Intelligent Simulations Will Demand New Extreme-scale Computing Capabilities

Ian Foster

The University of Chicago

Argonne National Laboratory

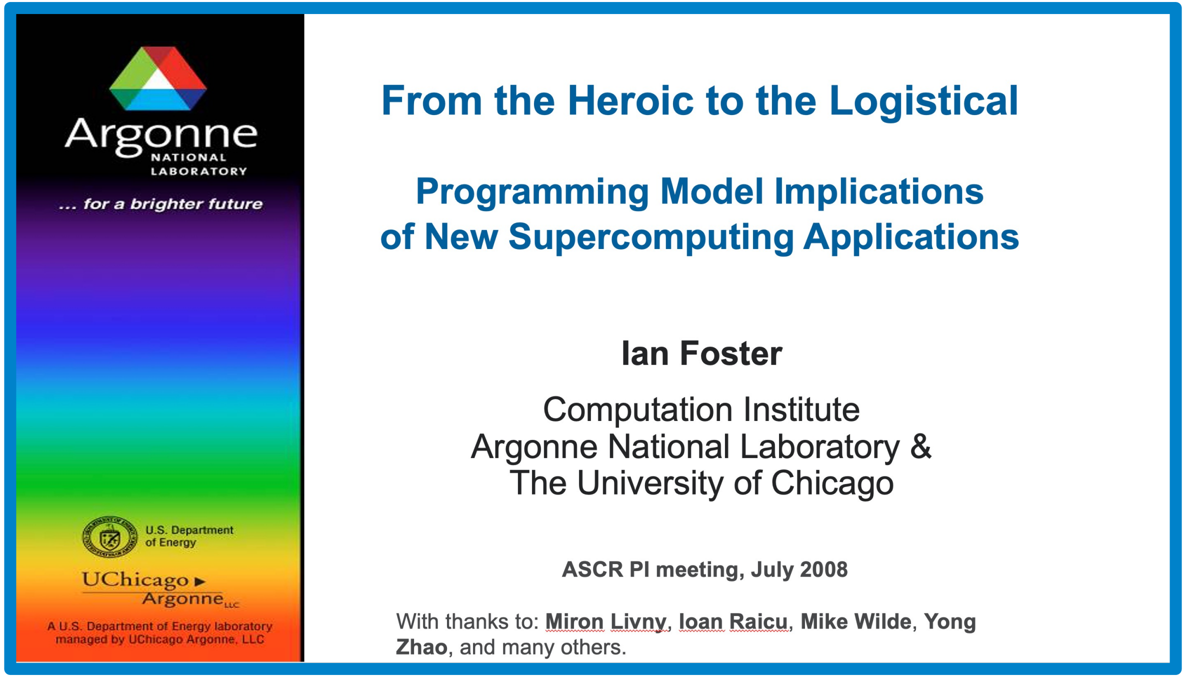
globus labs

foster@uchicago.edu, @ianfoster

Crescat scientia; vita excolatur

A talk that I gave in 2008

“What will we do with exascale computers?”



Argonne NATIONAL LABORATORY

... for a brighter future

U.S. Department of Energy

UChicago ▸ Argonne LLC

A U.S. Department of Energy laboratory managed by UChicago Argonne, LLC

From the Heroic to the Logistical

Programming Model Implications of New Supercomputing Applications

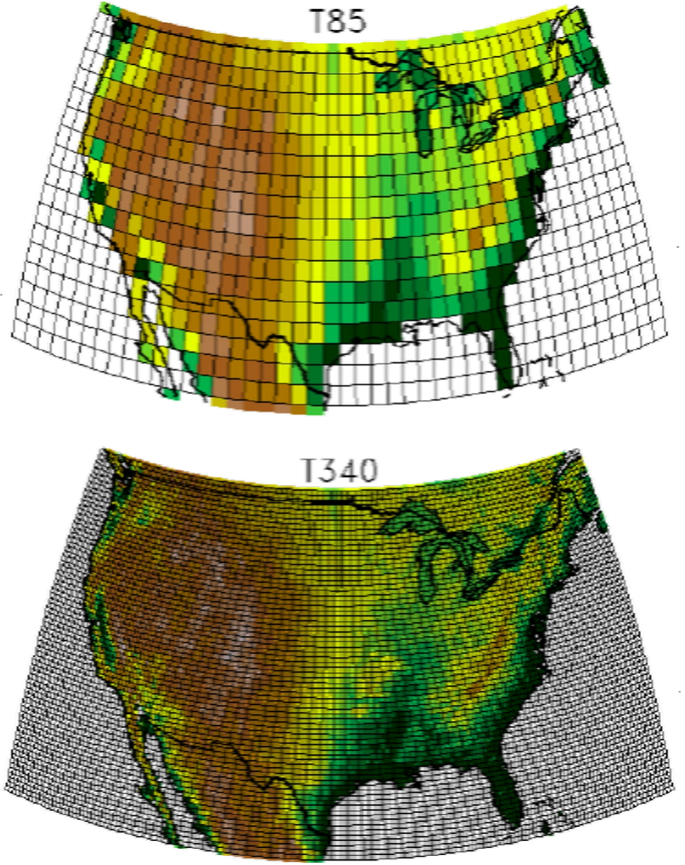
Ian Foster

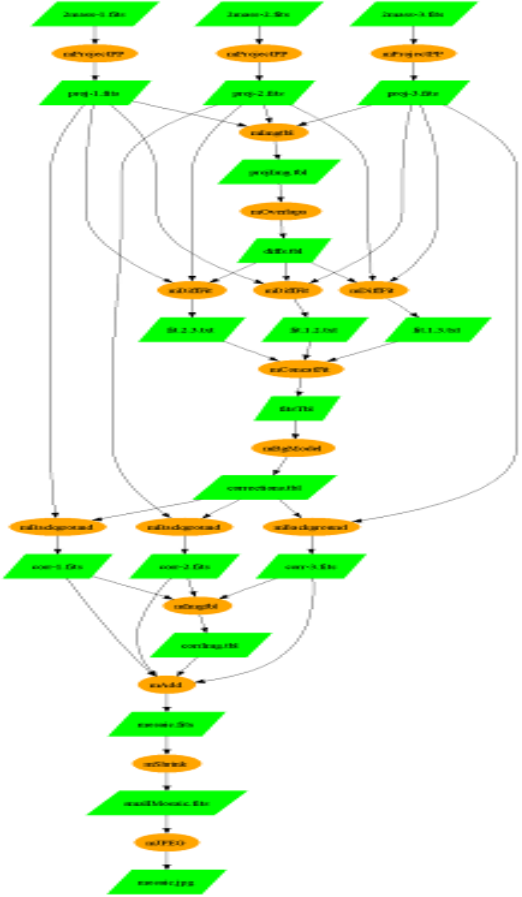
Computation Institute
Argonne National Laboratory &
The University of Chicago

ASCR PI meeting, July 2008

With thanks to: **Miron Livny**, **Ioan Raicu**, **Mike Wilde**, **Yong Zhao**, and many others.

