

Cloud Native AI on Openshift

ML Prague ~~2020~~ 2021

Václav Pavlín

AI Services, Red Hat

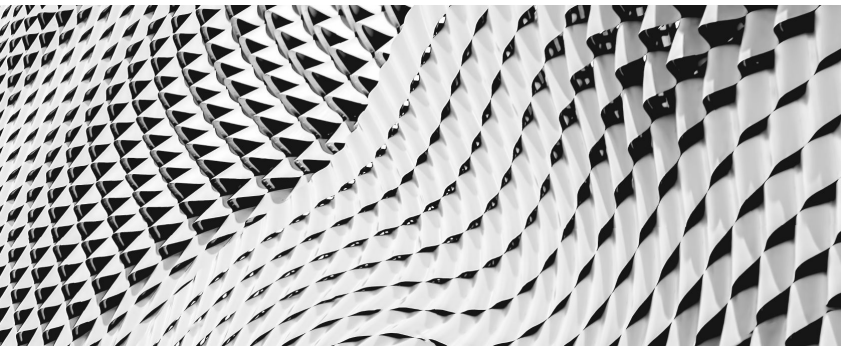
Francesco Murdaca

AI CoE, Red Hat

- Cloud Native AI
- Open Data Hub
- What is Project Thoth
- Thoth Integrations
- Operate First

Red Hat & AI

Where does Red Hat fit in
the AI story?

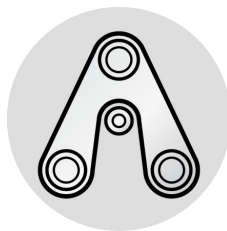


How Red Hat Sees AI



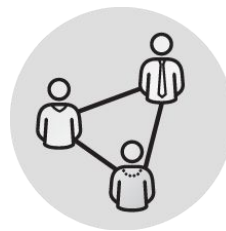
AI as a Workload

Represents a workload requirement for our **platforms** across the hybrid cloud.



AI Accelerates Open Source Development

Applicable to Red Hat's existing core business in order to increase **Open Source** development and production **efficiency**.



AI Enhances RH Products & Services

Valuable to our customers as specific services and product capabilities, providing an **Intelligent Platform** experience.



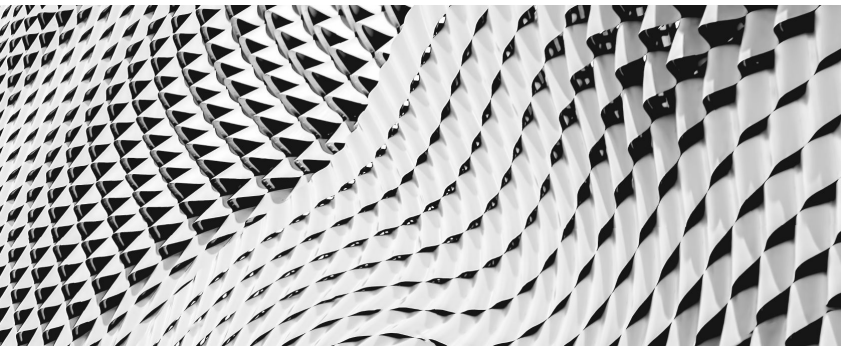
AI Enables Customers

to build **Intelligent Apps** using Red Hat products as well as our broader partner ecosystem.

010110
101010

Data as the Foundation

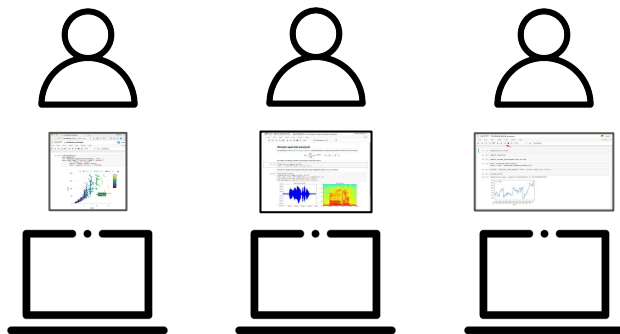
It's Better to Share



Easy **sharing** is key when
it comes to data science.

One Person Show

Working on a personal infrastructure brings complications



Look what I made!

Sharing results of the work might be harder if all the work has been done on your personal machine.

Can I get more ...

