

Arquitetura computacional baseada em computação GRID, aplicada a sistemas de informação geográfica na gestão de risco e alerta da bacia do rio Itajaí

Marcos Rodrigo Momo¹
Julio Cesar Refosco¹

¹ Universidade Regional de Blumenau – FURB
Centro de Operações do Sistema de Alerta da Bacia Hidrográfica do Rio Itajaí - CEOPS
Blumenau - SC, Brasil
{momo, refosco}@furb.br

Abstract. GIS became a very important component to risk and emergency management and every day new applications can be viewed. As new requirements are developed in the risk management and emergence management as well as in the hazard management, they demand new technologies related to its tools including GIS. This ongoing evolution continuous demands higher computing capacities, security and processing as well as redundancy against fails and lot of times require remote web connections. The simple client server models have important problems to achieve those requisites. Thus, it is perceived the necessity of a collaborative dynamic application in the base for the GIS and in this case the middleware model has emerged. In the other hand, GRIDs services represent an extended form of web, therefore, beyond making possible the automatic communication between applications, it also allows that computational resources can be shared and be offered as services of transparent form to the user. The objective of this paper is to propose a computation architecture model, based on distributed high level systems, specifically using grid technologies in order to supply the monitoring and alert system for Itajaí basin. There are some preliminary results, restricted to the conceptual field. Until the present moment, the conceptual model has been developed and improved to make front to the institutional demands.

Palavras-chave: geographic information system, grid computing, risk management, Itajaí basin, sistema de informação geográfica, computação grid, gestão de risco, bacia do Itajaí.

1. Introdução

Os sistemas de informática desenvolvem-se rapidamente à medida que novas aplicações vêm impondo avanços, tais como, acesso web, segurança, capacidade computacional, armazenamento, entre tantos outros. Aplicações modernas de SIG (Sistema de Informação Geográfica) seguem esta tendência e demandam cada vez mais carga computacional, gerando grandes bancos de informação aplicados nas mais diversas áreas. Um dos exemplos mais marcantes e atuais desta demanda é a *e-Science* (FOSTER, 2005), termo que denota o sistemático desenvolvimento de métodos para explorar o pensamento computacional avançado e tem sido aplicado a desenvolvimentos em diversas áreas de conhecimento. Tais métodos possibilitam acesso a recursos distribuídos usando a idéia da computação distribuída. Os recursos incluem coleções de dados, recursos de computação em grande escala, instrumentos científicos e sistemas de alto desempenho em visualização.

Tradicionalmente as aplicações baseada no modelo cliente servidor, ou seja, onde há uma separação de clientes e de servidor, interconectados por uma rede de computadores, apresenta diversos inconvenientes, tais como, sobrecarga no servidor, falta de robustez no processamento, custo de aquisição e manutenção geralmente é bastante oneroso. Sobretudo a arquitetura de cliente servidor comum, não contempla plenamente os requisitos demandados pelos sistemas de emergência e alerta de catástrofes naturais. Neste tipo de aplicação requisitos de disponibilidade, interoperabilidade e redundância a falhas são imprescindíveis, principalmente durante a ocorrência de um evento climático extremo no qual a tomada de decisão deve ser urgente.

Para atender tais requisitos, neste trabalho se propõe um modelo de arquitetura computacional baseado em computação distribuída de alto desempenho, utilizando mais

especificamente tecnologias baseadas em computação GRID aplicada a SIG para gestão de risco e alerta da bacia hidrográfica do rio Itajaí.

Os problemas, de longa data, que compreendem o histórico das enchentes sucedidas na Bacia Hidrográfica do Itajaí em Santa Catarina, somados ao processo de degradação do solo, foram sendo construídos junto à história da colonização ocorrida a partir do ano de 1850. O Centro de Operações do Sistema de Alerta da Bacia Hidrográfica do Rio Itajaí – CEOPS (CEOPS, 2010), criado em 1984 pela Universidade Regional de Blumenau, desenvolve diversas atividades relacionadas à gestão de situações de alerta (MOMO et al., 2010; SILVA, 2010), especialmente contra enchentes. Dentre as atividades desenvolvidas destacam-se a previsão do tempo e nível do rio Itajaí-açu, monitoramento do nível dos rios e clima, previsão hidrológica, e outras atividades relacionadas a estes temas.

