Challenges in Dynamic Deployment of Condor Across Distributed Environments

Andrew Pavlo
Computer Sciences Department
University of Wisconsin-Madison
pavlo@cs.wisc.edu
http://www.cs.wisc.edu/~pavlo/

Problem Statement

> Difficult to allocate reliable resources across multi-sites:

- Batch Systems (Scheduling)
- Network (Public vs Private, Firewalls)
- Availability
- Capabilities
- Etiquette

Overlay Grid Network

> Create custom global Condor pool using Glidein technologies.

> Global fair share at user and group level.

> Uniformity across all Grids (OSG, EGEE)

> "Reduces grid-related errors by 50%"

Condor

CRONUS

> ATLAS Virtual Computing Cluster

> Condor-G Glideins

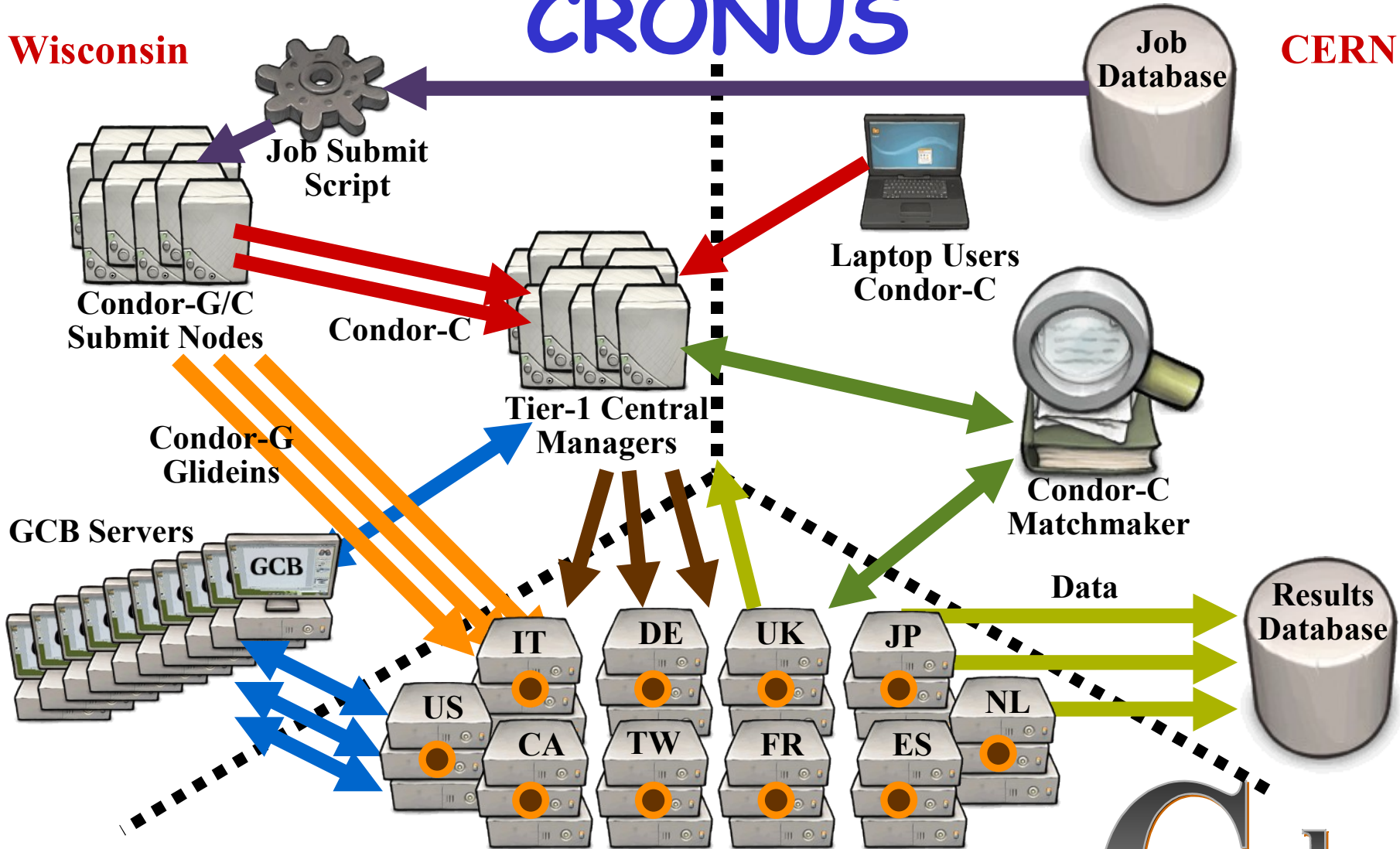> Condor-C Job Submissions

> GCB Network Nodes

> Goal: +10,000 jobs

Sanjy Padhi
HEP @ University of Wisconsin

CRONUS

Wisconsin

CERN

Job Database

Job Submit Script

Condor-G/C Submit Nodes

Condor-C

Laptop Users Condor-C

Tier-1 Central Managers

Condor-C Matchmaker

Condor-G Glideins

GCB Servers

GCB

IT　DE　UK　JP

US

CA　TW　FR　ES　NL

Data

Results Database

Tier-1 Clouds
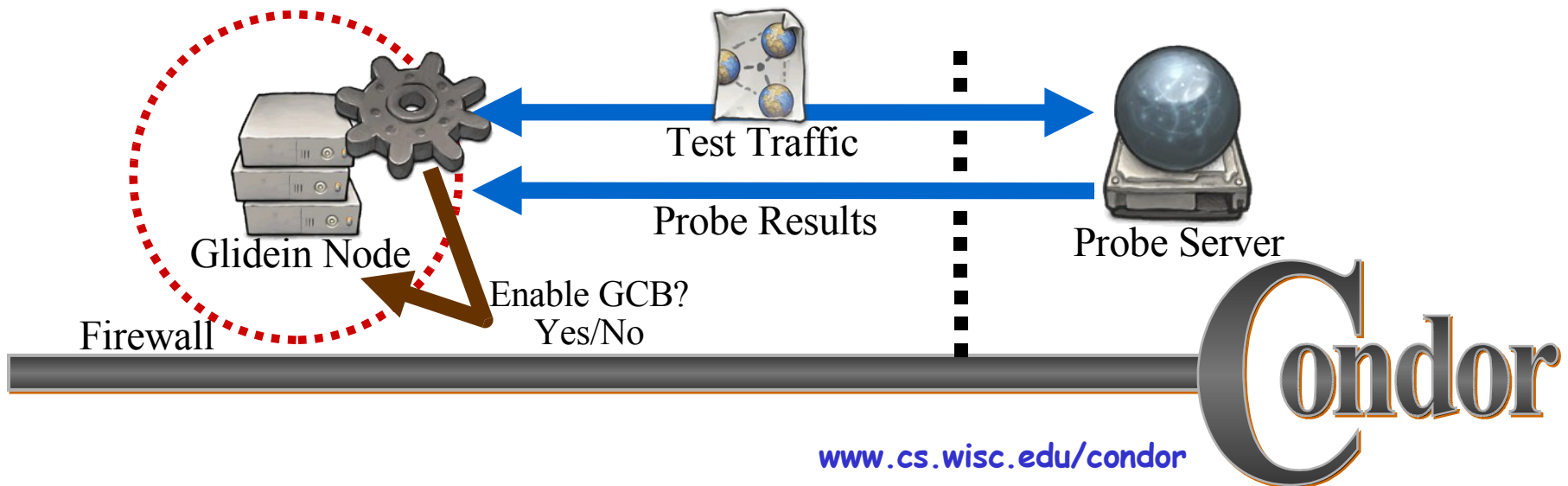
Condor

www.cs.wisc.edu/condor

Deployment Challenges

> Unknown Network Capabilities

> Cleaning Up on Execution Node

> Retrieving Job Attributes

> Scalability Issues

Unknown Network Capabilities

> **Problem**: How can we determine the network environment of execute nodes?

> Firewalls, Public vs. Private IPs

> GCB mitigates problem, but is error prone.

Solution: Network Probe

> Contact Condor servers @ Wisconsin to determine network information.

> Only enable GCB if needed.

> Source code is available!

Glidein Node

Firewall

Test Traffic

Probe Results

Enable GCB?
Yes/No

Probe Server

