Análise comparativa entre NIS e LDAP

Augusto Peixoto Bueno, Alexandre Carissimi

Instituto de Informática – Universidade Federal do Rio Grande do Sul (UFRGS)
Caixa Postal 15.064 – CEP: 91501-970 – Porto Alegre – RS – Brasil

apbueno@inf.ufrgs.br, asc@inf.ufrgs.br

Resumo. *O* Lightweight Directory Access Protocol *(LDAP) vem sendo usado em um número cada vez maior de aplicações distribuídas porém pouco se sabe sobre seu desempenho em ambientes de produção. Este trabalho apresenta uma análise do LDAP e suas vantagens em substituição ao* Network Information Services (NIS) *para a autenticação de usuários do serviço de e-mail (POP/SMTP AUTH) na rede do Instituto de Informática da UFRGS.*

1. Introdução

Um dos desafios no desenvolvimento de sistemas distribuídos é localizar e identificar recursos de forma não ambígua. Entende-se por recursos, computadores, serviços e até mesmo usuários vinculados a uma certa aplicação. Para que isso seja possível é necessário armazenar e, posteriormente, consultar uma série de informações sobre esses recursos. Essas informações devem, portanto, serem organizadas e armazenadas para permitir de maneira rápida e confiável a sua recuperação. Nesse contexto dois serviços são largamente utilizados: serviço de nomes e serviço de diretório.

O *Lightweight Directory Access Protocol* (LDAP) é um protocolo de acesso a diretórios que define várias operações para recuperação, atualização e remoção de informações em um diretório.Apesar do crescimento no seu uso, pouco se encontra sobre desempenho do LDAP em ambientes de produção. Alguns estudos foram feitos, mas, em sua maioria, são com o objetivo de marketing de um produto[5] ou foram feitos em ambientes controlados [1,6,8], poucos são situações reais de uso [2,4]. Este trabalho possui como objetivo avaliar o uso e o desempenho do LDAP no ambiente de rede do Instituto de Informática da UFRGS, quando comparado ao NIS (*Network Information Service*).

Este trabalho está organizado da seguinte forma. Na seção 2 são abordados alguns conceitos básicos sobre serviços de nomes e de diretórios. Na seção 3 é feito uma breve descrição do que é o LDAP e na seção 4 sobre o NIS. Na seção 5 é analisado comparativa entre o NIS e o LDAP mostrando os resultados obtidos neste estudo. Finalmente, na seção 6, discute-se as principais conclusões obtidas neste trabalho e os trabalhos futuros.

2. Serviços de Nomes e Diretórios

Muitas vezes um serviço de diretório é confundido com um serviço de nomes, mas esses são, essencialmente, diferentes. Um serviço de nomes é responsável pela tradução de um nome em seu atributo. Já um serviço de diretório é muito mais flexível, pois permite a obtenção de quaisquer informações sobre um recurso a partir de um atributo (característica) ou de um conjunto de atributos. O DNS (*Domain Name System*), o NIS (*Network Information System*) são, respectivamente, exemplos de serviços de nomes e de diretórios comumente empregados em ambientes de rede de muitas organizações.

Um conceito importante a ambos é o de espaço de nomes. Um espaço de nomes é a coleção de todos os nomes válidos em um determinado escopo e dividem-se em duas categorias: hierárquicos e *flat* (planos). Um espaço de nomes hierárquico é, por definição, infinito. Isso ocorre através da criação de diversos níveis, como por exemplo, em uma estrutura de arquivos com diretórios e subdiretórios. Já, um espaço *flat* é limitado pelo tamanho máximo de um nome. Entretanto, basta não limitar esse tamanho para tornar o espaço de nomes *flat* também infinito. Um espaço hierárquico facilita o armazenamento e a organização de informações o que possibilita, em relação ao *flat*, delegar a administração de diferentes contextos de nomes a diversas pessoas ou organizações.

É comum confundir um serviço de diretório com um banco de dados, já que ambos guardam informações. De fato, um diretório é um tipo de banco de dados, mas com uma finalidade específica, focada para sua aplicação. Assim, para cada uso que um diretório pode ter, é necessário criar uma estrutura única para esse uso. Isso difere dos bancos de dados convencionais, onde a estrutura é padronizada, cabendo a quem desenvolve a aplicação relacionar adequadamente as entidades. Além disso, um banco de dados é otimizado para tratar tanto leituras quanto escritas, enquanto um diretório é voltado para leitura de informações, considerando que atualizações são menos freqüentes. Outra diferença entre diretórios e bancos de dados é que estes últimos normalmente suportam o conceito de transações, ao passo que os diretórios não o fazem obrigatoriamente.

