

Análise de Perfis de Execução de Submissão de Trabalhos em Ambientes Multi-Agregados Usando o Portal ICE*

Laércio Lima Pilla, Rodrigo da Rosa Righi e Philippe Olivier Alexandre Navaux
Grupo de Processamento Paralelo e Distribuído (GPPD) - Instituto de Informática - UFRGS
{llpilla, rodrigo.righi, navaux}@inf.ufrgs.br

Resumo

The ICE - Integrated Cluster Environment - Web Portal offers access to various clusters. It hides technical details about the tools that effect job submission functionality on clusters. In ICE environment, the interactions among the Web Portal and the clusters are made with Web Services technology. The Web Services usage has been growing in the construction of distributed environments, once that it treats with protocol standardization and offers interoperability. However, its performance is questioned, because Web Services use messages in clear text (SOAP-XML) and embracing a container on the service implementation side. In this direction, this paper shows an analysis of the profiling of job submissions in different clusters using the ICE Portal. The idea is to verify the cost added by the ICE and the job submission performance using different schedulers.

1. Introdução

O ambiente ICE[3] é um Portal Web que age como um facilitador no acesso a agregados (*clusters*). Ele abstrai para o usuário as diferentes características das ferramentas de escalonamento de trabalhos presentes nos agregados. Esse Portal foi construído com base numa Arquitetura Orientada a Serviços e foi viabilizado através de Web Services[1].

A tecnologia de Web Services se destaca pelas características de fraco acoplamento, uso de protocolos padrão (SOAP-XML) e interoperabilidade. Empregada no ICE, ela oferece estensibilidade em dois níveis: (i) adição de novas funcionalidades de agregado (como gerenciamento de trabalhos e monitoramento); (ii) implementação de vários Web Services para uma dada funcionalidade. Entretanto, a duração de uma chamada Web Service é criticada em sistemas que visam alto desempenho. Isso porque os dados são transferidos usando mensagens SOAP-XML sobre o protocolo HTTP. Além disso, o Web Service está associado a um servidor (*container*) que agrega mais custo a chamada.

Para analisar o sobrecusto causado pelo Portal ICE, foi usada a técnica que mede o perfil de execução (*profiling*)[2]. À partir desta técnica, o presente artigo avalia a submissão de trabalhos pelo Portal ICE em dois agregados. O primeiro deles é o Labtec, no Instituto de Informática da UFRGS, e outro é o LCP, no Instituto Coppe da UFRJ. Os objetivos do trabalho incluem: (i) a comparação dos dados do perfil com a execução direta do escalonador (em linha de comando); (ii) a observação do desempenho de diferentes escalonadores de trabalhos e; (iii) a medição do impacto de uma chamada Web Service para diferentes domínios administrativos.

2. Perfil de Execução no ICE

O perfil de execução da submissão de trabalho engloba tempos no Portal Web e no Web Service presente nos agregados. O tempo total de execução é calculado à partir dos diversos tempos capturados no perfil. Estes foram colhidos através do método `System.nanoTime()` presente na linguagem Java desde a sua versão 1.5.

Neste experimento são considerados 6 tempos diferentes. Os tempos de T1 a T5 são referentes às atividades desempenhadas na máquina do Portal Web. O T1 refere-se a captura dos dados da submissão de trabalho. O tempo T2 trata do tempo de verificação das permissões de acesso do usuário. No tempo T3 encontra-se a construção da mensagem SOAP-XML a ser enviada ao agregado. O tempo T4 mede a duração de uma chamada síncrona ao Web Service. O último tempo, T5, refere-se ao desempacotamento do retorno do serviço e a gravação de informações no repositório.

O tempo T6 refere-se ao serviço no agregado e é sub-dividido em quatro intervalos. Primeiramente é medido o tempo de desempacotamento da mensagem SOAP-XML (T6D). O tempo de mapeamento do usuário ICE para usuário do agregado corresponde a medida T6M. A sub-divisão T6S encapsula o tempo gasto pelo escalonador na submissão de um trabalho em lote (*batch*). É importante frisar que esse tempo não engloba a duração da execução do trabalho. Por fim, a duração do empacotamento da resposta está inclusa no tempo T6E.

* Pesquisa parcialmente financiada pelos Órgãos Finep e CNPq

Ambiente	Tempos capturados no ICE					Tempos capturados no agregado			
	T1	T2	T3	T4	T5	T6D	T6M	T6S	T6E
LCP	1695.71	3538.85	97.28	235691.14	26075.28	889.71	10301.78	150692.61	34.15
Labtec	1792.85	3686.14	95.85	1017236.10	28044.67	1451.42	31779.42	924576.71	62.28

Tabela 1. Tempos (em microssegundos) de uma submissão de trabalho em diferentes ambientes

3. Resultados

Para realizar o perfil de execução foram submetidas 50 requisições de trabalhos para cada um dos agregados. Com base nos tempos capturados, foram calculadas suas médias aritmética e seus desvios padrão. A cada submissão de trabalho, o Portal ICE comunica-se com a máquina de entrada (*front-end*) do agregado escolhido. Tal máquina no LCP possui um processador Pentium 4 de 2.40 GHz e 1GB de memória, enquanto que no agregado Labtec possui um processador Pentium 3 de 866MHz e 768MB de memória.

