**UNIVERSIDADE FEDERAL DE SANTA CATARINA**
**DEPARTAMENTO DE INFORMÁTICA E ESTATÍSTICA**
**BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO**

**Hendri Nogueira**

# Infraestrutura de Autenticação Única em Instituições de Ensino e Pesquisa Brasileiras

Trabalho de Conclusão de Curso submetido à Universidade Federal de Santa Catarina como parte dos requisitos para a obtenção do grau de Bacharel em Ciência da Computação.

Prof. Ricardo Felipe Custódio, Dr.
Orientador

André Bereza Júnior
Co-Orientador

Florianópolis, dezembro de 2010

# Infraestrutura de Autenticação Única em Instituições de Ensino e Pesquisa Brasileiras

Hendri Nogueira

Este Trabalho de Conclusão de Curso foi julgado adequado para a obtenção do título de Bacharel em Ciência da Computação e aprovada em sua forma final pelo Departamento de Informática e Estatística da Universidade Federal de Santa Catarina.

Prof. Luís Fernando Friedrich, Dr.
Coordenador do Curso

Banca Examinadora

Prof. Ricardo Felipe Custódio, Dr.

André Bereza Júnior

Prof. Olinto José Varela Furtado, Dr.

Káthia Regina Lemos Jucá, M.Sc.

*“Vai estudar moleque!”*
*Sebastião Sávio Nogueira*

Dedico este trabalho à minha Família, principalmente aos meus pais, que proporcionaram ótimas condições de educação, vida e sabedoria. Aos meus irmãos, que por serem mais velhos, me ensinaram tudo aquilo que não se devia fazer. E especialmente à minha namorada, por estar sempre ao meu lado durante essa longa trajetória, nos bons e nos maus momentos, independente da distância.

# Agradecimentos

Agradeço ao professor Ricardo Felipe Custódio pelo acolhimento, incentivo, e ensinamentos que proporcionaram a realização deste trabalho.

Um agradecimento especial aos meus amigos, principalmente do LabSEC, que me ajudaram na trajetória desses anos e se tornaram uma segunda família.

Agradeço ainda à Káthia Regina Lemos Jucá, pela ajuda e atenção na realização deste trabalho.

# Sumário

# Lista de Figuras

# Lista de Siglas

| | |
|---|---|
| **CAFe** | Comunidade Acadêmica Federada |
| **CSV** | Comma-separated values |
| **DAP** | Directory Access Protocol |
| **DIT** | Directory Information Tree |
| **DN** | Distinguished Name |
| **ETL** | Extract Transform Load (Extração Transformação Carga) |
| **IAA** | Infraestrutura de Autenticação e Autorização |
| **IANA** | Internet Assigned Numbers Authority |
| **IETF** | Internet Engineering Task Force |
| **JDBC** | Java Database Connectivity |
| **LDAP** | Lightweight Directory Access Protocol |
| **LDIF** | LDAP Data Interchange Format |
| **NPD** | Núcleo de Processamento de Dados |
| **OASIS** | Organization for the Advancement of Structured Information Standards |
| **ODBC** | Open Data Base Connectivity |
| **OSI** | Open Systems Interconnection |
| **RDN** | Relative Distinguished Name |
| **RFC** | Request for Comments |
| **RNP** | Rede Nacional de Ensino e Pesquisa |
| **SAML** | Security Assertion Markup Language |
| **SASL** | Simple Authentication and Security Layer |
| **SSL** | Secure Sockets Layer |
| **UFMG** | Universidade Federal de Minas Gerais |
| **UFSC** | Universidade Federal de Santa Catarina |
| **URL** | Uniform Resource Locator |
| **WAYF** | Where Are You From |
| **XSLT** | Extensible Stylesheet Language Transformations |

# Resumo

Os diversos serviços providos pela internet necessitam de bases de dados com informações sobre seus usuários, e os usuários necessitam de um cadastro para cada serviço que quiserem acessar.

Para amenizar o excesso de cadastros, criou-se o conceito de federação acadêmica no qual as informações pessoais são mantidas em uma base única (gerida por sua instituição de vínculo).

Toda federação acadêmica possui uma infraestrutura de autenticação e autorização (IAA). Ela define o modelo de comunicação entre as partes, os que provém a identidade e aqueles que fornecem os serviços.

No Brasil, surgiu a Federação CAFe, possuindo um conjunto de ideologias de várias outras federações espalhadas pelo mundo. Dentro das tecnologias utilizadas, se destacam o uso de serviço de diretório (como o LDAP) para armazenar e organizar as informações, e do Shibboleth que realiza a comunicação e autenticação única (*single sign-on*).

Sua estrutura permite cadastrar usuários, armazenando e gerenciando seus dados em uma única base de dados localizado na instituição (aquela no qual ele pertence) e acessar os recursos providos por outros domínios, criando assim uma autenticação única para todos os serviços providos dentro da federação.

**Palavras chave:** Federação CAFe, CAFe, Provedor de Identidade, Federação Acadêmica, single sign-on, Shibboleth.

# Abstract

The various services provided on the Internet requires databases with information about its users, and users need to register themselves for each service they want to have access.

To decrease the excess of entries, was created the concept of academic federation in which personal information is kept in a single database (managed by your institution bond).

Every federation has an academic authentication and authorization infrastructure. It defines the communication model between the parties, the ones that provide identity and those who provide services.

In Brazil, was created the CAFe Federation, getting a set of ideologies of many others federations around the world. Among the technologies used, the most important is the use of directory service (like LDAP) to store and organize information, and the use of Shibboleth to carry out the communication and a single authentication.

The structure allows the user register, store and manager of their data into a single database located in the institution (the one in which he belongs) and access the resources provided by other domains, creating a single authentication for all services provided in the federation.

**Key words:** CAFe Federation, CAFe, Identity Provider, Academic Federation, single sign-on, Shibboleth.

# Capítulo 1

# Introdução

A infraestrutura de autenticação única foi modelada para atender o conceito de Federação Acadêmica e integrar todas as instituições de ensino e pesquisa do país.

## 1.1 Contextualização

Federação CAFe (Comunidade Acadêmica Federada) é o resultado dos primeiros esforços para a implantação de uma federação acadêmica no Brasil. Sua meta é congregar todas as universidades e instituições de pesquisa brasileiras em uma rede de confiança, na qual cada instituição é responsável por autenticar e prover informações de seus usuários (alunos, professores, técnicos, funcionários, etc.) para provedores de serviços (aqueles que oferecem serviços de acesso restrito) autorizados.

## 1.2 Objetivos

### 1.2.1 Gerais

Este trabalho tem como objetivo geral, construir uma infraestrutura de autenticação e autorização dentro da UFSC (Universidade Federal de Santa Catarina), tornando-a um provedor de identidade na Federação CAFe.

### 1.2.2 Específicos

Os objetivos específicos deste trabalho são:

- Construir uma infraestrutura de acordo com o *Acordo de Participação na Federação (Provedor de Identidade)* [RNP 10b].
- Implementar uma base de dados centralizada.
- Fornecer o modo de autenticação única para os usuários.
- Tornar a UFSC um membro da Federação CAFe.
- Gerar documentação técnica a respeito do que é uma federação.

## 1.3 Motivação

As motivações para realização deste trabalho são:

- Uso de um único sistema de controle de acesso para serviços internos e externos à instituição.
- Uma única conta (login) por usuário para acesso aos serviços.
- Integração com todas as instituições cadastradas na Federação CAFe.
- Garantia de privacidade dos dados pessoais no controle de acesso aos serviços.

## 1.4 Metodologia

Para realizar este trabalho, foi necessário em primeiro lugar a ambientação do autor à federação em questão e à estrutura que a instituição possua, a fim de conhecer as exigências da Federação CAFe, funcionalidades das ferramentas utilizadas e as aplicações providas pela instituição.

Inicialmente, houve a realização de um conjunto de estudos sobre a federação em si, seus serviços oferecidos, ferramentas que serão utilizadas e as obrigações exigidas pela RNP (Rede Nacional de Ensino e Pesquisa) [RNP 10k]. Como exemplo destas ferramentas tem-se o LDAP, EID, EID2LDAP e Shibboleth.

Após o estudo, se inicia a construção da infraestrutura, com as instalações e configurações dos servidores e ferramentas, análise e construção de um novo banco de dados LDAP com base no banco de dados atualmente utilizado pela universidade.

O processo de configuração da infraestrutura é acompanhado com reuniões quinzenais pelo departamento responsável da RNP em conjunto com diversas instituições nacionais que também estão ingressando (ou já ingressaram) na Federação CAFe.

A parte de construção e população das bases de dados com informações das pessoas são supervisionados pelos responsáveis do NPD (Núcleo de Processamento de Dados) da UFSC [NDP 10].

Após a conclusão do ingresso da UFSC na federação, é necessário divulgar a nova estrutura para a comunidade acadêmica permitindo assim o seu uso.

## 1.5 Limitações do Trabalho

Uma das limitações do trabalho está nas condições em como as diversas bases de dados da UFSC estão espalhadas e interligadas. A descentralização das bases de dados e a dificuldade de obter algumas informações necessárias de todas as pessoas vinculadas à UFSC, como o e-mail, exigiu em primeira instância sua não utilização na importação da base LDAP.

O manuseio dos dados de diversas pessoas é um processo complicado, pois existem alguns atributos restritos quanto à permissão, é o caso das senhas de usuários. Isso não permitiu um funcionamento imediato da infraestrutura para utilização pelos usuários, deixando a resolução deste problema para o futuro.

# Capítulo 2

# Serviço de Diretório

Um diretório é uma lista de informações sobre objetos organizados ou catalogados em uma ordem, e fornece o acesso aos dados dos objetos. Permite que os usuários ou aplicações possam encontrar recursos no ambiente com características necessárias para um tipo de tarefa particular [TRI 07].

Para exemplificar, imagine uma lista telefônica, onde os objetos listados são pessoas. Os nomes são organizados em ordem alfabética e os endereços e o número de telefone são os detalhes fornecidos sobre cada pessoa.

Um diretório é um banco de dados especializado, que pode ser chamado de repositório de informação, cujos registros de dados são definidos em forma de objetos e armazenados de forma ordenada. Sua principal característica é a forma em que os registros (objetos) e suas informações são acessados, sendo o acesso de leitura e pesquisa maior que o de escrita, diferente de um banco de dados relacional, no qual, os dados são constantemente atualizados, adicionados ou excluídos.

Diretórios contêm otimizações para suportar grande volume de acesso de leitura, e seu acesso à escrita deve ser limitado à administradores de sistema ou ao proprietário de cada parte da informação. Permitem também, que pessoas ou aplicações localizem usuários, recursos, serviços e informações em ambientes distribuídos.

Os diretórios podem diferir entre si no modo como a informação é representada e acessada, na flexibilidade com que a informação pode ser pesquisada e

como pode ser estendida ou atualizada. Além da capacidade de controlar o acesso e autenticação, gerenciando o acesso às informações contidas nele.

Deste modo, um serviço de diretório é toda infraestrutura capaz de disponibilizar a informação contida no diretório. Esta infraestrutura é representada por softwares, hardwares, processos e políticas utilizadas para acessar e administrar a informação.

## 2.1 LDAP

LDAP (Lightweight Directory Access Protocol) é um protocolo de serviço de diretório executado em TCP/IP, ou seja, na rede. É um padrão aberto, produzido pela IETF [IET 10] (Internet Enginneering Task Force) e a RFC 2251 [GRO 97a] define esse protocolo.

Sendo o LDAP um protocolo de comunicação, ele define o transporte e o formato das mensagens usadas por um cliente para acessar os dados de um servidor de diretório de tipo X.500 [EHL 04].

Por definir um método para acessar e atualizar informações em um diretório, o LDAP obtém ampla aceitação como um método de acesso aos diretórios da internet, dando suporte para várias aplicações Web, intranet, navegadores, servidores IMAP, banco de dados, etc., além de fornecer mais funcionalidades para a rede de forma integrada com vários outros serviços como: Proxy, FTP, Apache, Samba, servidores de email, entre outros.

O primeiro protocolo criado para este fim foi o *Directory Access Protocol* (DAP). O DAP fazia parte das especificações X.500, desenvolvidas pela ITU Telecommunication. O DAP foi baseado no modelo OSI, que era extremamente difícil de ser implementado e resultava em aplicações complexas, lentas e de alto custo. Em 1993, a Força Tarefa de Engenharia da Internet (IETF) padronizou o LDAP como uma alternativa de acesso aos diretórios do X.500.

Existem várias implementações de servidores de diretórios, como por exem-

plo, MS Active Directory, IBM Lotus Domino e OpenLDAP, mas nem todas são compatíveis entre si, podendo apenas servir a um determinado *software* e possuir restrições de uso ou característica exótica.

O OpenLDAP, implementação mantida pela Fundação OpenLDAP, é um servidor LDAP de código aberto e de uso geral, ou seja, não agrega nenhum outro serviço que não tenha relação com a administração do diretório. Fundado em 1998, o projeto OpenLDAP foi baseado em uma implementação de servidor LDAP feita pela Universidade de Michigan. Deste modo, o OpenLDAP foi escolhido como servidor de diretório para o projeto e-AA [RNP 10j].

O projeto e-AA é coordenado pela RNP e seu objetivo foi implantar a Federação CAFe com soluções técnicas e ferramentas desenvolvidas tanto no contexto do projeto como também em iniciativas anteriores apoiadas pela RNP.

O LDAP define quatro modelos básicos que descrevem por completo a sua operação, quais informações podem ser armazenadas em diretórios LDAP e o que pode ser feito com essas informações. São eles:

- Modelo de Informação: define o tipo de informação que pode ser armazenada em um diretório LDAP.
- Modelo de Nomes: define como a informação no diretório LDAP pode ser organizada e referenciada.
- Modelo Funcional: define o que pode ser feito com a informação no diretório LDAP e como ela pode ser acessada e alterada.
- Modelo de Segurança: define como a informação no diretório LDAP pode ser protegida de acessos ou modificações não autorizadas.

### 2.1.1 Modelo de Informação

As informações no LDAP são estruturadas de forma hierárquica. As buscas sempre começam por um elemento raiz e são percorridas até achar o nó filho onde se

encontra a informação desejada. A raiz e os ramos são chamados de diretórios, que podem conter outros diretórios ou elementos chamados de entradas.

As entradas representam objetos de interesse no mundo real, como pessoas, servidores ou organizações. São compostas de coleções de atributos que contêm informações sobre o objeto. Todo atributo tem um tipo e um ou mais valores. O tipo do atributo está associado com uma sintaxe que especifica o tipo de valor que pode ser gravado. Por exemplo, uma entrada deve ter um atributo, e a sintaxe associada ao tipo do atributo deve especificar os valores possíveis para este atributo. Em adição, na definição dos dados que podem ser guardados como os valores de um atributo, uma sintaxe de atributo também define como estes valores se comportarão durante pesquisas e outras operações.

Alguns atributos possuem apelidos (*alias*) que podem ser utilizados como os nomes reais dos mesmos. Por exemplo, *brEduAffiliation* e *braff* representam o mesmo atributo, sendo *braff* o *alias* para *brEduAffiliation*.

Os valores dos atributos podem ser limitados quanto ao número que podem ser guardados ou para limitar o tamanho total de um valor, criando assim um vínculo de associação com tipos de atributos. Um atributo para guardar um número de CPF pode ser limitado a um valor único e um atributo que contém uma imagem pode ser limitado quanto ao seu tamanho.

### 2.1.2 Esquemas e Classes de Objeto

Os atributos LDAP podem armazenar qualquer tipo de informação, como nomes identificadores de usuários, senhas, e-mails, fotos, servidores, etc. Os nomes dos atributos e os tipos de informações suportados pelas entradas necessitam de algumas regras para se determinar e organizar a estrutura LDAP. O arquivo que contém as definições e a estrutura das entradas LDAP é denominado de esquema (*schema*).

Um esquema é como uma planta-baixa, uma definição da estrutura das entradas e dos atributos que podem ser inseridos nelas. Nos esquemas são definidas as classes de objetos (*objectClass*), que define os atributos ou o tipo de itens de dados

contidos em um tipo de objeto. Classe de objetos é um termo LDAP que determina o tipo de objeto que é representado por uma entrada do diretório ou registro. Alguns exemplos de objetos (entradas) são: *person*, *organization*, *organizationUnit*, *domainComponent*, *groupOfNames*.

O esquema pode usar *objectClasses* padrões como aqueles já definidos para países, organizações, grupos e pessoas, ou pode criar novas para atender requisitos específicos.

A definição de atributos é independente da definição de classe de objetos. Alguns exemplos são atributos típicos como *cn*, *sn*, *givenName*, *mail*, *uid* e *userPassword*. Como as classes de objetos, os atributos são definidos com OIDs únicos, com cada atributo contendo também um único número OID ligado a ele.

Os métodos de definição de classe de objetos para LDAPv3 são descritos nas RFCs 2251[GRO 97a] e 2252[GRO 97b]. A forma genérica de definição de classes de objetos é mostrada abaixo:

```
objectclass ( <OID da classe de objeto>
    [ NAME <nome da classe de objetos> ]
    [ DESC <descrição da classe de objeto> ]
    [ OBSOLETE ]
    [ SUP <OID da classe de objeto ancestral> ]
    [ ( ABSTRACT | STRUTURAL | AUXILIARY ) ]
    [ MUST <atributos obrigatórios> ]
    [ MAY <atributos opcionais> ] )
```

Cada classe de objeto começa com uma sequência de números delimitados por pontos. Estes números são referenciados como OID[1] (Object ID – número de controle) registrado na IANA[2] [IAN 10]; WHSP é uma abreviação de "white space" e apenas indica a necessidade de um espaço. Depois do OID, está o nome da classe

[1]OID é um identificador único de um objeto.

[2]IANA (Internet Assigned Numbers Authority) é a organização mundial reponsável pela coordenação global do DNS raíz, endereçamento IP, e outros protocolos dos recursos da Internet.

(NAME) seguido por uma descrição (DESC). Se a classe é subordinada a outra, a classe superior (SUP) é listada. Finalmente, a definição da classe de objetos especifica os atributos obrigatórios (MUST) e os opcionais (MAY).

Uma classe de objetos é declarada como abstrata (ABSTRACT), estrutural (STRUTURAL) ou auxiliar (AUXILIARY). O tipo abstrato é usado como modelo para criação de outras classes. Uma entrada do diretório não pode ser instanciada por uma classe de objeto abstrata e sim por classes estruturais. Uma classe de objetos auxiliar fornece um método para estender classes estruturais sem mudar a definição do esquema desta classe estrutural. Deste modo, uma classe auxiliar não pode ser a única a instanciar uma entrada do diretório. É obrigatório que em uma entrada do diretório haja ao menos uma classe estrutural.

Diferentes classes podem prescrever alguns atributos que se sobrescrevem, ou são redundantes com atributos de outras classes. Muitas classes de objetos são definidas em uma ordem hierárquica, onde uma classe é dita herdeira de outra superior. Considere o objeto LDAP, que é definido com as classes de objetos:

```
objectclass:  top
objectclass:  person
objectclass:  organizationalPerson
objectclass:  inetOrgPerson
objectclass:  posixAccount
```

A ordem mostrada para as classes de objetos acima indica uma relação hierárquica entre elas, mas não necessariamente. A classe *top* está no topo da hierarquia. Muitas outras que não são subordinadas a nenhuma outra, têm *top* como classe superior. *Person* é subordinada de *top* e requer que os atributos *cn* e *sn* sejam populados, permitindo vários outros atributos opcionais. Abaixo mostra a descrição do *objectClass person*:

```
( 2.5.6.6 NAME 'person'
DESC 'RFC 2256:  a person'
```

```
SUP top STRUCTURAL
MUST ( sn $ cn )
MAY ( userPassword $telephoneNumber $ seeAlso $
description ) )
```

*OrganizationalPerson* é uma subclasse de *person*, portanto uma classe herdeira, assim como a classe *inetOrgPerson*. Como exemplo, a classe *posixAccount* é subordinada à *top* e requer que os atributos *cn* e *uid*, dentre outros, sejam atribuídos. Perceba que isso se sobrepõe aos requerimentos para *cn* da classe *person*.

