

big data architecture at facebook

Ashish Thusoo  FB track : architectures you're always wondered about

scope
very simple reporting
index generation
one new job every second

2007
traditional EDW
web clusters
mysql clusters
with scribe mid-Tier
Summarization cluster write into RDBMS data warehouse

pain points
daily ETL > 24 hours
lots of indexes/tunings
lots of hardware planning
compute close to storage
very rigid
scribe and mysql on NAS filers

2009
democratizing data
building tools around the Hadoop dataWH
hipal : adhoc queries UI
databee & chronos : data pipeline
scrapes : configuration driven
nectar
json encoded short term data
decomposing json for long term storage
cons of democratization
isolation to protect against bad jobs
fair sharing : what priorities ?

2011
New Requirements
more real time requirements for aggregations
optimize
beyond hadoop
puma for real time analytics
Peregrine for simple and fast queries
between ptail and hbase

figures
25 PB of compressed data
150 PB uncompressed
400 TB/day uncompressed of new data

illustration
friend map - Paul Butler

2008
move to hadoop
hadoop/Hive data warehouse
pb : java
to make cluster available to everyone : develop layer

immediate pros
data science at scale became possible
all of the instrumented data could be held online
use cases expanded

controlling chaos
isolation
reducing operational overhead

2010
better resource utilization
split Warehouse into 2 parts
platinum WH
Hive replication
Silver Warehouse

RCFile : row colonar format for compressing hive tables
hive optimizations to save CPU

ops efficiency
HDFS accessed by web clusters
ptail parallel tail with consumers

monitoring
per job stat rolled to owner/group/team
expected time of arrival v.s. actual time of arrival of data

other challenges
moving datacenters
moving sustainably fast
2-3 moths to accomplish the entire move
moved 20 PB of data
blog post by paul yang : moving an elephant