UNIVERSIDADE CATÓLICA DE PELOTAS
CENTRO POLITÉCNICO
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA

Explorando Modelos Semânticos para Descoberta de Recursos na Computação Ubíqua

por
Renato Marques Dilli

Trabalho Individual I
TI-2009/1-002

Orientador: Prof. Dr. Adenauer Corrêa Yamin
Co-orientador: Prof. Dr. Luiz Antônio Moro Palazzo

Pelotas, agosto de 2009

AGRADECIMENTOS

Ao Instituto Federal Sul-rio-grandense pelo incentivo e apoio investidos em mim. Agradeço também ao Prof. Adenauer, meu orientador, sempre bem disposto a auxiliar e presente em todos os momentos.
Agradeço aos colegas do PPGINF que me receberam muito bem e caminham comigo neste percurso.
Aos funcionários e professores do Centro Politécnico da Universidade Católica de Pelotas pela seriedade e compromisso.

SUMÁRIO

LISTA DE FIGURAS

LISTA DE TABELAS

LISTA DE ABREVIATURAS E SIGLAS

API	Application Programming Interface
DAML	DARPA Agent Markup Language
DL	Description Logic
EXEHDA	Execution Environrnent for High Distributed Applications
FIPA	Foundation for Intelligent Physical Agents
GRR	Generic Rule Reasoner
ISAM	Infra-estrutura de Suporte à Aplicações Móveis
LAN	Local Area Network
MANETs	Mobile Ad hoc Networks
MDS	Monitoring and Discovery System
OWL	Web Ontology Language
PDA	Personal Digital Assistant
PerDis	Pervasive Discovery Service
RDF	Resource Description Framework
RDQL	RDF Data Query Language
RMI	Remote Method Invocation
SGML	Standard Generalized Markup Language
SPARQL	SPARQL Protocol and RDF Query Language
SQL	Structured Query Language
UBICOMP	Ubiquitous Computing
UPnP	Universal Plug and Play
URI	Universal Resource Identifier
VO	Virtual Organization
W3C	World Wide Web Consortium
WAN	Wide Area Network
XML	eXtensible Markup Language

RESUMO

Em ambientes ubíquos os recursos devem estar compartilhados para que possam ser acessados de qualquer lugar, e a qualquer momento. Nesta abordagem o processo de descoberta de recursos assume um importante papel em satisfazer adequadamente as requisições por recursos. Este trabalho realiza um estudo sobre mecanismos de descoberta de recursos no contexto da Computação Ubíqua e tecnologias de Web Semântica com a intenção de apontar as características iniciais de um modelo de descoberta de recursos baseado em *matching* semântico para o *middleware* EXEHDA. Para isto, são relacionados conceitos de arquitetura dos mecanismos de descoberta de recursos, visando selecionar características que irão compor o modelo a ser desenvolvido. Em trabalhos como de Tangmunarunkit, demostra que o uso de ontologias aumenta a expressividade no processo de descoberta de recursos, reafirmando a importância das tecnologias de Web Semântica na organização, e entendimento semântico entre as partes envolvidas no processo de consulta por recursos.

Palavras-chave: Computação Ubíqua, Computação Pervasiva, Descoberta de Recursos, Web Semântica.

TITLE: “EXPLORING SEMANTIC MODELS FOR RESOURCE DISCOVERY IN UBIQUITOUS COMPUTING”

ABSTRACT

In ubiquitous environments resources must be shared to be accessed from anywhere, at any time. In this approach the process of discovery of resources is an important role in adequately satisfy the requests for resources. This paper performs a study on mechanisms of discovered resources in the context of ubiquitous computing, and semantic web technologies with the intention of pointing the initial characteristics of a model for discovery of resources based on semantic matching for EXEHDA middleware. For this, concepts of architecture are related mechanisms for discovery of resources, to select features that will compose the model to be developed. In work such as Tangmunarunkit shows that the use of ontologies increases the expression in the discovery of resources, reaffirming the importance of Semantic Web technologies in the organization, and semantic understanding between the parties involved in the consultation process for resources.

Keywords: keyword-one, keyword-two, keyword-three, keyword-four.

1 INTRODUÇÃO

A computação Ubíqua foi citada pela primeira vez por Mark Weiser, cientista chefe do Centro de Pesquisa da Xerox, em 1991. Em seu artigo "O Computador para o Século Vinte e um"(WEISER, 1991) Weiser previa que no futuro o foco dos usuários ficaria voltado para a realização de seu trabalho, e não para a ferramenta utilizada, utilizando-se de computação sem perceber ou necessitar de conhecimentos técnicos.

Atualmente o acesso a rede está presente em praticamente qualquer lugar, desde escolas, universidades, cafés, aeroportos e até em praças. A evolução tecnológica atual tem concebido equipamentos de informática portáteis cada vez menores, com autonomias estendidas, possibilitando ao usuário ficar conectado por várias horas ininterruptas, facilitando sua mobilidade. Pode-se dizer que atualmente estamos com recursos tecnológicos para prover a computação transparente, presente em todo lugar e a qualquer momento. A fim de prover serviços ao usuário final de forma transparente, independentemente de tempo e espaço, sistemas ubíquos exploram os sensores e redes disponíveis (KINDBERG; FOX, 2002).

A abstração deste ambiente computacional ubíquo pode ser feita através de uma camada de software chamada *middleware*. Este software realiza o gerenciamento das aplicações, bem como o contexto atual e as adaptações necessárias ao contexto, devido as alterações que ocorreram no ambiente ubíquo.

Uma das funções dos *middlewares* é gerenciar os recursos disponíveis e dispersos pelo ambiente. Os recursos podem ser serviços e/ou dispositivos que podem entrar e sair do ambiente a qualquer momento. Cabe ao mecanismo de Descoberta de Recursos gerenciá-los.

A Descoberta de recursos em um ambiente ubíquo é de extrema importância para o funcionamento do ambiente, pois o *middleware* precisa saber quais recursos estão presentes no contexto e quais atendem as necessidades da aplicação para que possam ser alocados e descartados sem interferência do usuário.

1.1 Tema

Este trabalho visa o estudo de mecanismos de descoberta de recursos, revisando principalmente sua arquitetura, linguagem de descrição e consulta de recursos.

A computação ubíqua utiliza-se de recursos dispersos e heterogêneos, de forma muito dinâmica. Estas características fazem da descoberta de recursos um desafio a ser pesquisado para atender as diversidades do ambiente ubíquo.

As tecnologias de Web Semântica também serão estudadas neste trabalho para

agregar recursos semânticos em mecanismos de descoberta de recursos para promover um *matching* semântico, estendendo a capacidade de interpretar consultas sobre recursos existentes no ambiente.

1.2 Motivação

O grande número de recursos que devem ser gerenciados em um ambiente ubíquo é fortemente heterogêneo e dinâmico (SOLDATOS et al., 2006). O compartilhamento de recursos e informações em ambientes distribuídos faz parte das particularidades da computação ubíqua. Neste sentido, a maneira que estes recursos são descobertos e disponibilizados tem sido alvo de pesquisa em vários trabalhos atuais (SCHAEFFER, 2005),(LOPES, 2005),(ALLEMAND, 2006). Estudos atuais tem apontado o uso de tecnologias de Web Semântica no aprimoramento dos componentes do ambiente ubíquo, entre eles a consciência e adaptação de contexto e a descoberta de recursos. O serviço de descoberta é uma função chave para implementar aplicações conscientes de contexto para usuários móveis em ambientes de computação ubíqua (ZHU; MUTKA; NI, 2005).

Ontologias estão sendo utilizadas para prover um entendimento semântico entre duas ou mais partes, expandindo a compreensão de termos sub-entendidos, através das relações entre as palavras.

Nesta perspectiva nota-se a importância de integrar mecanismos de Web Semântica no processo de descoberta de recursos da computação ubíqua.

1.3 Objetivos

O objetivo principal deste trabalho é avaliar a utilização de tecnologias de Web Semântica e mecanismos de descoberta de recursos, no contexto da computação ubíqua, avaliando suas principais características e desafios de pesquisa. Será criado um modelo ontológico para ser utilizado pelo mecanismo de descoberta de recursos do *middleware* EXEHDA. No processamento deste modelo ontológico será possível a utilização de um domínio bem definido, evitando ambiguidade de conceitos e principalmente a utilização de semântica e inferência na consulta por recursos.

Os objetivos específicos são:

- estudar fundamentos teóricos sobre computação ubíqua;
- estudar fundamentos teóricos sobre descoberta de recursos, identificando tecnologias relacionadas;
- revisar os principais projetos em computação ubíqua, sistematizando os diferentes aspectos relativos aos mecanismos de descoberta de recursos utilizados;
- estudar o projeto EXEHDA, revisando seus fundamentos e a concepção dos diversos módulos de sua arquitetura;
- relacionar tecnologias de Web Semântica que podem ser integradas a um mecanismo de descoberta de recursos, a ser integrado ao EXEHDA;
- definir as linhas gerais a serem perseguidas na modelagem do mecanismo de descoberta de recursos a ser proposto.

1.4 Estrutura do Texto

O texto é composto por sete capítulos. O capítulo um descreve uma pequena introdução, o tema selecionado, a motivação para pelo tema escolhido e os objetivos propostos para o desenvolvimento deste trabalho.

No capítulo dois são apresentados conceitos sobre computação ubíqua, bem como alguns projetos relevantes.

No capítulo três é feito um estudo detalhado sobre tecnologias de Web Semântica, com atenção especial às ontologias e sua utilização.

O capítulo quatro descreve as características essenciais dos mecanismos de descoberta de recursos, e requisitos que devem ser satisfeitos na computação ubíqua pelos mecanismos de descoberta.

O capítulo cinco relaciona alguns mecanismos de descoberta de recursos, sendo classificados por utilizar ou não recursos de semântica na descrição e/ou consulta por seus recursos.

No capítulo seis são levantadas algumas características apontadas no estudo realizado que indicam ser adequadas de serem sugeridas ao mecanismo de descoberta de recursos do *middleware* EXEHDA.

No capítulo sete é feito o fechamento deste trabalho com algumas considerações finais.

2 COMPUTAÇÃO UBÍQUA: PRINCIPAIS CONCEITOS E PROJETOS

Este capítulo resume o estudo desenvolvido a respeito de computação ubíqua e relaciona alguns projetos de computação ubíqua. O *middleware* EXEHDA (YAMIN, 2004) é apresentado em mais detalhes porque o modelo de descoberta de recursos semântico que será proposto está direcionado ao ambiente ubíquo disponibilizado por este *middleware*.

Com a evolução dos Sistemas de Informação Distribuídos, com possibilidades de acesso a redes sem fio, através de diversas tecnologias, juntamente com a Computação Móvel em crescente avanço tecnológico e a disseminação de pequenos dispositivos com com acesso à Internet possibilitam o acesso de Sistemas de Informação a qualquer hora em praticamente qualquer lugar. Mark Weiser (WEISER, 1991) definiu o termo Computação Ubíqua pela primeira vez, através de seu artigo "*The Computer for the 21st Century*". Weiser teorizou que no futuro os usuários utilizariam recursos de computação sem perceber ou necessitar de conhecimentos técnicos. Ele também descreveu uma proliferação de dispositivos de diferentes tamanhos indo desde os dispositivos portáteis até grandes dispositivos compartilhados. Essa proliferação de dispositivos realmente aconteceu com os dispositivos utilizados normalmente e o desenvolvimento e produção da infra-estrutura necessária para suportar uma computação presente em qualquer lugar, a qualquer momento, está chegando.

O termo "Computação Pervasiva" muitas vezes é utilizado como sinônimo à Computação Ubíqua.

A computação "invisível" promovida pela computação ubíqua é realizada através de interfaces de comunicação diferente do tradicional teclado e mouse. As aplicações precisam ser sensíveis ao contexto e adaptarem seu comportamento através da informação capturada do ambiente.

A utilização de interfaces naturais, que possibilitem uma maior capacidade nas comunicações entre pessoas e computadores. O objetivos dessas interfaces naturais é suportar formas comuns de expressão humana. Esforços anteriores se focaram em interfaces de reconhecimento de voz e escrita, mas estas interfaces ainda não lidam de forma adequada com erros que ocorrem naturalmente com estes sistemas. Além disso estas interfaces são muito difíceis de serem implementadas. A computação ubíqua inspira o desenvolvimento de aplicações que não utilizam o *desktop*. Implícito a isto está a consideração que a interação física entre humano e computadores serão bem diferentes do *desktop* atual com teclado, mouse, monitor, e será mais parecida com a maneira que os humanos interagem com o mundo físico. Interfaces que suportem formas de computação humanas mais naturais (fala, escrita e gestos) estão começando a substituir os dispositivos mais tradicionais.

Estas interfaces se sobressaem por causa da sua facilidade de aprendizado e de uso. Além disso elas podem ser usadas por pessoas com deficiência física, para quem o tradicional mouse e teclado são menos acessíveis.

As aplicações para a computação ubíqua precisam ser sensitivas ao contexto, adaptando o seu comportamento baseando-se na informação adquirida do ambiente físico e computacional. Primariamente tivemos muitos avanços na área de localização e reconhecimento de identidade, mas ainda existem numerosos desafios na criação de representações de contexto reutilizáveis, e de reconhecimento de atividades. Duas demonstrações de computação ubíqua foram produzidas o laboratório de pesquisa da Olivetti e o Xerox Parctab, ambas demonstrações com aplicativos sensíveis a localização. Esses dispositivos forneciam a localização do usuário e proviam serviços interessantes como mapas, "sigame"automático e etc. Apesar da conexão entre dispositivos computacionais e o mundo físico não ser nova, (sistemas de controle de satélite e de mísseis são outros exemplos) esta simples aplicação sensível a localização é talvez a primeira demonstração ligando uma atividade humana implícita com serviços computacionais. As aplicações mais abrangentes são a navegação baseada em sistema GPS para carros e dispositivos portáteis que variam o conteúdo mostrado dando ao usuário a localização física dentro de uma área. Outra parte importante é o reconhecimento de objetos pessoais. Antigamente, sistemas se focavam no reconhecimento de algum tipo de código de barras ou etiqueta de identificação enquanto os trabalhos recentes incluem o uso de reconhecimento de imagem. Apesar de terem sido demonstrados vários sistemas que reconhecem a identidade da pessoa e sua localização eles ainda são difíceis de serem implementados.

Finalmente um grande número de aplicações na computação ubíqua dependem da captura automática de experiências reais, e, depois disso prover acesso flexível e universal para estas experiências. Uma grande parte da nossa vida é gasta escutando e gravando mais ou menos precisamente os eventos que estão a nossa volta e depois tentando lembrar partes importantes das informações adquiridas nesses eventos. Há um claro valor e um perigo em potencial em usar recursos computacionais para ajudar na "falta de memória"dos seres humanos, especialmente quando há múltiplas sequências de informação relacionada que são virtualmente impossíveis de serem absorvidas como um todo. Ferramentas que suportem gravação e acesso automático a experiências reais podem remover o fardo de fazer alguma coisa que os humanos não são bons de modo que possamos focar a atenção em atividades que nós somos bons.

A computação Ubíqua (UbiComp) pode ser definida como a integração entre a mobilidade, sistemas de reconhecimento de contexto e computação distribuída de forma invisível ao usuário.

2.1 Projetos

Para compor esta seção foram selecionados projetos representativos da área de computação ubíqua. Os mesmos tem como elemento comum na implementação de suas propostas uma camada de software denominada *middleware*. Esta camada tem o objetivo de facilitar o desenvolvimento das aplicações em ambientes distribuídos e ubíquos. O *middleware* é uma camada intermediária entre o ambiente de execução e as aplicações.

2.1.1 OXYGEN

O projeto (MIT PROJECT OXYGEN: PROJECT OVERVIEW, 2008) está sendo desenvolvido no *Massachusetts Institute of Technology* (MIT) e defende que no futuro a computação será disponível gratuitamente, em todo o lugar, como oxigênio no ar. O objetivo é prover computação pervasiva, centrada no usuário através de dispositivos móveis e estacionários conectados por uma rede auto-configurável. O projeto usa dispositivos baseados em computação embarcada e PDAs. Toda a interação com o usuário é feita por visão e voz, ao invés de teclado e mouse. Nas aplicações de Oxygen, é dada ênfase especialmente ao acesso automático e personalizado à informação, adaptando as aplicações às preferências e necessidades do usuário.

2.1.2 GAIA

O projeto GAIA (ROMAN et al., 2002) tem origem na Universidade de Illinois em Urbana-Champaign. Ele consiste de um *middleware* distribuído que coordena entidades de software e redes de dispositivos heterogêneos. A idéia do projeto é criar um metasistema operacional (Gaia OS) que abstrai os espaços e recursos como uma única entidade programável. O Gaia OS provê os principais serviços de um sistema operacional: execução de programas, operações de E/S, sistema de arquivos, comunicação, detecção de erros e alocação de recursos.

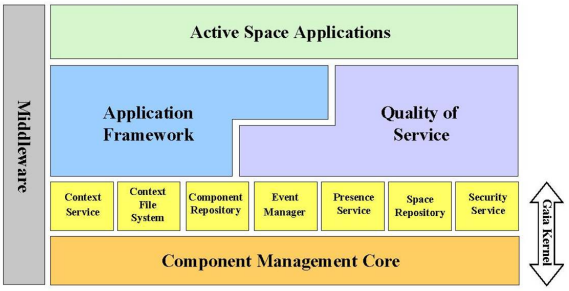

Figura 2.1: Arquitetura GAIA (ROMAN et al., 2002)

2.1.3 NEXUS

NEXUS (MICHAL JAKOB; GHANEA-HERCOCK, 2007) é um projeto do Ministério da Defesa, UK, Defense Technology Center, Nexus é impulsionada pelos objetivos da Network Enabled Capability (NEC).

Ele fornece suporte para descoberta, estruturação e fusão de informações principais permitindo a visão NEC. Ele é construído em cima de três conceitos: computação baseada em web services, uma arquitetura ponto-a-ponto e um serviços de composição automática. Combinando esses três conceitos, Nexus entrega uma ágil fusão de informação através de um conjunto de informações e fusão de serviços implantados em serviços de rede. Em adição aos três conceitos NEXUS interage diretamente com a camada de serviços, contendo os serviços disponibilizados pelo serviço de rede e a camada de aplicação que consiste de uma aplicação cliente do *middleware* NEXUS.

