

MILENA CONSTANTINO CAIRES

**AVALIAÇÃO E IMPLEMENTAÇÃO DE PROPOSTAS DE MELHORIA
PARA O PROTOCOLO IRIS BASEADAS EM TECNOLOGIAS DE
WEB SEMÂNTICA**

**São Paulo
2007**

MILENA CONSTANTINO CAIRES

**AVALIAÇÃO E IMPLEMENTAÇÃO DE PROPOSTAS DE MELHORIA
PARA O PROTOCOLO IRIS BASEADAS EM TECNOLOGIAS DE
WEB SEMÂNTICA**

Dissertação apresentada ao Departamento de Engenharia de Computação e Sistemas Digitais da Escola Politécnica da Universidade de São Paulo, como parte dos requisitos para obtenção do título de Mestre em Engenharia.

**São Paulo
2007**

MILENA CONSTANTINO CAIRES

**AVALIAÇÃO E IMPLEMENTAÇÃO DE PROPOSTAS DE MELHORIA
PARA O PROTOCOLO IRIS BASEADAS EM TECNOLOGIAS DE
WEB SEMÂNTICA**

Dissertação apresentada ao Departamento de Engenharia de Computação e Sistemas Digitais da Escola Politécnica da Universidade de São Paulo, como parte dos requisitos para obtenção do título de Mestre em Engenharia.

Área de Concentração: Sistemas Digitais

Orientadora: Profa. Dra. Graça Bressan

**São Paulo
2007**

Este exemplar foi revisado e alterado em relação à sua versão original, sob responsabilidade única do autor e com a anuência de seu orientador.

São Paulo, 5 de agosto de 2007.

Assinatura do autor

Assinatura do Orientador

FICHA CATALOGRÁFICA

Caires, Milena Constantino

Avaliação e implementação de propostas de melhoria para o protocolo IRIS baseadas em tecnologias de Web semântica / M.C.Caires. – ed. rev. – São Paulo, 2007.

p.141

Dissertação (Mestrado) – Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia de Computação e Sistemas Digitais.

1. Protocolos de comunicação (Análise). I. Universidade de São Paulo. Escola Politécnica. Departamento de Engenharia de Computação e Sistemas Digitais. II.t.

DEDICATÓRIA

Dedico este trabalho à minha querida mãe Altéia.

AGRADECIMENTOS

À minha família pelo constante apoio e estímulo transmitido durante todo o trabalho, mesmo quando um oceano nos separava.

Ao meu querido Alexander Schutz que sempre esteve ao meu lado em todos os momentos difíceis, com a solidária e providencial ajuda para a conclusão do trabalho.

À professora Dra Graca Bressan, pela orientação e pelo conhecimento transmitido sobre metodologia de pesquisa.

Ao professor Dr. Demi Getschko e ao NIC.br pelo apoio e pela oportunidade de apresentar meu artigo científico no congresso internacional WWW/Internet 2006 em Múrcia, na Espanha.

Ao meu supervisor Dr. Siegfried Handschuh pela oportunidade de realizar o intercâmbio na Universidade Nacional da Irlanda em Galway, onde foram adquiridos os conhecimentos específicos sobre a Web Semântica.

Aos meus amigos Andreas Harth, Christiane Meiler Baptista, Ian Korolkovas, Klaus Steding-Jessen, Marcelo Gardini do Amaral, Raquel Plec Manao, e a todos que colaboraram direta ou indiretamente, na execução deste trabalho.

RESUMO

O objetivo desta dissertação é avaliar se as tecnologias de Web Semântica podem contribuir para o desenvolvimento do protocolo *Internet Registry Information Service Protocol* (IRIS). IRIS é um novo protocolo para o serviço de informação sobre registros da Internet. Ele ainda encontra-se em fase de desenvolvimento por um grupo de trabalho do *Internet Engineering Task Force* (IETF). O objetivo do grupo de trabalho é desenvolver e padronizar um novo protocolo para substituir o protocolo Whois. Whois é o protocolo padrão utilizado atualmente para serviços de informação sobre registros da Internet, por exemplo nomes de domínios, endereços *Internet Protocol* (IP), sistemas autônomos, dentre outros. A principal motivação para o desenvolvimento do novo protocolo foi a crescente preocupação com a segurança dos dados armazenados na base de dados Whois pois o protocolo Whois não provê nenhum mecanismo de segurança. Outro motivo foi a ausência de suporte a base de dados distribuída porque o protocolo Whois foi desenvolvido para uma base de dados centralizada e, conseqüentemente, não atende aos requisitos padrões para protocolos da Internet. Até agora, o grupo de trabalho abordou e solucionou dois dos principais problemas do protocolo Whois: (1) segurança e (2) suporte à base de dados distribuída. Entretanto, o desenvolvimento de um novo padrão envolve um grande investimento da comunidade, em particular com respeito a políticas baseadas em consenso. Além disso, existe uma grande barreira a ser vencida para a adoção do novo protocolo: a adoção pelos usuários. O novo protocolo deve ter longevidade sem necessidade de atualização ou substituição por outro protocolo. Para atingir esse objetivo, é preciso não apenas satisfazer necessidades imediatas, como segurança, mas prever necessidades futuras. Este estudo envolveu as seguintes atividades de pesquisa: (1) análise comparativa dos atuais protocolos de busca de informação sobre registros da Internet, (2) o estudo aprofundado do protocolo IRIS e (3) a avaliação de novas tecnologias que pudessem ser incorporadas ao novo protocolo, em particular tecnologias de Web Semântica. Os resultados deste estudo demonstraram que as tecnologias de Web Semântica garantiriam a flexibilidade e extensibilidade necessárias para que o protocolo possa se adaptar às necessidades atuais e futuras. Para validar os resultados teóricos do estudo foi implementado um protótipo baseado na especificação do protocolo IRIS utilizando tecnologias de Web semântica. Dois tipos de experimentos foram conduzidos: (1) experimentos comparando os desempenhos do protótipo e do cliente Whois e (2) avaliação de desempenho do protótipo baseada em testes de carga. Finalmente, a implementação do protótipo e subseqüentes experimentos serviram como prova de conceito de que as tecnologias de Web Semântica podem contribuir para o sucesso do protocolo IRIS.

Palavras-Chave: Web Semântica. Protocolo IRIS. Protocolo Whois.

ABSTRACT

The aim of this thesis is to evaluate whether Semantic Web technologies can contribute to *Internet Registry Information Service Protocol* (IRIS) protocol development. IRIS is a new protocol for providing an information service for Internet resources. It is currently still under development by an *Internet Engineering Task Force* (IETF) working group. The objective of the working group is to develop and standardize a new protocol to replace the Whois protocol. Whois is the standard protocol used today by information services for Internet resources, i.e. domain names, *Internet Protocol* (IP) addresses, autonomous systems, amongst others. The motivation to develop a new protocol was based on increasing concerns regarding the security of data stored in the Whois database as the Whois protocol does not provide any security mechanism. Another motivation was the absence of support for distributed databases as the Whois protocol was developed for a centralized database, hence it no longer meets the standard requirements for Internet protocols. So far, the working group has tackled and solved two main issues concerning the Whois protocol: (1) security and (2) support for distributed databases. However, the development of a new standard demands a great investment from the community, in particular with respect to consensus-based policies. Additionally, there is one major barrier against adopting the new protocol: the users adoption. The new protocol must have longevity without being updated or replaced by another protocol. To reach this goal, it is necessary to meet not only the current requirements, such as security issues, but to cater also for future requirements. This thesis is concerned with the following research activities: (1) comparative analysis of the current protocols used to provide information services on Internet resources, (2) the IRIS protocol analysis and (3) the evaluation of new technologies that could be incorporated in the new protocol, in particular Semantic technologies. The results demonstrate that Semantic Web technologies could provide the necessary flexibility and extensibility to meet the current and future requirements of IRIS. To validate the theoretical results a prototype based on the IRIS specification was implemented using Semantic Web technologies. Two types of experiments were conducted: (1) experiments comparing the Whois and the prototype performance and (2) performance evaluation of the prototype based on load tests. Finally, the prototype implementation and subsequent experiment results serve as a proof-of-concept that Semantic Web technologies could contribute towards the IRIS protocol success.

Keywords: Semantic Web. IRIS Protocol. Whois Protocol.

LISTA DE FIGURAS

Figura 1– Diagrama dos blocos funcionais do protocolo IRIS	72
Figura 2– Diagrama entidade-relacionamento	97
Figura 3– Modelo RDF em formato de Grafo simplificado: apenas definição de classes.	97
Figura 4– Modelo RDF em formato de Grafo simplificado: definição de uma das propriedades da classe <i>swhois:Domínio</i>	98
Figura 5– Interação do protótipo com o usuário, os servidores Whois e outras aplicações	111

LISTA DE TABELAS

Tabela 1– Análise comparativa dos protocolos de acesso a informações de Registros da Internet através dos critérios funcionais básicos . . .	66
Tabela 2– Análise comparativa dos protocolos de acesso a informações de Registros da Internet através dos critérios específicos de registros de domínios	66
Tabela 2– Análise comparativa dos protocolos de acesso a informações de Registros da Internet através dos critérios específicos de registros de domínios	67
Tabela 3– Correspondência entre os nomes e suas respectivas URIs para a Figura 3	98
Tabela 4– Correspondência entre os nomes e suas respectivas URIs para a Figura 4	99
Tabela 5– Servidores Whois escolhidos para o protótipo	110
Tabela 6– Resultados dos Experimentos para Análise de Desempenho do protótipo <i>swhois</i>	115
Tabela 6– Resultados dos Experimentos para Análise de Desempenho do protótipo <i>swhois</i>	116
Tabela 7– Lista de todas propriedades da classe <i>swhois: dominio</i>	124
Tabela 7– Lista de todas propriedades da classe <i>swhois: dominio</i>	125
Tabela 8– Lista de todas propriedades da classe <i>swhois: Contato</i>	125
Tabela 9– Lista de todas propriedades da classe <i>swhois: ServidorDNS</i>	126

LISTA DE ABREVIATURAS

AfriNIC	<i>African Network Information Center</i>
APNIC	<i>Asia Pacific Network Information Service</i>
APCCIRN	<i>Asia Pacific Coordination Committee for Intercontinental Research Networks</i>
APNG	<i>Asia Pacific Network Group</i>
ARIN	<i>American Registry for Internet Numbers</i>
ARPA	<i>Advanced Research Projects Agency</i>
AS	<i>Autonomous System</i>
BEEP	<i>Blocks Extensible Exchange Protocol</i>
BGP	<i>Border Gateway Protocol</i>
ccTLD	<i>Country Code Top Level Domain</i>
CIDR	<i>Classless InterDomain Routing</i>
DNS	<i>Domain Name System</i>
DTD	<i>Document Type Definition</i>
DOM	<i>Document Object Model</i>
ENUM	<i>TElephone NUmber Mapping</i>
EPP	<i>Extensible Provisioning Protocol</i>
FIRS	<i>Federated Internet Registry Service</i>
FOAF	<i>Friend Of A Friend</i>
FTP	<i>File Transfer Protocol</i>
gTLD	<i>Generic Top Level Domain</i>
IANA	<i>Internet Assigned Numbers Authority</i>
ICANN	<i>Internet Corporation For Assigned Names and Numbers</i>
IDN	<i>Internationalized Domain Names</i>
IETF	<i>Internet Engineering Task Force</i>

IP	<i>Internet Protocol</i>
IR	<i>Internet Registry</i>
IRIS	<i>Internet Registry Information Service Protocol</i>
IRR	<i>Internet Routing Registry</i>
ISP	<i>Internet Service Provider</i>
CRISP	<i>Cross-Registry Information Service Protocol Working Group</i>
LACNIC	<i>Latin America and Caribbean Internet Addresses Registry</i>
LDAP	<i>Lightweight Directory Access Protocol</i>
LIR	<i>Local Internet Registry</i>
LWZ	<i>Lightweight UDP Transfer Protocol</i>
N3	<i>Notation 3</i>
NAPTR	<i>Naming Authority Pointer</i>
NCC	<i>Network Coordination Centre</i>
NIC	<i>Network Information Center</i>
NIR	<i>National Internet Registry</i>
NRO	<i>Number Resource Organization</i>
NSF	<i>Network Science Foundation</i>
NSI	<i>Network Solutions Inc.</i>
OIL	<i>Ontology Inference Layer</i>
OWL	<i>Web Ontology Language</i>
PIR	<i>Public Interest Registry</i>
RDBMS	<i>Relational Database Management System</i>
RDF	<i>Resource Description Framework</i>
RFC	<i>Request for Comments Internet Engineering Task Force Standard Specification</i>
RARE	<i>Reseaux Associes pour la Recherche Europeenne</i>

RIPE	<i>Reséaux IP Européens</i>
RIR	<i>Regional Internet Registry</i>
RPSL	<i>Routing Policy Specification Language</i>
RWHOIS	<i>Referral Whois</i>
SAX	<i>Simple API for XML</i>
S-NAPTR	<i>Straightforward-Naming Authority Pointer</i>
SRI	<i>Stanford Research Institute</i>
TCP	<i>Transmission Control Protocol</i>
TLS	<i>Transport Layer Security</i>
TriX	<i>Triples in XML</i>
UDP	<i>User Datagram Protocol</i>
URI	<i>Uniform Resource Identifier</i>
W3C	<i>World Wide Web Consortium</i>
XML	<i>Extensible Markup Language</i>
XPC	<i>XML Pipelining with Chunks</i>
XSL	<i>Extensible Stylesheet Language</i>

SUMÁRIO

1 INTRODUÇÃO	19
1.1 OBJETIVOS	19
1.2 MOTIVAÇÃO	20
1.3 JUSTIFICATIVA	20
1.4 TRABALHOS RELACIONADOS	21
1.4.1 VERISIGN	21
1.4.2 NIC.br	22
1.4.3 RIPE	22
1.5 EVOLUÇÃO	23
1.5.1 Protocolo Whois++ (whois++)	23
1.5.2 Referral Whois (rwhois)	23
1.5.3 Protocolo LDAP-Whois (ldap-whois)	24
1.5.4 Universal Whois (uwhois)	24
1.5.5 Joint Whois (jwhois)	24
1.5.6 Extensible Provisioning Protocol (EPP)	25
1.6 METODOLOGIA	25
1.6.1 Estudo bibliográfico	26
1.6.2 Análise dos Protocolos de Informação	28
1.6.3 Análise Científica	29
1.6.4 Implementação	29
1.7 ESCOPO	30
1.8 ORGANIZAÇÃO	31
2 CONCEITOS SOBRE REGISTROS DA INTERNET	33

2.1	BREVE HISTÓRIA DA INTERNET	33
2.2	EVOLUÇÃO DO SISTEMA DE REGISTROS	34
2.2.1	Primeira Etapa: Alocação de Endereços Centralizada	34
2.2.2	Segunda Etapa: Institucionalização	35
2.2.3	Terceira Etapa: Descentralização e Internacionalização	35
2.3	REGISTROS DA INTERNET	36
2.3.1	RIPE (Reséaux IP Européens)	36
2.3.2	APNIC (Asia Pacific Network Information Service)	37
2.3.3	ARIN (American Registry for Internet Numbers)	37
2.3.4	LACNIC (Latin America and Caribbean Internet Addresses Registry)	38
2.3.5	AfriNIC (African Network Information Center)	39
2.4	REGISTROS DE DOMÍNIOS	39
2.5	REGISTROS DE ROTEAMENTO	40
3	PROTOCOLOS PARA SERVIÇOS DE INFORMAÇÃO SOBRE REGISTROS DA INTERNET	41
3.1	COMUNIDADE VIRTUAL	41
3.1.1	Registros de Domínios	42
3.1.2	<i>Registrars</i>	42
3.1.3	Registros Regionais da Internet	43
3.1.4	Registros de Roteamento da Internet	43
3.1.5	Registros de Contato para a Coordenação de Incidentes	43
3.1.6	Desenvolvedores	44
3.1.7	Registrante	44
3.1.8	Provedores de Serviço e Operadores de Rede	44
3.1.9	Proprietário de Direitos Autorais	44
3.1.10	Poder Judiciário	45
3.1.11	Autoridade Certificadora	45
3.1.12	Usuários <i>Domain Name System</i> (DNS)	45

3.1.13	Outros atores	45
3.2	ESTADO DA ARTE DO PROTOCOLO WHOIS	46
3.2.1	História	46
3.2.2	Tipos de Informações	46
3.2.3	Usuários	47
3.2.4	Deficiências do Protocolo	47
3.3	PROTOCOLO JOINT WHOIS	49
3.3.1	Solução a curto prazo	49
3.3.2	Escopo	49
3.3.3	Transparente ao Usuário	49
3.3.4	Processo de consulta	50
3.4	PROTOCOLO IRIS	50
4	ANÁLISE COMPARATIVA DE PROTOCOLOS	54
4.1	CRITÉRIOS DE AVALIAÇÃO	54
4.1.1	Critérios Funcionais Básicos	54
4.1.2	Critérios Funcionais Específicos de Registros de Domínios	57
4.2	ANÁLISE COMPARATIVA	60
4.2.1	Critérios Funcionais Básicos	60
4.2.2	Critérios Funcionais Específicos de Registros de Domínios	63
4.3	RESULTADOS	65
4.4	DISCUSSÃO	67
4.5	CONCLUSÃO	68
5	ESTRUTURA DO PROTOCOLO IRIS	70
5.1	EVOLUÇÃO	71
5.2	ESTRUTURA DO PROTOCOLO	72
5.2.1	Transporte	74
5.2.2	Especificação dos Dados	75

5.3	SISTEMA DE DESCOBERTA DE SERVIÇOS	76
5.3.1	Resolução Direta	76
5.3.2	Resolução Bottom-Up	77
5.3.3	Resolução Top-Down	78
5.3.4	Mecanismo S-NAPTR: Configuração do Servidor DNS	79
5.3.5	Mecanismo S-NAPTR: Procedimento Cliente para Descoberta de Serviço	81
5.4	SOLUÇÃO DE CO-HABITAÇÃO COM O PROTOCOLO WHOIS	82
5.4.1	Caso de Uso S-NAPTR: Serviço de informação sobre domínios e IPs . .	83
5.4.2	Caso de Uso S-NAPTR: Delegar Serviço de Informação a Terceiros . . .	83
5.5	CONCLUSÕES	84
6	WEB SEMÂNTICA: XML VERSUS RDF	85
6.1	INTRODUÇÃO	85
6.2	ONTOLOGIA	87
6.3	XML	88
6.4	RDF	90
6.5	Representação de conhecimento	91
6.5.1	XML	92
6.5.2	RDF	93
6.6	Futuro da Web Semântica	94
7	INTEGRAÇÃO DE TECNOLOGIAS DE WEB SEMÂNTICA AO PROTOCOLO IRIS	95
7.1	INTRODUÇÃO	95
7.2	MODELO DE ONTOLOGIA	96
7.3	ESTUDO DE CASO	100
7.3.1	Escolher um servidor Whois	101
7.3.2	Consultar um domínio	101
7.3.3	Resposta mapeada em RDF sem extensão	103

7.3.4	Resposta mapeada em RDF com extensão	105
7.4	IMPLEMENTAÇÃO	108
7.4.1	Escolher quatro servidores whois	109
7.4.2	Coletar uma amostra de saída de cada um servidores	110
7.4.3	Mostrar como cada uma delas poderia ser mapeada em RDF	110
7.4.4	Implementação de um protótipo de servidor whois	111
7.5	Avaliação de Desempenho	112
7.5.1	Análise de Desempenho: Whois x SWhois	112
7.5.2	Análise de Desempenho: Teste de Carga swhois	114
7.6	Resultados Experimentais	114
7.7	Discussão	117
8	CONCLUSÃO	119
	Referências Bibliográficas	120
	Apêndice A – TABELAS DE PROPRIEDADES DAS CLASSES	124
	Apêndice B – MODELO RDF SERIALIZADO NA LINGUAGEM N3	127
	Apêndice C – CONSULTAS A SERVIDORES WHOIS UTILIZADAS DURANTE O DESENVOLVIMENTO DO PROTÓTIPO	136
C.1	REGISTRO BRASILEIRO DE NOMES DE DOMÍNIOS	136
C.2	REGISTRO IRLANDÊS DE NOMES DE DOMÍNIOS	137
C.3	REGISTRO DE NOMES DE DOMÍNIOS .ORG	138
C.4	REGISTRO DE NOMES DE DOMÍNIOS .COM	140

1 INTRODUÇÃO

Esta pesquisa tem por objetivo propor melhorias na especificação do protocolo *Internet Registry Information Service Protocol* (IRIS). IRIS é um novo protocolo em fase de desenvolvimento por um grupo de trabalho do *Internet Engineering Task Force* (IETF) com o objetivo de estabelecer um novo padrão para a busca de informações sobre recursos na Internet: nomes de domínios, endereços *Internet Protocol* (IP), sistemas autônomos, dentre outros. O grupo de trabalho tem a intenção de substituir o protocolo Whois, que é o atual padrão para a busca de informações sobre recursos da Internet. O atual protocolo Whois foi definido há muito tempo e por isso encontra-se em parte obsoleto. Por essa razão, o novo protocolo busca suprir as deficiências do protocolo Whois.

1.1 OBJETIVOS

O objetivo do trabalho consistiu em propor a utilização de tecnologias de Web Semântica no protocolo IRIS. A especificação atual do protocolo ainda não está finalizada e possui um mecanismo para troca de dados baseado em formato *Extensible Markup Language* (XML) para garantir interoperabilidade entre as aplicações. Entretanto o estudo de novas tecnologias de Web Semântica demonstrou que esta não é a melhor forma de garantir interoperabilidade na troca de dados entre aplicações na Internet. *Resource Description Framework* (RDF) é uma tecnologia chave para garantir interoperabilidade na Web Semântica e poderia contribuir para o sucesso do novo protocolo a curto, médio e longo prazo. Para demonstrar os benefícios da nova tecnologia foi desenvolvido um protótipo de serviço Whois para a Web Semântica. O protótipo serve como prova de conceito que a nova tecnologia é capaz de suprir todas as funcionalidades do serviço de informação Whois atual, facilitar a interoperabilidade na troca de dados entre aplicações na Internet e ainda possibilitar o desenvolvimento de aplicações capazes de processar e compartilhar dados de forma automatizada no futuro.

1.2 MOTIVAÇÃO

A motivação deste estudo consiste em demonstrar os benefícios que as tecnologias de Web Semântica podem proporcionar ao protocolo IRIS. A principal motivação para o desenvolvimento do protocolo IRIS foi a preocupação com a segurança dos dados armazenados na base de dados Whois. Por sua vez, o protocolo IRIS satisfaz todos requisitos de segurança necessários com base em mecanismos de autenticação e autorização. Além disso, o protocolo também propõe uma solução bastante eficiente para localizar o servidor autoritativo sobre determinado recurso, viabilizando o suporte a base de dados distribuída. Desta forma, o protocolo IRIS soluciona duas das principais deficiências do protocolo Whois. Entretanto, apesar da preocupação com as deficiências do protocolo atual, é preciso considerar o impacto do novo protocolo no futuro. Dada a grande barreira para a adoção de novas tecnologias e o investimento que o desenvolvimento de um novo padrão demanda, o resultado deve ser ótimo a curto, médio e longo prazo. No presente e, principalmente, no futuro as tecnologias de Web Semântica podem contribuir para o sucesso do novo protocolo.

1.3 JUSTIFICATIVA

Novas pesquisas apontam a tendência de crescimento da Web Semântica e a integração das tecnologias de Web Semântica ao protocolo IRIS poderia trazer benefícios a ambos. Se o novo protocolo for adotado com as novas tecnologias, pode facilitar a disseminação da Web Semântica. Por outro lado, o sucesso da Web Semântica pode levar o novo padrão a obsolescência, se ele não puder se adaptar a nova realidade. Além disso, a adoção das novas tecnologias poderia trazer benefícios para o protocolo IRIS no presente e no futuro. A curto prazo, os principais benefícios consistem em facilitar a manutenção e a interoperabilidade entre aplicações na Internet. A longo prazo, possibilitaria o desenvolvimento de agentes capazes de processar e compartilhar dados de forma automatizada.

1.4 TRABALHOS RELACIONADOS

O interesse da comunidade no desenvolvimento do novo protocolo vem sendo demonstrado pela participação nas listas de discussão, nas reuniões do grupo de trabalho durante os encontros do IETF e nas iniciativas de desenvolvimento de protótipos para o novo protocolo. Três instituições destacaram-se na implementação de protótipos e participação ativa nas discussões do grupo: Verisign, NIC.br e *Reséaux IP Européens* (RIPE).

1.4.1 VERISIGN

Atualmente, a Verisign é o maior registro de nomes de domínios. Por essa razão, a empresa possui interesse no desenvolvimento de novos padrões para substituir o protocolo Whois. O próprio autor do IRIS, Andrew Newton, trabalhou para a Verisign durante a maior parte do tempo dedicado ao desenvolvimento do protocolo IRIS. Ele foi o responsável pela criação de ferramentas de suporte para o desenvolvimento de aplicações IRIS, disponíveis para o público no *website* do IRIS mantido pela Verisign¹.

Dentre as ferramentas desenvolvidas destacam-se:

- Servidor IRIServ: Servidor baseado em *Relational Database Management System* (RDBMS);
- Servidor Flowerbed: Servidor baseado em banco de dados XML;
- Cliente Pimmit: Cliente escrito em Java;
- Cliente DREGquery: Cliente específico para a consultas a nomes de domínios;
- Biblioteca Net *Blocks Extensible Exchange Protocol* (BEEP) Lite: Biblioteca para o desenvolvimento de aplicações cliente-servidor IRIS usando protocolo de transporte BEEP;
- Biblioteca PYSNAPTR: Biblioteca para descoberta de serviços usando consulta *Domain Name System* (DNS) (*record Straightforward-Naming Authority Pointer* (S-NAPTR)) escrita em Python;

¹<http://iris.verisignlabs.com>

- Biblioteca JSNAPTR: Biblioteca para descoberta de serviços usando consulta DNS (record S-NAPTR) escrita em Java;
- VL BEEP Profiles: Biblioteca escrita para adicionar funcionalidades na biblioteca Beepcore-J escrita em Java para o protocolo de transporte BEEP. BEEP é um protocolo de transporte padrão do IETF e uma das alternativas de transporte do protocolo IRIS.

1.4.2 NIC.BR

O Registro.br é o departamento do NIC.br² responsável pelo registro brasileiro de nomes de domínios e provisionamento de endereços IP e sistemas autônomos no Brasil. O Registro.br participou ativamente do desenvolvimento e especificação do novo protocolo. Com destaque para a especificação da segunda versão do protocolo IRIS para nomes de domínios escrita por Andrew Newton em parceria com Frederico Neves, diretor de tecnologia do NIC.br. A nova versão do protocolo foi desenvolvida para suprir algumas das funcionalidades existentes no serviço Whois atual e que não poderiam ser oferecidas com a primeira versão do protocolo. Exemplo: informação sobre o status dos servidores de nomes para cada domínio, resultado da última consulta DNS efetuada durante um procedimento periódico de checagem DNS.

1.4.3 RIPE

O Registro Regional da Internet na Europa, conhecido como RIPE³, foi responsável pelo desenvolvimento de um protótipo do protocolo IRIS para provisionamento de informações sobre endereços IP e sistemas autônomos. Além da participação ativa nas discussões do grupo de trabalho, tanto nas listas de discussões, como nos encontros do IETF.

²<http://nic.br/>

³<http://ripe.net>

1.5 EVOLUÇÃO

Houve tentativas anteriores ao protocolo IRIS de criação de um novo protocolo padrão para substituir o protocolo Whois. De certa forma, essas tentativas falharam porque não foram adotadas pela comunidade e atualmente encontram-se em parte obsoletas em termos tecnológicos. O sucesso de um novo padrão para Internet depende de sua adoção e os fatores que determinam sua adoção são diversos e muitas vezes imprevisíveis. Entretanto, não se pode dizer que esses trabalhos tenham sido em vão, pelo contrário, sua pesquisa e desenvolvimento contribuíram para uma melhor compreensão do domínio do problema. Além disso, a participação da comunidade durante a fase de desenvolvimento do protocolo levando em conta os erros e acertos do passado pode contribuir muito para que ele seja o mais completo possível.

1.5.1 PROTOCOLO WHOIS++ (WHOIS++)

O protocolo Whois++ (DEUTSCH et al., 1995) foi definido como padrão da Internet mas se encontra obsoleto (REYNOLDS et al., 2001). Ele foi especificado como uma extensão do protocolo Whois para permitir o provisionamento de informações mais estruturadas. Foram incluídas funcionalidades para buscas mais avançadas com certo grau de internacionalização, com suporte a múltiplos idiomas e conjuntos de caracteres. O protocolo também inclui um mecanismo de autenticação opcional para proteger integralmente ou parcialmente as informações da base de dados Whois.

1.5.2 REFERRAL WHOIS (RWHOIS)

O protocolo *Referral Whois* (RWHOIS) (WILLIAMSON et al., 1997) foi definido em caráter informativo, ou seja, não tem o intuito de tornar-se um padrão da Internet. Esse protocolo busca principalmente desenvolver mecanismos para estender o protocolo Whois para que ele seja capaz de lidar com uma base de dados distribuída. O desenvolvimento deste protocolo teve grande influência da estrutura do protocolo DNS e conceitos de serviços de diretórios do protocolo X.500. O protocolo RWHOIS define a arquitetura da estrutura de diretórios e o protocolo de acesso a informações

armazenadas nesta estrutura.

1.5.3 PROTOCOLO LDAP-WHOIS (LDAP-WHOIS)

O protocolo LDAP-Whois (NEWTON, 2003) foi definido com objetivo de provisionar um sistema de informação sobre nomes de domínios baseado em *Lightweight Directory Access Protocol* (LDAP) (WAHL; HOWES; KILLE, 1997). O autor do LDAP-Whois, Andrew Newton, é o mesmo autor do IRIS. O protocolo não foi desenvolvido com o intuito de ser transformado em padrão da Internet. Na época, já havia uma proposta semelhante para endereços IP e sistemas autônomos (HOWARD, 1998). O projeto nunca deixou de ser experimental e foi publicado como *Request for Comments Internet Engineering Task Force Standard Specification* (RFC) na categoria experimental.

1.5.4 UNIVERSAL WHOIS (UWHOIS)

O Universal Whois (uwhois) foi um projeto do departamento de pesquisa da Verisign como resultado de um acordo entre a Verisign e a *Internet Corporation For Assigned Names and Numbers* (ICANN) estabelecido em 2001, que se fosse bem sucedido seria encaminhado ao IETF para padronização. Os autores do projeto eram: Andrew Newton (autor do IRIS), Leslie Daigle e Mark Kosters. O projeto acabou sendo interrompido e atualmente a documentação é praticamente inexistente, salvo algumas apresentações do grupo de pesquisa (NEWTON; DAIGLE; KOSTERS, 2001). Na primeira RFC do grupo de trabalho *Cross-Registry Information Service Protocol Working Group* (CRISP) (NEWTON, 2004), onde são definidos os requisitos do protocolo, o autor faz referência ao Universal Whois e agradecimentos aos colaboradores Leslie e Mark.

1.5.5 JOINT WHOIS (JWHOIS)

O Joint Whois (jwhois) representa uma solução a curto prazo para as deficiências do protocolo Whois mas não foi desenvolvido com o objetivo de tornar-se um novo pa-

drão da Internet. Na verdade, o projeto é bem simples, ele apenas define um proxy para o servidor Whois capaz de redirecionar a consulta ao servidor autoritativo sobre o recurso desejado. O projeto tem escopo limitado a consultas de endereços IP e sistemas autônomos. O Joint Whois encontra-se em produção desde de setembro de 2005 no *Latin America and Caribbean Internet Addresses Registry* (LACNIC), que é a organização responsável pelo provisionamento de endereços IP e sistemas autônomos na América Latina e Ilhas Caribenhas.

