Share Technical Conference
San Francisco   2/99
Share Session 2591

OS/390 Unix
Performance
Update
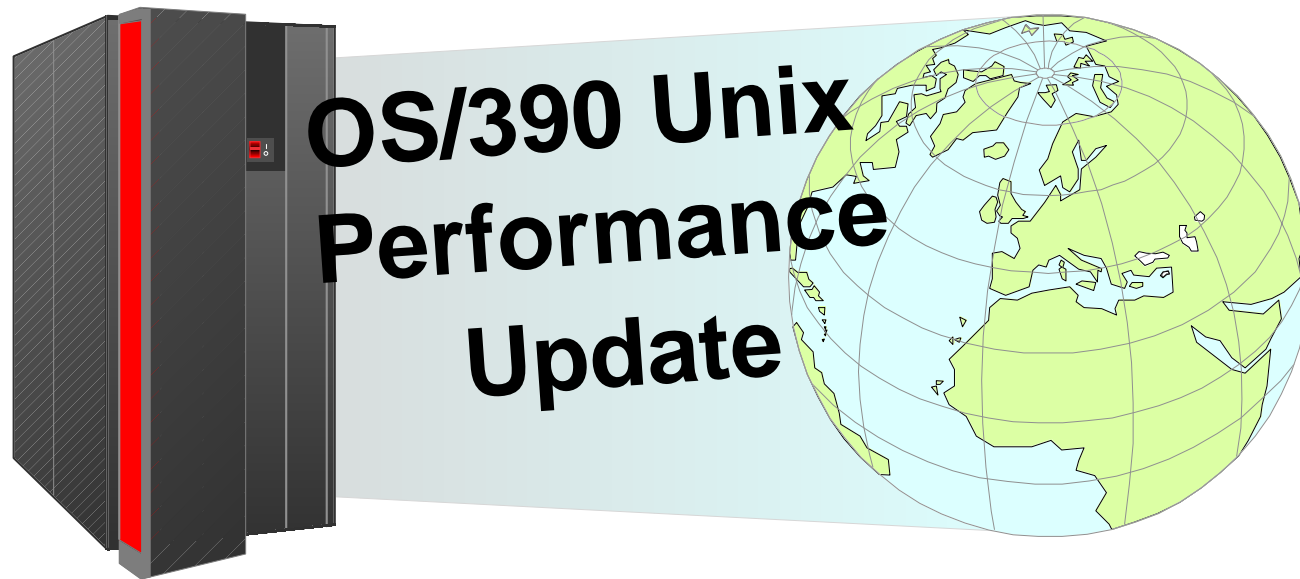
Carl Parris

OS/390 Design and Performance Analysis
Mid-Hudson Valley Programming Lab
914-435-6287
parris@us.ibm.com

Your Mileage May Vary Pt.1

Your Mileage May Vary Pt.2 **IBM**

Some of the measurements quoted in this presentation may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

UNIX is a registered trademark in the United States and other countries, licensed exclusively through X/Open Company, Limited. Other company, product and service names may be trademarks or service marks of others.

Trademarks or Registered trademarks of International Business Machines Inc.

ACF/VTAM tm	MVS tm
CICS tm	OpenEdition tm
DB2 tm	OS/390 tm
DFSMS tm	RACF tm
ESCON tm	RMF tm
IBM tm	S/390 tm
IMS tm	

Your Mileage May Vary Pt.3

Unless otherwise noted all performance results
contained in this presentation are from IBM
internal workload measurements.

Agenda

IBM

- 'Why OS/390 UNIX' Overview

- Technical View of Performance
  - ► TCP/IP
  - ► File System
  - ► NFS
  - ► C/C++ Compiler
  - ► Kernel/RTL
  - ► DB2
  - ► Web
  - ► Net.commerce

Overview

Today's Distributed Environment

IBM

Major Data Moves

- End users access application (and data) via distributed process
- OS/390 and other servers transform data when both process it
- ADSM / MVS backsup data from LAN file/database servers

