CASC

Center for Applied Scientific Computing

Autonomous Multiscale Simulations – Embedded Machine Learning for Smart Simulations

Peer-Timo Bremer

Center for Applied Scientific Computing

Center for Extreme Data Management Analysis and Visualization

CEDMAV

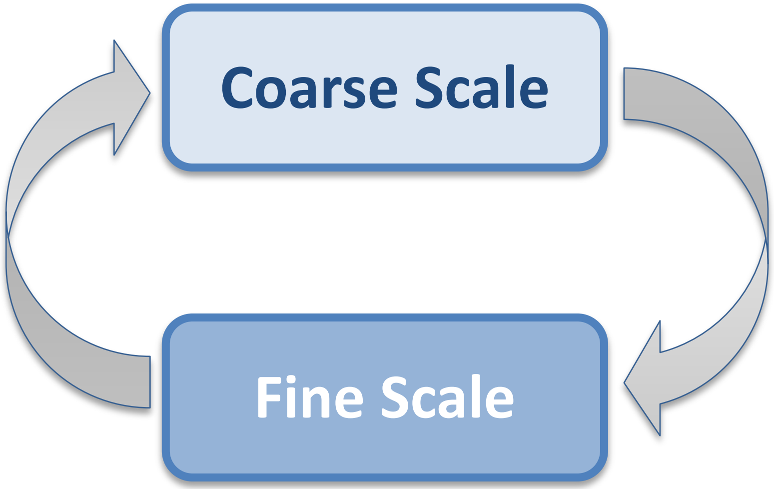
LLNL-PRES-844010

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344. Lawrence Livermore National Security, LLC

Lawrence Livermore National Laboratory

Predictive Simulations have Become the Third Branch of Science and are Key for Hypothesis Generation, Exploration, and Validation

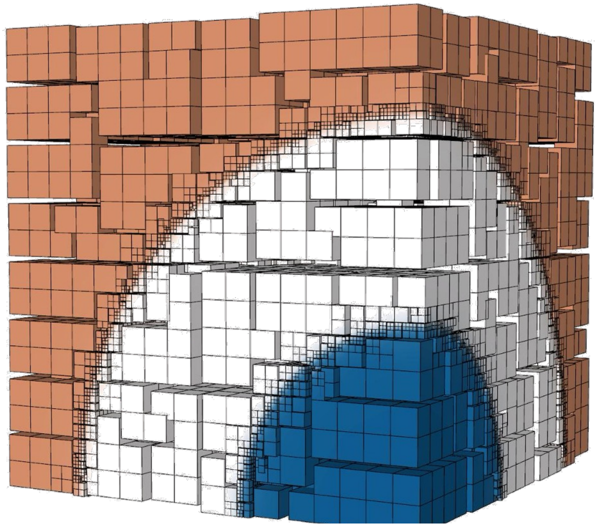
- In most cutting edge science applications we face a choice between
  - Coarse scale simulations:
    - ☑ Device scale simulations
    - ☑ Enable design optimization and UQ
    - ☒ Cannot represent all salient features
    - ☒ Do not resolve all necessary physics
  - Fine scale simulations:
    - ☑ Can represent any important feature
    - ☑ Resolve all necessary (known) physics
    - ☒ Are severely limited in size
    - ☒ Are exceedingly expensive

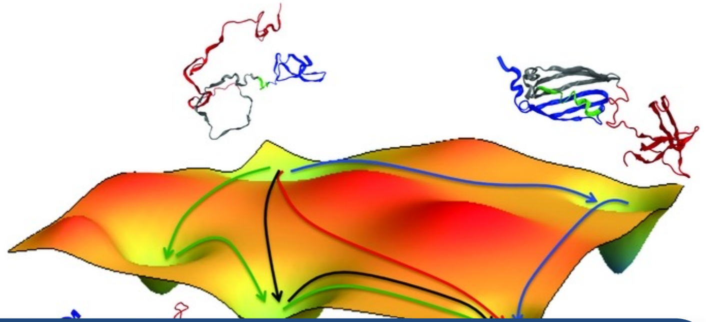


Multiscale simulations combine the best of both world

Multiscale is a Widely Used Term With Different Meanings in Different Communities and With a Long History of Ideas

- Multi-grid techniques explicitly express simulations hierarchically to cover longer length- and timescales
  - Solve the same equations faster by focusing on key degrees of freedom
- Accelerated molecular dynamics explore multiple timescales by modifying the energy landscape or approximate force fields
  - Specific solution for the sampling problem of MD simulations
- Similar solutions depend on unified representations that fundamentally do not bridge scales
  - Coarse scale equations are based on coarse scale physics
  - Fine scale representations grow beyond reasonable bounds



