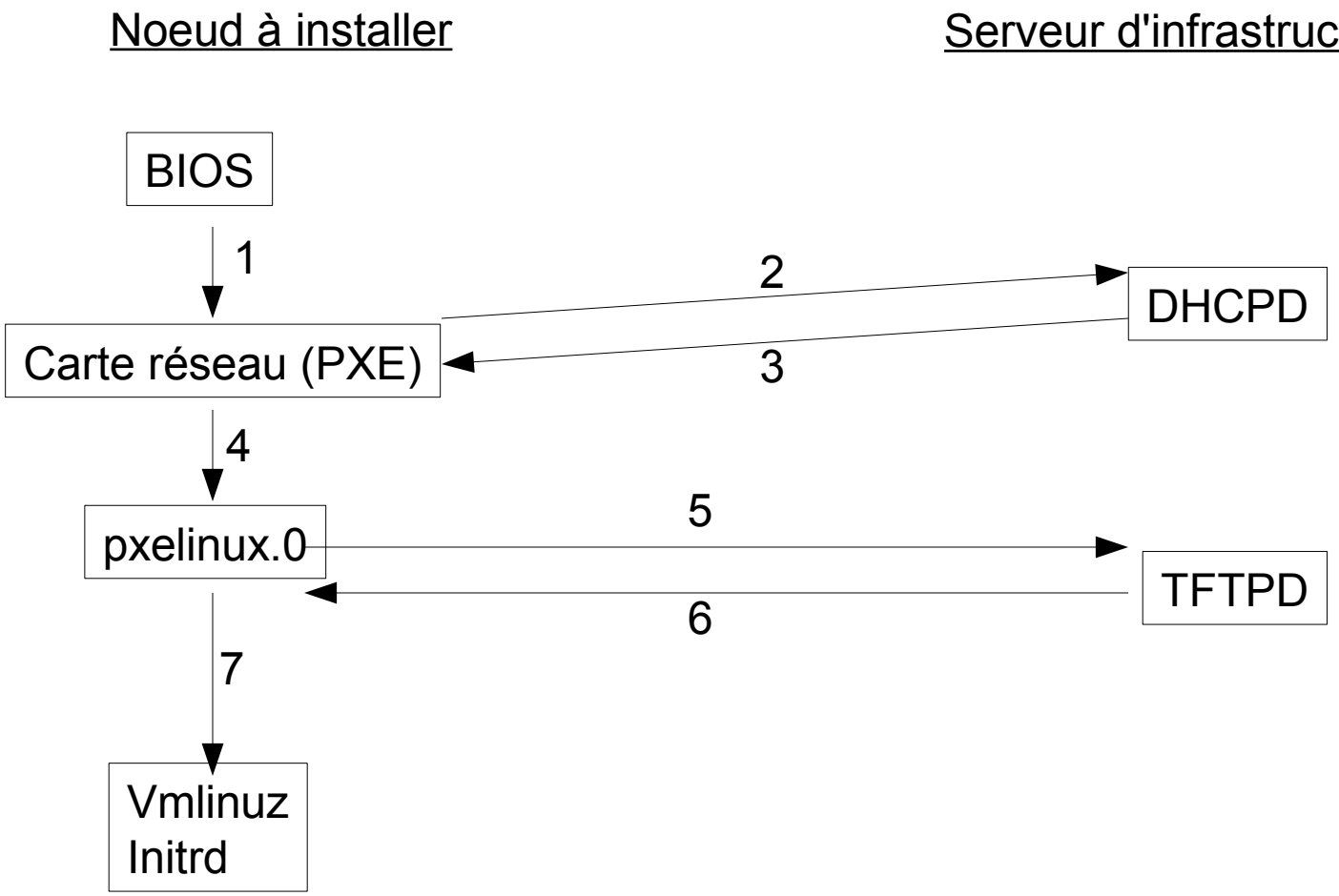
- ou

```
DEFAULT SCT
label SCT
   kernel vmlinuz
   append initrd=initrd.img
            ks=nfs:192.168.42.1:/opt/ks.cfg ksdevice=eth1
            method=nfs:192.168.42.1:/centos53/
```


SERVI WARE
A Bull Group Company

BULL

Séquence de boot réseau



SERVI WARE
A Bull Group Company

BULL

Séquence de boot réseau

- Note

  - Il existe d'autres méthodes de boot réseau
  - Dépendant du matériel
  - Par exemple: Bootp (Bootstrap Protocol)
  - Même principe: donner la main à un système de base qui réalisera l'installation
  - Pareil pour d'autres OS (BSD, Windows, ...)

