Integration Validation System

Suchandra Thapa
OSG @ University of Chicago
http://twiki.mwt2.org/bin/view/ITB/WebHome

Software Tools Group Meeting
2/17/09

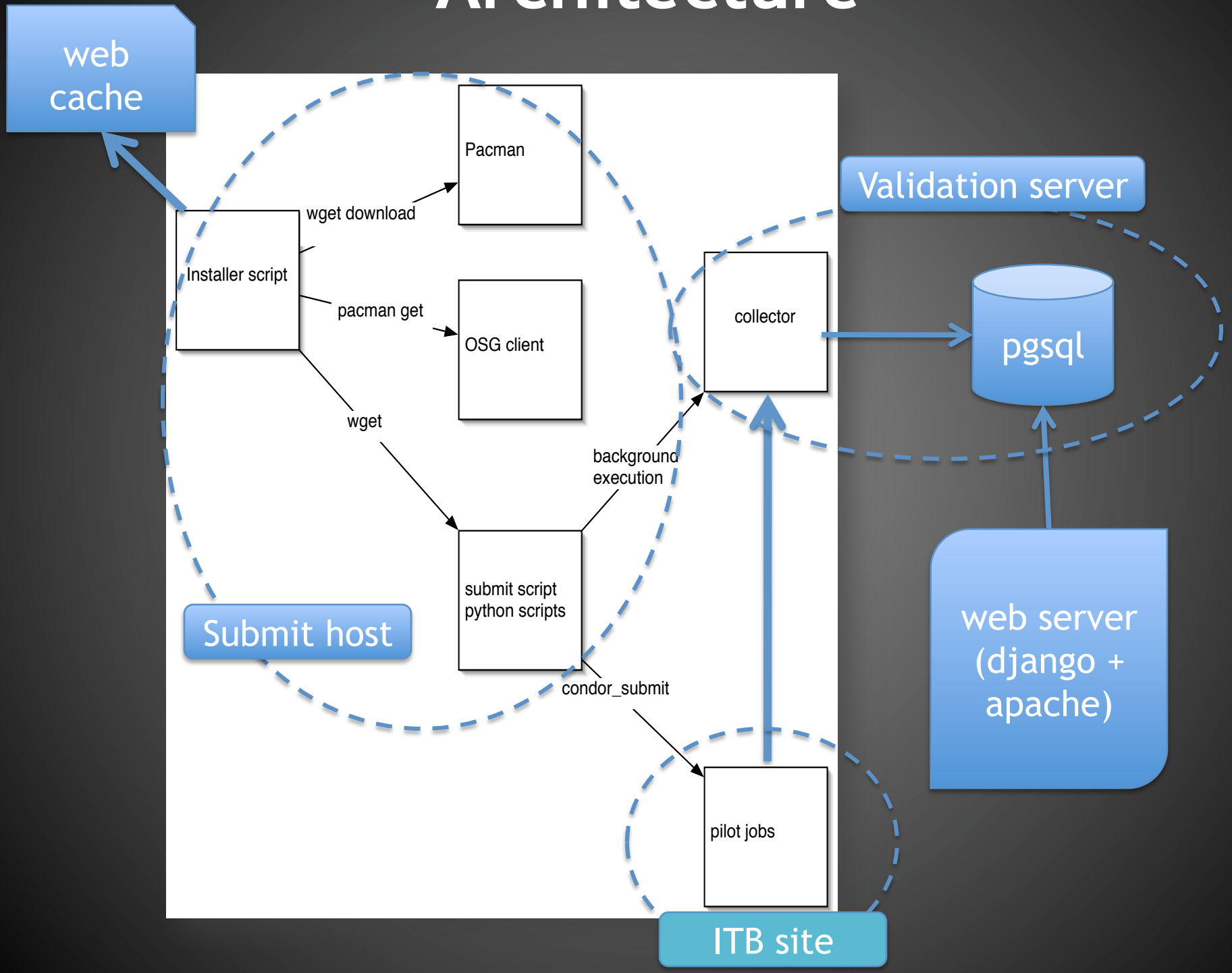
Background

- OSG Integration project has as a goal speeding the release process for OSG software components
- Connected to this is providing automation to the currently labor-intensive ITB validation process
- Need for capturing current status and validation history (by VO, by service, by site) during and at conclusion of ITB cycle
- Need for generating automatic workloads – a synthetic job load system as discussed in OSG Blueprint meetings
- People:
  - Suchandra Thapa
  - Robert Veitch
  - Rob Gardner

