

UNIVERSIDADE FEDERAL DE CAMPINA GRANDE - UFCG
CENTRO DE ENGENHARIA ELÉTRICA E INFORMÁTICA - CEEI
DEPARTAMENTO DE SISTEMAS E COMPUTAÇÃO - DSC

PLANO DE ESTÁGIO

Avaliação qualitativa das principais tecnologias atuais
de grades computacionais oportunistas

Ricardo Araújo Santos
Curso de Bacharelado em Ciência da Computação

Campina Grande, Março de 2009

Sumário

| | | |
|-----------|---|-----------|
| 1 | Informações Pessoais | 2 |
| 2 | Ambiente de Estágio | 3 |
| 2.1 | Estrutura Física | 3 |
| 3 | Supervisão (Acadêmica e Técnica) | 4 |
| 4 | Resumo do Problema | 5 |
| 5 | Objetivos | 5 |
| 6 | Metodologia | 6 |
| 7 | Atividades Planejadas | 7 |
| 8 | Resultados Esperados | 8 |
| 9 | Cronograma | 8 |
| 10 | Aprovação | 10 |

1 Informações Pessoais

Nome: Ricardo Araújo Santos

Matrícula: 20511120

Endereço Residencial: Rua Olegário de Azevedo, 268 / ap 01 - São José - Campina Grande - PB

Endereço Profissional: Av. Aprígio Veloso, 882 Bodocongó, Bloco CO
58109-970, Campina Grande, PB

Fone: +55 (83) 3310 1640 , (83) 8875 7392

Fax: +55 (83) 3310 1498

E-mail: ricardo [AT] lsd.ufcg.edu.br

2 Ambiente de Estágio

O estágio será desenvolvido no Laboratório de Sistemas Distribuídos (LSD) do Departamento de Sistemas e Computação (DSC) da Universidade Federal de Campina Grande.

O laboratório foi criado em 1996, como forma de aglutinar pesquisadores e alunos do DSC e de outros departamentos em torno de projetos na área de Sistemas Distribuídos. Atualmente o LSD é coordenado pelo professor Francisco Brasileiro. As pesquisas do LSD estão concentradas em Grades Computacionais, Sistemas Peer-to-Peer, Cloud Computing, Tolerância a Falhas, Desenvolvimento de Software Distribuído e Aplicações Industriais.

2.1 Estrutura Física

O LSD está instalado num prédio com $550m^2$ de área e conta com a colaboração de dezenas de alunos desenvolvendo trabalhos de doutorado, mestrado e iniciação científica, além de vários pesquisadores. Os diferentes projetos em execução estão distribuídos em 8 (oito) salas climatizadas, cada uma contanto com quadro branco, pincéis e um acervo bibliográfico de diversos temas em computação e engenharia. Cada pessoa possui um posto de trabalho individual, com máquinas conectadas à Internet via POP-PB da Rede Nacional de Pesquisa (RNP).

Endereço: Universidade Federal de Campina Grande
Departamento de Sistemas e Computação
Laboratório de Sistemas Distribuídos
Av. Aprígio Veloso, 882 - Bloco CO
Bodocongó, CEP 58109-970 Campina Grande - PB, Brasil
Fone: +55 83 3310 1365
Fax: +55 83 3310 1498

3 Supervisão (Acadêmica e Técnica)

A supervisão acadêmica será efetuada pela professora Raquel Vigolvino Lopes, pesquisadora do Laboratório de Sistemas Distribuídos (LSD) e professora do DSC/UFCG. A supervisão técnica será efetuada pelo aluno de mestrado Marcus Williams Aquino de Carvalho.