Condor

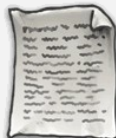# Cleaning Up on Execution Node

> **Problem**: How do we make sure that our Glideins are actually doing work and not wasting cycles?

> Must handle severed network connections.

Condor

Solution: Shutdown Exprs.

> New expressions allow Condor daemons to shutdown individually and not be restarted by the Master.

```
STARTD.DAEMON_SHUTDOWN = \
   State == "Claimed" && \
   Activity == "Idle" && \
   (CurrentTime - EnteredCurrentActivity) > 600
MASTER.DAEMON_SHUTDOWN = \
   STARTD_StartTime == 0
```

Glidein Condor Configuration File

Retrieving Job Attributes

> **Problem**: How can we get additional information about Condor-C jobs when they are executing on Glideins?

> Use only existing, reliable Condor mechanisms.

Solution: Copy Attributes List

> Provide a list of attributes to copy back to Condor-C job's ClassAd on submit node.
> Resolves $$(<Parameter>) at runtime.

```
CONDORC_ATTRS_TO_COPY = \
    MATCH_FileSystemDomain, \
    MATCH_UidDomain, ....
```
Submit Side Condor Configuration File

```
+Remote_Env = \
    "FileSystemDomain=$$(FileSystemDomain)"
```
Condor-C Submission File

Scalability Issues

> **Problem**: How can we increase the number of jobs per central manager and GCB node?

> Preliminary tests showed only 1,000 jobs could reliably be submitted for each Tier-1 central manager.

Solution: Internal Improvements

> Improved core ClassAd library: faster attribute look-ups and parsing.

> Re-factored scheduling algorithms.

> Increased scalability of GCB libaries.

> Localhost communication optimizations.

> Effort is still ongoing...

Summary

> Network Probe

> Daemon Shutdown Expressions

> Condor-C Copy Attributes List

> Scalability Improvements

> Questions?