

VMware vFabric Data Director

Database as a Service in Your Cloud

AT A GLANCE

VMware® vFabric™ Data Director is a database-virtualization and lifecycle-management platform that powers database as a service in your cloud, increasing agility and reducing database TCO. vFabric Data Director reduces capex through database-aware virtualization on VMware vSphere®, increases agility through automated lifecycle management, and accelerates analytics and application development through policy-based self-service. Designed for both enterprises and service providers, vFabric Data Director empowers administrators to securely automate and delegate routine tasks, including heterogeneous data engine provisioning, backup and cloning.

vFabric Data Director currently supports Oracle 10gR2, 11gR2, SQL Server 2008 R2, SQL Server 2012, VMware vFabric Postgres 9.1.6, and Hadoop 1.0 distributions, with more data engines to be supported in the future. Hadoop support is powered using Serengeti, an open source project initiated by VMware for deployment and management of Hadoop clusters in virtual environments.

KEY BENEFITS

- Reduce database hardware and license costs through database-aware virtualization on vSphere.
- Increase agility through automated database lifecycle management.
- Accelerate analytics and application development through policy-based self-service.

What is VMware vFabric Data Director?

Enterprises and service providers are rapidly implementing next-generation, self-service cloud platforms that enable IT to quickly provision virtual machines and other core infrastructure components, proving the benefits of cloud computing. Now users want similar benefits from the more-complex software solutions that are powered by the core infrastructure. Data engines are the next frontier of virtualization, including databases and Hadoop clusters. Databases are being virtualized to address issues of growing database infrastructure, long lead times for developers, overworked DBAs, and excessive operational cost of databases. Hadoop is the upcoming technology for analyzing petabytes of data, i.e., big data analytics. Current implementations on physical hardware are extremely costly due to low utilization, excessively time consuming, and inherently unreliable due to many single points of failure in Hadoop.

VMware vFabric Data Director is a software solution that enables you to implement database-aware virtualization on vSphere, and provide database as a service in your cloud. It automates deployment, management and governance of heterogeneous data engines, and enables policy-based self-service database management for application developers. vFabric Data Director combines the convenience and ease of use of public cloud services with the enterprise-grade security, flexibility, control and compliance required to build your own private and hybrid clouds.

How Is vFabric Data Director Used?

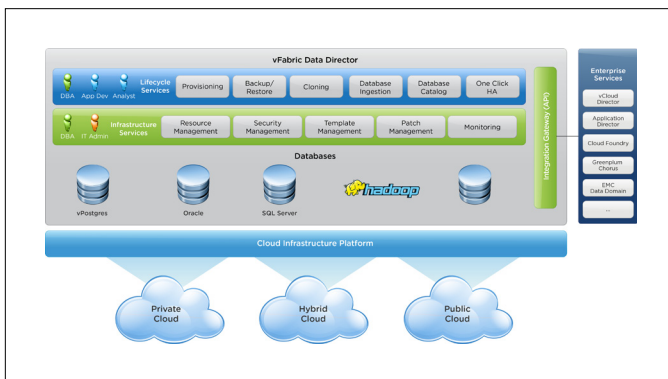
Enterprises and service providers use vFabric Data Director as a unified platform to enable database as a service for heterogeneous data engines. vFabric Data Director is installed on top of vSphere 5.0 Enterprise or Enterprise Plus.

What Are the Key Capabilities of vFabric Data Director?

vFabric Data Director is a unified virtualization platform for managing the lifecycle of heterogeneous data engines. vFabric Data Director currently supports Oracle 10gR2, 11gR2, SQL Server 2008 R2, SQL Server 2012, VMware vFabric Postgres 9.1.6, and Hadoop 1.0 distributions including Apache Hadoop 1.0.x, Cloudera CDH3, Greenplum HD 1.1 and 1.2, and Hortonworks HDP-1.

vFabric Data Director delivers three sets of capabilities:

- **Database-aware virtualization**
- **Automated lifecycle management**
- **Self-service**



VMware vFabric Data Director powers Database-as-a-Service for Your Cloud.

