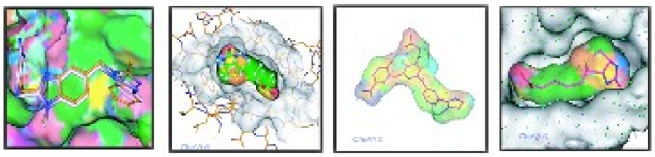


eHiTS® on the Cell B./E. ™
Revolutionary Hardware Technology Opens New Frontiers in Molecular Modeling

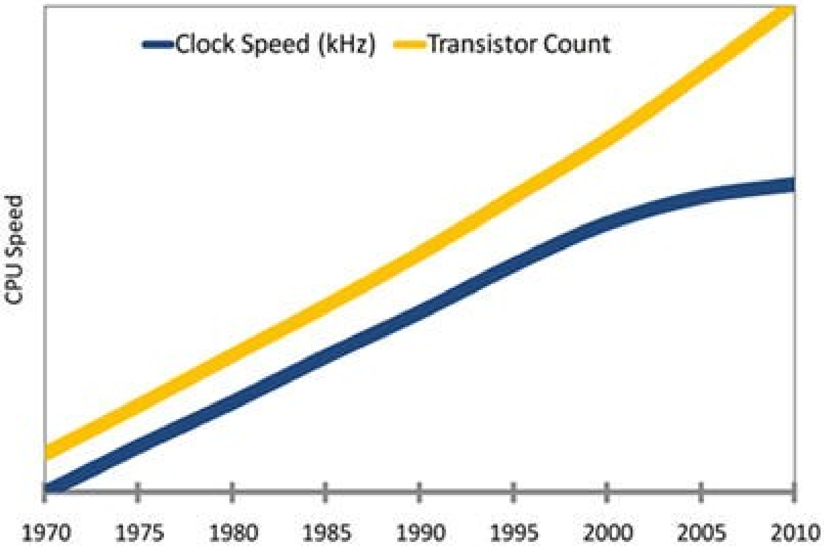
Table of Contents

1. Introduction

Bigger is better. More hard drive space, more RAM, more GHz. We have become obsessed with the components and functions crammed into our electronic devices whether they be cameras, phones, music players or, of course, computers. While the general populace should not necessarily be concerned with the latest and greatest performance from their electronics, in our world of high performance computing algorithms we do have to care about the technologies. Our focus for our users has been to provide algorithms and visualization tools facilitating the best performance with the greatest ease of use. We believe that our algorithms produce the highest accuracy in performance despite a related trade-off in time. With an intention to deliver the best performance in the shortest time we have investigated opportunities available through new hardware technologies, specifically the investigation of the new Cell Broadband Engine™ (Cell B./E.). While you may not be aware of the Cell B. E. you *are* likely aware of the Playstation PS3 game station. The performance of this highly successful platform offers game players access to improved experiences via faster calculations and visualizations specifically via improved physics engines. The Cell B./E. processor is core to these improvements. We will discuss how the processor technologies used today in many of your living rooms are being utilized to produce improved performance in your laboratories.

Historical background of the Cell B./E.: The Cell B./E. is a microprocessor architecture jointly developed by Sony, Toshiba, and IBM since 2000. The first major commercial application of the Cell B./E. was in Sony's PlayStation 3 (PS3) game console. The Cell B./E. architecture enabled the PS3 to speed up physics simulation in order to catch up with the 3D graphics rendering speeds provided by advanced graphics processing units (GPUs). Toshiba has announced plans to incorporate Cell B./E. in high definition television sets and there are many other applications to come. With innovative features such as the XDR memory subsystem and coherent Element Interconnect Bus (EIB) interconnect appear to position the Cell B./E. for future applications in the supercomputing space.

