Development of Distributed Geoscience Workflows with WMS-light

A Short Hands-on Tutorial

Presented by: Alexey Cheptsov (cheptsov@hlrs.de)
Workflow Development with WMS-light

CONTENT

1 Research Background
2 Workflows for Geoscience-Applications
3 Getting started with WMS-light
4 Running a Demo-Workflow
1 Research Background
2 Workflows for Geoscience-Applications
3 Getting started with WMS-light
4 Running a Demo-Workflow
ChEESE: Center of Excellence in Solid Earth

- A EU-funded preparatory action for the upcoming Exascale supercomputers
  - Establishes a new Center of Excellence (CoE) in the domain of Solid Earth (SE)
  - Addresses 15 scientific, technical, and socio-economic Exascale Computational Challenges (ECC) in the domain of SE.
  - Develops 12 Pilot Demonstrators (PD) and enable services oriented to society on critical aspects of geohazards like hazard assessment, urgent computing, and early warning forecast.
  - Integrate around HPC and HDA transversal European institutions in charge of operational geophysical monitoring networks, Tier-0 supercomputing centers, academia, hardware developers, and third-parties from SMEs, Industry and public governance bodies (civil protection).
- www.cheese-coe.eu
Research Background

ChEESE: Need of Workflows

- Many geoscience applications are built of coupled codes, running on distributed HPC and Cloud resources
  - Pre- and post-processing
  - Simulation
  - Visualization
Workflow Development with WMS-light

CONTENT

1 Research Background
2 Workflows for Geoscience-Applications
3 Getting started with WMS-light
4 Running a Demo-Workflow
Deployment and Execution Challenge

(1) Use of external databases / storage locations
Workflows for Geoscience Applications

Deployment and Execution Challenge

(1) **Use of external databases / storage locations**
(2) **Distributed Computing- and Data-Infrastructure**
Workflows for Geoscience Applications

Deployment and Execution Challenge

(1) Use of external databases / storage locations
(2) Distributed Computing- and Data-Infrastructure
(3) Need to perform/track multiple experiments (e.g., parametric studies)
Workflow Development with WMS-light

CONTENT

1 Research Background
2 Workflows for Geoscience-Applications
3 Getting started with WMS-light
4 Running a Demo-Workflow
Workflow Management Solution of ChEESE

• **WMS-light**
  • Set of light-weight Java components and shell scripts for launching/management/tracking of the execution of component-based, data- and control-flow interconnected distributed applications (*workflows*)
  • Allows automation of the everyday’s routine operations (submission of applications to HPC schedulers, execution, copying data, etc.), which are frequently performed manually and are thus very time-consuming.
  • On-the-fly deployment on any supported infrastructure (incl. HPC) due to 0-inference into the system software layer of the targeted infrastructure.
  • All middleware runs on the client side → allows integration with almost any compute infrastructure with a minimum of performance overhead
Getting Started with WMS-light

Workflow Management Solution of ChEESE

- WMS-light Architecture

User-Defined Workflows

Workflow Management System

Federated Infrastructure
Getting Started with WMS-light

Workflow Management Solution of ChEESE

- **Major Specifications**
  - To be provided in flexible JSON-format

---

**Component Model**
- Execution options
- Scalability properties
- Data-dependencies

**Infrastructure Model**
- Access details
- Resource Manager information
- Hardware properties

**Deployment Model**
- Deployment properties of components on the infrastructure resources

---

(1) Workflow Specification

(2) Execution Request

(3) Real-Time Monitoring

**UEID (Unique Execution ID)**
e.g. „2020-03-31T19:00:00.000“

**Execution Statistics for all components,**
e.g. status, start and end time, etc.
Getting Started with WMS-light

Let’s try it out

• **WMS-light’s Docker Container**
  
  • Contains all pre-installed software for running WMS-light on the local machine
  
  • Downloadable from: https://fs.hlrs.de/projects/cheese/Dockerfile
  
  • Requires Docker software
  
  • If Docker cannot be installed for some reason, a manual installation is also possible (but is a more time-consuming option).
Installing WMS-light with Docker

- Download the Docker container file
  
  "wget https://fs.hlrs.de/projects/cheese/Dockerfile"

- Build the container
  
  "sudo docker build --rm -t wmslight:0.3 ."

- Run a container instance
  
  "sudo docker run --privileged -ti -e container=docker -v 
  /sys/fs/cgroup:/sys/fs/cgroup wmslight:0.3"
Getting Started with WMS-light

Running a Demo-Workflow

• Download the Docker container file
  
  "wget https://fs.hlrs.de/projects/cheese/Dockerfile"

• Build the container
  
  "sudo docker build --rm -t wmslight:0.3 ."

• Run a container instance
  
  "sudo docker run --privileged -ti -e container=docker -v 
   /sys/fs/cgroup:/sys/fs/cgroup wmslight:0.3"
Getting Started with WMS-light

Running a Demo-Workflow

• Inside the container:
  “cd WMS-light/Demo; ./run_demo.sh”

• Check all the workflow specifications in the
  “WMS-light/Demo/Simple” directory