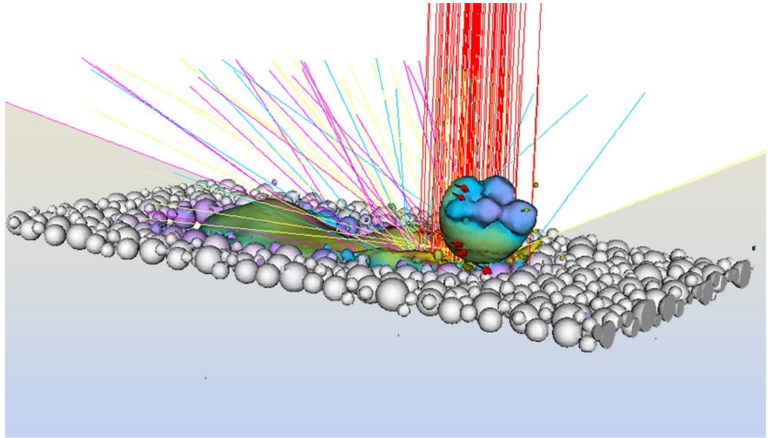


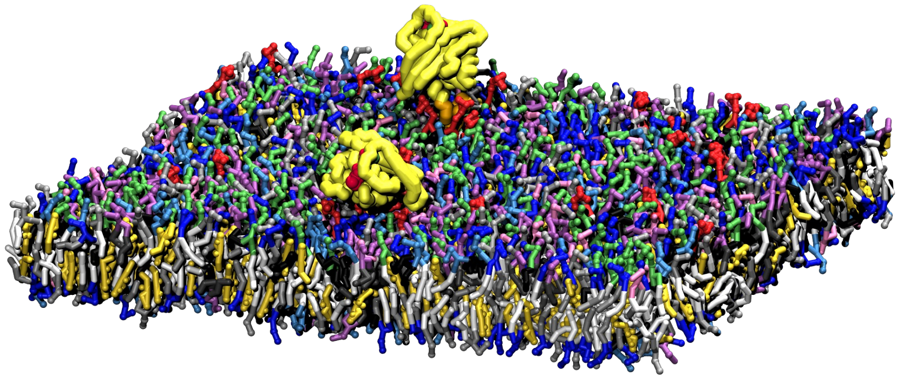
Many of the most interesting problems require coupling different representations

https://cen.acs.org/articles/90/i29/Simulations-Peg-Protein-Folding.html

Different Representations can Cover Different Physics, Different Length Scales, Different Time Scales, etc., and can be Nested

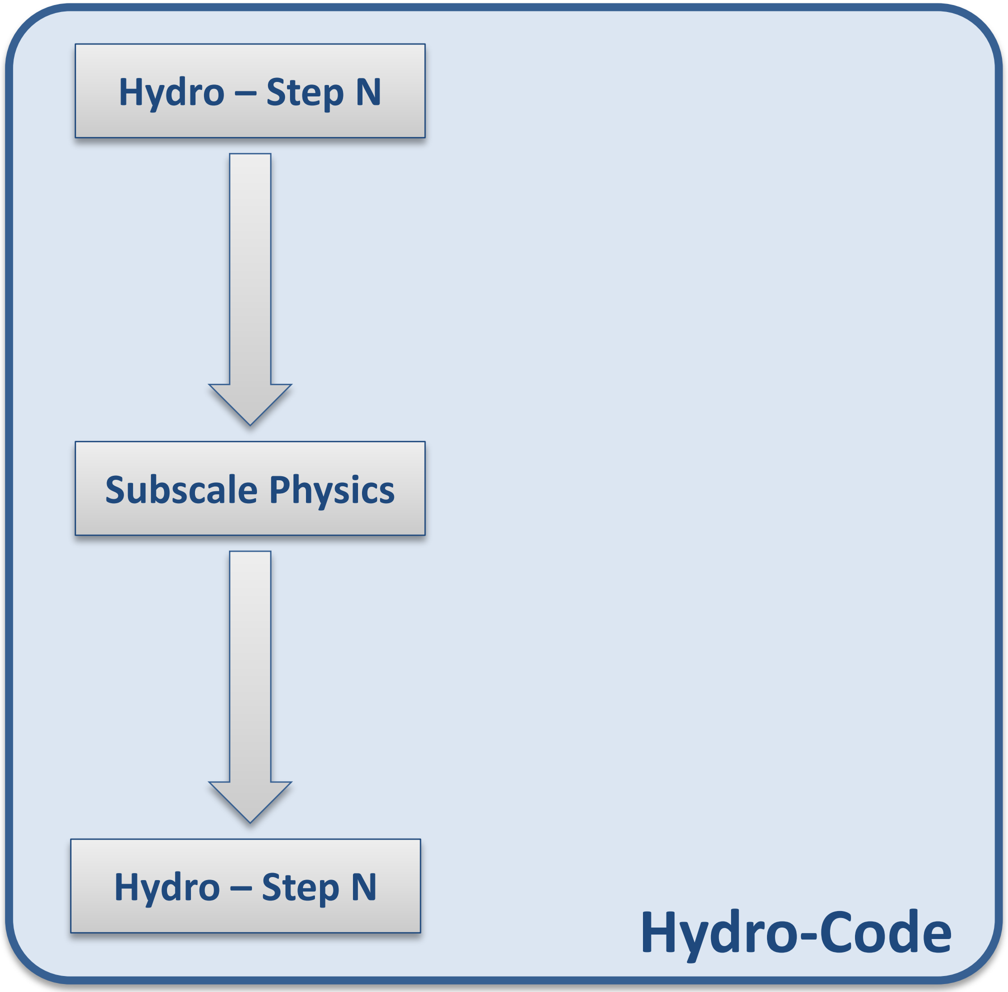
- Melt pools in manufacturing:
  - Continuum ⟺ Phase field
- HE materials
  - Homogeneous material ⟺ grain scale responses
- NIF holhraums
  - Hydrodynamics ⟺ Atomic energy model
- Cancer signaling chain
  - Continuum membrane ⟺ Bead level MD ⟺ Atomistic
- Climate simulations
  - Ocean modeling ⟺ ice sheets





