



BUILDING FOR THE ROAD AHEAD

# DETROIT 2022

# Computational Fluid Dynamics (CFD) analysis with Kubernetes, Kubeflow, and OpenFOAM



BUILDING FOR THE ROAD AHEAD

**DETROIT 2022**



**KubeCon**



**CloudNativeCon**

North America 2022

BUILDING FOR THE ROAD AHEAD

**DETROIT 2022**

**October 24-28, 2021**



**Erik Jacobs**

*Red Hat*

# Thank You



**Eduardo Arango**  
*Red Hat*



**Aldo Culquicondor**  
*Google*

BUILDING FOR THE ROAD AHEAD

**DETROIT 2022**



© Warner Brothers

<https://cloud.redhat.com/blog/how-to-use-kubeflow-and-the-mpi-operator-on-openshift>



[https://en.wikipedia.org/wiki/Grumman\\_F-14\\_Tomcat#/media/File:US\\_Navy\\_051105-F-5480T-005\\_An\\_F-14D\\_Tomcat\\_conducts\\_a\\_mission\\_over\\_the\\_Persian\\_Gulf-region.jpg](https://en.wikipedia.org/wiki/Grumman_F-14_Tomcat#/media/File:US_Navy_051105-F-5480T-005_An_F-14D_Tomcat_conducts_a_mission_over_the_Persian_Gulf-region.jpg)



Danger Zone  
Obscure Memes  
Pop Culture  
Bad Jokes  
Poor Taste



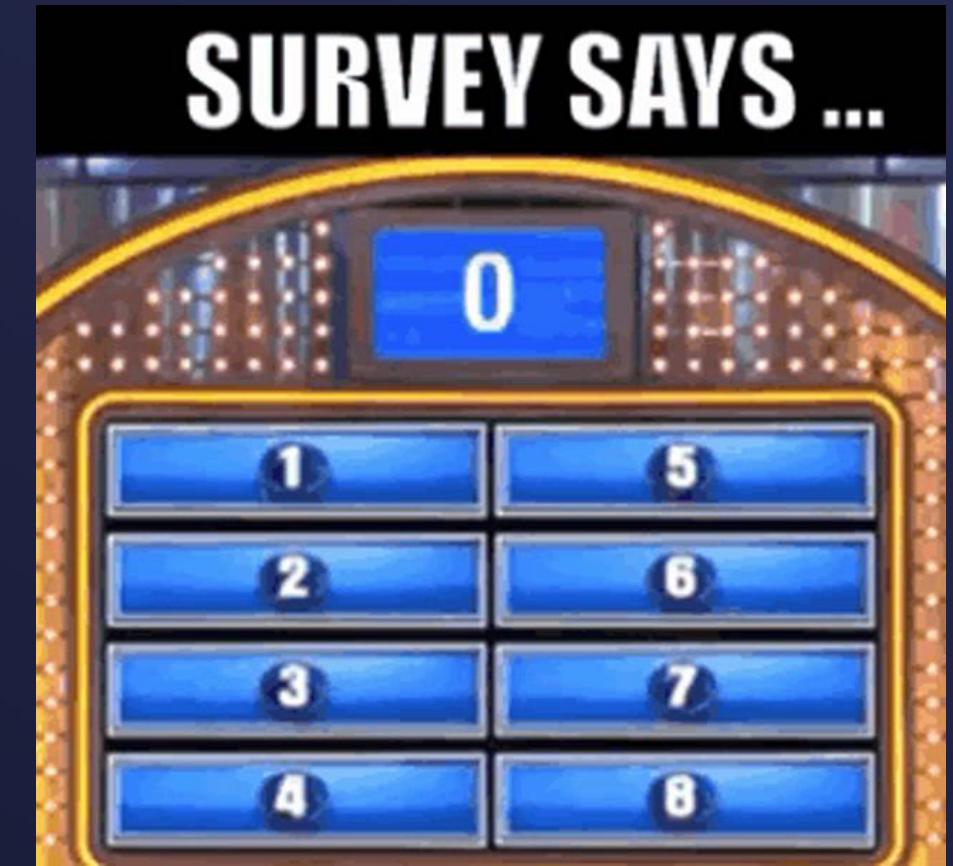
BUILDING FOR THE ROAD AHEAD

**DETROIT 2022**



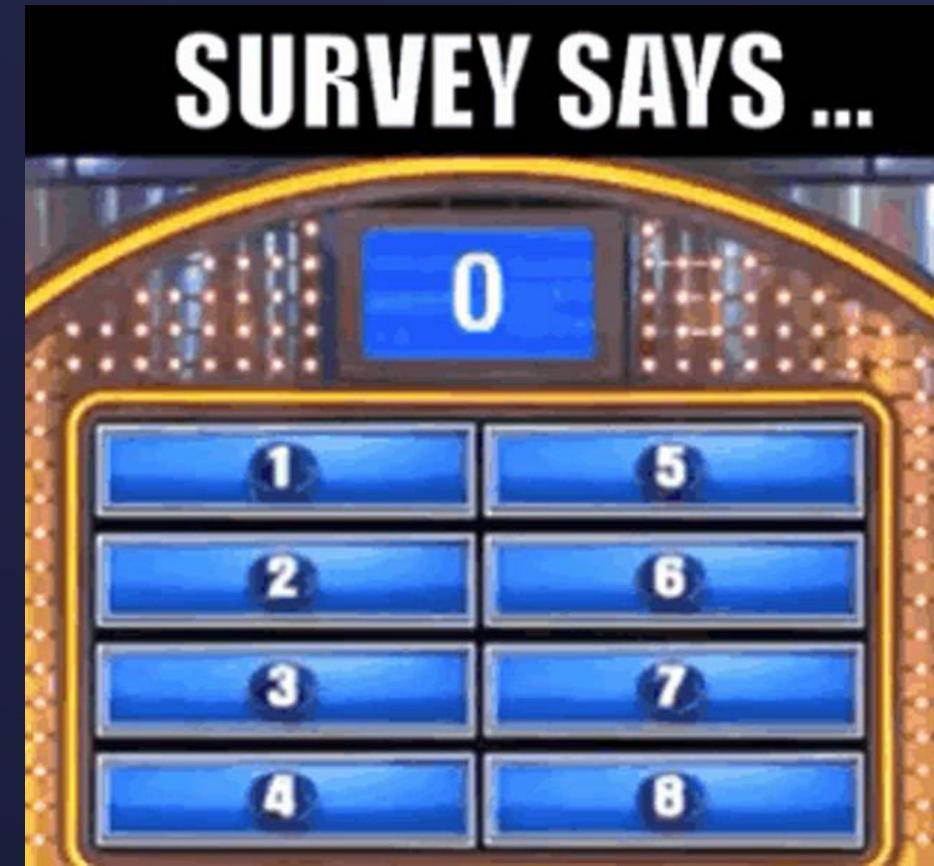
<https://www.flickr.com/photos/jessicafm/2767435428>

**DETROIT 2022**



© ABC Television

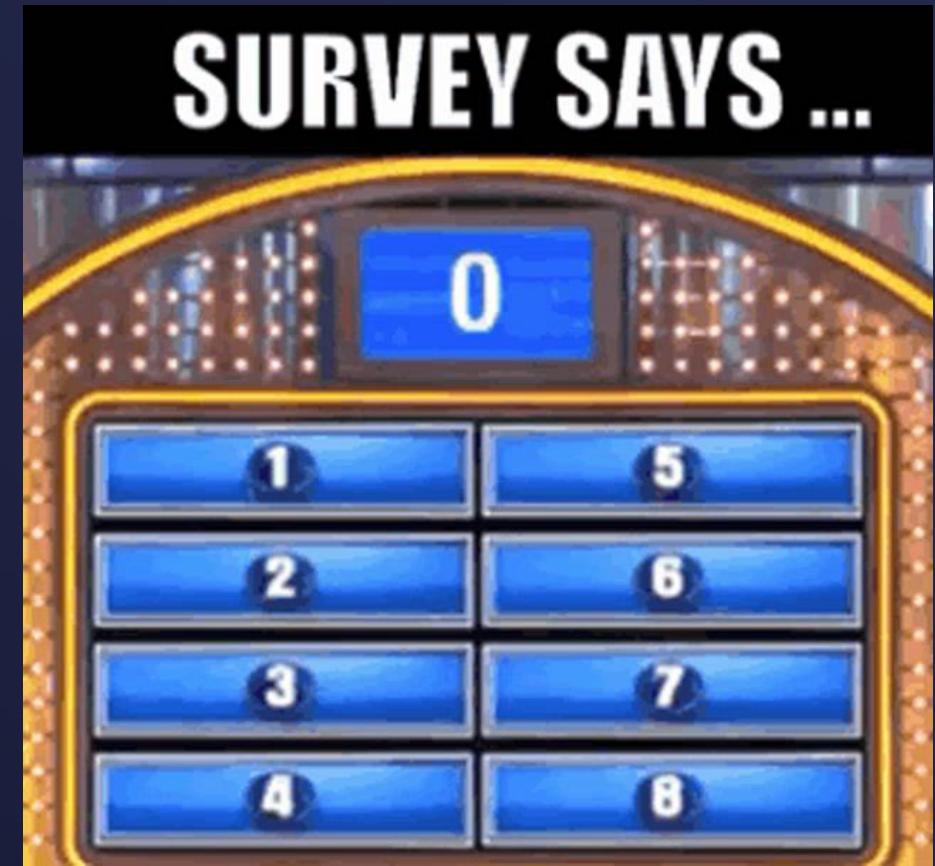
**Kubernetes user?**



© ABC Television

**DETROIT 2022**

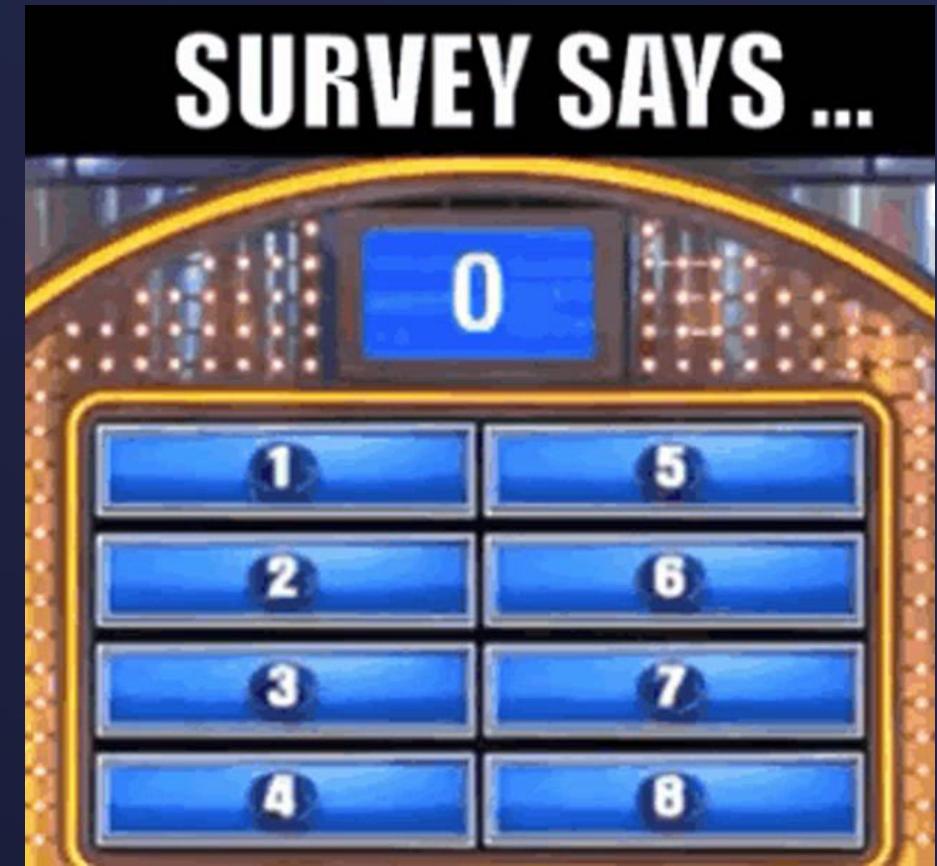
HPC user?



© ABC Television

**DETROIT 2022**

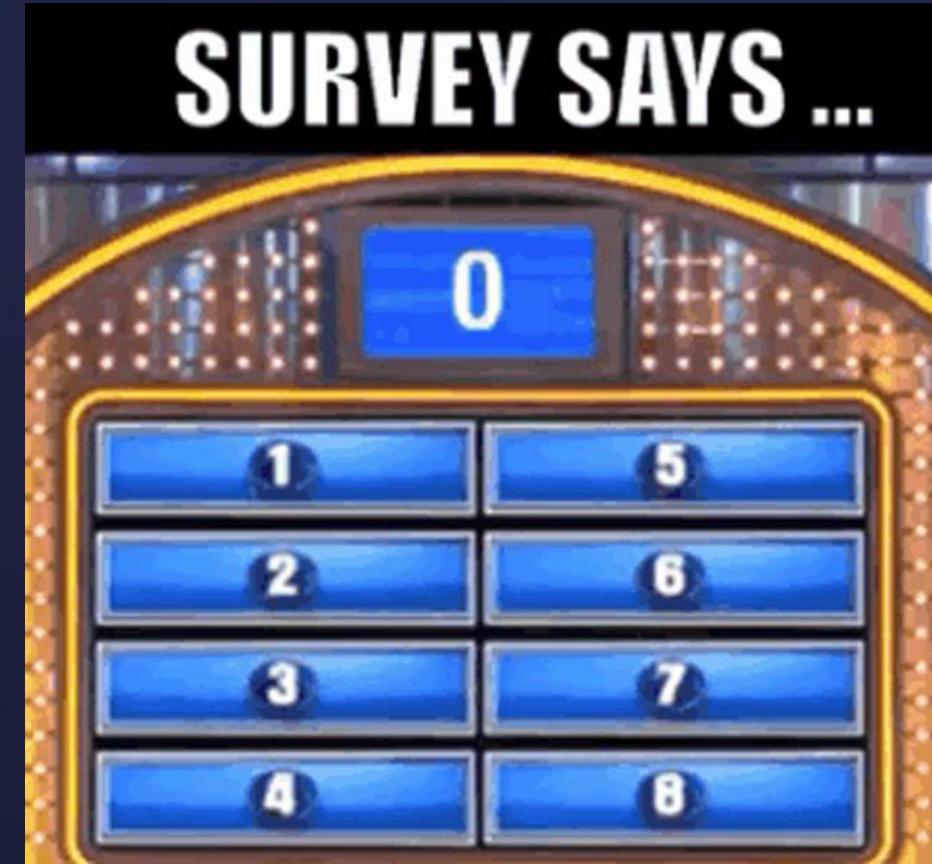
**MPI user?**



© ABC Television

**DETROIT 2022**

**OpenFOAM user?**



© ABC Television

**DETROIT 2022**

# How Did We Get Here?

- 14 years at Red Hat
- Not a software engineer
- Not a Kubernetes, Kubeflow, or OpenFOAM developer
- Not an aerodynamicist
- But...

# About Me



# How it started...

## AUTOMAKER

Hi

Thank you for your reponce.

At first [REDACTED] has already converted some OSS based solver (OpenFOAM) to container. And they have already estimated the performance of the container based application.

Some application is based on Windows. This is the really why they need to use VM. And they want to manage all VM and container with the same method.

Best Regards.

