# Architecture Overview

The Palantir server architecture was designed to be robust, scalable, open, and extensible. A system is considered robust if it can handle the full range of inputs and situations in a given environment. The Palantir architecture is able to handle a wide range of data sources.

Scalability refers to a system's ability to perform as the size of the problem it is solving increases. Palantir scales as you add users (or clients) to the system. It also scales as you add more data to it. For example, searches continue to be fast whether a user is searching 1000 or 100,000 records.

Palantir Technologies also sought to build an architecture that is open and maintainable. Palantir is open in that it works within your existing network infrastructure. It uses standard protocols over your standard network. Additionally, Palantir provides a rich set of APIs that you can use to access the features of Palantir software either from your own applications or from third-party applications.

The following diagram depicts the Palantir server architecture. The diagram does not show a Phoenix cluster. The *Palantir Phoenix Guide* illustrates the Phoenix architecture.

Palantir Server Architecture
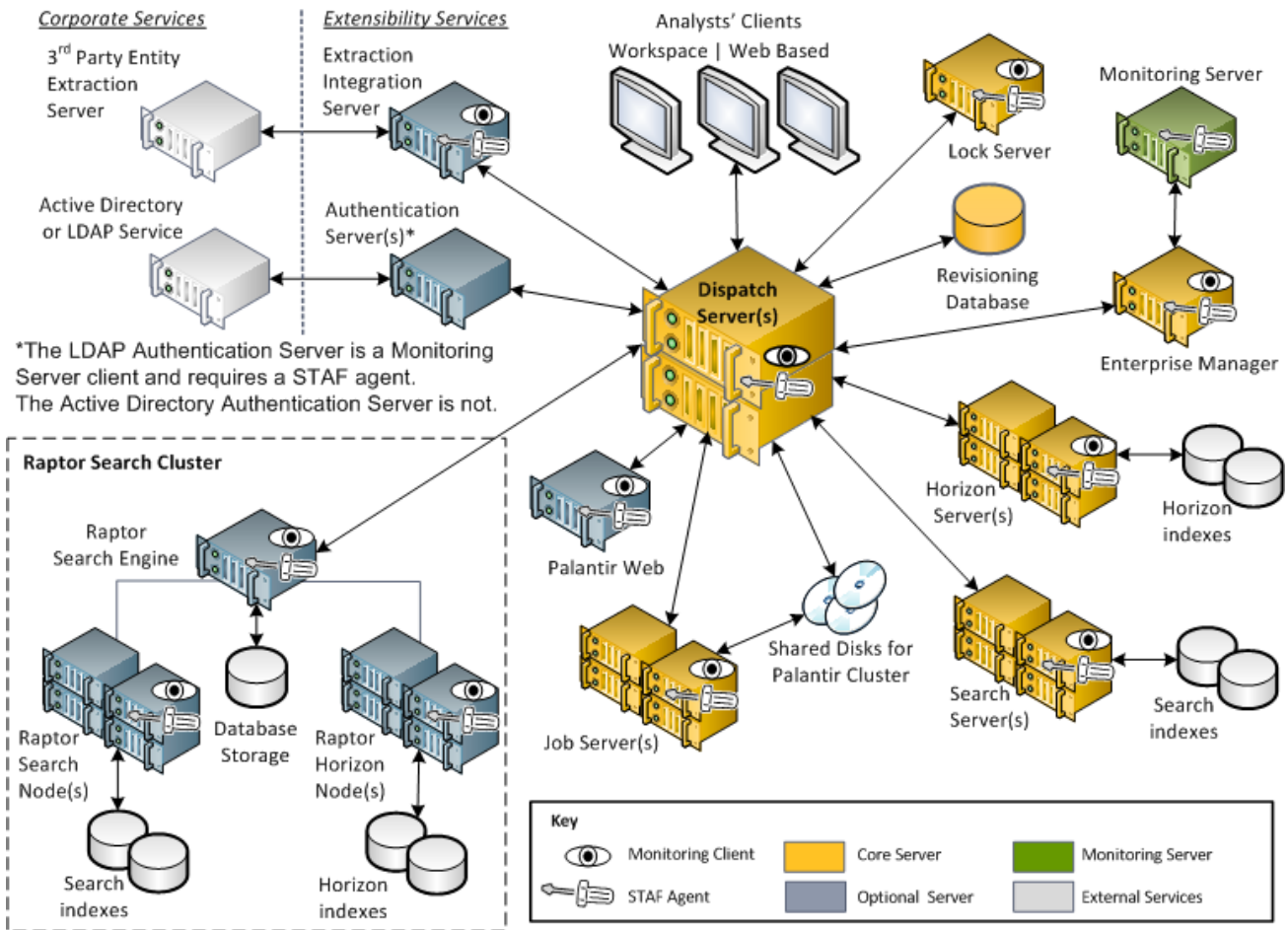
The Palantir core product servers can search and analyze data that has been imported into, or created in, Palantir. The core product consists of a Dispatch Server, a Job Server, a Search Server, a Horizon Server, and the Palantir Enterprise Manager.