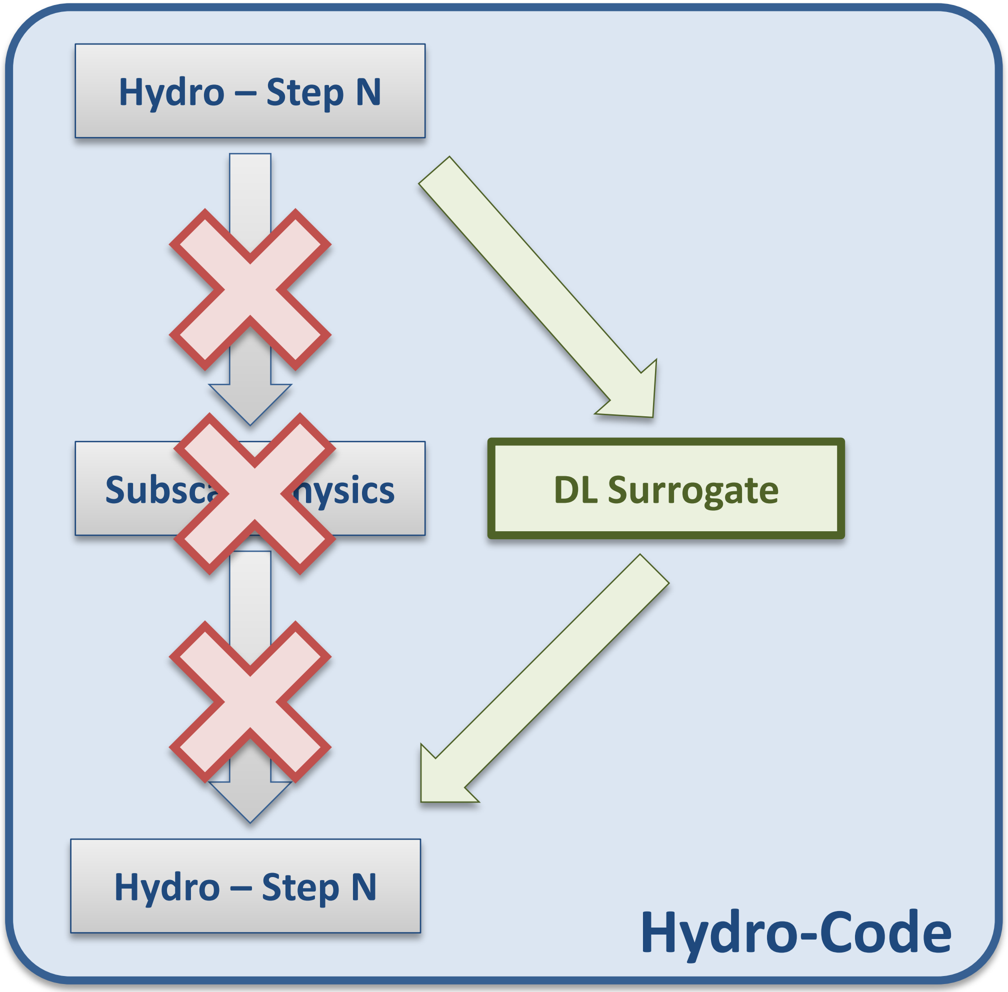
Many Physics Simulations Directly Couple Either Multiple Scales or Multiple Types of Physics to Improve Accuracy

- Hydrodynamics vs. radiation transport
- Transport vs. reaction
- One particularly common pattern are “subgrid” models:
  - At each coarse time-step
  - At each mesh point
  - Call fine-scale physics (EOS, kinetics, rad. transport , …)
- Depending on the fidelity subgrid models quickly dominate the run time



Deep Learning is Enabling us to Build Surrogate Models with Arbitrary Inputs and Outputs to Replace the Fine Scale Solution

- Advantages:
  - Massive expected improvements in performance
- Challenges:
  - Collect training data
  - Guarantee sufficient accuracy
  - Report potential failures
- Existing Solution
  - Execute a simulation of interest with the original physics to collect data
  - Train reliable surrogate model
  - Execute new simulation with surrogate model
  - Check for coverage and accuracy of the model
    - If problems are found – **REPEAT** until convergence

