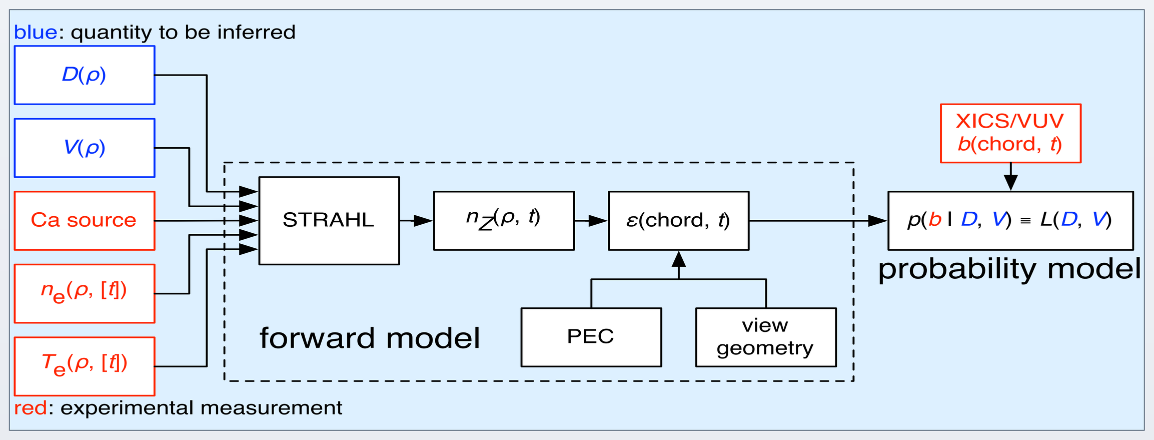
Data Management and UQ in Extreme Scale XGC Simulation

R. Moser[1], M.Greenwald[2], M. Parashar[4], S. Klasky[3],
V. Carey[1], M.Chilenski[2], S. Ku[5], J.Choi[3] and the rest of EPSI Team
U. Texas[1], MIT[2], ORNL[3], Rutgers[4], PPPL[5]

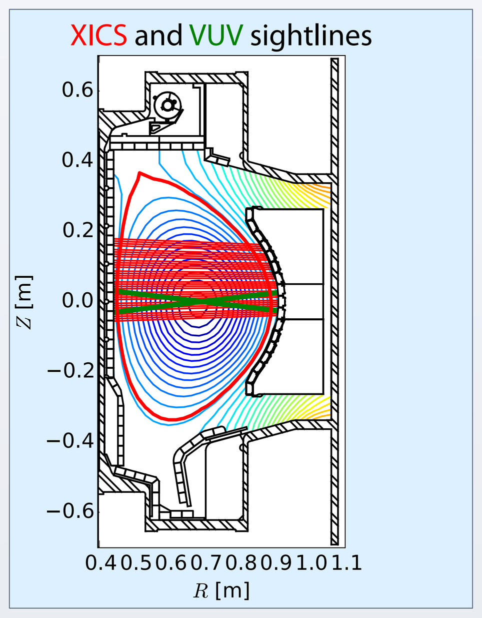
Uncertainty Quantification

Towards a Bayesian analysis of impurity transport data (MIT) for XGC validation

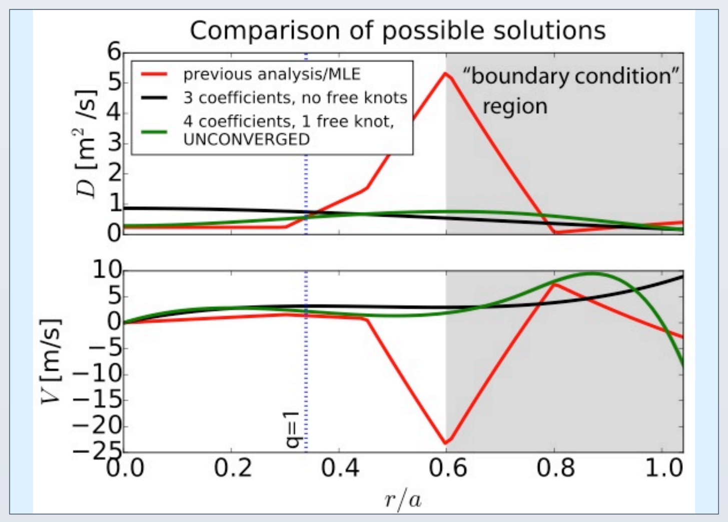
- Objective is to find *D*, *V* profiles that best reproduce the observed brightnesses *b* on each of the diagnostics.
- Nonlinear inverse problem: Key issues are existence, uniqueness and stability of the solution.



