

Deploying CoolThreads™ Technology for HPC Workloads

Performance, scalability, and high efficiency

Getting more answers more quickly, or being able to ask more and bigger questions, can make all of the difference for both scientific and commercial organizations. The high-performance computing (HPC) infrastructure used to answer these important questions must address a range of demands, including:

- Floating-point and integer performance
- Scalability to handle very large computational problems and large datasets (both in storage and in memory)
- Multiple on-going projects at any one time
- Programmability for large and complex applications

scalability, and optimal processor utilization. Best of all, these systems take the form of general-purpose computing platforms with a straightforward and open programming model.

Innovative CoolThreads technology

As applications and systems have grown in size and complexity, memory latency has emerged as the single largest impediment to good performance across a wide range of applications. Conventional processors with complex superscaler designs have not provided much relief, even when deployed in multicore versions.

Highlights

- Innovative UltraSPARC® T2 and T2 Plus processors with CoolThreads™ technology effectively hide memory latency so that processors stay busy performing computations.
- Memory-intensive integer applications such as string matching scale linearly as processors are added, handily beating competitors' laboriously hand-coded implementations.
- With up to eight floating-point units per processor, applications such as linear algebra also scale well with optimal processor utilization.
- Sun™ Logical Domains (LDom)s allow the partitioning and virtualization of the considerable computational resources of Sun's chip multithreading (CMT) systems.

Sun SPARC® Enterprise systems utilizing CoolThreads technology can provide extremely high processor utilization and efficiency for memory-intensive integer and floating-point workloads

Innovative, balanced, and efficient systems design is ultimately essential for HPC infrastructure. Even the most demanding floating-point intensive workloads must also perform a range of integer, memory, and I/O operations. Beyond mere power consumption, efficiency means keeping processing elements as busy as possible doing productive work.

Systems based on the UltraSPARC T2 and T2 Plus processors are demonstrating new applicability to scientific and other HPC workloads. Innovative processor and system design lets these systems operate extremely efficiently, providing performance,

Sun's UltraSPARC T2 and T2 Plus processors with CoolThreads™ technology employ proven chip multithreading. With support for up to eight cores and up to 64 threads per processor, this technology yields considerable advantages for HPC applications, including:

- The ability to effectively hide memory latency
- Rapid context switching between a large number of threads
- An open and general-purpose programming model
- The ability to fully utilize processing capabilities for real-world applications

Chip multithreading systems for HPC

Each UltraSPARC T2 and T2 Plus processor provides considerable computational resources. Up to eight cores are provided in each processor, with each core able to rapidly switch between up to eight threads (up to 64 threads per processor) as they block for memory access (Figure 1). In addition, each core provides two integer execution units and a floating-point unit, so that a single UltraSPARC T2 or T2 Plus processor core is capable of executing up to two threads at a time (up to 16 threads per processor).

Sun provides a broad family of CMT systems that use these processors, all with large memory support and significant I/O capabilities:

- Sun SPARC® Enterprise T5120 and T5220 servers support an UltraSPARC T2 processor, up to 64 threads, and up to 64 GB of memory
- Sun SPARC Enterprise T5140 and T5240 servers support up to 2 UltraSPARC T2 Plus processors, up to 128 threads, and up to 128 GB of memory
- Sun SPARC Enterprise T5440 servers support up to four UltraSPARC T2 Plus processors, up to 256 threads, and up to 512 GB of memory

In addition to rackmount configurations, Sun CMT processing capabilities are also available in modular (blade) form factors through the Sun Blade™ 6000 modular system to facilitate dense and scalable high-performance computing deployments.