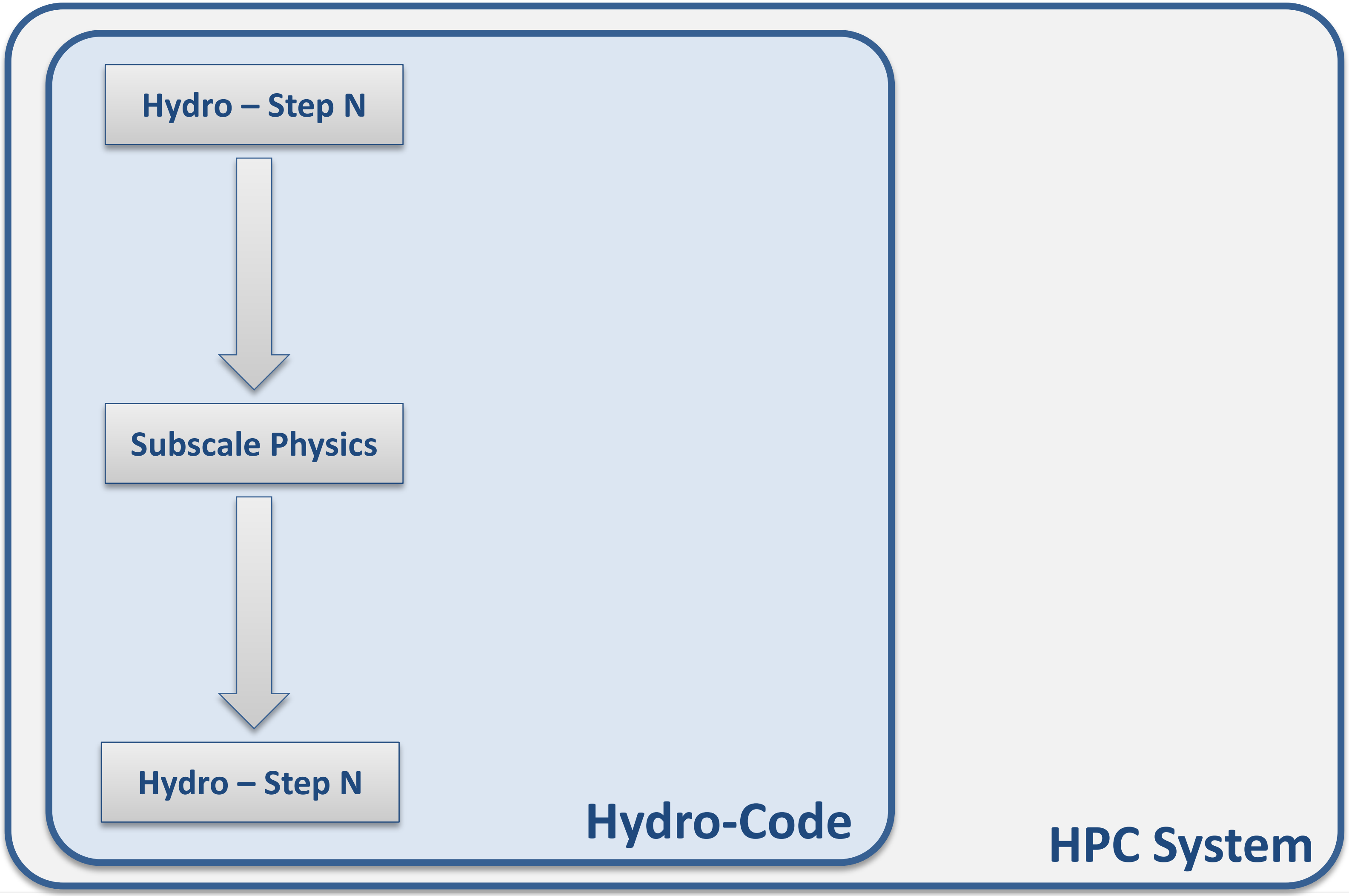


Autonomous Multiscale Simulations (AMS) Describe a Next Generation Multiscale Framework to Address these Problems

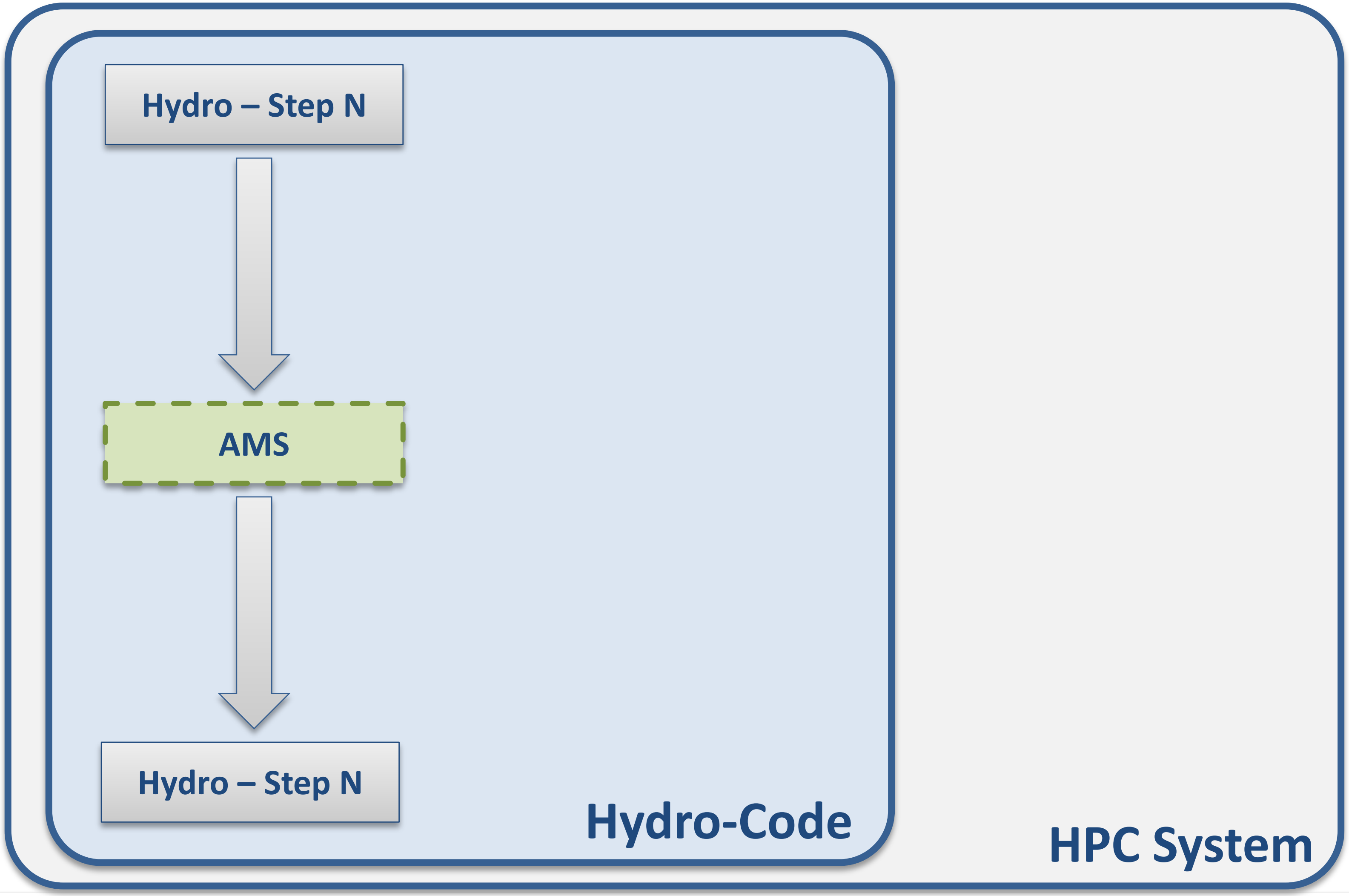
- Surrogate models without the need for exhaustive training data
- Reliable models with well vetted defaults
- Persistent and continuously improving models
- Ability to exploit ensemble calculations
- Potential for nested multiscale without exploding system complexity

