POLITECNICO DI MILANO





PoliCloud
Architecture & Applications

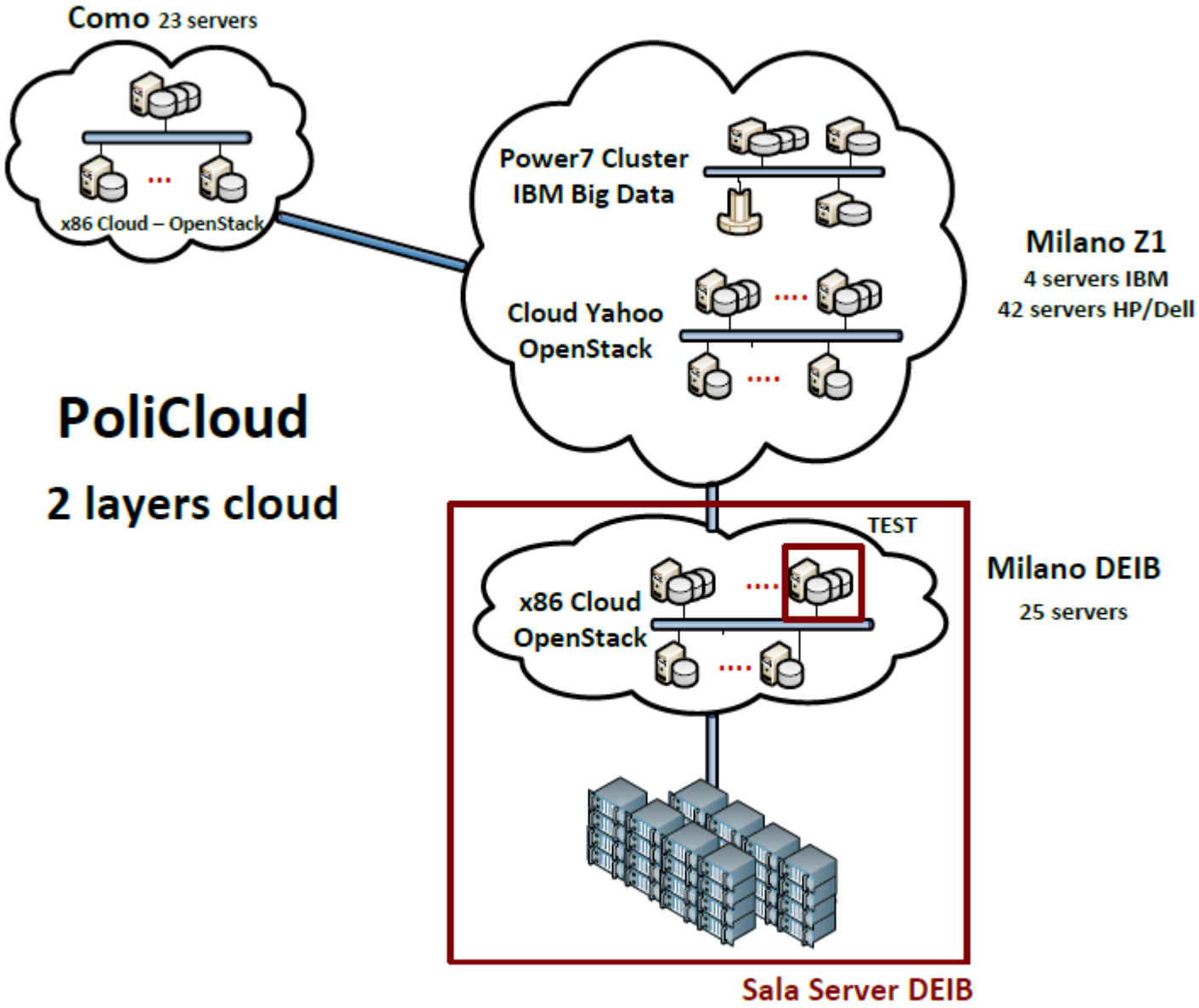
G. Serazzi

May 7, 2015

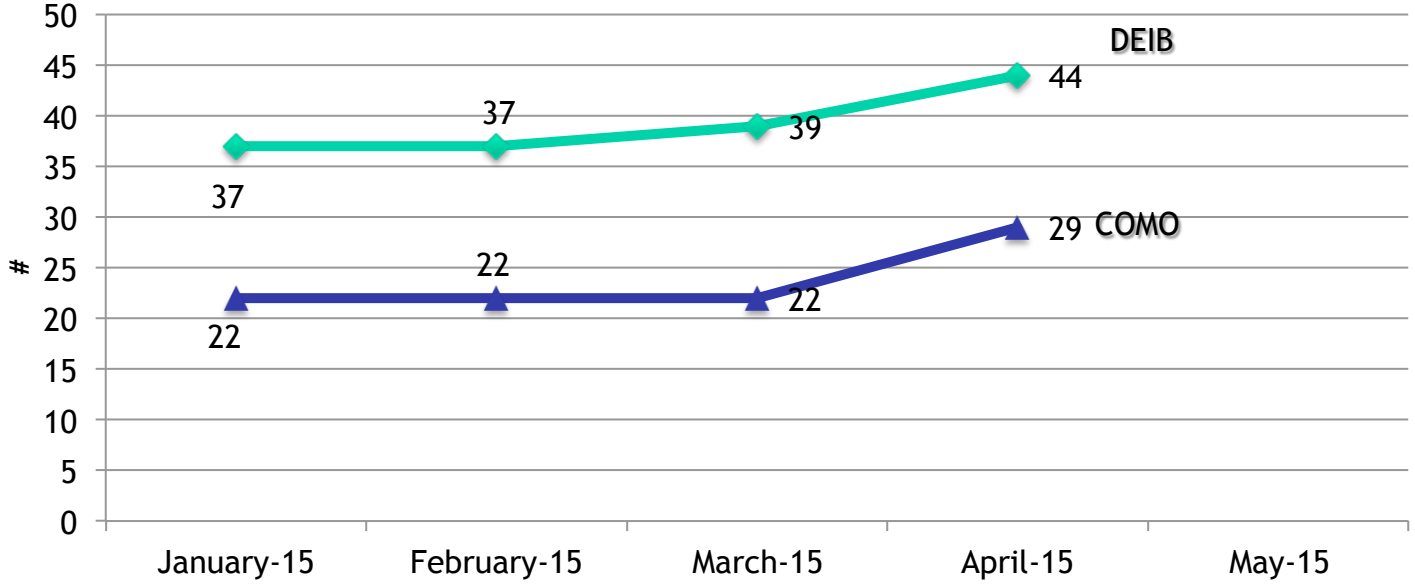
Outline

- PoliCloud: Architecture & Applications. G. Serazzi
- Cloud architecture and its evolution: A. Frossi
- BigData architecture and its evolution: M. Marchente
- Performance profiling of Big Data applications: D. Ardagna
- PoliCloud App: call traffic visualizer: R. Pagano
- PoliCloud and MODAClouds: E. Di Nitto
- PoliCloud Apps: R. Fedorov
- Energy optimization and storage reliability: M. Gribaudo

