Estrutura de diretório LDAP multi-campi utilizando OpenLDAP

Clever de Oliveira Júnior*, Paulo E. M. Almeida*

**Departamento de Recursos em Informática*
Centro Federal de Educação Tecnológica de Minas Gerais
Av. Amazonas, 5253, Belo Horizonte, MG, Brasil.

Email: clever@dri.cefetmg.br, pema@dri.cefetmg.br

Resumo – Este trabalho apresenta um modelo de distribuição de um diretório institucional entre os diversos campi de uma Instituição utilizando Software Livre. A medida que uma organização cresce, sua estrutura física, o número de sistemas e de unidades fisicamente distintas também tende a crescer, o que agrava problemas relacionados ao armazenamento de informações tais como perda, duplicidade, corrupção e indisponibilidade de dados, além de gerar retrabalho e esforço computacional para tentar manter a ordem com recursos improvisados. A partir desse cenário, houve a necessidade de armazenar informações de maneira mais eficiente, segura e com alta disponibilidade. Para solução deste problema foi adotado o OpenLDAP, que se mostrou uma solução em Software Livre bastante estável, robusta e segura.

Termos de Indexação – ldap, disponibilidade, unificação, estruturação, distribuição, software livre.

1. INTRODUÇÃO

Diante das mudanças de comportamento das sociedades e organizações, a partir da explosão das tecnologias de comunicação e informação, observa-se um crescimento na quantidade e diversidade de informações, muitas vezes inesperado, o que ocasiona vários problemas relativos a desorganização de dados.

Acompanhando esse fenômeno e o crescimento das instituições, o volume de dados gerado por uma instituição aumenta mais ainda e requer ações diretas para seu armazenamento apropriado, a fim de manter sua integridade, consistência e disponibilidade.

Com relação às instituições de ensino, o impacto desse crescimento pode ser ainda maior, visto que a distribuição desorganizada de informações entre as diversas unidades,

setores e coordenações é comum, o que causa a duplicidade, dificulta a integridade, a consistência e a segurança dos dados, relatado também por Kreutz et al [1]. A gerência torna-se difícil e além desses prejuízos, o esforço pessoal e computacional para coletar as informações fragmentadas nos diversos locais com diferentes formatos é grande e, necessariamente, não garante a integridade e consistência final dos dados. Outro problema é a pluralidade de contas com senhas diferentes, dificultando a utilização de sistemas com *login* único.

Conhecendo a importância vital das informações para as organizações e, sendo de responsabilidade da área de Tecnologia da Informação o seu armazenamento e disponibilização segura e íntegra, este trabalho apresenta uma implementação de uma estrutura distribuída e unificada de identificação de usuários (denominada conta institucional – *login*), utilizando diretório LDAP a fim de reduzir os impactos danosos causados pelo crescimento das organizações e seus dados.

Este artigo está organizado como se segue: na seção 2 será descrito informações relativas a estrutura unificada e na seção 3 a distribuição dessa estrutura, implementada em LDAP, pelas unidades do CEFET-MG. As questões sobre gerenciamento das informações armazenadas no diretório são tratadas na seção 4 e considerações sobre a estrutura são discutidas na seção 5.

2. ESTRUTURA UNIFICADA

A estrutura desejável para o armazenamento de informações institucionais é aquela que garanta, de maneira unificada, a integridade, a segurança e a disponibilidade dos dados. Entende-se, por uma base unificada, um repositório de dados hierarquicamente organizado, indexado, no qual cada objeto possui um identificador único, e que provê um canal único de acesso a esses objetos pelo usuários.

Em ambientes bastante diversificados como instituições de ensino, o desejo torna-se necessidade e a busca por uma estrutura unificada, que armazene as informações de todas as suas unidades, setores, coordenações e que são acessadas por um único canal é cada vez maior. No escopo desse trabalho, essas informações resumem-se em contas institucionais que agregam atributos como, por exemplo, nome completo do usuário, identificação única (*login*), senha, pasta pessoal, informações sobre correio eletrônico e controlador de domínio, entre

outras.

A implementação de uma estrutura unificada para contas evita duplicação, provê conta e senha única para diversos sistemas e serviços, melhora a disponibilidade e integridade dos dados, facilita cópias de segurança (*backup*), disponibiliza controle de acesso eficaz, otimiza a gerência e reduz retrabalho e esforços computacionais.

