

SpringerBriefs in Electrical and Computer Engineering

For further volumes:
<http://www.springer.com/series/10059>

Rao Mikkilineni

Designing a New Class of Distributed Systems

Rao Mikkilineni
Kawa Objects Inc.
Los Altos, CA
USA
e-mail: rao@kawaobjects.com

ISSN 2191-8112

ISBN 978-1-4614-1923-5

DOI 10.1007/978-1-4614-1924-2

Springer New York Dordrecht Heidelberg London

e-ISSN 2191-8120

e-ISBN 978-1-4614-1924-2

Library of Congress Control Number: 2011940831

© The Author(s) 2011

All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Springer Science+Business Media, LLC, 233 Spring Street, New York, NY 10013, USA), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden. The use in this publication of trade names, trademarks, service marks, and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

A new computing model, extends the von Neumann stored program control (SPC) computing model to program and execute self-configuring, self-monitoring, self-healing, self-protecting and self-optimizing (in short, self-managing or self-*) distributed software systems. As opposed to self-organizing systems that evolve based on probabilistic considerations, this approach focuses on the encapsulation, replication, and execution of distributed and managed (regulated) tasks that are specified precisely. Its innately parallel architecture (non-von Neumann), and its architectural resiliency of cellular organisms, are ideally suited to exploit the many-core architecture and low-latency bandwidth networks emerging in the new generation of data centers to improve the price/performance of IT infrastructure by orders of magnitude.

The new computing model (known as the Distributed intelligent managed element (DIME) Network Architecture) consists of a signaling network overlay over the computing service network and allows parallelism between the control (setup, monitoring, analysis and reconfiguration based on workload variations, business priorities and latency constraints) and the computing functions of the distributed software components. A workflow is implemented as a set of tasks, arranged or organized in a directed acyclic graph (DAG) and executed by a managed network of DIMEs. These tasks, depending on user requirements are programmed and executed as loadable modules in each DIME. The distributed software components along with associated profiles defining their use and management constraints are executed by DIMEs endowed with self-management and signaling-enabled-control architecture. The profiles are used as blueprints to setup, execute and control the down-stream DAG at each node based on global and local policies which depend on business priorities, workload fluctuations and latency constraints. A new class of distributed systems with the architectural resilience of cellular organisms are possible using the DIME network architecture with signaling to monitor and regulate the execution of computational workflows with self-* properties.

We describe two proofs-of-concept implementations; one using a new native operating system called Parallax which is specially designed using the DIME

computing model to monitor and control multi-core and many-core processors in an Intel processor based server; another using a conventional operating system (Linux) to encapsulate a process as a DIME. The results suggest that the service management can be decoupled from the underlying hardware infrastructure management by utilizing signaling based dynamic configuration of DIME network architecture, and potentially reduce the layers of management software in developing next generation highly-scalable, parallel, distributed virtual service creation, delivery, and assurance platforms.

The purpose of this research brief is to introduce some new ideas based on the study of cellular organisms, human organizational networks and telecommunication networks to improve the resiliency, efficiency and scaling of distributed systems executing various computational workflows. The ideas extend the current von-Neumann computing model by separating the services management from services execution exploiting the parallelism and performance offered by the new class of many-core processors.

The DIME network architecture is a departure from conventional wisdom currently being pursued by the universities and corporate research & development. It adds monitoring and control to each Turing computing node and a parallel signaling enabled network to implement the management of temporal behavior of workflows executed as directed acyclic graphs using a network of managed Turing machines.

The concept of a parallel signaling channel is foreign to the current generation of IT professionals, except for those with telecommunications or voice over IP experience, who are by now either retired or dead. Signaling allows establishing equilibrium patterns and monitor and control exceptions system-wide. It allows contention resolution based on system-wide view and eliminates race conditions and other common issues found in current distributed computing practice. In systems with strong dynamic coupling between various elements of the system, where each change in one element continually influences other element's direction of change, signaling in the computational model helps implement system-wide coordination and control based on system-wide priorities, workload fluctuations and latency constraints.

