Reliable digital twin simulation development and execution on HPC infrastructures

Marek Kasztelnik⁽¹⁾, Tomasz Gubała^(1,2), Piotr Nowakowski^(1,2), Jan Meizner^(1,2), Piotr Połeć⁽¹⁾, Maciej Malawski^(1,2), Marian Bubak^(1,2) ⁽¹⁾ ACC Cyfronet AGH University of Science and Technology, Kraków, Poland ⁽²⁾ Sano Centre for Computational Medicine, Kraków, Poland

Collapse sidebar

Motivation

- Creating simulations and testing them is time-consuming
- HPC systems are not intuitive, introducing an additional layer of complexity
- The standard way of using HPC systems does not provide repeatability and reproducibility; it also lacks modern versioning (such as is provided by Git)

Sample use case

🔶 🚍 Menu		🛛 👻 🔍 Q. Search GitLab	
demo steps	Until then, shared runners will be unavailable. Validate your account or	f use your own runners-	
ect information	cyfronet > MEE demo steps		
epository	MEE domo al		
ues O	M Project ID: 31599627	Ceps ⊕	r 🖄 Star 0 💱 Fork 0
ge requests 0	o 22 Campite 12 E Brancher	A There III THE RE Broket Charges	
.D	······································	🗸 V rags 🖂 778 KB Project scorage	
turity & Compliance	Example MEE demo steps		
Deployments		enter la	
rages & Registries	main v mee-	Find file Web lot	
rastructure	A Merge branch 'update	e artifact-script' into 'main'	0787ed7d
tics	piotrpolec authored 1 w	week ago	31016010 13
iki			
pets	README Add LICENS	SE 🗉 Add CHANGELOG 🗄 Add CONTRIBUTING 🔃 Add Kubernetes cluster	Set up CI/CD
gs	Configure Integrations		
	Name	Last commit	Last update
	O_generate_input.job.bash.	Lliquid Remove repository from clone_repo	5 months ago
	🗢 0_generate_input.rb	Generate only 5 random numbers	2 years ago
	1_artifact_sort.job.bash.liq	quid Update 1_artifact_sort.job.bash.liquid	1 week ago
	1_sort.job.bash.liquid	Remove repository from clone_repo	5 months ago
	📥 1_sort.py	Added parentheses to prints in python script	2 years ago
	🖸 2_visualization.bash	Generate gif animation with 5 frames per s	2 years ago
	2_visualization.job.bash.liqu	quid Remove repository from clone_repo	5 months ago

General information		Name :			
Pipeline step points to the Gitlab repository template files are located.	y where computation code and Slurm	Generate animation			~
Citranacitan		Repository *			
MEE uses Git repositories to store scripts	which will be started on the	primageproject/mee/demo-steps			~
Prometheus supercomputer. If you don't h	have a repository already prepared, you	File :			
Clone template repository		2_visualization.job.bash.liquid			~
Cronie template repository					
Git configuration override		Override organization git configuration Default Git configuration is defined on the organization level, but you	can override it and provide step specific	one.	
Grant type srant type is needed to differentiate betwe ind ones that require GPUs.	een calculations that can run on CPUs	Grant type _ cpu			~
Required inputs		Required file types			
Here you can specify what kinds of files an defined in this step. If you cannot find the r	required to start computations required type, you can always create a	Sorting steps ×			
new one.					
New data file type					
Step parameters		Model version parameter			
You can customize computation execution by defining parameters.	Select parameter	Name :		Key *	
Value of the parameter can be injected into slurm starting script. To	String	Model version	~	tag-or-branch	~
see the examples see step details		Mint			

- It is difficult to monitor tasks
- Terminal access to HPC requires some knowledge

Solution

The Model Execution Environment can run simulations on HPC from a web browser. It is integrated with Git repositories, where the simulation source code should be stored. The user has the ability to choose which branch/version to use; then MEE uploads and executes the code. During execution, status is monitored, and result files can be downloaded afterwards. Pipelines can have many steps that can be launched automatically when all required inputs are present, or executed in parallel if possible. MEE also takes care of the underlying file structure, making it easy to upload and download simulation data and results.

Architecture

Fig.1 Git repository is used to store pipeline step code



Fig.2 Patient view with pipelines and their current status (completed; available for execution; missing inputs)

M PRIMAGE				💽 admin admin 🕶
RESEARCH Patients	😂 / 167 / 🏟 demo Demo with number generation (automatic pipeline)		Edit pipe	Time Remove pipeline
Pipelines	⊘ Generate numbers ⊘ Sort numbers ⊘ Generate animation			
PIPELINES' BLUEPRINTS				
Flows	Generate animation finished successfully, results stored in the outputs directory.			
Steps	Computation details	Saved parameter values		
ADMINISTRATION				
Organization's details	Start time 11 Aug 06:36	Model version	master	
Data file types	Revision b5d971ee7511edd5fcffe325f5a9d4fb73ede8fc	Grant	plgprimage4	
Grants	Execution time 00h 00m 28s			
Licenses	Outputs stdout, stderr			
Users	Status Finished			
HELP				
JWT				
Patients API	demo pipeline inputs Choose File No file chosen Upload	demo pipeline outputs		
Pipelines API	Empty	File name Data type	Ac	tions
Pipeline steps				

- Patient virtual space to run different calculations on patient data.
- Flow Blueprint for creating pipelines quickly it represents a collection of steps (or instructions) that comprise the pipeline.
- Step Single instruction or program. Used to assemble flows. Each step can specify what type of input is required to start the calculation.
- Pipeline Simulation created for a given patient. Interfaces HPC and delivers output files. Can specify which version of code will be used. Can be either automatic or manual.
- Model set of scripts/source code stored in a Git repository.

