

CyberIntegrator: A Highly Interactive Scientific Process Management Environment to Support Earth Observatories

¹Peter Bajcsy, ¹Luigi Marini, ¹Rob Kooper, ¹Jim Myers and ²Barbara Minsker

¹National Center for Supercomputing Applications (NCSA)

and

²Civil and Environmental Engineering (CEE) Department
University of Illinois at Urbana-Champaign (UIUC)

POC Email: pbajcsy@ncsa.uiuc.edu

Phone: 217-265-5387

Fax: 217-244-7396

Address: 1205 W. Clark St, Room 1008, MC-257, Urbana, IL 61801, USA

Abstract

This poster presents CyberIntegrator, a component of the Environmental Cyber-Infrastructure Demonstration (ECID) project at the National Center for Supercomputing Applications (NCSA), which is exploring cyberinfrastructure for environmental observatories with an emphasis on supporting exploratory analysis and end-to-end productivity. CyberIntegrator is a novel workflow-based system that supports interactive workflow creation, connection to external data and event streams, provenance tracking, and incorporation of workflow fragments and functionality from other systems and applications, all of which enable the types of tasks expected in observatories. The key features of the CyberIntegrator system are described below:

- **Interactive human-computer interface:** CyberIntegrator editor provides a user-friendly interface for browsing registries of data, tools and computational resources; creating workflows in a step-by-step exploration mode; re-using and re-purposing workflows; executing process flows locally or remotely; aiding research explorations using a provenance-to-recommendation pipeline; and incorporating heterogeneous tools and linking them transparently.
- **Integration of heterogeneous tools:** CyberIntegrator engine can execute tools from multiple heterogeneous software packages while hiding integration complexity. It passes the outputs of a software tool to the input of another tool transparently to an end user.
- **Event-driven execution:** Multiple CyberIntegrator instances and other software applications can communicate using a message enterprise bus and generate and consume events occurring within or externally to the CyberIntegrator to support event-driven distributed analyses. For example, analyses of streaming data coming from sensors create events that could trigger execution of tools and workflows.
- **Service execution:** A webservice execution engine allows publishing of meta-workflows created on the desktop to remote machines to share them with the community or run them on dedicated resources. Using a simple web interface the user can then run those workflows or remote resources and download shared meta-workflows to the desktop.

CyberIntegrator has been demonstrated with several environmental use cases including (a) modeling fecal coliform concentrations in Copano Bay, TX; (b) detecting anomalous data from sensors in Corpus Christi Bay, TX; (c) discovering vegetation variability from large remotely sensed images at the US continental scale and (d) predicting algal biomass in Illinois streams.