

Desafios para a computação pervasiva no futuro cenário tecnológico

Luigi Carro e Flávio Rech Wagner
PPGC - UFRGS

O problema

No cenário mundial há uma enorme demanda por computação portátil, ubíqua e de alto poder computacional. Exemplos clássicos são as evoluções das áreas de entretenimento e telecomunicações, onde celulares e set-top-boxes demandam cada vez mais operações por segundo e enormes quantidades de memória, mas por preços cada vez menores. Para equipamentos de telecomunicações 4G, por exemplo, prevê-se a necessidade de Tflops de capacidade computacional e de Tbytes de memória (<http://public.itrs.net/HomeStart.htm>). Começam a aparecer também exemplos significativos de novos mercados para a computação portátil, principalmente na área da saúde, pela possibilidade de monitoração remota e uso de grandes bancos de dados para populações com pouco acesso a recursos de grandes centros. Brasil, Índia e China estão neste modelo.

A preocupação básica, no entanto, é que a lei de Moore está perdendo velocidade, pelos simples limites físicos do átomo. Já existem diversas pesquisas para o desenvolvimento de novas tecnologias que substituirão ou trabalharão complementarmente ao silício, como Finfets, SET ou outros dispositivos (<http://public.itrs.net/HomeStart.htm>). Todas estas novas tecnologias, porém, sofrem de um mesmo mal: embora possuam enorme capacidade de integração, permitindo muito paralelismo, são bem mais lentas do que as tecnologias atuais. Além disto, todas elas terão confiabilidade baixa e alta sensibilidade a defeitos e ruídos, que hoje ainda não são dominantes no projeto digital. Finalmente, a maioria dos equipamentos será movida a bateria, e a capacidade destas não tem aumentado de maneira exponencial, provocando um gargalo na energia que pode ser consumida.

Neste cenário, e com a demanda pelos novos produtos, as técnicas de projeto de hardware e software precisarão ser alteradas. O modelo atual, em que o projetista de software utiliza uma abstração de alto nível do hardware, terá de ser modificado, pois deve-se tirar proveito da computação massivamente paralela mas sem descuidar do consumo de energia e da confiabilidade, hoje tipicamente considerados apenas no projeto nos baixos níveis de abstração. Isto parece indicar a necessidade do abandono do projeto do software sobre uma abstração do hardware. No entanto, não se pode perder de vista o porte do software que se quer projetar, devido à complexidade computacional crescente dos problemas em mãos, por exemplo relacionados aos diversos tipos de tratamento de imagens (um telefone celular atual já contém milhares de linhas de código). Isto parece sinalizar a direção aparentemente oposta, que é o projeto em níveis ainda mais altos de abstração e utilizando doses crescentes de automação, indispensáveis num cenário de competitividade exacerbada com um *time-to-market* bastante exíguo. Como se poderá conciliar, no entanto, a automação e a abstração, requeridas para lidar com a complexidade e os requisitos de *time-to-market* do projeto de software, com o projeto conjunto e otimizado do software e do hardware, necessário para a obtenção de eficiência energética e de compensação da baixa confiabilidade das novas tecnologias?

O desafio tendo em vista o cenário

Neste cenário tecnológico e de mercado, o esforço de quem projetará os novos sistemas pervasivos pode ser modelado como um triângulo, onde os três vértices correspondem à tecnologia falha com componentes lentos, ao paralelismo massivo para lidar com os problemas tecnológicos de maneira eficaz, e à abstração adequada para o software, envolvendo sua geração automática para exploração de um enorme espaço de projeto.

O primeiro vértice do triângulo concerne o estabelecimento de um modelo de computação que, em várias camadas, e sempre com atenção à potência consumida, use os recursos de uma tecnologia pouco confiável para implementação de programas confiáveis. O modelo de falhas para as tecnologias futuras será muito diferente daquele usado nos produtos atuais. Além das falhas naturais do processo de fabricação, falhas transitórias derivadas de ruído eletromagnético começarão a ser muito importantes nas tecnologias formadas por apenas poucas camadas de átomos. Cada modelo de falhas (permanente ou transitória) exigirá estratégias diferentes de solução. As falhas permanentes provavelmente demandarão redundância sob forma de reconfiguração dinâmica, enquanto que falhas transitórias exigirão mecanismos mais sofisticados em várias camadas do hardware ao software, incluindo as pilhas de abstrações que terão de ser construídas (sistema operacional, middleware e APIs).

