Survival Tips for Big Data Impact on Performance

Share Pittsburgh Session 15404

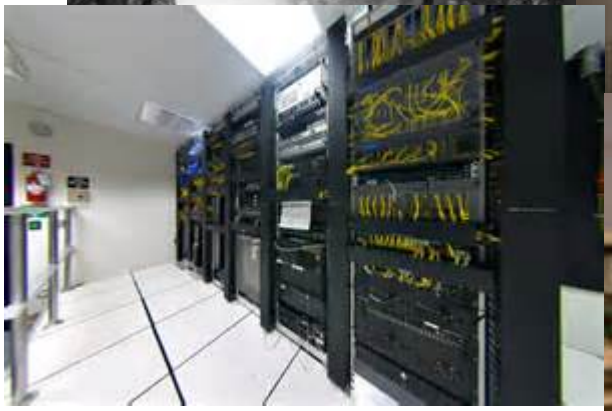
Laura Knapp
WW Business Consultant
Laurak@aesclever.com
ipv6hawaii@outlook.com

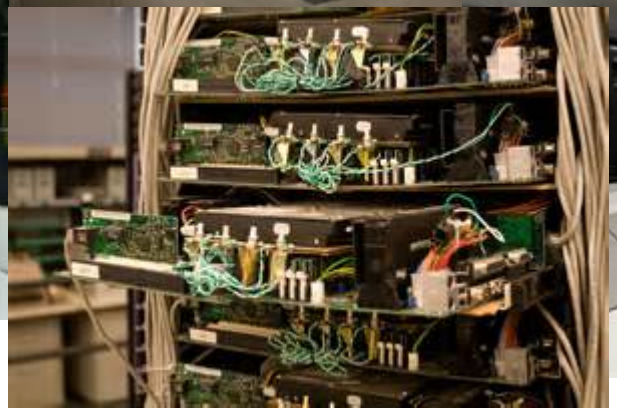
Background

Hadoop Example

Closing Thoughts

The Past





New Device Types

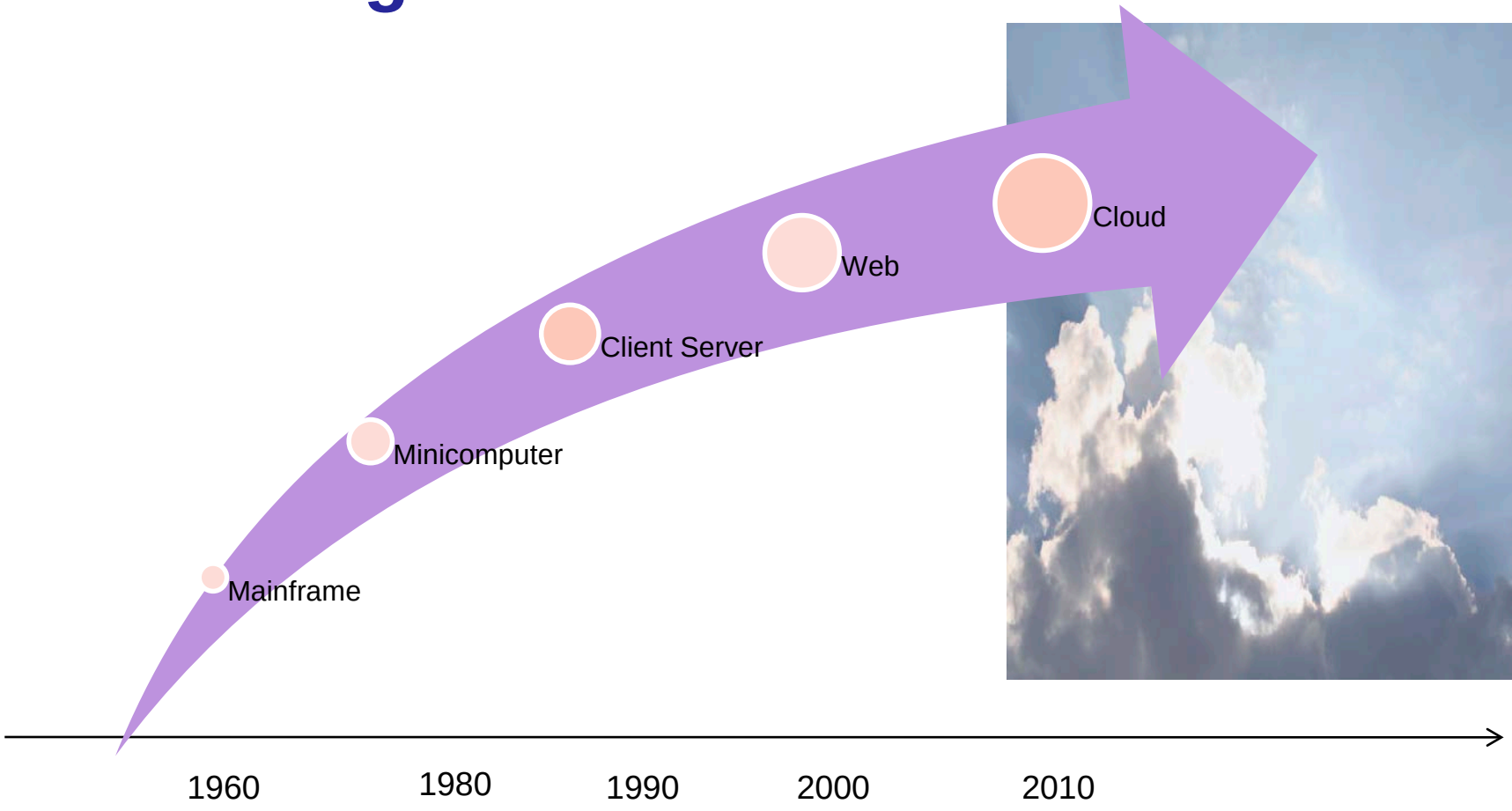
Always Connected

Big Data is the Intersection of Cloud, Social, and Mobile

Data Means More

1990's	*2000's*	*2010's*
Generated Internally	Initial Internet Generation	Generated Externally
Key to Operational Efficiency	Key to Business Transformation	Key to Competitive Advantage
	Product Expansion	Source of Product Innovation
		Integration from Multiple Sources
		Real-Time
		Changing our Lives

Path to Big Data



Three Characteristics of Big Data

Volume

- Typical PC has 50 gigabytes of data
- Facebooks ingests 500 terabytes of new data everyday
- Boeing 737 generates 240 terabytes of flight data during a singe flight
- Infrastructure and sensors generate massive log data in real-time
- Add smart phones, sensors, internet of things – becomes mindboggling

Velocity

- Clickstreams capture user data at millions of events per second
- Stock trading algorithms reflect market changes in microseconds
- Machine-machine processes exchange data between billions of devices
- On-line gaming systems support millions of concurrent users each producing multiple inputs per second

Variety

- Geospatial data
- 3D data
- Audio
- Video
- Unstructured
- GPS
- RFID
- Structured

