

Résumé

Le groupe système du *CCIN2P3* administre plus de 1800 serveurs en plus des unités de stockage. On distingue deux axes dans l'administration : matérielle et système (d'exploitation). Ceux-ci se traduisent respectivement en deux *canaux* du point de vue de l'accès : le *canal de gestion* et le *canal de données*. Le *canal de gestion* se matérialise par un dispositif physique et direct (e.g. clavier, port série, ...) faisant office d'Entrée/Sortie (E/S). Il peut être maîtrisé de manière déportée (télé-administration) en utilisant différents logiciels par le biais de protocoles standardisés et de matériels embarqués dans les serveurs. Les pannes matérielles sont ainsi détectées automatiquement et génèrent des alertes dans *RLS* ou par mail. Aussi, tous les workers peuvent être administrés, éteints ou allumés électriquement à distance. L'implémentation actuelle (logicielle et matérielle) est présentée en détail, ainsi que les travaux en cours et les évolutions du côté des constructeurs.

Télé-administration système

Piloter une salle machine à distance

Fabien Wernli

Centre de Calcul
IN2P3

Vendredi 13 Mars 2009

Télé-
administration
système

Fabien Wernli

Introduction

Quelques chiffres
Administration
Méthodes

Accès direct

Le *KVM*
Port série
Inconvénients

Accès déporté

Principe
Implémentation

Monitoring

Conclusions

Futur
Conclusion

- 1▶ Introduction
 - ▶ Quelques chiffres
 - ▶ Administration
 - ▶ Méthodes
- 2▶ Accès direct
 - ▶ Le *KVM*
 - ▶ Port série
 - ▶ Inconvénients
- 3▶ Accès déporté
 - ▶ Principe
 - ▶ Implémentation
- 4▶ Monitoring
- 5▶ Conclusions
 - ▶ Futur
 - ▶ Conclusion

Télé-
administration
système

Fabien Wernli

Introduction

Quelques chiffres
Administration
Méthodes

Accès direct

Le *KVM*
Port série
Inconvénients

Accès déporté

Principe
Implémentation

Monitoring

Conclusions

Futur
Conclusion

1 ▶ Introduction

- ▶ Quelques chiffres
- ▶ Administration
- ▶ Méthodes

2 ▶ Accès direct

- ▶ Le *KVM*
- ▶ Port série
- ▶ Inconvénients

3 ▶ Accès déporté

- ▶ Principe
- ▶ Implémentation

4 ▶ Monitoring

5 ▶ Conclusions

- ▶ Futur
- ▶ Conclusion



- ▶ **1146 Machines *Linux***
 - ▶ +252 workers *IBM iDPx dx360* recus fin Janvier 2009
- ▶ **368 Machines *Solaris***
 - ▶ dont 302 Thumpers/Thors (14496 disques, 2416 rams, 604 CPUs)
- ▶ **100 Machines *AIX***
- ▶ **Capacité stockage bande/disque : 30/8 Po**
 - ▶ 8 Administrateurs système (3 *Linux*, 2 *Solaris*, 1 *AIX*, +2)



- ▶ 1146 Machines *Linux*
- ▶ +252 workers *IBM iDPx dx360* recus fin Janvier 2009
- ▶ 368 Machines *Solaris*
- ▶ dont 302 Thumpers/Thors (14496 disques, 2416 rams, 604 CPUs)
- ▶ 100 Machines *AIX*
- ▶ Capacité stockage bande/disque : 30/8 Po
- ▶ 8 Administrateurs système (3 *Linux*, 2 *Solaris*, 1 *AIX*, +2)

Question

	2005-06	2009-02	2010
serveurs	~ 900	1861	?
Disques	~ 300 <i>To</i> (< 1000u)	8 <i>Po</i> (18558u)	?
Sysadmins	6	7 ± 1	?

Comment gérer autant d'unités ?

Question

	2005-06	2009-02	2010
serveurs	~ 900	1861	?
Disques	~ 300 <i>To</i> (< 1000u)	8 <i>Po</i> (18558u)	?
Sysadmins	6	7 ± 1	?

Comment gérer autant d'unités ?

Question

	2005-06	2009-02	2010
serveurs	~ 900	1861	?
Disques	~ 300 <i>To</i> (< 1000u)	8 <i>Po</i> (18558u)	?
Sysadmins	6	7 ± 1	?

Comment gérer autant d'unités ?

- ▶ Administration matérielle \neq système
- ▶ *Canal de gestion \neq canal de données*
- ▶ Accéder physiquement au serveur
 - ▶ Redémarrage physique
 - ▶ Contrôle des fonctions vitales du matériel
 - ▶ Surveillance
 - ▶ Tout cela peut être fait serveur éteint (mais sous tension)

- ▶ Administration matérielle \neq système
- ▶ *Canal de gestion \neq canal de données*
- ▶ Accéder physiquement au serveur
 - ▶ Redémarrage physique
 - ▶ Contrôle des fonctions vitales du matériel
 - ▶ Surveillance
 - ▶ Tout cela peut être fait serveur éteint (mais sous tension)

- ▶ Mise en marche (Première installation) 
- ▶  Arrêts de la salle machine 
- ▶ Surcharge / Plantage du système

```
Kernel panic - not syncing : Attempted to kill init!
```
- ▶ Intervention *single-user*

```
resize2fs /dev/rootvg/var
```
- ▶ Panne réseau

```
Ping RENATER-Paris down 17 :51 :39 jeudi 5 février
```
- ▶  Panne matérielle

```
0000000F  
00000003
```

- ▶ Ecran + clavier (+ souris) + boutons
- ▶ *KVM*
- ▶ Port série
- ▶ *IPMI*
- ▶ Serial Over LAN (*SOL*)

Télé-
administration
système

Fabien Wernli

Introduction

Quelques chiffres
Administration
Méthodes

Accès direct

Le *KVM*
Port série
Inconvénients

Accès déporté

Principe
Implémentation

Monitoring

Conclusions

Futur
Conclusion

- ▶ Introduction
 - ▶ Quelques chiffres
 - ▶ Administration
 - ▶ Méthodes

- ▶ Accès direct
 - ▶ Le *KVM*
 - ▶ Port série
 - ▶ Inconvénients

- ▶ Accès déporté
 - ▶ Principe
 - ▶ Implémentation

- ▶ Monitoring

- ▶ Conclusions
 - ▶ Futur
 - ▶ Conclusion

Télé-
administration
système

Fabien Wernli

Introduction

Quelques chiffres
Administration
Méthodes

Accès direct

Le KVM
Port série
Inconvénients

Accès déporté

Principe
Implémentation

Monitoring

Conclusions

Futur
Conclusion





les +

- ▶ Universel *x86* (même *Windows™*)
- ▶ Intervention techniciens
- ▶ Sur roulettes



les +

- ▶ Universel x86 (même *Windows™*)
- ▶ Intervention techniciens
- ▶ Sur roulettes

les -

- ▶ Pas de journalisation
- ▶ Pas très fiable





les +

- ▶ Universel x86 (même Windows™)
- ▶ Intervention techniciens
- ▶ Sur roulettes

les -

- ▶ Pas de journalisation
- ▶ Pas très fiable



- ▶ Se perd facilement

Télé-administration système

Fabien Wernli

Introduction

Quelques chiffres
Administration
Méthodes

Accès direct

Le KVM
Port série
Inconvénients

Accès déporté

Principe
Implémentation

Monitoring

Conclusions

Futur
Conclusion



les +

Google Maps

KVM

Find businesses, addresses and places of interest. [Learn more](#)

