Использование XEN для учебной грид-инфраструктуры

Попова Е., Бережная А., Котляр В.
ИФВЭ

Виртуализация

- Понятие виртуализации представляет собой сокрытие настоящей реализации какого-либо процесса или объекта от истинного его представления для того, кто им пользуется.
- Понятие виртуализации условно можно разделить на две фундаментально различающиеся категории:

- **виртуализация платформ**
  Продуктом этого вида виртуализации являются виртуальные машины - некие программные абстракции, запускаемые на платформе реальных аппаратно-программных систем.

- **виртуализация ресурсов**
  Данный вид виртуализации преследует своей целью комбинирование или упрощение представления аппаратных ресурсов для пользователя и получение неких пользовательских абстракций оборудования, пространств имен, сетей и т.п.

Виды виртуализации платформ

Эмуляция

Полная виртуализация

Паравиртуализация

Виртуализация на уровне ОС

Виды виртуализации ресурсов

- Объединение, агрегация и концентрация компонентов
- Кластеризация компьютеров и распределенные вычисления (grid computing)
- Разделение ресурсов (partitioning)
- Инкапсуляция

Проекты виртуализации для Linux

Проект	Тип	Лицензия
Bochs	Эмуляция	LGPL
QEMU	Эмуляция	LGPL/GPL
VMware	Полная виртуализация	Проприентарное
z/VM	Полная виртуализация	Проприентарное
Xen	Паравиртуализация	GPL
UML	Паравиртуализация	GPL
Linux-VServer	Виртуализация на уровне ОС	GPL
OpenVZ	Виртуализация на уровне ОС	GPL

Виртуальные машины

- Одна физическая система с памятью, процессорами, устройствами ввода-вывода
- Множество сред исполнения, каждая из которых отображается в виде идентичной физической системы.
- ОС, запущенная в виртуальной машине, «не знает» что использует разделяемые ресурсы.
  - ВМ должны быть изолированы друг от друга, не смотря на то что они используют аппаратные средства совместно
  - Работа одной ВМ не может повредить другим ВМ
  - Накладные расходы должны быть приемлемыми
- ВМ должны использовать аппаратные средства на равных условиях
- ВММ – монитор виртуальных машин

