www.bsc.es

Microsoft

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

ALOJA-ML: A Framework for Automating Characterization and Knowledge Discovery in Hadoop Deployments

Josep Ll. Berral, Nicolas Poggi, David Carrera,
Aaron Call, Rob Reinauer, Daron Green

ACM SIGKDD – KDD 2015, 10th – 13th August
Sydney, Australia

Brief Introduction

- Hadoop
  - Environment for parallel task execution (Map–Reduce)
  - Complex distributed runtime executions
- Initial Motivation
  - Hadoop optimization requires run and examine multiple executions
  - Modeling behaviors to estimate execution performance
  - … then observe Hadoop behavior without running lots of executions

The ALOJA Project

- ALOJA framework:

"Which Hadoop & data-center configuration is the best in cost-effectiveness terms?"

  - Provide **expert-guided to automated** analysis
    - User schedule Hadoop executions
    - Framework runs and collect monitor information
    - Information is shown to the user for comparing configs. & deployments

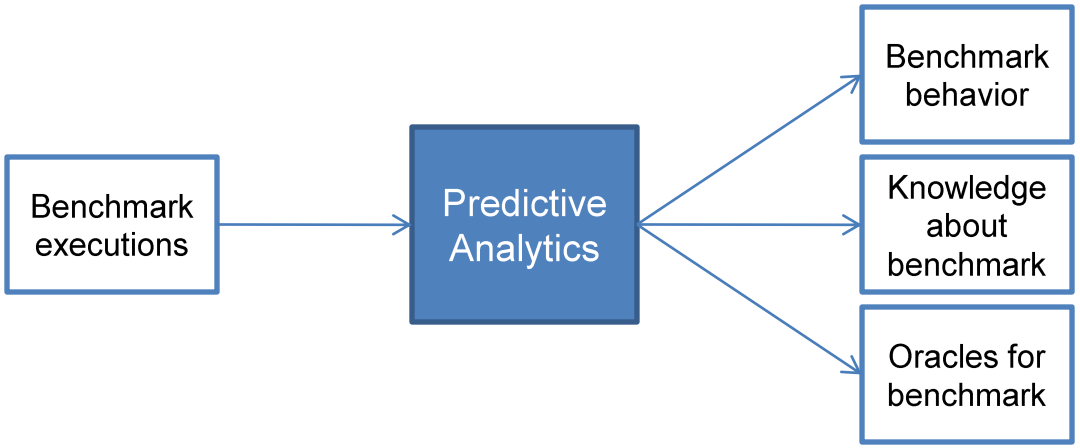
- The project:
  - Joint initiative Barcelona Supercomputing Center – Microsoft Research
  - Towards comparing Big-Data deployment providers and configurations
  - Seeking to provide knowledge and tools to the community

Benchmarks and Predictive Analytics

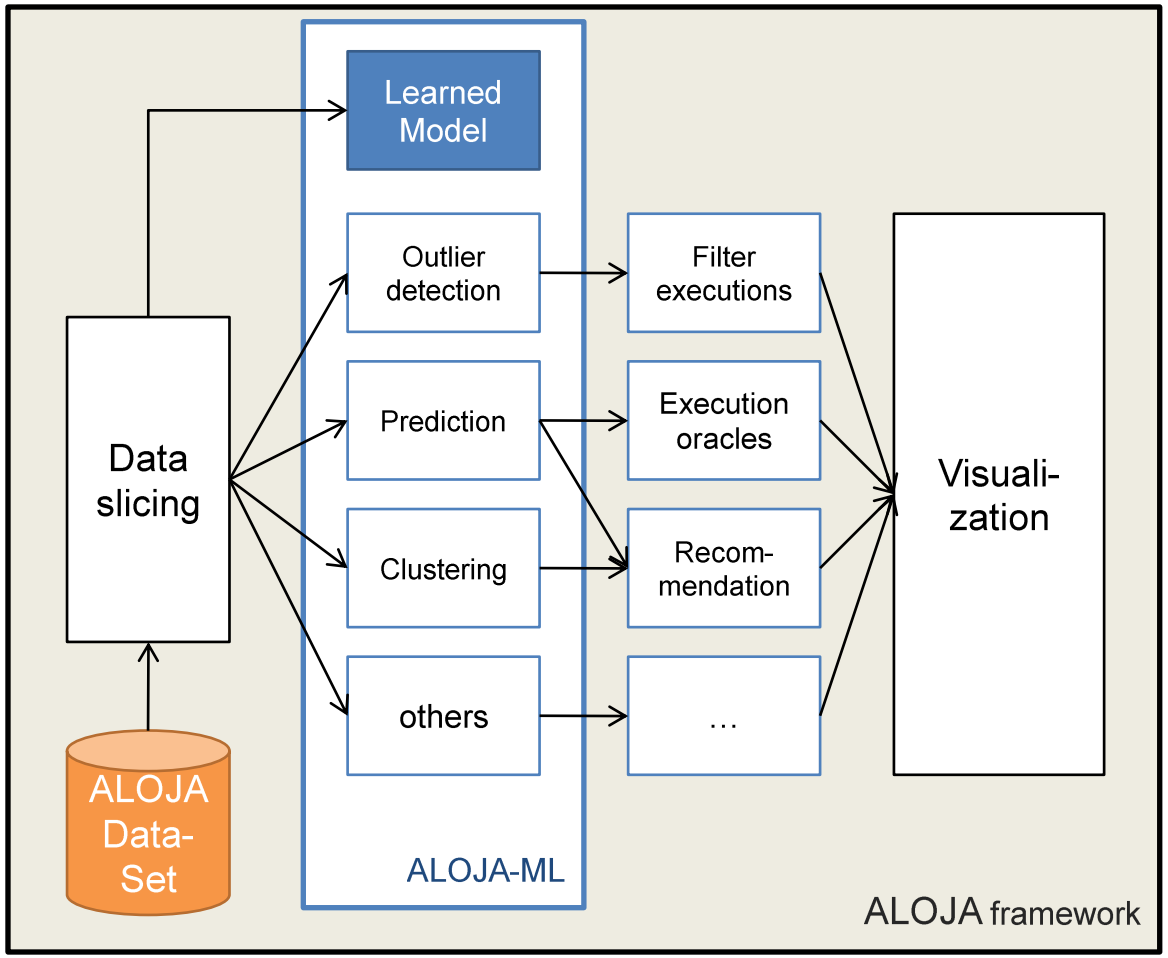
- Predictive Analytics
  - Deploy **model-based methods** enhancing analysis
  - Predict behavior elements and apply them to extract knowledge

