



Sergio Traldi
INFN Padova

Big Data

Mesos, Spark, Hadoop

Cosa si intende per Big Data

- Il termine **big data** (**grandi dati** o **grossi dati**) è fuorviante, farebbe pensare all'enorme quantità di dati oggi disponibili;
- per “rivoluzione Big Data” si intendono le opportunità oggi disponibili di avere così tante informazioni al servizio business;
- esistono settori dove i dati, per quanto ve ne siano davvero in ingenti quantità, non sono sempre disponibili a tutti e, soprattutto, non vengono sempre condivisi;

IL VERO SIGNIFICATO:

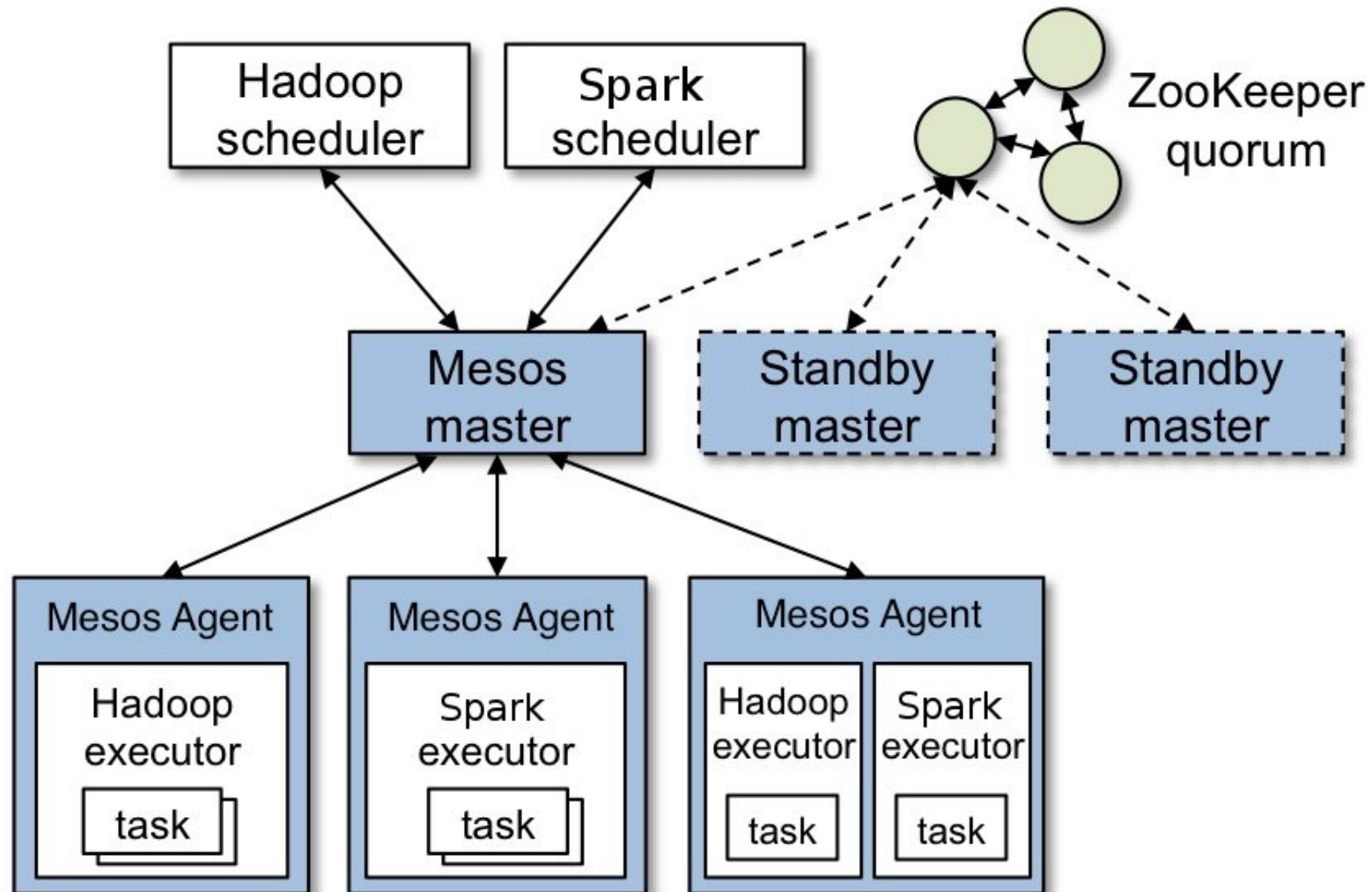
- la capacità di usare tutte queste informazioni per elaborare, analizzare e trovare riscontri oggettivi su diverse tematiche.

