

Scaling Messaging Infra @ Twitter

An operational view

@sijieg



500,000,000

1.5 trillion

17.5 PB

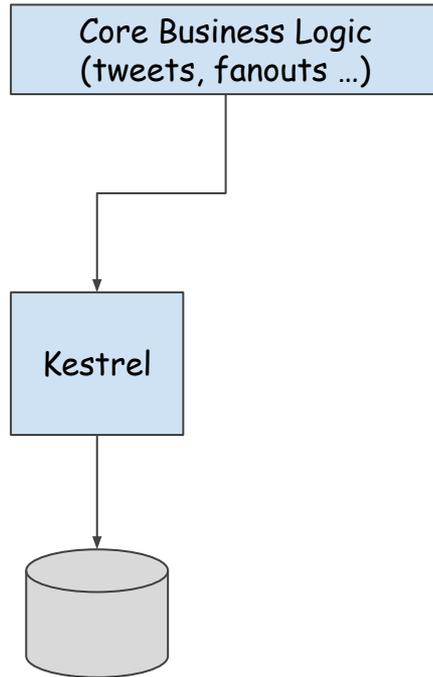


Messaging

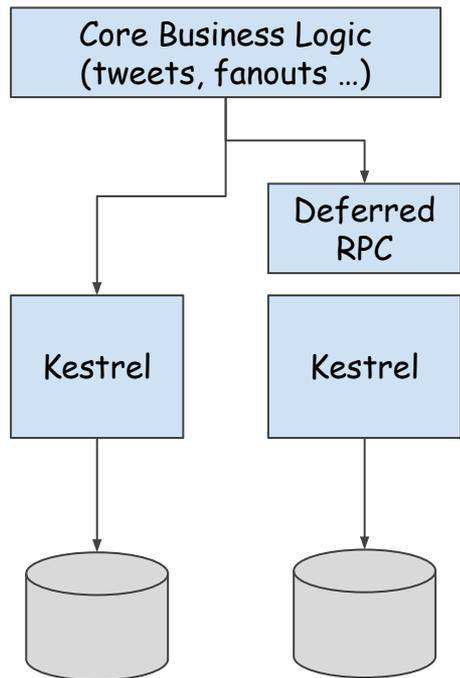
- Online Service - 10s of milliseconds
 - WAL, Transaction Log, Queues, RPC
- Real-time Processing - 100s of milliseconds
 - Change propagation, Stream Computing
- Data delivery for analytics - seconds to minutes
 - Log collection, aggregation



