Fault Detection and Diagnosis in the EGEE grid

C. Germain-Renaud, X. Zhang, M. Sebag – LRI
C. Loomis – LAL
CNRS & Paris-Sud University

Manchester 9-11 May 2007

DEMAIN
Des Donnés Massives
Aux Interprétations

www.eu-egee.org

LABORATOIRE DE L'ACCÉLÉRATEUR LINÉAIRE

Information Society and Media

Outline

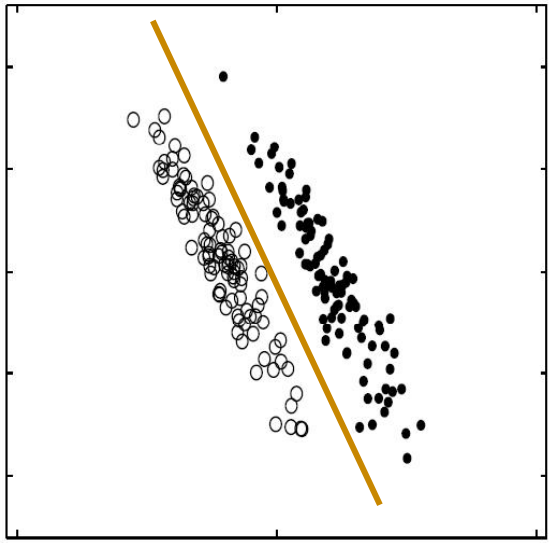
- **Mining the Logging and Bookkeeping data**
  - Dataset: L&B logs of the LAL site October 2004 - October 2005
- **Blackhole (and other failures) detection in Torque logs**
  - Dataset: Torque logs of the LAL site March 2006-February 2007

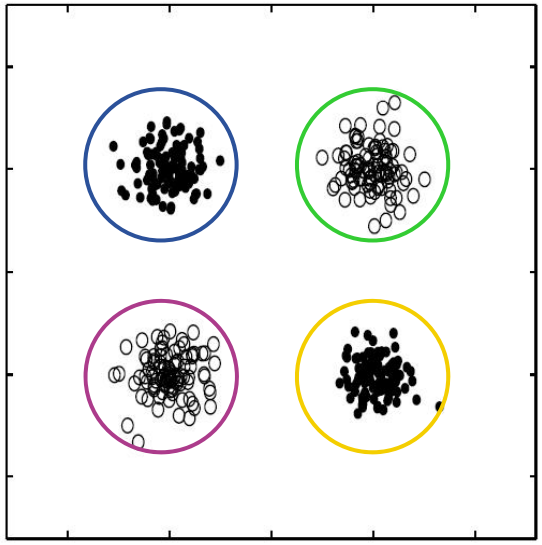
Pre-processing

- **Circa 300K jobs, 3Mevents, 2GB**
- **Operational log: a lot of information in blobs = incremental verbatim of the reports from the various services**
- **LB2F: A software suite for filtering and conversion towards a propositional vector**
  - Flatten compound attribute e.g. requirements
  - Tag with the job outcome
  - Prune attributes values
  - Normalize numerical atts. (dates)
  - Automatic identification of functional dependencies and trivial correlations
  - Anonymization
  - 408 attributes

Issues

- Simple classifiers fail
  - Feature construction
  - Integration of weak learners may produce good results
- No gold standard
- Probably not linear
  - Unsupervised clustering
- High variability following users and date
  - Aggressive subsampling





Constructive feature induction

- 36 users-consistent slices and 47 week-consistent slices
- Each slice has a lower variability, so something can be learned
- Here we use the linear learner ROGER: *ROC based genetic learner*
- Technically optimization of the Area Under ROC Curve, equivalent to Wilcoxon-Man-Whitney statistics
- Maps the boolean features onto the real-valued learned hypothesis

$$x = (x_1, x_2, \ldots, x_n) \rightarrow h(x) = w.x$$

Constructive feature induction (cont'd)

