facebook

HBase at Facebook

Jonathan Gray
Hadoop User Group UK
November 19, 2010

About Me
Jonathan Gray

- Software Engineer at Facebook

- Member of the Facebook Open Source Team
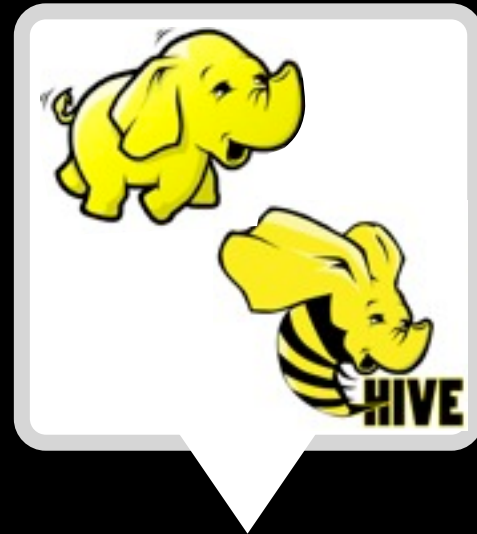
- HBase Committer for several years

Data at Facebook

500 Million active monthly users

›**500 Billion** page views per month

25 Billion pieces of content per month

Cache

Data analysis

OS

Web server

Database

Language

Hadoop & Hive Cluster

Cluster Usage

- 100TB ingested and 1PB processed per day

- 10-25k jobs per day

- New job submitted every 10 seconds

300-400 users submit a job each month

Hadoop & Hive Cluster

Cluster Statistics

- Largest capacity Hadoop cluster in the World

- Multiple MR clusters (SILVER and PLATINUM)

- 3000 high density nodes
  - Each with 8 cores, 32 GB RAM, 12 TB of disk

24k cores, 96TB RAM, 36PB disk

Hadoop Improvements
Disk Capacity

- HDFS limited by disk capacity not bandwidth

- **HDFS-RAID**
  - Saves 25% without lose in FT

- **Columnar Compression**
  - Saves 15-25% for Hive tables

Hadoop Improvements
NameNode

- NameNode on one node

- **AvatarNode**
  - Active/passive hot standby pair of NNs
  - Failover 20PB across 65M files in 10 seconds

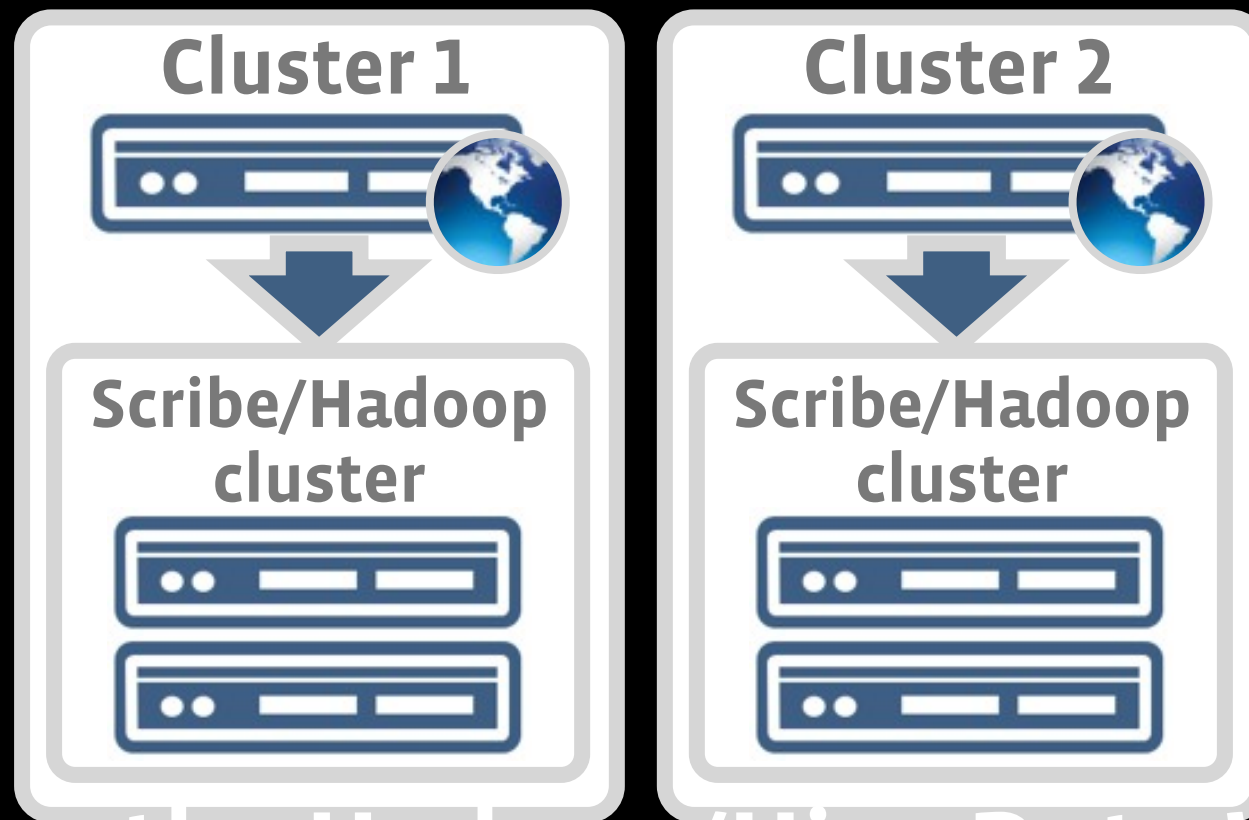- **Sharded NameNode, Hadoop HighTide**
  - Works in progress