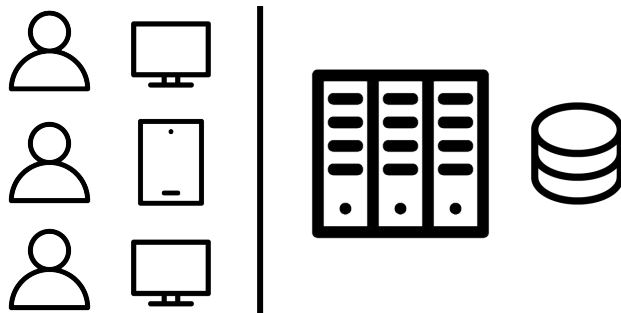
Personal laptops and desktops tend to suffer from the lack of resources when it comes to big data and AI

Works on my machine

Productizing the work can become complicated due to the differences between dev and prod environments

Shared Infrastructure

Providing shared infrastructure often helps



Collaboration for the win

Collaboration gets a boost when sharing is made easy.

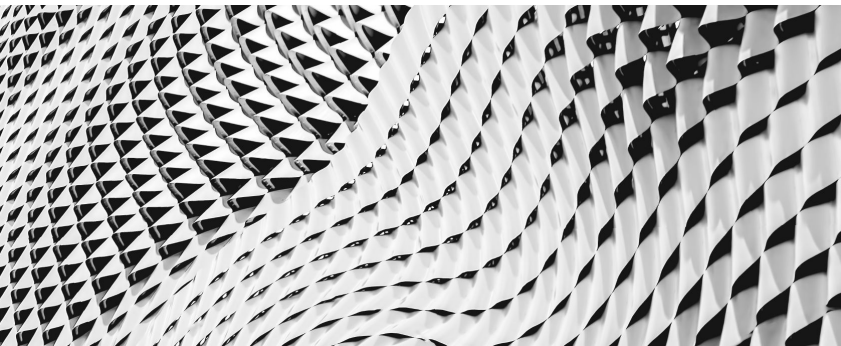
Fluent resource reallocation

Shared infrastructure can serve many teams and projects and resources can be allocated based on real needs.

Product level tools for everyone!

Whether development, testing or production, the tools should not differ too much.

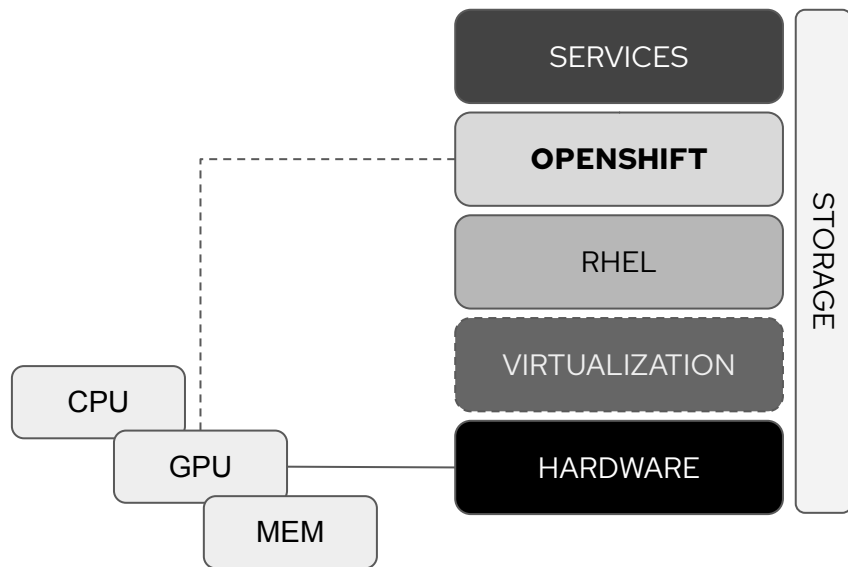
HW & SW Stack Integration



Solid **integration** of all the components in the stack is fundamental.

