Booz | Allen | Hamilton

SCAPE FRAMEWORK

Leverage Advances in Big Data to Make Multi-Sensor Data Discoverable

MARCH 2017

CONSULTING | ANALYTICS | DIGITAL SOLUTIONS | ENGINEERING | CYBER

AGENDA

1: CHALLENGES WITH GROUND SYSTEMS

2: BIG DATA ARCHITECTURE

3: APPLIED TO GROUND SYSTEM CHALLENGES

4: SCAPE ARCHITECTURE AND TECHNOLOGY

5: LESSONS LEARNED

CHALLENGES WITH CURRENT GROUND SYSTEMS

- **Data and Systems Are Closed**
  - Costly and difficult to modify (O&M)
  - Acquiring organization is often tied to a single provider/integrator/gatekeeper for that ground system

- **We Lack an Enterprise or Holistic View of Sensor and Data Environment**
  - Increasingly problems or decisions require more than a single sensor's data to solve

- **We Must Adjust Rapidly to the Changing World**
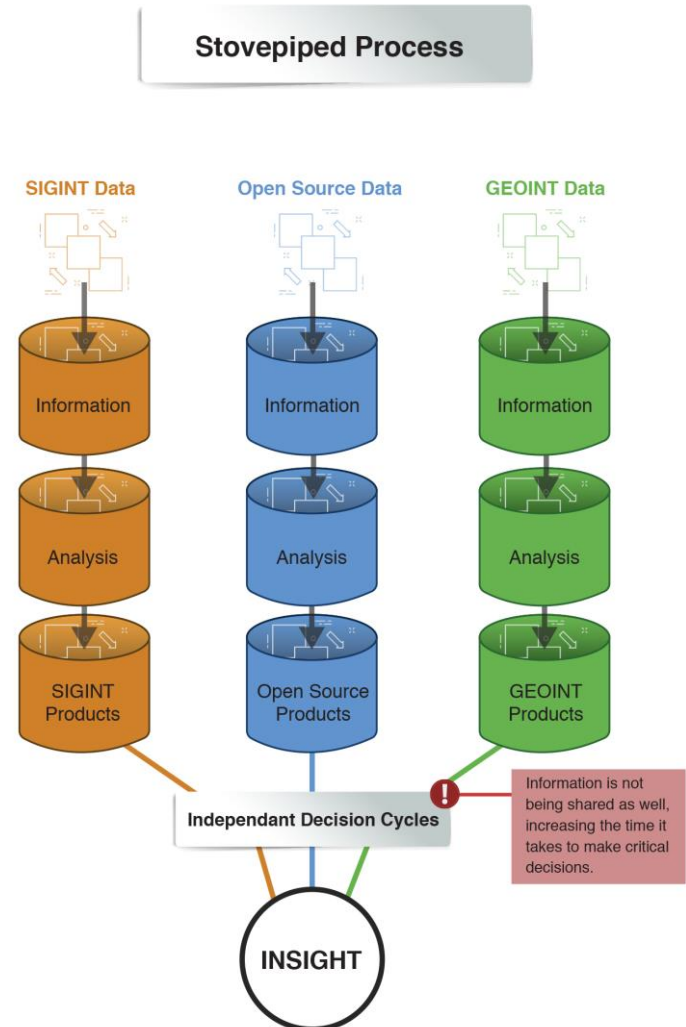
BIG DATA APPROACHES

- An enterprise consists of multiple streams of data coming from sensors that eventually make their way into a Data Lake. We can't observe the state of the enterprise until after the data arrives in the lake.
    - *We have a really good perspective on the past--*


