Sizing Exchange Servers

Steve Tramack
Sr. Solutions Engineering Manager
HP

HP WORLD 2003
Solutions and Technology Conference & Expo

Objectives

- Make Exchange 2000 scream, baby!
- Understand performance basics
- Get the most out of your servers
- Understand key capacity planning and design principals for "max headroom"
- First Look: Exchange 2003 sizing and performance considerations

Topics

- Sizing Process
- Server Capacity Planning
- Exchange Server Sizing and Design
- Monitoring and Tuning
- Best practices

Why Are You *Really* Here?

"OK, so, how many users will this server support?"

"If 2P is good, and 4P is better, what about 32?"

"I have a limited budget; where should I focus?"

"With VSS support, should I finally consolidate?"

"Does Exchange 2003 support more users/server than Exchange 2000?"

What's Out There?

- Benchmarks
  - …Are not the best source of sizing data
- Why?
  - Single server
  - Unrealistic configuration
  - One workload
  - No additional software or redundancies
- Intended to compare server models, not provide sizing guidelines

What Else Is Available?

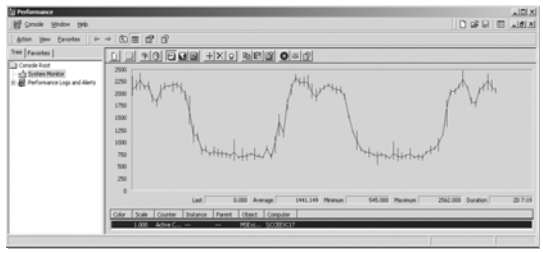
- Sizing tools
  - Varying levels of complexity and comprehensiveness
  - Risk of GIGO as solution grows
- White papers
  - Source reference materials
  - Based on lab testing or "notes from the field"
- Consultants

Where Do I Start?

- Know thyself
  - Identify your definition of a "user"
  - Understand workload characteristics
  - Identify best practices, business considerations, SLAs, planned architecture
- Establish baselines
- Gather data
  - Personalized benchmarking
  - Understand peaks versus percentiles
  - Pilot, if possible, and monitor

Usage Patterns

- Active client logins – 2 day period
  - MSexchangeIS Mailbox \ Active Client Logons



Topics

Sizing Process

Server Capacity Planning
- Bottlenecks
- Key Resources
- Exchange-specific concerns

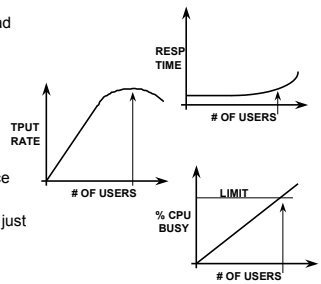
Exchange Server Sizing and Design

Monitoring and Tuning

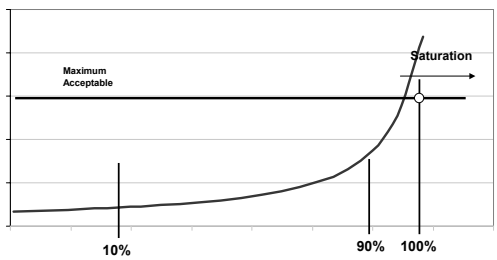
Best practices

Measuring Server Capacity

- Response Time
  - Server's reaction to load
  - Only part of the picture
- Throughput
  - Messages/sec
  - Transactions/sec
- What is a Bottleneck
  - High Demand Resource
  - Workload Dependent
  - Never really eliminate; just move



Finding Acceptable Response Times Benchmarking versus Characterization



Key System Resources

- CPU
- Memory
- I/O
  - Disks
  - Network
- Scale up versus Scale Out
- SLAs

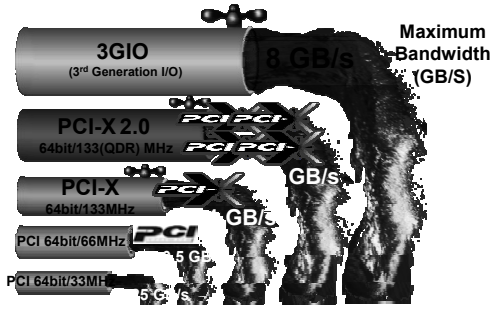
CPU

- Single and Multi-processors
  - Prestonia (Xeon DP)
  - Foster and Gallatin (Xeon MP)
- Communicate with other components by the means of a Front-Side Bus
  - Memory
  - PCI Bridges (I/O)
  - Other CPUs
- Chipset
- IA32 versus IPF versus Opteron

