



# Apache Flink

Concrete Architecture

Group Small

# In Today's Presentation

- Concrete Architecture
- Subsystem Functionality and Interactions
- Derivation Process
- Lsedit Demo
- Reflexion Analysis
- ResourceManager Subsystem
- Use Case
- Lessons Learned
- References



# Concrete Architecture

## Subsystems

- JobManager
- TaskManager
- JobMaster
- Dispatcher
- ResourceManager

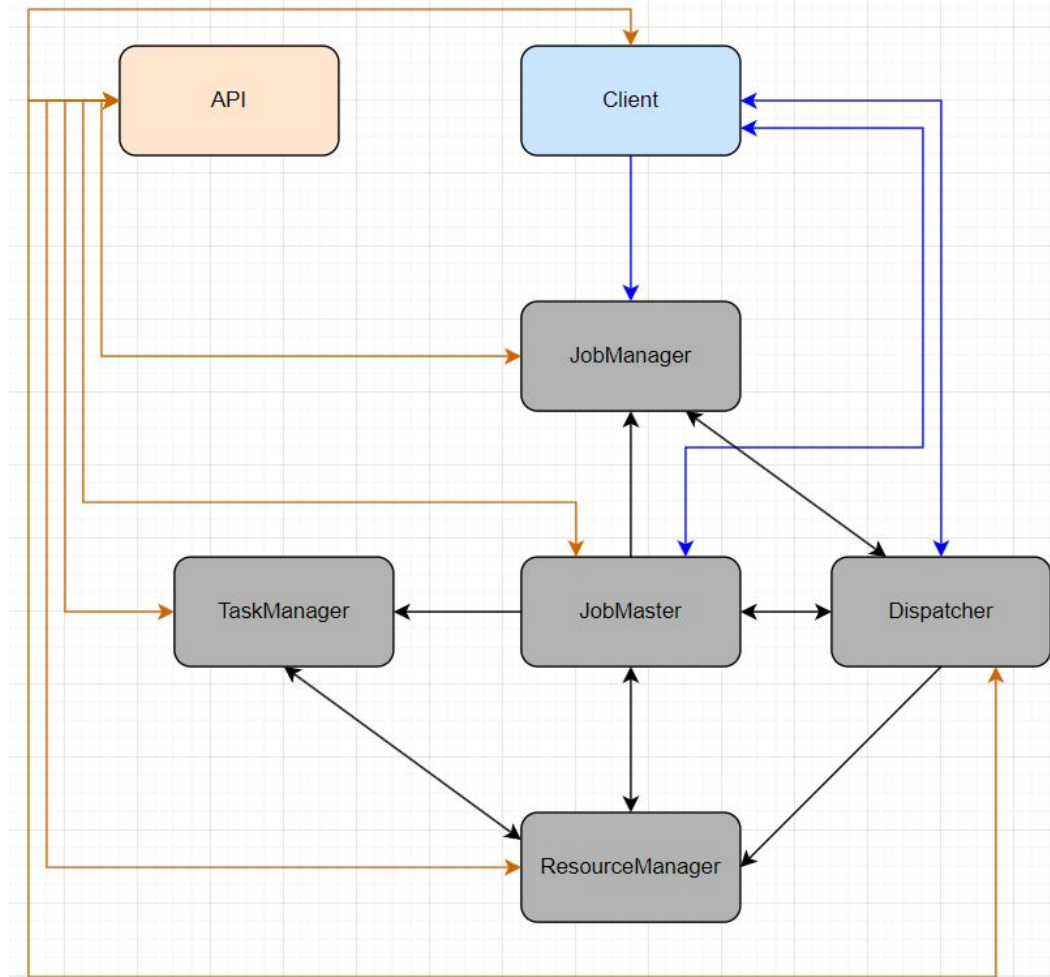


Figure 1: Concrete architecture of Apache Flink

# Derivation Process

## Final Process

- Understand
- Scripts
- Lsedit

## Previous Processes

- Elimination
- Source Folder

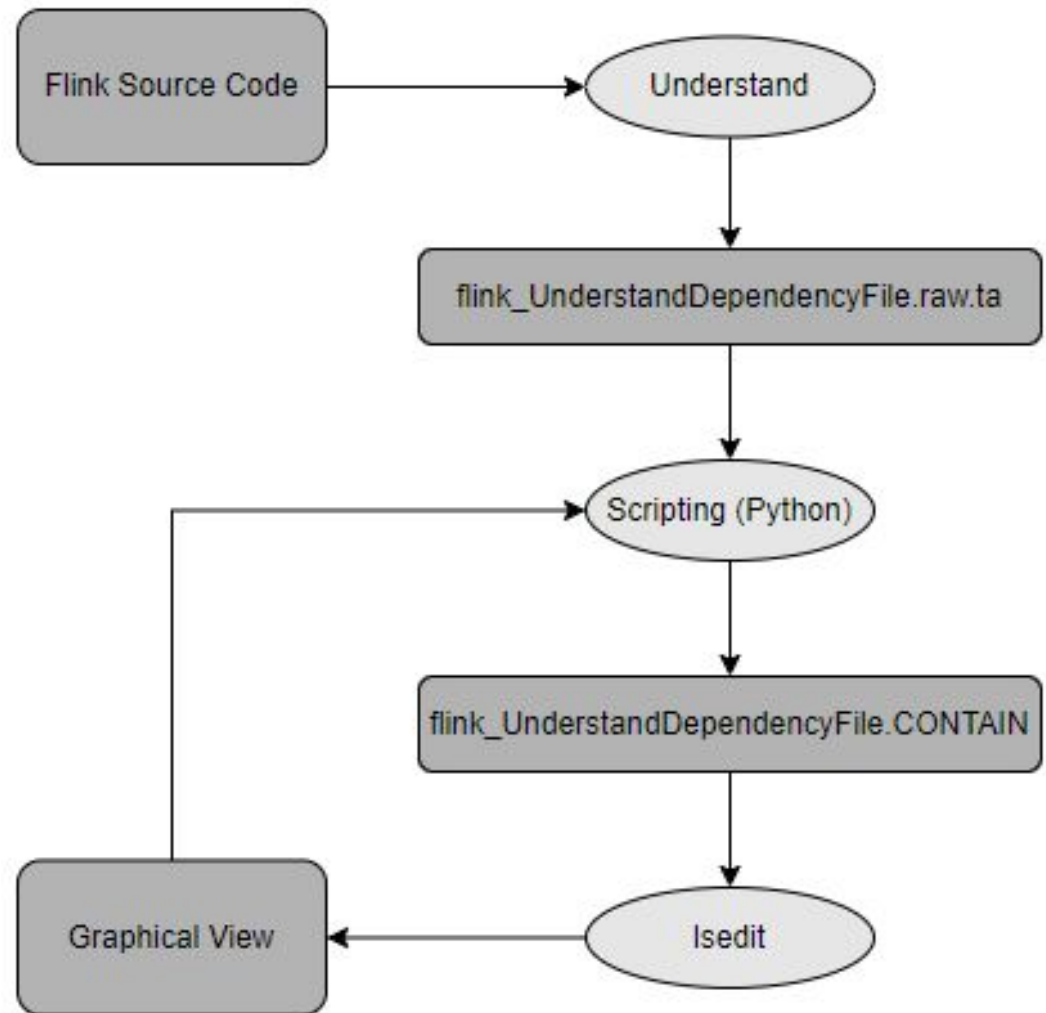


Figure 2: Derivation process diagram

