

LA-UR- 11-00403

Approved for public release;  
distribution is unlimited.

*Title:* Institutional Computing (IC) Information Session

*Author(s):* Ken Koch, HPC-DO  
Bryan Lally, CCS-7

*Intended for:* LANL web posting



Los Alamos National Laboratory, an affirmative action/equal opportunity employer, is operated by the Los Alamos National Security, LLC for the National Nuclear Security Administration of the U.S. Department of Energy under contract DE-AC52-06NA25396. By acceptance of this article, the publisher recognizes that the U.S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or to allow others to do so, for U.S. Government purposes. Los Alamos National Laboratory requests that the publisher identify this article as work performed under the auspices of the U.S. Department of Energy. Los Alamos National Laboratory strongly supports academic freedom and a researcher's right to publish; as an institution, however, the Laboratory does not endorse the viewpoint of a publication or guarantee its technical correctness.

## **Abstract**

### **Institutional Computing (IC) Information Session**

The LANL Institutional Computing Program (IC) will host an information session about the current state of unclassified Institutional Computing at Los Alamos, exciting plans for the future, and the current call for proposals for science and engineering projects requiring computing. Program representatives will give short presentations and field questions about the call for proposals and future planned machines, and discuss technical support available to existing and future projects. Los Alamos has started making a serious institutional investment in open computing available to our science projects, and that investment is expected to increase even more.

Intended audience: Current and Potential Institutional Computing Users

Date & Time: Tuesday January 18, 1:00pm - 2:00pm

Location: Cochiti room, Oppenheimer Study Center

URL: [http://int.lanl.gov/orgs/adtsc/Institutional\\_Computing/inst\\_comput.shtml](http://int.lanl.gov/orgs/adtsc/Institutional_Computing/inst_comput.shtml)

# Institutional Computing (IC) Information Session

January 18, 2010

Ken Koch, HPC-DO, LANL

Bryan Lally, CCS-7, LANL

Andy White (IC Program Director)

Ken Koch & Bob Tomlinson (IC planning and support)

Bryan Lally & Mike Wall (programming support contacts, [ic-help@lanl.gov](mailto:ic-help@lanl.gov))



Operated by the Los Alamos National Security, LLC for the DOE/NSA



## Outline

---

- Summary of Institutional Computing Resources and Future Plans
- Summary of the IC 2011 Call for Proposals
- Programming Support
- Questions and Answers



Operated by the Los Alamos National Security, LLC for the DOE/NSA



## IC Computing Resources

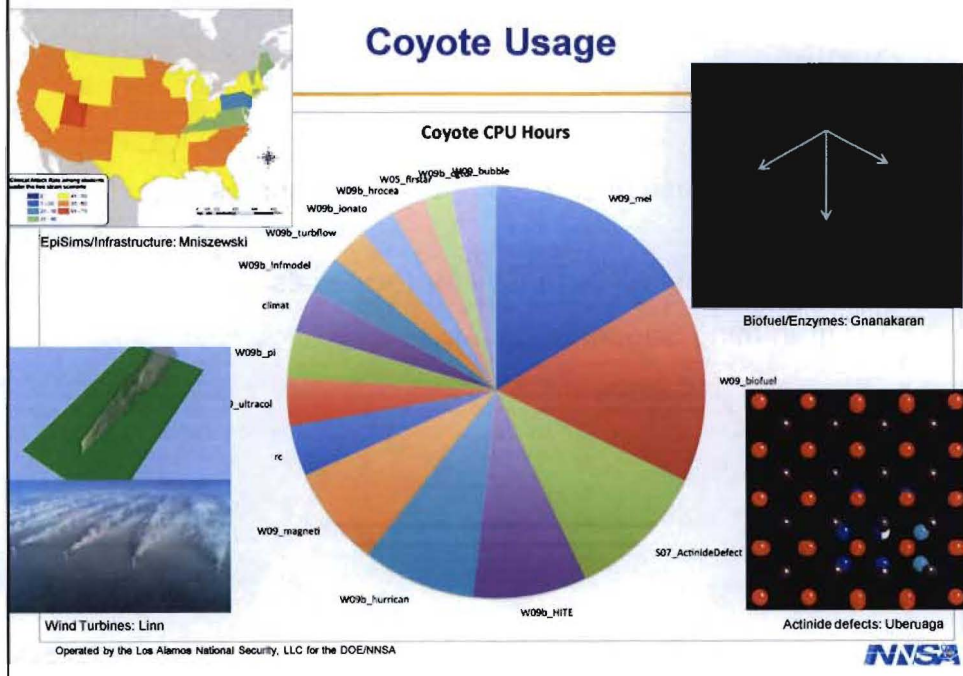
All IC machines are in the LANL Turquoise network!

