

Enterprise Steam

Secure, Self-Service Artificial Intelligence Environments with Comprehensive IT Control

Enterprise Steam is a service to securely start and connect to H2O YARN jobs in a Hadoop environment. Encouraging adoption among Data Science users is a key goal, so ease of use is paramount. Key user personas include Data Scientists, Hadoop Admins, Enterprise Architects, and IT Security Specialists. Enterprise Steam offers **security**, **resource control** and **resource monitoring** out-of-the-box in a **multi-tenant architecture** so that organizations can focus on the core of their data science practice.

Enterprise Steam equips the stakeholders in AI practices with the capabilities required to perform their tasks without interfering with each other. Simply put, Enterprise Steam enables streamlined H2O adoption in a secure manner that complies with company policy. Administrators can easily control, monitor and measure H2O usage. This further enables use cases such as internal chargeback, internal cloud deployment, and H2O Platform as a Service (PaaS).

Enterprise Steam provides the following benefits to Artificial Intelligence (AI) practitioners:

Data Scientists

- Self-Service
 - Enterprise Steam provides easy R/Python APIs and a Web UI for starting H2O YARN jobs.
 - Without having to become Hadoop experts, Data Scientists can manage H2O clusters and connect to them using a stable service at a known IP address and port.
- Familiar Interface
 - Data Scientists can work in the comfort of familiar environments such as RStudio and Jupyter notebooks, without ever needing a terminal prompt.

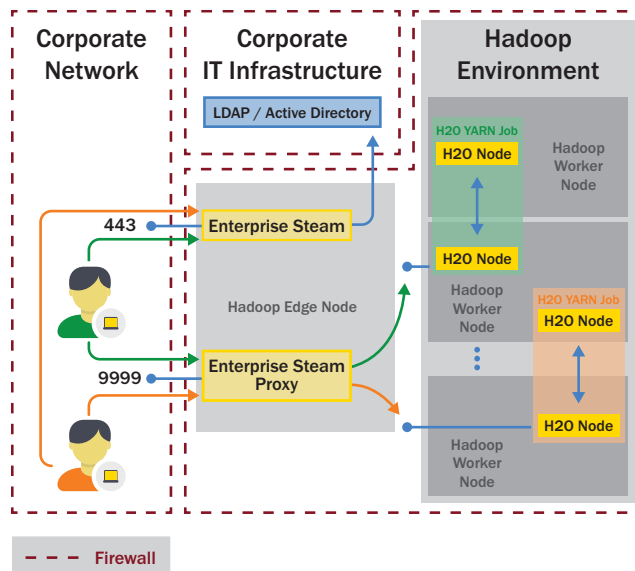


Figure 1: Enterprise Steam Deployment Diagram

- No SSH to Hadoop edge node required.
- Data Scientists can work directly from their laptops on the insecure side of a firewall.

Hadoop Admins, Enterprise Architects, IT Security Specialists

- Security
 - Enterprise Steam enforces full control of H2O YARN job security for administrators automatically without having to rely on the Data Scientist.
 - Enterprise Steam provides role-based Access Control for Admin Users and Standard Users.
 - Encrypted connections (SSL/TLS).
 - LDAP and Active Directory login authentication.
 - Kerberos authenticated YARN job submission.
 - Enterprise Steam offers the ability to put Hadoop clusters behind a firewall.
- Multi-Tenancy
 - Enterprise Steam is multi-tenant and prevents