Windows Server 2003-Terminal Services Scalability with HP ProLiants

Allen D. Northcutt, Engineering Manager
Server Based Computing Solutions
Industry Standard Servers Division
August 2003

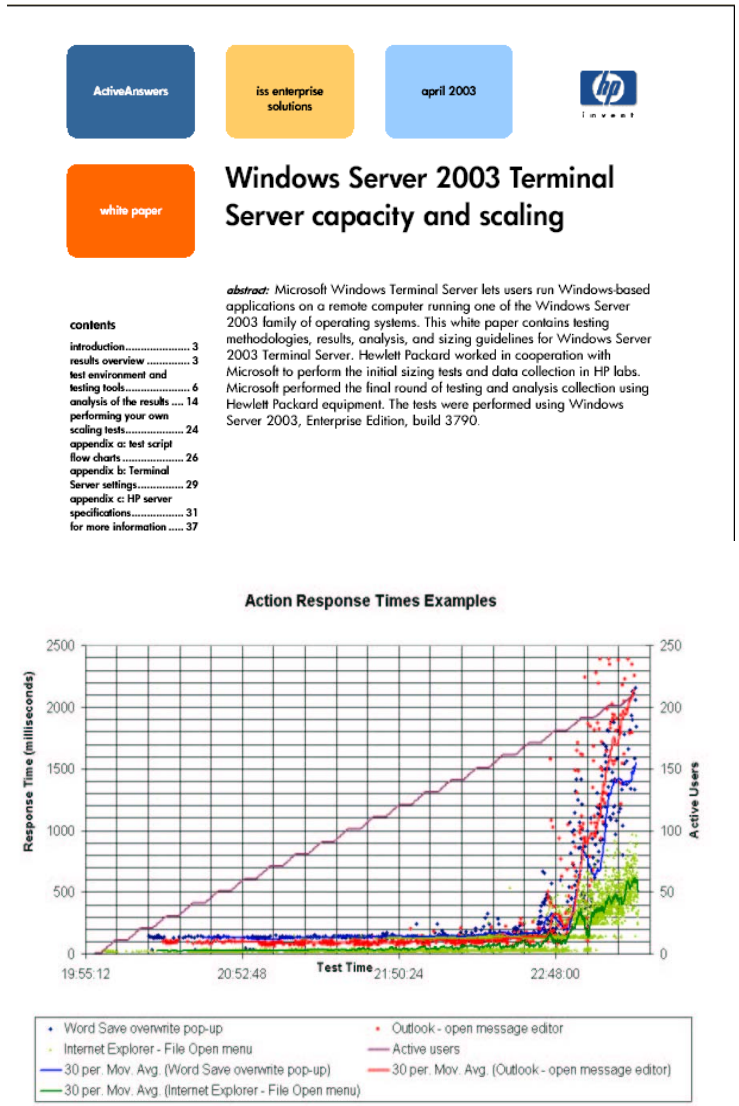
HP WORLD 2003
Solutions and Technology Conference & Expo

Agenda

- Introduction
- Goals
- What was tested?
- Where was the testing performed?
- Scalability results
- Sizing recommendations
- Major Advances in kernel memory management
- Key findings in respect to kernel memory management
- SBC server sizing recommendations
- Online sizer
- A look back in time
- HP ProLiant server family
- BL v. DL
- Important links

Goals of Windows Server 2003 Scalability Testing

- Determine scalability of HP DL and BL server platforms
- Determine scalability of Windows Server 2003 Terminal Services
- Determine relative differences between Windows Server 2003 and Windows 2000 Server
- Evaluate impacts on performance of new features in Terminal Services
- Develop white paper describing the testing and the results
- Provide sizing guidance to the field through online sizer

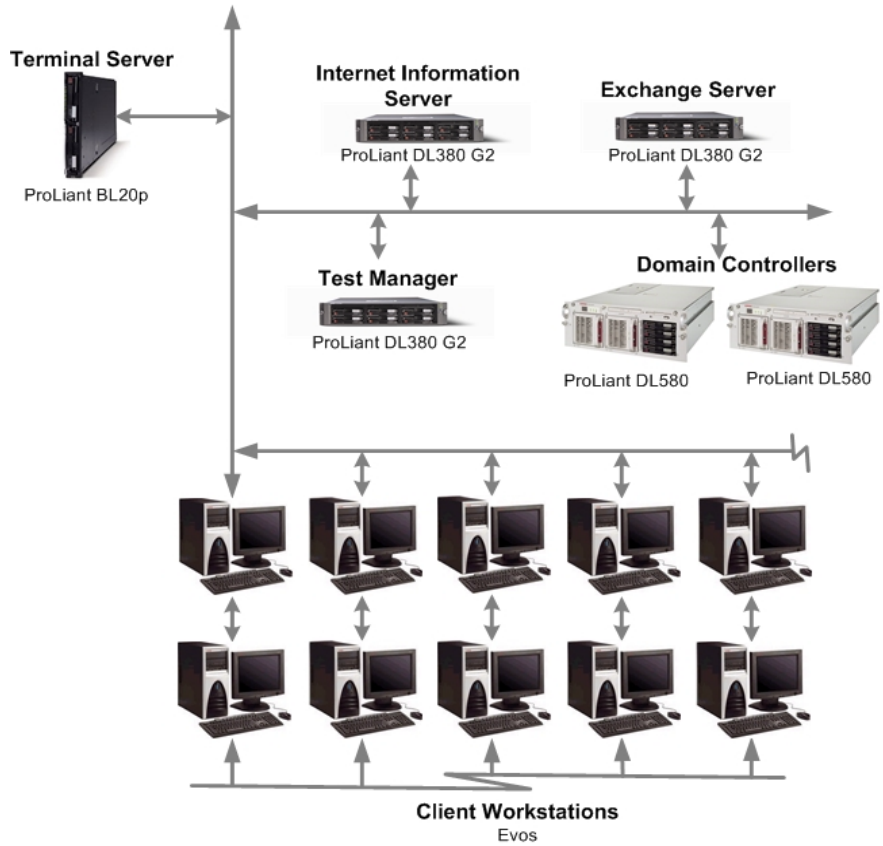


What was tested?

- DL360 G3
- DL380G3
- DL560
- DL580G2
- DL740
- DL760G2
- BL10e
- BL20p
- BL40p
- Legacy Support (DL360G1 and ProLiant 6400R)
- Windows 2000 w/TS
- Windows Server 2003 w/TS Enterprise Edition build 3790
- Windows Server 2003 w/TS

Where was it tested?

- HP SBC Solutions Lab
  - Houston, TX
  - 150 clients capable of driving 3000 terminal server sessions
  - Supplies performance characteristics for most BL and DL servers
    - Average 12 benchmarks per year
    - Supplies data for online sizer and performance briefs
  - Hosted Microsoft in March 2003 for Windows Server 2003 scalability effort.



Scalability Results

Server configuration	ProLiant server model	Structured worker (Heavy)[5]	Knowledge worker (Medium)	Data entry worker (Light)
4xIntel Xeon Processors MP 2 GHz 2 MB L2 Cache 4096 MB	DL560	TBD	270[1,2]	520[3]
2xIntel Xeon Processors 2.4 GHz 2 MB L2 Cache 4096	DL360G3	TBD	200[3]	440
1xIntel Xeon Processors 2.4 GHz 2 MB L2 Cache 4096 MB	DL360G3	TBD	140	200
1xIntel Xeon Processors 2.4 GHz 2 MB L2 Cache 4096 MB	DL380G3	TBD	200[3]	440
2xIntel Xeon Processors 2.4 GHz 2 MB L2 Cache 4096 MB	BL20pG2	TBD	200	440
1xIntel Ultra Low Voltage Pentium III 900 MHz 1024 MB	BL10e	21	50	120
4xIntel Xeon Processors MP 2.0 GHz 2 MB L2 Cache 4096 MB	BL40p	TBD	240[3]	Not Tested
Pentium III Xeon 550 MHz 2 MB L2 Cache 4096 MB	6400R[4]	Will Not Test	170	Not Tested

[1] Using PAE

[2] System PTE Exhausted

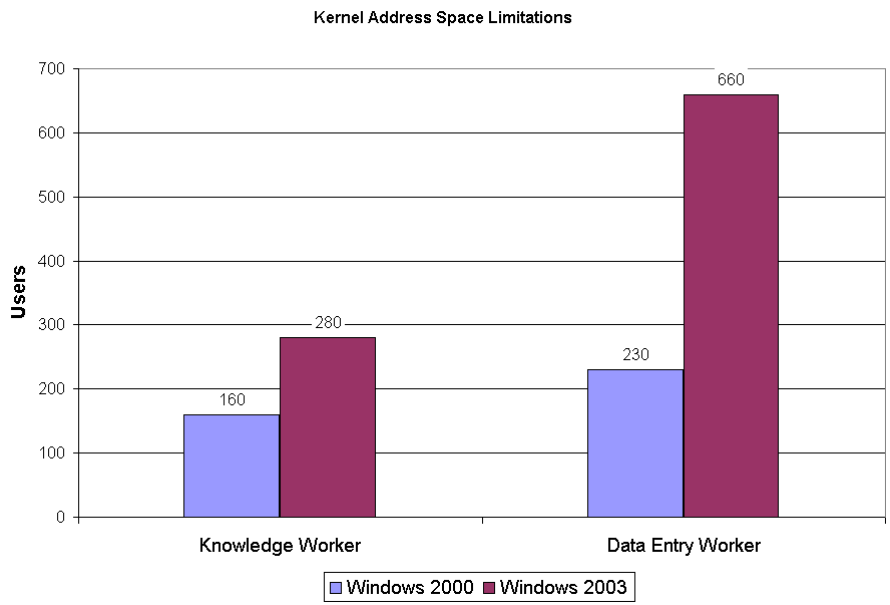
[3] Response time degradation due to cache hit ratio drop

[4] Retired platform utilized for legacy testing

[5] HP developed script post joint engagement

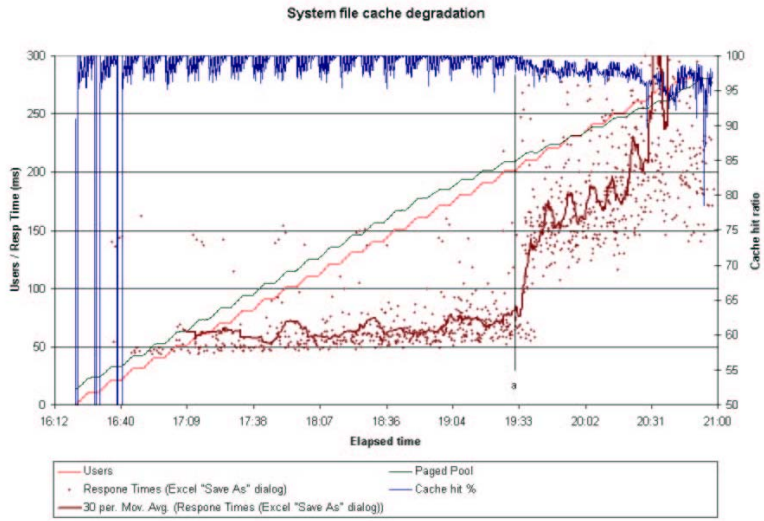
Major Advances in Kernel Memory Management

