

Performance Management Update

Part 2 of 2

2001 Announcements
ITSO Technical Overview
May 2001

IBM @server. For the next generation of e-business.

Performance Update

- General V5R1 Performance Expectations
- Processors-based Updates
 - Relative Processor Ratings
 - Benchmarks Update
- TCP/IP-based Update
- Java-based Update
- I/O-based Update
- Disk Arm Requirements Calculator
- Additional Sources of Performance Information
- Processor Technology

General V5R1 Performance Expectations

IBM @server. For the next generation of e-business.

Innovative iSeries Technology

IBM  server iSeries

Mid-market industry leading server performance

Refreshed iSeries product line - CPW, CIW

- New processors for 840, 820, 270 models
- New upgrades into Capacity Upgrade on Demand

DSD and Domino - MCU

- NotesBench
- Complementary Workloads

Java™

- VolanoMark, SPECjBB, Trade2

TPC-C

JD Edwards - high watermark test

Deliverables

- V5R1 Performance Capabilities Reference Guide
- Workload Estimator (WLE) Enhancements
- Monthly Newsletter from Competitive Analysis



IBM  server. For the next generation of e-business.

Notes: Innovative iSeries Technology

This foil is a good overview of iSeries performance, based upon a technology view.

Additional details are included in this presentation. However, it is worth noting that new SStart technology-based processor features demonstrates IBM's continuing commitment to iSeries. We expect to get continued performance upgrades in environments where processor power is the primary performance component, such as TPC-C, Domino workloads, Java-based workloads and a planned update on JD Edwards "high water mark" testing (for example, 4000 active users).

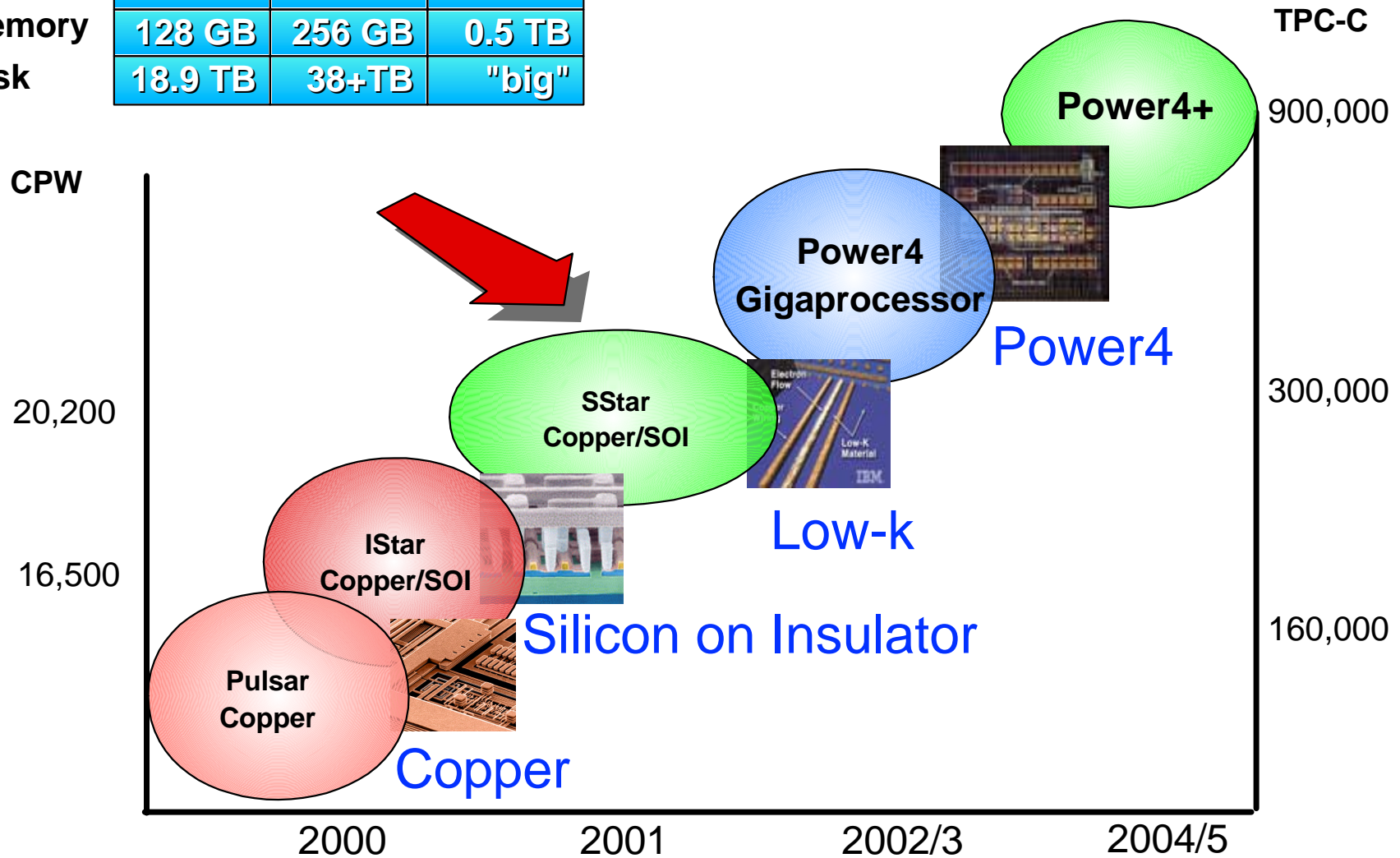
Here you see an update on the key "performance deliverables" available in previous releases:

- Performance Capabilities Reference manual (Internet web site at
 - <http://publib.boulder.ibm.com/pubs/html/as400/online/chgfrm.htm> (What's New Topic)
- Workload Estimator
- IBM Intranet performance newsletter ("ca-web" website)

iSeries Processor Roadmap

2000/1 2002/3 2004/5

SMP	24	32	64
Memory	128 GB	256 GB	0.5 TB
Disk	18.9 TB	38+TB	"big"



IBM  server. For the next generation of e-business.

Notes: iSeries Processor Roadmap

With SStar processors and V5R1, iSeries now delivers its seventh generation of 64-bit processors.

The roadmap shown here is the roadmap for IBM eServer. Both iSeries and pSeries customers benefit from IBM's ability to deliver innovative technologies that provide industry leading performance, based on exploiting scalable 64-bit symmetric multiprocessing. With support for Linux on PowerPC, the IBM eServer advantage extends to give customers robust and scalable alternatives for implementing the next generation of e-business applications.

iSeries was the first server in the industry to deliver Silicon-On-Insulator and, combined with Copper chip technology, has forged a new role as an enterprise server.

IBM has now announced another milestone in semiconductor manufacturing: a new method for building microchips that can deliver up to a 30 percent boost in computing speed and performance. This new manufacturing technique uses a material technologists refer to as a "low-k dielectric" to meticulously shield millions of individual copper circuits on a chip. IBM is the first to use the low-k dielectric technique with copper wiring. For more information on IBM's chip technology including Copper, SOI, Low-k and Power4, [see http://www.chips.ibm.com/bluelogic/showcase/](http://www.chips.ibm.com/bluelogic/showcase/)

While iSeries, with its 24-way SMP, sets impressive performance benchmarks at the high end, it is also able to scale down, bringing the same processor and robust operating system advantages to small businesses. With SStar, iSeries is again able to bring the latest Copper and SOI technology down to smaller servers and, in doing so, reaffirms its strategy for entry products designed for small business.

The iSeries roadmap includes future exploitation of Power4 and Power4+ processors, with over 1 GigaHertz clock speed. Of course, iSeries is designed with business applications requirements for balanced system performance in mind. Therefore, along with new 64-way processor technologies, our system roadmap includes designs for supporting up to .5 TB of memory and massive disk storage capabilities to match the growth expectations of large corporations.

IBM  server. For the next generation of e-business.

Notes: iSeries Processor Roadmap-2

Of course, while any roadmap is subject to change, AS/400 and iSeries has an impressive track record of delivering on past roadmaps. Unlike some technology competitors, AS/400 and iSeries has delivered on its promise of exploiting 64-bit processors and has done so without forcing customers to rewrite their applications. Note that Intel announced Merced in 1995, the same year AS/400 provided customers a seamless transition to its first generation of 64-bit computing.

IBM  server. For the next generation of e-business.

Release-to-Release Expectations

Minimal degradation or improvement in existing areas:

- Within + or 5% V4R5
 - Collect current release performance information before upgrading to any new release (for example, V5R1)
- General code path improvements in TCP/IP and Database
- Largest single release in iSeries history

New V5R1 Dedicated Server for Domino algorithm for existing models

Notes: Release-to-Release Expectations

IBM  server iSeries

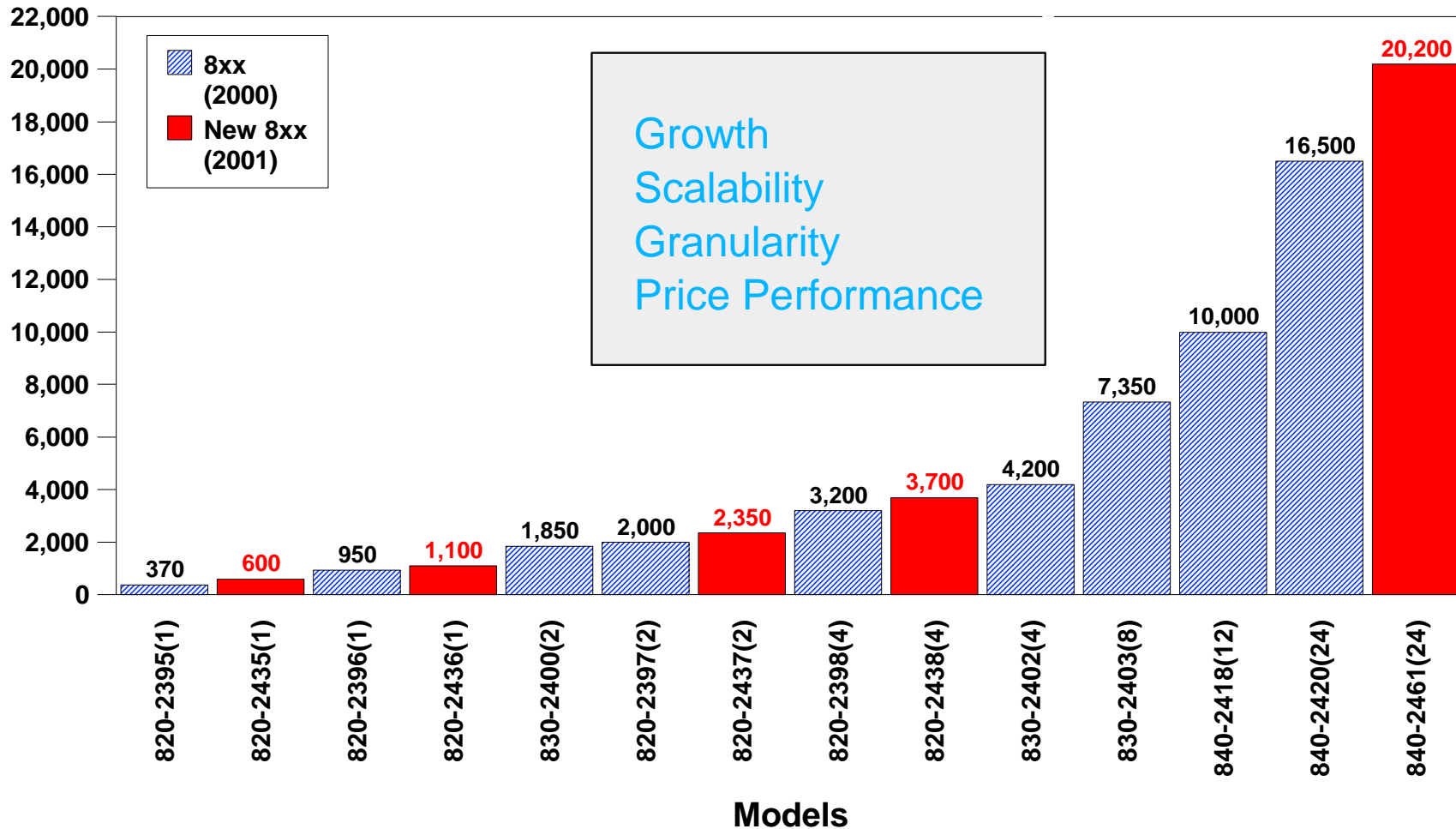
In the April 2001 announcements, under Dedicated Servers for Domino (DSD) we indicated an algorithm improvement for DSDs that will treat Java and WebSphere-based jobs/tasks as Domino work. We discuss more on this later in the DSD section.

The next foils starts a section showing where the new processors announced in April 2001 "fit" with the iSeries CPW-based performance ratings.

IBM  server. For the next generation of e-business.

