



Deploying DataStax Enterprise with Mesosphere Datacenter Operating System

August 2016



Table of Contents

TABLE OF CONTENTS	2
OVERVIEW	3
INTRODUCTION	3
About Mesosphere Enterprise DC/OS	3
About DataStax	3
About DataStax Enterprise	3
DATASTAX ENTERPRISE WITH MESOSPHERE DC/OS	4
DataStax Enterprise	4
Mesosphere Enterprise DC/OS	4
Any Infrastructure	4
Mesosphere Enterprise DC/OS and DataStax Enterprise: Easily deploy and scale a first-class enterprise distributed database	5
Easy Data Infrastructure with Mesosphere Enterprise DC/OS	5
DEPLOYING DATASTAX ENTERPRISE THROUGH MESOSPHERE DC/OS UI	6
Simple Installation	6
Prerequisites	6
Deployment Steps	6
Advanced Installation	8
Check the Status of the Deployment	10
Deployment Deep Dive	11
Service Configurations	11
Node Configurations	12
Troubleshooting Logs	13
DEPLOYING SINGLE DC DSE THROUGH MESOSPHERE DC/OS CLI	14
Prerequisites	14
Deployment Steps	14
CQLSH Connection	15
ABOUT DATASTAX	15
ABOUT MESOSPHERE	15

Overview

Apache Cassandra™ is a distributed datastore that has become the industry standard database for massive-scale data pipelines. DataStax Enterprise builds on Cassandra with additional features and capabilities that companies like Netflix, Intuit, ING and eBay are using to power their mission critical, data-intensive applications.

However, provisioning of the underlying infrastructure can be difficult, as it is not automated with modern orchestration nor highly elastic. Mesosphere Enterprise DC/OS makes it easy to deploy and scale DataStax Enterprise, while running it on the same shared infrastructure as your containers and other big data services such as Kafka and Spark.

Introduction

About Mesosphere Enterprise DC/OS

Mesosphere Enterprise DC/OS makes it easy to build and run modern distributed applications in production at scale, by pooling resources across an entire datacenter or cloud. With Enterprise DC/OS you can confidently run containers and microservices at scale with a rock-solid platform powering today's production hyperscale datacenters, and easily install and manage big data services like Kafka, Spark and Cassandra that power many of today's Internet of Things and Big Data stacks.

About DataStax

DataStax, the leading provider of database software for cloud applications, accelerates the ability of enterprises, government agencies, and systems integrators to power the exploding number of cloud applications that require data distribution across datacenters and clouds, by using our secure, operationally simple platform built on Apache Cassandra™.

About DataStax Enterprise

DataStax Enterprise (DSE) accelerates the ability to deliver real-time value at epic scale by providing a comprehensive and operationally simple data management layer with a unique always-on architecture built on Apache Cassandra™. DataStax Enterprise provides the distributed, responsive and intelligent foundation to build and run cloud applications.

