



KnowARC

Grid Computing with Debian, Globus and ARC

Mattias Ellert, Uppsala Universitet (.se)

Steffen Möller, Universität zu Lübeck (.de)

Anders Wäänänen, Niels Bohr Institutet (.dk)



- ❖ Seamless integration of distributed computing and storage resources from the user's point of view
- ❖ Computing grid vs. power grid analogy
 - ❖ Power grid: users plug in their electrical devices and don't need to care which power plant provides the electricity (unless they want to)
 - ❖ Computing grid: the user prepares a computing task and sends it to the "grid" and doesn't need to care which cluster performs the calculation (unless it wants to)



❖ Volunteer Computing: BOINC

- ❖ “single regular users fetch prepared workunits”
- ❖ regular Debian client package
- ❖ unofficial server packages

❖ Computational Grids

- ❖ “big compute clusters wait for arbitrary jobs”
- ❖ no previous packages for any Linux distributions
- ❖ common IT backbone for High Energy Physics



❖ Network of trust

❖ Users trust sites

- Data security, validity of installations

❖ Sites trust users

- All usage can be traced back to the user

❖ X.509 certificates

❖ Certificate Authorities (CAs) guarantee identities

❖ User creates time-limited variants of these certificates (proxies) to delegate their rights to jobs



- ❖ International Grid Trust Federation (IGTF)
 - ❖ CAs that trust each other's policies
 - ❖ Users with a user certificate issued by a member CA can authenticate to resources that have host certificates issued by any other member CA
- ❖ Virtual organisations
 - ❖ Clusters in the grid delegate decision over admissions to virtual organisations
 - ❖ Easiest: a website collecting the individuals' certificates and descriptive names



✦ Submission of Job

- ❖ Task should be described in a job description – executable, input data, output data, software and hardware requirements, ...

✦ Status information

- ❖ Query the state of clusters and jobs

✦ Retrieval of results

- ❖ Download to client or (if specified in the job description) automatically upload to storage

✦ Data management

- ❖ Keep track of large sets of input and output files



- ✦ Make grid access easier
 - ❖ Local vs. grid accounts
- ✦ Increase flexibility
 - ❖ Migration of jobs
 - ❖ Preparation of runtime environments
- ✦ Increase public awareness
 - ❖ Universities and research groups
 - ❖ Industry
 - ❖ Computer clubs
 - ❖ Presentations like this one ;-)



✦ Globus

- ❖ can be used as a complete grid middleware
- ❖ is a library of core functionalities for many

✦ Unicore

- ❖ both Grid and Grid Infrastructure

✦ EGEE

- ❖ uses the gLite grid middleware and Globus

✦ NorduGrid

- ❖ with or without Globus
- ❖ compatible with the others

- ❖ A set of libraries and tools for grid computing used by many grid projects
 - ❖ Globus security infrastructure (GSI)
 - Authentication and authorization based on short lived proxy certificates
 - standardized as RFC 3820
 - ❖ GridFTP
 - Extensions to the FTP protocol to support GSI authentication, third-party transfers, multiple data channels for parallel transfers, partial file transfers
 - “proposed recommendation” document in the Global Grid Forum (GFD-R-P.020)



- ❖ Source
 - ❖ Distributed as >100 MB tarball
 - ❖ Contains ~300 inter-dependent packages within
- ❖ Split into individual packages to become manageable
 - ❖ Strong consistency between Globus and Debian packages
- ❖ Build uses the Grid Packaging Toolkit (GPT)
- ❖ Patches communicated back to upstream



- ❖ Redundancies with system libraries are all eliminated from the source tree
 - ❖ e.g. openssl, openldap, libltdl
- ❖ Glue packages are provided instead
 - ❖ providing GPT metadata information for system packages to satisfy build dependencies
- ❖ Status
 - ❖ First packages uploaded to Debian new queue, also uploaded to Fedora



- ❖ Regular package for Grid Package Toolkit
- ❖ Use GPT packaging metadata information to autogenerate Debian folders in source code management system
- ❖ Manual curation of these folders
 - ❖ preparation of patches
 - ❖ provisioning of better descriptions



- ✦ Advanced Resource Connector
- ✦ Grid middleware built on top of the Globus libraries, with higher level services
- ✦ Used by the Nordic Data Grid Facility (NDGF) to provide computing resources for
 - ❖ High Energy Physics researchers at the CERN Large Hadron Collider
 - ❖ Bioinformatics
 - ❖ Quantum chemistry
 - ❖ ...