8xx Processor Performance

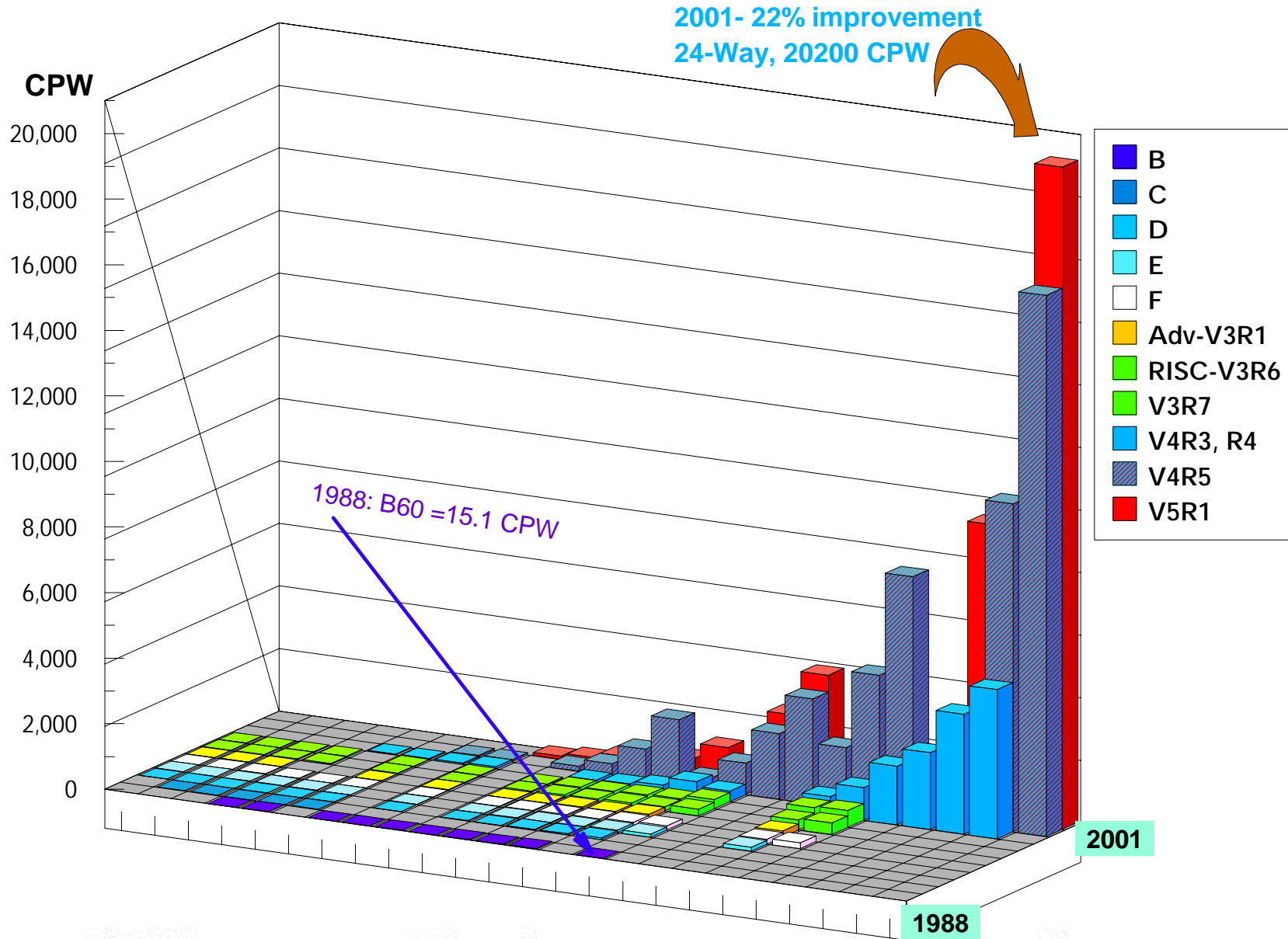
Processor CPW



Processor Commercial Processor Workload (CPW) values are used. CPW is a relative measure of performance of iSeries and AS/400 families of processors. Performance in customer environments may vary. The value is measured on maximum configurations.

IBM  server. For the next generation of e-business.

iSeries and AS/400 Scalable Performance



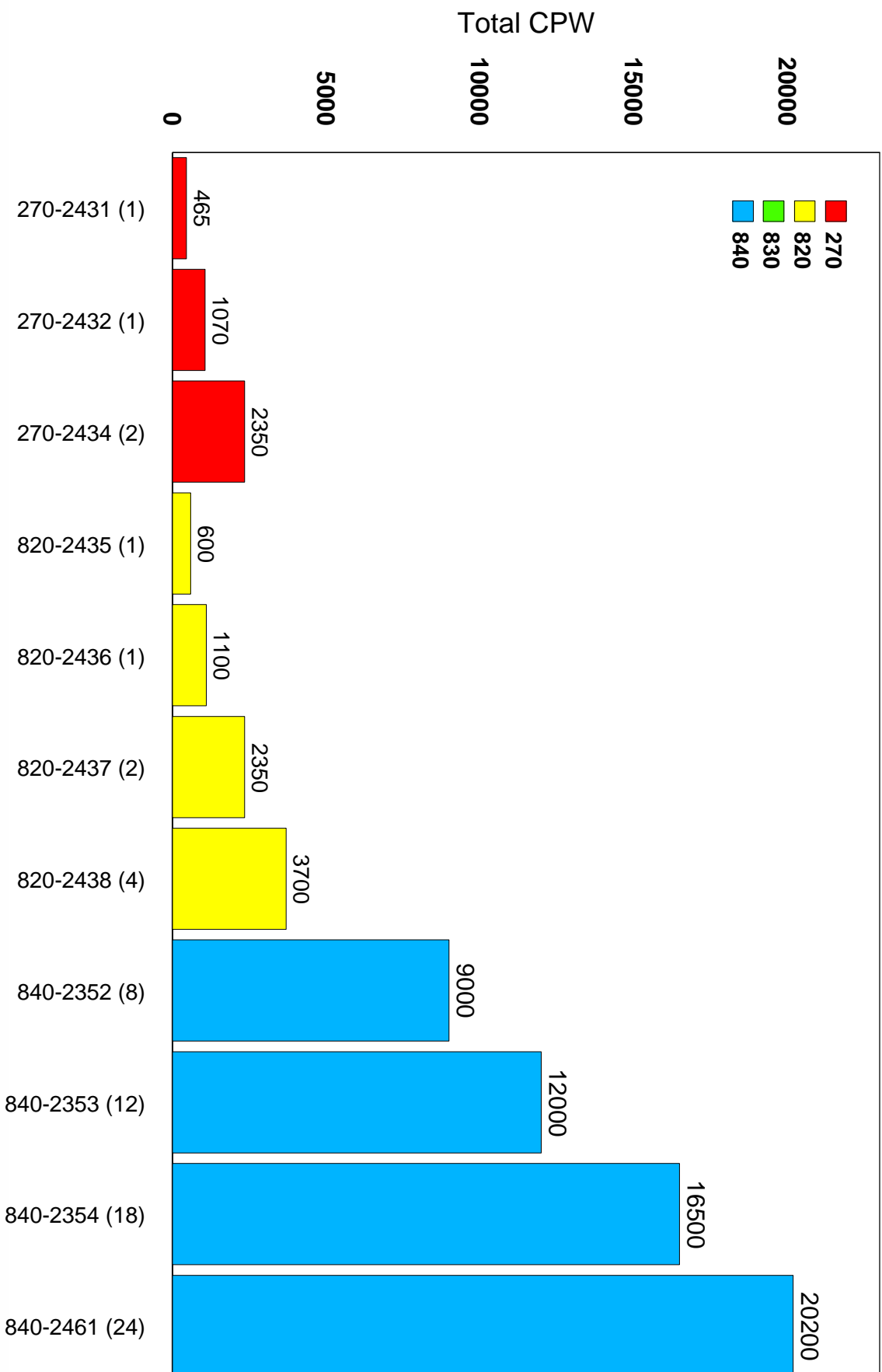
IBM  server. For the next generation of e-business.

Relative Processor Ratings

IBM @server. For the next generation of e-business.

V5R1 SStar Processors (CPW)

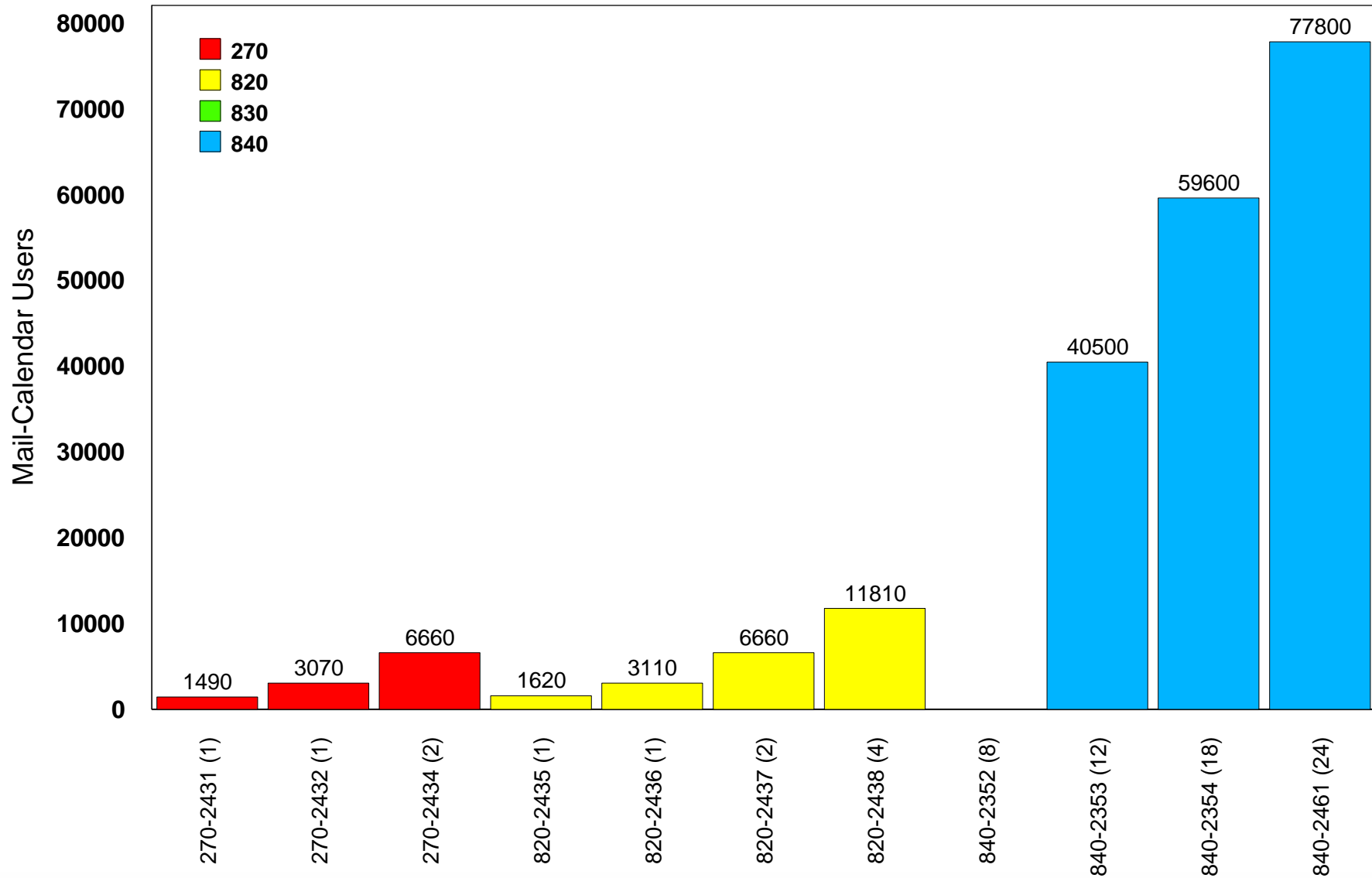
IBM  server iSeries



IBM  server. For the next generation of e-business.

