

Corujah: Um Modelo para a Computação Distribuída na Internet utilizando Java e HTTP

Vanessa Barcellos Vargas, Marcos José Brusso

Universidade de Passo Fundo, ICEG
Campus I, BR 285, Bairro São José
Passo Fundo (RS) - Brasil - CEP: 99001-970 - Caixa Postal: 611
Telefone: (54) 316-8354, Fax: (54) 316-8346
vanessabv@correios.com.br, brusso@upf.br

Introdução

A computação distribuída é uma tecnologia que divide a carga de processamento de uma determinada aplicação entre diversos computadores. Com os avanços da tecnologia e da Internet, tornou-se possível construir ambientes de computação distribuída com enorme poder computacional, através da capacidade de processamento dos computadores que se encontram atualmente conectados Internet. Estes ambientes computacionais são capazes de processar uma grande quantidade de dados que levaria um tempo inaceitável para sua conclusão se implementado sobre um único dispositivo de computação [LEN 01].

A principal vantagem de utilizar a computação distribuída através da Internet é a sua grande capacidade de escalabilidade, isto, devido à facilidade de aumentar o poder computacional destes ambientes através da agregação de mais computadores.

Diversos projetos têm sido desenvolvidos para tirar proveito desta tecnologia, como o SETI (*Search for Extraterrestrial Intelligence*) [SET 02], o RC5 [DIS 02] e o GIMPS (*Great Internet Mersenne Prime Search*) [GIM 02], os quais têm atingido alta capacidade de processamento. Normalmente, estes projetos fazem o uso de interfaces e formas de comunicação proprietárias, utilizando a Internet apenas como meio de comunicação. Desta forma, estes projetos não abrangem a todos os usuários da Internet devido a vários tipos de plataformas existentes na rede e a diversidade da utilização dos softwares desenvolvidos.

Este artigo tem por objetivo apresentar um modelo de computação distribuída, denominado Corujah, que visa tirar proveito da Internet como um computador distribuído virtual, utilizando *applets* Java como ferramenta de programação, devido a sua independência de plataforma e a sua facilidade de uso através de *browsers Web*, e o protocolo HTTP como forma de comunicação, por ser um dos protocolos padrão utilizados na Internet.

Modelo Corujah

O Corujah (Computação Distribuída na Internet utilizando Java e HTTP)¹ é um modelo para ser utilizado no desenvolvimento de aplicações distribuídas na Internet

¹ O site deste projeto está disponível em <http://inf.upf.br/compadi/corujah/index.html>.

através da implementação de *applets* Java com comunicação via protocolo HTTP. Este modelo visa, principalmente, facilitar a implementação de aplicações as quais possibilitem que os computadores conectados na rede possam participar de um sistema distribuído, utilizando como meio de acesso os *browsers* Web com suporte a Java, existentes em praticamente todas as máquinas.

A principal motivação de utilizar *applets* no desenvolvimento deste modelo é simplificar a utilização e execução do mesmo, dispensando aos usuários a necessidade de realizar o *download* e a instalação do software de forma manual.

O modelo utiliza o HTTP como forma de comunicação entre o servidor e o cliente, por ser um protocolo padrão utilizado pela Internet. Desta forma, todo o controle da comunicação entre o cliente e o servidor não precisa ser codificada pelo desenvolvedor das aplicações que forem implementadas a partir deste modelo, ficando a cargo do *browser*, onde o *applet* é executado, e do servidor Web. Além disso, o desenvolvimento da aplicação que será executada no servidor independe da linguagem utilizada, pois o servidor HTTP suporta qualquer linguagem script.

A estrutura deste modelo é composta, em um mais alto nível, por um servidor e por diversos clientes, conforme ilustrado na Figura 1.

