



PROCESSAMENTO DISTRIBUIDO EM LARGA ESCALA DE FLUXOS DE EVENTOS

Isabelly Lourêdo Rocha¹, Andrey Elísio Monteiro Brito²

RESUMO

Avanços tecnológicos recentes permitem que dispositivos eletrônicos dos mais diversos tipos possam estar conectados através de uma rede de computadores, como exemplo disso temos os celulares que podem gerar informações periódicas sobre sua localização geográfica e a produção constante de conteúdo em sistemas de redes sociais. A disponibilidade desses dados que são produzidos continuamente tem motivado o desenvolvimento de aplicações de processamento de fluxo de eventos que têm como objetivo processar, em tempo real, os dados que são gerados continuamente de forma a detectar situações de interesse. Buscando-se encontrar uma forma eficiente de processar dados que são gerados continuamente e em larga escala, desenvolvemos algumas dessas aplicações de processamento de fluxo de eventos adaptando-as a diferentes ambientes de processamento com o objetivo de analisar e comparar o desempenho de cada uma delas nos diferentes ambientes observados. As ferramentas utilizadas pelos experimentos de análise foram o OurGrid, uma outra ferramenta de processamento de dados desenvolvida usando um arcabouço próprio que possui uma versão paralela e outra distribuída e a ferramenta StreamMine3G. Desse modo, observamos que cada uma das alternativas propostas têm suas vantagens e desvantagens, bem como suas limitações dependendo do tipo de aplicação que precisamos implementar.

Palavras-chave: Aplicações de Processamento de Fluxo de Eventos, OurGrid, StreamMine3G

DISTRIBUTED PROCESSING LARGE-SCALE FLOW OF EVENTS

ABSTRACT

Recent technological advances allow electronic devices of all kinds to be connected via a computer network, as an example we have phones that can generate periodic reports on their geographical location and the constant production of content in social network systems. The availability of these data that are continuously produced has motivated the development for applications processing event streams, which aim at processing, in real time, data generated continuously and in large scale in order to detect situations of interest. Looking for a efficient way of processing data that are generated continuously and on a large scale, we developed some event stream processing applications and adapted them to different processing environments in order to analyze and compare the performance of each in these different environments. The tools used for the experiments of analysis were OurGrid, another data processing tool designed to use our own framework, which has a parallel and a distributed version, and the StreamMine3G tool. Thus, we observed that each of the proposed alternatives have advantages and disadvantages, as well as limitations which depend on the type of application to be implemented.

Keywords: Applications For Event Stream Processing, OurGrid, StreamMine3G

¹ Aluna do Curso de Ciência da Computação, Unidade Acadêmica de Sistemas e Computação, UFPG, Campina Grande, PB, E-mail: isabelly.rocha@ccc.ufpg.edu.br

² Ciência da Computação, Professor. Doutor, Unidade Acadêmica de Sistema e Computação, UFPG, PB, E-mail: andrey@dsc.ufpg.edu.br