3. O *Lightweight Directory Access Protocol* (LDAP)

O LDAP [7] é um protocolo assíncrono baseado em troca de mensagens. Apesar de ser referenciado como serviço de diretório, o LDAP é, de fato, um protocolo de comunicação da camada de aplicação. O LDAP segue o modelo cliente-servidor e, de uma maneira geral, a comunicação entre clientes e servidores LDAP ocorre da seguinte forma. Inicialmente, o cliente estabelece uma sessão com o servidor. Esse processo é chamado de *binding* (mapeamento). Nesse momento, o cliente pode ser ou não autenticado pelo servidor, assim como negociar um método de criptografia para tornar a comunicação segura. Uma vez autenticado o cliente executa operações de leitura e escrita sobre os dados do diretório, o que possibilita a recuperação e atualização das informações. Ao concluir suas operações, o cliente encerra a sessão com o servidor através da operação *unbinding*. O LDAP é implementado sobre o protocolo de transporte TCP.

Modelos do LDAP: Os serviços fornecidos por diretórios LDAP são organizados em modelos. A RFC2251 define dois: modelo de protocolo e modelo de dados. Entretanto, é comum se empregar a definição de quatro modelos proposta por Howes [3], a saber:

1. Modelo de informação: fornece a estrutura das informações armazenadas em um diretório LDAP. A unidade de informação é a entrada (objeto). A estrutura que define quais objetos existem para que as entradas possam ser instanciadas e os atributos de cada objeto é chamada *Schema*. É no *Schema* que também está associada a sintaxe de cada atributo, se é multi ou monovalorado, se o valor é obrigatório ou não.
2. Modelo de nomes: define a organização das entradas no diretório de forma hierárquica em uma estrutura denominada DIT (*Directory Information Tree*). As entradas são dispostas de acordo com seu *distinguished name* (DN) que serve como um identificador único. Um DN é uma seqüência de *relative distinguished names* (RDN) que juntos apontam para uma única entrada. Cada RDN é um ponto único na DIT.

3. Modelo funcional: define três categorias de operações para a manipulação de entradas do diretório independentemente de linguagens de programação: pesquisa, atualização e autenticação. A partir do LDAPv3 são incluídas operações estendidas.
4. Modelo de segurança: provê mecanismos que permitem aos usuários provar sua identidade (autenticação) e ao servidor controlar o acesso dos usuários autenticados (autorização). Inclui acesso anônimo, uso de par usuário e senha (senha "aberta" na rede) e a partir do LDAPv2 a utilização de Kerberos. O LDAPv3 introduz o uso de SASL além da definição de uma operação estendida do LDAP relativa à segurança, a *Extension for Transport Layer Security for LDAPv3*, baseada em SSL.

Replicação: A replicação em serviços de diretórios é importante por aumentar sua capacidade de expansão bem como seu desempenho e tolerância a falhas. Na especificação original do LDAP a replicação não era prevista., porém, atualmente, existe um grupo de trabalho do *IETF* padronizando do sistema de replicação do LDAP.

4. Network Information Service (NIS)

O NIS é um repositório de arquivos que normalmente é usado como ferramenta de administração de ambientes de rede cujo principal objetivo é manter a consistência das informações administrativas e gerenciais de uma rede. O NIS mantém suas informações em uma estrutura denominada *map*, organizada na forma de um par {*chave,valor*}. Ao ser utilizado no gerenciamento de uma rede, os mapas contêm informações sobre diversos aspectos dessa, tais como: usuários, senhas, grupos, usuários, entre outros. Os mapas são derivados de arquivos de configuração. Por exemplo, a informação contida em */etc/hosts* é usada para criar um mapa que possui o nome do *host* como chave e o endereço IP como valor. Os pares chave-valor, também conhecidos como registros, extraídos de */etc/hosts* criam o mapa *hosts.byname*. O NIS usa um espaço de nomes *flat*.