# Lsedit Demo

# Architectural Style

## Client-Server Style: RPC

1. Client Request
2. Server Implementation
3. Remote Execution

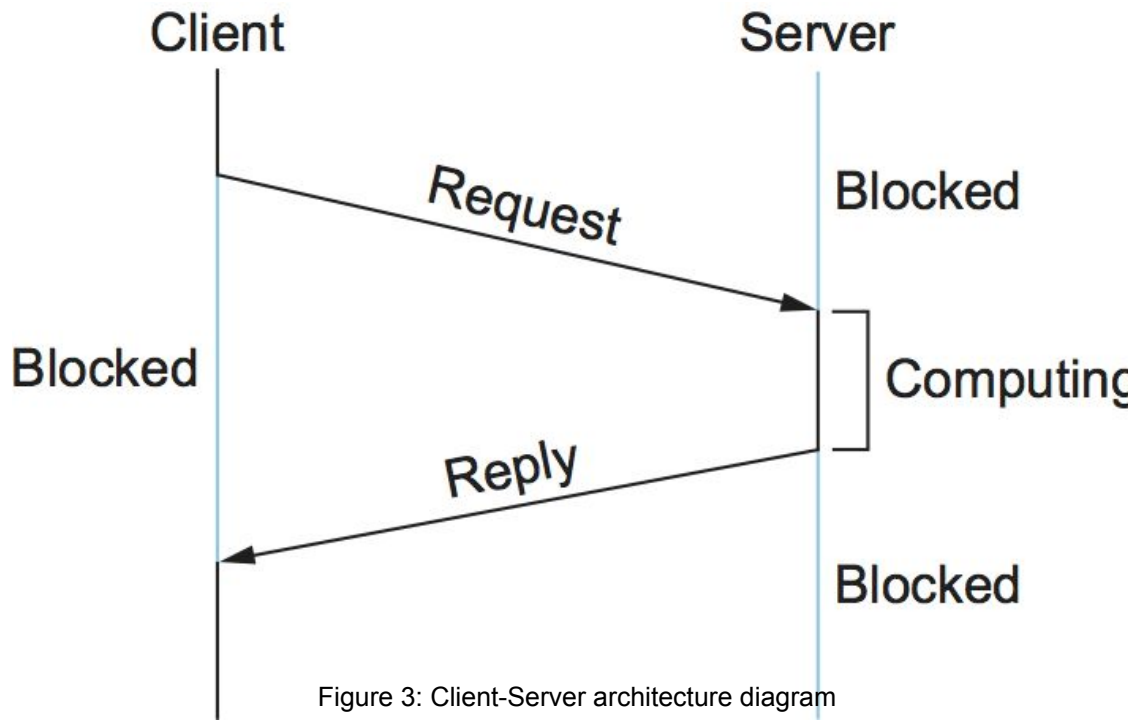


Figure 3: Client-Server architecture diagram

# Reflexion Analysis

Conceptual Architecture

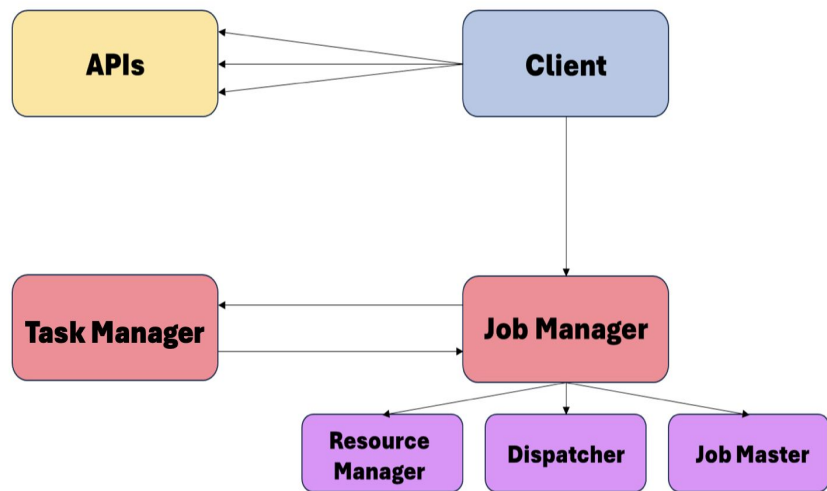


Figure 4: Conceptual architecture of Apache Flink

Concrete Architecture

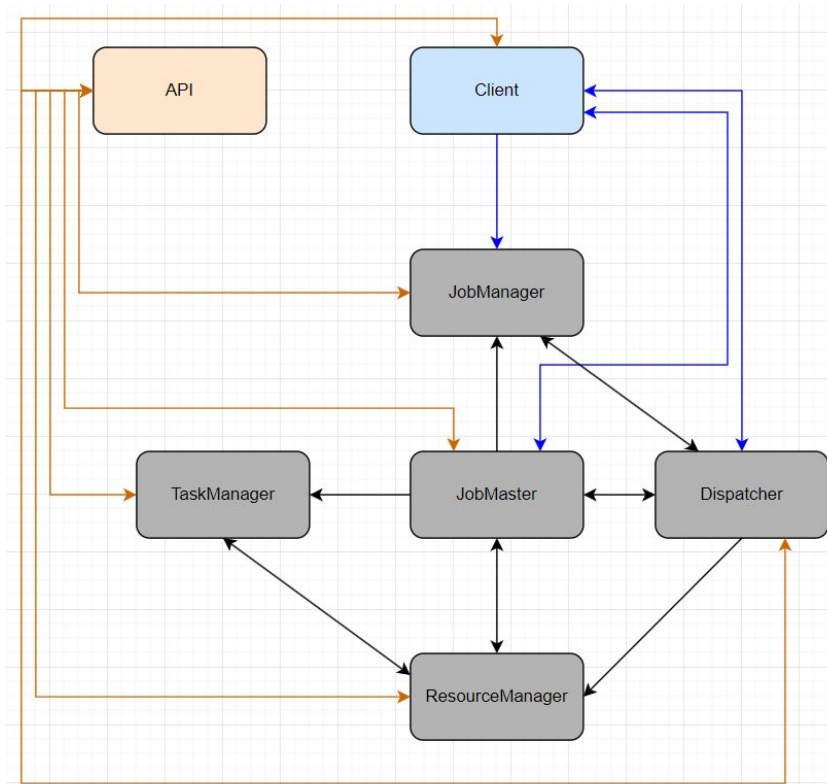


Figure 1

# Divergence (ResourceManager $\longleftrightarrow$ JobMaster)

WHAT	connectToResourceManager (flink-runtime/.../runtime/jobmaster/slotpool/SlotPool.java) depends on ResourceManagerGateway (flink-runtime/.../runtime/resourcemanager/ResourceManagerGateway.java)