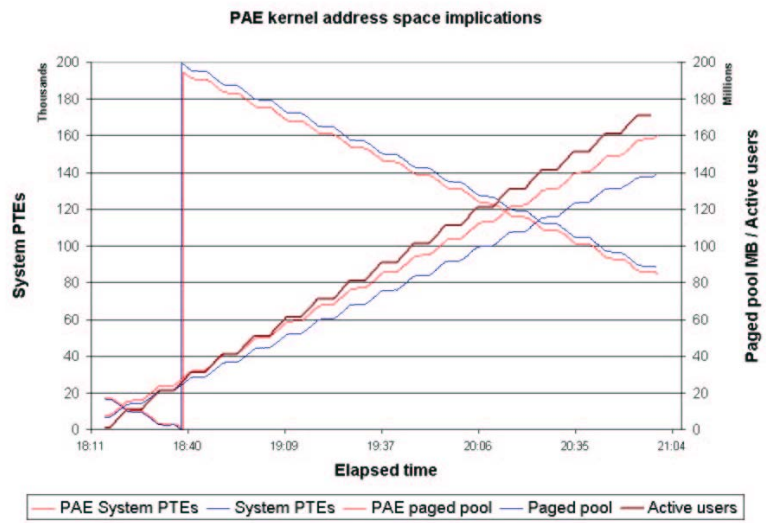
- 32 bit operating systems can address 2^{32} or 4 GB of Physical Memory
- By Default, Windows divides ½ to kernel and ½ to user-mode processes
- Kernel area divided between paged pool area, system page table entries (PTEs), and system file cache area
- Due to better kernel memory management, tests yielded 75% and 286% for knowledge and data entry worker respectively



Key Findings in Respect to Kernel Memory Management

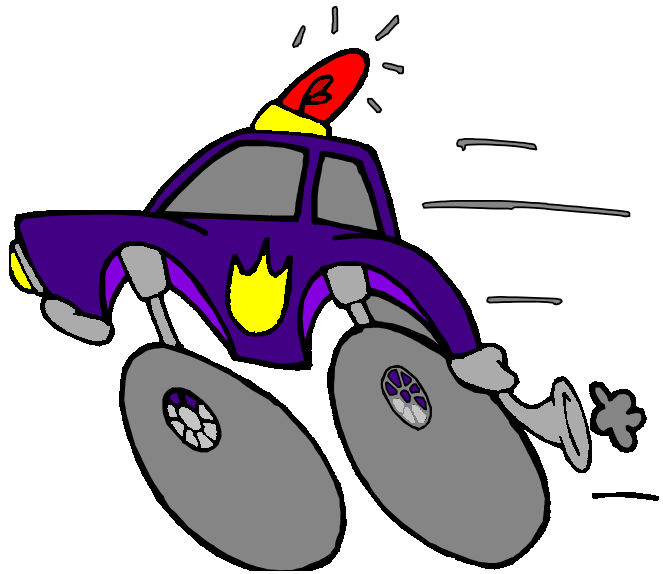
- Copy Read Hits % below 99% indicated the onset of server congestion
- /PAE may help some memory hungry applications
- Key ratio to monitor is System PTEs to paged pool areas. Ratio can be modified via HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Control\Session Manager\Memory Management





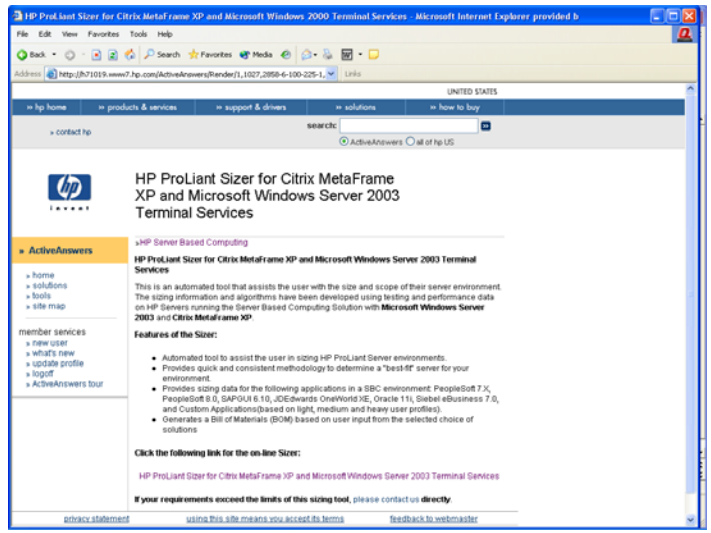
SBC Server Sizing Recommendations

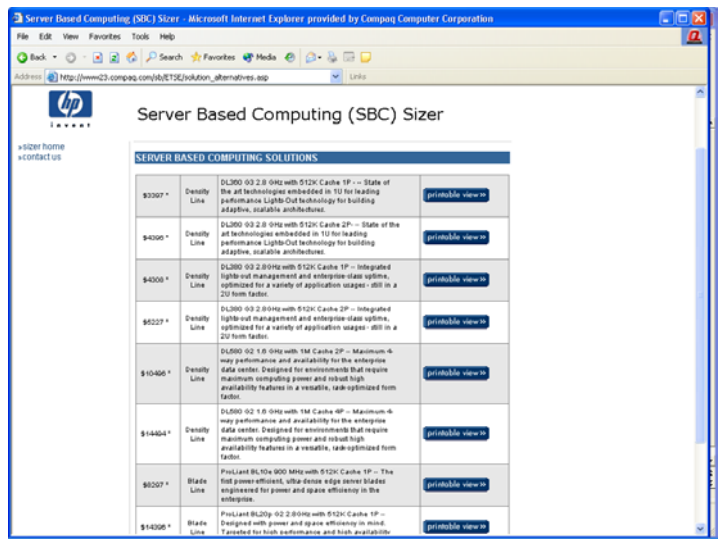
- Purchase the fastest clocking rate available
- Cache sizes above 1 MB have historically yielded minimal gains
- 2P servers have near linear scaling, 4P and beyond drop off considerably
- Now that servers are supporting greater user densities, I/O is becoming more of a concern
  - Pay attention to controller cache
  - Pay attention to I/O “top 4”: page file, profiles, user data, and system/application binaries
- Leave room to grow and for other SW agents/applications.



SBC Online Sizer (via ActiveAnswers)

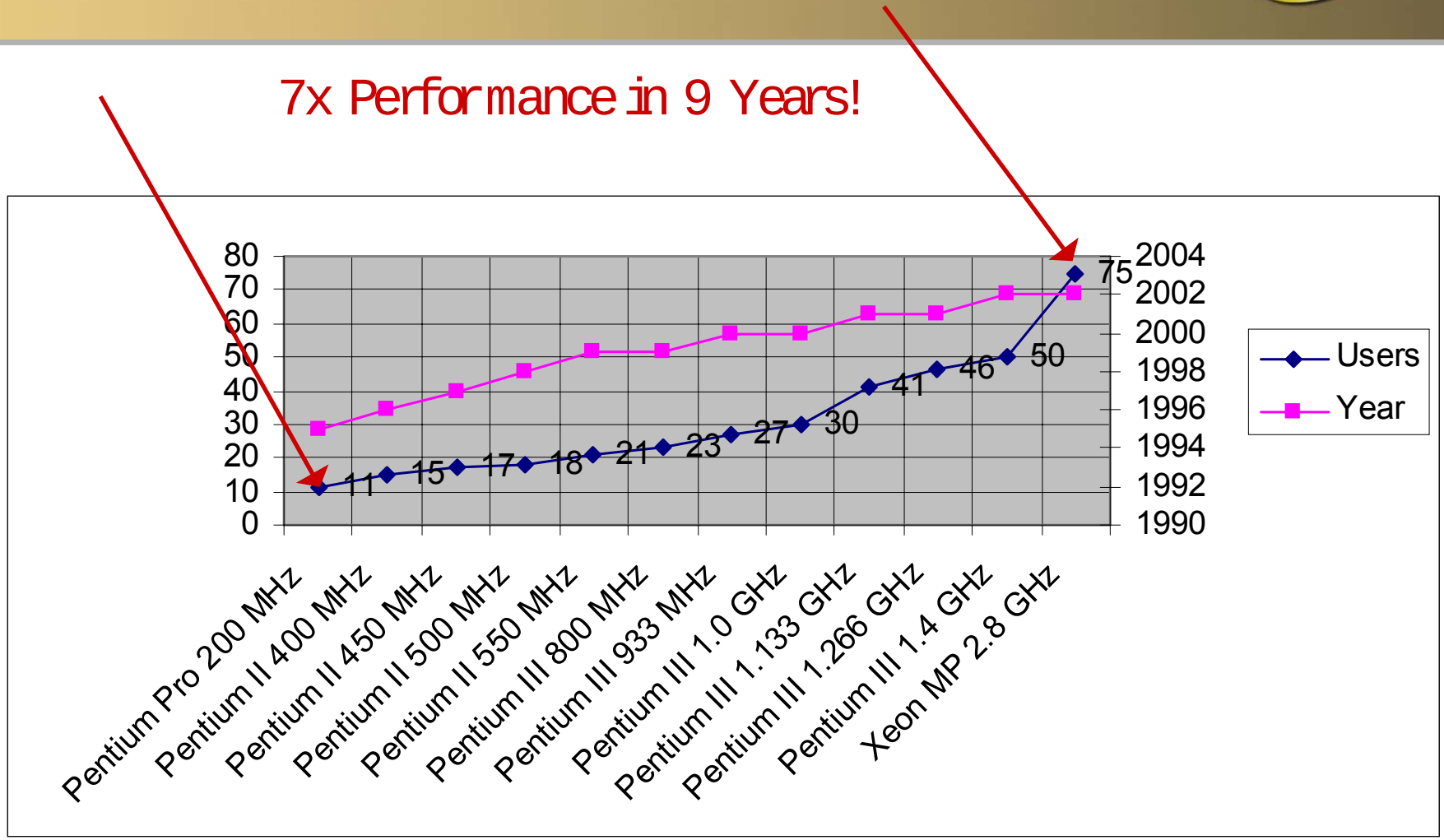
- **Features**:
  - Provides sizings for MSOFFICE, JDEdwards, PeopleSoft, SAP, Oracle, and Siebel
  - Server support for both BL and DL server lines
  - Provides consolidated Bill Of Material (p/n, pricing, and required configuration)
  - Thin client support
  - Advanced user inputs for the experienced IT professional
  - Support for Citrix MetaFrame and Windows Terminal Services
  - http://www.hp.com/solutions/activeanswers <check right pane>





- **Benefits**:
  - HP benchmarks server platforms so the customer does not have to
  - Shortens customer deployment cycle
  - Eliminates the over/under purchase risk
  - Automatically generates necessary purchase information

SBC Scalability 1995 - 2003

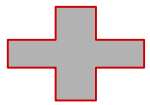


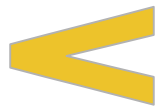
Simple replacement of older server with new

- More user manage per administrator
  - 4.6 x improvement of users per processor (16 to 75 user)
- Fewer servers to manage
  - 1000 users would require 16 –6400Rs vs. 7 –DL360s
- Fewer racks – less power and floor space
  - 1000 users 64 U (1.5 racks) to 7 U (.17 racks)

