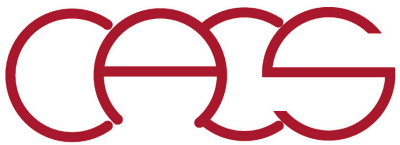
Lattice-Boltzmann (LB) Fluid Simulation on a Playstation3 (PS3) Cluster

Ken-ichi Nomura, Liu Peng & Richard Seymour

Collaboratory for Advanced Computing & Simulations
Dept. of Computer Science, Dept. of Physics & Astronomy,
Dept. of Chemical Engineering & Materials Science
University of Southern California

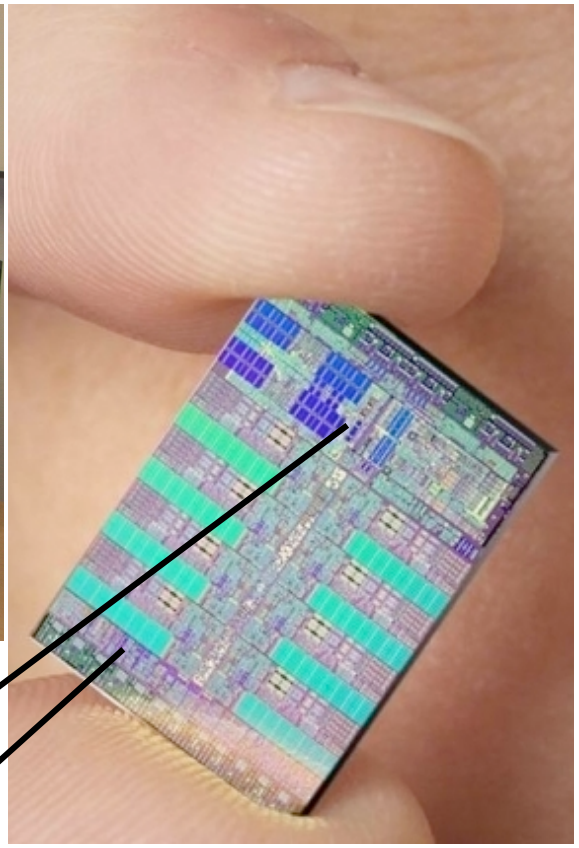
Email: (knomura, liupeng, rseymour)@usc.edu

"Parallel lattice Boltzmann flow simulation on a low-cost PlayStation3 cluster," K. Nomura, S.W. de Leeuw, R.K. Kalia, A. Nakano, L. Peng, R. Seymour, L.H. Yang, & P. Vashishta, *Int'l J. Comput. Sci.* **2**, 437 ('08)