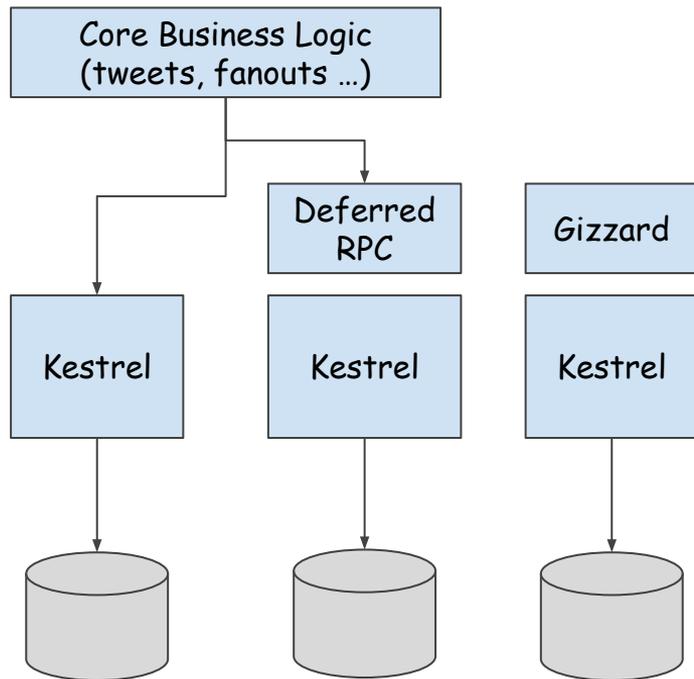
2012 - Online Services



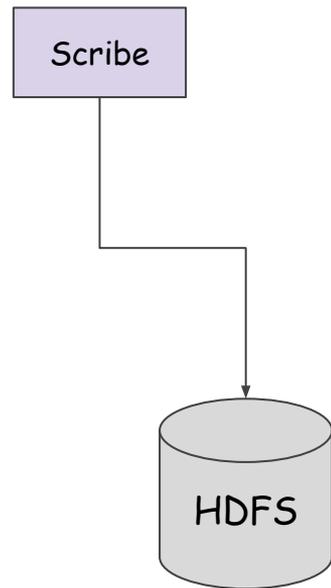
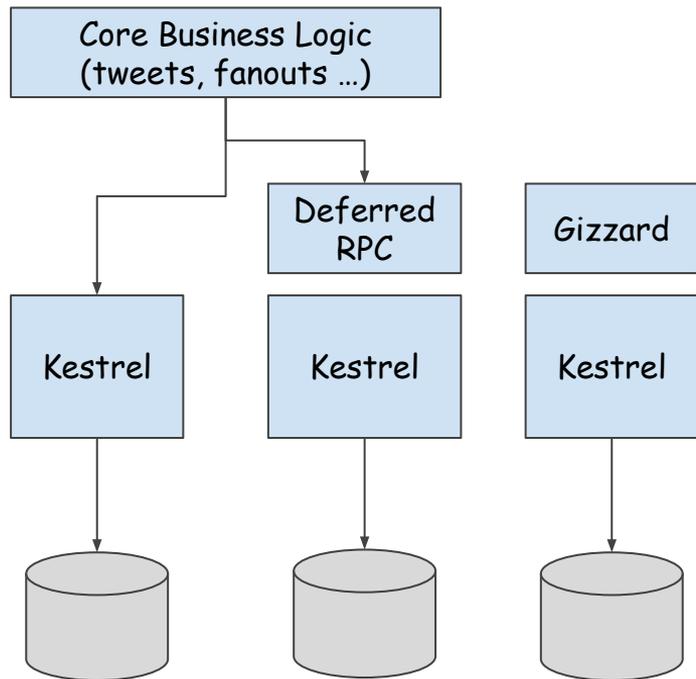
2012 - Online Services