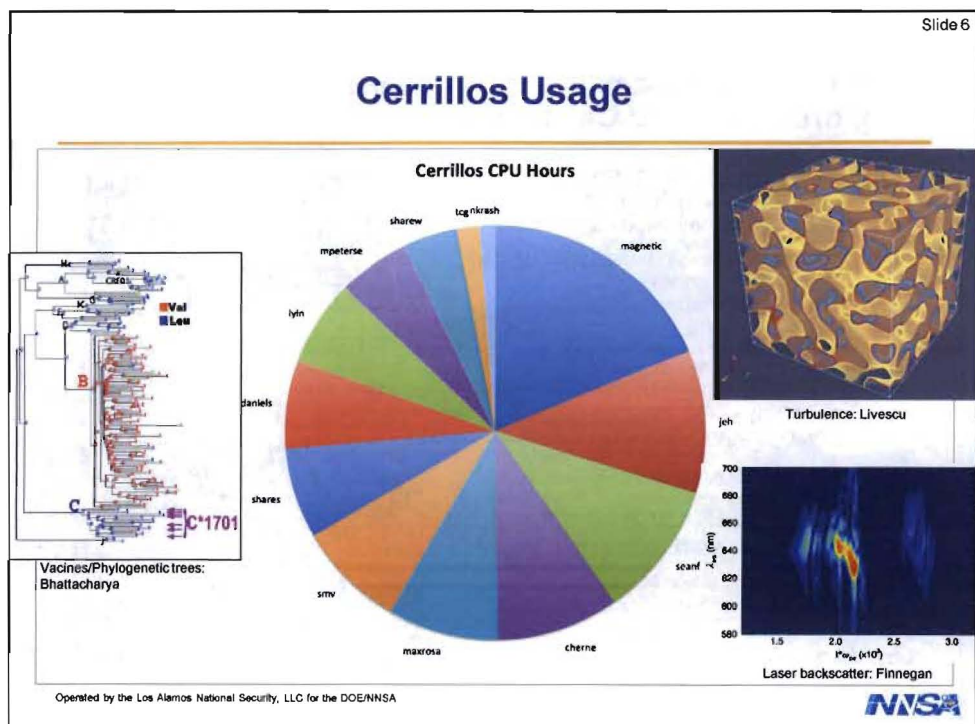
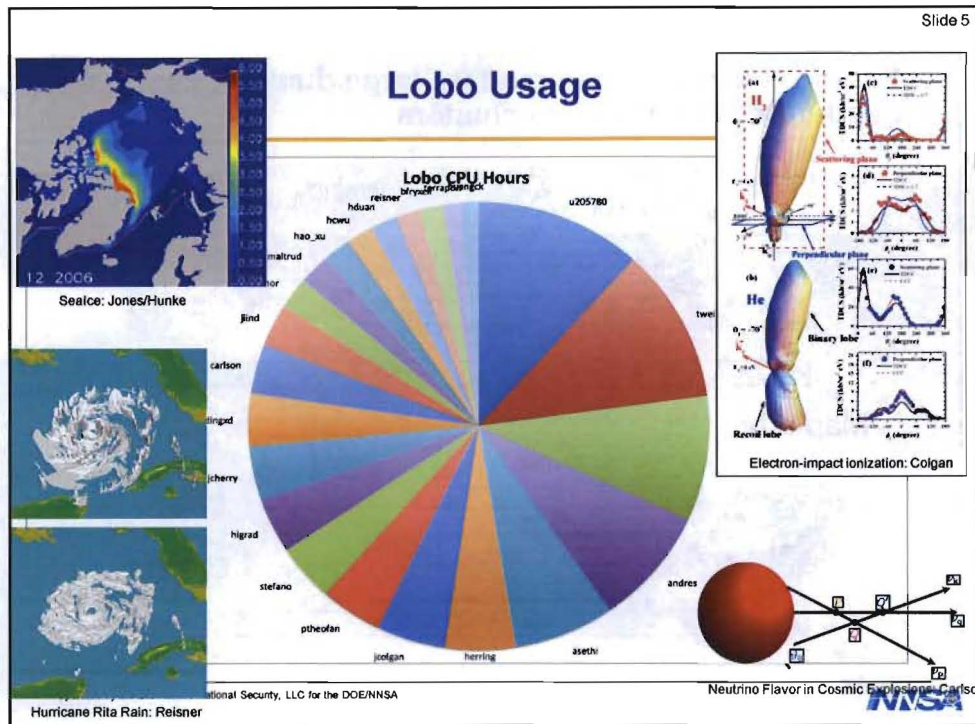


