IBM eServer J iSeries J

GP02 V5R2 Performance Update Part 1: Performance Update

The next generation iSeries...
simplicity in an on demand world

January 2003

ITSO iSeries Technical Forum

Agenda

- **Performance Update (Part 1)**
  - Hardware
  - Database Enhancements
  - HTTP web serving
  - Integrated xSeries Server File Backup
  - Miscellaneous Enhancements
- **Performance Management (Part 2)**
  - Management Central monitors and Graph History
  - Performance Tools for iSeries
  - BMC software, inc. Patrol for iSeries - Predict
  - Workload Estimator, PM eServer iSeries
  - Miscellaneous performance tools

For latest performance information, refer to http://www.ibm.com/eserver/iseries/perfmgmt

Notes: Agenda

The Performance Update is separated into two parts:

- Performance Update (facts, test results,tips, and so forth)
- Management (various performance measurement , sizing, and capacity planning tools)

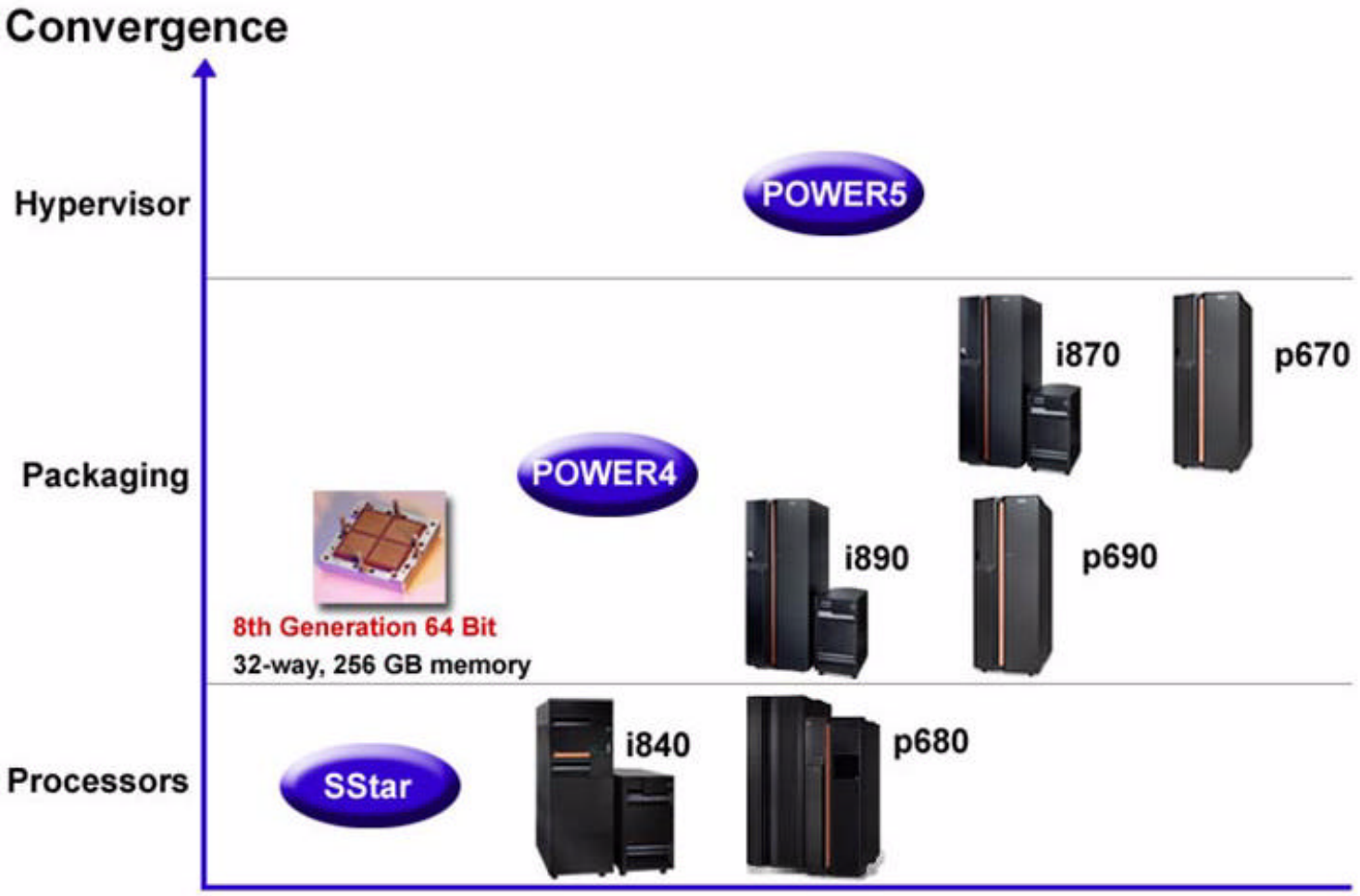
Note that in January 2003 PM/400 was renamed PM eServer iSeries.

Hardware

Hardware

- iSeries and pSeries Power Processor Convergence
- Processor Architecture
- Performance & Scalability
- New I/O enhancements

iSeries and pSeries Convergence



Notes: iSeries and pSeries Convergence

This foil reminds us of the iSeries and pSeries Powerx-based technology convergence and where we are in 2003.

Note that this convergence is a hardware-based convergence. The two primary operating systems, their own user interfaces and tools, along with applications will continue to build the strengths of each one.

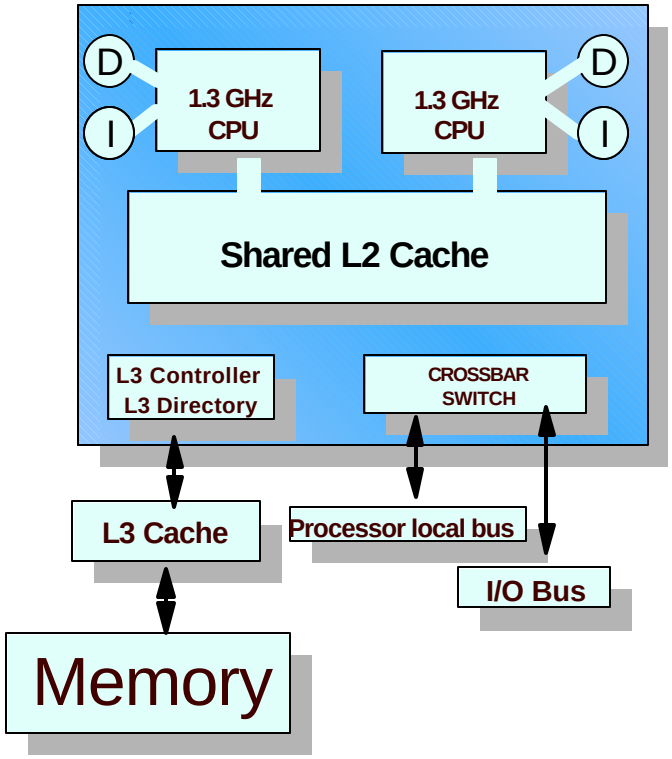
Integrating several operating systems on the same physical hardware has been a statement of direction that includes Linux (already available), AIX, and Windows operating systems (already available).

In fact, the POWER4 is the 8th generation of 64-bit processors brought to market by IBM on the AS/400 and iSeries since 1995.

The AS/400 and iSeries IStar and SStar processors such as the 250, 270, 820, 830, 840, and 890 (models introduced April 2002) continue to be available for specific price performance requirement such as would be satisfied by the 250 and specific Interactive Feature (5250 applications) requirements available in the 270, 820, 830, and 840 servers. They also support both OS/400 V5R2 and V5R1.

i825, i870, i890 with up to 32-way POWER4 Processors

- IBM POWER4 delivers unmatched iSeries growth on i825*, i870*, i890* 32-way
  - 37,400 CPW, up to 1.85 x growth over i840
  - All processor features enabled for Capacity Upgrade on Demand (permanent and temporary)
  - 1.3 GHz (i825 1.1 GHz) POWER4 microprocessors
    - 174 million transistors on a chip
  - Up to 534 MB of L2/L3 cache Per 8 processors MCM
  - 2X or more memory and I/O capacity
    - Up to 256 GB memory
    - **Up to 144 Terabytes disk****
    - **Up to 48 IXS, 60 IXA****
- Common IBM eServer processor roadmap for iSeries and pSeries J



* Requires V5R2
** i870, i890-2497/2498

Notes: iSeries POWER4 Microprocessor

POWER4 cannot be considered only a chip, but rather an architecture of how a set of chips are designed together to realize a system. As such, POWER4 can be considered a technology in its own right. In that light, systems are built by interconnecting POWER4 chips to form up to 32-way symmetric multiprocessors. The interconnect topology, referred to as a Distributed Switch, is new to the industry. Finally, no system discussion would be complete without some view of the reliability, availability and serviceability (RAS) features and philosophy incorporated into POWER4 systems. The RAS design is pervasive throughout the system and is as much a part of the design as is anything else.

The POWER4 design can handle a varied and robust set of workloads. This is especially important as the e-business world evolves and data intensive demands on systems merge with commercial requirements. The need to satisfy high performance computing requirements with its historical high bandwidth demands and commercial requirements with its data sharing and SMP scaling requirements dictated a single design to address both environments.

The chip as it is shown on this foil has two processors on board. Included in what we are referring to as the processor are the various execution units and the split first level instruction and data caches. The two processors share a unified second level cache (L2), also onboard the chip, through a Core Interface Unit. The Core Interface Unit is a crossbar switch between the L2, implemented as three separate, autonomous cache controllers, and the two processors. Each L2 cache controller can operate concurrently and feed 32 bytes of data per cycle. The Core Interface Unit connects each of the three L2 controllers to either the data cache or the instruction cache in either of the two processors.

The directory for a third level cache (L3) and logically its controller are also located on the POWER4 chip (L3 Directory Control in the picture). The actual L3 is on a separate chip. A separate functional unit, referred to as the Fabric Controller, is responsible for controlling data flow between the L2 and L3 controller for the chip and for POWER4 communication.

With the POWER4 architecture, this is also where an interface to a switch for clustering multiple POWER4 nodes can be attached.

A GX controller is the name for the chip that controls the I/O bus. It connects to either the MCM interconnect switch to other MCM modules (sometimes referred to as an "EADS boundary") as described above or to the external I/O chip. This external I/O chip supports up to four intermediate I/O buses (HSL), which connect to a HSL-to-PCI bridge chip.

Notes: iSeries POWER4 Microprocessor -2

A single processor on a chip has all of the L3 resources attached to the module, and the full L2 onboard the chip. Four such modules can be interconnected to form a 32-way SMP. To accomplish this, each chip has five primary interfaces. To communicate to other POWER4 chips on the same module, there are logically four 16-byte buses. Physically, these four buses are implemented with six buses, three on and three off.

IBM plans to continue to exploit its low-k technology employed in the 0.13 micron lithography process. IBM plans to increase processor frequencies to the 2+ GHz range while maintaining the system balance the current design offers. The current design introduces parallelism throughout the system so as to overcome the memory latencies resulting from high frequency operations. Enormous levels of bandwidth and concurrency contribute to superior performance across a broad range of commercial and high performance computing environments. These performance levels are achieved by a total system design that exploits IBM's leading technologies.

Some additional details on the i825 (1.1 GHz) and i870 (1.3 GHz) and i890 (2497/2498) processors:

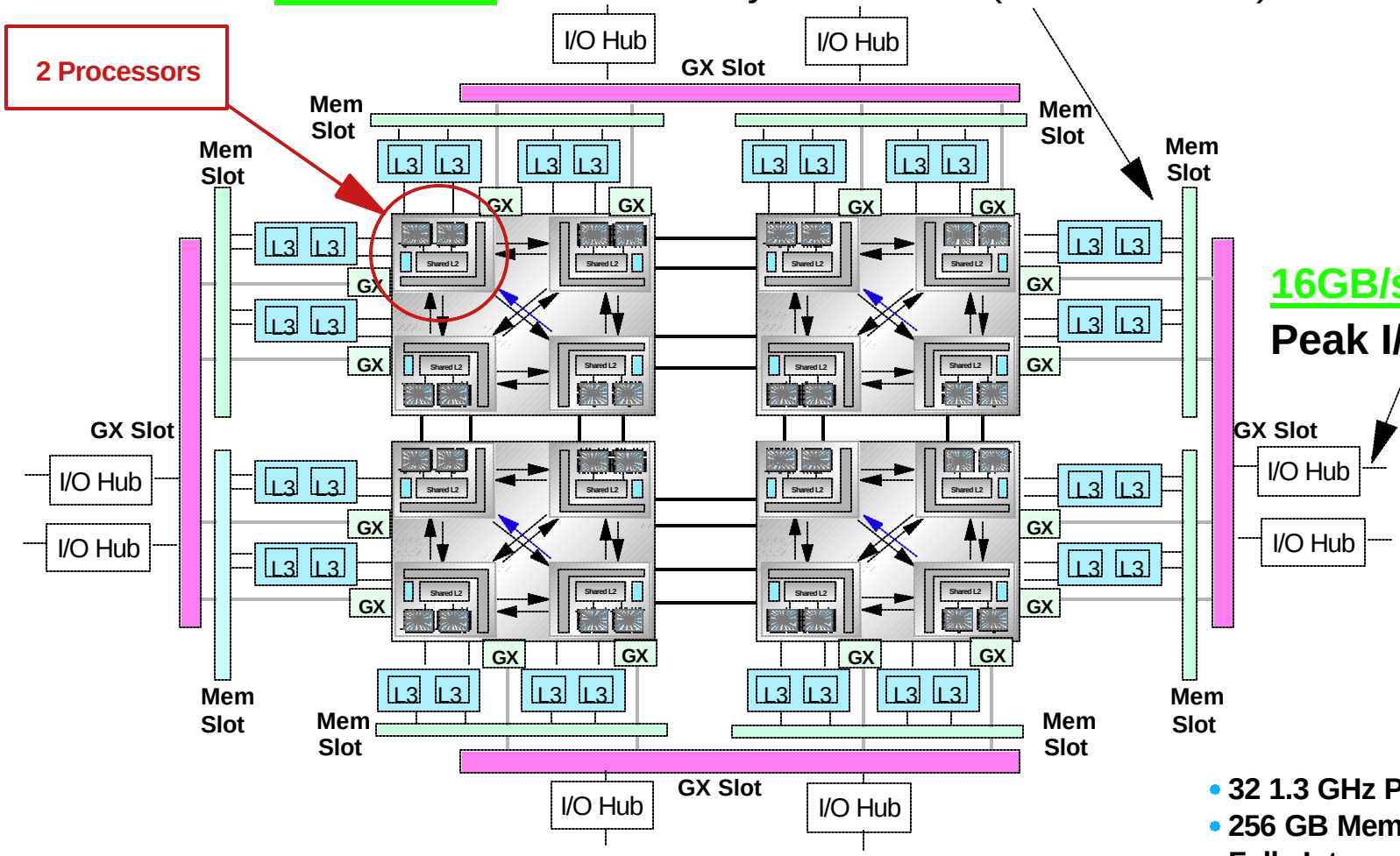
- Copper SOI Technology
  - 174 million transistors per chip
  - 0.18 micron lithography
- Superscalar
  - 8 Execution Units (pipelines)
  - Up to 5 parallel issues
- Cache Architecture
  - Total L2/L3 per MCP 534 MB
  - On chip L2 cache - 1.42 MB per 2 processors
  - Off chip L3 cache - 32 MB per 2 processors, 128MB per 8 processors (MCM)
- Packaging
  - 2 processors per chip
  - 8 processor Multi Chip Module (MCM)

Notes: iSeries POWER4 Microprocessor -3

- 1.3 GHz POWER4 microprocessors
  - 174 million transistors on a chip
- 2X or more memory and I/O capacity
  - Up to 256 GB memory (same as i890s announced April 2002)
  - **Up to 144 Terabytes disk on i870 and i890 (2497/2498): 2x i890s announced April 2002**
  - **Up to 48 IXS, 60 IXA on i870 and i890 (2497/2498): 32 on i890s announced April 2002**

iSeries i890 processors, memory, I/O band width

205 GB/sec Peak Memory Bandwidth (12.8GB/s x16)



16GB/sec (2GB/s x 8)
Peak I/O Bandwidth

- **32 1.3 GHz Processors**
- **256 GB Memory**
- **Fully Interconnected on Module**
- **4 Bus Ring Connecting Modules**
- **Distributed Memory & I/O**

Notes: POWER4 processors, memory, I/O band width

This foil shows pictorially represents where the MCM is placed relative to the memory slots and I/O attachments. You can see all the places the GX controller is placed - to control data flow between the MCM and L3 cache, and next to the I/O Hub connections.

All buses interconnecting POWER4 chips, whether or not on- or off-module, operate at half processor speed. As future technology is exploited, allowing chip size to decrease and operating frequencies to increase, system balance is maintained, since bus speeds are no longer fixed but are geared to processor frequency.

The multiple MCM configuration provides a worst-case memory access latency of slightly greater than 10% more than the best-case memory access latency maintaining the flat memory model, simplifying programming.

The eight-way MCM is the building block for the system. It is only available with four chips, each with its attached L3. A single processor on a chip has all of the L3 cache resources attached to the module, and the full L2 cache on the chip itself.

The next foils give some more internal performance "facts" on 8xx and 270 "capacities."

For additional information on POWER4 processor architecture, visit:

http:http://www1.ibm.com/servers/eserver/pseries/hardware/whitepapers/power4.html

http://www.research.ibm.com/journal/rd46-1.html

Notes: POWER4 processors, memory, I/O band width -2

Summary information for 270, 800, 810, 820, 825, 830, 840, 850, 890 fastest model feature cross-bar switch (processor-cache-memory) speed peak capacity is summarized in the following table on this page.

The 800 and 810 have the same bandwidth as the 270 - 8.2 GB/s using the same assumptions as below.

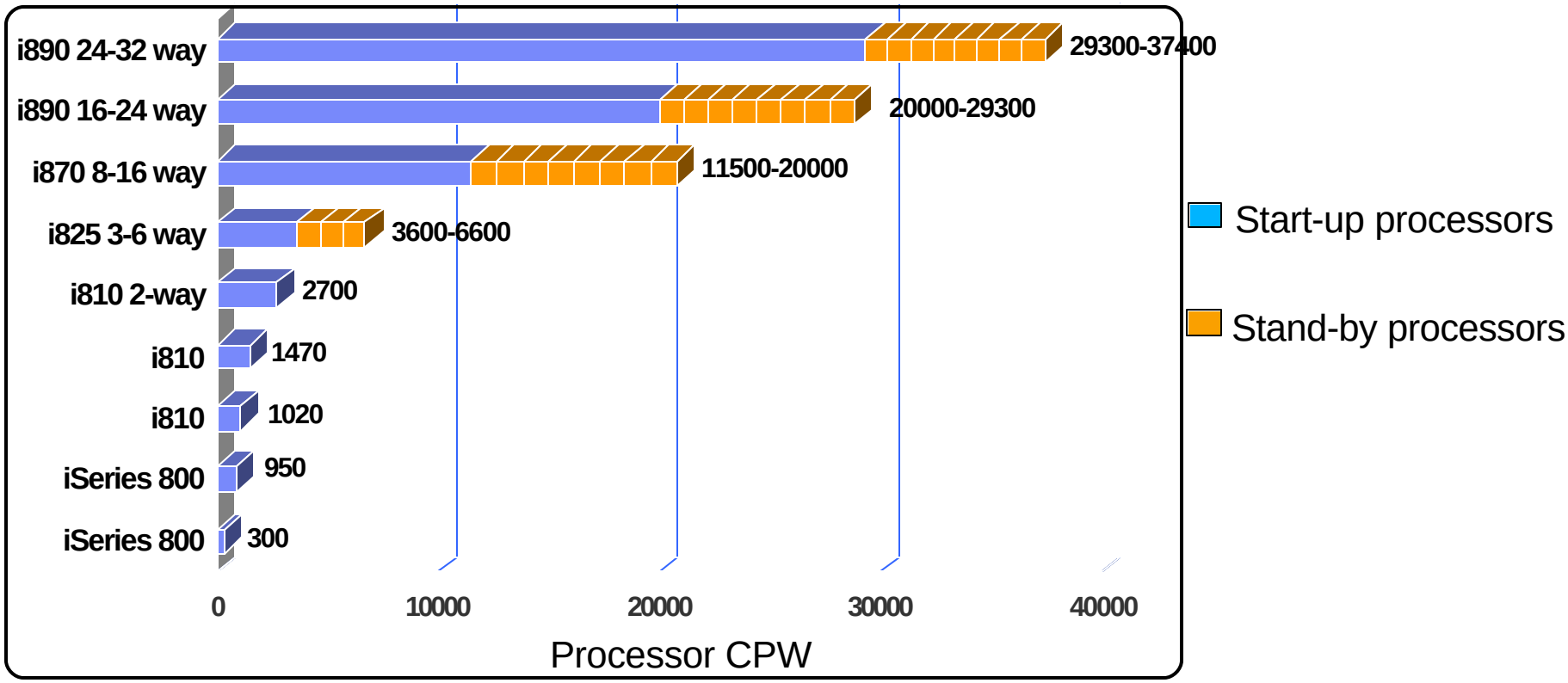
The 825 and 870 and 890 are Power4 systems and have a completely different memory system than the 270, 800, 810, 820, 830, and 840 systems. So the I/O "cross-bar switch" technology is a little different and any picture depicting the 840 cross-bar switch would be different.

Basically, on these models, the memory is distributed with each processor and there is an I/O connection to each processor chip as well. The interconnect of the three processor chips is three busses that form a ring (A to B, B to C, and C to A). If one adds up the bandwidth to the three pieces of memory, the three GX busses to I/O and the three fabric busses, one gets 39.6 GB/second. This is considerably higher than the iStar/SStar systems because these systems were built to support Scientific and Engineering workloads. (This bandwidth would probably only be used for the large government lab workloads like ASCI blue).

The 870 is similar, but has 8 chips, so there are 8 GX busses (8B @ 3:1) and 8 busses to memory (16B @ 3:1). The fabric between the chips is a little different and not explained here in a picture. Counting 8 busses (8B @ 2:1) between the chips out of the 32 total busses totals to 124.8 GB/s.

Model		Processor-Memory-I/O switching
270		8.2 GB/sec
iSeries 800		8.2 GB/sec
810		8.2 GB/sec
820		8.2 GB/sec
825		39.6 GB/sec
830		19.44 GB/sec
840		43 GB/sec
870		124.8 GB/sec
890		205 GB/sec

January 2003 iSeries iSeries 800, i810, i825, i870, i890



- Capacity Upgrade on Demand (temporary & permanent)
- Standard Edition ("Limited Administration 5250 CPW")
- Enterprise Edition ("Maximum 5250 CPW")

Note: Processor Commercial Processor Workload (CPW) values are used. CPW is a relative measure of performance of iSeries processors. Performance in customer environments may vary. The value is measured on maximum configurations.

A CPW and Pricing Tier View

2002

*		
890	**CPW**	**Tier**
24/32 Way	29300-37400	P50/**60**
16/24 Way	20000-29300	P50/**60**
840		
18/24 Way	16500-20200	P40/50
12/18 Way	12000-16500	P40/50
8/12 Way	9000-12000	P40/50
830		
4/8 Way	4200-7350	P30/40
2-Way	1850	P20/30
820		
4-Way	3700	P30/40
2-Way	2350	P20/30
1-Way	1100	P20/30
1-Way	600	P10/20
1-Way	370	P10/20
270		
2-Way	2350	P20
1-Way	1070	P10
1-Way	465	P10
1-Way	150	P05

2003

890	**CPW**	**Tier**
24/32 Way	29300-37400	P50
16/24 Way	20000-29300	P50
870		
8/16 Way	11500-20000	P40
825		
3/6 Way	3600-6600	P30
810		
2-Way	2700	P20
1-Way	1470	P10
1-Way	1020	P10
800		
1-Way	950	P10
1-Way	300	P05

* These processor features remain orderable throughout 2003. In most cases 800, 810, 825, 870, 890 (2497/2498) offer improved price/performance

A CPW and Processor Technology View

2002

- **890**
  - 24/32 Way (29300/37400 CPW)
  - 16/24 Way (20000/29300 CPW)
- **840**
  - 16/24 Way (16500/20200 CPW)
  - 12/16 Way (12000/16500 CPW)
  - 8/12 Way (9000/12000 CPW)
- **830**
  - 4/8 Way (4200/7350 CPW)
- **820**
  - 4-Way (3700 CPW)
  - 2-Way (2350 CPW)
  - 1-Way (1100 CPW)
  - 1-Way (370 CPW)
- **270**
  - 2-Way (2350 CPW)
  - 1-Way (1070 CPW)
  - 1-way (465 CPW)
  - 1-Way (150 CPW)

Medium to Large Enterprises

Small to Medium Enterprises

2003

- **890**
  - 24/32 Way (29300/37400 CPW)
  - 16/24 Way (20000/29300 CPW)
- **870**
  - 8/16 Way (11500/20000 CPW)

POWER4 processors
Capacity Upgrade on Demand
OS/400 V5R2 Required
5250 OLTP (0 or Max CPW)

- **825**
  - 3/6 Way (3600/6600 CPW)

- **810**
  - 2-Way (2700 CPW)
  - 1-Way (1470 CPW)
  - 1-way (1020 CPW)

SStar processors
OS/400 V5R2 Required
5250 OLTP (0 or Max CPW)

- **800**
  - Advanced (950 CPW)
  - Standard (300 CPW)
  - Value (300 CPW)

SStar processors
OS/400 V5R2 Required
5250 OLTP (Fixed CPW)

Notes: A CPW and Processor Technology View

This chart shows the performance lineup of the new servers, including the ratings of each processor added on a CUoD server, such as the i825, i870 and i890 (2497/2498).

Processor CPW by the Number of Processors

# Proc	CPW
1	300
1	950

iSeries 800

# Proc	CPW
1	1020
1	1470
2	2700

i810

# Proc	CPW
3	3600
6	6600

i825

# Proc	CPW
8	11500
16	20000

i870

# Proc	CPW
16	20000
24	29300
32	37400

i890-2497/2498

* These are estimates. Each processor increment is approximately = (total processor's CPW - start up processor's CPW) / number of stand by processors

Notes: Processor CPW by the Numbers of Processors

This foil give and indication of the incremental CPW values as stand-by processors are activated. Given the L2 and L3 cache sharing notes for the foil titled "POWER4 processors, memory, I/O band width," there will be processor activations, depending on 3-way to 6-way, or 8-way to 16-way, or 16-way to 24-way, or 24-way to 32-way, that will have slightly more or less "CPW per processor" improvements than those listed on this foil.

The general rule of thumb for estimating CPW increment per CUoD processor activation is to subtract the start-up CPW rating from the full activation of all stand-by processors and divide the result by the number of stand-by processors, as described at the iseries/ondemand website.

CPW values for each activated processors are not published. Therefore the formula give above is only an approximation.

System Performance Properties 800, i810, i825, i870, i890

Model	Processor Feature	Min-Max # Processors/ Technology / MHz / L2 - L3 in MB/Mem GB / Arms/ DASD	Processor CIW	5250 /Processor CPW	Mail/Calendar Users@70% CPU estimate
iSeries 800	2463	1 / 540 / 0 - 0 8 GB / 63 / 4.4 TB	NA*	25 300	NA*
iSeries 800	2464	1 / 540 / 2 - 0 8 GB / 63 / 4.4 TB	350	50 950	2900
i810	2466	1 / 540 / 2 - 0 16 GB / 108 / 7.6 TB	380	1020 0 / 1020	3100
i810	2467	1 / 750 / 4 - 0 16 GB / 108 / 7.6 TB	530	1470 0 / 1470	4200
i810	2469	2 / 750 / 4 - 0 16 GB / 108 / 7.6 TB	975	2700 0 / 2700	7900
i825	2473	3-6 / 1100 / 1.42 - 32** 48GB / 825 / 58 TB	1570 - 2890	3600 - 6600 0 / 3600 - 6600	8700 (3-W) 11600 (4-W) 17400 (6-W)
i870	2486	8-16 / Power4 / 1300 / 1.42 - 32** 128 GB / 2047 / 144 TB	5280 - 9100	11500 -20000 0 / 11500 -20000	29600 - 57600
i890	2497	16-24 / Power4 / 1300 / 1.42 - 32** 192GB / 2047 / 144 TB	8840 - 12900	20000 - 29300 0 / 20000 - 29300	57600 - 84100
1890	2498	24-32 / Power4 / 1300 / 1.42 - 32** 256 GB / 2047 / 144 TB	12900 - 16700	29300 - 37400 0 / 29300 - 37400	84100 - 108900

**Each processor chip contains two processors and 1.42 MB of L2 cache per chip; each MCM has 128 MB of L3 cache; each processor can access 32 shared or dedicated L3 cache, depending on the numbers of processors activated on each chip

The Domino Mail/Calendar number of users estimates are not formally validated NotesBench benchmark numbers. *NA means not available as of January 28, 2003. Refer to http://www.ibm.com/eserver/iseries/perfmgmt for latest information.

CPW is an internal iSeries benchmark used as a relative measure of performance among iSeries processors. CIW (Compute Intensive Workload) is another internal iSeries benchmark that is significantly more CPU intensive than the CPW benchmark. Performance in customer environments may vary. The values are measured on maximum configurations. If you have a CPU intensive application environment consider using the relative CIW rating rather than CPW rating among iSeries processors to size the server for acceptable performance. See the Performance Capabilities Reference manual for more information.

Notes: System Properties January 2003 8xxs

This chart gives an overview of the specifications for each iSeries server models announced January 2003. Subsequent charts list the older 8xx and 270 models which continue to be marketed at least through December 31, 2003.

CIW (Compute Intensive Workloads)

CIW, or Compute Intensive Workload, is a modeled projection based on the characteristics of workloads where:

- The majority of the processing time for the application is spent in the processor.
- Little or no I/O wait time, caused by, e.g. database activity, is considered in this workload. Instead, most time is spent on string handling or compute intensive application code.

The CIW rating is shown on this and the following foils even though CPW is the commonly used performance metric for sizing systems. The CIW values are developed by the Rochester performance measurement group do focus on CPU intensive workloads (high CPU utilization relatively few disk I/Os) compared to CPW which focuses on expected 5250-like work - utilizes moderate CPU and disk I/Os workloads.

The CIW is discussed later in this presentation relative to moving to a new release or processor technology.