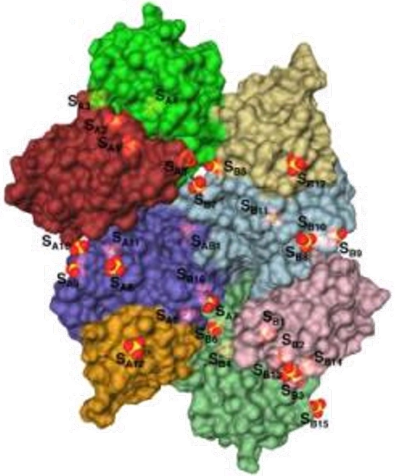
Bigger problems … or … more complex problems



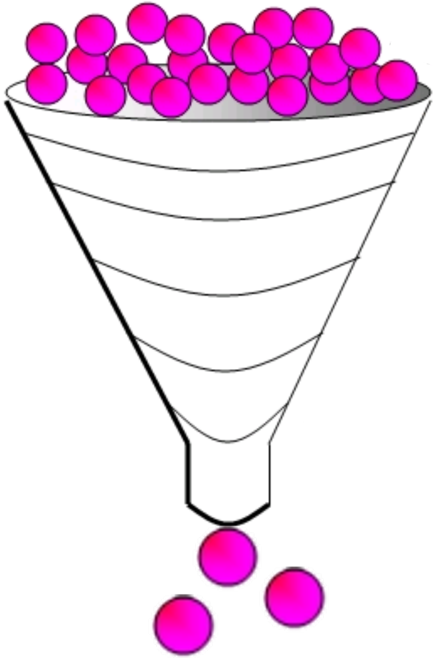


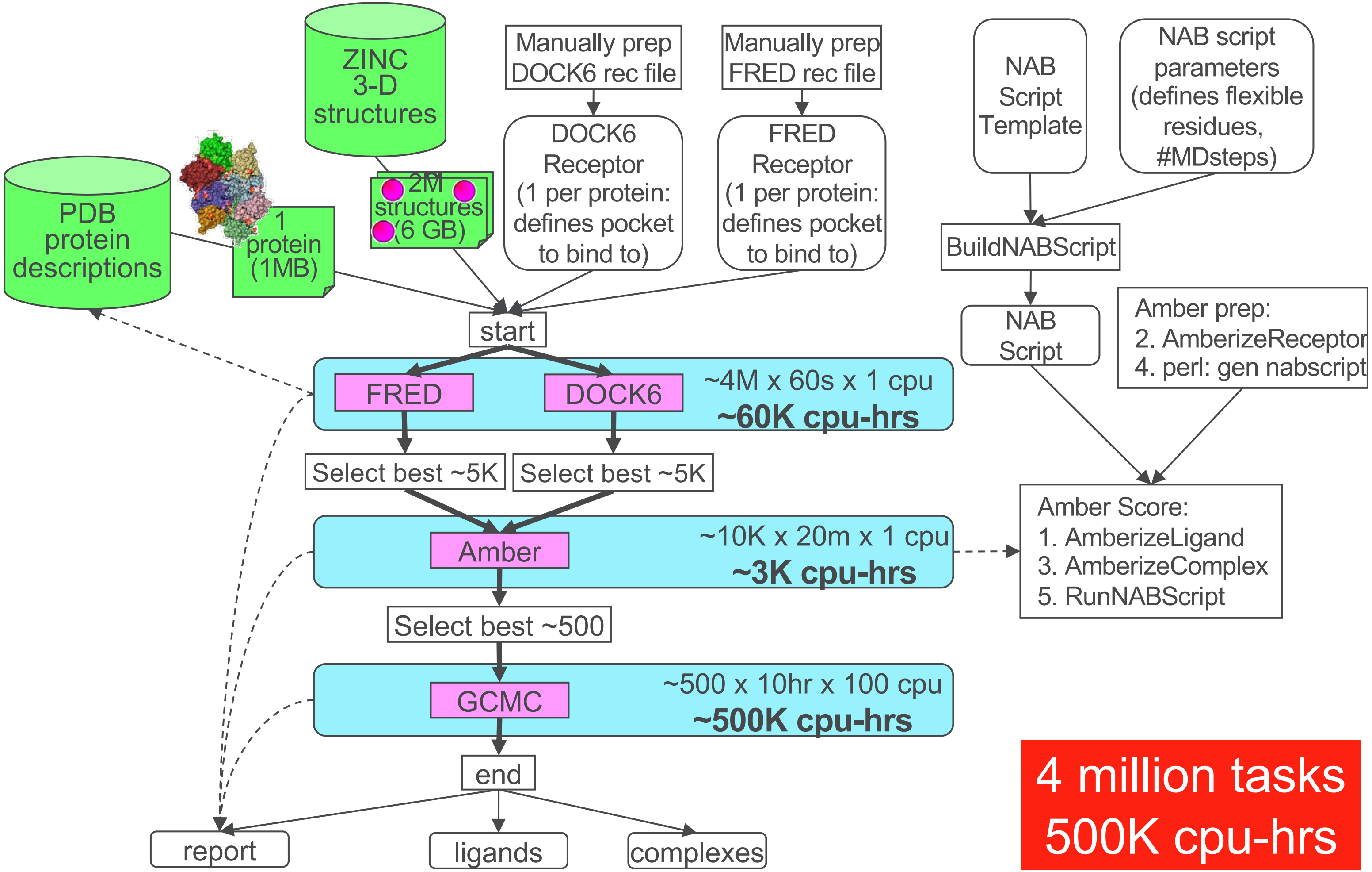
Example: Identifying potential drug targets

Protein
target(s)



2M+ ligands





ZINC 3-D structures

2M structures (6 GB)

PDB protein descriptions

1 protein (1MB)

Manually prep DOCK6 rec file

DOCK6 Receptor (1 per protein: defines pocket to bind to)

Manually prep FRED rec file

FRED Receptor (1 per protein: defines pocket to bind to)

NAB Script Template

NAB script parameters (defines flexible residues, #MDsteps)