1.5.6 EXTENSIBLE PROVISIONING PROTOCOL (EPP)

O protocolo *Extensible Provisioning Protocol* (EPP) (HOLLENBECK, 2004) é um novo padrão do IETF para provisionamento de recursos da Internet. Ele merece atenção especial porque o protocolo IRIS sofreu grande influência de seus conceitos e tecnologias. Apesar do protocolo EPP possuir mecanismos de busca ele não é especializado em consultas, o foco principal do protocolo são as operações de transformação de recursos: inserção, remoção e atualização. Uma diferença fundamental entre os dois é que o IRIS foi desenvolvido para uma base de dados distribuída e o EPP para uma base de dados centralizada. Provavelmente por lidar com operações de transformação e não apenas consultas, o protocolo EPP preocupa-se mais com questões de segurança e por este motivo pode servir de exemplo para o IRIS com relação a medidas efetivas contra a varredura de dados.

1.6 METODOLOGIA

A metodologia utilizada para o desenvolvimento do trabalho consistiu das seguintes atividades: levantamento bibliográfico relevante aos assuntos abordados, estudo aprofundado dos temas, análise dos trabalhos relacionados, análise comparativa de protocolos para serviço de informação sobre registros da Internet, análise científica do protocolo IRIS, busca de novas tecnologias que pudessem agregar valor ao novo protocolo, em particular tecnologias de Web Semântica, e a implementação de um protótipo associada a experimentos de avaliação de desempenho.

1.6.1 ESTUDO BIBLIOGRÁFICO

Na etapa inicial foi realizado um estudo bibliográfico, que consistiu em grande parte na leitura de RFCs que especificam o protocolo IRIS, os protocolos relacionados ou que estão indiretamente ligadas ao tema. O estudo bibliográfico incluiu, principalmente, a leitura de RFCs na forma de *drafts* que especificam o protocolo IRIS, visto que o protocolo IRIS ainda está em fase de desenvolvimento e padronização.

As principais referências para o protocolo IRIS foram:

- (NEWTON, 2006b) – Artigo científico publicado pelo autor do protocolo IRIS – Andrew Newton – sobre a substituição do protocolo Whois pelo protocolo IRIS;
- (NEWTON, 2004) – RFC que especifica os requisitos do protocolo IRIS;
- (NEWTON; SANZ, 2005b) – RFC que especifica o funcionamento básico do protocolo IRIS que pode ser estendido para prover serviço de informação sobre qualquer registro da Internet;
- (NEWTON; SANZ, 2005a) – RFC que descreve a extensão do protocolo IRIS para nomes de domínios;
- (NEWTON; GUNDUZ; KERR, 2006) – RFC que descreve a extensão do protocolo IRIS para endereços IP e sistemas autônomos;
- (NEWTON; NEVES, 2006) – *Draft* de RFC que descreve a extensão do protocolo IRIS para nomes de domínios em uma segunda versão, se aprovada esta extensão substituiria a primeira;
- (NEWTON, 2006a) – RFC que descreve a extensão do protocolo IRIS para registros *TElephone NUmber Mapping* (ENUM);
- (NEWTON; SANZ, 2005c) – RFC que especifica o mecanismo de transporte BEEP aplicado ao protocolo IRIS;
- (NEWTON, 2007b) – *Draft* de RFC que especifica o mecanismo de transporte *XML Pipelining with Chunks* (XPC) do protocolo IRIS;
- (NEWTON, 2007a) – *Draft* de RFC que especifica o mecanismo de transporte *Lightweight UDP Transfer Protocol* (LWZ) do protocolo IRIS; e
- (NAGAHASHI; YOSHIDA; KONDO, 2005) – *Draft* de RFC que descreve a extensão do protocolo IRIS para informação de roteamento.

Para o estudo do protocolo Whois, as RFCs que especificam o protocolo foram a base do estudo (DAIGLE, 2004), (HARRENSTIEN; WHITE, 1982) e (HARRENSTIEN; STAHL; FEINLER, 1985). Para o protocolo Joint Whois, as fontes são as listas de discussão⁴ e encontros do IETF.

Na etapa seguinte foi feito um estudo bibliográfico dos trabalhos relacionados. Foram vários os trabalhos anteriores que buscaram uma solução para as deficiências do protocolo Whois. Entretanto, a obsolescência de muitos deles sem que tenham sido difundidos demonstra como pode ser árdua a tarefa de substituir um padrão da Internet amplamente difundido e bem sucedido apesar de suas deficiências. As principais referências para os trabalhos relacionados foram: (DEUTSCH et al., 1995), (WILLIAMSON et al., 1997), (HOWARD, 1998) e (NEWTON, 2003).

Sobre a questão da dificuldade de adoção do protocolo existe uma solução. A solução é a existência de uma fase planejada de transição em que as duas tecnologias coexistem durante um certo período de adaptação. O autor do protocolo já confrontou esta questão antes e o resultado é uma RFC em forma de *draft* (NEWTON; DAIGLE, 2003) que trata exclusivamente da fase de transição do protocolo Whois para o Protocolo IRIS.

Posteriormente, foi feito um estudo de novas tecnologias que pudessem ser incorporadas ao protocolo IRIS. Em particular, tecnologias relacionadas a Web Semântica que muito podem contribuir para o processamento automático de consultas e respostas principalmente para um protocolo especificado usando XML como é o caso do IRIS. Em termos de novas tecnologias XML e novos padrões do *World Wide Web Consortium* (W3C), a especificação do IRIS está atrasada em diversos aspectos.

O conceito de Web Semântica é um conceito novo fora do meio acadêmico, apesar de já ter algumas tecnologias padronizadas pelo W3C, como a *framework* RDF. Algumas grandes empresas participaram do processo de padronização da *framework* RDF pelo W3C e já possuem aplicações em linguagem semântica em produção, como é o caso da Xerox, da Adobe, dentre outras. Além disso, as recomendações do W3C para Web Semântica, principalmente os documentos de especificação da *framework* RDF contribuíram bastante para o trabalho.

A maior parte das referências relativas a Web Semântica são provenientes de artigos científicos, ao contrário das referências ao protocolo IRIS e aos trabalhos relacionados que em sua maioria eram RFC ou *drafts* de RFCs. Este fato não é mero acaso, a comunidade científica sempre esteve envolvida no desenvolvimento de novos padrões para a Web Semântica. Muitos acreditavam que se tratava de uma utopia mas com as recentes padronizações pelo W3C e adoção de linguagens semânticas por grandes

⁴<https://www1.ietf.org/mailman/listinfo/crisp>

empresas esta realidade está mudando.

As principais referências sobre Web Semântica foram:

- (BERNERS-LEE; HENDLER; LASSILA, 2001) – Um dos primeiros artigos publicados sobre a Web Semântica, o autor deste artigo é considerado o pai da Web Semântica (Tim Berners-Lee);
- (BERNERS-LEE, 1999) – Livro sobre a Web Semântica escrito por Tim Berners-Lee;
- (DECKER; MITRA; MELNIK, 2000) – Artigo em forma de tutorial sobre a *framework* RDF;
- (BROEKSTRA et al., 2001) – Artigo sobre como estender a *framework* RDF;
- (HOLLENBECK; ROSE; MASINTER, 2003) – RFC que especifica como utilizar XML em protocolos do IETF;
- (DECKER et al., 2000) – Artigo sobre as vantagens do uso de RDF e as desvantagens de XML;
- (HORROCKS; PATEL-SCHNEIDER, 2003) – Artigo sobre representação de dados na Web Semântica;

Outras fontes de referência para a Web Semântica correspondem a recomendações do W3C para a Web Semântica, como a descrição de RDF ⁵.

1.6.2 ANÁLISE DOS PROTOCOLOS DE INFORMAÇÃO

Os protocolos de informação relacionados são muitos e uma análise que buscasse relacionar todos eles em um único estudo comparativo poderia ser mais confuso do que esclarecedor. Por essa razão foi realizada uma pesquisa sobre o estado da arte em termos de protocolos de busca de informações sobre recursos da Internet. Como resultado desta pesquisa foram selecionados três protocolos para uma análise comparativa.

A análise consistiu na comparação dos protocolos Whois, Joint Whois e IRIS, levando em consideração as peculiaridades de cada protocolo. O protocolo Whois foi escolhido

⁵<http://www.w3.org/TR/rdf-primer/>

por ser o atual padrão *de facto*, sendo um padrão do IETF amplamente difundido e presente na maioria dos sistemas operacionais. O Joint Whois foi escolhido por estar em produção apesar de seu escopo limitado, não sendo um protocolo padrão mas um projeto resultante da iniciativa da *Number Resource Organization* (NRO), proposto como uma solução a curto prazo para as deficiências do protocolo Whois e sem o intuito de tornar-se um padrão no futuro. Finalmente, o IRIS por ser a promessa de solução a longo prazo para as deficiências do protocolo Whois, sendo um protocolo em fase de desenvolvimento e padronização pelo IETF.

Para esta análise foram definidas métricas e critérios de avaliação para que fosse possível a análise comparativa dos protocolos. Os critérios foram baseados nas características presentes nos protocolos atuais e características desejáveis para os futuros protocolos. Entretanto o estudo dos trabalhos relacionados não se limitou a análise comparativa entre os protocolos Whois, Joint Whois e IRIS. Pelo contrário, todos os trabalhos citados foram levados em consideração ao longo do estudo possibilitando a identificação de possíveis falhas e viabilizando propostas de melhorias.

1.6.3 ANÁLISE CIENTÍFICA

Posteriormente foi feito um estudo aprofundado do protocolo IRIS, suas vantagens e desvantagens, além da busca de novas tecnologias que pudessem agregar valor ao novo protocolo. Nesta fase mais avançada do estudo foi possível detectar pontos na especificação do protocolo que poderiam ser otimizados se levassem novas tecnologias em consideração. Uma das novas tecnologias analisadas é a *framework* RDF que poderia contribuir para viabilizar processamento automático de consultas e respostas. Além disso, RDF garante a interoperabilidade entre aplicações da Internet possibilitando o compartilhamento de dados mais eficiente. Um dos objetivos do estudo é demonstrar como a *framework* RDF poderia ser aplicada ao protocolo IRIS e quais benefícios poderia trazer ao novo protocolo.

1.6.4 IMPLEMENTAÇÃO

Este estudo inclui a implementação de um protótipo como prova de conceito de que a *framework* RDF pode ser aplicada ao protocolo IRIS. Foi possível mostrar que não

existe necessidade de um grande custo de implementação, pois já existem bibliotecas prontas, de código aberto e livre, para dar suporte a RDF. Além disso, demonstrou-se que todas as funcionalidades do sistema de informação Whois atual podem ser provisionadas utilizando a *framework* RDF, ou seja, não haverá perda de funcionalidade. Adicionalmente o protocolo poderia contar com todas as facilidades de RDF para processamento automático e compartilhamento de dados entre aplicações na Internet. A implementação foi feita na linguagem de programação Java 1.5 usando o sistema operacional Kubuntu 6.06 LTS (Dapper Drake), a IDE escolhido foi Eclipse, *subversion* para controle de versão, em outras palavras, apenas *software* livre. O protótipo teve um escopo limitado a consultas nomes de domínios, embora possa ser estendido para suportar outros tipos de recursos, como endereços IP e sistemas autônomos. A implementação envolveu o desenvolvimento de uma aplicação cliente capaz de transformar o resultado da consulta de domínios em quatro servidores Whois de formato texto para RDF, utilizando um modelo RDF baseado na especificação do protocolo IRIS. Ao final do estudo sobre a implementação foram realizados testes de comparação do desempenho do protótipo e do cliente Whois, que demonstraram que o desempenho é muito semelhante, ou seja, o protótipo é uma prova de conceito de que é possível introduzir tecnologias de Web Semântica sem perda de desempenho. A curto prazo, o benefício seria facilitar o desenvolvimento e a manutenção das aplicações. No futuro, viabilizaria o desenvolvimento de aplicações capazes de compartilhar e processar a informação de maneira automatizada.

1.7 ESCOPO

O escopo deste estudo está restrito a um único tipo de operação sobre recursos da Internet. O único tipo de operação é consulta de informações sobre recursos. As operações de atualização, remoção ou inserção de novos registros estão fora do escopo deste estudo.

Na verdade, este escopo corresponde ao escopo do próprio protocolo IRIS pois as funções de provisionamento de recursos, como são chamadas as operações de inserção, atualização ou remoção de recursos, são tratadas por outro protocolo recentemente padronizado pelo IETF, o protocolo EPP. Recentemente, o protocolo EPP foi implementado e se encontra em produção no sistema brasileiro de registro de nomes de domínios: o Registro.br.

Além da restrição relativa ao tipo de operação, o estudo tem escopo limitado a busca

de informação sobre alguns tipos específicos de recursos da Internet. O escopo do estudo está restrito a nomes de domínios. Entretanto, o protocolo IRIS potencialmente pode ser usado para prover um sistema de informação sobre qualquer tipo de recurso, por exemplo endereços IP, sistemas autônomos e informações de roteamento.

1.8 ORGANIZAÇÃO

O documento foi organizado em nove partes principais:

- Introdução;
- Conceitos sobre registros da Internet;
- Estado da arte em protocolos de informação sobre recursos da Internet;
- Análise comparativa de protocolos;
- Web Semântica: XML versus RDF;
- Tecnologias de Web Semântica aplicadas ao Protocolo IRIS;
- Discussão;
- Conclusão.

Na introdução são definidos os objetivos, motivação, justificativa, trabalhos relacionados, metodologia, escopo e organização do trabalho. Na segunda parte, intitulada Conceitos sobre registros da Internet, são apresentados separadamente um breve histórico da Internet, dos registros da Internet, dos registros de domínios e dos registros de roteamento da Internet. Na terceira parte, é apresentado o estado da arte em termos de protocolos de acesso a informações de registros de recursos da Internet: a comunidade virtual envolvida com os protocolos, o protocolo Whois, o protocolo Joint Whois e o protocolo IRIS. Na análise comparativa, correspondente a quarta parte do trabalho, são estabelecidos os critérios de comparação entre os protocolos, seguidos pela comparação dos três protocolos de acesso descritos na terceira parte, os resultados da comparação são apresentados na forma de tabelas para melhor visualização, em seguida são discutidos os resultados e é feita a conclusão da análise comparativa dos protocolos. Na quinta parte do estudo será realizada uma análise científica detalhada do protocolo IRIS envolvendo sua estrutura, protocolos de transporte (LWZ

e XPC), objetos de consultas (nome de domínios, contatos, IP e *Autonomous System* (AS)), considerações sobre uma solução de co-habitação com o protocolo atual. Em seguida, será feito um estudo sobre tecnologias de Web Semântica, principalmente, as vantagens de RDF em relação a XML. Baseado no estudo das tecnologias de Web Semântica, em especial RDF, foi feito um estudo de caso da aplicação da *framework* RDF para o serviço Whois do registro brasileiro de nomes de domínios e a implementação de um protótipo capaz de mapear o resultado de uma consulta a nomes de domínios de formato texto para RDF. Ao final do estudo sobre a implementação foram realizados testes de comparação de desempenho do protótipo e de um cliente Whois, que demonstraram que o desempenho é muito semelhante, ou seja, o protótipo é uma prova de conceito de que é possível introduzir tecnologias de Web Semântica sem afetar o desempenho. Além disso traria, benefícios para o desenvolvimento e a manutenção das aplicações, e viabilizaria o processamento automatizado de dados. Finalmente, será feita uma discussão sobre a pesquisa e as soluções propostas, seguida de uma conclusão do trabalho.

2 CONCEITOS SOBRE REGISTROS DA INTERNET

Esta seção tem o intuito de fornecer os conceitos básicos para a compreensão do contexto em que se insere os mecanismos de busca de informações sobre recursos da Internet.

O sistema responsável pelo registro de recursos da Internet também é responsável pelo acesso às informações sobre os mesmos. Este sistema evoluiu a partir de um sistema extremamente centralizado para um sistema distribuído e globalizado. O desafio atual é fornecer um sistema único de informação apesar da informação permanecer distribuída.

O sistema de registro e informação sobre recursos da Internet se desenvolveu paralelamente a história da própria Internet. Este sistema influenciou e foi influenciado pelo desenvolvimento da rede mundial. Por esta razão é importante conhecer um pouco da história da Internet e da evolução do sistema de registro para a melhor compreensão do estado da arte apresentado na próxima seção.

2.1 BREVE HISTÓRIA DA INTERNET

No início da década de 60, a *Advanced Research Projects Agency* (ARPA) iniciou a implantação de uma rede experimental de computadores de longa distância que conectava importantes centros de pesquisa norte-americanos, a ARPAnet. O objetivo original da ARPAnet era permitir que os centros de pesquisa pudessem compartilhar os recursos computacionais escassos e caros na época.

No início da década de 80, o protocolo *Transmission Control Protocol* (TCP) foi desenvolvido e rapidamente se tornou o protocolo padrão da ARPAnet. Mais tarde, com a inclusão do protocolo no sistema operacional desenvolvido pela Universidade de Berkeley, o BSD Unix, o protocolo TCP se popularizou. A rede assim formada cresceu a partir de escassos computadores para mais de dez mil hosts, formando um backbone de redes locais e regionais, constituindo a base para a Internet atual.

Em 1988, a ARPA decidiu que o experimento havia terminado e deu início ao seu

desmembramento. Ao mesmo tempo, outra rede denominada NSFNET, fundada pela *Network Science Foundation* (NSF), começou a substituir a ARPAnet e constituir o novo backbone da Internet. Mais recentemente, em 1995, a Internet sofreu uma nova transição, quando deixou de ser uma rede pública para ser um backbone constituído de múltiplos backbones comerciais.

Atualmente, a Internet conecta milhões de hosts ao redor do mundo. Alguns dos novos backbones comerciais conseguem carregar um volume de dados de centenas de megabits, algumas dezenas de vezes a largura de banda da ARPAnet original. Mais de dez milhões de pessoas utilizam a rede para comunicação e colaboração diariamente.

2.2 EVOLUÇÃO DO SISTEMA DE REGISTROS

A responsabilidade de alocação de endereços IP e sistemas autônomos é uma função da *Internet Assigned Numbers Authority* (IANA) sob a supervisão do IETF. A IANA delega a função de alocação destes recursos a entidades regionais conhecidas como Registros Regionais da Internet ou *Regional Internet Registry* (RIR). Entretanto para compreender melhor a atual estrutura de delegação de endereços e sistemas autônomos é preciso acompanhar a evolução deste sistema.

A evolução dos Registros da Internet foi dividida em três partes: inicialmente a função de alocação de endereços era extremamente centralizada e informal; na segunda etapa esta função torna-se mais institucionalizada; e na terceira etapa ocorre a internacionalização do processo de alocação de endereços e sistemas autônomos. Os estudos sobre a evolução dos registros da Internet podem ser encontrados no artigo *Development of the Regional Internet Registry System* em (KARRENBERG et al., 2001).

2.2.1 PRIMEIRA ETAPA: ALOCAÇÃO DE ENDEREÇOS CENTRALIZADA

Desde o início da Internet, a tarefa de designar endereços era necessária para evitar que duas ou mais redes tentassem utilizar o mesmo endereço de rede na Internet. Essa era uma tarefa elementar, que consistia em manter uma lista dos endereços de rede designados e era exercida de forma voluntária por Jon Postel, que de acordo com

a lenda usava um simples bloco de papel para anotar essas informações.

2.2.2 SEGUNDA ETAPA: INSTITUCIONALIZAÇÃO

No início da década de 80, com crescimento acelerado da Internet a tarefa de manter a lista de endereços de rede designados crescia vertiginosamente. Para solucionar esta questão, a IANA foi criada e junto com ela o Registro da Internet ou *Internet Registry* (IR). Com a criação do IR, a função do bloco de notas de Jon Postel foi transferida através de um contrato da NSF para um centro de informações localizado na Califórnia denominado *Network Information Center* (NIC).

Durante todo este tempo, as redes eram alocadas de forma liberal a qualquer organização que satisfizesse requisitos mínimos. Entretanto, durante o final da década de 80, o crescimento da Internet gerou dois problemas: a rápida alocação do espaço de endereçamento disponível devido a política de alocação não se preocupar com a conservação do espaço de endereçamento e o conseqüente crescimento descontrolado da tabela de roteamento da Internet.

A falta de uma política de alocação de endereços mais rígida, com garantia de agregação e conservação do espaço de endereçamento da Internet, associada ao crescimento acelerado da Internet forçou as autoridades a desenvolver a forma de endereçamento IP, para tentar solucionar os problemas encontrados. Foram criados: o conceito de subredes, classes de endereçamento (A, B e C) e finalmente o *Classless InterDomain Routing* (CIDR). Mais recentemente, foi desenvolvido o protocolo IPv6 cuja principal motivação foi a solução destes problemas.

2.2.3 TERCEIRA ETAPA: DESCENTRALIZAÇÃO E INTERNACIONALIZAÇÃO

Com o passar do tempo surgiram duas necessidades, a primeira de caráter técnico era a necessidade de designação topológica do espaço de endereçamento. A segunda de caráter administrativo era a necessidade de modernização dos mecanismos administrativos inerentes ao gerenciamento do espaço de endereçamento.

O sistema centralizado vigente até então demonstrava sinais de obsolescência diante dos novos desafios: grande volume de informações, distância entre a administração

e os usuários do espaço de endereçamento, falta de uma estrutura global, falta de participação das comunidades locais nas políticas de delegação de recursos.

A solução encontrada foi a criação do conceito de RIR que consiste em uma autoridade responsável pela distribuição de endereços IP e sistemas autônomos em uma determinada região geográfica com proporções continentais. Sendo assim, por volta de 1990 começaram a aparecer os primeiros sinais de preocupação da comunidade internacional com a necessidade de descentralização e internacionalização da administração do espaço de endereçamento de Internet. Como resultado foram publicadas duas RFCs: RFC1174 (CERF, 1990) e RFC1366 (GERICH, 1992) – descrevendo as políticas para criação e gerenciamento dos RIRs.

2.3 REGISTROS DA INTERNET

A responsabilidade de alocação de endereços IP e sistemas autônomos é uma função de organizações denominadas Registros Regionais da Internet. Estes registros foram criados gradualmente a medida em que a necessidade de controle regional do espaço de endereçamento crescia. A seguir cada um dos Registros Regionais são apresentados na ordem em que foram criados.

2.3.1 RIPE (RESÉAUX IP EUROPÉENS)

O RIR europeu foi o primeiro a ser criado em 1990, quando os operadores da rede IP RIPE perceberam a necessidade de gerenciar as funções de registro de recursos da Internet em sua região. A RFC 1174 foi criada logo em seguida definindo o funcionamento do RIPE *Network Coordination Centre* (NCC). Entretanto foi somente dois anos após a publicação da RFC que o primeiro RIR se tornou efetivamente operacional.

O RIPE surgiu como uma parte independente da rede de pesquisa européia *Reseaux Associes pour la Recherche Européenne* (RARE), mas recebeu contribuições tanto da rede de pesquisa quanto de redes comerciais emergentes. Atualmente, ele está localizado em Amsterdã nos Países Baixos e sua comunidade envolve 109 países da Europa, Oriente Médio, Ásia central e países do norte da África acima da linha do Equador.

Mais informações sobre o RIPE NCC, podem ser encontradas em seu site institucional¹.

2.3.2 APNIC (ASIA PACIFIC NETWORK INFORMATION SERVICE)

O segundo RIR foi criado em 1993, o *Asia Pacific Network Information Service* (APNIC) localizado em Tóquio no Japão. Na verdade, o APNIC começou como um projeto piloto da rede de pesquisa *Asia Pacific Coordination Committee for Intercontinental Research Networks* (APCCIRN), atual *Asia Pacific Network Group* (APNG). O RIR recém criado permaneceu em fase de experimental durante 10 meses até ser oficialmente reconhecido como RIR.

A sede do APNIC foi transferida em 1998 para Brisbane na Austrália. O APNIC possui atualmente cerca de 700 membros de 39 economias diferentes, da Ásia Central e Meridional até a Oceania. Dentre os membros do APNIC estão os cinco *National Internet Registry* (NIR) da região: do Japão, da China, Taiwan, Coréia e Indonésia. Os NIRs executam funções semelhantes ao APNIC em nível nacional e seus membros compõem grande parte do APNIC.

Mais informações sobre o APNIC, podem ser encontradas em seu site institucional².

2.3.3 ARIN (AMERICAN REGISTRY FOR INTERNET NUMBERS)

Pelo fato das atividades de gerenciamento do espaço de endereçamento estarem inicialmente centralizadas nas redes de pesquisa norte-americanas, demorou um certo tempo para as autoridades gradualmente perceberem a necessidade de criação de um RIR para o continente americano, conhecido como *American Registry for Internet Numbers* (ARIN).

Em 1991, a função de Registro de Internet foi transferida do NIC para a Network Solutions Inc, cujas funções envolviam além da distribuição de endereços IP e sistemas autônomos: o registro e suporte de domínios, registro de usuários, serviços online e serviços de atendimento ao público. Porém apesar da mudança o RIR americano

¹<http://ripe.net/>

²<http://apnic.net/>

ainda não havia sido criado.

Em pouco tempo, o governo dos Estados Unidos decidiu separar o suporte de rede para a Internet Comercial do Departamento de Defesa dos Estados Unidos devido ao crescimento acelerado da Internet neste período. Assim, foi criado em 1993 um projeto chamado InterNIC, resultado de uma parceria da *Network Solutions Inc.* (NSI) e NSF. O projeto acumulava as seguintes funções: registro de domínios e distribuição de endereços IP.

Mais tarde, a partir de um acordo do IETF, IANA, RIPE NCC, APNIC e NSF, foi decidido que seria melhor separar o registro de domínios da distribuição de endereços IP, que é uma atividade que exige grande rigor no intuito de conservar o espaço de endereçamento IP. Como resultado deste acordo foi criado o ARIN em 1997, como uma organização sem fins lucrativos, independente e aberta, onde as entidades e indivíduos interessados podem ser membros e contribuir.

O ARIN se localiza em Chantilly no estado da Virgínia nos Estados Unidos. Na época de sua criação, ARIN envolvia os países da América do Norte, América do Sul e as ilhas caribenhas e os países africanos abaixo da linha do Equador. Atualmente, o ARIN presta seus serviços aos países da América do Norte e países africanos abaixo da linha do Equador pois os países da América Latina e ilhas caribenhas são gerenciadas pelo recém criado LACNIC.

Mais informações sobre o ARIN podem ser encontradas em seu site institucional³.

2.3.4 LACNIC (LATIN AMERICA AND CARIBBEAN INTERNET ADDRESSES REGISTRY)

O quarto e mais recente Registro Regional de Internet é o LACNIC, o RIR para América Latina e Caribe, reconhecido oficialmente em outubro de 2002. Contudo, já havia sido aprovado em caráter provisório em março do mesmo ano. O LACNIC é uma organização em fins lucrativos, baseada em sociedade e estabelecida juridicamente no Uruguai, sua função consiste na administração do espaço de endereçamento IP e números de sistemas autônomos para a região da América Latina e Caribe em nome da comunidade Internet.

Mais informações sobre o LACNIC podem ser encontradas em seu site institucional⁴.

³<http://arin.net/>

⁴<http://lacnic.net/>

2.3.5 AFRINIC (AFRICAN NETWORK INFORMATION CENTER)

O mais novo RIR é o *African Network Information Center* (AfrinIC) que assumiu a responsabilidade sobre a alocação de endereços IP e sistemas autônomos para o continente africano. Mais informações sobre o AfrinIC podem ser encontradas em seu site institucional.

Os RIRs atuais servem países fora de sua região principal para prover cobertura global. Sendo assim, é esperado que novos RIRs devam emergir num futuro próximo para suprir essas necessidades regionais. Porém é esperado um número pequeno de RIRs emergentes porque por definição um RIR deve ter uma área de abrangência de dimensões continentais.

Mais informações sobre o AfrinIC podem ser encontradas em seu site institucional⁵.

2.4 REGISTROS DE DOMÍNIOS

Antes do surgimento dos Registros de Domínios da Internet o registro de nomes era feito através da atualização e distribuição de um arquivo conhecido como HOST.TXT. Durante toda a década de 70, a ARPAnet foi uma pequena comunidade de algumas centenas de hosts e um único arquivo, HOSTS.TXT, continha toda a informação necessária sobre estes hosts. Para cada host conectado à ARPAnet ele continha um mapeamento do nome do host em seu endereço IP. A tabela de hosts UNIX foi compilada a partir do arquivo HOSTS.TXT.

O arquivo HOSTS.TXT original era mantido pelo NIC na Califórnia. Os administradores da ARPAnet tipicamente enviavam qualquer mudança ao NIC por email e, periodicamente, buscavam atualizar as suas cópias do arquivo HOSTS.TXT via o protocolo *File Transfer Protocol* (FTP). As mudanças no arquivo de hosts eram feitas semanalmente mas a medida que a ARPAnet cresceu este esquema se tornou inviável.

O problema essencial do esquema baseado no arquivo HOST.TXT é o fato de não ser escalável. Os membros da ARPAnet iniciaram uma investigação de um sucessor do mecanismo HOSTS.TXT. O objetivo da investigação era criar um novo sistema que permitisse administração local de dados e ainda fosse capaz de disponibilizar os dados de maneira global. A descentralização da administração eliminaria o gargalo do

⁵<http://afrinic.net/>

sistema e aliviaria o problema de tráfego.

Paul Mockapetris foi responsável por desenvolver a arquitetura do novo sistema. Em 1984, ele publicou as RFCs 882 (MOCKAPETRIS, 1983a) e 883 (MOCKAPETRIS, 1983b), que descreviam o DNS. Estas RFCs foram substituídas pelas RFCs 1034 (MOCKAPETRIS, 1987a) e 1035 (MOCKAPETRIS, 1987b), que contêm as especificações atuais para o sistema DNS, definindo um espaço de nomes de domínios hierárquico com garantia de unicidade dos hosts.

Depois da definição do sistema de nomes de domínios, teve início a formação dos registros de domínios da Internet. Ao contrário dos Registros Regionais de Internet, os Registros de Domínios são numerosos, havendo potencialmente um para cada país do mundo, devido a definição da extensão *Country Code Top Level Domain* (ccTLD)s, por exemplo: .us, .br e .uk. Além dos Registros de Domínios de cada país, existem registros responsáveis por nomes de domínios gerais ou *Generic Top Level Domain* (gTLD), exemplo: .com, .org e .edu.

2.5 REGISTROS DE ROTEAMENTO

Atualmente, o único serviço que os Registros de Roteamento da Internet ou *Internet Routing Registry* (IRR) provêem é um mecanismo de validação do conteúdo de uma sessão *Border Gateway Protocol* (BGP) através do protocolo *Routing Policy Specification Language* (RPSL) (ALAETTINOGLU et al., 1999) ou o mapeamento de um sistema autônomo em um conjunto de redes. Entretanto, o protocolo BGP funciona mesmo sem utilizar os serviços destes registros razão pela qual em muitos casos seus serviços não são efetivamente utilizados.

Os IRRs são a próxima geração de base de dados em desenvolvimento envolvendo participantes de muitas organizações internacionais. Os dados da base de dados dos Registros de Roteamento da Internet podem ser usadas por qualquer pessoa ao redor do mundo para corrigir erros, configurar e gerenciar roteamento e endereçamento na Internet.

Os IRRs surgiram no início de 1995, um tempo em que os provedores de serviço estavam se preparando para o fim do serviço do backbone da NFSNET e o nascimento da Internet comercial. Originalmente, haviam cinco bases de dados, incluindo a operada pela Merit, o RIPE NCC, a ANS(atual UUNet), internetMCI(atual Cable & Wireless) e a Bell Canada. Atualmente, existem mais de vinte bancos de dados de Registros de Roteamento da Internet.