A estrutura adotada pelo CEFET-MG e relatada nesse trabalho utiliza o OpenLDAP, Software Livre que implementa um serviço leve de acesso a diretórios (*Lightweight Directory Access Protocol*) [4]. Diretórios LDAP são estruturas hierarquicamente organizadas, de propósito geral, otimizadas para leitura, que disponibilizam os requisitos exigidos para a estrutura proposta.

3. DISTRIBUIÇÃO MULTI-CAMPI

Instituições de ensino, que geralmente possuem diversos setores, coordenações e unidades fisicamente distantes, requerem necessariamente, além da unificação das informações, a sua distribuição entre locais e o correto controle de acesso, de acordo com as responsabilidades de cada administrador local.

Por ter um ambiente com alguns setores e coordenações que possuem uma estrutura local de usuários e sistemas, além das unidades fisicamente distintas, o CEFET-MG demanda uma estrutura unificada e distribuída para armazenar de forma íntegra, consistente, segura e disponível as suas contas institucionais, com controle de acesso rígido. A **Figura 1** exibe um fragmento da árvore do diretório institucional do CEFET-MG com suas unidades organizacionais. As permissões de acesso são definidas de acordo com a responsabilidade dos administradores locais e delegadas pelo orgão responsável pela Tecnologia da Informação da instituição, o Departamento de Recursos em Informática (DRI). Então, esforços foram somados para definir essa estrutura que será descrita a seguir.

Figura 1: Fragmento da árvore institucional do CEFET-MG

Cada setor ou coordenação que possui um número suficiente de funcionários e uma estrutura mínima para manter o funcionamento de sistemas e serviços é definido como ramo distribuído e é responsável pelos dados contidos nesse ramo. Uma amostra das informações contidas em um ramo é exibida na **Figura 2**.

Figura 2: Fragmento do ramo de um departamento

Operações como inclusão, alteração e remoção de dados serão exclusivamente realizadas pelo administrador do ramo. Utilizando um recurso nativo de replicação do OpenLDAP, as informações desse ramo serão replicadas para a raiz da instituição e acessível por qualquer local do CEFET-MG, em uma relação provedor/consumidor (*Provider/Consumer*), conforme a **Figura 3**. Como característica desse tipo de relação, os dados replicados para o consumidor (servidor raiz) permanecerão nele com permissão somente leitura, o que ajuda a garantir que os dados permanecerão íntegros e consistentes em relação aos dados armazenados nos nós locais – provedores (servidores de ramo), pois os administradores raiz não terão acesso de escrita nessas bases, assim como os administradores dos servidores de ramo também não terão acesso de administração ao banco de dados das outras bases, pois elas não estarão armazenadas em suas máquinas.

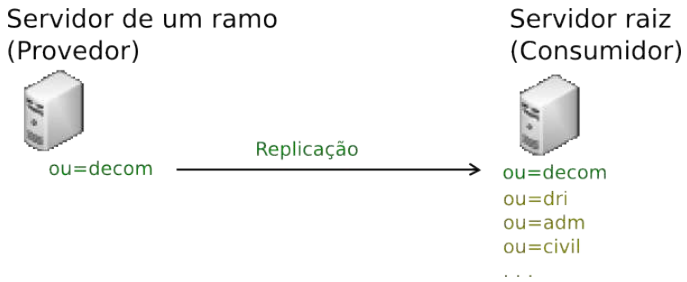

Figura 3: Replicação de dados entre Provedor e Consumidor

A raiz da instituição também é distribuída, mas somente entre as unidades fisicamente distintas, que possuem uma cópia de todos os objetos, tanto locais quanto replicados, veja **Figura 4**. Essa distribuição também é do tipo provedor/consumidor, ou seja, alterações realizadas serão propagadas para os consumidores, que exigirão um grau maior de responsabilidade, pois armazenarão todos os objetos dos usuários, além de disponibilizarem acesso aos serviços da organização.

