GridVine: Building Internet-Scale Semantic Overlay Networks

Karl Aberer, Philippe Cudré-Mauroux, Manfred Hauswirth,
and Tim Van Pelt
School of Computer and Communication Sciences
Swiss Federal Institute of Technology(EPFL), Switzerland
Department of Computer and Information Science
Linköpings Universitet, Sweden

Presented by Lan Tian
Advised by Matthias Bender

Proseminar Peer-to-Peer Information Systems
Max-Planck-Institute for Informatik

November.21, 2004

Contents

1 Introduction

In a peer-to-peer (P2P) system, nodes typically connect to a small set of random nodes (their neighbors), and queries are propagated along these connections. Such query flooding tends to be very expensive. So the node connections are proposed to be influenced with content. Thus the semantically related nodes form a Semantic Overlay Network, is a flexible network organization that improves query performance while maintaining a high degree of node autonomy. Recently research on semantic overlay networks in P2P architectures has received a lot of attention on the emerging area of peer-to-peer data management. The base of these approaches is a generalization of the concept of federated databases. The database is the integration of databases that have been independently developed. In such Database the peers store data or metadata according to their local schemas and freely can define mapping from their schemas to those of other peers. Then a network is constructed where peers are semantically connected through schema mappings and where query can be propagated to other peers with different schemas. The main thing is that query should be consistently answered and the reconciliation of different mappings in the semantic overlay network.

A distributed, scalable indexing structure is built up to route search requests, for example, Chord and P-Grid. Such decentralized P2P systems can be employed in order to respond to simple keyword-based queries. They could also support the efficient operations of a semantic overlay network. But the research on how to take advantage of this potential is in its infancy.

In order to enable semantic interoperability we introduce an architecture and implementation of a large-scale semantic overlay network with the help of structured overlay networks. At first we will explain what semantic overlay networks mean. The original P2P systems offer the potential for low cost sharing of information, autonomy and privacy. However, query processing in current P2P systems is very inefficient and does not scale well. The inefficiency arises because most P2P systems create a random overlay network. There queries are blindly forwarded from node to node. Therefore semantic overlay networks are proposed to improve query performance and nevertheless maintain a high degree of node autonomy. With Semantic Overlay Networks, nodes with semantically similar content are clustered together. To illustrate, consider Figure 1 which shows eight nodes, A to H are connected by solid lines. The nodes are connected to the other that have semantically similar content. Node A,B,C all have Rock Songs, so they establish connections among them. Similarly, nodes C,E, and F have Rap songs, so they cluster close to each other. In a semantic overlay network System, queries are processed by identifying which semantic overlay networks are better suited to answer it. Then the query is sent to a node in those semantic overlay networks and the query is forwarded only to the other members of that semantic overlay network. In this way, a query for Rock songs will go directly to the nodes that have Rock content. Almost as important, nodes outside the rock Semantic Overlay Network are not bothered with that query.

Another definition we should understand is Schema mapping. Many modern applications such as data warehousing, global information systems and electronic commerce need to take existing data with a particular structure or schema, and re-use it in a different form. These applications start with an

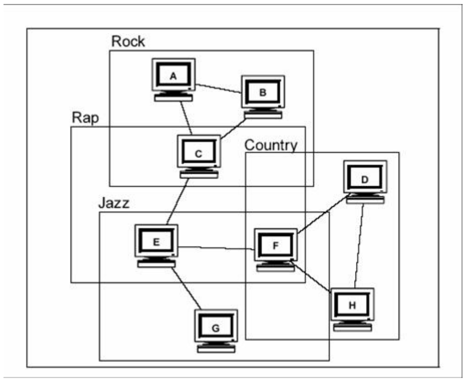
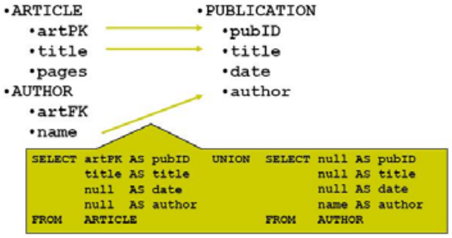


