PURDUE UNIVERSITY

TeraGrid™

Condor to Every Corner of Campus

Condor Week 2010

Preston Smith
Purdue University

ITaP
INFORMATION TECHNOLOGY AT PURDUE

PURDUE UNIVERSITY

Outline

- Campus Grids
- Condor Nests at Purdue
  - Central Clusters
  - Computing Labs
  - Departmental Resources
  - TeraGrid
- Budget Realities in 2010
  - IT Cost Reduction
- Spreading Wings Across Campus
  - Making Condor Easy for IT
  - Dashboard
- Virtualization
  - VMGlide

ITaP INFORMATION TECHNOLOGY AT PURDUE

Campus Grids

- Open Science Grid Campus Grids Workshop held in January
  - Identified themes common to many Campus Grid implementations
    - Barriers are often diplomatic rather than technological
- At the core, a campus grid is a way for an institution to **share resources**, and **maximize its investment** in computing
  - Many different ways to share resources
    - Purdue, FermiGrid, GLOW, others all implement in their own way

http://www.isgtw.org/?pid=1002447

Community Clusters

- Purdue's model for resource sharing begins here
- 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

PURDUE UNIVERSITY

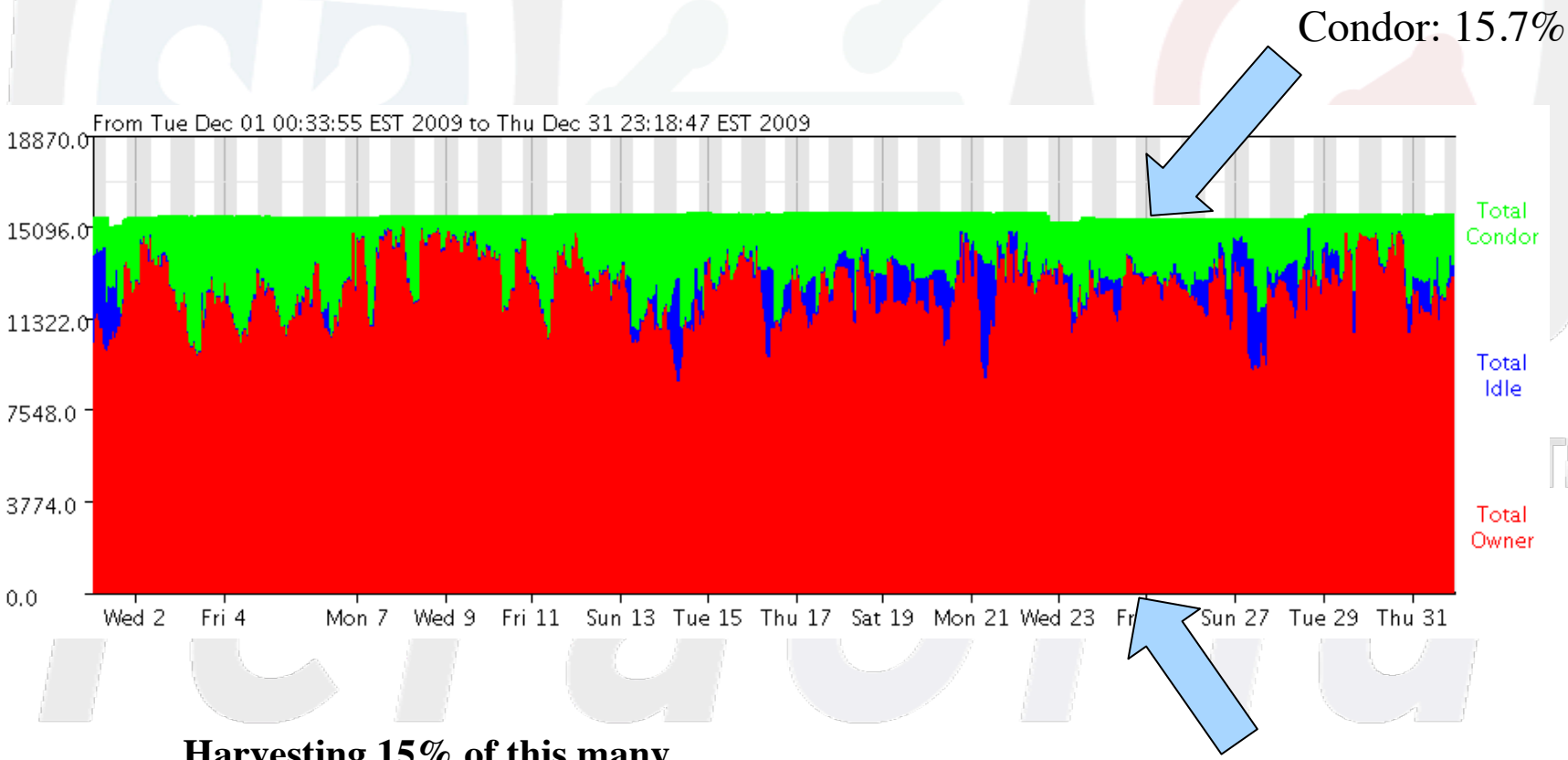
Community Clusters

- 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

ITaP INFORMATION TECHNOLOGY AT PURDUE

Central Cluster Usage

- Maximizing value from investment



Harvesting 15% of this many machines' availability is 22 million potential hours per year!

PURDUE UNIVERSITY

Student Labs

