

为邮件系统选择合适的硬件，不仅有助于软件安全可靠的运行，更能够为企业节省资金，降低系统长期的总体拥有成本。

免责声明

- ◆ 本讲座上的任何言论仅代表个人观点，讲师并不提供任何形式的担保。讲座中所提到的第三方产品、设备和硬件供应商仅仅是为了演示和举例的目的，讲师无法完全验证所有信息的可靠和真实性，也不代表微软公司推荐或鼓励听众采用讲座中所提到的任何第三方产品。
- ◆ 硬件使用的环境错综复杂，讲师无法保证同样的硬件设备可以在听众的企业中达到课程中标称的性能指标和投资回报。对任何根据本讲座所进行的硬件采购和投资引起的负面后果，讲师和微软公司不承担任何责任。

硬件选型策略

- ◆ **首要考虑的不是性价比！**
- ◆ 考虑TCO(包括后期运营和维护的成本)和长期的投资回报
- ◆ 可管理性，易于维护
- ◆ 是否满足工业标准？
 - Designed for Windows
 - <http://www.microsoft.com/whdc/winlogo/default.mspix>
- ◆ 供货商的售后服务和故障维修(备件替换周期)
- ◆ 整个公司尽量使用统一的标准采购硬件

具体需要考虑的硬件配置

- ◆ CPU ✓
 - AMD Opteron
 - Intel Xeon with EM64 Technology
- ◆ 远程控制卡
- ◆ TOE，双网卡和Teaming
- ◆ x86, x64, IA64
- ◆ 砸锅卖铁也要上RAID,拒绝IDE，SATA系统
- ◆ 苗条的服务器受欢迎
- ◆ 阵列卡
- ◆ 磁盘柜
- ◆ 不要忽视网线

网卡支持TOE，可以大幅度降低CPU的负载，把一些流量在硬件层面来完成。

- ◆ 根据服务器承担的角色，对硬件有不同的要求
 - 后端的Store Server
 - 前端服务器，OWA
 - SMTP网关
 - 杀毒网关
- ◆ 可以考虑使用虚拟机
- ◆ How to choose server hardware for Exchange 2003 that can be effectively re-utilized for Exchange 12?
- ◆ <http://blogs.technet.com/exchange/archive/2006/03/13/421914.aspx> 