  - Used also for recommendation (of configurations/set-ups)
  - Or even anomaly detection mechanisms

- ALOJA-ML: the ALOJA predictive analytics component for modeling benchmarks



Modeling Hadoop – P.A. Layer



Modeling and prediction as a service layer in the framework

Modeling Hadoop – Data-sets

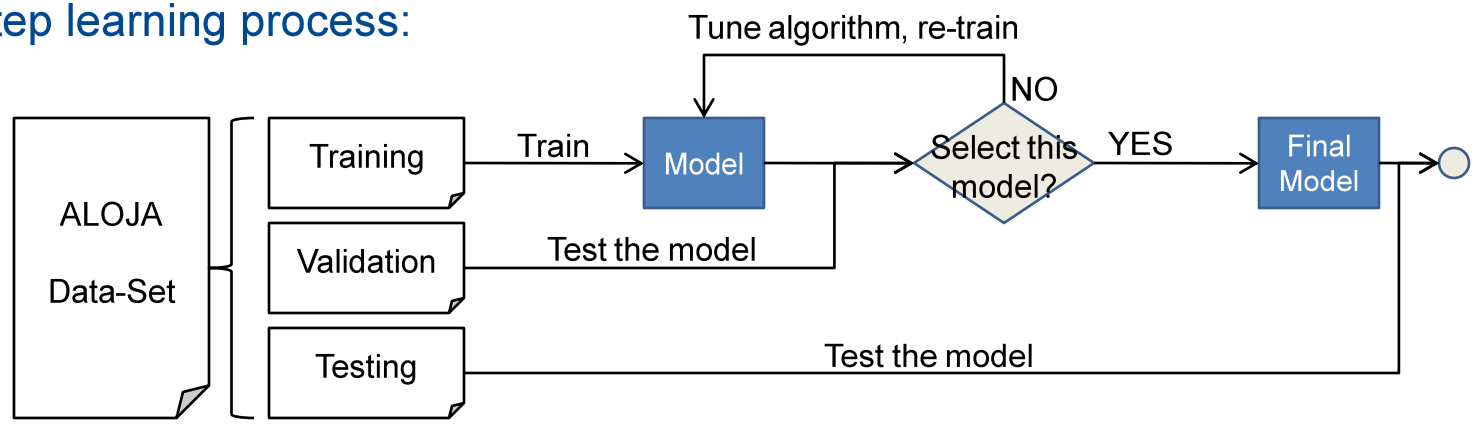
- The ALOJA data-set
  - Over +40.000 Hadoop benchmark executions
  - Input features: Benchmark info, Configuration info, Deployment info, …
  - Output features: Service Level Objects, Used Resources, …
- Hadoop Executions:
  - …from different Hadoop versions
  - …from different underlying infrastructure
  - …with different input sizes (100GB–1TB)
  - …with some anomalous executions
  - …

Benchmarks	
bayes, terasort, sort, wordcount, kmeans, pagerank, dfsioe_read, dfsioe_write	
Hardware Configurations	
Network	Ethernet, Infiniband
Storage	SSD, HDD, Remote Disks {1-3}
Cluster	# Data nodes, VM description
Software Configurations	
Maps	2 to 32
I/O Sort Factor	1 to 100
I/O File Buffer	1KB to 256KB
Replicas	1 to 3
Block Size	32MB to 256MB
Compression Algs.	None, BZIP2, ZLIB, Snappy
Hadoop Info	Version

Configuration parameters on data-set

Modeling Hadoop – Methodology

- Methodology
  - 3-step learning process:



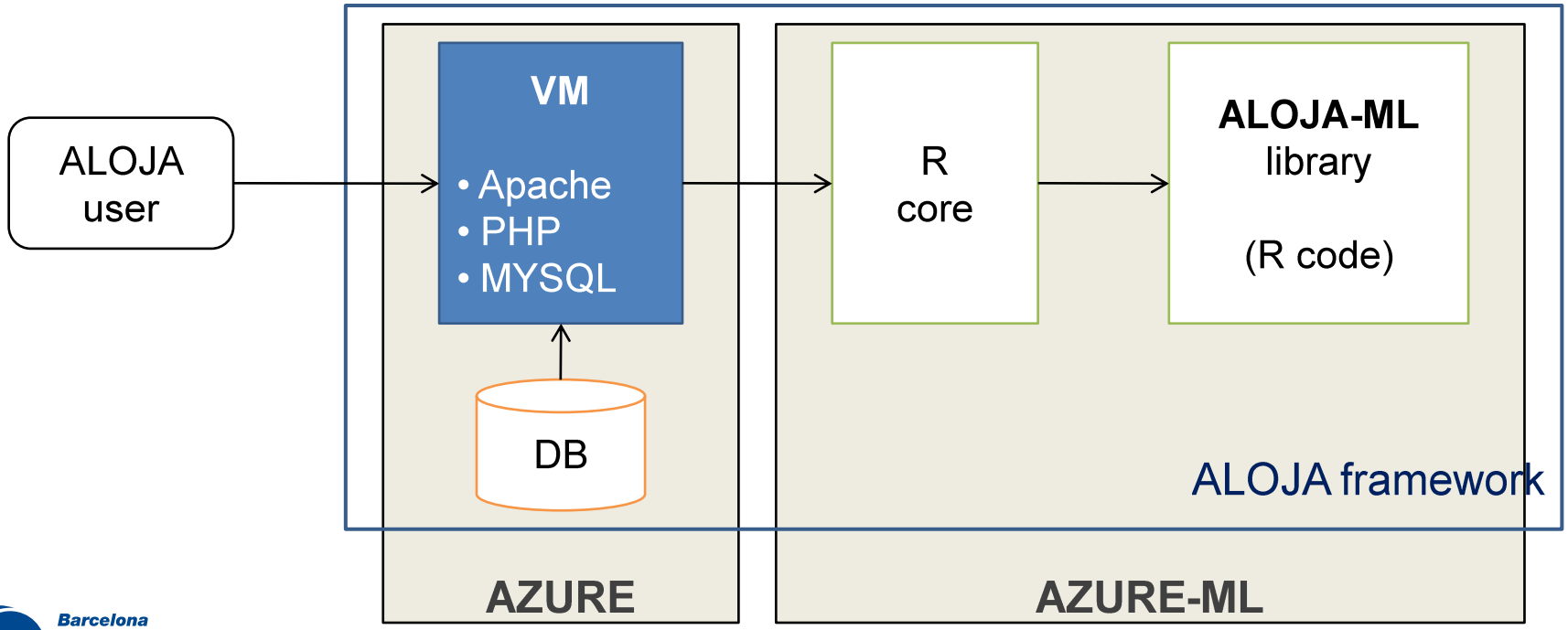
  - Different split sizes tested: (10% ≤ training ≤ 50%)
  - Different learning algorithms: Regression trees; Nearest-neighbors methods; Linear/Multinomial regressions; Neural networks

- Learning results
  - Mean Absolute Errors ~250s (ranges in [100s, 6000s])
  - Relative Absolute Errors between [0.10, 0.25]
    - Depend on benchmark and # of examples per benchmark
    - Some executions are/may be anomalies

Implementation and Technology

- Software infrastructure
  - Environment (ALOJA): Hadoop + LAMP + Vagrant
  - Data mining tools (ALOJA-ML): R-cran + Java (for RWeka)

