

Endorsing and Using The Virtual Data Toolkit As a Common Approach.

News Update – 5 Mar 2004



Grid Physics Network
Particle Physics Data Grid
International Virtual Data Grid
Laboratory
the Trillium Collaboration

What is the VDT?

The Virtual Data Toolkit (VDT) is an ensemble of grid middleware that can be easily installed and configured. The goal of the VDT is to make it as easy as possible for users to install, maintain and use grid middleware. Ideally, you just type a single-command and you can consume grid resources or make your resources available to others. In reality, it is a bit more work than that, but not much.

The History of VDT

VDT was originally created to serve as delivery channel for grid technologies developed and hardened by the NSF funded GriPhyN and iVDGL projects. These two projects are the primary sources of funding for VDT activities. However, the role of VDT has expanded over the last year and now supports the LHC Computing Grid Project (LCG) and the Particle Physics Data Grid (PPDG). U.S. CMS and U.S. ATLAS were early adopters of VDT and played an important role in enhancing the functionality and robustness of the toolkit components. Grid2003 has provided a 26-site scale installation of a common version of the toolkit. All Experiments on PPDG are transitioning to the use of VDT.

What Components are in VDT?

VDT contains three kinds of middleware, roughly speaking

- **Basic Grid services:** Includes Condor-G and the Globus Toolkit.
- **Virtual Data Tools:** Includes software tools developed by the Griphyn project to work with virtual data, particularly the Virtual Data System.
- **Utilities:** Includes a wide variety of tools, such as GSI-Enabled OpenSSH, software to update GSI certificate revocation lists, and monitoring software like MonaLisa.

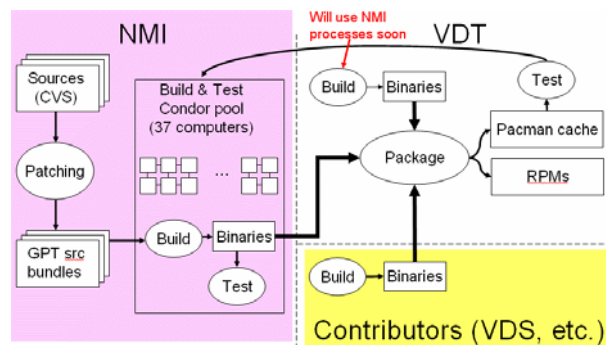
The PACMan utility is used to package and deploy the toolkit. Detailed contents of the latest VDT release are given on the next page.

How is VDT built and tested?

Most VDT components are built using machinery (software and computers) developed and deployed by the NSF National Middleware Initiative (NMI). The picture below illustrates how the VDT is built and packaged.

National Middleware Initiative

NMI builds several components for VDT: Globus Toolkit, Condor, MyProxy, KX509, GSI OpenSSH, PyGlobus, and UberFTP. NMI checks the software out of the appropriate CVS repositories, then patches the software. Each of these software packages uses the Grid Packaging Toolkit (GPT) utility, so GPT source bundles are then created. These bundles are built and tested on the NMI build pool, then given to VDT.



The NMI build pool is a cluster of nearly forty computers. There are a wide variety of computer architectures and operating systems in this pool. When a VDT build is done, Condor runs the same build on three architectures: RedHat 7, RedHat 9, and RedHat 7 with the gcc 3 compiler. After the build, NMI automatically deploys the software and does basic verification testing to ensure that the software works. After the build completes, the VDT team imports the software into the VDT cache. There is close collaboration between the VDT and NMI groups.

What is in VDT 1.1.13?

Starting VDT 1.1.13 VDT-Server and VDT-Client have been unified to form a single component called VDT. This merge was necessary to allow for easier install, upgrade and uninstall feature for individual components. VDT-SDK still remains as an individual component. The SDK contains libraries to develop new software.

VDT

- [Chimera Virtual Data System, v1.2.3](#)
- [Condor/Condor-G 6.6.1](#)
 - [VDT Condor configuration script](#)
- [Fault Tolerant Shell, v2.0.2](#)
- [Globus Toolkit, v2.4.3](#) + [patches](#) (gatekeeper, MDS, GridFTP, etc.)
 - Alternative Globus Gatekeeper with [authorization callout](#)
 - Globus 2.4.3 replica management
 - [VDT Globus configuration script](#)
 - CA signing policies / tools for the [Globus, DOE and EDG CAs](#)
 - Glue CE, SE, CESEBind Schemas, and information providers
 - Glue CE Schema 1.1 (CVS version 1.2, 23-June-2003)
 - Glue SE Schema 1.1 (CVS version 1.5, 29-July-2003)
 - Glue CESEBind Schema 1.1 (CVS version 1.2, 29-July-2003)
 - Glue basic information providers 1.1 (CVS version 1.79, 25-July-2003)
 - A basic Glue static information setup
 - [More information about Glue](#)
 - EDG mkggridmap 2.1.0 and related perl modules
 - EDG CRL Update
 - [RLS, v2.1.4](#) (MySQL/ODBC optional)
- [GSI OpenSSH, v3.0](#)
- [Java JDK-1.4.1](#)
- [KX509](#)
- [MonaLisa, v0.95](#)
- [MyProxy 1.11](#)
- [PyGlobus, v1.0](#)
- [UberFTP-Client, v1.3](#)
- [VDT Version](#)
- [VDT System Profiler](#)

VDT SDK

- [Globus 2.4.3](#) SDK + advisories
- [ClassAds 0.9.5](#)
- [RLS 2.1.4](#) SDK
- [PyGlobus, v1.0](#)
- [Netlogger 2.2](#)
- A Netlogger-instrumented globus-url-copy look-alike program (written in python).
- [VDT Version](#)
- [VDT System Profiler](#)

Other optional packages (actual Pacman package names specified in **bold**)

- **Globus-LSF-Setup** - configures GRAM jobmanager/reporter for LSF
- **Globus-PBS-Setup** - configures GRAM jobmanager/reporter for PBS
- **Globus-RLS-Server-Setup-MySQL** - installs [MySQL](#), ODBC and sets up MySQL databases for RLS
- **FBSNG** - [Farms Batch System Next Generation](#)
- **VDT-Test** - [VDT Certification Tests](#)

<http://www.lsc-group.phys.uwm.edu/vdt/contents-1.1.13.html>