(2) 6400Rs 500 MHz supported 122 Users

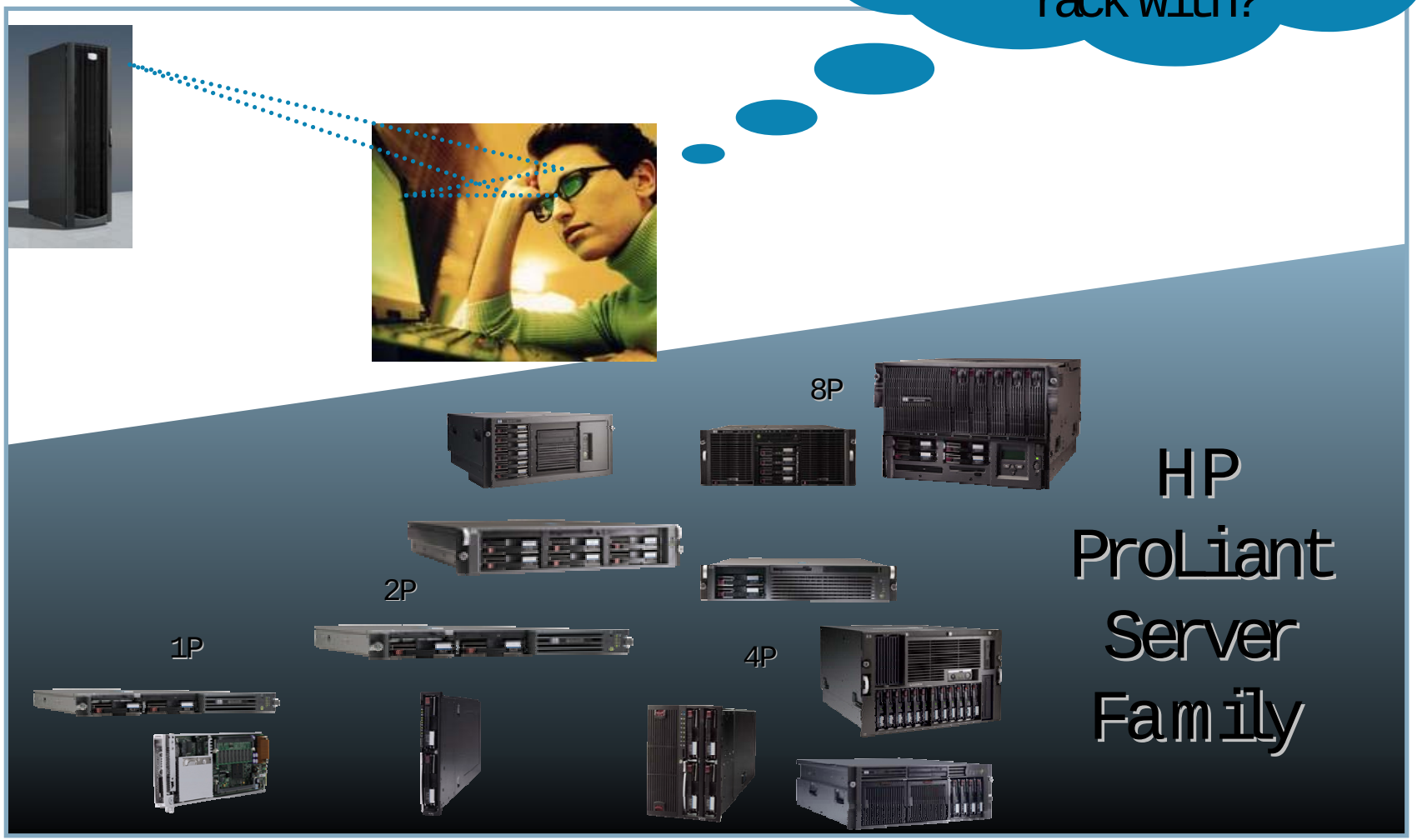
(1) DL360G3 2.8 GHz Supports 150 Users





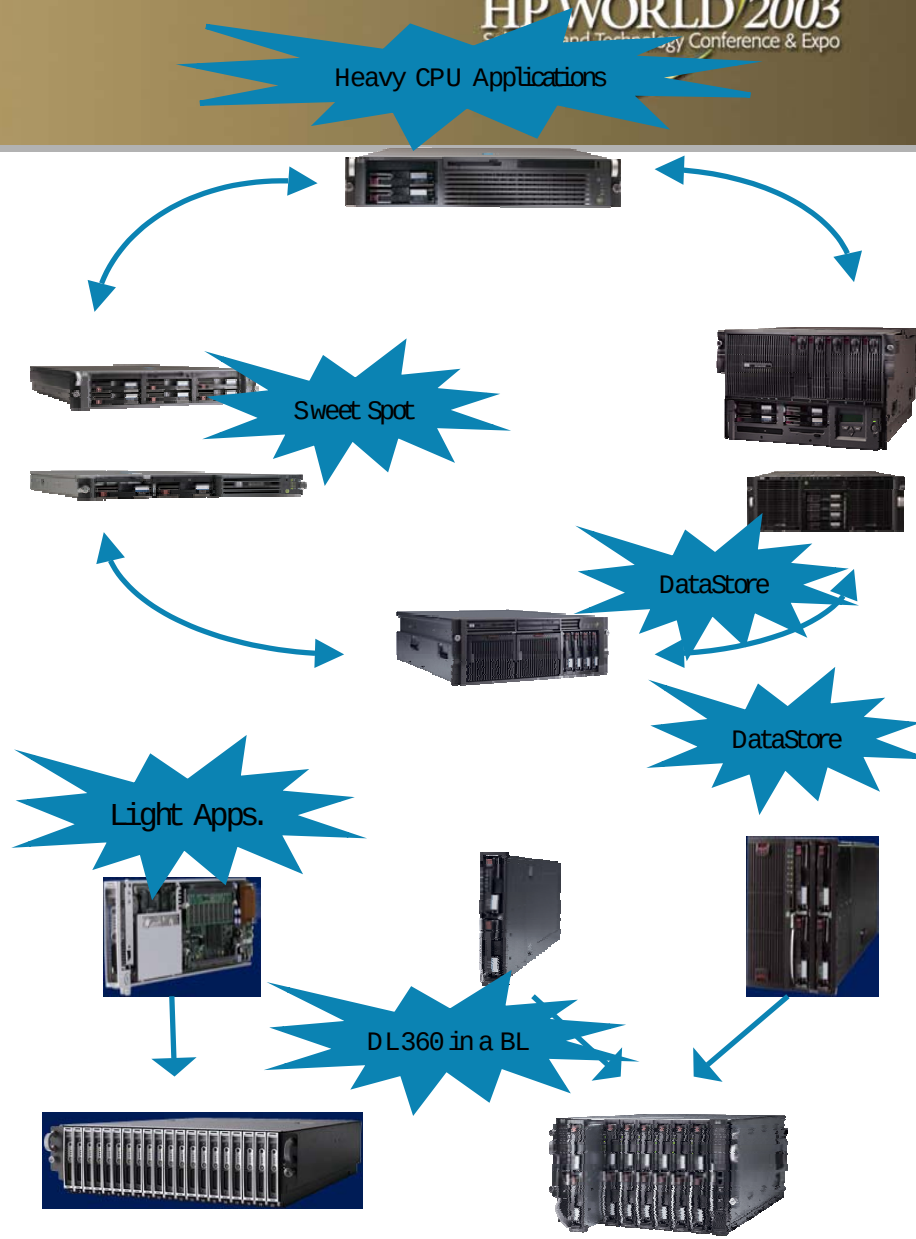
Lower total cost of ownership

HP ProLiant Server Family



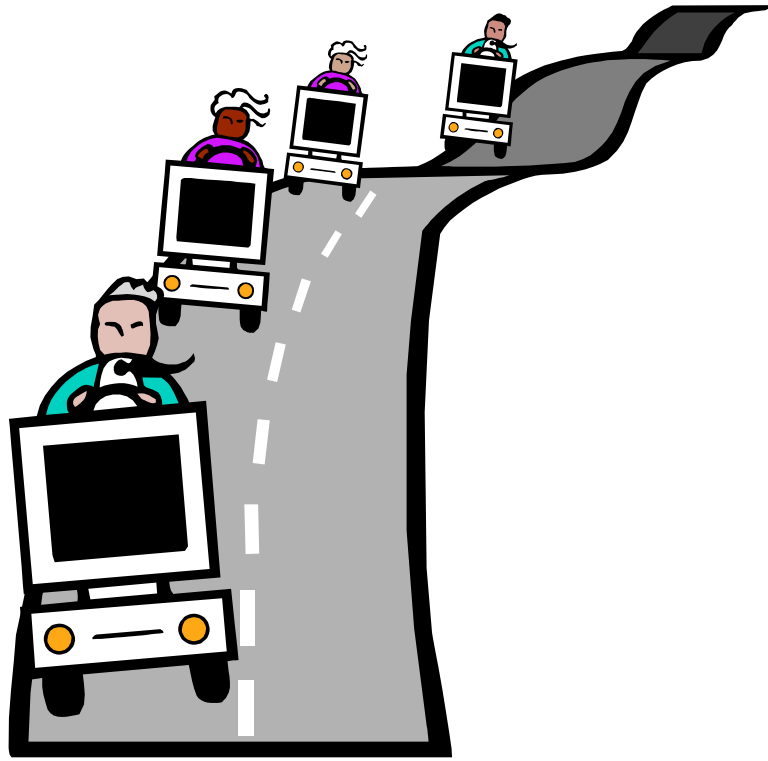
BL vs. DL

- DLs
  - Expandability (Disks, Slots, NICs, and More)
  - High Availability Features
- BLs
  - Density
  - Reduced Power Requirements
  - Cable Management
  - Floor Space Reduction



Important Links

- SBC solutions content: http://www.hp.com/solutions/activeanswers/hpsbc
- SBC online sizer: http://www.hp.com/solutions/activeanswers/proliant-sizer-term-services
- Windows server 2003 whitepaper: http://h71019.www7.hp.com/ActiveAnswers/Render/1,1027,6297-6-100-225-1,00.htm



Summary

- Windows Server 2003 scaled equally to Windows 2000 Server on a per CPU basis
- Windows Server 2003 scaled much better in kernel memory constrained situations
- /PAE will help certain memory hungry applications
- Three key kernel areas (System Cache, Page Pool, and System PTE) can be further tuned
- HP publishes scalability data on a monthly basis via ActiveAnswers

IIS6 Performance and Scalability on HP *ProLiant* server platforms

Kevin P. Kenefic, Engineering Manager
Ecommerce Integration Solutions
Industry Standard Servers Division
August 2003

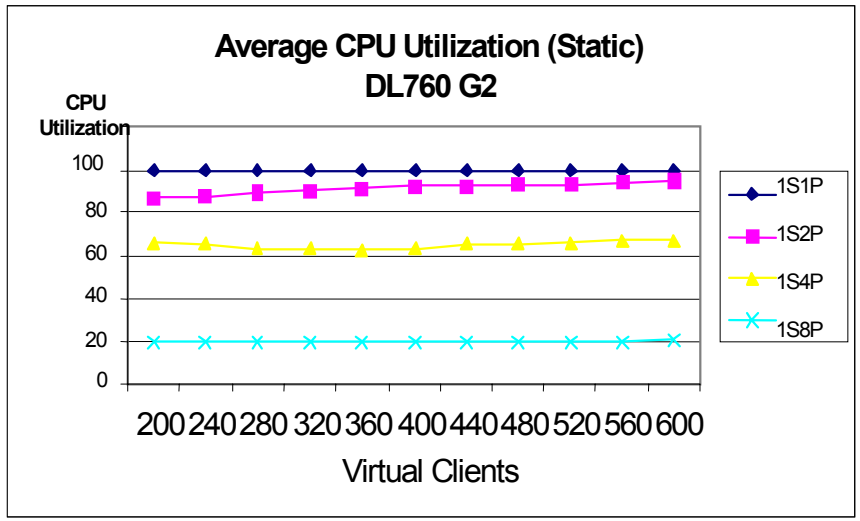
HP WORLD 2003
Solutions and Technology Conference & Expo

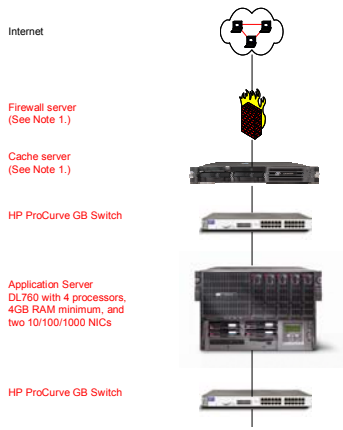
Agenda

- Introduction
- Goals
- Where it was tested
- What was tested
- How it was tested
- “The Numbers”
- Conclusions
- Recommendations
- Questions??????
- Contact information and related links
- Back-Up slides

Goals of Windows Server 2003 Scalability Testing

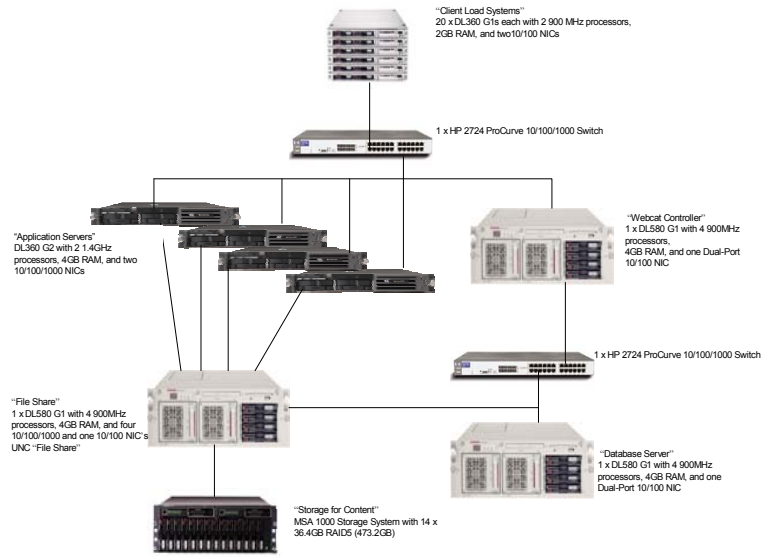
- Test the performance and scalability characteristics of IIS6 on HP DL and BL server platforms
- Write a Performance White paper detailing the test findings
- Publish a Recommended Configuration White Paper designed to provide timely configuration suggestions
- Develop an On-Line sizing tool that will assist customers and SIs size their specific site





Where was it tested?

