PURDUE UNIVERSITY

TeraGrid™

Condor on Campus: Purdue University

Oct 28, 2010

Preston Smith, Ben Cotton
Purdue University

ITaP
INFORMATION TECHNOLOGY AT PURDUE

PURDUE UNIVERSITY

Outline

- Resource Sharing
  - Community Clusters
- Evangelism around campus
- Politics
- Support and Staffing

ITaP
INFORMATION TECHNOLOGY AT PURDUE

Resource Sharing – Community Clusters

- Peace of Mind
  - Professional systems administration so faculty and graduate students can concentrate on research.
- Low Overhead
  - Central data center provides infrastructure such as networking, storage, racks, floor space, cooling, and power.
- Cost Effective
  - Works with vendors to obtain the best price for computing resources, pooling funds from different disciplines to leverage greater group purchasing power.
    - Large purchases also leveraged for departmental server acquisitions
- Get more than you pay for
  - Opportunistic access to other resources if faculty buy in

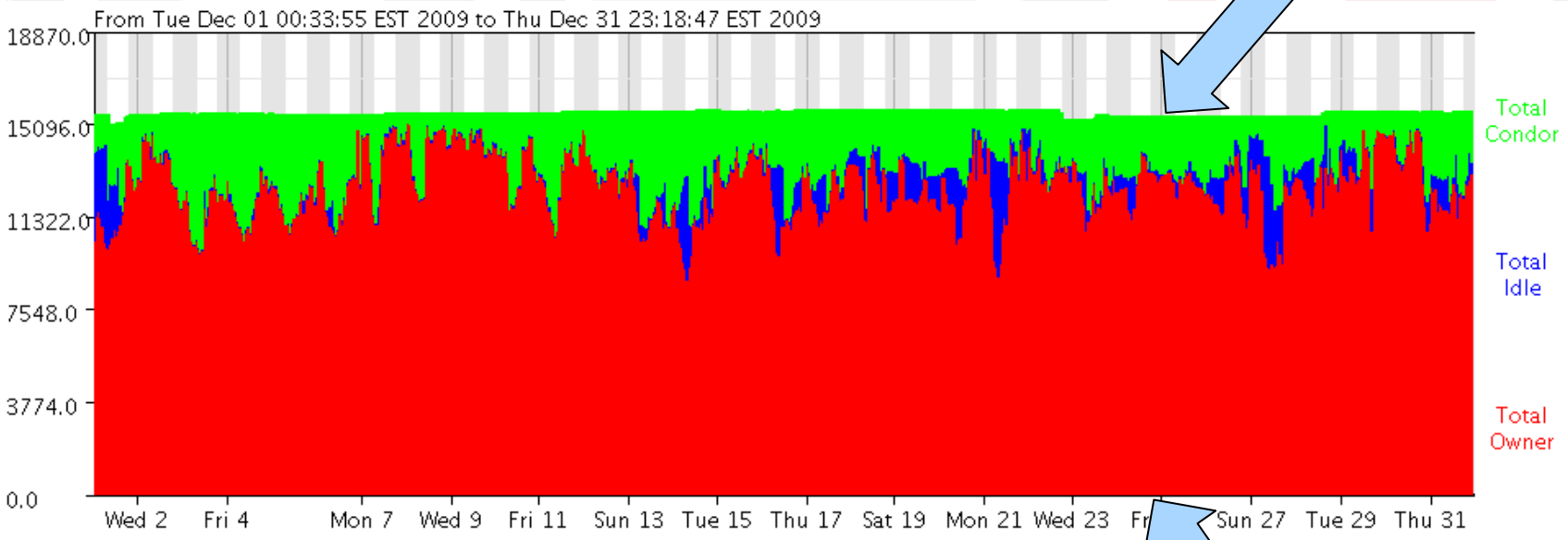
Community Clusters

- "Steele"
  - 902 nodes (7216 cores) Xeon E5410
  - GigE interconnect
- "Coates"
  - 993 nodes (7944 cores) Opteron 2380 Shanghai
  - 10Gb Ethernet interconnect
- "Rossmann"
  - 356 nodes (8544 cores) 24-core Opteron 6172
  - 10 Gb Ethernet interconnect

Community Clusters -> Condor

- Backfilling on idle HPC cluster nodes
  - Condor runs on idle cluster nodes (nearly 24,000 cores today) when a node isn't busy with PBS (primary scheduler) jobs

Central Cluster Usage



Condor: 15.7%

PBS: 81%

Centrally Operated Condor

- To date, the bulk of campus grid cycles are provided by ITaP, Purdue's central IT
  - Submission is handled centrally – on login nodes operated by RCAC
  - OSG and TeraGrid gatekeepers for Condor and community clusters
- Centrally operated Linux clusters provide approximately 24k slots
- Centrally operated student labs provide 7k Windows slots
- Centrally supported workstations have Condor available for install through SCCM.
- That's actually a lot of slots now, but there's more around a large campus like Purdue
  - 27, 317 machines, to be exact
  - Can the campus grid cover most of campus?

On-Campus Evangelism

- **What about non-centralized IT?**
- Less than half of Purdue's IT staff is centralized (ITaP)
  - Of 27,317 machines, relatively few are operated by ITaP!
- Outreach to distributed IT organizations – Many colleges and departments operate over 1000 machines each
  - Agriculture, Computer Science, Engineering, Management, Physical Facilities, Liberal Arts, Education
- Educate IT leadership around campus about what Condor can do for their faculty
- Provide preconfigured, managed packages to ease deployment burden for IT organizations (RPM, deb, .exe)

Campus evangelism is not a technology problem, but a people problem!

Politics

- Campus grid is supported at the highest levels
  - CIO funded additional staff position to support the campus grid, and evangelize
    - » President, Provost, Treasurer all support campus grid
- IT Reorganizing
  - Distributed IT being reorganized – creating reporting lines to CIO
    - » Potentially meaning more members of campus grid?
- Cost reduction
  - University requiring $15M of cost-savings from IT over 3 years
  - Power reduction in IT counts towards savings
  - Condor being piloted to manage machine hibernation around campus

Staffing

- 1 FTE systems engineering, evangelism
- 1.5 FTE advanced user support
  - This is the most important!
- .25 FTE software development, user interface creation
- Very little time required from distributed IT staff contributing machines
  - "Here, install this".
  - If IT has methods to manage many machines, adding Condor on top of it is little additional work
    - **BUT: the department must have investment in making it successful – end users, etc**

The End

Questions?

http://www.rcac.purdue.edu/boilergrid