WHO	Stephen Ewen (Senior Programmer, Co-Creator)
WHEN	Created 08/02/2017      Resolved 08/29/2017
WHY	<p>The following are <b>core aspects of the ResourceManager</b> design:</p> <p>The ResourceManager no longer has a resource pool size, but <b>receives individual requests for slots</b>. That way, jobs can request TaskManagers of different resources (Memory/CPU).</p> <p><b>/**</b></p> <ul style="list-style-type: none"><li>* Connects the SlotPool to the given ResourceManager. After this method is called, the SlotPool</li><li>* will be able to request resources from the given ResourceManager.</li><li>*</li><li>* @param resourceManagerGateway The RPC gateway for the resource manager.</li></ul>



# Divergence (Resource Manager )

WHAT	MiniClusterITCase.testHandleStreamingJobsWhenNotEnoughSlot, ZooKeeperLeaderElectionConnectionHandlingTest, HAJobRunOnHadoopS3FileSystemITCase, depend on finegrained_resource_management
WHO	Matthias Pohl (Contributor, Jr Dev)
WHEN	Created: 11/Oct/23 14:31 Resolved:16/Oct/23 09:28
WHY	The dedicated finegrained_resource_management stage was removed because the fine-grained ResourceManager became the default implementation

# Resource Manager

- Allocating and deallocating resources
- Manages Task Slots
- Distributes the slots of TMs
- Slot caching
- Slot sharing