EST. 1943  
Operated by the Los Alamos National Security, LLC for the DOE/NNSA



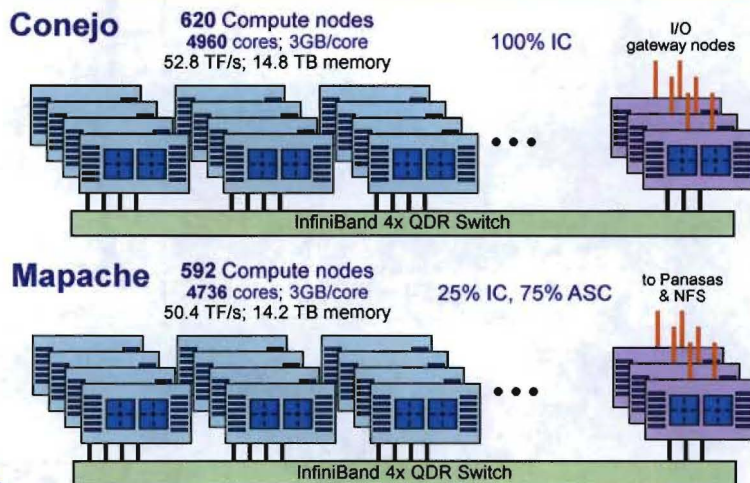
## Coyote Usage







## Conejo & Mapache are modern large dual-socket quad-core Intel Nehalem clusters



Los Alamos  
NATIONAL LABORATORY  
EST. 1943

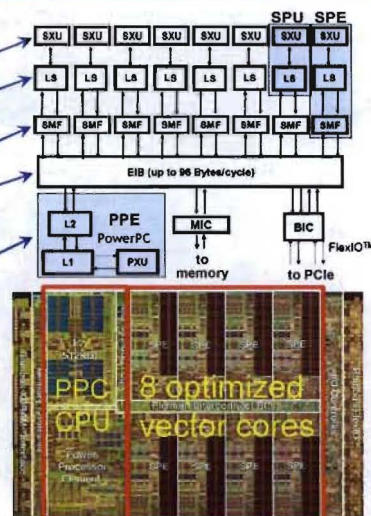
Operated by the Los Alamos National Security, LLC for the DOE/NSA

NSA

## The IBM PowerXcell 8i Cell processor is the key to Roadrunner's and Cerrillos' performance

- Cell Broadband Engine (CBE\*) was developed by Sony-Toshiba-IBM
  - used in Sony PlayStation 3
- 8 Synergistic Processing Elements (SPEs)
  - 128-bit **vector cores**
  - 256 kB **local memory** (LS = Local Store)
  - Direct Memory Access (**DMA**) engine (25.6 GB/s each)
  - Chip interconnect (**EIB**)
  - Run SPE-code as POSIX threads (SPMD, MPMD, streaming)
- 1 PowerPC **PPE** runs Linux OS

**Design: SPEs provide optimal flop/s per watt in minimal area**  
**This is an Exascale trend**



Los Alamos  
NATIONAL LABORATORY  
EST. 1943

\* trademark of Sony Computer Entertainment, Inc.

