Big Data Engineering

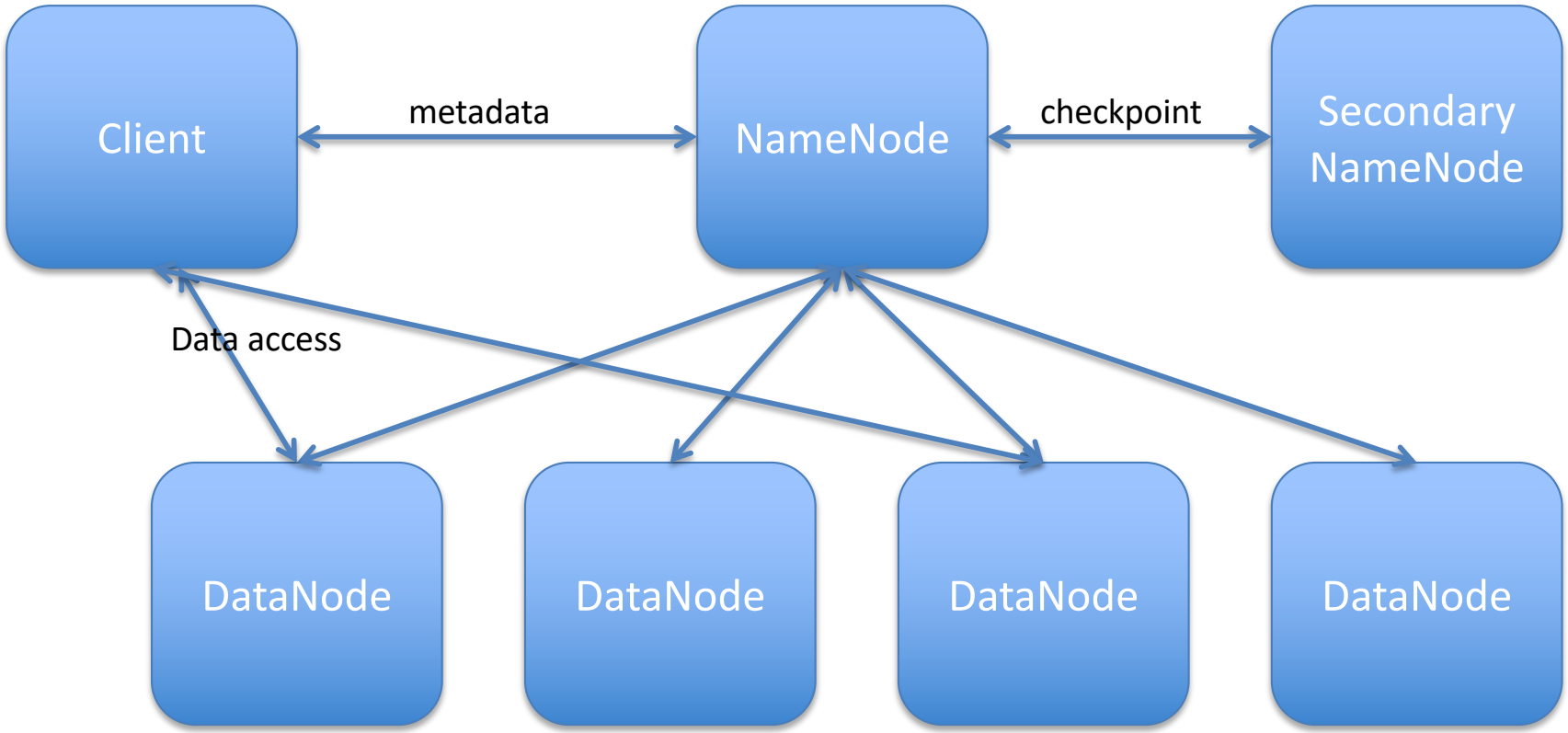
Other Tools and Libraries

Julie Weeds

March 2020

© Paul Fremantle 2015. This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License
See http://creativecommons.org/licenses/by-nc-sa/4.0/

