

Lidando com Demanda e Variedade de Serviços em Grades Computacionais

Marcelo Silva Santos¹, Álvaro Vinícius de Souza Coêlho¹

¹Departamento de Ciências Exatas e Tecnológicas - DCET

Universidade Estadual de Santa Cruz - UESC

Rod. Ilhéus-Itabuna, Km. 16 – 45.662-900 – Ilhéus – BA – Brasil

Resumo. *Grades computacionais são largamente estudadas como alternativa para computação de alto desempenho. Quando as consideramos em ambientes baseados em reciprocidade, a seleção dos serviços a serem oferecidos é a única forma se obter receita e, assim, manter o interesse dos nós em permanecer no sistema. Ocorre que computar o melhor portfolio de serviços possível é intratável, de forma que os nós implementam heurísticas de seleção. Este trabalho explora o desempenho de diferentes heurísticas submetidas a diferentes condições no tocante à demanda e variedade de serviços.*

1. Introdução

A distribuição e o paralelismo vem sendo atalhos bastantes estudados quando se caminha rumo à computação de alto desempenho. Neste sentido, as grades computacionais são uma alternativa atraente, possibilitando o uso de uma federação de recursos para execução de aplicações paralelas [Andrade et al. 2007]. Todavia, estes recursos são disponibilizados por diferentes domínios administrativos (os nós da grade), e cada um deles é autônomo. Logo, é necessário conciliar interesses diversos para que eles permaneçam no sistema. Para alcançar este objetivo, o ganho obtido por cada nó precisa superar seus custos.

Os nós de uma grade valoram os serviços diferentemente uns dos outros [Mowbray et al. 2006]. Por esse motivo, para cada nó, a alocação de recursos precisa ter custos e receitas em vista, de forma que se perceba um lucro. Para isso, pode-se empregar modelos de negociação baseados em teorias de mercado [Buyya and Vazhkudai 2000], mas estes modelos impõem uma série de mecanismos de controle como os que se vê nas operações comerciais e bancárias em geral, gerando custos marginais transacionais (contratos, auditoria, contabilidade, cobrança e registros bancários). Alternativamente, emprega-se modelos baseados em reciprocidade, prescindindo destes controles [Andrade et al. 2007]. Em ambientes assim, a única maneira de um nó conseguir receber recursos (e receita) é sendo recíproco¹. Logo, os nós devem encontrar a melhor seleção possível de serviços. Como isto é computacionalmente intratável, usa-se *heurísticas de seleção* [Coêlho et al. 2009]. Ocorre que essas heurísticas não podem ser avaliadas em condições reais de operação por conta da complexidade e do indeterminismo desses ambientes. Neste sentido, Coêlho e Brasileiro apresentam uma metodologia que permite avaliar heurísticas sob diferentes condições [Coêlho and Brasileiro 2009], e exploraram o desempenho de algumas delas considerando diferentes orçamentos médios disponíveis para os nós. Neste artigo, exploramos outros aspectos: o tamanho do universo de serviços disponíveis e a quantidade de serviços diferentes consumidos pelos nós.

¹Por simplificação, consideramos que as estratégias cuidarão de marginalizar os *free riders*

2. Descrição do Problema

Consideramos ambientes em que os nós oferecem e consomem serviços computacionais, como a execução de um software, a disponibilização de algum conjunto de dados, etc. usando um modelo de reciprocidade. Além disso, é importante observar que os serviços não são consumidos indistintamente pois cada nó consome alguns serviços específicos, desinteressando-se pelos demais. Andrade et.al. [Andrade et al. 2007] definiu *Favor Típico (FT)* como sendo as diferentes proporções em que cada serviço é consumido por um nó. Portanto, para estabelecer uma parceria recíproca com um nó p qualquer², é necessário oferecer algum serviço s que esteja em $FT(p)$. Por isso, cada conjunto de serviços gera um resultado diferente, tanto nos aspectos de custo quanto nos de receita, que virá pela reciprocidade dos demais [Coêlho et al. 2008]. Do ponto de vista computacional, porém, é impraticável determinar a melhor escolha de serviços devido às possíveis combinações de serviços, mais o indeterminismo inerente ao ambiente. Com isso, a seleção de serviços precisa ser feita baseando-se em heurísticas [Coêlho et al. 2009]. Além disso, o ambiente que consideramos neste trabalho se compõe de múltiplos aspectos distintos, que podem descortinar cenários que favorecem de maneira diferenciada as estratégias que os nós implementam para a seleção de serviço. Coêlho et al. exploraram um desses aspectos: o orçamento médio disponível para os nós. Neste trabalho, fazemos uma extensão daqueles resultados, explorando outros aspectos do ambiente, medindo de que maneira eles impactam no desempenho das heurísticas.

