



# LANL Roadrunner Summary

3/12/2007

Ken Koch

Computer, Computational, and Statistical  
Sciences Division

Los Alamos National Laboratory



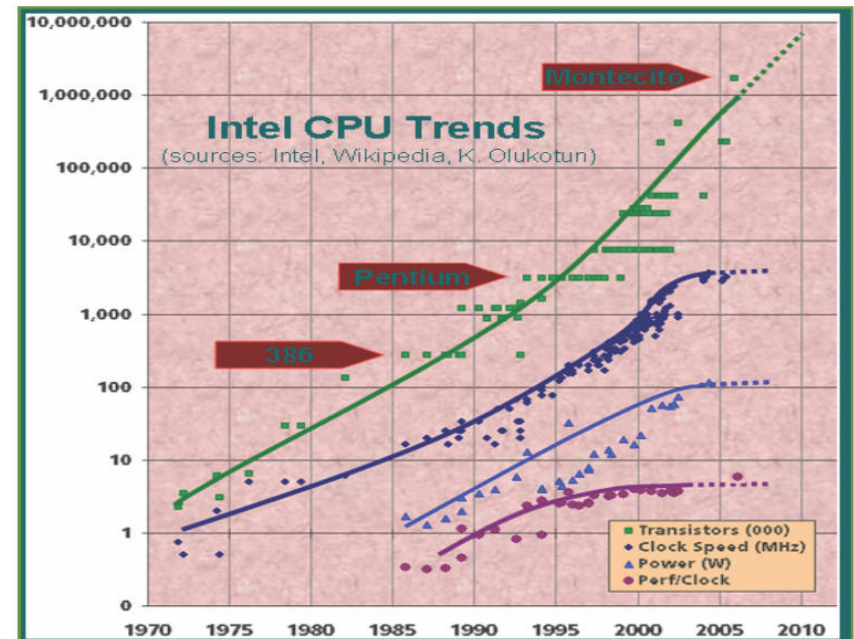
# Microprocessor Trends

- **New implications of Moore's law**

- Moore's law still holds but is now being realized differently
  - Clock frequency, chip power, & instruction-level-parallelism (ILP) have all plateaued
  - Multi-core is here today and manycore ( $\geq 32$ ) looks to be the future
  - Complexity of shared memory and cache coherency for multi-core designs is likely not scalable to manycore designs
  - Memory bandwidth and memory capacity per core are headed downward (predominantly caused by increased core counts)

- **Key references:**

- IDC report #205025, January 2007
- UC Berkeley UCB/EECS-2006-183
- LASCI-06 Burton Smith keynote "Reinventing Computing"



From Burton Smith, LASCI-06 keynote, with permission

# Hybrid Computing Trends

---

- **Atypical and/or truly heterogeneous chips are appearing and being planned**
  - IBM Cell , AMD Fusion, Intel Polaris, NVidia G8800
  - From TGDaily article: ( [http://www.tgdaily.com/2007/02/11/intel\\_80\\_core/](http://www.tgdaily.com/2007/02/11/intel_80_core/) )
    - *According to Intel, the "manycore" design of the chip would be able to house different processor cores, including "general purpose" cores that are necessary, for example, to efficiently run traditional applications such as an operating system. In this view, Intel's approach has some similarity to AMD's design approach of the "Fusion" processor, which is expected to merge general purpose cores with graphics and cores. Intel, however, did not say whether it plans to integrate graphics capability into its 80-core chip.*
  - From Intel: ( [ftp://download.intel.com/research/platform/terascale/terascale\\_overview\\_paper.pdf](ftp://download.intel.com/research/platform/terascale/terascale_overview_paper.pdf) )
    - *Legacy software and inherently single-threaded algorithms present another challenge to tera-scale architectures. To address that issue, Intel researchers are exploring ways to incorporate heterogeneous general-purpose cores (both single-thread and multithread optimized cores) into tera-scale architectures. In addition, researchers are exploring the use of ensembles of simple cores to accelerate single threads through thread-level speculation and other advanced techniques.*
- **Hybrid technologies:**
  - Accelerators
    - FPGAs, GPGPU, Clearspeed CSX600, IBM Cell, XtremeData XD1000, Nvidia G8800, AMD Stream Processor
  - Connection standards
    - AMD Torrenza, Intel/IBM Geneseo, AMD HyperTransport Initiative
  - Programming
    - RapidMind, Peakstream, Impulse C, Stanford's Sequoia, NVidia CUDA, Clearspeed C, Mercury MFC, IBM Roadrunner ALF & DaCS libraries, stream programming