- Il modello di analisi di CMS prevede l'utilizzo di ROOT per archiviare, analizzare e trasformare i dati provenienti dal rivelatore sia offline che durante l'esecuzione di un "run";
- L'idea è di convertire questi dati in un formato che o Scala o pyspark riescano a leggere;
- Esaminare le ntuple e estrarre informazioni utili e in caso utilizzare delle librerie di tipo histogram per il plot di tali risultati;
- Sfruttare algoritmi di tensorflow e machine learning per ottimizzare la ricerca degli eventi significativi;

Creazione di un testbed 1/2

- Inizialmente avevamo creato primo testbed composto da:
 - spark stand-alone
 - un spazio hdfs (hadoop) per la memorizzazione dei dati e dei risultati;
- Successivamente per rendere scalabile e più flessibile l'applicazione spark abbiamo introdotto:
 - mesos/zookeeper/marathon:
 - che permettono di avere un cluster in HA
 - da una semplice interfaccia grafica viene fatto il deploy della propria applicazioni nel cluster
 - permettono di eseguire più applicazioni differenti sul cluster e danno una priorità di esecuzione e risorse variabili secondo la richiesta.

Creazione di un testbed 2/2



Template heat

heat_template_version: 2013-05-23

parameters:

image_centos_7:

type: string

default: 5938e89b-1ff7-4caf-b090-79c949da7ba76

....

fixed_ip_nodomhs_1:

type: string

default: "10.64.12.40"

...

resources:

secgroup-bigdata_secgroup:

type: OS::Neutron::SecurityGroup

properties:

name: "secgroup-bigdata"

rules: [{"direction": ingress, "remote_ip_prefix":

0.0.0.0/0, "port_range_min": 7077,

...

nodomhs1_server_instance:

type: OS::Nova::Server

properties:

name: "nodomhs1"

...,

user_data:

....

cat > /etc/hosts << EOF

\$IP_NODE1 nodomhs1.novalocal nodomhs1

\$IP_NODE2 nodomhs2.novalocal nodomhs2

...

EOF

yum localinstall -y

http://repos.mesosphere.com/el/7/noarch/RPMS/mesosp
here-el-repo-7-1.noarch.rpm

....

yum install -y mesos marathon chronos

zookeeper-server

cat > /etc/mesos/zk << EOF

zk://

\$IP_NODO1:2181,\$IP_NODO2:2181,\$IP_NODE3:2181/

mesos

EOF

tar -zxf hadoop-2.6.0.tar.gz

source spark-2.1.0-bin-hadoop2.6/conf/spark-env.sh

...

nodomhs2_server_instance:

