

REPORT DOCUMENTATION PAGE

AFRL-SR-AR-TR-05-

0242

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, gathering existing data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503

1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE June 20, 2005	3. REPORT TYPE AND DATES COVERED Final Apr 15, 2003 – Apr 14, 2005	
4. TITLE AND SUBTITLE A Testbed for Highly-Scalable Mission-Critical Information Systems			5. FUNDING NUMBERS F49620-03-1-0263	
6. AUTHOR(S) Birman, Kenneth P.				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Cornell University 120 Day Hall Ithaca, NY 14853			8. PERFORMING ORGANIZATION REPORT NUMBER 43058	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) AFOSR 4015 Wilson Boulevard Room 713 Arlington, VA 22203-1954			10. SPONSORING / MONITORING AGENCY REPORT NUMBER	
11. SUPPLEMENTARY NOTES				
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for Public Release; distribution is Unlimited				12b. DISTRIBUTION CODE
13. ABSTRACT (Maximum 200 Words) The DURIP cluster has rapidly become a mainstay of our research in Quicksilver, to include scalable services architecture, time critical services, and scalable reliable event delivery. The cluster is anticipated to increase in usage over the next several years to include more members of the Systems groups here at Cornell University. We also anticipate a greater interaction with our colleagues at AFRL in Rome, NY.				
14. SUBJECT TERMS Scalable, Distributed, Gossip, Time-Critical, Event Delivery, Data Mining, Web Services			15. NUMBER OF PAGES 4	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT U	18. SECURITY CLASSIFICATION OF THIS PAGE U	19. SECURITY CLASSIFICATION OF ABSTRACT U	20. LIMITATION OF ABSTRACT UU	

20050705 044

Henk Gtz *Final DURIP Report*

DURIP-03 -
A TESTBED FOR HIGHLY-SCALABLE MISSION CRITICAL INFORMATION SYSTEMS

Final Report

PRINCIPAL INVESTIGATOR: PROF. KENNETH P. BIRMAN
INSTITUTION: CORNELL UNIVERSITY
SPONSORED PROGRAMS
115 DAY HALL
ITHACA, NY, 14853
AWARD NUMBER: F49620-03-1-0263
PROJECT TITLE: (DURIP 03) A TESTBED FOR HIGHLY-SCALABLE MISSION
CRITICAL INFORMATION SYSTEMS

DISTRIBUTION STATEMENT A
Approved for Public Release
Distribution Unlimited

YEARLY PROGRESS REPORT

The DURIP cluster has rapidly become a mainstay of our research in the following areas:

QuickSilver. This effort is building a new system for scalable distributed computing. The basic problem is common in GIG and NCES systems, where an acute need has arisen for simple tools to assist the developer of a distributed service that will be shared by huge numbers of client systems in a networked environment. Headed by Professor Ken Birman, the project is exploring a novel fusion of classical protocols for reliable multicast communication with a new style of peer-to-peer protocol called scalable "gossip". The basic idea is to implement a communication platform using these new protocols, and then integrate the platform with standard Web Services tools and technologies to achieve a uniquely easy to use, scalable and robust solution.

QuickSilver currently has three sub-efforts that rely heavily on the cluster. The first project focuses on what we are calling a "scalable services architecture." In this work we explore a novel new approach to building high performance, scalable, self-managed distributed services that can be (more or less) dragged and dropped onto the cluster. The developer doesn't need to have any special distributed computing experience or skills: our software takes a conventional Web Services application (or a set of them) as input, uses the WSDL document to deduce a replication strategy, and then automatically handles load balancing, restart after failure, workload partitioning and other management and control tasks. The cluster is our primary development target: we expect to reach a point at which relatively unskilled developers are able to put new mission-critical applications on the cluster at the click of a button. This work is in C on Linux.

Key technical ideas include the use of gossip for dynamic membership tracking, inconsistency discovery and state reconciliation, partitioning, chain replication and fault-discovery. The cluster is key here: our work seeks to build the runtime environment, demonstrate the methodology and evaluate it experimentally in realistic military GIG scenarios under stress typical of real deployments. We'll make the solutions available to our colleagues at AFRL in Rome NY.

A second project adopts a similar approach but with a focus on time-critical services. Using a new form of forward error correction, this activity seems to support a new kind of time-critical or real-time replication technology that includes support for deadline-driven communication, periodic communication, and guaranteed low-latency responsiveness even in the face of load bursts or failures. We believe we can slash response times by as much as two orders of magnitude. Again, the methodology is heavily experimental. This work is in Java on Linux, and again uses Web Services paradigms and standards.

A third project focuses on scalable reliable event delivery, messaging and notification. Working with the virtual synchrony model, and implementing in C# on Windows .NET, we are building a tremendously scalable high speed communications infrastructure that remains stable and self-managed even under intense stress and even with vast numbers of clients (most likely running on Windows PCs using Web Services interfaces and standards).

Scalable event filtering and data mining. This work is just getting underway (it required a large storage capability, and only just became practical). The emphasis is on experimentally exploring new technologies for building high speed, scalable, data processing systems in which a variety of probabilistic techniques are exploited to improve scalability and performance. We expect to transition our solutions to projects like the AFRL JBI, which centers on a massive information repository and requires a high speed content filtering technology to achieve good scalability and performance.

STATUS OF EFFORT

The hardware has arrived in its entirety and has been installed.

ACCOMPLISHMENTS

The final stages of hardware implementation were completed within the last 6 months. This is including the software design with a focus on the overall control environment of nodes and switches, including network and node configuration according to the specification of the emulation.

PERSONNEL SUPPORTED

No personnel are supported under this grant.

Personnel associated with the research effort:

- Prof. Kenneth P Birman
- Prof. Alan Demers
- Prof. Fred B. Schneider
- Prof. Johannes E. Gehrke
- Prof. Emin Gun Sirer
- Dr. Robbert van Renesse
- Dr. Werner H. Vogels

PUBLICATIONS

Slingshot: Time-Critical Multicast for Clustered Applications. Mahesh Balakrishnan, Stefan Pleisch, Ken Birman. To Appear in IEEE Network Computing and Applications 2005 (NCA 05). Boston, MA.

At this time we have one paper that has utilized the cluster, this is due in large part to the cluster's relatively new existence coupled with the configuration and troubleshooting period during the installation phase. However, we do currently have multiple users planning on using the technology and we fully expect to publish numerous papers over the next several years.

INTERACTION/TRANSITIONS

During this period we continued interaction with Dr. Mark Linderman to achieve 3 goals:

- Involvement of the JBI team in the requirements and design phase of the system
- Investigation of experimentation scenarios executed by Cornell researchers that are relevant to the JBI effort.
- Investigation of ways that the JBI team can access the Testbed for running unclassified scalability scenarios.