Não há a necessidade de guardar o atributo *cn* duas vezes pois ambas as classes requerem a presença de um atributo *cn*. Não é possível adicionar atributos sem valor ou apenas preenchidos com espaço, não havendo restrição em relação ao valor contido ou existência de uma exclusividade de atributos em relação às classes.

Descrição da objectClass *posixAccount*:

```
( 1.3.6.1.1.1.2.0 NAME 'posixAccount'
DESC 'Abstraction of an account with POSIX attributes'
SUP top AUXILIARY
MUST ( cn $ uid $ uidNumber $ gidNumber $ homeDirectory )
MAY ( userPassword $ loginShell $ gecos $ description ) )
```

### 2.1.3 Modelo de Nomes

As entradas são organizadas em forma de estrutura de árvore invertida, chamada de DIT (*Directory Information Tree*) ou árvore de informação do diretório que está exemplificada pela figura 2.1. O modelo de nomes define como estas entradas são identificadas unicamente, através de um DN (*Distinguished Name*).

Um DN é um nome único que identifica sem ambiguidade uma entrada específica. Tradicionalmente essa estrutura refletia os limites geográficos ou organizacionais. Entradas representando países aparecem no topo da árvore.

Aquelas encontradas abaixo delas representam os estados (Unidades Federativas) e em seguida podem ser encontradas entradas representando unidades organizacionais, pessoas, impressoras, documentos, etc.

DNs são feitos de sequências de RDN (*Relative Distinguished Name*), ou nome distinto relativo. Cada RDN em um DN corresponde a um ramo em uma DIT saindo da raiz até a entrada do diretório. Cada RDN é derivado de atributos de entradas de diretório.

De forma simplificada, um RDN tem a forma `<nome do atributo> = <valor>`. Um DN é composto de uma sequência de RDNs separados por vírgulas. RDNs podem ser multivalorados:

```
atributo = valor + atributo = valor.
```

As entradas em um diretório LDAP são identificadas por seus nomes e possuem as seguintes características:

- Possuem duas formas, uma representação por cadeias de caracteres e uma URL.
- Sintaxe uniforme.
- O limite do espaço de nomes não é evidente.

O nodo folha representado acima, cuja entrada “João Santos”, possui um RDN igual a `cn=João Santos` e seu DN é descrito assim:

```
cn=João Santos,ou=funcionarios,ou=pessoas,dc=ufsc,dc=br
```

Esta entrada, pode também possuir vários outros atributos, por exemplo:

```
cn:  João Santos
sn:  Santos
uid:  jsantos
mail:  joaosantos@mail.br
```

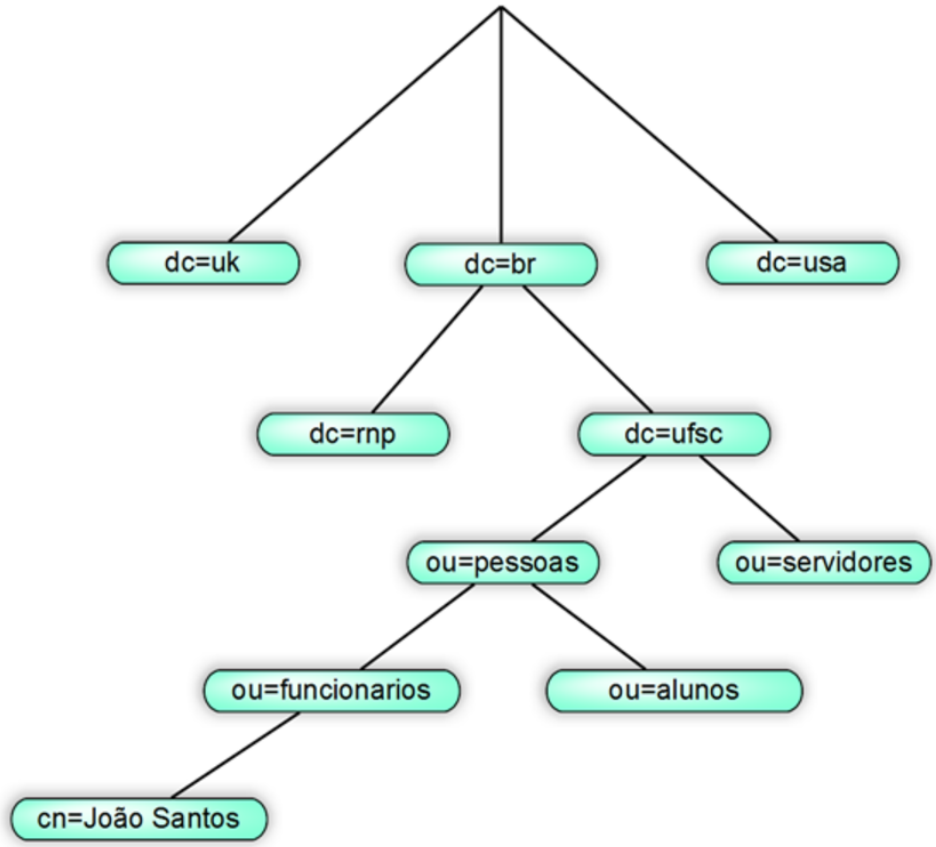


**Figura 2.1:** Árvore de Informação de Diretório – DIT

A sintaxe exata para nomes é definida na RFC 2253 [GRO 97c]. Os exemplos seguintes são DNs válidos escritos na forma de *string*:

```
cn=Joao Santos,dc=ufsc,dc=br
```

Este é um nome contendo dois RDNs:

```
ou=pessoas + ou=funcionários,o=ufsc
```

O modelo de nomes também tem uma representação por URL. O formato da URL LDAP tem forma geral descrita assim:

```
ldap://<host>:<porta>/<caminho>,
```

onde `<caminho>` tem a forma:

<dn>[?<atributos>[?<escopo>?<filtro>]]].

O <dn> é um nome distinto LDAP usando a representação em string. O <atributo> indica os atributos que devem ser retornados da(s) entrada(s). Se for omitido, todos os atributos serão retornados.

O <escopo> especifica o propósito da busca a ser feita. Pode ser uma entrada, um nível, entrada e filhos imediatos, ou uma sub-árvore inteira.

O <filtro> especifica o filtro de busca a ser aplicado às entradas dentro do escopo especificado durante a busca. O formato de URL permite a clientes de internet, por exemplo, navegadores web, terem acesso direto ao protocolo LDAP, e consequentemente ao diretório.

### 2.1.4 Modelo Funcional

O modelo funcional LDAP é composto por três categorias de operações que podem ser feitas em um servidor LDAPv3:

- Autenticação – Operações de *Bind*, *Unbind* e *Abandon* usadas para conectar a um servidor LDAP ou desconectar-se dele, estabelecer direitos de acesso e proteger a informação.

- Pesquisa – *Search* e *Compare* para pesquisar ou comparar entradas de acordo com o critério especificado.

- Atualização – *Add* para adicionar uma entrada, *Delete* para excluí-la, *Modify* para modificá-la e *ModifyRDN* para modificar seu RDN.

Existe uma operação de comparação que é utilizada para verificar se as entradas possuem um atributo com determinado valor. Se a entrada tem o valor, a operação *Compare* retorna VERDADEIRO; caso contrário, retorna FALSO.

A operação mais comum é a de pesquisa, bastante flexível e com algumas opções mais complexas, permitindo uma pesquisa através de alguma porção da DIT,

procurando informações de acordo com o critério especificado e listando os resultados. Não há distinção entre ler e listar.

As pesquisas podem ser gerais ou específicas. Elas permitem especificar um ponto de início dentro da DIT, a profundidade da busca, os atributos que uma entrada deve obter para ser considerada compatível, quais atributos retornados e ainda se os valores destes atributos devem ser retornados ou não.

Por exemplo, pode-se procurar na sub-árvore cuja raiz é `dc=ufsc,dc=br` pessoas cujos nomes sejam ”João Santos”, recuperando o endereço de e-mail de cada entrada achada. O LDAP permite que isso seja feito facilmente.

Para realizar uma busca ou pesquisa, os seguintes parâmetros devem ser especificados:

- Base;
- Escopo;
- Filtro de busca;
- Atributos para serem retornados;
- Limites.

A base é um DN que define o ponto de início da busca, ou seja, o nó da árvore dentro da DIT. O escopo especifica a profundidade da busca. Pode-se escolher entre três tipos: *baseObject*, *singleLevel* e *wholeSubtree*. *BaseObject* examina somente o objeto base, o *singleLevel* apenas as entradas filhas do objeto e o *wholeSubtree* o objeto base e todos os seus descendentes.

O filtro de busca serve para determinar um critério em que uma entrada deve se encaixar para que seja retornada na pesquisa e pode-se selecionar quais atributos devem ser retornados das entradas que se encaixam no critério de busca.

Um filtro é determinado da seguinte maneira:

```
<Atributo> <operador> <valor>
```

O `<operador>` são operadores lógicos que podem ser:

```
&   ⇒ (&(cn=joao)(sn=santos))
|   ⇒ (|(uid=joao)(uid=santos))
!   ⇒ (!(uid=joao))
=   ⇒ gidNumber=100
≃   ⇒ sn≃sant
>=  ⇒ uidNumber>=5000
<=  ⇒ Sn<=santos
*   ⇒ *
```

Os filtros também podem ser combinados com operadores lógicos para formar filtros mais complexo. Sua sintaxe é:

```
=   igualdade
>=  maior igual
<=  menor igual
≃   aproximação
=*  quaisquer caracteres
```

### 2.1.5 Modelo de Segurança

LDAP fornece um mecanismo para o cliente autenticar-se, ou comprovar sua identidade para um serviço de diretório e também suporta serviços de segurança de dados (integridade e confidencialidade).

As operações de autenticação são usadas para estabelecer e finalizar uma sessão entre um cliente e um servidor LDAP. A sessão pode estar segura em vários níveis, desde uma anônima, insegura (o cliente identifica-se pelo fornecimento de login e senha), ou até cifrada com mecanismos de SASL ou SSL.

A operação *bind* de autenticação é quando inicia uma sessão LDAP entre um cliente e um servidor assim que o cliente se identifica para o servidor. A operação *unbind* termina a sessão cliente-servidor e *abandon* permite ao cliente realizar o pedido de cancelamento de uma operação ao servidor.

A segurança deve ser considerada para cobrir os seguintes aspectos:

- Autenticação: garantir que o contato realmente é quem diz ser.
- Integridade: garantir que a informação enviada não se altere no meio do caminho.
- Confidencialidade: garantir que a informação seja revelada somente ao destinatário desejado.
- Autorização: garantir ao solicitante a permissão para realizar o que está solicitando. Até a versão 3 do LDAP, esse aspecto não fazia parte da especificação.

### 2.1.6 LDIF

LDIF (*LDAP Data Interchange Format* ou Formato de Intercâmbio de Dados do LDAP) é usado para representar entradas LDAP em um formato de texto simples. A RFC 2849 [GRO 00b] especifica seu formato.

Esse formato de arquivo é apropriado para descrever informações de diretório ou modificações realizadas. Ele é tipicamente usado para importar ou exportar informações de diretório entre servidores LDAP, ou para descrever um conjunto de modificações que é aplicado a um diretório.

Existem várias situações em que um formato de intercâmbio padrão é desejável. Por exemplo, exportar uma cópia dos conteúdos de um servidor para um arquivo, mover o arquivo para outra máquina e importar os conteúdos em um segundo servidor de diretório.

Além disso, usando um formato bem definido, o desenvolvimento de ferramentas de importação de dados é facilitado.

A forma básica de uma entrada LDIF é:

```
dn:  <nome distinto>
<atributo>:  <valor>
<atributo>:  <valor>
```

Existem duas construções para um LDIF:

- Descrição de conjuntos de entradas.
- Descrição de sentenças de atualização.

Um LDIF, cuja estrutura é a de conjuntos de entradas, contém todas as informações das entradas nele contidas, isto é, todos os atributos e seus respectivos valores estão presentes em cada uma de suas entradas:

```
dn:  o=ufsc
objectclass:  top
objectclass:  organization
o:  UFSC
description:  Universidade Federal de Santa Catarina

dn:  cn=Joao Santos,dc=ufsc,dc=br
objectclass:  top
objectclass:  person
cn:  Joao Santos
sn:  Santos
cn:  IGJlZ2lucyB3aXRoIGEgc3BhY2U=
cn:  <file:///tmp/arquivo
```

Quando se utiliza o LDIF para modificar uma entrada, ela é sobrescrita, ou seja, todas as informações da entrada no diretório são substituídas pelas informações no LDIF quando é feita uma operação de atualização. Os atributos que não existirem no arquivo LDIF, mas que existem na entrada do diretório, serão apagados.

Uma analogia deste tipo de operação é similar a sobrescrever um arquivo de um sistema operacional por outro arquivo, em que as informações do arquivo antigo deixam de existir e as novas informações assumem o posto.

Um LDIF estruturado em sequências de atualização possui apenas as informações relevantes para as modificações necessárias a uma entrada do diretório. Ao

invés de realizar a modificação da entrada como um todo, ele realizará a modificação em um único atributo de uma entrada. Sua forma básica segue abaixo:

```
dn:  <nome distinto>
changeType:  <Tipo da operação>
<operação>:  <atributo>
<atributo>:  <valor>
-
<operação>:  <atributo>
<atributo>:  <valor>
-
```

Para todos os tipos de LDIF, as entradas são separadas por uma linha em branco, enquanto para um LDIF de sequência de atualização, cada operação em um atributo diferente é separada por uma linha contendo um hífen (-).

```
dn:  cn=Joao Santos,dc=ufsc,dc=br
changetype:  add
objectclass:  person
objectclass:  inetorgperson
cn:  Joao
cn:  Joao Santos
sn:  Santos

dn:  cn=Joao Santos,dc=ufsc,dc=br
changetype:  modify
add:  givenName
givenName:  jo
givenName:  Joao
-
replace:  description
description:  Funcionario Joao
```

A sequência de comandos acima, demonstra um arquivo LDIF em que tem-se uma entrada que será modificada e o tipo de modificação que se deseja realizar.

Hoje em dia, há diversas ferramentas gráficas para a realização do gerenciamento de um servidor LDAP de forma simples e intuitiva, ao invés de realizar sequências de comandos Shell para cada operação.

Uma das ferramentas é o Apache Directory Studio. ApacheDS é um cliente LDAP feito em uma plataforma Eclipse, possui uma série de plugins e é uma ferramenta completa que pode ser utilizada para acessar um servidor LDAP local ou remotamente. Seu LDAP Browser[3] permite não apenas mostrar os dados como também criar, modificar, editar e remover entradas.

Outra ferramenta interessante é o phpLDAPadmin. Ele é definido em seu site oficial[4] como um "um navegador LDAP baseado na Web para gerenciar o seu servidor LDAP". Essa ferramenta é um cliente LDAP implementado em PHP, que pode ser acessada por navegadores Web. A sua visualização da DIT e sua avançada funcionalidade de busca ajudam a tornar mais intuitiva a administração do diretório LDAP.

[3]Ferramenta de navegação e visualização de um diretório LDAP

[4]http://phpldapadmin.sourceforge.net/

# Capítulo 3

# Federação Acadêmica

Federação acadêmica é um conceito que visa minimizar a manutenção de informações usadas para autenticação e autorização de pessoas para terem acesso aos serviços disponibilizados pelas instituições de ensino e pesquisa. Com a federação acadêmica, as informações sobre uma pessoa serão mantidas em uma única base, criando um vínculo entre a pessoa e a instituição.

Cada instituição necessita estabelecer seu modelo de gestão de identidade, ou seja, determinar a forma como as informações sobre as pessoas serão mantidas, atualizadas e os métodos de autenticação.

Aqueles que fornecem um serviço de acesso restrito precisam confiar no modelo de gestão de identidade das instituições para disponibilizar seus serviços.

Existem diversas federações acadêmicas implementadas em diversos países como: InCommon [INC 10], Feide [FEI 10], Switch [SWI 10] e UK [JAN 10].

## 3.1 Federação CAFe

A Comunidade Acadêmica Federada (Federação CAFe) [RNP 10f] teve início em 2008, como projeto piloto e com a meta de reunir todas as universidades e instituições de ensino e pesquisa brasileiras em uma rede de confiança. O projeto de criação da Federação CAFe inclui ainda o estudo, a proposição, a análise e a validação

de políticas para regular o funcionamento da federação (requisitos mínimos que provedores de identidade e de serviço deverão cumprir).

Assim como nas federações representadas nos diversos países, a Federação CAFe segue protocolos bem definidos na troca de mensagens entre provedores de identidade e serviço. O SAML (*Security Assertion Markup Language*) [MAL 04] é um protocolo adotado em várias federações inclusive na CAFe, que em conjunto com o software Shibboleth [INT 10], vem se firmando como um padrão *ad hoc* que permite a troca de informações de autenticação e autorização entre provedores de identidade e de serviço. A ferramenta Shibboleth será detalhada na seção 4.4.

Através de um componente denominado de WAYF (*Where Are You From*), que é centralizado e mantido pela RNP, os provedores de serviços poderão ser implantados por membros externos, ou seja, atuar apenas como provedores de serviços.

A rede de confiança permite que um usuário seja autenticado em sua instituição de origem e consiga acessar recursos e serviços oferecidos via *web* tanto pela própria instituição quanto por outros membros da federação, através de um único *login*.

### 3.1.1 Benefícios

Dentre os mais diversos benefícios oferecidos pela estrutura da federação, a participação de uma instituição como um provedor de identidade envolve:

- Uso de um único sistema de controle de acesso para serviços internos e externos à instituição.
- Uma única conta (login) por usuário para acessar todos os serviços.
- Manutenção de dados pessoais restrita à instituição de origem do usuário.
- Garantia de privacidade: apenas informações necessárias dos usuários serão passadas para os provedores de serviço.

Para provedores de serviços, os benefícios incluem:

- Simplificação do procedimento de controle de acesso, devido à:
  - Autenticação e disponibilização de informações sobre os usuários que são realizados pelos provedores de identidade, eliminando a necessidade de manu-
    tenção dessas informações no provedor de serviço.
  - Autorização para acesso a um recurso por um usuário que pode ser realizada pela característica do usuário como, tipo de vínculo, ou outro atributo disponibilizado pelo provedor de identidade.
- Redução de requisitos de suporte aos usuários.

## 3.2 Infraestrutura de Autenticação e Autorização

Uma federação possui uma infraestrutura de autenticação e autorização interdomínios ou também chamada de infraestrutura de autenticação e autorização federada. Seu objetivo é cadastrar um usuário, armazenando e gerenciando seus dados em uma única instituição (aquela no qual ele pertence), mas podendo ter acesso aos recursos oferecidos por outros domínios.

Uma infraestrutura de autenticação e autorização (IAA) federada é formada por dois tipos de provedores:

- Provedores de Identidades;
- Provedores de Serviços.

Os provedores de identidade são responsáveis por criar e manter cadastros e atualizar informações sobre as pessoas vinculadas à instituição, entre eles, os dados pessoais (nome, data de nascimento, CPF, nomes dos pais, sexo, etc.) e os vínculos internos (data de admissão, cargo ocupado, número de matrícula, número VoIP, etc.). O provedor de identidade necessita estabelecer seu método de autenticação interno e deve garantir que cada pessoa possua um identificador único.