3 PROTOCOLOS PARA SERVIÇOS DE INFORMAÇÃO SOBRE REGISTROS DA INTERNET

Com o crescimento da Internet a função de registro passou de um centro de informações tradicionalmente centralizado para se tornar responsabilidade de vários registros de Internet. Por esta razão, houve a necessidade de criação de uma autoridade para estabelecer políticas de gerenciamento da Internet que fossem comuns no mundo todo. A autoridade estabelecida para esta função é a ICANN.

A ICANN é responsável pelo gerenciamento e coordenação do sistema DNS e para assegurar que cada endereço é único e que todos os usuários da Internet possam encontrar todos os endereços válidos. Ela garante isto ao supervisionar a distribuição única de endereços IP e nomes de domínios, também garante que cada nome de domínio seja mapeado em um endereço válido.

Dentre as suas funções a ICANN gerencia as atividades de várias organizações que juntas compõem o sistema de registros de recursos da Internet e o acesso a informações sobre estes recursos. Sua estrutura organizacional envolve uma hierarquia com várias outras autoridades, por exemplo a IANA.

3.1 COMUNIDADE VIRTUAL

Esta seção tem o intuito de apresentar a comunidade da Internet beneficiada pelos protocolos de acesso a informações de Registros de Internet. Assim como demonstrar a importância do acesso a estas informações para a comunidade mundial. Os membros desta comunidade são:

- Registros de Domínios
- Registrars
- Registros Regionais da Internet
- Registros de Roteamento da Internet
- Registros de Contato para a Coordenação de Incidentes

- Desenvolvedores
- Registrante
- Provedores de Serviço e Operadores de Rede
- Proprietário de Direitos Autorais
- Poder Judiciário
- Autoridades Certificadoras
- Usuários DNS
- Outros atores

A seguir cada uma categoria de membro da comunidade virtual é descrito em detalhes.

3.1.1 REGISTROS DE DOMÍNIOS

Os Registros de Domínios são responsáveis pelo processo de registro de domínio para uso com o DNS e resolução de nomes de hosts em endereços IP. Um registro de domínio gTLD ou ccTLD pode ser uma entidade governamental, não-governamental, não comercial, ou uma entidade comercial. Registros de domínios seguem dois tipos de modelo para o processo de registro: modelo *thick* e *thin*. O modelo *thick* é mais tradicional, neste modelo o Registro contém ambos: os dados técnicos (nomes e endereços IP dos servidores de nomes relacionados ao domínio) e os dados de contatos, informações sobre os contatos técnico, cobrança, administrativo ou contatos adicionais do domínio (nome, email, endereço, telefone e etc). No modelo *thin*, o usuário detentor de um domínio ou que deseja possuir um domínio não entra em contato direto com o Registro, ele contacta uma entidade intermediária denominada *Registrar*, neste modelo as informações técnicas sobre o domínio ficam armazenadas no Registro enquanto as informações sobre contatos ficam armazenadas no *Registrar*.

3.1.2 REGISTRARS

Os *registrars* aceitam representar o usuário final na função de registrar um domínio junto ao Registro (ou *Registry*), em ambos os casos no modelo *thin* e *thick*. No mo-

delo *thin* o *Registrar* passa somente as informações técnicas para o Registro mas no modelo *thick* passa tanto as informações técnicas quanto as de contato.

3.1.3 REGISTROS REGIONAIS DA INTERNET

Os Registros Regionais da Internet, conhecidos como RIRs, administram a alocação do espaço de endereçamento IP e número de sistemas autônomos. Cada RIR serve uma região geográfica específica de dimensões continentais e juntos eles servem a Internet inteira. Em alguns casos existem *Local Internet Registry* (LIR) ou NIR que coordenam as mesmas funções dos RIRs para regiões geográficas menores e mais específicas. Exemplo de NIR: Registro.br

3.1.4 REGISTROS DE ROTEAMENTO DA INTERNET

Os Registros de Roteamento da Internet, conhecidos como IRRs (Internet Rounting Registries), são as bases de dados de políticas de roteamento. O propósito é fornecer informação útil na administração de roteadores da Internet. A sintaxe e o conteúdo são definidos pelo protocolo RPSL (ALAETTINOGLU et al., 1999).

Os IRRs são operados por organizações acadêmicas, governamentais dentre outras, incluindo os próprios RIRs. Ao contrário dos dados normalmente armazenados nos RIRs(endereço IP e AS), os dados dos IRRs são frequentemente duplicados em diferentes organizações.

3.1.5 REGISTROS DE CONTATO PARA A COORDENAÇÃO DE INCIDENTES

Os Registros de Coordenação e Comunicação de Incidentes fornecem um meio para que operadores de rede possam notificar a comunidade virtual em caso de incidente. Usando este tipo de registro, os operadores de rede são providos com informação para contactar outro operador de rede caso suspeite que algum incidente possa estar ocorrendo.

3.1.6 DESENVOLVEDORES

Desenvolvedores de software cliente são frequentemente afiliados a grandes operadores de rede, a Registros da Internet ou entidades comerciais, sendo muitos destes software de código aberto e gratuito. Já os desenvolvedores de software servidor são frequentemente afiliados a entidades comerciais especializados no desenvolvimento de software para Registros da Internet.

3.1.7 REGISTRANTE

Entidades às quais foi dada autoridade sobre um recurso da Internet via compra, aluguel, ou permissão de uso de um registro da Internet, diretamente por um Registro da Internet ou via um Registrar.

3.1.8 PROVEDORES DE SERVIÇO E OPERADORES DE REDE

Os provedores de serviço e operadores de rede interagem com todos os tipos de Registro da Internet citados neste documento em uma base frequente e consistente. Por exemplo, operadores de rede utilizam os serviços de diretório dos registros da Internet para obter informação de contato de recursos que estejam com problemas técnicos.

3.1.9 PROPRIETÁRIO DE DIREITOS AUTORAIS

Pessoas físicas ou jurídicas que possuam direitos autorais sobre uma sequência de caracteres alfanúmericos acessam as informações dos Registros de Domínios para propósitos de manutenção e defesa dos direitos sobre nomes de domínios consistentes com as leis e regras aplicáveis.

3.1.10 PODER JUDICIÁRIO

Os agentes do poder judiciário acessam as informações dos Registros de Domínios para encontrar informações necessárias para o encaminhamento de ações judiciais dentro de suas jurisdições.

3.1.11 AUTORIDADE CERTIFICADORA

As autoridade certificadoras (CAs) acessam as informações dos Registros da Internet como parte do processo de verificação ao emitir certificados para hosts com nome na Internet.

3.1.12 USUÁRIOS *DOMAIN NAME SYSTEM* (DNS)

Todos os usuários da Internet possuem software cliente que resolve nomes de domínios em endereços IP e vice versa. Frequentemente, quando ocorrem problemas no processo de resolução DNS estes usuários solucionam estes problemas com ajuda da informação proveniente de Registros da Internet.

3.1.13 OUTROS ATORES

O serviço de informação dos Registros da Internet são utilizados por outras entidades como governos, organismos não governamentais responsáveis pelo estabelecimento de políticas e dentre outros. Além disso, este serviço de informação é frequentemente torna-se alvo de usuários abusivos. Exemplo: uso de informações de contatos de domínios ou endereços IP para envio de emails não solicitados.

3.2 ESTADO DA ARTE DO PROTOCOLO WHOIS

Nesta seção será realizado um estudo do estado da arte do protocolo Whois. O atual padrão para busca de informações sobre recursos da Internet, sejam nomes de domínios, endereços IP, sistemas autônomos, informação de roteamento e outros recursos dependendo das políticas locais.

3.2.1 HISTÓRIA

O protocolo Whois foi especificado pela primeira vez na RFC812 (HARRENSTIEN; WHITE, 1982) de 1982, escrita por membros do NIC do instituto *Stanford Research Institute* (SRI). Em 1985, foi publicada a RFC954 (HARRENSTIEN; STAHL; FEINLER, 1985) que substituiu a RFC812, descrevendo a nova especificação do protocolo. Atualmente todo os tipos de registros da Internet oferecem serviços de informações utilizando o protocolo whois. O título da RFC954 é NICNAME/WHOIS. O número da porta 43 está registrado junto à IANA como *nicname* mas o protocolo é mais conhecido como whois. Devido a sua história, o protocolo whois se estende além do escopo de registros de domínios. Sendo assim, é muito difícil determinar uma política para o escopo completo do whois como ele é usado atualmente.

3.2.2 TIPOS DE INFORMAÇÕES

Existem pelo menos três tipos principais de registros mantidos nos servidores whois: registros de nomes de domínios, alocação de endereços IP e sistemas autônomos. Alguns servidores whois contém um único tipo de registro, outros possuem todos os 3 tipos de informação e alguns contém informações que não se encaixam em nenhuma das três categorias como informações sobre incidentes de segurança.

3.2.3 USUÁRIOS

Para cada tipo de servidor whois existe uma série de usuários e devido a grande variedade de serviços de informações para os quais o protocolo whois é usado e muito difícil listar todos os tipos de usuários deste protocolo. É possível entretanto traçar uma lista de categorias dos principais usuários do protocolo.

3.2.4 DEFICIÊNCIAS DO PROTOCOLO

O protocolo whois foi definido há bastante tempo, antes mesmo do protocolo DNS ter sido definido. Desde então a realidade da Internet mudou bastante, as necessidades da comunidade virtual também. O sistema de registro de recursos da Internet se desenvolveu bastante durante todo este tempo, o sistema evoluiu a partir de um modelo centralizado para um modelo distribuído mundialmente. A natureza distribuída da informação sobre registros de recursos: nomes de domínios, endereços IP e sistemas autônomos, dificulta a localização da informação sobre estes recursos.

Um problema difícil de resolver é a questão de localização de informações sobre a entidade detentora de um determinado domínio. O próprio nome do domínio contém informações sobre a localização do endereço IP, pois o sistema DNS foi estruturado de maneira hierárquica e oferece de maneira rápida e eficiente uma resolução do nome de um domínio em um endereço IP. Entretanto o sistema DNS não oferece informações sobre a entidade detentora do domínio, esta informação só pode ser obtida através do protocolo whois.

Outro problema clássico é a questão de localização de informações sobre um endereço IP, como informação de contato com os administradores e informações sobre a entidade responsável por aquele endereço. A entidade responsável pelo endereço IP é responsável por um ou mais blocos contínuos de endereços IP, ou seja, ela é responsável por todos os endereços dentro de um intervalo, entre um determinado endereço inicial e um endereço final. Utiliza-se a notação de prefixos de rede para indicar estes blocos contínuos de endereços. Por exemplo, suponha que se queira contactar a entidade responsável por endereço contido no bloco 131.0.0.0/8, a localização do RIR é dificultada pois a informação está distribuída da seguinte maneira:

- 131.178.0.0/16 informações no LACNIC, alocado para o Instituto Tecnológico y

de Estudios Superiores de Monterrey;

- 131.179.0.0/16 informações no ARIN, alocado para a University of California;
- 131.180.0.0/16 informações no RIPE NCC, alocado para a Technische Universiteit Delft;
- 131.181.0.0/16 informações no APNIC, alocado para a Queensland University of Technology.

Este bloco de endereços IP assim como outros foram alocados antes da criação dos RIRs e por esta razão não foram alocados de acordo com o critério de agregação que rege os RIRs, de acordo com o qual os endereços em uma mesma região geográfica devem ser alocados de forma mais ou menos contínua e endereços para regiões geográficas distintas devem pertencer a blocos de endereços distintos.

A solução para o problema de localização de informações provavelmente é a criação de um único sistema de serviço de informações padronizado para o oferecimento da informação que está efetivamente armazenada de forma distribuída nas bases de dados de cada um dos registros da Internet. No início da Internet, a informação era centralizada. Em seguida, surgiu a necessidade de descentralização para melhor administração dos recursos. Atualmente, a necessidade exige um sistema distribuído com um serviço de informações único para prover o acesso às informações pelos usuários em qualquer parte do mundo sobre um recurso cujos dados também podem estar armazenados em qualquer outra parte do mundo.

Além do problema de localização de informações sobre recursos da Internet, existem outras questões que surgiram com a evolução da Internet. Um exemplo é a crescente preocupação com segurança nos sistemas de computação sujeitos a contantes ataques de usuários abusivos. Existem usuários que utilizam as informações do protocolo whois para fins maliciosos, um caso não crítico mas preocupante é o fato de usuários abusivos utilizarem as informações fornecidas pelo protocolo whois para o envio de correspondência não autorizada, os famosos *spammers*. Um possível solução para as preocupações com a exposição excessiva destas informações seria a criação de um mecanismo de autenticação e autorização dos usuários dos serviços de informação dos registros da Internet.

3.3 PROTOCOLO JOINT WHOIS

3.3.1 SOLUÇÃO A CURTO PRAZO

O protocolo Joint Whois corresponde à solução a curto prazo para algumas das deficiências do protocolo Whois. O Joint Whois é uma iniciativa do NRO (Number Resource Organisation), uma organização formada pelos registros da Internet para solucionar problemas de interesse comum. O protocolo não têm a intenção de se tornar um padrão do IETF mas prover uma solução a curto prazo para o problema de localização de recursos.

3.3.2 ESCOPO

O Joint Whois tem um escopo bem mais limitado que o protocolo Whois original, ele atende somente às informações armazenadas dos acRIRs, ou seja, prove somente acesso a informações sobre endereços IP e sistemas autônomos.

3.3.3 TRANSPARENTE AO USUÁRIO

Do ponto de vista do usuário, o processo de consulta é completamente transparente independente de estar consultando um servidor Whois padrão ou um servidor Joint Whois. Isso porque o Joint Whois será um serviço de informações com a mesma interface e também oferecido na porta 43 do protocolo TCP. O protocolo estará de acordo com a RFC 954 que definiu o protocolo Whois. A diferença em relação ao protocolo whois é que qualquer servidor Joint Whois será capaz de responder a todas as consultas de informações sobre endereços IP e sistemas autônomos não importando em qual RIR a informação está armazenada.

3.3.4 PROCESSO DE CONSULTA

Para atingir tal objetivo, a NRO vai oferecer um servidor Joint Whois. Este servidor ao receber uma consulta vai buscar em uma base de dados local, qual o RIR responsável por aquela informação. Uma vez determinado qual o RIR responsável pela informação, o próprio servidor Joint Whois vai fazer a consulta no servidor Whois do RIR responsável e devolver o resultado da consulta para o usuário.

Entretanto o servidor da NRO, não é muito conhecido e provavelmente os usuários continuariam consultando os servidores whois de cada um dos RIRs individualmente até encontrar a solução como fazem atualmente. Como esta é uma solução temporária não faz sentido divulgar amplamente o servidor Joint Whois da NRO para depois substituir o Joint Whois por um novo protocolo.

Por essa razão, cada um dos RIRs vai possuir um proxy para o seu servidor Whois, este proxy vai consultar uma base de dados local para descobrir qual RIR é responsável por aquela informação. Se for ele próprio o proxy encaminha a consulta ao servidor Whois do RIR ao qual ele pertence como ocorreria normalmente. Entretanto se for outro RIR, o proxy vai fazer a consulta no servidor Whois do RIR responsável pela informação e vai devolver o resultado para o usuário.

A base de dados local tanto do Joint Whois da NRO quanto dos proxies dos RIRs será populada e atualizada periodicamente a partir dos arquivos de estatísticas de cada um dos RIRs. Atualmente, estes arquivos já são gerados periodicamente pelos próprios RIRs em formato padronizado e são usados para construir gráficos informativos sobre a alocação de recursos. Estes arquivos possuem toda a informação necessária para a localização de informações sobre endereços IP e sistemas autônomos. As informações para descrição do protocolo Joint Whois, as fontes são as listas de discussão¹ e encontros do IETF².

3.4 PROTOCOLO IRIS

O protocolo IRIS está em desenvolvimento por um grupo de trabalho do IETF de mesmo nome, dentro da Área de Aplicações do IETF. O grupo tem a intenção de desenvolver um padrão para substituir o protocolo Whois atual, uma solução a longo

¹<https://www1.ietf.org/mailman/listinfo/crisp>

²<http://www3.ietf.org/meetings/meetings.html>

prazo. Os trabalhos foram divididos em duas partes: uma destinada a consultas de nomes de domínios e outra a consulta de endereços IP, sistemas autônomos e informações de roteamento.

Os trabalhos relativos a informações armazenadas por registros de domínios estão bem avançados e o grupo deve publicar as novas RFCs padronizando o novo protocolo nos próximos meses. Já os trabalhos relacionados às consultas de informações dos RIRs: endereço IP e sistemas autônomos estão menos adiantados mas em desenvolvimento intenso.

A razão para a decisão de desenvolver os trabalhos em duas frentes se deve a oposição inicial dos RIRs em participar do desenvolvimento do novo protocolo, para isso haviam desenvolvido a proposta do Joint Whois. Entretanto, frente ao sucesso do grupo no desenvolvimento de uma padrão para substituir o protocolo Whois, os RIRs decidiram depois dos trabalhos terem começado que iriam participar do desenvolvimento do novo protocolo.

Por esta razão, o desenvolvimento e implementação do servidor Joint Whois não foi descartada pelos RIRs. Ao contrário, foi adotada como uma solução a curto prazo, a ser utilizada até que o grupo conclua seus trabalhos e o padrão seja aprovado pelo IETF.

Os requisitos do novo protocolo, principalmente aqueles relativos às consultas de informações sobre domínios, foram definidos na RFC3707, de caráter informativo, escrita por Andrew Newton. Dentre os requisitos principais podemos citar: prevenção de varredura das bases de dados, autorização em níveis de acesso, máxima reutilização de recursos técnicos(ou mínima reinvenção técnica), utilização de estruturas padronizáveis e extensíveis, descentralização dentre outros.

O protocolo tem a intenção de ser compatível com os protocolos já existentes. Entretanto surge a questão se ele será capaz de se adaptar as diferentes políticas de cada um dos registros da Internet. Esta é uma função dos próprios registros da Internet, que deveriam tomar conhecimento da RFC3707 (NEWTON, 2004) e contribuir com sugestões e comentários que possam adicionar ou modificar as características do protocolo de forma a torná-lo compatível com as políticas regionais. Principalmente no atual estágio, quando ainda está em fase de desenvolvimento.

O grupo de trabalho IRIS, assim como os demais grupos de trabalho do IETF, trabalham para desenvolver e dar suporte a um padrão aberto. O grupo trabalha para o desenvolvimento para um novo protocolo para o acesso às informações dos RIR e informações sobre o registro de domínios. Muitas vezes são encontradas referências ao IRIS como o novo protocolo WHOIS. Durante o desenvolvimento o grupo avaliou a possibilidade de fazer atualizações no protocolo mas ele se mostrou limitado para

atender às necessidades atuais, por isso foi criado um novo protocolo com nome diferente para representar esse rompimento com o passado.

Os três maiores desafios encontrados pelo grupo foram:

- A grande diversidade de usuários e de suas plataformas;
- A necessidade de padronizar uma informação comum a todos os registros da Internet que na prática ainda não tem nenhum padrão;
- Padronizar as fontes dessa informação que assim como os RIRs estão distribuídos globalmente sob políticas e legislações bem diferentes.

O grupo de trabalho pretende com esse esforço criar a próxima geração de whois para qualquer tipo de registro: nome de domínio, endereço IP, sistema autônomo, etc. Além de permitir a consulta a qualquer tipo de registro, o protocolo também permite localizar o servidor autoritativo para aquele registro em qualquer parte do mundo.

Os objetivos do grupo de trabalho são:

- Ter um conjunto de consultas padrão;
- Prover um serviço global padronizado;
- Atender os requisitos para responder ao conjunto de consultas padrão;
- Permitir aos usuários encontrar respostas autoritativas para suas consultas;
- Permitir aos provedores da informação obedecer às políticas regionais.

Resumindo o objetivo é prover uma ferramenta que pode ser usada para prover um serviço de informação de registro global, que consiste em tarefa puramente técnica. O IETF não tem a pretensão de criar uma política global para o registro da Internet.

Partindo do princípio que o grupo não pretende estabelecer uma política de registro global, estes requisitos se referem apenas ao conjunto de ferramentas para que os registros regionais possam aplicar suas próprias políticas e, apesar disso, se comunicarem através de um protocolo padrão. O documento de especificação dos requisitos também define uma série de funções que permitem aos registros regionais aplicarem suas políticas regionais, funções estas que podem ser utilizadas ou desabilitadas.

As informações relativas ao desenvolvimento do grupo de trabalho CRISP podem ser encontradas no artigo *A perspective on the role of the IETF CRISP working group*

(DAIGLE; NEWTON,) ou na página do grupo³. Detalhes da especificação dos requisitos do protocolo podem ser encontrados na RFC3707 (NEWTON, 2004) publicada pelo grupo.

³<http://www.ietf.org/html.charters/crisp-charter.html>

4 ANÁLISE COMPARATIVA DE PROTOCOLOS

É possível fazer uma análise comparativa entre os diferentes protocolos de acesso a informações dos registros da Internet. Para isso é necessário estabelecer alguns critérios de comparação entre eles. Finalmente, estabelecidos os critérios é possível estabelecer uma base comparação dos diferentes protocolos. Esta análise comparativa serviu de base para a publicação do artigo "Whois, Joint Whois and CRISP"(CAIRES, 2006) na conferência internacional WWW/Internet em Múrcia, na Espanha.

4.1 CRITÉRIOS DE AVALIAÇÃO

Os critérios de avaliação foram obtidos através das características funcionais do protocolo de acesso atual: o protocolo whois. Assim como, os critérios desejáveis nos futuros protocolos de acesso diante das necessidades atuais não atendidas pelo protocolo utilizado. Os critérios foram estabelecidos tendo em vista a necessidade de criação de um sistema único de acesso a informações de registros da Internet. As fontes utilizadas foram: especificação do protocolo NICNAME/WHOIS (DAIGLE, 2004), especificação dos requisitos do protocolo IRIS (NEWTON, 2004), os *drafts* de definição dos protocolos que fazem parte do protocolo IRIS candidatos a se tornarem padrões, a lista de discussão do grupo de trabalho e os fóruns de discussão. Os critérios assim estabelecidos se dividem em duas classes assim como os trabalhos do grupo CRISP. A primeira se refere aos requisitos funcionais básicos e comuns a todos os tipos de registro. A segunda diz respeito aos requisitos específicos do registro de domínios.

4.1.1 CRITÉRIOS FUNCIONAIS BÁSICOS

Os critérios funcionais, chamados básicos pois são comuns a todos os tipos de Registros da Internet, são os seguintes:

1. **Prevenção contra a varredura de dados:** de forma a evitar a aquisição inapropriada de dados de um registro da Internet, muitos servidores limitarão a quantidade de dados que pode ser retornada em um intervalo fixo de tempo, de um servidor para um mesmo cliente. Os limites estabelecidos em diferentes tipos de dados e aplicativos dependendo do status de acesso podem diferir de um servidor para outro baseado na política e na necessidade.
2. **Mínima reinvenção técnica:** os protocolos não devem empregar soluções tecnológicas únicas nas camadas acima da camada de rede e de transporte a não ser quando necessário. Para as camadas de rede e de transporte os protocolos devem necessariamente utilizar os padrões definidos pelo IETF, sem nenhuma reinvenção tecnológica. Entretanto os protocolos podem implementar mecanismos de controle de congestionamento.
3. **Facilidade de localização de recurso:** a dificuldade de localização de recursos foi bem exemplificada na descrição das deficiências do protocolo Whois. Tanto no caso de informações de domínios quanto de endereços de rede e sistemas autônomos.
4. **Autenticação:** atualmente o protocolo de acesso a informações não oferece nenhum tipo de mecanismo de autenticação.
5. **Autorização:** sem autenticação as políticas de autorização são inexistentes e todas as tentativas de restrição de acesso a informações ficam limitadas ao controle de acesso a informações cujas consultas são provenientes de endereços IP aos quais foi associado nível mais baixo (por já ter ocorrido um ou mais incidentes utilizando este endereço) ou mais alto de acesso (exemplo quando o endereço é utilizado por um determinado *Internet Service Provider* (ISP) no exercício de suas funções) em relação a um usuário anônimo utilizando um IP desconhecido.
6. **Estruturas de dados padronizadas e extensíveis:** atualmente os registros da Internet não possuem nenhuma padronização das informações armazenadas em seus bancos de dados. Apesar de armazenarem informações sobre os mesmos tipos de recursos as estruturas de dados utilizadas são diferentes o que dificulta a troca de informações entre eles.
7. **Padronização de consultas e Respostas:** atualmente os Registros da Internet não possuem nenhum padrão quanto ao tipo de consulta aceita e também não possuem padrão para a resposta enviada ao cliente. Muitas vezes registros do mesmo tipo de recurso possuem alguns tipos de consultas que variam, nos tipos de entradas aceitos por exemplo, e as respostas nunca não seguem um padrão.

8. **Processamento cliente:** as respostas devem ser processadas pelo cliente. Atualmente a resposta a uma consulta sobre um recurso a um registro da Internet deve ser interpretada pelo usuário. Ela não pode ser processada pelo cliente pois não existe um padrão para consultas e respostas.
9. **Referência a entidades:** atualmente um servidor que não possua informações sobre um determinado recurso em sua base de dados não possui referência nenhuma ao servidor que possa ter autoridade sobre o recurso consultado. Neste caso, o servidor simplesmente envia uma mensagem de erro padrão, possivelmente contendo uma referência a um servidor padrão que pode ou não ser o autoritativo sobre o recurso consultado.
10. **Descentralização:** o serviço de acesso a informações de recursos de Registros da Internet supõe que a base de dados é centralizada mas na verdade é distribuída, assim quando não encontra o dado retorna uma mensagem de erro. Isso ocorre porque quando o serviço de informações foi desenvolvido a base de dados era efetivamente centralizada, atualmente a base é distribuída e o serviço de acesso a informações obsoleto.
11. **Permissão de acesso a consultas:** se o sistema de acesso a informações de registros da Internet possuir um mecanismo de autenticação e autorização pode ser desejável ter um mecanismo através do qual um cliente possa determinar se ele tem permissão para fazer uma consulta antes da consulta ser submetida. Obviamente, isto só será possível nos casos em que essa informação não for confidencial.
12. **Distribuição da autenticação:** num futuro próximo, quando o sistema de acesso a informações de Registros da Internet possuir um mecanismo de autenticação e autorização será desejável ter um sistema de distribuição de autenticação. Isso porque as bases de dados serão distribuídas e o sistema de informação único, não faz sentido um usuário se autenticar várias vezes, uma para cada nível de indireção, ao fazer uma única consulta.
13. **Padronização das respostas de erro:** não existe nenhuma padronização das consultas, das respostas ou dos códigos erros no sistema de informação atual. Entretanto, a existência de padronização dos códigos de erro é desejável no serviço de informações. Assim um cliente pode tratar o erro e decidir se o erro é temporário (por exemplo: servidor sobrecarregado) então deve refazer a consulta ou se o erro é permanente (por exemplo: permissão negada) e não deve tentar de novo. Este requisito poderia desta forma possibilitar um mecanismo de controle de congestionamento.

14. **Referências do lado do cliente ou dos servidores:** já foi descrito o processo executado por um servidor ao ser consultado sobre um recurso que não consta em sua base de dados. O servidor atualmente não possui uma referência exata ao servidor que detém aquela informação e retorna uma mensagem padrão com referência a um servidor padrão que pode ou não possuir a informação desejada. De qualquer maneira, a referência ao segundo servidor é feita dentro de uma resposta e não segue nenhum padrão, por esse motivo não poderia ser tratada de forma automática pelo software cliente apenas pelo usuário. A referência ao servidor que possivelmente possui a informação desejada é chamada de referral.
15. **Atributos de privacidade:** seria desejável em um futuro próximo poder atribuir níveis de privacidade à informação oferecida a um cliente autenticado, por exemplo: informação pública ou informação confidencial não divulgar e assim por diante. Apenas no protocolo IRIS um servidor autoritativo é capaz de informar o nível de privacidade da informação concedida. Ao retornar um conjunto de resultados a um usuário autorizado o servidor marca a informação como pública, não redistribua, e etc.

4.1.2 CRITÉRIOS FUNCIONAIS ESPECÍFICOS DE REGISTROS DE DOMÍNIOS

Nesta seção serão descritos os requisitos dos protocolos de acesso a informações de registros de domínio da Internet. Os requisitos específicos de outros tipos de registro estão fora do escopo deste estudo. Além disso, o protocolo Joint Whois não se aplica a registros de domínios por isso os critérios estabelecidos nesta seção não se aplicam a este protocolo. É importante atentar para o fato de que um servidor pertencente a qualquer um dos tipos de registros não precisa atender todos os requisitos de todos os tipos de registros da Internet. Ao contrário, um servidor só precisa implementar os requisitos mínimos necessários ao tipo de registro ao qual ele pertence.

1. **Lookups:** lookups correspondem a um tipo de busca exata em que a entrada identifica unicamente uma entidade, que pode ser um domínio, um contato ou pessoa responsável pelo domínio ou um servidor de nomes. Sendo assim, o conjunto de informações resultantes desta consulta refere-se a uma única entidade ou nenhuma, caso a entidade não exista e o resultado da consulta seja vazio. Dependendo da política e da necessidade de um registro de nomes ele pode permitir todos ou pelo menos algum dos tipos de lookup. Pelo mesmo motivo, um registro pode não retornar parte da informação ou simplesmente negar

qualquer informação sobre o recurso consultado.

2. **Lookup de contato:** lookup de contato corresponde a busca das informações de contato sobre uma pessoa responsável por um determinado domínio dada uma entrada que o identifique, por exemplo, seu nome completo.
3. **Lookup de servidor de nomes:** lookup de um servidor de nomes corresponde a busca de informações de um determinado servidor de nomes baseado em uma entrada que pode ser seu endereço IP ou o nome completo do servidor.
4. **Lookup de domínio:** lookup de um domínio é o tipo de busca mais comum. Dado um nome de domínio completo (fully qualified domain name) esta consulta retorna todas as informações relativas a este domínio ou um conjunto vazio caso ele não exista.
5. **Buscas:** as buscas são um tipo de consulta mais flexível do que os lookups. A consulta deve aceitar uma entrada que identifique um único recurso : nome de domínio, contato de domínio ou servidor de nomes : assim como deve aceitar uma entrada que identifique um conjunto de recursos, dentro de um certo limite.
6. **Busca de nome de domínio:** uma busca por nome de domínio deve aceitar um nome de domínio completo como entrada, assim como deve aceitar uma parte do nome do domínio. No primeiro caso, o resultado da consulta vai conter informações relativas a um único domínio ou um conjunto vazio caso o domínio não exista. No segundo caso, o resultado da consulta vai retornar informações sobre todos os domínios que possuam a entrada como parte de seu nome caso este conjunto não exceda um determinado limite, ou a consulta pode retornar um conjunto vazio caso não exista nenhum domínio com a entrada como parte de seu nome. Assim como nos *lookups*, nas buscas um servidor pertencente
7. **Busca do detentor do domínio:** é desejável ter uma busca pelo detentor de um domínio tendo como entrada seu nome completo ou parcial. O resultado da consulta deve conter as informações de uma determinada pessoa, empresa ou organização detentora de um ou mais domínios, ou de vários detentores de domínios cujo nome ou razão social contém o valor fornecido como entrada.
8. **Busca de domínios hospedados em determinado servidor:** é desejável também ter uma busca dos domínios hospedados em um determinado servidor de nomes dado o nome completo do servidor ou seu endereço IP.
9. **Padronização de resposta:** a estrutura de dados padrão deve ser capaz de expressar as seguintes informações de domínios: o status (ativo ou inativo), o detentor do domínio, os contatos administrativo, técnico e de cobrança, registro

de domínio do qual o domínio foi registrado e registrar, caso exista. A estrutura de dados deve ser flexível suficiente para poder armazenar informações adicionais e opcionais na forma de texto, como por exemplo uma notificação explicando o status do domínio, notícias de propaganda e referências na forma de URLs (Universal Resource Location). Mais uma vez, não é esperado que todo registro de domínio forneça todas as informações descritas acima, as estruturas de dados empregadas pelo protocolo de acesso devem ser capazes de armazenar todas estas informações entretanto só serão utilizadas se o registro de domínios em questão precisar provê-las.