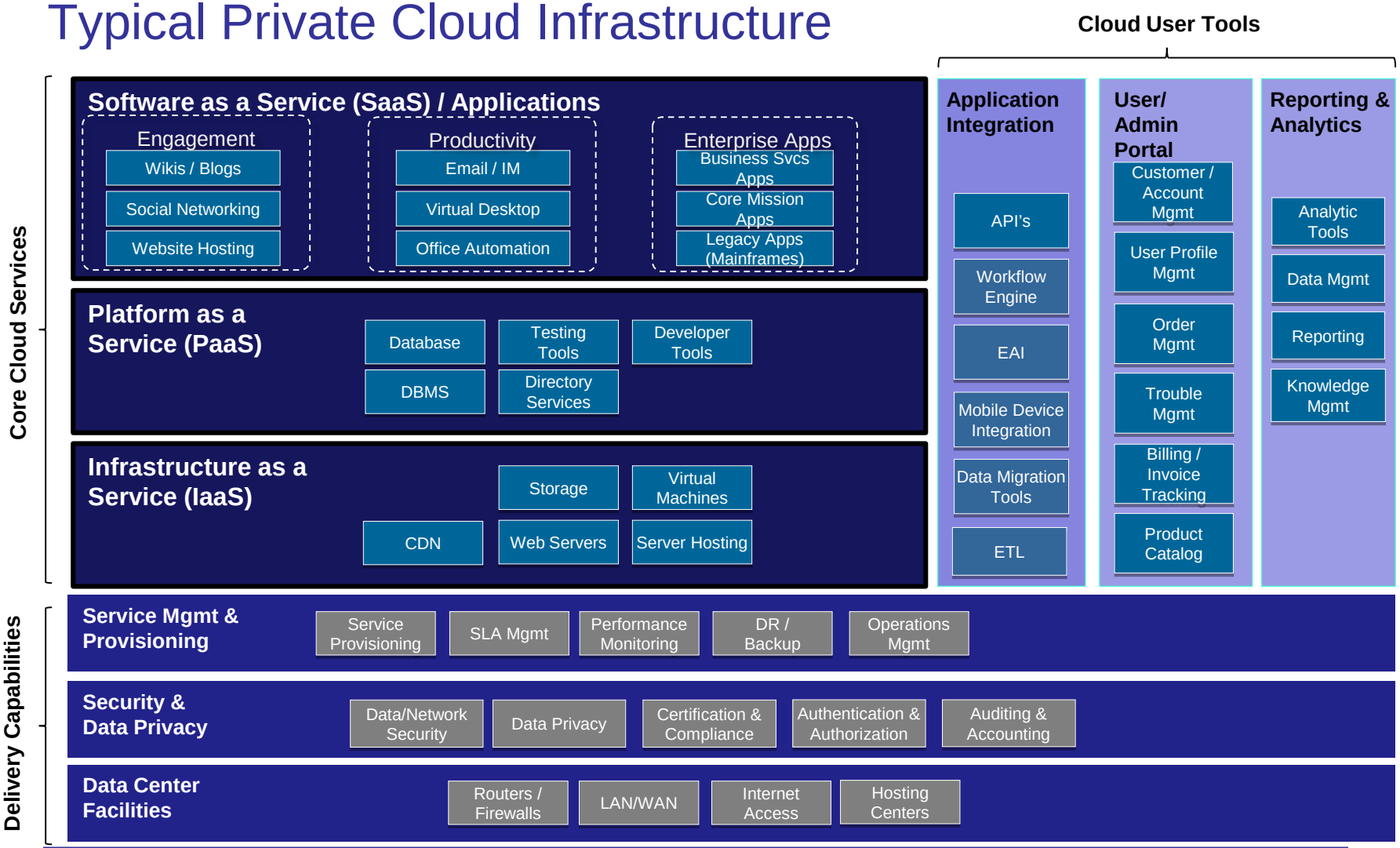
Types of Data

- Structured Data
  - Schema Based
  - Defined Semantics by Raw/Column Oriented Data Structure
  - Records Retrieved Through key-value pair
- Unstructured Data & Semi-structured Data
  - Text, Logs, Web-Clicks, Sensors Data, Picture, Video
  - Each have structure or semi-structure, e.g., TXT file has space, tab, semi-colon etc..
  - When various data-types combined in one data set, it become unstructured in sense of semantics
- Data warehouse & RDBMS
  - Oracle, Sybase, SQL are structured relational data bases
  - ERP and CRM Feeds
  - Traditional Transactional (OLTP) & Reporting(OLAP) via ETL, API Access

Cloud Services for Big Data

Model	Capability Provided	Example Services
SaaS (End User)	To use the provider's applications running on a cloud infrastructure and accessible from various client devices through a thin client interface such as a Web browser	▪ Citizen Engagement (Wikis, Blogs, Data.gov) ▪ Government Productivity (Cloud based tools) ▪ Business Enablement (Salesforce.com) ▪ Enterprise Applications (Core Mission & Business Svcs
PaaS (Application Developers)	To deploy onto the cloud infrastructure consumer-created applications using programming languages and tools supported by the provider (e.g., java, python, .Net)	▪ Database and Database Management Systems ▪ Developer / Testing Tools ▪ Virtual Environments
IaaS (Network Architects)	To provision processing, storage, networks, and other fundamental computing resources where the consumer is able to deploy and run arbitrary software, which can include operating systems and applications	• Computing • Storage • Application hosting

