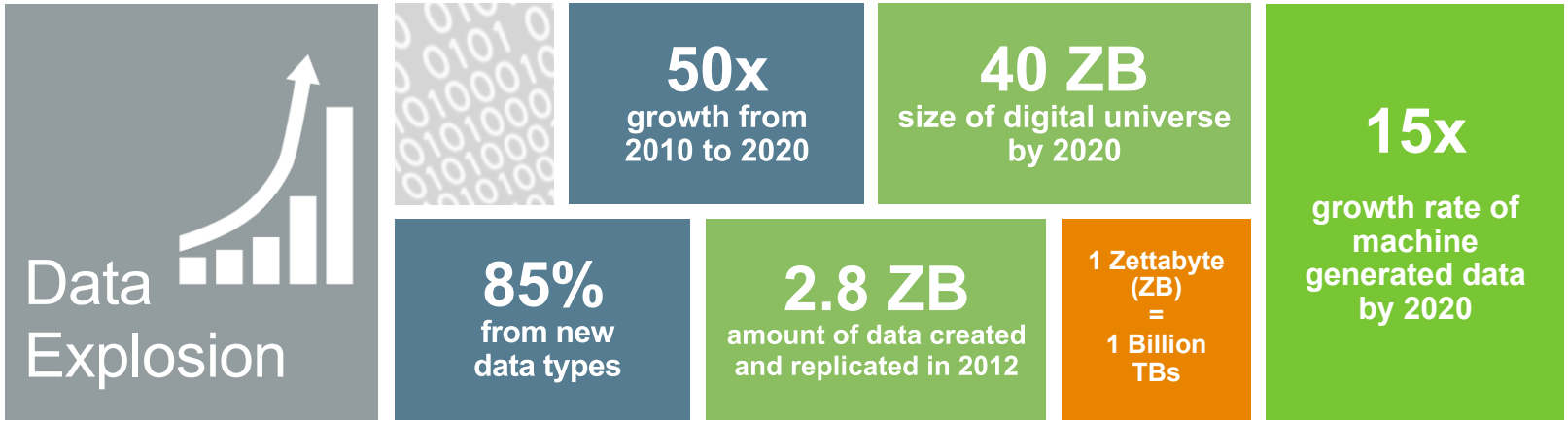
Hortonworks

Stinger Initiative: Introduction

Interactive Query on Hadoop

Chris Harris
E-Mail : charris@hortonworks.com
Twitter : cj_harris5

The World of Data is Changing



Source: IDC

By 2015, organizations that build a modern information management system will outperform their peers financially by 20 percent.

– Gartner, Mark Beyer, “Information Management in the 21st Century”

What is Hadoop?

Hadoop is a new data platform,

that can store more data, more kinds of data, and perform more flexible analyses

Hadoop is open source

and runs on industry standard hardware,

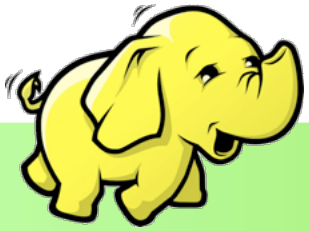
so it's 1-2 orders of magnitude more economical than conventional data solutions

Hadoop provides more cost effective storage, processing, and analysis. Some existing workloads run faster, cheaper, better

Hadoop Enterprise Use Cases

Vertical	Use Case	Data Type
Financial Services	**New Account Risk Screens**	**Text, Server Logs**
	Fraud Prevention	**Server Logs**
	Trading Risk	**Server Logs**
	Maximize Deposit Spread	**Text, Server Logs**
	Insurance Underwriting	**Geographic, Sensor, Text**
	Accelerate Loan Processing	**Text**
Telecom	**Call Detail Records (CDRs)**	**Machine, Geographic**
	Infrastructure Investment	**Machine, Server Logs**
	Next Product to Buy (NPTB)	**Clickstream**
	Real-time Bandwidth Allocation	**Server Logs, Text, Sentiment**
	New Product Development	**Machine, Geographic**
Retail	**360° View of the Customer**	**Clickstream, Text**
	Analyze Brand Sentiment	**Sentiment**
	Localized, Personalized Promotions	**Geographic**
	Website Optimization	**Clickstream**
	Optimal Store Layout	**Sensor**
Manufacturing	**Supply Chain and Logistics**	**Sensor**
	Assembly Line Quality Assurance	**Sensor**
	Proactive Maintenance	**Machine**
	Crowdsourced Quality Assurance	**Sentiment**

A Brief History of Apache Hadoop



Apache Project Established — 2004

Yahoo! begins to Operate at scale — 2006, 2008

Hortonworks Data Platform — 2010, 2012

2013 Enterprise Hadoop

2005: Yahoo! creates team under E14 to work on Hadoop

Focus on INNOVATION

2008: Yahoo team extends focus to operations to support multiple projects & growing clusters

Focus on OPERATIONS

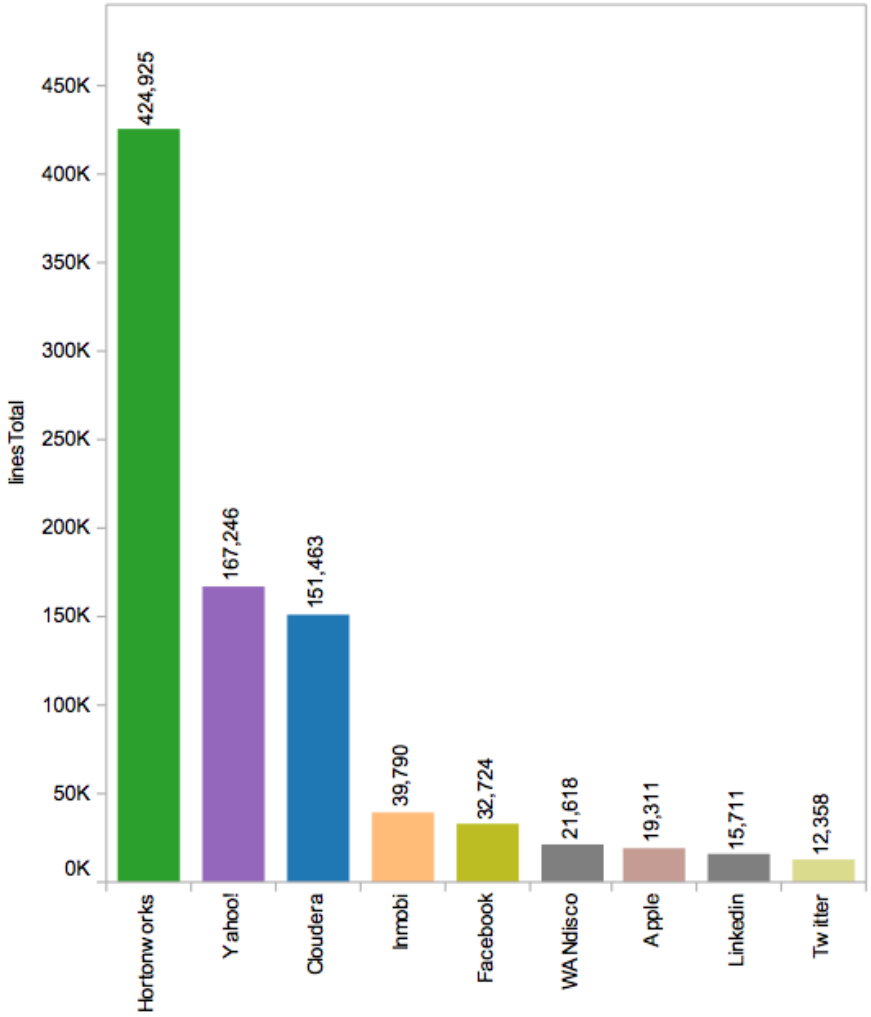
Hortonworks

2011: Hortonworks created to focus on "*Enterprise Hadoop*". Starts with 24 key Hadoop engineers from Yahoo