Appl - Data transform

Data Source

Applic users

TCP/IP

TCP/IP

APPLIC ebcdic

TCP/IP SNA

ADSM
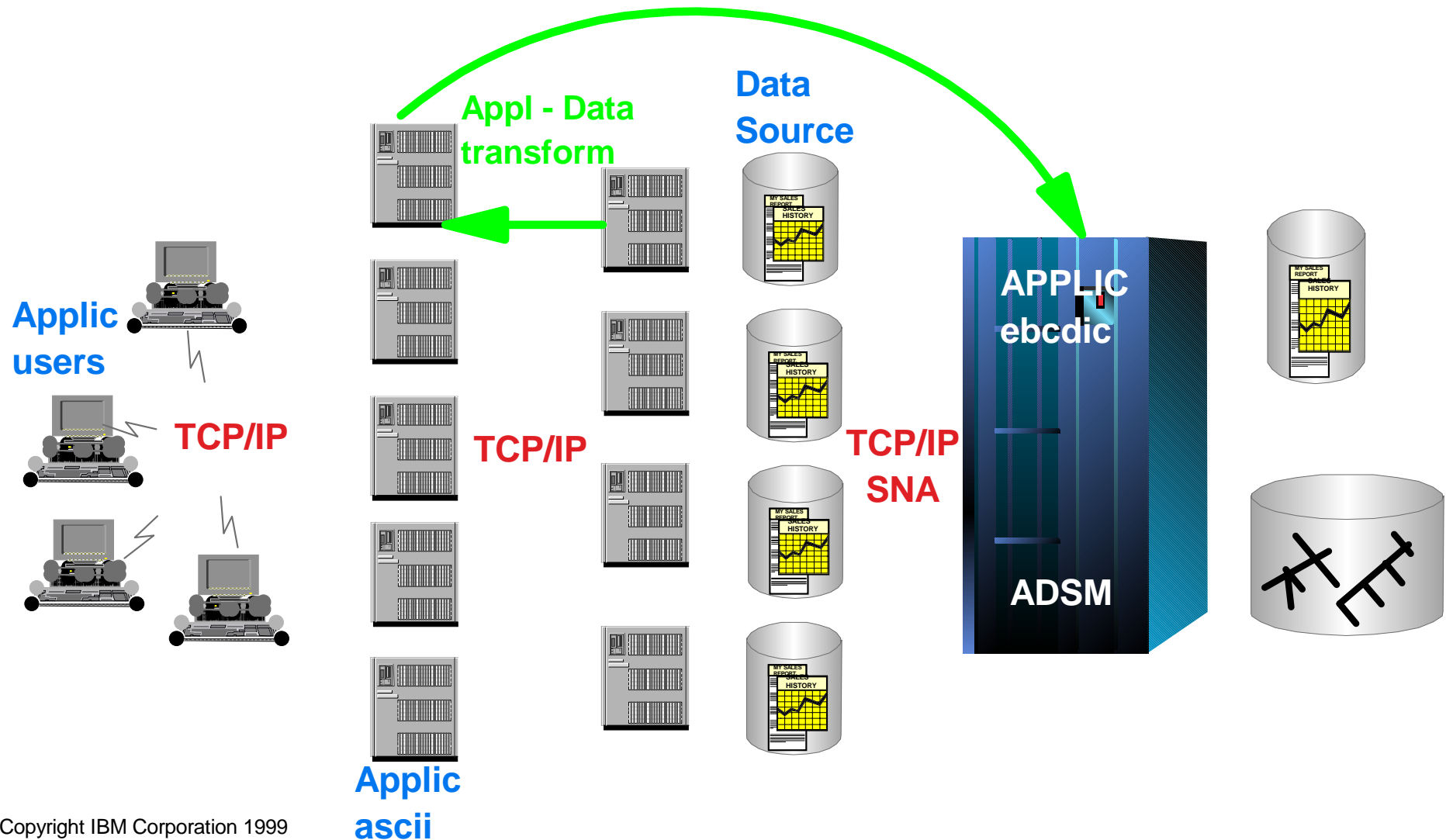
Applic ascii

Application Consolidation - End Step

IBM

Applic users

TCP/IP

APPLIC

Data

ADSM

MY SALES REPORT
SALES HISTORY

MY SALES REPORT
SALES HISTORY

MY SALES REPORT
SALES HISTORY

MY SALES REPORT
SALES HISTORY
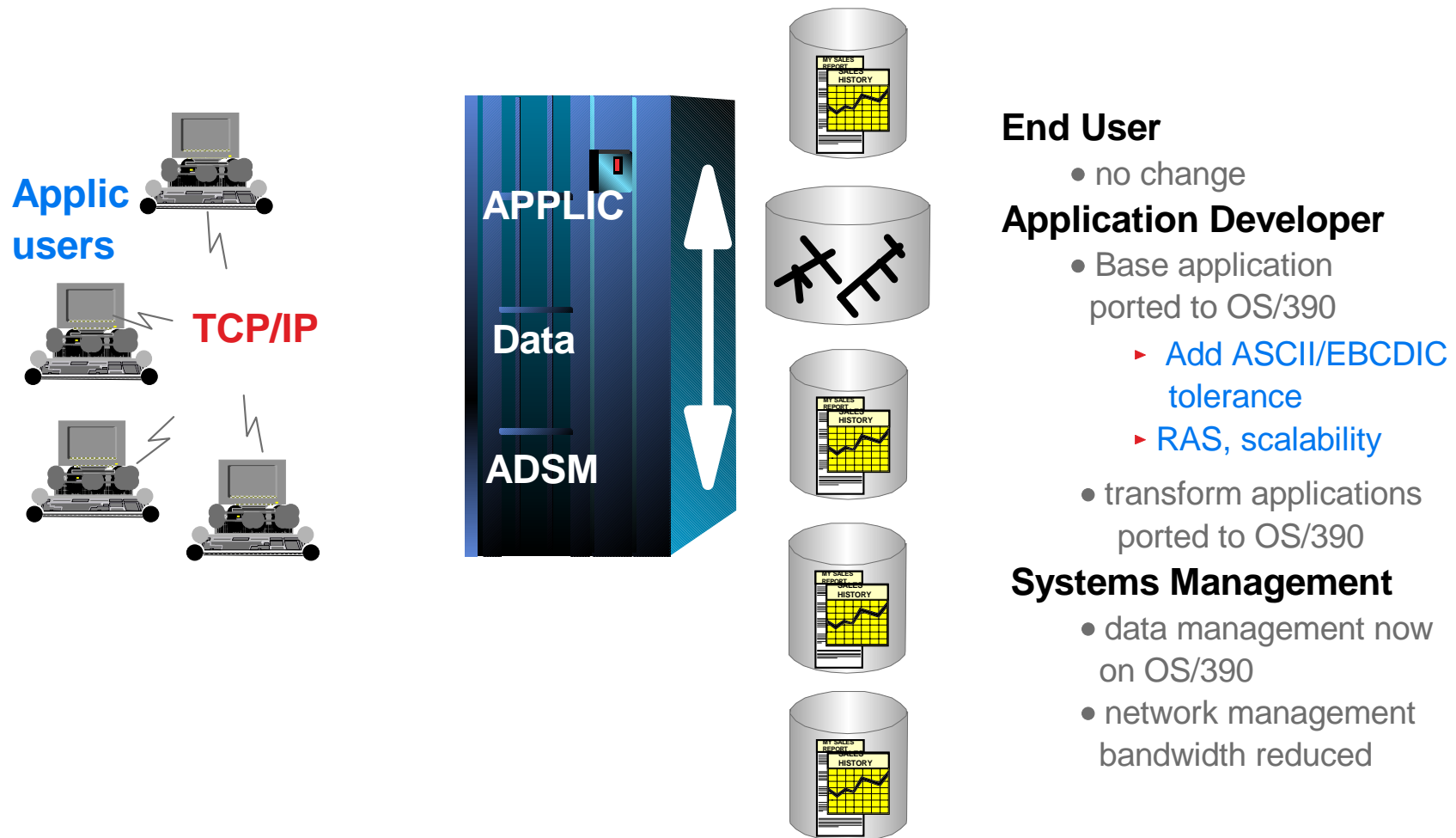
End User
- no change

Application Developer
- Base application ported to OS/390
  - ▸ Add ASCII/EBCDIC tolerance
  - ▸ RAS, scalability
- transform applications ported to OS/390

Systems Management
- data management now on OS/390
- network management bandwidth reduced

Major Data Moves
- End users access applications (and data) via distributed protocols

Data Moves Avoided
- ADSM / MVS backups from LAN file/database server
- OS/390 and other servers transform data when both process it
- Application to Data move now occurs within OS/390

Configuration Strategies

Application and Data Consolidation

Application migration using OS/390 to eliminate tiers of servers to Reduce:

- movement of data
- protocol handshakes
- data transforms
- points of failure
- complexity
- system management costs

and provide Improved:

- large scale deployment
- price/performance
- availability
- security
- resource balancing and management
- configuration management

Technical View

Development Emphasis

- OS/390 UNIX implementation is highly integrated in base operating system

- UNIX 95 Branding is functionally robust

- Performance Improvements coming regularly

- Development organization view performance
  as 'key' focus area

- Significant performance gains in '97, '98,'99

CS for OS/390 V2R6 for TCP/IP

IBM

- Protocol stack re-written for performance
- Improved SMP scalability
- RAS improvements

CS OS/390 V2R5 & V2R6 performance compared to MVS TCP/IP V3R2 on S/390 MP 3-way and 10-way configurations:

9672-RX53-Way:

Workload	CS OS/390 V2R5	CS OS/390 V2R6
Interactive	4.84X	5.66X
Web	4.16X	7.40X
File Tfr	1.64X	2.35X

9672-RY5 10-Way:

Workload	OS/390 V2R5	OS/390 V2R6
Interactive	11.51X	14.97X
Web	9.72X	14.22X
File Tfr	2.89X	5.31X

Interactive Workload

Throughput (Interactions/sec) vs Number of CPUs (9672-RX5)

