

Spring Cloud Data Flow Server for Cloud Foundry

1.0.0.BUILD-SNAPSHOT

Copyright © 2013-2016Pivotal Software, Inc.

Copies of this document may be made for your own use and for distribution to others, provided that you do not charge any fee for such copies and further provided that each copy contains this Copyright Notice, whether distributed in print or electronically.

Table of Contents

I. Spring Cloud Data Flow for Cloud Foundry	1
1. Spring Cloud Data Flow	2
2. Spring Cloud Stream	3
3. Spring Cloud Task	4
4. Deploying on Cloud Foundry	5
4.1. Provision a Redis service instance on Cloud Foundry.	5
4.2. Provision a Rabbit service instance on Cloud Foundry.	5
4.3. Download the Spring Cloud Data Flow Server and Shell apps:	5
4.4. Deploying the Server app on Cloud Foundry	5
4.5. Running the Server app locally	6
4.6. Running Spring Cloud Data Flow Shell locally	7
5. Configuration Reference	8
II. Appendices	9
6. Building	10
6.1. Basic Compile and Test	10
6.2. Documentation	10
6.3. Working with the code	10
Importing into eclipse with m2eclipse	10
Importing into eclipse without m2eclipse	11

Part I. Spring Cloud Data Flow for Cloud Foundry

This project provides support for deploying Spring Cloud Stream and Spring Cloud Task applications to Cloud Foundry.

1. Spring Cloud Data Flow

Spring Cloud Data Flow is a cloud-native programming and operating model for composable data microservices on a structured platform. With Spring Cloud Data Flow, developers can create and orchestrate data pipelines for common use cases such as data ingest, real-time analytics, and data import/export.

The Spring Cloud Data Flow architecture consists of a server that deploys [Streams](#) and [Tasks](#). Streams are defined using a [DSL](#) or visually through the browser based designer UI. Streams are based on the [Spring Cloud Stream](#) programming model while Tasks are based on the [Spring Cloud Task](#) programming model. The sections below describe more information about creating your own custom Streams and Tasks

For more details about the core architecture components and the supported features, please review Spring Cloud Data Flow's [core reference guide](#). There're several [samples](#) available for reference.

2. Spring Cloud Stream

Spring Cloud Stream is a framework for building message-driven microservice applications. Spring Cloud Stream builds upon Spring Boot to create standalone, production-grade Spring applications, and uses Spring Integration to provide connectivity to message brokers. It provides opinionated configuration of middleware from several vendors, introducing the concepts of persistent publish-subscribe semantics, consumer groups, and partitions.

For more details about the core framework components and the supported features, please review Spring Cloud Stream's [reference guide](#).

There's a rich ecosystem of Spring Cloud Stream [Application-Starters](#) that can be used either as standalone data microservice applications or in Spring Cloud Data Flow. For convenience, we have generated RabbitMQ and Apache Kafka variants of these application-starters that are available for use from [Maven Repo](#) and [Docker Hub](#) as maven artifacts and docker images, respectively.

Do you have a requirement to develop custom applications? No problem. Refer to this guide to create [custom stream applications](#). There're several [samples](#) available for reference.

3. Spring Cloud Task

Note

This feature is actively being developed and it is yet to be supported in Spring Cloud Data Flow's Cloud Foundry server.

Spring Cloud Task makes it easy to create short-lived microservices. We provide capabilities that allow short-lived JVM processes to be executed on demand in a production environment.

For more details about the core framework components and the supported features, please review Spring Cloud Task's [reference guide](#).

There's a rich ecosystem of Spring Cloud Task [Application-Starters](#) that can be used either as standalone data microservice applications or in Spring Cloud Data Flow. For convenience, the generated application-starters are available for use from [Maven Repo](#). There are several [samples](#) available for reference.

4. Deploying on Cloud Foundry

Spring Cloud Data Flow can be used to deploy modules in a Cloud Foundry environment. When doing so, the server application can either run itself on Cloud Foundry, or on another installation (e.g. a simple laptop).

The required configuration amounts to the same in either case, and is merely related to providing credentials to the Cloud Foundry instance so that the server can spawn applications itself. Any Spring Boot compatible configuration mechanism can be used (passing program arguments, editing configuration files before building the application, using [Spring Cloud Config](#), using environment variables, etc.), although some may prove more practicable than others when running *on* Cloud Foundry.

Note

By default, the [application registry](#) in Spring Cloud Data Flow's Cloud Foundry server is empty. It is intentionally designed to allow users to have the flexibility of [choosing and registering](#) applications, as they find appropriate for the given use-case requirement. Depending on the message-binder of choice, users can register between [RabbitMQ](#) or [Apache Kafka](#) based maven artifacts.