MORLIND ENGINEERING SERVICES

ABOUT PORTFOLIO CONTACT SIGN IN

## ENGINEERING SOLUTIONS

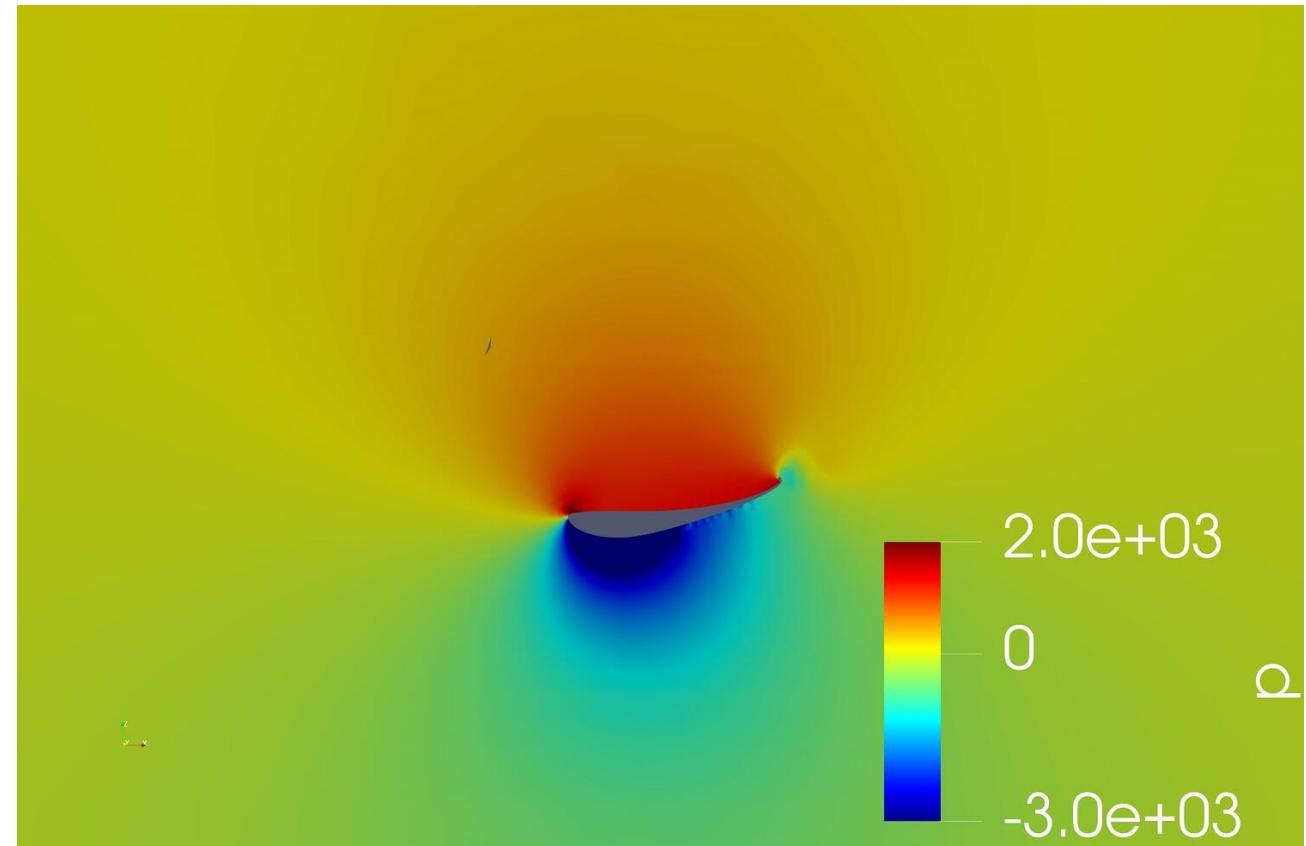
One-stop shopping for design, development,  
simulation, prototyping, and testing



# Experiment time!



© Warner Brothers



**DETROIT 2022**

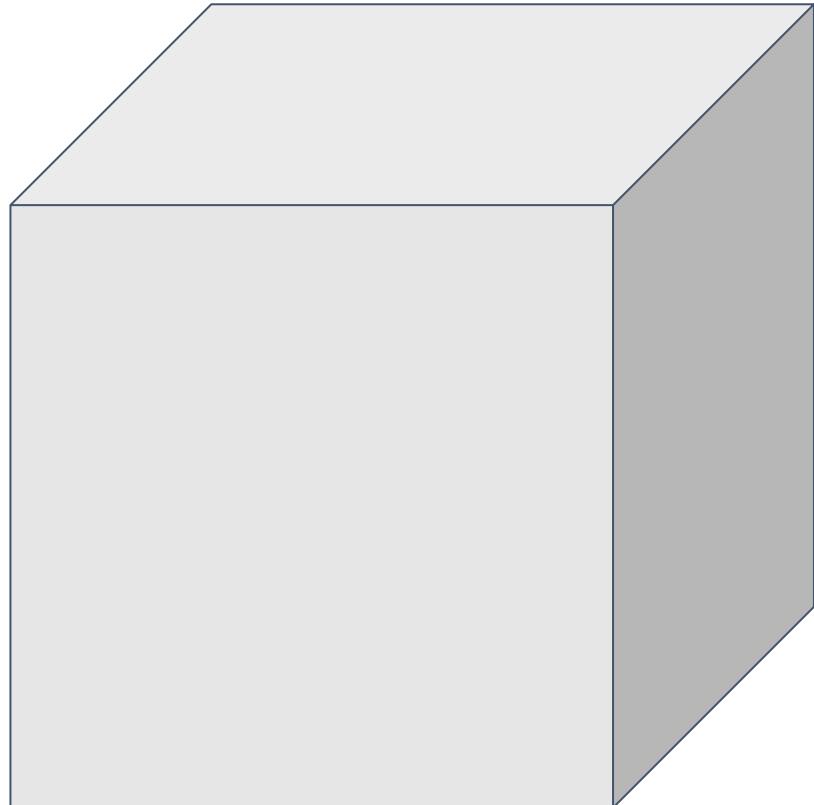
# OpenFOAM & MPI



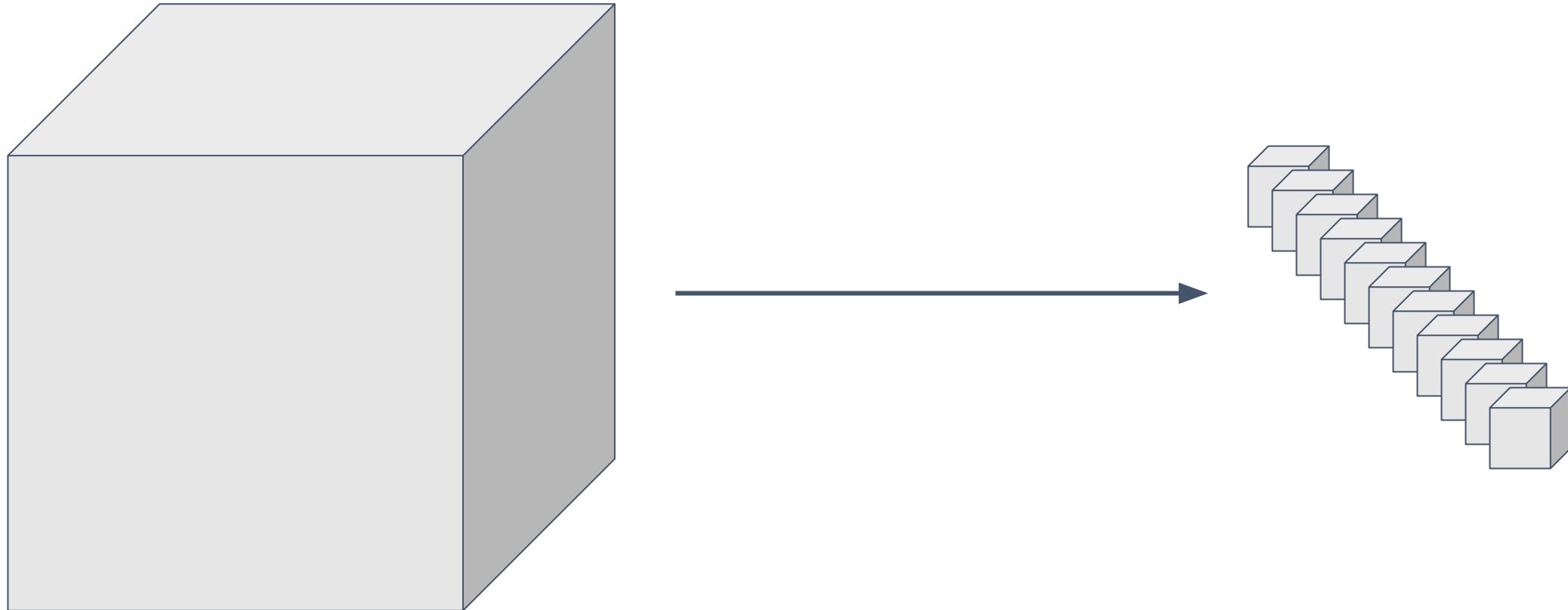
MPI is a standardized API typically used for **parallel and/or distributed computing**. Open MPI is an open source, freely available implementation of the MPI specifications.



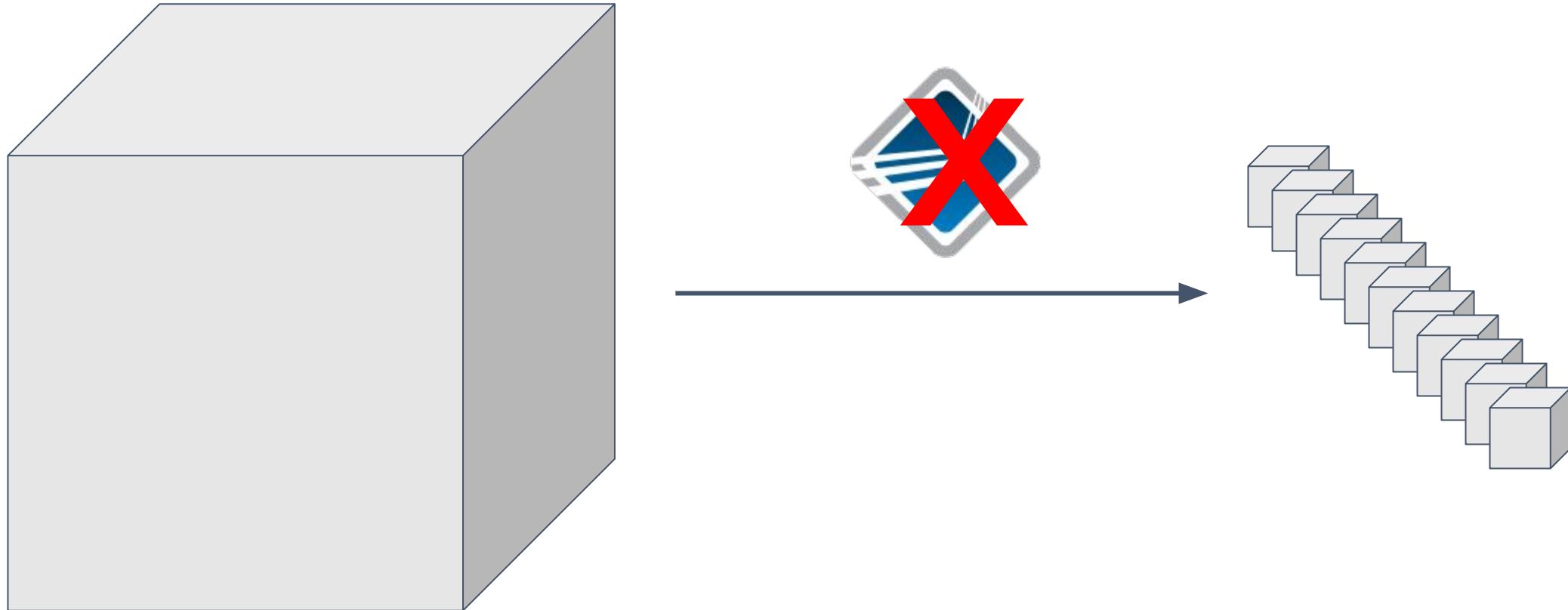
OpenFOAM is an open-source toolkit for performing **computational fluid dynamics** (CFD) analysis.



