Lesson Proper for Week 13

Peer-to-Peer Systems: The Present and the Future

The client-server paradigm has long been the mainstay of most applications. Peer-to-peer (P2P) is a new paradigm that supports applications that provide file-sharing, a content exchange like music, movies, and programs, but also distributed computing and Internet-based telephony. A peer-to-peer (P2P) system is a self-organizing system of equal, independent entities (peers) that intends to share scattered resources in a networked environment without relying on central services. Peer-to-peer is a system with decentralized self-organization and resource consumption.

Unlike the client-server paradigm, the peer-to-peer notion is entirely decentralized and self-organizing. Peer-to-peer concepts face several obstacles, such as durable and scalable distributed systems and innovative services. Peer-to-peer apps account for 50% of Internet traffic, and in extreme circumstances up to 75%. The expanding Internet, consumers, and bandwidth necessitate more applications. The client-server paradigm needs a lot of effort and resources. Internet-based applications identify three main characteristics:

- Scalability.
- Security and reliability.
- Flexibility and quality of services.

Client-server programs struggle to keep up with Internet innovation. Due to its location within the network architecture, the client-server centralized model is easily attacked and difficult to modify. All of the above indicates a paradigm shift from client-server to peer-to-peer.

Unstructured Peer-To-Peer Systems

Was the first iteration of unstructured peer-to-peer file sharing. One of these was Napster, which used a meta server and servers to locate data items, before sending them directly between peers. It works by flooding all peers in the system until the relevant data is located. Peer-to-peer networks don't rely on single transport infrastructure. A peer-to-peer system uses TCP or HTTP connections to create an overlay structure for content distribution. A central server stores and provides material in client-server systems. Peer-to-peer search finds the needed content at a peer and provides the peer's IP address to the searching peer. This content is downloaded via a different connection. Clients only request content or services from the server in a client-server system. Peers offer all resources in peer-to-peer systems, referred to as either clients or servers (first syllable of the term server and the second of the term client). Someone employed a centralized method in the first generation of P2P systems. The server is still available, but unlike the client-server model, it only saves IP addresses of peers that have content available, minimizing server burden (Napster is an example).

Gnutella 0.4 and Freenet were decentralized alternatives to the centralized method described above. These techniques rely on flooding the network with the desired content identification. Those who share content will respond. Flooding the request generates a lot of traffic. To avoid this, Gnutella 0.6 provides a super-peer hierarchy that stores the content available to connected peers along with their IP addresses. These Super-peers' main goal is to eliminate search hops, therefore minimizing network traffic. The above schemes are unstructured peer-to-peer because the content stored on a given node and its IP address are unrelated and do not follow any structure. Examples of unstructured peer-to-peer systems are Napster, Gnutella, FastTrack, eDonkey, and Freenet.

Structured Peer-To-Peer Systems

The difficulty of developing scaled unstructured Peer-to-Peer applications attracted researchers. Because of the benefits and potential of decentralized self-organizing systems, researchers have concentrated on Distributed Hash Tables (DHTs) (DHT). Its main features are distributed indexing, scalability, and fault tolerance. Using DHT, a data item from the network can be retrieved in O (logN). Unlike unstructured peer-to-peer systems, which normally have linear search complexity, the underlying network and peer count can grow without affecting the distributed application's efficiency. Four of the most interesting and representative mechanisms for routing messages and locating data for structured content distribution systems are:

- **Freenet** is a loosely structured system that uses file and node identifiers to produce an estimate of where a file may be located, and a chain mode propagation approach to forward queries from node to node.
- **A chord** is a system whose nodes maintain a distributed routing table in the form of an identifier circle on which all nodes are mapped and an associated finger table is built.
- **CAN** is a system using n-dimensional cartesian coordinate space to implement the distributed location and routing table, each node is responsible for a zone in the coordinate space.
- **Tapestry** (and Pastry and Kadmelia) are based on Plaxton mesh data structure, which maintains pointers to nodes in the network whose IDs match the elements of a tree-like structure or ID prefixes up to a digit position.

Self Organization

For example, self-organization can be defined as autonomous self-maintenance and optimization.

