Pivotal Big Data Suite

Product Suite

Overview

DATA AND DIGITAL TRANSFORMATION

Great software companies leverage Big Data to fundamentally change the user experience, reach new levels of profitability and efficiency, and even pioneer entirely new business models. They understand how to put data to work, delivering information in context for maximum impact.

How are companies that have software as a core competency able to achieve this? There are four major trends underlying modern, cloud-native data architectures:

- **Open source data management innovation:** Providing extreme scale and performance advantages built for the cloud.

- **Cloud-native application platforms:** Enabling continuous delivery, automated operations and agile software development.

- **Data microservices:** Loosely coupled services architecture bounded by context to surface insights within applications.

- **Machine learning at scale:** Advanced analytics and data science against large volumes of data to glean predictive insights.

Today traditional enterprises across vertical markets — from retail and manufacturing to healthcare and finance — have the opportunity to likewise use software, data and analytics to disrupt their own industries.

Enterprises that take advantage of this opportunity and embrace software and analytics as core competencies are the ones that will become the leaders in today's increasingly digital economy.

Pivotal BDS is available under a single, flexible license that covers three offerings – Pivotal Greenplum, Pivotal HDB and Pivotal GemFire.

PIVOTAL GREENPLUM

- **Based on Open Source and Open Platforms:** Flexibility of deployment options, including appliance, software for on-premises or cloud environment, or as a managed service.

- **Scalable and Performant:** Linear scalability, query optimization for complex SQL with concurrent and mixed workloads, and hybrid row/columnar massively parallel processing architecture and petabyte-scale data loading.

- **Mission Critical:** Built-in security, authentication, authorization and permissions model, connects to Kerberos or LDAP, and has fine grained permissions and group management.

- **Rich Analytics:** Full SQL compliance, PostGIS geospatial analytics; Apache MADlib for machine learning; GPText, NLTK, OpenNLP; pl/Python, pl/R, pl/Java for custom analytics.

- **Robust Data Management Framework:** Servers can be added while database remains online and fully available. Performance monitoring framework allows separation of hardware and software issues. One unified framework for monitoring, administration and workload management.

Pivotal Big Data Suite

GemFire   Greenplum   HDB

Pivotal

ANALYZE AND OPERATIONALIZE BIG DATA

To take advantage of this opportunity requires a modern, cloud-native data platform that allows enterprises to leverage open source software innovation, lower the cost of asking questions of their data, and support event-driven architectures. The Pivotal Big Data Suite (BDS) is the only comprehensive platform of open data management solutions that meets these demanding requirements.

Pivotal BDS allows enterprises to modernize their data architecture, discover more insights from their data with advanced analytics, and build analytic applications at scale. Pivotal BDS enables enterprises to:

- Analyze data of all types - structured, semi-structured and unstructured data.

- Leverage existing SQL skills and related tools to perform complex analytics and interactive queries against petabyte-scale data.

- Build, run and iterate machine learning and data science algorithms at scale to glean predictive insights.

- Develop and support smart data-driven applications that operationalize Big Data insights by delivering information in context.

THE PIVOTAL BIG DATA SUITE

The Pivotal BDS can be deployed on bare metal, in virtualized environments or in the public cloud, and is made up of three primary solutions delivered under a single, flexible license.

- Pivotal HDB is a Apache Hadoop-native analytical database powered by Apache HAWQ (incubating) that combines exceptional MPP-based analytics performance with robust ANSI SQL compliance. It enables high performance ad hoc queries and predictive analytics on data stored in HDFS using SQL syntax and related tools, as well as in-database machine learning with Apache MADlib (incubating.)

- Pivotal Greenplum is an advanced, fully featured, open source MPP analytical database for powerful and rapid analytics on petabyte scale data volumes. Uniquely geared toward Big Data analytics and machine learning, Greenplum enables in-database machine learning with Apache MADlib (incubating) and delivers high analytical query performance on large data volumes due to the world's most advanced cost-based query optimizer.

- Pivotal GemFire is a distributed, in-memory data grid powered by Apache Geode (incubating) that is designed to support high-volume, low- latency, mission-critical, data-driven operational and transactional applications. GemFire-powered applications operationalize Big Data insights by notifying applications of new and updated data, and can process many simultaneous operations and maintain sub-second response time at linear scale.

PIVOTAL HDB

- **Advanced Analytics Performance:** Interactive SQL query execution against complete datasets of virtually any size.

- **Most Complete Language Compliance:** Broad list of standard SQL analytic functions, supporting ANSI SQL-92, SQL-99, and SQL-2003. Extensible framework supports user-defined functions in multiple languages such as PL/R, PL/Python, PL/C and more.

- **Integrated Machine Learning:** Supports advanced, highly scalable, predictive analytics algorithms running in-database on full datasets (vs sample data sets) for better models and results.

- **Federated Data Query:** Analyze disparate datasets and formats in Apache Hadoop from different locations or formats including external data in HDFS, Hive, and HBase with easy reference to HCatalog tables.

- **Best-in-class Query Optimizer:** Purpose-built Cost-Based Optimizer selects the best possible query execution model for the best possible performance.

- **Elastic Architecture for Hybrid Cloud:** Scale-up/down or scale-in/out in any on-premises or cloud computing environment.

Pivotal®

In addition, the Big Data Suite includes support for Spring XD, a distributed and unified runtime to develop and orchestrate stream and batch processing pipelines.

SUMMARY

At Pivotal, our mission is to transform how the world builds software. With the Pivotal BDS, enterprises have all the tools they need to put their data to work. So together, let's build something meaningful.

PIVOTAL GEMFIRE

- **Predictable Low Latency:** In-memory, horizontally scalable architecture for low latency application requirements. Grid-aware queries and operations routed to nodes holding relevant data for processing.

- **Elastic Scale-Out:** Scale out horizontally and scale back down again gracefully to keep steady state runtime costs down and maximize efficiency. Adding nodes increases capacity predictably.

- **Real-Time Event Notifications:** Applications can subscribe to real-time events to react to changes immediately as data comes into the system, while reducing the overhead on your SQL database.

- **High Availability and Business Continuity:** Automatic failover to other nodes in the cluster in case of failures. Grids resiliently rebalance and reform if nodes leave or join the luster. WAN replication allows for multi-site, global-scale and disaster recovery deployments.

- **Durability:** Data is durable through in-grid, in-memory replication, shared-nothing architecture and persistent write-optimized disk stores. Multiple checks ensure transactional consistency of data.

Pivotal transforms how the world builds software. Pivotal combines the Silicon Valley state of mind, modern approach, and infrastructure with organizations' core expertise and values.

We enable the leading companies in the world to innovate by employing an approach focused on building—not buying—software.

Pivotal 3495 Deer Creek Road Palo Alto, CA 94304 **pivotal.io**

Pivotal®