# OpenFOAM & MPI



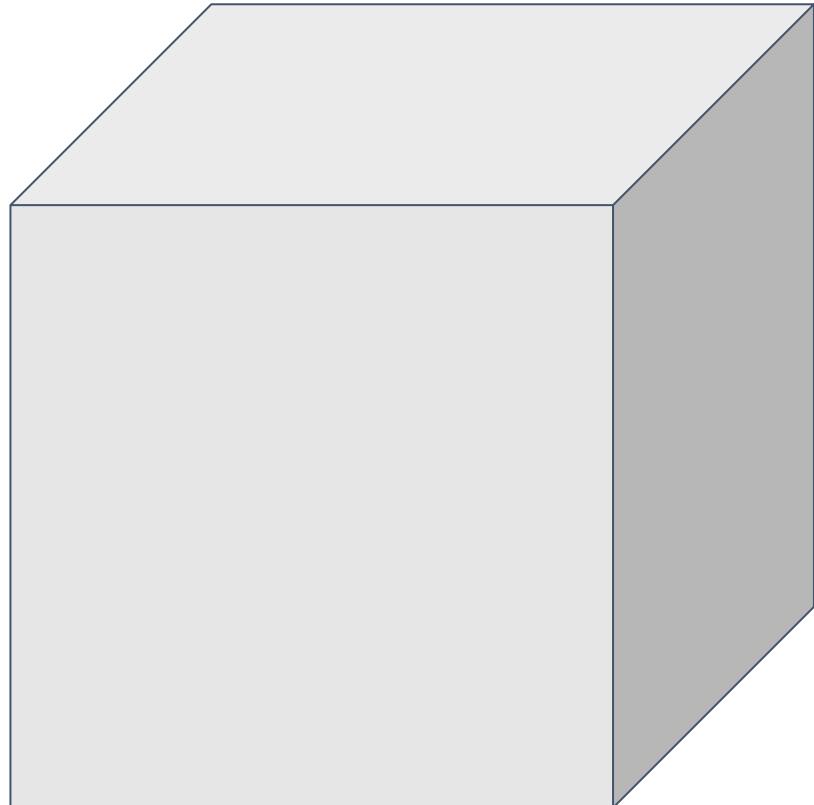


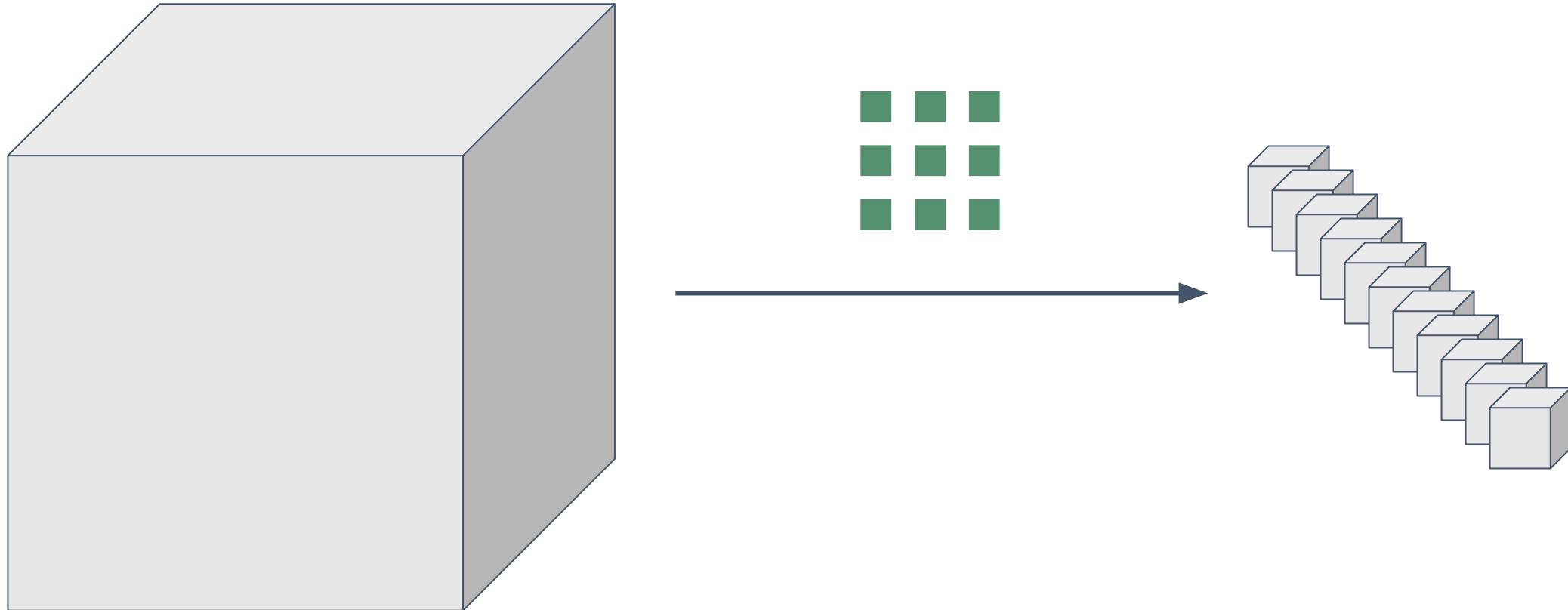




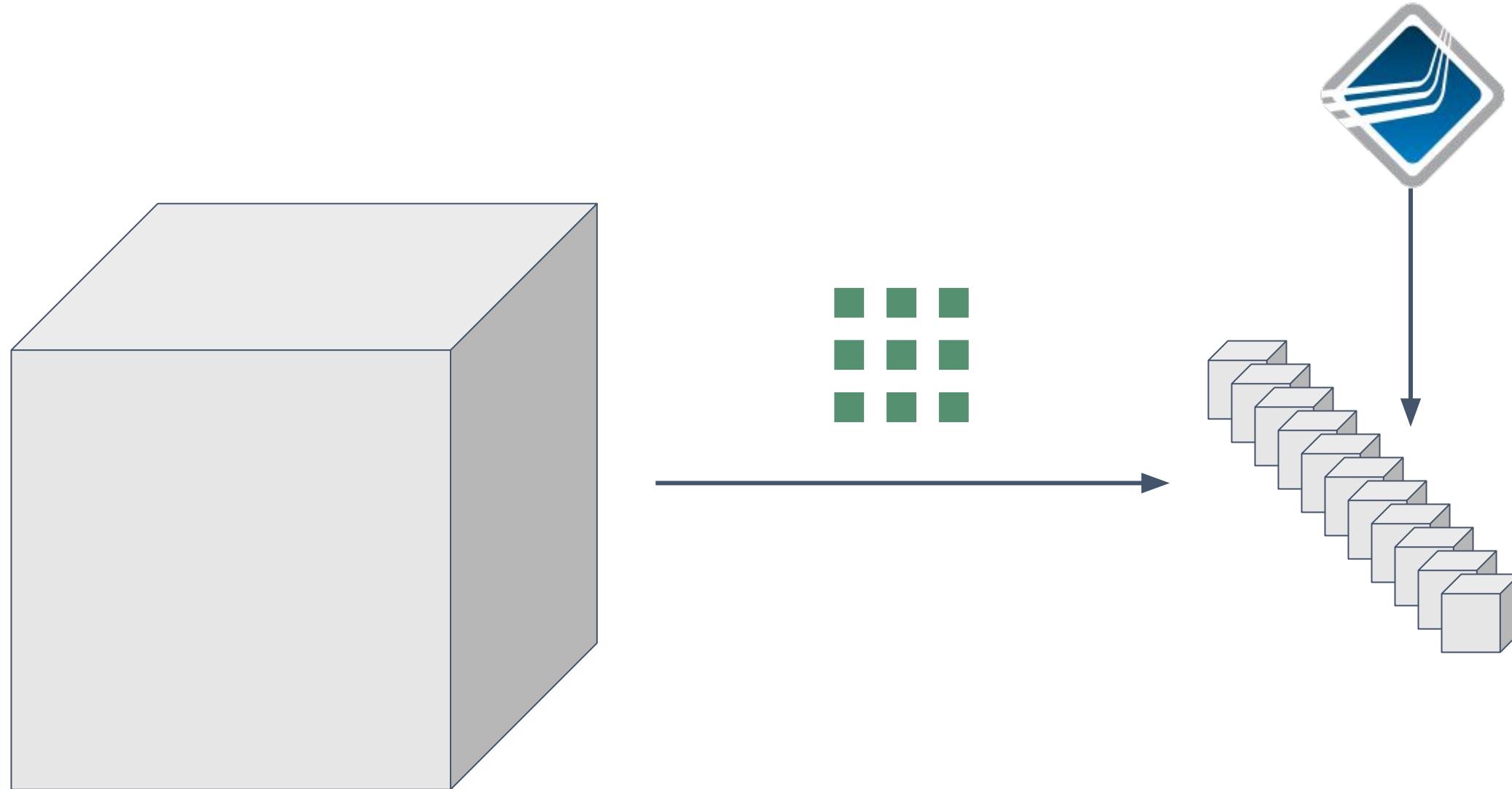




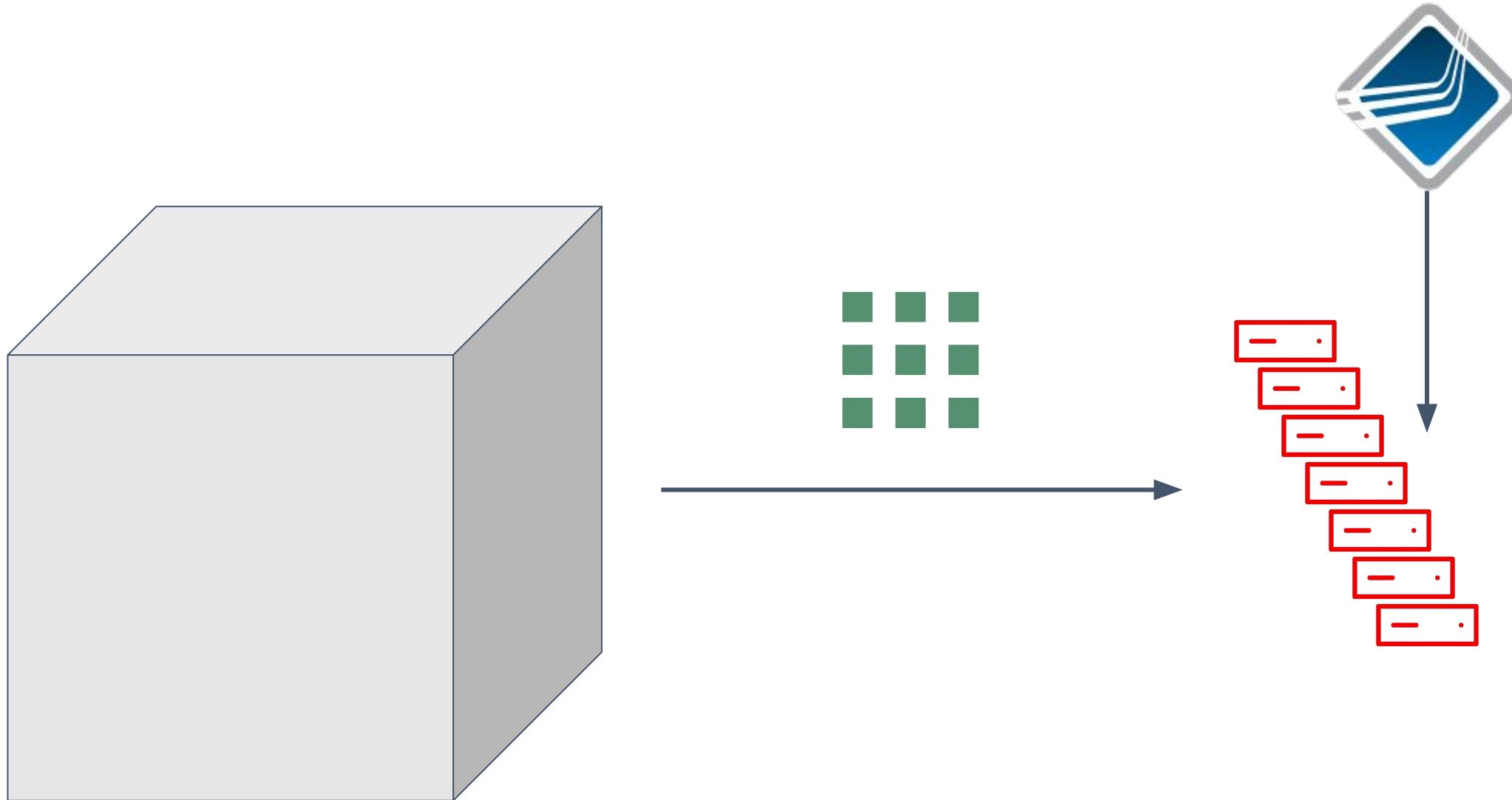




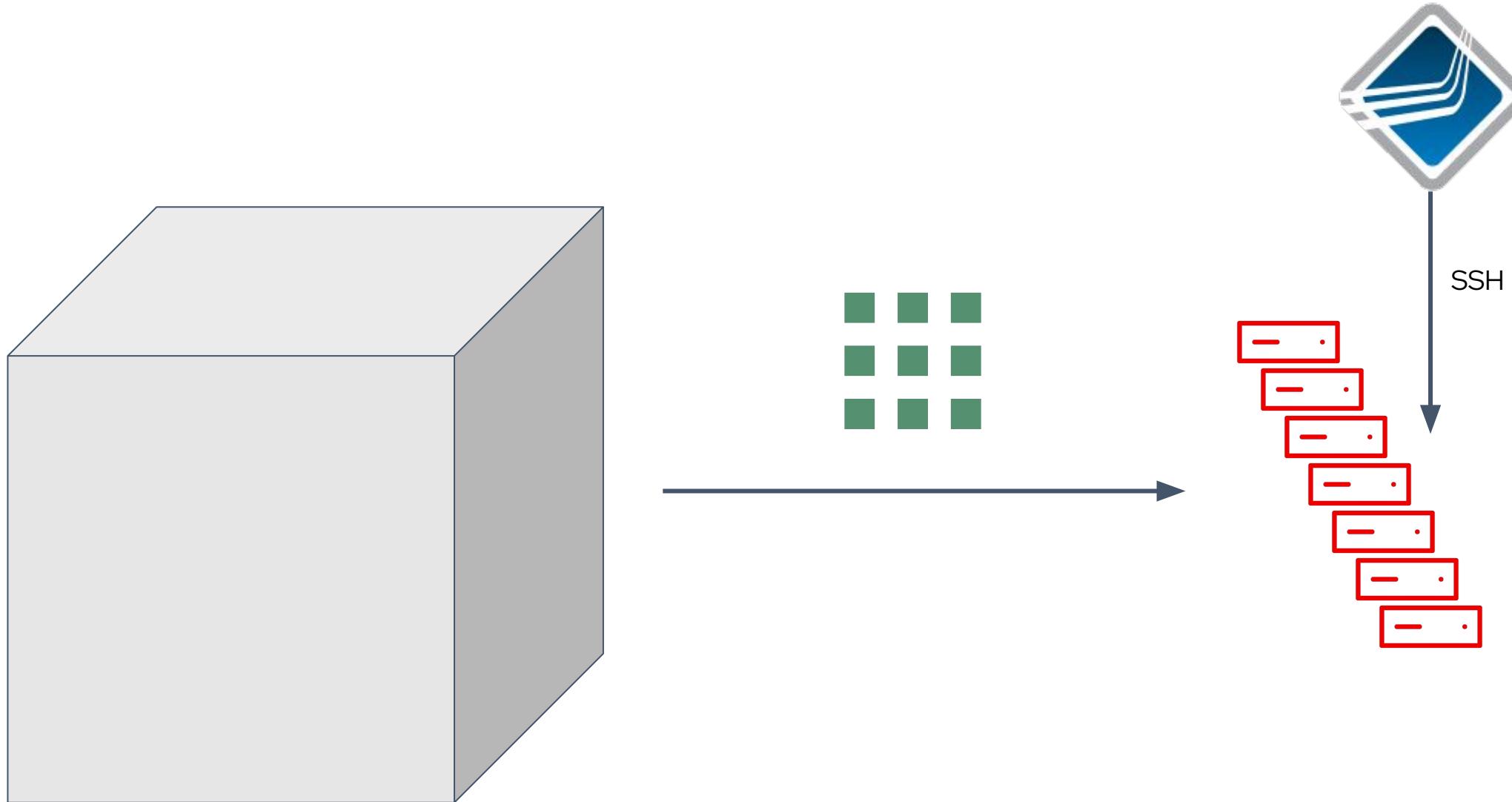
# OpenFOAM & MPI



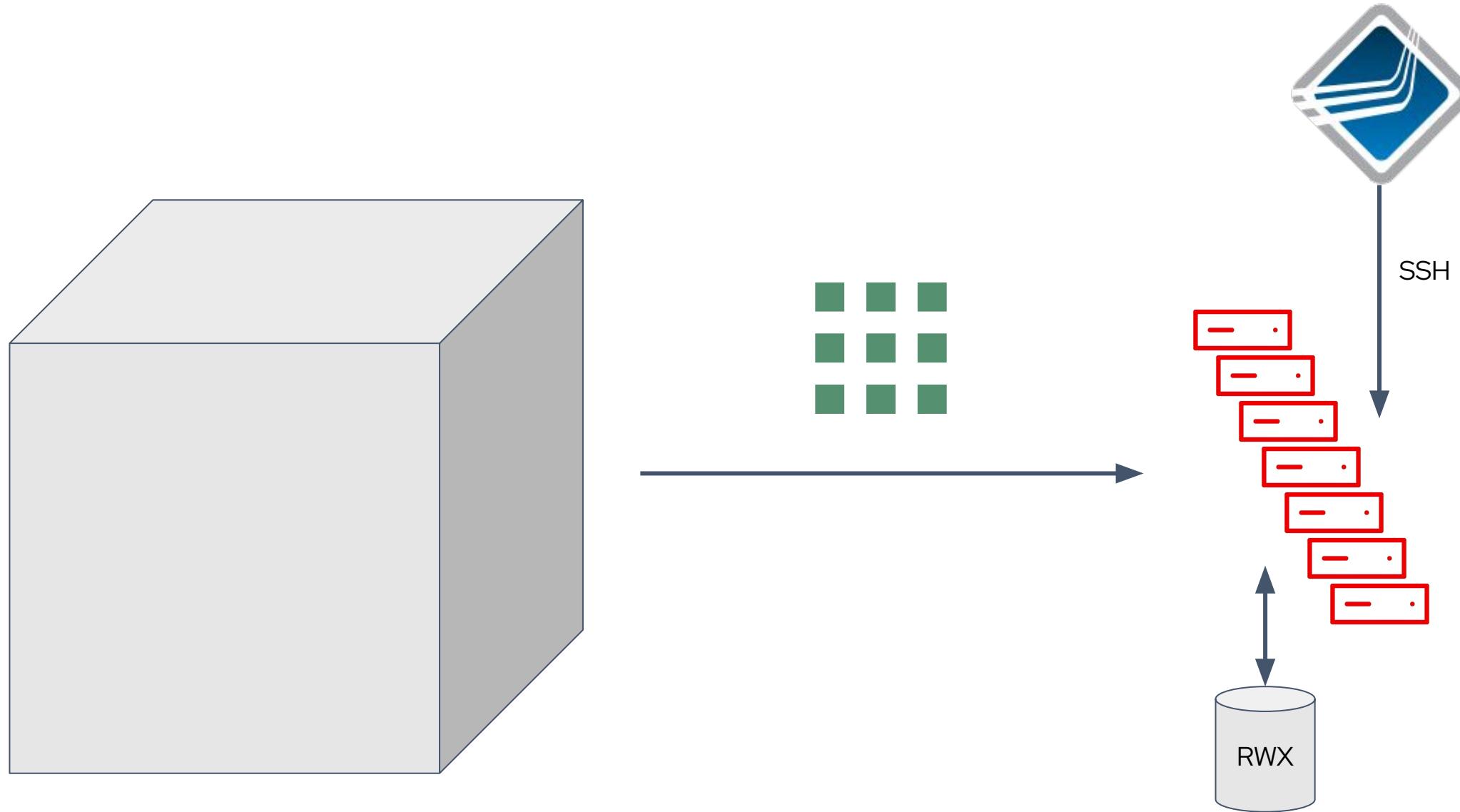
# OpenFOAM & MPI



# OpenFOAM & MPI



# OpenFOAM & MPI

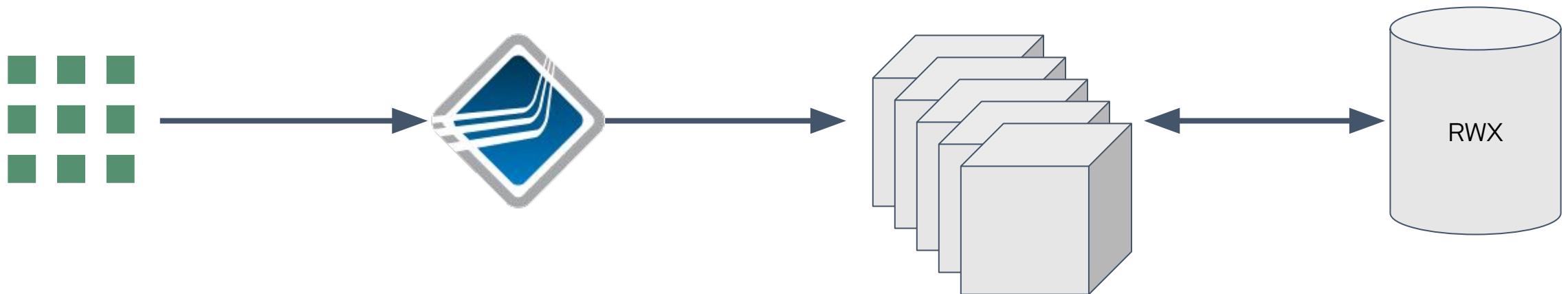


© New Line Cinema

**DETROIT 2022**



# For the sake of simplicity



**DETROIT 2022**

# Clusters and (are?) Kubernetes

# Clusters and (are?) Kubernetes



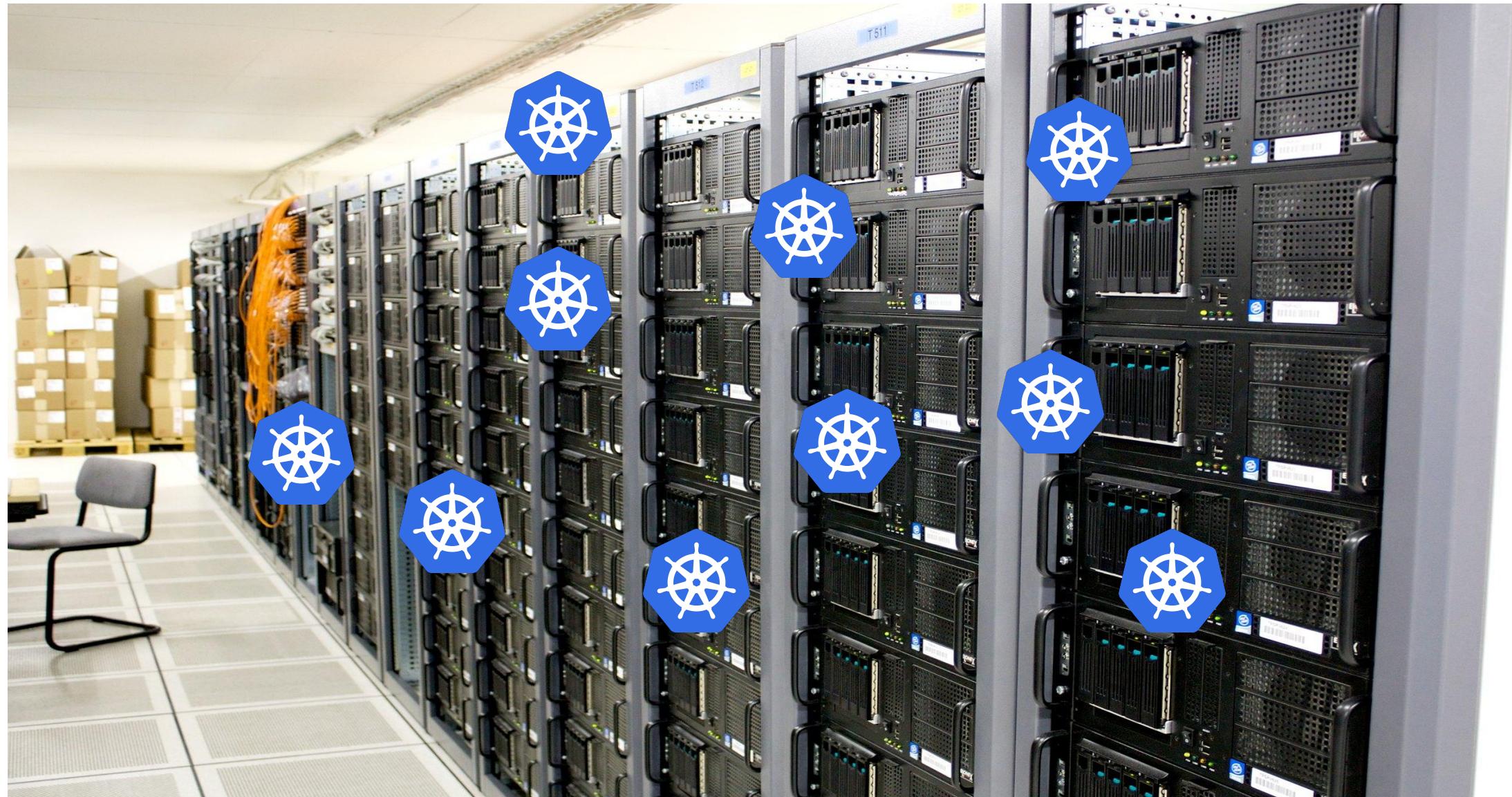
<https://www.flickr.com/photos/torkildr/3462607995>