A atividade de previsão meteorológica é caracterizada por três etapas: 1) coleta dos dados meteorológicos e das imagens de satélite; 2) análise dos modelos matemáticos de previsão e 3) elaboração da previsão do tempo. A bacia do rio Itajaí-açu é monitorada por dezesseis estações telemétricas, que coletam, em tempo atual, o nível dos rios e a precipitação ocorrida. Com base nessas informações, são realizadas as previsões de nível para Blumenau, com até oito horas de antecedência. O CEOPS desenvolve estudos meteorológicos, climatológicos e hidrológicos com o objetivo de aperfeiçoar o conhecimento dos principais sistemas meteorológicos que atuam no sul do Brasil (SEVERO, 1994).

Estes estudos também fornecem subsídios para o aperfeiçoamento dos modelos matemáticos de previsão do nível do rio Itajaí em Blumenau (CORDERO et al., 1998 e outras cidades do vale do Itajaí (CORDERO et al., 2010; CORDERO, 1989). Estas atividades são de grande interesse à comunidade em geral e especialmente ao setor de administração de defesa civil do município de Blumenau, já que as informações do CEOPS são utilizadas para a tomada de decisão.

Os sistemas de informação geográfica apresentam grande desenvolvimento, sempre em busca de melhorias. O trabalho com grandes e complexas bases de dados espaciais e não espaciais, necessitam de ferramentas específicas para tarefas desta natureza. Uma das tendências de desenvolvimento é o processamento distribuído de informação. A informação espacial tende a ser espacial e logicamente distribuída, assim como muitas funções e serviços do SIG. Além disso, as análises espaciais tem se tornado cada vez mais complexas e computacionalmente intensivas. Assim, percebe-se a necessidade de uma aplicação dinâmica colaborativa como base para o SIG e neste caso os modelos “*Middleware*” tem sido destacados.

Os tipos de *middleware* baseados em grid computacional têm sido considerados como uma solução bastante interessante para a próxima geração do SIGs. Basicamente o conceito de GRID pode trazer facilidades para a coordenação dos recursos do SIG bem como para a solução de problemas em organizações dinâmicas, multiorganizacionais através da conexão de recursos em redes de alto desempenho (AHMED, 2008).

O processamento paralelo para SIGs trás como principais vantagens as seguintes:

- Ganho em tempo
- Possibilidade de solução de problemas de grande dimensão
- Capacidade de simultaneidade de processamento e solução de problemas diversos ao mesmo tempo
- Uso de recursos remotos em rede e mesmo em internet quando os recursos locais forem escassos
- Redução de custos em função do uso de recursos de computação baratos
- Superacao de constrições de recursos de memória dos computadores individuais (AHMED, 2008; XIAOSHENG et al., 2009)

2. Metodologia de trabalho

2.1. Arquitetura Orientada a Serviços (SOA)

A arquitetura orientada a serviços (SOA – *Service Oriented Architecture*) (PAPAZOGLU, 2003) é um paradigma computacional que flexibiliza o desenho, o desenvolvimento e a execução de aplicações distribuídas. SOA utiliza os serviços como o elemento básico para a composição das aplicações.

Serviço é um componente de software mantido por uma organização provedora que oferece funcionalidades para seus potenciais clientes do serviço (SZYPERSKI, 2003). Os serviços dos distintos provedores podem ser utilizados e integrados com o objetivo de oferecer outros serviços, neste caso são chamados de serviços compostos (BOTE-LORENZO et al., 2004).

A Figura 1 ilustra o modelo de interação entre cliente e os serviços. Este modelo possibilita ao cliente entrar em contato e utilizar os serviços que oferecem uma funcionalidade desejada.

