

Comunicação Assíncrona e Adaptativa em Grades Computacionais

Rodrigo da Rosa Righi, Philippe O. A. Navaux

Universidade Federal do Rio Grande do Sul – UFRGS
{rrighi, navaux}@inf.ufrgs.br

Contextualização

Segundo Aumage [Aumage, 2002], os experimentos envolvendo a comunicação em grades computacionais (*grids*), potencialmente formada por recursos de rede heterogêneos, na sua maioria se concentram em duas categorias. A primeira delas é a metacomputação [Resch, 2000] que geralmente é organizada sobre sistemas padrão como o Globus (que fornece gerenciamento, segurança e alocação de recursos). Comumente essa área envolve a comunicação entre os computadores através de redes tradicionais usando protocolo TCP/IP, que impõem sobrecargas de *software* para garantir confiabilidade na transmissão de dados. Além disso, essa categoria pode envolver vários domínios administrativos, onde as características de segurança e tolerância a falhas são desejáveis. A segunda categoria envolve a interconexão de agregados (*clusters*), formando a arquitetura de multi-agregado [Barreto et al., 2000]. Essa arquitetura representa um ambiente mais leve, onde não são requeridos tratamentos como aqueles necessários em grades de escala mundial [Aumage, 2002]. Nela, por exemplo, cada agregado pode usar técnicas de troca de mensagens otimizadas internamente ao seu domínio sem se preocupar com segurança.

Um tópico importante de pesquisa em grades envolve a linguagem e a interface de programação pela qual a sua infraestrutura será disponibilizada. Observa-se a expansão de Java, assim como a utilização da interface já conhecida para agregados MPI para trabalhar sobre as grades. Existem alguns sistemas Java que, diferente de RMI, conseguem trabalhar de forma assíncrona e são mais flexíveis quanto a comunicação (cobrindo também redes de sistema). Como exemplo, pode-se citar iniciativas como o ProActive, Ibis, NetIbis e JavaSymphony. Além desses sistemas, também existem implementações de MPI para grades que tratam a heterogeneidade de rede, como o MPICH/Madeleine III e o PACX-MPI. Ambos tratam das necessidades de multi-agregados. Em uma escala maior há o MPICH-G2, que usa as características do Globus para gerir a infraestrutura de comunicação. Fazendo uma pesquisa sobre esses sistemas, pode-se inferir que eles não cobrem a comunicação transparente e eficiente para ambas categorias de grades citadas acima.

Objetivos e Metodologia da Pesquisa

Baseado no contexto apresentado, pretende-se desenvolver um modelo para comunicação em grades que consiga abranger os multi-agregados e a metacomputação, provendo comunicação adaptativa e assíncrona. A adaptabilidade remete a idéia da exploração dinâmica, transparente e eficiente das vantagens dos recursos disponíveis (*hardware* e *software*) no sistema distribuído. O assincronismo é importante para aumentar a eficiência da aplicação, uma vez que o transmissor não fica bloqueado nas interações pela rede. O

seu emprego contribui para esconder a latência da rede e é especialmente interessante nas bordas do sistema. Em última análise, objetiva-se um modelo para grades o qual servirá de base para o desenvolvimento de um *middleware* que visa desempenho, buscando minimizar o tempo de resposta pelos resultados de uma aplicação guardando a precisão desejada. Em adição, o *middleware* será útil para aplicações dinâmicas, onde processos podem ser alocados em tempo de execução, ou mesmo migrados entre plataformas.

Estão sendo pesquisados sistemas em duas frentes, Java e MPI, envolvendo os nomes relacionados na seção anterior. Além desses, pesquisar-se-á os sistemas MPI na sua norma 2 das distribuições LAM e MPICH buscando observar o que cada um tem a oferecer para a computação em grade. Para cada um dos sistemas, são abordadas as suas capacidades de comunicação no que tange os seguintes itens: tolerância a falhas, comunicação assíncrona, segurança, confiabilidade no transporte de dados, tratamento da heterogeneidade em nível de rede, suporte a redes de sistema e propriedades de adaptação. Para cada um desses itens é observado como ele é tratado e se pode ser desabilitado da estrutura do sistema, ou seja, se ele é facultativo na infraestrutura de comunicação. Levando em consideração os sistemas Java, a pesquisa até o momento constatou que o ProActive se destaca pelo assincronismo em alto nível (usando mecanismos de espera pela necessidade e objetos futuros) e pela configuração de tolerância a falhas e lançamento de processos em arquivos XML. O Ibis possui uma interface IPL (*Ibis Portability Layer*) que separa sistemas de alto nível (RMI e Satin) das questões de comunicação e serviços de grade. Já as principais contribuições do JavaSymphony são a sua noção de hierarquia de computadores e o tratamento de eventos. Além das características, verificar-se-á a arquitetura de cada sistema a fim de identificar idéias de adaptação em tempo de execução. Seguindo essa área, será pesquisado o arcabouço Fractal [ObjectWeb Open source Middleware, 2005], que possibilita desenvolver programas e componentes separando conceitos relacionados com a implementação da idéia da modelagem e conceitos funcionais (segurança e configuração). Por fim, também está sendo estudada a linha de pesquisa que trata da computação autônoma [Kephart and Chess, 2003]. Ela trata da adaptabilidade automática do ambiente seguindo informações locais e dos outros nós do sistema distribuído.

Referências

- [Aumage, 2002] Aumage, O. (2002). Heterogeneous multi-cluster networking with the madeleine iii communication library. In *Parallel and Distributed Processing Symposium*, pages 85–96. Los Alamitos USA, IEEE Computer Society.
- [Barreto et al., 2000] Barreto, M. E., Ávila, R. B., and Navaux, P. O. A. (2000). The MultiCluster model to the integrated use of multiple workstation clusters. In *Proceedings of the Parallel and Distributed Processing*, pages 71–80. Berlin: Springer-Verlag.
- [Kephart and Chess, 2003] Kephart, J. O. and Chess, D. M. (2003). The vision of autonomic computing. *Computer*, 36(1):41–50.
- [ObjectWeb Open source Middleware, 2005] ObjectWeb Open source Middleware (2005). The fractal project. <http://fractal.objectweb.org/index.html>. Visitado em agosto de 2005.
- [Resch, 2000] Resch, M. M. (2000). Metacomputing in a high performance computing center. In *ICPP '00: Proceedings of the 2000 International Workshop on Parallel Processing*, page 165, Washington, DC, USA. IEEE Computer Society.