BuildNABScript

NAB Script

Amber prep:
2. AmberizeReceptor
4. perl: gen nabscript

start

FRED

DOCK6

~4M x 60s x 1 cpu

~60K cpu-hrs

Select best ~5K

Select best ~5K

Amber

~10K x 20m x 1 cpu

~3K cpu-hrs

Amber Score:
1. AmberizeLigand
3. AmberizeComplex
5. RunNABScript

Select best ~500

GCMC

~500 x 10hr x 100 cpu

~500K cpu-hrs

end

report

ligands

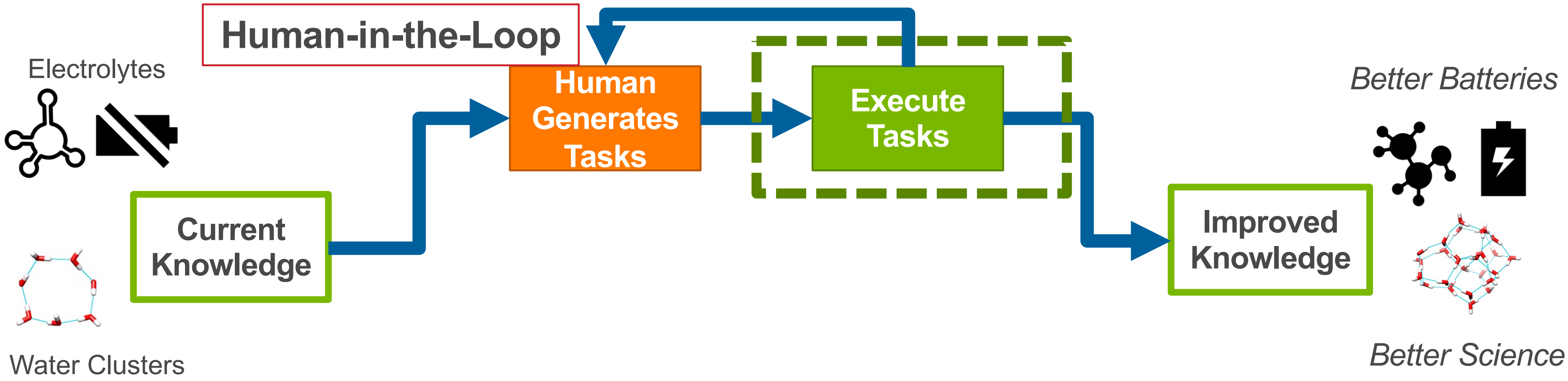
complexes

4 million tasks
500K cpu-hrs

We need to make **smarter** choices

We need to make **smarter** choices

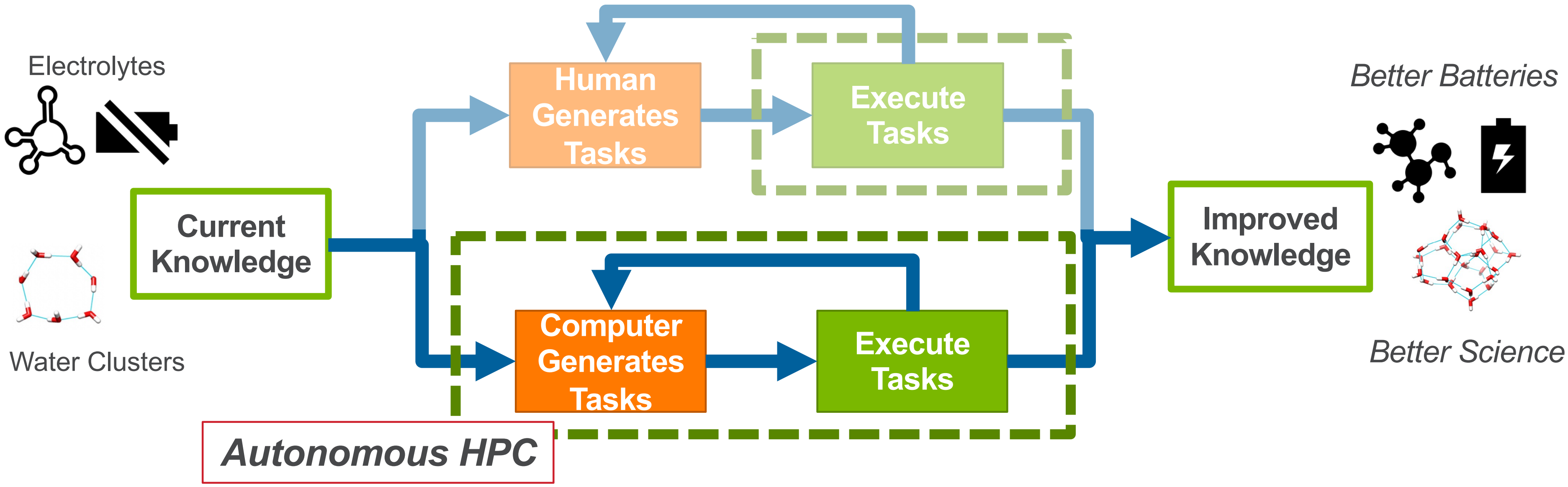
Ask a human? Humans steer HPC, HPC performs simulations



We need to make **smarter** choices

Ask a human? Humans steer HPC, HPC performs simulations

But: Humans are slow and are not getting faster



We want HPC to steer itself

Substitute AI “agents” for human as decision-maker

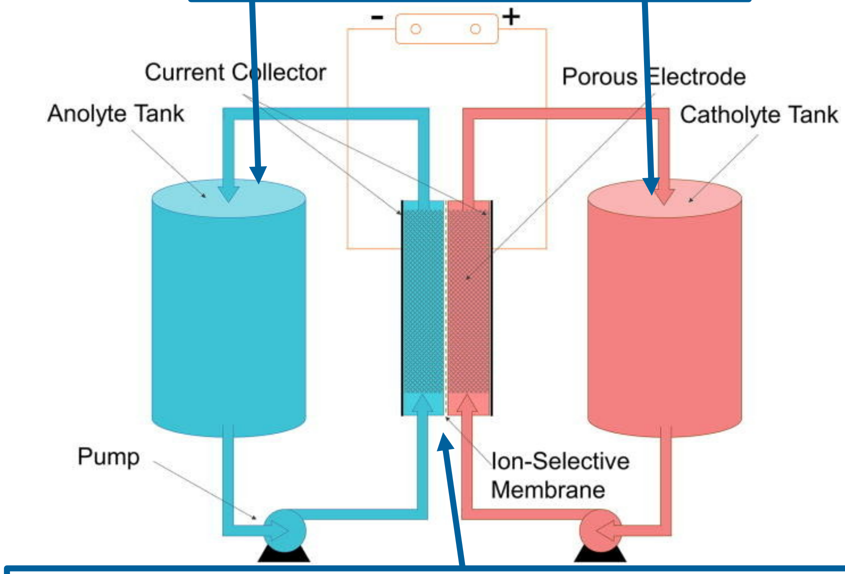
1) AI agents are **trained on all available data** prior to computational experiment
   - E.g., data from scientific literature, results of previous simulations

2) AI agents are **updated as computational experiment proceeds**
   - They gets “smarter” as more data are acquired
   - Requires periodic retraining of AI models

3) Updated model makes **smarter choices over time**
   - Active learning, Bayesian optimization, surrogate optimization, optimal experimental design

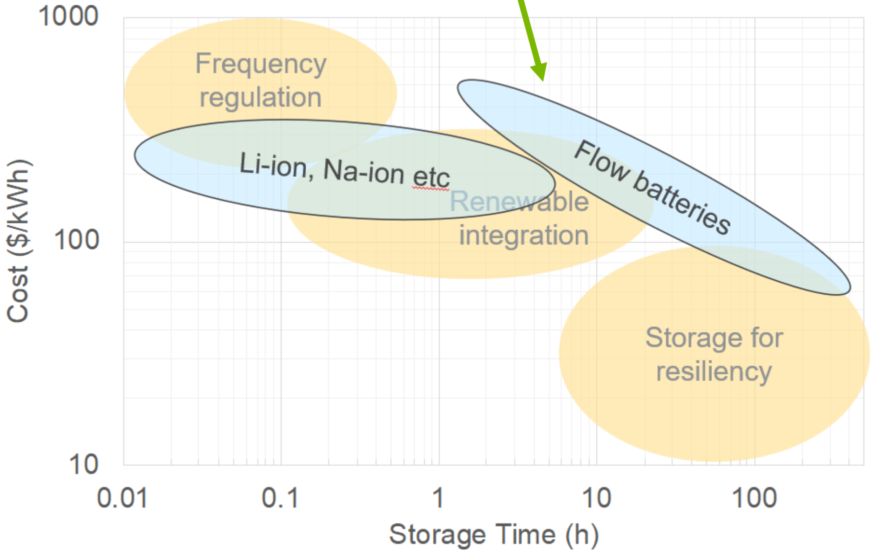
Example 1: Redox flow batteries

Energy stored in molecules that hold electric charge: “gas tank”



Store/release energy at the current collector: “engine”

Cheaply upgrade storage capacity – Just buy a bigger gas tank and more gas



Key problem: What molecules do I use to hold electric charge? (“fuel”)

Figures: (left) Wikipedia, (right) V. Srinivasan (Argonne)

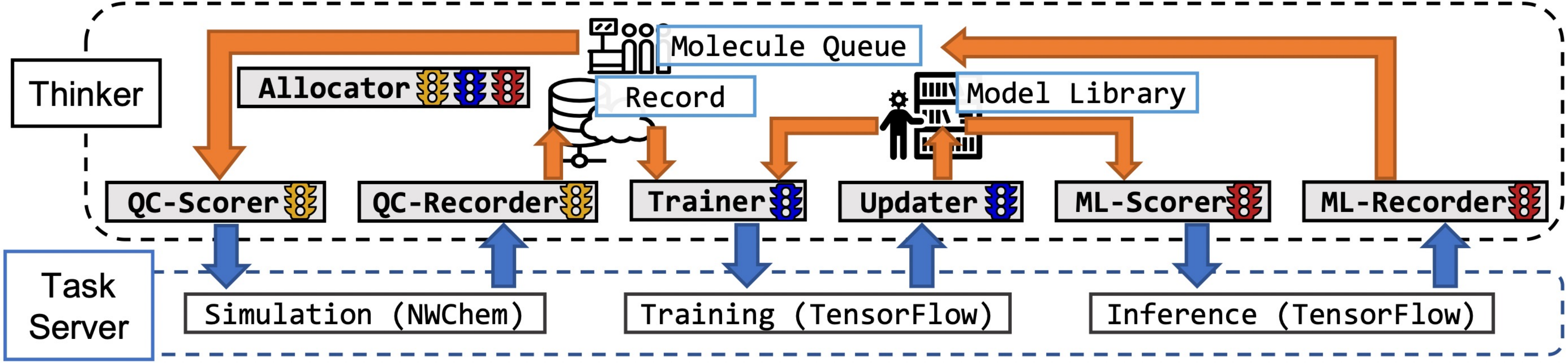
Simplified design problem

Entities: 10^5 molecules in QM9 **Resources**: 1024 KNL nodes, ALCF Theta

Possible Tasks:

1. **Simulation**: Ionization potential (NWChem, B3LYP/3-21g, 6 node-hr/mol)
2. **Inference**: Estimate ionization potential (MPNN, 3×10^{-6} node-hr/mol)
3. **Training**: Retrain MPNN with latest dataset

Objective Function: # molecules with high ionization potential (IP > 10V)



Building ML-guided applications: The **Colmena** framework

Problem: We have many policy ideas, e.g.:

- *Submit a new simulation* ***once another completes***
- *Retrain a model* ***after each 8 successful computations***
- ***Allocate more nodes to inference*** *after models finish training*

Event-triggered

Conditional logic

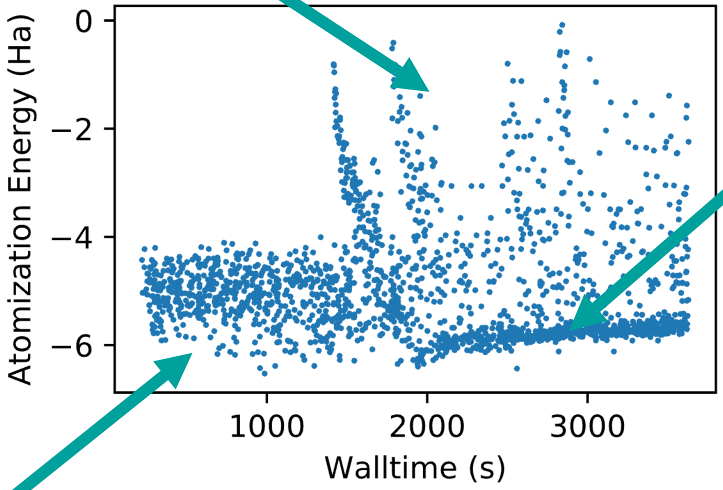
Resource management

Solution: Program **agents** to encode such policies

1. Can **react to events**
2. Can **hold state**
3. Can **re-allocate resources** between pools
4. Separate **agent** from **how to run tasks** and **interface with HPC**