- ITaP operates nearly 2000 lab machines used in classrooms, general student labs, and for departments.
  - Nearly 6000 cores among those 2000 machines

ITaP INFORMATION TECHNOLOGY AT PURDUE

PURDUE UNIVERSITY

Distributed IT

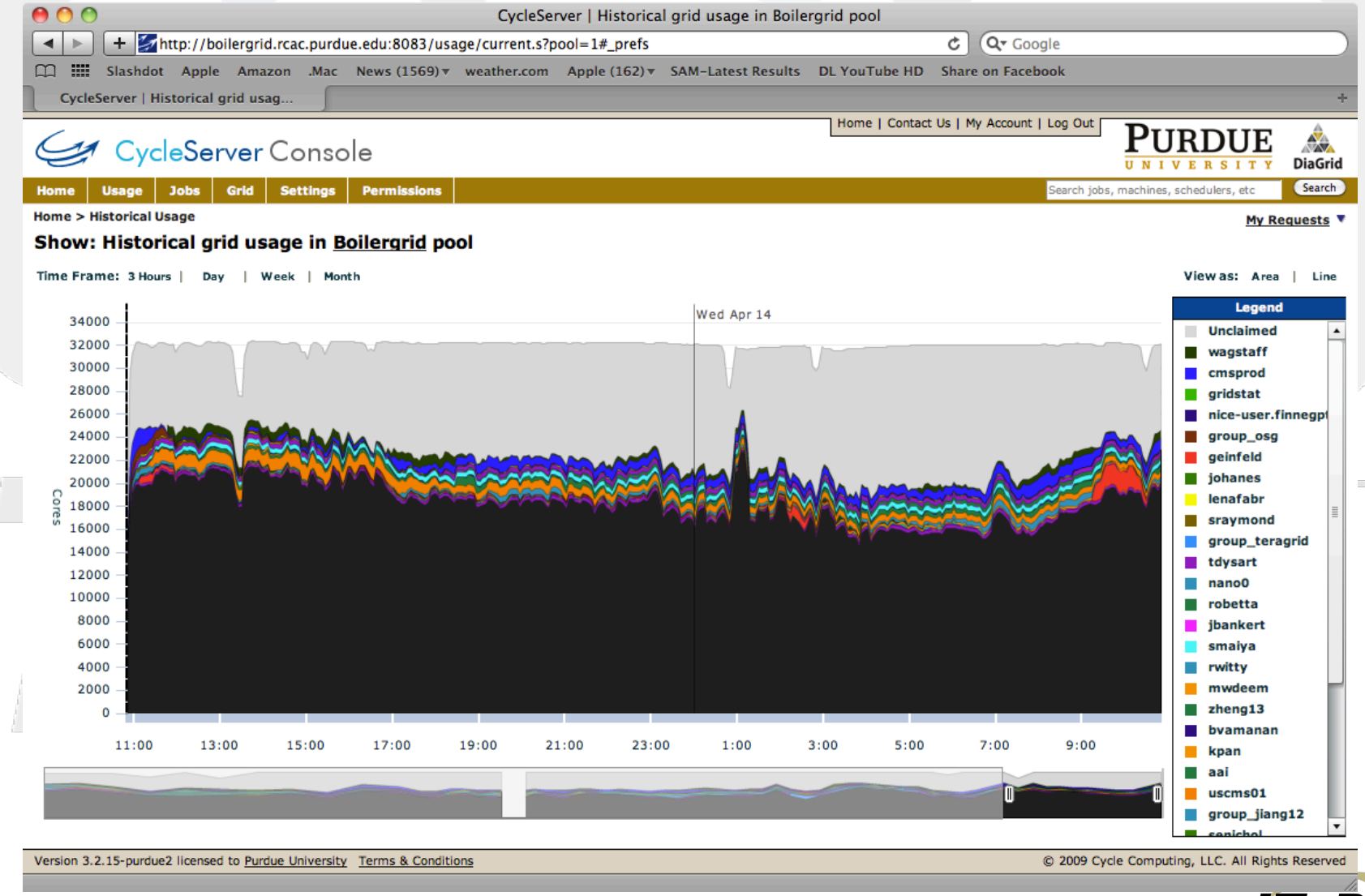
- Less than half of IT at Purdue is centralized
  - Remainder is in individual colleges and departments
  - 27,317 desktop machines at West Lafayette, relatively few of which are operated by ITaP
- Many of these islands of IT are quite large
  - Agriculture, Computer Science, Engineering, Management, Physical Facilities, Liberal Arts, Education
  - 1000+ machines each
    - Many of these IT organizations are in the Condor Grid already
    - But many are not…