Memory

- Interleaving
- 200/266/333MHz DDR
- 32-bits physical address space = 4GB
- >4GB is addressed by
  - Hardware: Physical Address Extensions
  - Software: Address Windowing Extensions
- Who needs more than 4GB?!
  - Memory-bound applications
  - In-memory databases

I/O Bus Evolution



I/O: Disks

- Disks have improved
  - $/GB
  - GB/ft
  - MB/s
  - NOT I/O/s
- Still considered slow devices
  - 50-70ns to memory
  - 5-15ms to disks
  - Additional CPU overhead

I/O: Network

- 10-100Mb/s prevalent
- 1Gb/s
  - E2K backup over network
  - iSCSI
- Requires processing of the OSI stack
  - Some NICs will offload
- NIC teaming

Storage

- Capacity versus I/O
- Use Cache for low latency
  - Read versus Write
- Use RAID for high transaction rates
  - Impact of RAID levels
- Avoid saturation
  - Similar response between 10% and 90% of the workload

Sizing – No Silver Bullets

- User Workload
- Hardware/Software Configuration
- Software Stack
- Active versus Subscribed Users
- Client Access Protocols
- Architecture
  - Clustering
  - FE/BE

Topics

Sizing Process

Server Capacity Planning

Exchange Server Sizing and Design
- Server roles
- Back-end server sizing guidelines
- I/O planning
- Tools

Monitoring and Tuning

Best practices

Determine The Server Roles

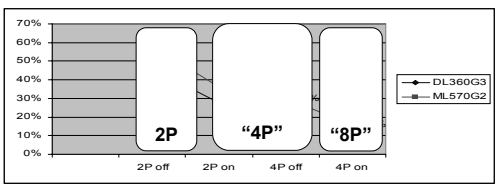
- Different requirements based on role
  - Mailbox Servers
  - Bridgehead Servers
  - Front End Servers
  - Active Directory/Global Catalog
  - Expansion Servers
  - Free/Busy
  - Public Folders

Sizing Exchange: CPU

- Back-end (Exchange 2000) servers scale well through 4P
  - Sweet Spot moving down the chain?
  - 4 to 8 processors? Exchange 2003?
  - Above 8 processors?
- Scalability factors
  - L2, iL3 cache
  - Front-side bus impact
  - Hyperthreading impact

Hyper-Threading Scaling

- Common workload (5,000 "heavy" LoadSim users)
  - Exchange 2003 post-RC1
  - Windows 2003 RTM
- Common configuration variables (4GB, SG/MDB, spindles, network)
- ProLiant DL360G3: 2.8GHz Xeon
- ProLiant ML570G2: 2.0GHz Xeon MP



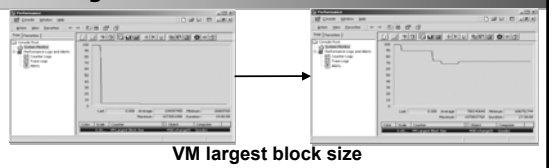
Sizing Exchange: Memory

- 3 – 4GB physical
- Virtual Memory
  - Introduction of /USERVA=xxx to fine-tune the kernel mode space size in Win2K3 (recommended: xxx=3030)
- Impact of memory speeds, interleaving
- Cache Area
  - DBA introduced in Exchange 5.5
  - More efficient in RAM than in the controller or VA
  - Exchange does not use AWE: >4GB is useless