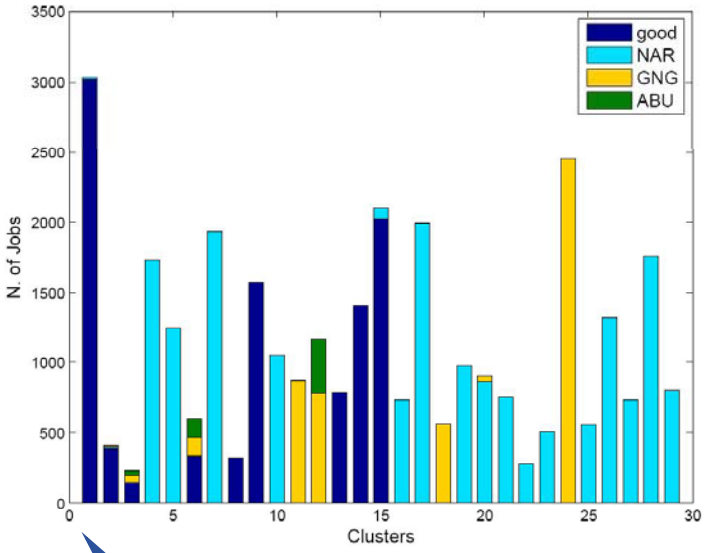
- 36 users consistent slices and 47 week consistent slices
- Maps the boolean features onto the real-valued learned hypothesis

$$x = (x_1, x_2, \ldots, x_n) \rightarrow h(x) = w.x$$

- Because the optmization is stochastic, multiple hypotheses must be kept: $l = 50$
- U-representation: $h_{i,u}(x)$ with i=1...l and u varying in the set of users: 1800 features
- W-representation: $h_{i,w}(x)$ with i=1...l and w varying in the set of weeks: 2350 features

Clustering

- **Meaningful features but need to eliminate useless redundancy and keep the useful ones**
- **Double clustering (Slonim & Tishby 2000):**
  - first clustering: “compress” the features along the examples
  - second clustering: cluster the examples along the synthetic features
- **K-means algorithm: discover a pre-defined number of clusters**
  - T feature clusters, K example clusters
  - Empirical optimization of K and T



