

THE NEW POWER EQUATION

IBM i Performance Management - 6.1 Update



Dawn May
dmmay@us.ibm.com



Presentation created by Gottfried Schimunek, GottfriedSchimunek@us.ibm.com

THE NEW POWER EQUATION

Performance Disclaimer

- “it depends ...”
- Performance information and recommendations in this presentation is based on measurements, analysis, and projections in a controlled environment for specific performance workloads.
- Your results may vary significantly and are dependent on the application and configuration.
- This information is provided along with general recommendations for you to better understand system performance.
- Information is provided ***AS IS*** without warranty of any kind. Any “guideline values” that may be presented are good starting points but not hard rules to be applied for every application and workload.

Agenda

- POWER6 compared to POWER5
- Key 6.1 Performance areas
 - Java, WebSphere, DB2, encrypted ASP
- External Storage
 - IOPlless Fibre Channel SAN controller
- IBM i Client to VIOS Virtual I/O on the IBM JS22 BladeCenter
- Other performance topics

Performance Management Resource Library

<http://www-03.ibm.com/systems/i/advantages/perfmgmt/>

IBM Power Systems software > IBM i > Advantages >

Performance Management

on IBM i

Overview Resources Sizing

Resource Library

A wealth of reference information for the IBM i operating system: white papers, articles, the latest tool documentation, and more.

General Performance Resources

- IBM i V6.1 Performance Capabilities Reference - October 2008 (1.19MB)
- V5R4 Performance Capabilities Reference - January/April/July 2007 (1.63MB)
- V5R4 Performance Capabilities Reference - January/May/August 2006 (1.63MB)
- V5R3 Performance Capabilities Reference - April/July 2005 Edition (1.02MB)
- V5R3 Performance Capabilities Reference - May/July/Aug/Oct 2004 Edition (1.10MB)
- Sizing IBM i5/OS Work on IBM System i Partitions
- iSeries Performance Management Tools Redpaper
- iSeries Performance Overview - V5R3 (1.97MB)
- System i Performance Overview - V5R4 (2.00MB)

Performance Benchmark Reports

- Lawson M3 System Foundation 7.1 using IBM Technology for Java on System i (203KB)
- Benchmark Validation Test for IBM System i with JD Edwards EnterpriseOne (172KB)
- SPECjbb2000 #1 Benchmark Results
- Intentia One Box System i5 Benchmark (79.7KB)
- IBM eServer i5 570 NotesBench R6Mail Benchmark
- IBM eServer i5 570 NotesBench R6iNotes Benchmark
- eServer i5 Model 520 R6Mail Benchmark

Performance Papers and Articles

- Of GigaHertz and CPWs (293KB)
- Sizing System i5 and iSeries Disk with IBM Workload Estimator (334KB)
- Converting iSeries Disk Arm Requirements Based on Protection Type (58.1KB)
- LPAR Performance on Power Systems with POWER4, POWER5 and POWER6 (320KB)

To IBM Systems Workload Estimator tool

Notes

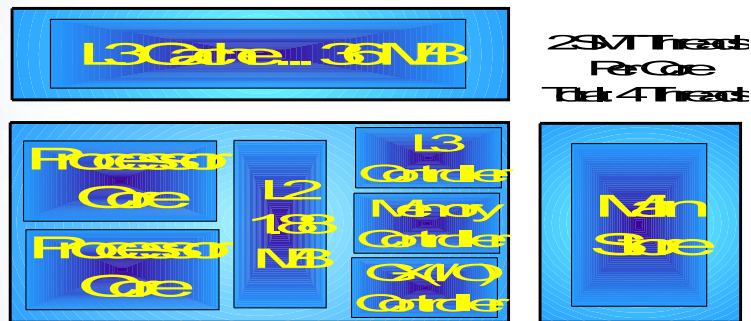
Before starting slides on the technical content of this presentation we first remind reviewers of sources for System i, Power Systems, and IBM i based performance information. We have selected a Resources tab from the first page found at

<http://www-03.ibm.com/systems/i/advantages/perfmgmt/>

Much of the performance information contained within this presentation is derived from information available at the performance management Web site and especially the October 2008 IBM i 6.1 Performance Capabilities Reference manual.

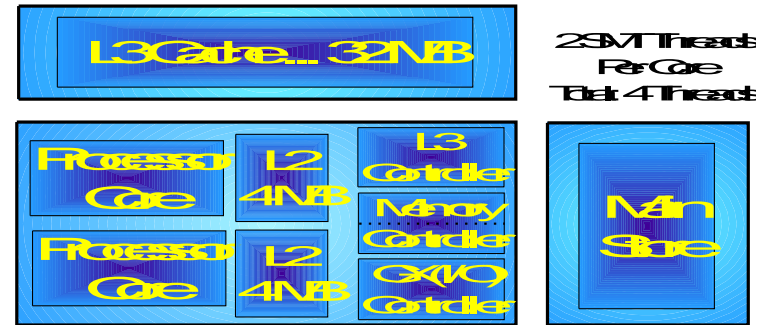
System i Performance: POWER 5 and POWER6

POWER5™



System i various: 1.5-2.3 GHz
Latest model 570: 2.2 GHz

POWER6™



Initial model 570: 4.7 Ghz
Power Systems: 3.5 – 5.0 Ghz

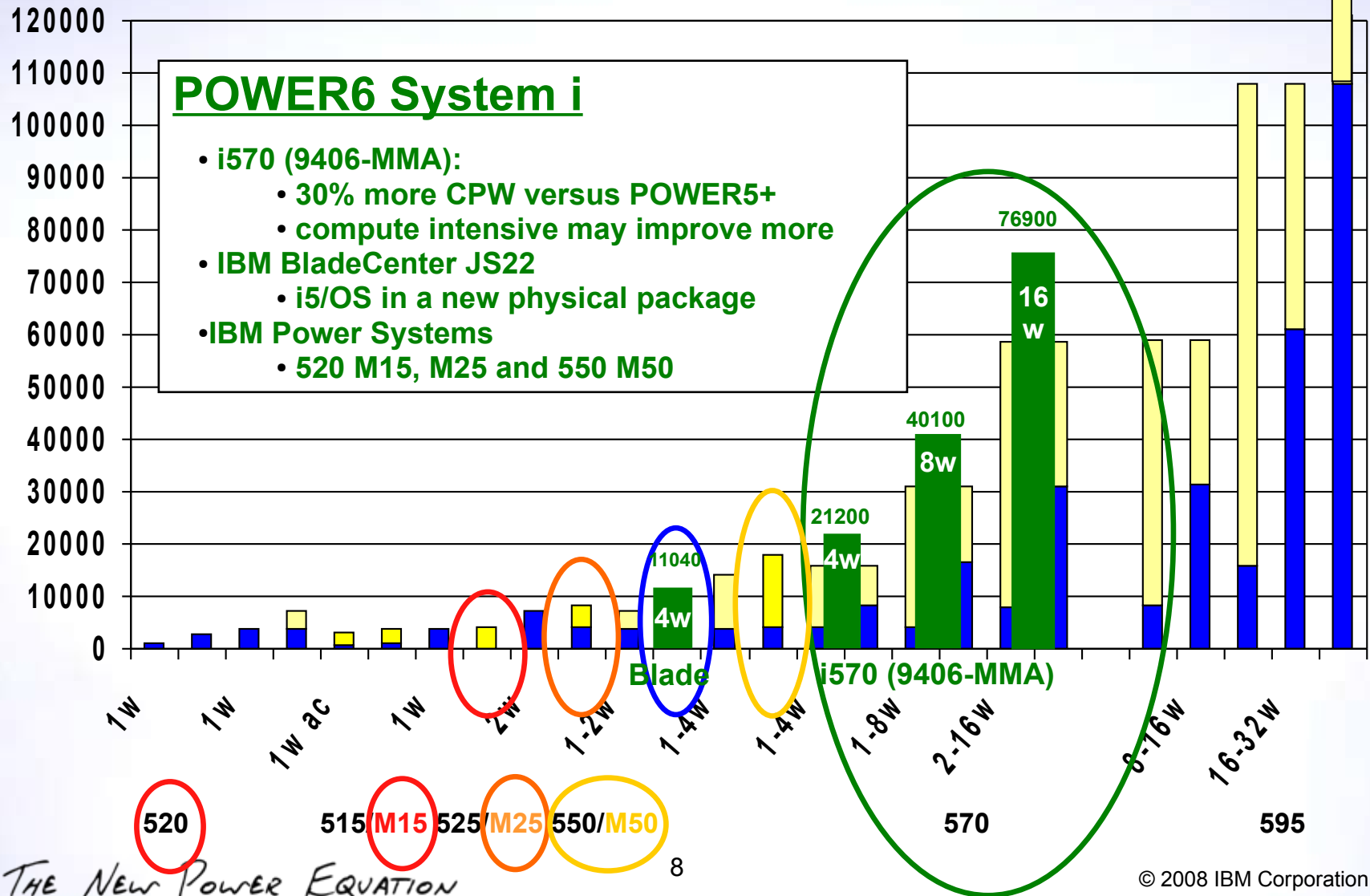
Notes:

In addition to faster processor clock speed the POWER6 technology includes a significantly enhanced L2 cache packaging.