Memory Tuning

- Buffer Area
  - Data kept only in memory
  - Attempt to gather slow devices accesses
  - I/O pattern changes from 5.5 to Exchange 2000
  - Size is small by default
  - Too large: Massive transaction log access
  - Too small: Frequent transaction log access
    - Impact on the transaction commit latency
    - Linked to user response time

Virtual Memory Improvements In Exchange 2000 SP3



VM largest block size

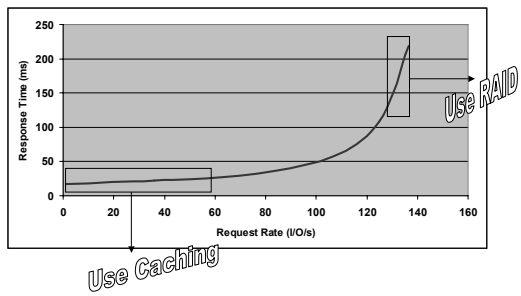
VM total large free block bytes

Sizing Exchange: I/O

- I/O Profiles
  - EDB versus STM versus Logs
  - Mailbox versus Bridgehead
  - Impact of Exchange, third-party elements
- Store and log files on separate RAID volumes
  - Store: RAID 0+1 (recommended) or RAID 5
  - Log files: Recommend RAID 1 with write-back caching
- Separate page file partition on system disk
- Use Hardware RAID controllers
- Virtualization changing the landscape?

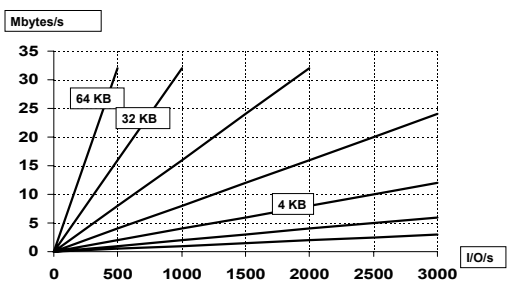
I/O Performance Fundamentals

What are we trying to achieve?



I/O Performance Fundamentals

Why is the I/O Profile Important?



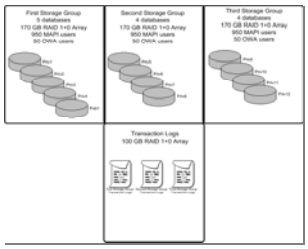
Performance Design
I/O Profiling by RAID level
Play Along at Home!

- How to Analyze Disk I/O (Formulas)
  - RAID0: READS + WRITES = I/Os per second
  - RAID1 / 0+1: READS + (2*WRITES) = I/Os per second
  - RAID5: READS + (4*WRITES) = I/Os per second
- BASIC Disk Subsystem Capabilities (per spindle)
  - Random I/O: 80-100 I/Os per second
  - Sequential I/O: 150-200 I/Os per second
- I/O Profile Example
  - 12 x 36GB drives in a RAID5 Array for IS Store (Random I/O)
  - PerfMon: Reads/Sec: 265; Writes/Sec: 198
  - Disk Array Capability: 12*80 = 960 I/Os per second
  - Exchange Server Demand: 265+(4*198) = 1060 I/Os/sec

Reference: Compaq RAID Technology for Database Servers White Paper

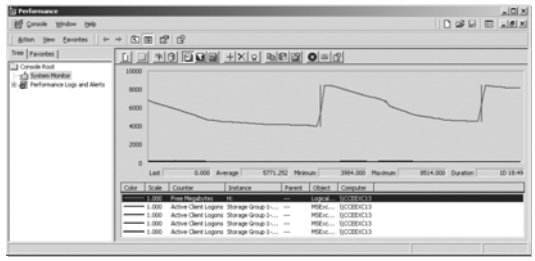
Database Design

- Log placement based on resiliency to failure
- Impact of Storage Groups and MDBs
  - Read/write mix
  - Loss of SIS ratio
  - VM hit
- Example:
  3000 users



Transaction Log Growth

- Roughly 4GB/day for 1,500 corporate users



Storage Alternatives

- Direct-Attached Storage
  - Traditional SCSI cabinets and backplane RAID controllers
- Networking Technologies
  - Promise of common fabric
  - Bandwidth increasing at rapid pace
  - Improvements in network stack
    - Storage protocols = high bandwidth and reliability, low latency
    - Network protocols = low bandwidth and reliability, high latency