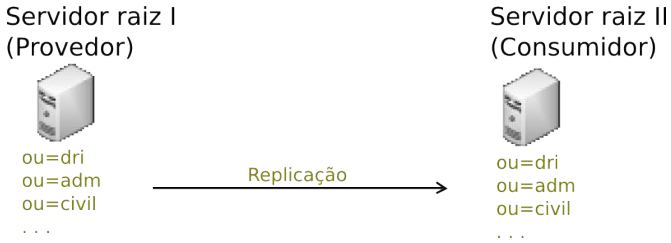

Figura 4: Replicação de dados entre servidores raiz

Para a implementação da estrutura completa, foi necessário dividir as bases locais e remotas em contextos separados para promover a possibilidade de escrita nos ramos locais do servidor raiz. No caso padrão do OpenLDAP, as bases são armazenadas em banco de dados BDB (Berkeley *Database*) [2] e cada tipo de base foi direcionada para diretórios distintos, ou seja, bancos de dados individuais. Os servidores de ramo possuem somente as suas bases BDB e os servidores raiz, além das suas bases, têm também as bases de todos os ramos replicados, ilustrado na **Figura 5**. Esse arranjo provê um isolamento de dados em relação aos servidores de ramo, aumentando a segurança e, possivelmente, o desempenho, pois os índices são menores do que em uma base única.

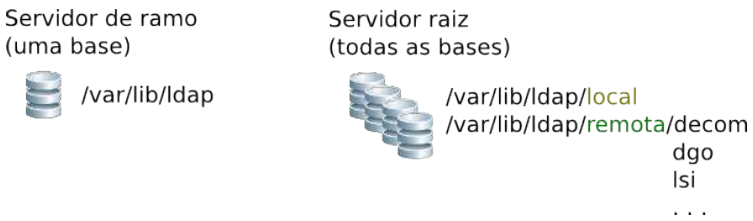

Figura 5: Distribuição das bases de dados

Adotando a replicação provedor/consumidor e isolamento das bases de dados, pretende-se satisfazer, com a estrutura completa, vários requisitos como a autonomia no gerenciamento de informações pelas unidades, setores e departamentos, a unificação, a integridade, a segurança e a disponibilidade dos dados armazenados. A estrutura unificada e distribuída é exemplificada na **Figura 6**.

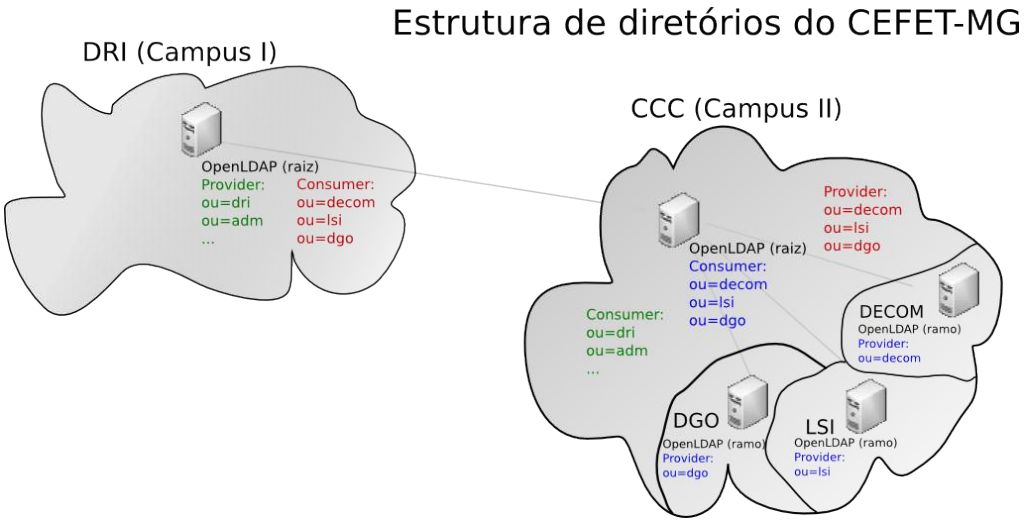

Figura 6: Estrutura unificada e distribuída

Alguns ajustes podem ser realizados para compatibilizar o modelo à estrutura da instituição. Dependendo do tamanho e necessidade de um determinado setor, o servidor de ramo pode sincronizar seus dados com um outro, o que aumenta a disponibilidade do diretório. Utilização de *no-breaks* e realização sistemática de *backups* também são recomendados para manter a integridade e disponibilidade das informações.