- HP Ecommerce Integration Solutions Lab
  - Houston, TX
  - 20 Physical client servers capable of generating 600 simultaneous virtual clients
  - Generated performance data for the majority of the BL and DL servers models
  - Performance data was used to develop online sizer, recommended configurations and the performance white paper



What was tested?

HP Hardware

DL360 G3

What was tested?

HP Hardware

DL360 G3

DL580G2

What was tested?

HP Hardware

DL360 G3

DL580G2

DL760G2

What was tested?

HP Hardware

DL360 G3

DL580G2

DL760G2

BL20p

What was tested?

HP Hardware

DL360 G3

DL580G2

DL760G2

BL20p

BL40p

What was tested?

HP Hardware

DL360 G3

DL580G2

DL760G2

BL20p

BL40p

What was tested?

HP Hardware

HP WORLD 2003
Solutions and Technology Conference & Expo

DL360 G3

DL580G2

DL760G2

BL20p

BL40p

Windows Server 2003

+

IIS 6

What was tested?

HP Hardware

DL360 G3

DL580G2

DL760G2

BL20p

BL40p

Windows Server 2003

+

IIS 6

+

WebCat

(Microsoft's script based test tool)

What was tested?

LCW2 Web Site: (Using Microsoft's WebCat test tool)

- 120GB of static and dynamic data

What was tested?

LCW2 Web Site: (Using Microsoft's WebCat test tool)

- 120GB of static and dynamic data
- 100% Static web data scenario

What was tested?

LCW2 Web Site: (Using Microsoft's WebCat test tool)

- 120GB of static and dynamic data
- 100% Static web data scenario
- 100% Dynamic web data scenario

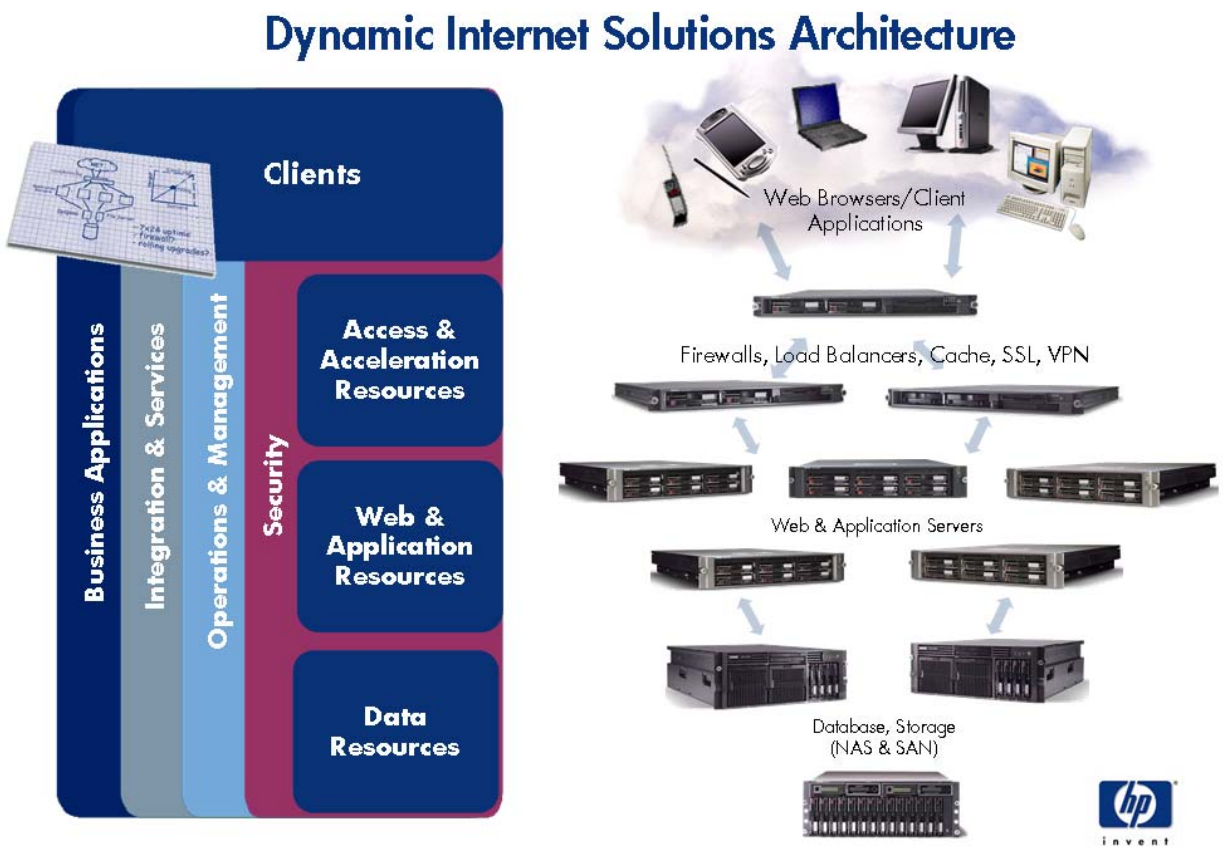
What was tested?

LCW2 Web Site: (Using Microsoft's WebCat test tool)

- 120GB of static and dynamic data
- 100% Static web data scenario
- 100% Dynamic web data scenario
- Fully mixed web data scenario

How was it tested? Using DISA!!!!

The Dynamic Internet Solutions Architecture (DISA) test configuration was used:

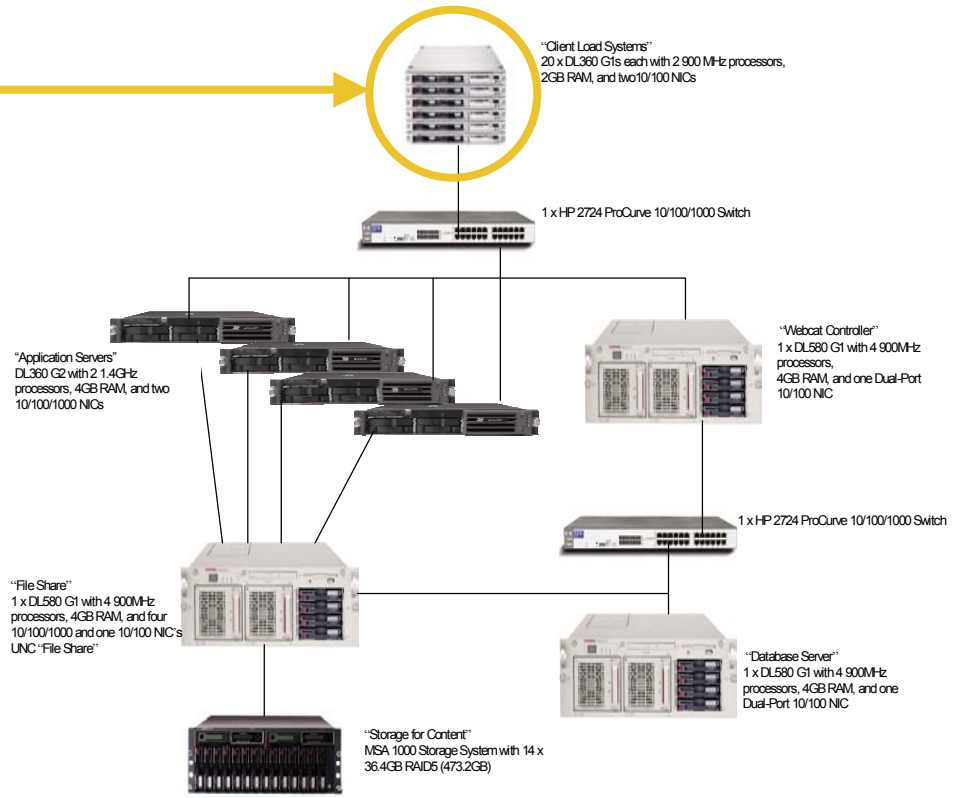


http://www.hp.com/solutions/disa

How was it tested?

Load Generation Clients:

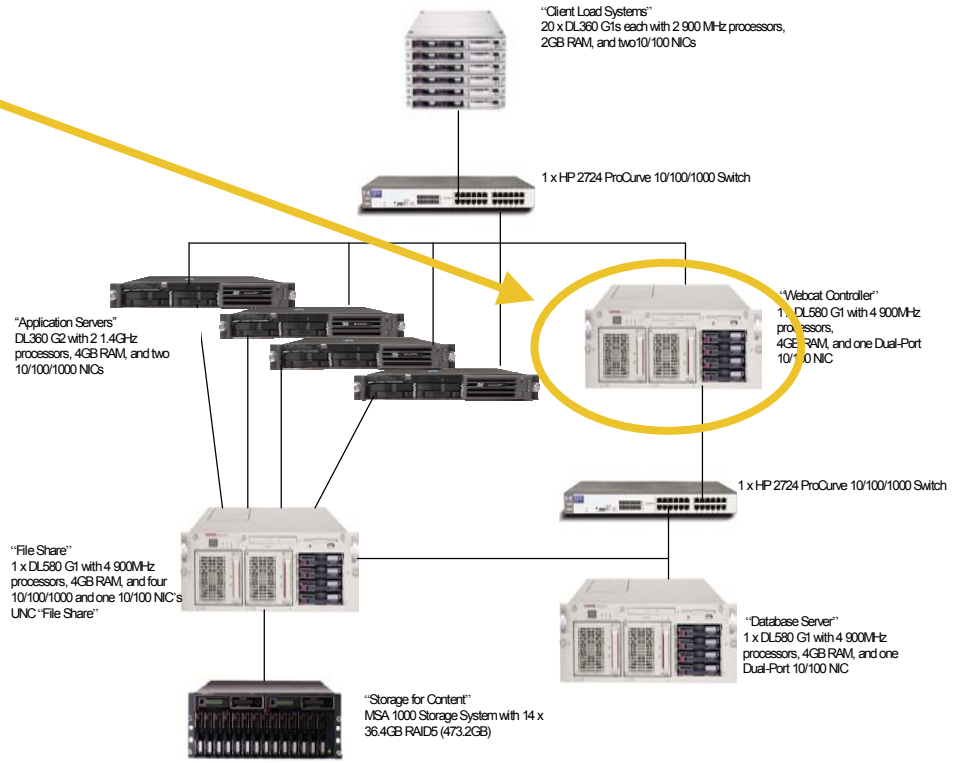
- 20 *ProLiant* DL360 G1s
- Two (2) 800 MHz Processors
- 2 GB RAM
- Dual Integrated 10/100 NICs
- WebCat test tool client
- Windows Server 2000



How was it tested?