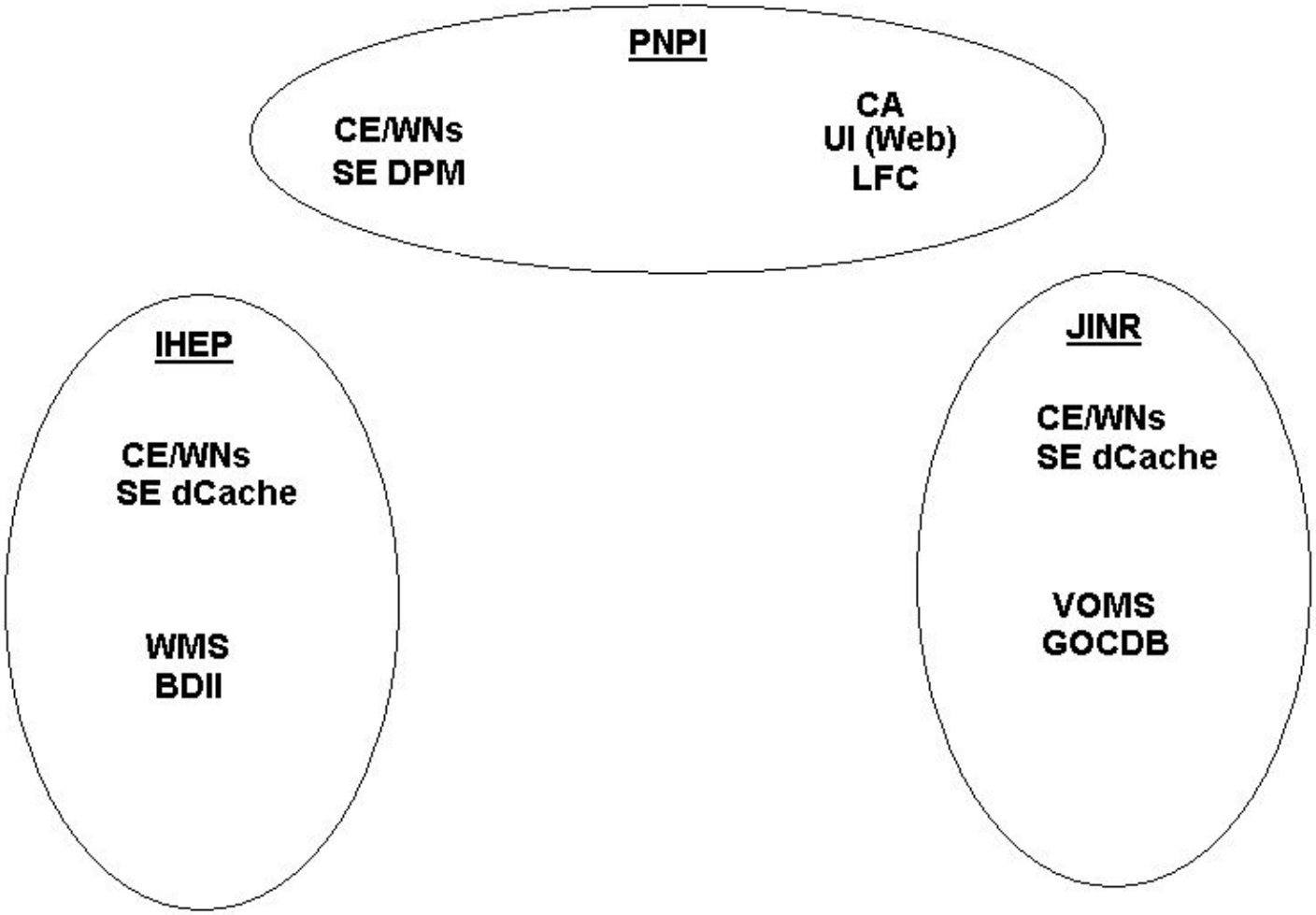
Монитор виртуальных машин

- Гипервизор
- Обеспечивает интерфейс для совместного существования множества ВМ.
- Может запускать несколько ОС на одном компьютере.
- Гарантирует стабильность, даже если одна из ОС выйдет из строя, другие продолжат функционировать
- Устраняет необходимость содержания отдельных машин для разных ОС
- Изолирует экземпляры ОС и распределяет физические ресурсы между работающими ВМ.

Учебная грид-инфраструктура

- В проекте EGEE учебная грид-инфраструктура (t-infrastructure) предназначена для распространения знаний о грид-технологиях и дает возможность пользователям и системным администраторам приобрести непосредственные опыт работы с грид-технологиями.