4.1 Provision a Redis service instance on Cloud Foundry.

Use `cf marketplace` to discover which plans are available to you, depending on the details of your Cloud Foundry setup. For example when using [Pivotal Web Services](#):

```
cf create-service rediscloud 30mb redis
```

4.2 Provision a Rabbit service instance on Cloud Foundry.

Use `cf marketplace` to discover which plans are available to you, depending on the details of your Cloud Foundry setup. For example when using [Pivotal Web Services](#):

```
cf create-service cloudamqp lemur rabbit
```

4.3 Download the Spring Cloud Data Flow Server and Shell apps:

```
wget http://repo.spring.io/snapshot/org/springframework/cloud/spring-cloud-dataflow-server-cloudfoundry/1.0.0.BUILD-SNAPSHOT/spring-cloud-dataflow-server-cloudfoundry-1.0.0.BUILD-SNAPSHOT.jar
wget http://repo.spring.io/milestone/org/springframework/cloud/spring-cloud-dataflow-shell/1.0.0.M3/spring-cloud-dataflow-shell-1.0.0.M3.jar
```

You can either deploy the server application on Cloud Foundry itself or on your local machine. The following two sections explain each way of running the server.

4.4 Deploying the Server app on Cloud Foundry

Push the server application on Cloud Foundry, configure it (see below) and start it.

Note

You must use a unique name for your app; an app with the same name in the same organization will cause your deployment to fail

```
cf push dataflow-server --no-start -p spring-cloud-dataflow-server-cloudfoundry-1.0.0.BUILD-SNAPSHOT.jar
cf bind-service dataflow-server redis
cf bind-service dataflow-server rabbit
```

Note

If you are pushing to a space with multiple users, for example on PWS, there may already be a route taken for the applicaiton name you have chosen. You can use the options `--random-route` to avoid this when pushing the app.

Now we can configure the app. The following configuration is for Pivotal Web Services. You need to fill in {org}, {space}, {email} and {password} before running these commands.

Note

Only set 'Skip SSL Validation' to true if you're running on a Cloud Foundry instance using self-signed certs (e.g. in development). Do not use for production.

```
cf set-env dataflow-server SPRING_CLOUD_DEPLOYER_CLOUDFOUNDRY_API_ENDPOINT https://api.run.pivotal.io
cf set-env dataflow-server SPRING_CLOUD_DEPLOYER_CLOUDFOUNDRY_ORGANIZATION {org}
cf set-env dataflow-server SPRING_CLOUD_DEPLOYER_CLOUDFOUNDRY_SPACE {space}
cf set-env dataflow-server SPRING_CLOUD_DEPLOYER_CLOUDFOUNDRY_DOMAIN cfapps.io
cf set-env dataflow-server SPRING_CLOUD_DEPLOYER_CLOUDFOUNDRY_SERVICES redis,rabbit
cf set-env dataflow-server SPRING_CLOUD_DEPLOYER_CLOUDFOUNDRY_USERNAME {email}
cf set-env dataflow-server SPRING_CLOUD_DEPLOYER_CLOUDFOUNDRY_PASSWORD {password}
cf set-env dataflow-server SPRING_CLOUD_DEPLOYER_CLOUDFOUNDRY_SKIP_SSL_VALIDATION false
```

We are now ready to start the app.

```
cf start dataflow-server
```

Alternatively, you can run the Admin application locally on your machine which is described in the next section.

4.5 Running the Server app locally

To run the server application locally, targeting your Cloud Foundry installation, you need to configure the application either by passing in command line arguments (see below) or setting a number of environment variables.

To use environment variables set the following:

```
export SPRING_CLOUD_DEPLOYER_CLOUDFOUNDRY_API_ENDPOINT=https://api.run.pivotal.io
export SPRING_CLOUD_DEPLOYER_CLOUDFOUNDRY_ORGANIZATION={org}
export SPRING_CLOUD_DEPLOYER_CLOUDFOUNDRY_SPACE={space}
export SPRING_CLOUD_DEPLOYER_CLOUDFOUNDRY_DOMAIN=cfapps.io
export SPRING_CLOUD_DEPLOYER_CLOUDFOUNDRY_SERVICES=redis,rabbit
export SPRING_CLOUD_DEPLOYER_CLOUDFOUNDRY_USERNAME={email}
export SPRING_CLOUD_DEPLOYER_CLOUDFOUNDRY_PASSWORD={password}
export SPRING_CLOUD_DEPLOYER_CLOUDFOUNDRY_SKIP_SSL_VALIDATION=false
```

You need to fill in {org}, {space}, {email} and {password} before running these commands.

Note

Only set 'Skip SSL Validation' to true if you're running on a Cloud Foundry instance using self-signed certs (e.g. in development). Do not use for production.