L. Ward et al., MLHPC Workshop, 2021: https://arxiv.org/abs/2110.02827

Colmena system guiding exploration of electrolyte design space



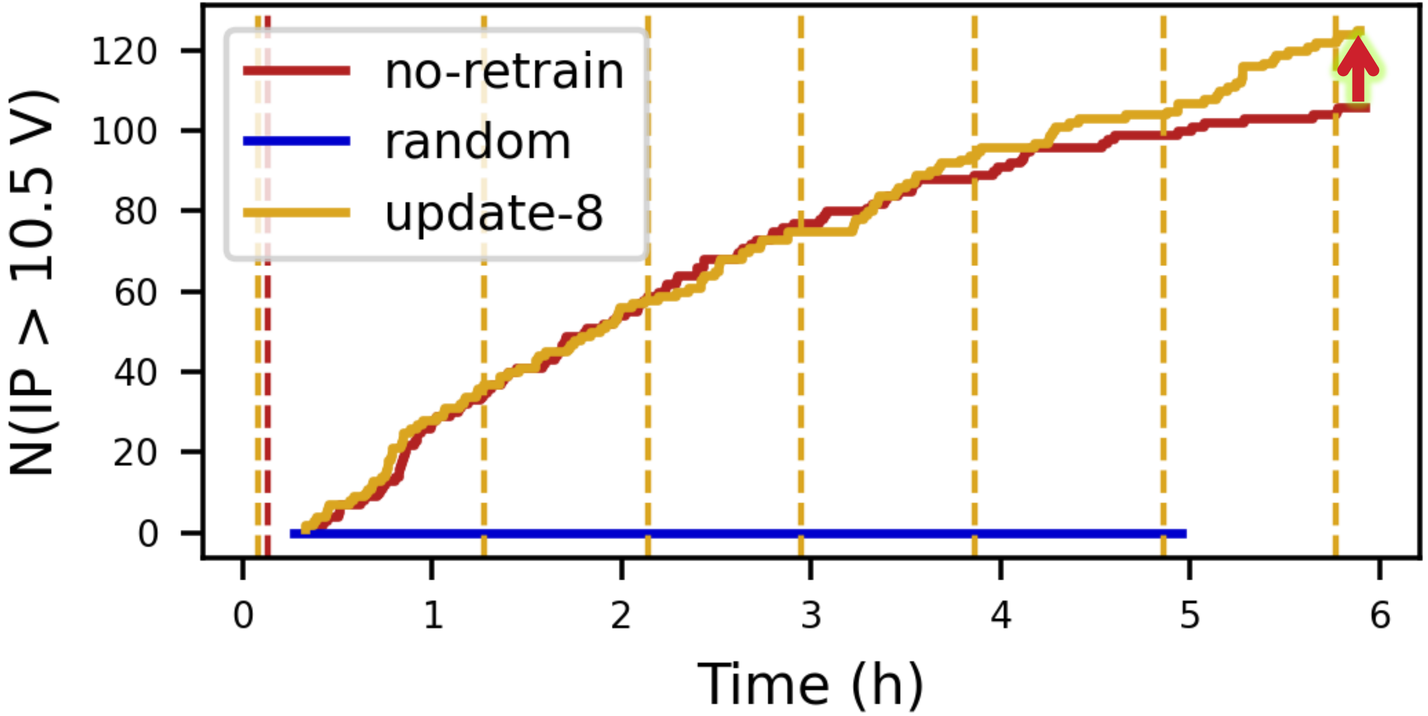
Running “exploratory” simulations

Better performance after scoring completes

Running random guesses at first

L. Ward et al., MLHPC Workshop, 2021: https://arxiv.org/abs/2110.02827

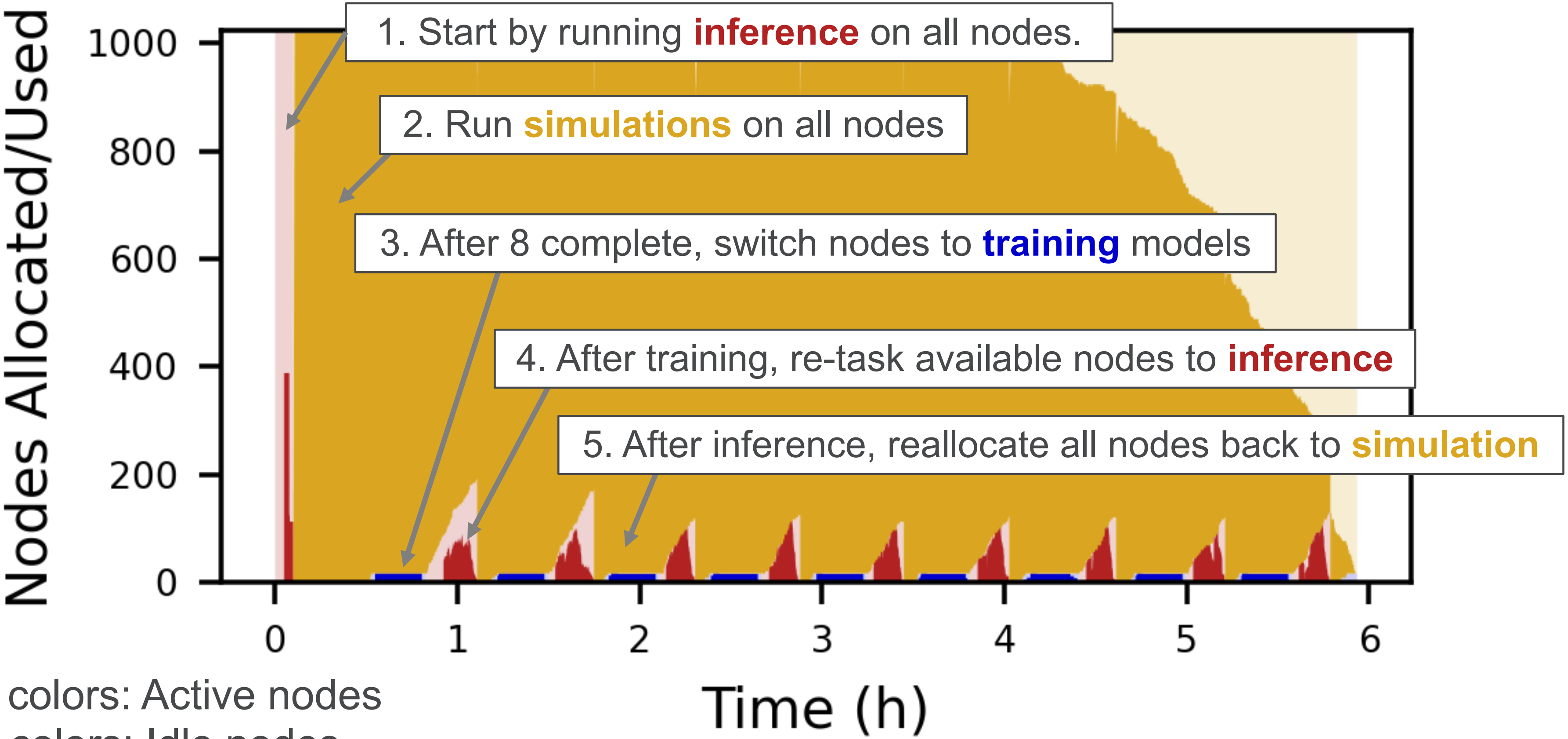
Even on this simple problem, good scientific performance



Found 10% more high-performing molecules with same allocation size

L. Ward et al., MLHPC Workshop, 2021: https://arxiv.org/abs/2110.02827

Policies guide dynamic behaviors



1. Start by running **inference** on all nodes.

2. Run **simulations** on all nodes

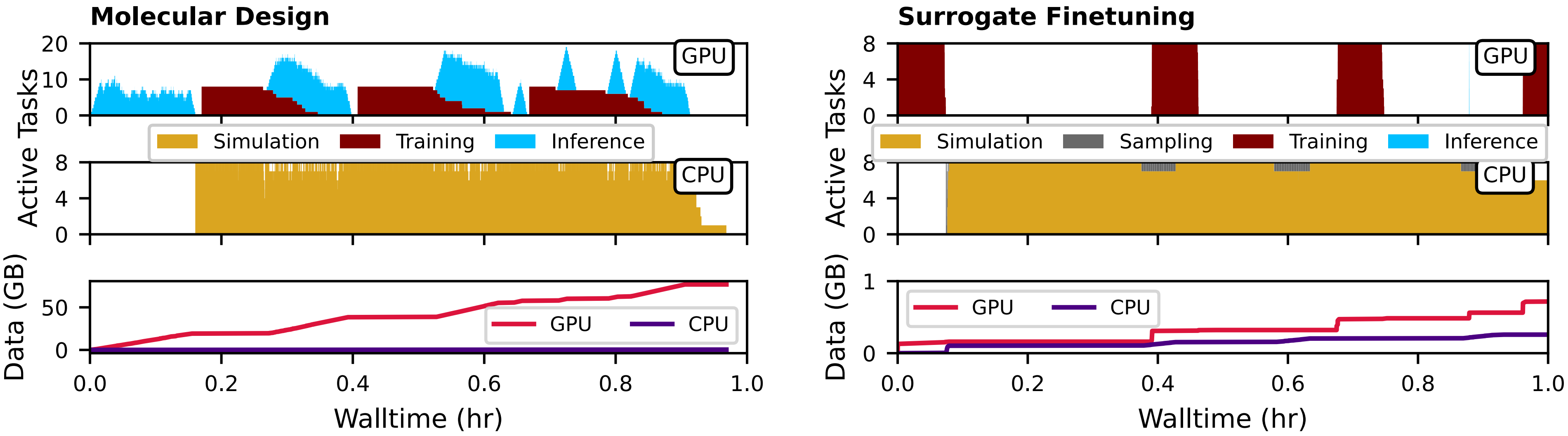
3. After 8 complete, switch nodes to **training** models

4. After training, re-task available nodes to **inference**

5. After inference, reallocate all nodes back to **simulation**

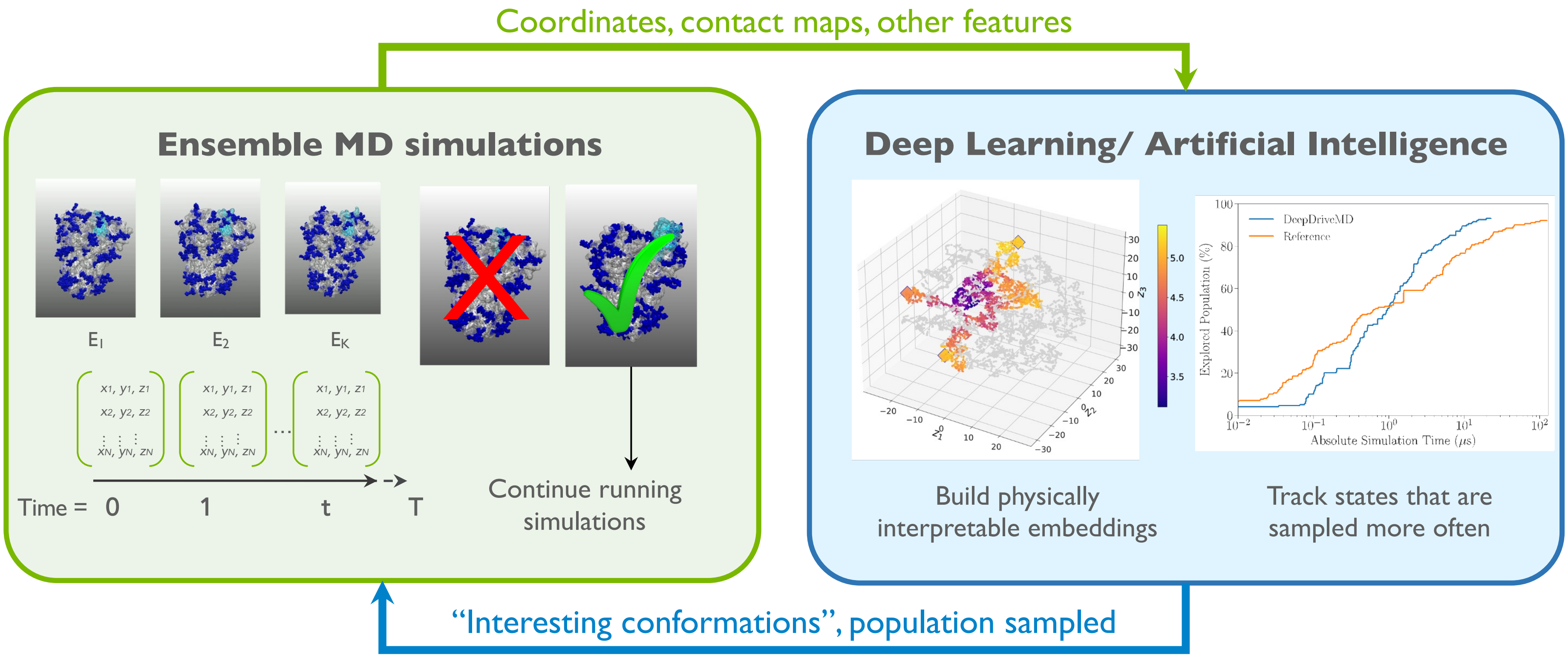
Dark colors: Active nodes
Light colors: Idle nodes

Exploiting heterogeneous & distributed computers



Logan Ward et al.