Warning – Work in Progress

Autonomous Multiscale Simulations

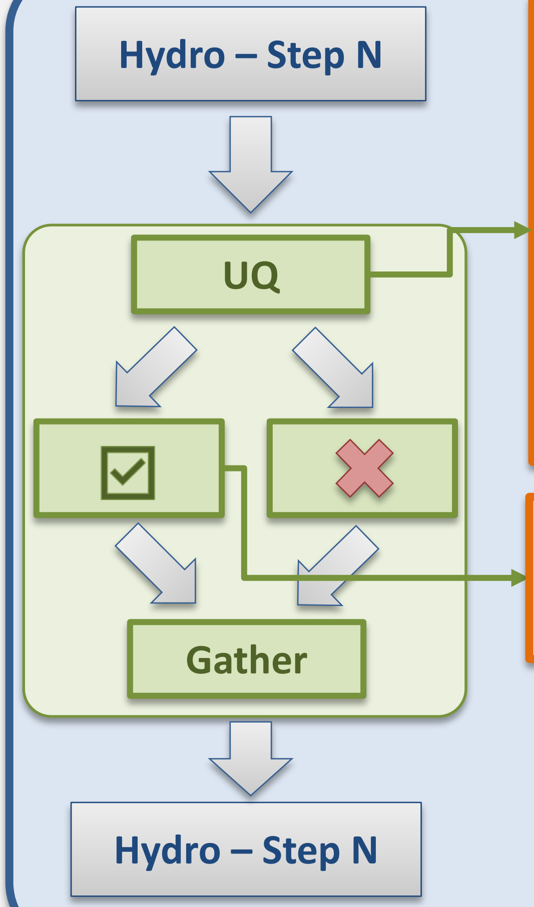


Autonomous Multiscale Simulations – Integration



- Assume standard MPI code
- Aiming for a single line replacement

Autonomous Multiscale Simulations – Turn Bulk-Parallel Subscale Physics into Dynamic Workflow



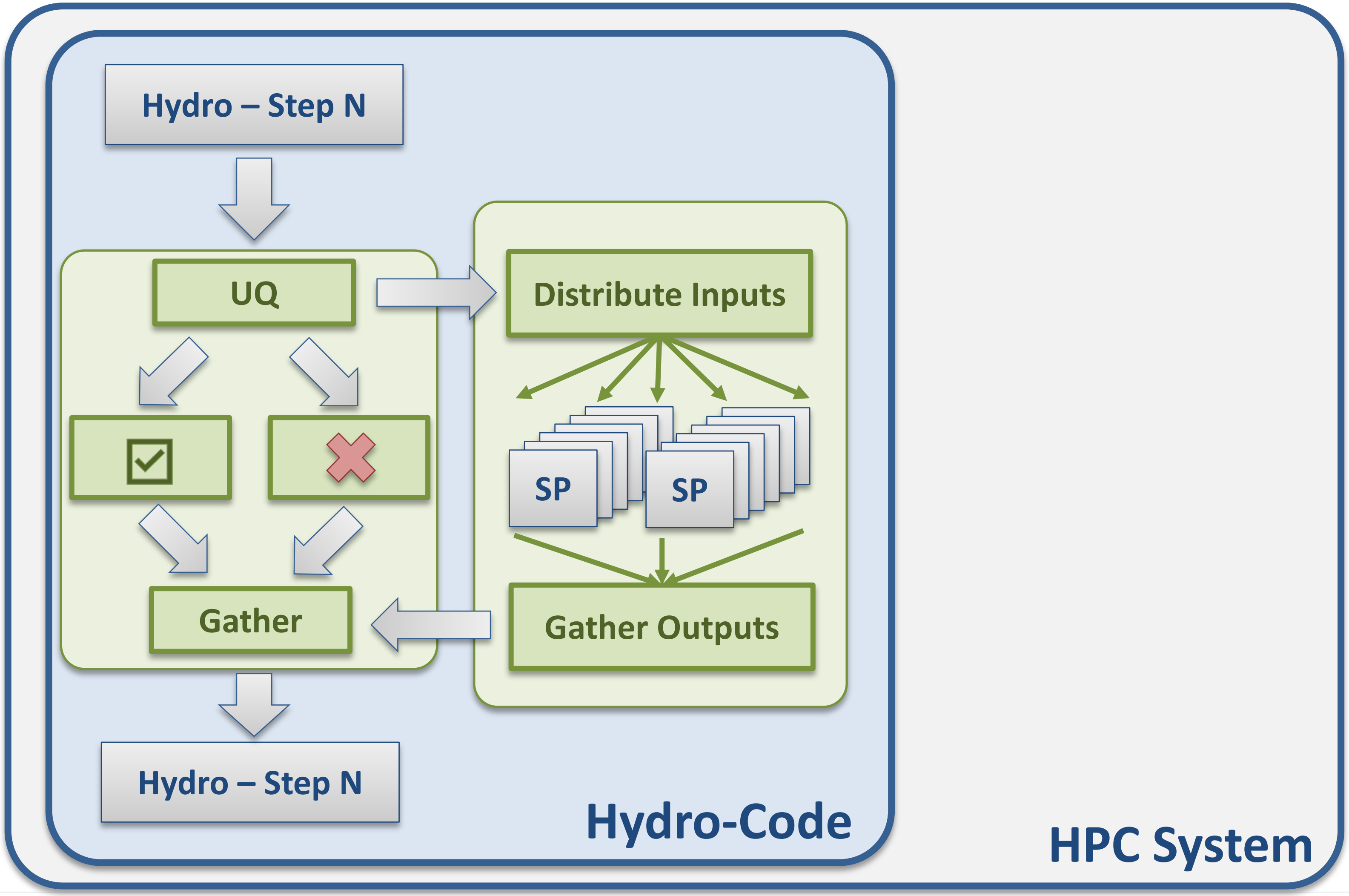
- Confidence in the current model with current input
  - **Input based** – How far away from the training data is the current inputs
  - **DeltaUQ** – Uncertainty estimates for arbitrary neural networks
    https://paperswithcode.com/paper/
    single-model-uncertainty-estimation-via

