Porting and Performance of the Community Climate System Model (CCSM3) on the Cray X1

CUG05

May 19, 2005

Albuquerque, NM

ornl

NCAR

Co-Authors

- National Center for Atmospheric Research (NCAR)
  - George R. Carr Jr, gcarr@ucar.edu
- Oak Ridge National Laboratory (ORNL)
  - John B. Drake, drakejb@ornl.gov
  - Michael W. Ham, hammw@ornl.gov
  - Forrest M. Hoffman, forrest@climate.ornl.gov
  - Patrick H. Worley, worleyph@ornl.gov
- Cray, Inc. (CRAY)
  - Matthew J. Cordery, mcordery@cray.com

… and too many others to list

ornl

NCAR

Acknowledgements

- The National Center for Atmospheric Research is funded by the National Science Foundation.
- ORNL research sponsored by the Atmospheric and Climate Research Division and the Office of Mathematical, Information, and Computational Sciences, Office of Science, U.S. Department of Energy under Contract No. DE-AC05-00OR22725 with UT-Battelle, LLC.
- These slides have been coauthored by a contractor of the U.S. Government under contract No. DE-AC05-00OR22725. Accordingly, the U.S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or allow others to do so, for U.S. Government purposes
- Oak Ridge National Laboratory is managed by UT-Battelle, LLC for the United States Department of Energy under Contract No. DE-AC05-00OR22725.

ornl

NCAR

Overview

- CCSM3 Introduction
- Cray X1 Introduction
- An Orientation to the Port and Evaluation Process
- Status of the Validation Process
- Future Activities

ornl

NCAR

CCSM Introduction

- CCSM, the Community Climate System Model, is a coupled model for simulating the earth's climate system.
  - Developed at NCAR with significant collaborations with DOE, NASA and the university community
- Components in CCSM3 include
  - Atmospheric Model – CAM 3.0

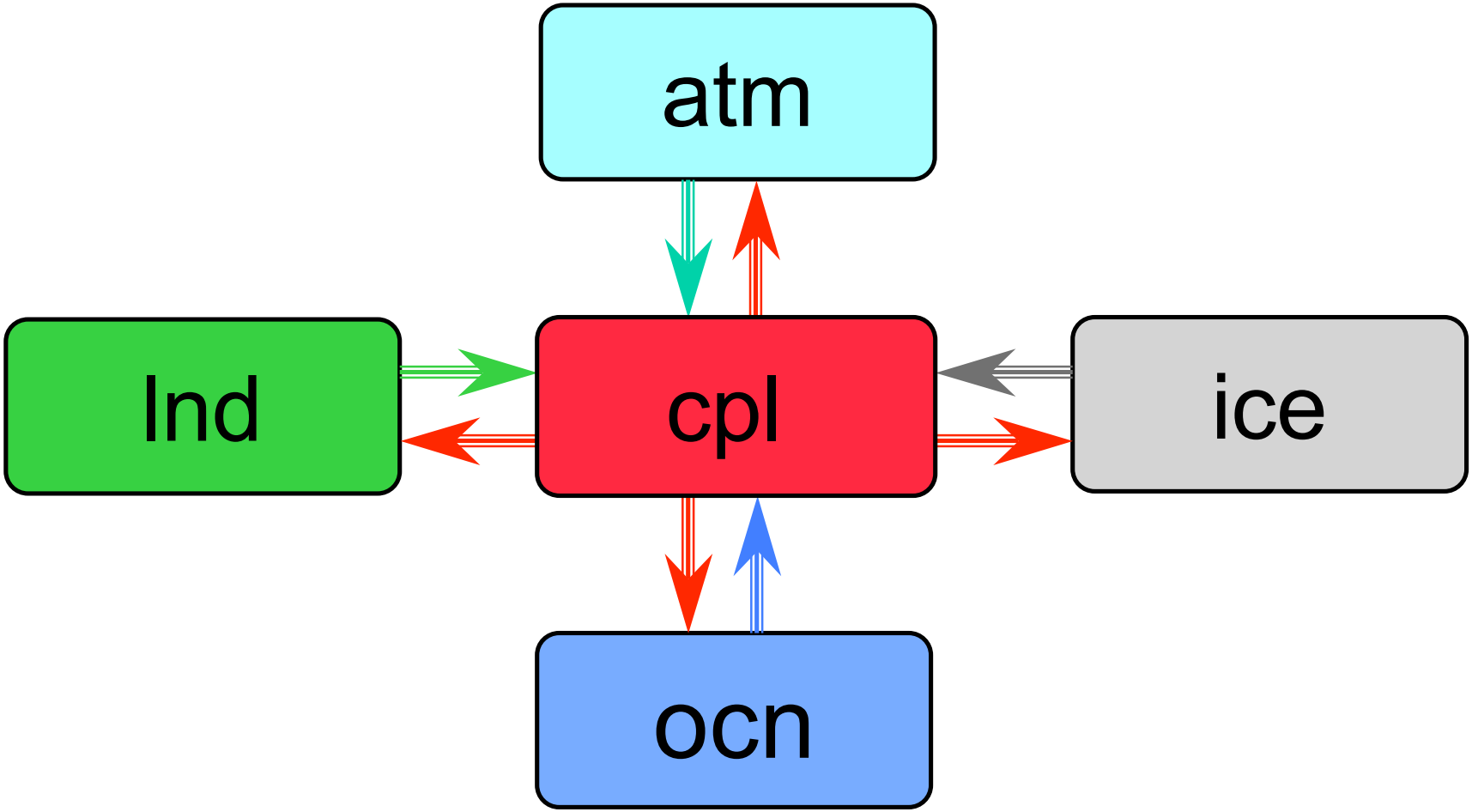
    T31: (48 x 96 x 26) **T42**: (64 x 128 x 26) **T85**: (128 x 256 x 26)
  - Ocean Model – modified version of POP 1.4.3

    3 degree: (100 x 116 x 25) **1 degree**: (320 x 384 x 40)
  - Sea Ice Model – CSIM5 - grid matches ocean
  - Land Model – CLM3 - grid matches atmosphere
  - Coupler - CPL6

ornl

NCAR

CCSM Hub and Spoke



ornl

NCAR

Supported Machines

- IBM Power3, Power4 - fully validated
- Xeon Linux Clusters (GigE and Myrinet) - validated T31x3
- Cray X1 - recently validated T31x3, just starting T85x1
- SGI Altix - baseline validation complete for T31x3
- Earth Simulator - fully validated on pre-release, update planned
- Opteron Linux Clusters (Myrinet) - work well underway
- Xeon Linux Clusters (InfiniBand) - work continuing
- Cray XT3 and XD1 - work begun

Approximate Timelines (1 of 2)

=> December 2003:

- Component model vectorization

=> April, 2004:

- Merge of vector versions into development branch, including basic support for the X1
- CAM/CLM2 standalone model (spectral Eulerian dycore) validated on the Earth Simulator and X1

=> June, 2004:

- CCSM validated on Earth Simulator and achieves required percentage of vectorization
- CCSM3 released, including basic support for X1

ornl

NCAR

Approximate Timelines (2 of 2)