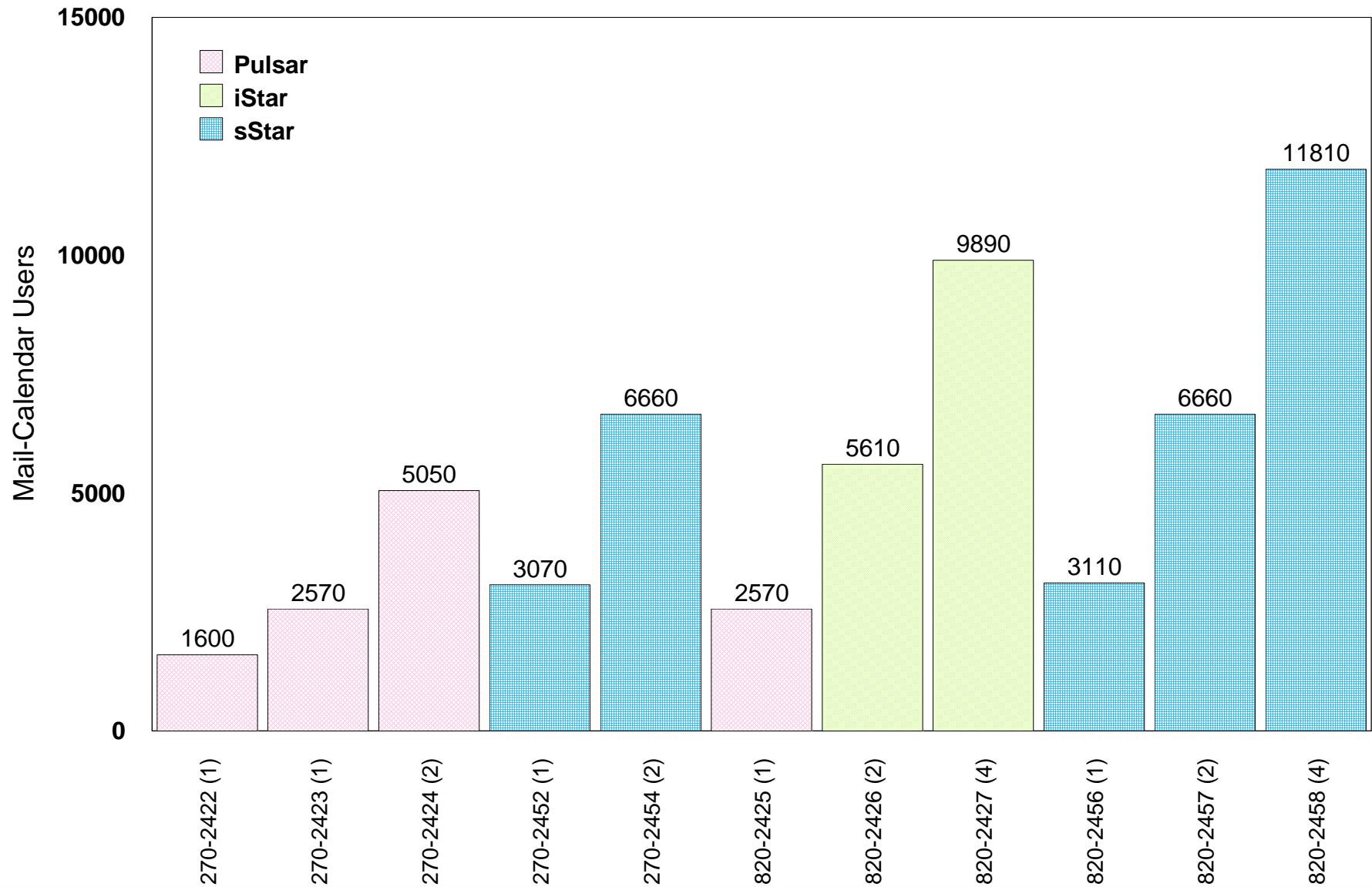
© 2001 IBM Corporation

V5R1 SStar Processors (Domino/Mail-Calendar) IBM server iSeries



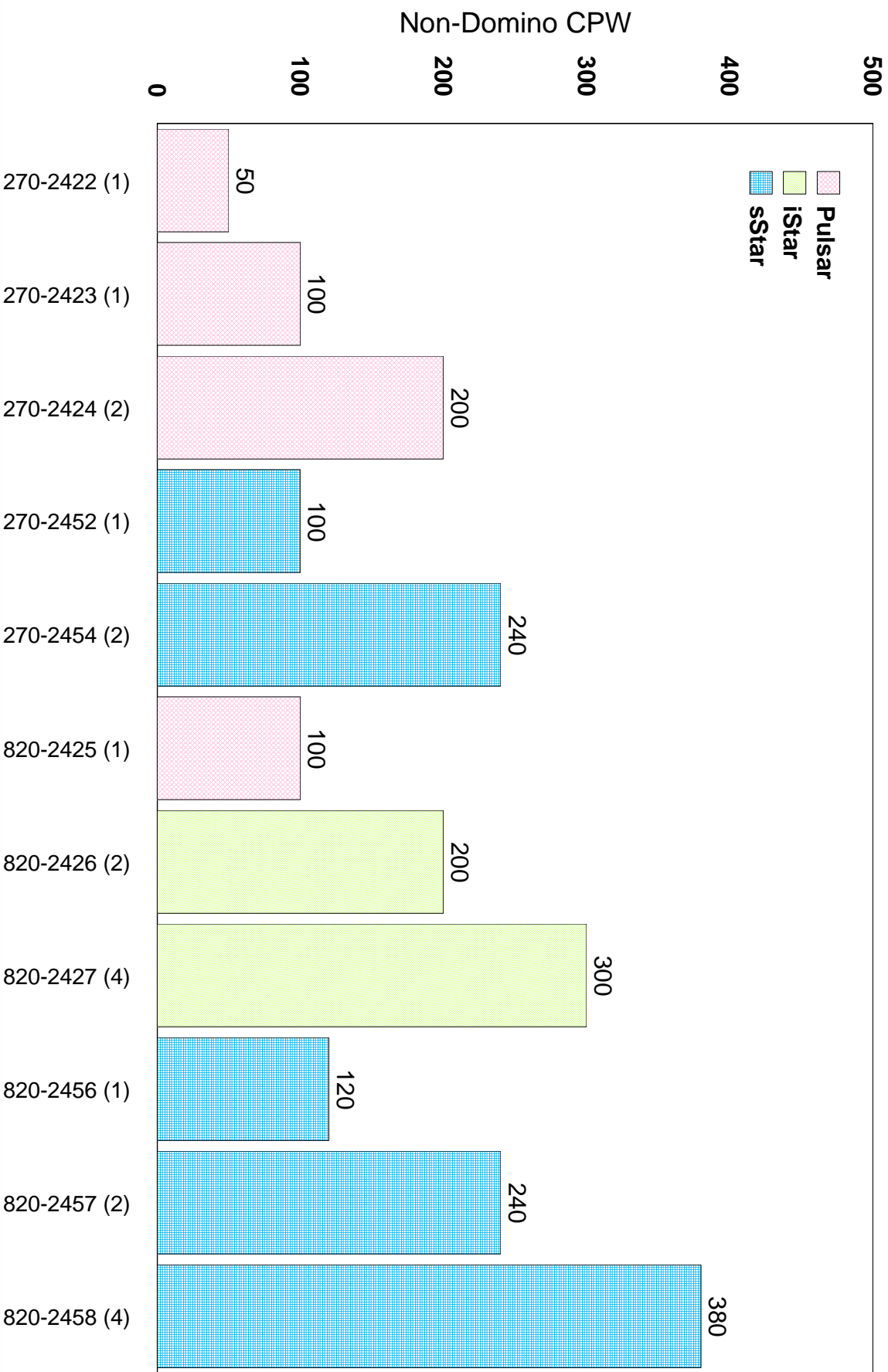
 **@server.** For the next generation of e-business.

V5R1 Domino-Dedicated Servers (Mail-Calendar) IBM @server iSeries



IBM @server. For the next generation of e-business.

V5R1 Domino-Dedicated Servers (non-Domino)



 **@server.** For the next generation of e-business.

Domino for AS/400 Performance - DSDs

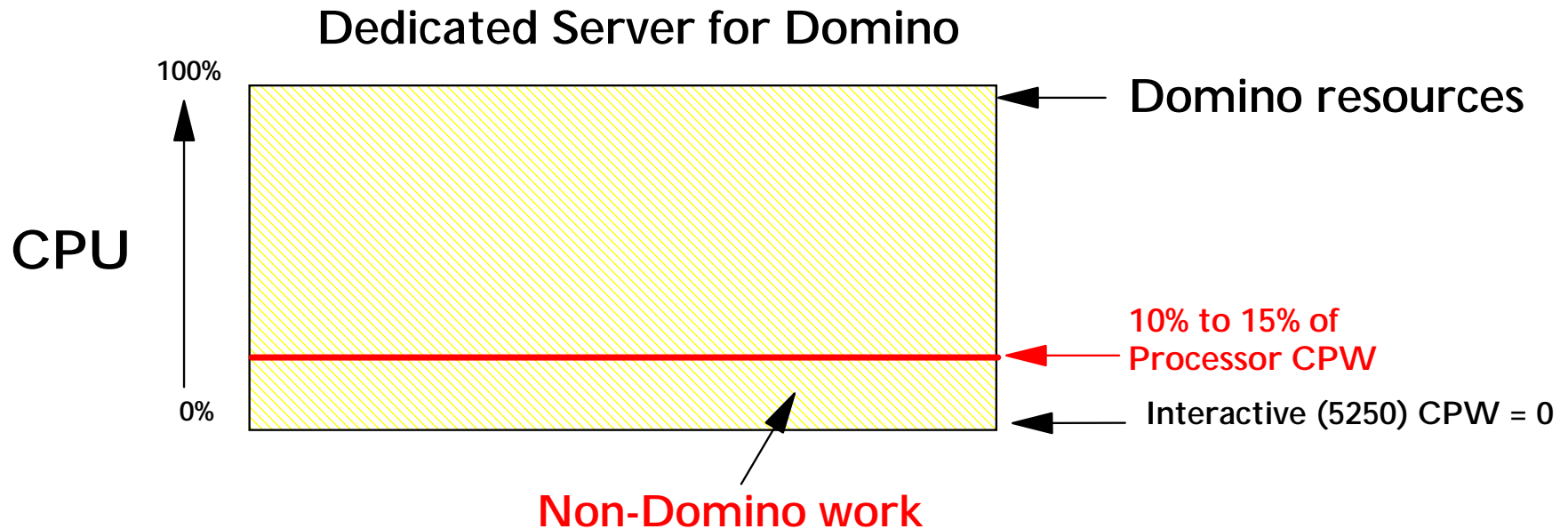
Dedicated Server for Domino

Model/Feature	Simple Mail	Mail/Calendar	Typical
820-2458 4-W	17720	18100	5910
820-2427 4-W	14840	9890	
820-2457 2-W	9990	6660	3330
820-2426 2-W	8420	5610	
820-2456	4670	3110	1560
820-2425	3860	2570	
270-2454 2-W	9990	6660	3330
270-2424 2-W	7580	5050	
270-2452	4610	3070	1540
270-2423	3860	2570	
270-2422	2400	1600	

Note: Results projected (70% CPU utilization) here should not be compared to audited NotesBench Results. Use these figures for comparing among iSeries, AS/400s only.

Managing work on a Dedicated Server for Domino IBM @server iSeries

CPU Resources Optimized 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.

IBM @server. For the next generation of e-business.

Notes: DSD Performance Considerations

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

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, and Java servlets) will be designated non-Domino workloads and will be managed to a maximum of 10-15% of the CPU.

Notes: DSD Performance Considerations

SYSTEM ADMINISTRATION -- iSeries and AS/400e Dedicated Server for Domino can be effectively managed when there is no Interactive CPW available for application workloads. System Administration activities performed in any of the following ways can leverage all of the capacity of the AS/400e Dedicated Server represented by the Processor CPW value.

- Any job started from a console (green screen) to perform system administration function is not considered interactive work if it is the ONLY interactive job running (single interactive job exception).
- Any system administration job submitted to batch is not considered interactive work.
- Any use of Operations Navigator (GUI administration functions) is not considered interactive work.

INTERACTIVE APPLICATION WORKLOADS (Interactive CPW=0) -- Any task started through a 5250 session (5250 device or 5250 emulation) where the user waits for a response is considered OS/400 interactive work. Interactive CPW=0 represents the maximum amount of 5250 processing capability available for customer applications.

- Application processing that uses 5250 interfaces has no available capacity on the Dedicated Server for Domino.
- 5250 print has no available capacity on the Dedicated Server for Domino.

Use of the Workload Estimator is recommended for sizing Domino workloads that have no OS/400 performance collection data available. If you have collected OS/400 performance data running the Domino application environment you can use BEST/1, part of the the Performance Tools/400 licensed program, 5769-PT1 or 5722-PT1, for capacity planning.

Important Software Considerations

New DSDs require OS/400 V5R1 and Lotus Domino for AS/400 server release 5.0.6a or later

Increased capacity for Domino applications using Java Servlets and WebSphere Application Server to be available September 28, 2001:

- Requires V5R1 at a specific level number (to be announced)
- Requires reinstall of V5R1 on systems already running V5R1
- New systems shipped after September 28, 2001 will include this support

DSDs and Java Workload Update

V4R5 and Domino Processors

- Non-Domino workloads should be managed to a maximum of 10-15% of the CPU (including DB2 Universal Database access)

V5R1+ and "old" Domino Processors

- 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, should be managed to a maximum of 10-15% of the CPU
- **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 "new" Domino Processors

- 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**

Domino for AS/400 Performance - 8xx, 270

IBM  server iSeries

Domino Mail on 2xx/8xx Models

Model/Feature	Simple Mail	Mail/Calendar	Typical User
840 - 2420 24-W	83420	55610	27810
840 - 2418 12-W	42650	28430	14220
830 - 2403 8-W	31370	20910	10460
830 - 2402 4-W	16020	10680	5340
830 - 2400 2-W	6060	4490	2250
820 - 2398 4-W	14840	9890	4950
820 - 2397 2-W	8420	5610	2810
820 - 2396	3860	2570	1290
820 - 2395	2480	1650	830
270 - 2434 2-W	9990	6660	3330
270 - 2253 2-W	6810	5050	2530
270 - 2252	3860	2570	1290
270 - 2250	2400	1600	800
270 - 2248	1220	810	410

Note: Results projected (70% CPU utilization) here should not be compared to audited NotesBench Results. Use these figures for comparing among iSeries and AS/400s only.

IBM  server. For the next generation of e-business.

Notes: Domino Mail on 2xx/8xx Models

This chart applies starts a showing of some "number of users" ratings for simple mail, mail and calendaring, and typical users for iSeries and older AS/400 models. These are **estimates** to be used only when comparing among AS/400 and iSeries systems. These values should be used only for early sizing estimates. In most cases the "Typical Workload" number of users is the proper balance among the work you think the users will do and the amount of work they actually do.

In general, take the Simple Mail users and divide by 3 to get "Typical Users." Alternatively you can divide Mail and Calendaring number of users by 2 to get "Typical Users."