Figure 1: Semantic Overlay Network

understanding of how data will be used and viewed. That is, they start by determining a target schema. They then must create mappings between this target and the schemas of the underlying data sources. We could construct a schema mapping from a set of value correspondences. See figure 2. That means also that we could also retrieve the data items in different schemas.



```
SELECT artPK AS pubID   UNION  SELECT null AS pubID
       title AS title                 null AS title
       null  AS date                  null AS date
       null  AS author                name AS author
FROM   ARTICLE                 FROM   AUTHOR
```

Figure 2: Schema Mapping

The approach that we introduce is based on the concept of federated databases. The database is the integration of databases that have been independently developed. In such Database the peers store data or metadata according to their local schemas and freely can define mapping from their schemas to those of other peers. Then a network is constructed where peers are semantically connected through schema mappings. And query can be propagated to other peers with different schemas. The main thing is that query should be consistently answered and the reconciliation of different mappings in the semantic overlay network.

The key of this approach is seperating a logical from a physical layer. At the logical layer we support the operations necessary for maintenance and use of semantic overlay network and to support semantic interoperability. And we also provide a specific schema reconciliation technique, Semantic Gossiping. At the physical layer we provide efficient realization of the operations using a structured overlay network, P-Grid. So we need mapping of operations and data to the

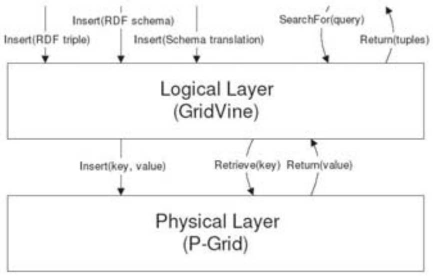


Figure 3: Two-layer model

physical layer. Important aspects of this mappings are:

- One can find a specific namespace for resources in the peer space.
- The mapping of data and metadata to routable keys. Data consists of RDF statements, whereas metadata are described with schemas and schema mappings.
- We can realize the forwarding of queries according to schema mappings. It's interesting to implement this approach with two strategies at the structured overlay network that are totally different from normal solutions. The solutions follow the structure of semantic overlay networks. One of these strategies is iterative and the other is recursive.

2 Overview of our Approach

2.1 Data Independence

The approach follows the principle of data independence by separating a logical layer, the semantic overlay for managing and mapping data and metadata schemas, from a physical layer consisting of a structured peer-to-peer overlay network for efficient routing of messages. The physical layer is used to implement various functions at the logical layer, including attribute-based search, schema management and schema mapping management. The separation of a physical from a logical layer allows us to process logical operations in the semantic overlay using different physical execution strategies. In particular we identify iterative and recursive strategies for the traversal of semantic overlay networks as two important alternatives. At the logical layer we support semantic interoperability through schema inheritance and semantic gossiping. Thus our system provides a complete solution to the implementation of semantic overlay networks supporting both scalability and interoperability.

P-Grid is an efficient, self-organizing and fully decentralized access structure that often use distrubuted hash table(DHT). Two basic functions of P-Grid are used by GridVine:

- Insert(key,value) store new data items based on a key
- Retrieve(key) retrieve data items according to the key

We build a semantic overlay network on the top of P-Grid so that we could take advantage of the two limited functions. This way we can also add metadata, schemas and schema mappings. In addition the system allows end-users to retrieve semantic information using appropriate query language. Here RDF/RDFS are chosen as encoding language. They are general-purpose language for representing information in the Web and probably become standard.

2.2 Decentralized Semantics

We can use the same concepts if we have global categories of resources and definition of vocabularies. On the one hand only centrally predefined schemas can be inherited. Such systems are not fit for capturing information sources in a network of autonomous and heterogeneous parties. On the other hand we cannot sense the power of P2P. And it's a pleasure if users can use their own schemas. In order to realize global semantic interoperability and search capabilities, we provide schema inheritance and Semantic Gossiping mechanisms.

