Zettaset Secure Data Warehouse and Hadoop

The Zettaset Secure Data Warehouse (SDW)™ solution provides the foundation for a secure, robust, and scalable data warehouse in today's Hadoop-based, distributed processing environments.

The Evolving Data Warehouse Environment

Data warehousing supports the long term structured storage of data for eventual access and analysis. A data warehouse is no longer defined as a DBMS running a centralized repository. A data warehouse is first and foremost a data integration and rationalization engine, and as such requires multiple technologies in order to operate. The need to extract value from Big Data is driving organizations to seek advanced capabilities from data warehouse providers, including extended query-driven analytics, time series analysis and causal analysis of extended data structures beyond relational data.

Distributed processing, which is represented by MapReduce and solutions that emulate this type of processing, such as Hadoop, are in high demand. Organizations that were previously content to complete such distributed processes on separate clusters of servers and storage devices now expect this capability to be integrated with existing data warehouse architectures and analytic engines.

Unique Security Challenges and Requirements

Data warehousing poses its own set of challenges for security. One major challenge is that enterprise data warehouses are often very large systems, serving many user communities with varying security needs. Thus, while data warehouses require a flexible and powerful security infrastructure, the security capabilities must operate in an environment that has stringent performance and scalability requirements.

Many of the basic requirements for security are well-known, and apply equally to a data warehouse as they would to any other system: Unauthorized users must be prevented from accessing or modifying data; the applications and underlying data must not be susceptible to data theft by hackers; the data

must be available to the right users at the right time; and the system must keep a record of activities performed by its users.

These requirements are perhaps even more important in a data warehouse because a warehouse contains data consolidated from multiple sources, and therefore, a data warehouse can be one of the most lucrative hacking targets in an enterprise.

A robust security infrastructure which offers protection can vastly improve the effectiveness and reduce the costs associated with a data warehouse environment.

Data Warehouse Security - Use Case Scenarios

Typical customer scenarios for data warehouse security include the following:

- An enterprise is managing a data warehouse that will be widely used by many divisions and subsidiaries. This enterprise needs a security infrastructure that ensures the employees of each division will be able to view only the data that is relevant to their own division, while also allowing employees in its corporate offices to view data for all divisions and subsidiaries.

- An enterprise's data warehouse stores personal information. Privacy laws may govern the use of such personal information. The data warehouse must handle data so as to adhere to these laws.

- An enterprise sells data from a data warehouse to its clients. Those clients may view only the data to which they have purchased or to which they have subscribed. They should never be permitted to see the data of other clients.

Zettaset Secure Data Warehouse Features and Benefits

Granular User Access Control for Compliance and Policy Management

Zettaset Secure Data Warehouse (SDW) includes Role-based Access Control (RBAC) to augment Kerberos authentication. This provides robust group and user access controls which further strengthen the security offering. RBAC also enables distributed cluster security to be integrated into existing security policy frameworks managed through LDAP/AD, and helps Hadoop clusters address compliance requirements.

High Availability and Reliability

Automated backup and failover of all critical services within the Hadoop cluster infrastructure ensures high availability and zero data loss. Service monitors are self-healing and alert administrators of any cluster health or utilization issues. Each key service in the system is monitored and mirrored for consistency and redundancy, while status information from each node is collected and aggregated.

Ease of Use and Deployment

Web-based interfaces minimize user error by providing a high level dashboard view of the cluster, while eliminating unnecessary complexities. A powerful web-based API supports all major programming and scripting languages to ease integration with existing data warehouse applications. An installation wizard simplifies and speeds up deployment.

Zettaset Secure Data Warehouse (SDW) Solution

The sheer volume of Big Data has magnified the data warehouse security challenge. In response, Zettaset has developed an enterprise software solution designed for the distributed architectures emerging in today's data warehouse environments, and has built security into every component of the system. The Zettaset SDW solution is based on reliable, scalable and secure Hadoop technology.

Hadoop functions as the back-end data store and in a distributed processing environment, and is able to process both structured and unstructured data for subsequent analysis. Hadoop provides additional benefits to the environment by enabling the data warehouse to store much more data then it normally would, using the legacy central repository approach.

With Hadoop's distributed processing architecture, a data warehouse can easily store data at the petabyte level, well beyond the typical terabyte or gigabyte levels that exist today.

CLOUD PLATFORM SERVICE	amazon web services, rackspace the open cloud company
BUSINESS ANALYTICS APPLICATION	ORACLE, Dell, Microsoft
BIG DATA EXTRACT/ TRANSFORM/LOAD (ETL)	INFORMATICA, pentaho, PERVASIVE
HADOOP CLUSTER MANAGEMENT	Zettaset
OPERATING SYSTEM	ORACLE LINUX, redhat, CentOS
HADOOP DISTRIBUTION	intel, cloudera Ask Bigger Questions, Hortonworks, IBM, hadoop
CLUSTER SERVERS	IBM, HITACHI Inspire the Next, FUJITSU, Dell, hp
DATA CENTER/ INTERCONNECT SWITCHES	BROCADE, Mellanox Technologies, ARISTA, CISCO
HIGH DENSITY STORAGE CLUSTER	violin MEMORY, FUSION-io
STORAGE ARRAY	NetApp, EMC² where information lives

The Zettaset Secure Data Warehouse (SDW) solution addresses the need for secure, high-performance Hadoop cluster management for Big Data deployments within a "best-of-breed" data warehouse reference architecture.

Data Warehouse Reference Architecture

The reference architecture within the data warehouse may vary as well as the supporting vendors, but foundationally a common approach is necessary. This includes a combination of complementary and interoperable systems: From the storage array, to the interconnect, to the Hadoop clusters and the business intelligence (BI) and analytics applications.

A New Approach to Securing the Data Warehouse

The Zettaset SDW solution provides a secure environment that can scale with business needs. It solves the technical problem of log aggregation in a distributed network, and provides components that are necessary to meet enterprise security, audit and compliance needs. Zettaset SWD is built upon the Zettaset Orchestrator™ enterprise management solution for Hadoop Cluster management.

Zettaset Orchestrator is an enterprise software solution that automates, accelerates, and simplifies Hadoop installation and cluster management for Big Data deployments, and delivers faster time to value. Zettaset has created a solution that meets the exacting requirements of enterprises for security, high availability, and performance within the Hadoop cluster environment. This is particularly critical in highly regulated verticals such as financial services, healthcare, and government.

Security and Performance… without Compromise

Security is at the core of the Zettaset value proposition. The Zettaset Secure Data Warehouse (SDW) solution comes with all critical levels of protection enabled by default and delivers this protection without compromising performance, security, or control. Zettaset SDW can be run as a standalone solution or seamlessly integrated within an existing environment so that data can be structured for analysis and potential threats identified.

About Zettaset

Zettaset is an enterprise software company and the leader in Big Data management and security. Zettaset provides its critical enabling technology to end-user customers through a network of strategic partners, including Brocade, IBM, and Synnex-Hyve. **For more information, contact sales@zettaset.com**

Zettaset

1945 Landings Drive, Mountain View, CA 94043 // USA: +1.650.314.7920 // Fax: +1.650.314.7950
sales@zettaset.com // www.zettaset.com