The Workload Estimator includes these "number of users" sizing techniques.

Domino for AS/400 Performance - 7xx, 170

IBM  server iSeries

Domino Mail on 170/7xx Models

Model/Feature	Simple Mail	Mail/Calendar	Typical User
740 - 2070 12-W	26790	17860	8930
740 - 2069 8-W	18530	12350	6180
730 - 2068 8-W	16440	10960	5480
730 - 2067 4-W	10350	6900	3450
730 - 2066 2-W	5290	3530	1760
730 - 2065	2810	1870	940
720 - 2064 4-W	7380	4920	2460
720 - 2063 2-W	3850	2570	1280
720 - 2062	2010	1340	670
720 - 2061	1410	940	470
170 - 2388 2-W	4730	3150	1580
170 - 2385	2330	1550	780
170 - 2292	1410	940	470
170 - 2291	670	450	220
170 - 2290	420	280	140
170 - 2289	290	190	100

Note: Results projected (70% CPU utilization) here should not be compared to audited NotesBench Results. Use these figures for comparing among iSeries and AS/400s only.

IBM  server. For the next generation of e-business.

Domino for AS/400 Performance - 6xx, Snn

Domino Mail on 6xx, Snn Models

Traditional Model	Traditional Feature	Server Model	Server Feature	N-way	Estimated Mail Users
600	2129			1	100*
	2134			1	130*
	2135	S10	2118	1	200*
	2136		2119	1	350*
		170	2290	1	400*
			2291	1	600*
			2292	1	1300*
			2385	1	2300*
			2386	1	2300*
			2388	2	4300*
620	2179			1	370*
	2180	S20	2161	1	500*
	2181		2163	1	900*
	2182		2165	2	2000*
			2166	4	3400*
640	2237	S30	2257	1	1400*
	2238		2258	2	2400*
	2239		2259	4	4600*
650	2240		2240	8	8700*
	2188	S40	2207	8	16500*
	2243		2261	12	12500*
	2189		2208	12	23900*

* Results projected (70% CPU utilization) here should not be compared to audited NotesBench Results.
Use these figures for comparing among iSeries and AS/400s only.

 For the next generation of e-business.

Recommended Starting Model/Processor Configurations

IBM @server. For the next generation of e-business.

Notes: Recommended Model Configurations

The marketing material has some good "starting configuration" recommendations for various workloads. This information is reproduced here. Remember, these are "starting configurations," which means if you are proposing a smaller capacity configuration, you need to understand the "steady state" of the workload so as to not exceed the performance capacities of the configuration. Also, if you are going to have many peaks and valleys of resource consumption you need to consider a higher capacity configuration.

Here are the workloads used for the recommended configurations

- Application and Web Serving Workloads - describes compute (CPU) intensive processing, often requiring minimal database processing. There is no requirement for 5250 interface support typically used in many traditional transaction based applications.
- Transaction Workloads - this environment requires the traditional .iSeries balance of server resources ... a fairly even mix of CPU and I/O processing.
- Domino Workloads - Domino workloads consist of the following:
 - 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 local relational database access
- Mixed Workloads - this business environment requires an overall balance of server resources. A server truly capable of a mixture of applications simultaneously running at optimal performance...no matter the type of workloads.

The next foils shows the recommended configurations

Recommended Model/Processor Configurations

IBM  server iSeries

For Application and Web Serving Workloads

Model	270	820
Processors	2432, 2434 (#1520)	0150/0151/0152
N-Ways	1 - 2	1 - 4
Processor CPW range	1070 - 2350	1100 - 3700
Interactive CPW range	0	0

For Transaction Workloads

Model	270	820
Processors	2248	2395
N-Ways	1	1
Processor CPW range	150	370
Interactive CPW range	25	35 - 240

For Domino Workloads

Model	270	820
Processors	2452, 2454	2456, 2457, 2458
N-Ways	1 - 2	1 - 4
Mail, Calendaring Users	3070 - 6660	3110, 6660, 11810
Interactive CPW range	25	35 - 240

For Mixed Workloads

Model	270	820	830	840
Processors	2431, 2432, 2434	2435, 2436, 2437	2400, 2402, 2403	2418, 2461
N-Ways	1 - 2	1 - 4	2 - 8	12 - 24
Processor CPW range	465 - 2350	100 - 3700	1850 - 7350	10000 - 20200
Interactive CPW range	30 - 70	35 - 2000	70 - 4550	120 - 20200

IBM  server. For the next generation of e-business.

iSeries Models (V5R1)

IBM  server iSeries

Mdl	Feat	Tech	MHz	L2 (MB)	Thrds	CPW	n-way	M-C Users
270	2431	SStar	540	0	2	465	1	1490
	2432	SStar	540	2	2	1070	1	3070
	2434	SStar	600	4	4	2350	2	6660
	2452*	SStar	540	2	2	100*	1	3070
	2454*	SStar	600	4	4	240*	2	6660
	2248	Pulsar	400	0	1	150	1	810
@	2250	Pulsar	400	0	2	370	1	1600
@	2252	Pulsar	450	2	2	950	1	2570
@	2253	Pulsar	450	4	4	2000	2	5050
	2422*	Pulsar	400	0	1	50*	1	1600
	2423*	Pulsar	450	2	2	100*	1	2570
	2424*	Pulsar	450	4	4	200*	2	5050
820	0150	SStar	600	2	2	1100	1	3110
	0151	SStar	600	4	4	2350	2	6660
	0152	SStar	600	4	8	3700	4	11810
	2435	SStar	600	2	1	600	1	1620
	2436	SStar	600	2	2	1100	1	3110
	2437	SStar	600	4	4	2350	2	6660
	2438	SStar	600	4	8	3700	4	11810
	2456*	SStar	600	2	2	120*	1	3110
	2457*	SStar	600	4	4	240*	2	6660
	2458*	SStar	600	4	8	380*	4	11810

Mdl	Feat	Tech	MHz	L2 (MB)	Thrds	CPW	n-way	M-C Users
820	2395	Pulsar	400	0	1	370	1	1600
@	2396	Pulsar	450	2	2	950	1	2570
@	2397	IStar	500	4	4	2000	2	5610
@	2398	IStar	500	4	8	3200	4	9890
	2425*	Pulsar	450	2	2	100*	1	2570
	2426*	IStar	500	4	4	200*	2	5610
	2427*	IStar	500	4	8	300*	4	9890
830	2400	IStar	400	2	4	1850	2	4490
	2402	IStar	540	4	8	4200	4	10680
	2403	IStar	540	4	16	7350	8	22900
840	2461	SStar	600	16	48	20200	24	77800
	2352	SStar	600	16	16/24	9000	8(12)	
	2353	SStar	600	16	24/36	12000	12(18)	40500
	2354	SStar	600	16	36/48	16500	18(24)	
	2418	IStar	500	8	24	10000	12	28430
	2420	IStar	500	8	48	16500	24	55610
	2416	IStar	500	8	16/24	7800	8(12)	
	2417	IStar	500	8	24/36	10000	12(18)	28430
	2419	IStar	500	8	36/48	13200	18(24)	

***** Dedicated Server for Domino

@ Still orderable but recommend ordering SStar Models

Note: All Performance Estimates at 70% CPU Utilization



IBM  server. For the next generation of e-business.

Benchmark Update

IBM @server. For the next generation of e-business.

Last year...2000



#2 TPC-C 152,300/163,400 tpmC
(iSeries 400 #1 TPC-C Leadership from 7/10 through 11/27/00)



#1 NotesBench 75,000 mail users



#1 VolanoMark 108,000 messages/sec



#1 SPECjBB2000 spec 80,000 ops/sec

*Note: Benchmark rankings as of 09/07/00 & 11/27/00 from <http://www.ideasinternational.com>

IBM @server. For the next generation of e-business.

TPC-C Benchmarks

- No new TPC-C for new V5R1
- "Assume pSeries" - 24W, SStar



● About TPC-C

- ▶ Most reported midrange commercial benchmark in the marketplace
- ▶ Traditional OLTP Order Entry/Inventory scenario
- ▶ Heavy I/O and journaling requirements
- ▶ "Business Model Definition" is provided, not specific code
- ▶ SMP and Clustering configurations are allowed
- ▶ Reporting metric is number of transactions per second (tpm-c) and the configuration price per transaction (\$/tpm-c)

VolanoMarkJ Benchmark

IBM  server iSeries

First Java™ server-side Internet benchmark avail, Measures a "chat room" environment server <http://www.volano.com>



VolanoMark, a 100 percent pure Java server benchmark, has become increasingly important as e-business customers rapidly deploy Java environments as a key component of their server applications. The VolanoMark network scalability test measures how many client computers a server can support, while the local performance test determines how fast a Java Virtual Machine runs on a particular server.

Reporting metric is messages per second

Rank	Vendor/Server	Oper Sys	# of CPUs	Date	Messages/sec at 200 connect	Msgs/sec at 9,000 cont	Msgs/sec at 20,000 cont
1	IBM iSeries M840	OS V4R5	24w 12w	6/2000	108,153	- 39,529	- 23,942
2	IBM RS/6000 M80	AIX	8w	5/2000	46,370		
3	IBM RS/6000 S80	AIX	6w	1/2000	33,906		
4	Sun E6500	Solaris	22w	3/2000	25,131	4847	

Notes: Partial List of performance results for the 200 Connection VolanoMark LCC 2.1.2 benchmark

<http://www.as400.ibm.com/news/05jun2000.htm>

http://www.rs6000.ibm.com/resource/pressreleases/2000/May/ann_509.html on 5/11/00

<http://www.rs6000.ibm.com/resource/pressreleases/2000/Jan/volanomark.html>

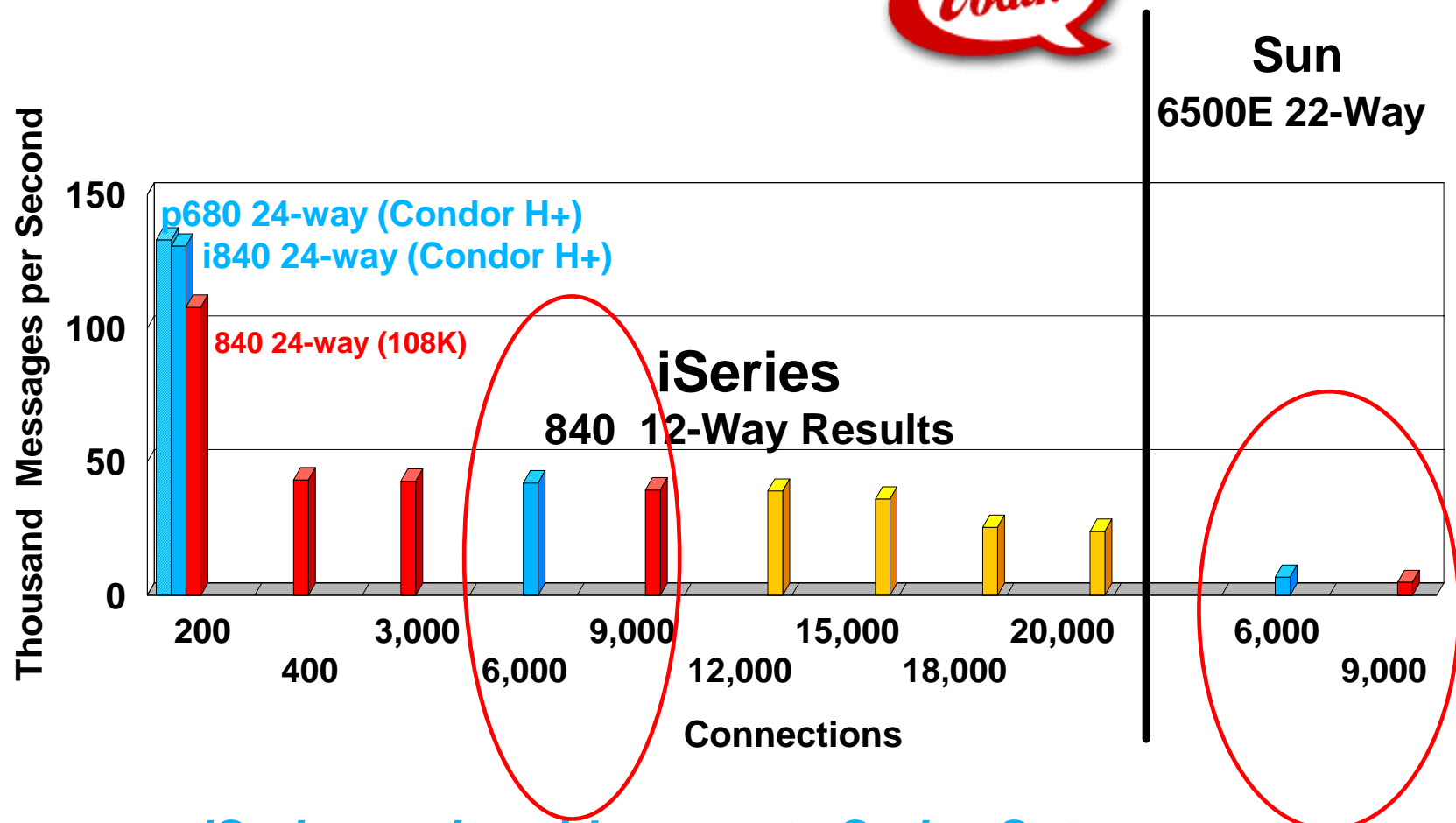
<http://www.sun.com/software/solaris/java/benchmarks.html> on 5/8/00

AS/400: Great Java Performance - secure, stable and reliable.

