



ALOJA: Cost-effective Big Data deployments

Nicolas Poggi, Senior Researcher



EXCELENCIA
SEVERO
OCHOA



Microsoft Research
Centre

February 2015



INSTITUTIONAL ABOUT THE PROJECT

Barcelona Supercomputing Center (BSC)

« 22 year history in Computer Architecture research

- European Center for Parallelism of Barcelona (CEPBA)
 - Based at the Technical University of Catalonia (UPC) in 1991
- ~~Received funding with IBM in within the Autonomic & Parallelism~~

« Led by Mateo Valero

- ACM fellow, Eckert-Mauchly award 2007, Goode award 2009
- Active research staff with 1000+ publications
- Large ongoing life science computational projects
 - Computational Genomics
 - Molecular modeling & Bioinformatics
 - Protein Interactions & Docking
- In place computational capabilities
 - Mare Nostrum Super Computer



MareNostrum Supercomputer

« Prominent body of research activity around Hadoop since 2008

- Previous to ALOJA
 - SLA-driven scheduling (Adaptive Scheduler), in memory caching, etc.
- Research group page: <http://www.bsc.es/autonomic>



BSC-MSRS Centre and ALOJA



Microsoft Research
Centre

- « Long-term relationship between
 - BSC and Microsoft Research and Microsoft product teams
- « Previous research at the intersection of computer architecture, language implementation, and systems software, and performance profiling
- « Open model:
 - **No patents, public IP, publications and open source main focus**
 - **87 publications, 4 Best paper awards**
- « **ALOJA** is the latest phase of the engagement
- « With intent to explore:
 - upcoming hardware architectures
 - and building automated mechanism
- « for deploying cost-effective Hadoop clusters.



BSC-MSRS Centre



Motivation

- « The Hadoop framework implements a complex distributed execution model
 - Over 100 interrelated config parameters
 - Requires manual iterative benchmarking and tuning
 - « Early results show that Hadoop's price/performance
 - are affected by relatively simple SW >3x
 - and HW configuration choices > 3x
 - « Commodity HW no longer low-end
 - new affordable hardware from original design (ie., SSDs)
 - Hadoop performs poorly on scale-up
 - or low power HW
 - « New Cloud services for Hadoop
 - IaaS and PaaS
 - Direct vs. remote attached volumes
 - « Spread Hadoop ecosystem
 - Dominated by vendors
 - Lack of verifiable benchmarks

Hadoop – Map / Reduce – Overview of I/O

Original InputFile

File split 0 → Map 0 → Spill files → imc_0_0, imc_0_1, imc_0_2 → HTTP → irc_0_0, irc_1_0, irc_2_0, irc_3_0, irc_4_0, irc_5_0, irc_6_0, irc_7_0, irc_8_0, ...

File split 1 → Map 1 → Spill files → imc_1_0, imc_1_1, imc_1_2 → HTTP → irc_0_1, irc_1_1, irc_2_1, irc_3_1, irc_4_1, irc_5_1, irc_6_1, irc_7_1, irc_8_1, ...

File split 2 → Map 2 → Spill files → imc_2_0, imc_2_1, imc_2_2 → HTTP → irc_0_2, irc_1_2, irc_2_2, irc_3_2, irc_4_2, irc_5_2, irc_6_2, irc_7_2, irc_8_2, ...

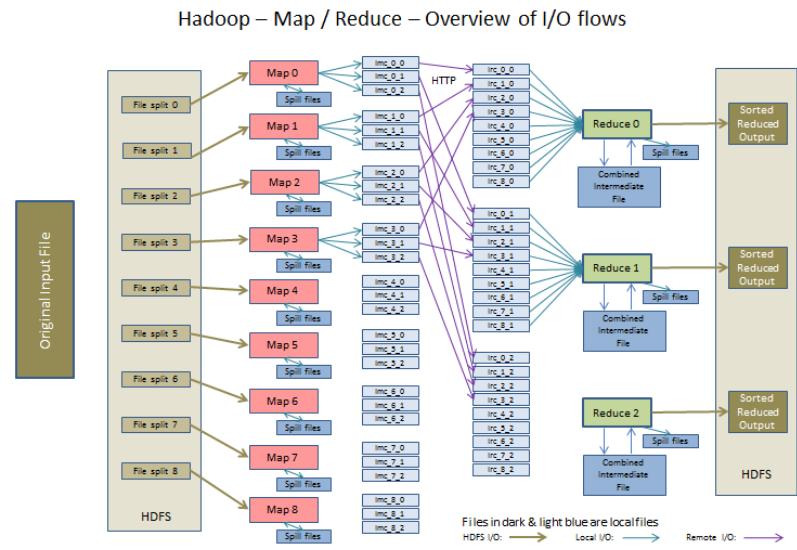
File split 3 → Map 3 → Spill files → imc_3_0, imc_3_1, imc_3_2 → HTTP → irc_0_3, irc_1_3, irc_2_3, irc_3_3, irc_4_3, irc_5_3, irc_6_3, irc_7_3, irc_8_3, ...

File split 4 → Map 4 → Spill files → imc_4_0, imc_4_1, imc_4_2 → HTTP → irc_0_4, irc_1_4, irc_2_4, irc_3_4, irc_4_4, irc_5_4, irc_6_4, irc_7_4, irc_8_4, ...

File split 5 → Map 5 → Spill files → imc_5_0, imc_5_1, imc_5_2 → HTTP → irc_0_5, irc_1_5, irc_2_5, irc_3_5, irc_4_5, irc_5_5, irc_6_5, irc_7_5, irc_8_5, ...

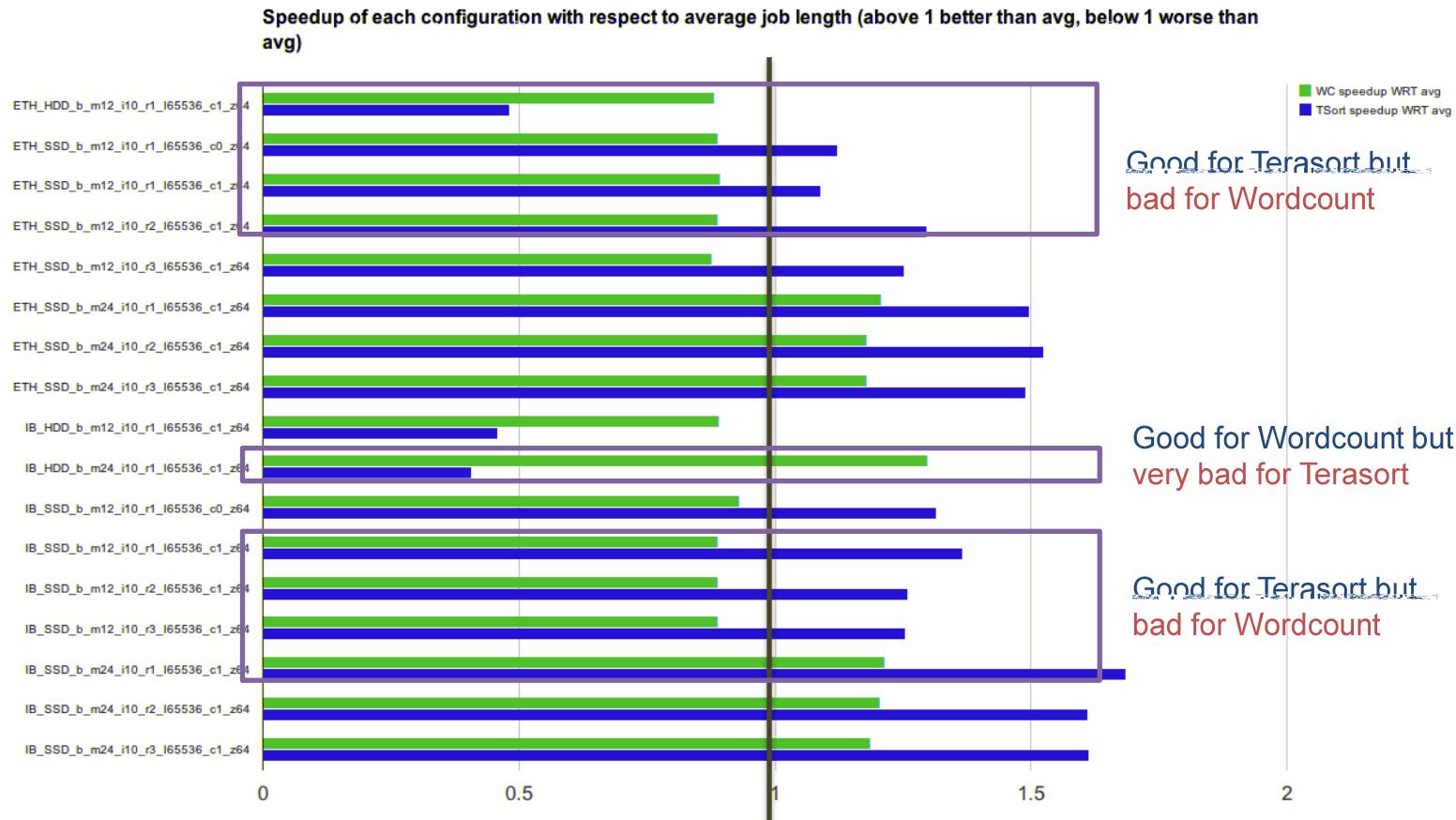
File split 6 → Map 6 → Spill files → imc_6_0, imc_6_1, imc_6_2 → HTTP → irc_0_6, irc_1_6, irc_2_6, irc_3_6, irc_4_6, irc_5_6, irc_6_6, irc_7_6, irc_8_6, ...

File split 7 → Map 7 → Spill files → imc_7_0, imc_7_1, imc_7_2 → HTTP → irc_0_7, irc_1_7, irc_2_7, irc_3_7, irc_4_7, irc_5_7, irc_6_7, irc_7_7, irc_8_7, ...



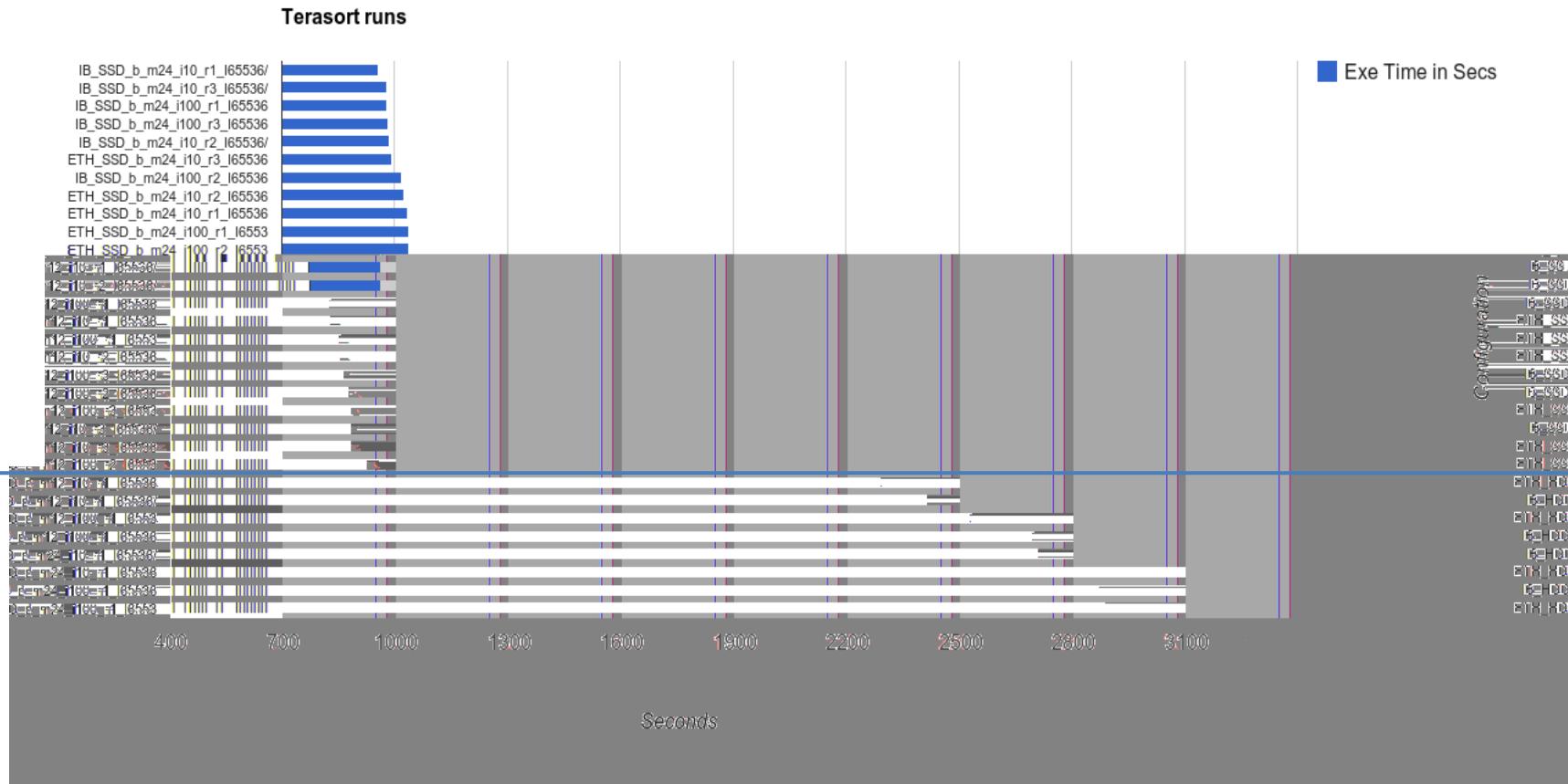
Early exploration: Job resource configs

Is there one software configuration iteration that fits everybody?