All the Layers in the Stack Matter

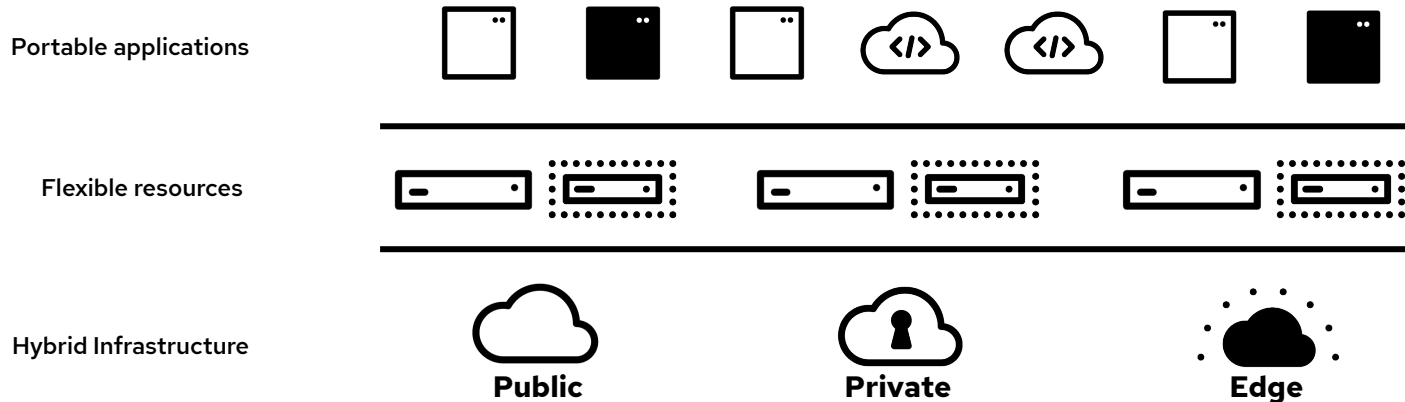
It's like an operating system, but for the cloud



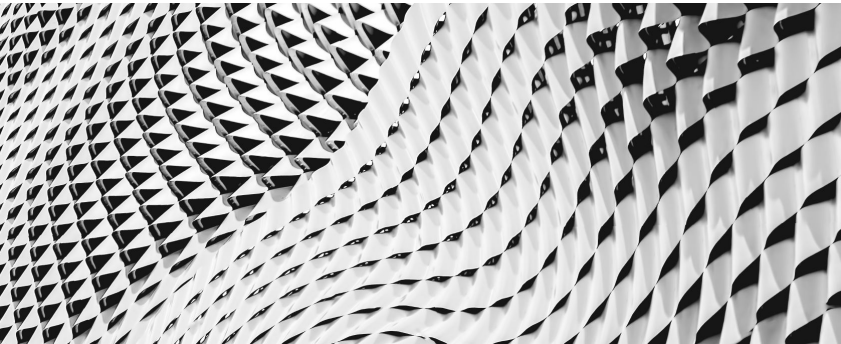
From hardware to the custom applications, all the layers have to cooperate to provide cohesive and integrated experience and seamlessly enable all the (scarce) resources through the whole stack.

Anything Anywhere

An application should run virtually on any hardware, whether it's public, private or at the edge.



Automation is King



Automation provides time and energy for the work that business successful.

Operators

Running a service on OpenShift is just a click away with OperatorHub

"The goal of an Operator is to put operational knowledge into software."



Partner Ecosystem

"It's better to share" applies to automation as well

Wide ecosystem of partners providing easily installable and manageable tools.

AI/Machine Learning

Application Runtime
Big Data
Cloud Provider
Database
Developer Tools
Integration & Delivery
Logging & Tracing
Monitoring
Networking
OpenShift Optional
Security
Storage
Streaming & Messaging

Install State
☐ Installed (2)
☐ Not Installed (32)

Provider Type
☐ Red Hat (0)
☐ Certified (8)
☐ Community (5)
☐ Marketplace (7)

Provider
☐ Red Hat (1)
☐ Ansible (2)
☐ Ansible (2)
☐ Ansible (2)

Show 132 more

Capability Level
☐ Basic Install (16)
☐ Seamless Upgrades (14)
☐ Full Lifecycle (2)
☐ Deep Insights (0)
☐ Auto Pilot (4)

Filter by keyword...