Running simulations

			numbers.bd	Numbers to be sorted	Delete
167 patient inputs	Choose File No file chosen	Upload	steps.bd	Sorting steps	Delete
Empty			plot.gif	Sorting animation	Delete
Notes					

Fig.3 Pipeline view. All details about a specific simulation - its status and the version that was ran. The user can download output files.

Uses so far

- Lung disease classification
- Differential expression profile analysis; variant search
- Forensic genotyping
- Computational fluid dynamics for cardiac simulations
- Cardiac uncertainty analysis
- Cardiac pressure volume loop comparison
- Cardiac reduced order model

🗢 EurValve				Marek Kasztelnik 👻	🏶 EurValve		Marek Kasztelnik 🝷
RESEARCH	Patients			+ Add a new patient	RESEARCH Patients	Set / EURV_S_M_0255_04 /	Edit pipeline Remove pipeline
Pipelines	LURV_S_A_1231_08	Files: 153, Pipelines: 4	Last pipeline: O O O O		Pipelines	O Segmentation Computational Fluid Dynamics simulation Parameter Optimization O D Model x4 O Uncertainty Analysis	
PIPELINES' BLUEPRINTS	LURV_S_A_1257_09	Files: 22, Pipelines: 1	Last pipeline: 🔿 🔗 🔗 🔿		Flows Steps	Computational Fuld Dynamics simulation finished successfully, results stored in the outputs directory.	

- Check if all required inputs are present
- Fetch model starting script template from Git repository
- version specified by user
- Generate model starting script basing on inputs provided by the user
- Submit job to HPC queuing system (Slurm)
- Upload required inputs to HPC
- Monitor job execution
- Fetch results from HPC

ps	LURV_S_M_0140_16	Files: 22, Pipelines: 2	Last pipeline: O O O O	Ŵ	ADMINISTRATION
NISTRATION	& EURV_S_M_0145_19	Files: 44, Pipelines: 1	Last pipeline: 🔿 🖉 🖉 📿		Organization's d
ganization's details	LURV_S_M_0173_11	Files: 23, Pipelines: 1	Last pipeline: O O O O O		Data file types Grants
ta file types	LURV S M 0255 04	Files: 44, Pipelines: 1	Last pipeline: O O O O O		Licenses
ants	* EUDU O 14 0000 10	Files 88 Displace 4	1	-	Users
ers	▲ EURV_5_M_0202_18	Files: 22, Pipelines: 1	Last pipeine: 000000	•	HELP
	LURV_S_M_0363_14	Files: 22, Pipelines: 1	Last pipeline: O O O O O	*	Patients API
т	LURV_S_M_0443_12	Files: 22, Pipelines: 1	Last pipeline: O O O O O	¥	Pipelines API
tients API	LURV_S_M_0447_20	Files: 22, Pipelines: 1	Last pipeline: O O O O O		Pipeline steps
elines API	LURV_S_M_0549_02	Files: 22, Pipelines: 1	Last pipeline: O O O O O		
en e arepa	& EURV_S_M_0737_15	Files: 90, Pipelines: 3	Last pipeline: $\oslash \oslash \oslash \oslash$		
	& EURV_S_M_0776_06	Files: 44, Pipelines: 1	Last pipeline: O O O O O		

Start time	15 Jan 10:03			Model version	master	
Revision	8c7c5b92324f480d22c39196645b455d3b	6c7161				
Execution time	01h 40m 31s					
Outputs	stdouit, stderr					
Status	Finished					
Run1 pipeline inputs	Choose	a file No file chosen	Upload	Run1 pipeline outputs		
Run1 pipeline inputs	Choose	e file No file chosen	Upload	Run1 pipeline outputs	Data type	Actions
Run1 pipeline inputs	Choose	file No file chosen	Upload	Run1 pipeline outputs File name ValveChar.dat	Data type Pressure drops	Actions
Run1 pipeline inputs Empty EURV_S_M_0255_04 patie	Choose ent inputs Choose	e file No file chosen	Upload	Run1 pipeline outputs File name ValveChar.dat ProvFile.bt	Data type Pressure drops Provenance log	Action: Delete
Run1 pipeline inputs Empty EURV_S_M_0255_04 patie File name	Choose ent inputs Choose	e file No file chosen	Upload	Run1 pipeline outputs File name ValveChar.dat ProvFile.bd OutFileGA.csv	Data type Pressure drops Provenance log Parameter optimization result	Actions Delete Delete
Run1 pipeline inputs Empty EURV_S_M_0255_04 patie File name EURV_S_M_0255_04_4B_M0_4cl	Choose ent inputs Choose h_2ch_Zoom_MV_pre_op_003.Trunc.off	e file No file chosen e file No file chosen e file No file chosen Data type Truncated 3D mesh	Upload Upload Actions Delose	Run1 pipeline outputs File name ValveChar.dat ProvFile.bd OutFileGA.csv Vol3.png	Data type Data type Pressure drops Provenance log Parameter optimization result Visual image	Actions Delete Delete

References:

1. M Bubak, K Czechowicz, T Gubała, DR Hose, M Kasztelnik, M Malawski, J Meizner, P Nowakowski, S Wood The EurValve model execution environment, Interface Focus 11 (1), 20200006, 2021

Acknowledgments:

This work was supported by the EU project PRIMAGE H2020-EU.3.1.5.3 826494 and by the PL-Grid Infrastructure (www.plgrid.pl).



PRIMAGE

Medical imaging **Artificial intelligence** Childhood cancer research