NEIGHBOUR R NEIGHBOUR REPLICA AFFIRMATIVE ADAPTIVE FAILURE DETECTION AND AUTONOMOUS RECOVERY

AHMAD SHUKRI BIN MOHD NOOR

A thesis submitted in
fulfillment of the requirements for the award of the
Doctor of Philosophy.

Faculty of Computer Science and Information Technology
Universiti Tun Hussein Onn Malaysia

NOVEMBER 2012

All Praise and thanks be to Allah.

We praise and thank Him, ask Him for His Help and Forgiveness.This dissertation is a humble effort to augment the body of knowledge in the field of computer science particularly distributed system. I could never complete a PhD dissertation unless helped by Allah. I would like to thank Allah almighty for helping me in every aspect of this dissertation.

ACKNOWLEDGEMENTS

There are many people deserving acknowledgment for their roles in helping me through this work. Foremost I would like to thank my supervisor Professor Dr. Mustafa Bin Mat Deris and Faculty of Computer Science and Information Technology, UTHM for all support given through out the duration of this research.

Special gratitude to my parents; Mohd Noor Ibrahim and Hasmah Ali, my wife Alia; Mastura Alias, my son; Arif Mustaqim Ahmad Shukri, my mother in-law; Azizah Endut and my father in-law; Alias Sembok, for everything that they did for me during this time. Their encouragement, invaluable spiritual help, prayers, support and help was always there for me whenever I needed them.

Many thanks to my postgraduate and working colleagues Rabie Mamat, Ahmad Nazari Mohd Rose, Isa Awang, Zailani Abdullah, Mohamad Nor Hassan, Dr. Masila Abdul Jalil and Sapiee bin Jamel for their help and all the existential discussions, which happened to be the most useful ones.

To each of the above, I extend my deepest appreciation.

ABSTRACT

High availability is an important property for current distributed systems. The trends of current distributed systems such as grid computing and cloud computing are the delivery of computing as a service rather than a product. Thus, current distributed systems rely more on the highly available systems. The potential to fail-stop failure in distributed computing systems is a significant disruptive factor for high availability distributed system. Hence, a new failure detection approach in a distributed system called Affirmative Adaptive Failure Detection (AAFD) is introduced. AAFD utilises heartbeat for node monitoring. Subsequently, Neighbour Replica Failure Recovery(NRFR) is proposed for autonomous recovery in distributed systems. AAFD can be classified as an adaptive failure detector, since it can adapt to the unpredictable network conditions and CPU loads. NRFR utilises the advantages of the neighbour replica distributed technique (NRDT) and combines with weighted priority selection in order to achieve high availability, since automatic failure recovery through continuous monitoring approach is essential in current high availability distributed system. The environment is continuously monitored by AAFD while auto-reconfiguring environment for automating failure recovery is managed by NRFR. The NRFR and AAFD are evaluated through virtualisation implementation. The results showed that the AAFD is 30% better than other detection techniques. While for recovery performance, the NRFR outperformed the others only with an exception to recovery in two distributed technique (TRDT). Subsequently, a realistic logical structure is modelled in complex and interdependent distributed environment for NRDT and TRDT. The model prediction showed that NRDT availability is 38.8% better than TRDT. Thus, the model proved that NRDT is the ideal replication environment for practical failure recovery in complex distributed systems. Hence, with the ability to minimise the Mean Time To Repair (MTTR) significantly and maximise Mean Time Between Failure (MTBF), this research has accomplished the goal to provide high availability self sustainable distributed system.

