



Realise the promise of big data.

Simon Gregory
sgregory@hortonworks.com

Hadoop core concepts.

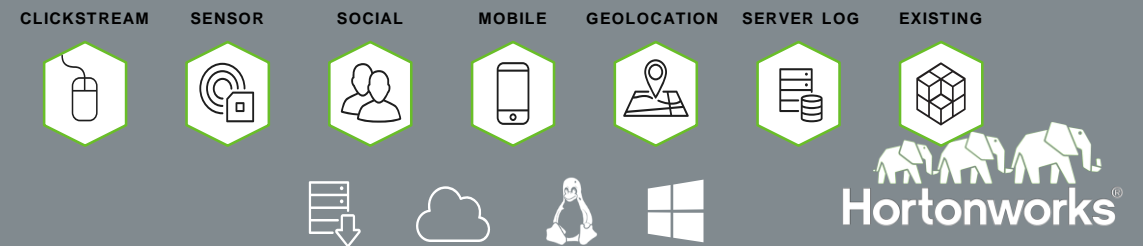
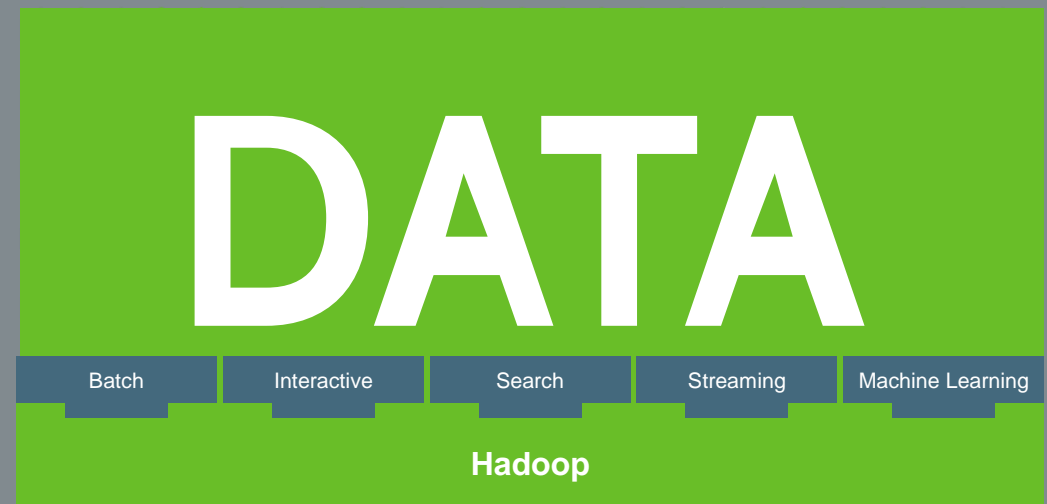
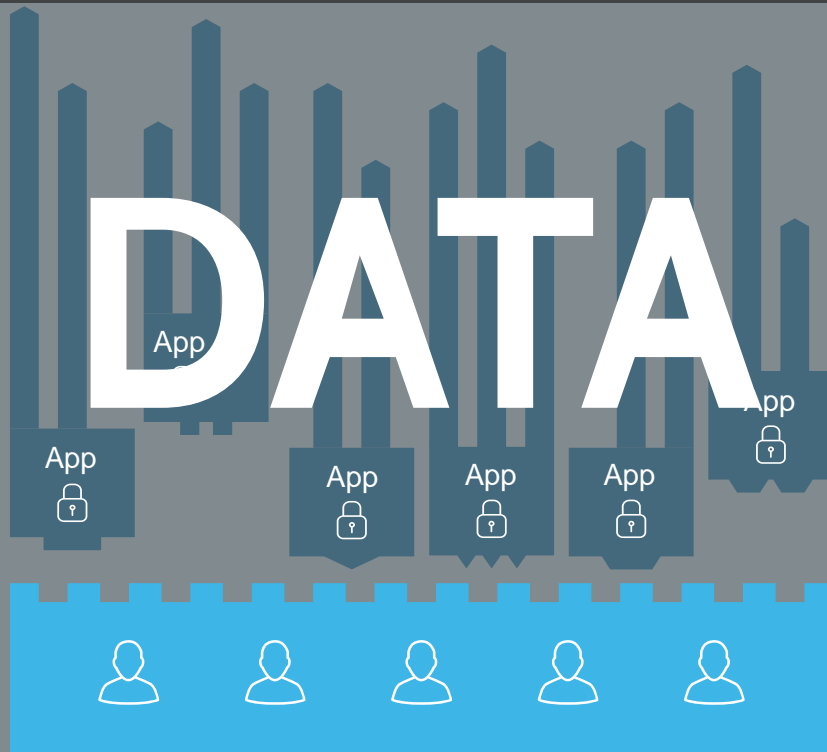
Purpose and outcome.