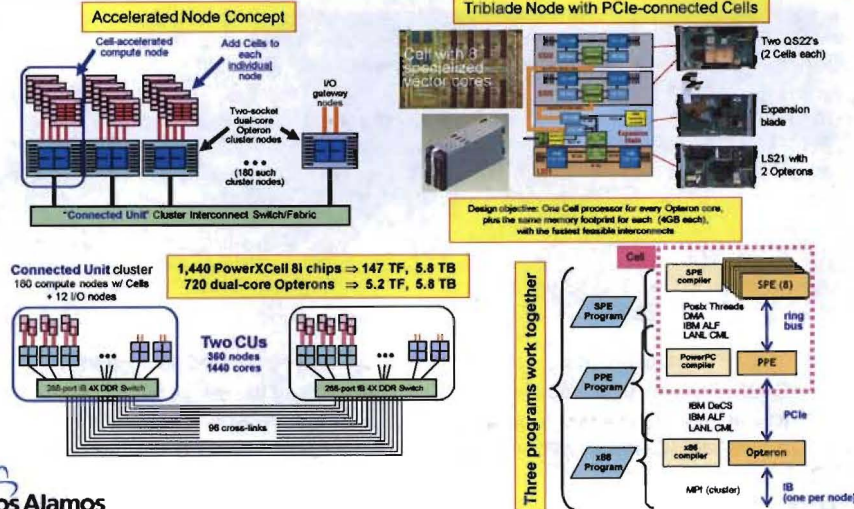
Operated by the Los Alamos National Security, LLC for the DOE/NSA

NSA

## Cerrillos is a small Roadrunner (12%)

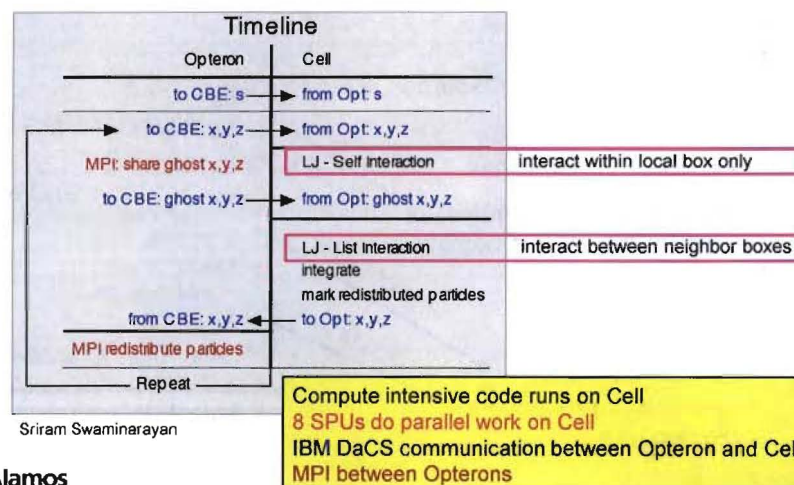


Slide 9



## Roadrunner Program Example: SPaSM Molecular Dynamics with Lennard-Jones

Slide 10





## IC Resources and Anticipated Transition Timelines

	Coyote	Lobo	Cerrillos	Conejo	Mapache
CPU Cores	2580 total (5 * 516)	4352	1440	4960	4736
IC CPU Hours per Year †	22,600,800	9,530,880 28,592,640 (ASC)	12,614,400	43,449,600	10,371,840 31,115,520 (ASC)
Timeline	Retire in March	Moved to Red this spring	Utilized thru FY13 or FY14?	Existing IC users by early- February New projects by mid-to-late February	Existing IC users by early- February New projects by mid-to-late February

† COSIM (climate program) gets 17.2% of all cycles for their funding support  
 CASL (DOE multi-lab nuclear energy hub) gets 10% of the remaining cycles  
**IC will allocate the rest to new selected proposals from this call**  
 - 32.4M (Conejo), 7.7M (Mapache), & 9.4M (Cerrillos) CPU-hrs per year



Operated by the Los Alamos National Security, LLC for the DOE/NSA



## IC has an Aggressive Future Strategy, Motivated by an Increased Interest and Support for Open Computing by the Lab Director

- Target 10% of the largest national open science resources

- This is the scale of resources available to ASCR Leadership Computing Facilities: ORNL & ANL

- Procure Mustang, a 200+ TF/s system for deliver this coming August/September!

