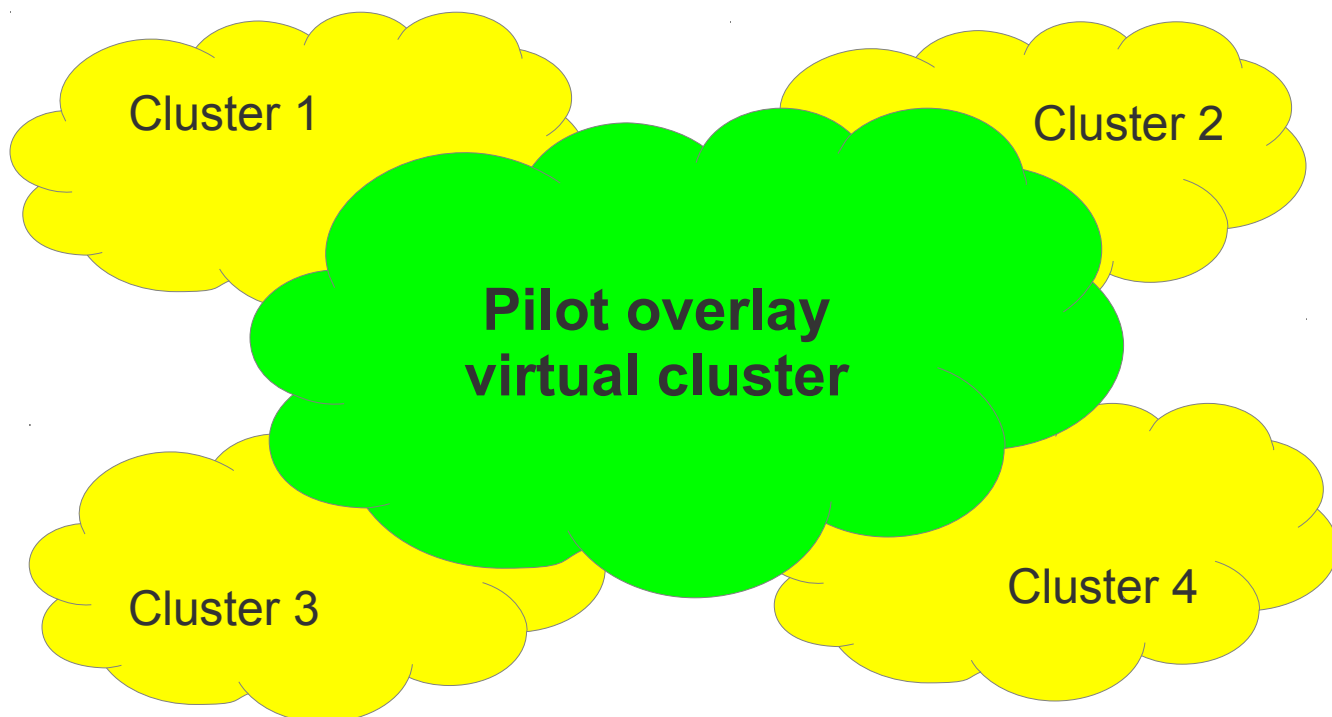


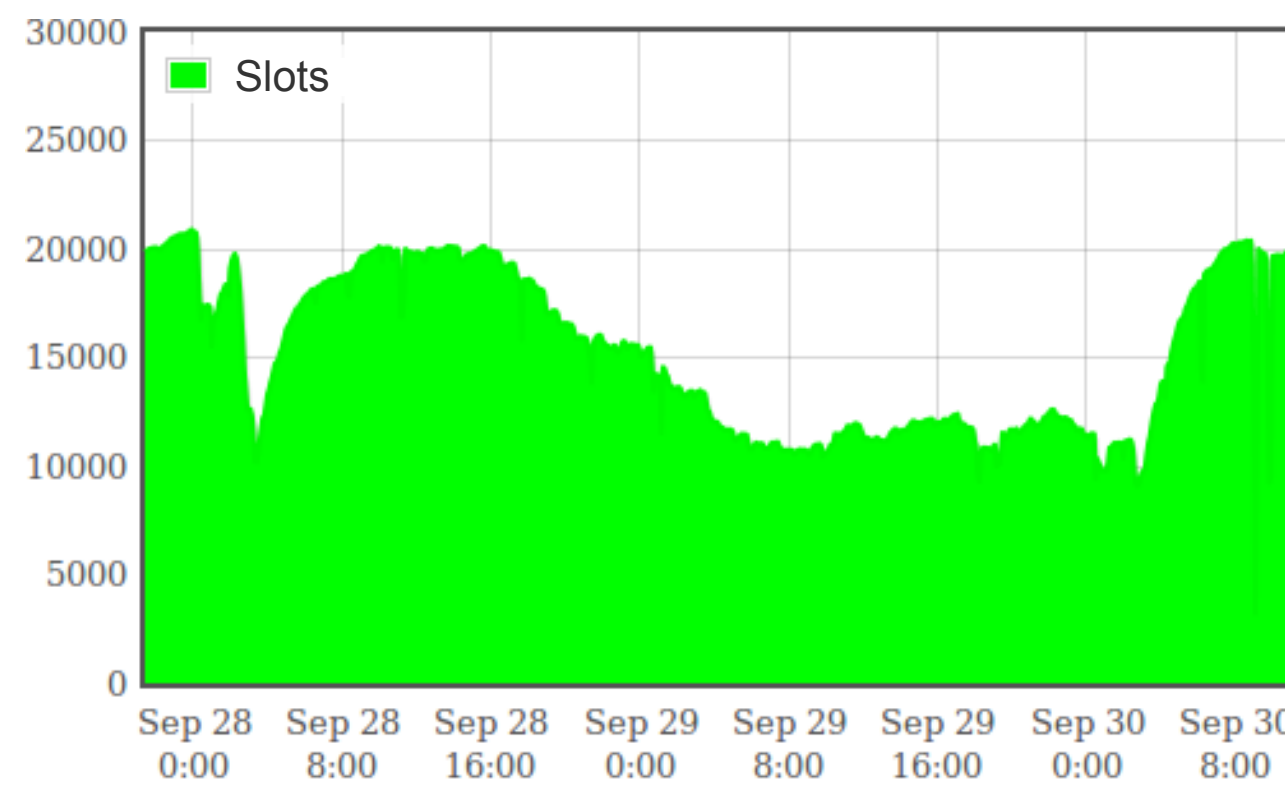
Estimating job runtime for CMS analysis jobs

by
I Sfiligoi and E Arias-Castro
 University of California San Diego, La Jolla, CA 92093, USA