Validation System

- Features
  - System to populate ITB sites with computational tasks according to a random and adjustable job model
  - Message reporting system to track job states, users & site performance
  - Database backend populated with messages passed from job submitters and execution pilots
  - Web front-end for viewing validation progress and report generation
  - Accommodate other functional tests (service checks, RSV probes, etc) to correlate job and RSV-like data
    - in a manner to create automated site validation table
  - Provide downloadable kit containing client tools and instrumented submitter and pilot so that any VO member can quickly access and monitor synthetic jobs on the ITB

Architecture



Infrastructure Pieces

- Web cache
  - Starter script
- SVN server
  - Submit host prep script
  - Pilots
  - Pilot submission script
  - Synthetic job payloads
  - Validation job payloads
- Validation server host
- Database server
- Web framework server

Starter & Submission Scripts

- Starter script
  - Tiny shell script user grabs from web & executes
  - Checks out main submitter & job pilot scripts from SVN which are expected to remain under active development
  - No privileges required
- Submit host preparation & exec scripts
  - Grabs Pacman, installs locally
  - Pacman installs VDT-Client and prepares for condor-G submission
  - Script for validating the submit host installation
  - Script for submitting to “active” ITB sites (kept up to date by server)
- Dependencies
  - Supported VDT platform
  - Grid user certificate & membership in a supported VO

Job Pilot

- Python script downloaded by starter script
- Reports job state information to validation server
- Extensible interface to allow new payloads to be added

Synthetic Job Models

- The payload of the job
- Current payloads include:
  - Matrix inversion (synthetic cpu load)
  - Information probe (gets system information)
- Planned payloads
  - Testing I/O (network, disk, etc.)
  - Validating variables (OSG_DATA, OSG_SCRATCH, etc.)
- Selectable from submit host script
- Incidentally allows VO users to check whether sites support their VO