Учебная грид-инфраструктура



Xen

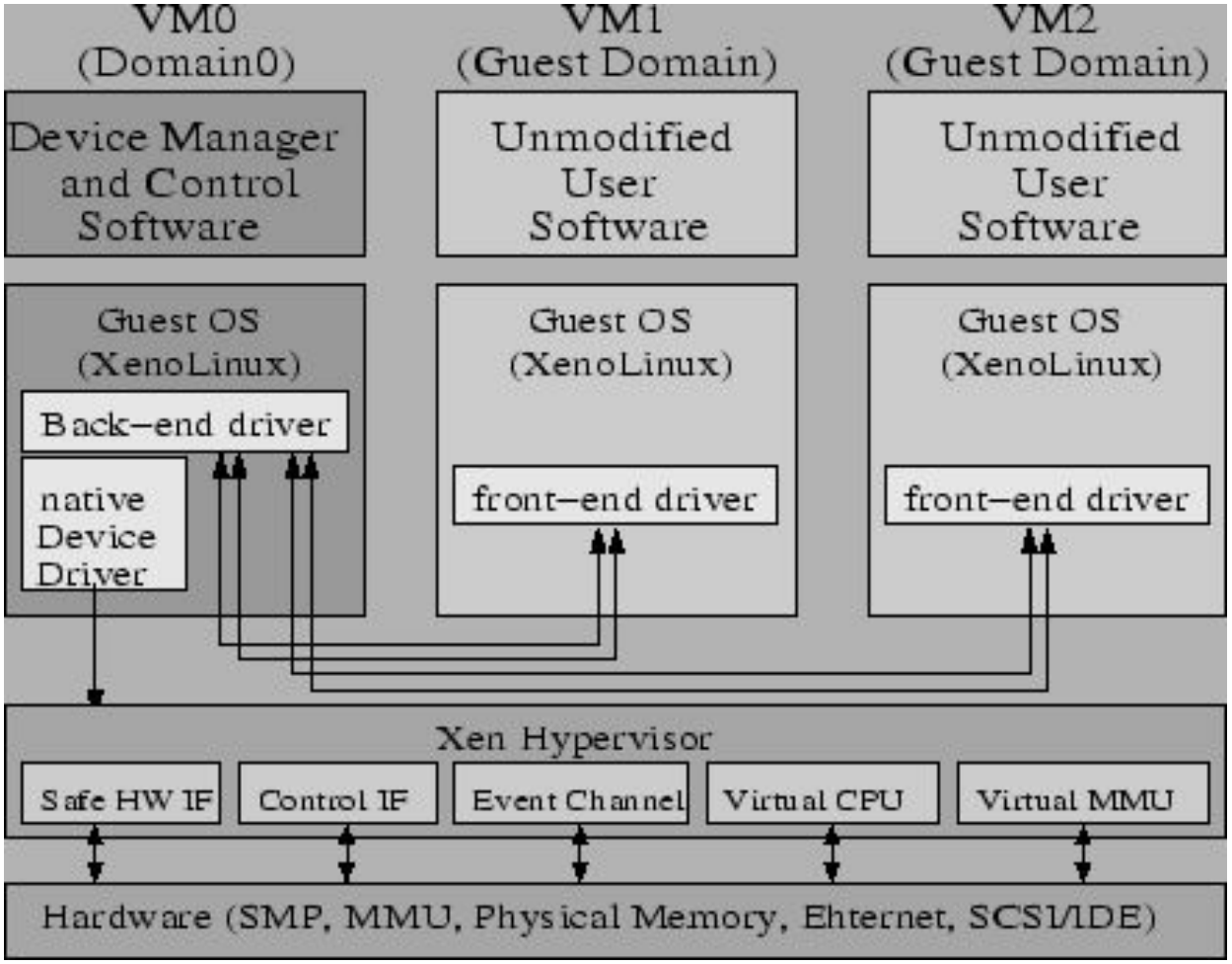
- Монитор виртуальных машин (гипервизор) с поддержкой паравиртуализации (para-virtualization) для процессоров x86 архитектуры, распространяющийся с открытым исходным кодом (opensource)

- Варианты использования
  - Консолидация серверов
  - Независимость от аппаратного обеспечения
  - Запуск множества различных ОС
  - Разработка ядра ОС
  - Кластерные системы
  - Аппаратная поддержка для новых ОС.