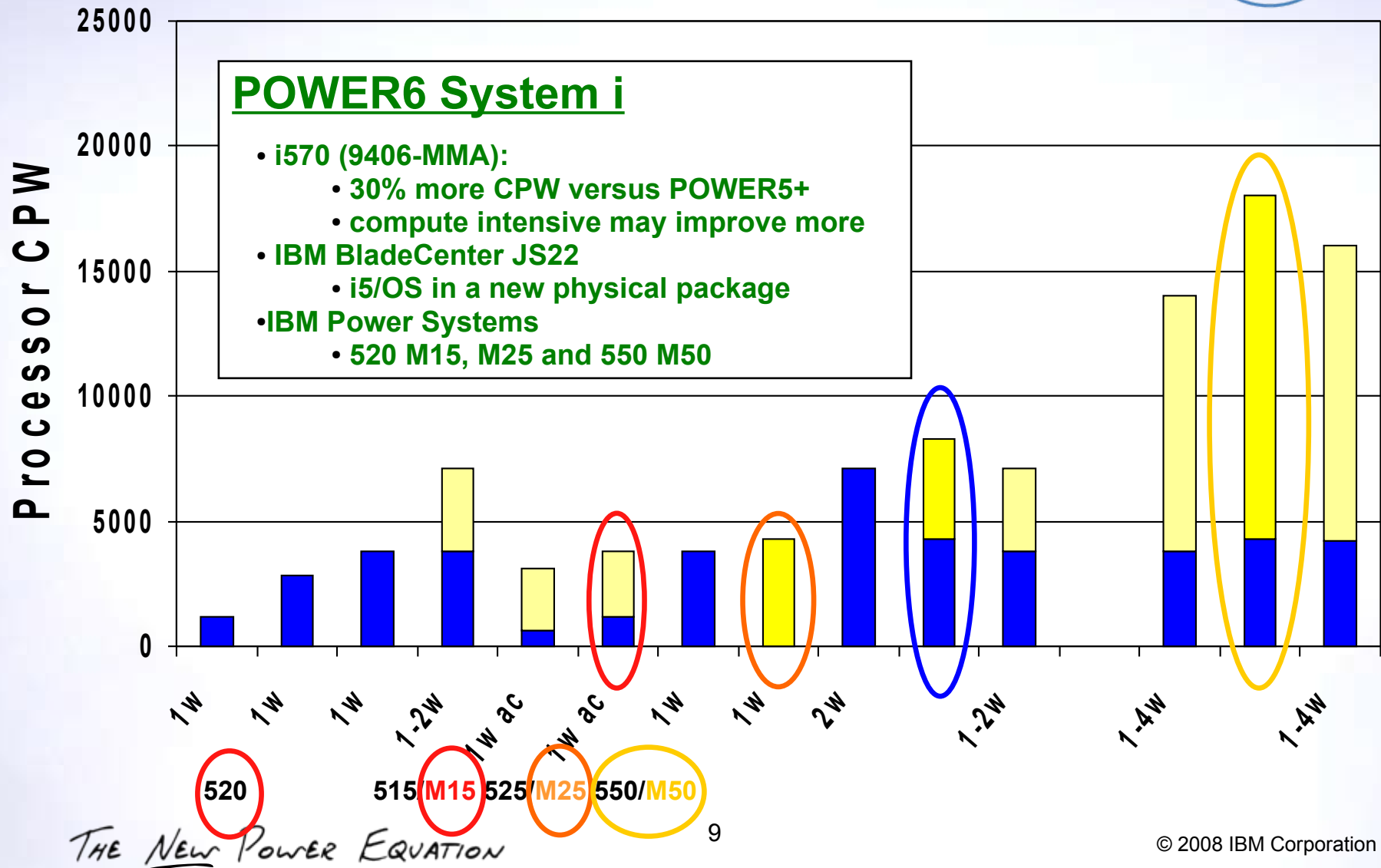


POWER5 and POWER6 System i

Processor CPW



POWER5 and POWER6 System i



Notes:

This slide superimposes the CPW ratings for the System i 570 technology POWER6 processor configurations, including the POWER6 BladeCenter on top of the fastest System i POWER5 technology systems. A later section will have additional performance information for 6.1 running on a BladeCenter, which requires some processor capacity assigned to a partition running Virtual I/O Server (VIOS).

Some notes:

- The blue portion of each bar represents the base number of processors CPW rating.
- The yellow/cram color portion of each bar represents the maximum CPU using the maximum number of processors that can be activated.
- The label 1 W ac represents the 520 model single processor rating and the maximum CPW rating with the processor accelerator feature activated.

System i570 (9406-MMA) Performance

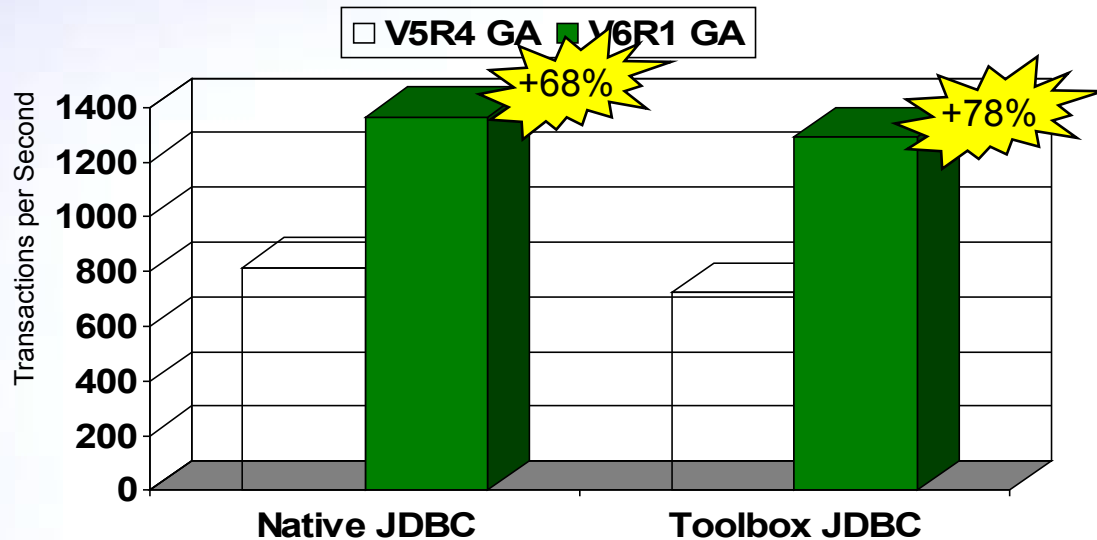
Workload	POWER5 1.65GHz	POWER5+ 2.2 GHz	POWER6 4.7 GHz
CPW 16-core	44,700 CPW	58,500 CPW +31% POWER5	76,900 CPW +72% POWER5 +31% POWER5+
Trade 6 JDBC 2- tier (WebSphere, 32 bit J9) 4-core	1,517 trans/sec	2,045 trans/sec +35% POWER5	2,980 trans/sec +96% POWER5 +46% POWER5+
SAP BI Datamart* 4-core	N/A	51,875	92,716 +79% POWER5+
MCU (Domino Mail and Calendar Users) 4-core	26,600 MCU	35,500 MCU +33% POWER5	47,500 MCU +79% POWER5 +34% POWER5+

Improvements shown above are valid only for these specific workloads -- Your results may vary and are dependent on the application, system model, and i5/OS release.

IBM System i 570, 2 processors / 4 cores / 8 threads, POWER6, 4.7 GHz, 128 KB L1 cache and 4 MB L2 cache per core, 32 MB L3 cache per processor, 64 GB main memory i5/OS V5R4M5 with DB2 for i5/OS. *SAP BI Data Mart standard NetWeaver 7.0 (2004s) two tier application benchmark performed June 1, 2007 in Rochester MN 92716 Query Navigation Steps The SAP certification number can be found at the following Web page: www.sap.com/benchmark.

THE NEW POWER EQUATION

IBM i 6.1 Performance Improvements Java™ and WebSphere®



Test Environment

- IBM Technology for Java VM – 32 bit
- WebSphere Application Server 6.1
- DB2 for i
- 2-tier environment
- POWER5+ 2.2Ghz 2-core partition
- Trade 6.1 (stock trading) workload

- **Improvements primarily in JDBC, DB2® for i, and Java**
 - Includes 64K pages (only available on P5+ and beyond)
- **Some improvements available via V5R4 PTFs**
 - But majority of improvements available only in 6.1

Improvements shown above are valid only for these specific workloads -- Your results may vary and are dependent on the application.

THE NEW POWER EQUATION

Notes:

The internal IBM Java/WebSphere “stock market” benchmark Trade6.1 was run using i5/OS V5R4 IBM 32-bit IBM Technology for Java and then IBM i 6.1 IBM 32-bit IBM Technology for Java. Each i5/OS release benchmark was run twice, differing in database access using the Native JDBC and then the Toolbox for Java JDBC database interfaces.

The test environment was:

- WebSphere Application Server 6.1
- DB2 for i5/OS
- 2-tier environment
- POWER5+ 2.2Ghz 2-core partition
- Trade 6.1 (stock trading) workload

This internal IBM Java/WebSphere TRADE6.1 benchmark running 32-bit IBM Technology for Java has over 60% performance improvement on 6.1 over V5R4. Approximately 35% of the overall improvement can be available in V5R4 PTFs, 65% available only in 6.1. Primary contributors of the improved performance were in the areas of 6.1 Java, DB2 for IBM i and WebSphere Application Server V6.1.

Some of the improvements to JDBC and database would also apply to the Classic JVM, but not to the same degree.

Java and IBM i 6.1 performance

- No Direct Execution optimization under IBM i. All Java runs JIT
- Tune Java programs according to recommendations
- i5/OS V5R4:
 - 64-bit Classic JVM (default JVM is Classic 1.4)
 - Classic 1.3, 5.0, and 6.0
 - 32-bit IBM Technology for Java JVM (less main storage requirements)
- IBM i 6.1:
 - 64-bit Classic JVM supports Java versions 1.4, 5.0, and 6.0
 - 32-bit IBM Technology for Java 5.0 (JDK V1.5) (default JVM in 6.1)
 - 64-bit IBM Technology for Java 6.0 (JDK V1.6), with March 2008 PTFs
 - 64-bit Classic JVM supports Java versions 1.4, 5.0, and 6.0
- Comparing i5/OS V5R4 and 6.1 Java performance, in general:
 - Classic 64-bit JVM will have same performance
 - 32-bit IBM Technology for Java 5.0 (JDK V1.5): approximately 10% improvement with 6.1
 - Greater improvement if the -Xlp64k flag is set to enable 64K pages
 - Moving from 32-bit IBM Technology for Java 5.0 (JDK V1.5) to 32-bit IBM Technology for Java 6.0 (JDK V1.6) should demonstrate "substantial performance improvements"
 - 32-bit IBM Technology for Java 6.0, 64-bit IBM Technology for Java 6.0 performance (see notes and Performance Capabilities Reference manual)

Notes:

The Trade 6.1 performance results slide shown earlier, specifically mentions 64k pages as one contributor to improved Java performance. Improved memory management running under i5/OS PASE (under IBM Technology for Java 32-bit and now 64-bit JVM environments) also improve performance for Java applications, particularly in cases where the heap size is larger than 1 GB.

i5/OS reaps the benefits of moving to IBM Technology for Java by taking advantage of the research and development IBM-wide on improving JVM performance across all of IBM's platforms. This enables i5/OS to focus on other performance areas unique to i5/OS.

Some specific performance improvements or recommended "good practices" for improving Java performance follow.

If your application also runs on IBM JVMs on other platforms, such as AIX, then you might consider trying the Garbage Collection (GC) parameters from those platforms as a starting point when using IBM Technology for Java on i5/OS. The IBM Technology for Java JVMs on all IBM platforms use the same tuning parameters. You may need to make i5/OS adjustments but the tuning should be easier than when running under the 64-bit Classic JVM under i5/OS.

If you are testing a new application, or aren't certain about the performance characteristics of an existing application running in the Classic 64-bit VM, start by running the application with the default heap size parameters (currently an initial heap size of 4 MB and a maximum of 2 GB). Run the application and see how large the heap grows under a typical peak load. The maximum heap size can be set to this value (or perhaps slightly larger). Then the initial heap size can be increased to improve performance. The optimal value will depend on the application and several other factors, but setting the initial heap size to about 25% of the maximum heap size often provides reasonable performance.