Schema inheritance provides users the ability to reuse basic schemas. And a subset of schemas gain popularity while the others remain confidential. Thus we foster interoperability by reusing sets of conceptualizations.

Semantic Gossiping can be applied to foster semantic interoperability in decentralized systems. With this approach we want to develop global forms of agreement in an evolutionary and completely decentralized process that solely relies on pair-wise, local interactions: Participants provide translations between schemas they are interested in and can learn about other translations by routing queries (gossiping). To support the participants in assessing the semantic quality of the achieved agreements we develop a formal framework that takes into account both syntactic and semantic criteria. The assessment process is incremental and the quality ratings are adjusted along with the operation of the system. Ultimately, this process results in global agreement, i.e., the semantics that all participants understand.

Query forwarding can be implemented using several approaches. Iterative forwarding, where peers process series of translation links repeatedly. And recursive forwarding, where peers delegate the forwarding to other peers.

3 The P-Grid P2P system

GridVine uses our P-Grid P2P system as its physical layer. The main characteristics of P-Grids are

- they are completely decentralized, there exists no central infrastructure, all peers can serve as entry point to the network and all interactions are strictly local.
- they use randomized algorithms for constructing the access structure, updating the data and performing search; probabilistic estimates can be given for the success of search requests, and search is robust against failures of nodes.
- they scale gracefully in the total number of nodes and the total number of data items present in the network, equally for all nodes, both, with respect to storage and communication cost.

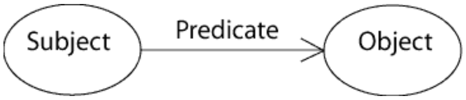


Figure 4: Graph

We use a binary tree to construct a P-Grid system. This is not a fundamental limitation as a generation of P-Grid to k-ary structures has been introduced, but will simplify the presentation. Every peer is represented as a leaf of the tree. And each leaf corresponds to a binary string. Each peer holds part of the overall tree. Every participating peer's position is determined by its path, that is, the binary bit string representing the subset of the tree's overall information that the peer is responsible for. And each peer stores a set of data items.

P-Grid supports two basic operations: Retrieve(key) and Insert(key, value). We could conclude from other available results of researches that these two operations are efficient also in general implmentation.

4 Semantic Support

4.1 Metadata Storage

Here we should introduce some background knowledge of RDF. The underlying structure of any expression in RDF is a collection of triples, each consisting of a subject, a predicate and an object. A set of such triples is called an RDF graph. This can be illustrated by a node and directed-arc diagram, in which each triple is represented as a node-arc-node link (hence the term "graph").

Each triple represents a statement of a relationship between the things denoted by the nodes that it links. Each triple has three parts:

- a subject,
- an object, and
- a predicate (also called a property) that denotes a relationship.

The direction of the arc is significant: it always points toward the object. The nodes of an RDF graph are its subjects and objects.

The assertion of an RDF triple says that some relationship, indicated by the predicate, holds between the things denoted by subject and object of the triple. The assertion of an RDF graph amounts to asserting all the triples in it, so the meaning of an RDF graph is the conjunction (logical AND) of the statements corresponding to all the triples it contains.

We introduce P-Grid specific URI schemas p-grid://, for resources, and p-grids://, for schema-elements in order to implement an application specific addressing space. This does not exclude other URI schemas in conjunction with the specific ones.

If all resources are identified by P-Grid URIs, a typical situation would be like that. The Subject is identified by a P-Grid key, while the predicate and domain refer to P-Grid specific RDF schemas. Then we can constrain the use of resources and manage the architecture easily. An example of such a statement would be the P-Grid resources 11110101(subject)is entitled(predicate) Rain, Stream, Speed(object). It would be translated into the XML syntax of RDF.

```
<?xml version="1.0"?>
<rdf:RDF xmlns="http://www.w3.org/1999/02/22-rdf-syntax-ns#">
<rdf:Description rdf:about="pgrid://11110101">
<Title xmlns="pgrids://01001101:bmp#">Rain, Stream and Speed</Title>
</rdf:Description>
</rdf:RDF>
```

Since most RDF query language is based on constraint searches on the triple's subject, predicate or object, we reference individual triple three times. That means we get seperate keys for them.