ABSTRAK

Kebolehsediaan yang tinggi ialah satu ciri penting untuk sistem teragih semasa. Kecenderungan sistem-sistem teragih masakini seperti *grid computing* dan *cloud computing* ialah penyedian pengkomputeran sebagai satu perkhidmatan berbanding sebagai satu produk. Oleh itu, sistem teragih semasa sangat memerlukan sistem yang mempunyai kebolehsediaan yang tinggi. Potensi untuk gagal-berhenti dalam sistem pengkomputeran teragih adalah faktor yang memyebabkan gangguan kepada kebolehsediaan yang tinggi. Oleh itu, tesis ini mencadangkan pengesanan kegagalan yang afirmatif serta adaptif (AADF). AAFD menggunakan *heartbeat* untuk pemantauan nod. Seterusnya pemulihan kegagalan replika kejiranan (NRFR) dicadangkan untuk pemulihan secara autonomi. Oleh kerana AAFD dapat mengadaptasi dengan ketidaktentuan rangkaian dan CPU, ia boleh diklasifikasikan sebagai pengesan kegagalan yang adaptif. NRFR menggunakan kelebihan teknik replika kejiranan teragih (NRDT) dan menggabungkan pemilihan keutamaan berdasarkan pemberat. Seterusnya AAFD dan NRFR dinilai melalui pelaksanaan *virtualisation*. Hasil keputusan menunjukkan, secara puratanya AAFD adalah 30% lebih baik dari teknik-teknik yang lain. Manakala bagi prestasi pemulihan, NRFR mengatasi yang lain kecuali untuk pemulihan didalam teknik replika berdua (TRDT). Seterusnya, struktur logik yang realistik dan praktikal bagi kebolehsediaan tinggi dalam persekitaran teragih yang komplek dan saling bergantungan dimodelkan untuk NRDT dan TRDT. Model ini membuktikan bahawa kebolehsediaan NRDT adalah 38.8% lebih baik. Oleh yang demikian, model ini membuktikan NRDT adalah pilihan terbaik untuk memulihkan kegagalan di dalam sistem teragih yang komplek. Oleh itu, dengan kebolehan meminimumkan Mean Time To Repair (MTTR) dan memaksimumkan Mean Time Between Failure (MTBF), kajian ini mencapai matlamat untuk menyediakan sistem teragih yang mampan dan kebolehsediaan tinggi.

PUBLICATIONS

1) Ahmad Shukri Mohd Noor , Mustafa Mat Deris and Tutut Herawan Neighbour-Replica Distribution Technique Availability Prediction in Distributed Interdependent Environment. International Journal of Cloud Applications and Computing (IJCAC) 2(3), 98-109, IGI Global , 2012
2) Ahmad Shukri Mohd Noor, Mustafa Mat Deris, Tutut Herawan and Mohamad Nor Hassan. On Affirmative Adaptive Failure Detection. LNCS 7440 pp. 120-129 Springer-Verlag Berlin Heidelberg 2012.
3) Ahmad Shukri Mohd Noor and Mustafa Mat Deris. Fail-stop-proof fault tolerant model in distributed neighbor replica architecture. Procedia-Computer Science. Elsevier Ltd 2011. (Accepted to be published).
4) Ahmad Shukri Mohd Noor and Mustafa Mat Deris. Deris Failure Recovery Mechanism in Neighbor Replica Distribution Architecture. LNCS 6377, pp. 41–48, 2010. Springer-Verlag Berlin Heidelberg 2010.
5) Ahmad Shukri Mohd Noor and Mustafa Mat Deris. Extended Heartbeat Mechanism for Fault Detection Service Methodology CCIS 63, pp. 88–95, 2009. Springer-Verlag Berlin Heidelberg 2009.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

AAFD	-	Affirmative Adaptive Failure Detection
CH	-	Cluster Head
CPU	-	Centre Processing Unit
DC	-	Data Collector
FTP	-	File Transfer Protocol
GHM	-	Globus Heartbeat Monitor
HB	-	Heartbeat
HBM	-	Heartbeat Monitor
IS	-	Index Server
MTBF	-	Mean Time Between Failures
MTTF	-	Mean Time To Failure
MTTR	-	Mean Time To Repair
NRDT	-	Neighbour Replication Distributed Technique
OS	-	Operating Systems
ROWA	-	Read-One Write-All
SLAs	-	Service Level Agreements
SPOF	-	Single Point Of Failure
SSH	-	Secure shell protocol
TRDT	-	Two-Replica Distributed Technique
TQ	-	Tree Quorum
VT	-	Voting

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

In this chapter, the background of the research is outlined, followed by problem statements, objectives, contributions, scope of the research and lastly, the organization of the thesis.