# Clusters and (are?) Kubernetes



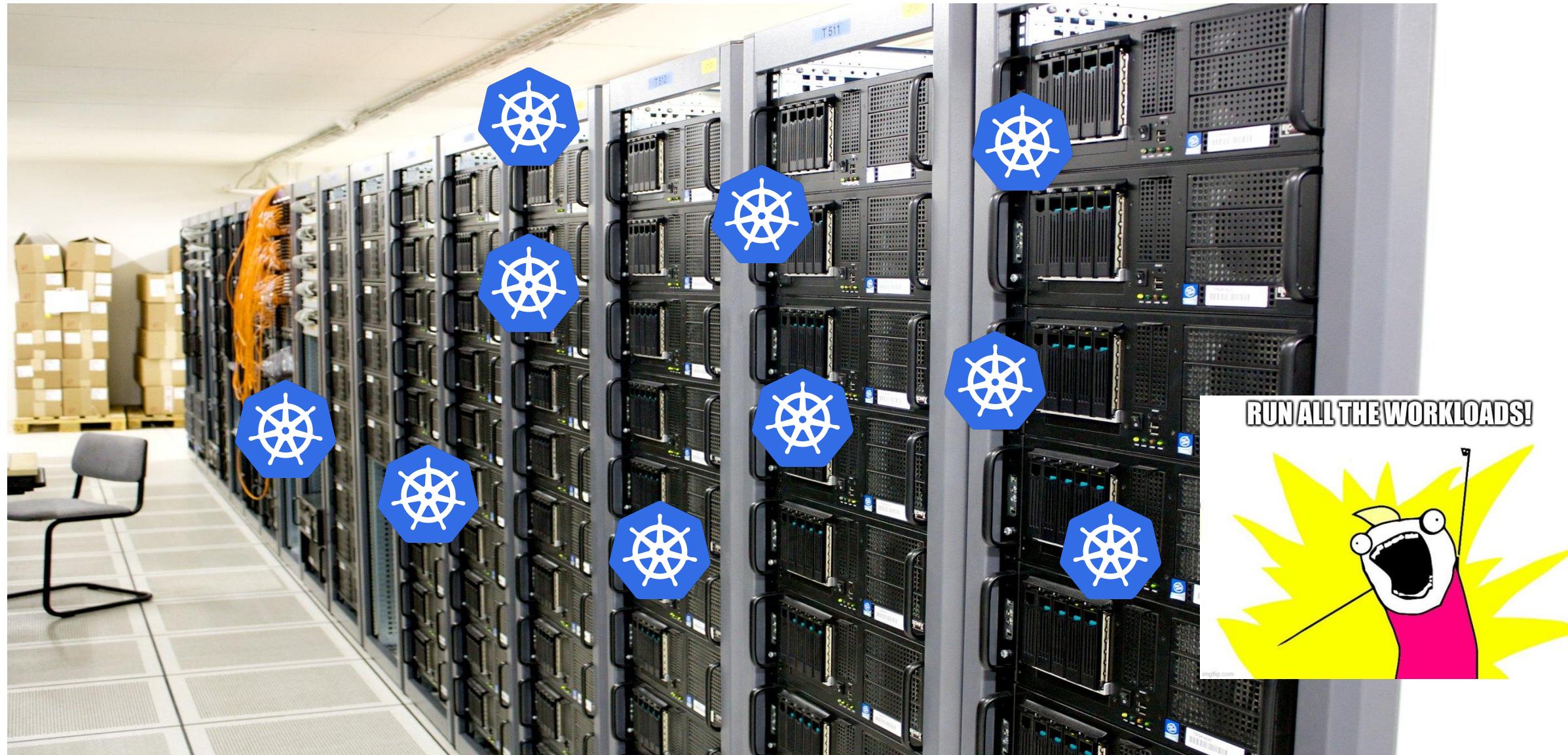
7 June 2014

# Clusters and (are?) Kubernetes



<https://www.flickr.com/photos/torkildr/3462607995>

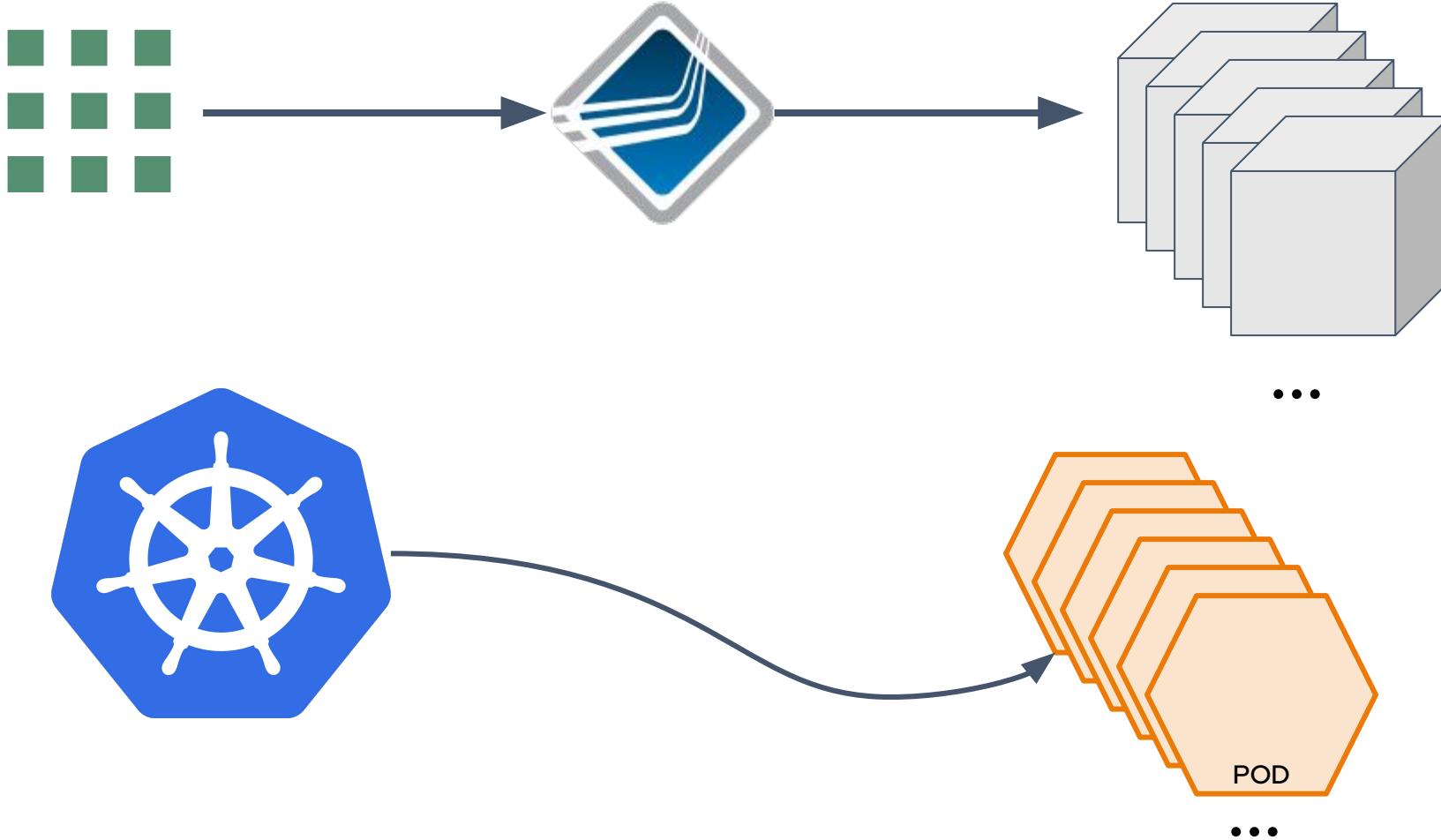
# Clusters and (are?) Kubernetes



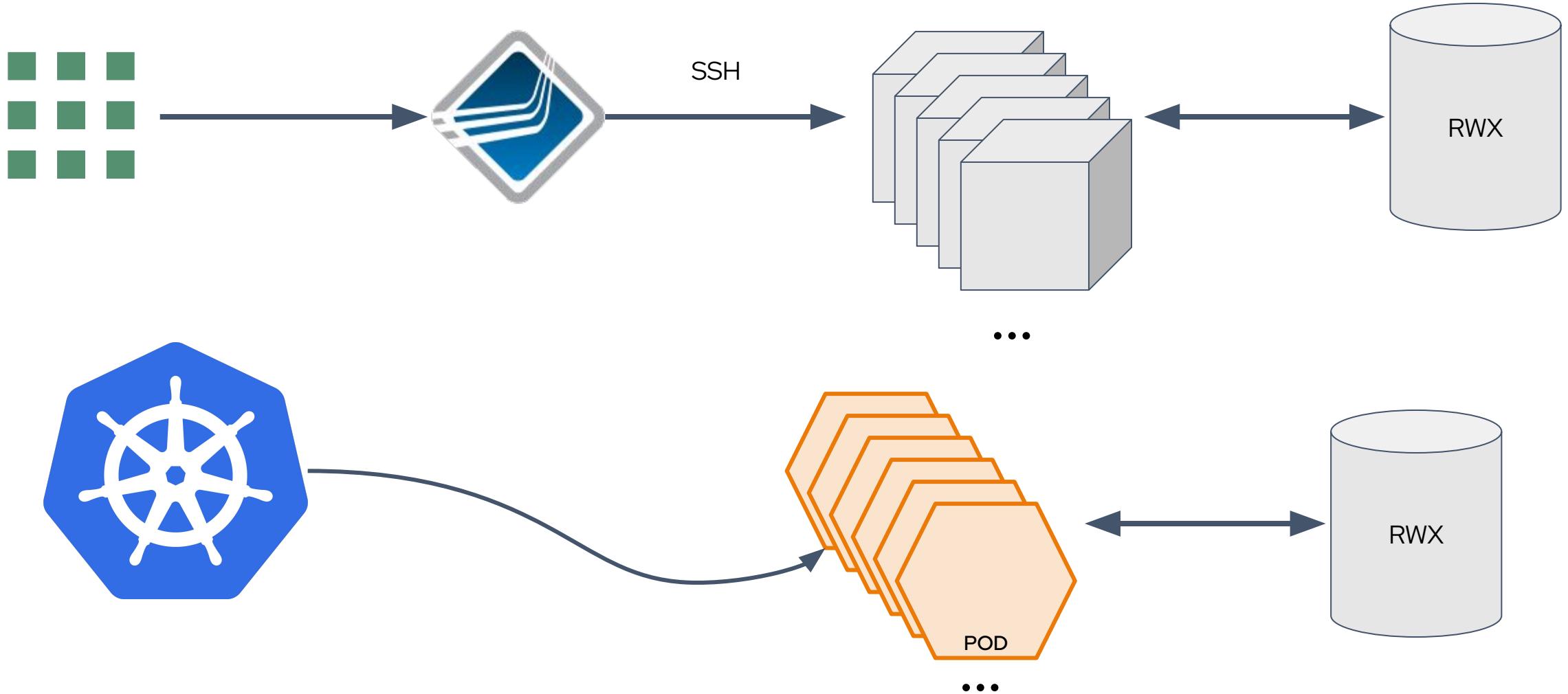
# Clusters and (are?) Kubernetes



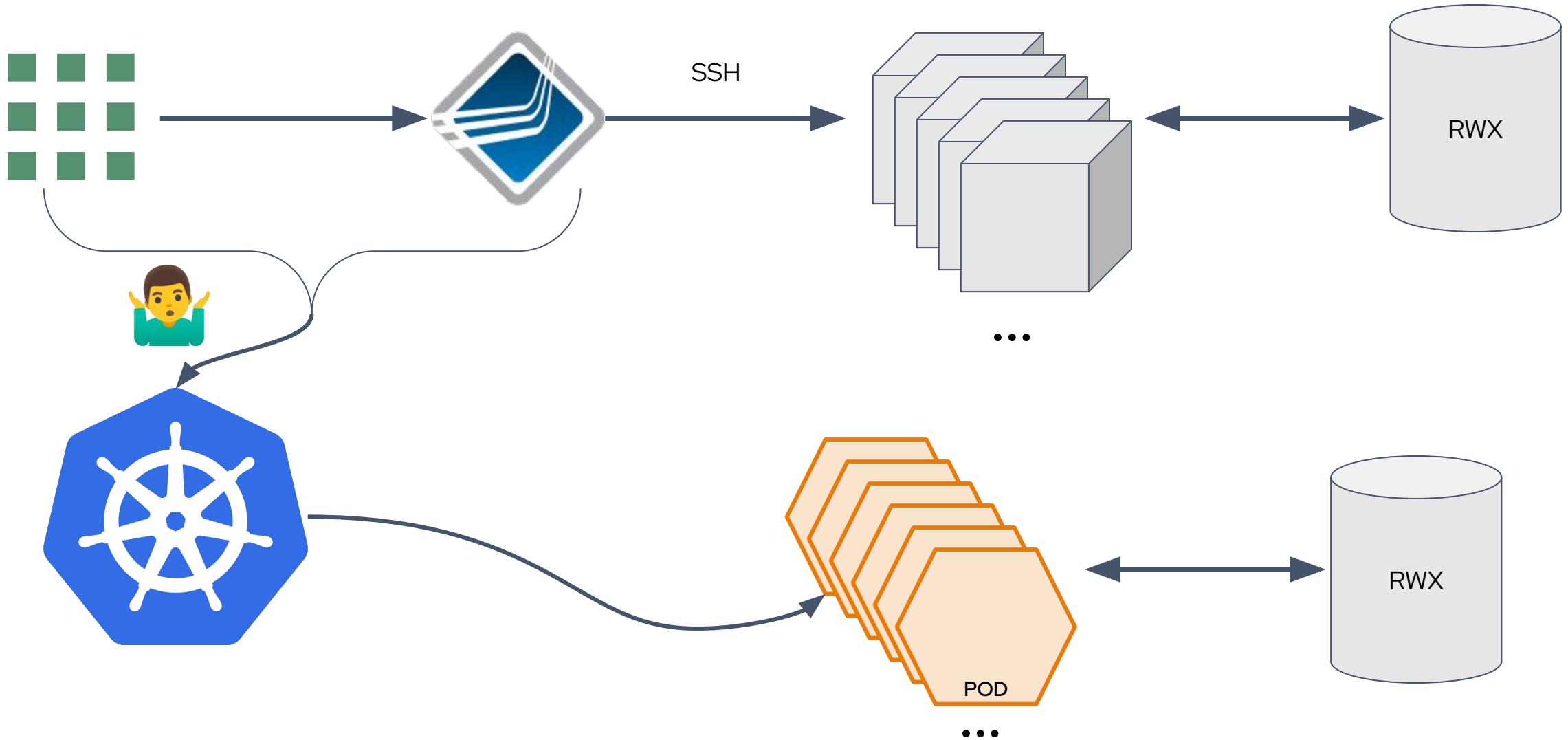
# Clusters and (are?) Kubernetes



# Clusters and (are?) Kubernetes



# Clusters and (are?) Kubernetes



**DETROIT 2022**



© Warner Brothers

**DETROIT 2022**

# Kubeflow



The Kubeflow project is dedicated to making deployments of machine learning (ML) workflows on Kubernetes simple, portable and scalable.



The Kubeflow project is dedicated to making ~~deployments of machine learning (ML) workflows~~ on Kubernetes simple, portable and scalable.



Kubeflow

Search this site...

## Documentation

- ▶ About
- ▶ Getting Started
- ▼ Components
  - ▶ Central Dashboard
  - ▶ Kubeflow Notebooks
  - ▶ Kubeflow Pipelines
  - ▶ Katib
- ▼ Training Operators
  - TensorFlow Training (TFJob)

[Documentation](#) / [Components](#) / [Training Operators](#) / [MPI Training \(MPIJob\)](#)

# MPI Training (MPIJob)

Instructions for using MPI for training

## Alpha

This Kubeflow component has **alpha** status with limited support. See the [Kubeflow versioning policies](#). The Kubeflow team is interested in your [feedback](#) about the usability of the feature.

This guide walks you through using MPI for training.

The MPI Operator makes it easy to run allreduce-style distributed training on Kubernetes. Please check out [this blog post](#) for an introduction to MPI Operator and its industry adoption.

MPI is MPI. It doesn't have to be "training". End of AI/ML speech.

