AliECS

A New Experiment Control System for the ALICE Experiment

Teo Mrnjavac
CERN EP-AID-DA
on behalf of the ALICE O²/FLP project
9 October 2019
The ALICE Online-Offline computing system

- Multiprocess data flow and processing framework
- 100,000s of processes, ~1000 machines
- Synchronous and asynchronous (grid-like) workflows
- One computing system, 2 types of node arranged in 2 clusters:
  - FLP - First Level Processors
  - EPN - Event Processing Nodes
- Operations start in 2021
ECS and $O^2$ cluster control

- Manage the lifetime of thousands of stateful processes in the $O^2$/FLP cluster (control of $O^2$/EPN delegated to a specialized $O^2$/EPN cluster control)
- Minimize the waste of beam time by reusing processes and avoiding time-consuming process restart operations
- Interface with the LHC, the trigger system, the Detector Control System and other systems through common APIs
“Program against your datacenter like it’s a single pool of resources.”
Managing a cluster with Apache Mesos

“Program against your datacenter like it’s a single pool of resources.”

- Mesos acts as a unified **distributed execution environment** which streamlines how AliECS manages its components, resources and tasks inside the O²/FLP farm.

- Benefits:
  - **knowledge** of what runs where,
  - **resource management** (ports, CPU, RAM, ...),
  - **transport** for control messages,
  - task event **notification** (dead, failed to launch, ...),
  - node attributes, high availability, checkpointing, ...
AliECS in a nutshell
AliECS in a nutshell

• Components:
  • AliECS core (incl. Apache Mesos scheduler)
  • AliECS executor
  • $O^2$ control and configuration FairMQ plugin ($\text{FairMQPlugin}_\text{OCC}$)
  • $O^2$ control and configuration library ($\text{libOCC}$)
  • AliECS control and configuration utility ($\text{coconut}$)
  • Single-node OCC debug utility ($\text{peanut}$)

• Also available:
  • The web-based AliECS GUI
  • AliECS deployment mechanism
AliECS concepts

- AliECS schedules, configures and controls **tasks**
- A task represents a stateful process, which implements a **role**
- A collection of AliECS roles (arranged in a workflow) along with their configuration is an **environment**
- Tasks, roles and environments have their own **state machines**
- An environment represents the collective state of its constituent roles
- When an environment reaches the **RUNNING** state, it is granted a unique **run number** which remains valid until the **RUNNING** state exits
AliECS workflow and task configuration

- Based on **Git**, multiple repositories per AliECS instance
- Task descriptors and workflow templates are **YAML** (plus template system)
- Once loaded, every task type and workflow is **uniquely identified** by git repository + task/workflow file name + git revision

Documentation: https://github.com/AliceO2Group/Control/blob/master/coconut/doc/coconut_repository.md
AliECS workflow and task configuration

1. Workflow template
2. Task references
3. Configuration references
4. Consul
5. YAML
6. Task descriptors

AliECS core

Environment

role

role

role

role

role

role

role

Task configuration

port assignments

task-specific key-values

Tasks

launch tasks (Mesos)

