

武神

新巅峰 新势力 新网游

网络游戏基础系统架构与监控策略

武神世纪网络技术有限公司

曹世军

caoshijun@gmail.com

2009.8.29

提纲

- 游戏服务器架构需求
- 游戏服务器架构介绍
- 监控策略及实例
- 经验点滴

游戏服务器架构需求

- 苛刻的数据安全及系统安全要求
- 超乎寻常的网络延时敏感性及南北互联问题
- 故障硬件的快速替换及远程控制的可靠性
- 快速部署能力以满足游戏产品的爆发性增长导致的资源需求
- 快速且简单的实现大批量服务器的程序版本升级，以减少例行维护导致的服务不可用
- 快速且简单的合并、分散压力以应对频繁的合服、劈服操作
- 自动化运维实现的方便性

游戏服务器架构介绍

- 帐号数据中心
- 游戏服务器组
- 游戏辅助服务器组
- 官方网站及论坛
- 游戏服务器操作系统
- 更深入的理解你的硬件

帐号数据中心

- 存储游戏客户的帐号密码、点卡信息、充值记录、密码保护等重要信息
- 目前一般都是多款游戏产品共用一个数据中心，所有游戏产品的登录验证都在帐号数据中心进行，充值金额也可以在多款游戏产品中使用。
- 一般部署在可靠性较高的多线机房
- 爆发性访问比较常见，日常负载不高
- 游戏企业的命脉，安全要求最大化