Note that "NA" on this and any of the following foils means "Not Available when this presentation was published." See the Performance Management website for the latest information at:

- http://www.ibm.com/servers/eserver/iseries/perfmgmt

System Performance - i890 (features available May 2002)

i890 Base Processors

Model ----------	Orderable Processor FC#	Orderable Interactive FC#	Processor N-way / MHz / L2,L3 in MB* MemGB / Arms/ DASD	Processor CIW	Processor CPW	5250 CPW	Mail/Calendar Users@70% CPU estimate
i890 POWER4	0198	----------	32 / 1300/ 1.44, 32 256GB / 2047 / 72TB	16700	37400	----------	108900
i890 POWER4	0197	----------	24 / 1300/ 1.44, 32 192GB / 2047 / 72TB	12900	29300	----------	84100

i890 Standard Processors

Model ----------	Orderable Processor FC#	Orderable Interactive FC#	Processor N-way / MHz / L2,L3 in MB* Mem GB / Arms/ DASD	Processor CIW	Processor CPW	5250 CPW	Mail/Calendar Users@70% CPU estimate
i890 POWER4	2488	1576 - 1591	24-32 / 1300/ 1.44, 32 256GB / 2047 / 72TB	12900 - 16700	29300 - 37400	120, 240, 560, 1050, 2000, 4550, 10000, 16500, 20200 - 37400	84100 - 108900
i890 POWER4	2487	1576 - 1588	16-24 / 1300/ 1.44, 32 192GB / 2047 / 72TB	8840- 12900	20000 - 29300	120, 240, 560, 1050, 2000, 4550, 10000, 16500, 20200	57600 - 84100

**Each processor chip contains two processors and 1.44 MB of L2 cache per chip; each MCM has 128 MB of L3 cache; each processor can access 32 shared or dedicated L3 cache, depending on the numbers of processors activated on each chip

The Domino Mail/Calendar number of users estimates are not formally validated NotesBench benchmark numbers

CPW is an internal iSeries benchmark used as a relative measure of performance among iSeries processors. CIW (Compute Intensive Workload) is another internal iSeries benchmark that is significantly more CPU intensive than the CPW benchmark. Performance in customer environments may vary. The values are measured on maximum configurations. If you have a CPU intensive application environment consider using the relative CIW rating rather than CPW rating among iSeries processors to size the server for acceptable performance. See the Performance Capabilities Reference manual for more information.

System Performance Properties 840

i840 Base Processors

Model ----------	Orderable Processor FC#	Orderable Interactive FC#	Processor N-way / MHz / L2 in MB* MemGB / Arms/ DASD	Processor CIW	Processor CPW	5250 CPW	Mail/Calendar Users@70% CPU estimate
i840 SStar	0159	----------	24 / 600/ 16 128GB / 1080 / 38TB	10950	20200	----------	77800
i840 SStar	0158	----------	12 / 600 / 16 128GB / 1080 / 38TB	5700	12000	----------	40500

i840 Standard, Capacity on Demand Processors

Model ----------	Orderable Processor FC#	Orderable Interactive FC#	Processor N-way / MHz / L2,L3 in MB* MemGB / Arms/ DASD	Processor CIW	Processor CPW	5250 CPW	Mail/Calendar Users@70% CPU estimate
i840 SStar	2461	1540 - 1548	24 / 600/ 16 128GB / 1080 / 38TB	10950	20200	120, 240, 560, 1050, 2000, 4550, 10000, 16500, 20200	77800
i840 SStar CUoD	2354	1540 - 1548	18-24 / 600/ 16 128GB / 1080 / 38TB	8380 - 10950	16500 - 20200	120, 240, 560, 1050, 2000, 4550, 10000, 16500, 20200	59600 - 77800
i840 SStar CUoD	2353	1540 - 1547	12- 18 / 600 / 38B	5700 - 8380	12000 - 16500	120, 240, 560, 1050, 2000, 4550, 10000, 16500	40500 - 59600
i840 SStar CUoD	2352	1540 - 1546	8 - 12 / 600 / 38B	3850 - 5700	9000 - 12000	120, 240, 560, 1050, 2000, 4550, 10000	27400 - 40500

* L2 cache is per processor

The Domino Mail/Calendar number of users estimates are not formally validated NotesBench benchmark numbers

CPW is an internal iSeries benchmark used as a relative measure of performance among iSeries processors. CIW (Compute Intensive Workload) is another internal iSeries benchmark that is significantly more CPU intensive than the CPW benchmark. Performance in customer environments may vary. The values are measured on maximum configurations. If you have a CPU intensive application environment consider using the relative CIW rating rather than CPW rating among iSeries processors to size the server for acceptable performance. See the Performance Capabilities Reference manual for more information.

System Performance Properties 830

i830 Base Processors

Model ----------	Orderable Processor FC#	Orderable Interactive FC#	Processor N-way / MHz / L2 in MB* MemGB / Arms/ DASD	Processor CIW	Processor CPW	5250 CPW	Mail/Calendar Users@70% CPU estimate
i830 IStar	0153	----------	8 / 540 / 4 64GB / 630 / 22TB	3220	7350	----------	20910

i830 Standard, Capacity on Demand Processors

Model ----------	Orderable Processor FC#	Orderable Interactive FC#	Processor N-way / MHz / L2 in MB* MemGB / Arms/ DASD	Processor CIW	Processor CPW	5250 CPW	Mail/Calendar Users@70% CPU estimate
i830 IStar	2349	1531 - 1537	4 - 8 / 540 / 4 64GB / 630 / 22TB	3220	4200 - 7350	70, 120, 240, 560, 1050, 2000, 4550	20910
i830 IStar	2403	1531 - 1537	8 / 540 / 4 64GB / 630 / 22TB	3220	7350	70,120, 240, 560,1050, 2000, 4550	20910
i830 IStar	2402	1531 - 1536	4 / 540 / 4 64GB / 630 / 22 TB	1630	4200	70, 120, 240, 560, 1050, 2000	10680
i830 IStar	2400	1531 - 1535	2 / 400 / 2 64GB / 630 / 22TB	580	1850	70, 120, 240, 560, 1050	4490

* L2 cache is per processor

The Domino Mail/Calendar number of users estimates are not formally validated NotesBench benchmark numbers

CPW is an internal iSeries benchmark used as a relative measure of performance among iSeries processors. CIW (Compute Intensive Workload) is another internal iSeries benchmark that is significantly more CPU intensive than the CPW benchmark. Performance in customer environments may vary. The values are measured on maximum configurations. If you have a CPU intensive application environment consider using the relative CIW rating rather than CPW rating among iSeries processors to size the server for acceptable performance. See the Performance Capabilities Reference manual for more information.

System Performance Properties 820

i820 Base Processors

Model ----------	Orderable Processor FC#	Orderable Interactive FC#	Processor N-way / MHz / L2 in MB* Mem GB / Arms/ DASD	Processor CIW	Processor CPW	5250 CPW	Mail/Calendar Users@70% CPU estimate
i820 ** SStar	0152	----------	4 / 600 / 4 32GB / 237 / 8TB	1670	3700	----------	11810
i820 ** SStar	0151	----------	2 / 600 / 4 32GB / 237 / 8TB	840	2350	----------	6660
i820 ** SStar	0150	----------	1 / 600 / 2 32GB / 237 / 8TB	385	100	----------	3110

i820 Standard Processors

Model ----------	Orderable Processor FC#	Orderable Interactive FC#	Processor N-way / MHz / L2 in MB* Mem GB / Arms/ DASD	Processor CIW	Processor CPW	5250 CPW	Mail/Calendar Users@70% CPU estimate
i820 SStar	2438	1521 - 1527	4 / 600 / 4 32GB / 237 / 8TB	1670	3700	35, 70, 120, 240, 560, 1050, 2000,	11810
i820 SStar	2437	1521 - 1526	2 / 600 / 4 32GB / 237 / 8TB	840	2350	35, 70, 120, 240, 560,1050	6660
i820 SStar	2436	1521 - 1525	1 / 600 / 2 32GB / 237 / 8 TB	385	1100	35, 70, 120, 240, 560	3110
i820 SStar	2435	1521 - 1524	1 / 600 / 2 32GB / 237 / 8TB	200	600	35, 70, 120, 240	1620
i820 Pulsar	2395	1521 - 1524	1 / 400 / 0 4GB / 237 / 8TB	not available	370	35, 70, 120, 240	1600

* L2 cache is per processor

**Corresponding Dedicated Server for Domino models are 2458, 2457, 2456

The Domino Mail/Calendar number of users estimates are not formally validated NotesBench benchmark numbers

CPW is an internal iSeries benchmark used as a relative measure of performance among iSeries processors. CIW (Compute Intensive Workload) is another internal iSeries benchmark that is significantly more CPU intensive than the CPW benchmark. Performance in customer environments may vary. The values are measured on maximum configurations. If you have a CPU intensive application environment consider using the relative CIW rating rather than CPW rating among iSeries processors to size the server for acceptable performance. See the Performance Capabilities Reference manual for more information.

System Properties i270

i270 Base Processors

Model ----------	Orderable Processor FC#	Orderable Interactive FC#	Processor N-way / MHz / L2 in MB* MemGB / Arms/ DASD	Processor CIW	Processor CPW	5250 CPW	Mail/Calendar Users@70% CPU estimate
i270 ** SStar	2434	1516	2 / 600 / 4 16GB / 24 / 843GB	840	2350	0	6660
i270 ** SStar	2432	1516	1 / 540 / 2 8GB / 24 / 843GB	380	1070	0	3070

i270 Standard Processors

Model ----------	Orderable Processor FC#	Orderable Interactive FC#	Processor N-way / MHz / L2 in MB* MemGB / Arms/ DASD	Processor CIW	Processor CPW	5250 CPW	Mail/Calendar Users@70% CPU estimate
i270 SStar	2434	1520	2 / 600 / 4 16GB / 24 / 843GB	840	2350	70	6660
i270 SStar	2432	1519	1 / 540 / 2 8GB / 24 / 843GB	380	1070	50	3070
i270 SStar	2431	1518	1 / 540 / 0 8GB / 24 / 843GB	185	465	30	1490
i270 Pulsar	2248	1517	1 / 400 / 0 4GB / 24 / 843G	200	150	25	810

* L2 cache is per processor

**Corresponding Dedicated Server for Domino models are 2454, 2452

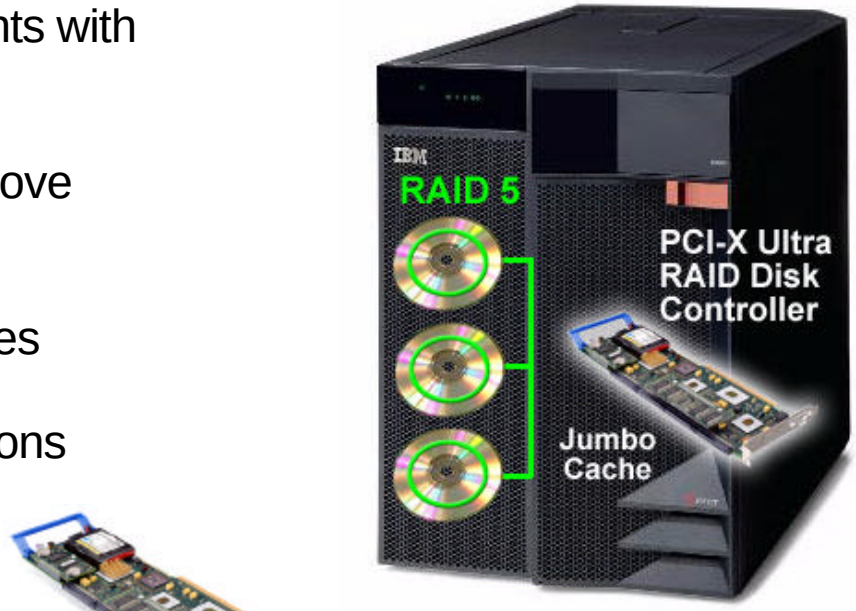
The Domino Mail/Calendar number of users estimates are not formally validated NotesBench benchmark numbers

CPW is an internal iSeries benchmark used as a relative measure of performance among iSeries processors. CIW (Compute Intensive Workload) is another internal iSeries benchmark that is significantly more CPU intensive than the CPW benchmark. Performance in customer environments may vary. The values are measured on maximum configurations. If you have a CPU intensive application environment consider using the relative CIW rating rather than CPW rating among iSeries processors to size the server for acceptable performance. See the Performance Capabilities Reference manual for more information.

I/O Performance

High Performance Direct Attach Storage Options

- Achieve up to 3X performance improvements with enhanced PCI-X I/O options*
- Reduce disk storage requirement and improve performance with IBM's optimized RAID-5
- New 35 GB and 72 GB 15K RPM disk drives
- New PCI-X I/O towers with rack mount options



New PCI-X Ultra RAID Disk Controllers	Write Cache in MB (Maximum)	Min / Max drives per RAID set
#2757 - High Performance**	757	3 /18
#2782 - Lower cost alternative for SME servers	40	3 / 12

* Maximum performance requires new PCI-X #2844 IOP, PCI-X # 2757 Ultra SCSI Disk Controller, and 15K Ultra SCSI Disk Units

** Note: Available February 2003 with 160 MB/second support

Notes: High Performance Direct Attach Storage Options

For maximum Disk performance (shown on the following foils), the following I/O features must all be installed. You can install several lesser combinations, but performance will be throttled back.

- Input Output Processor #2844
- Disk Controller #2757
- 15K rpm disks

The 2757 RAID controller using new RAID striping implementation that reduces the amount of disk storage to contain the striping information.

There are three options when you set up RAID configuration. The following is summary of the new RAID implementation

- *CAPACITY: If you select *CAPACITY, the arrays are reduced to favor disk capacity. For example, with 15 disk units - you will most likely get 1 array only.
- *BALANCE: With *BALANCE, depending on the configuration, the system will most likely configure 2 arrays (9 + 6) .
- *PERFORMANCE: With *PERFORMANCE, you get the configuration optimized for performance - therefore more arrays. In the case of 15 disk drives in I/O tower - the system will utilize 3 arrays (5 each in an array and assuming all disks are alike).

The January 2003 Hardware Technical Overview presentation contains additional RAID implementation information.

See also the DASD chapter in the January 2003 Performance Capabilities Reference manual, available at

- http://www.ibm.com/eserver/iseries/perfmgmt

2757 - x778 DASD Controller Comparison

Feature	2778 / 4778	2757	Improvements
SCSI bus	80 MB/s	**160 MB/s ***	**2x faster**
# SCSI buses	3	**4**	**1.25x more**
Max PCI Burst Rate	133 MB/s	**532 MB/s**	**4x more**
Processor Speed	80 MHz	**500 MHz**	**6.25x faster**
Compressed Write Cache	104 MB	**757 MB**	**7x larger**
Min/Max drives in RAID5 array	4 / 10 disks	**3 / 18 disks**	**Optimized**
SCSI bus tagged command queuing	N/A	**Yes**	**Faster Response Time (under heavy I/O load)**
Array parity checking and memory scrubbing	Yes	**Yes - New HDW assist**	**5x faster**
RAID Configuration	Enable or Disable	**Capacity, Performance, Balance or disable**	**Greater Flexibility**

* NOTE: Shipped February 2003 with 160 MB/second support.

Notes: 2757 - x778 DASD Controller Comparison

The DASD controller plays an important role in determining the I/O busy time for a given process. Each time am I/O request has to be staged to online storage, the DASD controller is one of the prime components to provide the components and the code to provide persistence of the data between main storage and online storage. The key elements that reduce the time spent for this operation are:

- **Bandwidth** - the new 2757 controller has a 8.8 time larger bandwidth on the PCI bus and 3.3 on the SCSI bus than the previous controller, meaning that transfer of large blocks or transfer of multiple small blocks will improve its speed.
- **Processing capacity** - the use of a Power PC processor, running at 500 MHz and 256K of L2 cache allows faster execution (6.25 times) of all storage transfer commands, including the parity checking and data scrubbing.
- **Cache memory** - the performance of I/O commands can greatly benefit from staging data into temporary RAM, especially if the memory has enough capacity to store large segments. Using data compression allows to store write buffers in cache memory of up to 757 MB, in comparison with the 104 MB on the x778 controller.
- **Number of SCSI buses** - a 2757 now has support for 4 SCSI buses (3 for internal DASD, one for external removable media), which allows to better distribute the disk units over the controller, thus optimizing the data flow.
- **RAID support** (see also the *Hardware Overview* of the January 2003 Technical Overview set of presentations) - optimized data and parity striping using subarrays greatly improves disk response time since the updates of both data and parity buffers runs more efficiently and requires no longer seeks between both stripes: intermixing data stripes with parity data, especially for update/write functions will yield in a considerable gain of performance. At this time, we have no information about the difference in throughput between the settings (Balance, Performance, Capacity) for a RAID configuration.

DASD Specifications

DASD Feature	Size (GB)	Speed (RPM)	Read Seek Time (ms)	Write Seek Time (ms)	Latency (ms)	Interface Speed (MB/s) in 5074	Interface Speed (MB/s) in 5094
4317 / 6717	8	10K	5.3	6.3	3	80	80
4318 / 6718	18	10K	4.9	5.9	3	80	80
4319 / 6719	35	10K	4.7	5.3	3	80	160
4326 /	35	15K	3.6	4.0	2	Not Supported	160*
4327	70	15K	3.6	4.0	2	Not Supported	160*

* Note: Shipped February 2003 with 160 MB/second capacity

Notes: DASD Specifications

This chart highlights the major technical changes from the existing 10K RPM disks as compared with the new 15 K RPM drives. Read/write seek times have improved with a factor of 25 to 32 percent, while the average latency improves also by 33%.

The interface speed has one to 160 MB/second speed (February 2003 availability hardware level) thus reducing a possible bottleneck while transferring large blocks of data, as typically generated by heavily I/O oriented batch jobs, copy or save/restore operations.

The following foils are a series of "different looks" at various supported combinations of existing disk controllers, existing 10K RPM disks with new disk controller and new 15K RPM disks - highlight the advantages of these new disk I/O capabilities.

Notes:

All the following disk controller and disk performance charts are based upon the CPW workload, which is primarily an interactive workload that would not take maximum advantages of the improved controller processor speed, write cache and disk RPM speed.

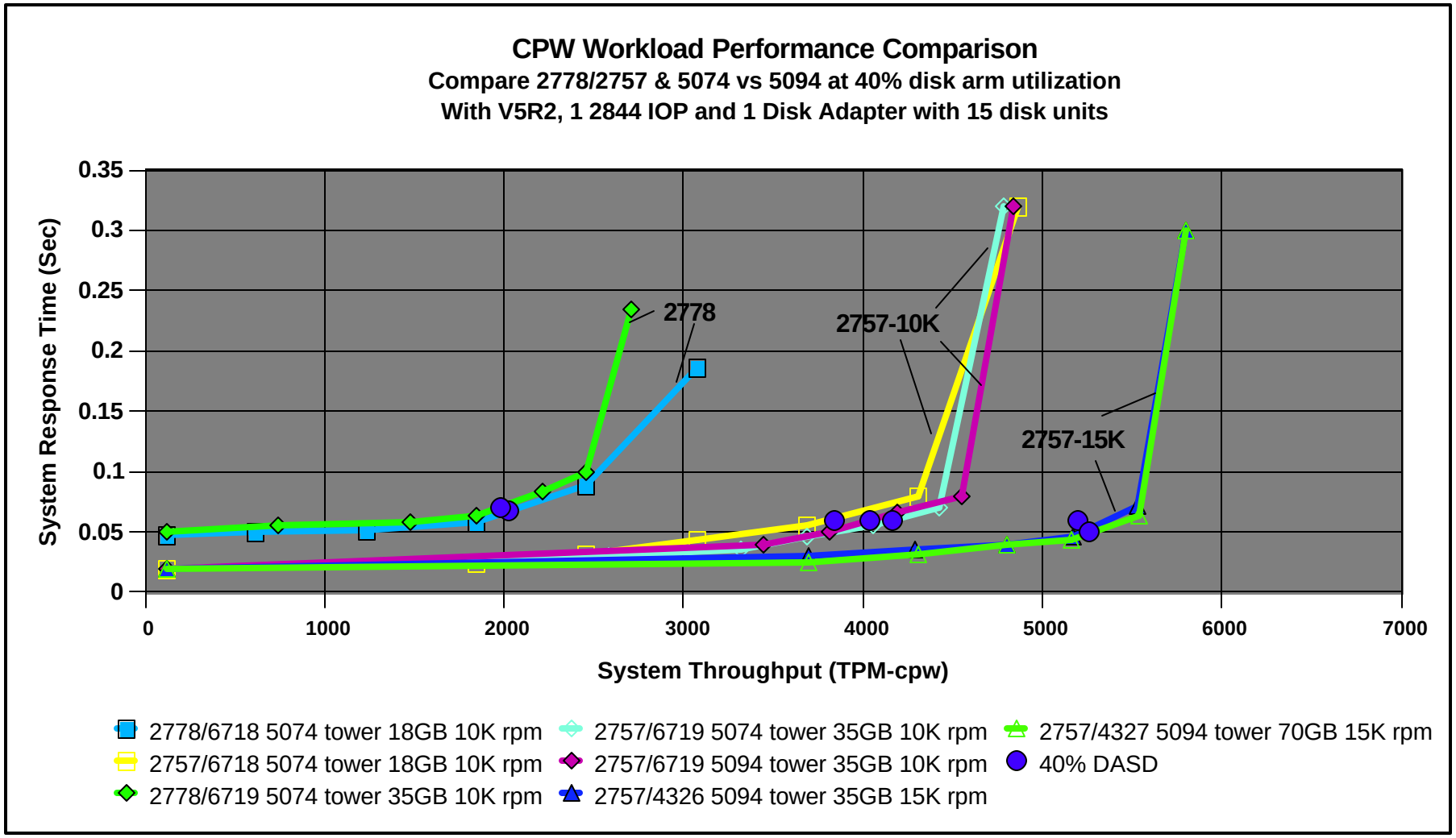
All performance numbers shown were at disk percent busy of 40% (shown as purple circle in the graph).

At the time this presentation was developed there were no test results available with batch applications. Consult the performance management website for current information:

- http://www.ibm.com/servers/eserver/iseries/perfmgmt
- See also the DASD chapter in the January 2003 Performance Capabilities Reference manual, available at this website

Disk Controller Comparison: 2778/4778 <-> 2757

Includes old, new I/O towers



Notes: Disk Controller Comparison: 2778/4778 <-> 2757

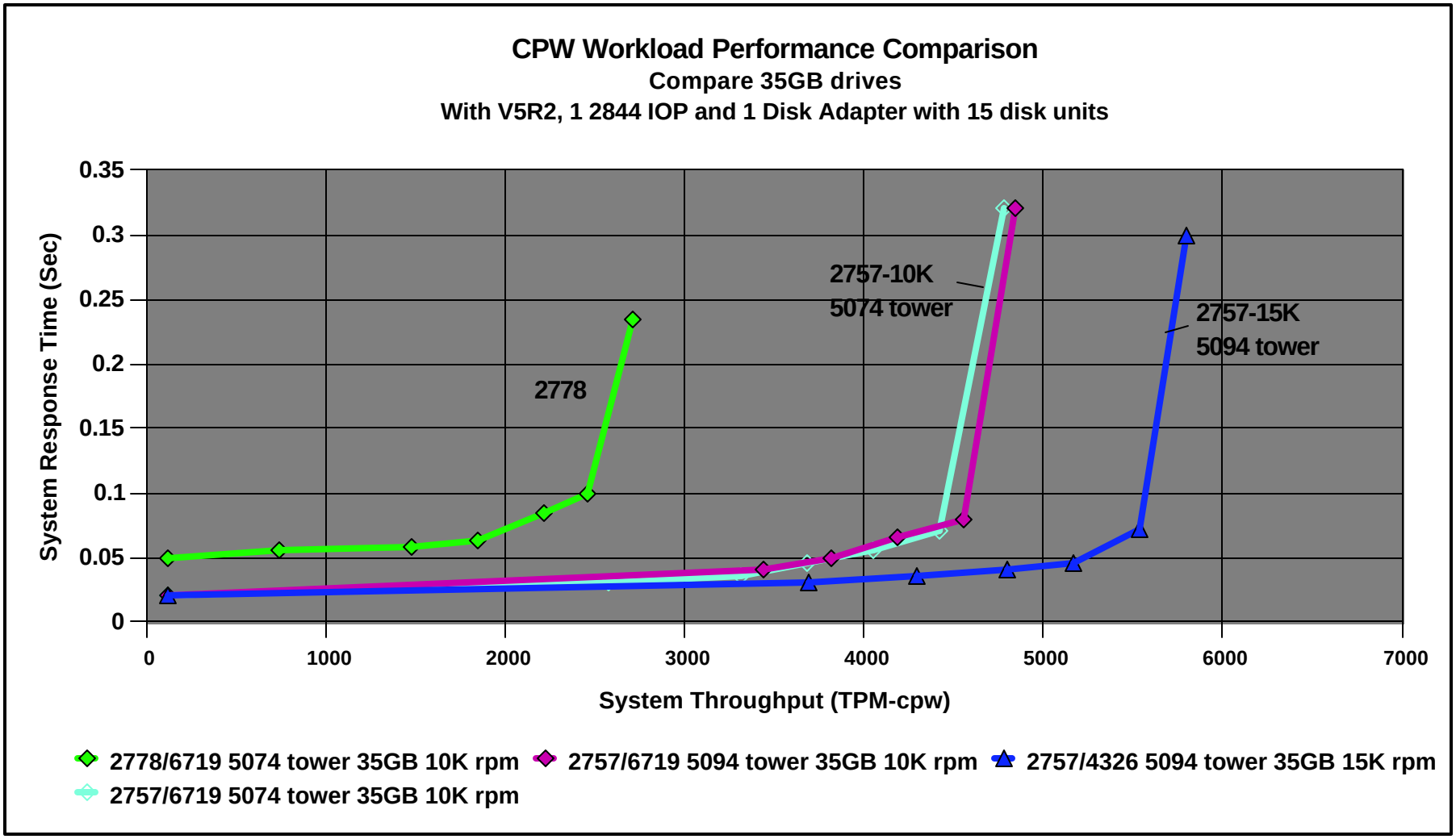
This foil compares the existing 2778/4778 disk controller with the new 2844 IOP, 2757 disk controller and combinations of 10K RPM and 15K RPM disks, along with the I/O towers in which they can be attached. The best performance results were conducted with one 2757 attached to the 2844 in this foil and the following foils.

When combining the new Input Output Processor (IOP) #2844 and new 2757 with the new 15K RPM disk drives, you can achieve a global gain of performance up to three when running an identical workload, such as CPW, on a given configuration. These results reflect a simulated workload on a 8xx-xxxx with xx GB of main storage, running the CPW workload on a set of disks in either a 5074 or 5094 tower, using a single IOP, one RAID controller and 15 disk units

Progressing from left to right, you can see the System (Disk) Response time and the CPW metric progress from the existing 2778 controller and 10K RPM disk drives, improve with the new 2757 disk controller with 10K RPM disk drives, and improve further with the combined new 2757 controller and 15K RPM disk drives. This is as you would expect. However, we note that CPW transaction throughput improvement is quite dramatic even with the existing 10K RPM disk drives on the new 2757 controller.

The following performance charts "breakout" different performance metrics combinations (disk controller, disk RPM speeds, and I/O Tower information included in this chart.

Performance Comparison: 35GB Disks



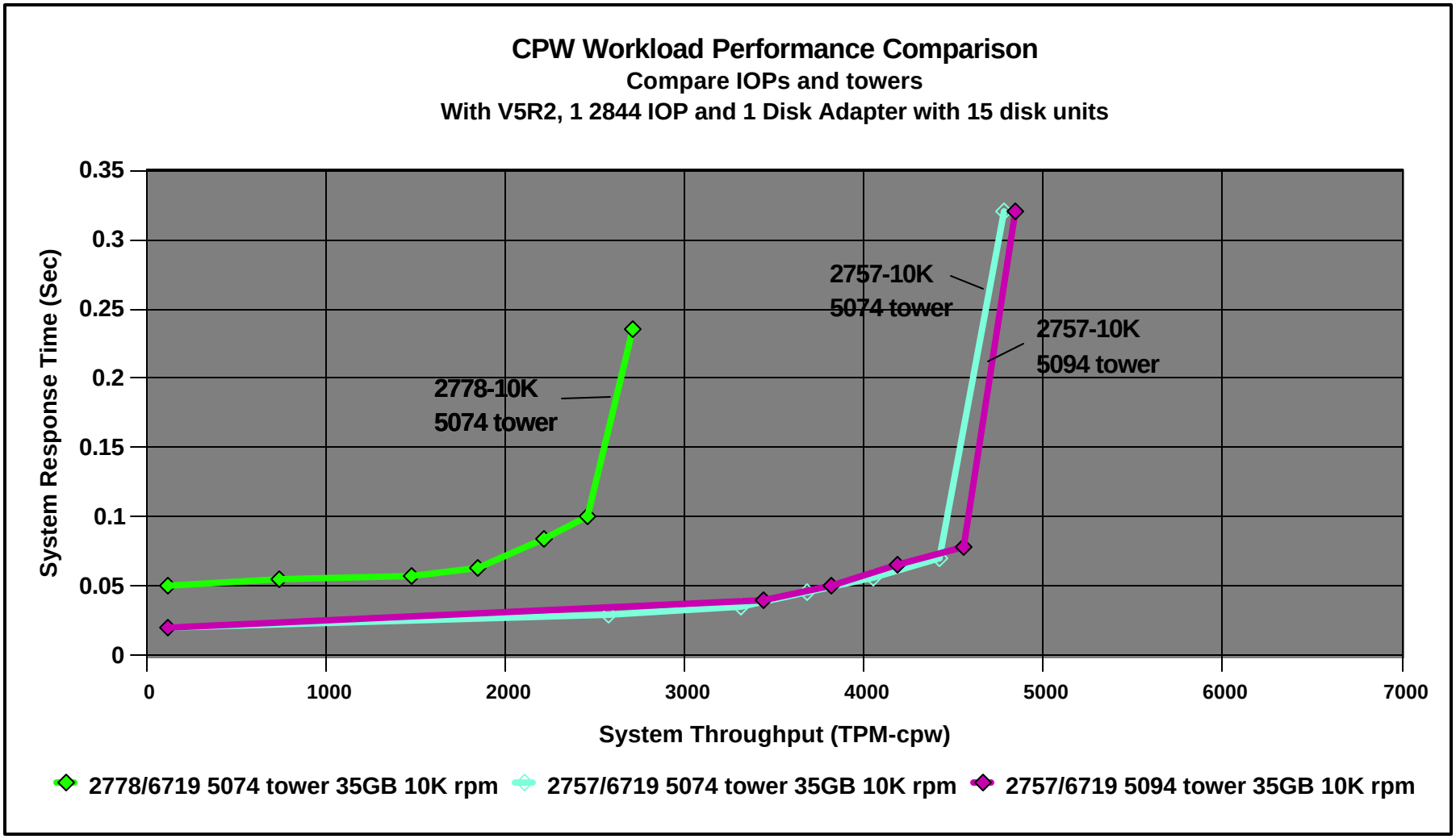
Notes: Performance Comparison: 35GB Disks

In this chart we separated out the CPW performance graph lines from the original "all combinations" foil to show:

- the 35 GB 10K RPM drives attached to the 2778 controller
- the 35 GB 10K RPM drives attached to the new 2757 in the 5074 I/O tower
- the 35 GB 10K RPM drives attached to the new 2757 in the 5094 I/O tower
- the 35 GB 15K RPM drives attached to the new 2757 in the 5094 I/O tower

This chart demonstrates the clear performance improvement in #2844 IOP/#2757 controller (IOA) and the 15K RPM drives under the CPW workload.

