MariaDB performance on Linux on Power

Performance overview of MariaDB Enterprise on Linux on Power featuring the new IBM POWER8 technology

Axel Schwenke
MariaDB Corporation

Hari Reddy
IBM Systems ISV Enablement

September 2015

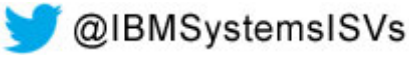

Table of contents

Abstract

This white paper describes the performance of MariaDB version 10.0.19 on IBM Power Systems servers featuring the new IBM POWER8 processor technology. The target audience is users and system integrators interested in using Linux on Power and MariaDB.

Introduction

The performance results in this paper demonstrate how the MariaDB application and IBM® Power Systems™ perform using the sysbench 0.5 benchmark suite. This white paper details the results on the IBM Power® System S822L server running MariaDB 10.0.19 on RHEV/RHEL with IBM PowerKVM 2.1.1 and sysbench 0.5* as the benchmark test suite.

The innovative design of the IBM POWER8™ processor technology is for very demanding, data-intensive applications making it well-suited for MariaDB customers.

Built with a processor designed for such data workloads, the IBM Power Systems design combines the computing power, memory bandwidth, and I/O in ways that are easier to consume and manage; building on strong resiliency, availability and security, demonstrated by:

- Computing power with 50% more cores and smart acceleration enabled by the Coherent Accelerator Processor Interface (CAPI).
- Massive memory with over twice the bandwidth of prior generation servers to process data faster and achieve greater speed and efficiency for transactional applications such as MariaDB.
- Systems that are easy to deploy and manage with open source technologies such as OpenStack, kernel-based virtual machine (KVM), simplified virtualization management, and flexible capabilities to drive rapid adoption and dramatically simplify IT consumption.
- Better cloud economics for scale-out infrastructures, with price-performance advantages and security to confidently move data-centric applications to the cloud.

The subsequent sections provide clear examples of the advantages of MariaDB, and its performance advantages over other platforms on Linux on Power.

Advantages of MariaDB on Power Systems

The key advantages of MariaDB running on IBM Power Systems include:

- **Performance**: Up to 2.1 times the system-level performance of Intel® x86-64 based systems (Haswell)
- **Economics**: Fewer cores are required compared to other platforms and this in turn means, reduced software and support costs

* For more information about sysbench refer to: https://launchpad.net/sysbench

MariaDB architecture

The MariaDB architecture consists of the following components and all ran in a single hardware partition along with the sysbench online transaction processing (OLTP) benchmark component used as a workload driver to MariaDB.

