On demand computing solutions for life sciences
September 2004

IBM *e*server
BladeCenter JS20
1.6 GHz PPC 970
HMMER Benchmarks

Contents

This solution offers a range of popular open source applications and tools—all optimized for the reliable, low-cost @server BladeCenter.

Executive summary

Offering the power of a supercomputer at a fraction of the cost, the IBM @server® BladeCenter™ consolidates servers, storage and networking "blades" into a densely packaged and highly-scalable infrastructure. It balances high performance with physical integration and reduced power consumption, heating and system cost. Flexible and easy to manage, the IBM @server BladeCenter can utilize both Intel® Xeon® and IBM PowerPC® 970 processors, as well as Linux®, Microsoft® Windows® and AIX 5L™ V5.2 operating environments—enabling companies to deploy the best technology for any given application or workload.

The server used in this benchmark study is the IBM @server BladeCenter JS20, a two-way, 1.6 GHz PowerPC 970 processor derived from the IBM 64-bit POWER4™ processor. It is the first IBM PowerPC chip with AltiVec™ extensions, the Single Instruction Multiple Data (SIMD) operations that help accelerate data-intensive processing tasks.

The IBM @server BladeCenter JS20 is the core of the IBM @server BladeCenter for Bioinformatics Solution. Designed specifically for high-throughput environments, this solution offers a range of popular open source applications and tools—all optimized for the reliable, low-cost @server BladeCenter. HMMER, the application reported on here, is just one of the applications ported to and optimized for the @server BladeCenter JS20.

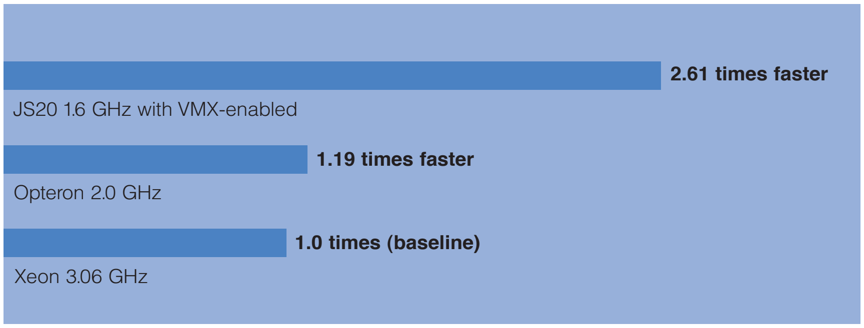
HMMER benchmark results

Profile hidden Markov models (profile HMMs) perform sensitive database searches to identify similarities at the protein primary structure level. HMMER is a freely distributable implementation of profile HMM software for protein sequence analysis and multiple sequence alignment to build a statistical model (a "hidden Markov model"). The model is used to query a sequence database to find (and/or align) additional homologues of the sequence family.

Testing was performed with the hmmsearch program from the HMMER suite Version 2.3.2 with a single execution thread, SWISS-PROT Release 41.9, 44x10E6 residues, Sequence PF02610 and length = 500. This version of HMMER exploits the power of the JS20 blade's VMX functionality. VMX is a Vector Multimedia eXtension to the PowerPC 970 processor architecture.

Testing was performed with HMMER Version 2.3.2 with a single execution thread, SWISS-PROT Release 41.9, 44x10E6 residues, Sequence PF02610 and length = 500. This version of HMMER capitalizes on power of the JS20 blade's VMX functionality.

The 1.6 GHz PowerPC 970 JS20 blade completed searches up to 2.61 times faster than a 3.06 GHz Xeon processor-based server. Such impressive outcomes can result in faster problem resolution (this is a small part of the research process), or in the ability to tackle larger and previously unsolvable problems.



HMMER was run on the following system configurations using compilers as specified:

System/ processor	Frequency (GHz)	No. of CPUs	RAM (GB)	Operating system	Compilers
JS20 PPC 970	1.6	2	2*	Linux/SUSE LINUX SLES 8 Kernel: 2.4.21-111-pSeries64 SMP	GNU GCC 3.4 (VMX enabled)
Xeon	3.06	2	4	Linux/Red Hat EL 3 WS Kernel: 2.4.21-9 SMP	GNU GCC 3.3**
Opteron	2.0	2	4	Linux/SUSE LINUX SLES 8 Kernel: 2.4-19 SMP	GNU GCC 3.3*

*This particular benchmark requires much less than 2GB of system RAM, so the difference in installed memory between the systems tested should not affect the test results.

**SSE-enabled HMMER code was not available at the time of the benchmark testing.

The benchmarks and values shown here were derived using specific, well-configured, development-level computer systems. Buyers should consult other sources of information to evaluate the performance of systems they are considering and ask for or conduct relevant application-oriented testing. For additional information about the benchmarks, values and systems tested, contact your local IBM office or IBM authorized reseller.

VMX potential

The VMX, or Vector Multimedia eXtension to the PowerPC 970 processor architecture, is the PowerPC implementation of the SIMD operation set, designed jointly by IBM and Motorola for Apple. The IBM VMX, Motorola AltiVec and Apple Velocity Engine™ are all names for the same instruction set.

Results using Vector Multimedia eXtension (VMX) to the PowerPC 970 processor architecture can provide a performance between four and six times faster at the instruction level.

