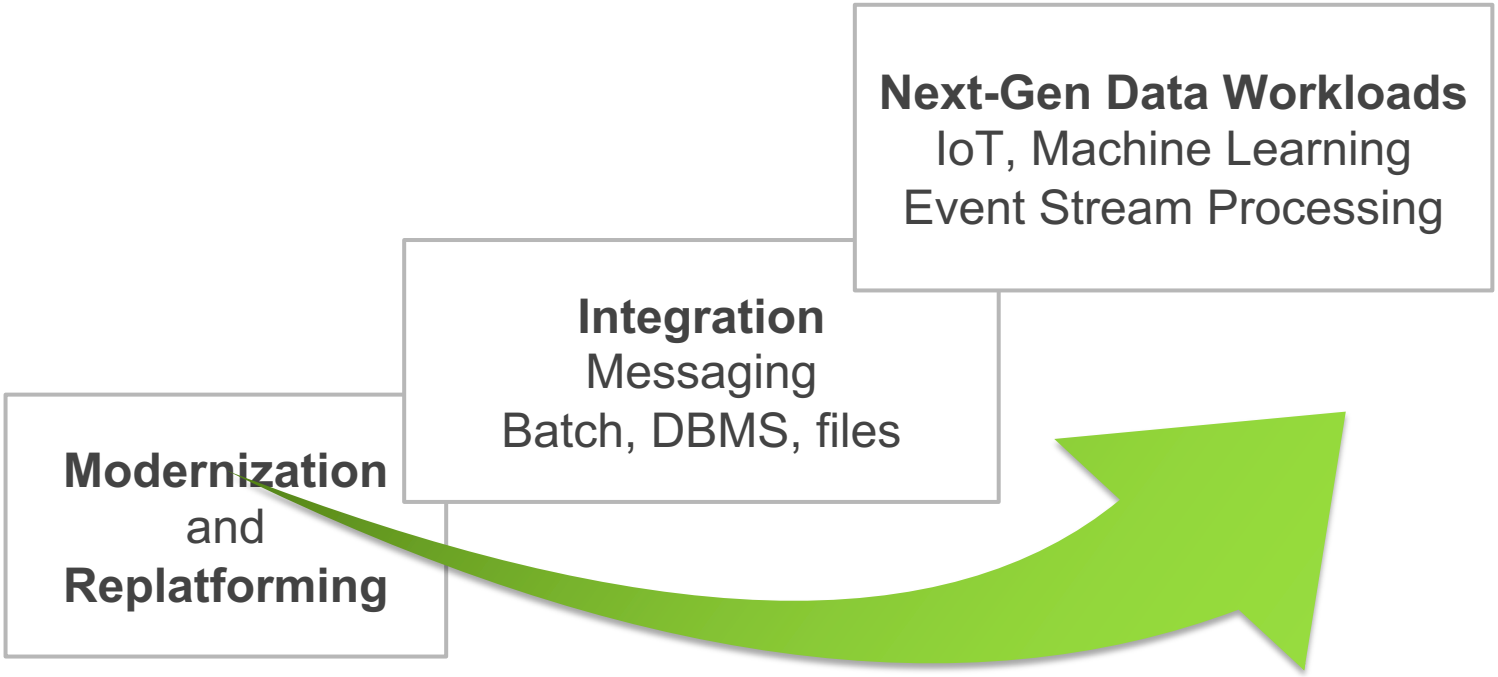
Spring Cloud Data Flow

An Overview

vmware

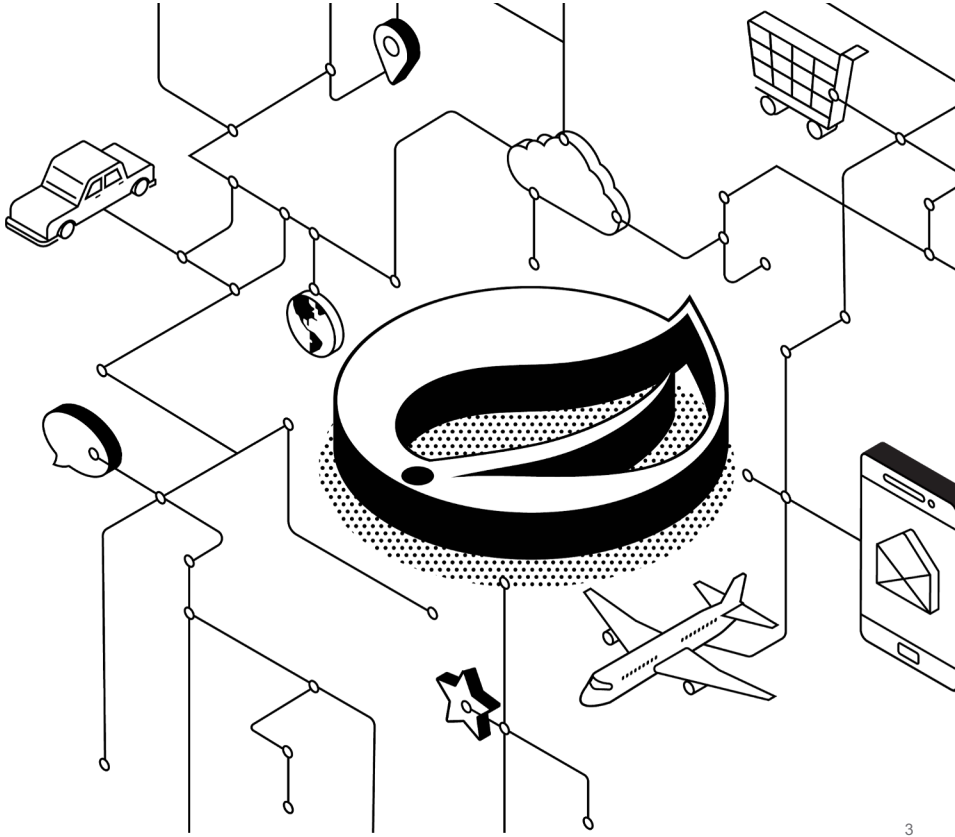
Pivotal

What is Spring Cloud Data Flow Used For?