# Decades of Hybrid Computers

---

- **Before Roadrunner there was:**
  - Floating Point Systems FPS Array Processors (AP-120B, FPS-164/264) (circa 1976-1982)
    - [http://en.wikipedia.org/wiki/Floating\\_Point\\_Systems](http://en.wikipedia.org/wiki/Floating_Point_Systems)
  - Deep Blue for chess (IBM SP-2: 30 RS6K + 480 chess chips) (circa 1997)
    - [http://en.wikipedia.org/wiki/Deep\\_Blue](http://en.wikipedia.org/wiki/Deep_Blue)
  - Grape-6 for stellar dynamics w/ custom chips) (circa 2000-2004)
    - <http://grape.astron.s.u-tokyo.ac.jp/~makino/grape6.html>
  - Various FPGA supercomputers from system vendors:
    - SRC-6 (w/ MAP)
    - Cray XD1 (w/ Application Acceleration)
    - SGI Altix (w/ RASC)
  - Titech TSUBAME (w/ some Clearspeed) (2006)
    - <http://www.gsic.titech.ac.jp/English/Publication/pressrelease.html.en>
  - RIKEN MDGrape-3 “Protein Explorer” (w/ custom chips) (2006)
    - <http://mdgrape.gsc.riken.jp/modules/tinyd0/index.php>
  - Terra Soft’s Cell E.coli/Amoeba PS3 Cluster (cluster of 1U PlayStation 3 development systems) (2007)
    - <http://www.hpcwire.com/hpc/967146.html>