- PyTorch model inference

Hydro-Code

HPC System

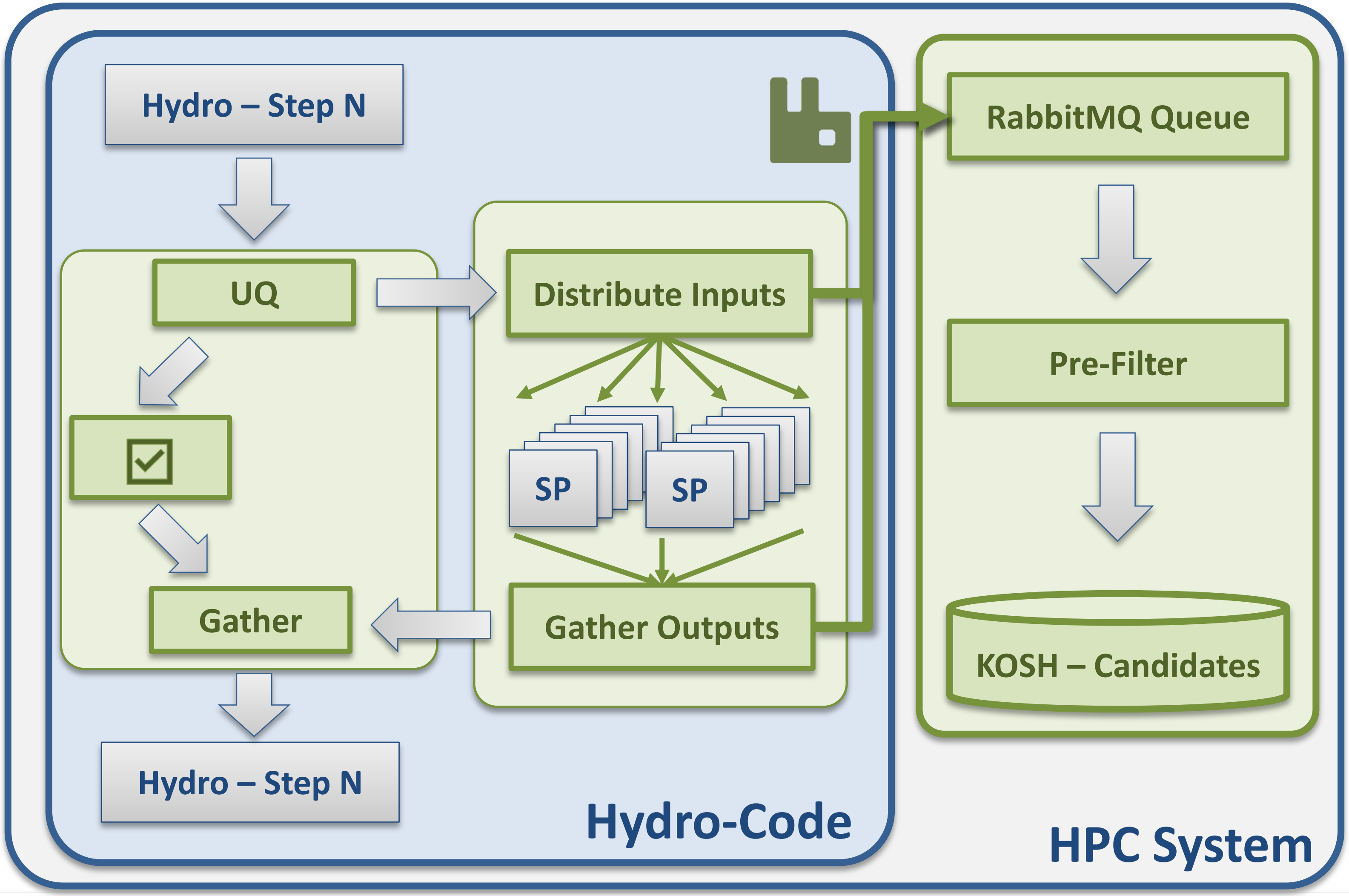
Autonomous Multiscale Simulations – Mitigate Inevitable Bottlenecks in Subscale Physics via Automatic Load Balancing