1.1 Research background

Availability is one of the most important issues in distributed systems (Renesse & Guerraoui, 2010; Deris *et al*., 2008; Bora, 2006). With greater numbers of computers working together, the possibility that a single computer failure can significantly disrupt the system is decreased (Dabrowski, 2009). One of the benefits of a distributed system is the increase of parallelism for replication (Renesse & Guerraoui, 2010). Replication is a fundamental technique to achieve high availability in distributed and dynamic environments by masking errors in the replicated component (Noor & Deris, 2010; Bora, 2006). Thus, replication is very important in providing high availability and efficient distributed system. Distributed systems can therefore lend themselves in providing high availability (Mamat *et al*., 2006).

A fail-stop system is one that does not produce any data once it has failed. It immediately stops sending any events or messages and does not respond to any messages(Arshad,2006). This type of failures is common in today's large computing systems. When a fail-stop failure occurs, a prompt and accurate failure detection with minimum time to recover are critical factors in providing high availability in distributed systems. If these factors can efficiently and effectively be handled by a failure detection and recovery technique, it can provide a theoretical and practical high availability solution for a distributed system.

Since current distributed computing such as grid computing and cloud computing become larger, increasingly dynamic and heterogeneous. These distributed systems become more and more complicated. Failures or errors are arising due to the inherently unreliable nature of the distributed environment include hardware failures, software errors and other sources of failures. Many failure detection and recovery techniques have been adopted to improve the distributed system availability. In addition to the outstanding replication technique for high availability, failure detection and recovery is an important design consideration for providing high availability in distributed systems (Dabrowski, 2009; Stelling *et al*., 1998; Abawajy, 2004b; Flavio, 2006).

Therefore, failure detection and recovery in distributed computing has become an active research area (Dimitrova & Finkbeiner, 2009; Siva & Babu 2010; Khan, Qureshi & Nazir, 2010; Montes, Sánchez & Pérez, 2010; Costan *et al*., 2010). Research in failure detection and recovery distributed computing aims at making distributed systems high availability by handling faults in complex computing environments. In order to achieve high availability, an autonomous failure detection and recovery service need to be adopted. An autonomous failure detection and recovery service is able to detect errors and recover the system without the participation of any external agents, such as human. It can be restored, or has the ability of self-healing, then back to the correct state again (Arshad, 2006). If no failure detection and recovery is provided, the system cannot survive to continue when one or several processes fail, and the whole program crashes.

Failure detection (or fault detection) is the first essential phase for developing any fault tolerance mechanism or failure recovery (Avizienis *et al*., 2004). Failure detections provide information on faults of the components of these systems (Stalin *et al*., 1998).

Failure recovery is the second phase in developing any recovery mechanism (Avizienis *et al*., 2004). Replication is one of the core techniques that can be utilised for failure recovery in distributed and dynamic environments (Bora, 2006). Exploitation of component redundancy is the basis for recovery in distributed systems. A distributed system is a set of cooperating objects, where an object could be a virtual node, a process, a variable, an object as in object-oriented programming, or even an agent in multi-agent systems. When an object is replicated, the application has several identical copies of the object also known as replicas (Helal, Heddaya & Bhargava , 1996; Deris *et al*., 2008). When a failure occurs on a replica, the failure is masked by its other replicas, therefore availability is ensured in spite of the failure. Replication mechanisms have been successfully applied in distributed applications. However, the type of replication mechanisms to be used in the application is decided by the programmer before the application starts. As a result, it can only be applied statically. Thus, the development of autonomous failure detection and recovery model with suitable replication technique and architectural design strategy is very significant in building high availability distributed systems.

1.2 Problem statements

A study has found fault-detection latencies covered from 55% to 80% of non-functional periods (Dabrowski *et al*., 2003). This depends on system architecture and assumptions about fault characteristics of components. These non-functional periods happened when a system is uninformed of a failure (or failure detection latency) and periods when a system attempts to recover from a failure (failure-recovery latency) (Mills *et al*., 2004). Even though the development of fault detection mechanism in large scale distributed system is subject to active research, it still suffers from some weaknesses (Dabrowski, 2009; Pasin, Fontaine & Bouchenak, 2008; Flavio, 2006).

i) Failure detection trade-offs between accuracy and completeness. Current failure detection approaches suffer from the weaknesses of either fast detection with low accuracy or completeness in detecting failures with a lengthy timeout. Inaccurate detection may result in the recovery malfunction while delays in detecting a failure will subsequently delay the recovery action. These trade-offs need to be improved.