10. **Suporte a serialização:** as estruturas de dados utilizadas pelo protocolo devem ser serializáveis. A serialização permite a execução de uma série de operações importantes como backup e recuperação de falhas e balanceamento de carga.
11. **Limite no tamanho da resposta:** o protocolo deve prover um mecanismo a ser usado de acordo com o critério de cada registro , que permita ao servidor estabelecer um limite no tamanho da respostas a consulta de um cliente. O protocolo deveria fazer 3 distinções: (1) resposta vazia, (2) resposta truncada com o propósito de evitar gargalos provocados por congestionamento na rede e (3) resposta truncada para evitar varredura de dados.
12. **Referência a delegação DNS:** o protocolo deve usar o modelo de delegação DNS como meio padrão para determinar a fonte autoritativa da informação referente a domínios, servidores de nomes, endereço IP e outros objetos quando aplicáveis. Em outras palavras, quando um recurso é naturalmente mapeável por DNS o comportamento desejável é consultar o sistema DNS para encontrar o servidor autoritativo contendo a informação sobre aquele recurso.
13. **Distribuição para diferentes modelos de registros de domínio:** dependendo do modelo de registro em uso, dados técnicos de um domínio podem residir em um servidor enquanto os dados de contato para o mesmo domínio podem residir em um servidor administrativo em outra entidade. Entretanto em muitos casos esta situação não ocorre. O protocolo de acesso a informações de registros deve fornecer tanto as informações técnicas quanto administrativas não importando o tipo de modelo de registro.
14. **Omissão de dados:** quando um valor em uma resposta não puder ser concedido devido a uma restrição política o protocolo deve ser capaz de expressar o valor de uma das três formas: (1) completa omissão do valor sem explicação, (2) uma indicação de que o valor não pode ser visualizado devido a autorização insuficiente e (3) uma indicação de que o valor não pode ser visualizado devido à política de privacidade sem mencionar a falta de permissão.

15. **Internacionalização:** as estruturas de dados que definem os recursos relacionados a domínios devem seguir a RFC2277 (ALVESTRAND, 1998) com relação aos dados tipo texto. Particularmente, a estrutura de dados deve indicar o conjunto de caracteres e o idioma em uso com dados tipo texto não estruturados. O protocolo deve ser capaz de suportar múltiplas representações de dados de contato pois as representações variam de um registro de domínio para outro.

4.2 ANÁLISE COMPARATIVA

Os resultados da análise comparativa dos protocolos Whois, Joint Whois e IRIS serão apresentados nesta seção. Os resultados serão apresentados na mesma seqüência em que foram apresentados os critérios de avaliação.

4.2.1 CRITÉRIOS FUNCIONAIS BÁSICOS

1. **Prevenção contra a varredura de dados:** as especificações dos protocolos Whois e Joint Whois não possuem nenhum mecanismo de controle de varredura de dados mas a especificação dos requisitos do protocolo IRIS possui. Na prática, muitos servidores Whois restringem o controle de varredura de dados ao limitar o número de consultas provenientes de um mesmo endereço IP em um determinado intervalo de tempo.
2. **Mínima reinvenção técnica:** o protocolo Whois não é novo por isso este requisito não se aplica. O protocolo Joint Whois não possui nenhuma reinvenção técnica pois se baseia quase que totalmente no protocolo Whois. Finalmente, um dos principais requisitos do protocolo IRIS é a mínima reinvenção técnica nas camadas acima da camada de rede do modelo TCP/IP e para as camadas de rede e de transporte a regra é uso dos padrões definidos pelo IETF. Os protocolos não devem empregar soluções tecnológicas únicas nas camadas acima da camada de rede e de transporte a não ser quando necessário. Para as camadas de rede e de transporte os protocolos devem necessariamente utilizar os padrões definidos pelo IETF, sem nenhuma reinvenção tecnológica. Entretanto os protocolos podem implementar mecanismos de controle de congestionamento.

3. **Facilidade de localização de recurso:** o protocolo Whois não prevê nenhum mecanismo para localização de recursos caso a informação consultada não esteja armazenada na base de dados local. O protocolo Joint Whois foi desenvolvido especificamente para solucionar este problema mas restringe seu escopo a consultas de endereços IP e sistemas autônomos, ou seja, informações dos RIRs. Este requisito foi uma das principais motivações para o desenvolvimento do protocolo IRIS, que por sua vez não restringe seu escopo ao registro de endereços IP e sistemas autônomos mas abrange: nomes de domínios, endereços IP, sistemas autônomos e prevê mecanismos de extensão para suporte a outros tipos de recursos.
4. **Autenticação:** as especificações dos protocolos Whois e Joint Whois não possuem nenhum mecanismo de autenticação mas a especificação dos requisitos do protocolo IRIS possui. Na prática, muitos servidores Whois aplicam o controle do número de consultas provenientes do mesmo endereço IP em determinado intervalo de tempo a endereços IP pertencentes a ISPs por exemplo.
5. **Autorização:** como as especificações do Whois e Joint Whois não exigem nenhum mecanismo de autenticação, eles também não possuem nenhum mecanismo de autorização. Por outro lado, a especificação dos requisitos do IRIS prevê o oferecimento de mais de um mecanismo de autenticação, com opção de escolha entre eles ou escolha pela não utilização de nenhum mecanismo de autorização dependendo da política regional.
6. **Estruturas de dados padronizadas e extensíveis:** as especificações dos protocolos Whois e Joint Whois não possuem definição de estrutura de dados padrão e não são extensíveis. O protocolo Joint Whois apenas aproveita a facilidade da padronização já existente das estruturas de dados dos arquivos de estatísticas dos RIRs para construir o banco de dados utilizado pelo servidor Joint Whois ou pelo proxy dos servidores Whois de cada RIR. A especificação dos requisitos do IRIS enfatiza a necessidade desta padronização para o sucesso do novo protocolo.
7. **Padronização de consultas e respostas:** o protocolo Whois não possui padronização das consultas e respostas oferecidas, cada registro da Internet implementa o protocolo e oferece os serviços da maneira que lhe for mais conveniente. O protocolo Joint Whois define um conjunto restrito de consultas e respostas. As consultas e respostas do protocolo IRIS são padronizadas e estruturadas de forma que possam ser processadas pelo software cliente, pois se cada protocolo whois respondessem a um conjunto de consultas diferentes e de

forma diferente, seria praticamente impossível o oferecimento de um serviço de informações único com base de dados distribuída.

8. **Processamento cliente:** os protocolos Whois e Joint Whois não prevêem nenhum mecanismo de suporte ao processamento cliente. O primeiro retorna uma mensagem de erro não padronizada caso a informação consultada não esteja armazenada na base de dados local. O segundo concentra todo o processamento do lado do servidor. O IRIS segue a tendência de concentrar o processamento no lado do cliente para evitar tráfego na rede e a sobrecarga do servidor.
9. **Referência a entidades:** no protocolo Whois não existe nenhuma especificação de referência a uma informação que não esteja armazenada em sua base de dados local. No protocolo Joint Whois esta referência fica armazenada em uma base de dados separada populada a partir dos arquivos de estatísticas dos RIRs. O protocolo IRIS prevê um mecanismo padrão para que uma entidade armazenada em um servidor seja referenciada por uma entrada em outro servidor, de forma que o segundo possa informar ao cliente a referência ao primeiro, para que o cliente possa refazer a consulta e obter a resposta desejada.
10. **Descentralização:** o protocolo Whois é essencialmente centralizado uma vez que supõe que a informação consultada está na base de dados local, caso contrário, retorna uma mensagem de erro. O Joint Whois não exige que os dados estejam em um repositório central. Assim como o IRIS, pois um dos requisitos do protocolo é a não exigência que os dados estejam armazenados em um repositório central.
11. **Permissão de acesso a consultas:** o protocolo Whois e Joint Whois não implementam autenticação, conseqüentemente não implementam a autorização. Por outro lado, o IRIS além de oferecer mecanismos de autenticação e autorização, ainda deve fornecer meios para que um cliente seja capaz de determinar se uma consulta será negada antes mesmo dela ser submetida. No IRIS, este é um mecanismo opcional, assim como a autorização e autenticação, ou seja, o operador pode desabilitar se sua política não permitir.
12. **Distribuição da autenticação:** não se aplica aos protocolos Whois e Joint Whois. No protocolo IRIS, apesar de oferecer o mecanismo de autenticação e autorização eles devem ser opcionais, assim como ao oferecer um serviço de informações envolvendo diversos registros de Internet o protocolo não deve exigir autenticação de nenhum dos participantes do serviço de informação. Entretanto deve oferecer um serviço opcional de distribuição de autenticação caso seja necessário.

13. **Padronização das respostas de erro:** não se aplica aos protocolos Whois e Joint Whois. O protocolo IRIS possui um conjunto mínimo obrigatório de códigos de erros padronizados que todos os servidores devem retornar em caso de erro, isso permite que o software cliente tome a melhor decisão baseado no código de erro específico.
14. **Tratamento de referências:** caso um servidor não possua a informação consultada ele normalmente faz uma referência ao servidor autoritativo sobre aquela informação, caso ele não tenha informação exata no servidor autoritativo retorna uma referência a um provável servidor. No protocolo Whois esta referência é sempre um provável servidor autoritativo pois o protocolo não possui a informação exata e a resposta é padrão para todos os casos, além disso, a resposta é somente tratada por um ser humano pois o software cliente não é capaz de entender a referência e refazer a consulta para o servidor referenciado. No protocolo Joint Whois, caso o servidor não possua a informação consultada ele mesmo se encarrega de buscar a referência ao servidor autoritativo e refazer a consulta para o mesmo, ou seja o tratamento da referência é feita no lado do servidor, completamente transparente ao software cliente e ao usuário. O mecanismo de referência é considerado *server-side*. No protocolo IRIS, existe um mecanismo de referência padrão e estruturado de forma que a aplicação cliente consiga entender e seja capaz de refazer a consulta ao servidor referenciado de maneira automática.
15. **Atributos de privacidade:** apenas no protocolo IRIS um servidor autoritativo é capaz de informar o nível de privacidade da informação concedida em resposta a uma consulta. Ao retornar um conjunto de resultados a um usuário autorizado o servidor marca a informação como pública, não redistribua, e etc.

4.2.2 CRITÉRIOS FUNCIONAIS ESPECÍFICOS DE REGISTROS DE DOMÍNIOS

Uma observação importante a ser feita é que o protocolo Joint Whois não se aplica a registros de domínios, por isso os critérios estabelecidos nesta seção não se aplicam a este protocolo. Sendo assim, a análise comparativa dos protocolos através dos critérios específicos de registros de domínios é apresentada a seguir:

1. **Lookups:** o protocolo Whois em sua especificação não padroniza o conjunto de consultas que um servidor Whois deve aceitar. Por essa razão, cada servidor

responde a um conjunto de consultas que pode variar em número e sintaxe, assim como no tipo de entrada e no resultado obtido. O conjunto de consultas varia de um registro para outro de acordo com a política e necessidade local. O critério não se aplica ao protocolo Joint Whois. Por sua vez, o protocolo IRIS possui dentre os seus requisitos necessários os três tipos de consulta citados anteriormente.

2. **Buscas:** assim como foi descrito sobre lookups, o protocolo Whois não padroniza o conjunto de consultas que um servidor Whois deve aceitar, as implementações variam de um registro de domínio para outro o que dificulta a criação de um sistema único de acesso a informações de registros. O critério não se aplica ao protocolo Joint Whois. Por outro lado, o protocolo IRIS especifica o conjunto de consultas que um servidor IRIS deve aceitar e as consultas são: busca de nome de domínio, detentor de domínio e domínios hospedados em determinado servidor de nomes de acordo com as especificações da seção 4.1.
3. **Padronização de resposta:** o protocolo Whois não possui padronização das consultas fornecidas por um servidor Whois conseqüentemente também não padroniza as informações retornadas em resposta às consultas suportadas. O critério não se aplica ao protocolo Joint Whois. Entretanto, o protocolo IRIS padroniza as informações fornecidas em resposta a cada uma das consultas do conjunto padronizado de consultas suportadas por um servidor IRIS. Desta forma, a padronização das informações retornadas em cada tipo de consulta associada a padronização do conjunto de consultas suportadas viabiliza a criação de um sistema único de acesso a informações de registros de domínios.
4. **Suporte a serialização:** os protocolos Whois e Joint Whois não possuem suporte a serialização. Entretanto, este é um requisito importante da especificação do protocolo IRIS que, conforme discutido na seção 4.1, viabiliza o processo de backup, balanceamento de carga e recuperação de falhas.
5. **Limite no tamanho da resposta:** geralmente, os servidores Whois estabelecem um limite para o tamanho da resposta do servidor para um determinado cliente em uma mesma consulta. Porém este limite não faz parte da especificação do protocolo. O protocolo Joint Whois não se aplica a registros de domínios e o protocolo IRIS deve fornecer três opções de limites para o resultado de uma consulta: a resposta vazia, resposta truncada para evitar congestionamento e resposta truncada para evitar varredura de dados.
6. **Referência a delegação DNS:** o protocolo Whois não prevê a utilização da delegação DNS para obtenção da informação autoritativa sobre um determinado

recurso pois se o próprio servidor não tiver a informação sobre o recurso em sua base de dados não se responsabiliza por fornecer a informação ou uma referência de onde encontrá-la. O protocolo IRIS deve utilizar a delegação DNS para determinar o servidor autoritativo sobre determinado recurso e retornar uma referência deste servidor para que o cliente possa refazer a consulta ao servidor autoritativo e obter a resposta desejada.

7. **Distribuição para diferentes modelos de registros de domínio:** cada registro de domínio implementa o protocolo Whois de acordo com seu modelo de registro e suas necessidades por isso o protocolo se adapta aos diferentes modelos mas não possui uma padronização que possa ser aplicada a todos os registros de domínio independente de seu modelo. Este é um requisito necessário para a viabilização de um sistema único de informações. O IRIS estabelece um protocolo padrão aplicável a qualquer registro de domínio independente de seu modelo.
8. **Omissão de dados:** devido a falta de padronização no protocolo Whois cada registro de domínio tem uma política para a omissão de dados e muitas vezes o mesmo nível de omissão é implementado de forma diferente e o usuário recebe uma resposta diferente de cada um dos registros. No IRIS a padronização é a palavra de ordem em todos os seus requisitos pois um dos seus objetivos é a viabilização de um sistema de acesso a informações único para todos os registros da Internet, sendo assim ele deve implementar os três tipos de omissão de dados descritos anteriormente.
9. **Internacionalização:** no protocolo Whois seguir ou não os critérios de internacionalização é uma questão de política do registro de domínio em questão. No IRIS é mais um dos requisitos padronizados para todos os registros de domínios como descrito anteriormente.

4.3 RESULTADOS

Sendo assim, os resultados da comparação dos três protocolos através dos critérios funcionais básicos definidos na seção anterior, sendo denominados básicos pois são comuns a todos os tipos de Registro, podem ser visualizados na tabela 1:

Tabela 1: Análise comparativa dos protocolos de acesso a informações de Registros da Internet através dos critérios funcionais básicos

Critério	Whois	Joint Whois	IRIS
Prevenção contra Varredura de Dados	Não possui	Não possui	Possui
Reinvenção Técnica	Não	Não	Mínima
Facilidade de Localização de Recursos	Nenhuma	Sim (IPs e ASs)	Sim (sem restrição)
Autenticação e Autorização	Não	Não	Sim
Estruturas Padronizadas e Extensíveis	Não	Não	Sim
Padronização de Consultas e Respostas	Não	Não	Sim
Processamento no Cliente	Não	Não	Sim
Referência a Recursos Remotos	Não	Sim (IPs e ASs)	Sim (sem restrição)
Descentralização	Não	Sim (IPs e ASs)	Sim (sem restrição)
Consulta de Permissão de Acesso	Não	Não	Sim
Distribuição de Autenticação	Não	Não	Sim
Padronização das Respostas de Erro	Não	Não	Sim
Tratamento de Referências	Não	Sim (servidor)	Sim (cliente)
Indicadores de Privacidade	Não	Não	Sim
Linguagem Semântica	Não	Não	Não

Tabela 2: Análise comparativa dos protocolos de acesso a informações de Registros da Internet através dos critérios específicos de registros de domínios

Critério	Whois	Joint Whois	IRIS
Lookups	Sim (políticas locais)	Não se aplica	Sim (padrão)
Buscas	Sim (políticas locais)	Não se aplica	Sim (padrão)

continua...

Tabela 2: Análise comparativa dos protocolos de acesso a informações de Registros da Internet através dos critérios específicos de registros de domínios

Critério	Whois	Joint Whois	IRIS
Respostas	Sem padrão	Não se aplica	Com padrão
Serialização	Não	Não se aplica	Sim
Limite para respostas	Sim (políticas locais)	Não se aplica	Sim (padrão)
Delegação DNS	Não utiliza	Não se aplica	Utiliza
Diferentes modelos de registro	Não	Não se aplica	Sim
Omissão de Dados	Sim (políticas locais)	Não se aplica	Sim (padrão)
Internacionalização	Não	Não se aplica	Sim

4.4 DISCUSSÃO

Como pode-se observar pelos resultados da comparação entre os três protocolos, de uma maneira geral: o protocolo Whois apesar de ser o único dos três protocolos atualmente padronizado pelo IETF se apresenta obsoleto em vários aspectos, já o protocolo Joint Whois não têm o intuito de se tornar um padrão e tem escopo restrito a acesso a informações de endereços IP e sistemas autônomos, porém o IRIS tem previsão de se tornar o novo padrão do IETF para protocolo de acesso a informações de registros de Internet nos próximos meses e foi desenvolvido para suprir todas as deficiências do protocolo Whois e atender a novas necessidades da comunidade virtual.

Quanto aos requisitos funcionais, o protocolo Whois, pelo fato de ter uma especificação muito aberta, dá margem para que cada registro da Internet implemente seu próprio servidor Whois de acordo com suas políticas, sendo possível implementar todos os requisitos funcionais básicos sem violar o padrão. Entretanto, o fato da especificação ser aberta dificulta qualquer tentativa de padronização do serviço de informações e inviabiliza a criação de um sistema único de acesso a informações de registros da Internet.

O protocolo Joint Whois têm um escopo bem mais restrito em relação aos protocolos Whois e o protocolo IRIS. Dentro de seu escopo, que envolve apenas as informações de consultas às informações de endereços IP e sistemas autônomos, ele atende alguns requisitos além dos atendidos pelo protocolo Whois mas não atende todas as novas necessidades da comunidade virtual. Ele provê um sistema único de acesso informações mas não provê nenhum mecanismo de autenticação e autorização.

O protocolo IRIS possui um escopo amplo pois tem o intuito de atender registros de domínio e endereços IP, e ser extensível a outros tipos de recursos e registros da Internet. Além disso, ele atende todas as necessidades da comunidade virtual e apresenta novas funções desejáveis, como marcar a informação fornecida a um cliente autorizado com um determinado nível de privacidade.

Entretanto, apesar do protocolo IRIS estar em um estágio de desenvolvimento bem avançado, principalmente no desenvolvimento de sistemas de acesso a informações de registros de domínio, ainda existem algumas dificuldades a serem enfrentadas. Cada um dos registros da Internet, deve se envolver com o desenvolvimento do novo protocolo contribuindo com sugestões para que o protocolo possa ser adaptado às suas políticas regionais.

Além disso, para que um sistema único de acesso a informações de registro possa ser criado todos os registros devem adotar o novo padrão e não há nenhuma regra que exija dos registros da Internet a adoção de um novo protocolo, tudo depende da cooperação entre eles. A adoção do novo protocolo certamente vai impactar os servidores de acesso a informações de todos os registros, na forma como seus dados são armazenados e apresentados.

O protocolo IRIS pode ser tão eficiente ao viabilizar um sistema único de acesso a informações de registros da Internet quanto revolucionário na forma como cada um deles oferece este serviço. Cada registro deve ser capaz de rever, revisar documentar cada uma de suas políticas, de forma a continuar exercendo suas políticas regionais mesmo utilizando o novo protocolo.

4.5 CONCLUSÃO

Durante a análise comparativa dos protocolos Whois, Joint Whois e IRIS foi possível constatar as deficiências do protocolo atual: o protocolo Whois. Assim como foi possível constatar que apesar do protocolo Joint Whois se propor a solucionar alguns problemas do protocolo Whois, ele possui um escopo bem mais limitado que o primeiro e

não atende todas as necessidades da comunidade virtual por isso o Joint Whois pode ser entendido como uma solução a curto prazo. Por outro lado, o protocolo CRISP foi desenvolvido para suprir todas as deficiências do protocolo Whois e viabilizar um sistema único de acesso a informações de registros da Internet e se apresenta como a melhor solução a longo prazo para as deficiências do sistema de informações atual.

O desenvolvimento do protocolo IRIS têm sido um desafio e a implementação de um sistema único de acesso a informações de registros utilizando o novo protocolo é um desafio ainda maior a ser enfrentado nos próximos meses e anos. Apesar destes desafios e segundo os dados apresentados o protocolo IRIS tem grandes perspectivas de ser o melhor sucessor para o protocolo Whois.

5 ESTRUTURA DO PROTOCOLO IRIS

Quando surgiam problemas de interconexão na ARPANET no início da década de 80, os operadores de rede precisavam do número do telefone de outro operador de rede para solucionar o problema e para tal executavam o seguinte comando:

```
\$whois !nancy
```

e recebiam a seguinte resposta:

```
Fischer, Nancy C. (NANCY)    FISCHER@SRI-NIC SRI International  
Telecommunication Sciences Center  
333 Ravenswood Avenue, EJ289  
Menlo Park, California 94025  
Phone: (415) 859-2539  
MILNET TAC user
```

Então o operador possuía o número de telefone para ligar. Por essa razão, todos os operadores eram orientados a fornecer suas informações de contato para a base de dados central do SRI NIC e essa informação poderia ser acessada por qualquer outro operador de rede via protocolo Whois. O protocolo Whois foi desenvolvido para acessar informação de contatos para a ARPANET pertencente ao Departamento de Defesa dos Estados Unidos. As aplicações clientes eram configuradas para executar as consultas diretamente no servidor do SRI NIC.

Atualmente, o SRI NIC não existe mais, a ARPANET se transformou na Internet e as informações sobre contatos encontram-se distribuídas nas bases de dados dos diversos registros. Entretanto o protocolo Whois praticamente não mudou desde sua criação. A informação continua disponível para quem desejar acessar, incluindo *spammers* e outros indivíduos com más intenções. Contudo, usuários legítimos encontram dificuldade em descobrir o servidor autoritativo sobre a informação desejada pois a base de dados encontra-se distribuída. Além disso, muitas vezes eles se confundem com a diversidade no formato e conteúdo das consultas e respostas. Por essa razão e por todos os outros motivos levantados durante a análise comparativa dos protocolos realizada anteriormente foi necessário desenvolver um novo protocolo.

5.1 EVOLUÇÃO

Depois de uma série de sessões para discutir a melhor forma de aprimorar o protocolo Whois, o IETF decidiu criar um grupo de trabalho denominado CRISP¹ para desenvolver um novo protocolo para substituí-lo. A especificação do protocolo Whois é tão simplificada que melhorar o protocolo exigiria o mesmo trabalho necessário para desenvolver um novo protocolo com a desvantagem de exigir compatibilidade com a versão anterior. O objetivo do grupo foi suprir as necessidades não atendidas pelo protocolo Whois, sendo inicialmente restrito a nomes de domínios e, posteriormente, sendo estendido para endereços IP e sistemas autônomos. Atualmente, o protocolo IRIS é genérico suficiente para se aplicar a qualquer recurso.

Inicialmente o grupo de trabalho CRISP avaliou duas abordagens:

- O protocolo *Federated Internet Registry Service* (FIRS) baseado em LDAP e especificado na forma de *draft* de RFC (HALL, 2003).
- O protocolo IRIS baseado em XML sobre BEEP e especificado na forma de RFC (NEWTON; SANZ, 2005b).

O grupo concluiu que as duas abordagens seriam razoáveis mas acabaram optando pela segunda. A princípio desenvolver um protocolo baseado em LDAP parecia ser a abordagem mais prática dada a disponibilidade de aplicações LDAP. Porém baseado na experiência da Verisign, que é atualmente o maior registro de nomes de domínios, ao desenvolver um protótipo de serviço Whois utilizando LDAP, adaptar o LDAP para busca de nomes de domínios como é realizada no protocolo Whois não era uma tarefa tão simples.

Além de não ser uma tarefa nada trivial os próprios operadores de rede não consideravam uma vantagem a forma de busca realizada com LDAP, em que eles precisam realizar uma busca na árvore de diretórios ao invés de simplesmente especificar uma consulta. Os operadores possuem rotinas de gerenciamento de incidentes adaptadas ao tipo de consulta realizada com o protocolo Whois e a utilização do LDAP causaria um impacto indesejado nas rotinas.

Várias mudanças no cliente e servidor LDAP seriam necessárias para o funcionamento do sistema de informação. O servidor LDAP precisaria ser flexível suficiente para abranger as políticas locais de acesso. Os clientes LDAP também precisariam implementar uma série de modificações, por exemplo, para ter suporte a nomes de domí-

¹<http://ietf.org/html.charters/crisp-charter.html>

nios internacionalizados *Internationalized Domain Names* (IDN) (FALTSTROM; HOFFMAN; COSTELLO, 2003), o cliente precisaria distinguir entre nomes de domínios ASCII e o equivalente internacionalizado. Devido ao excessivo número de alterações necessárias a proposta LDAP foi descartada. O grupo percebeu que desenvolver um novo protocolo seria a melhor alternativa para adequar o protocolo as particularidades do serviço de informação Whois.

5.2 ESTRUTURA DO PROTOCOLO

O protocolo IRIS consiste em um protocolo da camada de aplicação composto de múltiplos blocos funcionais. O novo protocolo utiliza outras tecnologias como componentes funcionais para prover flexibilidade e extensibilidade. A configuração inicial dos blocos funcionais era formado por três partes principais: a especificação do conjunto padrão de consultas e respostas, um protocolo de transporte denominado BEEP e um sistema para descobrir serviços dinamicamente. Entretanto a estrutura do protocolo já se mostrou flexível suficiente para que fosse possível substituir o protocolo de transporte BEEP, por duas outras formas alternativas de transporte: o protocolo LWZ e o protocolo XPC.



Figura 1: Diagrama dos blocos funcionais do protocolo IRIS

Cada bloco funcional é responsável pela especificação e implementação de uma funcionalidade específica e cada um dos blocos funcionais foi especificado em uma RFC separada:

- **IRIS:** A camada denominada IRIS, corresponde ao protocolo *IRIS Core*, que consiste em um protocolo da camada de aplicação para representação genérica dos tipos básicos de consultas e respostas para serviços de informação de registros

da Internet, como especificado na RFC3981 (NEWTON; SANZ, 2005b). A especificação consiste fundamentalmente na descrição do *Schema* XML utilizado como formato padrão na troca de dados entre aplicações cliente e servidor.

- **DREG:** Consiste na especificação e descrição de um *Schema* XML para o protocolo IRIS com o objetivo de prover serviço de informação sobre nomes de domínios, como especificado na RFC3982 (NEWTON; SANZ, 2005a). O DREG estende as operações de consulta e resposta especificadas no protocolo IRIS para atender os requisitos do serviço de informação sobre nomes de domínios oferecido pelos registros de domínios e *registrars*.
- **AREG:** Consiste na especificação e descrição de um *Schema* XML para o protocolo IRIS com o objetivo de prover serviço de informação sobre endereços IP e sistemas autônomos, como especificado na RFC4698 (NEWTON; GUNDUZ; KERR, 2006). O AREG estende as operações de consulta e resposta especificadas no protocolo IRIS para atender os requisitos do serviço de informação sobre endereços IP e sistemas autônomos oferecido pelos registros da Internet.
- **RREG:** Consiste na especificação e descrição de um *Schema* XML para o protocolo IRIS com o objetivo de prover serviço de informação sobre roteamento na Internet, como especificado no *draft* de RFC (NAGAHASHI; YOSHIDA; KONDO, 2005). O RREG estende as operações de consulta e resposta especificadas no protocolo IRIS para atender os requisitos do serviço de informação sobre roteamento na Internet oferecido pelos registros de roteamento.
- **EREG:** Consiste na especificação e descrição de um *Schema* XML para registros ENUM, como especificado na RFC4414 (NEWTON, 2006a). Registros ENUM consistem em nomes de domínios, correspondentes a números de telefone com raiz em e164arpa, como por exemplo 1.2.1.2.5.5.5.3.0.7.1.e164.arpa para o telefone +1-703-555-1212. O EREG estende as operações de consulta e resposta especificadas no protocolo IRIS para atender os requisitos do serviço de informação sobre registros ENUM.
- **BEEP:** O protocolo BEEP consiste em um mecanismo de transporte dentro da camada de aplicação utilizado pelo protocolo IRIS para o transporte de dados, o protocolo BEEP é um protocolo padrão do IETF e foi especificado nas RFCs: RFC3080 (ROSE, 2001b) e RFC3081 (ROSE, 2001a). A interação entre o protocolo BEEP e o protocolo IRIS foi especificada na RFC3983 (NEWTON; SANZ, 2005c). O IRIS utiliza o protocolo BEEP para prover serviço de informação orientado a conexão baseado em TCP sobre *Transport Layer Security* (TLS). A motivação para o uso do protocolo BEEP ao invés de utilizar IRIS diretamente

sobre TCP consiste no requisito do protocolo IRIS de trabalhar com *referrals*. *Referral* consiste em um mecanismo para que um servidor, caso não seja autoritativo sobre a informação consultada, possa responder com uma referência a outro servidor mais indicado para ser o servidor autoritativo, permitindo que o cliente possa refazer a consulta a outro servidor baseado em uma referência dada pelo primeiro servidor. Este mecanismo é transparente ao usuário final.

- **XPC:** O protocolo XPC consiste em um mecanismo de transporte dentro da camada de aplicação utilizado pelo protocolo IRIS para o transporte de dados, o protocolo XPC consiste em uma alternativa ao protocolo BEEP e foi especificado em um draft de RFC (NEWTON, 2007b). Assim como o protocolo BEEP, o protocolo XPC oferece um mecanismo de transporte baseado em TCP sobre TLS. Além disso, o protocolo XPC oferece um mecanismo de *pipelining* que não existe no protocolo BEEP, esse mecanismo é baseado no envio de pequenas unidades (ou blocos) de dados entre aplicações, permitindo que elas possam processar os dados antes da transferência completa. A motivação para o desenvolvimento do protocolo XPC consistiu na dificuldade encontrada por desenvolvedores na implementação do protocolo BEEP.
- **LWZ:** O protocolo LWZ consiste em um mecanismo de transporte dentro da camada de aplicação utilizado pelo protocolo IRIS para o transporte de dados, o protocolo LWZ consiste em uma alternativa de transporte não orientado a conexão e foi especificado em um draft de RFC (NEWTON, 2007a). O protocolo LWZ consiste em um mecanismo de transporte não orientado a conexão baseado em *User Datagram Protocol* (UDP). Este protocolo usa um único pacote para cada consulta ou resposta, e opcionalmente emprega um mecanismo de compressão para permitir que uma quantidade maior de dados seja transmitida em cada pacote. É muito importante notar que esse mecanismo de transporte não suporta nenhum mecanismo de segurança, que os mecanismos de transporte orientados a conexão (BEEP e XPC) suportam, como autenticação e autorização baseadas em TLS.

