



服务器虚拟化 在数据中心的 综合应用

上海交通大学网络信息中心

黄保青

2008.1



上海交大网络信息中心系统部简介

- § 服务于全校6000名教工，40000名学生
- § 管理100多台服务器：
IBM、Sun小型机、PC服务器
- § 服务器操作系统：
AIX, Solaris, Linux, Windows, FreeBSD
- § 互联网基础服务：
Mail、DNS、NTP、LDAP、RADIUS、虚拟主机
- § 数字化校园相关的应用系统



服务器虚拟化实施状况

§ 最初目的：提高服务器效率，减少服务数量

§ 2004年开始关注并引入虚拟化技术

- 2004年 3月应用VMware GSX
- 2005年 应用Solaris 10的 zone，评估Virtuozzo for Windows
- 2006年评估Xen 3.0，开始使用VMware Server

0011 0010 § 2007年 6月引入OpenVZ

- 2008年准备使用Virtuozzo for Windows

§ 以操作系统虚拟化技术为主

- Solaris: zone
- Linux: OpenVZ
- Windows: Swsoft Virtuozzo



服务器虚拟化的技术分类

§ 硬件级虚拟化

- IBM PowerPC, Sun Domain

§ 全虚拟:

- VMware, Qmenu, VirtualServer

§ 半虚拟(para-virtual)

- Xen

§ 操作系统虚拟化

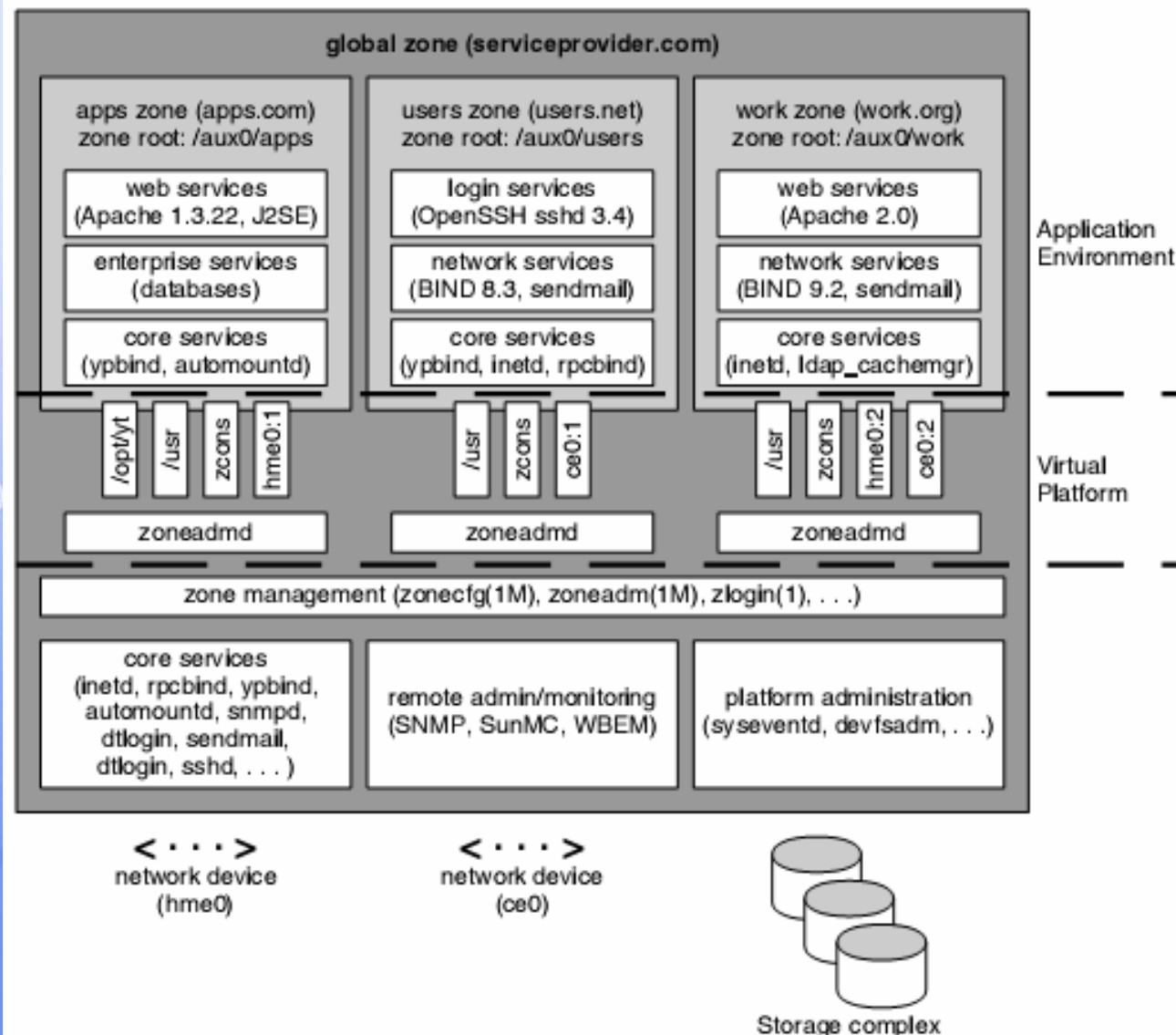
- FreeBSD Jail
- User Mode Linux, Linux Virtual Server, OpenVZ, Virtuozzo for Linux
- Solaris Zone
- Virtuozzo for Windows