Dados do supervisor acadêmico

Nome: Raquel Vigolvino Lopes

Endereço Profissional: Av. Aprígio Veloso, 882 Bodocongó, Bloco CO
58109-970, Campina Grande, PB

Fone: +55 (83) 3310 1643

Fax: +55 (83) 3310 1498

E-mail: raquel [AT] dsc.ufcg.edu.br

Dados do supervisor técnico

Nome: Marcus Williams Aquino de Carvalho

Endereço Profissional: Av. Aprígio Veloso, 882 Bodocongó, Bloco CO
58109-970, Campina Grande, PB

Fone: +55 (83) 3310 1640

E-mail: marcuswac [AT] lsd.ufcg.edu.br

4 Resumo do Problema

A pesquisa científica em algumas áreas tem demandado cada vez mais poder computacional, seja para a realização de simulações como para a execução de experimentos. Uma saída natural é a aquisição de supercomputadores ou máquinas dedicadas (clusters) exclusivamente à computação paralela, o que nem sempre é possível dado seu elevado custo. As grades computacionais (BERMAN; FOX; HEY, 2003; KESSELMAN; FOSTER, 1998) surgiram com a idéia de resolver tais problemas computacionais de maneira eficiente e com baixo custo, através do compartilhamento de recursos.

Uma forma comum de compartilhamento de recursos faz uso de poder computacional ocioso, formando o que se chama no contexto desse trabalho de grades computacionais oportunistas. Exemplos de grades oportunistas são: Condor (THAIN; TANNENBAUM; LIVNY, 2005), BOINC (ANDERSON, 2004), XtremWeb (FEDAK et al., 2001) e OurGrid (CIRNE et al., 2006).

Muito tem sido publicado em conferências e jornais sobre essas ferramentas. No entanto, não são conhecidos estudos empíricos que visem uma caracterização do uso das ferramentas disponibilizadas, impossibilitando identificar o estado-da-prática desta área. Acredita-se que um estudo prático, em contrapartida aos estudos teóricos conhecidos, é de grande importância tanto para identificação de lacunas quanto para seleção do melhor serviço em diferentes contextos.

5 Objetivos

O principal objetivo desse trabalho é comparar de forma qualitativa o uso das principais tecnologias de grades desktops ou oportunistas da atualidade.

Como objetivos específicos espera-se:

- Realizar um levantamento bibliográfico sobre as grades desktop ou oportunistas mais usadas atualmente;
- Levantar métricas que possam ser usadas nesse estudo comparativo qualitativo;
- Comparar qualitativamente, com base nas métricas definidas, as grades identificadas no início do estágio.

6 Metodologia

Uma abordagem qualitativa é mais adequada uma vez que existem características próprias dessas grades que não podem ser medidas, somente observadas. Inicialmente, pretende-se construir uma base de informações sobre as ferramentas a serem estudadas, tais como artigos e documentação das ferramentas. No entanto, a base de nossa metodologia é a prática, de forma que, serão determinadas algumas métricas com relação às características das ferramentas escolhidas e a partir da análise dos dados coletados será possível concluir algo sobre o estado da prática da área. As métricas levarão em conta as diversas fases do uso das tecnologias como a instalação e o uso da grade em si, como também custos de manutenção e preparação de aplicações para execução na grade, por exemplo.

Dessa forma, pode se dizer que a metodologia deste trabalho se divide nas seguintes etapas:

- Formulação do problema e dos procedimentos da pesquisa:
 - Definição do problema;
 - Identificação das ferramentas a serem analisadas;
- Planejamento das métricas avaliadas:
 - Determinar o que deve ser estudado nas ferramentas;
 - Identificar aspectos comuns nas ferramentas;
- Determinação os cenários de experimentação:
 - Determinar a seleção de amostras;
 - Verificar e validar a seleção de amostras no tocante a representar a população;
- Realização os experimentos e coleta de dados;
- Desenvolvimento de um plano de análise da amostra:
 - Filtrar dados relevantes;
- Processar e analisar os dados obtidos.

7 Atividades Planejadas

As seguintes atividades serão realizadas:

Levantamento bibliográfico: Apesar da natureza prática da avaliação qualitativa, se faz necessário uma investigação bibliográfica a fim de se identificar as mais importantes grades computacionais de desktop usadas atualmente;

Elaboração das métricas de interesse para a avaliação: Essa atividade visa identificar os pontos nos quais as tecnologias serão comparadas e contrastadas;

Projeto de experimentos: Planejar os experimentos é necessário para se cobrir os detalhes de interesse das métricas identificadas;

Preparação do ambiente experimental: A avaliação prática requere instalar cada uma das ferramentas identificadas;

Realização dos experimentos: Com a finalidade de coletar observar as métricas e coletar dados sobre elas, quando possível;

Apresentação de seminário com resultados: Para divulgar na universidade para laboratórios interessados a pesquisa realizada e os dados coletados;

Escrita de artigo para possível publicação: Para divulgar para a comunidade científica;

Escrita do relatório de estágio: Documentar a pesquisa realizada;

Preparação da apresentação de defesa do estágio: Preparar apresentação da pesquisa e conclusões obtidas.