1-way: 912, 1682, 2946, 4891
3-way: 1312, 3839, 6352, 7430
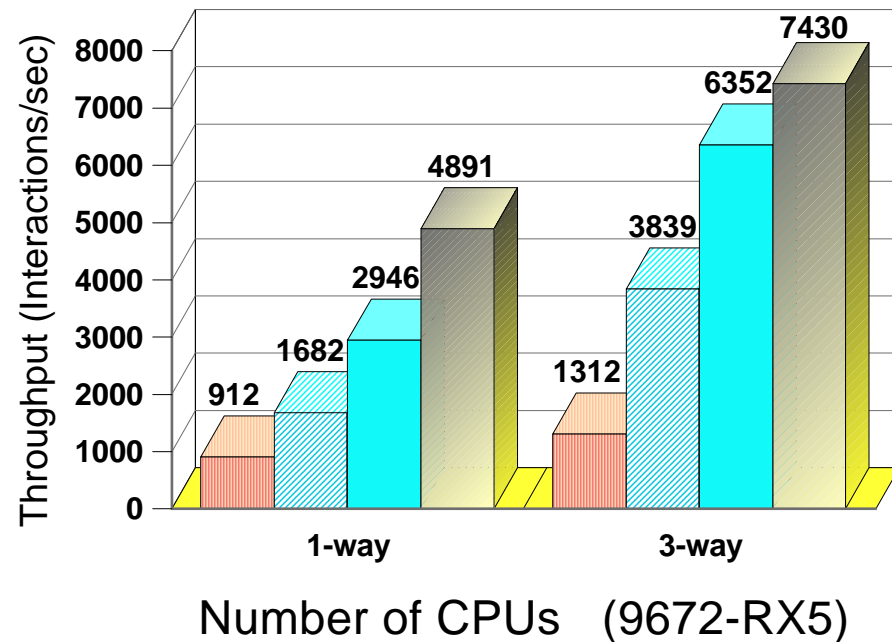
Legend:
- TCP/IP V3R2*
- CS OS/390 V2R4
- CS OS/390 V2R5
- CS OS/390 V2R6

*Note: TCP/IP V3R2 tests were with performance enhancement PTFs applied. The performance difference would have been even greater without them.

HFS Performance

- UNIX Byte File Expectations
  - ► Large numbers of small files opened and closed frequently
  - ► Metadata I/O written on every file access
  - ► Synch daemon for deferred/asynchronous write

- Original HFS Implementation
  - ► Hardened all File Writes
  - ► File I/O synchronous

- New HFS Implementation
  - ► Deferred/asynchronous I/O metadata and file data
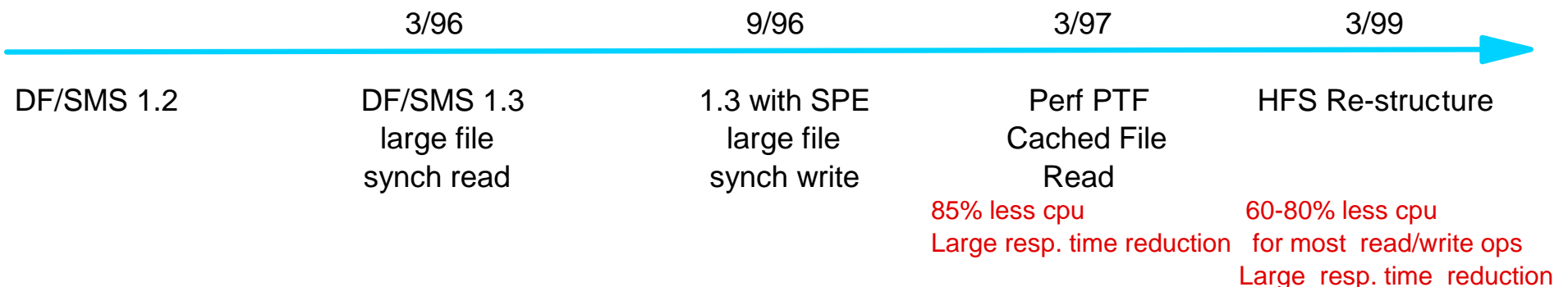 - ► Index manager cache for index pages
  - ► Large fixed file data buffers
  - ► Pathlength reductions

	3/96	9/96	3/97	3/99
DF/SMS 1.2	DF/SMS 1.3 large file synch read	1.3 with SPE large file synch write	Perf PTF Cached File Read	HFS Re-structure
			85% less cpu Large resp. time reduction	60-80% less cpu for most read/write ops Large resp. time reduction

HFS Performance Restructure

DFSMS1.5 GA- 3/99

Runs on OS/390 R5,R6,R7

WRITE - throughput

up to 40X increase

READ - throughput

up to 30X increase

HFS1.4

HFS1.5
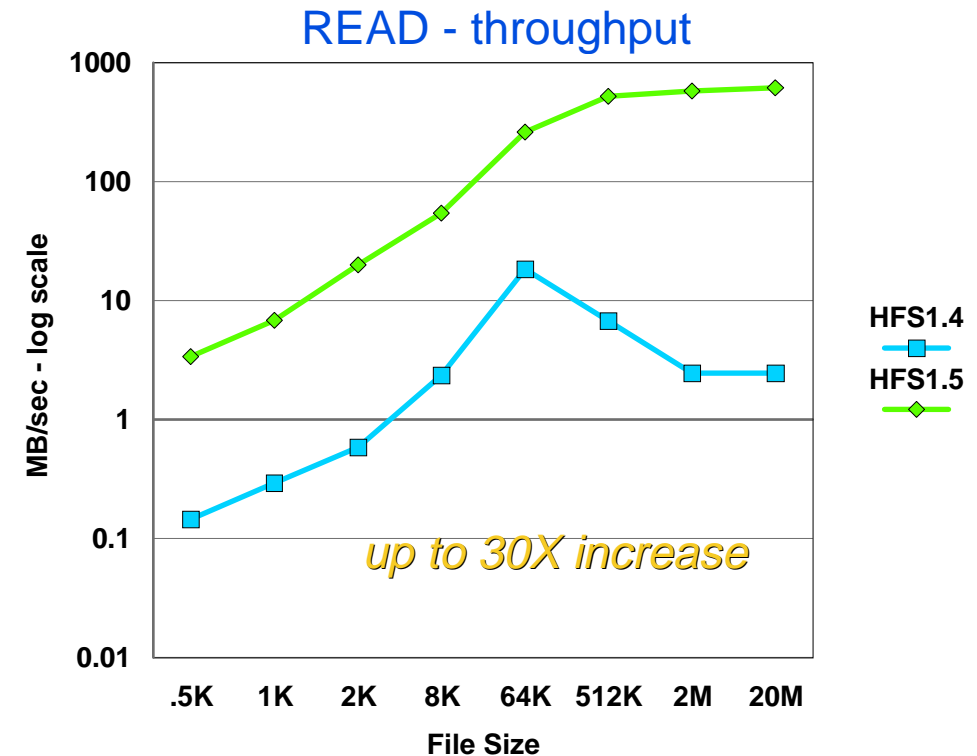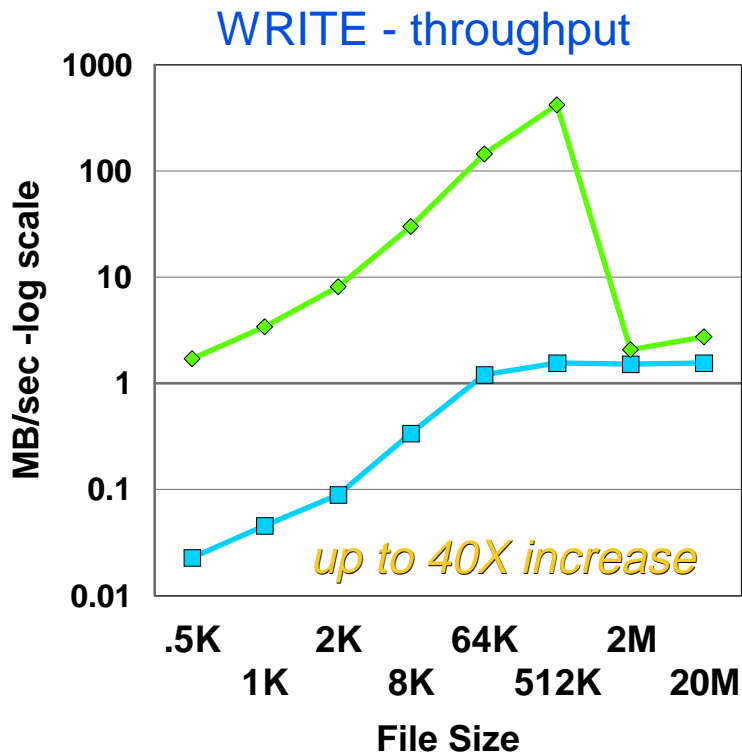
➤ Up to 85% reduction in CPU time