Using VMX can produce performance that is between four and six times faster at the instruction level. That is because, unlike traditional Single Instruction Single Data (SISD) systems, SIMD-enabled systems can execute the same instruction on multiple data values in parallel. At the core of the VMX unit is a set of 32 x 128-bit vector registers that can store from 4 x 32-bit to 16 x 8-bit data values. VMX also introduces a set of 162 instructions. The instructions can operate on from 4 to 16 data values at once—compared with traditional SIMD operations that utilize only one data value—helping to deliver a significant boost in performance.

VMX offers the potential for dramatic performance gains. Applications in which the algorithms and data structures align with the characteristics of the vector unit can realize significant performance speed-ups. Applications in which 10 percent of the code consumes 80 percent of the CPU in heavy computational and data loads can exploit data parallelism, and if the computationally-intensive section is also SIMDizable, can also see jumps in performance.

VMX is ideal for matrix and array-oriented codes. Application categories that can realize significant advantage from VMX include bioinformatics, molecular dynamics, seismic processing, signal processing, 3D graphics and applications requiring a Fast Fourier Transform function.

IBM is exploring additional application categories for VMX beyond traditional high performance computing, including basic data movement and manipulation. It is important to note that VMX hardware supports single precision floating-point instructions only. In some situations, it can be as simple as replacing a math library with a VMX-enabled one. In other situations, some degree of code modification may be required.

IBM is exploring other applications categories beyond traditional high performance computing, including basic data movement and manipulation.

VMX has proven to be superior to other SIMD extensions, such as Intel MMX/SSE/SSE2 running on x86 processors. Specifically, VMX boasts maximum throughput of 8 Flops/cycle on the PowerPC 970 compared with only 4 Flops/cycle on X86 derivatives in 32-bit mode. VMX does not interfere with floating point registers and offers 32 vector registers compared with 8 for X86 SIMD. While VMX is just another short vector/SIMD extension, it boasts significant advantages over MMX/SSEx.

The SUSE™ and Red Hat Linux distributions will include vector type definitions and support for VMX in the future. Developers can use these definitions to write VMX-optimized code. Compilers will then convert the high-level code to machine-level binaries.

Resources and links	
IBM @server BladeCenter Solution for Bioinformatics	**http://www-1.ibm.com/industries/lifesciences/doc/content/solution/1012978221.html** or e-mail **LS@us.ibm.com**, SUBJ: "**Bio Blades**"
HMMER	**http://hmmer.wustl.edu/**
HMMER VMX-enabled binaries	**http://www.ciri.upc.es/cela_pblade/HMMER.htm**
HMMER benchmark results	**http://www.pc.ibm.com/ww/eserver/bladecenter/benchmarks/index.html**
PowerPC 970	**http://www-306.ibm.com/chips/techlib/techlib.nsf/products/PowerPC_970_and_970FX_Microprocessors**

This document was developed for IBM offerings in the United States as of the date of publication. IBM may not make these offerings available in other countries, and the information is subject to change without notice. Consult your local IBM Healthcare and Life Sciences representative for information about offerings in your area.

- Information in this document concerning non-IBM products was obtained from the suppliers of these products or other public sources. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.
- IBM may have patents or pending patent applications covering subject matter in this document. The furnishing of this document does not give you any license to these patents. Send license inquires, in writing, to IBM Director of Licensing, IBM Corporation, New Castle Drive, Armonk, NY 10504-1785 USA.
- All statements regarding IBM future direction and intent are subject to change or withdrawal without notice and represent goals and objectives only.
- The information contained in this document has not been submitted to any formal IBM test and is provided "AS IS" with no warranties or guarantees either expressed or implied.
- All examples cited or described in this document are presented as illustrations of the manner in which some IBM products can be used and the results that may be achieved. Actual environmental costs and performance characteristics will vary depending on individual client configurations and conditions.
- IBM is not responsible for printing errors in this document that result in pricing or information inaccuracies.
- IBM hardware products are manufactured from new parts, or new and serviceable used parts. Regardless, our warranty terms apply.
- Any performance data contained in this document was determined in a controlled environment. Actual results may vary significantly and are dependent on many factors including system hardware configuration and software design and configuration. Some measurements quoted in this document may have been made on development-level systems. There is no guarantee these measurements will be the same on generally available systems. Some measurements quoted in this document may have been estimated through extrapolation. Users of this document should verify the applicable data for their specific environment.

© Copyright IBM Corporation 2002
IBM Corporation
Route 100
Somers, NY 10589
U.S.A.

Printed in the United States of America
9-04
All Rights Reserved

IBM, the IBM logo, @server, AIX 5L, BladeCenter and BladeCenter for Bioinformatics, POWER and PowerPC are trademarks or registered trademarks of IBM Corporation in the United States, other countries, or both. A full list of U.S. trademarks owned by IBM may be found at: ibm.com/legal/copytrade.shtml.

AMD Opteron is a trademark of Advanced Micro Devices, Inc.

AtliVec is a trademark of Motorola.

Intel, Itanium and Pentium are registered trademarks and Xeon and MMX are trademarks of Intel Corporation in the United States other countries, or both.

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

Microsoft and Windows are registered trademarks of Microsoft Corporation in the United States, other countries, or both.

SuSE is a trademark of SuSE Linux AG in the United States, other countries or both.

Velocity Engine is a trademark of Apple in the United States, other countries or both.

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

GH10-4091-00