=> October, 2004:

- Successful initial T31x3 X1 validation

=> March/April, 2005:

- Failed second T31x3 X1 validation

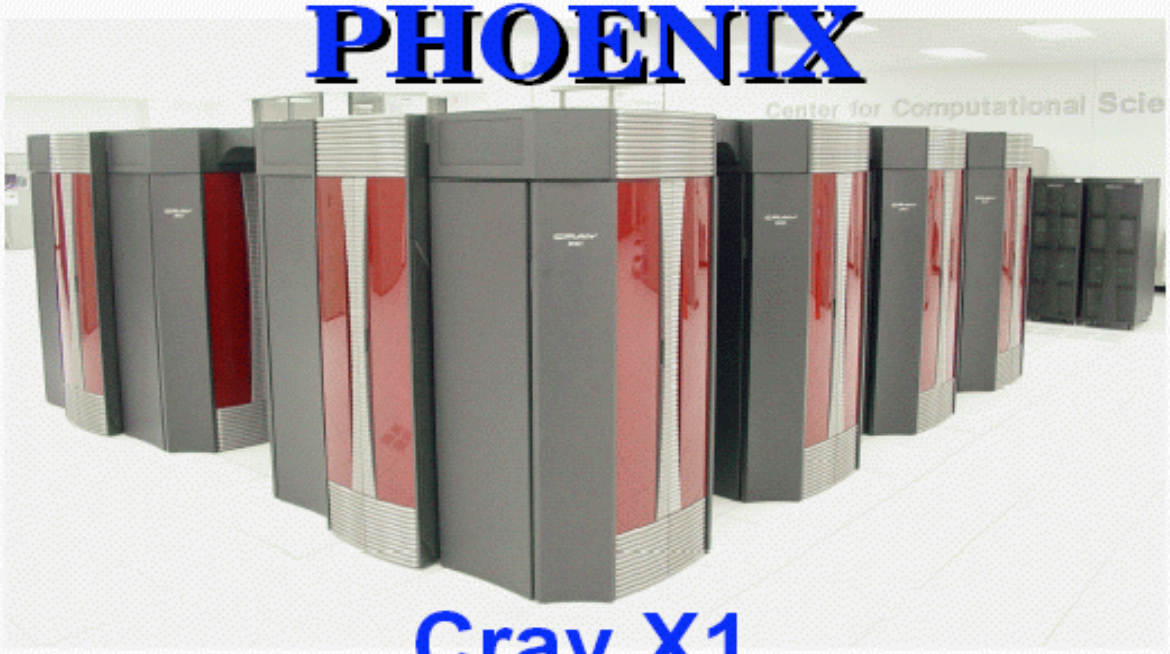
=> May, 2005:

- Recast second T31x3 X1 validation attempt
- CAM T170 runs begun

ornl

NCAR

PHOENIX



Cray X1

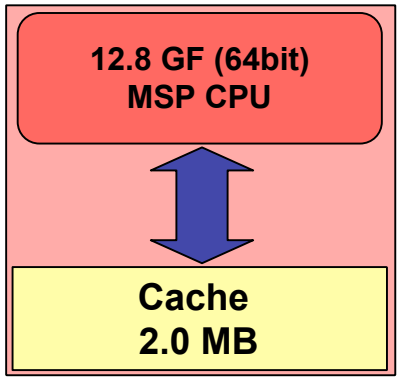
- 128 SMP nodes
- 4 Multi-Streaming Processors (MSPs) per node
- 4 Single Streaming Processors (SSPs) per MSP
- Two 32-stage, 64-bit wide vector units running at 800 MHz and one 2-way superscalar unit running at 400 MHz per SSP
- 2 MB E-cache per MSP
- 16 GB of memory per node

512 processors (MSPs), 2048 GB of memory, and 6400 GFlop/s peak

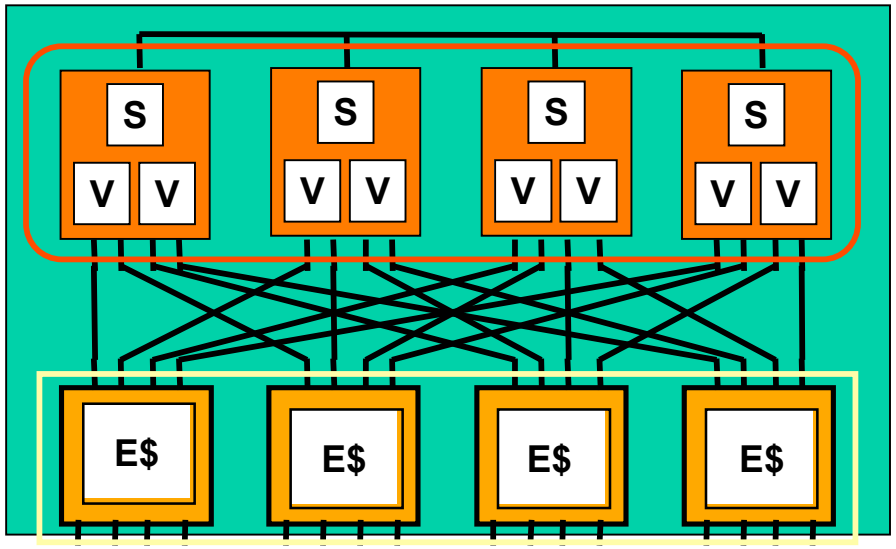
ornl

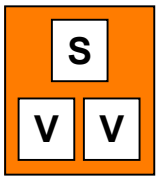
NCAR

Multistreaming Processor



=





SSP – Single-Streaming Processor

- Two vector pipe units
- One 4-way superscalar processor

ornl

NCAR

CCSM Validation Process

- Get code to build and run
- CAM Perturbation Growth test (PERGRO)
- Atmospheric Diagnostics test on CAM and/or CCSM CAM history files
- CCSM 100 year statistical test
- Harden the scripts and do CCSM component load balancing

ornl

NCAR

CAM PERGRO

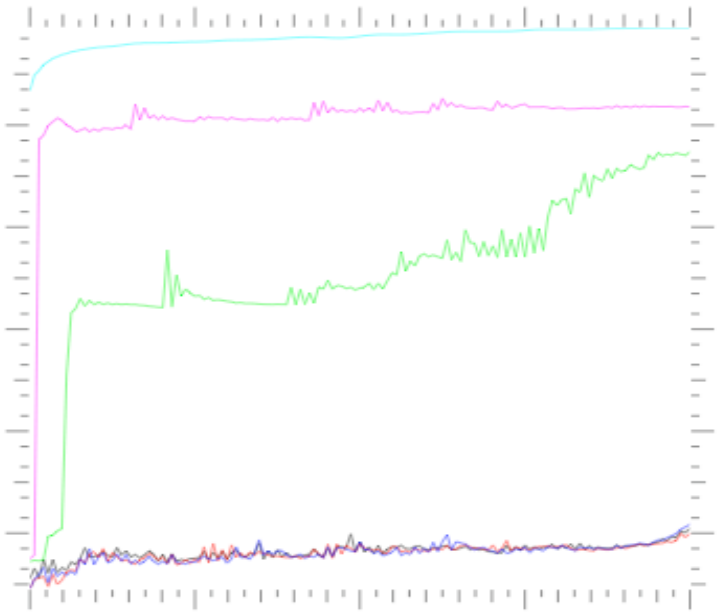
- Establish known baseline
- Perturb initial data
- Roundoff errors:
  - Within PERGRO limits: great
  - Close to PERGRO: probably ok
  - Way off: reject