type: OS::Nova::Server

depends_on: nodomhs1_instance_wait

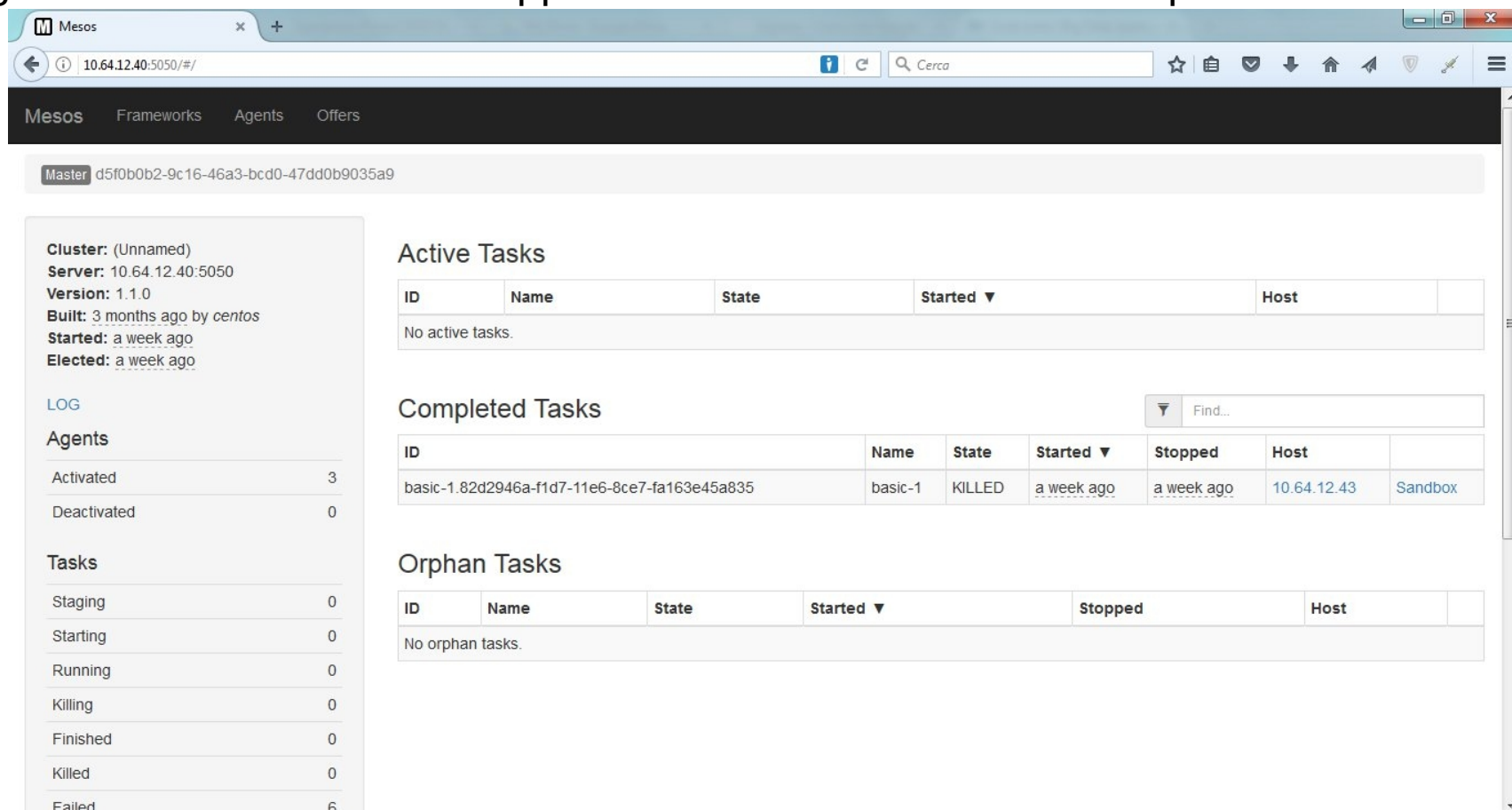
...

outputs:

root_pw

Apache Mesos

- è un software open-source che gestisce delle risorse in un cluster;
- fornisce un isolamento efficiente delle applicazioni eseguite nel cluster:
- gestisce la condivisione di applicazioni nel cluster dando delle priorità di esecuzione:



The screenshot shows the Apache Mesos web interface in a browser. The address bar shows the URL `10.64.12.40:5050/#/`. The interface has a dark navigation bar with links for Mesos, Frameworks, Agents, and Offers. Below this, a status bar shows the Master ID: `d5f0b0b2-9c16-46a3-bcd0-47dd0b9035a9`.

Cluster: (Unnamed)
Server: 10.64.12.40:5050
Version: 1.1.0
Built: 3 months ago by centos
Started: a week ago
Elected: a week ago

LOG

Agents

State	Count
Activated	3
Deactivated	0

Tasks

State	Count
Staging	0
Starting	0
Running	0
Killing	0
Finished	0
Killed	0
Failed	6

Active Tasks

ID	Name	State	Started ▼	Host
No active tasks.				