Performance Comparison: 2778<->2757, I/O Towers



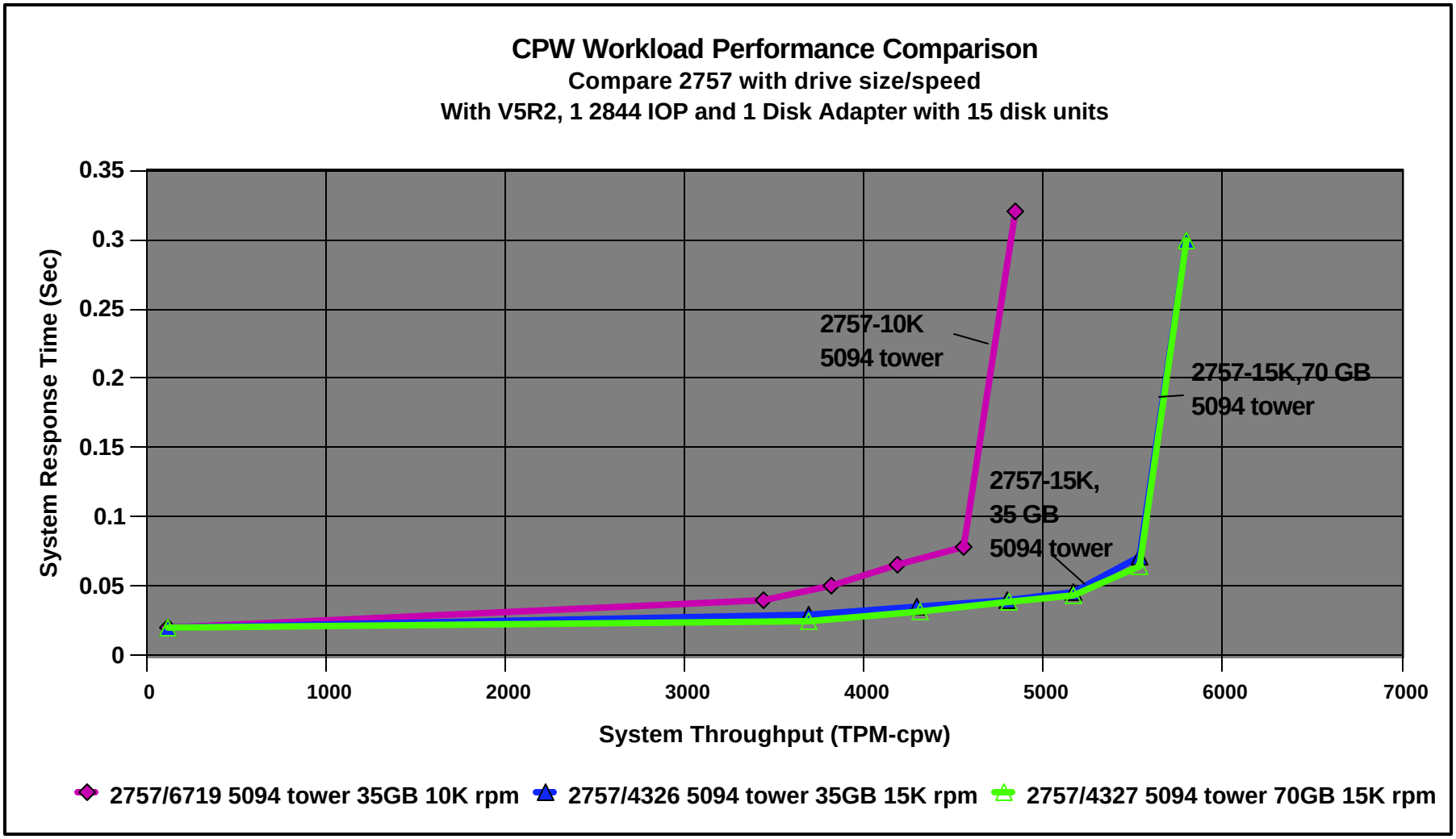
Notes: Performance Comp: 2778<->2757, I/O Towers

In this chart we separated out the CPW performance graph lines from the original "all combinations" foil to show:

- 2778 with 35GB 10K RPM disks in 5074 Tower
- 2757 with 35GB 10K RPM disks in 5074 Tower
- 2757 with 35GB 10K RPM disks in 5094 Tower

This chart highlight the measurable performance improvement with the 5094 tower using the new 2844 IOP/2757 disk controller under the CPW workload.

Performance Comparison: 2757 with 10K, 15K RPM Disks



Notes: Performance Comp: 2757, 10K, 15K RPM Disks

In this chart we separated out the CPW performance graph lines from the original "all combinations" foil to show:

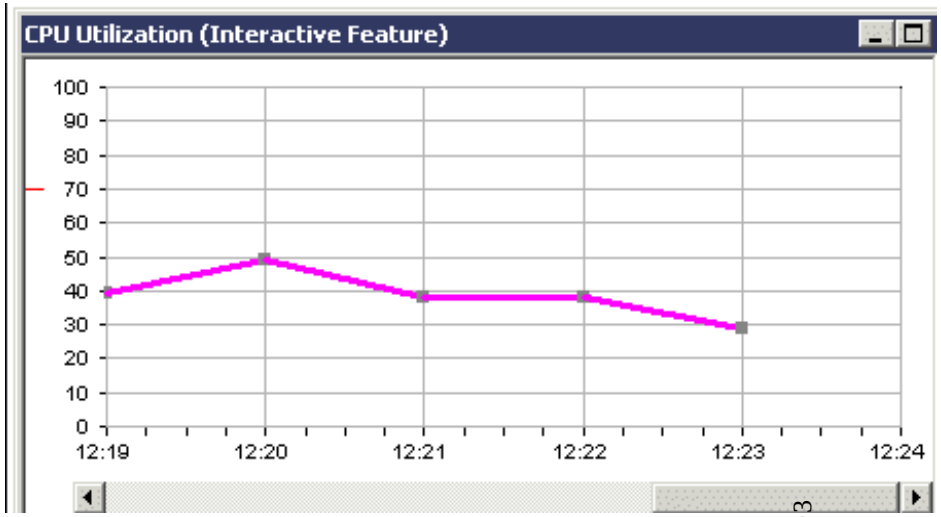
- 2757 with 35GB 10K RPM disks in 5094 Tower
- 2757 with 35GB 15K RPM disks in 5094 Tower
- 2757 with 70 GB 15K RPM disks in 5094 Tower

This chart highlights the measurable performance improvement on the #2844 IOP/#2757 controller (IOA) between 10K and 15K RPM disks and shows essentially no difference between the 35GB and 70 GB 15K RPM disk drives under the CPW workload.

5250 OLTP CPW and special workloads update

5250 OLTP work definition

- **Any task that uses a 5250 Data Stream is considered 5250 OLTP work and requires some amount of 5250 CPW to process. . . no matter how the task was started**
  - A task submitted through a 5250 session (5250 device, 5250 emulation, 3270 remote attach) that does display or printer I/O requires 5250 CPW
  - A batch job that does display or printer I/O requires 5250 CPW
  - A task submitted through a 5250 session (5250 device, 5250 emulation, 3270 remote attach) as a "batch" job is not considered 5250 OLTP work and will not require any 5250 CPW unless the task does display or printer I/O
  - Zero 5250 OLTP CPW is available with the **new Standard Package or on older Base (0 Interactive CPW)** models for a system administrator who wishes to use 5250 display device I/O to manage various aspects of the server.
    - **Caution: Multiple administrative jobs will quickly exceed this capability**



Notes: 5250 OLTP work definition

With the January 2003 announcement, there is increased interest in "defining" the type of work performed by a 5250 On-Line Transaction Processing OLTP) application.

From this point forward the iSeries will focus on using the term "5250 OLTP," rather than use of the term "interactive" - when discussing CPU utilization and system resource consumption, as "interactive" is considered a more generic description of an application interface characterized by user input and the associated response. Thus, "interactive" can refer to the familiar 5250 keyboard input and display device response, but also can refer to PC workstation-based dialogues such as through a Windows operating system interface, a Domino mail and calendaring interface, or a web browser interface.

Existing V5R2 (and earlier) system screens and V5R2 and older release documentation still use the interactive term generally to mean 5250 applications. This foil uses the term 5250 OLTP and clearly states what causes a job or thread to be considered as 5250 OLTP - which is tied directly to the Interactive Feature and 5250 CPW ratings. Jobs doing 5250 I/O operations but with middleware that translates that 5250 datastream to an HTTP-based datastream understood by browsers, remain classified as 5250 OLTP jobs. This includes "screen scrapers" and WebSphere Host Publisher data exchange, and WebSphere Development Studio (WDS) Client for iSeries WebFacing server on all iSeries and AS/400 systems **except the iSeries 800, i810, i825, i870, and new i890 (features 2497, 2498) models.**

The WebFacing implementation (only) is not considered as 5250 OLTP CPW work when running on the iSeries 800, i810, i825, i870, and new i890 (features 2497, 2498) models. These are termed the January 2003 models.

Remember, the January 2003 announcement includes the Enterprise Edition. This edition offers as much 5250 CPW power as the active number of processors on a specific model permit. For example, assuming you have an 825 and with 3 Start up processors and have purchased the Enterprise package your 5250 jobs have access to the power of all 3 startup processors, subject to any partial processor assignment per logical partition.

When you take advantage of Capacity Upgrade on Demand additional processors, your 5250 applications have full access to the total additional processor power, again limited by any partition and partial processor assignment constraints.

For the January 2003 new models with a Standard Edition, as well as the previously available "base" 270, 820, 830, 840, 890 models, the single "administrative job" exception continues to apply. That is, a single active 5250 OLTP job can use significantly more CPU resource than one would expect, given a "zero interactive" CPW rating. Since OS/400 does not enforce any "single job active" requirement, multiple administrative jobs active and consuming significant CPU utilization could quickly exceed this capability.

Notes: 5250 OLTP work definition -2

Note that these definitions regarding 5250 OLTP include:

- java programs started from a 5250 session **as not being counted as a 5250 OLTP task** since the Java program actually runs in a separately "spawned" job within the same subsystem that 5250 job is running in.
- Printer work that passes through twinax media is also considered 5250 work - even though there is no "user interface". This is true regardless of whether the printer is working in dedicated mode or is printing spool files from an output queue.

- Chapter 2 in the Performance Capabilities Reference Guide explains more about 5250 OLTP workloads. The V5R2 Performance Capabilities Reference Guide (.pdf format) is accessible via the InfoCenter or through the iSeries Performance Management web page at:
  - http://www-1.ibm.com/servers/eserver/iseries/perfmgmt/
  - Select ---> Resource Library

On this foil we use the System monitor for the Interactive Feature metric example as a way to monitor Interactive Feature (think 5250 OLTP CPW rating) utilization real-time. Collection Services also includes this metric in its performance data collection. Performance database files (QAPMcccc) generated from Collection Services data can be queried or used as input to the Performance Tools for iSeries licensed program 5722-PT1, which can produce report information that includes Interactive Feature utilization.

Interactive feature utilization is the only iSeries performance metric that includes all jobs and tasks that perform I/O operations to 5250 displays and printers. For example, Work with Active Jobs (WRKACTJOB) command and Performance Tools for iSeries, 5722-PT1, reports identify as "interactive" those jobs from signed on 5250 display workstations. Submitted jobs that perform 5250 display or printer I/O operations would appear as "batch" and a spool writer job appears as a "writer" job.

Migrating to a new release or higher processor rating

Migrating to a new release and not using the i810, i825, i870, i890 Enterprise Package

Before moving:

- Consider CIW rating as well as CPW rating. CIW would be more important for CPU intensive workloads (5250 or non-5250)
- Collect Collection Services performance data, especially if running high CPU utilization for 5250 interactive workloads
- Review Interactive Feature Utilization (Performance Tools for iSeries, 5722-PT1 System Report or Management Central System Monitors or PM/400 since V4R5)
  - **Total CPU Utilization (Interactive Feature) . .: 50 OR 70% makes a difference**
- Use Collection Services collection objects and / or PM eServer iSeries data for the performance critical time period
  - Examples: frequent peaks in interactive workload or batch job run time during night

Move to software release first if going to upgrade to new processor and collect, review performance data

Notes: Migrating to a new release, higher processor rating

This foil applies to all iSeries systems **except those announced January 2003 (that is, the iSeries 800, I810, i825, i870, i890 (#2497/#2498)) that are ordered with the Enterprise Edition package, which enables all of the processor CPW to be available to any combination of 5250 OLTP and other applications.**

This graph shows the new interactive work management algorithm used on the 270 and 8xx servers starting with V4R5. Essentially when the microcode determines that your 5250 work is achieving its rated interactive performance - at the vertical bar (Max/Knee) the 5250 work does not continue to take available CPU from the non-interactive (Client/Server) work on the server. Instead the 5250 work contends with itself at this time. Over time, as the 5250 work consumes lower CPU utilization, the OS/400 Dynamic Priority Schedule support (active starting with RISC V3R6) 5250 work will be managed normally.

In real-life scenarios the system does tolerate high CPU utilization by 5250 jobs before putting into effect the algorithm discussed here. That is, the system permits 5250 OLTP CPW utilization to be higher than your interactive feature for a period of time before downgrading the priority of the 5250 jobs doing this work. Monitoring this activity, the system relatively quickly detects 5250 work that lowers to "within 5250 CPW capacity level" and readjusts the job run priority value back to its original setting. Extended peaks of high CPU utilization by 5250 work can impact other, lower priority work. If the peak is short, normal job/thread prioritization resumes. If high interactive CPU utilization persists, then the algorithm discussed here goes into effect. The easiest way to tell this is going on, is to use the Work with System Activity (WRKSYSACT) command available only through the Performance Tools for iSeries, 5722-PT1 product. By refreshing WRKSYSACT screens over time you can identify when the "interactive work manager component" is "angered" by noting when 5250 jobs show a run priority value equal to or greater than 35.

Using an 820 -2435 with processor CPW of 600 and interactive CPW of 70 as an example, you can generally expect up to 11-14 percent Interactive CPU utilization when running the maximum amount of interactive work at 70 CPW. This is roughly 70 CPW / 600 CPW.

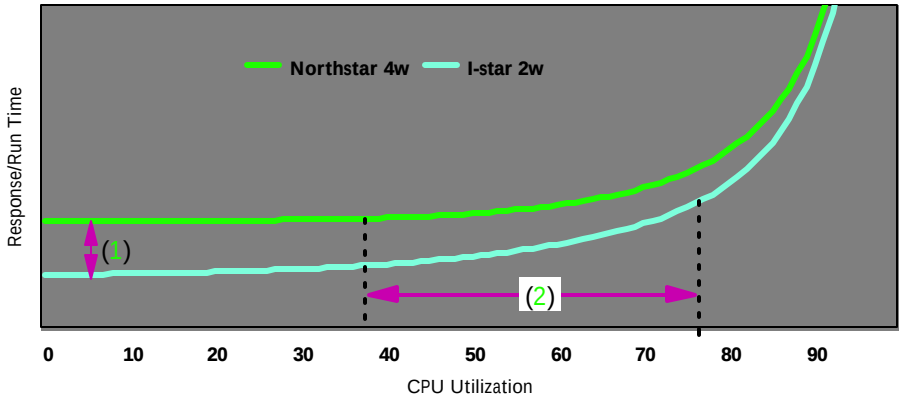
Note that the metric - **Total CPU Utilization (Interactive Feature)** that shows up on the Performance Tools for iSeries System Reports and per job in the Component Report or can be monitored for with a Management Central System monitor, is a more direct way of measuring and monitoring your utilization of the hardware's Interactive feature capability. When your application environment is consuming 100 percent of that feature you are getting full utilization of that feature. If you are consuming 70% of that feature and appear to be increasing that utilization, you need to begin planning on upgrading to a higher rated Interactive Feature. There is some debate on whether your Interactive Feature utilization at 100 percent (rather than 70 percent should be an acceptable operational environment. At the time this presentation was published, the best recommendation we can make is that on a single processor system or partition, you may begin noticing your interactive response time is starting to degrade but still get acceptable interactive response time. Certainly on a system or partition with 4 to 32 processors per "partition" you should get very good response time as you approach or exceed 100 per cent of your Interactive Feature.

You can monitor this activity yourself with 5722-PT1 reports or Management Central System monitors where you must determine any increased resource utilization trend or **let PM eServer iSeries (formerly PM/400) do this for you**!

CPW and CIW workloads

- CPW values are not representative of all workloads
  - CPW values are a statement of total capacity (Balanced Server Resources)
  - Low response, shorter run times are often more important for CPU-intensive workloads (CIW)
- iSeries servers offer dramatically improved response, run times for CPU-intensive workloads
  - Processors used in 270, 8xx are at least twice as fast as 170, 7xx models
  - If a processor is not busy, a job/task could complete in half the time (1)
- On 270, 8xxs, lower response/run times can be maintained at high CPU utilization (2)

If your application is CPU intensive, uses lots of main storage, performs 1000s of random database or storage accesses per transaction, the 2xx and 8xx servers are for you!



Notes: CPW and CIW workloads

This foil reminds us the processor or interactive CPW ratings are not the only way to look at performance. The new 270 and 8xx servers do, in general have twice the Megahertz ratings of the 170 and 7xx servers. If your application can be "profiled" to have a fairly even distribution of CPU, main storage access and database access the 170/7xx servers can deliver quite nice performance.

Note, if your interactive response time is already under 1 second, it is difficult to notice improved response time. Also, if your application is overly database I/0 or storage access oriented, but cannot take advantage of the improved main storage sizes or memory-processor-I/O switching speeds in the new 270/8xx servers, the 170/7xx servers also offer very acceptable performance.

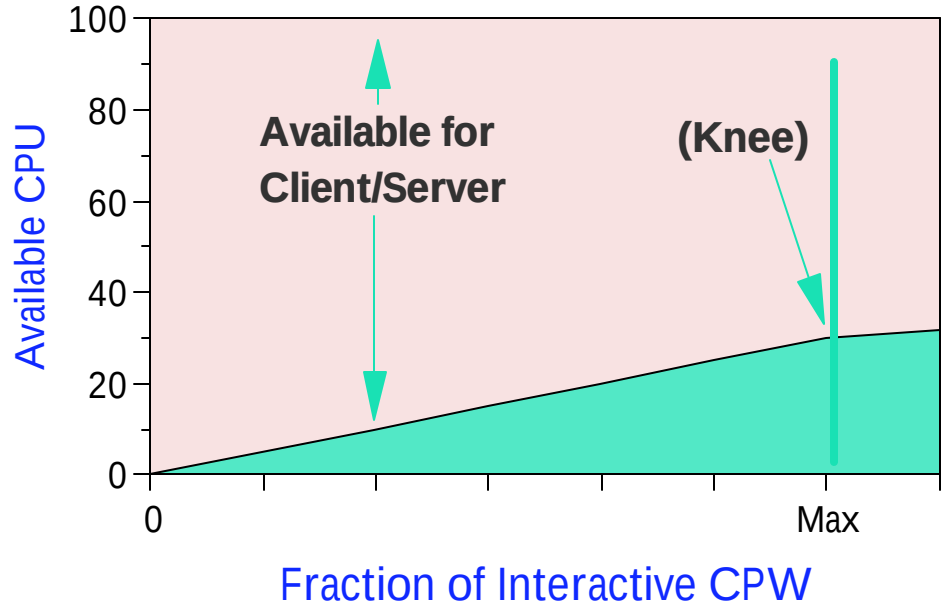
However, with the new technology 270/8xx with twice the Megahertz ratings, larger L2 cache (and with the i890 both L2 and L3 cache), and larger main storage (that can be taken advantage of by reducing physical disk I/O operations), applications that are overly concentrated in one or two of these areas can gain faster run times or response times by running on the 270/8xx servers.

Do not forget the Workload Estimator, which is discussed later in this presentation for sizing workloads.

Managing Interactive workloads: 270, 820, 830, 840, 890*

Interactive Performance Model - What's Different?

Model 8xx and Model 270 CPU
Distribution vs Interactive Utilization



V4R5: Significant advantage over earlier AS/400 models/OS/400 releases where system overhead reduces client/server capacity after exceeding the interactive threshold

available
interactive

*Not applicable to iSeries servers announced January 2003 with Enterprise Package feature

When interactive threshold is exceeded...

- Interactive workload will contend with other interactive workload for % of CPU available to interactive processing (determined by optional level of interactive CPW purchased)
- Interference from system overhead caused by excess interactive workload is virtually eliminated for any/all client/server workloads

Notes: Interactive workloads: 270, 820, 830, 840, 890*

This foil applies to all iSeries systems **except those announced January 2003 (that is, the iSeries 800, I810, i825, i870, i890 (#2497/#2498)) that are ordered with the Enterprise Edition package, which enables all of the processor CPW to be available to any combination of 5250 OLTP and other applications.**

This graph shows the new interactive work management algorithm used on the 270 and 8xx servers starting with V4R5. Essentially when the microcode determines that your 5250 work is achieving its rated interactive performance - at the vertical bar (Max/Knee) the 5250 work does not continue to take available CPU from the non-interactive (Client/Server) work on the server. Instead the 5250 work contends with itself at this time. Over time, as the 5250 work consumes lower CPU utilization, the OS/400 Dynamic Priority Schedule support (active starting with RISC V3R6) 5250 work will be managed normally.

In real-life scenarios the system does tolerate high CPU utilization by 5250 jobs before putting into effect the algorithm discussed here. That is, the system permits 5250 OLTP CPW utilization to be higher than your interactive feature for a period of time before downgrading the priority of the 5250 jobs doing this work. Monitoring this activity, the system relatively quickly detects 5250 work that lowers to "within 5250 CPW capacity level" and readjusts the job run priority value back to its original setting. Extended peaks of high CPU utilization by 5250 work can impact other, lower priority work. If the peak is short, normal job/thread prioritization resumes. If high interactive CPU utilization persists, then the algorithm discussed here goes into effect. The easiest way to tell this is going on, is to use the Work with System Activity (WRKSYSACT) command available only through the Performance Tools for iSeries, 5722-PT1 product. By refreshing WRKSYSACT screens over time you can identify when the "interactive work manager component" is "angered" by noting when 5250 jobs show a run priority value equal to or greater than 35.

Using an 820 -2435 with processor CPW of 600 and interactive CPW of 70 as an example, you can generally expect up to 11-14 percent Interactive CPU utilization when running the maximum amount of interactive work at 70 CPW. This is roughly 70 CPW / 600 CPW.

Note that the metric - Total CPU Utilization (Interactive Feature) that shows up on the Performance Tools for iSeries System Reports and per job in the Component Report or can be monitored for with a Management Central System monitor, is a more direct way of measuring and monitoring your utilization of the hardware's Interactive feature capability. When your application environment is consuming 100 percent of that feature you are getting full utilization of that feature. If you are consuming 70% of that feature and appear to be increasing that utilization, you need to begin planning on upgrading to a higher rated Interactive Feature. There is some debate on whether your Interactive Feature utilization at 100 percent (rather than 70 percent should be an acceptable operational environment. At the time this presentation was published, the best recommendation we can make is that on a single processor system or partition, you may begin noticing your interactive response time is starting to degrade but still get acceptable interactive response time. Certainly on a system or partition with 4 to 32 processors per "partition" you should get very good response time as you approach or exceed 100 per cent of your Interactive Feature.

You can monitor this activity yourself with 5722-PT1 reports or Management Central System monitors where you must determine any increased resource utilization trend or let PM eServer iSeries (formerly PM/400) do this for you!

Notes: Interactive workloads: 270, 820, 830, 840, 890* -2

In actual measurements, you can exceed 100 percent of the Interactive Feature utilization for periods of time. If, either in the System Report or the System monitor of the Interactive Feature metric, you detect consistent system-wide utilization values in the 100% or higher range, you are using "more than rated 5250 CPW" (corresponds to Interactive Feature nnnn). Because the V4R5 - V5R2 microcode monitoring 5250 work support 5250 CPW utilization to be higher for a short periods, your environment can successfully operate during peaks of time. The time period of peak Interactive Feature utilization tolerated by this microcode increases based upon the relative number of processors and their CPW.

Logical Partition considerations:

In a logical partition, where a percentage of Processor CPW and Interactive Feature CPW has been assigned, you need to manually account for the percentage relative to total system-wide capability. The Collection Services data used by both 5722-PT1 and the Interactive Feature system monitor, automatically adjusts for these percentages.

For example, assume you have an i820 2435 partition set up to be:

- 70% of 600 Processor CPW - 420 CPW
- 70% of 70 Interactive CPW - 49 CPW

Then the performance data would show 100 percent CPU utilization when your partition's total CPU capacity of 420 CPW is being consumed. For monitoring the interactive feature of that partition, you would know that 100 percent Interactive Feature utilization would be 49 CPW.

Additional shared processor LPAR partition considerations:

If you are running in a *shared processor LPAR partition* (for example, a partition has 3.5 processors) a new SLIC task appears - HVLPTASK, **that does not consume actual CPU utilization,** but appears to do so in either the Performance Tools reports, WRKSYSACT command displays, or the System monitor - CPU utilization metric.

The system is managing the work appropriately but you need to understand how HVLPTASK works and essentially discard the CPU utilization that appears to be consumed by this task in the performance tools we have just listed. There is an iSeries support description of how HVLPTASK works at

http://www-1.ibm.com/servers/eserver/iseries/lpar/hvlptask.htm

Notes: Interactive workloads: 270, 820, 830, 840, 890* -3

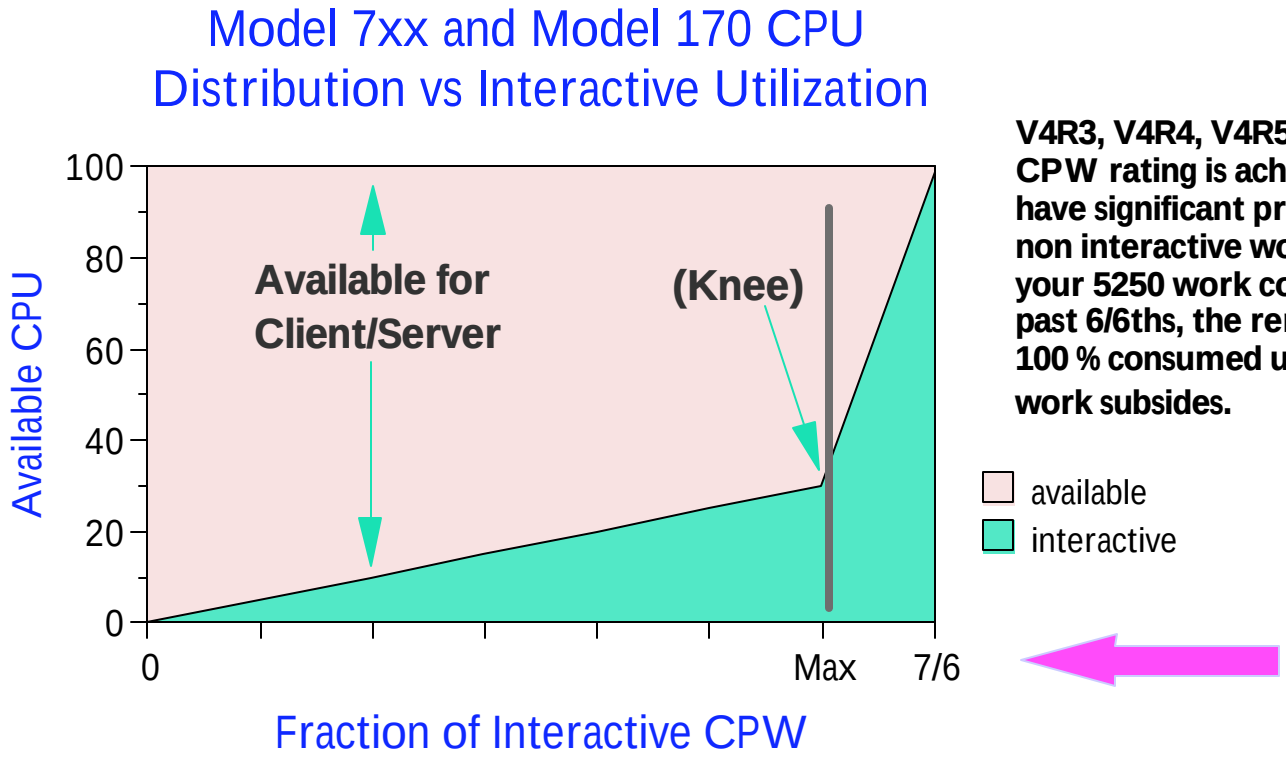
The CPU time reported for HVLPTASK is a function of the processing capacity assigned to the partition - "processor CPW." The CPU time charged to HVLPTASK scales with the amount of work being done by real jobs - thus making the system CPU percent utilization behave appropriately - going from 0 to 100 in direct proportion to the amount of customer work going on.