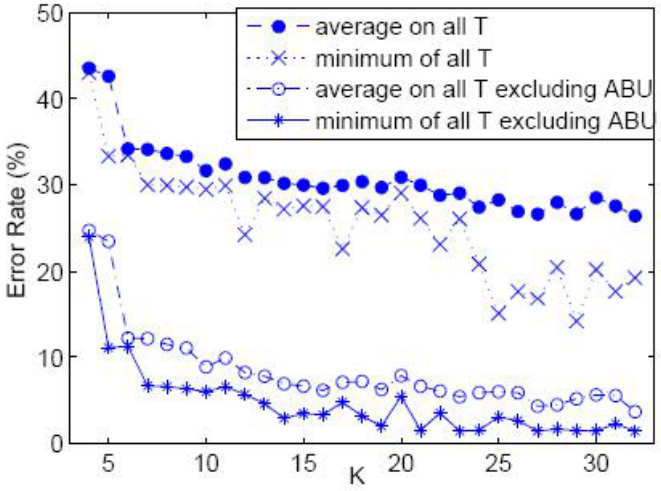
T=30

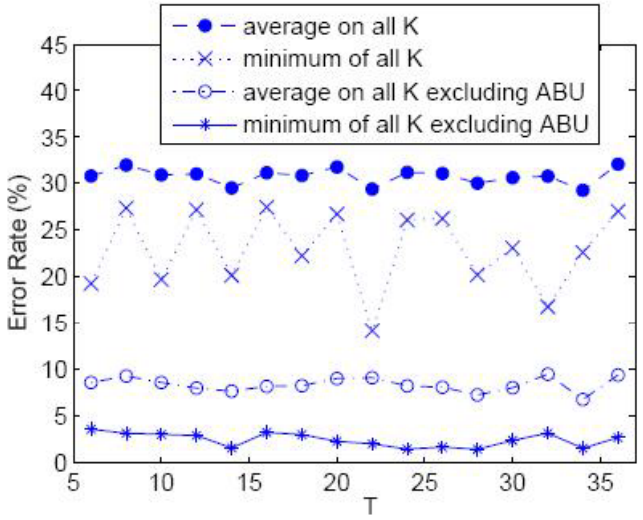
K=29

W-rep

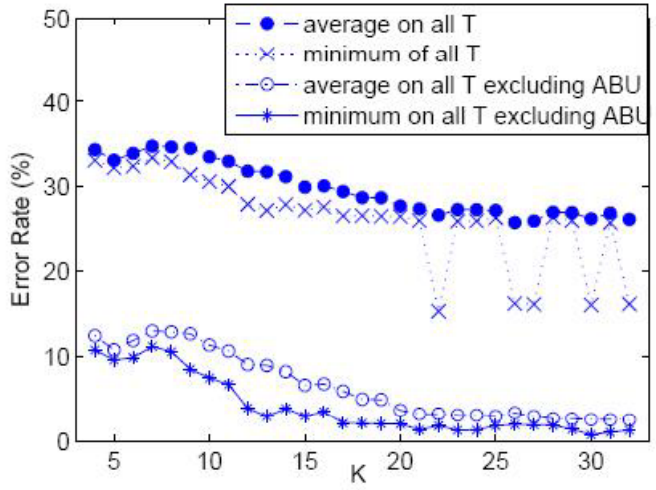
- **Mostly pure clusters**
- **Natural use for detection**
- **Diagnostic ?**