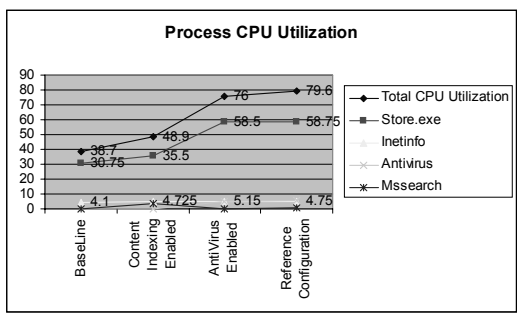
Storage Alternatives

- Storage Area Network (SAN)
  - Consolidation and Availability
    - Multi-node, application fabric
    - Volume Shadow Copy Services with Windows Server 2003, Exchange 2003
    - RAIS (SAN booting)
  - Virtualization benefits (I/O performance)
- iSCSI
  - Bridge between SAN and TCP/IP network
  - Microsoft driver expected in June 2003
- Network-Attached Storage (NAS)
  - KB Article (Q317173) describes block-mode requirements

Sizing Exchange: Software

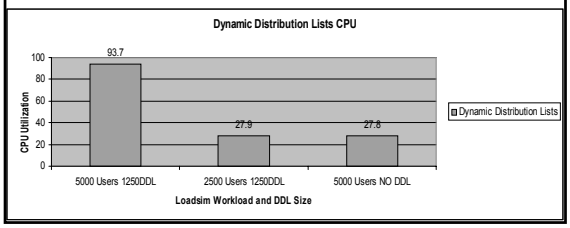
- Software component impacting CPU
  - Anti-Virus
  - Content indexing
  - Exchange 2003-specific
    - DDLs
    - Smart Folders
- Mobile device support
- Refer to Performance Tuning Guide for Exchange-specific registry key settings

Example Of Software Impact



Exchange 2003 DDLs

- Example shows DDLs of 1,250 users against different user populations
- DL membership is based on LDAP query result returning all members of a single MDB



Topics

Sizing Process

Server Capacity Planning

Exchange Server Sizing and Design

Monitoring and Tuning
- Tools
- Objects/Counters
- Percentiles versus Averages

Best practices

Monitoring Tools

- PerfMon
  - The good, the raw, and the ugly
- MOM
  - Exchange Management Pack for MOM
  - Included with Exchange 2003
- Third Parties

Baseline Analysis

- Contextual analysis user interviews
  - Client type
- Tracking log analysis
- Performance counters analysis
- Necessary to start your trend analysis
  - Helps you to pro-actively manage your capacity

Perfmon Basics Objects and counters

- Main Exchange Objects
  - Database(s)
  - MSExchangeIS series
  - MSExchangeMTA
  - SMTP Server
  - Exchange Web Mail
  - Process
    - STORE
    - MTA
    - Internet Information Server

If You Were To Pick One Counter?

- MSExchangeIS Mailbox(_Total)\Send Queue Size
- Others
  - Process\STORE\CPU
  - Epoxy\<xxx> Que Len
  - SMTP Server\Categorizer Queue Length
    - Indicates if Active Directory is not handling the demanded workload

      Something that is not always obvious to determine

Measurement: CPU And Memory

- CPU: Processor Object
  - System Object
    - Context switches per second
    - Processor Queue Length +++
- Memory: Memory Object
  - Can hardly do anything
  - Page faults per second DOES NOT mean access to the pagefile
    - Page Reads/sec
    - Page Writes/sec