➤ Order of magnitude improvement in throughput for files < 512KB

➤ Optimized for multi-user access

➤ Make time improved by 75%

Make test

-7%

-75%

CPU time   Elapsed time

WRITE - response time

Miliseconds - log scale

File Size: .5K 1K 2K 8K 64K 512K 2M 20M 40M 80M 160M

READ - response time

Miliseconds-log scale

File Size: .5K 1K 2K 8K 64K 512K 2M 20M 40M 80M 160M

WRITE - cpu time

Miliseconds - log scale

File Size: .5K 1K 2K 8K 64K 512K 2M 20M 40M 80M 160M

Read - cpu time

Miliseconds - log scale

File Size: .5K 1K 2K 8K 64K 512K 2M 20M 40M 80M 160M
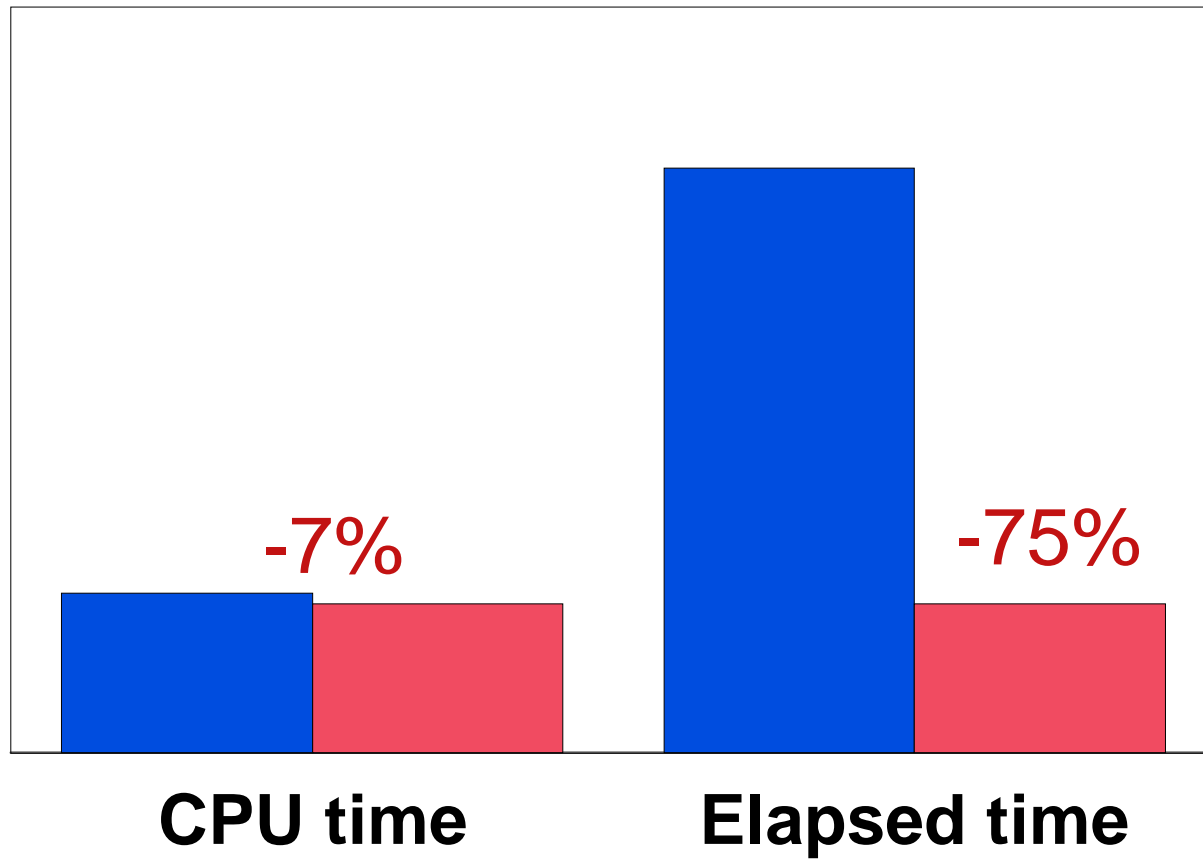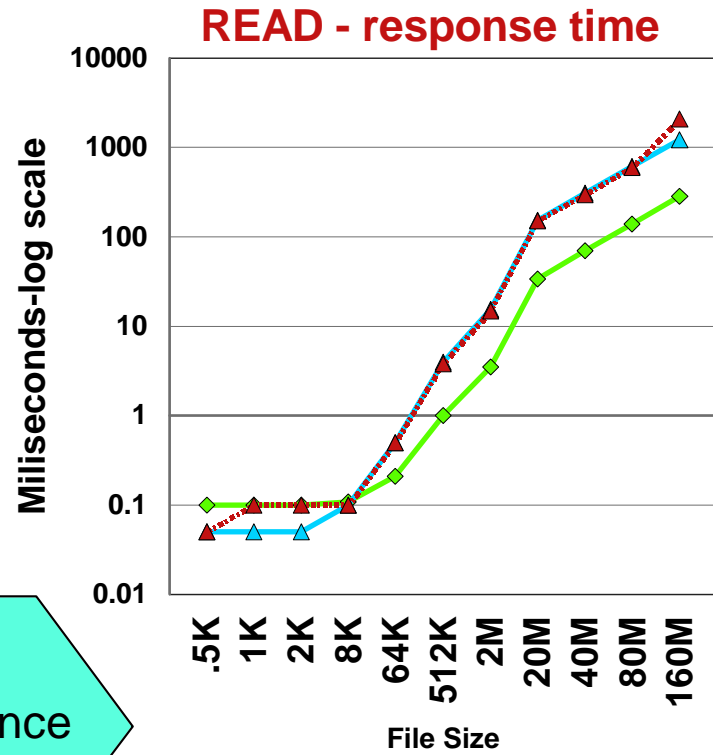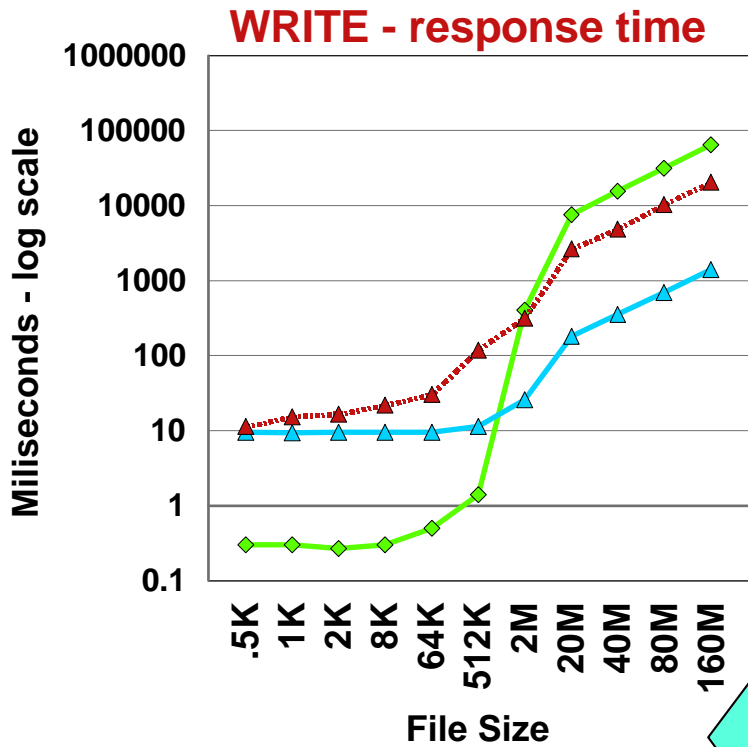
HFS Performance Restructure

