Roadrunner: Science, Cell and a Petaflop/s

International Supercomputing Conference

Press Conference, June 18, 2008

Andy White

Los Alamos

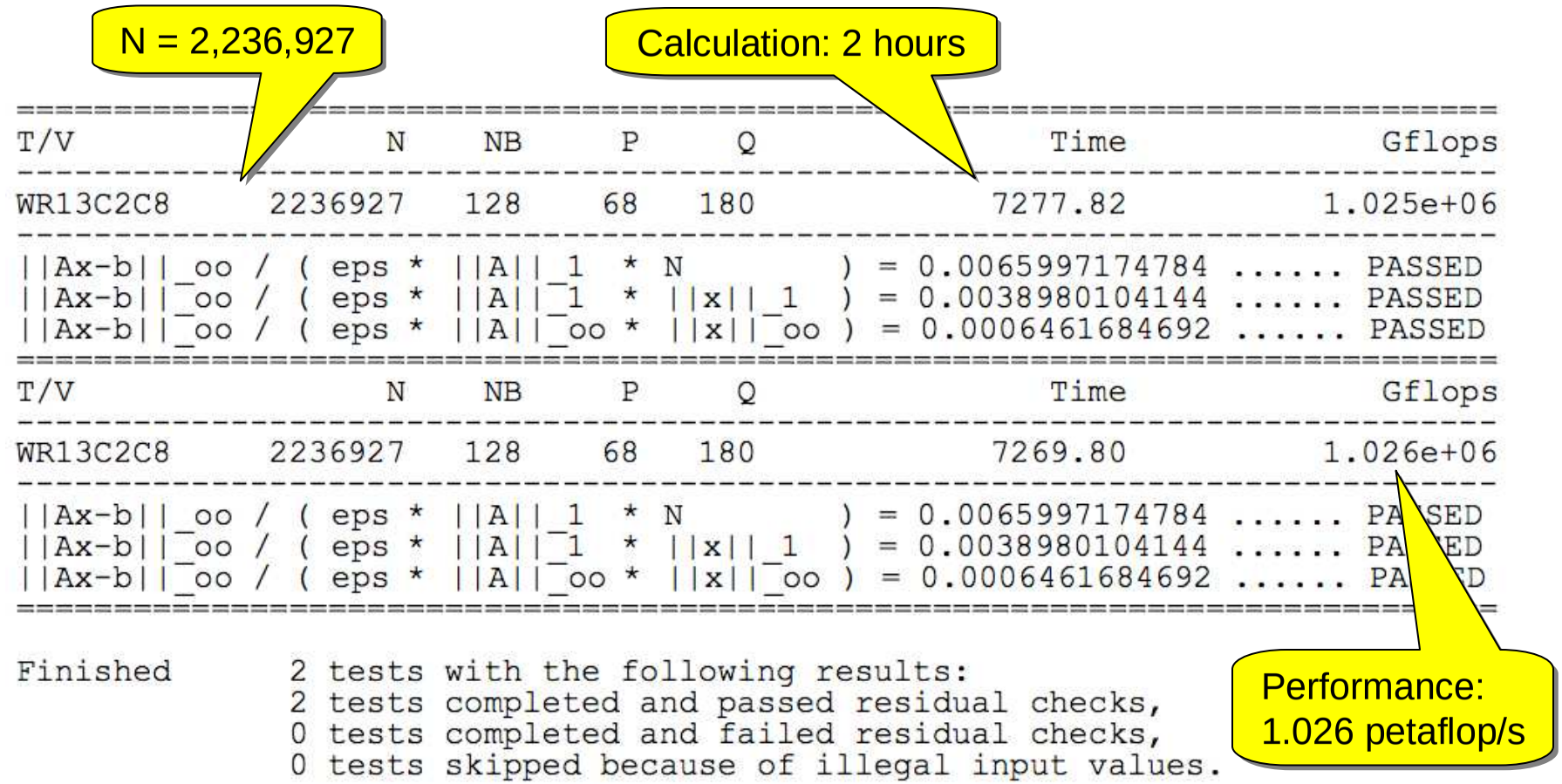
Don Grice

IBM

The partnership between IBM and Los Alamos made Roadrunner possible.