Marketplace Anaconda Team Edition provided by Anaconda, Inc. Operator for Anaconda Team Edition	Marketplace Anaconda Team Edition provided by Anaconda, Inc. Operator for Anaconda Team Edition	cnvrg.io cnvrg-operator provided by cnvrg.io Operator for deploying Cnvr AI OS	Marketplace cnvr.io cnvr-operator provided by cnvr.io Operator for deploying Cnvr AI OS	Marketplace Cortex Certifi Operator provided by CognitiveScale Cortex Certifi empowers enterprises to identify and mitigate risk and vulnerabilities...	Marketplace Cortex Certifi Operator provided by CognitiveScale Cortex Certifi empowers enterprises to identify and mitigate risk and vulnerabilities...
Marketplace Cortex Fabric provided by CognitiveScale Collaborative platform for building, deploying, and managing trusted AI systems.	Marketplace Cortex Fabric provided by CognitiveScale Collaborative platform for building, deploying, and managing trusted AI systems.	Marketplace Cortex Hub Operator provided by CognitiveScale Discover and manage high quality AI building blocks to speed up development, deployment and...	Marketplace Dotsi Deployment Operator provided by Dotsi ML model deployment operator	Marketplace Driverless AI Operator provided by H2O Operator that allows users to spin up a Driverless AI server in OpenShift environment	Marketplace EDDI Operator provided by LASS AI Basic installation of EDDI, upgrade to new versions
Marketplace EDDI Operator provided by LASS AI Basic installation of EDDI, upgrade to new versions	Marketplace Fedorator provided by Prophetor Data Services, Inc. Federator Operator provides easy configuration and management of AI-based...	Community Fedorator provided by Prophetor Data Services, Inc. Federator Operator provides easy configuration and management of AI-based...	Marketplace Fedorator provided by Prophetor Data Services, Inc. Federator Operator provides easy configuration and management of AI-based...	Marketplace FP Predict Plus Operator provided by Feasibility Sciences, LLC This operator will deploy FP-Predict-Plus in the cluster.	Marketplace FP Predict Plus Operator provided by Feasibility Sciences, LLC This operator will deploy FP-Predict-Plus in the cluster.
Marketplace InsightEdge Enterprise Operator provided by Gigaspaces InsightEdge is an in-memory real-time analytics platform for instant insights to action. It's an always-on...	Marketplace InsightEdge Enterprise Operator provided by Gigaspaces InsightEdge is an in-memory real-time analytics platform for instant insights to action. It's an always-on...	Community NVIDIA GPU Operator provided by NVIDIA Corporation Automate the management and monitoring of NVIDIA GPUs.	Community Open Data Hub Operator provided by Open Data Hub Open Data Hub Operator for...	Marketplace PerceptiLabs Operator provided by PerceptiLabs AI platform which lets you Build, Train and Analyze	Marketplace PerceptiLabs Operator provided by PerceptiLabs AI platform which lets you Build, Train and Analyze

Administrator

Home

Overview

Projects

Search

Explore

Events

Operators

OperatorHub

Installed Operators

Workloads

Pods

Deployments

Deployment Configs

Stateful Sets

Secrets

Config Maps

Cron Jobs

Jobs

Daemon Sets

Replica Sets

Replication Controllers

Horizontal Pod Autoscalers

Networking

Storage

Builds

Project: opendatahub

OperatorHub

Discover Operators from the Kubernetes community and Red Hat partners, curated by Red Hat. You can purchase commercial software through [Red Hat Marketplace](#). You can install Operators on your clusters to provide optional add-ons and shared services to your developers. After installation, the Operator capabilities will appear in the [Developer Catalog](#) providing a self-service experience.

All Items

AI/Machine Learning

AI/Machine Learning

Filter by keyword...

42 items

Application Runtime

Big Data

Cloud Provider

Database

Developer Tools

Integration & Delivery

Logging & Tracing

Monitoring

Networking

OpenShift Optional

Security

Storage

Streaming & Messaging

Install State

- ☐ Installed (0)
- ☐ Not Installed (42)

Provider Type

- ☐ Red Hat (0)
- ☐ Certified (19)
- ☐ Community (6)
- ☐ Marketplace (16)
- ☐ vpvain-operators (1)

Provider

- ☐ Red Hat (1)
- ☐ 2ndQuadrant (0)
- ☐ Alcid (0)

Marketplace



Anaconda Team Edition
provided by Anaconda, Inc.

Operator for Anaconda Team Edition



Anaconda Team Edition
provided by Anaconda, Inc.

Operator for Anaconda Team Edition



can-operator
provided by Avanseus Technologies Pvt. Ltd.

Avanseus Cognitive Assistant for Networks (CAN), applies AI and machine learning to predict whe...



can-operator
provided by Avanseus Technologies Pvt. Ltd.

Avanseus Cognitive Assistant for Networks (CAN), applies AI and machine learning to predict whe...

cnvrg.io

cnvrg-operator
provided by cnvrg.io/new

Operator for deploying Cnvrg AI OS

Marketplace

cnvrg.io

cnvrg-operator
provided by cnvrg.io/new

Operator for deploying Cnvrg AI OS



Cortex Certifai Operator
provided by CognitiveScale

Cortex Certifai empowers enterprises to identify and mitigate risk and vulnerabilities...



Cortex Certifai Operator
provided by CognitiveScale

Cortex Certifai empowers enterprises to identify and mitigate risk and vulnerabilities...



Cortex Fabric
provided by CognitiveScale

Collaborative platform for building, deploying, and managing trusted AI systems.



Cortex Fabric
provided by CognitiveScale

Collaborative platform for building, deploying, and managing trusted AI systems.



Cortex Hub Operator
provided by CognitiveScale

Discover and manage high quality AI building blocks to speed up development, deployment and...