Os provedores de serviços oferecem serviços em que seus acessos são restritos e ainda podem determinar privilégios de acessos baseados em atributos adicionais sobre o usuário, como por exemplo, o vínculo dele com a instituição (aluno, professor, técnico, etc.). Cabe ao provedor de serviço solicitar as informações adicionais (caso necessite) ao provedor de identidade para completar a autenticação.

O acesso de usuários de diferentes instituições nos provedores de serviços é comum, então há a necessidade de redirecionar os usuários para os respectivos provedores de identidade. O serviço centralizado e responsável por obter as informações sobre os provedores de identidade cadastrados na federação e seus respectivos redirecionamentos é o WAYF (*Where Are You From*). Nele o usuário seleciona sua instituição de origem, e passa a interagir com o seu provedor de identidade para fornecer as suas credenciais.

O WAYF contém o serviço de *Metadata*, do qual possui um arquivo metadado responsável por concentrar as informações dos provedores pertencentes à federação. Este é um arquivo de configuração padronizado e compartilhado entre os provedores de identidade e de serviço da federação.

O arquivo de metadado é disponibilizado pela federação e a sua função é estabelecer a relação de confiança entre os provedores, utilizando certificados digitais ou chaves públicas. Nesse arquivo são indicadas, para o provedor de identidade, as informações pertinentes sobre os provedores de serviço (e vice-versa). Desta forma, garante-se a segurança e a autenticidade na comunicação entre os provedores.

Além disso, o arquivo de metadados disponibiliza as informações relevantes para a comunicação entre os provedores, como identificadores, URLs e protocolos utilizados.

Ao acessar um serviço através do *browser* pela primeira vez, ele será redirecionado para WAYF e deverá selecionar o provedor de identidade (instituição) de origem. Após a seleção, navegador será redicionado para o *site* de autenticação de sua instituição. Após a autenticação do usuário, o provedor de identidade repassa o resultado da autenticação ao provedor de serviço e cria uma sessão de uso associada

ao usuário, de forma que acessos a novos serviços dentro de um determinado intervalo de tempo não gerem novas requisições de autenticação (*single sign-on*).

Para ilustrar melhor este processo, a figura 3.1 representa este processo de autenticação e informa os seguintes passos:

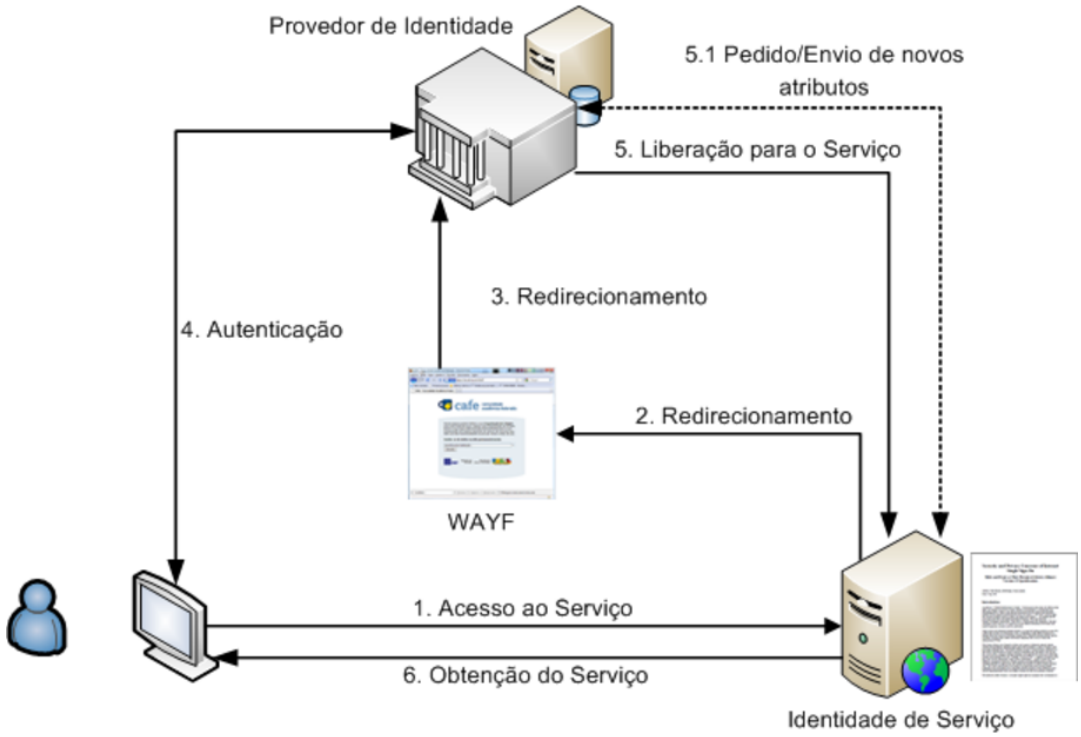


**Figura 3.1:** Fluxo de informações

1. O usuário acessa a um serviço desejado através do browser.

2. O servidor redireciona o navegador para o serviço de descoberta da federação (WAYF).

3. O usuário seleciona uma instituição oferecida na lista de instituições cadastradas, e seu navegador o redireciona para a página de autenticação da seleção.

4. O provedor de identidade da instituição envia ao navegador a página de autenticação do usuário, o qual se autentica através de *login* e senha.

5. O provedor de identidade gera um *handle* e o envia ao navegador, que o encaminha ao provedor de serviço, e assim obtém a prova de autenticação do usuário. Para algumas aplicações, isso é suficiente para autorizar o acesso do usuário ao serviço.

   (a) Opcionalmente, o provedor de serviço pode enviar um pedido para obtenção de novos atributos ao provedor de identidade, utilizando o *handle* para especificar o usuário em questão.

6. Devidamente autenticado, o provedor de serviço retorna a solicitação do usuário.

### 3.2.1 Responsabilidades

As responsabilidades que um provedor de identidade deve assumir para participar da federação são descritas em: *CAFe – Acordo de Participação na Federação (Provedores de Identidade)* [RNP 10b] . A instituição deve seguir algumas obrigações, tais como:

- Instalar e atualizar os *softwares* requisitado pela Federação no documento *Requisitos técnicos para membros da Federação CAFe* [RNP 10l].
- Garantir que dados fornecidos estejam corretos.
- Fornecer metadados corretos.
- Obter e instalar os metadados atualizados da federação a cada 30 dias, no mínimo.
- Aceitar as normas estabelecidas no documento *CAFe – Gorvernança da Federação* [RNP 10d].
- Indicar um representante oficial da instituição junto à federação.
- Disponibilizar informação sobre sistemas de gestão de identidade, atualizando-a no mínimo a cada renovação deste acordo, através do preenchimento do formulário *on-line*[1] .

[1] Disponível em: http://www.rnp.br/cafe/docs/questGI.php.

- Manter um sistema de gestão de identidade com os requisitos mínimos estabelecidos no documento *CAFe – Requisitos Mínimos de Gestão de Identidade* [RNP 10e].

- Receber e auxiliar equipe designada pela Federação CAFe para realizar auditoria na instituição.

Já os provedores de serviço devem assumir para participar da federação os requisitos descritos em *CAFe – Acordo de Participação na Federação (Provedores de Serviço)* [RNP 10c]. Em particular, o provedor deverá comprometer-se a respeitar a privacidade das informações recebidas de participantes da Federação, não armazenando, divulgando ou usando essas informações para propósitos diferentes do controle de acesso a recursos, a menos que explicitamente autorizado a tal por esses participantes.

# Capítulo 4

# Estrutura do Provedor de Identidade

Um provedor de identidade deve construir uma infraestrutura apta para cumprir com todas as regras determinadas pela Federação CAFe e que possa ter um ótimo funcionamento.

Dentre os diversos aplicativos contidos nela, esta seção descreverá sobre aqueles com maior importância: esquema brEduPerson usando no LDAP, os aplicativos EID, EDI2LDAP, Shibboleth.

## 4.1 Estrutura do Esquema brEduPerson

BrEduPerson é um esquema modificado para armazenar informações específicas para a realidade do país, como informações genéricas de qualquer cidadão residente no Brasil (CPF), informações gerais sobre os membros de uma instituição (e-mail, cargo, etc.), além de informações específicas sobre os funcionários e alunos destas instituições.

Este esquema é também utilizado integrado com os esquemas *inetOrgPerson*, *eduPerson* e *SCHAC*.

Dentro de uma instituição de ensino e pesquisa, há a necessidade de modelar relacionamentos entre conjuntos de informações. Deve suportar modelagens de pessoas que possam desempenhar diferentes papéis de aluno, a cada um dos quais está

associada uma data de ingresso, um código de curso, e outras informações, ou que uma mesma pessoa pode obter vários números VoIP, cada um com características distintas. Esse modelo de relacionamento é realizado de forma hierárquica.

Os nós em um diretório LDAP formam uma árvore. Cada nó, independentemente de ser pai de algum outro nó na árvore, é uma entrada com suas próprias informações (atributos). Esses nós são chamados de *containers* na terminologia X500.

Uma pessoa com alguma ligação com a instituição é o item principal, contendo informações genéricas, como nome, CPF, gênero, data de nascimento, e as demais informações de relacionamento dela com a organização (aluno, professor, técnico, etc.) são tratados como *containers* interligados abaixo dele.

Além disso, abaixo da entrada com os dados gerais, podem aparecer diversas entradas descrevendo telefones VoIP, dados biométricos, etc.

Este modelo de estrutura tem a consequência de que as classes que em princípio não precisariam ser definidas como estruturais, passam a ser definidas dessa forma para permitir que suas informações apareçam em uma entrada independente.

A entrada principal de cada pessoa é definido pelo objeto de classe estrutural *inetOrgPerson* e possuindo classes auxiliares *schacPersonalCharacteristics*, *eduPerson* e *brPerson*.

A classe *inetOrgPerson* é definida na RFC 2798 [GRO 00a] e *schacPersonalCharacteristics* é parte da proposta da TERENA (*Trans-European Research and Education Networking Association*) [TER 09]. *EduPerson* é mantida pelo grupo MACE da Internet2 [IMAC 08] .

*BrPerson* tem o objetivo de capturar atributos específicos de pessoas que vivem no Brasil. Abaixo dessa entrada principal, é recomendado existir pelo menos uma entrada de classe estrutural *brEduPerson* que descreve um vínculo da pessoa com a instituição.

Cada vínculo é descrito por uma entrada separada, e pode haver um número arbitrário de tais entradas, refletindo em diferentes vínculos. Classes auxiliares podem vir a ser definidas para incluir atributos relativos a cada tipo de vínculos. A figura 4.1

demonstra o modelo de nomes da estrutura proposta pela Federação CAFe.

A figura A.1 , localizada na seção apêndice, apresenta um exemplo mais detalhado de um *container* da estrutura hierárquica. Cada objeto de classe *brEduPerson* representa os vínculos (aluno, funcionário, professor, pesquisador, etc.) e ainda as posições temporárias (coordenador, diretor, etc.).

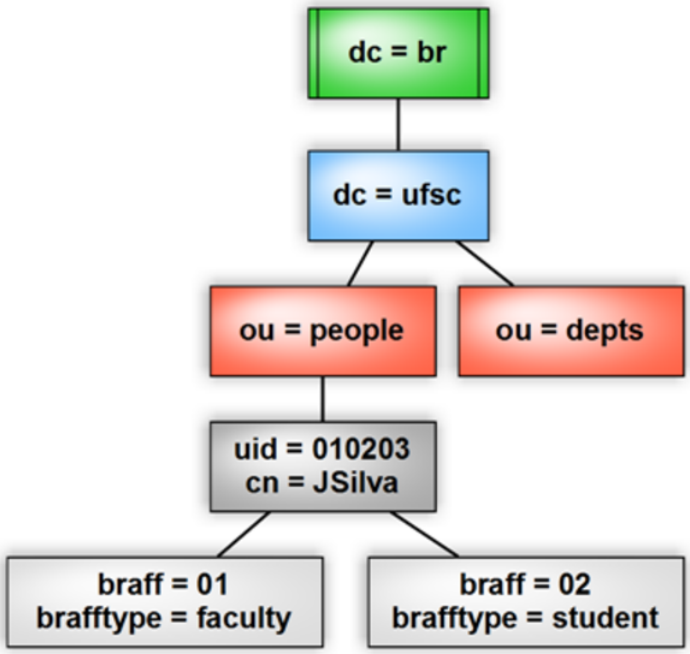


**Figura 4.1:** Modelo de nomes da estrutura da CAFe

O atributo *brEduAffiliationType*, cujo apelido é *brafftype* informa o nome do papel que esta pessoa realiza na instituição, podendo seu valor ser: *student*, *faculty*, *employee*, *alum*, *other* (contrato de outro tipo), *position* (coordenador, reitor, etc.) e *scholarshioAwardee* (iniciação, pós, etc.).

## 4.2 EID

Diretórios que possuem um baixo fluxo de cadastros de pessoas podem ser facilmente gerenciados pela inclusão e exclusão manual de registros. Já os diretórios com muitos usuários e comportamento dinâmico necessitam de um esforço maior na

manutenção, o que deixa seu gerenciamento manual mais complicado, como os diretórios acadêmicos.

O objetivo do EID (Export Import Directory) [UFM 10a] é facilitar a integração de dados de diversos sistemas para construir um metadiretório (uma base relacional intermediária entre as fontes efetivas dos dados e o diretório) e em seguida um ou mais diretórios, ou seja, ferramenta para auxiliar nas funcionalidades de importação e exportação de dados para outras fontes. O conceito de metadiretório será descrito melhor na seção 4.2.1 .

Desenvolvido pelo Grupo São Tomé da UFMG, teve base na ferramenta PingifesImport[1], ferramenta de extração, transformação e carga (ETL), utilizadas pelas instituições de ensino superior para alimentação do modelo de dados (PingIFES) definido pelo MEC.

A ferramenta EID permite importar dados de bancos relacionais, desde que exista um drive JDBC ou ODBC para fazer a conexão e/ou arquivos de texto CSV. O EID disponibiliza um serviço *web* para exportação e consulta de dados, o que facilita o acesso por aplicações que utilizem tecnologias diversas.

O serviço *web* serve de base para outras ferramentas de exportação também, como o EID2LDAP. Os dados importados são associados às pessoas e os registros completos podem ser facilmente recuperados através da interface *web* do EID.

O EID utiliza o conceito de Objeto (*EidObject*) para representar as informações que armazenam. Um objeto é uma entidade que possui um identificador único e um conjunto de atributos, sendo a unidade mínima de armazenamento de informações. São considerados objetos: pessoas e definições de grupos.

Os atributos são mapeamentos nome-valor, onde o valor possui um tipo ou um domínio definido. Os nomes e os tipos dos atributos são especificados em entidades denominadas classes. As classes são definições de agrupamentos de atributos. Cada classe pode ser considerada uma definição de um tipo de dado composto.

Atribuir valores aos atributos definidos pela classe e sua associação a um

[1] http://pingifes.mec.gov.br/pingifes/

objeto é o processo denominado de instanciação da classe. Um objeto pode estar associado a várias instâncias de uma mesma classe ou de classes diferentes, mas não aos atributos individualmente. As classes podem ser definidas de acordo com as necessidades do utilizador.

Todo objeto é identificado globalmente através do GUID, que é gerado automaticamente pela ferramenta. Esse atributo é definido pela classe especial *EidObject*.

O GUID é gerado ao importar a primeira classe denominada de Identificação (possui informações básicas referente às pessoas). Sua geração é realizada a partir da escolha de um atributo que seja único para cada indivíduo, tornando assim um ponto de referência na base.

As demais classes extraídas usarão o atributo único para conciliar com o GUID determinado, interligando o registro principal sobre a pessoa com as funções e seus atributos, como funções que ela exerce na instituição (aluno, professor, técnico, etc.), e-mails que a pessoa possui, endereços, telefones para contato, VoIPs, etc.

Toda classe criada na aplicação gera uma tabela no banco de dados. A classe *EidObject* se relaciona com as demais classes do sistema denominadas *EidClass*. Essas classes agregam a um objeto EID e seus atributos específicos. A figura 4.2 demostra a estrutura do *EidObject*.

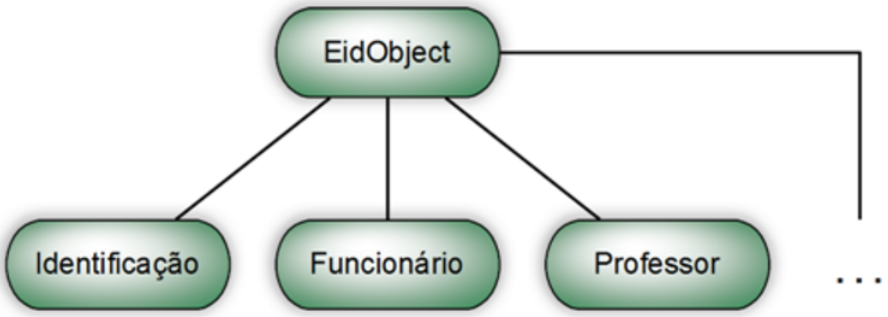


**Figura 4.2:** Estrutura EidObject

O EID já fornece algumas classes que podem ser usadas para alimentar diretórios LDAP sem nenhuma configuração adicional, isso porque existe uma

conversão pré-configurada para a ferramenta EID2LDAP que definem os atributos necessários para o esquema *brEduPerson*. Outras classes podem ser definidas pela própria organização, para suprir suas necessidades. Estas modificações deverão também ser realizadas na conversão utilizada pelo EID2LDAP.

### 4.2.1 Metadiretório

Metadiretório [SAN 07] é uma base de dados intermediária para construção do diretório e seu modelo independe do esquema final. Um metadiretório ideal permite ao administrador realizar alterações em um repositório e prover a atualização da informação em todos os diretórios ligados a ele.

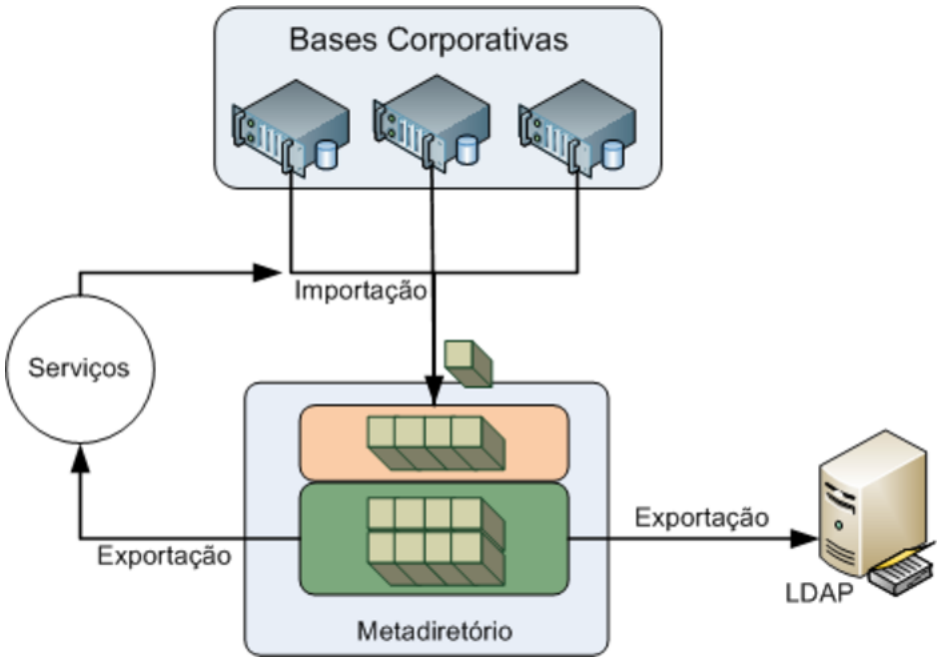


**Figura 4.3:** Fluxo do Metadiretório

O fluxo de informações acontece a partir dos dados das bases corporativas que serão importados para o metadiretório, e em seguida esses dados podem ser exportados para o LDAP, sendo utilizados, por exemplo, para uma autenticação. A figura 4.3 demonstra o fluxo dos dados.