Le déploiement: principes de bases

- 2 types d'installations
  - Installation automatique
  - Copie d'une image de référence

SERVI WARE
A Bull Group Company

BULL

Le déploiement: principes de bases

- Installation automatique

  - Jouer une installation classique comme avec le DVD
  - Pré-répondre à toutes les questions de manière automatique
    - Partitionnment des disques
    - Points de montages
    - Listes des paquets à installer
    - Password
    - ...
  - Mécanisme fourni par la distribution Linux

SERVI WARE
A Bull Group Company

BULL

Le déploiement: principes de bases

- Installation automatique
  - Avantages
    - Informations résumant l'installation de petite taille (un fichier texte en général)
    - Utilisation des kernel/initrd des distributions (validés/supporté par les fabriquants de matériels)
    - Reste au plus prêt d'une installation par DVD
  - Inconvénients
    - Lié fortement à un type de distribution linux
    - Problèmes pour le passage à l'échelle
      - Utilisation des protocoles NFS, HTTP, FTP
    - Plus lent qu'une installation par image (même pour un seul noeud)

Le déploiement: principes de bases

- Copie d'une image de référence

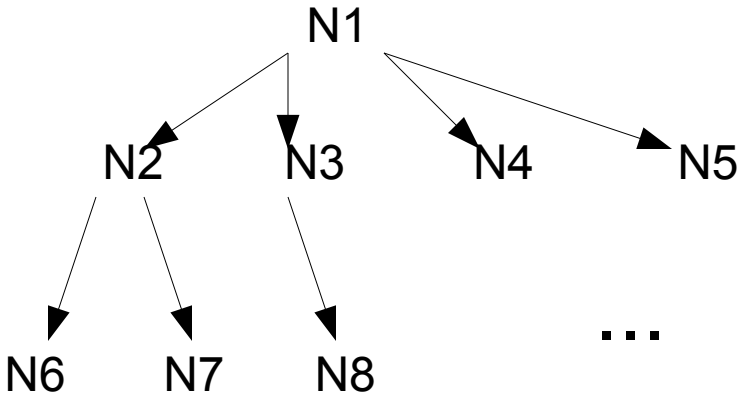
  - Réaliser une image de référence sur un noeud (manuellement avec un DVD ou par la méthode précédente)
  - Ce noeud est généralement dénommé: « Golden node »
  - En extraire une image comme un backup/snapshot (ex: tar)
  - Diffuser cette image sur les noeuds cibles
    - Réaliser un boot réseau
    - Utilisation d'un kernel/initrd spécifique qui implémente la méthode de diffusion de l'outil choisi

SERVI WARE
A Bull Group Company

BULL

Le déploiement: principes de bases

- Diffusion d'une image de référence
  - Plusieurs méthodes
    - NFS, HTTP, FTP comme pour l'installation automatique
    - Multicast
    - BitTorrent
    - Algorithmes de diffusion en arbres

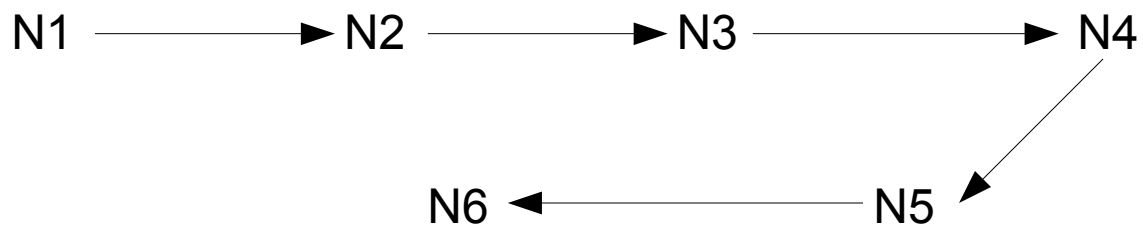


SERVIWARE
A Bull Group Company

BULL

Le déploiement: principes de bases

- Diffusion d'une image de référence
  - Plusieurs méthodes
    - Chaîne de transferts (pipe)



    - Hybride: arbre + chaîne
  - Les méthodes peuvent très être très performantes
  - Possibilité de prendre en compte la topologie réseau du cluster

Le déploiement: principes de bases

- Copie d'une image de référence

  - Avantages
    - Pas de fichier de configuration, peut être perçu comme plus simple
    - Possibilité d'algorithmes de diffusion performants
    - Non dépendant des distributions
  - Inconvénients
    - Peut être difficile à maintenir sans « versionning » précis (retour arrière)
    - Les images peuvent être volumineuses à manipuler
    - Génération de kernel/initrd spécifiques pour la diffusion (drivers)
    - Ajout d'un paquet == refaire une image (en général)

Le déploiement: principes de bases

- Conclusion personnelle

  - Installation automatique
    - Redhat / Suse / centos / fedora

  - Installation à base d'images
    - Debian / Ubuntu / Gentoo

SERVI WARE
A Bull Group Company

BULL

Post-configuration

- Objectifs
  - Transformer une installation type de serveur en un noeud de calcul « prêt à l'emploi »
  - Effectuer des opérations non possibles lors de l'installation automatique
- Comment
  - Scripts
  - Outils de configuration pour cluster
  - Configuration automatique : dhcp / dns / YP / ldapp

Post-configuration

- Exemples

  - Configurer l'adresse IP de la carte de management
  - Installer la partie noeud du batch scheduler (enregistrement auprès du serveur)
  - Installation de GPFS (enregistrement auprès du serveur)
  - Installation des drivers Infiniband

Plan

- Introduction sur les outils de déploiement
- **Les distributions standards et leurs méthodes de déploiement intégrées**
- Les différents outils de déploiement
- Les distributions spécialisées
- Configuration par outils automatisés
- Perspectives

SERVI WARE
A Bull Group Company

BULL

Les distributions standards et leurs méthodes de déploiement intégrées

- Redhat/CentOS : kickstart
- Suse : yast (xml)
- Debian : FAI
- Solaris: JumpStart
- ...

