GEON Software Architecture
C. Baru, K. Bhatia, S. Chandra, K. Lin, A. Memon, and C. Youn
SDSC

Scientist in the GEON network are linked to distributed resources via the GEONgrid system. The term “grid” implies a single authentication domain for this environment, though the individual scientists are typically in different administrative domains. Authentication is via GSI (Grid Security Infrastructure). The implementation of the GEON software architecture is based on the Web services model.

The GEONgrid consists of a collection of “sites”, representing each PI or participating institution. Each site runs a GEONgrid node, which provides an entry point, or “point of presence (PoP)”, into the GEONgrid system. This “PoP node” is a Rocks node (based on Redhat Enterprise Linux) with a reference “GEON software stack”, which enables easier interoperation among nodes. Some GEON sites may provide additional capability either as “Data nodes” and/or “Compute nodes”. Data nodes provide additional disk capacity and host data sets that have been submitted to GEON, with possible replication of data among such nodes. Compute nodes provide additional computing capability. A single certificate authority (CA) provides authentication in this environment.

The services, tools, and applications developed in GEON are accessible from clients as Web services, but using GSI authentication. The GEON Portal provides authenticated access to this environment as well. The portal is implemented using GridSphere and a portlet-based framework, and provides access to a rich set of functionality including the ability to register data and ontologies into the system; smart search capability; access to a variety of geoscience-related tools; access to Grid-enabled geoscience applications; and a customizable private work space, referred to as “myGEON”, from which users have access to the Kepler scientific workflow environment, which allows them to easily author and execute processing “pipelines” for data analysis, modeling, and visualization.

The GEON software architecture supports a number of novel and pioneering distributed system concepts including, the Grid Account Management Architecture (GAMA), the notion of “Distributed Portals”, Data and Ontology Registration, Knowledge-based GIS Map Integration, and ad hoc data integration using “integration carts”, which are analogous to the notion of “shopping carts” on the Web.

The GEONgrid system also provides a number of administrative capabilities including data backup (from Data Nodes to long-term archives), system monitoring (using the Inca framework that was developed as part of the TeraGrid system), and performance benchmarking (using software “probes” that were developed as part of the GRASP project at SDSC).