 with support from the CMS Analysis Operations team

A pilot system creates an overlay batch system on top of other resources.



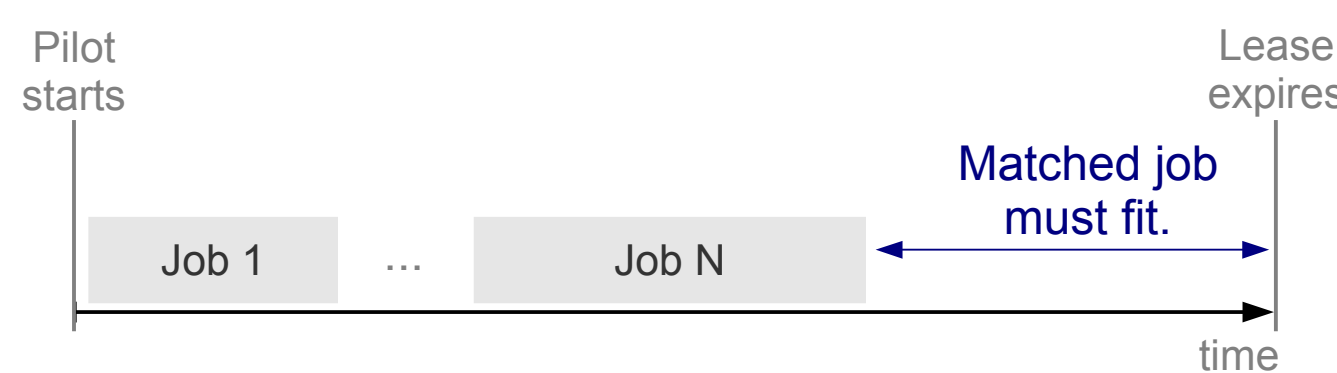
The size of the overlay system changes in time.