STABILITY

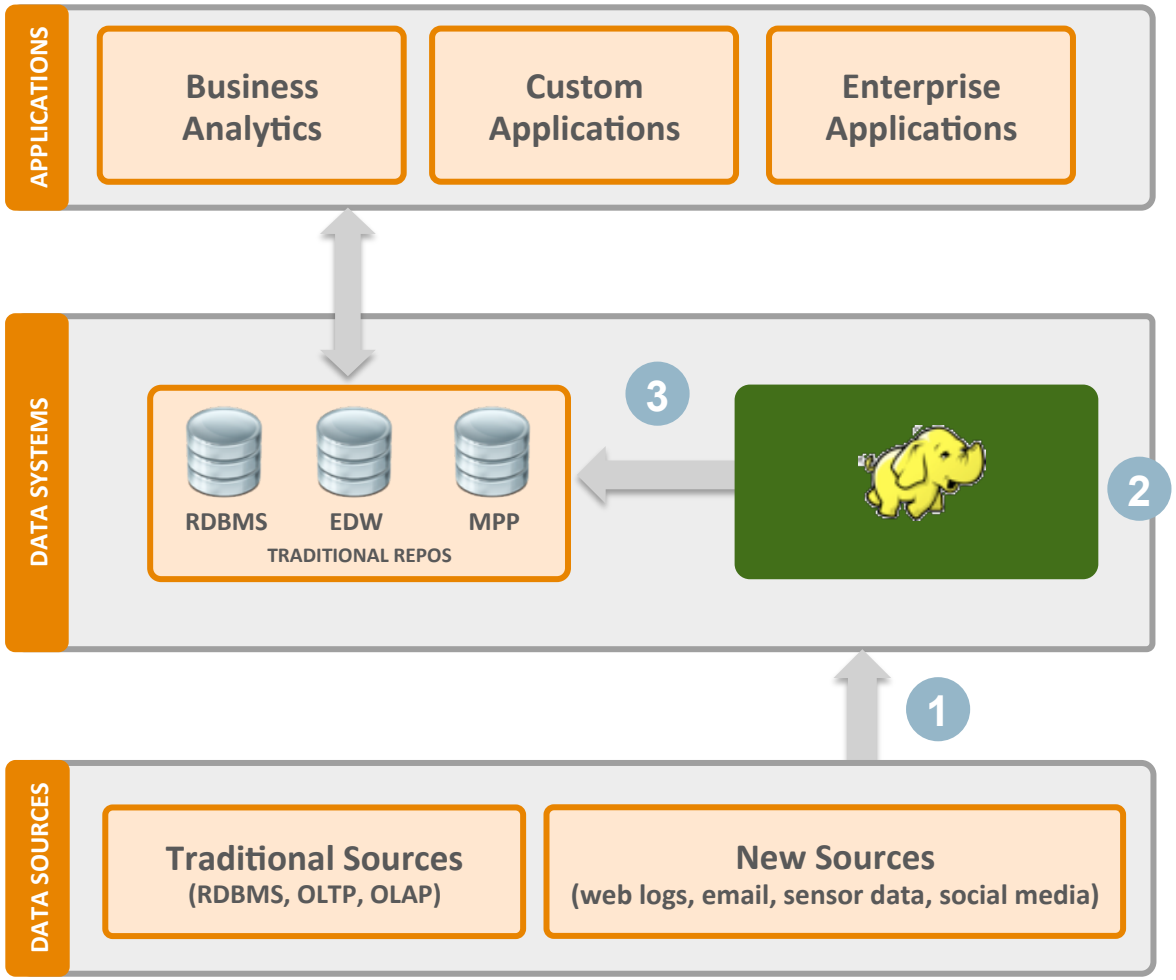
Leadership that Starts at the Core

- **Driving next generation Hadoop**
  - YARN, MapReduce2, HDFS2, High Availability, Disaster Recovery
- **420k+ lines authored since 2006**
  - More than twice nearest contributor
- **Deeply integrating w/ecosystem**
  - Enabling new deployment platforms
    - (ex. Windows & Azure, Linux & VMware HA)
  - Creating deeply engineered solutions
    - (ex. Teradata big data appliance)
- **All Apache, NO holdbacks**
  - 100% of code contributed to Apache



Operational Data Refinery

Refine | Explore | Enrich

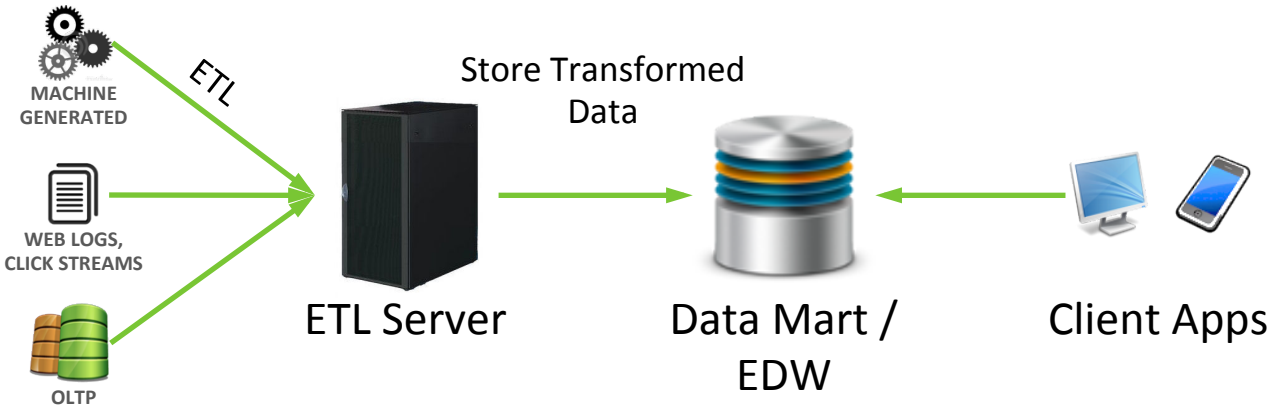


Collect data and apply a known algorithm to it in trusted operational process

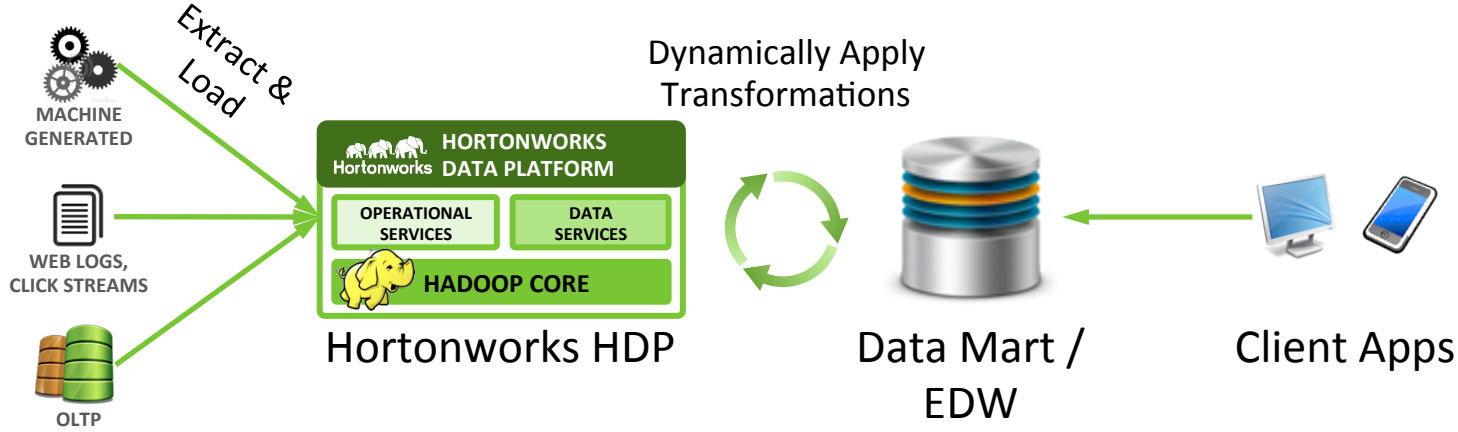
1. **Capture**
   Capture all data
2. **Process**
   Parse, cleanse, apply structure & transform
3. **Exchange**
   Push to existing data warehouse for use with existing analytic tools

Key Capability in Hadoop: Late binding

With traditional ETL, structure must be agreed upon far in advance and is difficult to change.