O NIS segue um modelo cliente-servidor onde os servidores são responsáveis pelo armazenamento dos mapas e por atender às requisições dos clientes. Existem dois tipos de servidores: mestre e escravo. O servidor mestre possui autoridade sobre os mapas e o(s) servidor(es) escravo(s) mantém réplicas. Além de aumentarem a disponibilidade, os servidores escravos ajudam no aumento do desempenho, pois são capazes de responder às mesmas requisições que o servidor mestre. Um NIS cliente faz requisições de consulta a mapas para os servidores e espera as respostas. Estas requisições são feitas via RPC (*Remote Procedure Call*) e não fazem distinção entre servidor mestre e servidor escravo.

5. Análise Comparativa NIS e LDAP

O objetivo é comparar o NIS e o LDAP para autenticação de usuários do serviço de e-mail (POP/SMTP AUTH) usando como métricas carga na rede e tempo médio de resposta. Essas métricas são usadas por serem dois parâmetros importantes de desempenho em qualquer protocolo de rede. Considera-se ainda o espaço ocupado em disco.

A rede do Instituto de Informática historicamente autentica seus usuários via NIS. Para empregar o LDAP converteu-se a base NIS existente para um diretório LDAP usando ferramentas de migração disponibilizadas no OpenLDAP. Nessa conversão foi utilizado o *schema* padrão *PosixAccount* [7] que oferece suporte as informações NIS. A base de usuários possui cerca de 1700 entradas (objetos), cada uma composta por 400 bytes.

Metodologia: nesse estudo considera-se apenas a carga gerada pelo servidor de e-mail quando os usuários do Instituto de Informática acessam seus e-mails via POP ou quando enviam mensagens usando SMTP AUTH (SMTP Autenticado). Outras autenticações (*login*, impressão, etc) continuam sendo feitas no servidor NIS original (servidor a parte). Inicialmente, com o servidor de e-mail do Instituto de Informática autenticando os usuários via NIS, foi feita a captura dos pacotes de dados e controle relacionados com o NIS. Após isso, o servidor de e-mail teve o método de autenticação alterado para usar LDAP e os dados foram novamente coletados. Nesse caso, duas coletas distintas foram feitas: com e sem a utilização do índice para os atributos *uid, uidNumber* e *gidNumebr* (a indexação é descrita mais adiante). Os dados foram coletados em horários considerados de pico de utilização, gerando, em média, quatro requisições de autenticação por segundo, tanto para o NIS quanto para o LDAP. Os resultados obtidos nessa experiência são validados estatísticamente com um intervalo de confiança de 95%.

Plataforma experimental: O servidor LDAP/NIS executa em um DELL PowerEdge 600sc, Pentium IV 2.4GHz, 1 GB RAM, 40 GB de disco rígido (SCSI) e barramento de dados de 533 MHz. O cliente é o servidor de e-mail do Instituto de Informática: DELL PowerEdege6400, biprocessador Xeon 700 MHz, 512 MB RAM, 27 GB de disco rígido (SCSI) em RAID 5 e um barramento de dados de 133MHz. O servidor LDAP utilizado foi o OpenLDAP 2.1.23 sobre uma máquina FreeBSD 4.9-stable. A base de dados usada foi a BDB (Berkeley DB v4.1). A mesma máquina executa o servidor NIS v1.31.2.1. O servidor de e-mail usa RedHat Linux 8 (kernel 2.4.23) e o daemon de mail é o *sendmail* 8.12.8. Todas as máquinas utilizadas estão ligadas a um switch de 100 Mbit/s.

Funcionamento básico do NIS e do LDAP para autenticação de usuários: O NIS utiliza RPC para troca de mensagens. Inicialmente uma conexão TCP é feita para requisitar ao serviço *portmap* a informação da porta a ser usada nas requisições RPC. Em seguida, o cliente faz uma ou mais requisições do tipo *match,* usando UDP, para recuperar as informações necessárias a autenticação, entre elas o mapa *group.byname*. Já, no LDAP, toda comunicação é feita em uma conexão. O cliente inicia uma sessão executando a operação *bind*. Em seguida, o cliente requisita ao servidor as informações necessárias para autenticação do usuário. Em resposta, o servidor envia apenas os atributos requisitados pelo cliente e não o contéudo inteiro da entrada associada a solicitação (como no NIS). Pode ser necessário, para atender todo o processo de autenticação, que o cliente faça mais de uma requisição mudando apenas o atributo desejado. O cliente encerra a sessão executando a primitiva *unbind*. Esses procedimentos são ilustrados na figura 1. No caso específico do LDAP existe a possibilidade de utilizar índices para otimizar o procedimento de busca de informações no diretório (análogo ao uso de chaves primárias e secundárias em banco de dados). A definição se um atributo é indexado ou não é feita no momento da definição da estrutura de diretório.