Definitions

- **System:** A system is a set of components that have relations between each other and form a unified whole. A system distinguishes itself from its environment.
- **Complexity:** This term is used to denote the existence of system properties that make it difficult to describe the semantics of a system's overall behavior in an arbitrary language, even if complete information about its components and interaction is known.
- **Feedback:** The return to the input of a part of the output of a machine, system, or process (as for producing changes in an electronic circuit that improve performance or in an automatic control device that provide self-corrective action).
- **Emergence:** Refers to unexpected global system properties, not present in any of the individual subsystems, that emerge from component interactions.
- **Complex Systems:** Complex systems are systems with multiple interacting components whose behavior cannot simply be inferred from the behavior of the components.
- **Criticality:** An assembly in which a chain reaction is possible is called critical, and is said to have obtained criticality.
- **Hierarchy:** In this context, hierarchy is defined as a rooted tree.
- **Heterarchy:** A heterarchy is a type of network structure that allows a high degree of connectivity. By contrast, in a hierarchy, every node is connected to at most one parent node and zero or more child nodes. In heterarchy, however, a node can be connected to any of the surrounding nodes.
- **Stigmergy:** Stigmergy defines a paradigm of indirect and asynchronous communication mediated by an environment.
- **Perturbation:** A perturbation is a disturbance that causes an act of compensation, whereby the disturbance may be experienced positively or negatively.

Characteristics of self-organization

Based on the above definitions, the self-organization of systems could be characterized as follow:

Self-determined Boundaries:

The border between system and environment is defined by the system itself.

Independence of identity and structure:

The distinction between identity and structure allows for explaining flexibility and adaptability.

Maintenance:

A self-organizing system must try to maintain itself. Feedback and heterarchy: If a system is perturbed,

it tries to restructure to maintain itself, so it needs cross-linked relations with its neighborhood.

Self-determined reaction to perturbation:

A self-organizing system reacts when a perturbation occurs, but it needs metrics for detecting and evaluating the perturbation.

For example, boundaries, reproduction, mutability, organization, metrics, and adaptivity are all self-organizing system features that can be extended to P2P systems. Aside from the degree of conformity to these requirements, every system has an identity or principal purpose that defines it. A P2P system's identity is imposed by outsiders (developers) and not self-determined.

Information

This section is explained how P2P networks are deployed in areas of information.

Presence Information: Presence information is vital in P2P applications. It tells you about available peers and resources. This is important for system self-organization. Information is also used to share CPU cycles because the system knows which processors are overworked. Peers act as information brokers for their peers.

Document management: Document management systems are usually centrally organized to share data storage, management, and use. Creating a central index of important papers takes time and effort. Documents developed in a firm are dispersed among desktop PCs without a central repository knowing about them. P2P networks come in handy here.

Collaboration: P2P allows for closed working group document management.

Files

A characteristic of file-sharing is that sometimes peers ares clients when they download files and sometimes servers when they upload files (*sevents*). A central problem in P2P systems is the searching for the contents or files required (lookup problem). In the context of file-sharing, three different models have been developed: the flooding request model

(Gnutella), the centralized directory model (Napster) and document routing model (Freenet).

Bandwidth

The traffic on networks is constantly rising, mainly in a large volume of multimedia data, file-sharing, so the effective use of the bandwidth has suffer an important increment. When data are centralized and a spontaneous increment of demand arises, the bandwidth becomes a bottleneck. The P2P approach increases load balancing without any kind of additional administration, by taking advantage of transmission routes that are not fully exploited. This concept is applied in the areas of streaming. A shared use of the bandwidth is also very well exploited by splitting big files into smaller blocks that are downloaded by the requesting peers, BitTorrent is an implementation using this principle.

Storage Space

With P2P storage networks, only a portion of the disk space available on desktop PC will be used. A P2P storage network is a cluster of computers, based on existing networks, which share all the storage available

in the network. Examples are PAST, Pasta, CFS, Oceanstore, Farsite, and Intermemory.

Processor Cycles

There are requirements for high-performance computing, at the same time there is computing power unused, this is an incentive for using P2P applications to bundle that computer power. In this way, it is possible to achieve computing power cheaper than a supercomputer can provide. This is effected by forming a cluster of independent, networked computers, in which a single computer is transparent and all the networked nodes are merged into a single logical computer. An example is SETI@home.

Applications Based On Peer-To-Peer

Some applications based on P2P follows:

- **Application-Layer Multicast** In the early days the size of the Internet, certainly limited, permitted broadcasting a single packet to every possible node. On the present Internet, this technique of broadcasting is very expensive. Now is necessary for a selective broadcast, such as multicast. In this field P2P technology has helped, in its unstructured networks, to reach *unlimited* scalability.