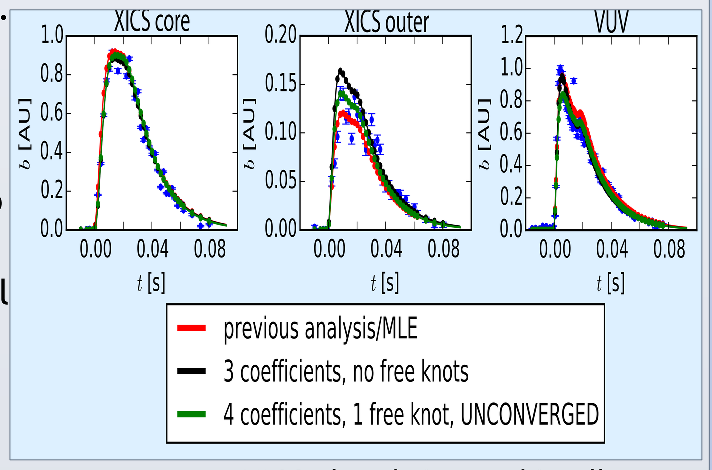
STRAHL takes electron profiles and initial guesses on impurity transport coefficients D and V and iterates to match exp. data. D and V are highly sensitive to profile uncertainties.



X-ray Imaging Crystal and Vacuum UV spectrometer sightlines in Alcator C-Mod



Previous analysis:
- Piecewise linear basis functions.
- Maximum Likelihood Estimate (MLE)without estimate of width of the posterior distribution.
- Behavior in r/a > 0.6 thought to be only weakly constrained.
- But, uncertainty there too small to be consistent with this.



Bayesian Analysis:
- Cubic spline basis functions.
- Adaptive Parallel Tempering(APT) to handle multiple maxima, width of posterior distribution.
- Uncertainty estimate in r/a > 0.6 still too small to be consistent with assumed lack of knowledge there.
- Cases shown are likely overconstrained.

- Agreement on core XICS chords is good in all cases; outer XICS chords shows widest variation
- Agreement on VUV spectrometer is reasonable.
- This shows the importance of accounting for the possibility of multiple solutions.
- Next steps: include n_e, T_e profile uncertainties.

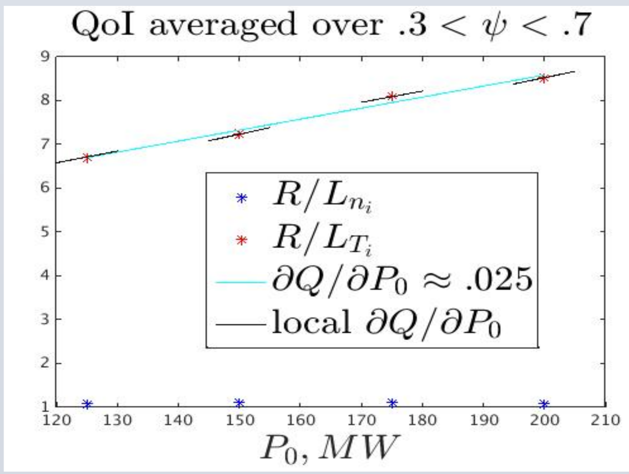
Towards UQ for Extreme Scale XGC Simulations(UT-Austin, PPPL)

Challenges: Global nonlinear system with scale-inseparable and nonlinearly self-organizing multiphysics.