Better

HFS1.5 (G5)
AIX4.3.2(S7A)
UE10000
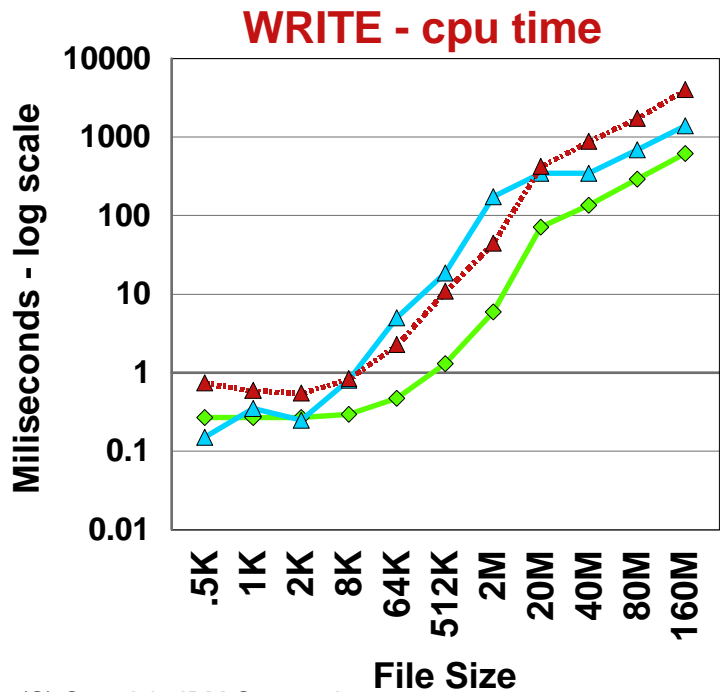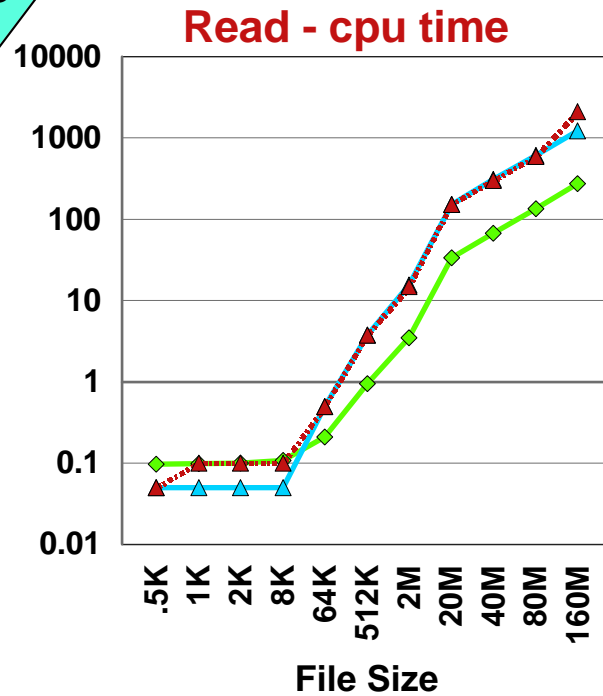
NFS Performance Improvements

IBM
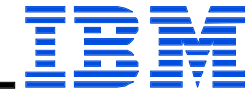
- **Pathlength reduction - R6,R7**
- **Improved scaleability - R6,R7**
- **Full multi-tasking   - R7 SPE**

Mount File A
Mount File B
Copy A,B

File A

TCP/IP

MVS data

NFS Server

HFS data

File B
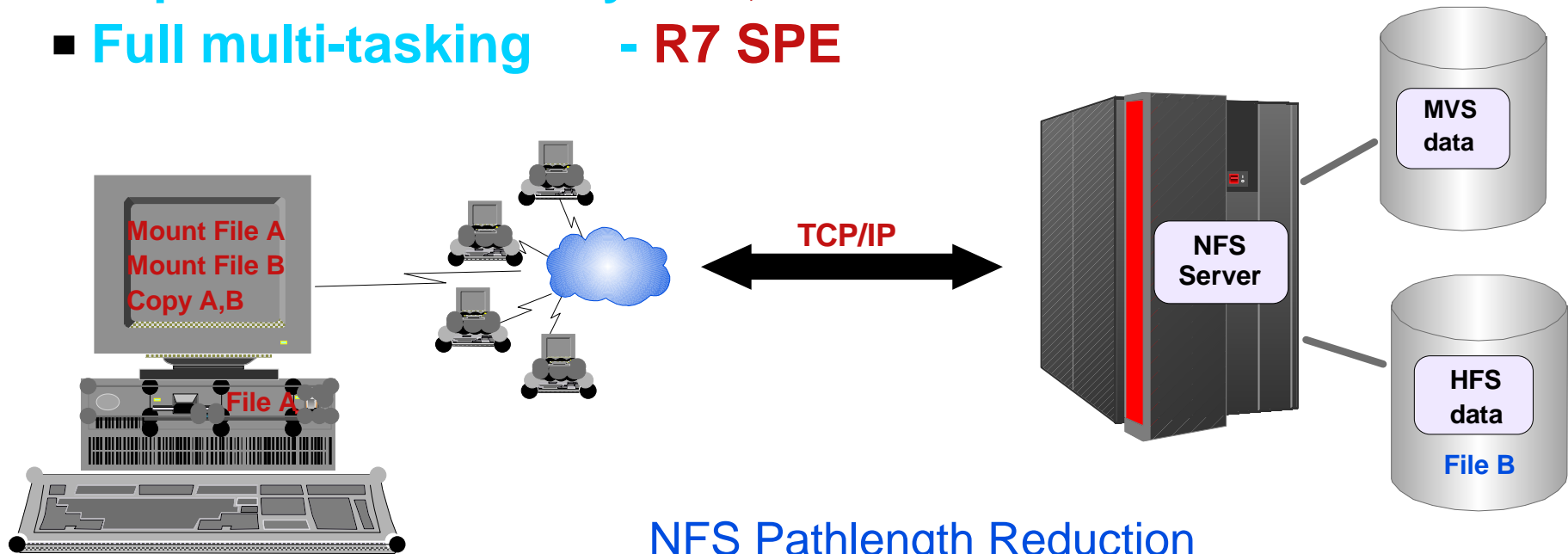
NFS Pathlength Reduction

OS/390 Rel	R3	R6	R7(SPE)
Pathlength	Base	80%	90%

NFS Throughput Gains - R6 only

File Size	48K	256k	20MB	100MB
Single Client Read	55%	70%	85%	30%
Single Client Write	70%	10%	5%	30%
Multi-client Read	*	*	180%	*
Multi-client Write	*	*	50%	*