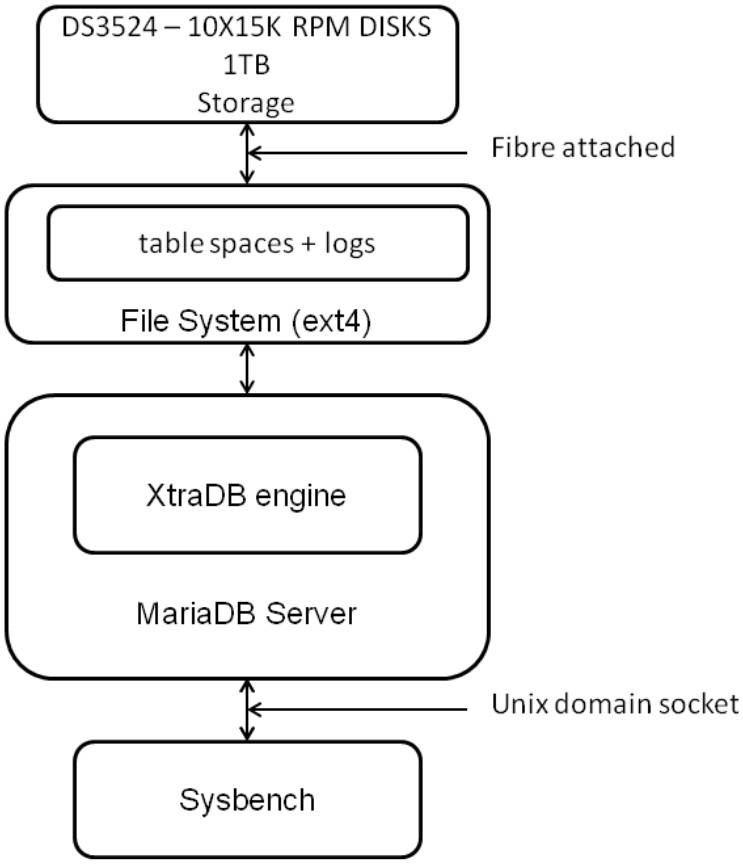

Figure 1: MariaDB and sysbench architecture and test topology

sysbench OLTP benchmark

sysbench is a very common benchmark in the MySQL / MariaDB world. It was written by a MySQL AB employee many years ago and is developed in public. sysbench contains many benchmarks, but only the use of the online transaction processing (OLTP) database benchmark is described in this paper.

There are two fundamentally different sysbench branches available. Most Linux distributions include the earlier sysbench 0.4.12 release. This has the following severe limitations:

- It can use only a single SQL table
- The OLTP workload is hardcoded
- There are no progress messages during the benchmark run

The newer sysbench 0.5 branch fixes all of those. In addition, it integrates a Lua interpreter, which allows to write the OLTP benchmark in Lua. This makes it straightforward to implement arbitrary benchmark workloads.

sysbench OLTP operation modes and query mix

The sysbench OLTP benchmark mimics the database workload from an OLTP application. That means transactions are rather short, tables are well-indexed, and the required behavior is high throughput and low latency.

The OLTP benchmark can be run in two modes: read-only and read/write. In the read-only mode, each database transaction does a certain number of point selects and range selects. In the read/write mode, it does additional updates and deletes and reinserts one row. The read/write mode is the default.

It is possible to tune sysbench OLTP to a different read/write ratio. In the case of the benchmarks mentioned later in this paper, a 90%:10% mix of read and write operations was used by requesting 28 point selects and two of each range select. This gives 36 read operations in addition to the default value of four write operations.

Note: The default query mix consists of 10 point selects and four different range selects in read-only mode. Additionally, two updates, one delete, and one insert in read/write mode. This results in a 14:4 = 78%:22% mix of read and write operations in the read/write mode.

MariaDB performance

The main measure of performance used in the sysbench benchmark study is the transactions per second (TPS) count. Additional metrics such as average response time per transaction, 99 percentile response time, and processor utilization are also measured. All these measures can help provide a comprehensive view of the behavior of the implementation of MariaDB on a given architecture.

Performance of IBM Power System S822L and HP ProLiant DL380 Gen9

Both IBM Power® System S822L and HP ProLiant DL380 Gen9 servers are equipped with two sockets each.

Read-only performance

IBM Power S822L with the guest partition running RHEV/RHEL 7.1 is over 2.6 times faster per core than HP ProLiant DL380 with the guest partition running RHEV/RHEL 7.1 for read-only mode of the sysbench benchmark (see Figure 2). System-level performance of IBM Power S822L is 2.15 times faster than HP ProLiant DL380 (see Figure 3).

Read/write performance

When considering the performance of these configurations on the read/write performance, which introduces the effects of a virtual hypervisor and its I/O characteristics, the Power S822L server continues to demonstrate strong performance, which is 2.2 times faster per core than the HP ProLiant DL380 server (see Figure 2). System-level performance of IBM Power S822L is 1.8 times faster than HP ProLiant DL380 (see Figure 3).

Scalability

For read-only workloads, one to two socket scalability of IBM Power S822L is 1.8 times faster compared to 1.5 times for HP ProLiant DL380. For read/write workloads also IBM Power S822L results in better scalability. Due to higher scalability and excellent speed-up over HP ProLiant DL380, performance of a single socket (10 cores) of IBM Power S822L exceed the performance of HP ProLiant DL380 configured with two sockets (24 cores). For more details, refer to “Appendix A: Benchmark run details”.

The conclusion drawn from this testing is that the Power S822L server is a capable and high-performing system that is well suited for MariaDB workloads and the growing data-intensive demands of its customers.