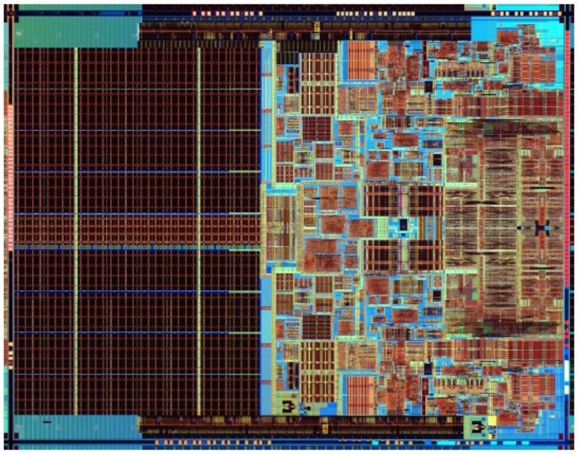
2. Advancements in CPU performance



Moore's Law describes an important trend in the history of computer hardware: that the number of transistors (**transistor count** in the image on the right) that can be inexpensively placed on an integrated circuit is increasing exponentially, doubling approximately every two years. This observation was first made by Intel co-founder Gordon E. Moore in a 1965 paper. This trend has continued for more than half a century and this exponential growth of transistor density has directly translated into exponential performance increase of the CPU's, but only up to a few years ago. As shown in the figure above, the CPU Clock Speed growth has slowed down since about year 2000.

This Clock Speed performance lag is mainly due to three different performance limiting obstacles:

- *The Frequency Wall*: the standard speed up technique of processor instruction pipelining is not returning any more improvements in speed because of code branching
- *The Memory Wall*: the processor frequency has now surpassed the speed of the DRAM and the present workaround of using multi-level caching leads to increased memory latency
- *The Power Wall*: both heat generation and power consumption can become unmanageable



The slow main memory access on traditional Intel/AMD architectures creates a data flow bottleneck causing processor idle times. This results in much lower sustained performance than the theoretical peak of the CPU. To combat the bottleneck, state-of-the-art processors have significant cache, commonly several MBytes on the processor chip. This uses up space that would otherwise be available to allow more dense packing of transistors and as a results more processing power. In the picture of the Intel Core2 duo chip shown to the left the large dark area is occupied by the L2 cache memory. This "wasted" area is one explanation for why Moore's law no longer translates into equivalent performance increases.

Alternatively, special purpose computational hardware such as graphics processors (GPU) on 3D video cards and various hardware accelerators continue to increase performance exponentially. These processors use streaming system architecture, many simple co-processor cores, wide data parallelism and localized dedicated data storage to avoid the previously highlighted performance walls.

The following table provides a performance comparison between Intel or AMD dual core 2.4 GHz processors versus the Cell B./E. developed in collaboration between IBM, Sony and Toshiba:

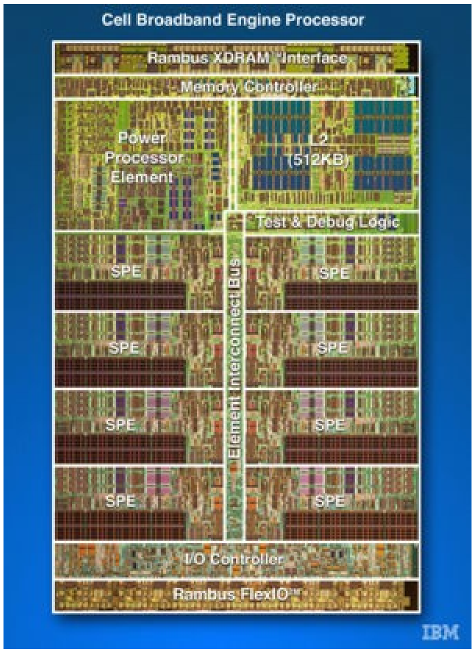
Feature	Intel/AMD	Cell B./E.	Speedup
Maximum instructions per cycle	2	66	33x
Theoretical peak GFlops @ 4GHz*	8	264	33x
Sustained practical GFlops @ 4GHz*	3	210	70x

* GFlops=GigaFlops: billion **FL**oating point **O**perations **P**er **S**econd

The significant difference between the peak and sustained performances of the traditional CPUs stems from the difference in memory access speed as well as latency and cache issues. It is not possible to push enough data through the processor fast enough to keep it busy. This problem is even more striking with the latest quad-core processors where the computational capability is doubled but the memory bandwidth remains the same thereby creating an even more serious bottleneck.

