

Grisbi 29

SIDUS, une autre approche de la
machine "à discrétion"

Emmanuel Quémener

Ce que SIDUS n'est pas...

Premièrement : SIDUS n'est pas SIDIOUS !

Darth **SIDIOUS**
alias Palpatine



SIDUS : en latin
constellation



SIDUS : Single Instance Distributing Operating System

Différence entre Sidious & SIDUS : IO (Input/Output)

Ce que SIDUS n'est pas...

Mais ce que SIDUS partage avec eux !

Ce que SIDUS n'est PAS !

- LTSP : Linux Terminal Server Project
 - Un serveur avec la “charge”, administration simplifiée du client
- FAI, Kickstart, Debian Installer Preseed :
 - « Lorsque la machine remplace l'humain dans la procédure d'installation... »
- LiveCD par réseau :
 - Une image ISO distribuée par le réseau

Mais ce que SIDUS partage avec eux

Boot PXE, TFTP, NFSroot, (AUFS)

Ce dont SIDUS se rapproche...

Ce que c'est finalement...

- Faddej
- DRBL

Et concrètement, c'est quoi ?

- Ce n'est pas un logiciel (pas de paquet à installer, tout existe)
 - Pas de paquet à installer !
 - Juste une documentation à suivre
- C'est une approche : partager simplement & efficacement

Les deux principales propriétés de SIDUS

Reproductibilité dans l'espace-temps

- Unicité de configuration
 - Deux clients SIDUS : le même OS au bit près !
 - Exploitation des ressources locales
 - Processeurs & RAM (& extra...) exploités : ceux des clients !
- Reproductibilité ? Pour un SIDUS inchangé
 - Stabilité dans le temps (pour un client défini)
 - 2 démarrages consécutifs sur 1 machine : même système
 - Stabilité dans l'espace (pour deux ou plus clients différents)
 - 2 clients démarrant simultanément : même système
 - 2 machines démarrant ne peuvent pas ne pas avoir un OS identique

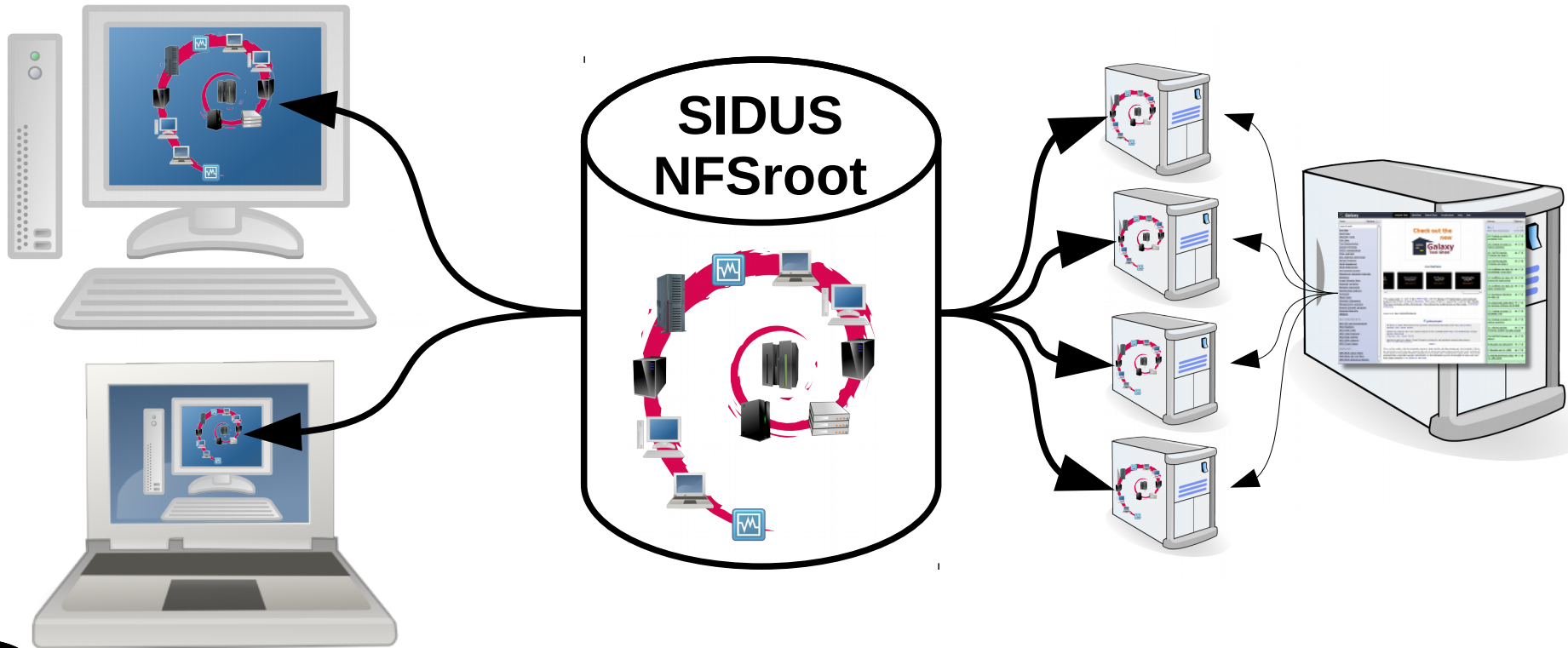
Quelques cas d'usage typiques COMOD & ANOD

COMOD

Compute On My Own Device

ANOD

Another Node On Demand



SIDUS en 7 Questions : CQQCOQP (Ben) Pourquoi ?

