

Aplicação web para auxiliar na gestão de logs distribuídos utilizando métodos e técnicas de Recuperação de Informação

Fabricio Kalbusch¹, Raphael G. Krauss², Flavio Ceci³

1. Graduado em Sistemas de informação pela Universidade do sul de Santa Catarina (UNISUL).

*fabriciokalbusch@gmail.com

2. Graduado em Sistemas de informação pela Universidade do sul de Santa Catarina (UNISUL).

3. Possui graduação em Ciência da Computação pela Universidade do Sul de Santa Catarina, mestrado e doutorado em Engenharia do Conhecimento pela Universidade Federal de Santa Catarina. Atua como professor titular dos cursos de graduação em Ciência da Computação e Sistemas de Informação e dos programas de pós-graduação em Governança de TI e Gerência de Projetos de TI da Universidade do Sul de Santa Catarina.

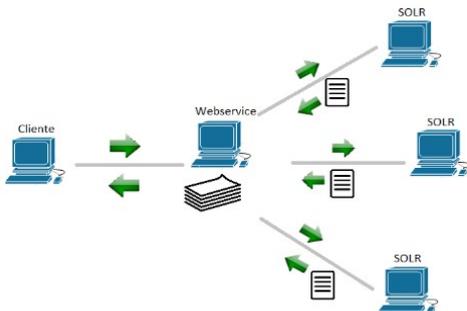
Introdução

Os sistemas distribuídos geralmente possuem uma grande quantidade de componentes que podem ser subsistemas, ferramentas e rotinas auxiliares espalhadas por diversos computadores na rede. Para tornar possível a manutenção, é essencial que estes componentes gerem logs, que são uma forma de registrar as operações realizadas de modo a saber qual o estado da aplicação. Gupta (2005, p. 1) define log como “uma maneira sistemática e controlada de representar o estado de uma aplicação de forma legível”. Geralmente, estes logs são gerados em forma de arquivo no disco do servidor, o que gera uma grande quantidade de barreiras para que os administradores de sistemas possam visualizar essas informações de forma rápida, fácil e sem causar impactos no ambiente. O objetivo é implantar uma aplicação web que gere e recupere informações de logs, que contenham dados semiestruturados e não estruturados, de sistemas distribuídos de uma organização.

Resultados e Discussão

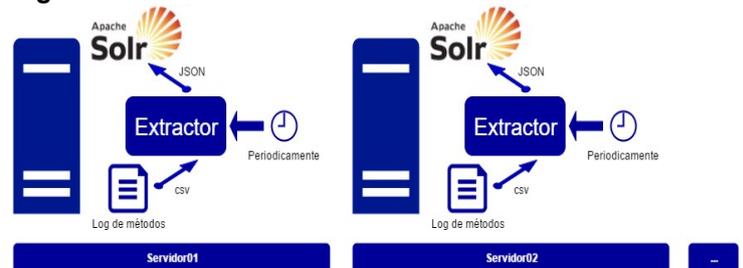
Foi criada uma página web denominada consulta centralizada de logs, que possibilita aos usuários pesquisas em um único local, sem que se preocupem com as barreiras para acesso aos logs. A solução possui alguns componentes que são apresentados a seguir. A **consulta centralizada** de logs é uma página web responsável por enviar os filtros de pesquisa de logs a um webservice de integração e apresentar os resultados. O **webservice de integração** recebe os filtros na requisição de pesquisa. Ele descobre quais servidores tem os logs solicitados e faz chamadas paralelas para cada servidor SOLR. Depois, os resultados são concatenados e retornados a página web, como pode ser visto na figura 1.

Figura 1 - Funcionamento do webservice de integração



O **Solr** é um projeto open source que implementa um servidor de buscas de alta performance. Ele foi utilizado para armazenar e indexar os logs. O Solr responde ao webservice de integração com o resultado dos logs solicitados. Com a finalidade de coletar os logs e enviar ao SOLR (<http://lucene.apache.org/solr/downloads.html>), foi criada uma aplicação chamada **extractor**, que tem seu funcionamento apresentado na imagem a seguir.

Figura 2 - Funcionamento do extractor



Conclusões

Existem diversos problemas que tornam o gerenciamento de logs um desafio. Os principais acontecem por falta de permissão de acesso ou dificuldade em encontrar o exato local onde o log foi gerado. Estes problemas e suas variações impõem barreiras para consultar os logs e, por consequência, geram perda de tempo na cadeia produtiva. Utilizando os princípios de recuperação de informação, com foco em armazenamento, organização e acesso aos itens de informação, o protótipo denominado consulta centralizada de logs foi desenvolvido, provando que é possível consultar os logs de forma centralizada. O protótipo foi avaliado por meio de uma entrevista aplicada, onde ficou evidente os problemas relatados e também ficou claro que os usuários acreditam que a ferramenta pode solucionar estes problemas.

Palavras-chave

Recuperação da informação, Logs, Solr.

Agradecimentos

Agradecemos ao professor Flávio Ceci que sempre nos incentiva e eleva nosso potencial.

Referências

GUPTA, Samudra. **Pro Apache Log4j**: Java application logging using the open source Apache log4j API. 2. ed. New York: Apress, 2005. 211 p.