Towards a Comprehensive Peer-to-Peer Communication Model

Jesus Pajaro
Department of Computer Science
The University of Texas at El Paso
El Paso, Texas 79968
jesuspaj@utep.edu

Brian J. d'Auriol
Department of Computer Science
The University of Texas at El Paso
El Paso, Texas 79968
dauriol@acm.org

Abstract ***Peer-to-Peer computing is a model consisting of cooperating distributed processing units terms peers. Characteristics of this computing model include lack of centralized control and hierarchical organization of the peers. Instead, peers take on roles of client/server, provider/requester or spotter (router) as needed. Peers need to be equipped with hardware and software capabilities that provide these functions. Peer-to-Peer computing is a relatively recent and emerging computing model. One among the interesting research issues of this computing model is the communication requirements induced by Peer-to-Peer applications. Peer-to-Peer applications can be described in terms of a service layer model that abstracts the required services needed to perform all aspects of that application together with a physical layer that abstracts the required physical components involved in the services. This paper proposes the two respective models of Service Model and Physical Model to provide these abstractions in a comprehensive and cohesive manner. The main point of emphasis is the two-layer modeling of communication requirements induced by Peer-to-Peer applications.***

Keywords: Peer-to-Peer, Communication, Services

1 Introduction

Peer-to-Peer systems have become attractive for providing the following advantages: cost reduction due to the elimination of expensive servers, for example, server installation, hardware, software, and server administrative and maintenance costs might be avoided [1]; increased cooperation, scalability, robustness and fault tolerance due to the distributed nature of the control [2]; and increased aggregation of computing capability due to the emergent virtual computing resources that may, for example, consist of many users [3]. One highly interesting and unique characteristic is that Peer-to-Peer computing often results in a computing community that allows peers to join and leave the network at any time, thereby creating highly dynamic and loose network structures.

As an emerging area of research, Peer-to-Peer computing raises many interesting and challenging research problems in distributed systems. One of these is the lack of identification and characterization of the physical components and communication processes that are involved in a Peer-to-Peer network. The lack of a characterization methodology makes difficult the description and evaluation of communication costs involved in a Peer-to-Peer network. As Peer-to-Peer computing matures, there is a need to characterize Peer-to-Peer applications in terms of the processes involved during communication. It is reasonable to consider that such a characterization would lead to a methodology that could be applied to Peer-to-Peer communication cost modeling.

Peer-to-Peer applications can be described in terms of a service layer model that abstracts the required services needed to perform all aspects of that application together with a physical layer that abstracts the required physical components involved in the services. This paper proposes the two respective models of Service Model and Physical Model to provide these abstractions in a comprehensive and cohesive manner. The main point of emphasis is the two-layer modeling of communication requirements induced by Peer-to-Peer applications.

This paper is organized as follows. Section 2 describes a basic and typical Peer-to-Peer computing model. In Section 3, we propose the Peer-to-Peer Characterization Model (P2PCM) that consists of the physical and service layers. Conclusions are given in Section 4.

2 Peer-to-Peer Computing

Peer-to-Peer (P2P) is a computing model that is described as a collective computing environment formed by P2P systems [4]. P2P systems are classified as distributed systems without any centralized control or hierarchical organization, where the role of every peer has equivalent functionality [5]. In other words, a given peer may act as a server or as client at any given time in a P2P system.

The P2P computing model is a distributed system where every peer has two flexible roles at any given time: server and client [2]. However, in the P2P computing model, peers are allowed to switch roles. At times, a peer may act as a server or as a client. This immitates the functions of a Client/Server network model but provides additional flexibility and freedom to the network structure. A peer can also interact with other peers on the P2P network in more complex ways, for example, by partitioning a complex computational problem into a set of subtasks that can run concurrently in the network to accomplish a given common task [4].