Completed Tasks

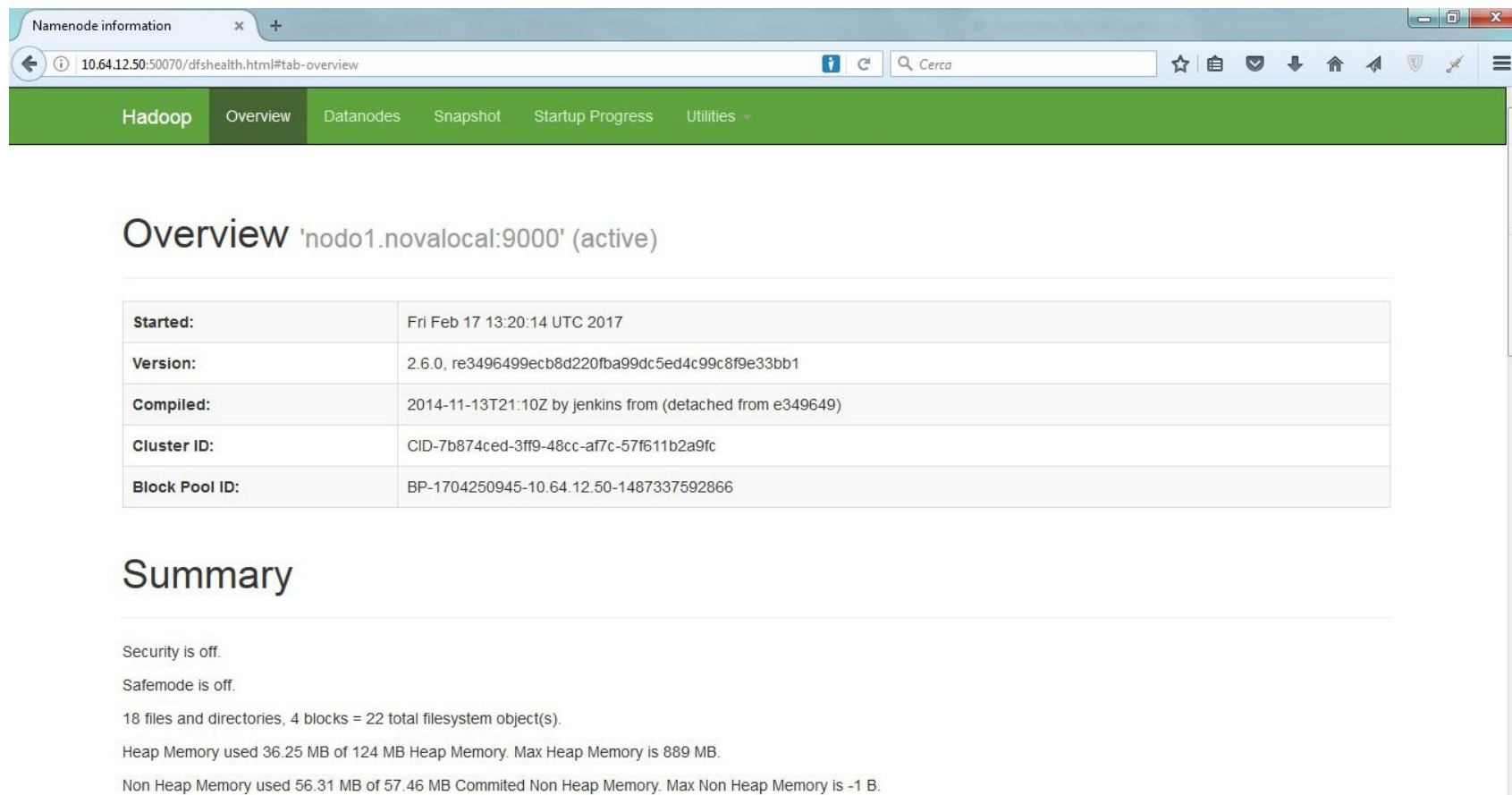
Find...

ID	Name	State	Started ▼	Stopped	Host	
basic-1.82d2946a-f1d7-11e6-8ce7-fa163e45a835	basic-1	KILLED	a week ago	a week ago	10.64.12.43	Sandbox

Orphan Tasks

ID	Name	State	Started ▼	Stopped	Host
No orphan tasks.					

- è un framework opensource che supporta applicazioni distribuite con elevato accesso ai dati
- quello che noi utilizziamo in questo momento è solo HDFS (Hadoop Distributed File System) che è un file system distribuito, portabile e scalabile scritto in Java.



