Pontifícia Universidade Católica
do Rio de Janeiro

Rafael Oliveira Vasconcelos

A Dynamic Load Balancing Mechanism for Data Stream Processing on DDS Systems

Dissertação de Mestrado

Dissertation presented to the Programa de Pós-Graduação em Informática of the Departamento de Informática, PUC-Rio as partial fulfillment of the requirements for the degree of Mestre em Informática.

Advisor: Prof. Markus Endler

Rio de Janeiro
January 9th, 2013

Rafael Oliveira Vasconcelos

A Dynamic Load Balancing Mechanism for Data Stream Processing on DDS Systems

Dissertation presented to the Programa de Pós-Graduação em Informática of the Departamento de Informática, PUC-Rio as partial fulfillment of the requirements for the degree of Mestre em Informática. Approved by the following commission.

Prof. Markus Endler
Advisor
Departamento de Informática – PUC-Rio

Prof. Noemi Rodriguez
Departamento de Informática – PUC-Rio

Prof. Renato Fontoura de Gusmão Cerqueira
Departamento de Informática – PUC-Rio

Prof. José Eugenio Leal
Coordinator of the Centro Técnico Científico – PUC-Rio

Rio de Janeiro, January 9th, 2013

Rafael Oliveira Vasconcelos

BSc. in Computer Science at UNIT (Universidade Tiradentes) – 2010.

Ficha Catalográfica

Vasconcelos, Rafael Oliveira

A dynamic load balancing mechanism for data stream: processing on DDS systems / Rafael Oliveira Vasconcelos ; advisor: Markus Endler. – 2013.
74f ; 30 cm

Dissertação (mestrado)–Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Informática, 2013.
Inclui bibliografia

1. Informática – Teses. 2. Balanceamento de carga. 3. Processamento de fluxo de dados. 4. DDS. 5. SDDL. 6. Publish/Subscribe. 7. Computação autonômica. I. Endler, Markus. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Informática. III. Título.

CDD: 004

Acknowledgments

First and foremost , I have to thank my parents for all unconditional and eternal support, which includes not only financial support but also loving and educational ones. They are the pillars that hold me up.

Apart from the efforts of myself, the success of any thesis depends largely on the encouragement and guidelines of many others. I take this opportunity to express my gratitude to the people who have been instrumental in the successful completion of this project. I would like to show my greatest appreciation to my advisor Markus Endler and all my LAC friends. I can't say thank you enough for their tremendous support and help. Without their encouragement and guidance this project would not have materialized. The guidance and support received was vital for the success of this work. I am grateful for their constant support and help.

I also would like to extend my sincere thanks to Twin Oaks Computing, in special to Nina Tucker, for the support provided during this journey.

Lastly, I am grateful for all other people that helped me direct or indirectly.

Abstract

Vasconcelos, Rafael Oliveira; Endler, Markus. **A Dynamic Load Balancing Mechanism for Data Stream Processing on DDS Systems.** Rio de Janeiro, 2013. 74p. MSc. Dissertation – Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

This thesis presents the Data Processing Slice Load Balancing solution to enable dynamic load balancing of Data Stream Processing on DDS-based systems (Data Distribution Service). A large number of applications require continuous and timely processing of high-volume of data originated from many distributed sources, such as network monitoring, traffic engineering systems, intelligent routing of cars in metropolitan areas, sensor networks, telecommunication systems, financial applications and meteorology. The key concept of the proposed solution is the Data Processing Slice (DPS), which is the basic unit of data processing load of server nodes in a DDS Domain. The Data Processing Slice Load Balancing solution consists of a load balancer, which is responsible for monitoring the current load of a set of homogenous data processing nodes and when a load unbalance is detected, it coordinates the actions to redistribute some data processing slices among the processing nodes in a secure way. Experiments with large data stream have demonstrated the low overhead, good performance and the reliability of the proposed solution.

Keywords

Load Balancing; Data Stream Processing; DDS; SDDL; Publish/Subscribe; Autonomic Computing.

Resumo

Vasconcelos, Rafael Oliveira; Endler, Markus. **Um Mecanismo de Balanceamento de Carga Dinâmico para Processamento de Fluxo de Dados em Sistemas DDS.** Rio de Janeiro, 2013. 74p. Dissertação de Mestrado – Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

Esta dissertação apresenta a solução de balanceamento de carga baseada em fatias de processamento de dados (Data Processing Slice Load Balancing solution) para permitir o balanceamento de carga dinâmico do processamento de fluxos de dados em sistemas baseados em DDS (Data Distribution Service). Um grande número de aplicações requer o processamento contínuo de alto volume de dados oriundos de várias fontes distribuídas., tais como monitoramento de rede, sistemas de engenharia de tráfego, roteamento inteligente de carros em áreas metropolitanas, redes de sensores, sistemas de telecomunicações, aplicações financeiras e meteorologia. Conceito chave da solução proposta é o Data Processing Slice, o qual é a unidade básica da carga de processamento dos dados dos nós servidores em um domínio DDS. A solução consiste de um nó balanceador, o qual é responsável por monitorar a carga atual de um conjunto de nós processadores homogêneos e quando um desbalanceamento de carga é detectado, coordenar ações para redistribuir entre os nós processadores algumas fatias de carga de trabalho de forma segura. Experimentos feitos com grandes fluxos de dados que demonstram a baixa sobrecarga, o bom desempenho e a confiabilidade da solução apresentada.

Palavras-chave

Balanceamento de carga; processamento de fluxo de dados; DDS; SDDL; Publish/Subscribe; computação autonômica.

Contents

List of Figures

List of Tables