- In the cloud, we can observe the entire state of the enterprise upon ingest, before data enters the Data Lake. Our cloud-based analytic system needs interactive, real time, predictive, and batch analytics.
    - *SCAPE adds a real time analytic capability to our cloud-base analytic platform.*
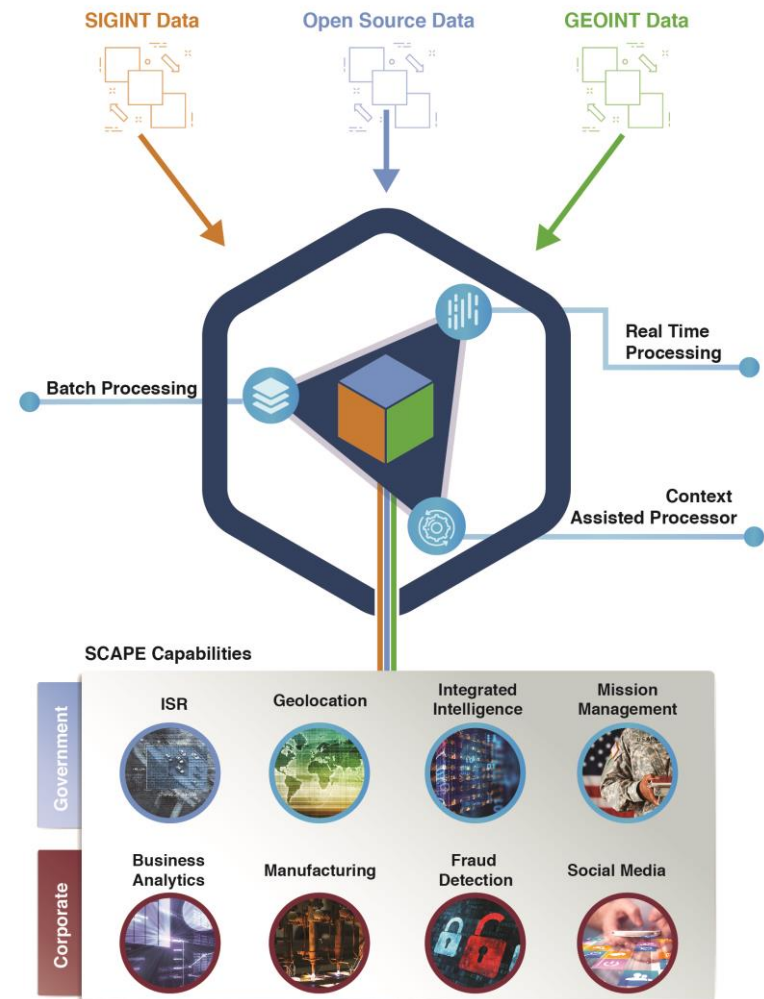
BIG DATA ARCHITECTURE

- The Booz Allen Cloud Analytic Reference Architecture provides a holistic approach to people, processes, and technology in four integrated layers

- Simple, well-documented integration points make it easy to continuously evolve Big Data systems to support new sensors, data streams, and analytics

- CARA uses the Lambda design pattern to combine streaming and batch data, leveraging context from the data lake to improve real-time analytics
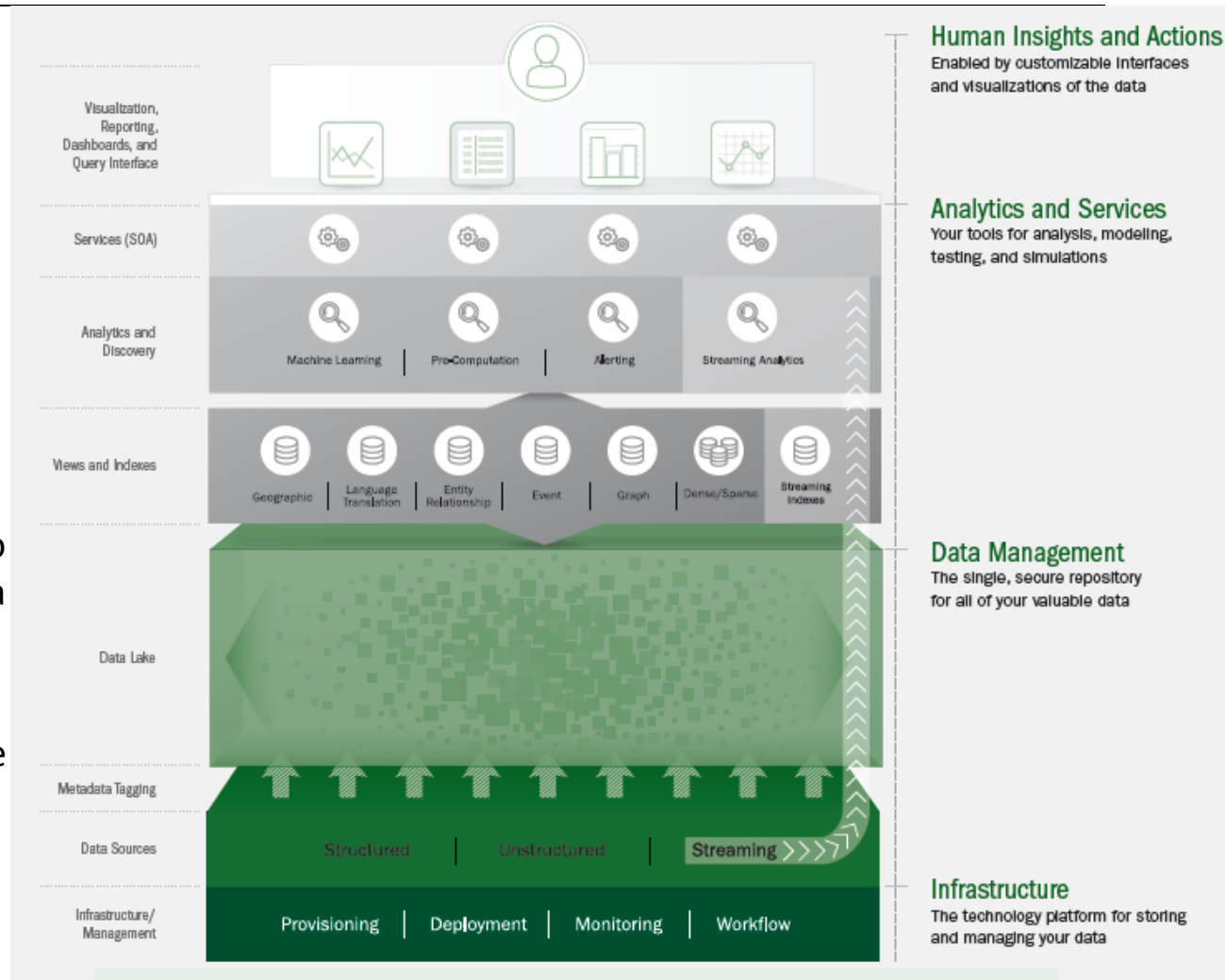
**Booz Allen Cloud Analytic Reference Architecture**