Solaris Zone

§ Global

§ Non-Global





Zone虚拟化特点

- § 性能损失小，数100个zone并存
- § 独立的用户和文件系统，进程隔离，网络数据包隔离
- § 对绝大多数应用透明
 - Apache, OpenLDAP, Oracle, Squid, Proftpd
- § 通过Resource Management实现资源调度
- § 共享文件系统，减少存储容量
- § 支持IPv6

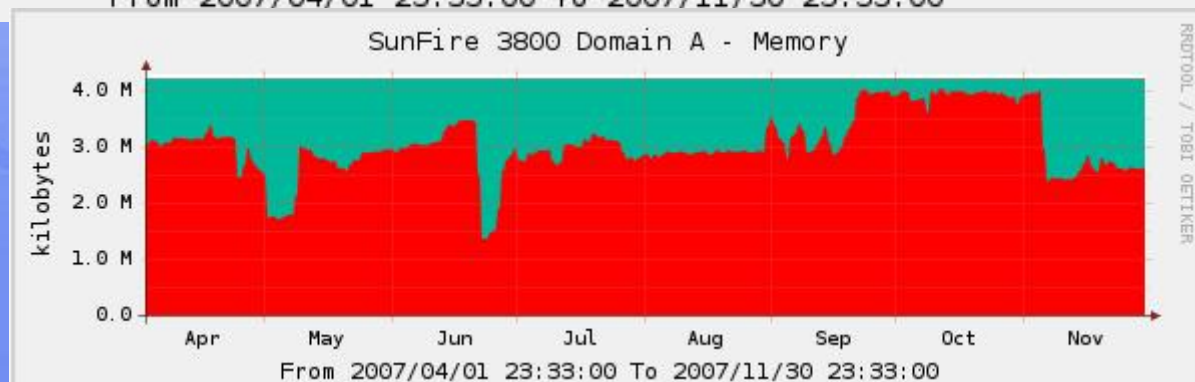
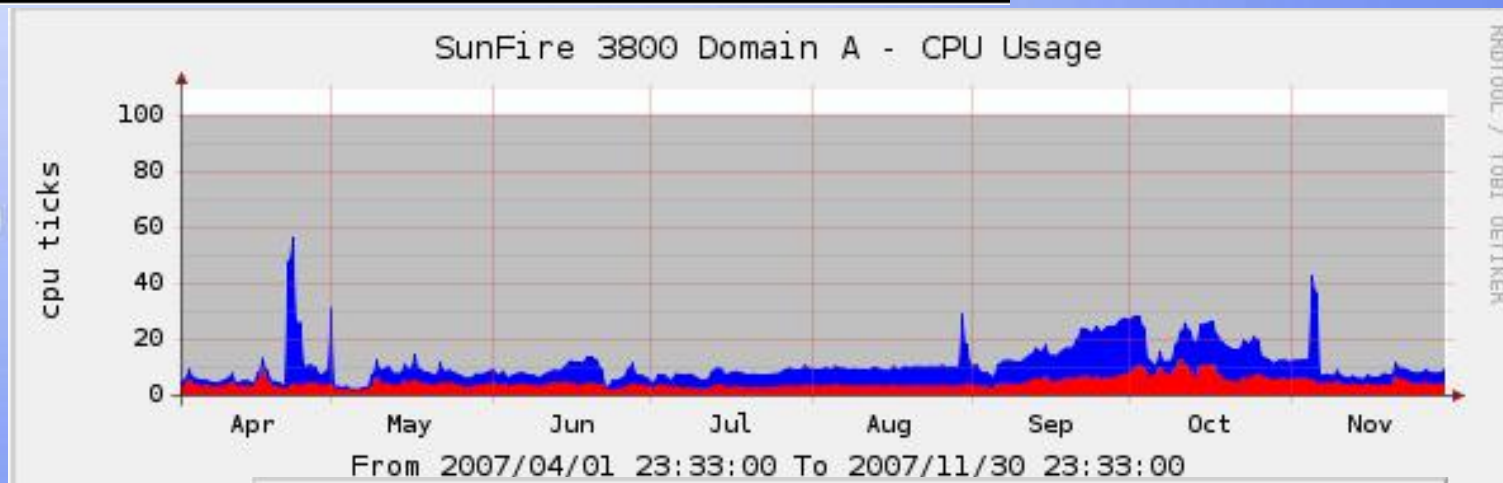
```
# ifconfig -a6
lo0: flags=2002000849<UP,LOOPBACK,RUNNING,MULTICAST,IPv6,VIRTUAL> mtu 8252 index
1
    inet6 ::1/128
lo0:1: flags=2002000849<UP,LOOPBACK,RUNNING,MULTICAST,IPv6,VIRTUAL> mtu 8252 ind
ex 1
    zone sunlab1
    inet6 ::1/128
hme0: flags=2000841<UP,RUNNING,MULTICAST,IPv6> mtu 1500 index 2
    inet6 fe80::203:baff:fe02:4c4c/10
    ether 0:3:ba:2:4c:4c
hme0:1: flags=2080841<UP,RUNNING,MULTICAST,ADDRCONF,IPv6> mtu 1500 index 2
    inet6 2001:da8:8000:1:203:baff:fe02:4c4c/64
hme0:2: flags=2000841<UP,RUNNING,MULTICAST,IPv6> mtu 1500 index 2
    zone sunlab1
    inet6 fe80::ca78:208:3/10
hme0:3: flags=2000841<UP,RUNNING,MULTICAST,IPv6> mtu 1500 index 2
    zone sunlab1
    inet6 2001:da8:8000:1::208:3/64
```




Zone运行实例

```
# zoneadm list -cv
```

ID	NAME	STATUS	PATH
0	global	running	/
1	████████jtu.edu.cn	running	/var/zones/████████jtu.edu.cn
3	████████	running	/data/zones/████████
4	████████sjtu.edu.cn	running	/var/zones/████████sjtu.edu.cn
5	ipv6cj	running	/var/zones/ipv6cj
6	sunlab1	running	/var/zones/sunlab1
-	████████	installed	/data/zones/████████





Zone的一些不足

- 我们使用www.blastwave.org的包管理软件
 - 在global zone中设置 PKGADDFLAGS=-G
- § 在u3 版后才能较好支持Zone的迁移