Sensitivity of Plasma Gradients to Applied Heating

Expense of global gyrokinetic full-f simulation: build surrogate response

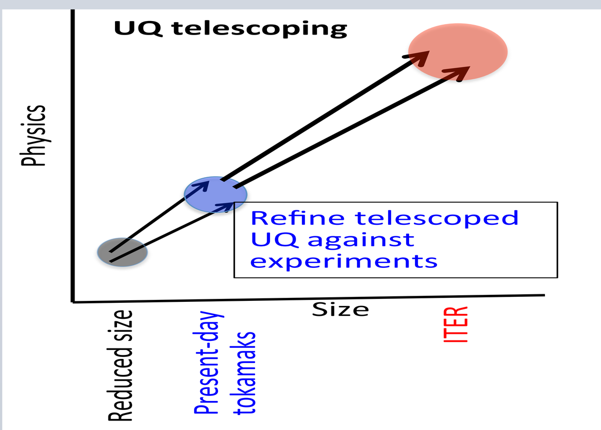
- Enrich surrogate(aka response surface) with local sensitivity information
- Perturbational local approach for flexibility, efficiency
- Figure on Right: R/L_t sensitivity to heating power P_0 computed via a global surrogate (curve fit) or local perturbation.
- Local perturbation faster; computable from restart file of "steady-state" high-fidelity XGC1 run(flexibility).
- Currently examining perturb. approach in wedge mode XGC1, XGC1-XGCa coupling.



"Steady-state" gradient sensitivities with respect to heating power computed on a high temperature version of the CYCLONE base case. Simulations performed on NERSC Edison and Hopper.

The UQ "telescoping" approach in EPSI

1. UQ in reduced size tokamaks-more details (e.g. gradients) and samples to construct resp. surface ✔
2. Scaling-up the UQ results to larger sizes ✔(*telescoping*): limited number of studies
   - Calibration/enrichment of the telescoped response surface.
   - Validation of predictions against present-day experiments.(including UQ on experimental data). Extension/Numerical version of experimental scaling laws for predicting ITER performance.
   - Use of experimental data to inform surrogate model.
3. Telescope further to ITER-scale, compare prediction against ITER-scale XGC1 simulations



Telescoping and calibration considerations:

- Simulation and data-driven separation of scale-independent and -dependent quantities.
- Physics guidance important-response surface should inform, not dictate, higher-fidelity studies.
- Negative telescoping results also useful-identify key regimes in parameter space for high-fidelity simulation(e.g. bifurcations in parameter space), compatible with Expected Information Gain(EIG) base approaches.

Initial "Telescoping" Studies Details

- Builds on mod. CYCLONE heating study-provide Q, dQ/dP at various heating power
- Telescoping parameter is $1/\rho^*$ (or a/ρ_i). Scaling is achieved by increasing magnetic field
- Compute Q,dQ/dP at range of ρ^* values, as well as multiple values of P_0.
- Compute krieging surface in the scaling parameter and the uncertain model parameter(s).
- Predict Q,dQ/dP at a higher value of a/ρ_i. How valid is "reduced-size" data?
- As number of model parameters increase, leverage QUEST software: *QUESO, GPMSA*