邮箱服务器主要要求高内存，高I/O磁盘。

前端服务器主要对CPU的负载比较大。

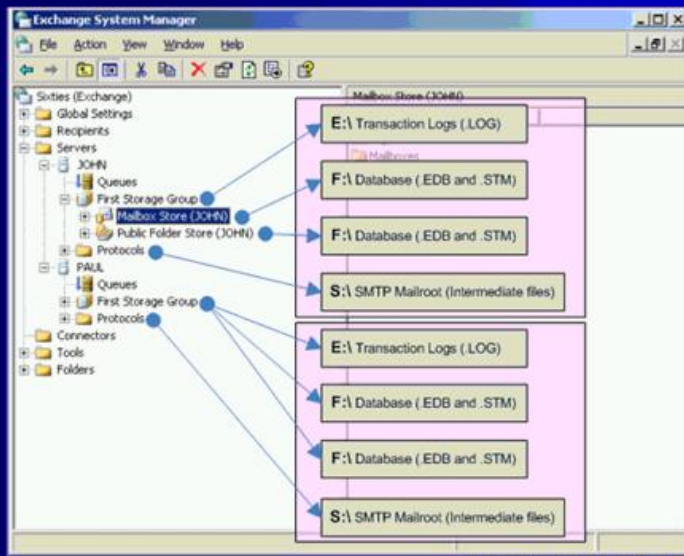
SMTP网关对CPU的处理能力，网络的吞吐能力要求比较高。

一台典型邮件服务器硬件的报价单

A	B	C	D	E	F	G
Product Number	Description	Qty	Unit Price (RMB)	Unit Price (RMB)	Total Amount (RMB)	TAT*
HP DL380G4						
379696-AA1	DL380R04 X3.0/800-1M PRC	1	19615.38	22950	22950.00	22WDS
380040-B21	X3.0/800-1M 370/380 G4 ALL	1	4230.77	4950	4950.00	22WDS
343055-B21	1GB REG PC2-3200 2x512 ALL RAM	1	3350.43	3920	3920.00	22WDS
343056-B21	2GB REG PC2-3200 2x1GB ALL RAM	1	7923.08	9270	9270.00	22WDS
286778-B22	72GB Universal Ultra 320 15K 1" HDD	3	3760.68	4400	13200.00	22WDS
346914-B21	SA6i upgrade kit	1	1367.52	1600	1600.00	22WDS
364507-B21	DL380G4 Floppy	1	376.07	440	440.00	22WDS
355892-B21	DL380G4 Power	1	1256.41	1470	1470.00	22WDS
263825-B21	iLO Remote control software kit	1	1025.64	1200	1200.00	22WDS
293048-B21	DL 380G3/G4 Fans	1	512.82	600	600.00	22WDS
U4545A	3 years support, 24*7	1	1376.07	1610	1610.00	22WDS
Sub Total:					61210.00	
Options						
273915-B21	Smart Array 6402/128 ALL	2	4188.03	4900	9800.00	22WDS
273913-B21	256MB Upgrade-SA6402 ALL	2	1794.87	2100	4200.00	22WDS
302969-B21	MSA30 S-B ALL	2	16188.03	18940	37880.00	22WDS
371535-B21	146G 10K HP U320 ML1XX ALL	14	3760.68	4400	61600.00	22WDS
286776-B22	36GB 15K U320 UNI HDD ALL	8	2008.55	2350	18800.00	22WDS
C9572CB	HP StorageWorks Ultrium Tape Autoldr	1	42735.04	50000	50000.00	22WDS
C7971A	HP ultrium 200GB data cartridge	50	341.88	400	20000.00	22WDS
Sub Total:					202280.00	
Total Amount(RMB) for the above products and					263490.00	
Total Amount(USD) for the above products and					31745.78	

存储系统占用exchange硬件投资的50%强，对存储的规划和设计决定了性能。

System Admin Viewpoint



虚拟化应用的场景：1) 服务器的CPU和内存配置非常高，跑一个应用有点浪费；2) 对于一些负载不高或者不是关键的角色可以考虑虚拟化。

网卡teaming有很小的概率会导致exchange出现网络问题。

RAID Levels and Arithmetic

- ◆ **RAID 0 (2 drives or more)**
 - $READS + WRITES = I/Os / sec$
- ◆ **RAID 1 (2 drives)**
 - $READS + (2 \times WRITES) = I/Os / sec$
- ◆ **RAID 5 (3 drives or more)**
 - $READS + (4 \times WRITES) = I/Os / sec$
- ◆ **RAID10 = RAID1 with more than 2 drives**
- ◆ **In reality, cache-centric arrays and virtualization break the above rules**
 - Using arithmetic will lead you to over-engineering: can you afford it?

IOPS计算实例

“I need 800 I/O per sec (70% read)”

- ◆ **RAID0**
 - 800 disk I/O
 - 140 I/O per disk → 6 disks
- ◆ **RAID1**
 - $800 \times 0.7 + 2 \times (800 \times (1-0.7)) = 1,040 \rightarrow 8 \text{ disks}$
- ◆ **RAID5**
 - $800 \times 0.7 + 4 \times (800 \times (1-0.7)) = 1,520 \rightarrow 11 \text{ disks}$
- ◆ **In reality**
 - Virtualization will “cloud” your results because your disks might be busy elsewhere
 - Cache will “cloud” your results because I/O is decoupled → writes are cached, possibly aggregated, resulting in lower effective IOPS

上面的公式是比较粗糙的计算，实际情况还要考虑到用户数量。

Best (?) Practices

RAID Levels

- ◆ **R1 vs. R5**
 - R1 is more efficient than R5 for IOPS/\$
 - R5 is more efficient than R1 for GB/\$
- ◆ **R1 vs. R10**
 - R10 is stripping of mirrors (many R1's together) – used for databases
- ◆ **R1 for logs**
 - Dedicated disks for logs is “one” way of achieving the real intent: FAST WRITES
- ◆ **Cache memory**

使用高性能的磁盘

- ◆ **Small (36 GB-72 GB) disks**
- ◆ **SCSI**
 - FC or SCSI electrical media
- ◆ **Fast rotation speed (15krpm)**
- ◆ **Separate bus/paths**
 - Enable parallel transfers
- ◆ **Do not use**
 - **SATA = smaller duty time, smaller MTBF**
 - **High capacity to make up the required volume (e.g., 300 GB or 400 GB)**

考虑什么样的备份设备

磁带机备份的速度

推荐以存储组为单位进行备份，一个存储组最好备份到同一个磁带，方便恢复。

- ◆ **The network is slow for modern tape libraries**
 - Some are > 30 MB/s (100 GB/hr for a single tape)
- ◆ **Sustain data rate is required for best speed**
 - A buffer underrun causes tape rewind which can significantly reduce your tape throughput
- ◆ **Use large I/O**
 - NTBACKUP has buffer tuning that improves throughput
- ◆ **Use SAN libraries (tape or virtual) or array-to-media technologies (NDMP or similar)**