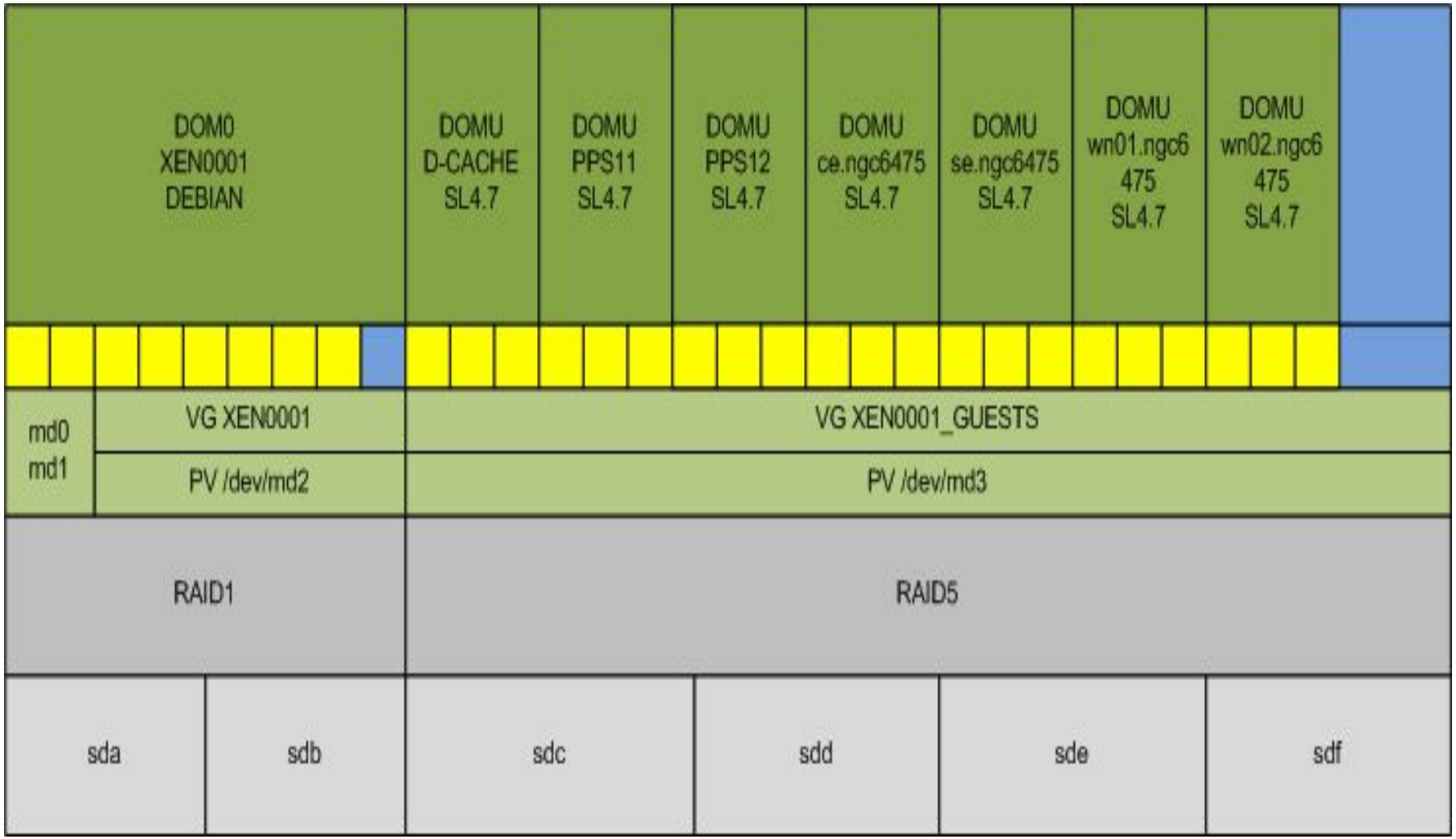
Основы Xen

- Поддержка немодифицированных бинарных приложений
  - Полезно для пользователей
- Поддержка полнофункциональных, многозадачных гостевых ОС
- Использование паравиртуализации для обеспечения производительности и высокого уровня изоляции ресурсов.
- Гостевые ОС должны быть модифицированы чтобы запускаться Монитором ВМ.
- Гостевые ОС больше не выполняются в нулевом кольце защиты, т.к. это кольцо занято гипервизором. Поэтому ОС должны быть изменены

Структура Xen-системы



Xen-сервер


DOM0
XEN0001
DEBIAN
DOMU
D-CACHE
SL4.7
DOMU
PPS11
SL4.7
DOMU
PPS12
SL4.7
DOMU
ce.ngc6475
SL4.7
DOMU
se.ngc6475
SL4.7
DOMU
wn01.ngc6
475
SL4.7
DOMU
wn02.ngc6
475
SL4.7
md0
md1
VG XEN0001
VG XEN0001_GUESTS
PV /dev/md2
PV /dev/md3
RAID1
RAID5
sda
sdb
sdc
sdd
sde
sdf

Шаблон ФС для DomU

```
xen0001:~# cat /etc/xen/SL4.7_001.orig
kernel = "/boot/vmlinuz"
ramdisk="initrd.img"
name = "SL4.7_001"
memory = 256
selinux = 0
disk = [ 'phy:xen0001_guests/SL4.7_001,xvda,w' ]
vif = [ 'ip=10.165.1.10' ]
#root = "/dev/xen0001_guests/SL4.7_001 ro"
pci = ['06.02.0']
#bootloader="/usr/lib/xen-3.0.3-1/bin/pygrub"
on_poweroff = 'destroy'
on_reboot   = 'restart'
on_crash    = 'restart'
```

```
/usr/sbin/xm create SLC4Xguest -c \
        kernel=vmlinuz \
        ramdisk=initrd.img \
        extra="method=http://linuxsoft.cern.ch/cern/slc4X/i386/"
```

```
xm shutdown -w SL4.7_001
kpartx -d /dev/xen0001_guests/SL4.7_001
mount /dev/mapper/xen0001_guests-SL4.7_001p1 /mnt/tmp/
dump 0 -L FULL_boot -f BACKUP-XEN/SL4.7_001_boot.colddump /mnt/tmp/
restore -rf /root/BACKUP-XEN/SL4.7_001_boot.colddump
....
```