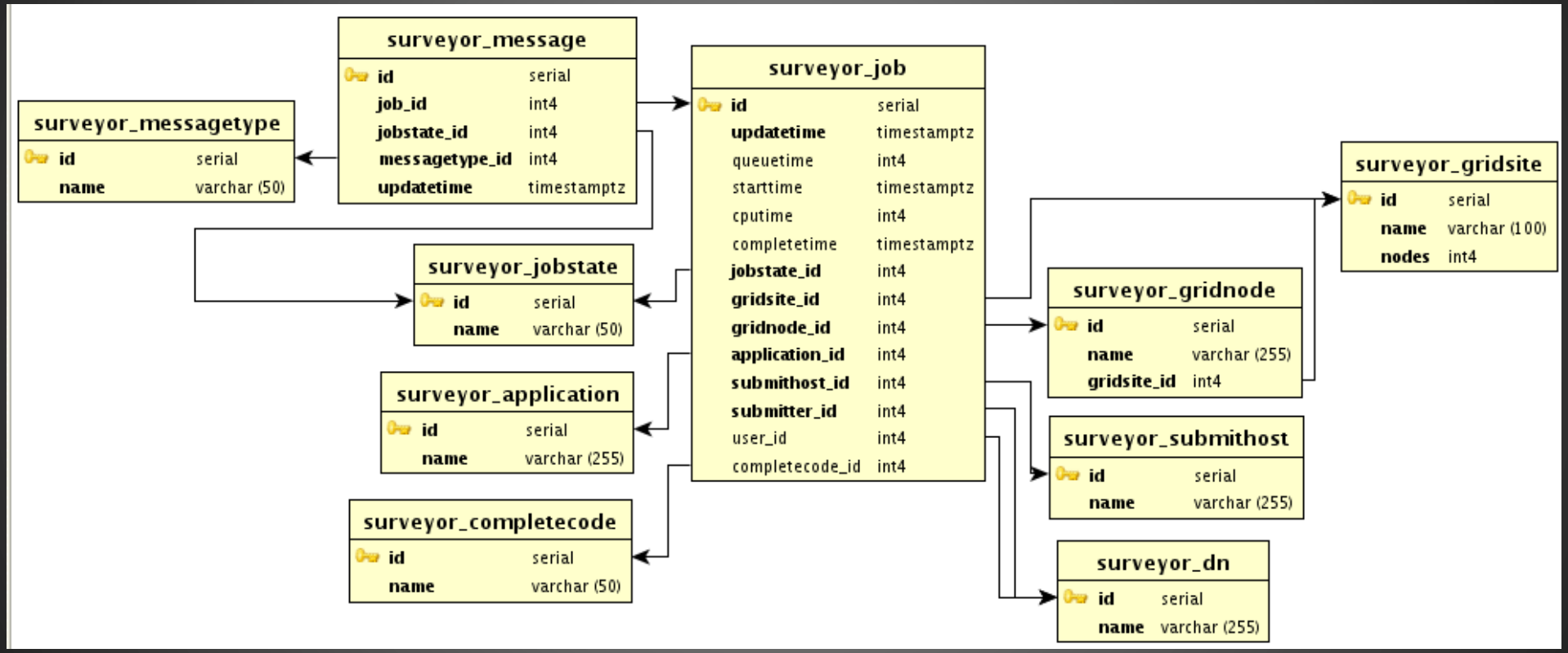
Service Validation Jobs

- Jobs targeted to test services available at ITB sites
- Will run on UC_ITB site but can be run on other ITB sites
- Will query services to create record of ITB availability

Validation Server and Database

- Built on pilot infrastructure
- Validation jobs are special cases of pilot jobs
- Information stored in database with other pilot information
- Application id allows validation information to be easily selected
- Website will have pages showing validation of itb resources
- Eventually this could be used as a semi-automated tool to validate itb releases

