Validated Designs for Analytics Solution Brief

Validated Design for VMware Tanzu Greenplum

How to build and deploy an all-in-one solution for analytics at scale

Why VMware Tanzu Greenplum?

Powered by massively parallel Postgres, VMware® Tanzu™ Greenplum® handles SQL-based analytics of structured and unstructured data.

Administrators can easily extend Tanzu Greenplum with analytical tools, such as Apache® Spark® and commercial data science engines.

Build and maintain your data advantage

Businesses are generating data at an exponential rate. Modern analytics applications consume and generate terabytes — even petabytes — of data to fuel accurate insights. To remain competitive, businesses need to store, collect and analyze these massive amounts of data at speed. They also need the scalability to keep up as the business continues to collect and analyze data.

However, most IT organizations spend a great deal of time managing a highly siloed data center infrastructure, leading to a number of drawbacks:

- **Higher cost:** Redundant replication of data across silos increases hardware budgets and day-to-day administrative requirements.
- **Data inconsistency:** Data is replicated and transformed based on various application needs, resulting in inconsistency between different system reports.
- **Security concerns:** Enterprises have a broad attack surface for bad actors, further complicated by various requirements, access rights, user counts and physical security. Securing and tracking access across silos is difficult.
- **Maintenance requirements:** Basic administration of data infrastructure is tedious, often involving maintaining numerous copies of the same data.

In a data-driven market, siloed, single-purpose solutions are no longer tenable. Many enterprises are building data strategies based on systems designed to service a comprehensive range of analytics and AI requirements across a wide variety of use cases.

A combination of VMware Tanzu Greenplum, VMware virtualization, Dell EMC PowerEdge servers and PowerScale storage can provide a flexible, all-in-one solution that answers these needs. This combination is also cost-effective, easy to build and simple to manage.

A scale-out solution for enterprise analytics

VMware Tanzu Greenplum with Dell EMC PowerScale and Dell EMC PowerScale Isilon storage provides a scale-out solution for analytics that is space-efficient, performant, and easy to scale. This combination has proven to be flexible and addresses several industry-standard use cases — such as landing zones for data ingestion, data lake storage, data science, data warehouses and data marts — in a single design that's easy to deploy and support.

Use cases

- Landing zones for data ingestion
- Data lake storage
- Data science
- Data warehouses
- Data marts

Learn more

- Validated Design
- tanzu.vmware.com/greenplum-reference-architecture
- delltechnologies.com/vmware
- Analytics InfoHub
- delltechnologies.com/analytics

Elastic infrastructure for VMware Tanzu Greenplum

The all-in-one elastic infrastructure described in the Validated Design creates a modern solution for abstracting and managing infrastructure with a robust and cloud-ready approach to enterprise data needs. The Validated Design outlined in the paper is an example of how these technologies can work together to streamline scalability, provide better integration into data systems and evolve systems to be more cloud-ready.

Reference architecture technical specifications

The document describes a four-node VMware vSAN™ cluster connected to a pair of switches in a high availability configuration designed to meet enterprise standards. Each server has four network interfaces, and these are split off into the VMware services network and the Greenplum network.

Compute	Storage	Virtualization	Data services
Dell EMC PowerEdge R740xd vSAN Ready Nodes	Dell EMC PowerScale Isilon H500 storage	• VMware vSphere® • VMware vSAN	• VMware Tanzu Greenplum • Apache Spark • Apache Kafka®

Together, these components provide a data ecosystem that can scale to nearly unlimited data sizes and comes with enterprise management capabilities to simplify management and lower the total cost of ownership.

Expert help when you need it

Dell Technologies analytics experts are available at every step of the analytics journey. From workshops and planning sessions to IT solutions created from our extensive portfolio of workstations, servers, storage, networking, software and services, Dell Technologies and VMware can help you unlock the value of your data. Connect with one of the worldwide Dell Technologies Customer Solution Centers for proofs of concept, design sessions and/or technical deep dives today.

vmware®

DELLTechnologies