**DETROIT 2022**



© Warner Brothers

**DETROIT 2022**

# Operators

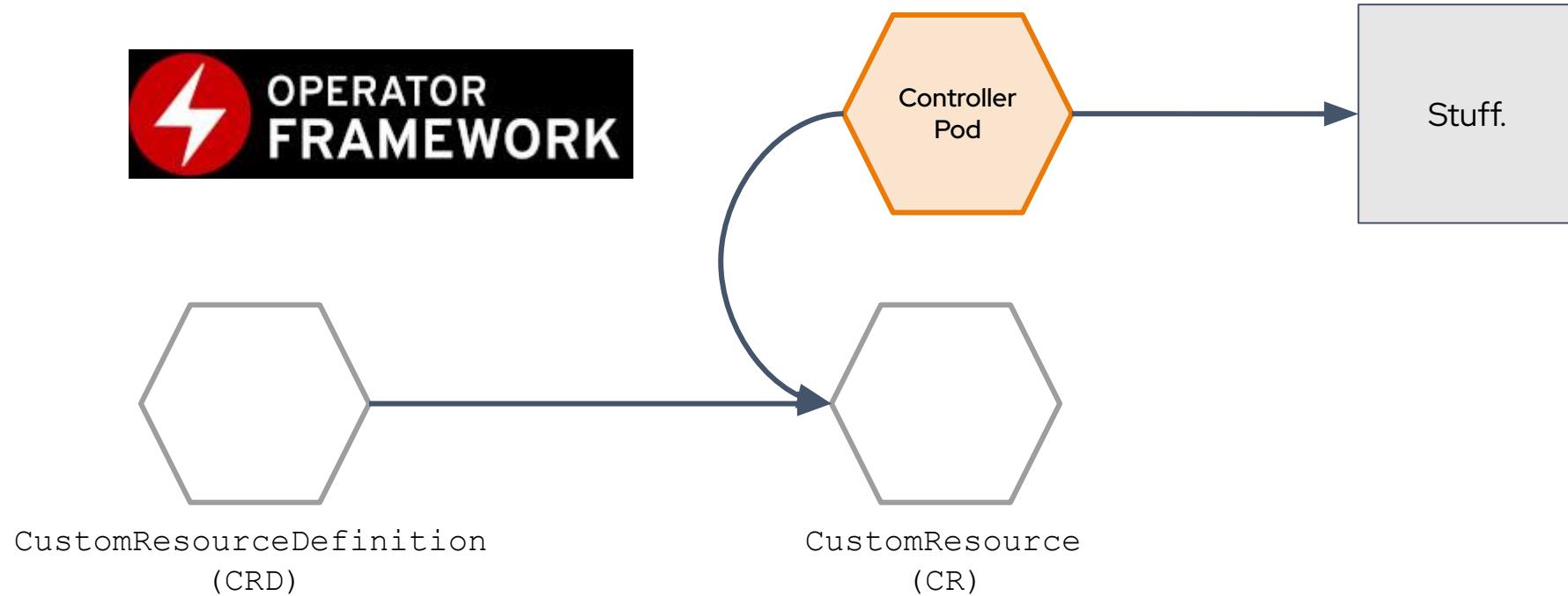


Operators implement and automate common Day-1 (installation, configuration, etc) and Day-2 (re-configuration, update, backup, failover, restore, etc.) activities in a piece of software running inside your Kubernetes cluster, by integrating natively with Kubernetes concepts and APIs.

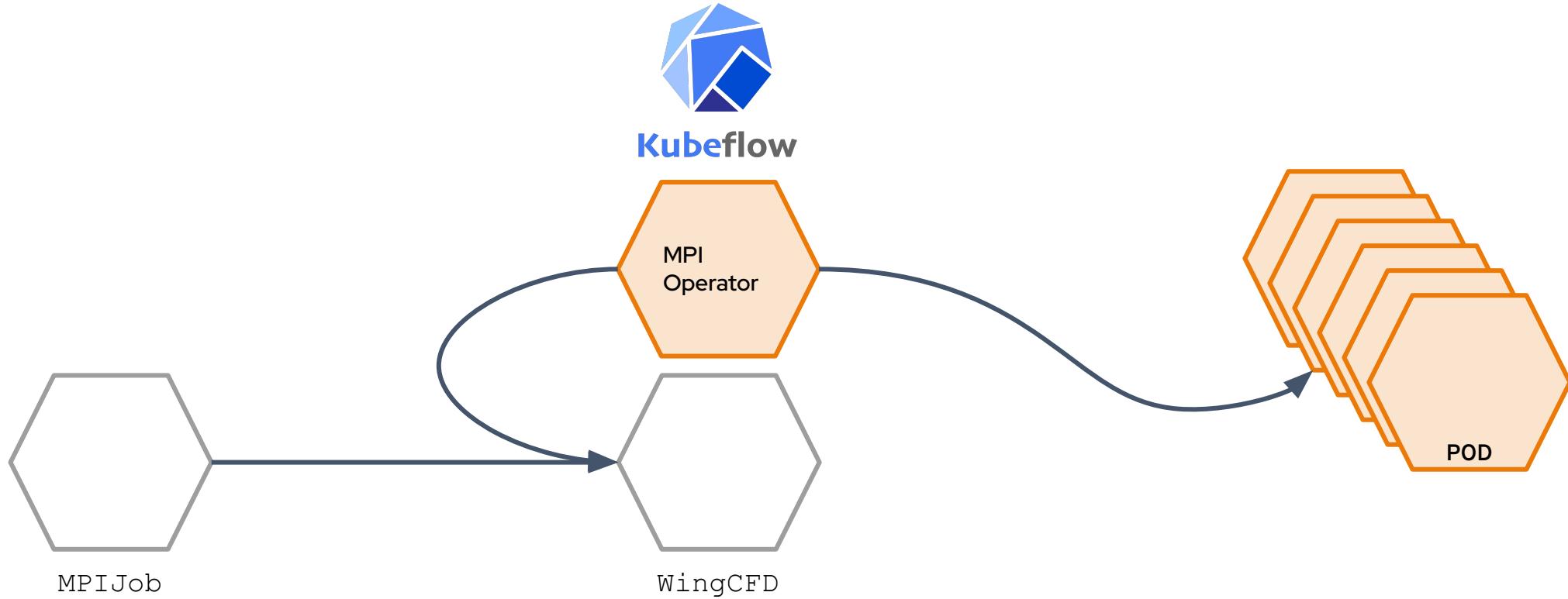


~~Operators implement and automate common Day 1 (installation, configuration, etc) and Day 2 (re-configuration, update, backup, failover, restore, etc.) activities in a piece of software running inside your Kubernetes cluster, by integrating natively with Kubernetes concepts and APIs.~~

