Personal Condor

Todd Tannenbaum
Computer Sciences Department
University of Wisconsin-Madison
tannenba@cs.wisc.edu
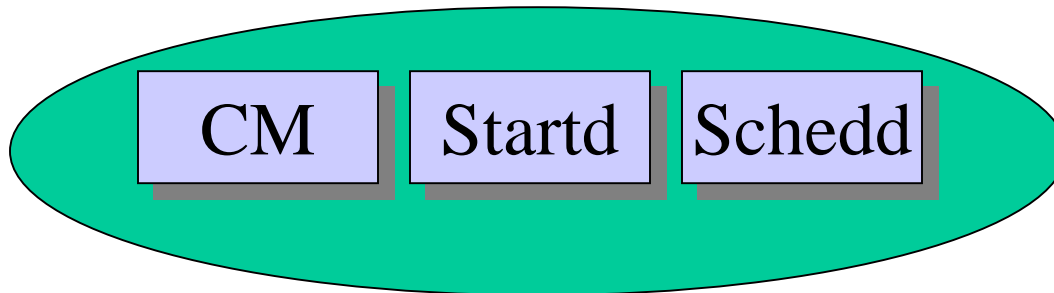http://www.cs.wisc.edu/~tannenba

Condor

What is Personal Condor?

> Condor re-focused and re-packaged to emphasize

- Use by an individual
  - … without scores of machines local to their organization
  - … without sysadmin experience
  - … without sysadmin authority
- Some Specific Condor Mechanisms
  - Flocking
  - Globus Job Universe
  - GlideIn

Condor

A Pool of One

> Typical Condor Installation

Startd

Schedd

CM

Startd

Schedd

Startd

Schedd
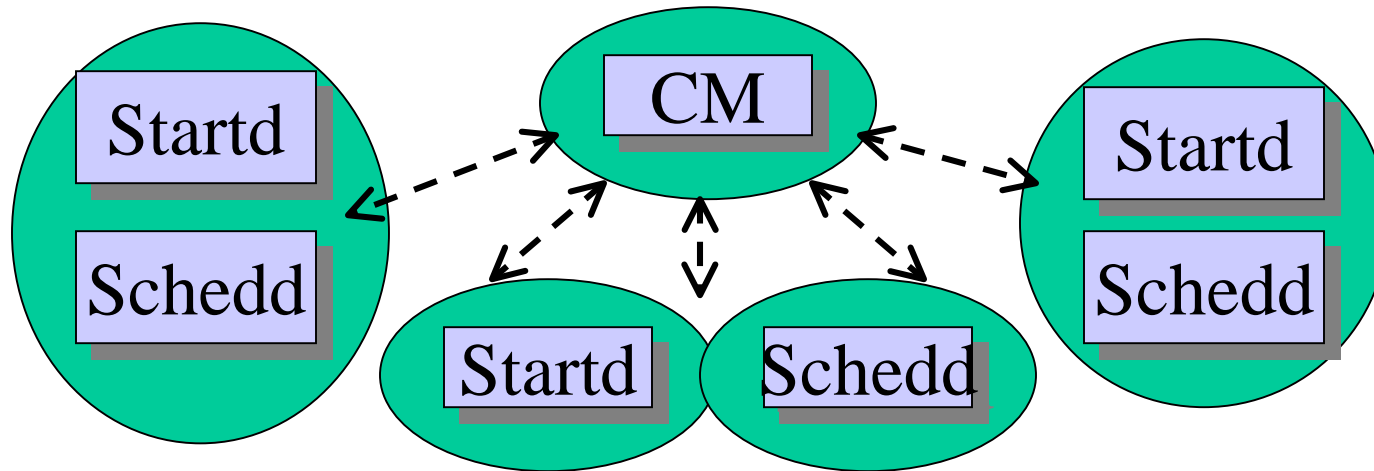
> Typical Personal Condor Installation

CM   Startd   Schedd

Personal Condor?!

# What's the benefit of a Condor "Pool" with just one user and one machine?

Your Personal Condor will ...