The numbers in the last row of the table are based on the details outlined in section number 4. These numbers relate to the success stories of others as well as our own measurements detailed in section number 7: Flexible docking speed: eHiTS® on the Cell B./E.

3. The Cell Broadband Engine™



How can the Cell B./E., shown in the figure below, achieve such a tremendous performance gain over contemporary processors? The improvements in performance are due to the fact that two different layers of hardware parallelism are invoked:

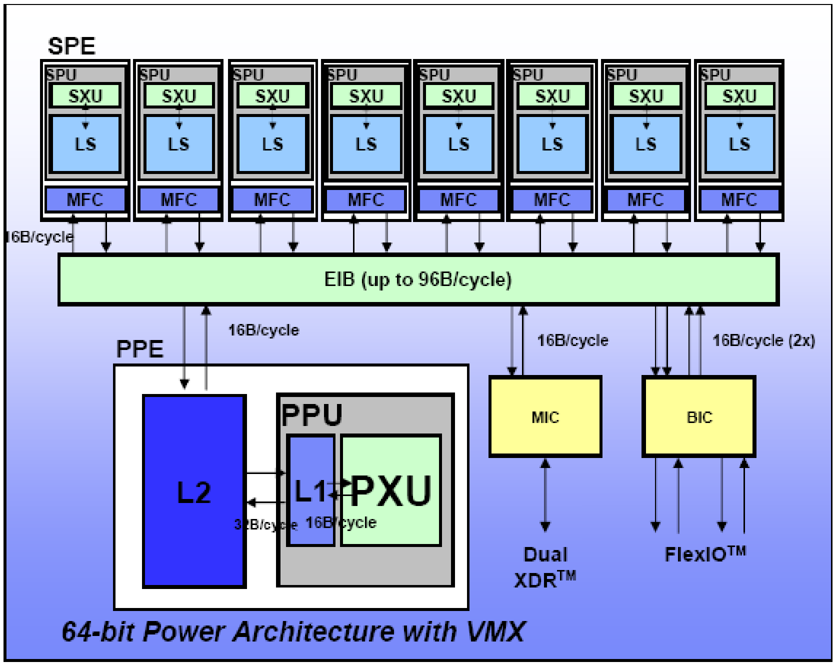
- Multiple processor cores: 1 hyperthreaded Power Processing Element (PPE) plus 8 Synergetic Processing Elements (SPE), each of them with its own DMA controller unit and local fast memory on-die.

- The SPEs are 128 bit vector processors with Single Instruction Multiple Data (SIMD) architecture using dual instruction pipes and are therefore capable of executing 8 floating point operations per clock cycle, each with 128x128bit registers and 256KB local store.

Another important design difference from the traditional Intel/AMD/PPC processors is the memory architecture using dedicated on-die local memory storage for each SPE with separate Memory Flow Controller (MFC) unit for managing data transfer plus the use of XDR RAMBUS memory that can operate at CPU clock speeds (>3GHz versus <1GHz of DDR RAM used in PCs) thereby eliminating the need for large L2 cache memory and data latency.

Performance specification highlights of the Cell B./E. processor:

- 241 million transistors
- 9 cores, 10 parallel execution threads
- >200 GFlops single precision
- Up to 25 GB/s memory bandwidth
- Up to 75 GB/s I/O bandwidth
- >300 GB/s bandwidth on the Element Interconnect Bus (EIB)
- Top frequency > 4GHz



Cell B./E. Block Diagram; Courtesy of International Business Machines Corporation.