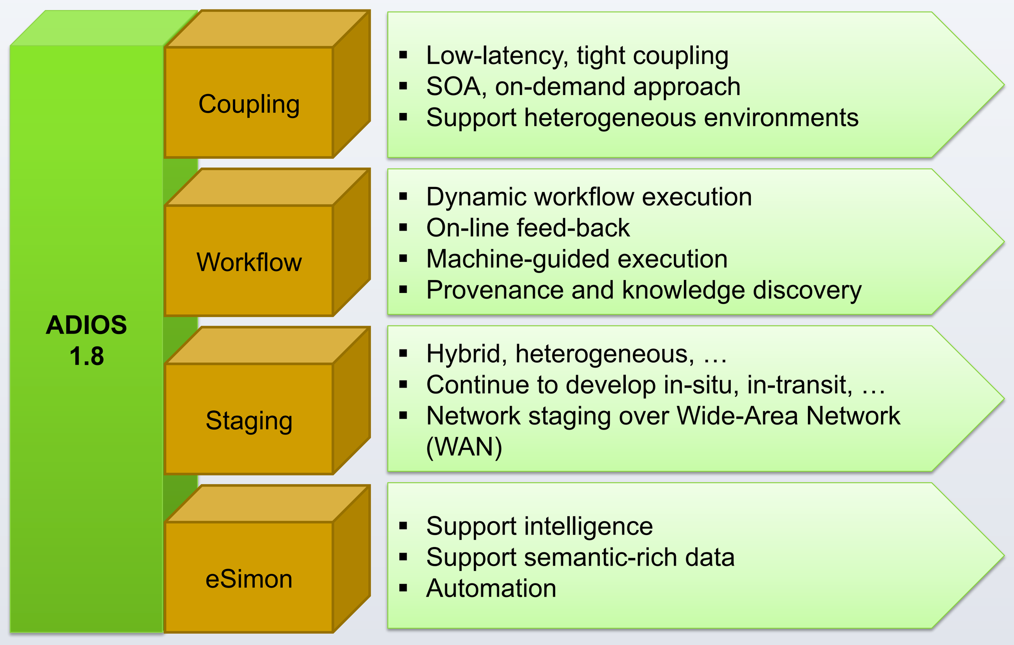
Future "Telescoping" Work

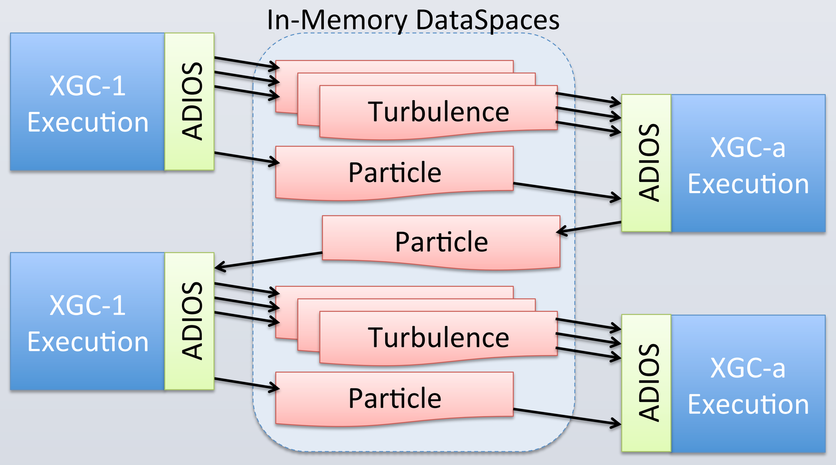
- Approach validated in for $.5 < \rho^* < 1.5$ CYCLONE case values.
- Adding additional QoI(thermal diffusivity)
- Add additional parameters present in heating model (torque, source profiles)
- Incorporate uncertainty in magnetic data into prediction response surface.
- Derive approx. uncertainty estimates for future use in response surfaces.
- Move response surface simulations and scaling closer to available exp. data.

Data Management

ADIOS in EPSI

- Data staging with low-latency, tight coupling execution environments through in-memory data exchanges between different codes.
- Service Oriented Architecture (SOA) for on-demand coupling executions with support of dynamic workflow invocation.
- Selection and chunked reads to enable schedule optimization.
- Hierarchical data management with staging for multiple I/O resources and network staging is under development.

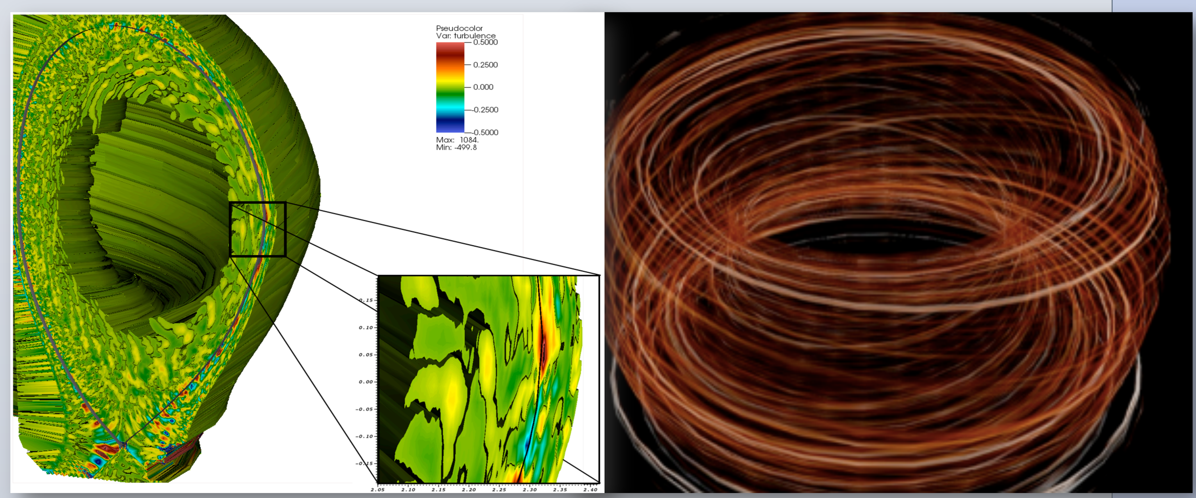




Data staging for XGC1-XGCa workflow: In-memory coupling with DataSpaces

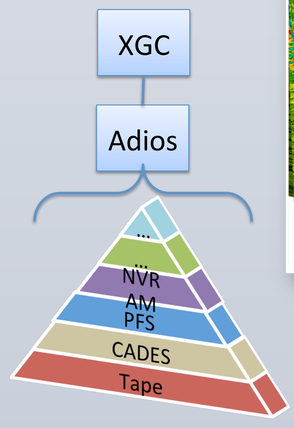
Data-centric integrated execution environment

Our focus is to support EPSI simulation by providing integrated data-centric execution environments for tight code coupling, staged data process, and monitoring system with a support of dynamic workflow system.