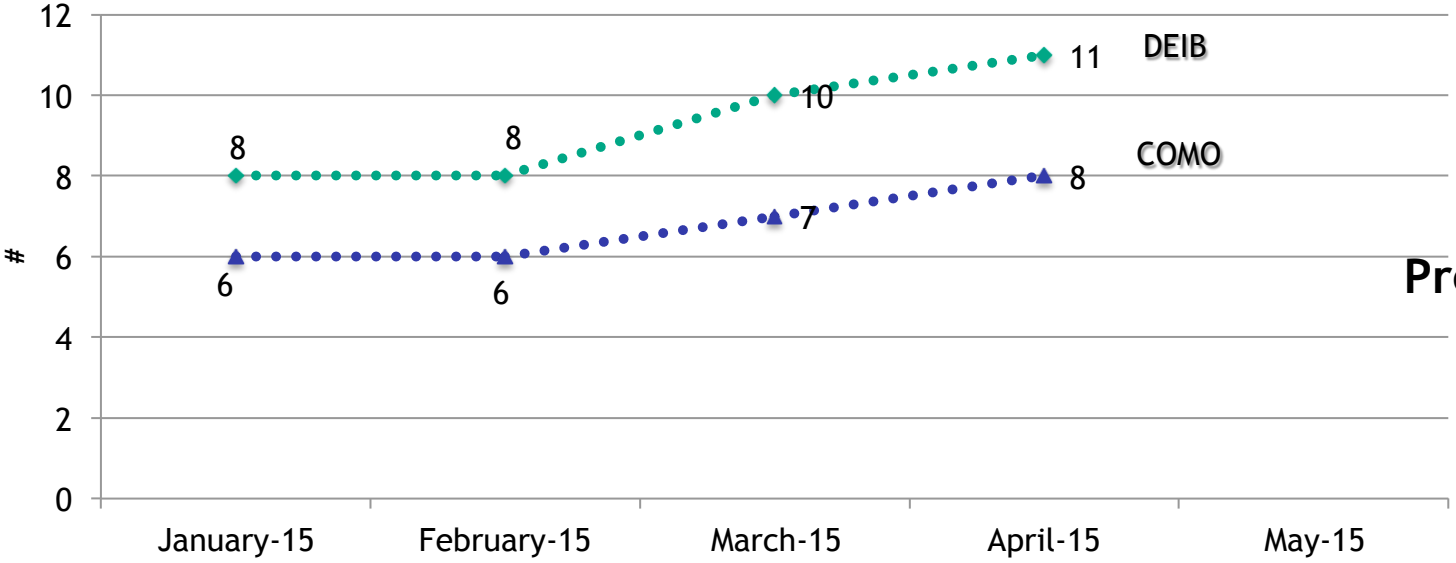
the PoliCloud structure



Users and Projects



Users



Projects

Alessandro Frossi

07/05/2015

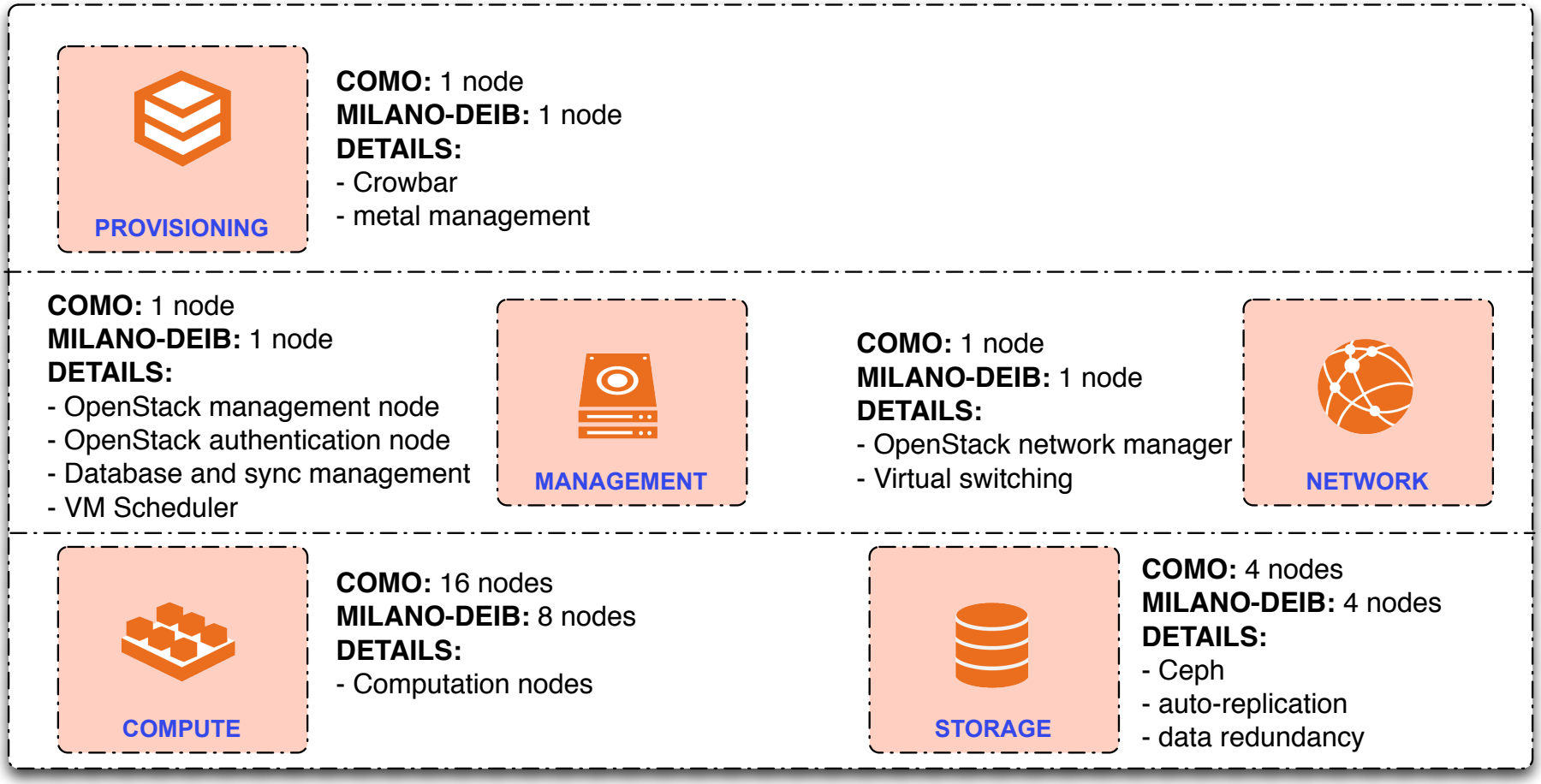
POLITECNICO DI MILANO





Cloud Architecture and its Evolution

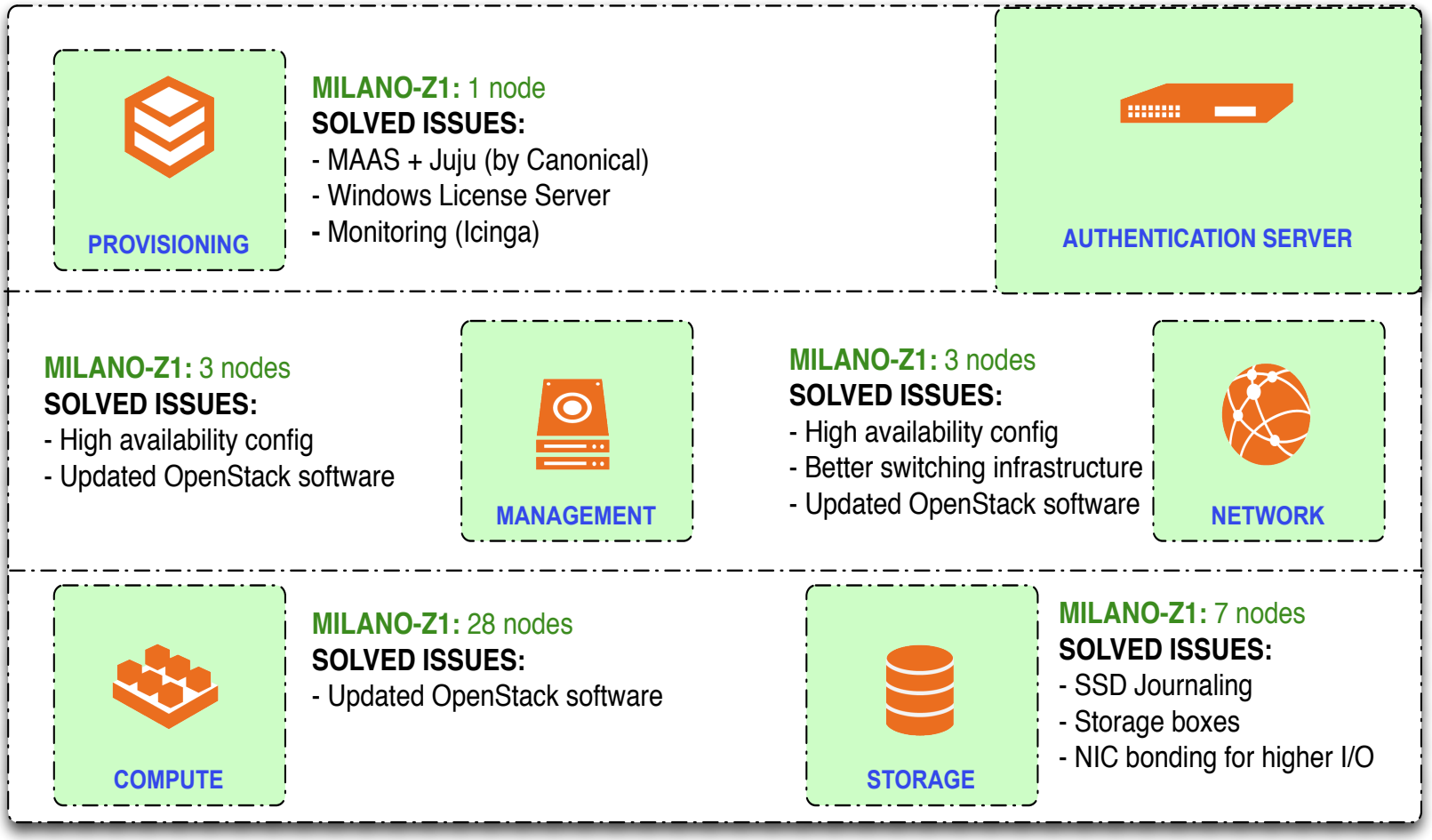
Current Cloud Architecture



Pros and Cons

PRO	CONS
- Fast deployment - Centralized access management	- Dell discontinued support - Limited customizability
- It works as expected	- Single Point of Failure - Out of date
- Network works	- Single Point of Failure - Out of date - Limited scalability
- Computation works	- Out of date
- Great scalability	- No SSD journaling - Limited throughput

Future Cloud Architecture



Numbers and Expectations

Ambient	Nodes	Cores	RAM	Storage	VMs
MILANO-DEIB	15 -> 19	88	250GB	18TB	~20[1]
COMO	23 -> 27	128	330GB	26TB	~20[1]
MILANO-Z1	42	304	624GB	162TB	~100[2]

[1] Mean number of VMs active last month
[2] Estimated number of medium size VMs in the system

Future projects:

- Shibboleth integration
- Automatic hosting platform (Heroku-style)
- Repository platform (GitLab style)

BigData environment at PoliCloud Lab.

Mario Marchente

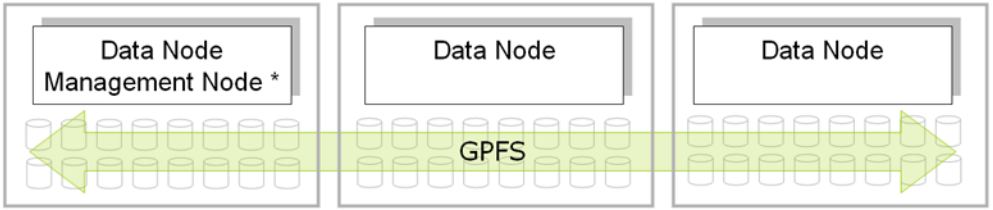
DEIB - Politecnico di Milano

07/05/2015

BigData Infrastructure

IT Infrastructure for IBM InfoSphere BigInsights and DB2 on Power processor

A 3-nodes Cluster for Infosphere BigInsight on PowerLinux



Sistema IBM Power Linux 7R2

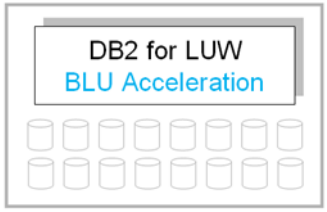
- 16 core POWER7+ 4.2 Ghz / 128 GB RAM
- Linux RedHat EL 6

Storage System Expansion

* The following management functions need to be hosted

- ✓ Hadoop JobTracker
- ✓ Hadoop NameNode
- ✓ Hadoop Secondary NameNode
- ✓ BigInsights Console
- ✓ NameNodes not necessary if using GPFS file system in place of Hadoop HDFS file system

Single DB2 node for "In-memory computing"