The Figure above shows the architecture diagram of the Cell B./E. processor. The PPE accesses the main memory via the L1 and L2 cache in the traditional way, thus accessing the whole main memory. There is however a latency penalty caused by two layer caching. However, the SPEs can only operate directly on their own 256KB local memory. Each SPE has a programmable MFC that allows fast data transfer between the main memory and the local store. The SPEs can also communicate with each other and with the PPE via the very high speed EIB. Most of the processing power of the Cell B./E. comes from the SPEs since each of them is capable of executing up to 8 floating point or 32 bit integer operations per cycle, a total of 64 operations in a single cycle. Alternatively, they can execute 128 * 16 bit operations or 256 * 8 bit operations in parallel. A good application design for the Cell B./E. uses the PPE for main control and disk I/O tasks, while all computation intensive tasks are distributed to the SPEs.

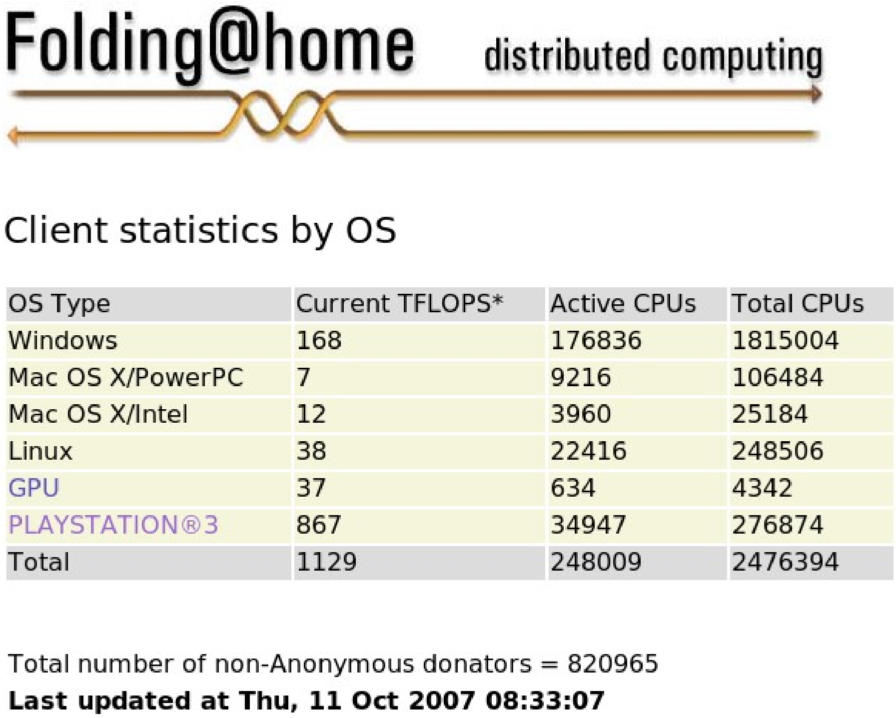
4. Success stories

One of the reasons to pursue development of the Cell B./E. processor was that the speed of physics simulation in video games was lagging far behind the graphics performance of 3D rendering. The realism level of 3D games was no longer limited by the capability of the graphics processing unit being able to display the dynamically changing environment, but instead the CPUs were not able to compute the movement of objects or particles with realistic physics simulations as fast as the graphics could render it. With the introduction of the Cell B./E. the situation is reversed: RapidMind has a crowd simulation demo (see picture on the right, copyright of RapidMind) running on the Cell B./E. simulating a chicken farm. Each individual chicken has its own behavior model interacting with other birds. The simulator was demonstrated to provide real-time (30fps) performance with several thousand chickens. In fact, when the number of chickens was increased to a total of 15,000 birds the Cell B./E. processor was still able to perform the simulation with interactive speed, but

the graphics rendering was not able to keep pace, even on a state-of-the-art NVidia GPU and started dropping 2 out of 3 frames, resulting in 10fps "sluggish" video output.

While interactive gaming relies on GPUs to render 3D virtual reality graphics, Hollywood has been striving for photo-realistic rendering to be used in CGI effects for movies dependent on special effects (Star Wars, Lord of The Rings etc.). To achieve photorealistic rendering CGI studios use a technique called ray tracing that is based on modeling the path taken by light by following rays of light as they interact with optical surfaces. Such modeling calculations are so demanding that on a single traditional CPU it takes several hours to render a single image, i.e. one video frame. Movies require 24 or 30 frames per second, so CGI studios use large CPU farms to perform rendering for the films. The Cell B./E. processor has been demonstrated to perform real time raytracing on both the Sony PS3 (~30fps) and the IBM QS20 blade (~60fps).

