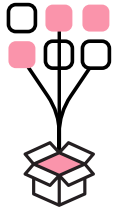


GLOSSARY FOR TALKING ABOUT HADOOP SECURITY



Yarn (n)

[jɑ:n] / noun

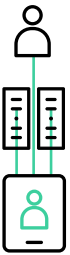
1. a resource manager for execution; assigns locations for memory and computation.



Impersonation (n)

[ɪmpərsə'neiʃən] / noun

1. doing an action on a cluster on behalf of another user.



Kerberos (n)

['kerberɒs] / noun

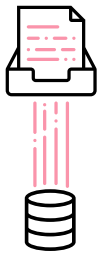
1. the main secure authorization system between users and services used by Hadoop.



Apache Knox (n)

[ə'pætʃi nɒks] / noun

1. secures API calls; used to secure web services built on Hadoop.

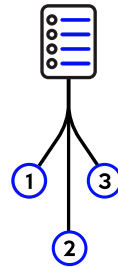


HDFS (n)

Hadoop distributed file system

[eɪf-di-ɛf-ɛs] / noun

1. a distributed file system that stores data on standard servers connected so that they form a cluster.

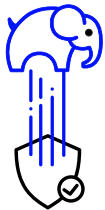


ACLs (n)

Access control lists Protocol

[eɪ-si:-ɛlz] / noun

1. lists of rules that define for each user or group who has access to file and which right each and every entity is allowed to do (i.e., read/write/execute).



Sentry (n)

['sɛntri] / noun

1. centralized authorization database management system; alternative to Ranger.



Ranger (n)

['reɪndʒər] / noun

1. centralized authorization database management system; alternative to Sentry.

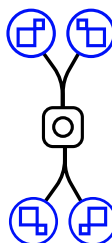


LDAP (n)

Lightweight Directory Access Protocol

[ɛl-di-ɛɪ-pi] / noun

1. a large database that contains the identity of all user groups and services in an organization or entity; the LDAP authenticates all clients.



Metastore (n)

['mitəstɔːr] / noun

1. a location that contains the metadata -- the central structure -- of each and every file in the system.



Hive, Pig, Impala, Tez, Spark, MapReduce (n)

[haɪv - piɡ - ɪm'palə - Tez spark - məp rɪ'dʒuːs] / noun

1. principal execution engines; a way of interpreting a language or instructions in order to process the calculation, e.g. translating instructions in one language into another.