

# Linking GitOPS towards fast innovation over BigHPC

Speaker: Samuel Bernardo (LIP) on behalf of BigHPC consortium











- The BigHPC consortium
- Overview of BigHPC platform and development challenges
- Why GitOPS and how to get it in practice
- GitOPS applied to BigHPC
- Next steps

#### **BigHPC consortium**



Austin, USA	Portugal
UT Austin	Wavecom
TACC	INESC TEC
	LIP Austin, USA Portugal
	MACC
	B-PC High PERFORMANCE COMPUTING

# BigHPC platform overview





# Overview of BigHPC development challenges



- Software quality
- Components integration
- Deploy more often
- Easy error recovery
- Keep teams together in same direction
- Share knowledge between teams
- Documented deliveries with complete changes history

#### What is GitOPS





#### Why GitOPS





#### Current status: implemented workflow



**HPC** 

## Review Job log directly from web interface



D Default		lint-test-job Retry
Project information Repository	<pre>ruby:2.5 with digest ruby@sha256:ecc3e4f5da13d881a415c9692bb52d2b85b090f38f4ad99ae94f932b359844 4b </pre>	Duration: 24 seconds
<ul> <li>D Issues</li> <li>D Merge requests</li> <li>0</li> </ul>	<ul> <li>Running on runner-ed2dce3a-project-30845538-concurrent-0 via runner-ed2dce3a-srm-1635421744-8d 13a1cd</li> <li>11 Getting source from Git repository</li> </ul>	Timeout: 1h (from project)       ⑦         Runner: #380987 (ed2dce3a) shared-runners-manager-6.gitlab.com
CI/CD     Pipelines	<pre>12 \$ eval "\$CI_PRE_CLONE_SCRIPT" 13 Fetching changes with git depth set to 50 14 Initialized empty Git repository in /builds/bighpc/ci-cd/pipeline-templates/default/.git/</pre>	<b>Commit</b> fab3eb66 fa Update README.md
Editor Jobs Schedules	15 Created fresh repository. 16 Checking out fab3eb66 as main 17 Skipping Git submodules setup ✓ 19 Executing "step script" stage of the job script 000:10	O Pipeline #397447084 for main <sup>™</sup>
<ul><li>♥ Security &amp; Compliance</li><li>● Deployments</li></ul>	20 Using docker image sha256:27d049ce98db4e55ddfaec6cd98c7c9cfd195bc7e994493776959db33522383b for ruby:2.5 with digest ruby@sha256:ecc3e4f5da13d881a415c9692bb52d2b85b090f38f4ad99ae94f932b359844 4b	→ ⊙lint-test-job
<ul><li>Monitor</li><li>Infrastructure</li></ul>	<pre>21 \$ echo "Linting code This will take about 10 seconds." 22 Linting code This will take about 10 seconds. 23 \$ sleep 10</pre>	⊘ security-dev-job
Packages & Registries In Analytics	24 \$ echo "No lint issues found."         25 No lint issues found.         ✓ 27 Cleaning up project directory and file based variables         00:01	⊘ unit-test-job
« Collanse sidehar	29 Job Succeeded	

# Attach the Environment with the Infrastructure





- GitOPS push based strategy is implemented using Gitlab Agent
- agentk communicates to the Gitlab Agent Server (KAS) to perform GitOPS operations
- Changes applied to Kubernetes cluster anytime developer changes a manifest file managed within Gitlab
- This approach keeps Gitlab platform isolated from real infrastructure

## GitOPS aware of infrastructure state





- Argo CD provides a web user interface that allows to check application real time activity
- Health status analysis of application resources

#### Next steps



- GitOPS insight
  - Improve GitOPS implementation, extending Gitlab CI/CD agent with kubernetes
  - Improve team collaboration using Gitlab platform
  - Test deployment error recovery taking the advantage of git
- Pilot
  - Test platform components with real use cases
  - Identify application performance issues





#### **Partners:**



	. 11		1	 	Minho
1111	1111	11 1	1 11	 	
11 111		11 1	1 11	11	Advanced
11 11			1 11	11	Computing
11	11	11 1	i ii	 -ii II	computing
	- 11		1 I	 	Center









#### Funding:













#### User and administrator interfaces

User and ac	Iministrato	r interfa	aces				B B C	PC
Jupyter <mark>hub</mark>								<b>5</b> ⊖
≡ Jupyter	Jupyter   admin's server	r 💽 On						
El Admin 4	Files   Running							ı.
	A / Select items to perform actions or	a them.				2 Upload	3 New → 2	
(1)	■   ▼ Name \$			Status	Last Modified 🖨			
X +	@Recycle				3 months ago			
SAR	jupyter_example	•			21 days ago			
Jupyter <mark>hub</mark>								Ģ
⊜ Jupyter	Server Management							
🛱 Admin	Jupyterhub created a server for each NA	S user. You can view the use	r list and manage their server stat	us.				
	Stop all servers   All Server : 5							
	Username <sup>•</sup>	Description	Last Seen ≑	Server Status 🖨		Action		
	admin	Admin	a few seconds ago	On On		Access Server		
	anir	User	7 hours ago	Off Off				