- PERGRO is a simple two simulation day run for each configuration being tested.
- See http://www.ccsm.ucar.edu/models/atm-cam/port/

ornl

NCAR

CAM PERGRO Example



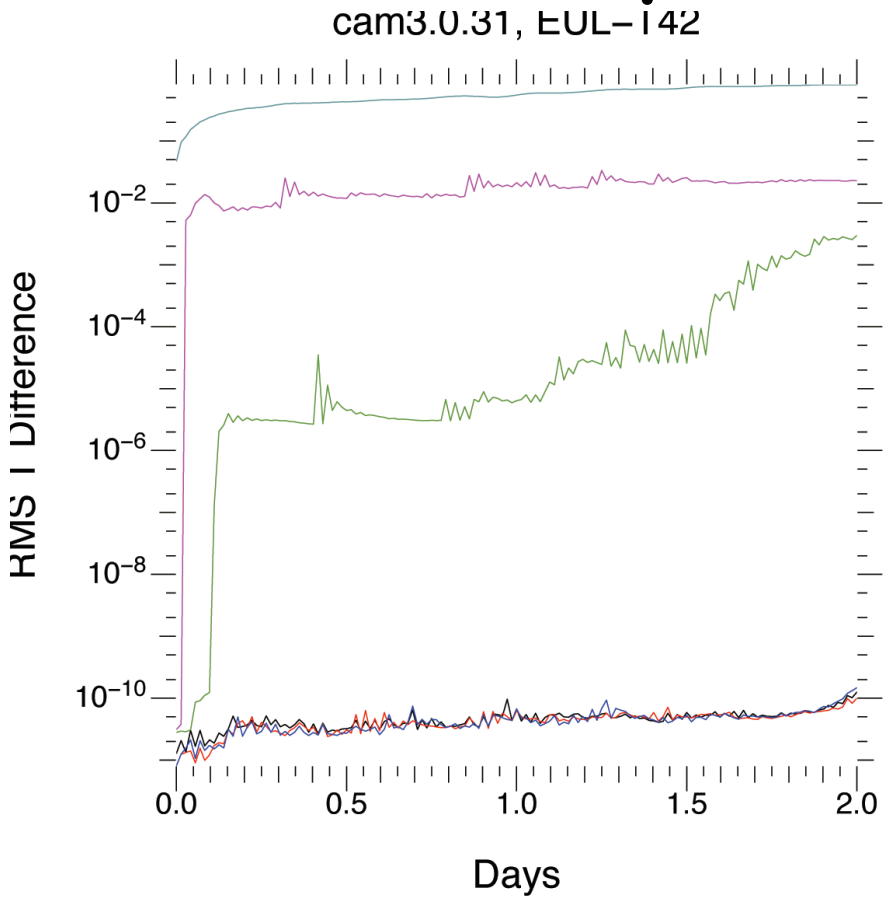
Days

perturbation (ibm–P4)
difference: ibm – sgi–O3800
difference: ibm – x86(lf95)
difference: ibm – x86(pg–513 –O1)
difference: ibm – amd64(pg64–524 –O1)
difference: ibm – amd64(pg64–524 –O0)

ornl

NCAR

CAM PERGRO Example



perturbation (ibm–P4)
difference: ibm – sgi–O3800
difference: ibm – x86(lf95)
difference: ibm – x86(pg–513 –O1)
difference: ibm – amd64(pg64–524 –O1)
difference: ibm – amd64(pg64–524 –O0)

ornl

NCAR

Atmospheric Diagnostic Tests

- Compare 5-100 years of test configuration to 100 years of accepted baseline
- Computes monthly averages for test and baseline
- Monthly averages used to generate large number of plots and graphs
- Find Oracle to perform evaluation

- See http://www.ccsm.ucar.edu/models/atm-cam/port/

ornl

NCAR

CCSM Statistial Test

- Requires 100 years of baseline
- Requires 100 years of test configuration
- Set of rigorous statistical analysis of the two data sets

ornl

NCAR

Two Kinds of "Load Balancing"

- CCSM load balancing: assigning right number of processors for each component
- Classic load balancing: moving processing around to even out execution times

ornl

NCAR

The CCSM MPMD Balancing Act

- Each component has different scaling attributes in part based on different grid sizes
- System architecture/configuration constraints

ornl

NCAR

Load Balancing Example - X1

T31x3	OCN	ATM	ICE	LND	CPL	Tot	Yrs/Day
Case 1	4	16	8	8	4	40	20.76
Case 2	2	16	2	8	8	36	22.12

Case 2 used fewer processors and got better performance

ornl

NCAR

Current X1 Validation Activities

- Move to PE5302 and minimal change set (nomodinline) and addition of fp1 - validated
- Add CAM physics load balance option changes - ok
- Better CAM vector length ("pcols") - ok
- Better CLM vector length ("clump" or CSD) - must choose
- CLM use of vector3, scalar3, stream3 appears ok
- CAM limited use of vector3, scalar3, stream3 might be ok, may not be ok for CAM as a whole
- Change to ice_transport_remap postponed