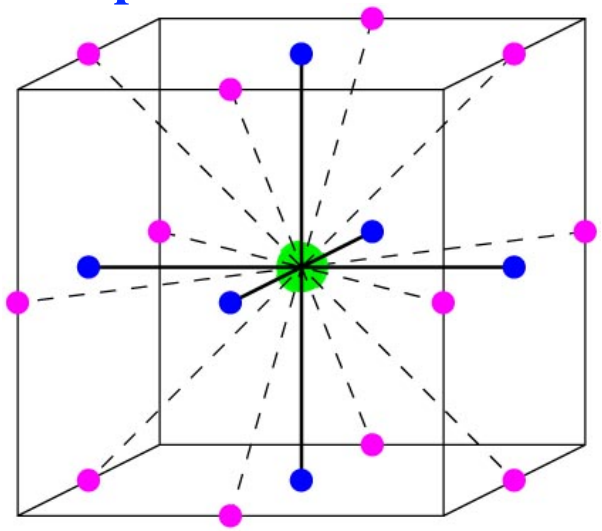
CACS Playstation3 (PS3) Cluster

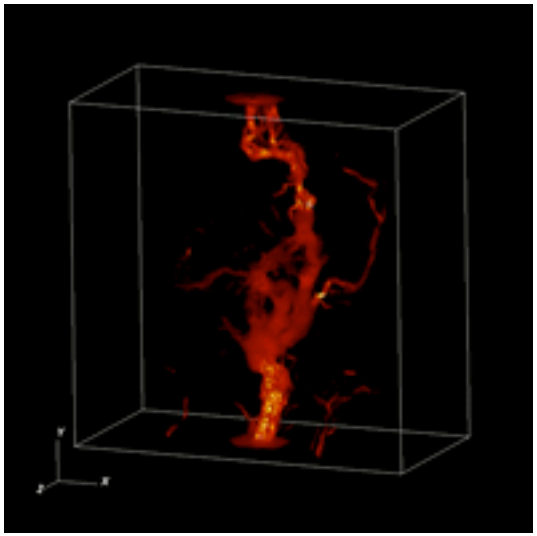
- 9 PS3's connected via a Gigabit Ethernet switch
- 3.2×2×2×2×9 = 230 sp-Gflops peak on CPU (1 PPE + 8 SPEs*) & 1.8 Tflops on CPU+GPU per box @$500 (*6 available to user)
- 2.1 Tflops on CPUs & 16.2 Tflops on CPUs+GPUs on the 9-PS3 cluster @5K** (**including a Gigabit Ethernet switch)



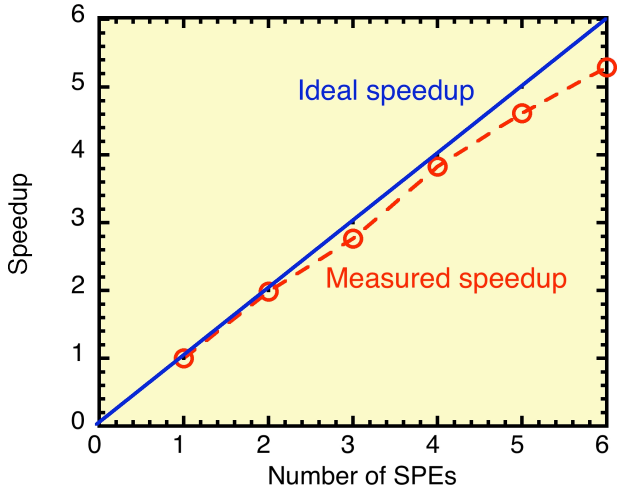
Lattice-Boltzmann (LB) on a PS3 Cluster

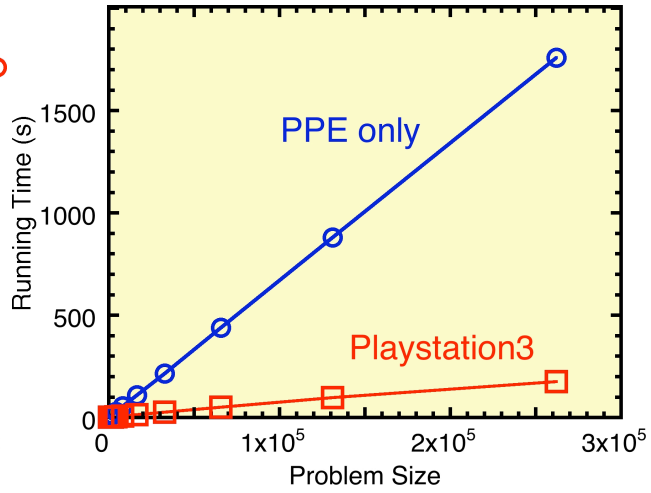
- **Developed a parallel LB flow-simulation code on a PS3 cluster**

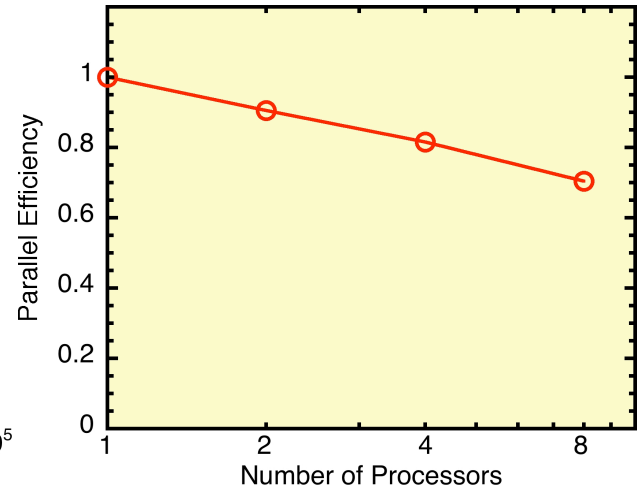




- **Achieved performance improvement of factor 13.2 over a PowerPC per PS3 box as well as good (0.882) multi-threading & reasonable (0.705) inter-console parallel efficiency over an inexpensive Ethernetswitch**