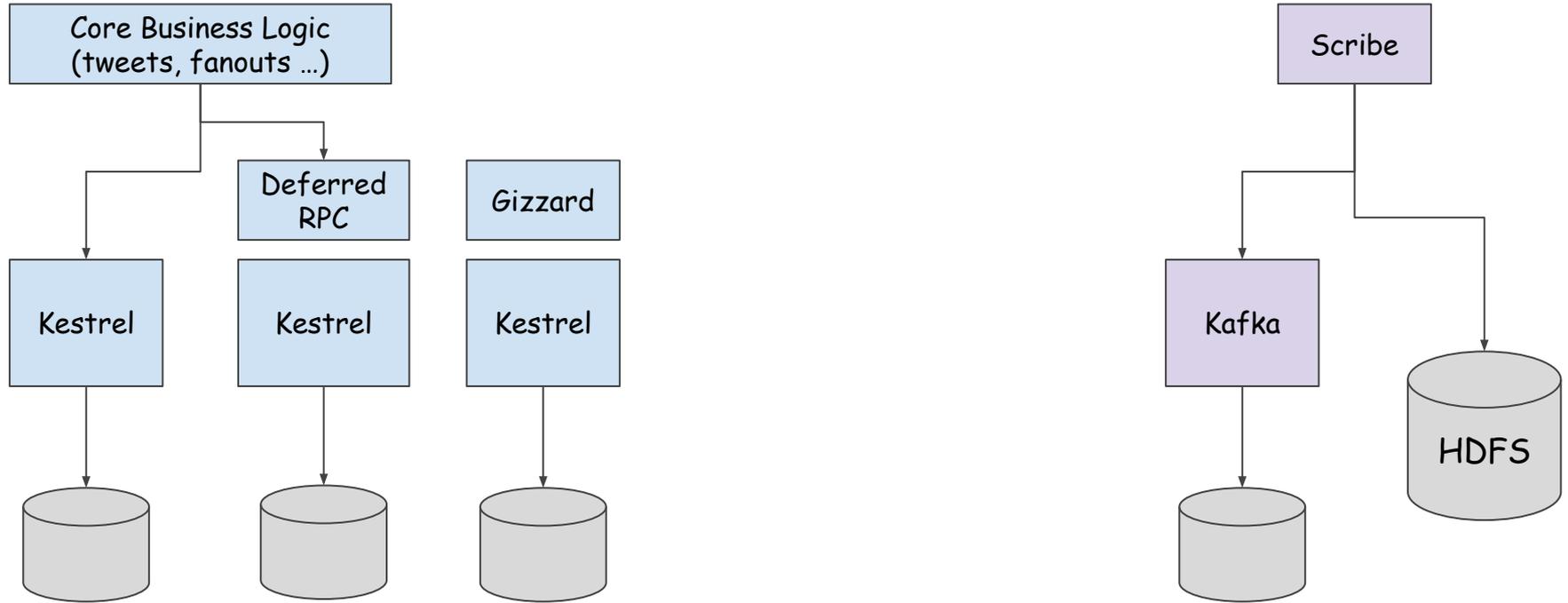
2012 - Online Services



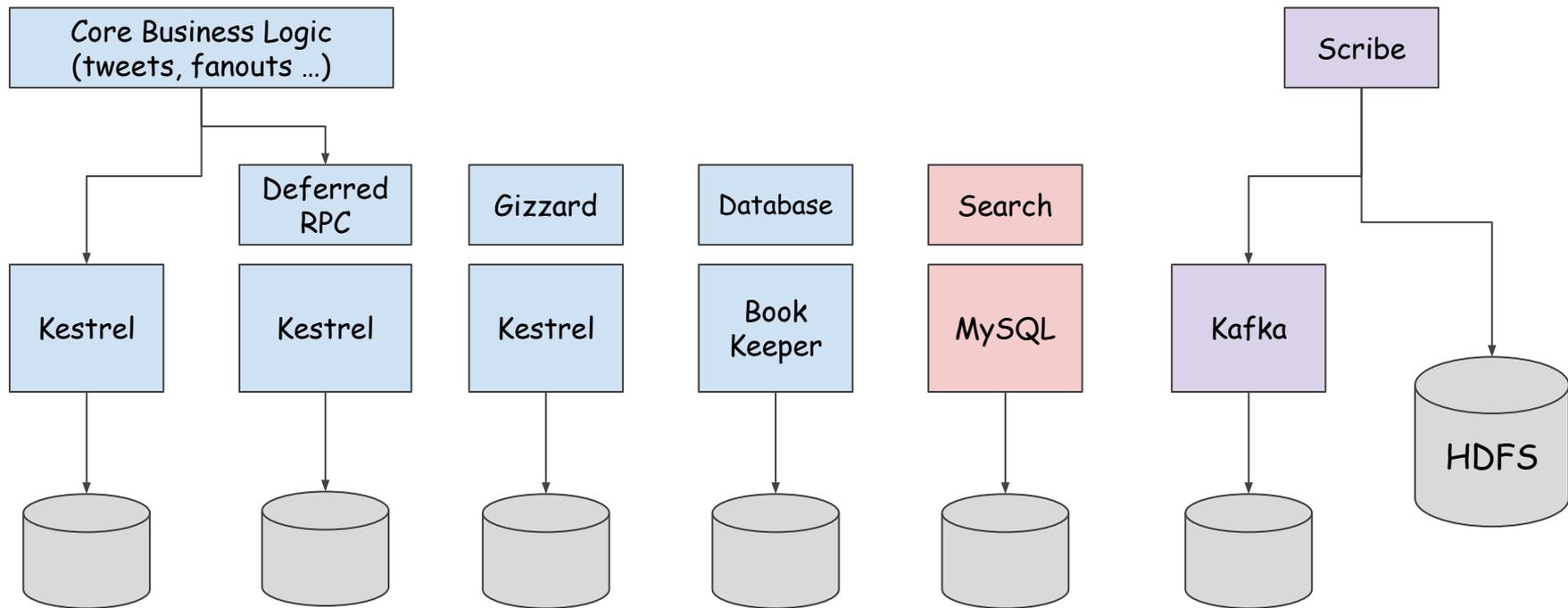
2012 - Analytics



2012 - Stream Computing



2012 - Transaction Log



Problems

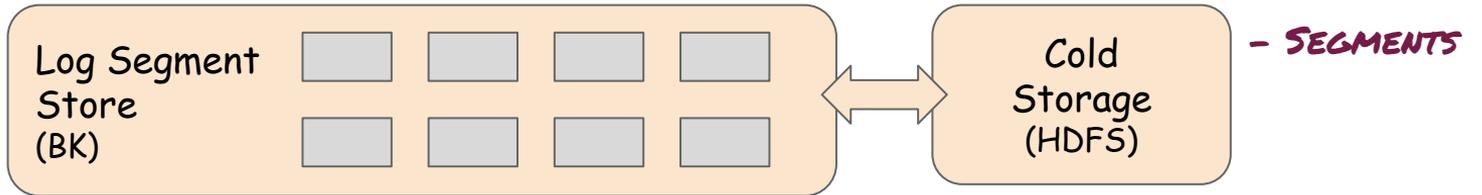


Operation Overhead

- Software Components
- Tools and Optimization
- Technical know-how



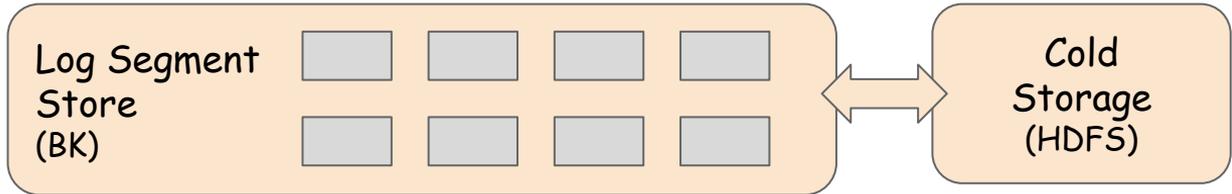
Storage



Storage



- RAW STREAMS