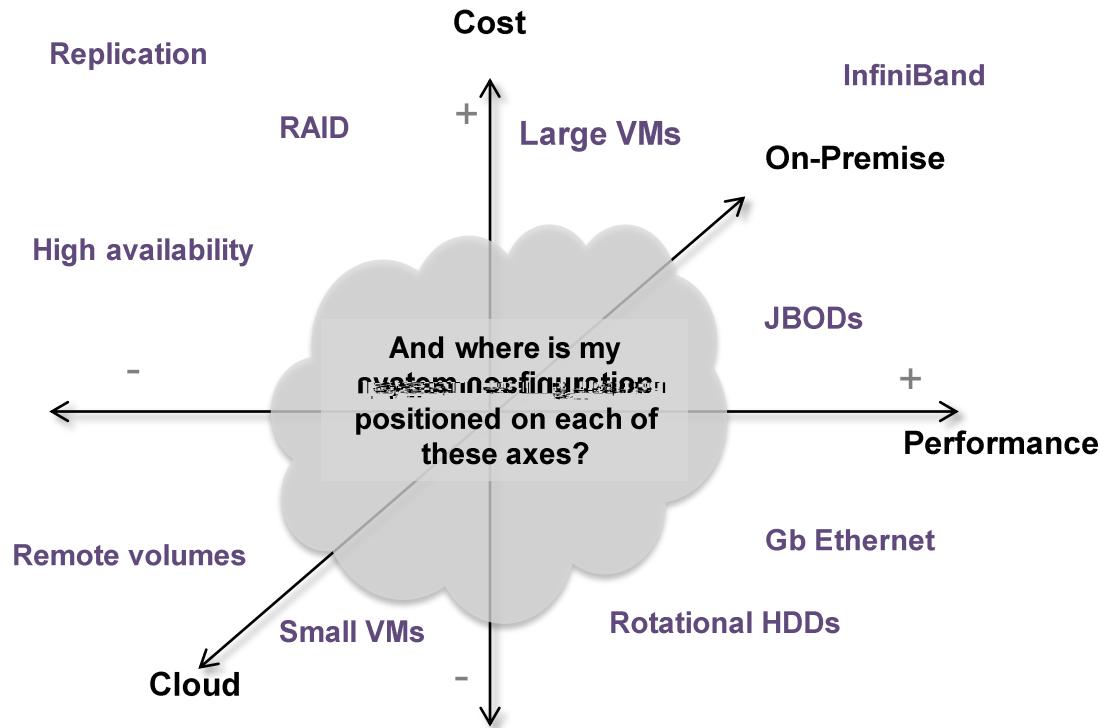
Early exploration: HW technology impact

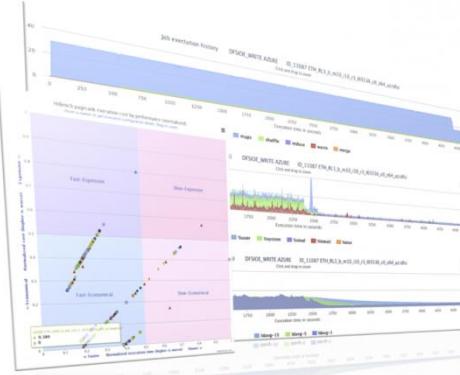
Impact of SSDs to running time of Terasort



Current scenario and problematic

- « What is the most cost-effective configuration for my needs?
 - Multidimensional problem





- « Joint initiative to produce mechanisms for an
 - automated characterization of cost-effectiveness
 - of Big Data deployments
- « Results from, of a growing, need of the community to understand job execution details
- « Explore different configuration deployment options and their tradeoffs
 - Both software and hardware
 - Cloud services and on-premise
- « Seeks to provide knowledge, tools, and an online service
 - to which users make better informed decisions
 - reduce the TCO for their Big Data infrastructures
 - Guide the future development and deployment of Big Data clusters and applications

ALOJA Project phases

1. Systematic study of Hadoop runtime executions
 - across a range of hardware components
 - software parameters, job types,
 - and deployment patterns
2. Analytical models of Hadoop cost-effectiveness
 - Price vs. performance evaluation
 - Expand exploration in Cloud: IaaS (HDI) vs. PaaS, storage
 - VM flavor and cluster characterization
3. Automation and Prediction
 - Modeling and Prediction of executions
 - Minimize # of executions
 - Job similarity characterization
 - Online learning of configuration and recommendation
 - ...

} We are here



PLATFORM

ALOJA Platform: Evolution and status

« Benchmarking, Online Repository and Analytics tools for Big Data »



« Composed of open-source

- Benchmarking, provisioning and orchestration tools,
- high-level system performance metric collection,
- low-level Hadoop instrumentation based on **BSC Tools**
- and Web based data analytics tools
 - And recommendations

« Online Big Data Benchmark repository of:

- 8000+ runs (from HiBench)
- Sharable, comparable, repeatable, verifiable executions

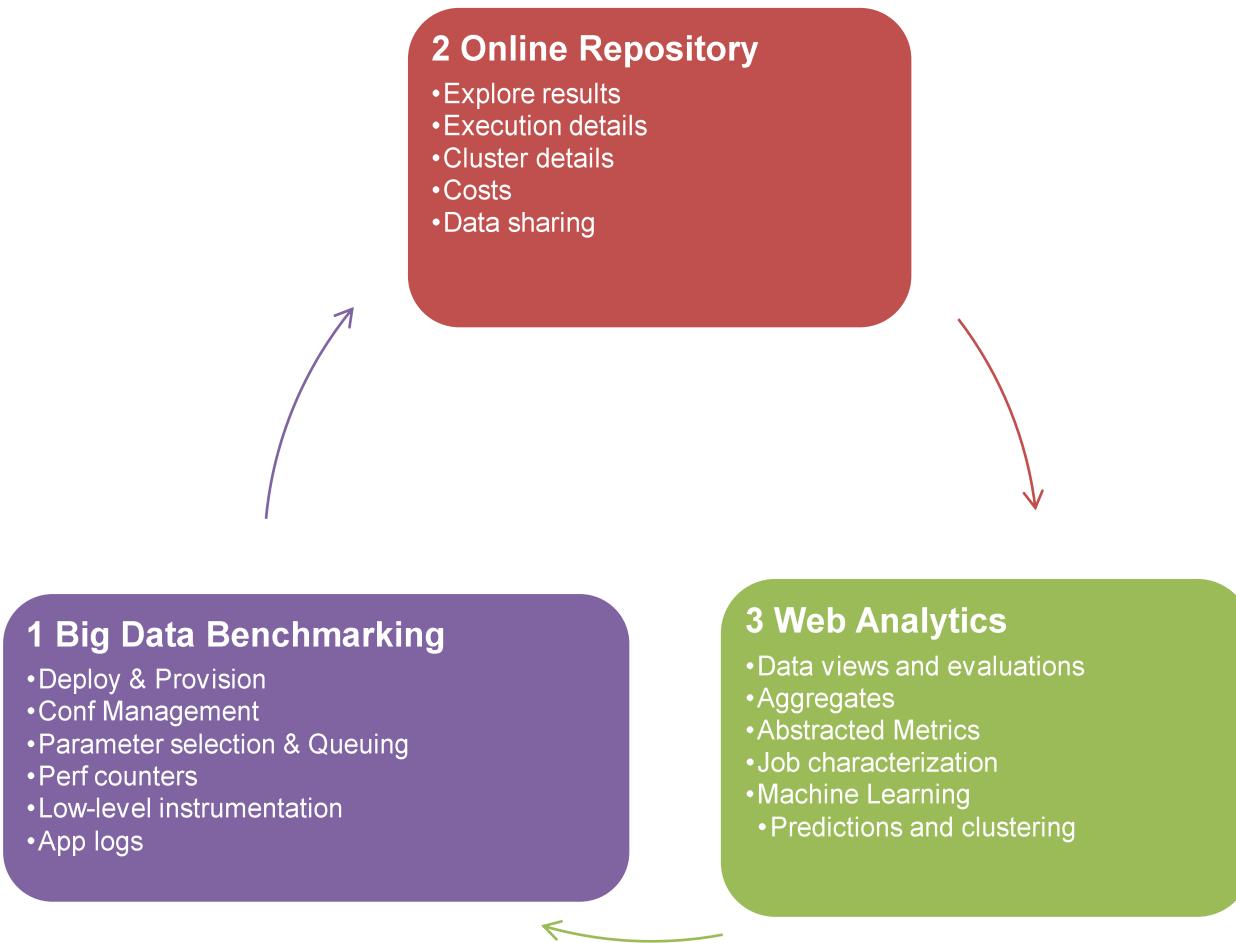
« Abstracting and leveraging tools for BD benchmarking

- ~~Abstracting from them~~, not re-inventing
- most current BD tools designed for production, **not for benchmarking**
- leverages current compatible tools and projects
 - When possible

« Dev VM toolset and sandbox

- via Vagrant

ALOJA Platform main components





1.) BIG DATA BENCHMARKING TOOLS

1.) Big Data Benchmarking

« ALOJA-DEPLOY Composed of scripts to:

- Automatically create, stop, delete clusters in the cloud
 - From a simple and abstracted node and cluster definition files
 - Both for Linux and Windows
 - IaaS and PaaS (HDInsight)
 - Abstracted to support multiple providers
- Provision and configuration of base software to servers
 - Both for cloud based as on premise
 - Composed of portable configuration management scripts
 - Designed for benchmarking needs
- Orchestrate benchmark executions
 - Prioritized job queues
 - Results gathering and packaging

« ALOJA-BENCH

- Multi-benchmark support
- Flexible performance counter options
- Dynamic SW and HW configurations

```
Done 00000283_conf_ETH_RL3_b_m4_i10_r2_I32768_c0_z128_al-08_terasort
Executing: 00000284_conf_ETH_RL3_b_m4_i10_r2_I32768_c0_z256_al-08_terasort
20150122_161449 7949: INFO: Loading benchmarks_defaults.conf
20150122_161449 7949: INFO: loading /home/pristine/share/shell/common/..../conf/cluster_a-08.conf
20150122_161449 7949: Starting ALOJA deploy tools for Provider:
20150122_161449 7949: INFO: loading /home/pristine/share/shell/common/..../aloja-deploy/providers/azure.sh
20150122_161449 7949: INFO: loading /home/pristine/share/shell/common/common_benchmarks.sh
20150122_161449 7949: INFO: loading /home/pristine/share/shell/common/common_hadoop.sh
1421943289 : STARTING EXECUTION of 20150122_161449_conf_ETH_RL3_b_m4_i10_r2_I32768_c0_z128_al-08
```

Node and Cluster definitions

- « The configuration files can be used with the scripts to automate the definition of clusters (at provisioning stage)
- « Example: **cluster_al-14.conf**

```
defaultProvider="azure"
clusterID='14' #from 03 0 99
clusterName="al-${clusterID}"
numberOfNodes="8" #starts at 0 (max 99) 0 is assigned to master
size 'vmSize' 'extralarge' #extralarge are A4s
attachedVolumes="3"
diskSize="512"
queueJobs="true" #enable on cluster config to queue benchmarks after deploy
vmCores="8"
vmRAM="14GB"
clusterCostHour="2.664"
clusterType="IaaS"
```