Setting Up a PS3 Linux Cluster

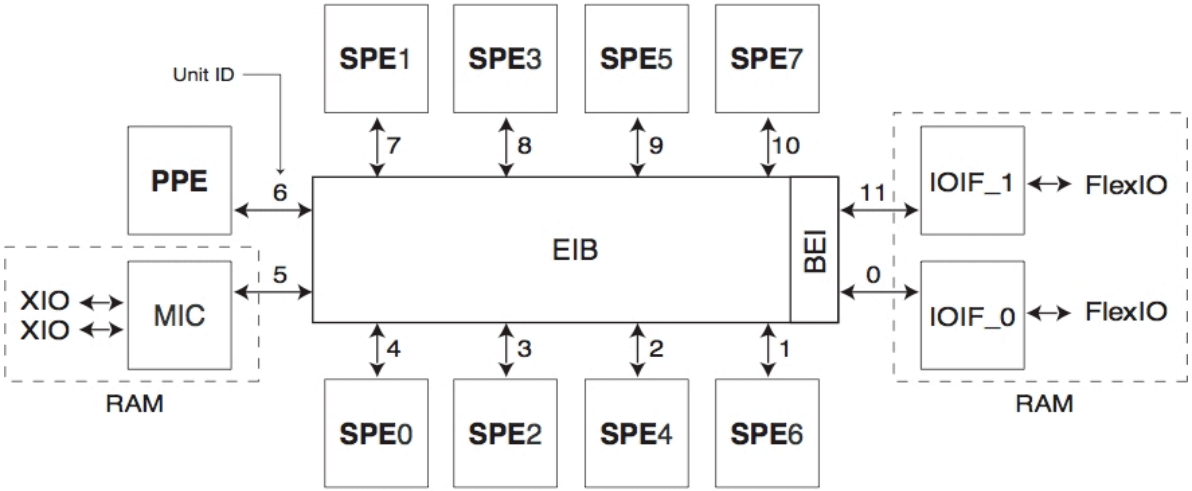
- **Fedora Core 6 Linux OS distribution with libraries & infrastructure to support the IBM Cell Software Development Kit (SDK) version 2.1**
- **The SDK offers an IBM compiler & the GNU compiler collection (gcc) for the Cell processor**
- **Installation instructions at**

  `http://www.ps3coderz.com/index.php?option=com_content&task=view&id=73&Itemid=31`

- **PS3's connected together via a gigabit Ethernet switch (each PS3 has an gigabit Ethernet port)**
- **Message passing interface (MPI) installed as in a standard Linux cluster**

PS3-Cluster Parallel Programming

- **Within each PS3 box, a main program runs on the PowerPC processor element (PPE): 3.2 GHz, 256MB RAM, 32KB L1, 512KB L2**
- **The PPE program spawns Posix threads that run on multiple synergistic processor elements (SPEs): 3.2GHz, 256KB RAM**
- **Direct memory access (DMA) for data transfer between the PPE & SPEs**
- **MPI for inter-box message-passing programming**