2.1.4 TSPACES

TSpaces (LEHMAN; MCLAUGHRY; WYCKOFF, 1999) é um *middleware* de rede para computação ubíqua. Uma sucinta descrição de TSpaces pode ser definido como um *buffer* de comunicação em rede com funcionalidades de banco de dados. Ele permite comunicação entre aplicações e dispositivos em uma rede com computadores e sistemas operacionais heterogêneos. TSpaces provê grupo de serviços de comunicação, serviços de banco de dados, serviços de transferências de arquivos baseados em URL, e serviços de notificação de eventos. Ele está implementado na linguagem de programação Java e, automaticamente possui rede ubíqua através da independência de plataforma, bem como um tipo padrão de representação para todos tipos de dados.

O sistema TSpaces é apropriado para qualquer aplicação que tem requisitos de distribuição ou armazenamento de dados. Ele pode realizar muitas das funções de um banco de dados relacional sem ser excessivamente restritivo (e primitivo).

O TSpaces pode ser considerado um sistema de banco de dados para computação diária que não gera consultas SQL complexas, mas possibilita um armazenamento confiável em rede.

2.1.5 ISAM/EXEHDA

O projeto está em desenvolvimento no II/UFRGS, sendo o único projeto de grande porte nessa área no país. O projeto ISAM – Infra-estrutura de Suporte às Aplicações Móveis (ISAM, 2009) foi concebido para execução de aplicações móveis distribuídas com comportamento adaptativo e considerando a computação pervasiva. O paradigma de programação empregado no projeto ISAM é o Holoparadigma. A execução das aplicações no ISAM é gerenciada pelo *middleware* EXEHDA.

O módulo de acesso pervasivo a dados e códigos disponibiliza o ambiente pervasivo denominado ISAMpe (ISAM *pervasive environment)*.

A organização do ISAMpe, apresentada na figura 2.2, contém as seguintes abstrações básicas (YAMIN, 2004):

- célula: representa a área de atuação de uma base. É composto por nodos e a própria base;
- base: é o módulo responsável pelos serviços básicos do ISAMpe e sua disponibilização aos nodos;
- nodo: responsáveis pela execução das aplicações. São todos os equipamentos que são empregados na execução das aplicações.

O EXEHDA (*Execution Environment for Highly Distributed Applications*) (YAMIN, 2004) é o *middleware* que disponibiliza a abstração do ambiente pervasivo deste trabalho. O EXEHDA, figura 2.3 é um componente da arquitetura ISAM. A arquitetura ISAM foi modelada baseada nos conceitos do modelo Holoparadigma, remodelado para o ambiente pervasivo.

As principais funcionalidades do EXEHDA são:

- gerenciar aspectos funcionais e não funcionais da execução das aplicações;
- suportar adaptações dinâmicas na execução das aplicações;

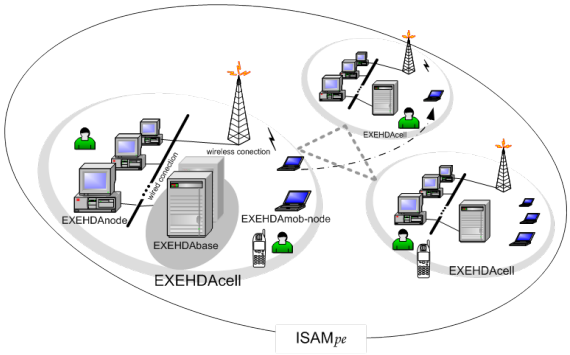

Figura 2.2: ISAMpe (ISAM, 2009)

- fornecer mecanismos para construir, gerenciar e disseminar informações de contexto;
- usar as informações de contexto em seus mecanismos de tomada de decisão;
- decidir, juntamente com as aplicações, a respeito de ações de adaptação;
- oferecer as usuários um comportamento que expresse a semântica *siga-me* das aplicações pervasivas, onde a aplicação segue o usuário em sua movimentação pelo ambiente ubíquo..

O núcleo mínimo do EXEHDA é composto por dois componentes:

- *ProfileManager:* responsável por interpretar as informações contidas nos perfis de xecução, tornando esses dados disponíveis para outros serviços em tempo de execução;
- *ServiceManager:* realiza a ativação de serviços em um nó, baseado nas informações providas pelo ProfileManager.

Os recursos da infra-estrutura física são mapeados para três abstrações básicas, as quais são utilizadas na composição do ISAMpe:

- EXEHDAcel: denota a área de atuação de uma EXEHDAbase, e é composta por esta e por EXEHDAnodes;
- EXEHDAbase: é o ponto de contato para os EXEHDAnodes. Uma EXEHDAbase é responsável por todos os serviços básicos de uma célula de execução e, embora constitua uma referência lógica única, seus serviços, sobretudo por aspectos de escalabilidade, poderão estar distribuídos entre vários equipamentos;
- EXEHDAnode: são os equipamentos de processamento disponíveis no ISAMpe, sendo responsáveis pela execução das aplicações. Um subtipo dessa abstração é o EXEHDAmob-node. Esses são os nós do sistema com elevada portabilidade, tipicamente dotados de interface de rede para operação sem-fio e, neste caso, integram a célula a qual seu ponto-de-acesso está subordinado. São funcionalmente análogos aos EXEHDAnodes, mas tipicamente são recursos com uma capacidade mais restrita (por exemplo, PDAs).

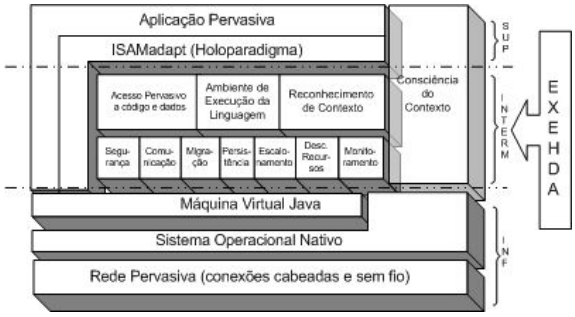

Figura 2.3: Middleware EXEHDA (YAMIN, 2004)

Os serviços do EXEHDA estão organizados em quatro grandes subsistemas: execução distribuída, adaptação, comunicação e acesso pervasivo. O serviço de descoberta de recursos do EXEHDA é realizado pelo Discoverer localizado no subsistema de execução distribuída. Este subsistema é composto também pelos seguintes serviços:

- *Executor:* realiza o disparo de aplicações, criação e migração de seus objetos.
- *Cell Information Base (CIB):* implementa a base de informações da célula, mantendo os dados estruturais da EXEHDACell, tais como, informações sobre os recursos que a compõe, informação de vizinhança, e atributos que descrevem as aplicações em execução.
- *ResourceBroker:* faz o controle da alocação de recursos às aplicações.
- *Gateway:* faz a intermediação das comunicações entre os nós externos a uma célula e os recursos internos a ela, podendo alternar a visibilidade dos recursos de uma célula quando vistos de fora dela.

O processo de gerenciamento de cada célula é autônomo em relação às outras células, e cada célula é responsável por gerenciar e prover acesso aos componentes computacionais locais, os quais podem ser dados, código, dispositivos, serviços ou outros recursos. Cada célula tem associada uma Cell Information Base (CIB), que mantém controle de toda a informação estática e dinâmica originada internamente À célula. Além disso, cada célula tem um conjunto dinâmico de outras células conhecidas no ISAMpe, o qual compõe sua vizinhança.

2.2 Considerações sobre o Capítulo

Este capítulo apresentou uma breve introdução à Computação Ubíqua, bem como requisitos que devem ser atendidos para caracterizar o ambiente ubíquo. Alguns projetos que visam criar uma abstração do ambiente pervasivo/ubíquo foram apresentados. No próximo capítulo serão apresentadas algumas tecnologias de Web Semântica para a criação e processamento de ontologias.

3 REVISANDO TECNOLOGIAS DE WEB SEMÂNTICA

Considerando o interesse central deste trabalho de explorar modelos semânticos para descoberta de recursos na computação ubíqua, o estudo da Web Semântica e de suas diferentes tecnologias se mostram um aspecto central para qualificação das soluções a serem buscadas. Este capítulo sintetiza o estudo feito envolvendo tecnologias de Web Semântica para a construção e processamento de ontologias.

A Web Semântica descrita por Tim Berners-Lee (BERNERS-LEE; HENDLER; LASSILA, 2001), é uma extensão da web atual, na qual a informação tem um significado bem definido, permitindo pessoas e computadores trabalharem em cooperação. Na Web atual a informação é processada pelos computadores a nível sintático, no futuro da Web Semântica será possível os computadores processar e raciocinar as informações no nível semântico. A Web Semântica faz uso de várias tecnologias para possibilitar esta automatização semântica. O W3C (*World Wide Web Consortium*) define e mantém uma arquitetura em camadas para Web Semântica, conforme a figura 3.1.

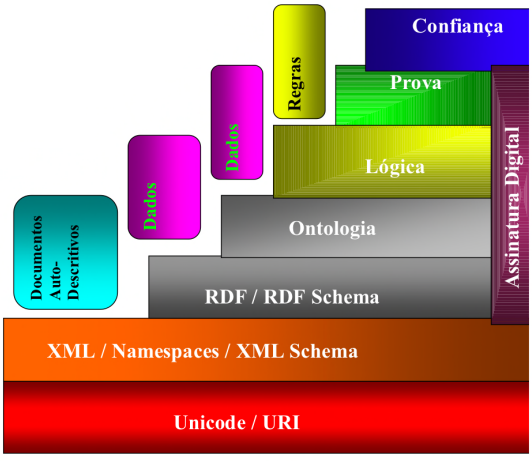

Figura 3.1: Camadas da Web Semântica proposta pela W3C

A camada denominada ***Unicode / URI*** fornece a interoperabilidade em relação à codificação de caracteres e ao endereçamento e nomeação de recursos da Web Semântica.

A camada denominada de ***XML / Namespace / XML Schema*** fornece a interoperabilidade em relação à sintaxe de descrição de recursos da Web Semântica.

A camada denominada ***RDF / RDF Schema*** fornece um *framework* para representar informação (metadados) sobre recursos.

A camada denominada de ***Ontologia*** fornece suporte para a evolução de vocabulários e para processar e integrar a informação existente sem problemas de indefinição ou conflito de terminologia. A linguagem RDF-Schema permite a construção de ontologias com expressividade e inferência limitadas, pois fornece um conjunto básico de elementos para a modelagem, e poucos desses elementos podem ser utilizados para inferência. A *Web Ontology Language* (OWL) estende o vocabulário da *RDF Schema* para a inclusão de elementos com maior poder com relação a expressividade e inferência.

A camada denominada ***Lógica*** fornece suporte para a descrição de regras para expressar relações sobre os conceitos de uma ontologia, as quais não podem ser expressas com a linguagem de ontologia utilizada.

As camadas denominadas de ***Prova e Confiança*** fornecem o suporte para a execução das regras, alem de avaliar a correção e a confiabilidade dessa execução. Essas camadas ainda estão em desenvolvimento e dependem da maturidade das camadas inferiores.

3.1 Lógica de Descrição

A Lógica de Descrição, ou *Description Logic* (DL) é uma evolução dos formalismos de representação do conhecimento baseados em objeto, ao qual corresponde um subconjunto estruturado da lógica de primeira ordem. Em termos gerais as lógicas de descrições são formalismos para representar conhecimento e raciocinar sobre ele. A OWL é baseada em lógica de descrição, havendo uma correspondência entre as linguagens XML para expressar ontologias e uma lógica de descrição.

Em DL, o conhecimento é representado por conceitos (predicados ou classes) e papéis (relacionamentos binários). Conceitos e papéis podem ser construídos utilizando construtores disponibilizados pela linguagem. Uma base de conhecimento DL consiste de uma TBox (*Terminological Knowledge*) e uma ABox (*Assertional Knowledge*). TBox contém definições de conceitos e axiomas (definem como conceitos e papéis estão relacionados). ABox contém o conhecimento extensional, que especifica indivíduos do domínio. Ele é a instanciação da estrutura de conceitos.

Na Figura 3.2 são ilustrados os construtores da lógica descritiva ALC (*Attribute Language with Complement*) (operador de conjunção/interseção, disjunção/união, negação, e os quantificadores “para todo” e “existe”).

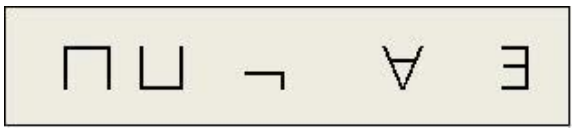

Figura 3.2: Construtores da Lógica de Descrição

A utilização de DL como método de representação do conhecimento possibilita a utilização de sistemas de raciocínio. Sistemas de raciocínio são sistemas que tem como objetivo processar conhecimento representado explicitamente e encontrar informações implícitas nestas informações, através de mecanismos específicos.

3.2 Ontologias

Uma ontologia é um modelo de dados que representa um conjunto de conceitos compartilhados (GRUBER, 1993). Ontologias são utilizadas para compartilhar informações de um domínio, composto de um vocabulário bem definido e com um entendimento comum e não ambíguo dos termos e conceitos utilizados pelas aplicações. Os elementos que formam uma ontologia são definidos pela tripla: Classes, Atributos e Relacionamentos. Um grande benefício no uso de ontologias é a possibilidade de descrever os relacionamentos entre os objetos. O conjunto destes relacionamentos é chamado de semântica. Este modelo semântico facilita o entendimento dos significados dos objetos pelas máquinas.

As ontologias não apresentam sempre a mesma estrutura, mas existem características e componentes básicos comuns presentes em grande parte delas. Mesmo apresentando propriedades distintas, é possível identificar tipos bem definidos.

Os componentes básicos de uma ontologia são classes (organizadas em uma taxonomia), relações (representam o tipo de interação entre os conceitos de um domínio), axiomas (usados para modelar sentenças sempre verdadeiras) e instâncias (utilizadas para representar elementos específicos, ou seja, os próprios dados) (Gruber, 1996; Noy & Guinness, 2001).

Algumas das propostas definem tipos de ontologias relacionando-as à sua função (MIZOGUCHI; VANWELKENHUYSEN; IKEDA, 1995), ao grau de formalismo de seu vocabulário (Uschold & Gruninger, 1996), à sua aplicação (Jasper & Uschold, 1999) e à estrutura e conteúdo da conceitualização (Van-Heijist, Schreiber & Wielinga, 1997), (Haav & Lubi, 2001).

Em (GUARINO, 1998) propõe uma classificação de ontologias sobre três aspectos:

- Pelo nível de detalhes:
  - Ontologias de referência (off-line)
  - Ontologias compartilháveis (on-line)
- Pelo nível de dependência de uma tarefa ou ponto de vista:
  - Ontologias de alto nível
  - Ontologias de domínio
  - Ontologias de tarefas
  - Ontologias de aplicações
- Ontologias de representação.

3.3 Construção e Processamento de Ontologias

Gruber (GRUBER, 1993) define quatro componentes que devem ser definidos para especificação de uma ontologia: classes, relações, funções e axiomas. Estes componentes são definidos da seguinte forma:

- Conceitos/Classes: ontologias são organizadas em taxonomias que podem ser interpretadas como sendo a identificação e a classificação dos termos presentes na ontologia;

- Relações: representam um tipo de interação entre conceitos e o domínio. Exemplos de relações binárias: “subclasse de” e “conectado a”;

- Funções: consistem de um caso especial de relações, onde o n-ésimo elemento do relacionamento é único para os n-1 elementos precedentes. Exemplo de funções: “Mãe-de”, associando um ou vários filhos à sua mãe;

- Axiomas: são sentenças formais que possuem o resultado lógico sempre verdadeiro.

Alguns outros termos importantes definidos no trabalho de (LOPES, 2005):

- Taxonomia: consiste de um conjunto de termos ou conceitos que organizados em uma hierarquia formam uma ontologia;

- Slots/Papéis/Propriedades (*Slots/Roles/Properties*): representam as várias características e atributos de um conceito;

- *Facets*: descrevem restrições nos *slots*;

- Instâncias: representam elementos.

3.3.1 XML

XML (eXtensive Markup Language) é uma linguagem de marcação recomendada pela W3C. Em meados da década de 1990, o World Wide Web Consortium (W3C) começou a trabalhar em uma linguagem de marcação que combinasse a flexibilidade da SGML (Linguagem Padronizada de Marcação Genérica) com a simplicidade da HTML. O XML é um formato para criação de documentos com dados organizados de forma hierárquica.

3.3.2 RDF/RDF Schema

O *Resource Description Framework* (RDF) é um padrão recomendado pelo W3C para descrever recursos da Web, desde fevereiro de 2004. Ele foi projetado para ser lido e entendido por computadores, é escrito XML e não foi construído para ser visualizado pelas pessoas.

RDF (KLYNE; CARROLL, 2004) identifica coisas usando identificadores Web (URIs), e descreve recursos com propriedades e valores das propriedades.

Um modelo RDF possui os seguintes objetos:

- Recursos: é qualquer coisa que tem uma URI, por exemplo, "http://www.w3schools.com/rdf"

- Propriedades: é um Recurso que tem um nome, por exemplo "homepage"

- Literais: é o valor de Propriedade é o valor da propriedade, por exemplo "htt://www.w3schools.com"

- Declaração: é a declaração de um recurso mais as propriedades desse recurso e o valor dessas propriedades.

A combinação de um Recurso, uma Propriedade e um Valor de Propriedade forma uma Declaração, também conhecido como sujeito, predicado e objeto.

Para a visualização da representação do RDF, pode-se utilizar grafos, a figura 3.3 mostra como é feita essa representação.

Figura 3.3: Representação RDF em grafo

3.3.3 OWL

Web Ontology Language (OWL) é construída no topo da RDF, utilizada para processamento de informações da web, foi construída para ser interpretada por computadores, e não para ser lida por pessoas. OWL (BECHHOFER et al., 2004) é escrita em XML, e é um padrão recomendado pelo W3C.