Example 2: AI-enabled molecular dynamics (MD)



Coordinates, contact maps, other features

Ensemble MD simulations

Continue running simulations

Deep Learning/ Artificial Intelligence

Build physically interpretable embeddings

Track states that are sampled more often

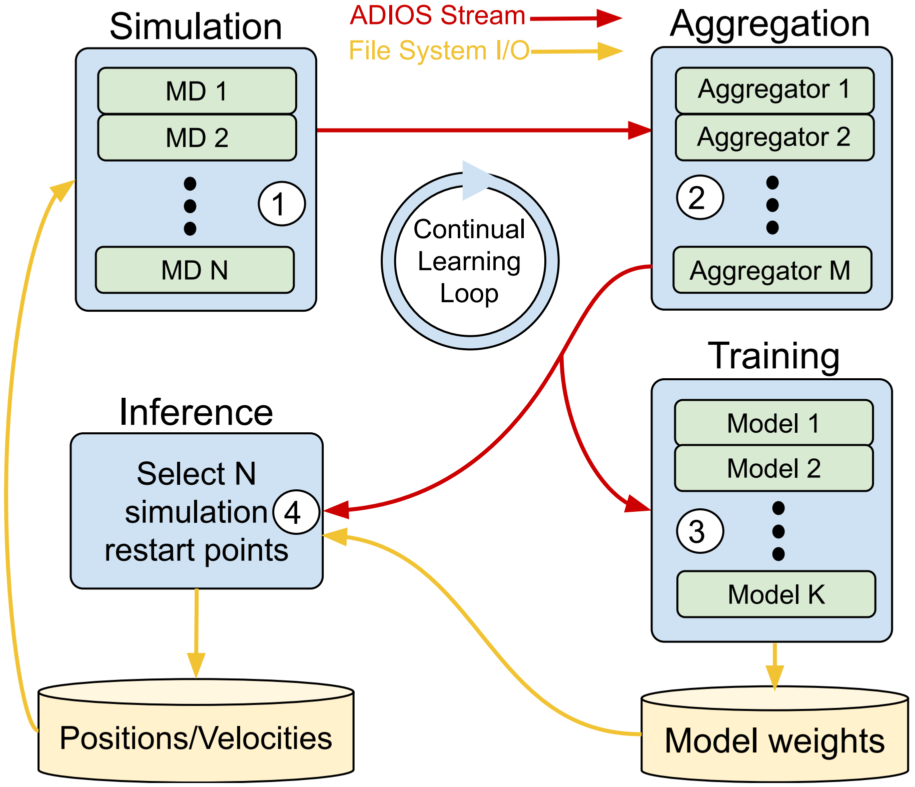
“Interesting conformations”, population sampled

Arvind Ramanthan et al.

DeepDriveMD framework for ML steering of MD simulations

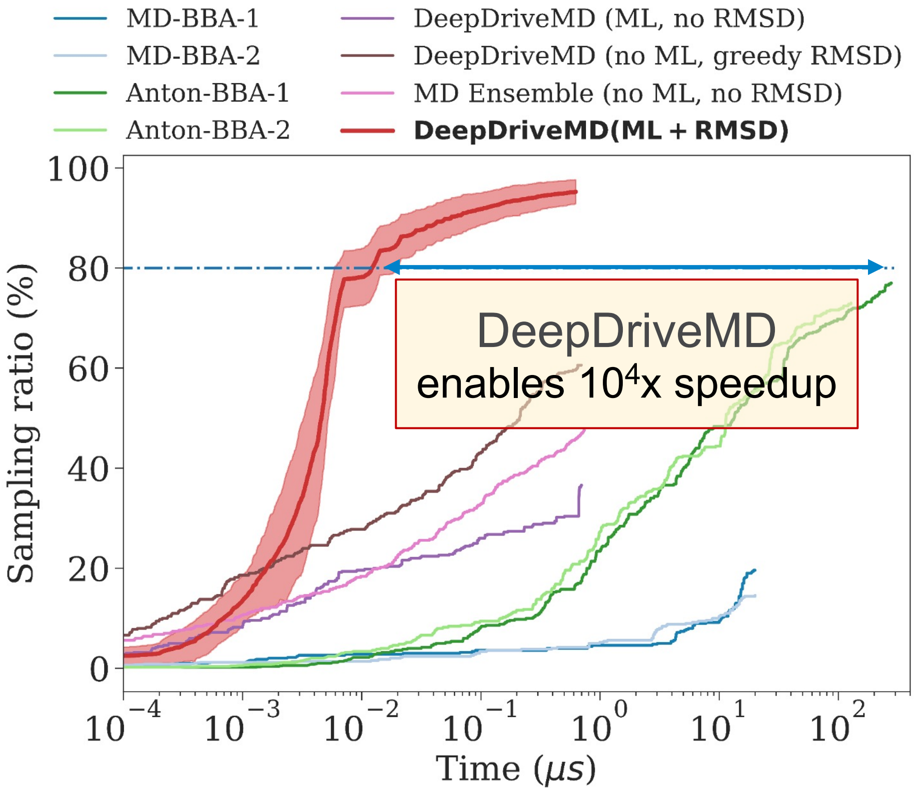
Link MD ensembles and ML training in a continual learning loop

- Blue: DeepDriveMD components
- Green: Tasks, managed by Radical Cyber Tools (Jha et al.)
- Red: ADIOS streams
- Yellow: File system.



A. Brace et al., https://arxiv.org/abs/2104.04797

DeepDriveMD enables 10^4x acceleration of sampling effectiveness for FSD-EY ($\beta\beta\alpha$) folding



DeepDriveMD enables 10^4x speedup

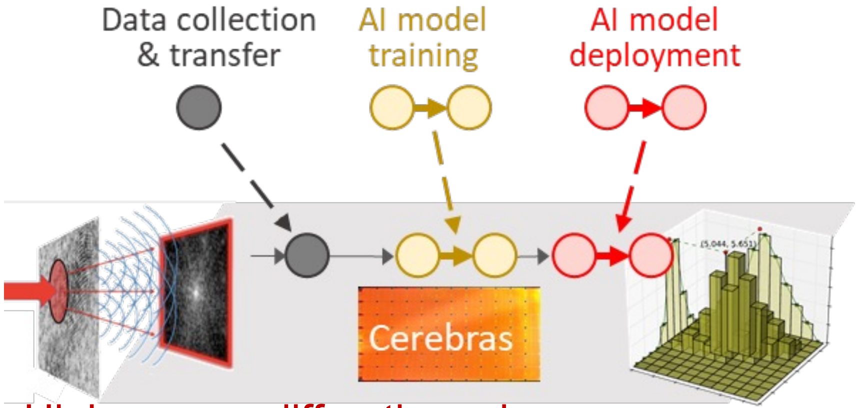
Embedding states into the VAE latent space and clustering with k-means keeps a constant definition of the number of states sampled enabling fair comparison between simulations

The ML + RMSD strategy reaches **80% sampling more than 1000x faster** (in total simulated time) than Anton-1 simulations

Note: Uncertainty from 10 trials in light red

A. Brace et al., https://arxiv.org/abs/2104.04797

Increasingly diverse data + compute “flows” … linking HPC with the computing continuum