DataStax Enterprise with Mesosphere DC/OS

DataStax Enterprise

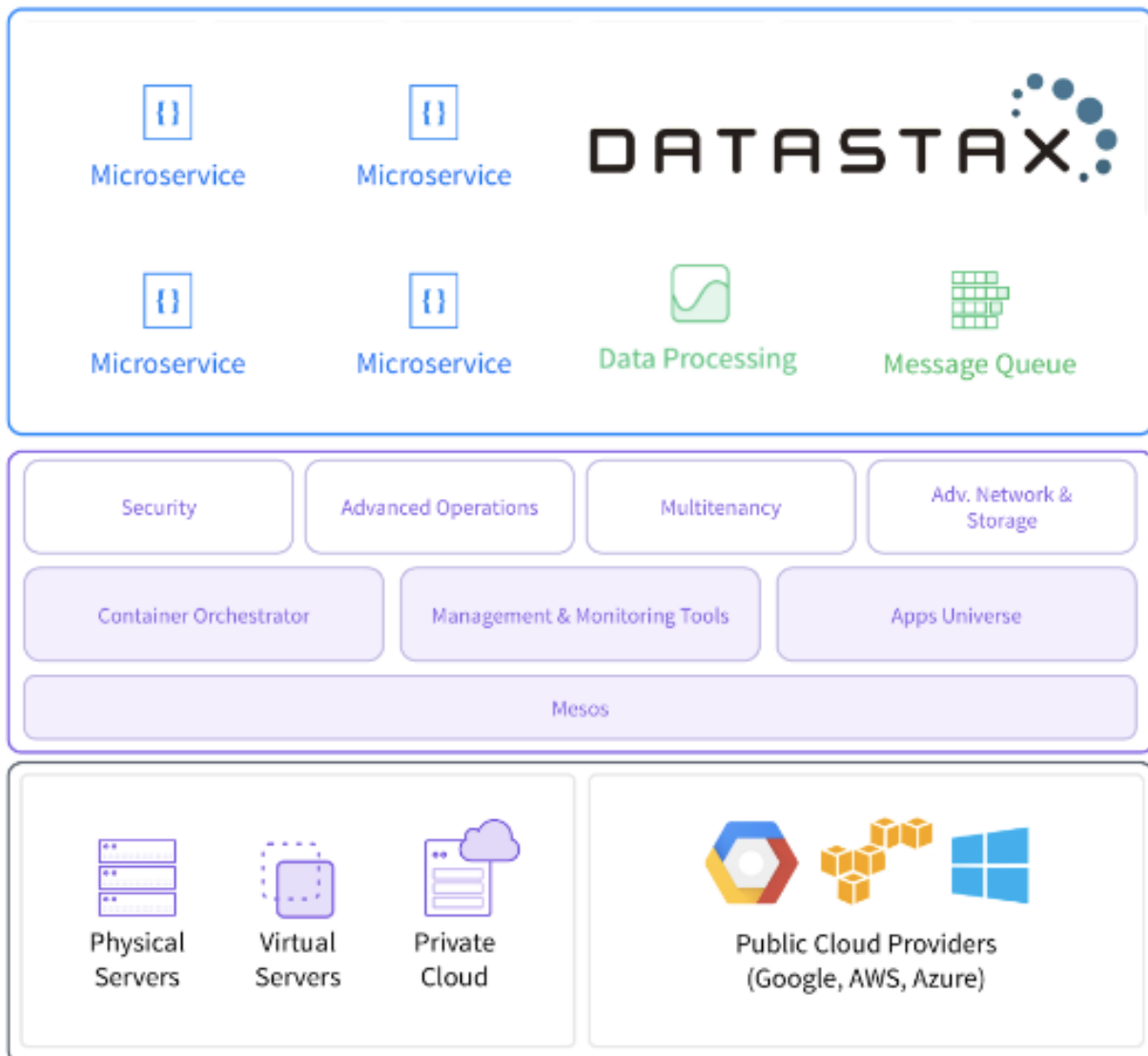
DataStax Enterprise is the only operational database, which can scale linearly in excess of 1,000 nodes, with no single point of failure, and is capable of providing real-time active-everywhere replication across many datacenters and cloud providers.

Mesosphere Enterprise DC/OS

The only production proven and fully supported platform that elastically runs containerized apps and data services at scale. Effortlessly power IoT, predictive analytics, and personalization applications with an elastic, resilient data infrastructure.

Any Infrastructure