Change in number of slots triggered by two different mechanisms:

- The pilot provisioning system acting on user demand
- Lease on provisioned resource expires

Pilot resources have a **limited lifetime.**



Expected **job runtime needed** during scheduling.

Several studies showed that asking users to provide an accurate estimate is hopeless.

For example: doi:10.1007/11407522_14

We need an automated system for job runtime prediction.

Use historical information to predict future.

Generic automated job runtime prediction has been studied with mixed results.

See doi:10.1109/TPDS.2007.70606 for an overview.

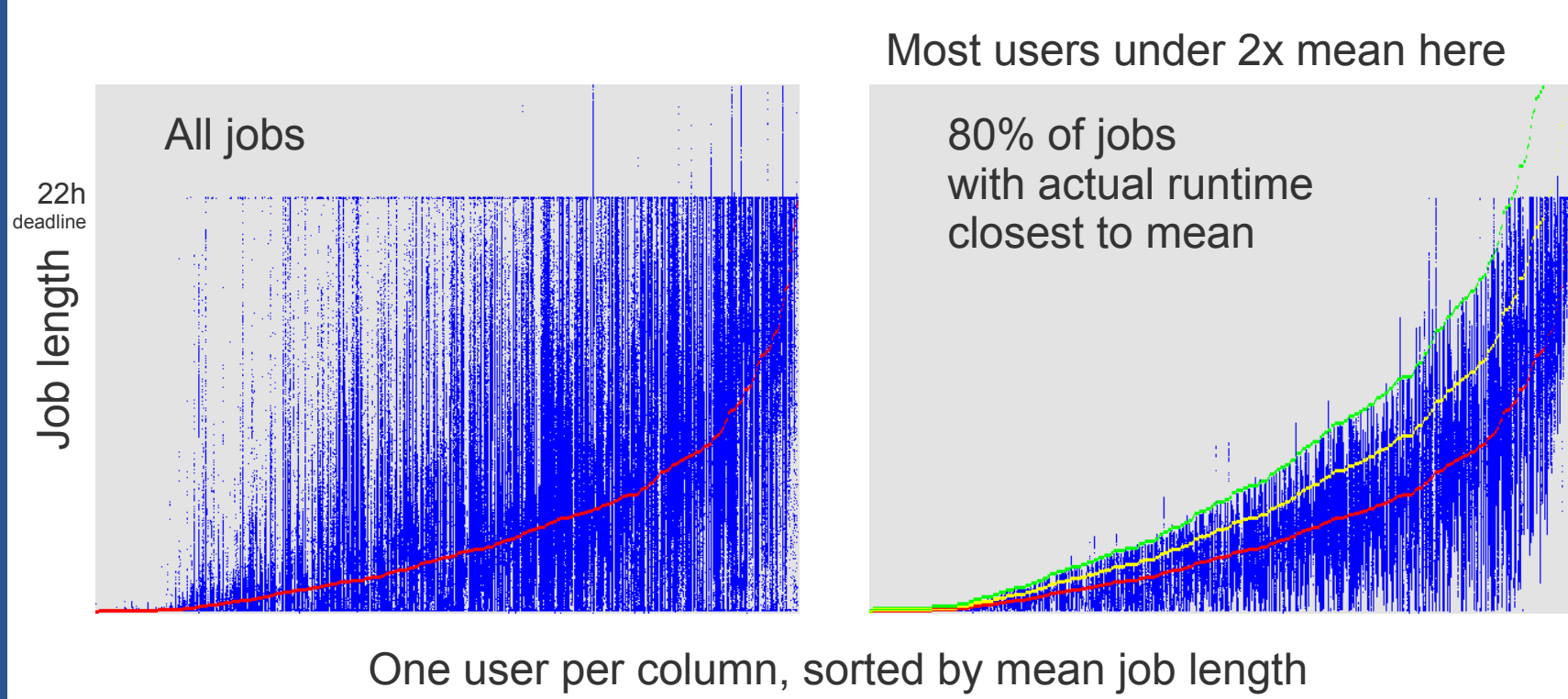
Production systems don't use it out of the box.

Can we still get a useful prediction inside CMS AnaOps?