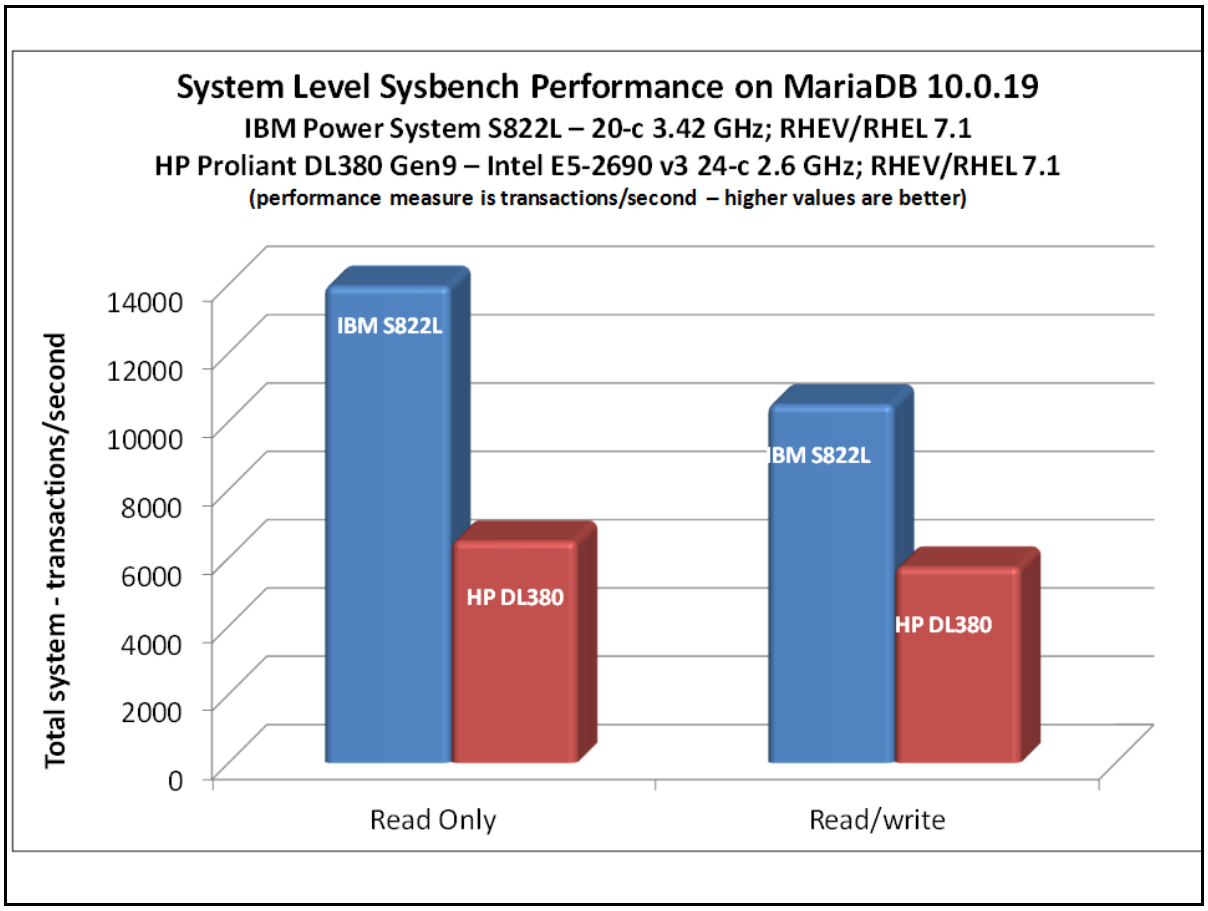

Figure 2: System-level MariaDB performance on IBM Power System S822L and HP Proliant DL380 Gen9