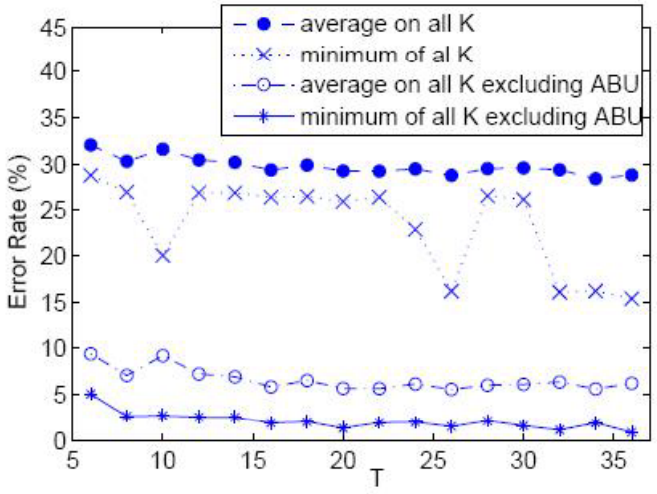
Classification





U-rep





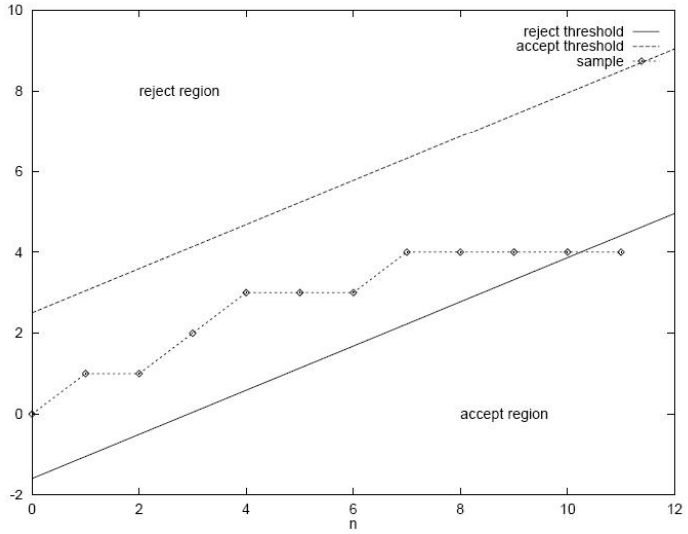
W-rep

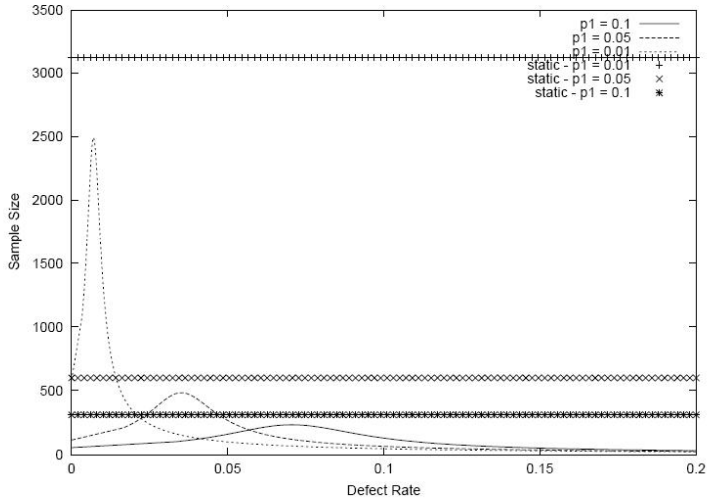
Blackhole detection

- What is a blackhole?
- A site fault which results in an ultra-fast (erroneous) execution
- Goal: on-line detection of blackholes – alarm
- Quantitative measurements
  - Anomalous job arrival rate and job service rate
  - And regular users and queues distributions?

Changepoint detection

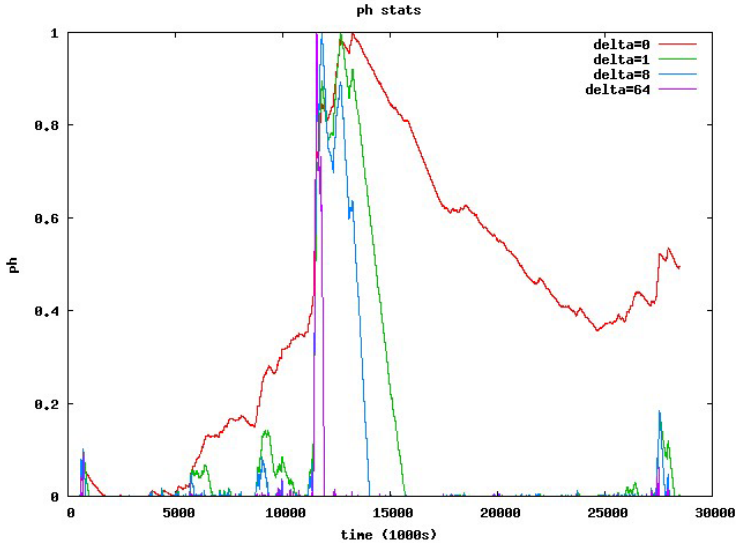
- Page-Hinckley statistics
- Time-sequential version of Wald's statistics – also known as CUSUM
- Provides an « intelligent threshold » test
- Minimizes the expected time before a change detection for a fixed false positive rate

