The Performance of the Czech National Grid Infrastructure after Major Reconfiguration of Job Scheduling System

Dalibor Klusáček Šimon Tóth

Faculty of Informatics, Masaryk University

October 27, 2014

The Czech National Grid

The Czech National Grid

- 10 000 CPU cores
- 12GB..6TB RAM per-node
- 1PB permanent storage
- 27PB long-term storage

Motivation

- workloads evolve throughout time
- small changes no longer sufficient
- three major shifts:
  1. memory heavy workloads
  2. increased average job length
  3. increasingly frequent wide jobs

Original Configuration

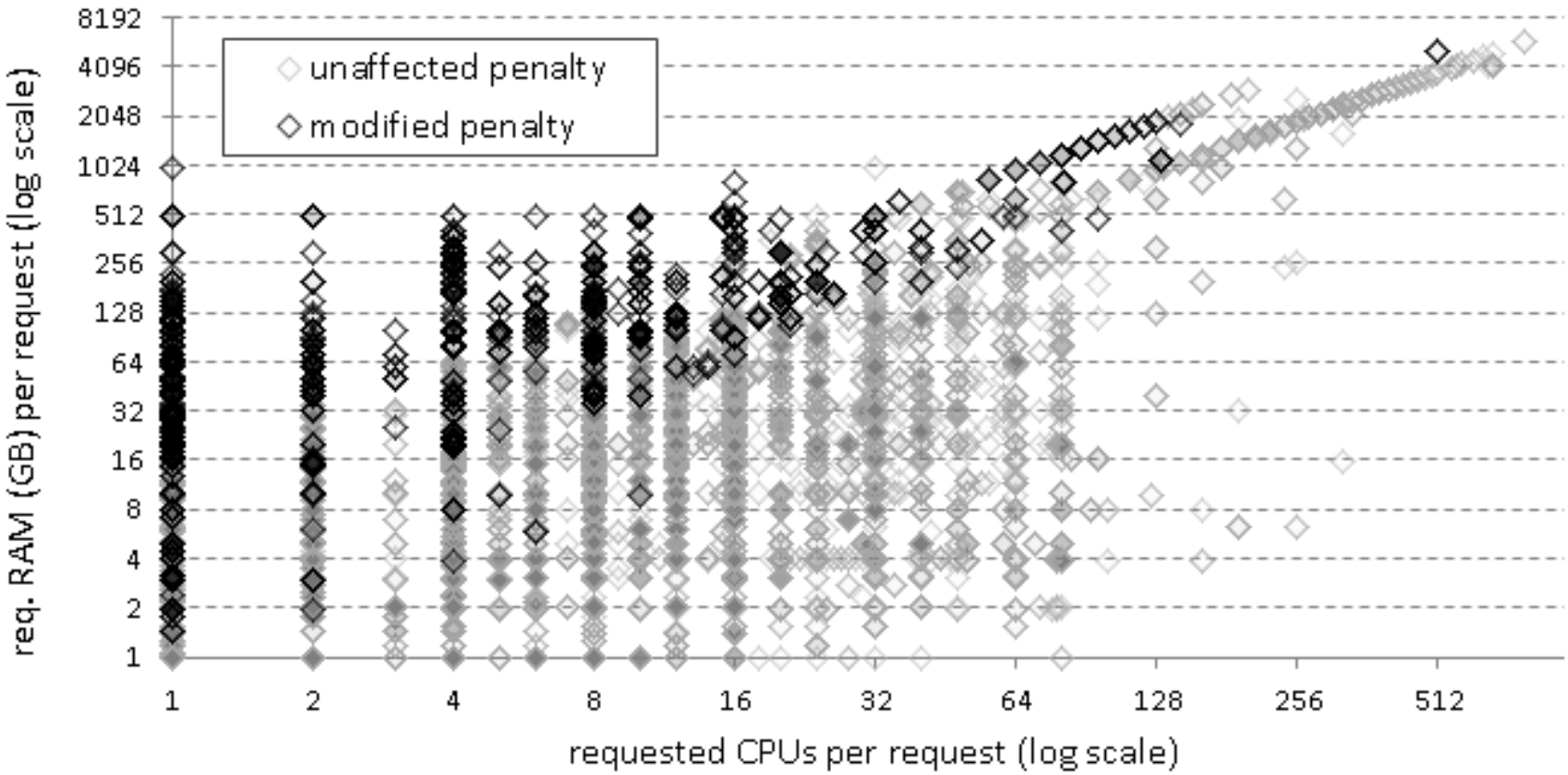
- designed by experts to fit original workload
- full support for very long jobs (up to 30 days)
- fair allocation of resources across users (long-term)
- fair wait times across users