Store an unlimited quantity and variety of data in a single place, regardless of it's format or origin.

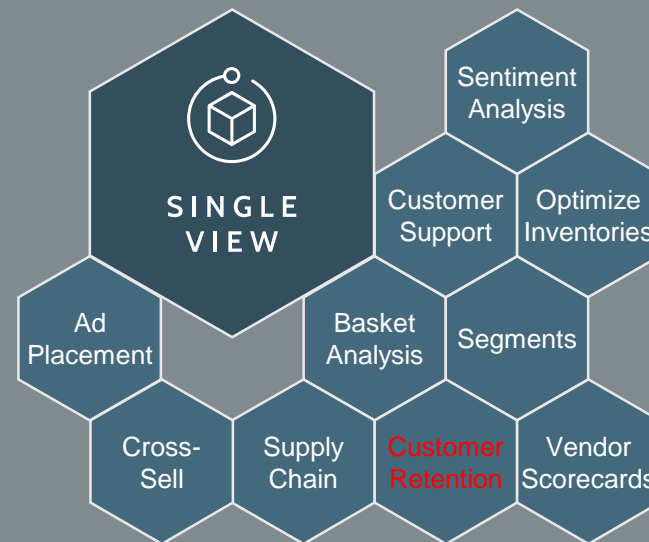
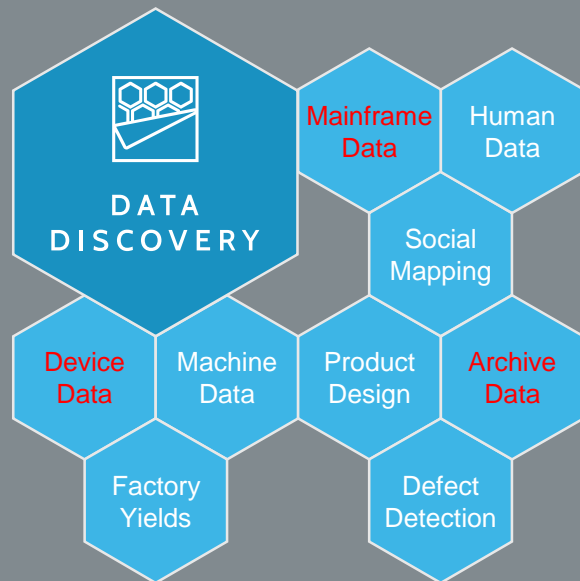
Process - merge, interact, sort, refine, normalise, index, expose and ANALYSE that data.

Operationalise that data for your benefit.

Hadoop core concepts: Part of the modern data architecture.



Why Hadoop & SAS? Challenges & Opportunities.





Audience understanding:

Find – Understand – Engage - Measure

Telco

Rogers Media is a subsidiary of Rogers Communications, which owns Canada's largest publishing company.

Why Hadoop and SAS?

Single View of the customer.

Rogers' Media Audience Platform: Integration and analysis of all data collected across the organization

Query all data in one location Blend of online and offline data, subscription, ecommerce, loyalty programs, etc.

Land massive click stream log files, 100+ M records / day, 30 million unique IDs / month

Use 100% of the data for Analysis and Visualization instead of smaller random samples (over sampling)

Identified and modeled more than 600 relevant web characteristics out of a field of 75,000 with SAS

Deliver products and solutions with relevance and timeliness.



Insurance Healthcare

Large US medical
insurer

Why Hadoop and SAS?

Data Discovery

Operationalise manual processes at scale:

Difficulty identifying coding errors among 300K daily claims

Using Analytics and Hadoop to improve reimbursement revenue and health outcomes

HDP + SAS: Marrying and analyzing numerous pool of data stored in multiple silo's—including gross margins, taxes, customer claims and policy premiums—to determine the company's potential exposure.