Example

- Partition A has a capacity of 0.9 processor units, partition B has a capacity of 0.1 processor units and partition C has a capacity of 1.5 processor units. Partition A is defined to use one virtual processor, partition B is defined to use one virtual processor and partition C is defined to use two virtual processors.
- Partition A: When the partition is idle, HVLPTASK shows consuming 0% of the CPU time. As the CPU time consumed by real jobs in the partition goes from 0 to 0.9 processor units (that is the maximum CPU resource allowed), reported system CPU utilization for the partition will go from 0% to 100% . Reported CPU utilization for HVLPTASK will go from 0% to 10% and reported CPU utilization by real jobs will go from 0 to 90%.
- In an actual test scenario, we observed with WRKSYSACT in a .9 processor partition, that as real jobs started taking a total of 10-15 percent CPU utilization in that partition, HVLPTASK was shown as taking 50-60 percent CPU utilization. When we increased CPU utilization by real jobs into the 50 to 60 percent CPU utilization the CPU utilization shown for HVLPTASK was significantly lowered - supporting the target maximum CPU utilization "taken" by HVLPTASK of 10 percent for that specific partition configuration.

Interactive workload management: 170, 7xx

Interactive Performance Model - V4R3, V4R4, V4R5 170, 7xx



V4R3, V4R4, V4R5: When 5250 CPW rating is acheived you still have significant processor left for non interactive work. However, if your 5250 work continues to grow past 6/6ths, the remaining CPU is 100 % consumed until the 5250 work subsides.

When interactive threshold is exceeded...

- Interactive workload can continue, based upon job priorit,y and consume more than rated CPW.
- If left to continue, interactive work could consume 100 percent of available processor power

Notes: Interactive workload management: 170, 7xx

As you can see from this graphic on the older technology systems, you could exceed you 5250 rated capacities for a time period - with direct impact to any non-interactive work that is running at a run priority above 35. (50 is the typical default).

If this is your environment, which would be indicated, for example in the Interactive Feature utilization under 5722PT1 System Report, Management Central - System Monitor, and V5R1 Graph History, or PM/400e statistics, is in the 5/6 to 7/6 range you may experience interactive performance degradation when moving to an 270 or 8xx with Interactive CPW ratings equal to or slightly higher than the corresponding 170 or 7xx rating.

iSeries for Domino - January 2003 Announced Servers

- **All jobs and tasks have access to system resources except 5250 OLTP applications***
  - Limited 5250 CPW is available for a system administrator who wishes to use 5250 display device I/O to manage various aspects of the server

825 and 810 Server Feature Codes are Domino Specific

825 and 810 Processor Feature Codes are not Domino Specific

Standard Package Features* (not Enterprise Packages)

	Mail/Calendar Users*	Server FC	Processor FC	Package FC	Min Disk (GB)	Min Mem (GB)	Minimum Processor Activations	Minimum Notes or iNotes Licenses	Minimum Domino Server Licenses
825 6 way	17400	0773	2473	7416	945	12	3	2000	1
825 4 way	11600	0772	2473	7416	560	6	1	1000	1
810 2 way	7900	0771	2469	7428	525	5.5	N/A	200	1
810 1 way	4200	0770	2467	7410	315	3.5	N/A	100	1
810 1 way	3100	0769	2466	7407	105	1.5	N/A	100	1

6 way = 3/6 way + 3 processor activations (comes with a a total of six OS/400 licenses

4 way = 3/6 way + 1 processor activation (comes with a a total of four OS/400 licenses

* MCU values are estimates requiring more than the minimum memory and disk

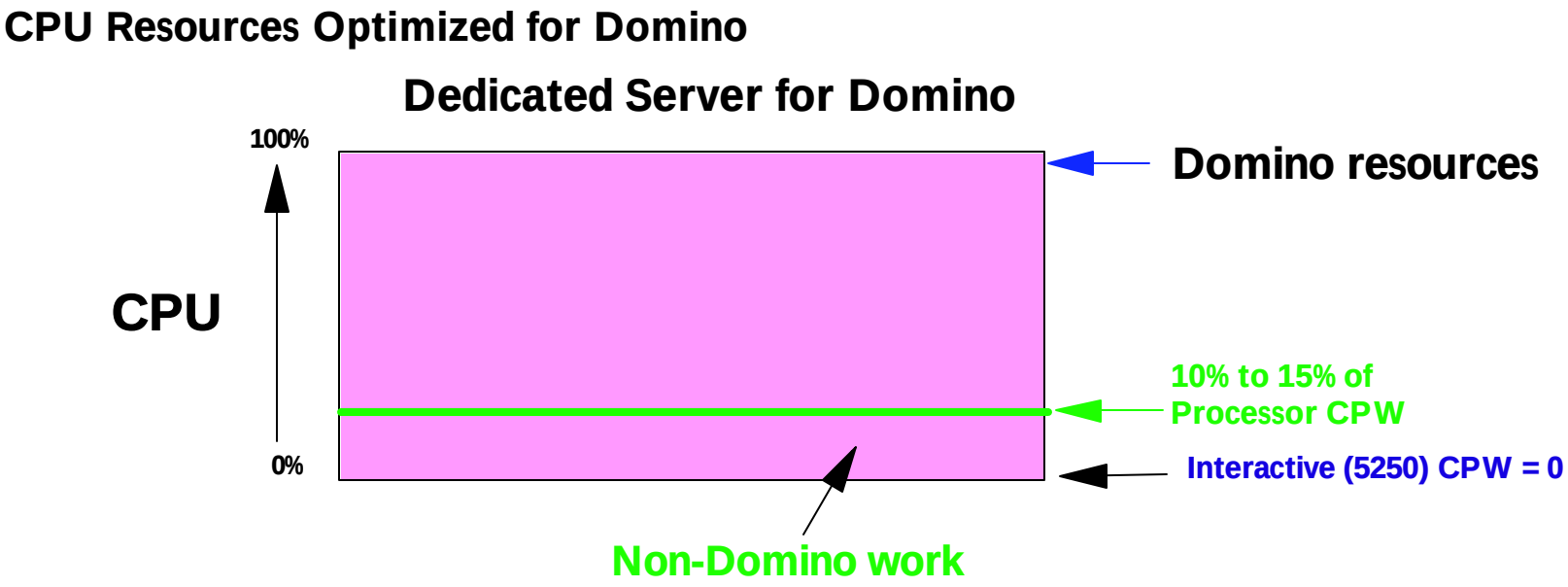
Notes: iSeries for Domino - January 2003 Servers

This foil focuses on the new iSeries for Domino servers announced January 2003. These are "Server feature coded" versions of the newly announced i810 and i825. With Server feature codes (which require verification of the minimum iNotes licenses and minimum Domino server licenses) 0769, 0770, 0771, 0772, 0773, you get a server that can run all workloads except 5250 OLT applications with full access to all system processor capabilities.

These Domino Server feature codes are associated with the Standard Package feature codes (as shown) available with the i810 and i825. As with all i810, i825, i870, and i890 (#2497, 2498) servers announced January 2003, the Standard package offers enough 5250 capability for an Administrative job to run from a 5250 workstation.

On these iSeries for Domino i810 and i825 models, full CPU processor capability is available to non-5250 jobs with or without Domino active. The older DSD models require Domino work to be available.

Managing work on a Dedicated Server for Domino*



- Domino workloads can utilize 100% of the Dedicated Server for Domino ... it is designed to be a "dedicated" Domino server
- Non-Domino workloads
  - Recommendation: Keep non-Domino workload CPU utilization less than 15% of CPU
  - Use single 5250 job for system administration functions only
- Java Servlets and WebSphere Application Server, are considered complementary when used with Domino on V5R1 Dedicated Server for Domino models shipped after September 28, 2001 or running V5R2

**** Does not apply to i800 and i825 iSeries for Domino servers announced January 2003***

Notes: Managing work on a Dedicated Server for Domino*

This DSD Domino and non-Domino workload management chart has applied since the first DSDs were announced.

The algorithm discussed here does not apply to the new iSeries for Domino i810 and i825 servers announced January 2003. See the foil "iSeries for Domino - January 2003 Announced Servers" for information on these servers.

The Dedicated Domino Processor Features are designed specifically to support Lotus Domino workloads. They provide sufficient capacity for non-Domino workloads in support of a Domino environment (e.g., remote DB2 access, file serving, Integrated Netfinity Server, TCP/IP).

Lotus Domino Workloads -- defined as out-of-the-box functions such as:

- E-mail
- Calendaring and scheduling
- Web serving
- Standard Lotus Domino template applications (discussion database, workflow) and custom developed applications written with Domino Designer that perform no external program calls, local relational database access, or Java integration

Non-Domino Workloads -- Do not propose the AS/400e Dedicated Server for Domino for customers planning significant use of non-Domino workloads. Instead, propose the AS/400e 270 or 8xx server with standard processor features. Following are the restrictions for non-Domino workloads on the AS/400e Dedicated Server for Domino:

- Any and all workloads not designated as Domino workloads are considered to be non-Domino workloads (capacity is published as Processor CPW) and should be managed to a maximum of 10-15% of the CPU.
- Application integration functions used as components of a Domino application (e.g., local DB2 Universal Database access, external RPG/COBOL program calls) will be designated non-Domino workloads and will be managed to a maximum of 10-15% of the CPU
- V5R1+ and "new" Domino Processors shipped starting October, 2001
  - Non-Domino workloads, that are not used with Domino, should be managed to a maximum of 10-15% of the CPU
  - Non-Domino workload, DB2 Universal Database access, is strictly limited ... up to 15% of the CPU is available
  - **Non-Domino workloads, such as Java Servlets and WebSphere Application Server, are considered complementary when used with Domino ... thus, 100% of CPU is available** .

V5R1 and later task dispatching algorithm

Starting with V5R1

In under 90-100 per cent system-wide/partition-wide CPU utilization environment when sufficient number of processors are available:

- Lower numbered processor(s) reserved for high priority microcode tasks
- Higher numbered processors used for normal work

Example - 8-WAY 830:

```
Average CPU Utilization . . . . . .:  43,0
CPU 1 Utilization . . . . . .:   3,8
CPU 2 Utilization . . . . . .:   7,4
CPU 3 Utilization . . . . . .:  42,5
CPU 4 Utilization . . . . . .:  43,5
CPU 5 Utilization . . . . . .:  45,4
CPU 6 Utilization . . . . . .:  43,3
CPU 7 Utilization . . . . . .:  45,2
CPU 8 Utilization . . . . . .:  44,5

Total CPU Utilization (Interactive Feature) . .:   0,0
Total CPU Utilization (Database Capability) . .:  20,2
```

Notes: V5R1 and later task dispatching algorithm

Current V5R1 cumulative software package and V5R2 at general availability implement an algorithm that determines when there are a sufficient number of processors available per system/partition and the total CPU utilization is in the 90-100 PERCENT range, the lower processor or processors are assigned to only microcode high priority tasks.

Using an 8-WAY 830 with 8 processors as an example, 1/4 of the processors available are considered "low processors" (processors 1 and 2 in our 830 example). In an under 90-100 percent total CPU utilization system/partition environment, the higher 3/4 processors (3-8 in your 8-Way 830) are used for "normal work."

The "average 43% CPU utilization" shown in this example illustrates this utilization of the higher-numbered processors.

Benchmarks update

Benchmark	iSeries Results	Status	Significance and Currency: JANUARY 28, 2003)
VolanoMark (Java)	**1st Overall** 310, 000 iSeries 890 32-way	Published June 2002	Establishes iSeries server as leading Java server. Demonstrates scalability of i890 in Java - over twice the throughput of iSeries 840 running V5R1
Intentia Movex (Java)	**1st Overall** 2.38M trans/hr iSeries 890 32-way 2X Improve	Published June 2002	Demonstrates iSeries 890 transaction processing scalability through a leadership position in this leading ISV Java application benchmark. Once again, 2X improvement over iSeries 840 running V5R1
NotesBench R5-Notes Client	**1st Overall** 150,000 Users iSeries 890 32-way	Published August 2002	An iSeries server is the leader of this key industry benchmark, besting previous best of breed by 40%. Demonstrates leadership performance for the non-Intel market share leader, iSeries.
NotesBench R5iNotes	**1st Overall** 40,200 users iSeries 890 32-way	Published September 2002	NotesBench R5iNotes workload running with the Domino 6 server code. Web Access users with Lotus Domino 6 Release on IBM OS/400 V5R2. The audit establishes a new NotesBench record response time of 58 milliseconds, beating a previous record also held by and iSeries server.
SPECweb99 (e-Commerce)	**1st Apache** (3rd overall) 12,900 simultaneous connections/sec iSeries 890 16-way	Published September 2002	1st time an iSeries server has appeared in this critical industry benchmark, and in a leadership position as the only Apache-based Webserver in the list. Demonstrates tremendous performance improvement in V5R2 web serving capability and puts iSeries on the web serving map.
SPECweb99SSL (secure e-Commerce)	**1st Overall** 4,497 simultaneous connections/sec iSeries 890 16-way	Published January 2003	An iSeries server leads in this industry benchmark, asserting its unparalleled SECURE e-business scalability. Results are over two times better than previous best of breed, and demonstrate that iSeries is the best Webserver for iSeries customers to use.

See Notes page for references to websites providing additional details

Notes: Benchmark update summary

This foils lists the January 28, 2003, status of iSeries benchmarks for:

- VolanoMark (number of messages switched)
- NotesBench Mail and iNotes
- Intentia MoveX
- SPECWEB99
- SPECSWEB99SSL

Use the following URLs to obtain more details information.

- http://www.ideasinternational.com/benchmark/spec/specjbb2000.html
- http://www.notesbench.org
- http://www.ideasinternational.com/benchmark/lotus/mailonlyR5.html
- http://www.intentia.com
- http://www.volanomark.com

Apache Powered SPECweb benchmark update

SPECwebSSL*

	HTTP Server	Processor	Simultaneous Connections
IBM eServer iSeries 890	HTTP Server Powered by Apache	16-way POWER4	4,497*
HP Server rp8400	Zeus V4.1R3	16-way PA 8700	4,400
IBM eServer iSeries 890	HTTP Server Powered by Apache	16-way POWER4	3,600

SPECweb99

	HTTP Server	Processor	Simultaneous Connections
IBM eServer pSeries 690 HPC	Zeus 4.0	16-way POWER4	21,000
HP rp8400	Zeus 3.4r3	16-way PA 8700	15,000
IBM eServer iSeries 890	HTTP Server Powered by Apache	16-way POWER4	12,900

- Delivering industry-leading Apache based Web serving performance
- 64% of production Web sites based on Apache, <2% based on **Zeus**** servers

* See http://www.spec.org/osg/web99ssl
**See http://www.netcraft.com/survey

Notes: Apache Powered SPECweb benchmark update

A 16-way IBM eServer i890 set a new industry record with 4,497 secure socket layer (SSL) simultaneous connections in SPECWeb99 SSL benchmark testing, eclipsing both its previous record of 3,600 connections and HP's record of 4,400 connections. These benchmark results are as of January 7, 2003, as shown on the SPECWeb99SSL web site at:

- http://www.spec.org/osg/web99ssl

Note that the HP benchmark numbers are using a Zeus HTTP server - not the more popular Apache-based web server installed on so many more web servers than Zeus.

Webserving Performance Topics

HTTP Server for iSeries

Collection Services

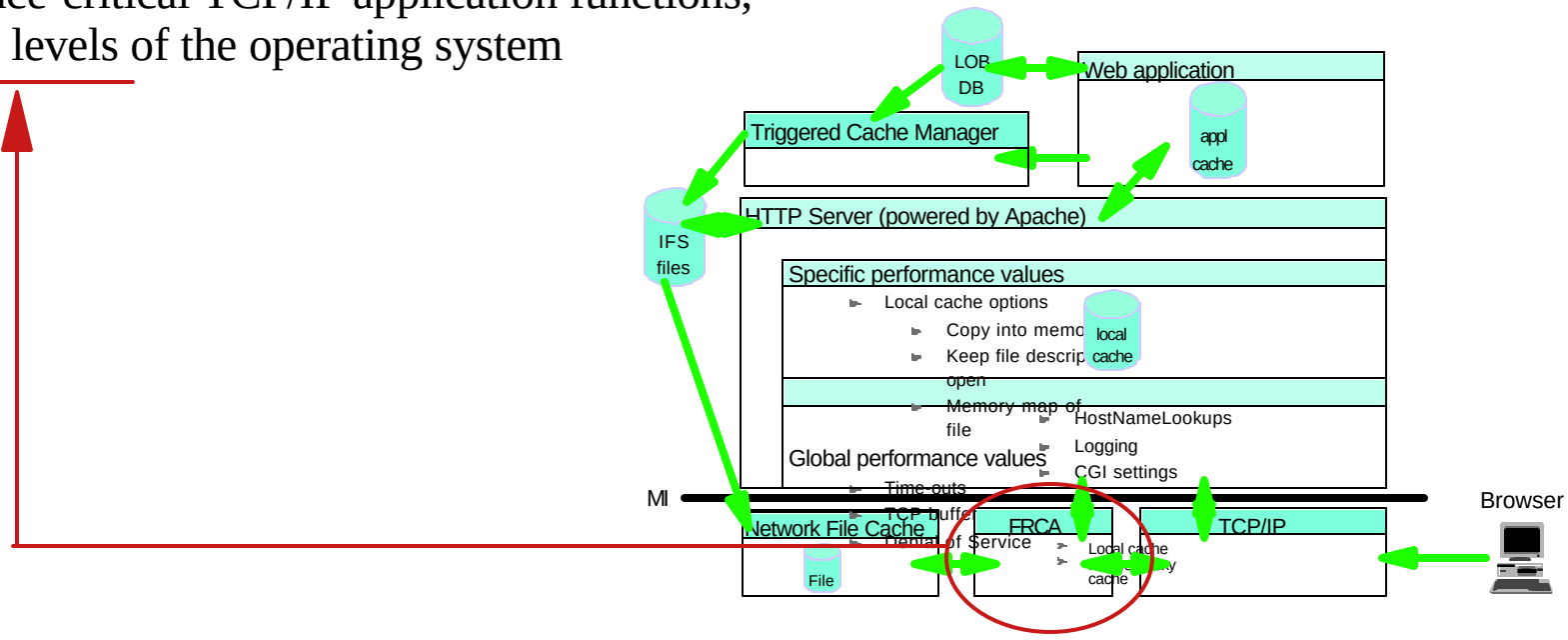
- Analyze HTTP server activity and transaction types

Performance Tools Reports

- Information on the transactions processed by HTTP Server jobs

Fast Response Cache Accelerator (FRCA)

- Performance-critical TCP/IP application functions, into lower levels of the operating system



Notes: HTTP Server for iSeries

Web-based transaction processing and webserving environments continue to grow in importance. We are enhancing the performance of these environments by providing improvements to SSL, implementing Fast Response Cache Accelerator (FRCA), and continued work with asynchronous I/O.

HTTP data collection category to contain HTTP performance data for Collection Services. The HTTP performance data can then be queried to analyze HTTP server activity and better understand what types of HTTP transactions are being processed by the iSeries (for example, static files, CGI, or Java Servlets).

Performance Tools for iSeries, 5722-PT1 has been enhanced to generate reports based on the HTTP performance data. The reports contain information on the transactions processed by HTTP Server jobs.

Refer to the Performance Management section of this presentation for an overview of the report enhancements.

FRCA (Fast Response Cache Accelerator)

Responding to the growing need for improved speed and performance of Web servers, IBM research has defined the Adaptive Fast Path Architecture (AFPA). AFPA has been implemented on several server platforms including Windows NT and Windows 2000, OS/390, AIX and most recently Linux. The external product name is most commonly known as Fast Response Cache Accelerator (FRCA). AFPA is a software architecture that dramatically improves the capacity of Web and other TCP servers. The architecture defines interfaces that allow these generic mechanisms to be exploited to accelerate a variety of application protocols, with the focus on HTTP. The architecture is general purpose and applicable to many TCP servers, including FTP, NFS, DNS and Domino.

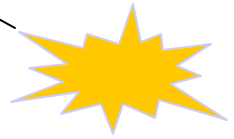
For OS/400 V5R2, this architecture is implemented as the FRCA feature with the HTTP Server (powered by Apache). FRCA is an IBM research cache based on a software architecture (AFPA) that dramatically improves the capacity/performance of Web and other TCP servers. FRCA, in V5R2, is used for the HTTP Server (powered by Apache) only. The architecture includes a network file cache that serves non-secure static (via a local cache option) and dynamic (via proxy) content directly from beneath the Machine Interface (MI).

Note that SSL/TLS (Secure Sockets Layer/Transport Layer Security) is not supported for the FRCA enabled sessions/ports and no authentication protection for the file is available with FRCA (NFC) cached pages. The page contents should be for public access under FRCA. Also, no NLS code page conversion performed - IFS files are read in binary and loaded into the NFC cache as is.

For more information on FRCA refer to the detaile**d e-business presentation.**

HTTP Serving Test Results: Hits per Second estimates

Transaction Type	Non Secure Trans/sec per CPW	Non Secure Trans/sec per CPW	Secure Trans/sec per CPW	Secure Trans/sec per CPW
	V5R2	V5R1	**V5R2**	V5R1
Static Page (IFS, non-cached	**1.75**	1.51	**1.47**	0.69
Static Page (cached)	**2.79**	2.05	**2.23**	0.82
Static Page (FRCA)	**13.01**	not available	not applicable	not applicable
CGI - new activation	0.06	0.08	0.06	0.08
CGI - named activation	0.37	0.40	0.36	0.31
Persistent CGI	0.28	0.32	0.26	0.26
Net.Data	0.17	0.22	0.17	0.20
User Module	**3.15**	1.21	**2.69**	0.62



Notes:

- IBM HTTP Server Powered by Apache or iSeries
- Based on 270 with moderate web server load measurements
- Data assumes no access logging, no name server interactions, Keep Alive ON, LiveLocalCache OFF
- Secure testing: 128-bit RC4 symmetric cipher and MD5 message digest with 1024-bit RSA public/private keys
- CPWs are "Relative System Performance Metrics" used to rate iSeries servers
- Web server capacities are based on total CPU available. Capacities may not necessarily scale by CPW. Actual results may differ significantly.
- Transactions using more complex programs or serving very large files (pages) will have lower capacities than shown in this table
- See the Performance Capabilities Reference manual, Network chapter, for WAS primitives performance

For more information refer to http://www.ibm.eserver.iseries/perfmgmt. Select the V5R2 Performance Capabilities Reference manual dated January 2003.

Notes: HTTP Serving Test Results

This table shows an excerpt from the V5R2 Performance Capabilities Reference manual. This manual discusses the internal IBM benchmarks (CGI program, Net.Data and so forth transaction processing and also has some test results showing how larger file (web page) sizes can further lower the CPW multiplier. Other Performance Capabilities Reference manual information also shows an alternate way of estimating hits per second capacities of an iSeries - CPW required per transaction per second estimated values.

Review the notes and disclaimers on this foils. The numbers shown assume the entire processor (CPW capacity) is available for web serving.

Important considerations include:

- V5R2 test results showed mostly improved (higher CPW number the more hits per second estimated) CPW multipliers for static (read only, no program processing) and WebSphere servlet and User Module (uses Apache APIs) transaction types.
- Use of V5R2 **Fast Response Cache Accelerator for unsecured static (not changed) data shows a significantly higher CPW multiplier rating.** FRCA can certainly support dynamic pages, but this table shows only static page content for all entries. Please note that FRCA does not apply to SSL data.
- Examples using this table:
  - Assume an i810-2466 with 1050 CPW rating with 100 percent of the processor rating is available for HTTP web servings. Assuming your application is similar to the IFS, non-cached transaction tested, yo u multiply 1.75 times 1050 and get an estimated number of hits per second value of **1837 transaction**s (hits) per second.
  - Assume the same full processor rating of 1050 and same non-cached transaction, but with only 20% of the system available for web serving, you multiply 1.75 by 1050 by .20 to get 367 **transactions** per second.
- Follow the performance tips in the Performance Capabilities Reference manual for optimal performance. These tips include using FRCA wherever possible (requires configuring HTTP Server powered by Apache, which is included with no charge 5722-DG1), use HTTP server fixed or dynamic cache where FRCA is not possible (no SSL or National Language Support), and other tuning tips at the TCP/IP protocol level (such as using Persistent Connections - set Keep Alive on, and, if using Java servlets, any tuning tips for the Java servlet manager, such as IBM's WebSphere Application server.
  Currently there is a WebSphere Application Server 4.0 for iSeries Performance Considerations document at
  - http://www-1.ibm.com/servers/eserver/iseries/software/websphere/wsappserver/product/PerformanceConsiderations.html .

The following foil summarizes the various HTTP server caching options supported on iSeries and some important "performance benefiting" parameters.

HTTP Server Features: Performance and Cache

As of V5R2

- Both **Asynchronous I/O**
- Both **KeepAliveTimeout (works with persistent connections)**
- Both **Denial of service**
- Both **Local memory cache**
- pbA **Fast Response Cache Accelerator (FRCA)**
- Both **Proxy caching**
  - pbA Reverse Proxy cache
- Both **Triggered Cache Manager (TCM)**

Both HTTP Server (original) and HTTP Server powered by Apache

pbA HTTP Server powered by Apache

Notes: HTTP Server Features: Performance and Cache

Asynchronous I/O: With the HTTP Server (powered by Apache and original) implementation, the HTTP Server processes communications requests asynchronously. In this asynchronous I/O model, threads are only involved in processing when there is work to be done. Threads are dispatched to perform work as required and when not performing work, the threads are returned to a pool of available threads making the server process more efficient and improving performance by better utilizing the thread resources. Asynchronous I/O also makes the server more scalable to support a high number of users especially when combined with persistent connections.

KeepAliveTimeout: When the server runs with persistent connections, the KeepAliveTimeout setting determines the number of seconds the server waits for subsequent requests before closing the connection. If this value is too low, the server could be impacted in terms of performance as connections could be closed frequently. If this value is too high, the Web server could have many connections open and the server could run out of resources. In this case the use of asynchronous I/O can alleviate (but not eliminate) the problem of running out of resources.

Denial of service: The denial of service attribute is equally a performance setting as well as a security setting. This setting allows us to identify, based on the data frame size, the possibility of an attack. The HTTP server may identify an attack because the frame size differs to the one it expects. Although this setting impacts the server performance as each request is tracked, it allows you to prevent a more dangerous performance degradation when dealing with a type of attack that may intentionally slow down or even completely paralyze your server.

Local memory cache: You can provide a caching service for files on your site using the local memory cache configuration options.

To use a local memory cache, you identify an amount of memory to allocate and a set of files to be cached. When the IBM HTTP Server for iSeries is started, the files are read into the local memory cache up to the limit of the amount of memory allocated or the limit of the number of files that you allow to be cached.

When a request is received at your IBM HTTP Server for iSeries, the local memory cache is checked first to determine if it has a copy of the requested file. If so, the file is served from the cache, which is significantly faster than if the file is retrieved from disk storage.

Fast Response Cache Accelerator (FRCA): With the HTTP Server (powered by Apache) FRCA provides a cache mechanism that dramatically improves the file serving performance of the HTTP Server (powered by Apache). Once a file is loaded into the below the MI cache called Network File Cache (NFC), as described earlier in this presentation, the second request for that file can be served from the NFC. This eliminates open, read and close actions for the file and is entirely handled below the MI by specially written SLIC code. FRCA can handle both a static file caching and a dynamic reverse proxy caching.

Notes: HTTP Server Features: Performance and Cache -2

Proxy caching: The IBM HTTP Server for iSeries can be configured as a non-caching or caching proxy server. When used as a non-caching proxy, the primary benefit of enabling proxy services is that the IP addresses used on the internal network are not sent out of your network. The proxy service forwards the request from your internal network using the IP address of the proxy server, not the address of the original requester. When the proxy server receives the response, it forwards the response to the original requester.

With caching enabled, the proxy server can act as a high-speed local store of previously accessed Web pages. When you configure the server as a proxy caching server, you can improve performance. You can also allow users of your internal network to access documents on the Internet. For example, if you frequently access the same set of Web pages from one or more sites, it may be advantageous to activate the caching feature. The retrieved Web page is stored locally on your iSeries server. Any subsequent accesses to the page occur at LAN speed, rather than at Internet speed.

Web pages can be encoded with a “no-cache” attribute or a specific expiration date. You can also configure the IBM HTTP Server for iSeries proxy service so that it periodically performs “garbage collection” to remove expired files from the cache.

Another use of the proxy service (with or without caching) is to log client requests. Some of the data available includes:

- Client IP address
- Date and time
- URL requested
- Byte count
- Success code

With this information, you can construct reports to account for the use of your Web site. For example, the information can be used in a charge-back system to understand and track marketing trends.

Reverse Proxy Caching - Apache server only

A reverse proxy is another common form of a proxy server and is generally used to pass requests from the Internet, through a firewall to isolated, private networks. It is used to prevent Internet clients from having direct, unmonitored access to sensitive data residing on content servers on an isolated network, or intranet. If caching is enabled, a reverse proxy can also lessen network traffic by serving cached information rather than passing all requests to actual content servers. Reverse proxy servers may also balance workload by spreading requests across a number of content servers. One advantage of using a reverse proxy is that Internet clients do not know their requests are being sent to and handled by a reverse proxy server. This allows a reverse proxy to redirect or reject requests without making Internet clients aware of the actual content server (or servers) on a protected network.