A OWL deriva das linguagens OIL (*Ontology Inference Layer*) (FENSEL; HORROCKS; VAN HARMELEN, 2000) e DAML (DARPA *Agent Markup Language*). A OIL foi a primeira linguagem direcionada à Web Semântica, sendo unida a linguagem DAML+OIL.

A linguagem OWL possui as características do RDF e um vocabulário maior que a DAML+OIL, oferendo mais recursos. A linguagem OWL foi contruída para ser utilizada por aplicações que necessitam realizar o processamento do significado das informações antes de apresentá-las aos usuários. Seu vocabulário permite a descrição de classes e propriedades, relacionamentos entre as classes, cardinalidade, igualdade, tipos e características de propriedades, entre outras funcionalidades.

Os recursos que a linguagem OWL oferece são divididos em três sublinguagens:

- OWL Lite: fornece uma hierarquia de classificação e funcionalidades de restrições simples. Por exemplo: a OWL Lite suporta cardinalidade, mas só permite os valores 0 e 1. Ela se torna mais fácil de ser implementada em uma ferramenta e faz com que a transição de outros modelos de vocabulários e taxonomias para OWL seja mais rápida.

- OWL DL: fornece uma maior expressividade e, ao mesmo tempo, seus sistemas mantêm a completude (garantia que todas as conclusões serão executadas) e decidibilidade (todos os cálculos terminarão em tempo finito) do sistema. A OWL DL inclui todos os artefatos da linguagem OWL, mas impõem restrições quanto à sua utilização.

- OWL Full: fornece a máxima expressividade e liberdade sintática do RDF. A OWL Full permite que uma ontologia aumente o significado do vocabulário (RDF ou OWL) predefinido. Não é esperado que nenhum software suporte todas as características da OWL Full.

OWL Full pode ser vista como uma extensão da RDF, enquanto OWL Lite e OWL DL são extensões de uma visão delimitada de RDF. Portanto, todo documento OWL é um documento RDF e todo documento RDF é um documento OWL Full. Mas somente alguns documentos RDF serão válidos em OWL Lite ou OWL DL.

OWL faz parte da "Visão de Web Semântica"(BERNERS-LEE; HENDLER; LASSILA, 2001), pois a informação disponibilizada na Web possui significado exato, pode ser processada por computadores, e computadores podem integrar a informação da web.

OWL difere de RDF por conter uma linguagem mais completa com maior capacidade de interpretação. OWL contém um grande vocabulário e sintaxe mais completa que RDF.

3.3.4 Protégé

Protégé (BIOMEDICAL INFORMATICS RESEARCH, 2009) é uma plataforma de código aberto, livre, que disponibiliza um conjunto de ferramentas para construir modelos e aplicações baseadas em conhecimento com ontologias. Protégé implementa um conjunto de estruturas e ações que suportam criação, visualização, e manipulação de ontologias em vários formatos de representação.

A plataforma Protégé suporta duas principais formas de modelar ontologias:

- Protégé-Frames editor: habilita usuários construir e popular ontologias que são baseadas em frames. Neste modelo, uma ontologia consiste de um conjunto de classes organizadas em uma hierarquia para representar um domínio de conceitos, um conjunto de slots associados às classes descrevem suas propriedades e relacionamentos, em um conjunto de instâncias destas classes.

- Protégé-OWL editor: habilita usuários construir ontologias para a Web Semântica, em particular em OWL. Uma ontologia OWL pode incluir descrições de classes, propriedades e suas instâncias.

3.3.5 SPARQL

SPARQL (SEABORNE, 2008) é uma linguagem de consulta e protocolo de acesso a dados em RDF, recomendada pela W3C. Seu nome é um acrônimo recursivo que significa SPARQL *Protocol and* RDF *Query Language*.

A especificação SPARQL funciona com outras tecnologias de web semântica da W3C. Entre elas, está a RDF, para representar dados; a RDF SChema; a Web Ontology Language (OWL), para construir vocabulários; e a Gleaning Resource Descriptions Dialects of Languages (GRDDL), para extração automática de dados semânticos de documentos.

Outros padrões, como o WSDL (Web Service Definition Language), também podem ser usados pela SPARQL.

3.3.6 Jena Toolkit

Jena (DICKINSON, 2008) é um Framework Java, de código aberto, desenvolvido por Brian McBride da HP Labs Semantic Web Programme, para desenvolvimento de aplicações com suporte à Web Semântica. A Jena suporta as linguagens RDF, RDFS, OWL e SPARQL, incluindo uma máquina de inferência baseada em regras.

O Framework Jena inclui:

- API para RDF
- Leitura e escrita RDF em RDF/XML, N3 e N-Triples
- API para OWL
- Armazenamento em memória ou de foram persistente
- Máquina de pesquisa SPARQL

A Jena permite criar e manipular grafos RDF, representados pelos recursos, propriedades e literais, formando tuplas que dará origem aos objetos criados pelo java. Esse conjunto de objetos é usualmente chamado de *model*. O *Model* é o conjunto de declarações que forma o grafo completo. Essa relação é demonstrada por (VERZULLI, 2001) na figura 3.4.

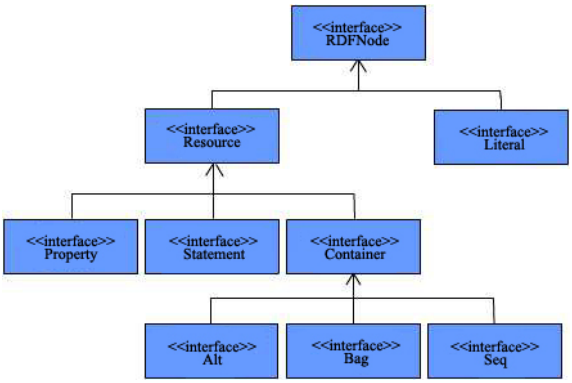

Figura 3.4: Hierarquia de Interfaces da API Jena (VERZULLI, 2001)

A arquitetura base da API Jena, figura 3.5, é composta por três camadas:

- *Ontology Model*: contém todas as classes necessárias para trabalhar com ontologias descritas em OWL, DAML, OIL ou RDFS. Neste módulo a classe mais relevante é a OntModel que representa um modelo ontológico. Esta classe no entanto é uma interface;
- *Graph Interface Reasoner*: permite fazer inferências sobre modelos OWL. O uso das inferências sobre modelos semânticos é permitir obter informação adicional (inferida) sobre as ontologias.
- *Graph Interface Base RDF*: OWL definida sobre RDFs. O Jena usa a API de RDF para manipular as ontologias.

As interfaces para programação são descritas da seguinte maneira:

- RDFNode: interface que possui a função de fazer a ligação de todos os outros elementos do RDF, a fim de formar as triplas para a criação dos elementos.
- Resource: interface que representa objetos que possuam um URI.

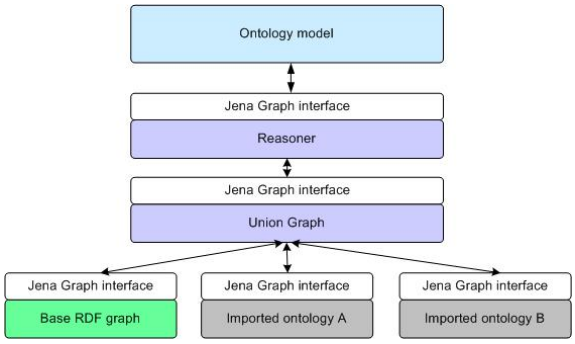

Figura 3.5: Camadas da API Jena (DICKINSON, 2008)

- Literal: interface que representa valores usados como objetos em triplas (sujeito, predicado, objeto). A interface Literal possui métodos para converter valores para vários tipos de dados em Java, como String, int e double.
- Property: interface que representa as propriedades em uma tripla.
- Statement: interface que representa uma tripla.
- Container: interface que representa um conjunto de objetos (Alt, Bag, Seq) em uma tripla.

O Jena fornece alguns motores de inferência e possibilita a criação de novos motores quando necessário, ou a possibilidade de estender os já existentes, abaixo é exibida uma breve descrição dos reasoners disponibilizados pelo Jena (REYNOLDS, 2009):

- *Transitive reasoner*: Disponibiliza suporte para armazenar e percorrer classes e propriedades ligadas. Implementa apenas as propriedades transitivas e simétricas de rdfs:subPropertyOf e de rdfs:subClassOf;
- RDFS *rule reasoner*: Implementa um subconjunto configurável das implicações RDFS;
- OWL, OWL Mini, OWL *Micro Reasoners*: Implementação incompleta da linguagem OWL-Lite;
- DAML *micro reasoner*: Usada internamente para viabilizar o uso da API legada de DAML, fornece uma capacidade mínima de inferência;
- *Generic rule reasoner*: Raciocinador baseado em regras que suporta a criação de regras definidas pelo usuário. Suporta encadeamento para frente (*forward chaining*), encadeamento para trás (*tabled backward chaining*) e estratégias de execução híbridas.

O Generic Rule Reasoner (GRR) é o mais independente dos motores de inferência do Jena, por isso terá uma descrição mais detalhada. Ele foi utilizado tanto para implementar o reasoner RDFS quanto o OWL, mas também possibilita que o programador importe as regras dos outros reasoners existentes. Desta maneira ele se torna o reasoner mais abrangente do Jena. Uma regra é definida como uma instância da classe Rule que

contém uma lista de premissas e uma lista de conclusões sobre as mesmas, opcionalmente a regra definida possui um nome e um sentido. Uma premissa ou uma conclusão pode ser uma tripla, uma tripla estendida ou uma chamada a procedimento primitivo (chamado *builtin*). O Jena (REYNOLDS, 2009) também disponibiliza um parser para checar a legalidade das regras definidas seguindo a sintaxe original, mas também permite que um outro parser seja definido pelo usuário para se obter um melhor diagnóstico dos erros encontrados, já que o parser disponibilizado pelo Jena não se demonstra muito eficiente nesse diagnóstico.

Como mencionado anteriormente, o GRR possui suporte a três tipos de mecanismos de ativação de regras:

- *Forward Chaining Engine* (Para frente): No momento em que o modelo de inferência recebe a primeira consulta, todos os dados relevantes do modelo são enviados para o mecanismo de regras. Quando uma regra causa a criação de triplas extras, novas regras podem ser disparadas. Nesse momento, se as regras não forem bem definidas, pode acontecer um loop infinito. Cada vez que são criadas ou removidas triplas do modelo pelos próprios métodos da API, as regras podem ser ativadas. A inferência acaba quando as regras pararem de ser ativadas. O algoritmo utilizado por este motor de inferência trabalha de forma incremental;

- *Backward Chaining Engine* (Para trás): No modo para trás o *reasoner* usa uma estratégia de execução parecida com o mecanismo do Prolog. No momento em que o modelo de inferência recebe uma consulta, ele a transforma em um objetivo, e o motor de inferência aplica as regras de modo a tentar atingir esse objetivo, unindo as triplas armazenadas com as regras de *backward chaining*. Neste caso, o motor de inferência não trabalha incrementalmente, isto é, sempre que os dados originais forem alterados, todo o processamento realizado é perdido;

- *Hybrid* (híbrido): Utiliza os dois mecanismos acima de forma conjunta. O mecanismo para frente executa primeiro e guarda um conjunto de deduções. Caso uma regra para frente crie novas regras para trás, ela vai instanciá-la de acordo com as variáveis guardadas nas deduções e depois irá passar as regras instanciadas para o mecanismo LP para trás. Todas as consultas são resolvidas posteriormente pelo LP engine usando a mistura dos dados brutos e das deduções que vieram do mecanismo para frente.

3.3.7 Raciocinadores

Os *reasoners*, (raciocinadores) são mecanismos computacionais criados para realizar inferências lógicas a partir de um conjunto de fatos ou axiomas, norteiam-se nas regras pré-definidas pela ontologia, descobrindo inconsistências, dependências escondidas e eventuais redundâncias. Através dos *reasoners* é possível que novos conhecimentos, além dos contidos explicitamente nas ontologias, sejam agregados e apresentados de forma transparente.

3.4 Considerações sobre o Capítulo

Este capítulo apresentou algumas tecnologias necessárias para a criação e manipulação de ontologias. Estas tecnologias possibilitam um entendimento semântico entre

máquinas, estendendo a capacidade de reconhecer conceitos e termos que não estão explicitamente definidos. No próximo capítulo são apresentados fundamentos de mecanismos de descoberta de recursos na visão de vários autores e características que estes mecanismos devem possuir para satisfazer as necessidades do ambiente ubíquo.

4 DESCOBERTA DE RECURSOS: FUNDAMENTOS E REQUISITOS

O processo de descoberta e organização de recursos no ambiente ubíquo é essencial para o compartilhamento dos recursos no ambiente heterogêneo e dinâmico.

O início do capítulo apresenta a visão de vários autores sobre as características, objetivos, arquitetura, linguagem de descrição e o formato utilizado para realizar as consultas.

Alguns trabalhos de descoberta de recursos são apresentados e classificados como baseados em *matching* sintático ou *matching* semântico.

Atualmente não existe um padrão na terminologia utilizada pelos mecanismos de descoberta para referenciar os recursos. Segundo McGrath (MCGRATH, 2000), dispositivos e serviços são recursos equivalentes.

Conforme (MARIN-PERIANU; HARTEL; SCHOLTEN, 2005) os protocolos de descoberta de recursos possuem os seguintes objetivos:

- descoberta: A habilidade de encontrar um provedor de serviço na rede de acordo com os critérios descritos pelo cliente. Para concretizar este requisito os protocolos necessitam utilizar uma linguagem de descrição para facilitar o processo de descoberta, onde as requisições de serviços (recursos) também são expressas utilizando esta linguagem. Os protocolos devem também armazenar as informações dos recursos. O armazenamento pode ser centralizado, distribuído ou híbrido. A procura pelos recursos devem ser endereçadas a um diretório ou disseminada na rede;

- transparência: A rede precisa organizar e disponibilizar informação sobre seu conteúdo sem intervenção humana. Para isto os protocolos precisam manter atualizadas as descrições dos recursos, alterando quando houver alterações nos recursos. A consistência deve ser mantida verificando a disponibilidade dos serviços.

Na computação pervasiva o mecanismo de descoberta deve atender aos seguintes requisitos (SCHAEFFER, 2005): utilização de informações do contexto de execução, utilização de estratégias para manutenção automática de consistência, expressividade na descrição de recursos e critérios de pesquisa, possibilidade de interoperabilidade com outras estratégias de descoberta, suporte à descoberta de recursos em larga-escala e utilização de preferências por usuário.

4.1 Revisando a Descoberta de Recursos na Computação Ubíqua

Um serviço é uma interface para uma aplicação ou dispositivo, pelo qual o cliente pode acessar a aplicação ou dispositivo. A localização e interação com serviços em um ambiente ubíquo é de responsabilidade do serviço de descoberta. Segundo (THOMPSON, 2006) o serviço de descoberta pode ser dividido em tarefas de descrição, disseminação, seleção e interação. Cada componente tem um conjunto único de responsabilidades para o realizar o serviço de descoberta.

Com cada vez mais dispositivos movendo-se dentro de um ambiente ubíquo, um grande número de dispositivos podem estar presentes em um dado ambiente a qualquer momento. A conectividade global, remove o senso de localização do mundo físico. Um dispositivo localizado do outro lado do mundo pode ser acessado da mesma forma se ele estivesse localizado na sala ao lado. por estas razões o número de dispositivos que o cliente pode acessar é muito grande, e cabe ao serviço de descoberta selecionar os recursos viáveis para serem utilizados pelo cliente.

Funções dos Componentes do Serviço de Descoberta:

- Descrição: responsável por balancear entre informação útil com o custo de armazenar e processar;
- Disseminação: precisa eficientemente distribuir informação, gerenciar recursos, e distribuir informação numa quantidade de tempo aceitável enquanto os recursos estão disponíveis;
- Seleção: precisa pegar a lista de opções aceitáveis e determinar qual é a melhor opção ou opções para completar a tarefa;
- Interação: provê comunicação significativa entre o cliente e os serviços.

Segundo (NABRZYSKI; SCHOPF; WEGLARZ, 2003) o processo de descoberta de recursos é executado em três passos:

- Passo 1: *Authorization Filtering* - Neste passo o usuário procura no sistema quais recursos ele tem acesso, como serviços básicos. Ao final deste passo o usuário terá uma lista de máquinas ou recursos que possui acesso;
- Passo 2: *Application requirement definition* - O usuário especifica um conjunto mínimo de requerimentos a fim de continuar a filtrar o conjunto de recursos viáveis;
- Passo 3: *Minimal requirement filtering* - Recursos que não preenchem os requisitos mínimos que foram dados no passo 2 são eliminados.

Na descoberta de recursos baseada em *query* deve-se prestar atenção nos seguintes fatores:

- *Query Language and Advertising Language:* Linguagem com suporte a formulação de consultas, bem como, descrição de serviços e recursos é crucial para o processo de descoberta. Isto é importante para avaliar a expressividade da linguagem e sua facilidade em formular consultas e publicações utilizadas durante o processo de descoberta. Atenção especial deve ser dada para o suporte semântico que a linguagem fornece para expressar diferentes aspectos de publicação e consultas;

- *Scalability:* É importante analisar como um sistema reage quando há um crescimento em uma ou mais de suas dimensões. A escalabilidade do mecanismo de descoberta é afetado pela escalabilidade de subcomponentes como processadores e dispositivos de armazenamento;

- *Reasoning support:* Um mecanismo de descoberta automático pode ser melhorado significadamente se o processamento das descrições da máquina são melhorados. Estas descrições são checadas se elas são equivalentes umas com as outras. Novo conhecimento pode ser inferido baseado em fatos existentes e este conhecimento pode ser adicionado durante o processo de descoberta.

4.1.1 Arquitetura

Uma estratégia de descoberta de recursos costuma apresentar uma arquitetura com dois ou três componentes. No primeiro caso, figura 4.1, são empregados apenas Resource Components (RCs) e User Components (UCs). No segundo caso, figura 4.2, acrescenta-se a uma terceira estrutura chamada Directory Component (DC).

