

Accelerate your Big Data Strategy

Execute faster with Capgemini and Cloudera's Enterprise Data Hub Accelerator

Executive Summary



Capgemini and Cloudera have collaborated to build a comprehensive approach to help organizations accelerate and execute on Big Data initiatives.

The Enterprise Data Hub Accelerator is an execution framework for Big Data, built around Cloudera Enterprise, an Apache Hadoop-based open-source enterprise data management platform. It helps you to define your first projects, execute them well, and grow them into a fully-defined and sustainable Big Data strategy, from all angles, because Big Data is not only about technology.

In this white paper, we illustrate how the Enterprise Data Hub Accelerator enables you to get started rapidly and cost-effectively with an action-focused plan for your first Big Data project.

The Big Data revolution is underway

Leading organizations are changing the way they think about data, transforming it from a cost to a strategic asset – and a way to gain an edge on the competition.

They are using Big Data solutions to get to a complete view of their business and markets and to increase customer interaction and understanding, grow revenue, improve operations and profitability.

To get meaningful and actionable insights, however, business leaders are demanding integration of all their data sources, including:

- **External sources** – third party data, competitive data, open data, social media
- **New data sets** – machine log data, sensor data
- **Traditional data sources** – online transactional data such as those from CRM and ERP systems, as well as data warehouses and data marts

These data are currently not well integrated. They are stored in different formats and in different places. They are worlds apart.

What's more, traditional business intelligence solutions suffer from serious short-comings, including:

1. **Data “blindness”** - silos create a lack of visibility to all data
2. **Delays** - analytical architectures are too rigid and slow
3. **Cost** - adding more data requires significant additional infrastructure spend
4. **Complexity** – making it nearly impossible to give business users real insight

Together, Capgemini and Cloudera can address these challenges for you and help you take action now.

Cloudera's enterprise data hub – the core of next generation data management platforms

How to extract new value from data and deliver tangible business benefits?

Cloudera's enterprise data hub is the core of the solution. Based on the open source technology of Apache Hadoop and delivered via Cloudera Enterprise, Cloudera's enterprise data

hub gives you one centralized platform for processing all your data and managing it in an agile, cost-efficient, and sustainable way.

Instead of moving data around between multiple or point-solutions, it provides a single store for all data, for as long as it is needed. The data is stored “as is” and transformed when necessary for a specific purpose. This means the raw data can be reused as necessary without additional movement.

An enterprise data hub is open and flexible. It can be integrated with existing infrastructure and applications (ETL engines, reporting and advanced analytics tools) to run a variety of enterprise workloads – including batch processing, interactive SQL, enterprise search, stream processing, machine learning, and advanced analytics.

It is also enterprise-ready with robust security, governance, data protection, and management capabilities. And it is proven: organizations such as financial institutions and government agencies have been running this technology for many years, with the appropriate security and data privacy measures.

Your core technologies are already certified on Cloudera's platform

Cloudera's open platform and rich partner ecosystem - more than 900 partners from the “existing BI” as well as the “new Big Data” worlds - give you maximum flexibility to integrate your existing and new systems, tools, and applications (ETL, analytic, and reporting tools). This openness ensures that your enterprise data hub investment is there for the long run, capable of integrating data from and to legacy technologies as well as emerging solutions.

The partners involved in this ecosystem include:

HP, IBM, Informatica, MicroStrategy, MongoDB, Oracle, Pentaho , QlikTech, Revolution Analytics, SAP, SAS, Splunk, Syncsort, Tableau, Talend, Teradata, TIBCO.

How to put Big Data to work for your organization

How do you progressively build an enterprise data hub for your organization?

Capgemini and Cloudera have partnered to build an execution framework to help you do this. This complete approach is built around:

- **Five pillars for Big Data execution:** Business Drivers, Governance, Analytics, Data, and Platform, to address all dimensions of how your organization manages data.
- **Four stages of maturity** - Initiate, Scale, Transform and Run, to enable progressive maturation while showing results quickly.

The Five Pillars of Information-Driven Business Transformation



Pillar 1 – Identify your Business Drivers

The end game for Big Data is to help transform an organization's business. Big Data technologies can be used to support the creation of new value-added services (eg. Connected Devices), improve customer loyalty (eg. Customer Experience), or develop new streams of revenue or costs optimization.

To enable this, setting up business drivers early on is key. Setting the bar at the right level for the first step is also important to be able to show tangible success quickly and visibly. As your Big Data strategy matures, business drivers will be refined and optimized, but without them it is difficult to engage business teams and IT teams together in a fruitful way.