* not measured

C/C++ Compiler Improvements

IBM

Run Time

| 96 | 97 | 98 | 99 |

OS/390 R2

C Compiler Run Time Improvements
15-30%

- Larger benefits to application intensive
code with multiple compile units
- Improved optimizer logic
- Enablement for new hardware
-5% improvement

OS/390 R3

C++ Compiler Run Time Improvements
15-30%

- Larger benefits to application intensive
code with multiple compile units
- Improved optimizer logic

OS/390 R4

Aggressive DLL loading
25% potential run time improvement

OS/390 R6

Binary Coded Decimal perf
enhancements for C++

OS/390 R6 PTF

IEEE Floating Point

OS/390 R10

Reduce linkage overhead
Up to 20% run time improvement
Smaller load modules

Compile Time

| 96 | 97 | 98 | 99 |

OS/390 R2
Pre-compiled headers
30%+

OS/390 R3 PTF
Metadata I/O
HFS fix

OS/390 R6
- C/C++ Compiler built
with IPA
- Improved utilities through
IPA build

OS/390 R7
New HFS

New Processor architecture on CMOS for C compiler

State of the Art Optimizer Technology Now!

Kernel/RTL Improvements

IBM

Application

Compiler 20%

HFS - File Access

Libcalls

X Syscalls

TCP/IP

Syscalls
10 - 60%

Syscall
Layer

MVS4.3.0=6000 inst
MVS5.1.0=3000 inst
MVS5.2.2=1000 inst
OS/390R2= 600 inst
OS/390R3= 500 inst
OS/390R4= 100 inst
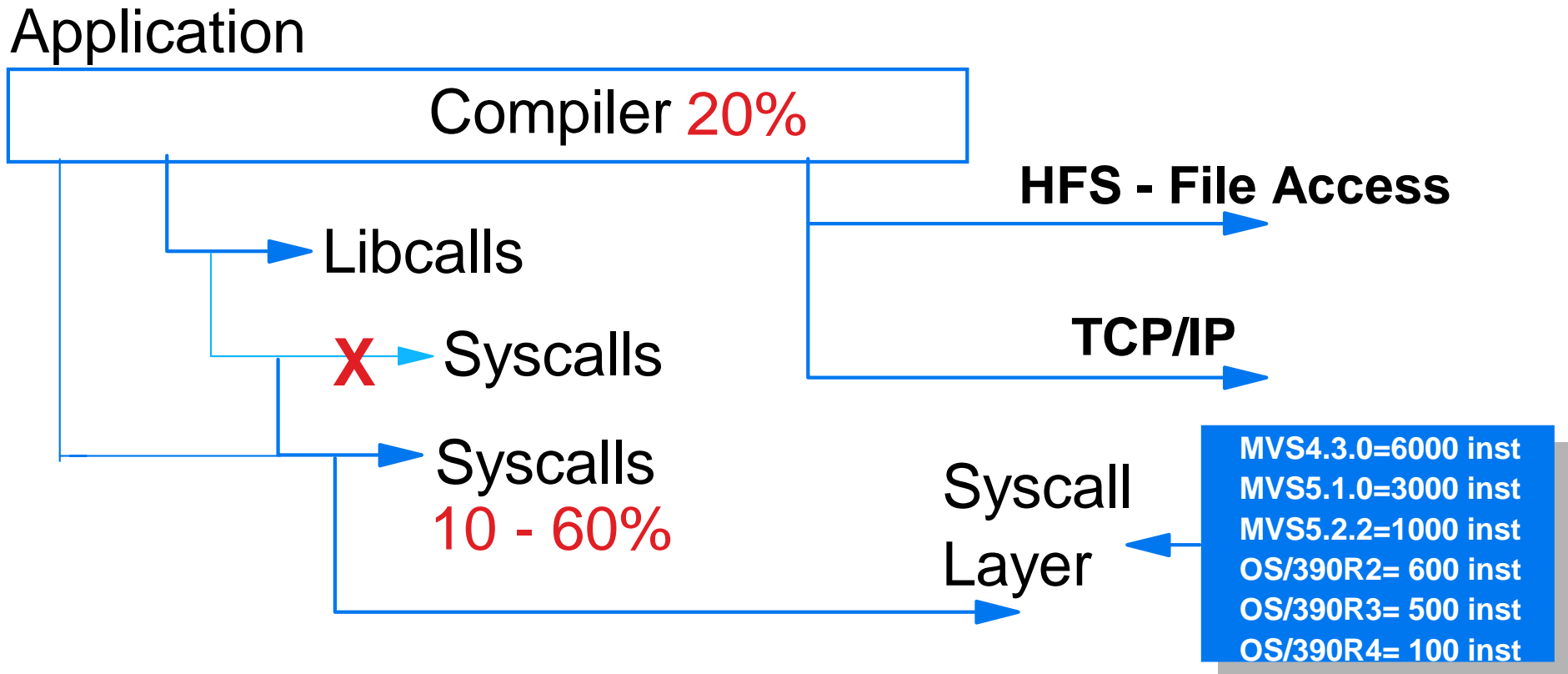
Focus on Pathlength Reduction for High Frequency Services

Other Kernel Performance Improvements

IBM
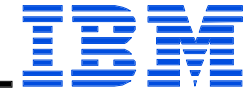
- Multiproc/Multiuser Support
  - Improves performance for DB2 Stored Procedures and Web
    - R6 with support back to R4
- Binary Semaphores using the PLO instruction
  - improved locking performance where GRS lock contention existed
    - R6 with support back to R4
- File system bundled syscalls - reduces pathlength
  - _open_stat()
  - _accept_and_recv() - benefits the web
  - send_file() - benefits the web and ftp type servers
    - R6
- Improved shared memory processing
  - greatly reduced ESQA and LSQA memory usage
    - R6
- Signaling performance enhancements
    - R6
- Optimized security checking in perf. sensative syscalls
  - Mesage queues, semaphores, File system(stat,open, lookups), ps cmd
    - R7
- Fast CGI Support for Web
    - R7

Other Kernel Perf Improvements cont...