Memory heavy workloads

- fairshare fairness model
  - only CPU accounted
- users with CPU heavy workloads penalized
- exploited by some of our users

Multi-resource fairness model

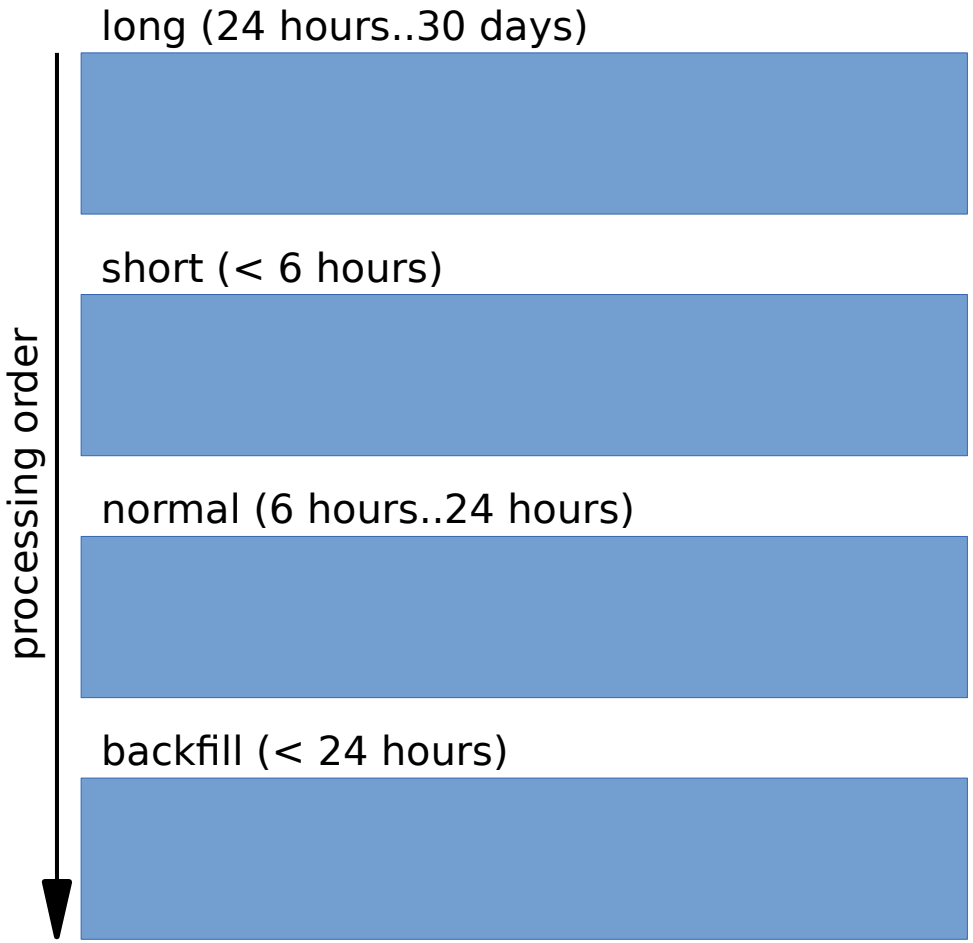
- based on dominant resource model
- 40% of users penalized