Notes:

Keep in mind that the maximum heap size for the 32-bit VM is 3328 MB. Attempting to use a larger value for the initial or maximum heap size will result in an error. The maximum heap size is reduced when using IBM Technology for Java's "Shared Classes" feature or when files are mapped into memory (via the java.nio APIs).

The maximum heap size can also be impacted when running large numbers of threads, or by the use of native code running in i5/OS PASE, since the memory used by this native code must share the same 32-bit address space as the Java JVM. As a result, many applications will have a practical limit of 3 GB (3072 MB) or even less. Applications with larger heap requirements may need to use one of the 64-bit JVMs (either IBM Technology for Java or the Classic JVM).

When heap requirements are not a factor, the 64-bit version of IBM Technology for Java will tend to be slightly slower (on the order of 10%) than 32-bit with a somewhat larger (on the order of 70%) memory footprint. Thus, the 32-bit VM should be preferred for applications where the maximum heap size limitation is not an issue.

See the Java Performance chapter in the October 2008 IBM i 6.1 Performance Capabilities Reference manual for more detailed coverage.

Leading Java™ Performance with IBM Power Systems

- IBM Power Systems is #1, #2, #3, #4, and #5 for SPECjbb®2005 *

Rank	Vendor	System	GHz	Cores	JVMs	Operating System	Result (BOPS)	Result (BOPS/JVM)	Date
1	IBM	System p 570	4.7	8	4	AIX 5.3	346,742	86,686	Jun 07
2	IBM	System p 570	4.7	8	4	Red Hat Enterprise Linux Advanced Platform 5.1	335,424	83,856	Nov 07
3	IBM	System p 550	4.2	8	5	AIX 6.1	333,779	83,445	Jan 08
4	IBM	System p 550	4.2	8	4	Red Hat Enterprise Linux Advanced Platform 5.1	328,343	82,086	Jan 08
5	IBM	System i 570	4.7	8	4	i5/OS V6R1	311,505	77,876	Jan 08
6	Dell	PowerEdge 2950 III	3.16	8	4	Microsoft Windows Server 2003 Enterprise x64 Edition	303,130	75,783	Nov 07

* Comparison is based on best-performing 8 processor core systems as of 1/29/08. IBM Power570 SPECjbb2005 result was submitted for publication by SPEC 1Q 2008. SPEC® and the benchmark name SPECjbb®2005 are registered trademarks of the Standard Performance Evaluation Corp. (SPEC). Results from <http://www.spec.org/jbb2005/results/>.

Notes:

This slide shows as of January 29, 2008, the Standard Performance Evaluation Corp. (SPEC) SPECjbb2005 benchmark ratings. The top 5 ranking positions are held by IBM POWER6 System p and System i configurations running AIX 5.3, AIX 6.1, IBM i 6.1, and Red Hat Enterprise Linux Advanced Platform 5.1 operating system levels.

The IBM System p 550 models are the 8-Way POWER6 models announced January 29, 2008.

The remainder of the text, describing the SPECjbb2005 Java benchmark in these notes pages is based upon information available at the SPECjbb2005 Web site listed on this slide:

<http://www.spec.org/jbb2005/results>

The SPECjbb2005 benchmark was introduced June 2005. It is an update to the previous SPECjbb2000 benchmark - for evaluating the performance of servers running typical Java business applications. The benchmark can be used across several versions of UNIX, Windows, Linux and other operating systems.

SPECjbb2005 was developed by SPEC's Java subcommittee, with contributions from BEA, Darmstadt University of Technology, HP, IBM, Intel and Sun. The benchmark's workload represents an order processing application for a wholesale supplier. Systems integrators and end users can use the benchmark to evaluate performance of hardware and software aspects of Java Virtual Machine (JVM) servers. Hardware vendors can use SPECjbb2005 to analyze their platform's scalability when running Java applications. Software vendors can use it to evaluate the efficiency of their JVMs, just-in-time compilers (JITs), garbage collectors, thread implementations, and operating systems.

It also measures the performance of CPUs, caches, memory hierarchy and the scalability of shared memory processors (SMPs). SPECjbb2005 provides an enhanced workload, implemented in a more object-oriented manner to reflect how real-world applications are designed and introduces new features such as XML processing and BigDecimal computations to make the benchmark a more realistic reflection of today's applications.

DB2 for i:

6.1 Performance Enhancements

- SQE now supports queries with the following attributes:
 - NLSS/CCSID translation between columns
 - User-defined table functions
 - Sort sequence
 - Lateral correlation
 - UPPER/LOWER functions
 - UTF8/16 Normalization support (NORMALIZE_DATA INI option of *YES)
 - LIKE with UTF8/UTF16 data
 - Character based substring and length for UTF8/UTF16 data
- 6.1 default value for QAQQINI file option IGNORE_DERIVED_INDEX changed to *YES
- More queries will run under SQE even if there is a select/omit logical file index created over any of the tables processed by the query
- Faster Full Open Optimization
 - 45% time reduction on average for a mix of complex queries
 - 30% time reduction to retrieve an existing plan from the SQE plan cache

Notes:

SQL queries with the attributes listed in the slide will be processed by the SQE optimizer and engine in 6.1.

Due to the SQE optimizer potentially choosing a better plan along with the more efficient query engine processing, there is the potential for better performance with these queries than was experienced in releases previous to 6.1.

SQL queries which continue to be routed to CQE in IBM i 6.1 have the following attributes:

- INSERT WITH VALUES statement or the target of an INSERT with subselect statement
- Logical files referenced in the FROM clause
- Tables with Read Triggers
- Read-only queries with more than 1000 dataspace or updateable queries with more than 256 dataspace
- DB2 Multisystem tables

Note, support of decimal float data type for a column introduces some “performance considerations.” In general, more CPU is used to process data of this type versus decimal or floating-point data. Given the additional processing time needed for decimal floating-point data, the recommendation is to use this data type only when the increased precision and rounding capabilities are needed. We also recommend avoiding conversions to and from this data type, whenever possible. In most application environments it should not be necessary to migrate existing packed or zoned decimal fields within a mature data base file to the new decimal floating point data type.

See the October 2008 Performance Capabilities Reference manual for more details.

With generally improved DB2 for i and SQL performance improvements in 6.1, it is anticipated JDBC-based and ODBC-base performance should be noticeably improved.

Notes:

Note the following:

There are two JDBC drivers for IBM i:

The Native JDBC driver is a type 2 driver. It uses the SQL Call Level Interface (CLI) for database access and is bundled in the System i Developer Kit for Java.

The JDBC Toolbox driver is a type 4 driver which is bundled in the System i Toolbox for Java.

In general, the Native driver is chosen when running directly under the i partition. The Toolbox driver is typically chosen when accessing data on the on an i partition from another partition or system. The Toolbox driver is typically used when accessing i data from a Windows machine, but it could be used when accessing the i database from any Java capable software.

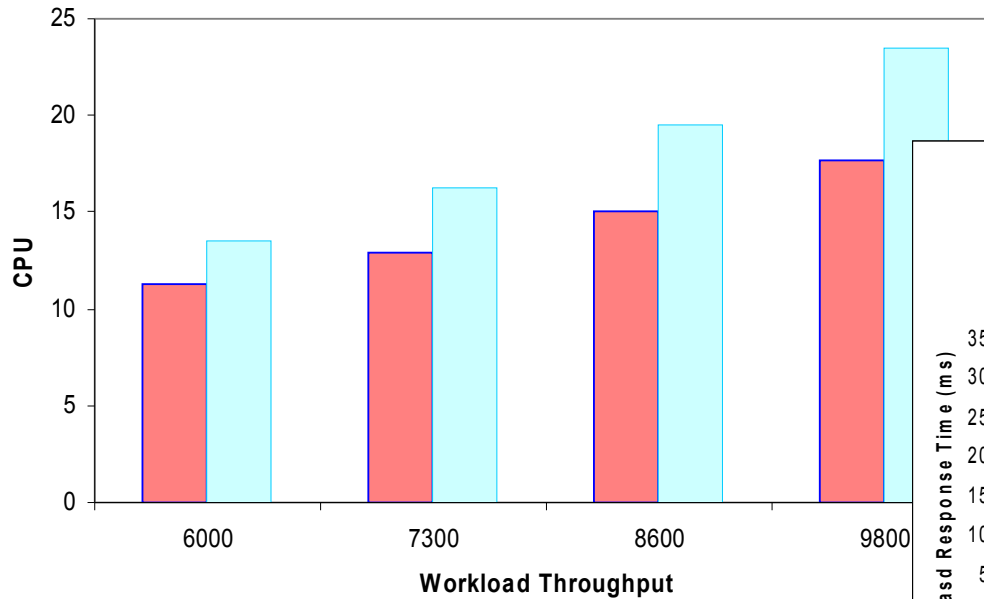
JDBC performance depends on many factors ranging from generic best programming practices for databases to specific tuning which optimizes JDBC API performance. Tips for both SQL programming and JDBC tuning techniques to improve performance are included in the Performance Capabilities Reference manual. Assuming your applications is implementing the recommended techniques just by moving to IBM i 6.1, your database access performance should improve noticeably.

Encrypted ASP Performance

CPU utilization and transaction throughput

Non Encrypted ASP vs Encrypted ASP

9406 MMA 4 Way 571F with 24 DASD Non Encrypted ASP
 9406 MMA 4 Way 571F with 24 DASD Encrypted ASP

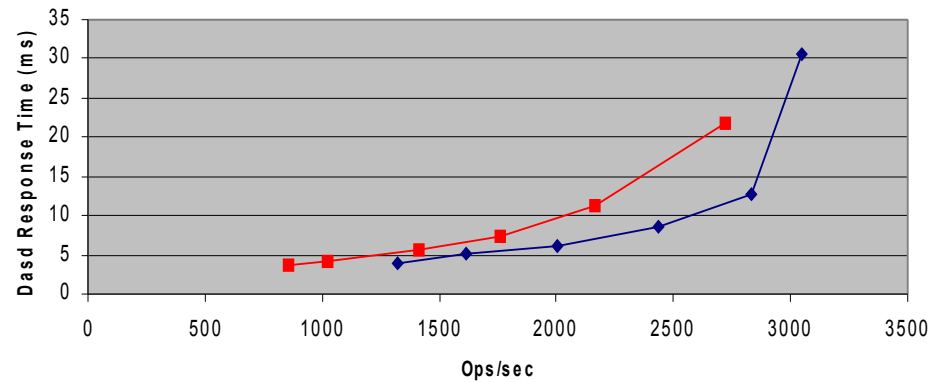


571F CCIN value ordered as #5739, 5778, # 5781, #5782, #5799, #5800

Disk response time and operations per second

Non Encrypted ASP vs Encrypted ASP

9406 MMA 4 Way 571F with 24 DASD Non Encrypted ASP
 9406 MMA 4 Way 571F with 24 DASD Encrypted ASP



Notes

This slide is a composite of figures from the October 2008 Performance Capabilities Reference manual. Since it is relatively easy to create a 6.1 ASP (dependent or independent pools) as encrypted, and all applications run unchanged against a non encrypted ASP or encrypted ASP set of data, you need to plan ahead for the possible performance impact accessing data in an unencrypted ASP compared to an encrypted ASP.

