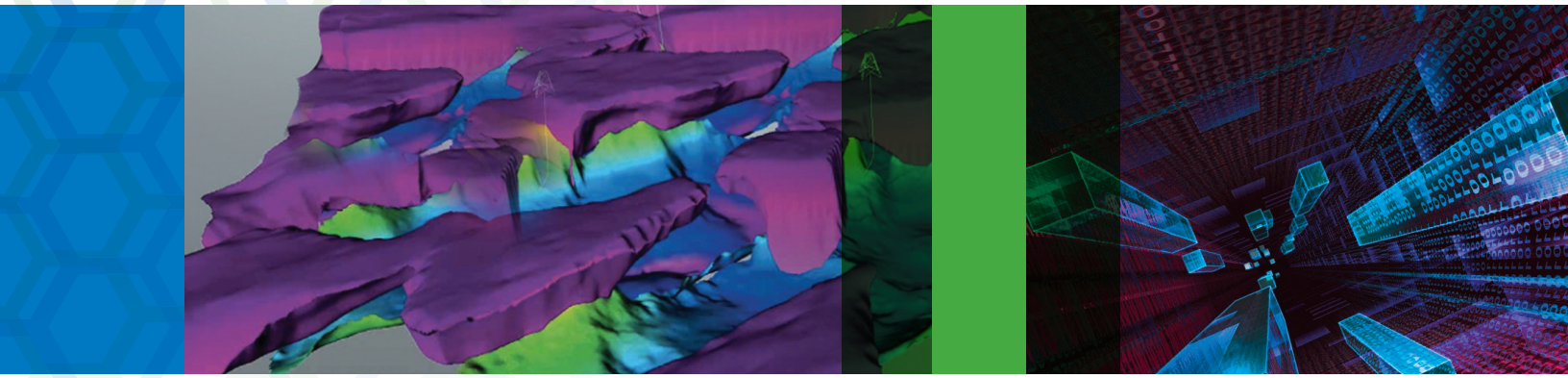
SGI Hadoop Clusters

Data Analysis in Real-Time, Not Next Time

Key Features

Hadoop starter-kits are factory integrated solutions that plug-in and go!

Design-to-order capabilities create unprecedented speed and scale solutions from the leader in Hadoop scale



SGI Big Data Leadership

Hadoop is the preferred method of managing and analyzing large unstructured data. SGI has been one of the leading suppliers of Apache Hadoop servers since the technology was first introduced – including both large commercial enterprise deployments and optimized clusters in key government agencies. SGI has also been a leader in complementary high speed, high-capacity data ingest and large-scale in-memory analytics using SGI UV and previous-generation technologies for over 15 years. Perhaps most importantly, SGI is ideally positioned to deliver clusters that fit an organization's unique requirements, and has built and deployed Apache Hadoop installations as large as 40,000 nodes and individual clusters as large as 4,000 nodes.

- **Plug In and Go!** SGI understands that organizations need solutions that are ready to deploy and that come with enterprise-class 7x24 services. SGI provides true factory-integrated Hadoop solutions that include servers, storage, networking, and installed Cloudera software, delivered ready to run.
- **End-to-end solutions for data analytics.** With a broad range of systems and storage solutions, SGI provides systems for end-to-end data analytics. SGI can provide complementary solutions for essential functions such as data ingest and archive, and systems with both horizontal (scale-out) and vertical (scale up) scalability.
- **Delivering unprecedented speed and scale.** With a strong background in high-performance and technical computing, SGI pushes the envelope in terms of both performance and scalability, as evidenced by continuous completion of world-record benchmarks. SGI provides a choice of high-performance Ethernet or InfiniBand backbone networking, and superior cluster manageability at scale using SGI Management Center.
- **The world's only power-optimized Hadoop cluster.** With growing power costs, data centers increasingly require power management solutions that optimize performance at the lowest possible power consumption. SGI provides the world's only power-optimized Hadoop cluster, offering fine-grained, dynamic, and policy-driven power management that avoids the hassle and risk of powering down individual servers.

A Flexible and Complete Hadoop Solution

SGI systems for Apache Hadoop combine enterprise ready distributions from Cloudera with flexible deployment options and an ecosystem of data analytics applications.

Apache Hadoop from Cloudera and an Expanding SGI Analytics Ecosystem

Individual organizations have their own unique analytics objectives that must be served by their Apache Hadoop infrastructure. By partnering with Cloudera, SGI provides the premier Apache Hadoop-based distribution in both commercial and non-commercial environments.

- *Cloudera's Distribution Including Apache Hadoop* (CDH) delivers a streamlined path for putting Apache Hadoop to work solving business problems in production. It is ideal for enterprises seeking an integrated and tested Apache Hadoop-based system without proprietary vendor lock-in.
- *Cloudera Enterprise* is a subscription offering that enables data-driven enterprises to run Apache Hadoop environments cost effectively in production with repeatable success.

Complementing Cloudera's offerings, SGI has expanding partnerships with industry-leading analytics, visualization, and tool vendors. Full feature trial applications from Datameer, Kitenga, Quantum4D, and Pentaho are all pre-installed on SGI Hadoop solutions (Figure 1), and can be easily extended for production use.

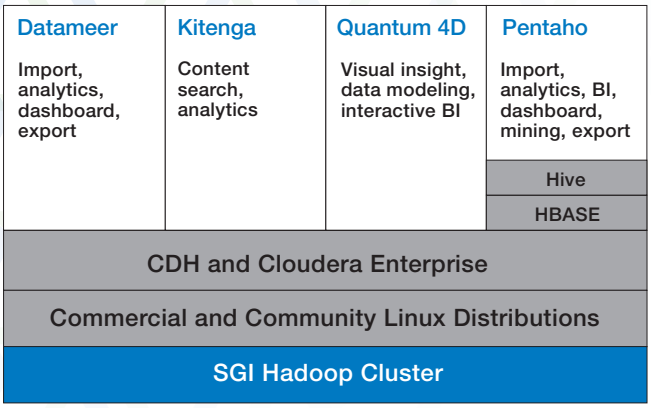


Figure 1: SGI provides a complete hardware and software stack for Apache Hadoop, including Cloudera software and popular analytics components.

Flexible Apache Hadoop Deployment Options

In addition to flexible software choices, SGI lets organizations choose deployment options that best fit their needs, including SGI Hadoop Starter Kits or full custom Design to Order (DTO) clusters (Table 1).

Table 1. SGI provides flexible Hadoop cluster options.

SGI Hadoop Cluster Option	Number of Servers	Capacity
Half Rack Starter Kit	• 20	• 160 TB
Full Rack Starter Kit	• 40	• 320 TB
Multi-rack	• Any	• Petabytes

- **SGI Hadoop Starter Kits.** Offering the fastest time to production, SGI Hadoop Starter Kits are pre-configured, factory integrated, and certified by Cloudera. With both half-rack and full-rack options based on SGI Rackable servers using the highest-performance Intel® Xeon® processor E5 family, these kits are purpose-designed to run Apache Hadoop.
- **Design to Order Hadoop Clusters.** For multi-rack deployments, SGI Design to Order (DTO) clusters are custom designed to meet exact customer needs. Custom considerations can include performance, density, power, cooling, networking, and cost requirements.

Global Sales and Support: sgi.com/global

©2012 Silicon Graphics International Corp. All rights reserved. SGI and the SGI logo are registered trademarks or trademarks of Silicon Graphics International Corp. or its subsidiaries in the United States and/or other countries. All other trademarks are property of their respective holders. 13072012 4381