ii) Choosing the right replication architectural design strategies are very crucial in providing high availability and efficient distributed system. This is because keeping all of the replicas requires extra communication as well as processing and may delay the recovery process. This will cause the system to be down for a considerable period of time. In contrast, insufficient replicas can jeopardise the availability of the distributed system.

iii) Although the idea and theory of replication is convincing and robust, practical implementation of replication technique is difficult to be modelled in real distributed environment (Christensen, 2006). This is due to the complexity in the implementation of replication and check pointing techniques. Therefore they have been studied more theoretically through the use of simulation technique (Khan, Qureshi & Nazir, 2010). Thus, most of them only discussed the simulation of the theories rather than its implementation.

iv) Many existing failure recovery techniques have a considerable period of downtime associated with them. This downtime can cause a significant business impact in terms of opportunity loss, administrative loss and loss of ongoing business. There is a need not just to reduce the downtime in the failure recovery process but also to automate it to a significant degree in order to avoid errors that are caused by manual failure recovery techniques.

1.3 Objectives

The main objectives of this dissertation can be summarized as follows:

i) To propose new approaches for failure detection and an autonomous failure recovery in distributed system by introducing;

- A new framework for continuous failure detection,
- A new framework for automated failure recovery

ii) To implement failure detection and autonomous failure recovery based on the proposed approach.

iii) To compare and analyse the performance of the proposed method with existing approaches.

1.4 Scope

The focus of this research is to continuously monitor the failure detection and to automate the failure recovery in an unpredictable network within Neighbour Replica Distributed environment with the assumption that failure model is fail-stop failure.

1.5 Contributions

There are four major contributions in this thesis;

i) Introduced new continuous failure detection approach. The approaches have improved the detection accuracy and completeness as well as reducing detection time.

ii) Proposed an autonomous failure recovery approach in a neighbour replica distributed system that can reduce computation time for failure recovery. The failure recovery approach also has the capability to determine and select the neighbour with the best optimal resources which can optimise the system availability.

iii) The implementation of continuous failure detection and autonomous failure recovery frameworks using Linux Shell script and tools in the neighbour replica distributed system. The implementation results showed that affirmative adaptive failure detection (AAFD) is able to achieve a complete and accurate detection with prompt timing while neighbour replica failure recovery NRFR can minimise the recovery time. Hence, by reducing failure detection latency and recovery processing time, the proposed approaches are able to reduce the Mean Time To Repair (MTTR) significantly as well as maximise the system availability or Mean Time Between Failure (MTBF). In addition, the implementation demonstrated that the proposed failure detection and recovery is theoretically sound as well as practically feasible in providing high availability distributed system.

iv) Modelled a realistic and practical logical structure for high availability in complex and interdependent distributed environment. This model provided availability predictions for neighbour replica distribution technique (NRDT) and two replica distribution technique (TRDT).

1.6 Thesis organisation.

The work presented in this dissertation is organized into six chapters. The rest of this document is organized as follows. Chapter two describes preliminary concepts and related works that are selected from related research. Chapter three proposed a methodology for failure detection and failure recover in neighbour replica distributed architecture. This chapter discusses in detail the proposed methodology. The implementation of proposed failure detection and recovery is presented in Chapter four. Chapter five presents the results and analysis of the proposed approach implementation and provide in-depth discussion of the implementation results. Lastly, Chapter six describes the conclusions and possible future work in relation to this dissertation.

CHAPTER 2

LITERATURE REVIEW

This chapter describes related background knowledge and reviews existing literature on failure detection and recovery. The background knowledge would provide the information on failure detection metrics, the behaviour of failed systems and interaction policies. Furthermore, this chapter also discusses and reviews existing related researches on failure recovery in distributed system which includes, check-pointing and replication techniques. Since one of the objectives of this thesis is to automate failure recovery, this chapter will provide detailed review of replication techniques that best suited the high availability distributed system with self recovery characteristics. This includes the costs of resources and communication for replication as well as architectural complexity which will affect the recovery time. It also highlights the advantages and disadvantages of recent work that have been done in these fields.