数据恢复考虑事项

数据恢复和日志回滚

- ◆ **Recovery is about:**
 - Restore from media
 - Log playback
- ◆ **Ensure:**
 - Fast restore from media (why should it be half of the backup time?)
 - Fast log playback
 - Increase transaction log buffers to 9,000 for playback of many logs (ESE tuning)
 - Log playback can take hours

硬盘首要考虑的是IOPS的要求。在满足IOPS要求的前提下，再去考虑选择大硬盘还是小硬盘，同时考虑预算的多少。

如果使用的是硬件的阵列卡，那在进行磁盘RAID的扩容的时候，基本是不需要停机的。

做一个称职的Exchange专家

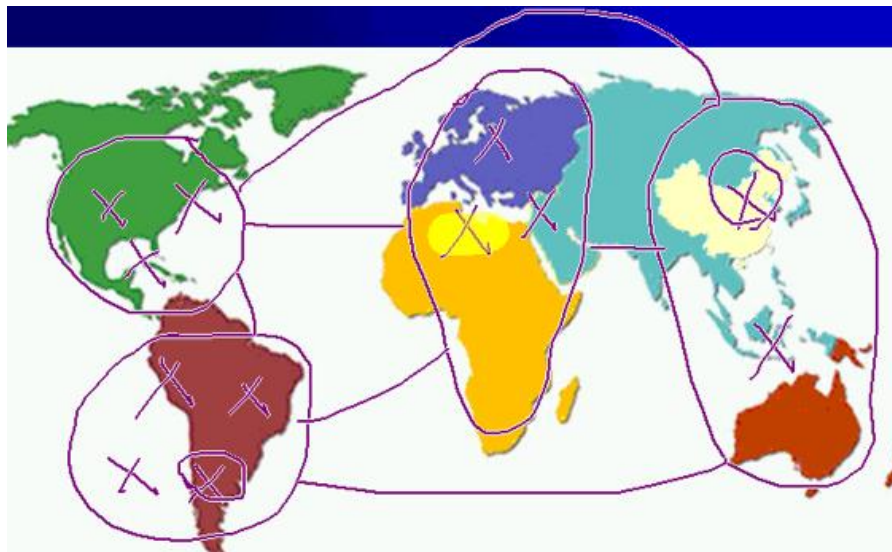
- ◆ 精通软硬件
- ◆ 精通活动目录
- ◆ 紧跟最近技术和潮流
- ◆ **Business Sense**
- ◆ 能够将技术转化为企业的竞争力
- ◆ **128道微软Exchange考题在网站可下载**
 - www.yuyong.net
 - 考考你自己

路由组的设计

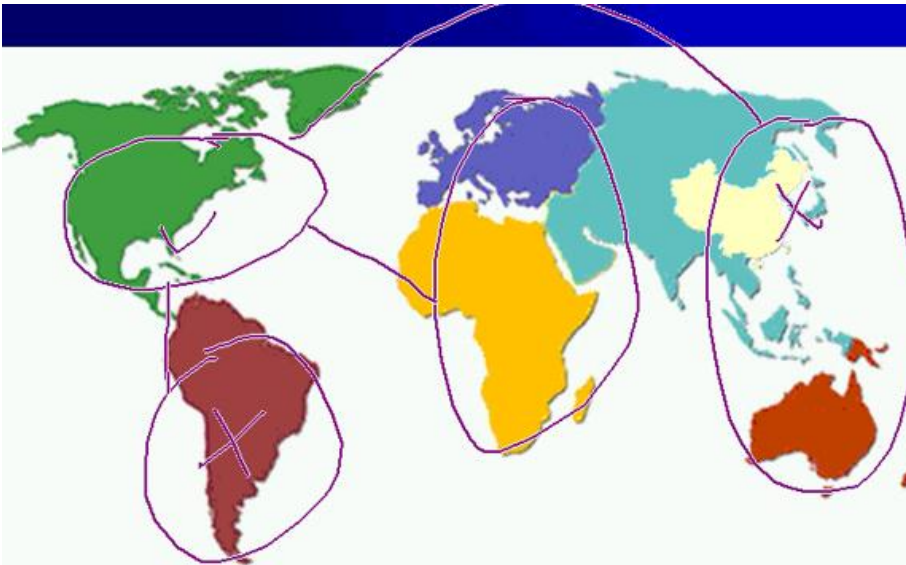
- ◆ 什么时候需要使用路由组？
 - 服务器之间的网络连接慢(小于128Kbps)，或者不稳定
 - 需要在服务器或者不同物理位置之间安排邮件传输计划
 - 管理或者业务上的需要
 - 控制公共文件夹的访问
- ◆ 路由拓扑的设计要点
 - 简单，为了避免路由组过于复杂，必要的时候升级网络
 - 反映网络的带宽情况
 - 集中式(星型拓扑)
 - 分散式(网状拓扑)
 - 避免过多的路由组，管理成本高
 - 与活动目录的Site相吻合
- ◆ 选择什么类型的路由组连接器？
 - Routing Group Connector是首选
- ◆ 如何设计冗余的路由组结构？
 - 多个桥头堡服务器
 - 尽量做到同样的cost
- ◆ 桥头堡服务器的选择
- ◆ 路由组之间的一些控制
- ◆ 参考读物
 - Exchange传输组件大揭秘(上/中/下)
 - 规划 Exchange Server 2003 邮件系统