## 4.3 EID2LDAP

O EID2LDAP [UFM 10b] é uma ferramenta que busca informações de diretórios armazenados no metadiretório EID e as transfere para servidores LDAP, além de possuir uma interface *web*.

A ferramenta EID possibilita extrair e incorporar ao seu metadiretório, dados de diversas bases institucionais e os disponibilizam em arquivos XML através de um serviço *web*. O EID2LDAP por sua vez acessa estas informações, utilizando a marcação XSLT (*Extensible Stylesheet Language Transformations*) para especificar a transformação dos dados para o formato LDIF (*Ldap Data Interchange Format*) compatível com o formato do servidor LDAP de destino, e realiza a atualização do servidor LDAP.

O XSLT introduz flexibilidade no EID2LDAP, permitindo ao usuário definir como se dará o mapeamento entre os dados do EID e o formato do LDAP.

Assim como o EID, o EID2LDAP permite o agendamento periódico das exportações. Para cada exportação são atualizados apenas os registros modificados, ou inseridos, ou desativados desde a última importação.

A exportação dos dados se inicia quando o algoritmo de transformação determina o momento em que o tempo do agendamento foi alcançado:

- EID2LDAP acessa o EID via *Web Service*;
- Busca registros modificados/inseridos/desativados;
- Todos os registros são transformados em LDIF e depois enviados;
- Cria um novo agendamento caso o modo de repetição esteja acionado;
- Registros são requisitados e processados de 100 em 100;
- Caso ocorra erro, o processamento será interrompido;
- O próximo agendamento reiniciará a partir da série de registros em que o erro ocorreu;

- O que foi enviado no intervalo antes do erro ao LDAP não será desfeito, mas reescrito na próxima iteração.

O LDAP possui uma sintaxe rígida com alguns atributos, como *mail*, *telephoneNumer*, etc, e podem ocorrer erros durante a exportação de dados malformados importados das fontes. O EID2LDAP possui um tratador onde a correção do dado da fonte pode ser realizada através de scripts de conversão executados ao inserir o valor do atributo ao seu destino.

Arquivo XML fornecido pelo EID contém informações sobre pessoas e grupos, mas apenas objetos novos (recém-inseridos no metadiretório), aqueles que sofreram algum tipo de alteração ou que foram removidos. Não há marcação no XML para indicar o atributo alterado, nem para diferenciar um objeto novo de um alterado. Sempre é enviado todo o conteúdo do objeto.

Alguns exemplos dos formatos para informações sobre pessoas:

Pessoas e seus atributos:

```
<eid-object type="person">....</eid-object>
```

Grupos:

```
<eid-object type="group">...</eidobject>
```

Membros do grupo:

```
<member> <eid-object>...</eid-object>...</member>
```

Pessoas e grupos desativados:

```
<eid-object type="person" removed="true">
```

No XML do EID, as várias classes existentes para a pessoa são recuperadas em elementos *attributes* e seus atributos dispostos em elementos `attribute`, contendo nome e valor de cada um como exemplificado abaixo:

```
1  <!--Pessoa ou Grupo -->
2  <eid-object type="person" guid="EHBBCXKA-YLHXBAAA"
3     serial="148048">
4    <!-- Classe -->
5    <attributes class="Identificacao" id="52347">
6      <attribute name="nomeCompleto" key="03812882698"
7        ><![CDATA[JOAO SILVA]]></attribute>
8      <attribute name="nomeSolteiro"
9        ><![CDATA[ ]]></attribute>
10     <attribute name="cpf"
11       ><![CDATA[01212222222]]></attribute>
12   </attributes>
13
14   <attributes class="Email" id="72201">
15     <attribute name="email"
16       ><![CDATA[silva@mail.com]]></attribute>
17   </attributes>
18 </eid-object>
19 <!--Membros de Grupos -->
20 <member>
21   <eid-object >...</eid-object>
22 </member> }
```

**Algoritmo 4.1:** Exemplo de um registro no EID em formato XML

No momento da exportação, caso o registro já exista no LDAP (identificado pelo DN gerado no LDIF), o LDIF é modificado para aplicar operações de alteração (*modify*) no registro do LDAP.

Apenas os *objectClasses* representados no LDIF serão substituídos no LDAP, isto é, os *objectClasses* no LDAP passarão a ter os atributos com os mesmos valores do EID, enquanto outros *objectClasses* permanecerão com seus atributos inalterados. Isto possibilita que outras aplicações alimentem diretamente o diretório sem a necessidade de passar pelo EID.

## 4.4 Shibboleth

O Shibboleth é um sistema baseado em padrões para *web single sign-on*, ou seja, o usuário necessita se autenticar apenas uma vez, para posteriormente, ter acesso automaticamente aos outros serviços com a mesma política de autenticação dentro das fronteiras organizacionais. É um projeto da Internet2 Middleware Initiative, código aberto e permite que serviços *web* informem decisões de autorização para acessos individuais de recursos protegidos *on-line* de forma que preserve a privacidade do usuário.

O software consiste na implementação de padrões amplamente utilizados para autenticação e autorização federada via *web*, principalmente o SAML (*Security Assertion Markup Language*) criado pela OASIS (*Organization for the Advancement of Structured Information Standards*).

Um usuário se autentica com sua credencial referente à sua organização de origem. A organização (provedor de identidade) passa a mínima quantidade de informações de identificação necessárias para que o gerente de serviço habilite a decisão de autorização.

Shibboleth é composto por dois módulos: Provedor de Identidade e Provedor de Serviço. O primeiro é responsável pela autenticação e o segundo pela autorização. Numa federação Shibboleth, normalmente apresenta mais dois componentes: o serviço WAYF (*Where Are You From*), usado para localizar o provedor de identidade do usuário e o serviço de *Metadata*, responsável em concentrar as informações dos provedores pertencentes à federação.

O Shibboleth foi desenvolvido para tratar os seguintes desafios:

- Acabar com as múltiplas senhas requeridas para múltiplas aplicações;
- Suportar a escalabilidade no gerenciamento de múltiplas aplicações;
- Melhorar a segurança associada ao acesso de serviços de terceiros;
- Aumentar a privacidade dos dados dos usuários;

- Suportar a interoperabilidade dentro e entre organizações;
- Permitir a liberdade de escolha das tecnologias de autenticação para as instituições;
- Permitir o controle de acesso efetuado a partir dos provedores de serviço.

Hoje existem diversas federações espalhadas pelo mundo que fazem parte do projeto Shibboleth.

### 4.4.1 Shibboleth IdP

O Shibboleth IdP (Identity Provider) é a aplicação voltada para ser utilizada como provedor de identidade. Através do IdP, será firmado declarações de autenticação ou de atributos para as partes confiantes, neste caso os provedores de serviços. Os subcomponentes do IdP são mostrados pela figura 4.4 .

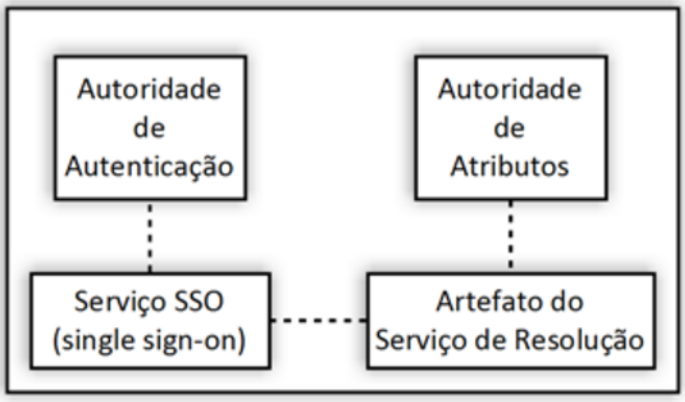


**Figura 4.4:** Subcomponentes do Shibboleth IdP

A autenticação e a entrega de atributos são realizadas da seguinte forma:

1. O usuário envia as suas credenciais, que são devidamente verificadas pelo provedor de identidade.
2. O provedor de identidade envia um handle para o provedor de serviço, atestando que o usuário foi autenticado.

3. O provedor de serviço envia este handle para o provedor de identidade, solicitando a entrega de atributos referentes ao usuário em questão.

4. O provedor de identidade envia esses atributos para o provedor de serviço.

A instalação padrão de um provedor de identidade da federação CAFe é composta por três elementos principais:

- Shibboleth Identity Provider – Serviço de *middleware*, responsável por intermediar a autenticação e o envio de atributos.

- Serviço de autenticação *Single Sign-on* – Serviço de autenticação *web single sign-on*, responsável pela interface de autenticação com o usuário. Na versão Shibboleth IdP 2.x, este serviço já vem incorporado ao serviço de provedor de identidade.

- OpenLDAP – Servidor de diretório, responsável por armazenar os atributos dos usuários e validar as suas credenciais.

Shibboleth IdP pode trabalhar com outros servidores de autenticação e atributos.

### 4.4.2 Shibboleth SP

O provedor de serviço é responsável por fazer a autorização do usuário e disponibilizar o acesso ao recurso, através da autenticação e dos atributos disponibilizados pelo provedor de identidade.

A autorização e o acesso ao recurso são realizados da seguinte forma:

1. O usuário solicita o acesso ao recurso.

2. O provedor de serviço solicita que ele se autentique no provedor de identidade da sua instituição.

3. O provedor de identidade envia um handle atestando a autenticação do usuário.

4. O provedor de serviço envia o handle para o provedor de identidade solicitando os seus atributos.

5. O provedor de serviço processa a autorização baseado nos atributos do usuário e disponibiliza o acesso ao recurso.

A instalação padrão de um provedor de serviço da federação CAFe é baseada no Shibboleth SP (Service Provider), que, por sua vez, é composto por dois elementos:

- Mod_shib: Módulo do Apache, responsável por controlar a autorização e o acesso ao recurso.
- Shibd: Daemon, responsável por intermediar a solicitação de autenticação e de atributos.

Shibboleth SP pode trabalhar com o servidor HTTP Microsoft IIS. Hoje, já existe uma vasta lista de aplicações que são compatíveis com o Shibboleth, por exemplo:

- Confluence Wiki
- eAcademy
- Horde
- Media Wiki
- Microsoft DreamSpark [2]
- WordPress
- Blackboard
- Moodle
- Google Apps

[2]Este serviço já faz parte da Federação CAFe e pode ser acessado em: https://www.dreamspark.com/

### 4.4.3 Funcionamento

Nesta seção, será explicado o funcionamento de uma autenticação e autorização de um usuário cadastrado na Universidade Federal de Santa Catarina – UFSC, acessando um recurso de um provedor de serviço qualquer.

A figura 4.5 representa o processo completo e representa um usuário acessando um serviço através do *browser*, e toda linha de comunicação entre o *browser* do usuário, provedor de serviço, WAYF e provedor de identidade.

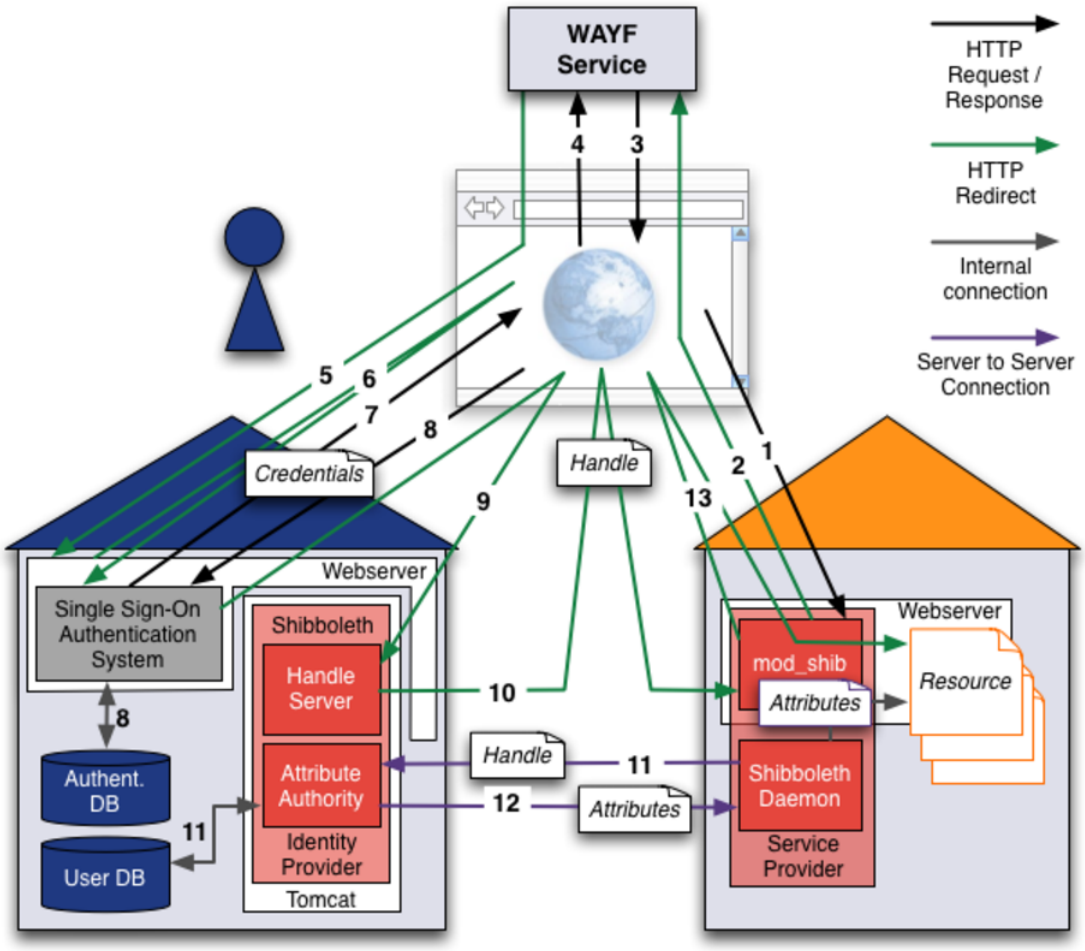


**Figura 4.5:** Processo completo de autenticação e autorização dentro da Federação CAFe

O processo é dividido em fases. A primeira fase é realizada a solicitação de acesso ao recurso e redirecionamento do usuário. As seguintes etapas são:

1. O usuário inicializa o *browser* e acessa o endereço `http://atlases.muni.cz/en/index.html` referente ao site da Atlases que fornece imagens de patologias. Lembrando que esse processo só funcionará para os provedores de serviço que estejam cadastrados na Federação CAFe.

2. Como o usuário ainda não está autenticado, o servidor *web* responde com um redirecionamento HTTP para o servidor WAYF (https://ds.cafe.rnp.br/WAYF). WAYF necessita saber qual provedor de serviço o usuário está tentando acessar, então as informações são enviadas como parâmetro GET.

3. O WAYF responde ao *browser* mostrando uma página em que o usuário deve selecionar a instituição de origem.

A seleção da instituição de origem é determinada pela segunda fase do processo. Sua escolha é armazenada por *cookies* de sessão no navegador do usuário.

Após esta etapa, inicia a fase de autenticação na instituição de origem. Nesta terceira fase ocorre:

4. O usuário envia para WAYF a seleção a partir de uma requisição HTTPS.

5. Após o envio da requisição do usuário, o WAYF responde com um redirecionamento HTTPS para o provedor de identidade do usuário. Os *cookies* são habilitados para lembrar a escolha do usuário para o checkbox “Lembrar a seleção nesta sessão do navegador”, ou seja, o *cookie* estará disponível somente durante a sessão atual do navegador. Então o *browser* do usuário, envia uma requisição HTTPS para o Shibboleth Handle Server da sua instituição de origem.

6. Para ocorrer o processo de autenticação, o servidor *web*, protegendo o acesso ao Handle Service, redireciona o navegador para o sistema de autenticação *single sign-on*.

7. O sistema de autenticação envia a página de *login* para o navegador e habilita os *cookies* do usuário.

8. Na página de *login*, o usuário deve fornecer suas credenciais de *login* e senha. Assim, o navegador envia uma nova solicitação para o sistema de autenticação *single sign-on*. Esse sistema verifica as credenciais do usuário através de uma pesquisa ao diretório LDAP.

9. Após o sucesso da autenticação, o navegador recebe um pedido de redirecionamento e *cookies* são enviados ao Handle Server do Shibboleth IdP.

10. O Shibboleth idP identifica que o usuário foi devidamente autenticado se baseando nos *cookies*, então o Handle Server cria um *handle* para o usuário. Esse *handle* é embarcado em um *hidden form*[3] que é enviado pelo navegador para o provedor de serviço.

Apesar do usuário estar autenticado, ele não está autorizado ainda. Na próxima fase, o provedor de serviço solicitará alguns atributos do usuário e decidirá sobre a sua autorização. O mod_shib do provedor de serviço examina as regras de acesso do Shibboleth. O seguinte fragmento do arquivo de configuração do Apache[4] habilita o acesso a qualquer usuário da federação com uma sessão válida:

```
<Directory /var/www/secure>
AuthType shibboleth
ShibRequireSession On
require valid-user
</Directory>
```

11. O Shibboleth SP, solicita ao provedor de identidade todos os atributos disponíveis para o usuário associado ao *handle* recebido no passo anterior.

12. Após a sessão HTTPS ser estabelecida entre o Shibboleth Daemon e o Attribute Authority do Shibboleth IdP, o Attribute Authority verifica a identidade do SP

[3]Forma no qual o processo ocorre sem ser percebido pelo usuário.

[4]Servidor *web* livre cuja funcionalidades são: transmissão de dados via *web*, processamento de dados e execução de aplicativos distribuídos.

com base no certificado enviado pelo Shibboleth Daemon. Uma vez que o Attribute Authority recebe a solicitação de atributos, ele verifica se o *handle* é o mesmo gerado pelo Handle Server no passo 10. Caso isso seja verdade, o Attribute Authority sabe qual usuário o *handle* se refere e então verifica o *Attribute Release Policy* (ARP) (arquivo XML responsável pelas regras que determinam quando um atributo de um determinado usuário pode ser enviado para um determinado provedor de serviço). Após esta verificação, o Attribute Authority envia para o provedor de serviço todos os atributos permitidos de acordo com o ARP.

13. Finalmente, o usuário recebe um *cookie* de sessão Shibboleth e é redirecionado para o recurso. Os atributos enviados pelo provedor de identidade são disponibilizados para aplicação *web* pelo mod_shib, na forma de variáveis de ambiente do servidor *web*. Desta forma, o recurso pode usar esses atributos para prover um nível de autorização mais granular, além de possibilitar funcionalidades extras na aplicação, baseado nestes atributos.

# Capítulo 5

# Implantação do Provedor de Identidade

Nesta seção será descrito todo o processo de instalação da infraestrutura do provedor de identidade aplicada na UFSC. Estes processos foram realizados com base nos tutoriais realizados pela RNP localizados na página wiki referente ao projeto CAFe [RNP 10n].

## 5.1 Instalação do Servidor

A instalação do servidor foi realizada através da criação de uma máquina virtual e com o sistema operacional Ubuntu Server 8.04 [1] . Este sistema foi escolhido pela recomendação do projeto e-AA [RNP 10j], pois disponibiliza pacotes Java versão 6 nativamente e terá suporte para atualizações de segurança até abril de 2013.

Além da instalação básica e padrão do sistema, foi realizado:

- Instalação do pacote OpenSSH [2] : conjunto de softwares que provém à criptografia em sessões de comunicação em uma rede de computadores usando o protocolo SSH/TLS.