2.1 Introduction

Schmidt (2006) defined availability as the frequency or duration in which a service or a system component is available for use. If this component is needed to provide the service, outage of a component is also applicable for service availability. In addition, any features that could help the system to stay operational despite the occurrences of failures will also be considered as availability.

REFERENCES

Abawajy J. (2004a). Fault-tolerant scheduling policy for Grid computing systems. *Proceedings of the 18th International Parallel and Distributed Processing Symposium*, April 2004. Los Alamitos, CA: IEEE Computer Society Press. pp. 238–244.

Abawajy, J. (2004b). Fault detection service architecture for Grid computing systems. *LNCS*, *3044.* Berlin, Heidelberg: Springer-Verlag. pp. 107–115.

Agrawal, D. and El Abbadi, A. (1992). The generalized tree quorum protocol: an efficient approach for managing replicated data. *ACM Trans. Database System.* pp. 689-717.

Ahmad, N. (2007). *Managing replication and transactions using neighbour replication on data grid Database design.* Ph.D. Thesis. Universiti Malaysia Terengganu.

Amazon.com Inc. (2010). Amazon Simple Storage Service (Amazon S3). Retrieved on September 12, 2010 from http://aws.amazon.com/s3

Andrieux, A., Czajkowski, K., Dan, A., Keakey, K., Ludwig, H., Nakata, T., Pruyne, J., Rofrano, J., Tuecke, S. & Xu, M. (2007). *Web services agreement specification (WS-Agreement). GFD.107*, Open Grid Forum.

Andrzejak, A., Graupner, S., Kotov, V. & Trinks, H. (2002). Algorithms for self-organization and adaptive service placement in dynamic distributed systems. *HPL-2002-259*, Hewlett Packard Corporation.

Arshad N. (2006). *A Planning-Based Approach to Failure Recovery in Distributed Systems*. Ph.D. Thesis. University of Colorado.

Avizienis, A., Laprie, J., Randell, B., & Landwehr, C. (2004). Basic concepts and taxonomy of dependable and secure computing. *IEEE Transactions on Dependable and Secure Computing, 1(1),* pp. 11–33.

Bertier, M. & Marin, P. (2002). Implementation and performance evaluation of an adaptable failure detector. *Proceedings of the Intl. Conf. on Dependable Systems and Networks.* pp. 354 – 363.

Bora S. (2006). A Fault Tolerant System Using Collaborative Agents. *LNAI, 3949*. Berlin, Heidelberg: Springer-Verlag. pp. 211– 218.

Boteanu A, Dobre C, Pop F & Cristea, V (2010). Simulator for fault tolerance in large scale distributed systems. *Proceedings of the 2010 IEEE 6th International Conference on Intelligent Computer Communication and Processing*. pp. 443-450.

Budati K, Sonnek J, Chandra A & Weissman J. (2007). RIDGE: Combining reliability and performance in open Grid platforms. *Proceedings of the 16th International Symposium on High Performance Distributed Computing (*ISHPDC2007). New York, USA: ACM Press. pp. 55–64.

Chen, W., Toueg, S. & Aguilera, M.K. (2000). On the quality of service of failure detectors. *Proceedings of the International Conference on Dependable Systems and Networks New York.* IEEE Computer Society Press.

Chen, W. Toueg, S. & Aguilera, M.K. (2002). On the QoS of failureDetectors. *IEEE Trans. Computers*, *51(5),* pp. 561–580.

Chervenak, A., Vellanki, V.& Kurmas, Z. (1998) Protecting file systems: A survey of backup techniques. *In: Proc. of Joint NASA and IEEE Mass Storage Conference.* Los Alamitos: IEEE Computer Society Press.

Christensen N. H. (2006). A formal analysis of recovery in a preservational data grid. *Proceedings of the 23rd IEEE Conference on Mass Storage Systems and Technologies*, College Park, Maryland USA.

Cooper M. (2008). Advanced Bash-Scripting Guide, An in-depth exploration of the art of shell, Retrieved on October 9, 2008 from http://theriver.com.

Costan, A. Dobre, C. Pop, F. Leordeanu, C. & Cristea, V. (2010). A fault tolerance approach for distributed systems using monitoring based replication. *Proceedings of the 2010 IEEE 6th International Conference on Intelligent Computer Communication and Processing*. pp. 451-458.