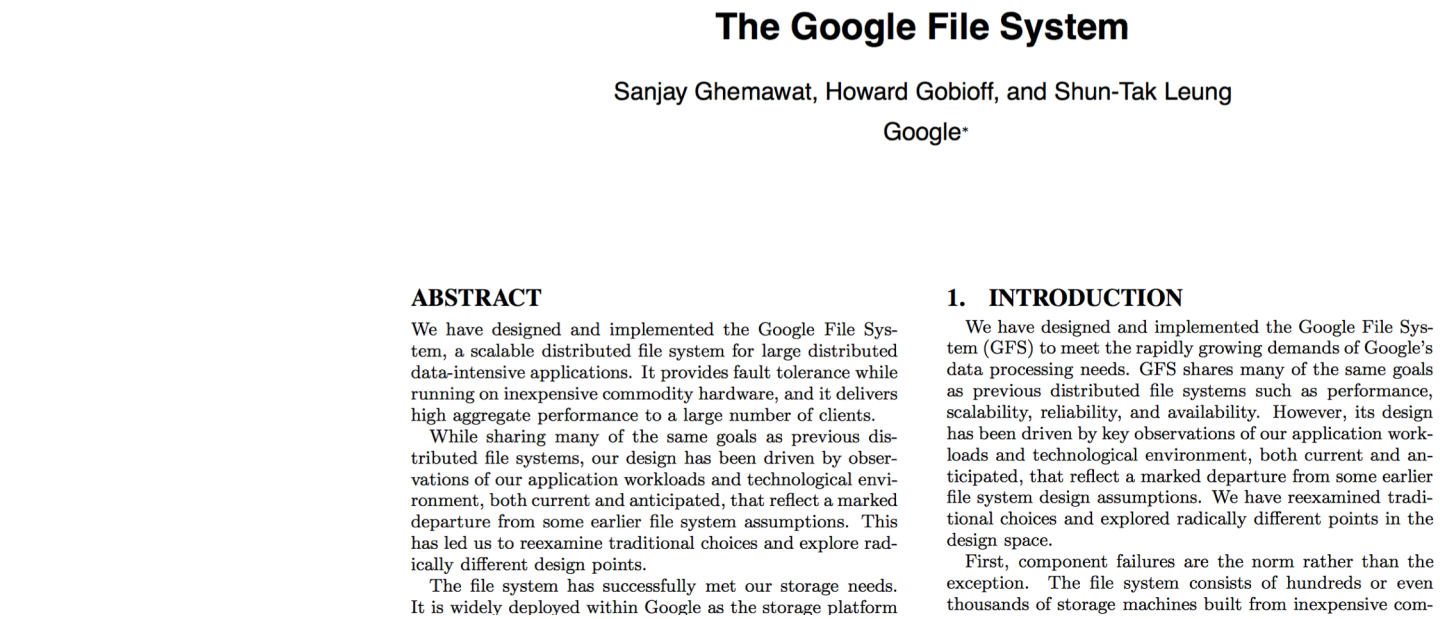
HDFS in a nutshell



© Paul Fremantle 2015. This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License
See http://creativecommons.org/licenses/by-nc-sa/4.0/

HDFS inspiration

- Google File System
  - Sanjay Ghemawat, Howard Gobioff, and Shun-Tak Leung. 2003. The Google file system. In Proceedings of the nineteenth ACM symposium on Operating systems principles (SOSP '03). ACM, New York, NY, USA, 29-43.



The Google File System

Sanjay Ghemawat, Howard Gobioff, and Shun-Tak Leung

Google*

ABSTRACT

We have designed and implemented the Google File System, a scalable distributed file system for large distributed data-intensive applications. It provides fault tolerance while running on inexpensive commodity hardware, and it delivers high aggregate performance to a large number of clients.

While sharing many of the same goals as previous distributed file systems, our design has been driven by observations of our application workloads and technological environment, both current and anticipated, that reflect a marked departure from some earlier file system assumptions. This has led us to reexamine traditional choices and explore radically different design points.

The file system has successfully met our storage needs. It is widely deployed within Google as the storage platform

1. INTRODUCTION

We have designed and implemented the Google File System (GFS) to meet the rapidly growing demands of Google's data processing needs. GFS shares many of the same goals as previous distributed file systems such as performance, scalability, reliability, and availability. However, its design has been driven by key observations of our application workloads and technological environment, both current and anticipated, that reflect a marked departure from some earlier file system design assumptions. We have reexamined traditional choices and explored radically different points in the design space.

First, component failures are the norm rather than the exception. The file system consists of hundreds or even thousands of storage machines built from inexpensive com-

© Paul Fremantle 2015. This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License
See http://creativecommons.org/licenses/by-nc-sa/4.0/

HDFS overview

- Good for streaming access to large files, reliability, scale
- Not good for random access, small files
- Blocks of data 64Mb in size (configurable)
- Each block can be replicated across multiple data nodes for High Availability (HA)

© Paul Fremantle 2015. This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License
See http://creativecommons.org/licenses/by-nc-sa/4.0/

HDFS Usage

- Spotify has 1600+ nodes, storing 60+ petabytes of data
  - https://www.usenix.org/system/files/conference/fast17/fast17-niazi.pdf
- One of Facebook's largest clusters (based on HDFS) holds more than
  100 PB of data, processing more than 60,000 Hive queries a day
  - https://www.facebook.com/notes/facebook-engineering/under-the-hood-scheduling-mapreduce-jobs-more-efficiently-with-corona/

© Paul Fremantle 2015. This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License
See http://creativecommons.org/licenses/by-nc-sa/4.0/

HopFS

- HopFS is a drop-in replacement for HDFS, based on HDFS v2.0.4.