ADIOS Vis Schema

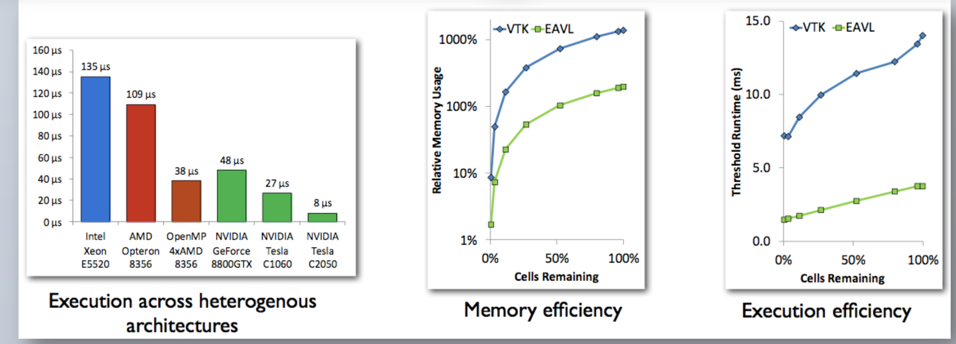
- Create an easy-to-use schema for ALL ADIOS codes.
- Facilitates data sharing without adding code complexity
- Visualization schema: Semantics of the data for the purpose of visualization.
- Describing visualization data for various tools (VTK, Matlab, ParaView).



Hierarchical data management

XGC Data Management

- Manage hierarchical data and resources
- Integrated staging services
- Support for ADIOS Vis Schema.



EAVL: Extreme Scale Analysis and Visualization Library

Data Staging

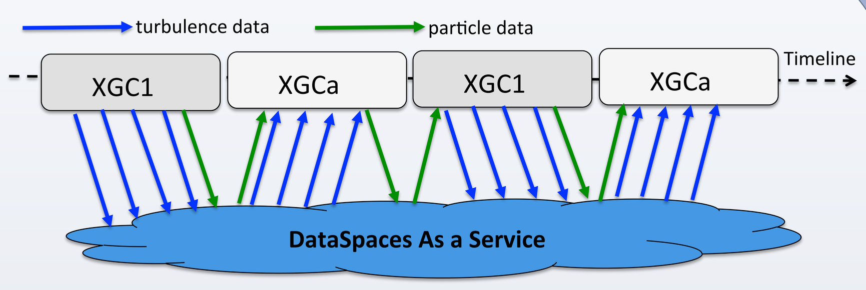


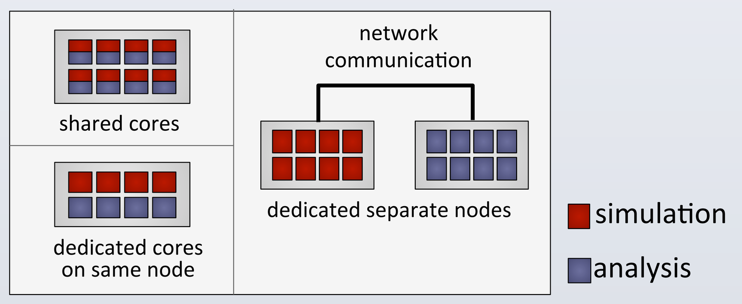
Illustration of using data staging service (DataSpaces-as-a-Service) to build the XGC1-XGCa workflow: connecting the different component applications, enabling data exchange and sharing.

