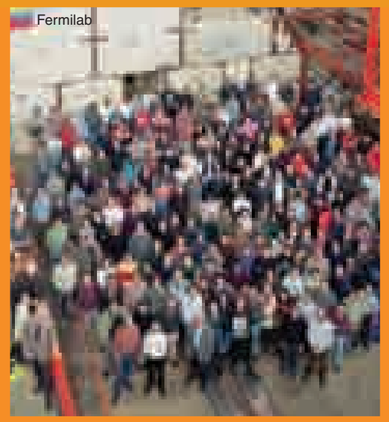
What is the OSG?


Fermilab

Q. Wei, D. Ebert and G. Klimeck, Purdue University

A National Cyberinfrastructure for Science

The Open Science Grid makes innovative science possible by bringing multidisciplinary collaborations together with the latest advances in distributed computing technologies. This shared cyberinfrastructure, built by research groups from United States universities and national laboratories, includes over 10,000 CPUs and many terabytes of data storage.

Scientists from diverse fields contribute manpower and resources to the OSG and benefit from easy access to local and remote resources; testing and production environments for middleware and applications; and a common computing infrastructure. The OSG includes the Integration Grid, a test bed for new grid technologies and applications, and the Production Grid, a fully supported environment for sustained applications. OSG software is based on the NSF Middleware Initiative distribution and the Virtual Data Toolkit, with enhancements targeted to meet the needs of science researchers around the world.

Enhancing Collaboration

Grid computing gives scientists an extraordinary range of computational and storage resources and tools to advance scientific collaboration. The OSG Consortium invites faculty and researchers from all scientific disciplines, middleware developers, resource providers, and national and international grid projects to join this unique distributed computing infrastructure. Over 20 member organizations representing more than 50 institutions and hundreds of researchers contribute to the OSG and benefit from access to shared resources worldwide.

Creating a Worldwide Federation of Grids

The OSG combines resources from many different computing infrastructures, and works jointly with its partners, such as Enabling Grids for E-sciencE (EGEE) and TeraGrid, to create a worldwide interoperable system for cutting-edge research. Gateways now under development allow scientists to run the same application on multiple grids, while computer science researchers build bridges between different infrastructures to allow them to communicate freely.

Science on the OSG

Grid computing and the OSG allow scientists to carry out research faster, more collaboratively and more efficiently, and to bring together data, ideas and tools to address complex problems in new ways. Scientists contribute to the OSG and benefit from access to a range of computational and storage resources, tools to advance national and international collaboration, and a framework to share resources and applications across scientific and regional boundaries.

Astronomy and Astrophysics

- The Sloan Digital Sky Survey used the OSG for the data- and processing-intensive task of combining several images of a southern portion of the sky to allow the detection of very faint and distant astronomical objects.
- Grid computing helps Laser Interferometer Gravitational-Wave Observatory (LIGO) collaborators find, transfer and analyze data to search for gravitational waves, and simulate the gravitational-wave sky.

Biology and Medicine

- Argonne National Laboratory's Genome Analysis and Database Update system (GADU) uses resources on the OSG to compare individual protein sequences against all annotated sequences in publicly available databases.
- OSG resources greatly reduce the time needed to analyze and normalize magnetic resonance imaging data at Dartmouth College.

Particle and Nuclear Physics

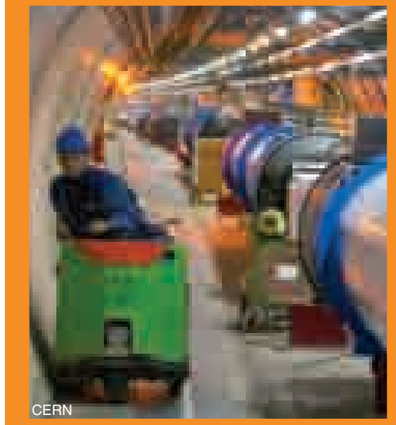
- United States participants in the ATLAS and CMS particle physics experiments use the OSG to simulate over 100 million events per year in the detectors, and will transfer and analyze petabytes of data when the Large Hadron Collider begins operation in 2007.
- OSG tools enable facilities around the globe to churn through 250 terabytes of data from Fermilab's DZero experiment, processing the raw events into a form usable by physicists in their search for the nature of matter and energy.
- The STAR GRID links collaborators worldwide to nuclear physics data collected at Brookhaven National Laboratory's Relativistic Heavy Ion Collider and to OSG resources.

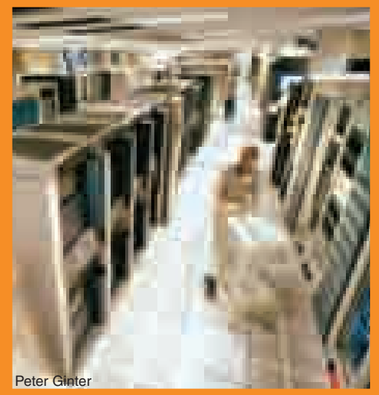
Computer Science

- The GriPhyN Virtual Data System and the Virtual Data Toolkit were developed with the collaboration of the iVDGL and GriPhyN projects, which participate in the OSG. The VDS provides a set of tools for expressing, executing, and tracking the results of workflows, and the VDT is a collection of grid middleware that can be easily installed and configured.

Sloan Digital Sky Survey

Gideon Caplovitz and James Dobson, Dartmouth College


CERN


Peter Ginter