# Chukwa: a large-scale monitoring system

Ari Rabkin, Andy Konwinski, Mac Yang, Jerome Boulon, Runping Qi, Eric Yang

chukwa

# The goal

- Data intensive debugging + analysis
- Around 3KB/node/sec of data to collect on Hadoop clusters.  (Logs + metrics)
- Want to store all this data, and analyze it
- Also want near-real time display of "cluster weather" -- load, storage available, etc.
  - Can be delayed by 5-10 minutes

- Need to use MapReduce to analyze

- Some sample uses:

  - Admin: Billing, accounting, provisioning

  - Development: SW log analysis and debugging

  - Ops: HW failures and performance

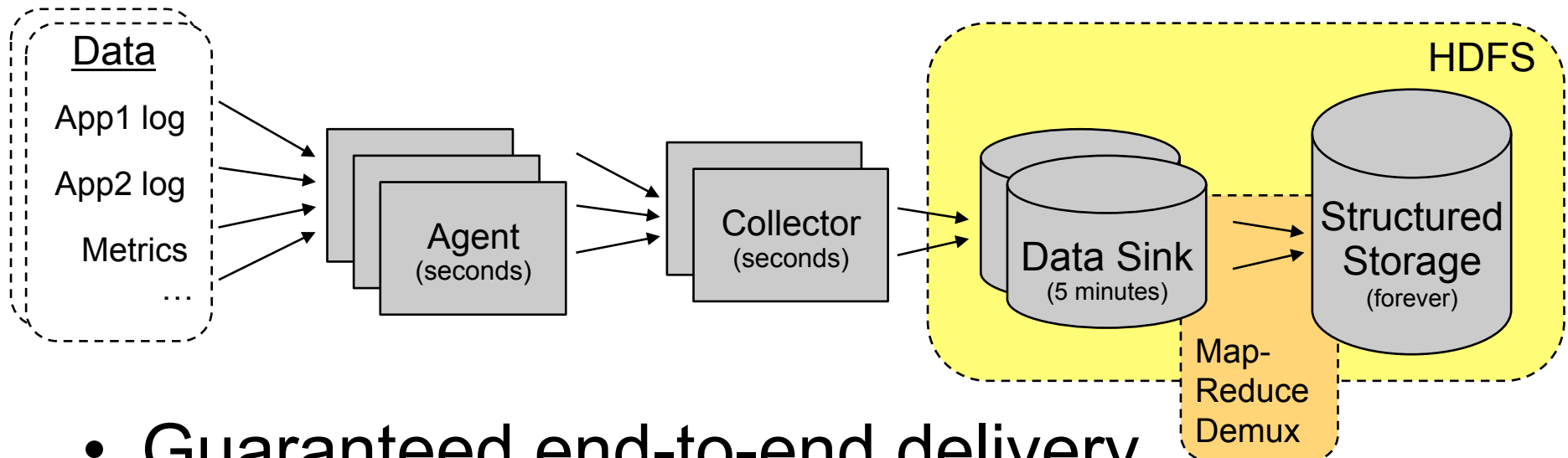  - Users: Want estimate of current resources

- Chukwa isn't a general-purpose streaming database.
  - Can't generate aggregates in the pipeline
  - No ad-hoc queries: MapReduce, not SQL
- Chukwa isn't a real-time system
  - We'd rather get everything eventually than get some of it right away
  - Don't want to re-implement Ganglia

# The solution

- Pipeline architecture



- Guaranteed end-to-end delivery
  - Failure tolerant, crash recovery
- Trade latency for scalability
  - Buffer data in temporary files
  - Use MapReduce to organize it

# Why build on Hadoop?

- Leverage existing code and expertise
- Gain from future Hadoop improvements
- It scales

# Why HDFS?

- Hadoop distributed filesystem (HDFS)…
  - Scales to petabytes
  - Has good performance for large reads/writes.
- Write: 20 MB/sec/client
  Read: 60 MB/sec/client
  - measured with 20-node HDFS and 5 clients
- Cons: some FS semantics and performance limitations
  - No appends, chokes on too many files

# Performance

- Goal: a 2000-node cluster generates ~5.5 MB/sec
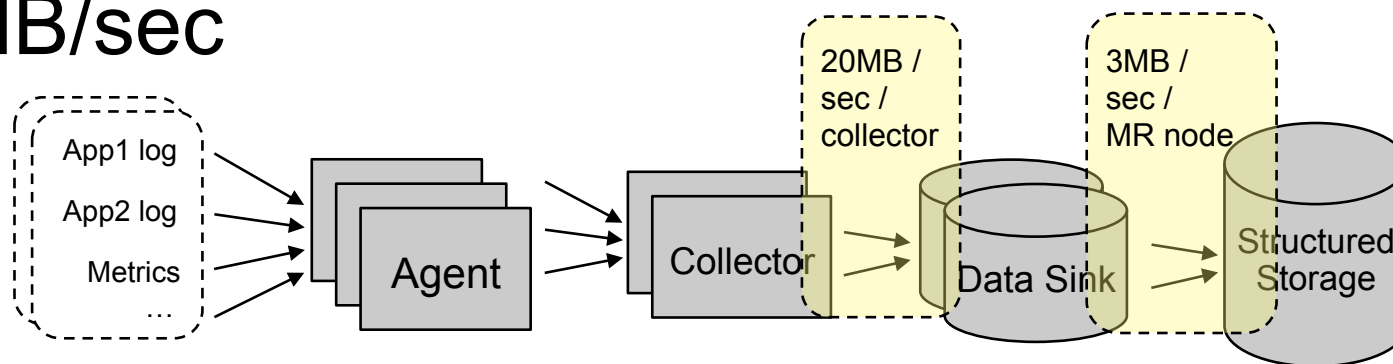


- Collectors can write at 20MB/sec/collector
  - No state at collectors, so easy to add more
- Demux MapReduce job is bottleneck, runs at 3 MB/sec/node
  - Can add nodes for speed
  - Hadoop will improve

- Entering service at Yahoo!
  - Currently on 500 grid nodes
  - 2000 nodes in next 2 months
- Open-source and available as a contribution to Hadoop

- More documentation
- Configuration/deployment needs polish
- Visualization tools not yet released (license issues)
- Bring latency down
  - Hadoop 0.19+ will make short jobs faster
- Real-time alarms + filters
- Better structured storage
  - Hive? Cassandra? Hypertable?

- http://wiki.apache.org/hadoop/Chukwa for

# Questions?

**Web:** http://wiki.apache.org/hadoop/Chukwa
**Email:** asrabkin@cs.berkeley.edu
andyk@cs.berkeley.edu