



# Condor on Campus: BoilerGrid, DiaGrid and Beyond

Jan 19, 2010

Preston Smith, Ben Cotton  
Purdue University



No Cycle Left Behind

No Byte Left Unexplored

*TeraGrid™*

OSG Campus Grids Workshop





At Purdue, we don't make a lot of the products in our grid.

We make a lot of the products in our grid run bigger.<sup>TM</sup>

The word "teragrid" in a large, stylized, lowercase font. The letters are light gray with a drop shadow, giving them a three-dimensional appearance. The "T" and "G" have a small trademark symbol (TM) at the top right corner.

## Outline

- Central Cyberinfrastructure
  - Community Clusters
- Evangelism
- Virtualization
- Managing a large campus grid
- Distributed Storage
- Support and Staffing
- Gratuitous Numbers

TeraGrid™

# 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

## Community Clusters

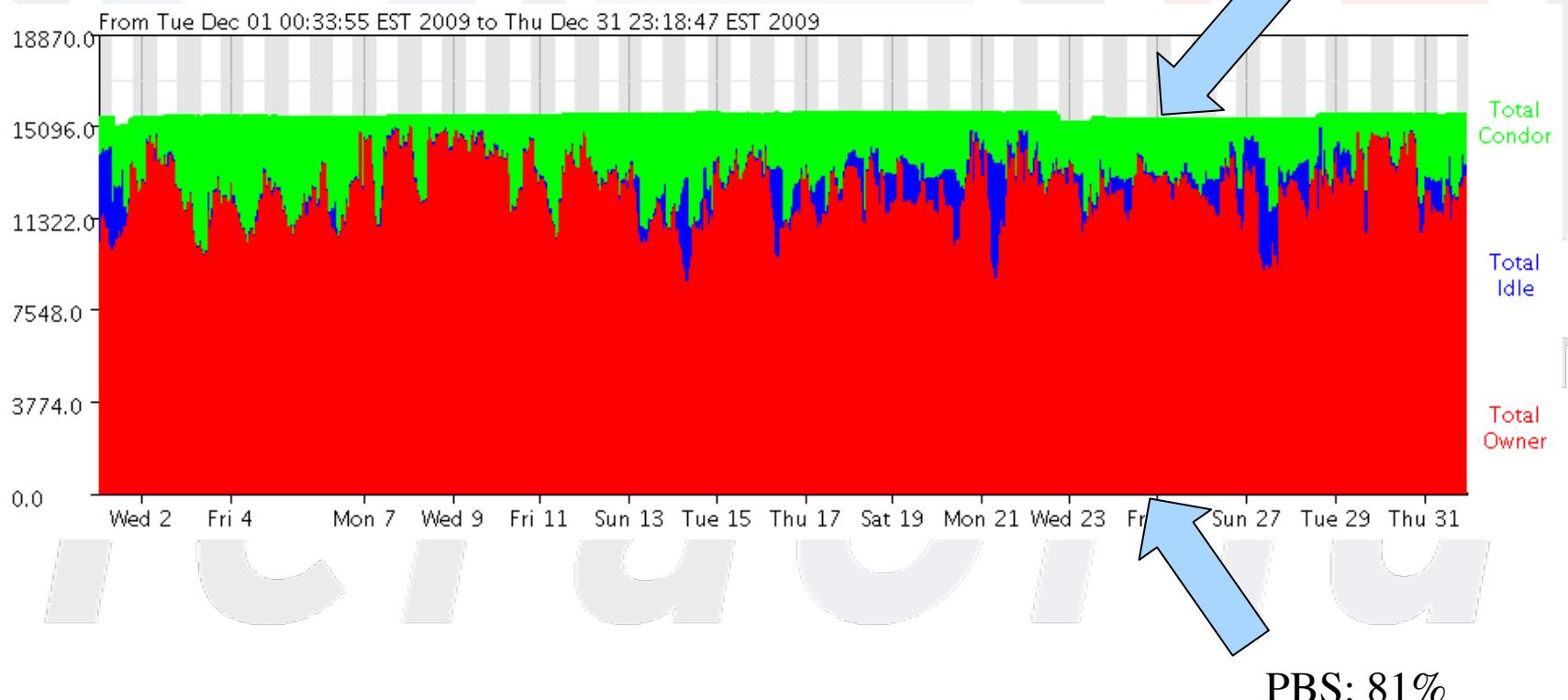
- “Steele”
  - 900 nodes (7200 cores) Xeon E5410
  - GigE interconnect
- “Coates”
  - 950 nodes (7200 cores) Opteron 2380 Shanghai
  - 10Gb Ethernet interconnect
- Coming in 2010 – “Abell”
  - Similar size
  - Named for Purdue staffer Vic Abell – author of “Isof”

## Community Clusters

- Backfilling on idle HPC cluster nodes
  - Condor runs on idle cluster nodes (nearly 16,000 cores today) when a node isn't busy with PBS (primary scheduler) jobs



# Central Cluster Usage



## Central Storage

- 1.3 PB Archival storage system (DXUL)
- Most file storage is NFS – BlueArc
  - ~200TB scratch
  - ~50TB for home directories
    - All centrally maintained clusters share filesystems

- LHC applications have .5PB dCache
- 20TB Hadoop distributed filesystem

## 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 16k slots
- Centrally operated student labs provide 6k 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?