  - Methods can run locally or at Microsoft **AZURE** and **AZURE-ML**

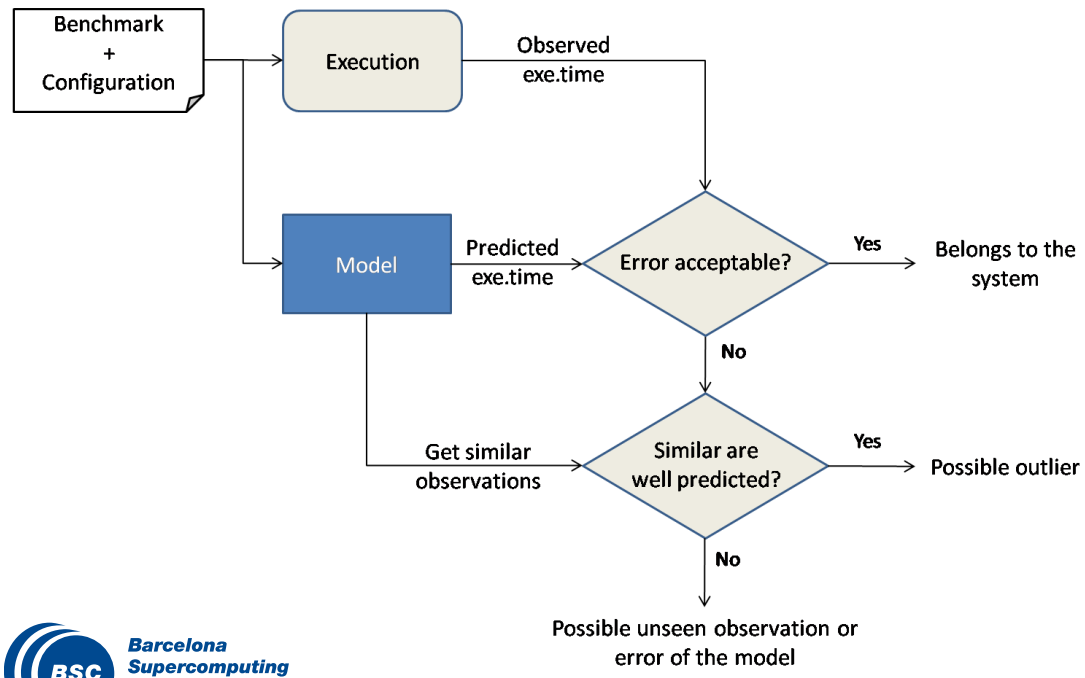


Case of use 1: Anomaly Detection

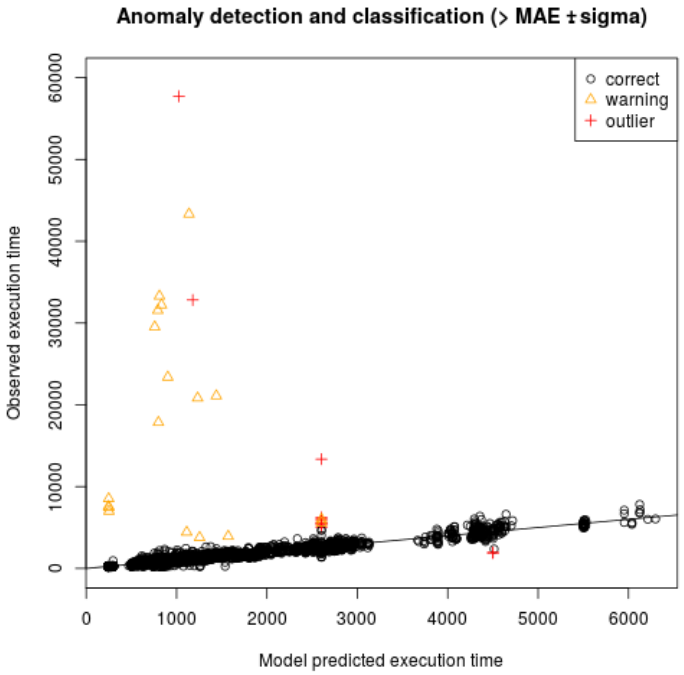
Anomaly Detection

- Model-based detection procedure
- Pass executions through the model
- Executions not fitting the model are considered “out of the system”

Anomaly detection procedure:

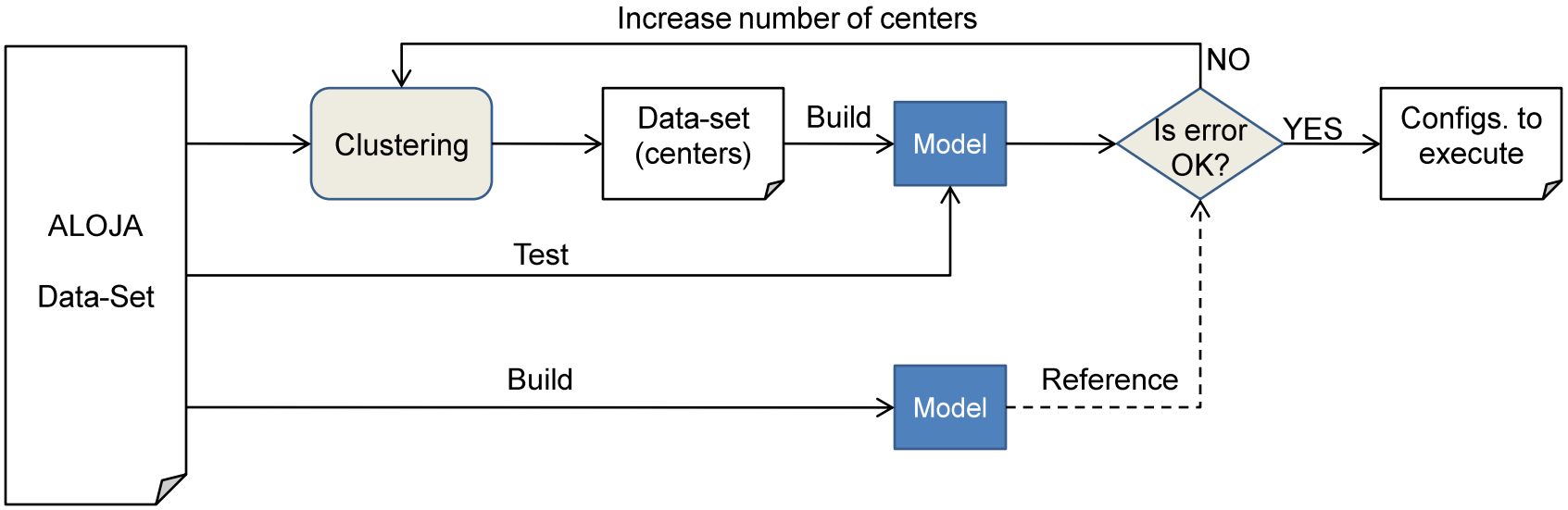


Testing ALOJA Data-set:



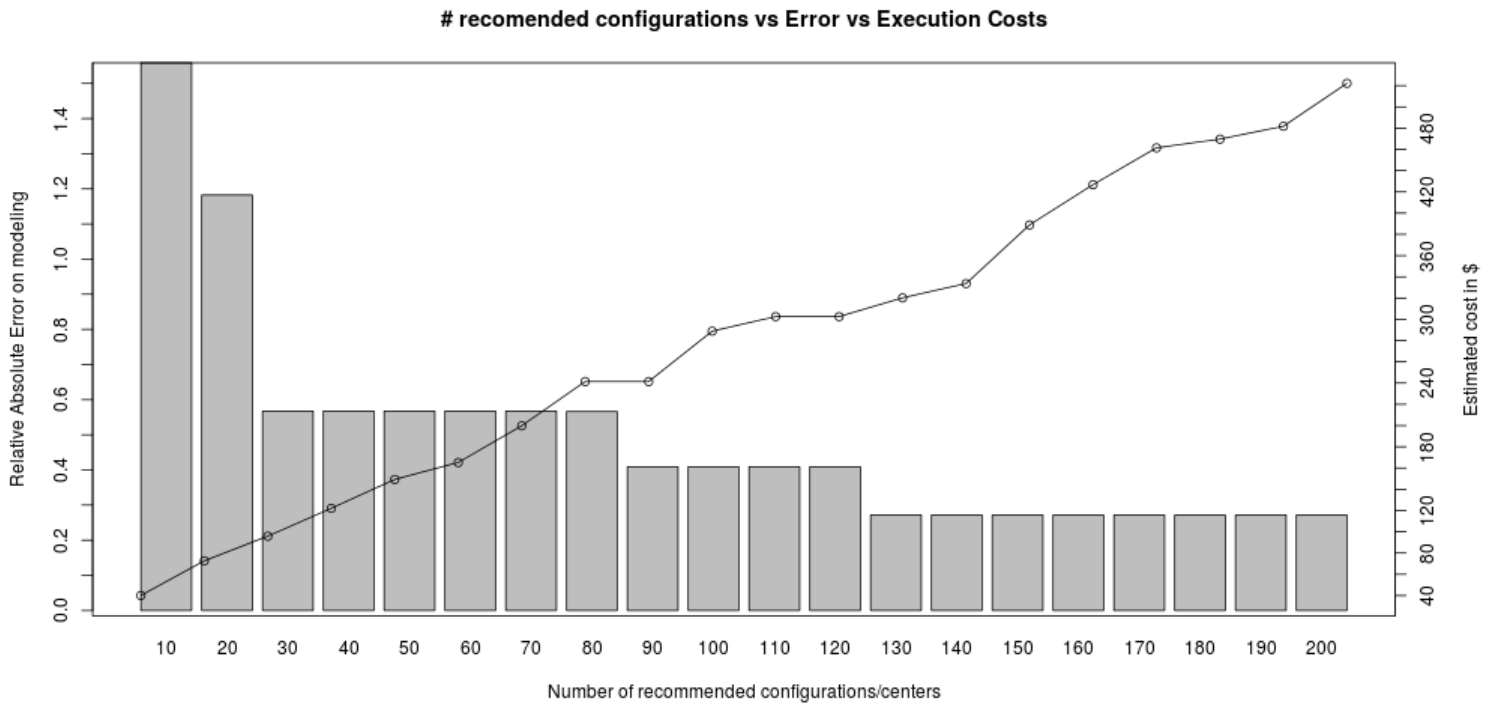
Case of use 2: Guided Benchmarking – Method

- Guided Benchmarking:
  - Best subset of configurations for modeling a Hadoop deployment
  - Clustering to get the “representative execution” for each similar subset of executions



Case of use 2: Guided Benchmarking – Trade-off

- Trade-off "# executions" vs. "deployment representation"
  - Clustering methods
    - # executions ~ running time/costs ~ fidelity to original benchmarking
    - k-means: iterate over "k" → trade-off "executions-accuracy"



Other cases

- Tools for treating data, observed and predicted
- Find best expected configurations
  - Use models to predict search sub-spaces and order by execution time
- Feature exploration:
  - Use models to unfold search sub-spaces and retrieve feature rankings
  - Then display that in a user-readable way

Net	Disk	IO.FBuf	Blk.Size	Prediction (s)
ETH	HDD	65536	128	2249.766
IB	HDD	65536	128	2737.112
ETH	SSD	65536	128	1036.366
IB	SSD	65536	128	1036.366
ETH	HDD	131072	128	2165.927
IB	HDD	131072	128	2653.273
ETH	SSD	131072	128	969.537
IB	SSD	131072	128	969.537
ETH	HDD	65536	256	2249.766
IB	HDD	65536	256	2737.112
ETH	SSD	65536	256	1036.366
IB	SSD	65536	256	1036.366
ETH	HDD	131072	256	2165.927
IB	HDD	131072	256	2653.273
ETH	SSD	131072	256	969.537
IB	SSD	131072	256	969.537
Terasort, 4 maps, sort factor 10, no comp				

```
Disk=SSD
    IO.FBuf=131072 -> 970s
    IO.FBuf=65536 -> 1036s
Disk=HDD
    Net=ETH
        IO.FBuf=131072 -> 2166s
        IO.FBuf=65536 -> 2250s
    Net=IB
        IO.FBuf=131072
            Blk.size=128 -> 2653s
            Blk.size=256 -> 2653s
        IO.FBuf=65536
            Blk.size=128 -> 2737s
            Blk.size=256 -> 2737s
```