Currently, several different technologies exist that provide a common set of functionalities and building blocks for constructing higher-level P2P applications. Two of these technologies are the Microsoft .NET Framework [6, 7] and the JXTA project [2, 8]. The Microsoft .NET Framework aims to provide a universal networking platform for distributed systems to integrate and connect resources and Web Services by allowing peers to use Internet standard protocols [6]. Somewhat similarly, project JXTA provides a set of standard protocols for building P2P applications. In addition, it provides a set of common protocols and mechanisms used for peers to build interoperable P2P applications on top of the basic and generic P2P building blocks [2].

3 Peer-to-Peer Characterization Model

The P2P Characterization Model (P2PCM) proposed in this paper consists of the physical layer and service layer that integrates the network components involved in the communication process and the functionality of the necessary services provided to P2P applications.

The Physical Model presented in Section 3.1 provides an abstraction of the essential communication processes that occur in P2P computing. Some of the components abstracted are software, spotter/gateway peers, and the provider and requester peer.

The Service Model presented in Section 3.2 abstracts the functions and services of every P2P application request. Application requested services are organized in a multi-layer hierarchy within the Service Model. Each of these layers are connected with the associated counterparts in the Physical Model.

Interactions between these two models are investigated in Section 3.3.

3.1 Physical Model

The Physical Model consists of the following components.

- Peers are the processing units that participate, collaborate and communicate to expose and consume P2P applications. Computers, laptops, PDAs or cellular phones may all act as peers. Peers can be classified according to their functionality into provider peers, requester peers and spotter peers. A provider peer is the device where a given application resides and executes. The provider peer advertises its resources and makes those resources available to other peers in the community. A requester peer is the device that has an executing application interested in consuming a particular resource in the community. The spotter peers are the devices which propagate published service advertisements or discover service resources over the network. Spotter peers could also be gateways.

- The P2P network is a collective environment of peers wherein every peer possesses a network port and the network interconnection to connect to other peers. The interconnection medium could be Ethernet-based or other similar medium. Specific allocation of ports is required. Network protocols such as TCP/IP, HTTP or SMTP are used.

- P2P Infrastructure Technology (e.g. .NET or JXTA) is the software that provides for a common platform capable of supporting the basic Peer-to-Peer functions of communication, message exchange and application interoperability.

  - Messages support the communication amongst peers. Messages should include information about the routing path, the specific spotter peers involved and the specific network ports utilized in the

transmission. XML is typically used for message representation. Messages are transmitted as a stream of bytes throughout the P2P network.

- Service advertisements are abstractions that encapsulate the functional description of a particular service provided. Advertisements are represented as messages.

Communication is the action of sending and receiving messages. It is defined as the interaction of peers along the network. Peers interact and communicate across the network establishing sequences of communication processes. These sequences occur in one of the two following contexts: (a) indexing and discovery of a service advertisement, and (b) requesting and utilization of such resources.

The following expression describes the sequence of communication processes that are performed by peers during the indexing and discovery of a service advertisement:

$$P_P \longrightarrow \underset{\circlearrowleft}{P_{Si}} \longrightarrow P_{Sn} \longleftrightarrow P_R$$

where P_P denotes the provider peer, P_{Si} for $1 \leq i \leq n-1$ as well as P_{Sn} denote spotter peers, P_R denotes the requester peer, and, the connecting arrows denote communication links. The indexing involves P_P and the spotter peers, P_{Si} and P_{Sn}. The discovery involves the spotter peers, P_{Si} and P_{Sn} and P_R. Note that the requester peer must initiate a message to the n-th spotter in order to have that spotter forward the service advertisement.

Similarly, the following expression describes the sequence of communication processes that are performed by the peers involved during the utilization of a given resource:

$$P_P \longleftrightarrow \underset{\circlearrowleft}{P_{Si}} \longleftrightarrow P_{Sn} \longleftrightarrow P_R$$

A physical model scenario is included to better study the sequencing interactions that take place during an application's service request. This scenario is depicted in Figure 1 and subsequently discussed.

1. A provider peer must create a P2P application that implements the basic set of P2P services (collaboration, content sharing and hardware resource sharing). Once the application is created and implemented in a physical machine, the core functionality of the service and the methods corresponding to the P2P services are exposed and thus, accessible to other peers. At this step, the provider peer P_P is the only active component in the communication sequence and thus, implying no communication processes.