This is where the room for growth is!

ITaP INFORMATION TECHNOLOGY AT PURDUE

PURDUE UNIVERSITY

Grid Overview



ITaP
INFORMATION TECHNOLOGY AT PURDUE

PURDUE UNIVERSITY

TeraGrid

- Purdue provides the campus Condor pool to the nation via the TeraGrid
  - 50% of jobs on TeraGrid in 2004-2006 were single-CPU
  - Of those, 64% ran for an hour or less
    - (Arvind Gopu of Indiana University – TG'07)
- Robetta gateway, many others regularly use Purdue Condor on TeraGrid
- Condor will continue to be a TeraGrid resource through the end of the TeraGrid project

ITaP INFORMATION TECHNOLOGY AT PURDUE

Blah Blah Blah

- You say "Sure, we've heard all this before.. What's new?"

State Budgets in 2010

- A common conversation on campuses today
  - Higher Ed in Indiana has been directed to reduce budget by 5.5%
- At Purdue, we have been given the following charge for IT:
  - "Identifying cost savings approaches that will generate at least $15M recurring over time while providing high quality information technology (IT) services to meet the University's strategic goals."
    - Data Centers
    - Computer Labs
    - Power Savings
    - Strategic Sourcing for Purchasing

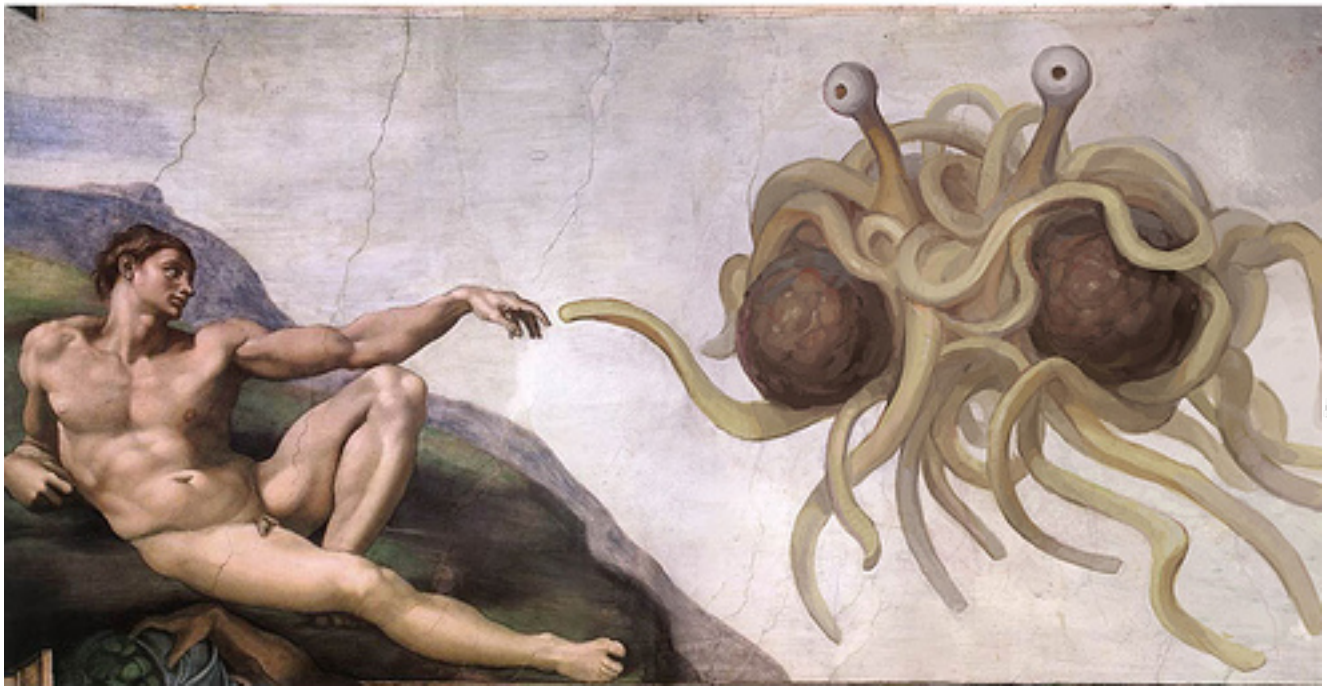
What does this have to do with Condor?

- The Campus Grid ties into several of these areas
  - Data Centers
    - Building community clusters instead of private ones, and then **maximizing usage** with Condor
  - Computer Labs
    - Centralize management of labs – and **run Condor** on the machines
  - Strategic Sourcing in purchasing
    - For example, **community cluster** purchase for good pricing
  - Power Savings
    - Virtualized data centers, **power off** idle computers
    - "Power credits" for running Condor

PURDUE UNIVERSITY

Power Off or Install Condor

- Recommendations from committee report
  - "Thou shalt turn off thy computer or install Condor and join the Campus Grid"
  - "Thou should power-save your machines and we should find tools to manage their waking-up"