Dispatch Server

The Dispatch Server is the gateway to the Palantir Platform. The Dispatch Server handles business logic, manages security and policy controls, and responds to user requests. The following diagram illustrates the components of the Dispatch Server.

Dispatch Server Components
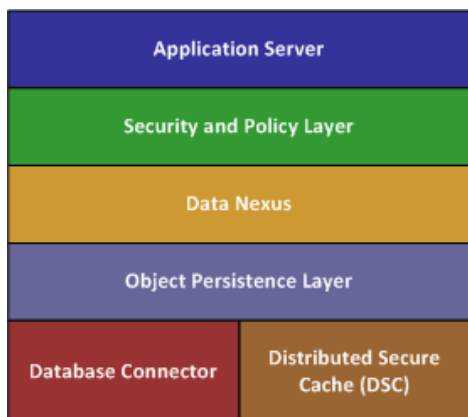
You can install two or more Dispatch Servers in a cluster. When clustered, the Dispatch Servers provide better responsiveness when you have fifty or more concurrent users.

The Dispatch Server connects to a database that holds the Palantir Revisioning Database. It maintains both the access control defined by the administrator and the ontology or schema you define. In fact, the Dispatch Server is the only way that clients can connect into the Revisioning Database, the Search Server, or Raptor.

The Dispatch Server sends queries to the Search Server for processing. It creates and schedules jobs through the Job Server.

Lock Server

If you have clustered Dispatch Servers, then you must use a Lock Server. The Lock Server manages in-memory locks for multiple Dispatch Servers against the Revisioning Database.

Job Server

Requests for job processing come into the Job Server from the Dispatch Server. The Dispatch Server stores the data for processing in shared storage. The Job Server accesses the shared storage, processes the data, and returns it back to the Dispatch Server. Effectively, the Job Server is a bulk client to the Dispatch Server. The following illustration shows the components of a Job Server.

Job Server Components
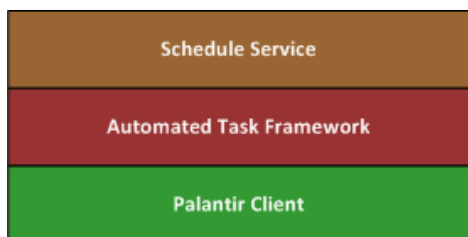
The Job Server processes bulk data import requests and other large data processing jobs. Bulk data import requests are typically back end import tasks. Bulk data imports occur when any of the following conditions are true:

- amount of data to be imported exceeds 5MB
- number of data sources to be imported is greater than 100
- number of estimated entities in the data is greater than 2000 (before any internal resolution)

Bulk data imports use Palantir customizations to Hadoop Map/Reduce services. The customized Hadoop services are installed automatically with the Job Server. To learn more about the Job Server, see the *Palantir Administration: Basic Configuration Guide*.

Finally, the Job Server is responsible for system operations, for example, an administrator-initiated job such as an LDAP-authentication sync. You can add multiple Job Servers to your installation as needed. To learn more about configuring multiple servers, see the *Palantir Installation Guide*.

Search Server and Horizon Server

These servers support queries against the Data Repository. The Dispatch Server handles requests and sends them to the optimal server depending on the nature of a search. Generally, the Search Server best handles searches against document text, while the Horizon Server optimally handles searches coming out of the Object Explorer application and the SearchAround feature.

To learn more about mirroring and partitioning Search and Horizon servers in a deployment, see the *Palantir Installation Guide*.

Search Server Operation

The Search Server indexes the content in the Revisioning Database and provides search capabilities. The Search Server is built on top of the Apache Lucene Information Retrieval Library. This library has excellent text processing capabilities that allow it to do full-text searches across structured data (for example, entities and events) and unstructured data such as documents. The following diagram illustrates the components of the Search Server.