2. A P2P application service provider exposes a given application functionality for others to consume. In order to make the application available to other peers, the provider must publish a service advertisement to the P2P community. For instance in .NET, a WSDL service description might be used as an advertisement which encapsulates the Web Service specification and the format of the messages to be exchanged among peers. The service advertisement has to be created by the provider and published to the P2P community through a spotter peer, P_{S1} as described in the following expression.

$$P_P \longrightarrow P_{S1}$$

3. The provider can use a centralized indexing mechanism such as UDDI [6] or a decentralized indexing mechanism such as CAN [9] or CHORD [5] to advertise the application's service. A decentralized indexing mechanism approach is illustrated in Figure 1 wherein P_{S1} is a spotter peer responsible to propagate the service advertisement specification to P_{S2}. Next P_{S2} is responsible to propagate the service advertisement specification to P_{S3}. The active components involved during this process are the chain of spotter peers responsible advertisement propagation. This is indicated in the following expression:

$$P_{S1} \longrightarrow P_{S2} \longrightarrow P_{S3}$$

4. A P2P application should be locatable by other peers in the network. "...without the knowledge of existence of a peer or a service on the network, there is no possibility for a device [peer] to engage that service" [2]. By using a centralized discover mechanism such as UDDI or a decentralized mechanism such as CAN or CHORD implemented using the spotter peers, the requester peer is capable of querying and retrieving the functionality of a given application described by the service advertisement

specification document as illustrated in the following expression:

$$P_{S3} \longleftrightarrow P_R$$

5. Once the service advertisement specification is retrieved from one of the spotter peers (P_{S3}) in the P2P community in Figure 1, the requester peer creates an application based on the specification document. This application utilizes messages based on the format already described by the service advertisement definitions. The application's interface which is exposed publicly to the community and becomes accessible via messages. At this step, the requester peer P_R is the only active component in the communication sequence.

6. The requester peer can establish a direct communication with the provider peer, by extracting the URL if available, from the service advertisement specification document as illustrated in Figure 1. However, a virtual channel may be established to the application endpoint through the path provided by the spotter/gateways peers, this is also illustrated in the figure: P_{S3}, P_{S2}, and P_{S1} respectively. The following expression shows the sequence of communications for the construction of a communication channel that is established throughout the spotter peers. This channel serves as a virtual link making the interaction possible between the requester and the provider peers ($i = \{1, 2\}$):

$$P_P \longleftarrow \underset{\circlearrowleft}{P_{Si}} \longleftarrow P_{S3} \longleftarrow P_R$$

7. The communication is established with the provider peer, and thus the P2P application located at the provider peer's endpoint is capable of receiving messages requesting a specific service and/or sending messages that encapsulate the corresponding results to the requester. The information transmitted from the requester peer to the application and vice-versa will flow using a communication protocol, for example, TCP/IP. The information consists of XML documents embedded into messages and transmitted using the communication protocols specified by the service application specifications. The following expression states the provider-requester peer communication interaction in this sequence of communications, ($i = \{1, 2\}$):

$$P_P \longleftrightarrow \underset{\circlearrowleft}{P_{Si}} \longleftrightarrow P_{S3} \longleftrightarrow P_R$$

3.2 Service Model

P2P applications may perform a variety of different tasks, however, they all preserve many of the same functions such as discovery of peers and/or services, service advertising, exchange of messages, etc. The Service Model abstracts the general behavior of P2P applications into a four layer model as illustrated in Figure 2. These layers are:

- The P2P Service Infrastructure Technology together with the indexing and discovery services provide the set of core services needed by any P2P application. P2P applications need sophisticated discovery mechanisms to enable peers to find, identify and communicate with other peers [10]. Indexing and discovery mechanisms can be based on, for example, Chord [10], Pastry [10], and Content Addressable Networks (CAN) [9, 10]. The green part of the inner layer in Figure 2 corresponds to this mechanism.*