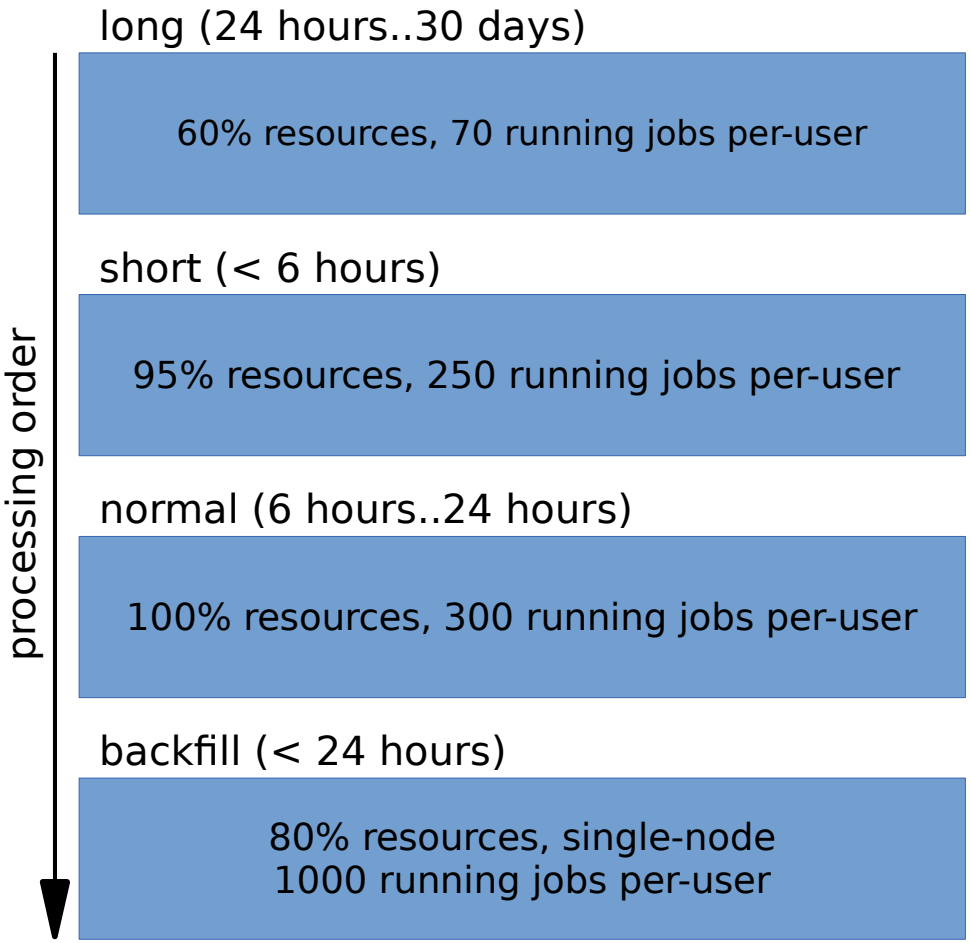
Increase in average job length

- originally unclear issue
- users experienced long wait times
- system experienced low utilization