- Load balancing
  - Cross-node
  - Cross-machine

- Proxy-App emulating hydro
- AMS interface integrated with MARBL

Autonomous Multiscale Simulations – Preparing for Dynamic, "Autonomous" Models via Training Data Collection

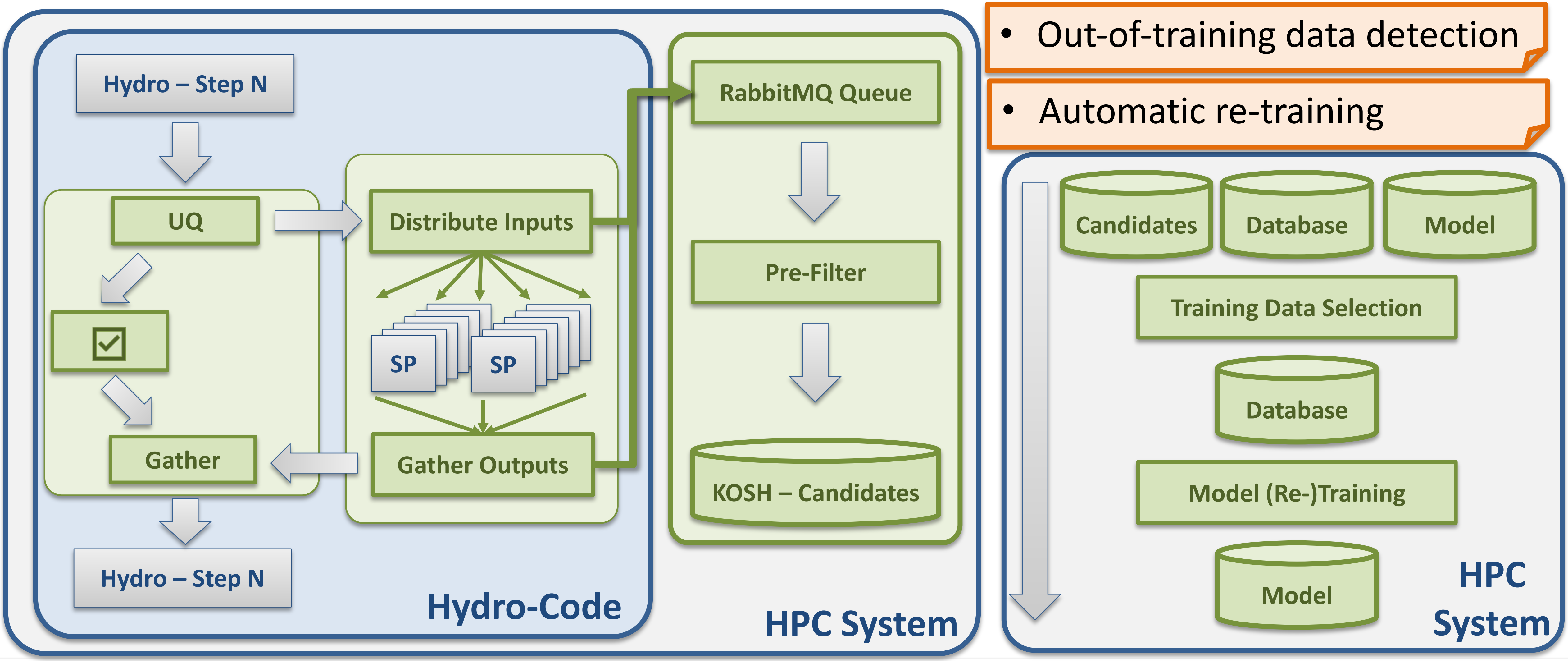


- Asynchronous queue of potential new training data

- Similarity based filter (FAISS)

