

Project Hop

exploring the future of data integration

Project Hop

The Hop Orchestration Platform aims to facilitate all aspects of data and metadata orchestration

Why Project Hop?

- Lots of great data platforms
- But:
 - Writing code is hard
 - Lots of (frequently changing) APIs
- There is a need for a platform that is
 - Up to date with the latest technologies
 - Independent (ASF)
 - Easy to develop, test, deploy, document, extend, ...
 - Easy to contribute to





Project Hop - Current Status

- Done
 - Forked from Kettle 8.2.0.7
 - Package, code refactoring
 - Dependency updates
 - Karaf out!
 - Infrastructure
 - Database types as plugins
- In Progress
 - ASF Incubation
 - GUI: remove XUL, rewrite DB GUI
 - Documentation setup
- Build: ~200MB, startup time: 2-3 seconds

Project Hop - Roadmap

- Start ASF incubation (Apache Hop)
- Next-gen UI
- Configurable runtimes (Beam, Airflow, ...)
- Documentation as first class citizen (asciidoc, Doc as Code)
- Everything is a plugin
- Multi project/environment (dev,prod,...)
- Everything can be tested (integration, unit, regression, data testing)
- Log, lineage, audit, monitoring

Project Hop - Architecture

What?

- Separate metadata from runtime code, tooling
- Minimal installation footprint
- Stateless (e.g. Docker, K8S)
- Facilitate execution on a variety of platforms (Beam, Airflow,...)
- Support different life cycles
 - development, test, acceptance, production
 - testing, CI/CD, ...

Project Hop - Architecture

How?

- Allow orchestration workloads defined as metadata
- Metastore as single interface for all expressions of metadata
- Pluggable persistence (XML, JSON, YAML)
- 100% UI coverage

Project Hop - Concept Analogies

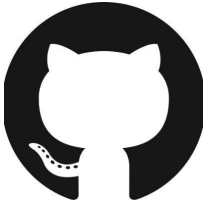
Kettle/PDI	Hop
Spoon	hop-gui
Kitchen, Pan, Maitre	hop-run
Carte	hop-server
Job	(Work)Flow
Transformation	Pipeline
Job Entry	Action
Transformation Step	Transform or Function
Hop	Hop



Together, We Create!

Project Hop - Available Tools

Project Hop: <https://www.project-hop.org>, <https://forums.project-hop.org>



Project Hop - Community is key!

Contributing > code!

- Answer forum questions
- Write documentation
- Test functionality, write test flows, pipelines
- Create bugs, feature requests
- Start local user groups, organize meetups
- Spread the word

Project Hop - Early Contributors

- PR (thanks, @pfabrici)
- Logo (thanks, @sramazzina)
- Docs translation
- Everyone who provided input/feedback on Slack, Twitter, ...



hop

HOP

HOP

Hop

Project Hop - Additional Plans

- After initial Hop release(s): LEAN!
- Lean Enterprise ANalytics
- Metadata driven analytics
- Pluggable
- Seamless integration with Hop