The background left graph shows relative CPU utilization differences for an internal performance benchmark number of transactions comparing encrypted ASP (blue/green bar metrics) to unencrypted ASP stored on the disks attached to our highest performing disk controller identified via the CCIN value 571E and its corresponding auxiliary write cache, CCIN value 575B.

The lower right graph (red line indicates the encrypted performance results) shows the corresponding disk physical operations per second and disk response time for corresponding CPU utilization and internal benchmark transaction rates upper left graph .

Per the October 2008 Performance Capabilities Reference manual, more CPU and memory may be needed to achieve the same performance once encryption is enabled.

The October 2008 (and earlier versions) of the Performance Capabilities Reference manual contain much more disk-related performance information and guidelines – especially highlighting the improved disk performance with the larger write (1.5 GB cache) and read cache (1.5 GB) disk controllers (IOAs).

You can access this publication under the Resource Library link at the URL:

<http://www-03.ibm.com/systems/i/advantages/perfmgmt/>

The next slides switch topics to demonstrate performance results using IBM i 6.1 with the new #5749 (PCIX card) and #5774 (PCIe card) SAN adapters to the attached DS8300 external disk storage configuration.

IBM i 6.1:

Internal Disks, External SAN Attached Disks

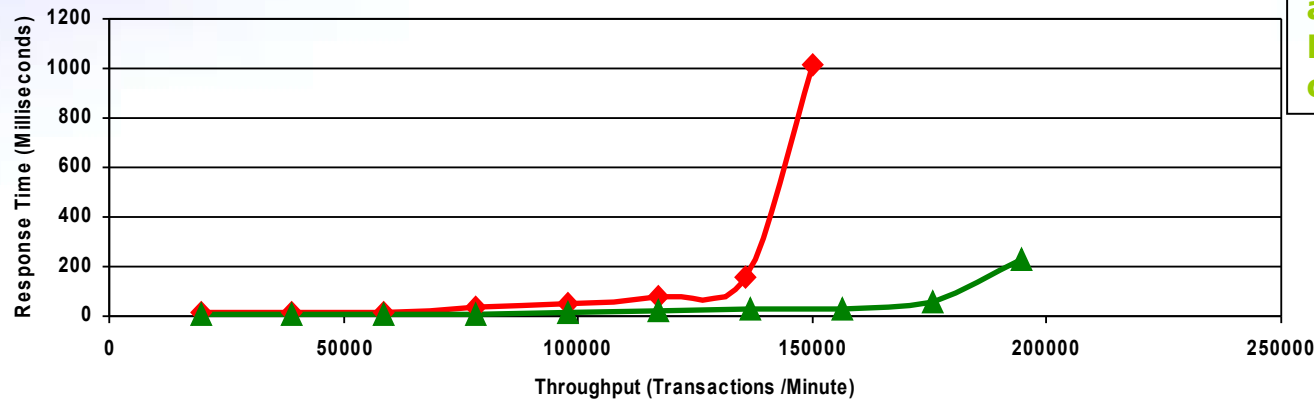
- Internal integrated disks attached to the fastest, largest cache disk controllers deliver best response time
 - CCIN 571E/574F (aux write cache): orderable features #5738, #5777, #5582, #5583
 - CCIN 571F/575B (aux write cache): orderable features #5739, #5778, #5781, #5782, #5799, #5800
- Transaction throughput essentially equal using either fastest internal disk configuration or using new SAN adapters (**#5749/#5774**)
- For external SAN-attached disks, use the disk protection provided by the DS8000 server:
 - For On-line Transaction Processing (OLTP) fastest response time use DSnnn RAID1 (mirroring) or RAID 10 versus RAID5 (assuming same number of disks)
 - 2 physical disks per RAID1 or RAID10 set = better performance
- In general do not share the same physical disks with multiple partitions

Notes

This slide is a summary of the performance tips and considerations listed in the October 2008 Performance Capabilities Reference manual. See that document for more details for the bullets shown on this slide.

SAN IOP Based versus SAN (#5749/#5774) IOP-less

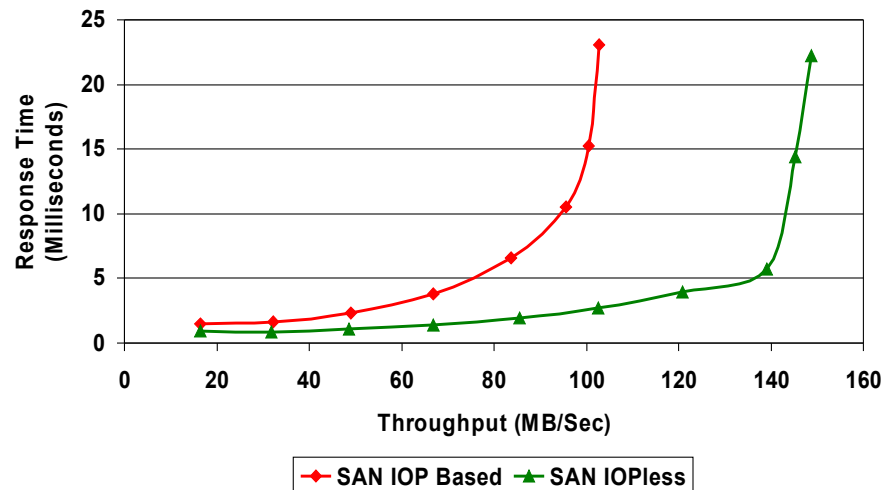
CPW Application



Test results using the Fiber Channel adapter attached to DS8300 configurations

◆ SAN IOP Based
▲ SAN IOPless

CPW DB Disk



◆ SAN IOP Based ▲ SAN IOPless

THE NEW POWER EQUATION

Notes

This slide is a composite of performance test results related to use of the new SAN adapters, #5749/#5774 for SAN attachment and the change I/O operations algorithm tested to the SAN attached disk drive subsystem DS8300.

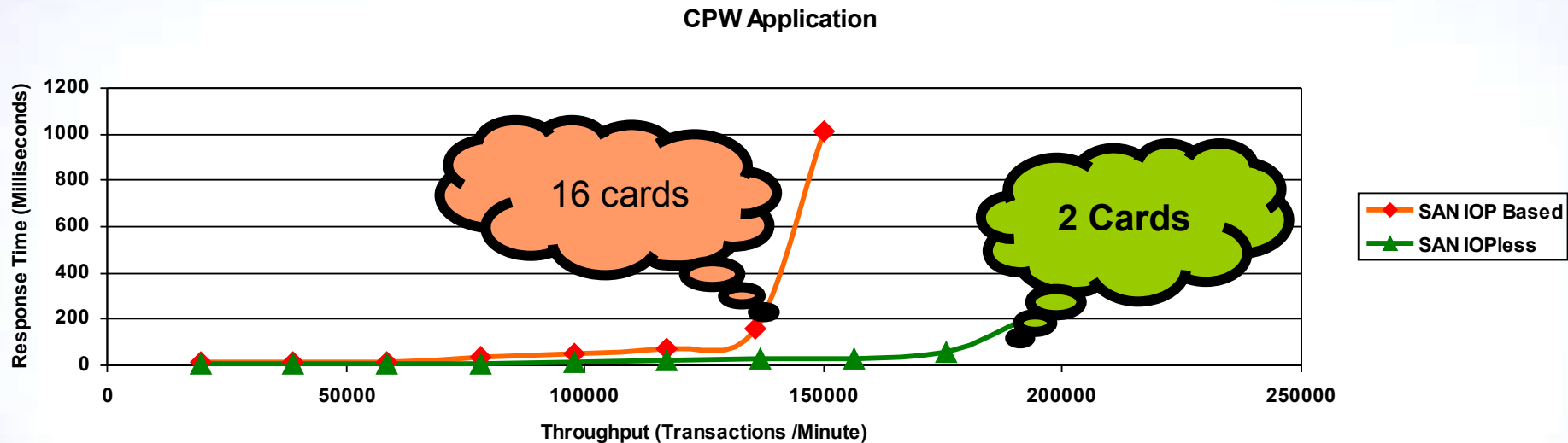
Note, IBM i 6.1 has implemented new algorithms in sending requests to the DS8000 and processing responses from the DS8000 family of IBM Systems Storage products. As of March 2008 these algorithms and the associated performance test results have been tested only to DS8000 configurations.

The objective here is fastest performance, which requires the DS8000 products internal processor and other I/O related capabilities. At this time to attach other external IBM disk hardware, we recommend continuing to use the current SAN IOA-IOP configurations if you wish to continue to attach to the DS6000 family of models. The DS4000 models are currently targeted toward configurations needing moderate to lower I/O operations per second.

These 2 graphs are separated and discussed on the following slides.

At the time this presentation was created, the schedule to include this information in the next update to the Performance Capabilities Reference manual was not available.

SAN IOP versus SAN (#5749/#5774): CPW comparison



Notes

This slide represents the CPW workload 5250 workstation transactions – response time and transaction per hour.

A following slide describes the 5250-focused CPW workload. But for here note that the individual CPU transaction requires very little CPU capacity so it is a very good workload for focusing on the disk component of performance.