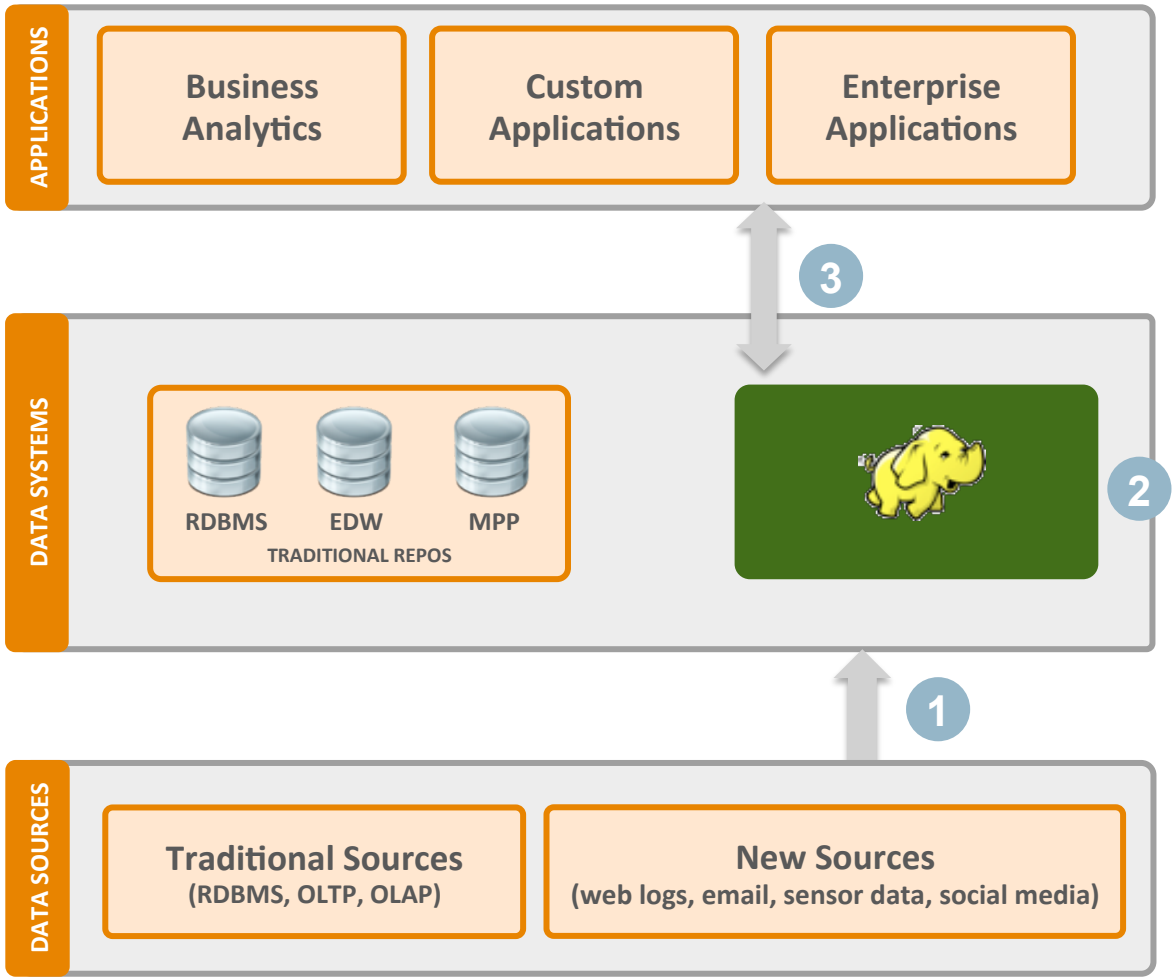


With Hadoop, capture all data, structure data as business need evolve.



Big Data Exploration & Visualization

Refine | Explore | Enrich



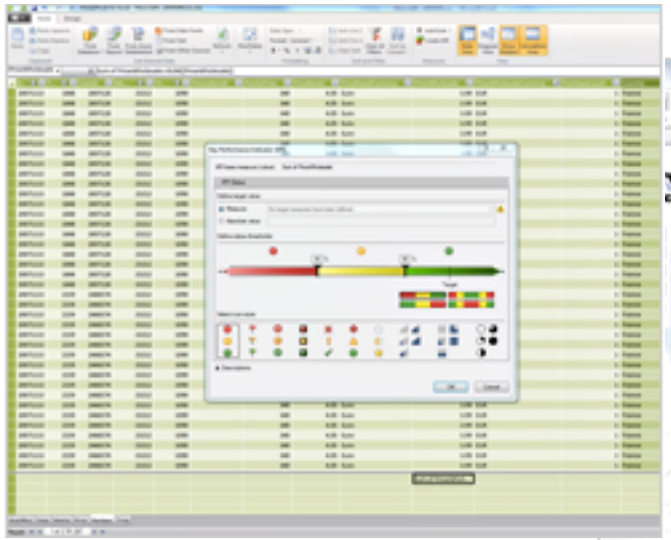
Collect data and perform iterative investigation for value

1. **Capture**
   Capture all data
2. **Process**
   Parse, cleanse, apply structure & transform
3. **Exchange**
   Explore and visualize with analytics tools supporting Hadoop

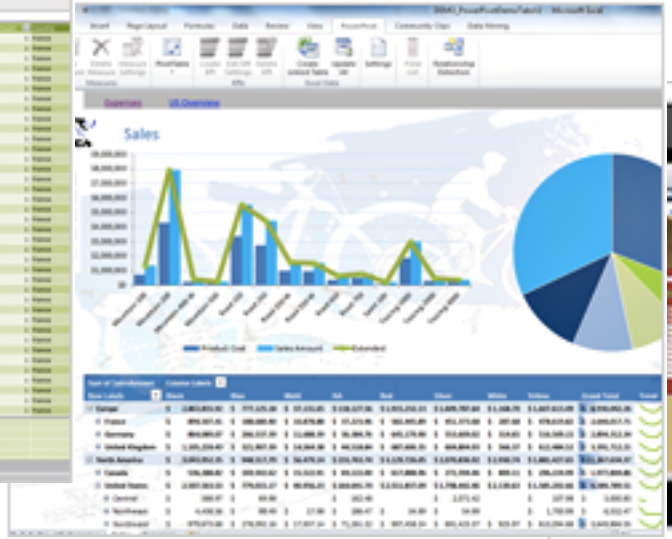
Visualization Tooling

- **Robust visualization and business tooling**
- **Ensures scalability when working with large datasets**

Native Excel support

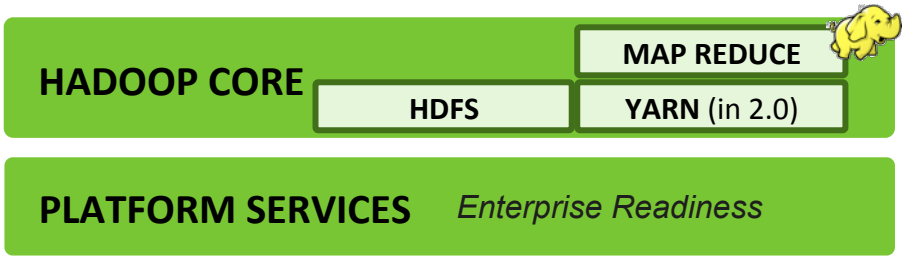


Web browser support



Mobile support

Enhancing the Core of Apache Hadoop

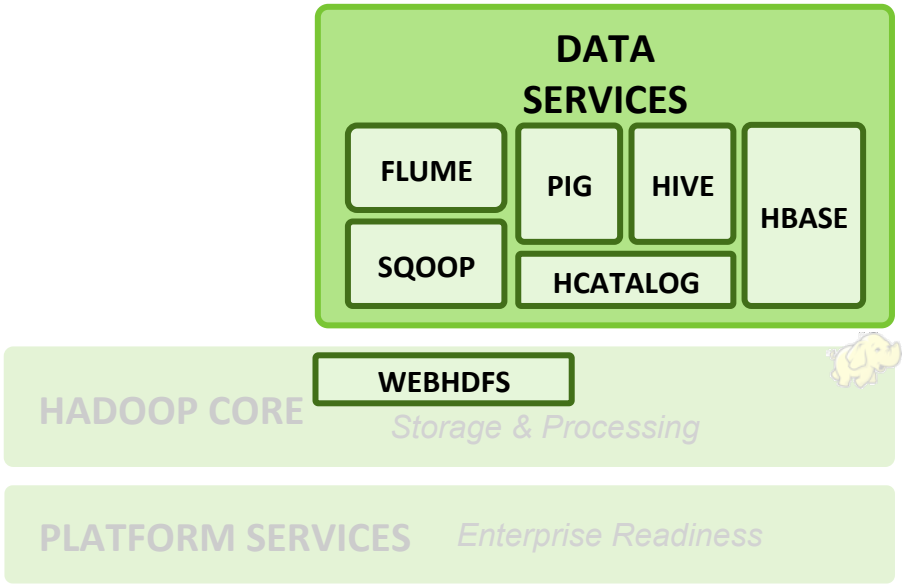


Deliver high-scale *storage & processing* with enterprise-ready *platform services*

Unique Focus Areas:

- **Bigger, faster, more flexible**
  Continued focus on speed & scale and enabling near-real-time apps
- **Tested & certified at scale**
  Run ~1300 system tests on large Yahoo clusters for every release
- **Enterprise-ready services**
  High availability, disaster recovery, snapshots, security, …