Sistema IBM Power 730

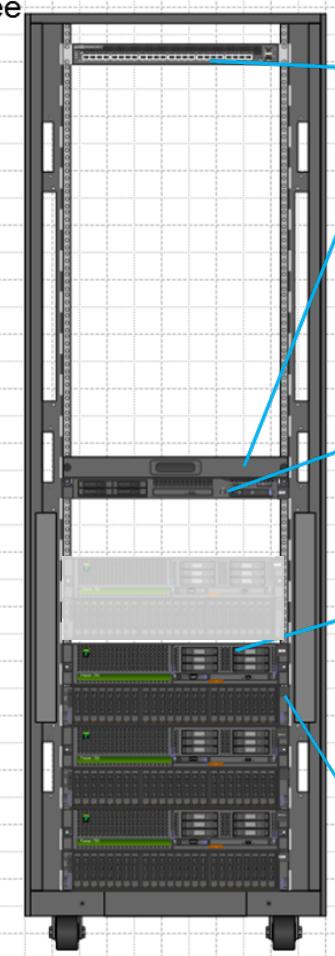
- 8 core POWER7+ 4.3 Ghz / 512 GB RAM
- AIX 7.1

Storage System Expansion

BigData Hw - BigInsights cluster

7014-T42 Rack (42 U)

- 19 U occupied
- 23 U free



Switch Ethernet

Server

x3

Storage Sub-System

x3

Lab Infrastructure: Physical lay-out (BigInsight Systems)

1455-24E Switch Ethernet - 24 port 10 Gb/s

- Active ports:
  - 8 sfp 10 Gb/s SX
  - 8 sfp 1 Gb/s TX

1U

7316-TF3 Flat Panel Console

1U

7042-CR7 Hardware Management Console

1U

8246-L2T IBM Power Linux 7R2

- 16 core POWER7+ 4.2 Ghz
- 128 GB RAM
- 1 Ethernet 4 porte 1 Gb/s Tx (RJ45)
- 1 Ethernet 2 porte 10 Gb/s Sx (Fibre)
- 6 disks 600 GB 10k rpm
- RedHat EL 6.5
- IBM POWERVM per Linux
- IBM GPFS

6U

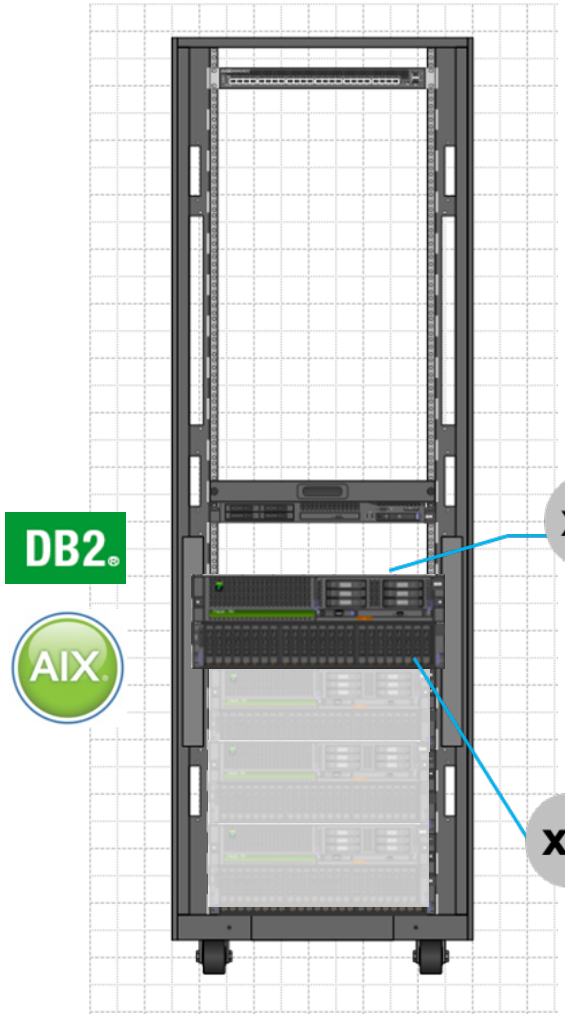
EL1S EXP24S SAS SFF Gen 2

- 16 SAS disks 600 GB/s 10k rpm

6U

BigData Hw – DB2 Server

Lab Infrastructure: Physical lay-out (DB2 LUW with BLU Acceleration)



x1 **Server**

8231-E2D IBM Power 730

- 8 core POWER7+ 4.3 Ghz
- 512 GB RAM
- 2 Ethernet 4 port 1 Gb/s Tx (RJ45)
- 1 Ethernet 2 port 10 Gb/s Sx (Fibre)
- 2 Fibre Channel 2 port 8Gb/s
- 6 disks 300 GB 15k rpm
- AIX 7.1
- IBM POWERVM EE for small server

2U

x1 **Storage Sub-System**

EL1S EXP24S SAS SFF Gen 2

- 24 SAS disks 300 GB 15k rpm

2U

BigData SW Configuration

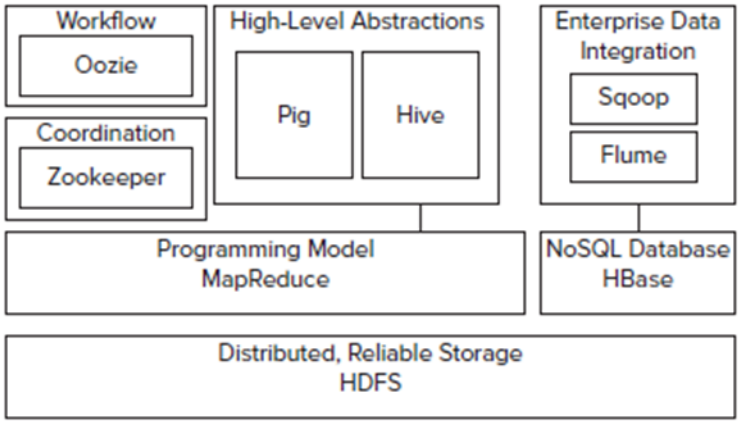
Now:

- IBM Infosphere BigInsights Enterpise Edition 2.1.0.1 (General Parallel File System – File Placement Optimizer (GPFS-FPO), Hadoop, MapReduce) – Linux RedHat 6.5.
- Data Integration.
- DB2 10.5 with BLU Accelerator – AIX 7.1 .

Next:

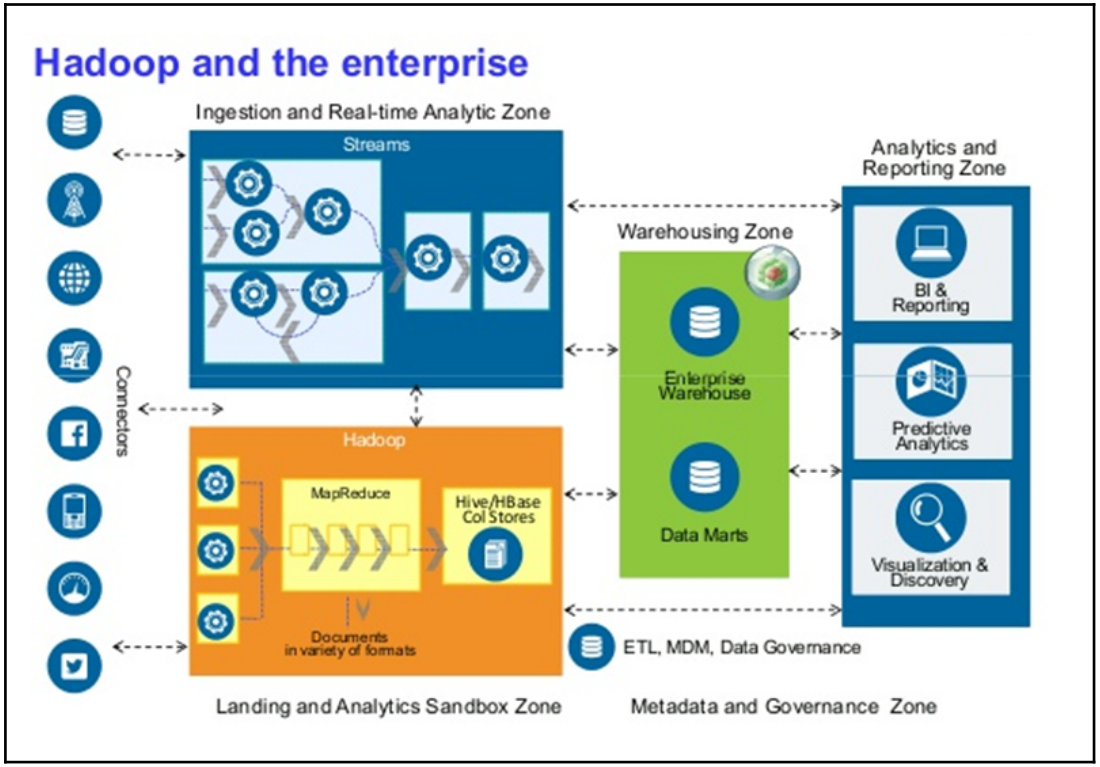
- IBM Infosphere BigInsights Enterpise Edition 3.0.2 (before) and then 4.0 .
- Cognos / SPSS.
- Ilog Cplex.

