Designing and implementing ethernet networks

Initial project report, HOD190

HØGSKOLEN I BERGEN

Joachim Tingvold

Bergen University College

ALICE HLT, CERN

Abstract

A brief explanation of the project to be carried out; that is to design a network for a computer cluster from scratch, both physically and logically. A description of my employer is presented, together with requirements – herein methods, software and hardware, to be needed/used. A preliminary progress-plan is also discussed.

Contents

1

Introduction

1.1 Employer

I work at CERN, which is an international organization whose purpose is to operate the worlds largest particle physics laboratory. It's situated on the border between France and Switzerland, near Geneva. It was established in 1954, and currently has twenty European member states.

1.1.1 ALICE

The experiment I'm working for, is called ALICE – which is an acronym for ***A L**arge **I**on **C**ollider **E**xperiment* – and is one of the largest experiments in the world devoted to research in the physics of matter at an infinitely small scale.

1.1.2 HLT

HLT, which is an acronym for ***H**igh **L**evel **T**rigger*, is a part of the ALICE-experiment. It combines and processes all the information from all the major detectors of ALICE in a large, high-performance computer cluster. It's task is to select the relevant part of the huge amount of incoming data. This high-performance cluster consists of several hundred servers, which is connected together with two types of network; ethernet and IP-over-InfiniBand (IPoIB).

1.2 The project

My project is to redesign, implement and configure the ethernet network for the high-performance computer-cluster in HLT. Usually you cannot make such large interventions, as it causes downtime, which isn't possible due to the fact that the ALICE-detector gathers data 24/7 from the LHC. However, since the LHC is being shut down for a longer period for the first time in a year, it gives us the opportunity to do larger interventions.

1.3 Web

The project has it's own homepage, and can be reached by visiting http://cern.jocke.no.

2

Aims

The aim is to end up with a network that performs without trouble, even under high load. It should have high availability and redundancy. We also want to introduce a separation, so that servers/services that are critical, has limited access, so that the chance of something (or someone) intervening is at an absolutely minimum.

2.1 History

Since a lot of the work on and around the ALICE-detector has been on-going since the 90's, there are also a lot of decisions that has been affected by this. A lot of things are the way they are today, due to the fact that it was planned a while ago. This is not always fortunate, since specific things should be overhauled now-and-then. Computer networks is one of those things. Since computer science is changing fast, one should adapt to these changes. Not necessarily at the same pace, but at least more often than once every decade. The current ethernet network in the HLT-cluster is therefore affected by this, and is also why this restructuring is really needed.

2.2 Current layout

The current network-layout is as flat and non-configured as you'll ever get it (it's literally plug-and-play). The switches has zero config, and the entire network is in the same subnet; *10.162.0.0/16*. Even if the cluster has 2^{16} IP-addresses available, they have problems finding free IP-addresses (even though they don't consume more than, at most, 2^{10} IP-addresses). All the switches are located in one place, and is the cause of a huge cable-mess, which makes it really hard to debug errors – at least if you have to find out the path of the cable(s). There is also a lot of old mess laying around, such as intermediate home-switches (that is, small 8-port switches) that are interconnected and daisy-chained. There is a saying that ”if it ain’t broke, don’t fix it”, which, together with time-constraints, probably is why the network is in the state that it currently is.

2.3 Main aims

There are a few, main aims that we’re going for. There is room for lots of improvements, but we’ll take the major improvements first, and when things are stabilizing, we can start looking into fine-tuning it.

2.3.1 Physical

The first thing we want to do, is to get rid of the centralized switch-placement. Instead of having all the switches in one place, we want to distribute them into the racks, and connect the servers directly to the switches (thus removing a source of error – the extra patching). This distribution also reduces the amount of cables needed at the core-switches with over 80%, which is a great deal!

2.3.2 Redundancy & availability

In addition to spread out the switches, we also want to increase the number of uplinks between all of the switches, and even more between the core-switches. This is both to ensure high bandwidth, but also to make sure we have enough redundancy; we should be able to lose any given cable, without loosing connectivity. Redundant paths is also something we want; we should be able to lose any given core-switch, without having noticeable downtime. We’re also going to install two new gateways, which will be fully redundant (which isn’t the case with the current gateways – they require manual intervention to be able to ”failover”).