Install Enterprise DC/OS on any public cloud or in your own private datacenter—whether virtualized or on bare metal. Create a consistent user experience and move your workloads with ease.



Mesosphere Enterprise DC/OS and DataStax Enterprise: Easily deploy and scale a first-class enterprise distributed database

Enterprise DC/OS modern orchestration automates app & cluster provisioning, management, & elastic scaling. Without Enterprise DC/OS, provisioning of infrastructure is manual.

Datastax Enterprise on Enterprise DC/OS allows you to easily deploy a first-class enterprise distributed database

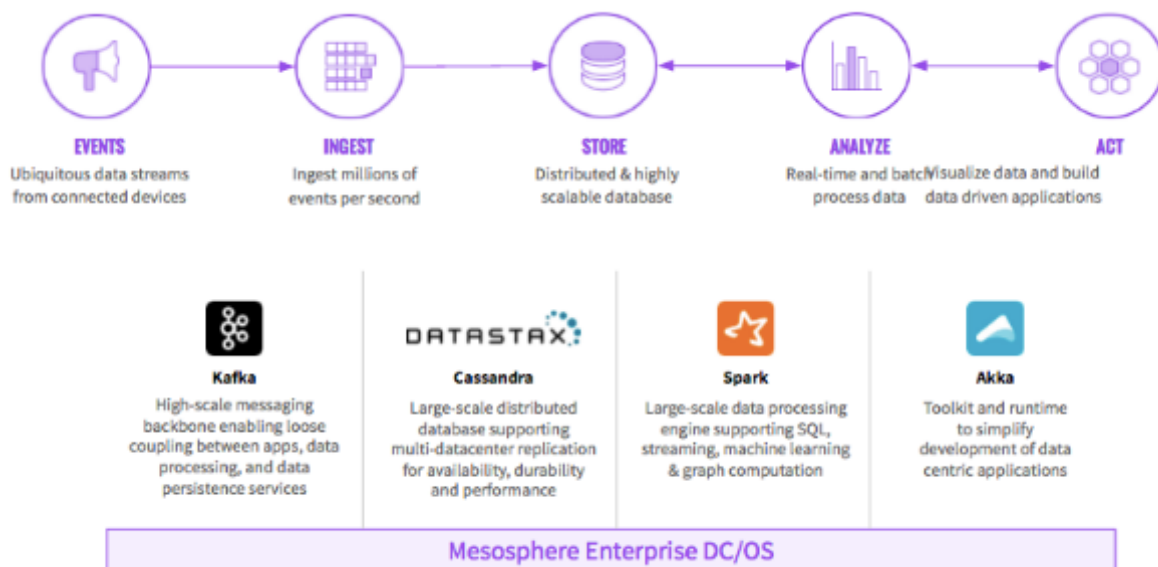
- Flexible, highly-configurable solution for enterprises
- Future releases are planned to add support for additional DSE features, DSE Max, and DSE Graph