Load Generation Controller
(WebCat Controller)

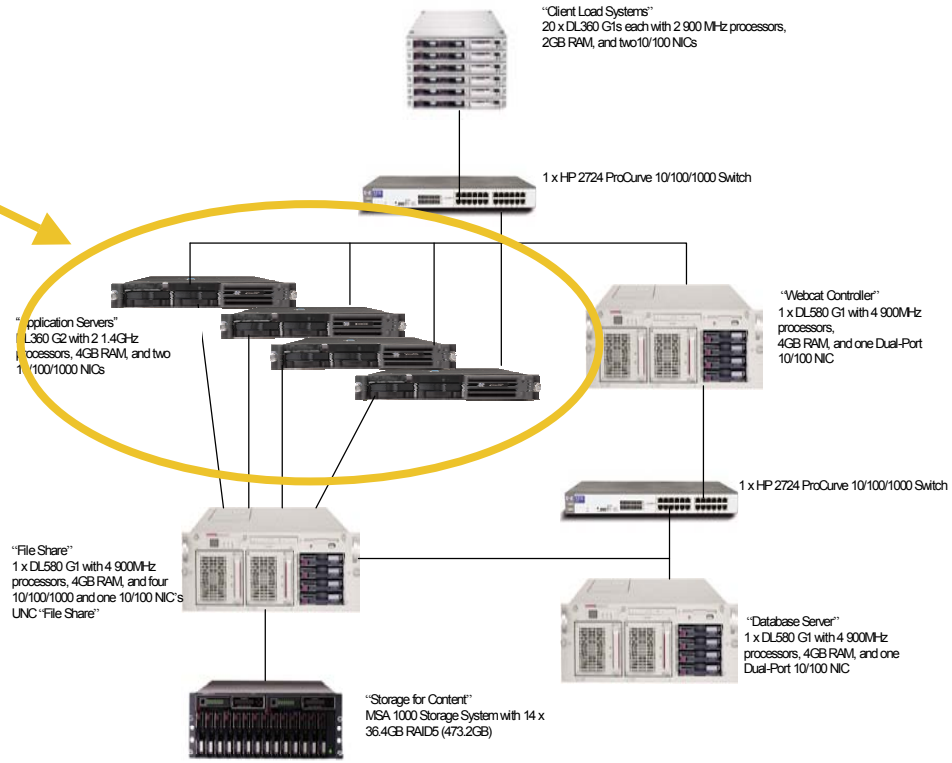
- HP *ProLiant* 6500
- Four (4) Pentium III 900 MHz-256KB cache Processors
- 2GB RAM
- Two (2) NC3131 Dual-Port Fast Ethernet NICs
- WebCat Controller software
- Windows Server 2000



How was it tested?

Application Web Servers
(Systems under Test)

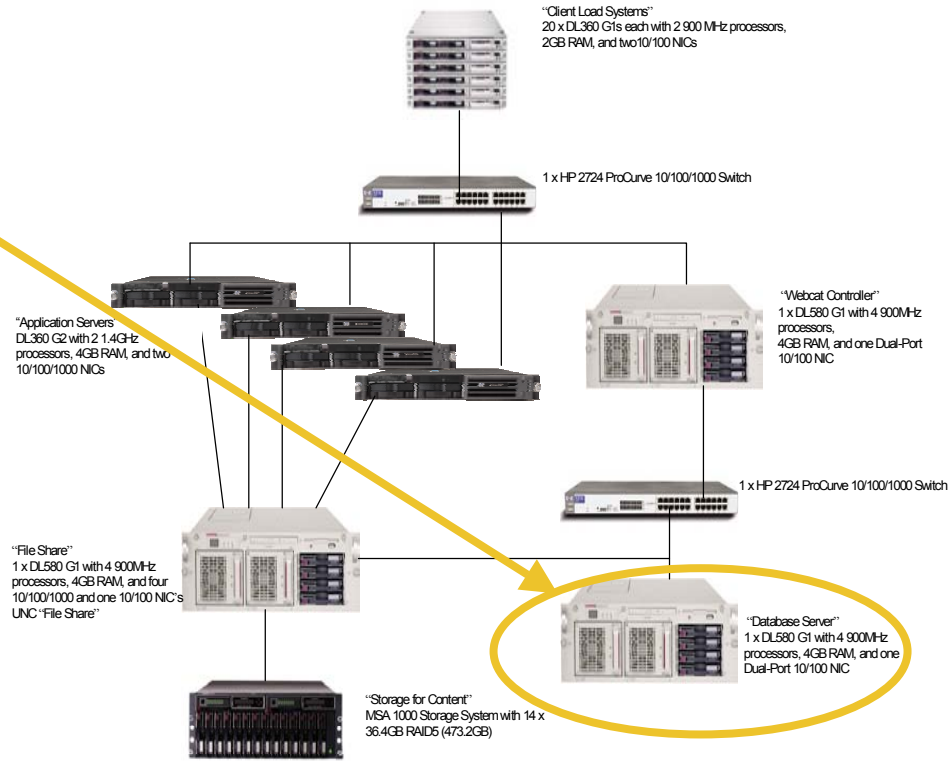
- DL360 G2
- DL580 G2
- DL760 G2
- BL20P Blades
- BL40P Blades
- LCW2 Web Site
- Windows 2003 Enterprise



How was it tested?

DataBase Backend

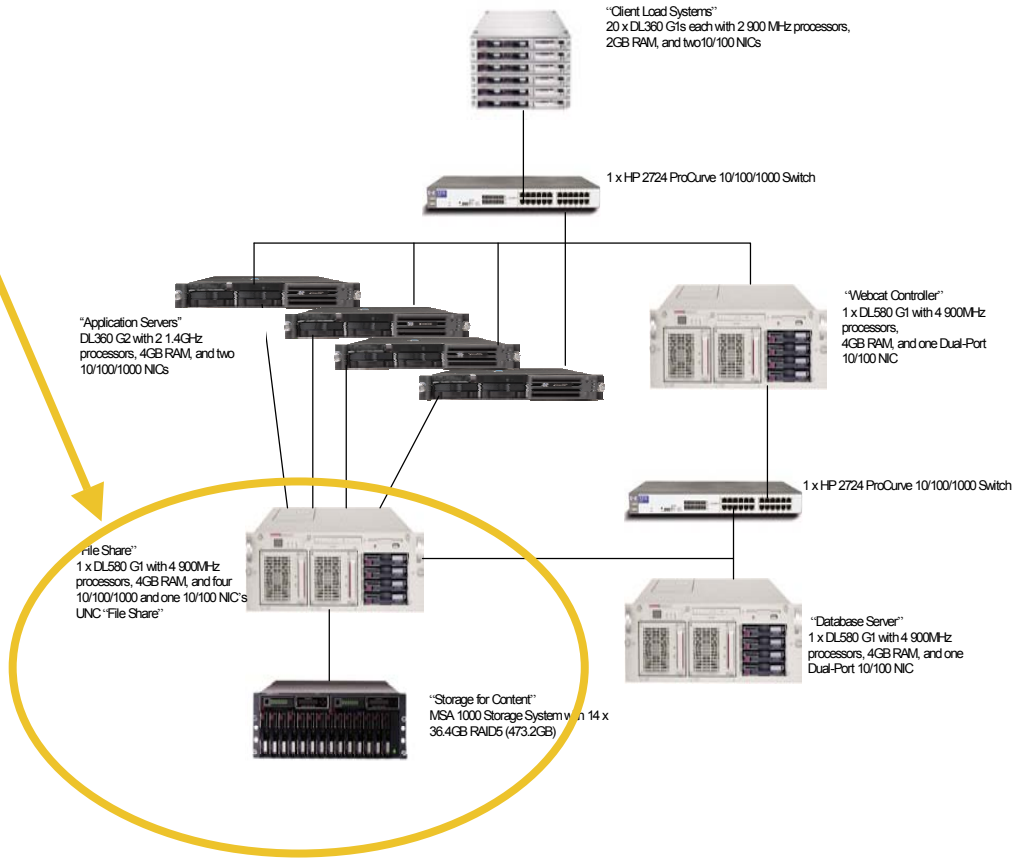
- HP *ProLiant* DL580 G1
- Four (4) Pentium III 900 MHz Processors
- 2GB RAM
- One (1) NC3131 Dual-Port Fast Ethernet NIC
- SQL 2000
- Windows 2003 Enterprise



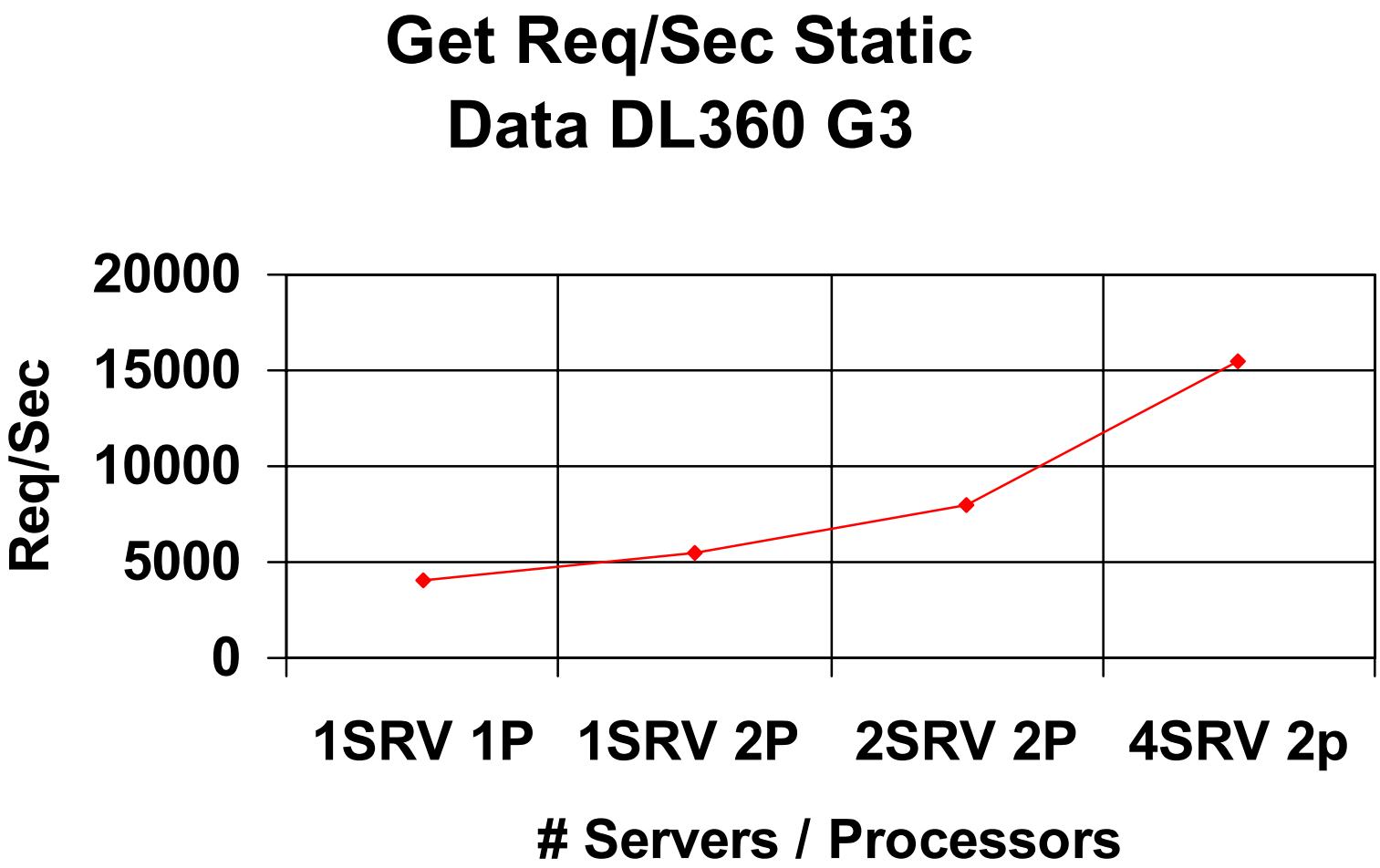
How was it tested?

