BOSCO Testing

Marco Mambelli – marco@hep.uchicago.edu

BOSCO-ITB meeting

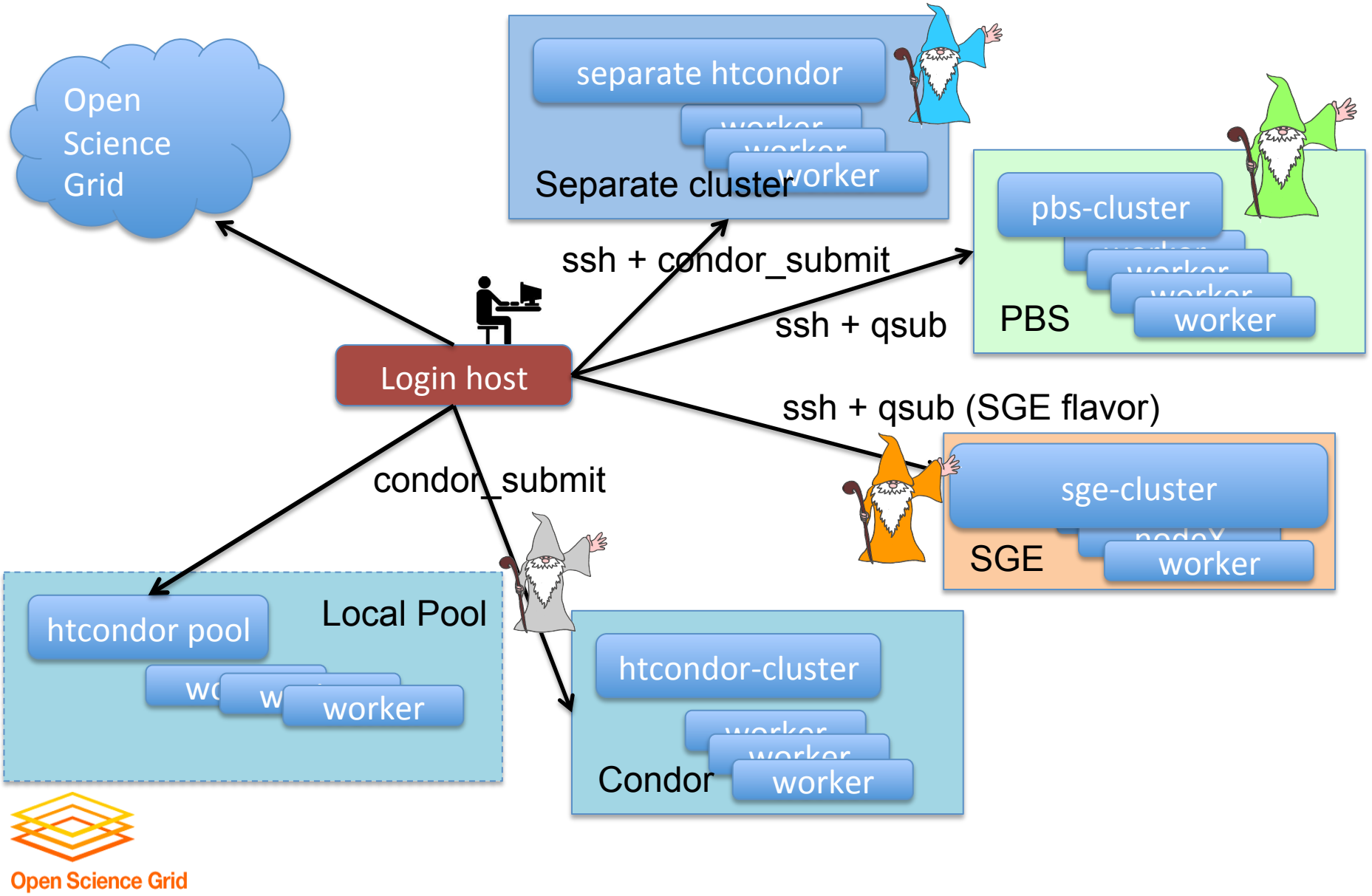
12/13/12

BOSCO

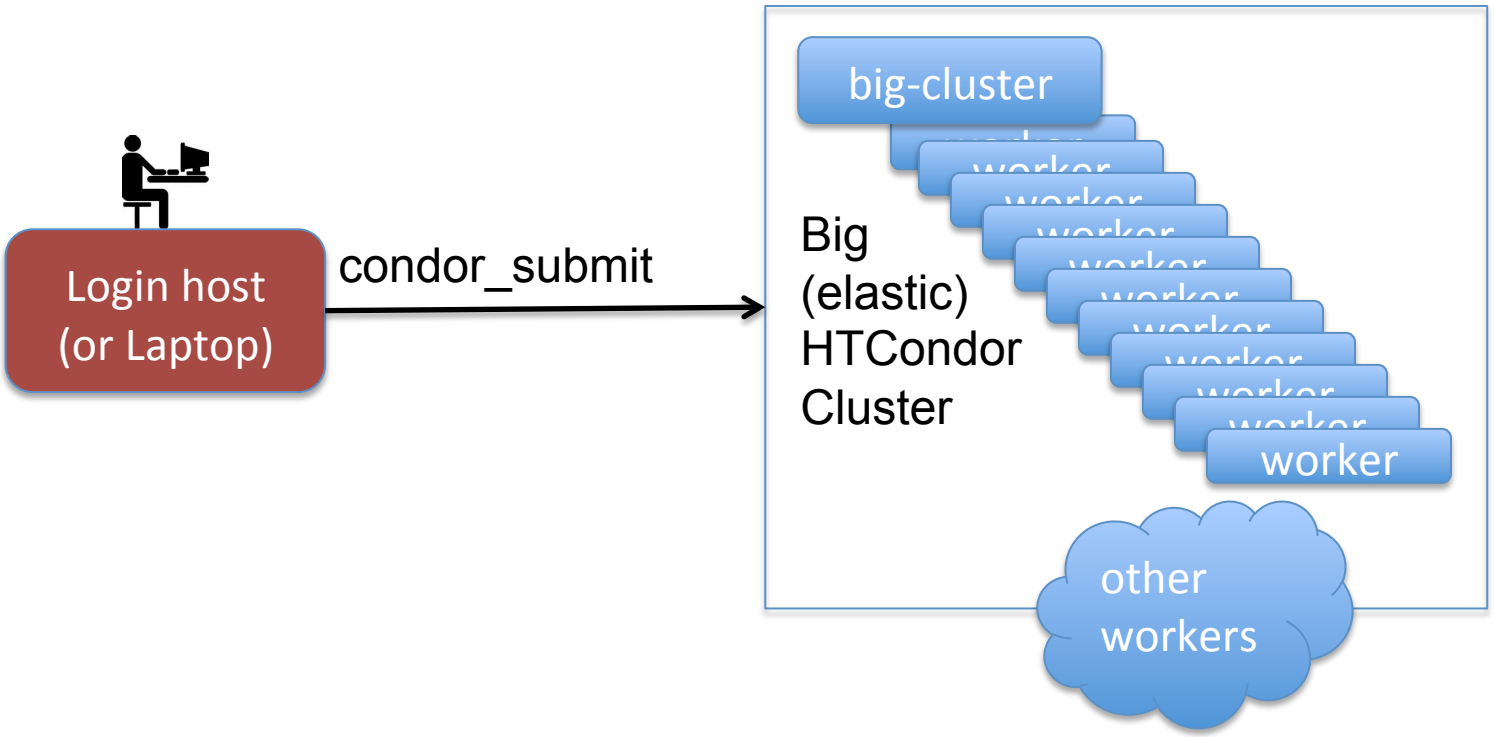
- Tool developed within Open Science Grid
- Easier High Throughput Computing
- Submit Locally (from your workstation)
- One Submit Model for Different Cluster Types

- http://bosco.opensciencegrid.org/
- https://twiki.grid.iu.edu/bin/view/CampusGrids/BoSCO

