

Entrega 3

Clouds: Análise de Artigos

Edgard Quirino dos Santos Neto
201000285

Eduardo Filipe Amaral Soares
200906159

Gonçalo Manuel Duarte Lourenço
200802456

Faculdade de Ciências da Universidade do Porto
Computação Grid 12/13
Maio de 2013

1 Introdução

Este curto relatório tem como objetivo apresentar curtas avaliações dos artigos utilizados para criar a apresentação sobre o tema *Clouds*, desenvolvido no âmbito da 3ª entrega da cadeira de Computação Grid. Cada secção aborda um artigo diferente e cada sub-secção avalia uma componente diferente. O 1º artigo foi analisado por Edgard Quirino, o 2º por Eduardo Soares e o último por Gonçalo Lourenço.

2 *A Quantitative Analysis of High Performance Computing with Amazon's EC2 Infrastructure: The Death of the Local Cluster?*

2.1 Sumário

O artigo em questão [1] compara a performance de computação da infraestrutura Amazon *Elastic Computing Cloud* (EC2) com a duma cluster de escala reduzida de laboratório. A introdução de recursos computacionais virtualizados *on-demand* a preços razoáveis, particularmente da Amazon EC2, fez com que vários utilizadores questionassem como estes recursos devem ser disponibilizados no futuro. Este tipo de oferta tem tido um impacto significativo no mercado e na comunidade de Tecnologias de Informação e disponibiliza alternativas atrativas e razoáveis a infraestruturas locais.

O uso deste tipo de infraestrutura para fins científicos tem sido posto em causa devido ao seu uso de virtualização e estrangulamento seletivo de tráfego de rede (*network shaping*), fatores que podem ter um impacto significativo na *performance*. Este artigo insere-se num panorama

onde a maioria das comparações realizadas prendem-se a *clusters* de *High-performance Computing* (HPC) que recorrem a ligações de alta velocidade entre os diferentes nós constituintes.

2.2 Critério de escolha

A escolha deste artigo prende-se única e exclusivamente com a recomendação feita pela Exm.^a Sr.^a Dr.^a Prof.^a no enunciado do trabalho.

2.3 Avaliação

- **Problema abordado:** Comparação da *performance* de computação entre a infraestrutura Amazon EC2 e uma cluster de tamanho reduzido, tipicamente encontrado em laboratórios;
- **Bibliografia:** A bibliografia é abrangente, contando com várias referências não só a artigos e projetos de carácter académico, mas também a serviços *IT Business* para além do próprio abordado pelo artigo 5, como se pode observar na referência 6. Um ponto a ser ressaltado é o facto de todas as hiperligações a páginas de projetos na Internet, excetuando a da referência 4, continuam disponíveis e têm sido constantemente atualizadas, mostrando que os projetos citados pelo artigo ainda têm continuidade;
- **Organização:** A organização do artigo encontra-se bem pensada e o texto é claro. O extrato está muito bem escrito, o que facilita a compreensão do artigo. O uso de gráficos e figuras contém muita informação, mas o posicionamento e a qualidade das imagens deixa muito a desejar;
- **Relevância:** O artigo está claramente desatualizado em termos de *hardware* descrito, o que se torna claro na secção "*Evaluation Setup and Evaluation*", onde as memórias tipo DDR e DDR2 são classificadas de arquiteturas novas de memória, quando, atualmente, as arquiteturas mais recentes são DDR3 e DDR4 (que ainda não se encontra comercialmente disponível), chegando a mais de 1600MHz. Isto é compreensível, visto que o artigo foi escrito em 2009. Porém, os conceitos e a informação que fornece sobre o modelo de computação *Cloud* e *clouds* científicas continuam atualizados;
- **Metodologia/Experimentação:** As experiências não podem ser replicadas, visto que recorreram a um serviço pago, o Amazon EC2, e o *cluster* usado no momento da elaboração do artigo já não existe. Quanto ao estado-da-arte da área, as experiências em questão apresentam resultados de alta qualidade, visto que demonstram as vantagens económicas e de performance do uso de serviços como o supramencionado através de testes controlados.