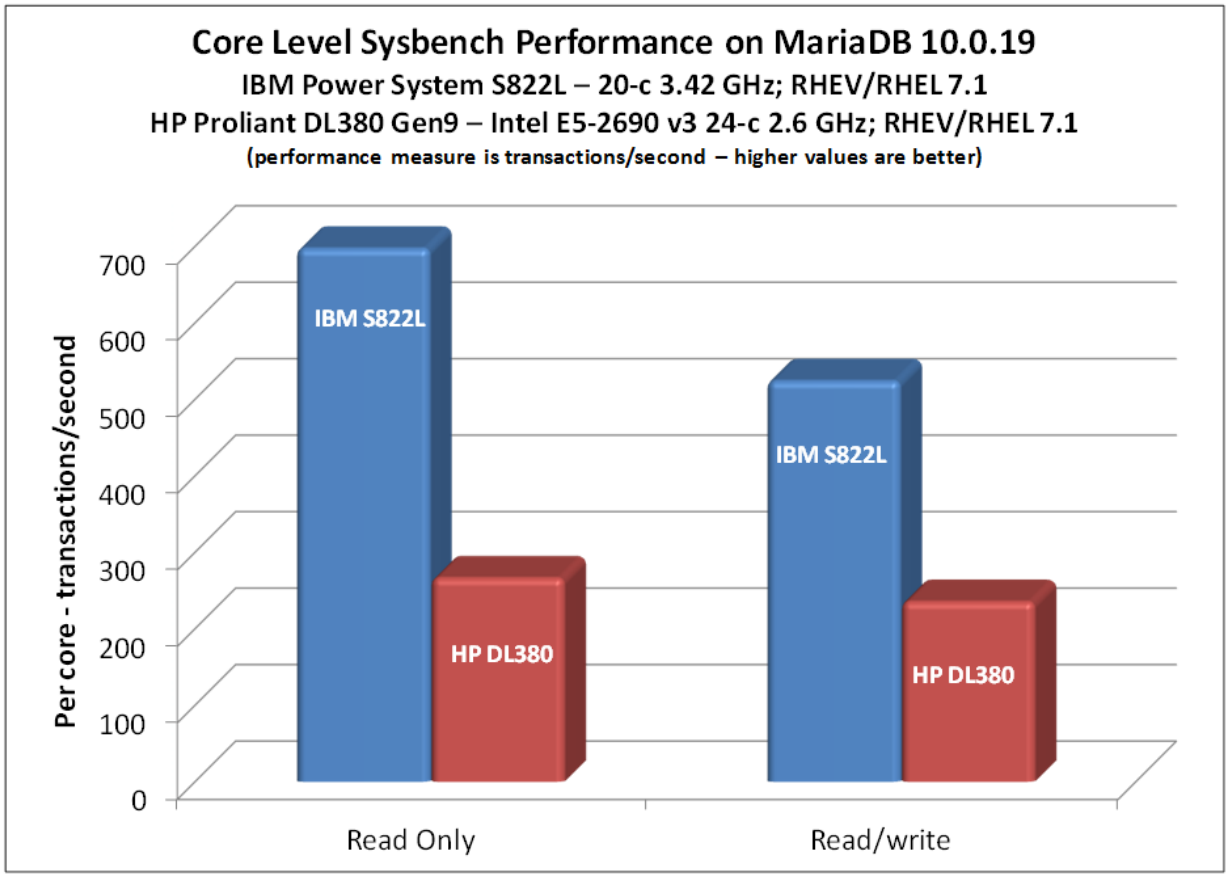

Figure 3: Per-core MariaDB performance on IBM Power System S822L and HP Proliant DL380 Gen9

Processor utilizations of both these systems were similar and range from 93% to 99%. The duration of the measurement was 600 seconds for read-only transactions and 1800 seconds for read/write transactions. Those times include the database warm-up phase also.

Refer to the “Tested configuration details” section for the description of the configurations used in this study. For more detailed benchmark data, refer to “Appendix A: Benchmark run details”.

Tested configuration details

In this section, the hardware and software configurations used for the testing are described.

Configuration of the IBM Power System S822L server with RHEV/RHEL

Hardware

- Processor: IBM POWER8, DCM, 20-core, 3.42 GHz, 96 MB L3 Cache, SMT 8
- Memory: 256 GB, sixteen 16 GB 1866 MHz DIMMs
- Storage: IBM System Storage® **DS3524** Express Dual Controller Storage System
- One logical unit number (LUN) of 10 15K RPM disks, RAID5
- Storage connectivity: QLogic Corp. ISP2532-8Gbs Fibre Channel to PCI Express (PCIe) HBA (rev 02)

Base software

- OS: RHEV 3.4.4; RHEL 7.1 guest; Linux® kernel: 3.13.0-35-generic
- Virtualization: PowerKVM 2.1.1
- File system: ext4

Application software

- sysbench 0.5 with scalability patches
- Application: MariaDB 10.0.19 (Power optimized)

For instructions to download and install MariaDB 10.0.19 (Power optimized) and download and build sysbench 0.5 with scalability patches, refer to *IBM Power Systems and MariaDB* web page at: **ibm.com**/partnerworld/page/stg_com_sys-ibm-power-systems-and-mariadb.