Sources: <https://github.com/Aloja/aloja/tree/master/shell/conf>

Provisioning scripts

<https://github.com/Aloja/aloja/tree/master/aloja-deploy>

| | | |
|--|---|---|
|  cache | Adding cache dir | 2 months ago |
|  include | Changed variable name | a month ago |
|  providers | Added a comment to quickly edit host line | 23 hours ago |
|  README.md | Changes for multi cloud provider2 | 3 months ago |
|  connect_cluster.sh | Change in providers and default values | 3 months ago |
|  connect_node.sh | Changes for multi cloud provider2 | 3 months ago |
|  delete_cluster.sh | Calculating total time | 3 months ago |
|  delete_node.sh | Changes for multi cloud provider2 | 3 months ago |
|  deploy_cluster.sh | Renaming of global vars | 3 months ago |
|  deploy_node.sh | Download and install ARM JDK | 3 months ago |
|  start_cluster.sh | Calculating total time | 3 months ago |
| 3 months ago |  start_node.sh | Improvements in cluster deployments |
| 3 months ago |  stop_cluster.sh | Calculating total time |
| 3 months ago |  stop_node.sh | Multi provider cleanup |
| a month ago |  sync_node.sh | Added a new command to sync code changes without deploy |

Running benchmarks in ALOJA

↳ <https://github.com/Aloja/aloja/tree/master/shell>

- Example of submitting a job to run:
 - https://github.com/Aloja/aloja/blob/master/shell/run_benchmarks.sh

```
run_benchmarks.sh -C al-04 -n IB -d HDD -r 1 -m 12 -i 10 -p 3 -b -min -I 4096 -l wordcount -c 1
```

*al-04 cluster must be previously provisioned with the provisioning scripts

- Controls de Jobs in execution:
 - <https://github.com/Aloja/aloja/blob/master/shell/exeq.sh>

```
Done 00000283.conf_ETH_RL3_b_m4_i10_r2_I32768_c0_z128_al-08_terasort
Executing: 00000284.conf_ETH_RL3_b_m4_i10_r2_I32768_c0_z256_al-08_terasort
20150122_161449 7949: INFO: Loading benchmarks_defaults.conf
20150122_161449 7949: INFO: loading /home/pristine/share/shell/common/..../conf/cluster_a
l-08.conf
20150122_161449 7949: Starting ALOJA deploy tools for Provider:
20150122_161449 7949: INFO: loading /home/pristine/share/shell/common/..../aloja-deploy
/providers/azure.sh
20150122_161449 7949: INFO: loading /home/pristine/share/shell/common/common_benchmarks
sh
20150122_161449 7949: INFO: loading /home/pristine/share/shell/common/common_hadoop.sh
1421943289 : STARTING EXECUTION of 20150122_161449.conf_ETH_RL3_b_m4_i10_r2_I32768_c0_z
56_S8_al-08
```

Initial testing infrastructure

« High-End Cluster:

- ~~Doesn't have 16 cores, 128GB RAM, 8x SSDs, 2x GbE, 1Gb GbE (bonding)~~

« Mid-end Cluster:

- 18 nodes, 12 real cores, 64GB RAM, 1x SSD, 6x HDDs, 1Gb GbE
 - Evaluating different number of datanodes performance

« Cloud IaaS (Azure)

- ~~1 broad node~~ 8 datanodes of A3, A4, A6, A7 VMs

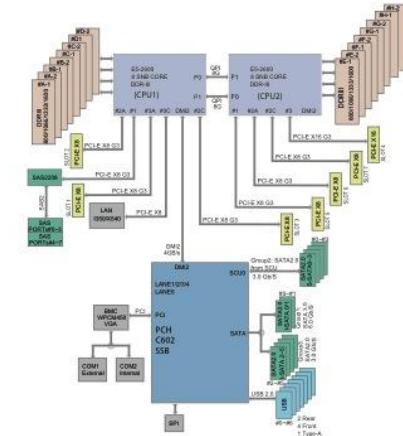
« Cloud PaaS (HDInsight)

- 4, 8, 16, 32 datanodes

« Low-powered cluster:

- 10-node ARM based cluster*

* Result numbers not online yet





ONLINE BENCHMARK REPOSITORY

2.) ALOJA-WEB Online Repository

« Entry point for explore the results collected from the executions

- Index of executions
 - Quick glance of executions
 - Searchable, Sortable
- Execution details
 - Performance charts and histograms
 - Hadoop counters
 - Jobs and task details

Available at: <http://hadoop.bsc.es>

« Data management of benchmark executions

- Data importing from different clusters
- Execution validation
- Data management and backup

« Cluster definitions

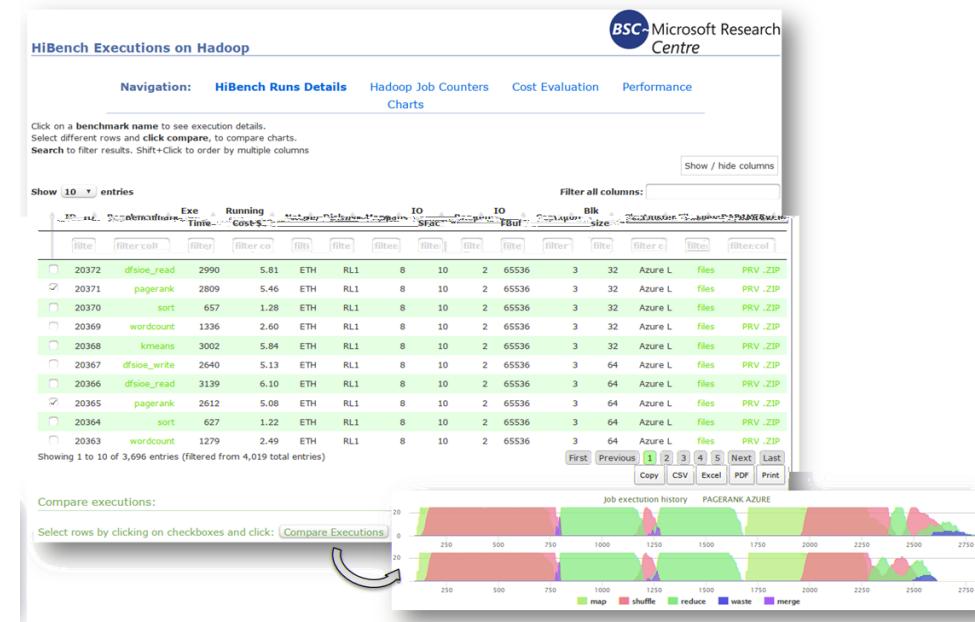
- Cluster capabilities (resources)
- Cluster costs

« Sharing results

- Download executions
- Add external executions

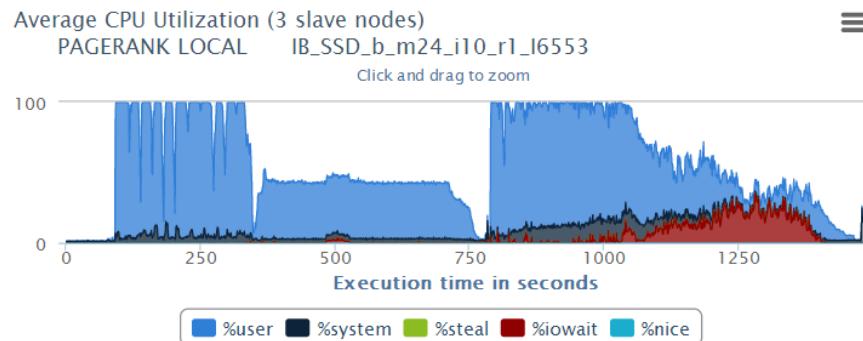
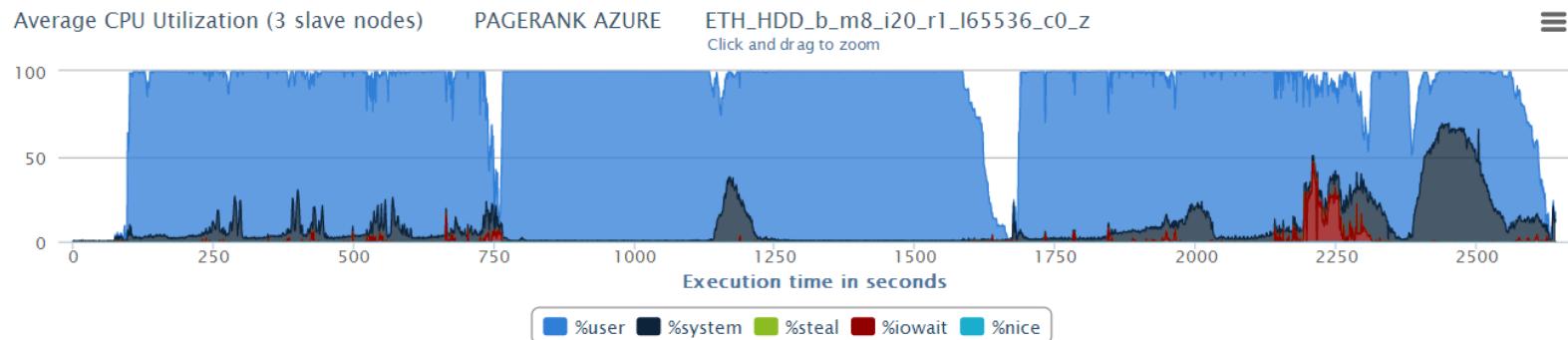
« Documentation and References

- Papers, links, and feature documentation



Benchmarks Execution comparisons

- « You can compare, side by side, all execution parameters:
 - CPU, Memory, Network, Disk, Hadoop parameters....



Example: 24 maps in parallel, SSD vs HDD vs ETH vs IB

Terasort



ETH+HDD

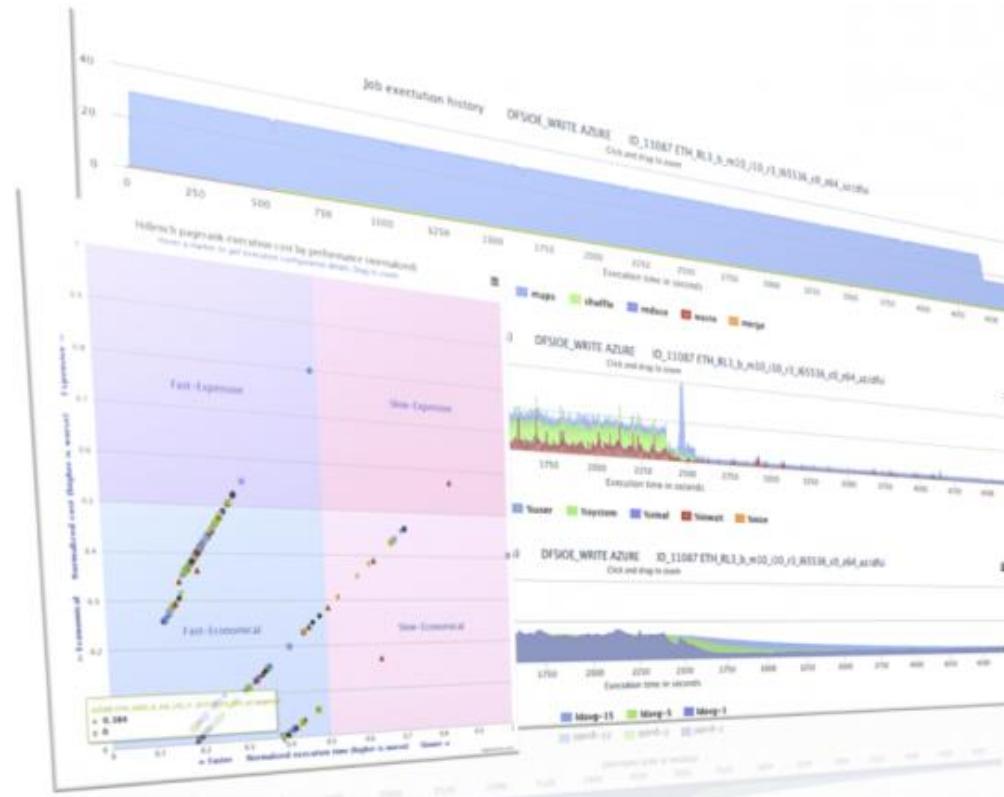
IB+HDD

ETH+SSD

IB+SSD

ALOJA-WEB

- « Entry point for explore the results collected from the executions.
 - Provides insights on the obtained results through continuously evolving data views.
- « Online **DEMO** at: <http://hadoop.bsc.es>