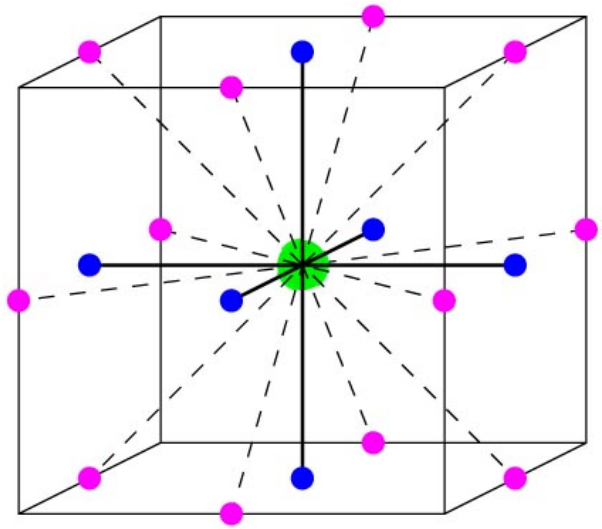


BEI	Cell Broadband Engine Interface	MIC	Memory Interface Controller
EIB	Element Interconnect Bus	PPE	PowerPC Processor Element
FlexIO	Rambus FlexIO Bus	RAM	Resource Allocation Management
IOIF	I/O Interface	SPE	Synergistic Processor Element
		XIO	Rambus XDR I/O (XIO) cell

Lattice-Boltzmann Simulation

- **Distribution function: number density of particles at lattice position x with (discretized) velocity v_i (i = 0,...,17) at time t**

$$f_i(\vec{x},t)$$



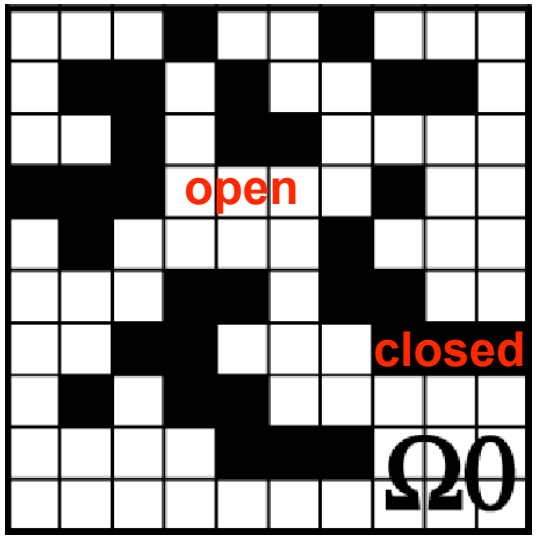
- **Time-stepping**
  - **Collision**

$$f_i(\vec{x},t^+) = f_i(\vec{x},t) - \frac{1}{\tau}\left(f_i(\vec{x},t) - f_i^{eq}(\vec{x},t)\right)$$

  - **Streaming**

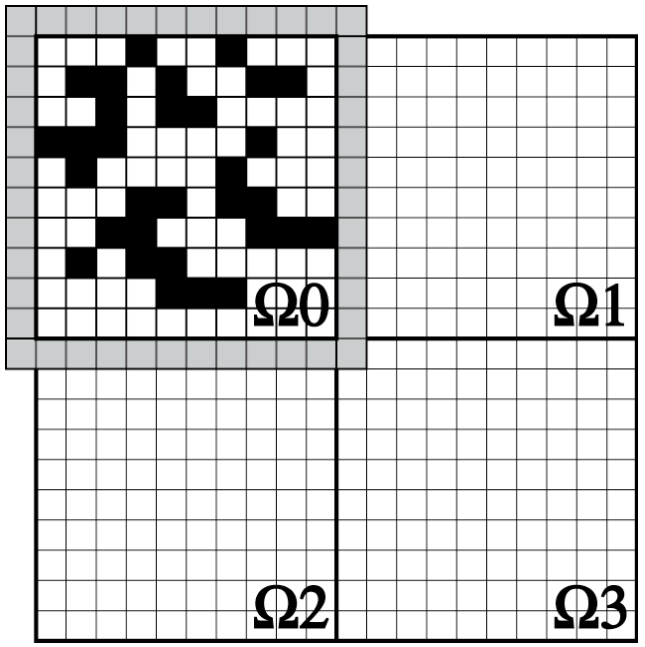
$$f_i(\vec{x}+\vec{e}_i,t+1) = f_i(\vec{x},t^+)$$