Data Services for Full Data Lifecycle



Provide *data services* to store, process & access data in many ways

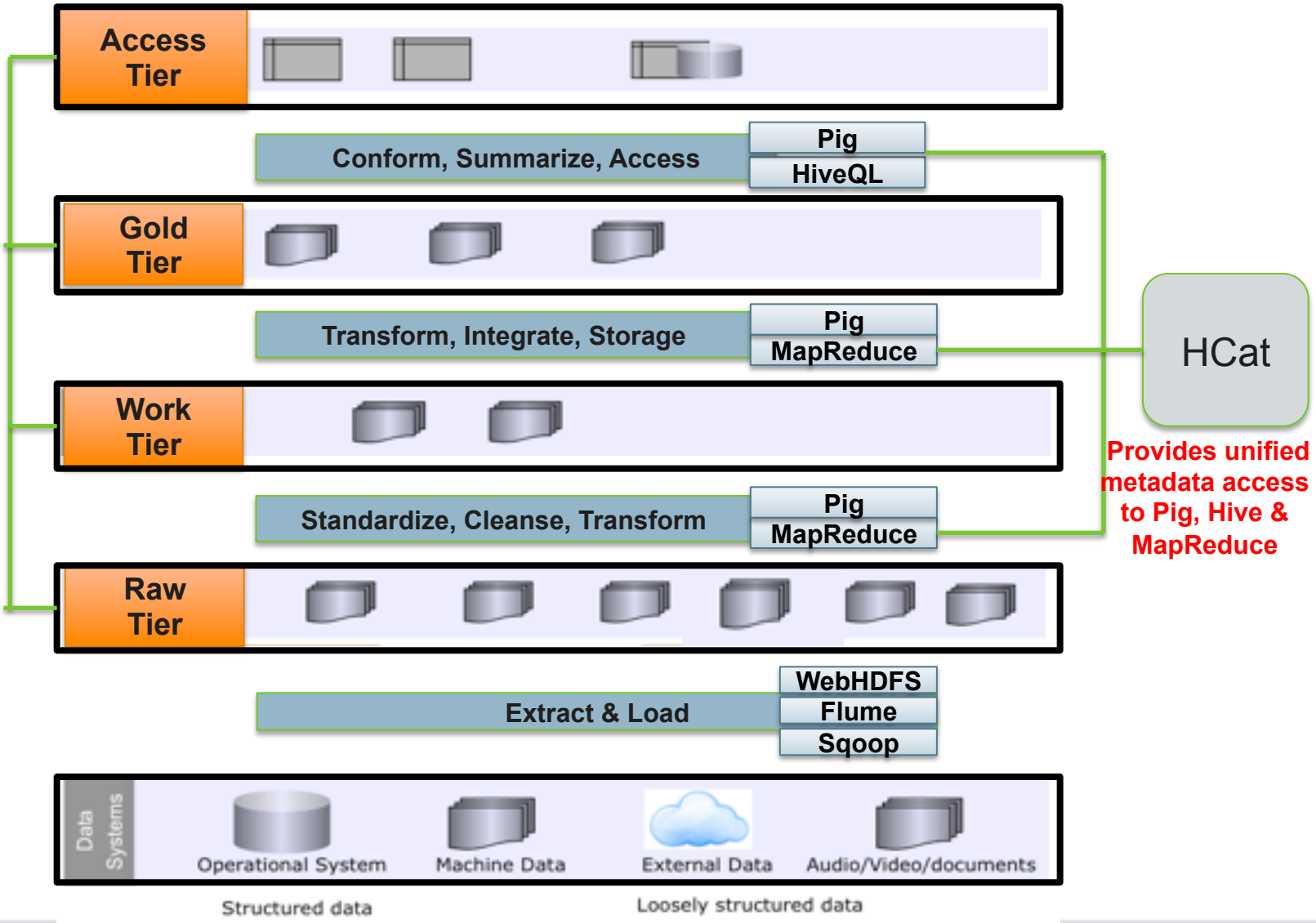
Unique Focus Areas:

- **Apache HCatalog**
  Metadata services for consistent table access to Hadoop data
- **Apache Hive**
  Explore & process Hadoop data via SQL & ODBC-compliant BI tools
- **Apache HBase**
  NoSQL database for Hadoop
- **WebHDFS**
  Access Hadoop files via scalable REST API

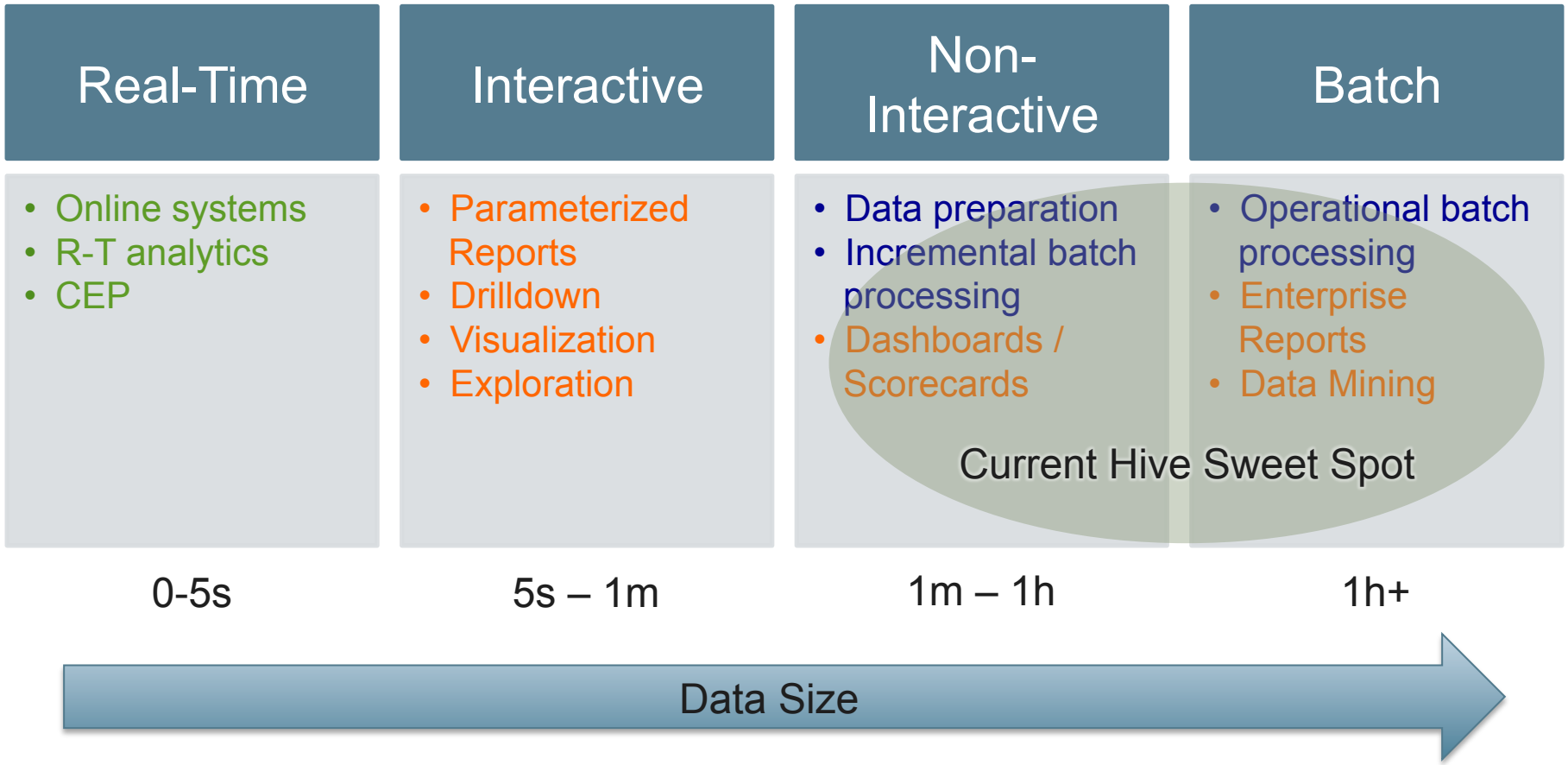
- **Talend Open Studio for Big Data**
  Graphical data integration tools