Linux OpenVZ



- § SWsoft 资助
- § Hardware node
- § Host OS
- § Virtual Environment
- § Template

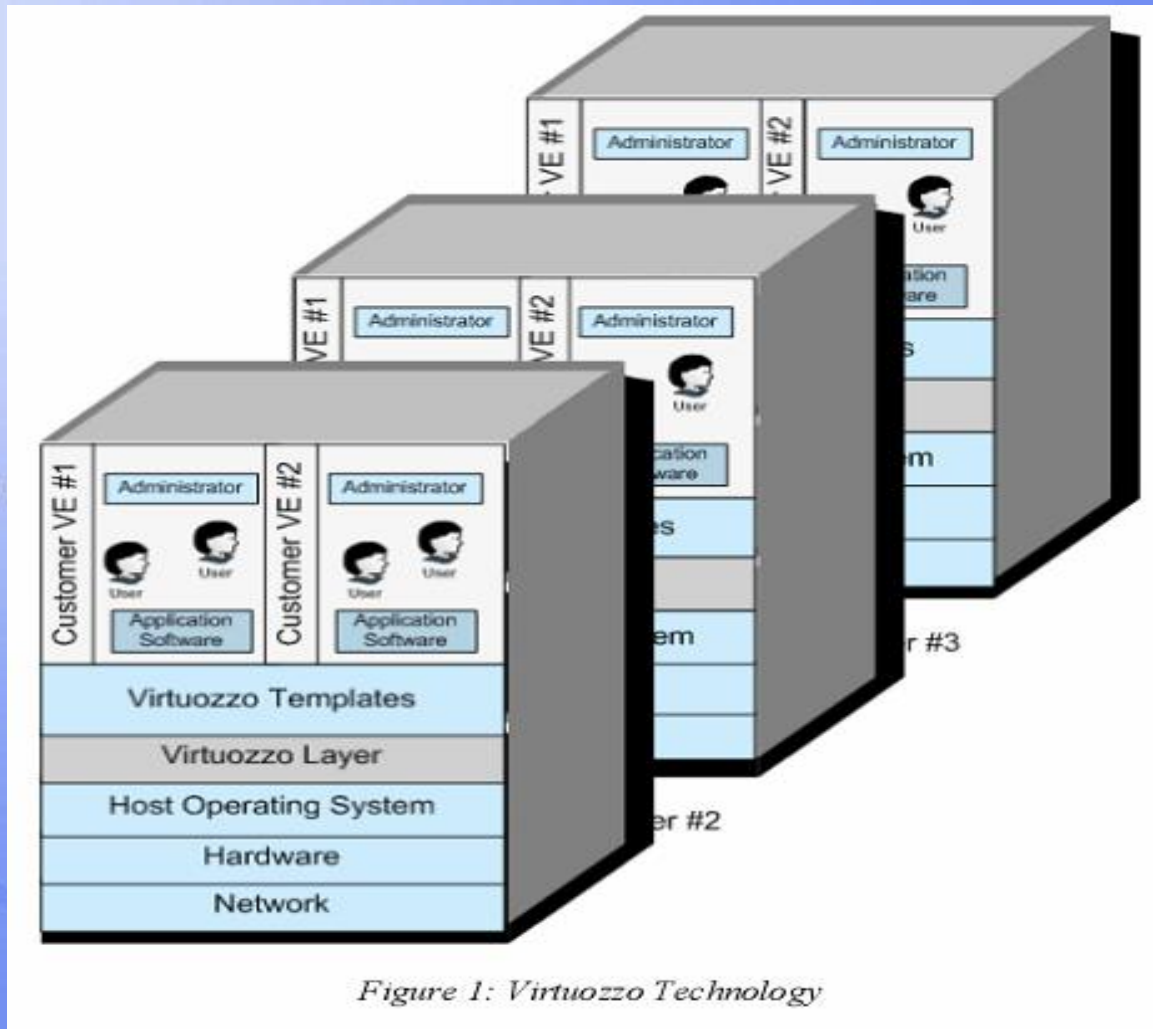


Figure 1: Virtuozzo Technology



OpenVZ特点

- § Linux 开源项目(www.openvz.org)
- § 使用同一Linux内核，不同发行版并存
- § 性能损失小，可支持数百VE
- § 支持资源调度
 - 基本配置分配的资源很少
- § 支持IPv6
- § 基本可实现动态迁移
- § 用template生成VE速度很快
- § 可以管理自己的IPTables
- § 可与VMware和Xen等虚拟化软件共存