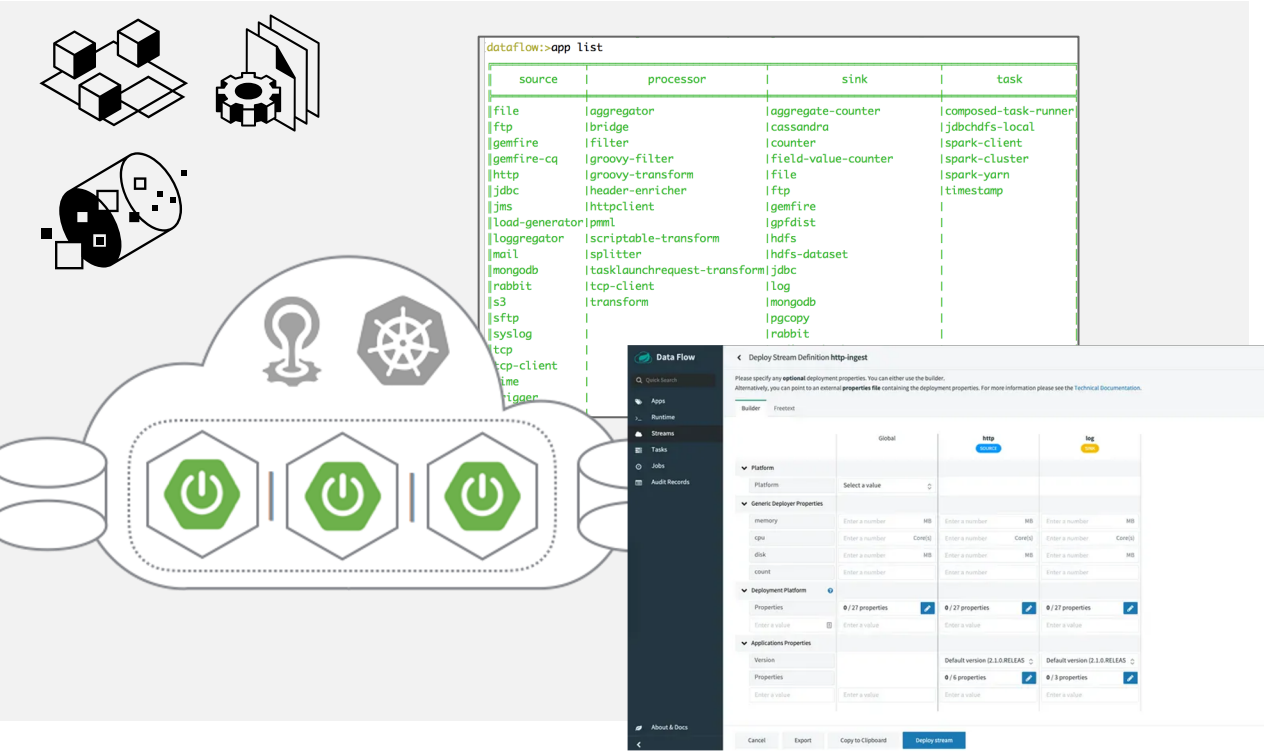


Trends

- Enterprises are adopting DevOps practices in their transition into software and data-driven businesses.
- ETL integration with existing systems, and modernization efforts are still very important.
- Continuous Event processing is becoming mainstream.
- Organizations are finding new ways to integrate IoT data flows and Machine Learning.



What Is Spring Cloud Data Flow



Spring Cloud Data Flow is a Microservices toolkit for building data pipelines.

Pipelines consist of Spring Boot apps, Spring Cloud Stream events or Spring Cloud Task processes.

The Data Flow server provides interfaces to compose and deploy pipelines onto Tanzu.

Microservices Are Not Just For REST & Web Apps

Spring Cloud Task

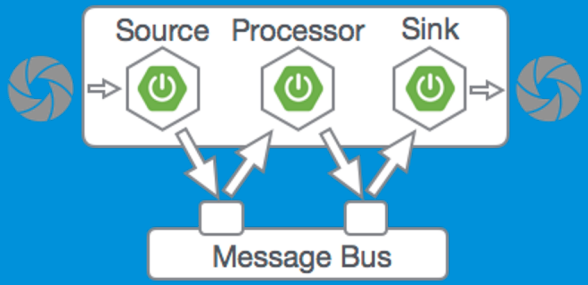
Tasks are finite Boot Microservices. They connect to data/storage. The system tracks invocations, exit-status