We have demonstrated the feasibility of this approach using two prototypes. However, in order to take these concepts to practical application in mission critical environments, the DIME network architecture requires its validation and acceptance by a larger community. This research brief presents the concept and the results for such validation and analysis.

Acknowledgments

I am extremely indebted to Dr. Giovanni Morana, Daniel Zito and Marco Di Sano from the University of Catania, for their contributions to the development of the DIME Network Architecture and the prototype in Linux operating system. I am also grateful to Ian Seyler from Return Infinity who developed the native operating system from scratch implementing the DIME network architecture and demonstrated self-repair, auto-scaling and live migration in multi-core servers without the use of Hypervisors. All the prototype work was carried out by individuals in their spare time, interested in exploring new ideas that go beyond conventional thinking in spite of the risk involved, with no funding from any organization. I also wish to acknowledge many valuable discussions with and encouragement from Kumar Malavalli, and Albert Comparini from Kawa Objects Inc.

About the Author

Dr. Rao Mikkilineni received his PhD from University of California, San Diego in 1972 working under the guidance of prof. Walter Kohn (Nobel Laureate, Chemistry, 1998). He later worked as a research associate at the University of Paris, Orsay, Courant Institute of Mathematical Sciences, New York and Columbia University, New York. He is currently the Founder and CTO of Kawa Objects Inc., California, a Silicon Valley startup developing next generation computing infrastructure. His past experience includes working at AT&T Bell Labs, Bellcore, U S West, several startups and more recently at Hitachi Data Systems.

Dr. Mikkilineni co-chairs the 2nd “Convergence of Distributed Clouds, Grids and their Management” conference track in IEEE International WETICE 2012 Conference. Current work on DIME network architecture was first presented in WETICE 2010 in Larissa, Greece based on the workshop discussions started in WETICE 2009 in Groningen, The Netherlands.

Contents

1	An Introduction to the Design of a New Class of Distributed Systems	1
	Communication, Collaboration, and Commerce at the Speed of Light	1
	Distributed Systems and their Management	2
	Evolution, Revolution and the Adoption of Disruptive Innovation.	5
	Organization of this Research Brief	7
	References	8
2	Understanding Distributed Systems and Their Management	9
	Stored Program Control Computing Model, Distributed Computing, and Management	9
	Cellular Organisms, Genes, Chromosomes and Distributed Computing.	16
	Human Networks and Distributed Computing	19
	Telecommunications and Distributed Computing.	21
	Client Server, Peer-to-Peer, Grid and Cloud Computing Management.	22
	Client–Server and Peer-to-Peer Models	22
	Grid and Cloud Computing.	23
	Hardware Upheaval and The von Neumann Bottleneck	24
	References	28
3	Distributed Intelligent Managed Element (DIME) Network Architecture Implementing a Non-von Neumann Computing Model.	29
	The DIME Network Architecture and the Anatomy of a DIME	29

DIME Network Architecture and the Architectural Resiliency of Cellular Organisms	34
DIME Network Architecture with a Native OS in a Multi-Core Processor.	36
DIME Network Architecture Implementation in Linux.	41
Grids, Clouds, DIMEs and Their Management	43
References	46
4 Designing Distributed Services Creation, Service Delivery and Service Assurance with the Architectural Resilience of Cellular Organisms	47
The Dial-Tone Metaphor and the Service Creation, Delivery and Assurance Platforms.	47
DIME Network Architecture and the Resilient Service Creation, Delivery and Assurance Environment.	49
DIMEs in Linux Implementation.	52
DIMEs in Multi-Core Server Using a Native Operating System Called Parallax	55
References	57
5 Dime Network Architecture: Future Research Directions and Conclusion	59
Designing a New Class of Distributed Systems Using DIME Network Architecture.	59
Future Research Directions.	61
Conclusion	63
References	65