As possíveis características ambientais se pode avaliar são muitas. Algumas, porém, merecem ser destacadas. A **quantidade de serviços consumidos pelos nós** em seus favores típicos, que chamaremos de *tamanho* do seu favor típico ($|FT|$), impacta na probabilidade encontrar serviços lucrativos, já que a chance de se encontrar um serviço que seja consumido por um nó p é diretamente proporcional a ($|FT(p)|$). Por outro lado, o **universo de serviços** afeta também a probabilidade de se encontrar serviços lucrativos, já que, havendo muitos nós e poucos serviços passíveis de serem oferecidos, espera-se maior probabilidade de se encontrar parcerias (serviços lucrativos), ao passo que no caso oposto espera-se que esta probabilidade diminua.

Acontece que, neste ambiente, a avaliação absoluta dessas heurísticas não é praticável computacionalmente porque é impossível compará-las com um seletor ótimo, já que ele não pode ser implementado por conta da complexidade e do indeterminismo do sistema. Desta forma, abordaremos o problema usando uma metodologia alternativa.

3. Metodologia de aferição de desempenho

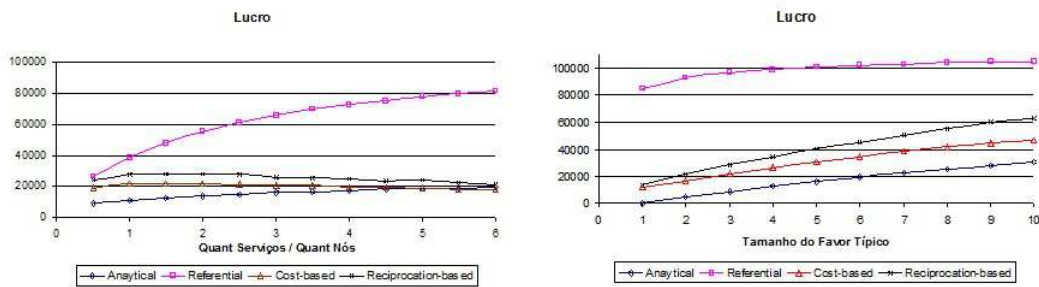
Coêlho e Brasileiro apresentam uma metodologia que mapeia o problema da seleção de serviços numa variação do *problema da mochila*, permitindo-se ajustar diferentes parâmetros do ambiente e assim gerar conjuntos de mochilas cujas características representam aspectos do ambiente e permitem avaliar o desempenho das heurísticas [Coêlho and Brasileiro 2009]. Este novo problema torna a seleção de serviços tratável metodologicamente, já que é possível se desenvolver um algoritmo de referência que busca a solução ótima, algo impraticável no problema da seleção de serviços original. Esta metodologia também permite se estimar analiticamente a performance de uma heurística “péssima” que seleciona os serviços de maneira aleatória.

²Considerando que p seja recíproco

Para comparar o desempenho das heurísticas em cenários variados, estabelecemos, usando a metodologia proposta, parâmetros que representam as seguintes características: uma comunidade de 50 nós com diferentes ($|FT|$), e diferentes tamanhos de universos de serviços, fixando o orçamento médio em 10, 0. Utilizando dados obtidos em experimentos reais usando características similares [Coelho et al. 2009], considerando que os custos são compostos de diferentes fatores, e de acordo com o Teorema do Limite Central (TLC), definimos os custos dos serviços a partir de uma distribuição $X \sim N(0.75, 0.25)$. Como a receita é oriunda da reciprocidade de diferentes nós, também de acordo com o TLC, definimo-na usando a média e o desvio padrão dos valores obtidos nos mesmos experimentos, como $X \sim N(10.34, 2.5)$.

