

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



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

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e do orientador.

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

1 Introdução	12
1.1 Motivação	12
1.2 Proposta da dissertação	13
1.3 Organização da dissertação	13
2 Gerenciamento de Log	15
2.1 Definições básicas	15
2.2 Logs como um fluxo de eventos	16
2.3 Caracterização e padrões de logs	17
2.4 Sistema de gerenciamento de logs	21
3 Eventos: definição e processamento	24
3.1 Definição de eventos	24
3.2 Processamento de eventos	25
3.3 Cenários de uso para o Processamento Complexo de Eventos	28
3.4 Processadores complexos de eventos	29
3.4.1 Processadores baseados em regras	29
3.4.2 Processadores baseado em consultas	30
4 Proposta de arquitetura	35
4.1 Requisitos de um Sistema de Gerenciamento de Logs	35
4.1.1 Requisitos Funcionais	35
4.1.2 Requisitos Não-funcionais	37
4.2 Conceito de Arquitetura Orientada a Eventos	38
4.3 Arquitetura Orientada a Eventos para Gerenciamento de Eventos de Log	40
5 Desenvolvimento do protótipo	47

5.1 Visão geral	47
5.2 Canal Principal de Comunicação de Eventos	48
5.3 Nó de primeira interação	50
5.4 Nó de correlação de eventos por consulta	55
5.5 Um caso de uso do protótipo	57
5.6 Testes funcionais do protótipo	62
6 Conclusão	71
7 Referências	73

Lista de figuras

Figura 1 - Logs como fluxos de eventos	17
Figura 2 – Exemplo de registro de evento de log no padrão <i>Common Logfile Format</i>	19
Figura 3 – Diagrama ilustrando os conceitos relacionados a eventos	28
Figura 4- Comparação entre modelos	31
Figura 5 – Estágios de interações fluxo de eventos	39
Figura 6 – Canal Principal de Comunicação de Eventos, grupos e Nós.	41
Figura 7 – Composição de um Nó	42
Figura 8 – Arquitetura proposta e seus componentes	43
Figura 9– Exemplo de configuração de segmento no Spread	49
Figura 10 – Configuração LogPP	51
Figura 11– Exemplo de regra no SEC	53
Figura 12 – Componentes do Nó de primeira interação	55
Figura 13 – Descrição de um fluxo de eventos de log	56
Figura 14 – Exemplo de consulta contínua	57
Figura 15 – Componentes do Nó de correlação de eventos por consulta	57
Figura 16 – Arquitetura de uma aplicação <i>Web</i>	59
Figura 17 – Filtros configurados no LogPP para prova de conceito	60
Figura 18 – Regras do SEC utilizadas no protótipo	61
Figura 19 – Consultas utilizadas no protótipo	62

Lista de tabelas

Tabela 1 – Níveis de severidade para eventos de log	18
Tabela 2 – Descrição de campos do <i>Common Logfile Format</i>	19
Tabela 3 – Diferentes formatos de mensagens de log com <i>log4j</i>	20
Tabela 4 – Diferentes formatos de eventos de log	45
Tabela 5 – Caso de teste Apache e LogPP	63
Tabela 6 – Caso de teste Apache e LogPP com vírgula	63
Tabela 7 – Caso de teste JbossWeb e LogPP	64
Tabela 8 – Caso de teste do SEC com padrão 404	65
Tabela 9 – Caso de teste do SEC com erro específico	66
Tabela 10 – Caso de teste de fila de mensagens	67
Tabela 11 – Caso de teste do Publicador	68
Tabela 12 – Caso de teste do Consumidor	69
Tabela 13 – Caso de teste do Inersor	70
Tabela 14 – Caso de teste do sistema completo	70

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