Organize Tiers and Process with Metadata

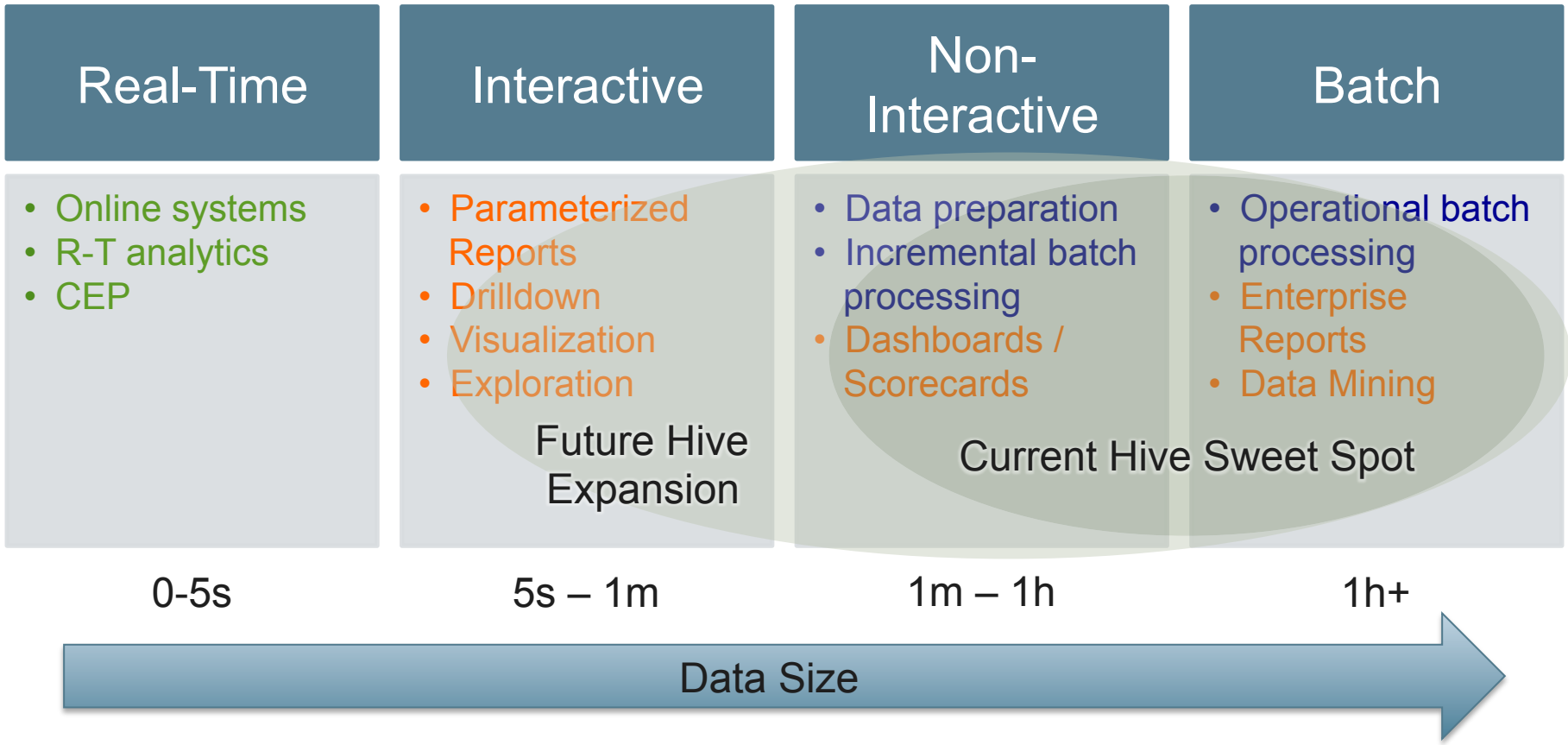
- Organize data based on source/derived relationships
- Allows for fault and rebuild process



Hive Current Focus Area



Stinger: Extending Hive's Sweetspot



Improve Latency & Throughput

- Query engine improvements
- New "Optimized RCFile" column store
- Next-gen runtime (elim's M/R latency)

Extend Deep Analytical Ability

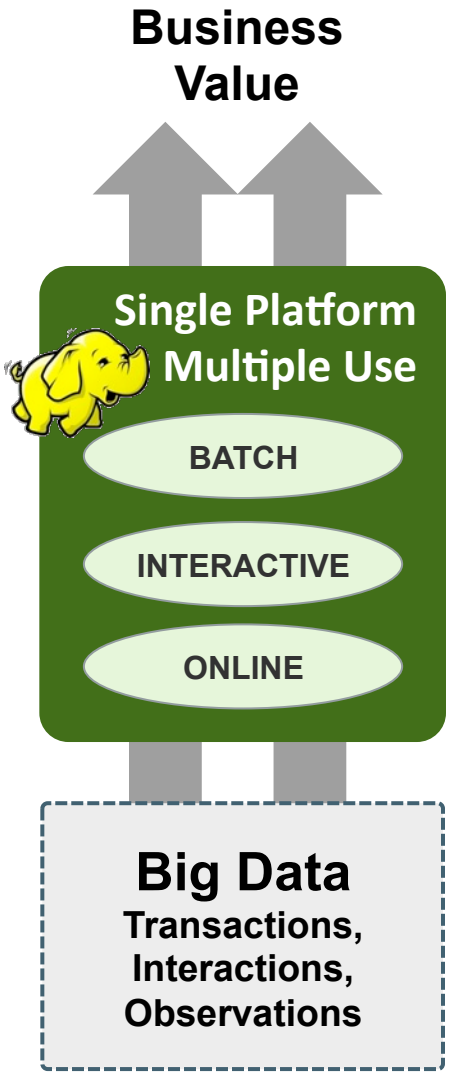
- Analytics functions
- Improved SQL coverage
- Continued focus on core Hive use cases

Hadoop 2.0… The Enterprise Generation

1.0 → Architected for the **Large Web Properties**

2.0 → Architected for the **Broad Enterprise**

Enterprise Requirements	Hadoop 2.0 Features
Mixed workloads	YARN
Interactive Query	Hive on Tez
Reliability	Full Stack HA
Point in time Recovery	Snapshots
Multi Data Center	Disaster Recovery
ZERO downtime	Rolling Upgrades
Security	Knox Gateway