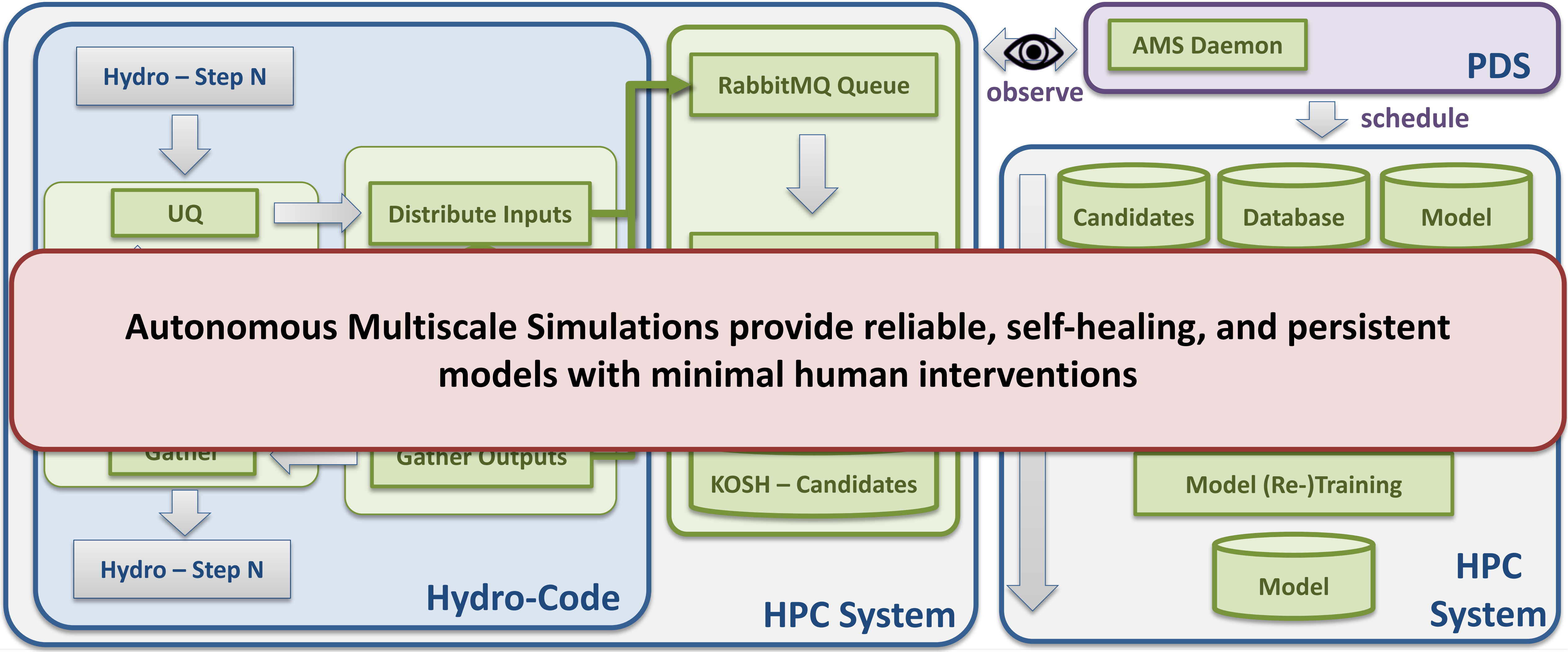
- File based KOSH database of potential candidates

Autonomous Multiscale Simulations – Asynchronous and Independent Training Data Selection and Model Update



- Out-of-training data detection
- Automatic re-training

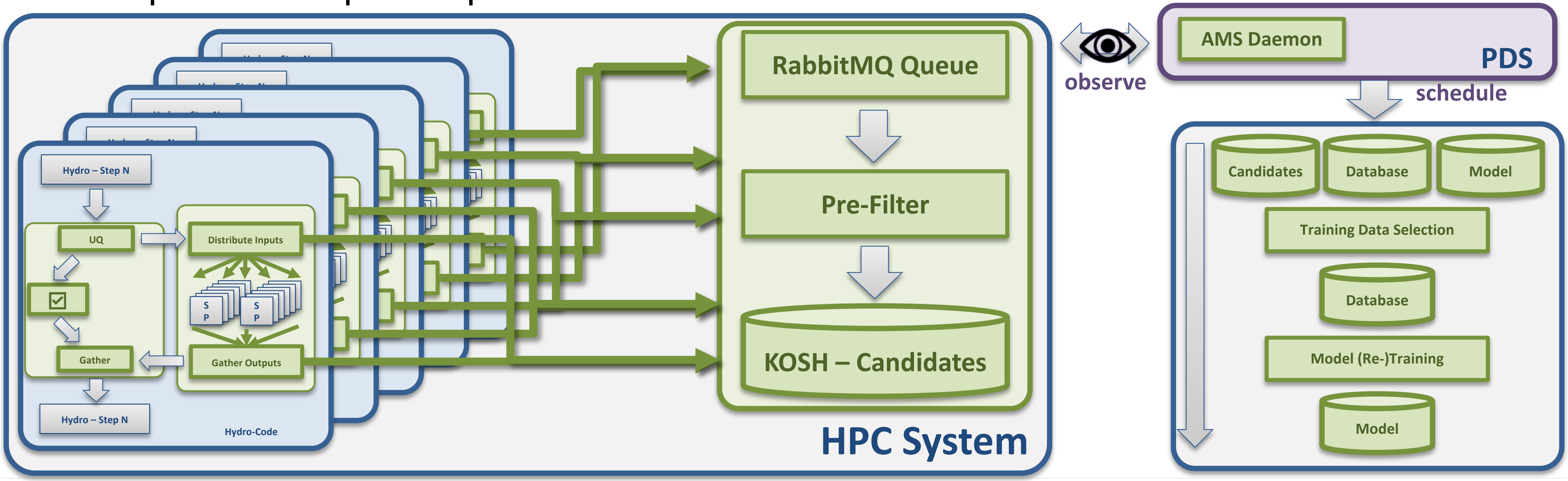
Autonomous Multiscale Simulations – Workflow Coordination using Existing Cloud Infrastructure



Autonomous Multiscale Simulations provide reliable, self-healing, and persistent models with minimal human interventions

Autonomous Multiscale Simulations are Structurally Different from the State-of-the-Art and Enable Fundamentally New Capabilities

- Interface for persistent models that will become major assets
- Implicit sharing of information and computing resources between related projects
- Super-linear speed-up of simulation ensembles



Autonomous Multiscale Simulations are Structurally Different from the State-of-the-Art and Enable Fundamentally New Capabilities

- Interface for persistent models that will become major assets
- Implicit sharing of information and computing resources between related projects
- Super-linear speed-up of simulation ensembles
- Interface to gradually increase accuracy of subscale physics as models are converging
  - Including the potential to integrate experimental data
- Simultaneously build multiple disjoint models
- Ability for nested multiscale modeling

First light computation of the single app version this Spring 2023

CASC

Center for Applied Scientific Computing

Lawrence Livermore National Laboratory