Benefits include:

- Easy to configure, deploy, scale, monitor & manage DSE clusters
- Efficiently leverage shared infrastructure

Easy Data Infrastructure with Mesosphere Enterprise DC/OS

With Mesosphere Enterprise DC/OS you can install data and analytics services with a single click, and elastically scale and operate them with ease, in the cloud or your own datacenter. Services are loosely coupled and composable, to evolve with you and enable you to adapt to new services as they emerge. Because services run on the Enterprise DC/OS platform, everything shares the same set of resources and can be monitored using the same dashboard.



Deploying DataStax Enterprise Through Mesosphere DC/OS UI

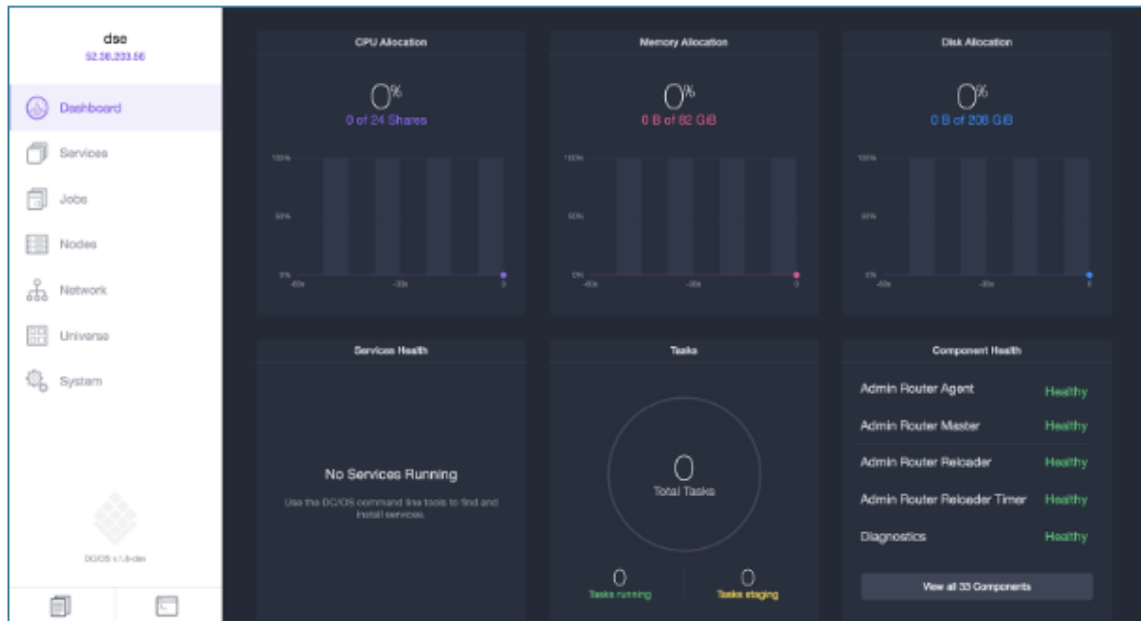
Simple Installation

Prerequisites

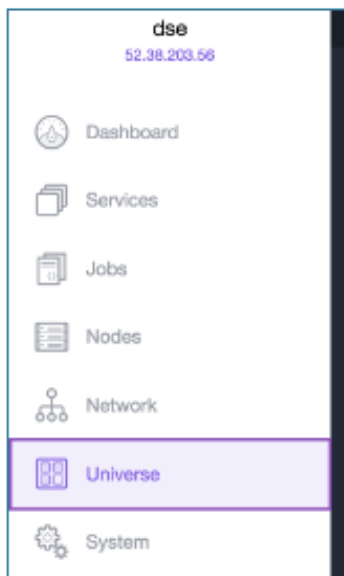
- DC/OS 1.8 Environment

Deployment Steps

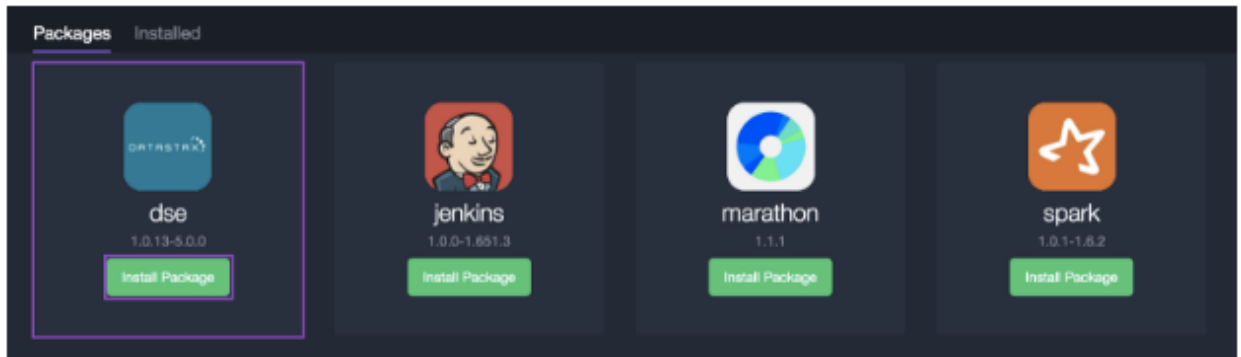
1. Launch DC/OS GUI



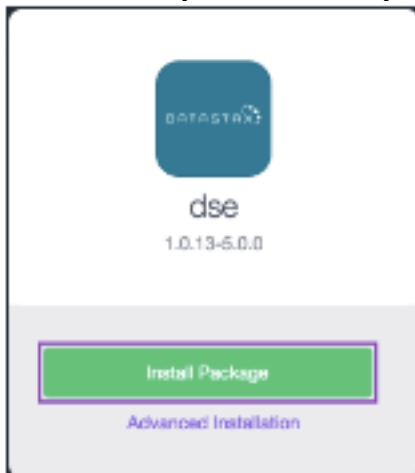
2. Click on the Universe Tab.



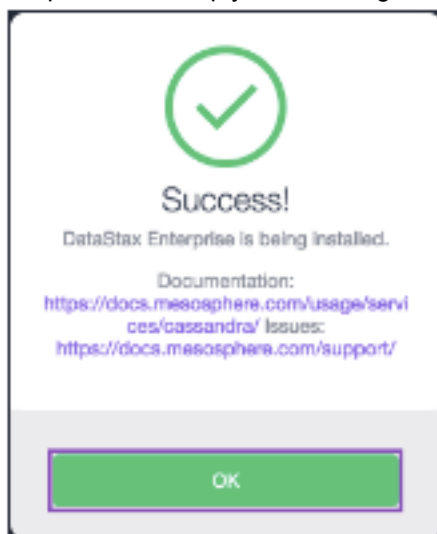
- In the packages window pane on the right, locate the **dse** package and click the **Install Package** button. You can search for the package as well.



- Click on **Install Package**. To configure installation parameters such as number of nodes, cpu cores, ram, dse yaml, cassandra yaml settings, please click on [Advanced Installation](#).