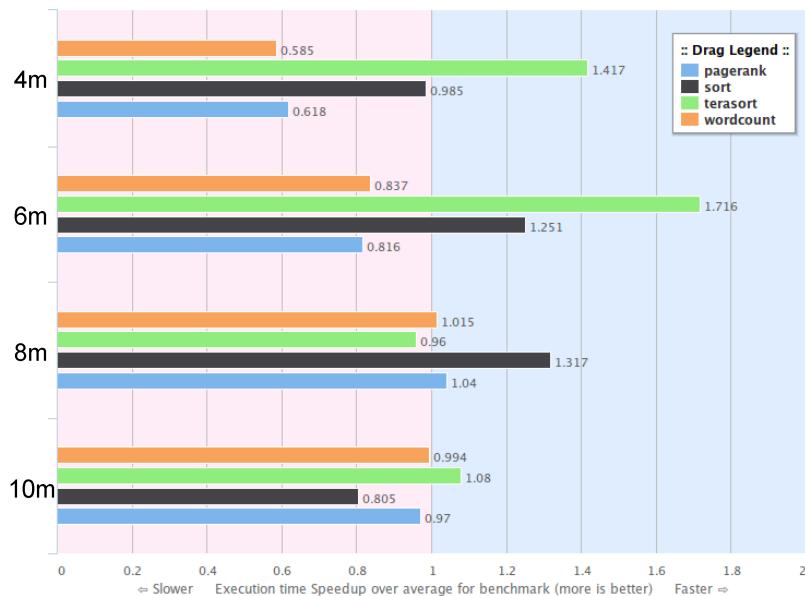




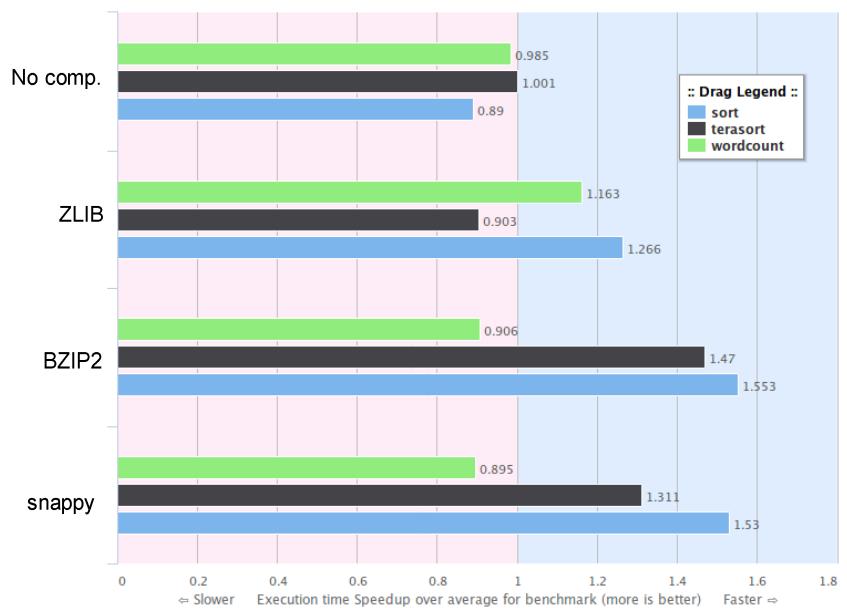
EARLY FINDINGS IN SW AND HW CONFIGURATIONS

Impact of SW configurations in Speedup

Number of mappers



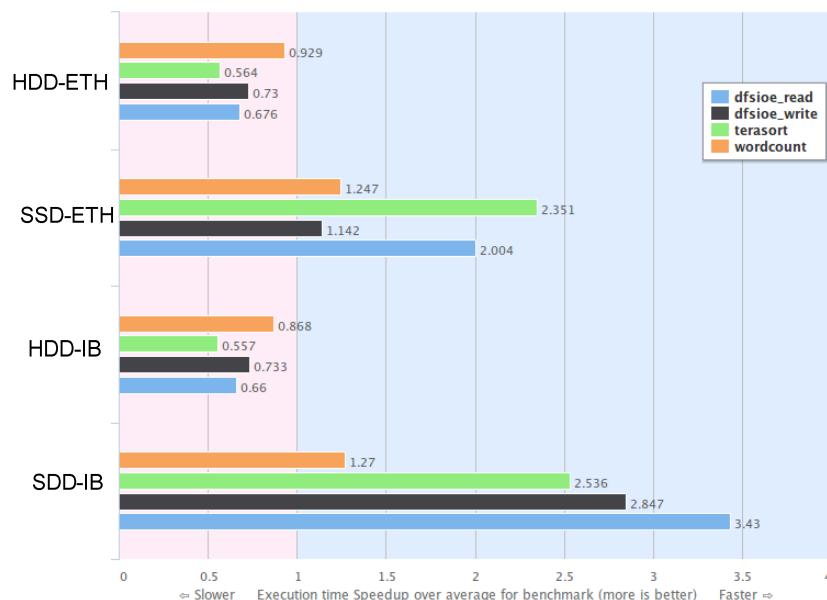
Compression algorithm



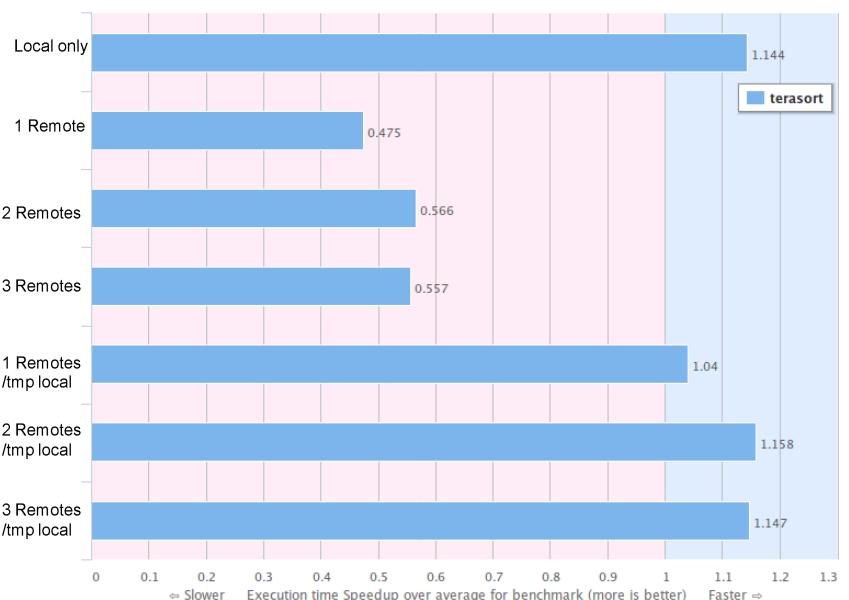
Speedup (higher is better)

Impact of HW configurations in Speedup

Disks and Network

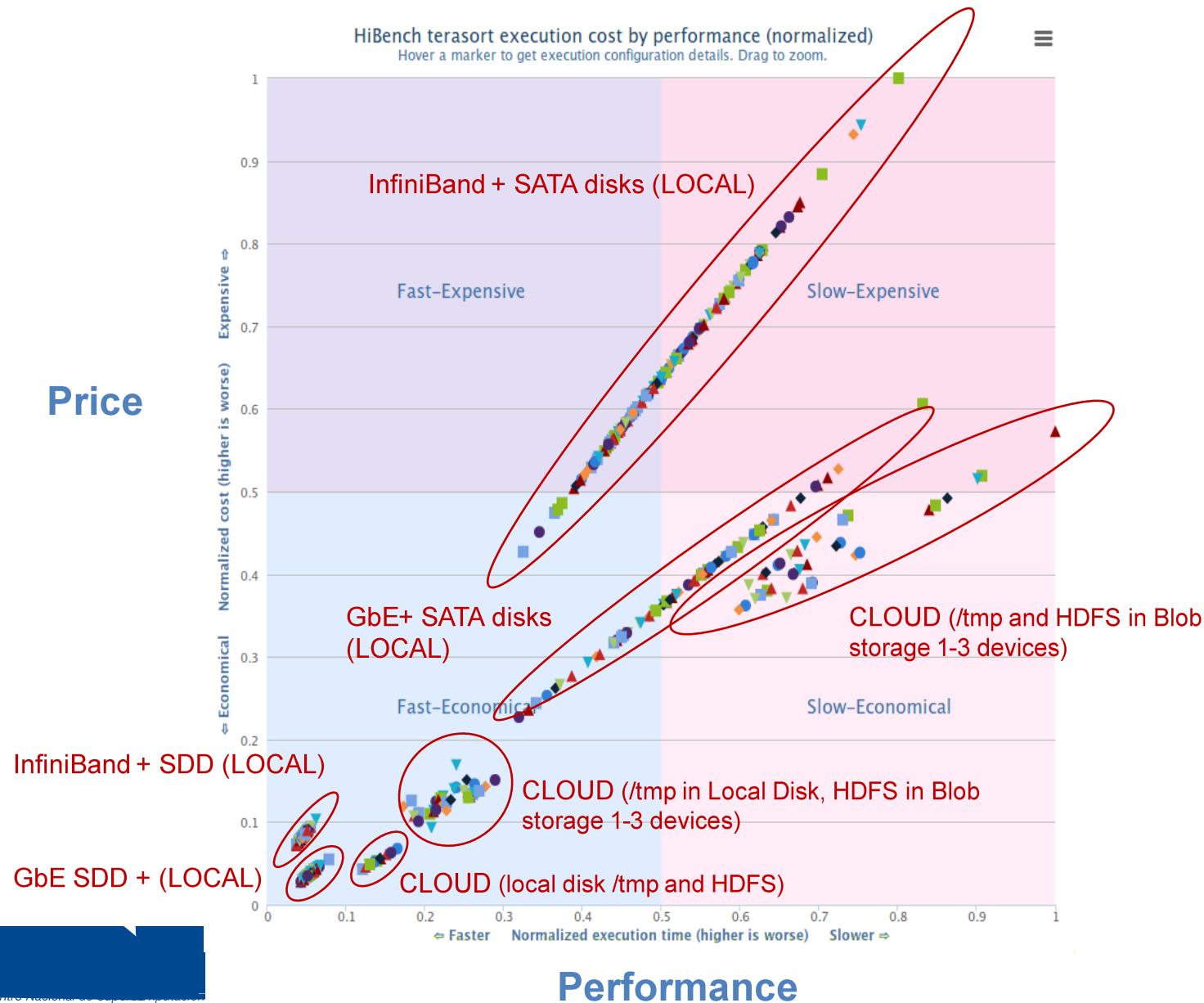


Cloud remote volumes



Speedup (higher is better)

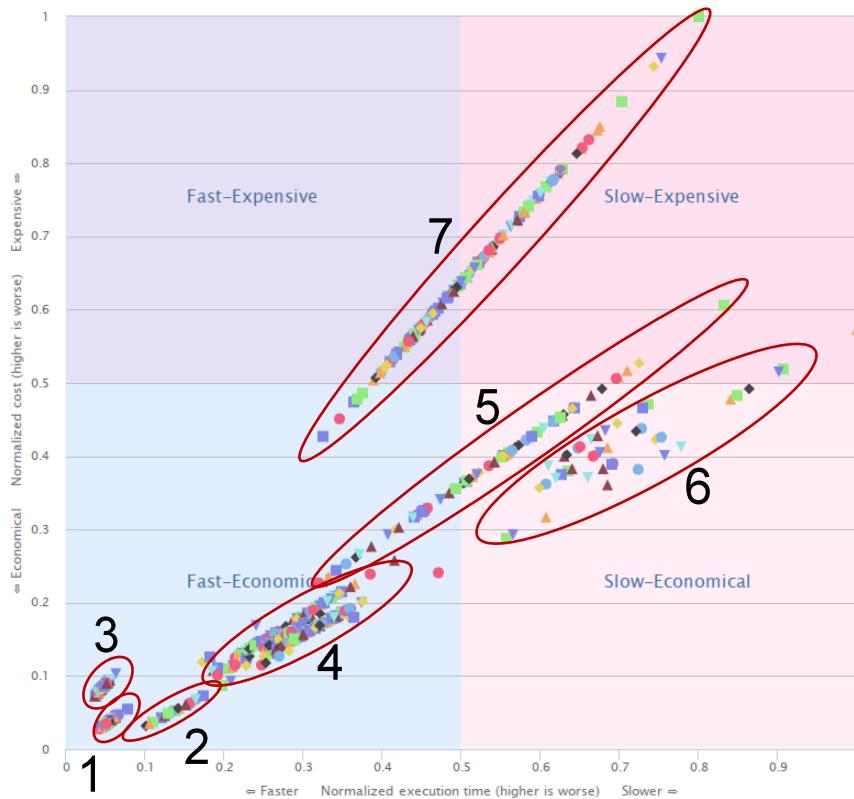
Cost-effectiveness of SW and HW (On-premise vs. Cloud)



