CS654 Project Proposal

Critical Examination of Service Discovery Techniques in Peer-to-Peer Systems

N.M. Mosharaf Kabir Chowdhury

(20262829)

nmmkchow@cs.uwaterloo.ca

June 7, 2007

Web services are the new paradigm of distributed computing because of their promise toward interoperability of applications and integration of large scale distributed systems. On the other hand, with the ever increasing number of computers all around the globe, peer-to-peer(P2P) networks have gained large popularity capitalizing on their open, distributed, inter-operable, cost-effective and resource-friendly nature. As a consequence, the problem of finding a way to effectively combine web services with P2P technologies has gained much attention because of its potential to be an effective solution to business integration problems. Hence, different large-scale service infrastructures utilizing P2P systems have been proposed by many research groups. Service discovery, an integral part of any large-scale service infrastructure, ensures that all the services provided by different providers within the infrastructure are collected, organized and published in a way that the clients can efficiently search and find their desired services.

Service discovery systems generally consist of three types of entities - *servers*, *clients* and *middle agents* [3]. This *'middle agent'* can be *matchmakers*, *directory* service, *blackboard* agents or *brokers*. This process of discovery has following general steps [2]: 1) Bootstrapping, 2) Service advertisement, 3) Querying, 4) Lookup and 5) Service handle retrieval. There are some other discovery systems that do not have any middle agents in between clients and servers.

Directory architecture adopted by different service discovery approaches can broadly be classified as *centralized* and *decentralized* [2] depending on how the directory/registry is stored and managed. Centralized UDDI is an example of such architecture. The problem

with this approach is that the centralized server becomes a performance bottleneck and forms a single point of failure. On the hand, there are several decentralized variants like *replicated* (e.g. INS), *distributed or partitioned* (e.g. SLP, Jini, UPnP) and *hybrid* (e.g. TWINE) [1]. The obvious advantage is the removal of the bottleneck. P2P networks is widely used in this regard because of its inherent distributed nature and favorable properties like autonomy and scalability. Advanced research in this context include DAML-S[5], MWSDI[7] etc.

The most challenging task in P2P based service discovery is to deal with multiple registries(replicated or partitioned). Apart from the issues regarding to autonomous behavior of the peers, synchronization between copies or efficiently partitioning and placement, just searching for a web service in this large distributed environment requires close attention. Most of the service discovery technologies use either some DHT-based routing like *CAN*, *CHORD*, *PASTRY*, *TAPESTRY* etc or non-DHT ones like *SkipNet*, *SkipGraph* etc in case of structured P2P networks and *Flooding* or *Random-Walk* for unstructured networks. One of the aims of this project is to evaluate and identify relative pros and cons of searching for a service using such techniques. An inherent problem with existing solutions is their inability to effectively do partial matching. As a result, service discovery for queries that match partially with the advertised services becomes a daunting task. Recently, there has been a research on partial matching in P2P systems that formalizes the problem as DPM [1] and proposes two solutions, DPMS and Plexus. This project aims to evaluate the ideas of DPM in the context of service discovery.

Another challenging task is the heterogeneity between services provided by the servers. Different kinds of heterogeneity has been identified as follows [6]: 1) Technological heterogeneity, 2) Ontological heterogeneity and 3) Pragmatic heterogeneity and different methods to handle such heterogeneities has also been proposed. Most important of the heterogeneities is the data model used in the services. Two alternative viewpoints[4] in this aspect are: 1) Information Retrieval Approach and 2) Semantic Approach. In recent years, the later approach has gained much attention. Study of different semantic approaches is also under consideration.

The current research plan includes the following milestones -

- Develop an understanding of the available frameworks for service discovery
- Critical examination of the pros and cons of the frameworks
- Extending DPM concept to service discovery problem to facilitate partial matching
- Study of the different viewpoints for data modeling used in service discovery

References

[1] R. Ahmed, *Efficient and flexible search in large scale distributed systems*, Ph.D. thesis, University of Waterloo, 2007.

[2] R. Ahmed, R. Boutaba, F. Cuervo, Y. Iraqi, T. Li, N. Limam, J. Xiao, and J. Ziembicki, *Service discovery protocols: A comparative study*, Proceedings of the IFIP/IEEE International Symposium on Integrated Network Management (IM) Application Sessions (Nice, France), 2005, pp. 22–37.

[3] K. Decker, K. Sycara, and M. Williamson, *Middle-agents for the internet*, Proceedings of the 15th International Joint Conference on Artificial Intelligence (Nagoya, Japan), 1997.

[4] John Garofalakis, Yannis Panagis, Evangelos Sakkopoulos, and Athanasios Tsakalidis, *Web service discovery mechanisms: Looking for a needle in a haystack?*

[5] Takuya Nishimura Massimo Paolucci, Katia Sycara and Naveen Srinivasan, *Using daml-s for p2p discovery*, Proceedings of the First International Conference on Web Services (ICWS'03), October 2003, pp. 203–207.

[6] Sven Overhage, *On specifying web services using uddi improvements*, 3rd Annual International Conference on Object-Oriented and Internet-based Technologies, Concepts, and Applications for a Networked World, 2002, pp. 535–550.

[7] Abhijit A. Patil, Swapna A. Oundhakar, Amit P. Sheth, and Kunal Verma, *Meteor-s web service annotation framework*, WWW '04: Proceedings of the 13th international conference on World Wide Web (New York, NY, USA), ACM Press, 2004, pp. 553–562.