Print Send Link

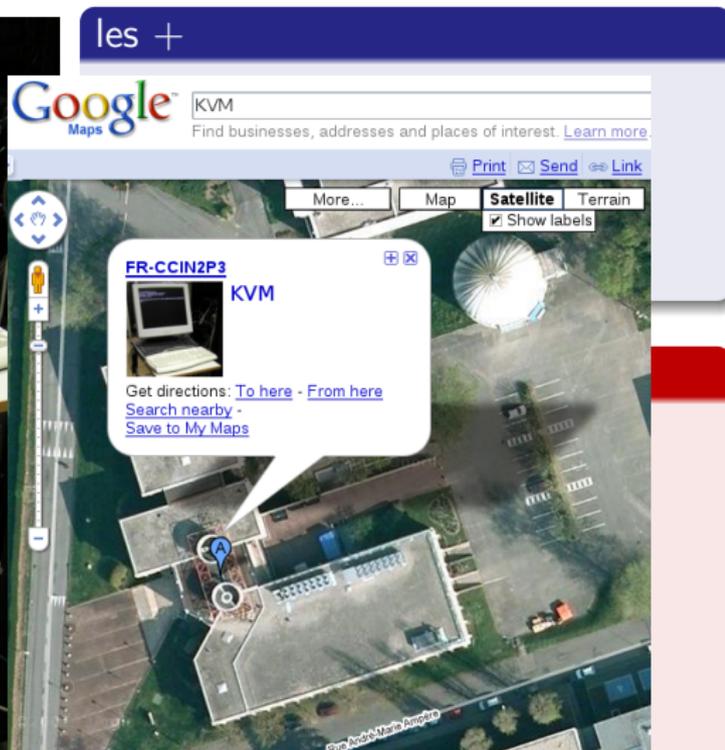
More... Map Satellite Terrain

Show labels

FR-CCIN2P3 KVM

Get directions: [To here](#) - [From here](#)

[Search nearby](#) - [Save to My Maps](#)



▶ Se perd facilement

Le "super" KVM

L'Ecran Clavier Souris partagé

Télé-
administration
système

Fabien Wernli

Introduction

Quelques chiffres
Administration
Méthodes

Accès direct

Le KVM
Port série
Inconvénients

Accès déporté

Principe
Implémentation

Monitoring

Conclusions

Futur
Conclusion



les +

- ▶ Universel x86
- ▶ Escamotable
- ▶ Une pièce par rack ou moins

les -

- ▶ Pas de journalisation
- ▶ Peu fiable sur le long terme
- ▶ Câblage additionel
- ▶ Inadapté pour un grand nombre d'unités

Télé-
administration
système

Fabien Wernli

Introduction

Quelques chiffres
Administration
Méthodes

Accès direct

Le KVM
Port série
Inconvénients

Accès déporté

Principe
Implémentation

Monitoring

Conclusions

Futur
Conclusion



les +

- ▶ Universel x86
- ▶ Escamotable
- ▶ Une pièce par rack ou moins

les -

- ▶ Pas de journalisation
- ▶ Peu fiable sur le long terme
- ▶ Câblage additionel
- ▶ Inadapté pour un grand nombre d'unités

Le "super" KVM

L'Ecran Clavier Souris partagé

Télé-
administration
système

Fabien Wernli

Introduction

Quelques chiffres
Administration
Méthodes

Accès direct

Le KVM
Port série
Inconvénients

Accès déporté

Principe
Implémentation

Monitoring

Conclusions

Futur
Conclusion



les +

- ▶ Universel x86
- ▶ Escamotable
- ▶ Une pièce par rack ou moins

les -

- ▶ Pas de journalisation
- ▶ Peu fiable sur le long terme
- ▶ Câblage additionel
- ▶ Inadapté pour un grand nombre d'unités



- ▶
- ▶ Connexion par console entrée/sortie...
- ▶ ...ou modem *RTC*

Configuration *Linux*

```
/etc/grub.conf : kernel /vmlinuz root=sda  
console=ttyS1,19200  
/etc/inittab :  
S1 :2345 :respawn :/sbin/agetty ttyS1 19200  
vt102  
/etc/securetty : ttyS1
```



- ▶
- ▶ Connexion par console entrée/sortie...
- ▶ ...ou modem *RTC*

Configuration *Linux*

```
/etc/grub.conf : kernel /vmlinuz root=sda
console=ttyS1,19200
/etc/inittab :
S1 :2345 :respawn :/sbin/agetty ttyS1 19200
vt102
/etc/securetty : ttyS1
```



- ▶
- ▶ Connexion par console entrée
- ▶ ...ou modem *RTC*



Configuration *Linux*

```

/etc/grub.conf : kernel /vmlin
console=ttyS1,19200
/etc/inittab :
S1 :2345 :respawn :/sbin/aget
vt102
/etc/securetty : ttyS1
    
```

les +

- ▶ Universel (OS “mode texte”, même AIX!)
- ▶ Ports multiples
- ▶ Très fiable

les –

- ▶ Configuration initiale
- ▶ Lenteur en 9600 baud

les +

- ▶ Universel (OS “mode texte”, même AIX!)
- ▶ Ports multiples
- ▶ Très fiable

les –

- ▶ Configuration initiale
- ▶ Lenteur en 9600 baud

- ▶ `ssh cconsrvX` (ayant beaucoup de ports série)



les +

- ▶ Centralisation^a
- ▶ Journalisation^b

^amodérée

^bvia *conserver*

les -

- ▶ Câblage additionel
- ▶ Inadapté pour un grand nombre d'unités
- ▶ 9600 baud obligatoire

- ▶ `ssh cconsrvX` (ayant beaucoup de ports série)



les +

- ▶ Centralisation^a
- ▶ Journalisation^b

^amodérée

^bvia *conserver*

les -

- ▶ Câblage additionel
- ▶ Inadapté pour un grand nombre d'unités
- ▶ 9600 baud obligatoire

- ▶ ssh cconsrvX (ayant beaucoup de ports série)



les +

- ▶ Centralisation^a
- ▶ Journalisation^b

^amodérée

^bvia *conserver*

les -

- ▶ Câblage additionnel
- ▶ Inadapté pour un grand nombre d'unités
- ▶ 9600 baud obligatoire