路由组太多的话，发送接收邮件的时候会在服务器之间经过特别多的路由，所以要控制路由组的数量。圈定路由组的范围。建议与活动目录的SITE吻合。

全网状的连接方式，通过路由组连接器。

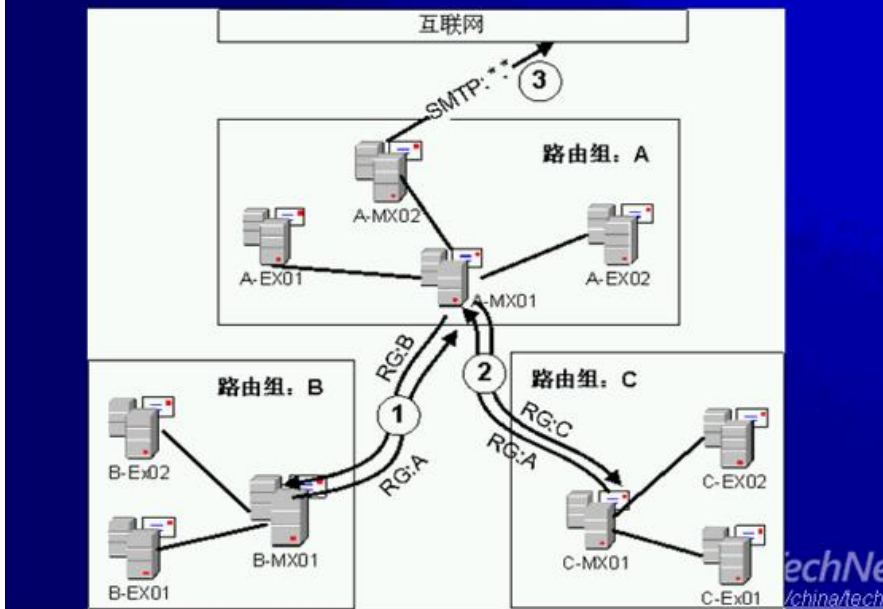


通过中枢节点进行连接，优点是路由组的数量比较少。一般出口会放在中心路由组。



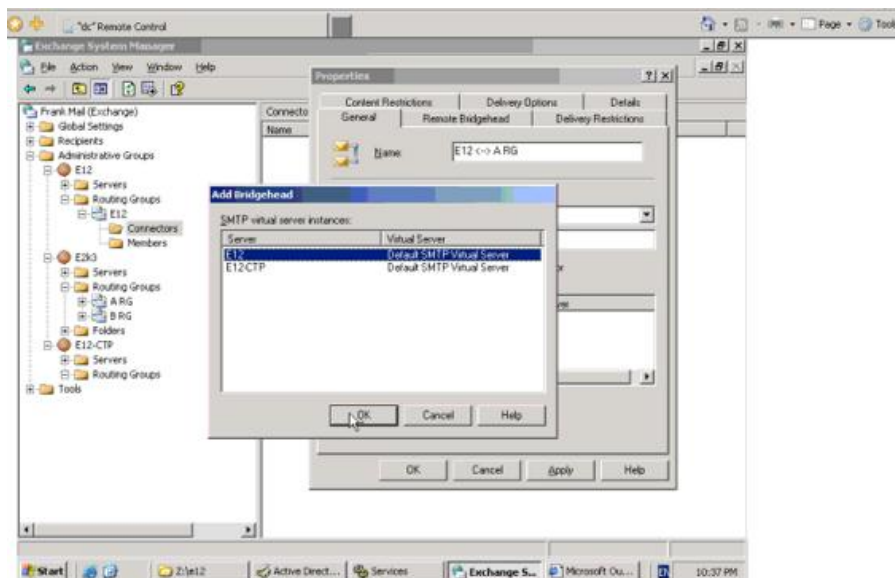
具体采用哪种形式，要看跨路由组的邮件是不是特别多。如果特别多，那可以用网状拓扑。如果是组织内部的连接，则建议使用路由组连接器。