Conclusions

- Predictive analytics enhancing benchmark cost-effectiveness visualization

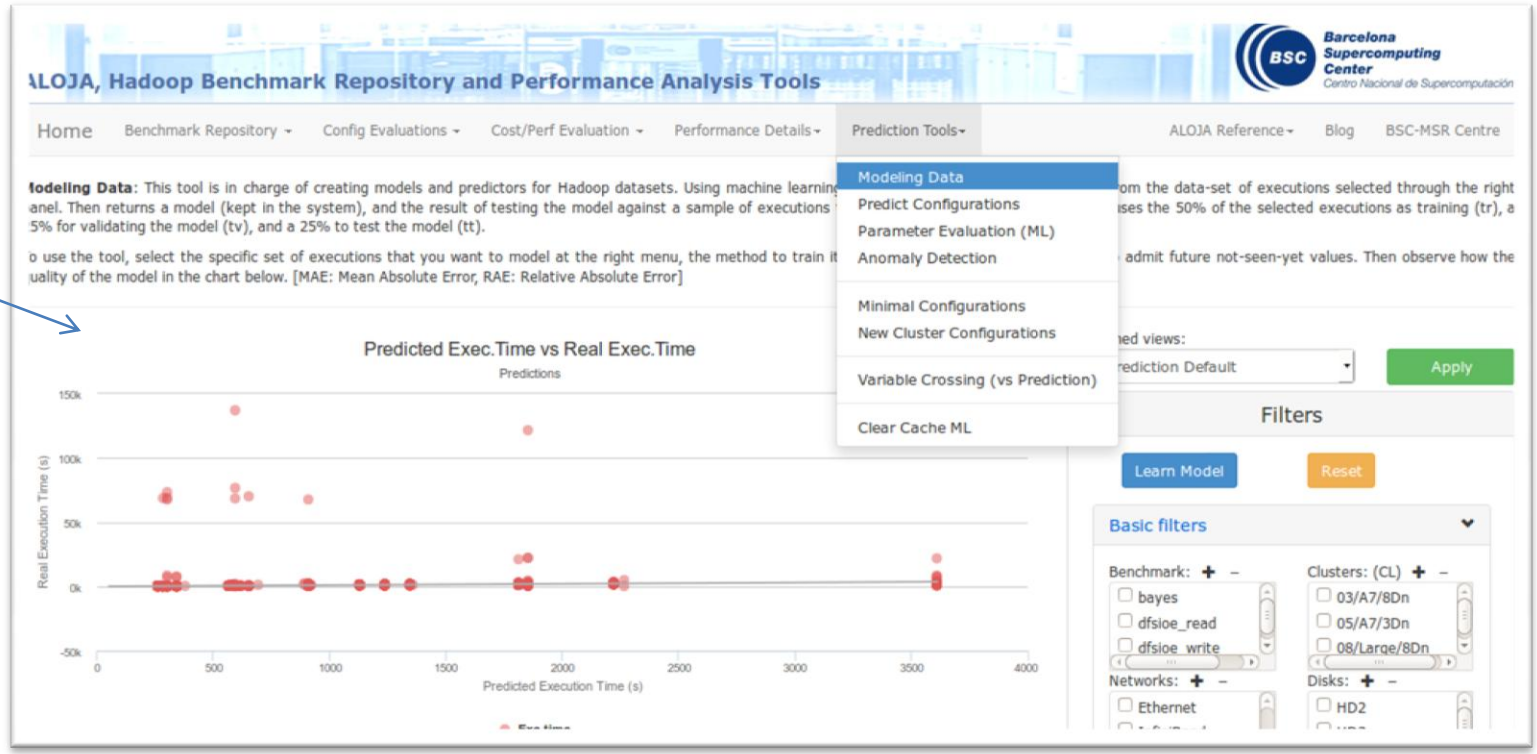
- Modeling behaviors → Realize predictions → Use in tools
  - Execution prediction [Planning, tuning configurations, ...]
  - Anomaly detection [Model-based + data-driven]
  - Recommendations [Parameters, HW deployments, data-center/providers]

- Next steps:
  - New tools into the ALOJA-ML framework (knowledge extraction, pattern mining, …)
  - Expand and detail features (benchmark information, HW properties, …)
  - New examples and addition of features into the Data-Set

Availability of the framework and the data-set

- Framework **on-line demo** available at http://aloja.bsc.es
  - Also downloadable to deploy locally
- Data-sets available with the framework
  - Also downloadable from the on-line demo

Check it out!