Télé- administration système

Fabien Wernli

Introduction

Quelques chiffres
Administration
Méthodes

Accès direct

Le KVM
Port série

Inconvénients

Accès déporté

Principe
Implémentation

Monitoring

Conclusions

Futur
Conclusion

Télé- administration système

Fabien Wernli

Introduction

Quelques chiffres
Administration
Méthodes

Accès direct

Le KVM
Port série

Inconvénients

Accès déporté

Principe
Implémentation

Monitoring

Conclusions

Futur
Conclusion



Télé- administration système

Fabien Wernli

Introduction

Quelques chiffres
Administration
Méthodes

Accès direct

Le KVM
Port série

Inconvénients

Accès déporté

Principe
Implémentation

Monitoring

Conclusions

Futur
Conclusion



Télé-administration système

Fabien Wernli

Introduction

Quelques chiffres
Administration
Méthodes

Accès direct

Le KVM
Port série

Inconvénients

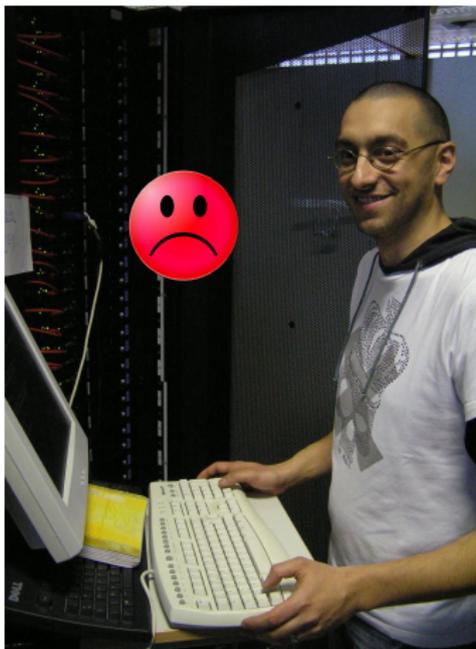
Accès déporté

Principe
Implémentation

Monitoring

Conclusions

Futur
Conclusion



Télé-administration système

Fabien Wernli

Introduction

Quelques chiffres
Administration
Méthodes

Accès direct

Le KVM
Port série
Inconvénients

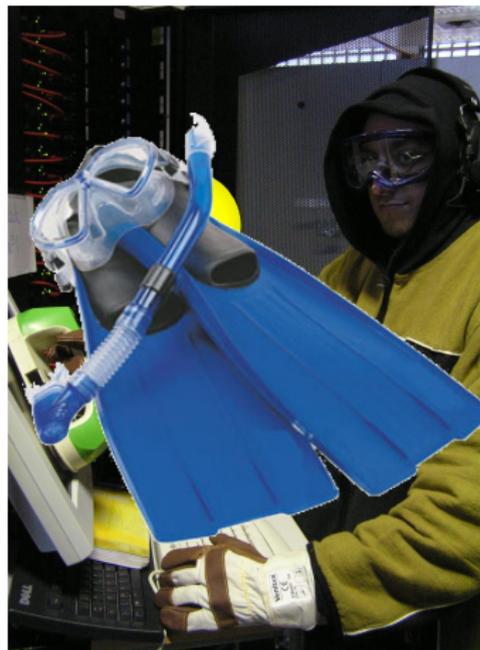
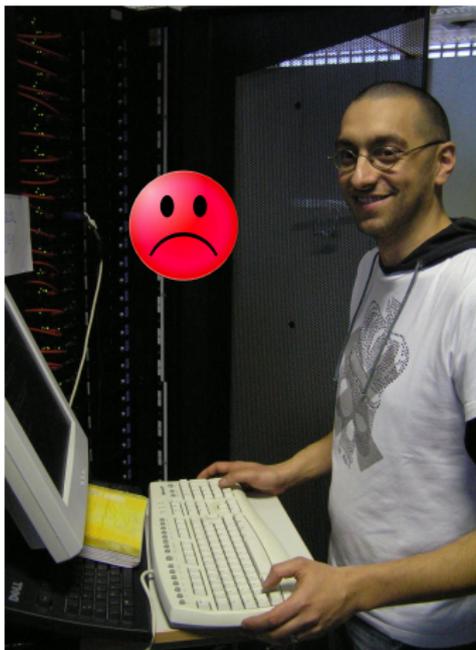
Accès déporté

Principe
Implémentation

Monitoring

Conclusions

Futur
Conclusion



les –

- ▶ Cablage supplémentaire
- ▶ Fiabilité médiocre
- ▶ *Non-scalabilité*
- ▶ Décentralisation
- ▶ *Input* limité au clavier (souris)

les –

- ▶ Cablage supplémentaire
- ▶ Fiabilité médiocre
- ▶ *Non-scalabilité*
- ▶ Décentralisation
- ▶ *Input* limité au clavier (souris)



Télé-
administration
système

Fabien Wernli

Introduction

Quelques chiffres
Administration
Méthodes

Accès direct

Le *KVM*
Port série
Inconvénients

Accès déporté

Principe
Implémentation

Monitoring

Conclusions

Futur
Conclusion

- 1 ▶ Introduction
 - ▶ Quelques chiffres
 - ▶ Administration
 - ▶ Méthodes
- 2 ▶ Accès direct
 - ▶ Le *KVM*
 - ▶ Port série
 - ▶ Inconvénients
- 3 ▶ Accès déporté
 - ▶ Principe
 - ▶ Implémentation
- 4 ▶ Monitoring
- 5 ▶ Conclusions
 - ▶ Futur
 - ▶ Conclusion

Canal de données \neq

OS \neq

```
ssh ccaliXY
```

\neq

\neq

Canal de gestion

lights-out management unit

Canal de données \neq
OS \neq

Canal de gestion

lights-out management unit

`ssh ccaliXY`

\neq



\neq

Canal de données \neq
OS \neq

Canal de gestion
lights-out management unit

`ssh ccaliXY`

\neq



\neq

- ▶ **Interface E/S** basé sur le port série



Rappel : les +

- ▶ Universel
- ▶ Robuste
- ▶ Journalisable

- ▶ **Transport** basé sur un réseau physique type *Ethernet*
 - ▶ Cable réseau séparé...

- ▶ **Interface E/S** basé sur le port série



Rappel : les +

- ▶ Universel
 - ▶ Robuste
 - ▶ Journalisable
 - ▶ Pas de câble supplémentaire
- ▶ **Transport** basé sur un réseau physique type *Ethernet*
 - ▶ Cable réseau séparé...
 - ▶ ... ou même **cable réseau** que POS

- ▶ **Interface** E/S basé sur le port série



Rappel : les +

- ▶ Universel
- ▶ Robuste
- ▶ Journalisable
- ▶ Pas de câble supplémentaire

- ▶ **Transport** basé sur un réseau physique type *Ethernet*

- ▶ Cable réseau séparé...

- ▶ ... ou même câble réseau
que l'OS



- ▶ **Gestion** Evénements

- ▶ Actions : boutons (power, reset, led)
- ▶ Alertes : températures, ventilateurs, disques, ...

- ▶ **Interface E/S** basé sur le port série



Rappel : les +

- ▶ Universel
- ▶ Robuste
- ▶ Journalisable
- ▶ **Pas de câble supplémentaire**

- ▶ **Transport** basé sur un réseau physique type *Ethernet*

- ▶ Cable réseau séparé...

- ▶ ... ou **même câble réseau que l'OS**



- ▶ **Gestion Événements**

- ▶ Actions : boutons (power, reset, led)
- ▶ Alertes : températures, ventilateurs, disques, ...