Apache Hadoop

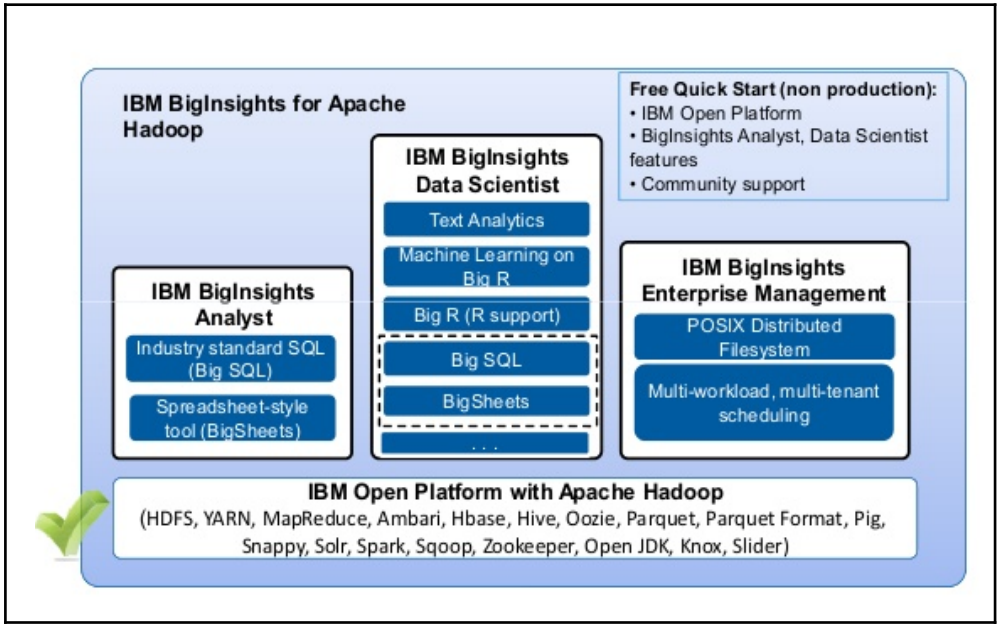


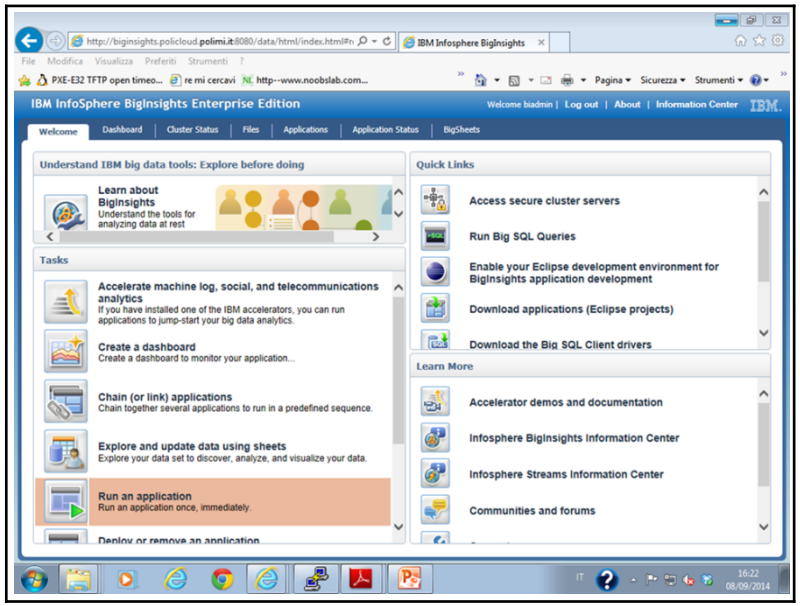
A project to develop open-source software library framework for reliable, scalable, distributed computing, that:

- allows for the distributed processing of large data sets across clusters of computers using simple programming models.
- is designed to scale up from single servers to thousands of machines, each offering local computation and storage.
- delivers a highly-available service detecting and handling failures at the application layer.



IBM Infosphere BigInsights Enterpise Edition



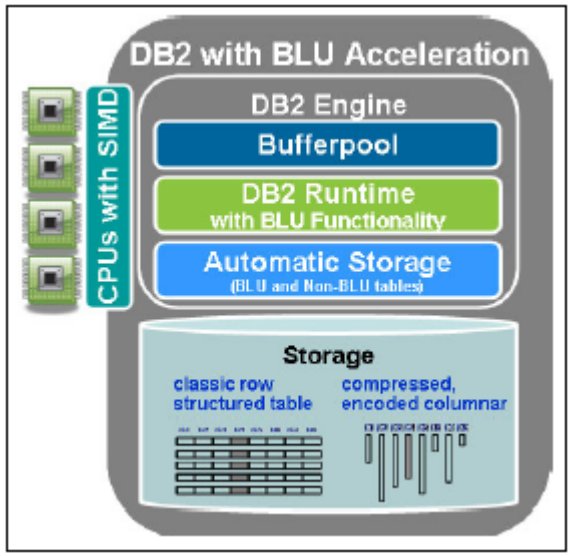


Ambari	1.7	Apache License, Version 2.0
Avro	1.7.7	Apache License, Version 2.0
Flume	1.5.2	Apache License, Version 2.0
Hadoop	2.6.0	Apache License, Version 2.0
Hbase	0.98.8	Apache License, Version 2.0
Hive	0.14.0	Apache License, Version 2.0
Knox	0.5.0	Apache License, Version 2.0
Oozie	4.1.0	Apache License, Version 2.0
Parquet-MR / Format	1.5.0 / 2.1	Apache License, Version 2.0
Pig	0.14.0	Apache License, Version 2.0
Slider	0.60.0	Apache License, Version 2.0
Solr	4.10.3	Apache License, Version 2.0
Spark	1.2.1	Apache License, Version 2.0
Sqoop	1.4.5	Apache License, Version 2.0
Zookeeper	3.4.6	Apache License, Version 2.0

IBM Db2 with BLU Acceleration

DB2 for Linux/Unix/Windows (LUW) is a database management server product developed by IBM. The last version 10.5 "offers accelerated analytic processing by introducing a new processing paradigm and data format within the DB2 database product". This version fully integrates the IBM BLU Acceleration collection technologies created by the IBM Research and Development Labs for analytical database workloads. The BLU Acceleration capabilities consist of:

- in-memory processing of columnar data,
- the ability to "act on the data while still compressed", also called "actionable compression" (which uses approximate Huffman encoding to compress and pack data tightly),
- cpu acceleration (which exploits Single Instruction Multiple Data (SIMD) technology and provides parallel vector processing)
- data skipping (data not useful to the current active workload are ignored).



Performance profiling of Big Data Applications

Danilo Ardagna

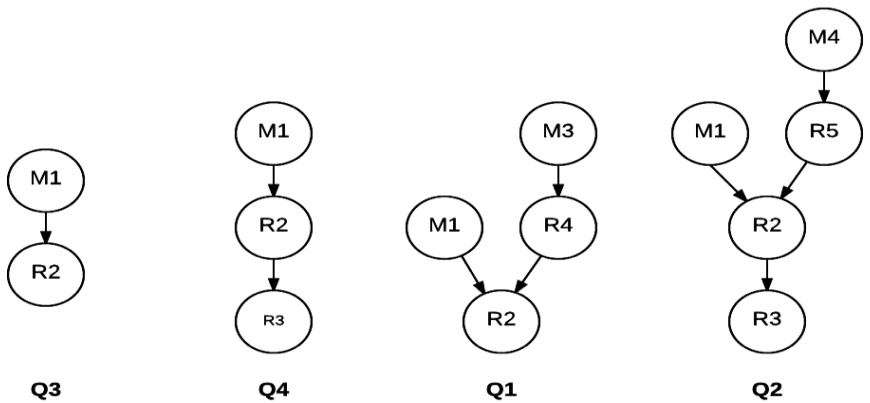
Research goals

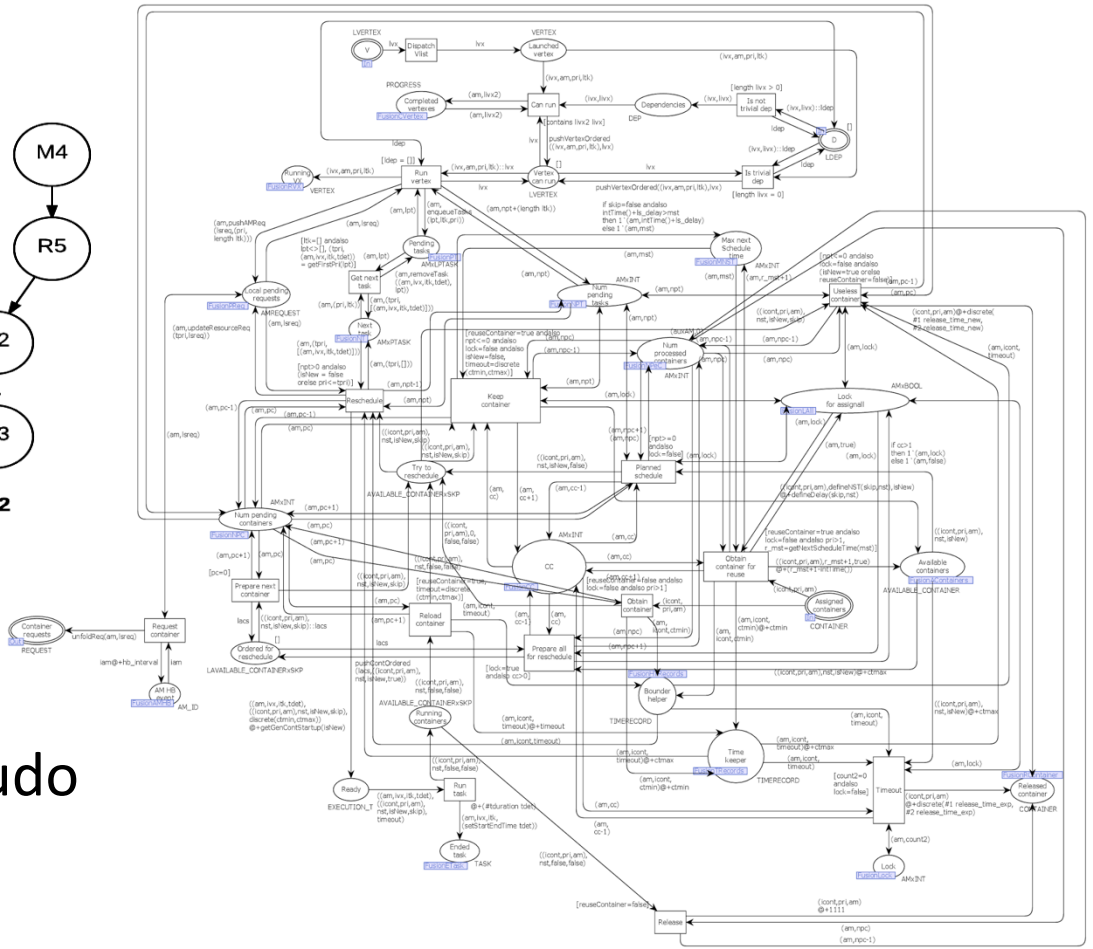
- Big Data Frameworks:
  - Hadoop Map Reduce: “Legacy” for large volume unstructured data
  - Spark: New “Map Reduce” up to 100x faster than Hadoop