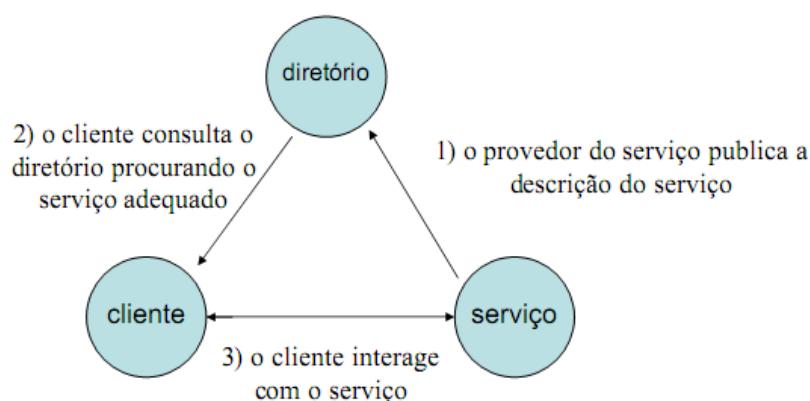


Figura 1. Modelo de interação SOA (PAPAZOGLU, 2003).

2.2. Computação GRID

Computação GRID ou computação em malha é um paradigma computacional que possibilita agregar recursos tanto de hardware como de software para obter capacidades de supercomputação (FOSTER, 2004). Um GRID computacional é uma infraestrutura distribuída com alta escalabilidade permitindo o compartilhamento de recursos computacionais entre múltiplas organização e capaz de virtualizar recursos em serviços. Em (DANTAS, 2005) afirma que os GRIDs representam uma forma estendida dos *web services*, pois além de possibilitar a comunicação automática entre aplicações, também permite que recursos computacionais possam ser compartilhados e oferecidos como serviços de forma transparente ao usuário.

Os *web service* são componentes de software programados para obter informações distribuídas, que possam ser acessadas a partir de protocolos Internet, com a finalidade de realizar a comunicação entre as aplicações. *Web service* integra aplicações e torna a lógica dessas aplicações disponíveis para outros clientes. Oferece interoperabilidade de interações “máquina-a-máquina” através de uma rede. Possui uma interface descrita em um formato processável pela máquina, especificamente WSDL (*Web services Description Language*). Outros sistemas interagem com o *web service* em uma maneira prescrita por sua descrição usando mensagens SOAP (*Simple Object Access Protocol*), tipicamente usando HTTP com uma serialização XML em conjunção com outros padrões relacionados à web (CURBERA, 2007).

Entre as características mais importantes de um GRID computacional é de grande escala, distribuição de recursos distintos e localizados geograficamente, compartilhamento de recursos de hardware e software, múltiplos domínios administrativos, acesso transparente.

Hoje em dia o uso de tecnologia baseada em computação GRID é empregado em diversos domínios, principalmente para oferecer capacidade de processamento computacional e qualidade de serviços não trivial.

2.3. Arquitetura computacional para apoiar aplicações SIG

Segundo (FOSTER, 2004) a arquitetura orientada a serviços é o modelo mais adequado para a implementação de uma arquitetura GRID, pois todas as funcionalidades correspondentes aos níveis de serviços locais, básicos e avançados devem ser oferecidos como serviços, de acordo com a definição do modelo SOA.

A Figura 2 ilustra uma visão geral do sistema descrito neste trabalho.