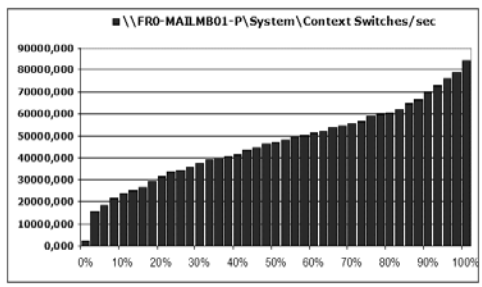
Measurement: Disk I/O

- Forget % Disk Time
- Forget averages
- Ignore Queue Length
- Key measurements
  - Request Rate: Read / Sec, Write / Sec
  - Data Rate: Bytes / Read, Bytes / Write
  - R/W Ratio: <none> (use Excel)
  - Latency: Second / Read, Second / Write
- No physical disk counters by default

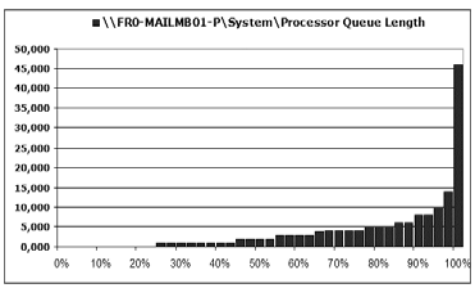
Reporting

- At sampling every second
  - +200MB of data for 8 hours
- Challenge
  - Make sense of this data
- Transactional system
  - Don't use averages
  - Use Percentile

Example: Context Switches Per Second



Corresponding CPU Load



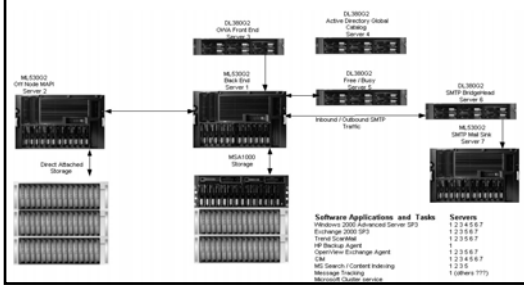
Topics

Sizing Process
Server Capacity Planning
Exchange Server Sizing and Design
Monitoring and Tuning

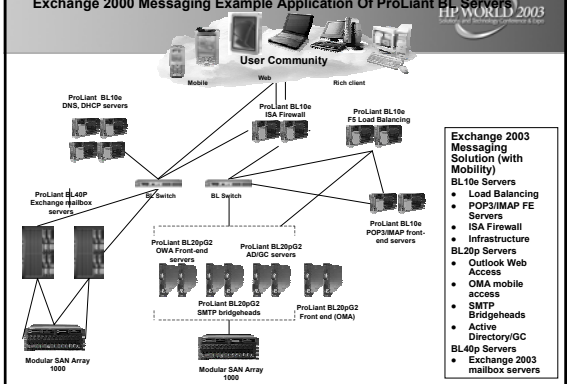
Best practices
Sample configurations
Summary

ERC Test Example

- Benchmark: 8,010 MMB2
- ERC Test: 3,000 "users" at ~80% CPU



Exchange 2000 Messaging Example Application Of ProLiant BL Servers



Best Practices

- Performance testing benefits
  - Basic MAPI or OWA workload analysis
    - Big changes in functionality, performance impact from Exchange 5.5 → Exchange 2003
  - Backup and restore performance
  - Online defragmentation impact
  - Allows to establish accurate service levels
    - Set expectations correctly!

Best Practices

- Treat registry changes with care
  - Further Service Pack may address some issues
  - Issues around deployment on many servers
  - Always room for mistakes/mis-handling
  - Latest service packs address much of the tweaking required

For Further Reference

- Microsoft Exchange Tech Resources Website
  - http://www.microsoft.com/exchange/techinfo/default.asp
  - http://www.microsoft.com/technet/default.asp
  - http://support.microsoft.com for KB articles
  - http://www.hp.com/solutions/activeanswers for white papers, sizing tools
- Books and publications
  - Exchange Administrator Newsletter
  - Pierre Bijaoui's excellent book on Exchange 2000 Performance

HP WORLD 2003
Solutions and Technology Conference & Expo

Interex, Encompass and HP bring you a powerful new HP World.

interex shared knowledge • shared power
encompass AN HP USER GROUP
hp invent