SERVI WARE
A Bull Group Company

BULL

Les distributions standards et leurs méthodes de déploiement intégrées

- Système de paquets de la distribution important
- Essentiellement 2 types

  - Ceux qui posent des questions (ex: Debian, Ubuntu, ...)
    - Permet d'avoir les fichiers de configuration des applications prêt à l'emploi à la fin de l'installation
    - Par contre il faut « pré-répondre » à toutes les questions
  - Ceux qui ne posent pas de question (ex: RedHat)
    - Pas besoin de préparer des réponses à l'avance
    - Nécessaire de copier des fichiers de configuration en post-configuration

RedHat: exemple d'installation par kickstart

- Après une installation « manuelle », le fichier kickstart correspondant est accessible dans

  /root/anaconda-ks.cfg

- Fichier à plat
- Possibilité d'inclure d'autres fichiers kickstart
- Section «%pre »: exécution de commandes sh avant l'installation
- Section «%post »: exécution de commandes sh après l'installation

RedHat: exemple d'installation par kickstart

- Exemple d'un fichier PXE récupéré lors du boot réseau d'un noeud

```
DEFAULT SCT
label SCT
    kernel vmlinuz
    append initrd=initrd.img
                ks=nfs:192.168.42.1:/opt/ks.cfg ksdevice=eth1
                method=nfs:192.168.42.1:/centos53/
```

- Arguments du kernel importants

SERVI WARE
A Bull Group Company

BULL

RedHat: exemple d'installation par kickstart

- Exemple d'un fichier kickstart

```
install
skipx
key --skip
lang C
keyboard us
rootpw --iscrypted $1$DYvF4kJf$C/icjdtJ/lrGIIT7bEwrN.
firewall --disabled
selinux --disabled
authconfig --enableshadow --enablemd5
timezone --utc Europe/Paris
bootloader --location=mbr --driveorder=sda,sdb,sdc --append="rhgb
quiet"
clearpart --all
zerombr yes
part /boot --fstype ext3 --size=100 --ondisk=sda
part /var --fstype ext3 --size=20480 --ondisk=sda
part swap --fstype swap --size=8192 --ondisk=sda
part / --fstype ext3 --size=40960 --ondisk=sda
part /scratch --fstype ext3 --size=100 --grow --ondisk=sda
reboot
```


SERVI WARE
A Bull Group Company

BULL

RedHat: exemple d'installation par kickstart

```
packages
@development-libs
@editors
@system-tools
@core
...
system-config-bind
system-switch-mail-gnome
system-config-boot
memtest86+
-qlvnictools
-ibvexdmtools
-srptools
-tvflash
-libibverbs
-openib
-sendmail-cf
-sendmail
...
```


SERVI WARE
A Bull Group Company

BULL

RedHat: exemple d'installation par kickstart

- La technique du kickstart est facilement intégrable dans des outils de plus haut niveau

SERVI WARE
A Bull Group Company

BULL

Debian: FAI

- Outil à base de « classes »
- C'est + que la méthode basique d'installation d'une Debian
- Utilise le mécanisme de paquets Debian (apt) avec un mirroir local

SERVI WARE
A Bull Group Company

BULL

Debian: FAI

- Arborescence des fichiers de configuration sur le serveur

  - class : contient la liste des classes FAI pour chaque machine
  - disk_config : partitionnement du disque en fonction de la classe
  - debconf : réponses aux questions des paquets debian
  - files : fichiers spécifiques à copier à la fin de l'installation
  - scripts : scripts exécutés à la fin de l'installation
  - package_config : liste des paquets debian à installer

Plan

- Introduction sur les outils de déploiement
- Les distributions standards et leurs méthodes de déploiement intégrées
- **Les différents outils de déploiement**
- Les distributions spécialisées
- Configuration par outils automatisés
- Perspectives

SERVI WARE
A Bull Group Company

BULL

Les différents outils de déploiement

- IBM : CSM et XCAT
- SUN : N1SM, SUN xVM
- HP : CMU
- Bull : Ksis
- SERVIWARE : SCT
- SCALI : Scali manage
- System imager
- Kadeploy
- ...