- Pour uniformiser de facto tous les clients
- Pour limiter les d'administration à un poste unique
- Pour comparer des matériels différents à une même base
- Pour récupérer des “fluides” (Watts & BTU)
- Pour rationaliser l'usage des stations de travail
- Pour investiguer des ressources de stockage sous anesthésie
- Pour s'assurer de la reproductibilité de l'OS & ses applications

SIDUS en 7 Questions : CQQCOQP

Pour Quoi ? Pour Qui ?

Pour Quoi ?

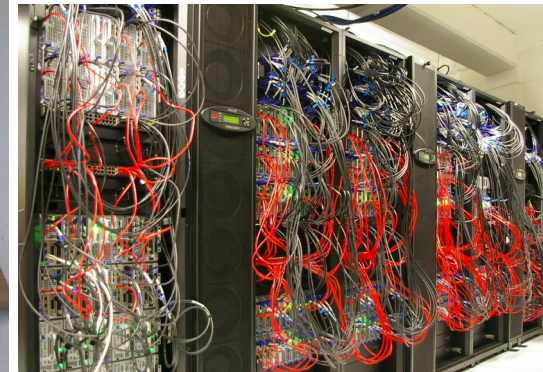
- Nœuds de cluster en HPC
- Postes de travail
- Stations graphiques
- Paillasse numériques
- COMOD

Pour Qui ?

- Chercheur en informatique
- Ingénieur en informatique
- Administrateur de salle info
- Formateur outils compliqués
- RSSI

SIDUS en 7 questions : CQQCOQP Où & Quand ?

- Centre Blaise Pascal, ENS-Lyon : salle informatique
 - 12 Neoware en 2010Q1, 24 stations 2016Q2
- Centre Blaise Pascal, ENS-Lyon : cluster
 - 24 nœuds in 2010Q1, 88 nœuds en 2016Q2
- Centre de calcul PSMN, ENS-Lyon
 - 100 nœuds 2012Q2, 500 nœuds 2016Q2
 - Tout Equip@Meso
- Laboratoires, ENS-Lyon
 - Chimie, IGFL, LBMC, UMPA, RDP
- École de physique des Houches
 - Editions de 2011 à 2016
- Institut Sciences Analytiques
 - Démonstrateur & formation 2016



SIDUS en 7 questions : CQQCOQP (Dit) Comment (ça marche) ?

AUFS : Another Union File System

- Agrégation de systèmes de fichiers : ruse de LiveCD
- 4 étapes :
 - Monter en NFSroot en lecture seule l'OS sur un dossier (point de montage)
 - Créer un système de fichiers temporaire TMPFS sur un second dossier
 - Lier les deux précédents dossier avec AUFS
 - Offrir le dossier résultant comme racine de l'OS
- Comportement d'un système de fichier en Lecture/Écriture
- Au redémarrage toute modification disparaît
- Un prérequis : chroot pour l'installation initiale (& administration)

SIDUS en 7 questions : CQQCOQP, la fin !

Comment installer : SIDUS en 7 étapes

- Création par Debootstrap d'une nouvelle racine exportée par NFS
- Création d'un “cordon ombilical” avec l'hôte
 - Montage des dossiers /proc /sys /dev/shm
- Installation (& purge de certains paquets non sollicités)
- Adaptation à l'environnement local
 - Fuseau horaire, clavier, localisation, serveur de fichiers utilisateurs, authentification.
- Création d'une séquence de démarrage avec AUFS
 - Modification du aufs-mount.sh cassé de Dracut par le SIDUS
 - Changement du bail DHCP passé à « forever » dans le script dclient-script.sh
 - Quelques ruses pour éviter des problèmes au démarrage (comme autofs avec dbus)
- Importation du noyau & du initrd spécifique sur le serveur TFTP
- Décrochage du “cordon ombilical” avec l'hôte

SIDUS en 7 questions : la fin !

Comment administrer ?

- Une limitation : le /proc doit être unique..
 - Grande vigilance sur les processus
 - Manipulation of Java, compilation with optimization, installation
- La truand :
 - Passage par le chroot, les opérations classiques marchent à 90%
- La brute :
 - Passage par le chroot, réglage du “cordon ombilical”
 - Opérations classiques, Classical operations, unset the cord
- La bonne :
 - Démarrage d'une machine avec accès en Lecture/Ecriture du NFSroot & administration

SIDUS en 7 questions : la tout fin !

Combien ça coûte ?