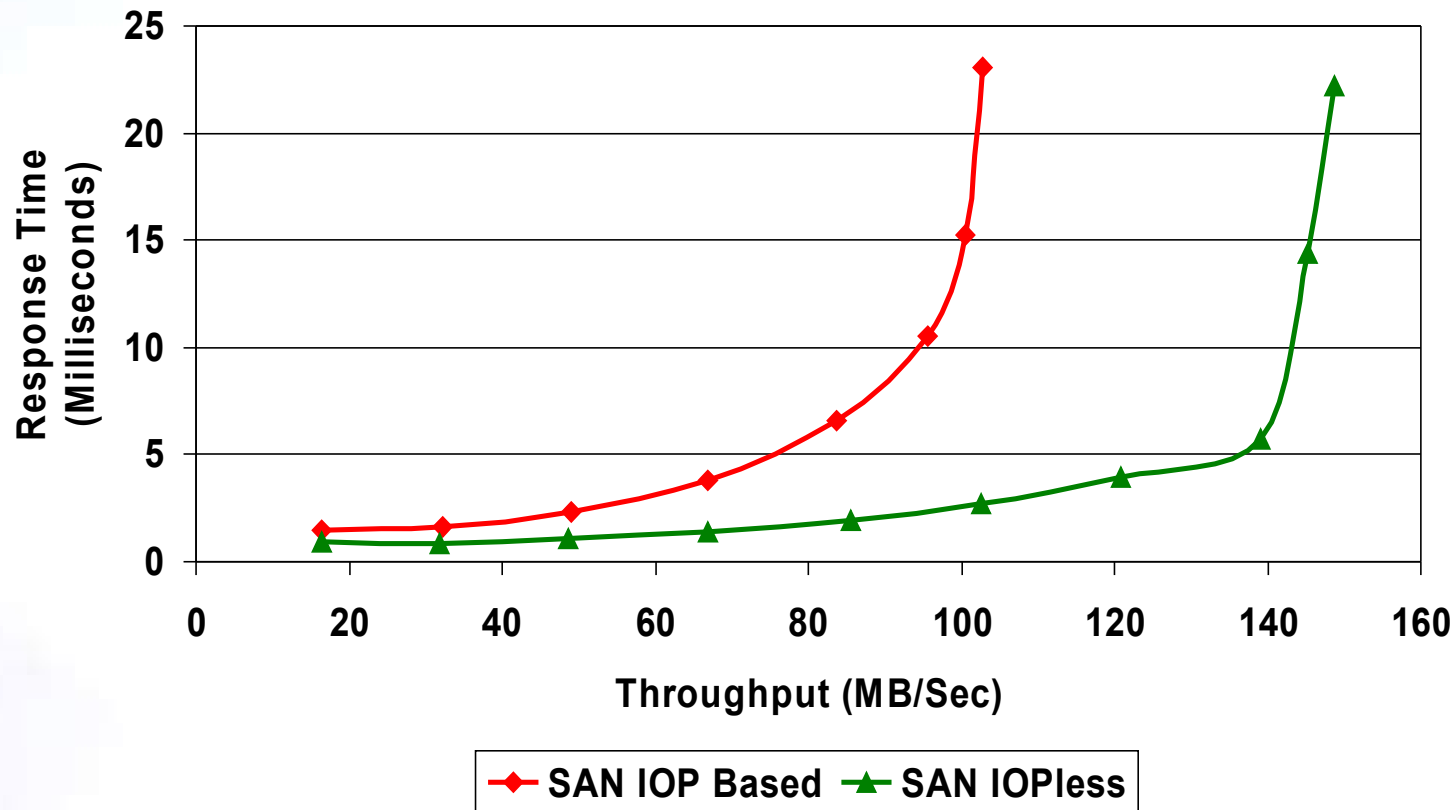
As you can see up to approximately 130,000 CPY transactions per second the response time and number of transactions per hour are essentially the same using the older highest speed IOP- disk IOA hardware compared to the new IOP-less 5749/#5774 SAN Fibre Channel adapters.

And, since the #5749/#5774 are dual port adapters, you can take advantage of multipathing, generally used for both availability and highest performance possible with external disks. To get the same number of “external disks” (actually logical LUNs) defined to i5/OS requires a much smaller number of FC IOAs than the previous highest speed FC IOA-IOP technology number of card slots.

Keep in mind, we are considering external disk performance, with the new 6.1 implementation and SAN IOA (same maximum speed rating as existing IOP-required IOAs), to be essentially equivalent to integrated disks and adapters. In general, however, the disk-based hardware costs are different and should be understood before choosing between internal and external disk configuration.

SAN IOP versus SAN (#5749/#5774): DB2 operations

CPW DB Disk Operations



Notes

This slide represents the CPW database requests physical disk I/O response time and Megabytes per second rates with the new #5749/#5774 IOAs.

Up to 5 millisecond second disk response times, applications similar to the CPW workload should perform the same. As the number of characters increases, you can see the benefits of the new disk SAN adapter and V6R1 implementation.

The next slide overviews the CPW workload used in the previous slides on the external disk performance.

CPW Workload Characteristics

- Heaviest workload: many 5250 workstation jobs performing simple to medium complexity database updates (non SQL I/O)
- Batch background job
- Journaling with commitment control
- Focused on a single steady-state, data base oriented commercial application with a large number of concurrent users .
- System ASP
 - Very light I/O
- DB IASP (Database physical I/O operations to an independent Auxiliary Storage Pool)
 - Random small block (average 6K – 8K) reads and writes
 - High percentage of skip writes
- JO IASP (Journal receiver physical I/O operations to an independent Auxiliary Storage Pool)
 - Nearly all medium sized (16K- 20K) sequential writes

Notes

The full description of CPW – Commercial Processing Workload, is contained in Appendix A of the IBM i Performance Capabilities Reference manual described in the beginning pages of this presentation.

Details: new Disk I/O algorithms using #5749, #5774 SAN adapters

Implementation areas new or rewritten

- **Tagged Command Queuing**

- Allows multiple ops to each LUN-connection
- Leverages DS8000 concurrency capabilities

- **Header Strip/Merge Improvements**

- Packages 8 sector headers into single direct memory access (DMA)
- Reduces traffic on HSL and PCIX
- Reduces overall IO latency

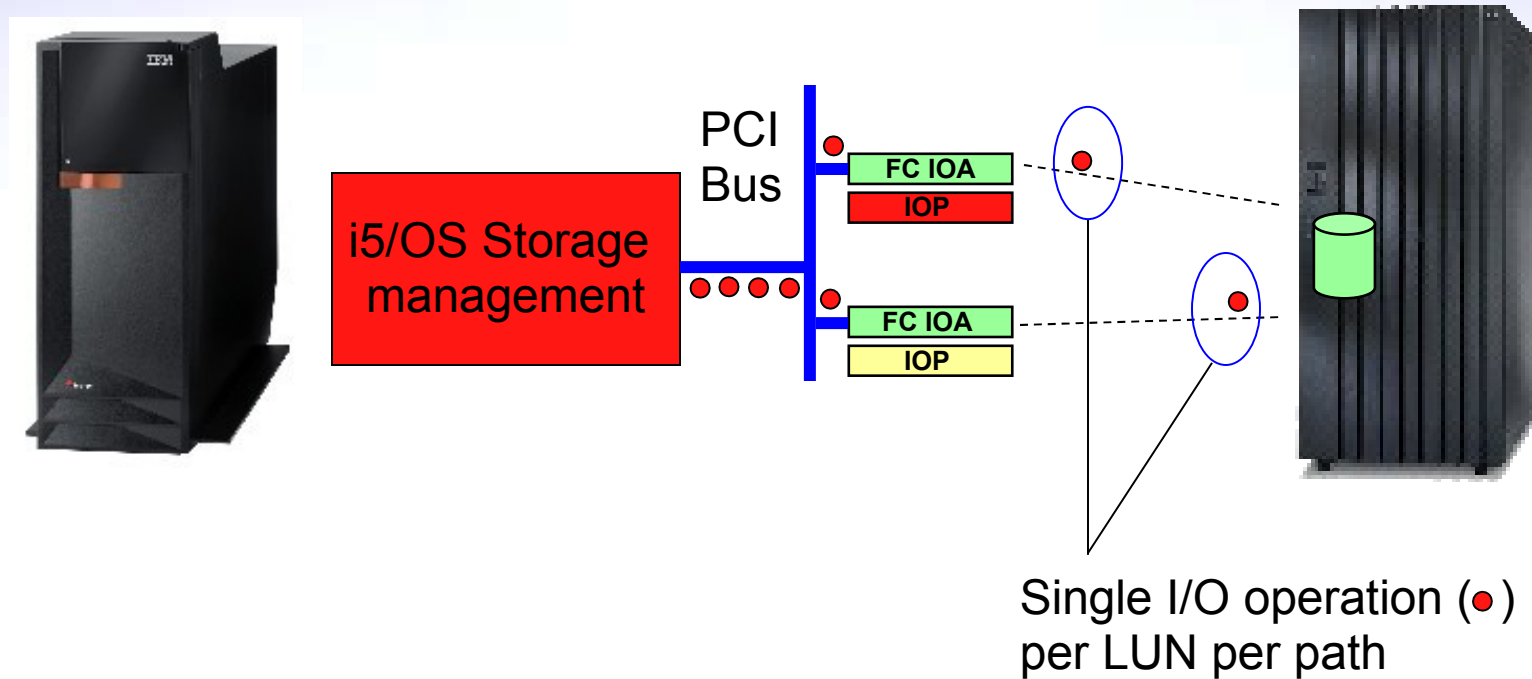
- **No IOP in the command path**

- Reduces IO latency – most significant on cache hits

Notes

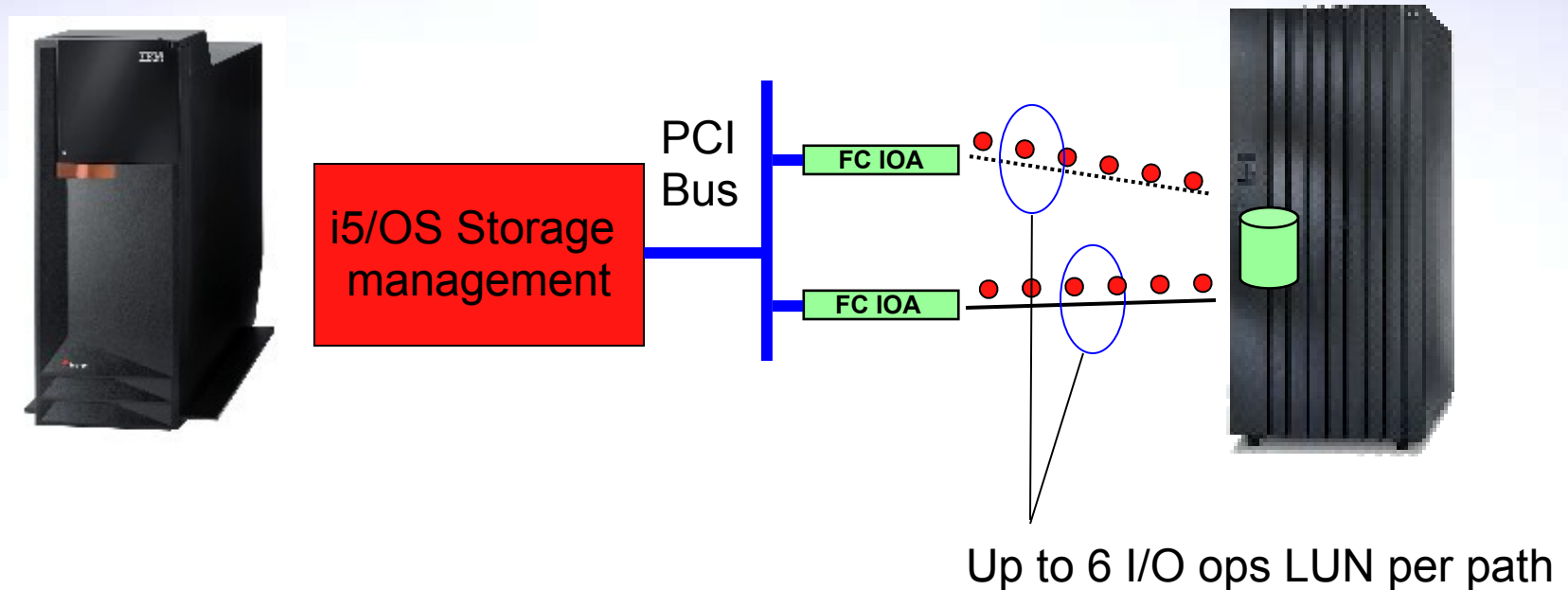
The following slides provide graphical representations of the implementation details.