Microsoft

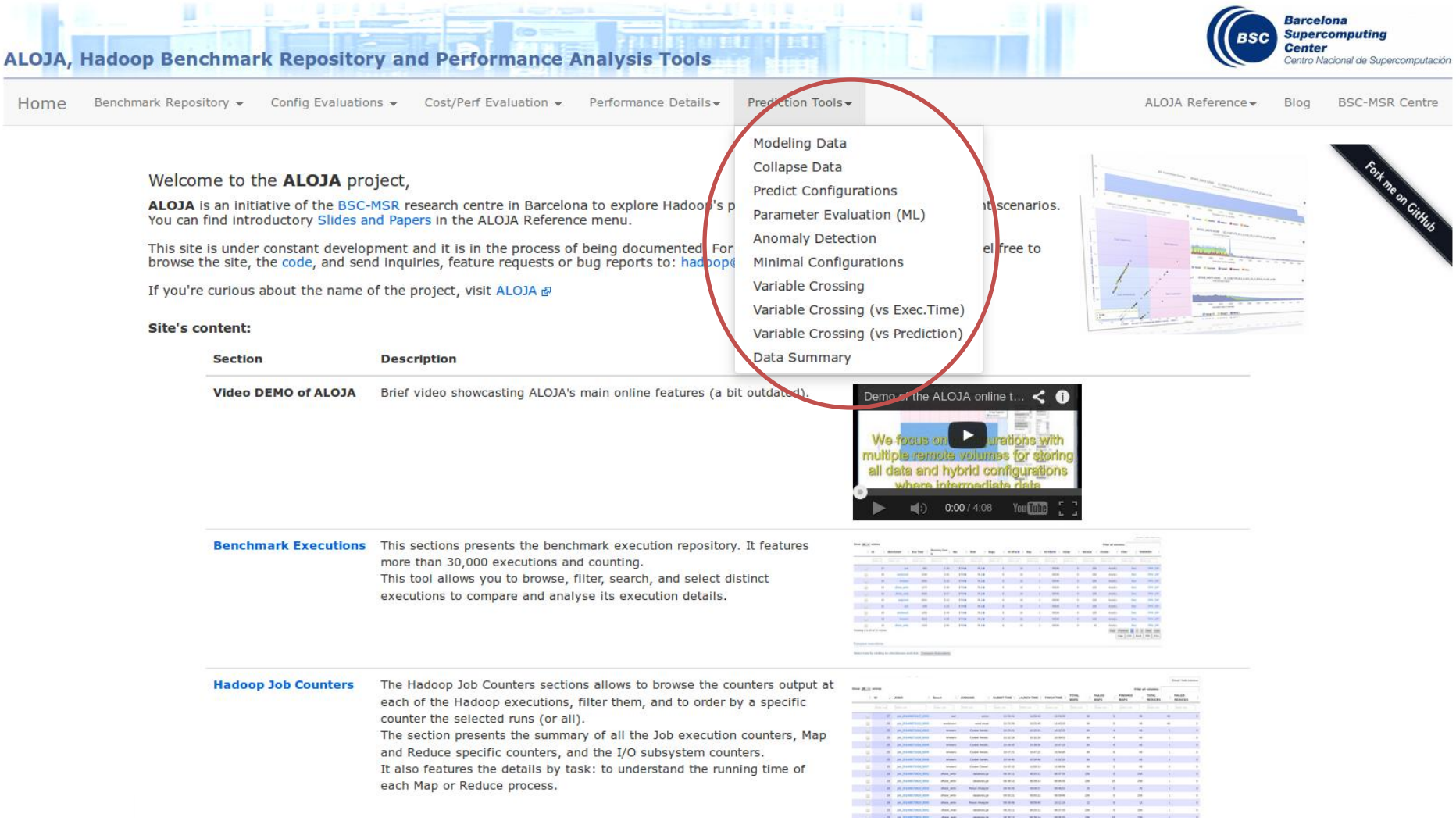
Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Thanks for your attention

Questions?

(I'll be at the Poster session, for more information and a live demo)

[Screen 1: Front page and ML tools]



[Screen 2: Modeling benchmark data-sets]

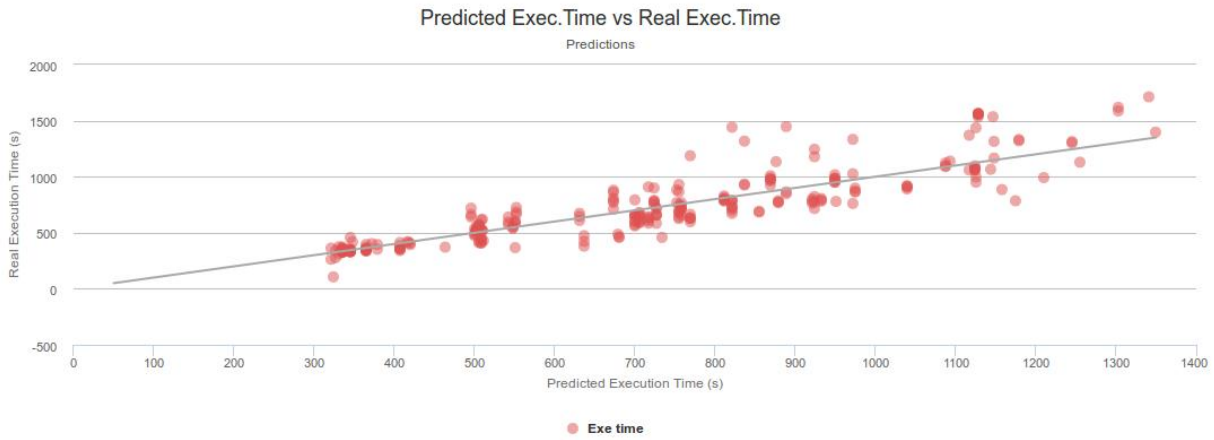
ALOJA, Hadoop Benchmark Repository and Performance Analysis Tools

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Home | Benchmark Repository | Config Evaluations | Cost/Perf Evaluation | Performance Details | Prediction Tools | ALOJA Reference | Blog | BSC-MSR Centre

Modeling Data: This tool is in charge of creating models and predictors for Hadoop datasets. Using machine learning algorithms, this tool creates a model from the data-set of executions selected through the right panel. Then returns a model (kept in the system), and the result of testing the model against a sample of executions to check its accuracy. Our methodology uses the 50% of the selected executions as training (tr), a 25% for validating the model (tv), and a 25% to test the model (tt).

To use the tool, select the specific set of executions that you want to model at the right menu, the method to train it, and as option if you want the model to admit future not-seen-yet values. Then observe how the quality of the model in the chart below. [MAE: Mean Absolute Error, RAE: Relative Absolute Error]



Dataset: tr => MAE: 101.4091250 RAE: 0.14609210568
Dataset: tv => MAE: 104.6996406 RAE: 0.13311951094
Dataset: tt => MAE: 85.4222769 RAE: 0.11794156231

Model information:

- Model cache ID: 32c3503f3b5b94abc83e34c89d38d3c9
- Instance: terasort,*,*,*,*,1,32768|65536|131072,Cmp0,*,*,*,*,*,*,*,*
- Model Info: bench ("terasort") net ("*") disk ("*") maps ("*") iosf ("*") replication ("1") iofilebuf ("32768","65536","131072") comp ("0") blk_size ("*") id_cluster ("*") datanodes ("*") bench_type ("*") vm_size ("*") vm_cores ("*") vm_RAM ("*") type ("*")