① ②: create environment
STANDBY state

③ ④: control environment
CONFIGURE transition

task configuration

push
### AliECS GUI

**Number of Tasks:** 7

**ID:** 6be3a3a3-cdb5-11e9-b3b9-fa163eb219f7

**Created:** 9/2/2019, 9:11:09 PM

**State:** RUNNING

**Root Role:** readout-qp-1

<table>
<thead>
<tr>
<th>name</th>
<th>locked</th>
<th>taskid</th>
<th>status</th>
<th>state</th>
<th>className</th>
<th>deploymentInfo</th>
<th>actions</th>
</tr>
</thead>
<tbody>
<tr>
<td>github.com/AliceO2Group/ControlWorkflows/tasks/readout@master#01f13a3-cdb5-11e9-b3b9-fa163eb219f7</td>
<td>true</td>
<td>6df1f03-cdb5-11e9-b3b9-fa163eb219f7</td>
<td>ACTIVE</td>
<td>RUNNING</td>
<td>github.com/AliceO2Group/ControlWorkflows/tasks/readout@master</td>
<td>teo-test19</td>
<td>More</td>
</tr>
<tr>
<td>github.com/AliceO2Group/ControlWorkflows/tasks/external-dpl-clock@master#ebf2f3ca37-cdb5-11e9-b3b9-fa163eb219f7</td>
<td>true</td>
<td>6df1c37-cdb5-11e9-b3b9-fa163eb219f7</td>
<td>ACTIVE</td>
<td>RUNNING</td>
<td>github.com/AliceO2Group/ControlWorkflows/tasks/external-dpl-clock@master</td>
<td>teo-test19</td>
<td>More</td>
</tr>
<tr>
<td>github.com/AliceO2Group/ControlWorkflows/tasks/readout-proxy@master#d9a3c8-db5-11e9-b3b9-fa163eb219f7</td>
<td>true</td>
<td>6df1a3c8-db5-11e9-b3b9-fa163eb219f7</td>
<td>ACTIVE</td>
<td>RUNNING</td>
<td>github.com/AliceO2Group/ControlWorkflows/tasks/readout-proxy@master</td>
<td>teo-test19</td>
<td>More</td>
</tr>
<tr>
<td>github.com/AliceO2Group/ControlWorkflows/tasks/Dispatcher@master#01f13a3-cdb5-11e9-b3b9-fa163eb219f7</td>
<td>true</td>
<td>6df1f03-cdb5-11e9-b3b9-fa163eb219f7</td>
<td>ACTIVE</td>
<td>RUNNING</td>
<td>github.com/AliceO2Group/ControlWorkflows/tasks/Dispatcher@master</td>
<td>teo-test19</td>
<td>More</td>
</tr>
<tr>
<td>github.com/AliceO2Group/ControlWorkflows/tasks/QC-TASK-RUNNER-daqTask@master#01f502-db5-11e9-b3b9-fa163eb219f7</td>
<td>true</td>
<td>6df1c502-db5-11e9-b3b9-fa163eb219f7</td>
<td>ACTIVE</td>
<td>RUNNING</td>
<td>github.com/AliceO2Group/ControlWorkflows/tasks/QC-TASK-RUNNER-daqTask@master</td>
<td>teo-test19</td>
<td>More</td>
</tr>
<tr>
<td>github.com/AliceO2Group/ControlWorkflows/tasks/daqTask-checker@master#01f113b-cdb5-11e9-b3b9-fa163eb219f7</td>
<td>true</td>
<td>6df1b113b-cdb5-11e9-b3b9-fa163eb219f7</td>
<td>ACTIVE</td>
<td>RUNNING</td>
<td>github.com/AliceO2Group/ControlWorkflows/tasks/daqTask-checker@master</td>
<td>teo-test19</td>
<td>More</td>
</tr>
<tr>
<td>github.com/AliceO2Group/ControlWorkflows/tasks/external-dpl-global-binary-fie-sink@master#01f502-db5-11e9-b3b9-fa163eb219f7</td>
<td>true</td>
<td>6df1c502-db5-11e9-b3b9-fa163eb219f7</td>
<td>ACTIVE</td>
<td>RUNNING</td>
<td>github.com/AliceO2Group/ControlWorkflows/tasks/external-dpl-global-binary-fie-sink@master</td>
<td>teo-test19</td>
<td>More</td>
</tr>
</tbody>
</table>
Conclusions

- The new ALICE O² computing system requires a new control system
- AliECS carries both ECS and O²/FLP cluster control duties
- Opportunity to leverage technologies such as Mesos and Go for a high performance, low latency ECS
  - Mesos gives us resource management, transport and much more
  - Minimize waste of beam time
  - Improved operational flexibility

AliECS on GitHub: github.com/AliceO2Group/Control
Configuration examples: github.com/AliceO2Group/ControlWorkflows