SERVI WARE
A Bull Group Company

BULL

xCAT : présentation

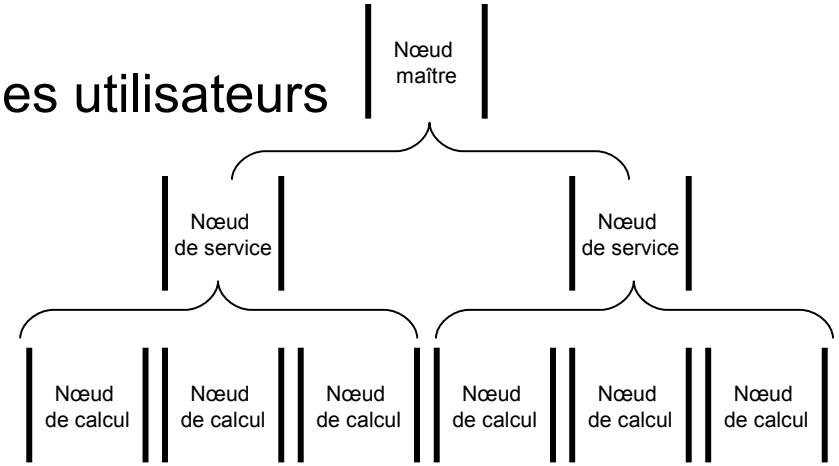
- **eXtreme Cluster/Cloud Administration Toolkit**
- Outil de déploiement et de gestion open-source
- Flexible, indépendant du matériel
- Développé (et supporté) par IBM
- Capable de déployer des milliers de nœuds
- Utilisé pour la mise en service de *Roadrunner*, Los Alamos (#1 Top500)

xCAT : fonctionnalités

- Interface unifiée pour le contrôle matériel (IPMI) : contrôle de l'alimentation, contrôle du boot, lecture des sondes matérielles, SoL, event logs…
- Matériel : toute machine IPMI
- OS : Linux (distrib RPM), AIX, Windows
- VM : Xen, KVM, VirtualBox (création, contrôle, migration)
- Installation des nœuds :
  - *stateful* (diskeless avec iSCSI, boot SAN ou diskful)
  - *stateless (*RAM-root, compressed RAM-root, NFS-root w/ RAM overlay)
  - clonage (imaging)
- Découverte automatique des nœuds (interrogation des switches par SNMP)
- Consoles et logs centralisés
- Monitoring (SNMP, Ganglia, RMC)
- Inventaire software/firmware (possibilité de mise à jour distante)

xCAT : architecture

- Architecture client/server
- Administration basée sur des rôles utilisateurs
- Gestion hiérarchique,
  pour le passage à l'échelle



- Choix du backend DB (SQLite, PostgreSQL, MySQL)
- Architecture modulaire (plugins)
- Description du cluster dans des tables : plan d'adressage et de nommage, caractéristiques des nœuds, groupes…

SERVI WARE
A Bull Group Company

BULL

xCAT : bilan

- Les plus
  - Nombreuses fonctionnalités, interface unifiée et centralisée
  - Description compacte des caractéristiques du cluster
  - Découverte automatique des nœuds
  - Auto-configuration des BMC
  - Possibilité de personnalisation (pre-/postscripts)
- Les moins
  - Documentation parfois incomplète ou obsolète
  - Quelques fonctionnalités manquantes par rapport aux versions 1.x
  - En cours de développement

HP: CMU

- Cluster Management Utility
- Outil propriétaire
- Installation à base d'image (golden node)
- Cartes de management ILO
- Ensemble d'outils (comme pdsh)
- Intègre une interface de monitoring
- Interface conviviale
- Support des machines HP