Predefined views:

MLPrediction Default | Apply

Filters

Learn Model | Reset

Basic filters

Hardware

Cluster datanodes: + −
- 8

Benchmark type: + −
- HiBench
- HiBench-1TB
- HiBench-min

VM size: + −
- general1-8
- io1-15
- io1-30
- performance1-8

VM #cores: + −
- 4
- 8

VM RAM: + −
- 8 GB
- 15 GB
- 30 GB

Cluster type: + −
- IaaS

[Screen 3: Predicting configurations]

Predicted Configurations

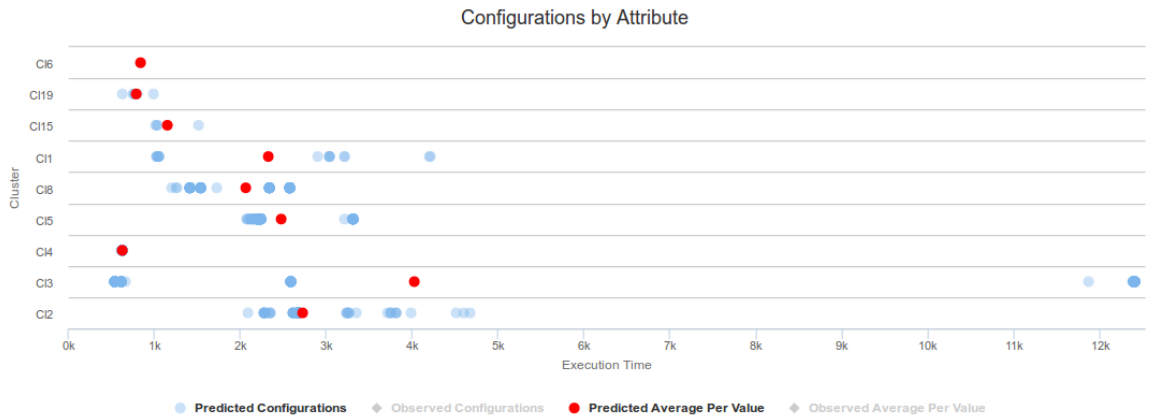
Show 10 entries Search:

	Benchmark	Net	Disk	Maps	IO.SFS	Rep	IO.FBuf	Comp	Blk.Size	Cluster	Bench.Type	Prediction
1	terasort	ETH	HDD	4	10	1	131072	Cmp0	128	CI15	HiBench	1019.1590000
2	terasort	ETH	HDD	4	10	1	65536	Cmp0	128	CI4	HiBench	618.2790000
3	terasort	ETH	HDD	4	10	1	65536	Cmp0	128	CI5	HiBench	2136.4190000
4	terasort	ETH	RR3	4	10	1	65536	Cmp0	128	CI6	HiBench	843.0730000
5	terasort	ETH	RR3	4	10	1	65536	Cmp0	128	CI19	HiBench	762.6890000
6	terasort	ETH	RR2	4	10	1	65536	Cmp0	128	CI19	HiBench	753.3910000
7	terasort	ETH	RR2	4	10	1	131072	Cmp0	128	CI19	HiBench	771.5810000
8	terasort	ETH	RR1	4	10	1	131072	Cmp0	128	CI19	HiBench	994.6670000
9	terasort	ETH	RL3	4	10	1	131072	Cmp0	128	CI8	HiBench	1203.0580000
10	terasort	ETH	HDD	4	10	1	65536	Cmp0	128	CI15	HiBench	1520.0000000

Showing 1 to 10 of 1,027 entries Previous 1 2 3 4 5 … 103 Next

Plot Executions by Class

Feature selector: Clusters Display Predicted Display Observed



Predefined views: MLFindAttrs Default Apply

Filters

Generate Predictions Reset

Basic filters

Benchmarks: terasort

Clusters: (CL) + −
- 02/A7/3Dn
- 03/A7/8Dn
- 04/A7/8Dn

Networks: + −
- Ethernet
- InfiniBand

Disks: + −
- HD2
- HD3
- HD4
- HD5

Hardware

Hadoop configurations

Machine learning

Model to use: 41ed1f3ff1dd72dac

Predict with unseen atributes ⚠

Advanced filters

ALOJA Related Publications

- ALOJA project: automatic characterization of cost-effectiveness on Hadoop deployments

  - Nicolas Poggi, David Carrera, Aaron Call, Rob Reinauer, Nikola Vujic, Daron Green and Jose Blakeley, et al. "ALOJA: a Systematic Study of Hadoop Deployment Variables to Enable Automated Characterization of Cost-Effectiveness". IEEE BigData 2014

- ALOJA-ML: Predictive analytics tools for benchmarking on Hadoop deployments
  - Prediction of benchmarking behavior, anomaly detection, ranking features…

    - Josep Ll. Berral, Nicolás Poggi, David Carrera, Aaron Call, Rob Reinauer, Daron Green. "ALOJA-ML: A Framework for Automating Characterization and Knowledge Discovery in Hadoop Deployments". ACM SIGKDD - KDD 2015

  - Study of modeling of benchmarks, specifics vs general models
    - Josep Ll. Berral, Nicolas Poggi, David Carrera. "A Case of Study on Hadoop Benchmark Behavior Modeling Using ALOJA-ML". Technical session on WBDB'15.