Current I/O Queuing and IBM i



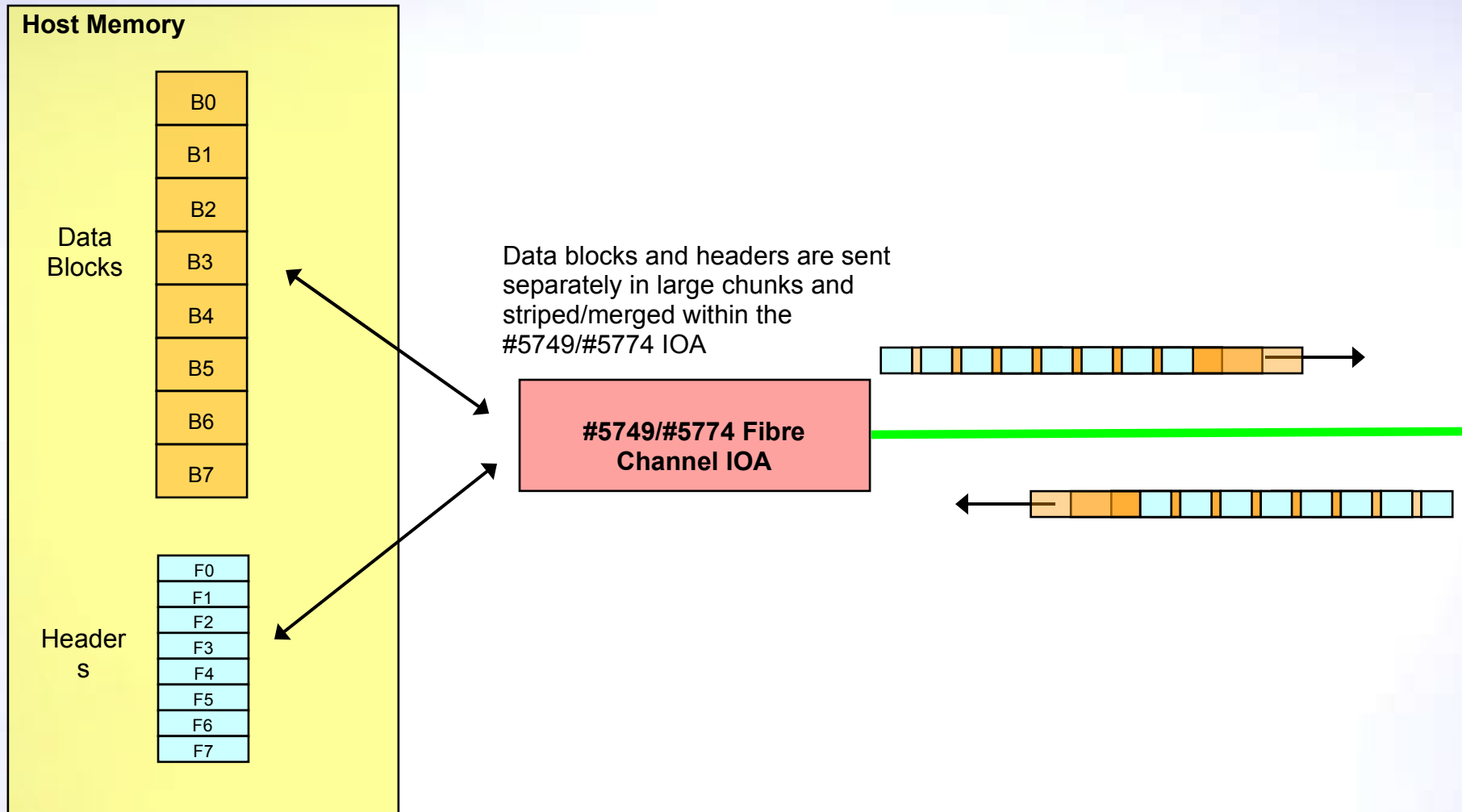
- Limited to one I/O operation (blue circle) per LUN per path
- Can improve the number of I/Os per LUN by using multi-path and multiple connections (two shown in this example). This implementation continues when using IOP-IOA based SAN controllers other than 6.1 with the #5749/#5774 IOAs.

Tagged Command Queuing with IBM i 6.1



- Limited to 6 I/O operations (**blue circles**) per LUN per path
- Can improve the number of I/Os per LUN with multi-path and multiple connections:
 - 2 connections = 12 I/Os per LUN
 - 3 connections = 18 I/Os per LUN

Header Strip/Merge function



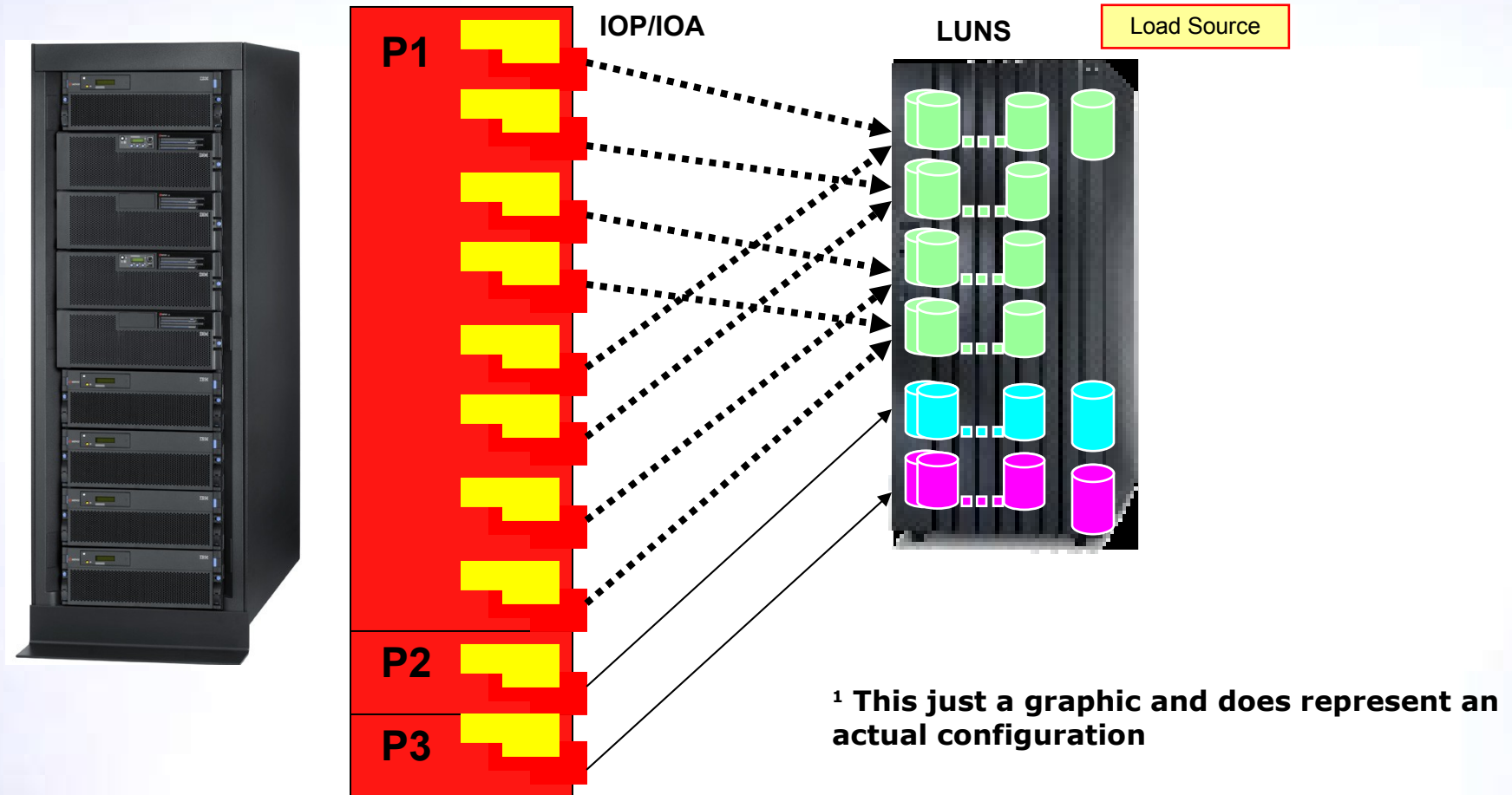
Notes

This slide illustrates the new Header Strip/Merge algorithm used with the #5749/#5774 IOAs.

As shown, data blocks and headers are sent separately in large chunks and striped/merged within the #5749/#5774 IOA. This reduces the amount of “data” exchanged over the SAN.

