CODAS

UKAEA
Restoring our Environment

JDN/H(06)042
Issue 1: 25/09/2006

UNCONTROLLED WHEN PRINTED

Installation Issues for the EFDA-JET Grid Site

Author: D.W.Robson

Abstract

This document outlines some of the problems encountered in the installation of the EFDA-JET Grid Site. These findings are to be presented to the EIROForum Grids Group by Dr. J.W.Farthing. EIROforum is a partnership of Europe's seven largest inter-governmental research organisations.

R.O.: D.W.Robson	Reviewers: K.Purahoo	Approver: J.W.Farthing

Distribution
CODAS, NetworkServTeam@jetnet.jet.uk

Table of Contents

1 Introduction 3
2 Documentation Problems 3
3 Networking Problems 3
4 Kernel Problems 4
5 Problems with the Default Configuration 4
6 Support Problems 4

1 INTRODUCTION

The EFDA-JET Grid Site was envisaged following Dr. J.W.Farthing's meetings with the EIROForum Grid Group during 2005. Initial planning began in January 2006, and the hardware was built in March. EFDA-JET was finally certified as a production site in the LCG grid in August 2006. Some of the issues arising from the installation of the site are outlined below.

2 DOCUMENTATION PROBLEMS

The first hurdle was trying to get an overall picture of EGEE/glite/LCG. We did not find the documentation very useful. The best we found was the LCG-2 Middleware Overview. However, this document is two years old, in a rapidly changing field, and is still at version 0.1 (first draft).

The other document we made much use of initially was "Generic Installation and Configuration". This was a very important document for us, and we monitored it during our installation problems. Unfortunately we missed some important information because the document had been updated, but not its version number. The last change of this document I am aware of was on September 5th, but it is still at the same version (3.0.0) as the copies we pulled of the web in March.

We eventually worked out which services we needed to run at our site, in which services we should need from Tier 1 or other sites. However, the documentation did not tell us which remote sites to use. It proved impossible to debug problems from our end alone, and we relied on the assistance of other site managers. Unfortunately this meant fixing one problem at a time via email exchanges, which was often lengthy during the holiday season, and because the other managers had their sites to maintain. In addition, we had to change the configuration to point to different remote servers as our site progressed through the different stages; uncertified -> candidate -> production. Maybe the Regional Operation Centres could provide more detailed templates for the configuration files, at the onset.

After considerable time installing LCG 2.7.0, all sites received a message saying that they had to upgrade to glite 3.0.0 within two weeks. Despite following the email lists for the last four weeks, there seemed to be no previous mention. We abandoned LCG 2.7.0 and began installing glite 3.0.0. Some longer-term notification about this would have been nice, to assist out planning. Perhaps a road map is required?

During the installation, we had to provide names of the types of nodes that we needed to build. It was surprisingly difficult to determine these "target" names, as they weren't listed in the documentation at that time. Other site managers pointed us to where we could locate them in the source code. Again, the documentation needs to keep up with the source code!

3 NETWORKING PROBLEMS

One major problem we encountered arose from the fact that the grid middleware is not very firewall friendly. In fact it requires more than five thousand ports to be open. Worse still, the majority of the ports are configurable at each site, so in order to guarantee connectivity with all the other sites

in the LCG, we were forced to open all incoming and outgoing ports between our cluster and the outside world.

A more difficult problem to overcome arises from the security components of the grid middleware, where many of the nodes have x509 certificates. The security components do not work well with Network Address Translation (NAT) firewalls. Most sites do not use NAT for their grid nodes, but instead assign external IP addresses for each node. This was not a viable solution at JET because of the very small number of IP addresses that we have available.

Instead, a more complex solution was adopted where the Worker Nodes had private IP addresses and the nodes that provided services across the internet had their interfaces configured to use public and private IP addresses. This approach raised problems with the information and monitoring services, as their openldap servers seemed to bind only to one IP address, as opposed to most services that bound to sockets with the wrong IP address rather than using the INADDR_ANY flag that binds to all of the IP addresses.

4 KERNEL PROBLEMS

Following the LCG documentation, we installed the recommended operating system release (Scientific Linux 3.0.5). This is in fact an ancient version, as its kernel is version 2.4.21, which was released in June 2003. Consequently, we were apprehensive that it wouldn't detect or properly manage the hardware in a modern PC. This fear proved to be correct, as it didn't detect our on-board Marvell gigabyte network interfaces. However, we did manage to install a driver that we downloaded.

Over the next few weeks, we noticed frequent hangs on nodes on the grid cluster. These were accompanied with messages to the console complaining about lost interrupts to the hard disk drive. These incidents occurred more often during time of high I/O load, particularly during the backups. After some investigation, we pinpointed it to problems with the APIC (Advanced Programmable Interrupt Controllers) module. These problems have been resolved with the newer kernels, such as the version we run on our internal linux cluster. There is a kernel option to run without APIC, but when we tried it the Marvell gigabit card failed to load. Finally, we had to buy netgear GA311 network cards for each node, and were able to run without the APIC controller.

5 PROBLEMS WITH THE DEFAULT CONFIGURATION

The installation and configuration of a site involves tailoring a configuration file. A template for this file comes with the installation tool, yaim. However, there were several problems with the default file. One example was the definition of the operating system. The format in the template was invalid and later caused our site to be flagged as failing the Site Functional Tests. Other problems arose because the template made assumptions that a CLASSIC_HOST node was being used, when in fact this type of node is obsolete.

6 SUPPORT PROBLEMS

One of the methods of getting assistance is to open a Global Grid User Support (GGUS) ticket. We did so to report a problem with a resource broker. GGUS documentation states “to ensure 24/7 support, GGUS has three teams in different time zones. “. However despite this promise of speedy response, our ticket was opened on the 9th September, and after 16 days (and counting), we have only had an automated response to say that the ticket has been received. This is despite the fact that our ticket was flagged as urgent, and indicated that our users could not submit jobs.