Resources in a Campus - diverse



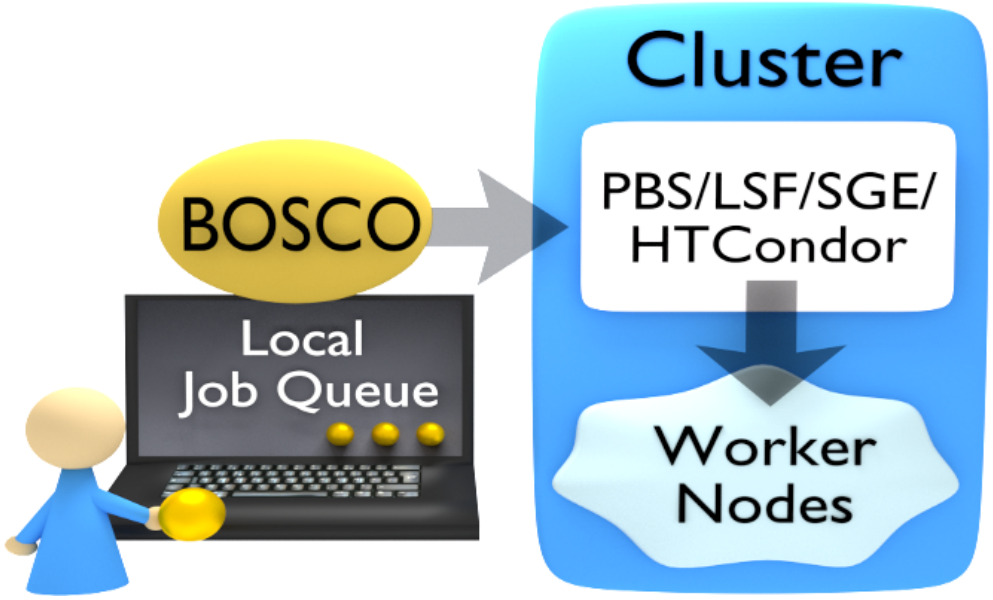
Submit locally, compute globally



Open Science Grid

Components of a BOSCO system

- BOSCO Submit node
- BOSCO Resources
  - Cluster Login Node
  - Cluster (worker nodes)



Courtesy of Soichi Hayashi

Use cases / Test cases

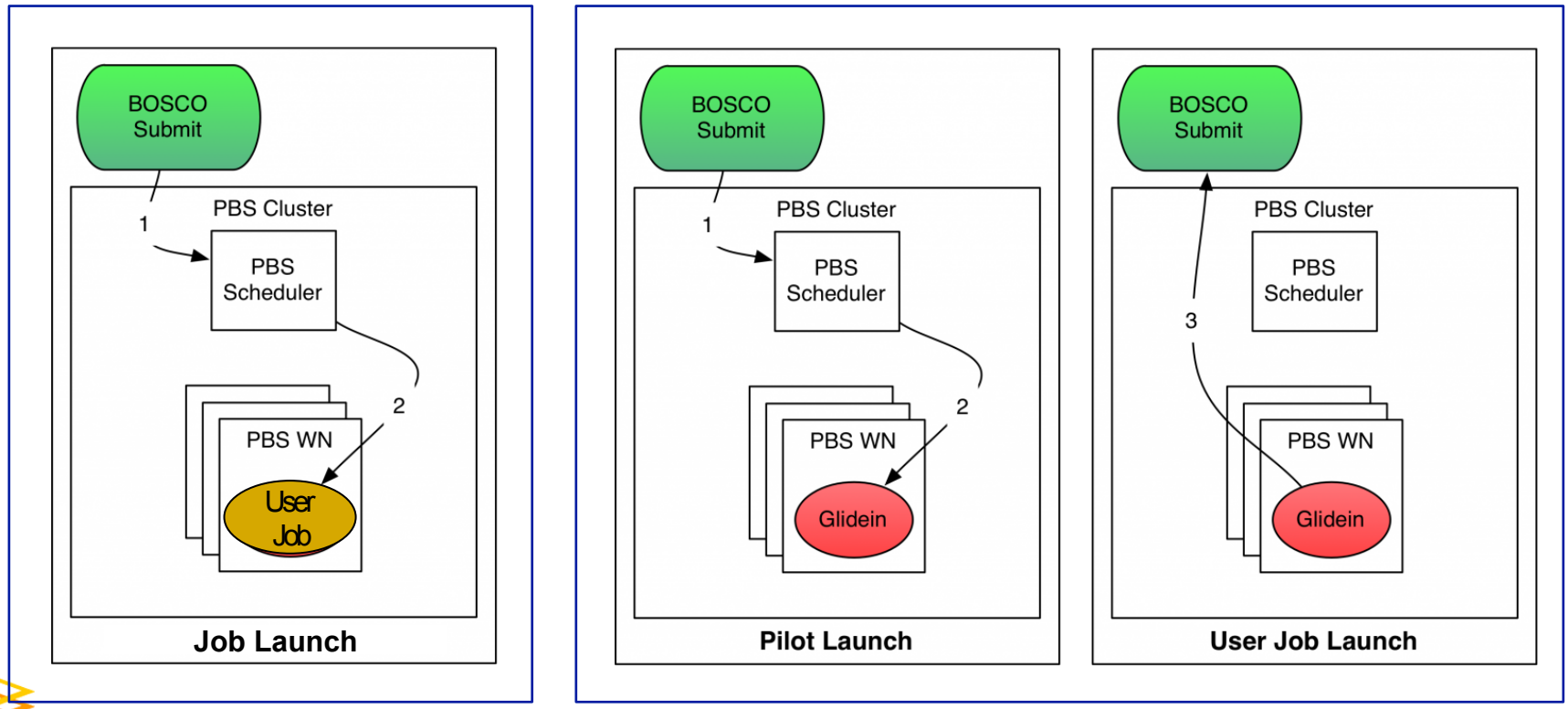
- Single user, single cluster
- Multi cluster (via Campus Factory)
- Multi user and integration with a HTCondor infrastructure