APPLYING CLOUD ANALYTIC ARCHITECTURES TO GROUND SYSTEMS

- SCAPE is an open framework and platform for ingesting and managing multi-sensor data with a focus on directing a multi-constellation enterprise

- Combines streaming real-time events with batch analytic results at scale for optimized interoperability

- Leverages Context Assisted Processor (CAP) subsystem to identify and apply contextual knowledge, enriching data value and real-time analytics

- Exposes data at multiple points within the processing pipeline for analysis and enrichment

SCAPE ARCHITECTURE AND TECHNOLOGY

- Apache Niagara Files (NiFi) – simplifies ingestion of structured data, provides many prebuilt types, with encryption and data provenance to enable secure solutions

- Apache SPARK – in memory big data processing, more complex processing faster, allows for combining stream and batch processing for powerful interactive applications

- Kafka Message Queues – enables data sharing and processing through the architecture

- Complex Event Processor – rules engine for responding to codified events in real time

- Data Registry – authoritative source for data, applications, services and relationships

Current SCAPE Architecture

6
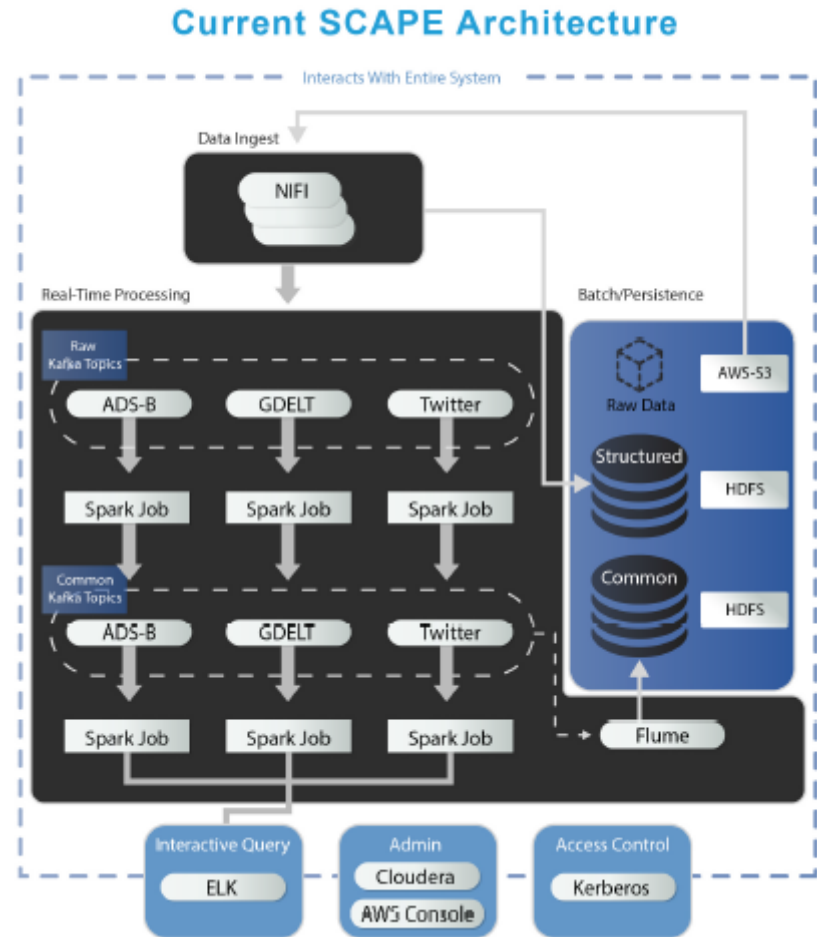
LESSONS LEARNED

- Leverage Open Architecture Principles to encourage Third Party Development of Applications and Analytics

- Design for self-service and discovery

- Use Commercially Viable Open Source Software as much as possible

- Favor Integration over Development

- Control the data – don't let it control you!

- Develop using Agile methodologies and DevOps practices where possible to match demand