Typical Private Cloud Infrastructure



Cloud Functions for Big Data

User Access

Web-based Joint

Collaboration

Real-time voice, text, video, application

Discovery

Locate specific information and services

Messaging

Real-time updates & alert notifications as data change

Metadata Discovery

Ability to discover, develop & reuse data semantics

Mediation

Exchange data with unanticipated users & formats

Business

Content Delivery

Responsive

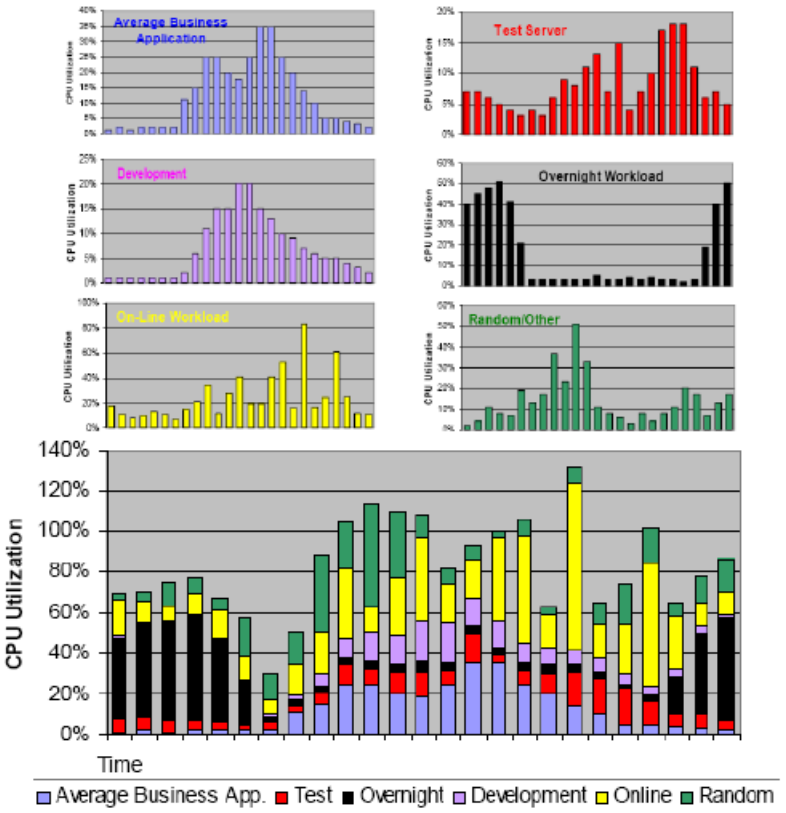
Enterprise Service Management

Monitors services availability, performance & reliability

Security

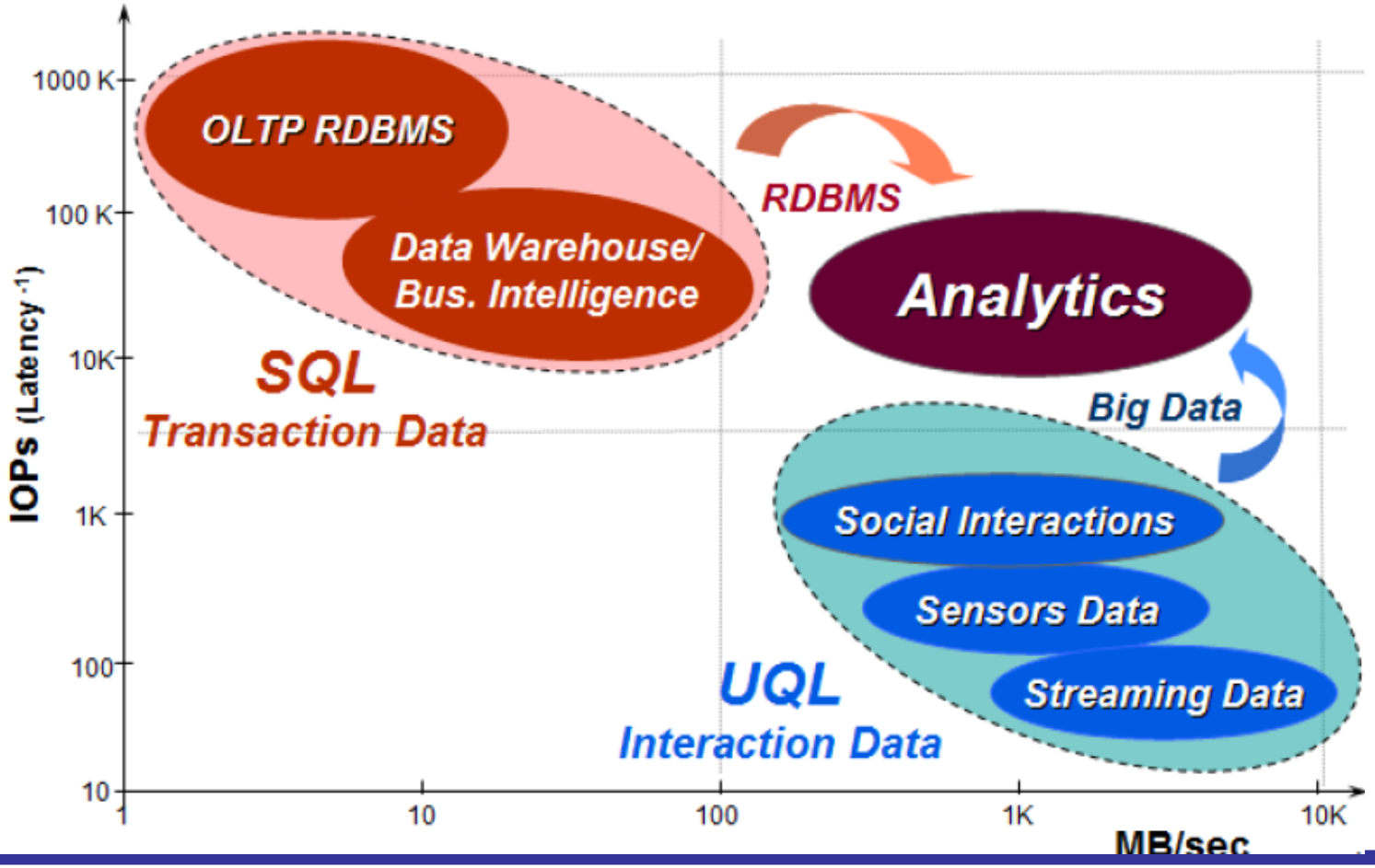
Ability to operate in a secure environment

Workload Considerations

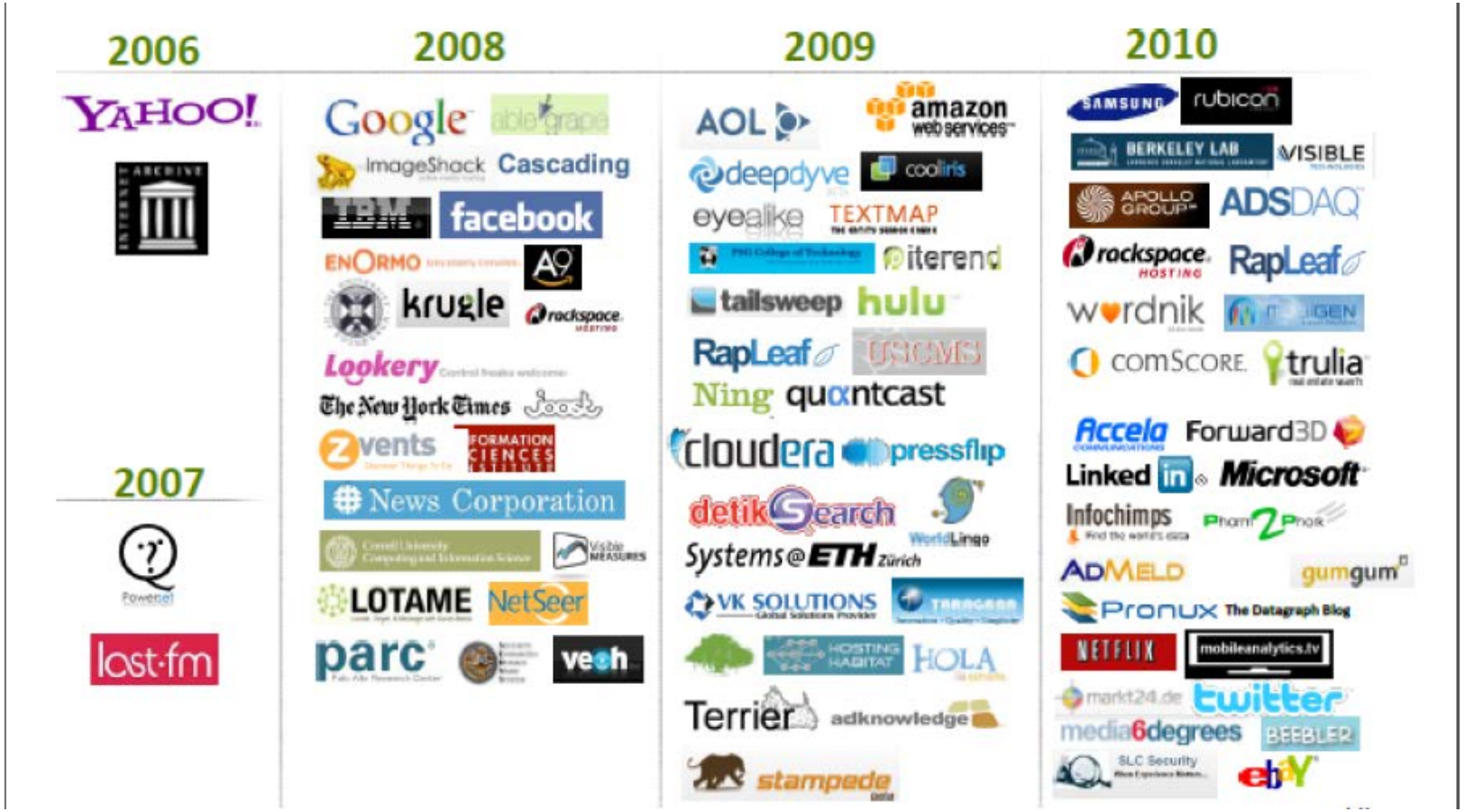


I/O usage
Inter-arrival times
Service demand
Classes of data
System Load (CPU, Buffers)
Frequency distribution
Geographic orientation
Number of arrivals
Number of executing programs
Resource usage