- Asynch I/O for sockets
  - multiple socket read/writes w/o waiting for I/O complete
    - R7
- Shell and Utilities
  - recompiled with OPT(2) and IPA - highest levels of optimization
  - performance tune up: vi, find, pax, tar, cpio, compress, uncompress,
  -                       zcat, grep, ar, ls, du
    - R7
- Message Queues using the PLO instruction
  - potential performance benefit depending on GRS latch contention rate
    - R8
- Additional asynch I/O support
  - avoids the inefficiency of SELECT - benefits Lotus
    - R8
- Major improvement in DBX performance
  - use of local spawn instead of fork/exec
    - R8

For more detail see Don Ault's presentation:

OS/390 Unix System Services
Update for
OS/390 V2R6, V2R7, and V2R8

http://www.s390.ibm.com/unix/

Focus on DB2 Connectivity

- **Stored Procedures:**
  - Decreases costs with less network traffic
  - May be used by local or remote applications
  - Protects sensitive portions of your applications from unauthorized use
  - Offers more programming and application design choices
  - Allows access to non-SQL sources
  - Provides greater flexibility for change
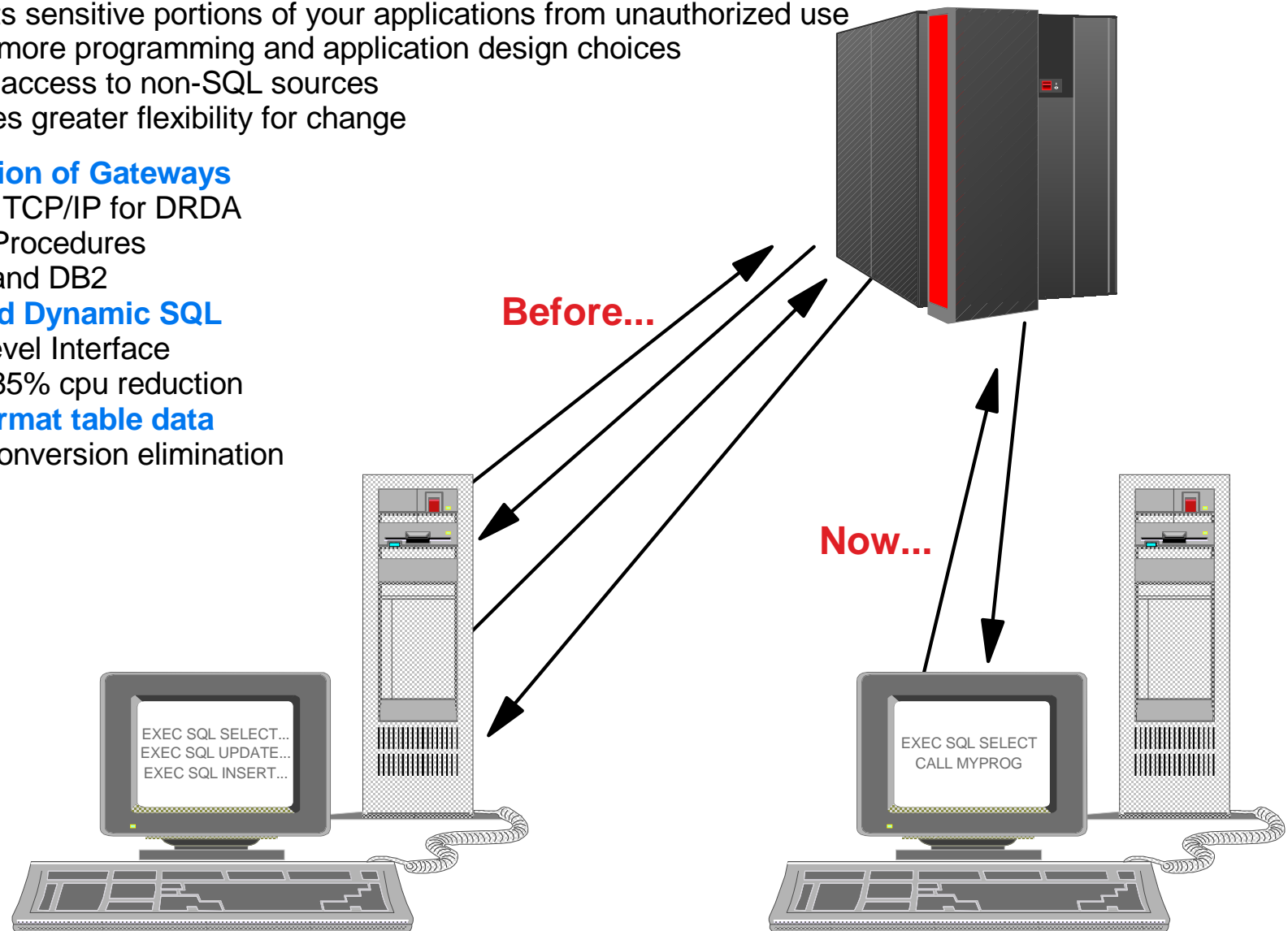
- **Elimination of Gateways**
  - Native TCP/IP for DRDA
  - Store Procedures
  - WLM and DB2
- **Improved Dynamic SQL**
  - Call Level Interface
  - Up to 85% cpu reduction
- **ASCII format table data**
  - Data conversion elimination

Before...

Now...

EXEC SQL SELECT...
EXEC SQL UPDATE...
EXEC SQL INSERT...
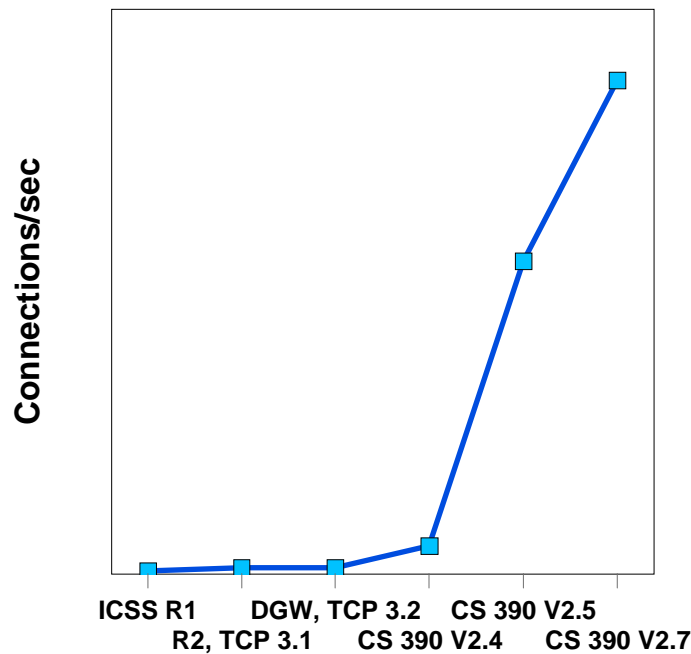
EXEC SQL SELECT
CALL MYPROG

Web Serving Performance

IBM

- Websphere Application Server, TCP/IP, Unix/kernel services, and Crypto HW improvements combine to provide large reductions in WEB serving pathlength

Webstone Throughput

Connections/sec (y-axis)

| ICSS R1 | DGW, TCP 3.2 | CS 390 V2.5 |
| R2, TCP 3.1 | CS 390 V2.4 | CS 390 V2.7 |

SSL Throughput

SSL Handshakes (Trans) Per Second (y-axis)