File Share (centralized content)

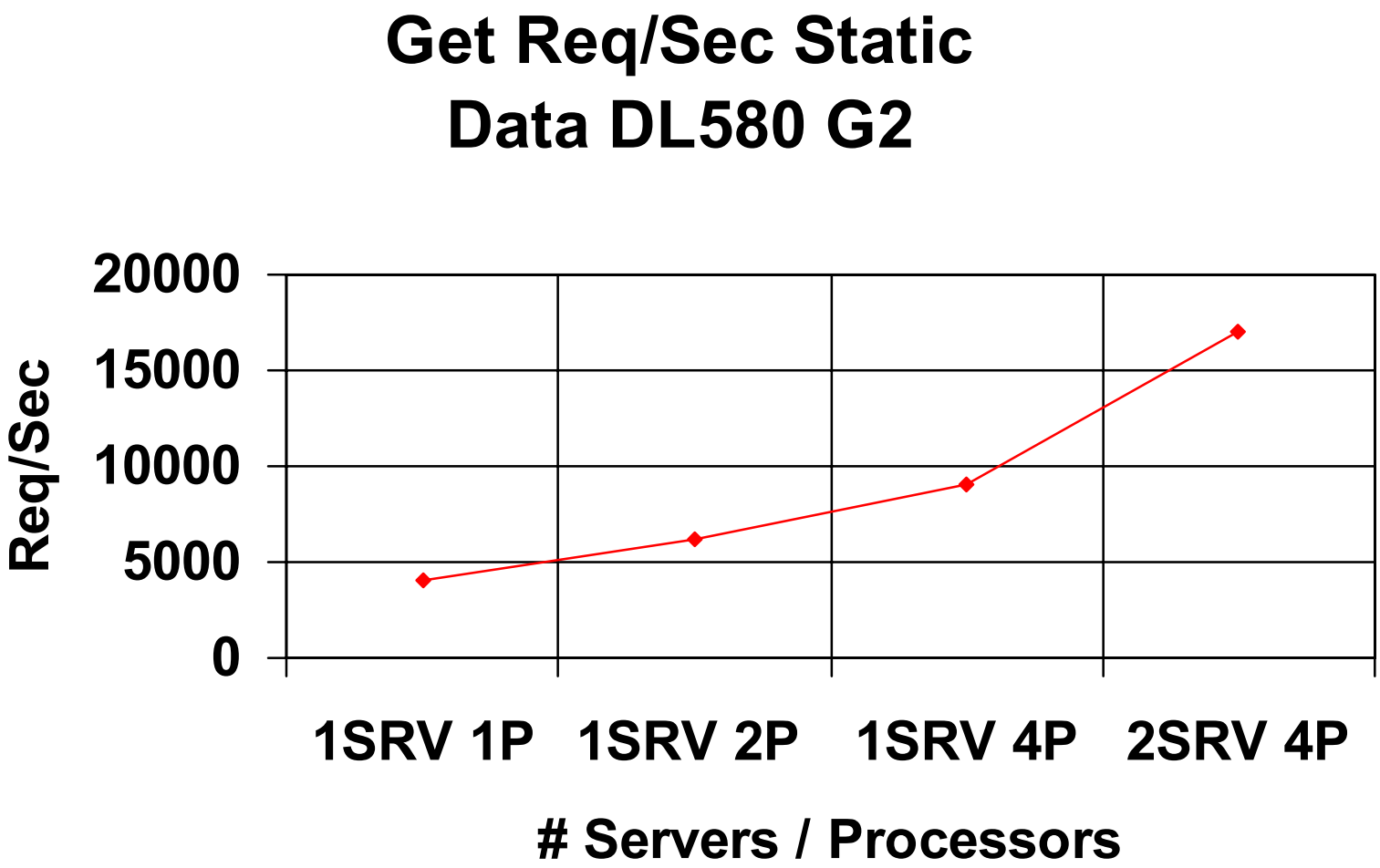
- HP *ProLiant* DL580 G1
- Four (4) Pentium III 900 MHz Processors
- 2GB RAM
- Four (4) NC3131 Dual-Port Fast Ethernet NICs
- One (1) 2GB Fibre Channel Host Bus Adapter (to the MSA 1000 SAN)
- Windows 2003 Enterprise
- MSA 1000 SAN



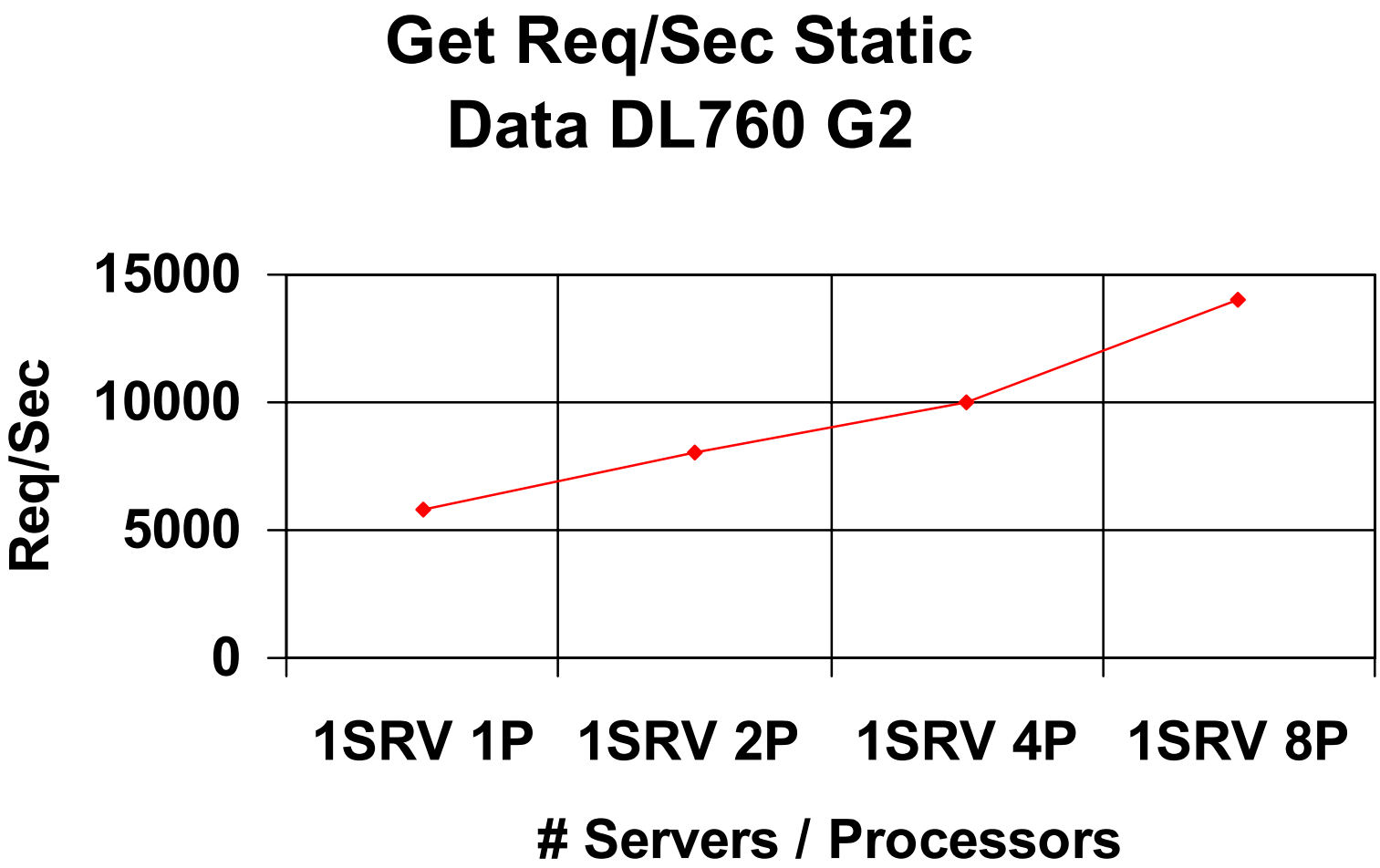
Performance Scalability: "The Numbers"



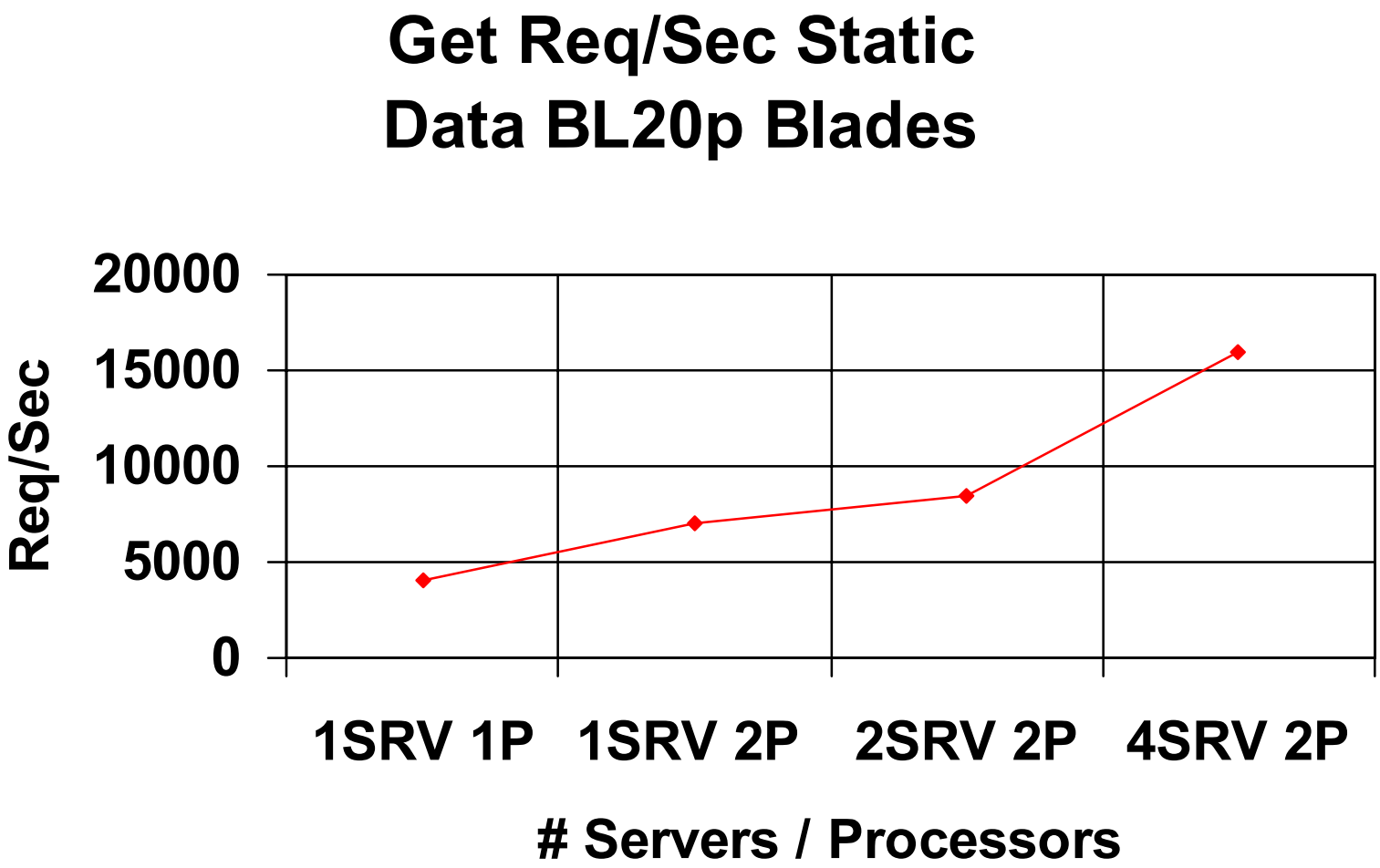
Performance Scalability: "The Numbers"