5.2.1 TRANSPORTE

A proposta inicial originalmente especificava o protocolo BEEP para transferência de dados porque é um protocolo robusto, que possui os mecanismos de autenticação e autorização necessários. Entretanto, as tentativas de implementação do mesmo mos-

traram que o excesso de funcionalidades, com destaque para arquitetura de múltiplos canais (desnecessária para o protocolo IRIS), tornavam a implementação uma tarefa árdua e contribuía para o surgimento de erros.

Para solucionar o problema o grupo de trabalho decidiu desenvolver dois novos protocolos de transporte visando facilitar a implementação do protocolo IRIS. Transações simples devem utilizar o protocolo LWZ baseado em UDP. Transações mais complexas, incluindo as transações que requerem autenticação e autorização devem utilizar o protocolo XPC baseado em TCP sobre TLS.

A abordagem do protocolo LWZ surgiu de um estudo sobre as consultas efetuadas nos servidores Whois, que demonstrou que 90% das consultas não-abusivas são buscas sobre um recurso específico, o que poderia ser executado em uma única consulta sobre UDP sem autenticação, utilizando opcionalmente compressão caso a resposta ultrapasse o limite máximo do tamanho de um pacote UDP.

A abordagem do protocolo XPC surgiu da necessidade de prover autenticação e autorização para alguns tipos de consultas. Além disso, o XPC é utilizado para enviar informações em blocos, dessa forma diminui o tempo de resposta para o cliente e diminui a memória alocada no servidor para responder a várias consultas, assim como diminui o período de latência da rede.

5.2.2 ESPECIFICAÇÃO DOS DADOS

Os dois níveis superiores do protocolo foram especificados utilizando XML, XML *namespaces* e *Schemas* XML. O nível denominado IRIS especifica definições comuns a todos os tipos de registros e formaliza a interação entre o cliente e o servidor. No topo da pilha está a camada específica de cada tipo de registro: nomes de domínios, endereços IP e roteamento. Por exemplo, um registro de domínio pode definir uma consulta para retornar todos os domínios com determinado servidor de nomes configurado. Por outro lado, um registro da Internet poderia definir uma consulta para encontrar todas as redes que possuem um determinado contato.

Atualmente, os seguintes registros foram especificados usando IRIS:

- Registros de Nomes de domínios
- Registros de Internet
- Registros ENUM (International Telecommunication Union E.164 telephone num-

ber)

- Registros de Roteamento

Os dois primeiros foram definidos pelo grupo de trabalho CRISP e o terceiro foi definido pelo grupo de trabalho ENUM do próprio IETF. Os estudos sobre registros de roteamento ainda não foram formalizados pelo IETF e encontram-se na forma de *draft* de RFC.

5.3 SISTEMA DE DESCOBERTA DE SERVIÇOS

O mecanismo para descoberta dinâmica de serviços do protocolo IRIS é baseada no mecanismo *Straightforward-Naming Authority Pointer* (S-NAPTR). Este mecanismo é a base para a solução de co-habitação entre o protocolo IRIS e o protocolo Whois, como definido pelo autor do protocolo (NEWTON; DAIGLE, 2003). O S-NAPTR consiste em um mecanismo que utiliza as consultas DNS, especificamente consultas a *records*² DNS dos tipos *Naming Authority Pointer* (NAPTR) e SRV, para divulgar serviços utilizando dois componentes: um rótulo para identificar a aplicação e outro para especificar o protocolo que a aplicação utiliza. Mais informações sobre o sistema S-NAPTR para descoberta de serviços serão descritas na seção seguinte. O mecanismo para descoberta de serviços pode utilizar três abordagens: resolução direta, resolução indireta *botton-up* e resolução indireta *top-down*. As três abordagens foram descritas para dois tipos de registros: nomes de domínios e endereços IP. Outros tipos de recursos estão fora do escopo deste trabalho mas também podem ser mapeados, de acordo com o mesmo mecanismo.

5.3.1 RESOLUÇÃO DIRETA

A resolução direta é a forma mais simples do cliente IRIS encontrar o servidor responsável por determinado recurso e que corresponde ao comportamento típico dos clientes que acessam a Internet.

²IMPORTANTE: Neste contexto o termo *record* será utilizado para designar registros DNS. Este termo será utilizado para distinguir registro DNS e registro da Internet, pois a palavra registro foi utilizada para designar registros da Internet, como nomes de domínios, endereços IP e sistemas autônomos.

Regras:

1. Na situação ideal o cliente possui: (1) IP servidor, (2) a porta e (3) o tipo de serviço (whois ou iris). A consulta é executada no IP obtido e a porta fornecida, utilizando o protocolo específico do serviço.
2. Se o cliente possuir o nome do servidor ao invés do IP, pode obter o IP através de consulta DNS pelos *records* DNS A e AAAA. Obtém o endereço IP e volta ao item 1.
3. Se o cliente não possuir o número da porta utiliza a porta padrão para o serviço. Obtém a porta e volta ao item 1.
4. Se o cliente não possuir o serviço ele tenta um e se falhar tenta o outro. Obtém o serviço e volta ao item 1.
5. Se o único dado fornecido for o recurso a ser consultado e trata-se de um nome de domínio, o cliente executa o mecanismo S-NAPTR para obter: (1) IP servidor, (2) a porta e (3) o tipo de serviço (whois ou iris). O mecanismo NAPTR será descrito em detalhes a seguir. Obtém os dados necessários e retorna ao item 1. Se não foi possível obter os dados necessários para este domínio através do mecanismo S-NAPTR vai para o passo seguinte.
6. Em último caso, se o único dado fornecido for o recurso a ser consultado, sendo que não se trata de um nome de domínio ou então não foi possível obter o resultado esperado na consulta S-NAPTR para esse domínio. O cliente deve resolver o recurso em um nome de domínio, através de resolução indireta (que pode ser *top-down* ou *bottom-up*) como será descrito nas seções seguintes. Então executa o mecanismo S-NAPTR para obter: (1) IP servidor, (2) a porta e (3) o tipo de serviço (whois ou iris). Obtém os dados necessários e retorna ao item 1.

5.3.2 RESOLUÇÃO BOTTOM-UP

Esta é a tentativa de encontrar o servidor autoritativo sobre determinado recurso que parte da referência mais próxima do registrante do que do registro de domínio ou de Internet responsável pelo recurso.

- Resolução para nomes de domínios O primeiro passo para resolução *bottom-up* de nomes de domínios é tentar o processo de resolução direta para o nome do

domínio (ex. `example.com.`). Se não houver sucesso, então a parte do domínio mais a esquerda é removida e o próximo passo é repetido com o restante do domínio (ex. `com.`). Se todos os componentes do domínio forem removidos e nenhum *record* foi encontrado, então é feita uma consulta DNS para tentar encontrar o número IP associado a este nome de domínio, presumindo que é um nome de servidor. Uma vez obtido o endereço IP é feita uma consulta para este endereço utilizando uma porta conhecida para o endereço desejado.

- Resolução para endereços IP: O endereço IP é convertido para o nome de domínio correspondente utilizando a estrutura de resolução DNS reversa. Por exemplo, `200.160.3.98` é convertido para `98.3.160.200.in-addr-arpa`. A resolução direta é executada para o nome de domínio obtido no passo anterior. Se nenhum *record* é encontrado, então o componente mais a esquerda do nome do domínio obtido no primeiro passo é removida, e tenta-se novamente a resolução direta (`3.160.200.in-addr-arpa.`). Se todos os componentes do nome do domínio obtido no primeiro passo foram removidos e todas as tentativas de resolução direta fracassaram, então o endereço IP original, presumindo que é o IP do próprio servidor, é utilizado juntamente com uma porta conhecida para o serviço de informação desejado.

5.3.3 RESOLUÇÃO TOP-DOWN

A resolução *top-down* é a tentativa de encontrar o responsável por determinado recurso que esteja mais próximo do registro de domínios ou de Internet do que do registrante.

- Resolução para nomes de domínios: O nome de domínio é reduzido a seu componente mais a direita, que é sempre o `."` independente de qual nome de domínio seja. A partir do `."` sendo realizada uma busca através de resolução direta. Caso nenhum *record* tenha sido encontrado, então o componente a esquerda do `."` sendo adicionado ao nome do domínio e a busca através de resolução direta é tentada novamente e assim sucessivamente (ex. `."`, depois `com.` e por fim `example.com.`). Se todos os componentes do domínio original estão presentes e nenhum *record* foi encontrado então é feita uma consulta para resolver o nome do domínio original em um endereço IP, consulta pelo *record* A ou AAAA para este domínio, presumindo que o nome do domínio é o nome de um servidor. Uma vez obtido o endereço IP é realizada uma consulta para este endereço

utilizando uma porta conhecida para o serviço desejado.

- Resolução para endereços IP: O endereço IP é convertido para o nome de domínio correspondente utilizando a estrutura de resolução DNS reversa. Por exemplo, 200.160.3.98 é convertido para 98.3.160.200.in-addr-arpa. A resolução direta é executada para o nome de domínio reduzido a seu componente mais à direita, que é sempre "in-addr-arpa." Independente do nome do domínio em questão. Se nenhum *record* é encontrado, então o componente a esquerda do componente mais à direita é adicionado ao nome do domínio e tenta-se novamente a resolução direta e assim sucessivamente (ex. em seqüência 200.in-addr-arpa., depois 160.200.in-addr-arpa., depois 3.160.200.in-addr-arpa. e finalmente 98.3.160.200.in-addr-arpa.). Se todos os componentes do nome do domínio obtido no primeiro passo foram adicionados e todas as tentativas de resolução direta fracassaram, então o endereço IP original é utilizado juntamente com uma porta conhecida para o serviço de informação desejado, presumindo que é o IP do próprio servidor.

5.3.4 MECANISMO S-NAPTR: CONFIGURAÇÃO DO SERVIDOR DNS

Um registro pode fornecer serviço de informação sobre nomes de domínios utilizando um servidor IRIS ou um servidor WHOIS, ou ambos. Neste cabe ao cliente selecionar o servidor que deseja acessar. Exemplo de configuração dos *records* DNS dos tipos NAPTR e SRV na zona DNS:

```
;; Exemplo de records NAPTR para a zona DNS example.com
example.com.
;;      order  pref  flags  service      re      replacement
IN NAPTR 100    10    ""     "XP-DN: whois" ""      _whois._tcp.example.com.
IN NAPTR 100    10    ""     "XP-DN: crispy" ""      _crispy._tcp.example.com.

;; Exemplo de records SRV para o servico de informacao whois
_whois._tcp.example.com
;;      pri     weight  port  target
IN SRV 10      0       43    big-whois.example.com.
IN SRV 20      0       43    small-whois.example.com.

;; Exemplo de records SRV para o servico de informacao IRIS
_crispy._tcp.example.com.
;;      pri     weight  port  target
IN SRV 10      0       34    big-crispy.example.com.
IN SRV 20      0       34    small-crispy.example.com.
```

Ou seja, da mesma forma que um cliente é capaz de descobrir um servidor IRIS ele

pode ser capaz de descobrir um servidor Whois, desde que a configuração da zona DNS esteja correta.

Descrição dos *records* NAPTR para a zona DNS *example.com*:

- **example.com.:** nome da zona DNS.
- **order:** indica a ordem crescente de preferência dos *records* NAPTR.
- **pref:** dentro de uma mesma ordem (*order*) também existe uma ordem crescente de preferência indicada por *pref*.
- **flags:** indica se a consulta é recursiva ou não. Se a flag estiver vazia (recursiva) indica que é preciso refazer a consulta DNS buscando novamente um *record* tipo NAPTR utilizando o nome indicado no campo *replacement*. Se a flag indicar *s* (não-recursiva) quer dizer que a próxima consulta DNS deve buscar um *record* do tipo SRV com o nome indicado no campo *replacement*.
- **IN:** indica a classe do *record* DNS, no caso *IN* indica a classe INTERNET, a outra possibilidade seria a classe *CH* que indicaria a classe caos (CHAOS).
- **NAPTR:** indica o tipo do *record* DNS.
- **service:** indica o tipo de serviço. No caso *XP* é uma abreviação para IRIS (ou CRISP) e *DN* uma abreviação para nome de domínio (*domain name*) e *whois* indica o protocolo Whois.
- **re:** uma expressão regular que é padrão do *record* NAPTR utilizada para outros propósitos mas é sempre vazia para nomes de domínios de acordo com a RFC3958.
- **replacement:** o nome que deve ser utilizado na próxima consulta DNS.

Descrição dos *records* SRV para o nome *_whois._tcp.example.com*:

- **_whois._tcp.example.com:** nome indicado no *record* NAPTR para a consulta DNS de *record* do tipo SRV.
- **pri:** indica a ordem crescente de prioridade a ser utilizado quando houver mais de um *record* SRV para o mesmo nome, como é o caso do exemplo.
- **weight:** indica a ordem crescente de preferência para *records* SRV com o mesmo nome e o mesmo valor para *pref*.
- **port:** indica a porta UDP/ TCP em que serviço está disponível.

- **target:** o nome da máquina onde o serviço foi disponibilizado.

A mesma descrição é válida para os *records* SRV para o nome *_crispy._tcp.example.com*.

5.3.5 MECANISMO S-NAPTR: PROCEDIMENTO CLIENTE PARA DESCOBERTA DE SERVIÇO

O procedimento do cliente para descobrir o serviço de informação disponível sobre determinado recurso é o seguinte:

1. O primeiro passo consiste em resolver o recurso a ser consultado em um nome que possa ser utilizado para consultas DNS buscando por *records* DNS do tipo NAPTR e SRV.
2. O cliente faz uma consulta DNS buscando todos os *records* DNS do tipo NAPTR para um determinado nome de domínio.
3. Para o uso S-NAPTR do *record* NAPTR o campo *re* relativo a expressão regular é sempre vazio. Então a primeira etapa do processo de seleção elimina todos os *records* NAPTR que não possuam o campo *re* vazio.
4. Para os demais *records* NAPTR o campo deve selecionar o *record* de mais baixa ordem com o tipo de serviço desejado (exemplo: tipo de serviço "XP-DN"). Uma flag vazia indica que a próxima etapa da busca consiste em uma consulta DNS por um *record* do tipo NAPTR utilizando o nome indicado no campo *replacement*.
5. Em seguida o cliente faz uma nova consulta DNS pelo *record* NAPTR com o nome fornecido pelo *record* NAPTR anterior e busca pelo resultado que corresponda ao tipo de serviço e protocolo desejado (por exemplo: "XP-DN: crispy").
6. O próximo passo é buscar o *record* NAPTR com a flag igual a "s", que indica que o próximo *record* DNS a ser consultado é do tipo SRV com nome igual ao fornecido no campo *replacement*. O cliente refaz a consulta DNS para o *record* SRV e seleciona o *record* SRV mais apropriado de acordo com os campos de prioridade (*priority*) e peso (*weight*) e usa o nome do servidor fornecido como entrada para a próxima consulta DNS para buscar o *record* do tipo A, ou seja, resolver o nome do servidor em um endereço IP.
7. Uma vez que o cliente obteve o número IP e porta do serviço desejado, ele faz a consulta do recurso (ex. nome de domínio ou endereço IP). Se a conexão for

recusada o cliente deve retornar ao passo anterior e selecionar outro servidor. Se todos os servidores fornecidos por *records* SRV do passo anterior tiverem sido testados sem sucesso então a aplicação cliente deve retornar um erro ao usuário.

5.4 SOLUÇÃO DE CO-HABITAÇÃO COM O PROTOCOLO WHOIS

A utilização do mecanismo S-NAPTR não se resume a encontrar o servidor para determinado tipo de serviço e protocolo. Ele possibilita encontrar o servidor autoritativo sobre determinado recurso, algo que não era possível no sistema anterior. Existem três formas diferentes de fazer esta consulta:

Atualmente, muitos registros da Internet oferecem o serviço Whois que será substituído por um novo serviço mais completo baseado no protocolo IRIS. Entretanto, será necessário oferecer o serviço de informação utilizando os dois protocolos durante uma fase de adaptação, que poderia ser chamada de fase de co-habitação dos dois protocolos.

Uma possível abordagem³ proposta pelo autor do protocolo IRIS, Andrew Newton, consiste em utilizar S-NAPTR para prover uma forma única para clientes IRIS encontrar e negociar o protocolo que eles são capazes de utilizar. O S-NAPTR é um dos mecanismos padrão para descoberta de serviço do protocolo IRIS e pode ser utilizado para prover uma solução de co-habitação dos dois protocolos. Se o cliente IRIS não obtiver o resultado da consulta através do protocolo IRIS pode obter através do protocolo Whois.

O S-NAPTR é um mecanismo para descoberta de serviços baseado em consultas DNS e não requer modificação no protocolo Whois ou no protocolo IRIS. Clientes sem suporte a S-NAPTR continuarão a se comunicar com os servidores Whois e os servidores Whois também não requerem modificação para responder a clientes com ou sem suporte a S-NAPTR. Os clientes e servidores IRIS também não precisam ser alterados antes, durante ou depois do período de co-habitação. A seguir uma descrição detalhada do funcionamento do mecanismo S-NAPTR e, conseqüentemente, do funcionamento da solução de co-habitação entre os dois protocolos.

Outros dois possíveis casos de uso para o mecanismo S-NAPTR consistem em: (1) permitir que um registro ofereça serviço de informação sobre diferentes tipos de regis-

³<http://tools.ietf.org/html/draft-newton-whois-crisp-cohabitation-00>

tros (como nomes de domínios e endereços IP) ou (2) permitir a delegação do serviço de informação a terceiros.

5.4.1 CASO DE USO S-NAPTR: SERVIÇO DE INFORMAÇÃO SOBRE DOMÍNIOS E IPS

Um registro pode oferecer serviço de informação sobre endereços IP e nomes de domínios, ele pode optar por configurar dois servidores, um para fornecer informações sobre IPs e outro para nomes de domínios. Exemplo da configuração dos *records* na zona DNS:

```
;; Exemplo de records NAPTR adicionais para delegar o servico de
;; informacao sobre enderecos IP para outra zona DNS:
example.com.
;;      order  pref  flags  service          re  replacement
IN NAPTR 100    10    ""     "XP-DN: whois: crispy" ""   example.com.
IN NAPTR 100    10    ""     "XP-IP: whois: crispy" ""   somebodyelse.com
```

5.4.2 CASO DE USO S-NAPTR: DELEGAR SERVIÇO DE INFORMAÇÃO A TERCEIROS

Ou um registro pode fornecer serviço de informação sobre IPs utilizando o protocolo Whois e delegar a função de informação utilizando o protocolo IRIS para outro servidor em outra zona DNS. Como exemplo de configuração da zona DNS bastaria substituir o seguinte *record* NAPTR:

```
example.com.
;;      order  pref  flags  service          re  replacement
IN NAPTR 100    10    ""     "XP-IP: whois: crispy" ""   somebodyelse.com
```

Por:

```
example.com.
;;      order  pref  flags  service          re  replacement
IN NAPTR 100    10    ""     "XP-IP: whois: crispy" ""   example.com.
IN NAPTR 100    10    ""     "XP-IP: whois"      ""   _whois._tcp.ex.com.
IN NAPTR 100    10    ""     "XP-IP: crispy"     ""   _crispy._tcp.other.com.
```

5.5 CONCLUSÕES

O grupo de trabalho do IETF teve longas discussões, via listas de discussão ou em encontros do IETF, para tentar produzir uma especificação capaz de atender o maior número de políticas locais possível sem perder a generalidade. As discussões foram importantes para determinar os requisitos genéricos do protocolo, ou seja, requisitos comuns a todos os provedores de serviço de informação sobre nomes de domínios, cuja especificação serviu de base para o desenvolvimento do modelo RDF neste estudo. Entretanto na fase de especificação do protocolo IRIS teria sido mais adequado e vantajoso utilizar RDF ao invés de XML puro como a base para garantir a interoperabilidade entre as aplicações. Além disso, o protocolo IRIS poderia contar com uma *framework* com facilidades de extensão como RDF. Assim cada servidor poderia utilizar uma extensão própria desde que respeitasse o modelo RDF original e todas as aplicações desenvolvidas para o modelo original não precisariam ser atualizadas para suportar novas extensões.

6 WEB SEMÂNTICA: XML VERSUS RDF

6.1 INTRODUÇÃO

Desde o princípio, o foco da Internet tem sido prover conteúdo em linguagem natural, ou seja, conteúdo para seres humanos. Entretanto a próxima geração da Internet, que Tim Berners-Lee chama de Web Semântica, está sendo desenvolvida para prover conteúdo em linguagem natural, assim como para viabilizar o processamento automático de dados. Com a mudança de foco, a Web Semântica tem o intuito de possibilitar o desenvolvimento de agentes capazes de processar informação de forma automatizada e, eventualmente, agir com base no conhecimento do domínio e na informação processada. O termo agente será utilizado como definido em (FRANKLIN; GRAESSER, 1996), podendo referir-se tanto a uma aplicação capaz de apenas processar informação de forma automatizada, quanto a um agente inteligente capaz de aprender, adaptar-se e inferir com base na informação processada.

Web Semântica é o termo utilizado para designar um conjunto de tecnologias que viabilizam o compartilhamento e processamento automático de informação entre aplicações na Internet. A Internet, como a conhecemos hoje, somente tornou-se possível devido um amplo conjunto de padrões desenvolvido para garantir interoperabilidade em vários níveis. Da mesma forma, a evolução da Internet somente será possível quando novos padrões permitirem interoperabilidade baseada no conteúdo, ao invés de interoperabilidade baseada exclusivamente no formato dos documentos (DECKER et al., 2000). Neste sentido, recentes trabalhos de padronização junto ao W3C, como XML, *Schema* XML, RDF e *Schema* RDF, formam a base para o desenvolvimento da Web Semântica. XML e RDF são atualmente os padrões para o estabelecimento da interoperabilidade semântica na Internet. Contudo XML trata apenas da estrutura do documento. Enquanto RDF facilita a interoperação porque provê um modelo de dados que pode ser facilmente estendido para ser utilizado com técnicas sofisticadas de representação semântica.

Tecnologias de Web Semântica podem ser usadas em uma grande variedade de aplicações. Alguns projetos de pesquisa englobam várias aplicações de tecnologia de

Web Semântica, por exemplo o projeto de pesquisa NEPOMUK ¹ (FRANK; DECKER, 2002) cujo objetivo é desenvolver um *desktop* otimizado para processamento automatizado de dados, visando principalmente o compartilhamento e troca de recursos. O projeto é mantido pela União Europeia e possui colaboradores em vários centros de pesquisa europeus. Todas as aplicações a seguir são utilizadas no projeto NEPOMUK:

- Integração de dados, onde dados provenientes de diversas fontes podem ser integrados em uma única aplicação;
- Descoberta e classificação de recursos para prover mecanismos de busca mais eficientes;
- Criação de catálogos para descrever o conteúdo e o relacionamento dos dados provenientes de diversas páginas na Internet ou de bibliotecas digitais;
- Permite que agentes inteligentes possam compartilhar e trocar conhecimento;
- Classificação de mídias digitais.

Para atingir os objetivos descritos, o passo mais importante consiste em definir e descrever o relacionamento entre dados (ou recursos). O relacionamento entre recursos é semelhante ao relacionamento definido por *links* conectando diferentes páginas na Internet. A diferença é que o relacionamento pode ser estabelecido entre quaisquer dois recursos e não existe conceito de página corrente. O relacionamento também possui um identificador único *Uniform Resource Identifier* (URI) e seu papel é definido de forma que tanto um ser humano quanto uma máquina é capaz de compreender. RDF é uma das tecnologias chave para a descrever formalmente o relacionamento entre recursos na Web Semântica.

É importante notar que as tecnologias de Web Semântica existem para agregar funcionalidade à Internet, facilitar compartilhamento de informações e processamento automático de dados. Entretanto, o funcionamento da Web Semântica é transparente para o usuário final, ele apenas obtém o resultado sem necessidade de ter conhecimento sobre o relacionamento entre os dados ou das regras de inferência usadas. No máximo, um usuário deveria ser capaz de definir relacionamento entre dados em um domínio específico, por exemplo medicina, usando uma ferramenta capaz de exportar o modelo para um formato que possa ser processado automaticamente, por exemplo RDF.

Como todas as tecnologias inovadoras, a Web Semântica começou sua evolução em laboratórios de pesquisa, então foi adotada pela comunidade responsável pelo desenvolvimento de padrões abertos, e atualmente, esta sendo adotada por empresas.

¹ <http://www.nepomuk.semanticdesktop.org>

Oracle, IBM, Adobe, Software AG e North Gruman são algumas das grandes empresas que já adotaram a tecnologia. Algumas empresas estão vendendo ferramentas, outras soluções completas de negócios.

A seguir serão descritas as vantagens da utilização da *framework* RDF ao invés de utilizar XML como base de interoperabilidade na troca de dados entre aplicações na Internet. É importante notar que RDF consiste na representação de um modelo de dados que pode ser serializado em formato XML ou em outro formato, como *Notation 3* (N3) (BERNERS-LEE,) e *Triples in XML* (TriX) (CARROLL; STICKLER,). Entretanto, como será mostrado a seguir a obtenção do modelo de dados utilizando a *framework* RDF é independente da sintaxe ou formato do documento trocado entre aplicações. Enquanto, a obtenção de um modelo de dados a partir de XML puro é dependente da sintaxe do documento trocado e requer atualização das aplicações a cada alteração de sintaxe.

6.2 ONTOLOGIA

Uma ontologia consiste na conceitualização formal de um domínio específico baseada em consenso. O objetivo de desenvolver ontologias é viabilizar o entendimento de informação que possa ser compartilhada entre pessoas e sistemas de aplicação. Além disso, o desenvolvimento de ontologias tem um papel crucial para viabilizar processamento, compartilhamento e reutilização de conhecimento entre aplicações na Internet. Ontologias são utilizadas em comércio eletrônico para promover a comunicação automatizada entre vendedores e compradores, além de permitir o reuso de descrições em diversos mercados. Ferramentas de busca usam ontologias para encontrar páginas contendo palavras-chave que são sintaticamente diferentes mas semanticamente similares.

Uma ontologia tipicamente contém uma hierarquia de conceitos (ou classes) em um domínio específico e a descrição das propriedades (ou atributos) relevantes através da associação de pares atributo-valor a cada conceito. Além disso, uma ontologia pode conter relacionamentos entre conceitos e regras lógicas de inferência. Exemplo simples: *Mamífero* é uma sub-classe da classe *Animal*; *Pessoa* é uma sub-classe da classe *Mamífero*; das duas regras lógicas pode-se inferir que *Pessoa* é uma sub-classe de *Animal*. Apesar do exemplo ser bastante simples, existem muitas regras lógicas sofisticadas que poderiam ser utilizadas. Entretanto o estudo detalhado de regras lógicas de inferência está fora do escopo deste trabalho.

Exemplo de ontologia utilizando a linguagem de representação de ontologias *Ontology Inference Layer* (OIL) (FENSEL et al., 2001):

```
class-def animal                // classe animal
class-def planta                // classe planta
  subclass-of NOT animal
class-def arvore                // classe arvore
  subclass-of planta
class-def galho
  slot-constraint is-part-of // galho faz parte de arvore
    has-value arvore
class-def folha
  slot-constraint is-part-of // folha faz parte de galho
    has-value galho
```

XML e RDF formam a fundação para a Web Semântica. Entretanto, RDF provê mecanismos mais apropriados para aplicar linguagens de representação de ontologias, como OIL. RDF é mais apropriado, principalmente, porque permite a representação de ontologias garantindo a interoperabilidade entre sistemas, como será descrito a seguir.

6.3 XML

XML foi padronizado pelo W3C e consiste em um formato texto simples e flexível derivado do SGML ², que por sua vez foi definido pela norma ISO 8879. Originalmente foi desenvolvido para provisionamento de publicação eletrônica em larga escala. Atualmente, XML vem sendo utilizado para a troca de dados em uma grande variedade de aplicações na Internet.

XML foi desenvolvido para prover formatação em documentos de estrutura arbitrária. Enquanto HTML foi desenvolvido para prover formatação em documentos com estrutura fixa. Um documento XML bem-formatado corresponde a uma estrutura de dados em forma de árvore, onde cada elemento pode conter um número ilimitado de pares atributo-valor.

Não existe um conjunto padrão de elementos ou atributos válidos para um documento XML. Sendo assim, a validade de um documento XML bem-formatado depende da aplicação. A validação de documentos XML em aplicações é realizada através de *Document Type Definition* (DTD) (BRAY et al.,), também padronizado pelo W3C. DTD

²<http://www.w3.org/MarkUp/SGML/>

permite definir a gramática a ser utilizada em determinada aplicação, os elementos e atributos permitidos e suas possíveis combinações. Recentemente, a validação usando DTD vem sendo substituída por *Schemas XML* (FALLSIDE; WALMSLEY,), outro padrão do W3C. Apesar de oferecer diversas vantagens em relação ao DTD, *Schemas XML* tem essencialmente o mesmo papel: definir a gramática para documentos XML.

Exemplo de serialização de uma parte da ontologia descrita anteriormente utilizando XML:

```
<class-def>
  <class name="planta"/>
  <subclass-of>
    <NOT><class name="animal"/></NOT>
  </subclass-of>
</class-def>
<class-def>
  <class name="arvore">
  <subclass-of>
    <class name="planta"/>
  </subclass-of>
</class-def>
<class-def>
  <class name="galho">
  <slot-constraint>
    <slot name="is-part-of"/>
    <has-value>
      <class name="arvore">
    </has-value>
  </slot-constraint>
</class-def>
```

É importante notar que o exemplo é apenas uma das várias possíveis sintaxes baseada em XML para a ontologia. XML é acima de tudo uma linguagem para definir gramáticas, porque gramáticas diferentes podem ser usadas para descrever o mesmo conteúdo. A mesma informação poderia ser expressada de forma bem diferente usando XML. Por exemplo, a classe *galho* poderia ser definida da seguinte forma:

```
<class-def>
  <name>galho</name>
  <slot-constraint>
    <name>is-part-of</name>
    <has-value>arvore</has-value>
  </slot-constraint>
</class-def>
```

XML pode ser utilizada para diversos propósitos:

1. Sintaxe de serialização para outras linguagens. Por exemplo, SMIL consiste basicamente de um DTD que define a estrutura de um documento SMIL, facilitando o entendimento do significado dos elementos em cada documento XML.
2. Formatação para páginas na Internet. Um documento XML pode ser usado em associação com *Extensible Stylesheet Language (XSL)* (BERGLUND,) para apresentar elementos de forma diferenciada de acordo com o estilo.
3. Formato uniforme para troca de dados. Um documento XML pode ser transferido como objeto de dados entre duas aplicações.

XML, DTD e *Schema XML* tem por objetivo definir exclusivamente convenções de sintaxe, ignorando qualquer conceito de semântica associado ao documento. Como foi visto no exemplo, é possível expressar uma mesma ontologia em XML utilizando diversas sintaxes. Sendo que os *parsers XML*, baseados na API *Document Object Model (DOM)* (APPARAO et al.,) ou *Simple API for XML (SAX)* (MEGGINSON,), fornecem uma estrutura de dados que corresponde exatamente a sintaxe do documento XML. Entretanto, toda vez que a sintaxe do documento muda, a aplicação que processa os documentos XML precisa ser atualizada porque a estrutura de dados é alterada.

6.4 RDF

RDF (DECKER; MITRA; MELNIK, 2000) é uma recente recomendação do W3C (KLYNE; CARROLL; MCBRIDE, b) (KLYNE; CARROLL; MCBRIDE, a) e consiste em uma linguagem para representar informação na Internet. RDF foi inicialmente desenvolvida para padronizar a definição e uso de meta-informação sobre recursos. Entretanto, ela pode ser usada de forma geral para representar dados e como tecnologia base para garantir interoperabilidade na troca de dados. RDF tem características que facilitam a integração de dados mesmo quando eles possuem diferentes esquemas de representação. Em particular, ela suporta a evolução dos esquemas de representação com o tempo sem necessidade de atualização das aplicações.