A Tabela 1 apresenta os tempos de submissão de trabalho capturados. A terminologia dos tempos nessa tabela pode ser vista na Seção 2. Dos tempos capturados ainda na máquina do Portal ICE, o que possui maior variação é o tempo de serviço (T4). Tal variação é devida à diferença nos tempos de execução do serviço nos agregados.

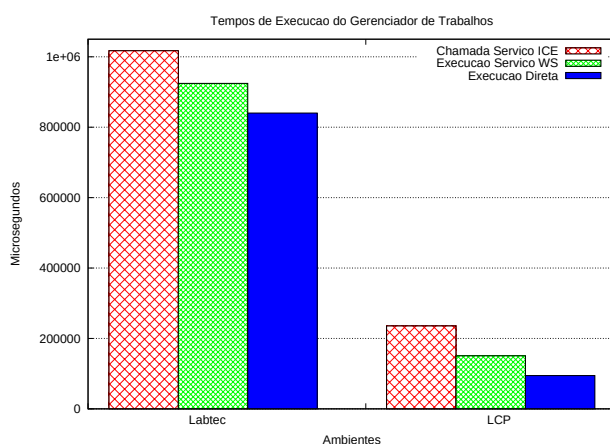


Figura 1. Tempos em cada ambiente

Outros resultados obtidos são os tempos parciais nos agregados. A desigualdade entre os tempos T6S pode ser atribuída ao uso de escalonadores de trabalho diferentes nos agregados (OAR no Labtec e SGE no LCP). Esse resultado deve-se ao fato que OAR grava suas informações em banco de dados, enquanto o SGE as grava mais rapidamente em modo texto. O mapeamento de usuário é medido no tempo T6M, e mostra-se mais custoso no Labtec. Isto deve-se ao fato de que a captura do diretório do usuário necessita de uma consulta à base de dados NIS, enquanto no LCP faz-se apenas uma leitura de arquivo (/etc/passwd).

Subtraindo do tempo T4 o tempo T6, obtém-se uma estimativa do tempo de rede e da sobrecarga do servidor de Web Service. Esse tempo é igual a 73772.89 μ s para o agregado LCP e 59366.27 μ s para o Labtec, um resultado esperado, pois o segundo agregado encontra-se no mesmo domínio administrativo da máquina do Portal ICE.

O gráfico da Figura 1 mostra três tempos de execução para cada ambiente testado. O primeiro deles refere-se a chamada síncrona ao Web Service pelo ICE (tempo T4). O segundo tempo é a execução da requisição do Web Service no agregado, medido no tempo T6S. Por fim, o último tempo caracteriza execução direta (em linha de comando) de uma submissão de trabalho.

Numa análise dos tempos de cada ambiente da Figura 1, observa-se que os tempos de submissão de trabalhos pelo Portal ICE são de aproximadamente 260 ms no agregado LCP, e 1000 ms no Labtec. Mesmo assim, em ambos ambientes a sobrecarga é pequena e as vantagens proporcionadas pelo ICE sobrepujam esse custo.

4. Conclusão

Comparando os tempos capturados pelo perfil com a execução direta do escalonador, percebe-se que o Portal ICE é viável e impõe um custo inferior a 1 segundo em ambos ambientes testados. O ICE acrescenta um tempo que é mais perceptível na submissão de trabalho ao agregado LCP, pois esse possui um escalonador mais rápido. Analisando os resultados obtidos, também é possível notar que o tempo de acesso (tempo de rede) do Portal ICE ao agregado LCP é maior que o tempo de acesso ao agregado Labtec. Isso porque este último está no mesmo domínio que o Portal ICE.

Referências

- [1] P. Kumar. Web services in it management. *ACM Queue*, 3(6), August 2005.
- [2] M. Oikonomakos, K. Christodoupoloulos, and E. Varvarigos. Profiling computation jobs in grid systems. In *CCGRID - Cluster Computing and the Grid*, pages 197–204, Rio de Janeiro - RJ, May 2007.
- [3] R. Righi, C. Marquezan, L. Schnorr, A. Carissimi, N. Mailard, and P. Navaux. Portal ice - uso de web services para atingir extensibilidade no gerenciamento de múltiplos agregados. In *Workshop em Sistemas Computacionais de Alto Desempenho*, pages 133–140, Ouro Preto - MG, Nov. 2006.