运行实例

多种发行版并存，64位与32位系统并存

```
[root@vhostlinux1 ~]# vzlist -a
  VEID      NPROC STATUS IP_ADDR      HOSTNAME
  151        -  stopped 202.120. [REDACTED] [REDACTED].idc.sjtu.edu.cn
  152       46  running 202.120. [REDACTED] [REDACTED].sjtu.edu.cn
  153        4  running 202.120. [REDACTED] [REDACTED].idc.sjtu.edu.cn
  154       26  running 202.120. [REDACTED] [REDACTED].sjtu.edu.cn
  155       58  running 202.120. [REDACTED] [REDACTED].sjtu.edu.cn
  156       15  running 202.120. [REDACTED] wiki.idc.sjtu.edu.cn
  157       76  running 202.120. [REDACTED] wiki32.idc.sjtu.edu.cn
```

0011 0010 1010 1101 0001 0100 1011

实际的应用

Bind9, OpenLDAP, Postfix, Tomcat, Apache, MySQL, Zope

45



Virtuozzo for Windows

- § SWsoft, 与OpenVZ概念类似
- § 在Windows 2003中虚拟多个2003
- § 性能损失小, 可支持数百VE
- § 支持应用隔离, 支持资源调度
- § VE可以在不同Hardware Node间快速迁移
- § Hardware Node之间可以实现High Availability
- § 有商业化支持



Virtuozzo管理界面

Virtuozzo Management Console 4.0

File Action View Help [Submit feedback on Virtuozzo 4.0 beta](#)

Local Server
+ File Manager
+ Monitor
+ Users and Groups
+ Services
+ Logs
+ Templates
+ Backups
+ Scheduled Tasks
+ Security
+ VE Samples
+ **Virtuozzo Virtual Environment**

ID	Name	Hostname	IP Address	Status	Resources	OS	Architecture
1046	IIS6	IIS6					
1047	IIS5	IIS5					

Action Name

Properties of Virtual Environment "IIS6"

General Network **Resources** Terminal Server Options Advanced

CPU parameters
Disk Quota
System parameters

Parameters

Name	Hard Limit	Description
CPU guarantee	Not limited	CPU guarantee
CPU limit	100 percents	CPU limit
CPU Units	1000 pcs	CPU Units
Number of CPUs	Not limited	Number of CPUs