Cost-effectiveness of SW and HW

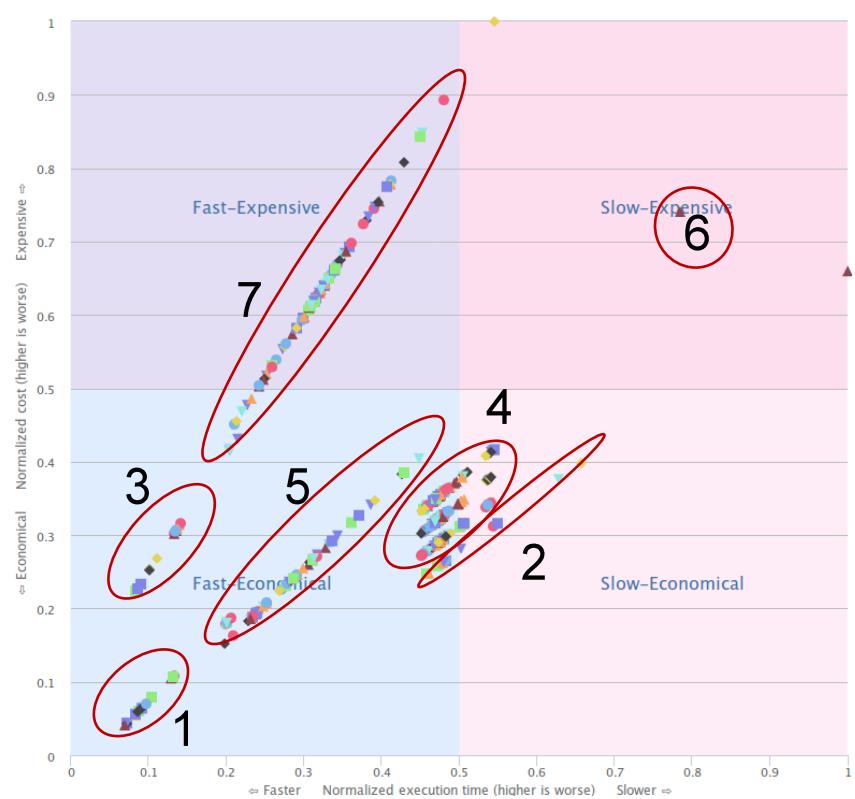
Point (0,0) represents most cost-effective execution

Terasort



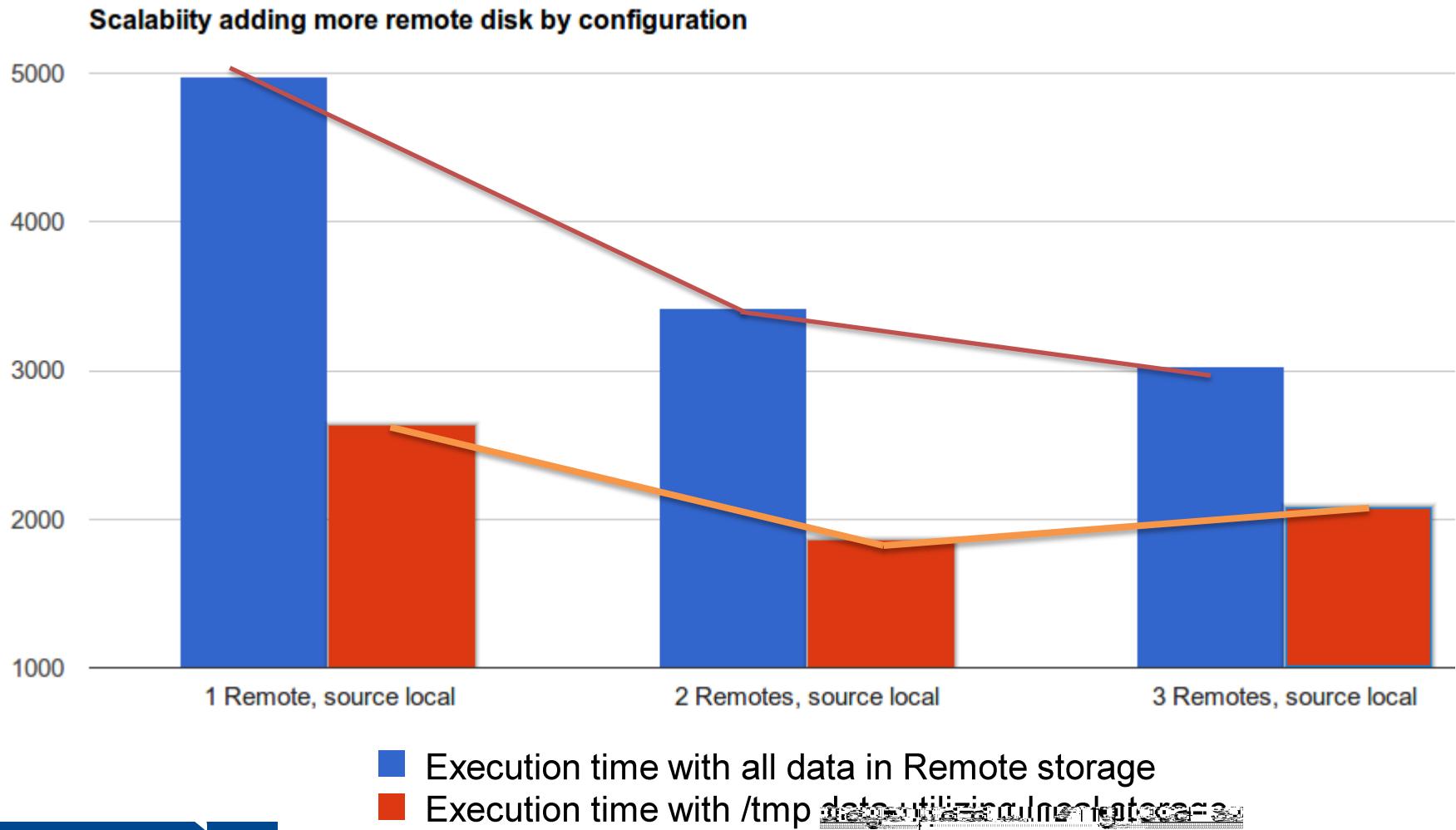
- 1) On-premise cluster: SSD disks + GbE.
- 2) Azure IaaS: Only local disk, virtualized SSD and GbE (baseline).
- 3) On-premise cluster: SSD disks + InfiniBand.

Wordcont

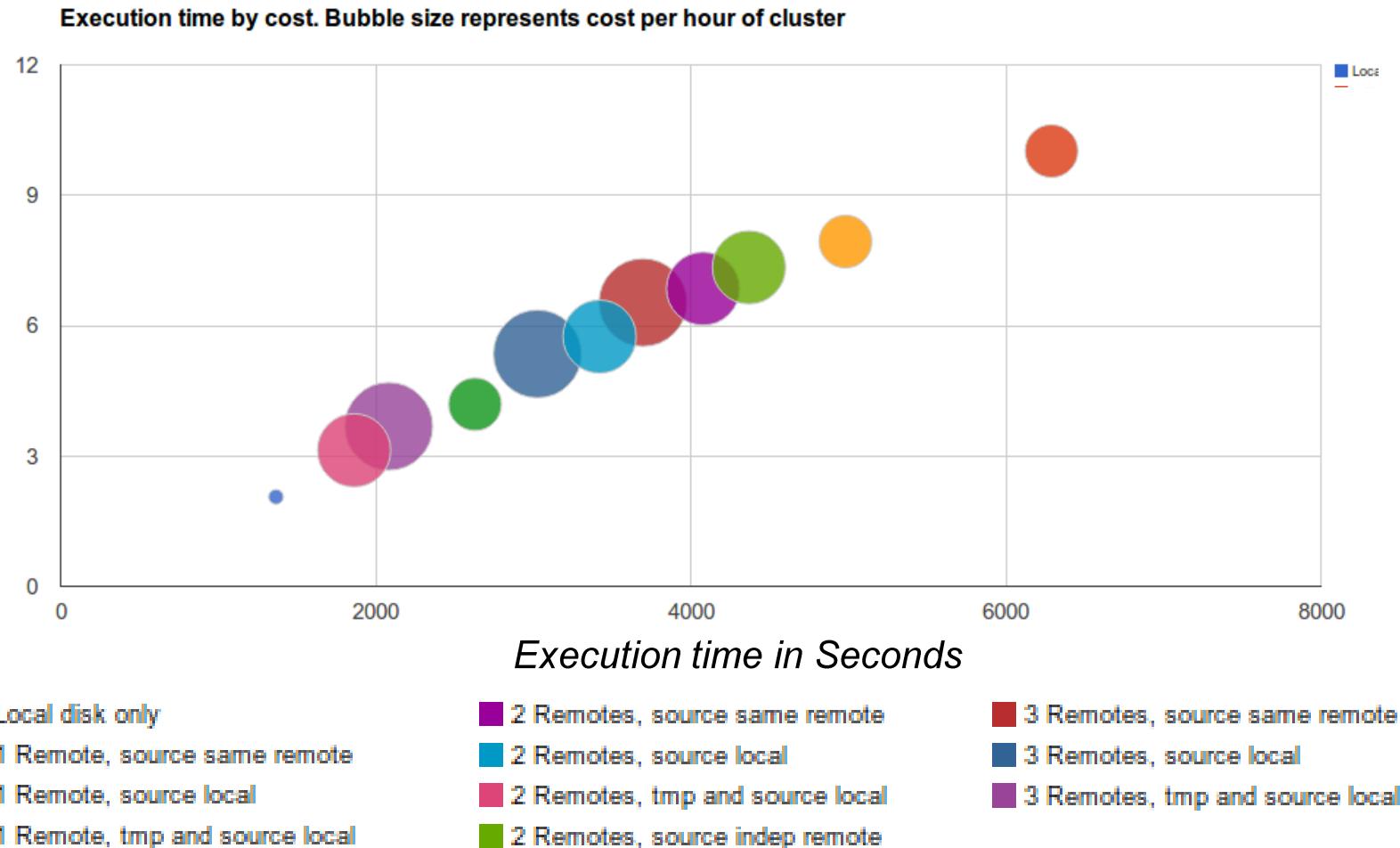


- 4) Azure IaaS: 1-3 remote vol. and Hadoop /tmp to local disk (SSD) and GbE
- 5) On-premise cluster: 1 SATA disk + GbE.
- 6) Azure IaaS: 1-3 remote volumes (Blob storage).
- 7) On-premise cluster: 1 SATA disk + InfiniBand.

Cloud IaaS impact of utilizing local vs. remote storage



Cost efficiency of different Cloud deployment options





3.) ADVANCED WEB ANALYTICS WITH ALOJA-ML

3.) ALOJA-ML: Advanced Web Analytics

« Data views and evaluations

- Best configuration recommendation
- Configuration improvement
- Parameter evaluation

« Cost / Performance analysis

« Aggregation and data filters

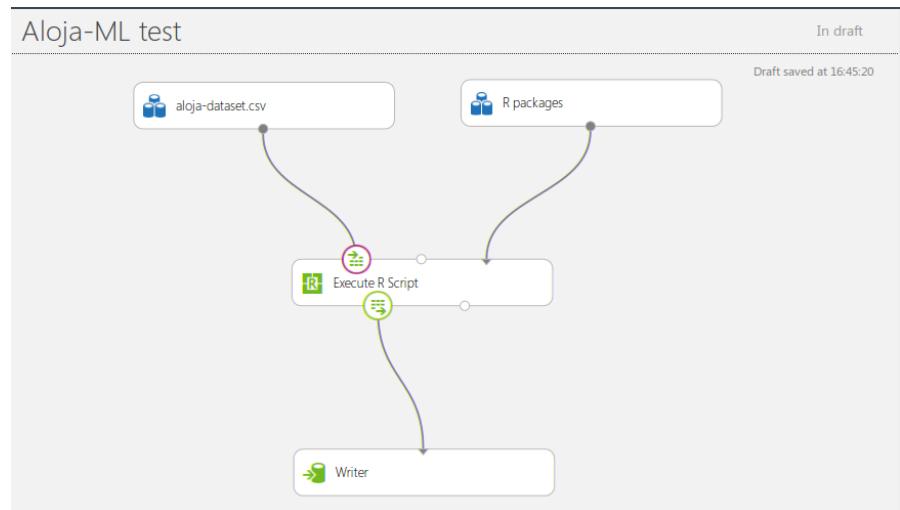
- Aggregated performance metrics
- Data filters
- Abstracted Metrics

« Job characterization

- Job resource consumption and bottlenecks
- Job execution characteristics

« Machine Learning

- Execution time prediction
 - By SW and HW configuration
- Estimation of missing values
- Clustering
 - DBScan and K-means for different views