- **Irregular geometry as obstacles**

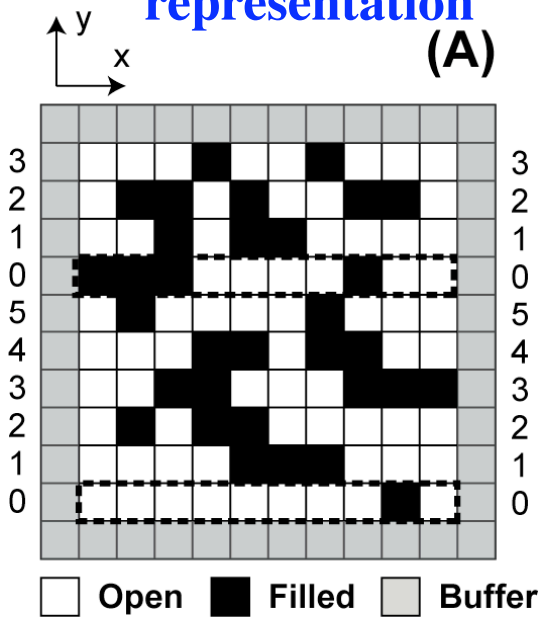


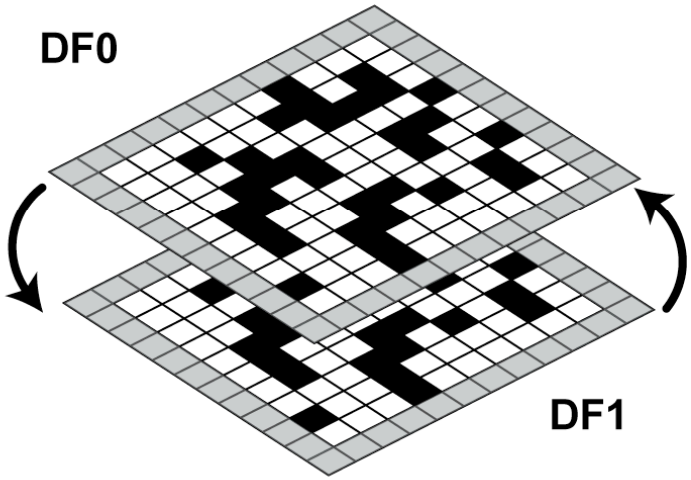
Parallel Lattice-Boltzmann Algorithm

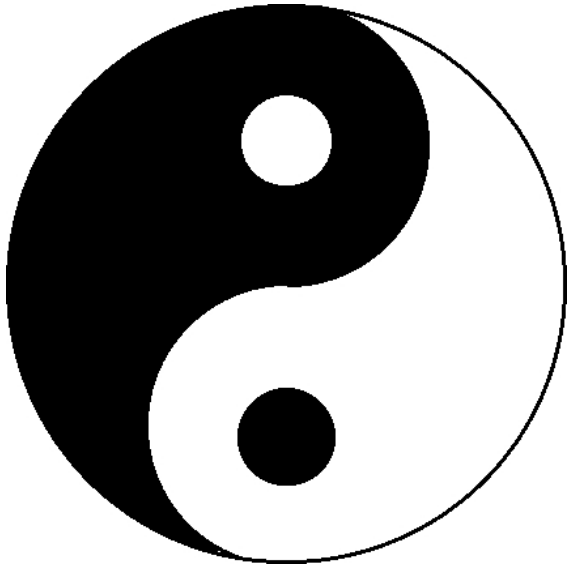
- **Hybrid thread+ message passing programming**
  - **1 Inter-box parallelization:**
    - **> spatial decomposition + message passing**
  - **2 Intra-box parallelization**
    - **> Multithread processing of interleaved rows**
    - **> Critical section-free, dual representation**