- A P2P community is defined as a dynamic collection of peers that collaborate with each other for the pursuit of a common set of interests. A virtual P2P community can be constructed to provide peer accessibility to the inherent set of P2P computing services available in the community. P2P communities provide peers the ability to publish, discover and consume services within a specific community. Communities are useful in structuring the information storage space, discovering resources and pruning the search space [11]. Communities are implicit, self-organizing, dynamic and constantly changing, forming, or breaking down due to the *ad hoc* nature of the network [11]. The yellow layer in Figure 2 represents the peer community and abstracts away from the necessary services to 'get the job done'.

- Existing P2P applications can be reduced to a set of basic characteristics [4]. Typically, P2P computing technology enables three main P2P

*Gray-scale rendering has lighter shades of gray towards the inner part of the figure, the service layer is the darkest gray-shaded region.

usage models. A usage model describes characteristics unique to specific P2P services. Three usage models are identified in [4, 12]: (a) content sharing, (b) collaborative computing, and (c) hardware resource sharing. These are illustrated in the orange layer in Figure 2. A wide variety of applications can be built using these three basic services. Applications may use overlapping basic services.

- P2P applications represent the outer-most layer colored in pink in Figure 2. Applications have the following characteristics:

  1. make use of the basic P2P services such as hardware resource sharing, content sharing and collaborative computing;
  2. participate and collaborate with other applications in a P2P community;
  3. locate and/or advertise services by using the indexing and discovery mechanisms; and
  4. make use of a specific P2P infrastructure technology.

3.3 Service and Physical Model Interactions

Figure 3 illustrates the integration of the Physical and Service models described previously. This combined approach allows: (a) characterization of the communications as carried out through the P2P network in the form of messages tied together with those services that are used by applications, (b) better understanding of the necessary supporting services and technologies for P2P applications, and (c) specific and exact metrics to be designed for later communication cost modeling. The interactions here are described in terms of the scenario that was developed earlier.

In the top part of Figure 3, the seven steps in the communication sequencing of the Physical Model are associated with their corresponding service layers. This determination arose from an analysis of the service communication requirements in the Physical Model. This analysis is shown in the lower part of the figure, the seven steps are, as before, shown. These interactions characterize the P2P Characterization Model from the point of view of the Physical Model.

Similarly, the physical components are color-identified by their associated service layers. In the figure, the applications layer (pink color) only involves the provider and requester. The discovery mechanism through the spotter peers is shown in the green color; note, the blue colored direct link between the two applications.

4 Conclusions

Peer-to-Peer systems have become attractive for providing some cost reductions due to the elimination of expensive servers and increased computing capability and scalability due to the distributed nature of the control. However, the nature of the distributed control, the requirements for infrastructure support, for example, indexing and discovery as well as the highly dynamic and loose network structures all contribute to issues about communications support.

A model called the P2P Characterization Model (P2PCM) is proposed in this paper to better describe the communication requirements of P2P applications. Peer-to-Peer applications are described in terms of a service layer model that abstracts the required services needed to perform all aspects of that application together with a physical layer that abstracts the required physical components involved in the services. These two models are component sub-models to the P2PCM. Details of each of these sub-models are presented.

Communications are described in terms of the services that are used by applications. To this end, this paper describes a sequencing component of the Physical Model that was later incorporated into the interaction analysis of the P2PCM. Figure 3 represents the primary results of this analysis, that is, the integration of the two sub-models.

The strengths of this paper lie in the comprehensive characterizations of application communication requirements. To the best of our understanding, this approach has not yet been considered in the literature. Hence, this paper establishes the beginning portions of a formal communications model. The weakness of the paper is that it does not go far-enough to describe metrics nor cost modeling of such communications. This motivates us to continue enhancing the P2PCM that was proposed in this paper towards a communication model for cost analysis.

References

[1] D. Liben-Nowell, H. Balakrishnan, and D. Karger, "Analysis of the evolution of peer-to-peer systems," in *Proc. of the twenty-first annual symposium on Principles of distributed computing*, (Monterey, California), pp. 233–242, July 2002.