- Identify performance models and develop techniques for capacity planning

Hadoop performance modelling

- Colored Petri Nets
- TPC-DS (3-40-250-500GB) benchmark, 4 ad-hoc Hive-Tez queries:





Joint work with Marco Gribaudo

Colored Petri Nets Results

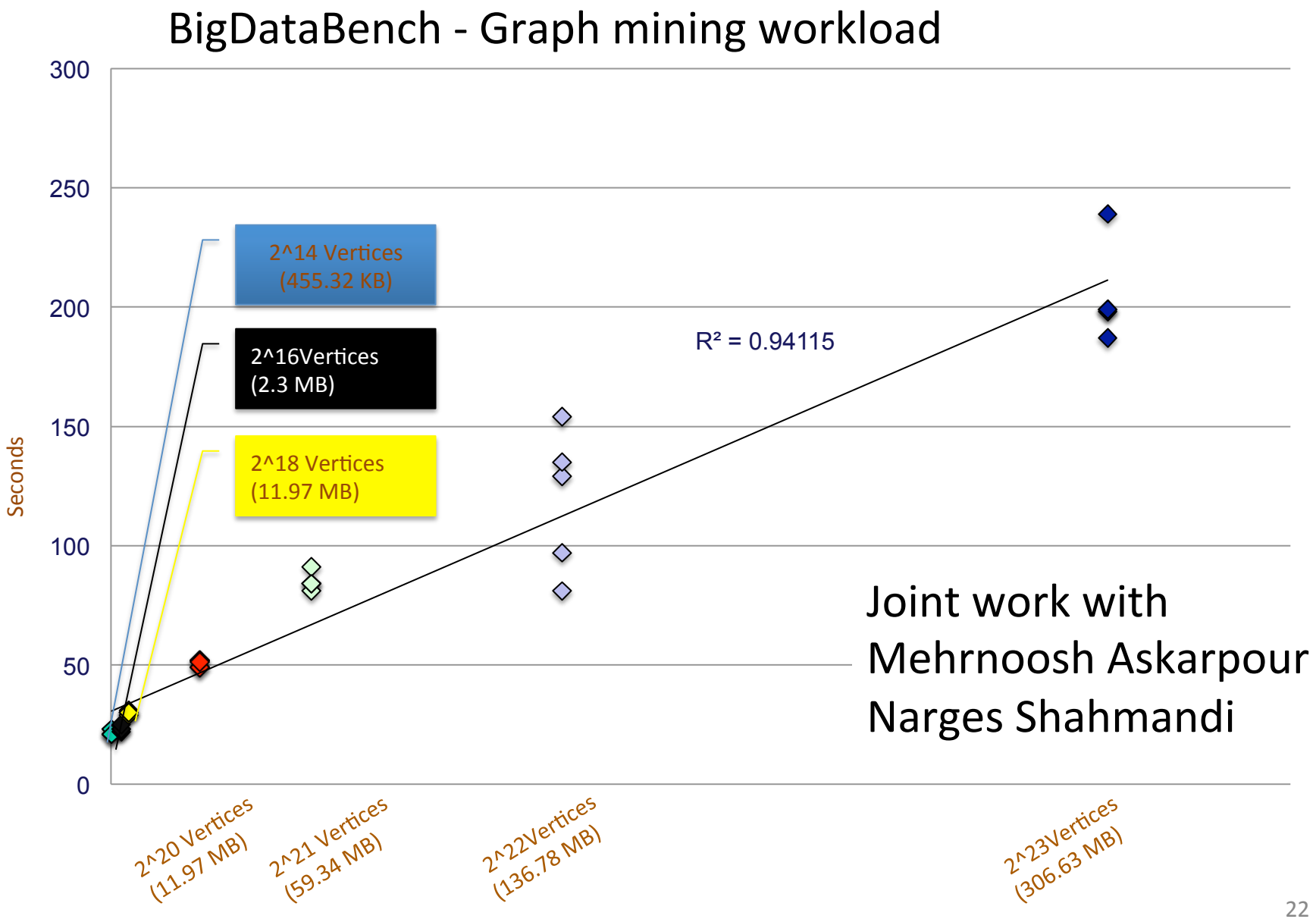
Single queries

Test case	Min error	Max error	Avg error
3GB - 5 nodes	1%	9%	3.5%
40GB - 5 nodes	5%	15%	10%
40GB - 30 nodes	7%	21%	12.5%
250GB - 30 nodes	2%	7%	5.3%
500GB - 30 nodes	2%	16%	8.6%

Concurrently running queries

Test case	Min error	Max error	Avg error
40GB - 30 nodes	8%	31%	17.25%
250GB - 30 nodes	12%	30%	18%
500GB - 30 nodes	24%	31%	26.3%

Spark job profiling results



- DICE: **D**eveloping Data-**I**ntensive **C**loud Applications with Iterative Quality **E**nhancements
- H2020 Research & Innovation Actions (small) n. 644869, ≈4MLN EU contribution
- Febru
- Collea Di Nitto, Matte

Define a **quality-driven** framework for **data-intensive** applications leveraging Big Data technologies in **private** or **public clouds**:

- A novel profile and for data-aware application development
- Quality driven tools
- Continuous testing and deployment