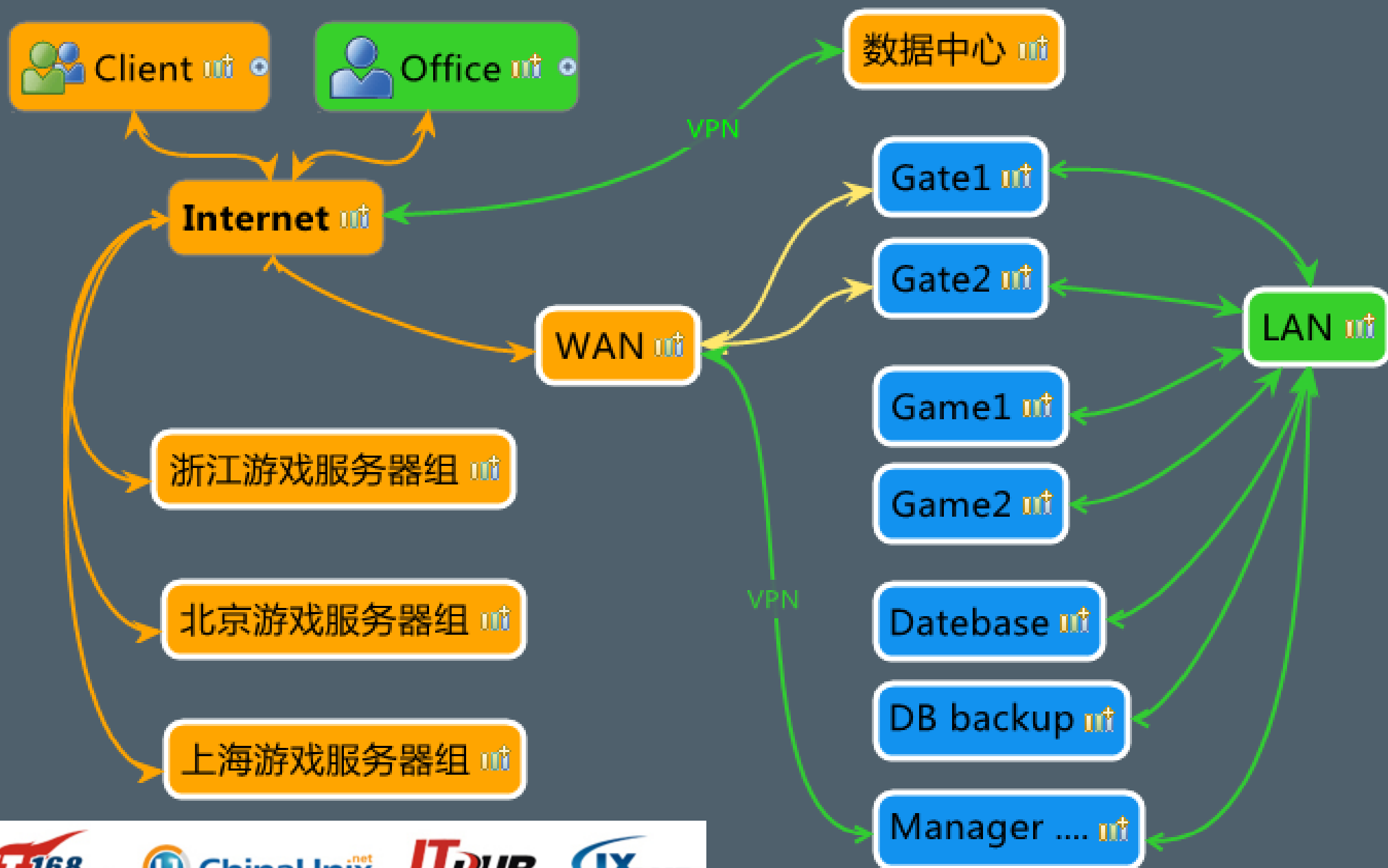
帐号数据中心示意图



游戏服务器组

- 以组分布，各组之间架构一致，便于横向扩展及管理，所有的游戏运算都在此进行。
- 一般由接入服务器、逻辑运算服务器、游戏数据库、日志服务器和其他特殊服务器构成。
- 对网络延时非常敏感，一般采用分布式部署，让服务器就近某一地区的客户，最大化减小客户端到服务器的网络延时。
- 分布式部署也解决了南北互联的问题。