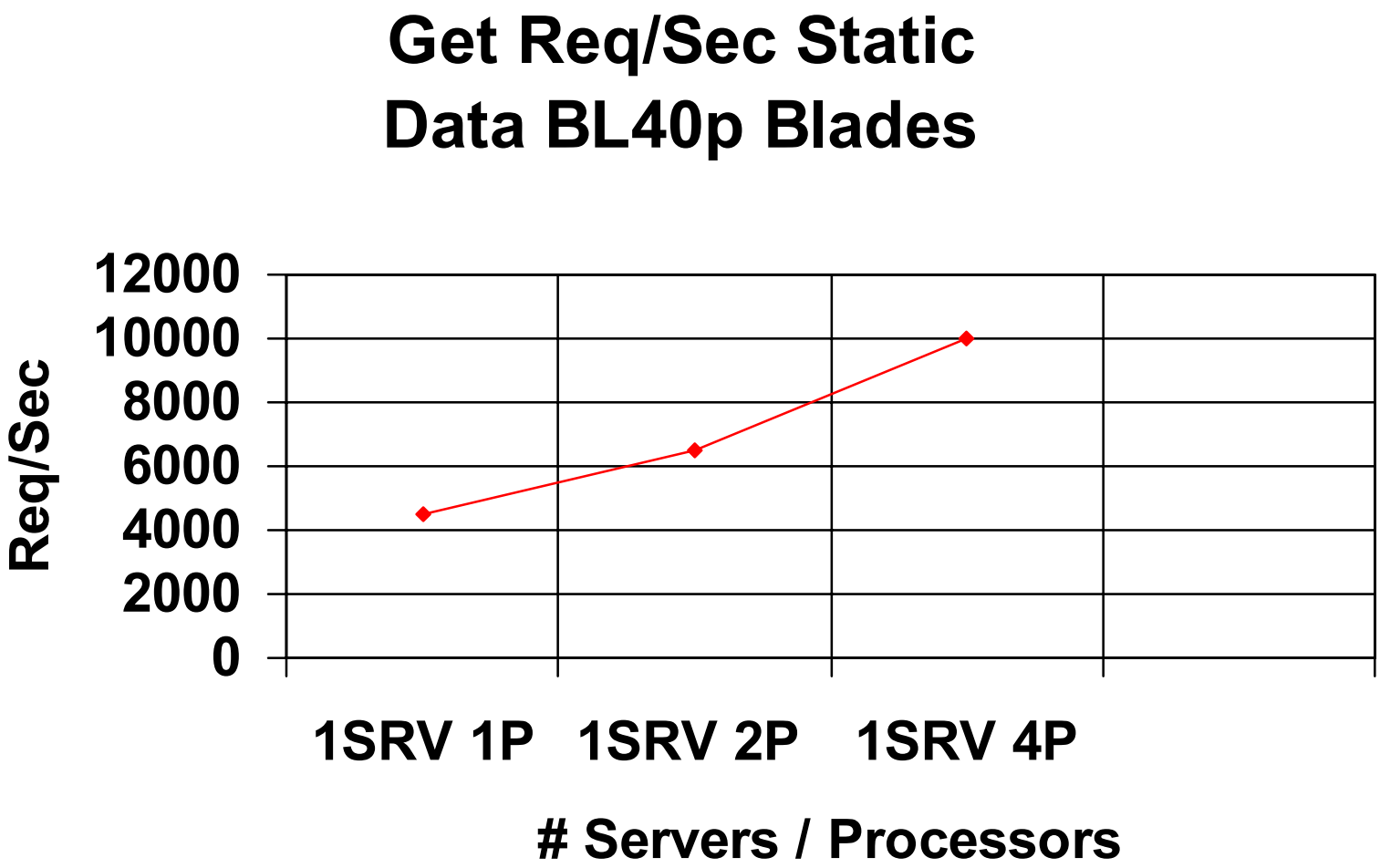
Performance Scalability: "The Numbers"



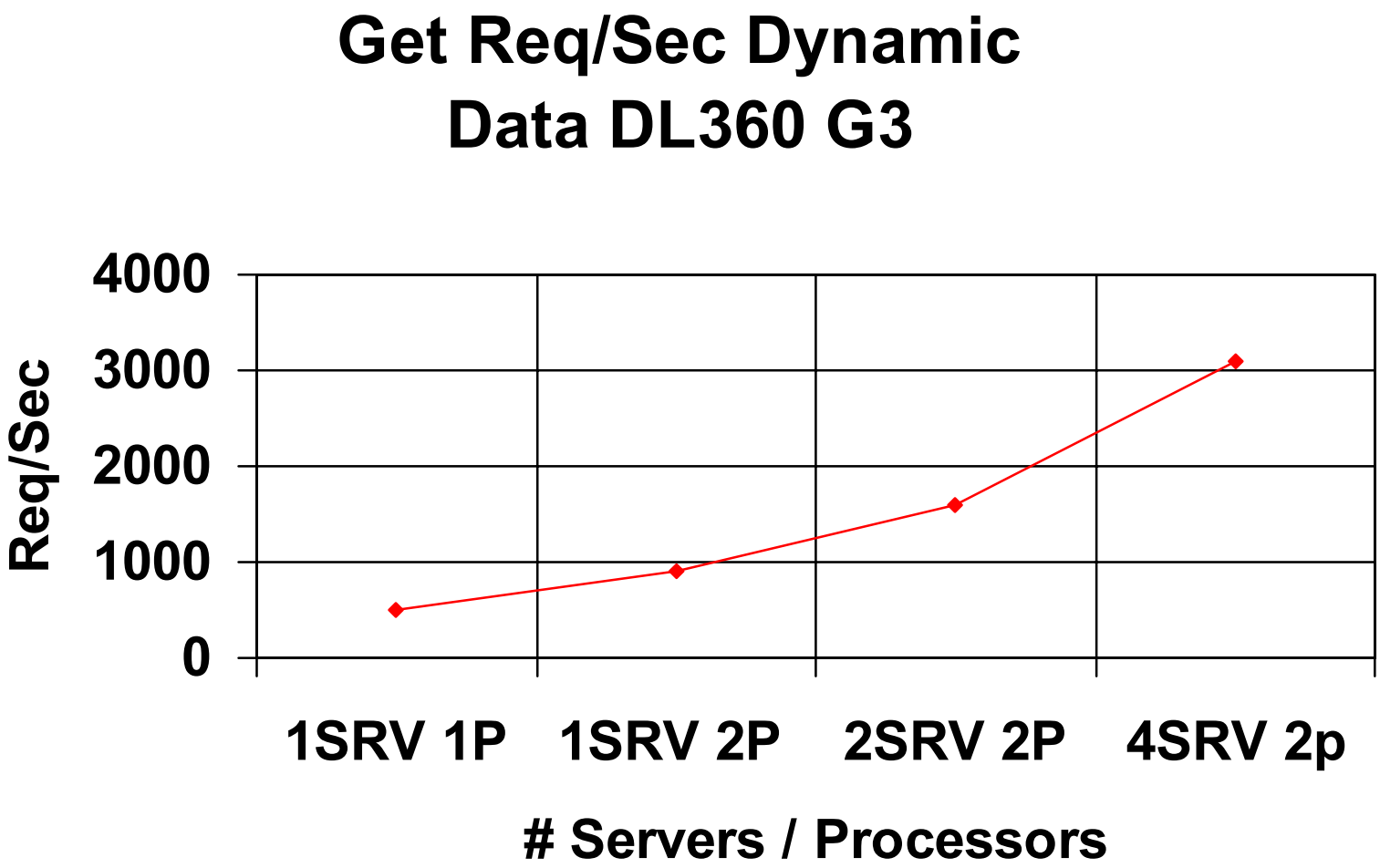
Performance Scalability: "The Numbers"



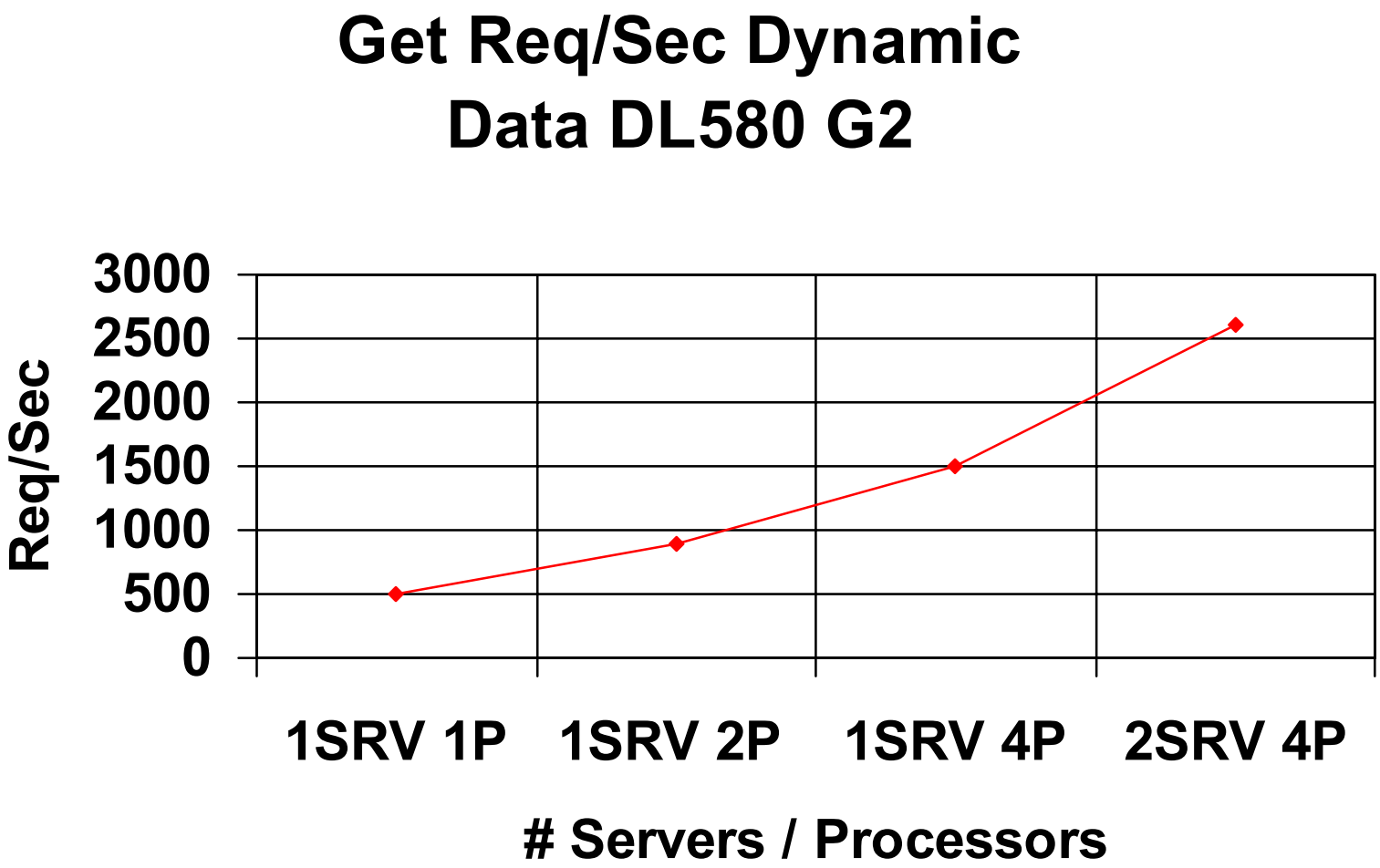
Performance Scalability: "The Numbers"