Workload Characterization



Big Data Adoption



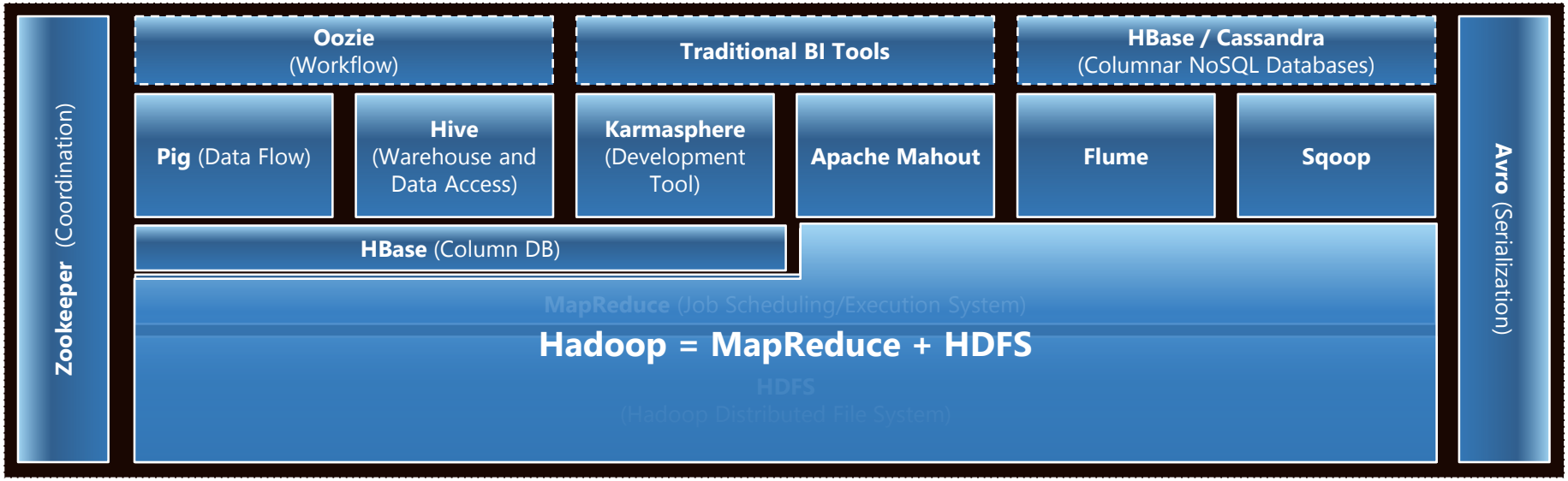
Background

Hadoop Example

Closing Thoughts

What is Hadoop?

New scale out system for processing peta-scale data
Ideal for unstructured data
Changes economics of large scale data processing
- Through the use of commodity servers



Hadoop Operations

Scalable & Fault Tolerant

Data **is not centrally located**

Data is stored across all data nodes in the cluster

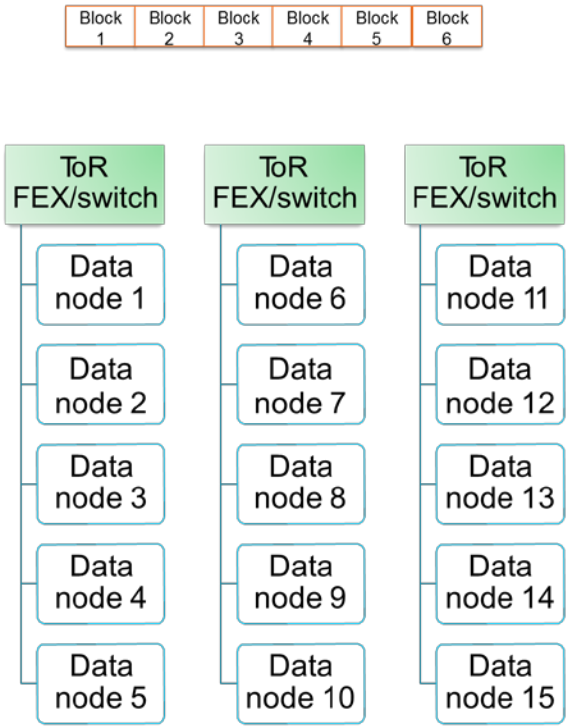
Data is divided in multiple **large blocks** – 64MB default, typical block 128MB

Blocks are not the related to disk geometry

Data **is stored reliably-** replicated 3 times

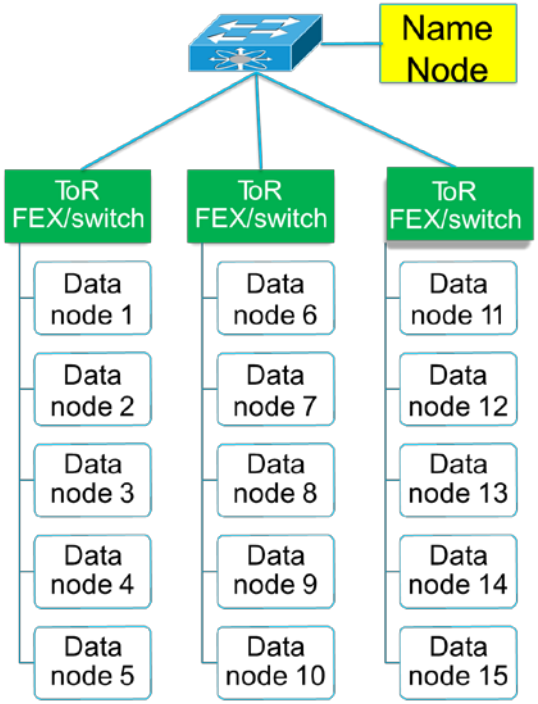
Types of Functions

- Name Node (Master) - Manages Cluster
- Data Node (Map and Reducer) – Carries blocks



Hadoop Components

- Name Node
  - Runs a scheduler – Job Tracker
  - Manages all data nodes, in memory
  - Secondary Name Node – Snapshot of meta data of HDFS cluster
  - Typically all three JVM can run on single node
- Data Node
  - Task Tracker Receives Job Info from Job Tracker(Name Node)
  - Map & Reducer Task Managed by Task Tracker
  - Configurable Ratio of Map & Reduce Task for various workload per Node/CPU/Core
  - **Data Locality** - IF data not available where the map task is assigned, a missing block be copied over the network



Hadoop and the Network

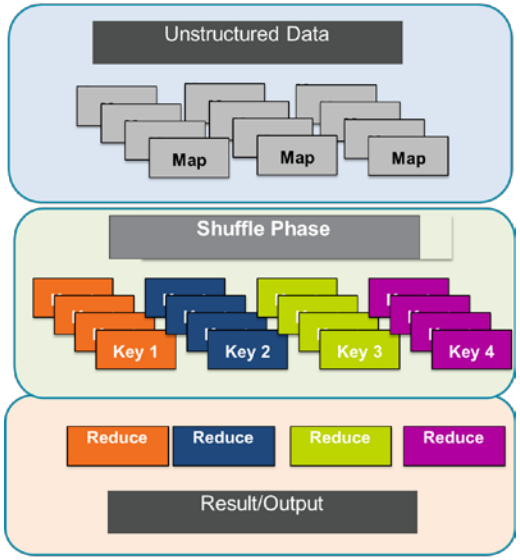
Data gathering and replication
External network activity

Map Phase -Raw data analyzed and converted to name/value pair
High I/O and compute functions

Shuffle Phase -Name/Value pairs sorted and grouped
Reducer is puller the data from the mapper nodes

Reducer Phase -All values associated with a key are processed

Result/Output Phase -Reducer replicates results to nodes
High network activity with usage depending on workload behavior



Characteristics that Affect Hadoop Clusters

- **Cluster Size**
  - Number of Data Nodes
- **Data Model & Mapper/Reduces Ratio**
  - MapReduce functions
- **Input Data Size**
  - Total starting dataset
- **Data Locality in HDFS**
  - Ability to processes data where it already is located
- **Background Activity**
  - Number of Jobs running
  - type of jobs
  - Importing
  - exporting

Characteristics of Data Node

- I/O, CPU, Memory etc.