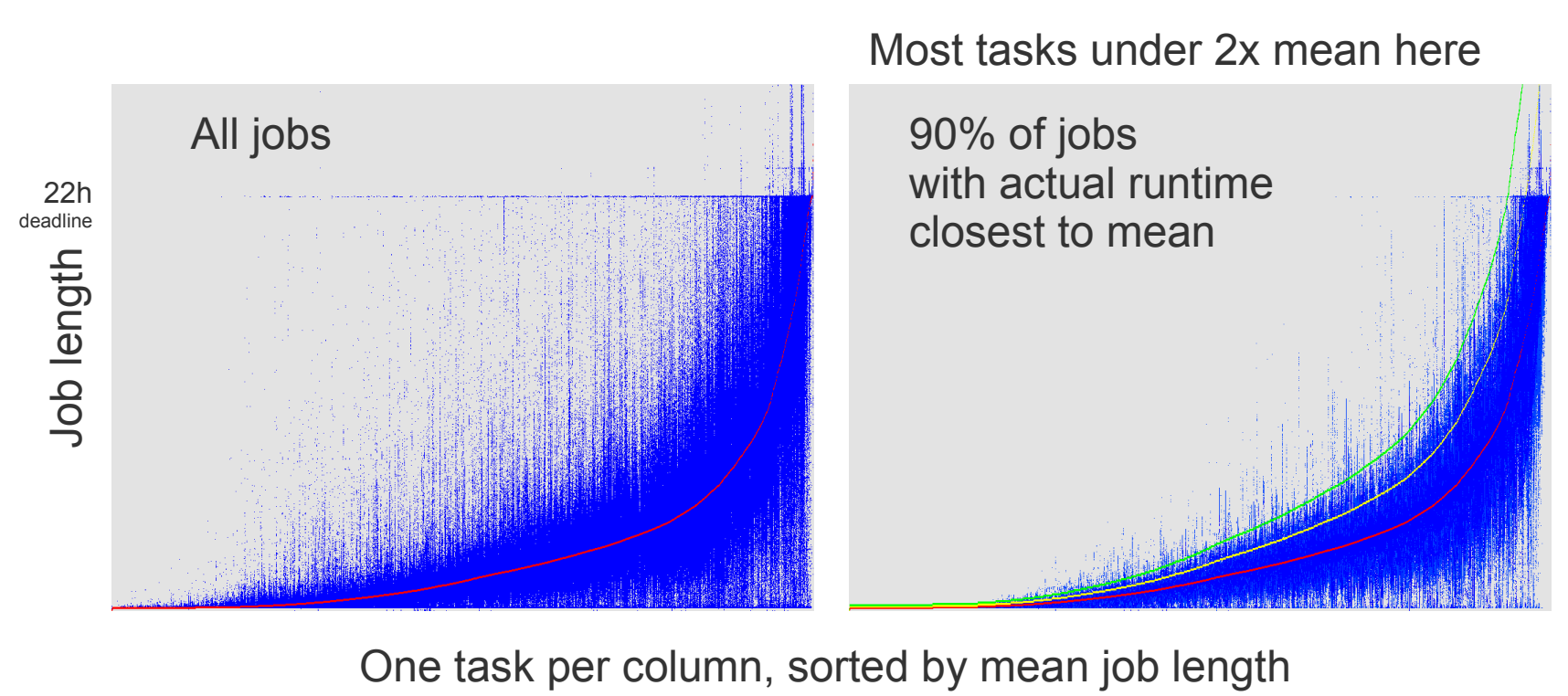
Analyzed the runtimes of CMS analysis jobs that have been submitted between Apr and Aug 2013 through UCSD's glideinWMS schedulers.

Number of tasks: 18k
 Number of jobs: 1.4M
 Number of users: 608

User alone already a decent discriminator



Task ID provides a very good estimate



User-based estimate can be used at any time

- Seems stable in time

Want to refine estimate as jobs in Task terminate

Predicted max job runtime already usable after few jobs from a Task

- Allowing to influence scheduling of many jobs

	Jobs shorter than 1.5x mean	Jobs shorter than 2x mean	Jobs still alive
First job	90%	94%	82%
First 10 jobs	94%	97%	63%
First 30 jobs	94%	98%	50%

CMS Analysis jobs exhibit clear patterns that can be used to predict max job runtime

- With high confidence level

Even at 2x mean, predicted max job runtime an order of magnitude lower than worst case for most jobs.

Wrong prediction for a few jobs acceptable

- Job will be automatically rescheduled
- Can use pessimistic estimate at that point