路由的过程

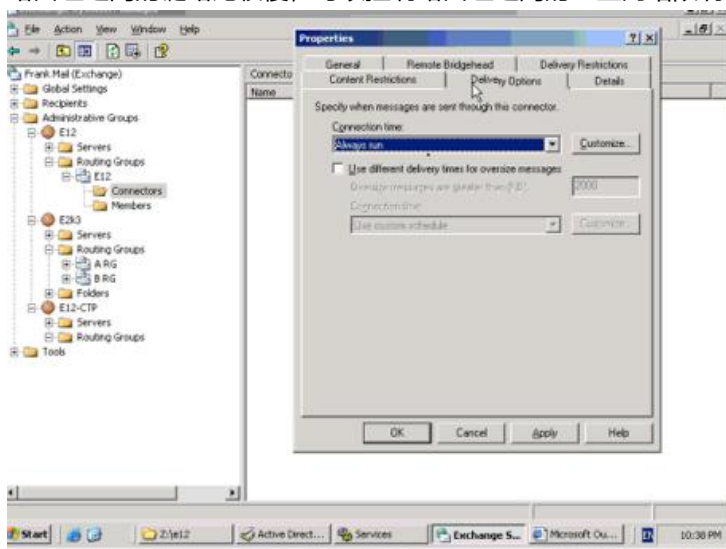


上图中B和C之间是没有连接的。

配置哪些服务器可以使用指定的连接器，如果不指定的话，则任何一个服务器都可以使用该连接器。

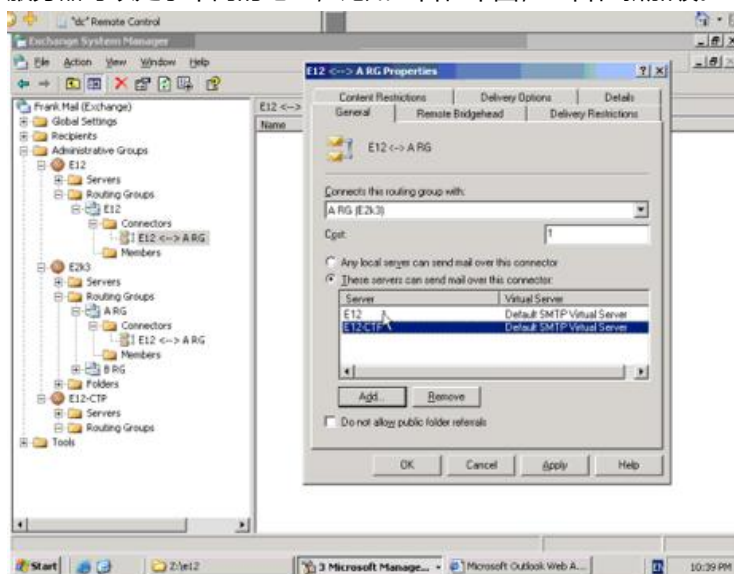


路由组之间的链路比较慢，可以控制路由组之间的一些网络限制。



连接器是双向的，我们在创建一个单向连接器的時候，会自动创建反向的连接器。

路由组之间的桥头堡服务器，建议至少指定两台，以避免某台出问题，一方面可以负载流量，另外一方面，这两台服务器可以处于不同的地区，比如一台在中国，一台在新加坡。



一个SMTP连接器可以指定多个桥头堡服务器，来进行外发的工作。很多企业只设置一个互联网站点的出口，可以统计管理，控制邮件的归档、外发、病毒检测等。

拓扑结构-SMTP网关的设计

- ◆ 逻辑上应该只有单一的网关
 - 一个负责外发的SMTP连接器
 - 为了冗余，这个连接器可以有多个桥头堡
- ◆ SMTP网关负责收发邮件，应该选择可靠的ISP链路，必要时选择双ISP或者异地站点进行冗余
 - 有些ISP的IP地址已经在SPAM黑名单上
 - 注意DNS MX记录和反向解析记录的注册
- ◆ 不要为了冗余或者防止链路故障而做多个SMTP连接器，这样增加管理成本，也增加出现故障的可能。
- ◆ 可以采用非Exchange服务器作SMTP网关
 - MiraPoint等硬件SMTP服务器
 - 速度快，维护成本低

