HDFS Multi-Tenant Controls

So far, we assumed that all data stored in HDFS can be accessed by all clients. Of course, in any production environment, you will have multiple users with different levels of privileges, all requiring data access. This is called running Hadoop in a multi-tenant environment, and HDFS is well equipped to deal with this situation. It allows data to be protected by security controls, and it allows per-user quotas, so that one user does not eat up all of the storage capacity; more specifically:

For security

- HDFS supports the notion of users and groups of users. Please note that these are different accounts from what you may have provisioned as Linux or LDAP users on the servers running HDFS services

- HDFS offers classic POSIX filesystem permissions for controlling who can read and write (e.g., `-rwxr-xr--`)

- HDFS also offers extended Access Control Lists (ACL) for richer scenarios

- Outside core HDFS, the Apache Ranger HDFS plugin offers centralized authorization policies and audit

For quotas

- HDFS provides easy-to-understand data size quotas

- HDFS also gives you additional options for controlling the number of files.

Learn About Verified Certificates