Cortex Hub Operator
provided by CognitiveScale

Discover and manage high quality AI building blocks to speed up development, deployment and...



Driverless AI Operator
provided by H2O.ai

Operator that allows users to spin up a Driverless AI server in OpenShift environment



EDDI Operator
provided by LABS.AI

Basic Installation of EDDI, upgrade to new versions



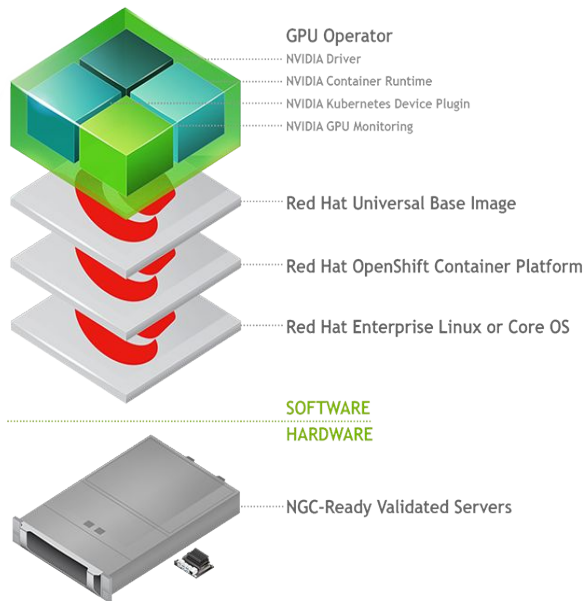
EDDI Operator
provided by LABS.AI

Basic Installation of EDDI, upgrade to new versions

Marketplace

Adding Special HW

Enabling expensive hardware has to be simple and fast

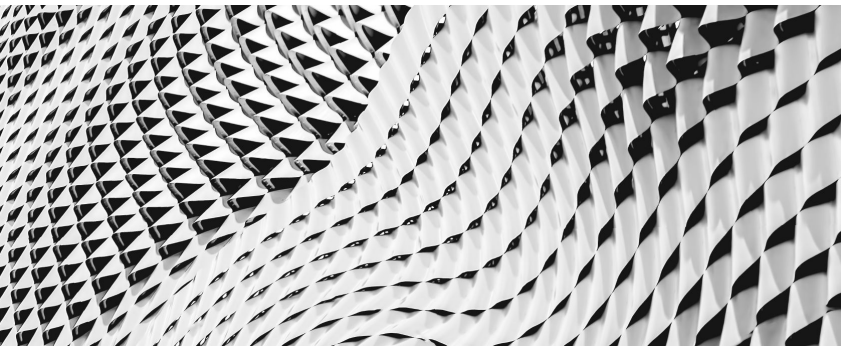


e.g. GPUs have to be enabled in operating system, pass through virtualization, be recognized by orchestration and assigned to enabled containers for applications to use them.

“An automation” should take care of it all.

Open Data Hub

The open source
As-as-a-service platform
to bring it all together.



AI-as-a-Service platform on OpenShift

The best of what we talked about packaged as a blueprint



Open Data Hub

Open Source

ODH is an open source community project leveraging popular AI/ML tools and enabling them on OpenShift

Meta-Operator

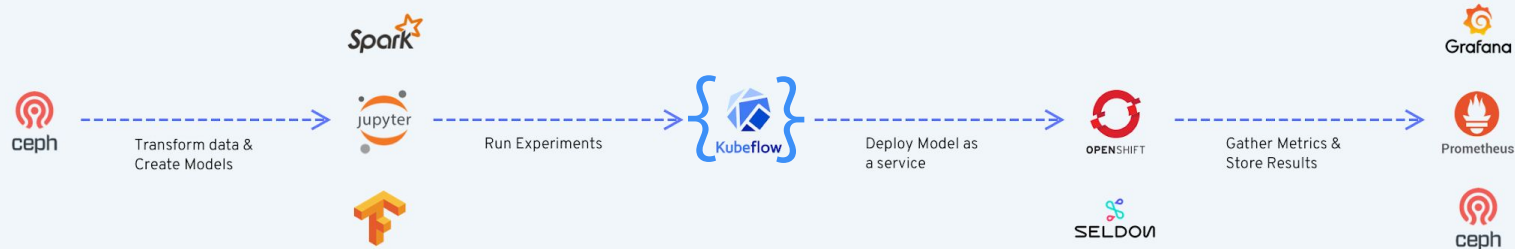
Manages the components through the operators

Production-Ready

Runs at Red Hat as a production service used by hundreds of users

The Flow

From exploration, through experimentation to deployment and evaluation



Leverage

ODH provides default configurations which work fine for PoCs and exploring the platform.

Modify

There is plenty to be customized based on specific use cases and requirements

Bring your own

There is too many use cases and technologies - integrate those that are important for you, or contribute them.

Learn more



opendatahub.io

Community

Join our [community meetings](#) and our [repositories](#).

Videos

Watch our previous [presentations and demos](#).

Docs

Learn more about Open Data Hub from our [documentation](#).

Project Thoth

ML Prague 2021

Project Thoth



Project Thoth

- Help developers in the selection of dependencies for their applications depending on their requirements
- Use bots to automate mundane work to offload humans work

what Thoth stores in the knowledge graph

Application Stack related:

- Buildtime and runtime environment
- Dependencies
- Performances

Software Package:

- Application Binary Interfaces (ABI)
- Security (CVE, analyzers)

Source Code Meta Information:

- Project features (TTR, TTCL, etc,..) from different software development platform (Github, GitLab, Pagure, etc...)

what Thoth stores in the knowledge graph

Application Stack related:

- Buildtime and runtime environment
- Dependencies
- Performances

Thoth solver

Analyze each Python package version

- Get its dependencies by installing it
- Aggregate information about possible installation errors
 - Analysis of installation logs
 - Package was not installed due to..

what Thoth stores in the knowledge graph

Application Stack related:

- Buildtime and runtime environment
- Dependencies
- Performances

Thoth Dependency Monkey and Amun

- Creates combination of software stacks and create inputs for Amun
- Run Performance Indicator to gather performance observations:
 - ML frameworks
 - Python Interpreter

what Thoth stores in the knowledge graph

Package-extract

- Extract dependencies from an image and image build logs

Software Package:

- Application Binary Interfaces (ABI)
- Security (CVE, analyzers)

CVE Update job

- checks for new CVE records in the safety-db provided by PyUp.io.
- store them in the database

Security workflow

- SI analyzers (SI-bandit, SI-cloc)

what Thoth stores in the knowledge graph

Src Ops Metrics

- Gather information from software development platforms:
 - GitHub
 - GitLab
 - Pagure

Source Code Meta Information:

- Project features (TTR, TTCL, etc,..) from different software development platform (Github, GitLab, Pagure, etc...)

How do we use this knowledge?

- Recommender system is called Adviser in Thoth.
- It uses Reinforcement Learning (RL).

Thoth Recommendation types

- Latest
- Stable
- Security
- Performance
- Testing

.thoth.yaml

```
host: {THOTH_SERVICE_HOST}
tls_verify: true
requirements_format: {requirements_format}

runtime_environments:
- name: '{os_name}:{os_version}'
  operating_system:
    name: {os_name}
    version: '{os_version}'
  hardware:
    cpu_family: {cpu_family}
    cpu_model: {cpu_model}
  python_version: '{python_version}'
  cuda_version: {cuda_version}
  recommendation_type: stable
  platform: '{platform}'
```

.thoth.yaml

```
host: {THOTH_SERVICE_HOST}
tls_verify: true
requirements_format: pipenv

runtime_environments:
  - name: 'cuda' # <<<
    operating_system:
      name: fedora
      version: '32'
    hardware:
      cpu_family: 6
      cpu_model: 94
    python_version: '3.8'
    cuda_version: '10.1' # <<<
    recommendation_type: stable
    platform: 'linux-x86_64'

  - name: 'no_cuda' # <<<
    operating_system:
      name: fedora
      version: '32'
    hardware:
      cpu_family: 6
      cpu_model: 94
    python_version: '3.8'
    cuda_version: null # <<<
    recommendation_type: stable
    platform: 'linux-x86_64'
```

.thoth.yaml

```
host: {THOTH_SERVICE_HOST}  
tls_verify: true  
requirements_format: pipenv  
overlays_dir: overlays
```

```
runtime_environments:  
  - name: 'fedora:33'  
    operating_system:  
      name: fedora  
      version: '33'  
      python_version: '3.8'  
  
  - name: 'ubi:8'  
    operating_system:  
      name: rhel  
      version: '8'  
      python_version: '3.8'
```

```
.  
├── app.py  
├── overlays  
│   ├── fedora:33  
│   │   ├── Pipfile  
│   │   └── Pipfile.lock  
│   └── ubi:8  
│       ├── Pipfile  
│       └── Pipfile.lock  
└── .thoth.yaml
```


How Thoth can help developers?

- Keep dependencies up to date.
- Maintain software stack secure avoiding CVE.
- Recommend the most performant software stacks for AI Apps.
- Integrate source metadata information related to the packages used in the software stack to give advice to users.
- Integrate Thoth in day-to-day developers/data scientists tools.
- ...

Thoth Integrations

Thoth Integrations

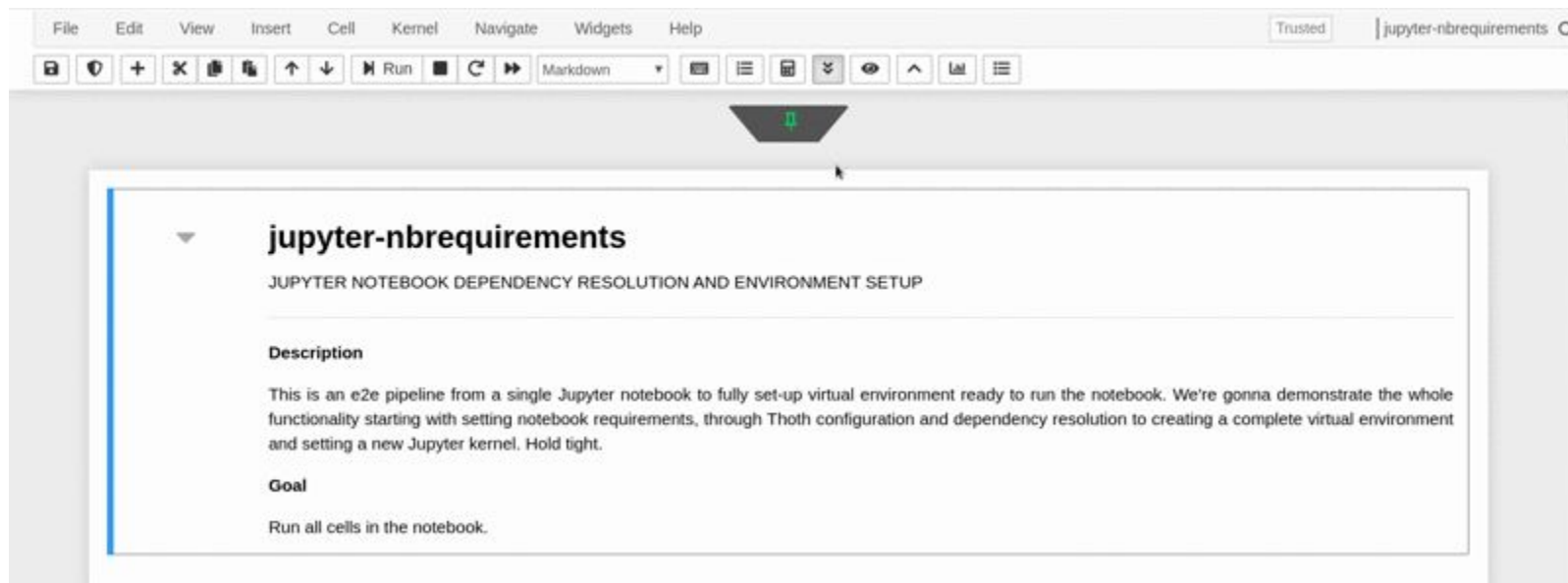
- Command line tool thamos (developer laptop)

~ `pip install thamos`

~ `thamos advise`

Thoth Integrations

- Jupyter Tools (jupyter-nbrequirements)



Thoth Integrations


- Jupyter Tools (jupyterlab-requirements)



Thoth Integrations

- GitHub App Qeb-Hwt (CI pipeline check)

Adjust link #7

 Merged sesheta merged 1 commit into `thoth-station:master` from `pacospace:remove-self` on Jun 25

 Conversation 2

 Commits 1

 Checks 1

 Files changed 1

 Adjust link 2910711 ▾

▾ Qeb-Hwt

▾ Thoth: Advise (Developer Previe...

Qeb-Hwt / Thoth: Advise (Developer Preview)

succeeded on Jun 25 in 2m 12s

Thoth's Advise

Thoth's adviser finished with conclusion: 'success'

Justification:

Report

Justifications

	message	type
0	No observations spotted for this stack in Thoth's database	INFO

See the document below for more details.

Thoth Integrations

- Cyborg Kebechet (pull request/issues creator)

Automatic update of dependencies by kebechet. #526

 Merged **sesheta** merged 1 commit into `master` from `kebechet-automatic-update` 3 days ago

 Conversation 2  Commits 1  Checks 1  Files changed 1



khebhut bot commented 3 days ago

Contributor  ...

Kebechet has updated the dependencies to the latest version 🚀
The direct dependencies updated in the pull request are -

Package Name	Old Version	Updated Version	Is Dev
thoth-storages	0.29.2	0.29.3	False

Kebechet Version: 1.1.4

Thoth Integrations

- Source-to-Image (container builder)

The screenshot shows the GitHub repository page for `thoth-station/s2i-thoth`. The repository is in the `master` branch, has 2 branches, and 54 tags. The repository description is "Thoth's addition to OpenShift's s2i Python builds to benefit from Thoth's recommendations in your application". The repository is licensed under the GPL-3.0 License. The repository has 189 commits and 54 releases. The repository is categorized under tags: openshift, s2i, thoth, software-stack, ai, python, python3, and hacktoberfest. The repository has a README and a CHANGELOG.md file. The repository is maintained by fridex, who provided hints on how to suppress s2i behavior (#102) 6e5bac0 5 days ago. The repository has 189 commits. The repository has 54 releases, with the latest release being v0.14.0:sparkles: base ima... on Jul 13. The repository has 53 releases in total.

thoth-station / s2i-thoth

Unwatch 5

<> Code 1 Issues Pull requests Actions Security Insights Settings

master 2 branches 54 tags Go to file Add file Code

fridex Provide hints on how to suppress s2i behavior (#102) 6e5bac0 5 days ago 189 commits

.github	added a 'tekton trigger tag_release pipeline issue'	6 months ago
f31-py37	Version 0.22.0 (#99)	7 days ago
f32-py38	Version 0.22.0 (#99)	7 days ago
ubi8-py36	Version 0.22.0 (#99)	7 days ago
ubi8-py38	Version 0.22.0 (#99)	7 days ago
.gitignore	added Fedora32 Python38 s2i-thoth	7 months ago
.thoth.yaml	Add section for Thoth runtime environment to .thoth.yaml	10 months ago
CHANGELOG.md	new release due to #27	13 months ago
LICENSE	Initial commit	16 months ago

About

Thoth's addition to OpenShift's s2i Python builds to benefit from Thoth's recommendations in your application

openshift s2i thoth software-stack ai python python3 hacktoberfest

Readme

GPL-3.0 License

Releases 54

v0.14.0:sparkles: base ima... Latest on Jul 13

+ 53 releases

Thoth Integrations

- Reproducible Deployment Pipeline
- Pipelines for reproducible builds
- Optimizing Deployment Pipeline

Thoth Integrations

- Command line tool thamos (developer laptop)
- Jupyter Tools (data scientist browser)
- GitHub App Qeb-Hwt (CI pipeline check)
- Cyborg Kebechet (pull request/issues creator)
- Source-to-Image (container builder)
- Optimizing Deployment Pipeline

- [Google Chat](#)
- **Thoth Station YouTube**
 - https://www.youtube.com/channel/UCIUIDuq_hQ6vlzmqM59B2Lw/videos
- **Website:**
 - <https://thoth-station.ninja/>
- **Twitter**
 - <https://twitter.com/thothstation>
- **GitHub**
 - <https://github.com/thoth-station>

Operate First

ML Prague 2021



“Operate First is an initiative to operate software in a production-grade environment - bringing users, developers and operators closer together.

Ideally Operate First becomes a partner to Upstream First as a basic tenet of our workflow.”

Operate First Is The Solution

With the Mass Open Cloud and Open Infra Labs, Red Hat is launching an effort to open source cloud operations at scale.



Upstream Projects Operate First

We will open the MOC to upstream communities who need a place to operate their services in order to develop them.

Red Hat Products Operate First

We will begin operating our own products on the MOC in the open, before we ship them to our customers.

Open Telemetry, Open Tracing, Open Ops

We will work with the community around the MOC to develop new tools in the open that will be the key to open, autonomous operations at scale.



CLOUD with full
visibility into the
operations center

The Power of Open Source

Turn Users into Contributors



Read Only Access to **all** the Data

Easy onboarding

Read → Report → Resolve

OpenDataHub

Operate First

Common history with Operate First Team

Young project

Users, Users, Users

e.g. ACM/OCP/M4D/IBM/CNV are other workloads



Call to Action

Community means collaboration



Access All Areas

Deploy Demos

Solve Issues

operate-first.cloud

- **Operate First:**
 - <https://www.operate-first.cloud/>
- **MOC**
 - <https://massopen.cloud/>

Open Reusable Workshop

ML Prague 2021

Thank you

Red Hat is the world's leading provider of
enterprise open source software solutions.
Award-winning support, training, and consulting
services make
Red Hat a trusted adviser to the Fortune 500.



linkedin.com/company/red-hat



youtube.com/user/RedHatVideos



facebook.com/redhatinc



twitter.com/RedHat