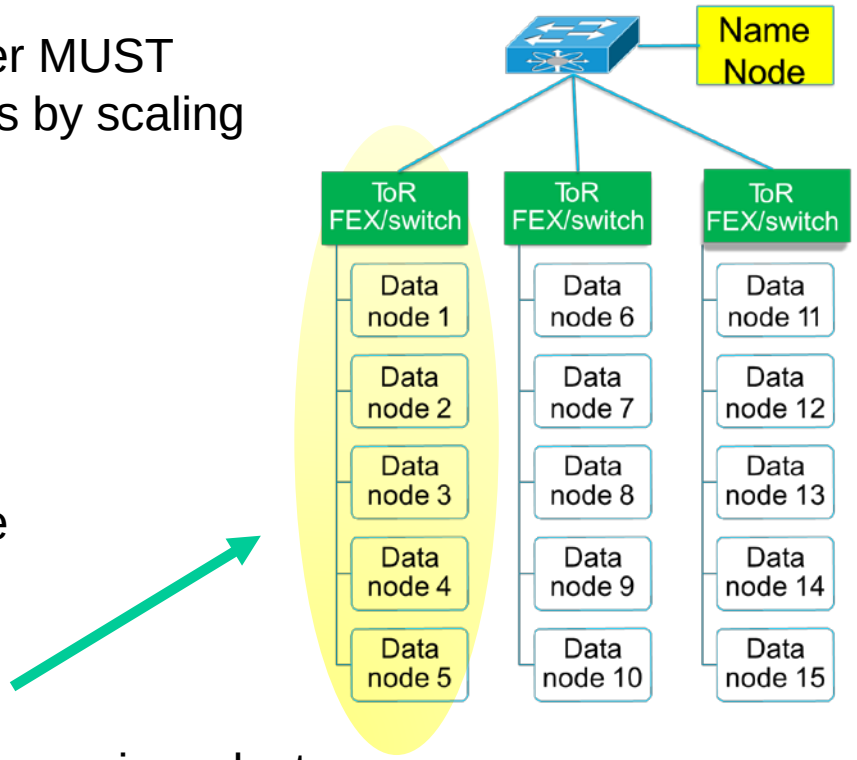
- **Networking Characteristics**
  - Availability
  - Buffering
  - Data Node Speed (1G vs. 10G)
  - Oversubscription
  - Latency

Cluster Size

An optimally configured cluster MUST decrease job completion times by scaling out the nodes

Sizing Depends on

- Workload Size & Type
- Job completion time
- Map/Reduce Ratio
- CPU, IO and local storage

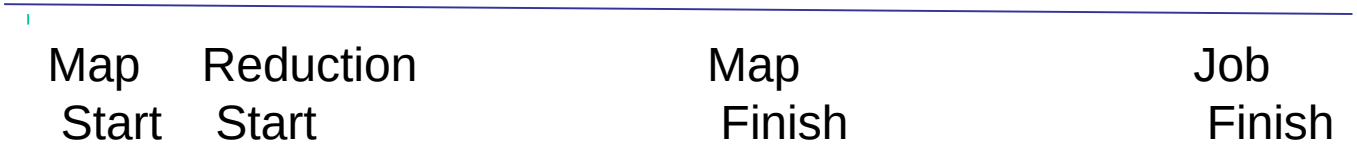


Cluster – more efficient data access in a cluster than going between racks

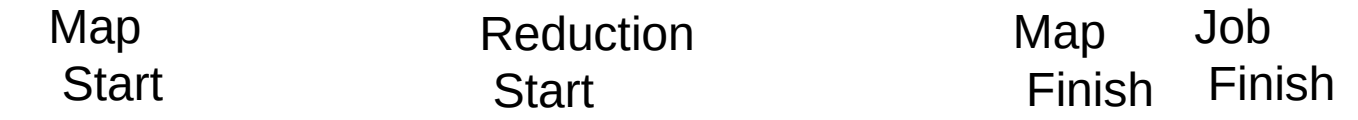
Map Reduce Model

Complexity of functions used in Map and/or Reduce has a large impact on the job completion time and network traffic

ETL Workload is very network intensive
Input, shuffle and output data size is the same



Shakespeare word count (BI workload)
Data set size varies with varying impact on the network



Traffic Types

Small Flows/Messaging

(Admin Related, Heart-beats, Keep-alive, delay sensitive application messaging)

Small – Medium Incast

(Hadoop Shuffle)

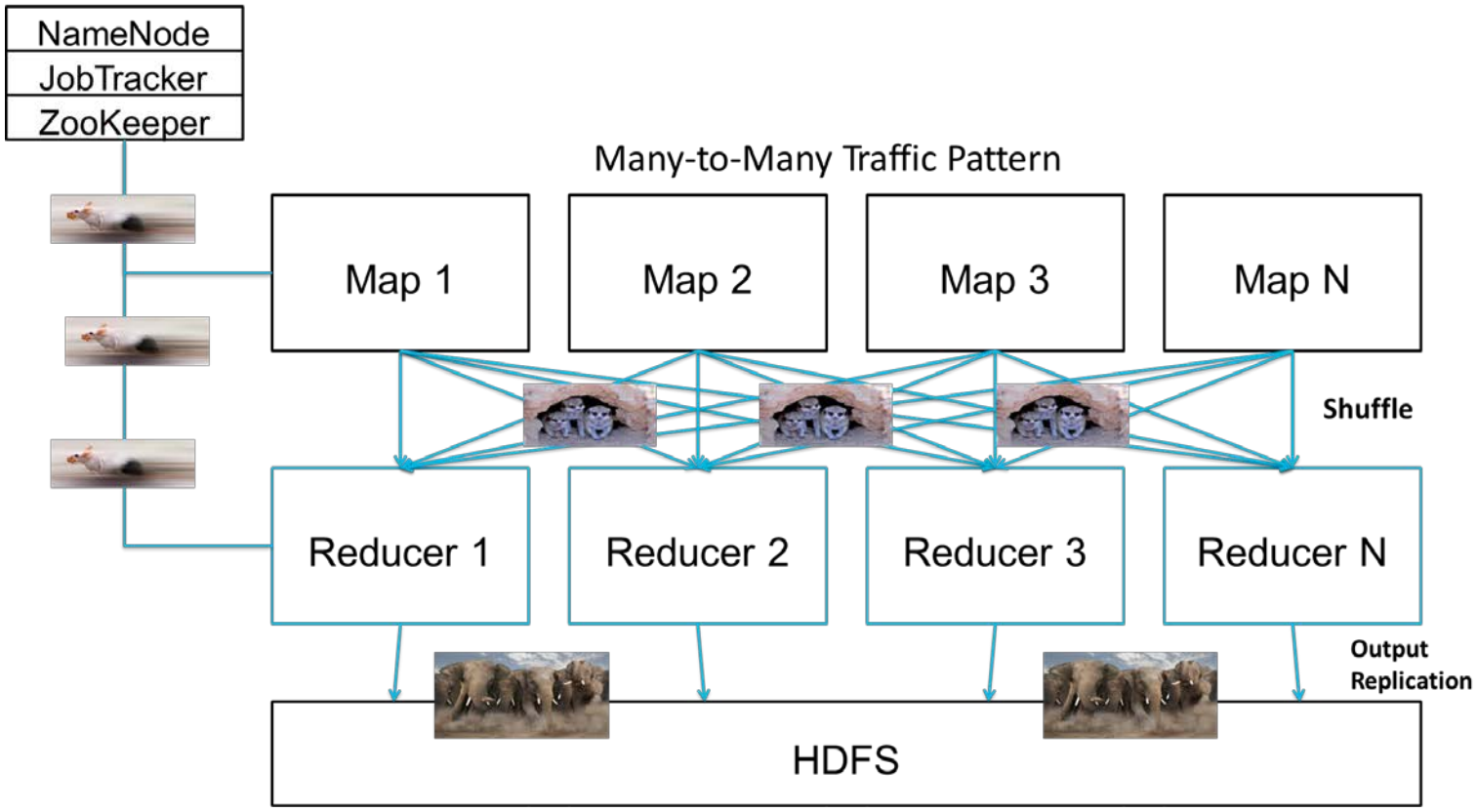
Large Flows

(HDFS gathering)