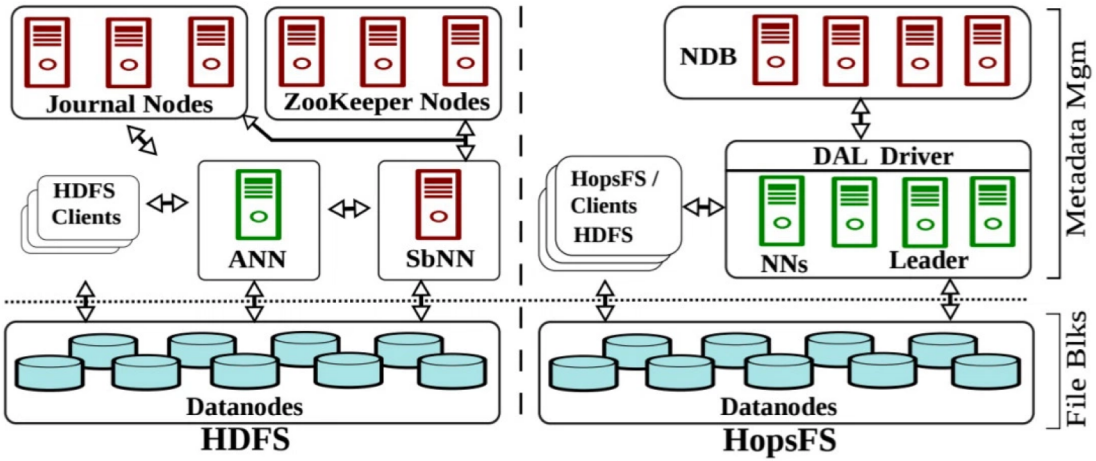


Figure 1: System architecture for HDFS and HopsFS. For high availability, HDFS requires an Active NameNode (ANN), at least one Standby NameNode (SbNN), at least three Journal Nodes for quorum-based replication of the write ahead log of metadata changes, and at least three ZooKeeper instances for quorum based coordination. HopsFS supports multiple stateless namenodes that access the metadata stored in NDB database nodes.

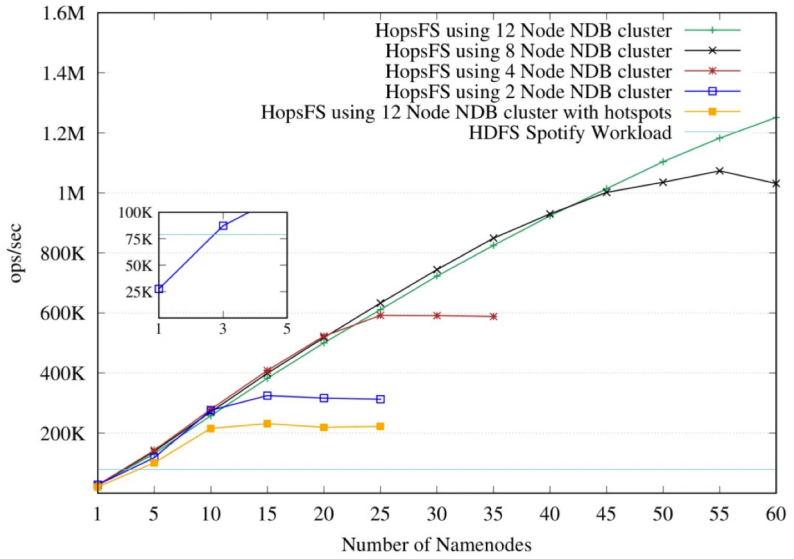
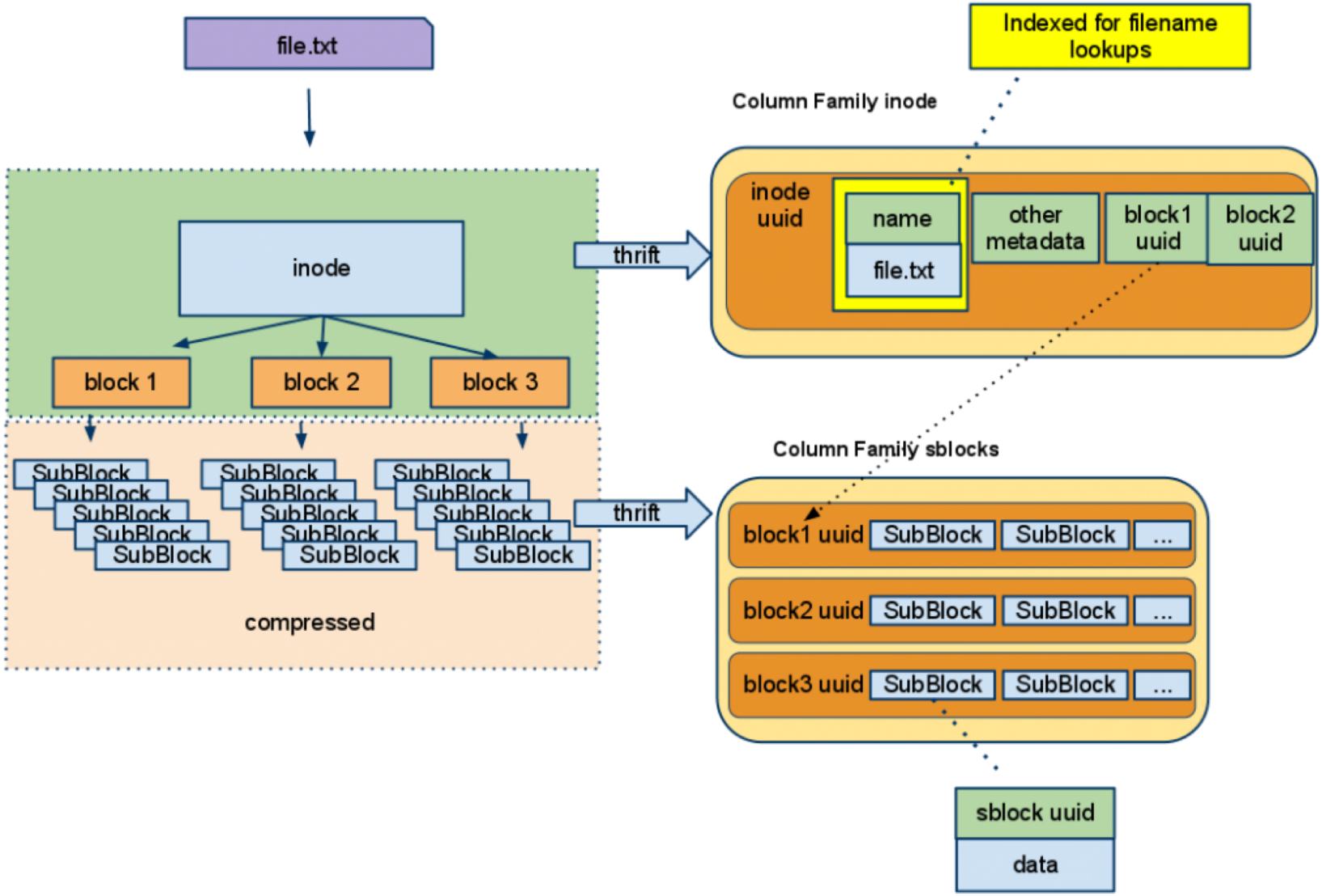


