

facebook

facebook

Hadoop Usage at Facebook

Dhruba Borthakur
Facebook Data Infrastructure Team
ApacheCon2008, New Orleans, 11/6/2008

Data Infrastructure Team

- **Cross functional team of 11 members**
 - **5 people working in Hive development**
 - **2 people on Hadoop development**
 - **2 people on Data Pipelines and Oracle Data Mart**
 - **1 Production Operations**

Hadoop Cluster Hardware

- **Production cluster**

- **2400 cores, 300 machines, 16GB per machine -- Oct 2008**
- **4800 cores, 600 machines, 16GB/8GB per machine – Nov 2008**
- **4 SATA disks of 1 TB each per machine**
- **2 level network hierarchy, 40 machines per rack**

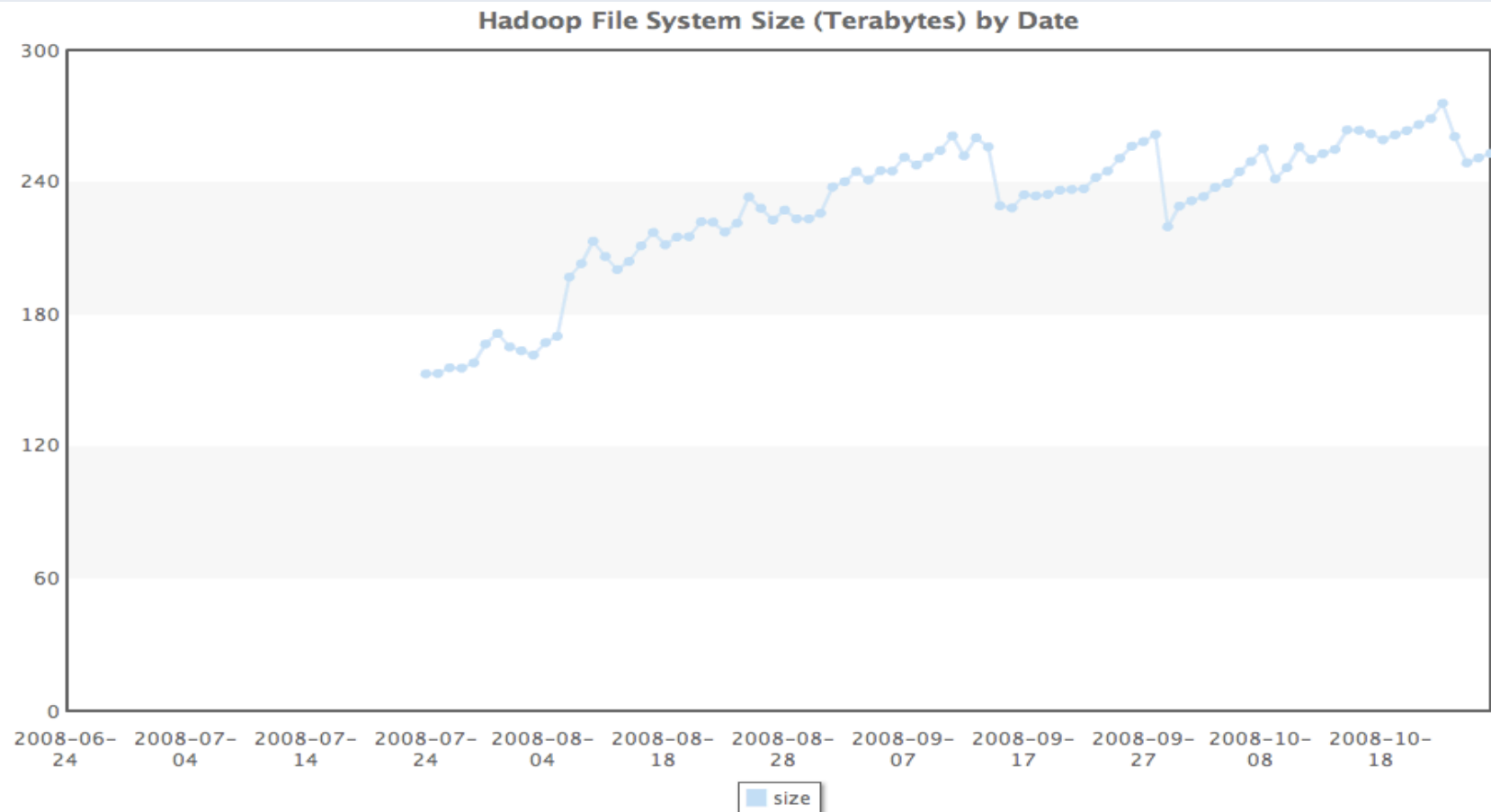
- **Test cluster**

- **800 cores, 16GB each**

HDFS Storage Cluster

- **Single HDFS cluster across all cores**
 - **Running 0.17.2 + patches**
 - **Over 1.2 PB raw capacity**
 - **Ingest rate is 2 TB compressed user-data per day**
 - **10 TB uncompressed**
 - **10 Million files**
 - **NameNode on 64 bit JVM with 20GB heap size**