[2] B. J. Wilson, *JXTA*. New Riders Publisher, first ed., June 2002.

[3] SETI@home, 2003. http://setiathome.ssl.berkeley.edu.

[4] K. Kant, R. Iyer, and V. Tewari, "A framework for classifying peer-to-peer technologies," in *Proc. of the 2nd. IEEE/ACM International Symposium on Cluster Computing and the Grid (CCGRID'02)*, 2002.

[5] I. Stoica, R. Morris, D. Karger, M. F. Kaashoek, and H. Balakrishnan, "Chord: A scalable peer-to-peer lookup service for internet applications," in *Proc. of the ACM-SIGCOMM 2001 Conference. Applications, Technologies, Architectures, and Protocols for Computer Communications*, (San Diego, California, USA), August 2001. vol.31, no.4.

[6] S. Short, *Building XML Web Services for the Microsoft .NET platform*. Redmont, Washington: Microsoft Press, 2002.

[7] ".NET framework developer's guide." http://msdn.microsoft.com.

[8] B. Traversat, M. Abdelaziz, M. Duigou, J.-C. Hugly, E. Pouyoul, and B. Yeager, "Project jxta virtual network," Feb. 2002. Sun Microsystems, http://www.jxta.org/project/www/docs/JXTAprotocols.pdf.

[9] S. Ratnasamy, P. Francis, M. Handley, R. Karp, and S. Shenker, "A scalable content-addressable network," in *ACM SIGCOMM'01*, (San Diego, California, USA), pp. 161–172, August 2001.

[10] M. Kelaskar, V. Matossian, P. Mehra, D. Paul, and M. Parashar, "Study of discovery mechanisms for peer-to-peer applications," in *Proc. of the 2nd IEEE/ACM International Symposium on Cluster Computing and the Grid(CCGRID'02)*, (Piscataway, NJ, USA), pp. 444–5, May 2002.

[11] M. Khambatti, K. Ryu, and P. Dasgupta, "Peer-to-peer communities: Formation and discovery," in *Proc. of the 14th IASTED International Conference Parallel and Distributed Computing and Systems*, (Cambridge, MA, USA), pp. 166–73, Nov. 2002.

[12] D. Barkai, "Technologies for sharing and collaborating on the net," in *Proc. of the First International Conference on Peer-to-Peer Computing*, (Linkoping, Sweden), pp. 13–28, Aug. 2001. IEEE.

