DAS: Core Concepts

Date of Publish: 2019-03-20

Contents

Data Analytics Studio overview

Data Analytics Studio (DAS) is an application that provides diagnostic tools and intelligent recommendations to make the business analysts self-sufficient and productive with Hive.

DAS helps you to perform operations on Hive tables and provides recommendations for optimizing the performance of your queries. With DAS, you can search for queries, compose and execute queries, manage databases to create and edit tables, view reports and recommendations to optimize the query performance.

About DAS and DAS-Lite

DAS has two variants: DAS and DAS-Lite. DAS-Lite is a subset of the full-featured version of DAS. The following features that are available in DAS are not available in DAS-Lite:

*   Query optimization recommendation
*   Query comparison
*   Table optimization recommendation
*   Read and write report
*   Join report

Starting 1.2.0 onward, you have an option to use DAS/DAS-Lite with or without the Hortonworks DataPlane Platform. DataPlane (DP) serves as a management layer for your data clusters. For older releases of DAS, you need a DP deployment to use DAS.
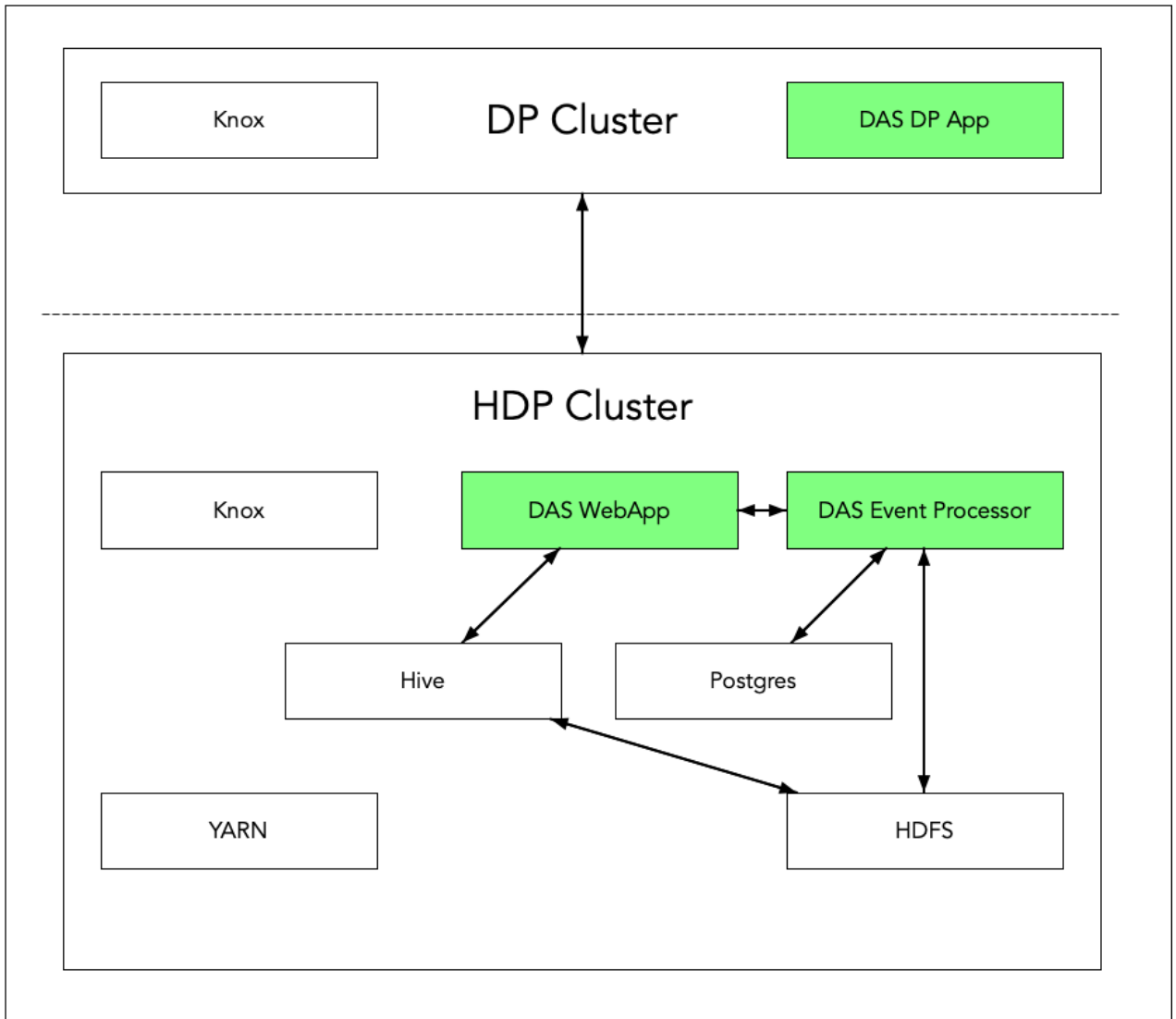
> **Note:** DAS and DAS-Lite have the same release versions. However, the binaries for DAS and DAS-Lite are different for a given version. You must download the appropriate binary (DAS or DAS-LIte) depending on your purchase from Hortonworks before starting the installation.