- SEGMENTS



Serving

Write - Ownership Tracking
Proxy - Batching, Compression

Record Cache - Read
Rate Limiting, Quota - Proxy

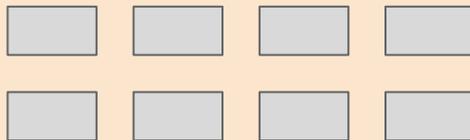
- PROXIES

Storage

Log Streams - Abstraction & Naming
- Data Management - Efficient Write & Read
- Intra-cluster & Geo Replication

- RAW STREAMS

Log Segment
Store
(BK)

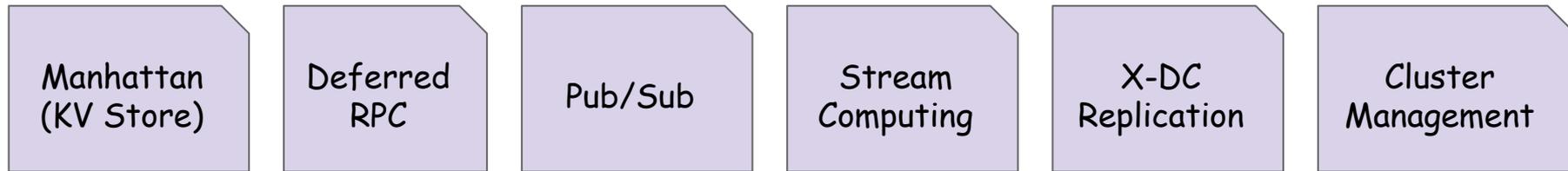


Cold
Storage
(HDFS)