拓扑结构-SMTP网关的设计

- ◆ 邮件内部发送和外部发送的区别
- ◆ 要非常清楚由于邮件收发对网络引起的负载
- ◆ SMTP虚拟服务器和SMTP连接器的配置
- ◆ 关于限制邮件收发的问题
 - 不允许特定用户发送邮件到外网
 - 不允许特定用户发送邮件到指定的domain
 - CheckConnectorRestrictions
 - Demo

exchange所有服务器都可以做桥头堡服务器，跟是不是前端没有关系。

前端服务器是用来连接客户端到后端数据库的建立访问。

前端服务器的访问时可以跨路由组的，外网的邮件先到前端服务器，要考虑前端服务器的带宽是不是够用。

前端服务器是要找后端服务器的，前端到后端的带宽也要够。

前端服务器一般会放在广域网所在的中心节点的位置。

前后端设计的要点

- ◆ 根据实际情况决定前端服务器的数量和位置
- ◆ 前端放在防火墙的非军事区，上面的角色尽可能的简单
- ◆ 控制防火墙的端口
- ◆ 不推荐使用前端服务器发布SMTP,POP3和IMAP4等服务
- ◆ 理想情况下，只对方提供443端口的加密HTTP访问
- ◆ 需要时，使用NLB做前端服务器的冗余

条件成熟的话，推荐集中式的设计。

后端服务器设计

- ◆ 集中还是分散？这是个问题！
- ◆ 集中式后端邮件服务器
 - 几个国家共享一套处在数据中心内的大型邮件服务器集群
 - 初期投入高，但是总体成本和可靠性高
 - 客户端全部访问远程服务器，对网络的可靠性要求高
 - 微软目前采用这种做法
 - 适用于网络结构健全，用户数量大的企业
 - 可以在实施之前先考虑网络的可靠性，比如全年的平均故障时间
- ◆ 分散式后端邮件服务器
 - 每个国家或者分支站点一个邮件服务器
 - 节省网络开销
 - 需要有额外的域控制器和全局编录服务器
 - 管理成本高
 - 总体可靠性下降
 - 冗余成本高

Microsoft Tech
<http://www.microsoft.com/china>

授权，进行成本控制

需要的软件License

- ◆ 前端服务器
 - Windows 2003 Std
 - Exchange 2003 Std
 - OS级别防病毒软件
- ◆ SMTP网关
 - Windows 2003 Std
 - Exchange 2003 Std
 - OS级别防病毒软件
 - 反垃圾邮件系统，SMTP过滤反病毒系统(可以是独立的服务器)
 - 或者独立的硬件SMTP系统
- ◆ 后端服务器
 - Windows 2003 Std或Ent
 - Exchange 2003 Std或Ent
 - 备份软件
 - OS级别防病毒软件
 - 邮件防病毒软件
- ◆ 如果使用SAN
 - SAN管理软件
 - 网络备份软件
- ◆ 监控系统(比如MOM)
- ◆ 活动目录服务器的OS和防病毒软件
- ◆ 前端服务器的SSL证书

Microsoft TechNet
<http://www.microsoft.com/china/tech>

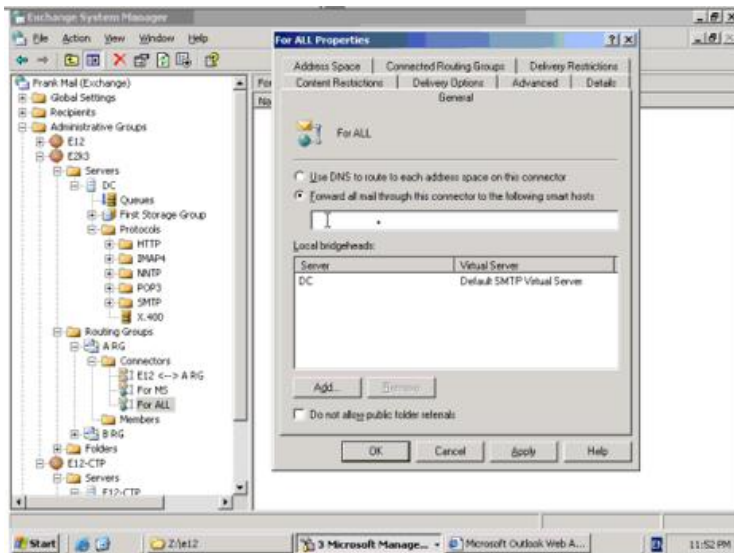
长痛不如短痛。保持架构的稳定，不要被非技术性的因素所左右。

邮件系统设计和选型的原则

- ◆ 简单就是美
- ◆ No dirty work
- ◆ 承上启下
 - 承上：符合现有网络和活动目录的架构
 - 启下：能够为今后邮件系统的监控、防病毒、反垃圾邮件、审计归档等等外围功能留有充分的余地
- ◆ 不大没准备的仗
- ◆ 长痛不如短痛
- ◆ 保持架构的稳定性，不要被非技术性的因素所左右