An ideal architecture for string matching

Real-time string searching is gaining in importance, with applications from viral intrusion detection to life sciences applications such as Basic Local Alignment Search Tool (BLAST). In searching for patterns, string matching algorithms represent an interesting challenge because they must

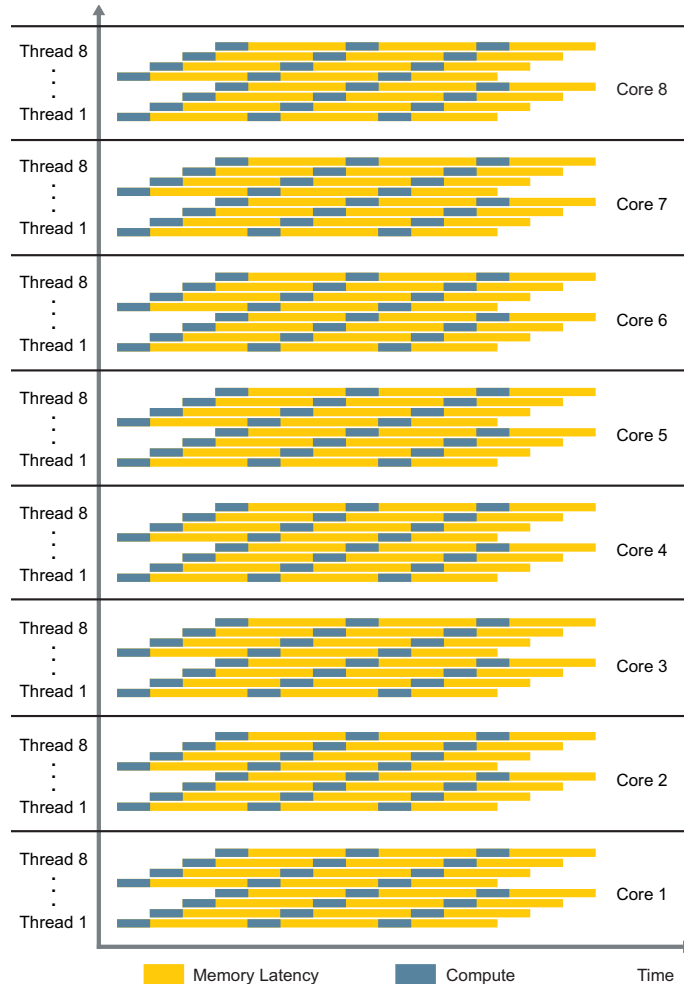


Figure 1: UltraSPARC T2 and T2 Plus processors hide memory latency so that processors stay busy doing real work.

traverse a large data structure in memory, sharing the data structure among multiple threads.

Much has been made regarding IBM's recent efforts to optimize code running on its Cell Broadband Engine (Cell/BE)¹. Unfortunately, with a non-uniform memory architecture, and a lack of per-core cache, the Cell/BE represents a highly-specific special-purpose architecture that requires considerable hand tuning and optimization to enhance

performance. Even with these efforts, The performance realized by these much-publicized efforts was ultimately lackluster, indicating that the Cell/BE is a rather poor choice for string matching infrastructure.

In contrast, Sun's CMT architecture as implemented in the UltraSPARC T2 and T2 Plus processors is ideal for string matching. With a shared 4 MB L2 cache, the UltraSPARC T2 and T2 Plus processors can fit even very large data structures entirely in cache.

1. See IEEE Computer, Volume 41, Number 4, pp. 42-50, April 2008.

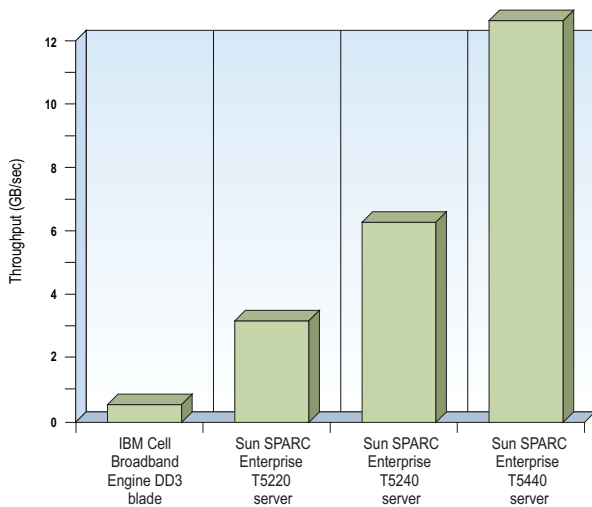


Figure 2: Sun SPARC Enterprise T5440 servers provide over 26 times the throughput of an IBM Cell Broadband Engine DD3 blade for integer-intensive and memory-intensive string matching.