Ability to crunch several terabytes of data, and then revise, recalculate and report on that data on a regular basis.

Removing costly manual intervention and increasing accuracy.

Unable to monetize data.

Fragmented data silo's.

Banking

One of the largest US banks

Why Hadoop and SAS?

Data Discovery

Data flows into one data lake with centralized security policies, reducing storage costs

Multiple data sets extracted from source platforms with single point of security & privacy for de-identification, masking, encryption, authentication and access control.

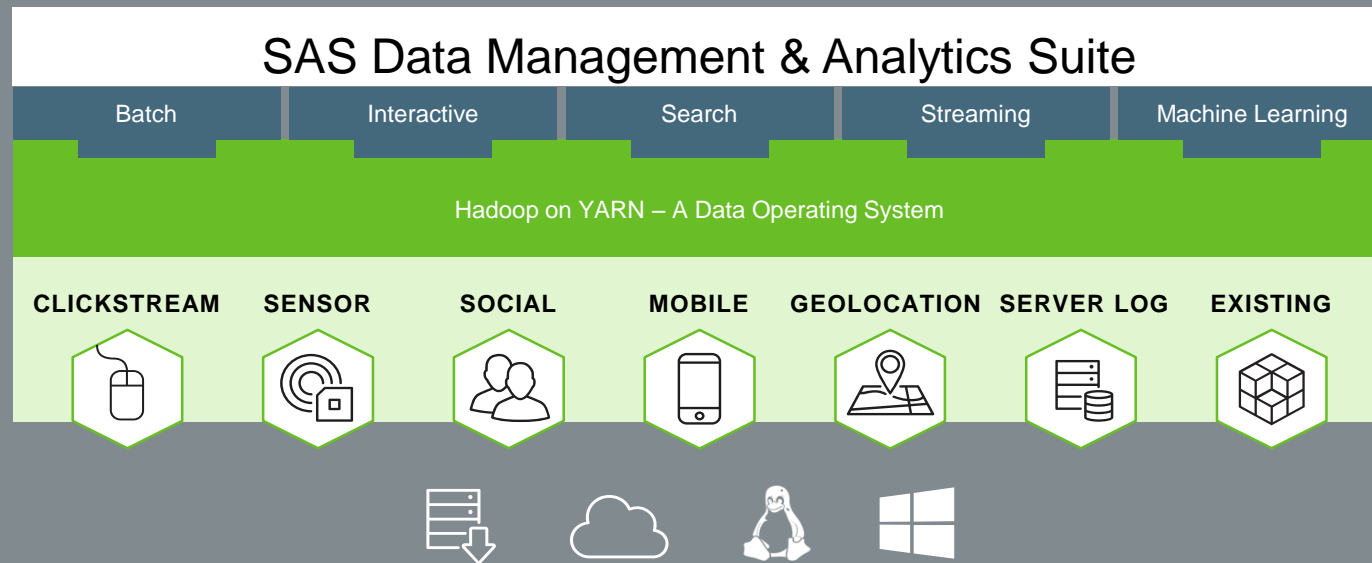
All departments have access to the same cross-sell data

Seamless integration with SAS

Hadoop & SAS: Why we do what we do...

Hortonworks HDP is Apache Hadoop.

SAS is the leader in analytics.



Skills & Processes.

Use existing skills and processes against net new data.

Data Management.

Ingest, transform, cleanse and tag data within the Hadoop eco-system.