中小企业在路由组的设计上会比较简单，可能只需要一个路由组就可以了。

如果当前的外发邮件服务器的IP被放到了黑名单，那可以配置把邮件外发到另外一个IP。



06Exchange Server白手起家系列之六：如何执行压力测试？

压力测试是指在一定约束条件下测试系统所能承受的并发用户数量、运行时间、数据量，以确定系统所能承受的最大负载压力。

压力测试有助于确认被测系统是否能够支撑性能需求，以及预期的负载增长。

压力测试的目的和原则

- ◆ 进行有效的具有前瞻性的容量管理
- ◆ 在真实环境下检测系统性能，评估系统性能以及服务等级的满足情况
- ◆ 预见系统负载压力承受力,在应用实际部署之前，评估系统性能
- ◆ 分析系统瓶颈、优化系统
- ◆ 为压力测试设定基准线
- ◆ 尽可能在投入生产环境之前完成压力测试
- ◆ 测试的类别
 - 并发性能测试
 - 疲劳强度测试
 - 大数据量测试

Microsoft TechNet
http://www.microsoft.com/china/technet

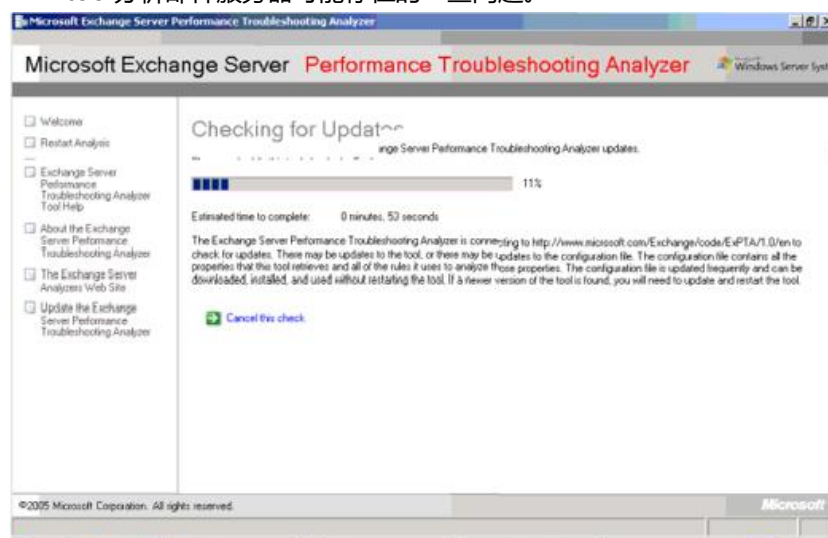
关键指标：IOPS、RPC latency、disk QUEUE Length

MSExchangeIS					
Active User Count	473				
RPC Averaged Latency	14				
RPC Operations/sec	33.003				
RPC Packets/sec	17.002				
RPC Requests	0				
RPC Requests Peak	27				
User Count	694				
Virus Scan Queue Length	0				
VM Largest Block Size	16392192				
VM Total 16-MB Free Blocks	0				
VM Total Free Blocks	479				
VM Total Large Free Block Bytes	0				
Physical Disk					
Current Disk Queue Length	2	0 E:	0 I:	0 F:	0 G:
Disk Reads/sec	370.034	0.000	0.000	0.000	0.000
Disk Transfers/sec	432.040	0.000	4.000	0.000	0.000
Disk Writes/sec	62.006	0.000	4.000	0.000	0.000
Process					
Virtual Bytes	293630616				
SMTP Server					
SMTP 1					
Cat: Address lookups/sec	0.000				
Cat: Categorizations completed/sec	0.000				
Cat: Categorizations failed (DS connection failure)	0				