ornl

NCAR

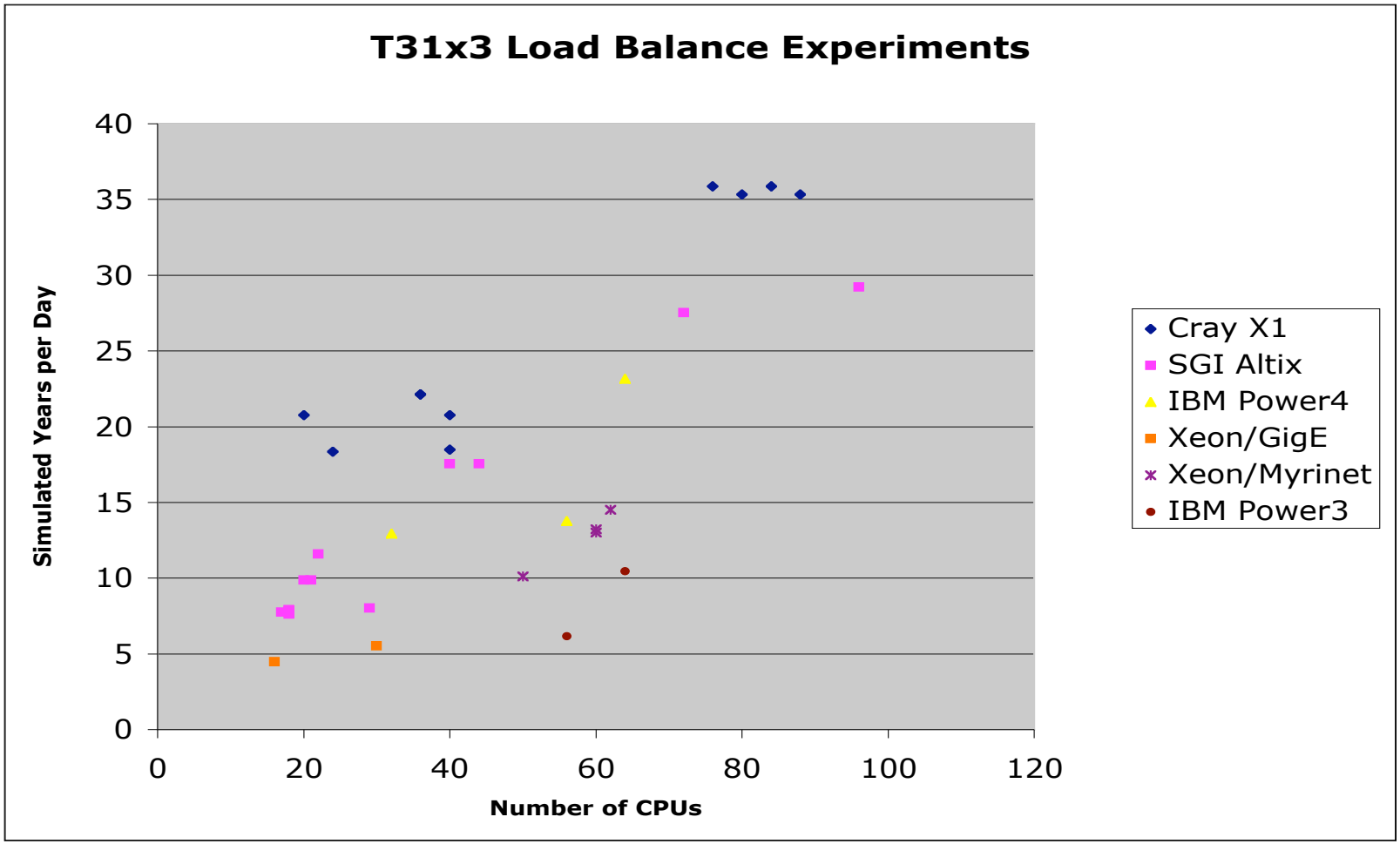
Performance Metrics

- Simulated years per wall clock day
  - Optimize for single job maximum performance (capability oriented)
- Simulated years per wall clock day per "cpu"
  - Optimize for system aggregate performance (capacity oriented)