SINGLE cluster vs. MULTI cluster

▪ Direct submission of “grid” universe jobs (GHAP via SSH), one resource at the time	▪ “vanilla” univers jobs using HTCondor glideins (submitted via GHAP via SSH)
▪ SSH to login node of the BOSCO resource	▪ SSH to login node of the BOSCO resource
▪ No network service	▪ Using port 11000 (no firewall blocking it, no other service using it)
▪ BOSCO submit node can have no inbound connectivity	▪ BOSCO submit node reachable from BOSCO resources
▪ No outbound connectivity from the BOSCO resource	▪ Worker nodes require outbound connectivity and the ability to reach port 11000 of the submit node
▪ Can transfer executable, input files and output files	▪ Can transfer executable, input files and output files

Differences in job submission

- Direct “grid” universe job batch
- “vanilla” universe job using HTCondor glideins: pilot + job



Open Science Grid

BOSCO SINGLE-user vs. BOSCO MULTI-user

- Installed by user
- User manages BOSCO
  - BOSCO started as User
  - Contributing clusters (BOSCO resources) added by User
- User must have SSH access on all BOSCO resources
- Only User can submit jobs to the HTCondor pool of BOSCO
- No choices because it must be easy to install and run for scientists without system administration experience

- Installed as administrator (root)
- Administrator manages BOSCO
  - BOSCO started as root
  - Contributing clusters added using a single service account
- SSH access via group service account (negotiated by admin)
- All the users on the system can submit jobs to the HTCondor pool of BOSCO
- More flexible because there may be more customization to add BOSCO in the Campus Grid

Open Science Grid

Requirements

- BOSCO Submit Node
  - RH5/6, Debian 6, Mac OS X 10.6 or later
  - Outbound SSH
  - Ability to receive connections on port 11000 *
- BOSCO Resources (Clusters)
  - RH5/6, Debian 6
  - Local Resource Manager: PBS: Torque and PBSPro, HTCondor 7.6+, LSF, SGE
  - Incoming SSH connection (Cluster Login Node)
  - Connecting back to the BOSCO Submit Node (Cluster Worker Nodes) *

* Requirements in green are only for “vanilla” jobs, using the Campus Factory component of BOSCO

BOSCO standard tests

- Functionality
  - Single cluster (and all the rest)
    - Install/un-install
    - Add/remove BOSCO resources
    - “Grid” universe job
  - Multi cluster only
    - “vanilla” universe job
  - Multi user only (in Campus)
    - Flocking
    - Monitoring
- Scalability and reliability
  - Completion of multiple jobs
  - Running continuously

Open Science Grid

Help us improving BOSCO!

- Test procedures and test results are available online: https://twiki.grid.iu.edu/bin/view/CampusGrids/TestBoSCO
- It's a wiki page and you are welcome to add your test results or comments

- Questions or suggestions?
  Write to bosco-discuss@opensciencegrid.org