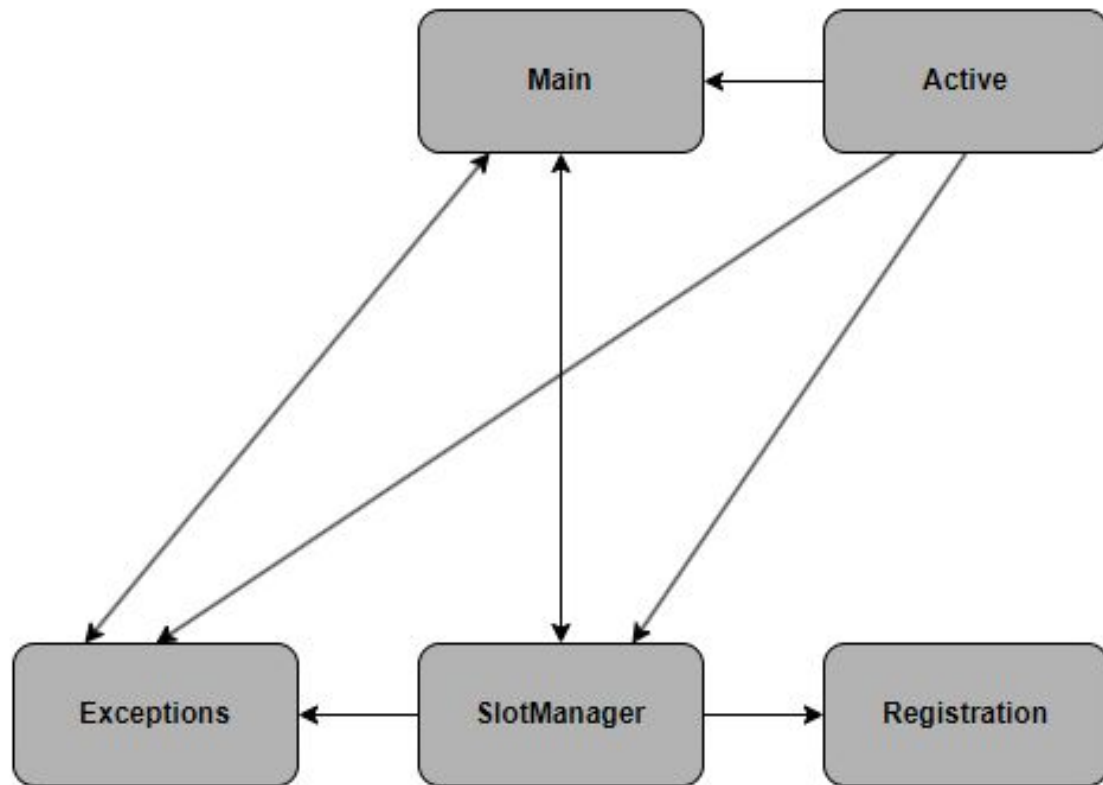


Figure 5 ResourceManager Components and Interactions

Slot Allocation

Starting TaskManagers

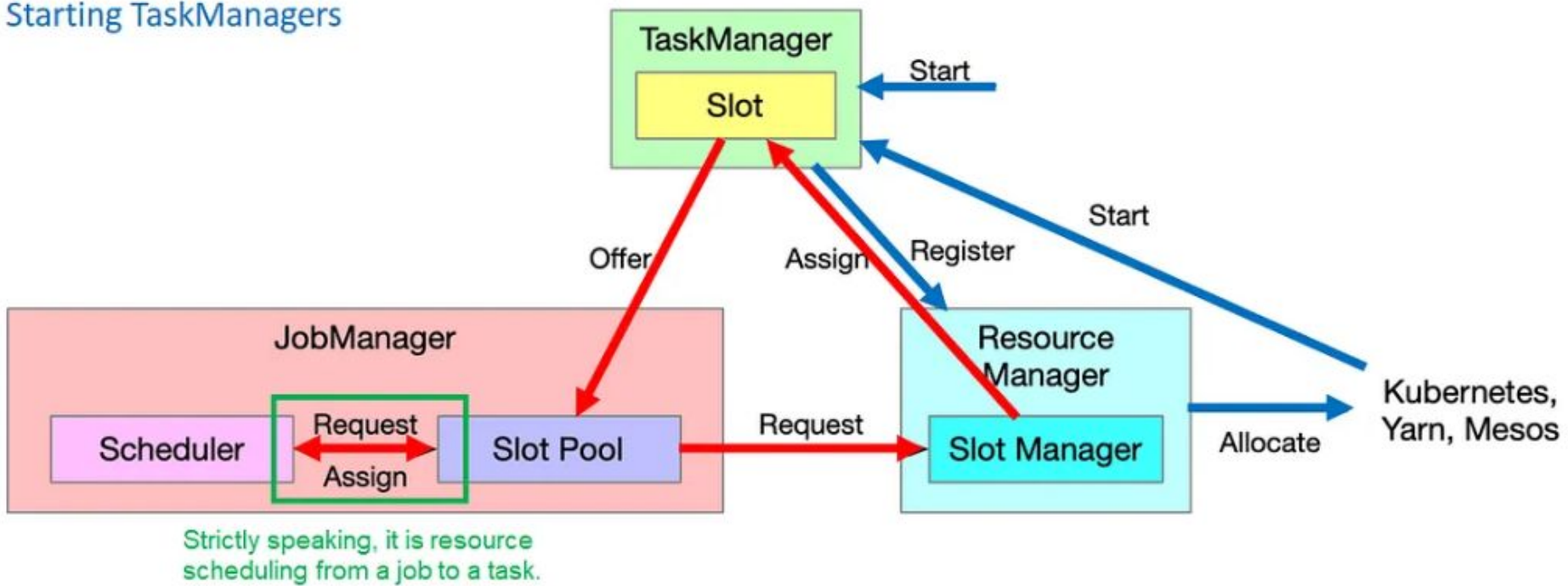


Figure 6  
Image taken from: <https://alibaba-cloud.medium.com/data-warehouse-in-depth-interpretation-of-flink-resource-management-mechanism-5c13b531abfa>