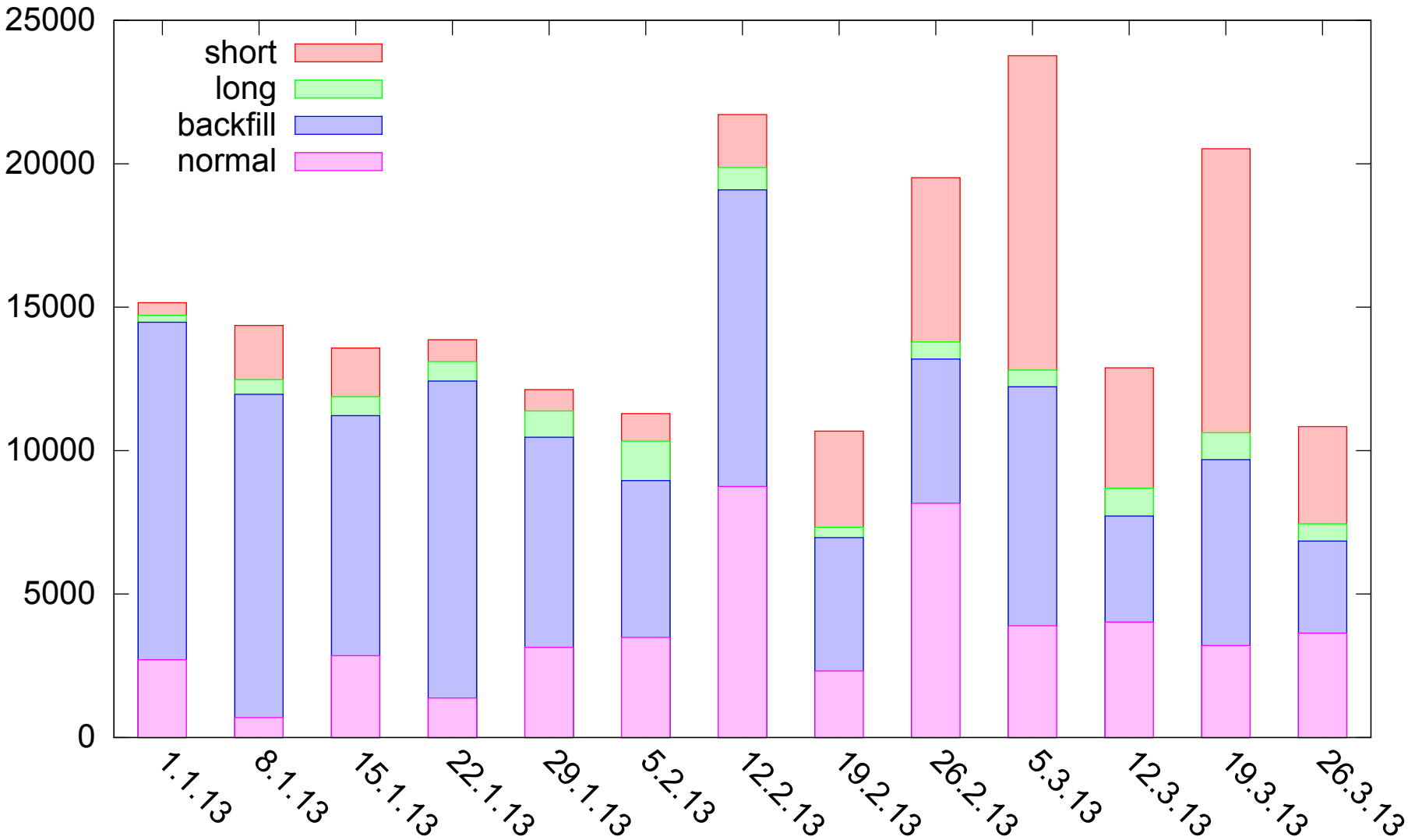
Original Configuration



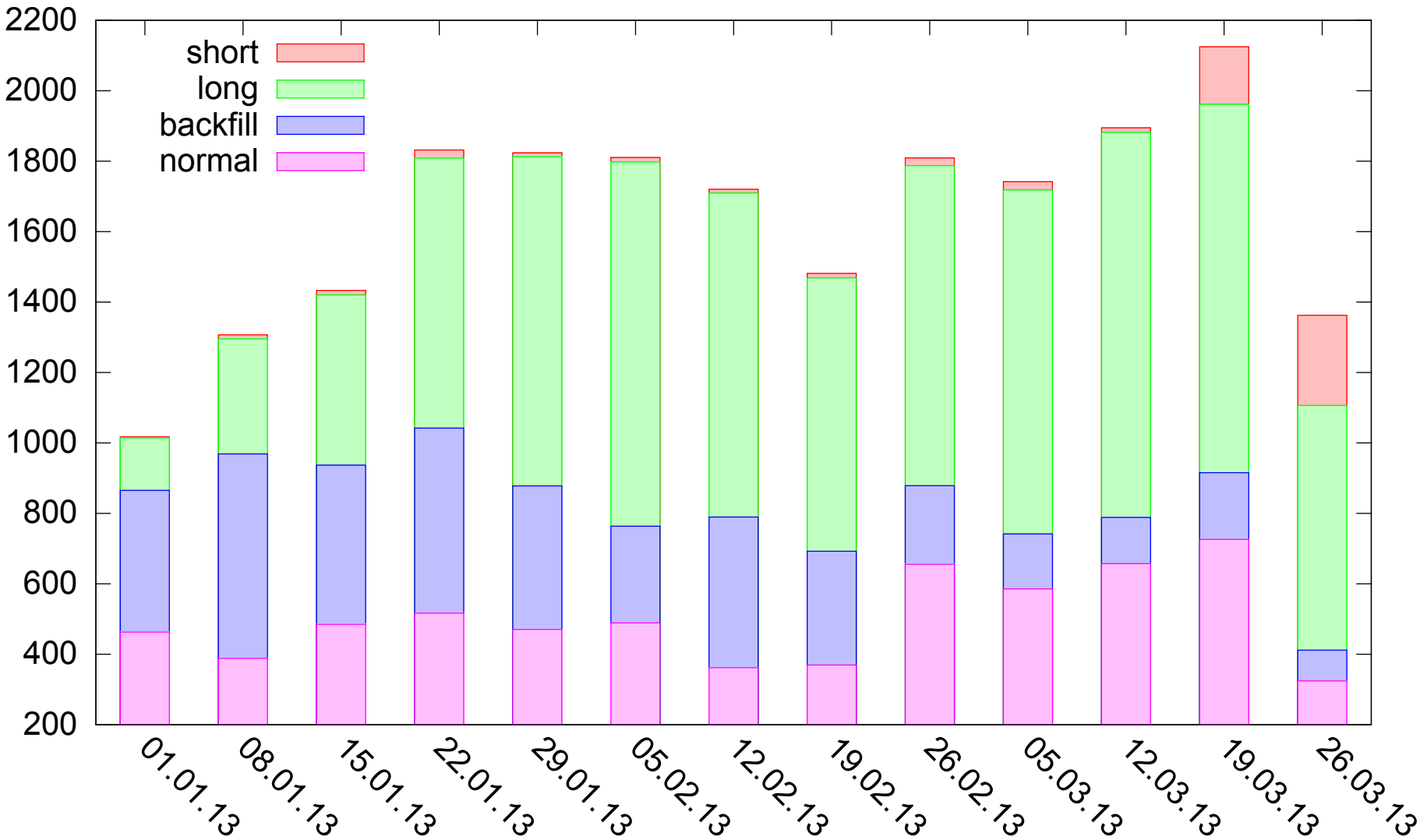
Original Configuration



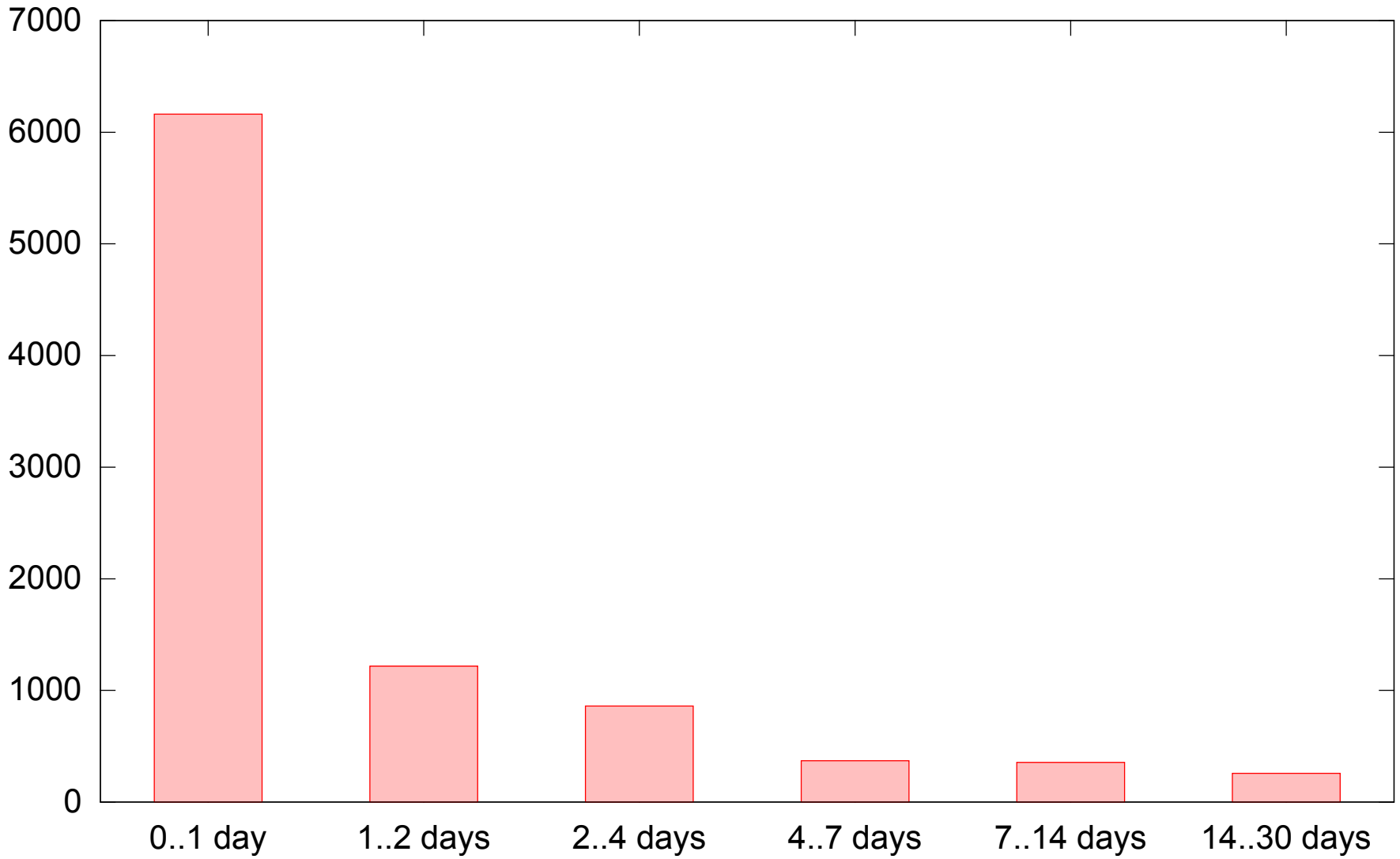
Analysis - job arrivals by queue



Analysis - CPU time by queue



Analysis - job runtime in long queue



Goals for new configuration

- increase resource pool for mid-range jobs (2-4 days)
- decrease resource pool for very long jobs (1 week +)