Figure 6: HopsFS and HDFS throughput for Spotify workload.

CassandraFS
(not open source)



© Paul Fremantle 2015. This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License
See http://creativecommons.org/licenses/by-nc-sa/4.0/

Amazon S3

- Simple Storage Service
- Unlimited storage of files
  - Up to 5 terabytes each
  - Stored in named “buckets”
  - Accessible via AWS APIs or HTTP
  - Authenticated or Public

© Paul Fremantle 2015. This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License
See http://creativecommons.org/licenses/by-nc-sa/4.0/

Spark packages

- A wide set of plugins
  - Currently 148 community donated plugins
- Data connectors
  - Cassandra, Couchbase, Mongo, CSV, etc
- Machine Learning, Neural networks
- Streaming
- etc

© Paul Fremantle 2015. This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License
See http://creativecommons.org/licenses/by-nc-sa/4.0/

Using Spark Packages

Automatic download from the web:

```
bin/spark-shell
    --packages com.databricks:spark-csv_2.11:1.2.0
```

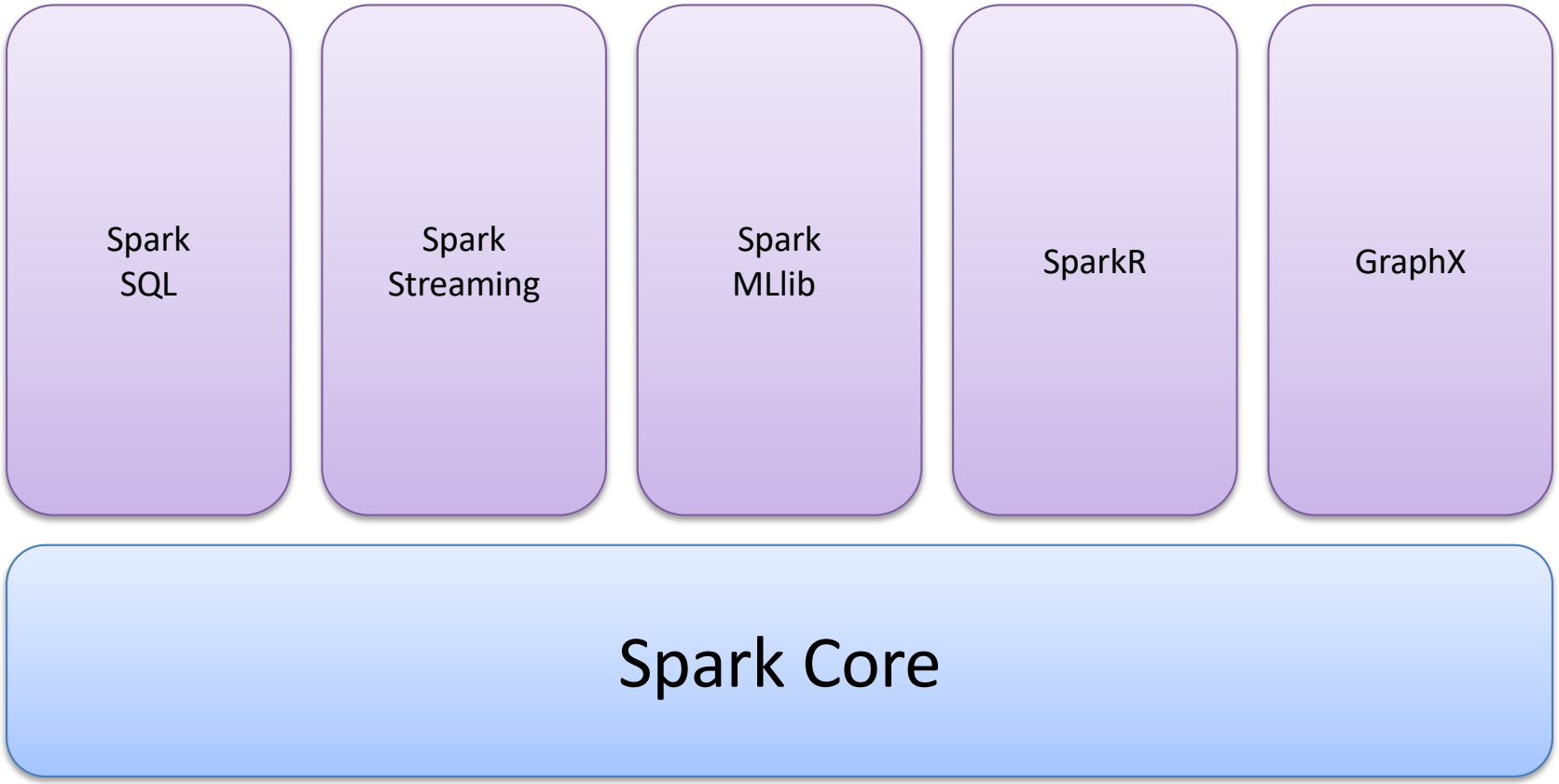
© Paul Fremantle 2015. This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License
See http://creativecommons.org/licenses/by-nc-sa/4.0/