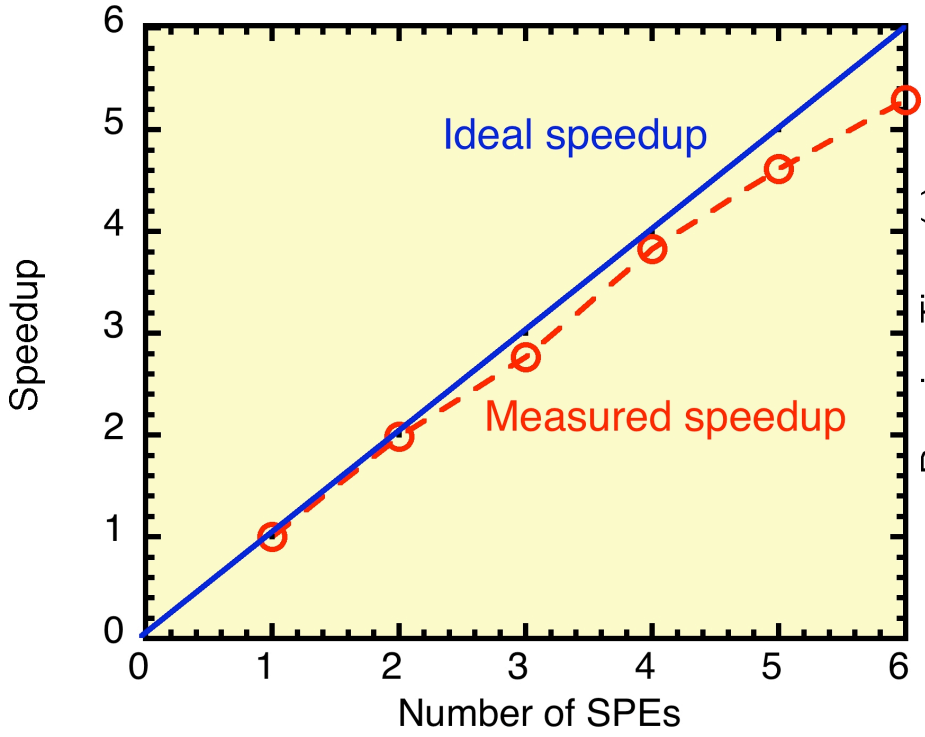
(B)