- Follow two advanced technology tracks

- GPU Accelerators
  - Low-power processors
  - None of the above

- Current

- Jaguar at ORNL (2.3 PF/s)

- FY12-13

- "OLCF-3/Titan" at ORNL
    - Cray: Opterons + nVidia GPUs
    - 20 PF/s
  - "MiRA" at ANL
    - IBM Blue Gene Q
    - 10 PF/s

- FY15-16 (DOE Exascale Initiative)

- 200 PF/s pre-exascale systems
  - Almost certainly accelerated
  - 10% = 20PF/s



Operated by the Los Alamos National Security, LLC for the DOE/NSA





## Other Current & Recent Institutional and Turquoise Network Improvements

- **Systems:**
  - **Viz: Lightshow** (replacing Phoenix) (ASC funded)
    - 24 nodes Intel Westmere, with 96 GB/node of memory & dual nVidia Quadro 6000's
  - **IS&T: New Data Intensive Super Computing (DISC) systems** (director funded)
    - Built from old HPC systems w/ added memory and disk and flash memory
  - Also in the open, but not an IC system: **New large Hybrid CPU+GPU testbed** cluster with nVidia & ATI GPUs (in CCS-7 & HPC-5)
- **Facilities:**
  - Upgrades to the LDCC are planned & future facilities discussions are starting
- **Turquoise Network Storage & Archive:**
  - **Panasas upgrades:**
    - 24 next-gen 40 TB shelves (960 TB @ 15 GB/s) (existing is 640 TB)
    - New "director blades" for improved access performance
    - All 10 GigE connected
  - **NFS upgrades:**
    - Moving from NetApps to GPFS
    - NFS v4 testbed (better data protection)
  - **Archive (tape) upgrades:** 2 PB → 3 PB
  - **I/O backbone** (PaScalBB) will be upgraded to dual-lanes when Coyote's gone
- **Network / bandwidth:**
  - **Much improved Yellow ↔ Turquoise bandwidth** through "Fast Data Pipe" project
    - Turq archive → Yellow Panasas: 440 – 485 MB/s
    - Yellow Panasas → Turq archive: 662 MB/s
  - **10 Gb/s ESnet** fully replacing Internet2

## IC 2011 Call for Proposals

## IC 2011 Call for Proposals

- Proposals are due back Monday, January 24th by midnight via email to [ic\\_proposal@lanl.gov](mailto:ic_proposal@lanl.gov)
  - 6 pages in PDF format using the Word template provided
  - Proposals are for 2 years duration unless a shorter term is requested
  - [https://int.lanl.gov/orgs/adts/Institutional\\_Computing/inst\\_comput.shtml](https://int.lanl.gov/orgs/adts/Institutional_Computing/inst_comput.shtml) or the LANL internal home page "What's Happening" link
- Proposals will be peer-reviewed and scored, and the strongest and best proposals selected
- Aggregate allocations will be made consisting of:
  - 32.4 million CPU-hrs per year on Conjeo for 2 years
  - 7.7 million CPU-hrs per year on Mapache for 2 years
  - 9.4 million CPU-hrs per year on Cerrillos for 2 years



Operated by the Los Alamos National Security, LLC for the DOE/NNSA



## IC Proposal Selection Outcomes

- Our process is competitive and our intent, as in the past, is to enable all reasonable proposed projects to succeed and improve.
- Our ability to accomplish this will depend on the totality of requests received relative to our resources.
- There are three levels of allocations:
  - *There will be full 2-year allocations for the best-reviewed proposals.*
  - *There will be 1-year allocations for some excellent proposals that did not make the above group and/or have significant promise, but need additional time or effort to realize this potential.*
  - *There will be modest allocations for proposals that require substantial additional computational science development in order to help them compete better in the future.*



Operated by the Los Alamos National Security, LLC for the DOE/NNSA



## Job CPU-hr Estimates

---

- A CPU-hour is also a core-hour of use.
  - *Conejo and Mapache jobs will all be run on whole nodes (8 cores) dedicated to that job*
    - 1 wall clock hour amounts to 8 CPU-hrs per node
  - *Cerrillos jobs are all run on whole nodes (4 Opteron cores plus 4 Cells as accelerators) dedicated to that job*
    - 1 wall clock hour amounts to 4 CPU-hrs per node (not 8)
- Conejo & Mapache should be faster than Coyote & Lobo
  - *Clock speeds are similar, yet the memory sub-systems are considerably faster in the new Intel Nehalem nodes*
  - *For CPU-hr estimates on a core-per-core basis use:*
    - 1.5x faster than Lobo (e.g. runs in 2/3 the time)
    - 2.0x faster than Coyote (e.g. runs in 1/2 the time)
    - YMMV (Your Mileage May Vary)
    - If you have Nehalem performance estimates for your code, by all means use them.