Im
Lo

exiant™
your cloud simplified

Universidad Zaragoza
1542

prodevelop
Integración de tecnologías

ATC
ATHENS TECHNOLOGY CENTER

European Commission

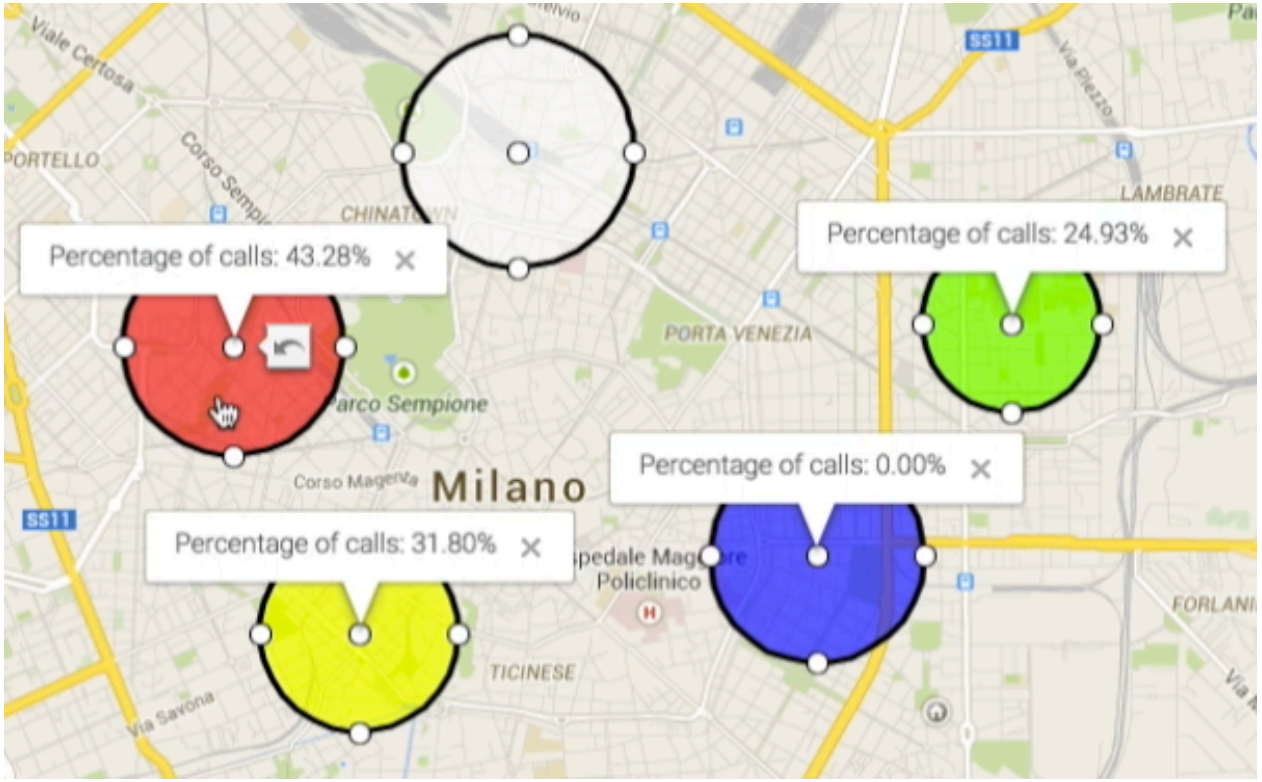
PoliCloud App: Call Traffic Visualizer

Paolo Cremonesi
Roberto Pagano
Mona Nasery
Massimo Quadrana

Project for Big Data PhD course

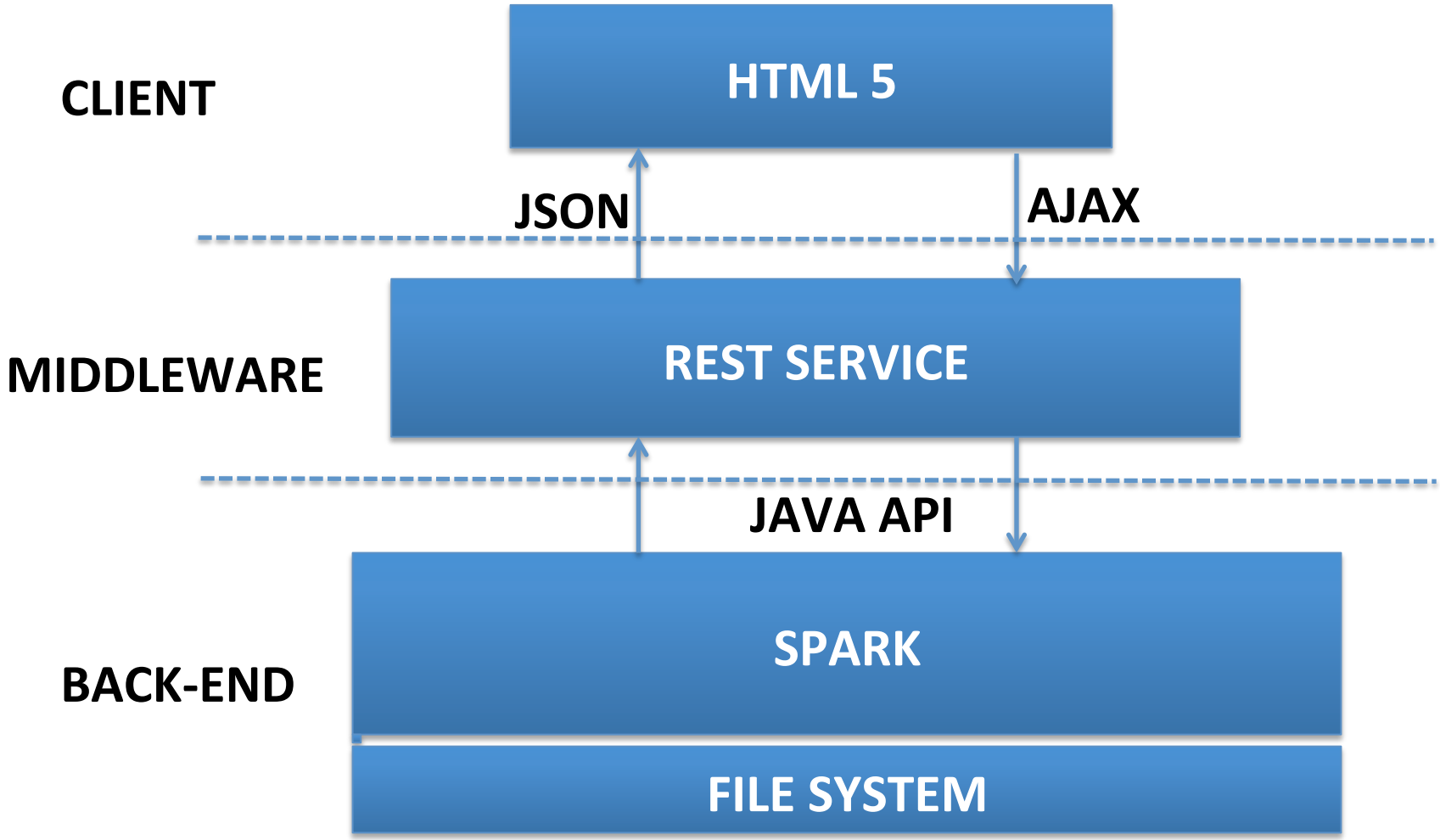
The Idea

To provide a visualizing tool to show traffic statistics of calls in different areas selected by users in Milano city in real time.

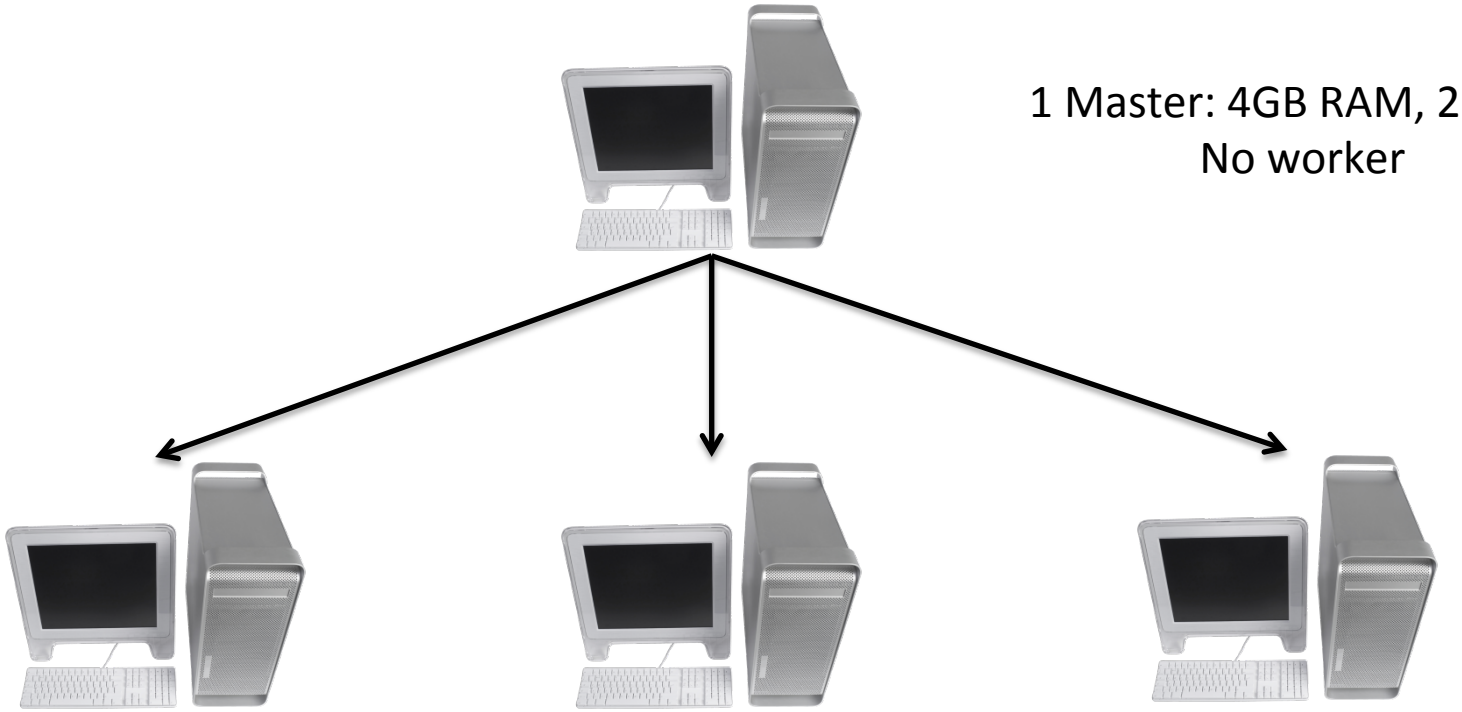


Dataset: Telecom Italia Challenge 2014

Architecture



SPARK Cluster Architecture



1 Master: 4GB RAM, 2 cores
No worker

3 Slaves: each 8GB RAM, 4 cores, Workers

Deployed on PoliCloud

MODAClouds (Improving Agility of Cloud Ecosystems) and PoliCloud

Danilo Ardagna, Michele Ciavotta, Elisabetta Di Nitto, Giovanni Gibilisco, Marco Miglierina, Marco Scavuzzo