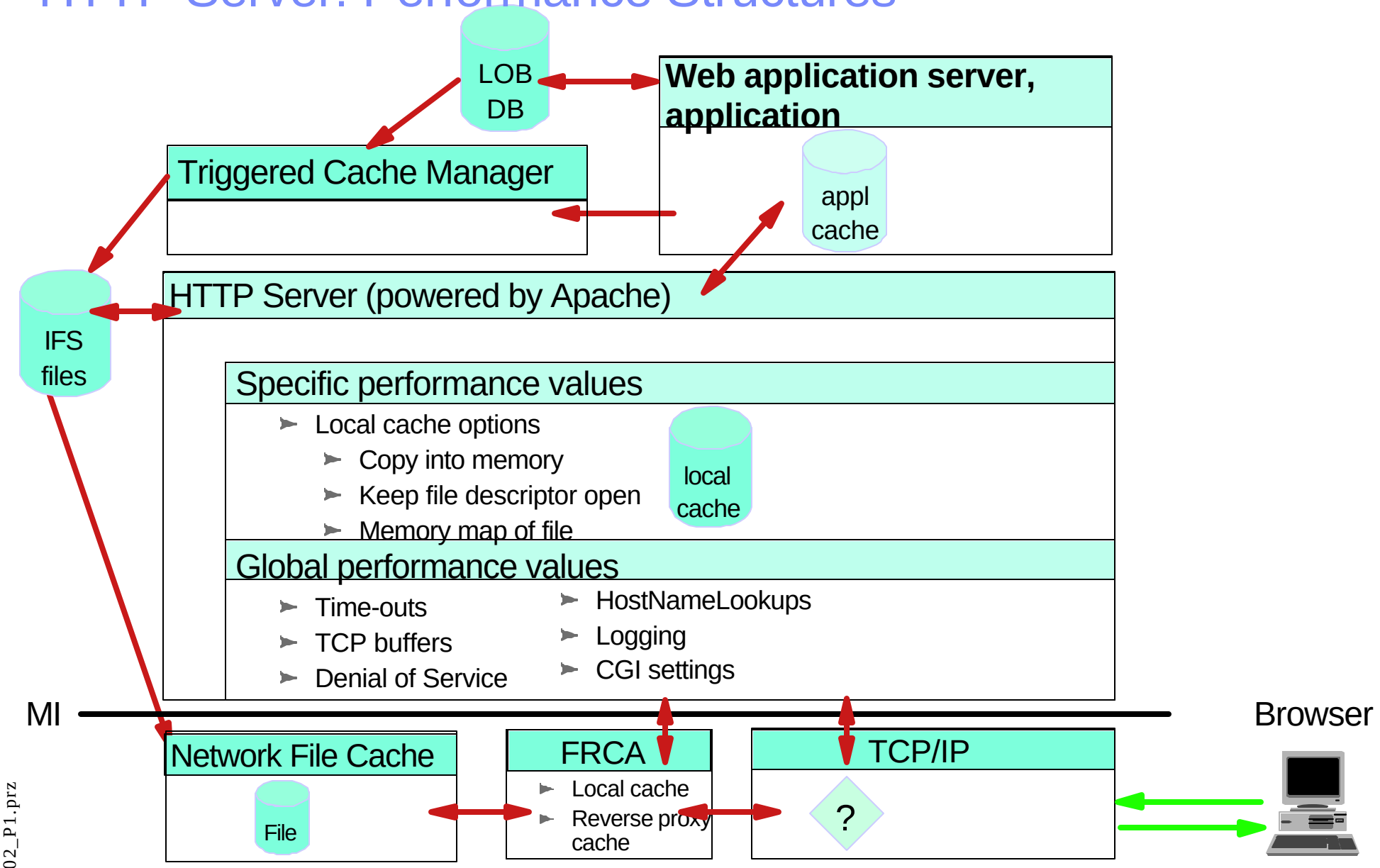
Notes: HTTP Server Features: Performance and Cache -3

Triggered Cache Manager (TCM): TCM provides a mechanism to manage dynamically-generated Web pages. TCM is a separate server that can be used in conjunction with the HTTP Server to allow a Web designer to build dynamic pages. It only updates the page in the IFS when the underlying data changes, thereby improving the performance of a Web site. TCM is not a cache - it is a cache manager, because it can update cache based upon knowing that a cached page has been changed.

Note: The Web application itself may have its own caching mechanisms. In the case of Web Application Servers, that is, Java servlet managers, may also have their own caching mechanisms.

See the next foil for a schematic of all these possible performance, including caching mechanisms working together.

HTTP Server: Performance Structures



Notes: HTTP Server: Performance Structures

This foil "puts together" several optionally specified caching algorithms and some "performance parameters" - illustrating how they all can work together. Not all of the performance parameters are discussed in this presentation. See the Performance Capabilities Reference manual.

The HTTP Server (powered by Apache) can improve the Web server performance in two different levels:

- Using global parameters that allow us to configure the attributes used by all the HTTP servers in your iSeries server.
- Using specific parameters based on the type of data the client is requesting. These specific parameters generally revolve around the concept of a local cache, proxy cache, or the FRCA (Fast Response Cache Accelerator) used by the HTTP Server (powered by Apache).

In addition, using a Triggered Cache Manager (TCM) server can dramatically improve the performance of your Web application by being proactive in creating dynamic web content and placing it in the iSeries Integrated File System (IFS) to be served at static document speeds.

The figure all the HTTP Server components available to improve the performance using the HTTP Server (powered by Apache).

When the client sends the request, configured global performance values are used to optimize the work performed by the HTTP Server (powered by Apache) Web server. The Web server then tries to process the request either using data cached in the FRCA associated Network File Cache, the local cache (memory) or to serve the file from the IFS.

If the content is to be dynamically generated, control is passed to a Web application that can optionally access Line of Business (LOB) databases to *reactively* create content. Note that the Web application can be managed by a web application server, such as WebSphere Application Server V4.0 managing Java servlets or enterprise java beans. A web application server or the web application itself could have its own caching mechanism as well.

Changes to the Line of Business (LOB) database or some other ‘triggering’ mechanism can cause the Web application to be proactive in the update of pages directly in the iSeries IFS. In this way, even before the first client request arrives at the iSeries server, the content (a web page with dynamic data pulled from the LOB database) can be generated and stored in the IFS - as a ‘static’ document to be quickly served by the HTTP server.

Clearly, the sooner cached data can be used during client request the fewer resources and time that is needed for the entire end-to-end transaction between client and server.

Cryptographic Accelerator - 2058-001

- Offloads selected portions of cryptographic processing from host processors
  - CRTDEVCRP DEVD(CRYACL) RSRCNAME(CMN002) PKAKEYFILE(*NONE) DESKEYFILE(*NONE) TEXT('2058 (iSeries Feat #4805)')
- iSeries SSL uses for SSL session "handshake" processing (just as SSL can do for older 4758 (#4801/#4802 Cryptographic Coprocessor)
- Single Cryptographic Accelerator can support up to 1000 full SSL handshakes per second
- V5R2 performance metric:
  - Collection Services counts SSL and non-SSL connections
  - Performance Tools for iSeries System Report shows connection counts

January 2003 Performance Capabilities Reference manual contains chapter on cryptography performance

Notes: Cryptographic Accelerator - 2058-001

This new Cryptographic Accelerator "card" is used by the system on an SSL connection handshake process (negotiates the level of SSL support and key exchange information to be used by an SSL session). You still use a digital certificate to perform the SSL processing and define the application (HTTP serving, Management Central, Telnet, and so forth) to use SSL or not.

The older 4758 technology and the new technology 2058 are used for SSL handshaking and not for the actual data transmission encryption. This is because the overhead of constantly passing data to the hardware to be encrypted on output and decrypted upon input negate the benefits of having the hardware perform the function.

The V5R1 and V5R2 Performance Capabilities Reference manual has test results showing the handshaking performance improvements for both the older 4758 technology and the new technology 2058. Note that the 4758 provides a broader range of encryption algorithms and financial industry encryption capabilities that are not supported with the 2058.

The detailed Security presentation contains additional information. See V5R2 Information Center -> Security for the most complete coverage of 4758 and 2058 capabilities.

In the Performance Management section of this presentation we show an example a new section of the Performance Tool for iSeries System Report showing the counts of SSL and non-SSL connections during the performance data collection period.

WebSphere V5.0 Express and Base

- Both support WebSphere Development Studio WebFacing
- WAS 5.0 Base
  - Includes J2EE 1.3 certified functions including Enterprise Java Beans (EJBs), XML Web services, Java Messaging Service (JMS) and more
  - Uses both Just In Time (JIT) interpretation and IBM's Mixed Mode Interpretation (MMI) instead of all Direct Execution (DE) to run the Java classes.
  - Up to approximately 10% performance improvement compared to WAS 4.0.3 Advanced Edition
  - Up to approximately 11% run time performance improvement compared to WAS 5.0 Express
- WAS 5.0 Express
  - Manages Java servlets, JSPs, not full function J2EE: -> designed for less sophisticated application environments than J2EE-based implementations
  - Tuned for lower processing power servers (lower CPW iSeries Series servers)
  - Minimum recommended CPW is 300 (iSeries 800-2463)
  - Uses DE:
    - Startup: Up to 16% faster than WAS 5.0 Base startup
    - Run time CPU consumption is up 10% over WAS 5.0

Notes: WebSphere V5.0 Base and Express

WebSphere Application Server 5.0 includes Express, Base, and Network Deployment editions.

Express is an attractive solution for many small and midsized customers. Whether the customer is deploying an application which requires a WebSphere solution (such as WebFacing), or the customer is looking to write their own solution to leverage the Web, Express can be a lower cost alternative to the full J2EE compliant WebSphere Application Server V5. There are features that the WebSphere Application Server Base edition has that Express does not have. The most notable of these features include Sun's J2EE functions such as Enterprise Java Bean (EJB)(*trademark) transaction recovery support and horizontal and vertical scaling using clones. All of the performance measurements documented in the January 2003 Performance Capabilities reference manual use the JDBC version of Trade2, rather then the EJB version. Underneath the covers, the WebSphere Express code base is the same as the WebSphere Application Server base code, with the only differences are the added features of base, and the way Express is packaged.

Note: WebSphere Application Server 5.0 Network Deployment edition includes Base implementation, but adds clustering, workload management, distributed security, systems management and directory functions.

Express is tuned for the lower CPW machines, using Direct Execution (or pre-compiled Java programs). This tuning is evidenced by faster startup time compared to other WAS versions and WAS 5.0 Base. In general WAS 5.0 Express has improved startup time compared to WAS 5.0 Base and earlier versions of WAS editions. The January 2003 Performance Capabilities reference manual has test results that show the Trade2 application server starts up over 16% faster with Express, WebSphere 5.0 Base. This is primarily due to the DE implementation.

However, the Performance Capabilities manual also shows test results indicating WebSphere Express delivers up to 11% less throughput then WebSphere 5.0 Base, while running the Trade2 application. This is primarily do to the fact that the JIT compiler has become faster runtime then DE recently, due to new technology. Thus, once the application is up and running, you can expect slightly worse performance with Express, then base WebSphere Application Server 5.0. This is due to the fact that JIT (Just-In-Time) compiled code generated on the fly can perform more optimizations then the static code generated by DE. It is important to note that your application will likely not see the same 11% degradation that Trade2 experienced, but will experience some degradation.

Notes: WebSphere V5.0 Base and Express -2

The IBM Workload Estimator does not recommend running WAS 5.0 Express on an iSeries with less than 300 processor CPW.

Note:

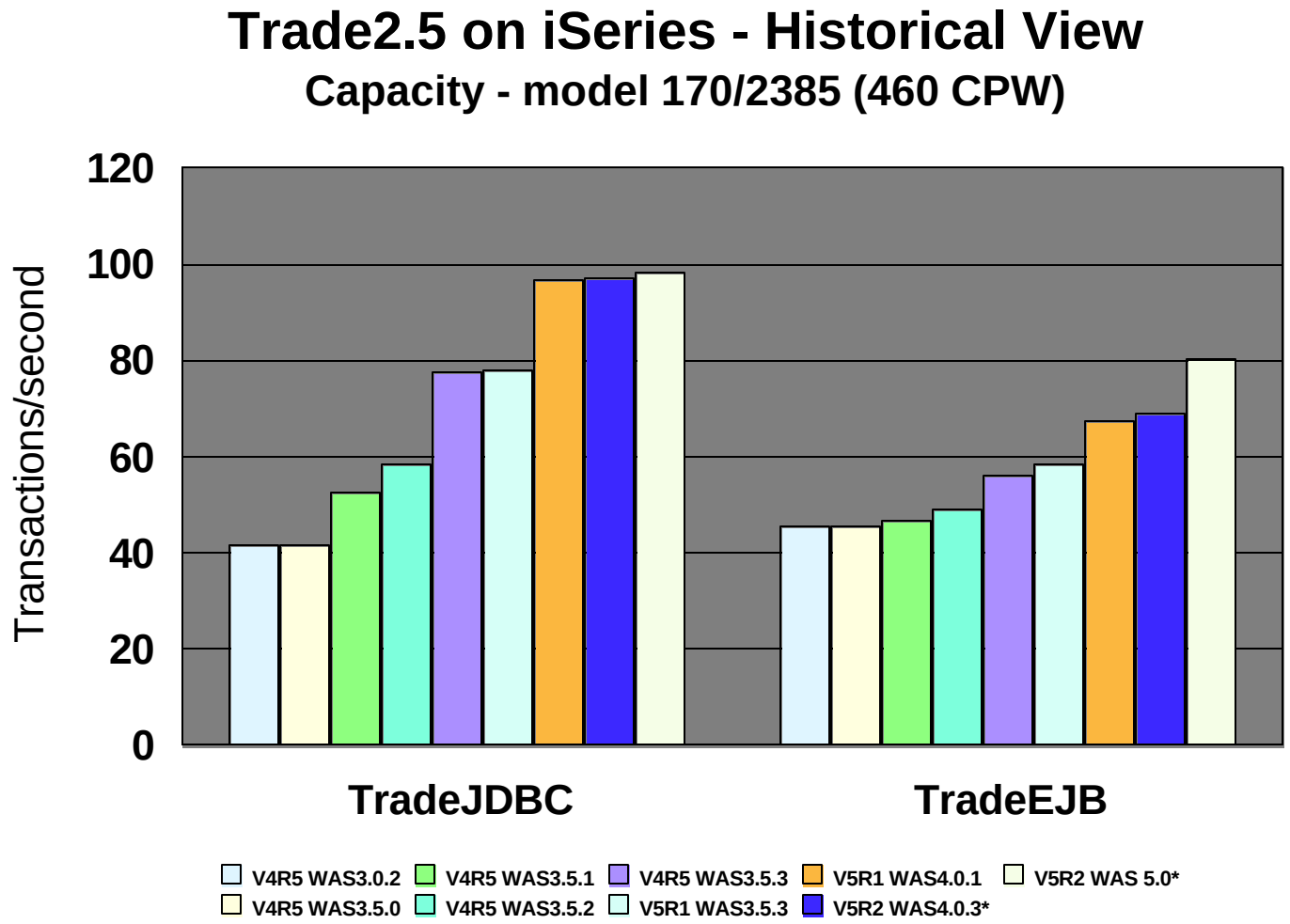
- The IBM JVM has incorporated an advanced selective compilation technology (which IBM named mixed-mode interpretation, or MMI) that significantly speeds application loading times. It does so by delaying the just-in-time compilation of certain methods until it is clear that the time spent to compile them is justified. For some methods, it may never be determined sufficiently "hot" to warrant compilation. The net effect is an overall improvement in application load time, without degrading overall processing performance.
- The Direct Execution (DE) environment is one where the Java methods that, prior to run time, have already been compiled into machine-specific hardware executable instructions.

The following charts show how WAS 5.0 Base running on V5R2 runtime performance compares to earlier OS/400 releases and WAS releases.

WebSphere Applic. Servers on iSeries: Small Systems

IBM Trade2 Benchmark (WebSphere eBusiness Benchmark)

- JDBC access (no EJB)
- EJB



* Results are projected from Trade2.7, which is 20% heavier then Trade 2.5

Notes: WebSphere Applic. Servers: iSeries Small System

The IBM-wide WebSphere Performance team has built a benchmark for characterizing performance of the WebSphere Application Server called the WebSphere eBusiness Benchmark.

The benchmark was derived from experiences with many customer environments. The WebSphere eBusiness Benchmark was built to emulate an online brokerage firm. Trade2, which is the actual application name, is a versatile test case designed to measure aspects of scalability, performance and competitiveness. The Trade2 application is a collection of Java classes, Java Servlets, Java Server Pages and Enterprise Java Beans which together form an application providing emulated brokerage services. Trade2 is a follow-on to the Trade and Broker benchmarks used in previous reports. Trade2 was developed using the VisualAge for Java and WebSphere Studio tools and each of the components are written to open web and Java Enterprise APIs making the Trade2 application portable across J2EE compliant application servers. Trade2 follows the "WebSphere Application Development Best Practices for Performance and Scalability".

The Trade2 application allows a user, typically using a web browser, to perform the following actions:

- Register to create a user profile, user ID/password and initial account balance
- Login to validate an already registered user
- Browse current stock price for a ticker symbol
- Purchase shares
- Sell shares from holdings
- Browse portfolio
- Logout to terminate the users active interval

To measure performance across various configuration options, the Trade2 application can be run in several modes. A mode defines the environment and components used in a test and is configured by modifying settings in a profile. For example, data object access can be configured to use JDBC directly or to use EJBs under WebSphere Advanced Edition by setting the Trade2 runtime mode. In a Trade2 "Sell function," operations are listed for the EJB runtime mode. If the mode is set to JDBC, the sell action is completed by direct data access through JDBC from the TradeAppServlet. Several testing modes are available and are varied for individual tests to analyze performance characteristics under various configurations.

Notes: Comparing WebSphere App.Servers on iSeries -2

This chart shows historical performance test results on several releases of WebSphere application serving software.

Results were measured on a 170-2385 system (Processor CPW = 460). Note that the Trade2 internal IBM benchmark is revised over time to reflect live customer feedback. Over time the Trade2 application will have newer versions, that typically use increasing amounts of system resources on a tested operating system.

The following notes apply to the software release levels used.

- Trade2 JDBC and Trade2 EJB benchmarks
- WebSphere 3.0.2, 3.5.0, 3.5.1, and 3.5.2 were on a V4R5 system
- WebSphere 3.5.3 was measured on both V4R5 and V5R1
- WebSphere 4.0 AE was measured on V5R1
- WebSphere 4.0.3 AE on V5R2 was estimated via measurements with WAS 4.0.2 with software enhancements to be included with WAS 4.0.3
- The IBM HTTP Server (powered by Apache) was used starting with the V5R2 measurements

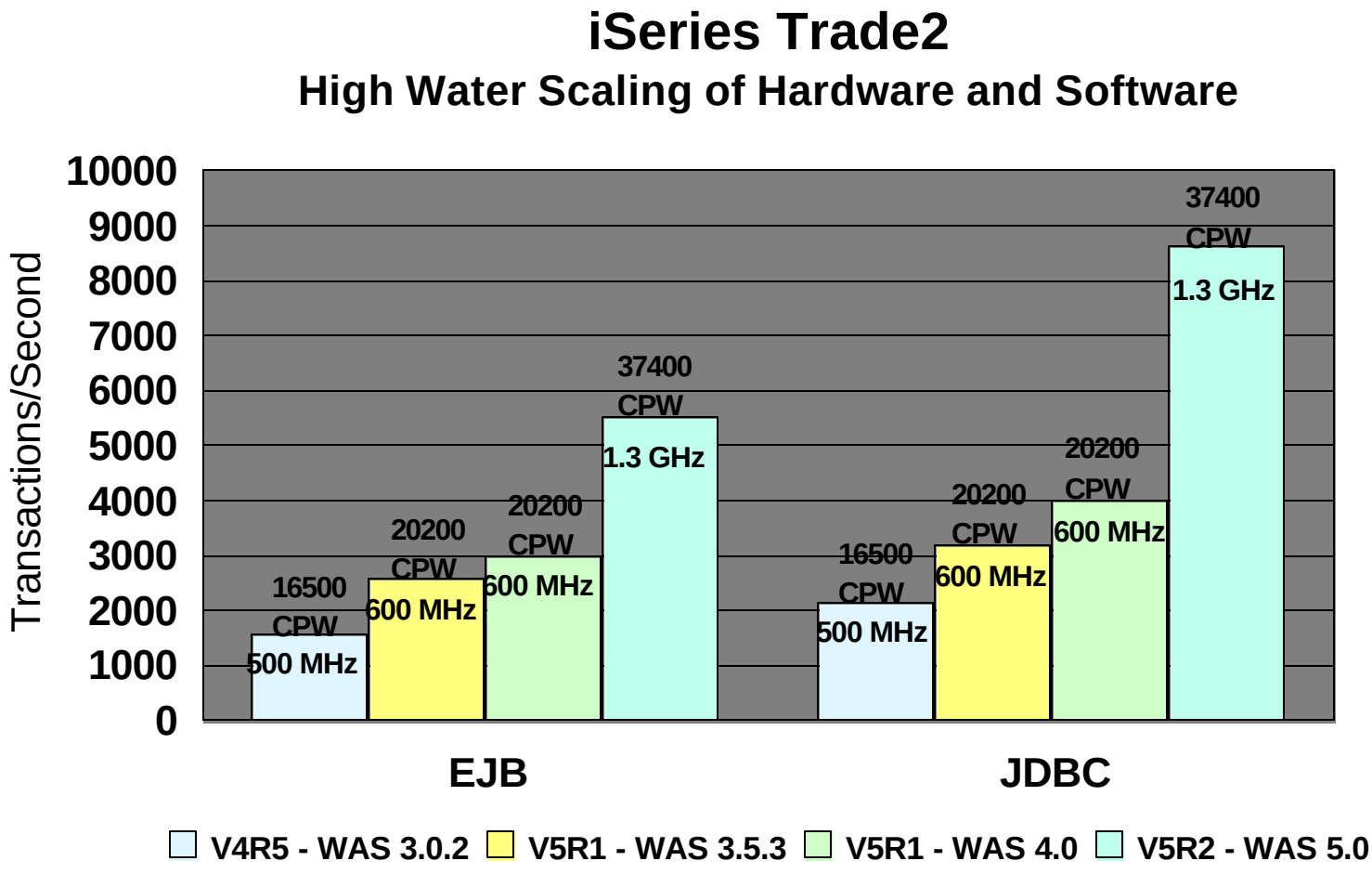
Note, the asterisk (*) shown for V5R2 WAS 4.0.3 and 5.0 test results are projected from Trade2.7, which is a 20% heavier workload than Trade 2.5.

Looking closely, you can see the V5R2 WAS 5.0 has marginal improvement over WAS 4.0.3 in the TRADEJDBC tests, but has approximately 10% throughput improvement when using EJBs.

The January 2003 Performance Capabilities Reference manual has additional details.

The next foil shows Trade2 testing on much faster iSeries systems than the 170-2385 used for the test results on this foil. This foil is used to show what can be done on a "low" CPW system. The next foils shows test results on a "high CPW" system, as an aid for understanding the scaling of the Trade2 (and similar e-business applications) across the iSeries and AS/400 family.

WebSphere Applic. Servers on iSeries: Large Systems



Notes: WebSphere Applic. Servers: Large Systems

This chart shows iSeries scalability, using historical performance test results on different releases of WebSphere application serving software as well as faster iSeries processors. Both EJB and JDBC modes are shown.

The chart includes CPW and Megahertz ratings for each test result because some of the performance between each OS/400 release and WebSphere Application Server release is attributable to improved processor speed - not just newer releases of WebSphere Application Server.

The V5R1 WAS 3.5.3 and V5R1 WAS 4.0 show improved performance using the same processor capacities.

In the January 2003 Performance Capabilities manual there are additional performance test results for "Trade2 primitives," using WAS 4.0.4, compared to WAS 5.0 using the same system - an i820-2397 (500 MHz, 2-Way). These primitives include:

- PingHtml: Provides access to Hello World static HTML page
- PingJSP: Tests a direct call to JavaServer Page providing server-side dynamic HTML through JSP scripting
- PingServlet2Servlet: Tests request dispatching. Servlet 1, the controller, creates a new JavaBean object forwards the request with the JavaBean added to Servlet 2. Servlet 2 obtains access to the JavaBean through the Servlet request object and provides dynamic HTML output based on the JavaBean data.
- PingServlet2JSP: Tests a commonly used design pattern, where a request is issued to servlet providing server side control processing. The servlet creates a JavaBean object with dynamically set attributes and forwards the bean to the JSP through a RequestDispatcher The JSP obtains access to the JavaBean and provides formatted display with dynamic HTML output based on the JavaBean data.

Testing primitives for WAS 4.0.3 and WAS 5.0 on the same processor capacity demonstrated performance improvements in all primitives between with WAS 5.0 . While these primitives are not complete e-business transactions, they did demonstrate WAS release to release performance expectations. The PingHtml primitive demonstrated the most improvement, about 3x. Other primitive improvements show measurable, but less improvement with WAS 5.0, compared to WAS 4.0.4.

WebFacing Performance Considerations

- Web-ize the applications that benefit from HTML-based user interfaces
- Refined 5250-based applications that make efficient use of 5250 hardware technology and DDS keywords use minimal system and communication resources
- Set realistic WebFacing (and other 5250-HTML tools) performance expectations compared to 5250 implementation
  - Select the right application and implement a "pilot web version" of the application
  - Depending on original 5250 application implementation, WebFaced version could require significantly more system resource
- Number of DDS fields processed directly affects performance
  - Similar consideration for direct 5250 implementation
  - Consider 5250 application Subfile implementation -> SFLPAG=SFLSIZ recommended
- January 2003 Performance Capabilities Reference manual documents test results and additional tips including use of the Workload Estimator for sizing.
  - http://www.ibm.com/eserver/iseries/perfmgmt
  - http://www.ibm.com/eserver/iseries/support/estimator

Notes: WebFacing Performance Considerations

WebSphere Development Studio Client for iSeries V5.0 includes a suite of development tools for applications that will run on iSeries servers.

Included is the IBM WebFacing Tool, which makes it easy to quickly and cost effectively Web enable your existing 5250 applications. The IBM WebFacing tool converts your 5250 application Data Description Specifications (DDS) display files, (User Interface Manager (UIM) menu source, and help files into Java Servlets, JSPs, JavaBeans, and JavaScript to allow your application to run in either WebSphere Application Server Version 5 or Version 4. Enterprise Java Beans are not used. This is an easy way to bring your application to either the Internet, or the Intranet, both quickly and inexpensively.

iSeries-specific wizards, additions to the Struts builder, and other extensions, make it easy to build e-business applications that reuse existing iSeries programs, data, and skills.

Also included in the IBM WebFacing Tool are:

- the option to generate a Web infrastructure based on the open-source Struts 3 standard that improves the extensibility, conversion-time extension 3 points to allow developers to replace the default HTML
- an option to use IBM and user-defined custom tags to enhance the customizability
- the ability to override command-key actions to invoke user-defined functions
- support for displaying and printing iSeries spool files from inside your Web-enabled applications generated using the IBM WebFacing Tool.

The interface is run under WebSphere - Express for iSeries, WebSphere Application Server V5 (Base and Network deployment editions), and WebSphere Application Server V4 Single server and Advanced Edition. Tomcat is not a supported application server.

These IBM WebFaced applications do not require 5250 CPW capacity when run on the Model 800, 810, 825, 870 and 890.

Applications that have been Web-enabled using the IBM WebFacing Tool generally perform better than applications refaced with other tools. Most other tools convert the 5250 data stream to a Web interface in a run-time conversion. This impacts the execution performance of these refaced applications. With the IBM WebFacing Tool, the Web interfaced is created at development time. During application execution, the data from the application is redirected to the Web interface created by the IBM WebFacing Tool. No 5250 data stream is created and there is no run time conversion. This up-front resource investment pays off with more efficient production operations.

Notes: WebFacing Performance Considerations - 2

It is important to note, that 5250 applications running with the new GUI interface will generally take significantly more CPU resource than if run with the original "green screen" interface under a 5250 OLTP environment. The 5250 data stream and 5250 OLTP is extremely efficient and has been fine-tuned over many years. The magnitude of the resource increase is very dependent upon the application. At the time this was written there is not enough experience with enough real, production applications to provide generalized guidelines. It is clear that the larger the percentage of time spent doing screen I/O in a 5250 application, the larger the CPW increase when running with a GUI interface. Initial results show the application will require several times the processor resource.

Note that many times the web-enabled version of the application has added functions or changed numeric data formats (for example use of floating point data, rather than packed decimal format), which require more system resource per "transaction" than the original 5250 implementation.

Understanding the 5250 application functions being brought to the web and piloting those parts of the application is recommended. For additional tips and considerations, refer to the Performance Capabilities Reference manual. This and other documentation can be found at:

- http://www.ibm.com/eserver/iseries/perfmgmt

Also, check the eServer Workload Estimator to help predict the system characteristics for these IBM WebFacing Tool enhanced 5250 applications. The Workload Estimator will ask you to specify a transaction rate (5250 panels per hour) for a peak time of day. It will further attempt to characterize your workload by considering the complexity of the panels and the number of unique panels that are displayed by the JSP. This site is updated frequently, based upon continuous feedback. You can access this tool via the web with URL: http://www.ibm.com/eserver/iseries/support/estimator.

WebFacing is performance sensitive to the number of input/output fields that are requested from WebFacing. Therefore minimizing the number of 5250 input/output fields to be processed is key - just as with basic 5250 application efficiency. Counting the number of fields processed by the original 5250 implementation is recommended. If using the subfile function, where the application reads and writes to a single screen "record, but OS/400 Work Station Data Management manages several sets of records with the same field/record definition, requires additional analysis..

Note - minimum CPW recommended for WAS applications:

- 370 CPW under WebSphere Application Server editions 4.0 or older and under WebSphere Application Server V5 Base or Networking editions.
- 300 CPW under WebSphere Applications Server V5 Express

Notes: WebFacing Performance Considerations - 3

5250 Subfile Support considerations

If the application is using workstation data management subfile support, the number of fields in a subfile significantly impacts WebFacing performance. The most important DDS keywords in this area are:

- SFLPAG - The number of rows shown on a 5250 display
- SFLSIZ - The number of rows of data extracted from the database.

When using a DDS subfile, there are 3 typical modes of operation:

- SFLPAG=SFLSIZ. In this mode, there are no records that are cached. When more records are requested, WebFacing will have to get more rows of data. This is the recommended way to run your WebFacing application.
- SFLPAG < SFLSIZ. In this mode, WebFacing will get SFLSIZ rows of data at a time. WebFacing will display SFLPAG rows, and cache the rest of the rows. When the user requests more rows with a page-down, WebFacing will not have to access the database again, unless they page below the value of SFLSIZ. When this happens, WebFacing will go back to the database and receive more rows.
- SFLPAG = (SFLSIZ) * (Number of times requesting the data). This is a special case of option 2 above, and is the recommended approach to run green screen applications. For the first time the page is requested, SFLPAG rows will be returned. If the user performs a page down, then SFLPAG * 2 rows will be returned. This is very efficient in 5250 applications, but less efficient with WebFacing.

For those unfamiliar with OS/400 5250 workstation subfile support. Here is a quick summary:

- Subfile support is intended to minimize the application logic required to manage "multiple instances" of the same record format shown on the screen as the same time. The program wants to process a single "ordered item" record at a time, but the interface to the workstation user enables the user to enter multiple item orders before sending data back to the program.
- Through subfile support, the application implementer uses DDS to define a single record (for example an order item) format and then specifies how many of these "subfile records" appear on a single screen image (SFLPAG). Optionally the DDS can specify a larger number of records to be kept within internal storage on the iSeries - the subfile size (SFLSIZ).
- When SFLSIZ is larger than SFLPAG, workstation data management, through workstation "roll up" and "roll down" keys, manages the displaying of different sets of subfile records without processing by the application itself.
- Through DDS, the end user is enabled to enter a specific function key that finally returns control to the application.
- The application deals with a single record at a time and uses "control information" to determine there are no more "changed" subfile records or a DDS keyword (SFLDSP) to identify when subfile records are to be displayed.

