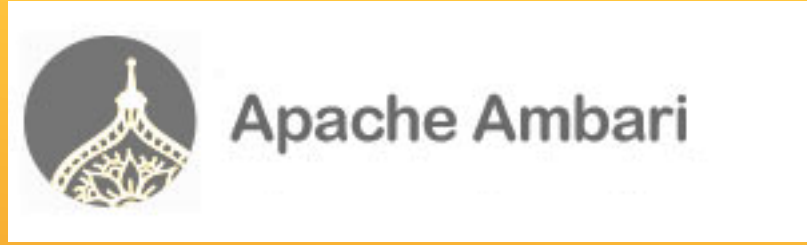


# UTILIZZO CLUSTER UNIROMA3



# AMBARI



The Apache Ambari project is aimed at making Hadoop management simpler by developing software for provisioning, managing, and monitoring Apache Hadoop clusters. Ambari provides an intuitive, easy-to-use Hadoop management web UI backed by its RESTful APIs.

# AMBARI



The screenshot shows the Apache Ambari dashboard for a cluster named "Clusterino". The top navigation bar includes "Dashboard", "Services", "Hosts", "Alerts", "Admin", and a user profile for "admin". The dashboard displays various metrics for the cluster's health and performance.

**Clusterino** 0 ops 2 alerts

Navigation: Metrics | Heatmaps | Config History

Metric Actions | Last 1 hour

Metric	Value
HDFS Disk Usage	1%
DataNodes Live	20/20
HDFS Links	NameNode, Secondary NameNode, 20 DataNodes
Memory Usage	4.6 GB
Network Usage	9.7 KB
CPU Usage	100% (graph)
Cluster Load	20 (graph)
NameNode Heap	11%
NameNode RPC	0.08 ms
NameNode CPU WIO	0.0%
NameNode Uptime	1.1 d
ResourceManager Heap	6%
ResourceManager Uptime	1.1 d
YARN Memory	6%
NodeManagers Live	20/20

YARN Links: ResourceManager, 20 NodeManagers

Left sidebar services: HDFS, YARN, MapReduce2, Tez, Hive, Pig, ZooKeeper, Ambari Infra, Ambari Metrics, Ranger, SmartSense, Spark, Slider.

# AMBARI FILE VIEW

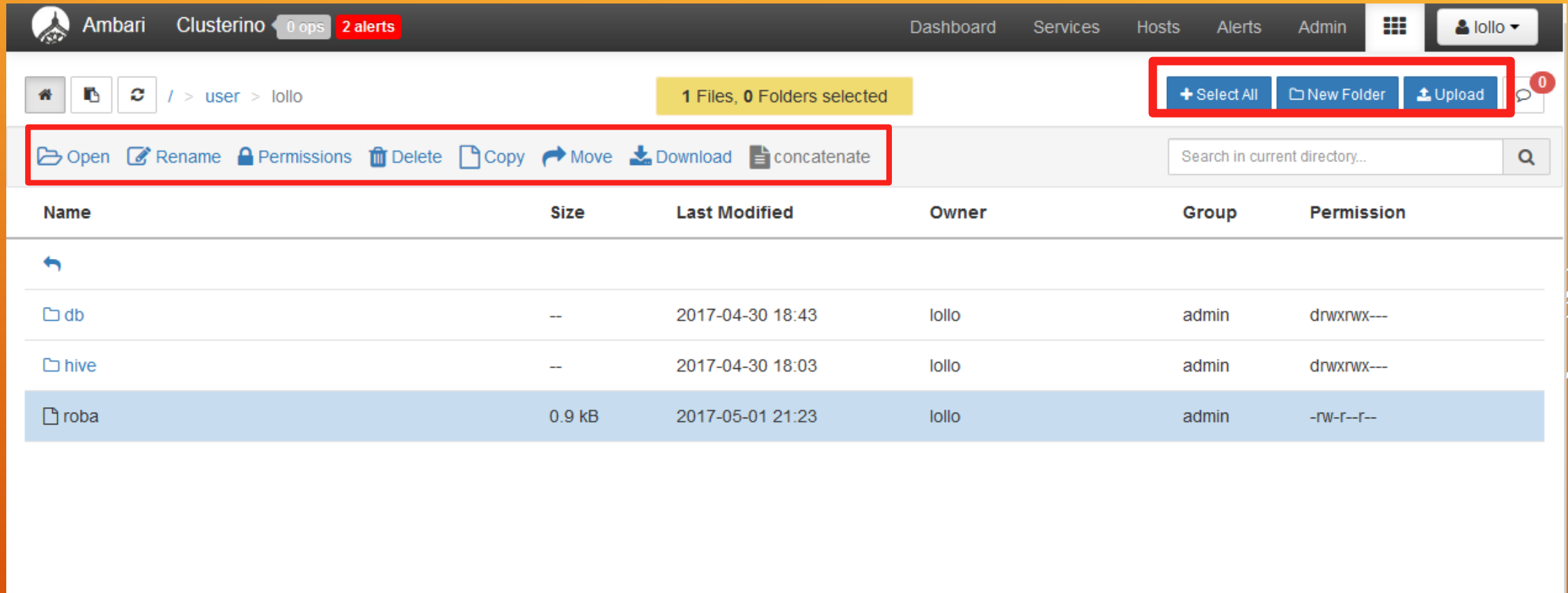


The screenshot shows the Ambari dashboard for a cluster named "Clusterino" with 0 ops. The top navigation bar includes "Dashboard", "Services", "Hosts", "Alerts", "Admin", and a user profile "lollo". A red box highlights the "Files View" and "Hive View 2.0" options in the top right corner, with a red arrow pointing to the "Files View" option.