4.2 Schema Definition And Storage

RDFS is used to encode the schematic information in GridVine. So we should know some basic concepts about RDF Schema. RDF's vocabulary description language, RDF Schema, is a semantic extension of RDF. It provides mechanisms for describing groups of related resources and the relationships between these resources. RDF Schema vocabulary descriptions are written in RDF using the terms described in this document. These resources are used to determine characteristics of other resources, such as the domains and ranges of properties. The RDF vocabulary description language class and property system is similar to the type systems of object-oriented programming languages such as Java. RDF differs from many such systems in that instead of defining a class in terms of the properties its instances may have, the RDF vocabulary description language describes properties in terms of the classes of resource to which they apply. This is the role of the domain and range mechanisms described in this specification. For example, we could define the eg:author property to have a domain of eg:Document and a range of eg:Person, whereas a classical object oriented system might typically define a class eg:Book with an attribute called eg:author of type eg:Person. Using the RDF approach, it is easy for others to subsequently define additional properties with a domain of eg:Document or a range of eg:Person. This can be done without the need to re-define the original description of these classes. One benefit of the RDF property-centric approach is that it allows anyone to extend the description of existing resources, one of the architectural principles of the Web.

Currently GridVine schemas allow to declare a specific class with an arbitrary number of properties. And they have the specific class as domain. The class *Schema* comes from the basic class *P-Grid Data Item.* Then we can process references to all addresable subjects. Each property derives from a generic property *P-GridDataItemProperty.* The whole relationship is represented in the following figure.

We create distinct RDFS files for various categories that derived by the users. The definition of a subclass and all its related properties are in the same group. It's necessary that we create a unique identifier for the category by concatenating the path π(p) of the category creator and the name of its class. It looks like that: Insert(rdf_schema) = Insert(Hash(π(p)):class_name), rdf_schema).

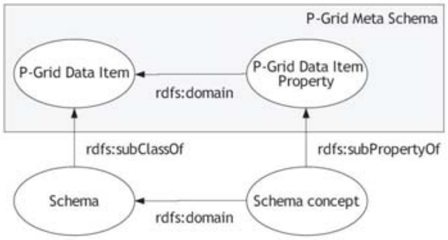


Figure 5: Relations of a P-Grid schema to the P-Grid meta schema

4.3 Resolving Queries in GridVine

An RDF model is graph, often expressed as a set of triples. An RDQL consists of a graph pattern, expressed as a list of triple patterns. Each triple pattern is comprised of named variables and RDF values (URIs and literals). An RDQL query can additionally have a set of constraints on the values of those variables, and a list of the variables required in the answer set.

For example, the following RDQL query:

SELECT ?y WHERE ($\langle p - grid : //01101000\rangle$, ?y, ?z)

would return all the predicates used to annotate date item 01101000. Such query will be called native query routing one message through the P-Grid infrastructure. We can clarify, how can that implemented. At first we have a message that consists of the query and the address of the peer p from which the query originates. Then the message will be routed as explained before to the peer q responsible for storing data item d with key(d) = 01101000. After receiving the message the peer q checks its local database and sends back to p the set of triples as answer.

More generally, triple patterns consist of expressions where the subject, the predicate and the object can all be replaced by variables which may be not bound.

5 Semantic Interoperability

5.1 Schema Inheritance

Users can freely derive categories from the existing ones. We impose that the new classes representing the subcategory subclasses the base category class. From rdf:subClassOf semantics we know the subcategory automatically inherit the properties of the base category. The properties are defined through the definition of the property ranges. That means, the instances of this class are also instances of the base class. At the same time subcategories may introduce the properties directly related to the subclass. Here is an example. That shows how is a category for annotating JPEG files derived from a more generic category of image files.

The basic concept is that a subcategory may in turn serve as a super-category for another category. Since super-category contains subcategories, subcategories of a give category can be used as instance of base category. And the operations

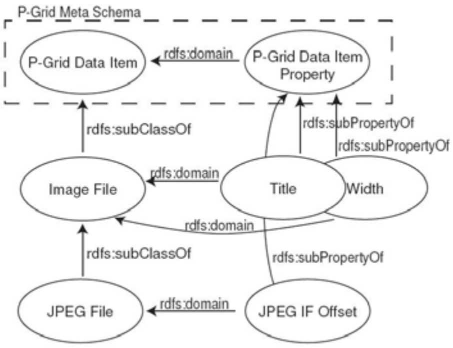


Figure 6: A simple example of category inheritance

on the base category influence the instances of subcategories. Thus, we derive a sets of semantically interoperable schemas from properties shared by all the derivations of base schema.

5.2 Semantic Gossiping

Here is the other method how the peer-to-peer system architectures can be applied to tackle the problem of semantic interoperability in the large, driven in a bottom-up manner by the participating peers.

We assume that there exists a communication facility among the participants that enables sending and receiving of information, i.e., queries, data, and schema information. The underlying system could typically be a P2P system, but also a federated database system or any system of information sources communicating via some communication protocol. We assume that the peers P are all using semantically heterogeneous schemas S to represent their information. The case where multiple peers share the same schemas leads to possible optimizations of the approach, but is not conceptually different; therefore we ignore it in the following. To be semantically interoperable, the peers maintain knowledge about the relationships among their schemas. This knowledge can be given in the form of views, for example.

For peers $p_1, p_2 \in P$, with schemas $S_1, S_2 \in S$, the relationship is given by a query $q_{1,2}$ applicable to schema S_2 and producing results according to schema S_1. We assume that skilled experts supported by appropriate mapping tools are able to provide these mappings. The direction of the mapping and the node providing a mapping are not necessarily correlated. For instance, both node p_1 or p_2 might provide a mapping from schema S_1 to schema S_2, and they may exchange this mapping upon discretion. During the operation of the system, each peer has the opportunity to learn about existing mappings and add new ones. This means that a directed graph of mappings will be built between the peers along with the normal operation of the system (e.g., query processing and forwarding in a P2P system). Such a mapping graph has two interesting properties:

(1) based on the already existing mappings and the ability to learn about mappings, new mappings can be added automatically by means of transitivity,

and

(2) the graph has cycles. The first property essentially means that we can propagate queries towards nodes to which we have no direct translation link. This is what we denote as semantic gossiping. While doing so, the nodes may check whether the translated query is worth to be propagated at all. It may occur that the translated query will return no or no meaningful results because of schema mismatches. This can be checked at a syntactic level. We do this by introducing the concept of syntactic distance, which analyses to what extent a query is preserved after translation, and use the syntactic distance as a criterion to decide whether or not to propagate a query. Following this approach, we allow peers in GridVine to create translation links mapping one schema onto another. Translation will be used to propagate queries from one semantic domain to another.

We need to retrieve all the schemas related by translation links starting from a given schema in order to forward queries properly. Thus, we index the translations based on their base schemas:

Insert(owl_file) = Insert(Hash($\Pi(p_{source}$):source_class_name),owl_file).

A query is forwarded from a given domain and get transformed following translation links, until the translation is considered as too different. After each translation a new query starts.

6 Related Work

Hyperion is an ongoing project which represents an architecture and a set of challenges for peer database management systems. These systems team up to build a network of nodes (peers) that coordinate at run time most of the typical DBMS tasks such as the querying, updating, and sharing of data. Such a network works in a way similar to conventional multidatabases. Conventional multidatabase systems are founded on key concepts such as those of a global schema, central administrative authority, data integration, global access to multiple databases, permanent participation of databases, etc. Instead, this proposal assumes total absence of any central authority or control, no global schema, transient participation of peer databases, and constantly evolving coordination rules among databases. In this work, we describe the status of the Hyperion project, present our current solutions, and outline remaining research issues.