3 *Grid and Cloud Computing: Opportunities for Integration with the Next Generation Network*

3.1 Sumário

O artigo em questão [2] estuda a possibilidade de uso de *grids* ou *clouds* nas *Next Generation Network* (NGN). As NGN são redes de próxima geração dos operadores de comunicação como a AT&T, Telefónica, TMN. Redes que têm de dar resposta às necessidades desses mesmos operadores para fornecer diversos serviços aos seus clientes. Alguns dos serviços são, por exemplo, comunicação de voz ou acesso a dados. Dadas as necessidades, estas redes têm que ser escaláveis e dinâmicas, de forma a conseguir dar resposta a qualquer quantidade de pedidos. Portanto, este tipo de redes têm necessidades e obstáculos comuns com os encontrados nas *grids*, com o ponto extra de uma forte necessidade de interoperabilidade.

Devido às NGN tentarem resolver problemas comuns com os das *grids* e das *clouds*, é uma boa ideia olhar para elas e tentar partilhar recursos e soluções para problemas comuns. Em particular, a utilização de *clouds* torna-se impossível devido à falta de interoperabilidade entre elas. Tipicamente, uma solução desenvolvida para uma *cloud* não corre noutra, o que se torna um problema. Criar uma solução comum ou pelo menos uma solução interoperável é objetivo nas NGN.

É discutido no artigo a possibilidade do uso de *grids* para a criação da NGN, mas devido à arquitetura muito própria das NGN, não seria fácil consegui-lo. Essa arquitetura passa por a divisão em duas componentes, uma de transporte e uma de serviços base às aplicações. A camada de transporte é responsável pela ligação a várias redes, como por exemplo os terminais móveis e a Internet. A camada de serviços fornece um conjunto base às aplicações, esses serviços são por exemplo identificação do utilizador, serviços de pagamento, serviços dedicados a *IPTV*, entre outros.

Um outro problema está contido naquilo que é um ponto forte nas redes de comunicação atuais, a interoperabilidade, isto é, a capacidade de dois sistemas distintos interagirem sem problemas entre si. Um exemplo disto é a possibilidade de efetuar um telefonema entre dois terminais de dois operadores distintos em que cada um usa o seu próprio equipamento, só que, do ponto de vista do uso do serviço, é totalmente transparente e de aparência homogénea.

Para resolver este problema nas *grids* e assim ajudar a aproximar as NGN às *grids*, o artigo discute ainda os problemas de interoperabilidade e metodologia de testes aos mesmos. O problema principal passa por a falta de standards e, no caso da sua existência, a forma geral como estão definidos deixando à interpretação deles possibilidades de diferenças, causando implementações incompatíveis. Um exemplo disso é a uso dos certificados *X.509*, um *standard* de enorme sucesso em todas as *grids*, mas cujo uso nas NGN é realizado de forma diferente, tornando assim incompatível.

3.2 Critério de escolha

A escolha deste artigo prende-se única e exclusivamente com a recomendação feita pela Exm.^a Sr.^a Dr.^a Prof.^a no enunciado do trabalho.

3.3 Avaliação

- **Problema abordado:** Possibilidade de integração nas *grids* ou *clouds* nas NGN;
- **Bibliografia:** A bibliografia do artigo é bastante variada contendo fontes de relevo como o *Open Grid Forum*, *Globus Alliance* e a ETSI;
- **Organização:** O artigo encontra-se organizado cinco secções, começando com uma introdução em que define NGN e explica as vantagens possíveis no uso de *grids* ou *clouds* para a sua construção. A segunda secção mostra os pontos comuns que as NGN têm com as *grids* e *clouds*, de seguida explica quais os problemas de interoperabilidade e apontar o caminho para conseguir este objetivo. Por fim mostra ainda as metodologias usadas para testar a interoperabilidade entre componentes;
- **Relevância:** Este artigo mostra um ponto importante das *grids*, a falta de interoperabilidade e aponta alguns dos problemas que causam isto e qual o caminho para os resolver. Dado o artigo ser de 2008 pode estar já um pouco desatualizado mas com certeza que dada a quantidade de *software* diferente e a variedade de componentes com certeza ainda hoje este ponto é importante e ainda estão a ser implementados ou melhorados standards para alcançar este objetivo;
- **Metodologia/Experimentação:** O artigo resume alguns standards e o suporte deles em vários *softwares* populares de *grids* e como tal pode ser testado, mas nenhum deles foi experimentado.