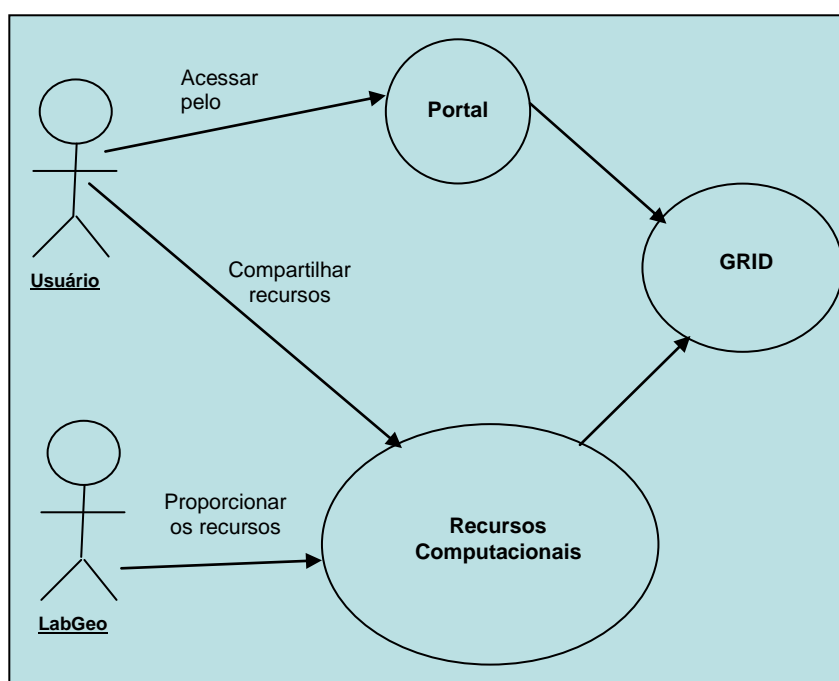


Figura 2. Funcionamento do sistema.

Este modelo proporciona ao usuário não apenas beneficiar-se dos recursos disponibilizados pelo GRID através do portal *web*, mas também contribuir com recursos computacionais para o GRID, ou seja, o usuário pode compartilhar seus próprios recursos para fazer parte do sistema.

Como uma alternativa para oferecer um ambiente com capacidade de supercomputação e atender aos requisitos das aplicações SIG, conforme descritos na Tabela 1, propõe-se um modelo de arquitetura baseado em GRID computacional. Através da utilização do *framework Globus Toolkit (GT)*, o qual proporciona um marco de desenvolvimento para GRID orientado a serviços (FOSTER, 2006), possibilitando desenvolver tecnologias e padrões para a construção de sistemas GRID.

Além de se implementar os serviços básicos da arquitetura GRID proporcionados pelo Globus Toolkit, deve-se implementar os serviços específicos para fazer cumprir os requisitos de aplicações baseadas em SIG. Na Figura 3, ilustra-se a arquitetura do sistema.

No nível inferior da arquitetura encontram-se os distintos recursos que são compartilhados e formam o GRID. Estes recursos são ofertados como serviços, ou seja, os recursos são

virtualizados como uma série de serviços web. Estes serviços se consideram serviços web locais básicos, e são concretamente serviços oferecidos por GT (FOSTER, 2006).

I) Serviços segurança

São compostos de serviços para autenticação, autorização e delegação para os usuários. GT proporciona interface de segurança GSI (*Grid Security Infrastructure*), que dispõe diversos serviços necessários para manter a segurança necessária do sistema.

- *Community Authorization (CA)* – estabelecer políticas de uso do GRID em nível de grupo;
- *Transport Security Layer (TLS)* – oferecer segurança nas transferências de dados;
- *Credential Management (CM)* – controlar quem é quem no GRID.