Configuration of HP ProLiant DL380 with RHEL/RHEV system

Hardware

- Processor: Intel Xeon E5-2690 v3 at 2.60GHz, 24-core, 30 MB L3 Cache, HT
- Memory: 384 GB, twenty-four 16 GB 1866 MHz DIMMs
- Storage: IBM System Storage **DS3524** Express Dual Controller Storage System
- One LUN of 24 15 K RPM disks, RAID5
- Storage connectivity: QLogic Corp. ISP2532-8Gbs Fibre Channel to PCIe HBA (rev 02)

Base software

- OS: RHEV/RHEL 7.1; Linux Kernel: 3.13.0-32-generic
- Virtualization: Ubuntu / KVM
- File system: ext4

Application software

- Application: MariaDB 10.0.19 (public release)
- sysbench 0.5 with scalability patches

For instructions to download and build sysbench 0.5 with scalability patches, refer to IBM and MariaDB web page at: **ibm.com**/partnerworld/page/stg_com_sys-ibm-power-systems-and-mariadb

Summary

The performance results in this paper demonstrated how the MariaDB application and IBM Power Systems servers perform using the sysbench benchmark suite.

The tested performance results of IBM Power S822L indicate that the Power S822L server is a capable and high-performing system that is well-suited for MariaDB workloads and the growing data-intensive demands of its customers. The results also demonstrate that IBM Power Systems running under a virtualized environment can perform very well and use resources efficiently.

These systems can deliver over two times the per-core performance of similarly configured virtualized servers on x86 systems. This can clearly be observed when comparing the total throughput of the one-socket IBM POWER® processor-based server to that of the two-socket Intel processor-based system. This directly translates into significant reduction in the total cost of ownership for the customer.

Appendix A: Benchmark run details

The following two architectures were used in the benchmark runs:

- IBM Power S822L server running RHEV/RHEL 7.1 guest
- Intel HP ProLiant DL380 system running RHEV/RHEL 7.1 guest

The following are the descriptions of the tables containing the results of the benchmark runs.

- Table 1: Results of running the sysbench benchmark in the read-only mode with varying thread counts
- Table 2: Results of running the sysbench benchmark under read/write mode with varying thread count
- Table 3: Peak performance for read-only runs from Table 1 for each architecture
- Table 4: Peak performance for read-only from Table 2 for each architecture

Note: Some software thread counts were not tested in middle range as they did not represent peak values for that specific hardware configuration and, therefore, did not offer additional technical meaning, and are represented with an asterisk (*) symbol in the following tables.

Details of runs used to determine peak read-only performance for each system (measure is transactions/sec - higher values are better)								
Threads	IBM Power S822L – RHEV/RHEL 7.1				HP ProLiant DL380 RHEV/RHEL 7.1			
	One socket		Two socket		One socket		Two socket	
	system	per core	system	per core	system	per core	system	per core
12		*		*	3312	276	3288	135
20	4485	449	6059	303	*	*		*
24		*		*	4319	360	5444	222
40	6133	613	8744	437	4285	357	6506	267
48		*		*	4334	361	6129	249
72		*		*	4317	360	5893	237
80	7762	776	11027	551			5955	236
96		*		*			5567	227
120	7744	774	11453	573				
160	7629	763	13721	686				
200			13966	698				
240			13859	693				