New feature: DBSCAN

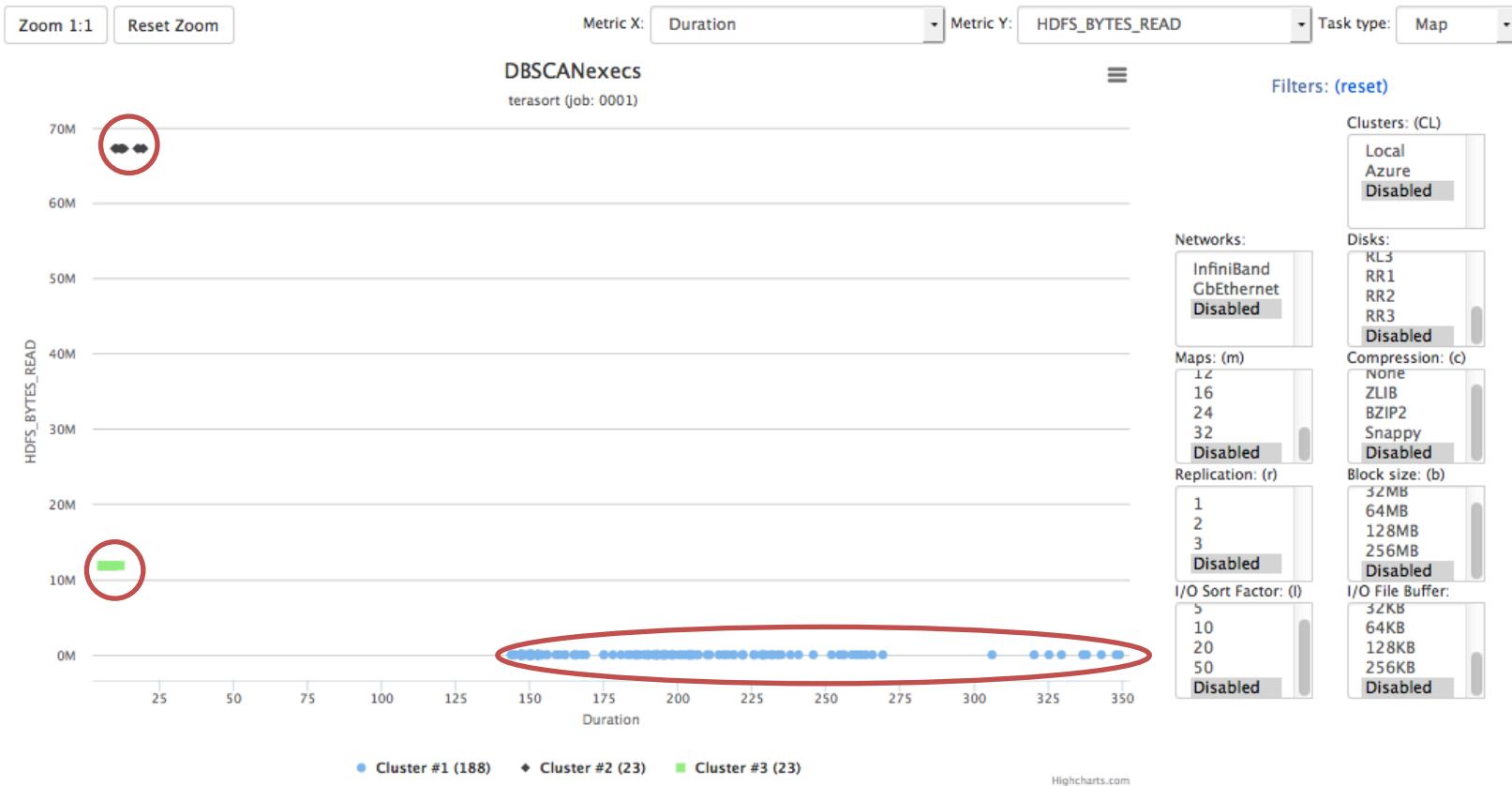
DBSCAN is a data clustering algorithm

- It finds a number of clusters starting from the estimated density distribution of corresponding nodes

General overview of all executions of a benchmark

- Select two metrics, Optionally filter desired parameters

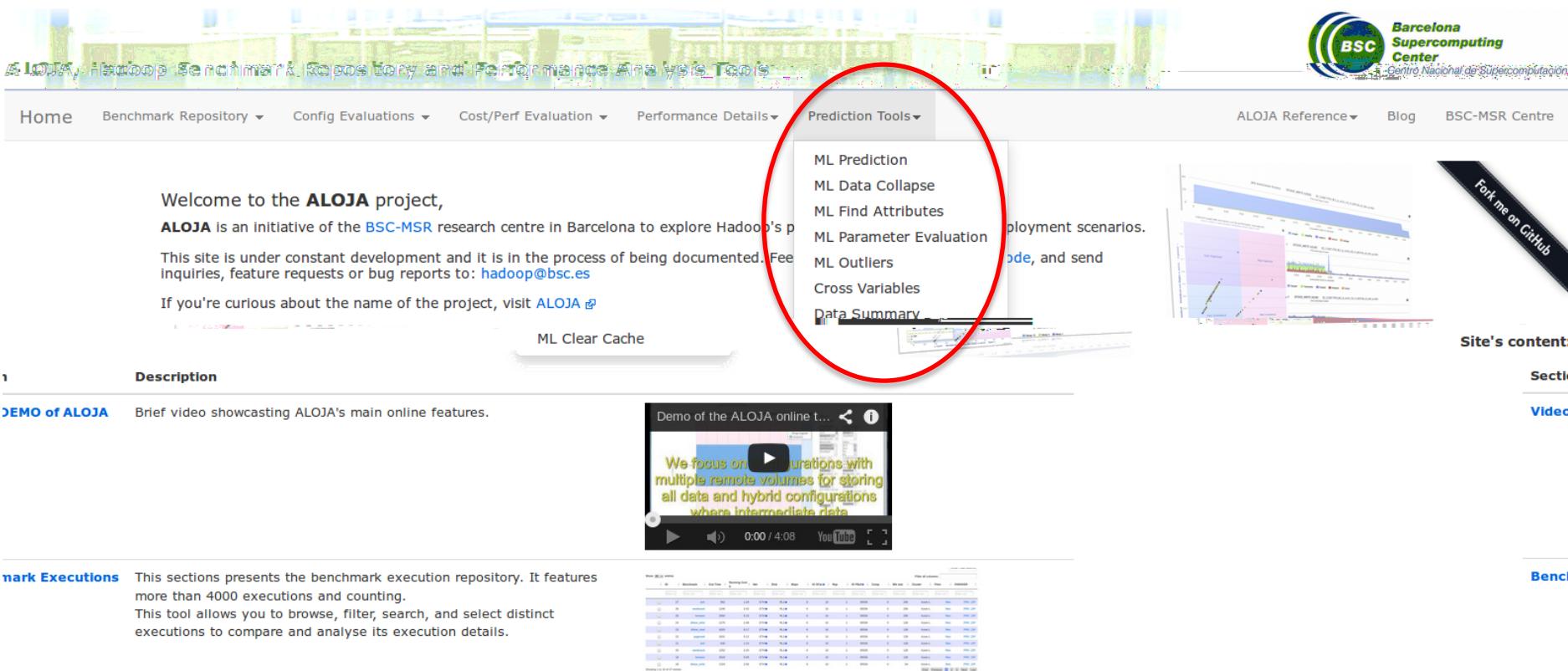
Automatic clustering of similar executions



Upcoming component : ALOJA-ML

ALOJA-ML

- Provides automatic means to characterize job executions and clusters



The screenshot shows the ALOJA web interface. At the top, there is a banner with the text "ALOJA: Hadoop Benchmark Repository and Performance Analysis Tools". The navigation bar includes links for Home, Benchmark Repository, Config Evaluations, Cost/Perf Evaluation, Performance Details, and Prediction Tools. The Prediction Tools dropdown is open, showing options like ML Prediction, ML Data Collapse, ML Find Attributes, etc., with "ML Prediction" being the selected item. A red circle highlights this dropdown menu. To the right of the menu, there is a section titled "Data Summary" with a small screenshot of a visualization. On the far right, there is a "Fork me on GitHub" button and a "Site's content" section.

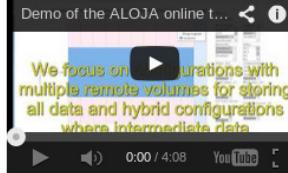
Welcome to the **ALOJA** project,
ALOJA is an initiative of the [BSC-MSR](#) research centre in Barcelona to explore Hadoop's p
This site is under constant development and it is in the process of being documented. Fee
inquiries, feature requests or bug reports to: hadoop@bsc.es

If you're curious about the name of the project, visit [ALOJA](#)

ML Clear Cache

Description

DEMO of ALOJA Brief video showcasing ALOJA's main online features.

Demo of the ALOJA online t... 

Benchmark Executions This sections presents the benchmark execution repository. It features more than 4000 executions and counting.
This tool allows you to browse, filter, search, and select distinct executions to compare and analyse its execution details.

Sections

Video

Benchmarks

The ALOJA-ML tool-set

1. Modeling and Prediction

- From ALOJA dataset → Find a model for $\langle \text{WorkId}, \text{Conf} \sim \text{Exe.Time} \rangle$

2. Configuration recommendation

- Rank (un)seen confs. for a benchmark from their expected Exe.Time

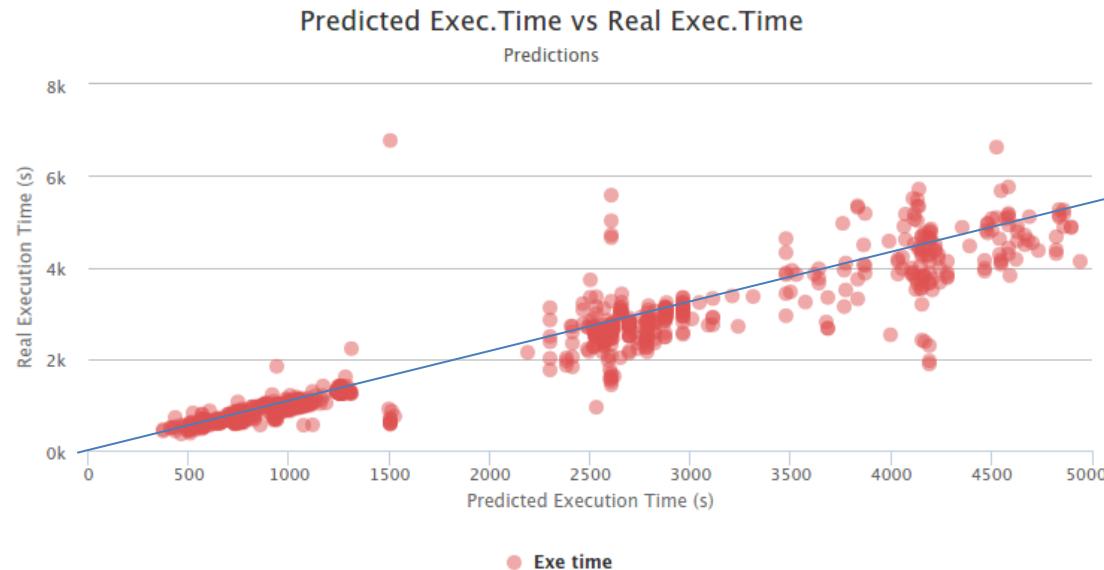
3. Anomaly detection (outliers)

- Statistic + Model-based detection of anomalous executions

4. ~~Deviation and statistic information for behavior mode~~

- Aggregate variables around the ones we want to observe
- Show frequency, percentiles and other useful information from ALOJA datasets

Modeling and Prediction



Real vs. Predicted execution times

Current available methods:

- *Regression Trees*
- *Nearest Neighbors*
- *FFA Neural Networks*
- *Multinomial Regression*

Prediction capabilities

« Different techniques used

- *Regression Trees, Nearest Neighbors, FFANNs, Multinomial Regs...*
- Mean Absolute Errors around **200s** [ranges from 100 to 6000]
- Relative Absolute Errors of 0.12 to 0.15 (that's actually good!)

