

Introduction to Pilot Jobs

Mats Rynge, rynge@isi.edu

USC / Information Sciences Institute

Igor Sfiligoi, isfiligoi@ucsd.edu

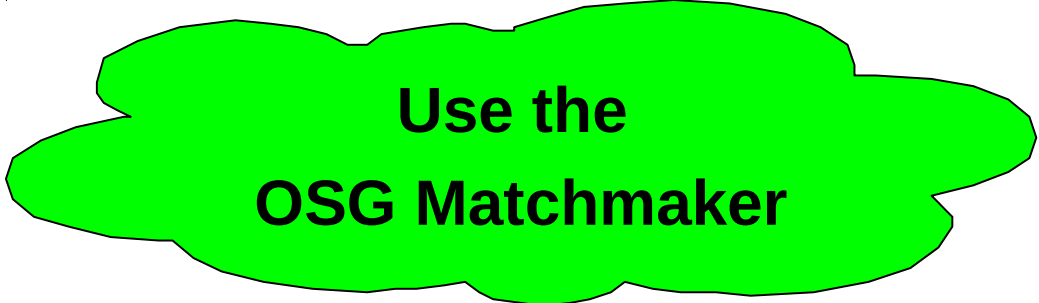
University of California San Diego

What you learned until now

- Scientific problems need a lot of CPU time
 - Using a batch system like Condor is a must
- Local compute resources are often not enough
 - Get some/most of the needed CPU cycles from the Grid

Moving to the Grid

- What resources are out there?
- How do I know if they have the needed software installed?
- How do I know where will my jobs finish the soonest?



**Use the
OSG Matchmaker**

Using OSG MM/Condor-G

- Limited functionality
 - Need to specify output file names
 - Limited monitoring
 - No checkpointing
- Very heterogeneous environment
 - Discover which apps are installed
 - Discover which libs are installed

Can we make life easier?

- More Condor (without G) like?

Yes



Pilot systems

What is a pilot system?

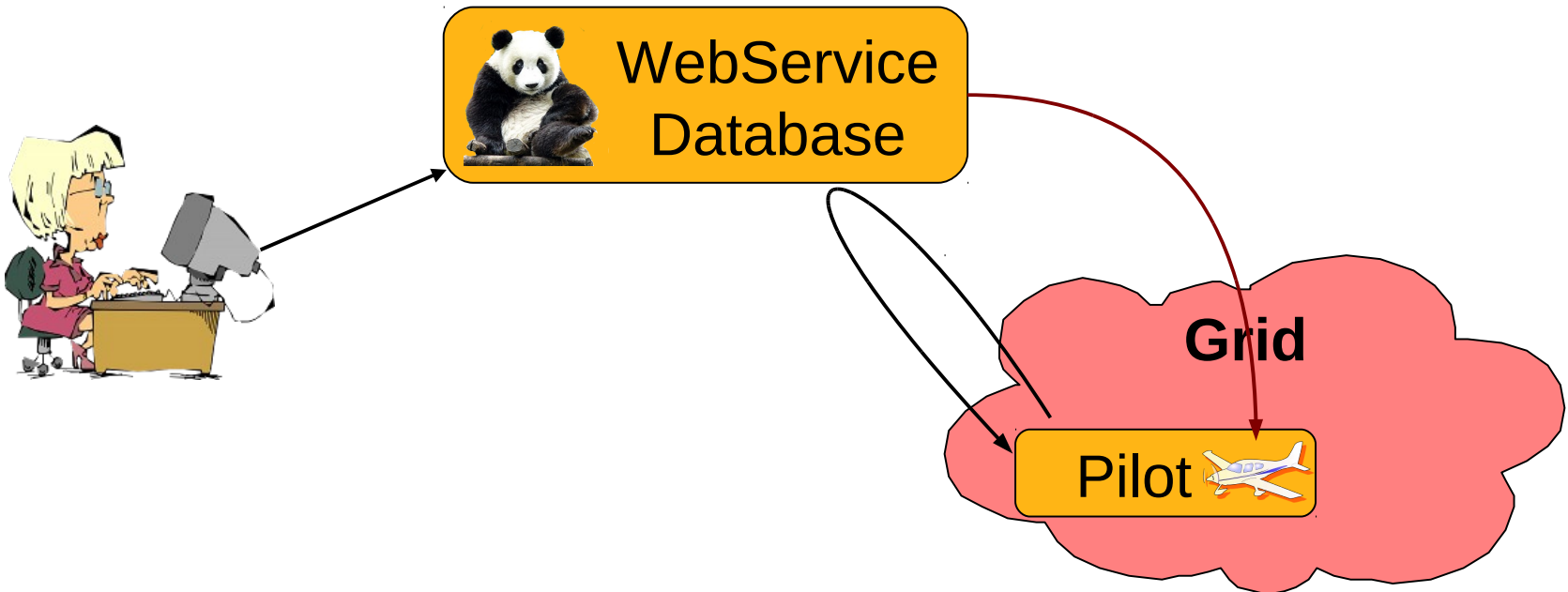
- An infrastructure that creates a virtual-private batch system
 - It hides “the Grid” from the users
- For users like a local system
- All Grid interaction taken care of in a separate layer (hidden from users)