Locality

- Spark understands the locality of data:
  - Already in memory
  - HDFS location
  - Cassandra location

© Paul Fremantle 2015. This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License
See http://creativecommons.org/licenses/by-nc-sa/4.0/

Spark Extras



© Paul Fremantle 2015. This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License
See http://creativecommons.org/licenses/by-nc-sa/4.0/

Spark Extras

- Spark Streaming
  - Realtime analysis in Spark
- Spark MLLib
  - Like Mahout – Machine learning in Spark
- GraphX
  - Graph processing in Spark
- SparkR
  - R statistical analysis on Spark

© Paul Fremantle 2015. This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License
See http://creativecommons.org/licenses/by-nc-sa/4.0/

Spark MLlib

- Simple stats and correlation testing
- Classification and regression
- Collaborative Filtering
  - Alternating Least Squares
- Clustering
  - k-means, etc
- Frequent Pattern Mining
- Plus more

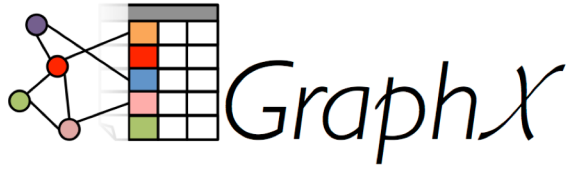
© Paul Fremantle 2015. This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License
See http://creativecommons.org/licenses/by-nc-sa/4.0/

MLlib example

```
from pyspark.mllib.fpm import FPGrowth

data = sc.textFile("data/mllib/sample_fpgrowth.txt")

transactions = data.map(lambda line: line.strip().split(' '))

model = FPGrowth.train(transactions, minSupport=0.2,
    numPartitions=10)

result = model.freqItemsets().collect()
for fi in result:
    print(fi)
```

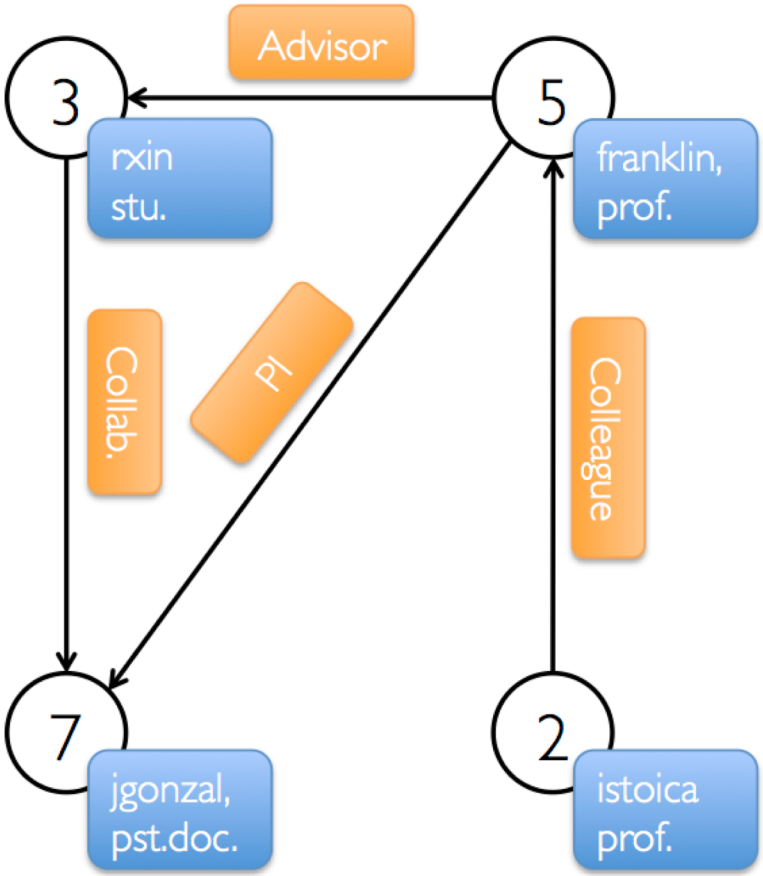
© Paul Fremantle 2015. This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License
See http://creativecommons.org/licenses/by-nc-sa/4.0/

GraphX



-

Property Graph



Vertex Table

Id	Property (V)
3	(rxin, student)
7	(jgonzal, postdoc)
5	(franklin, professor)
2	(istoica, professor)

Edge Table

SrcId	DstId	Property (E)
3	7	Collaborator
5	3	Advisor
2	5	Colleague
5	7	PI

© Paul Fremantle 2015. This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License
See http://creativecommons.org/licenses/by-nc-sa/4.0/

R

- R is an open source system for statistics and graphics
  - Based on the S language from AT&T Bell Labs
- Supports a wide variety of statistical techniques and graphing tools
- An extensible set of packages that provide extra functions via CRAN
  - The Comprehensive R Archive Network