[1] Pode ser encontrado em: http://pacotes.ufrgs.br/ubuntu/ubuntu-8.04.4-server-i386.iso

[2] *SSH Communications Security* (http://www.openssh.com/)

- Configuração do `hostname` e do arquivo `/etc/hosts`. Para determinar o nome local do servidor e o IP do qual o servidor se reconhecerá. É importante que a máquina esteja utilizando um IP fixo para evitar eventuais problemas decorrentes da indisponibilidade do servidor DHCP [3] .

- Configuração do *firewall* adicionando linhas de comando para liberar apenas as portas dos principais serviços que serão usados, como: SSH, NTP (Serviço de sincronização do relógio), Apache, Tomcat.

- Configuração do NTP para acessar um servidor de relógio e sempre mantê-lo sincronizado.

## 5.2 Instalação do Diretório com Esquema BrEduPerson

Após a configuração completa do sistema, foi instalado o serviço de diretório (LDAP) e também configurado para usar o esquema brEduPerson pelos seguintes passos:

- Instalação do pacote slapd.

- Abertura das portas no *firewall* que serão usadas pelo serviço slapd.

- Edição do arquivo de configuração do servidor ldap (`slapd.conf`): É importante citar que nesse arquivo foram incluídos os seguintes esquemas LDAP:

  - Core
  - Cosine
  - Nis

[3] *Dynamic Host Configuration Protocol* é um protocolo de serviço TCP/IP que oferece configuração dinâmica de terminais, com concessão de endereços IP de host e outros parâmetros de configuração para clientes de rede.

  - InetOrgPerson
  - EduPerson
  - BrEduPerson
  - Schac

- Geração dos certificados SSL para o LDAP com as seguintes características: tamanho da chave de 2048 bits e validade de 730 dias.

- Inclusão de parâmetros no arquivo `slapd.conf` para suportar conexão segura usando o certificado criado.

- Edição do arquivo de configuração do cliente ldap (`ldap.conf`).

- Criação de listas de permissões de leitura e escrita no arquivo `sldap.conf` para o administrador da base, acesso de leitura para o serviço shibboleth, e o bloqueio de todas as outras permissões.

- Inclusão da estrutura inicial da base LDAP, do usuário administrador e usuário de leitura usado pelo Shibboleth, como por exemplo:

```
dn:  dc=ufsc,dc=br
objectClass:  top
objectClass:  dcObject
objectClass:  organization
o:  ufsc.br
dc:  ufsc

dn:  ou=people,dc=ufsc,dc=br
objectClass:  organizationalUnit
objectClass:  top
ou:  people
```

```
dn:  cn=admin,dc=ufsc,dc=br
objectClass:  simpleSecurityObject
objectClass:  organizationalRole
cn:  admin
description:  Administrador da base LDAP
userPassword:  secret

dn:  cn=leitor-shib,dc=ufsc,dc=br
objectClass:  simpleSecurityObject
objectClass:  organizationalRole
cn:  leitor-shib
description:  Leitor da base para o shibboleth
userPassword:  secret
```

## 5.3 Extração dos dados para o Metadiretório

A extração dos dados para o metadiretório é realizado pelo EID e por isso há a necessidade de sua instalação. Para isso, foi necessário:

- Instalar o Java, o Tomcat [4] e o MySQL [5].
- Criar as bases de dados eid e pcollect para a utilização do EID.
- Descompactar e compilar o arquivo do EID [UFM 10a].
- Criar o arquivo `/etc/tomcat6/Catalina/localhost/eid.xml` [RNP 10h] e editar de acordo com as informações locais: esse arquivo é responsável pela configuração e deploy do EID.

Esta extração necessita que o administrador tenha alguns conhecimentos, como:

[4] http://tomcat.apache.org/
[5] http://www.mysql.com/

- A organização das bases de dados da instituição que serão utilizadas, como bases de alunos, professores, recursos humanos, etc.

- Verificar as classes e os atributos necessários para o esquema brEduPerson.

O EID necessita se conectar ao servidor de onde serão retirados estes dados. Com isso precisa-se cadastrar cada repositório, informando o tipo de banco de dados (relacional ou via arquivo CSV [6]), a URL, o *driver* utilizado, o usuário com permissão ao banco e sua senha.

Após os cadastramentos dos bancos de dados, deve-se criar as extrações para alimentar as classes definidas. É importante saber que a primeira classe a ser carregada deve ser a classe Identificação, pois essa classe será o container principal de uma pessoa (na base LDAP) e a partir dele que será gerado o GUID.

Na configuração da extração da classe principal, deverá determinar um atributo que seja único para cada indivíduo e utilizá-lo para ser o ponto de referência da geração do GUID.

A partir do identificador único de cada pessoa serão criados os relacionamentos às demais classes que informarão a conta (login e senha) e a(s) sua(s) função(ões) dentro da universidade (aluno, professor, funcionário).

Extraído a classe principal Identificação, poderá extrair as demais. A conciliação delas com a principal é baseada no atributo GUID. No momento da extração de cada um das demais classes, há a necessidade de realizar o *select* do mesmo atributo único da classe Identificação e conciliar com o atributo GUID.

O banco de dados do EID irá verificar se existe aquele atributo único na forma de GUID, e caso exista, realizará a conciliação do novo registro com o registro principal.

Caso o atributo único não exista em forma de GUID, a extração acusará um erro dizendo que não foi encontrado o valor em questão no banco e apenas funcionará quando o registro principal estiver incluído.

[6] Um arquivo CSV (*Comma Separted Value*) é um formato de arquivo que normalmente é usado para troca de dados entre aplicações diferentes.

A segunda classe a ser extraída deve ser a classe Conta. Apenas depois que uma pessoa, ou seja, um registro tiver sido conciliado, entre pelo menos a classe Identificação e Conta, o EID2LDAP estará apto para importar os dado do metadiretório EID para a base LDAP.

A configuração das extrações das classes é realizada com a criação de uma ETC. Numa ETC, será determinado uma consulta SQL no repositório cadastrado para referenciar os atributos da classe com os resultados obtidos.

O EID possibilita que os dados adquiridos da seleção sejam tratados por meio scripts em Javascript para serem convertidos de acordo com a necessidade do administrador.

Alguns desses scripts são por exemplo na transformação dos hash SHA-1 e MD5 das senhas na sequência hexadecimal para base64 B.1 . Outro script de transformação é realizado no atributo e-mail que pode estar com sintaxe errada (com caracteres acentuados) e deve-se removê-los de forma automática.

Para cada ETC criada, tem-se que criar processos de extração. Esses processos servem para que o agendamento saiba o que será executado. A criação do agendamento permite escolher qual ETC deseja executar e também o modo de repetição (nenhum, diário, mensal, anual, etc.), a data e horário da execução.

## 5.4 Alimentação do Diretório a partir do Metadiretório

Esta seção apresenta os passos necessários para realizar a instalação da ferramenta EID2LDAP no qual exporta os registros do metadiretório EID para a base LDAP. Para isso, teve-se que:

- Criar uma base de dados que será utilizada pelo EID (seu nome deve ser eid2ldap).
- Fazer a carga da base com o arquivo eid2ldap-1.1.1.sql [7] .
- Adquirir a versão mais recente do EID2LDAP [UFM 10b].

[7] Pode ser localizado em http://pacotes.ufrgs.br/ubuntu/hardy/instalacao-eid2ldap/eid2ldap-1.1.1.sql.

- Descompactar o arquivo adquirido.
- Criar o arquivo `/etc/tomcat6/Catalina/localhost/eid2ldap.xml` [RNP 10g] que será responsável pela configuração e deploy do EID.

Após a instalação e configuração de todos os requisitos anteriores, pode-se acessar o serviço EID2LDAP através do endereço: http://hostname:8080/eid2ldap/Home.faces [8] . Realizando o login do serviço, deve-ser determinar as seguintes ações:

- Ajustar o endereço do *web service* do EID em http://hostname:8080/eid/services/EidService?wsdl.
- Iniciar o agente exportador através do menu Agendamento / Agente Gerenciador De Agendamento.
- Configurar qual servidor LDAP será usado para importar os dados do metadiretório para base LDAP, no menu Configuração / Servidor Ldap: informar o usuário (na forma de DN) e sua senha de acesso.
- Agendar para a realização da atualização do diretório.

A conclusão dos processos de importação pelo agendamento pode ser visualizado pelos logs gerados em cada agendamento concluído. Tanto o EID quanto EID2LDAP permitem esse monitoramento. Caso seja necessário fazer a verificação do dados na base LDAP, pode-se realizar uma pesquisa por meio de algum programa de gerenciamento LDAP ou através do comando *ldapsearch*.

## 5.5 Instalação do Shibboleth IdP

O Shibboleth IdP representa o principal item de um provedor de identidade, pois é ele quem habilitará a conversa com os provedores de Serviço e WAYF. A versão instalada foi o Shibboleth-IDP 2.x e em uma nova máquina virtual. O processo

[8] Hostname é o nome completo do servidor, mas podendo ser também o ip.

completo de sua instalação pode ser visualizado na página wiki referente ao projeto CAFe e na parte de instalação Shibboleth-IDP 2.x [RNP 10i]. Os principais processos de instalação são:

- Instalação do Java, Tomcat e Apache.
- Editar o arquivo de configuração de segurança do java (`/etc/java-6-sun/security/java.security`) e adicionar quatro bibliotecas.
- Editar o arquivo `/etc/tomcat6/server.xml` para definir que o Shibboleth-IDP irá receber conexões pela porta 8443, ou seja, conexões cifradas.
- Criar o arquivo `/etc/tomcat6/Catalina/localhost/idp.xml` contendo o caminho para `auto-deploy` do Shibboleth-IDP.
- Configurar o Apache para habilitar o portal de autenticação da instituição.
- Criar o arquivo `/etc/apache2/conf.d/idp.conf` para configurar a ligação entre o Apache e Tomcat.
- Obter o Shibboleth-IDP [9] .
- Obter o arquivo `bcprov-jdk16-144.jar` [10] que é uma biblioteca que implementa protocolos criptográficos em Java.
- Instalar o Shibboleth-IDP.
- Editar o arquivo `/opt/shibboleth-idp/conf/handler.xml`.
- Editar o arquivo `/opt/shibboleth-idp/conf/relying-party.xml`.
- Configurar a resolução/liberação [11] de atributos recomendados pelo projeto CAFe [RNP 10m].

[9] Disponível em http://shibboleth.internet2.edu/downloads/shibboleth/idp/.

[10] Disponível em http://polydistortion.net/bc/index.html.

[11] Conjunto de configurações que possibilita que atributos sejam buscados de diversas formas e em diversas bases de dados, e que sejam definidos quais provedores de serviço são autorizados a buscá-los.

- Configurar o arquivo `/opt/shibboleth-idp/conf/login.config` para permitir que o usuário para o uso do shibboleth criado na base LDAP possa acessar os dados da base.

- Gerar um par de chaves criptográficas com tamanho de 2048 bits e validade de 1095 dias para ser usado na conexão segura do Shibboleth-IdP.

- Gerar um par de chaves criptográficas com tamanho de 2048 bits e validade de 1095 dias para ser usado na conexão segura do Apache.

- Editar o arquivo `/opt/shibboleth-idp/metadata/idp-metadata.xml` para incluir o certificado criado para Shibboleth-IDP e permitir que o *metadata* se identifique com ele.

O serviço de autenticação *web* pré-configurada através de uma página de *login*, é instalada automaticamente por esse processo de instalação. Apesar disso, há a possibilidade de personalizar a página de *login* da instituição realizando a mudança do *template* nos arquivos localizados no diretório `/src/main/webapp/` da instalação do Shibboleth-IDP.

A página de *login* da UFSC foi modificada de forma simples incluindo a identidade visual da universidade e assim visualizando o funcionamento da mudança.

## 5.6 Entrada na Federação CAFe

Após a instalação de todos os pré-requisitos e os softwares necessários, o Shibboleth-IdP ainda não estará pronto para ser reconhecido pela federação, pois é necessário que o provedor de identidade envie o seu metadado para ser cadastrado no WAYF e assim permitir o seu reconhecimento.

A RNP possui uma federação de teste denominado de Federação Chimarrão. Antes de um provedor (identidade ou serviço) entrar na Federação CAFe, é necessário entrar na Federação Chimarrão e seu serviço WAYF necessita cadastrar o metadado do provedor em questão.

Para validar a conclusão da etapa de entrada na Federação Chimarrão, é necessário realizar um teste padrão. Este teste consiste em acessar o endereço `https://chimarrao.ufrgs.br/homologa` e selecionar a IdP da instituição de teste. Será realizado um redirecionamento para a página de autenticação da instituição e após preencher o login e a senha com algum usuário cadastrado na base LDAP, deve-se observar se houve sucesso no retorno dos atributos do usuário autenticado. Se os atributos foram visualizados sem nenhum problema, então o teste foi bem sucedido.

Agora, para entrar na Federação CAFe, há a necessidade de configurar o shibboleth para o reconhecimento da federação [RNP 10a] . Os principais passos são:

- Atualizar o certificado para a conexão HTTPS de `ds.cafe.rnp.br` no keytool do sistema para que os metadados sejam atualizados.
- Modificar o arquivo `/opt/shibboleth-idp/conf/relying-party.xml` para incluir o reconhecimento do metadado referente à CAFe.
- Certificar que o arquivo `/opt/shibboleth-idp/conf/attribute-filter.xml` está configurado para liberar os mesmos atributos para as federações Chimarrão e CAFe.

Com isso, poderá realizar a solicitação da entrada do provedor de identidade na Federação CAFe. A instituição estará automaticamente cadastrada tanto na Federação Chimarrão quanto na CAFe e assim poderá usufruir dos serviços oferecidos por todos os provedores de serviço cadastrados em ambas as federações.

# Capítulo 6

# Dificuldades

Diante deste trabalho, foram encontrados várias dificuldades que exigiram espaços de tempo diferentes para sua resolução. Apesar de existir um processo descritivo sobre a implementação da infraestrutura de um provedor de identidade no site da RNP, isso não tornou o trabalho mais fácil.

As dificuldades se devem a falta de documentação, ou documentação insuficiente, exigindo assim um processo de dedução e com ajuda de terceiros. Também se enquadram nesta categoria a falta de conhecimento nas estruturas de bases de dados da UFSC e na sua imensidão, necessitando tempo e discussões com as partes responsáveis dentro da instituição.

## 6.1 Dificuldades de Instalação

As ferramentas EID e EID2LDAP não são conhecidas e usada por todos, pois suas aplicações são somente para estes fins, e isso trouxe alguns problemas tanto de conflitos entre os serviços requisitados (Java e Tomcat), quanto no próprio funcionamento interno deles.

A demanda de maior tempo de dedicação e a necessidade de discutir com os criadores do software para obter um conhecimento mais profundo conseguiram solucionar esses problemas.

## 6.2 Dificuldades Estruturais

O processo de instalação e configuração da infraestrutura exigiu a necessidade do entendimento e levantamento da estrutura da base de dados da instituição (UFSC), e determinar qual melhor estrutura para o metadiretório EID e posteriormente da configuração do EID2LDAP.

Uma base acadêmica possui alguns dados restritos, por exemplo senhas de usuários. Para não violar a integridade e nem a segurança do mesmo, não houve a utilização dos valores deste dado para compor a base LDAP.

Durante este trabalho, foram realizadas duas estruturas para que a UFSC fosse capaz de prover identidade. A primeira foi realizada de forma compacta onde visava aprender o processo de instalação e suas ferramentas, o ingresso da instituição na Federação CAFe seguindo todas as regras com nível mínimo de exigência. Isso demandou que a base LDAP tivesse apenas uma pequena porcentagem de pessoas interligadas à UFSC.

A segunda estrutura foi criada depois de um processo de análise sobre a melhor forma do metadiretório e os atributos utilizados, e com a instalação da versão atual do EID.

Uma base de dados acadêmica possui muitos registros decorrente das informações de milhares de pessoas que possuem ou possuíram algum vínculo com a instituição. Devido a essa imensa quantidade de dados, o EID e o EID2LDAP necessitaram desde vários minutos a até algumas horas para poder realizar o processamento da atualização dos dados no metadiretório, a conciliação de todos os registros percorridos e a importação dos novos dados na base.

Por isso, há um cuidado no agendamento dos processos em que eles devem ser calculados para existir um tempo de processamento e conciliação grande o suficiente entre o agendamento de um processo e outro. Caso este tempo seja curto, poderá ocorrer um acúmulo de recursos de processamento em que o servidor poderá travar seus processos. Assim os serviços realizados pelo EID e EID2LDAP não funcionará corretamente.

# Capítulo 7

# Considerações Finais

A infraestrutura de autenticação e autorização criada pela RNP e o conceito de federação acadêmica proporciona que todas as instituições de ensino e pesquisa se interliguem por meio de uma rede de confiança e possibilitem a troca de serviços entre elas.

Além desta conexão mutua entre as partes, criou-se uma estrutura com o uso de uma base única de dados (pelo menos a UFSC que não possuía), ou seja, uma base centralizada em que os dados são obtidos a partir de diversas outras fontes. Deste modo, a parte de gerenciamento do dados (adição, remoção, atualização) é feita de forma automática por estar conectada com as outras.

Das duas infraestruturas criadas, como citada anteriomente, a primeira teve um metadiretório com cerca de 7210 registros dentre eles, professores, técnicos.

Já na segunda infraestrutura foi criado uma base mais detalhada com cerca de 42927 registros inclusos no metadiretório. A base possui pessoas que estão vinculadas com a UFSC pelas as seguintes funções: aluno de graduação, aluno de pós-graduação, professor ou técnico-administrativo.

Foram realizadas as extrações dos dados referentes às classes padrões do EID (Identificação, Conta, Aluno, Professor, Técnico, Email), e a criação de uma nova classe denominada de Comunidade para incluir pessoas que possuem algum vínculo com a universidade por meio de outras atividades, como aluno especial, ouvinte ou

possuindo alguma matrícula isolada.

Os serviços podem ser providos por qualquer entidade, basta que siga as regras da federação. Uma instituição pode interligar todos os seus principais serviços, como e-mail, internet, *wireless*, acesso a livros na biblioteca, periódicos, VoIP e até acesso ao restaurante universitário, para autorizar com apenas umas autenticação.

O conceito de federação proporciona uma maior segurança tanto na comunicação entre as partes quanto no acesso aos dados de uma autenticação. Toda comunicação é feita de forma cifrada em que todos os provedores (IdP e SP) necessitam possuir certificados e configurar o Shibboleth para utilizá-los na comunicação.

A segurança na autenticação vem do princípio em que os provedores de identidade apenas irão repassar atributos necessários do usuário para os provedores de serviço, indicando uma maior privacidade dos dados do usuário. Os serviços ainda poderão determinar diferentes níveis de acessos para a autorização, baseando-se apenas nos valores dos atributos.

Os provedores de serviço e identidade possuem responsabilidades distintas e bem definidas, o que torna o processo de comunicação mais segura. Quanto maior for o número de instituições cadastradas na Federação CAFe, melhor será, pois o acesso aos serviços não terá fronteiras.

## 7.1 Trabalhos Futuros

Diante de uma estrutura já criada do provedor de identidade, ainda há a necessidade de realizar melhorias no gerenciamento de senhas. Para não quebrar a segurança das senhas já cadastradas no banco de dados, temos a necessidade de criar novas senhas para este novo fim. Então poderá criar um serviço de gerenciamento de senhas de cada usuário vinculado à UFSC, em que o usuário irá primeiramente cadastrar sua senha, e posteriormente poderá realizar o pedido de modificação da mesma.

Além da UFSC ser um provedor de identidade, há também a necessidade de torná-la um provedor de serviço para poder integrar todos os serviços internos.

Alguns dos principais serviços que inicialmente poderão ter suporte para este tipo de autenticação seriam os serviços de e-mails e moodle. Estes dois serviços em especial, já possuem compatibilidade e integração com os serviços do Shibboleth, facilitando o processo.

