

Acelerando o Fluxo Global de Tarefas no Processo de Impressão Distribuída

Rafael Nemetz, Mateus Raeder,
Mariana Kolberg, Luiz Gustavo Fernandes

Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS)
Av. Ipiranga, 6681 - Prédio 32 - PPGCC
{rafael.nemetz, mateus.raeder}@acad.pucrs.br, {mariana.kolberg, luiz.fernandes}@pucrs.br

Introdução

Com o avanço tecnológico das impressoras industriais, a demanda por publicação em papel (livros e publicidade) sofreu grande crescimento. Este crescimento vem sendo atendido através da informatização e automação de todo o processo de impressão. Surgem, então, formatos digitais que permitem descrever documentos em alto nível, como o PS (*PostScript*) e o PDF (*Portable Document Format*). Estes formatos, no entanto, não são reconhecidos pelas impressoras e, antes da impressão, precisam passar por um processo de conversão bastante pesado em termos de custos computacionais. Dentre as diversas etapas que envolvem a impressão de documentos, esta etapa de conversão (chamada rasterização) é mais lenta que as demais, prejudicando o fluxo geral do sistema.

A investigação de técnicas de alto desempenho que possam aumentar a velocidade da rasterização traz ganhos para todo o sistema e merece atenção individual. Na busca por este desempenho, são empregadas técnicas de processamento paralelo. Diferentes formas de melhorar a utilização dos *clusters* responsáveis pela rasterização de documentos PDF foram sugeridas e testadas em estudos anteriores [NGK 09]. Em [NRK 09], foi desenvolvida uma ferramenta capaz de extrair as informações mais relevantes de cada página de um PDF. Esta ferramenta possibilitou que fossem encontradas métricas que expressam a complexidade de cada tarefa. Conhecendo a complexidade das tarefas, puderam ser testados diferentes algoritmos de escalonamento que objetivam encontrar um balanceamento de carga ideal para determinado *cluster*. Visando um balanceamento de carga ainda mais justo, estão sendo testados algoritmos que, baseados nas informações extraídas de cada documento, podem quebrá-lo em partes de complexidades equivalentes.

Escalonamento global

Além do trabalho direcionado ao ganho de desempenho em um âmbito local, dentro de cada *cluster*, o cenário apresentado pelas casas de impressão atuais permite também a realização de uma análise mais global do fluxo de tarefas. Esta análise pode evidenciar novos aspectos a serem otimizados. Normalmente, existe mais de um *cluster* responsável pela etapa de rasterização de documentos. A otimização do gerenciamento da fila das tarefas que esperam pelo processamento, assim como do escalonamento e balanceamento de carga entre os *clusters* presentes (podem ser heterogêneos), poderá aumentar ainda mais a velocidade do processo de impressão como um todo.

Este trabalho consiste na busca por um escalonador global, uma vez que este escalonamento ainda é realizado manualmente. O escalonador deve ser capaz de lidar com

as prioridades, restrições de tempo e outras necessidades que possam ser impostas pelos clientes às suas tarefas. Com o intuito de prover informações úteis ao escalonador, um mecanismo de previsão irá estimar o tempo que determinada tarefa levará para ser rasterizada em determinado *cluster*. Como entrada, este mecanismo utilizará a complexidade da tarefa (obtida com as ferramentas citadas anteriormente) e informações sobre recursos disponíveis nos *clusters*. Para a busca de recursos e para o escalonador propriamente dito, serão propostos algoritmos capazes de funcionar sob o JDF (*Job Definition Format*), protocolo específico de sistemas de impressão [BUC 06].

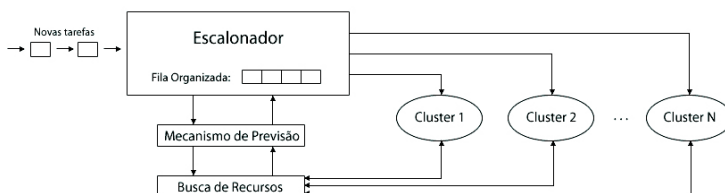


Figura 1: Ilustração simplificada do sistema.

O algoritmo de escalonamento envolverá uma investigação das possíveis necessidades dos clientes e das propriedades físicas de suas tarefas. Tais necessidades englobam fatores que podem ser de interesse ao serem especificadas características da tarefa, como tempo para término, prioridades, espaço de cores, resolução, entre outros. As propriedades físicas são aquelas extraídas diretamente do documento, como a presença de transparência, a quantidade de objetos reusáveis e o tamanho dos arquivos. Tarefas iguais serão processadas mais rapidamente se estiverem juntas, uma vez que as máquinas armazenam em cache os objetos rasterizados. Podem haver, também, máquinas otimizadas para transparência e aumentar suas cargas de tarefas transparentes pode ser vantajoso. Assim, mantendo a fila de tarefas organizada de forma inteligente, o escalonador distribuirá os trabalhos de uma forma mais adequada (Figura 1).

Para que o sistema seja compatível com o JDF, um estudo aprofundado do protocolo será necessário. A realização de testes requerirá, ainda, o desenvolvimento de um ambiente simulado que contenha todas as entidades presentes no sistema e que seja capaz de reconhecer os protocolos por elas utilizados.

Referências

- [BUC 06] Buckwalter, C. Integrating Systems in the Print Production Workflow: Aspects of Implementing JDF. Digital Media Division, 2006.
- [NGK 09] Nunes, T. ; Giannetti, F. ; Kolberg, M. ; Nemetz, R. ; Cabeda, A. ; Fernandes, L. G. . Job Profiling in High Performance Printing. In: 9th ACM Symposium on Document Engineering, 2009, Munique. Proceedings of the 9th ACM Symposium on Document Engineering (ACM DOCENG). Nova Iorque, NY (EUA) : ACM Press, 2009. p. 109-118.
- [NRK 09] Nunes, T. ; Raeder, M. ; Kolberg, M. ; Fernandes, L. G. ; Cabeda, A. ; Giannetti, F. . High Performance Printing: Increasing Personalized Documents Rendering through PPML Jobs Profiling and Scheduling. In: 12th IEEE CSE - International Conference on Computational Science and Engineering, 2009, Vancouver (Canadá). Proceedings of the 12th IEEE International Conference on Computational Science and Engineering (CSE). Los Alamitos, CA (EUA): IEEE Computer Society, 2009. p. 285-291.