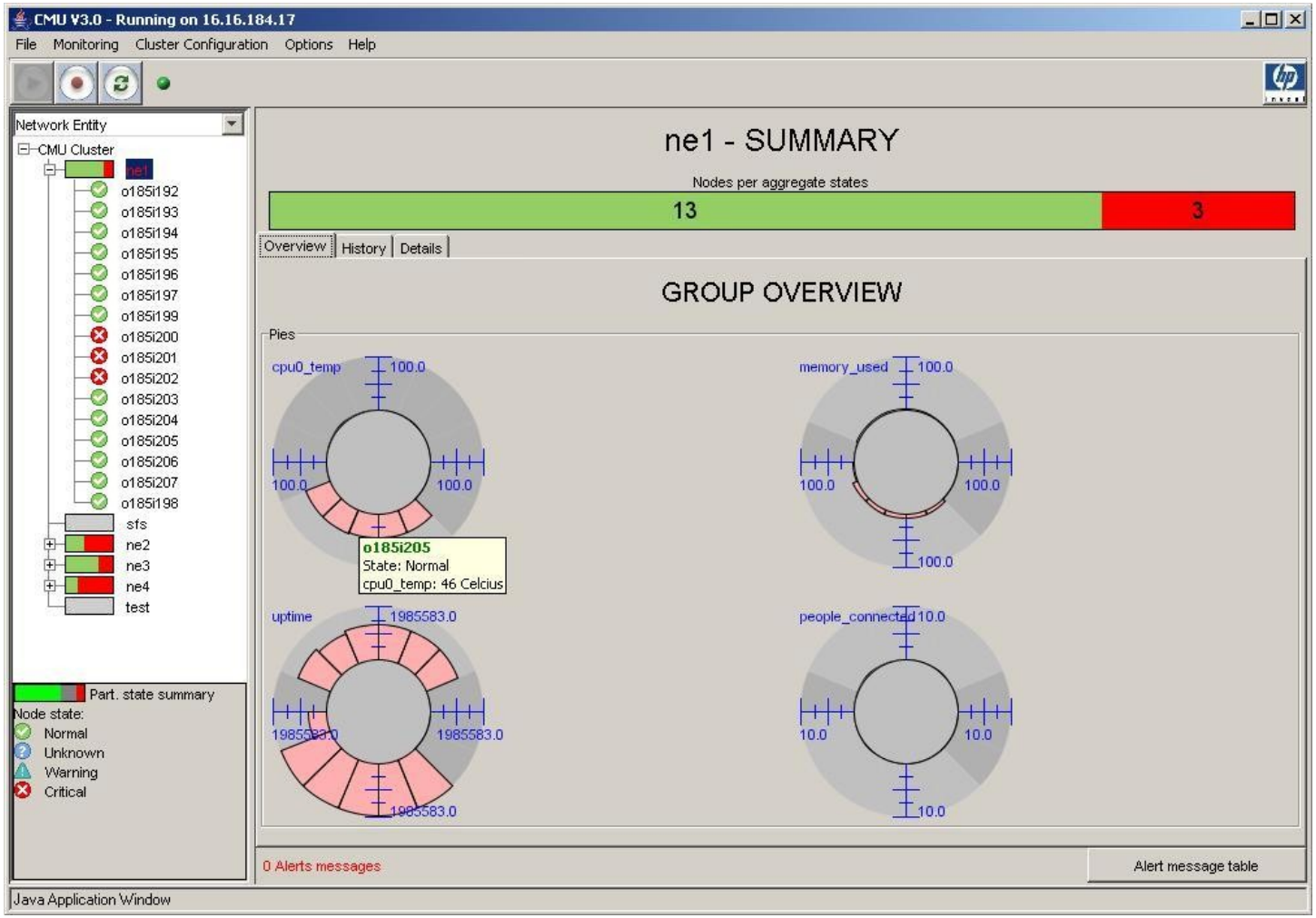
HP: CMU

- Distributions supportées
  - RedHat
  - Suse
  - Debian
- Ne supporte pas le RAID soft ni le LVM sur les noeuds de calcul
- Possibilité de diskless depuis peu de temps

SERVIWARE
A Bull Group Company

BULL

HP: CMU



SERVIWARE
A Bull Group Company

BULL

SystemImager

- Outil opensource
- Diffusion d'image (golden node)
- Fourni un kernel/initrd de base (i386)
- Outils de génération de kernel/initrd pour le boot réseau (UYOK: Use Your Own Kernel)
- Possibilité d'utiliser comme transports
  - BitTorrent
  - Multicast
  - SSH

BULL XBAS: ksis

- Diffusion par images
- Basé sur SystemImager pour certains éléments
- Système de « versionning » sur les images
- Possibilité d'envoyer des différences d'images (seulement)
- Mise au point du noeud référence via QEMU

- Découverte automatique des noeuds de calcul
- Découverte de la topologie réseau et/ou renseignement manuel

  → BD

SERVI WARE
A Bull Group Company

BULL

XBAS: ksis

- Diffusion avec passage à l'échelle
  - Utilisation des informations sur la topologie
  - Technique de chaînes + hiérarchies
- Fonctionnalités implémentées à cours terme:
  - support du déploiement de machines virtuelles
  - Boot des machines en diskless

KaDeploy

- Opensource
- Outil développé dans le cadre de Grid5000
- Diffusion d'image
- Base de données permettant le versionning des images
- Approche utilisateur (un utilisateur peut déployer son environnement)
- Algorithme de diffusion: chaîne (rapide)

Serviware: SCT

- Ensemble de scripts développés par Serviware
- À base de kickstart
- Très forte souplesse d'utilisation/configuration
- Permet la synchronisation automatique d'un ensemble de fichiers de configuration
- Pas d'interface graphique

SUN xVM

- Successeur de N1SM
- À base de kickstart
- Interface WEB conviviale
- Outil assez récent

Plan

- Introduction sur les outils de déploiement
- Les distributions standards et leurs méthodes de déploiement intégrées
- Les différents outils de déploiement
- **Les distributions spécialisées**
- Configuration par outils automatisés
- Perspectives

SERVI WARE
A Bull Group Company

BULL

Les distributions spécialisées HPC

- Xbas
- Rocks
- Oscar
- ...

SERVI WARE
A Bull Group Company

BULL

XBAS

- Distribution Linux basée sur RedHat
- Intégration au plus proche des machines Bull
- Automatisation de l'installation d'un cluster
- Outils intégrés/configurés:
  - Management
  - Monitoring
  - Job scheduler
  - MPI
  - Suite de développement
  - ...

XBAS

- Base de données: stockage centralisé de toutes les informations de la plateforme

  → informations utilisée par tous les outils de XBAS

- Uniquement sur du matériel Bull (pour l'instant)
- Environnement validé pour un usage HPC

SERVI WARE
A Bull Group Company

BULL

Rocks

- Distribution basée sur RedHat
- Base de données
- Fichiers de configuration en XML
- Directif (couche d'abstraction cluster)
- À base de ROLL (comme des modules applicatifs)
- Outil de diffusion à base de kickstart
- Peu manquer de souplesse dans certaines conditions

Oscar

- Open Source Cluster Application Resources
- Ensemble de paquets
- Repository pour plusieurs distributions
  - RPM
    - RedHat
    - CentOS
    - Opensuse
    - Suse Enterprise
  - DEB
    - Debian
- Utilise SystemImager pour le déploiement

Oscar

- Paquets
  - Torque/MAui
  - SGE
  - SystemImager
  - Openmpi
  - Mpich
  - Heartbeat
  - Ganglia
  - Lam
  - ...