The dashboard displays various metrics and charts:

- HDFS Disk Usage:** 1%
- DataNodes Live:** 20/20
- HDFS Links:** NameNode, Secondary NameNode, 20 DataNodes
- Memory Usage:** 4.6 GB
- Network Usage:** 9.7 KB
- CPU Usage:** 100% (max), 50% (current)
- Cluster Load:** 20 (max), 10 (current)
- NameNode Heap:** 2%
- NameNode RPC:** 0.08 ms
- NameNode CPU WIO:** 0.0%
- NameNode Uptime:** 1.1 d
- ResourceManager Heap:** 8%
- ResourceManager Uptime:** 1.1 d
- YARN Memory:** 6%
- NodeManagers Live:** 20/20
- YARN Links:** ResourceManager, 20 NodeManagers

# AMBARI FILE VIEW



The screenshot shows the Ambari File View interface for a cluster named "Clusterino". The breadcrumb path is "/ > user > lollo". A yellow status bar indicates "1 Files, 0 Folders selected". A red box highlights the action menu containing "Open", "Rename", "Permissions", "Delete", "Copy", "Move", "Download", and "concatenate". Another red box highlights the top right controls: "+ Select All", "New Folder", and "Upload". A search bar is present with the text "Search in current directory...".

Name	Size	Last Modified	Owner	Group	Permission
↳					
db	--	2017-04-30 18:43	lollo	admin	drwxrwx---
hive	--	2017-04-30 18:03	lollo	admin	drwxrwx---
roba	0.9 kB	2017-05-01 21:23	lollo	admin	-rw-r--r--

NB Quando si elimina un file ricordarsi di mettere la spunta a «Delete permanently»

# AMBARI HIVE VIEW 2.0

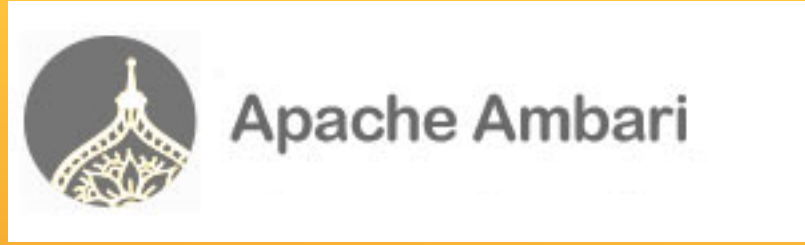


A screenshot of the Apache Ambari Hive View 2.0 interface. The top navigation bar includes "Ambari Clusterino" with status indicators for "0 ops" and "3 alerts", and menu items for "Dashboard", "Services", "Hosts", "Alerts", and "Admin". The main header shows "HIVE" with buttons for "+ NEW JOB" and "+ NEW TABLE". Below this is a navigation menu with "QUERY", "JOBS", "TABLES", "SAVED QUERIES", "UDFs", and "SETTINGS". The main workspace is titled "Worksheet1" and contains a "DATABASE" section with a search input field containing "lollobd" (highlighted with a red box) and a "Browse" button. Below the search field is a large text area for writing queries. To the right of the text area is a sidebar showing the selected database "lollobd" and a list of tables: "1999\_2006" and "num". At the bottom of the interface are buttons for "Execute", "Save As", "Insert UDF", and "Visual Explain", along with a footer menu with "RESULTS", "LOG", "VISUAL EXPLAIN", and "TEZ UI".

Verificare sempre che sia selezionato il database corretto, altrimenti le query non andranno a buon fine!



# LIMITI



Per garantire un utilizzo corretto del cluster le vostre utenze sono limitate a:

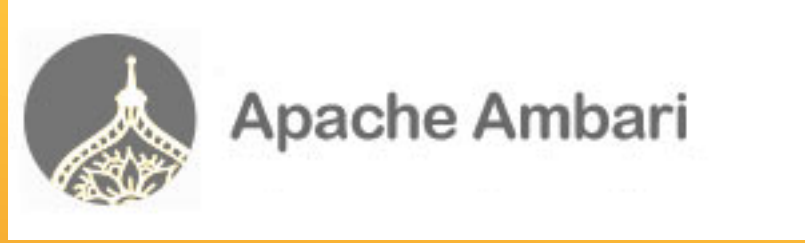
- HDFS 30GB
- Disco su macchina 50GB

# HADOOP MAP-REDUCE

```
$> yarn jar wordcount.jar /user/lollo/fileToCount /user/lollo/output
```



# LINK E INDIRIZZI UTILI



- Il cluster è visibile **ESCLUSIVAMENTE** dall'interno della rete universitaria
- Ambari: <http://cluster.inf.uniroma3.it:8080>
- SSH: `ssh cluster.inf.uniroma3.it`
- Hadoop History <http://cluster.inf.uniroma3.it:8088>
- Spark History <http://cluster.inf.uniroma3.it:18080>
- Cartella utente `/user/userX/`

Se volete cambiare la password dell'account le operazioni da effettuare sono **DUE**:

- Da SSH con il comando «**passwd**»
- Dall'interno di Ambari