Now we are ready to start the server application:

```
java -jar spring-cloud-dataflow-server-cloudfoundry-1.0.0.BUILD-SNAPSHOT.jar [--option1=value1] [--option2=value2] [etc.]
```

4.6 Running Spring Cloud Data Flow Shell locally

Run the shell and optionally target the Admin application if not running on the same host (will typically be the case if deployed on Cloud Foundry as explained [here](#))

```
$ java -jar spring-cloud-dataflow-shell-1.0.0.M3.jar
```

```
server-unknown:>dataflow config server http://dataflow-server.cfapps.io
Successfully targeted http://dataflow-server.cfapps.io
dataflow:>
```

5. Configuration Reference

The following pieces of configuration must be provided, e.g. by setting variables in the apps environment, or passing variables on the Java invocation:

```
# Default values cited after the equal sign.
# Example values, typical for Pivotal Web Services, cited as a comment

# url of the CF API (used when using cf login -a for example), e.g. https://api.run.pivotal.io
# (for setting env var use SPRING_CLOUD_DEPLOYER_CLOUDFOUNDRY_API_ENDPOINT)
spring.cloud.deployer.cloudfoundry.apiEndpoint=

# name of the organization that owns the space above, e.g. youruser-org
# (For Setting Env var use SPRING_CLOUD_DEPLOYER_CLOUDFOUNDRY_ORGANIZATION)
spring.cloud.deployer.cloudfoundry.organization=

# name of the space into which modules will be deployed
# (for setting env var use SPRING_CLOUD_DEPLOYER_CLOUDFOUNDRY_SPACE)
spring.cloud.deployer.cloudfoundry.space=<same space as server when running on CF, or 'development'>

# the root domain to use when mapping routes, e.g. cfapps.io
# (for setting env var use SPRING_CLOUD_DEPLOYER_CLOUDFOUNDRY_DOMAIN)
spring.cloud.deployer.cloudfoundry.domain=

# Comma separated set of service instance names to bind to the module.
# Amongst other things, this should include a service that will be used
# for Spring Cloud Stream binding
# (for setting env var use SPRING_CLOUD_DEPLOYER_CLOUDFOUNDRY_SERVICES)
spring.cloud.deployer.cloudfoundry.services=redis,rabbit

# username and password of the user to use to create apps (modules)
# (for setting env var use SPRING_CLOUD_DEPLOYER_CLOUDFOUNDRY_USERNAME and
#  SPRING_CLOUD_DEPLOYER_CLOUDFOUNDRY_PASSWORD)
spring.cloud.deployer.cloudfoundry.username=
spring.cloud.deployer.cloudfoundry.password=

# Whether to allow self-signed certificates during SSL validation
# (for setting env var use SPRING_CLOUD_DEPLOYER_CLOUDFOUNDRY_SKIP_SSL_VALIDATION)
spring.cloud.deployer.cloudfoundry.skipSslValidation=false
```

Part II. Appendices

6. Building

6.1 Basic Compile and Test

To build the source you will need to install JDK 1.8.

The build uses the Maven wrapper so you don't have to install a specific version of Maven. To enable the tests for Redis you should run the server before building. See below for more information on how to run Redis.

The main build command is

```
$ ./mvnw clean install
```

You can also add '-DskipTests' if you like, to avoid running the tests.

Note

You can also install Maven (>=3.3.3) yourself and run the `mvn` command in place of `./mvnw` in the examples below. If you do that you also might need to add `-P spring` if your local Maven settings do not contain repository declarations for spring pre-release artifacts.

Note

Be aware that you might need to increase the amount of memory available to Maven by setting a `MAVEN_OPTS` environment variable with a value like `-Xmx512m -XX:MaxPermSize=128m`. We try to cover this in the `.mvn` configuration, so if you find you have to do it to make a build succeed, please raise a ticket to get the settings added to source control.

The projects that require middleware generally include a `docker-compose.yml`, so consider using [Docker Compose](#) to run the middleware servers in Docker containers. See the README in the [scripts demo repository](#) for specific instructions about the common cases of mongo, rabbit and redis.

6.2 Documentation

There is a "full" profile that will generate documentation. You can build just the documentation by executing

```
$ ./mvnw package -DskipTests=true -P full -pl spring-cloud-dataflow-server-cloudfoundry-docs -am
```

6.3 Working with the code

If you don't have an IDE preference we would recommend that you use [Spring Tools Suite](#) or [Eclipse](#) when working with the code. We use the [m2eclipse](#) eclipse plugin for maven support. Other IDEs and tools should also work without issue.

Importing into eclipse with m2eclipse

We recommend the [m2eclipse](#) eclipse plugin when working with eclipse. If you don't already have m2eclipse installed it is available from the "eclipse marketplace".

Unfortunately m2e does not yet support Maven 3.3, so once the projects are imported into Eclipse you will also need to tell m2eclipse to use the `.settings.xml` file for the projects. If you do not do this

you may see many different errors related to the POMs in the projects. Open your Eclipse preferences, expand the Maven preferences, and select User Settings. In the User Settings field click Browse and navigate to the Spring Cloud project you imported selecting the `.settings.xml` file in that project. Click Apply and then OK to save the preference changes.

Note

Alternatively you can copy the repository settings from [.settings.xml](#) into your own `~/.m2/settings.xml`.

Importing into eclipse without m2eclipse

If you prefer not to use m2eclipse you can generate eclipse project metadata using the following command:

```
$ ./mvnw eclipse:eclipse
```

The generated eclipse projects can be imported by selecting `import existing projects` from the `file` menu.