Computação GRID

# Paper presentation

## Security in grid computing: A review and synthesis

## Grid Security: Next Steps


Miguel Lobo nº 200707399

Pedro Oliveira nº 200501071

## Paper – Security in grid computing: A review and synthesis

### 1.Escolha do Paper

Escolhemos este paper, já que resume várias soluções de segurança em grid apresentadas noutros trabalhos, dando uma boa vista geral sobre o *state of the art* das diferentes abordagens possíveis em segurança em grids à data de escrita do paper, bem como as respectivas vantagens e desvantagens.

Este foi um dos papers sugeridos pela docente, tendo bastado uma pesquisa pelo seu título para o acesso ao mesmo ser facilmente obtido.

### 2.Principal problema tratado

O problema tratado neste paper são as diferentes abordagens possíveis quanto a segurança em grids, a sua adequação às exigências de performance de uma grid e às suas peculiaridades em relação a outras configurações de rede mais usuais – i.e. o facto de os seus utilizadores e nós terem uma transiência na rede bastante superior à verificada numa rede centralizada.

O paper começa por comparar as características de diferentes redes e respectivos desafios no que toca a segurança. As redes consideradas são uma rede centralizada, uma rede P2P e uma rede de computação em grid.

A conclusão atingida é que devido à transiência dos seus utilizadores e nós, ao facto de organizações diferentes poderem fazer parte da mesma grid e à sua componente distribuída, não é possível aplicar mecanismos de segurança já existentes, pelo menos sem os modificar para respeitarem o modo de funcionamento de uma grid e visarem os seus desafios e vulnerabilidades específicas, e.g. a necessidade de manter identidades de utilizadores sobre redes locais e globais, as relações de confiança entre entidades e utilizadores e a gestão de certificados na grid.

O paper divide as possíveis soluções em 3 grandes classes: soluções de sistemas, soluções de comportamento, e soluções híbridas.

As soluções de sistemas procuram focar-se em manipular o hardware e software da grid directamente para atingir maior segurança.

As soluções de comportamento tentam atingir segurança através da implementação de políticas administrativas, ao invés de uma segurança física – e.g. controlo de acessos.

As soluções híbridas têm uma abordagem que considera ambas as aproximações acima definidas como válidas e procura misturá-las ou usar ambas em simultâneo.

- **Soluções de sistemas**

Nesta classe é proposto o sistema Entropia. O sistema Entropia procura tornar os recursos na grid mais seguros com recurso a sandboxing. Com esta técnica, os nós da grid (possivelmente máquinas de utilizadores a serem usadas em períodos idle) são protegidos por isolação da aplicação a executar do sistema em si (a aplicação corre num ambiente "caixa de areia", não conseguindo actuar fora da mesma). Assim, evitam-se ataques maliciosos ao sistema do nó ou aos seus ficheiros, bem como efeitos nefastos de aplicações mal conseguidas, tais como *memory leaks* ou alta utilização de disco.

Apesar destas serem características de segurança importantes, esta é uma segurança incompleta, já que não visa o aspecto da autenticação e identificação dos utilizadores.

- **Soluções de comportamento**

Esta secção lida com segurança através de políticas administrativas. Estas políticas podem afectar todas as áreas de grid computing, incluindo a selecção de utilizadores autorizados a fazer determinadas acções, procedimentos de autenticação e controlo de acessos e definições de segurança locais e globais.

Este tipo de solução oferece propostas de políticas para gerir grupos de utilizadores. Uma delas cria o conceito de CUG (Closed User Group), que recorre a controlo de acessos por políticas para suportar grupos de utilizadores geograficamente dispersos. Nós de administração no CUG controlam o acesso ao grupo através da emissão de certificados.

Este tipo de solução de segurança suporta melhor as considerações de overhead intrínsecas a um ambiente grid, já que a segurança obtida não tem base num software ou encriptação de segurança pesada, mas sim à instauração de políticas no sistema. Também respeita a heterogeneidade e a autonomia dos componentes da grid, dado o suporte a políticas locais e globais que cada nó pode exercer.

Ainda nesta classe de soluções, insere-se a subclasse de soluções baseadas em confiança. Esta classe contem algoritmos que conseguem baixar ainda mais o overhead devido a segurança na grid ao atribuir a cada entidade um valor associado à confiança na mesma.