Análise de resultados: A tabela I fornece, com base nos dados coletados, a carga média na rede (em Mbits/s) gerada pelo NIS e pelo LDAP. No caso do LDAP foram considerados os acessos ao diretório com e sem a utilização dos índices para os atributos *uid, uidNumber* e *gidNumber*. Esses atributos foram escolhidos porque seus valores são os consultados no procedimento de autenticação de usuários.

Ao analisar a tabela I, percebe-se que o protocolo LDAP, tanto não-indexado como indexado, é mais eficiente na utilização da rede, apresentado uma carga média na rede

cerca de 75% menor que a do NIS. Isso se deve ao fato do tamanho médio dos pacotes LDAP ser cerca 71% menor que os pacotes NIS. Essa diferença, em parte, é devido ao NIS executar a cada autenticação de usuário uma *procedure* chamada *ALL*, a qual retorna todo o mapa *group.byname* (dados que representa todos os grupos de usuário existentes na rede). Já o servidor LDAP retorna apenas os atributos requisitados pelo cliente.

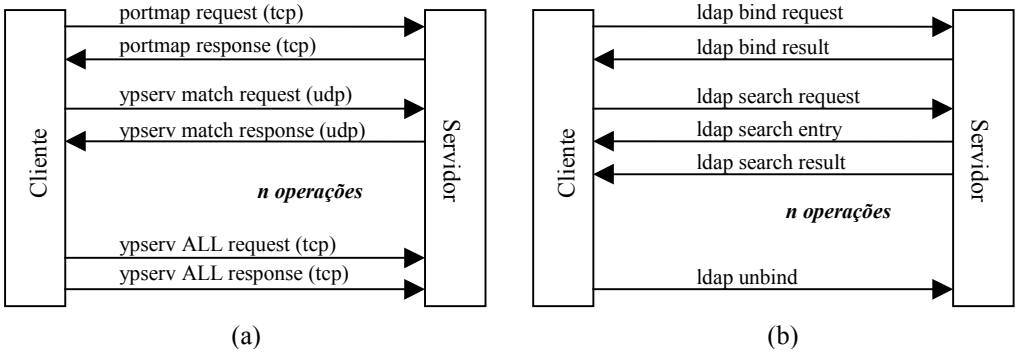

Figura 1 – Comunicação entre cliente e servidor: (a) NIS e (b) LDAP

Tabela I – LDAP versus NIS: carga média na rede em Mbits/s

LDAP (indexado)	LDAP (não indexado)	NIS
0.069	0.076	0.32

A figura 2 analisa o mesmo conjunto de dados em termos de tempo médio de resposta. Novamente para o LDAP são considerados dois casos: com (LDAP indexado) e sem (LDAP não indexado) a utilização de índices para os atributos *uid, uidNumber* e *gidNumber*. Percebe-se que o LDAP indexado teve um desempenho em torno de dez vezes maior que o NIS. Porém, o LDAP não-indexado apresenta um desempenho aproximadamente 26 vezes menor que o caso indexado e 2,6 vezes menor que o NIS. Fica evidente o interesse em identificar quais atributos serão os mais frequentemente acessados em um diretório LDAP para torná-los atributos do tipo indexado.

Em relação ao espaço em disco, a base LDAP é composta por um conjunto de arquivos que representam a base de dados propriamente dita (entradas) e os arquivos de índices dos atributos indexados. Neste estudo, o tamanho total ocupado pelo LDAP foi de cerca de 12 Mbytes contra 7.5 Mbytes ocupados pelas informações mantidas pelo NIS.