- **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!**

## On-Campus Evangelism

- Host periodic on-campus Condor “Boot camp” for users and sysadmins
- One-on-one conversations with distributed IT leadership
  - Security questions
  - Configurability
  - Scoreboard: “My Dean wants to know how much work our machines have provided”
  - **Machine owners need to be confident that they remain in control of how their machines are used.**
    - Condor is perfect for this!

- Working with Indiana Higher Education Telecommunication System (IHETS) to train Universities in Indiana to run a campus grid, and partner in DiaGrid.
- Deliver presentations on building a campus grid at campus IT conferences such as EDUCAUSE, LabMan.

# Scoreboard

CycleServer | Purdue University Virtual Grid

<http://boilergrid.rcac.purdue.edu:8083/dashboard.s>

Most Visited ▾ Getting Started Latest Headlines

[Home](#) | [Contact Us](#)

 CycleServer Console

### Purdue University Virtual Grid

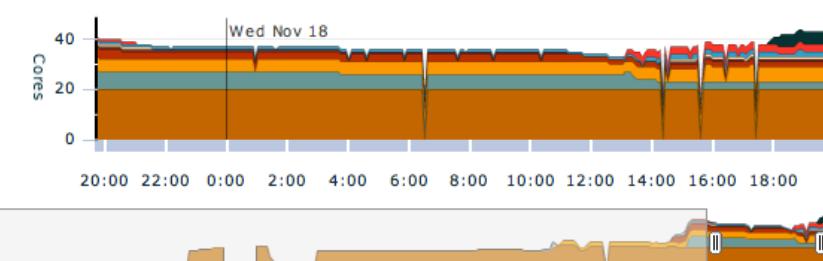
#### VM Contributor Ranking

Contributor	Curr Slots	Avg Slots	CPU Time Provided
Purdue University	20	16	90d 6h 15m
Cycle Computing	6	4	9d 6h 55m
HP	3	5	6d 21h 20m
Indiana University	2	2	3d 19h 30m
University of Iowa	2	1	1d 11h 35m
SC 09 Exhibits Office	1	1	1d 4h 5m
Two Guys and a Cluster	3	2	1d 1h 40m
Clemson University	1	1	13h 40m
Anonymous	0	5	11h 45m

#### VM Contributions

Time Frame: 3 Hours | Day | Week | Month

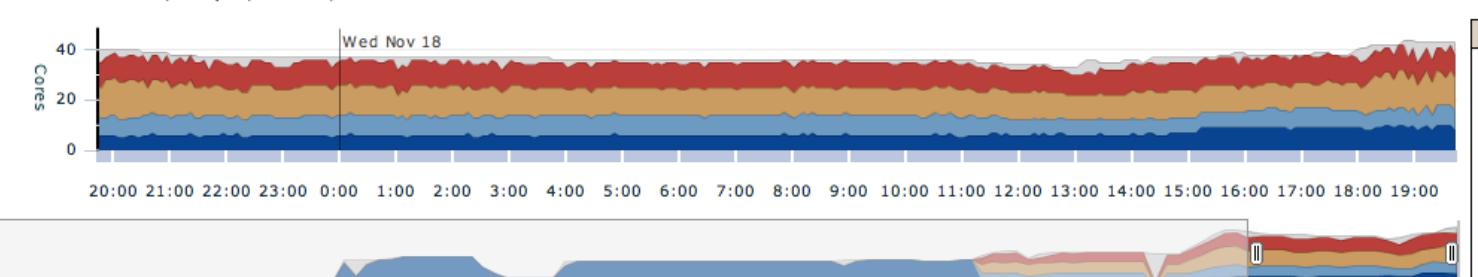
View as: Area | Line



#### Total Grid Usage

Time Frame: 3 Hours | Day | Week | Month

View as: Area | Line



Legend:

- Used by owner
- Varian, Inc
- Bioinformatics
- nextGenSequencing
- Computational Chemistry
- Unclaimed

Version 3.2.5sc [Terms & Conditions](#)

Transferring data from boilergrid.rcac.purdue.edu...

© 2009 Cycle Computing, LLC. All Rights Reserved

## On-Campus Evangelism

- **For example:**
- Engineering is Purdue's largest non-central IT organization – 4000 machines
  - Already a campus grid partner, providing nearly 1000 cores of Linux cluster nodes
- But what about desktops? What about Windows?
  - Engineering is interested... But...

- Engineering leadership wants the ability to sandbox Condor away from systems holding research or business data.
- Can we do this?