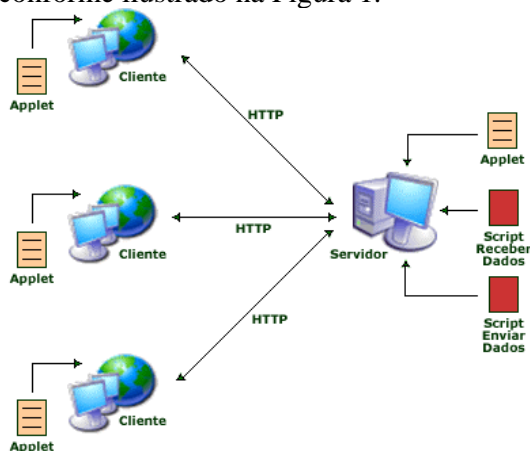


Figura 1 – Estrutura do Corujah

Os Clientes Corujah

Cada cliente irá acessar, através de uma URL, o *applet* que se encontra armazenado no servidor, no qual será descarregado e disparado para ser executado localmente no cliente. O *applet*, ao ser carregado, estabelece uma conexão com servidor, solicitando através da utilização dos métodos GET ou POST (dependendo da aplicação), o envio dos dados para serem processados. O servidor responde a esta requisição do cliente com uma mensagem de resposta contendo os dados a serem processados no seu corpo. Com isso, o *applet* realiza todo o processamento necessário com estes dados e quando os resultados forem concluídos, o cliente estabelece novamente uma conexão com o servidor através de uma mensagem contendo os resultados em seu corpo ou cabeçalho, dependendo do método empregado. Os processos de requisição e submissão de dados ao servidor são realizados enquanto o *applet* não for destruído. Sendo assim, após o armazenamento dos dados processados pelo cliente, o *applet* realiza novamente outra solicitação de dados e assim, sucessivamente.

O Servidor Corujah

A aplicação do servidor HTTP é composta pela implementação de dois *scripts*. O primeiro é responsável por controlar e enviar os dados a serem processados pelos clientes, e o segundo é responsável por receber e armazenar os resultados já processados. Estes processos executados no servidor não são persistentes, pois são disparados pelo servidor *Web* para atender a uma requisição, desta forma, eles não podem usar a memória principal para manter o estado dos dados a serem distribuídos. Com isso, o armazenamento destes dados podem ser realizado através de arquivos comuns em disco ou através da utilização de um gerenciador de banco de dados, como por exemplo, o *Oracle*, o *MySQL*, o *PostgreSQL*, o *Interbase*, entre outros.

Os scripts podem ser desenvolvidos em quaisquer linguagens que o servidor HTTP suporte, tais como, PHP, JSP, C/C++ (na forma de *scripts CGI*), ASP, entre outras. Desta forma, o programador pode utilizar a linguagem que preferir para desenvolver a sua aplicação.

Um aspecto que deve ser considerado em relação ao desenvolvimento destes *scripts* é que os mesmos devem conter um controle do tempo de resposta dos clientes para o servidor. Caso este tempo seja ultrapassado, o servidor reenvia estes dados para que outro cliente realize o processamento dos mesmos, desconsiderando o primeiro envio. Este tempo se faz necessário porque, por se comunicarem através do protocolo HTTP, os clientes não mantêm uma conexão persistente com o servidor.

Estudo de caso: Corujah Nqueens

Com o objetivo de validar o modelo Corujah foi desenvolvida uma aplicação de computação distribuída utilizando a Internet, denominada *Corujah nQueens*. *NQueens* é um estudo científico realizado desde 1850, proposto por Carl Friedrich Gauss, originalmente conhecido como “*The NQueens Problem*”. Este estudo propõe encontrar disposições de N rainhas do jogo de xadrez em um tabuleiro $N \times N$, de tal modo que nenhuma rainha ataque às outras rainhas, de acordo com as regras do jogo. Para que satisfaça esta condição significa que no máximo pode haver uma rainha em cada linha, coluna ou diagonal, e exatamente N rainhas no tabuleiro [LET 03].