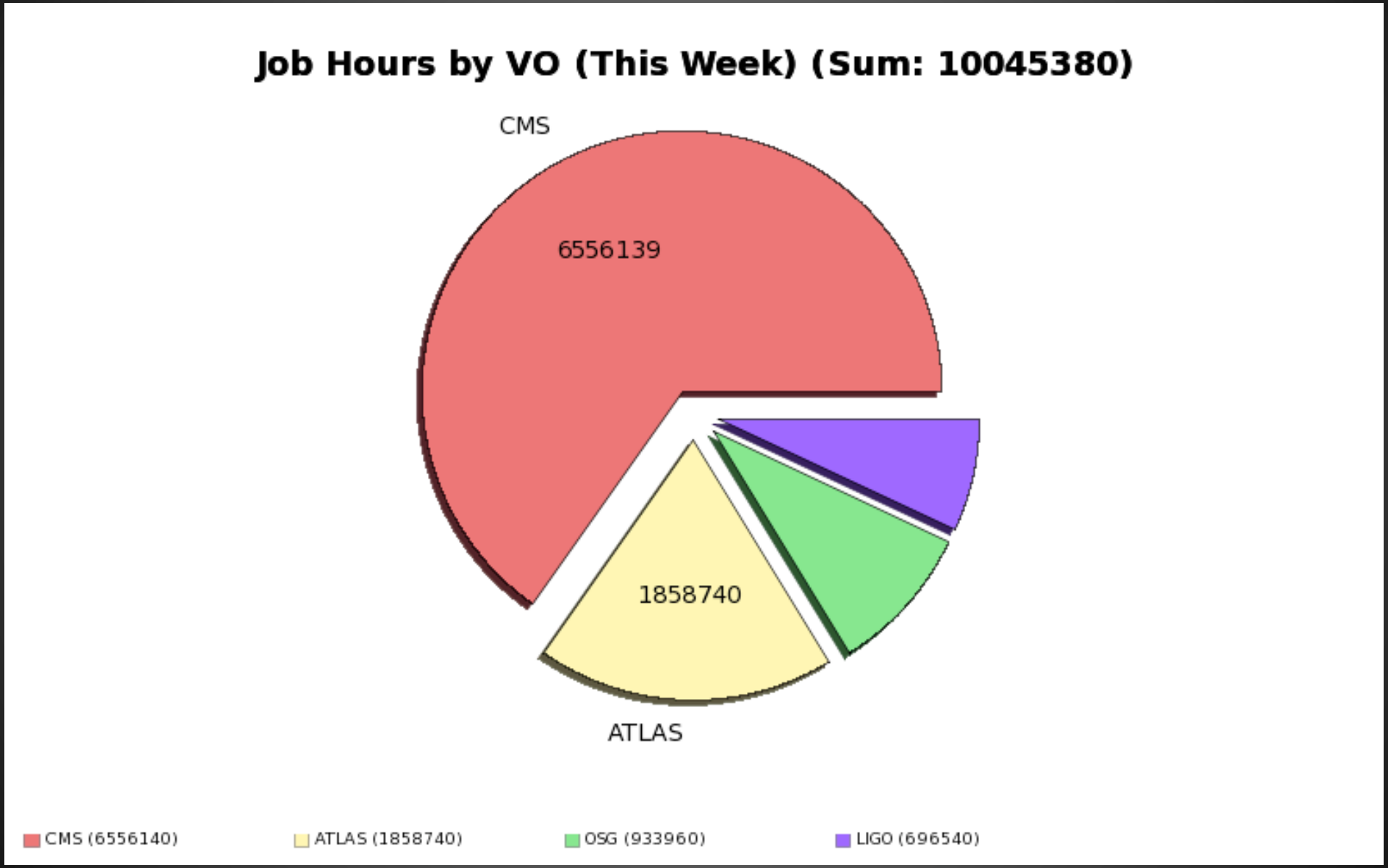
Database Schema



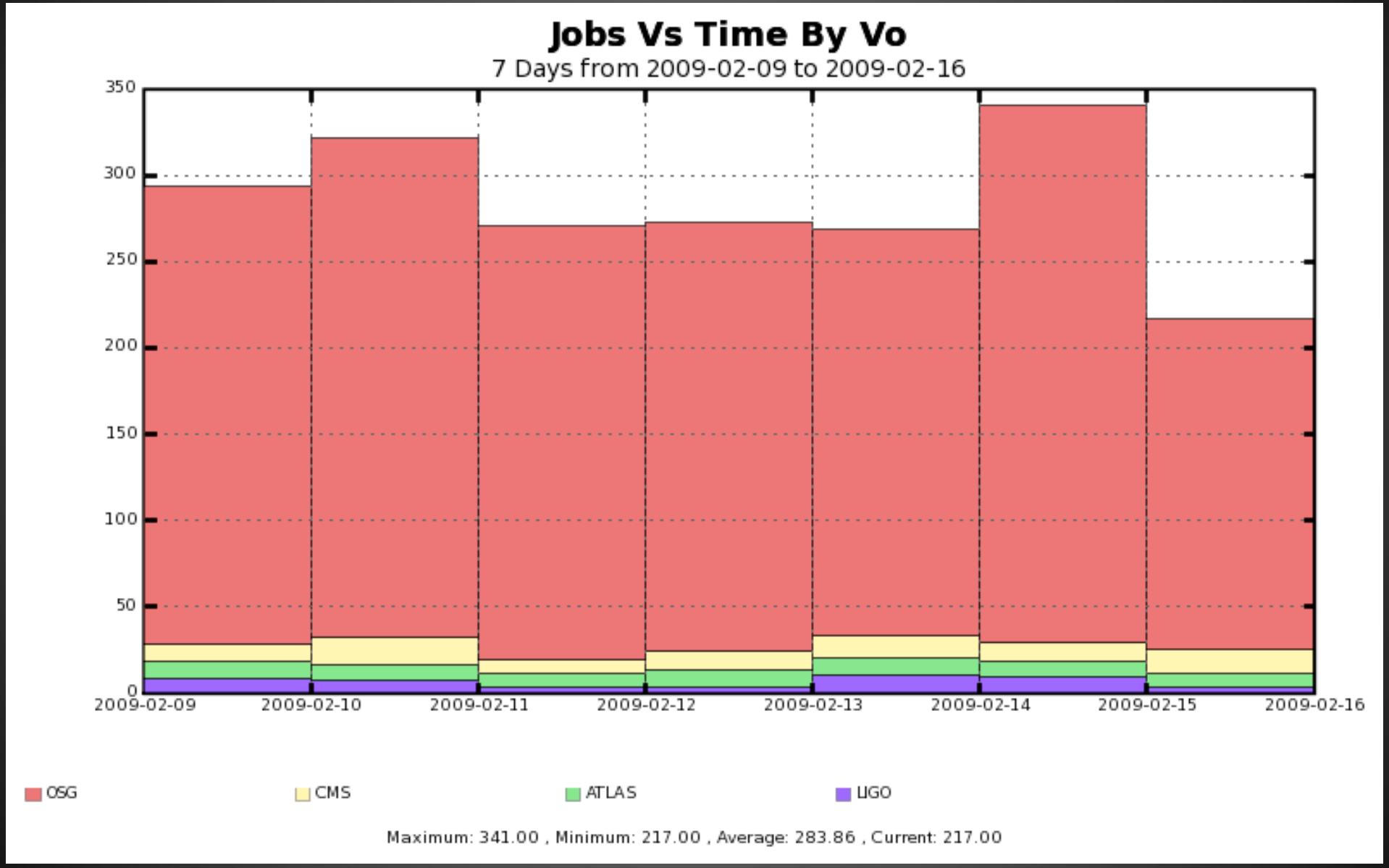
Web framework (Surveyor)