Scale... Properties

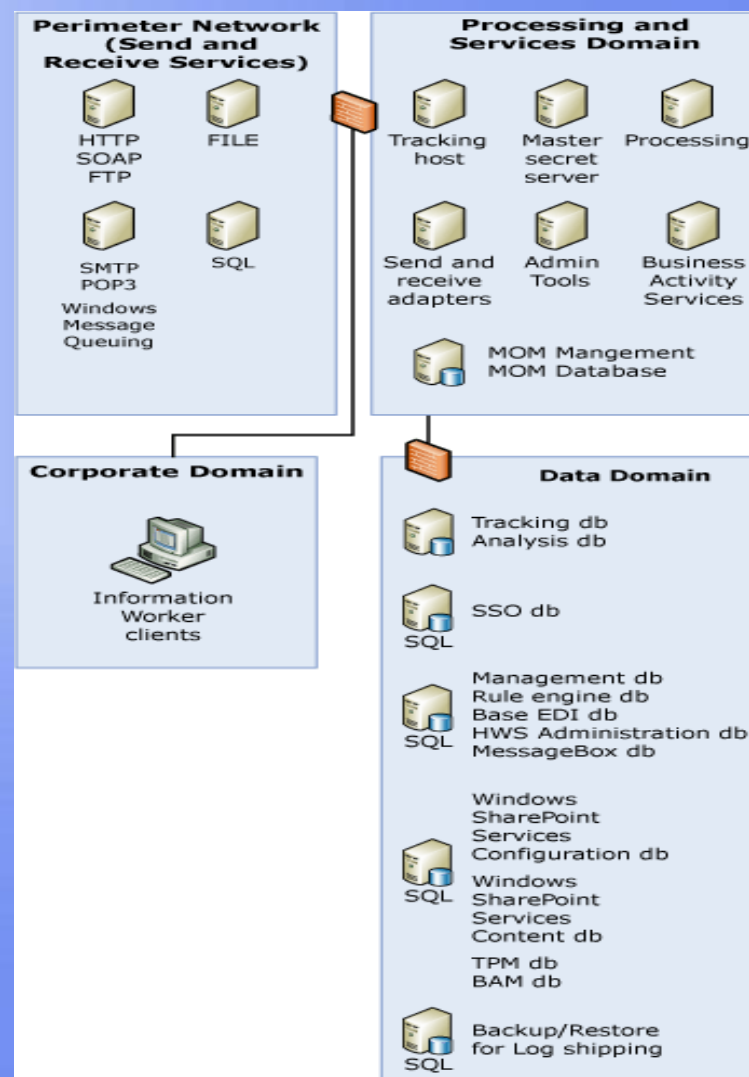
Help OK Cancel Apply



应用

- § SQLServer 2005 & 2000并存
- § 虚拟主机服务
- § Biztalk服务器归并

0011 0010 1010 1101 0001 0100 1011





Virtuozzo for Windows 一些不足

§ VE不能运行内核级的服务如：

- 文件共享，病毒查杀

§ 最好使用SWsoft的Windows 更新

§ IPv6 ?

0011 0010 1010 1101 0001 0100 1011

12
45



VMware

- § 更多的用于测试
- § 支持的操作系统种类多
- § 支持的网络类型多
- § 安装操作容易

0011 0010 1010 1101 0001 0100 1011

12
45



虚拟机的资源控制

§ 操作系统虚拟化容易相互影响

	VMware		Xen		Solaris	
	<i>Good</i>	<i>Bad</i>	<i>Good</i>	<i>Bad</i>	<i>Good</i>	<i>Bad</i>
Memory	0	91.30	0.03	DNR	DNR	DNR
Fork	0	DNR	0.04	DNR	90.03	87.80
CPU	0	0	0	0.03	0	0
Disk Intensive	0	39.80	1.37	11.53	1.48	1.23
Network Server Transmits	0	52.9	0.01	0.33	4.00	3.53
Network Server Receives	0	0	0.03	0.10	1.24	1.67

Table 1: Summary of Stress Test Results Percent of degradation in good response rate. For each test, the percent degradation for both the bad or misbehaving VM is shown, as well as, the average degradation across the three good or well-behaving VMs. DNR indicates the SPECweb client reported only an error and no actual results because of the unresponsiveness of the server it was testing.



不同虚拟技术的混用

§ OpenVZ+VMware

- 在Linux中虚拟Windows

§ VMWare + OpenVZ 或 Virtuozzo for Windows

- 用于测试环境

§ Xen + 其它虚拟化技术

- 最好有Intel VT 或 AMD-pacifica的支持

§ OpenSolaris xVM(Xen)

- 可实现Zone + Linux + Windows的虚拟化



对数据中心建设的作用

- § 提高服务器利用率，减少数量
 - 刀片 vs 多核 + 虚拟化
- § 虚拟机与硬件无关，方便安装、迁移
- § 重启时间短，提高了服务器的可用性

0011 0010 1010 1101 0001 0100 1011

12
45



需要注意的问题

§ 谨慎增加，不要滥用

§ 虚拟机的安全

- 虚拟机的主机安全
- 虚拟机服务程序的安全

§ 还没有统一的管理界面

- 传统的管理软件看不到虚拟机
- 虚拟机管理软件不管理应用
- 不同虚拟机有不同的管理方法

§ 内存与与存储的配合

§ 支持虚拟化的CPU的功能利用

§ 虚拟化技术有不适用的应用，如：网络抓包



谢 谢!

0011 0010 1010 1101 0001 0100 1011

12
45