Simulation

- Alea Simulator

  `http://www.fi.muni.cz/~xklusac/alea/`

  - complex job specification
  - multi-queue configurations including operational limits
  - high-resolution output
- historical 5-month workloads
  - 376 722 jobs
  - 302 users

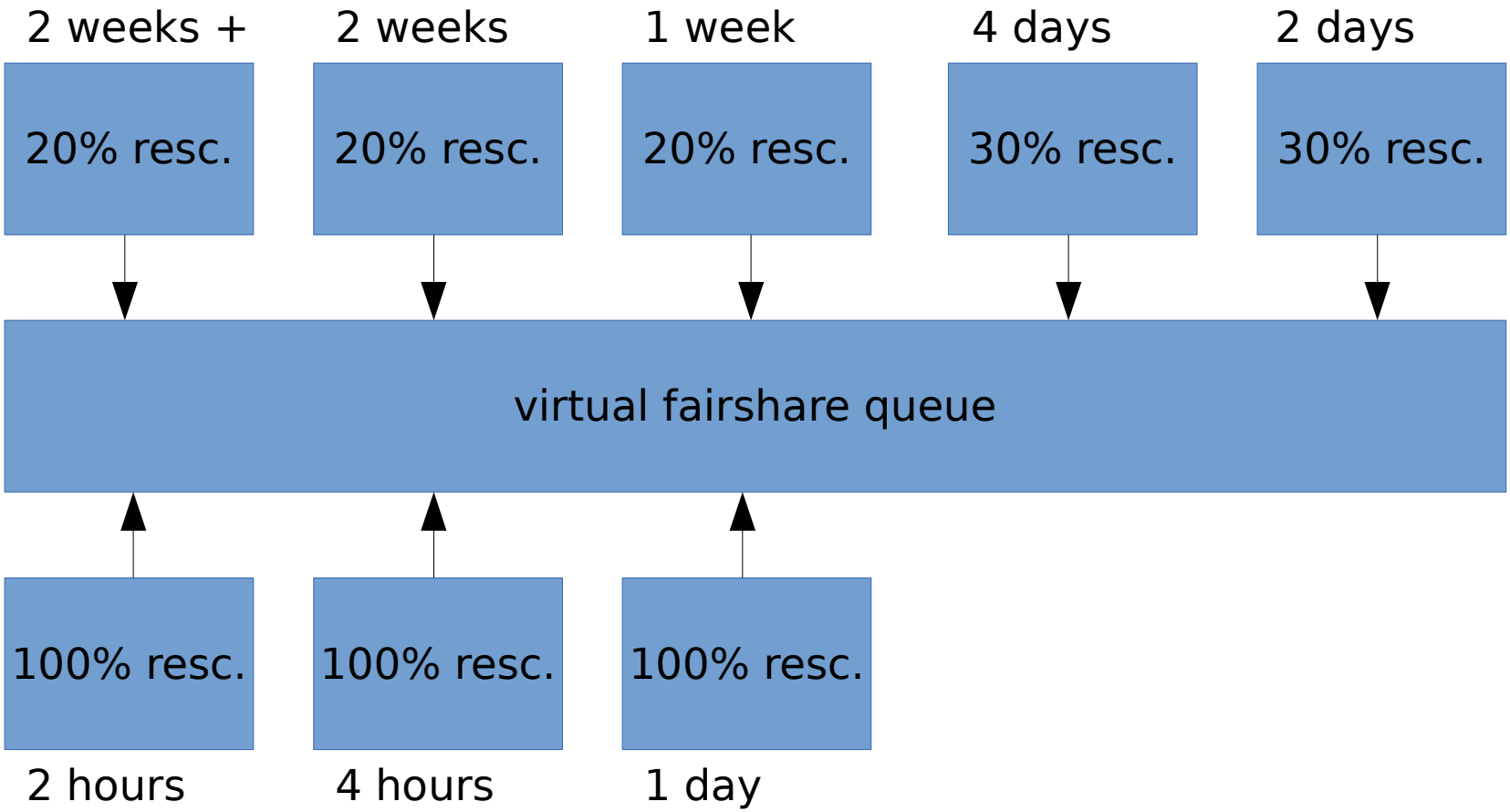
Simulation results

- 26.5% users with improved average wait time
- 19.2% users with deteriorated average wait time
- average improvement of wait time by 6.7 hours
- average deterioration of wait time by **55.7 hours**