Dabrowski, C., Mills, K. & Rukhin, A. (2003). A Performance of Service-Discovery Architectures in Response to NodeFailures, *Proceedings of the 2003 International Conferenceon Software Engineering Research and Practice (SERP'03)*: CSREA Press. pp. 95-10.

Dabrowski, C. (2009). Reliability in grid computing, *Concurrency Computation: Practice and Experience.* Wiley InterScience.

Deris, M.M. (2001). *Efficient Access of Replicated Data in Distributed Database Systems*. Ph.D. Thesis, Universiti Putra Malaysia.

Deris, M.M. , Abawajy, J. H. & Mamat, A. (2008). An efficient replicated data access approach for large-scale distributed systems. *Future Generation Comp. Syst. 24(1)*, pp. 1-9.

Deris M.M., Ahmad N, Saman M. Y., Ali N. & Yuan Y. (2004). High System Availability Using Neighbor Replication on Grid. *IEICE Transactions 87-D (7)*, pp. 1813-1819.

Dimitrova, R. & Finkbeiner, B. (2009). Synthesis of Fault-Tolerant Distributed Systems *LNCS, 5799*. Berlin Heidelberg: Springer-Verlag. pp. 321–336.

Elhadef, M & Boukerche, A. (2007). A Gossip-Style Crash Faults Detection Protocol for Wireless Ad-Hoc and Mesh Networks. *Proceedings of Int. Conf. IPCCC.* pp. 600-602.

European Power Supply Manufacturers Association. (2005). Guidelines to Understanding Reliability Prediction. Wellingborough, Northants, U.K.

Felber, P D´efago, X. Guerraoui, R. & Oser, P. (1999). Failure detectors as first class objects. *Proceedings of the 9th IEEE Int'lSymp. on Distributed Objects and Applications.* pp. 132–141.

Fetzer C. Raynal M. & Tronel F. (2001). An adaptive failure detection protocol. *Proceedings of the 8th IEEE Pacific Rim Symp. on Dependable Computing.* pp. 146 - 153.

Figgins, S. Siever, E. & Weber, A. (2003). Linux in a Nutshell, 4th Edition O'Reilly, USA.

Flavio, J. (2006). *Coping with dependent failures in distributed systems*. Ph.D. Thesis, University of California, San Diego.

Genaud, S. & Rattanapoka, C. (2007). P2P-MPI: A peer-to-peer framework for robust execution of message passing parallel programs on Grids. *Journal of Grid Computing, 5(1)*, pp. 27–42.

Gillen, M. Rohloff, K. Manghwani, P. & Schantz, R (2007). Scalable, Adaptive, Time-Bounded Node Failure Detection. *Proceedings of the 10th IEEE High Assurance Systems Engineering Symposium (HASE '07)*. DC, USA.

Goodale. T, Allen, G., Lanfermann, G., Masso, J., Radke, T., Seidel, E., Shalf, J. (2003). The cactus framework and toolkit: Design and applications. *LNCS, 2565*. Berlin Heidelberg: Springer. pp. 15–36.

Hayashibara, N. Defago, X. Yared R.& Katayama T. (2004).The f accrual failure detector. *In 23rd IEEE International Symposium on Reliable Distributed Systems(SRDS'04)*: IEEE Computer Society. pp. 66–78.

Hayashibara, N. & Takizawa, M. (2006).Design of a notification system for the φ accrual failure detector. *Proceedings of the 20th International Conference on Advanced Information Networking and Applications - Volume 1 (AINA'06).* pp. 87-97.

Helal, A., Heddaya, A. & Bhargava, B. (1996). Replication Techniques in Distributed Systems: Kluwer Academic Publishers.

Hwang, S. & Kesselman, C. (2003). Introduction Requirement for Fault Tolerance in the Grid, Related Work. A Flexible Framework for Fault Tolerance in the Grid. *Journal of Grid Computing 1,* pp. 251-272.

ITEM Software, Inc.(2007) , Reliability Prediction Basics, Hampshire ,U.K.

Jia, W. & Zhou, W. (2005), Distributed Network Systems: From Concepts to Implementations. Springer Science and Business Media.

