instaclustr

Apache Spark™ is a high performing engine for large-scale analytics and data processing, While Apache Spark™ provides advanced analytics capabilities, it requires a fast, distributed back-end data store. Apache Cassandra® is the most modern, reliable and scalable choice for that data store. Spark when fully integrated with the key components of Cassandra, provides the resilience and scale required for big data analytics.

Instaclustr's Managed Apache Spark provides a reliable and managed platform, collocated with your Apache Cassandra® data store, to leverage the power of Apache Spark for stream or batch analytics.

Features

Collocated Data Engine
Spark engine resides in the operational database, so there is no need for extracting, transforming and loading into a new environment.

Performance
Spark has an advanced DAG execution engine that supports cyclic data flow and in-memory computing. It runs programs up to 100x faster than Hadoop MapReduce in memory, or 10x faster on disk.

Monitoring
We provide integrated Spark management and monitoring through the Instaclustr Management Console as well as provide 24x7 monitoring and support to our Spark customers.

Unified Memory Management
Spark comes with the shared memory pool for execution and catching.

Reliability
Instaclustr Managed Spark architecture and support offering enables you to use the power of Spark from your application, with the confidence to meet your availability and processing requirements.

Automated Provisioning of Spark Jobserver and Apache Zeppelin
To provide easy access to your Spark processing engine, Instaclustr's Spark cluster can include Spark Jobserver (REST API) and Apache Zeppelin (analyst notebook UI).

About Instaclustr

Instaclustr delivers reliability at scale through our integrated data platform of open source technologies such as Apache Cassandra®, Apache Kafka®, Apache Spark™ and Elasticsearch.
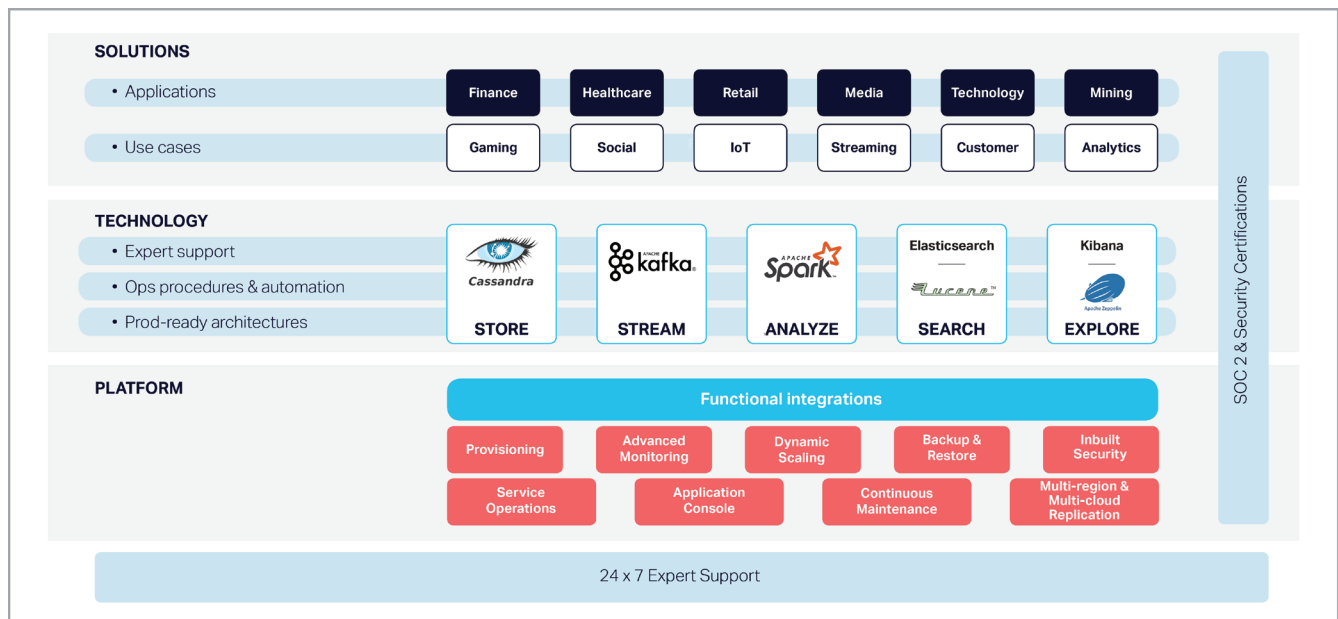
Our expertize stems from delivering more than 25+ million node hours under management, allowing us to run the world's most powerful data technologies effortlessly.

We provide a range of managed, consulting and support services to help our customers develop and deploy solutions around open source technologies. Our integrated data platform, built on open source technologies, powers mission critical, highly available applications for our customers and help them achieve scalability, reliability and performance for their applications.

Our Services

Instaclustr provides range of consulting and support services to help our customers develop and deploy solutions that use Apache Cassandra and related technologies.

Managed Solutions	Our managed solutions enable our customers to run seamlessly in the cloud with confidence that our team is monitoring and managing their clusters.
Consulting Services	Our team of experts have extensive experience with a wide range of use cases and deployed architectures.
Enterprise Support	Our dedicated 24x7 support team of experts have hundreds of thousands of hours of operational engineering and support experience.