© Paul Fremantle 2015. This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License
See http://creativecommons.org/licenses/by-nc-sa/4.0/

SparkR

- A lightweight approach to use Spark from within R
- Also works with MLlib for machine learning
- Allows complex statistical analysis to be done on a Spark cluster

© Paul Fremantle 2015. This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License
See http://creativecommons.org/licenses/by-nc-sa/4.0/

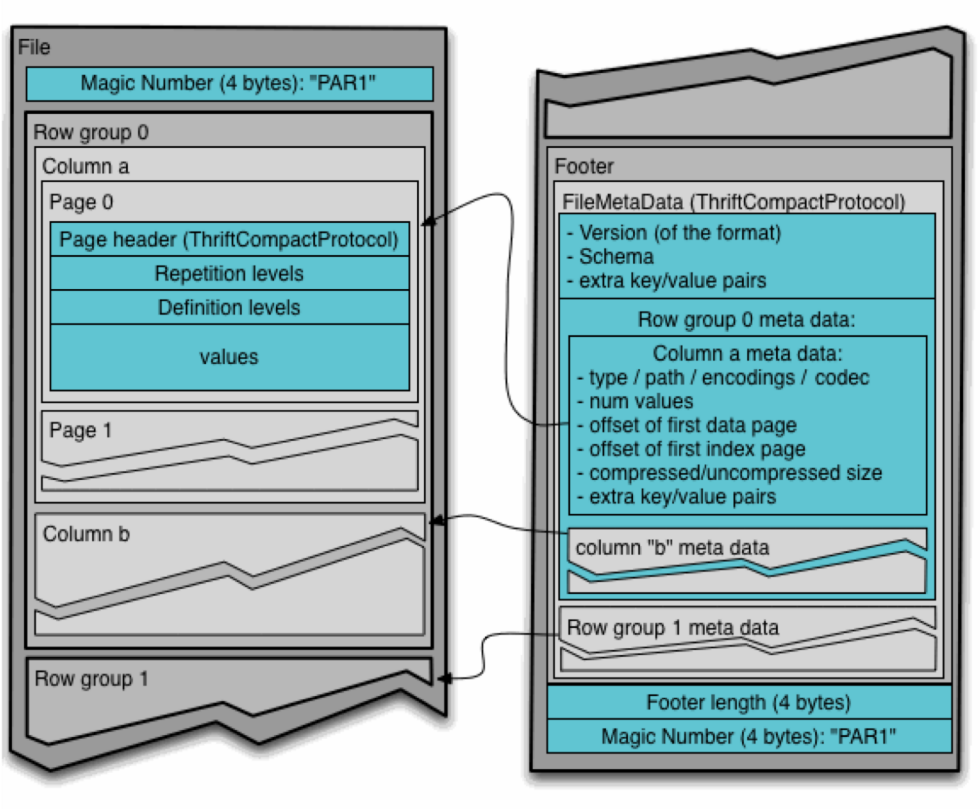
Apache Avro

- A compact data storage and transmission system
  - Uses schemas of data to ensure it can be read by the receiver
  - Supports dynamic typing
- Used by RPC or data collection systems
  - Fast binary protocols
- Also supports storage
  - Hence used by many Big Data apps including Hadoop and Spark

© Paul Fremantle 2015. This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License
See http://creativecommons.org/licenses/by-nc-sa/4.0/

Apache Parquet

- Apache Parquet is a columnar data storage model
  - Works with Hadoop, Spark and many others
  - Efficient storage of data
  - Based on another Google system called Dremel



© Paul Fremantle 2015. This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License
See http://creativecommons.org/licenses/by-nc-sa/4.0/

Cluster management systems for Big Data

- YARN
  - Part of Hadoop but significantly rebuilt since Hadoop 1
- Mesos
  - Popular Apache project
  - Built to be a resource manager for a complete datacenter
    - Supports many workloads (e.g. Docker as well as Spark)