A estimativa de horas a se realizar as atividades pode ser vista na tabela 1.

| | Atividades | Horas estimadas |
|----|---|-----------------|
| A1 | Levantamento bibliográfico | 30 |
| A2 | Elaboração das métricas de interesse para a avaliação | 30 |
| A3 | Projeto de experimentos a serem realizados usando as grades de desktops | 20 |
| A4 | Preparação do ambiente experimental | 20 |
| A5 | Realização de experimentos e análise de resultados | 100 |
| A6 | Apresentação de seminário com resultados | 20 |
| A7 | Escrita de artigo para possível publicação em uma conferência nacional | 50 |
| A8 | Escrita do relatório de estágio | 50 |
| A9 | Preparação da apresentação de defesa do estágio | 10 |
| | Total | 330 |

Tabela 1: Atividades e horas estimadas

8 Resultados Esperados

Como resultado principal do estágio, espera-se obter uma tabela que contemple as grades avaliadas e as compare, com base nas métricas identificadas ao longo do estudo. Através de tal tabela, será possível identificar claramente pontos positivos e negativos de cada tecnologia.

9 Cronograma

Pretende-se realizar as atividades enumeradas na seção 7 segundo o cronograma representado na tabela 2:

| | Março | Abril | | | | Maio | | | | Junho | | | | Julho | |
|----|-------|-------|--|--|--|------|--|--|--|-------|--|--|--|-------|--|
| | | | | | | | | | | | | | | | |
| A1 | | | | | | | | | | | | | | | |
| A2 | | | | | | | | | | | | | | | |
| A3 | | | | | | | | | | | | | | | |
| A4 | | | | | | | | | | | | | | | |
| A5 | | | | | | | | | | | | | | | |
| A6 | | | | | | | | | | | | | | | |
| A7 | | | | | | | | | | | | | | | |
| A8 | | | | | | | | | | | | | | | |
| A9 | | | | | | | | | | | | | | | |

Tabela 2: Cronograma

Referências

ANDERSON, D. P. BOINC: A System for Public-Resource Computing and Storage. In: *GRID '04: Proceedings of the 5th IEEE/ACM International Workshop on Grid Computing*. Washington, DC, USA: IEEE Computer Society, 2004. p. 4–10. ISBN 0-7695-2256-4.

BERMAN, F.; FOX, G.; HEY, T. *Grid Computing: Making The Global Infrastructure a Reality*. [S.l.]: John Wiley & Sons, 2003. Hardcover. ISBN 0470853190.

CIRNE, W. et al. Labs of the world, unite!!! *Journal of Grid Computing*, v. 4, n. 3, p. 225–246, 2006.

FEDAK, G. et al. XtremWeb: a generic global computing system. In: *Cluster Computing and the Grid, 2001. Proceedings. First IEEE/ACM International Symposium on*. [S.l.: s.n.], 2001. p. 582–587.

KESSELMAN, C.; FOSTER, I. *The Grid: Blueprint for a New Computing Infrastructure*. [S.l.]: Morgan Kaufmann Publishers, 1998. Hardcover. ISBN 1558604758.

THAIN, D.; TANNENBAUM, T.; LIVNY, M. Distributed computing in practice: the condor experience. *Concurrency and Computation: Practice and Experience*, Computer Sciences Department, University of Wisconsin-Madison, 1210 West Dayton Street, Madison, WI 53706, U.S.A., v. 17, n. 2-4, p. 323–356, 2005.

10 Aprovação

Declaro para os devidos fins que aprovo o planejamento das atividades descritas neste documento como plano de estágio do aluno Ricardo Araújo Santos, matrícula 20511120.

Marcus Williams Carvalho de Aquino
Supervisor Técnico

Raquel Vigolvino Lopes
Supervisora Acadêmica

Joseana Macêdo Fachine
Coordenadora da Disciplina Estágio Integrado