IBM  server. For the next generation of e-business.

VolanoMark Results

Projected VolanoMark Client/Server Benchmark



*iSeries parity with current pSeries Sstar
Should expect Public stmt at ANN*

Source: <http://www.as400.ibm.com/news/05jun2000.htm>, and <http://www.sun.com/software/solaris/java/benchmarks.html> on 7/17/00

IBM  server. For the next generation of e-business.

Notes: VolanoMark Results

Since the VolanoMark benchmark is primarily a message switching benchmark, do not expect as great a percentage of performance improvement with the 840-2061 as there would be with a CPU intensive benchmark which can get make use of processor speed and size of L2 cache increases.

SPECjBB2000 Benchmark

Java emulated 3-tier typical business application, Input (tier 1) to business logic (tier 2) call database (tier 3)

Totally self contained and self driving, no data base I/O

Loosely follows the TPC-C specification for its schema, input generation, and operation profile

SPECjbb2000 is memory resident, uses totally different data set sizes, mix of workloads, performs no I/O to disks, and has no think times.

Reporting metric is business operations are performed per second

Effectively exercise the implementation of the Java Virtual Machine (JVM), Just-in-time compiler (JIT), garbage collection, threads and some aspects of the operating system.

Benchmark measures the performance of CPUs, caches, memory hierarchy and the scalability of Shared Memory Processors (SMPs)

iSeries performed 2,4,8 and 24-way measurements - Check: <http://www.spec.org>

- iSeries delivered significantly better response time, with asynchronous garbage collection, and heap size

iSeries is #1, 2, 3 in SPECjBB2000

IBM  server. For the next generation of e-business.

SPECjbb2000 Results

iSeries Java™ Server-side leadership


- Measurement nearing completion
- Fundamental iSeries architectural advantages
- Expect public statement at V5R1 announcement

840-2461:
Approximately 60%
Improvement

iSeries and AS/400 is a leader in the industry in SPECjbb2000 benchmark

System	# CPUs	JAVA JDK Level	Trans/second	Max Trans Time (SEC)
IBM iSeries 840-2461 (600 Mhz)	24	1.3.0	132,322	0.18
IBM iSeries 840-2461 (600 Mhz)	24	1.2.2	130,791	0.20
IBM iSeries 840-2461 (600 Mhz)	24	1.1.8	131,479	0.15
Sun Fire 6800 (750 Mhz)	24	1.3.1	109,146	2.80
IBM iSeries 840-2420 (500 Mhz)	24	1.2.2	80,348	0.46
IBM iSeries 840-2420 (500 Mhz)	24	1.3.0	79,316	0.35
IBM iSeries 840-2420 (500MHz)	24	1.1.8	79,250	0.35
Sun Fire 6800 (750 Mhz)	12	1.3.1	62,463	1.77
IBM pSeries 680 (600MHz)	12	1.2.2	56,834	1.11
Sun Fire 6800 (750 Mhz)	8	1.3.1	43,353	1.17
HP N4000	8	1.2.2	40,192	1.66

<http://www.ideasinternational.com>, or <http://www.spec.org>
SPECJBBB200 as of April 25, 2001

IBM  **server. For the next generation of e-business.**

Trade 2 Benchmark

Trade 2 Benchmark - IBM Internal Server-side Java Benchmark

The Trade 2 benchmark simulates an online stock trading site using WebSphere Application Server Advanced Edition 3.02.

Trade 2 uses a web browser to interact with simulated user, trans/sec metric

The benchmark uses JSPs and servlets to interact with EJBs for database access

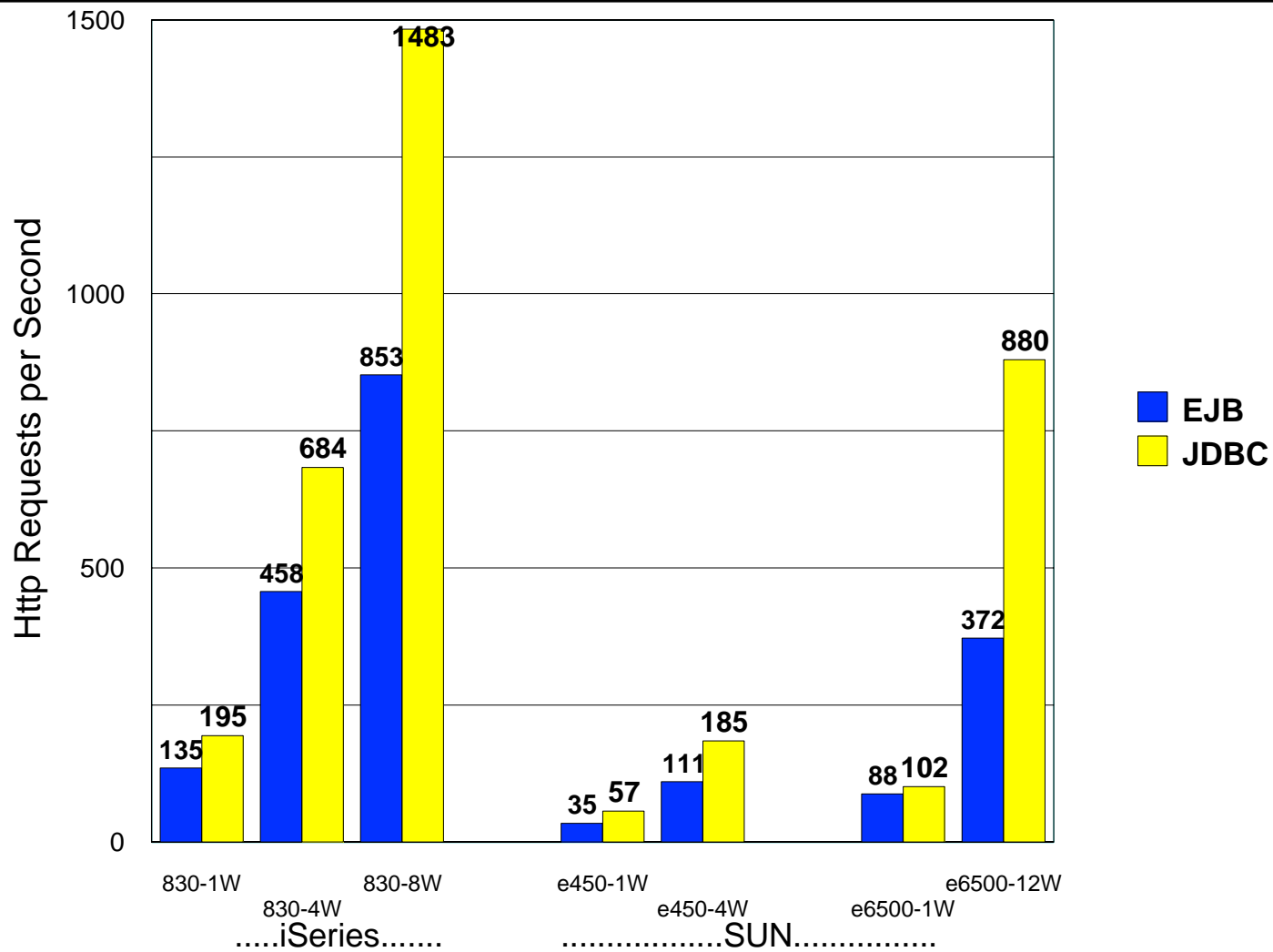
It contains transactions for the following:

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

Internal cross platform and development positioning

iSeries expecting continued leadership (internal use only reports)

Trade2 SMP Scaling



- iSeries 830/2403 540 MHz 8-way, 32 GB memory, w/WAS 3.5.3 Driver, IHS, JDK 1.2.2
 - Test Driver: iSeries-LoadRunner-EJB, 1w-15, 4w-25, 8w-100; JDBC, 1w-30, 4w-90, 8w-200, <1000K data transmitted.
 - Sun e450 300 MHz 4-way, 2 GB memory, w/Solaris 2.7, WASAdv 3.5, IHS, JDK 1.2.2
 - Sun e6500 400 MHz 12-way, 16 GB memory, w/Solaris 2.7, WAS Adv 3.5.2, IHS, JDK1.2.2
 - Test Driver: Sun-AKstress; 100 client load, 25,000 page hits, <1K data Transmitted per page hit

Data from IBM Laboratory measurements - 03/01

IBM  server. For the next generation of e-business.

System Performance TPC Benchmark Update

TPC-W Description and Action Plan

Newest Webserving Benchmark

- iSeries "should" do well
- Server-side Java
- Heavy Amazon.com scenario with database updates
- Few reported results (IBM Netfinity, Unisys) to date
- Work-in-progress

IBM server. For the next generation of e-business.

SAP

- SD 2-tier results published 4Q2000

JD Edwards

- Public Press Release (May 14, 2001)
 - 3 x 24-ways running WebSphere and JDE OW Xe
 - 6,000+ concurrent web clients doing real work, with a response time of 1.16 seconds
 - ▶ At 5850 users response time of 0.93 seconds
 - Previous best was Compaq/NT doing 4550 users many Intell systems

Linux Benchmarking

- Perf Considerations Tips & Techniques in Performance Capabilities Reference

SPECWeb99

- Under consideration

JD Edwards OneWorld Xe High Water Benchmark

Objective: Produce an industry leading benchmark of xxxx users running JD Edwards OneWorld Xe software (JAS Server) using WebSphere 3.5 on iSeries.

Approach: 3 Stage Plan

1. Functional/Scalability Test - 740 12W (WAS 3.02), target 500-1000 users
2. Scalability/Sizing Test - 830 4W + 740 12W (WAS 3.5), target 2000 users
3. Scalability Test - 3 x 840 24W + 840 12W (WAS 3.5, optionally V5R1), target xxxx users

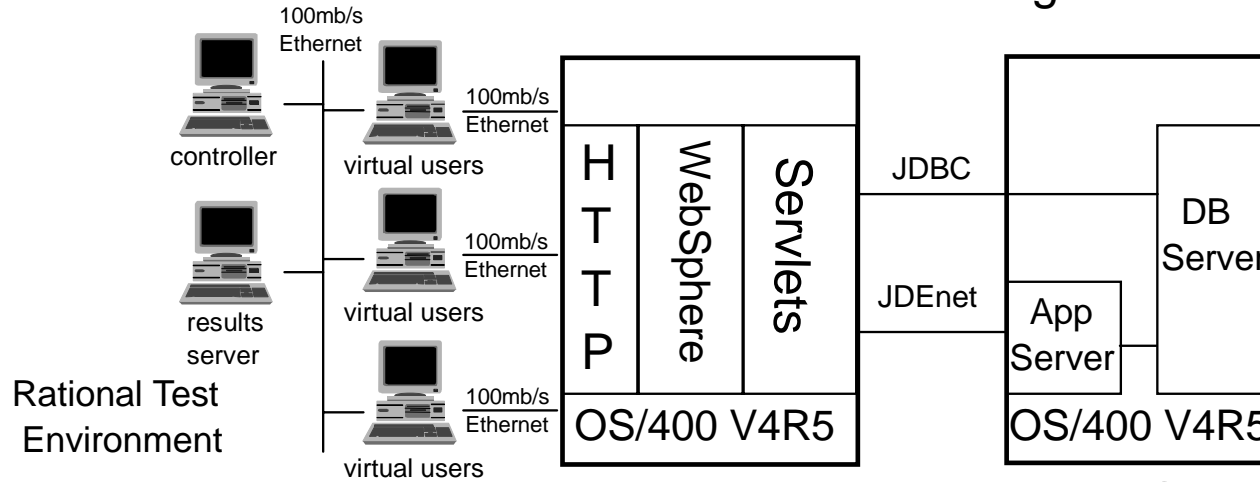
▶ **Results:** Market Statement

- ▶ demonstrating scalability of iSeries and WebSphere
- ▶ industry leading ERP application (JDE has mind/market share)
- ▶ showcase e-business solutions (E2E e-business environment WS and iSeries)
- ▶ target - results available for V5R1 GA (5/25) and JDE Focus (6/11-14)

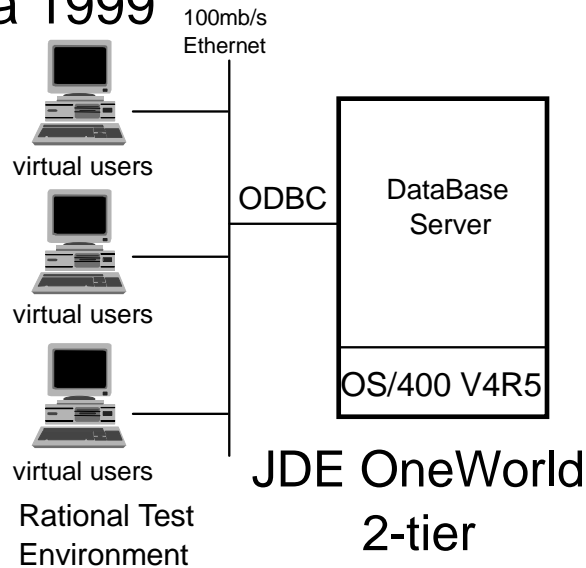
 **For the next generation of e-business.**