- **GRID Computing** The basic objective of GRID computing is to support resource sharing among individuals and institutions (organizational units), or resource entities within a networked infrastructure. Grids are structured and have standards, but not the capacity for self-organizing, fault tolerance, and scalability. On the other hand, P2P systems are self-organizing, have fault tolerance, and react very well to transient populations of peers but lack standards. All the efforts of researching in these fields are to merge the best of the two worlds. Indeed the question of how the two concepts converge is still open.

Summary: The Present And The Future

There was a lot of work done and there is a lot of work to do in the field. It is possible to classify and summarize all the activities in applications and research, present and future.

Applications

- *The Present*

From 2004 up today

- Support for different communications forms
- Telephony.
- Streaming
- Scalable and flexible naming systems.
- Personal communications (e.g.e-mail).
- Inter-organization resource sharing.
- Context/content-aware routing.

- *The Future*

Challenges in the future of applications

- Video conference.
- Distribution of learning material.
- Location-based services in Mobile Ad Hoc Networks (MANET), distributed and centralized.
- Context-aware service.
- Trustworthy computing.

Drawbacks

Reasons against peer-to-peer.

The Present

Up today.

- Lawsuits against users.
- Software patents.
- Intellectual properties.
- P2P requires *flat rates* access.
- Still low bandwidth end nodes.
- Digital right management.
- Best effort service is insufficient for most applications.

The Future

- Lack of trust.
- Commercialization as the end of P2P.
- P2P is integrated into other topics.

Research Focus

What are the present research efforts and what is the research work to do?

Nowadays

Points of research.

- Semantics integration of different information types in the specific peer database.

- Quality of services criteria (consistency, availability, security, reliability).
- Legacy support in overlays.
- P2P and non-request reply interactions.
- Highly adaptive DHTs.
- Overlay optimization.
- P2P signaling efficiency.
- Data dissemination.
- Resource allocation (mechanism and protocols) and guaranteeing the quality of services in
- P2P systems.
- Self-determination of information source.
- Accounting incentive.
- Realistic P2P simulator.
- Decentralize reputation mechanism.
- Semantics queries.
- Efficient P2P content distribution.
- Content-based search queries, metadata.
- Reduction of signaling traffic.
- Data-centric P2P algorithm.
- Content management.
- Application/data integration.
- Security trust, authentication transmission.
- Incentive market mechanism.
- Reliable messaging.
- P2P in mobile cellular/ad-hoc.

Future Challenges

- Anonymous but still secure e-commerce.

- Interoperability and/vs standards.
- ReaP2P for business information systems.
- Real-time P2P data dissemination.
- P2P file systems.
- Concept of trust and dynamic security.
- Dynamic content update.
- Distributed search mechanism.
- P2P technologies in MANET.
- Mobile P2P.
- Intelligent search.
- Service differentiation.
- P2P-GRID integration.

Certainly, there is a lot of work to do, this paper has no conclusions (nothing is over) because all is just beginning. The fields of applications are huge. There are excellent readings that should be used for researching and teaching.

References

https://www.seminarstopics.com/

◄ Preliminary Activity for Week 13 Jump to... Analysis, Application, and Exploration for Week 13 ►

Navigation

Home
Dashboard
Site pages
My courses
Capstone Project 2
Network Defense and Remote Access Configuration

OJT/Practicum 2

Seminars and Tours

Participants

General

12 - Midterm Examination

13 [Enter Module Title Here]

Preliminary Activity for Week 13

Lesson Proper for Week 13

Analysis, Application, and Exploration for Week 13

Generalization for Week 13

Evaluation for Week 13

Assignment for Week 13

Courses

Fair Warning

NOTICE: Please be reminded that it has come to the attention of the Publishing Team of eLearning Commons that learning materials published and intended for ***free use only by students and faculty members within the eLearning Commons network were UNLAWFULLY uploaded in other sites without due and proper permission***.

PROSECUTION: Under Philippine law (Republic Act No. 8293), copyright infringement is punishable by the following: Imprisonment of between 1 to 3 years and a fine of between 50,000 to 150,000 pesos for the first offense. Imprisonment of 3 years and 1 day to six years plus a fine of between 150,000 to 500,000 pesos for the second offense.

COURSE OF ACTION: Whoever has maliciously uploaded these concerned materials are hereby given an ultimatum to take it down within 24-hours. Beyond the 24-hour grace period, our Legal Department shall initiate the proceedings in coordination with the National Bureau of Investigation for IP Address tracking, account owner identification, and filing of cases for prosecution.

Activities

Assignments
Forums
Quizzes
Resources