O aumento da segurança é sempre um fator importante no que se diz a respeito à segurança da informação. O uso de certificação digital [HOU 01] no processo de autenticação dificulta ainda mais na quebra de segurança da senha e também possibilitando o uso de um *token*[1] ou *smartcard*.

[1]Tanto um token quanto um smartcart são *hardwares* criptográficos, utilizados principalmente para guardar e proteger a(s) chave(s) criptográfica(s) do usuário e utilizado no processo de autenticação. Possui vários mecanismos de proteção.

# Referências


[EHL 04] EHLENBERGER, A. et al. **Understanding LDAP - Design and Implementation**. IBM, June, 2004.

[FEI 10] FEIDE. **Feide**. Disponível em <http://www.feide.no/>. Acesso em: 7 de julho de 2010.

[GRO 97a] GROUP, N. W. **Lightweight Directory Access Protocol (v3)**. Disponível em <http://www.ietf.org/rfc/rfc2251.txt>. Acesso em: 7 de julho de 2010.

[GRO 97b] GROUP, N. W. **Lightweight Directory Access Protocol (v3): Attribute Syntax Definitions**. Disponível em <http://www.ietf.org/rfc/rfc2252.txt>. Acesso em: 7 de julho de 2010.

[GRO 97c] GROUP, N. W. **Lightweight Directory Access Protocol (v3): UTF-8 String Representation of Distinguished Names**. Disponível em <http://www.ietf.org/rfc/rfc2253.txt>. Acesso em: 7 de julho de 2010.

[GRO 00a] GROUP, N. W. **Definition of the inetOrgPerson LDAP Object Class**. Disponível em <http://www.ietf.org/rfc/rfc2798.txt>. Acesso em: 7 de julho de 2010.

[GRO 00b] GROUP, N. W. **Request for Comments: 2849**. Disponível em <http://ftp.rfc-editor.org/in-notes/rfc2849.txt>. Acesso em: 7 de julho de 2010.

[HOU 01] HOUSLEY, R.; POLK, T. **Planning for PKI: Best Practices Guide for Deploying Public Key Infrastructure**. New York, NY, USA: John Wiley & Sons, Inc., 2001.

[IAN 10] IANA. **Internet Assigned Numbers Authority**. Disponível em <http://www.iana.org/>. Acesso em: 7 de julho de 2010.

[IET 10] IETF. **The Internet Engineering Task Force**. Disponível em <http://www.ietf.org/>. Acesso em: 7 de julho de 2010.

[IMAC 08] INTERNET2 MIDDLEWARE ARCHITECTURE COMMITTEE, F. E. **Internet2-mace-dir-eduperson-200806**. Disponível em <http://middleware.internet2.edu/eduperson/>. Acesso em: 7 de julho de 2010.

[INC 10] INCOMMON. **InCommon Federation**. Disponível em <http://www.incommonfederation.org/>. Acesso em: 7 de julho de 2010.

[INT 10] INTERNET2. **Shibboleth - federated single sign-on software**. Disponível em <http://shibboleth.internet2.edu>. Acesso em: 7 de julho de 2010.

[JAN 10] JANET(UK). **UK federation**. Disponível em <http://www.ukfederation.org.uk/>. Acesso em: 7 de julho de 2010.

[MAL 04] MALER, E.; HUGHES, J. **Security Assertion Markup Language(SAML) 2.0 Technical Overview**. Disponível em <http://www.oasis-open.org/committees/documents.php?wg_abbrev=security>. Acesso em: 7 de julho de 2010.

[NDP 10] NDP. **Núcleo de Processamentos de Dados**. Disponível em <http://setic.ufsc.br/>. Acesso em: 7 de julho de 2010.

[RNP 10a] RNP. **Alterar configuração do IdP para migrar para Federação CAFe**. Disponível em <http://wiki.rnp.br/pages/viewpage.action?pageId=41616243>. Acesso em: 19 de outubro de 2010.

[RNP 10b] RNP. **CAFe – Acordo de Participação na Federação (Provedores de Identidade)**. Disponível em <http://www.cafe.rnp.br/cafewiki/images/b/bc/CAFe_AcordoIdP_v1.0.pdf>. Acesso em: 7 de julho de 2010.

[RNP 10c] RNP. **CAFe – Acordo de Participação na Federação (Provedores de Serviço)**. Disponível em <http://www.cafe.rnp.br/cafewiki/images/0/0b/CAFe_AcordoSP_v1.0.pdf>. Acesso em: 12 de outubro de 2010.

[RNP 10d] RNP. **CAFe – Governança da Federação**. Disponível em <http://www.cafe.rnp.br/cafewiki/images/b/b6/CAFe_Governanca_v1.0.pdf>. Acesso em: 12 de outubro de 2010.

[RNP 10e] RNP. **CAFe – Requisitos Mínimos de Gestão de Identidade**. Disponível em <http://www.cafe.rnp.br/cafewiki/images/b/b3/CAFe_ReqGestaoIdentidade_v1.0.pdf>. Acesso em: 12 de outubro de 2010.

[RNP 10f] RNP. **Federação CAFe**. Disponível em <http://www.cafe.rnp.br/wiki/Federação_CAFe>. Acesso em: 7 de julho de 2010.

[RNP 10g] RNP. **Instalação do EID2LDAP**. Disponível em <http://wiki.rnp.br/pages/viewpage.action?pageId=42143530>. Acesso em: 19 de outubro de 2010.

[RNP 10h] RNP. **Instalação EID**. Disponível em <http://wiki.rnp.br/pages/viewpage.action?pageId=41190354>. Acesso em: 19 de outubro de 2010.

[RNP 10i] RNP. **Instalação Shibboleth-IDP 2.x**. Disponível em <http://wiki.rnp.br/pages/viewpage.action?pageId=41616303>. Acesso em: 19 de outubro de 2010.

[RNP 10j] RNP. **Projeto e-AA**. Disponível em <http://www.cafe.rnp.br/wiki/Projeto_e-AA>. Acesso em: 7 de julho de 2010.

[RNP 10k] RNP. **Rede Nacional de Ensino e Pesquisa**. Disponível em <http://www.rnp.br>. Acesso em: 7 de julho de 2010.

[RNP 10l] RNP. **Requisitos Técnicos para Membros da Federação CAFe**. Disponível em <http://www.cafe.rnp.br/cafewiki/images/f/f6/CAFe_ReqTecnicos_v1.0.pdf>. Acesso em: 12 de outubro de 2010.

[RNP 10m] RNP. **Resolução e liberação de atributos**. Disponível em <http://wiki.rnp.br/pages/viewpage.action?pageId=41616318>. Acesso em: 19 de outubro de 2010.

[RNP 10n] RNP. **Roteiro de Atividades para Entrada de um IDP na CAFe**. Disponível em <http://wiki.rnp.br/pages/viewpage.action?pageId=41616281>. Acesso em: 19 de outubro de 2010.

[SAN 07] SANTOS, A. **Gerenciamento de identidades**. Rio de Janeiro: Brasport, 2007.

[SWI 10] SWITCH. **Switch**. Disponível em <http://www.switch.ch/>. Acesso em: 7 de julho de 2010.

[TER 09] TERENA. **SCHAC (SCHema for ACademia) – attribute definitions for individual data**. Disponível em <http://www.terena.org/activities/tf-emc2/schacreleases.html,>. Acesso em: 7 de julho de 2010.

[TRI 07] TRIGO, C. H. **OpenLDAP: uma abordagem integrada**. São Paulo: Novatec, 2007.

[UFM 10a] UFMG, L. **EID**. Disponível em <http://sourceforge.net/projects/eid/>. Acesso em: 19 de outubro de 2010.

[UFM 10b] UFMG, L. **EID2LDAP**. Disponível em <http://sourceforge.net/projects/eid2ldap/>. Acesso em: 19 de outubro de 2010.

# Apêndice A

# Diagramas

## A.1 Modelo de nomes LDAP

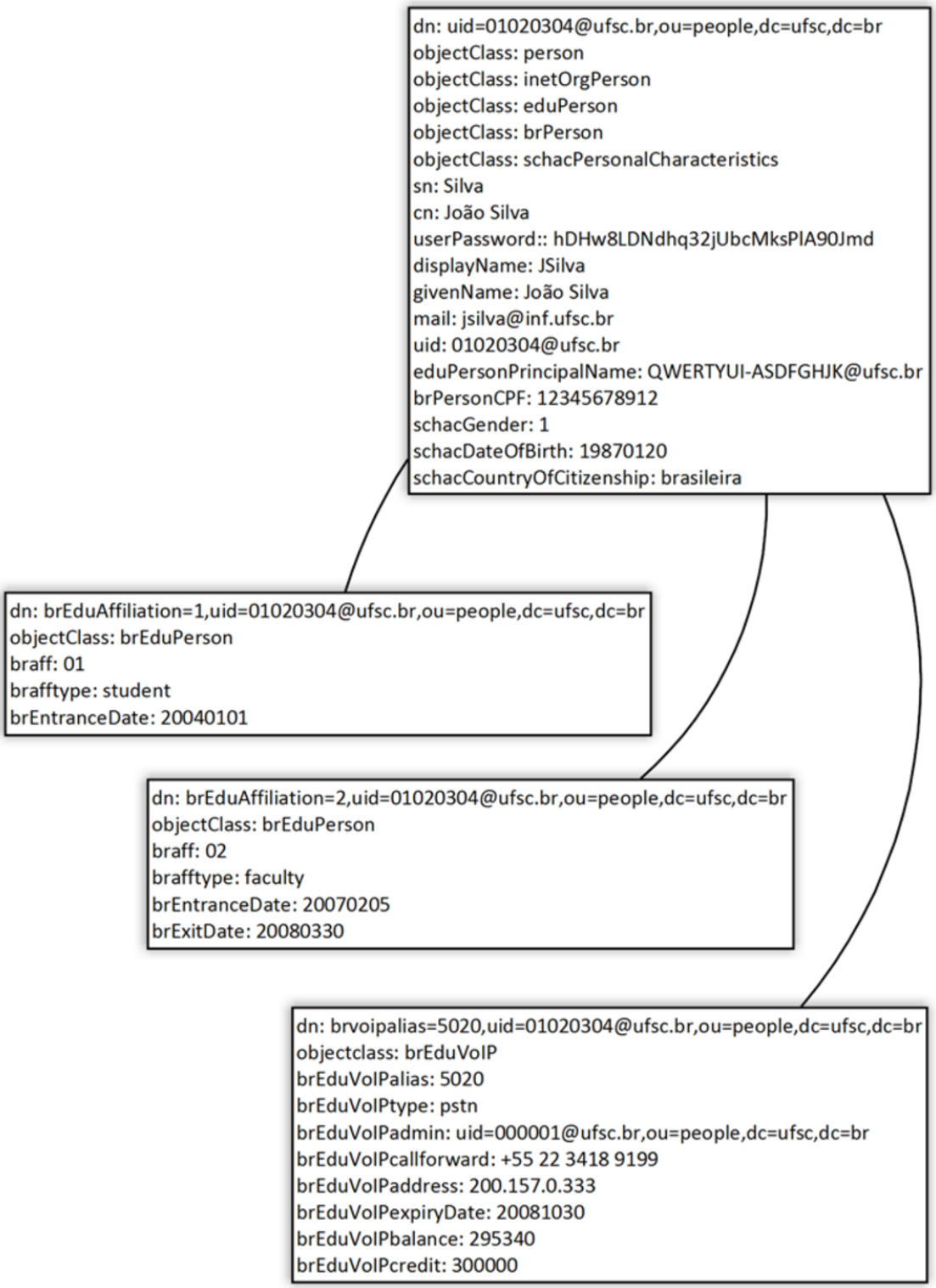


**Figura A.1:** *Container* da estrutura hierárquica

# Apêndice B

# Javascripts

## B.1 Javascript para manipulação de senhas

```
1  result = null;
2  public void execute(){
3          String pass = String.valueOf(senha);
4          java.security.MessageDigest md = java.security.MessageDigest
               .getInstance( "SHA1" );
5          md.update( pass.getBytes() );
6          java.math.BigInteger hash = new java.math.BigInteger( 1, md.
               digest() );
7          String password = hash.toString( 16 );
8          byte[] sha1 = new byte[password.length() / 2];
9          for (int i=0; i< sha1.length; i++){
10                 sha1[i] = (byte) Integer.parseInt(password.substring
                       (2*i, 2*i+2), 16);
11         }
12         com.mindprod.base64.Base64 base64 = new com.mindprod.base64.
               Base64();
13         result = base64.encode(sha1);
14 }
```

**Algoritmo B.1:** Método para transformar senha hexadecimal para base64

## B.2 Javascript para manipulação de e-mail

```
String result = null;
public void execute() {
        String novoemail = String.valueOf(email);
        if (novoemail != null) {
                java.util.regex.Pattern patternEmail = java.util.
                    regex.Pattern.compile("^([a-zA-Z0-9_\\-\\.]+)@
                    ((\\[[0-9] {1,3}\\.[0-9] {1,3}\\.[0-9]{1,3}\\.)
                    |(([a-zA-Z0-9\\-]+\\.)+))([a-zA-Z
                    ]{2,4}|[0-9]{1,3})(\\]?)$");
                java.util.regex.Matcher matchEmail = patternEmail.
                    matcher(novoemail);
                if (matchEmail.find()) {
                        result = novoemail;
                }
        }
}
```

**Algoritmo B.2:** Método para retirar os acentos do atributo e-mail

# Apêndice C

# Layouts

## C.1 Layout do Provedor de Serviço


**Figura C.1:** Site de acesso ao serviço da Atlases

## C.2 Layout do serviço WAYF


**Figura C.2:** Página de acesso ao WAYF da Federação CAFe

# Apêndice D

# Artigo

# Infraestrutura de Autenticação Única em Instituições de Ensino e Pesquisa Brasileiras

**Hendri Nogueira**

[1]Departamento de Informática e Estatística (INE)
Universidade Federal de Santa Catarina (UFSC) – Florianópolis, SC – Brasil

jimi@inf.ufsc.br

***Abstract.*** *CAFe Federation was created with the purpose to bring all the Brazilian institutions of education and research in a trusted network, in which each of them is responsible for authenticating and providing information of its users for authorized service providers. With a goal to ease the excess entries of users (and also login and passwords) to access a restricted access, there is a need to create an infrastructure that allows users to register, storing and managing your data into a single database, located in the institution and to access the resources provided by other domains, creating a single authentication for all services provided within the concept of academic federation.*

***Resumo.*** *A Federação CAFe surgiu com a meta de congregar todas as instituições de ensino e pesquisa brasileiras em uma rede de confiança, na qual cada instituição é responsável por autenticar e prover informações de seus usuários para provedores de serviços autorizados. Com objetivo de amenizar o excesso de cadastros dos usuários (e também login e senhas) para acessar um serviço de acesso restrito, há a necessidade de se criar uma infraestrutura que permita cadastrar usuários, armazenando e gerenciando seus dados em uma única base de dados localizado na instituição e acessar os recursos providos por outros domínios, criando assim uma autenticação única para todos os serviços providos dentro do conceito de federação acadêmica.*

## 1. Introdução

A infraestrutura de autenticação única foi modelada para atender o conceito de Federação Acadêmica e integrar todas as instituições de ensino e pesquisa do país.

### 1.1. Contextualização

Federação CAFe (Comunidade Acadêmica Federada) é o resultado dos primeiros esforços para a implantação de uma federação acadêmica no Brasil. Sua meta é congregar todas as universidades e instituições de pesquisa brasileiras em uma rede de confiança, na qual cada instituição é responsável por autenticar e prover informações de seus usuários (alunos, professores, técnicos, funcionários, etc.) para provedores de serviços (aqueles que oferecem serviços de acesso restrito) autorizados.

### 1.2. Objetivos

Este trabalho tem como objetivo geral, construir uma infraestrutura de autenticação e autorização dentro da UFSC (Universidade Federal de Santa Catarina) de acordo com o

*Acordo de Participação na Federação (Provedor de Identidade)* [RNP 2010b], tornando-a um provedor de identidade na Federação CAFe.

Dentre os objetivos específicos, se destacam: implementar uma base de dados centralizada; fornecer o modo de autenticação única para os usuários; tornar a UFSC um membro da Federação CAFe; gerar documentação técnica a respeito do que é uma federação.

### 1.3. Motivação

As motivações para realização deste trabalho são:

- Uso de um único sistema de controle de acesso para serviços internos e externos à instituição.
- Uma única conta (*login*) por usuário para acesso aos serviços.
- Integração com todas as instituições cadastradas na Federação CAFe.
- Garantia de privacidade dos dados pessoais no controle de acesso aos serviços.

### 1.4. Metodologia

Para realizar este trabalho, foi necessário em primeiro lugar a ambientação do autor à federação em questão e à estrutura que a instituição possua, a fim de conhecer as exigências da Federação CAFe, funcionalidades das ferramentas utilizadas e as aplicações providas pela instituição.

Inicialmente, houve a realização de um conjunto de estudos sobre a federação em si, seus serviços oferecidos, ferramentas que serão utilizadas e as obrigações exigidas pela RNP (Rede Nacional de Ensino e Pesquisa) [RNP 2010h]. Como exemplo destas ferramentas tem-se o LDAP, EID, EID2LDAP e Shibboleth.

Após o estudo, se inicia a construção da infraestrutura, com as instalações e configurações dos servidores e ferramentas, análise e construção de um novo banco de dados LDAP com base no banco de dados atualmente utilizado pela universidade.

O processo de configuração da infraestrutura é acompanhado com reuniões quinzenais pelo departamento responsável da RNP em conjunto com diversas instituições nacionais que também estão ingressando (ou já ingressaram) na Federação CAFe.

A parte de construção e população das bases de dados com informações das pessoas são supervisionados pelos responsáveis do NPD da UFSC [NDP 2010].

Após a conclusão do ingresso da UFSC na federação, é necessário divulgar a nova estrutura para a comunidade acadêmica permitindo assim o seu uso.

### 1.5. Limitações do Trabalho

Uma das limitações do trabalho está nas condições em como as diversas bases de dados da UFSC estão espalhadas e interligadas. A descentralização das bases de dados e a dificuldade de obter algumas informações necessárias de todas as pessoas vinculadas à UFSC, como o e-mail, exigiu em primeira instância sua não utilização na importação da base LDAP.

O manuseio dos dados de diversas pessoas é um processo complicado, pois existem alguns atributos restritos quanto à permissão, é o caso das senhas de usuários. Isso não permitiu um funcionamento imediato da infraestrutura para utilização pelos usuários, deixando a resolução deste problema para o futuro.

### 1.6. Organização deste Trabalho

Na próxima seção será descrito sobre a definição de Federação Acadêmica, que no Brasil existe a Federação CAFe. Sua descrição segue na seção 2.1, no qual em seguida, um breve detalhamento sobre a ferramenta de organização e armazenamento de dados na infraestrutura.

A infraestrutura de autenticação e autorização e seus elementos são ilustrados na seção 3. Após será detalhado sobre as ferramentas usadas para o auxílio na criação da estrutura do Provedor de Identidade implantada dentro da UFSC.

## 2. Federação Acadêmica

Federação acadêmica é um conceito que visa minimizar a manutenção de informações usadas para autenticação e autorização de pessoas para terem acesso aos serviços disponibilizados pelas instituições de ensino e pesquisa. Com a federação acadêmica, as informações sobre uma pessoa serão mantidas em uma única base, criando um vínculo entre a pessoa e a instituição.

Cada instituição necessita estabelecer seu modelo de gestão de identidade, ou seja, determinar a forma como as informações sobre as pessoas serão mantidas, atualizadas e os métodos de autenticação.

Existem diversas federações acadêmicas implementadas em diversos países como: InCommon [InCommon 2010], Feide [Feide 2010], Switch [Switch 2010] e UK [JANET(UK) 2010].