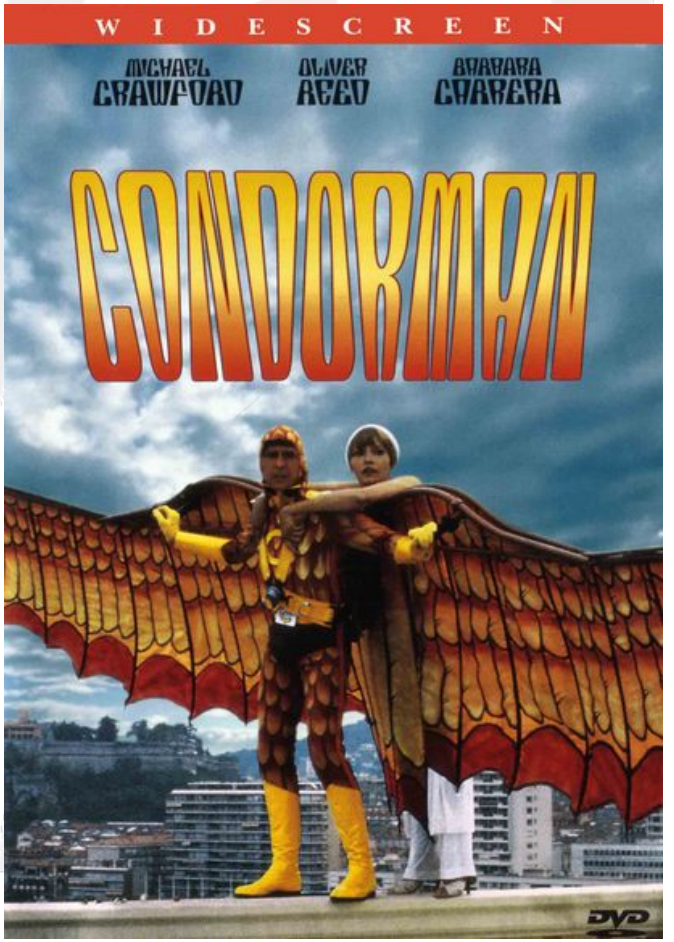
id™

ITaP
INFORMATION TECHNOLOGY AT PURDUE

Power?

- The Blue-Ribbon committee making recommendations probably didn't know this..

- But this **also** sounds like a job for Condor!
  - Killing two "birds" with one stone
    - Add machines to the Campus Grid – harvest the cycles, as already recommended
    - Power-save machines by a policy
    - Wake them up when needed



Now What?

- Currently a "recommendation", not quite a "policy" yet
- What happens when it becomes a policy?
  - The bet is that IT folks won't want to shut their machines down overnig ups/ software distr
  - Expect a tsun

How to Prepare?

- Host periodic on-campus Condor "Boot camp" for users and sysadmins
  - These are very much like the User and Administrator tutorials many of you were in yesterday
    - (Thanks to the Condor team for letting us base things off of their materials)
- Ease of deployment
  - Provide pre-configured binaries
    - Windows, Linux (RHEL, Debian, Ubuntu, Fedora)
- Configurability
  - Centrally managing Condor Configurations on machines with distributed ownership
  - … while leaving configuration also in the hands of the machine's owners
- **Machine owners need to be confident that they remain in control of how their machines are used.**
  - **Condor is perfect for this!**
- Scoreboard:
  - "My Dean wants to know how much work our machines have provided"
  - The president asks how much work her individual machine has done
- Security questions?