> ... keep an eye on your jobs and will keep you posted on their progress

> ... implement your policy on when the jobs can run on your workstation

> ... implement your policy on the execution order of the jobs

> ... add fault tolerance to your jobs

> ... keep a log of your job activities

600 Condor
jobs
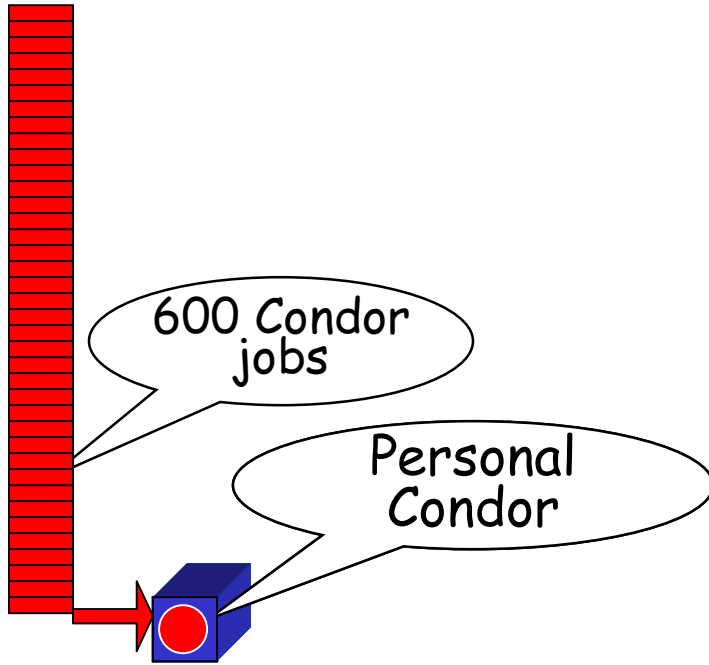
Personal
Condor

There's more! Build a Personal Cluster

> Install Condor on any other machines you have

> Point them at your Personal Condor installation machine
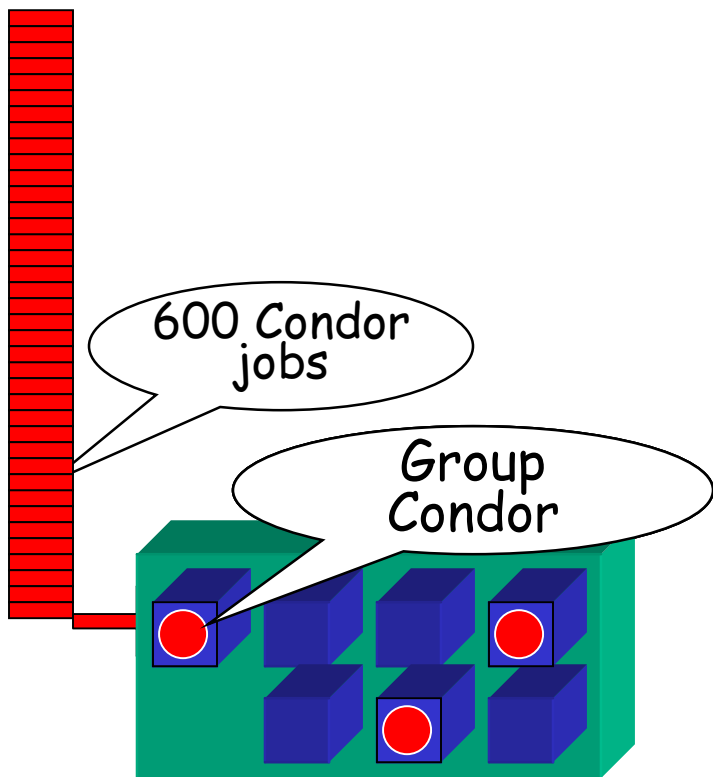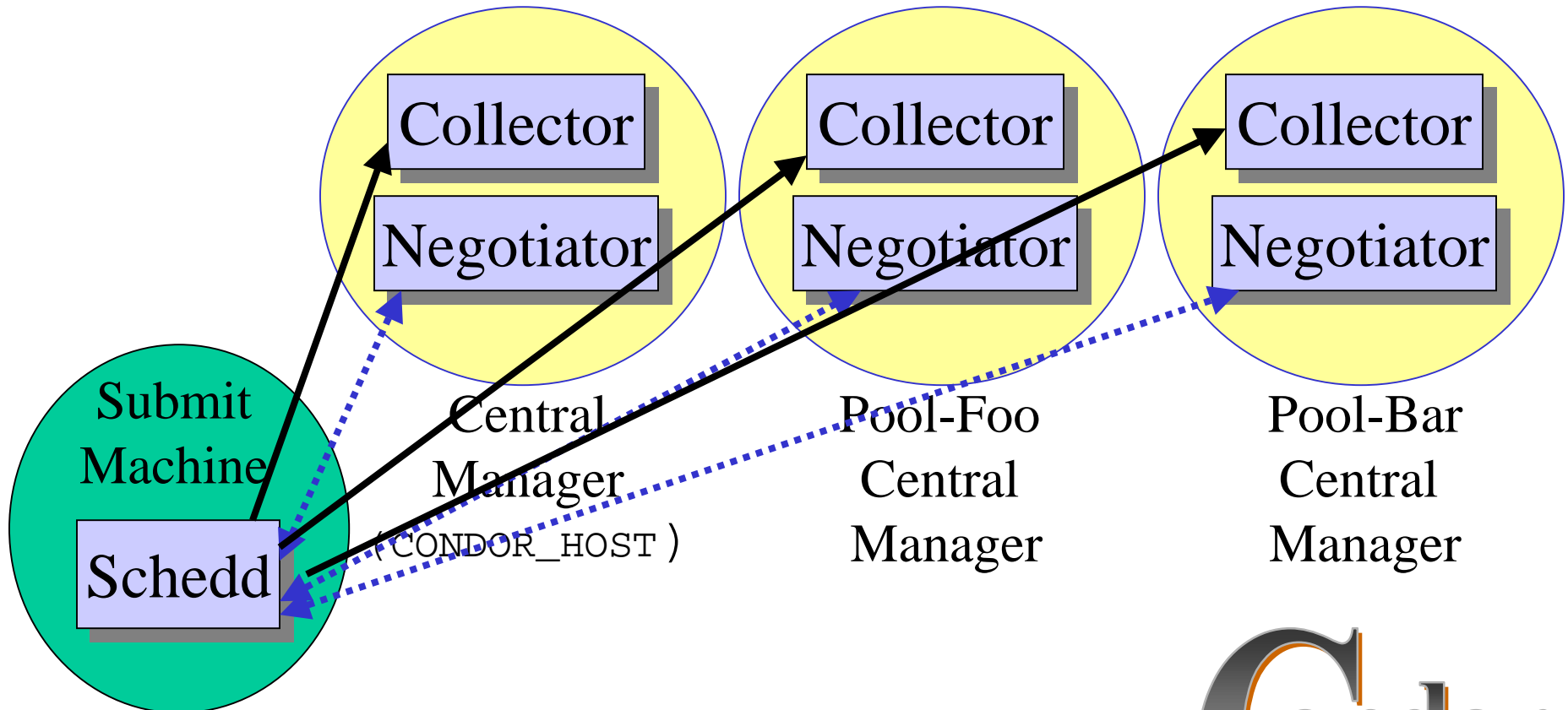
There's more! Take advantage of your friends