Current SAN implementation: Multipath

- Multipath IOP/IOA (single port) combination connection to external storage ¹



Notes

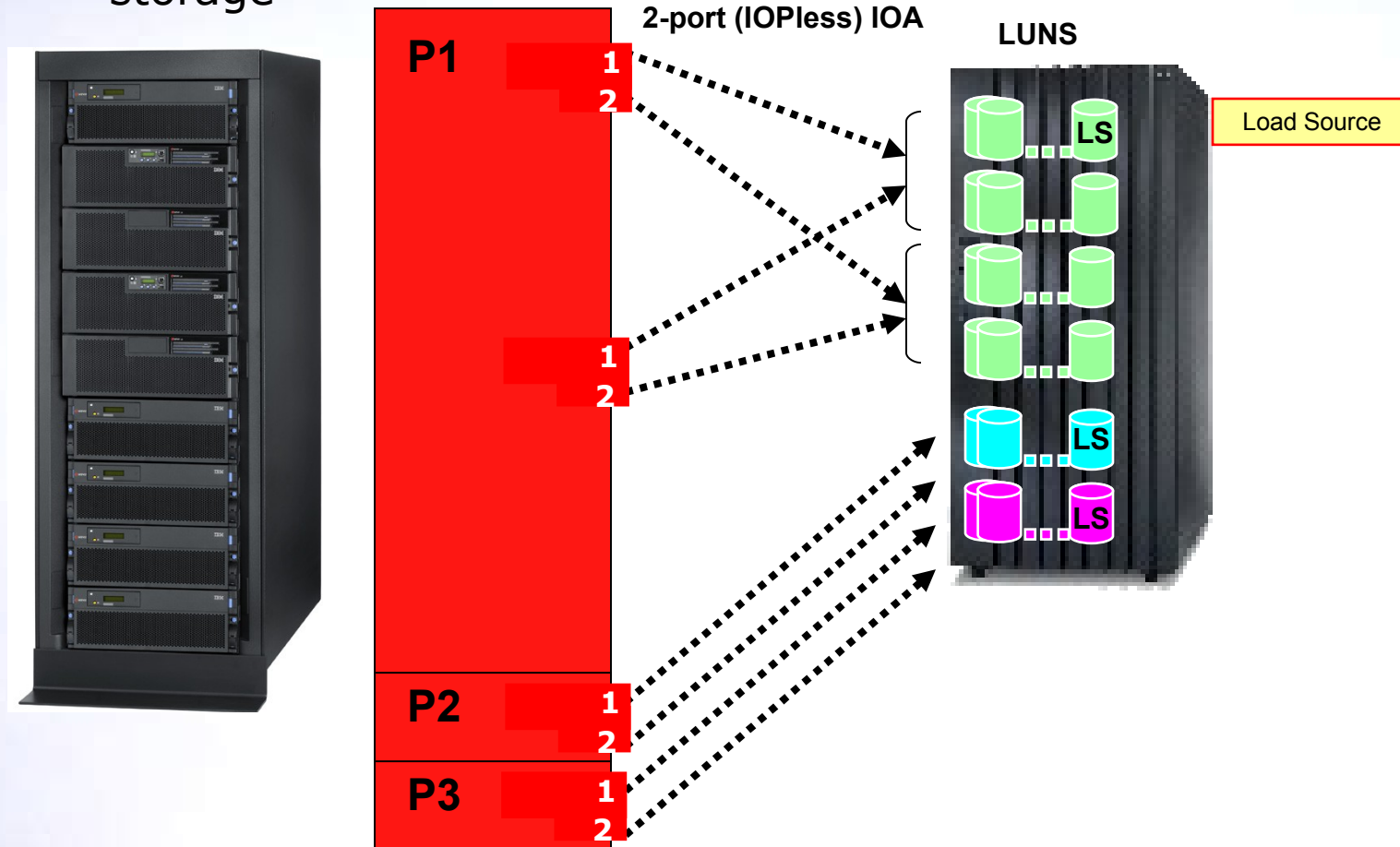
In this slide, the yellow graphic represents an IOP. The red graphic represents the IOA. In V5R1, the iSeries added support for connecting to an external storage server (ESS) DASD (disk) subsystem via a Fibre Channel connection (the device type each "virtual disk" was 2105).

This initial support provided for a single connection per Logical Unit (LUN) – logical disk configured within the ESS DASD Subsystem. Note that the iSeries may have more than one connection to a physical ESS box, each to a different LUN, and may have multiple LUNs on a single connection, but each LUN has only one connection to the iSeries.

Licensed internal code (LIC) level V5R3M5 introduced support for load source.

POWER6 and 6.1 SAN implementation: Multipath

- Multipath #5749/#5774 IOA combination connection to external storage



Sizing “disk I/Os” capacity with an External Storage System

- Size all external DSnnnn storage using Disk Magic Tool provided by IntelliMagic Corporation for IBM
 - For IBM i:
 - Input from Performance Tools for i System and Component Report disk statistics
 - Input from the Workload Estimator tool
 - Manually entered “disk statistics”

Notes

Disk Magic is a powerful and versatile performance modeling product for IBM, EMC and Hitachi Enterprise Disk Subsystems, provided by IntelliMagic corporation. It is available for download only to IBMers and authorized business partners, by searching the PartnerWorld or internal IBM Systems Sales Websites with "Disk Magic."

IBM has licensed Disk Magic from IntelliMagic since 1994 and used it at thousands of sites around the world to analyze and predict the performance of customer workloads on new disk subsystems and technologies. In 2006, IntelliMagic released the end-user version of Disk Magic which now also supports popular disk subsystem configurations from EMC, Hitachi Data Systems and HP in addition to IBM.

Disk Magic offers complete and current support for disk subsystems such as the IBM Enterprise Storage Server, DS4000, DS6000 and DS8000 family of products, SAN Volume Controller (SVC), for the EMC DMX and 8000 range, for the HDS 9000V and USP, and for the HP XP512, XP128, XP1024, XP10000 and XP12000. All modeling is based on your own installation's recorded I/O load, server configuration and service level targets.

For IBM hardware, Disk Magic supports disk subsystems attached to any platform: System z, System i, System p, System x, and Unix (includes AIX), Linux, and Windows servers. Modeling of z/OS workloads is also supported on the EMC disk subsystems 8230, 8530, 8830, DMX Generations 1, 2 and 3, on the HDS disk subsystems 9900 and 9900V and the Universal Storage Processor (USP) models, as well as on the HP XP models.

Disk Magic helps you to select the most cost-effective configuration options, such as the disk drive technology, the cache size, the number of Fibre channels and more.

Notes

The I/O operations per second and card replacement values shown are good starting points for “sizing a configuration” using the new #5749/#5774 adapters. However, use of Disk Magic will help prevent wrong assumptions on your part on what the current I/O workload (before installing the new SAN adapters) is.

Specific System i and i5/OS capabilities within Disk Magic include:

- Disk Magic accepts Performance Tools for iSeries, 5721-PT1 (i5/OS V5R4 and earlier) and Performance Tools for i5/OS, 5761-PT1 (i5/OS V6R1 and later) input via the System Report and Component Report spooled output.
- The IBM Systems Workload Estimator (for IBM i) supports inputting its disk sizing output into Disk Magic
- i5/OS support for IOP-based SAN adapters and DS6000 and DS8000 products is supported by Disk Magic
- Disk Magic March-April 2008 enhancements are planned to include support for IBM i 6.1 support of the DS8000 models attaching via the #5749/#5774 IOAs and Virtual I/O Server (VIOS) supporting DS4700 and DS4800 with an i 6.1 client partition.

Note, VIOS is based upon AIX which supports the #5774 (PCIe form factor card), but not the #5749 (PCIX form factor card).

The next slide gives a short summary of the steps for using Disk Magic to size hardware requirements to estimate disk performance of a DSnnnn configuration attached to System i Fiber Channel adapters.

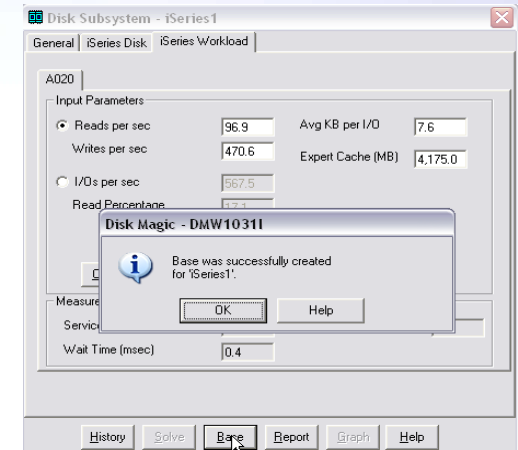
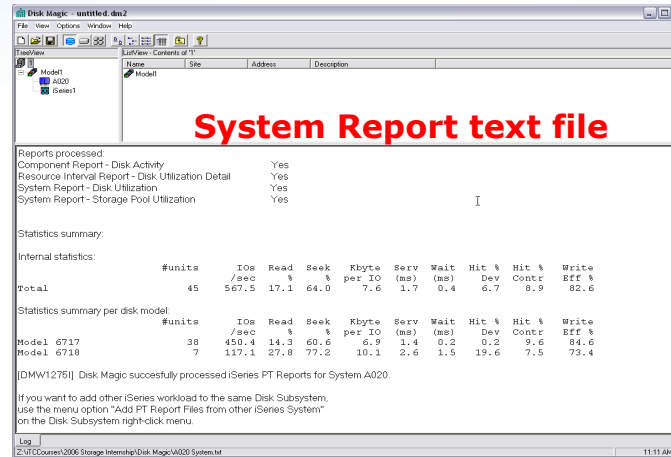
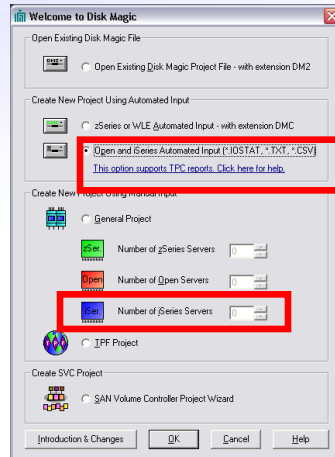
Note, it is beyond the scope of this presentation to provide complete coverage of how to use Disk Magic for an i5/OS partition. This takes a level of expertise and experience using Disk Magic, as well as in use of the DSnnnn configuration and disk LUN configuration.

Sizing with Disk Magic

Performance
Tools Reports

Workload
Estimator
Input

Manual Input
of Disk
Workload

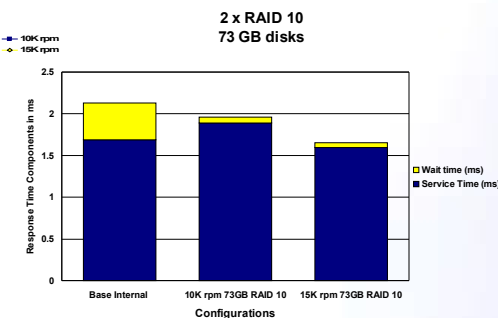
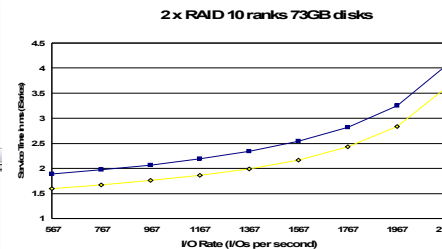
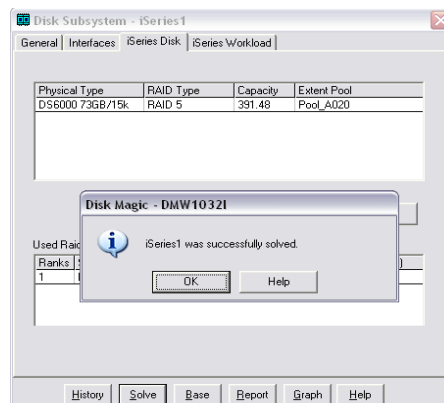
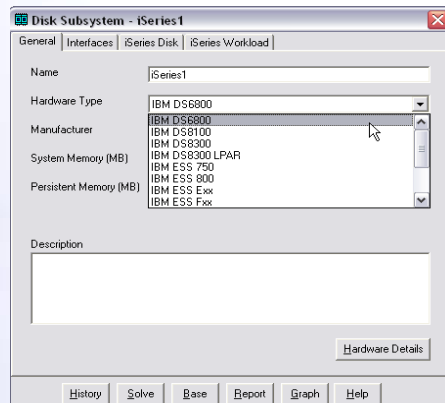


1) Gather Input Data

2) Start Disk Magic

3) Read in or Key in Base System Data

4) Create the Base System Model



6) "Solve" the Model

7) Print Results – Reports + Graphs showing Response Time + Growth Projections

THE NEW POWER EQUATION

Notes:

As stated Disk Magic accepts as input disk performance metrics from the Performance Tools reports (as .txt files), Workload Estimator disk "results information, and manually entered data into Disk Management screens.