HBase at Facebook

HBase

Key Properties

- Linearly **scalable**

- **Fast** indexed writes

- Tight integration with **Hadoop**

Bridges gap between *online* and *offline*

Augmenting the Hadoop/Hive Data Warehouse

Data Warehousing
Incremental Updates

- **Currently**
  - Nightly dumps of UDBs into Warehouse

- **With HBase**
  - UDB replication logs into HBase

UDB to Warehouse in *minutes*

Realtime Analytics

High Frequency Counters

- **Currently**
  - Scribe to HDFS, periodically aggregate to UDB

- **With HBase**
  - Scribeh to HBase, read in realtime via Thrift or MR

Storage, serving, and analysis *in one*

Facebook Messaging

Facebook Messaging
Now in Production

- **Launched Project Titan this Past Monday**
  - New messaging product (email, sms, im)

- **Largest Engineering Effort in History of FB**
  - 15 engineers over the course of a year
  - Incorporates over 20 infrastructure technologies
    - Haystack, Apache ZooKeeper, mail tools, etc...

Facebook Messaging
What is it?

- **Seamless messaging (SMS, chat, e-mail, facebook)**

- **Your own @*facebook.com* e-mail address**

- **Conversation History**

- **Social Inbox**

Facebook Messaging

Heavy Requirements

- **Lots of communication on Facebook**
  - More than 350 million users using messaging
  - 15 billion person-to-person messages / month
  - 120 billion instant messages / month

- **New version of messaging expected to increase usage**
  - Even more data: more reads, more writes

Facebook Messaging
Deciding on a database

- **Early this year, engineers at FB compared databases**
  - Apache Cassandra, Apache HBase, Sharded MySQL

- **Laid out requirements and compared performance, scalability, and features**
  - Small set of temporal, volatile data
  - Ever-growing set of rarely accessed data

Facebook Messaging
Deciding on a database

- **Sharded MySQL**
  - Did not handle long-tail of data well

- **Apache Cassandra**
  - Eventual consistency hard to reconcile for messages

Apache HBase

Facebook Messaging
Why HBase?

- **Good performance and scalability**

- **Strong consistency works well for messaging**

- **Already supported many features we needed**

- **Integration with Hadoop and inherits HDFS properties**

Conclusion

- **HBase can be used in both online and offline contexts**

  - **Add random read/write capabilities to warehouse**

  - **Efficient counting for realtime analytics**

  - **Fast and scalable writes with strong consistency make it a good match for some user-facing applications**

Thank You!

Questions?

facebook