- SEGMENTS



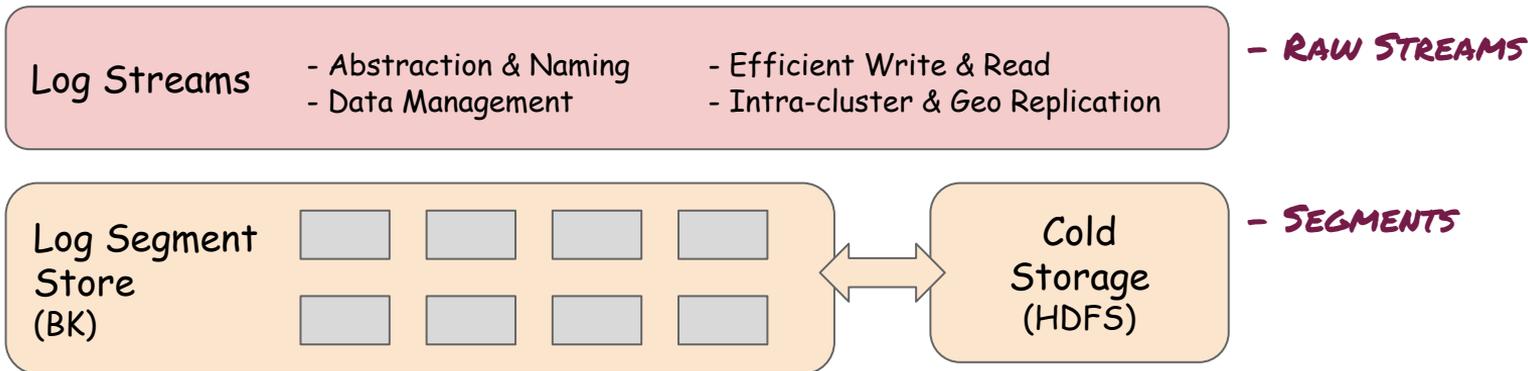
Application



Serving



Storage



Unified Stack

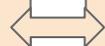
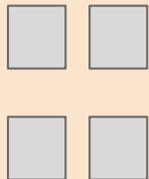
Serving

Apache DistributedLog
(write & read proxies)

Storage

Apache DistributedLog
(core library)

Apache
BookKeeper
(warm data)



HDFS
(cold data)



Hardware + Capacity Planning

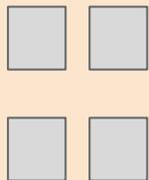
Serving

Apache DistributedLog
(write & read proxies)

Storage

Apache DistributedLog
(core library)

Apache
BookKeeper
(warm data)



HDFS
(cold data)



Deployment - Aurora + Mesos

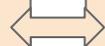
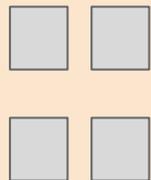
Serving

Apache DistributedLog
(write & read proxies)

Storage

Apache DistributedLog
(core library)

Apache
BookKeeper
(warm data)



HDFS
(cold data)



Apache
AURORA™



MESOS



Deployment - Auto Scaling

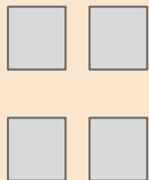
Serving

Apache DistributedLog
(write & read proxies)

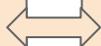
Storage

Apache DistributedLog
(core library)

Apache
BookKeeper
(warm data)



HDFS
(cold data)



Apache
AURORA™



Configuration - ConfigBus

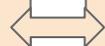
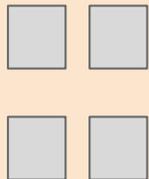
Serving

Apache DistributedLog
(write & read proxies)

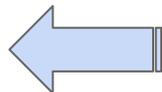
Storage

Apache DistributedLog
(core library)

Apache
BookKeeper
(warm data)



HDFS
(cold data)



System Configuration,
Customer Configuration,
Quota, Rate limiting,
Replication Settings,
...



Configuration - Decider

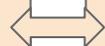
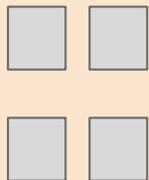
Serving

Apache DistributedLog
(write & read proxies)

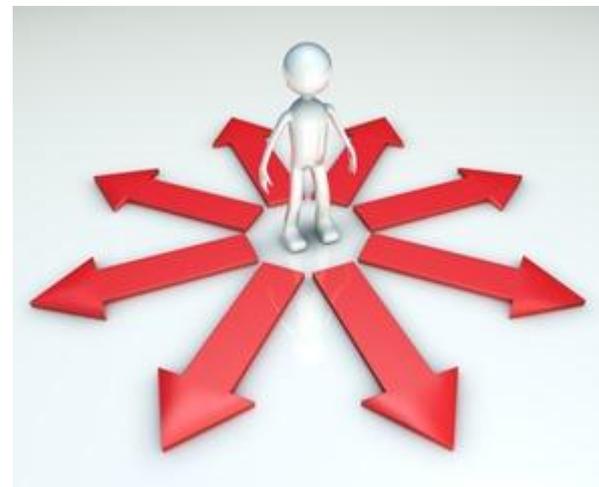
Storage

Apache DistributedLog
(core library)