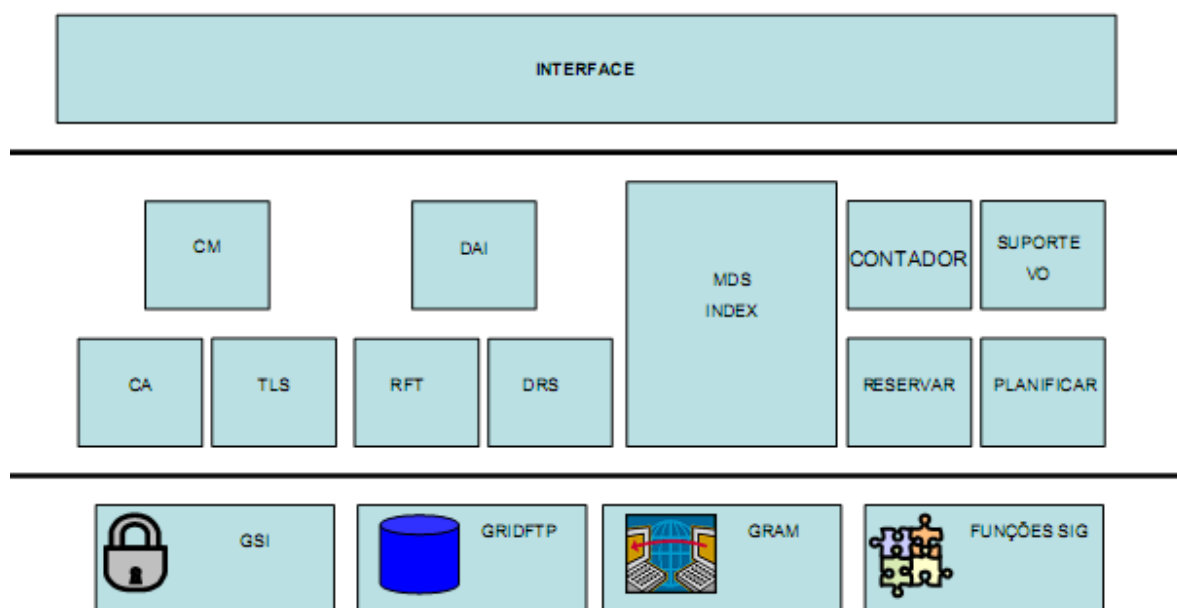


Figura 3. Arquitetura do sistema.

II) Serviços gestão de dados

As transferências de dados para processamento devem ser rápidas e confiáveis. As transferências são realizadas mediante GRIDFTP, oferecido por GT. Para proporcionar este requisito a arquitetura dispõe dos seguintes serviços:

- *Reliable File Transfer (RFT)* – caso ocorra um erro durante a transmissão de dados, este requisito se encarrega de guardar o estado da transferência;
- *Data Replication Service (DRS)* – permite replicar dados existentes em um entorno GRID;
- *Data Access and Integration (DAI)* – este serviço viabiliza o acesso e a integração dos recursos de dados.

III) Serviços gestão de execução

Para realizar o gerenciamento da execução de trabalho, o GT oferece GRAM (*Grid Resource Allocation and Management*), que proporciona serviços para enviar, monitorar e cancelar a execução de trabalhos nos diversos recursos do GRID. Cria um ambiente para o trabalho submetido, ser executado através de planejamento, monitorando a execução e disponibilizando informações de mudanças de estados. Permite ao cliente, através de interfaces XML, entre outras coisas, especificar o tipo e a quantidade de recursos a serem

utilizados, os dados de entrada (a enviar) e saída (a receber), assim como as credenciais que podem utilizar.

- *Monitoring and Discovery System Index (MDS)* – este serviço é responsável por coletar as informações de monitoramento e descobrimento dos recursos no GRID. As informações são publicadas no *Index* que é a localização ou diretório onde estão todas as informações de monitoramento e descobrimento de recursos no GRID, o *Index* possibilita uma única localização e de forma integrada;

IV) Serviços específicos

São os serviços específicos para o sistema SIG aplicado na gestão de risco e alerta da bacia hidrográfica do rio Itajaí, como reserva, planificação, contabilidade e suporte a organização virtual.

- *Reservar* – possibilita ao usuário do GRID fazer reservas de recursos através de serviços de reservas. Este serviço comunica ao planejador de tal forma que no momento da execução do *job*, o serviço planejador possa aceitar de forma prioritária a execução dos trabalhos enviados pelo usuário que efetuou a reserva;

- *Planificar* – existem vários tipos de planejadores, entre os mais conhecidos estão Condor, LSF, OpenPSB. Estes planejadores podem ser integrados ao GRAM de forma que o GT trabalhe com o CSF (*Community Scheduler Framework*) que proporciona uma interface única para diferentes planejadores;

- *Contador* – é um serviço que permite saber quem e quanto se utiliza o GRID. Com esta informação é possível realizar o uso efetivo do sistema, compensar os usuários que mais contribuem para GRID e saber quem realiza o uso ineficiente, de forma que possam ser sancionados.