SERVI WARE
A Bull Group Company

BULL

Outils des constructeurs et Dépendance au matériel

- Pourquoi et quoi?
  - Problème de compatibilité des cartes de management
  - Problème de driver dans le « premier noyau »
  - Support de versions très précises des distributions
  - Pas forcement fait volontairement
- Que faire?
  - Préférer les outils génériques (ils peuvent être fournit par certains constructeurs)
  - Changer d'outils au fur et à mesure de l'évolution du cluster : pénible

Installation atypiques (pour l'instant)

- Boot on LAN (NFSROOT) / Boot on SAN

  - Avantages
    - Pas besoin d'installer les noeuds
    - Pas besoin de disques durs sur les noeuds
  - Inconvénients
    - Passage à l'échelle
    - 1 export NFS par noeud ou obligation de gérer les fichiers «changeants » en RAM (ou syslog redirigé sur le réseau)

SERVI WARE
A Bull Group Company

BULL

Installation atypiques (pour l'instant)

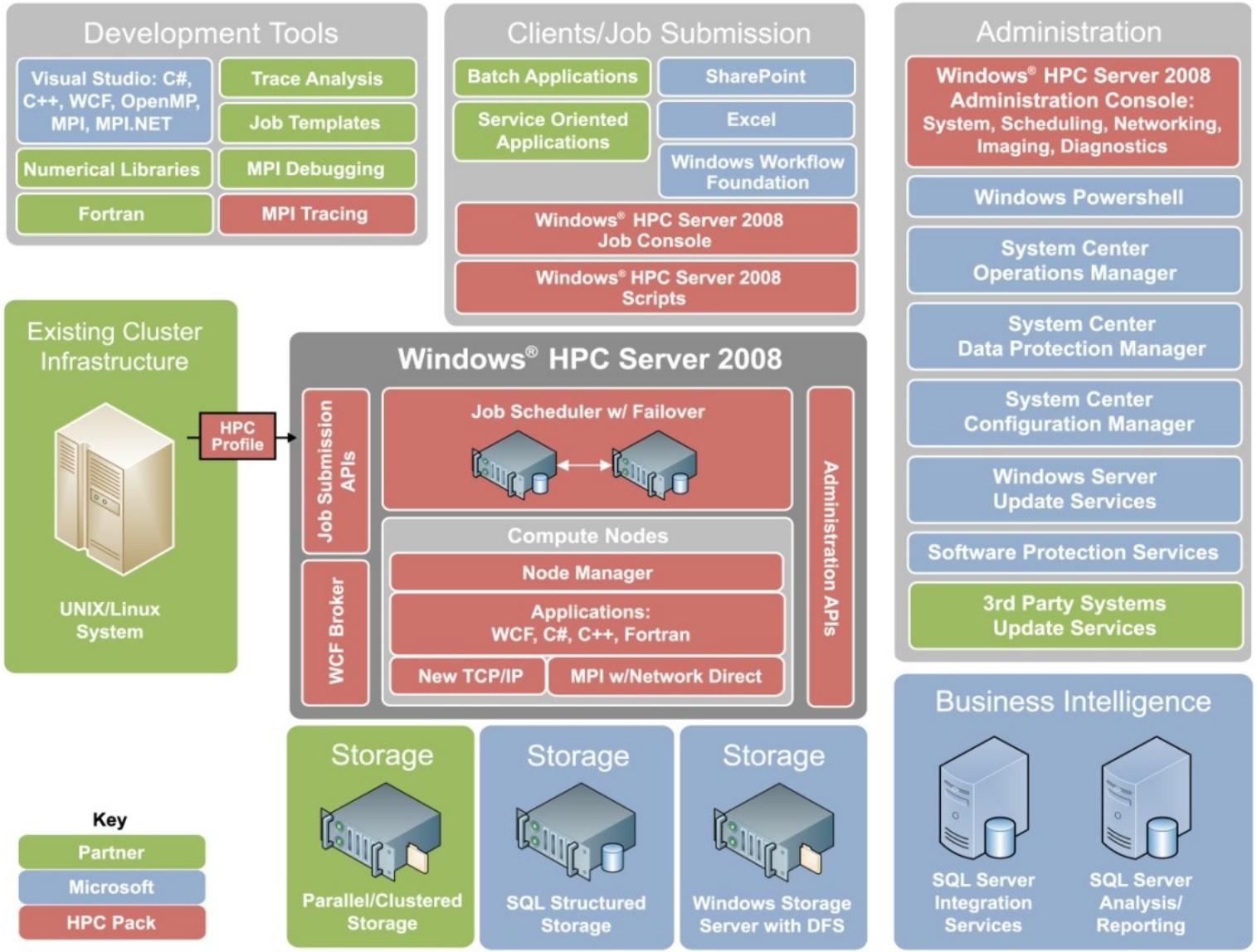
- Diskless « agressif »
  - Quantités de mémoire de + en + grandes sur chaque noeud
  - Transfert de tout l'OS à chaque boot de la machine
  - Tout dans un tmpfs
  - Avantages
    - Pas besoin d'installer les noeuds
    - Pas besoin d'avoir de disques durs sur les noeuds
    - Passage à l'échelle possible
  - Inconvénients
    - Mise au point de l'image
    - Prend un peu de place

Cluster Windows ?