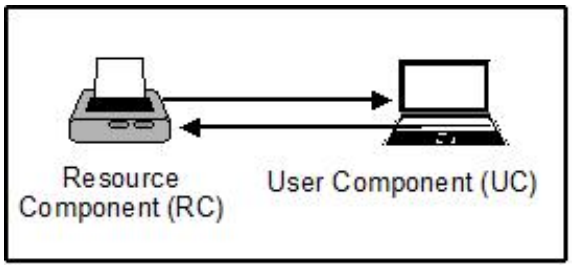

Figura 4.1: Descoberta de Recursos: Arquitetura de 2 componentes

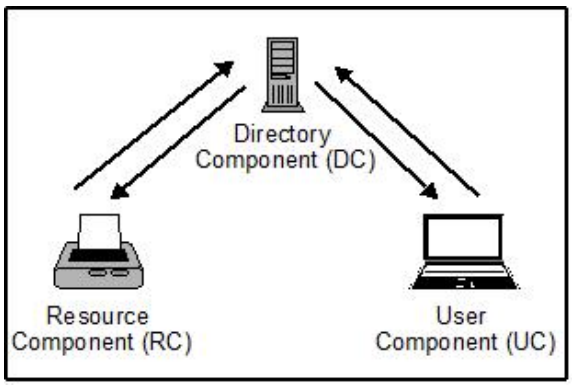

Figura 4.2: Descoberta de Recursos: Arquitetura de 3 componentes

O diretório (DC) é uma entidade que gerencia as informações dos serviços disponíveis na rede. Os mecanismos de descoberta podem adotar uma abordagem centralizada, baseada em diretórios, os quais são responsáveis pelo processamento de anúncios e consultas em nome de todos os dispositivos da rede, ou distribuída, onde cada dispositivo é responsável por manter as informações sobre os serviços que disponibiliza ou conhece.

As arquiteturas centralizadas podem utilizar um ou mais diretórios para tratar o registro das informações de serviços e as solicitações dos clientes, subdividindo-se em

estruturas hierárquicas e planas. no modelo hierárquico, os diretórios estão organizados em uma estrutura em árvore, similar à estrutura adotada pelo DNS. No modelo plano, os diretórios trocam informações entre si sobre os serviços que cada um gerencia.

A vantagem desta técnica é que o gerenciamento dos recursos distribuídos é centralizada, o controle de autoridade e qualidade é facilmente gerenciada e a descrição do serviço é consistente no estilo e apresentação.

Como desvantagem está a pouca escalabilidade, pois há um limite do número de recursos que podem ser gerenciados e registrados no registro global. Pode ocorrer também um alto tempo de resposta com acessos paralelos ao registro central.

As arquiteturas distribuídas baseiam-se na interação direta entre clientes e provedores de serviços, não existindo a sobrecarga da administração do diretório, como ocorre nas arquiteturas centralizadas. Nas arquiteturas distribuídas, as entidades produzem um maior volume de tráfego de mensagens de controle e o processamento das requisições é transferido dos diretórios para os clientes do sistema. Alguns protocolos oferecem suporte tanto à arquitetura centralizada quanto à distribuída.

(MARIN-PERIANU; HARTEL; SCHOLTEN, 2005) define uma arquitetura híbrida da arquitetura centralizada e distribuída. Cada VO (Virtual Organization) tem seu próprio registro local e toda organização é gerenciada por um registro global centralizado.

Esta técnica tem o benefício de compartilhar o gerenciamento e a administração dos registros locais e globais. As VOs são responsáveis pelo gerenciamento de seus próprios recursos, mas o registro global é de responsabilidade de gerenciamento de alto nível. Quando um serviço está ausente no registro local dentro de uma VO, a requisição é enviada para o registro global requisitando onde encontrar este serviço. Ele então pode contactar diretamente a VO responsável.

Esta técnica escala bem até o limite da capacidade de armazenagem do registro global.

Os recursos são administrados duas vezes, nos registros locais de cada VO e também no registro global. Se o registro global falhar, não será possível comunicar entre as VOs. Entretanto, um registro global replicado é necessário para garantir um sistema totalmente funcional.

4.1.2 Escopo da Descoberta

Entende-se por "Escopo da Descoberta" o conjunto de serviços que podem ser descobertos por um cliente. O conceito de escopo possibilita que os serviços disponíveis em uma rede possam ser observados sob diferentes perspectivas.

A definição do escopo da descoberta pode ser feita com base na topologia da rede, nas permissões de acesso embutidas nos diferentes usuários de um domínio administrativo e nas informações de contexto.

Em protocolos de descoberta distribuídos, o escopo é geralmente definido em função da topologia da rede e o alcance do mecanismo de descoberta é determinado, em parte, pelas regras de roteamento.

Nos protocolos de descoberta centralizados, baseados na estrutura de diretórios, o escopo pode ser expandido interconectando-se diretórios em diferentes domínios administrativos com relações de confiança previamente estabelecidas.

4.1.3 Descrição dos Recursos

Descrição de recurso é uma abstração das características e facilidades inerentes a um recurso. Essa abstração torna-se necessária em diferentes etapas do processo de descoberta:

- Quando um recurso, na arquitetura centralizada, é registrado pelo seu provedor junto ao diretório;
- Quando um provedor, na arquitetura distribuída, divulga na rede, através de mensagens de anúncios, a disponibilidade dos serviços que oferece;
- Quando o serviço é requisitado por outros dispositivos ou mesmo por outros recursos.

Gonzalez-Castillo et al. em (GONZALEZ-CASTILLO; TRASTOUR; BARTOLINI, 2001) identificou os seguintes requisitos que devem ser atendidos pela linguagem para expressar as descrições dos serviços:

- Alto grau de flexibilidade e expressividade: É a capacidade da linguagem permitir que a publicidade seja composta de vários graus de complexidade e completude. Dependendo das propriedades de um serviço que precisa de ser descrito, algumas propriedades podem ser expressas com simples pares atributo/valor, outros podem necessitar de mais estruturas. Os provedores de serviços podem querer descrever determinados aspectos do serviço a um grande nível de detalhamento, mas podem querem não especificar certas propriedades porque não são conhecidas, não aplicáveis ou porque precisam de ser negociadas mais tarde;
- Suporte a relacionamentos entre conceitos: Para o *matcmaking* realizar *matches* complexos, baseado em relacionamentos, a linguagem precisa permitir o relacionamento entre conceitos;
- Suporte a tipos de dados: Atributos como quantidades, datas e preços podem fazer parte da descrição do serviço. Por isso a linguagem de descrição precisa suportar tipos de dados para facilitar a expressividade e o emparelhamento (*matching*) de atributos no serviço de descrições;
- Expressar restrições e limitações: Serviços e requisições precisam ser descritos com definições conceituais de instâncias aceitáveis ao invés de uma simples instância do serviço. A linguagem de descrição deve facilitar a descrição entre limites sobre certos parâmetros;
- Nível de concordância semântico: Para cada *matchmaker* entender e comparar diferentes descrições de serviços, eles precisam compartilhar a mesma semântica. Para isto é necessário o uso de ontologias comuns na descrição dos serviços, permitindo interoperabilidade entre diferentes provedores de serviços e requisitores.

Clientes procuram por um determinado recurso preenchendo valores nos campos do serviço de descrição. Alguns protocolos utilizam pares de atributos-valores hierárquicos, como (BALAZINSKA; BALAKRISHNAN; KARGER, 2002), muitos descrevem seus recursos em XML, como (SCHAEFFER, 2005). Alguns utilizam ontologias para

uma melhor classificação e expressividade, tal como, (TANGMUNARUNKIT; DECKER; KESSELMAN, 2003).

Consultas de usuários compostas de pares de atributos e valores requerem que o conteúdo da comparação satisfaça todos os pares presentes na requisição (MARIN-PERIANU; HARTEL; SCHOLTEN, 2005). Se um atributo não é especificado, geralmente considera-se qualquer valor. Alguns protocolos possibilitam a seleção com vários critérios, na qual resulta uma coleção de nodos que satisfazem mais de uma condição.

4.1.4 Gerenciamento da Informação

O método de consulta utilizado em um protocolo de descoberta de serviços pode ser entendido como um processo de busca atrelado a um casamento (*matching*) entre as requisições de descoberta (demandas) e as descrições de serviços locais e remotos (ofertas), estas últimas divulgadas através de anúncios ou registradas em diretórios. O algoritmo de *matching* utilizado pelo mecanismo de consulta representa uma função que aceita, como entrada, a demanda e um conjunto de descrições de ofertas, provendo, como resultado, o subconjunto das ofertas que satisfazem a demanda especificada.

Em (DECKER; WILLIAMSON; SYCARA, 1996) é destacada a diferença entre os termos *matchmaking* e *brokering*. *Matchmaking* é um processo que permite que um agente com certo objetivo possa se comunicar diretamente com outros agentes para satisfazer a este objetivo, envolvendo três diferentes agentes:

- Consumidor: um agente com um objetivo que deve ser atingido por algum outro agente;
- *Matchmaker*: um agente que sabe como se comunicar com outros agentes e sabe de suas capacidades;
- Provedor de Recursos: um *matchmaking*agente que disponibiliza suas capacidades para serem utilizadas por outros agentes.

O processo de *brokering* é formado por uma agente com determinado objetivo que pode ter este objetivo sendo alcançado por outro agente. O processo de *brokering* também envolve três agentes diferentes:

- Consumidor: um agente com um objetivo que deve ser atingido por algum outro agente;
- *Broker*: um agente que sabe como se comunicar com outros agentes e sabe de suas capacidades. É responsável por selecionar os agentes que melhor atendam à solicitação de um serviço;
- Provedor de Recursos: um agente que disponibiliza suas capacidades para serem utilizadas por outros agentes.

No *matchmaking* a decisão de qual agente irá prover o serviço é tomada pelo cliente, com este tendo a responsabilidade de negociar com o provedor de recursos diretamente para a realização do serviço, já no *brokering* esta seleção é feita pelo *broker*.

O algoritmo de *matching* pode empregar diferentes funções de comparação, conforme representado na figura 4.3: uma função simples, baseada na comparação de atributos (influenciada pela descrição baseada em pares atributo-valor), uma função semântica

(influenciada pela uso de linguagens de descrição) ou uma função dependente de linguagem de programação.

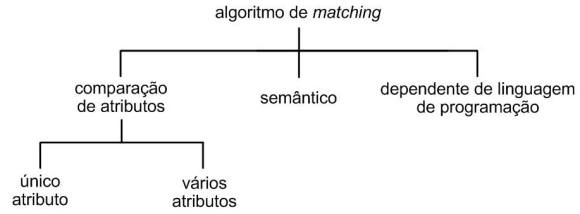

Figura 4.3: Algoritmos de matching para consulta de recursos

Algoritmos de *matching* baseados na comparação sintática de atributos podem ser implementados através da comparação de um único atributo, geralmente o identificador ou o tipo do serviço, ou de vários atributos. Algoritmos de *matching* baseados na função de comparação semântica beneficiam-se dos mecanismos de consulta embutidos nas linguagens de descrição, provenientes da utilização de padrões relacionados ao XML, os quais oferecem mecanismos para se representar a forma como os recursos se relacionam, incluindo propriedades como domínio e cardinalidade, e restrições de integridade, o que garante consultas mais poderosas. A representação semântica dos serviços consiste de um conjunto de informações que abrangem, mas não se limitam, a descrição de suas capacidades, funcionalidades, portabilidade e requisitos do sistema – como largura de banda, sistema operacional e processador. Uma função de *matching* semântico introduz a possibilidade de se obter resultados aproximados em resposta a uma requisição de serviço. Nesse caso, dependendo dos requisitos definidos na consulta, o resultado da função de *matching* pode ser satisfatório mesmo se uma, ou mais, características do serviço não sejam satisfeitas integralmente. Por exemplo, se uma requisição especifica um serviço que possa ser executado em sistemas baseados em um processador de uma determinada família (por exemplo, Intel Pentium) e uma instância desse serviço, compatível com sistemas baseados nessa família, for descoberta, um resultado aproximado é encontrado. Por fim, existem abordagens que atrelam a função de comparação à linguagem de programação utilizada. Este é o caso de Jini, que apresenta um forte acoplamento à linguagem de programação Java. Como resultado do processamento da requisição de descoberta pelo algoritmo de *matching*, são obtidas informações sobre o serviço requisitado, as quais permitirão a sua invocação e utilização. Essas informações são encaminhadas ao dispositivo que originou a requisição, encapsuladas em uma mensagem de resposta à solicitação do serviço.

4.1.5 Requisição e Anúncio

O mecanismo de descoberta pode consultar o diretório sobre um determinado recurso ou acessar o recurso desejado diretamente.

Basicamente, existem duas formas para se consultar informações sobre os serviços disponíveis na rede: a descoberta passiva, onde os provedores anunciam periodicamente os seus serviços para toda a rede, e a descoberta ativa, onde o cliente envia mensagens de descoberta para provedores ou diretórios com o intuito de obter informações sobre um serviço específico ou sobre todos os serviços disponíveis. Essas abordagens também são conhecidas, na literatura, como proativa e reativa, respectivamente, em uma alusão à classificação utilizada com os protocolos de roteamento planos para redes sem fio ad

hoc de saltos múltiplos. Grande parte dos protocolos de descoberta implementa ambas as abordagens.

4.1.6 Armazenamento das Informações

Informações de serviço são divulgadas pelo provedor do serviço (ou diretório) através de mecanismos de anúncio ou em resposta a mensagens de requisição. A informação de um serviço é utilizada para descrever, identificar e selecionar o serviço na rede. A informação pode incluir o nome do serviço, o seu identificador, o endereço do provedor do serviço, o protocolo que o cliente e o provedor devem utilizar para invocar o serviço, o período no qual a informação é válida, entre outros itens. As informações sobre os serviços devem ser armazenadas em repositórios, de modo que os demais dispositivos possam recuperá-las e contactar o provedor de um serviço particular. De acordo com o modo como esses repositórios são organizados, eles são classificados como centralizados ou distribuídos. Na primeira abordagem, as informações sobre todos os serviços disponíveis na rede são armazenadas em diretórios. Na segunda abordagem, todos os dispositivos da rede possuem o seu próprio repositório local, onde armazenam informações sobre os serviços que disponibilizam, assim como sobre os serviços divulgados na rede. A abordagem distribuída pode ainda ser subdividida em cooperativa, quando os nós mantêm informações parciais sobre os serviços da rede, que se complementam, e não cooperativa, quando os dispositivos armazenam as informações sobre todos os serviços da rede, mantendo uma visão global do sistema.

4.1.7 Métodos de Seleção e Invocação

Quando o resultado de uma requisição apresenta mais de um provedor do mesmo serviço é adequado que o melhor provedor seja selecionado sem a interferência do usuário. Os protocolos de descoberta podem oferecer opções de seleção manual, com a intervenção direta do usuário da aplicação, ou automática, através de um algoritmo implementado no lado cliente, selecionando a melhor opção em função das características da aplicação. Nos protocolos de descoberta que utilizam repositórios centralizados, o algoritmo de seleção pode ser implementado no próprio diretório, que se encarrega de fazer a seleção do melhor serviço, em nome do cliente, de acordo com critérios especificados na requisição. Quando a abordagem de seleção automática é implementada, deve-se considerar as métricas que irão definir qual é a melhor oferta. Uma métrica possível é a menor distância, em número de saltos, entre o dispositivo que originou a requisição e o provedor do serviço.

4.1.8 Suporte à Mobilidade

O suporte à mobilidade implica que a informação sobre os serviços disponíveis na rede, quer seja armazenada nos diretórios (centralizada) ou em cada dispositivo da rede (distribuída), deva ser atualizada em função das modificações na topologia da rede, provocadas pela mobilidade dos dispositivos. Se, em um dado grupo, um dispositivo armazena informações sobre os serviços dos demais dispositivos, espera-se que ele mantenha informações corretas, na medida do possível. Se o dispositivo altera a sua posição em relação ao grupo, ou se algum membro do grupo se desloca, as informações de serviços devem ser atualizadas o mais rapidamente possível. Somente dessa forma, pode-se esperar, com uma certa probabilidade, que um protocolo de descoberta consiga descobrir serviços em tempo

hábil. Caso as informações mantidas por um dispositivo sobre os serviços da rede sejam obsoletas, há uma grande chance desse dispositivo responder a uma requisição de serviço com informações desatualizadas. Ao receber a resposta a sua solicitação, o cliente tentará acessar o serviço e, só então, irá detectar a sua indisponibilidade, pois o provedor do serviço pode ter se deslocado ou se tornado inalcançável. Existem dois métodos principais para solucionar esse problema: proativos e reativos. No método proativo, os dispositivos mantêm uma visão atualizada das informações sobre os serviços disponíveis na rede com a troca periódica de mensagens de anúncio contendo dados mais recentes sobre o serviço. No método reativo, a informação é atualizada em razão da ocorrência de eventos na rede, como, por exemplo, a indisponibilidade de uma rota para um provedor ou a detecção de mudanças de estado informadas pelo próprio serviço.

4.2 Descoberta com Semântica

Um dos problemas em mecanismos de descoberta baseados em palavras-chave é que eles não conseguem capturar completamente a semântica da consulta do usuário porque eles não consideram os relacionamentos entre as palavras.

Um mecanismo de descoberta de recursos que utiliza tecnologias de web semântica, em especial ontologias, pode facilmente ser estendido. Adicionando-se um vocabulário e regras de inferência para incluir novos conceitos sobre recursos e aplicações.

No trabalho de (SHARMA A BAWA, 2007) é feito um comparativo de vários critérios, confrontando mecanismos de descoberta baseados em sintaxe com mecanismos de descoberta baseados em semântica, conforme tabela 4.1.