- Without going deeper, we can learn from < **1000** different random observations [Current tested dataset: 4400 instances]

« A model can be used to:

- Predict expected execution times for unobserved configurations
- Determine if an observation is an **outlier** (anomaly)
- Determine which configuration properties influences more a run

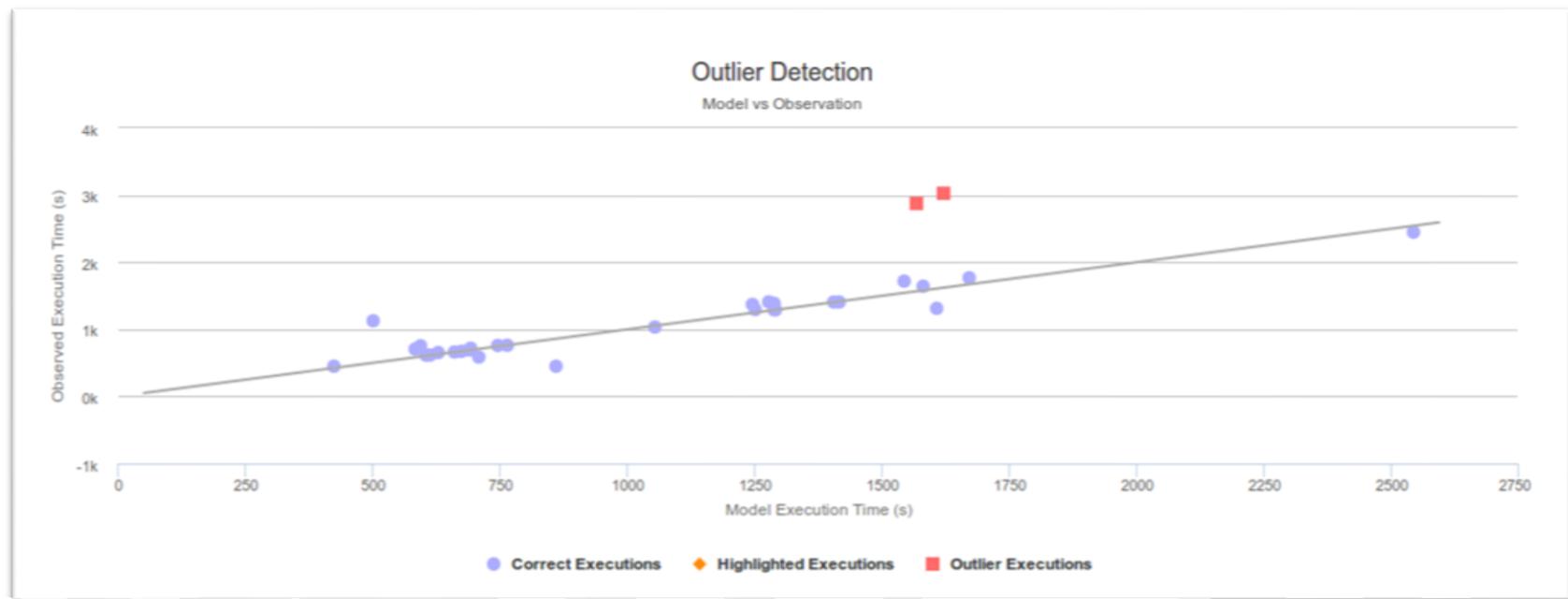
Rank and Recommend Configurations

- « Predict a range of configurations, previously seen or unseen
 - « Order them by predicted execution time and rank configurations
 - « Compare also predicted execution times vs. observed execution times, if there are any.

Anomaly Detection

« Anomaly and Outlier Detection

- Use of statistic and model-based outlier detections
- Highlight executions with high probability of anomaly
- Mark down executions with high probability of being errors



Data aggregation and Statistics

- Tools for data aggregation (also predicting their aggregates)
 - Find relevance or discard parameters

| Exe.Time: Observed / Estimated | | | |
|--------------------------------|------|----------|-----------|
| | sort | terasort | wordcount |
| HDD:Cmp0:1:131072 | 747 | 1622 | 1288 |
| HDD:Cmp0:1:32768 | 2810 | 3300 | 992 |
| SSD:Cmp0:1:131072 | 457 | 1022 | 718 |
| SSD:Cmp0:1:32768 | 454 | 967 | 674 |

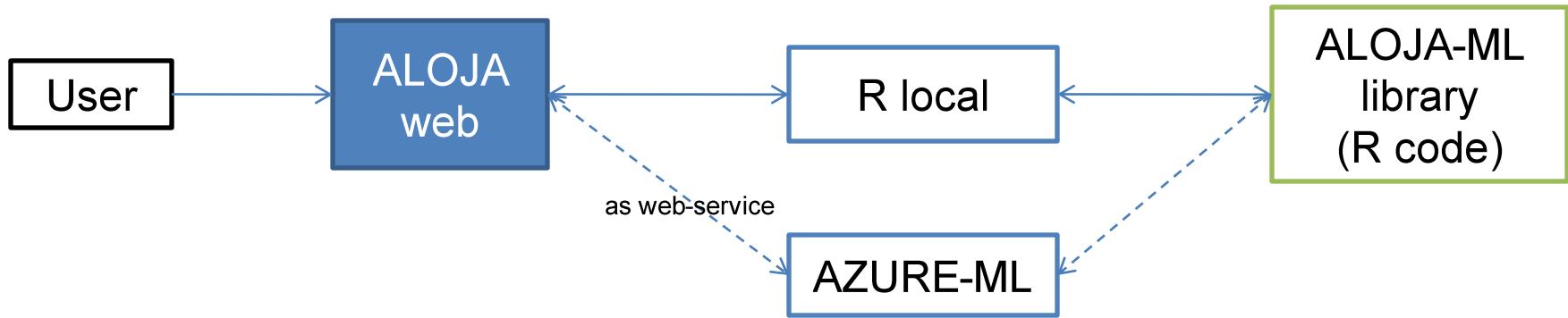
Show 10 entries Search:

Showing 1 to 4 of 4 entries Previous [1](#) Next

ALOJA-ML engine as a Cloud Service

« ALOJA-ML works with R

- AZURE-ML has incorporated recently R to its workbench
- We can run the ML engine locally, also use AZURE-ML as an option



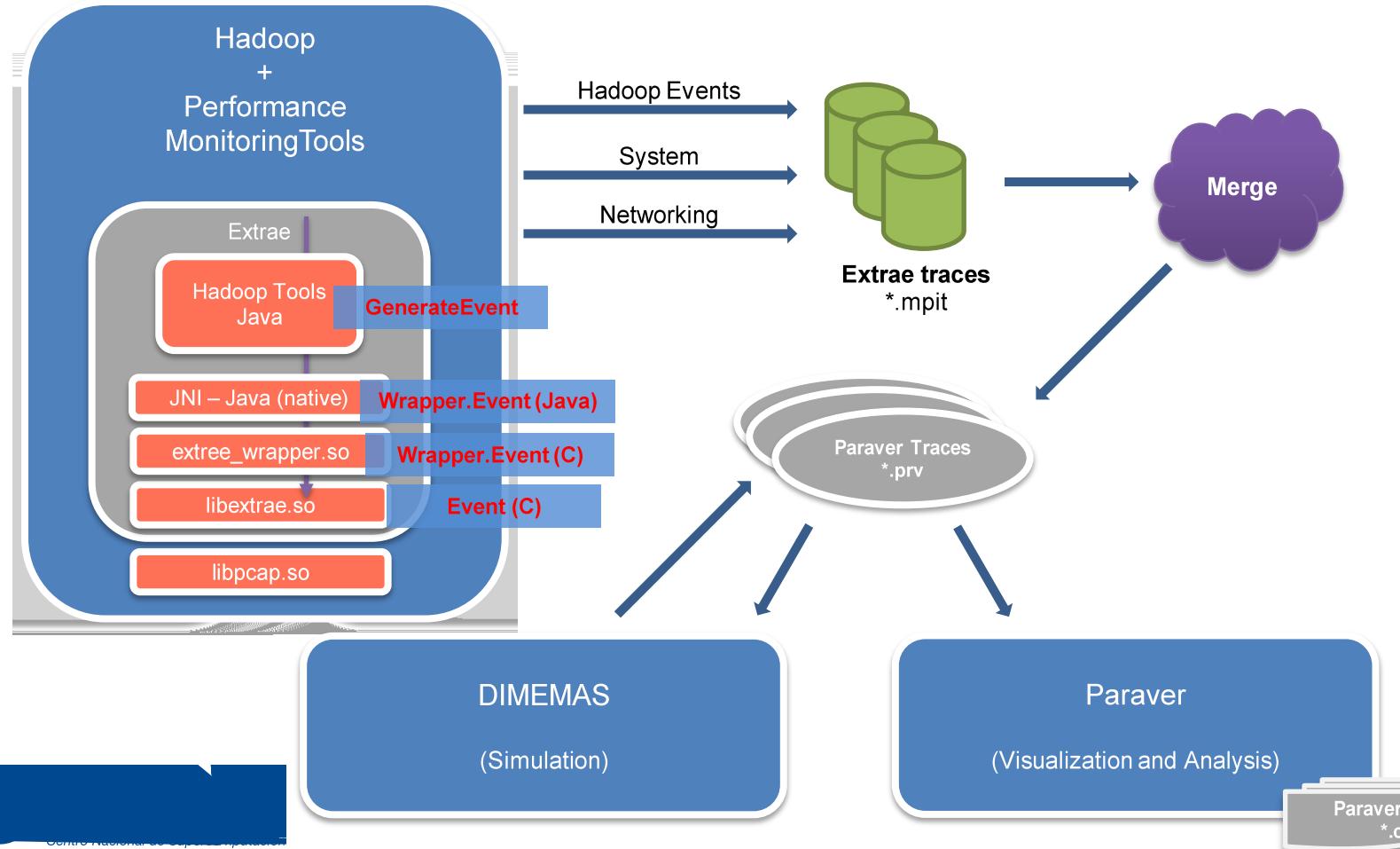
- « Most of the R code of ALOJA-ML library can run on AZML+R
- « The ML process can be delegated to a AZML web-service