- Windows HPC server 2008
- Intégration
- Support Infiniband
- Diffusion par images
  - Windows Deployment Service: WDS
  - Support du multicast
- powershell

- possible manque de souplesse

Cluster Windows ?



Plan

- Introduction sur les outils de déploiement
- Les distributions standards et leurs méthodes de déploiement intégrées
- Les différents outils de déploiement
- Les distributions spécialisées
- **Configuration par outils automatisés**
- Perspectives

SERVI WARE
A Bull Group Company

BULL

Configuration par outils automatisée

- Une seule image sans personnalisation
- Outils qui le permettent:
  - Dhcp
  - Yp
  - LDAP
  - Cfengine
  - Automount
  - NFS
  - ...

SERVI WARE
A Bull Group Company

BULL

Configuration par outils automatisée

- Ce qu'il est possible de faire
  - Réseau
  - User/ Group
  - Montage NFS → installation d'applications
  - Fichiers de configuration
- Les limites
  - Installation / configuration de drivers spécifiques
  - Scripts de custo évolués
- Goulot d'étranglement (passage à l'échelle de ces différents services)
  - Ex: Openldap

→Trouver le bon compromis

Les mises à jours après déploiement

- Problématique
  - Des outils rendent les changement de configuration sur les noeuds de calcul très simple
    - Kash (Taktuk)
    - Dsh
    - Gexec
    - ...
  - Yum / apt-get: installation de logiciels facilement
  - Que se passe-il en cas de panne d'un noeud et de réinstallation???
  - Idée: un outil pour gérer ce genre de modifications

Cfengine

- **Historique:** logiciel libre créé en 1993 par Mark Burgess, professeur à l'Université d'Oslo. Cfengine en est aujourd'hui à la version 3.
- **GNU Public License** (GPL v3) et **Commercial Open Source License** (COSL)
- **Cfengine** 3 existe en 4 versions
  - **Community Edition**: version de base libre et gratuite
  - **Nova**: version de base commercialisée
  - **Constellation**: version améliorée commercialisée pour grandes entreprises (disponible en 2010)
  - **Galaxy**: version commercialisée pour très grandes entreprises (disponible en 2011)

Cfengine

- **Concept de base**
  - administration de parc hétérogène automatisée.
  - politique établie sous formes de règles
  - permet de centraliser un comportement d'assez haut niveau plutôt que d'avoir à définir les tâches en détail pour tous les cas possibles de machines.
  - puissant outil, écrit en C
  - Agents sur les noeuds qui communiquent avec un serveur

Cfengine

Principe opérationnel:

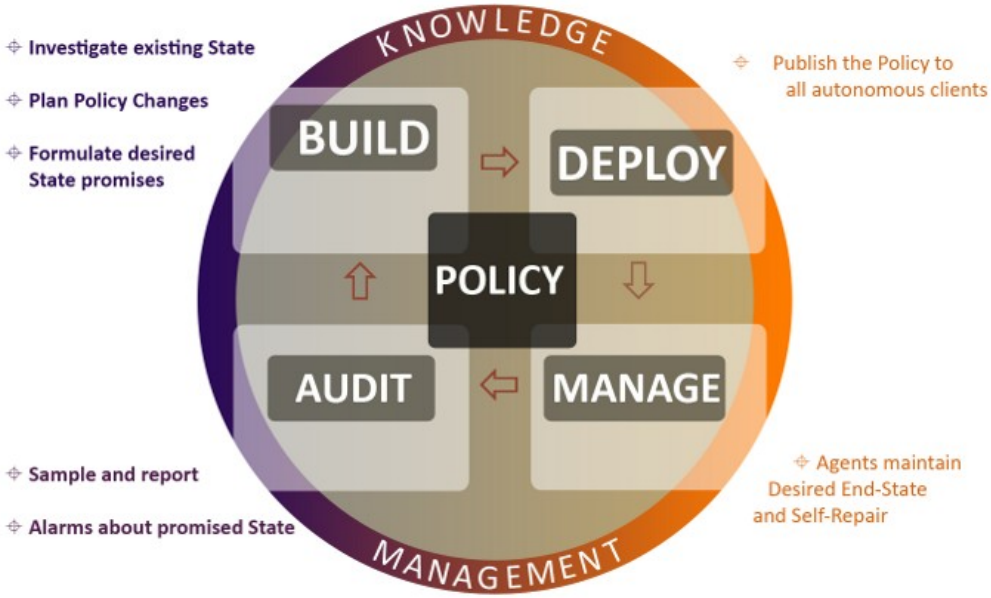
- créer un ensemble de fichiers de configuration qui vont décrire une procédure d'installation des machines du réseau
- les configurations sont modifiées sur le serveur (fichiers **cfagent.conf** et **update.conf**), puis envoyés et exécutés vers les clients
- on peut déclarer des classes (par exemple des groupes de machines)

