

# AI and Scientific Research Computing with Kubernetes Software Stack, Applications, and Services

A tutorial at PEARC24

July 22, Providence, Rhode Island

Presented by Mahidhar Tatineni and Dmitry Mishin  
University of California San Diego – San Diego Supercomputer Center

Ref: Tutorials at PEARC, SC, 5NRP by Igor Sfiligoi, Dmitry Mishin, and Mahidhar Tatineni



# docker

## Find your container images

Explore Docker's Container Images

hub.docker.com/search?q=PyTorch&image\_filter=official%2Cstore

docker hub

PyTorch

Sign In Sign up

Filters (2) [Clear All](#)

Products

- ☐ Images
- ☐ Extensions
- ☐ Plugins

Trusted Content

- ☒ Docker Official Image
- ☒ Verified Publisher
- ☐ Sponsored OSS

Categories

- ☐ API Management
- ☐ Content Management System
- ☐ Data Science
- ☐ Databases & Storage
- ☐ Developer Tools

1 - 19 of 19 results for PyTorch.

Docker Official Image Verified Publisher

Best Match

**bitnami/pytorch**  
By VMware · Updated 15 days ago  
Bitnami container image for PyTorch  
DATA SCIENCE LANGUAGES & FRAMEWORKS MACHINE LEARNING & AI

↓ 500K+ · ☆ 72  
Pulls: 2,796  
Last week  
  
[Learn more](#)

**graphcore/pytorch**  
By Graphcore · Updated a year ago  
The Poplar® SDK components required to run Pytorch on IPUs. Designed for production.  
LANGUAGES & FRAMEWORKS MACHINE LEARNING & AI

↓ 100K+ · ☆ 4  
Pulls: 5  
Last week  
  
[Learn more](#)

**intel/intel-optimized-pytorch**  
By Intel Corporation · Updated 13 days ago  
Containers for running PyTorch workloads on Intel® Architecture.

↓ 100K+ · ☆ 14  
Pulls: 200  
Last week  
  
[Learn more](#)



docker

Find your  
container images

Wasm is a fast, light alternative to Linux containers — try it out today in the [Docker+Wasm Technical Preview](#)

docker hub  [Sign In](#) [Register](#)

↓ Pulls 50M+

Sort by

TAG	OS/ARCH	COMPRESSED SIZE
<a href="#">latest</a> Last pushed 2 months ago by <a href="#">tensorflowpackages</a>	linux/amd64	429.94 MB

docker pull tensorflow/tensorflow



docker

Community  
maintained  
images

hub.docker.com/u/bioconductor

dockerhub

bioconductor/bioconductor\_docker

Sign In

**Bioconductor** Sponsored OSS

Community Organization Bioconductor Boston, MA, USA <http://bioconductor.org/> Joined November 19, 2014

Repositories

Displaying 1 to 25 of 49 repositories

**bioconductor/yes-jupyter**   
By [Bioconductor](#) • Updated 9 hours ago

↓ 531 • ☆ 0

Pulls: 11  
Dec 3 to Dec 9

[Learn more](#)

**bioconductor/rstudio\_ycsds**   
By [Bioconductor](#) • Updated 9 hours ago

↓ 553 • ☆ 0

Pulls: 5  
Oct 29 to Nov 4

Vendor  
optimized and  
signed images

<https://catalog.ngc.nvidia.com/>  
(Look for public images)

The screenshot shows the NVIDIA NGC Catalog interface. The left sidebar contains navigation links: Explore Catalog, Collections, Containers, Helm Charts, Models, and Resources. The main content area is titled 'Containers > NVIDIA HPC SDK' and 'NVIDIA HPC SDK'. It features the NVIDIA HPC logo and a description: 'The NVIDIA HPC SDK is a comprehensive suite of compilers, libraries and tools essential to maximizing developer productivity and the performance and portability of HPC applications.' Below the description are fields for Publisher (NVIDIA), Latest Tag (24.5-devel-cuda\_multi-ubuntu22.04), Modified (June 30, 2024), Compressed Size (10.41 GB), and Multinode Support. On the right, there are tabs for Overview, Tags, Layers, Security Scanning, and Related Collections. The Tags tab is active, showing a search bar and a list of tags. A blue banner at the top of the tags section states: 'A public key is required to validate the signed images below. [View all public keys](#) maintained by NVIDIA.' The list of tags includes: 24.5-devel-cuda\_multi-ubuntu22.04 (Signed, 06/07/2024 6:00 PM, 10.41 GB, 2 Architectures), 24.5-runtime-cuda12.4-ubuntu22.04 (Signed, 06/07/2024 5:59 PM, 2.42 GB, 2 Architectures), and 24.5-devel-cuda\_multi-centos7 (Signed, 06/07/2024 5:58 PM, 10.45 GB, 1 Architecture).

catalog.ngc.nvidia.com/orgs/nvidia/containers/nvhpc/tags

**NVIDIA** NGC Catalog

Explore Catalog  
Collections  
Containers  
Helm Charts  
Models  
Resources

Containers > NVIDIA HPC SDK  
NVIDIA HPC SDK

**Description**  
The NVIDIA HPC SDK is a comprehensive suite of compilers, libraries and tools essential to maximizing developer productivity and the performance and portability of HPC applications.

**Publisher**  
NVIDIA

**Latest Tag**  
24.5-devel-cuda\_multi-ubuntu22.04

**Modified**  
June 30, 2024

**Compressed Size**  
10.41 GB

**Multinode Support**

Overview **Tags** Layers Security Scanning Related Collections

Search tags...

A public key is required to validate the signed images below. [View all public keys](#) maintained by NVIDIA.

**24.5-devel-cuda\_multi-ubuntu22.04** **Signed** nvcr.io/nvidia/nvhpc:24.5-devel-cuda\_multi-ubuntu22.04  
06/07/2024 6:00 PM 10.41 GB 2 Architectures

**24.5-runtime-cuda12.4-ubuntu22.04** **Signed** nvcr.io/nvidia/nvhpc:24.5-runtime-cuda12.4-ubuntu22.04  
06/07/2024 5:59 PM 2.42 GB 2 Architectures

**24.5-devel-cuda\_multi-centos7** **Signed** nvcr.io/nvidia/nvhpc:24.5-devel-cuda\_multi-centos7  
06/07/2024 5:58 PM 10.45 GB 1 Architecture



docker

Find your  
container images

Your community  
may have a  
standard image.  
Best to stick with it.

Can create derivative  
image with  
customizations

The screenshot shows the Docker Hub interface for the TensorFlow image. At the top, there's a navigation bar with links for Explore, Pricing, Sign In, and a Register button. Below this, the TensorFlow image is listed with a star icon and a pull count of 50M+. The image description mentions it's a machine learning framework. The 'Tags' tab is selected, showing a list of tags with 'latest' as the current tag. It also shows the last push time (2 months ago) and the image digest (7f9f23ce2473). On the right side, there's a section for 'COMPRESSED SIZE' which is 429.94 MB.

Wasm is a fast, light alternative to Linux containers — try it out today in the [Docker+Wasm Technical Preview](#)

Explore Pricing Sign In Register

★ Pulls 50M+

machine learning framework TensorFlow (<http://www.tensorflow.org>)

Overview Tags

Sort by Newest

TAG

[latest](#)

Last pushed 2 months ago

DIGEST

[7f9f23ce2473](#)

tensorflow/tensorflow

COMPRESSED SIZE 429.94 MB



# Explore the rich ecosystem

Kubernetes is quite powerful by itself

- But it is the surrounding ecosystem that makes it so valuable!

Kubernetes has become an industry standard

- Many projects build on top of it to deliver additional capabilities

# Helm package manager

Helm has become the main package manager

- Many pods can be bundled in a coherent package
- Known as “helm charts”

<https://helm.sh>

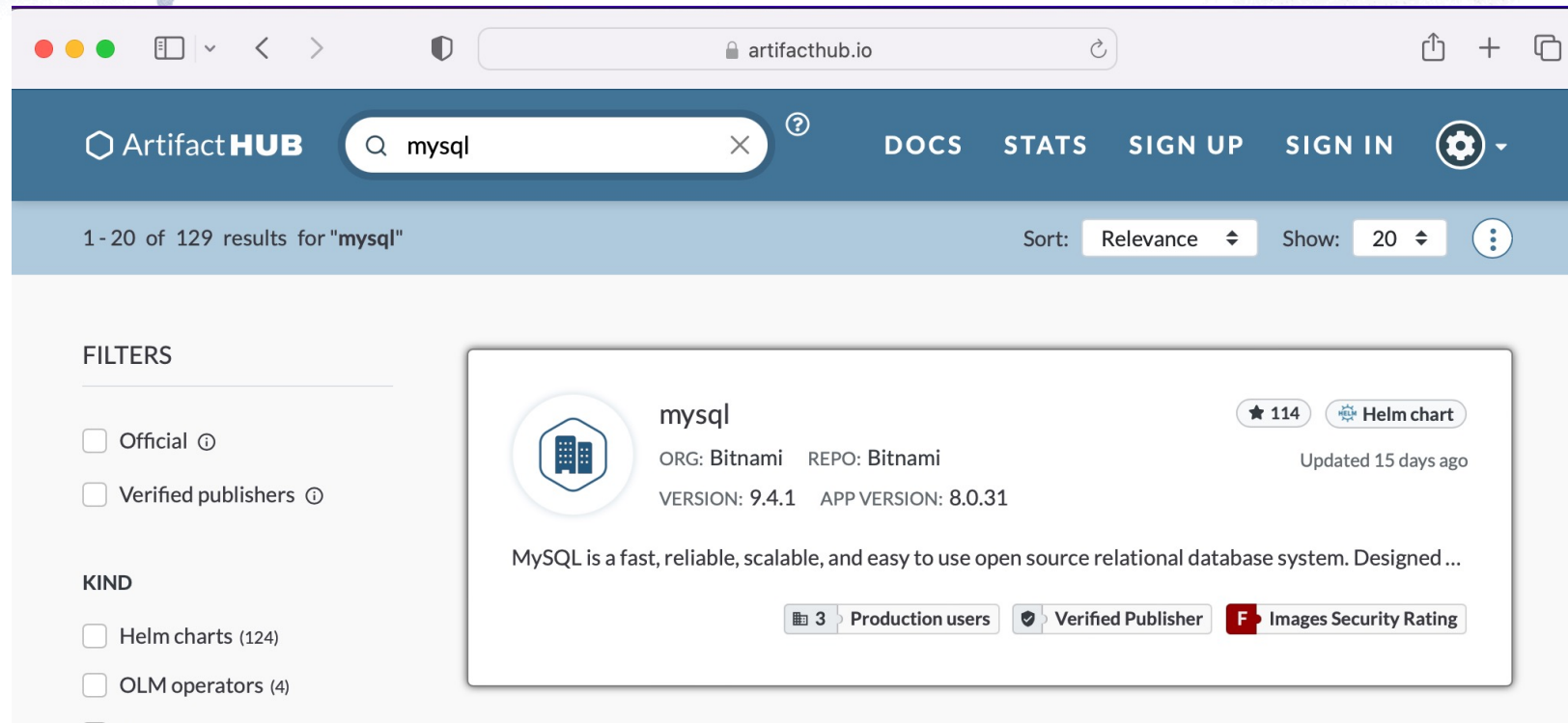


Typically used to deploy new capabilities

- e.g. multi-pod services
- But can be used for complex science workflows, too
- We will have a LLM as a service deployment example in hands-on session

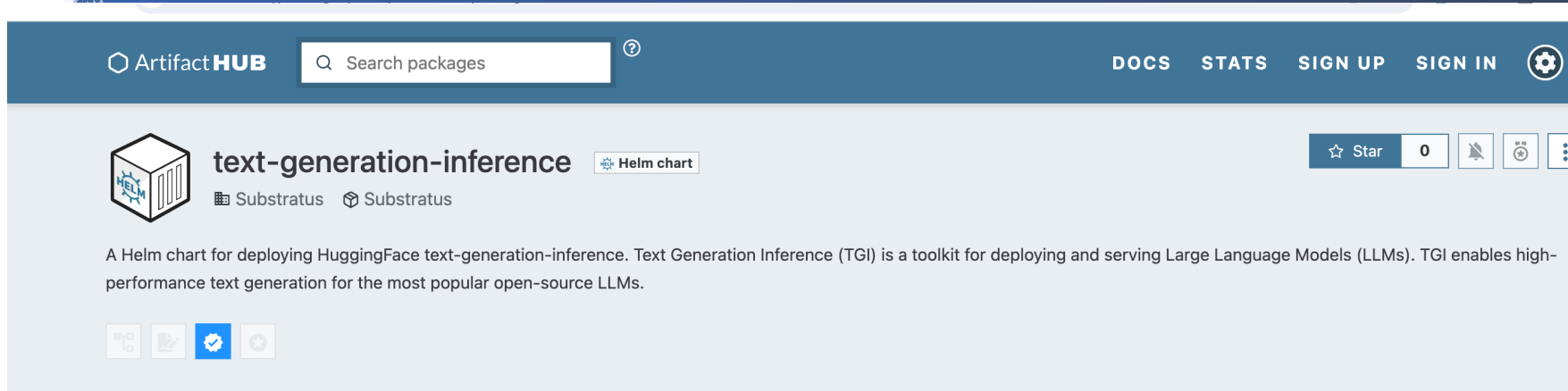


# Large catalogue of services



<https://artifacthub.io/>

# Large catalogue of services



The screenshot shows the top section of the Artifact Hub page for the 'text-generation-inference' Helm chart. The header includes the 'Artifact HUB' logo, a search bar, and navigation links for 'DOCS', 'STATS', 'SIGN UP', and 'SIGN IN'. The chart title 'text-generation-inference' is displayed with a 'Helm chart' badge and a 'Star' button showing 0 stars. Below the title, it lists 'Substratus' as the provider. A description states: 'A Helm chart for deploying HuggingFace text-generation-inference. Text Generation Inference (TGI) is a toolkit for deploying and serving Large Language Models (LLMs). TGI enables high-performance text generation for the most popular open-source LLMs.' At the bottom of this section are icons for GitHub, Docker, and Helm.

<https://artifacthub.io/>

## text-generation-inference Helm Chart

A Helm chart for deploying HuggingFace text-generation-inference. Text Generation Inference (TGI) is a toolkit for deploying and serving Large Language Models (LLMs). TGI enables high-performance text generation for the most popular open-source LLMs.

### Usage

Basic usage:

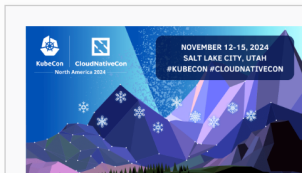
```
helm repo add substratusai https://substratusai.github.io/helm
helm install mistral-7b-instruct substratusai/text-generation-inference
# Note by default the resource limit is set to 1 GPU
helm install mistral-7b-instruct substratusai/text-generation-inference \
  --set model=mistralai/Mistral-7B-Instruct-v0.1
```

INSTALL

TEMPLATES

DEFAULT VALUES

CHANGELOG



## And if K8S still does not look appealing

- If you do not like YAML and Pods, it is easy to create an **overlay batch system** on top of Kubernetes
  - Kubernetes capabilities are really a superset of most batch systems (if we ignore scheduling capabilities)
- We have plenty of experience using **HTCondor** as an overlay batch system for Kubernetes (in collaboration with OSG)
  - Users never see Kubernetes, just HTCondor
  - Resources provisioned by Kubernetes given to HTCondor for management (the pilot concept)

<https://doi.org/10.48550/arXiv.2205.01004>

# MPI jobs

- Kubernetes does not natively support MPI jobs
  - Or gang-scheduling in general
- Several external projects help getting a cluster to support them, e.g.
  - kubeflow:  
<https://www.kubeflow.org/docs/components/training/mpi/>
  - kube-openmpi:  
<https://github.com/everpeace/kube-openmpi>

# MPI jobs

- Kubeflow MPI allows for specs on both launcher (where mpirun command runs) and worker nodes (where the actual MPI tasks land)
- There are pods on the launcher node and each of the worker nodes
- stdout/stderr from the MPI job is in the launcher logs as that is where the mpirun is executed
- Example job showing the launcher and worker pods below

```
mahidhar@vgr-1-20:~$ kubectl get pods -n default | grep mpiexample
mpiexample-mahi-launcher-nqbth      1/1      Running      0          35s
mpiexample-mahi-worker-0            1/1      Running      0          35s
mpiexample-mahi-worker-1            1/1      Running      0          35s
mahidhar@vgr-1-20:~$
```



# Acknowledgements

This work was partially funded by  
US National Science Foundation (NSF) awards  
OAC-1541349, OAC-1826967, OAC-2112167,  
OAC-2030508, OAC-1841530 and CNS-1730158.





The end