游戏服务器架构示意图



游戏辅助服务器

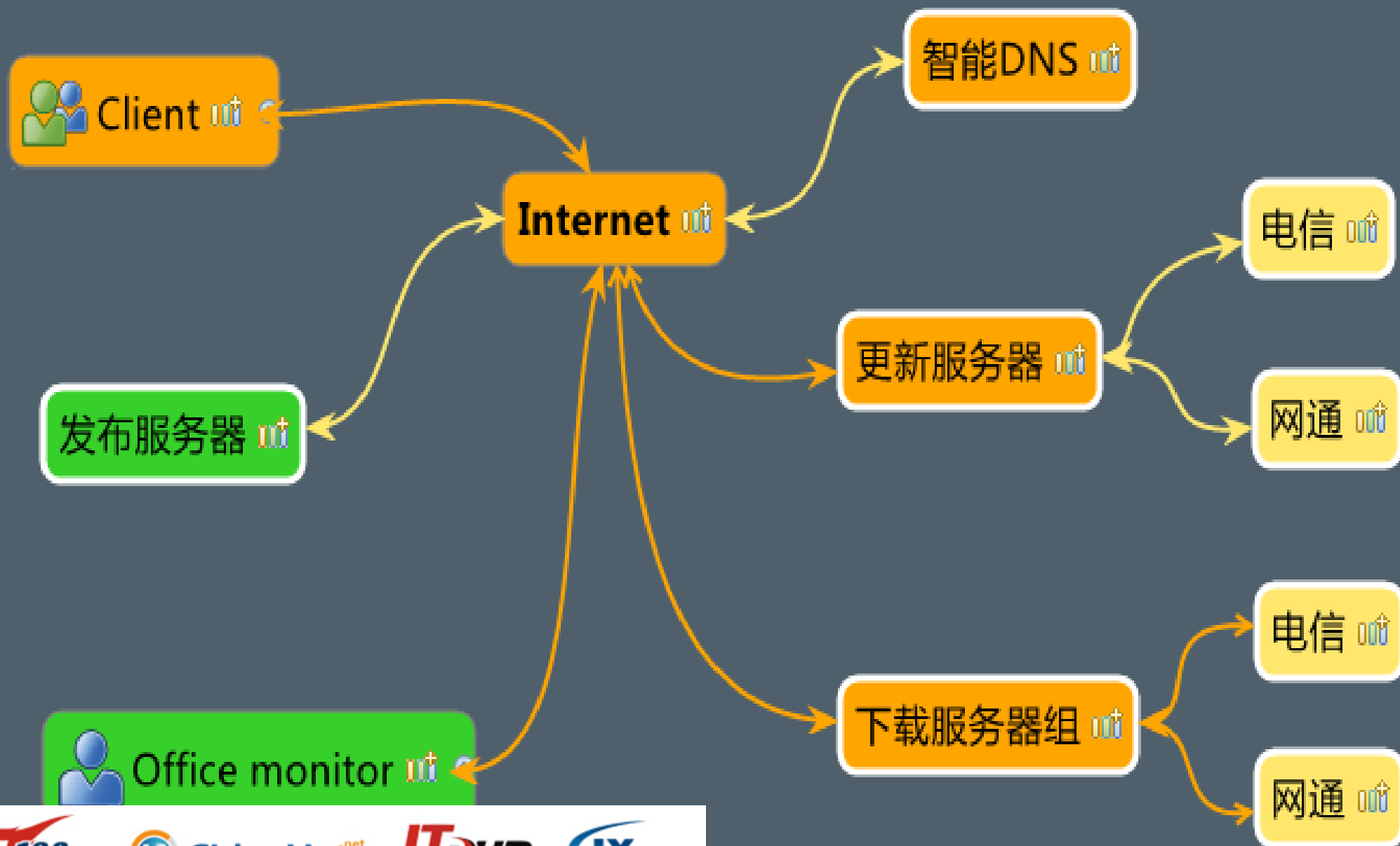
- 下载服务器

- 提供游戏客户端，客户端补丁升级包的下载
- 对网络的延时不敏感，爆发性流量较大
- 联通，电信分开部署

- 自动更新服务器

- 提供游戏客户端的自动升级
- 对网络的延时不敏感，爆发性流量较大
- 联通，电信分开部署

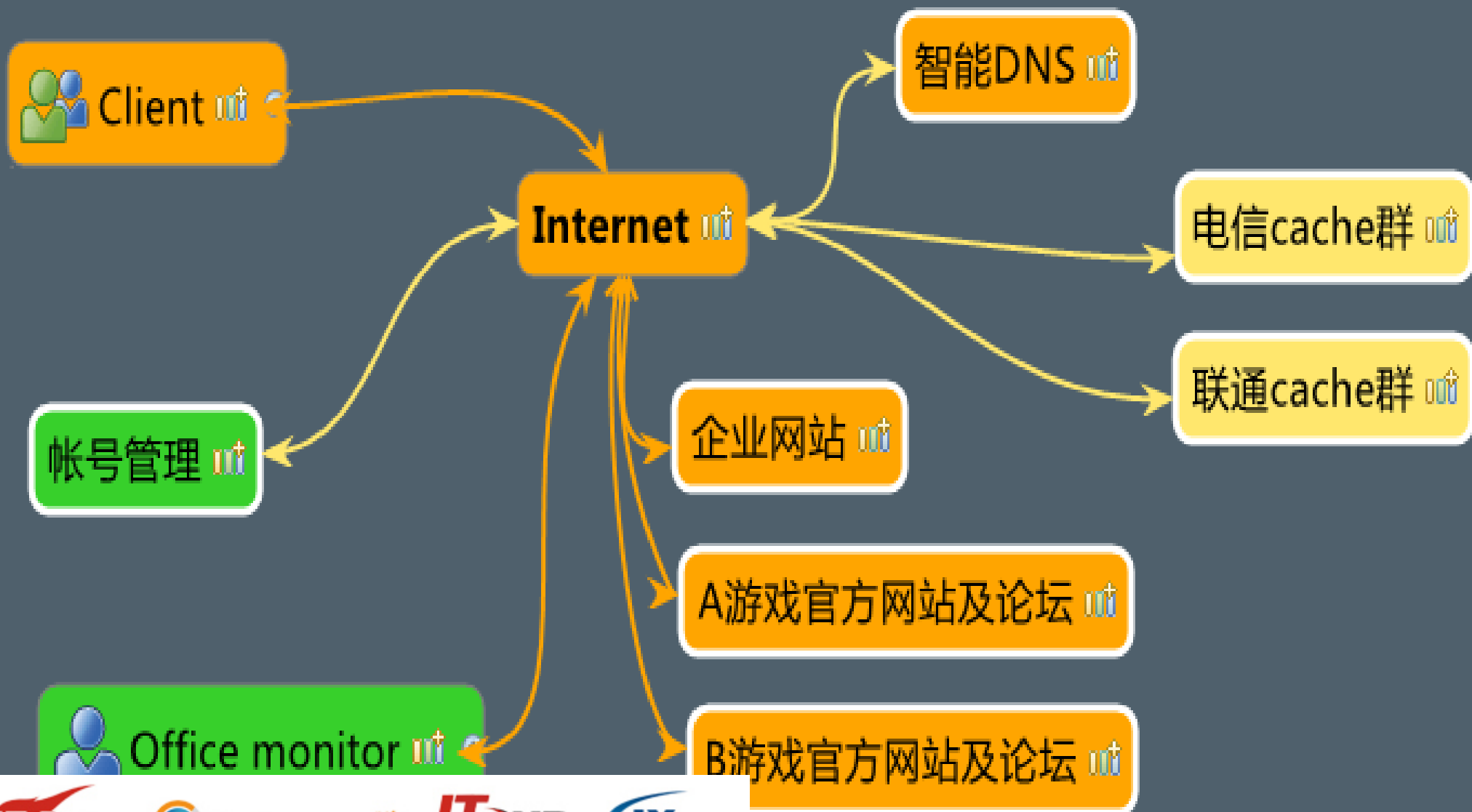
游戏辅助服务器示意图



官方网站及论坛

- 官方网站包含企业网站、游戏产品网站、帐号管理中心。
- 企业网站一般都是用作企业形象宣传。
- 游戏产品网站及对应的论坛一般都是针对某一款游戏产品做宣传。
- 帐号管理中心提供各个游戏产品用户的帐号注册、秘保设定、密码找回等。
- 该类应用在游戏停机维护期间突发流量大。

游戏官方网站及论坛



游戏服务器操作系统

- 现实世界的实际情况
 - 幸福的人都是相同的，不幸的人各有各的不幸
- 没错，最好就是用 Linux!
 - 和其他 Unix like 相比人才更好找
 - 和其他 Unix like 相比驱动程序更好找
 - 和 windows 相比更灵活，考虑下在 Windows 上自动化运维的代价。
- 可以的话用 64 位吧
- 实在不行的话 Windows 也不是不可以
- 我们是自己定制的 Debian X86_64 无盘系统

更深入的理解你的硬件

- 内存是新一代的磁盘
- 磁盘是新一代的磁带
- 无处不在的要命缓存
- 更快的 CPU 解决了很多程序语言的效率问题

更深入的理解你的硬件 - 带外管理

- 带外管理的历史
- BMC/IPMI
- IDC 机房能确保的资源：电源、网络
- 系统磐机，操作系统失去响应，服务器需要冷重启的时候只能打电话由机房操作吗？
- 调整机房里服务器的 BIOS 非要出差到机房所在地去吗？
- 重装服务器的操作系统必须由机房实施或者出差到机房所在地吗？
- 硬件日志查询（机箱盖，温度，风扇转速）
- Linux 下 Console 重定向结合 IPMI 的神奇