Pilot principles

- Never send user jobs directly to the Grid
 - Send pilots instead
- After a pilot lands on a Grid resource
 - Detect local resources
 - Stop execution if resource is bad
- Pilot fetches a user job
 - Fetch currently the most important job (late binding)
 - Start and monitor the job

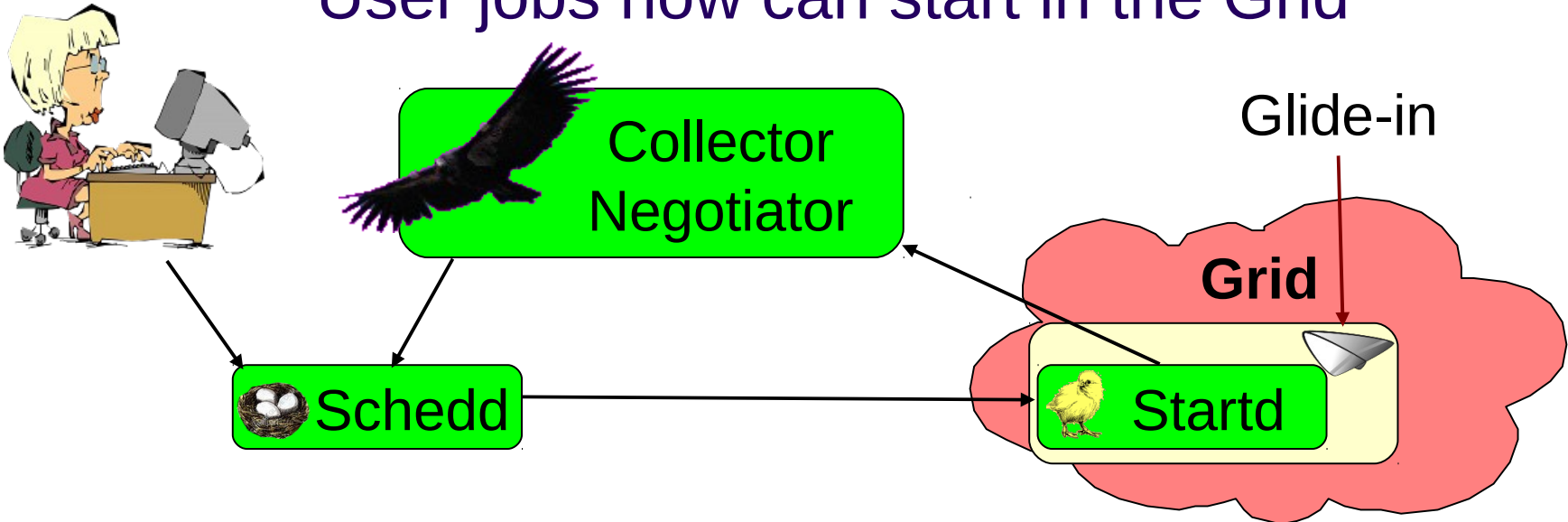
PANDA

- A Web Services based pilot system
 - Database driven
- Same concept, different implementation



Glide-ins

- Glide-ins are Condor-based pilots
- A glide-in is a Condor startd sent as a Grid job
 - When it starts, it joins the Condor pool
 - User jobs now can start in the Grid



Pilot system drawbacks

- Requires networking from the worker nodes
 - Pilots need to “call home”
 - Regular Grid jobs can live without
- Potentially wasting resources
 - When a pilot starts, there may not be any more user jobs waiting to be run

What's next?

- PANDA
- GlideinWMS
- GlideinWMS exercises