DAS architecture

DAS cluster agent comprises of a DAS Webapp and an Event Processor. These components are installed on the Ambari cluster to use with DAS. Additionally, you can install the DAS-DP App within the DP environment, and it communicates with the DAS components within the cluster.

Following is a block diagram showing the various components in a typical DAS deployment:

DP Cluster

| Knox | | DAS DP App |

HDP Cluster

| Knox | DAS WebApp | DAS Event Processor |

| Hive | Postgres |

| YARN | HDFS |

The DAS Webapp serves the DAS UI, and it should be on a machine that is accessible to your end users, or you should have a proxy setup to access the DAS UI.

Understanding and obtaining DAS binaries

You must download the DAS binaries from the Hortonworks Customer Portal by following the instructions provided as part of the subscription fulfillment process.

Make sure that you download the DAS/DAS-Lite binaries according to the HDP version and operating system as given in the following table. If you choose to use DP, then you'll need to download an additional tarball as outlined in a separate table.

Table 1: Obtaining binaries: DAS deployed in a standalone cluster

Versions and Operating Systems	CentOS 7	Ubuntu	Debian
HDP 2.6.5	• HDP 2 MPack for CentOS • HDP 2 DAS tarball for CentOS	• HDP 2 MPack for Ubuntu • HDP 2 DAS tarball for Ubuntu	• HDP 2 MPack for Debian • HDP 2 DAS tarball for Debian
HDP 3.x	• HDP 3 MPack for CentOS • HDP 3 DAS/DAS Lite tarball for CentOS	• HDP 3 MPack for Ubuntu • HDP 3 DAS/DAS Lite tarball for Ubuntu	• HDP 3 MPack for Debian • HDP 3 DAS/DAS Lite tarball for Debian

DAS is provided as RPMs in tarball repositories and as an Ambari Management Pack (MPack).

The DAS cluster components are installed in to your cluster using the Ambari MPack. The DAS cluster components are available for RHEL/CentOS/OEL 7, Debian 9 (HDP 3.x), and Ubuntu 16 platforms.

Optioanlly, you can install the DAS-DP App in to the DataPlane environment and is available for RHEL/CentOS/OEL 7 and is for use on the DP Platform 1.2.0 or higher.

Table 2: Obtaining binaries: DAS managed through DP

Versions and Operating Systems	CentOS 7	Ubuntu	Debian
HDP 2.6.5	CentOS 7 DAS tarball.	CentOS 7 DAS/DAS-Lite tarball. It is mandatory to download this file if you want to install and use the DAS-DP App.	CentOS 7 DAS/DAS-Lite tarball. It is mandatory to download this file if you want to install and use the DAS-DP App.
HDP 3.x	CentOS 7 DAS tarball.	CentOS 7 DAS/DAS-Lite tarball. It is mandatory to download this file if you want to install and use the DAS-DP App.	CentOS 7 DAS/DAS-Lite tarball. It is mandatory to download this file if you want to install and use the DAS-DP App.

Table 3: Supported platforms - DAS cluster components

DAS/DAS-Lite	CentOS 7	Debian 9	Ubuntu 16
DAS Cluster Components (DAS Event Processor and DAS Webapp)	Yes	Yes (HDP 3.x)	Yes

Table 4: Supported platform - DAS-DP application

DAS/DAS-Lite	CentOS 7
DAS-DP App	Yes

Installation roadmap

The DAS cluster components comprise of the DAS Webapp and a DAS Event Processor. These components are installed on the Ambari cluster to use with DAS. Optionally, you can install the DAS-DP App within the DataPlane environment. The DAS-DP App communicates with the DAS components within the cluster.