文件(E) 编辑(E) 查看(V) 终端(T) 帮助(H)

Entering Setup
F10 = Utility Mode
F11 = Boot Menu
F12 = PXE Boot

Phoenix ROM BIOS PLUS Version 1.10 2.4.3
Copyright 1985-1988 Phoenix Technologies Ltd.
Copyright 1990-2008 Dell Inc.
All Rights Reserved

Dell System PowerEdge 1950
www.dell.com

Two 2.33 GHz Quad-core Processors, Bus Speed: 1333 MHz, L2 Cache: 2x6 MB

System Memory Size: 8.0 GB, System Memory Speed: 667 MHz

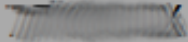
Dell SAS 6 Host Bus Adapter BIOS
MPTBIOS-6.22.03.00 (2008.08.06)
Copyright 2000-2008 LSI Corporation.

Initializing..|

caojun@caojun: ~/mp3

文件(E) 编辑(E) 查看(V) 终端(T) 帮助(H)

Dell Inc. (www.dell.com) - PowerEdge 1950
BIOS Version 2.4.3

Service Tag: 

Asset Tag:

System Time 16:22:31
System Date Wed Aug 26, 2009

Memory Information <ENTER>
CPU Information <ENTER>

SATA Port A Off

Boot Sequence <ENTER>
Boot Sequence Retry Enabled

Integrated Devices <ENTER>
PCI IRQ Assignment <ENTER>

Serial Communication <ENTER>
Embedded Server Management <ENTER>

caojun@caojun: ~/mp3

文件(E) 编辑(E) 查看(V) 终端(T) 帮助(H)

GRUB Loading stage1.5.

GRUB loading, please wait...

Press `ESC' to enter the menu... 0



文件(E) 编辑(E) 查看(V) 终端(T) 帮助(H)

```
pci 0000:03:00.0: PME# supported from D3hot D3cold
pci 0000:03:00.0: PME# disabled
pci 0000:00:1e.0: transparent bridge
ACPI: PCI Interrupt Link [LK00] (IRQs 3 4 5 6 7 10 *11 12)
ACPI: PCI Interrupt Link [LK01] (IRQs 3 4 5 6 7 10 11 12) *0, disabled.
ACPI: PCI Interrupt Link [LK02] (IRQs 3 4 5 6 7 10 11 12) *0, disabled.
ACPI: PCI Interrupt Link [LK03] (IRQs 3 4 *5 6 7 10 11 12)
ACPI: PCI Interrupt Link [LK04] (IRQs 3 4 5 *6 7 10 11 12)
ACPI: PCI Interrupt Link [LK05] (IRQs 3 4 5 6 7 *10 11 12)
ACPI: PCI Interrupt Link [LK06] (IRQs 3 4 5 6 7 10 11 12) *0, disabled.
ACPI: PCI Interrupt Link [LK07] (IRQs 3 4 5 6 7 10 11 12) *0, disabled.
SCSI subsystem initialized
usbcore: registered new interface driver usbfs
usbcore: registered new interface driver hub
usbcore: registered new device driver usb
PCI: Using ACPI for IRQ routing
hpet0: at MMIO 0xfed00000, IRQs 2, 8, 0
hpet0: 3 comparators, 64-bit 14.318180 MHz counter
pnp: PnP ACPI init
ACPI: bus type pnp registered
pnp: PnP ACPI: found 12 devices
ACPI: ACPI bus type pnp unregistered
system 00:08: ioport range 0x800-0x87f has been reserved
system 00:08: ioport range 0x800-0x87f has been reserved
```

游戏企业的监控目标

- 报警
 - 及时准确的发现并初步判断问题
 - 实现迅速可靠的多途径报警
 - SMS
 - IM
 - E-Mail
- 趋势分析及预警
 - 网络流量
 - 系统性能 (cpu,mem,load,disk,etc)
 - 由历史数据分析出未来将要面临的问题
 - 由历史数据分析出配置变更的效果