Definindo um patamar mínimo de confiança, é possível dispensar medidas de segurança adicionais para entidades que igualem ou superem o patamar mínimo definido. O nível de confiança associado a uma entidade é actualizado com base nas suas acções e também

possivelmente de uma função de decay para que a confiança em entidades que não tenham interagido com a rede durante um período alargado desça naturalmente ao longo do tempo.

- **Soluções híbridas**

Esta classe de soluções implementa uma combinação da abordagem das classes anteriores, especialmente para o campo da autenticação e autorização de utilizadores.

Por exemplo uma solução de sistemas apresentada é o uso do sistema SHARP (Secure Highly Available Resource Peering). Neste sistema, os agentes e os resource managers possuem certificados de chave pública e as suas acções no sistema são assinadas criptograficamente, de forma a que não podem ser repudiadas. Não existe nenhuma agência de certificação central, cada organização local age como a sua própria agência de certificação. Isto evita o bottleneck associado à centralização da CA, bem como o único ponto de ataque que a mesma apresenta.

Uma solução de comportamento apresentada é a "Delegation Logic". O objectivo desta solução é apenas lidar com a autorização de utilizadores. A autorização ocorre por um método de "prova de compatibilidade" – uma entidade deve mostrar credenciais que provem que passou um determinado requisito para poder executar uma acção (os requisitos são determinados pelo recurso ou pela própria grid).

**3.Análise da Bibliografia**

O paper cita adequadamente todas as suas fontes, que apresentam uma boa visão sobre o panorama de segurança em grids à data da sua publicação. No entanto, a fonte mais recente citada neste paper data de 2006, e a mais antiga de 1999, pelo que não se pode considerar que este paper apresente uma visão fidedigna do *state of the art* neste campo.

**4.Organização**

O paper está razoavelmente bem organizado no geral. Divide adequadamente as soluções exploradas em três grandes campos, o que permite analisá-las em detalhe numa única secção, sem dispersão da informação. Também condensa alguma informação em tabelas, oferecendo uma boa vista geral sobre a mesma.

No entanto, detectámos algumas frases menos bem conseguidas e alguns erros ortográficos. Também achamos o paper algo confuso e em algumas secções algo perdido em detalhes, o que tornou a leitura algo difícil.

A apresentação de várias perspectivas sobre o mesmo tema é uma ideia interessante, mas por vezes teria sido bom explorar e detalhar melhor o funcionamento de algumas ideias ou algoritmos. Devido a isto, no final da leitura ficamos com uma ideia muito geral sobre as possibilidades e limitações das implementações de segurança em grid, mas não aprendemos muito sobre nenhuma em particular.

**5.Relevância do tema tratado**

A segurança é um tema particularmente interessante e relevante nos dias de hoje, em que cada vez mais são descobertos novos ataques e descobertas *0-day vulnerabilities* que tornam crucial e indispensável o desenvolvimento de soluções para manter a rede segura e íntegra.

Nas grids em particular, a componente de segurança para atacantes internos é tão importante como a que é dirigida a atacantes externos, o que torna este tipo de redes um desafio de segurança particularmente interessante, na nossa opinião. O facto de muitas vezes as grids correrem simulações e tarefas com dados empresariais e/ou confidenciais, e estarem expostas à Internet, torna qualquer pessoa com um computador com acesso à Internet um potencial atacante. Destes argumentos, concluímos que a necessidade de implementar segurança em Grids é inquestionável e de suma importância.

## Paper - Grid Security: Next steps

### 1. Escolha do paper

Escolheu-se analisar e discutir este paper porque, além de ser um dos papers sugeridos, fala de "Case Studies" reais, onde a segurança na GRID tornou-se um "must-have", para evitar possíveis problemas, tanto no uso de recursos, como no abuso de confiança entre utilizadores.

Para conseguirmos encontrar o paper referido, bastou fazer uma simples pesquisa pelo nome do paper.

### 2. Principal problema tratado

O principal foco deste paper é notar as diferenças relativamente a segurança dos utilizadores, equipamentos (recursos) e dados entre dois tipos de GRID (Globus-based Grids e Grids do tipo SETI@home ou climateprediction.net). O Paper também aborda questões importantes e pertinentes, tais como "O que acontece com os dados que coloco na GRID?", e a resistência dos utilizadores em utilizarem sistemas de computação como este.