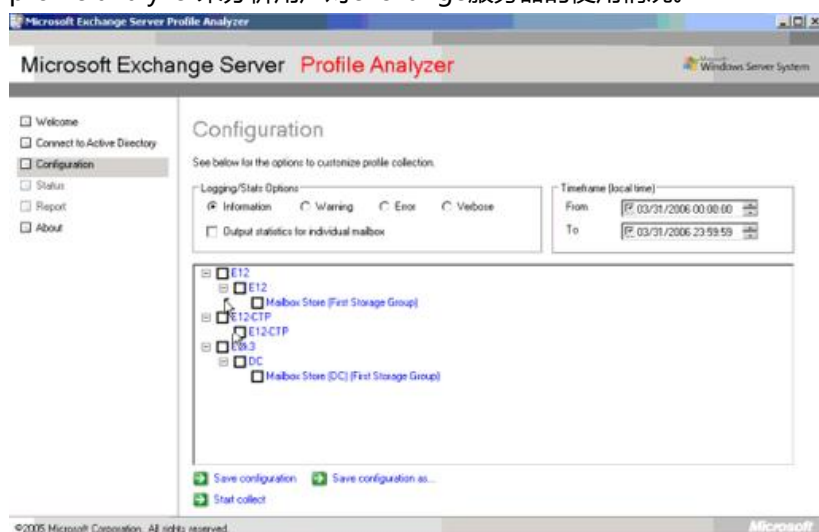
exchange server 2003相关的性能分析工具

- ◆ Performance Monitor
- ◆ Microsoft Exchange Server Performance Troubleshooting Analyzer Tool
- ◆ Microsoft Exchange Server Profile Analyzer
- ◆ Microsoft Exchange Server User Monitor
- ◆ Microsoft Exchange Server Jetstress Tool
- ◆ Microsoft Exchange Server 2003 Load Simulator (LoadSim)
- ◆ Exchange Server Stress and Performance 2003

PTA tool分析邮件服务器可能存在的一些问题。



profile analyzer来分析用户对exchange服务器的使用情况。



jetstress tool主要用来测试硬件的性能，模拟exchange的一些东西，而且不要求安装exchange服务器，即可以运行。比如测试当前服务器可以允许支持多少用户，可以模拟exchange的运行状态。

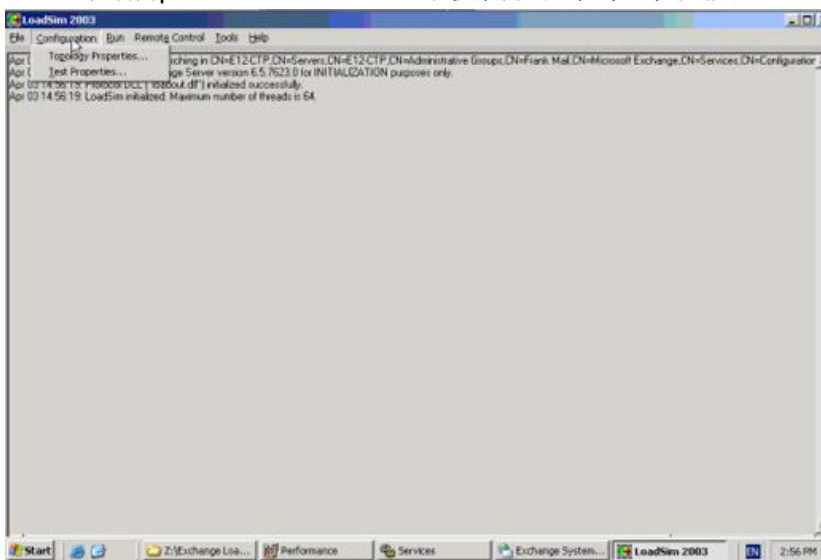
Microsoft Exchange Server Jetstress Tool

- ◆ <http://www.microsoft.com/downloads/details.aspx?familyid=94b9810b-670e-433a-b5ef-b47054595e9c&displaylang=en>
- ◆ Use Jetstress to verify the performance and stability of a disk subsystem prior to putting an Exchange server into production. Jetstress helps verify disk performance by simulating Exchange disk Input/Output (I/O) load. Specifically, Jetstress simulates the Exchange database and log file loads produced by a specific number of users. You use Performance Monitor, Event Viewer, and ESEUTIL in conjunction with Jetstress to verify that your disk subsystem meets or exceeds the performance criteria you establish. After a successful completion of the Jetstress Disk Performance and Stress Tests in a non-production environment, you will have ensured that your Exchange 2003 disk subsystem is adequately sized (in terms of performance criteria you establish) for the user count and user profiles you have established. It is highly recommended that the Jetstress user read through the tool documentation before using the tool.
- ◆ Jetstress is only supported when running with Exchange 2000 or Exchange 2003 or newer versions of ESE.DLL. Also, due to this support limitation, Jetstress is only supported on Windows 2000 Server, Windows Server 2003, Advanced Server, Datacenter and newer Windows OS platforms (Windows NT 4.0 and older builds are not supported).



自带performance monitor的关键指标，不需要单独创建性能计数器，建议跑24小时。

=====
loadsim主要是测试软件瓶颈的。模拟存储组增加新用户的数量。还可以测试exchange server同AD之间的工作情况。可以结合performance monitor来收集测试过程中的性能情况。



ESP工具。结合了jetstress和loadsim的优点。可以测试POP3/owa等协议的访问请求。测试exchange server不同模块。

测试现有环境中已经存在的数据库服务器。