监控策略

- 明确需求并分清主次
 - 要达到什么目标？要监控何种内容？
 - 按照被监控对象在企业应用中的重要性划分主次
- 选用合适的软件
 - 监控目标和内容决定了你应该采用何种监控软件
- 选择合适的架构
 - 选择适合自己的监控平台架构，分布式 or 集中式
- 逐渐添加监控目标
- 养成每天看图的习惯

实例 -1

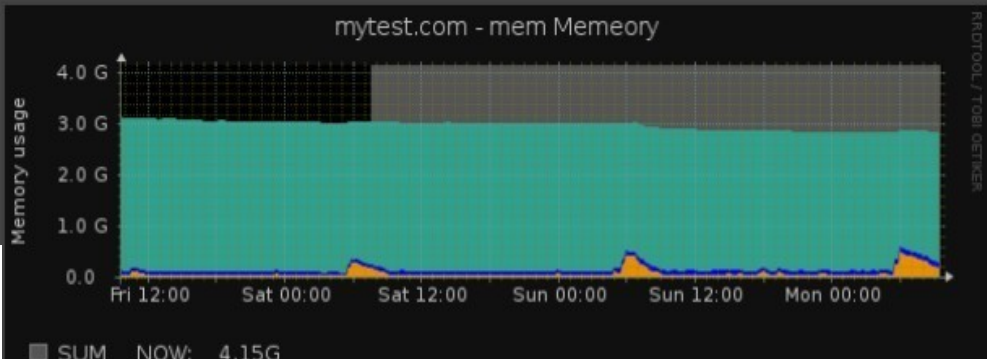
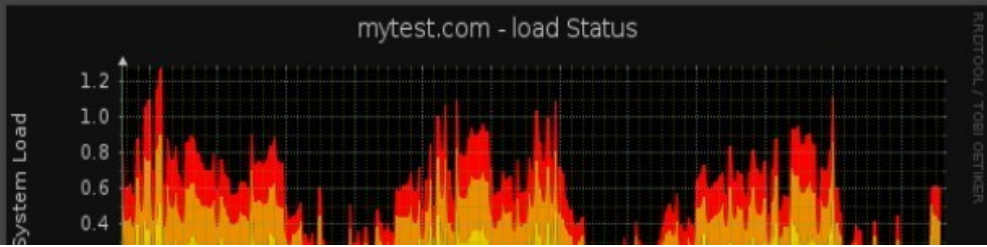
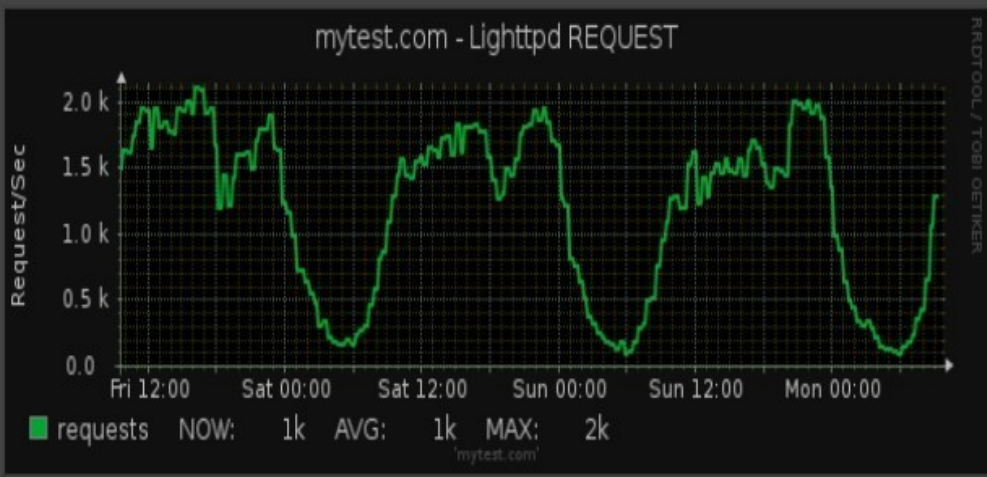
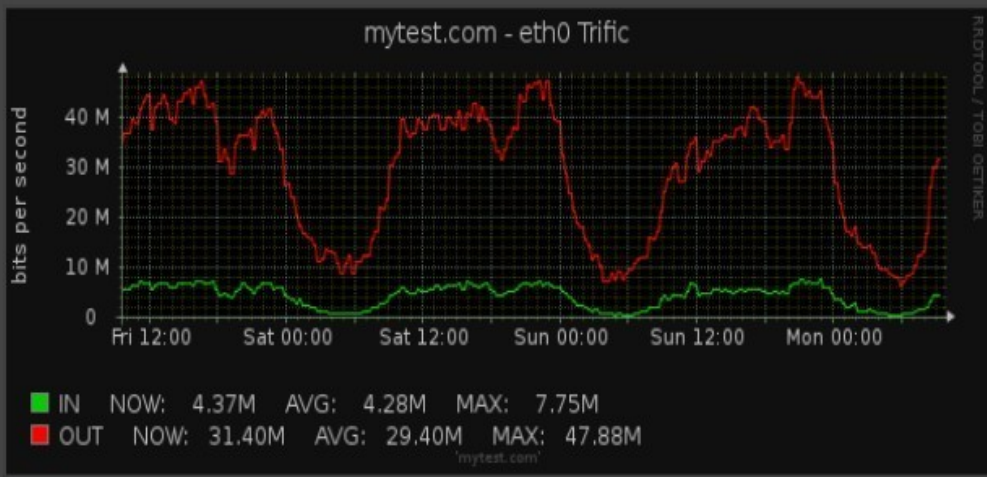