WebSphere Portal Server

- At the time this presentation was published there was no specific performance information available for WebSphere Portal Server V4.1.3
- See the January 2003 Performance Capabilities reference manual for the latest information.
- Use the eServer Workload Estimator for sizing
  - http://www.ibm.com/eservers/iseries/support/estimator

WebSphere Commerce

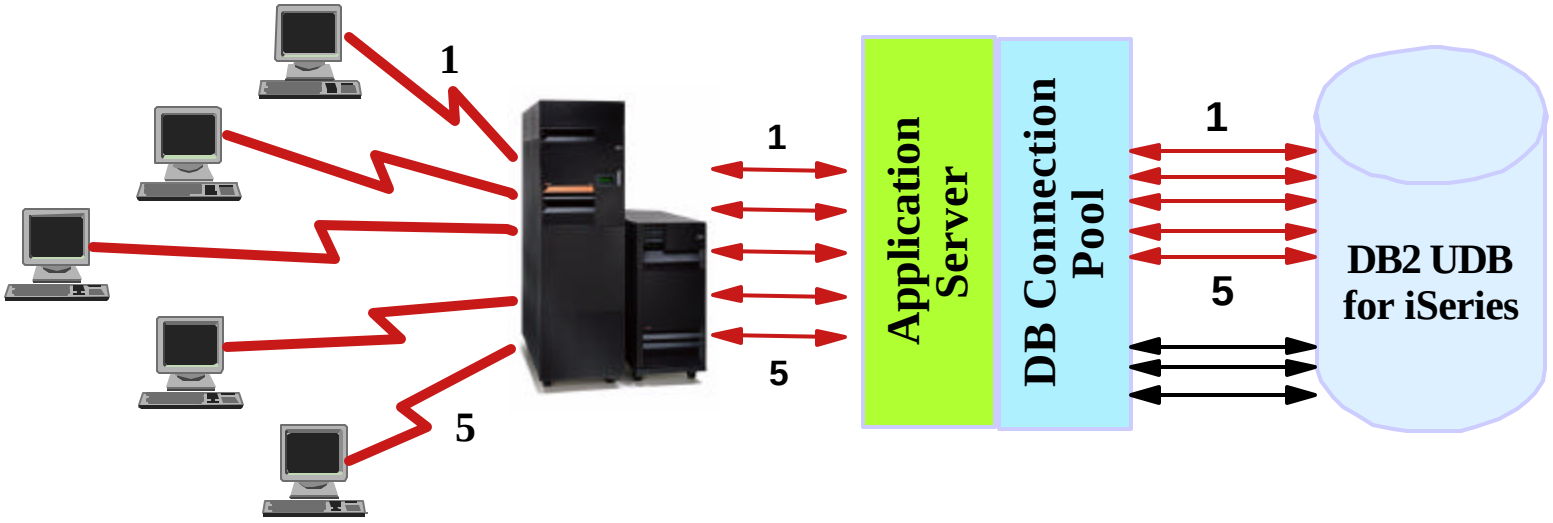
- At the time this presentation was published there was no specific performance information available for WebSphere Commerce
- See the January 2003 Peformance Capabilties reference manual for the latest information.
- Use the eServer Workload Estimator for sizing
  - http://www.ibm.com/eservers/iseries/support/estimator

Database Enhancements

Database Topics

- Transaction Services
- Multiple database "name space" support
- Statistics Manager and SQL Query performance improvements
- Index Advisor (Visual Explain)
- Database Technology Enhancements
- SQL - Identity Column Attribute
- Journal Processing

Current Application / Transaction Support



- 1 to 1 relationship between client transaction and system resources
  - One connection associated with each commit definition
- Resources not reusable until client completes transaction

Notes: Current Application / Transaction Support

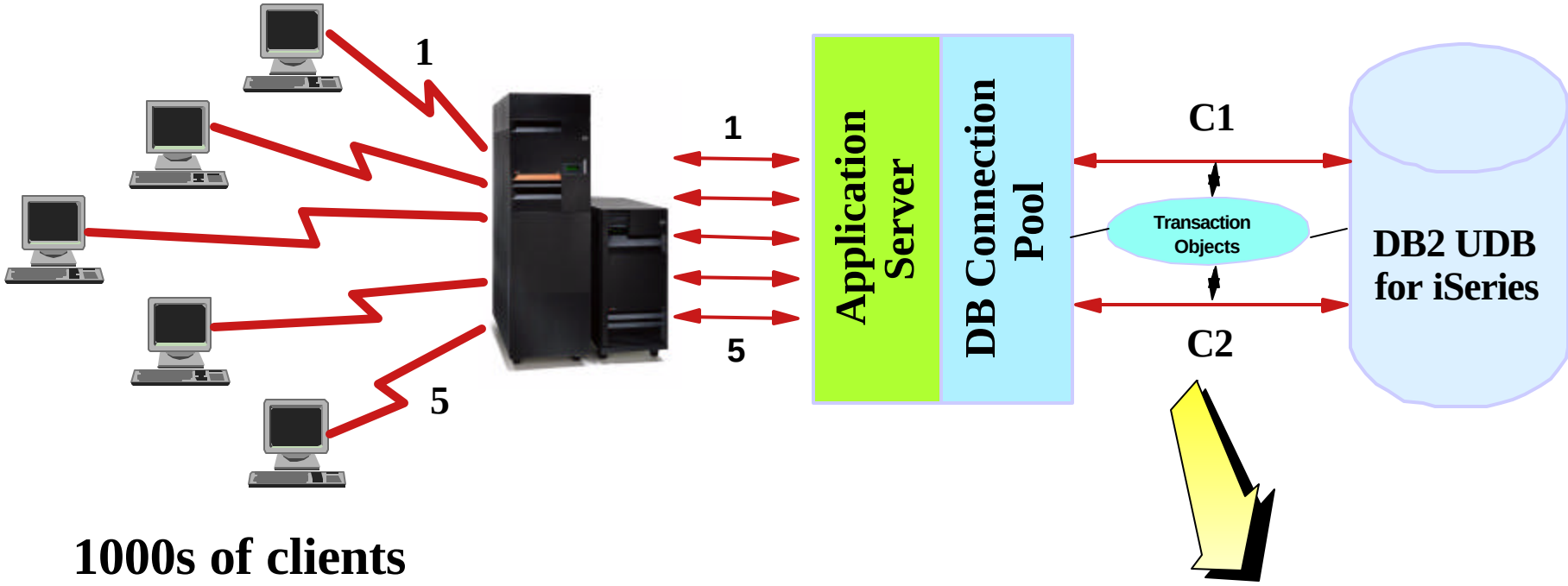
The current infrastructure in the iSeries to provide access to shared resources, such as DB2, is performed via host server jobs - QSQSRVR job names. These jobs are called whenever a request is brokered against the shared resource. This implies that the transaction boundaries are known only within the instance of the server job that will perform the entire access to the shared resource, even if this would require the job to access another platform. Each transaction is linked with the atomicity of the server job in which it is running.

Although this ;model requires little or no work to set up and has given proof of stability, it nevertheless has some implications on transaction management, compatibility and scalability; the major drawbacks being that there is a one-to-one relation between a client transaction and system resources and that resources, locked up by one transaction, cannot be made available before the transaction has terminated.

In our example we have five application server jobs each doing a "remote" connection with one of five QSQSRVR jobs. Each QSQSRVR job is responsible for all functions associated with a single connection - commit cycle transaction.

In the next foil you see how transaction management can be done starting with V5R2 of OS/400.

New Transaction Service Implementation



- Multiple connections per transaction in SQL server job or
  Multiple transactions per connection in SQL server job
- Implemented using industry standard transaction specifications (JTA, XA, CORBA, JTS)
- WebSphere Application Server, Tuxedo provide transaction management

Notes: New Transaction Service Implementation

When JTA and JDBC are used together, there are a series of steps between them to accomplish transactional work. Support for XA is provided through the XADataSource class. This class contains support for setting up connection pooling exactly the same way as its ConnectionPoolDataSource superclass. With an XADataSource instance, you can retrieve an XAConnection object. The XAConnection object serves as a container for both the JDBC Connection object and an XAResource object. The XAResource object is designed to handle XA transactional support. XAResource handles transactions through objects called transaction IDs (XIDs).

The XID is an interface that you must implement. It represents a Java mapping of the XID structure of the X/Open transaction identifier. This object contains three parts:

- A global transaction's format ID
- A global transaction ID
- A branch qualifier

The JTA API is designed to decouple transactions from JDBC connections. This API allows you to have either multiple connections work on a single transaction or a single connection work on multiple transactions concurrently. This is called multiplexing and many complex tasks can be performed that cannot be accomplished with JDBC alone - and take less system resource to perform a function.

A key implementation object is a new internal *transaction object* that maintains the transaction status, object and row/record locks, and other "commitment definition" information necessary to maintain transaction integrity. A single transaction object can be processed by multiple threads or a single thread can process a different transaction object that is passed to it. Under OS/400 the V5R2 transaction object is managed across connections and one or more QSQSRVR jobs/threads.

Use of the new transaction model is optional. It can be enabled through the use of specific XA APIs by an application program or it can be enabled by the application transaction manager - on behalf of the application. When enabled by the transaction manager, there is no change to the application program as the resource management implementation can be independent of the application programming.

Transaction managers known to support this implementation include WebSphere Application Server Advanced Edition 40. or later and Tuxedo.

For information on using the XA APIs on iSeries refer to iSeries Information Center and use the search word "xa."

Multiple Database Support

User Profile

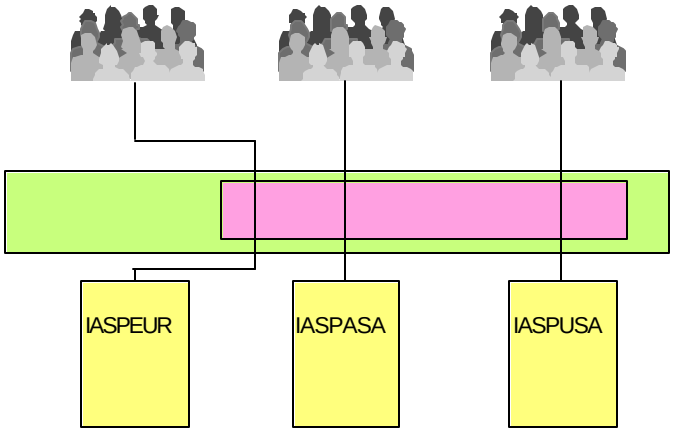
- Initial Program
- Job Description

Initial Program:

- Identifies DB libraries to be worked with, uses SETASPGRP

Job Description, per IASP user(s):

- INLASPGRP
- Library List = *SYSVAL



System Value QUSRLIBL:

- PGMGLAPAR
- PGMOE
- PGMPROD
- PGMMISC
- ...

Code Libraries in SYSBAS:

- PGMGLAPAR
- PGMOE
- PGMPROD
- PGMMISC

...DB Libraries in SYSBAS:

- MASTER

IASP Definition

Libraries per IASP DB:

- DBGLIB
- ...

Notes: Multiple Database Support

This figure shows details of using multiple databases (via multiple Independent Auxiliary Storage Pools (IASPs) on a single system. This example is intended only as what we believe to be an implementation adhering to the principle of separate placing of processes and data.

In this example three IASPs are represented in the graphic in the upper right of this foil- IASPEUR, IASPASA, and IASPUSA.

By definition, V5R2 uses the term *SYSBAS to refer to the namespace that collectively addresses all objects in the system storage pool ("system ASP" or ASP 1) and any optionally created user ASPs - now called a *base disk pool* in V5R2. In this example, library MASTER contains the database tables/files to be accessed by an application that may also access databases in one or more independent ASP.

The base assumption is to keep *as much as possible the code libraries* (programs, modules, environment descriptions) separated in the system ASP, while the data libraries are contained in separate IASPs, as required by the business entities.

The system value QUSRLIBL contains the names of the common code libraries and is used in the job descriptions associated with the different user profiles; each user profile has an initial program that sets up the data libraries the user will work with.

In a multilingual environment, you can add the libraries, containing the MRI or any other culturally dependent objects, in the system ASP also. These might also reside in the IASP.

If you write an application to access files and libraries in a disk pool group, you must specify how to access that specific database. Some of your options include:

- Use the Set ASP Group (SETASPGRP) command.
- In an SQL environment use CONNECT to specify the right database. To achieve the fastest performance, make sure that the database to which you perform an SQL CONNECT corresponds with your current library name space. You may need to use the SETASPGRP command first to achieve this. If the SQL CONNECT function is not operating within the same library name space, the application will use Distributed Relational Database Architecture(TM) support which can affect performance.
- Use the Change Job Description (CHGJOBD) command to set the initial ASP group in the job description for a user profile.

Multiple Database Performance Considerations

- Minimize number of database objects in *SYSBAS relative to IASP name space system catalog maintenance
- If connected database does not match current name space, then a remote database connection will be used
  - Current library name space can be set with SETASPGRP command or job description INLASPGRP parameter
- Minimize switches to different independent ASP groups
  - Each SQL program or SQL package access plan is unique, including database location
  - Each switch to a different database may require access plan to be rebuilt
  - Relative performance impact when access plan has to be rebuilt:
    - CRTSQLxxx (static, pre-created access plans) = 3 (highest impact)
    - Extended Dynamic (ODBC application specifies Extended Dynamic in its data source option) =2
    - Dynamic SQL (which includes all CLI applications) = 1

Notes: Multiple Database Performance Considerations

Each Independent ASP will hold its own name space objects, though we focus primarily on database objects in this topic. From each database, the catalog of the schema (collections) residing in *SYSBAS is accessible: each catalog in each database in its IASP contains also the entries of the database in the system ASP. This means that if a user or application is connected, to a database called EUROPE, the user or program can also perform operations against the database MASTER which resides in the system ASP, since these are catalogued as a single entity.

First, the vary on of an IASP takes time to correlate and rebuild the necessary name space information. If you have a large number of database objects in *SYSBAS and your IASP, the longer the vary on will take. This "catalog information" is also maintained synchronously and dynamically at run time when table changes are made. Therefore we recommend you minimize the number of database objects you have in *SYSBAS.

In an SQL environment use CONNECT to specify the right database. To achieve the fastest performance, make sure that the database to which you perform an SQL CONNECT corresponds with your current library name space. You may need to use the SETASPGRP command first to achieve this. If the SQL CONNECT function is not operating within the same library name space, the application will use Distributed Relational Database Architecture (TM) support which can affect performance.

It is important to note, that associated with an application program accessing a database through SQL there is an access plan which serves as a performance aid in accessing the database data. Access plans can be stored in programs, SQL packages, and internal job or system storage depending on the environment. SQL packages are only used for DRDA and Extended Dynamic applications.

The access plan was built knowing the location of database information for the SQL operations you have in the program or package. Changing to a different IASP could require a rebuild of this access plan. Therefore you need to either carefully control the frequency and timing of switching to a different IASP - before performance critical application processing, or consider having a duplicated "application object," that contains the SQL program or package in each IASP. The duplicate SQL program or package would more likely have an access plan that does not require rebuilding when a job or thread initiates the connection to the database.

For programs that are generated from a Create SQL xxx (RPG, COBOL, and so forth) program, the associated SQL package is generated at program creation time, if the CRTSQLxxx command specifies an RDB. An SQL package can also be created and selected by an application at run time. This would typically be the case when the iSeries server is accessed by a client using the ODBC interface. An ODBC application gets an SQL package by specifying the Extended Dynamic option on the data source.

Notes: Multiple Database Performance Considerations -2

Alternatively your program could use the SQL Call Level Interface (CLI) to perform SQL processes against tables/files. The CLI interface is used by several OS/400 system functions, including Management Central, and typically by Java applications performing SQL functions. When using the CLI interface, the system implicitly creates the information that corresponds to the SQL package/access plan the first time the using application is run.

You need to either carefully control the frequency and timing of switching to a different IASP - before performance critical application processing, **or** consider having a duplicated "application object," such as your SQL program and/or SQL package in each IASP. The duplicate SQL program/package would more likely have an access plan that does not require rebuilding.

Note the performance impact would be worse for a program with Static SQL when frequent rebuilds of the access plan are required. Dynamic SQL (which includes all CLI applications) uses internal job storage and/or system-wide statement cache available for each IASP name space. Because it is geared to real time updates, performance impact should be minimal for applications using this interface. Performance impact for Extended Dynamic should be between Static SQL and Dynamic SQL.

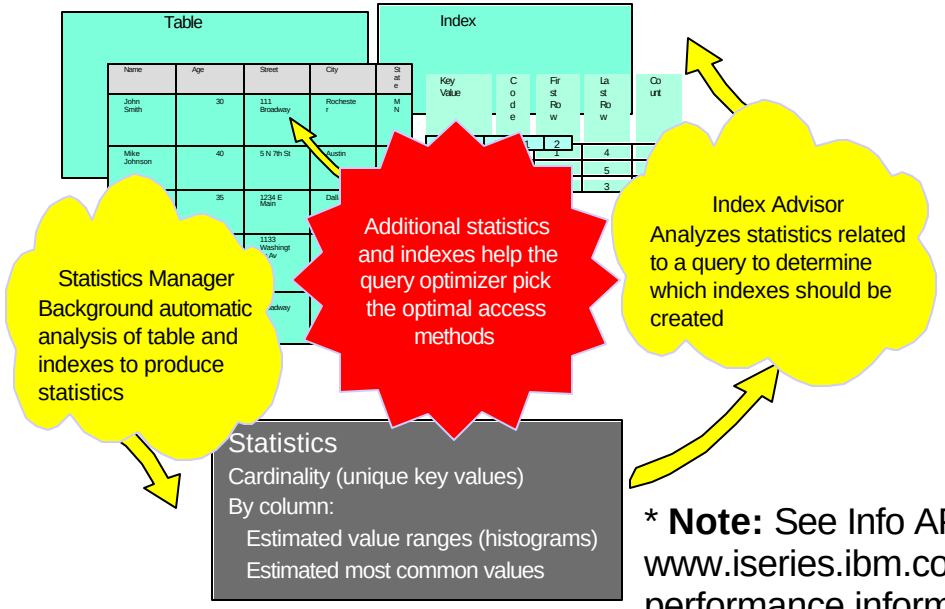
Statistics Manager

OS/400 collects statistics on column values

New system value *Database file statistics collection* (QDBFSTCCOL) specifies how the collection task will run (*NONE - *USER - *SYSTEM - *ALL (default))

Statistical usage is input for the Query Optimizer*

Management of statistical collection is done via iSeries Navigator



* **Note:** See Info APAR II13486 or www.iseries.ibm.com/db2/sqe.html for related performance information.

Notes: Statistics Manager

Statistical information and other factors can be used by the query optimizer to help determine the best access plan for a query. To be of value, this statistical information must be accurate and complete. Since the query optimizer bases its choice of access plan on the statistical information found in the table, it is important that this information be current. On many platforms, statistics collection is a manual process that is the responsibility of the database administrator. With iSeries servers, the database statistics collection process is handled automatically, and it is rarely necessary for the statistics to be manually updated, even though it is possible to manage statistics manually.

In this release, the database statistics function of iSeries Navigator gives you the ability to manage statistical information for a table.

Note: See Info APAR II13486 or www.iseries.ibm.com/db2/sqe.html for related performance information.

If you decide to collect statistics manually, and you set the statistics in iSeries Navigator to be maintained manually, not allowing the system to perform automatic updates, or if you want to speed up the automatic update process, then statistics should be updated when:

- A table is loaded or reorganized.
- A significant number of rows have been inserted, updated, or deleted.
- A new column has been added to the table.
- The Statistics Advisor in Visual Explain recommends that statistics should be created or updated.

This option to reset statistics comes in very helpful when loading tables onto a DB2 UDB for iSeries using an SQL interface that uses an access plan, as contained in a SQL package or an SQL program or procedure; initially, when the table is empty, the Query Optimizer may have chosen a table scan, while after a number of updates, an appropriate index might prove more effective.

With the *Database file statistics collection* (QDBFSTCCOL) system value you specify the types of statistics collection requests that are allowed to be processed in the background by the database statistics system job, QDBFSTCCOL. Statistical collections which are requested by either a user or automatically by the database manager to be processed in the foreground are not affected by this system value. When this system value is changed to a more restrictive value, background statistic collections in progress in the statistics system job, which are not allowed at the new value, will be ended. Their processing will be restarted when the system value is changed back to a value which would allow their processing. A change to this system value takes effect immediately.

The shipped default value is *ALL.

Self-Optimizing SQL Query Availability

- **New State of the Art SQL Query Engine (SQE)**
  - **Phased Approach providing**
    - **Improved SQL Performance**
    - **Reduced System Resources during SQL Processing**
  - **Delivery Method - see Informational APAR II13486**
    - **DB2 PTFs planned to be available 1H 2003**

Notes: Self-Optimizing SQL Query Availability

The query optimizer of previous releases is being improved in V5R2.

Note that while some improvements were included in the original V5R2 general availability software, the "total package" of performance improvements is planned to be made available though software fixes (PTFs) during 1H 2003.

In V5R2, DB2 UDB for iSeries redesigned the query engine, which may provide performance improvement for SQL read-only queries. Individual results may vary, but many workloads run within the lab up through January 21, 2003, have shown an up to 2X performance gain. Individual query performance may degrade. When the new DB2 PTFs become available, there will be significant information available providing details on the performance improvements, which types of queries can take advantage of the redesign, and how to aid the optimizer in taking advantage of the new improvements.

These sources of information include:

- Info APAR II13486 describes how to order the improvements
- www.ibm.com/eserver/iseries/db2/sqe.html on preparing for the new optimizer
- Performance Capabilities Reference Manual V5R2 updates, available at http://www.ibm.com/eserver/iseries/perfmgmt
- Red Draft (formerly Redpiece) SG24-6598, to be updated 1H 2003

Important: The query optimizer of previous releases handles queries from a number of different interfaces. The V5R2 SQE query optimizer only handles queries from SQL interfaces. For example, queries from OPNQRYF and Query/400 are not supported by SQE. OPNQRYF and Query/400 will continue to use the current Optimizer implementation.

Ensure you read the information referenced in this presentation when the support becomes available.

V5R2 Performance Summary with 1H 2003 PTFs

- Performance of some read-only queries involving complex joins and selection logic has improved
- Some simple queries showed a slight degradation
- On average, most workloads tested showed up to a 2 X improvement
- Individual results may vary
- See the following sources of performance information:
  - Info APAR II13486 describes how to order the improvements
  - www.iseries.ibm.com/db2/sqe.html on preparing for the new optimizer
  - Performance Capabilities Reference Manual V5R2 updates, available at http://www.ibm.com/eserver/iseries/perfmgmt
  - Red Draft (formerly Redpiece) SG24-6598, to be updated 1H 2003

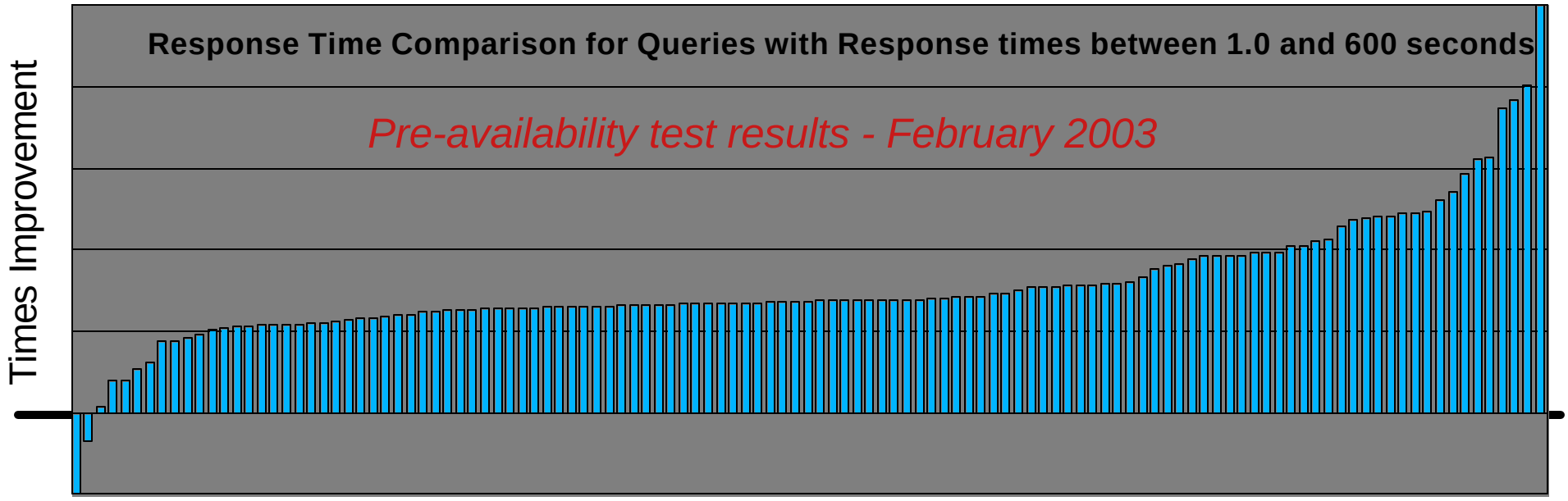
Notes: V5R2 Performance Summary with 1H 2003 PTFs

This is a summary of expectations when the new Query Optimizer set of PTFs becomes available which is planned for the first half of 2003. Many queries will improve, but some will not. A few may actually degrade.

You need to consult the available documentation listed on this chart when the PTFs become available to ensure you have appropriate expectations and instructions on how to use the new support.

The next foils show Rochester laboratory test results as of February, 2003.

Response Time Comparison - Longer Running BI Type Queries



119 queries exercising a wide range of function were selected for study. In this set of test cases:

- **Most queries demonstrated significant response time improvement, but 2 queries showed degradation**
- **The average response time improvement on a per query basis was 2.6 times across the query group**
- **The most improved query showed a response time improvement of 18 times**
- **The most degraded query showed a degradation of 2 times**

All equal or degraded queries are being examined

Consult the available performance documentation before installing PTFs when they become available

Notes: Long Running Query Response Time Comparison

This chart provides a view of the new query engine/optimizer performance by comparing response times between the new engine/optimizer and the current query engine/optimizer - for long running queries.

119 longer running (up to 600 seconds) Business Intelligence (BI) type queries were measured and a comparison of response time between the new engine/optimizer and the old engine/optimizer is graphed in the chart.

- When "Times Improvement" is positive, the new query engine/optimizer outperformed the current support by the amount shown
- When "Times Improvement" is negative, the old engine and optimizer perform better than the new query engine /optimizer by the amount shown

Here are important observations to set proper perspectives for the new support coming 1H 2003. The data shows across the range of 119 queries:

- The new query engine/optimizer in general outperforms the existing engine/optimizer with all but two of the queries performing better
- On a per query basis, the new engine/optimizer response times were on average 2.6 times faster than the current

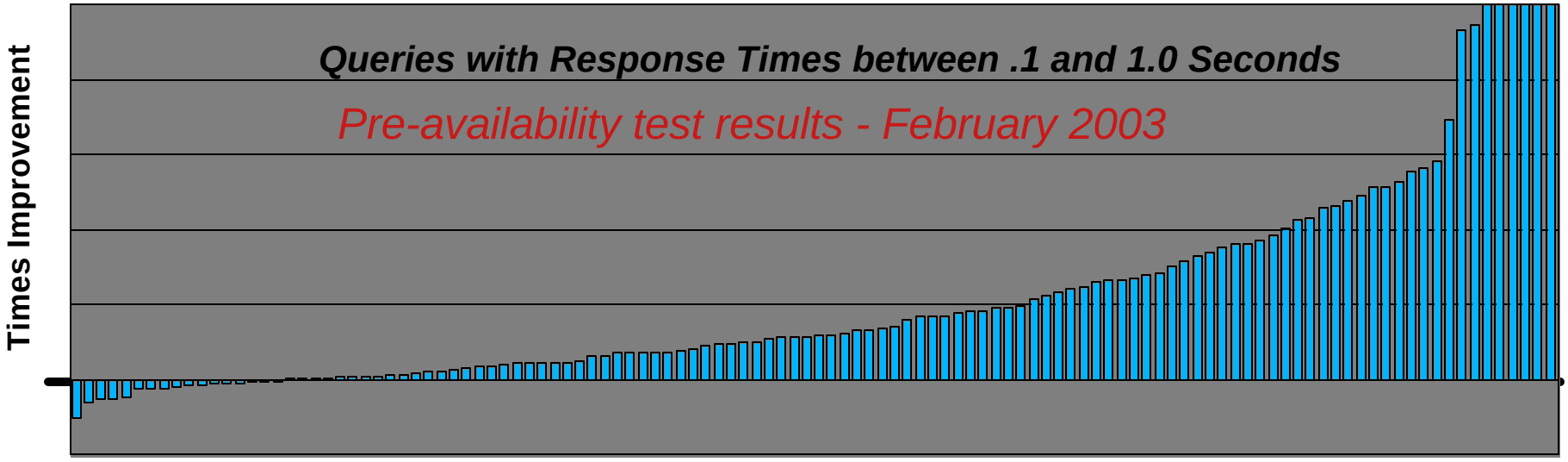
The query comparisons were made in the following environment. Results in different environments may vary:

- These queries were run in a single user environment where there was only one job active
- Individual queries were run sequentially, one query at a time
- The queries were run in a 2 Gigabyte memory pool with Query Degree set to *Optimize
- There were two CPUs present in the environment
- All of the data needed by the queries was on disk (as is typical in most BI Environments)

As stated in earlier foils, refer to the following sources for more up to date information when the new support becomes available:

- Info APAR II13486 describes how to order the improvements
- www.ibm.com/eserver/iseries/db2/sqe.html on preparing for the new optimizer
- Performance Capabilities Reference Manual V5R2 updates, available at http://www.ibm.com/eserver/iseries/perfmgmt
- Red Draft (formerly Redpiece) SG24-6598, to be updated 1H 2003

Short Running Query Response Time Comparison



118 queries exercising a wide range of function were selected for study. In this set of test cases:

- **91 queries showed significant response time improvement.**
- **19 queries showed approximately equal performance**
- **8 queries were degraded**
- **Of the 118 queries studied, the average improvement was 2.5 times**
- **The most improved query showed a response time improvement of 20 times**
- **The most degraded query showed a response time degradation of a 1/2 time (or 50%)**

All equal or degraded queries are being examined

Consult the available performance documentation before installing PTFs when they become available

Notes: Short Running Query Response Time Comparison

This chart provides a view of the new query engine/optimizer performance by comparing response times between the new engine/optimizer and the current query engine/optimizer - for short running queries.