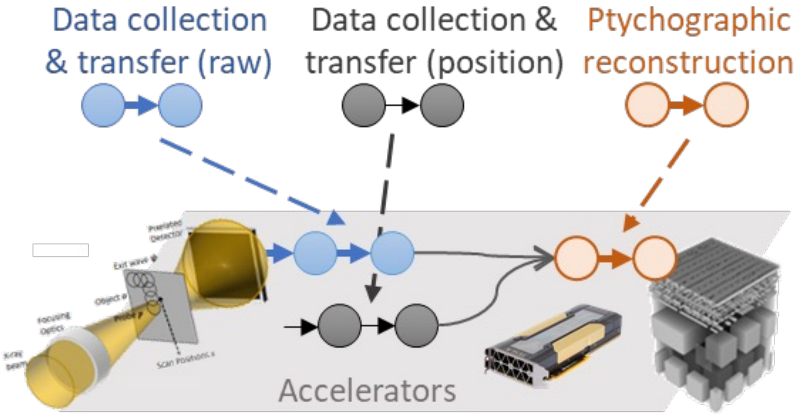
Data collection & transfer

AI model training

AI model deployment

Cerebras

High-energy diffraction microscopy



Data collection & transfer (raw)

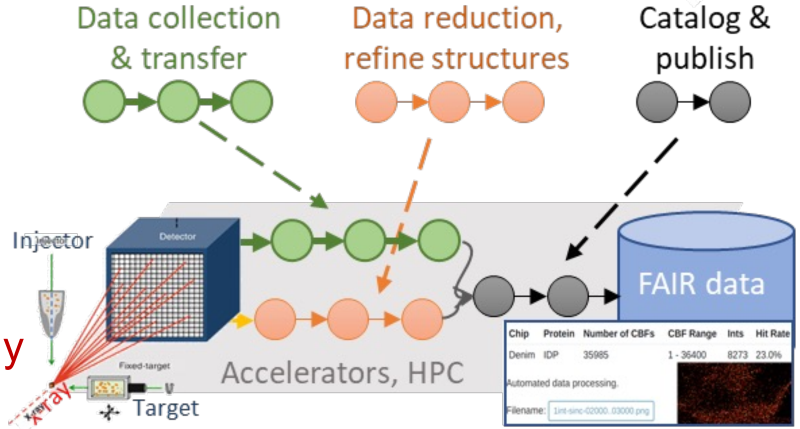
Data collection & transfer (position)

Ptychographic reconstruction

Accelerators

Ptychography

Serial synchrotron crystallography



Data collection & transfer

Data reduction, refine structures

Catalog & publish

Injector

Accelerators, HPC

FAIR data

Target

Chip	Protein	Number of CBFs	CBF Range	Ints	Hit Rate
Denim	IDP	35985	1 - 36400	8273	23.0%

Automated data processing.

Filename: 1int-sinc-02000..03000.png

“Metacomputing” revisited

10^{10} x faster

10^5 x more tasks

10^6 x more data

Link HPC, AI, instruments

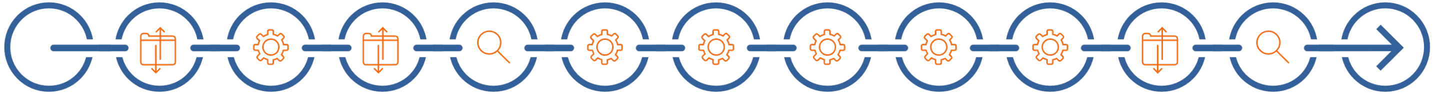
c still 3 x 10^8 m/s ☹

Globus Automation Services

Reusable **flows** composed from an extensible set of **actions**

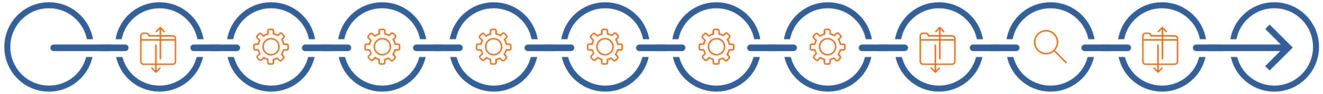
Built on **global auth, compute, data fabric**

XPCS



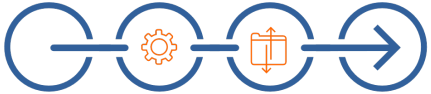
TRANSFER	COMPUTE	TRANSFER	SEARCH	COMPUTE	COMPUTE	COMPUTE	COMPUTE	COMPUTE	TRANSFER	SEARCH
Transfer to HPC	Aggregate Initial Metadata	Transfer Metadata to Storage	Load Metadata in Catalog	Prepare Nodes	XPCSBoost Analysis	Correlation Plots	Extract Results Metadata	Aggregate Metadata and Data	Transfer Metadata and Data	Publish Results

SSX-Stills



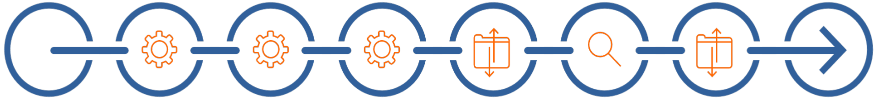
TRANSFER	COMPUTE	COMPUTE	COMPUTE	COMPUTE	COMPUTE	COMPUTE	TRANSFER	SEARCH	TRANSFER
Transfer to HPC	Wait for Files	Create DIALS Phil	DIALS Stills	Gather Execution Data	Plot Chip Image	Gather Metadata	Transfer Metadata	Publish results	Transfer Results to Instrument

SSX-Prime



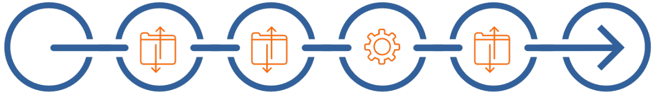
COMPUTE	TRANSFER
Refine Structure	Transfer Results to Instrument

SSX-Publish



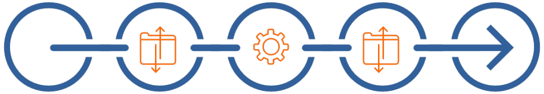
COMPUTE	COMPUTE	COMPUTE	TRANSFER	SEARCH	TRANSFER
Gather Data	Tar	DIALS Plot Histogram	Transfer Metadata	Ingest Metadata	Transfer Results

Ptycho



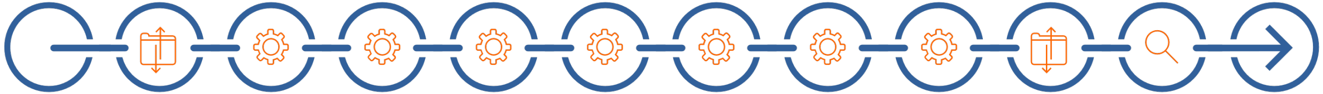
TRANSFER	TRANSFER	COMPUTE	TRANSFER
Transfer Raw Data to HPC	Transfer Position Data to HPC	Ptychographic R econstruction	Transfer Results to Instrument

PeakNN



TRANSFER	COMPUTE	TRANSFER
Transfer Data to AI Testbed	ML Model Training	Transfer Trained Model