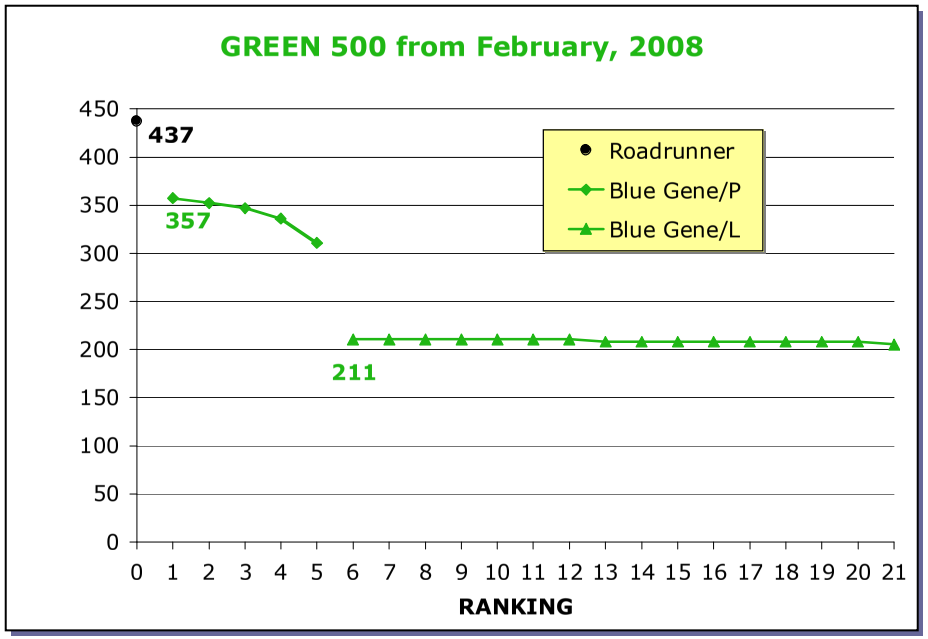
- Los Alamos began working with IBM in 2002 on the possibilities of the Cell processor
- Roadrunner was selected via a competitive procurement (2006) for a petascale supercomputer
- Over the last year we have proven that Roadrunner has great potential.
- Beginning May 23, we have begun realize that potential.

Roadrunner first achieved a petaflop/s at 3:30 am, Monday, May 26.



Achieving a petaflop/s in less than 3 days demonstrates the stability of the Roadrunner system.

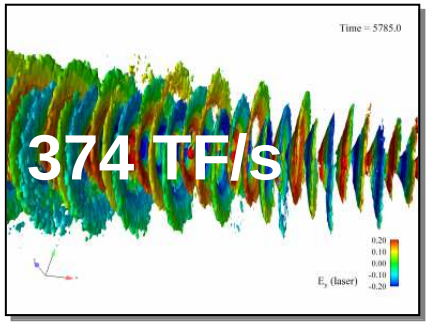
- **Full system available**
  - 8:30 am, Friday, May 23
- **Full system job launch tests begin**
  - 3:00 pm, Friday, May 23
- **First full system LINPACK launch**
  - 8:30 pm, Friday, May 23
  - Node failure after running an hour
- **Successful LINPACK runs**
  - 5:45 pm, Saturday, May 24 (879 TF/s)
  - 2:45 pm, Sunday, May 25 (945 TF/s)
  - 1:10 am, Monday, May 26 (997 TF/s)
  - 3:30 am, Monday, May 26 (Petaflop/s)

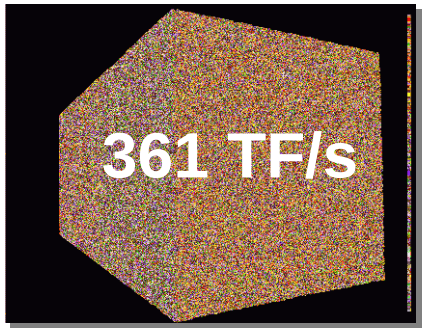


Roadrunner is also very energy efficient, 437 MF/s per watt.

We have focused on important application codes.

Code	Description
VPIC (8.5K lines)	**Fully-relativistic, charge-conserving, 3D explicit particle-in-cell code.**
SPaSM (34K lines)	**Scalable Parallel Short-range Molecular Dynamics code, orig. developed for the CM-5.**
Milagro (110K lines)	**Parallel, multi-dimensional, object-oriented code for thermal x-ray transport via Implicit Monte Carlo on a variety of meshes.**
Sweep3D (2.5K lines)	**Simplified 1-group 3D Cartesian discrete ordinates (Sn) kernel representative of the PARTISN neutron transport code.**



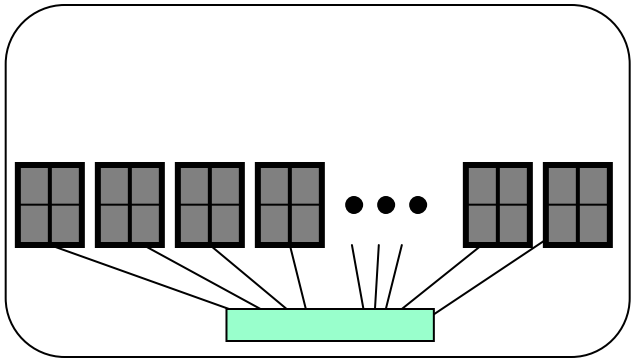


Petavision achieved 1.144 petaflop/s.

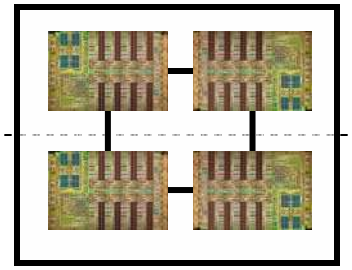
http://www.lanl.gov/news/

Roadrunner anticipates the future of supercomputing.

