

Decaf: High-Performance Decoupling of Tightly Coupled Data Flows

Tom Peterka (PI), Argonne National Laboratory, tpeterka@mcs.anl.gov
 Franck Cappello (co-PI), Argonne National Laboratory, cappello@mcs.anl.gov
 Jay Lofstead (co-PI), Sandia National Laboratories, gflfst@sandia.gov

The need to distill enormous amounts of data into useful knowledge is pushing the limits of computational science. In many cases, to achieve high performance, programmers tightly couple the data analysis with data generation -- making the analysis interdependent and closely coordinated with the computation. Often, however, tight coupling limits the flexibility provided by individual modules. To address this issue, this project explores a hybrid approach that combines both types of coupling---tight and loose---in effect decoupling tightly coupled applications. The proposed work will result in a library of four data flow primitives: selection, aggregation, pipelining, and buffering. The research includes a software stack of three layers: a high-level data description of scientists' datasets, a data flow model built from the basic primitives, and a transport layer that will move data. A set of resilience strategies cross-cuts all three layers to enable the system to continue operating properly in the event of a failure of a component. The project will produce a method for automatically constructing data flows from these primitives, designed as a generic solution that other workflow and coupling tools can use. The software will be evaluated in three science applications: fluid dynamics, superconductivity, and cosmology. The aim is to improve performance, reduce power, mitigate errors, and enhance usability.

The proposed research, named Decaf, targets in situ methods and workflows in order to **improve performance, reduce power, add fault tolerance, and enhance usability**. It does not sacrifice performance or increase power: on the contrary, intelligently loosening rigid flows enables connectivity of a simulation to analysis tools that would otherwise default to expensive postprocessing. *Automatic pipelining* supports shallow copy adapters that transform data models, one chunk of data at a time, rather than in their entirety, relieving pressure at all levels of memory. *Automatic aggregation* reduces interconnect congestion. The ability to *select and permute* time-varying data in situ reduces the need to read entire datasets back into memory for temporal analysis. Similarly, the ability to *handle faults* during coupling avoids expensive restarts in case of errors. *Automatic buffering* allows bursty data production and consumption to be managed and faults to be contained. Because storage is just another producer or consumer in this design, it can be augmented with the same data primitives

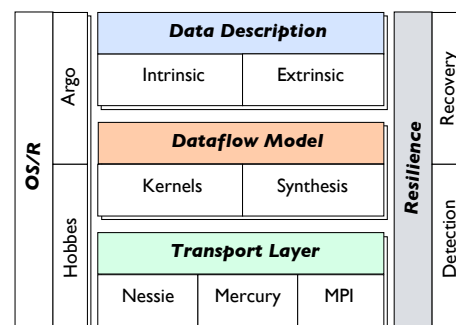


Figure 1: Overall project organization

and operate on data where they reside, further reducing postprocessing data round trips.

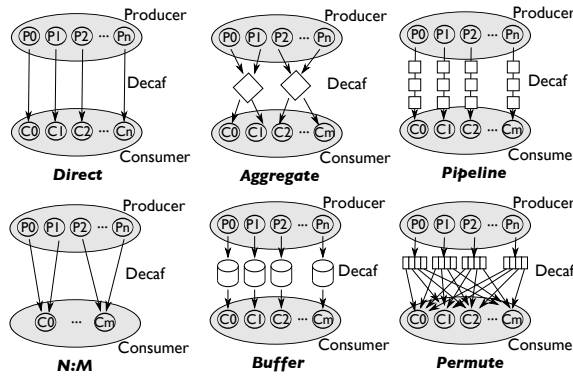


Figure 2: Decaf dataflow modes

Decaf integrates three levels of abstraction and software (Figure 1). A high-level data description captures data properties. The transport layer moves data. The dataflow model optimizes the performance of each link between producer and consumer (Figure 2). Cross-cutting the three layers are low-overhead resilience strategies combining detection and recovery techniques for fault tolerance. The computing environment is described by the

other cross-cutting component, the operating system and runtime, that provides capabilities such as system-wide control.

Three common use cases---data streams, data pipelines, and data networks---drive Decaf research. Data streaming will be decoupled in the in situ analysis of DOE simulations to trace particles in nuclear reactor Navier-Stokes computations. Data pipelining will be studied with several independent steps of in situ analysis to identify features in Ginzburg-Landau superconductivity simulations. Dataflow networks will be investigated in a workflow that transforms particles, meshes, and grids in Vlasov-Poisson N-body cosmology codes (Figure 3). Decaf is designed to integrate with and contribute to other ASCR-funded efforts through the principal investigators' involvement in SciDAC, co-design, and exascale research initiatives.

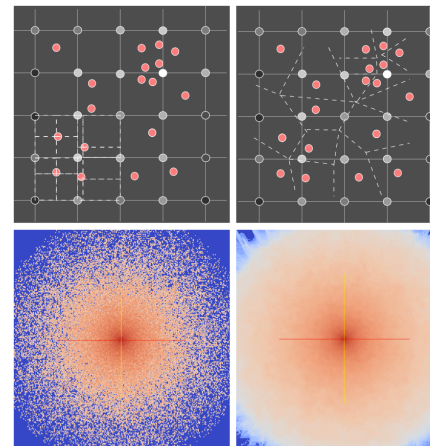


Figure 3: Density estimation of particle data from N-body cosmology simulations