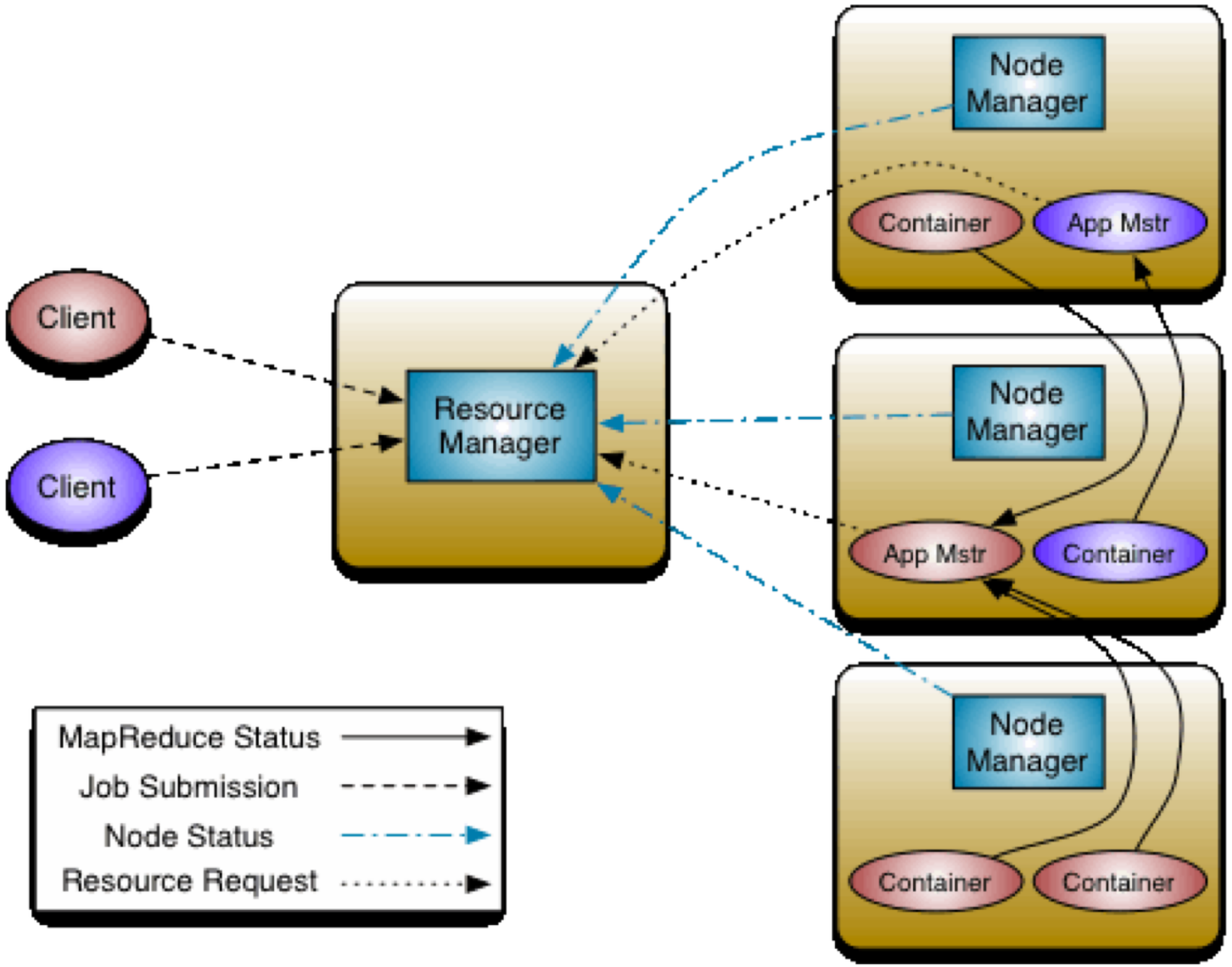
© Paul Fremantle 2015. This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License
See http://creativecommons.org/licenses/by-nc-sa/4.0/

What is YARN?

- **YARN is the system that runs your code on multiple nodes**
- Hadoop 2.0 replacement for the cluster manager
  - Basically a model to distribute and manage workloads
  - Not just MapReduce but supports other workloads

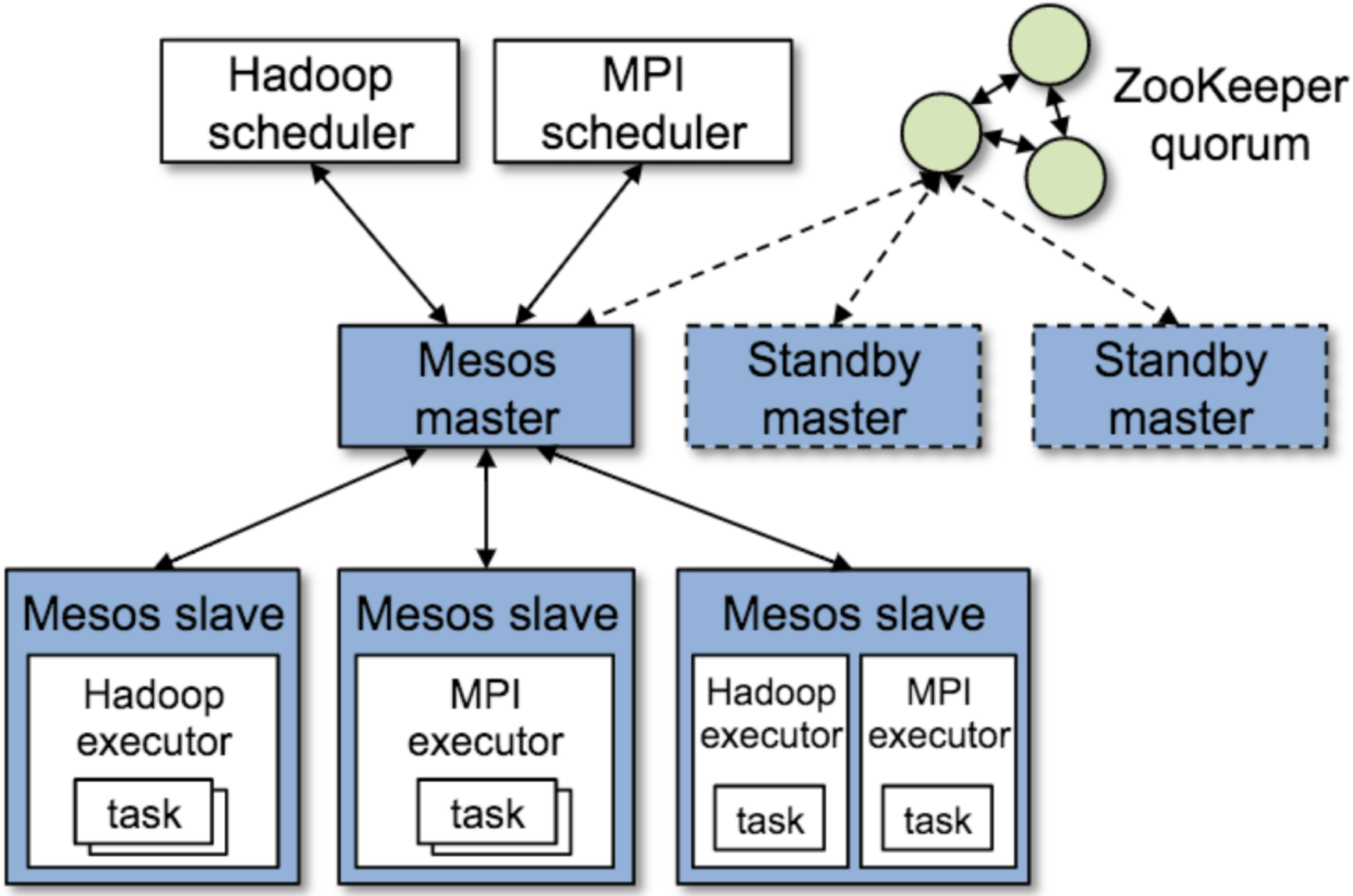
© Paul Fremantle 2015. This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License
See http://creativecommons.org/licenses/by-nc-sa/4.0/