Useful for ETL e.g. between DBMS and analytics clusters like Hadoop

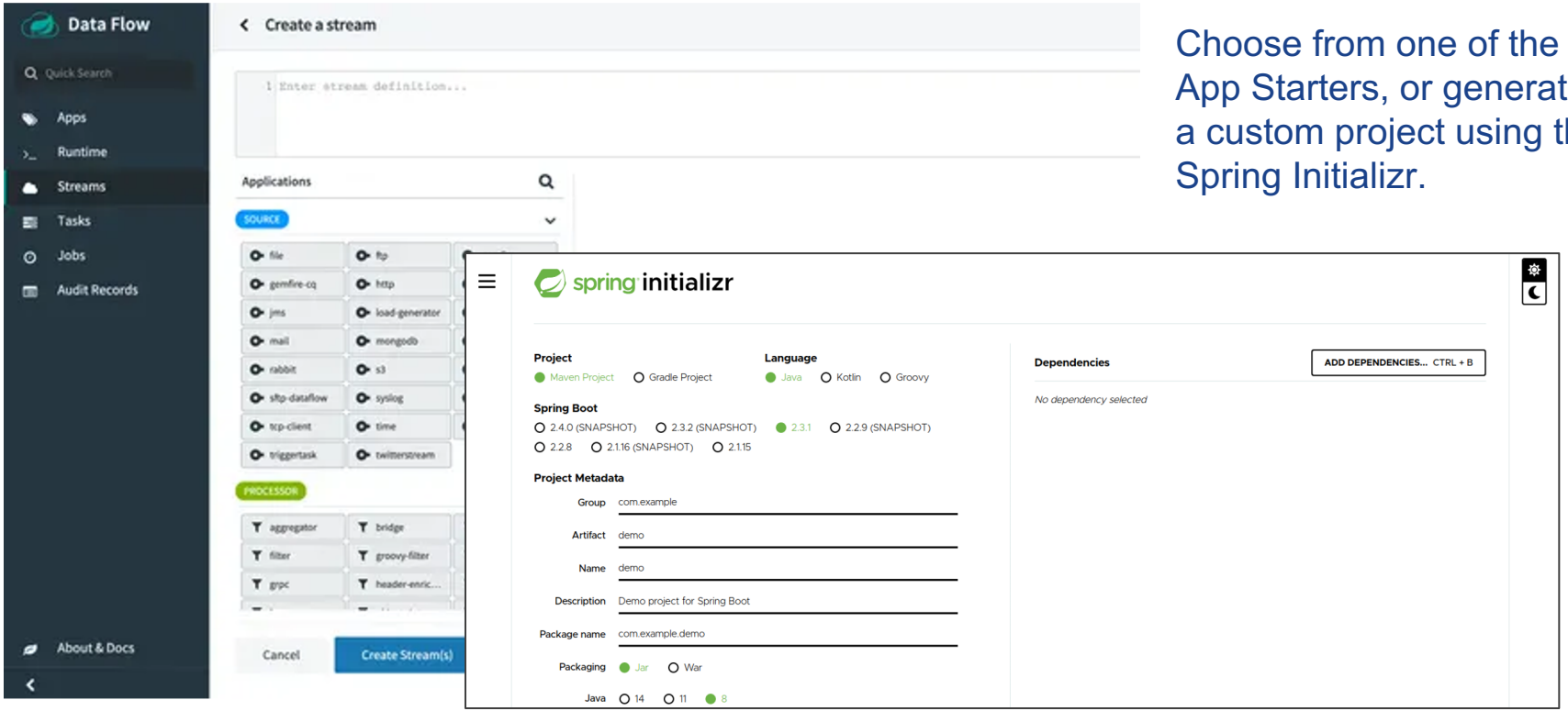
Spring Cloud Stream

Message based Microservices. Loose coupling via pub/sub topics. Pluggable message bus Kafka or RabbitMQ

For integration and stream processing



Don’t Start From Scratch



Choose from one of the App Starters, or generate a custom project using the Spring Initializr.

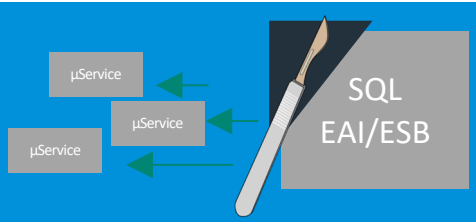
Replatforming

Move Java applications running integration workloads

From App servers on dedicated bare-metal hosts or VMs

To event-driven, or on-demand task oriented Java Microservices, deployed on a Cloud Native platform.

Modernization



Refactor Monolithic ETL workloads

Migrate SQL stored procedures

Rewrite Shell scripts

From EAI and ESB servers

To Java Microservices, deployed on a Cloud Native platform.

Why SCDF on Tanzu?

Spring Boot Microservices
Build production-ready Java apps
Choose from curated starters

The Spring Ecosystem
Mature, actively maintained libraries
Trusted, helpful community

Pivotal Cloud Foundry
Open source PaaS for DevOps & Ci/CD
Heavy Lifting: Elastic scale, HA, Secure
For on-premise & cross-cloud

Spring Cloud Data Flow Sucess Stories

Telecommunications	Data ingest, stream processing, integration with Spark and Hadoop
Real Estate Analytics	Dataflow modernization, consolidation of legacy apps onto PCF (details...)