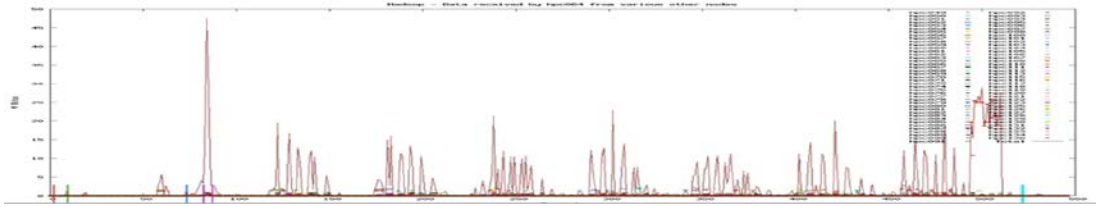
Large Incast

(Hadoop Replication)

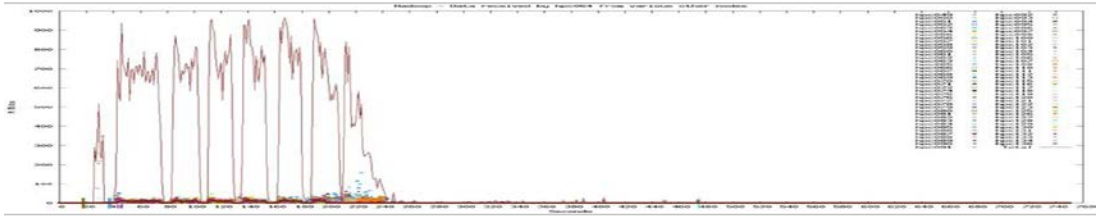
Map and reduce Traffic



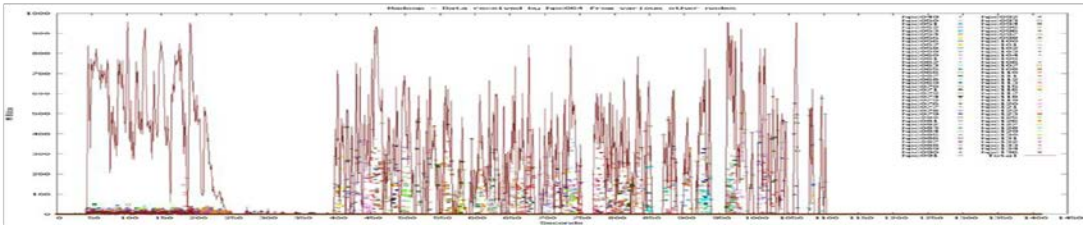
Job Patterns Impact on Network Utilization



Analyze a Word Count BI

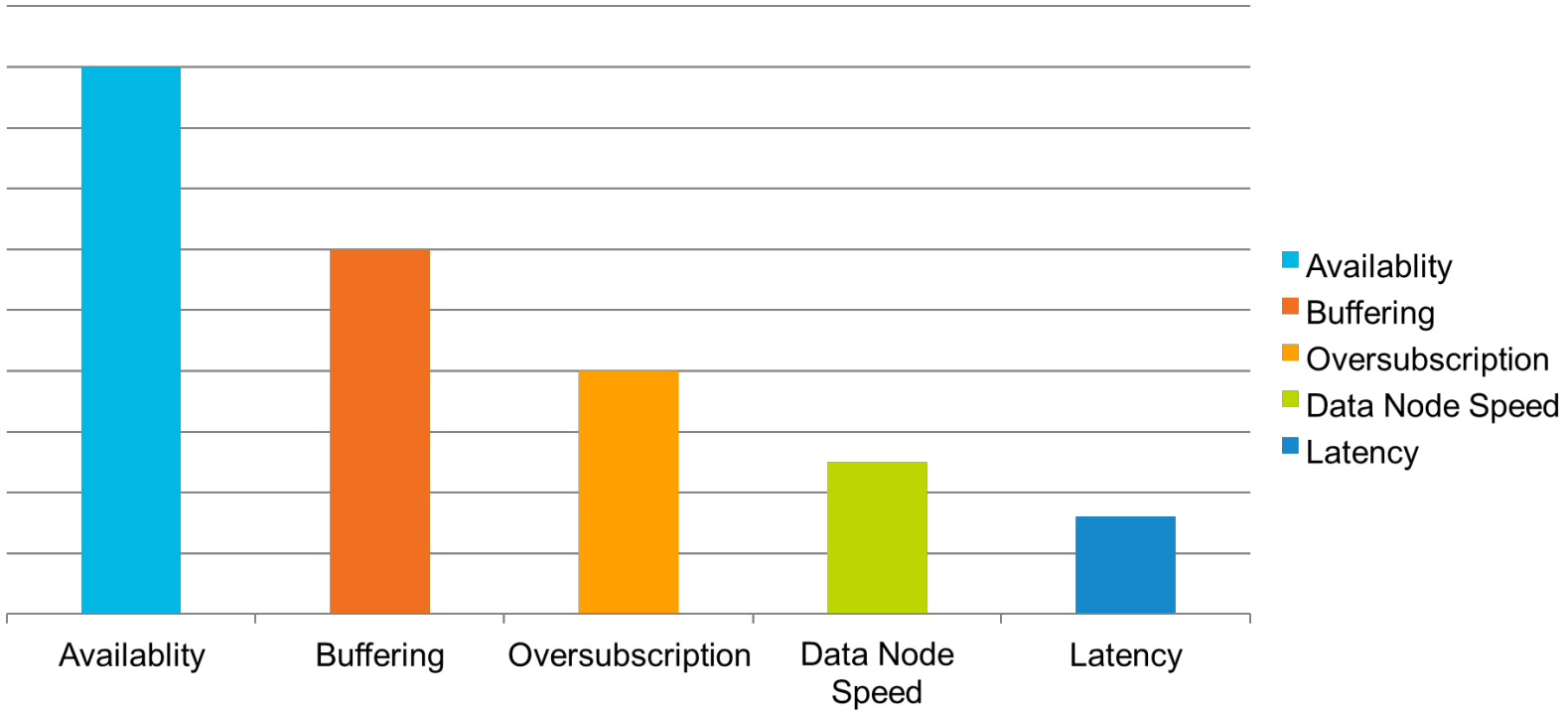


ETL



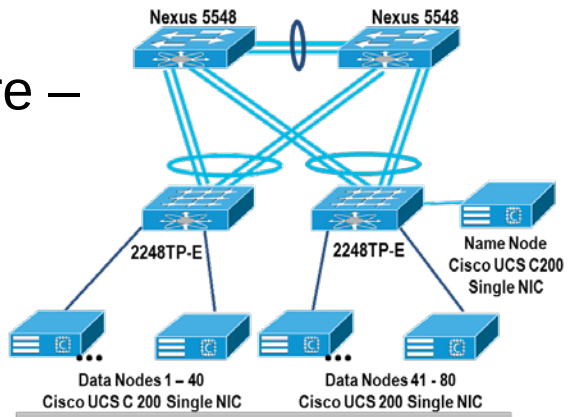
ETL with output replication

Network Characteristics and Hadoop Clusters

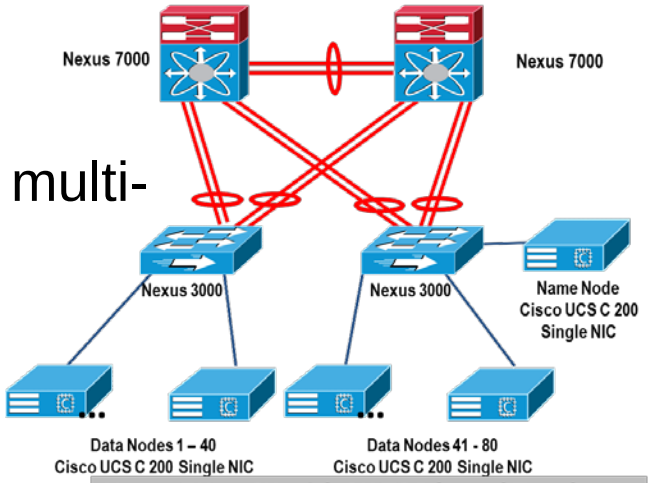


Hadoop Network Topologies

- Integration with Enterprise architecture – essential pathway for data flow
  - Architecture
  - Consistency
  - Management
  - Risk-assurance
  - Enterprise grade features
- Consistent Operational Model
  - NxOS, CLI, Fault Behavior and Management
- Over the time it will have multi-user, multi-workload behavior
  - Need enterprise centric features
  - Security, SLA, QoS etc.
- Big Data is just another app