The top leftmost window has an option identified that tell Disk Magic will be sizing for a System and will be using System and Component report files in .txt format. The middle top row screen is a summary of the input indicating Performance Tools for i5/OS data was used for two internal disk drives. The rightmost top row screen shows in the background average disk I/Os per second and size (in kilobytes) of each I/O and, if any, i5/OS Exert Cache size, and the window saying the "base model" was created successfully. You can manually change the values shown assuming you know why you would be changing the values.

Note, in this simple example, only 2 disk drives are shown in the report example.

In the leftmost lower row screen one of the tabs for modifying the "base model" is shown. The General window shows selecting a DS6000 configuration model. Moving to the right, the "second" screen capture shows one physical DS6000 disk, with RAID5 protection will be used for DS6000 storage pool A020. In a real sizing effort you would probably use more than one physical disk. Click OK and the "disk configuration" is "solved" window appears.

The two graphs on the lower right show" the "solved results:"

Graph 1: x axis shows I/Os per second, y axis average disk service time. The yellow graph line is for 10K disk arms and the dark blue graph line is for 15K rpm disks.

Graph 2: shows, total disk average response time broken out into average disk wait time (yellow) and average service time (dark blue) for:

- Base internal disk configuration (measured data for 2 physical disks)
- DS6000 with 10K rpm 73 GB with RAID 10
- DS6000 with 15K rpm 73 GB with RAID 10

The output assumes you are not sharing physical DS6000 disk drives among multiple "host systems."

VIOS with IBM i 6.1

IBM i 6.1 as a Client Partition

- VIOS as the hosting partition on non-BladeCenter partitions
- DS4700 and DS4800 on POWER BladeCenter with an i 6.1 client partition

VIOS hosting an IBM i 6.1 partition



- Performance tips from the October 2008 Performance Capabilities Reference manual
 - Choosing the right initial configuration is most important
 - Use IBM Workload Estimator (WLE) and Disk Magic for sizing both the VIOS partition and the i client partition, especially SAN disk and tape attachment using new Fibre Channel adapters #5749 (PCIX) and #5774 (PCIE)
 - Use “best practices” documentation already existing for SAN device attachment
 - When using VIOS, best performance is with external SAN disks than internal disks managed via VIOS
 - Consider using more than just 2 virtual disks per i5/OS partition – more disks mean more concurrent disk I/O operations can be issued by i
 - For BladeCenter hosting VIOS partitions note the maximum of up to 16 LUNs
 - If journaling to disks managed by VIOS, use a separate ASP for journaling to separate journal entry targeted disks from “data disks”
 - In general use a VIOS partition defined to use dedicated processor units rather than shared processor, processor units.

Notes

- See the performance tips for SAN attached disk drives. They also apply to VIOS partitions with SAN attached DSnnnn disk drives. See the slide **IBM i 6.1: internal disks - external SAN attached disks.**

Notes

The October 2008 Performance Capabilities Reference manual has extensive coverage of performance tests with various numbers of VIOS attached DS4800 DDMs (disk arms) configured to be virtual disks to one or more client IBM i 6.1 partitions. Various processor unit combinations were assigned to the VIOS partition and the i client partition.

As with the January 2008 i support on the POWER6 JS22 7998-61X BladeCenter H chassis announcement, the disks configuration was fixed to include only DS4700 and DS4800 configurations. Other DSnnn configurations can be “plugged in,” but DS4700 and DS4800 are indicators of the performance expectations that should be assumed. That is, if you attach DS8000 level hardware, the i, VIOS and BladeCenter hardware configuration will not achieve DS8000 performance maximums.

The following slides show two test result charts achieved using:

- 4 processor 7998-61X Blade in a Blade Center H chassis
- 32 GB of memory
- DS4800 with a total of 90 DDMs
 - 8 DDMs using RAID1 externalized in 2 LUNs for the system ASP
 - 6 DDMs in each of 12 RAID1 LUNs (a total of 72 DDMs) in the database ASP
 - 10 DDMs unprotected externalized in 2 LUNs for the journal ASP

The Blade partition physically had two Fibre Channel attachments to the DS4800 with half of the LUNs in each of the ASP's using controller A as the preferred path and the other half of the LUNs using controller B as the preferred path.

Recall, a description of the Commercial Performance Workload can be found in appendix A of the Performance Capabilities Reference.

Notes

The 7998-61X Blade 4 processor unit ran several partitions. As the number of partitions increases above 4, things get a little hard to manage. See the Performance Capabilities Reference manual and the following URL for an i5/OS-based whitepaper on LPAR.

<http://www-03.ibm.com/systems/resources/lparperf.pdf>

For most testing one i partition was utilized on the Blade configuration. Note that VIOS is the base operating system on the 7998-61X Blade, installed on the internal SAS Disk (one) and VIOS must Virtualize all S4800 LUNs and communication resources to the i5/OS partition, which itself resides totally on the DS4800 configuration.

VIOS and Integrated Virtualization Manager (IVM) must have some of the 7998-61X Blade's memory and processor resources for this virtualization. The amount of resources needed will be dependent on the physical hardware in the Blade Center and the number of partitions being supported on a particular Blade. For the testing done to produce the Performance Capabilities Reference manual graphics, testing indicated running the tests were not possible when attempting to run under 1 GB of memory. Therefore all testing was done using 2 GB of memory.

The number of processor units and main memory for the VIOS partition and memory were varied only for the shared processor pool the VIOS partition was using. The i partition was created always using dedicated processor units.

IBM JS22 BladeCenter Performance

- **4way BladeCenter CPW 11,040**
 - 4way POWER (JS22 (7998-61X)), processor feature 7998; 4.0 GHz processor
 - L2 cache: 4x4MB, dedicated L2 cache per processor core
 - Zero L3 cache
 - 11,040 CPW rating:
 - Based upon i5/OS 3 dedicated processor partition with a 1 processor dedicated to VIOS partition
 - Consider constraint introduced by maximum of 16 virtual LUNs available (comparable to System i POWER5 model 515 CPW rating with constraint of maximum of 8 real disks)
 - Compare to POWER5+ 520 2-Way with CPW range of 3800-7100 and POWER5 550 4-Way with CPW range of 3800-14000. These CPW ratings are achieved with maximum main storage and disk configurations.
- **Best CPW performance CPU configuration**
 - 3 processor dedicated System i partition
 - One processor for VIOS
 - If your application environment is one that performs very few disk I/O operations (in contrast to the CPW workload) using less than one processor capacity for your VIOS partition is recommended
- **Blade requires more DS4800 DDMs**
 - Maximum of 16 virtual LUNs available
 - Cannot use two disk RAID 10 arrays

Notes

See the October 2008 IBM i 6.1 Performance Capabilities Reference manual for more a wide range of performance test results using i 6.1 as a client partition to a VIOS server partition.

IBM i 6.1 Performance Data Investigator

IBM Systems Director Navigator

Investigate Data

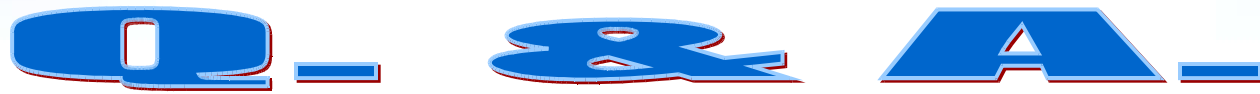
- **Web-based graphical analysis interface for i performance data**
 - Easy to find and use performance functions
 - Powerful, consolidated collection visualization
 - Supports analysis of data from
 - Collection Services (free with base operating system)
 - Disk Watcher (included with PT1 Manager option)
 - Job Watcher (new chargeable option of PT1)

Investigate Data

- **Visualize data in charts and tables**
 - Fully-interactive charts allow zooming/panning/selecting
 - Tables offer sorting, filtering, searching of entire data set.
- **Drill-down to related charts, retaining context-sensitive data**
 - Time range
 - Job names
 - Subsystems
 - Etc.
- **Save customized charts/tables for future reference**
 - Or share URL with others!
- **Extensible architecture to support other (IBM or ISV based) data collections**

Other Reference Material

- **Performance Management for System i and the IBM i 6.1 Performance Capabilities Reference**
<http://www-03.ibm.com/systems/i/advantages/perfmgmt/>
- **i5/OS LPAR Performance on POWER4 and POWER5 Systems**
<http://www-03.ibm.com/systems/resources/lparperf.pdf>
- **IBM Systems Workload Estimator**
<http://www.ibm.com/systems/support/tools/estimator>
- **eServer i5 Uncapped Partition CPU Utilization & Queuing**
http://www-03.ibm.com/systems/resources/systems_i_advantages_perfmgmt_pdf_lparcpu.pdf
- **Simultaneous Multi-Threading on eServer iSeries POWER5 Processors**
http://www-03.ibm.com/systems/resources/systems_i_advantages_perfmgmt_pdf_SMT.pdf
- **Enhanced eServer iSeries Performance Memory Affinity**
http://www-03.ibm.com/systems/resources/systems_i_advantages_perfmgmt_pdf_memaffin.pdf



Trademarks and Disclaimers

8 IBM Corporation 1994-2008. All rights reserved.

References in this document to IBM products or services do not imply that IBM intends to make them available in every country.

Trademarks of International Business Machines Corporation in the United States, other countries, or both can be found on the World Wide Web at <http://www.ibm.com/legal/copytrade.shtml>.

Intel, Intel logo, Intel Inside, Intel Inside logo, Intel Centrino, Intel Centrino logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium, and Pentium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

Linux is a registered trademark of Linus Torvalds 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.

IT Infrastructure Library is a registered trademark of the Central Computer and Telecommunications Agency which is now part of the Office of Government Commerce.

ITIL is a registered trademark, and a registered community trademark of the Office of Government Commerce, and is registered in the U.S. Patent and Trademark Office.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Java and all Java-based trademarks are trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.

Other company, product, or service names may be trademarks or service marks of others.

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

The 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 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.

Some information 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.

Prices are suggested U.S. list prices and are subject to change without notice. Starting price may not include a hard drive, operating system or other features. Contact your IBM representative or Business Partner for the most current pricing in your geography.

Photographs shown may be engineering prototypes. Changes may be incorporated in production models.