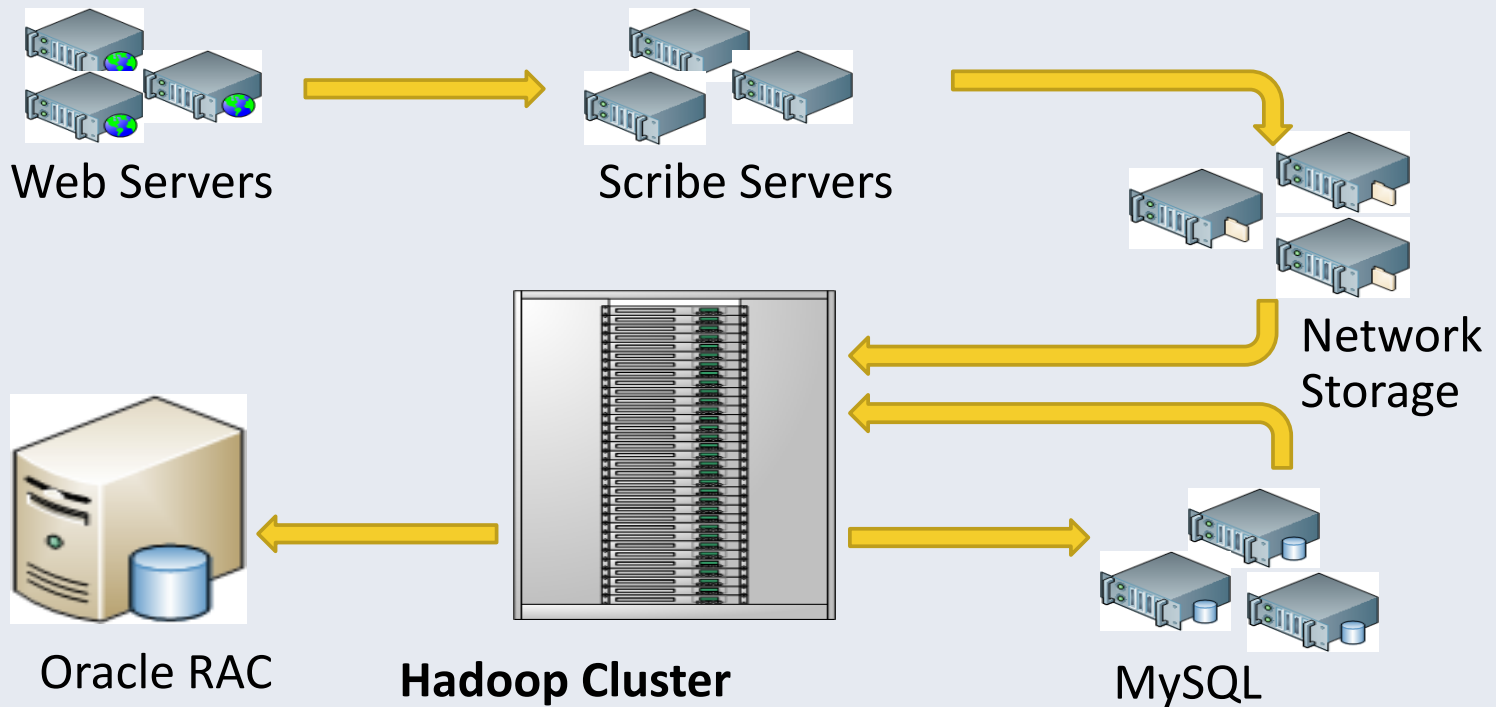
Rate of Growth of Hadoop Data



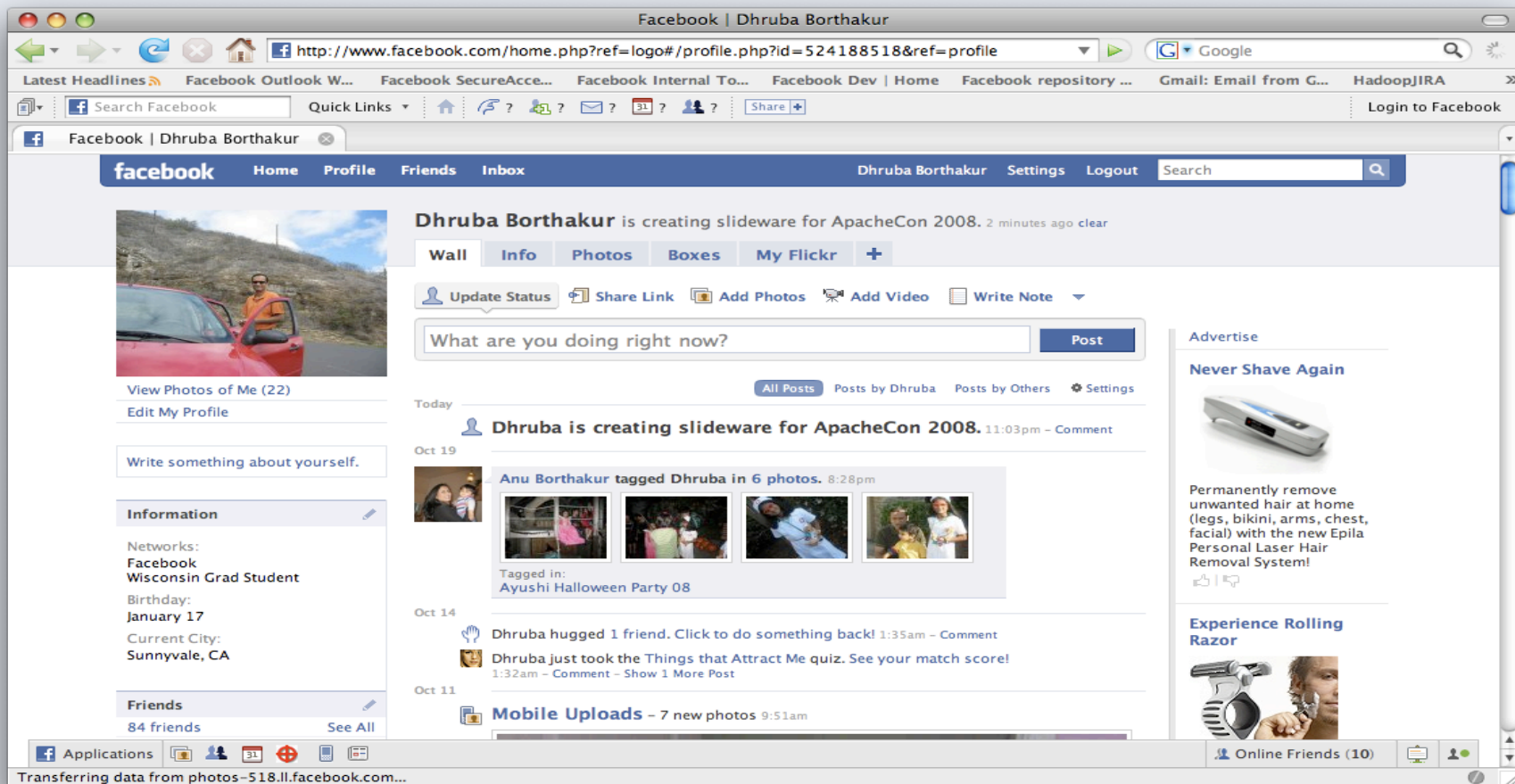
Map-Reduce Compute Clusters

- **3 static Map-Reduce clusters**
 - **Running 0.17.2 + patches**
 - **Main cluster has 2240 cores**
 - **Serves most users**
 - **Ads cluster has 80 cores**
 - **Dedicated for *Advertisement* related computations**
 - **Test cluster of 80 cores**
 1. **Test miscellaneous fixes**
- **Job Tracker(s) on 32 bit JVM with 3GB heap size**

Data Flow



A Facebook User has many dimensions



User Data

- 100M+ Active users
- Each user has 100+ dimensions
- 500K new user per day (one San Francisco every 36 hours)

Facebook Friend List

The screenshot shows a web browser window displaying the Facebook 'All Friends' page for a user named 'apachecon'. The browser's address bar shows the URL: <http://www.facebook.com/friends/?ref=tn#/friends/?id=524188518&flid=32603988518&view>. The page header includes the Facebook logo, navigation links (Home, Profile, Friends, Inbox), the user's name 'Dhruba Borthakur', and options for Settings and Logout. A search bar is also present.