EPSI coupling workflow

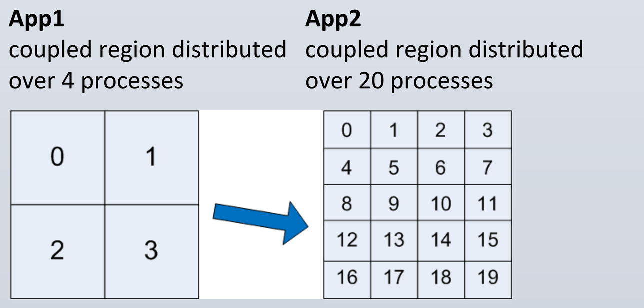
- Goal: To enable tightly coupled XGC1 and XGCa workflow using memory to memory coupling.

DataSpaces As a Service

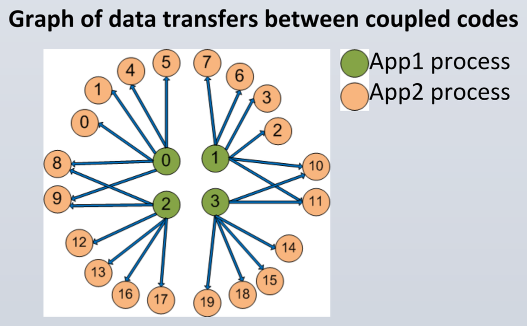
- Provides data staging as a service to applications – persistent data and staging services across application instances
- Allows XGC processes to dynamically connect/disconnect to/from the staging service
- New approach targets more complex and dynamic workflows with tighter coupling; provides more flexibility
- DataSpaces as a Service improves resource efficiency and increase I/O performance



Data-centric Mapping in DataSpaces



Locality-aware placement of coupled region processes on Cray XT5 12-cores computer nodes



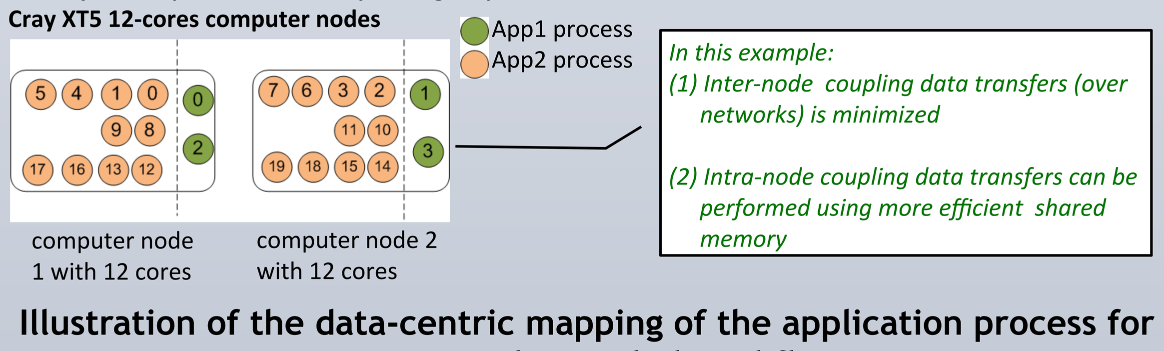
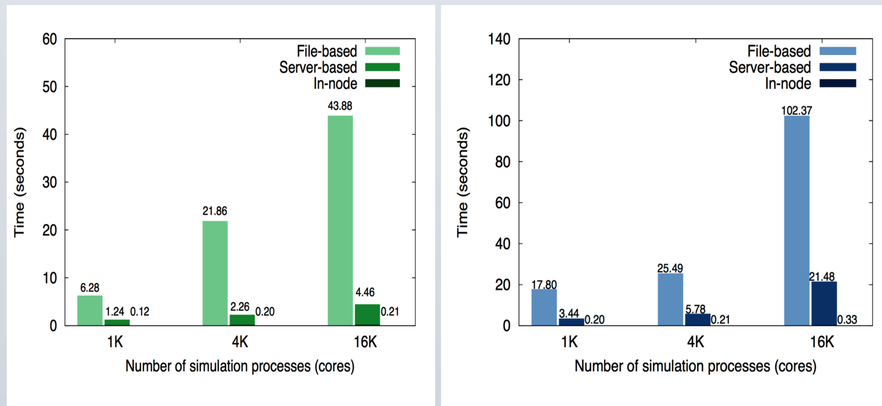


Illustration of the data-centric mapping of the application process for concurrently coupled workflow



Preliminary Results for In-node Staging

- Performance comparison with file-based staging and in-memory staging on dedicated servers
- Particle data read time reduced by ~98% compared to file-based and by ~90% compared to staging on dedicated servers
- Turbulence data read time reduced by ~99% compared to file-based and by ~96% compared to staging on dedicated servers