118 short running were measured and a comparison of response time between the new engine/optimizer and the old engine/optimizer is graphed in the chart.

- When "Times Improvement" is positive, the new query engine/optimizer outperformed the current support by the amount shown
- When "Times Improvement" is negative, the old engine/optimizer performed better than the new query engine/optimizer by the amount shown

These queries are similar (in function) to simple short running transactional queries. However these queries were run in a non-transactional environment where none of the data needed by these queries existed in memory. (In memory access is often the case in transactional environments.)

Here are important observations to set proper perspectives for the new support coming 1H 2003. The data shows that:

- The new query engine/optimizer, in general, outperforms the existing engine/optimizer across the range of 118 queries
- Although some queries performed worse with the new engine/optimizer, the new engine/optimizer was on average better
- On a per query basis, the new engine/optimizer response times were on average 2.5 times faster than the current engine/optimizer

The query comparisons were made in the following environment. Results in different environments may vary

- These queries were run in a single user environment where there was only one job active
- Individual queries were run sequentially, one query at a time
- The queries were run in a 2 Gigabyte memory pool with Query Degree set to *Optimize
- There were two CPUs present in the environment

As stated in earlier foils, refer to the following sources for more up to date information when the new support becomes available:

- Info APAR II13486 describes how to order the improvements
- www.ibm.com/eserver/iseries/db2/sqe.html on preparing for the new optimizer
- Performance Capabilities Reference Manual V5R2 updates, available at http://www.ibm.com/eserver/iseries/perfmgmt
- Red Draft (formerly Redpiece) SG24-6598, to be updated 1H 2003

Performance Considerations

- When performing SQL performance analysis or sizing:
  - Sizing is not useful until performance analysis techniques have been applied
  - Use of iSeries SQL-based performance analysis tools is often required
  - Application type examples:
    - Consider how your application uses JDBC, ODBC, SQL Call Level Interfaces, or other SQL-based interfaces
    - Consider how you have defined your database interfaces using SQL in a WebSphere Application Server instance
- iSeries query-based analysis tools include:
  - Query Optimizer information available through job debug, STRDBMON command
  - iSeries Navigator SQL Performance Monitors and Visual Explain
- SQL Optimization skills are not prevalent among iSeries application developers
  - Get iSeries-based SQL optimization skills
  - An excellent source is: http://www-1.ibm.com/servers/eserver/iseries/db2/db2educ_m.htm
    - Excellent set of presentations, labs, and web-based education resources
    - Includes:
      - DB2 UDB for iSeries: Coding for SQL Performance
      - DB2 UDB for iSeries: SQL Performance Basics
      - Piloting DB2 UDB for iSeries with iSeries Navigator on V5R2 or V5R1
      - Performance Tuning DB2 UDB for iSeries w/ Ops.Nav. & Visual Explain

Notes: Performance Considerations

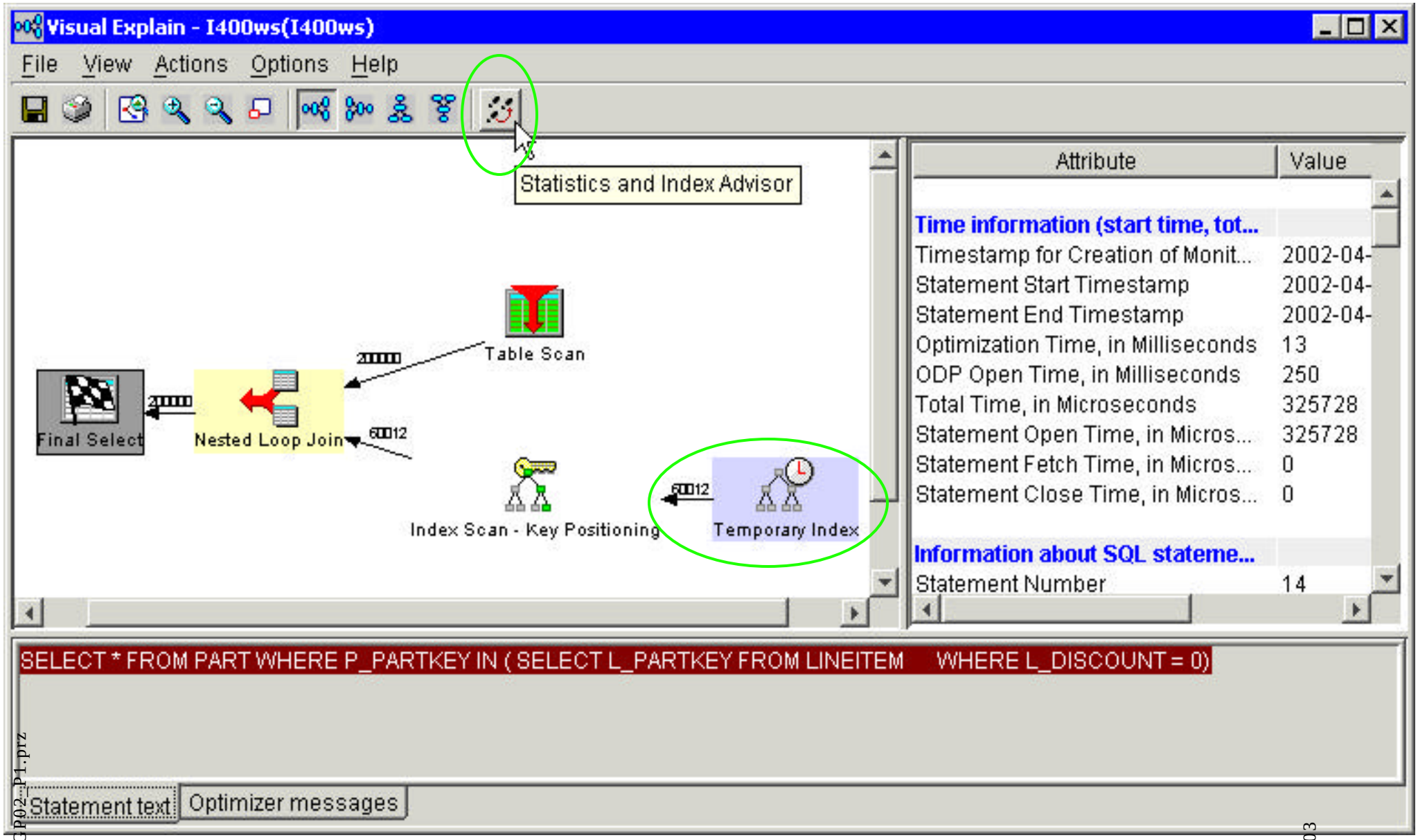
This foil lists some SQL-based application performance analysis and sizing considerations and performance analysis tools. The considerations listed lead to the subjects of iSeries SQL and query based performance analysis tools and iSeries specific SQL performance tuning techniques and education resources.

Further discussion on these tips and tools are beyond the scope of this presentation.

However, you are given a URL to an iSeries web page that has several SQL performance oriented presentations, classes, and workshops.

Note: There may be other application implementation inefficiencies than just SQL-based interfaces. For example, a database design that holds locks on rows of data for a period of time, causing other jobs/threads to wait, is independent of any interface to the database.

Visual Explain - Index Advisor



Notes: Visual Explain - Index Advisor

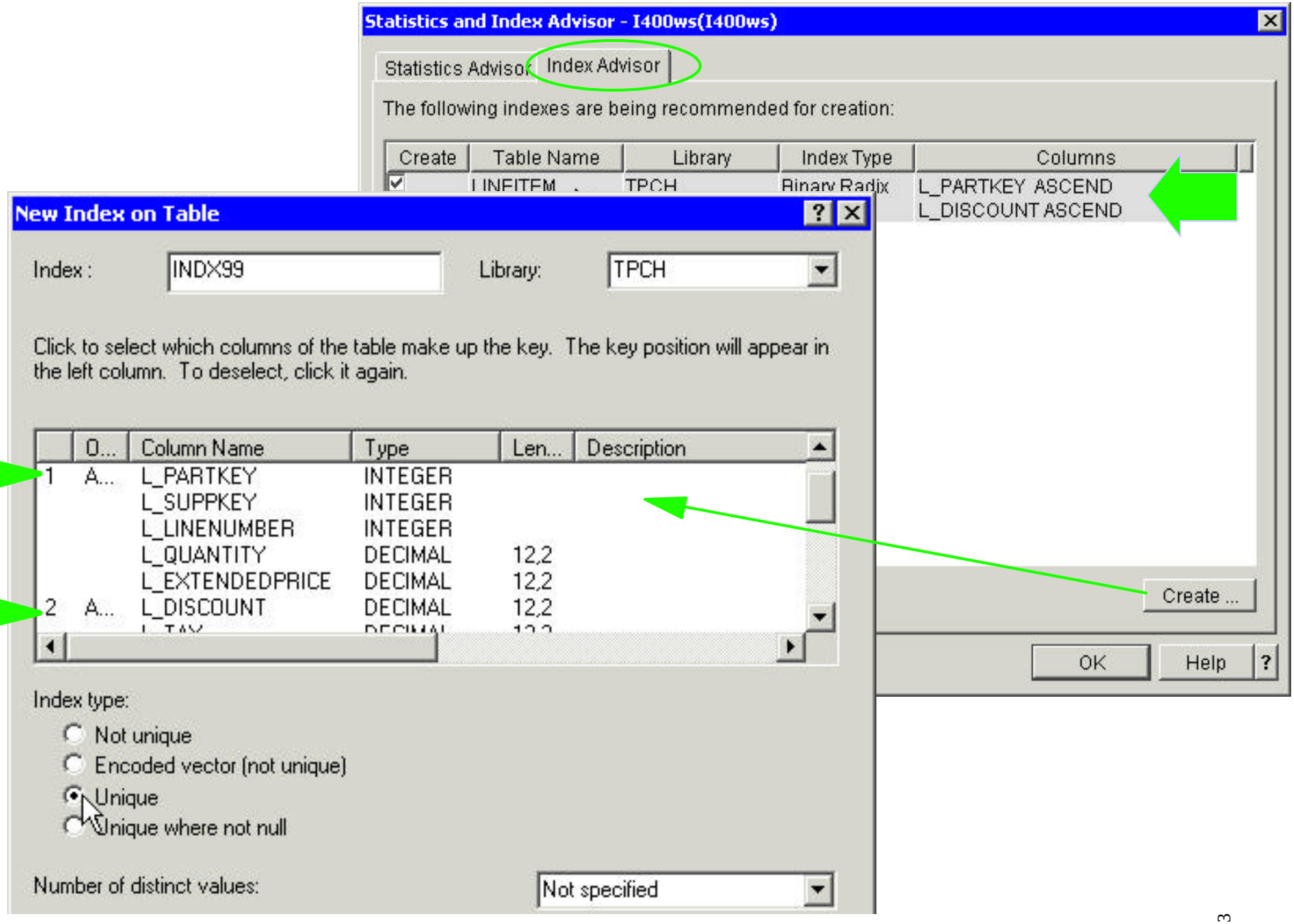
This example shows a sample Visual Explain graphic. In V5R1 you could see an "advised index set of columns" by:

Selecting a specific icon in the Visual Explain window - such as the circled temporary index on this foil. Then select the menu Actions -> Create Index

Selecting the menu View -> Highlight Index Advised and create an index based upon that advice

In V5R2 you can select the Statistics and Advisor icon as shown. This brings up the window on the next foil.

Visual Explain - Index Advisor -2



Notes: Visual Explain - Index Advisor -2

The new Statistics and Index Advisor window is shown in the upper right background. In this example, of all the possible indexes that could be created the one recommended by the system is shown. In this upper right window we see LINEITEM table name columns recommended.

By selecting Create on that first window you get the lower left window which shows the recommend key columns already selected by default for you. You enter the new index name and library, as well as select the Index type and also select the number of distinct values supported.

Database Technology Enhancements

- System Managed Access Paths Protection (SMAPP)
  - Enhanced performance and recovery options
- Encoded Vector Index Maintenance
- V5R2 Performance Gains Summary
  - Performance of some read-only queries involving complex join, selection, and grouping logic have shown improvement
  - Some queries performed the same or experienced slight degradation
  - On average, most workloads tested showed up to a 2X improvement
  - Individual results may vary
  - See the following sources of performance information:
    - Info APAR II13486 describes how to access the improvements
    - www.ibm.com/eserver/iseries/db2/sqe.html for the latest information and documentation
    - Performance Capabilities Reference Manual V5R2 at http://www.ibm.com/eserver/iseries/perfmgmt
    - SQL Query Engine RedDraft SG24-6598

Notes: Database Technology Enhancements

System Managed Access Paths

System-managed access-path protection (SMAPP) provides automatic protection for access paths. With SMAPP support, you do not have to use the journaling commands, such as STRJRNAP, to get the benefits of access path journaling. SMAPP support recovers access paths after an abnormal system end rather than rebuilding them during IPL. The shipped system automatically turns on SMAPP support.

New mechanisms to control System Managed Access Paths are available with V5R2 to allow more flexibility in what access paths are managed by the system to improve performance and reduce recovery times.

For new systems, the system-wide recovery time for access paths is 70 minutes, which is the default. If you move from a release that does not provide the SMAPP function to a release that does supports SMAPP, the system-wide recovery time for access paths is also set to 70 minutes. Otherwise, the recovery times remain as you have previously set them.

Encoded Vector Index Maintenance

As already stated with V5R1 there has been a continuing effort to improve on the maintenance of EVIs.

V5R2 Performance Gains

In V5R2, DB2 UDB for iSeries redesigned the query engine, which may provide performance improvement for SQL read-only queries. Individual results may vary, but most workloads ran internally showed a 2X performance gain. Individual query performance may be greater, degrade or stay the same. Significant information is available providing details on the performance improvements, which types of queries can take advantage of the redesign, and how to aid the optimizer in taking advantage of the new improvements. These include:

- Info APAR II13486 describes how to access the improvements
- www.ibm.com/eserver/iseries/db2/sqe.htmll for the latest information and documentation
- Performance Capabilities Reference Manual V5R2 at www.ibm.com/eserver/iseries/perfmgmt
- SQL Query Engine RedPiece SG24-6598

HA Journal Performance (OS/400 option 42)

- **The Journal Caching feature for performance**
  - PRPQ 84486 now part of OS/400 option 42 - **HA Journal Performance**
  - Not recommended if it is unacceptable to lose even one recent change
  - May not be suitable for interactive applications where single system recovery is the primary reason for using journaling

- **The Journal Standby (GHGJRN) parameter for Faster Switchover**
  - Used on the Backup system
  - Skip starting journaling for all objects
  - Not incurring the overhead of journaling
    - Most journal entries are not deposited when the journal is in standby state
    - No errors indicating that the entry was not deposited and no error messages are sent to the application

Notes: HA Journal Performance (OS/400 option 42)

For V5R2, there are a number of improvements and additions to journal management. For the most demanding high-availability clustering environments supported by our high availability Business Partners, Journal Standby Mode and Asynchronous Journaling capabilities enable faster failover and reduce performance bottlenecks. Both the Journal Caching feature and the Journal Standby feature are provided by installing OS/400 option 42.

The Journal Caching feature (PRPQ 84486 before V5R2)

The journal caching feature allows batch applications to substantially reduce the number of synchronous disk write operations performed thereby reducing overall elapsed batch execution time. You can specify journal caching with the JRNCACHE parameter on the Create Journal (CRTJRN) or Change Journal (CHGJRN) commands. Journal caching provides significant performance improvement for batch applications which perform large numbers of add, update, or delete operations against journaled objects. Applications using commitment control will see less improvement (commitment control already performs some journal caching). Journal caching modifies the behavior of traditional noncached journaling in batch. Without journal caching, a batch job waits for each new journal entry to be written to disk. Journal caching allows most operations to no longer be held up waiting for synchronous disk writes to the journal receiver. Journal caching is especially useful for situations where journaling is being used to enable replication to a second system.

- **Important**
  It is not recommended to use journal caching if it is unacceptable to lose even one recent change in the event of a system failure where the contents of main memory are not preserved. This type of journaling is directed primarily toward batch jobs and may not be suitable for interactive applications where single system recovery is the primary reason for using journaling.

The Journal Standby feature

You can use the CHGJRN command to put the journal in standby state. You might want to put a journal in standby state if the journal is on a backup system. By having the journal in standby state, a switchover to the target system can be accomplished more quickly because all objects on the backup system can be journaled thus allowing the switchover processing to skip the costly step of starting journaling for all objects. At the same time though, the backup system is not incurring the overhead of journaling because most journal entries are not deposited when the journal is in standby state.

With the journal state in standby, journal entries are not deposited into the journal. If an attempt is made to deposit an entry into the journal, there will be no errors indicating that the entry was not deposited and no error messages are sent to the application.. While in *STANDBY state, journaling can be started or stopped. However, using commitment control is not allowed while in *STANDBY state.

Note: See the Backup and Recovery information in the iSeries Information Center at http://www.iseries.ibm.com/infocenter for which journal entries are allowed to be deposited in these states.

Notes: HA Journal Performance (OS/400 option 42) -2

The Journal Caching PRPQ - 5799BJC is available for releases V4R4, V4R5, and V5R1.

The PRPQ is replaced with Option 42 for V5R2, which includes the additional STANDBY option for a remote journal. That is, option 42 is a superset of the PRPQ and includes some easier to use interfaces.

Contact IBM service to find out about a 60-day trial version prior to purchase. If necessary, you can contact Rochester IBMer Bob Gintowt who can provide more information on this.

In general, customers who have paid for a pre-V5R2 version of the PRPQ have two choices in V5R2:

- They can continue to use the pre-V5R2 version on a V5R2 machine (i.e. invoke QBJC/QJOSPEED API)
- They can upgrade to the V5R2 version - available via OS/400 Option 42). If they already have purchased the PRPQ, they get V5R2 Option 42 at a reduced price due to the fact that they've already paid for a portion of what Option 42 contains. Option 42 includes not only the journal caching support that PRPQ: 5799-BJC offers, but also broader caching support, easier interfaces on native CL commands, as well as the new High Availability offering - Journal Standby mode.

The 60-day trial version that requires a Test Fix. We haven't advertised this much and prefer the normal order process (buy up front and return it if you don't like it). But if the customer absolutely needs the trial version now, then here are the details about obtaining the trial version.

The cover letter for the Test Fix explains how to activate the journal caching. They'll need to load the first 3 PTFs (or a superceeding PTF) for the try-and-buy support. The test fix will not load without the proper SLIC and XPF prereq PTFs.

Here are the PTFs for V4R5:

- MF26772 (SLIC changes)
- SF65860 (XPF API changes)
- SA93215 (Test Fix containing QJOSPEED)

Note: SF65860 supercedes SF63192, which pre-reqs MF24870, MF24866, MF24863, MF25022, MF25078, MF25664, MF26189, MF26264, MF27116, and MF27787.

Here are the PTFs for V5R1:

- MF26785 (SLIC changes)
- SI01388 (XPF API changes)
- SE01985 (Test Fix containing QJOSPEED)

Note: Pre-req PTFs for V5R1 are MF26290, MF26295, MF26248, MF26255, MF26291, MF27167, and MF27789.

More Journal Enhancements...

- Support for library-capable independent disk pools
  - Use a secondary disk pool in a disk pool group for isolation
- Fixed-length options for journal entries
  - audit security related activity for journaled objects
- Delay automatic journal change
  - Wait before automatically attach a new journal receiver
- Delay the next attempt to delete a journal receiver

Notes: More Journal Enhancements...

Support for library-capable independent disk pools

Starting with V5R2 you can journal objects on library-capable independent disk pools.

Fixed-length options for journal entries

You can use the Fixed Length Data (FIXLENDTA) parameter of the Create Journal (CRTJRN) and Change Journal (CHGJRN) commands to audit security related activity for journaled objects on your system.

With the FIXLENDTA parameter, you can specify that the following data is included in the journal entries that are deposited into the attached journal receiver: Job name, User profile name, Program name, Program library name, System sequence number, Remote address, Thread identifier, Logical unit of work identifier, Transaction identifier

Delay automatic journal change

You can use the Manage Receiver Delay Time (MNGRCVDLY) parameter of the CHGJRN or CRTJRN commands to cause the system to wait the length of time that you specify before its next attempt to automatically attach a new journal receiver. See Manual versus system journal-receiver management for details in the iSeries Information Center.

Delay the next attempt to delete a journal receiver

Use the Delete Receiver Delay Time (DLTRCVDLY) parameter of the CHGJRN or CRTJRN commands to cause the system to wait the length of time that you specify before its next attempt to automatically delete a journal receiver. See Automatic deletion of journal receivers for details in the iSeries Information Center.

Miscellaneous topics

Miscellaneous Performance Topics

- Updated Guidelines for average CPU utilization of high priority ("interactive-like") jobs
- Reclaim Storage
- Linux
- Save, Restore
- LPAR
- Integrated File System (IFS)
- Integrated xSeries Server for iSeries save

High Priority Job CPU Utilization Guidelines

Number of Processors	Guideline - CPU utilization for "high priority jobs"
1-Way	70
2-Way	76
3-Way	79
4-Way	81
6-Way	83
8-Way	85
12-Way	87
16-Way	89
18-Way	90
24-Way	91
32-Way	93

Notes: High Priority Job CPU utilization guidelines

High Priority Job/Thread CPU Utilization Guideline Considerations

You should not wait until you have response time problems before monitoring performance issues. By monitoring CPU utilization for "high priority jobs/threads" you can be warned in advance of major problems and do something about .

Historically on iSeries and older AS/400e models we used the term "interactive job CPU utilization" for the guideline we are discussing on this chart. This originally focused only on 5250 applications. As the nature of applications being run on the system that were interactive from the end user's interface, included client/server queries, and graphical interfaces from client workstations, web serving applications, and more, we are using the term "high priority jobs" to include these newer applications as well as 5250 applications.

In any case queuing theory dictates that, depending on the resource required to respond to a request, at some point in time the utilization of that resource means response time is serious degraded. This chart focuses on the CPU utilization those of those jobs (and any associated threads) that are performing some request and associated response that is expected within a few seconds of making the request.

This topic can have some areas of debate, but for purposes of this presentation, consider the following. The guideline value for a set of jobs/threads at any given priority value then includes the cumulative utilization of **all** jobs/threads active at equal to or higher priority compared to the priority of the **selected** job(s)/thread(s). You are then monitoring the cumulative CPU utilization of all jobs/threads running at equal to or higher priority to that value. For example, we assume you want to monitor HTTP serving functions. Find out the run priority for the associated HTTP jobs/threads running under OS/400. Taking HTTP configuration defaults this would be OS/400 job/thread RUNPTY 25. By using a RUNPTY value of 25 you include the other OS/400 functions that run jobs and threads at RUNPTY values 0 (iSeries microcode) through 25. Typically this includes the system console job, 5250 workstation jobs and database serving jobs, that typically run at default priorities 10, 20, and 20, respectively.

Use a guideline value to alert you to possible response time degradation in the "near future" or a possible cause for "longer than normal" response times. The guideline values increase as the number of processors available in the system or partition (LPAR configuration) increases.

Recall that guidelines are just that -- guidelines. There may be instances when response time is very acceptable at values higher than guideline values shown. Alternatively, there may be other instances when response time is not acceptable and the CPU utilization guidelines are not exceeded.

The table shown is an updated set of guideline values for monitoring steady-state CPU utilization of "**high priority jobs**" running under OS/400 - where a response time is an important consideration. These guidelines apply to jobs/threads that perform work based upon input data and receiving a response. This work includes jobs doing 5250 transactions, web serving, Domino server functions, and SQL-oriented database file operations.

Notes: High Priority Job CPU utilization guidelines -2

Under OS/400 key high priority functions/applications with default RUNPTY values less than 25 include:

- 5250 jobs. These jobs default priority value is 20.
- QSQSRVR jobs that perform Call Level Interface DB2 UDB SQL functions used by IBM functions such as Management Central and typically by Java-based applications performing SQL. These jobs/threads default to running at priority value 10.
- QZDASOINIT jobs performing ODBC/JDBC SQL functions requested by client workstations. These jobs/threads default to running at priority 20.
- Domino for iSeries server jobs performing various Domino functions. These jobs default to running at priority 20.
- QRWTSRVR DDM Target jobs performing DDM requests from a DDM source application. These jobs default to running at priority 20.

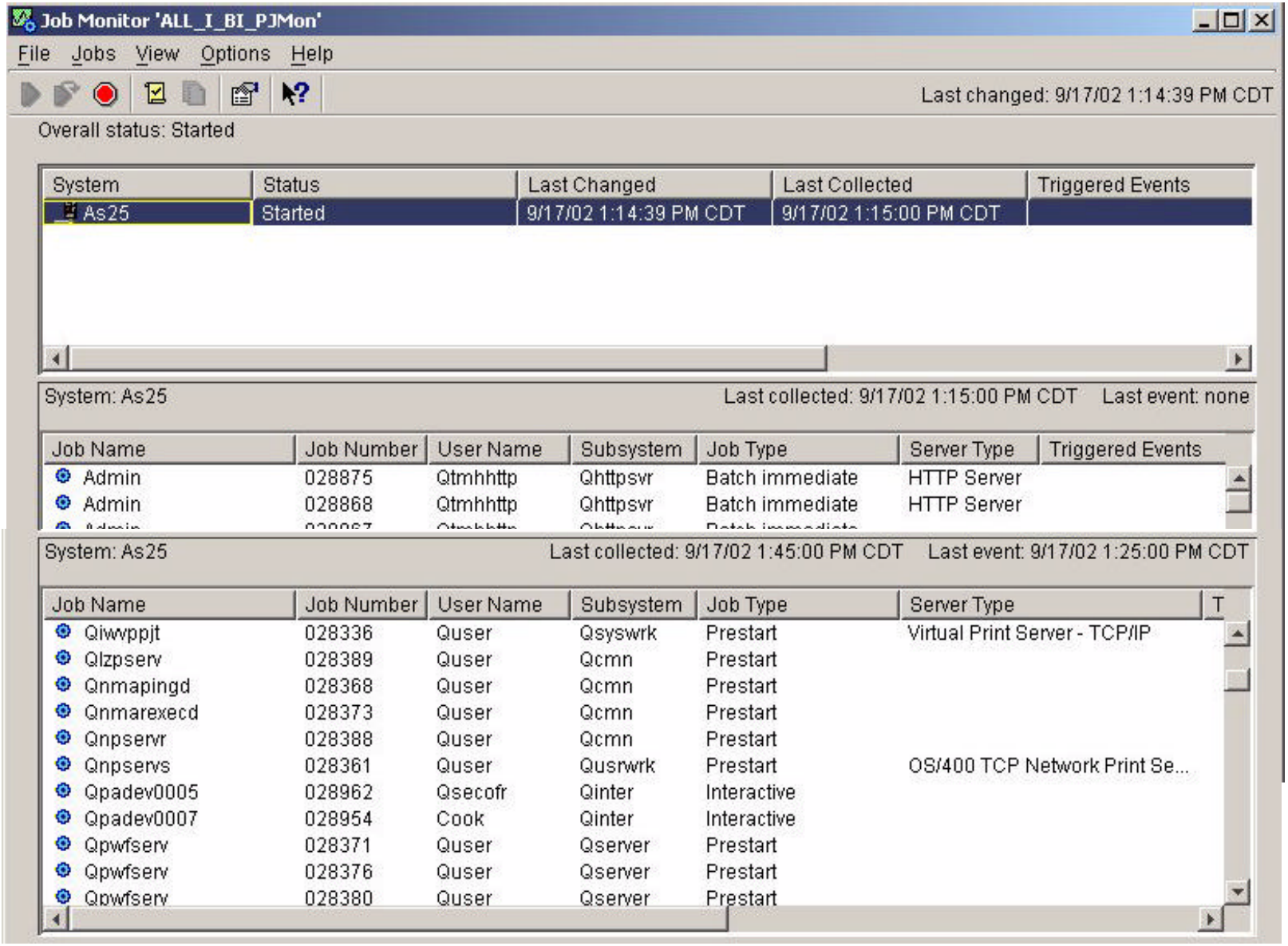
You can get an approximation of the cumulative CPU utilization according to job type run priority values from the Resource Utilization Expansion section of the System Report. This report is one of the functions available with the Manager feature of the Performance Tools for iSeries, 5722-PT1, licensed program. Performance data used by this report is collected using OS/400 Collection Services.

Starting with V5R1, iSeries Navigator (Operations Navigator) Management Central System monitor or Job monitor functions can assist you in real time monitoring of CPU utilization values for the various jobs/servers that you wish to monitor. Note that there is no specific real-time "high priority work" monitor function through V5R2. Here are some suggested techniques, all or some of which may be useful in your environment:

- Use Management Central System Monitor metric CPU Utilization (Interactive Jobs). This monitors average CPU utilization for all active jobs performing 5250 I/O operations. The percent is relative to the total processor capacity (processor CPW) of the system/partition. This does not include work performed by high priority job not performing 5250 I/O operations. For example, CPU utilization (Interactive Jobs) would not include CPU consumed by the QSQSRVR, QZDASOINIT, Domino, or HTTP -related jobs.
- Use a Management Central Job Monitor and select job types Interactive, Batch Immediate, and Prestart. Depending on your run time environment, this may cover only the jobs your are interested in. For example QSQSRVR, QZDASOINIT, and QRWTSRVR are prestarted jobs. The jobs that do most of the Domino server and HTTP server work run as Batch Immediate ("BCI") jobs. Note, there may be CPU consumed by other batch immediate and prestarted jobs in your environment.

Here are example windows of an active monitor defined for these job types. Note the various values under the Job Type and Server Type columns/

Notes: High priority CPU utilization guidelines -3



Notes: High priority CPU utilization guidelines -4

These "high priority CPU utilization" guideline values cannot be directly applied to Interactive Feature utilization. A complete discussion of a set of guideline values for the Interactive Feature utilization is beyond the scope of this tip.