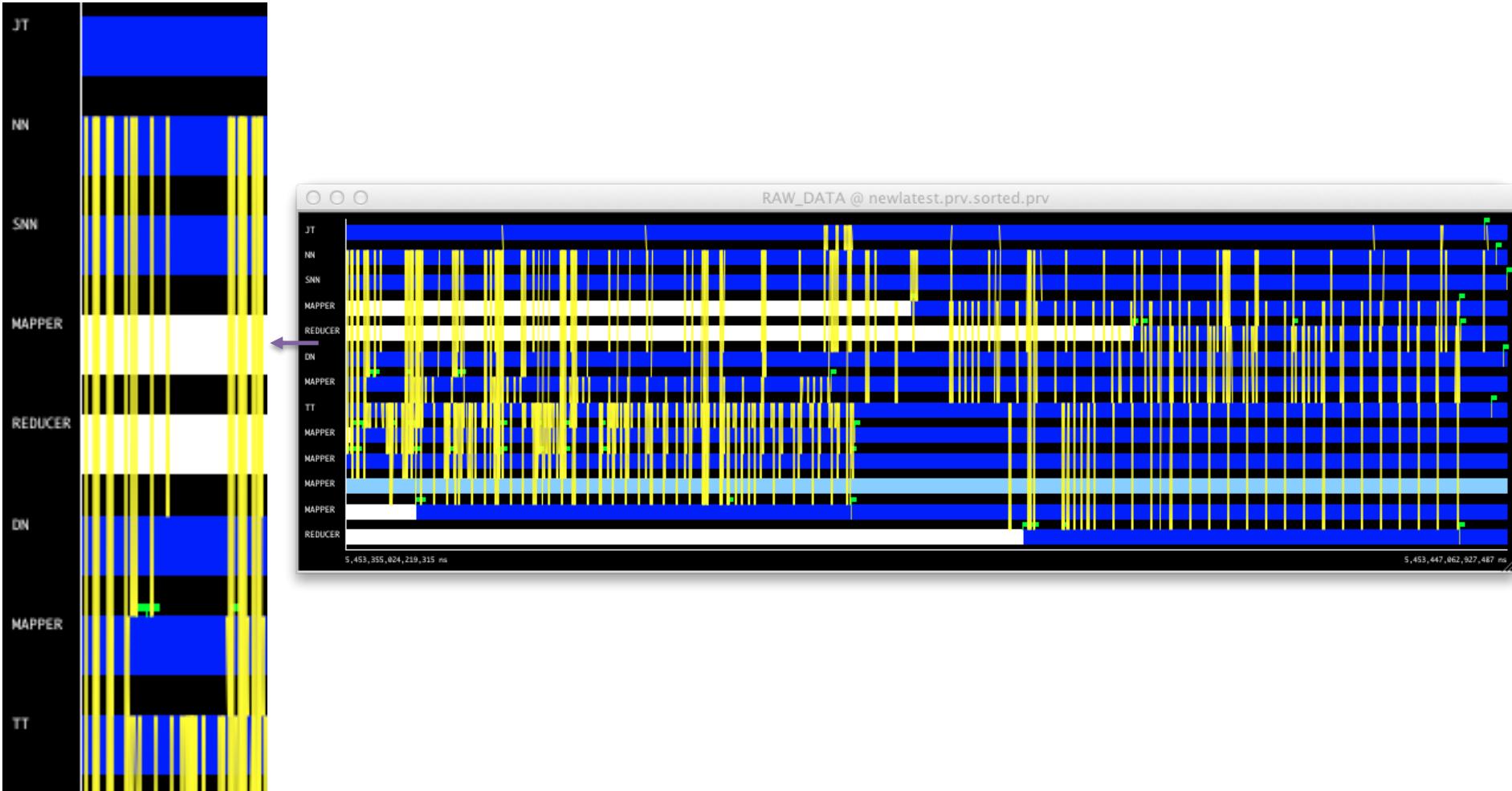


DEEP INSTRUMENTATION WITH BSC HPC TOOLS

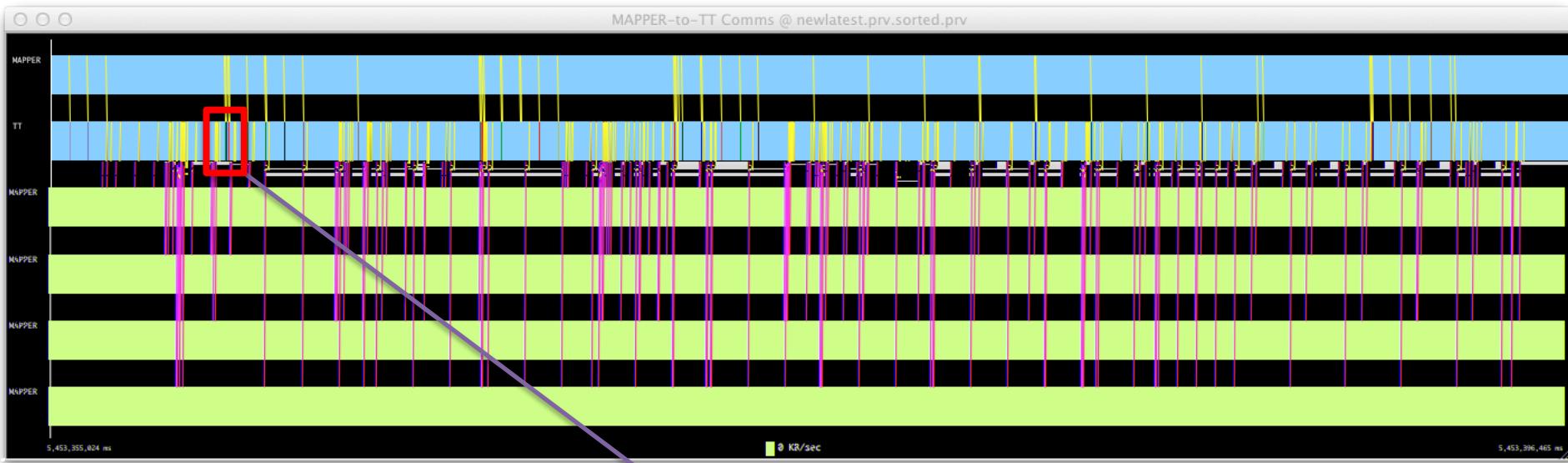
Overview

« Hadoop Analysis Toolkit and BSC tools



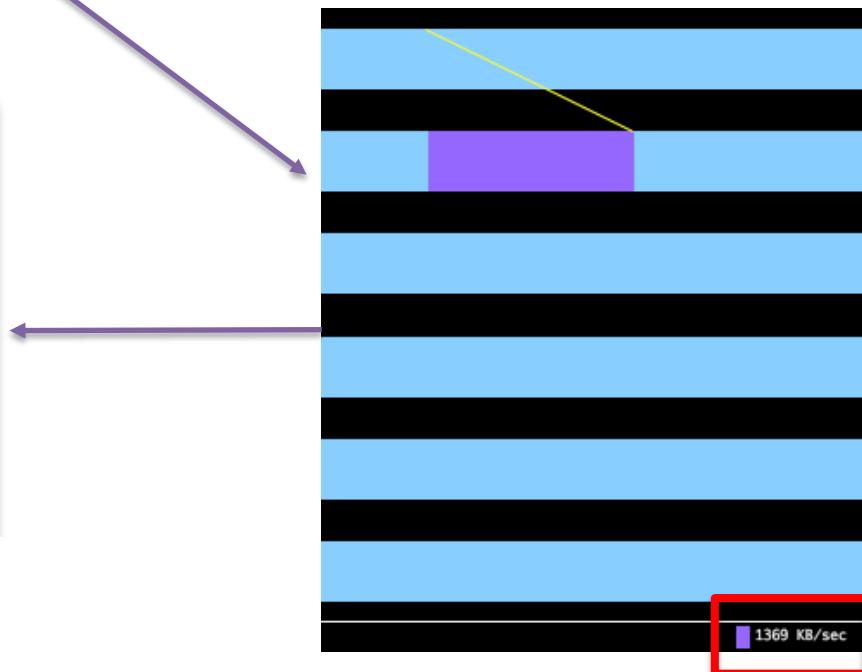


Example: Packet level communications

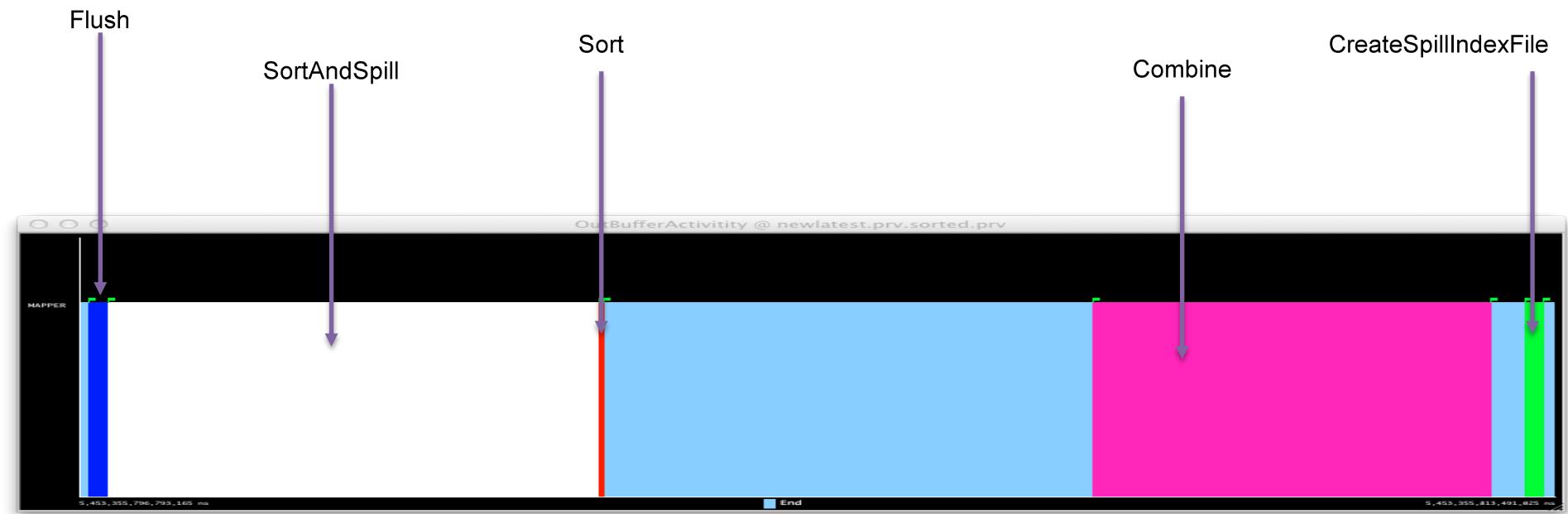


Data Sent to TT by Mappers @ newlatest.prv.sorted.prv

| | MAPPER | MAPPER | MAPPER | MAPPER | MAPPER | MAPPER |
|---------|------------|------------|------------|------------|------------|------------|
| JT | - | - | - | - | - | - |
| NN | - | - | - | - | - | - |
| SNN | - | - | - | - | - | - |
| MAPPER | - | - | - | - | - | - |
| REDUCER | - | - | - | - | - | - |
| DN | - | - | - | - | - | - |
| MAPPER | - | - | - | - | - | - |
| TT | 792,866.50 | 453,508.20 | 450,403.59 | 528,267.08 | 457,878.90 | 484,062.35 |
| MAPPER_ | - | - | - | - | - | - |
| MAPPER_ | - | - | - | - | - | - |
| MAPPER_ | - | - | - | - | - | - |
| MAPPER_ | - | - | - | - | - | - |
| REDUCER | - | - | - | - | - | - |
| Total | 792,866.50 | 453,508.20 | 450,403.59 | 528,267.08 | 457,878.90 | 484,062.35 |



Example: Low-level Hadoop events

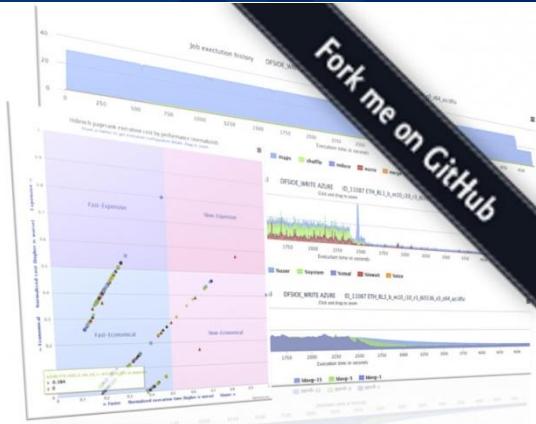




PROJECT DISSEMINATION, REFERENCE, AND CONCLUDING REMARKS

Extending and collaborating in ALOJA

1. Install prerequisites
 - vagrant
2. git clone <https://github.com/Aloja/aloja.git>
3. cd aloja/vagrant
4. vagrant up
5. Open your browser at: <http://localhost:8080>



Community engagement and PR

Recurrent

- Blogging on Big Data
- Presentations at local IT Big Data meetup events

Upcoming

- Cloudscape DEMO
- Strata EU, Hadoop Summit (TBC)

Research Center

- Barcelona Big Data Center of Excellence
- ... seeking EU funding

In talks/tests with Cloud providers to expand search

- Rackspace, Amazon, Google...

The screenshot shows a news article titled "BSC-Microsoft Research Center announces launch of project to optimize the performance of Big Data infrastructure". The article discusses the launch of Project Aloja, a joint research project between BSC and Microsoft Research. It highlights the goal of providing automated optimization for Hadoop infrastructure deployments. The page includes a sidebar with navigation links like "Home", "About BSC", "Computer Sciences", "Earth Sciences", etc., and a photo of several researchers.

The screenshot shows a blog post titled "Unravelling Hadoop Performance Mysteries" by Alex Woodie. The post discusses the complexity of Hadoop and the challenges of tuning it. It features a large "hadoop" logo. Below the post, there's a snippet of text about predicting Hadoop performance. To the right, there's a sidebar for "HADOOP PERFORMANCE REPOSITORY @BSC" which includes sections for "WORKING ON HADOOP PERFORMANCE FROM BARCELONA SINCE 2006", "BIGDATA MEETUP IN BARCELONA", and "NEW PAPER ACCEPTED ABOUT HADOOP SCHEDULING!". The sidebar also lists "RECENT POSTS" and "SPONSORS".



Additional reference and publications

« Online repository and tools available at:

- <http://hadoop.bsc.es>

« Publications: <http://hadoop.bsc.es/publications>

- Project description on:

- "ALOJA: a Systematic Study of Hadoop Deployment Variables to Enable Automated Characterization of Cost-Effectiveness"

- Upcoming:

- ALOJA-ML for KDD15'
 1. ALOJA-ML: A first dive into Hadoop behavior using Machine Learning
 - Working on:
 2. The Economics of Hadoop in the Cloud
 - » An evaluation of cost-effectiveness of Hadoop in the Cloud
 3. Cluster performance Characterization for Big Data
 - » A performance modeling comparison of Hadoop in different cluster sizes and OS configurations

Concluding remarks

- « The early findings of the project already show significant value in understanding Hadoop's runtime
 - for optimizing executions times
 - understanding the cost-effectiveness of different configuration and deployment options

- « Our intent is that researchers and organizations evaluating ~~Hadoop~~ the Hadoop stack will benefit
 - from this growing database of performance results and configuration guidance

BSCMSCR Team



Microsoft Research
Centre

« ALOJA Team members:

- « David Carrera Senior Researcher, Barcelona Super Computing Center (BSC)
Associate Professor, Universitat Politecnica de Catalunya (UPC)
- « Nicolas Poggi Post Doctorate Researcher, BSC
- « Aaron Call Research support engineer, BSC
- « Josep Lluis Berral Post Doctorate Researcher, BSC/UPC
- « Josep Cugat Research support engineer, BSC
- « Fabrizio Gagliardi Senior Strategy Consultant, BSC
Distinguished Research Director, UPC
Chairman, ACM Europe Council
- « Jordi Torres Research Manager, BSC
Professor, UPC
- « Rob Reinauer Partner Systems Architect, Microsoft SQL Server
- « Jose Blakeley Partner Software Architect, Microsoft SQL Server
- « Nikola Vujic Software Development Engineer, Microsoft HDInsight
- « Daron Green Sr. Director, Regional Research, Microsoft Research
- « J. Eduardo Campos Director, Business Strategy, Microsoft Emerging Markets



Thanks!

Q&A