ornl

NCAR

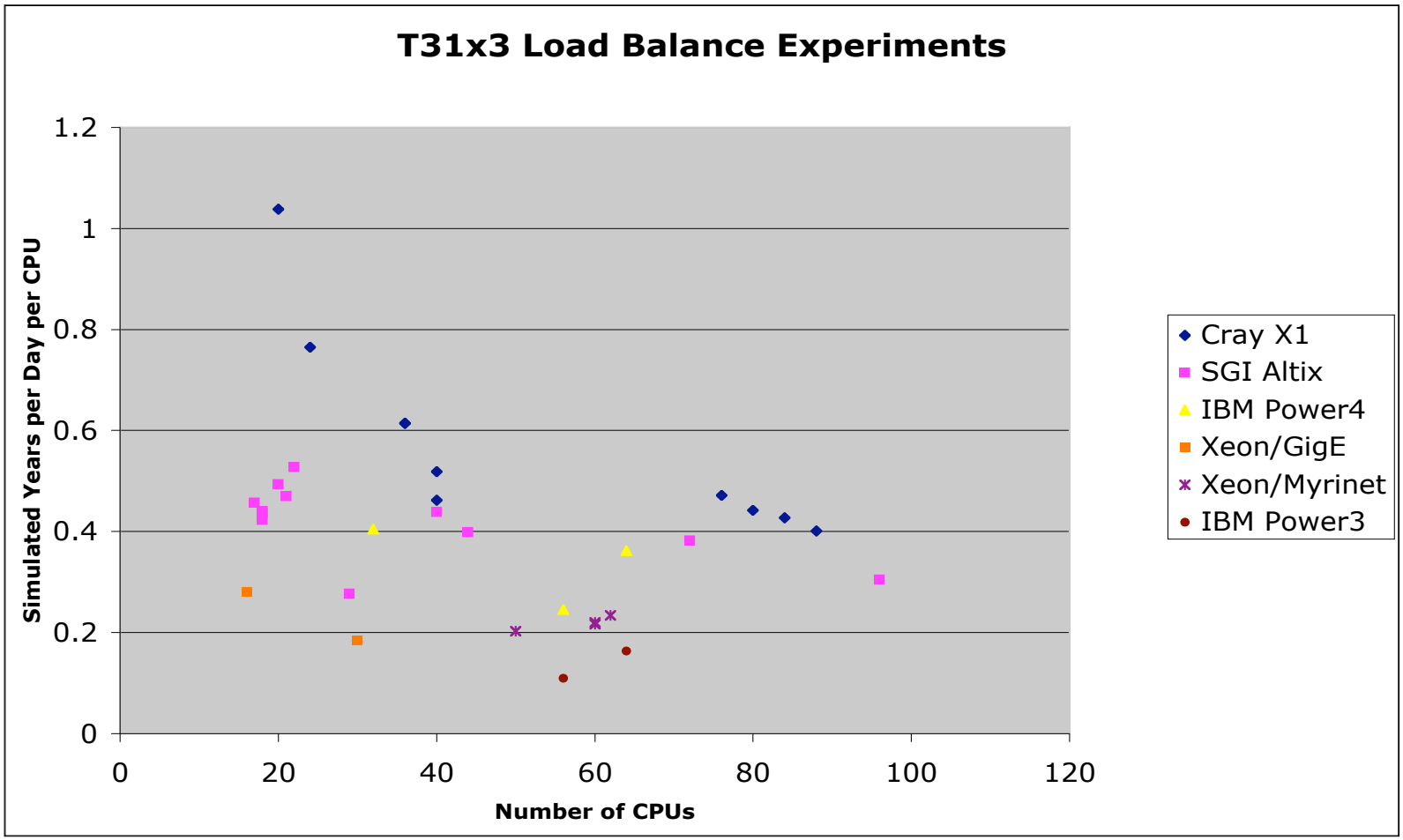
T31 Performance



ornl

NCAR

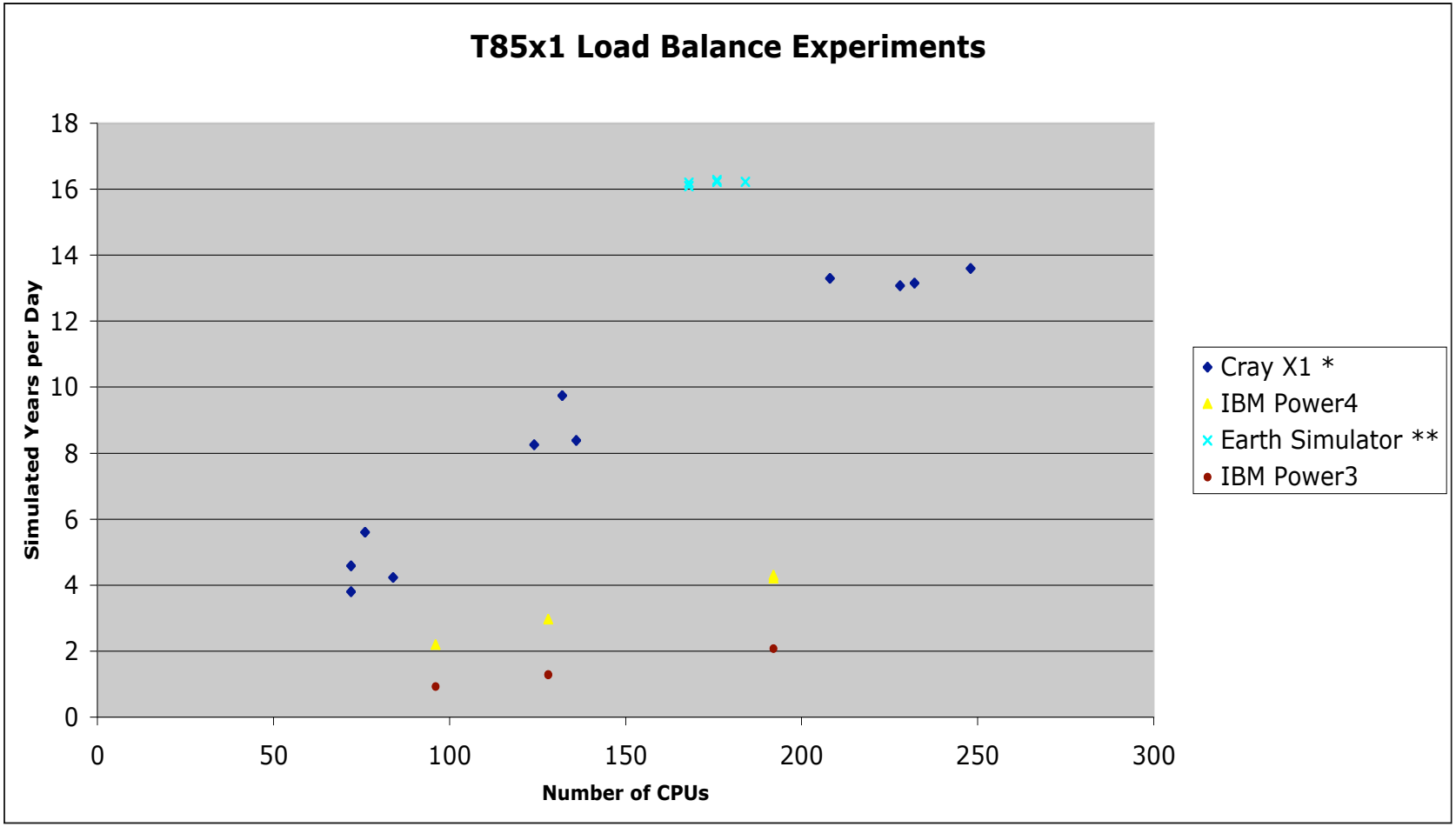
T31 Efficiency



ornl

NCAR

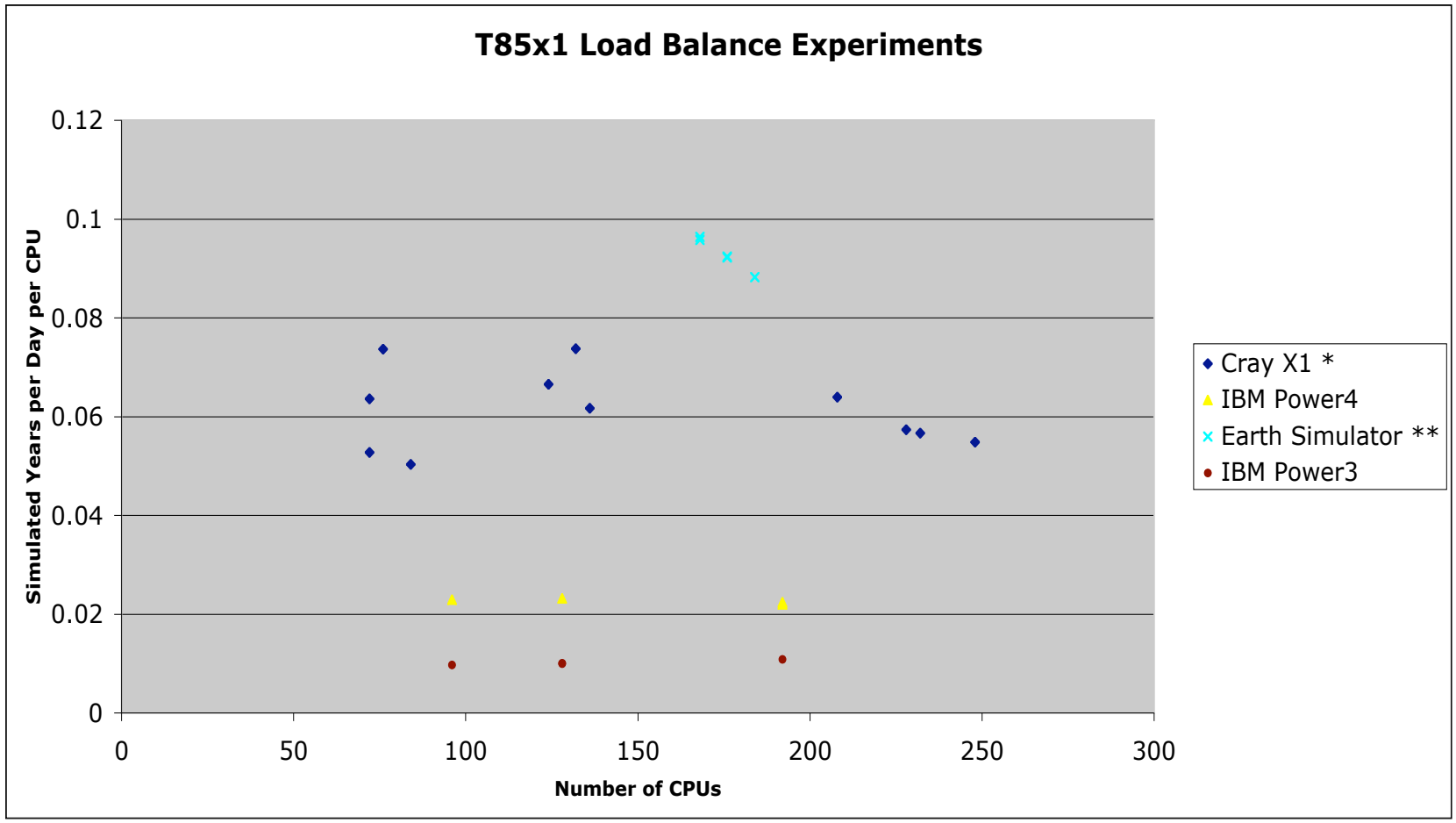
T85 Performance



ornl

NCAR

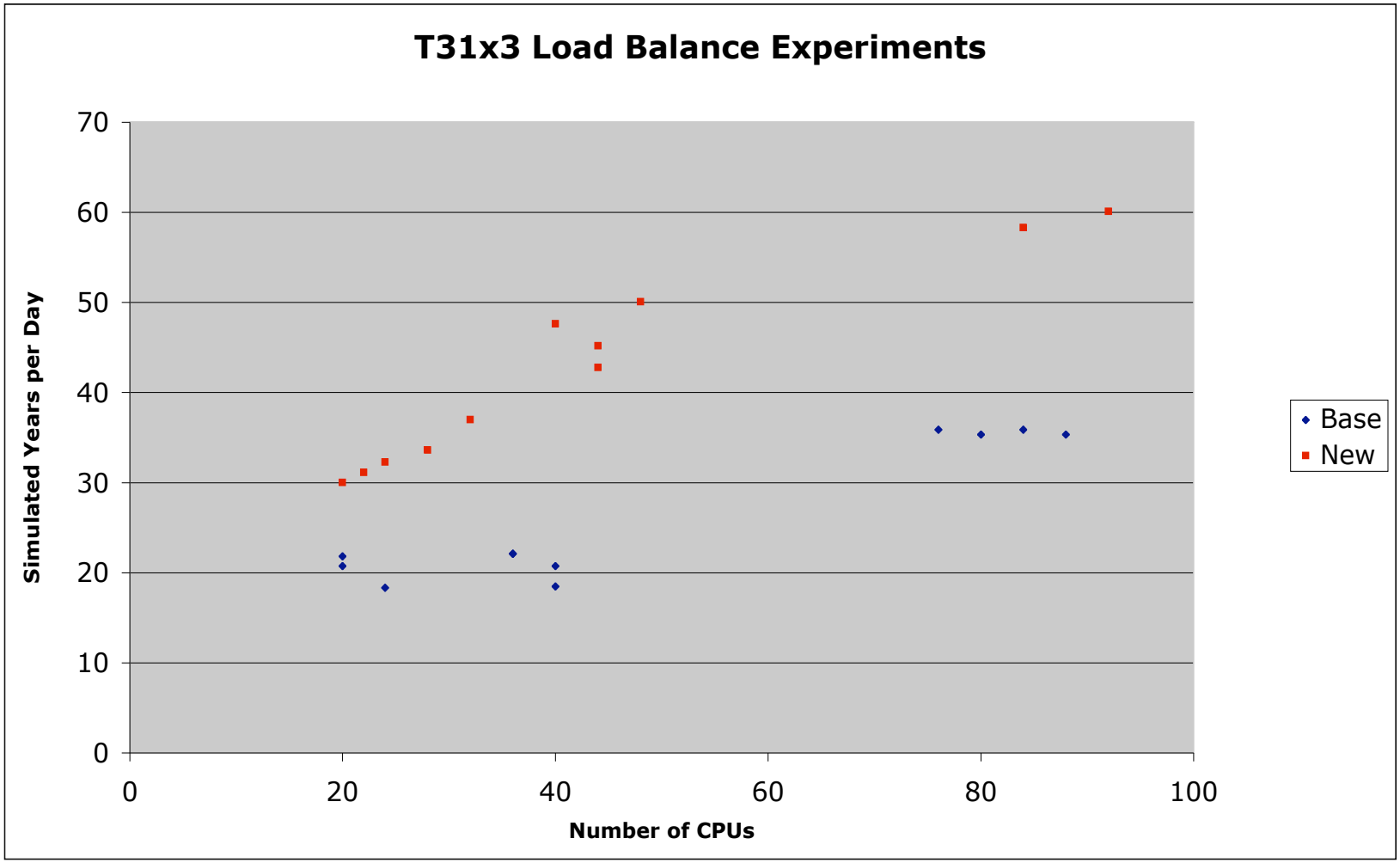
T85 Efficiency



ornl

NCAR

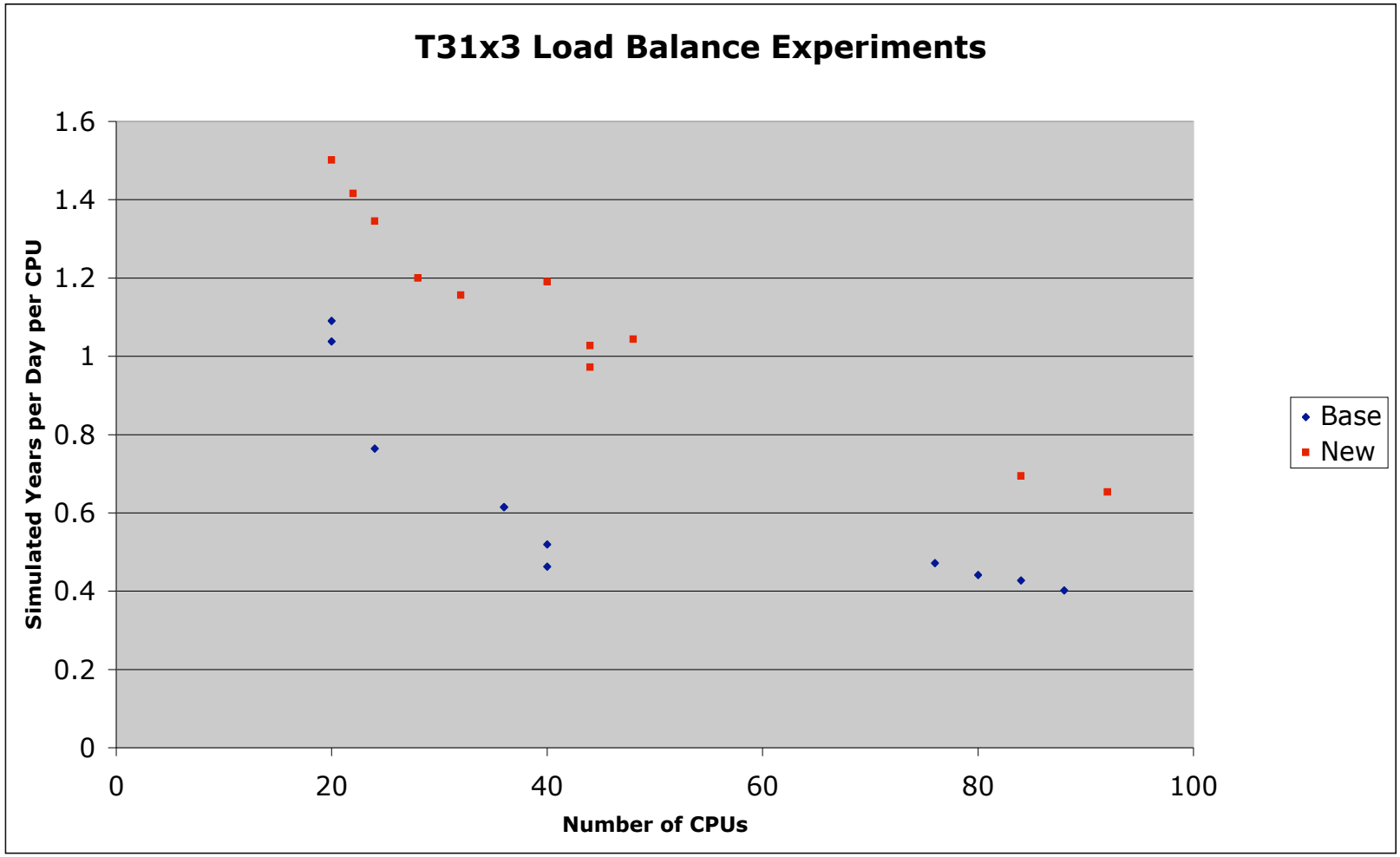
Anticipated T31x3 Improvement



ornl

NCAR

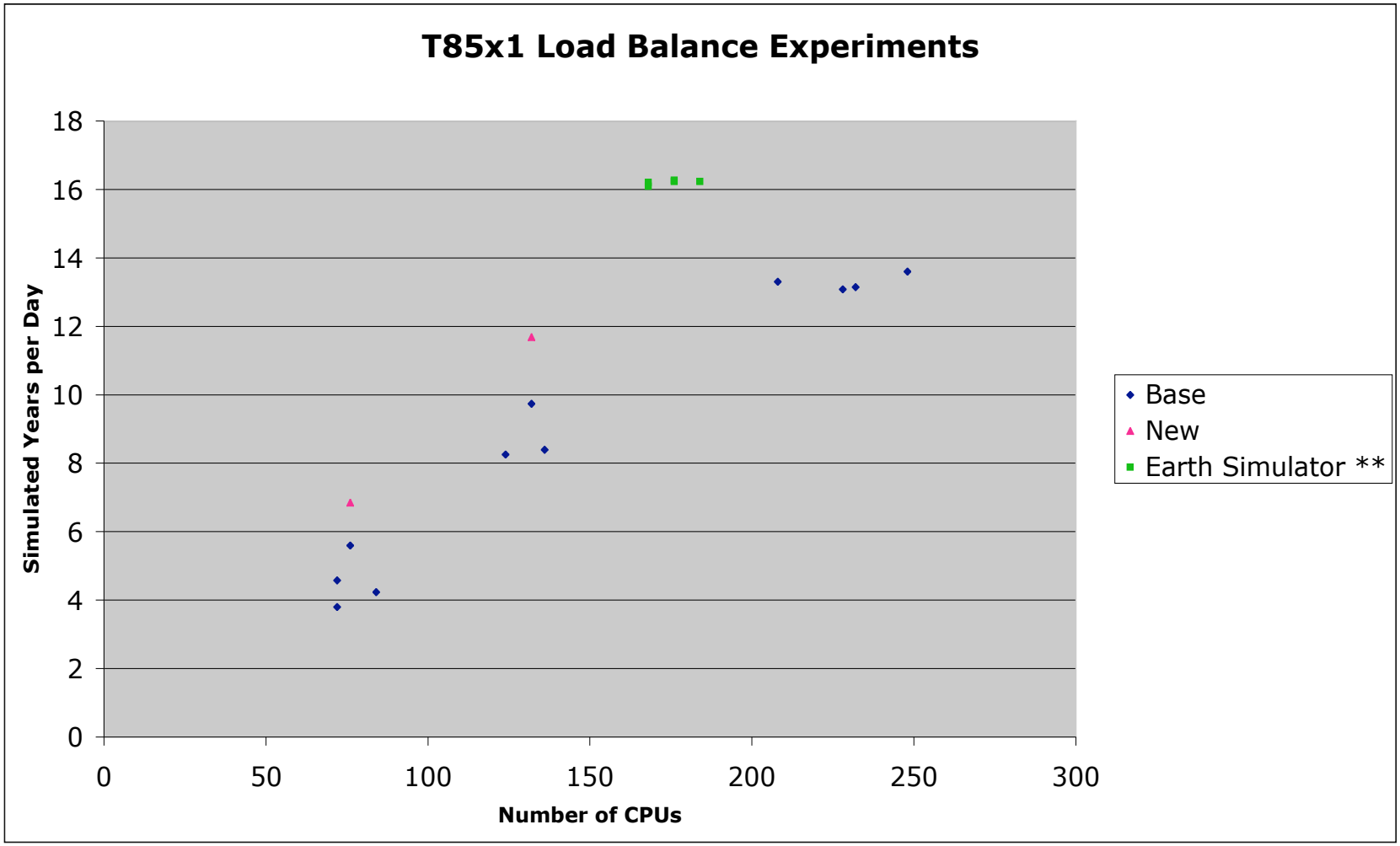
Anticipated T31x3 Efficiency



ornl

NCAR

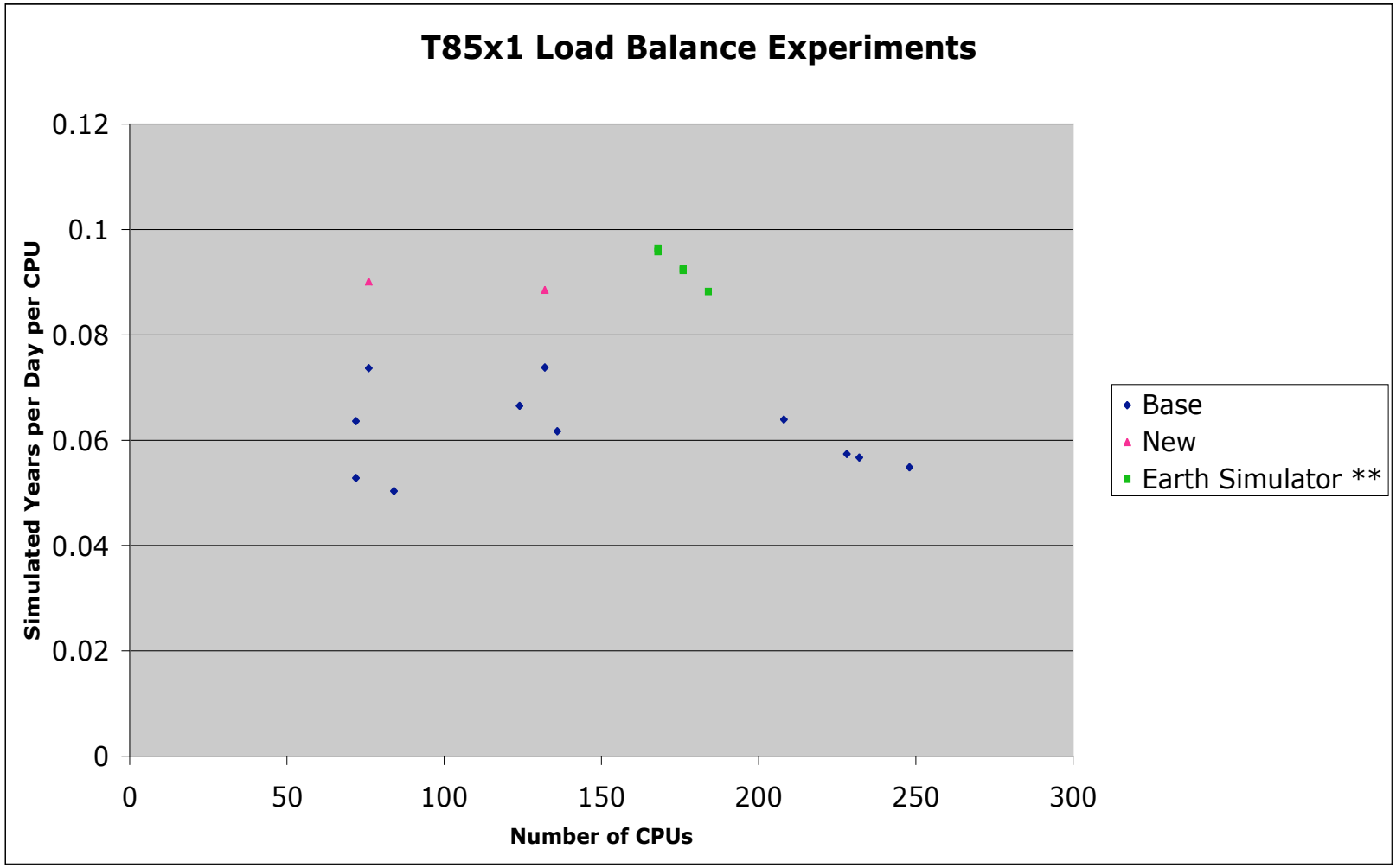
Anticipated T85x1 Improvement



ornl

NCAR

Anticipated T85x1 Efficiency



ornl

NCAR

Opteron Status

- NCAR has own Opteron/Myrinet cluster
- Work on XD1 and XT3 also begun
- Possible issues:
  - MPMD launch