Apache
BookKeeper
(warm data)



HDFS
(cold data)



Runtime Configurations

- Traffic Routing
- Enable/Disable Features
- Change Quota
- ...



Monitoring - Metrics

Serving

Apache DistributedLog
(write & read proxies)



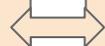
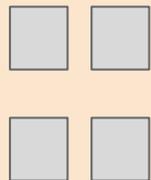
Cuckoo
(TSDB)



Storage

Apache DistributedLog
(core library)

Apache
BookKeeper
(warm data)



HDFS
(cold data)



Monitoring - Alerts

Serving

Apache DistributedLog
(write & read proxies)



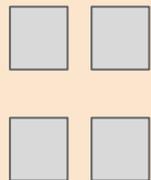
Cuckoo
(TSDB)



Storage

Apache DistributedLog
(core library)

Apache
BookKeeper
(warm data)



HDFS
(cold data)

MonAlerts



Monitoring - Mechanic

Serving

Apache DistributedLog
(write & read proxies)

Storage

Apache DistributedLog
(core library)

Apache
BookKeeper
(warm data)

HDFS
(cold data)

Cuckoo
(TSDB)

MonAlerts

Mechanic



Monitoring - Monitoring Jobs

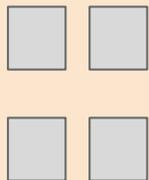
Serving

Apache DistributedLog
(write & read proxies)

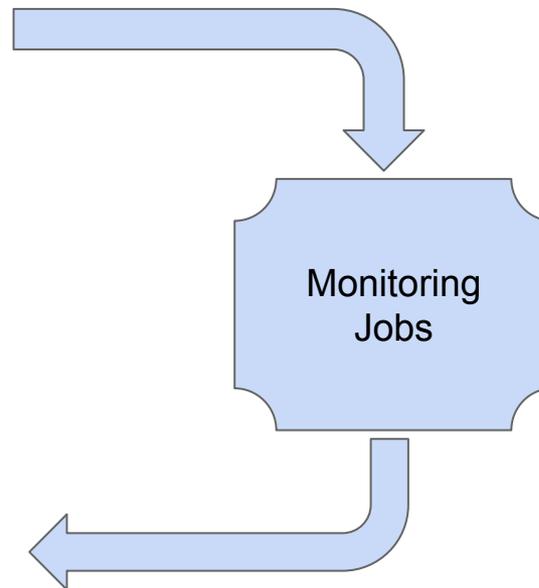
Storage

Apache DistributedLog
(core library)

Apache
BookKeeper
(warm data)



HDFS
(cold data)



Self-Serve

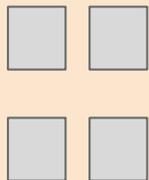
Serving

Apache DistributedLog
(write & read proxies)

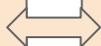
Storage

Apache DistributedLog
(core library)

Apache
BookKeeper
(warm data)



HDFS
(cold data)



Provisioning,
Configurations,
Customer Alerts,
...



Metrics,
Dashboards,
Alerts,
...



Aurora Workflow

- Layered Architecture - Accelerate Development
- Mono Repo
- CI uploads packages on code merged
- Aurora workflow updates jobs
 - Client - Monitoring Jobs, Replicators, ...
 - Proxies
 - Storage (rarely)



Team

- SWE + SRE
- SWE
 - Development, Deployment, Operations, Oncall Duties
- SRE
 - Automate everything
 - Deployment, Operations, Oncall Duties
 - Capacity Planning
- Oncall - Primary + Secondary



Support

- Self-Serve Portal
- Helpdesk (wiki, hipchat, mail lists)
- Office Hours

