IceProd 2

David Delventhal
University of Wisconsin-Madison
IceCube Collaboration
IceCube Computing

Medium size collaboration

- 2 data centers and several smaller clusters
- Most CPU compute is opportunistic
- Diverse computing infrastructure
Requirements

1. Cannot require root permissions.
2. Need more than just glideins.
3. Must support multiple OS’s and schedulers.
4. Local disk is usually not available.
5. Need to permanently record all job statistics.
6. Allow for both individual users and production.
7. Provide offline production environment.
8. Documentation of configuration parameters
IceProd 1

- Site-local python daemons backed by a central database
- Plugins for each batch system
- gridftp/http directly to/from main data center
- XMLRPC to communicate with jobs
- Store configuration and statistics for every job
IceProd 1 Issues

- A long series of patches to increase functionality
- Early design decisions hampering new ideas
- Multiple bottlenecks appearing:
  - DB - locking issues, queries per second
  - Queue - designed for 100 jobs per hour, not 10k
IceProd 2

Complete Rewrite

- Support for Python 3
- Focus on unit tests
- Local SQLite database at each site
- JSON Web API
- HTML based UI
- User Accounts
Scalable Web Server

**High-volume communications**
Web server can handle more than 10000 connections per minute on a single thread.

![Graphs showing query performance](image-url)
Pilot allocates extra memory in case job exceeds memory usage
Pilot can also run multiple jobs on the same node
Logs

Browsable logs on website
Includes links to entire logs.
IceProd 2: Status

Currently in internal beta

Plan to switch to production by year’s end

Upcoming features:

• Monitoring
• UI improvements
• File pre-staging
Questions?