- ▶ **Interface E/S** basé sur le port série



Rappel : les +

- ▶ Universel
- ▶ Robuste
- ▶ Journalisable
- ▶ **Pas de câble supplémentaire**

- ▶ **Transport** basé sur un réseau physique type *Ethernet*

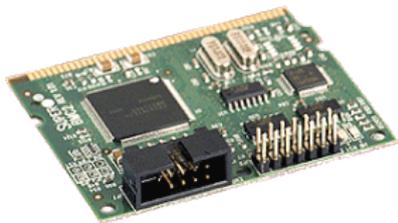
- ▶ Cable réseau séparé...

- ▶ ... ou **même câble réseau**
que l'OS



- ▶ **Gestion Événements**

- ▶ Actions : boutons (power, reset, led)
- ▶ Alertes : températures, ventilateurs, disques, ...



- ▶ Baseboard Management Controller (*Intel* et *x86*)
- ▶ Service Processor (*Sun Microsystems x86*)
- ▶ Integrated/Advanced Lights Out Manager (*Sun Microsystems SPARC*)
- ▶ *IBM Remote Supervisor Adapter (I et II)*
- ▶ *Dell Remote Access Controller*
- ▶ *Apple Xserve Lights Out Management*
- ▶ *Fujitsu Integrated Remote Management Controller*
- ▶ *hp Integrated Lights Out*



- ▶ Baseboard Management Controller (*Intel* et *x86*)
- ▶ Service Processor (*Sun Microsystems x86*)
- ▶ Integrated/Advanced Lights Out Manager (*Sun Microsystems SPARC*)
- ▶ *IBM Remote Supervisor Adapter (I et II)*
- ▶ *Dell Remote Access Controller*
- ▶ *Apple Xserve Lights Out Management*
- ▶ *Fujitsu Integrated Remote Management Controller*
- ▶ *hp Integrated Lights Out*



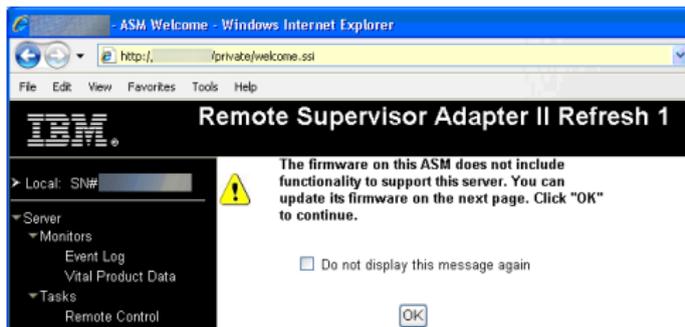
- ▶ Baseboard Management Controller (*Intel* et *x86*)
- ▶ Service Processor (*Sun Microsystems x86*)
- ▶ Integrated/Advanced Lights Out Manager (*Sun Microsystems SPARC*)
- ▶ *IBM* Remote Supervisor Adapter (I et II)
- ▶ *Dell* Remote Access Controller
- ▶ *Apple Xserve* Lights Out Management
- ▶ *Fujitsu* Integrated Remote Management Controller
- ▶ *hp* Integrated Lights Out



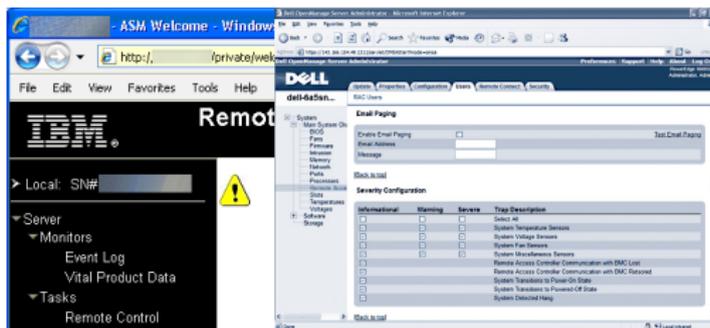
- ▶ Baseboard Management Controller (*Intel* et *x86*)
- ▶ Service Processor (*Sun Microsystems x86*)
- ▶ Integrated/Advanced Lights Out Manager (*Sun Microsystems SPARC*)
- ▶ *IBM* Remote Supervisor Adapter (I et II)
- ▶ *Dell* Remote Access Controller
- ▶ *Apple Xserve* Lights Out Management
- ▶ *Fujitsu* Integrated Remote Management Controller
- ▶ *hp* Integrated Lights Out



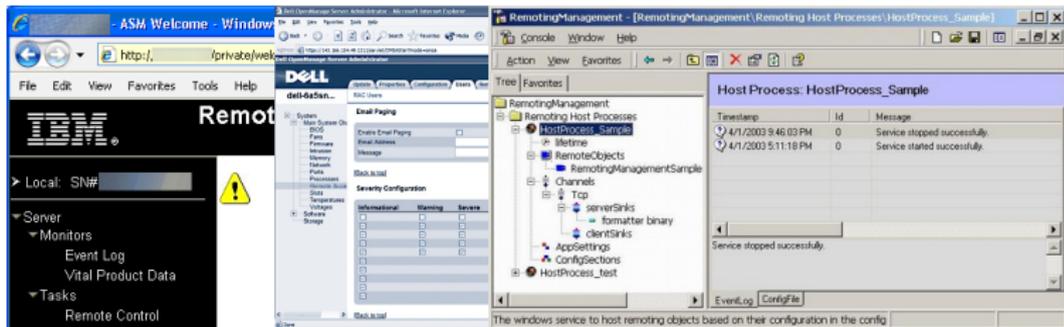
- ▶ Baseboard Management Controller (*Intel* et *x86*)
- ▶ Service Processor (*Sun Microsystems x86*)
- ▶ Integrated/Advanced Lights Out Manager (*Sun Microsystems SPARC*)
- ▶ *IBM* Remote Supervisor Adapter (I et II)
- ▶ *Dell* Remote Access Controller
- ▶ *Apple Xserve* Lights Out Management
- ▶ *Fujitsu* Integrated Remote Management Controller
- ▶ *hp* Integrated Lights Out



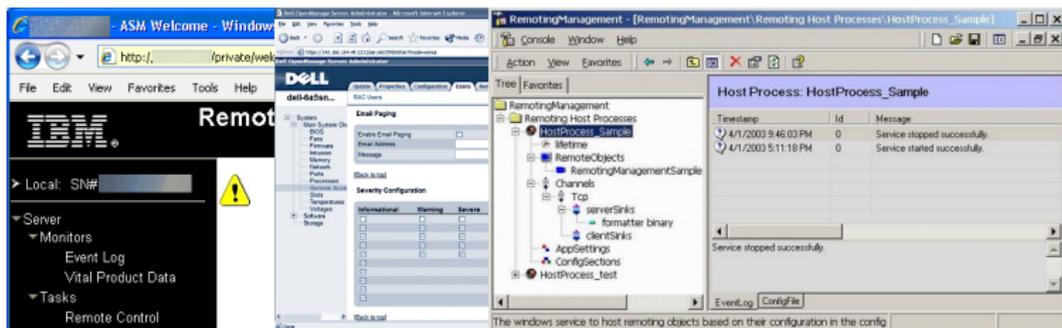
- ▶ *RSA de IBM : CLI, SNMP, console graphique sur interface web/java*
- ▶ *Dell DRAC : interface web, CLI*
- ▶ *Fujitsu IRMC : console graphique sur interface web/java*
- ▶ *hp iLO : console graphique, CLI*
- ▶ *Bladecenter IBM : interface web partagé sur les lames, console partagée*
- ▶ *ILOM ou ALOM de Sun : CLI sur ssh* poweron, showfru, ...



