Pontifícia Universidade Católica
do Rio de Janeiro

Ricardo Gomes Clemente

Uma arquitetura para processamento de eventos de log em tempo real

Dissertação de Mestrado

Dissertação apresentada como requisito parcial para obtenção do título de Mestre pelo Programa de Pós-Graduação em Informática da PUC-Rio.

Orientador: Prof. Marco Antonio Casanova

Rio de Janeiro
Agosto de 2008

Pontifícia Universidade Católica
do Rio de Janeiro

Ricardo Gomes Clemente

Uma arquitetura para processamento de eventos de log em tempo real

Dissertação apresentada como requisito parcial para obtenção do título de Mestre pelo Programa de Pós-Graduação em Informática da PUC-Rio. Aprovada pela Comissão Examinadora abaixo assinada.

Prof. Marco Antonio Casanova
Orientador
Departamento de Informática - PUC-Rio

Prof. Karin Koogan Breitman
Departamento de Informática - PUC-Rio

Prof. Marcelo Tílio M. Carvalho
TecGraf – PUC-Rio

Prof. José Eugenio Leal
Coordenador Setorial do Centro
Técnico Científico – PUC-Rio

Rio de Janeiro, 21 de agosto de 2008

Ricardo Gomes Clemente

Graduou-se em Engenharia Elétrica: Ênfase em Eletrônica pela UFRJ (Universidade Federal do Rio de Janeiro) em Julho de 2006. Trabalha com desenvolvimento e projeto de infra-estrutura de aplicações para *Internet* na Globo.com desde 2005.

Ficha Catalográfica

Clemente, Ricardo Gomes

Uma arquitetura para processamento de eventos de log em tempo real / Ricardo Gomes Clemente ; orientador: Marco Antonio Casanova. – 2008.

77 f. : il.(color.) ; 30 cm

Dissertação (Mestrado em Informática)– Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2007.

Inclui bibliografia

1. Informática – Teses. 2. Log. 3. Sistemas de gerenciamento de log. 4. Correlação de eventos. 5. Arquitetura orientada a eventos. I. Casanova, Marco Antonio. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Informática. III. Título.

CDD: 004

Agradecimentos

Agradeço primeiramente a Deus.

Agradeço ao meu orientador, Prof. Casanova, pela sua imensa capacidade de ensinar e orientar através de longas e agradáveis conversas.

Agradeço à PUC-Rio pela bolsa concedida para a realização desta pesquisa.

Aos demais professores e funcionários da pós-graduação com quem tive a oportunidade de conviver nestes dois anos, em especial ao Prof. Luiz Fernando, com quem iniciei minha pesquisa.

Agradeço à Globo.com por incentivar e apoiar meus estudos concedendo tempo de trabalho para minha dedicação ao mestrado.

Aos meus pais, minha família e minha namorada por me darem apoio e motivação sempre.

Resumo

Clemente, Gomes Ricardo; Casanova, Marco Antonio. **Uma arquitetura para processamento de eventos de log em tempo real.** Rio de Janeiro, 2008. 77p. Dissertação de Mestrado - Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

Logs são, atualmente, riquíssima fonte de informação para administradores de sistemas e analistas de negócio. Em ambientes com grande volume de acesso e infra-estrutura de centenas de servidores, processar toda a informação gerada e correlacioná-la com o objetivo de identificar situações de interesse técnico e de negócio em tempo real, é considerado um grande desafio. Nesse sentido, são explicados tanto os conceitos relacionados aos arquivos de log e aos sistemas que se propõem a gerenciá-los, quanto os métodos e ferramentas de correlação de eventos em tempo real, para que, então, seja proposta uma arquitetura de sistema capaz de lidar com o desafio citado. Por fim, um protótipo é desenvolvido e uma prova de conceito baseada em um caso real de uso é realizada.

Palavras-chave

log, sistemas de gerenciamento de log, correlação de eventos, arquitetura orientada a eventos.

Abstract

Clemente, Gomes Ricardo; Casanova, Marco Antonio. **An architecture for real time log events processing.** Rio de Janeiro, 2008. 77p. MSc. Dissertation - Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

Logs are, nowadays, a rich source of information for system administrators and business analysts. In environments with a high access volume and hundreds of servers, to process every generated information and correlate it, in order to identify interesting technical and business situations in real time, is considered a challenge. Considering that, concepts related to log files and systems that aim to manage it, besides methods and tools for real time event correlation are presented, in order to propose a system architecture capable of overcoming the stated challenge. At last, a prototype is developed and a concept prove based on a real case is done.

Keywords

log, log management systems, event correlation, event-driven architecture

Sumário

Lista de figuras

Lista de tabelas

Lista de siglas

API	Application Program Interface
BAM	Business Activity Monitoring
CEP	Complex Event Processing
CLF	Common Log Format
CSV	Comma Separated Values
DSMS	Data Stream Management System
EDA	Event Driven Architecture
EPN	Event Process Network
ESP	Event Stream Processing
FIFO	First-in First-out
HTTP	Hypertext Transfer Protocol
IPC	Inter-Process Communication
LMS	Log Management System
RFC	Request for Comments
SEC	Simple Event Correlator
SQL	Structured Query Language
XML	Extensible Markup Language
XMPP	Extensible Messaging and Presence Protocol