Traditional DC Design Nexus 55xx/2248



Nexus 7K or 6K with 3K TOR based Topology

Cisco Public

Enterprise Dependencies

Evaluate overall system availability
- Hadoop designed with failure in mind
- Network failures can span many nodes causing rebalancing

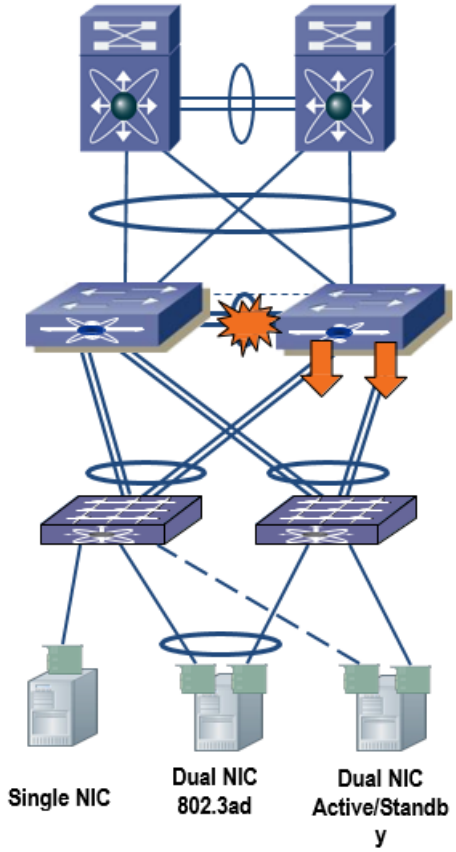
Evaluate redundancy paths and load sharing schemes

Ensure consistent management and operation

Dual homing (active-active) network connections preferred

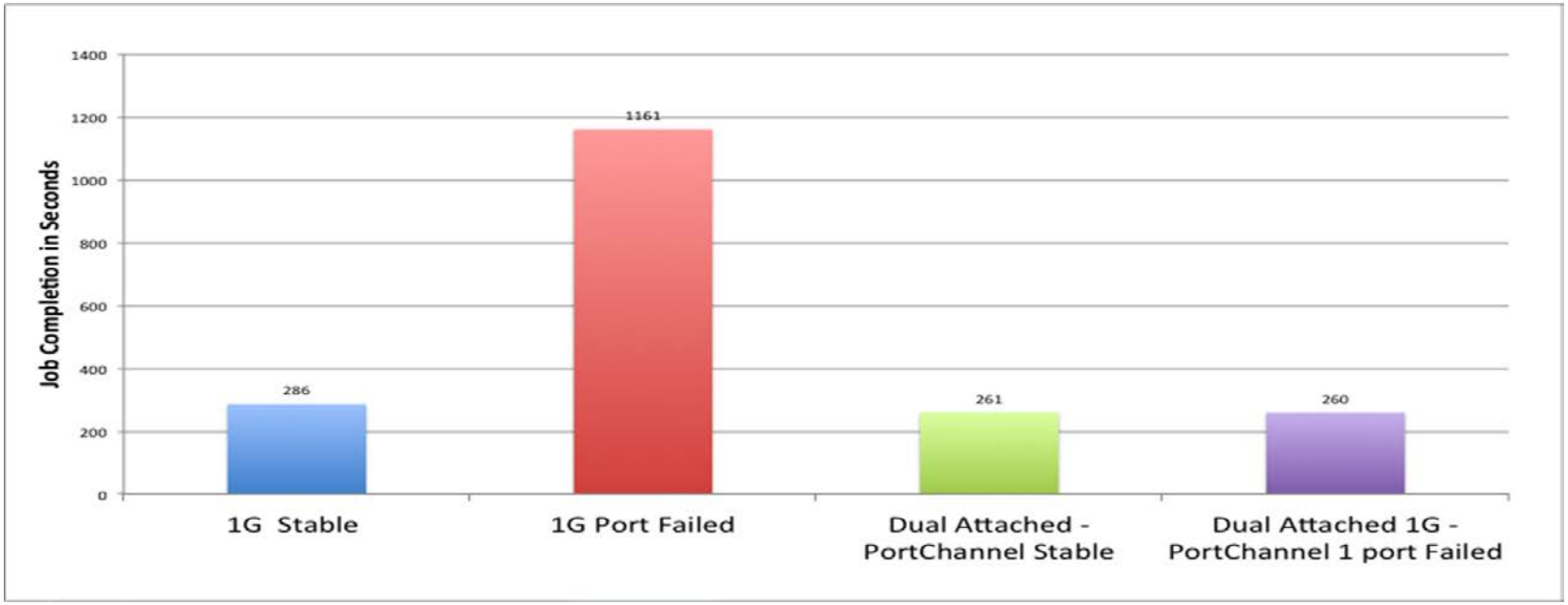
Dual homing FEX avoids single point of failure

Look at vendor specific functions like CISCO's EvPC



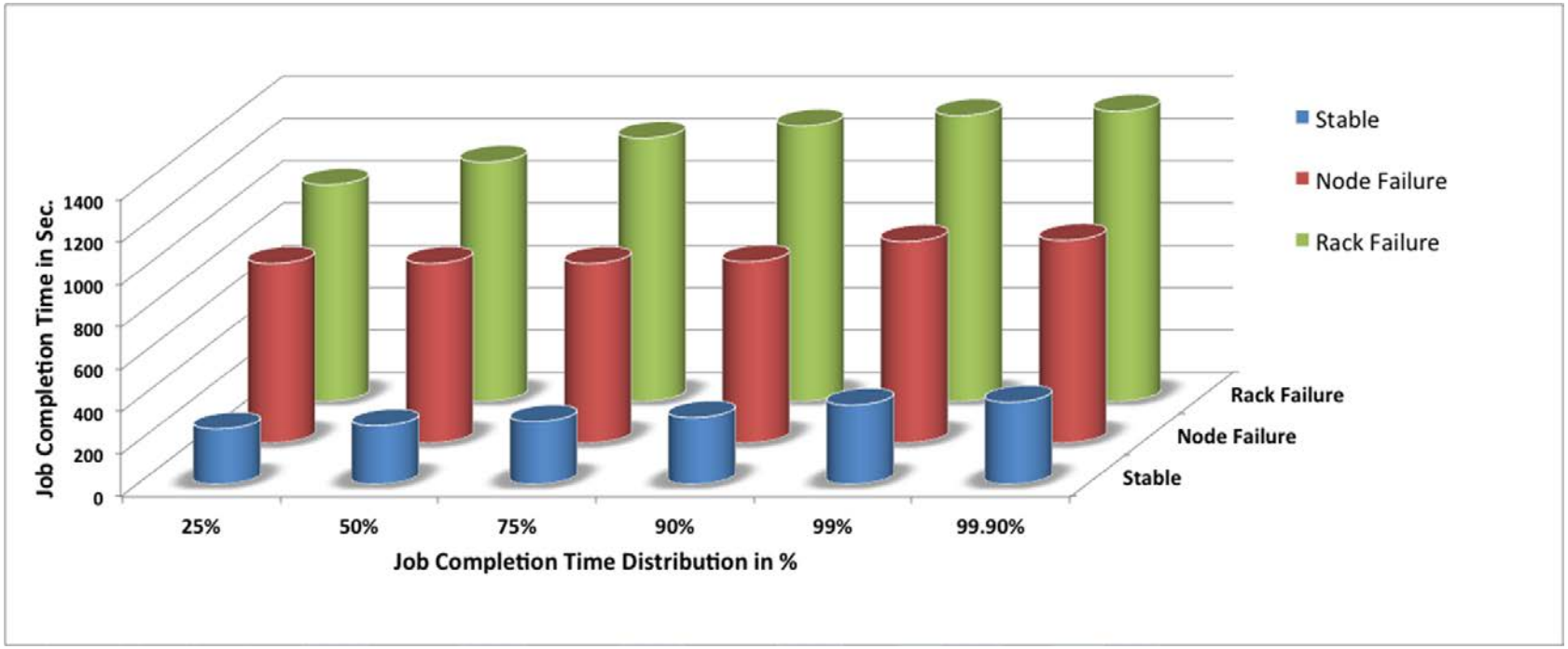
Single Attached vs. Dual Attached Node

- No single point of failure from network view point
- No impact on job completion time
- Effective load-sharing of traffic flow on two NICs