Diversos pesquisadores e cientistas estão realizando este estudo como exemplo de algoritmos de busca em profundidade, geração de permutação, algoritmos de divisão e conquista, entre outros. Desta forma, já foram encontradas as soluções para tabuleiros de até 23×23 [RUS 03]. Com base neste estudo realizado sobre o problema das N rainhas, verificou-se que esta aplicação é adequada para a validação do modelo Corujah. Desta forma, foi desenvolvida a aplicação *Corujah nQueens*, que utiliza o processamento dos computadores que se encontram conectados na Internet, conforme o propósito do modelo. Para esta aplicação, foi escolhida a busca em um tabuleiro de 24×24 , isto porque as soluções onde $N < 24$ já haviam sido encontradas [RUS 03].

A aplicação foi desenvolvida utilizando o JDK versão 1.1.8 para o desenvolvimento do *applet* que será executado no cliente e a linguagem PHP versão 4.2.3 para implementar os *scripts* que estarão rodando no servidor. Os dados foram armazenados em uma base de dados do *MySQL* versão 3.23.46.

Resultados obtidos

A aplicação *Corujah nQueens* foi disponibilizada para o acesso a partir do dia 30 de maio de 2003 e em 20 de junho de 2003 foram obtidos os resultados apresentados abaixo. Como esta aplicação não foi divulgada ao público em geral por estar em período de validação, contou-se, principalmente, com a colaboração dos computadores dos Laboratórios de Pesquisas e do Laboratório de Informática da Universidade de Passo Fundo. Foram utilizados computadores com sistemas operacionais diferentes, tais como, *Linux*, *Solaris*, *Windows 98* e *Windows NT*.

Durante este período de 20 dias foram encontradas 14.990.915.565 soluções em um tempo acumulado de execução de 4.824 horas. Esta quantidade de soluções é o resultado de apenas 0,05% dos conjuntos existentes para serem processados. Devido ser inviável calcular o tempo seqüencial desta aplicação, não foi possível fazer uma comparação com o tempo obtido pelo *Corujah nQueens*.

Conclusão

O presente trabalho apresentou um modelo de computação distribuída na Internet, denominado *Corujah*, utilizando *applets* Java com comunicação via protocolo HTTP. Este modelo possibilita desenvolver aplicações que fazem o uso da distribuição dos dados para serem processados em computadores ociosos que se encontram conectados na Internet.

Com a implementação da aplicação *Corujah nQueens* verificou-se que o modelo *Corujah* pode ser utilizado por diversas aplicações que necessitam deste tipo de solução para processarem seus dados com mais facilidade e rapidez. Sem muita divulgação do projeto já contabilizou 4.824 horas de processamento apenas nos 20 primeiros dias e em torno de 15 bilhões de soluções já foram encontradas.

Referências

- [DIS 02] DISTRIBUTED. **Distributed.net**. Disponível em: <<http://www.distributed.net>>. Acesso em: 04 abr. 2003.
- [GIM 02] GIMPS. **Source Code**. Disponível em: <<http://www.mersenne.org/source.htm>>. Acesso em: 30 abr. 2002.
- [LEN 01] LENGAUER, C.; PANCAKE, C. High-Performace Java. **Communications of the ACM**, New York, v. 44, n. 10, p. 98-101, out. 2001.
- [LET 03] LETAVEC, C.; RUGGIERO, J. The n-Queens Problem. **INFORMS Transcations on Education**, v. 2, n. 3, p. 101-103, mai. 2002, Disponível em: <<http://ite.informs.org/vol2no3/LetavecRuggiero/letavecruggiero.pdf>>. Acesso em: 25 maio 2003.
- [RUS 03] RUSKEY, F.; RUSKEY, S. **Information on the n Queens problem**. Disponível em: <http://www.schoolnet.ca/vp-pv/amof/e_queeI.htm>. Acesso em: 05 maio 2003.
- [SET 02] SETIATHOME. **Search for Extraterrestrial Intelligence**. Disponível em: <<http://setiathome.ssl.berkeley.edu/>>. Acesso em: 08 fev. 2002.