Analysis

- impact of fairness policy heavily mitigated
  - fairshare ordering policy overridden by hard priorities
- impact of anti-starvation policy amplified
  - reservations blocking most resources

Experimental configuration



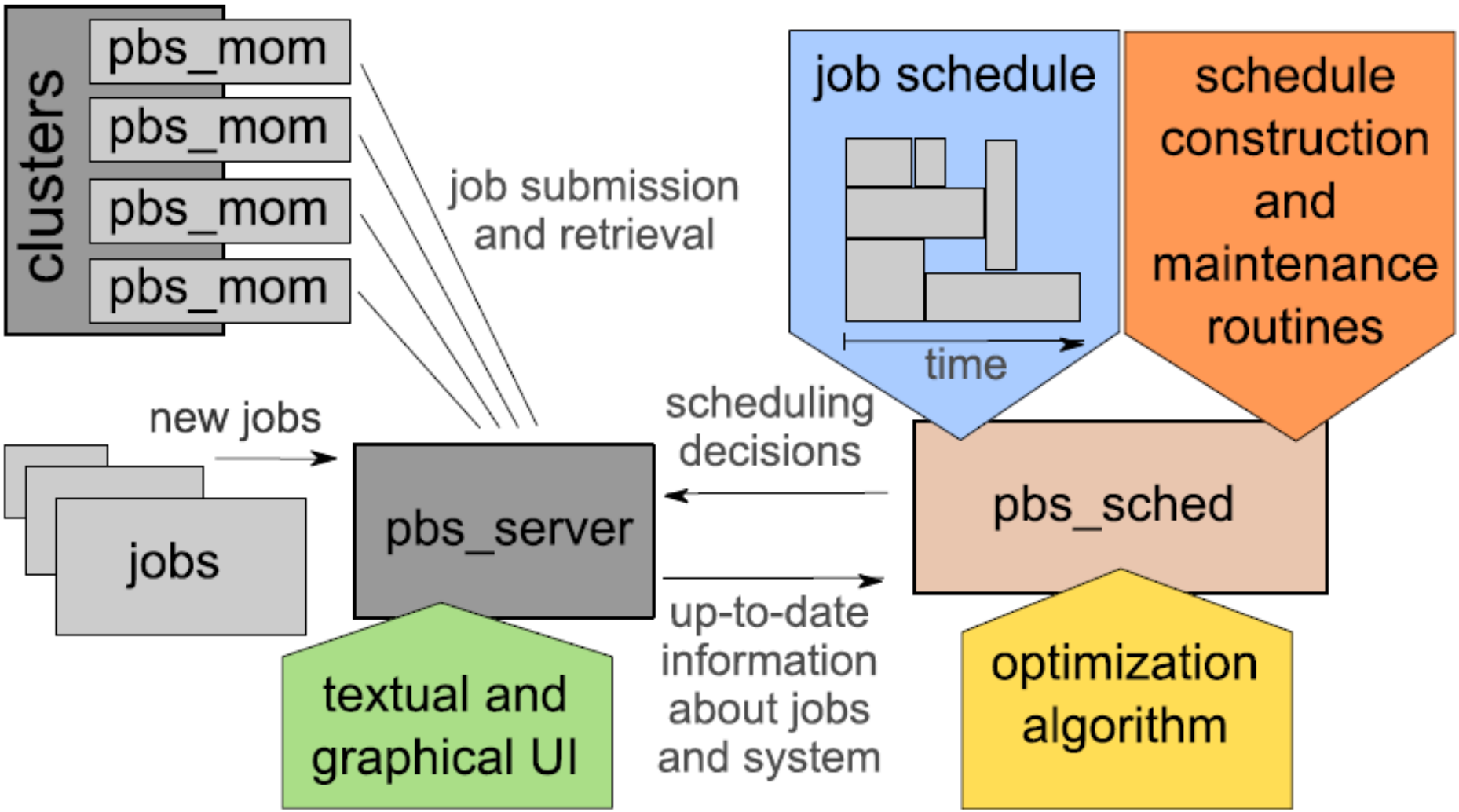
Simulation results

- 31.1% users with improved average wait time
- 13.9% users with deteriorated average wait time
- average improvement of wait time by 7.2 hours
- average deterioration of wait time by 2.1 hours

Increasingly frequent wide jobs

- parallel jobs are problematic in general
- freeing up resources can cause gaps in utilization
- problems with runtime estimates