Node and Rack Failure

Almost all jobs impacted with single node failure
With multiple jobs running concurrently node failure as significant as rack failure



Single Node/Port Failure Long Job Times

- The MAP job are executed parallel so unit time for each MAP tasks/node remains same and more less completes the job roughly at the same time.
- However during the failure, set of MAP task remains pending (since other nodes in the cluster are still completing their task) till ALL the node finish the assigned tasks.
- Once all the node finish their MAP task, the left over MAP task being reassigned by name node, the unit time it take to finish those sets of MAP task remain the same(linear) as the time it took to finish the other MAPs – its just happened to be NOT done in parallel thus it could double job completion time. This is the worst case scenario with Terasort, other workload may have variable completion time.
- Type of workload will have affect the impact of single port/node failure

Traffic Bursts Handling

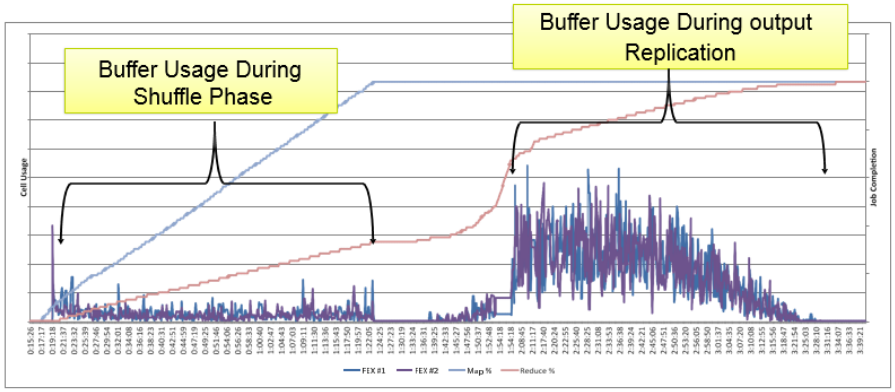
Several HDFS operations and phases of Map Reduction jobs are very bursty in nat

If the network cannot handle traffic bursts then packets will drop

Optimal buffering is required in network devices

TCP as it now exists will collapse at some point with traffic bursts

Proposals under works to change TCP to handle traffic bursts

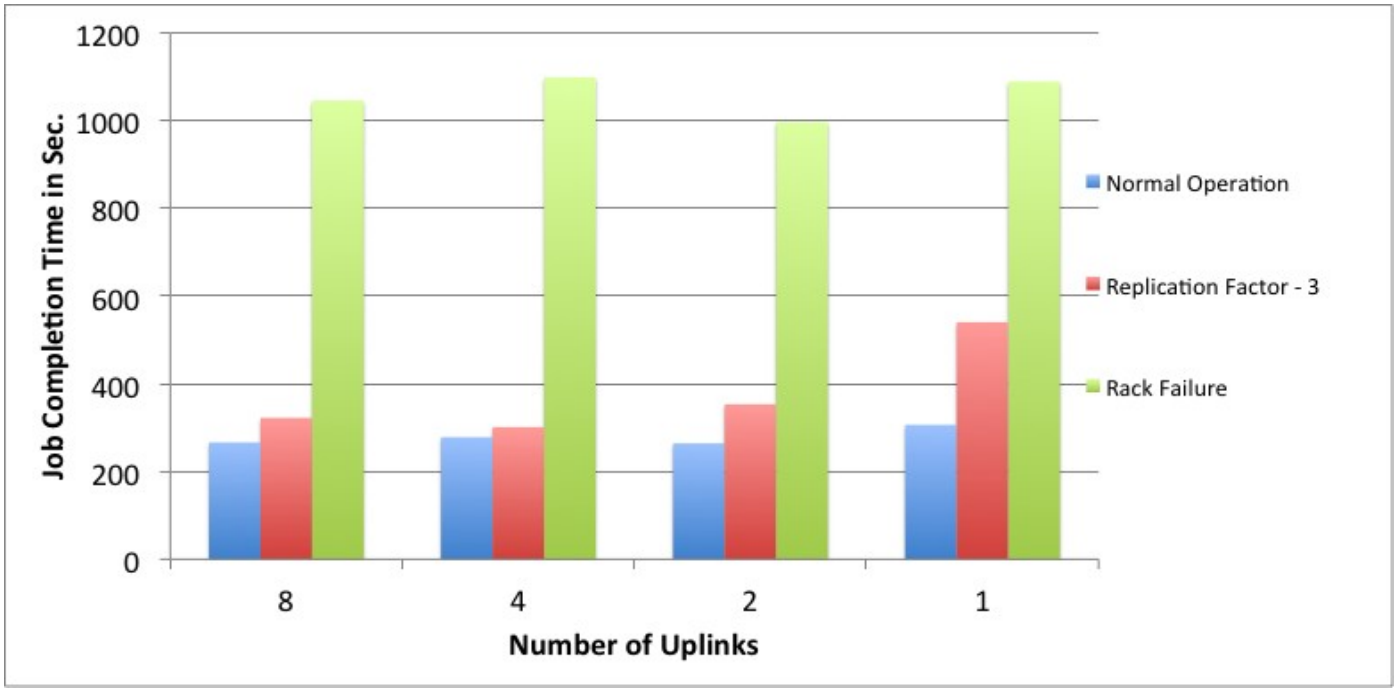


Hadoop Oversubscription

- Hadoop is a parallel batch job oriented framework
- Primary benefits of hadoop is to reduce the time required by workload that otherwise would require long time to meet the SLA e.g. Pricing, log analysis, Join-only job etc.
- Typically oversubscription is high with 10 G server access then at 1Gbps
- Non-blocking network is NOT a requirement, however degree of oversubscription matters for
  - Job Completion Time
  - Replication of Results
  - Oversubscription during rack or FEX failure

Network Oversubscription Impact

- Load to the the network is limited by IO of the compute node
- Oversubscription in the network is a reasonable trade-off
- 1GE and 10GE are normal
- 10Ge reduces the buffer issue



Network Latency

Network latency requires consistency

Not a major factor for Hadoop Clusters

Background

Hadoop Example

Closing Thoughts

Considerations

Big Data technologies – They all respond differently and put different workloads on your network

Bandwidth – Combination of Big Data and traditional traffic can overwhelm an enterprise network

Latency – Big Data by definition is real-real-real time and requires low latency to achieve optimal performance

Capacity – Plan for massive amounts of storage supporting a wide variety of information types

Processing – Big Data can strain computational, memory, and storage systems if proper planning is not done in these areas

Secure data access – Sensitive information from various sources expands the access protection required

Key Steps to Prepare for Big Data

Determine how Big Data will change your business

Gain awareness of the variety of vendor offerings

Define requirements for YOUR Big Data implementation to meet business objectives

Evaluate existing service management strategy for Big Data support

Establish training objectives across the business

Execute a pre-deployment assessment

Implement a test environment at the system level

Ensure an organizational structure to handle Big Data once implemented

ありがとうございました

Grazie

Danke

Merci

ขอบคุณ

धन्यवाद
Hindi
Gracias

laurak@aesclever.com
www.aesclever.com
650-617-2400

감사합니다

நன்றி
Tamil

Obrigado

: