

Semantisches Matchmaking und Semantische Ressourcenbeschreibung

Seminar
Semantic Grid
WS 2004/05

Inhalt

- Matchmaking (MM) und Ressourcenbeschreibung
- Traditionelles Matchmaking
- Semantisches Matchmaking
 - Ontologie-basierter Matchmaker
 - Verteilter P2P-Ansatz

Das Grid

- eine Plattform, die es dynamischen, multi-institutionellen Organisationen ermöglicht Ressourcen zu teilen und Probleme koordiniert zu lösen
- **Parteien:**
 - Provider
 - Kunden

Resource Matching

- Nicht jede Anwendung läuft auf jeder Ressource
- Ressource muss den Ansprüchen entsprechen
- das Finden einer Ressource anhand von einem Ressourcen-Gesuch
- Voraussetzung:
 - Beschreibung der Ressourcen

Traditionelles Matchmaking

- Condor Matchmaker
- Ressourcen Beschreibung
 - „Attribut-Wert“-Paare
- Mechanismus beruht auf
 - bilateralem, (beidseitiger Match)
 - symmetrischem, (gleiches Vokabular)
 - Attribut-basiertem (Syntax)

Matching

Nachteile der traditionellen Mechanismen

- Keine Semantik
 - Angebot: OpSys = “SOLARIS251”;
 - Gesuch: OpSys = “UNIX”;
- Im Vorfeld auf bestimmte Vokabeln einigen
- Attributlisten sind schwer erweiterbar

Semantic Matchmaking

- Syntax-Matching nicht ausreichend
- 2 semantische Mechanismen
 - Ontologie-basierter Matchmaker
 - Basiert auf Semantic Web Technologien
 - Verwendet Ontologien um Ressourcen zu beschreiben
 - Verteilter P2P-Ansatz
 - Das Wissen über die verfügbaren Ressourcen und das Hintergrundwissen ist auf die Teilnehmer des Netzwerks verteilt
- Ressour.-Beschreibung, Matchmaking

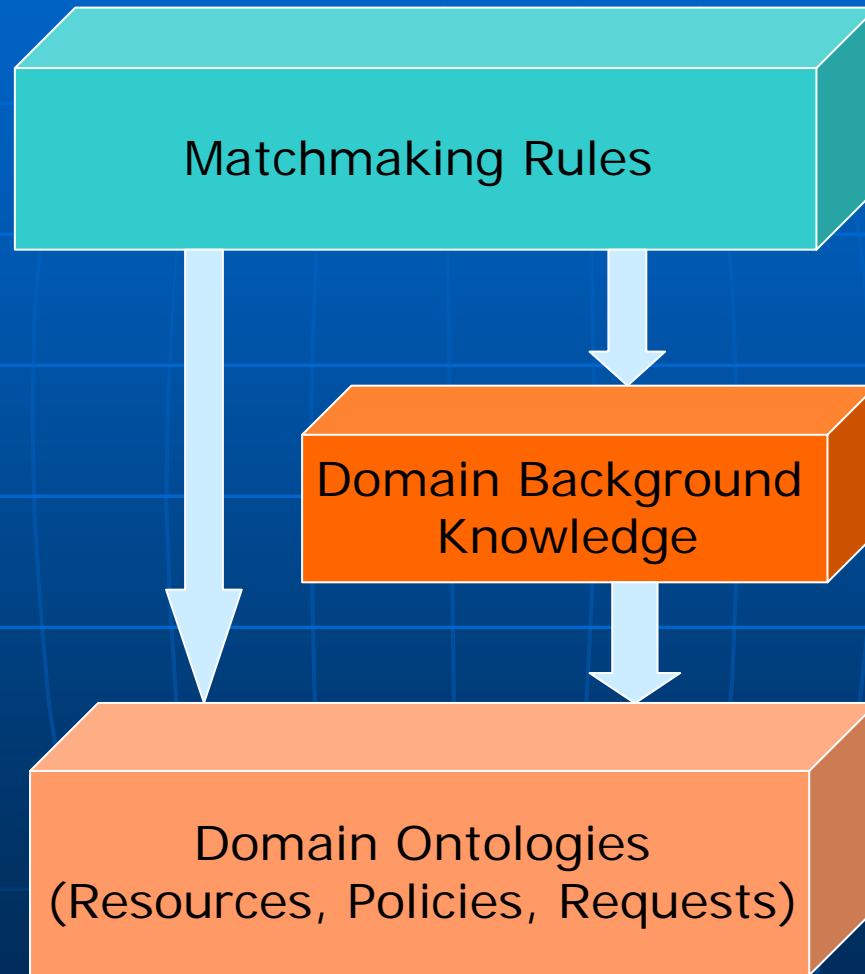
Ontologie-basierter Matchmaker

- Verwendet Ontologien um Ressourcen zu beschreiben
- Features
 - Asymmetrisches, (unabhängige Ontologien)
 - Bilaterales Matching
 - Matching preferences
 - Integrity check

Architektur des OMM

- Besteht aus 3 Komponenten
- Ontologien
 - Erfassen Domain Modell
 - Beinhalten Vokabeln
- Domain Background Knowledge
 - Bieten zusätzliches Wissen
- Matchmaking Rules
 - Beschreiben, wann ein „Match“ vorliegt

Architektur des OMM



Ontologien (1)

- Mechanismus basiert auf asymmetrischem Matching
- 3 unterschiedliche Ontologien
- Jede Ontologie definiert
 - Objekte
 - Eigenschaften dieser Objekte
 - Beziehungen zwischen diesen Objekten

Ontologien (2)

- Resource Ontology
 - Ressourcen, Fähigkeiten und Beziehungen
- Resource Request Ontology
 - Gesuch, Informationen über den Kunden
- Policy Ontology
 - Autorisation, Zugriffsbedingungen

Domain Background Knowledge

- Beinhaltet zusätzliches Wissen über die Domäne
- Dieses Wissen wird während des Matchmaking-Prozesses benötigt
- Wissen dargestellt in Form von Regeln
 - Regelsprache TRIPLE (basier auf Horn-Logik und F-Logik)

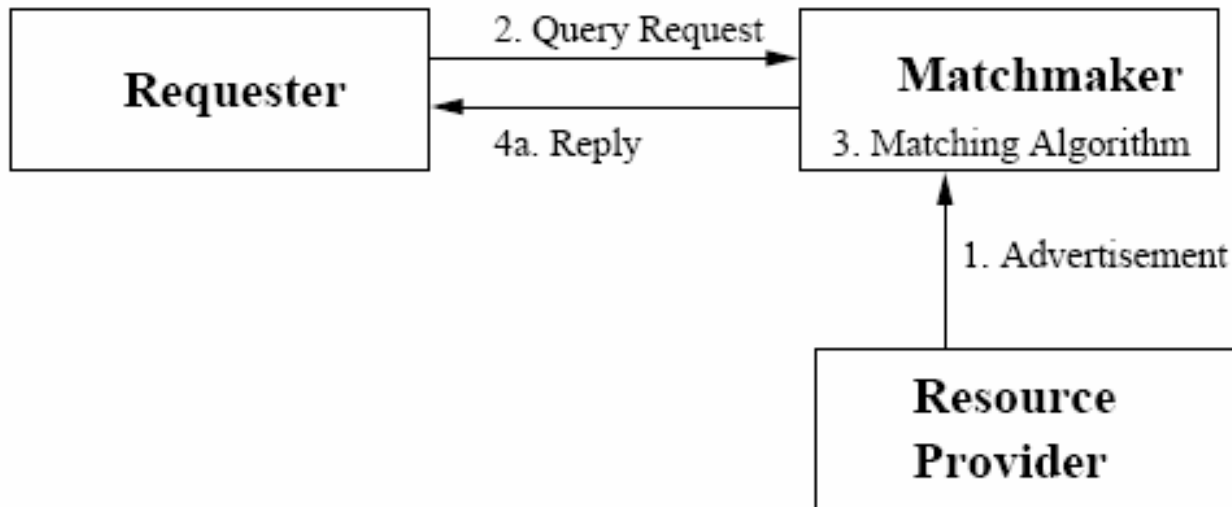
Beispiel DBK

```
@gridBackground { // specifies grid background knowledge
Linux[rdfs:subClassOf->GR:OperatingSystem].
Unix[rdfs:subClassOf->GR:OperatingSystem].
Debian[rdf:type->Linux].
Redhat[rdf:type->Linux].
SunOS[rdf:type->Unix].
Linux[rdf:type->Unix].
// transitivity axiom
FORALL X,Y,Z X[compatibleWith->Z] <- X[compatibleWith->Y] AND
Y[compatibleWith->Z].
// identity axiom
FORALL X X[compatibleWith->X].
//symmetry axiom
FORALL X,Y X[compatibleWith->Y]<- Y[compatibleWith->X].
FORALL X,Y,Z X[substitutes->Z] <- (Y[rdf:type->Z] and X[substitutes->Y])
or X[compatibleWith->Z].
}
```

Matchmaking Rules

- Definieren wann es zu einem Treffer kommt
- Dargestellt mit TRIPLE
- Matchmaking Rules nutzen die Ontologien und das Background Knowledge um den MM-Prozess durchzuführen

Der Matchmaking-Prozess



(a) Matchmaking Service

Schwächen des OMM

- Schwächen der traditionellen MM
- Erweiterbarkeit bei OMM?
 - Ontologie-Modellierung ständiger Prozess
 - Wissen über Ressourcen wächst
 - Ontologien leicht erweiterbar?
 - MM muss Zugriff auf komplettes Wissen haben

P2P-Ansatz

- Erweiterbarkeit bei bestehenden MM nicht gelöst
- Ansatz, bei dem Wissen bei Bedarf erworben werden kann
- Teilnehmer des P2P Netzwerks
 - Verwalten den Ressourcen-Katalog
 - Bieten Ressourcen an
 - Liefern Hintergrundwissen
 - Können Ressourcen suchen

Semantische Beschreibung

- Keine allgemeingültige Ontologie
 - Jeder Teilnehmer hat eigene, unvollständige Ontologie
- Description Logic (DL)
 - Wissen repräsentieren
 - Das Wissen ist in 2 Teile aufgeteilt
 - T-BOX (taxanomical)
 - A-BOX (assertional)

T-BOX (1)

- Enthält Background Knowledge
- Konzepte
 - Zum Aufbau des „Wissens“
 - Teile eines Individuums der Domäne
 - Mit ihnen kann man alle Aspekte einer Ressource beschreiben
 - Konzept: LINUX, Konzept: PENTIUM, ...

T-Box (2)

- Beziehung zwischen Konzepten
- Classification DAG
 - Knoten → Konzepte
 - Kanten → Beziehungen
- Jeder Knoten besitzt eigene Ontologie (erweiterbar)
- Virtual Classification DAG → „Gesamtwissen“

Node: 2.3.4.5

A-BOX

- Repräsentiert konkretes Wissen über Individuen der Domäne
- Ressourcen-Beschreibung
 - A-BOX-assertments (Instanzen bilden)
 - Eindeutige ID
 - Mehrere Eigenschaften → Instanzen von mehreren Konzepten

NODE: 1.1.2.1

Grundlagen

- Unvollständige Ontologien
- „Wissen“ wird über das Netzwerk verteilt
 - Nicht „geflutet“
- Jeder Knoten speichert Informationen über ein Konzept
 - Zugehöriges Superkonzept
 - Instanzen dieses Konzepts

| Concept | PENTIUM | Peer 3 |
|---------|----------------------------|--------|
| A-BOX | 1.1.2.5 [3] 1.1.2.8 [5] | |
| T-BOX | INTEL | |

Distributed Hash Tables (DHT)

- Konzepte werden mit Hilfe des DHT-Algorithmus über das Netzwerk verteilt
- Zu einem Schlüssel liefert der Algorithmus Adresse des Teilnehmers
- Als Schlüssel dienen die Konzepte

Beitritt eines Knotens

- Knoten tritt dem Netzwerk bei
 - Lokalen Graph veröffentlichen
 - Nachricht an Knoten
 - Die über neue Superkonzepte informiert werden müssen
 - Die für die Konzepte zuständig sind, für die dieser Knoten Instanzen hat (Ressourcen veröffentlichen)

Weitergabe der Information

- Neue Information übernehmen
- Betroffene Knoten informieren

- Neue Instanz → Nachricht an das Superkonzept

| | |
|-----------------|-------------|
| Concept PENTIUM | Peer 3 |
| A-BOX | 1.1.2.5 [3] |
| T-BOX | INTEL |

- Neues

Superkonzept

des eigenen Konzepts → Instanzen an dieses SK

- Antwort: Liste von Superkonzepten des eigenen Konzepts

Ausstieg eines Knotens

- Kein plötzliches Verlassen des Netzwerks
- Betroffene Knoten informieren
 - A-BOX Wissen wird gelöscht
Zieht sein Angebot zurück
 - T-BOX Wissen bleibt erhalten

Matchmaking

- Suche nach Ressource mit OS „UNIX“
 - Bestimmung des Knotens der verantwortlich für dieses Konzept ist (mit DHT-Algorithmus)
 - Antwort: Liste der Instanzen
 - Kontaktieren dieser Teilnehmer
- Ressourcen mit mehreren Eigenschaften
 - Komplexe Formel aufstellen; Konzepte mit UND, ODER und NICHT verknüpfen
 - Instanzenlisten dieser Konzepte erfragen
 - Formel auswerten

Schwächen

- „Nur die ersten Schritte...“
- Abfragen besser gestalten
- Knoten müssen sich abmelden
- T-BOX-Wissen bleibt immer erhalten
- Kein Ranking der Treffer
- Zusätzliche Beziehungen zwischen Konzepten; Verträglichkeiten