4 *Cluster, Grid and Cloud Computing: A Detailed Comparison*

4.1 Sumário

O artigo em questão [3] aborda três modelos diferentes de computação, nomeadamente *Cluster*, *Grid* e *Cloud*, e expõe a definição de cada um, a origem deles, os desafios que os utilizadores enfrentam, exemplos de utilização real de cada um e ainda uma comparação detalhada do que constitui cada modelo.

4.2 Critério de escolha

Senti uma forte necessidade de encontrar um artigo que me ajudasse a introduzir o tema, tanto numa perspectiva histórica como numa perspectiva de comparação com os outros modelos de computação em uso atualmente. Para tal, fui ao motor de pesquisa **IEEE Xplore** e pesquisei recorrendo às keywords *grid*, *cluster* e *cloud*.

4.3 Avaliação

- **Problema abordado:** O problema abordado é o de comparação detalhada entre os modelos de computação *Cluster*, *Grid* e *Cloud*;

- **Bibliografia:** A bibliografia é muito variada, tendo referências tanto a artigos fundamentais para cada um dos modelos, como também a artigos publicados apenas um ano antes deste. Nem todas as referências a páginas de Internet ainda se encontram válidas. De todos as hiperligações testadas, as referências 12, 17, 19, 20 e 28 encontram-se atualmente indisponíveis. Referência 26 contém um erro de compilação de \LaTeX , mas a hiperligação em si está correta;
- **Organização:** O artigo encontra-se dividido em 7 secções. A introdução dá uma perspetiva histórica geral, assim como uma definição introdutória de cada modelo. A segunda secção, " *Three Computing Models*", dá uma definição detalhada de cada um dos modelos. A terceira secção, " *Challenges in Cluster, Grid and Cloud Computing*", descreve vários desafios enfrentados pelos utilizadores de cada modelo. A quarta secção, " *Comparison of Cluster, Grid and Cloud Computing*", apresenta uma tabela que detalha a implementação de vários aspetos em cada um dos modelos, permitindo uma análise rápida de cada. A quinta, " *Projects and Applications in Cluster, Grid and Cloud Computing*", apresenta vários casos reais de projetos e aplicações de cada modelo. A sexta, " *Tools and Simulation Environment*", enuncia várias ferramentas úteis para cada modelo. A conclusão enuncia o que foi detalhado em cada secção. Não detetei nenhum erro ortográfico e a escrita é de fácil leitura. O único conteúdo gráfico é a tabela da quarta secção, que se revela essencial para uma análise rápida do conteúdo abordado nessa parte do texto;
- **Relevância:** O problema continua a ser muito relevante e ter um artigo que consiga apresentar as diferenças-chave entre estes diferentes modelos de forma sucinta, mas detalhada, é extremamente útil e essencial para a sua resolução;
- **Metodologia/Experimentação:** O tema abordado por este artigo pode ser facilmente reproduzido. Os resultados são de boa qualidade.

Referências

- [1] Zach Hill and Marty Humphrey. A Quantitative Analysis of High Performance Computing with Amazon's EC2 Infrastructure: The Death of the Local Cluster? *10th IEEE/ACM International Conference on Grid Computing*, pages 26–33, Outubro 2009.
- [2] Julian Gallop Jens Grabowski Tatiana Kovacikova Stephan Schulz Thomas Rings, Geoff Caryer and Ian Stokes-Rees. Grid and Cloud Computing: Opportunities for Integration with the Next Generation Network. *Journal of Grid Computing*, 7(3):375–393, Agosto 2009.
- [3] Naidila Sadashiv and S. M. Dilip Kumar. Cluster, grid and cloud computing: A detailed comparison. *6th International Conference on Computer Science & Education (ICCSE)*, pages 477–482, Agosto 2011.