Mercury Computer Systems reported a 15x to 30x speedup of Fast Fourier Transform (FFT) algorithms on the Cell B./E. compared to the best commercially available substitute processors. It is projected that they will produce up to a 100x improvement for actual customer applications.



Client statistics by OS

OS Type	Current TFLOPS*	Active CPUs	Total CPUs
Windows	168	176836	1815004
Mac OS X/PowerPC	7	9216	106484
Mac OS X/Intel	12	3960	25184
Linux	38	22416	248506
GPU	37	634	4342
PLAYSTATION®3	867	34947	276874
Total	1129	248009	2476394

Total number of non-Anonymous donators = 820965

Last updated at Thu, 11 Oct 2007 08:33:07

Folding@Home is a distributed computing project designed to perform computationally intensive simulations of protein folding and other molecular dynamics simulations. It was launched on October 1, 2000, and is currently managed by the Pande Group, within Stanford University's Chemistry department. As of June 2007 fifty scientific research papers have been published using the project's work. The PS3 firmware version 1.60 (released in March 2007) allows for Folding@home software to be used on the Sony PlayStation 3. The intent was that gamers would be able to contribute to the project by merely "contributing electricity", leaving their PlayStation 3 consoles running the client while not playing games. Within a month PS3 users were delivering nearly 400 teraflops, achieving a total computing power of over 700 teraflops at a single moment. This was more than double the computing capacity of the network before PS3 joined the program. The current client statistics segregated by operating system is available here. A snapshot of the table for October 11th is shown below. It shows that about 35 thousand PS3s are delivering 867 TeraFlops out of the total 1129 TeraFlops provided by about 250 thousand total CPUs. The PS3 systems represent only 14% of the CPUs but provide 77% of the computing power! Note that the statistics are collected by labeling CPUs active if they have contributed any work within the last 5-10 days, but it does not mean that they have been performing folding calculations all the time because the project uses idle times only.

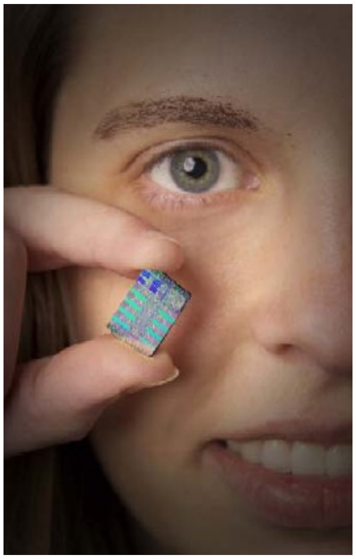
5. Coding differences

Considering all the positive comments regarding the Cell B./E. processor are there any downsides to this processor? The performance exhibited by the Cell B./E. is not delivered simply by recompiling existing C/C++ code for the Cell B./E. processor. There are several key differences in coding for the Cell B./E. processor versus traditional CPUs:

- The PPE and the SPE are not binary compatible. The code needs to be compiled with a different compiler to generate code fragments that run on the SPE. Simple posix threads cannot be scheduled by the OS to run on the SPE.
- A thread running on an SPE can only directly access the 256KB local storage of that SPE. This small amount of memory must contain the code, the data and the execution stack at any one time. Of course, data or code can be shuffled back and forth between the main memory and the local store via DMA calls, but those have to be explicitly progammed and managed by the code. Furthermore, the DMA calls in the code have to be designed to use double buffering or similar tricks to streamline data and avoid stalling due to latency.
- The power of the SPEs (256 GFlops out of a total of 264 GFlops) comes from SIMD vector operations, where the same instruction is executed for multiple data entries. An SPE using dual-pipe can execute 8 instructions per cycle. These are 2 different instructions each executed on 4 parallel data units and those data units *must* reside in a consecutive block of memory or in a single 128 bit register. High performance can therefore only be reached if the data is organized in a very specific way that is suitable for SIMD calculations.
- SPEs do not have branch prediction hardware. This was one of the simplifications compared to traditional CPUs that allowed the engineers to pack more computation cores into the chip. The downside of this is that branches in the code are more costly and lead to a loss in pipeline efficiency. A single branch-miss costs an 18 cycle delay and 144 (from 18x8) potential instruction executions are lost! Therefore, code must be written by minimizing branches. If a choice needs to be made, it is often worth computing both alternatives and selecting the most appropriate one from the results rather than inserting a branch to compute only what is necessary.

