

**FLEXIBLE AND EXTENSIBLE
INFRASTRUCTURE MONITORING
ARCHITECTURE FOR COMPUTING GRIDS
WITH INFRASTRUCTURE AWARE JOB
MATCHING**

R.M.K.D Wijethunga

228045C

Dissertation submitted in partial fulfillment of the requirements for the
degree
Degree of Master of Science (Research)

Department of Computer Science and Engineering
University of Moratuwa

University of Moratuwa
Sri Lanka

October 2023

DECLARATION

I declare that this is my own work and this Dissertation does not incorporate without acknowledgement any material previously submitted for a Degree or Diploma in any other University or Institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text. I retain the right to use this content in whole or part in future works (such as articles or books).

Signature: *UOM Verified Signature*

Date: 2023.10.05

The supervisors should certify the Dissertation with the following declaration.

The above candidate has carried out research for the Degree of Master of Science (Research) Dissertation under our supervision. We confirm that the declaration made above by the student is true and correct.

Name of Supervisor: Prof. Indika Perera

Signature of the Supervisor: *UOM Verified Signature*

Date: 05.10.2023

Name of Supervisor: Dr. Gayashan Amarasinghe

Signature of the Supervisor: *UOM Verified Signature*

Date: 05/10/2023

ACKNOWLEDGEMENT

I would like to take this opportunity to express my heartfelt gratitude and extend my sincere appreciation to my esteemed supervisors, Prof. Indika Perera and Dr. Gayashan Amarasinghe. Their unwavering support and guidance throughout the entire duration of this project work have been invaluable. Their insightful suggestions and constructive comments have played a pivotal role in resolving numerous challenges encountered during the project's development.

I am also immensely grateful to Dr. Adeesha Wijayasiri and Dr. A. Shehan Perera, who served as progress review panel members, for their unwavering dedication and valuable input. Their comments and suggestions have significantly contributed to the refinement of my work.

Furthermore, I would like to express my gratitude to Dr. Kutila Gunasekara, our MSc research degree coordinator, for assisting me in navigating the formalities associated with obtaining the degree. The guidance provided by Dr. Gunasekara has been immensely helpful throughout this process.

I would also like to extend my appreciation to all the lecturers at the Faculty of Computer Science and Engineering, University of Moratuwa. Their valuable advice and expertise have played a crucial role in shaping my academic journey.

Lastly, I am indebted to my parents, my brother, and my friends for their unwavering support and continuous motivation throughout my career. Their encouragement has been a constant source of inspiration, and I am truly grateful for their presence in my life.

ABSTRACT

Many research experiments with large data processing requirements rely on massive, distributed Computing Grids for their computational requirements. A Computing Grid is built by combining a large number of individual computing sites distributed globally. These Grid sites are maintained by different institutions across the world and contribute thousands of worker nodes possessing different capabilities and configurations. Developing software for Grid operations that works on all nodes while harnessing the maximum capabilities offered by any given Grid site is challenging without knowing what capabilities each site offers in advance. This research focuses on developing an architecture-independent Grid infrastructure monitoring design to monitor the infrastructure capabilities and configurations of worker nodes at sites across a Computing Grid without the need to contact local site administrators. The design presents a highly flexible and extensible architecture that offers infrastructure metric collection without local agent installations at Grid sites. The resulting design is used to implement a Grid infrastructure monitoring framework called “Site Sonar v2.0” that is currently being used to monitor the infrastructure of 7,000+ worker nodes across 60+ Grid sites in the ALICE Computing Grid. The proposed design is then used to introduce an improved Job matching architecture for Computing Grids that allows job matching based on any infrastructure property of the worker nodes. This dissertation introduces the proposed architecture for a highly flexible and extensible Grid infrastructure monitoring design and an improved job design for Computing Grids and the implementation of those designs to derive important findings about the infrastructure of ALICE Computing Grid while improving its job matching capabilities. This work provides a significant contribution to the development of distributed Computing Grids, particularly in terms of providing a more efficient and effective way to monitor infrastructure and match jobs to worker nodes.

Keywords: Grid computing, Grid monitoring, Grid infrastructure, infrastructure monitoring, Site Sonar, Job Matching, Infrastructure aware

TABLE OF CONTENTS

Declaration of the Candidate & Supervisor	i
Acknowledgement	ii
Abstract	iii
Table of Contents	iv
List of Figures	vii
List of Tables	viii
List of Abbreviations	viii
List of Appendices	x
1 Introduction	1
1.1 Overview	1
1.2 Background	2
1.2.1 Grid Site	2
1.2.2 Computing Grid	2
1.2.3 CERN	2
1.2.4 ALICE experiment	2
1.2.5 ALICE Computing Grid	3
1.2.6 Grid Infrastructure Monitoring	3
1.2.7 Jobs	4
1.2.8 Pilot Jobs	4
1.2.9 Job Matching	4
1.3 Motivation	5
1.4 Problem Statement	6
1.5 Research Objectives	6
1.6 Research Outcomes	6
1.7 Publications	6
1.8 Organization of Thesis	6