Khan, F.G., Qureshi, K. & Nazir, B. (2010). Performance evaluation of fault tolerance techniques in grid computing system. *Computers & Electrical Engineering Volume 36, Issue 6*, Elsevier B.V. pp. 1110-1122.

Khilar, P. Singh, J.& Mahapatra, S. (2008). Design and Evaluation of a Failure Detection Algorithm for Large Scale Ad Hoc Networks Using Cluster Based Approach. *International Conference on Information Technology 2008*, IEEE.

Koren, I. & Krishna C. M. (2007), Fault-Tolerant Systems. San Francisco, CA: Morgan-Kaufman Publishers.

Parziale, L., Dias, A., Filho, L.T., Smith, D., VanStee, J. & Ver, M. (2009). *Achieving High Availability,on Linux for System z with Linux-HA Release 2.* International Business Machines Corporation (IBM).

Lac C & Ramanathan S.(2006). A resilient telco Grid middleware. *Proceedings of the 11th IEEE Symposium on Computers and Communications.* Los Alamitos, CA: IEEE Computer Society Press. pp. 306–311.

Lanfermann, G, Allen G, Radke T, Seidel E. Nomadic. (2002). Fault tolerance in a disruptive Grid environment. *Proceedings of the 2nd IEEE/ACM International Symposium Cluster Computing and the Grid.* Los Alamitos, CA: IEEE Computer Society Press. pp. 280–282.

Li, M. (2006) Fault Tolerant Cluster Management, Ph.D. Thesis in Computer Science University of California, Los Angeles.

Limaye, K., Leangsuksum, B., Greenwood, Z., Scott, S., Engelmann, C., Libby, R. & Chanchio, K. (2005). Job-site level fault tolerance for cluster and Grid environments. *Proceedings of the IEEE International Conference on Cluster Computing.* Los Alamitos, CA: IEEE Computer Society Press. pp. 1–9.

Love, R. (2007). Linux System Programming, O'Reilly Media, United States of America.

Luckow, A. & Schnor, B. (2008). Migol: A Fault-Tolerant Service Framework for MPI Applications in the Grid. *Future Generation Computer Systems –The International Journal of Grid Computing: Theory, Methods and application, 24(2)*, pp. 142–152.

Ma, T. (2007). Quality of Service of Crash-Recovery Failure Detectors, Ph.D. Thesis, Laboratory for Foundations of Computer Science School of Informatics University of Edinburgh.

Mamat, A, Deris, M. M. Abawajy, J.H. & Ismail, S. (2006). Managing Data Using Neighbor Replication on Triangular-Grid Structure. *LNCS, 3994.* Berlin Heidelberg: Springer-Verlag. pp. 1071 – 1077.

Mamat, R., Deris, M.M. & Jalil, M. (2004). Neighbor Replica Distribution Technique for cluster server systems. *Malaysian Journal of Computer Science, 17(.2)*, pp. 11-20.

Mills, K., Rose S., Quirolgico, S., Britton, M. & Tan, C. (2004). An autonomic failure detection algorithm. *SIGSOFT Softw. Eng. Notes, 29(1)*, pp. 79–83.

Montes, J., Sánchez, A. & Pérez, M.S. (2010) "Improving Grid Fault Tolerance by Means of Global Behavior Modeling," Ninth *Parallel and Distributed Computing, International Symposium on.* pp. 101-108.

Natrajan, A., Humphrey, M. & Grimshaw, A. (2001). Capacity and capability computing in legion. *Proceedings of the International Conference on Computational Sciences, Part I.* Berlin Heidelberg: Springer-Verlag. pp. 273–283.

Noor, A.S.M. & Deris, M.M. (2010). Failure Recovery Mechanism in Neighbor Replica Distribution Architecture. *LNCS, 6377*. Berlin Heidelberg: Springer-Verlag. pp. 41–48.

Ozsu, M.T. & Valduriez, P.(1999). *Principles of Distributed Database Systems,*2nd Ed., Prentice Hall,

Pasin, M., Fontaine, S. & Bouchenak S. (2008). Failure Detection in Large-Scale Distributed Systems: A Survey. *In 6th IEEE Workshop on End-to-End Monitoring Techniques and Services (E2EMon 2008),* Brazil.

Platform Computing Corporation. (2007). Administering Platform Process Manager, version 3.1, USA.