http://multiclouddevops.com

http://www.modaclouds.eu

MODAClouds objectives

Multi-Cloud Dev&Ops Management

increase

Deployment speed

Flexibility

Elasticity/Adaptability

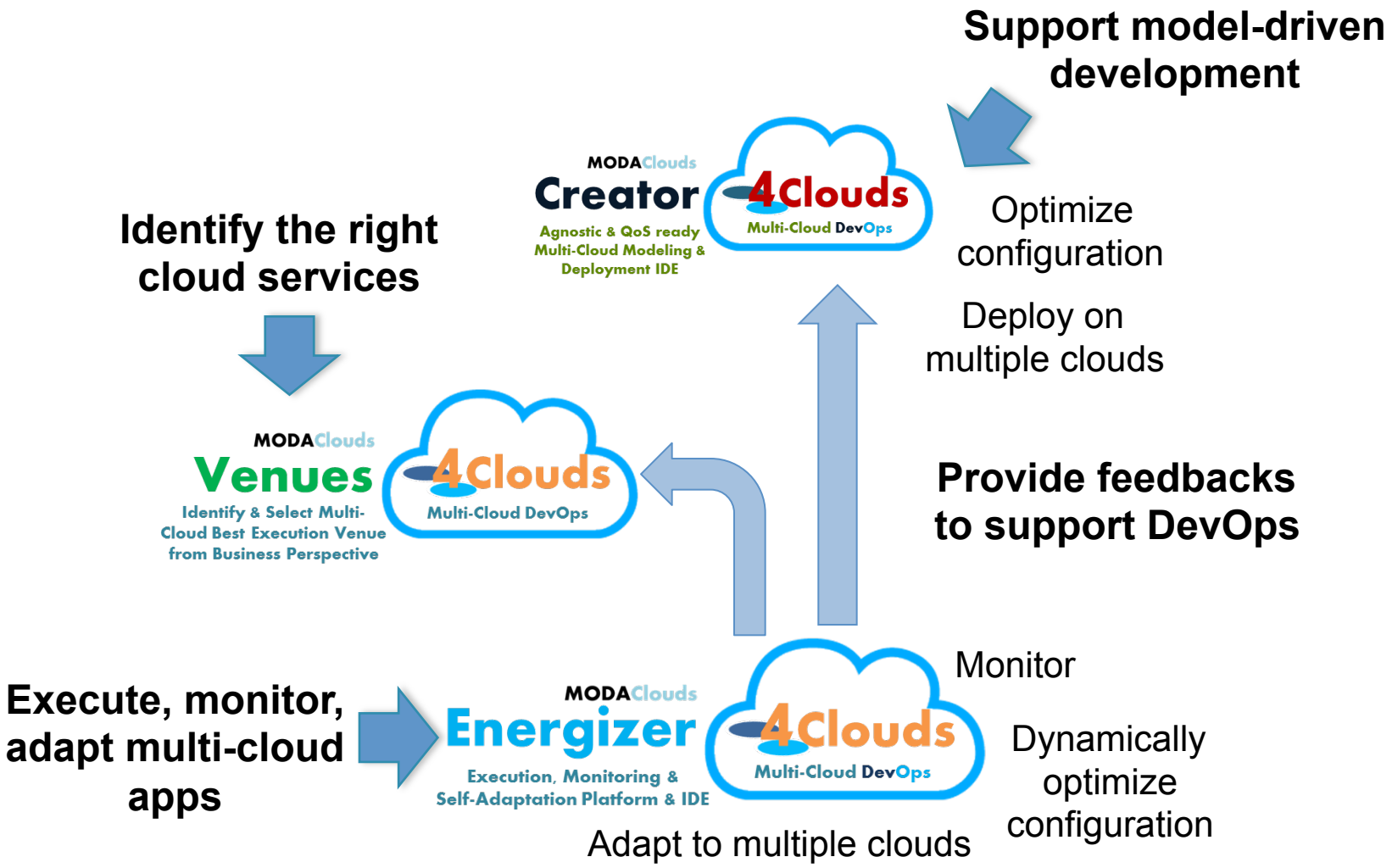
Set of services to support dev

reduce

Lock-in

Unpredictability of performance

MODAClouds OS assets



How MODAClouds exploits PoliCloud

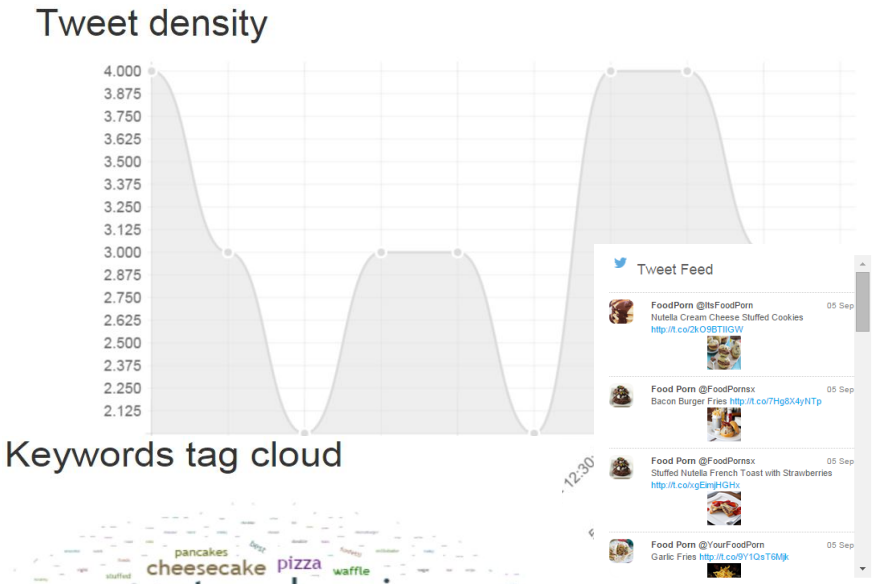
- Use it as a testing environment
- Make our tools available as a Service on PoliCloud
  - Venues4Clouds to support the cloud services selection process
  - Hegira4Clouds to support data migration
- Understand if we could exploit some of our tools to enrich the management mechanisms offered by PoliCloud
  - Tower4Clouds to support application monitoring

PoliCloud Apps

Fraternali, Fedorov, Ciceri, Bislimovska, Pasini, Baroffio, Bernaschina, Castelletti, Cominola, Giuliani

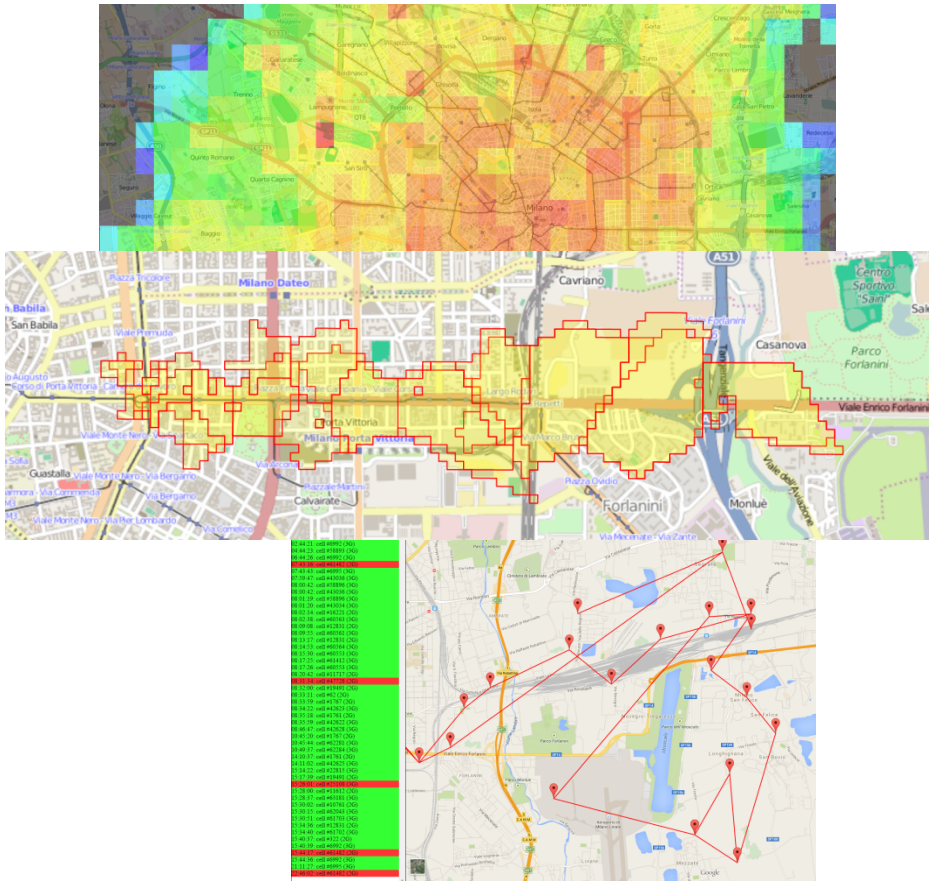
TwitterAnalytics

- Project: http://www.proactiveproject.eu/, http://smarth2o-fp7.eu/
- Data: tweets and associated images
- Goal: finding social network influencers based on text and image content
- Use of cloud: massive crawling, tweet and image classification, influence estimation and prediction. Over 2.600.000 processed tweets and 600 GB of used storage (~60k tweets/day)



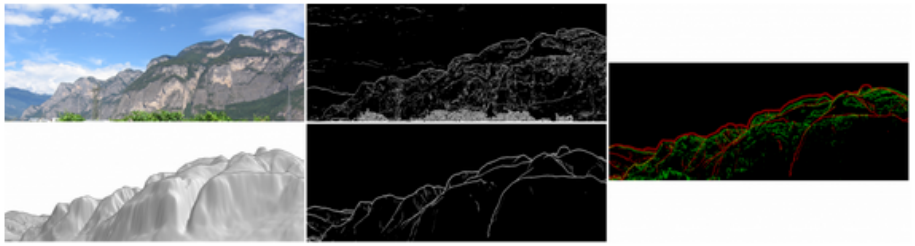
PeopleOnTheMove

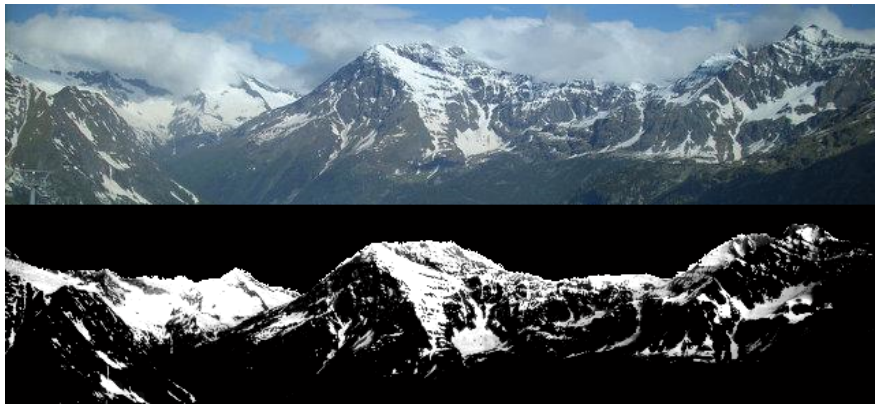
- Project: http://www.proactiveproject.eu/
- Data: cellular phone traces, metropolitan area of Milan
- Goal: predicting the presence of people in space and time
- Use of cloud: data collection and processing.
- ~250.000.000 of cellular events every day



SnowWatch

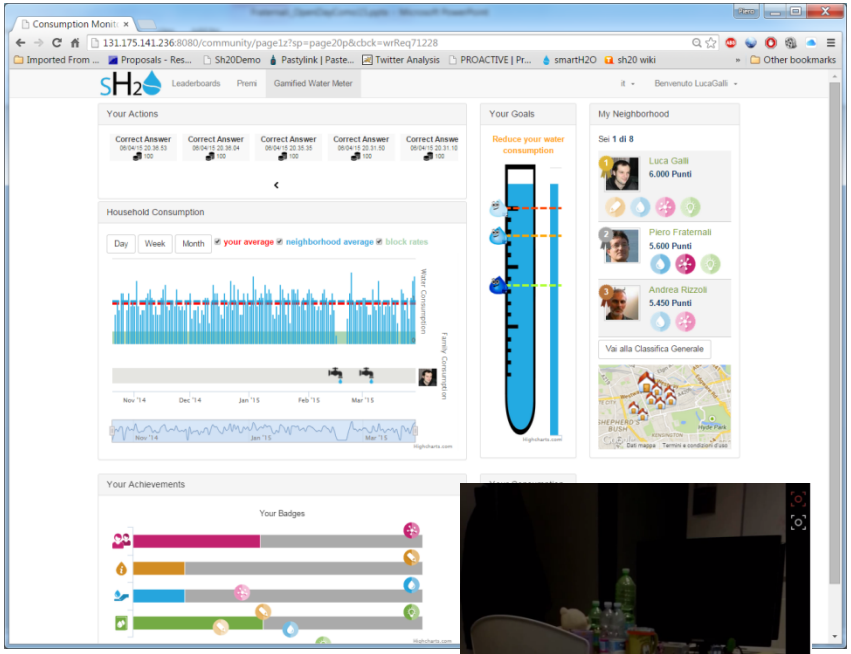
- Project: http://www.proactiveproject.eu/
- Data: images crawled from social media and web cams
- Goal: analyzing the status of snow cover to predict water availability
- Use of cloud: massive image crawling and storage, image processing (classification. peak recognition, snow extraction) predictive modeling.
- ~2.000 webcams crawled at maximum frequency, 10GB of data downloaded and stored daily





SmartH20

- Project:http://smarth2o-fp7.eu/
- Data: smart meter readings, user generated content, game play traces
- Goal: monitoring user's actions and recommend sustainable behaviors
- Use of cloud: IoT integration, data fusion, user modeling, demand prediction, gamification of actions, action recommendation