- ▶ *RSA de IBM : CLI, SNMP, console graphique sur interface web/java*
- ▶ *Dell DRAC : interface web, CLI*
- ▶ *Fujitsu IRMC : console graphique sur interface web/java*
- ▶ *hp iLO : console graphique, CLI*
- ▶ *Bladecenter IBM : interface web partagé sur les lames, console partagée*
- ▶ *ILOM ou ALOM de Sun : CLI sur ssh* poweron, showfru, ...



- ▶ *RSA de IBM* : CLI, SNMP, console graphique sur interface web/java
- ▶ *Dell DRAC* : interface web, CLI
- ▶ *Fujitsu IRMC* : console graphique sur interface web/java
- ▶ *hp iLO* : console graphique, CLI
- ▶ *Bladecenter IBM* : interface web partagé sur les lames, console partagée
- ▶ *iLOM ou ALOM de Sun* : CLI sur ssh `poweron, showfru, ...`



- ▶ *RSA de IBM* : CLI, SNMP, console graphique sur interface web/java
- ▶ *Dell DRAC* : interface web, CLI
- ▶ *Fujitsu IRMC* : console graphique sur interface web/java
- ▶ *hp iLO* : console graphique, CLI
- ▶ *Bladecenter IBM* : interface web partagé sur les lames, console partagée
- ▶ *ILOM ou ALOM de Sun* : CLI sur ssh `poweron, showfru, ...`

- ▶ Intelligent Platform Management Interface
- ▶ Standard interfaces de gestion
 - ▶ **Contrôle** : power on/off
 - ▶ **Inventaire** : numéros de série, versions, etc.
 - ▶ **Sondes** : température, voltage, ventilateurs, leds
 - ▶ Redirection **console** (Port série)

IPMI

“defines the messages and system interface to platform management hardware”

- ▶ Adopté par une douzaine de constructeurs
- ▶ *IPMI* v2.0 révision 3 : 15 Février 2006

- ▶ Intelligent Platform Management Interface
- ▶ Standard interfaces de gestion
 - ▶ **Contrôle** : power on/off
 - ▶ **Inventaire** : numéros de série, versions, etc.
 - ▶ **Sondes** : température, voltage, ventilateurs, leds
 - ▶ Redirection **console** (Port série)

IPMI

“defines the messages and system interface to platform management hardware”

- ▶ Adopté par une douzaine de constructeurs
- ▶ *IPMI* v2.0 révision 3 : 15 Février 2006

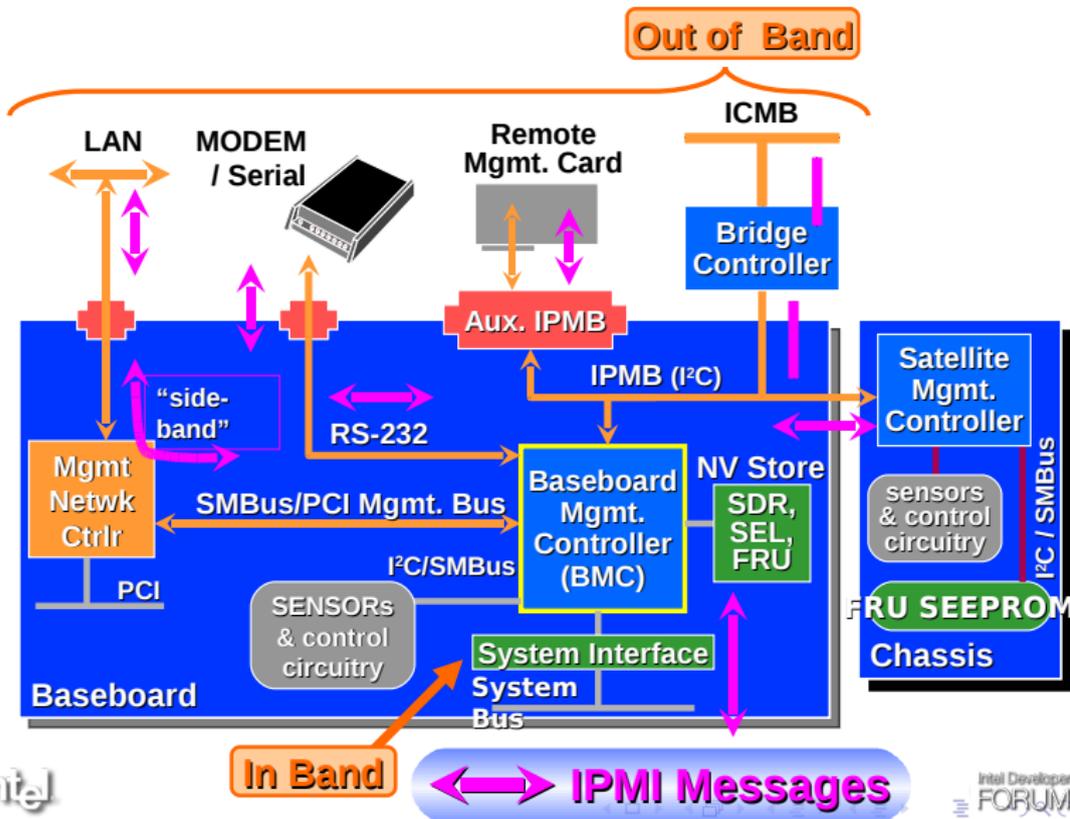
- ▶ Intelligent Platform Management Interface
- ▶ Standard interfaces de gestion
 - ▶ **Contrôle** : power on/off
 - ▶ **Inventaire** : numéros de série, versions, etc.
 - ▶ **Sondes** : température, voltage, ventilateurs, leds
 - ▶ Redirection **console** (Port série)

IPMI

“defines the messages and system interface to platform management hardware”

- ▶ Adopté par une douzaine de constructeurs
- ▶ *IPMI* v2.0 révision 3 : 15 Février 2006

- Télé-administration système
- Fabien Wernli
- Introduction
 - Quelques chiffres
 - Administration
 - Méthodes
- Accès direct
 - Le KVM
 - Port série
 - Inconvénients
- Accès déporté
 - Principe
 - Implémentation
- Monitoring
- Conclusions
 - Futur
 - Conclusion