- Python + Apache
  - Mod_python back end
  - Django web framework
  - Easy integration with postgresql back end
  - Web caches and accelerator possibly if needed
- Graphtool
  - Used for graphing and displays
  - Generated using cron jobs at regular intervals
- Postgresql backend
  - Provides transactional support
  - Provides facilities for data consistency and integrity

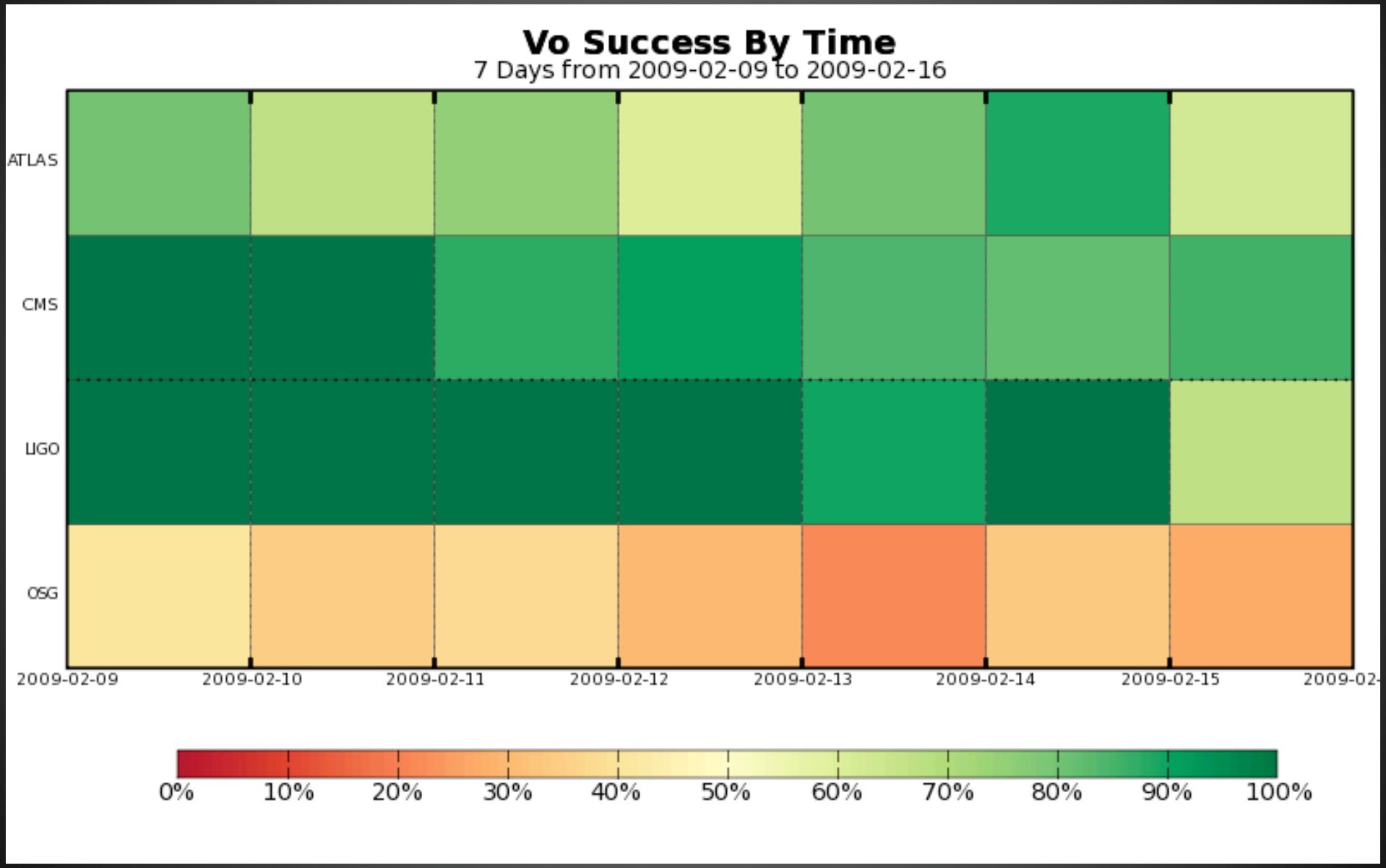
Simulated results



Simulated results



Simulated results



Next Steps

- Complete pilot infrastructure / make client tools more user friendly
- Run tests on itb sites and get some initial data
- Get first set of graphs and solicit feedback