  - Compiler questions
- None validated

Opteron Compiler Status

- PGI status (long term user)
  - Xeon clusters validated with 5.1-3 and 5.1-6 and -O1 (past -O2 issue)
  - Versions 5.2 and 6.0-2 fail PERGRO with -O1 and -O0
  - Bug identified. Bug fix released (6.0-4). Vendor testing with -fastsse (-O2++) shows promise.
- Pathscale status (very recent user)
  - Using latest 2.x version
  - Able to use -O2 to pass PERGRO and partial diagnostics
  - Identified byteswap bug impacts ability to complete validation and (if successful) support production

ornl

NCAR

Plans

- X1
  - Complete current validation
  - T85x1 validation leading to large number of production runs
  - Look at coupler performance
  - The next Programming Environments
  - Need to generalize vector length controls in scripts
- NCAR Opteron cluster
  - Test new PGI compiler
  - Begin first full validation attempt
- T1E/XT3/XD1 (all very important platforms)
  - Continue to work the validation process
- Improved timers (PAPI), CCSM performance characterization (Tau), performance regression testing
- CCSM entering time of massive changes: New Science
  - Change from Eulerian spectral dynamical core to Finite Volume
  - Bio-Geo-Chem changes
  - Increasing resolutions and time steps (T170, .1°, and more)

ornl

NCAR

Summary

- Significant speedup from initial baseline
- Entering significant X1 usage phase
  - CCSM3 official release use by community
  - New science development and experiments