IBM High Water Testing

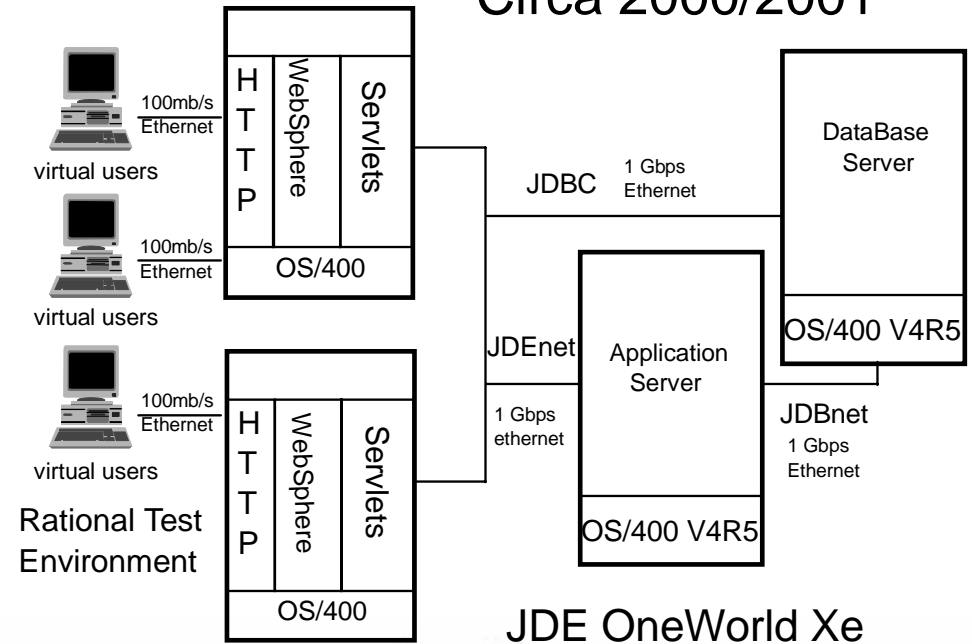
JDE OneWorld Xe Paradigm



Circa 1999



Circa 2000/2001



IBM  server. For the next generation of e-business.

TCP/IP-based Update

IBM @server. For the next generation of e-business.

TCP/IP-based Performance Update

General code path improvements in:

- FTP
- Telnet
- SMTP

HTTP Server for iSeries

- Triggered Cache Manager (TCM)
- Powered by Apache metrics in V5R1 Performance Capabilities Reference Manual

Java-based Update

IBM @server. For the next generation of e-business.

Moderate Performance Improvements: Toolbox for Java:

- Unicode <--> EBCDIC conversion
- CommandCall and ProgramCall within the iSeries JVM stay within thread
- IFS file list function

Moderate Performance Improvements: Developer Kit for Java (5722-JV1):

- Garbage Collection parallelism
- Code optimization
- JDBC and floating point processing

Notes: Java-based Update

Performance enhancements in V5R1 include:

- Improved garbage collection (GC) parallelism
- Improved code optimizations and heap allocations, including cross jar inlining
- JDBC and floating-point improvements; improvements of up to 30% have been seen in some workloads
- Several limits to growth have been removed. Most notably, the maximum object size has increased from 16 MB to 4 GB.

I/O Performance Update

IBM @server. For the next generation of e-business.

Rollback (Commitment Control) improved

Aggressive Journal Caching PRPQ (5799-BJC): reduces number of synchronous journal physical writes

- Lab tests: up to 20% shorter batch run time

New Journal option MINENTDTA: journal entry contains only updated columns in a row

- If few columns updated: less physical journal space: fewer writes, less "change journal receiver" situations

New RAID Disk Unit Controller

New PCI RAID disk controller: FC # 4778 or 2778

Up to 10% system performance improvement

- Biggest performance boost in a heavy write environment.

Replacement for 4748/2748 PCI RAID disk controller

- Effectively up to 4 x previous controller's write cache
 - Compression techniques provides up to 104 MB vs previous 26 MB
- Same read cache

Requires V5R1

For use in

- 270, 8xx (FC#4778)
- 6xx, Sxx, 7xx (FC#2778 in PCI I/O Tower FC# 5065/5066)

Fibre-Attached Tape Performance

** Preliminary Results **

	KB/Hour								
	NSRC.5GB	NSRC1GB	NUMX1GB	NUMX3GB	NUMX12GB	2GB	4GB	8GB	16GB
3590E SCSI									
SAVE	10000	14000	56000	71000	73000	100000	108000	110000	114000
RESTORE	7600	8800	35000	46000	46000	100000	111000	112000	115000
3590E FIBER									
SAVE	18000	18000	61000	78000	78000	119000	130000	130000	136000
RESTORE	8500	9000	37000	50000	50000	115000	125000	125000	132000
3580 SCSI									
SAVE	15000	15000	44000	61000	68000	77000	90000	99000	104000
RESTORE	7900	8500	31000	43000	43000	80000	92000	100000	104000

- Restricted state measurement
- If primarily DB data --> 2GB workload
- If primarily many small HTML pages --> NUMX1GM...--> NSRC2GB
- 3580 Fiber LTOgenerally equal to 3590E Fibre

IBM @server. For the next generation of e-business.

Notes: Tape performance

V5R1 supports Fibre Channel attached tapes. The devices supported are the 3590 E11/E1A with FC and the 3584 LTO with FC. We have measured the 3590 performance and completed the analysis for the V5R1 Newsletter. The numbers are shown below along with SCSI 3590 and SCSI LTO. They are expressed in KB/hr. (knock off the zeros and you have GB/hr, then divide by 3.6 and you have MB/s.)

The Fibre LTO numbers have not been completed as of publication of this presentation. Early test results indicate 14.2 MB/s with no compression and up to 45 MB/s with 3:1 compression. The numbers in the chart above were achieved with 2.7:1 compression.

We anticipate the LTO numbers slightly faster than 3590 at 3:1 compression. Early measurements indicated almost 43 MB/second with 3590 compared to the LTO at 45 MB/seconds.

Look for updates in the Performance Capabilities Reference manual.

Notes: Tape Workload Descriptions

The following work loads were designed to help evaluate the performance of save operations and restore operations. Familiarization with the makeup of the work loads will help clarify the differences in the save rates and restore rates that are presented below.

- **NUMX** - The old User Environment workload consists of 4 libraries. The first library contains 4 source files (for a total of 1204 members) that comprise about 39 MB of space. The second library consists of 28 database files, ranging in size from 2 MB to 200 MB, which total 470 MB in size. The third library consists of 200 program objects, with an average size of about 100 KB, for a total size of 20 MB. The fourth library is 12 MB in size and consists of 2156 objects of various types. The NUMX workload consists of about 556 MB of data.
- **NUMX6G** - The User mix data has been pulled into one library to create a larger sampling that can be used to evaluate concurrent and parallel save and restore operations on the newer high speed tape drives. All data was duplicated equally to create a balance between the old work load (NUMX) and the new (NUMX6G). The new workload sampling size is 6 GB or about 12 times the old sampling.
- **NSRC** - The NSRC workload consists of the 4 source files that are in the first library of the NUMX. These source files occupy about 39 MB of space and contain a total of 1204 members.

Notes: Tape Workload Descriptions-2

■ Large File

- 4GB, 8GB, 32GB - The 32GB workload is a single member database file that is about 32 GB in size. The old sampling size was 2 GB which was not a large enough sampling to evaluate parallel performance or the performance of the newer high speed tape drives. DLO The DLO workload consists of 8 folders with 3700 documents in the folders. The documents range in size from 53K to 233K with a combined size of 396MB. All of the documents reside in the first level folder structure. Integrated File System The following describes save and restore rates that a customer might see depending upon their data and its compaction capabilities. Take a system with an even mixture of client programs, such as, Lotus Notes databases and Web home pages. This example should save and restore in the range of the NUMX workload described in our charts. If the data stored on the system is largely made up of database files, the save/restore rates will probably tend toward the 2GB file type of workload, depending on the size and number of database files. If the data is largely made up of Web files, which tend to be numerous small HTML files such as small home pages, the save rates will tend downward from NUMX toward the NSRC workload. Web objects can be large images and client databases, just as Lotus Notes database files can be numerous empty or near empty mail files. This would reverse the description above. In all situations the actual data will dictate the save/restore rates and the customer will need to know the type of data they have on their systems in order to estimate the save or restore rates.

Disk Space

- Based on Customer Database Size
- Database Growth
- Temporary Space Requirements (by Application)

Number of Disk Arms

- Performance Consideration
 - Minimum Requirement - Refer "Disk Arm Consideration" - White Paper
 - Based on
 - ▶ Processor Capacity
 - ▶ Workload Characteristics

Tools

- Online Disk Arm Calculator
 - Adequate # of Arms for Processor
- Workload Estimator
 - Adequate # of Arms for specific workloads

As higher capacity disk (DASD) devices (8 or 18GB) for the iSeries 400 become available, fewer arms are needed to satisfy the capacity requirements of the more powerful 2xx/8xx models. This can lead to configuring too few disk arms (actuators) to meet the workload placed on them. A lack of disk arms can bottleneck the processor's performance. To avoid such a bottleneck, a minimum number of disk devices is needed for optimum performance on each AS/400 processor level. This number is independent of the quantity of drives needed to meet the desired storage capacity.

Online Disk Arm Calculator (ODAC) - We are about to make available a On-line Disk Arms Calculator Tool (ODAC). ODAC comes out in response of numerous field requests to automate disk arms estimation processes. The previously published V4R5 Disk Arm Consideration White Paper algorithms have been converted into a Lotus Script tool that will be available on the Internet. For the past two years the white paper has been a key document in helping selling additional DASD arms and avoiding negative (critsit) disk constraint installations and upgrades. The V4R5 Disk Arm Consideration White Paper found at:

<http://www-1.ibm.com/servers/eserver/series/perfmgmt/diskarm.htm>

ODAC should be available with V4R5 data in April and future release characteristics will be included as new products are announced. Watch the iSeries Performance Management web site for a link soon at:

<http://www.ibm.com/eserver/series/perfmgmt/>

Remember that the ODAC/Disk Arms White Paper answers a different question than the Workload Estimator and therefore both tools will generally provide different number of disk arms.

- ODAC answers the question of "how many disks are required to assure that a selected CPU will not be bottleneck" by insufficient disk arms being available.
- WLE answers the question of "how many disk arms are required for "good performance" for specific workloads" like Domino, Java, or WebSphere when the CPU has not already been selected. WLE will select both the CPU and the number of disk arms for "good performance". Somewhere in between the two is where your customer will probably be most happy and your personal experience will guide them.

New Performance Metric being developed - CIW

- CIW stands for " **Compute Intensive Workload**" ... this performance metric is representative of the new breed of workloads running on the iSeries. More and more, the type of applications that are being used on the iSeries are less dependent on overall balanced server resources ... measured and reported today as CPW (Commercial Processing Workload).
- We are introducing this new performance metric on the BASE Processor Features to establish a capability reference point for CPU-intensive workload processing ... as we move forward, this new metric will become the one we rely on to plan for growth and determine year-to-year performance improvements ... the Processor CPW metric will become less important over time.

CPW and CIW for 270, 8xx

Mdl	Feat	Tech	MHz	L2 (MB)	Threads	N-way	CPW	CIW	M-C Users
270	2248	Pulsar	400	0	1	1	150		810
	2431	SStar	540	0	2	1	465	185	1490
	2432	SStar	540	2	2	1	1070	380	3070
	2434	SStar	600	4	4	2	2350	840	6660
820	2395	Pulsar	400	0	1	1	370	180	1600
	2435	SStar	600	2	1	1	600	200	1620
	2436	SStar	600	2	2	1	1100	385	3110
	2437	SStar	600	4	4	2	2350	840	6660
	2438	SStar	600	4	8	4	3700	1670	11810
830	2400	IStar	400	2	4	2	1850	1630	4490
	2402	IStar	540	4	8	4	4200		10680
	2403	IStar	540	4	16	8	7350	3220	22900
840	2352	SStar	600	16	24	12	12000	5700	
840	2353	SStar	600	16	36	18	16500	8380	40500
840	2354	SStar	600	16	48	24	20200	10950	59600