Considerações gerais: O LDAP é um padrão, o que facilita sobremaneira seu emprego em ambientes heterogênos. Por utilizar um espaço de nomes hierárquico, o LDAP possui uma escalonabilidade maior que o NIS (que emprega um espaço de nomes *flat*), além de permitir um particionamento da base de dados (subárvores) em diversas máquinas o que contribui para uma descentralização de serviços e conseqüente balanceamento de carga entre servidores. Aproveitando características similares de diretórios a banco de dados, a centralização de informações permite um melhor controle sobre usuários e recursos da rede. Um usuário cadastrado em uma base LDAP pode herdar, por ser um objeto, uma série de configurações básicas que, se bem aproveitadas, auxiliam em aspectos de segurança da rede, com por exemplo, permissões, controle de acesso a recursos, etc. Finalmente, o LDAP suporta facilmente criptografia através do emprego de TSL/SSL, ao passo que o NIS trafega todas suas informações em texto claro

na rede. É verdade que o NIS+, agrega proteção a comunicação entre clientes e servidores porém, além de não ser facilmente configurável, apresenta alguns problemas de compatibilidade em ambientes UNIXes de diferentes fabricantes. A principal desvantagem do LDAP é a necessidade de planejar a estrutura do diretório (DIT), suas entradas e índices. Uma estrutura muito específica pode dificultar modificações e gerar, com diversos níveis, um custo de processamento mais elevado para atender uma requisição.

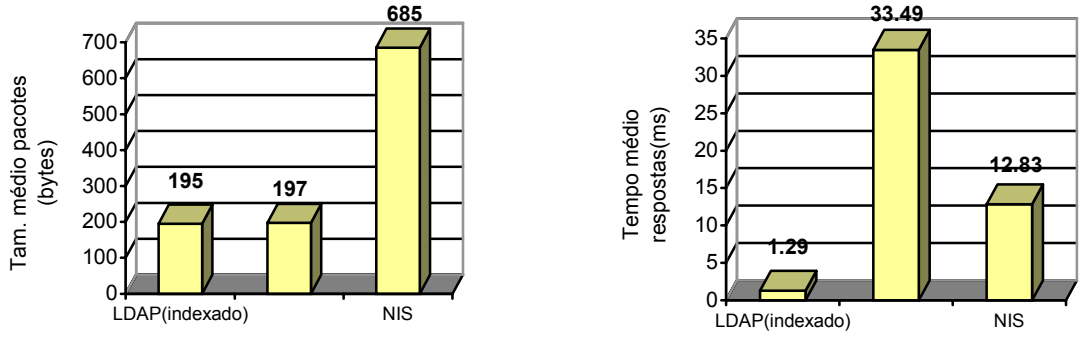

Figura 2 – LDAP versus NIS: tamanho médio dos pacotes

6. Conclusões

Este trabalho faz uma análise do LDAP como opção ao NIS na autenticação de usuários do serviço de e-mail. O LDAP apresentou melhor tempo de resposta, menor tráfego na rede porém, um consumo de espaço em disco pouco maior para armazenar as informações de usuários da rede. Foi mostrado que a indexação do diretório deve ser planejada com cuidado, pois o desempenho do serviço de diretório LDAP é afetado pela escolha dos índices. Como trabalhos futuros pretende-se investigar a carga sobre o servidor e o impacto do custo de processamento no uso de criptografia .

7. Referências Bibliográficas

[1] Choi, J.; et alli. Performance Evaluation of Two OpenLDAP Directory Server Back-end Designs. ftp://ftp.openldap.org/incoming/perfeval.zip (Acesso em: 04/01/04)

[2] Dixon, W.;. et alli. An Analysis of LDAP Performance Characteristics. http://www.crd.ge.com/cooltechnologies/pdf/2002grc154.pdf (Acesso em: 29/12/03)

[3] HOWES, T.; SMITH, M.; GOOD, G. Understanding and Deploying LDAP Directory Services. 1st ed. Macmillan Technical Publishing, 1999

[4] Klasen, N. Directory Services for Linux in comparison with Novell NDS and Microsoft AD. http://www.daasi.de/staff/norbert/thesis/html/thesis.html. (Acesso em: 03/01/04).

[5] Mindcraft Inc. LDAP Directory Server Comparison: Netscape and Novell. http://www.mindcraft.com/perfreports/ldap/netscape/dirserver10.html (Acesso: 4/01/04).

[6] Snyder, J. Sizing up LDAP servers. Network World, 2000. Disponível em: http://www.nwfusion.com/reviews/2000/0515rev2.html. (Acesso em: 28/12/03).

[7] WAHL, M.; HOWES, T.; KILLE, S. Lightweight Directory Access Protocol (v3): RFC 2251. [S.L]: Network Working Group, 1997.

[8] WANG, X.; Schulzrinne, H; Kandlur, D.; Verma, D. Measurement and Analysis of LDAP Performance. International Conference on Measurement and Modeling of Computer Systems (SIGMETRICS'2000), Santa Clara, CA, 2000, p. 156-165.