4. GERENCIAMENTO

Um fator muito relevante em estruturas de armazenamento de dados é o seu gerenciamento. A solução adotada para gerenciar os objetos armazenados no diretório distribuído foi o GOsa (Gonicus *Server Administration*) [3] que, além de ser Software Livre e ter um projeto ativo, reúne muitas funcionalidades. É desenvolvido em PHP e sua configuração é simples, entretanto não possui um pacote oficial de tradução para Português do Brasil. O Departamento de Recursos em Informática do CEFET-MG realizou a tradução das telas utilizadas pela instituição, criando um pacote de tradução[1] e submetendo-o ao desenvolvedor, que manifestou interesse em incluí-lo na próxima versão estável do *software*.

O GOsa organiza em abas a gerência de um conjunto de atributos relativos a um mesmo propósito, como pode ser visto na **Figura 7**. Entre as diversas abas disponíveis, pode-se citar o gerenciamento de usuários UNIX, correio eletrônico, controlador de domínio, telefonia, etc.

1 *Ticket* nº 611 aberto no *site* do projeto. Disponível em: https://oss.gonicus.de/labs/gosa/ticket/611. Acesso em: 18 mar 2009.

Figura 7: Gerência de atributos pelas abas

Ele suporta vários contextos, podendo administrar ramos do diretório separadamente com autenticação de usuário, recurso utilizado pelo CEFET-MG para disponibilizar acesso às unidades e setores. A **Figura 8** registra a tela de *login* e seleção do ramo, personalizada pela instituição.

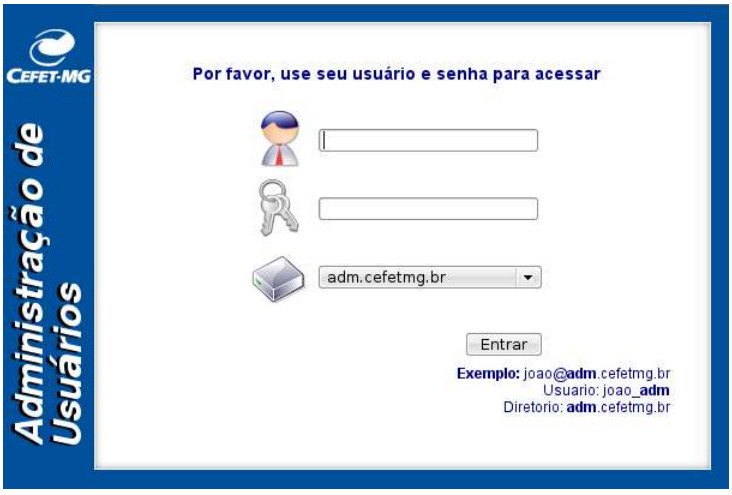

Figura 8: Tela de login e seleção do ramo

Para manter uma boa disponibilidade, a estrutura unificada e distribuída é administrada pelo GOsa hospedado com balanceamento de carga entre dois servidores WEB localizados em campus fisicamente distantes.

5. CONSIDERAÇÕES FINAIS

Apesar de ainda não estar em produção, a estrutura unificada e distribuída proposta por esse trabalho foi exaustivamente testada em laboratório, inclusive distribuída entre locais fisicamente distantes. Entre os diversos testes, foram realizadas simulações de falta de energia elétrica, falha na rede, lentidão causada por congestionamento de *links* e alterações *offline* de dados.

Por demonstrar estabilidade, robustez e segurança, o Departamento de Recursos em Informática irá migrar sua base LDAP para essa nova estrutura, que será distribuída, inclusive, entre as unidades do interior. Para garantir um melhor funcionamento e disponibilidade dessa proposta, o CEFET-MG está investindo na melhoria da rede de interconexão de dados, nos servidores e nos *no-breaks* de suas unidades.

REFERÊNCIAS

[1] KREUTZ, Diego Luís et al. **Unificando o Gerenciando de Informações de uma Universidade**. In: Encontro Nac. de Eng. de Produção, 24., 2004, Florianópolis.

[2] **GOsa** (Gonicus *Server Administration*). Disponível em: http://www.gosa-project.org. Acesso em: 18 mar 2009.

[3] **BDB** (Berkeley *Database*). Disponível em: http://www.oracle.com/database/berkeley-db. Acesso em: 18 mar 2009.

[4] **OpenLDAP** (*Open Lightweight Directory Access Protocol)*. Disponível em: http://www.openldap.org. Acesso em: 17 mar 2009.