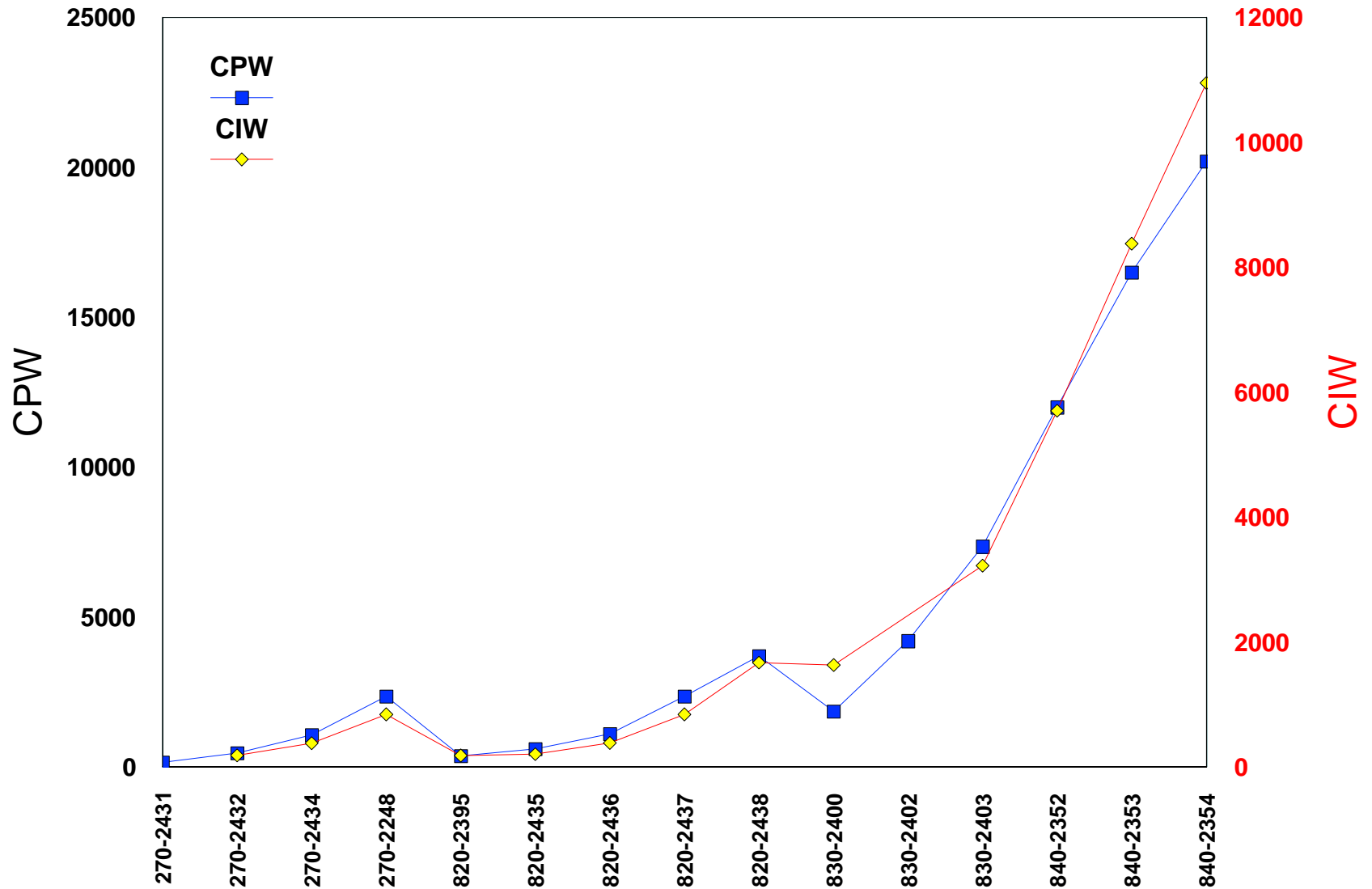
Note: Performance Estimates at 70% CPU Utilization



IBM  server. For the next generation of e-business.

CPW and CIW Graphic Comparison

IBM  server iSeries



 **e**server. For the next generation of e-business.

Notes: CPW and CIW Graphic Comparison

IBM  server iSeries

The Performance Capabilities reference manual introduces a new iSeries comparison metric - Compute Intensive Workload (CIW). This new metric will become more important in future Domino and e-business (Java/WebSphere) applications.

Though the CPW and CIW metrics are on different scales, this graph shows CPW and CIW track similarly for processors with similar MHz processors. The 830 area is different because the CIW workload takes advantage of the 540 MHz processor speed. Also, the top end belongs to the 600 MHz 840-2461 processor.

IBM  server. For the next generation of e-business.

System Performance Management Web Site

IBM @server iSeries

iSeries Performance Mgmt Web Page - new in 2001

<http://www.iseries.ibm.com/perfmgmt/>

Additional info on performance management or WLE is available via the public Internet.

The screenshot shows a Netscape browser window displaying the IBM @server Performance Management web page. The browser's address bar shows the URL <http://www.as400.ibm.com/perfmgmt/>. The page features the IBM logo at the top left, a navigation menu with links for Home, Products, Consulting, Industries, News, and About IBM, and a search bar. The main content area is titled "IBM @server Performance Management" and includes a "Shortcuts" section with a dropdown menu set to "Choose a topic". Below this, there is a section titled "Performance Management - What Is It?" which provides a brief overview of the capabilities and lists several key features: Real-time Monitoring, History/Trends of Performance/Capacity, Capacity Planning, Performance Tuning Knobs, Performance Problem Isolation, Workload Balancing, and Application "Design for Performance" Tools. To the right of this section is a "Performance Mgmt" sidebar with links for "What's New", "In the News", and "Resource Library". At the bottom of the page, a paragraph states: "To support these capabilities, products and tools, data collection facilities, and instrumentation are required. Size and complexity of customer computing environments - client/server, cooperative and distributed applications, distributed data, local and wide area networks, Web/intranet/internet, third-party".

WLE
→
ODAC

←
Perf White
Papers

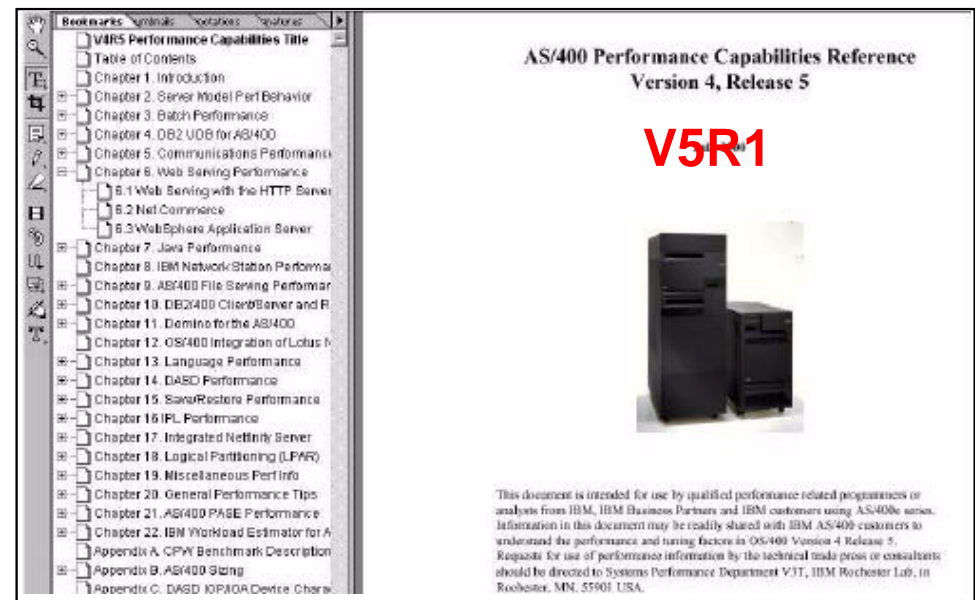
IBM @server. For the next generation of e-business.

V5R1 Perf Capabilities Reference

IBM  server iSeries

Performance Deliverables

- V5R1 Perf Capabilities Reference (Available 04/23/01, CA-Web and iSeries Internet)
- Publish tips/techniques, measurements complete, on Internet <http://publib.boulder.ibm.com/pubs/html/as400/online/chgfrm.htm> (What's New Topic)
- Updated information on Interactive Features, DSD, others
- Adobe Acrobat (.pdf) with Internet links
- (bold but tasteful use of color)
- White papers for special situations



IBM  server. For the next generation of e-business.

Performance Newsletter and Web Site

IBM  server iSeries

Performance Deliverables

- Performance Web, PID, Serv/Support Updated

CA-WEB/Perform site for IBMers

- <http://ca-web.rchland.ibm.com/perform/perfmenu.htm>
- PID and Serv/Support Internet site replication of key documents



Monthly Perf Newsletters, White Papers

- 650+ subscribers (IBMers Only) and growing
- Subscribe via mail to: rjodell@us.ibm.com
- Hundreds of papers and foil sets

AS/400 Performance Marketing Performance Team *Break-thru Performance Our Goal!*

<http://ca-web.rchland.ibm.com/perform/best1/vprpt.htm>

From: Dick Odell/Rochester/IBM @ IBMUS on 03/24/2000 05:20 PM
To: Performance-US, Performance-Non-US
Cc: Mike Tomasek/Rochester/IBM @ IBMUS, Theresa Backlund/Rochester/IBM @ IBMUS, Michael Maier/Somers/IBM @ IBMUS, Kevin Moe/Rochester/IBM @ IBMUS, Kathy Cebell/Rochester/IBM @ IBMUS, John Gressett/Rochester/IBM @ IBMUS, Mark Zagaria/Rochester/IBM @ IBMUS, Dean Ascherman/Rochester/IBM @ IBMUS, Hilary Melville/Rochester/IBM @ IBMUS, Sandee Johnson/Rochester/IBM @ IBMUS, Deb Smith/Rochester/IBM @ IBMUS, Craig Johnson/Rochester/IBM @ IBMUS, Maggie Blayney/Rochester/IBM @ IBMUS, John Quarantello/Dallas/IBM @ IBMUS, Kelly Schmitzer/Cleveland/IBM, Steve Ransom/Rochester/IBM @ IBMUS, Gary Brun/Rochester/IBM @ IBMUS, Kevin Gibson/Rochester/IBM @ IBMUS, Mahdad Majd/Rochester/IBM @ IBMUS, Patrick Keane/Rochester/IBM @ IBMUS, Ron Stevenson/Rochester/IBM @ IBMUS, Kent Milligan/Rochester/IBM @ IBMUS, Bob Gintow/Rochester/IBM @ IBMUS, DEPTHDL, DEPTH4, DEPTHX, Michael Cain/Rochester/IBM @ IBMUS, Jennifer Clarke/Atlanta/IBM @ IBMUS, Meg Thomson/UK/IBM @ IBMUS, Mark Olson/Rochester/IBM @ IBMUS, Sam Ellis/Rochester/IBM @ IBMUS, Doug Luebbe/Rochester/IBM @ IBMUS, Gregory Hoffa/Rochester/IBM @ IBMUS, Robert Elliott/Rochester/IBM @ IBMUS, Terry Reed/Rochester/IBM @ IBMUS, Jim Anderson/Rochester/IBM @ IBMUS, Mike Smith/Rochester/IBM @ IBMUS, Dave R Nelson/Rochester/IBM @ IBMUS
Subject: Performance Update AS/400 Performance Team - March '00

March '00 - Performance Update
Topic of the Month: VPN Performance Flash and Workload Estimator Webcast (4/11)

Performance Flash:
VPN Performance PTF - MF23279 offers improved VPN performance using new security protocols. Authentication performance using MD5 has been measured to improve by up to 42% and RC4 encryption is up to 64% faster than DES encryption. (01-20)
<http://ca-web.rchland.ibm.com/perform/best1/vprpt.htm>

Benchmarks:
COMPAG 227,079 tpm-c - Compaq ProLiant 8500-96P 12 nodes of 8-ways (total 96 processors) to reach 227,079.15 tpmC and 319.12 per tpm-c. Remember that cluster results of the TPC-C Benchmarks are easily scaled as only a minimal amount of transaction activity occurs between the nodes. So virtually any vendor willing to spend the money to assemble the huge configurations could achieve similar or higher transaction volumes. The 96 processors were Xeon 550 MHz and total config costs were \$4.3 million.
http://www.tpc.org/new_resultc-result1.idc?id=100021702

IBM  server. For the next generation of e-business.

Processor Technology

IBM @server. For the next generation of e-business.

Balanced System: Model 840

Model 840

- More off-chip cache
- Faster memory bus
- Extremely high speed data crossbar switch complex
 - with speeds up to 43.2 Gigabyte per second
- Faster connections from the I/O Hub structures to the I/O devices
 - 1.0 GB/sec connections (2x times 640, S40 or 740)
- HSL link to a Migration Tower
 - Speeds of approximately 350 MB per second
- Higher PCI Bus speed
 - up to 400 MB per second (10x faster than SPD bus on earlier systems)

Notes: Balanced System: Model 840

With the new processors in V5R1, we have much more off-chip cache, a faster memory bus, an extremely high speed data crossbar switch complex with speeds up to 43.2 Gigabyte per second for the new 840 Server. (The memory system bandwidth scales linearly with the frequency of the processor: the 36GB/s switch on the 500MHz processor is now 43.2GB/s). This switch is at the heart of delivering "data" to a processor fast enough so the high speed processor is not spending much time waiting for "data" to process.

In addition, we have faster DASD and much faster Input Output Processors and Adapters. The new Servers all have faster connections from the I/O Hub structures to the I/O devices. With the 840 each of the 1.0 GB/sec connections is 2 times the speed of the fastest predecessor (640, S40 or 740) design which uses the .5 GB/sec RIO technology interconnection access to I/O devices external to the primary system frame as far as 15 meters away. These 1.0 GB/second connections are up to 8 times faster than speed of the links to the I/O towers on the medium and low end models, using the 1,063 Mb/sec (132+ GB/sec) Optical Link Processor.

The HSL link to a Migration Tower (if there is one) is capable of speeds of approximately 350 MB per second.

In addition to the paths into and out of the high speed memory-processor-I/O interconnect switch, the PCI Bus speed can achieve up to 400 MB per second. This is 10 times faster than the 40 MB per second capacity of the SPD bus available on earlier systems.

