THE DATA INFRASTRUCTURE
YOUR AGENCY HAS BEEN WAITING FOR

Data consolidation and data fusion have long been key goals for IT managers. Thanks to emerging technology, the enterprise data hub architecture will help your agency change the way you store, manage and gain new insights from your data. Our infographic explores how.

CHANGING THE BUSINESS OF GOVERNMENT

A recent GovLoop survey of 256 public sector employees explored the benefits of big data in government and found:

0 20 40 60 80 100

83% cited improved efficiency and productivity via data.

81% said big data drives improved decision-making.

73% cited transparency and accountability as a reason they use data.

73% said they use big data to manage resources, budgets and controls costs.

60% said big data helps control waste, fraud and abuse.

BIG DATA CHALLENGES

Although there are many benefits to big data, organizations face challenges in leveraging information. Our survey participants cited the following challenges:

Unclear mission and goals.

Lack of clarity on metrics.

Lack of education.

Lack of leadership support.

Data governance.

Data locked in legacy systems.

THE ENTERPRISE DATA HUB: A PLATFORM FOR INNOVATION

To overcome these challenges and capitalize on big data, your agency must explore the enterprise data hub architecture. Agencies today must lie information-driven while managing risk and cost. Cloudera's enterprise data hub, powered by Apache Hadoop™, provides a centralized, low-cost platform to efficiently and securely store, process, analyze, govern, archive, and serve any and all data.

Why an enterprise data hub "EDH" enables your agency to speak a common data narrative.

What is an EDH? With an EDH, data can be stored in its original fidelity, accessed through open computing frameworks, and integrated with existing infrastructures and applications. This provides agencies with a dynamic and growing range of computing capabilities needed for various workloads. Critical mission workloads such as batch processing, interactive SQL, faceted and full-text search, and advanced analytics all share the same infrastructure, with the same security, governance and management frameworks, yet only relying on single source data sets.

This frees data sequestered in standalone data silos, which leads to faster time-to-discovery and time-to-action and improves decision-making and mission-driven outcomes. Most importantly, the convergence of data and computing builds a foundation – and culture – for data innovation.

Whether you are a chief data officer, analyst or a manager, making good decisions depends on your data. An EDH connects you to information like never before, helping you gain valuable insights and make stronger data-driven decisions.

Benefits for CIO & CDO

- Creates a common data narrative across agency
- Sets a strategic vision
- Consolidates systems
- Promotes data stewardship

Benefits for IT Managers

- Integrated, holistic security
- Audit trail of data and activity
- Comprehensive resource management
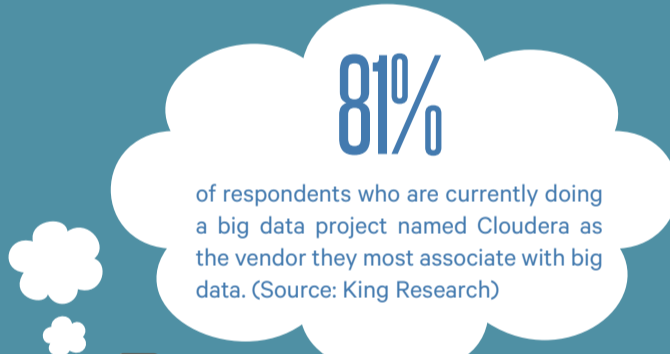- Compliance-ready

Benefits for Analysts

- Freedom of analysis
- More types of and more visibility of data
- Use any and all data for improved decisions
- Agile modeling and processing

HOW CLOUDERA CAN HELP

Cloudera will help you transform how you leverage data with products like Cloudera Manager and Cloudera Navigator and with solutions like Cloudera Search, Cloudera Impala, and Apache Spark ™.

81% of respondents who are currently doing a big data project named Cloudera as the vendor they most associate with big data. (Source: King Research)

Managing the EDH

Although an EDH presents opportunities to gain new insights from data, leaders must have the confidence they can manage, govern, and secure the EDH. Cloudera brings native systems and data management applications for Hadoop.

Cloudera Manager

Cloudera Manager supports the deployment, configuration, and management of an enterprise data hub to ensure quality of service and performance and streamline operations - all at scale.

- Deployment and Configuration Services
- System Health and QoS/SLA Notification
- Business Continuity and Cluster Maintenance Services
- Root Cause Discovery, Analysis, and Solutions
- Enterprise Security, Monitoring, and 3rd Party Integration

Cloudera Navigator

Cloudera Navigator is the only native end-to-end data governance solution for Hadoop - providing data management capabilities for users, administrators, and auditors, helping them to govern, explore, and secure the data within an EDH.

- Data and Services Usage and Permissions Audit
- Metadata, Data Discovery, and Data Exploration
- Data Lineage and Provenance
- Data Lifecycle Management
- Data Protection and Key Management

New Computing Capabilities

Cloudera and an EDH provide new computing capabilities to help meet your mission in transformative ways.

Cloudera Search

Cloudera Search brings full-text, faceted, and interactive search and scalable, flexible indexing to an enterprise data hub.

- Get Answers Quickly with Intuitive, Mission-Specific Language
- Accelerate Data Exploration and Data Fusion
- Improve Efficiency of Indexing and Storage
- Streamline Operations with other EDH Workloads

Cloudera Impala

Cloudera Impala is the the leading open source analytic database that runs natively in Apache Hadoop.

- Bring Business Intelligence to Big Data with analytic SQL
- Broaden Analysis to All Historical and Operational Data
- Use Data through Compliance-Ready Security and Role-Based Access
- Share Data with Other EDH Workloads without Replication

Apache Spark

Apache Spark is an open source, parallel data processing framework that works with Apache Hadoop storage and resource engines.

- Meet Multiple Mission Needs with a Single Computing Engine
- Build Powerful, Fast Parallel Applications with Familiar Tools
- Accelerate Data Pipelines and Machine Learning and Modeling
- Expand the Capabilities of the EDH to New Computing Approaches

LEARN MORE ABOUT CLOUDERA AND THE EDH

The EDH for the Public Sector

Cloudera offers public sector entities a secure and cost-efficient place to store and analyze all their data, empowering them to derive new insights and confidence while extending the value of existing investments. Visit Cloudera's website to discover the advantages of an EDH for:

- Data & Application Consolidation
- Advanced Data Fusion & Analytics
- Improved Security Intelligence & Fraud Detection
- Infrastructure & Skills Optimization

Try Apache Hadoop Now

Cloudera Live is the fastest and easiest way to get started with Hadoop and it now includes self-guided, interactive demos and tutorials. With a one-button deployment option, agencies can spin up a four-node cluster of CDH, Cloudera's open source Hadoop platform, within minutes. This free, cloud-based Hadoop environment lets you:

- Learn the basics of Hadoop (and CDH) through pre-loaded, hands-on tutorials
- Plan your Hadoop project using your own datasets
- Explore the latest features in CDH
- Extend the capabilities of Hadoop and CDH through familiar partner tools and applications

clouderagovt.com

cloudera.com/live

cloudera | **govloop**

Sources:
http://www.marketwired.com/press-release/survey-says-cloudera-leads-in-big-data-1850491.htm
http://www.slideshare.net/cloudera/leading-future-of-big-data-32682662
http://www.esd.doc.gov/Reports/fostering-innovation-creating-jobs-driving-better-decisions-value-government-data