- *Suporte VO* – este serviço se encarrega de oferecer o suporte necessário para que os usuários possam contribuir com recursos para o GRID. Estes serviços se limitam a dar suporte ao processo de adição de novos recursos, entidades e usuários no GRID computacional SIG.

Do ponto de vista tecnológico, estes serviços visam cumprir todos os requisitos demandados pelas aplicações SIG aplicado a gestão de risco e alerta da bacia hidrográfica do rio Itajaí. Na Tabela 1, descrevem-se os principais requisitos das aplicações SIG e os correspondentes serviços GRID associados.

Tabela 1. Requisitos do sistema SIG e prestações através de serviços GRID.

Requisitos SIG	Prestações por serviços Grid
Integração de dados de fontes diversas e transferências seguras	Gridftp, dai, rft Grid services
Integração de informações de fontes espacialmente e logicamente distribuídas e transferências seguras	Gridftp, dai, rft Grid services
<i>Data mining</i> de informações espaciais e réplicas de base de dados	Gridftp, drs Grid services
Capacidade de armazenamento	Compartilhamento de recursos
Capacidade de processamento	Compartilhamento de recursos
Ambiente colaborativo e contribuição de recursos	Compartilhamento de recursos Funções sig, contador, suporte vo
Segurança (autenticação, comunicação e autorização através de políticas da comunidade bem como permissão de controle local)	Gsi, cm, ca, tls
Paralelização de processos e gestão de processos	Funções sig, planificar Gram, mds index
Qualidade de serviços não trivial	Reservar

3. Resultados e Discussão

Este trabalho encontra-se em desenvolvimento e os resultados são parciais, restritos ainda, ao campo conceitual. Até o presente momento, o modelo conceitual tem sido desenvolvido e melhorado para fazer frente às demandas institucionais.

O modelo descrito visa oferecer uma ambiente computacional com grande capacidade de processamento, com baixo custo de implantação e aquisição de equipamentos, viabilizado através de um sistema de software recursos para armazenamento e compartilhamento de bases de dados e recursos de hardware de maneira segura e transparente para o usuário.

4. Conclusões

Neste trabalho apresentou-se um modelo de arquitetura computacional baseada em computação GRID, aplicada na gestão de risco e alerta da bacia hidrográfica do rio Itajaí. O modelo é baseado em GT e oferece ao usuário, requisitos de qualidade de serviços (QoS) através dos serviços que possibilitam reserva de recursos. Aspectos de segurança, gestão de execução e movimentação de dados são cobertos pelos serviços básicos oferecidos pelo GT. O usuário poderá beneficiar-se do GRID para realizar tarefas complexas de processamento e armazenamento de dados. O compartilhamento de recursos possibilitados pela tecnologia GRID, permite explorar a redundância de recursos, principalmente aqueles de uso esporádicos ou subutilizados, baseado nisso o sistema prevê a possibilidade de o usuário do GRID compartilhar seus próprios recursos.

Através do compartilhamento de recursos computacionais é possível obter capacidade computacional com requisitos de segurança, interoperabilidade, transparência e redundância a falhas, comparado aos grandes servidores, onde muitas vezes o custo de aquisição e de manutenção é proibitivo.

Aplicações SIG são típicos problemas que podem ser resolvidos através da paralelização de processos, permitindo distribuir processamento entre os recursos pertencentes no GRID. Este trabalho buscou-se uma solução para viabilizar execução de aplicações SIG de forma rápida, segura e transparente para os usuários.

5. Agradecimentos

Parte deste trabalho está sendo financiado pelo CNPq e pela FAPESC.

6. Referências Bibliográficas

Ahmed, Z. I. Grid Computing in Distributed GIS. **GIS Lounge**, 2008. Disponível em: <<http://ezinearticles.com/?Grid-Computing-In-Distributed-GIS&id=1180713>>. Acesso em: 11 nov. 2010.