RDF é uma linguagem que busca aplicar os conceitos da Internet para representar recursos. Para atingir esse objetivo, RDF estende a estrutura de interligação da Internet para utilizar URIs na identificação de relacionamentos entre recursos, assim como para identificar os próprios recursos, formando triplas sujeito-predicado-objeto

(ou objeto-atributo-valor). A estrutura resultante consiste em um grafo, onde os nós representam recursos e os arcos representam relacionamentos.

RDF pode ser representado através de:

- um conjunto de triplas (sujeito-predicado-verbo); ou
- um grafo; ou
- um documento em XML, N3, TriX ou outras linguagens.

Usando este modelo simples, ele permite que dados estruturados ou semi-estruturados sejam integrados, compartilhados e trocados entre diversas aplicações, independente da sintaxe do documento.

Assim como *Schema XML*, *Schema RDF* permite a definição de vocabulário de dados. Em outras palavras, ele provê um sistema de tipos de dados para modelos de dados RDF. O sistema de tipos possui alguns tipos pré-definidos, como *Class*, *subPropertyOf* e *subClassOf*. Um objeto RDF pode ser definido como instância de uma ou mais classes. A propriedade *subClassOf* permite a criação de hierarquias de classes, assim como a propriedade *subPropertyOf* faz o mesmo para propriedades. *RDFs:domain* e *RDFs:range* são restrições impostas em propriedades. Assim, é possível traduzir uma linguagem de representação de ontologias para RDF.

RDF e *Schema RDF* permitem que uma ontologia seja mapeada em um modelo de dados completamente independente da sintaxe do documento utilizado para a troca de dados. Os *parsers* de documentos RDF processam o documento independente da linguagem (XML, N3, TriG, N-Triple, etc) e fornecem como resultado uma estrutura de dados que corresponde ao modelo de dados e é completamente independente da sintaxe do documento. Esta característica facilita a atualização e manutenção das aplicações.

6.5 REPRESENTAÇÃO DE CONHECIMENTO

A Internet é um meio único para troca de dados e possui requisitos próprios:

- Universalidade. Um formato para troca de dados na Internet deve ser capaz de expressar dados provenientes de qualquer outro formato.

- Interoperabilidade sintática. Aplicações devem ser capazes de processar dados independente da plataforma ou tecnologia utilizada.
- Interoperabilidade semântica. O mais importante requisito é a interoperabilidade semântica, ou seja, a aplicação deve ser capaz de processar dados e analisar o conteúdo. O objetivo da análise pode ser a integração dos dados processados com dados provenientes de outras fontes de dados; ou inferir informação com base nos dados processados; ou compartilhar os dados.

6.5.1 XML

XML satisfaz os requisitos de universalidade e interoperabilidade sintática mas não satisfaz o requisito de interoperabilidade semântica. Com relação a universalidade, qualquer dado para o qual possa ser criada uma gramática pode ser expresso em formato XML. No que se refere à interoperabilidade sintática, programas são capazes de processar XML independente da aplicação utilizando bibliotecas existentes, por exemplo *Xerces*³ (Java), *Xerces-C*⁴ (C++) e *expat*⁵ (C). Quanto a interoperabilidade semântica: não é possível reconhecer uma unidade semântica porque XML não impõe nenhuma regra interpretação dos dados contidos em um documento.

Por exemplo existem duas aplicações tentando comunicar entre si. Primeiro, ambas as partes concordam em um modelo do domínio em termos de objetos e relações, por exemplo utilizando um modelo entidade-relacionamento. Então, ambos precisam concordar em uma sintaxe de documento XML definida utilizando DTD que represente do modelo do domínio, porque existem inúmeras maneiras de mapear um modelo de domínio em formato XML. Finalmente, as duas aplicações podem comunicar entre si. É importante notar que modelo de um domínio pode parecer bastante semelhante ao modelo de ontologia a princípio, embora modelar uma ontologia possa envolver a definição de regras de inferência e sofisticadas técnicas de representação de conhecimento.

A solução pode funcionar bem para duas aplicações tentando comunicar entre si. Entretanto, a principal desvantagem desta solução consiste em se tratar de uma comunicação um-a-um entre partes com um acordo prévio sobre o significado da informação trocada em formato XML. Em outras palavras, esta solução negligencia a realidade da

³<http://xerces.apache.org/xerces-j/>

⁴<http://xml.apache.org/xerces-c/>

⁵<http://expat.sourceforge.net/>

Internet porque ela requer comunicação entre múltiplas partes que mudam frequentemente, algumas partes são adicionadas e outras são removidas. Por essa razão, qualquer solução deveria ser flexível suficiente para se adaptar constantemente.

Neste sentido, o custo da adição de um novo parceiro no sistema de comunicação pode ser grande. Para trocar documentos XML, os mapeamentos de domínios em formato XML devem ser traduzidos utilizando transformações XSL (CLARK,). Então, é preciso reestruturar o modelo do domínio para que possa ser consistente com o DTD resultante. Finalmente, é preciso regerar os procedimentos utilizados para processar o conteúdo do documento XML baseado na nova estrutura.

6.5.2 RDF

RDF satisfaz os três requisitos: universalidade, interoperabilidade sintática e semântica. No que se refere a universalidade, a estrutura RDF é baseada em triplas (objeto-atributo-valor ou sujeito-predicado-objeto) que possui poder de expressão universal, sendo que RDF pode ser codificado em diversos formatos: XML, N-Triple, Turtle, TriG, TriX, etc. Quanto a interoperabilidade sintática, existem bibliotecas capazes de processar RDF independente da aplicação e do formato, da mesma forma que existem bibliotecas para processar XML, por exemplo *Jena*⁶ (Java), *Sesame*⁷ e *HP ARP*⁸. Em relação a interoperabilidade semântica, RDF tem diversas vantagens em relação ao XML, como será descrito a seguir.

Para interoperabilidade semântica, a estrutura RDF (baseada em triplas sujeito-predicado-objeto) provê naturalmente unidades semânticas porque todos os recursos (sujeito, objeto ou predicado) são entidades independentes. O modelo do domínio, que define entidades e relacionamentos entre entidades, pode ser representado naturalmente em RDF, por isso não são necessários passos de tradução como em XML. Além disso, para integrar dois modelos RDF e possível aplicar diretamente técnicas de representação de ontologias, que em sua maioria são técnicas desenvolvidas na área de inteligência artificial. Da mesma forma, as aplicações desenvolvidas com base no modelo RDF, não requerem atualizações quando ocorrem mudanças na sintaxe do documento utilizado para troca de dados, porque RDF é independente de sintaxe.

RDF tem alta flexibilidade e extensibilidade. A linguagem é altamente flexível porque

⁶<http://jena.sourceforge.net/>

⁷<http://www.openrdf.org/>

⁸<http://www-uk.hpl.hp.com/people/jjc/arp/>

possui tecnologia de representação de conhecimento suficiente para expressar modelos originalmente escritos em diversas outras linguagens. A linguagem é extensível porque, apesar de sua estrutura ser bastante simples, ela pode ser estendida para representar modelos de conhecimento mais complexos. Por exemplo, a linguagem *Web Ontology Language* (OWL) foi desenvolvida como uma extensão de RDF para facilitar processamento automático de dados baseado em regras de inferência.

6.6 FUTURO DA WEB SEMÂNTICA

Atualmente, a comunidade confia no uso de XML para troca de dados e integração semântica. Devido as limitações descritas anteriormente, a troca de dados em formato XML puro não tem escalabilidade para integração semântica. A solução funciona bem quando existem apenas duas aplicações tentando comunicar entre si e tem um acordo prévio sobre o modelo do domínio e o formato dos dados. Entretanto, este modelo não pode ser aplicado a Internet pois o ambiente é completamente dinâmico, o que requer flexibilidade, interoperabilidade e extensibilidade para atender requisitos futuros. RDF atende esses requisitos e seu uso sistemático poderia viabilizar a Web Semântica num futuro próximo.

Algumas organizações responsáveis pela padronização na Internet são extremamente conservadoras no que se refere a adoção de novas tecnologias. O conservadorismo é importante para impedir que novos padrões sejam baseados em tecnologias com futuro incerto, o que certamente afetaria a credibilidade do processo de padronização. Entretanto, conservadorismo tem a grande desvantagem de negligenciar tecnologias novas e promissoras. O IETF apóia o uso de XML e *Schema* XML como formato para troca de dados e negligencia a existência de RDF e *Schema* RDF que poderiam contribuir para a padronização de protocolos mais robustos. Por essa razão, o protocolo IRIS é baseado em XML ao invés de RDF. Por outro lado, o W3C não é tão conservador e tem alta credibilidade, no caso tanto XML quanto RDF são tecnologias recomendadas pelo W3C.

7 INTEGRAÇÃO DE TECNOLOGIAS DE WEB SEMÂNTICA AO PROTOCOLO IRIS

7.1 INTRODUÇÃO

O desenvolvimento de um conjunto de padrões para garantir a interoperabilidade em um ambiente altamente heterogêneo foi um fator determinante para o surgimento da Internet. Da mesma forma, a Web Semântica depende do surgimento de novos padrões para estabelecer interoperabilidade semântica. Este fato, assim como novas tecnologias em geral, devem ser considerados durante o processo de desenvolvimento de novos padrões. O protocolo IRIS é um novo padrão em fase de desenvolvimento pelo IETF e cabe a comunidade acadêmica contribuir durante este processo. A própria Internet surgiu no meio acadêmico, por isso a contribuição acadêmica para sua constante atualização é mais do que esperada.

Entretanto existe uma grande barreira a ser vencida para a adoção do novo protocolo. Houve várias tentativas de adoção de novos padrões para substituir o protocolo Whois e nenhuma delas foi efetivada. Os provedores e usuários do serviço Whois estão acostumados ao antigo protocolo e impõe resistência para mudar sua rotina. A motivação inicial para o desenvolvimento do novo protocolo foi a preocupação com a segurança dos dados armazenados na base de dados Whois. Entretanto segurança não é um requisito funcional e é difícil de ser medido ou demonstrado. Usuários de forma geral só se preocupam com segurança após sofrer algum incidente desagradável. Em outras palavras, é difícil convencer usuários que o principal motivo que levou especialistas a desenvolverem um novo padrão justifica a mudança em sua rotina.

Além disso, a preocupação com segurança pode ter sido a motivação inicial para o desenvolvimento do novo protocolo mas é preciso considerar o impacto do novo protocolo no futuro. Dada a grande barreira para a adoção de novas tecnologias e o grande esforço que o desenvolvimento de um novo padrão demanda, espera-se que o resultado seja ótimo a curto, médio e longo prazo. Novas pesquisas apontam a tendência de crescimento da Web Semântica e a integração de tecnologias de Web Semântica ao novo protocolo daria força a ambos. Se o novo protocolo for adotado com as novas tecnologias, pode facilitar a disseminação da Web Semântica. Por outro lado, o sucesso da Web Semântica pode levar o novo padrão a obsolescência, se ele

não puder adaptar-se a nova realidade.

O protocolo IRIS foi especificado usando *Schemas XML*. Como foi descrito anteriormente, o uso de XML como tecnologia base para garantir interoperabilidade na troca de dados entre aplicações pode ser uma tendência atual mas não tem escalabilidade para integrar diversas aplicações ou fontes de dados na Internet. Ao invés de usar XML, o protocolo IRIS poderia ser facilmente modificado para utilizar RDF enquanto ainda está em fase de desenvolvimento. A mudança seria pequena, sem necessidade de introduzir sofisticadas técnicas de representação de ontologias, mas permitindo que agentes inteligentes possam utilizar dados fornecidos pelo serviço de informação sobre registros da Internet no futuro. Independente de ampliar as possibilidades de aplicação no futuro, a mudança de paradigma de XML puro para RDF tem um impacto prático e imediato.

IMPORTANTE: Neste contexto o termo *swhois* será utilizado para designar o serviço de informação sobre registros da Internet utilizando o protocolo IRIS com algumas modificações para integrar tecnologias de Web Semântica.

7.2 MODELO DE ONTOLOGIA

Um modelo RDF pode ser obtido a partir de um modelo UML, ou de um modelo entidade-relacionamento. Tendo como base a especificação do protocolo IRIS, tem-se uma descrição clara e objetiva do domínio, baseada em consenso obtido durante anos em listas de discussão e, principalmente, em encontros do IETF. Os participantes, em geral, são membros de registros da Internet, de *registrars* ou ISPs, embora as discussões sejam abertas a todos os interessados. O modelo entidade-relacionamento baseado na especificação do protocolo IRIS é bem simples, como pode ser visto na Figura 2. Possui apenas três entidades e os relacionamentos entre elas: domínio, contato e servidor de nomes. A complexidade do modelo consiste na extensa série de atributos que cada uma das três entidades possui e nos relacionamentos entre elas.

O modelo RDF se concentra no aspecto descritivo do domínio e cabe a cada aplicação tratar o aspecto funcional. Sendo assim, quando o objetivo é extrair o modelo RDF do modelo entidade-relacionamento, o foco deve ser na definição de entidades e atributos. Por outro lado, os métodos não são relevantes, pois cabe a cada aplicação tratar o aspecto funcional. É importante notar que o escopo do modelo foi limitado a nomes de domínios, outros recursos como endereços IP e sistemas autônomos foram excluídos do modelo, embora o modelo possa ser estendido para englobar estes e

outros recursos.

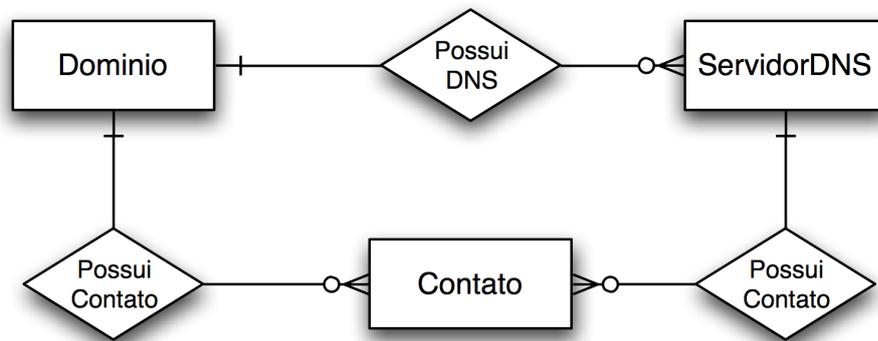


Figura 2: Diagrama entidade-relacionamento

O modelo RDF completo possui a definição das classes e das propriedades de cada uma delas. A definição de classes (Figura 3) é bastante semelhante ao modelo entidade-relacionamento (Figura 2). Entretanto, o modelo RDF na forma de grafo da Figura 3, é uma versão bastante simplificada, pois todas as propriedades foram omitidas para fins didáticos. Versões simplificadas de grafos RDF são comuns pois grafos completos são complexos e difíceis de entender. O termo técnico para uma visão simplificada de um grafo RDF é *Graph View*.

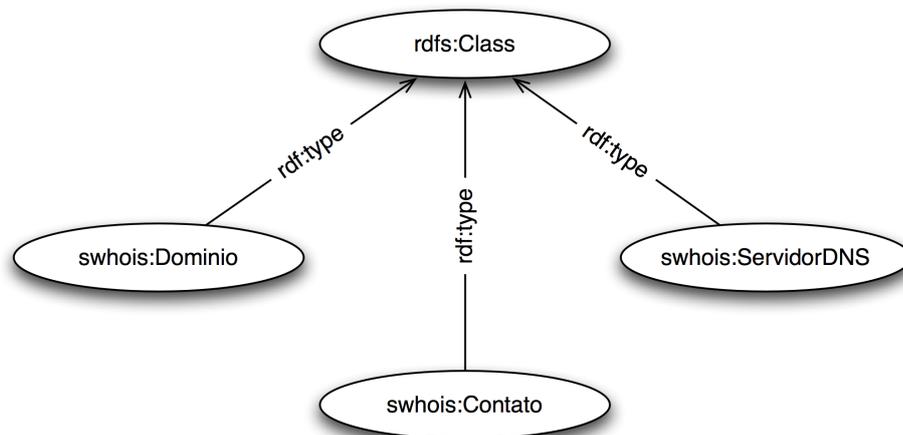


Figura 3: Modelo RDF em formato de Grafo simplificado: apenas definição de classes.

Tabela 3: Correspondência entre os nomes e suas respectivas URIs para a Figura 3

Nome	URI correspondente
swhois:Domínio	http://sw.deri.ie/swhois#Domínio
swhois:Contato	http://sw.deri.ie/swhois#Contato
swhois:ServidorDNS	http://sw.deri.ie/swhois#ServidorDNS
rdfs:Class	http://www.w3.org/2000/01/rdf-schema#Class
rdf:type	http://www.w3.org/1999/02/22-rdf-syntax-ns#type

Analisando a Figura 3 em detalhe, é possível observar os identificadores das classes: *swhois:Domínio*, *swhois:Contato* e *swhois:ServidorDNS*. O termo *swhois* utilizado corresponde a abreviação do espaço de nomes utilizado. A correspondência entre os identificadores das classes em sua versão expandida e abreviada são mostrados na Tabela 3. Todas as classes e propriedades são identificadas por URIs. O fato de utilizar um identificador único para todos os recursos é uma característica própria de RDF e da Web Semântica. Entretanto, utilizar URIs em algumas situações pode dificultar a visualização e compreensão por humanos, por isso é comum o uso de URIs em formato abreviado.

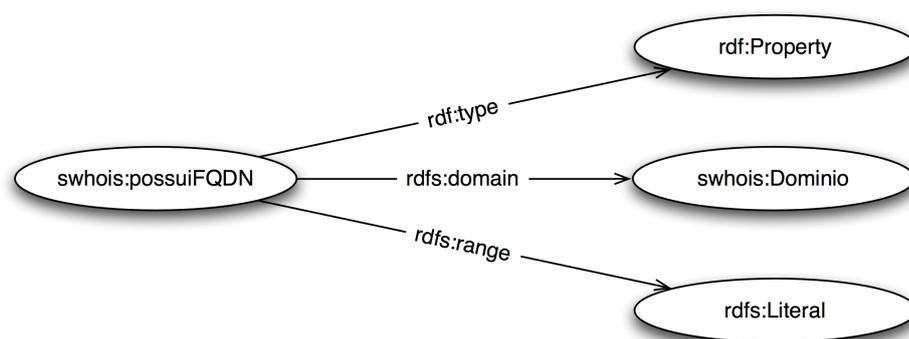


Figura 4: Modelo RDF em formato de Grafo simplificado: definição de uma das propriedades da classe *swhois:Domínio*.

Tabela 4: Correspondência entre os nomes e suas respectivas URIs para a Figura 4

Nome	URI correspondente
<code>swhois:possuiFQDN</code>	<code>http://sw.deri.ie/swhois#possuiFQDN</code>
<code>rdf:type</code>	<code>http://www.w3.org/1999/02/22-rdf-syntax-ns#type</code>
<code>rdf:Property</code>	<code>http://www.w3.org/1999/02/22-rdf-syntax-ns#Property</code>
<code>rdfs:domain</code>	<code>http://www.w3.org/2000/01/rdf-schema#domain</code>
<code>swhois:Domínio</code>	<code>http://sw.deri.ie/swhois#Domínio</code>
<code>rdfs:range</code>	<code>http://www.w3.org/2000/01/rdf-schema#range</code>
<code>rdfs:Literal</code>	<code>http://www.w3.org/2000/01/rdf-schema#Literal</code>

A Figura 4 mostra o grafo RDF correspondente a definição de uma propriedade da classe domínio, a propriedade *swhois:possuiFQDN*. Todas as propriedades da classe domínio possuem três triplas. A primeira tripla define o tipo do recurso (identificado por URI) como sendo uma propriedade. A segunda tripla especifica que este recurso é uma propriedade de domínio (*rdfs:domain*). A terceira tripla define os valores possíveis (*rdfs:range*) para esta propriedade. Sendo assim, todas as propriedades da classe *swhois:Domínio* tem em comum duas triplas:

- ($\langle URI \rangle, \langle rdf : type \rangle, \langle rdf : Property \rangle$)
- ($\langle URI \rangle, \langle rdfs : domain \rangle, \langle swhois : Domínio \rangle$)

O que difere as propriedades da classe *swhois:Domínio* são as triplas:

- ($\langle URI \rangle \langle rdfs : range \rangle \langle TipoEspecifico \rangle$)

O mesmo raciocínio pode ser aplicado às propriedades das classes *swhois:Contato* e *swhois:ServidorDNS*. Para mais detalhes sobre as propriedades das classes que não puderam ser expressadas na forma de grafo, elas estão listadas em três tabelas juntamente com a sua descrição no apêndice. Também no apêndice, pode ser encontrada a serialização do modelo RDF completo serializado nas linguagens N3 e XML. O modelo foi obtido através da utilização da ferramenta Protégé de código aberto. Protégé é uma ferramenta específica para modelagem de ontologias que permite que o modelo seja exportado como RDF e serializado em formato XML. A partir do modelo

em formato XML obtido pelo Protégé foi obtido o modelo em formato N3, para isso foi desenvolvido um programa de conversão simples em Java utilizando a biblioteca *Jena*.

7.3 ESTUDO DE CASO

Estudo de caso foi a forma escolhida para explicar a aplicação do modelo RDF. É preciso ter em mente que o protocolo IRIS foi especificado de forma genérica, pois um dos principais requisitos consistiu em generalizar a informação fornecida pelo serviço Whois para que fosse independente de políticas locais. Desta forma, é preciso identificar os requisitos genéricos que podem ser diretamente mapeados no modelo. Além disso, é necessário detectar quais requisitos locais não puderam ser mapeados diretamente. Neste caso, é preciso elaborar uma extensão do modelo específica para os requisitos locais.

RDF é flexível suficiente para que possa ser estendida sem que seja necessário atualizar aplicações. Uma aplicação swhois cliente seria capaz de processar uma resposta RDF, independente de qualquer extensão que possa estar contida, desde que a extensão seja consistente com o modelo original. No caso as extensões para o modelo RDF consistiriam na adição de propriedades específicas em cada uma das classes existentes para atender políticas locais. As propriedades genéricas ou específicas podem ser detectadas em tempo de execução, sem necessidade de atualização do código fonte.

A facilidade de extensão de RDF é uma das grandes vantagens em relação ao uso de XML puro. O processamento de um documento XML usando *parsers* tradicionais (API SAX ou DOM) fornece uma estrutura de dados que corresponde à sintaxe do documento. Por essa razão, toda vez que a sintaxe XML muda, por exemplo quando existe necessidade de extensão para atender políticas locais, a aplicação precisa ser atualizada. Enquanto o processamento de documentos RDF com *parsers* específicos fornece uma estrutura de dados que corresponde ao modelo de dados, que é completamente independente da sintaxe do documento original. Pequenas alterações, como a adição de propriedades em classes existentes, não requerem nenhuma alteração na aplicação. O impacto da extensão ou atualização do modelo original na manutibilidade da aplicação possui um papel fundamental na escolha da tecnologia a ser utilizada.

Os passos para o estudo de caso foram os seguintes:

1. Escolher um servidor whois
2. Consultar um domínio
3. Mostrar a resposta mapeada em RDF sem extensão
4. Mostrar a resposta mapeada em RDF com extensão

7.3.1 ESCOLHER UM SERVIDOR WHOIS

O servidor Whois escolhido foi o servidor Whois do registro brasileiro de nomes de domínios: Registro.br. Dados do servidor:

- Nome do Servidor: registro.br
- Porta: TCP/43
- Sintaxe: whois *<nome – do – dominio>*

7.3.2 CONSULTAR UM DOMÍNIO

A consulta para o domínio *usp.br* foi feita utilizando a *interface* em linha de comando do servidor Whois do Registro.br. Na prática o resultado foi obtido através da execução do comando *telnet* para o servidor *registro.br* na porta 43. O resultado obtido é mostrado a seguir:

```
% Copyright (c) Nic.br
% The use of the data below is only permitted as described in
% full by the terms of use (http://registro.br/termo/en.html),
% being prohibited its distribution, comercialization or
% reproduction, in particular, to use it for advertising or
% any similar purpose.
% 2007-05-15 11:27:29 (BRT -03:00)

domain:      usp.br
owner:       UNIVERSIDADE DE SAO PAULO
ownerid:     063.025.530/0001-04
```

responsible: TEREZA CRISTINA MELO DE BRITO CARVALHO
owner-c: JCF286
admin-c: JCF286
tech-c: ANG149
billing-c: JCF286
nserver: bee.uspnet.usp.br 143.107.253.3
nsstat: 20070514 AA
nslastaa: 20070514
nserver: reno.uspnet.usp.br 143.107.51.2
nsstat: 20070514 AA
nslastaa: 20070514
nserver: bee08.uspnet.usp.br 143.107.253.5
nsstat: 20070514 AA
nslastaa: 20070514
nserver: ns1.ansp.br
nsstat: 20070514 AA
nslastaa: 20070514
created: 19960101
changed: 20060704
status: published

nic-hdl-br: ANG149
person: Andre Gerhard
e-mail: cceadmin@usp.br
created: 20011217
changed: 20020411

nic-hdl-br: JCF286
person: Jairo Carlos Filho
e-mail: jcarlosf@usp.br
created: 20011217
changed: 20060424

% Security and mail abuse issues should also be addressed to
% cert.br, <http://www.cert.br/>, respectively to cert@cert.br
% and mail-abuse@cert.br
%
% whois.registro.br accepts only direct match queries. Types
% of queries are: domain (.br), ticket, provider, ID, CIDR
% block, IP and ASN.

7.3.3 RESPOSTA MAPEADA EM RDF SEM EXTENSÃO

A resposta para a consulta anterior mapeada em RDF sem extensão, ou seja, de acordo com o modelo RDF original, e serializada na linguagem N3 seria a seguinte:

```
@prefix rdf:      <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs:     <http://www.w3.org/2000/01/rdf-schema#> .
@prefix swhois:  <http://sw.deriv.org/swhois#> .
```

```
<http://swhois.registro.br/domain/usp-br>
  a          swhois:Domain ;
  swhois:possuiFQDN "usp.br" ;
  swhois:possuiContatoAdm
    <http://swhois.registro.br/contato/JCF286> ;
  swhois:possuiContatoTec
    <http://swhois.registro.br/contato/ANG149> ;
  swhois:possuiContatoCob
    <http://swhois.registro.br/contato/JCF286> ;
  swhois:possuiOutroContato
    <http://swhois.registro.br/contact/KAV22> ;
  swhois:possuiRegistrante
    <http://swhois.registro.br/contato/JCF286> ;
  swhois:possuiRegistro
    "Registro .br" ;
  swhois:possuiServidorDNS
    <http://swhois.registro.br/servidorDNS/bee> ,
    <http://swhois.registro.br/servidorDNS/reno> ,
    <http://swhois.registro.br/servidorDNS/bee08> ;
  swhois:possuiStatus "published" ;
  swhois:possuiDataDelegInicial
    "1996-01-01T00:00:00.000Z" .
```

```
<http://swhois.registro.br/contato/ANG149>
  a          swhois:Contato ;
  swhois:possuiDataCriacao
    "2001-12-17T00:00:00.000Z" ;
  swhois:possuiDataUltAlt
    "2002-04-11T00:00:00.000Z" ;
  swhois:possuiEmail "cceadmin@usp.br" ;
  swhois:possuiNome "Andre Gerhard" .
```

```
<http://swhois.registro.br/contato/JCF286>
  a          swhois:Contato ;
  swhois:possuiDataCriacao
    "2001-12-17T00:00:00.000Z" ;
  swhois:possuiDataUltAlt
```

```

                "2006-04-24T00:00:00.000Z" ;
swhois:possuiEmail "jcarlosf@usp.br" ;
swhois:possuiNome "Jairo Carlos Filho" .

<http://swhois.registro.br/servidorDNS/reno>
a          swhois:ServidorDNS ;
swhois:possuiDataUltVer
            "2007-05-14T00:00:00.000Z " ;
swhois:possuiIPv4 "143.107.51.2" ;
swhois:possuiNomeSrv
            "reno.uspnet.usp.br" .

<http://swhois.registro.br/servidorDNS/bee>
a          swhois:ServidorDNS ;
swhois:possuiDataUltVer
            "2007-05-14T00:00:00.000Z " ;
swhois:possuiIPv4 "143.107.253.3" ;
swhois:possuiNomeSrv
            "bee.uspnet.usp.br" .

<http://swhois.registro.br/servidorDNS/bee08>
a          swhois:ServidorDNS ;
swhois:possuiDataUltVer
            "2007-05-14T00:00:00.000Z " ;
swhois:possuiIPv4 "143.107.253.5" ;
swhois:possuiNomeSrv
            "bee08.uspnet.usp.br" .

<http://swhois.registro.br/servidorDNS/ns1>
a          swhois:ServidorDNS ;
swhois:possuiDataUltVer
            "2007-05-14T00:00:00.000Z " ;
swhois:possuiNomeSrv
            "ns1.ansp.br" .

```

Pela comparação entre a resposta Whois original e a resposta mapeada em RDF, de acordo com o modelo RDF, pode-se perceber que nem todas as informações puderam ser mapeadas. A maior parte das propriedades de domínio foram mapeadas, mas houve exceções que não estavam previstas no modelo original. Por outro lado, todas as propriedades de contato foram mapeadas sem nenhuma exceção. Com relação aos servidores DNS, algumas propriedades foram mapeadas e outras não. As propriedades da classe *swhois:Domínio* que não puderam ser mapeadas no modelo RDF original foram:

- Responsável;
- ID do proprietário;
- Nome do proprietário;
- Data de última alteração do domínio.

Por sua vez, as propriedades da classe *swhois:ServidorDNS* que não puderam ser mapeadas no modelo RDF original:

- Resposta da última verificação DNS;
- Data da última verificação DNS que confirmou o servidor como sendo o servidor autoritativo sobre o domínio.

7.3.4 RESPOSTA MAPEADA EM RDF COM EXTENSÃO

Como foi visto anteriormente, a resposta da consulta Whois não pôde ser completamente mapeada em RDF, de acordo com o modelo RDF original. O desafio consiste em utilizar as facilidades de extensão da *framework* RDF para obter um resultado mais satisfatório pois um dos requisitos do protocolo IRIS é agregar novas funcionalidades sem deixar de provisionar nenhuma das funcionalidades atuais do protocolo Whois. Embora não seja um item obrigatório foi criado um espaço de nomes dedicado a extensão do modelo RDF para o registro de domínios brasileiro, cujo prefixo é bastante sugestivo *br*. Desta forma, é possível enfatizar e visualizar melhor quais elementos fazem parte da extensão e quais elementos pertencem ao modelo RDF original.

As novas propriedades da classe *swhois:Domínio* que foram criadas como parte da extensão para registro brasileiro são:

- *br:possuiResponsavel*;
- *br:possuiPropId*;
- *br:possuiProprietario*;
- *br:possuiUltAltDom*.

As propriedades da classe *swhois:ServidorDNS* que não existiam no modelo RDF original e foram criadas para atender as necessidades do registro brasileiro:

- *br:possuiNsStat* – resposta da última verificação DNS;
- *br:possuiDataUltAA* – data da última verificação DNS que retornou AA, ou seja, servidor autoritativo sobre o domínio.

Nenhuma propriedade foi adicionada à classe *swhois:Contato*.

É importante notar que as aplicações cliente que tenham sido desenvolvidas com base no modelo RDF original continuariam a funcionar sem necessidade de atualização. Esta característica consiste em uma grande vantagem de RDF sobre XML. Cada servidor poderia possuir sua própria extensão para seguir políticas locais e as aplicações continuariam processando os dados de forma automatizada sem necessidade de atualização. Tecnicamente isso é possível pela forma como RDF foi desenvolvido, em que os *parsers* RDF são capazes de processar as novas propriedades de domínio em tempo de execução e, desta forma, podem incluir as novas propriedades na estrutura de dados.