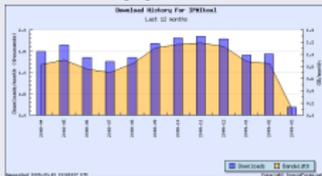


```
Commands :
raw          Send a RAW IPMI request and print response
i2c          Send an I2C Master Write-Read command and print response
spd          Print SPD info from remote I2C device
lan          Configure LAN Channels
chassis      Get chassis status and set power state
power        Shortcut to chassis power commands
event        Send pre-defined events to MC
mc           Management Controller status and global enables
sdr          Print Sensor Data Repository entries and readings
sensor       Print detailed sensor information
fru          Print built-in FRU and scan SDR for FRU locators
sel          Print System Event Log (SEL)
pef          Configure Platform Event Filtering (PEF)
sol          Configure and connect IPMIv2.0 Serial-over-LAN
tsol         Configure and connect with Tyan IPMIv1.5 Serial-over-LAN
isol         Configure IPMIv1.5 Serial-over-LAN
user         Configure Management Controller users
channel      Configure Management Controller channels
session      Print session information
sunoem       OEM Commands for Sun servers
kontron      OEM Commands for Kontron devices
picmg        Run a PICMG/ATCA extended cmd
fwum         Update IPMC using Kontron OEM Firmware Update Manager
firewall     Configure Firmware Firewall
shell        Launch interactive IPMI shell
exec         Run list of commands from file
set          Set runtime variable for shell and exec
hpm          Update HPM components using PICMG HPM.1 file
```

```
Commands :
raw          Send a RAW IPMI request and print response
i2c          Send an I2C Master Write-Read command and print response
spd          Print SPD info from remote I2C device
lan          Configure LAN Channels
chassis      Get chassis status and set power state
power       Shortcut to chassis power commands
event       Send pre-defined events to MC
mc           Management Controller status and global enables
sdr         Print Sensor Data Repository entries and readings
sensor      Print detailed sensor information
fru         Print built-in FRU and scan SDR for FRU locators
sel         Print System Event Log (SEL)
pef         Configure Platform Event Filtering (PEF)
sol         Configure and connect IPMIv2.0 Serial-over-LAN
tsol        Configure and connect with Tyan IPMIv1.5 Serial-over-LAN
isol        Configure IPMIv1.5 Serial-over-LAN
user        Configure Management Controller users
channel     Configure Management Controller channels
session     Print session information
sunoem      OEM Commands for Sun servers
kontron    kontron    OEM Commands for Kontron devices
picmg       Run a PICMG/ATCA extended cmd
fwum       Update IPMC using Kontron OEM Firmware Update Manager
firewall    Configure Firmware Firewall
shell       Launch interactive IPMI shell
exec        Run list of commands from file
set         Set runtime variable for shell and exec
hpm         Update HPM components using PICMG HPM.1 file
```

les +

- ▶ Opensource (License *BSD*)
- ▶ Développement actif (30 Write par mois sur le *CVS*)



- ▶ Supporte un grand nombre de matériels

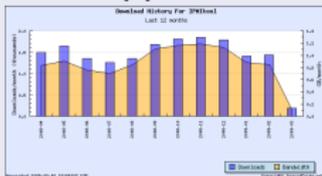
les -

- ▶ Monotâche
- ▶ Configuration initiale des clients (user/passwd)
- ▶ Compatibilité des clients
- ▶ *SOL* souvent mal (ou pas) supporté

Démo

les +

- ▶ Opensource (License *BSD*)
- ▶ Développement actif (30 Write par mois sur le *CVS*)



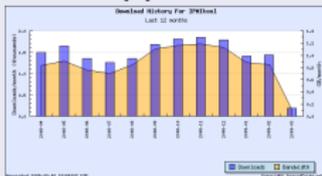
- ▶ Supporte un grand nombre de matériels

les -

- ▶ Monotâche
- ▶ Configuration initiale des clients (*user/passwd*)
- ▶ Compatibilité des clients
- ▶ *SOL* souvent mal (ou pas) supporté

les +

- ▶ Opensource (License *BSD*)
- ▶ Développement actif (30 Write par mois sur le *CVS*)



- ▶ Supporte un grand nombre de matériels

les -

- ▶ Monotâche
- ▶ Configuration initiale des clients (*user/passwd*)
- ▶ Compatibilité des clients
- ▶ *SOL* souvent mal (ou pas) supporté

Démo

- ▶ Lancer n commandes `ipmitool` simultanément
- ▶ Développé au CC
- ▶ Sortie formatée
- ▶ la *dB smurf* contient les spécificités de chaque client
- ▶ Timeout

Exemple

```
$ pipmi -c 'power status' $(select_machine -rack 'iDataplex 2')
ccw10556 Chassis Power is on
ccw10520 Error : Unable to establish IPMI v2 / RMCP+ session
ccw10520 Unable to get Chassis Power Status
ccw10508 Chassis Power is on
ccw10517 Chassis Power is on
ccw10545 Chassis Power is off
ccw10567 Chassis Power is on
ccw10543 Chassis Power is on
[...]
```

▶ Démo

- ▶ Lancer n commandes `ipmitool` simultanément
- ▶ Développé au CC
- ▶ Sortie formatée
- ▶ la *dB smurf* contient les spécificités de chaque client
- ▶ Timeout

Exemple

```
$ pipmi -c 'power status' $(select_machine -rack 'iDataplex 2')
ccw10556 Chassis Power is on
ccw10520 Error : Unable to establish IPMI v2 / RMCP+ session
ccw10520 Unable to get Chassis Power Status
ccw10508 Chassis Power is on
ccw10517 Chassis Power is on
ccw10545 Chassis Power is off
ccw10567 Chassis Power is on
ccw10543 Chassis Power is on
[...]
```

- ▶ Démo

▶ http://download.intel.com/design/servers/ipmi/IPMIv2_0rev1_0.pdf

“Serial Over LAN (SOL) is the name for the redirection of baseboard serial controller traffic over an IPMI session. This can be used to enable asynchronous serial-based OS and pre-OS communication over a connection to the BMC [...]”

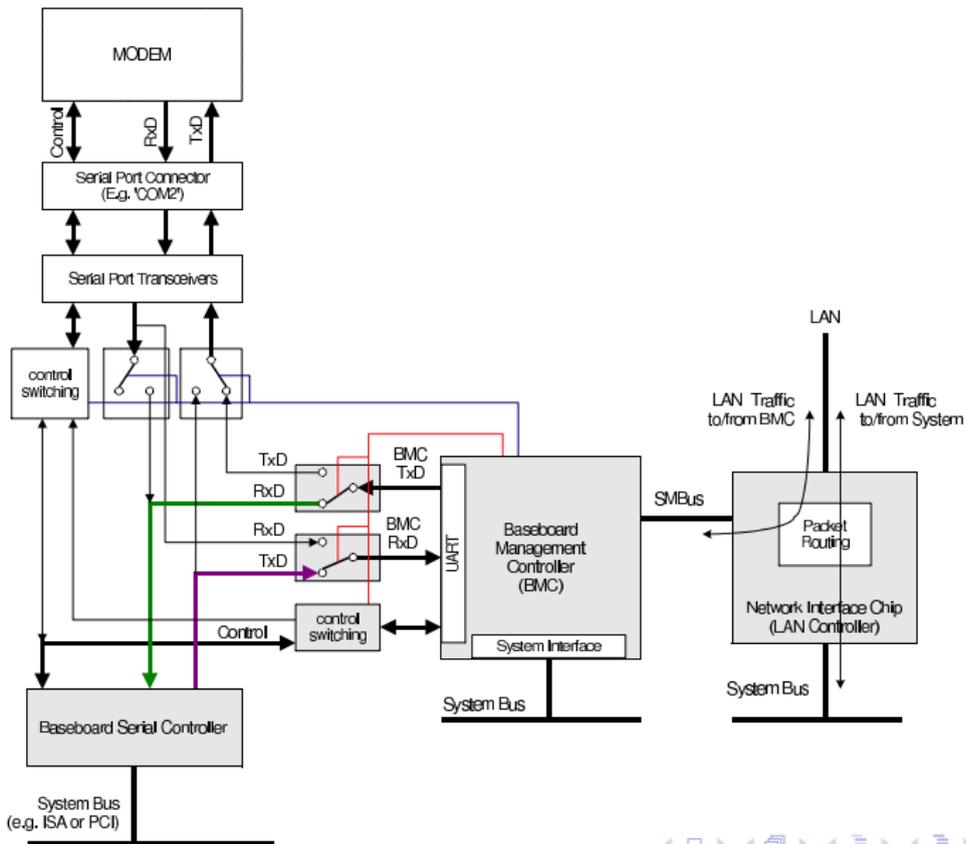
- ▶ Remplace l'Ecran-Clavier
- ▶ Permet l'accès à l'environnement pre-démarrage (e.g. BIOS)
- ▶

```
ipmitool [...] sol activate
```
- ▶ Demo