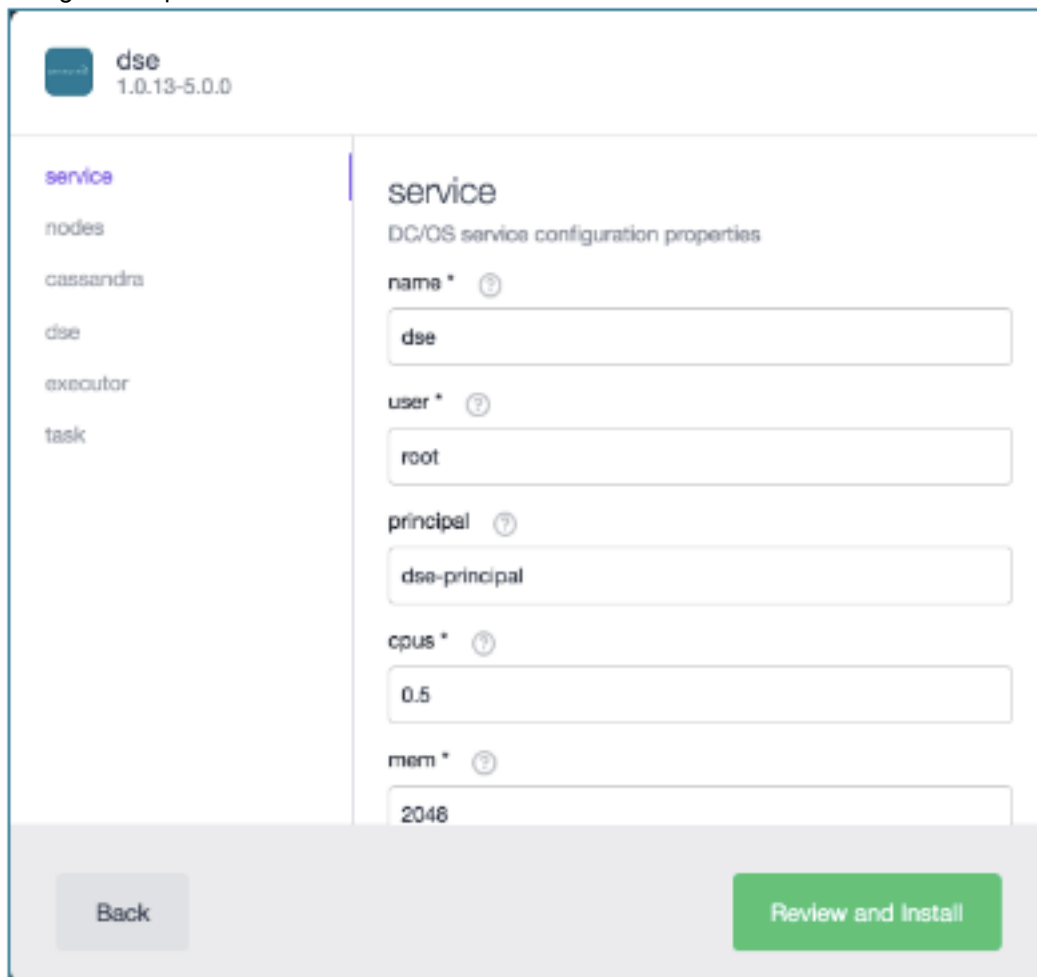


- Upon clicking install, you will receive a success. This does not indicate that the deployment is complete, this simply is indicating that the package installation service has successfully started.



Advanced Installation

1. Configure the parameters and click on **Review and Install**.



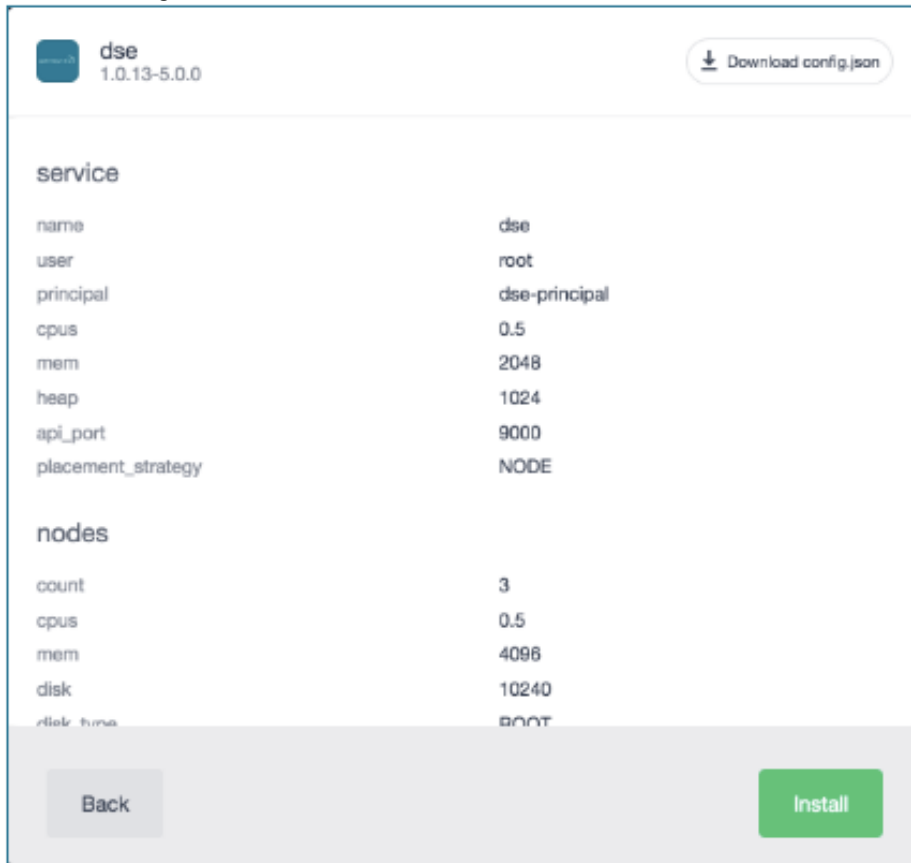
The screenshot shows the DSE configuration interface for version 1.0.13-5.0.0. The 'service' tab is selected in the left sidebar. The main panel displays 'DC/OS service configuration properties' with the following fields:

- name ***: dse
- user ***: root
- principal**: dse-principal
- cpus ***: 0.5
- mem ***: 2048

At the bottom, there are two buttons: 'Back' and 'Review and Install'.

Configuration Tab	Description
Services	DC/OS service parameters. Defaults recommended.
Nodes	Configure number of nodes to deploy as well as per node configurations.
Cassandra	Cassandra Yaml settings. For further information on the configuration settings you can search the setting here .
DSE	DSE.yaml settings. For further information on the configuration settings you can search the setting here .
Executor	DC/OS service parameters. Defaults recommended.
Task	DC/OS service parameters. Defaults recommended.