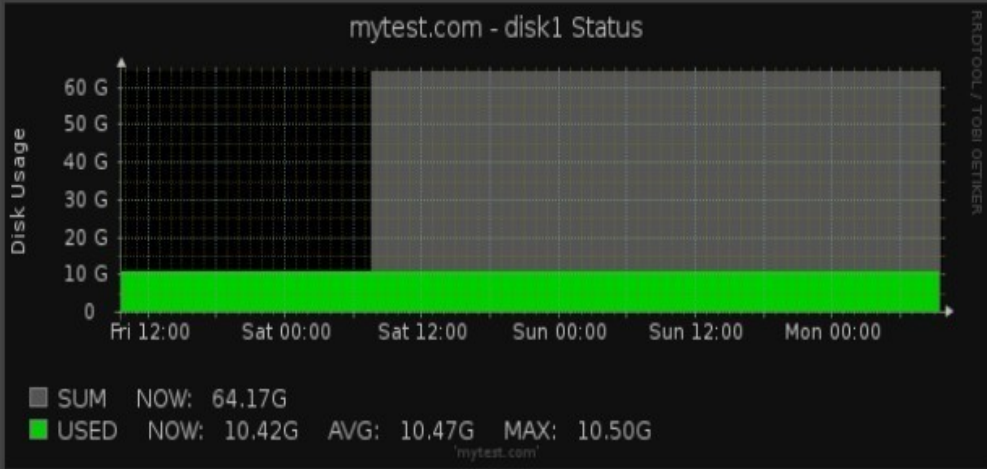
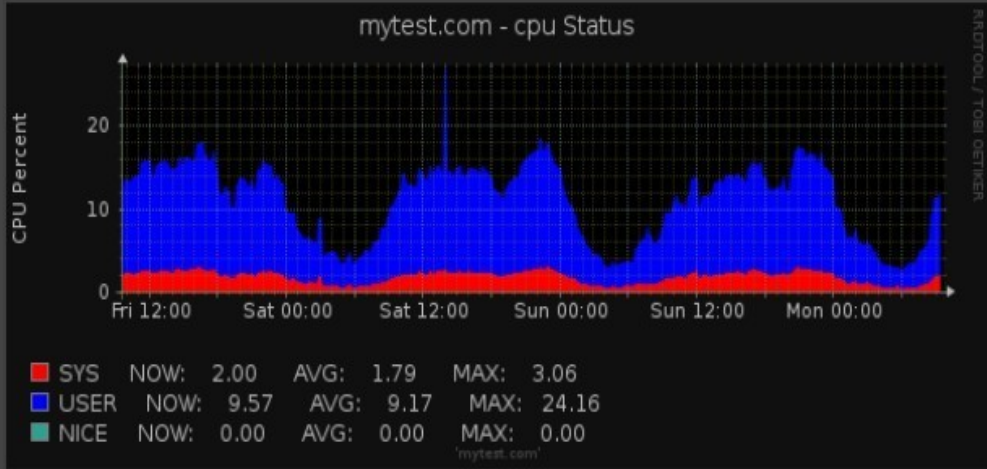
- 监控目标：
 - 游戏各类应用状态，系统状态，网络状态
- 划分等级：
 - 数据中心 > 游戏服务器 > 官方网站及论坛 > 游戏辅助服务器
- 软件使用：
 - Nagios 用于服务状态监测及报警，是自动化运维的基础，日常使用侧重于报警。
 - Cacti 用于数据记录及历史数据展现，是资源调度及数据分析的基础，日常使用侧重于数据记录。
 - 用 RRDTOOL 结合 Shell 实现了一套基础监控软件