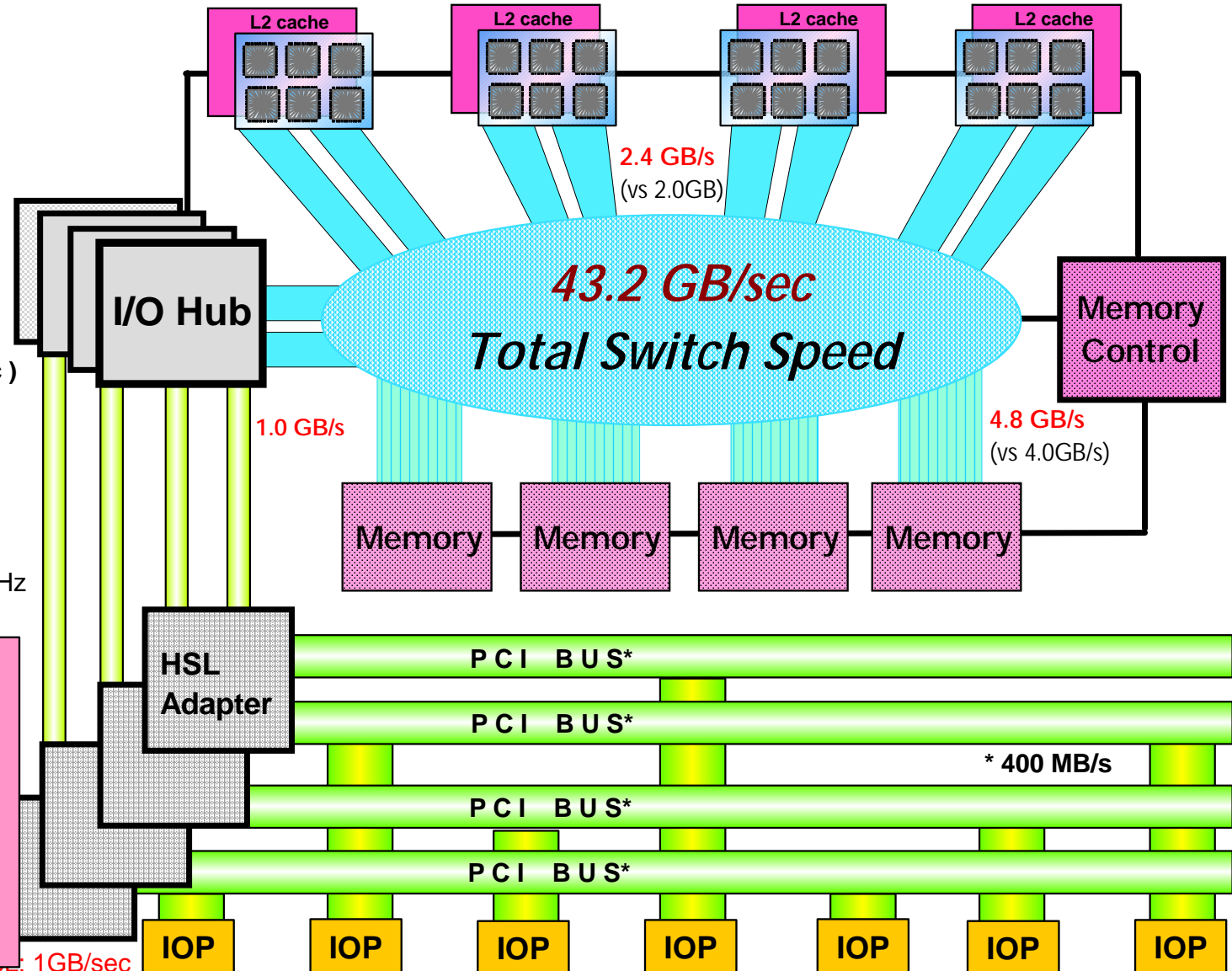
These improved speeds result in much greater capacity to satisfy the I/O demands of the added work that the new Server processors can accomplish.

To complete the balanced system concept, new and faster I/O hardware that connects to the PCI I/O buses of the new AS/400 Servers is also introduced with V5R1.

IBM  server. For the next generation of e-business.

Model 840 (600MHz vs 500MHz)

- ★ **24 x 600 MHz Processors**
(vs 500MHz)
- ★ **128 GB Main Memory**
(vs 96GB-128 GB with RPQ 847109)
- ★ **43.2 GB/s Crossbar Switch**
(vs 36GB/s)
- ★ **Processor Cache**
L1: 256 KB
L2: 16 MB
(vs 8MB)
- ★ **32 bytes x 600 MHz = 19.2 GB/sec**
(vs 32 bytes x 500 MHz = 16 GB/sec)
- ★ **14 buses**
- ★ **Superscaler: 4**
- ★ **Pipeline: 4**



4 I/O Hubs
I/O Pipe:
16Bytes x125MHz
=2.0 GB/sec

Transfer Rate
Memory to Processor is 4.8 GB/sec
One memory pipe
can feed two processor pipes
Memory Pipe:
64 Byte x 75 MHz (1/2 MHz) = 4.8
GB/sec
Processor Pipe:
16 Byte x 150 MHz = 2.4 GB/sec
Total Memory: 128 GB

1 GB/sec
4 per Hub

Notes: Model 840 (600MHz vs 500MHz)

In this chart the high-level overview of the structure of the 840 Server is shown. Central to this is the extremely high speed memory-processor-I/O interconnect ("crossbar") switch at 43.2 GB/sec. This allows for transfers of data between components connected to the switch to proceed at a very high rate. Separate connections to the switch are provided for each group of 6 processors, each of 4 banks of memory, and 2 connections for I/O traffic. These independent paths result in far superior performance than previous designs that have a single common shared memory bus path, which all components contended for. With the new design, there are many paths to the high speed switch which can operate separately. Its high bandwidth characteristics allow for connecting source and destination together quickly and efficiently to allow for significant benefit in overall processor speed and system performance enhancement.

One of the common problems in sophisticated advanced servers is how to read-write shared data. In a multiprocessor system such data can exist in the cache of only one processor to prevent multiprocessors from overwriting each other. To enforce this, the logical block, shown in this picture as memory control, implements a coherency protocol between the caches. This block is a logical representation of the functions present in the processors and the memory block using their knowledge of the content of all of the caches when main storage is accessed to determine if another processor has a more recently modified copy of the main storage location in its cache. If so, the data is moved to the requesting processor rather than fetched from main storage.

Notes:

- 128 GB of main storage (an RPQ was required in V4R5).
- Actual aggregate speeds realized in customer production mode may not achieve all the maximums described on this chart. However, actual throughput rates achieved will be significantly greater than on previous AS/400 systems, assuming there are no programming performance problems within the applications being used.
- Up to four I/O Hubs are supported. Each one can connect 4 High Speed Links. HSLs are connected to I/O devices Each one can connect 4 High Speed Links. HSLs are connected to I/O devices.
- The memory-processor-I/O switch architecture shown here is very similar to that available with the RS/6000 S85 model. The RS/6000 also has a maximum internal connect switch speed of 43.2 GB/second.

Notes: Model 840 (600MHz vs 500MHz)

- As stated, the high speed switch shown in the diagrams is designed to work as a common component for both the RS/6000 S80 and S85 and AS/400 840 model systems. The design accommodates a range of processor frequencies from the 450 MHz Copper-technology processors up to the 600 MHz Copper-SOI processors. For technical reasons, the switch cannot run at the identical speed for all processor frequencies in this range. Only selected "switch frequency settings" are usable relative to the particular MHz of the processors in the machine. The design is optimized so that excellent bandwidth and latency characteristics are provided for all AS/400 workloads running on the Copper-SOI processors with the switch design running at a base MHz rate. These design considerations are also carried over, then, to the RS/6000 implementation.
- From an AS/400 viewpoint the key is that the 43.2 GB/second speed is far beyond the capabilities within the 7xx servers and the 43.2 GB/second switch is delivered on the new 270 and 8xx servers. Competitors say their corresponding high speed switches are "on their way to the market place."

There is more logic behind the 'switch' and the 'memory control' than shown in this chart. For example:

- The AS/400 7xx and 8xx have super scalar processors, which enables multiple execution units per processor - more is better.
- A "pipeline" represents the number of stages a machine instruction takes to execute. The fewer stages, the better. Each execution unit is pipelined. Once the pipeline is full, then one instruction can be processed each machine cycle. The depth of the pipeline is important the more task switching or logic switching the processor must perform per unit of time. For example, with a task switch, the pipeline is flushed and needs to be refilled. The longer it takes to feed instructions through the pipeline to the processor, the longer the processor waits. Almost all commercial applications can be characterized as having very frequent task and logic switching. So the shorter the depth of the pipeline, the quicker this switching can be performed.
- If you want to learn more about architecture and technology used in the previous models go to <http://www.as400.ibm.com/beyondtech/>. **Reviewing the document : *V4R4 architecture, processor complex and performance*** will make it easier to compare the previous design with the technology used in the new for V5R1 AS/400 servers.

Comparison to non-IBM Systems

IBM @server. For the next generation of e-business.

Notes: Comparison to non-IBM Systems

The following pages are slightly updated comparisons to the non-IBM systems that were included in V4R5 presentation material.

With newer processors available from some of these vendors these comparisons are not as dramatic as during the year 2000. The charts are primarily included here for only as additional information in assisting in evaluating the iSeries technology portion of a "decision to buy." Certainly characteristics as mixed workloads and availability are at least as important as processor technology.

iSeries 400 model 840 vs Sun E10000

IBM  server iSeries

	Metrics	IBM PowerPC iSeries 840	Sun Enterprise 10000 *	Largest Server iSeries is
Processor	MHz	600	400	1.5X Faster
	# of Processors	24	64	2.6X Fewer
	SuperScaler	4	4	Same
	Pipeline	4	9	2.25X Better
Cache Per processor	Transfer Rate	19.2 GB/sec	3.2 GB/sec	6X Faster
	L1	256KB	32 KB	8X Bigger
	L2	16 MB	4 MB	4X Bigger
Memory To/From Switch	Size	128GB	64 GB	2X Bigger
	Transfer Rate	4.8 GB/sec	1.6 GB/sec	3X Faster
	Bandwidth	16 GB/sec	12.8 GB/sec	1.25X Faster
I/O To/From Switch	Transfer Rate	2 GB/sec	200 MB/sec	10X Faster
	Bandwidth	4 GB/sec	6.4 GB/sec	1.6X Slower
	Link Rate	1 GB/sec	100 MB/sec	10X Faster

* Previous to new TPC-C #1 rating model

Note: Server architectures are different. Figures are at rated speeds. Transfer rates are for one operation.

Source: Microprocessor Report 05/10/99 at <http://sun.com/servers/highend/10000/spec.html> on 5/12/00

**AS/400 wins
11 out of 13**

IBM  server. For the next generation of e-business.

iSeries 400 model 840 vs Compaq 8500

IBM  server iSeries

Largest Server

	Metrics	IBM PowerPC iSeries 840	Intel Pentium III Xeon Compaq 8500	iSeries is
Processor	MHz	600	800*	1.33X Slower
	# of Processors	24	8	3X Bigger
	SuperScaler	4	3	1.3X Bigger
	Pipeline	4	12/14	3X Better
Cache Per processor	Transfer Rate	19.2 GB/sec	6.4 GB/sec	3X Faster
	L1	256KB	32 KB	8X Bigger
	L2	16 MB	2 MB	8X Bigger
Memory To/From Switch	Size	128GB	16 GB	8X Bigger
	Transfer Rate	4.8 GB/sec	.8 GB/sec	6X Faster
To/From Switch I/O To/From I/O	Transfer Rate	2 GB/sec	.8 GB/sec	2.5X Faster
	Link Rate	1 GB/sec	100 MB/sec	10X Faster
Switch	# of Buses	14	5	2.8X More
	Switch Rate	43.2 GB/sec	4 GB/sec	10.8X Faster

* Estimated High End Compaq ProLiant Server specifications for 2H 2000. Figures are at rated speeds. Transfer rates are for one operation.

Source: Intel Product Brief: <http://support.intel.com/design/pentiumIII/xeon/prodbref/#4> on 3/10/00

<http://www.compaq.com/products/servers/ProLiant8500/overview.html> on 3/9/00

<http://www.intel.com/eBusiness/server/resources/pentiumiii/xeon/8way/Profusion.pdf>

**AS/400 wins
12 out of 13**

IBM  server. For the next generation of e-business.

iSeries 400 model 270 vs Compaq ML370

2-Way Server

Processor	iSeries Model 270	Intel Pentium III Compaq ML 370	iSeries wins 10 out of 12	
MHz	600	866	1.4X Slower	
# of Processors	2	2	Equal	
SuperScaler	4	3	1.3X Bigger	
Pipeline	4	12/14	3X Better	
Transfer Rate	9.6 GB/sec	6.9 GB/sec	1.39X Faster	
Cache Per processor	L1	256 KB	32 KB	8X Bigger
	L2	4 MB	256 KB	16X Bigger
	Transfer Rate	2.4 GB/sec	1.06 GB/sec	2.3X Faster
Memory	Size	16 GB	4 GB	4 X Bigger
I/O	Link Rate	1 GB/sec	100 MB/sec	10X Faster
Bus	# of Buses	2	1	2X More
Switch	Switch Rate	4.8 GB/sec	1.06 GB/sec	4.5X Faster

Figures are at rated speeds. Transfer rates are for one operation. Comparison made between 2-way low end servers.

Source: <http://www.compaq.com/products/servers/ProLiantml370/index.htm> 1/17/2001, Microprocessor Report 4/99

IBM  server. For the next generation of e-business.