The main content area is titled 'apachecon' and includes a link to 'Find people you know on Facebook'. Below this, there are tabs for 'Showing: Status Updates, Recently Updated, Phonebook, Everyone'. A search box for friends is also visible.

The page indicates that the list contains 5 friends. Below this, there is an 'Add to List' section with a search box and a 'Select Multiple Friends' button.

The friend list contains the following entries:

- Jeff Hammerbacher** is shocked that MacTeX is over a gig.
Status and Pages
Profile updated 20 hours ago
- Owen O'Malley** is back in Sunnyvale.
on Friday
- Benjamin Reed** is frantically cleaning the house before Carolina gets back.
on Saturday
- Mike Speiser** Forget The Tipping Point. Focus on Joe the Plumber.: “Things on a very small scale behave like not.. <http://tinyurl.com/6pwts>.
on Saturday - via Twitter
- Ding Zhou** Go Penn State!
Posted Items and Status
Profile updated on Friday

On the right side of the page, there is an 'Advertise' section for 'JESS3 Blog' with a small image and a description: 'JESS3 is a creative agency that specializes in web design, branding and social media PR. We blog about creative inspiration.' Below this is a 'Chat with Wendigo33' section with a message: 'I've got a funny picture of her in the locker room. I'll forward it 2U.' and a small image of a cat.

The bottom of the browser window shows the 'Done' bar with various application icons and a notification for 'Online Friends (8)'.