- **Thanks** Cray and ORNL for great support!

ornl

(that's a big thanks)

NCAR

For Further Information

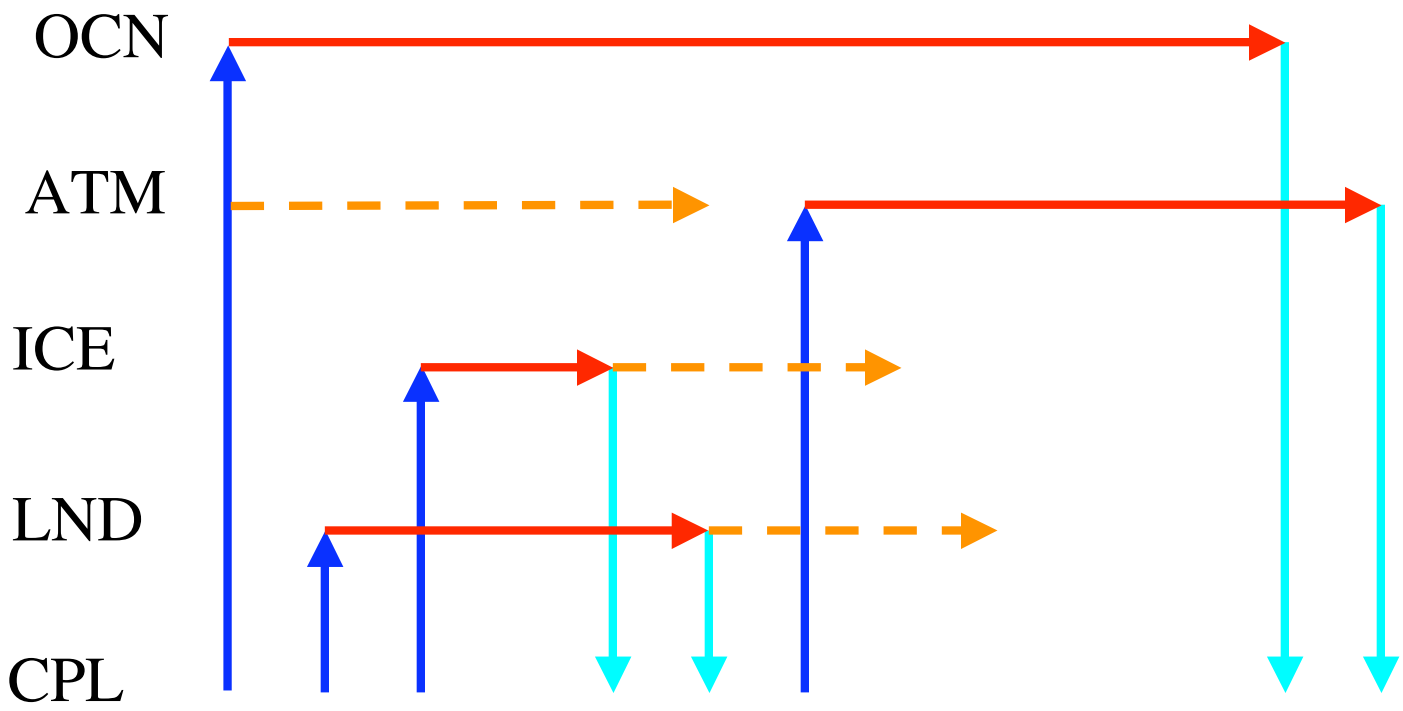
- CCSM web pages
  - http://www.ccsm.ucar.edu/ccsm3
  - http://www.ccsm.ucar.edu/support_model
    - See CCSM User's Guide
    - See Scripts Tutorial
  - http://www.ccsm.ucar.edu/support_model/mach_support.html
- CCSM Bulletin Board
  - http://bb.cgd.ucar.edu
- ORNL X1 evaluation web pages
  - http://www.csm.ornl.gov/evaluation/PHOENIX
- gcarr@ucar.edu

ornl

NCAR

Supplemental Charts

CCSM3 Process Flow



CPL sending data to component (state 1)

CPL receiving data from component (state 3)

Component processing data (state 2)

Component processing (state 4)