Tabela 4.1: Comparativo de Descoberta de Recursos baseado em Sintaxe e Semântica

Critério	Baseada em Sintaxe	Utilizando Ontologias
Descrição do Recurso e Requisição	Simétrico	Assimétrico
Manutenção e Compartilhamento da Ontologia	Difícil de manter e compartilhar	Fácil de manter e compartilhar
Preferências de Matching	Muito limitado, baseado em sintaxe	Facilmente adicionado nos recursos e requisições
Checagem da Integridade	Sem checagem de integridade	Feito na base do conhecimento do domínio
Expressividade	Menos Expressivo	Muito expressivo, a requisição pode ser modelada especificamente para aplicações específicas do domínio
Flexibilidade e Extensibilidade	Menos flexível e extensível	Novos conceitos e restrições podem ser facilmente adicionadas dentro da Ontologia a qualquer momento

4.3 Considerações sobre o Capítulo

Este capítulo apresentou as características básicas dos mecanismos de descoberta de recursos, assim como características arquiteturais que devem possuir para obter o máximo de desempenho em ambientes heterogêneos e dispersos. O próximo capítulo apresenta alguns mecanismos de descoberta de recursos classificados pela utilização ou não de semântica na descrição e consulta por recursos.

5 TRABALHOS EM DESCOBERTA DE RECURSOS

Nesta seção estão relacionados alguns mecanismos de descoberta de recursos e a forma que estes mecanismos utilizam para armazenar e pesquisar pelos recursos descobertos. Ao final é apresentado o PerDis, Pervasive Discovery Service, um mecanismo de descoberta de recursos proposto ao middleware EXEHDA.

5.1 Baseados em Sintaxe

Esta seção apresenta mecanismos de descoberta de recursos que realizam o *matching*, ou seja, a consulta por recursos utilizando a comparação de exata de atributos. Alguns utilizam operadores lógicos para expandir um pouco os resultados desejados. Estes mecanismos utilizam uma sintaxe bem definida, mas não utilizam nenhum tipo de semântica na descrição e consulta dos recursos.

5.1.1 INS/TWINE

O INS/Twine (BALAZINSKA; BALAKRISHNAN; KARGER, 2002) é um serviço para descoberta de recursos de forma escalável, onde entidades chamadas *resolvers* colaboram como *peers* para distribuir as informações relacionadas aos recursos e responder consultas, conforme figura 5.1.

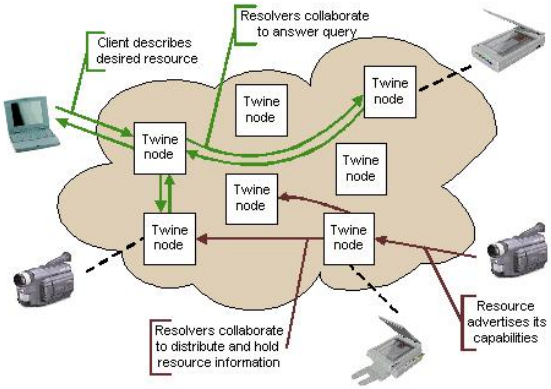

Figura 5.1: Arquitetura INS/Twine (BALAZINSKA; BALAKRISHNAN; KARGER, 2002)

O INS/Twine não possui expressividade para especificação de consultas. Seu esquema de representação de consultas é definido por um documento XML, mas ele realiza apenas combinações exatas de atributos e valores.

5.1.2 UPnP

O *Universal Plug-aNd-Play* (UPNP FORUM, 2008) possui uma arquitetura para conexão de dispositivos através de uma rede *peer-to-peer*. Através do UPnP o dispositivo pode juntar-se a uma rede de forma dinâmica, obter um endereço IP, anunciar e descobrir recursos. Utiliza o protocolo SSDP, Simple Service Discovery Protocol, para descoberta de recursos. O dispositivo envia uma mensagem de anúncio (ssdp:alive) multicast dos seus serviços para os possíveis clientes do serviço. O SSDP envia mensagem de busca (ssdp:discover) multicast quando um novo dispositivo é adicionado a rede, todos os dispositivos que escutarem a mensagem devem respondê-la através de unicast para quem enviou a consulta. O UPnP utiliza XML para descrever as características do dispositivo. A mensagem de anúncio contém a URL de um XML que permite a descrição precisa das características do dispositivo UPnP.

5.1.3 Allia

O Allia é um framework para descoberta de recursos em ambientes ad-hoc desenvolvido por pesquisadores da University of Maryland Baltimore Country (RATSIMOR et al., 2002). Ele é baseado em agentes e governado por políticas, utilizando caching em uma rede peer-to-peer. Não é especificado nenhum mecanismo para descrições durante o processo de descoberta de recursos. Recursos podem seguir a especificação Jini e serem descritos e acessados através de proxies Jini.

5.1.4 Jini

O Jini é uma arquitetura para descoberta de recursos baseada em Java (JINI, 2009), e desenvolvida pela Sun Microsystems. Utiliza mensagens multicast para localizar Lookup Service e registra os recursos no Lookup Service através do procedimento Join. Quando o cliente não encontra o Lookup Service ele faz a requisição diretamente ao recurso (Peer-lookup), figuras 5.2, 5.3, 5.4 e 5.5.

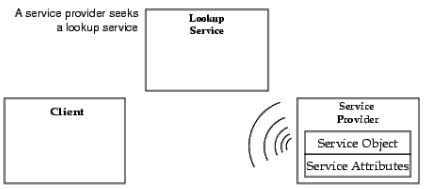

Figura 5.2: Jini - Descoberta (JINI, 2009)

Dispositivos e aplicações se registram com a rede Jini utilizando um processo chamado Descoberta e Adesão (*Discovery and Join*). Para se unir a uma rede Jini, um dispositivo ou aplicação se registra na Tabela de Consulta (*Lookup Table*) de um Serviço de Consulta (*Lookup Service - LS*), que é a base de dados para todos os serviços na rede (similar ao DA no SLP). Além de ponteiros para os serviços, a Tabela de Consulta de Jini pode armazenar trechos de código Java associados a esses serviços. Isso significa que

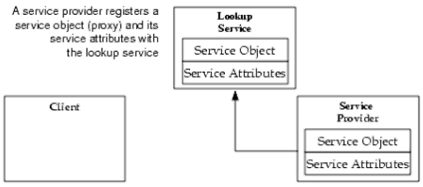

Figura 5.3: Jini - Join (JINI, 2009)

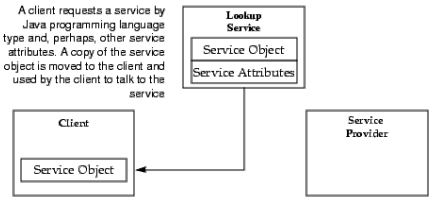

Figura 5.4: Jini - Lookup (JINI, 2009)

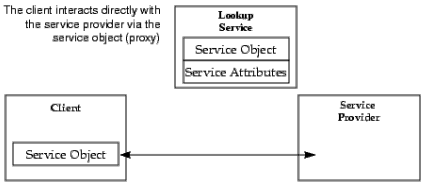

Figura 5.5: Jini - Cliente (JINI, 2009)

serviços podem armazenar *drivers* para os dispositivos, uma interface, e outros programas que ajudem o usuário a ter acesso aos serviços. Quando um cliente quer utilizar o serviço, o código objeto é carregado da Tabela de Consulta para a máquina virtual do cliente. Enquanto uma consulta de serviço em SLP devolve uma URL de Serviço, o código objeto Jini oferece acesso direto ao serviço usando uma interface conhecida pelo cliente. Essa mobilidade de código substitui a necessidade de pré-instalar *drivers* para o cliente.

A requisição e a seleção devem ser implementadas como objetos Java, seguindo as interfaces Jini específicas. Quando o serviço é localizado, Jini fornece um *stub* Java RMI para acessá-lo. O protocolo Jini é baseado em licenças (*leases*), desta forma todos os anúncios e registros são considerados válidos apenas por um período de tempo específico e relativamente curto. Clientes e serviços que são executados por muito tempo devem renovar suas licenças periodicamente, assim entidades que falham são removidas automaticamente de todos os serviços de busca quando a licença expira. Além disso, a renovação periódica de licenças pelas entidades é capaz de reconstruir o estado global no caso do travamento de um LS. Jini não suporta operação sem diretório, e não pode interoperar com nenhum outro protocolo ou linguagem. Uma vez que o protocolo depende do uso de stubs Java, um dispositivo deve implementar uma máquina virtual Java (JVM) ou usar um proxy. Esses requisitos fazem com que Jini não seja ideal para a utilização em pequenos dispositivos.

A expressividade na representação de propriedades de recursos é limitada devido ser baseada no processo de filtragem nos valores de atributos de uma interface Java

5.1.5 SLP

O SLP (*Service Location Protocol*) é um protocolo desenvolvido dentro do Internet Engineering Task Force (IETF), e é uma especificação independente de linguagem. O protocolo define três tipos de agentes e as mensagens de requisição e resposta trocadas entre eles (GROUP, 1999). SLP é descentralizado, leve, escalável e extensível para ser utilizado dentro de uma organização. Ele permite, mas não obriga uma administração centralizada. Sua arquitetura é constituída de três componentes principais: Agentes do Usuário (UA), que executam a descoberta do serviço em nome do cliente; Agentes de Serviço (SA), que anunciam a localização e características dos serviços, em nome dos próprios serviços e Agentes de Diretório (DA), que coletam endereços de serviços e informações recebidas dos SAs em suas bases de dados e respondem às requisições de serviços dos UAs. Quando um novo serviço se conecta à rede, o SA se comunica com o DA para anunciar sua existência (registro de serviço). Quando um usuário precisa de um certo serviço, o UA consulta os serviços disponíveis na rede a partir do DA (solicitação de serviço). Depois de receber o endereço e as características do serviço desejado, o usuário pode finalmente utilizar o serviço. Antes que o cliente (UA ou SA) possa se comunicar com o DA, ele deve descobrir a existência do DA. Há três métodos diferentes para a descoberta do DA: estático, ativo e passivo. Na descoberta estática, os agentes do SLP obtêm o endereço do DA através de DHCP (*Dynamic Host Configuration Protocol*). Servidores DHCP distribuem os endereços dos DAs para as aplicações que desejarem. Na descoberta ativa, UAs e SAs enviam solicitações de serviço para o endereço de grupo *multicast* (239.255.255.253). Um DA ouvindo neste endereço vai receber essas solicitações em algum momento e respondê-las diretamente ao agente solicitante. No caso de descoberta passiva, DAs periodicamente enviam anúncios *multicast* para seus serviços. UAs e SAs ficam sabendo o endereço do DA a partir desses anúncios e podem se comunicar diretamente com ele.

A linguagem de consulta provida pelo SLP é mais poderosa que a usada por Jini e UPnP, porém é baseada em uma sintaxe própria (baseada em string e não em XML), o que dificulta uma possível interoperabilidade com outras especificações.

5.1.6 Globus Toolkit

O Globus Toolkit é um software livre utilizado para construir sistemas e aplicações em grade (TOOLKIT, 2009). Ele está sendo desenvolvido pela Globus Alliance. O Globus possui um serviço chamado MDS (Monitoring & Discovery System) responsável por descobrir os recursos disponíveis e seu status. Adota o padrão LDAP para representação e procura (query) de recursos. Sua estrutura é descentralizada e permite escalabilidade. O MDS possui uma estrutura hierárquica que consiste de três componentes principais, conforme figura 5.6:

- *Information Providers (IPs):* coletam os dados do recurso e se comunicam com o GRIS (sensores);
- *Grid Resource Information Service (GRIS):* executa em um recurso e atua como um gateway de informações para esse recurso;
- *Grid Index Information Service (GIIS):* provê um diretório agregado (indexador) de dados que facilita a descoberta e monitoração.

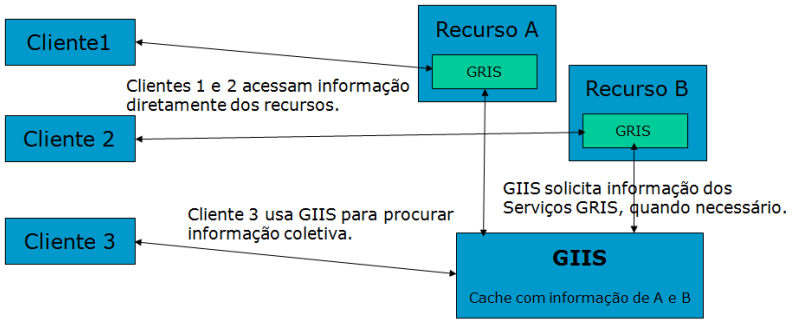

Figura 5.6: Globus Toolkit (TOOLKIT, 2009)

5.1.7 PerDis

O PerDis, Pervasive Discovery Service (SCHAEFFER, 2005), é um serviço proposto para o EXEHDA responsável pela descoberta de recursos que interage com os serviços de reconhecimento e adaptação de contexto do middleware. Esta interação incorpora ao mecanismo de descoberta o tratamento de informações relacionadas ao contexto dos usuários e recursos. A arquitetura do PerDis, figura 5.7, atende aos requisitos necessários a uma estratégia para descoberta de recursos na computação pervasiva. Estes requisitos abordam os seguintes aspectos:

- utilização de informações do contexto de execução;
- utilização de estratégias para manutenção automática de consistência;
- expressividade (parcial) na descrição de recursos e critérios de pesquisa;
- possibilidade de interoperabilidade com outras estratégias de descoberta;
- suporte à descoberta de recursos em larga-escala;
- utilização de preferências por usuário.

Os componentes que compõe a arquitetura do PerDis são os seguintes:

- *Resource Component (RC):* responsável por informar as características relacionadas a um determinado recurso, que permitam a um usuário optar ou não por sua utilização. É composto pelos módulos *Resource descriptor*, responsável por manter as propriedades descritivas do recurso e *Lease renewer* que realiza o gerenciamento da renovação do *lease*, utilizado para indicar periodicamente a disponibilidade do recurso;
- *User Component (UC):* habilita o usuário a realizar o processo de descoberta, interagindo com os demais componentes da arquitetura. Seus módulos são: *Query builder*, interface para montagem de consultas a recursos e *Result Browser*, permite ao usuário inspecionar os recursos retornados em uma consulta;

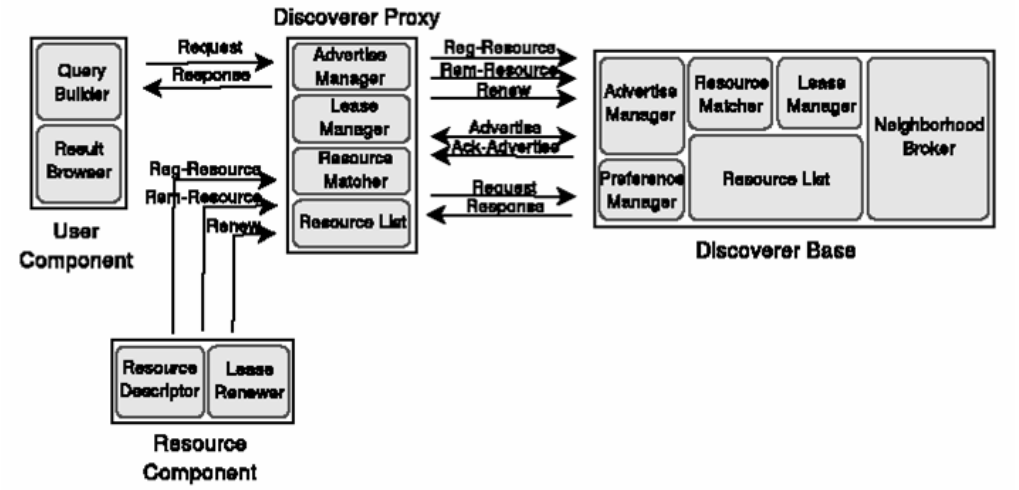

Figura 5.7: Arquitetura PerDis (SCHAEFFER, 2005)

- *Discoverer Base (DB) e Discoverer Proxy (DP):* Atuam como catálogo de recursos, abrangendo os recursos disponibilizados em uma célula de execução. Os principais módulos comuns a essas duas entidades são:
  - *Advertise manager:* realiza anúncio *multicast* da instância do serviço de descoberta no momento em que este for disparado, a fim de que a localização e a comunicação entre as instâncias desses componentes possam ser feitas de forma automática;
  - *Resource List:* mantém a lista de recursos gerenciados pela instância do serviço de descoberta;
  - *Lease Manager:* responsável pelo gerenciamento do controle de *lease*, realizando o descarte de recursos que não tenham sido renovados há um determinado período de tempo;
  - *Resource Matcher:* realiza a comparação de critérios de pesquisa.

O *discoverer base* ainda possui alguns módulos adicionais não presentes em *discoverer Proxy*:

- *Preference manager:* aplica preferências de usuários;
- *o Neighborhood broker:* usado na comunicação peer-to-peer entre células que compõe o Ambiente Pervasivo.

5.1.8 Condor

O Projeto Condor (THAIN; TANNENBAUM; LIVNY, 2005) nasceu na Universidade de Wisconsin, com objetivo de criar um ambiente capaz de prover grande poder computacional a médio e longo prazo. Para isso, utiliza-se de recursos computacionais ociosos de equipamentos não dedicados, apesar de ser possível utilizá-lo em *clusters* dedicados.

O Condor introduziu o conceito de *High-Thoughput Computing*, ou seja, Computação de Alta Vazão.

Em muitos projetos de pesquisa ou de engenharia é comum a necessidade de execução de programas que necessitem de semanas ou meses de execução para serem concluídos. Para este tipo de aplicação que o Condor foi desenvolvido.

O funcionamento do Condor baseia-se na execução de jobs. O usuário submete um Job ao sistema. Condor, por sua vez, encontra os recursos disponíveis para a execução. O sistema Condor monitora continuamente a execução dos jobs. Quando uma máquina executando um job torna-se indisponível (por exemplo, pelo retorno do usuário), Condor pode realizar um Checkpoint do job e migrá-lo para uma outra máquina que esteja disponível. Dessa forma, a execução na outra máquina poderá prosseguir do ponto em que havia parado.