However, here is some information that may assist you with setting up your own guideline values for the Interactive Feature utilization:

- Each system/partition with a low interactive CPW capacity with a single 5250 job consuming high CPU utilization over an extended period is considered an administrative job and may use "more CPU" than you may expect with a low interactive CPW capacity.
- Interactive Feature utilization applies where there are two or more 5250 jobs consuming high CPU utilization over an extended period of time. Interactive Feature utilization varies, based on the time interval of consistently high CPU utilization by jobs/threads performing 5250 input and output operations and the higher the interactive CPW capacity for the partition/system.
- In general, the Interactive Feature utilization is computed by the microcode and the algorithm can be characterized as slow to anger (high 5250 jobs CPU utilization over several minutes time span) and quick to forgive (lowered 5250 job/thread CPU utilization over a "short" time span)
- At this time, the general recommended guideline value for Interactive Feature utilization is a value of 70%. As with the high priority CPU utilization guideline there may be times of acceptable response time when Interactive Feature utilization approaches 100% and there may be times of degraded response time when less than 70% of the Interactive Feature is being utilized over an extended period of time.
- As the percent of Interactive CPW rating approaches 50% or more of the partition's (or total system) Processor CPW rating, guideline values higher than 70% may be applied.
- Additional information requires further investigation

This information assumes V5R1 or later. If you are running V4R5 you can use Management Central System monitor support, but not Job monitor support (new in V5R1). V4R5 or later Management Central System monitor support includes monitoring Interactive Feature utilization. V4R5 or later OS/400 Collection Services collects Interactive Feature utilization. After associated performance database files (QAPMcccc) have been generated, the Manager feature of Performance Tools for iSeries, 5722-PT1 can produce System and Component reports that contain Interactive Feature utilization values.

Additional information:

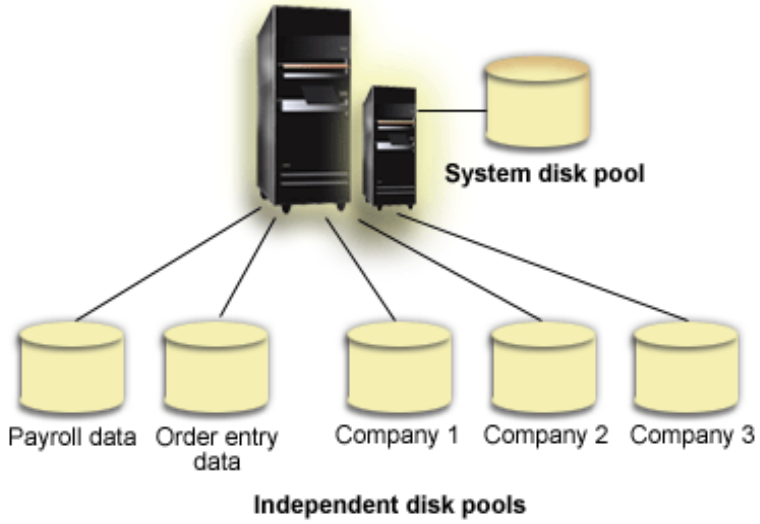
- Redpiece *Managing OS/400 with Operations Navigator V5R1 Volume 5: Performance Management*, SG24-6565, provides information on Collection Services and detailed information on using a System or Job monitors. This Red Draft can be found at http://www.ibm.com/redbooks. Search for this document by number.
- iSeries Information Center at http://www.ibm.com/eserver/iseries/infocenter. Select Systems Management -> Performance.

Reclaim Storage

Run Reclaim Storage (RCLSTG) on an IASP (introduced with V5R1)

- While the rest of the system keeps running
- Multiple ASP devices can be reclaimed in parallel

Requires that there are no users active in the IASP



Reclaim Storage

With the introduction of IASPs comes the capability to run Reclaim Storage (RCLSTG) on an IASP while the rest of the system keeps running. This implies that multiple IASP RCLSTG processes can execute simultaneously. V5R1 functional changes to the RCLSTG command added to support IASPs are:

- *SYSBAS values
  If the *SYSBAS value is specified for the ASP device, the Reclaim Storage operation runs as it does on systems prior to V5R1. The reclaim operation is performed on the system and on traditional user-defined ASPs. If the value specified is an ASP device name, then that ASP is reclaimed.
- Reclaim Storage for an ASP device (that is, an IASP) can be run without the system being in restricted state. Multiple jobs can be submitted, each performing RCLSTG on a different ASP device. Multiple ASP devices can be reclaimed in parallel.

Note: Reclaiming an auxiliary storage pool device requires that there are no active users of the ASP device that is the subject of the reclaim.

Save, Restore

- Windows File Level Back Up to OS/400 Tape
  - Signifcant V5R2 improvement
  - Over Virtual LAN rather then over the older 16 Mb Token-Ring internal LAN
- Save-While-Active performance improvements
  - Reduced checkpoint processing time
    - One test with a sample SAP application library showed that checkpoint processing that previously took 19:36 minutes now took only 1:36 minutes
- New 30/60 GB Internal QIC Tape Drive
  - New SLR60 technology
  - Up to 10x speed of 4GB tape drive, up to 2x speed of 25GB tape drive
  - 30GB (uncompressed) or 60GB (with 2x compression) per tape
- Enhanced save and restore times to save files
  - At the time this presentation was published there were no updated Save/Restore performance measurements available demonstrating improved throughput using the new Disk controller and disk devices.
- See January 2003 Performance Capability Reference Manual for details

Notes: Save, Restore

There are with V5R2 several changes made to the integration code to save Windows servers and the Virtual LAN that result in better performance for the following areas.

- Windows App to OS/400 App
- Windows to Windows Communication
- Windows File Level Back Up to OS/400 Tape
  File level backup can now be done over virtual LAN rather then over the older 16 Mb Token-Ring internal LAN. Several test scenarios showed huge time improvements. The changes to backup application (not virtual LAN support) will be PTFed back into V5R1. One test that was done is described here after, although it is a preliminary test and certainly no benchmark, is intended to give you some idea of the improvements:
  The test data used was a storage space with many subdirectories each having thousands of objects some small, some big, and many subdirectories. The total was approximately 360 MB in size. For the testing we saved data to a save file and we saved directly to a 3580 Ultrium tape drive. **All of these scenarios showed huge time improvements compared to previous OS/400 releases. Saving one of the directories (approximately 360 MB) using the previousl release took 10 minutes 6 seconds. Saving the exact same data using V5R2 took 1 minute 25 seconds.**
- The performance improvements for the save code will be PTFed back into V5R1. (Not Virtual LAN for integrated Windows servers). The series of PTFs for V5R1 (or their supercedes) that need to be applied to get the full benefit of these performance enhancementsare:
  - 5722-WSV SI04503 and for WSV as a windows service pack the SI04504
  - 5722-WSV: SI04504 Windows (W2K) service pack
  - 5722-WSV: SI04503 (OS/400)
  - 5722-SS1: SI04511 (OS/400)

Save-While-Active performance improvements

The time for checkpoint processing has been reduced significantly. One test with a sample SAP library showed that checkpoint processing that previously took 19:36 minutes now took only 1:36 minutes.

New 30/60 GB Internal QIC Tape Drive

The new 30/60 GB QIC tape drive is expected to be available June 14th, 2002. It can go up to 2x speed of a 25GB tape drive.

Enhanced save and restore times to save files

There are much more improvements to save and restore times. Please refer to the V5R2 Performance Capability Reference Guide, chapter 15.

LPAR

- Minimum required Interactive allocation for all partitions is 0%
  - Used to be approximately 1.5%
  - Primary must be running V5R2
  - V5R1 secondary partitions might donate some interactive to another secondary
    - Keep the interactive allocation to the V5R1 partition at or below the old minimum, or
    - Keep the interactive allocation of all partitions at or above the old minimum
- WRKSYSACT screen of Performance Tools
  - Current processor capacity
  - Helpful with Shared Processors Pool

```
                         Work with System Activity                    AS25
                                                          04/29/02  14:10:59
Automatic refresh in seconds  . . . . . . . . . . . . . . . . . .     5
Elapsed time . . . . . :   00:00:02    Average CPU util . . . . :      42.3
Number of CPUs . . . . . . :       4    Maximum CPU util . . . . . :    87.1
Overall DB CPU util  . . . :    53.4    Minimum CPU util . . . . . :      .1
                                        Current processing capacity:     4.00
```

Notes: LPAR

As long as the primary is running V5R2, the minimum required interactive allocation for all secondary partitions is 0%.

Secondaries running V5R1 (or any earlier releases that are allowed with a V5R2 primary) don't know about the 0% minimum, so if they see that other partitions are below the old minimum, they may attempt to donate some of their interactive capacity to the partitions which they perceive as having below-minimum allocations.

This donation mechanism is just a redundant backup in case LPAR fails to enforce the old minimum allocation, it is not documented and has no external indication.

Example of correct planning

840 12-way with 10,000 CPWs in total of which 120 is interactive CPW.

Primary partition = V5R2, and can have all 120 or some interactive CPW assigned

Secondary partition 1= V5R1 and can have 0 interactive CPW assigned

Secondary partition 2= V5R2 and can have 100 interactive CPW assigned.

Redundant donation mechanism

If you swapped the secondary interactive allocations, the V5R1 secondary with 100 CPW interactive would try to donate some interactive capacity to the V5R2 0 CPW secondary. The V5R2 secondary would be satisfied with 0 interactive, and would ignore the donation. In this case, the donated capacity would go unused. If both secondaries were running V5R1, then the below-minimum secondary would use the donated capacity.

The ways to prevent a V5R1 secondary from donating interactive capacity are to either

- keep the interactive allocation to the V5R1 partition at or below the old minimum (so it has no excess to donate)
- or keep the interactive allocation of all partitions at or above the old minimum.

WRKSYSACT screen of Performance Tools

The WRKSYSACT screen of Performance Tools has been enhanced to reflect the current processor capacity. Especially for servers with logical partitions this screen update can be very helpful.

V5R2 Integrated File performance enhancements

- In general: V5R2 File Server performance is same as V5R1 for same file type, same directory structure
- Enhanced directory performance and scalability for new directory structure
  - New *TYPE2 directory type for root, QOpenSys, and UDFS directories
  - One of a planned set of performance improvments
  - Improvement not yet significant in some environments
- TYPE 2 Byte Stream File in root, QOpenSys, and UDFS directories
  - Type 2 byte stream files introduced in V4R4
  - New auxiliary storage management attributes
  - New main storage management attributes

Notes: V5R2 Integrated File performance enhancements

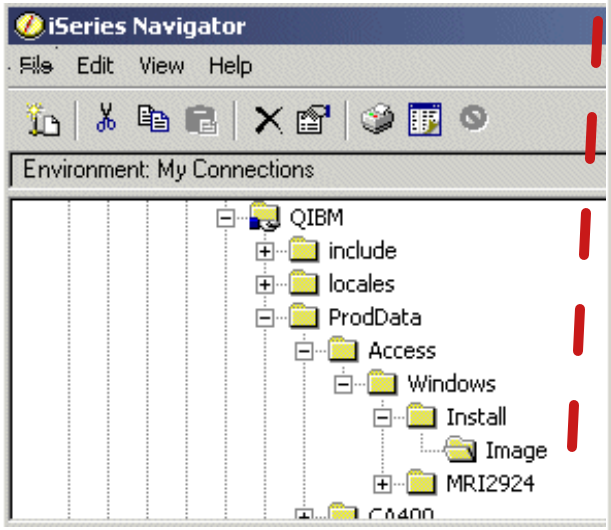
In general, when using the same byte stream file type and directory type, performance for IFS functions is equivalent between V5R1 and V5R2. However, V5R2 brings new byte stream file type 2 attributes and a new *TYPE2 directory structure that can lead to improved performance. These improvements apply to directories and files in the root (/), QOpenSys, and user-defined file systems (UDFS).

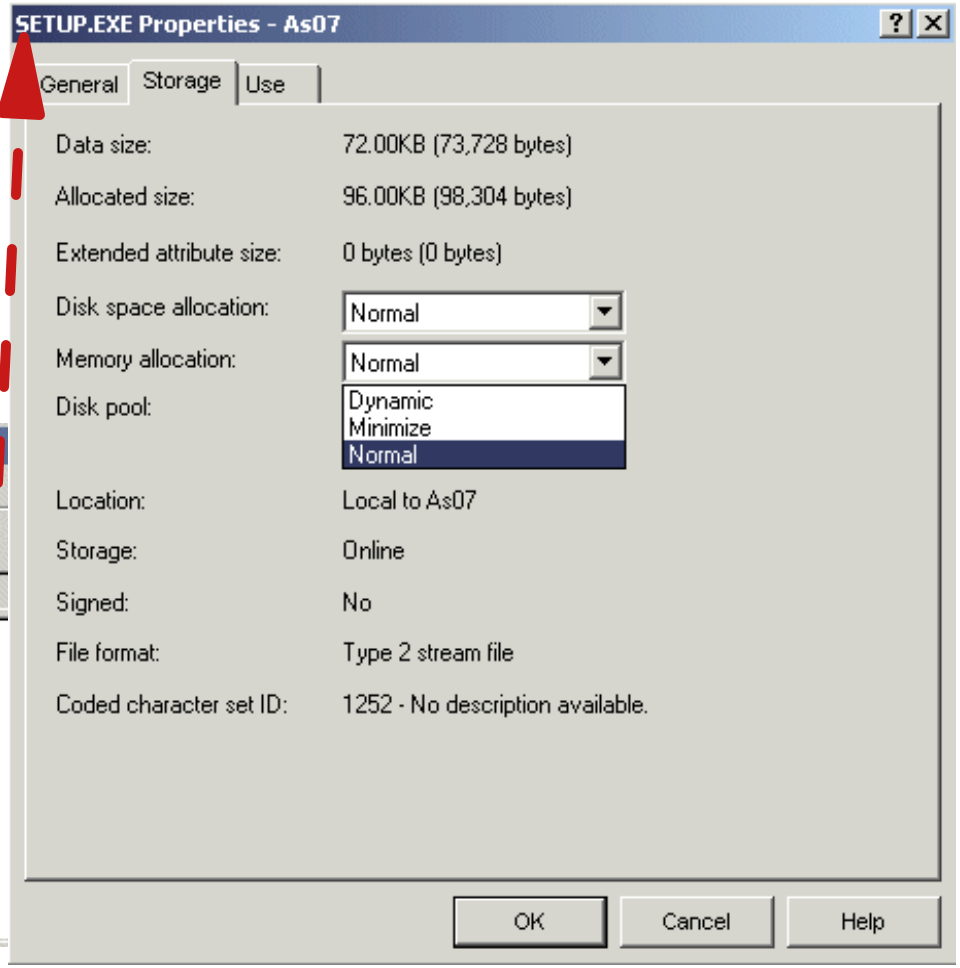
Note: Improvement is not yet significant for all accesses to the IFS "files." For example with Domino accesses to the IFS using Type1 or Type2 directories had essentially the same performance.

See the following foils for additional details.

Byte stream Type 2 file enhancements

- TYPE 2 Byte Stream File
  - *DISKSTGOPT: auxiliary storage management attributes
  - *MAINSTGOPT: main storage management attributes
  - CHGATR command, iSeries Navigator IFS interfaces
  - V5R2 first stage of performance improvements





Notes: Byte stream Type 2 file enhancements

Two new file attributes are introduced with V5R2 to influence the main storage, and auxiliary storage algorithms used by the file system. These changes are being offered as attributes to allow applications to obtain the desired affect at the file level with a minimum of software changes.

The improvements provide an adaptive, system tuned, option to influence the main storage algorithms used by the File System based on current system workload. Basically, we are introducing an Expert Cache approach on a file basis.

The CHGATR command The CHGATR command has new values for the attributes parameter. These values enable the customer or application developer to select the main storage and/or auxiliary storage algorithms used by the file system. These attributes allow applications, such as Domino, to select the desired internal file system algorithm with a simple localized change. **The two attributes are *DISKSTGOPT and *MAINSTGOPT.**

As shown on the foil (iSeries Access setup.exe file in the root directory), these attributes can also be set through iSeries Navigator.

These attributes apply to files in the root, QOpenSys, and a user defined file system (UDFS) directories.

Both options will influence the storage algorithms used to process a file. The values for these attributes are *NORMAL, *MINIMIZE and *DYNAMIC. For all details on these options, please refer to the help text of the CHGATR command or iSeries Navigator on-line help when viewing the properties of a type 2 byte steam file in the supported directories. The Dynamic values for both attributes are described here:

- *MAINSTGOPT *DYNAMIC: The system will dynamically determine the **optimum main storage allocation** for the object depending on other system activity and main storage contention. That is, when there is little main storage contention, as much storage as possible will be allocated and used to minimize the number of disk I/O operations. And when there is significant main storage contention, less main storage will be allocated and used to minimize the main storage contention. *This option only has an effect when the storage pool's paging option is *CALC.* When the storage pool's paging option is *FIXED, the behavior is the same as *NORMAL. When the object is accessed through a file server, this option has no effect. Instead, its behavior is the same as *NORMAL.
- *DISKSTGOPT *DYNAMIC: The system will dynamically determine the **optimum auxiliary storage allocation** for the object, balancing space used versus disk I/O operations. For example, if a file has many small extents, yet is frequently being read and written, then future auxiliary storage allocations will be larger extents to minimize the number of disk I/O operations. Or, if a file is frequently truncated, then future auxiliary storage allocations will be small extents to minimize the space used. Additionally, information will be maintained on the stream file sizes for this system and its activity. This file size information will also be used to help determine the optimum auxiliary storage allocations for this object as it relates to the other objects sizes.

Related commands

The WRKLNK and the DSPLNK command will display the values of the *DISKSTGOPT and *MAINSTGOPT attributes of a stream file.

At the time this presentation was created, specific performance test results were not available. Go to the Performance Management website for the latest information -> http://www.ibm.com/eserver/iseries/perfmgmt

IFS *TYPE2 Directories

- Supported for "root" (/), QOpenSys, and user-defined file systems (UDFS)
- Improved performance
  - Up to 4x for Delete, Read directory; up to 12x for Create directory (customer environments may vary)
- Improved reliability
  - If system abnormally ends:
    - *TYPE2 directories are completely recovered unless there has been an auxiliary storage failure
    - *TYPE1 directories sometimes require the use of the Reclaim Storage (RCLSTG) command
- Added functionality
  - *TYPE2 directories support renaming the case of a name in a monocase file system (for example, renaming from A to a)
  - An object in an *TYPE2 directory: up to one million links compared to 32,767 links for *TYPE1 directories.
  - iSeries Navigator displaying the list of entries: entries sorted in binary order for an *TYPE2 format directory
- Less auxiliary storage space (in many cases)
- To use an *TYPE2 on supported directories:
  - Install V5R2: new directories are *TYPE2; vary on IASP containing a UDFS , use CVTDIR command for *TYPE1 root, QOpenSys, UDFS file system directories not in an IASP already on the system

Notes: IFS *TYPE2 Directories

V5R2 supports a new IFS *TYPE2 directory structure and associated implementation for the root, QOpenSys, and User Defined File Systems (UDFSs). This delivers improved performance, reliability, some additional functionality, and for directories having less than 350 objects, reduced storage requirements.

As shown the create, delete, and read a directory functions are significantly faster than directories using the *TYPE1 directory structure available on releases prior to V5R2. The values shown are based upon development lab tests. Actual improvements in customer environments may differ.

*TYPE2 directories offer more reliability than *TYPE1 directories after a system abnormally terminates. For *TYPE2 directories there is internal journaling so that a *TYPE2 directory can be completely recovered unless there has been an auxiliary storage failure. Depending on activity on *TYPE1 directories at the time of failure, a Reclaim Storage (RCLSTG) command may be required to recover completely.

*TYPE2 directories provide the following added functionality:

- *TYPE2 directories support renaming the case of a name in a monocase file system (for example, renaming from A to a). Rename works for either *TYPE1 or *TYPE2 directory structures. However, if you rename, changing between upper and lower case characters, the rename on an *TYPE2 directory works as expected. Such a rename attempt on an *TYPE1 directory completes successfully, but the name remains unchanged.
- An object in a *TYPE2 directory can have up to one million links compared to 32,767 links for *TYPE1 directories. This means you can have up to 1 million hard links to a stream file, and an *TYPE2 directory can contain up to 1 million subdirectories.
- Using iSeries Navigator, the list of entries are automatically sorted in binary order when you open an *TYPE2 directory

Typically, *TYPE2 directories that have less than 350 objects require less auxiliary storage than *TYPE1 directories with the same objects. *TYPE2 directories with more than 350 objects are ten percent larger (on average) than *TYPE1 directories with the same objects.

There are several ways to get *TYPE2 directories on your system after V5R2 has been installed:

- A user-defined file system (UDFS) in an independent auxiliary storage pool (ASP) (33-233) is automatically converted to *TYPE2 format the first time the independent ASP is varied on ("made available" under iSeries Navigator) to a system installed with OS/400 V5R2
- All other supported file systems directories except those in UDFSs on independent ASPs must be converted to *TYPE2 by using the Convert Directory (CVTDIR) command, which is new with V5R2
- A brand new iSeries server that is preloaded with OS/400 V5R2 has *TYPE2 directories. No conversion is needed for "root" (/), QOpenSys, and UDFSs in ASPs 1-32
- A scratch install of OS/400 V5R2 on an iSeries server has *TYPE2 directories. No conversion is needed for "root" (/), QOpenSys, and UDFSs in ASPs 1-32

Notes: IFS *TYPE2 Directories -2

The V5R2 Convert Directory (CVTDIR) command can provide information on converting the supported integrated file system directories from *TYPE1 format to *TYPE2 format, or perform the conversion. The information provided includes estimates of the amount of time a conversion will take, the current directory format of the file systems and disk storage requirements of the conversion.

Only directories in root (/) and QOpenSys file systems, or a UDFS and in basic User ASPs can be converted or estimated. The system must be in a restricted state if the directories in the root (/) or QOpenSys file systems are being converted. Running the Reclaim Storage (RCLSTG) command is recommended shortly before performing a conversion to *TYPE2 format.

For additional information, refer to documentation at the following websites:

- iSeries Information Center. There is excellent information describing this support, including information on planning to convert from *TYPE1 directories to *TYPE2 directories.
  - http://www.ibm.com/eserver/iseries/infocenter. Select File systems and management -> Integrated file system -> Concepts -> Directory
- The MidrangeServer (non-IBM) web site has an excellent article on *TYPE2 directory capabilities that includes more complete performance information
  - http://www.midrangeserver.com/tfh/tfh082902-story05.html (Get Ready to Boost Your IFS Performance with V5R2)

*TYPE2 directory support became available September 2002 on V5R1. View Informational APAR II13161 for the availability notice and a pointer to any corresponding documentation for converting directories to *TYPE2 format.

Integrated xSeries Server File Backup

- Up to a 5 times improvement in the file-level backup SAV performance throughput
  - Included with V5R2
  - Requires V5R1 PTFs

Test Configuration Summary	Test Results prior to V5R1 PTFs	Test Results after V5R1 PTFs applied, V5R2 installed
Series model	270	270
Processor / CPW	2253 (2-way) / 2000	2250 (1-way) / 370
Memory	2 GB	512 MB
Disk	12 x 8 GB (10K RPM)	6 x 8 GB
Integrated xSeries Server	700 MHz	700 MHz
LAN	16 MB Token Ring	16 MB Token Ring
OS/400 Release Level	V4R5	V5R1
Files, Directories saved	722 of various sizes	1686 of various sizes
Total bytes saved	500 MB	633 MB
Saved data destination	Save file	Save file
SAV Performance Throughput	**2.2 GB per hour**	**11.6 GB per hour**

Notes: Integrated xSeries Server File Backup

V5R2 and V5R1 with PTFs deliver significantly improved file level throughput for Windows 2000 files and directories backup. The OS/400 SAV commands for files/directories in file system QNTC were used. This table compares a set of test scenario results using V5R1 to V4R5 performance test results documented in the redbook *Consolidating Windows 2000 Servers in iSeries: An Implementation Guide for the Integrated xSeries Servers for iSeries,* SG24-6056.

As the chart shows V5R1 file/directory level backup throughput can be as high as 5 times improved over V4R5 throughput.

The system used in V5R1 achieved this level of throughput increase with a less powerful (lower CPW rating) 270 server.

Performance results in your environment will vary according to specific hardware configuration and load on the system.

To achieve similar performance improvement you must have either V5R1 or V5R2 installed. V5R1 requires the following PTFs be applied to the iSeries server:

- 5722SS1 VRM510 SI04511 (OS/400 PTF)
- 5722WSV VRM510 SI04503 (OS/400 side 5722-WSV PTF)
- 722WSV VRM510 SI04504 (Service pack for 5722-WSV option 2: Windows 2000)

All 3 PTFs and their prereqs/coreqs must be applied and the special instructions in the cover letters must be followed. Then service pack PTF SI04504 must be loaded on the Windows 2000 server.

iSeries PTF SI03758 for supporting the QNTC file system must also be applied, if not already applied.

Notes: Integrated xSeries Server File Backup -2

You must also do the following configuration steps to achieve optimal file-level throughput:

- The TCP/IP send/receive buffer sizes need to be much larger than the default. We recommend no smaller than 65535. The default is 8192. The maximum size is 8 MB, which is much larger than needed. Change these values as follows:
  - Run command CFGTCP
  - Select option 3 "Change TCP/IP attributes"
  - Modify TCP receive and send buffer sizes to 65535
  - Press Enter to apply the changes
  - The changes take effect immediately.
- The latest QNTC PTF should be installed and configured - if not already done. The latest for V5R1 is SI03758. This PTF must be applied delayed. Before IPLing the machine, also create the system environment variable QZLC_SERVERLIST with a value of '1'. The environment variable name must be all uppercase: AENVVAR ENVVAR(QZLC_SERVERLIST) VALUE('1') LEVEL(*SYS)
  This environment variable is required to prevent QNTC from attempting to contact every server on the NetServer domain list when the QNTC path is traversed. If there are many servers on the list, this can talk a long time. The system must be re-IPLed for this environment variable to take effect, so we suggest it be set prior to the IPL for the QNTC PTF, if applicable to your situation.

See the following website for additional details on this support - http://www-1.ibm.com/servers/eserver/iseries/windowsintegration/. Select Product Information, then select the Windows IXS and IXA Backup Performance items.

Performance Documentation and Web Sites

- V5R2 Performance Capabilities Reference
  - Available via public Web site, Adobe Acrobat
  - Latest Java, WebSphere, Linux, MQSeries, DASD, Save/Restore ..internal test results and tips
- Performance Monthly Newsletter and Website for IBMers
  - http://ca-web.rchland.ibm.com/perform/perfmenu.htm
- Public Performance Management Web site for Customers and BPs
  - http://www.ibm.com/eserver/iseries/perfmgmt
  - Includes access to Performance Capabilities Reference manuals V5R2 - V4R3 sizing tips, access to Workload Estimator, ...
- Workload Estimator and PM eServer iSeries via Web site
  - http://www.ibm.com/eserver/iseries/support/estimator
  - http://www.ibm.com/eserver/iseries/pm400
- iSeries performance management and using Performance Tools for iSeries, 5722-PT1
  - http://www.ibm.com/eserver/iseries/infocenter
    - Systems management -> Performance
- PM eServer iSeries (Performance Manager/400e)
  - http://www.ibm.com/eserver/iseries/pm400/
- Examples using iSeries Navigator performance management functions
  - Managing OS/400 with Operations Navigator V5R1: Performance Management, SG24-6565 (contains appendix on V5R2 functions)

Trademarks

References in this document to IBM products or services do not imply that IBM intends to make them available in every country.
The following terms are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both:

400	BRMS	Host Integration Series	JustMail	Payment Manager	Stylized
ADSTAR	Client Series	Host on Demand	MQSeries	Payment Server	SystemView
Advanced Function Printing	ClusterProven	Host Publisher	MQSeries Integrator	PCOM	VisualAge for Java
AFP	CODE/400	HTTP Server for AS/400	Net.Commerce	PowerPC	VisualAge for RPG
AIX	DataGuide	IBM	Net.Data	PowerPC AS	WebSphere
AnyNet	DB2	IBM Logo	Netfinity	Print Service Facility	WebSphere Advanced Edition
Application Development	DB2 Extenders	IBM Network Station	NetView	pSeries	WebSphere Commerce Suite
APPN	DB2 UDB for AS/400	Information Warehouse	NUMA-Q	PSF	WebSphere Development Tools for AS/400
AS/400	DB2 Universal	Integrated Language Environment	OfficeVision	S/390	WebSphere Standard Edition
AS/400e	e-business logo	Intelligent Printer Data Stream	OS/2	San Francisco	Workpad
AT	e(logo) Server,	IPDS	Operating System/400	Screen Publisher	xSeries
BrioQuery	Enterprise Storage Server	iSeries	OS/400	SmoothStart	

cc:Mail, Domino.Doc, Freelance, LearningSpace, Lotus, Lotus Domino, Lotus Notes, iNotes, QuickPlace, Sametime, and Word Pro are trademarks of Lotus Development Corporation in the United States, other countries, or both.
Tivoli and NetView are trademarks of Tivoli Systems Inc. in the United States, other countries, or both.
C-bus is a trademark of Corollary, Inc. in the United States, other countries, or both.
Java and all Java-based trademarks and logos are trademarks or registered trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.
Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.
PC Direct is a trademark of Ziff Communications Company in the United States, other countries, or both and is used by IBM Corporation under license.
ActionMedia, LANDesk, MMX, Pentium and ProShare are trademarks of Intel Corporation in the United States, other countries, or both.
IBM's VisualAge products and services are not associated with or sponsored by Visual Edge Software, Ltd.
Linux is a registered trademark of Linus Torvalds.
UNIX is a registered trademark of The Open Group in the United States and other countries.
SET and the SET Logo are trademarks owned by SET Secure Electronic Transaction LLC.
Other company, product and service names may be trademarks or service marks of others.

Information is provided "AS IS" without warranty of any kind.

All customer examples described are presented as illustrations of how those customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics may vary by customer.
Information in this presentation concerning non-IBM products was obtained from a supplier of these products, published announcement material, or other publicly available sources and does not constitute an endorsement of such products by IBM. Sources for non-IBM list prices and performance numbers are taken from publicly available information, including vendor announcements and vendor worldwide homepages. IBM has not tested these products and cannot confirm the accuracy of performance, capability, or any other claims related to non-IBM products. Questions on the capability of non-IBM products should be addressed to the supplier of those products.
All statements regarding IBM future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only. Contact your local IBM office or IBM authorized reseller for the full text of the specific Statement of Direction.
Some information in this presentation addresses anticipated future capabilities. Such information is not intended as a definitive statement of a commitment to specific levels of performance, function or delivery schedules with respect to any future products. Such commitments are only made in IBM product announcements. The information is presented here to communicate IBM's current investment and development activities as a good faith effort to help with our customers' future planning.
Performance is based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput or performance that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput or performance improvements equivalent to the ratios stated here.
Photographs shown are of engineering prototypes. Changes may be incorporated in production models.