Table 1: MariaDB 10.0.19 read-only mode run details

Details of runs used to determine peak read/write performance for each system (in TPS – higher values are better)								
	IBM Power S822L – RHEV/RHEL 7.1				HP ProLiant DL380 RHEV/RHEL 7.1			
	One socket		Two socket		One socket		Two socket	
	One socket	per core	Two socket	per core	One socket	per core	Two socket	per core
Threads								
12		*		*	2895	241	2879	120
20	4209	421	5178	259		*		*
24		*		*	3853	321	4660	195
40	5347	535	7392	370	3791	316	5620	235
48		*		*	3816	318	5745	236
72		*		*	3811	318	5524	226
80	6674	667	9048	452			5427	217
96		*		*				210
120	6774	677	9543	477				
160	6716	672	10472	524				
200			10500	525				
240			10241	512				

Table 2: MariaDB 10.0.19 read/write mode run details

Peak performance summary – read-only mode				
	IBM Power S822L – RHEV/RHEL 7.1		HP ProLiant DL380 – RHEV/RHEL 7.1	
	One socket	Two sockets	One socket	Two sockets
Threads	80	200	48	40
Per system transaction rate (TPS)	7762	13966	4334	6506
Per core transaction rate (TPS)	776	698	361	267
Average response time (ms)	10	14	12	17
99% percentile response (ms)	12	48	13	60
Processor utilization (in percentage)	97	98	98	97

Table 3: Peak performance summary for read mode

Peak performance summary – read/write mode				
	IBM Power S822L – RHEV/RHEL 7.1		HP ProLiant DL380 – RHEV/RHEL 7.1	
	One socket	Two sockets	One socket	Two socket
Threads	120	200	24	48
Per system transaction rate (TPS)	6774	10500	3853	6506
Per core transaction rate (TPS)	677	525	321	236
Average response time (ms)	17	19	21	23
99% percentile response (ms)	42	90	83	101
Processor utilization (in percentage)	95	98	96	97

Table 4: Peak performance summary for read/write mode

Appendix B: Resources

The following websites provide useful references to supplement the information contained in this paper:

- IBM Power Systems and MariaDB
  ibm.com/partnerworld/page/stg_com_sys-ibm-power-systems-and-mariadb

- IBM Systems on PartnerWorld
  ibm.com/partnerworld/systems

- IBM Power Systems
  ibm.com/systems/in/power/?lnk=mhpr

- Linux on IBM Power Systems – resources
  ibm.com/systems/power/software/linux/resources.html

- MariaDB Foundation
  www.MariaDB.org/

- MariaDB Corporation official website
  www.MariaDB.com/

- IBM Power Systems Hardware Documentation
  http://publib.boulder.ibm.com/infocenter/powersys/v3r1m5/index.jsp

About the authors

Hari Reddy is a technical consultant in IBM Systems and Technology Group ISV Enablement organization. You can reach Hari at hnreddy@us.ibm.com.

Axel Schwenke is a performance analyst in MariaDB Corporation. You can reach Axel at axel@mariadb.com.

Acknowledgements

Thanks to the following people for their contributions to this paper**.**

Thomas Chen is a performance analyst in IBM Systems and Technology Group, ISV Enablement organization. You can reach Thomas at wtchen@us.ibm.com.

Jenifer Hopper is a performance analyst in IBM Systems and Technology Linux Technology Center. You can reach Jenifer at jhopper@us.ibm,com.

Mark Nellen is a program manager in IBM Systems and Technology Group, ISV Enablement organization. You can reach Mark at mnellen@us.ibm.com.

Basu Vaidyanathan is a performance analyst in IBM Systems and Technology Group. You can reach Basu at basu@us.ibm.com.

Sergey Vojtovich is a software developer in the MariaDB Foundation. You can reach Sergey reach at svoj@mariadb.org.

Trademarks and special notices

References in this document to IBM products or services do not imply that IBM intends to make them available in every country. Information is provided "AS IS" without warranty of any kind.

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. If these and other IBM trademarked terms are marked on their first occurrence in this information with a trademark symbol (® or ™), these symbols indicate U.S. registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at "Copyright and trademark information" at www.ibm.com/legal/copytrade.shtml.

Linux is a trademark of Linus Torvalds in the United States, other countries, or both.

Intel, Intel Inside (logos), MMX, and Pentium are trademarks of Intel Corporation in the United States, other countries, or both.

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

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.

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.

Any references in this information to non-IBM websites are provided for convenience only and do not in any manner serve as an endorsement of those websites. The materials at those websites are not part of the materials for this IBM product and use of those websites is at your own risk.