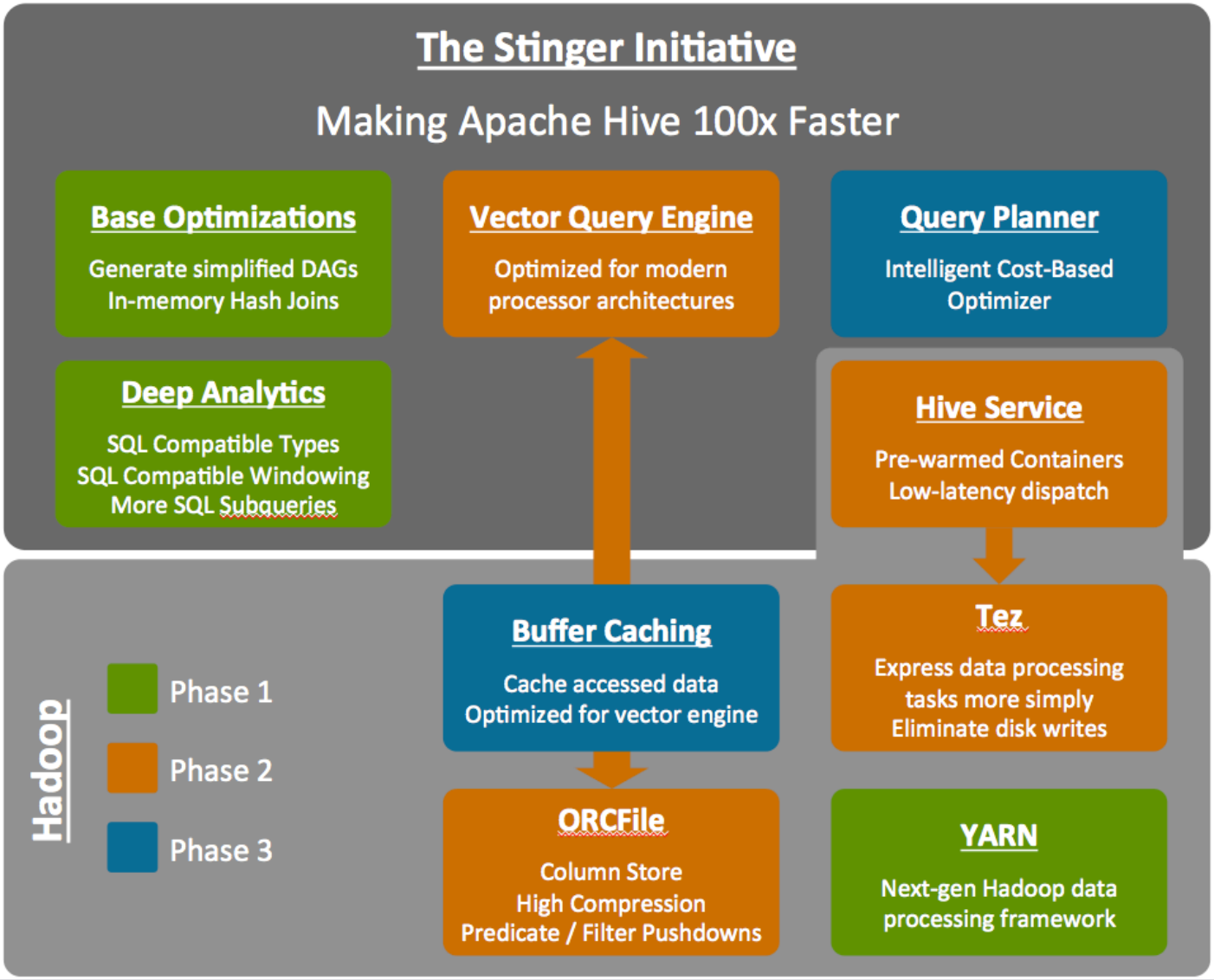


Hortonworks

The Stinger Initiative

Interactive Query on Hadoop

Stinger Initiative At A Glance

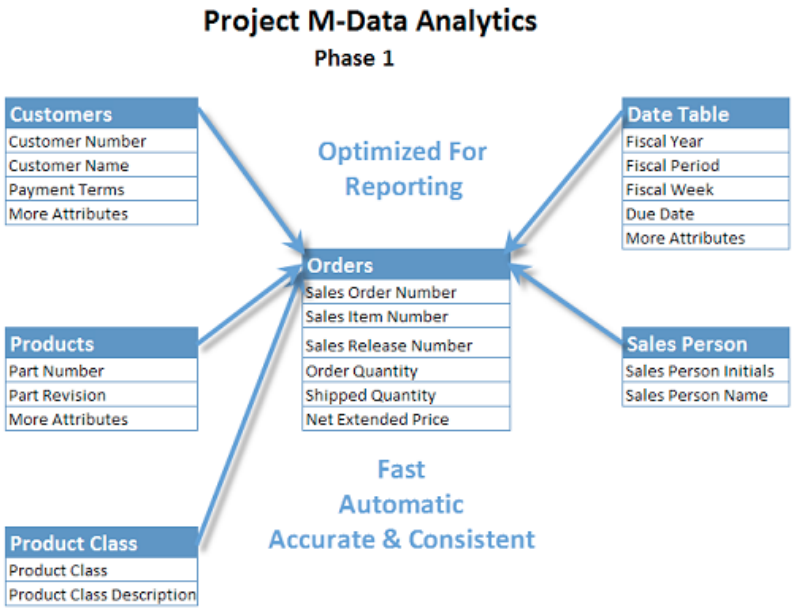


Base Optimizations: Intelligent Optimizer

- **Introduction of In-Memory Hash Join:**
  - For joins where one side fits in memory:
  - New in-memory-hash-join algorithm.
  - Hive reads the small table into a hash table.
  - Scans through the big file to produce the output.
- **Introduction of Sort-Merge-Bucket Join:**
  - Applies when tables are bucketed on the same key.
  - Dramatic speed improvements seen in benchmarks.
- **Other Improvements:**
  - Lower the footprint of the fact tables in memory.
  - Enable the optimizer to automatically pick map joins.

Dimensionally Structured Data

- **Extremely common pattern in EDW.**
- **Results in large "fact tables" and small "dimension tables".**
- **Dimension tables often small enough to fit in RAM.**
- **Sometimes called Star Schema.**



A Query on Dimensional Data

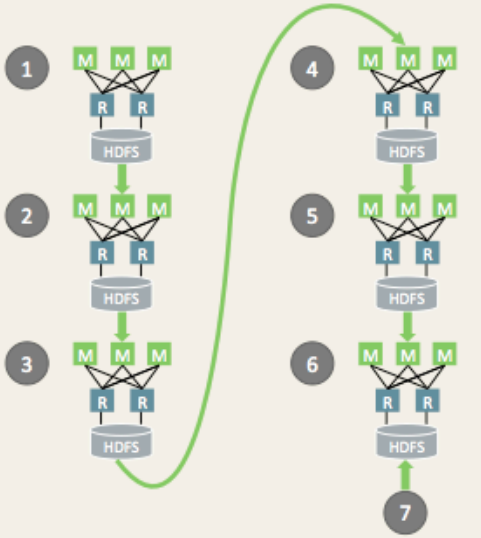
- **Derived from TPC-DS Query 27**

```
SELECT col5, avg(col6)
FROM fact_table
    join dim1 on (fact_table.col1 = dim1.col1)
    join dim2 on (fact_table.col2 = dim2.col1)
    join dim3 on (fact_table.col3 = dim3.col1)
    join dim4 on (fact_table.col4 = dim4.col1)
GROUP BY col5
ORDER BY col5
LIMIT 100;
```

- **Dramatic speedup on Hive 0.11**

Star Schema Join Improvements in 0.11

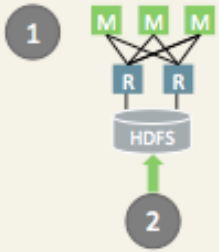
Star Schema Join: Hive 0.10 without hints.



1. Join `fact_table` to `dim1`. Persist results.
2. Join results to `dim2`. Persist results.
3. Join results to `dim3`. Persist results.
4. Join results to `dim4`. Persist results.
5. Group results by `col5`. Persist results.
6. Order results by `col5`. Persist results.
7. Client reads results.

BEFORE

Star Schema Join: Hive 0.11 without hints.



1. Load `dim1` – `dim4` into memory on all nodes. (HIVE-3784)
   Perform map-side joins. (HIVE-3952)
   Collapse ORDER BY and GROUP BY into single reducer. (HIVE-2340)
   Persist results.
2. Client reads results.

AFTER

ORCFile - Optimized Column Storage

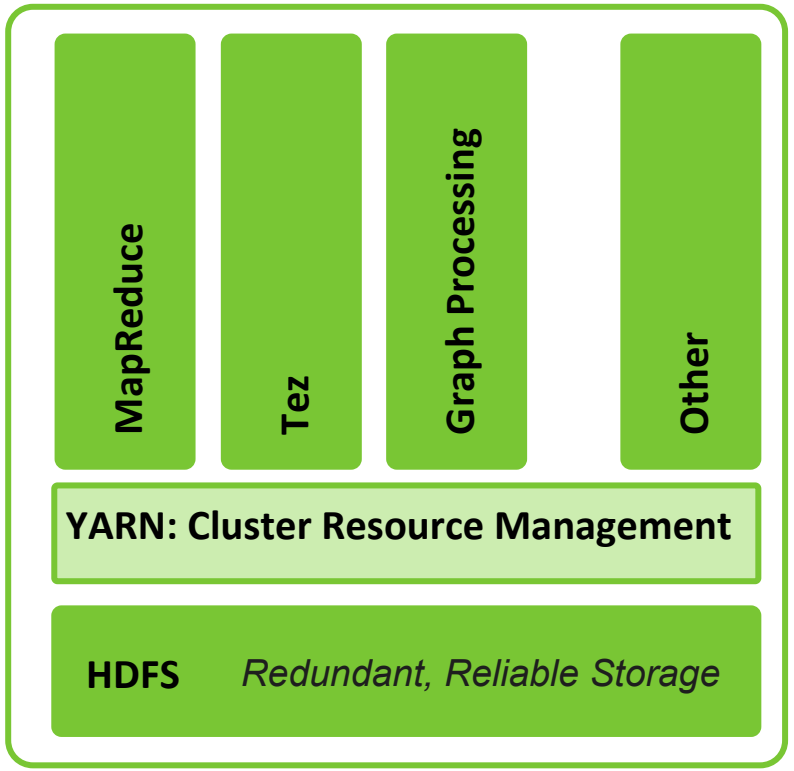
- **Make a better columnar storage file**
  - Tightly aligned to Hive data model
- **Decompose complex row types into primitive fields**
  - Better compression and projection
- **Only read bytes from HDFS for the required columns.**
- **Store column level aggregates in the files**
  - Only need to read the file meta information for common queries
  - Stored both for file and each section of a file
  - Aggregates: min, max, sum, average, count
  - Allows fast access by sorted columns
- **Ability to add bloom filters for columns**
  - Enables quick checks for whether a value is present

Yarn

Moving Hive and Hadoop beyond MapReduce

Hadoop 2.0 Innovations - YARN

- **Focus on scale and innovation**
  - Support 10,000+ computer clusters
  - Extensible to encourage innovation
- **Next generation execution**
  - Improves MapReduce performance
- **Supports new frameworks beyond MapReduce**
  - Low latency, Streaming, Services
  - Do more with a single Hadoop cluster