2. Review Configurations and Click **Install**.



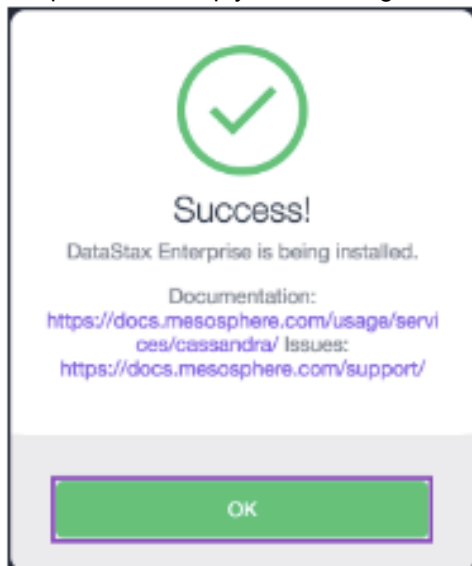
The screenshot shows the configuration page for DataStax Enterprise (DSE) version 1.0.13-5.0.0. At the top right, there is a button labeled "Download config.json". Below this, the configuration is divided into two sections: "service" and "nodes".

service	
name	dse
user	root
principal	dse-principal
cpus	0.5
mem	2048
heap	1024
api_port	9000
placement_strategy	NODE

nodes	
count	3
cpus	0.5
mem	4096
disk	10240
disk_type	ROOT

At the bottom of the page, there are two buttons: "Back" on the left and "Install" on the right.

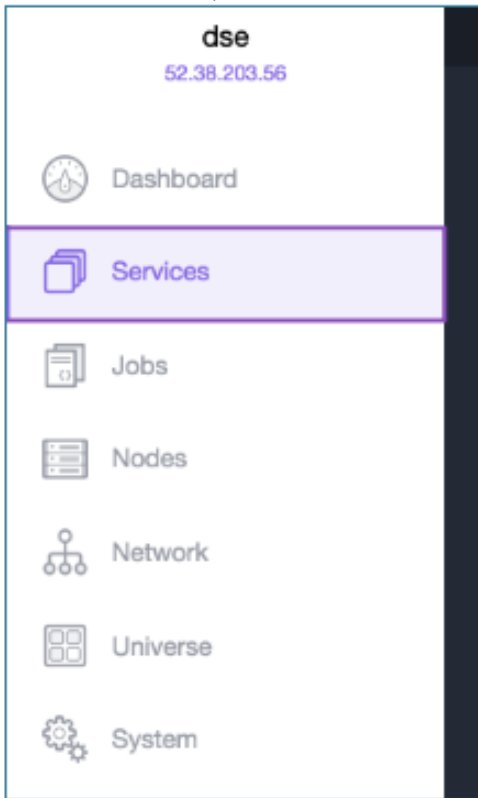
3. Upon clicking install, you will receive a success. This does not indicate that the deployment is complete, this simply is indicating that the package installation service has successfully started.



The screenshot shows a success dialog box with a green checkmark icon at the top. Below the icon, the text reads "Success!". Underneath, it says "DataStax Enterprise is being installed." followed by "Documentation:" and two links: <https://docs.mesosphere.com/usaga/services/cassandra/> and "Issues:" with a link <https://docs.mesosphere.com/support/>. At the bottom of the dialog, there is a green button labeled "OK".



Check the Status of the Deployment

1. In the DC/OS GUI, click on the **Services** Tab on the left panel.





2. Check the Status of the **dse** service. The status should denote “Deploying” if the package deployment is still running and “Running” if completed.

Deploying:

NAME ^	STATUS ?	CPU	MEM	DISK
 dse	 Deploying (1/1)	0.5	2 GiB	0 B

Completed:

NAME ^	STATUS ?	CPU	MEM	DISK
 dse	 Running (1/1)	0.5	2 GiB	0 B

Deployment Deep Dive

The following section outlines steps for checking the health of the nodes as well as checking individual node configurations, files, and logs.

Service Configurations

1. In the **Services** section, click on the **dse** deployment service.

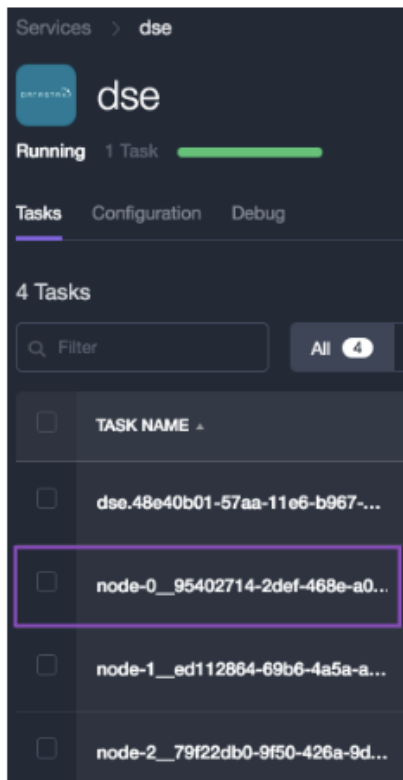
The screenshot shows the AWS Management Console interface for the 'dse' service. At the top, there are buttons for 'Scale', 'Edit', and 'More'. Below this, a progress bar indicates 'Running 1 Task'. The 'Tasks' tab is active, showing a list of 4 tasks. A filter bar at the top of the task list shows 'All 4', 'Active 4', and 'Completed 0'. The task list table has columns for 'TASK NAME', 'HOST', 'CPU', 'MEM', 'UPDATED', and 'VERSION'. Each task row shows a 'Running' status, which is highlighted by a purple box in the image.

