

## **14 Million Sequences Analyzed with Argonne's Grid-Enabled System**

GADU/GNARE, the Genome Analysis and Database Update system developed at the Mathematics and Computer Science division of Argonne National Laboratory, has successfully analyzed more than 14 million sequences in the past year – an amount of data equal to 3,500 average bacterial genomes. The system is one of the world's first Grid-enabled high-throughput computational environments designed to provide researchers with up-to-date sequence information.

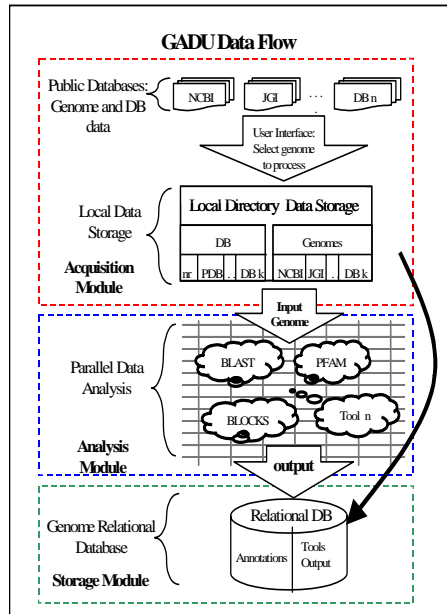
GADU/GNARE was developed as a collaborative effort between Argonne's bioinformatics group, led by Natalia Maltsev, and the Globus Alliance group, led by Ian Foster. Globus Toolkit software and related components are used to harness hundreds of distributed processors from the GRID2003, NSF TeraGrid, and the DOE Science Grid, as well as from Argonne and Pacific Northwest national laboratories, in order to perform high-throughput computations.

According to Foster, the Grid-enabled system allows efficient automation of the major steps of genome analysis, including data acquisition from diverse genomic databases and data analysis by various bioinformatics tools and algorithms. GADU/GNARE also expedites the process of storing the results of analyses and annotations.

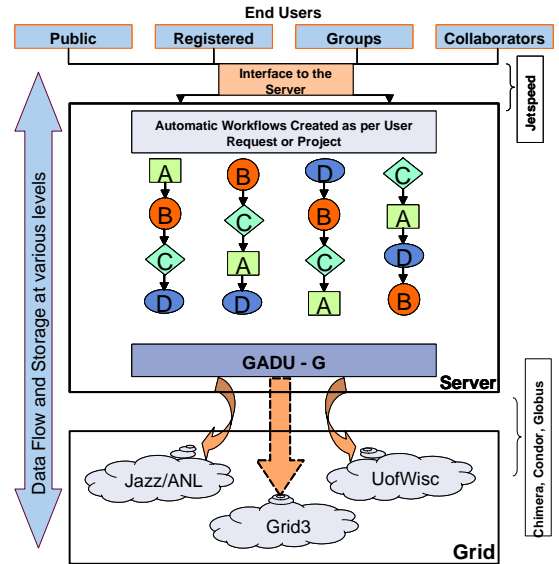
The flexible architecture enables users to tailor the genome analysis process to their individual needs, it can function in an automated mode as well as interactively through a Web-based interface

GADU/GNARE has been used extensively by the computational biology group at Argonne as well as by the University of Chicago SEED project, the NIH Midwest Center for Structural Genomics, the NIH Great Lakes Regional Center of Excellence for biodefense and emerging infections, and several DOE Microbial Genomes projects.

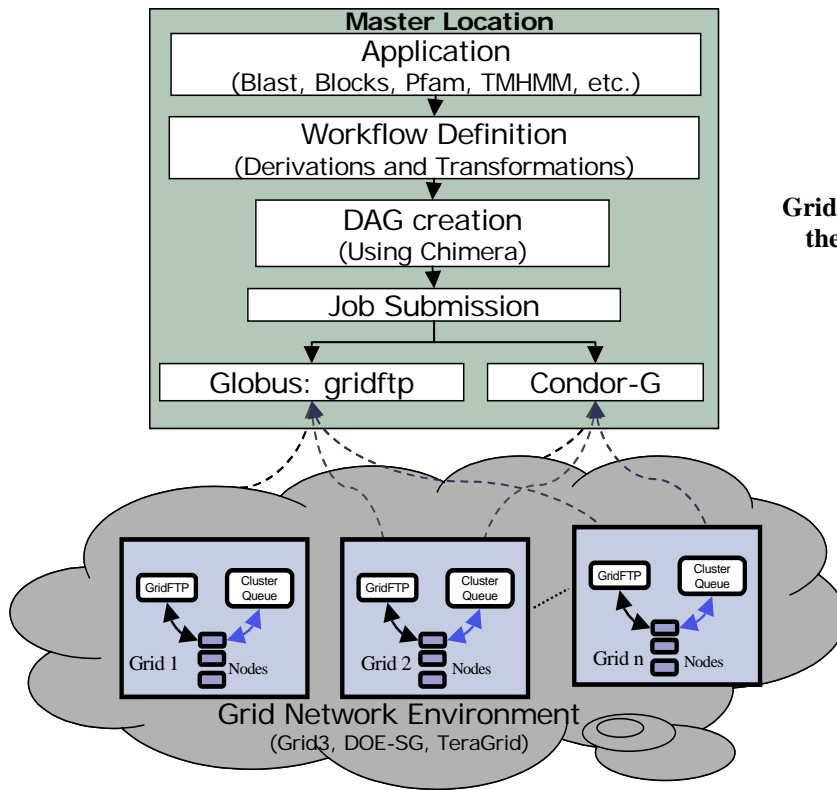
GADU is funded by the NSF National Computational Science Alliance as part of the Data Quest expedition, led by Ian Foster (ANL, [foster@mcs.anl.gov](mailto:foster@mcs.anl.gov)) and Robert Brunner (NCSA, [rb@ncsa.uiuc.edu](mailto:rb@ncsa.uiuc.edu)) and the Scientific Workplaces of the Future expedition, led by Rick Stevens (ANL, [stevens@mcs.anl.gov](mailto:stevens@mcs.anl.gov)).



**GADU workflow, showing acquisition, analysis, and storage modules.**



**GADU system environment with user interface**



**Grid execution pattern for the GADU application**