> Get permission from "friendly" Condor pools to access their resources

> Configure your Personal Condor to "flock" to these pools

600 Condor jobs

Group Condor

Friendly Condor
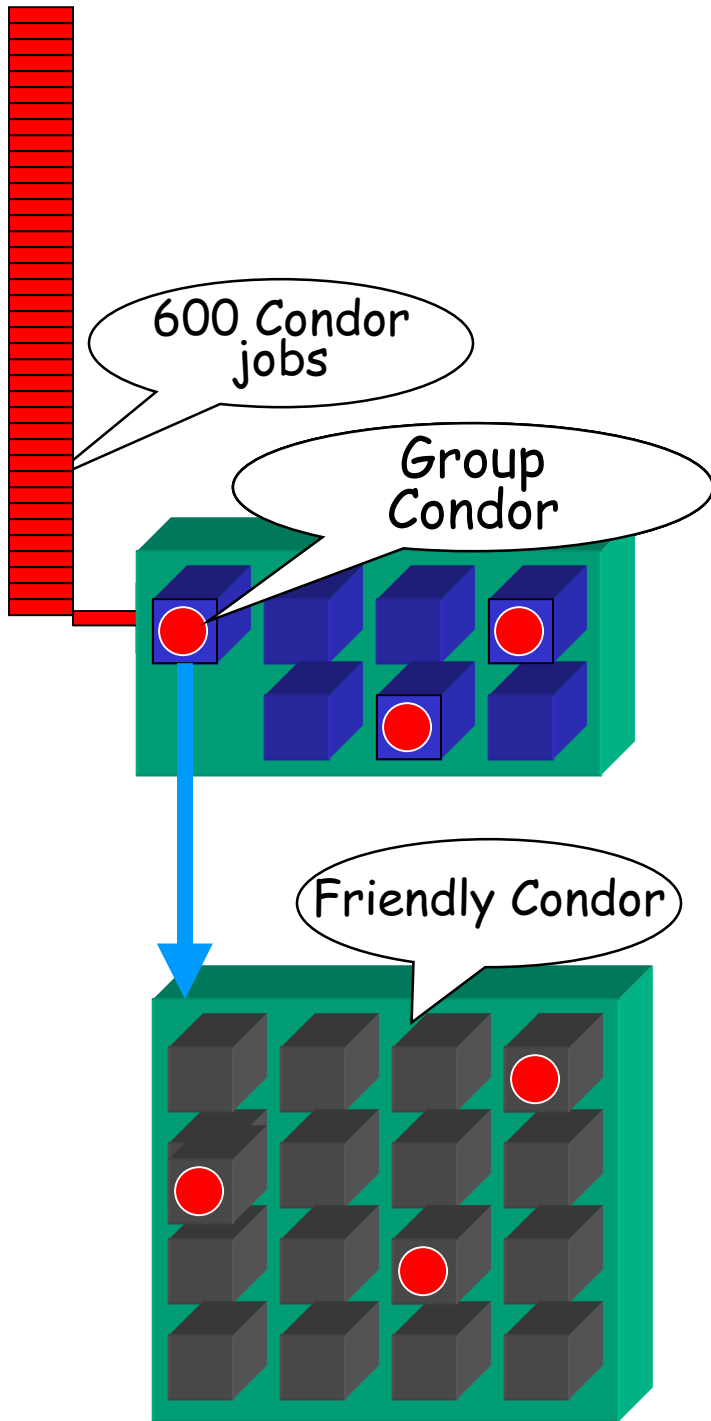
How Flocking Works

> Put in condor_config :

FLOCK_HOSTS = Pool-Foo, Pool-Bar

Flocking: Pros & Cons

> Pros
- Flock hosts are tried *in the order specified* until jobs are satisfied
- Property of the Schedd, not the CM
  - Different users can Flock to different pools
- User priority system is "flocking-aware"

> Cons
- Not yet well-integrated with the tools
- Job Rank does not work across pools

www.cs.wisc.edu/condor

Condor

There's More! Access the National Technology Grid

› National Technology Grid
  - Collaborative effort underway by many Government labs, National Science Foundation, and several Universities
  - Centralized "virtual machine room" for account creation, etc.

› Globus Toolkit
  - Grid Certificate (X.509)
  - Globus Gatekeepers

www.cs.wisc.edu/condor

Condor

Submit a job in Condor to run on Globus

> In submit description file given to *condor_submit*, specify:

- Universe = Globus

- Which Globus Gatekeeper to use

- Location of file containing your Globus certificate

Globus Grid

GK

GK

PBS   LSF

Condor Submit
Machine

Submit Jobs to Gatekeepers

Monitor Job Progress

Schedd

Globus_shadow

Globus_shadow

Globus_shadow

Condor

GK

Condor

www.cs.wisc.edu/condor
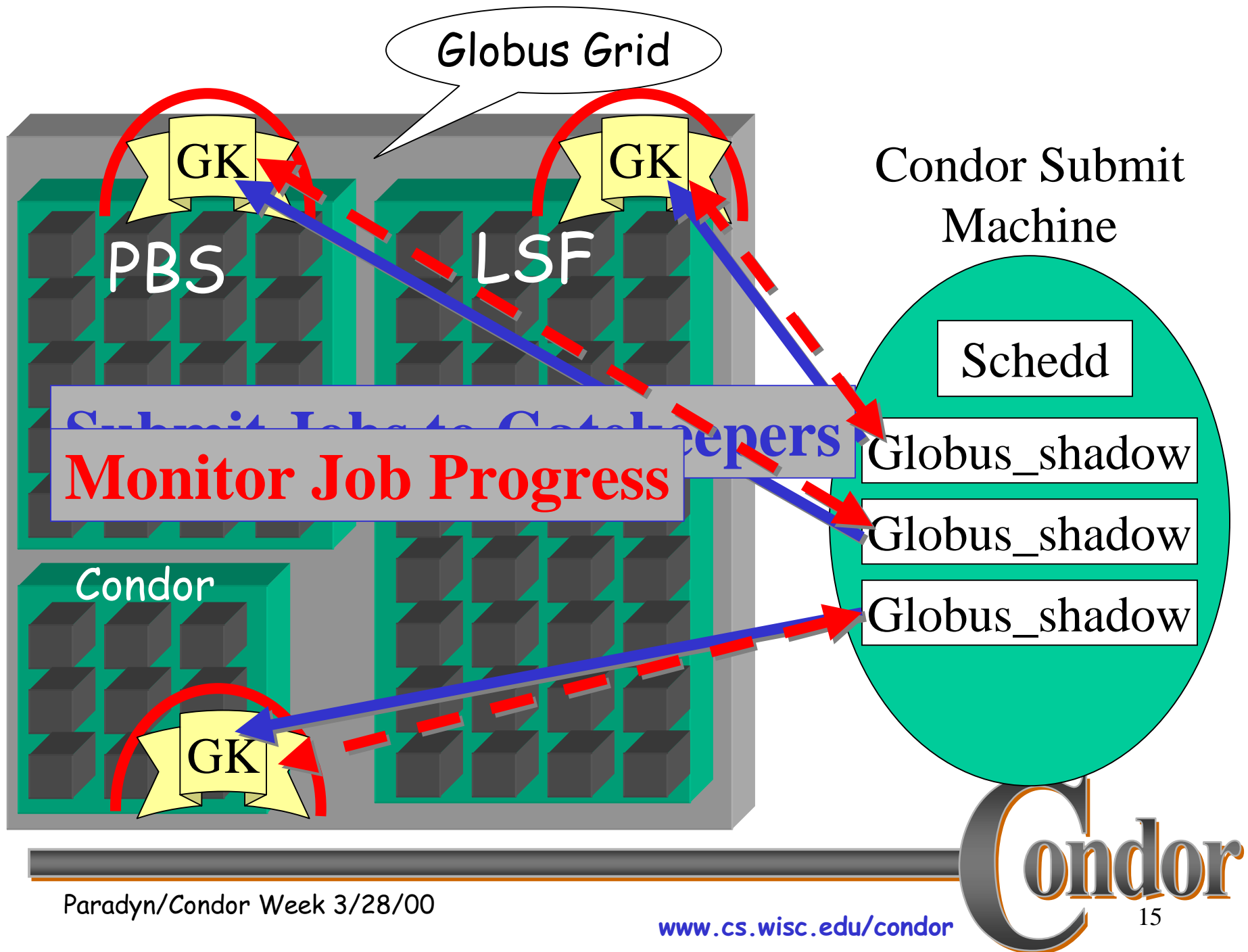
Condor's Globus Universe Pros/Cons

> Pros

- Persistent queue for jobs destined to run on Globus resources

- Jobs submitted to run on Globus resources are managed
 - details on progress, order of execution, DAGMan, …

> Cons

- Must specify a particular gateway

- What about Standard Universe features?
  - Checkpoint/Restart, Remote system calls, etc.

One Solution: Condor GlideIn

› Submit your jobs as regular Condor jobs (Standard, Vanilla, or PVM Universe…)

› Expand your Personal Condor pool by submitting "GlideIn Jobs"

  • A Globus Universe job which consists of the Condor Daemons (master, startd, starter).

› When your GlideIn Jobs run on the Globus resource, *that resource will join your Personal Condor Pool!*