DomUs

```
xen0001:~# cat /etc/xen/ce.ngc6475
#kernel = "/boot/vmlinuz-2.6.9-78.0.1.EL.cernxenU"
#ramdisk="/boot/initrd-2.6.9-78.0.1.EL.cernxenU.img"
name = "ce.ngc6475"
memory = 1024
maxmem = 1024
extra = "mem=1073741824"
cpu_weight = 512
cpu_cap =0
cpus = ""
vcpus = 2
selinux = 0
disk = [ 'phy:xen0001_guests/ce.ngc6475.boot,xvdc1,w',
         'phy:xen0001_guests/ce.ngc6475.root,xvdc2,w',
         'phy:xen0001_guests/ce.ngc6475.swap,xvdc3,w']
root = "/dev/xvdc2 ro"
vif = [ 'ip=194.190.165.216' ]
pci = ['06.02.0']
bootloader="/usr/lib/xen-3.0.3-1/bin/pygrub"
on_poweroff = 'destroy'
on_reboot   = 'restart'
on_crash    = 'restart'
```

CE.NGC6475

```
xen0001:~# cat /etc/xen/se.ngc6475
#kernel = "/boot/vmlinuz-2.6.9-78.0.1.EL.cernxenU"
#ramdisk="/boot/initrd-2.6.9-78.0.1.EL.cernxenU.img"
name = "se.ngc6475"
memory = 512
maxmem = 512
cpus = ""
vcpus = 2
selinux = 0
disk = [ 'phy:xen0001_guests/se.ngc6475.boot,xvdd1,w',
         'phy:xen0001_guests/se.ngc6475.root,xvdd2,w',
         'phy:xen0001_guests/se.ngc6475.storage,xvdd4,w',
         'phy:xen0001_guests/se.ngc6475.storage1,xvdd5,w',
         'phy:xen0001_guests/se.ngc6475.storage2,xvdd6,w',
         'phy:xen0001_guests/se.ngc6475.swap,xvdd3,w']
root = "/dev/xvdd2 ro"
vif = [ 'ip=194.190.165.176' ]
pci = ['06.02.0']
bootloader="/usr/lib/xen-3.0.3-1/bin/pygrub"
on_poweroff = 'destroy'
on_reboot   = 'restart'
on_crash    = 'restart'
```

SE.NGC6475

Name	ID	Mem(MiB)	VCPUs	State	Time(s)
Domain-0	0	128	2	r-----	29015.3
ce.ngc6475	43	768	2	r-----	45848.9
d-cache	1	96	1	-b----	1309.2
pps11	2	96	1	-b----	1705.3
pps12	30	96	1	-b----	1411.1
se.ngc6475	44	512	2	r-----	330594.2
wn01.ngc6475	34	128	1	-b----	1840.7
wn02.ngc6475	31	128	1	-b----	2005.4

Заключение

- Xen это программный монитор виртуальных машин (гипервизор)
- Позволяет множеству ОС работать параллельно на одной и той же машине
- Требует модификации ОС для работы с гипервизором.
- Обеспечивает изоляцию между доменами.
- Xen сегодня…
  - Open-source проект с GPL
  - Текущая версия 3.0.
  - NetBSD, Linux (некоторые дистрибутивы, включая SuSE, Fedora, RHEL, Mandrake), FreeBSD, Plan9, и NetWare поддерживают Xen. Работа над WindowsXP была успешна, но из-за запрета лицензии не используется

- Виртуализация на аппаратном уровне
  - Intel выпускает новую линию процессоров с поддержкой виртуализации
  - 2 вида операций ЦП
    - Добавлено новое кольцо защиты (-1), где сможет функционировать гипервизор.
    - Гостевые ОС продолжат работать на уровне нулевого кольца, что не требует переработки кода ядра

Источники

- Xen and the Art of Virtualization By Paul Barham, Boris Dragovic, Stevan Hand, Tim Harris, Alex Ho, Rolf Neugebauer, Ian Pratt, and Andrew Warfield.Presented by Diana Carroll
- http://www.osp.ru/os/2008/03/5015349/
- http://www.ibm.com/developerworks/ru/library/l-linuxvirt/
- http://xgu.ru/xen/manual/