Tez

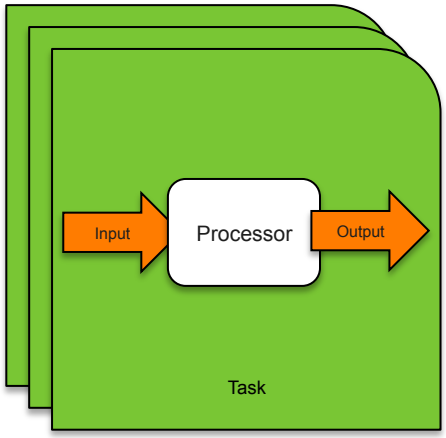
Moving Hive and Hadoop beyond MapReduce

Tez

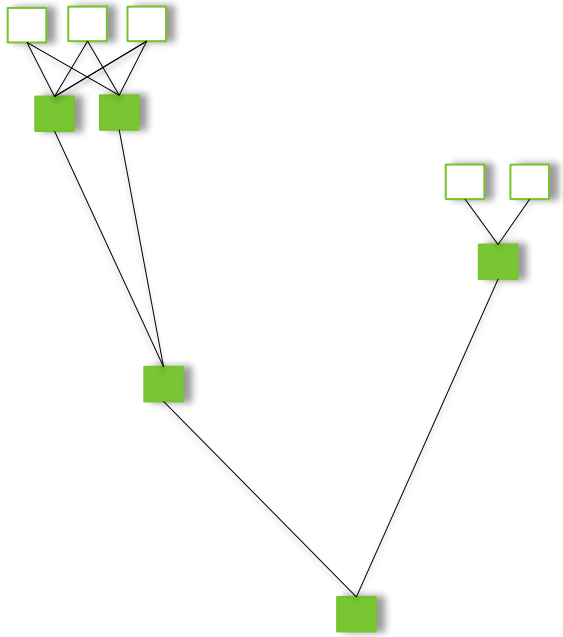
- **Low level data-processing execution engine**
- **Use it for the base of MapReduce, Hive, Pig, Cascading etc.**
- **Enables pipelining of jobs**
- **Removes task and job launch times**
- **Hive and Pig jobs no longer need to move to the end of the queue between steps in the pipeline**
- **Does not write intermediate output to HDFS**
  - Much lighter disk and network usage
- **Built on YARN**

Tez - Core Idea

Task with pluggable *Input*, *Processor* & *Output*



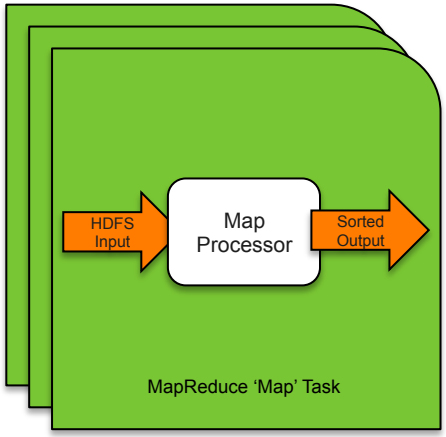
Tez Task - <*Input, Processor, Output*>



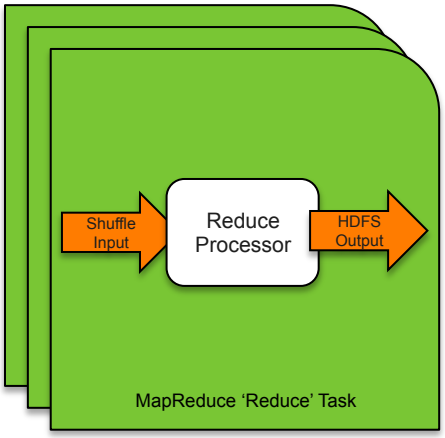
YARN ApplicationMaster to run DAG of Tez Tasks

Tez – Blocks for building tasks

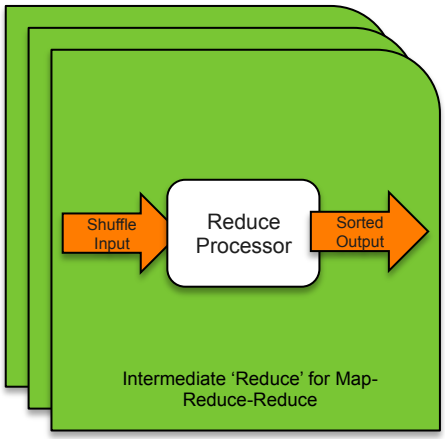
MapReduce 'Map'



MapReduce 'Reduce'

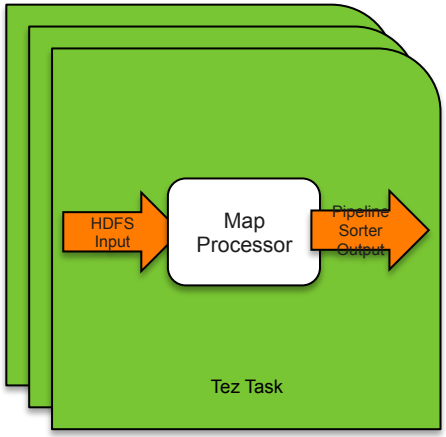


Intermediate 'Reduce' for Map-Reduce-Reduce

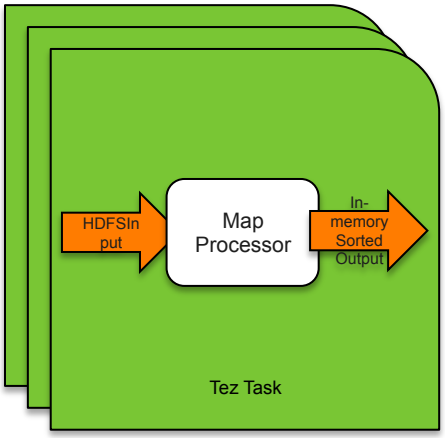


Tez – More tasks

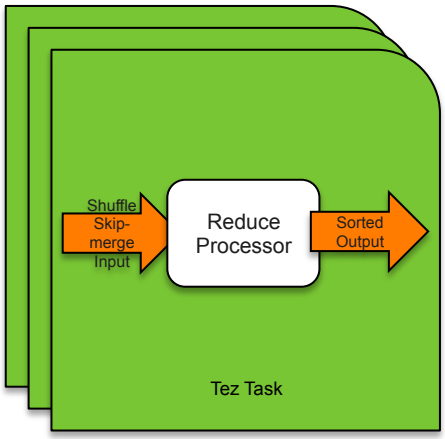
Special Pig/Hive 'Map'



In-memory Map

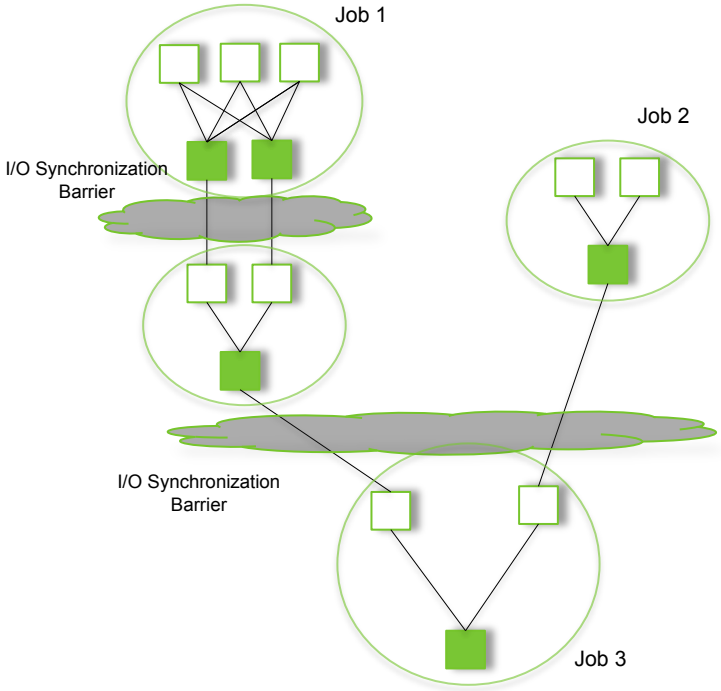


Special Pig/Hive 'Reduce'