HEDM



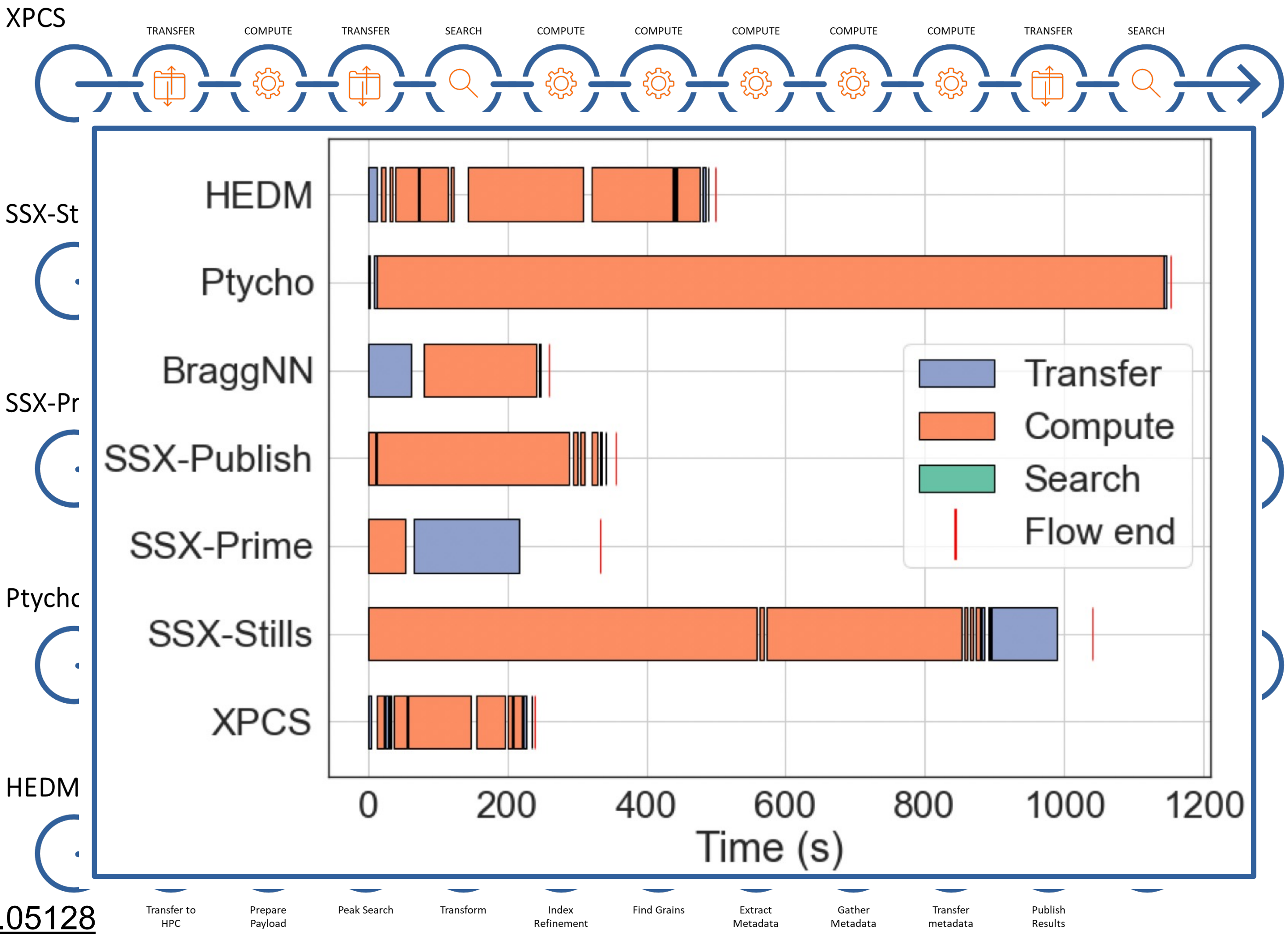
TRANSFER	COMPUTE	COMPUTE	COMPUTE	COMPUTE	COMPUTE	COMPUTE	COMPUTE	TRANSFER	SEARCH
Transfer to HPC	Prepare Payload	Peak Search	Transform	Index Refinement	Find Grains	Extract Metadata	Gather Metadata	Transfer metadata	Publish Results

https://arxiv.org/abs/2204.05128

Globus Automation Services

Reusable **flows** composed from an extensible set of **actions**

Built on **global auth, compute, data fabric**

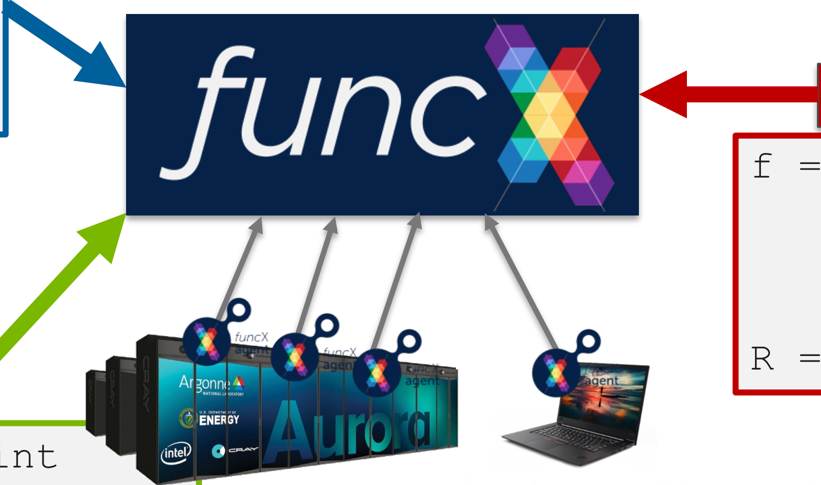


https://arxiv.org/abs/2204.05128

funcX: A managed research acceleration service that implements a **universal computing fabric**

```
def F(in_args):
 # do something
 return results

fxc.register_function(F)
```

Register functions



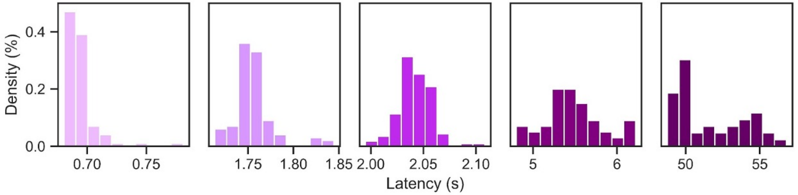
Run functions

```
F(ep, "A")
```

```
f = fxc.run("A",
      endpoint_id=ep,
      function_id=F)

R = fxc.get_result(f)
```

Deploy funcX agents

```
$ pip install funcx-endpoint

$ funcx-endpoint configure myep

$ funcx-endpoint start myep
```



(a) tabular file extraction (b) MNIST digit prediction (c) DIALS stills process (d) tomographic preview (e) correlation spectroscopy

Z. Li et al., https://doi.org/10.48550/arXiv.2209.11631 https://funcx.org

AI + HPC: Implications and opportunities

- Many important problems cannot be addressed via simple scaling of resolution, realism, timescale, number of ensemble members

  → Need **data-informed “intelligence”** to guide exploration of large search spaces and/or produce custom approximations for expensive computations

- New challenges for AI:
  - Representing complex search spaces
  - Rapid integration of data of varying degrees of accuracy

- Important implications for HPC hardware and software systems:
  - Dynamic creation and management of **many tasks**
  - **Heterogeneous workloads**: simulation, training, inference
  - Many data-intensive, latency-sensitive computations
  - New services needed to **link HPC with computing continuum**

- Implications for discovery processes:
  - Documenting and validating results; the role of human judgement

Thanks to wonderful colleagues

- **Colmena**: Logan Ward, Ganesh Sivaraman, Greg Pauloski, Yadu Babuji, Ryan Chard, Naveen Dandu, Paul Redfern, Rajeev Assary, Kyle Chard, Larry Curtiss, Rajeev Thakur
- **DeepDriveMD**: Alexander Brace, Hyungro Lee, Heng Ma, Anda Trifan, Matteo Turilli, Igor Yakushin, Todd Munson, Shantenu Jha, Arvind Ramanathan
- **Globus**: Rachana Ananthakrishnan, Kyle Chard, Lee Liming, and many others
- **Globus automation services**: Rafael Vescovi, Ryan Chard, Nickolaus Saint, Ben Blaiszik, Jim Pruyne, Tekin Bicer, Alex Lavens, Zhengchun Liu, Michael Papka, Suresh Narayanan, Nicholas Schwarz, Kyle Chard
- **funcX**: Kyle Chard, Dan Katz, Ryan Chard, Yadu Babuji, Anna Woodard, Zhuozhao Li, Tyler Skluzacek, Ben Blaiszik, etc.

And funders

DOE

NSF

NIST

NIH

Globus subscribers