To install DAS/DAS-Lite, review the installation roadmap and follow the steps. We strongly encourage you to read the support requirements and the pre-installation tasks before you start the installation.
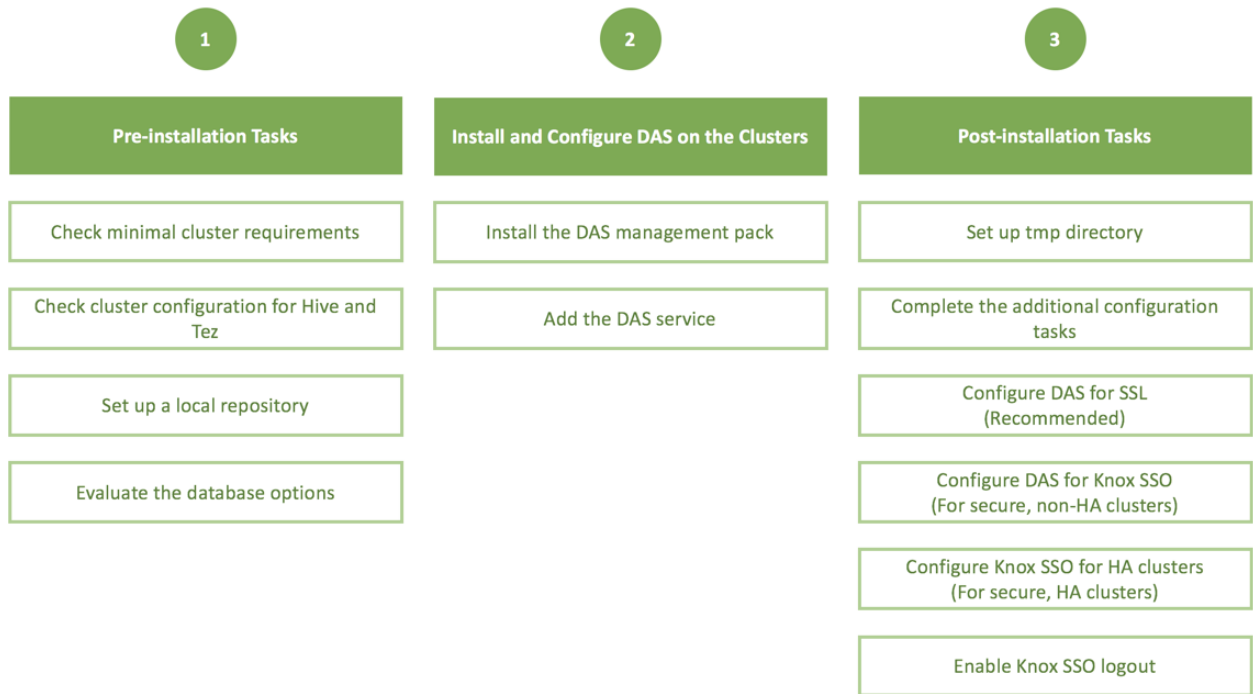
1	2	3
Pre-installation Tasks	**Install and Configure DAS on the Clusters**	**Post-installation Tasks**
Check minimal cluster requirements	Install the DAS management pack	Set up tmp directory
Check cluster configuration for Hive and Tez	Add the DAS service	Complete the additional configuration tasks
Set up a local repository		Configure DAS for SSL (Recommended)
Evaluate the database options		Configure DAS for Knox SSO (For secure, non-HA clusters)
		Configure Knox SSO for HA clusters (For secure, HA clusters)
		Enable Knox SSO logout

Table 5: DAS/DAS-Lite Installation Roadmap

No.	Task	Description
Pre-installation tasks		
1	Check the minimal cluster requirements.	To install and set up the necessary components and services, see Install the prerequisite components and services.
2	Check the cluster configuration for Hive and Tez.	Check the configuration settings for Hive and Tez in Ambari. • For HDP 2.6.5, see Check cluster configuration for Hive and Tez (HDP 2.6.5) • For HDP 3.x or higher, see Check cluster configuration for Hive and Tez
3	Create and set up a local repository.	Hortonworks does not host any public repository for DAS. Therefore, you need to setup a local repository to install the binaries. See Create and set up a local repository.
4	Evaluate the database options.	DAS requires a PostgreSQL database for storing query event information. See Evaluate your database options.
Installing and configuring DAS on the HDP clusters		
5	Install the DAS management pack.	Install DAS cluster components (the DAS Webapp and the DAS Event Processor) on the same machine on which you have installed the Ambari server, using an Ambari management pack (MPack). See Installing the DAS cluster components..
6	Add the DAS service.	After installing the DAS MPack through the Ambari console, add the DAS service through the Ambari UI. See Adding the DAS Service through the Ambari UI.
Post-installation tasks		
7	Set up tmp directory.	You need write permission on the /tmp directory to write logs for secure clusters. See Setting up the tmp directory.
8	Complete the additional configuration tasks.	See Additional configuration tasks.
9	Configure DAS for SSL.	It is recommended that you configure DAS for SSL. You can configure SSL if your HDP cluster is SSL-enabled. See (Recommended) Configuring DAS for SSL/TLS.
10	Configure DAS for Knox SSO.	To enable DAS to work with the HDP cluster SSO, configure the Knox settings as described in Configuring DAS for Knox SSO.
11	Configure Knox SSO for HA clusters.	To enable Knox SSO for HA clusters, see Configuring Knox SSO for HA clusters.
12	Enable Knox SSO logout.	See Enabling logout option for secure clusters.