TASK NAME	HOST	CPU	MEM	UPDATED	VERSION
dse.48e40b01-57aa-11e6-b967~...	10.0.3.96:9000	0.5	2 GiB	7 minutes ago	7/31/2016, 10:39:15 PM
node-0_95402714-2def-468e-a0...	N/A	0.5	4 GiB	6 minutes ago	
node-1_ed112864-69b6-4a5a-a...	N/A	0.5	4 GiB	4 minutes ago	
node-2_79f22db0-9f50-426a-9d...	N/A	0.5	4 GiB	3 minutes ago	

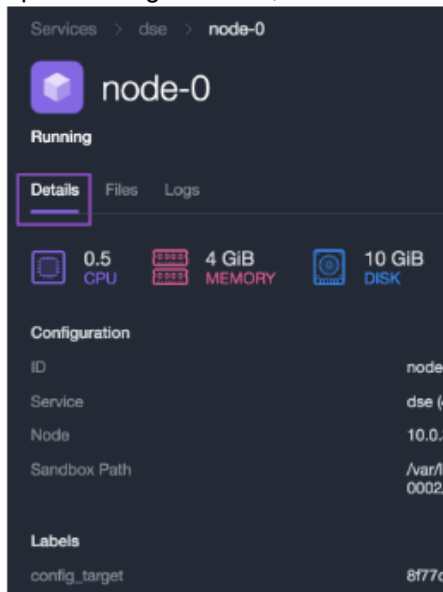
Tabs	Description
Tasks	List of all the nodes configured for service deployment and their status. Actions: <ol style="list-style-type: none"> 1. Suspend or Destroy the service using the More drop down menu. 2. Edit service configurations using Edit button. 3. Add more instances/nodes using Scale button.
Configuration	List of all service configuration settings
Debug	Health of dse service. Example: Last Changes, Last Task Failure, etc.

Node Configurations

1. In the **Services** section, click on the **dse** deployment service.
2. In the **dse** service window, click on the preferred node.



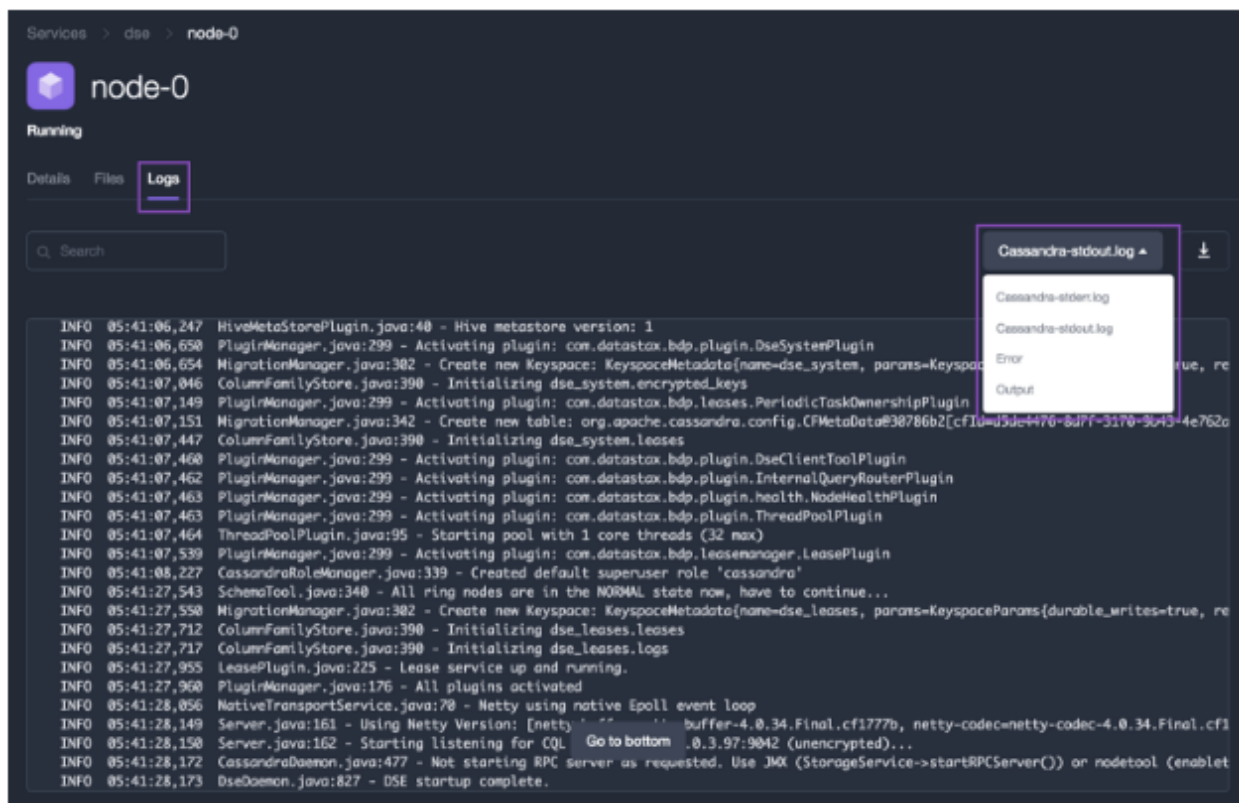
3. Upon clicking the node, the node view will appear as the screenshot below.