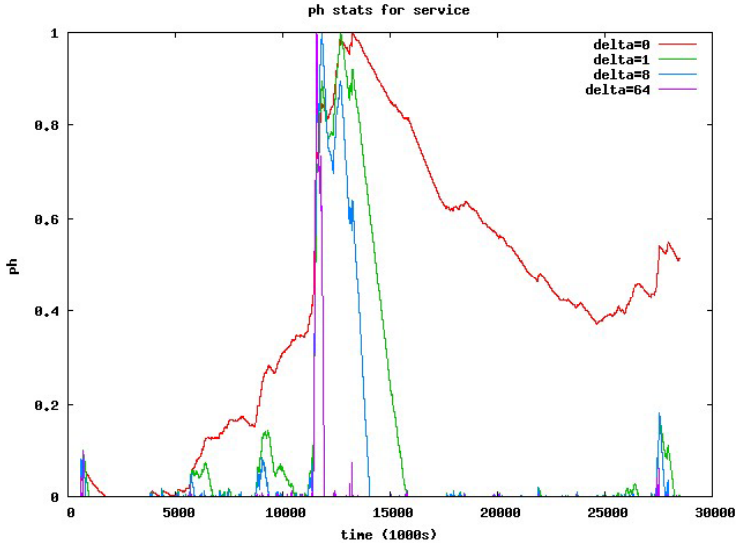




Changepoint detection

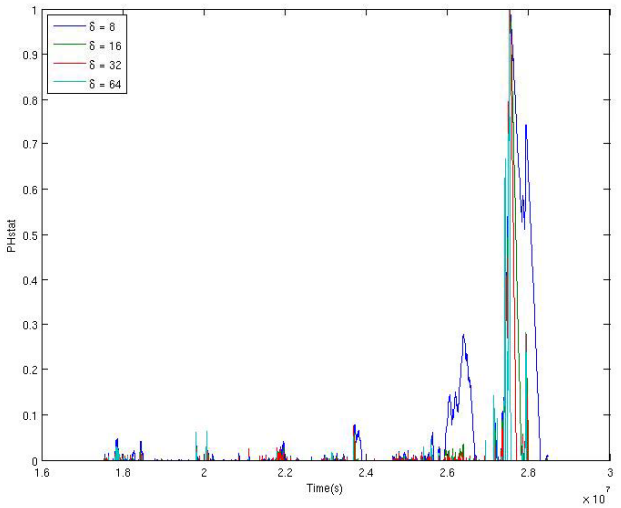
- Page-Hinckley statistics
- Time-sequential version of Wald's statistics – also known as CUSUM
- Provides an « intelligent threshold » test
- First event: VO software bug

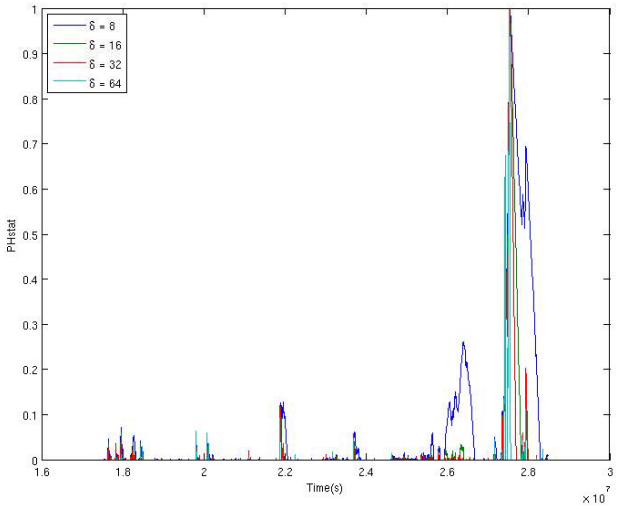




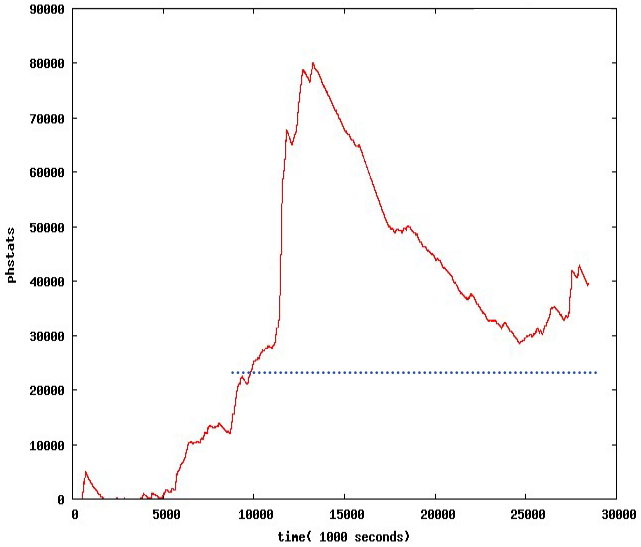
Changepoint detection with Page-Hinckley