The Edutella project addresses these shortcomings of current P2P applications by building on the W3C metadata standard RDF. The project is a multi-staged effort to scope, specify, architect and implement an RDF-based metadata infrastructure for P2P-networks based on the recently announced JXTA framework. The initial Edutella services will be Query Service (standardized query and retrieval of RDF metadata), Replication Service (providing data persistence / availability and workload balancing while maintaining data integrity and consistency), Mapping Service (translating between different metadata vocabularies to enable interoperability between different peers), Mediation Service (define views that join data from different meta-data sources and reconcile conflicting and overlapping information) and Annotation Service (annotate materials stored anywhere within the Edutella Network).

There are many other related work. For example, PeerDB,Piazza peer data management project etc. All the above approaches address semantic interoperability but offer limited scalability. Other approaches address scalability but do not deal with semantic interoperability.

7 Summary

GridVine is the first semantic overlay network based on an scalable, efficient and totally decentralized access structure. And semantic interoperability is provided although users have their own schemas. And it is important that logical layer and physical layer are seperate so that we could base the semantic overlay network on any other physical infrastructure.

References

[1] K.Aberer. Efficient Search in Unbalanced, Randomized Peer-To-Peer Search Trees. Technical Report IC/2002/79, Swiss Federal Institute of Technology, Lausanne(EPFL), 2002. http://www.p-grid.org/Papers/TR-IC-2002-79.pdf

[2] K. Aberer, P. Cudré-Mauroux, A. Datta, Z. Despotovic, M. Hauswirth, M. Punceva, and R. Schmidt. P-grid: A self-organizing structured p2p system. ACM SIGMOD Record, 32(3), 2003.

[3] K. Aberer, P. Cudré-Mauroux, and M. Hauswirth. Starting making sense: The Chatty Web approach for global semantic agreements. Journal of Web Semantics, 1(1), 2003

[4] K. Aberer, P. Cudré-Mauroux, and M. Hauswirth. The Chatty Web: Emergent Semantics Through Gossiping. In International World Wide Web Conference(WWW), 2003

[5] K. Aberer and M. Punceva. Efficient Search in Structured Peer-to-Peer Systems: Binary v.s. k-ary Unbalanced Tree Structures. In International Workshop On Databases, Information Systems and Peer-to-Peer Computing. collocated with VLDB 2003.,2003

[6] M. Arenas, V. Kantere, A. Kementsietsidis, I. Kiringa, R. J. Miller, and J. Mylopoulos. The Hyperion Project: From Data Integration to Data coordination. SIGMOD Record, 32(3), 2003.

[7] M. Cai and M. Frank. RDFPeers: A Scalable Distributed RDF Repository based on A Structured Peer-to-Peer Network. In International World Wide Web Conference(WWW), 2004.

[8] Karl Aberer(ed.). Special issue on peer to peer data management. ACM SIGMOD Record, 32(3), 2003.

[9] M. Ehrig, P. Haase, R. Siebes, S. Staab, H. Stuckenschmidt, R. Studer, and C. Tempich. The SWAP Data and Metadata Model for Semantics-Based Peer-to-Peer Systems. In Multiagent System Technologies(MATES), 2003.

[10] J. M. Hellerstein. Toward network data indepence. ACM SIGMOD Record, 32(3), 2003.

[11] R. Huebsch, J. M. Hellerstein, N. Lanham, B. T. Loo, S.Shenker, and I. Stoica. Querying the Internet with PIER. In Conference On Very Large Data Bases(VLDB), 2003.

[12] B. C. Ooi, Y. Shu, and K. L. Tan. Relational Data Sharing in Peer-based Data Management Systems. ACM SIGMOD Record, 32(3),2003.

[13] World Wide Web consortium. Semantic web activity at the w3c. http://www.w3.org/2001/sw/