Database-aware Virtualization

By using database virtualization, businesses are achieving cost savings and operational benefits through consolidation. Virtualizing databases requires a series of steps, and DBAs need to be familiar with the virtual infrastructure and operations. vFabric Data Director simplifies this process by providing a unified platform with integrated capabilities. It enables DBAs to easily virtualize databases on the proven vSphere infrastructure:

- **Resource management and isolation** – vFabric Data Director offers flexible resource management and isolation at the tenant and database level, thereby eliminating the “noisy neighbor” problem—interference caused by multiple databases sharing the same pool of resources.
- **Integrated template management** – Administrator-controlled database templates help to guarantee corporate compliance, enforce standardization and enable fast database provisioning.
- **Database ingestion** – DBAs can quickly copy the database from a physical server into a virtualized database managed by vFabric Data Director, reducing the complexity usually associated with migrating a physical database into a virtual environment.

Automated Lifecycle Management

vFabric Data Director has a rich set of capabilities that help DBAs automate lifecycle management and improve productivity:

- **Powerful Web-based management console** – IT operators can manage thousands of databases through a single pane, significantly reducing management overhead.
- **Database provisioning in minutes** – Using database templates, vFabric Data Director helps you provision new data engines, including Oracle, SQL Server and Hadoop clusters, within minutes instead of days.
- **Innovative cloning** – The average production database uses six clones for development and test purposes. The linked database clone feature in vFabric Data Director reduces cloning time from days or weeks to minutes, regardless of the source database size.
- **Policy compliance** – It is important to ensure that sensitive data is not exposed to unauthorized users during copy creation, and to destroy each copy on a timely basis to control sprawl and release unused resources. With postclone scripts, DBAs can enforce procedures such as data masking to ensure compliance, and can implement retention policies that enforce deletion of copies.
- **Database-aware high availability** – vFabric Data Director provides single-click high availability for all databases by leveraging the underlying vSphere High Availability capabilities.

Self-Service

Database as a service becomes a reality with vFabric Data Director—a single platform that helps deliver self-service capabilities for heterogeneous data engines:

- **Self-service database provisioning** – Self-service database provisioning provides developers with database access in minutes while reducing IT operational costs.
- **Self-service database management** – Developers need the flexibility to perform basic database operations without involving DBAs. Using self-service capabilities, developers can perform common operations such as cloning, backup, snapshots and point-in-time restore with just a click.
- **Robust role-based access control (RBAC)** – RBAC restricts system access to authorized users. By assigning granular privileges, IT administrators can enable self-service while maintaining control and compliance.

Provision and Manage Highly Available, Elastic, Multi-tenant Hadoop in Minutes

Provisioning a Hadoop cluster typically require large amount of hardware, complex setup and can take days to weeks to provision. vFabric Data Director leverages Serengeti, an open source project initiated by VMware, to provision, configure, and scale a Hadoop cluster on vSphere with a simple command or a few clicks. Data Director supports major distributions of Hadoop including Apache Hadoop 1.0.x, Cloudera CDH3, Greenplum HD 1.1 and 1.2, and Hortonworks HDP-1.

With Data Director, customers can run Hadoop with other workloads on existing vSphere clusters, and scale Hadoop on demand.

Increase Availability of Hadoop Clusters

Although Hadoop is known to provide reliability via data replication, there are certain major components that are single points of failure in the system. Examples include the ‘namenode’, the ‘jobtracker’ and other supporting components. Data Director can address the high availability needs of all these components in a generic way with vSphere vMotion™, High Availability (HA) and Fault Tolerance (FT) features, keeping your Hadoop system running efficiently on your virtual infrastructure, with minimal or no downtime.

Find Out More

For information or to purchase VMware vFabric Data Director, call 855 TRY VFABRIC (879 8322) or +1 650 427 3100, visit www.vmware.com/vfabric, send email to vfabric@vmware.com, or search online for an authorized reseller. For detailed specifications and requirements, refer to the product documentation.