› Then your regular Condor jobs get matched and run on Globus resources as usual.
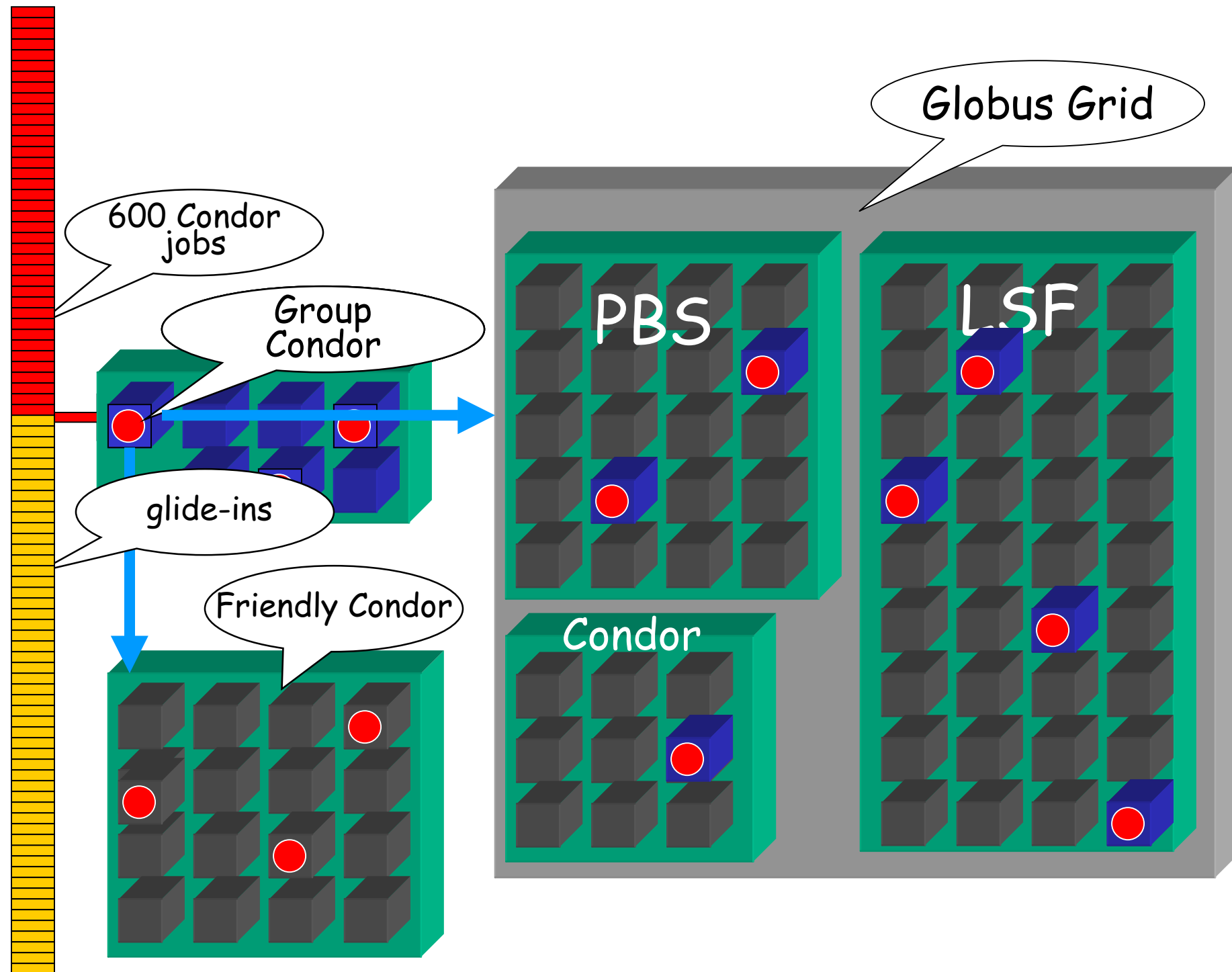
Some Problems and some Solutions

What if all your jobs completed before a given GlideIn job starts running?

Solution: If the Condor daemons which have glided-in are not matched with a job in 10 minutes, they terminate.

Condor

Some Problems and some Solutions, cont.

My Personal Condor is flocking with a bunch of Solaris machines, and also doing a GlideIn to a Silicon Graphics O2K. I do not want to statically partition my jobs.

<u>Solution</u>: In your submit file, say:

```
Executable = myjob.$$(OpSys).$$(Arch)
```

The "$$(xxx)" notation is replaced with attributes from the machine ClassAd which was matched with your job.

In Review

With Personal Condor you can…

- … manage your compute job workload
- … access local machines
- … access remote Condor Pools via flocking
- … access remote compute resources on the Grid via Globus Universe
- … carve out your own personal Condor Pool from the Grid with GlideIn technology.

www.cs.wisc.edu/condor

Condor

Current Status

> Initial Personal Condor implementation exists

- Includes Flocking, Globus Uni., GlideIn
- Powered several demonstrations at SC99 (SuperComputing 1999 Conference)
- In use by some collaborators
- Demonstration tomorrow

www.cs.wisc.edu/condor

Condor

Current Status, cont.

> Work currently in progress

- Enhance robustness, scalability

- Add features missing from Condor's Globus Universe

- Streamline the process for the user

- Personal Condor distribution

  - Simplified Installation Procedure

  - "Grid-Access Only" mode

Questions
and
Thank You!