YARN architecture



© Paul Fremantle 2015. This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License
See http://creativecommons.org/licenses/by-nc-sa/4.0/

Apache Mesos



Amazon Web Services

Compute

- **EC2** — Virtual Servers in the Cloud
- **EC2 Container Service** — Run and Manage Docker Containers
- **Elastic Beanstalk** — Run and Manage Web Apps
- **Lambda** — Run Code in Response to Events

Storage & Content Delivery

- **S3** — Scalable Storage in the Cloud
- **CloudFront** — Global Content Delivery Network
- **Elastic File System** PREVIEW — Fully Managed File System for EC2
- **Glacier** — Archive Storage in the Cloud
- **Import/Export Snowball** — Large Scale Data Transport
- **Storage Gateway** — Integrates On-Premises IT Environments with Cloud Storage

Database

- **RDS** — Managed Relational Database Service
- **DynamoDB** — Predictable and Scalable NoSQL Data Store
- **ElastiCache** — In-Memory Cache
- **Redshift** — Managed Petabyte-Scale Data Warehouse Service

Developer Tools

- **CodeCommit** — Store Code in Private Git Repositories
- **CodeDeploy** — Automate Code Deployments
- **CodePipeline** — Release Software using Continuous Delivery

Management Tools

- **CloudWatch** — Monitor Resources and Applications
- **CloudFormation** — Create and Manage Resources with Templates
- **CloudTrail** — Track User Activity and API Usage
- **Config** — Track Resource Inventory and Changes
- **OpsWorks** — Automate Operations with Chef
- **Service Catalog** — Create and Use Standardized Products
- **Trusted Advisor** — Optimize Performance and Security

Security & Identity

- **Identity & Access Management** — Manage User Access and Encryption Keys
- **Directory Service** — Host and Manage Active Directory
- **Inspector** PREVIEW — Analyze Application Security
- **WAF** — Filter Malicious Web Traffic

Internet of Things

- **AWS IoT** BETA — Connect Devices to the cloud

Mobile Services

- **Mobile Hub** BETA — Build, Test, and Monitor Mobile apps
- **Cognito** — User Identity and App Data Synchronization
- **Device Farm** — Test Android, Fire OS, and iOS apps on real devices in the Cloud
- **Mobile Analytics** — Collect, View and Export App Analytics
- **SNS** — Push Notification Service

Application Services

- **API Gateway** — Build, Deploy and Manage APIs
- **AppStream** — Low Latency Application Streaming
- **CloudSearch** — Managed Search Service
- **Elastic Transcoder** — Easy-to-use Scalable Media Transcoding
- **SES** — Email Sending Service
- **SQS** — Message Queue Service
- **SWF** — Workflow Service for Coordinating Application Components

Enterprise Applications

© Paul Fremantle 2015. This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License
See http://creativecommons.org/licenses/by-nc-sa/4.0/

EC2 / AWS main functions

- EC2 (Elastic Compute Cloud)
  - Instances
    - Servers of various sizes
  - AMIs (Amazon Machine Images)
    - Server images
  - Elastic Block Storage (EBS)
    - Virtualized Hard drives
  - VPC (Virtual Private Cloud)
    - Secure network space
- S3 (Simple Storage Solution)
  - “Buckets” of data
  - Longer term storage of data

© Paul Fremantle 2015. This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License
See http://creativecommons.org/licenses/by-nc-sa/4.0/

AWS - EMR

- EMR (Elastic Map Reduce)
  - runs on top of EC2
  - provides straightforward interface for launching clusters and running Spark
  - lab will go through the process of starting a cluster with Spark and other python packages installed, running Jupyter notebooks on the cluster, and introduce SSH connection for more intensive python scripts

© Paul Fremantle 2015. This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License
See http://creativecommons.org/licenses/by-nc-sa/4.0/

Questions?

© Paul Fremantle 2015. This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License
See http://creativecommons.org/licenses/by-nc-sa/4.0/