Pillar 2 - Enhance your data culture with the right Governance

Creating a well-established but agile governance model has traditionally been a major challenge. This is where Big Data solutions like Cloudera Enterprise can help by unleashing the possibilities of modern data management and analytics.

Enabling the right amount of agility and exploration, along with the right processes around data management (eg. data quality, data privacy) will guarantee that your Big Data initiatives move in the right direction and at the right speed.



Pillar 3 - Do more with your data – try new Analytics

Big Data initiatives are about looking at new solutions and techniques to get more insight out of all data. Analytics is central to this process:

- Improve prediction efficiency - by being able to store more data and see a longer history.
- Get more agile - analyzing data in-place. You can focus on the pure analytics part of the job, as data provisioning is done only once.

Taking your analytics processes to the next level with the capabilities of the new platform is a key dimension of the journey to the enterprise data hub.



Pillar 4 – Get more Data, get it ready

Cloudera Enterprise reduces the cost and complexity of data storage. It also stores all data in its original format for as long as you need it. By migrating and integrating both your internal and external data into an enterprise data hub you will have made a significant step towards solving the issue of data silos.

Our approach will help you draft the right roadmap to provision all the data needed as fast as possible, and set up the right processes to bring new datasets in as needed.



Pillar 5 - A data management Platform at the service of your business, not the other way around

An enterprise data hub lets you stop defining your business needs based on the limitations of your data management technology. Cloudera Enterprise makes it easier, faster, and more cost effective to manage all data in the long run, without worrying constantly about capacity planning.

The Four Stages of Maturity on your Road to Big Data

Capgemini and Cloudera have come together to help clients act now on the promises of Big Data - helping them to start small, explore the potential, test the waters, and move forward with confidence. We have clearly defined four key stages on the road towards implementing Big Data solutions across the enterprise.

Initiate

Kick-start the execution of your first Big Data project and build your first success by taking advantage of Capgemini accelerators. These help you identify the right projects while ensuring they deliver measurable benefits.

Scale

Grow your first projects on a bigger scale and demonstrate that success is repeatable and reliable, and that the foundations are solid for the long run.

Transform

Transform your operations to take advantage of the insights delivered by the new analytics such as machine learning and discovery algorithms.

Run

Deliver Big Data insights throughout your organization by focusing on industrialization and optimizing costs while increasing agility and capability.

Capgemini and Cloudera help you get value out of Big Data

Capgemini has collaborated extensively with Cloudera to build an integrated approach to leveraging Big Data. The Big Data revolution is here and the opportunities are at hand. But how do you "do it"?

Together we have already helped clients in telecommunications, the public sector, consumer products, retail, manufacturing and financial services. We have achieved this by utilizing Cloudera's expertise in Big Data platforms combined with Capgemini's global capabilities: more than 9,600 professionals in Business Information Management.

Capgemini and Cloudera help you become a truly innovative, information-centric company. Together, we address the full potential for Big Data within your organization, and we make it work for you.

We can get you started right now.

cloudera®

Cloudera is revolutionizing enterprise data management by offering the first unified Platform for Big Data, an enterprise data hub built on Apache Hadoop™. Cloudera offers enterprises one place to store, process and analyze all their data, empowering them to extend the value of existing investments while enabling fundamental new ways to derive value from their data. Only Cloudera offers everything needed on a journey to an enterprise data hub, including software for business critical data challenges such as storage, access, management, analysis, security and search. As the leading educator of Hadoop professionals, Cloudera has trained over 20,000 individuals worldwide. Over 1,000 partners and a seasoned professional services team help deliver greater time to value. Finally, only Cloudera provides proactive and predictive support to run an enterprise data hub with confidence. Leading organizations in every industry plus top public sector organizations globally run Cloudera in production.

More information is available at
www.cloudera.com



About Capgemini

With more than 130,000 people in over 40 countries, Capgemini is one of the world's foremost providers of consulting, technology and outsourcing services. The Group reported 2013 global revenues of EUR 10.1 billion.

Together with its clients, Capgemini creates and delivers business and technology solutions that fit their needs and drive the results they want. A deeply multicultural organization, Capgemini has developed its own way of working, the Collaborative Business Experience™, and draws on Rightshore®, its worldwide delivery model.

For further information visit

www.capgemini.com/bim

or contact us at

bim@capgemini.com