The paradigm differences described above require significant data structure and algorithmic modifications, redesigns and innovative thinking. Traditionally, optimized algorithms would reduce the necessary number of calculations by utilizing smart data structures such as lookup grids, search trees and so on. These techniques typically require large memory with random data access patterns and lots of branches in the code but these are not suitable for high performance Cell B./E. computation.

6. Cost savings



Returning to the positive aspects of the Cell B./E. processor we find that another major benefit, apart from the performance increase, is that it significantly lowers the cost of computation. There are several sources of cost savings to the users of Cell B./E. compared with traditional HPC solutions such as clusters, CPU farms or supercomputers. These are:

- Initial hardware cost (per GFlops) is much lower
- Electricity cost - Cell B./E. requires far less power
- Cooling cost - heat emission is much lower
- Space saving - smaller CPU and much fewer boards

The same advantages also apply to other hardware accelerator technologies, e.g. GPGPU or FPGA, but the Cell B./E. offers an order of magnitude more savings than the FPGA solution as demonstrated in the following table, with 3-year running cost calculations listed.

Costs	400 CPU cluster	FPGA	Cell B./E. (8xPS3)
Hardware purchase	\$200K-\$400K	\$60K	\$4K
Electricity (power+cooling)	\$180K-\$360K	\$6K	\$3K
Total cost	\$380K-\$760K	\$66K	\$7K

7. Flexible docking speed: eHiTS® on the Cell B./E.

The eHiTS flexible ligand docking software is being ported to run natively on the Cell B./E. fully utilizing the vector processing power of the SPEs. The port is currently in beta-testing phase and commercial release is planned for the third quarter of 2008. The following table shows the execution times and the speed improvement reached for one example program, i.e. the Lennard Jones 6-12 potential calculation, part of the eHiTS program.

System	**PPU only (no SPUs) ~ Power Mac**	**Intel ~ dual core, 2.4 GHz**	**PS3**	**IBM QS21**
Time	3m 34s	1m 14s	2.6s	1.0s
Speed up	1x	3x	82x	214x

The next table shows more examples, with speedup factors achieved in various calculations on the Cell B./E. compared to traditional CPUs at equivalent clock speeds. Since the code is not yet fully tuned or optimized, these are preliminary numbers. The final results are expected to be even more significant.

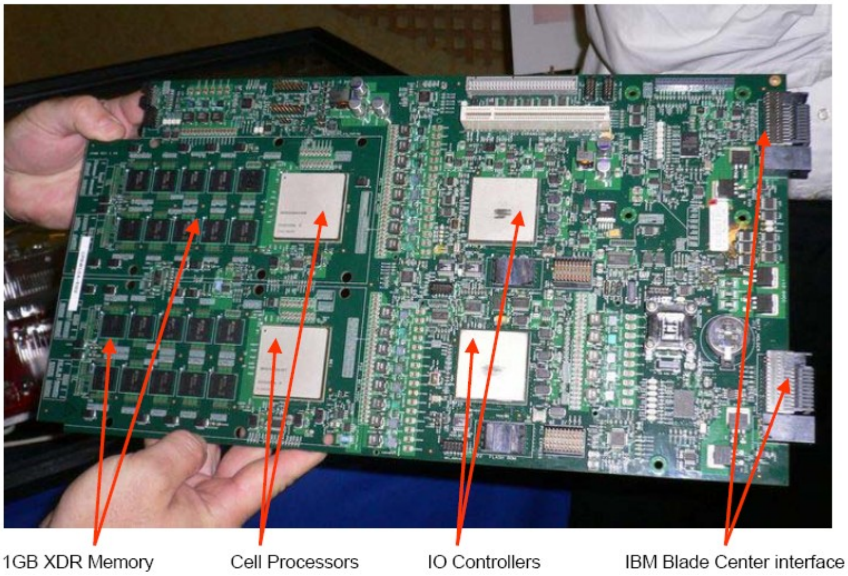
Function ported and tested	**Speed-up factor**		
	PS3	**CAB**	**QS20**
eHiTS Scoring (with rotamer optimization)	20x	27x	55x
Rigid Fragment Docking	21x	28x	56x
Pose Matching	7x	9x	18x
Conformation Minimization	24x	32x	65x
Final Optimization	12x	16x	33x
Total (Complete Flexible Docking)	**13x**	**18x**	**36x**