Configuration Tab	Description
Details	Node Hardware Configurations
Files	All DSE configuration files, Cassandra and system logs. Note: Files can be downloaded by clicking on them.
Logs	Cassandra-stdout, Cassandra-stderr, stdout,stderr log files output. For more details, click here .

Troubleshooting Logs

- Click on **Logs** Tab in the Node view.
 - On the right side, click on the drop down menu to select and view the preferred log file.



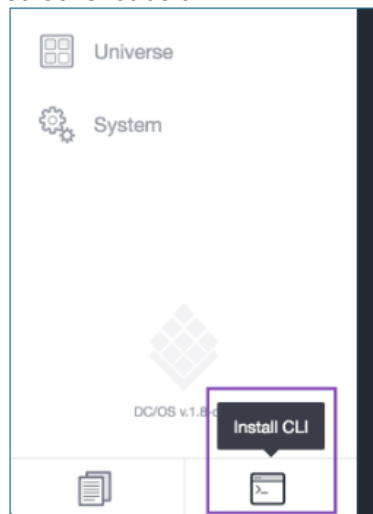
Deploying Single DC DSE Through Mesosphere DC/OS CLI

Prerequisites

- DC/OS 1.8 Environment

Deployment Steps

1. Browse to the DC/OS UI.
2. Install DC/OS CLI on local machine. Installation instructions can be acquired by following the instructions below.
 - Click on the “Install CLI” button on the bottom left corner of the UI. Reference the screenshot below.



3. Install **virtualenv** locally using the instructions found [here](#).
4. Install DataStax Enterprise using the following command. This command will deploy a single datacenter with 3 DSE nodes.

```
dcos package install dse
```

Note: For advanced installation options you can append an `--options` flag to the install command above. Example below. For more details please check the documentation [here](#).

```
dcos package install --options=sampleconfig.json dse  
sampleconfig.json:
```

```
{  
  "nodes": {  
    "count": 10,  
    "seeds": 3  
  }  
}
```

5. Verify Installation by running the command below and checking the 'status'.
 - `curl http://<dcos-url>/service/dse/v1/plan`

CQLSH Connection

1. Run the command below to add ssh Identity.
 - `ssh-add ~/path-to-ssh-key/<your-ssh-key>.pem`
2. Grab the IP for one of the DSE nodes.
 - `dcos dse connection`
3. SSH into the DC/OS master node.
 - `dcos node ssh --master-proxy --leader`
4. CQLSH into one of the nodes.
 - `docker run -ti cassandra:3.0.7 cqlsh <dse-node-ip>`

About DataStax

DataStax, the leading provider of database software for cloud applications, accelerates the ability of enterprises, government agencies, and systems integrators to power the exploding number of cloud applications that require data distribution across datacenters and clouds, by using our secure, operationally simple platform built on Apache Cassandra™.

With more than 500 customers in over 50 countries, DataStax is the database technology of choice for the world's most innovative companies, such as Netflix, Safeway, ING, Adobe, Intuit, Target and eBay. Based in Santa Clara, Calif., DataStax is backed by industry-leading investors including Comcast Ventures, Crosslink Capital, Lightspeed Venture Partners, Kleiner Perkins Caufield & Byers, Meritech Capital, Premji Invest and Scale Venture Partners. For more information, visit DataStax.com or follow us [@DataStax](https://twitter.com/DataStax).

About Mesosphere

Mesosphere (Twitter: [@Mesosphere](https://twitter.com/Mesosphere)) is leading the enterprise transformation toward distributed computing and modern applications with its Datacenter Operating System, and is the principal founding member of the DC/OS open source project. DC/OS is a production-proven datacenter-scale platform for container operations and simple installation of complex distributed systems—including HDFS, Apache Spark, Apache Kafka, Apache Cassandra and more. Mesosphere was founded in 2013 by the architects of hyperscale infrastructures at Airbnb and Twitter, along with the co-creator of Apache Mesos. Backed by A Capital, Andreessen Horowitz, Data Collective, Fuel Capital, Hewlett Packard Enterprise, Khosla Ventures, Kleiner Perkins Caufield & Byers, Triangle Peak Partners and Microsoft, Mesosphere is headquartered in San Francisco with additional offices in New York and Hamburg, Germany.

In March 2016, Mesosphere announced that it closed \$73.5 million in a Series C funding round led by Hewlett Packard Enterprise, with Microsoft joining as a new strategic investor. The round brought Mesosphere's total funding to date to nearly \$126 million and supports the company's accelerating growth as DC/OS becomes the de facto standard for enterprises operating containers and distributed systems in production.