Manageability

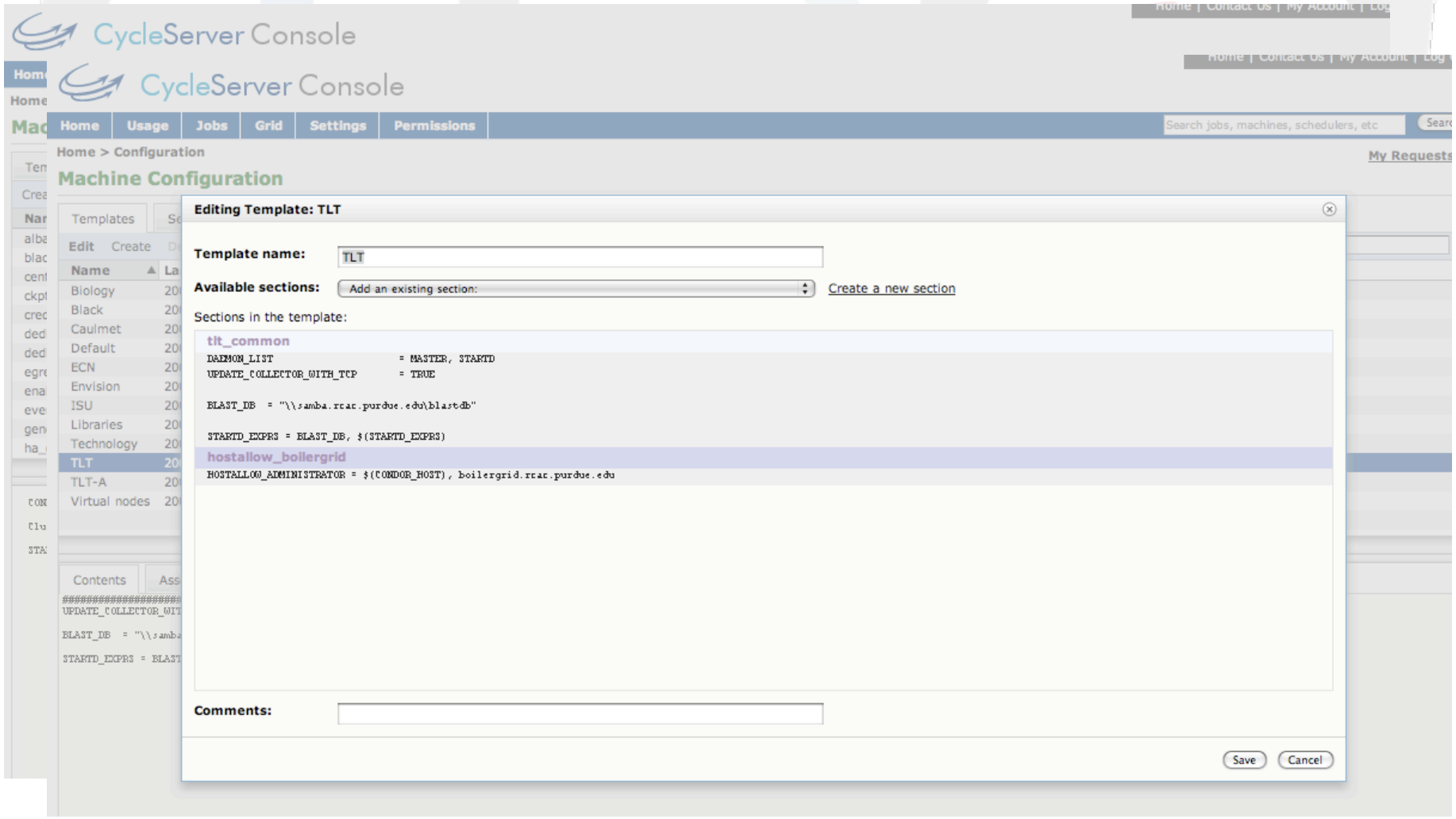
- So, given:
  - Thousands upon thousands of Windows lab machines, or all sorts of machines around campus that my staff don't administratively control..
- How do we manage Condor on them?

PLUG ALERT

- We use Cycle Computing's CycleServer
- VM appliances are configured to report in to CycleServer for management
- As are the native OS installers that we distribute