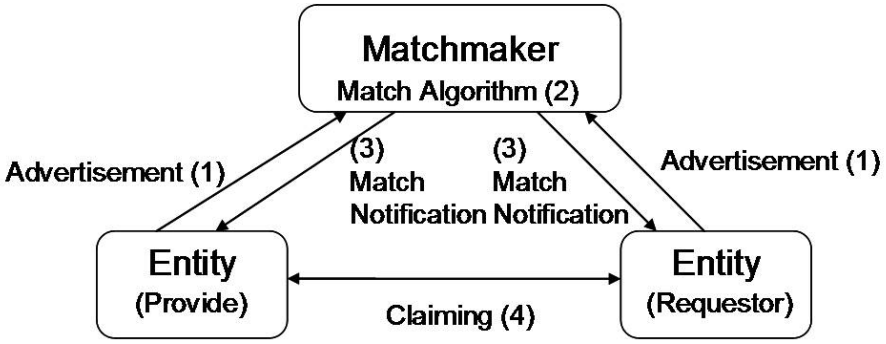

Figura 5.8: Arquitetura Condor (THAIN; TANNENBAUM; LIVNY, 2005)

O usuário Condor tem a ilusão de que as tarefas remotas são executadas localmente. De forma a tornar possível o acesso à maior quantidade possível de máquinas, Condor não assume que as máquinas para a execução remota montam os mesmos sistemas de arquivos da máquina do usuário. Para solucionar esta questão, Condor realiza o redirecionamento das chamadas de sistema da máquina que está executando o job para a máquina base do usuário. Condor provê uma biblioteca de redirecionamento para esta função.

Organização Básica:

- Sistema organizado em aglomerados
- Um aglomerado, tipicamente, compreende as máquinas em uma rede local
- O aglomerado é composto por três tipos de nós (fornecedor de recursos, consumidor de recursos e gerenciador do aglomerado)

As informações são disseminadas através da linguagem ClassAd e as entradas são do tipo atributo e valor.

5.1.9 Salutation

Salutation é formado por um consórcio aberto de indústrias, chamado Consórcio Salutation, é responsável pelo desenvolvimento de sua arquitetura, que define um modelo abstrato com três componentes: Cliente, Servidor e Salutation Manager (SLM). O Salutation Manager gerencia toda a comunicação, e faz a ponte através de diferentes meios

de comunicação. Serviços registram suas capacidades com um SLM, e clientes consultam o SLM quando necessitam de um serviço. Depois de descobrir um serviço desejado, clientes podem solicitar a utilização do serviço através do SLM. Salutation define seu protocolo baseado no SunRPC. O modelo pode ser implementado com ou sem um serviço de diretório (SLM) separado, e os diretórios podem ser organizados como uma hierarquia ou algum outro grafo de diretórios cooperativos. Quando implementado sem um diretório, clientes e serviços podem localizar uns aos outros diretamente, através de *broadcast* local. Essa configuração sem diretório permite ao Salutation trabalhar corretamente em uma rede sem administração, como por exemplo uma rede residencial ou automotiva. Salutation é uma protocolo bem estabelecido, com algumas implementações comerciais para Windows (95/98 and NT) e previsões de implementações adicionais para Palm OS e Windows CE planejadas para um futuro próximo. Exceto pelo SunRPC, Salutation é composto por tecnologias neutras. Aléem disso, Salutation foi projetado para ser compatível com tecnologias sem fio, e já existem adaptações de Salutation para IrDA e Bluetooth.

5.2 Baseados em Semântica

Esta seção apresenta mecanismos de descoberta de recursos que realizam o *matching*, ou seja, a consulta por recursos utilizando mecanismos semânticos. Através da semântica é possível obter resultados mais expressivos, pois o mecanismo pode obter entendimento de conceitos que não estavam declarados explicitamente.

5.2.1 OMM (Ontology-Based MatchMaker)

Ontology-Based MatchMaker (TANGMUNARUNKIT; DECKER; KESSELMAN, 2003) é um serviço de *matchmaking* em ambiente de grade. Este projeto utiliza tecnologias de Web Semântica para solucionar o problema de *matching* de recursos no ambiente distribuído. Foi desenvolvido um protótipo de um seletor de recursos baseados em ontologias que utiliza, além de ontologias, bases de conhecimento e regras para solucionar o *matching* de recursos no ambiente distribuído.

No OMM são criadas ontologias separadas para a descrição de recursos e suas requisições, o que permite ao cliente do recurso não conhecer como o recurso foi descrito exatamente, sendo realizado o *match* semântico através dos termos definidos nestas ontologias.

Principais características do OMM:

- descrição assimétrica de recursos e requisições: a descrição do recurso e a requisição são modelados separadamente, mas é utilizado o *match* semântico para a associação entre eles;
- compartilhamento e manutenção de ontologias;
- restrições bilaterais através de utilização de políticas de uso;
- habilidade em especificar preferências de *matching*;
- verificação de integridade através de regras de restrições;
- expressividade na descrição das requisições;

- flexibilidade e extensibilidade.

A arquitetura do OMM é formada por três módulos:

- ***Matchmaking Module***: consultas são realizadas através da linguagem TRIPLE e o sistema de inferência escrito em java e C, baseado em Jini. Este módulo realiza *matchmaking* e retorna o resultado como um objeto java contendo a lista de recursos que satisfazem a consulta;

- ***Request Handling***: um serviço persistente é utilizado para suportar vários clientes simultaneamente. Clientes submetem requisições em RDF, e recebem uma lista de recursos que satisfazem a consulta. Usuários podem expressar suas preferências para recursos selecionados utilizando uma função de *ranking*. Clientes podem indicar o número de recursos retornados que irão ser ordenados baseado em seus valores de *ranking*.

- ***Resource Discovery***: é o módulo chave do OMM, responsável por coletar dinamicamente informação dos recursos de várias fontes, transformando a informação para a ontologia, e atualiza a base de conhecimento.

5.2.2 A Grid Service Discovery Matchmaker based on Ontology Description

No artigo de (LUDWIG; SANTEN, 2002) é proposto um *framework* para o *matching* na descoberta de serviços para um ambiente de grade, baseado em ontologias.

Na justificativa apresentada no trabalho para utilização do *matching* semântico é o fato de que a maioria dos mecanismos de descoberta realizam o processo de descoberta através de comparação baseada em *string* e inteiros.

A ontologia é utilizada para que o provedor dos serviços e o usuário deste serviço possam compartilhar um entendimento comum das capacidades do serviço. Esta ontologia foi desenvolvida com a linguagem DAML-S.

O *framework* desenvolvido é baseado no *matchmaker* LARKS. O mecanismo de *matching* é composto de três estágios de filtragem, que são: contexto, sintático e semântico, onde a ontologia fornece a base do conhecimento. A diferença em realação ao LARKS é o filtro semântico.

Três componentes são necessários para o SD *matching*, o provedor de recursos, o solicitador de serviço e o serviço *matchmaker*. O provedor de serviços registra a descrição de seus serviços no serviço *matchmaker*. O solicitador de serviços requisita um serviço e envia uma requisição ao serviço *matchmaker*. O serviço *matchmaker* retorna os resultados ao solicitador de serviço. E o solicitador de serviço decide qual serviço utilizar de acordo com a necessidade do cliente.

O *matchmaker* processa a requisição seguindo os seguintes passos:

- comparando a requisição com todos os anúncios registrados no banco de dados;

- decidindo qual provedor de serviço tem o melhor *matching* com relação à requisição, dependendo do algoritmo e da ontologia definida;

- provendo as informações ao requisitante.

São utilizados filtros em três estágios:

- *Matching* de Contexto: seleciona os anúncios no banco de dados auxiliar tal que possam ser comparados à requisição em contexto igual ou similar;
- *Matching* Sintático: compara a requisição com os anúncios selecionados no estágio anterior em três etapas que são: a comparação de perfis, *matching* de similaridade e *matching* de assinatura. O serviço de ontologia provê o modelo de serviço e o fundamento do serviço para a comparação de perfis. Os outros estágios focam na entrada e saída de restrições e declarações na especificação;
- *Matching* Semântico: checa se as restrições de entrada e saída dos pares de requisição e os anúncios se emparelham logicamente.

5.2.3 Ontologias Aplicadas à Descrição de Recursos em Ambientes Grid

No trabalho de (PERNAS; DANTAS, 2004), foi desenvolvida uma ontologia com o objetivo de facilitar a busca e seleção de recursos. Desta forma, é possível descrever os recursos disponíveis em um ambiente de grade de forma sintática e semântica através de um vocabulário comum ao domínio. No desenvolvimento desta ontologia, foi definido primeiramente o domínio da aplicação, em seguida este domínio foi transformado em um vocabulário e, finalmente, foram criados os axiomas, que guardam todas as regras a serem respeitadas durante a inclusão de novo vocabulário na ontologia. A ontologia foi desenvolvida utilizando a linguagem OWL. Também foi desenvolvido um serviço de grade de modo a externalizar a ontologia aos usuários do ambiente de grade.

Na arquitetura proposta, figura 5.9, a ontologia localiza-se praticamente em uma camada a parte do ambiente de grade, e as requisições vindas dos usuários devem primeiramente realizar uma consulta à ontologia, que por sua vez utiliza os metadados e as visões semânticas para obter maior informação a respeito dos recursos.

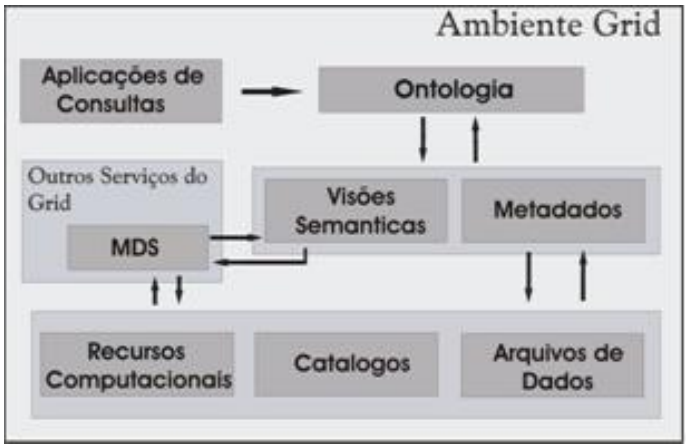

Figura 5.9: Arquitetura de Pernas e Dantas (PERNAS; DANTAS, 2004)

A ontologia proposta foi documentada com os seguintes componentes:

- Dicionário de Dados: agrega todas as classes e instâncias de classe da ontologia, juntamente com suas significações;

- Árvore de Classificação de Conceitos: agrupa todas as classes e subclasses da ontologia;
- Tabelas de Atributos de Classes e de Instâncias: apresentam, respectivamente, para cada instância e para cada classe da ontologia, todos os seus atributos;
- Tabelas de Instâncias: apresentam a descrição, atributos e valores de cada instância da ontologia;
- Árvores de Classificação de Atributos: mostram os atributos inferidos através da existência de outros atributos de hierarquia superior.

5.2.4 Serviço Baseado em Semântica para Descoberta de Recursos em Grade Computacional

A arquitetura proposta por (ALLEMAND, 2006) é composta por uma camada de Conhecimento com a finalidade de prover um serviço de descoberta de recursos computacionais de forma semântica, que está localizada acima das camadas Física e a de *Middleware* da grade. A camada de Conhecimento é composta por dois elementos básicos: o Repositório Semântico e o Raciocinador. O Repositório Semântico permite a descoberta semântica dos recursos da grade por meio do uso de um *template* ontológico. Para a edição deste *template* e tratamento dos diversos tipos de recursos computacionais no ambiente de grade foi utilizado o Protégé-OWL. O Raciocinador, segundo componente da camada de Conhecimento, interage com o Repositório Semântico para selecionar os recursos de forma apropriada. Neste trabalho foi utilizado o *middleware* de grade Globus Toolkit 4 (GT4) e seu serviço *Monitoring and Discovery System* (MDS4), juntamente com o Ganglia, para coletar automaticamente as informações sobre os recursos da grade computacional. Como motor de inferência empregou-se a ferramenta Pellet-OWL. Por meio da realização de um estudo de caso com uso do protótipo.

A linguagem de consulta de inferência utilizada é a RDQL (SEABORNE, 2004).

5.2.5 DReggie

Este trabalho, desenvolvido por (CHAKRABORTY et al., 2001), é um aperfeiçoamento do serviço de descoberta Jini. Ele aperfeiçoa a técnica de descrição de serviços e utiliza raciocinadores no mecanismo de *matching*. Em DReggie os serviços são descritos utilizando a DARPA Agent Markup Language (DAML). A ontologia em DAML foi criada para descrever serviços com suas propriedades e atributos. Os serviços de anúncio e requisição utilizam a mesma ontologia. O serviço de *lookup* do Jini é aprimorado com um raciocinador baseado em Prolog que é capaz de *matching* complexos baseados em descrições DAML.

5.3 Considerações sobre o Capítulo

Este capítulo apresentou alguns mecanismos de descoberta de recursos, com foco na forma que estes mecanismos realizam a descrição e consulta de recursos.

O próximo capítulo apresenta uma proposta inicial das características de um mecanismo de descoberta de recursos para o *middleware* EXEHDA com suporte semântico.

6 UMA CONTRIBUIÇÃO A DESCOBERTA DE RECURSOS COM MECANISMOS SEMÂNTICOS

Este capítulo tem por objetivo apontar as características que devem ser observadas quando da concepção de um mecanismo de descoberta de recursos que utilize tecnologias de Web Semântica para realizar *matching* semântico no ambiente ubíquo proposto pelo *middleware* EXEHDA (YAMIN, 2004). Deste modo, o capítulo introduz uma visão inicial da arquitetura proposta, bem como do Processador Semântico que estão sendo projetados.

Os recursos no ambiente ubíquo são heterogêneos e estão espalhados geograficamente, onde o usuário muitas vezes não tem uma visão real da localização desses recursos. Cabe ao *middleware* que promove o ambiente ubíquo compartilhar e gerenciar os recursos disponíveis.

O mecanismo de descoberta de recursos é fundamental para a descoberta e registro dos recursos do ambiente. Os recursos estão conectados por diversas tecnologias de rede e de vários tipos, como LANs, WANs e MANETs. A organização dos recursos pelo mecanismo de descoberta pode utilizar técnicas centralizadas, não centralizadas ou híbridas.

A escalabilidade do mecanismo de descoberta deve ser prevista, pois o ambiente ubíquo pode ter uma dimensão muito grande, envolvendo instituições, cidades, estados e até países. Por isso o modelo a ser proposto levará em consideração a escalabilidade, permitindo a interoperabilidade do mecanismo de descoberta entre as células do ambiente ISAMPe, promovido pelo EXEHDA.

O modelo que está sendo proposto visa qualificar o modelo PerDis proposto em (SCHAEFFER, 2005) para o *middleware* EXEHDA. O grande diferencial desta proposta para o modelo existente é na expressividade na descrição dos recursos. A descrição dos recursos será feita utilizando tecnologias de Web Semântica, promovendo uma melhor organização sintática e semântica na representação e consulta. Também será possível a criação de regras para restringir consultas incompatíveis, como exemplo, uma solicitação de um cliente por um determinado sistema operacional com sistema de arquivos incompatível com o sistema operacional solicitado.

O uso de ontologias na representação dos recursos tornará a consulta por recursos mais precisa, pois não haverá ambiguidade de conceitos no domínio da ontologia.

Em ambiente ubíquo envolvendo várias organizações, os recursos são gerenciados por mais de uma pessoa. Este procedimento pode acarretar em problemas, pois um mesmo

recurso pode ser nomeado de várias maneiras.

A utilização de *matching* sintático realiza apenas a comparação de atributos idênticos, descartando recursos que apesar de serem idênticos ao solicitado foram descritos de forma diferente ao passado pelo usuário/cliente.

No *matching* semântico a consulta é feita verificando instâncias disponíveis da classe do recurso solicitado, caso não haja nenhuma instância exata disponível no momento da classe do recurso solicitado, há a possibilidade de apresentar como resposta à consulta recursos semelhantes existentes em uma sub-classe ou numa superclasse.

Para concepção do mecanismo de descoberta de recursos para o *middleware* EXEHDA (YAMIN, 2004) com suporte semântico entende-se que se faz necessário atender as seguintes características:

- a descoberta de recursos deverá contemplar o ambiente celular promovido pelo ISAMPe;
- para atender a escalabilidade sem sobrecarregar a rede pretende-se utilizar redes *super peer*, já previstas no modelo de descoberta PerDis (SCHAEFFER, 2005);
- o catálogo de recursos de cada célula estará localizado no EXEHDABase de cada célula;
- a ontologia e a máquina de inferência também estarão localizadas no EXEHDABase de cada célula;
- haverá um gerenciador de recursos para cadastrar todos os recursos de cada célula;
- os recursos possuirão um atributo indicando seu *status*, indicando se está disponível ou não;
- todo recurso possuirá um *lease*, indicando seu tempo de vida;
- os recursos deverão enviar mensagens ao diretório de cada célula para renovar seu *lease*, caso contrário seu *status* será alterado para “indisponível”;
- a consulta por recursos será realizada primeiramente na célula em que se encontra o cliente/requisitante. Será feita uma consulta na ontologia local por recursos idênticos ou semelhantes;
- após a consulta por recursos na célula local será realizada a procura por recursos em células vizinhas;
- o recurso encontrado pela consulta poderá ser acessado diretamente pelo cliente.

A figura 6.1 apresenta a área de atuação do mecanismo de descoberta de recursos proposto. A figura demonstra um cliente realizando uma consulta por uma determinada impressora ao diretório (EXEHDA Base). O diretório encontrou uma instância de impressora na própria célula e retorna o resultado ao cliente. Com os dados do recurso solicitado o cliente acessa diretamente a impressora solicitada. A comunicação entre as células é realizada através dos EXEHDA Base de cada célula, formando uma topologia *super peer*. Em cada célula todos recursos são cadastrados no EXEHDA Base local.