# Operators



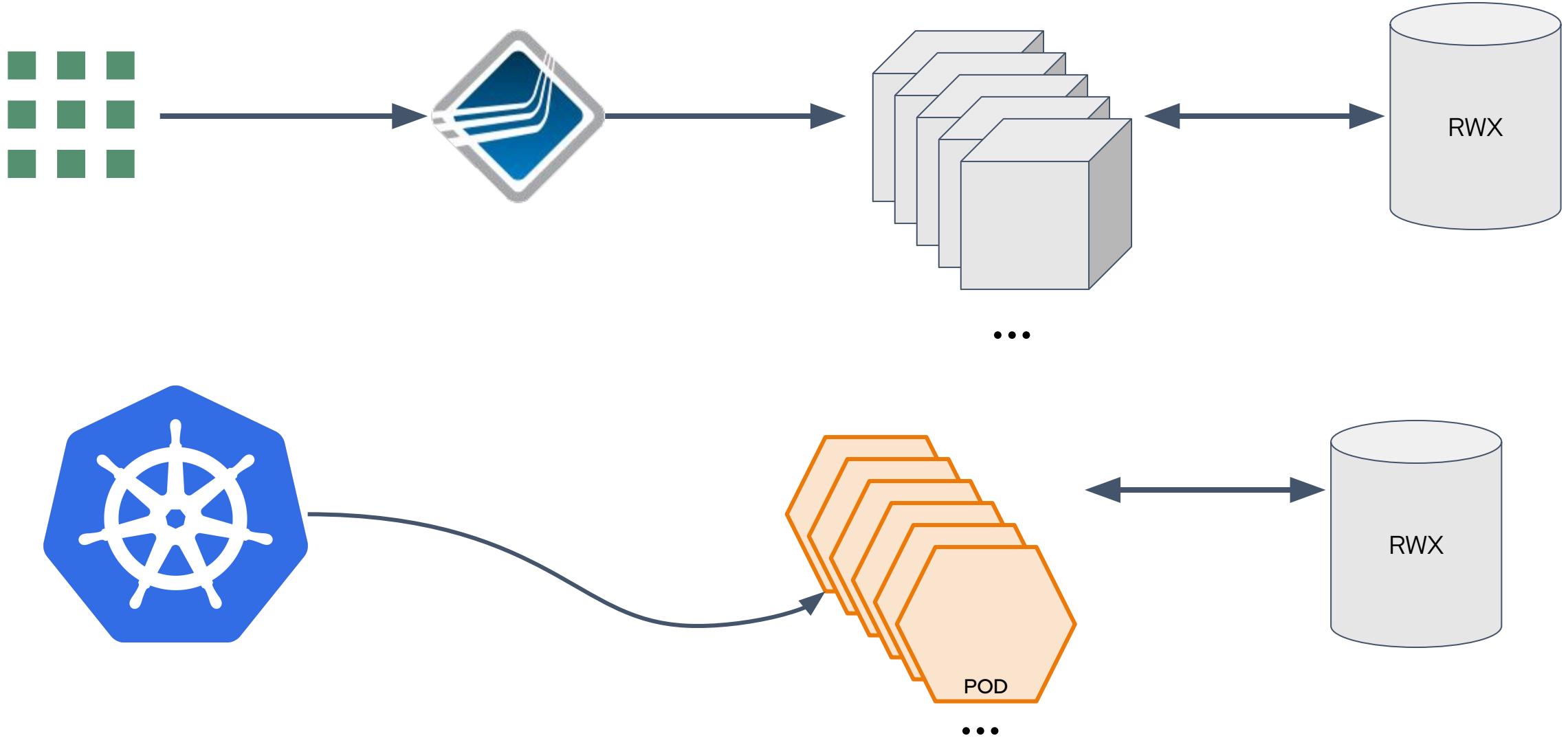
# Operators



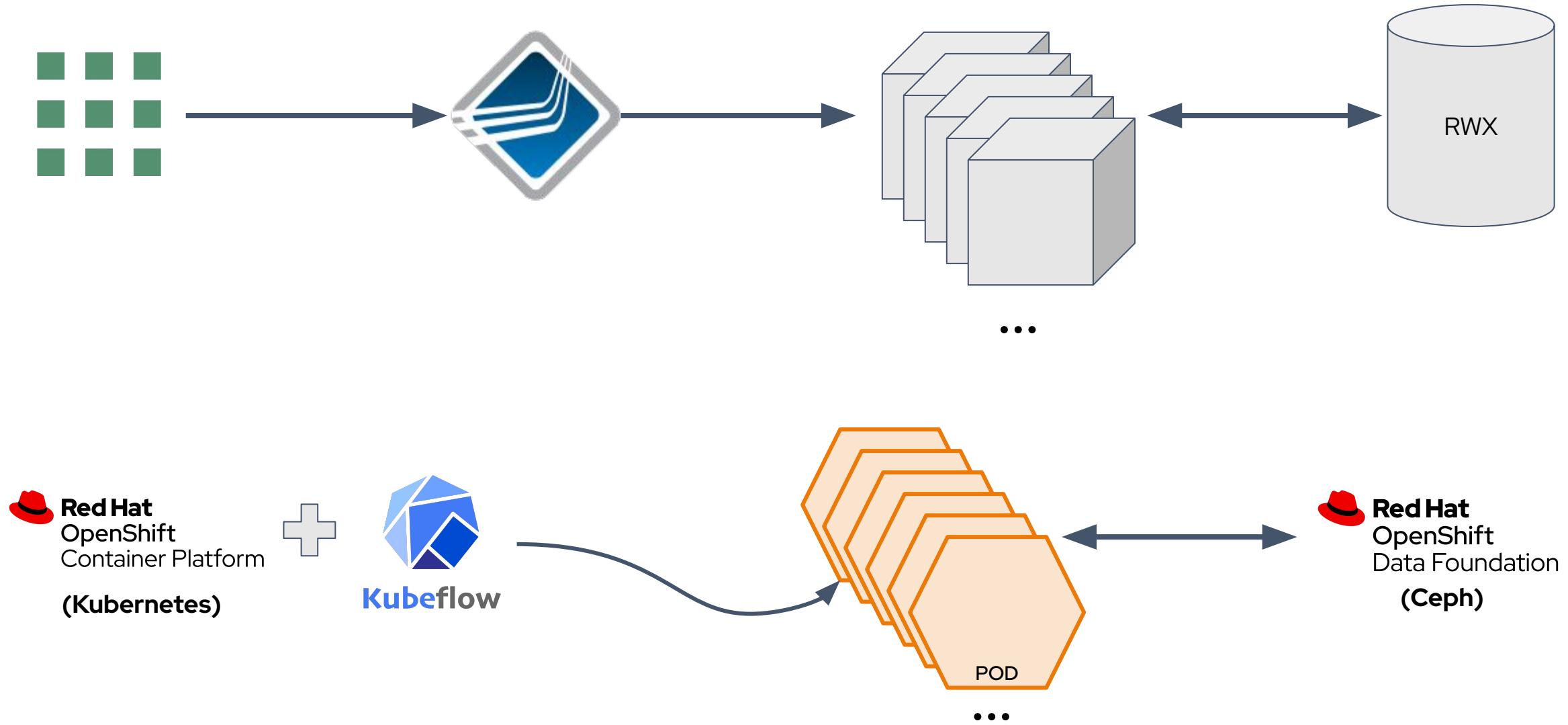
# DETROIT 2022



# Running CFD



# Running CFD

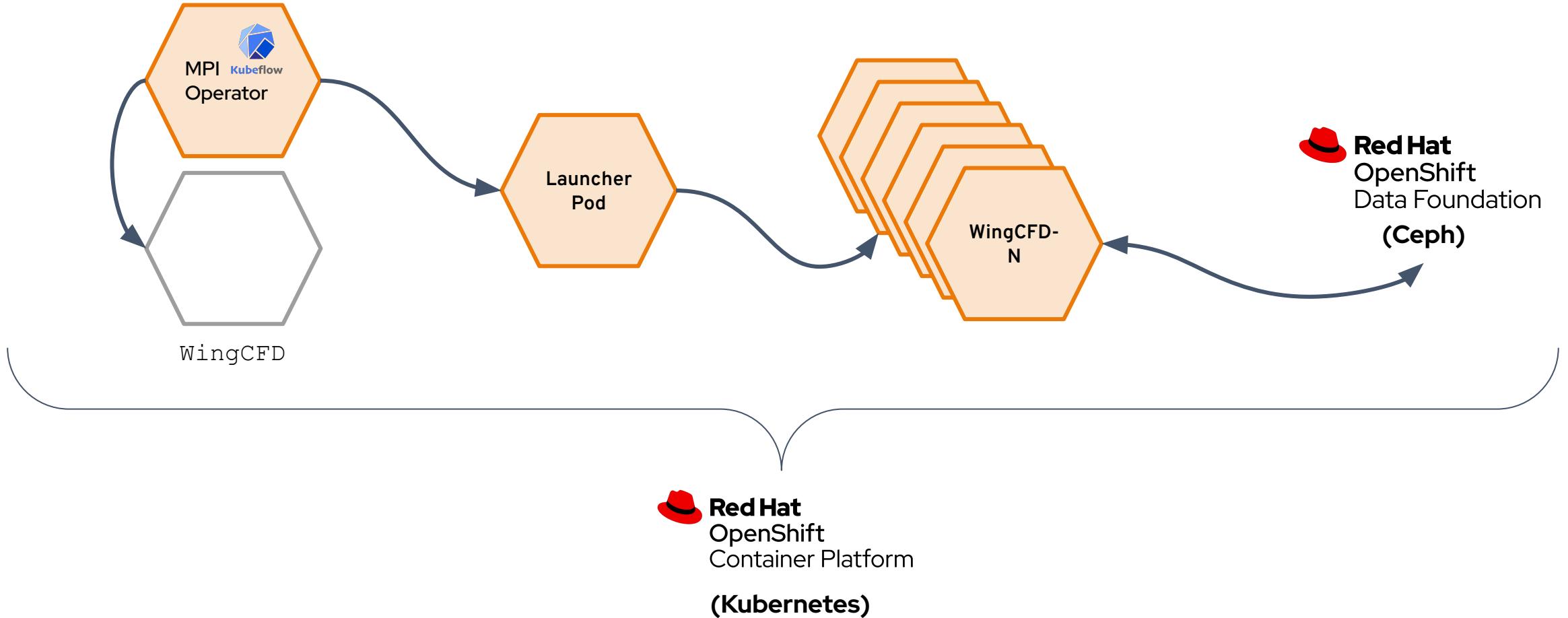


**DETROIT 2022**

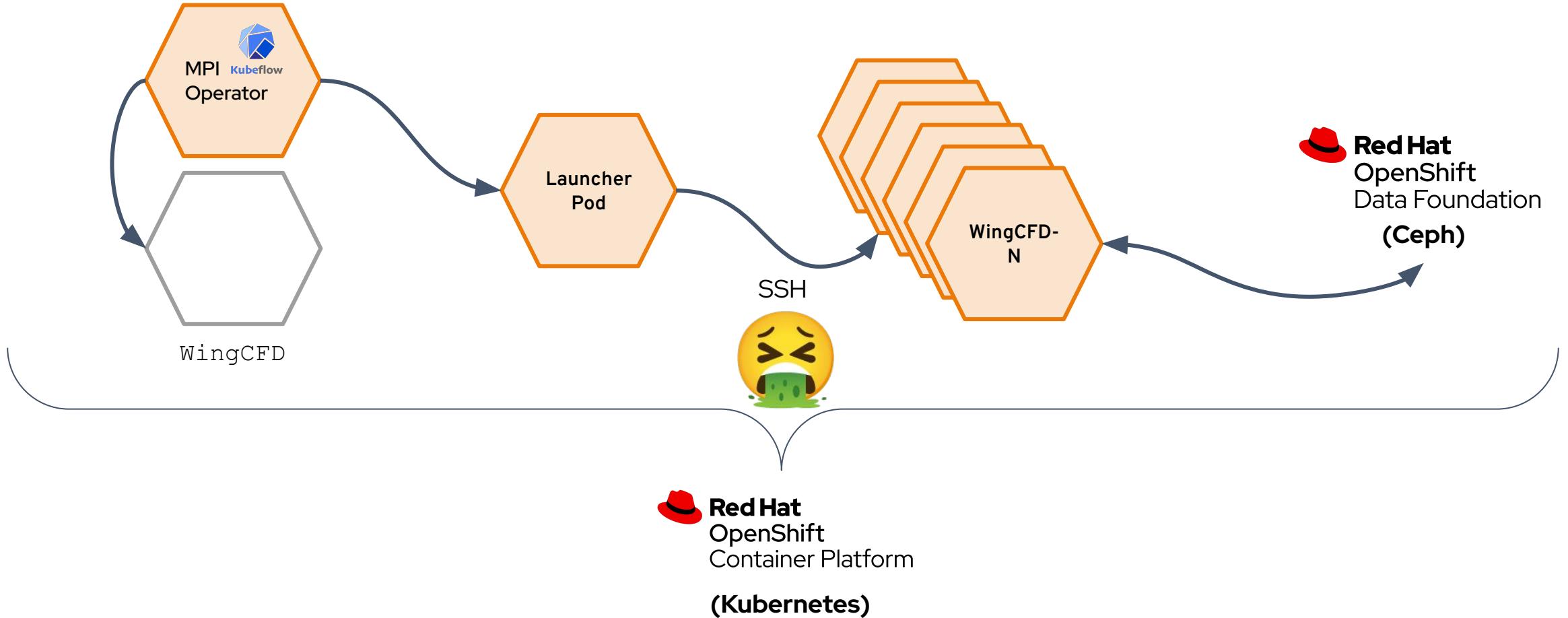


© Warner Brothers

# Running CFD



# Running CFD



**DETROIT 2022**



© Warner Brothers

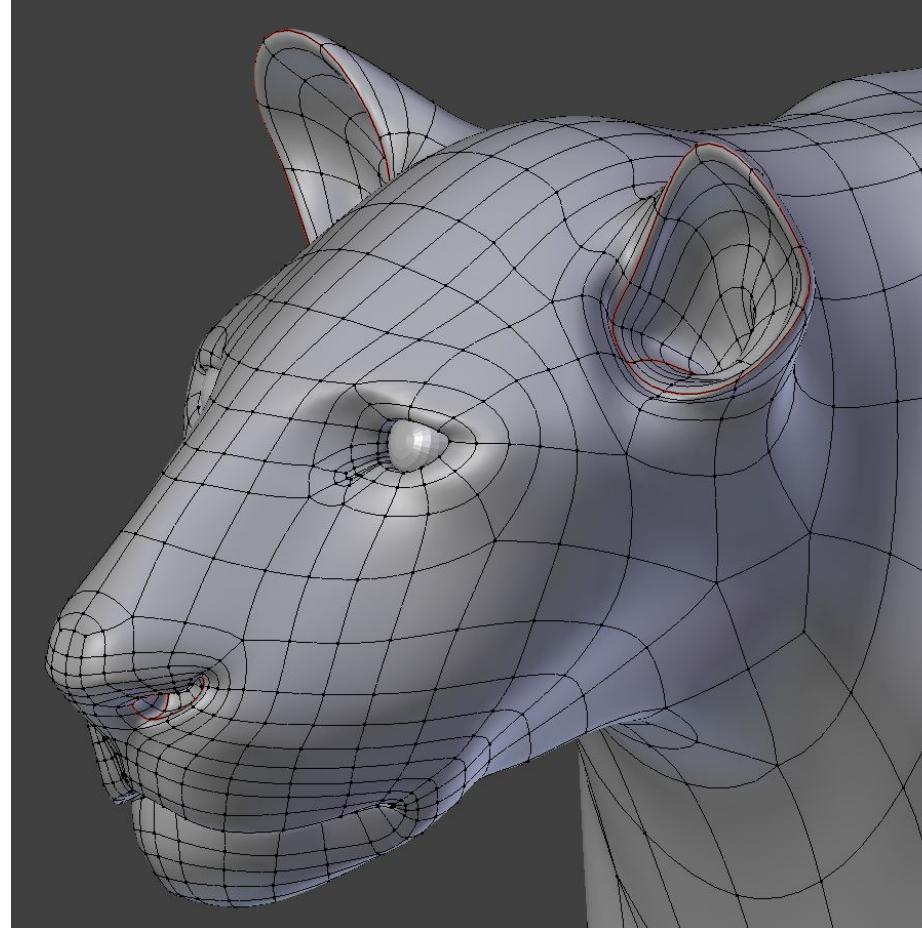
**DETROIT 2022**

# OpenFOAM Process

# DETROIT 2022

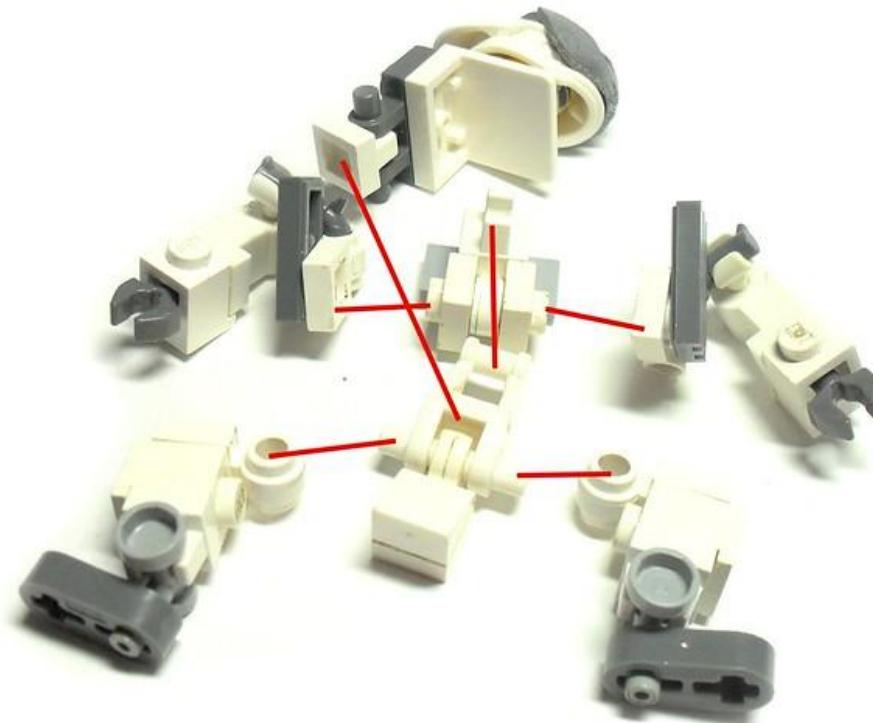


## Pre-processing: Meshing



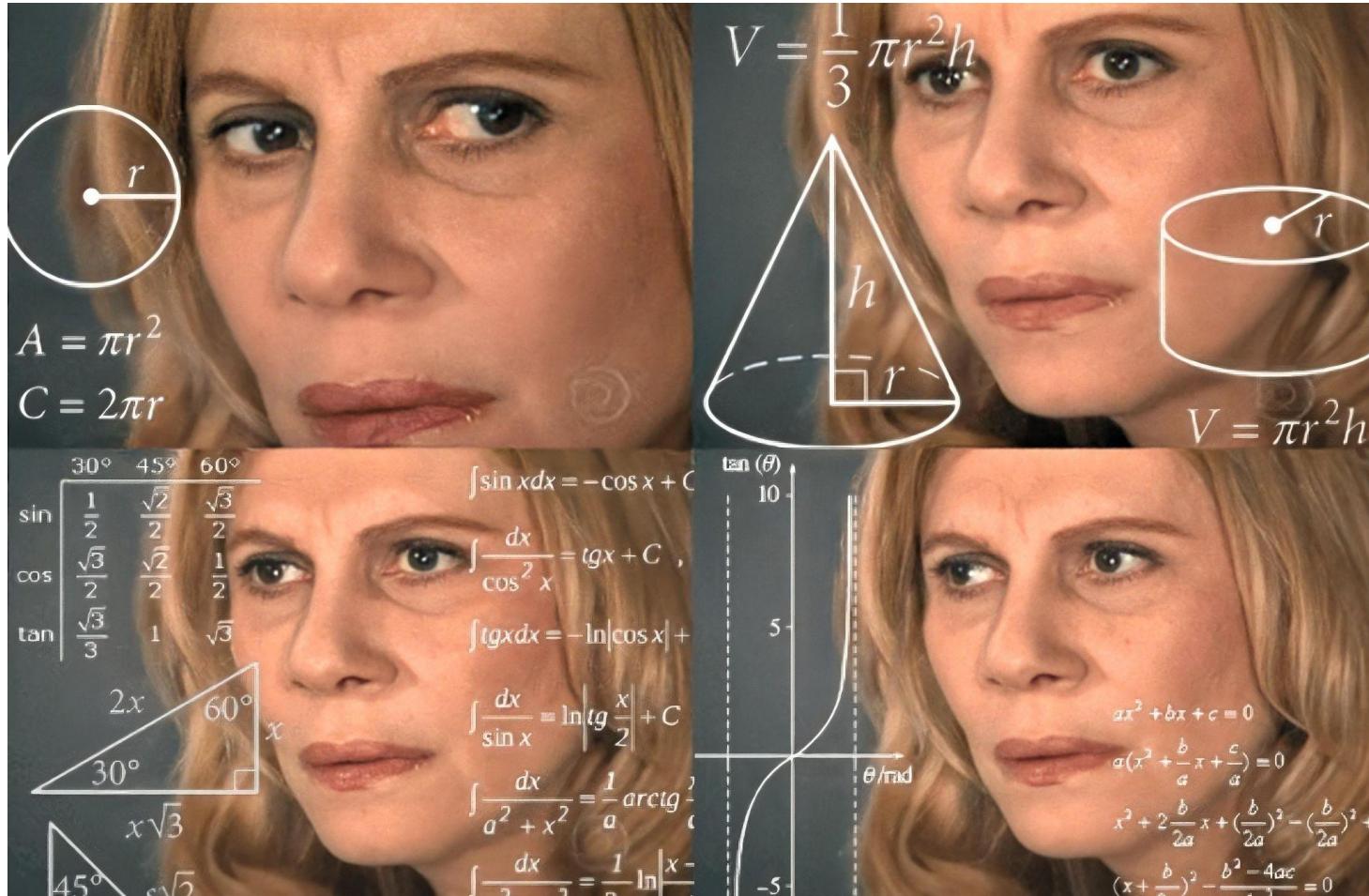
<https://www.flickr.com/photos/amagill/7059757185>

Pre-processing:  
Decomposition (for parallelization)



<https://www.flickr.com/photos/51917955@N08/5581064143>

## Processing: Solving



## Post-processing: Reconstituting



<https://www.flickr.com/photos/seanhobson/4380105315>

# DETROIT 2022



*Cutting corners to meet arbitrary management deadlines*



BUILDING FOR THE ROAD AHEAD

**DETROIT 2022**



*Essential*

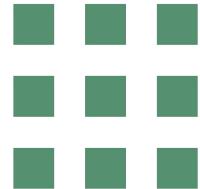
# Copying and Pasting from Stack Overflow

O'REILLY®

*The Practical Developer*  
@ThePracticalDev

*Cutting corners to meet arbitrary management deadlines*

**Tutorials!!1!!**

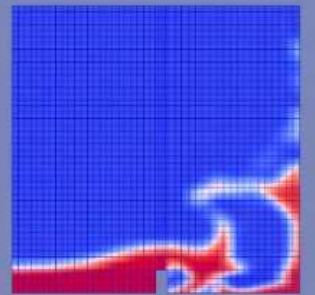


*Essential*

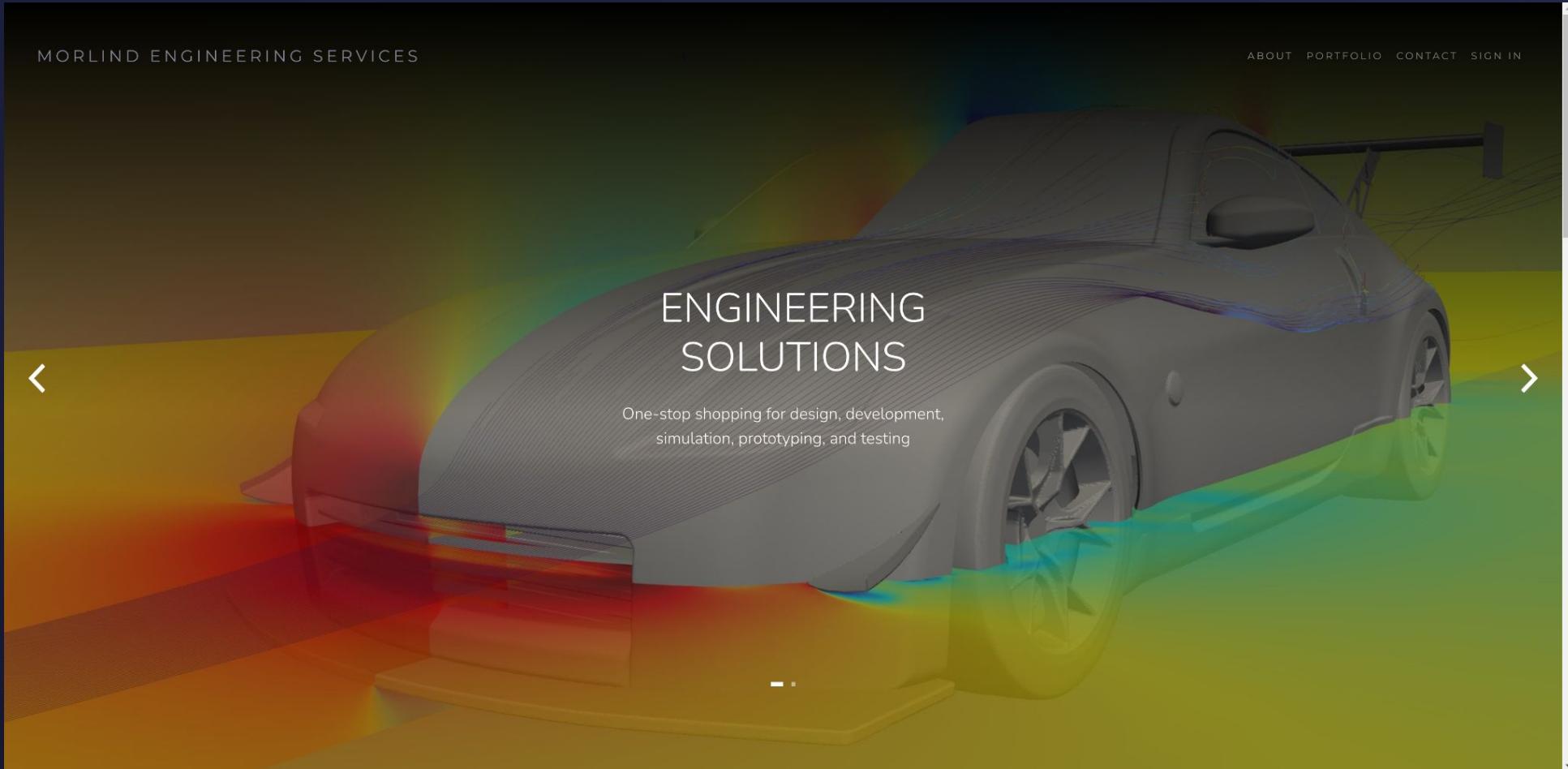
Copying and Pasting  
from Stack Overflow

O'REILLY®

*The Practical Developer*  
*@ThePracticalDev*



# DETROIT 2022



MORLIND ENGINEERING SERVICES

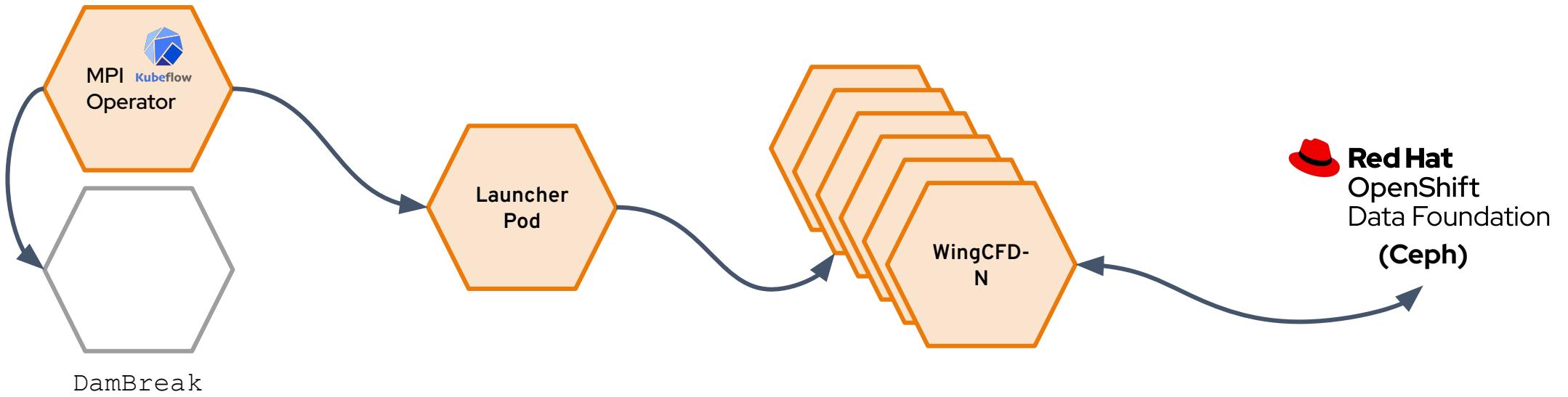
ABOUT PORTFOLIO CONTACT SIGN IN

**ENGINEERING  
SOLUTIONS**

One-stop shopping for design, development,  
simulation, prototyping, and testing