2.3.3 IP-scheme

The current IP-scheme was poorly designed. It had a logic system that relied on bit-flipping. This means you could, by just looking at an IP-address, figure out in what rack, and what position the server that used that IP-address (since each IP was statically "assigned" to a specific height and rack). This is all sweet and fluffy – but not really practical. Therefore a complete re-do of the IP-scheme is wanted. Divide it up into subnets of suitable sizes (but still allow for future growth), and have a subnet for each kind of server/service.

2.3.4 Security

Security is also a concern – the cluster has previously, on several occasions, been rooted. This is partly caused by bad network-security (or rather, the lack of it), and partly by outdated, and poorly configured, Linux-installations. This is something we're going to fix. Updating the Linux-installations, using more secure gateways (OpenBSD), and generally limit the connections to the outside world (a long way to go, due to all the nice hacks we have in the cluster – the one worse than the other regarding security).

2.3.5 Separation

Since the cluster consists of different "sections"; we have a production-cluster, development-cluster, infrastructure-machines, headnodes, etc. During physical runs, the production-cluster should be as isolated from the rest of the cluster as possible. This is to ensure that something, or someone, doesn't interrupt anything. This is known to happen with the current setup.

2.3.6 Avoid loops

Optimizations regarding layer 2 paths is also something they've had problems with; loops has occurred in the network several times. This is why we want to configure STP (no, it's not configured on HP-switches by default) properly on all the switches, and ensure that it's Working As Intended™.

2.3.7 Services

The current setup relies heavily on LDAP – users, DHCP, DNS, basically everything, is stored in the LDAP-database. This is fine, however, the learning-curve for LDAP is rather steep, so we're moving away from this, and towards flat-files. This simplifies things a lot, and we don't have to rely on yet another service that can fail. One less thing to maintain as well.

2.4 Methods

There are numerous ways one can achieve what we want. You could do most of the changes we want to do, without physically move the switches, as an example. However, the changes we want to do, is in many ways the most ideal, at least considering the equipment we currently have available.

2.4.1 Hardware

On the hardware-front, the only things we need, are some new power-cables for the switches (which is needed in order to move them). Except from that, we're basically covered. We have enough switches to do what we want. There are some old switches that should eventually be replaced, so that we have consistency (which eases up the work if we need to replace a faulty switch – especially since the switches has huge differences regarding the configuration syntaxes). Tools and network-cables we have plenty of, so there is no need for anything of that.

2.4.2 Software

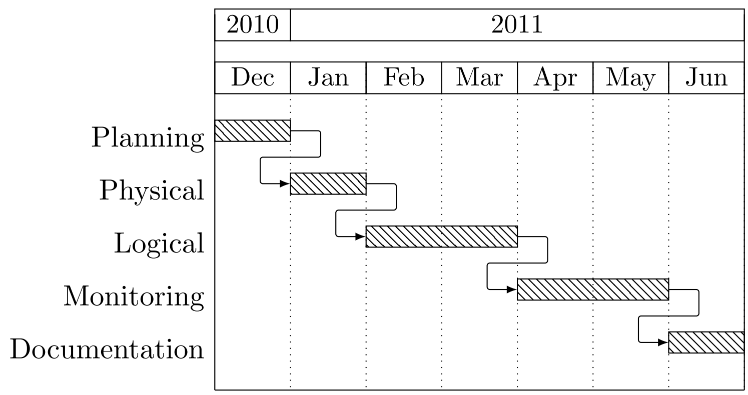
There isn't much software involved – the firmware of the switches will be updated, and we're going to use ISCs BIND9 and DHCPD with flat-file configurations.

3

Progress

3.1 Milestones

The project has several tasks (with a milestone at the end), all of which are to be completed in chronological order. Below is a brief Gantt-diagram showing the 5 planned tasks.



3.1.1 Planning

By the end of december 2010, all the fundamental planning of the network should be done. This includes the plans of how the physical intervention is to be done, and also how the logical setup of the network roughly is going to be. The latter is important to get a better understanding of how to physically place network equipment. Stuff we need for the next milestones should be known by the end of december, so that we can place orders for what we need.