Monitor of clusters contributing






ARC Grid Monitor

2009-02-06 CET 17:41:17

Prozesse: ■ Grid ■ Lokal

Land	Site	CPUs	Last (Prozesse: Grid+lokal)	In einer Queue
 Australien	Alfred (UniMelb)	124	<div style="width: 100%;"><div style="width: 29%; background-color: gray;"></div></div> 0+29	0+0
 Dänemark	Benedict - Aalborg pr>	52	<div style="width: 100%;"><div style="width: 30%; background-color: gray;"></div></div> 0+30	0+0
	Fyrkat (DCSC/AAU)	656	<div style="width: 100%;"><div style="width: 16%; background-color: gray;"></div></div> 0+160	0+584
	LSCF (NBI)	20	<div style="width: 100%;"><div style="width: 10%; background-color: green;"></div><div style="width: 90%; background-color: gray;"></div></div> 20+2	273+0
	Morpheus (DCGC/NBI)	13	<div style="width: 100%;"><div style="width: 0%; background-color: gray;"></div></div> 0+0	0+0
	Steno (DCSC/KU)	2296	<div style="width: 100%;"><div style="width: 44%; background-color: green;"></div><div style="width: 56%; background-color: gray;"></div></div> 1008+1097	427+ -417
 Deutschland	Uni Lübeck - INB	16	<div style="width: 100%;"><div style="width: 7.5%; background-color: gray;"></div></div> 0+12	0+0
 Finnland	Akaatti (M-grid)	200	<div style="width: 100%;"><div style="width: 41%; background-color: gray;"></div></div> 0+82	0+40
	Ametisti (M-grid)	260	<div style="width: 100%;"><div style="width: 54%; background-color: gray;"></div></div> 1+136	0+232
	Jaspis (M-grid, HIP)	14	<div style="width: 100%;"><div style="width: 0%; background-color: gray;"></div></div> 0+0	0+0
	Kiniini (CSC)	72	<div style="width: 100%;"><div style="width: 42%; background-color: green;"></div><div style="width: 58%; background-color: gray;"></div></div> 30+0	0+0
	Kvartsi (M-grid)	192	<div style="width: 100%;"><div style="width: 63%; background-color: gray;"></div></div> 0+123	1+10
	Liuske (CSC test)	8	<div style="width: 100%;"><div style="width: 0%; background-color: gray;"></div></div> 0+0	0+0
	Murska	2176	<div style="width: 100%;"><div style="width: 100%; background-color: gray;"></div></div> 0+2120	0+0
	Opaali (M-grid)	88	<div style="width: 100%;"><div style="width: 95%; background-color: gray;"></div></div> 0+84	3+123
	Sepeli	512	<div style="width: 100%;"><div style="width: 59%; background-color: green;"></div><div style="width: 41%; background-color: gray;"></div></div> 316+153	2843+0
	Spektroliitti (M-grid)	26	<div style="width: 100%;"><div style="width: 0%; background-color: gray;"></div></div> 0+0	0+0
	Topaasi (M-grid)	82	<div style="width: 100%;"><div style="width: 89%; background-color: gray;"></div></div> 0+70	0+0
 Island	Jotunn (Uol)	168	<div style="width: 100%;"><div style="width: 0%; background-color: gray;"></div></div> 0+0	0+0
	RHI-CSD	1	<div style="width: 100%;"><div style="width: 0%; background-color: gray;"></div></div> 0+0	0+0



- Available today from www.nordugrid.org
- version 0.6.x
 - ❖ “Production” release
 - ❖ full Globus dependency
 - ❖ Globus packages should be accepted first
- version 1.x
 - ❖ ongoing development
 - ❖ optional Globus dependency
 - ❖ Debian packages will offer the more compatible Globus-dependent version



- ❖ Increased connectivity
 - ❖ between users of Debian
 - ❖ in between clusters of Linux distributions
- ❖ Promotion as an extended concept of the Debian society
 - ❖ the sharing of packaging may be extended towards a sharing of resources
- ❖ Debian Technologies
 - ❖ packages are perfect descriptions for runtime environments
 - ❖ availability on many heterogeneous platforms



- ✦ KnowARC – www.knowarc.eu
 - ❖ European Commission 5th framework programme project
- ✦ NDGF – www.ndgf.org
- ✦ The developers of Globus – www.globus.org
 - ❖ Charles Bacon in particular, for his integration of patches
- ✦ The developers of NorduGrid ARC – www.nordugrid.org