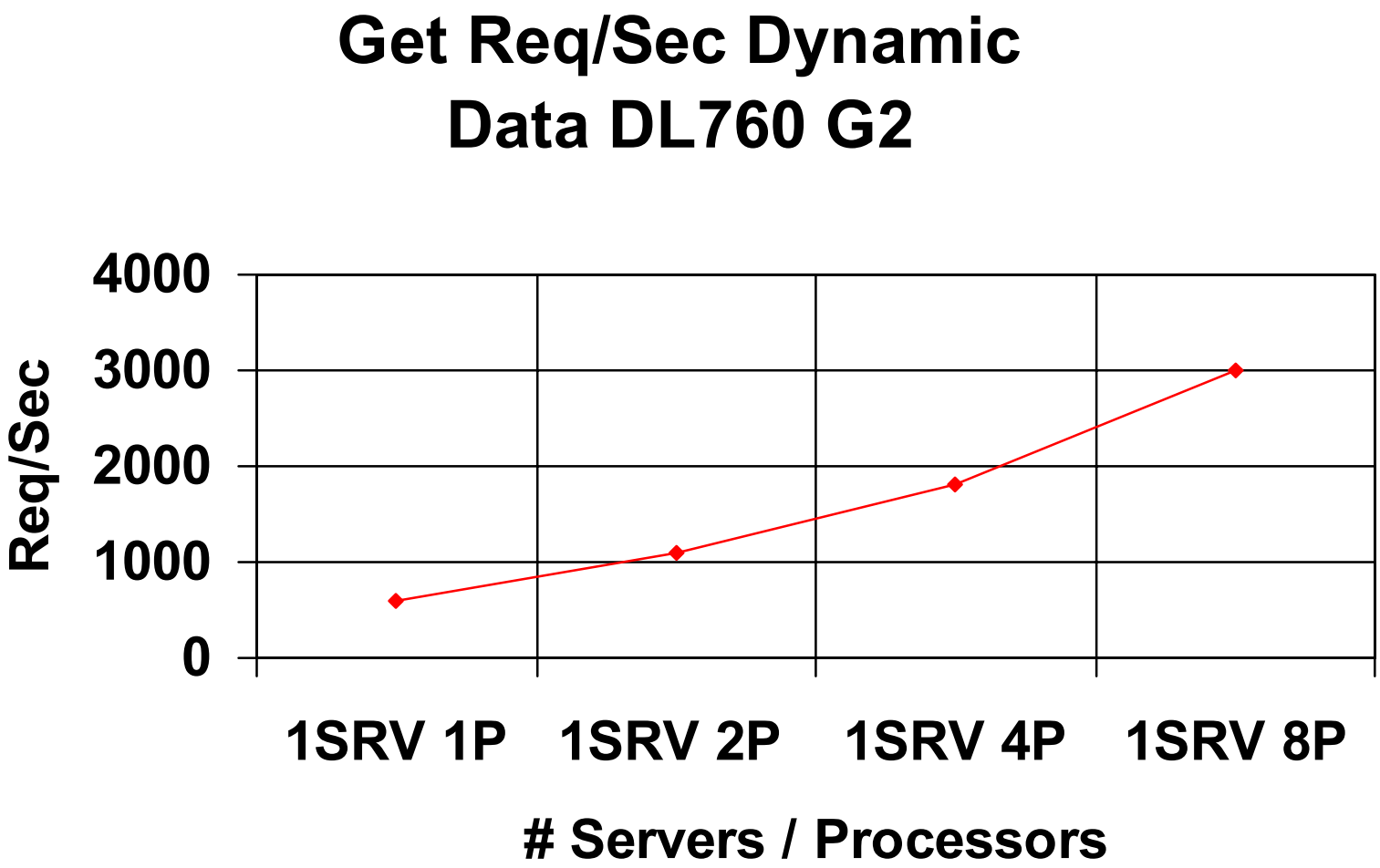
Performance Scalability: "The Numbers"



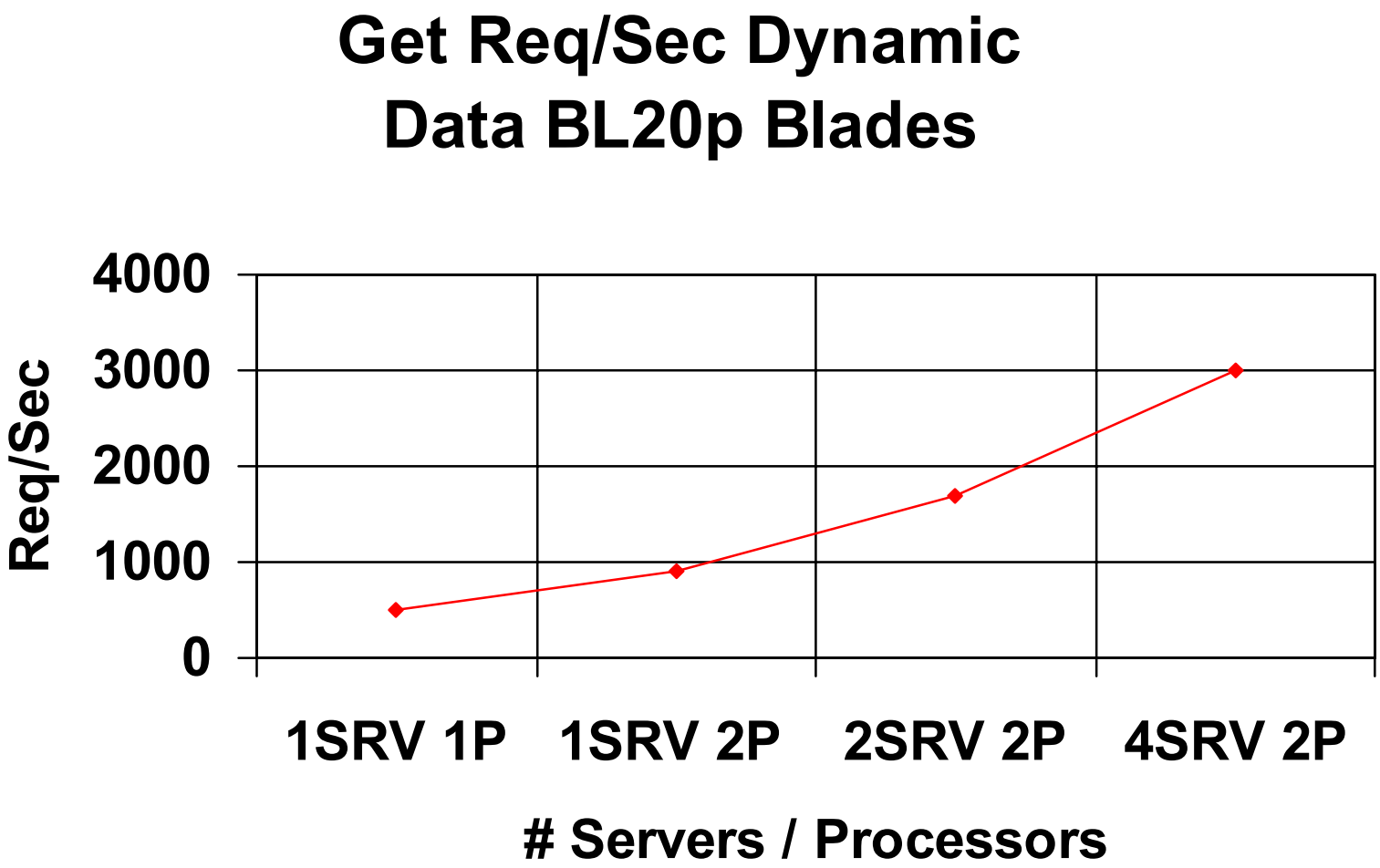
Performance Scalability: "The Numbers"



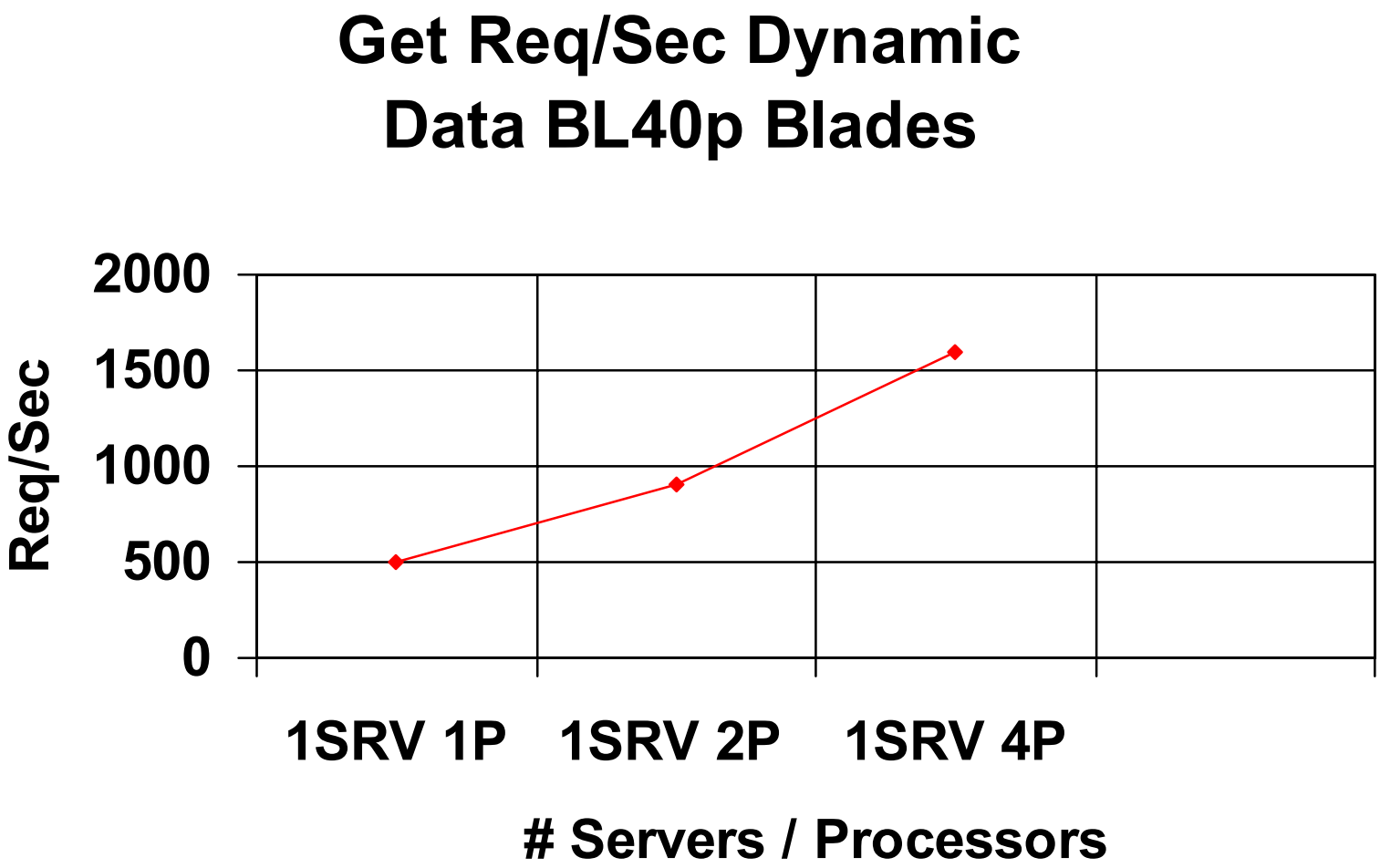
Performance Scalability: "The Numbers"



Performance Scalability: "The Numbers"



Performance Scalability: "The Numbers"



Performance Scalability Test Conclusions?

- IIS6 scales approximately 80-85% across properly load balanced servers (Horizontal Scaling). This is about the same from previous versions of IIS.

- IIS6 scales approximately 50-60% across multiple processors (Vertical Scaling). This is a big improvement over previous versions of IIS. (50-60% from 1-2 procs, 25-30% from 2-4 procs, and 5-10% from 4-8 procs.

The performance bottleneck is dependent on the type of content:

- **Dynamic Content:**
  1. **Processor (Faster is better. More is not necessarily efficient.)**
  2. **Memory (Cache as much content that is possible)**
  3. **Disk I/O**
  4. **NIC**

- **Static Content:**
  1. **NIC (NDIS Driver on the 1GB NICs maxed out at 35-40%)**
  2. **Memory**
  3. **Disk I/O (The more spindles the better)**
  4. **Processor**

Some Simple Recommendations:

- **Do not centralize your site content using a UNC share. (VERY BAD!!!)**
- **Use enough memory to enable you to cache your entire site + whatever the OS and other misc. apps running on the web server may need.**
- **When serving up static content, use multiple 1GB NICs**
- The more disk spindles the better (15K is Better than 10K)
- Use a device based load balancer when load balancing more than three servers. (Microsoft's NLB starts becoming a bottleneck after three).
- Use 2GB Fibre connectivity when connecting to a SAN

Questions?

Contact Information and links

Kevin P. Kenefic – kevin.kenefic@hp.com

Back Up Slides

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