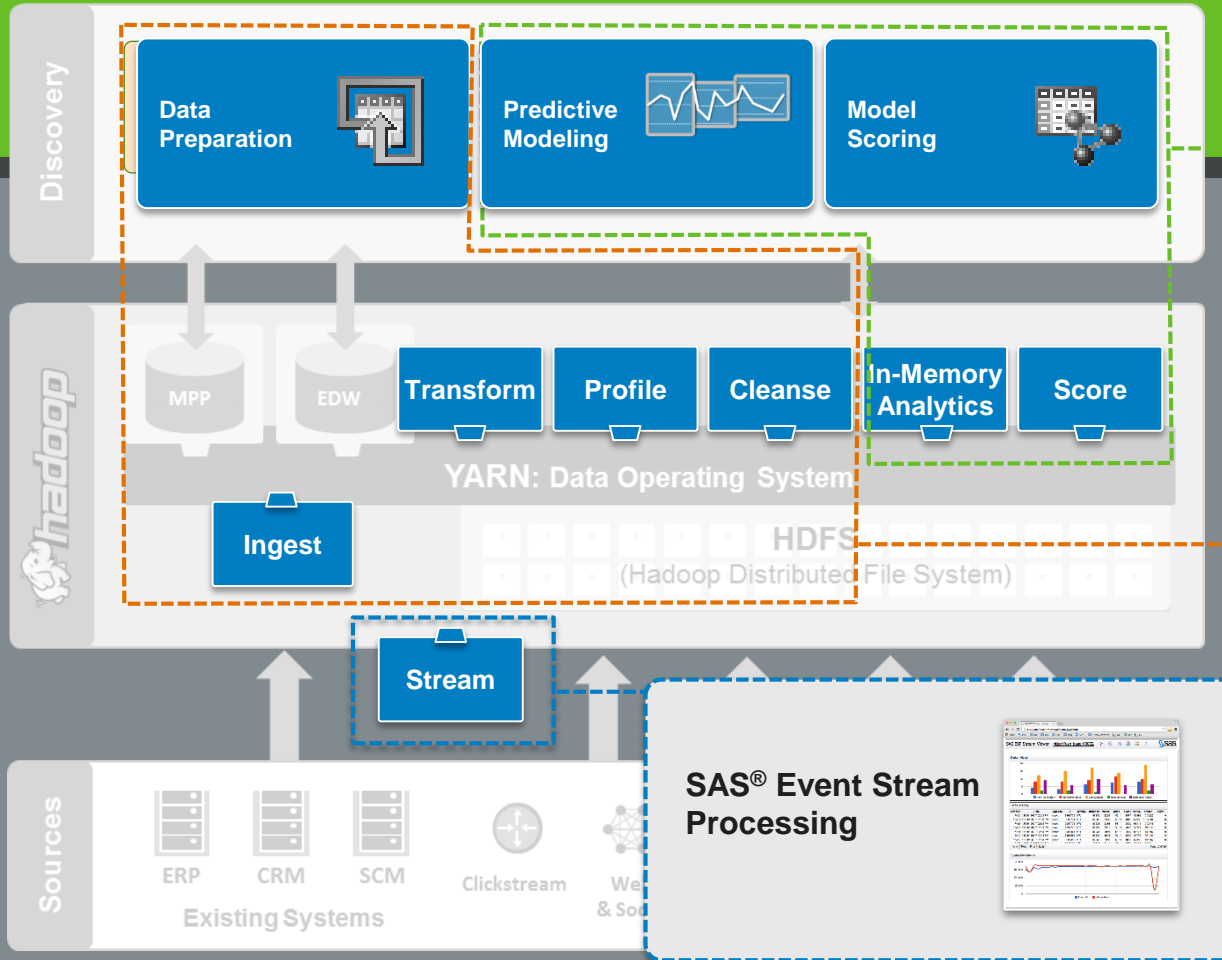
Operationalise Analytics.

Apply analytics and rules to pinpoint event relevance and urgency with continuous pattern detection.



Hortonworks & SAS combined...

Open Source Platform – Highly functional software solutions.



SAS® High-Performance Analytics

SAS® Visual Statistics

SAS® Scoring Accelerator

SAS® In-Memory Statistics

SAS/ACCESS® interface to Hadoop

SAS® Data Loader for Hadoop

The Hortonworks Methodology.

Open Source model defined

Architect

Introduce business requirements for planning future iterations of Hadoop

Support

Provide end user support to deployed Hadoop clusters, capturing new requirements



Develop

Work with the open source community, develop Hadoop projects to meet business requirements

Distribute

Distribute tested, packaged versions of Hadoop to end customers and their users

Summary

Why Hortonworks and SAS?

We provide our customers analytics against 100 % of the relevant data.

Provide customers a comprehensive insight into any entity

Provide customers the ability to make data derived decisions both pre & post transactional.

Questions?