



Cask Data Application Platform Administration

The **Cask Data Application Platform Administration** course will cover concepts and background that is required to run and manage CDAP and CDAP applications in a production environment. At the end of this course participants will be able to install, configure and operationalize CDAP.

Audience & Prerequisites

This course is designed for administrators who are looking to manage CDAP. Basic knowledge of scripting, HDFS and YARN is required.

Materials Required

- Laptop with Mac, Windows or Linux on 64-bit OS.
- Virtual Box installed (<https://www.virtualbox.org/wiki/Downloads>)
- CDAP Standalone VM installed (Instructions provided separately)

Course Overview

- CDAP Concepts & Capabilities
- Architecture deep-dive
- Installing and configuring CDAP
- CDAP Application overview
- Operational aspects of a CDAP application
- Health Checks and Monitoring
- Security in CDAP – Authentication, Keystore
- Security in CDAP – Authorization and Impersonation



Course Outline

<p>1. The Motivation to use CDAP</p> <ul style="list-style-type: none">• Context• The Hadoop Ecosystem Today• Challenges with Hadoop• The need for CDAP• CDAP benefits for developers <p>2. CDAP Overview</p> <ul style="list-style-type: none">• Functional view of CDAP• CDAP architecture• Building blocks <p>3. Understanding CDAP architecture - a deep dive</p> <ul style="list-style-type: none">• CDAP system components• Understanding data flow <p>4. CDAP Deployment architecture</p> <ul style="list-style-type: none">• CDAP components and requirements• Simplified deployment architecture• High-availability deployment <p>5. Installing CDAP</p> <ul style="list-style-type: none">• Preparing a cluster for CDAP• Installation using rpm/deb packages• Installation via Cloudera manager• Verification <p>6. Configuring CDAP</p> <ul style="list-style-type: none">• Understanding CDAP configuration• The run-down on important config settings <p>7. CDAP application building blocks</p> <ul style="list-style-type: none">• Overview of CDAP application components <p>8. Lifecycle management</p> <ul style="list-style-type: none">• Application lifecycle management• Program lifecycle management	<p>9. Health checks and monitoring</p> <ul style="list-style-type: none">• CDAP system service health checks• Monitoring CDAP components and applications using metrics <p>10. Transaction service operations</p> <ul style="list-style-type: none">• Impact of invalid transactions• Pruning invalid lists <p>11. Scaling programs</p> <ul style="list-style-type: none">• Scaling instances in services• Scaling Flowlet instances <p>12. Guaranteeing compute resources</p> <ul style="list-style-type: none">• Introduction to capacity scheduler• Managing CDAP application resources <p>13. Security - Authentication</p> <ul style="list-style-type: none">• Introduction to CDAP security• Enabling perimeter security• Custom authentication <p>14. Security - Keystore</p> <ul style="list-style-type: none">• Encrypting configuration in secure keystore• Using keystore in programs• Keystore and Runtime arguments• Demo <p>15. Security - Authorization</p> <ul style="list-style-type: none">• Security extensions for authorization• Enabling authorization with Sentry• Demo <p>16. Security – Cross Namespace access and Impersonation</p> <ul style="list-style-type: none">• Enabling accessing datasets from different namespace• Secure impersonation• Demo
--	---