#### Manage code and do debugging



exam	ple			
File I	Edit View Insert	Cell Kernel Help	Not Trusted Python 3 (GPU) 🔿	
-	× 42 B ↑ ↓	Run Cells		
		Run Cells and Select Below		
		Run Cells and Insert Below	NN GPU Version	
	# coding=utf-8	Run All		
		Run All Above		
	This demo is run on GF 	Run All Below	, please unmark the following two lines	
		Cell Type		
		Current Outputs		6
	import keras from keras.utils impor	All Output		
	from keras.models impo	ort Sequential		- A 1
	from keras.layers impo from keras.datasets im	port Dense, Activation, Convo uport mnist	lution2D, MaxPooling2D, Flatten, Conv2D	$\times$ NV.
	Using TensorFlow backend. /usr/local/Cellar/python3/3.6	5.3/Frameworks/Python.framework/Vers	ions/3.6/lib/python3.6/importlib/_bootstrap.py:219: RuntimeWarning: compileti	

return f(\*args, \*\*kwd

#### OnDemand: extend access to applications









- Its inception in 2017 by Weaveworks
- Uses developer common tools:
  - git version control system to track code changes
  - continuous deployment tools
- Have a git repository that contains infrastructure declarative description
- Production environment match the described state in the repository
- Deployment and updates means push changes to a git repository





- Always use same procedure to deploy applications without the burden of switching between required tools
- Lighter learning curve since everything happens on git with less effort
- Error recovery as easy as git revert
- Easier credential management since only required access to git repository
- No need to give developers direct access to the deployed endpoints
- Complete description of what is deployed
  - every changes goes through the git repository
  - complete history of every change made to the system
- Share knowledge with great commit messages where everybody can reproduce and find examples how to set up new systems

#### Deployment strategy: push based





- Implemented by Gitlab CI/CD, Jenkins, Circle CI or Travis CI
- Code updates trigger pipeline execution (external event)

- Environment configuration repository is updated with new deployment descriptors
- misses automatic check of environment desired state

## Deployment strategy: pull based





- Operator is introduced comparing with previous strategy
- Implemented by Argo CD, Flux
   CD and Gitlab starting from v15

 Operator checks consistency between deployed infrastructure, environment repository and image registry

# Multiple Applications and Environments





- GitOPS also support multiple build pipelines from different repositories
- Changes in each project generate an update to environment repository
- Multiple environments with GitOPS means separate branches in environment repository
- Operator will react on changes and run the associated environment pipeline

#### BigHPC Gitlab platform adoption



- Adopting GitOPS using Gitlab platform
  - Implement continuous deployment applying software quality best practices
  - Well documented deliveries with complete history of system changes
  - Promote shared knowledge and get teams together
  - Easy to reproduce the thought process of changing infrastructure
  - Get the examples to setup new systems

#### **BigHPC environments**



- Testbeds and Pilot
  - Review the hardware profiles and application requirements
  - Define execution environments for the required testbeds
  - Gather applications that covers the demanded use case requirements
  - Prepare the pilot to test platform components with collected applications over HPC
  - Collect the metrics and get the report from jobs submission

#### Current status: team collaboration



« Collanse sidebar

### Current status: team collaboration





submit workload request to the scheduler...

Commented on issue #1 "Platform evaluation using real applications" at BigHI

Packages & Registries

Settings

#### Gitlab stages for the project

- Development testbed
- Preview testbed
- Delivery to production (create a release)

Build	Development	Preview	Production
build-job	unit-test-job	integration-test-job	deploy-job
	lint-test-job	functional-test-job	
	security-dev-job	security-prev-job	



#### Create the pipeline configuration



	🔶 GitLab 🛛 ≡ Menu		Search GitLab Q	D	11 ~	$\square$	?•∽	
D	Default							-
-	Droject information	Browse templates						-
U	Project Information	14 # For mo	re information, see: <a href="https://docs.gitlab.com/ee/ci/yaml/index.html#stage">https://docs.gitlab.com/ee/ci/yaml/index.html#stage</a>	S				-
E	Repository	15						
	Issues	16 stages:	# List of stages for jobs, and their order of execution d					
-			lopment					
3	Merge requests 0	19 - prev	iew					
de		20 - prod	uction					
2	CITCD	21						
	Pipelines	22						
	Editory	23 ########						
	Editor	24 # BUILd	stage #					
	Jobs	25 <i>########</i> 26	****					
	Calculate	27 build-jo	b:					
	Schedules	28 stage:	build					
J	Security & Compliance	29 script						
_		30 - ec	ho "Compiling the code"					
Q	Deployments	31     - ec	ho "Compile complete."					
ų	Monitor	32						
~		34 ########	****					
ŝ	P Infrastructure	35 <i># Develo</i>	pment stage #					
-		36 ########	, <i>***********</i> **					
4	Collapse sidebar	37						
	compse sidebui	38 unit-tes	t-job:					