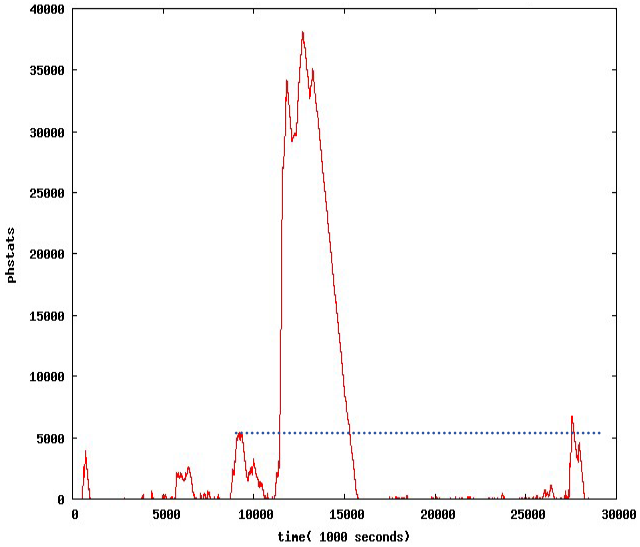
- Page-Hinckley statistics
- Time-sequential version of Wald's statistics – also known as CUSUM
- Provides an « intelligent threshold » test
- First event: VO software bug
- Second event: blackhole

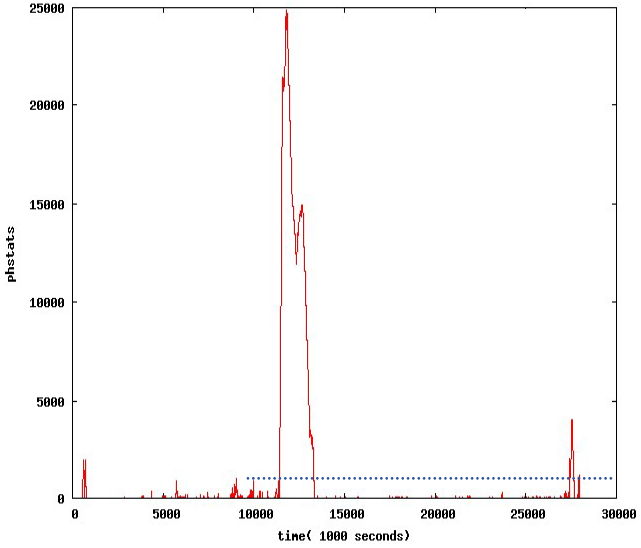


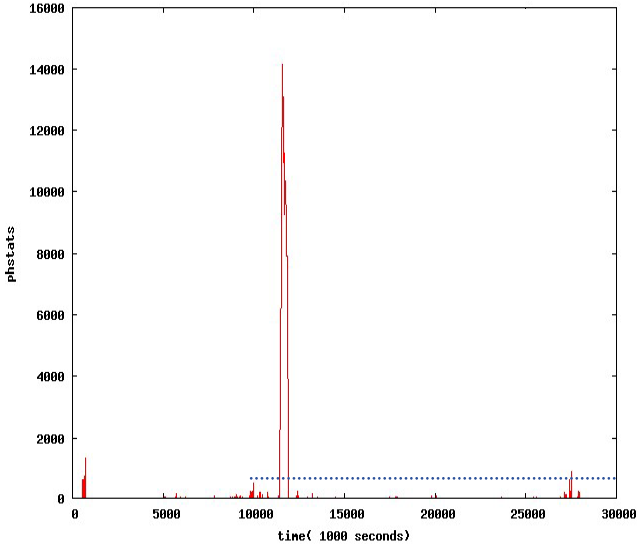


Details (unscaled)