ICSS 2.2   DGW4.6.1*   DGW 5.0 *

SSL Web Server Version

* with S/390 Crypto HW
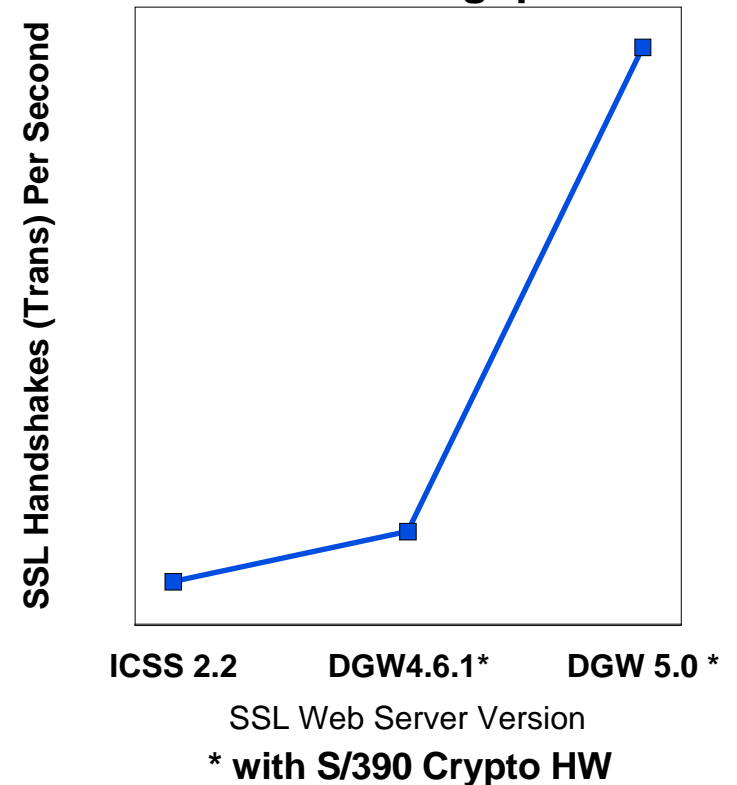
- **99% reduction in pathlength**
- **170X increase in connections/sec**

- **80% reduction in pathlength**
- **6X increase in connections/sec**

SPECWeb96 Published Results - 2/16/99

Rank	Company	System	Result	HTTP Version	# CPU
1	IBM Corporation	S/390 9672-YX6 G5	21591	IBM HTTP Server V5.1 for OS/390	10
2	Compaq	Compaq AlphaServer GS140 6/575	14263	Zeus 1.3.0	10
3	Hewlett-Packard	HP 9000 V2250	13811	Zeus 1.3.3	16
4	IBM	RS/6000 S70	12031	Zeus 1.3.3	12
5	Siemens	Primergy 870-40	10212	SWS 2.1 and SNCA	4
6	Siemens	Primergy 870-40	9199	SWS 2.1 and SNCA	4
7	Sun Microsystems Inc	Sun Enterprise 450	9115	SWS 2.1 and SNCA	4
8	IBM	RS/6000 S70 RS64-2	9081	Zeus 1.3.0	12
9	NCR Corporation	4400	9046	SWS 2.1 and SNCA	4
10	IBM	RS/6000 S70 RS64-2	8301	Zeus 1.3.0	12
11	IBM	Netfinity 7000 M10	7818	IBM HTTP Server 1.3.3	4
12	NCR Corporation	4400-4000-8090	7800	SWS 2.1 and SNCA	4
13	Hewlett-Packard	NetServer LXr 8000	7531	IIS 4.0 and SWC 1.1	4
14	Silicon Graphics	Origin 2000	7214	FastTrack 3.02	8
15	IBM	Netfinity 7000 M10	7180	IBM HTTP Server 1.3.3	4
16	IBM Austin	RS/6000 Model S70 RS64-2 Upgrade	7013	Zeus 1.3.0	12
17	Hewlett-Packard	NetServer LXr 8000	7006	Microsoft SWC 1.1 and IIS 4.0	4
18	Hewlett-Packard	9000/K580	6265	Zeus 1.3.3	4
19	Silicon Graphics	Origin 2000	6200	Zeus 1.3	8
20	IBM	Netfinity 7000 M10	5996	IBM HTTP Server 1.3.3	2

S/390 Results 50% better than previous BoB

Results from IDEAS International: www.ideasinternational.com
also check the SPEC Corp: www.specbench.org

Java Performance

- **JDK 1.1.4**
  - ‣ Single thread performance - 85% improvement
  - ‣ Multithread performance - 10X improvement
    - – available 5/98
- **JDK 1.1.6**
  - ‣ 2X increase in throughput for many applications vs JDK1.1.4
  - ‣ 100X increase in throughput for floating point intensive applications when using G5 IEEE FP hardware
    - – available 12/98
- **JDK 1.1.6p**
  - ‣ significant additional performance improvements
    - – available 3/99
- **JDK 1.1.8**
  - ‣ projected 2X performance improvement vs JDK1.1.6
    - – available summer '99

Watch this space !!!

OS/390 Unix Performance
Dramatic Improvements
Competitive Performance
More coming......

IBM

TCP/IP

WEB

NFS Server

MVS data

HFS data
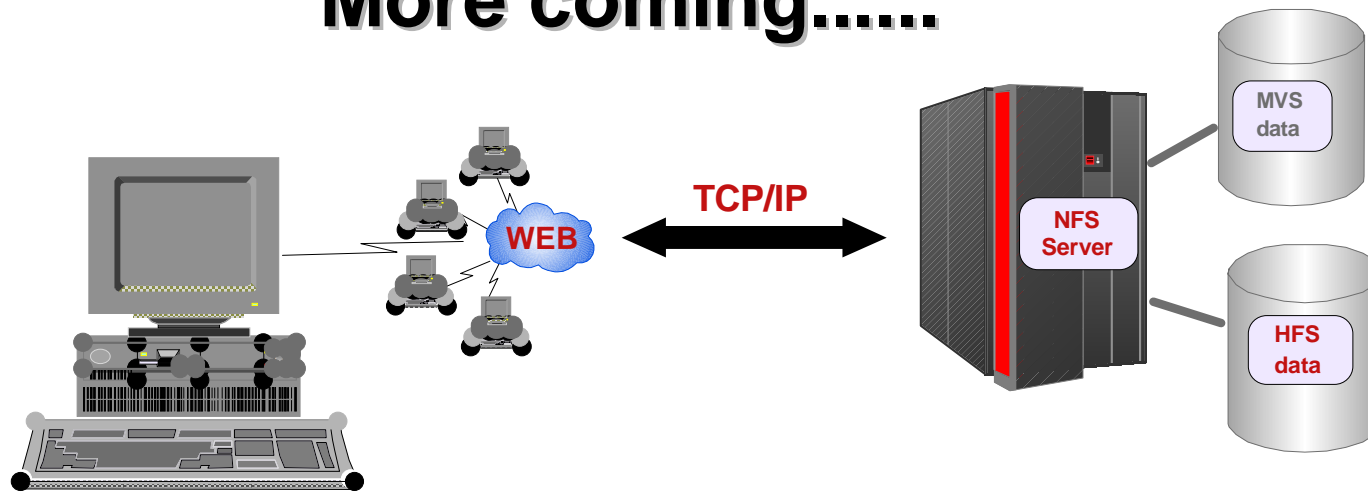
➤ **TCP/IP**	**2.3X-15X transactions /sec**
➤ **NFS**	**up to 2.8X multiple client file processing**
➤ **HFS**	**up to 40X file throughput processing**
➤ **Web serving**	**up to 170X connections/sec BoB SPECWeb96 results**
➤ **JAVA**	**up to 100X IEEE F.Pt. operations**