4. Resultados

Implementamos, segundo a metodologia, dois experimentos: um testando vários tamanhos de universo de serviços e outro testando diferentes quantidades médias de serviços consumidas pelos nós ($|FT|$). Como parâmetros de limite, usamos o algoritmo de referência (chamado de *Referential*), que conhece a lucratividade real dos serviços e implementa uma solução ótima, e uma avaliação analítica (chamada *Analytical*) baseada nas variáveis definidas na seção 3, representando o desempenho de um seletor aleatório³. Nestes experimentos, usamos duas heurísticas: uma, *Cost-based*, toma os serviços em ordem crescente de custos, maximizando a quantidade de serviços oferecidos para aumentar a receita total recebida; a outra, *Reciprocation-based*, toma os serviço de acordo com a ordem esperada de receita, algo factível em ambientes reais, onde se pode considerar a reciprocidade dos nós e os serviços por eles consumidos.



(a) Lucro variando o universo de serviços

(b) Lucro variando o tamanho $|FT|$

Figura 1. Lucro: diferentes universos de serviços e tamanhos de favor típico

Para avaliar o desempenho das heurísticas em diferentes universo de serviços, consideramos a razão entre a quantidade de serviços e de nós (ς). Note que a probabilidade de um serviço não gerar receita (não ser consumido) é diretamente proporcional à relação entre a quantidade de serviços e de nós. Para cada valor de ς , geramos um total de 1000 mochilas, produzindo resultados com 95% de confiança, e erro máximo de 0, 15%. A Figura 1(a) mostra o lucro obtido pelas heurísticas para diferentes valores de ς . Se o total de serviços é grande em relação ao de nós (200 serviços, $\varsigma = 4$) a heurística *Cost-Based* tem desempenho próximo do limite inferior. Para valores altos ($\varsigma \geq 5, 5$) praticamente não faz mais diferença usar qualquer heurística, pois a chance de se encontrar

³Consideramos que heurísticas piores do que a seleção aleatória são inúteis

serviços lucrativos se torna muito exígua. Por outro lado, com valores menores (ς entre 1 e 3,5) a heurística *Reciprocation-based* se destaca, já que há uma probabilidade maior de se encontrar serviços lucrativos (ou parcerias lucrativas). Neste intervalo, a heurística *Cost-based* é significativamente melhor que *Analytical*, devido à quantidade de serviços oferecidos. De modo geral, porém, *Cost-based* não se destaca, pois a quantidade de serviços oferecidos a mais acaba sendo proporcionalmente muito pequena. Note que todas as heurísticas tiveram desempenho muito aquém do algoritmo de referência, porque a probabilidade de um serviço não gerar receita (não ser consumido) é diretamente proporcional à relação entre a quantidade de serviços e de nós e, nos experimentos, chega a 72%. Isto explica o desempenho de *reciprocation-based* chegar muito próximo de *Referential* quando $\varsigma = 1$.

O outro experimento avaliou o desempenho das mesmas heurísticas variando o tamanho médio do favor típico ($|FT|$). Note que quanto maior for a quantidade de serviços que um nó consome, (seu $|FT|$) maior será a probabilidade de ele conseguir um doador e, assim, estabelecer uma parceria lucrativa. Parametrizamos mochilas representando ambientes em que os nós consomem diferentes quantidades médias de serviços: o valor de $(|FT|)$ variou em (1, 2, ..., 10) e, para cada um desses valores, gerou-se um total de 1000 mochilas, também segundo a metodologia descrita na seção 3, produzindo resultados com erro máximo de 0,15% com 95% de confiança. Na Figura 1(b) podemos ver que o desempenho de todas as heurísticas cresce com o tamanho de $(|FT|)$, conforme esperado. A heurística *Reciprocation-based* supera a heurística *Cost-based* porque consegue encontrar mais serviços lucrativos, mas esta última consegue um desempenho melhor do que *Analytical* em função da quantidade de serviços oferecidos a mais. O desempenho de todas as heurísticas tende a se estabilizar quando o valor de $(|FT|)$ cresce muito, o que faz sentido: no limite, se todos os nós consumirem todos os serviços, então não há sentido se implementar nenhuma heurística de seleção, e a lucratividade dos nós se equiparará.

5. Trabalhos Relacionados

Há trabalhos que tratam de compartilhamento recíproco de recursos. A *Network of Favors* (NoF) [Andrade et al. 2007] é um esquema que dispensa confiança ou negociação prévia, além de promover relações mutuamente lucrativas entre os nós. Satsiou e Tassioulas [Satsiou and Tassioulas 2007] propõem um *framework* onde se troca recursos usando um *vetor de reputações*, que computa a reputação dos demais nós da comunidade. Todavia, nenhum dos trabalhos considera mecanismos para selecionar serviços com vista na lucratividade, a fim de aumentar o interesse de os nós permanecerem no sistema.