### 2.1. Federação CAFe

A Comunidade Acadêmica Federada (Federação CAFe) [RNP 2010c] teve início em 2008, como projeto piloto e com a meta de reunir todas as universidades e instituições de ensino e pesquisa brasileiras em uma rede de confiança.

O projeto de criação da Federação CAFe inclui ainda o estudo, a proposição, a análise e a validação de políticas para regular o funcionamento da federação (requisitos mínimos que provedores de identidade e de serviço deverão cumprir).

Assim como nas federações representadas nos diversos países, a Federação CAFe segue protocolos bem definidos na troca de mensagens entre provedores de identidade e serviço.

O SAML (*Security Assertion Markup Language*) [Maler and Hughes 2004] é um protocolo adotado em várias federações inclusive na CAFe, que em conjunto com o software Shibboleth [Internet2 2010], vem se firmando como um padrão *ad hoc* que permite a troca de informações de autenticação e autorização entre provedores de identidade e de serviço. A ferramenta Shibboleth será detalhada na seção 4.3.

Através de um componente denominado de WAYF (*Where Are You From*), que é centralizado e mantido pela RNP, os provedores de serviços poderão ser implantados por membros externos, ou seja, atuar apenas como provedores de serviços.

A rede de confiança permite que um usuário seja autenticado em sua instituição de origem e consiga acessar recursos e serviços oferecidos via *web* tanto pela própria instituição quanto por outros membros da federação, através de um único *login*.

Dentre os mais diversos benefícios oferecidos pela estrutura da federação, a participação de uma instituição como um provedor de identidade envolve:

- Uso de um único sistema de controle de acesso para serviços internos e externos à instituição.
- Uma única conta (login) por usuário para acessar todos os serviços.
- Manutenção de dados pessoais restrita à instituição de origem do usuário.
- Garantia de privacidade: apenas informações necessárias dos usuários serão passadas para os provedores de serviço.

Para provedores de serviços, os benefícios incluem:

- Simplificação do procedimento de controle de acesso, devido à:
  - Autenticação e disponibilização de informações sobre os usuários que são realizados pelos provedores de identidade, eliminando a necessidade de manutenção dessas informações no provedor de serviço.
  - Autorização para acesso a um recurso por um usuário que pode ser realizada pela característica do usuário como, tipo de vínculo, ou outro atributo disponibilizado pelo provedor de identidade.
- Redução de requisitos de suporte aos usuários.

## 2.2. Serviço de Diretório

Um diretório é uma lista de informações sobre objetos organizados ou catalogados em uma ordem, e fornece o acesso aos dados dos objetos. Permite que os usuários ou aplicações possam encontrar recursos no ambiente com características necessárias para um tipo de tarefa particular [Trigo 2007].

Um diretório é um banco de dados especializado, que pode ser chamado de repositório de informação, cujos registros de dados são definidos em forma de objetos e armazenados de forma ordenada. Sua principal característica é a forma em que os registros (objetos) e suas informações são acessados, sendo o acesso de leitura e pesquisa maior que o de escrita, diferente de um banco de dados relacional, no qual, os dados são constantemente atualizados, adicionados ou excluídos.

Diretórios contêm otimizações para suportar grande volume de acesso de leitura, e seu acesso à escrita deve ser limitado à administradores de sistema ou ao proprietário de cada parte da informação. Permitem também, que pessoas ou aplicações localizem usuários, recursos, serviços e informações em ambientes distribuídos.

Os diretórios podem diferir entre si no modo como a informação é representada e acessada, na flexibilidade com que a informação pode ser pesquisada e como pode ser estendida ou atualizada. Além da capacidade de controlar o acesso e autenticação, gerenciando o acesso às informações contidas nele.

### 2.2.1. LDAP

LDAP (Lightweight Directory Access Protocol) é um protocolo de serviço de diretório executado em TCP/IP, ou seja, na rede. É um padrão aberto, produzido pela IETF [IETF 2010] (Internet Enginneering Task Force) e a RFC 2251 [Group 1997] define esse protocolo.

Sendo o LDAP um protocolo de comunicação, ele define o transporte e o formato das mensagens usadas por um cliente para acessar os dados de um servidor de diretório de tipo X.500 [Ehlenberger et al. 2004].

Por definir um método para acessar e atualizar informações em um diretório, o LDAP obtém ampla aceitação como um método de acesso aos diretórios da internet, dando suporte para várias aplicações Web, intranet, navegadores, servidores IMAP, banco de dados, etc., além de fornecer mais funcionalidades para a rede de forma integrada com vários outros serviços como: Proxy, FTP, Apache, Samba, servidores de email, entre outros.

O primeiro protocolo criado para este fim foi o *Directory Access Protocol* (DAP). O DAP fazia parte das especificações X.500, desenvolvidas pela ITU Telecommunication. O DAP foi baseado no modelo OSI, que era extremamente difícil de ser implementado e resultava em aplicações complexas, lentas e de alto custo. Em 1993, a Força Tarefa de Engenharia da Internet (IETF) padronizou o LDAP como uma alternativa de acesso aos diretórios do X.500.

Existem várias implementações de servidores de diretórios, como por exemplo, MS Active Directory, IBM Lotus Domino e OpenLDAP, mas nem todas são compatíveis entre si, podendo apenas servir a um determinado *software* e possuir restrições de uso ou característica exótica.

O OpenLDAP, implementação mantida pela Fundação OpenLDAP, é um servidor LDAP de código aberto e de uso geral, ou seja, não agrega nenhum outro serviço que não tenha relação com a administração do diretório. Fundado em 1998, o projeto OpenLDAP foi baseado em uma implementação de servidor LDAP feita pela Universidade de Michigan. Deste modo, o OpenLDAP foi escolhido como servidor de diretório para o projeto e-AA [RNP 2010g].

O projeto e-AA é coordenado pela RNP e seu objetivo foi implantar a Federação CAFe com soluções técnicas e ferramentas desenvolvidas tanto no contexto do projeto como também em iniciativas anteriores apoiadas pela RNP.

O LDAP define quatro modelos básicos que descrevem por completo a sua operação, quais informações podem ser armazenadas em diretórios LDAP e o que pode ser feito com essas informações. São eles:

- Modelo de Informação: define o tipo de informação que pode ser armazenada em um diretório LDAP.
- Modelo de Nomes: define como a informação no diretório LDAP pode ser organizada e referenciada.
- Modelo Funcional: define o que pode ser feito com a informação no diretório LDAP e como ela pode ser acessada e alterada.
- Modelo de Segurança: define como a informação no diretório LDAP pode ser protegida de acessos ou modificações não autorizadas.

## 3. Infraestrutura de Autenticação e Autorização

Uma federação possui uma infraestrutura de autenticação e autorização interdomínios ou também chamada de infraestrutura de autenticação e autorização federada. Seu objetivo

é cadastrar um usuário, armazenando e gerenciando seus dados em uma única instituição (aquela no qual ele pertence), mas podendo ter acesso aos recursos oferecidos por outros domínios.

Uma infraestrutura de autenticação e autorização (IAA) federada é formada por dois tipos de provedores:

- Provedores de Identidades;
- Provedores de Serviços.

Os provedores de identidade são responsáveis por criar e manter cadastros e atualizar informações sobre as pessoas vinculadas à instituição, entre eles, os dados pessoais (nome, data de nascimento, CPF, nomes dos pais, sexo, etc.) e os vínculos internos (data de admissão, cargo ocupado, número de matrícula, número VoIP, etc.). O provedor de identidade necessita estabelecer seu método de autenticação interno e deve garantir que cada pessoa possua um identificador único.

Os provedores de serviços oferecem serviços em que seus acessos são restritos e ainda podem determinar privilégios de acessos baseados em atributos adicionais sobre o usuário, como por exemplo, o vínculo dele com a instituição (aluno, professor, técnico, etc.). Cabe ao provedor de serviço solicitar as informações adicionais (caso necessite) ao provedor de identidade para completar a autenticação.

O acesso de usuários de diferentes instituições nos provedores de serviços é comum, então há a necessidade de redirecionar os usuários para os respectivos provedores de identidade. O serviço centralizado e responsável por obter as informações sobre os provedores de identidade cadastrados na federação e seus respectivos redirecionamentos é o WAYF (*Where Are You From*). Nele o usuário seleciona sua instituição de origem, e passa a interagir com o seu provedor de identidade para fornecer as suas credenciais.

O WAYF contém o serviço de *Metadata*, do qual possui um arquivo metadado responsável por concentrar as informações dos provedores pertencentes à federação. Este é um arquivo de configuração padronizado e compartilhado entre os provedores de identidade e de serviço da federação.

O arquivo de metadado é disponibilizado pela federação e a sua função é estabelecer a relação de confiança entre os provedores, utilizando certificados digitais ou chaves públicas. Nesse arquivo são indicadas, para o provedor de identidade, as informações pertinentes sobre os provedores de serviço (e vice-versa). Desta forma, garante-se a segurança e a autenticidade na comunicação entre os provedores.

Além disso, o arquivo de metadados disponibiliza as informações relevantes para a comunicação entre os provedores, como identificadores, URLs e protocolos utilizados.

Ao acessar um serviço através do *browser* pela primeira vez, ele será redirecionado para WAYF e deverá selecionar o provedor de identidade (instituição) de origem. Após a seleção, navegador será redirecionado para o *site* de autenticação de sua instituição. Após a autenticação do usuário, o provedor de identidade repassa o resultado da autenticação ao provedor de serviço e cria uma sessão de uso associada ao usuário, de forma que acessos a novos serviços dentro de um determinado intervalo de tempo não gerem novas requisições de autenticação (*single sign-on*).

Para ilustrar melhor este processo, a figura 1 representa este processo de autenticação e informa os seguintes passos:

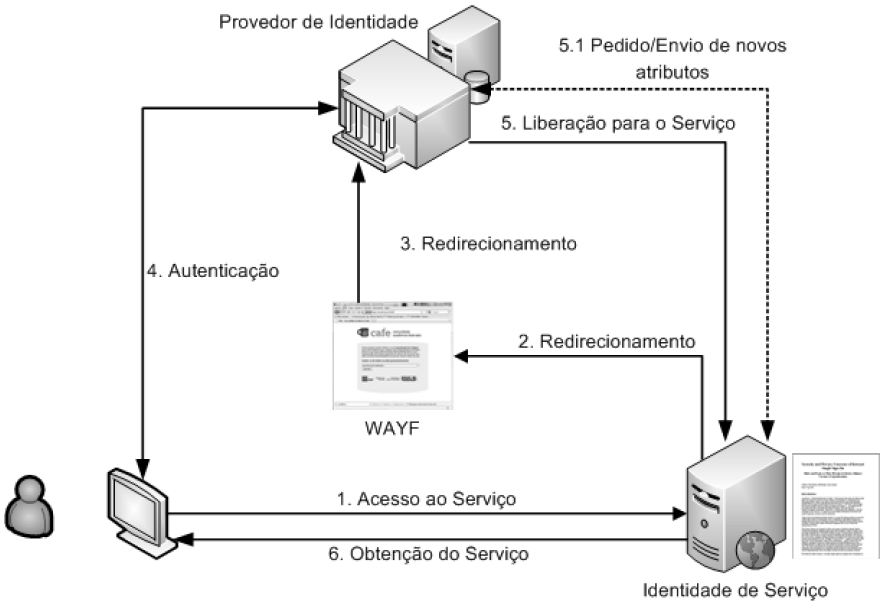


**Figura 1. Fluxo de informações**

1. O usuário acessa a um serviço desejado através do browser.
2. O servidor redireciona o navegador para o serviço de descoberta da federação (WAYF).
3. O usuário seleciona uma instituição oferecida na lista de instituições cadastradas, e seu navegador o redireciona para a página de autenticação da seleção.
4. O provedor de identidade da instituição envia ao navegador a página de autenticação do usuário, o qual se autentica através de *login* e senha.
5. O provedor de identidade gera um *handle* e o envia ao navegador, que o encaminha ao provedor de serviço, e assim obtém a prova de autenticação do usuário. Para algumas aplicações, isso é suficiente para autorizar o acesso do usuário ao serviço.
   (a) Opcionalmente, o provedor de serviço pode enviar um pedido para obtenção de novos atributos ao provedor de identidade, utilizando o *handle* para especificar o usuário em questão.
6. Devidamente autenticado, o provedor de serviço retorna a solicitação do usuário.

## 4. Estrutura do Provedor de Identidade

Um provedor de identidade deve construir uma infraestrutura apta para cumprir com todas as regras determinadas pela Federação CAFe e que possa ter um ótimo funcionamento.

Dentre os diversos aplicativos contidos nela, esta seção descreverá sobre aqueles com maior importância: esquema brEduPerson usando no LDAP, os aplicativos EID, EDI2LDAP, Shibboleth.

### 4.1. EID

Diretórios que possuem um baixo fluxo de cadastros de pessoas podem ser facilmente gerenciados pela inclusão e exclusão manual de registros. Já os diretórios com muitos

usuários e comportamento dinâmico necessitam de um esforço maior na manutenção, o que deixa seu gerenciamento manual mais complicado, como os diretórios acadêmicos.

O objetivo do EID (Export Import Directory) [UFMG 2010a] é facilitar a integração de dados de diversos sistemas para construir um metadiretório (uma base relacional intermediária entre as fontes efetivas dos dados e o diretório) e em seguida um ou mais diretórios, ou seja, ferramenta para auxiliar nas funcionalidades de importação e exportação de dados para outras fontes.

Desenvolvido pelo Grupo São Tomé da UFMG, teve base na ferramenta Pingifes-Import [PINGIFES 2010], ferramenta de extração, transformação e carga (ETL), utilizadas pelas instituições de ensino superior para alimentação do modelo de dados (PingIFES) definido pelo MEC.

A ferramenta EID permite importar dados de bancos relacionais, desde que exista um drive JDBC ou ODBC para fazer a conexão e/ou arquivos de texto CSV. O EID disponibiliza um serviço *web* para exportação e consulta de dados, o que facilita o acesso por aplicações que utilizem tecnologias diversas.

O serviço *web* serve de base para outras ferramentas de exportação também, como o EID2LDAP. Os dados importados são associados às pessoas e os registros completos podem ser facilmente recuperados através da interface *web* do EID.

Todo objeto é identificado globalmente através do GUID, que é gerado automaticamente pela ferramenta. Esse atributo é definido pela classe especial *EidObject*.

O GUID é gerado ao importar a primeira classe denominada de Identificação (possui informações básicas referente às pessoas). Sua geração é realizada a partir da escolha de um atributo que seja único para cada indivíduo, tornando assim um ponto de referência na base.

As demais classes extraídas usarão o atributo único para conciliar com o GUID determinado, interligando o registro principal sobre a pessoa com as funções e seus atributos, como funções que ela exerce na instituição (aluno, professor, técnico, etc.), e-mails que a pessoa possui, endereços, telefones para contato, VoIPs, etc.

Toda classe criada na aplicação gera uma tabela no banco de dados. A classe *EidObject* se relaciona com as demais classes do sistema denominadas *EidClass*. Essas classes agregam a um objeto EID e seus atributos específicos.

O EID já fornece algumas classes que podem ser usadas para alimentar diretórios LDAP sem nenhuma configuração adicional, isso porque existe uma conversão pré-configurada para a ferramenta EID2LDAP que definem os atributos necessários para o esquema *brEduPerson*. Outras classes podem ser definidas pela própria organização, para suprir suas necessidades. Estas modificações deverão também ser realizadas na conversão utilizada pelo EID2LDAP.

### 4.2. EID2LDAP

O EID2LDAP [UFMG 2010b] é uma ferramenta que busca informações de diretórios armazenados no metadiretório EID e as transfere para servidores LDAP, além de possuir uma interface *web*.

A ferramenta EID possibilita extrair e incorporar ao seu metadiretório, dados de

diversas bases institucionais e os disponibilizam em arquivos XML através de um serviço *web*. O EID2LDAP por sua vez acessa estas informações, utilizando a marcação XSLT (*Extensible Stylesheet Language Transformations*) para especificar a transformação dos dados para o formato LDIF (*Ldap Data Interchange Format*) compatível com o formato do servidor LDAP de destino, e realiza a atualização do servidor LDAP.

O XSLT introduz flexibilidade no EID2LDAP, permitindo ao usuário definir como se dará o mapeamento entre os dados do EID e o formato do LDAP.

Assim como o EID, o EID2LDAP permite o agendamento periódico das exportações. Para cada exportação são atualizados apenas os registros modificados, ou inseridos, ou desativados desde a última importação.

No momento da exportação, caso o registro já exista no LDAP (identificado pelo DN gerado no LDIF), o LDIF é modificado para aplicar operações de alteração (*modify*) no registro do LDAP.

Apenas os *objectClasses* representados no LDIF serão substituídos no LDAP, isto é, os *objectClasses* no LDAP passarão a ter os atributos com os mesmos valores do EID, enquanto outros *objectClasses* permanecerão com seus atributos inalterados. Isto possibilita que outras aplicações alimentem diretamente o diretório sem a necessidade de passar pelo EID.

### 4.3. Shibboleth

O Shibboleth é um sistema baseado em padrões para *web single sign-on*, ou seja, o usuário necessita se autenticar apenas uma vez, para posteriormente, ter acesso automaticamente aos outros serviços com a mesma política de autenticação dentro das fronteiras organizacionais. É um projeto da Internet2 Middleware Initiative, código aberto e permite que serviços *web* informem decisões de autorização para acessos individuais de recursos protegidos *on-line* de forma que preserve a privacidade do usuário.

O software consiste na implementação de padrões amplamente utilizados para autenticação e autorização federada via *web*, principalmente o SAML (*Security Assertion Markup Language*) criado pela OASIS (*Organization for the Advancement of Structured Information Standards*).

Um usuário se autentica com sua credencial referente à sua organização de origem. A organização (provedor de identidade) passa a mínima quantidade de informações de identificação necessárias para que o gerente de serviço habilite a decisão de autorização.

Shibboleth é composto por dois módulos: Provedor de Identidade e Provedor de Serviço. O primeiro é responsável pela autenticação e o segundo pela autorização. Numa federação Shibboleth, normalmente apresenta mais dois componentes: o serviço WAYF (*Where Are You From*), usado para localizar o provedor de identidade do usuário e o serviço de *Metadata*, responsável em concentrar as informações dos provedores pertencentes à federação.

O Shibboleth foi desenvolvido para tratar os seguintes desafios:

- Acabar com as múltiplas senhas requeridas para múltiplas aplicações;
- Suportar a escalabilidade no gerenciamento de múltiplas aplicações;
- Melhorar a segurança associada ao acesso de serviços de terceiros;

- Aumentar a privacidade dos dados dos usuários;
- Suportar a interoperabilidade dentro e entre organizações;
- Permitir a liberdade de escolha das tecnologias de autenticação para as instituições;
- Permitir o controle de acesso efetuado a partir dos provedores de serviço.

## 5. Implantação do Provedor de Identidade

Nesta seção será descrito todo o processo de instalação da infraestrutura do provedor de identidade aplicada na UFSC. Estes processos foram realizados com base nos tutoriais realizados pela RNP localizados na página wiki referente ao projeto CAFe [RNP 2010j].

### 5.1. Instalação do Servidor

A instalação do servidor foi realizada através da criação de uma máquina virtual e com o sistema operacional Ubuntu Server 8.04. Este sistema foi escolhido pela recomendação do projeto e-AA [RNP 2010g], pois disponibiliza pacotes Java versão 6 nativamente e terá suporte para atualizações de segurança até abril de 2013.