A resposta para a consulta anterior mapeada em RDF, com extensão para atender as particularidades do serviço de informação do registro de nomes de domínios brasileiro, e serializada na linguagem N3 seria a seguinte:

```
@prefix rdf:      <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs:     <http://www.w3.org/2000/01/rdf-schema#> .
@prefix swhois:  <http://sw.deri.org/swhois#> .
@prefix br:      <http://registro.br/swhois#> .
```

```
<http://swhois.registro.br/domain/usp-br>
  a          swhois:Domain ;
  swhois:possuiFQDN "usp.br" ;
  swhois:possuiContatoAdm
    <http://swhois.registro.br/contato/JCF286> ;
  swhois:possuiContatoTec
    <http://swhois.registro.br/contato/ANG149> ;
  swhois:possuiContatoCob
    <http://swhois.registro.br/contato/JCF286> ;
  swhois:possuiOutroContato
    <http://swhois.registro.br/contact/KAV22> ;
  swhois:possuiRegistrante
    <http://swhois.registro.br/contato/JCF286> ;
  swhois:possuiRegistro
    "Registro .br" ;
  swhois:possuiServidorDNS
    <http://swhois.registro.br/servidorDNS/bee> ,
    <http://swhois.registro.br/servidorDNS/reno> ,
    <http://swhois.registro.br/servidorDNS/bee08> ;
  swhois:possuiStatus "published" ;
```

```

swhois:possuiDataDelegInicial
    "1996-01-01T00:00:00.000Z" ;
br:possuiDataUltAltDom
    "2006-07-04T00:00:00.000Z" ;
br:possuiPropId "063.025.530/0001-04" ;
br:possuiProprietario
    "UNIVERSIDADE DE SAO PAULO" ;
br:possuiResponsavel
    "TEREZA CRISTINA MELO DE BRITO CARVALHO" .

```

```

<http://swhois.registro.br/contato/ANG149>
a      swhois:Contato ;
swhois:possuiDataCriacao
    "2001-12-17T00:00:00.000Z" ;
swhois:possuiDataUltAlt
    "2002-04-11T00:00:00.000Z" ;
swhois:possuiEmail "cceadmin@usp.br" ;
swhois:possuiNome "Andre Gerhard" .

```

```

<http://swhois.registro.br/contato/JCF286>
a      swhois:Contato ;
swhois:possuiDataCriacao
    "2001-12-17T00:00:00.000Z" ;
swhois:possuiDataUltAlt
    "2006-04-24T00:00:00.000Z" ;
swhois:possuiEmail "jcarlosf@usp.br" ;
swhois:possuiNome "Jairo Carlos Filho" .

```

```

<http://swhois.registro.br/servidorDNS/reno>
a      swhois:ServidorDNS ;
swhois:possuiDataUltVer
    "2007-05-14T00:00:00.000Z " ;
swhois:possuiIPv4 "143.107.51.2" ;
swhois:possuiNomeSrv
    "reno.uspnet.usp.br" ;
br:possuiDataUltAA "2007-05-14T00:00:00.000Z" ;
br:possuiNsStat "AA" .

```

```

<http://swhois.registro.br/servidorDNS/bee>
a      swhois:ServidorDNS ;
swhois:possuiDataUltVer
    "2007-05-14T00:00:00.000Z " ;
swhois:possuiIPv4 "143.107.253.3" ;
swhois:possuiNomeSrv
    "bee.uspnet.usp.br" ;
br:possuiDataUltAA "2007-05-14T00:00:00.000Z" ;
br:possuiNsStat "AA" .

```

```

<http://swhois.registro.br/servidorDNS/bee08>
  a      swhois:ServidorDNS ;
  swhois:possuiDataUltVer
    "2007-05-14T00:00:00.000Z " ;
  swhois:possuiIPv4 "143.107.253.5" ;
  swhois:possuiNomeSrv
    "bee08.uspnet.usp.br" ;
  br:possuiDataUltAA "2007-05-14T00:00:00.000Z" ;
  br:possuiNsStat "AA" .

<http://swhois.registro.br/servidorDNS/ns1>
  a      swhois:ServidorDNS ;
  swhois:possuiDataUltVer
    "2007-05-14T00:00:00.000Z " ;
  swhois:possuiNomeSrv
    "ns1.ansp.br" ;
  br:possuiDataUltAA "2007-05-14T00:00:00.000Z" ;
  br:possuiNsStat "AA" .

```

Desta forma, todos os dados fornecidos pela consulta do domínio *usp.br* no servidor Whois do registro brasileiro puderam ser mapeados em RDF com sucesso. Da mesma forma, esse mecanismo de extensão poderia ser utilizado para mapear o resultados de consultas de todos os outros servidores Whois.

7.4 IMPLEMENTAÇÃO

Prototipagem para sistemas de software é uma técnica bem conhecida para elucidação e validação dos requisitos de sistema. Protótipo é uma versão inicial do sistema final que está disponível na fase inicial do processo de desenvolvimento. Assim sendo, é fundamental que este seja desenvolvido rapidamente. Conseqüentemente, algumas funcionalidades serão sacrificadas com o objetivo de acelerar o desenvolvimento do protótipo. O protótipo permite demonstrar conceitos, opções de modelagem, aumentar o conhecimento sobre os problemas e sobre as possíveis soluções.

Para validar os resultados teóricos do estudo, foi desenvolvido um protótipo do que seria um serviço de informação sobre registros da Internet na Web Semântica. O protótipo teve o escopo limitado a nomes de domínios. Além disso, não houve intenção de demonstrar o mapeamento em RDF para todos os servidores Whois de nomes de

domínios. O intuito foi mostrar como poderia ser feito o mapeamento em RDF para quatro servidores diferentes com saídas bem distintas. Além disso, demonstrar que o modelo poderia ser estendido para outros servidores Whois sem restrições.

Os passos foram os seguintes:

1. Escolher quatro servidores whois
2. Coletar uma amostra de saída de cada um servidores
3. Mostrar como cada uma delas poderia ser mapeada em RDF
4. Implementação do protótipo de servidor whois para a Web semântica

Nas seções seguintes cada um dos passos será descrito em detalhes.

7.4.1 ESCOLHER QUATRO SERVIDORES WHOIS

Os quatro servidores Whois foram escolhidos de acordo com três critérios. Os únicos critérios foram a existência de diferença na semântica (conteúdo) da resposta, na sintaxe (forma) da resposta e a variabilidade no modelo de registro. Como foi visto, existem dois tipos de modelos, no primeiro modelo (*thick*) o registro de domínios mantém tanto informação de contato quanto informação sobre os servidores DNS. No segundo modelo (*thin*), o *registrar* mantém informação de contato e o registro mantém apenas informação sobre os servidores DNS. O modelo pode afetar o resultado da consulta a nomes de domínios no serviço Whois e por isso foi utilizado como critério de seleção. Na tabela 5 foram resumidos os dados de cada um dos servidores escolhidos.

A seguir uma breve explicação de cada um dos servidores escolhidos:

- Registro.br: registro brasileiro de nomes de domínios;
- Registry.ie: registro irlandês de nomes de domínios;
- *Public Interest Registry* (PIR): registro de domínios .org;
- Verisign: atualmente o maior registro de domínios, responsável pelo registro de domínios .com, .net dentre outros.

Tabela 5: Servidores Whois escolhidos para o protótipo

Organização	Nome do Servidor	Modelo	Tipo
Registro.br	registro.br	thick	ccTLD (.br)
Register.ie	whois.domainregistry.ie	thick	ccTLD (.ie)
PIR	whois.pir.org	thin	gTLD (.org)
Verisign	whois.verisign-grs.com	thin	gTLD (.com)

7.4.2 COLETAR UMA AMOSTRA DE SAÍDA DE CADA UM SERVIDORES

Como pode-se observar na tabela 5, o primeiro servidor Whois escolhido foi o mesmo servidor utilizado no estudo de caso: registro.br. Desta forma, toda a análise realizada durante o estudo de caso pôde ser reutilizada para a implementação do protótipo. Assim como foi realizado o estudo de caso, será feita uma análise semelhante para os demais servidores Whois. Para isso foi coletada uma amostra de saída de cada um dos servidores. Para mais detalhes sobre as amostras coletadas, ver o Apêndice.

7.4.3 MOSTRAR COMO CADA UMA DELAS PODERIA SER MAPEADA EM RDF

Assim como foi feito durante o estudo de caso, cada uma das saídas dos servidores Whois foi mapeada em RDF. Em seguida, foi realizado um mapeamento da informação fornecida por cada um dos servidores Whois no modelo RDF original, identificando quais propriedades não puderam ser mapeadas no modelo e exigiram extensão. Quando necessário o modelo RDF foi estendido. Assim como foi feito para o registro brasileiro de domínios durante o estudo de caso.

7.4.4 IMPLEMENTAÇÃO DE UM PROTÓTIPO DE SERVIDOR WHOIS

Na figura 5, é possível observar a interação entre o protótipo, o usuário e os quatro servidores Whois. Além disso, a figura mostra também uma possível interação entre esse protótipo e aplicações de Web Semântica, a serem desenvolvidas no futuro, que possam trocar dados em formato RDF e viabilizar o processamento automático da informação fornecida pelos servidores Whois. Afinal, este seria o objetivo final do desenvolvimento do serviço de informação sobre registros da Internet para a Web Semântica. Vale ressaltar que na interação entre o usuário e os servidores Whois, o protótipo funciona como um cliente Whois/IRIS. Na possível interação entre as demais aplicações e os servidores Whois o protótipo funcionaria como um servidor *proxy*.

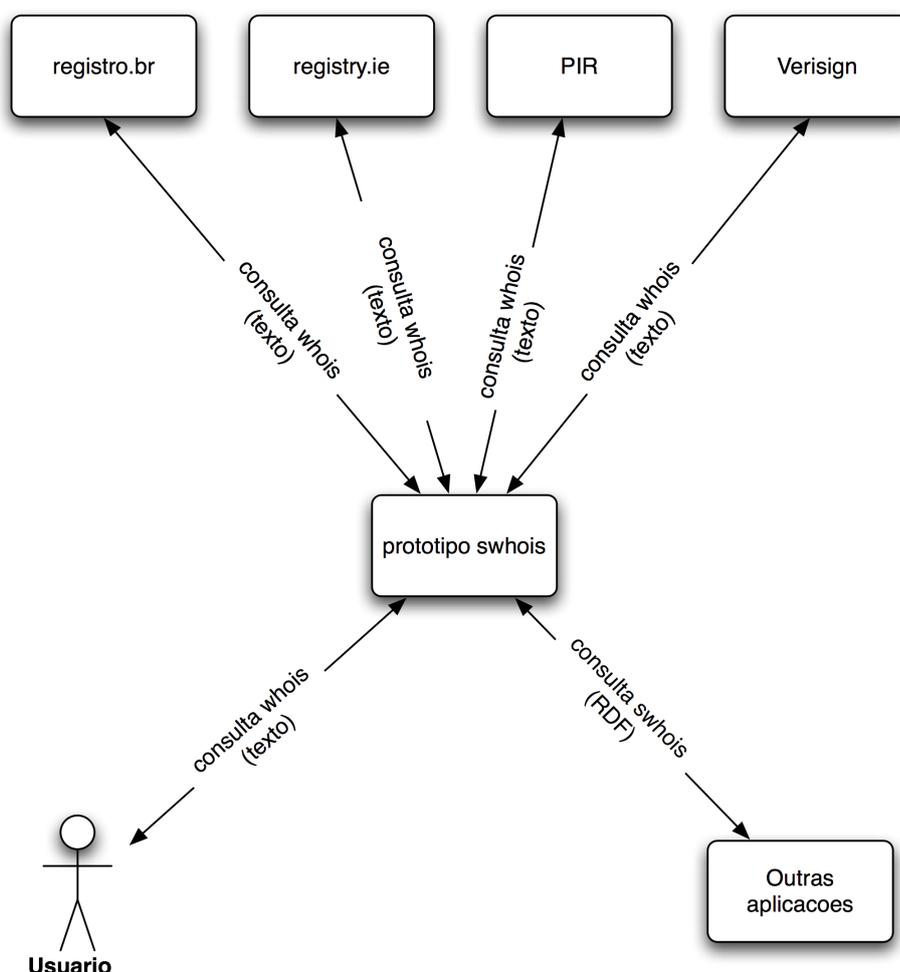


Figura 5: Interação do protótipo com o usuário, os servidores Whois e outras aplicações

O protótipo foi desenvolvido exclusivamente com tecnologia de código aberto e livre. A linguagem de programação utilizada foi Java(JDK1.5), com o uso biblioteca Jena

para processar RDF. Eclipse foi a IDE escolhida como ambiente de desenvolvimento. O sistema operacional escolhido foi Kubuntu (Ubuntu 4.0.3): Ubuntu com KDE. A máquina utilizada foi um *laptop* IBM ThinkPad modelo 2722, com 1 Giga bytes de memória e processador Intel Pentium 1400MHz. A linguagem de serialização escolhida foi XML mas poderia ter sido N3, TriX ou qualquer outra linguagem de representação suportada pela biblioteca Jena. O desafio consistiu mapear a sintaxe e semântica de cada servidor Whois para a ontologia. Então, estender o modelo RDF original quando necessário.

O protótipo demonstrou a facilidade de integração de diversas fontes de dados. Uma das principais vantagens da *framework* RDF consiste em facilitar a integração de diversas bases de dados sem a necessidade que elas sejam homogêneas. Principalmente, não há necessidade de converter as bases de dados relacionais em bases de dados RDF. O que certamente dificultaria a adoção do novo protocolo.

7.5 AVALIAÇÃO DE DESEMPENHO

Para complementar a validação dos resultados teóricos foram realizados dois tipos de experimentos utilizando o protótipo. O primeiro tipo de experimento realizado foi um experimento com o objetivo de comparar os desempenhos do protótipo e do cliente Whois. O segundo tipo de experimento teve como objetivo a análise de desempenho do protótipo com base em testes de carga. A seguir os experimentos serão descritos em detalhes.

7.5.1 ANÁLISE DE DESEMPENHO: WHOIS X SWHOIS

O experimento consistiu em realizar testes periódicos em que cada uma das aplicações fosse capaz de realizar 10 consultas de nomes de domínios para cada um dos quatro servidores para os quais o protótipo *swhois* tem suporte: *registro.br*, *registry.ie*, *whois.pir.org* e *whois.verisign-grs.com*. Para isso foi escrito um programa em Java capaz de ler um arquivo texto contendo um nome de domínio por linha, o arquivo utilizado possuía 40 nomes de domínios, sendo:

- 10 domínios .br

- 10 domínios .ie
- 10 domínios .org
- 10 domínios .com

A cada etapa o programa processava o arquivo de entrada, então para cada nome de domínio executava o programa Whois e o protótipo swhois e registrava o tempo de execução por domínio para cada um separadamente. Este experimento precisou ser dividido em etapas pois os servidores Whois consultados limitam o número de consultas por endereço IP em um determinado intervalo de tempo, que varia conforme políticas locais. Entre uma etapa e outra o endereço IP era alterado manualmente utilizando DHCP. Foram executadas 4 etapas por dia durante 5 dias, cada etapa corresponde a 40 consultas a cada um dos nomes de domínios para o protótipo e 40 consultas a nomes de domínios utilizando o cliente Whois. A cada etapa o protótipo e o cliente Whois executaram 10 consultas para cada um dos quatro servidores Whois, ou seja, cada etapa corresponde a 40 consultas. Ao final do experimento foram calculados os valores médios de tempo de execução por domínio para o protótipo e para o cliente Whois.

Duas considerações importantes para este experimento:

1. O ponto crítico para o desempenho do protótipo consiste no processamento da resposta para obter o modelo RDF e gerar a resposta em formato RDF serializado em XML;
2. O ponto crítico para o desempenho do cliente whois consiste no fato dele ter sido desenvolvido para todos os tipos de domínios e não apenas domínios .ie, .br, .org e .com. Seu desempenho é limitado principalmente pelo uso de expressões regulares para encontrar a correspondência entre o nome de domínio consultado e o servidor autoritativo em uma lista de servidores que o cliente mantém. Apesar disso, o cliente whois não processa a resposta apenas redireciona para a saída padrão.
3. O desempenho das duas aplicações é afetado por flutuações no desempenho da conexão a Internet. Esta foi outra razão pela qual o experimento envolveu diversas etapas distribuídas em dias e horários distintos. Além do fato dos servidores limitarem o número de consultas por IP em um dado intervalo de tempo.

7.5.2 ANÁLISE DE DESEMPENHO: TESTE DE CARGA SWHOIS

Para o teste de carga realizado com o protótipo, foi utilizado o mesmo arquivo de entrada do teste comparativo de desempenho do protótipo e do cliente whois. Durante a fase de preparação, para cada entrada do arquivo foi realizada uma consulta whois manual utilizando *telnet* e a resposta foi armazenada em um arquivo separado. O programa Java utilizado no teste anterior foi alterado para ler os arquivos de resposta whois durante a inicialização. Além disso, o programa de teste foi alterado para executar apenas o protótipo *swhois*. A cada iteração o programa de teste executava o protótipo fornecendo a resposta Whois e o protótipo apenas executaria a fase de processamento da resposta e geração do RDF serializado em XML correspondente. O objetivo deste experimento foi demonstrar que o processamento da resposta e geração do modelo RDF correspondente não teria um impacto significativo no desempenho do protótipo. Neste caso, o desempenho da rede não afetaria o resultado do experimento. Além disso, este teste carga não estaria limitado pelo número de consultas por IP imposto por cada um dos servidores Whois.

7.6 RESULTADOS EXPERIMENTAIS

Os resultados dos experimentos para análise comparativa de desempenho do protótipo *swhois* e do cliente whois, estão apresentados na tabela a seguir e demonstraram que o desempenho do protótipo nas condições especificadas apresentou tempo de execução 88,5% do tempo de execução do cliente Whois, ou seja, o desempenho do protótipo é mais de 10% melhor do que o desempenho do cliente Whois nas condições especificadas anteriormente. Os resultados do experimento de análise comparativa de desempenho estão apresentados nas colunas: whois (tempo de execução em milissegundos), *swhois* (tempo de execução em milissegundos) e *swhois/whois* (relação percentual do tempo de execução do protótipo em relação tempo de execução do cliente Whois em milissegundos).

Tabela 6: Resultados dos Experimentos para Análise de Desempenho do protótipo *swhois*

Nome do domínio	whois (ms)	swhois (ms)	swhois/ whois(%)	Carga swhois (ms)
usp.br	583	547	93,8	13,79
microsoft.com.br	484	485	100,2	0,93
verisign.com.br	502	481	95,8	2,18
cisco.com.br	477	499	104,6	0,91
lucent.com.br	502	488	97,2	2,17
nortel.com.br	475	477	100,4	2,59
milena.eng.br	482	479	99,0	0,90
demarest.adv.br	473	474	100,2	1,43
greenpeace.org.br	478	472	98,7	1,31
siemens.com.br	471	473	100,4	0,86
deri.ie	103	99	96,1	1,54
microsoft.ie	98	93	94,9	0,42
verisign.ie	103	91	88,0	0,41
cisco.ie	99	89	89,9	0,39
lucent.ie	101	89	88,1	1,36
nortel.ie	96	91	94,8	2,60
ibm.ie	104	90	86,5	0,41
apple.ie	100	89	89,0	0,31
sap.ie	106	94	88,7	2,07
bbc.ie	100	92	92,0	1,38
verisign.com	500	467	93,4	1,37
cisco.com	284	272	95,8	1,35
lucent.com	340	269	79,1	1,16
nortel.com	510	400	78,4	1,49
sap.com	357	271	75,9	1,38
bbc.com	318	281	88,4	1,36
britishairways.com	363	271	74,7	1,43
americanairways.com	312	270	86,5	1,36
boots.com	318	271	85,2	1,37
ryanair.com	350	275	78,57	1,37
deri.org	420	323	76,9	4,16
greenpeace.org	396	306	77,3	4,81

continua...

Tabela 6: Resultados dos Experimentos para Análise de Desempenho do protótipo *swhois*

Nome do domínio	whois (ms)	swhois (ms)	swhois/ whois(%)	Carga swhois (ms)
wwf.org	394	303	76,9	5,32
pir.org	396	303	76,5	2,21
sos.org	394	305	77,4	2,91
weightwatcher.org	393	300	76,3	2,78
gnu.org	397	315	79,3	2,00
ctan.org	397	316	79,6	2,07
sourceforge.org	280	300	107,1	1,96
slashdot.org	395	304	76,96	2,01
Média	336,2ms	297,9ms	88,5%	2,04 s

Para os mesmos domínios consultados durante a análise comparativa de desempenho entre o protótipo e o cliente Whois, foi realizado o teste de carga para o protótipo. A diferença é que no teste de carga o protótipo não abriu uma conexão TCP com cada um dos servidores Whois para obter a resposta pois as respostas foram obtidas previamente e foram armazenadas em arquivos em formato texto. O programa de testes leu os arquivos contendo as respostas durante a inicialização e forneceu ao protótipo para processamento e conversão em RDF. Desta forma foi possível isolar o processamento da resposta para o teste de carga, que é exatamente o processamento que não existe no protocolo Whois e poderia impactar o desempenho como um todo, caso seu desempenho isoladamente não fosse adequado. Além disso, não seria possível executar um teste de carga consultando um servidor Whois devido ao limite de consultas imposto por endereço IP. Cada simulação de consulta foi realizada dez mil vezes e o valor médio obtido por nome de domínio foi incluído na mesma tabela, na coluna denominada "Carga swhois". Ao final, foi calculado o valor médio de todos os domínios e, desta forma, chegou-se ao valor médio de 2,044 milissegundos para o processamento de uma resposta. Este valor corresponde a 0,6% do tempo médio do cliente Whois (336,2ms) para obter uma resposta para consulta de domínio. Os resultados dos dois experimentos mostraram que a inclusão de tecnologias de Web Semântica no serviço de busca de informação sobre registros da Internet não teria impacto no desempenho, ainda que durante a fase de transição os servidores não ofereçam a informação em formato RDF e haja necessidade de conversão de formato texto para o modelo RDF.

7.7 DISCUSSÃO

A mesma técnica aplicada aos quatro servidores Whois utilizados no protótipo poderia ser generalizada e aplicada a todos os demais servidores Whois. O grande desafio de um protótipo de serviço de informação sobre registros da Internet para a Web Semântica poderia ser a descoberta de serviço, ou seja, descobrir o servidor autoritativo para determinado domínio. Entretanto esse problema já foi resolvido pelo protocolo IRIS, que especifica como utilizar o mecanismo S-NAPTR para a descoberta de serviços. O mecanismo S-NAPTR consiste em utilizar consultas DNS para descoberta de serviço, no caso descoberta de servidores Whois ou servidores IRIS. Inicialmente este mecanismo foi introduzido para dar suporte a base de dados distribuída pois qualquer aplicação cliente seria capaz de encontrar o servidor Whois autoritativo sobre determinado recurso. A implementação e validação do protótipo foi o primeiro passo em direção ao serviço de informação sobre registros da Internet baseado no protocolo IRIS com tecnologias de Web Semântica, permitindo que diversas aplicações consultem o serviço, processem e compartilhem os dados obtidos de forma automatizada.

Os mecanismos de segurança poderiam ser os mesmos definidos pelo protocolo IRIS. O fato de utilizar RDF é completamente independente dos mecanismos de transporte e segurança que poderiam ser utilizados como especificados pelo protocolo IRIS. O protocolo IRIS define dois tipos de transporte: orientado a conexão baseado em TCP chamado (XPC) e não-orientado a conexão baseado em UDP chamado (LWZ). Os mecanismos de segurança só se aplicam ao transporte TCP, porque os mecanismos de autenticação e autorização são serviços orientados a conexão. O mecanismo de transporte LWZ existe apesar de não suportar mecanismos de segurança porque como o próprio nome diz ele é um mecanismo ideal para consultas Whois individuais e esporádicas porque elas envolvem a troca de uma pequena quantidade de dados.

O protótipo demonstrou algumas vantagens da framework RDF em relação ao XML:

- A facilidade de extensão para atender requisitos e políticas locais;
- A facilidade de integração de fontes de dados heterogêneas;
- A facilidade de manutenção das aplicações pois não há necessidade de alteração do código fonte e novas propriedades são detectadas em tempo de execução.
- A facilidade de processamento automatizado de documentos RDF e a disponibilidade de bibliotecas prontas de código aberto e livre para processar documentos RDF.

- Outra conclusão importante consiste no fato de não haver necessidade de conversão das bases de dados relacionais em bases de dados RDF, pois a conversão pode ser feita sob demanda o que facilitaria a adoção do novo protocolo.

Para o futuro, existem dois pontos principais que poderiam ser considerados para aprimorar o serviço de informação. Um dos princípios da Web semântica é a reutilização de ontologias existentes na modelagem de novas ontologias, por exemplo poderia ser utilizada a ontologia *Friend Of A Friend* (FOAF) ¹ para descrição de contatos. Um detalhe importante é que principal ferramenta atualmente disponível para a modelagem de ontologias Protégé ², que foi utilizada neste estudo, não suporta a reutilização de ontologias existentes e esta foi a principal dificuldade encontrada neste quesito. Outro fato é que existem novas pesquisas (CARROLL et al., 2005) dedicadas ao desenvolvimento de uma *framework* de segurança baseada em RDF para garantir a confiabilidade dos dados e poderia ser utilizada para garantir a confiabilidade nos dados fornecidos pelo serviço de informação sobre recursos da Internet. Nenhuma das duas propostas de aprimoramento invalida o estudo atual, pelo contrário poderia adicionar novas funcionalidades no futuro.

¹<http://www.foaf-project.org/>

²<http://protege.stanford.edu/>

8 CONCLUSÃO

As principais contribuições deste trabalho foram: (1) o trabalho de pesquisa aprofundado sobre o sistema de informação para registros da Internet; (2) o estudo do protocolo IRIS em comparação com seus predecessores – os protocolos Whois e Joint Whois; (3) estudo teórico sobre os benefícios das tecnologias de Web Semântica, em particular sobre *framework* RDF; (4) a modelagem de uma ontologia utilizando a *framework* RDF para o protocolo IRIS; (5) implementação em Java de um protótipo de aplicação para a ontologia desenvolvida; e (6) a validação de desempenho da aplicação desenvolvida. O estudo se limitou a consultas de nomes de domínios mas poderia ser facilmente generalizado para atender todas as consultas atualmente suportadas pelo protocolo Whois e outras a serem definidas. O que se desejou investigar é se as tecnologias de Web Semântica poderiam contribuir para o desenvolvimento do protocolo IRIS. Durante o estudo teórico foram demonstrados os benefícios qualitativos da adoção das tecnologias de Web Semântica: (1) a curto prazo a facilidade de extensão e manutenção das aplicações e (2) a longo prazo a possibilidade de processamento automatizado da informação entre as aplicações e a viabilização da criação de agentes inteligentes capazes de utilizar essa informação. Durante os experimentos foi possível demonstrar quantitativamente que o desempenho do sistema de informação não seria prejudicado pelo processamento adicional requerido pela *framework* RDF, em particular para conversão de uma resposta em formato texto para o modelo RDF em memória. Ao final do estudo concluiu-se que as tecnologias de Web Semântica poderiam contribuir para o sucesso do protocolo IRIS.

REFERÊNCIAS BIBLIOGRÁFICAS

- ALAETTINOGLU, C. et al. **Routing Policy Specification Language (RPSL)**. RFC2622. 1999.
- ALVESTRAND, H. **IETF Policy on Character Sets and Languages**. RFC2277. 1998.
- APPARAO, V. et al. **Document Object Model (DOM) Level 1 Specification. W3C Recommendation**. Disponível em: <<http://www.w3.org/TR/REC-DOM-Level-1>>. Acesso em: 05 ago. 2007.
- BERGLUND, A. **Extensible Stylesheet Language (XSL) Version 1.1. W3C Recommendation**. Disponível em: <<http://www.w3.org/TR/xsl11/>>. Acesso em: 05 ago. 2007.
- BERNERS-LEE, T. **Notation 3: An readable language for data on the Web**. Disponível em: <<http://www.w3.org/DesignIssues/Notation3>>. Acesso em: 05 ago. 2007.
- BERNERS-LEE, T. (Ed.). **Weaving the Web**. São Francisco: Harper. 1999.
- BERNERS-LEE, T.; HENDLER, J.; LASSILA, O. **The Semantic Web**. *Scientific American*, v. 284, n. 5, Maio 2001.
- BRAY, T. et al. **Extensible Markup Language (XML) 1.0. W3C Recommendation**. Disponível em: <<http://www.w3.org/TR/REC-xml>>. Acesso em: 05 ago. 2007.
- BROEKSTRA, J. et al. **Enabling Knowledge Representation on the Web by Extending RDF Schema**. *Proceedings of the 10th International Conference on World Wide Web*, Hong Kong: ACM, p. 467–478, Abril 2001.
- CAIRES, M. C. **WHOIS, JOINT WHOIS AND CRISP**. *Proceedings of the IADIS International Conference WWW/Internet 2006*, v. 2, p. 300–304, 2006.
- CARROLL, J.; STICKLER, P. **TriX: Triples in XML**. Disponível em: <<http://www.w3.org/2004/03/trix/>>. Acesso em: 05 ago. 2007.
- CARROLL, J. J. et al. **Trust Networks on the Semantic Web**. *Proceedings of the 14th international conference on World Wide Web*, ACM, Maio 2005.
- CERF, V. **IAB Recommended Policy on Distributing Internet Identifier Assignment and IAB Recommended Policy Change to Internet "Connected" Status**. RFC 1174. 1990.
- CLARK, J. **XSL Transformations (XSLT) Version 1.0. W3C Recommendation**. Disponível em: <<http://www.w3.org/TR/xslt>>. Acesso em: 05 ago. 2007.
- DAIGLE, L. **WHOIS Protocol Specification**. RFC 3912. 2004.

- DAIGLE, L.; NEWTON, A. **A perspective on the role of the IETFs CRISP working group**. Disponível em: <<http://www.icann.org/montreal/ietf-crisp-paper-22jun03.pdf>>. Acesso em: 05 ago. 2007.
- DECKER, S. et al. **The Semantic Web: The Roles of XML and RDF**. *IEEE Internet Computing*, IEEE Computer Society, v. 4, n. 5, p. 63–74, Setembro 2000.
- DECKER, S.; MITRA, P.; MELNIK, S. **Framework for the Semantic Web: An RDF Tutorial**. *IEEE Internet Computing*, IEEE Computer Society, v. 4, n. 6, p. 68–73, Novembro 2000.
- DEUTSCH, P. et al. **Architecture of the WHOIS++ service**. *RFC1835*. 1995.
- FALLSIDE, D. C.; WALMSLEY, P. **XML Schema Part 0: Primer. W3C Recommendation**. Disponível em: <<http://www.w3.org/TR/xmlschema-0/>>. Acesso em: 05 ago. 2007.
- FALTSTROM, P.; HOFFMAN, P.; COSTELLO, A. **Internationalizing Domain Names in Applications (IDNA)**. *RFC3490*. 2003.
- FENSEL, D. et al. : **An Ontology Infrastructure for the Semantic Web**. *IEEE Intelligent Systems*, IEEE Computer Society, v. 16, n. 2, p. 38–45, Março 2001.
- FRANK, M.; DECKER, S. **The Networked Semantic Desktop**. *ISWC 02. International Semantic Web Conference*, 2002.
- FRANKLIN, S.; GRAESSER, A. **Is it an Agent, or Just a Program? A Taxonomy for Autonomous Agents**. *Proceedings of the Workshop on Intelligent Agents III, Agent Theories, Architectures, and Languages*. Berlim: Springer-Verlag, 1996.
- GERICH, E. **Guidelines for Management of IP Address Space**. *RFC 1366*. 1992.
- HALL, E. **Defining and Locating DNS Domains in the Federated Internet Registry Service**. *RFC draft (draft-ietf-crisp-firs-dns-00)*. 2003.
- HARRENSTIEN, K.; STAHL, M.; FEINLER, J. **NICNAME/WHOIS version 2**. *RFC 954*. 1985.
- HARRENSTIEN, K.; WHITE, V. **NICNAME/WHOIS**. *RFC 812*. 1982.
- HOLLENBECK, S. **Extensible Provisioning Protocol (EPP)**. *RFC3730*. 2004.
- HOLLENBECK, S.; ROSE, M.; MASINTER, L. **Guidelines for the Use of Extensible Markup Language (XML) within IETF Protocols**. *RFC3470*. 2003.
- HORROCKS, I.; PATEL-SCHNEIDER, P. F. **Three Theses of Representation in the Semantic Web**. *Proceedings of the 12th International Conference on World Wide Web*, ACM, p. 39–47, Maio 2003.
- HOWARD, L. **An Approach for Using LDAP as a Network Information Service**. *RFC2307*. 1998.
- KARRENBERG, D. et al. **Development of the Regional Internet Registry System**. v. 4, n. 4, p. 17–29, *IP Journal*. Cisco. Dezembro 2001.