▶ http://download.intel.com/design/servers/ipmi/IPMIv2_0rev1_0.pdf

“Serial Over LAN (SOL) is the name for the redirection of baseboard serial controller traffic over an IPMI session. This can be used to enable asynchronous serial-based OS and pre-OS communication over a connection to the BMC [...]”

- ▶ Remplace l'Ecran-Clavier
- ▶ Permet l'accès à l'environnement pre-démarrage (e.g. BIOS)
- ▶ `ipmitool [...] sol activate`
- ▶ Demo



- ▶ Permet de gérer un ensemble hétérogène de *consoles*
- ▶ Plus anciennes : port série *R232* sur concentrateur (*Equinox ESP-16, perle IOLAN STS, ...*)
- ▶ Moins anciennes : *SOL* divers et variés

↳ *Support de la console IPMI 2.0 SOL* : wrapper `ipmitool`
↳ *Scripts* soumis sur [GitHub](#)

- ▶ Nouveaux serveurs *IPMI 2.0 SOL* : wrapper `ipmitool`
- ▶ Scripts soumis sur [GitHub](#)

- ▶ Permet de gérer un ensemble hétérogène de *consoles*
- ▶ Plus anciennes : port série *R232* sur concentrateur (*Equinox ESP-16, perle IOLAN STS, ...*)
- ▶ Moins anciennes : *SOL* divers et variés
 - ▶ scripts *expect* pour systèmes non *IPMI-proven*

Supermicro ipmicli, DellPoweredge 2850 solproxxy, IBM e326m IPMI 1.5, Sun X4x00 pas de SOL IPMI 2.0

- ▶ Nouveaux serveurs *IPMI 2.0 SOL* : wrapper *ipmitool*
- ▶ Scripts soumis sur [ipmitool.org](#)

- ▶ Permet de gérer un ensemble hétérogène de *consoles*
- ▶ Plus anciennes : port série *R232* sur concentrateur (*Equinox ESP-16, perle IOLAN STS, ...*)
- ▶ Moins anciennes : *SOL* divers et variés
 - ▶ scripts *expect* pour systèmes non *IPMI-proven*

Exemples

Supermicro ipmicli, DellPoweredge 2850 solproxy, IBM e326m IPMI 1.5, Sun X4x00 pas de SOL IPMI 2.0

- ▶ Nouveaux serveurs *IPMI 2.0 SOL* : wrapper `ipmitool`
- ▶ Scripts soumis sur [ipmiutils.org](#)

- ▶ Permet de gérer un ensemble hétérogène de *consoles*
- ▶ Plus anciennes : port série *R232* sur concentrateur (*Equinox ESP-16, perle IOLAN STS, ...*)
- ▶ Moins anciennes : *SOL* divers et variés
 - ▶ scripts *expect* pour systèmes non *IPMI-proven*

Exemples

Supermicro ipmicli, DellPoweredge 2850 solproxy, IBM e326m IPMI 1.5, Sun X4x00 pas de SOL IPMI 2.0

- ▶ Nouveaux serveurs *IPMI 2.0 SOL* : wrapper `ipmitool`
- ▶ Scripts soumis sur <http://conserver.com/Contrib>

- ▶ Permet de gérer un ensemble hétérogène de *consoles*
- ▶ Plus anciennes : port série *R232* sur concentrateur (*Equinox ESP-16, perle IOLAN STS, ...*)
- ▶ Moins anciennes : *SOL* divers et variés
 - ▶ scripts *expect* pour systèmes non *IPMI-proven*

Exemples

Supermicro ipmicli, DellPoweredge 2850 solproxy, IBM e326m IPMI 1.5, Sun X4x00 pas de SOL IPMI 2.0

- ▶ Nouveaux serveurs *IPMI 2.0 SOL* : wrapper **ipmitool**
- ▶ Scripts soumis sur <http://conserver.com/Contrib>

- ▶ Permet de gérer un ensemble hétérogène de *consoles*
- ▶ Plus anciennes : port série *R232* sur concentrateur (*Equinox ESP-16, perle IOLAN STS, ...*)
- ▶ Moins anciennes : *SOL* divers et variés
 - ▶ scripts *expect* pour systèmes non *IPMI-proven*

Exemples

Supermicro ipmicli, DellPoweredge 2850 solproxy, IBM e326m IPMI 1.5, Sun X4x00 pas de SOL IPMI 2.0

- ▶ Nouveaux serveurs *IPMI 2.0 SOL* : wrapper **ipmitool**
- ▶ Scripts soumis sur <http://conserver.com/contrib>

- ▶ Commande **unique** : `console`
- ▶ Gestion des *ACL* utilisateurs (e.g. Ingénieurs externes)
- ▶ *Partage* de console (1 rw, plusieurs ro)
- ▶ *Journalisation*
- ▶ Gestion hiérarchique des consoles
- ▶ Gestion des *break* e.g. *SysRq*
- ▶ Communications *cryptées* (*openssl*)

- ▶ Commande **unique** : `console`
- ▶ Gestion des **ACL** utilisateurs (e.g. Ingénieurs externes)
- ▶ **Partage** de console (1 rw, plusieurs ro)
- ▶ **Journalisation**
- ▶ Gestion hiérarchique des consoles
- ▶ Gestion des **break** e.g. `SysRq`
- ▶ Communications **cryptées** (`openssl`)

- ▶ Commande **unique** : `console`
- ▶ Gestion des **ACL** utilisateurs (e.g. Ingénieurs externes)
- ▶ **Partage** de console (1 rw, plusieurs ro)
- ▶ **Journalisation**
- ▶ Gestion hiérarchique des consoles
- ▶ Gestion des **break** e.g. `SysRq`
- ▶ Communications **cryptées** (`openssl`)

- ▶ **Plusieurs serveurs** interchangeable¹
- ▶ 6 serveurs de console + 2 desktop + 1 desktop telecom (cconsrvX + ccpctc)
- ▶ 1 seul fichier de configuration partagé par CVS
 - ▶ Commandes de contrôle *e.g. IPMI* également hébergées sur ces serveurs
 - ▶ *BMC* sur *VLAN* séparé (*IP=192.168.x.y*)
 - ▶ Serveur *DNS* master (cconsrv10) slave (cconsrvX) zone *sp.in2p3.fr*
 - ▶ La plupart des *BMC* partagent le port Ethernet avec le système