# Roadrunner's Goals

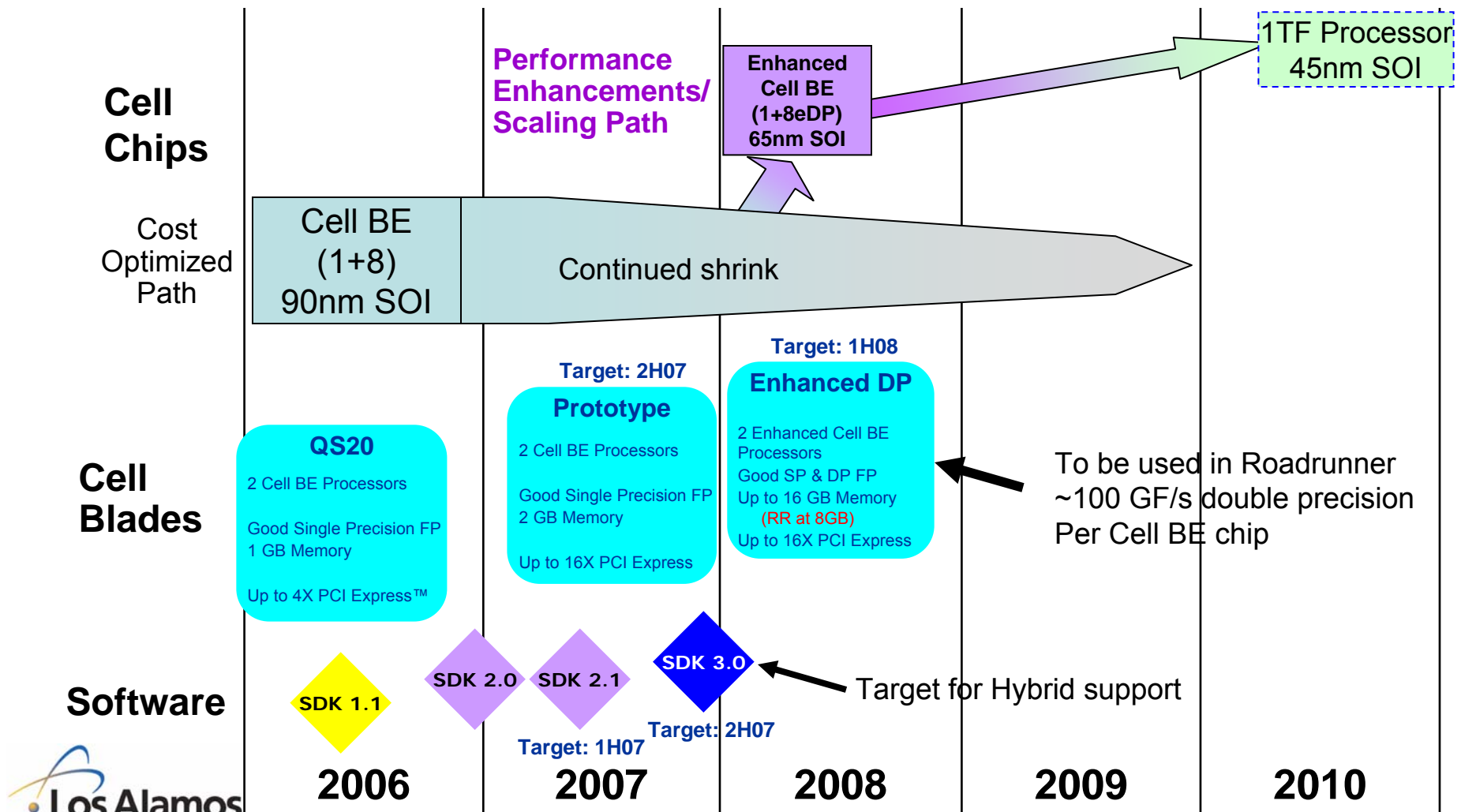
---

- **Provide a large “capacity-mode” computing resource for LANL weapons simulations**
  - Purchase in FY2006 and stand up quickly
  - Robust HPC architecture with known usability for LANL codes
  - Cycles for throughput of 64-way to 1000-way parallel jobs
- **Upgrade to petascale-class hybrid “accelerated” architecture in 2008**
  - Support future LANL weapons physics and system design workloads
  - Follow future trends toward hybrid/heterogeneous computers
    - Manycore and varied processors with special function units
  - Achieve a sustained PetaFlop
- **Roadrunner was designed to do both**
  - Obviously, the tension between “known usability” and “petascale hybrid accelerated” had some impact on the architecture design and timeline

# Roadrunner Timeline

- **Phase 1**      **Late-2006**      **Stage 1 Deployment**
  - Multiple non-accelerated clustered systems Oct. 2006
  - A large classified capacity at LANL
  - A small Advanced Architecture Initial System (AAIS) with 7 nodes with QS20 Cell blades
- **Phase 2: Technology Refresh & Assessment**      **Mid-2007**
  - Add 6 nodes of next generation Cell blades
  - Supports pre-Phase 3 assessment
- **Phase 3**      **2008**      **Stage 2 Deployment**
  - Populate entire system with 3<sup>rd</sup> generation Cell-eDP Blades
  - Run some accelerated weapons science codes initially
  - Achieve a **sustained** 1 PetaFlop Linpack
  - Contract Option