3.1.2 Physical

During the start of january, physical intervention can be started. At this point the logical structure of the network should be ready, and hence also a physical structure. Switches and cables is to be moved during this timeframe. Old hardware (both switches and cables) that isn't needed, should also be cleaned up.

3.1.3 Logical

Once the physical structure is done, we can configure the basic configuration of the switches on a logical level. This means, but is not limited to, VLANs, routing, hostnames, IP-addresses, username/passwords, etc. During this period, central network-services should also be configured and set up – mainly DHCP, DNS and gateway servers. Since the LHC is to be commissioned by middle of march, the production-part of the network should be fully operational by end of february/start of march.

3.1.4 Monitoring

Once the network is more or less fully operational, we can start to implement monitoring tools. This is to get a better overview over the performance of the network, and to discover (and correct) flaws and errors.

3.1.5 Documentation

When monitoring is in place, and all other aspects of the network is done, the documentation-process can start. Documentation has already been written "along-the-road", but a lot of it is just summaries, and not complete documentation. A lot has probably also changed since it was written down, so it needs to be updated.

4

Risks

When doing such a large intervention as this, there are lots of things that can go wrong. Seriously – it's a *lot* that can go wrong. Since we're basically pulling the plug on the network, and then build it from scratch, it can go wrong on both the logical and physical layer of the network, and you can also stumble upon weird software and/or hardware bugs. There are also other factors that can be an issue. I'll try to list most of them, and also give an estimation if they are risks that needs to be taken seriously or not.

Risk	**Probability vs. severity**	**Consequence**	**Comment**
Don't finish in time	10% / High	Can't join technical & physics-runs. Others might have to wait for us to be finished.	This should not happen. Have fallback-solution ready.
Cables break, physical connectivity-issues	40% / Medium	Machines and/or switches loses connectivity.	We have enough cables and patch-panels, so this is not an issue.
Switches/NICs fails	5% / High	Servers cannot communicate with the rest of the network.	We have spare switches and NICs, so this is not an issue.
Unforeseen requirements	30% / Medium	Can't proceed. Progress stops.	Hard to foresee everything in such a complicated setup as this.

Software, incompabilities, bugs	10% / Medium	Can cause unwanted things to happen. Things might not work as expected, and can cause things to take long time to figure out.	Not very much likely to happen, and not much we can prepare for.
Sub-optimal configuration	30% / Medium	Things doesn't work as expected. Bad performance.	Issues can take long time to figure out if one is not 100% familiar with how the given feature is working.
Human error	80% / High	One can forget to do critical things. One can misconfigure things, which can make things take longer time. One can do things that has no effect at that point, but might yield unwanted results at a later time.	Cannot be avoided, as something is bound to be configured wrong in such a complicated setup. It shouldn't pose any risk towards not being able to finish the project, but can give bumps along the road.

5

Literature

Redoing the network requires gaining a lot of knowledge (not only about the network, but the requirements of the cluster, and it's services). There is a lot of unknown things regarding the old setup, and how things work within the cluster. Some of this is documented. Some of it isn't. Some can be read online on wikis, whilst other things need to be "extracted" from previous system-administrators or users of the cluster.

5.1 "Internal" webpages

Some of the information has been gathered from our internal wiki-pages;

- http://wiki.kip.uni-heidelberg.de/ti/HLT/ – Old HLT wiki
- http://espace.cern.ch/alice-hlt/ – New HLT wiki

5.2 External webpages

A lot of information when setting up everything, has been gained by using Google to search for whatever is necessary at any given time.

- http://www.google.com/ – Our best friend (-:
- http://www.hp.com/ – Documentation/manuals/FAQ/KB for the switches

5.3 Old setup

A lot has been learnt by studying the old setup by looking at configuration files and the like. Some of the information in the wiki has also been somewhat outdated, whilst the configuration-files in a running system is as up-to-date as you can get it.

5.4 Books/papers

None so far. (-:

6

Completed

So far 3 of the 5 major tasks are more or less done. We are done planning the network-change. The switches has been moved physically, and all cable-restructuring is done. The switches has been configured, and all critical network-services has been restored. Network-monitoring tools are work-in-progress at the moment, but they will also soon be done.