Operated by the Los Alamos National Security, LLC for the DOE/NSA



## Programming Support



Operated by the Los Alamos National Security, LLC for the DOE/NSA



## User Group

---

- After discussions with several IC projects ...
  - *Create a User Group for IC projects and computing*
  - *To discuss and solve common problems*
  - *Begin with monthly meetings, adjust as needed*
  - *Communications by email, wiki, meetings*
  
- First Meeting: February 22<sup>th</sup> (time and place TBD)
  - *New project list*
  - *Machine updates*
  - *Fast data transfer to/from turquoise/yellow networks*



Operated by the Los Alamos National Security, LLC for the DOE/NSA



## Institutional Computing Help (ic-help)

---

- You have (some) support
- Support comes in the form of a group of people with experience programming large clusters, and multi-core, Cell-accelerated, and GPU-accelerated machines
- Many from CCS-7 (Applied Computer Science), but from elsewhere around CCS and LANL as needed
- Support is requested through [ic-help@lanl.gov](mailto:ic-help@lanl.gov)



Operated by the Los Alamos National Security, LLC for the DOE/NSA





## What support can and can't do

---

### What support **isn't**:

- Standard support provided by ICN help desk
  - *Can't log in? Need an account? Network is down? Queues are stuck? 665-4444 option 3!*
- Help with science – you are the domain experts



EST. 1943  
Operated by the Los Alamos National Security, LLC for the DOE/NSA



## Can and can't, continued

---

### What support **is**:

- Design – how to take advantage of large advanced architectures
  - *Cell*
  - *Multi-core*
  - *GPU*
  - *Languages and libraries such as Cell SDK, OpenCL, CUDA, pthreads, Domain Specific Languages, etc.*
- Help with implementation and porting to new machines – help getting codes up and running
  - *Most importantly, **running effectively!***



EST. 1943  
Operated by the Los Alamos National Security, LLC for the DOE/NSA



## Machine Summary

- Coyote going away in spring with Lobo to follow
- Cerrillos fully operational (small Roadrunner)
- Conejo & Mapache coming online in February
- Mustang operational December 2011
- Two additional systems targeted for 2012/2013 timeframe
  - Cray GPU-accelerated system (1PF/s)
  - IBM Blue Gene/Q (1 PF/s)



Operated by the Los Alamos National Security, LLC for the DOE/NNSA



## Contact Information

- IC Users Group email: [icug@lanl.gov](mailto:icug@lanl.gov)
- IC Wiki: <http://ic-help.lanl.gov/>
  - Has link to the Institutional Computing page that has the CFP and template
- IC Help email: [ic-help@lanl.gov](mailto:ic-help@lanl.gov)



Operated by the Los Alamos National Security, LLC for the DOE/NNSA



---

## Questions and Answers



EST. 1943  
Operated by the Los Alamos National Security, LLC for the DOE/NSA



---

## Some Q&A's already seen

- Can a group of investigators combine multiple projects into a single proposal?
  - *It is preferred that separate science areas submit separate proposals, even if the projects use the same code, or engage the same people. This helps in the selection process and in the allocation of resources.*
- Can post-docs be PIs?
  - Yes
- Do PIs need to be from LANL?
  - *The primary PI needs to be from LANL. Co-I's could be from other institutions.*



EST. 1943  
Operated by the Los Alamos National Security, LLC for the DOE/NSA



## Q&A

---

- Should I target a particular machine?
  - *Unless there is a compelling reason for needing Conejo versus Mapache, we'll likely try to provide projects allocation on one or the other.*
  - *Cerrillos clearly requires different programming skills and commitment and/or an already existing application. We expect fewer proposals to compete for Cerrillos allocations, and thus those projects targeting Cerrillos have a great opportunity for significant cycles.*



EST. 1943  
Operated by the Los Alamos National Security, LLC for the DOE/NSA



---

**The End**



EST. 1943  
Operated by the Los Alamos National Security, LLC for the DOE/NSA