Bote-Lorenzo, M. L.; Vaquero-Gonzalez, L. M.; Vega-Gorgojo, G.; Asensio-Perez, J. I.; Gomez-Sanchez, E.; Dimitriadis, Y. A. GRIDCOLE: a Grid Collaborative Learning Environment Cluster Computing and the Grid. IEEE International Symposium on, pp. 105-112, **Fourth IEEE International Symposium on Cluster Computing and the Grid (CCGrid'04)**, 2004.

CEOPS. **Centro de Operação do Sistema de Alerta da Bacia do Rio Itajaí**; CEOPS, 2010. Disponível em <<http://ceops.furb.br>>. Acesso em: 12 nov. 2010.

Cordero, A. Modelo de previsão de cheia em tempo atual para a cidade de Gaspar. **Anais do VIII Simpósio Brasileiro de Hidrologia e Recursos Hídricos**, 2(ABRH): p.62-73. 1989.

Cordero, A.; Momo, M. R.; Severo, D. L. **Previsão de cheias em tempo atual, com um modelo tipo ARMA para a cidade de Rio do Sul - SC**. Informe Técnico, 2010.

Cordero, A.; Teran, A. L.; Medeiros, P. A. Real-time flood forecasting with a stochastic model. in **International Workshop on non-structural flood control in urban areas**. São Paulo: IRTCUD-RCTC. 1998

Curbera, F. Component contracts in service-oriented architectures. **Computer**, v. 40, n. 11, p. 74-80, 2007

Dantas, M. **Computação Distribuída de Alto Desempenho**. Florianópolis: Axcel, 2005.

Foster, I. Globus toolkit version 4: Software for service-oriented systems. **Journal of Computer Science and Technology**, v. 21, n. 4, p. 513-520, 2006.

Foster, I. Service-oriented science. **Science**, v. 308, n. 5723, p. 814, 2005.

Foster, I.; Kesselman, C. **The grid: blueprint for a new computing infrastructure**. São Francisco: Morgan Kaufmann, 2004.

Globus Alliance, **Globus toolkit 5.0**. Disponível em: <<http://www.globus.org>>. Acesso em: 20 nov. 2010.

Momo, M. R., Silva, H. S.; Cordero, A.; Severo, D. L.; Tachini, M.; Refosco, J. C. Serviços Grid/Web para Sistemas de Emergência, in IV MIPE - Mostra Integrada de Ensino, Pesquisa e Extensão – MIPE, 2010, Blumenau. **Dynamis**, v. 16, n. 2, 2010.

Papazoglou, M. P. Service-oriented computing: concepts, characteristics and directions. Web Information Systems Engineering, 2003. WISE 2003. **Proceedings of the Fourth International Conference on**, p. 3-12, 2003.

Severo, D. L. Estudo de casos de chuvas intensas no estado de Santa Catarina. 1994. 118p. (INPE – 5682 – TDI / 568). Dissertação (Mestrado em Meteorologia) – Instituto Nacional de Pesquisas Espaciais – INPE. São José dos Campos, 1994.

Silva, H. S.; Cordero, A.; Momo, M. R.; Severo, D. L.; Tachini, M.; Refosco, J. C. Centro de Operações do Sistema de Alerta da Bacia do Rio Itajaí – operação, previsões e danos, in IV MIPE - Mostra Integrada de Ensino, Pesquisa e Extensão – MIPE, 2010, Blumenau. **Dynamis**, v. 16, n. 2, 2010.

Szyperski, C. Component technology: what, where, and how? In: 25th International Conference on Software Engineering, 2003, Portland. **Proceedings of...** Portland: IEEE Computer Society, Los Alamitos, p. 684-693, 2003.

Xiaosheng L.; Xiaobin H.; Zhiyong Z. Research on GIS Spatial Database Based on Grid Computing. In: International Conference on Communications and Mobile Computing, Kunming, 2009. **Proceedings of...** IEEE Computer Society, Los Alamitos, p. 156-159, 2009.