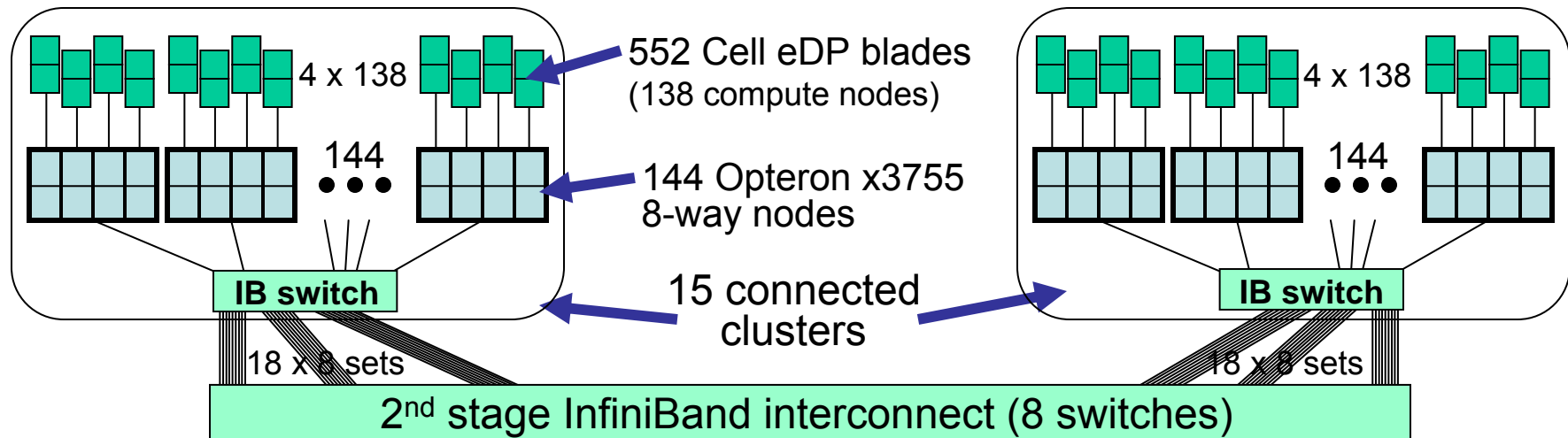
# Cell Technology Roadmap



# Roadrunner Architecture



## Plan-of-record 2008 Cell-Accelerated System



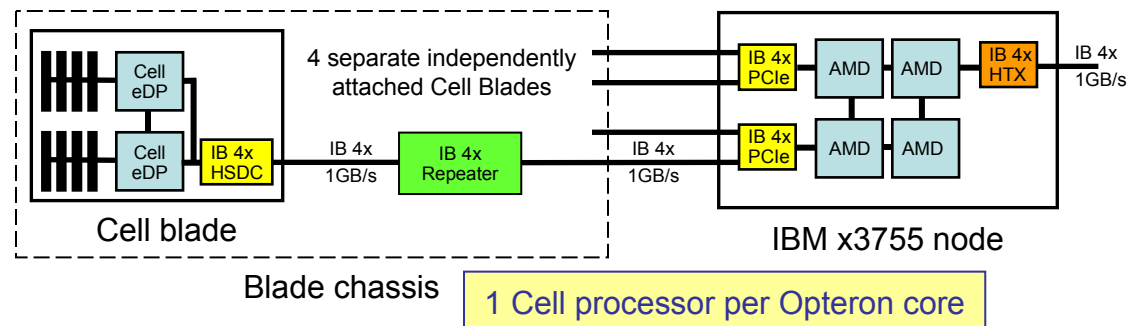
8,640 dual-core Opterons

⇒ **76 TeraFlops**

16,560 Cell eDP chips

⇒ **1.7 PetaFlops**

### One Accelerated Node



# Programming Roadrunner

- **Computational Library (ALF w/ IBM)**

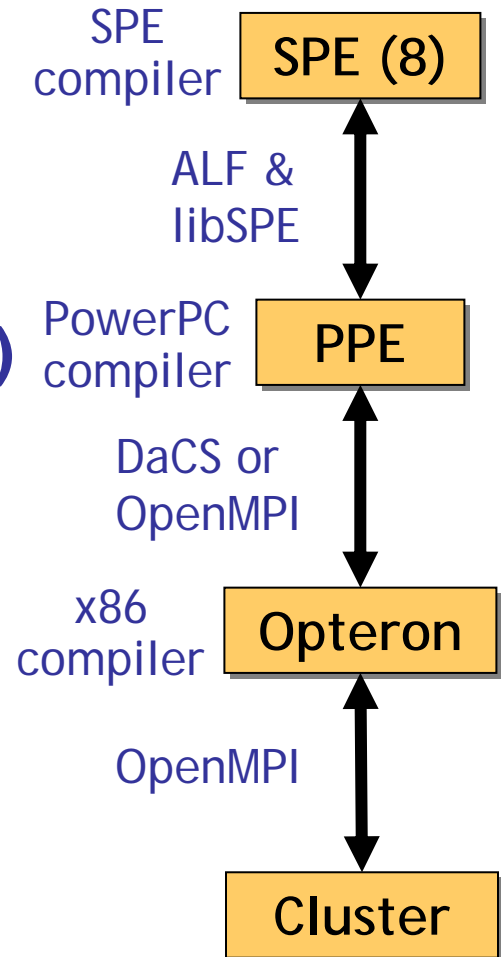
- Task & work block queuing & management
- Streaming & user-defined data partitioning
- Process management
- Error handling

- **Communication Library (DaCS w/ IBM)**

- Data movement & synchronization
- Process management & synchronization
- Topology description
- Error handling
- First implementation may leverage OpenMPI

- **Longer term**

- ALF & DaCS support in tools
- ALF from Opteron  $\Rightarrow$  Cell directly
- Compilers supporting some of this



# LANL Roadrunner URL

---

More information is available at:

<http://www.lanl.gov/orgs/hpc/roadrunner/>