Coelho et. al. exploraram aspectos de lucratividade, relacionando-a com a seleção de serviços [Coelho et al. 2008], bem como evidenciaram a variação de lucratividade em função da estratégia de seleção implementada pelos nós [Coelho et al. 2009], além de mostrar uma metodologia para aferir o desempenho de heurísticas em diferentes condições ambientais [Coelho and Brasileiro 2009]. Tais trabalhos, porém, focaram o desempenho das heurísticas submetidas a cenários que diferiam entre si em um único aspecto: o orçamento médio disponível para os nós.

6. Conclusões

Apesar de se posar como uma alternativa para se prover recursos computacionais diversificados, com alto desempenho e de forma escalável, as grades computacionais multi-

serviço em geral, e as baseadas em reciprocidade em particular, estão intrinsecamente ligadas a um ambiente muito complexo em seus diversos aspectos envolvidos. Para explorá-los há que se considerar tanto as variáveis ambientais quanto as estratégias comportamentais dos nós. Neste trabalho focamos-nos em duas dessas variáveis e consideramos duas estratégias. Em linhas gerais, corroboramos os resultados de trabalhos similares: a estratégia que traz melhores resultados, é aquela que busca promover relações mutuamente vantajosas entre os nós, oferecendo os serviços que são consumidos pelos nós mais recíprocos sendo, portanto, os serviços que se espera serem os mais lucrativos.

Entre trabalhos futuros poderemos explorar outros aspectos ambientais, bem como considerar duas variações de *Reciprocation-based* que implementam mecanismos de otimização progressiva, cujo desempenho, mostrou-se ligeiramente melhor em trabalhos anteriores. Além disso, sendo um ambiente cooperativo e iterativo, onde os nós podem coletar e usar informações ao mesmo tempo em que podem trabalhar em associação com outros, é interessante estudar a aplicação de métodos de aprendizagem de máquina, bem como estudar a implementação distribuída e colaborativa de técnicas de otimização combinatória como espalhamento de partículas, algoritmos genéticos, colônia de formigas, redes neurais, etc..

Referências

- Andrade, N., Brasileiro, F., Cirne, W., and Mowbray, M. (2007). Automatic grid assembly by promoting collaboration in peer-to-peer grids. *J. Parallel Distrib. Comput.*, 67(8):957–966.
- Buyya, R. and Vazhkudai, S. (2000). Compute power market: Towards a market-oriented grid. *The First IEEE/ACM International Symposium on Cluster Computing and the Grid (CCGrid 2001)*.
- Coêlho, A. and Brasileiro, F. (2009). On the evaluation of services selection algorithms in multi-services p2p grids. In *Fourth IEEE International Workshop on Business-driven IT Management (BDIM'09)*, pages 52–60, Piscataway, NJ, USA. IEEE Press.
- Coêlho, A., Brasileiro, F., and Maciel, P. D. (2009). Using heuristics to improve service portfolio selection in p2p grids. In *IM'09: Proceedings of the 11th IFIP/IEEE international conference on Symposium on Integrated Network Management*, pages 438–444, Piscataway, NJ, USA. IEEE Press.
- Coêlho, A., Maciel Jr., P. D., Figueiredo, F. d., Candeia, D., and Brasileiro, F. (2008). On the impact of choice in multi-service p2p grids. In *Third IEEE International Workshop on Business-driven IT Management (BDIM'08)*, pages 98–101, Salvador, Bahia, Brazil.
- Mowbray, M., Brasileiro, F., Andrade, N., Santana, J., and Cirne, W. (2006). A reciprocation-based economy for multiple services in peer-to-peer grids. In *P2P'06: Proceedings of the Sixth IEEE International Conference on Peer-to-Peer Computing*, pages 193–202, Washington, DC, USA.
- Satsiou, A. and Tassiulas, L. (2007). A trust-based exchange framework for multiple services in p2p systems. In *P2P '07: Proceedings of the Seventh IEEE International Conference on Peer-to-Peer Computing*, pages 45–52, Washington, DC, USA. IEEE Computer Society.