¹pour le *SOL* uniquement

- ▶ **Plusieurs serveurs** interchangeable¹
- ▶ 6 serveurs de console + 2 desktop + 1 desktop telecom (cconsvrX + ccpctc)
- ▶ 1 seul fichier de configuration partagé par CVS
- ▶ Commandes de contrôle e.g. *IPMI* également hébergées sur ces serveurs
- ▶ *BMC* sur *VLAN* séparé (*IP=192.168.x.y*)
- ▶ Serveur *DNS* master (cconsvr10) slave (cconsvrX) zone *sp.in2p3.fr*
- ▶ La plupart des *BMC* partagent le port Ethernet avec le système

¹pour le *SOL* uniquement

Todo

- ▶ Serveur de console spécial Technicien extérieur
- ▶ Failover automatique
- ▶ Alertes Journal (*swatch*)

Alternatives

`conman`, `conmux`

Todo

- ▶ Serveur de console spécial Technicien extérieur
- ▶ Failover automatique
- ▶ Alertes Journal (*swatch*)

Alternatives

`conman`, `conmux`

Télé-
administration
système

Fabien Wernli

Introduction

Quelques chiffres
Administration
Méthodes

Accès direct

Le *KVM*
Port série
Inconvénients

Accès déporté

Principe
Implémentation

Monitoring

Conclusions

Futur
Conclusion

- 1 ▶ Introduction
 - ▶ Quelques chiffres
 - ▶ Administration
 - ▶ Méthodes
- 2 ▶ Accès direct
 - ▶ Le *KVM*
 - ▶ Port série
 - ▶ Inconvénients
- 3 ▶ Accès déporté
 - ▶ Principe
 - ▶ Implémentation
- 4 ▶ **Monitoring**
- 5 ▶ Conclusions
 - ▶ Futur
 - ▶ Conclusion

▶ Serveur syslog central :

▶ `syslogd.conf`

```
*.notice @ccsyslog.in2p3.fr
```

▶ messages kernel

```
kernel.* @ccsyslog.in2p3.fr
```

```
kernel.* @ccsyslog.in2p3.fr
```

▶ Messages applicatifs

```
*.info @ccsyslog.in2p3.fr
```

▶ `traps SNMP` des BMC via démons `snmptrapd`

▶ alertes BMC via démons locaux (*in-band*) `ipmievd`

▶ Alertes mél et/ou *RLS(NG)* via `swatch`

▶ Serveur syslog central :

▶ `syslogd.conf`

```
*.notice @ccsyslog.in2p3.fr
```

▶ messages kernel

```
KWATCH NOTICE ccsrbi2 GENUNIX ERROR
```

```
KWATCH NOTICE ccwl1243 Rebooted
```

▶ Messages applicatifs

```
KWATCH ERROR ccdcamcli07 DCACHE POSTGRES FATAL : database fz does not exist
```

▶ *traps SNMP* des *BMC* via démons `snmptrapd`

```
snmptrapd : SUN-ILOM-PET-MIB (petTrapPowerSupplyStateDeasserted)
```

▶ alertes *BMC* via démons locaux (*in-band*) `ipmievd`

```
ipmievd : Memory sensor Mem ECC Warning - Transition to Non-critical from OK
```

▶ Alertes mél et/ou *RLS(NG)* via `swatch`

▶ Serveur syslog central :

▶ `syslogd.conf`

```
*.notice @ccsyslog.in2p3.fr
```

▶ messages kernel

```
KWATCH NOTICE ccsrb12 GENUNIX ERROR
```

```
KWATCH NOTICE ccwl1243 Rebooted
```

▶ Messages applicatifs

```
KWATCH ERROR ccdcamcli07 DCACHE POSTGRES FATAL : database fz does not exist
```

▶ *traps SNMP* des *BMC* via démons `snmptrapd`

```
snmptrapd : SUN-ILOM-PET-MIB (petTrapPowerSupplyStateDeasserted)
```

▶ alertes *BMC* via démons locaux (*in-band*) `ipmievd`

```
ipmievd : Memory sensor Mem ECC Warning - Transition to Non-critical from OK
```

▶ Alertes mél et/ou *RLS(NG)* via `swatch`

Télé-
administration
système

Fabien Wernli

Introduction

Quelques chiffres
Administration
Méthodes

Accès direct

Le *KVM*
Port série
Inconvénients

Accès déporté

Principe
Implémentation

Monitoring

Conclusions

Futur
Conclusion

- 1 ▶ Introduction
 - ▶ Quelques chiffres
 - ▶ Administration
 - ▶ Méthodes
- 2 ▶ Accès direct
 - ▶ Le *KVM*
 - ▶ Port série
 - ▶ Inconvénients
- 3 ▶ Accès déporté
 - ▶ Principe
 - ▶ Implémentation
- 4 ▶ Monitoring
- 5 ▶ Conclusions
 - ▶ Futur
 - ▶ Conclusion

- ▶ *AMT pour intel Core2 vPro*
- ▶ Gestion des versions de logiciels installés dans la “*BMC*” via agent logiciel
- ▶ Remote boot (à la *PXE*)
- ▶ Remote power on/off
- ▶ *SOL*
- ▶ Firewall
- ▶ Interrupteur permettant d’isoler un *PC compromis*
- ▶ Configuration déportée
- ▶ Accès déporté via *XML* et *SOAP*
- ▶ *AMD* semble développer l’équivalent (*Remote IT*)

:-)

- ▶ 1694 serveurs avec *IPMI* (437 sans)
- ▶ 1757 serveurs dans conserver (374 sans)

:-)

- ▶ Complexité croissante des *BMC* (boot *iLOM* : 3 min)
- ▶ Manque de conformité au standard *IPMI*
- ▶ Manque de documentation des constructeurs
- ▶ Support *IPMI* uniquement pour *x86* :-)

:-)

- ▶ 1694 serveurs avec *IPMI* (437 sans)
- ▶ 1757 serveurs dans conserver (374 sans)

:-(

- ▶ Complexité croissante des *BMC* (boot *ILOM* : 3 min)
- ▶ Manque de conformité au standard *IPMI*
- ▶ Manque de documentation des constructeurs
- ▶ Support *IPMI* uniquement pour *x86* :-)



Out-of-band management

wikipedia ▶ http://en.wikipedia.org/wiki/Out-of-band_management



Intelligent Platform Management Interface

Intel ▶ <http://www.intel.com/design/servers/ipmi/>



Xserve Lights Out Management

Apple ▶ <http://support.apple.com/kb/HT2773>



Active Management Technology

Intel ▶ <http://www.intel.com/technology/platform-technology/intel-amt/>



Conserver ▶ <http://www.conserver.com>

Plume ▶ <http://www.projet-plume.org/fr/fiche/conserver>



The Simple WATCHer of Logfiles

Swatch ▶ <http://swatch.sourceforge.net>



Simple Network Management Protocol

wikipedia ▶ <http://en.wikipedia.org/wiki/SNMP>