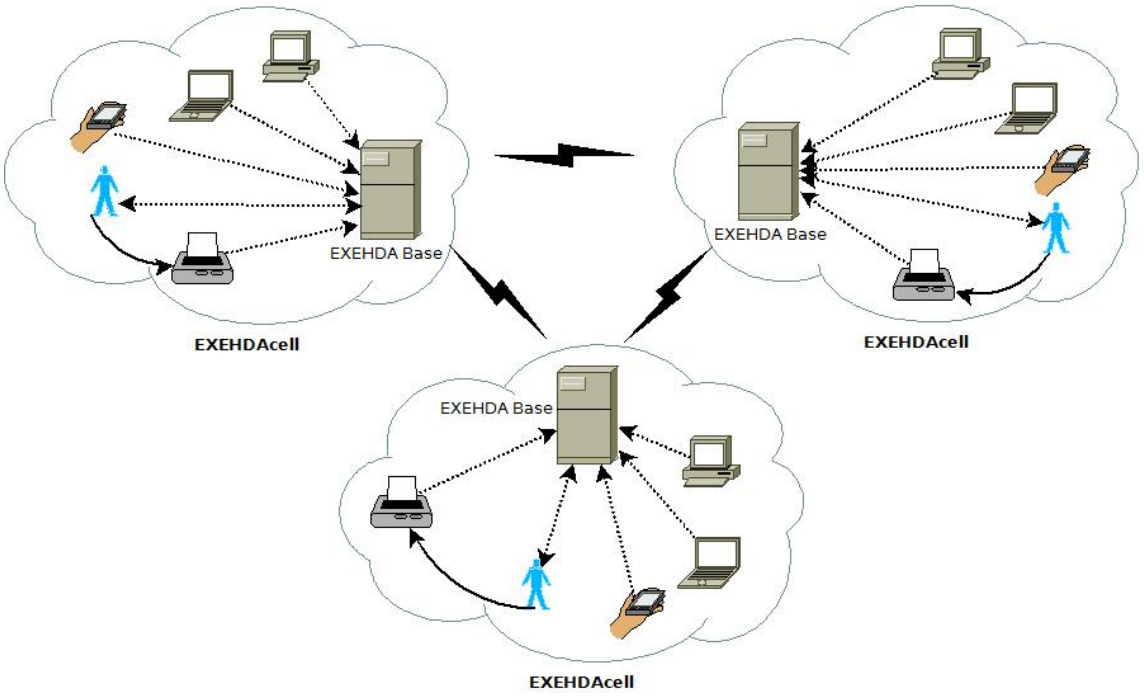

Figura 6.1: Ambiente Celular de Descoberta

A arquitetura proposta na figura 6.2 está organizada em três componentes distintos: diretório, provedor de recursos e cliente.

O diretório é formado por quatro componentes: Gerente de Recursos, Analisador de Consultas, Processador Semântico e Broker. O Gerente de Recursos é responsável pelo cadastro e manutenção do *status* do recurso. O Analisador de Consultas verifica a consistência e autorização e passa a consulta para o Processador Semântico para consultar a ontologia. O Broker é responsável pela comunicação entre células, provendo uma descoberta de recursos escalar.

O Processador Semântico é implementado em Java com API Jena e é formado por três componentes: Modelo Ontológico, Raciocinador e Persistência de Dados.

O Modelo Ontológico é responsável pela manutenção da ontologia do mecanismo de descoberta. Esta ontologia é desenvolvida na linguagem OWL.

O raciocinador é responsável por aplicar regras e processar as consultas realizadas em SPARQL. O raciocinador irá procurar por recursos idênticos e semelhantes ao solicitado.

O módulo de Persistência de Dados é responsável pelo armazenamento das triplas da ontologia no banco de dados MySQL.

O cliente possui um componente chamado Gerador de Consultas responsável pela formulação das consultas e apresentação de resultados.

O provedor de recursos possui o componente Gerente de atributos que mantém os atributos do recurso atualizados e é também responsável pela renovação do *lease*. O *lease* é o tempo de vida do recurso.

A figura 6.3 apresenta uma ontologia desenvolvida no Protégé, pelo grupo G3PD da Universidade Católica de Pelotas, para ser utilizada pelos mecanismos de consciência e adaptação de contexto e descoberta de recursos. Esta ontologia está em estágio inicial e será ampliada e detalhada no decorrer do trabalho.

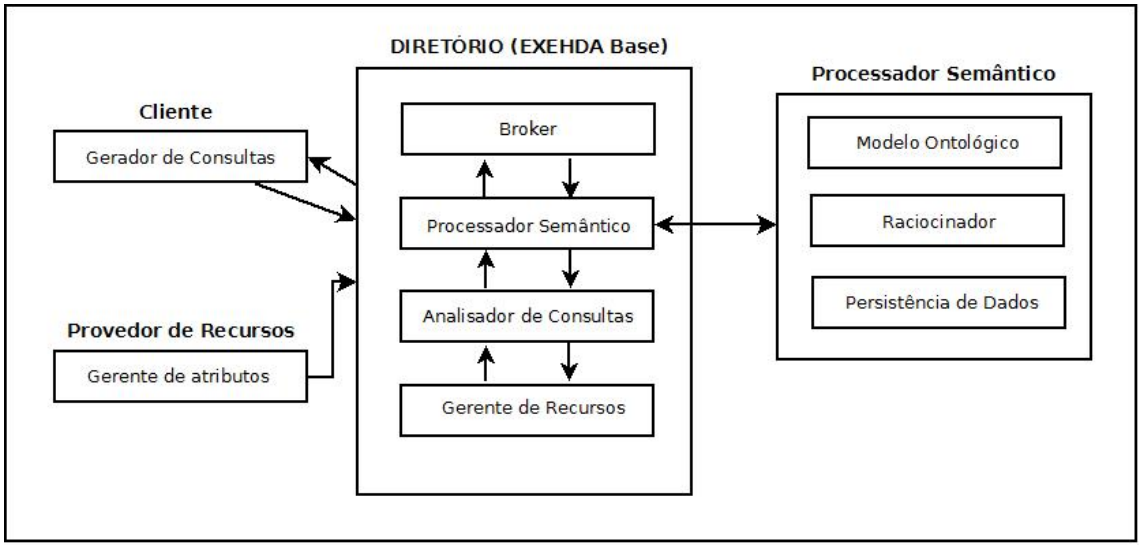

Figura 6.2: Modelo Proposto: Arquitetura Inicial

Na ontologia serão descritos todos os recursos do ambiente. Como ponto de partida, foram definidas as seguintes superclasses:

- Hardware: qualquer dispositivo que poderá ser utilizado no ambiente;
- Software: qualquer software que pode ser utilizado no ambiente;
- Locais: todos locais onde pode ser usado o modelo;
- Célula: contém informações sobre as células do EXEHDA, inclusive nodos e EXEHDA Base;
- Usuários: todos usuários que podem usar o ambiente.

Para a Classe Nodo foram criados os seguintes relacionamentos:

- SistOper: Um Nodo possui uma instância da Classe SO;
- Proces: Um nodo possui uma instância da Classe Processador.

Através de experimentos com a API Jena foi possível verificar o aumento de resultados numa consulta por recursos na ontologia proposta. Na ontologia foram instanciados alguns recursos através da ferramenta Protégé. Os recursos foram instanciados nas classes:

- Intel: Celeron, Pentium e Athlon;
- Unix: Solaris;
- Linux: Debian, Ubuntu e Kurumin;
- Windows: Windows_Vista_Premium, Windows_98 e Windows_XP;
- Nodo: Nodo_1, Nodo_9 e Nodo_6.

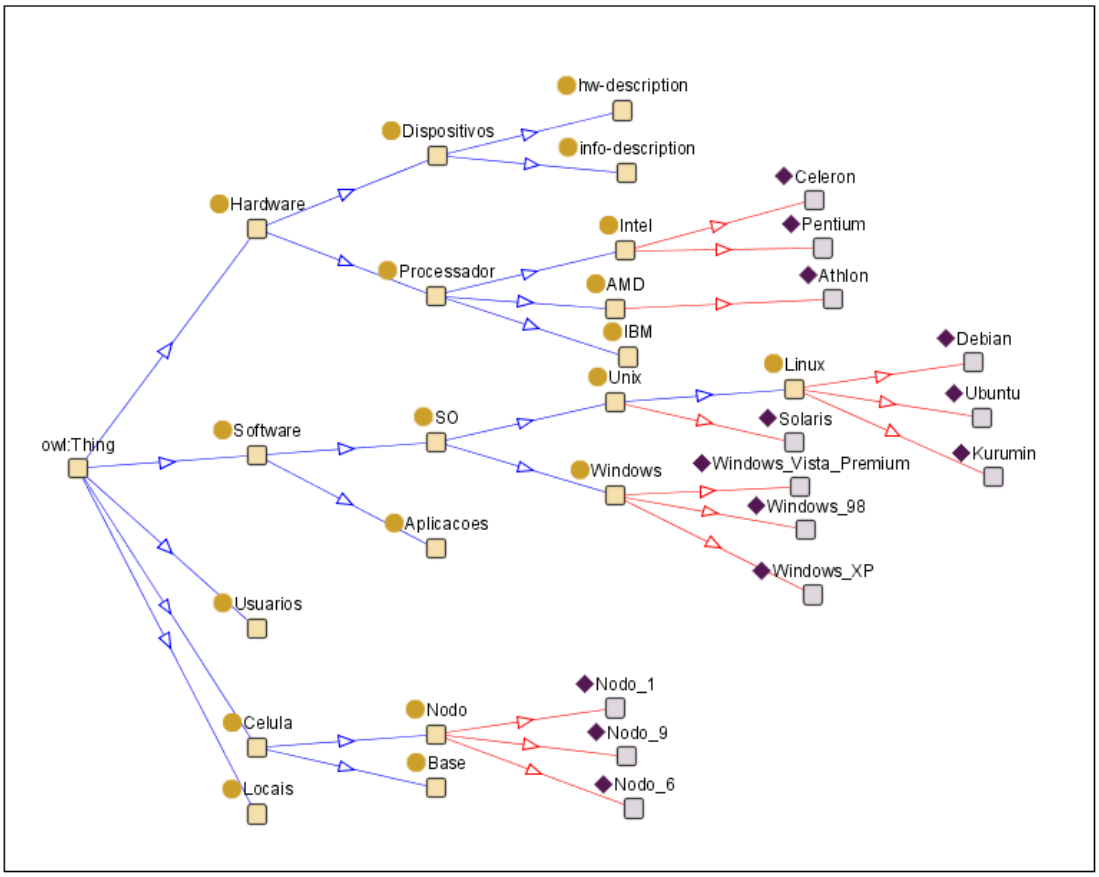

Figura 6.3: Ontologia do Mecanismo de Descoberta de Recursos

No primeiro experimento foi realizada a consulta por instâncias de sistemas operacionais UNIX sem o uso de raciocinadores. Esta consulta foi realizada através da linguagem SPARQL.

Listagem 6.1: Consulta SPARQL

```
PREFIX rdf: <http://www.w3.org/199/02/22-rdf-syntax-ns#>
PREFIX ont: <http://localhost/Ontologia.owl#>
SELECT *
        WHERE
        {?recurso rdf:type ont:Unix}
```

Esta consulta foi realizada sem auxílio do raciocinador da API Jena e apresentou a seguinte resposta:

Instâncias de Unix sem Raciocinador:

==========

<http://localhost/Ontologia.owl#Solaris>

A consulta retornou apenas o Sistema Operacional Solaris, como sistema operacional Unix disponível.

Observando a ontologia representada na figura 6.3 é possível verificar que existem instâncias de Linux disponíveis. O Linux é uma sub-classe de Unix. Como o raciocinador não estava ativo, o resultado não apresentou os conceitos similares ao solicitado.

No segundo teste foi ativado o raciocinador e realizada a mesma consulta, apresentando o seguinte resultado:

Instâncias de Unix com Raciocinador:

==========

<http://localhost/Ontologia.owl#Solaris>
<http://localhost/Ontologia.owl#Kurumin>
<http://localhost/Ontologia.owl#Debian>
<http://localhost/Ontologia.owl#Ubuntu>

O raciocinador apresentou também os sistemas operacionais compatíveis com Unix.

A segunda consulta realizada realiza uma procura por nodos com Sistema Operacional Unix e Processadores Intel. Sendo que Unix e Intel são classes que contém sub-classes.

Listagem 6.2: Consulta SPARQL 2

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX ont: <http://localhost/Ontologia.owl#>
SELECT *
WHERE
        {?nodo ont:SistOper ?SO .
         ?SO rdf:type ont:Unix .
         ?nodo ont:Proces ?Processador .
         ?Processador rdf:type ont:Intel }
```

O resultado obtido com a segunda consulta foi o seguinte:

<http://localhost/Ontologia.owl#Nodo_1> | <http://localhost/Ontologia.owl#Debian> | <http://localhost/Ontologia.owl#Pentium>

O Nodo_1 possui Sistema Operacional Debian e Processador Pentium, conforme pode ser verificado na ontologia representada pela figura 6.3.

Através da API Jena também foi possível validar a ontologia, verificando inconsistências nos dados. O código em java utilizado foi o seguinte:

Listagem 6.3: Validação da Ontologia

```
public static void Validar(String Ontology) {
        OntModel model = ModelFactory.createOntologyModel(OntModelSpec.
            OWL_MEM);
        model.read(Ontology);
        Reasoner reasoner = ReasonerRegistry.getOWLReasoner();
        reasoner = reasoner.bindSchema(model);
        InfModel owlInfModel = ModelFactory.createInfModel(reasoner,
            model);
        ValidityReport vrp1 = owlInfModel.validate();
        if (vrp1.isValid()) {
            System.out.println("Ontologia OK");
        } else {
            System.out.println("Ontologia com problemas...");
            for (Iterator i = vrp1.getReports(); i.hasNext(); ) {
            System.out.println(" - " + i.next());
        }

        }
```

```
        }
```

Para teste da validação foi alterado o valor da “Veloc_Proc” que está definido como “float” de 3.0 para a string “tres”.

Listagem 6.4: Arquivo OWL - nodo

```
<Nodo rdf:ID="Nodo_1">
    <URL rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
    >atlas.ucpel.tche.br</URL>
    <Num_Proc rdf:datatype="http://www.w3.org/2001/XMLSchema#int"
    >1</Num_Proc>
    <Nodo_Ativo rdf:datatype="http://www.w3.org/2001/XMLSchema#boolean"
    >true</Nodo_Ativo>
    <Veloc_Proc rdf:datatype="http://www.w3.org/2001/XMLSchema#float"
    >3.0</Veloc_Proc>
    <SistemaArquivos rdf:datatype="http://www.w3.org/2001/XMLSchema#
       string"
    >ext3</SistemaArquivos>
    <Memoria rdf:datatype="http://www.w3.org/2001/XMLSchema#int"
    >4000</Memoria>
    <SistOper rdf:resource="#Debian"/>
    <Proces rdf:resource="#Pentium"/>
</Nodo>
```

O resultado apresentado foi a inconsistência da ontologia gerada: Ontologia com problemas... - Error ("range check"): "Incorrectly typed literal due to range (prop, value)"
Culprit = http://localhost/Ontologia.owl#Nodo_1
Implicated node: http://localhost/Ontologia.owl#Veloc_Proc
Implicated node: ’tres’ http://www.w3.org/2001/XMLSchema#float

Através do suporte a persistência de dados da API Jena foi possível armazenar a ontologia no banco de dados MySQL, conforme o código java abaixo:

Listagem 6.5: Persistência de Dados

```
public static void Importar_Ont(String Ontology) {
        String url = "jdbc:mysql://localhost/jena";
        String driver = "com.mysql.jdbc.Driver";
        DBConnection conn = new DBConnection ( url, "root", "mysql", "
           mysql" );
        ModelRDB model;
        if( !conn.containsModel("DR")){
           System.out.println("Ontologia está sendo armazenada...");
           model = ModelRDB.createModel(conn,"DR");
        }
        else {
            System.out.println("Ontologia já existe no BD, abrindo...")
               ;
            model = ModelRDB.open(conn,"DR");
        }
        model.begin();
        model.read(Ontology);
        model.commit();
  }
```

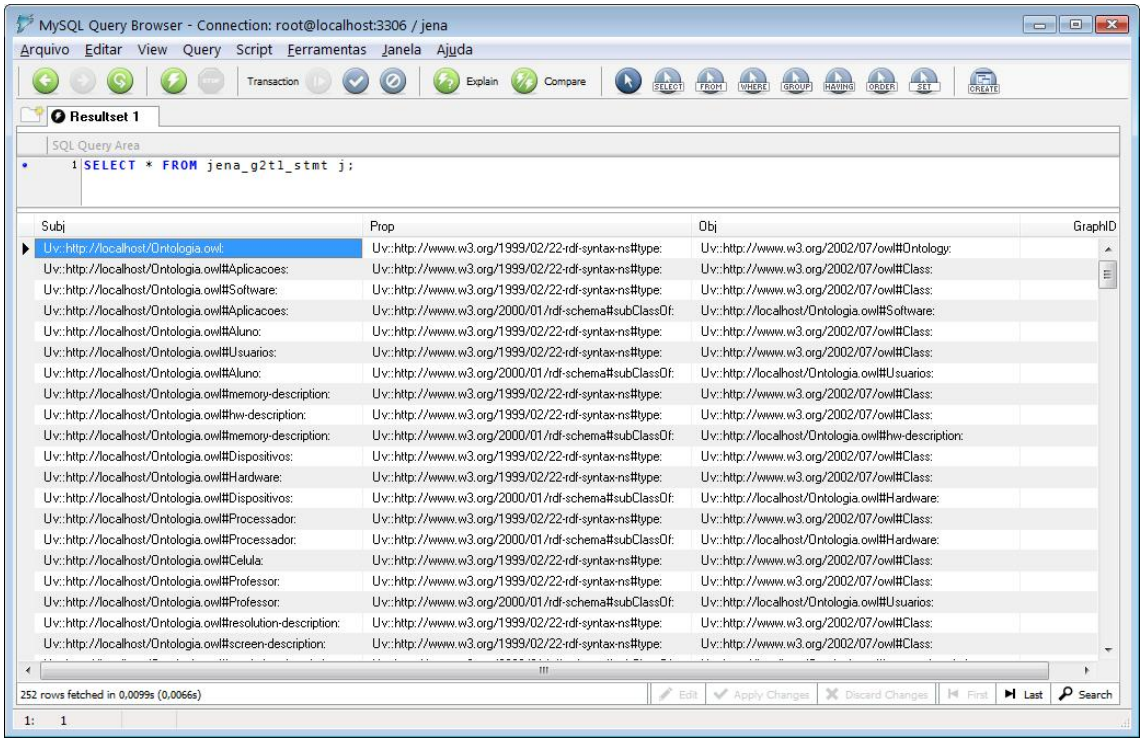

Figura 6.4: Ontologia armazenada no banco de dados MySQL

A figura 6.4 mostra as triplas que formam a ontologia, armazenadas no banco de dados MySQL.

Após armazenar as triplas da ontologia no banco de dados é possível abrí-lo e realizar consultas SPARQL diretamente no banco, conforme código Java abaixo:

Listagem 6.6: Consultas com Banco de Dados

```
public static void Consulta_BD(String Ontology) {
        String url = "jdbc:mysql://localhost/jena";
        String driver = "com.mysql.jdbc.Driver";
        DBConnection conn = new DBConnection ( url, "root", "mysql", "
            mysql" );
        ModelRDB model = ModelRDB.open(conn,"DR");
        String ocQuery = "PREFIX rdf: <http://www.w3.org/1999/02/22-rdf
            -syntax-ns#>" +
                         "PREFIX ont: <http://localhost/Ontologia.owl#>
                             " +
                         "SELECT * " +
                         "WHERE " +
                         "{?recurso rdf:type ont:Unix} ";
        QueryExecution ocQe = QueryExecutionFactory.create(ocQuery,
            model);
        System.out.println("Instâncias de Unix:");
        ResultSet ocResults = ocQe.execSelect();
        ResultSetFormatter.out(ocResults);
        ocQe.close();
    }
```

A utilização de regras permite aumentar o entendimento de termos que estão subentendidos, mas os computadores não conseguem entender. Para exemplificar foi criado

um arquivo com três regras definindo os termos *dualcore* (nodos com dois processadores), *quadcore* (nodos com quatro processadores) e *manycore* (nodos com mais de quatro processadores), da seguinte forma:

Listagem 6.7: Jena - Definição de Regras (Ontologia.rules)

```
@prefix ont: <http://localhost/Ontologia.owl#>.
@include <RDFS>.