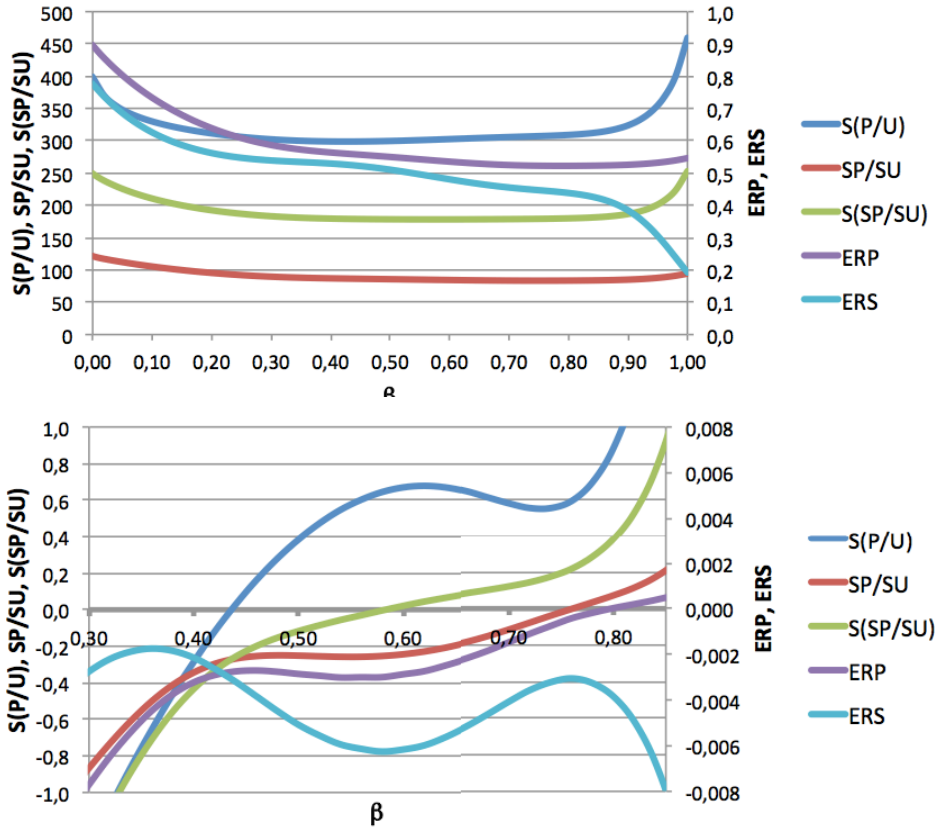


Energy optimization and storage reliability

Marco Gribaudo, Giuseppe Serazzi,
Davide Cerotti, Pietro Piazzolla,
Riccardo Pinciroli

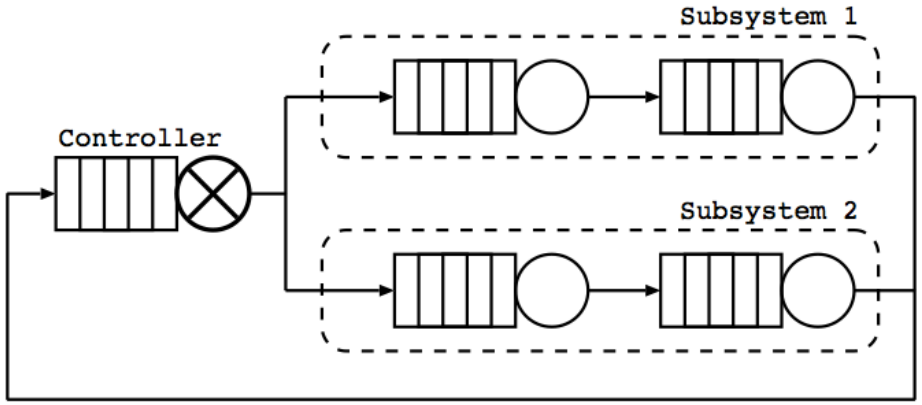
Energy optimization metrics

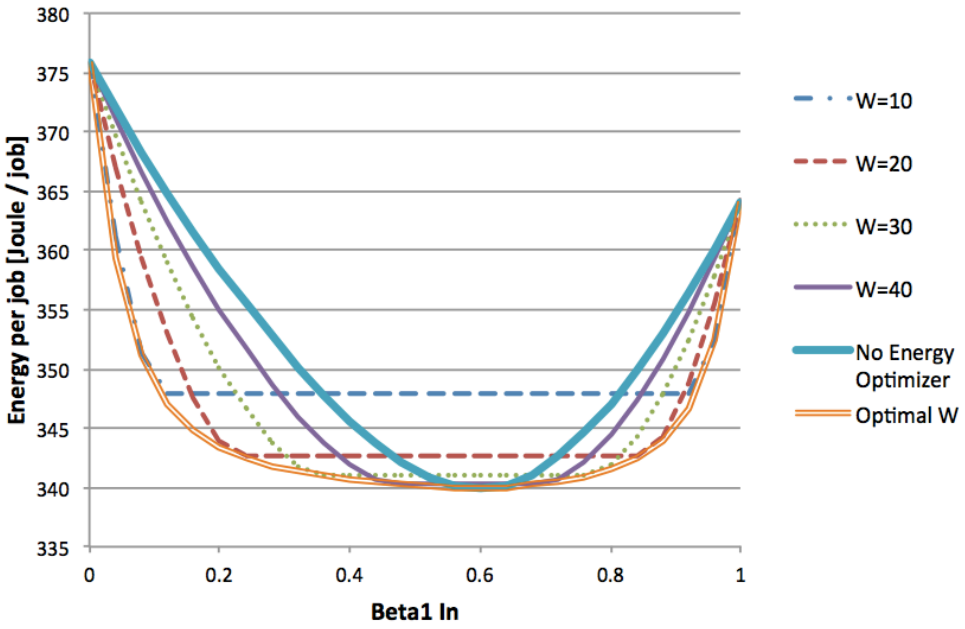
- Evaluating the cost / benefit tradeoff is well defined in single class workloads.
- A proper measure is no longer straight-forward in multi-class models.
- We have several alternatives: we want to test them in the PoliCloud environment.



Class-aware scheduling

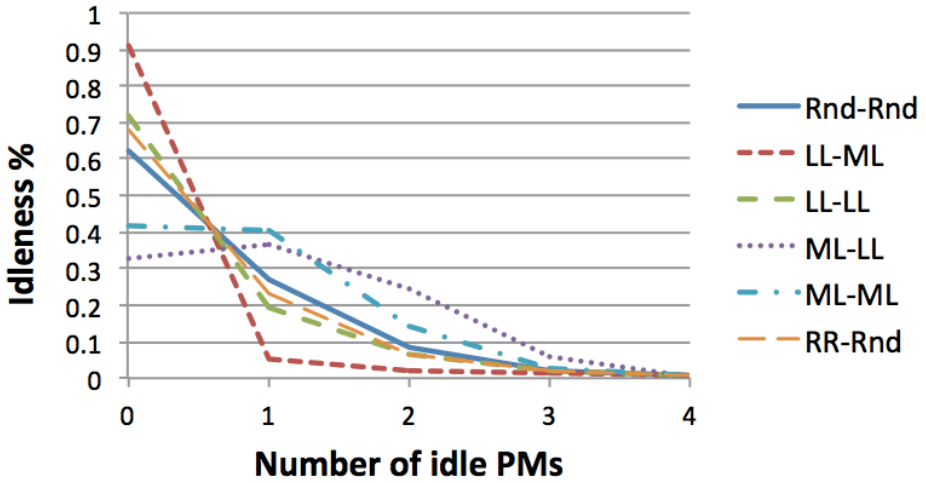
- When different workload are mixed on a set of servers, power consumption can be reduced by proper task assignment.
- We have developed new class-aware policies.
- We want to test them on the PoliCloud servers.





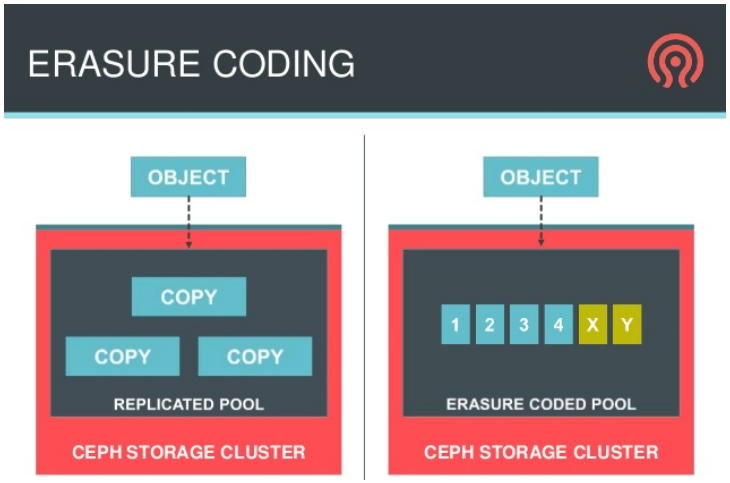
Increasing the probability of switching off servers.

- When considering replicated servers with auto-scaling features, allocation and de-allocation policies should be defined.
- Proper policies can increase the probability of having machines that can be switched off.
- We have studied several alternatives.
- We plan to test them on the PoliCloud.



Reliability in cloud storage

- Cloud storage systems like CEPH can exploit both replication and erasure coding to protect data while minimizing the required space, and without penalizing the performances.
- The combination of both techniques can increase the system reliability.
- We plan to test several possible configuration in the PoliCloud environment.



k	p	g	h	MTTF [hours]	Targ. occ. [blocks]	Av. max. occ. [blocks]	Av. tr. time [msec.]
1	0	0	0	102	5	5	100
2	0	0	0	598	10	15, 8	88.2
3	0	0	0	2283	15	30.5	81.3
4	0	0	0	6317	20	45.4	76.8
1	1	4	1	320	7	8	100.8
1	1	4	2	999	8	11.9	100.7
1	1	4	8	359531	15	22.6	99.6
2	1	4	2	16287	13	27	88.4
3	1	4	2	82531	18	39.8	81, 4
3	1	4	4	417187	20	42.8	81.46