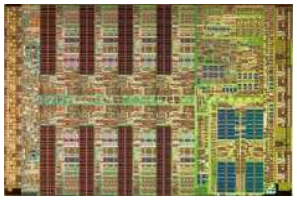
cluster



node



socket



Message Passing

Not Message Passing

Hybrid & many core technologies will require new approaches:

DaCS/ALF, libSPE,

OpenMP, PGAS, …

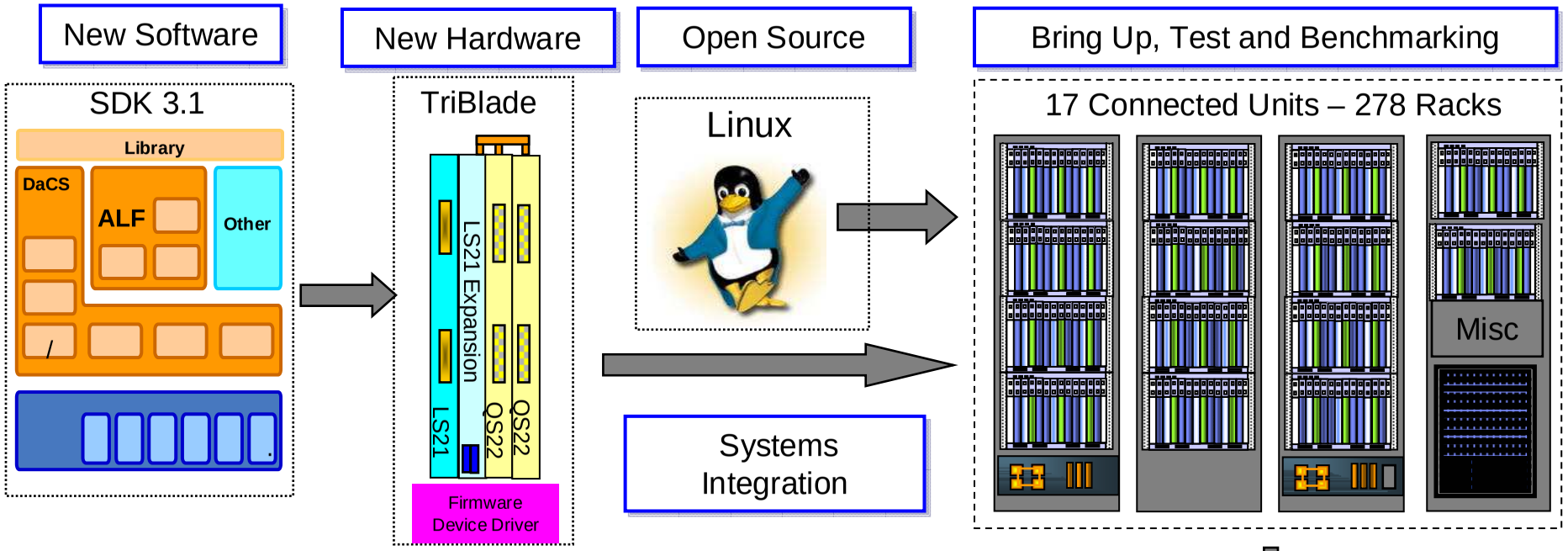
Los Alamos NATIONAL LABORATORY EST.1943

IBM

Roadrunner at a glance

- **Cluster of 17 Connected Units**
  - 12,240 IBM PowerXCell 8i accelerators
  - 6,120 AMD dual-core Opterons (comp)
  - 408 AMD dual-core Opterons (I/O)
  - 34 AMD dual-core Opterons (man)
  - 1.332 Petaflop/s peak (PowerXCell)
  - 44 Teraflop/s peak (Opteron-comp)
  - 1.026 Petaflop/s sustained Linpack
- **InfiniBand 4x DDR fabric**
  - 2-stage fat-tree; all-optical cables
  - Full CU bi-section bi-directional BW
    - 384 GB/s (CU)
    - 3.3 TB/s (system)
  - Non-disruptive expansion to 24 CUs
- **98 TB aggregate memory**
  - 49 TB Opteron
  - 49 TB Cell
- **408 GB/s peak File System I/O:**
  - 204x2 10G Ethernets to Panasas
- **RHEL & Fedora Linux**
- **SDK for Multicore Acceleration**
- **xCAT Cluster Management**
  - System-wide GigEnet network
- **2.35 MW Power (Linpack):**
  - 437 Megaflop/s per Watt
- **Other:**
  - 278 racks
  - 5200 ft^2
  - 500,000 lbs.
  - 55 miles of IB cables

Roadrunner Overview – Building Blocks



- A New Programming Model extended from standard, cluster computing
- Hybrid and Heterogeneous HW
- Built around BladeCenter and Industry IB-DDR

Next stop, exaflop?

www.lanl.gov/roadrunner

www.lanl.gov/news

www-03.ibm.com/press/us/en/pressrelease/24405.wss

www.ibm.com/deepcomputing