## A note about Windows

- Windows porting is definitely a hurdle to overcome.
- Some users are making the effort – Visual Studio ports of code are in use by some users.
- Now provide a centrally-operated Cygwin system configured to look like RCAC Linux server, can simplify porting that way
- Hmm. An OS porting hurdle? Machine owners interested in sandboxing?
  - This sounds like an opportunity...

## Virtual Condor

- All of this has happened before, and all of this will happen again.
- Some implementations exist
  - CoLinux (OU)
  - Grid Appliance (Florida)
  - Marquette's VirtualBox solution
- Some other ideas
  - Submit VM universe jobs as a "VM glide-in"



## Virtual Condor

- What we've implemented is a solution based on the Grid Appliance infrastructure from Florida's ACIS lab
- IPOP P2P network fabric
  - Solves NAT issues and IP space problems that come with bridged networking
  - No requirement for single VPN router to connect real network with the virtual overlay network.
- We only need to run IPOP services (a userland application) on all central submit nodes to access nodes in the virtual pool

## Virtual Condor

- For us **and** partners on campus, this is a win
  - Machine owners get their sandbox
  - Our support load to bring new machine owners online gets easier
- Much of the support load with new “sites”/departments is firewall and Condor permissions.
  - Virtual machines and the IPOP network makes that all go away.
- Not only native installers for campus users, but now a VM image
  - Install VMWare, and install systray app to connect keyboard and mouse to guest OS
- Not necessarily virtualization implementation dependent – we can prepare and distribute VM images with KVM, VirtualBox, Vmware, Xen, and so on.
  - Just VMWare currently
  - We’ll offer more in the future.

## "The Cloud"

- With Condor, virtualization is now an option – Virtual Machine Universe
  - VMware jobs are possible on Windows labs
  - KVM hypervisors available on cluster nodes
- Dedicated cloud resource "Wispy"
  - 32 nodes (96 cores) Xeon
  - Using Nimbus from Globus project
    - Resource being tapped for computer science undergrads and
  - Future work
    - Integrate Nimbus with Condor – "pilots" (an unfortunately overloaded word) to start machines with VM universe

## Manageability

- So, given:
  - Thousands of virtual Linux systems out there not under central cluster management framework
  - Thousands upon thousands of Windows lab machines that research computing staff don't administratively control..
- How do we manage Condor on them?
  - We use Cycle Computing's CycleServer
  - VM images are configured to report in to CycleServer for management
  - As are the native OS installers that we distribute



The screenshot shows the CycleServer Console interface. At the top, there are two tabs: "CycleServer Console" and "Machine Configuration". The "Machine Configuration" tab is active, showing a sub-menu with "Templates" selected. A modal dialog box titled "Editing Template: TLT" is open in the center. The dialog contains fields for "Template name:" (set to "TLT") and "Available sections:" (with a dropdown menu and a "Create a new section" link). Below these are sections of configuration code:

```

tit_common
DAEMON_LIST = MASTER, STARTD
UPDATE_COLLECTOR_WITH_TCP = TRUE

BLAST_DB = "\\\$samba.rcac.purdue.edu\blastdb"

STARTD_EXPRS = BLAST_DB, ${STARTD_EXPRS}

hostallow_boilergrid
HOSTALLOW_ADMINISTRATOR = $(CONDOR_HOST), boilergrid.rcac.purdue.edu

```

At the bottom of the dialog are "Comments:" and "Save" and "Cancel" buttons.

- Cycles are well understood, what about opportunistic bytes?
  - Our cluster nodes have 160+ GB boot disks.
  - Likewise with student labs
- That's over 300TB going unused
- Installation of Cycle's CloudFS
  - Based on Hadoop
  - With FTP interface, FUSE filesystem interface, and S3-like REST interface
- 30TB testbed today on "recycled" cluster
- Dedicated 32-node Hadoop Cluster
- Who's using?
  - CMS vaguely interested
  - Electrical Engineers Mapping and Reducing
  - Biologists interested in using distributed storage
  - TeraGrid team integrating with TeraGrid replicated storage service



- Our 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
- CIO's directive – power-save your idle computers, or run Condor and join the campus grid
  - Today, even the President of the University runs Condor on her PC

*teragrid*<sup>TM</sup>

- 2 FTE systems engineering, evangelism
- 2 FTE advanced user support
  - This is the most important!
- .5 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**

## Gratuitous Numbers

- 17M hours delivered to users in 2009
- 15M jobs completed in 2009
- ~30,000 slots today
- ***This is in addition to the Community Clusters!***
  - For Steele in 2009
    - 1.8 M jobs
    - 48 M hours
  - For Coates (and a couple small clusters) in 2009
    - 600k jobs
    - 37 M hours
  - In total
    - 17.4 M jobs
    - 102.3 M hours

*Condor Usage only 17% of HPC  
hours at Purdue*



The End



Questions?  
<http://www.rcac.purdue.edu/boilergrid>

OSG Campus Grids Workshop