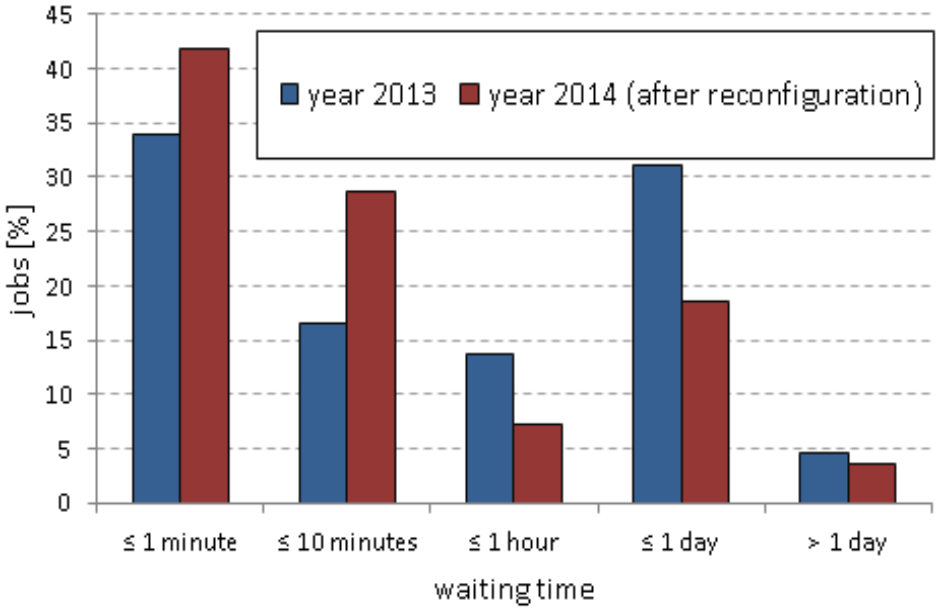
Planning & optimizing scheduler

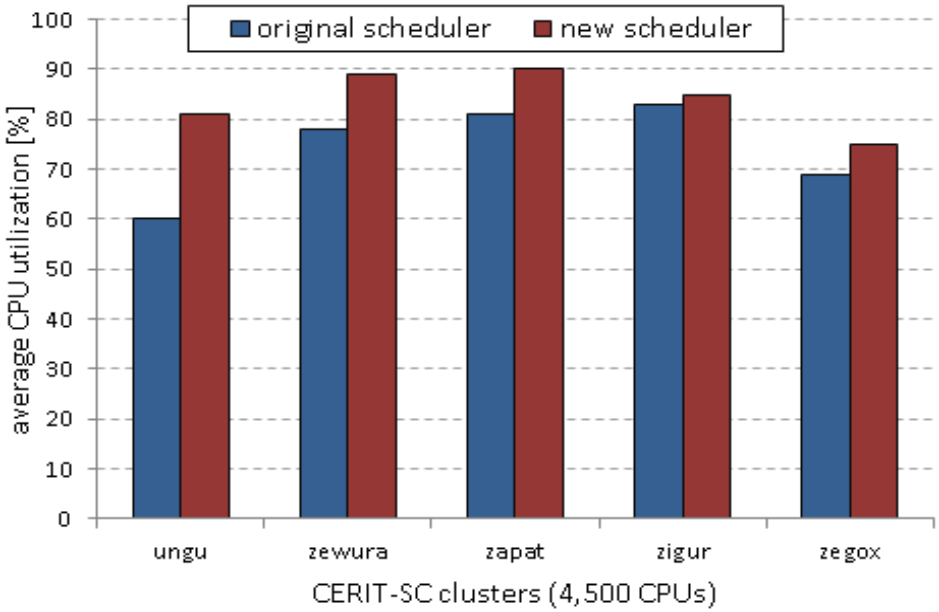


Planning & optimizing scheduler

- queue-less scheduler
- maintains a stable schedule
- schedule continuously optimized and adjusted
  - new jobs arrival
  - premature job completion
- currently managing 45% of grid resources

Results


45
40
35
30
25
20
15
10
5
0
jobs [%]
year 2013
year 2014 (after reconfiguration)
≤ 1 minute
≤ 10 minutes
≤ 1 hour
≤ 1 day
> 1 day
waiting time


100
90
80
70
60
50
40
30
20
10
0
average CPU utilization [%]
original scheduler
new scheduler
ungu
zewura
zapat
zigur
zegox
CERIT-SC clusters (4,500 CPUs)

Conclusion

- targeting fairness improved overall system performance
- user satisfaction improved as well
- required relaxation of some policies (e.g. anti-starvation)
- precise simulation tools are critical

Summary

- multi-resource fairness
- new queue configuration
- optimizing scheduler