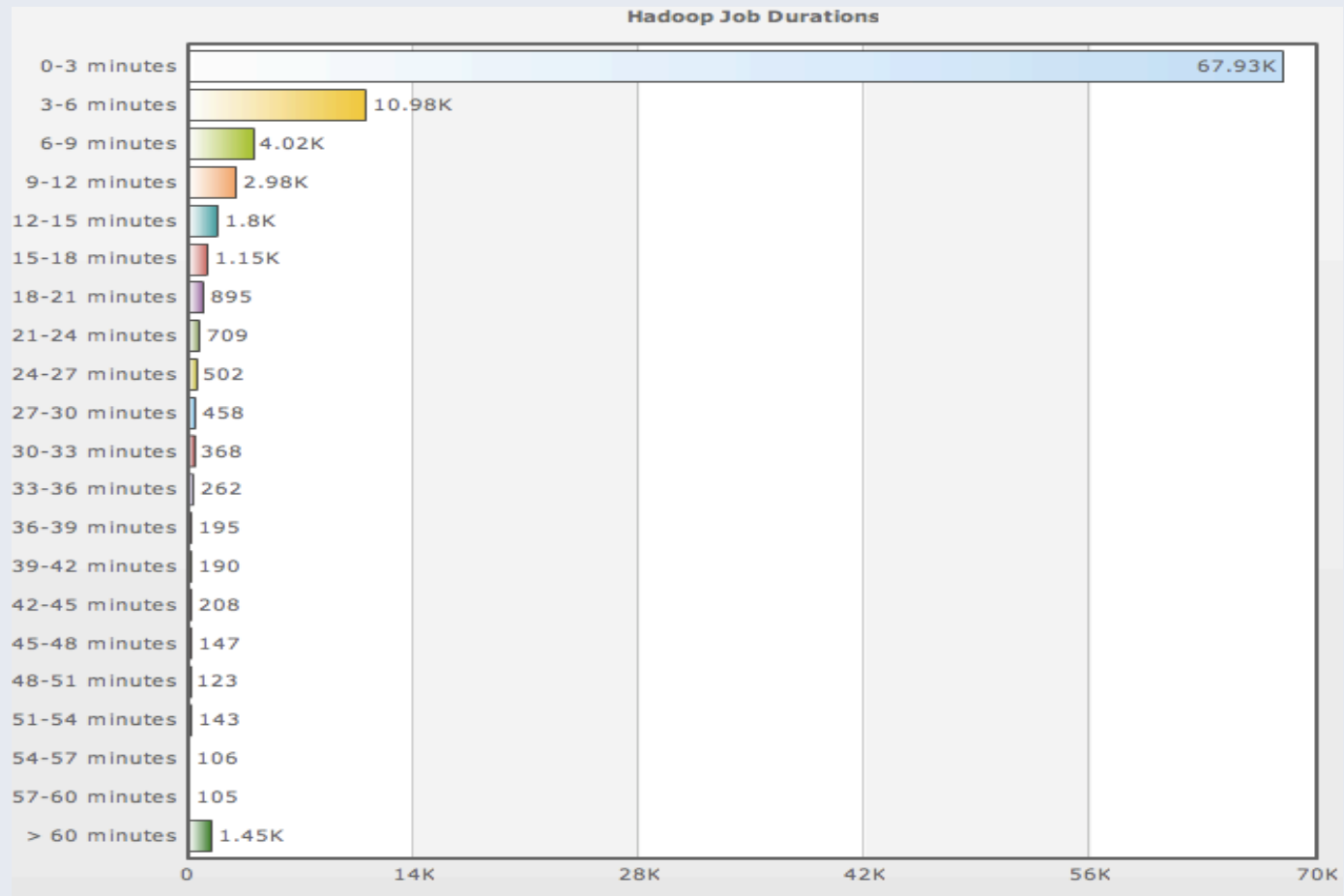
Social Graph Data

- 200+ friends per user (median)
- Interactions among users (100+ types of interactions)
- Interactions among users and application

Data Usage

- Statistics per day:
 - 55TB of compressed data scanned per day
 - 3200+ jobs on production cluster per day
 - 80M compute minutes per day
- Barrier to entry is significantly reduced:
 - All new engineers go through a Hadoop training sessions
 - 50+ engineers have run jobs on Hadoop platform
 - Analysts (non-engineers) starting to use Hadoop through Hive

Typical Job Durations



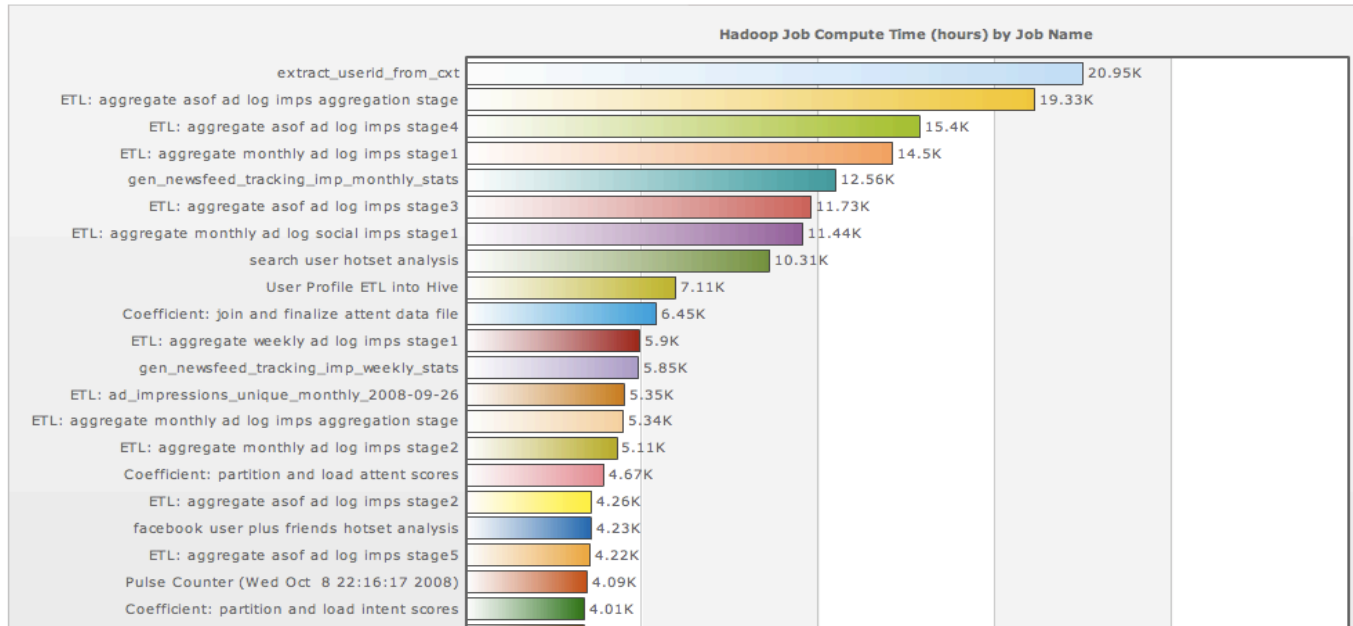
Cluster Usage Dashboard

- History logs are fed into a mySQL database
- A Dashboard displays cluster usage statistics from the database
- Displays cluster utilization, growth rates of cluster usage, etc
- HADOOP-3708