Além da instalação básica e padrão do sistema, foi realizado:

- Instalação do pacote OpenSSH: conjunto de softwares que provém à criptografia em sessões de comunicação em uma rede de computadores usando o protocolo SSH/TLS.
- Configuração do `hostname` e do arquivo `/etc/hosts`. Para determinar o nome local do servidor e o IP do qual o servidor se reconhecerá. É importante que a máquina esteja utilizando um IP fixo para evitar eventuais problemas decorrentes da indisponibilidade do servidor DHCP.
- Configuração do *firewall* adicionando linhas de comando para liberar apenas as portas dos principais serviços que serão usados, como: SSH, NTP (Serviço de sincronização do relógio), Apache, Tomcat.
- Configuração do NTP para acessar um servidor de relógio e sempre mantê-lo sincronizado.

### 5.2. Instalação do Diretório com Esquema BrEduPerson

Após a configuração completa do sistema, foi instalado o serviço de diretório (LDAP) e também configurado para usar o esquema brEduPerson pelos seguintes passos:

- Instalação do pacote slapd.
- Abertura das portas no *firewall* que serão usadas pelo serviço slapd.
- Edição do arquivo de configuração do servidor ldap (`slapd.conf`): É importante citar que nesse arquivo foram incluídos os seguintes esquemas LDAP: Core; Cosine; Nis; InetOrgPerson; EduPerson; BrEduPerson; Schac.
- Geração dos certificados SSL para o LDAP com as seguintes características: tamanho da chave de 2048 bits e validade de 730 dias.
- Inclusão de parâmetros no arquivo `slapd.conf` para suportar conexão segura usando o certificado criado.
- Edição do arquivo de configuração do cliente ldap (`ldap.conf`).
- Criação de listas de permissões de leitura e escrita no arquivo `sldap.conf` para o administrador da base, acesso de leitura para o serviço shibboleth, e o bloqueio de todas as outras permissões.
- Inclusão da estrutura inicial da base LDAP, do usuário administrador e usuário de leitura usado pelo Shibboleth.

### 5.3. Extração dos dados para o Metadiretório

A extração dos dados para o metadiretório é realizado pelo EID e por isso há a necessidade de sua instalação. Para isso, foi necessário:

- Instalar o Java, o Tomcat e o MySQL.
- Criar as bases de dados eid e pcollect para a utilização do EID.
- Descompactar e compilar o arquivo do EID [UFMG 2010a].
- Criar o arquivo `/etc/tomcat6/Catalina/localhost/eid.xml` [RNP 2010e] e editar de acordo com as informações locais: esse arquivo é responsável pela configuração e deploy do EID.

  Esta extração necessita que o administrador tenha alguns conhecimentos, como:

- A organização das bases de dados da instituição que serão utilizadas, como bases de alunos, professores, recursos humanos, etc.
- Verificar as classes e os atributos necessários para o esquema brEduPerson.

O EID necessita se conectar ao servidor de onde serão retirados estes dados. Com isso precisa-se cadastrar cada repositório, informando o tipo de banco de dados (relacional ou via arquivo CSV), a URL, o *driver* utilizado, o usuário com permissão ao banco e sua senha.

Após os cadastramentos dos bancos de dados, deve-se criar as extrações para alimentar as classes definidas. É importante saber que a primeira classe a ser carregada deve ser a classe Identificação, pois essa classe será o container principal de uma pessoa (na base LDAP) e a partir dele que será gerado o GUID.

Na configuração da extração da classe principal, deverá determinar um atributo que seja único para cada indivíduo e utilizá-lo para ser o ponto de referência da geração do GUID.

A partir do identificador único de cada pessoa serão criados os relacionamentos às demais classes que informarão a conta (login e senha) e a(s) sua(s) função(ões) dentro da universidade (aluno, professor, funcionário).

Extraído a classe principal Identificação, poderá extrair as demais. A conciliação delas com a principal é baseada no atributo GUID. No momento da extração de cada um das demais classes, há a necessidade de realizar o *select* do mesmo atributo único da classe Identificação e conciliar com o atributo GUID.

O banco de dados do EID irá verificar se existe aquele atributo único na forma de GUID, e caso exista, realizará a conciliação do novo registro com o registro principal.

Caso o atributo único não exista em forma de GUID, a extração acusará um erro dizendo que não foi encontrado o valor em questão no banco e apenas funcionará quando o registro principal estiver incluído.

A segunda classe a ser extraída deve ser a classe Conta. Apenas depois que uma pessoa, ou seja, um registro tiver sido conciliado, entre pelo menos a classe Identificação e Conta, o EID2LDAP estará apto para importar os dado do metadiretório EID para a base LDAP.

A configuração das extrações das classes é realizada com a criação de uma ETC. Numa ETC, será determinado uma consulta SQL no repositório cadastrado para referenciar os atributos da classe com os resultados obtidos.

O EID possibilita que os dados adquiridos da seleção sejam tratados por meio scripts em Javascript para serem convertidos de acordo com a necessidade do administrador.

Alguns desses scripts são por exemplo na transformação dos hash SHA-1 e MD5 das senhas na sequência hexadecimal para base64. Outro script de transformação é realizado no atributo e-mail que pode estar com sintaxe errada (com caracteres acentuados) e deve-se removê-los de forma automática.

Para cada ETC criada, tem-se que criar processos de extração. Esses processos servem para que o agendamento saiba o que será executado. A criação do agendamento permite escolher qual ETC deseja executar e também o modo de repetição (nenhum, diário, mensal, anual, etc.), a data e horário da execução.

### 5.4. Alimentação do Diretório a partir do Metadiretório

Esta seção apresenta os passos necessários para realizar a instalação da ferramenta EID2LDAP no qual exporta os registros do metadiretório EID para a base LDAP. Para isso, teve-se que:

- Criar uma base de dados que será utilizada pelo EID (seu nome deve ser eid2ldap).
- Fazer a carga da base de dados para preparação do EID2LDAP.
- Adquirir a versão mais recente do EID2LDAP [UFMG 2010b].
- Descompactar o arquivo adquirido.
- Criar o arquivo `/etc/tomcat6/Catalina/localhost/eid2ldap.xml` [RNP 2010d] que será responsável pela configuração e deploy do EID.

Após a instalação e configuração de todos os requisitos anteriores, pode-se acessar o serviço EID2LDAP através do endereço: http://hostname:8080/eid2ldap/Home.faces. Realizando o login do serviço, deve-ser determinar as seguintes ações:

- Ajustar o endereço do *web service* do EID em http://hostname:8080/eid/services/EidService?wsdl.
- Iniciar o agente exportador através do menu Agendamento / Agente Gerenciador De Agendamento.
- Configurar qual servidor LDAP será usado para importar os dados do metadiretório para base LDAP, no menu Configuração / Servidor Ldap: informar o usuário (na forma de DN) e sua senha de acesso.
- Agendar para a realização da atualização do diretório.

A conclusão dos processos de importação pelo agendamento pode ser visualizado pelos logs gerados em cada agendamento concluído. Tanto o EID quanto EID2LDAP permitem esse monitoramento. Caso seja necessário fazer a verificação do dados na base LDAP, pode-se realizar uma pesquisa por meio de algum programa de gerenciamento LDAP ou através do comando *ldapsearch*.

### 5.5. Instalação do Shibboleth IdP

O Shibboleth IdP representa o principal item de um provedor de identidade, pois é ele quem habilitará a conversa com os provedores de Serviço e WAYF. A versão instalada foi o Shibboleth-IDP 2.x e em uma nova máquina virtual. O processo completo de sua instalação pode ser visualizado na página wiki referente ao projeto CAFe e na parte de instalação Shibboleth-IDP 2.x [RNP 2010f]. Os principais processos de instalação são:

- Instalação do Java, Tomcat e Apache.
- Editar o arquivo de configuração de segurança do java (`/etc/java-6-sun/security/java.security`) e adicionar quatro bibliotecas.
- Editar o arquivo `/etc/tomcat6/server.xml` para definir que o Shibboleth-IDP irá receber conexões pela porta 8443, ou seja, conexões cifradas.
- Criar o arquivo `/etc/tomcat6/Catalina/localhost/idp.xml` contendo o caminho para `auto-deploy` do Shibboleth-IDP.
- Configurar o Apache para habilitar o portal de autenticação da instituição.
- Criar o arquivo `/etc/apache2/conf.d/idp.conf` para configurar a ligação entre o Apache e Tomcat.
- Obter o Shibboleth-IDP.
- Obter o arquivo `bcprov-jdk16-144.jar`, que é uma biblioteca que implementa protocolos criptográficos em Java.
- Instalar o Shibboleth-IDP.
- Editar o arquivo `/opt/shibboleth-idp/conf/handler.xml`.
- Editar o arquivo `/opt/shibboleth-idp/conf/relying-party.xml`.
- Configurar a resolução/liberação de atributos recomendados pelo projeto CAFe [RNP 2010i].
- Configurar o arquivo `/opt/shibboleth-idp/conf/login.config` para permitir que o usuário para o uso do shibboleth criado na base LDAP possa acessar os dados da base.
- Gerar um par de chaves criptográficas com tamanho de 2048 bits e validade de 1095 dias para ser usado na conexão segura do Shibboleth-IdP.
- Gerar um par de chaves criptográficas com tamanho de 2048 bits e validade de 1095 dias para ser usado na conexão segura do Apache.
- Editar o arquivo `/opt/shibboleth-idp/metadata/idp-metadata.xml` para incluir o certificado criado para Shibboleth-IDP e permitir que o *metadata* se identifique com ele.

O serviço de autenticação *web* pré-configurada através de uma página de *login*, é instalada automaticamente por esse processo de instalação. Apesar disso, há a possibilidade de personalizar a página de *login* da instituição realizando a mudança do *template* nos arquivos localizados no diretório `/src/main/webapp/` da instalação do Shibboleth-IDP.

A página de *login* da UFSC foi modificada de forma simples incluindo a identidade visual da universidade e assim visualizando o funcionamento da mudança.

### 5.6. Entrada na Federação CAFe

Após a instalação de todos os pré-requisitos e os softwares necessários, o Shibboleth-IdP ainda não estará pronto para ser reconhecido pela federação, pois é necessário que

o provedor de identidade envie o seu metadado para ser cadastrado no WAYF e assim permitir o seu reconhecimento.

A RNP possui uma federação de teste denominado de Federação Chimarrão. Antes de um provedor (identidade ou serviço) entrar na Federação CAFe, é necessário entrar na Federação Chimarrão e seu serviço WAYF necessita cadastrar o metadado do provedor em questão.

Para validar a conclusão da etapa de entrada na Federação Chimarrão, é necessário realizar um teste padrão. Este teste consiste em acessar o endereço `https://chimarrao.ufrgs.br/homologa` e selecionar a IdP da instituição de teste. Será realizado um redirecionamento para a página de autenticação da instituição e após preencher o login e a senha com algum usuário cadastrado na base LDAP, deve-se observar se houve sucesso no retorno dos atributos do usuário autenticado. Se os atributos foram visualizados sem nenhum problema, então o teste foi bem sucedido.

Agora, para entrar na Federação CAFe, há a necessidade de configurar o shibboleth para o reconhecimento da federação [RNP 2010a] . Os principais passos são:

- Atualizar o certificado para a conexão HTTPS de `ds.cafe.rnp.br` no keytool do sistema para que os metadados sejam atualizados.
- Modificar o arquivo `/opt/shibboleth-idp/conf/relying-party.xml` para incluir o reconhecimento do metadado referente à CAFe.
- Certificar que o arquivo `/opt/shibboleth-idp/conf/attribute-filter.xml` está configurado para liberar os mesmos atributos para as federações Chimarrão e CAFe.

Com isso, poderá realizar a solicitação da entrada do provedor de identidade na Federação CAFe. A instituição estará automaticamente cadastrada tanto na Federação Chimarrão quanto na CAFe e assim poderá usufruir dos serviços oferecidos por todos os provedores de serviço cadastrados em ambas as federações.

## 6. Conclusões

Diante deste trabalho, foram encontrados várias dificuldades que exigiram espaços de tempo diferentes para sua resolução. Apesar de existir um processo descritivo sobre a implementação da infraestrutura de um provedor de identidade no site da RNP, isso não tornou o trabalho mais fácil.

As dificuldades se devem a falta de documentação, ou documentação insuficiente, exigindo assim um processo de dedução e com ajuda de terceiros. Também se enquadram nesta categoria a falta de conhecimento nas estruturas de bases de dados da UFSC e na sua imensidão, necessitando tempo e discussões com as partes responsáveis dentro da instituição.

As ferramentas EID e EID2LDAP não são conhecidas e usada por todos, pois suas aplicações são somente para estes fins, e isso trouxe alguns problemas tanto de conflitos entre os serviços requisitados (Java e Tomcat), quanto no próprio funcionamento interno deles.

A demanda de maior tempo de dedicação e a necessidade de discutir com os criadores do software para obter um conhecimento mais profundo conseguiram solucionar esses problemas.

O processo de instalação e configuração da infraestrutura exigiu a necessidade do entendimento e levantamento da estrutura da base de dados da instituição (UFSC), e determinar qual melhor estrutura para o metadiretório EID e posteriormente da configuração do EID2LDAP.

Uma base acadêmica possui alguns dados restritos, por exemplo senhas de usuários. Para não violar a integridade e nem a segurança do mesmo, não houve a utilização dos valores deste dado para compor a base LDAP.

Uma base de dados acadêmica possui muitos registros decorrente das informações de milhares de pessoas que possuem ou possuíram algum vínculo com a instituição. Devido a essa imensa quantidade de dados, o EID e o EID2LDAP necessitaram desde vários minutos a até algumas horas para poder realizar o processamento da atualização dos dados no metadiretório, a conciliação de todos os registros percorridos e a importação dos novos dados na base.

Por isso, há um cuidado no agendamento dos processos em que eles devem ser calculados para existir um tempo de processamento e conciliação grande o suficiente entre o agendamento de um processo e outro. Caso este tempo seja curto, poderá ocorrer um acúmulo de recursos de processamento em que o servidor poderá travar seus processos. Assim os serviços realizados pelo EID e EID2LDAP não funcionará corretamente.

A infraestrutura de autenticação e autorização criada pela RNP e o conceito de federação acadêmica proporciona que todas as instituições de ensino e pesquisa se interliguem por meio de uma rede de confiança e possibilitem a troca de serviços entre elas.

Além desta conexão mutua entre as partes, criou-se uma estrutura com o uso de uma base única de dados (pelo menos a UFSC que não possuía), ou seja, uma base centralizada em que os dados são obtidos a partir de diversas outras fontes. Deste modo, a parte de gerenciamento do dados (adição, remoção, atualização) é feita de forma automática por estar conectada com as outras.

Foram realizadas as extrações dos dados referentes às classes padrões do EID (Identificação, Conta, Aluno, Professor, Técnico, Email), e a criação de uma nova classe denominada de Comunidade para incluir pessoas que possuem algum vínculo com a universidade por meio de outras atividades, como aluno especial, ouvinte ou possuindo alguma matrícula isolada.

Os serviços podem ser providos por qualquer entidade, basta que siga as regras da federação. Uma instituição pode interligar todos os seus principais serviços, como e-mail, internet, *wireless*, acesso a livros na biblioteca, periódicos, VoIP e até acesso ao restaurante universitário, para autorizar com apenas umas autenticação.

O conceito de federação proporciona uma maior segurança tanto na comunicação entre as partes quanto no acesso aos dados de uma autenticação. Toda comunicação é feita de forma cifrada em que todos os provedores (IdP e SP) necessitam possuir certificados e configurar o Shibboleth para utilizá-los na comunicação.

A segurança na autenticação vem do princípio em que os provedores de identidade apenas irão repassar atributos necessários do usuário para os provedores de serviço, indicando uma maior privacidade dos dados do usuário. Os serviços ainda poderão deter-

minar diferentes níveis de acessos para a autorização, baseando-se apenas nos valores dos atributos.

Os provedores de serviço e identidade possuem responsabilidades distintas e bem definidas, o que torna o processo de comunicação mais segura. Quanto maior for o número de instituições cadastradas na Federação CAFe, melhor será, pois o acesso aos serviços não terá fronteiras.

Diante de uma estrutura já criada do provedor de identidade, ainda há a necessidade de realizar melhorias no gerenciamento de senhas. Para não quebrar a segurança das senhas já cadastradas no banco de dados, temos a necessidade de criar novas senhas para este novo fim. Então poderá criar um serviço de gerenciamento de senhas de cada usuário vinculado à UFSC, em que o usuário irá primeiramente cadastrar sua senha, e posteriormente poderá realizar o pedido de modificação da mesma.

Além da UFSC ser um provedor de identidade, há também a necessidade de torná-la um provedor de serviço para poder integrar todos os serviços internos. Alguns dos principais serviços que inicialmente poderão ter suporte para este tipo de autenticação seriam os serviços de e-mails e moodle. Estes dois serviços em especial, já possuem compatibilidade e integração com os serviços do Shibboleth, facilitando o processo.

O aumento da segurança é sempre um fator importante no que se diz a respeito à segurança da informação. O uso de certificação digital [Housley and Polk 2001] no processo de autenticação dificulta ainda mais na quebra de segurança da senha e também possibilitando o uso de um *token* ou *smartcard*.

## Referências


Ehlenberger, A., Yang, C., Leiserson, J., Owen, N., Storrs, M., Gorthi, R., Macbeth, R., Tuttle, S., and Ranahandola, S. (2004). *Understanding LDAP - Design and Implementation*. IBM.

Feide (2010). Feide. http://www.feide.no/.

Group, N. W. (1997). Lightweight directory access protocol (v3). http://www.ietf.org/rfc/rfc2251.txt.

Housley, R. and Polk, T. (2001). *Planning for PKI: Best Practices Guide for Deploying Public Key Infrastructure*. John Wiley & Sons, Inc., New York, NY, USA.

IETF (2010). The internet engineering task force. http://www.ietf.org/.

InCommon (2010). Incommon federation. http://www.incommonfederation.org/.

Internet2 (2010). Shibboleth - federated single sign-on software. http://shibboleth.internet2.edu.

JANET(UK) (2010). Uk federation. http://www.ukfederation.org.uk/.

Maler, E. and Hughes, J. (2004). Security assertion markup language(saml) 2.0 technical overview. http://www.oasis-open.org/committees/documents.php?wg_abbrev=security.

NDP (2010). Núcleo de processamentos de dados. http://setic.ufsc.br/.

PINGIFES (2010). Pingifes. http://pingifes.mec.gov.br/pingifes/.

RNP (2010a). Alterar configuração do idp para migrar para federação cafe. http://wiki.rnp.br/pages/viewpage.action?pageId=41616243.

RNP (2010b). Cafe – acordo de participação na federação (provedores de identidade). http://www.cafe.rnp.br/cafewiki/images/b/bc/CAFe_AcordoIdP_v1.0.pdf.

RNP (2010c). Federação cafe. http://www.cafe.rnp.br/wiki/Federação_CAFe.

RNP (2010d). Instalação do eid2ldap. http://wiki.rnp.br/pages/viewpage.action?pageId=42143530.

RNP (2010e). Instalação eid. http://wiki.rnp.br/pages/viewpage.action?pageId=41190354.

RNP (2010f). Instalação shibboleth-idp 2.x. http://wiki.rnp.br/pages/viewpage.action?pageId=41616303.

RNP (2010g). Projeto e-aa. http://www.cafe.rnp.br/wiki/Projeto_e-AA.

RNP (2010h). Rede nacional de ensino e pesquisa. http://www.rnp.br.

RNP (2010i). Resolução e liberação de atributos. http://wiki.rnp.br/pages/viewpage.action?pageId=41616318.

RNP (2010j). Roteiro de atividades para entrada de um idp na cafe. http://wiki.rnp.br/pages/viewpage.action?pageId=41616281.

Switch (2010). Switch. http://www.switch.ch/.

Trigo, C. H. (2007). *OpenLDAP: uma abordagem integrada*. Novatec, São Paulo.

UFMG, L. (2010a). Eid. http://sourceforge.net/projects/eid/.

UFMG, L. (2010b). Eid2ldap. http://sourceforge.net/projects/eid2ldap/.