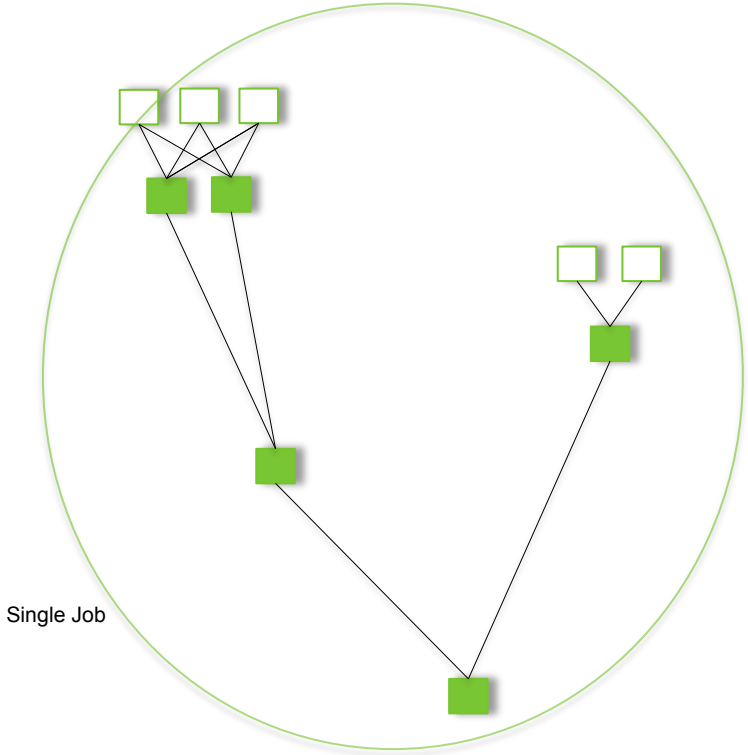


Pig/Hive-MR versus Pig/Hive-Tez

```
SELECT a.state, COUNT(*), AVERAGE(c.price)
FROM a
JOIN b ON (a.id = b.id)
JOIN c ON (a.itemId = c.itemId)
GROUP BY a.state
```

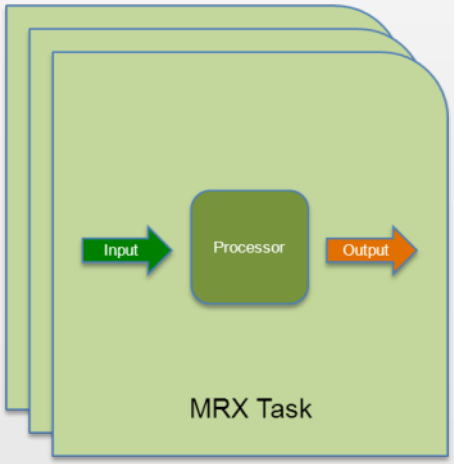


Pig/Hive - MR



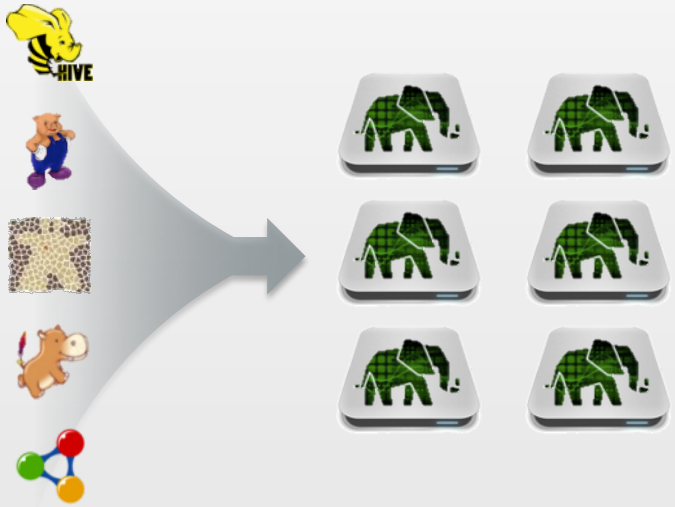
Pig/Hive - Tez

FastQuery: Beyond Batch with YARN



Tez Generalizes Map-Reduce

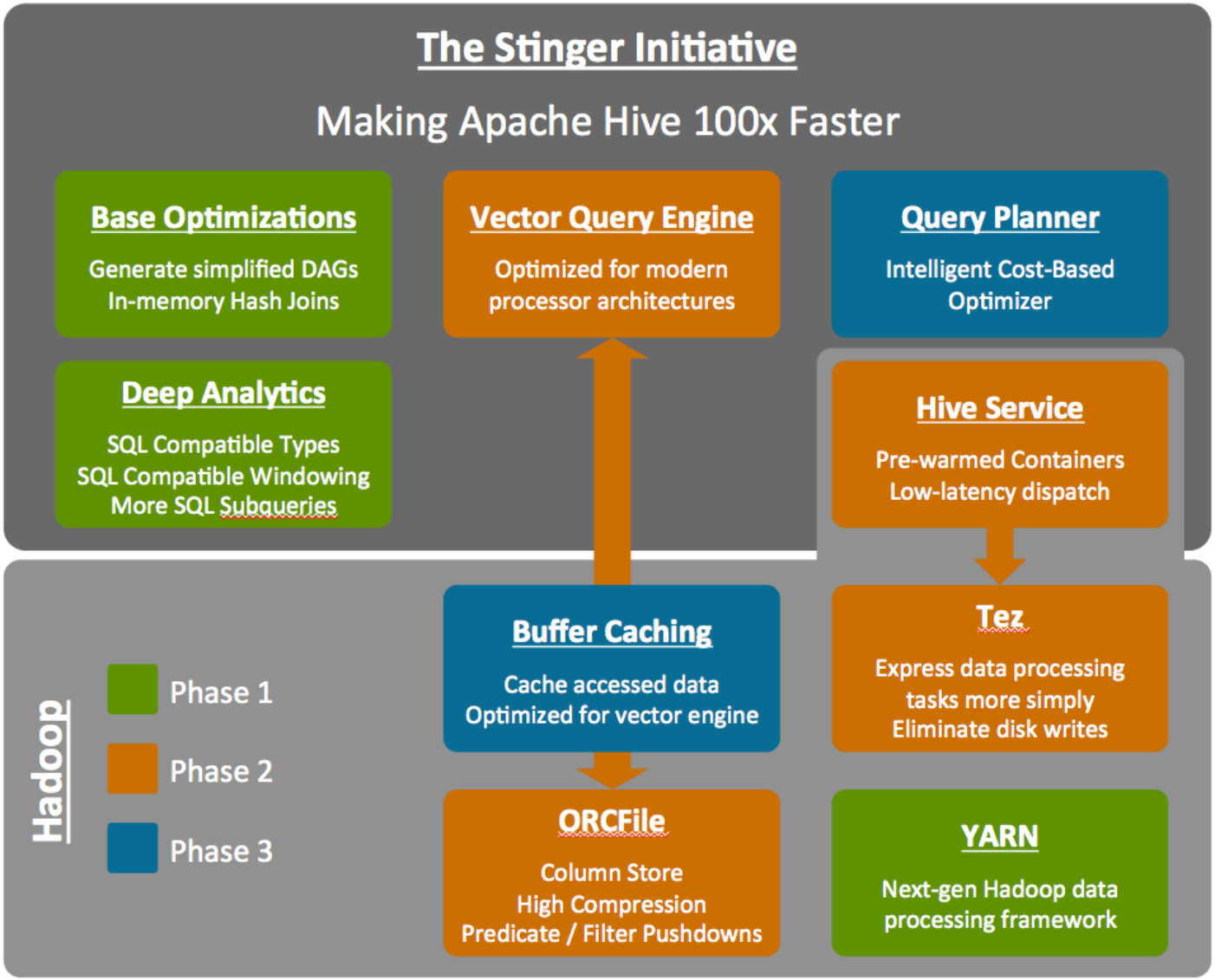
Simplified execution plans process data more efficiently



Always-On Tez Service

Low latency processing for all Hadoop data processing

Recap and Questions: Hive Performance



Hortonworks

Improving Hive's SQL Support

Stinger: Deep Analytical Capabilities

- **SQL:2003 Window Functions**
  - OVER clauses
    - Multiple PARTITION BY and ORDER BY supported
    - Windowing supported (ROWS PRECEDING/FOLLOWING)
    - Large variety of aggregates
      - RANK
      - FIRST_VALUE
      - LAST_VALUE
      - LEAD / LAG
      - Distrubutions

Hive Data Type Conformance

- Data Types:
  - Add fixed point NUMERIC and DECIMAL type (in progress)
  - Add VARCHAR and CHAR types with limited field size
  - Add DATETIME
  - Add size ranges from 1 to 53 for FLOAT
  - Add synonyms for compatibility
    - BLOB for BINARY
    - TEXT for STRING
    - REAL for FLOAT
- SQL Semantics:
  - Sub-queries in IN, NOT IN, HAVING.
  - EXISTS and NOT EXISTS

Questions?