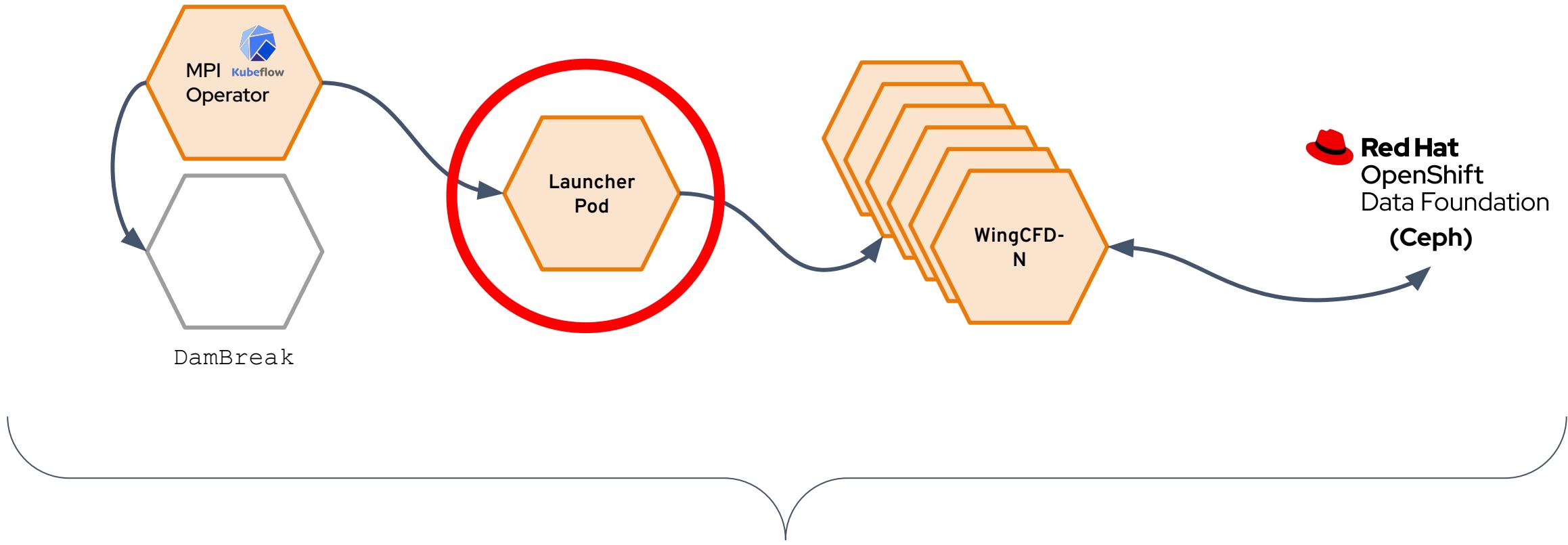
< >

# Running CFD



**Red Hat**  
OpenShift  
Container Platform  
**(Kubernetes)**

# Running CFD



**Red Hat**  
OpenShift  
Container Platform  
**(Kubernetes)**

BUILDING FOR THE ROAD AHEAD

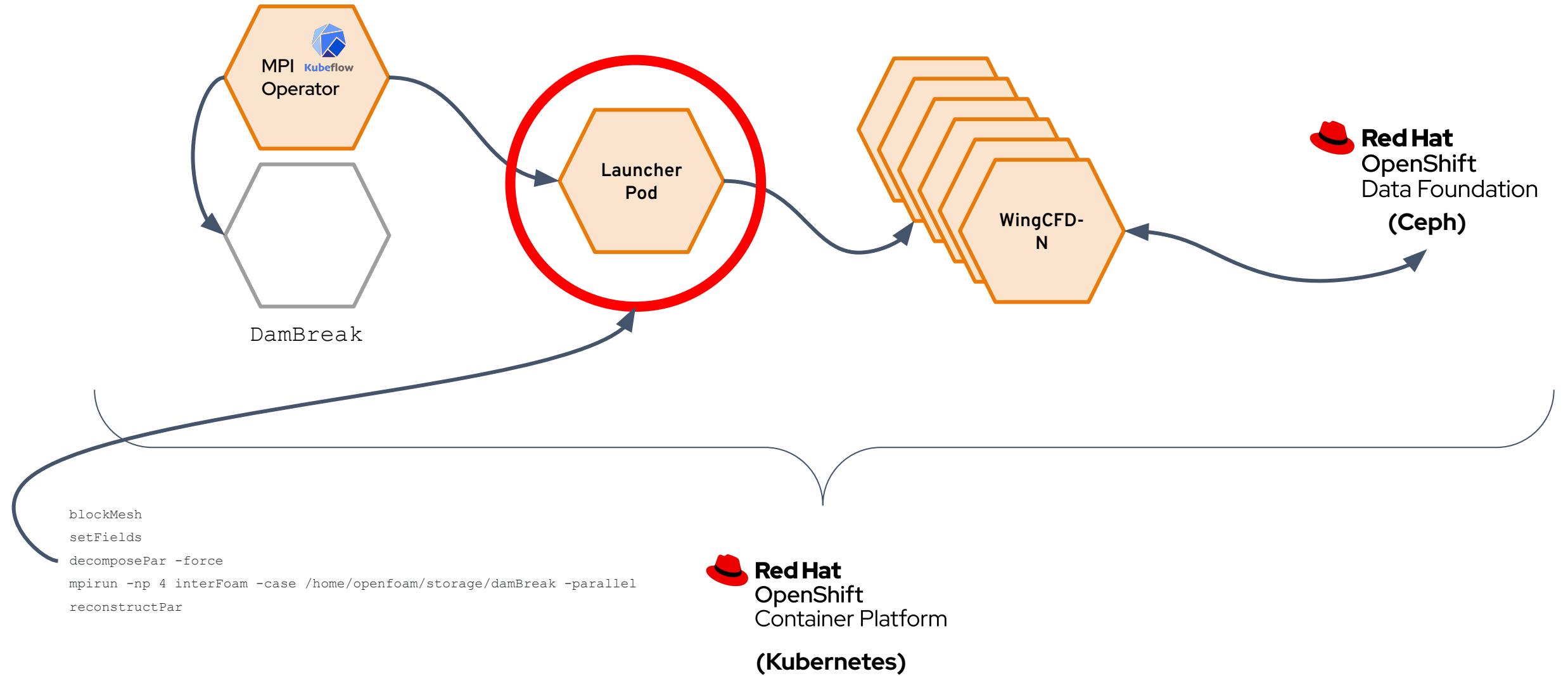
**DETROIT 2022**



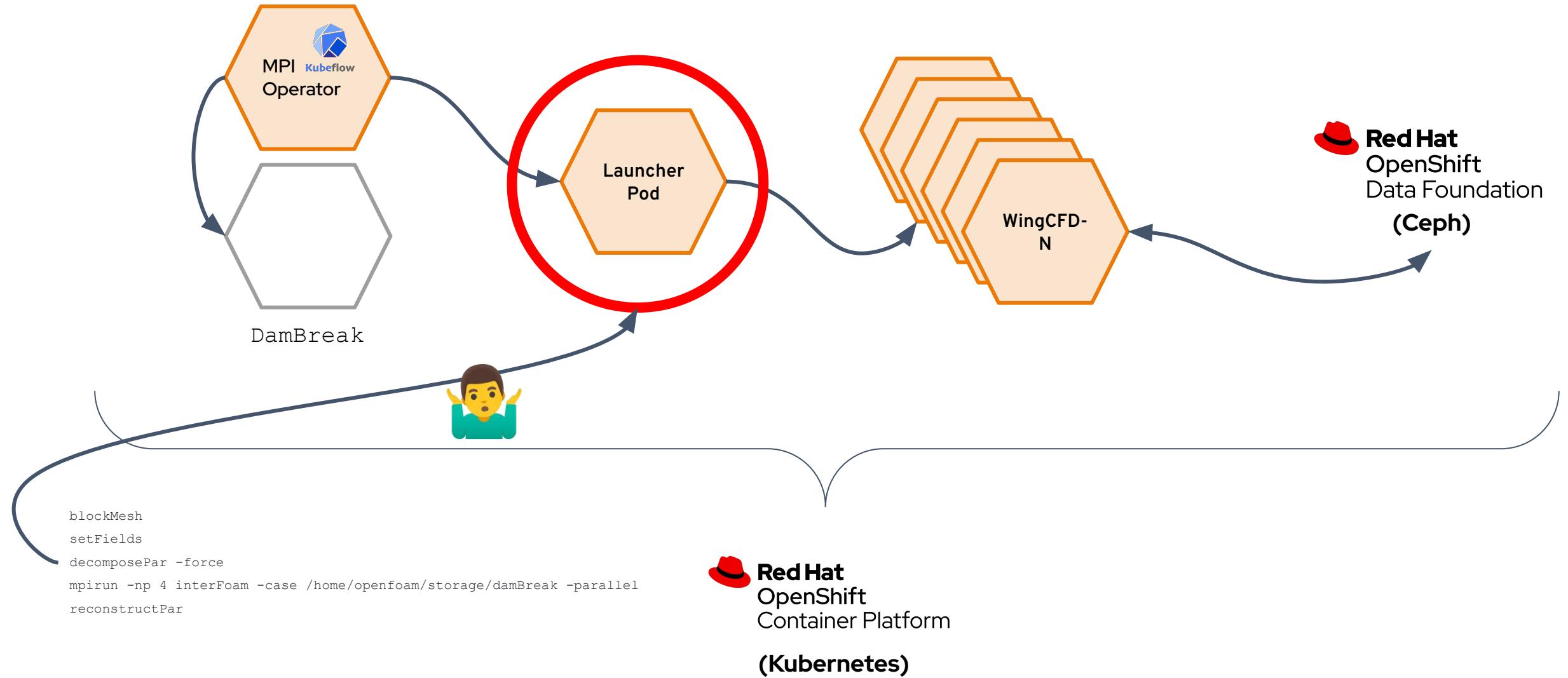
# damBreak Example

```
blockMesh
setFields
decomposePar -force
mpirun -np 4 interFoam -case /home/openfoam/storage/damBreak -parallel
reconstructPar
```

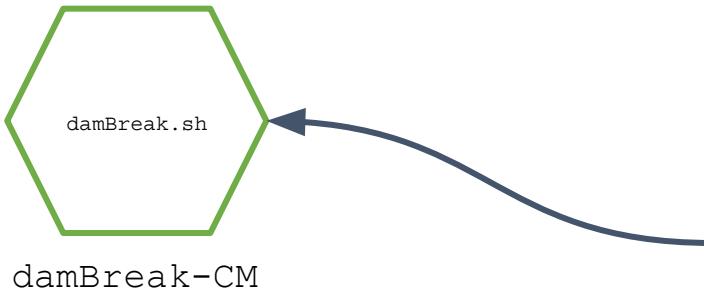
# damBreak Example



# damBreak Example

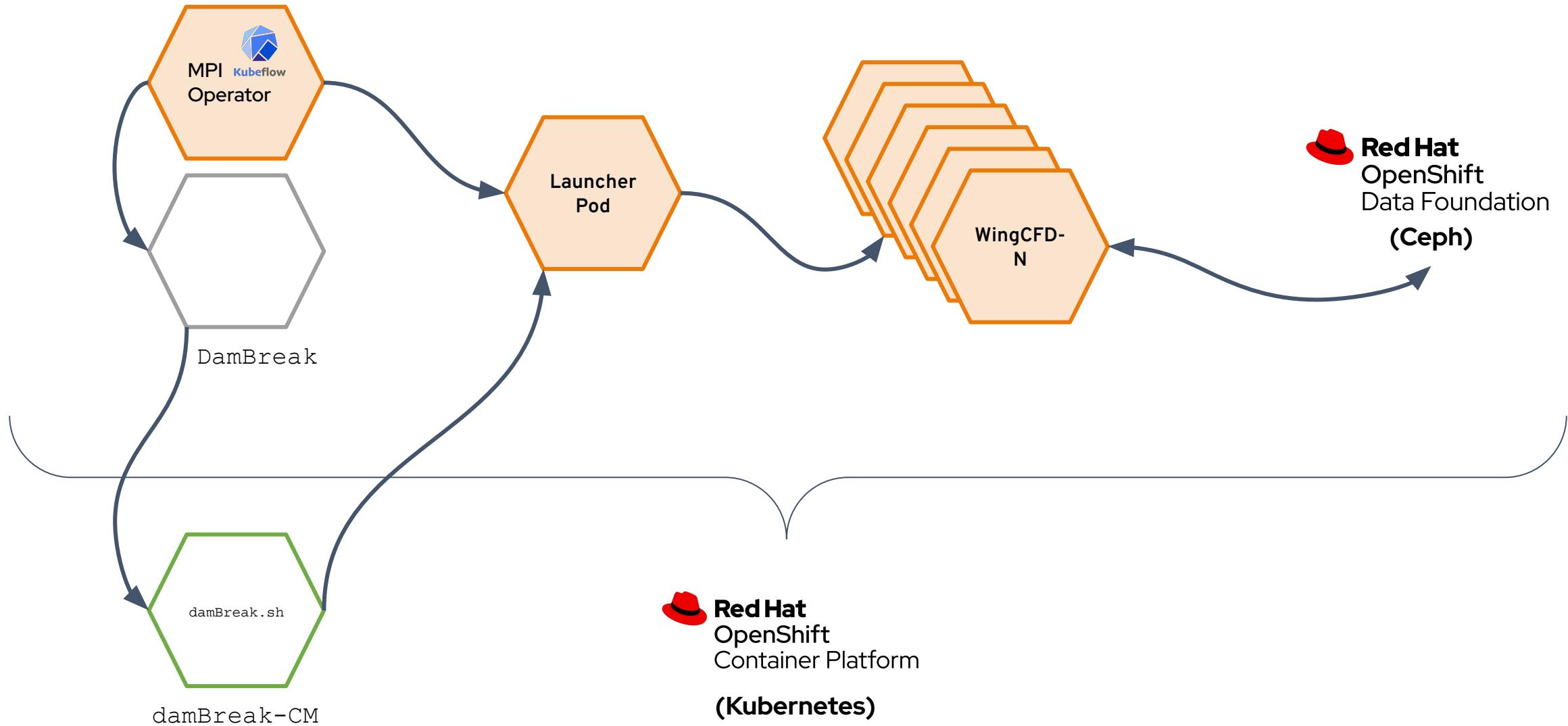


# damBreak Example



```
apiVersion: v1
kind: ConfigMap
metadata:
  name: dambreak-job
data:
  damBreak.sh: |
    #!/bin/bash
    echo "STARTING"
    date
    START=$(date +%s)
    source /opt/openfoam9/etc/bashrc
    cd /home/openfoam/storage/damBreak
    set -x
    blockMesh
    setFields
    decomposePar -force
    mpirun -np 4 interFoam -case /home/openfoam/storage/damBreak -parallel
    reconstructPar
    set +x
    date
    echo "ENDED"
    DURATION=$[ $(date +%s) - ${START} ]
    echo ${DURATION}
```

# damBreak Example



**DETROIT 2022**



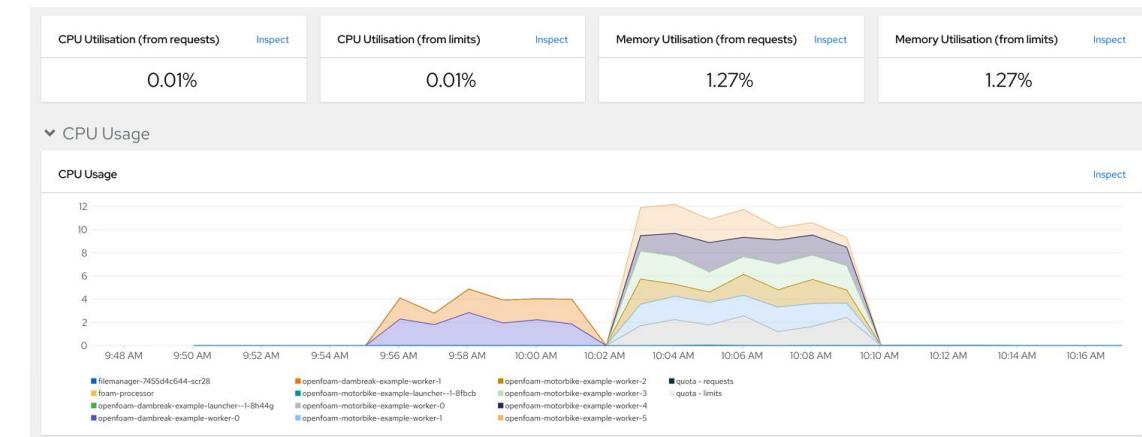
# It worked!