Pre, M.D. (2008). *Analysis and design of Fault-Tolerant drives*. Ph.D. Thesis, University of Padova.

Renesse R. V. & Guerraoui, R. (2010). Replication Techniques for Availability. *LNCS, 5959.* Heidelberg Berlin: Springer-Verlag. pp. 19-40.

Renesse, R. Minsky, Y. & Hayden, M. (1998). A Gossip-Style Failure Detection Service, Technical Report, TR98-1687.

Marechal, S. (2009). VMware Unveils VMware Tools as Open Source Software. Retrieved on 2009-07-01 from http://lxer.com/module/newswire/view/92570/index.html

Satzger, B. Pietzowski, A. Trumler, W.& Ungerer,T. (2007). A new adaptive accrual failure detector for dependable distributed systems. *SAC '07: ACM Symposium on Applied Computing.* New York, USA: ACM Press.

Satzger, B., Pietzowski, A., Trumler, W & Ungerer, T. (2008). A Lazy Monitoring Approach for Heartbeat-Style Failure Detectors, *Proceedings of the 3rd International Conference on Availability, Reliability and Security*. pp. 404-409.

Shen, H. H., Chen, S. M., Zheng, W. M. & Shi, S. M. (2001). A Communication Model for Data Availability on Server Clusters. *Proceedings of the Int'l. Symposium on Distributed Computing and Application.* Wuhan. pp. 169-171.

Siva, S.S. & Babu, K.S. (2010). Survey of fault tolerant techniques for grid. *Computer Science Review, 4(2)*: *Elsevier Inc.* pp. 101-120.

Siva, S.S., Kuppuswami, K. & Babu, S. (2007). Fault tolerance by check-pointing mechanisms in grid computing. *Proceedings of the International Conference on Global Software Development,* Coimbatore.

Schmidt, K. (2006). High Availability and Disaster Recovery: Concepts, Design, Implementation. Springer-Verlag.

So K. C.W. & Sirer, E.G. (2007). Latency and Bandwidth-Minimizing Failure Detector. Proceedings of the EuroSys.

Srinivasa, K. G. Siddesh, G. M. & Cherian, S. (2010).Fault-Tolerant middleware for Grid Computing . *Proceedings in the 12th IEEE International Conference on High Performance Computing and Communications.* Melbourne, Australia: pp. 635-640.

Stelling, P. Foster, I. Kesselman, C. Lee & C. Laszewski, G. (1998) A Fault Detection Service for Wide Area Distributed Computations. *Proceedings of the HPDC*. pp. 268-278.

Townend, P., Groth, P., Looker, N. & Xu, J. (2005). FT-grid: A fault-tolerance system for e-science. *Proceedings of the Fourth UK e-Science All Hands Meeting*. Engineering and Physical Sciences Research Council: Swindon, U.K.

Valcarenghi, L. & Piero, C. (2005). QoS-aware connection resilience for network-aware Grid computing fault tolerance. *Proceedings of the 7th International Conference on Transparent Optical Networks*, July 2005. Los Alamitos, CA: IEEE Computer Society Press. pp. 417–422.

Wang, Y. Li, Z. & Lin, W. (2007). A Fast Disaster Recovery Mechanism for Volume Replication Systems. *HPCC, LNCS, 4782,* pp. 732–743, 2007.

Weissman, J. & Lee, B. (2002). The virtual service Grid: An architecture for delivering high-end network services. *Concurrency and Computation*: *Practice and Experience, 14(4)*, pp. 287–319.

Verma. D, Sahu, S., Calo S., Shaikh, A., Chang, I. & Acharya, A. (2003). SRIRAM: A scalable resilient autonomic mesh. *IBM Systems Journal, 42(1).* pp. 19–28.

Wrzesinska, G. Nieuwpoort, R. G. Maassen, J. & Bal, H. E.(2005). Fault-tolerance, malleability and migration for divide-and-conquer applications on the grid. Proceedings *of* IEEE. *International Parallel and Distributed Processing Symposium*, IEEE.

Xiong, N., Yang, Y., Cao, M., He, J. & Shu, L. (2009). A Survey on Fault-Tolerance in Distributed Network Systems. *IEEE International Conference on Computational Science and Engineering, 2*, pp. 1065-1070.