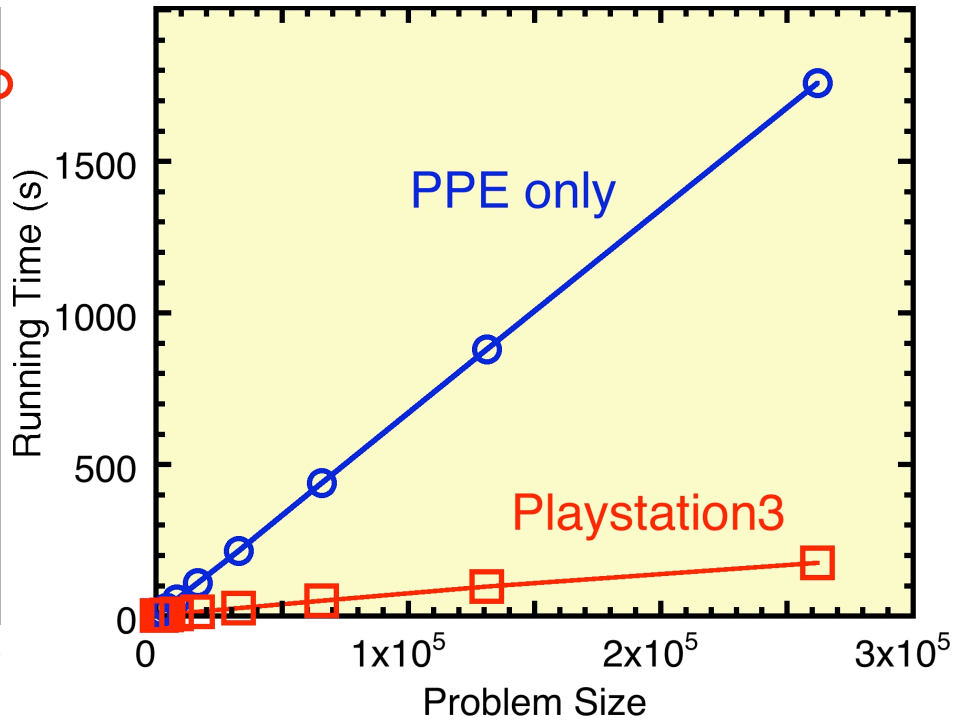






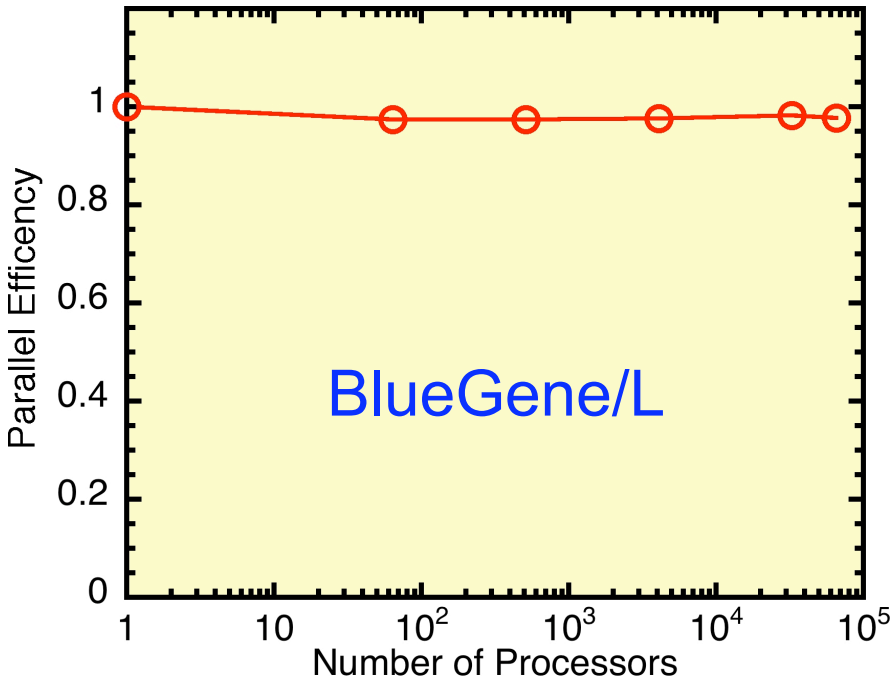
Multi-SPE Performance on a PS3

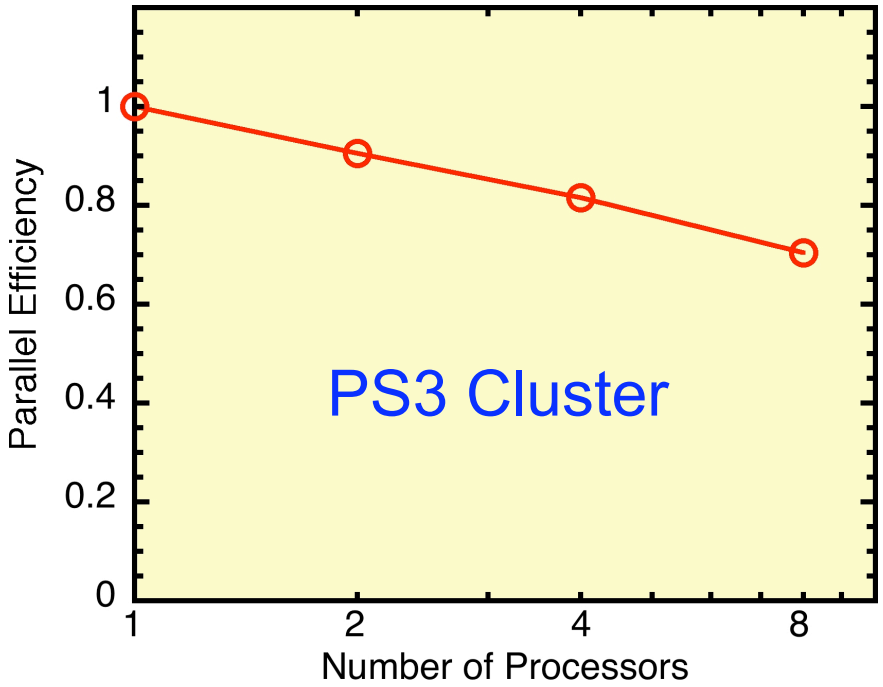




- **Multi-threading parallel efficiency 0.882 on 6 SPEs**
- **Performance improvement of factor 13.2 over an PowerPC using 6 SPEs**