- Un réseau idéal : Gigabit Ethernet au client, 10G au serveur
 - Mais ça tourne correctement sur un réseau 100 Mb/s !
- Un serveur idéal : 4 CPU, 16 GB RAM, 10G, SSD
 - Mais cela fonctionnait pour 330 nœuds sur un v(eau)40z au PSMN !
- Un client idéal : toutes les machines identiques
 - Mais cela fonctionne pour 16 types de machines au PSMN, 10 au CBP
- Un idéal (alias motivé) intégrateur & administrateur : ;-)
 - Déployé par L. Taulelle avec des rushs de documentation : PSMN
 - Déployé par T. Bellembois avec la documentation en ligne : IGFL

Futur de SIDUS

- Valorisation
 - Calcul scientifique, infrastructure de calcul scientifique
 - Gestion de parc, enseignement à la demande
- Simplification de l'installation & administration
 - Dédier une machine en Lecture/Écriture pour l'administration
 - Offrir une connexion SSH sur l'instance pour les opérations classiques
 - Utiliser Debian Preseed pour le processus d'installation
- Déploiement sur Mésocentre ou Grille
- SIDUS de partout
 - Lancer SIDUS de l'extérieur par VPN

D'autres informations ?

<http://www.cbp.ens-lyon.fr/sidus/>

Linux Journal
11/2013

Poster JRES
2013

Site Web CBP



Déduplication extrême d'OS avec SIDUS
Emmanuel Quémener & Loïs Taulelle
Centre Blaise Pascal & Pôle Scientifique de Modélisation Numérique, ENS-Lyon

Ce que SIDUS signifie :
Single Instance Distributing Universal System
Une instance unique distribuant un système d'exploitation universel

Ce que SIDUS n'est pas :

LTSP	FAI ou Kickstart	LiveCD réseau
LTSP : Linux Terminal Server Project Les plus • bon recyclage des vieux PC • migration vers distributions Les moins • toute la charge sur un seul serveur • périphériques locaux difficiles à intégrer	FAI : Fully Automatic Installation Les plus • automatisation de l'installation • processus nature et naturel Les moins • paramétrage initial • adaptation spécifique par outils	Une image ISO disponible sur le réseau... Les plus • unicité de la configuration • rapidité d'installation et de démarrage Les moins • personnalisation difficile • scalabilité quasi-nécessaire

Mais quelques composants que SIDUS partage :

- PXE : utilisation d'un démarrage en réseau
- TFTP : fourniture d'un noyau et d'un système de démarrage
- NFSROOT : système racine unique partagé par tous les clients

Ce que SIDUS propose :

- **Unité du système :** tous les clients démarrent exactement le même système (au bit près)
- **Usage des ressources locales :** les processeurs et mémoire vive exploités sont ceux des clients

SIDUS en 7 questions-réponses :

Pourquoi ?	Où & Quand ?	Comment ?
Uniforment de facto tous les postes • Limiter l'administration à un unique système • Assurer la reproductibilité • Rationaliser l'usage des postes de travail	Centre Blaise Pascal, ENS-Lyon : salle • 12 clients lignes bouclées en mars 2010 • 22 stations avec GPU différents fin 2013 Centre Blaise Pascal, ENS-Lyon : cluster • 24 nœuds en mars 2010 • 70 nœuds permanents fin 2013 Centre de calcul PSMN, ENS-Lyon : • 300 nœuds en 2012 en qualification • 330 nœuds en 2013 dont Equip@Homes Laboratoires, ENS-Lyon : • Laboratoire de Chimie : fin 2012 • Laboratoire LEMC & GFL : automne 2013 École de physique des Houches : • édition 2011, 2012, 2013 : jusqu'à 60 utilisateurs	Socle AUFS ► AUFS pour Another Union-Fix System ► Un système NFSroot en lecture seule ► Un système TMPFS en lecture/écriture ► AUFS comme glue entre les deux systèmes Installation en 8 étapes, 3 fondamentales 1 Formation d'un système racine par Débootstrap 2 Création de la séquence de démarrage (AUFS) 3 Importation des noyaux & intrinsèques sur serveur TFTP Administration simplifiée ► Passage dans l'instance par clout ► Application des commandes « standard » ► Montage des dossiers « système » au besoin
Nouveaux de cluster de calcul scientifique • Postes de salle libre-service • Stations de travail graphiques • Plateformes d'expérimentation numérique • COMODO pour Compute On My Own Device • démarrage d'un poste complet sous SIDUS • démarrage de SIDUS dans une machine virtuelle	Pour Qui ? ► Chercheur en informatique scientifique ► Ingénieur en calcul scientifique ► Gestionnaire de salle informatique ► Formateur exploitant des outils informatiques ► RSSI	Combien ? ► 1 réseau matériel 1 Gb/s (~400 Mbit/s) ► 1 serveur (virtuel) avec pour 100 clients : • 2 CPU • 4 Go de RAM • 50 Go d'espace par architecture complète ► 1 personne motivée pendant 1 journée

Pour en savoir plus : <http://www.cbp.ens-lyon.fr/sidus/>

<http://www.cbp.ens-lyon.fr/> & <http://www.ens-lyon.fr/7098/> emmanuel.quemener@ens-lyon.fr & lois.taulelle@ens-lyon.fr