## Inside MPIJob execution pods

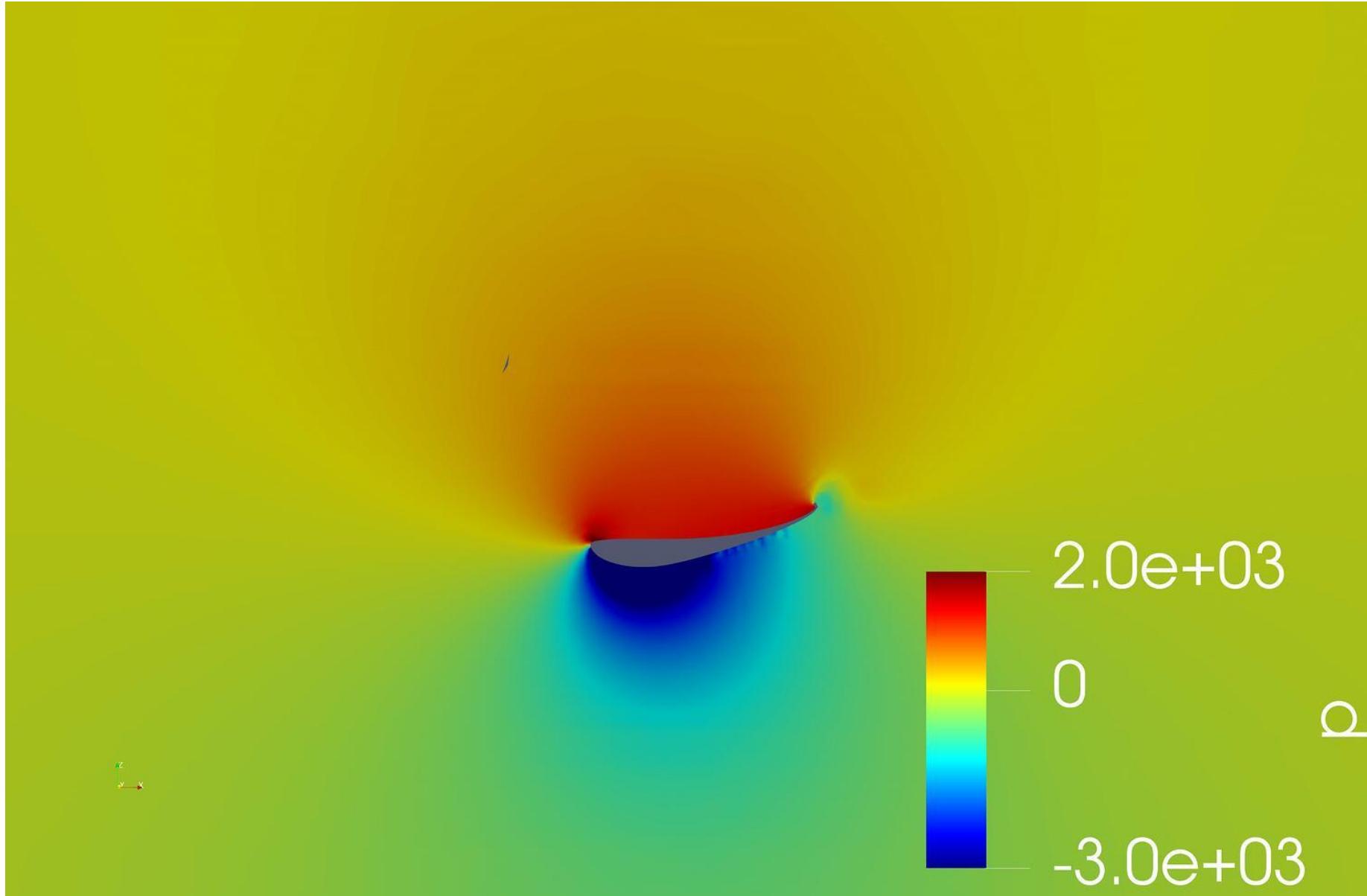
```
smoothSolver: Solving for alpha.water, Initial residual = 0.00193684, Final residual =
3.40587e-09, No Iterations 3
Phase-1 volume fraction = 0.124752 Min(alpha.water) = -2.76751e-09 Max(alpha.water) =
1
MULES: Correcting alpha.water
MULES: Correcting alpha.water
Phase-1 volume fraction = 0.124752 Min(alpha.water) = -2.76751e-09 Max(alpha.water) =
1
DICPCG: Solving for p_rgh, Initial residual = 0.0343767, Final residual = 0.0016133,
No Iterations 5
time step continuity errors : sum local = 0.000480806, global = -4.75907e-08,
cumulative = 7.68955e-05
DICPCG: Solving for p_rgh, Initial residual = 0.00181566, Final residual =
8.38192e-05, No Iterations 33
time step continuity errors : sum local = 2.42908e-05, global = 3.04705e-06, cumulative
= 7.99426e-05
DICPCG: Solving for p_rgh, Initial residual = 0.000251337, Final residual =
8.59604e-08, No Iterations 109
time step continuity errors : sum local = 2.49879e-08, global = -9.43014e-10,
cumulative = 7.99416e-05
ExecutionTime = 22.2 s ClockTime = 23 s
```

```
Courant Number mean: 0.076648 max: 1.01265
Interface Courant Number mean: 0.00417282 max: 0.936663
deltaT = 0.0010575
Time = 0.11616
```

## OpenShift User Interface



# Real World Things!



**DETROIT 2022**



**DETROIT 2022**





<https://www.flickr.com/photos/torkildr/3462607995>

# Not Reality!



<https://www.flickr.com/photos/torkildr/3462607995>

# Reality!\*



<https://www.flickr.com/photos/kboey/4363850853>



<https://www.flickr.com/photos/torkildr/3462607995>

!=



<https://www.flickr.com/photos/kboey/4363850853>

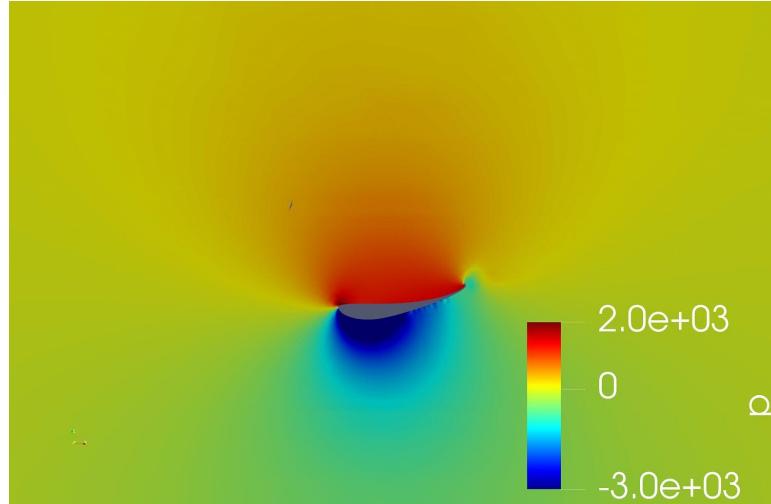


Really Reality.



# DETROIT 2022





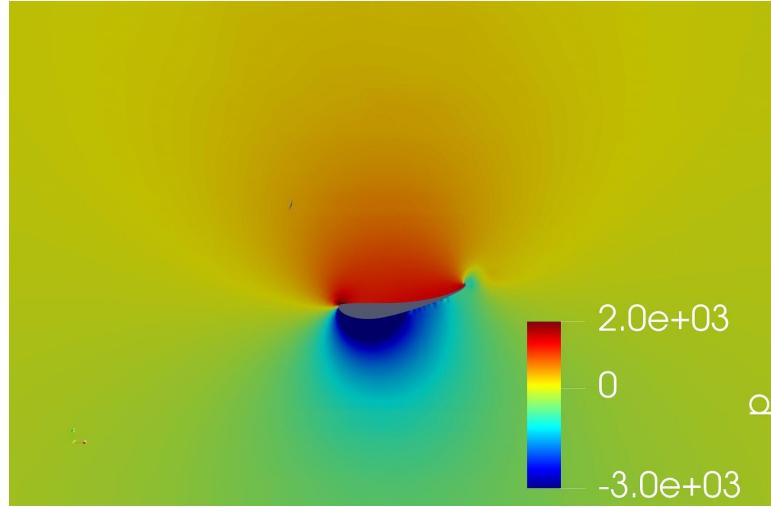
- Amazon m5a.4xlarge (not compute optimized)
- 2 CPUs per Pod
- 256 Pods
- 512 total CPUs
- Took about 5 hours
- **Way slower** than Morlind's on-premise small cluster

# Tuning! Tuning! Tuning!



<https://www.flickr.com/photos/andrewgustar/8548322749>

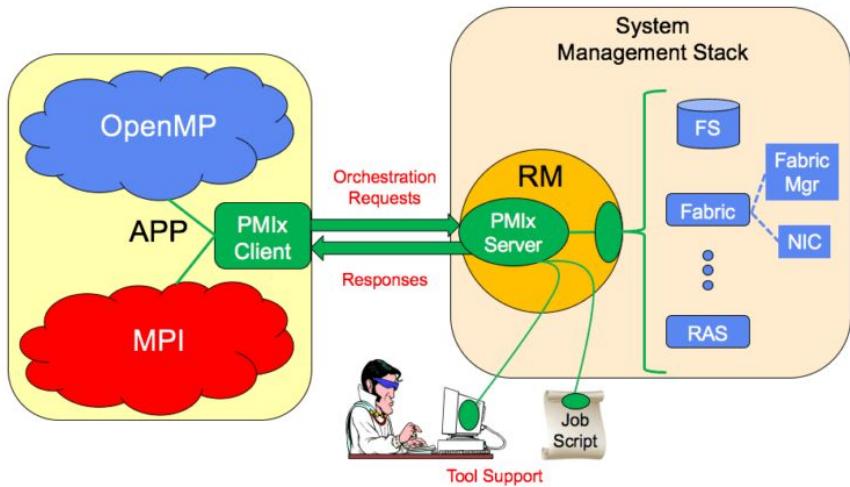
# Key Takeaways



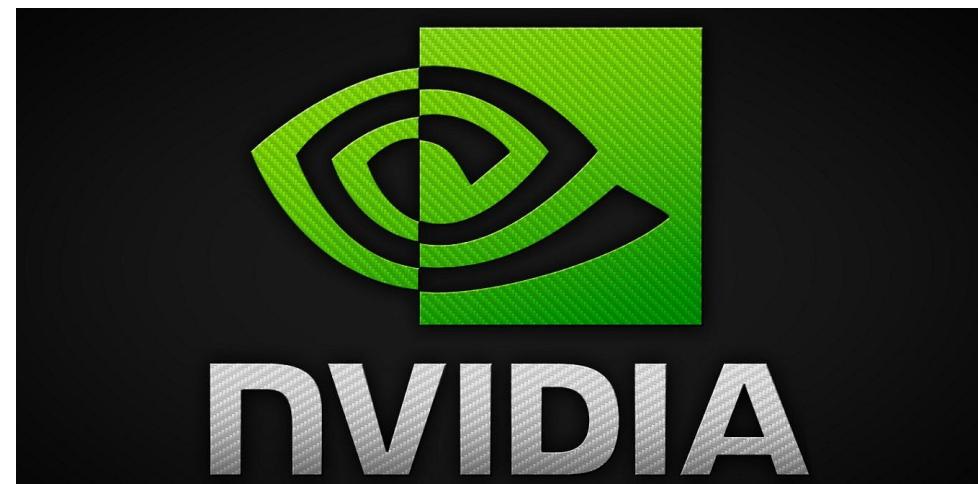
- Optimize instance type
- Probably (?) maximize pod CPU aligned to instance type
- Don't over-slice jobs (not too many solvers)
- **Know what you're doing**

# Where do we go from here?

## PMIx



<https://github.com/kubeflow/mpi-operator/issues/12>  
SSH 😞



# What's the point?

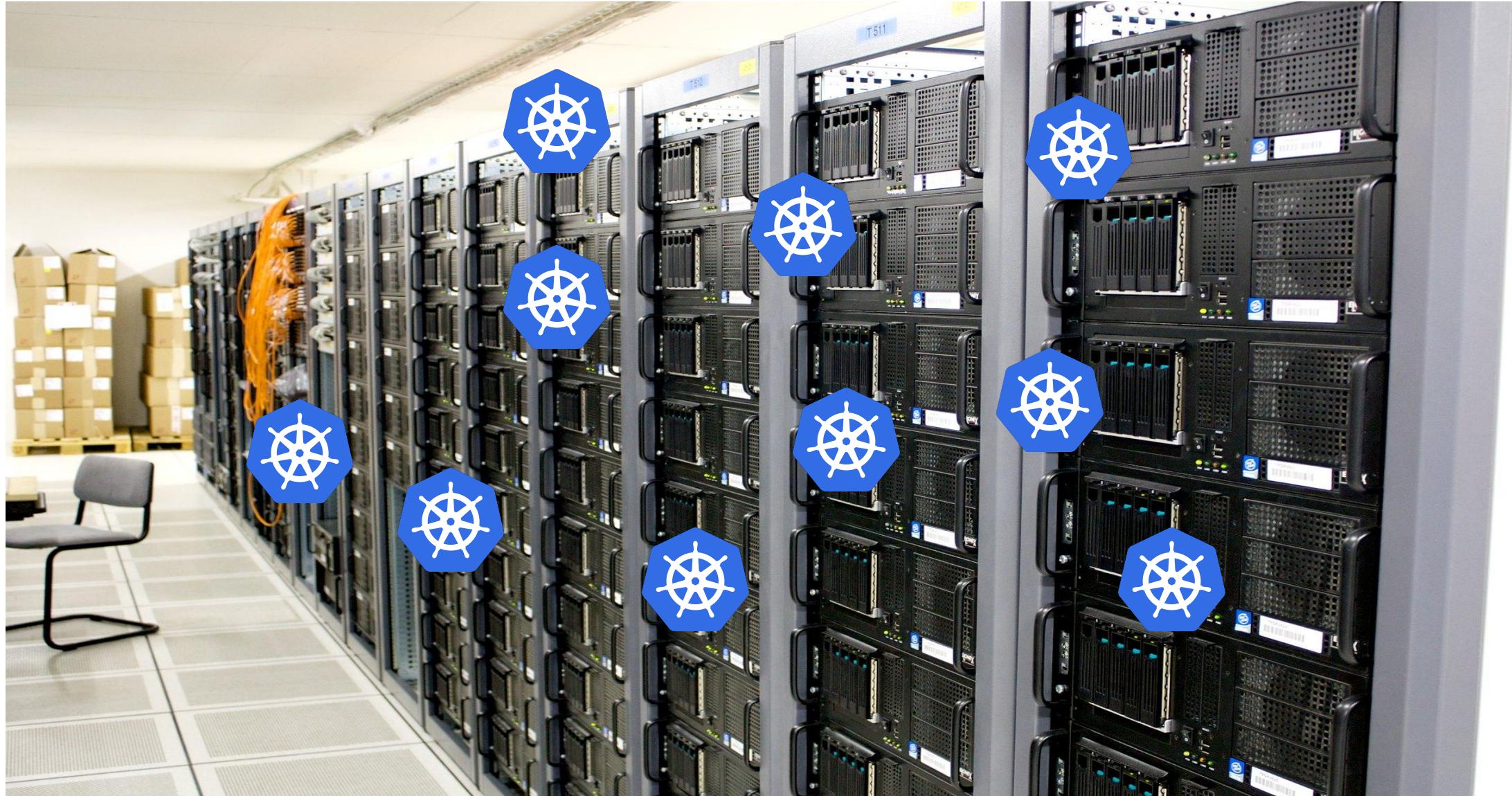


© Warner Brothers

What's the point?

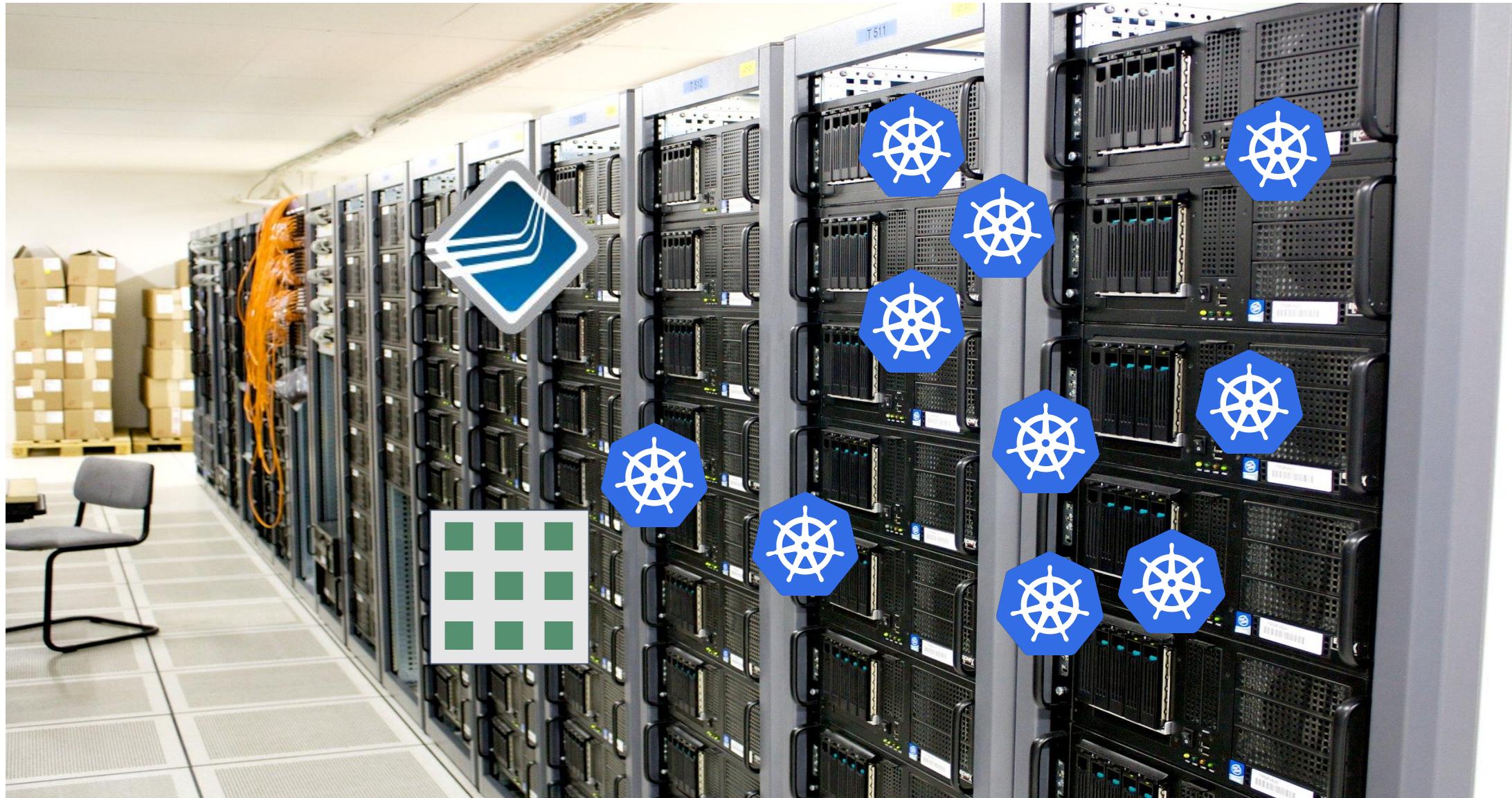


# What's the point?



<https://www.flickr.com/photos/torkildr/3462607995>

# What's the point?



<https://www.flickr.com/photos/torkildr/3462607995>



Please scan the QR Code above to  
leave feedback on this session