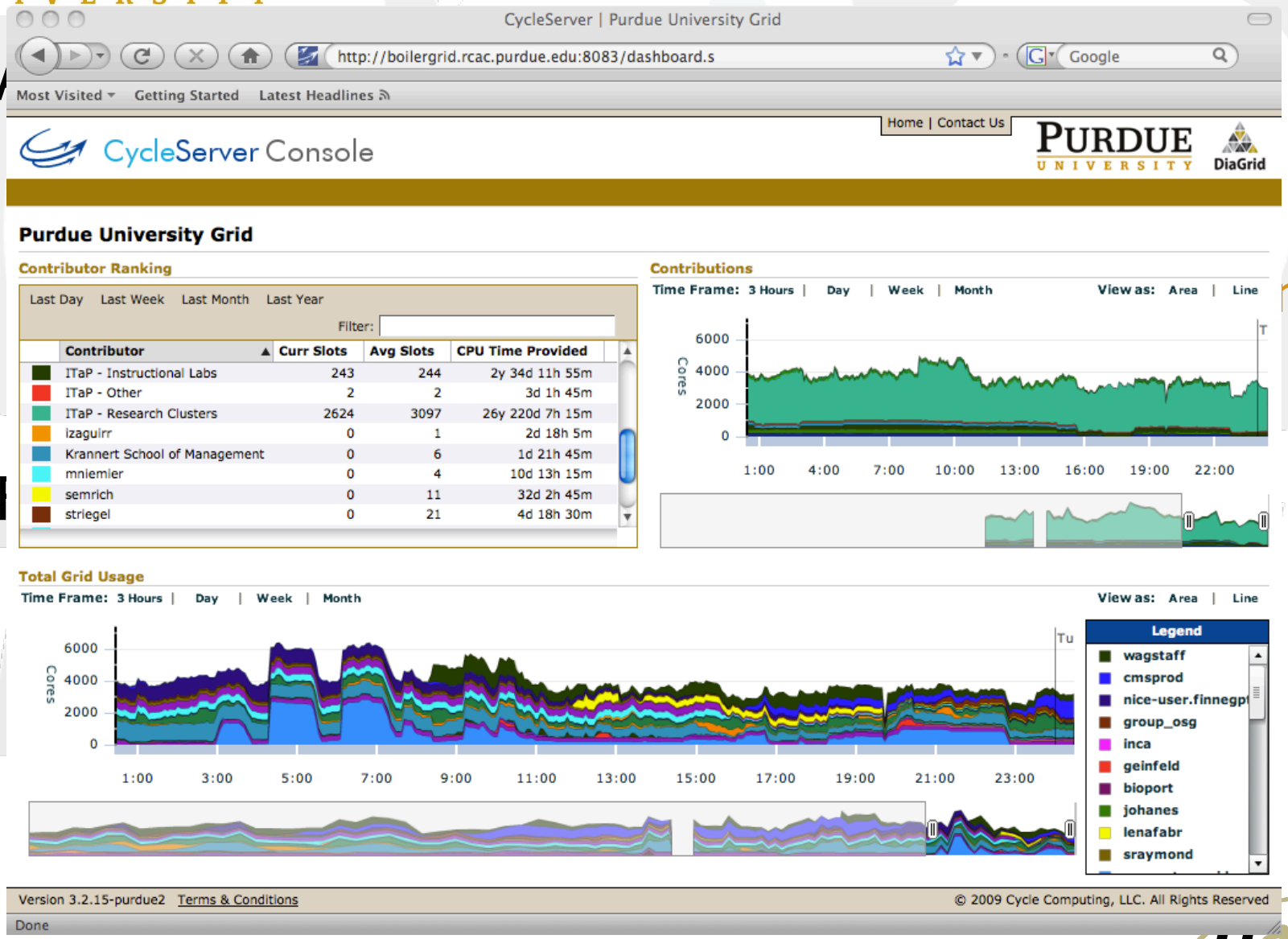
PURDUE UNIVERSITY

Managing Configurations Around Campus



ITaP
INFORMATION TECHNOLOGY AT PURDUE

Scoreboard


CycleServer | Purdue University Grid
http://boilergrid.rcac.purdue.edu:8083/dashboard.s
Most Visited
Getting Started
Latest Headlines
CycleServer Console
Home | Contact Us
PURDUE UNIVERSITY
DiaGrid
Purdue University Grid
Contributor Ranking
Last Day
Last Week
Last Month
Last Year
Filter:
Contributor
Curr Slots
Avg Slots
CPU Time Provided
ITaP - Instructional Labs 243 244 2y 34d 11h 55m
ITaP - Other 2 2 3d 1h 45m
ITaP - Research Clusters 2624 3097 26y 220d 7h 15m
izaguirr 0 1 2d 18h 5m
Krannert School of Management 0 6 1d 21h 45m
mniemier 0 4 10d 13h 15m
semrich 0 11 32d 2h 45m
striegel 0 21 4d 18h 30m
Contributions
Time Frame: 3 Hours | Day | Week | Month
View as: Area | Line
Cores
6000
4000
2000
0
1:00
4:00
7:00
10:00
13:00
16:00
19:00
22:00
Total Grid Usage
Time Frame: 3 Hours | Day | Week | Month
View as: Area | Line
Cores
6000
4000
2000
0
1:00
3:00
5:00
7:00
9:00
11:00
13:00
15:00
17:00
19:00
21:00
23:00
Legend
wagstaff
cmsprod
nice-user.finnegpt
group_osg
inca
geinfeld
bioport
johanes
lenafabr
sraymond
Version 3.2.15-purdue2
Terms & Conditions
© 2009 Cycle Computing, LLC. All Rights Reserved
Done

F A'd Qs about Security

- Can any ding-dong submit any code to my machine?
  - No – only specific machines with access limited to people with Purdue Career Accounts run a schedd
- Ok, fine, but what if they submit something nasty to our machine?
  - Then we know who they are and go club them with the appropriate IT policies.
- What about data on our faculty members' workstations Is it safe? Could a job steal it?
  - Well, maybe. Are their file permissions set appropriately?

Sandbox

- College of Engineering asks –
  - Can we sandbox Condor jobs away from the execution host?
- We think "sure" – and it'll also make those Windows boxes more generally useful.
  - **Maximizing investment again**

Condor in VM appliances

- Many ways to skin that cat
  - CoLinux from several years ago
  - Marquette from a few minutes ago
  - Condor as a virtual machine manager from a few minutes before that
- Some effort spent similar to what Marquette's doing
- But mostly on what we've dubbed VM-Glide
  - Using Condor to submit VM "nodes" as jobs to the lab machines