Principales fonctionnalités

- Gestion et configuration des interfaces réseaux
- Gestion de fichiers
- Gestion des droits et des permissions
- Gestion de nettoyage de fichiers obsolètes
- Automatisation des montages de systèmes de fichiers
- Gestion de contrôle des exécutions de scripts et commandes shell
- Gestions d'intégrité md5 des fichiers
- Gestion des processus démons
- Gestion de configuration des nouvelles installations selon une politique donnée
- Gestion de restauration des configurations systèmes modifiées accidentellement ou volontairement

http://www.cfengine.org/

- Linux
- Solaris
- Windows
- *BSD (OpenBSD, NetBSD, FreeBSD)
- AIX
- HP-UX
- Mac OS (Darwin)

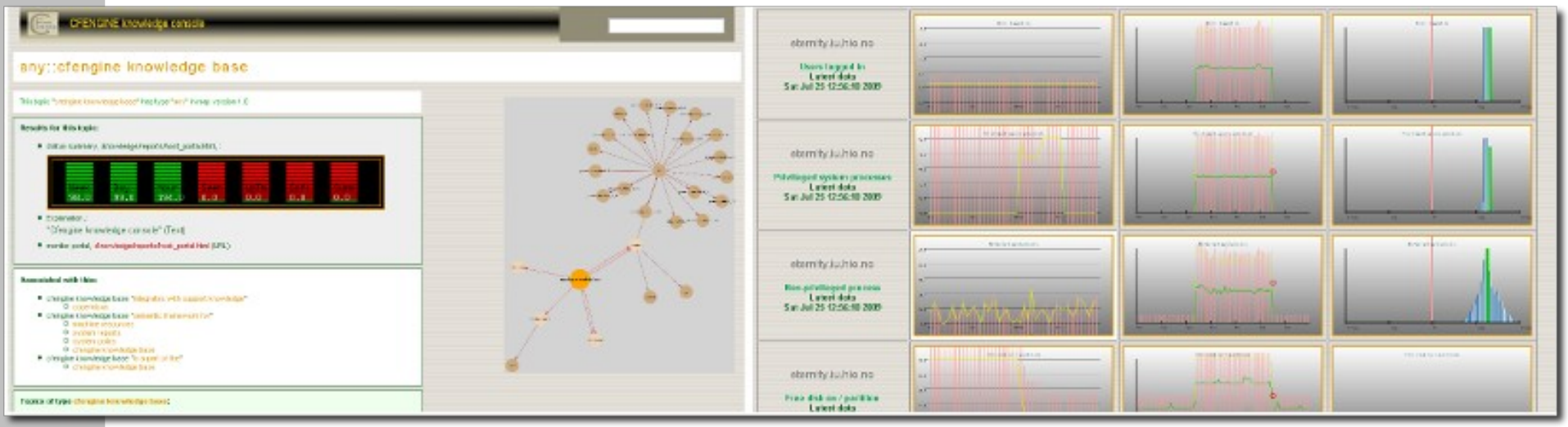


ServiWare
A Bull Group Company

BULL

Cfengine

- Exemple d'interfaces graphiques possibles



SERVIWARE
A Bull Group Company

BULL

Cfengine

- Il en existe d'autres
  - Arusha Project (ARK)
  - Bcfg2
  - Chef
  - DACS
  - ISconf
  - LCFG
  - OCS Inventory NG with GLPI
  - opsi (open pc server integration)
  - PCfengine
  - PIKT
  - Puppet
  - Quattor
  - Radmind
  - SmartFrog
  - STAF

http://en.wikipedia.org/wiki/Comparison_of_open_source_configuration_management_software

Plan

- Introduction sur les outils de déploiement
- Les distributions standards et leurs méthodes de déploiement intégrées
- Les différents outils de déploiement
- Les distributions spécialisées
- Configuration par outils automatisés
- **Perspectives**

SERVI WARE
A Bull Group Company

BULL

Perspectives

- Projet gPXE (Etherboot)
  - Opensource network bootloader
  - Pour remplacer la ROM propriétaire PXE des cartes réseau
  - Ou en boot chaîné après un boot PXE
  - Ou sur disquette, clé USB, CD
  - Fonctionnalités
    - Boot on HTTPD
    - Boot on SAN: ISCSI, AoE
  - Meilleur passage à l'échelle que TFTP

SERVI WARE
A Bull Group Company

BULL

Avenir : Virtualisation?

- QEMU, KVM, XEN, VMWARE, VirtualBox
  - Avantages
    - Plus besoin du boot réseau (rapidité)
    - Abstraction du matériel == images identiques sur plusieurs clusters
    - Performances HPC bonnes
  - inconvéniant
    - Performances IO (disk, reseau, etc...) mauvaises

Avenir : Virtualisation?

- Pourquoi ne pas faire une machine virtuelle pour HPC?
  - DOM 0 sans surcoût pour les IO
  - Une sorte de « superBIOS » (plus intelligent)
  - Permettrait d'avoir 1 OS par utilisateur ou par job...