[dualcore: (?n ?p ont:Nodo) (?n ont:Num_Proc 2) -> (?n rdf:type ont:
    dualcore)]
[quadcore: (?n ?p ont:Nodo) (?n ont:Num_Proc 4) -> (?n rdf:type ont:
    quadcore)]
[manycore: (?n ?p ont:Nodo) (?n ont:Num_Proc ?x) greaterThan(?x,4) ->
    (?n rdf:type ont:manycore)]
```

Para o processamento das regras foi criado um código em Java que adiciona os conceitos deduzidos pelas regras na ontologia e realiza a consulta através de SPARQL pelo novo conceito deduzido.

Listagem 6.8: Jena - Código Java com Regras

```
public static void Aplicar_Regras(String Ontology) {
        OntModel model = ModelFactory.createOntologyModel(OntModelSpec.
            OWL_MEM);
        model.read(Ontology);
        Reasoner reasoner = new GenericRuleReasoner(Rule.rulesFromURL("
            file:///d:/jena/Ontologia.rules"));
        InfModel infModel = ModelFactory.createInfModel(reasoner, model
            );
        Model model2 = infModel.getDeductionsModel();
        model.add(model2);
        reasoner = reasoner.bindSchema(model);
        String ocQuery =  "PREFIX rdf: <http://www.w3.org/1999/02/22-
            rdf-syntax-ns#>" +
                          "PREFIX ont: <http://localhost/Ontologia.owl
                              #>" +
                          "SELECT * " +
                          "WHERE " +
                             "{?nodo rdf:type ont:dualcore }" ;

        QueryExecution ocQe = QueryExecutionFactory.create(ocQuery,
            model);
        ResultSet ocResults = ocQe.execSelect();
        ResultSetFormatter.out(ocResults);
        ocQe.close();
    }
```

O resultado da consulta por nodos *dualcore* é apresentado abaixo:

Listagem 6.9: Jena - Resultado da Consulta

```
run:
-----------------------------------------
| nodo                                  |
=========================================
| <http://localhost/Ontologia.owl#Nodo_6> |
| <http://localhost/Ontologia.owl#Nodo_2> |
-----------------------------------------
```

CONSTRUÍDO COM SUCESSO (tempo total: 3 segundos)

A ontologia proposta beneficiará o modelo de descoberta de recursos por:

- possuir um domínio bem definido com classes e subclasses que serão utilizadas em todo o domínio do EXEHDA, evitando ambiguidade de conceitos sobre um mesmo recurso;
- possibilitar consultas semânticas;
- inferir novos conceitos sobre os recursos existentes;
- possibilitar a criação de regras.

Para o processamento da ontologia e integração com o mecanismo de descoberta de recursos a ser modelado, possivelmente será utilizado a API Jena. A Jena, como já foi descrita anteriormente, possui suporte a raciocinadores e persistência de dados com suporte a banco de dados MySQL.

7 CONSIDERAÇÕES FINAIS

Este trabalho apresentou uma revisão bibliográfica dos conceitos envolvidos em Computação Ubíqua, enfatizando a importância dos mecanismos de descoberta de recursos nesses ambientes.

Foram relacionados alguns projetos que realizam a abstração do ambiente ubíquo, procurando atingir a transparência e onipresença contidas na visão de Mark Weiser. Dentre os projetos relacionados, o projeto ISAM com o *middleware* EXEHDA foi estudado com mais atenção. O modelo do mecanismo de descoberta de recursos que será proposto será direcionado ao EXEHDA, qualificando trabalhos anteriores de descoberta de recursos que realizam o *matching* sintático, apenas.

Para delinear as características essenciais de um mecanismo de descoberta semântico foi realizado um estudo das características básicas, levando em conta a arquitetura, escopo da descoberta, descrição dos recursos, gerenciamento da informação, requisição e anúncio de recursos, armazenamento das informações, métodos de seleção e invocação e suporte a mobilidade. Foi observado a importância da escalabilidade dos mecanismos de descoberta, devido o contexto ubíquo.

A escolha da linguagem de descrição e consulta de recursos pode tornar o mecanismo de descoberta com maior ou menor expressividade. Muitas linguagens realizam o *matching* através de pares de atributos, comparando igualdades entre valores, limitando o número de respostas da consulta solicitada. Um mecanismo com maior expressividade utiliza linguagens de descrição e consulta com suporte semântico, descrevendo relacionamentos entre conceitos. Isto torna o entendimento de conceitos que estão implícitos inteligíveis pelas máquinas.

As tecnologias de Web Semântica também foram estudadas a fim de compreender a criação e a utilização de ontologias. As ontologias podem ser criadas através de diversas linguagens, entre elas, a OWL é a mais completa. A linguagem OWL contém os relacionamentos entre sujeito, predicado e objeto. Estes relacionamentos tornam o entendimento de conceitos bem mais precisos e organizados dentro do domínio da ontologia.

O processamento da ontologia pode ser realizado por várias ferramentas, este trabalho analisou a API Jena. Através da API Jena é possível processar as ontologias, consultar classes, subclasses, superclasses e instâncias. Também é possível armazenar a ontologia em banco de dados através da persistência de dados, disponível na API.

A possibilidade de utilizar raciocinadores expande os benefícios das ontologias. Com os raciocinadores é possível inferir conhecimento da ontologia, seja através de regras restritivas, ou através de novos conhecimentos inferidos. A consulta de instâncias da ontologia pode ser feita através da API Jena com a utilização da linguagem SPARQL.

Com a revisão das tecnologias de Web Semântica notou-se a importância do uso de ontologias e suas tecnologias para a realização do *matching* semântico desejado no mecanismo de descoberta.

Neste trabalho foram estudados vários mecanismos de descoberta de recursos. Muitos deles não atendem a heterogeneidade, dinamicidade e escalabilidade necessárias do ambiente ubíquo. A maioria dos mecanismos estudados realizam o *matching* através de pares de atributos, alguns poucos possibilitam a utilização de operadores lógicos, mas mesmo assim possuem uma baixa expressividade na representação e consulta de recursos.

Alguns trabalhos de descoberta de recursos com a utilização de semântica foram apontados aqui, mas a proposta destes mecanismos está direcionada à computação em grade. Alguns implementam o modelo ontológico em DAML, ou realizam consultas através da linguagem TRIPLE ou RQL.

O mecanismo de descoberta de recursos proposto neste trabalho está em desenvolvimento e irá utilizar ontologias escritas em OWL, realizar consultas em SPARQL, processar ontologias com API Jena, inferir conhecimento com raciocinadores da API Jena e utilizará persistência de dados em MySQL. As tecnologias citadas estão em estudo e poderão ser substituídas por outras equivalentes durante o desenvolvimento do trabalho.

O mecanismo proposto utilizará uma arquitetura de três componentes como já foi descrito e realizará a descoberta de recursos entre células do ambiente ISAMPe disponibilizado pelo *middleware* EXEHDA.

Acredita-se que o encaminhamento que foi tomado em direção à modelagem de um mecanismo de descoberta de recursos semântico permitirá qualificar o mecanismo de descoberta do *middleware* EXEHDA com respostas mais completas.

REFERÊNCIAS

ALLEMAND, J. N. C. **Serviço Baseado em Semântica para Descoberta de Recursos em Grade Computacional**. 2006. 120p. Dissertação de Mestrado — Departamento de Ciência da Computação, UnB, Brasília, DF.

BALAZINSKA, M.; BALAKRISHNAN, H.; KARGER, D. INS/Twine: A Scalable Peer-to-Peer Architecture for Intentional Resource Discovery. **Lecture Notes in Computer Science**, [S.l.], v.2414, p.195–??, 2002.

BECHHOFER, S.; HARMELEN, F. van; HENDLER, J.; HORROCKS, I.; MCGUINNESS, D. L.; PATEL-SCHNEIDER, P. F.; STEIN, L. A. **OWL Web Ontology Language Reference**. Disponível em: <http://www.w3.org/TR/owl-ref/>. Acesso em junho de 2007.

BERNERS-LEE, T.; HENDLER, J.; LASSILA, O. The Semantic Web. **Scientific American**, [S.l.], v.5, n.284, p.34–43, May 2001.

BIOMEDICAL INFORMATICS RESEARCH, S. C. for. **Protégé**. Disponível em: <http://http://protege.stanford.edu/>. Acesso em maio de 2009.

CHAKRABORTY, D.; PERICH, F.; AVANCHA, S.; JOSHI, A. **DReggie**: Semantic Service Discovery for M-Commerce Applications.

DECKER, K.; WILLIAMSON, M.; SYCARA, K. **Matchmaking and Brokering**.

DICKINSON, I. **Jena API - A Semantic Web Framework for Java**. Disponível em: <http://jena.sourceforge.net/ontology/>. Acesso em maio de 2009.

FENSEL, D.; HORROCKS, I.; VAN HARMELEN, F. OIL in a Nutshell. In: EUROPEAN WORKSHOP ON KNOWLEDGE ACQUISITION, MODELING AND MANAGEMENT, 12., 2000. **Anais...** ., 2000.

GONZALEZ-CASTILLO, J.; TRASTOUR, D.; BARTOLINI, C. **Description Logics for Matchmaking of Services**. [S.l.]: Hewlett Packard Laboratories, 2001. (HPL-2001-265).

GROUP, N. W. **SLPv2**. Disponível em: http://www.ietf.org/rfc/rfc2608.txt. Acesso em maio de 2009.

GRUBER, T. A Translation Approach to Portable Ontology Specifications. **Knowledge Acquisition**, [S.l.], p.199–220, 1993.

GUARINO, N. Formal ontology and information systems. In: GUARINO, N. (Ed.). **Formal Ontology in Information Systems**. Amsterdam: IOS Press, 1998. p.3–18.

ISAM. **Infra-estrutura de Suporte às Aplicações Móveis**. Disponível em: <http://www.inf.ufrgs.br/ isam/index.html>. Acesso em junho de 2009.

JINI. **Jini**. Disponível em: http://www.jini.org. Acesso em maio de 2009.

KINDBERG, T.; FOX, A. System software for ubiquitous computing. **Pervasive Computing Magazine**, [S.l.], 2002.

KLYNE, G.; CARROLL, J. **Resource Description Framework (RDF)**: Concepts and Abstract Syntax. Disponível em: <http://www.w3.org/TR/rdf-concepts/>. Acesso em janeiro de 2008.

LEHMAN, T. J.; MCLAUGHRY, S. W.; WYCKOFF, P. T Spaces: The Next Wave. In: HICSS, 1999. **Anais...** [S.l.: s.n.], 1999.

LOPES, J. G. **Matching Semântico de Recursos Computacionais em Ambientes de Grade com Múltiplas Ontologias**. 2005. 116p. Dissertação de Mestrado — Departamento de Ciência da Computação, UnB, Brasília, DF.

LUDWIG, S. A.; SANTEN, P. V. **A Grid Service Discovery Matchmaker Based On Ontology**.

MARIN-PERIANU, R.; HARTEL, P. H.; SCHOLTEN, J. **A Classification of Service Discovery Protocols**. [S.l.]: Centre for Telematics and Information Technology, Univ. of Twente, The Netherlands, 2005. Technical report. (TR-CTIT-05-25).

MCGRATH, R. E. **Discovery and Its Discontents**: Discovery Protocols for Ubiquitous Computing. [S.l.: s.n.], 2000.

MICHAL JAKOB, N. K.; GHANEA-HERCOCK, R. Nexus – Middleware for Decentralized Service-Oriented Information Fusion. , [S.l.], 2007.

MIT Project Oxygen: project overview. Disponível em: <<http:oxygen.csail.mit.edu//>. Acesso em junho de 2009.

MIZOGUCHI, R.; VANWELKENHUYSEN, J.; IKEDA, M. Task ontology for reuse of problem solving knowledge. **IOS Press**, Amsterdam. [S. 1.], p.p. 46–59, 1995.

NABRZYSKI, J.; SCHOPF, J. M.; WEGLARZ, J. (Ed.). **Grid resource management**: state of the art and future trends. Norwell, MA, USA, and Dordrecht, The Netherlands: Kluwer Academic Publishers Group, 2003. xxi + 574p. n.ISOR 64. (International series in operations research & management science).

PERNAS, A. M.; DANTAS, M. A. R. Ontologias Aplicadas a Descrição de Recursos em Ambientes Grid. **INFOCOMP Journal of Computer Science**, [S.l.], v.3, n.2, p.26–31, 2004.

RATSIMOR, O.; CHAKRABORTY, D.; JOSHI, A.; FININ, T. Allia: alliance-based service discovery for ad-hoc environments. In: SECOND ACM INTERNATIONAL WORKSHOP ON MOBILE COMMERCE (WMC-02), 2002, New York. **Proceedings...** ACM Press, 2002. p.1–9.

REYNOLDS, D. **Jena 2 Inference support**. Disponível em: <http://jena.sourceforge.net/inference/index.html>. Acesso em maio de 2009.

ROMAN, M.; HESS, C.; CERQUEIRA, R.; RANGANATHAN, A.; CAMPBELL, R. H.; NAHRSTEDT, K. A Middleware Infrastructure for Active Spaces. **IEEE Pervasive Computing**, [S.l.], v.1, n.4, p.74–83, 2002.

SCHAEFFER, A. E. PerDis: um Serviço para Descoberta de Recursos no ISAM Pervasive Environment. **Universidade Federal do Rio Grande do Sul**, [S.l.], 2005.

SEABORNE, A. **RDQL - A Query Language for RDF**. Disponível em: <http://www.w3.org/Submission/2004/SUBM-RDQL-20040109/>. Acesso em junho de 2009.

SEABORNE, A. **SPARQL - A Query Language for RDF**. Disponível em: <http://http://www.w3.org/TR/rdf-sparql-query/>. Acesso em junho de 2007.

SHARMA A BAWA, S. Resource Discovery using Ontology Approach in Grid. **COIT-2007**, [S.l.], p.63–67, Março 2007.

SOLDATOS, J.; PANDIS, I.; STAMATIS, K.; POLYMENAKOS, L.; CROWLEY, J. L. Agent based middleware infrastructure for autonomous context-aware ubiquitous computing services. **Journal of Computer Communications**, [S.l.], 2006.

TANGMUNARUNKIT, H.; DECKER, S.; KESSELMAN, C. Ontology-Based Resource Matching in the Grid - The Grid Meets the Semantic Web. In: INTERNATIONAL SEMANTIC WEB CONFERENCE, 2003. **Anais...** Springer, 2003. p.706–721. (Lecture Notes in Computer Science, v.2870).

THAIN, D.; TANNENBAUM, T.; LIVNY, M. Distributed computing in practice: the Condor experience. **Concurrency and Computation: Practice and Experience**, [S.l.], v.17, n.2-4, p.323–356, Feb. 2005.

THOMPSON, M. S. **Service Discovery in Pervasive Computing Environments**. 2006. 135p. Tese (Doctor of Philosophy in Computer Engineering) — Faculty of the Virginia Polytechnic Institute and State University, Blacksburg, Virginia.

TOOLKIT, G. **Globus Toolkit**. Disponível em: http://www.globus.org. Acesso em maio de 2009.

UPNP FORUM, M. of the. **UPnP™ Device Architecture 1.1**. Disponível em: http://www.upnp.org/specs/arch/UPnP-arch-DeviceArchitecture.pdf. Acesso em maio de 2009.

VERZULLI. **Using thr Jena API to Process RDF**. Disponível em: <http://www.xml.com/pub/a/2001/05/23/jena.html>. Acesso em maio de 2009.

WEISER, M. The Computer for the 21st Century. **Scientific American**, [S.l.], v.3, n.265, p.94–104, Setembro 1991.

YAMIN, A. **Arquitetura para um Ambiente de Grade Computacional Direcionado às Aplicações Distribuídas, Móveis e Conscientes do Contexto da Computação Pervasiva**. 2004. 195p. Tese (Doutorado em Ciência da Computação) — Instituto de Informática, UFRGS, Porto Alegre, RS.

ZHU, F.; MUTKA, M.; NI, L. Service Discovery in Pervasive Computing Environments. **IEEE Pervasive Computing**, [S.l.], v.4, n.4, p.81–90, 2005.