  - Lab machines run VMWare workstation –
    - Which is ok for Universities to use for "instruction and research" and for "grid and utility computing" if you enroll in a partner program

VM-Glide

- Our solution is 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.
  - See talk from Condor Week 2009
- We only need to run IPOP services (a userland application) on all central submit nodes to access nodes in the virtual pool

How well did this work?

- Set up a dedicated schedd with lots of disk to hand out VMs to student labs
  - (Fast) disk is important – checkpointing memory adds up!
- Configure lab machines to claim the entire machine when a single VM Universe job runs
  - Apparently users notice when 4 or more VMs try and evict when they sit down
- Now we're cooking – got nearly 1000 VMs running in labs over a weekend
  - All of which are running user jobs
  - IPOP fabric holds up great

So, you run this all the time now, right?

- Well, not quite
  - Even with just 1 VM per machine, vacating is still noticeable by the end users
  - Lab admins say: "Maybe it's the 100Mbit connection the machines are on".
- After we cried a little inside..
  - How to deal with this?
    - Use squids local to labs to cache VM images?
      - Nope, the lab network architecture doesn't lend to that
    - Pre-stage VMs on machines and just start with Condor?
      - Nope, VM-GAHP doesn't actually let you do that.
    - Upgrade network in labs?
      - Cost-prohibitive – switch gear is old enough that it's not gigabit capable

Next steps?

- Fortunately, a campus network upgrade is in progress
  - With new switches, will benchmark again
- Lab admins enabling vt support in BIOS
  - Allow for 64-bit VMs (more jobs want this)
  - Will probably make VMWare run faster, too
- Pre-stage VMs
  - Hack the VM-GAHP to start pre-staged VMs
  - Or use a file transfer plugin to copy from local hard drive

What's Next

- We expect to add Condor to machines from all across campus
  - And system-wide..
- We hope to use Condor as the tool to manage power on machines across campus
- Virtualization of compute environments will be a key characteristic of this environment
  - In labs and desktops, as well as on cluster nodes (KVM)

Thanks to the Condor Team for all the Software!

The End

Questions?

http://www.rcac.purdue.edu