Cluster Usage Dashboard

- Jobs
- Compute Time
- Map Time
- Reduce Time
- Job Durations
- Map Durations
- Reduce Durations
- Jobs by Date
- Compute Time by Date
- Task Time by Date
- I/O by Date
- HDFS Size
- HDFS Metadata
- Largest Hive Tables
- Largest Home Directories
- Largest Facebook Project Dirs

Days back: Break down by:



Confidential Materials — For Internal Use Only
 Hadoop Job Compute Time (hours) by Job Name



Add Comment

Fair Share Scheduler

- Short jobs finish fast while not starving longer jobs
- A pool has a guaranteed minimum allocation
 - Data loading jobs use a pool with large resource allocation
- Allows limiting the number of concurrent jobs per pool
- No preemption yet
- HADOOP-3746

Fair Share Scheduler

Pools

Pool	Running Jobs	Min Maps	Min Reduces	Running Maps	Running Reduces
akramer	1	0	0	214	0
dferrante	0	0	0	0	0
hwang	0	0	0	0	0
itamar	0	0	0	0	0
jssarma	1	0	0	1	0
lev	0	0	0	0	0
prha	0	0	0	0	0
realtime	0	50	50	0	0
root	5	1000	1000	7	2
viyer	0	0	0	0	0
default	0	0	0	0	0

Universal Access to HDFS

- Mount HDFS through fuse (HADOOP-4)
 - An user can use common utilities, e.g. find, ls on HDFS files
 - Goal to have HDFS fuse-mounted on all developer machines
- Use Thriftfs (HADOOP-3754)
 - Access HDFS through Python, PHP
 - Enables custom joins written in Python

WebUI

HiPal: an Online Tool for Querying Hive/Hadoop Data Warehouse

[+ Learn More about HiPal](#) + [Why am I on dev127?](#)

Query

Table Start Partition End Partition Data Size (bytes): Cat/Export Data
734,184,513,313 [Get 10 rows](#) | [Export the whole u_full](#)

Select Columns [\[All\]](#) [\[Clear\]](#)
 userid base affiliations last friends ext groups fbpages whs events photo_tags schools applications regionid

- [+ Join Options](#)
- [+ Group By Options](#)
- [+ Where Options](#)
- [+ Query Options](#)

```
CREATE TABLE tmp_hipal_<QUERYID> (userid STRING, friends STRING, schools STRING);  
FROM u_full TABLESAMPLE (BUCKET 1 OUT OF 1024) a  
INSERT OVERWRITE TABLE tmp_hipal_<QUERYID>  
SELECT a.userid, a.friends, a.schools
```

[\[Join HiPal User Mailing List\]](#)[\[Report problems or ask questions\]](#)

Job Status

Show all jobs enable Sort By In

QueryId	Submit Time	User	Query (Last Update: 2008-10-27 12:42:58 AM)	Time (sec)	Query Progress	Latest Hadoop Job
3393	2008-10-15 1:48:20 pm	dhruba	CREATE TABLE tmp_hipal_<QUERYID> (user STRING); FROM f_add_video TABLESAMPLE (BUCKET 1 OUT OF 1024) a INSERT OVERWRITE TABLE tmp_hipal_<QUERYID> SELECT a.user WHERE a.ds>='2007-10-27' AND a.ds<='2008-05-28'	57	<div style="width: 100%;"><div style="width: 100%; background-color: green;">100%</div></div>	Status

In the Works

- Hierarchical Storage Manager (HADOOP-4058)
 - Migrate data that will be rarely used in future
- Global scheduler
 - Ability to schedule more tasks to be rack-local
 - Ability to process multiple local splits by a single task
- Data Protection
 - HDFS Snapshots (HADOOP-3637)
 - HDFS Symlinks (HADOOP-4044)

Questions?

dhruba@facebook.com

Credits

Suresh Anthony

Zheng Shao

Prasad Chakka

Pete Wyckoff

Namit Jain

Raghu Murthy

Joydeep Sen Sarma

Rama Ramasamy

Matei Zaharia

Ashish Thusoo

Hao Liu

facebook

(c) 2008 Facebook, Inc. or its licensors. "Facebook" is a registered trademark of Facebook, Inc.. All rights reserved. 1.0