#### Developers integration task



- Create a new project in Gitlab
- Add pipeline from template



#### Check pipeline results



#### Click over status of pipeline after passing the tests

D Default	💌 BigHPC > 🚥	> Pipeline Template	es > Default > I	Pipelines			
Project information	All 4 Fin	ished Branche	es Tags			Clear runner caches	CI lint Run pipeline
Repository							
D Issues	Filter pipeli	nes				Q	Show Pipeline ID 🐱
រាំ Merge requests 🛛 💿							
🥠 CI/CD	Status	Pipeline ID	Triggerer	Commit	Stages	Duration	
Pipelines		#397447084	æ	<b>⊬ main </b> ↔ fab3eb66	~ ~ ~ ~		
Editor	⊘ passed	latest	6899 1	Update README.md	${}$	台 7 minutes ago	
Jobs							
Schedules	⊘ passed	#397446186	*	♥ main - 3695b7e1 ₩ Update .gitlab-ci.vml f	$\textcircled{\begin{tabular}{c} \hline \end{tabular}}$	00:04:04 〇 8 minutes ago	
👽 Security & Compliance		7				_ 0,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,	
	K	$\searrow$					

## Look into pipeline stages



D Default		AND STREET		
<b>1</b> Project information	U 8 jobs for main in 4 minutes	and 2 seconds		
	P latest			
Merge requests	-			
· CI/CD	ា No related merge requests f	ound.		
Pipelines				
Editor	Pipeline Needs Jobs 8 Tes	ts 0		
Jobs				
Schedules	Build	Development	Preview	Production
Security & Compliance				
ල Deployments	build-job	V lint-test-job	functional-t	deploy-job
🖳 Monitor		Security-de	integration	
✤ Infrastructure				
Packages & Registries		Unit-test-Job	security-pre	
🔟 Analytics				

# Check jobs state



		M			
Default	Pipeline Needs Jo	obs 8 Tests 0			
Project information	Status	Name	Job ID		Coverage
Repository	⊘ Build				
Issues   0     Merge requests   0	⊘ passed	build-job	#1725216939	⊘ 00:00:11 首 13 minutes ago	
CI/CD	-				
Pipelines	⊘ Development				
Editor Jobs	⊘ passed	lint-test-job	#1725216943	ō 00:00:24 台 13 minutes ago	
Schedules 7 Security & Compliance	(⊘ passed	security-dev-job	#1725216944	ō 00:01:12 音 12 minutes ago	
Deployments	(⊙ passed	unit-test-job	#1725216940	ō 00:01:24 音 12 minutes ago	
Infrastructure					
Packages & Registries	<ul> <li>Preview</li> </ul>				
<b>«</b> Collapse sidebar	⊘ passed	functional-test-job	#1725216947	♂ 00:01:13 럼 11 minutes ago	

# Attach the Environment with the Infrastructure



X Snippets

Wiki

Settings

General

Integrations

Webhooks

Repository

CI/CD

Monitor

Pages

Packages & Registries

Usage Quotas

≪ Collapse sidebar

- Gitlab Runners are agents that run CI/CD jobs from Gitlab
- Gitlab Runner implements executors used to run builds in different environments
- Environments comprehends different Operating Systems and available tools, such as docker and kubernetes
  - Each Gitlab Runner can have different access policies and can be limited to pipelines asking for specific tag

#### **Group runners**



#### Available group runners: 2

● #13935088 (UJU3fxiw) 合 bighpc runner on bighpc vm via docker bighpc-runner-02

● #13934510 (VL9E8yys) 合 Runner instaled on bighpc vm via rpm bighpc-runner-01

## Attach the Environment with the Infrastructure



### GitOPS aware of infrastructure state





- Argo CD implements the GitOPS pull based strategy
- This not only follows the repositories activity, but also the deployment state in the infrastructure
- Webhook integration with Gitlab
- Support of multiple config management and templating tools (Kustomize, Helm, ...)

#### Background execution behind the scene



- Develop Gitlab CI/CD configuration for the required pipelines
- Create a Gitlab project and submit the code for the defined testbeds
- Connect Gitlab project with Gitlab runner instance, providing the reports in Gitlab
- Create docker images to pack the required code
- Do the required experimental validation checking the automated results over the testbeds (development and preview)
- Deploy the release to the production infrastructure
- Present the Pilot for further validation