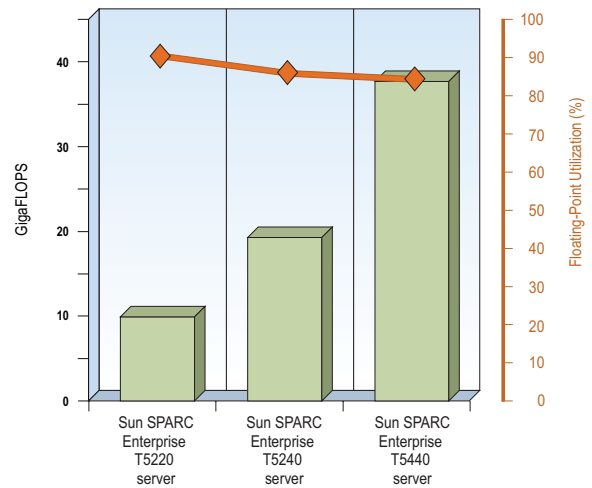


Figure 3: Sun SPARC Enterprise T5220, T5240, and T5440 servers yield near linear scalability for floating-point intensive linear algebra (matrix multiplication) and 85-90% floating point utilization.

In addition, because each processor core has uniform access to memory through an on-chip crossbar switch, memory is fewer cycles away from processing elements than with other architectures, reducing latency. Most importantly, efficient switching between threads allows the processor to rapidly move to execute new thread as an individual thread stalls, maximizing both processor utilization and performance.

To evaluate the strengths of CMT architecture for string matching, a straightforward implementation of the Aho-Corasick string-matching algorithm was tested across a range of Sun CMT systems. This memory-intensive dictionary-matching application is essentially a measure of how well a system can follow pointers through a graph in a large data structure. Sun compared the results of an Aho-Corasick string search using a range of systems based on UltraSPARC T2 and T2 Plus processors, to the published results of a similar search using an IBM Cell Broadband Engine (Cell/BE) DD3 Blade¹.

To test the performance of their processor, IBM used a 4.4 MB variant of the King James version of The Bible, using a dictionary of the 20,000 most used words in the English language (average word length of 7.59 characters). To approximate the dictionary and Bible that IBM used, Sun used a dictionary of 25,143 English words² (average word length: 7.2 characters) and a 4.6 MB variant of the King James version of the Bible.

Figure 2 graphs the throughput in GB/sec for the IBM system along with one-processor, two-processor, and four-processor CMT systems. Even the single-processor Sun SPARC Enterprise T5220 server bested the Cell/BE equipped system by over a factor of six. The Sun SPARC Enterprise T5440 server with four UltraSPARC T2 Plus processors provided over 26.7 times the throughput of the carefully hand-tuned IBM Cell Broadband Engine DD3 blade (16 cores).

These significant results demonstrate the importance of the balanced design of Sun's CMT systems. Linear scalability was observed

across one-processor, two-processor, and four-processor UltraSPARC T2 and T2 Plus based servers. This scalability indicates that the memory systems (and memory latency in particular) does not limit performance as additional processors and cores are added.

Additionally, 99.9% CPU utilization (as measured by the *top(1)* program) was measured on the Sun SPARC Enterprise T5440 server, using only 1,530 watts. These numbers indicate that the system is performing extremely efficiently, making the most of its computational resources and power budget.

Scalable performance in linear algebra

Integer performance and strong scalability may not be a surprise for CMT architectures. However, the UltraSPARC T2 and T2 Plus processors also offer considerable floating-point processing capabilities. In addition to two integer execution units, each of the up to eight cores features a dedicated floating-point processor.

1. IBM Cell Broadband Engine (Cell/BE) DD3 Blade (2x 3.2 GHz Broadband Engine, two chips, 16 cores, 0.475 GB/second)
 2. The OpenSolaris™ file cvs.opensolaris.org/source/xref/onnv/onnv-gate/usr/src/cmd/spell/list

Linear algebra, and specifically matrix multiplication represents a valid metric for evaluating memory-intensive floating-point application performance. While the core application is fundamentally arithmetic, large matrices must be able to fit into memory while being operated on by multiple threads.