2	Literature Review	8
2.1	Grid Infrastructure Monitoring	8
2.1.1	Grid Computing	8
2.1.2	Grid Monitoring	9
2.1.3	Existing Tools	9
2.1.4	Issues with Existing Tools	18
2.2	Job Matching	20
2.2.1	Issues with Existing Tools	21
2.2.2	Existing systems	23
3	Methodology	29
3.1	Data Collection	29
3.1.1	Initial data collection with a Job	29
3.1.2	Site Sonar v1.0 Implementation	30
3.1.3	Site Sonar v1.0 Drawbacks	30
3.1.4	Proposed Solution	32
3.2	Data Storage	34
3.2.1	SQL Storage	34
3.2.2	Post Data Filtering	35
3.2.3	NoSQL Storage	37
3.3	Data Visualization	37
3.3.1	No Code Visualizations	39
3.4	Proposed Monitoring Architecture	40
3.4.1	Data Collection Framework	40
3.4.2	Data Analysis Framework	43
3.5	Infrastructure Metrics Integration	44
3.5.1	Job Matching	45
3.5.2	Unlimited Infrastructure Constraints	46
3.6	Proposed Job Matching Architecture	48
4	Implementation	51
4.1	Site Sonar Architecture	51
4.1.1	Probe	51

4.1.2	Sonar	53
4.1.3	Central Services	55
4.2	Improved Job Broker	58
4.2.1	Site Sonar ELK Stack	60
4.2.2	Summary	64
5	Results	65
5.1	Analysis	65
5.1.1	Operating System Distribution	65
5.1.2	Singularity Support	67
5.1.3	Grid Overview	68
5.2	Findings	68
5.2.1	Sites running CentOS 6	69
5.2.2	Reusing hostnames on different nodes	70
5.2.3	Use of Site Sonar as a Grid debugging tool	71
5.3	Evaluation	71
5.3.1	Quantitative Evaluation	71
5.3.2	Qualitative Evaluation	78
5.3.3	Associated projects	80
6	Conclusion	82
6.1	Contribution	82
6.2	Limitations and Future Work	83
	References	84
	Appendix A Full data set collected from a worker node	89
	Appendix B Site Sonar v2.0 Component Template	103

LIST OF FIGURES

Figure	Description	Page
Figure 1.1	ALICE Grid in numbers	3
Figure 2.1	Foster's model of Grid Computing	8
Figure 2.2	Taxonomy of Grid monitoring systems	10
Figure 2.3	High level architecture of MonALISA	13
Figure 2.4	MONIT Architecture	15
Figure 2.5	Site Sonar v1.0 architecture	17
Figure 2.6	Site Sonar v1.0 Interface	18
Figure 2.7	PanDA WMS Overview	24
Figure 2.8	Components of JAliEn WMS	27
Figure 3.1	Query Analysis of Site Sonar v1.0	32
Figure 3.2	No. of nodes with given CPU Core count	36
Figure 3.3	CPU model distribution in the ALICE Grid	36
Figure 3.4	GridICE Interface	38
Figure 3.5	MONIT Interface	38
Figure 3.6	Proposed Architecture for Grid Infrastructure Monitoring Tool	40
Figure 3.7	Data flow from Job pilot to Elasticsearch cluster	44
Figure 3.8	Proposed Job Matching Architecture	49
Figure 3.9	Proposed Architecture Flow Diagram	50
Figure 4.1	Site Sonar v2.0 Architecture	52
Figure 4.2	List of Site Sonar probes	53
Figure 4.3	Site Sonar Integration with JAliEn	54
Figure 4.4	JAliEn Job Broker Constraint Matching Logic	59
Figure 4.5	Updated Job Broker Constraint Matching Logic	60
Figure 4.6	No. of documents injected per day	63
Figure 5.1	Operating System Monitoring Dashboard	66
Figure 5.2	Singularity Support Dashboard of ALICE Grid on 2023-04-21	68
Figure 5.3	Sites not supporting Singularity in ALICE Grid on 2023-04-21	69
Figure 5.4	ALICE Grid Overview on 2023-04-21	70
Figure 5.5	Time taken for Data Collection in Site Sonar	73
Figure 5.6	Core-hours wastage in Site Sonar v1.0 vs v2.0	75
Figure 5.7	Data Retrieval Time in Site Sonar v1.0 vs v2.0	77

LIST OF TABLES

Table	Description	Page
Table 2.1	Issues with existing Grid monitoring tools	20
Table 3.1	Constraints supported by different WMSs	47
Table 5.1	Operating system distribution of ALICE Grid as of 2022-08-30	67
Table 5.2	Operating system distribution of ALICE Grid as of 2023-04-22	67
Table 5.3	Time taken for Data Collection in Site Sonar	72
Table 5.4	No. of core hours spent for Data Collection in Site Sonar	74
Table 5.5	Simplified queries used for the evaluation	76
Table 5.6	Deployment time for adding a new job matching parameter in the existing design	77
Table 5.7	Deployment time for adding a new job matching parameter in the proposed design	78

LIST OF ABBREVIATIONS

Abbreviation	Description
ALICE	A Large Ion Collider Experiment
API	Application Programming Interface
ATLAS	A Toroidal LHC Apparatus
CE	Computing Element
CGroups	Control Groups
CMS	The Compact Muon Solenoid
CVMFS	CernVM File System
GUI	Graphical User Interface
I/O	Input/Output
JAliEn	Java ALICE Environment
JDL	Job Description Language
JSP	Java Server Pages
LHC	Large Hadron Collider
MDS	Globus Monitoring and Discovery Service
OS	Operating System
PanDA	Production and Distributed Analysis
TTL	Time To Live

LIST OF APPENDICES

Appendix	Description	Page
Appendix -A	Full data set collected from a worker node	89
Appendix -B	Site Sonar v2.0 Component Template	103