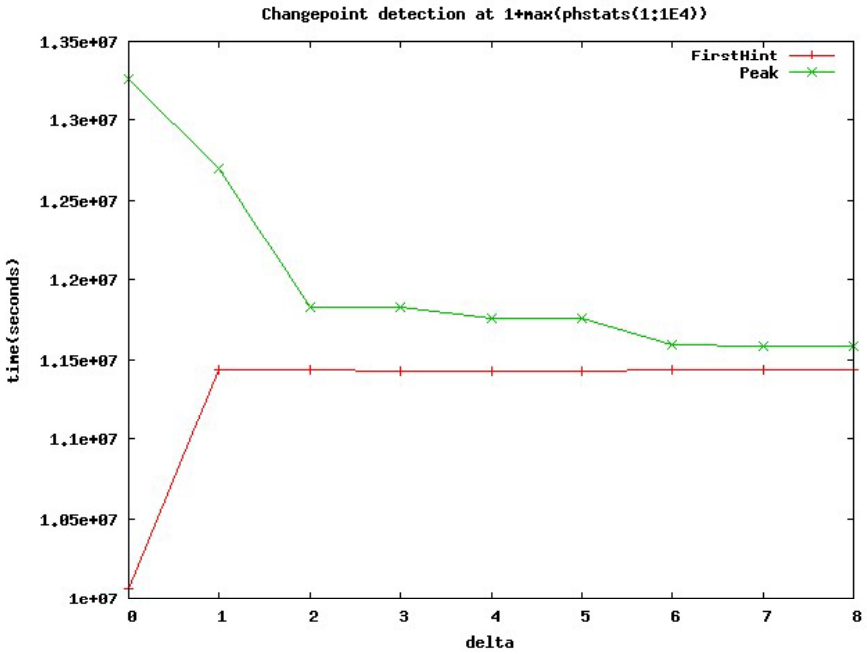




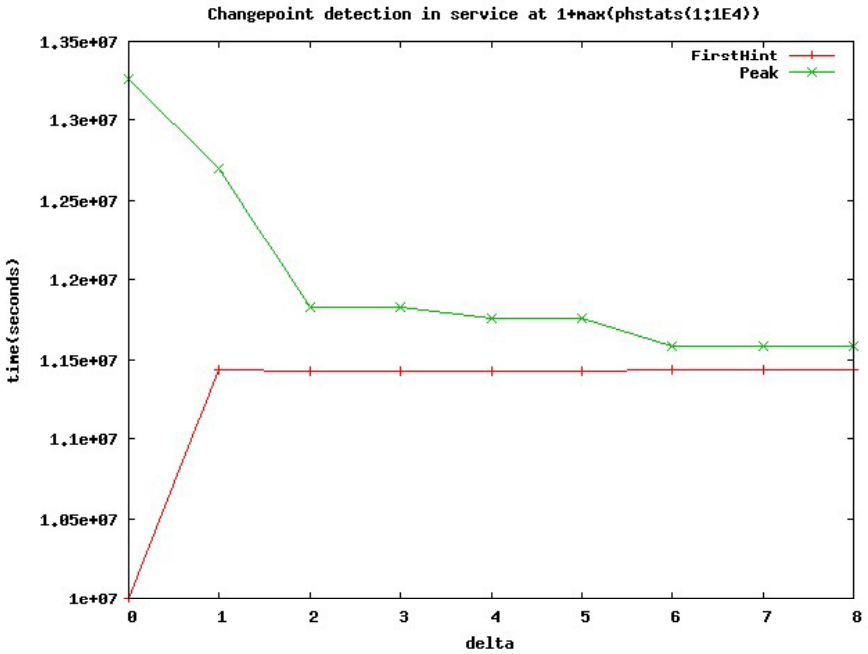
Robustness

Assume

- Everything OK until 10e4,
- Threshold =max(phstats) on this interval



- mean 1.1433E7
- std 28 minutes



- mean 1.1433E7
- std 33 minutes

Conclusion

- **Mining the Logging and Bookkeeping data**
  - Exemplifies the effectiveness and issues when applying advanced machine leaning workflows to grid data
- **Blackhole (and other failures) detection in Torque logs**
  - Classical statistical quality-control methods provide efficient on-line detection