The complete application speedup depends on the input structures (the size of the cavity versus the size of the rigid fragments, the number of rigid fragments etc.) and different cases require a different number of function calls of various types and therefore ranges between 26-fold and 60-fold.

It is important to note that due to the nature of the Cell B./E. the redesigned algorithms perform more brute force calculations versus smart cut-offs applied in the previous versions of the software. This paradigm shift leads to higher accuracy which is yet another significant benefit to the port.

8. Road map of the Cell B./E. processor

All benefits listed for the Cell B./E. are very attractive. However, the question should be asked: is it worth investing the time to develop for this infrastructure? Would Moore's law not guarantee that advancements in traditional CPU design will allow Intel/AMD to catch up to the Cell B./E. performance within a few years?



That would certainly be a valid concern if one assumed that the Cell B./E. remains stagnant and without further development, a potential situation if the Cell B./E. was only the basis of a single game console such as the PS3. However, the collaboration between IBM (who markets the Cell B./E. as an HPC board as pictured above), and Sony and Toshiba for the development of the Cell B./E. technology has been extended with a signed commitment for another 5 years with the following intentions:

Availability:	2006	2007	2008	2009-2010
SDK version	1.1	2.1	3.0	5.0
Processor cores	1PPE+8SPE	1PPE+8SPE	1PPE+8eDP-SPE	2PPE+32eSPE
SP performance	230 GFlops	264 GFlops	330 GFlops	1000 GFlops
DP performance	21 GFlops	109 GFlops	136 Gflops	500 GFlops
Memory	512 MB	1 GB	16 GB	?
Technology	90nm SOI	90nm SOI I/O BW enh.	65nm SOI PCI-Exp.x16	45nm SOI Next gen. memory technology

This commitment from the three hardware manufacturing giants guarantees that the currently observed 50-fold performance gain over state-of-the-art traditional CPU technologies is not merely a short term advantage but rather a persistent ratio will likely be maintained over the next five years.

9. Conclusions

In this white paper we have provided an overview of the Cell B./E. processor and the obvious potential that this advanced technology can offer to improved speed and performance of docking experiments. SimBioSys has been committed to providing the most innovative and optimized algorithms available to scientists. In order to provide high accuracy performance we have occasionally had to trade off accuracy for time associated with the calculations. Our commitment to accuracy has been demonstrated in the eHiTS program since its original release. Now, when we couple the improved performance we continue to demonstrate version to version with the throughput and performance of the Cell B./E. processor the potential impact on improved drug discovery is obvious. The coupling of superior software with high performance hardware offers a potential revolution in capabilities for scientists.

For more information please contact SimBioSys, Inc. by calling: +1 416 741 4263, or send e-mail to: info@simbiosys.ca, or visit our web site: www.simbiosys.ca

SimBioSys, Inc.
135 Queen's Plate Drive Univ 520
Toronto, ON Canada M9W 6V1
tel: +1 416 741 4263
fax: +1 416 741 5084
e-mail: info@simbiosys.ca
web: www.simbiosys.ca