Search Server Components
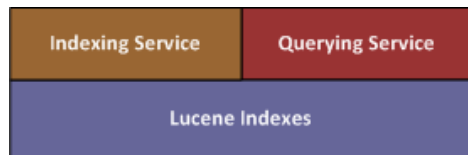
Palantir has heavily modified the Lucene implementation, adding the enterprise-level security and scaling characteristics of an enterprise-grade server. This server is integrated with the dynamic ontology, and it includes a number of customizations to the fuzzy-search capabilities of the information retrieval library.

The Palantir implementation of Lucene was designed to be database-agnostic. Palantir modified the index store, putting security data directly into the index. This allows Palantir to determine at query time whether a user is allowed to see certain data, thus improving query performance by eliminating a post-processing step.

Finally, the Palantir Search Server can better control scalability. You can scale your Search Server implementation horizontally by adding Search Server nodes. The Palantir architecture supports mirroring to address user scalability. To address data scalability, you can configure your installation to use clustered search partitioning.

Horizon Server Operation

Like the Search Server, the Horizon Server indexes the Revisioning Database and provides search capabilities. The Horizon Server provides greatly improved response times for certain kinds of searches. The Horizon Server can also benefit from clustered search partitioning in large-scale deployments.

Palantir Enterprise Manager

The Palantir Enterprise Manager is a collection of applications that install Palantir servers, monitor servers, administer authentication and database sources, set alerts and notifications, synchronize with other Palantir nexuses, and help diagnose problems. The Palantir Enterprise Manager works with the Palantir Monitoring Server to gather details about managed Palantir servers. Together, these services provide both a general system status and more detailed monitoring data.

Using the Palantir Enterprise Manager, you can:

- install, update, or uninstall servers
- start, stop, and restart servers
- update and distribute server configurations
- run diagnostics and collect log data
- customize and manage your dynamic ontology

The Palantir Enterprise Manager uses the Software Testing Automation Framework (STAF), which is a peer-to-peer framework.

Palantir Enterprise Manager Architecture depicts the architecture of the Palantir Enterprise Manager. A single machine on the network acts as the central Palantir Enterprise Manager. The Palantir Enterprise Manager uses daemon processes, called STAFProc Agents (or STAF agents), that run on individual machines on the network. The Palantir Enterprise Manager itself has a STAF agent that is a peer to the remote agents.

Each STAF agent communicates with the Palantir Enterprise Manager about the managed installation on the machine where the agent is installed. For example, if you plan to install the Palantir Dispatch Server on machine `eiger`, you must first install a STAF agent on that machine. If you try to install the Dispatch Server without first installing the agent, the server installation fails.

If you use the Palantir Enterprise Manager to manage the Palantir Platform, then each machine running a Palantir server also runs a light-weight STAF agent. The Palantir Enterprise Manager interacts with its local STAF agent to send commands to the remote agents. The remote STAF agents execute the server's commands, such as stopping a Palantir server or updating a Palantir server.

Palantir Enterprise Manager Architecture
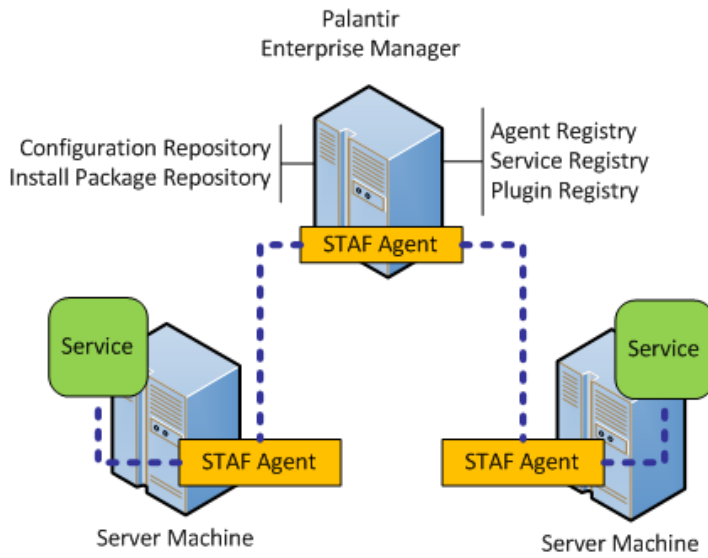
The Palantir Enterprise Manager maintains information locally about all the deployments it knows about. Through the Palantir Enterprise Manager, you can query the status of Palantir servers, obtain logging information, and push software updates.

Q Palantir
Technologies

Visit the Palantir Support Page