# Sequence Diagram

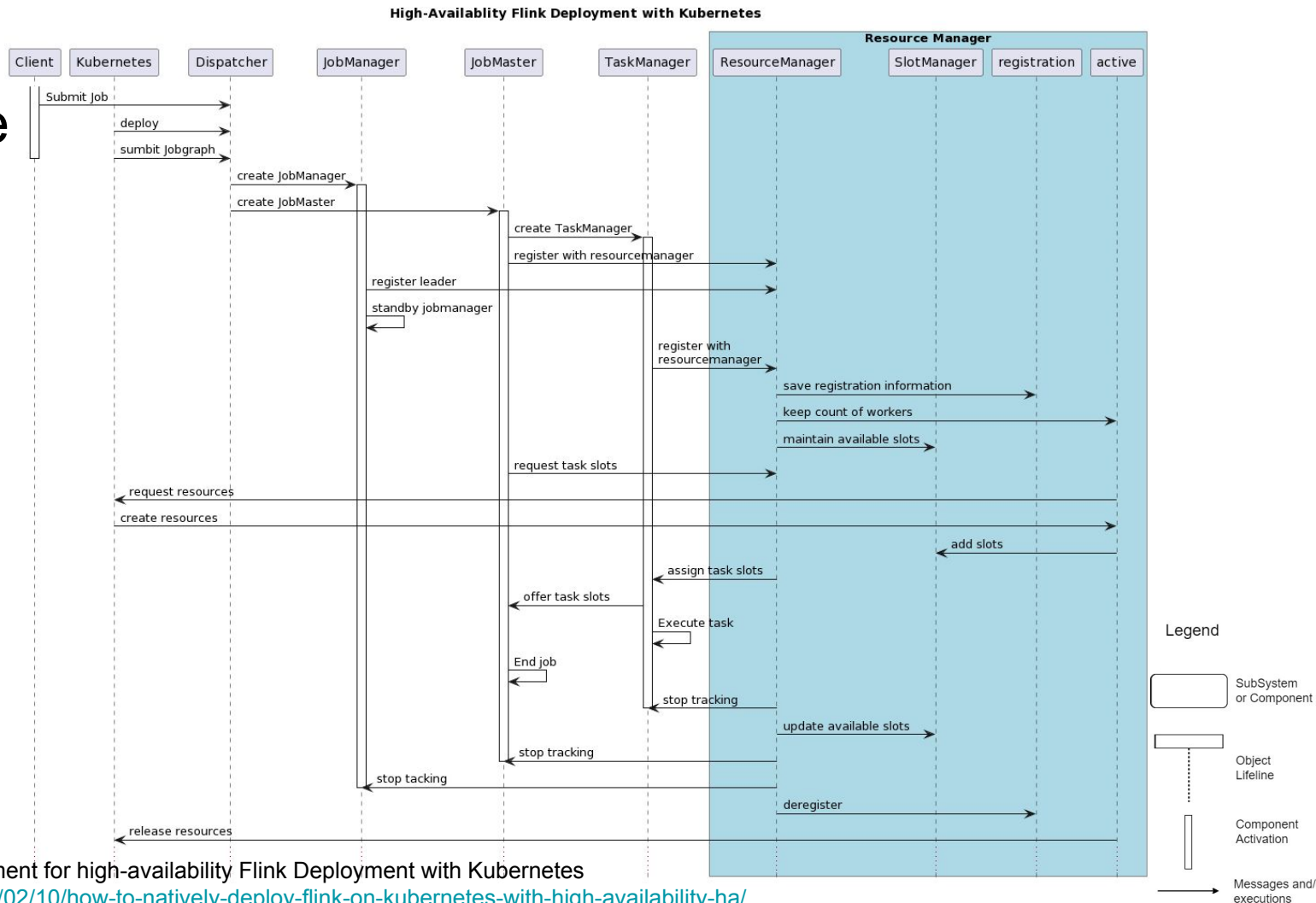


Figure 7 Resource management for high-availability Flink Deployment with Kubernetes

<https://flink.apache.org/2021/02/10/how-to-natively-deploy-flink-on-kubernetes-with-high-availability-ha/>