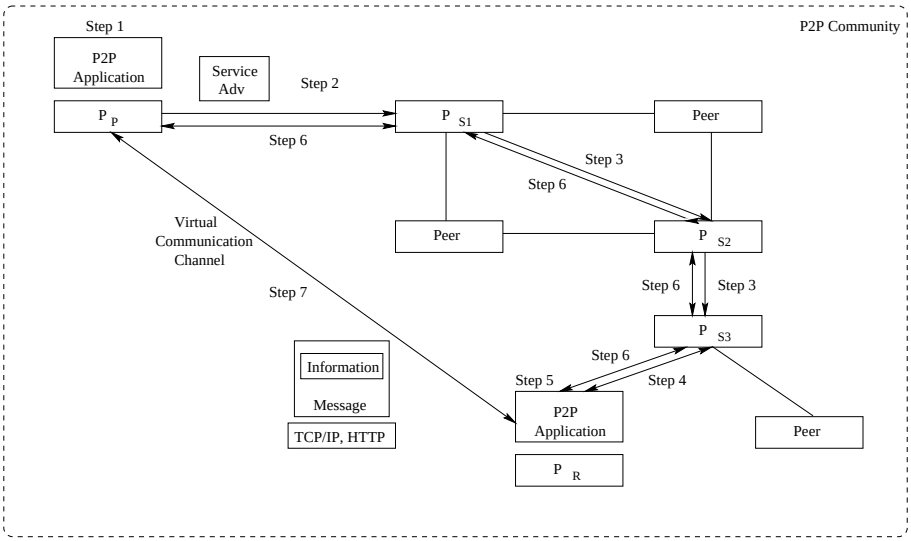


Figure 1: Physical Model Scenario

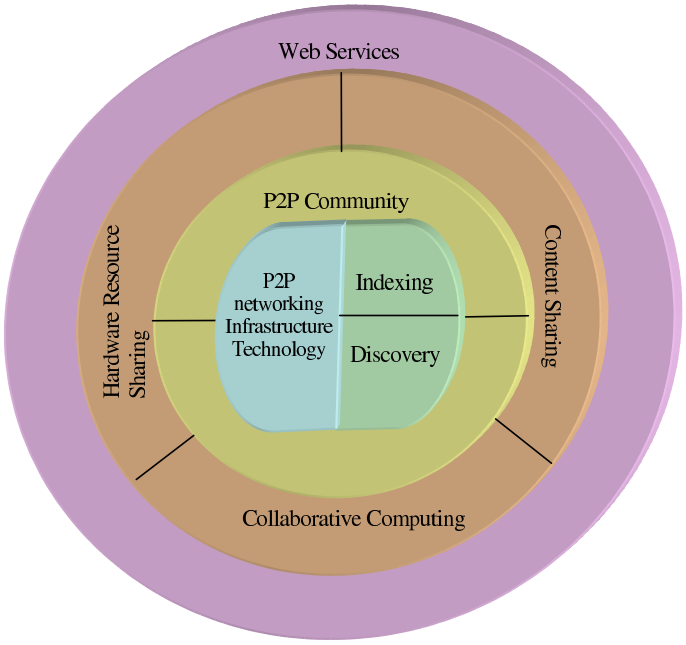
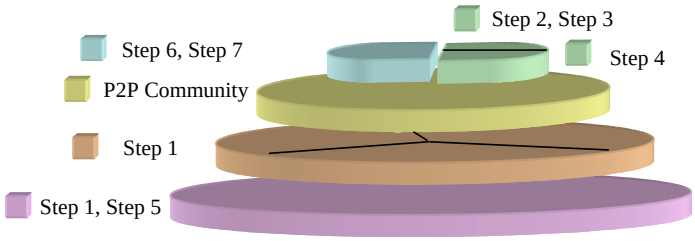


Figure 2: Service Model



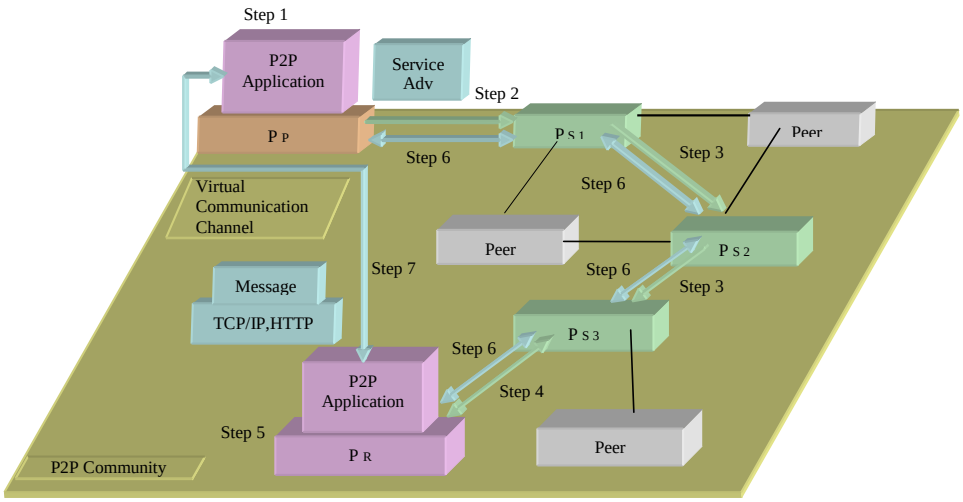


Figure 3: Peer-to-Peer Characterization Model