Sun SPARC Enterprise T5220, T5240, and T5440 servers were evaluated in Sun's testing. Execution times were compared for matrix-matrix multiplication as required for one iteration of the NAB (Nucleic Acid Builder) Newton-Raphson minimization of the 6,380-atom 1AKD molecule (available from the Brookhaven PDB databank). Figure 3 graphs the floating-point operations (in gigaFLOPS) achieved by the systems. The following multiplications were performed:

- A 6,380 by 6,380 matrix times a 6,380 x 6,380 matrix
- Two of a 6,380 by 3 x 6,380 matrix times a 3 x 6,380 by 6,380 matrix

Again, the testing exhibited near-linear scalability across one-processor, two-processor, and four-processor CMT systems respectively. These results convincingly demonstrate that memory contention and latency has minimal effect on this memory-intensive and floating-point intensive code.

Similar to the string matching results, processor utilization as measured by the *top(1)* program reached 99.8%, while consuming only 1,690 watts. Utilization of the floating-point processor was also extremely high. The right-hand axis of Figure 3 indicates that the servers realized very high floating-point processor utilization — with the systems achieving between 85% and 90% utilization of the theoretical peak performance of the installed processors.

These data highlight the ability of the CMT processor architecture to keep both integer and floating-point processors optimally busy doing real work for the application.

Virtualizing CMT resources with Sun Logical Domains and Solaris™ Containers

Computational infrastructure needs to be flexible so that resources can be shared across multiple groups, and agile so that resources can be adjusted to meet the needs of a range of diverse application demands. HPC applications have unique requirements, and not all applications can scale to take advantage of large numbers of processors, cores, and threads. Virtualization technology can help by allowing organizations to consolidate multiple slower systems onto more capable platforms.

To provide enhanced virtualization capabilities, UltraSPARC T2 and T2 Plus processors offer a fully-multithreaded hypervisor — a small firmware layer that provides a stable virtual-machine architecture that is tightly integrated with the processor. Sun provides multithreading at every layer of the technology stack (Figure 4).

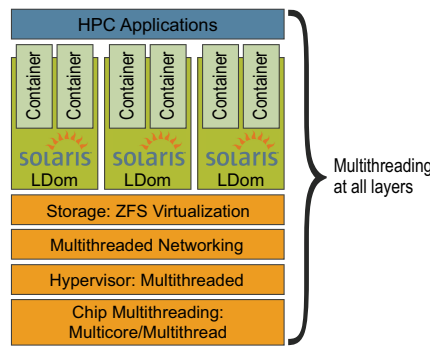


Figure 4: Parallelization and virtualization are provided at every level of the CMT technology stack

Multithreading is crucial, since the hypervisor interacts directly with the underlying chip-multithreaded processor. This architecture is able to context switch between multiple threads in a single core, a task that would require additional software and considerable overhead in competing architectures.

Supported in all Sun servers utilizing CMT technology, LDom provide full virtual machines that run an independent operating system instance. Available at no cost in an open-source distribution, LDom contain virtualized CPU, memory, I/O, storage, console, and cryptographic devices. In an HPC context, LDom can be used to partition CMT systems to fit multiple simultaneous applications. LDom can also be rapidly re-configured to suite new applications or changing workloads.

Providing virtualization at the operating system level, Solaris™ Containers consist of a group of technologies that work together to efficiently manage system resources, virtualize the environment, and provide a complete, isolated, and secure runtime environment. Together Solaris Containers and Sun Logical Domains allow CMT systems to be partitioned and virtualized to suit the needs of a wide range of HPC applications.

Together these virtualization technologies help make CMT systems flexible and agile to accommodate a range of HPC workloads. With ever-larger datasets, the clear advantages of CMT architecture for handling memory-intensive workloads presents a clear advantage. Though organizations should always evaluate their own applications, many integer and floating-point intensive workloads are well suited to running on CMT systems.