- KLYNE, G.; CARROLL, J. J.; MCBRIDE, B. **Resource Description Framework (RDF): Concepts and Abstract Syntax. W3C Recommendation.** Disponível em: <<http://www.w3.org/TR/rdf-concepts/>>. Acesso em: 05 ago. 2007.
- KLYNE, G.; CARROLL, J. J.; MCBRIDE, B. **Resource Description Framework (RDF): RDF Primer. W3C Recommendation.** Disponível em: <<http://www.w3.org/TR/rdf-primer/>>. Acesso em: 05 ago. 2007.
- MEGGINSON, D. **SAX: Simple API for XML.** Disponível em: <<http://www.sax.sourceforge.net/>>. Acesso em: 05 ago. 2007.
- MOCKAPETRIS, P. **DOMAIN NAMES - CONCEPTS and FACILITIES. RFC882.** 1983.
- MOCKAPETRIS, P. **Domain names: Implementation specification. RFC883.** 1983.
- MOCKAPETRIS, P. **Domain names - concepts and facilities. RFC1034.** 1987.
- MOCKAPETRIS, P. **DOMAIN NAMES - IMPLEMENTATION AND SPECIFICATION. RFC1035.** 1987.
- NAGAHASHI, K.; YOSHIDA, T.; KONDO, K. **IRIS - A Routing Registry (rreg) Type for the Internet Registry Information Service. RFC draft (draft-kengo-crisp-iris-rreg-00).** 2005.
- NEWTON, A. **Domain Administrative Data in Lightweight Directory Access Protocol (LDAP). RFC3663.** 2003.
- NEWTON, A. **Cross Registry Internet Service Protocol (CRISP) Requirements. RFC3707.** 2004.
- NEWTON, A. **An ENUM Registry Type for the Internet Registry Information Service (IRIS). RFC 4414.** 2006.
- NEWTON, A. **Replacing the Whois Protocol: IRIS and the IETF's CRISP Working Group. IEEE Internet Computing, IEEE Computer Society, v. 10, n. 4, p. 79–84, July 2006.**
- NEWTON, A. **A Lightweight UDP Transfer Protocol for the the Internet Registry Information Service. RFC draft (draft-ietf-crisp-iris-lwz-08).** 2007.
- NEWTON, A. **XML Pipelining with Chunks for the Information Registry Information Service. RFC draft (draft-ietf-crisp-iris-xpc-06).** 2007.
- NEWTON, A.; DAIGLE, L. **Cohabitation of the Nicname/Whois Protocol with the CRISP Protocol. RFC draft (draft-newton-whois-crisp-cohabitation-00).** 2003.
- NEWTON, A.; DAIGLE, L.; KOSTERS, M. **UWHO. Universal Whois.** 2001.
- NEWTON, A.; GUNDUZ, E.; KERR, S. **IRIS: An Address Registry (areg) Type IRIS: An Address Registry (areg) Type. RFC4698.** 2006.
- NEWTON, A.; NEVES, F. **Domain Registry Version 2 for the Internet Registry Information Service. RFC draft (draft-ietf-crisp-iris-dreg2-01).** 2006.

NEWTON, A.; SANZ, M. ***IRIS: A Domain Registry (dreg) Type for the Internet Registry Information Service (IRIS)***. RFC3982. 2005.

NEWTON, A.; SANZ, M. ***IRIS: The Internet Registry Information Service (IRIS) Core Protocol***. RFC3981. 2005.

NEWTON, A.; SANZ, M. ***Using the Internet Registry Information Service (IRIS) over the Blocks Extensible Exchange Protocol (BEEP)***. RFC3983. 2005.

REYNOLDS, J. K. et al. ***Internet Official Protocol Standards***. RFC3000. 2001.

ROSE, M. ***Mapping the BEEP Core onto TCP***. RFC 3081. 2001.

ROSE, M. ***The Blocks Extensible Exchange Protocol Core***. RFC 3080. 2001.

WAHL, M.; HOWES, T.; KILLE, S. ***Lightweight Directory Access Protocol (v3)***. RFC2251. 1997.

WILLIAMSON, S. et al. ***Referral Whois (RWhois) Protocol V1.5***. RFC2167. 1997.

Apêndice A – TABELAS DE PROPRIEDADES DAS CLASSES

Tabela 7: Lista de todas propriedades da classe
swhois:Dominio

Propriedade	rdfs:range
swhois:possuidDominio	rdfs:Literal
swhois:possuiFQDN	rdfs:Literal
swhois:possuiDN	rdfs:Literal
swhois:possuiVariante	rdfs:Literal
swhois:possuiContatoAbuse	swhois:Contato
swhois:possuiContatoAdm	swhois:Contato
swhois:possuiContatoCob	swhois:Contato
swhois:possuiContatoTec	swhois:contato
swhois:possuiContatoLegal	swhois:contato
swhois:possuiContatoSeg	swhois:contato
swhois:possuiContatoZona	swhois:contato
swhois:possuiOutroContato	swhois:Contato
swhois:possuiDataExp	xsd:Datetime
swhois:possuiDataUltVer	xsd:Datetime
swhois:possuiDataDelegInicial	xsd:Datetime
swhois:possuiDataUltAltContato	xsd:Datetime
swhois:possuiRespUltAltContato	swhois:Contato
swhois:possuiDataUltAltDNS	xsd:Datetime
swhois:possuiRespUltAltDNS	swhois:Contato
swhois:possuiDataUltAltDom	xsd:Datetime
swhois:possuiDataRenovacao	xsd:Datetime
swhois:possuiServidorDNS	swhois:ServidorDNS
swhois:possuiRegistrante	swhois:Contato
swhois:possuiRegistrar	swhois:Contato

continua...

Tabela 7: Lista de todas propriedades da classe *swhois: dominio*

Propriedade	rdfs:range
swhois:possuiRegistro	swhois:Contato
swhois:possuiRegistroRef	rdfs:Literal
swhois:possuiStatus	rdfs:Literal
swhois:possuiAddInfo	rdfs:Literal

Tabela 8: Lista de todas propriedades da classe *swhois:Contato*

Propriedade	rdfs:range
swhois:possuidContato	rdfs:Literal
swhois:possuidNeMail	rdfs:Literal
swhois:possuiEmail	rdfs:Literal
swhois:possuiNome	rdfs:Literal
swhois:possuidioma	rdfs:Literal
swhois:possuiOrganizacao	rdfs:Literal
swhois:possuiTelefone	rdfs:Literal
swhois:possuiFAX	rdfs:Literal
swhois:possuiEndereco	rdfs:Literal
swhois:possuiSIP	rdfs:Literal
swhois:possuiTradContato	rdfs:Literal
swhois:possuiTipoContato	rdfs:Literal
swhois:possuiDataCriacao	rdfs:Literal
swhois:possuiDataUltAlt	xsd:Datetime
swhois:possuiDataUltVerContato	xsd:Datetime
swhois:possuidNeMail	rdfs:Literal
swhois:possuiAddInfoContato	rdfs:Literal

Tabela 9: Lista de todas propriedades da classe *swhois:ServidorDNS*

Propriedade	rdfs:range
swhois:possuidServidor	rdfs:Literal
swhois:possuiNomeServidor	rdfs:Literal
swhois:possuiPv4	rdfs:Literal
swhois:possuiPv6	rdfs:Literal
swhois:possuiDataCriacaoSrv	xsd:Datetime
swhois:possuiContatoSrv	swhois:Contato
swhois:possuiInfoAddSrv	rdfs:Literal
swhois:possuiDataUltVerSrv	xsd:Datetime
swhois:possuiDataUltAltSrv	xsd:Datetime
swhois:possuiAddInfoContato	rdfs:Literal

Apêndice B – MODELO RDF SERIALIZADO NA LINGUAGEM N3

```
@prefix rdf:      <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs:     <http://www.w3.org/2000/01/rdf-schema#> .
@prefix xsd:     <http://www.w3.org/2001/XMLSchema#> .
@prefix swhois:  <http://sw.deri.org/swhois#> .

swhois:possuiDominioID
  a      rdf:Property ;
  rdfs:comment "Identificador unico do dominio no repositorio." ;
  rdfs:domain swhois:Dominio ;
  rdfs:label "Identificador unico de dominio" ;
  rdfs:range rdfs:Literal .

swhois:possuiAddInfo
  a      rdf:Property ;
  rdfs:comment "Informacao adicional sobre o dominio." ;
  rdfs:domain swhois:Dominio ;
  rdfs:label "Informacao Adicional" ;
  rdfs:range rdfs:Class .

swhois:possuiServidorDNS
  a      rdf:Property ;
  rdfs:comment "Servidor DNS." ;
  rdfs:domain swhois:Dominio ;
  rdfs:label "Servidor DNS do Dominio" ;
  rdfs:range swhois:ServidorDNS .

swhois:possuiContatoLegal
  a      rdf:Property ;
  rdfs:comment "Contato do dominio para assuntos juridicos." ;
  rdfs:domain swhois:Dominio ;
  rdfs:label "Contato Legal" ;
  rdfs:range swhois:Contato .

swhois:possuiRegistroRef
```

```

a      rdf:Property ;
rdfs:comment "Referencia de registro de dominio." ;
rdfs:domain swhois: dominio ;
rdfs:label "Referencia de Registro" ;
rdfs:range rdfs:Literal .

swhois:possuiRegistrante
a      rdf:Property ;
rdfs:comment "Registrante do dominio." ;
rdfs:domain swhois: dominio ;
rdfs:label "Registrante" ;
rdfs:range swhois:Contato .

swhois:possuiDataUltAlt
a      rdf:Property ;
rdfs:comment "Contact last modification date time." ;
rdfs:domain swhois:Contato ;
rdfs:label "Data de Ultima Alteracao do Contato" ;
rdfs:range xsd:dateTime .

swhois:possuiIPv6
a      rdf:Property ;
rdfs:comment "IPv6 do Servidor DNS." ;
rdfs:domain swhois:ServidorDNS ;
rdfs:label "IPv6 do Servidor DNS" ;
rdfs:range rdfs:Literal .

<http://sw.deri.org/shois#possuiAddInfoContato>
a      rdf:Property ;
rdfs:comment "Informacao adicional sobre o contato." ;
rdfs:domain swhois:Contato ;
rdfs:label "Informacao Adicional do Contato" ;
rdfs:range rdfs:Class .

swhois:possuiAddInfoSrv
a      rdf:Property ;
rdfs:comment "Informacao adicional sobre o servidor DNS." ;
rdfs:domain swhois:ServidorDNS ;
rdfs:label "Informacao Adicional do Servidor DNS" ;
rdfs:range rdfs:Class .

swhois:possuiEndereco
a      rdf:Property ;
rdfs:comment "Endereco do Contato." ;
rdfs:domain swhois:Contato ;
rdfs:label "Endereco" ;
rdfs:range rdfs:Literal .

```

```
swhois:possuiStatus
  a      rdf:Property ;
  rdfs:comment "Status." ;
  rdfs:domain swhois:Dominio ;
  rdfs:label "Status do Dominio" ;
  rdfs:range rdfs:Literal .

swhois:possuiVariante
  a      rdf:Property ;
  rdfs:comment "Variante de um dominio registrado." ;
  rdfs:domain swhois:Dominio ;
  rdfs:label "Variante de Dominio" ;
  rdfs:range rdfs:Literal .

swhois:ServidorDNS
  a      rdfs:Class ;
  rdfs:comment "Servidor de Nomes DNS (Domain Name System)." ;
  rdfs:isDefinedBy swhois: ;
  rdfs:label "Servidor DNS" .

swhois:possuiTelefone
  a      rdf:Property ;
  rdfs:comment "Telefone do contato." ;
  rdfs:domain swhois:Contato ;
  rdfs:label "Telefone" ;
  rdfs:range rdfs:Literal .

swhois:possuiFAX
  a      rdf:Property ;
  rdfs:comment "FAX do contato." ;
  rdfs:domain swhois:Contato ;
  rdfs:label "FAX" ;
  rdfs:range rdfs:Literal .

swhois:possuiEndereco
  a      rdf:Property ;
  rdfs:comment "Endereco do contato." ;
  rdfs:domain swhois:Contato ;
  rdfs:label "Endereco" ;
  rdfs:range rdfs:Literal .

swhois:possuiDataUltAltSrv
  a      rdf:Property ;
  rdfs:comment "Data da ultima alteracao do servidor DNS." ;
  rdfs:domain swhois:ServidorDNS ;
  rdfs:label "Data de Ultima Alteracao do Servidor" ;
  rdfs:range xsd:dateTime .
```

```
swhois:possuiContatoSrv
  a      rdf:Property ;
  rdfs:comment "Contato do Servidor DNS." ;
  rdfs:domain swhois:ServidorDNS ;
  rdfs:label "Contato do Servidor DNS" ;
  rdfs:range swhois:Contato .

swhois:possuiDataCriacao
  a      rdf:Property ;
  rdfs:comment "Data de criacao do contato." ;
  rdfs:domain swhois:Contato ;
  rdfs:label "Data de Criacao do Contato" ;
  rdfs:range xsd:dateTime .

swhois:possuiFDQN
  a      rdf:Property ;
  rdfs:comment "FQDN (Fully qualified domain name)." ;
  rdfs:domain swhois:Dominio ;
  rdfs:label "FQDN" ;
  rdfs:range rdfs:Literal .

swhois:possuiContatoCob
  a      rdf:Property ;
  rdfs:comment "Contato do dominio para assuntos de cobranca." ;
  rdfs:domain swhois:Dominio ;
  rdfs:label "Contato de Cobranca" ;
  rdfs:range swhois:Contato .

swhois:possuiDataCriacaoSrv
  a      rdf:Property ;
  rdfs:comment "Data de criacao do servidor DNS." ;
  rdfs:domain swhois:ServidorDNS ;
  rdfs:label "Data Criacao do Servidor DNS" ;
  rdfs:range xsd:dateTime .

swhois:possuiIdioma
  a      rdf:Property ;
  rdfs:comment "Contact Language." ;
  rdfs:domain swhois:Contato ;
  rdfs:label "Idioma do Contato" ;
  rdfs:range rdfs:Literal .

swhois:possuiContatoZona
  a      rdf:Property ;
  rdfs:comment "Contato do dominio responsavel pela delegacao DNS." ;
  rdfs:domain swhois:Dominio ;
  rdfs:label "Contato da Zona" ;
  rdfs:range swhois:Contato .
```

```
swhois:possuiDataUltAltDNS
  a      rdf:Property ;
  rdfs:comment "Data da ultima alteracao DNS do dominio." ;
  rdfs:domain swhois: dominio ;
  rdfs:label "Data Ultima Alteracao DNS" ;
  rdfs:range xsd:dateTime .
```

```
swhois:possuiTradContato
  a      rdf:Property ;
  rdfs:comment "Contato Traduzido." ;
  rdfs:domain swhois:Contato ;
  rdfs:label "Contato Traduzido" ;
  rdfs:range swhois:Contato .
```

```
swhois:possuiDataRenovacao
  a      rdf:Property ;
  rdfs:comment "Data da ultima renovacao do dominio." ;
  rdfs:domain swhois: dominio ;
  rdfs:label "Data da Ultima Renovacao" ;
  rdfs:range xsd:dateTime .
```

```
swhois:possuiNome
  a      rdf:Property ;
  rdfs:comment "Nome do contato." ;
  rdfs:domain swhois:Contato ;
  rdfs:label "Nome do contato" ;
  rdfs:range rdfs:Literal .
```

```
swhois:possuiDataUltVer
  a      rdf:Property ;
  rdfs:comment "Data da ultima verificacao DNS." ;
  rdfs:domain swhois: dominio ;
  rdfs:label "Data de Ultima Verificacao" ;
  rdfs:range xsd:dateTime .
```

```
swhois:possuiNomeServidor
  a      rdf:Property ;
  rdfs:comment "Nome do Servidor DNS." ;
  rdfs:domain swhois:ServidorDNS ;
  rdfs:label "Nome do Servidor DNS" ;
  rdfs:range rdfs:Literal .
```

```
swhois:possuiSIP
  a      rdf:Property ;
  rdfs:comment "SIP." ;
  rdfs:domain swhois:Contato ;
  rdfs:label "SIP" ;
```

```
rdfs:range rdfs:Literal .

swhois:possuiContatoSeg
  a      rdf:Property ;
  rdfs:comment "Contato do dominio para assuntos de seguranca." ;
  rdfs:domain swhois: dominio ;
  rdfs:label "Contato de Seguranca" ;
  rdfs:range swhois:Contato .

swhois:possuiDataUltVerSrv
  a      rdf:Property ;
  rdfs:comment "Result of the last DNS verification." ;
  rdfs:domain swhois:ServidorDNS ;
  rdfs:label "Data da Ultima Verificacao do Servidor" ;
  rdfs:range rdfs:Literal .

swhois:possuiIdContato
  a      rdf:Property ;
  rdfs:comment "Identificador Unico do contato no repositorio." ;
  rdfs:domain swhois:Contato ;
  rdfs:label "ID do Contato" ;
  rdfs:range rdfs:Literal .

swhois:possuiContatoTec
  a      rdf:Property ;
  rdfs:comment "Contato Tecnico." ;
  rdfs:domain swhois: dominio ;
  rdfs:label "Contato Tecnico" ;
  rdfs:range swhois:Contato .

swhois:possuiRespUltAltDNS
  a      rdf:Property ;
  rdfs:comment "Contato responsavel pela ultima alteracao DNS." ;
  rdfs:domain swhois: dominio ;
  rdfs:label "Responsavel pela Ultima Alteracao DNS" ;
  rdfs:range swhois:Contato .

swhois:possuiDataUltAltDominio
  a      rdf:Property ;
  rdfs:comment "Data da ultima alteracao do dominio." ;
  rdfs:domain swhois: dominio ;
  rdfs:label "Data da Ultima Alteracao DNS" ;
  rdfs:range xsd:dateTime .

swhois:possuiDataExp
  a      rdf:Property ;
  rdfs:comment "Data de expiracao do dominio." ;
  rdfs:domain swhois: dominio ;
```

```
rdfs:label "Data de Expiracao" ;
rdfs:range xsd:dateTime .

swhois:Dominio
  a      rdfs:Class ;
  rdfs:comment "Dominio." ;
  rdfs:isDefinedBy swhois: ;
  rdfs:label "Dominio" .

swhois:possuiOrganizacao
  a      rdf:Property ;
  rdfs:comment "Organizacao do contato." ;
  rdfs:domain swhois:Contato ;
  rdfs:label "Organizacao" ;
  rdfs:range rdfs:Literal .

swhois:possuiDataUltVerContato
  a      rdf:Property ;
  rdfs:comment "Data de Ultima Verificacao do Contato." ;
  rdfs:domain swhois:Contato ;
  rdfs:label "Data de Ultima Verificacao do Contato" ;
  rdfs:range xsd:dateTime .

swhois:Contato
  a      rdfs:Class ;
  rdfs:comment "Contato pode ser uma pessoa, grupo ou organizacao." ;
  rdfs:isDefinedBy swhois: ;
  rdfs:label "Contato" .

swhois:possuiRegistrar
  a      rdf:Property ;
  rdfs:comment "Registrar." ;
  rdfs:domain swhois:Dominio ;
  rdfs:label "Registrar" ;
  rdfs:range rdfs:Literal .

swhois:possuiContatoAbuso
  a      rdf:Property ;
  rdfs:comment "Contato do dominio para reportar casos de abuso." ;
  rdfs:domain swhois:Dominio ;
  rdfs:label "Contato para Casos de Abuso" ;
  rdfs:range swhois:Contato .

swhois:possuiContatoAdm
  a      rdf:Property ;
  rdfs:comment "Contato do dominio para assuntos administrativos." ;
  rdfs:domain swhois:Dominio ;
  rdfs:label "Contato Administrativo" ;
```

```
rdfs:range swhois:Contato .

swhois:possuiIDNeMail
  a      rdf:Property ;
  rdfs:comment "IDN (Internationalized Domain Name) email." ;
  rdfs:domain swhois:Contact ;
  rdfs:label "IDNeMail" ;
  rdfs:range rdfs:Literal .

swhois:possuiDataUltAltContato
  a      rdf:Property ;
  rdfs:comment "Data da ultima alteracao de contato do dominio." ;
  rdfs:domain swhois:Dominio ;
  rdfs:label "Data Ultima Alteracao de Contato" ;
  rdfs:range xsd:dateTime .

swhois:possuiRespUltAltContato
  a      rdf:Property ;
  rdfs:comment "Contato responsavel pela ultima alteracao do dominio." ;
  rdfs:domain swhois:Dominio ;
  rdfs:label "Contato Responsavel pela Ultima Alteracao" ;
  rdfs:range swhois:Contato .

swhois:possuiOutroContato
  a      rdf:Property ;
  rdfs:comment "Contato do dominio para qualquer outro assunto." ;
  rdfs:domain swhois:Dominio ;
  rdfs:label "Outro Contato" ;
  rdfs:range swhois:Contato .

swhois:possuiEmail
  a      rdf:Property ;
  rdfs:comment "Email do Contato." ;
  rdfs:domain swhois:Contato ;
  rdfs:label "Email do Contato" ;
  rdfs:range rdfs:Literal .

swhois:possuiTipoContato
  a      rdf:Property ;
  rdfs:comment "Tipo do Contato." ;
  rdfs:domain swhois:Contato ;
  rdfs:label "Tipo" ;
  rdfs:range rdfs:Literal .

swhois:possuiRegistro
  a      rdf:Property ;
  rdfs:comment "Registro." ;
  rdfs:domain swhois:Dominio ;
```

```
rdfs:label "Registro" ;  
rdfs:range rdfs:Literal .
```

```
swhois:possuiIDN  
  a      rdf:Property ;  
  rdfs:comment "IDN (Internationalized Domain Name)." ;  
  rdfs:domain swhois:Dominio ;  
  rdfs:label "IDN" ;  
  rdfs:range rdfs:Literal .
```

```
swhois:possuiIPv4  
  a      rdf:Property ;  
  rdfs:comment "IPv4 do Servidor DNS." ;  
  rdfs:domain swhois:ServidorDNS ;  
  rdfs:label "IPv4 do Servidor DNS" ;  
  rdfs:range rdfs:Literal .
```

```
swhois:possuiDelegInicial  
  a      rdf:Property ;  
  rdfs:comment "Data da delegacao inicial." ;  
  rdfs:domain swhois:Dominio ;  
  rdfs:label "Delegacao Inicial" ;  
  rdfs:range xsd:dateTime .
```

Apêndice C – CONSULTAS A SERVIDORES WHOIS UTILIZADAS DURANTE O DESENVOLVIMENTO DO PROTÓTIPO

C.1 REGISTRO BRASILEIRO DE NOMES DE DOMÍNIOS

```
% Copyright (c) Nic.br
% The use of the data below is only permitted as described in
% full by the terms of use (http://registro.br/termo/en.html),
% being prohibited its distribution, comercialization or
% reproduction, in particular, to use it for advertising or
% any similar purpose.
% 2007-05-16 17:41:11 (BRT -03:00)
```

```
domain:      usp.br
owner:       UNIVERSIDADE DE SAO PAULO
ownerid:     063.025.530/0001-04
responsible: TEREZA CRISTINA MELO DE BRITO CARVALHO
owner-c:     JCF286
admin-c:     JCF286
tech-c:      ANG149
billing-c:   JCF286
nserver:     bee.uspnet.usp.br 143.107.253.3
nsstat:      20070514 AA
nslastaa:    20070514
nserver:     reno.uspnet.usp.br 143.107.51.2
nsstat:      20070514 AA
nslastaa:    20070514
nserver:     bee08.uspnet.usp.br 143.107.253.5
nsstat:      20070514 AA
nslastaa:    20070514
nserver:     ns1.ansp.br
```

```

nsstat:      20070514 AA
nslastaa:   20070514
created:    19960101
changed:    20060704
status:     published

```

```

nic-hdl-br:  ANG149
person:      Andre Gerhard
e-mail:     cceadmin@usp.br
created:    20011217
changed:    20020411

```

```

nic-hdl-br:  JCF286
person:      Jairo Carlos Filho
e-mail:     jcarlosf@usp.br
created:    20011217
changed:    20060424

```

```

% Security and mail abuse issues should also be addressed to
% cert.br, http://www.cert.br/, respectively to cert@cert.br
% and mail-abuse@cert.br
%
% whois.registro.br accepts only direct match queries. Types
% of queries are: domain (.br), ticket, provider, ID, CIDR
% block, IP and ASN.

```

C.2 REGISTRO IRLANDÊS DE NOMES DE DOMÍNIOS

```

% Rights restricted by copyright;
% http://www.domainregistry.ie/copyright.html
% Do not remove this notice

```

```

domain:     deri.ie
descr:      National University of Galway
descr:      School/Educational Institution
descr:      Discretionary Name
admin-c:    AAI509-IEDR
admin-c:    AAY406-IEDR
tech-c:     RIH12-IEDR
renewal:    04-June-2007
status:     Active

```

```
nserver:      ns1.novara.ie
nserver:      ns2.novara.ie
source:       IEDR

person:       Christoph Bussler
nic-hdl:      AAI509-IEDR
source:       IEDR

person:       Christoph Bussler
nic-hdl:      AAI509-IEDR
source:       IEDR

person:       Register.ie Hostmaster
nic-hdl:      RIH12-IEDR
source:       IEDR
```

C.3 REGISTRO DE NOMES DE DOMÍNIOS .ORG

NOTICE: Access to .ORG WHOIS information is provided to assist persons in determining the contents of a domain name registration record in the Public Interest Registry registry database. The data in this record is provided by Public Interest Registry for informational purposes only, and Public Interest Registry does not guarantee its accuracy. This service is intended only for query-based access. You agree that you will use this data only for lawful purposes and that, under no circumstances will you use this data to: (a) allow, enable, or otherwise support the transmission by e-mail, telephone, or facsimile of mass unsolicited, commercial advertising or solicitations to entities other than the data recipient's own existing customers; or (b) enable high volume, automated, electronic processes that send queries or data to the systems of Registry Operator or any ICANN-Accredited Registrar, except as reasonably necessary to register domain names or modify existing registrations. All rights reserved. Public Interest Registry reserves the right to modify these terms at any time. By submitting this query, you agree to abide by this policy.

```
Domain ID:D52663-LROR
Domain Name:GREENPEACE.ORG
Created On:15-Jun-1992 04:00:00 UTC
Last Updated On:10-Mar-2007 21:45:01 UTC
Expiration Date:14-Jun-2012 04:00:00 UTC
Sponsoring Registrar:Network Solutions LLC (R63-LROR)
```

Status:CLIENT TRANSFER PROHIBITED
Registrant ID:22083437-NSI
Registrant Name:Greenpeace International
Registrant Organization:Greenpeace International
Registrant Street1:Ottho Heldringstraat 5
Registrant Street2:
Registrant Street3:
Registrant City:Amsterdam
Registrant State/Province:North-Holland
Registrant Postal Code:1066 AZ
Registrant Country:NL
Registrant Phone:+1.31205236
Registrant Phone Ext.:
Registrant FAX:+1.31205148151
Registrant FAX Ext.:
Registrant Email:Sjoerd.Jongens@GREENPEACE.ORG
Admin ID:22083438-NSI
Admin Name:Greenpeace International
Admin Organization:Greenpeace International
Admin Street1:Ottho Heldringstraat 5
Admin Street2:
Admin Street3:
Admin City:Amsterdam
Admin State/Province:North-Holland
Admin Postal Code:1066 AZ
Admin Country:NL
Admin Phone:+31.20514815
Admin Phone Ext.:
Admin FAX:+31.20514815
Admin FAX Ext.:
Admin Email:Sjoerd.Jongens@GREENPEACE.ORG
Tech ID:15325704-NSI
Tech Name:Greenpeace International
Tech Organization:Greenpeace International
Tech Street1:Keizersgracht 176
Tech Street2:
Tech Street3:
Tech City:1016 DW Amsterdam
Tech State/Province:
Tech Postal Code:99999
Tech Country:NL
Tech Phone:+1.31205236
Tech Phone Ext.:
Tech FAX:
Tech FAX Ext.:
Tech Email:Sjoerd.Jongens@GREENPEACE.ORG
Name Server:NS.GREENPEACE.ORG
Name Server:FW.GREENPEACE.ORG

Name Server:FWU.GREENPEACE.ORG
 Name Server:NSU.GREENPEACE.ORG
 Name Server:
 Name Server:

C.4 REGISTRO DE NOMES DE DOMÍNIOS .COM

Whois Server Version 2.0

Domain names in the .com and .net domains can now be registered with many different competing registrars. Go to <http://www.internic.net> for detailed information.

Domain Name: VERISIGN.COM
 Registrar: NETWORK SOLUTIONS, LLC.
 Whois Server: whois.networksolutions.com
 Referral URL: <http://www.networksolutions.com>
 Name Server: CARQUINEZ.VERISIGN.NET
 Name Server: GOLDENGATE-W2-INF6.VERISIGN.NET
 Name Server: NS1.CRSNIC.NET
 Status: clientTransferProhibited
 Updated Date: 24-apr-2007
 Creation Date: 02-jun-1995
 Expiration Date: 01-jun-2012

>>> Last update of whois database: Wed, 16 May 2007 20:45:18 UTC <<<

NOTICE: The expiration date displayed in this record is the date the registrar's sponsorship of the domain name registration in the registry is currently set to expire. This date does not necessarily reflect the expiration date of the domain name registrant's agreement with the sponsoring registrar. Users may consult the sponsoring registrar's Whois database to view the registrar's reported date of expiration for this

registration.

TERMS OF USE: You are not authorized to access or query our Whois database through the use of electronic processes that are high-volume and automated except as reasonably necessary to register domain names or modify existing registrations; the Data in VeriSign Global Registry Services' ("VeriSign") Whois database is provided by VeriSign for information purposes only, and to assist persons in obtaining information about or related to a domain name registration record. VeriSign does not guarantee its accuracy. By submitting a Whois query, you agree to abide by the following terms of use: You agree that you may use this Data only for lawful purposes and that under no circumstances will you use this Data to: (1) allow, enable, or otherwise support the transmission of mass unsolicited, commercial advertising or solicitations via e-mail, telephone, or facsimile; or (2) enable high volume, automated, electronic processes that apply to VeriSign (or its computer systems). The compilation, repackaging, dissemination or other use of this Data is expressly prohibited without the prior written consent of VeriSign. You agree not to use electronic processes that are automated and high-volume to access or query the Whois database except as reasonably necessary to register domain names or modify existing registrations. VeriSign reserves the right to restrict your access to the Whois database in its sole discretion to ensure operational stability. VeriSign may restrict or terminate your access to the Whois database for failure to abide by these terms of use. VeriSign reserves the right to modify these terms at any time.

The Registry database contains ONLY .COM, .NET, .EDU domains and Registrars.