Pela pouca confiabilidade da tecnologia, e ao mesmo tempo pelos recursos que existirão à disposição, o outro vértice concerne o estabelecimento de um modelo de computação que use os recursos de paralelismo massivo, mas sem obrigar o usuário a programar de modo muito diferente do atual. Isto não deve ser confundido com um simples retorno à pesquisa de programação paralela, pois os parâmetros de projeto irão se alterar radicalmente: as máquinas

disponíveis para o processamento pervasivo serão mais lentas e muito menos confiáveis, e com restrições de memória devido à potência consumida. Além disto, os sistemas serão baseados em processadores heterogêneos, otimizados para explorar paralelismo de diferentes e simultâneas maneiras (ILP, VLIW, SMT, dataflow através de arranjos reconfiguráveis, etc.), segundo necessidades de diferentes domínios de aplicação, como já se observa no processador Cell da IBM. Finalmente, as diferentes formas de paralelismo terão de ser escondidas do programador, primeiro para que se mantenham os modelos atuais para os quais todos estão treinados, e sobretudo porque o uso explícito e simultâneo de recursos paralelos e heterogêneos, medidos aos milhares, é impossível de ser dominado apenas pela mente humana.

O terceiro vértice do triângulo corresponde à complexidade do software. O desafio aqui está muito além da montagem de mais abstrações como vem sendo tradicionalmente feito na comunidade (orientação a objetos, UML, etc). O desafio é que se devem manter as abstrações para lidar com a complexidade funcional, mas ao mesmo tempo deve-se lidar com hardware pouco confiável, mas abundante, e com sérias limitações de consumo de energia. Com a perda de velocidade da lei de Moore, o uso de múltiplos processadores heterogêneos integrados, cada um fazendo uma tarefa específica, será fundamental para ter-se os produtos com a eficiência energética que o mercado demanda. Isto já pode ser visto nos telefones celulares, compostos de dois processadores diferentes. Parece muito difícil imaginar, num cenário futuro, a utilização de uma plataforma de propósitos gerais como o PC para a solução de todos os problemas computacionais demandados pela sociedade. A busca pela eficiência energética tem levado os fabricantes de sistemas a desenvolver plataformas multiprocessadas de hardware e software bastante distintas para cada domínio de aplicação (p.ex. OMAP para celulares da TI e Cell da IBM para a área de jogos 3D). Além disto, estes múltiplos processadores estarão conectados por redes comutadas dedicadas integradas dentro da pastilha (p.ex. rede Aetherial da Philips). Contudo, o desenvolvimento de software para estas plataformas multi-processadas é extremamente complexo, e muito do software ainda vem sendo feito em assembler, claramente contra a tendência de automação. Novas ferramentas de geração de código que suportem o desenvolvimento eficiente destas futuras plataformas serão fundamentais para seu sucesso comercial, com baixo tempo de projeto.

Cenário brasileiro

A fabricação de componentes nestas novas tecnologias exige hoje investimentos medidos em bilhões de dólares, e portanto haverá uma tendência à concentração destas fábricas em poucos centros. Assim, a fabricação das plataformas de hardware continuará fora do país, por razões de custo e volume. Contudo, o projeto e especialmente a programação eficiente destes componentes poderão ser tarefas mundiais. Quem chegar primeiro terá a chance de estabelecer padrões e ser copiado. Consequentemente, é possível ao país desenvolver tecnologia para suporte a este cenário futuro, considerando a competência existente em nossa comunidade de Computação, nas diferentes áreas envolvidas.

Recursos humanos

Um ingrediente essencial da competitividade futura será a disponibilidade de recursos humanos treinados para o projeto de sistemas neste novo cenário tecnológico e de mercado. O projeto independente das diversas camadas de software e hardware, tal como conhecemos hoje, está fadado ao rápido desaparecimento. Serão exigidas equipes e indivíduos com competências múltiplas, em temas como desenvolvimento de software aplicativo, sistemas distribuídos e paralelos, redes de processadores, sistemas operacionais, arquitetura de computadores, tolerância a falhas em vários níveis, tempo real e concepção de circuitos integrados. Pode-se prever o desaparecimento de muitas das fronteiras atuais entre estas disciplinas. As universidades brasileiras precisarão responder rapidamente a esta nova demanda em seus programas de formação.

Conclusão

O projeto de novas aplicações portáteis e pervasivas, diante de um cenário tecnológico bastante diverso do atual, passará a exigir uma visão fortemente integradora de muitas disciplinas hoje razoavelmente independentes e estanques, já mencionadas acima. Para desenvolvimento deste sinergismo, novas metodologias de projeto devem ser pesquisadas, com conseqüente impacto sobre as estruturas e competências dos grupos de pesquisa e currículos do país.

Biografias

Luigi Carro é doutor em Ciência da Computação (UFRGS, 1996), com pós-doutorado na UCSD, professor adjunto da UFRGS e bolsista I-D do CNPq. Já orientou 11 dissertações de mestrado e 3 teses de doutorado. Publicou dois livros didáticos e mais de 180 artigos completos em veículos qualificados, sendo 135 de âmbito internacional.

Flávio R. Wagner é doutor em Computação (Kaiserslautern, 1983), com pós-doutorado no INPG (França), professor titular da UFRGS e bolsista I-C do CNPq. Foi presidente da SBC e é membro dos comitês assessores da CAPES e CNPq. Já orientou 22 dissertações de mestrado e 6 teses de doutorado. Publicou 3 livros didáticos e mais de 110 artigos completos em veículos qualificados, sendo 63 de âmbito internacional.