实例 -2

- 监控架构的选择
 - 分布式和集中式两套监控系统
 - Cacti 和 Nagios 集中部署在公司内部机房，用于各个服务器的状态监测、报警及数据记录。
 - 每组游戏服务器都有一套自己写的组内服务器监控系统，部署在全国各地的游戏服务器上，用于网络不可达造成的集中监控失效的情况下，可以在恢复后通过分布式监控工具回查历史记录。
- 其他
 - 系统日志通过 syslog 汇聚存放，集中审查，重要消息多途径报警 Email,IM,SMS etc. 。

从2009-08-28 09:31:54到2009-08-31 09:31:54

mytest.com



经验点滴

- 一开始要简单
- 简单的同时要考虑横向扩展
- 以上两个实现自动化管理的基础
- 善用开源软件，不重复发明轮子，打磨轮子
- 重视标准
- 考虑硬件的可管理性
- 提高交流效率 (wiki)
- 在不断满足应用需求的同时遵循以上几点

参考资料

- “懒惰” 集群管理员的 11 个秘诀
 - <http://www.ibm.com/developerworks/cn/linux/l-11sysadtips/>
- 2009.3.28 CU 沙龙 北京 田逸 《开源监控利器 nagios 实战全解.rar》
 - <http://linux.chinaunix.net/bbs/attachment.php?aid=230127>
- 2009.3.28 CU 沙龙 北京 《使用 cacti 搭建企业监控平台》
 - <http://linux.chinaunix.net/bbs/attachment.php?aid=230128>
- 2007.12.8 CU 沙龙 西安 Cacti 在 Linux 上的部署和使用》
 - http://linux.chinaunix.net/salon/200712/files/cacti_xa.pdf
- 2007.12.8 CU 沙龙 西安 孙继东 《rrd 数据库及 rrdtool 简介》
 - http://linux.chinaunix.net/salon/200712/files/RRD_RRDTool_xa.pdf
- 2007.12.8 CU 沙龙 西安 李平 《nagios 在企业中的应用》
 - http://linux.chinaunix.net/salon/200712/files/nagios_xa.ppt
- 2009.3.28 CU 沙龙 北京 曹金城 《mysql 运行监控全解》
 - <http://linux.chinaunix.net/bbs/attachment.php?aid=230129>

Thanks!

Q & A