Em Globus-based Grids, o modelo se segurança usado é baseado na Grid Security Infrastructure (GSI), modelo esse que fornece políticas de segurança tanto locais como em VO's (Virtual Organizations), através de protocolos de autenticação e autorização. Essa autenticação deve ser persistente entre vários hosts, em vários domínios (para permitir escalabilidade), requerendo também a identidade dos utilizadores.

Estes requisitos levaram à adopção de autenticação mútua, através do uso de chaves públicas (PKI - Public Key Infrastructure). O uso de chaves públicas numa Grid obriga que tanto o utilizador como os resources registem a sua chave pública num CA (Certificate Authority) de confiança, para assim obterem um certificado, que é usado para autenticação perante a Grid.

Essa autenticação é feita em dois "tipos de agentes": Resource Brokers e Resource Providers. O Resource Provider, como o próprio nome indica, fornece recursos à Grid. O Resource Broker é responsável pela distribuição e agendamento de jobs aos recursos computacionais.

Quando a autenticação é feita perante os Resource Providers, estes criam uma conta local temporária para o utilizador, que é controlada conforme as políticas de segurança locais previamente definidas.

Quando a autenticação é feita perante os Resource Brokers, estes criam um proxy certificate, que permite ao Resource Broker agir em nome do utilizador. Desta forma, o utilizador pode submeter um grande conjunto de jobs, ou até mesmo submeter jobs que demorem muito tempo a serem executados, e o Resource Broker gere estas situações da melhor maneira, como por exemplo, escolher qual o melhor destino (Resources disponíveis) para enviar os jobs. Importante referir que os proxy certificates gerados normalmente têm um lifetime baixo, apropriado para a tarefa submetida.

No segundo Case Study presente no artigo (climateprediction.net), o modelo de Grid usado utiliza o poder computacional dos PCs dos utilizadores comuns, para computar modelos climáticos, libertando assim supercomputadores desse tipo de tarefas.

Os utilizadores, voluntariamente, instalavam um software no seu computador, tornando-se assim um Resource Provider. Se se multiplicar este gesto por milhares de máquinas existentes, cria-se assim um Grid bastante flexível e de baixo custo.

A principal ameaça a este tipo de topologia é o facto do tempo do projecto terminar sem que haja dados suficientes para para serem usados. Para resolver esta ameaça, e garantir igualmente a segurança do utilizador "provider", foi definido que os participantes deveriam juntar-se à Grid em grande número, e que tinham de se manter até o projecto terminar. Para encorajar e garantir essa permanência, foram desenvolvidas ferramentas de visualização de resultados, e assinados digitalmente os pacotes instalados, para que assim transmitisse alguma confiança em relação à entidade e veracidade do projecto.

Um outro problema deste tipo de Grid é o retorno de resultados falsos/adulterados por parte dos participantes. O que poderia motivar tal acto é, por exemplo, a competição entre os participantes sobre a quantidade de CPU time que cada um doou.

**3. Bibliografia**

As fontes e referências usadas estão bem escolhidas, pois estas permitem ter uma boa noção sobre o estado e segurança existente em Grids.

A fonte mais recente referida no paper é de Março de 2007, o que actualmente pode-se considerar já antiga, mas visto que a data de publicação do paper é de Maio de 2007, as fontes e referências usadas estavam actualizadas. Importante referir que o paper revisto não descreve o state of the art actual.

**4. Organização do paper**

No geral, consideramos que o paper está bem organizado e é de fácil leitura. Os Case Studies apresentados permitem ao leitor ter uma melhor noção da realidade do uso de Grids, e de uma boa parte de toda a envolvência que lhe está inerente.

**5. Relevância**

No mundo informático actual, o tópico da segurança é um tema interessante e essencial, pois cada vez mais torna-se crucial manter a privacidade tanto dos dados tratados, como dos seus utilizadores.

No conceito de Grid, o tema Segurança torna-se ainda mais importante, isto porque como os jobs e simulações executadas na Grid podem conter informação sensível/confidencial, e o facto de estarem visíveis na Internet, pode tornar qualquer pessoa com acesso à internet como um potencial atacante.

O paper revisto mostra algumas dessas possíveis falhas, e chama a atenção à importância da implementação de Segurança em Grids.