Health Care	ETL modernization, change data capture Event stream processing, GemFire (details...)
Banking	Integration pipelines Batch and near-real-time
Insurance	Rules-engine modernization (.NET) Offload DBMS and Data Warehouse to microservices

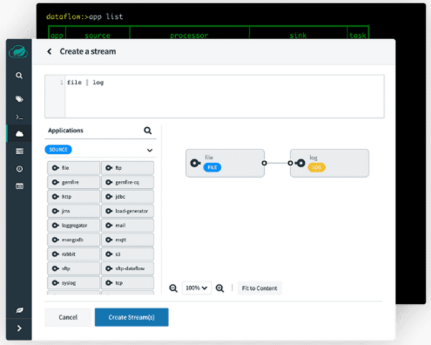
dataflow.spring.io

Spring Cloud **Data Flow** Features Documentation Getting Started Community Search

Microservice based **Streaming** and **Batch** data processing for Cloud Foundry and Kubernetes

- Develop and test microservices for data integration that do one thing and do it well
- Use prebuilt microservices to kick start development
- Compose complex topologies for streaming and batch data pipelines
- Open Source, Apache Licensed

Getting Started | View on Github



Flexible

Write Stream and Batch processing logic in multiple programming languages.
Use your favorite messaging middleware for Stream processing.
Interact with popular monitoring systems and dashboards.

Read more

Familiar Tools

Kick-start the solution for your use-case using our drag and drop designer.
Don't like designers? Use pipes and filters based textual Domain Specific Language instead.
Integrate using RESTful APIs.

Read more

Spring Opinionated

Are you already building microservices with Spring Boot?
Jump to the developer guide and extend the same learnings to build streaming and batch applications.

Read more

Thank You

Pivotal