The screenshot shows the Hadoop Namenode information page in a web browser. The browser address bar shows the URL `10.64.12.50:50070/dfshealth.html#tab-overview`. The page has a green navigation bar with tabs: Hadoop, Overview, Datanodes, Snapshot, Startup Progress, and Utilities. The 'Overview' tab is selected, showing the title 'Overview 'nodo1.novalocal:9000' (active)'. Below the title is a table with the following data:

Started:	Fri Feb 17 13:20:14 UTC 2017
Version:	2.6.0, re3496499ecb8d220fba99dc5ed4c99c8f9e33bb1
Compiled:	2014-11-13T21:10Z by jenkins from (detached from e349649)
Cluster ID:	CID-7b874ced-3ff9-48cc-af7c-57f611b2a9fc
Block Pool ID:	BP-1704250945-10.64.12.50-1487337592866

Below the table is a 'Summary' section with the following text:

Security is off.
Safemode is off.
18 files and directories, 4 blocks = 22 total filesystem object(s).
Heap Memory used 36.25 MB of 124 MB Heap Memory. Max Heap Memory is 889 MB.
Non Heap Memory used 56.31 MB of 57.46 MB Committed Non Heap Memory. Max Non Heap Memory is -1 B.

- Apache Spark è un framework open source progettato per il cluster computing
- Si differenzia dal paradigma MapReduce, basato sul disco a due livelli di Hadoop, in quanto usa delle primitive "in-memory" multilivello
- permette ai programmi utente di caricare dati in un gruppo di memorie e interrogarlo ripetutamente
- Spark è studiato appositamente per algoritmi di apprendimento automatico.
- Spark richiede un gestore di cluster e un sistema di archiviazione distribuita:
 - supporta nativamente un cluster Spark ma anche Hadoop YARN, o Apache Mesos
 - può interfacciarsi con Hadoop Distributed File System (HDFS), Apache Cassandra[3] , OpenStack Swift, Amazon S3, Apache Kudu

Apache Spark 2/2

Spark Master at spark://nodo1....
+
10.64.12.50:8080
Cerca



Spark Master at spark://nodo1.novalocal:7077

URL: spark://nodo1.novalocal:7077
REST URL: spark://nodo1.novalocal:6066 (cluster mode)
Alive Workers: 5
Cores in use: 18 Total, 0 Used
Memory in use: 17.0 GB Total, 0.0 B Used
Applications: 0 Running, 15 Completed
Drivers: 0 Running, 0 Completed
Status: ALIVE

Workers

Worker Id	Address	State	Cores	Memory
worker-20170217111915-10.64.12.51-42941	10.64.12.51:42941	ALIVE	2 (0 Used)	2.9 GB (0.0 B Used)
worker-20170217112006-10.64.12.55-49278	10.64.12.55:49278	ALIVE	4 (0 Used)	3.5 GB (0.0 B Used)
worker-20170217112026-10.64.12.54-46607	10.64.12.54:46607	ALIVE	4 (0 Used)	3.5 GB (0.0 B Used)
worker-20170217112108-10.64.12.53-59507	10.64.12.53:59507	ALIVE	4 (0 Used)	3.5 GB (0.0 B Used)
worker-20170217112124-10.64.12.52-38527	10.64.12.52:38527	ALIVE	4 (0 Used)	3.5 GB (0.0 B Used)

Running Applications

Application ID	Name	Cores	Memory per Node	Submitted Time	User	State	Duration
----------------	------	-------	-----------------	----------------	------	-------	----------

Completed Applications

Application ID	Name	Cores	Memory per Node	Submitted Time	User	State	Duration
app-20170221095303-0014	Spark shell	4	1024.0 MB	2017/02/21 09:53:03	andreett	FINISHED	13 min

- la scelta di mesos è stata fatta per permettere di bilanciare il carico tra hadoop, spark o altre applicazioni che potremmo istanziare sul cluster gestito da mesos
- la scelta di spark è stata fatta in quanto permette anche di leggere stream di dati (lettura online di dati provenienti da DAQ)
 - inoltre sembra più performante di yarn o mapreduce
- siamo ancora in una fase di testing e affinamento del deploy di spark nel cluster mesos
- stiamo iniziando ora a fare i primi test di deploy di applicazioni CMS nel cluster, usando librerie diana/hep, file avro o altre conversioni di file root e histogrammar per i plot dei risultati dell'analisi