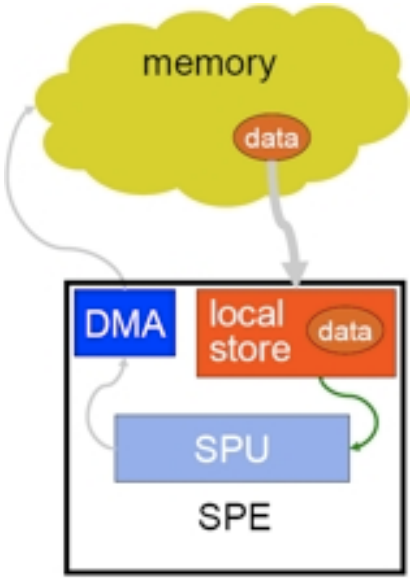
Inter-box Performance

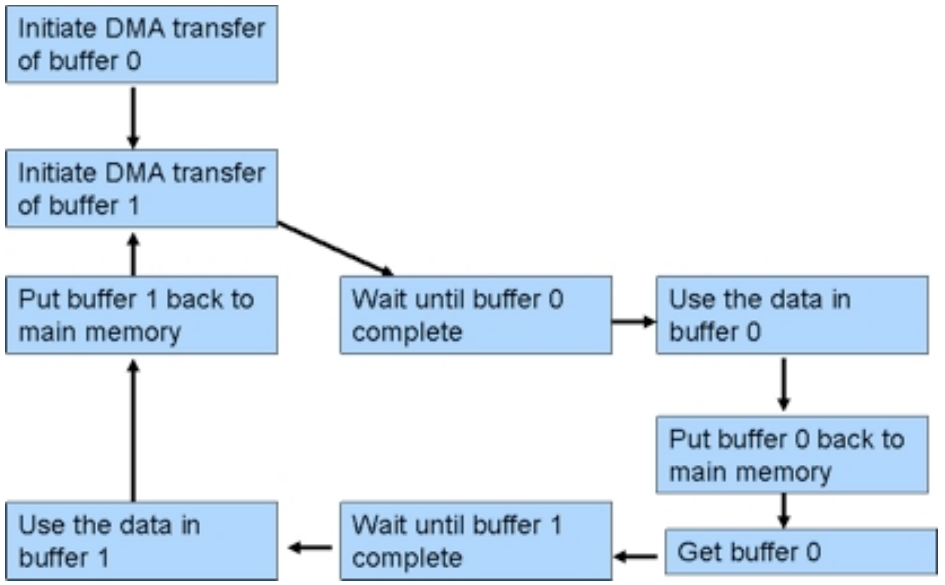


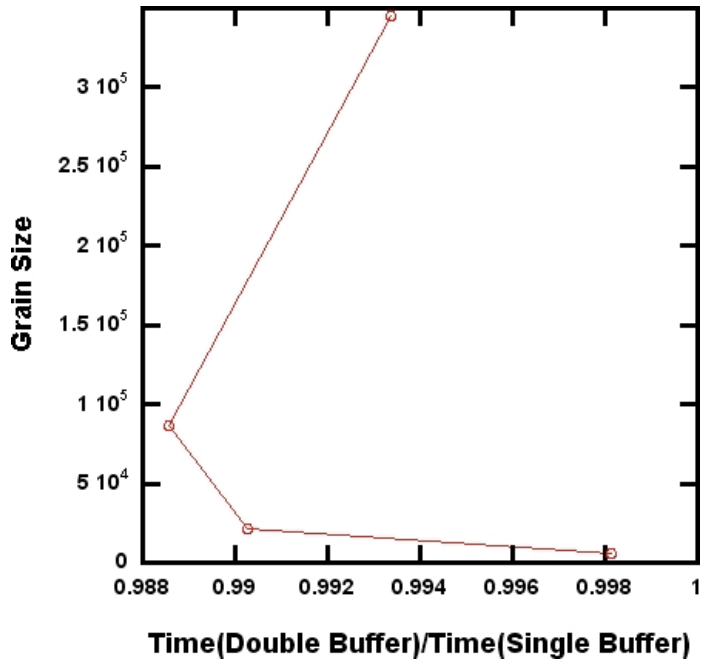


- **Good concurrency & data locality of the algorithm: parallel efficiency 0.977 on 65,536 IBM BlueGene/L nodes**
- **Modest parallel efficiency 0.704 on a PS3 cluster due to the limited bandwidth of the Gigabit Ethernet switch**

Double-Buffer DMA







In Progress: SIMD