

Effects of the EPNs and FLPs on the Information
Node for ALICE
Cern

M.Q Puls

2018

Title	Effects of the EPNs and FLPs on the Information Node for ALICE
Name	Mitchell Quinn Puls
Student Number	500659986
Phone	+31615050310
Email	mitchpuls@upcmail.nl mitch.puls@hva.nl
Place	Amsterdam
Date	2018
University	University of Applied Sciences Amsterdam
Department	TI
Mentor	C. J. Rijsenbrij
Company	University of Amsterdam Software for Science Wibautstraat 2-4 Amsterdam 0205995555
Company Supervisor	Dr. Marten Teitsma
Period	Feb. 2018 - Jul 2018

Preface

Contents

1	Summary	4
2	Introduction	5
2.1	ALICE	5
2.2	Load Balancing	5
2.3	Research	5
2.4	Research Model	6
3	Framework	7
3.1	O^2 Balancer	7
3.1.1	Devices	7
3.2	Raspberry Pi	8
3.3	FairMQ	8
3.4	Zookeeper	8
3.5	Fail-over	8
3.6	Blacklist Algorithm	8
4	Question	9
5	Experiments	10
6	Results	11
7	Conclusion	12
7.1	Appendix	12

Chapter 1

Summary

Chapter 2

Introduction

2.1 ALICE

ALICE stands for A Large Ion Collider Experiment, and is a detector mounted on the Large Hadron Collider at CERN. CERN is a European organization for nuclear research, situated in Geneva Switzerland. ALICE's main function is to study matter at extreme energy densities, where matter turns into a form called quark-gluon plasma.

2.2 Load Balancing

The data stream that comes from ALICE is equal to about 1.1 Terabyte per second. All of this data comes in what is known as a heartbeat. This heartbeat gets distributed over 268 First Level Processors and funneled through 1500 Event Processing Nodes. The efficient distribution of this process is what is known as Load Balancing. All of these computers would be monitored by an Information Node.

2.3 Research

This research is a continuation of a previous research done by Heiko van der Heijden. His results show that of the two algorithms tested, Re-initialization and Blacklist, that the Blacklist algorithm has fewer Time Frames lost. Even though the same ratio of FLPs to EPNs that is situated at CERN was used (1/6), there were fewer computers used than at CERN. Because of this it is not sure whether or not the Information Node is able to handle 1700+ computers as compared to the 15 computers used in the experiment. This research is focused around the capability of the Information Node monitoring a higher number of FLPs and EPNs and what the effects are on the results compared to the previous experiment.

2.4 Research Model

Contrary to the previous research, which used a cluster of computers situated at Nikhef Amsterdam, this research will be conducted using a cluster of Raspberry Pi's. The first step is to recreate the previous experiment which was focused around the various ticktimes of Zookeeper. After recreating the first experiment, the same experiment will be conducted with a higher numbers of computers to see what the results are.

All technical documentation of what everything is will be explained in the next section including the definition of the experiments. The next chapter will go more in depth of the prototype made for the experiment and it's difficulties that came with it. The following chapter looks at that the results from the executed experiments. After this follows an analysis of the results of the experiment. Finally a conclusion and recommendations.

Chapter 3

Framework

3.1 O^2 Balancer

O^2 Balancer is a framework of CERN used for simulation experiments for ALICE. The code is open source and licensed under the GNU General Public License V3.0.

3.1.1 Devices

The O^2 Balancer consist of a cluster of 1750 computers, divided in 250 First Level Processors (FLPs) and 1500 Event Nodes (EPNs) These computers are meant to process the data stream coming from ALICE. All of these computers are monitored using an Information Node.

First Level Processors

The FLPs are the first computers in the line. They recieve the data stream (approximately 1.1TB/s) from ALICE and need to distribute that to the next line of computer. In order to do that it takes the data received between two heartbeats, and compresses that into something that's called a Sub Timeframe (STF). A heartbeat lasts for about 20ms. It will then send this STF to the next line of computers which are the EPNs. Every EPN needs to get the same amount of STFs at the same time for recreation purposes.

Event Processing Nodes

The next line of computers are the EPNs. These receive the STFs from the FLPs and then compress them back into a time frame (TF). This compression reduces it's size by a factor of eight. These TFs are then stored for further use.

Information Node

There is one final computer which is the Information Node (IN). This computer keeps track of all the FLPs and EPNs that are online and makes sure that FLPs don't send data to offline EPNs.

3.2 Raspberry Pi

Raspberry Pi is a low cost small computer used for prototyping projects. These projects can reach go from small sensor applications, to bigger host-server applications.

3.3 FairMQ

The transport layer used for the O^2 Balancer is FairMQ. This is a transport layer from the larger framework FairRoot created by GSI Darmstadt. In order to accommodate the smaller processing size of the Raspberry Pi, a trimmed down version of FairRoot is used which is just FairMQ.

3.4 Zookeeper

Zookeeper is a program made by Apache to regulate the whole load balancing process. It is run on the Information Node and from there pings to all EPNs to check whether they are online or not. It then creates a list of online EPNs which it gives to the FLPs so that they know to what EPN to send data to. The frequency of these pings are called the Ticktime.

3.5 Fail-over

When an EPN goes offline it is called a Fail-over. When this happens, Zookeeper will know that it is offline and will notify the FLPs to not send any data to these EPNs anymore.

3.6 Blacklist Algorithm

The experiment used in the previous experiment is the Blacklist Algorithm. This algorithm constantly keeps a list of online channels which is updated by the Information Node using Zookeeper. Once Zookeeper realizes that an EPN is offline, it will update the list so that the algorithm will skip that offline EPN. With this list of online EPNs, the algorithm uses a Round Robin approach to distribute the STF over the EPNs. A way to implement the Blacklist algorithm is shown in listing x

Chapter 4

Question

What effect does the blacklist algorithm have on the Information Node with an higher amount of First Level Processors and Event Processing Nodes for ALICE?

Chapter 5

Experiments

Chapter 6

Results

Chapter 7

Conclusion

Bibliography

DIT IS NOG NIET APA WEET IK
<https://home.cern/about/experiments/alice>
<https://home.cern/about>
<https://fairroot.gsi.de/?q=about>
<https://github.com/FairRootGroup/FairRoot>

7.1 Appendix