# Use Case Diagram

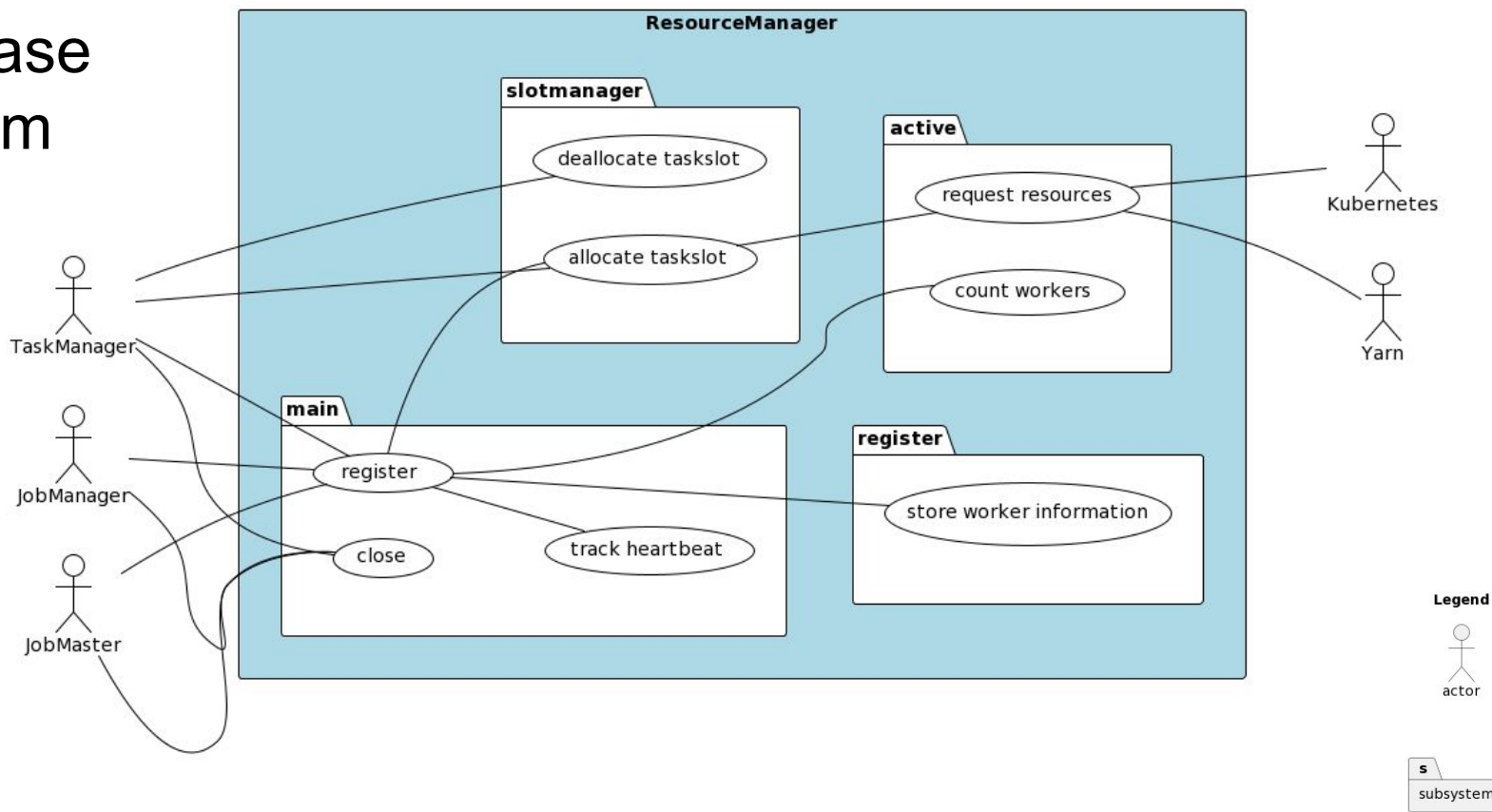


Figure 8 Resource management subsystem use cases

<https://flink.apache.org/2021/02/10/how-to-natively-deploy-flink-on-kubernetes-with-high-availability-ha/>

# Lessons Learned

- Grok
  - Convenience determining subsystems
  - jgrok
- Lab practice
  - Understanding .CONTAIN files
  - Familiarity with Isedit
  - Subsystems within subsystems



Figure 10: SWAG icon

<https://www.swag.uwaterloo.ca/tools/grok/>

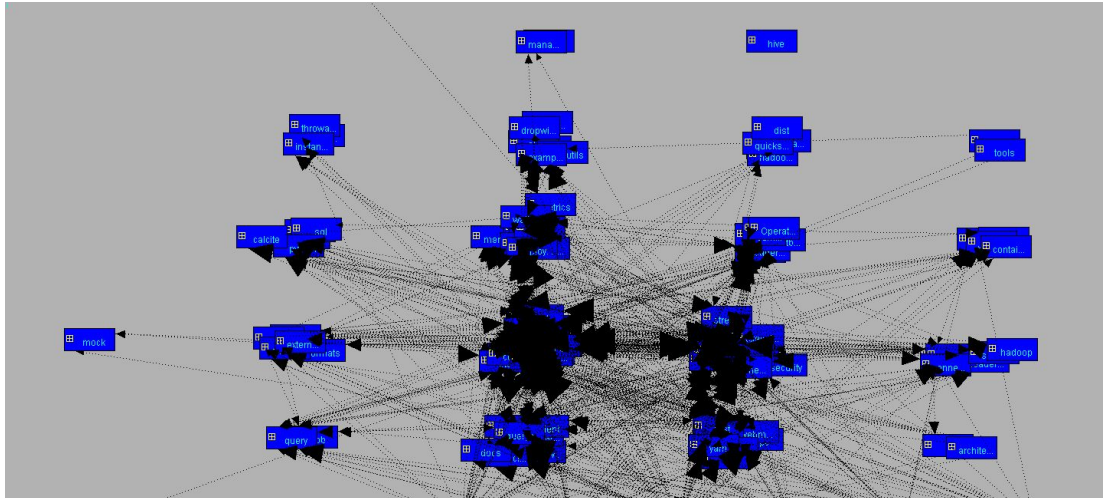


Figure 9: Itr1 of concrete architecture visualization

Thank You

# References

<https://nightlies.apache.org/flink/flink-docs-master/docs/concepts/flink-architecture/>

<https://nightlies.apache.org/flink/flink-docs-master/api/java/org/apache/flink/runtime/resourcemanager/ResourceManager.html>

<https://jbcodeforce.github.io/flink-studies/architecture/>

<https://alibaba-cloud.medium.com/data-warehouse-in-depth-interpretation-of-flink-resource-management-mechanism-5c13b531abfa>

<https://flink.apache.org/2021/02/10/how-to-natively-deploy-flink-on-kubernetes-with-high-availability-ha/>

<https://issues.apache.org/jira/projects/FLINK/issues/FLINK-32513?filter=allopenissues>

EECS

4314 Lecture Slides

EECS

4314 Lab Notes