DATA DISSEMINATION AND MANAGEMENT

Data dissemination on the internet is possible through many different kinds of communications protocols. The internet protocols are the most popular non-proprietary open system protocol suite in the world today. They are used in data dissemination through various communication infrastructures across any set of interconnected networks. Despite the name internet protocol, they are also well suited for local area networks (LAN) and wide area network (WAN) communication.

Using the internet, there are several ways data can be disseminated. The World Wide Web is an interlinked system where documents, images and other multimedia content can be accessed via the internet using web browsers. It uses a markup language called hyper text markup language (HMTL) to format disparate data into the web browser.

The Email (electronic mail) is also one of the most widely used systems for data dissemination using the internet and electronic medium to store and forward messages. The email is based on the Simple Mail Transfer Protocol (SMTP) and can also be used by companies within an intranet system so that staff could communicate with other.

The more traditional ways for data dissemination which are still in wide use today are the telephone systems which include fax systems as well. They provide fast and efficient ways to communicate in real time. Some telephone systems have been simulated in internet applications by using the voice over internet protocol (VoIP).

Through this protocol, hundreds of free or minimally charge international phone calls are already available. This simulated phone call is possible using the computer with microphone and speaker system or headphones. When a video camera is used, it could be possible to have video conferencing.

Of course, the use of non digital materials for data dissemination can never be totally eliminated despite the meteoric rise of electronic communication media. Paper memos are still widely used to disseminate data. The newspaper is still in wide circulation to communicate vital everyday information in news and feature items.

Despite the efficiency of electronic means of data dissemination, there are still drawbacks which may take a long time to overcome, if at all. Privacy is one of the most common problems with electronic data dissemination. The internet has thousands of loop holes where people can peep into the private lives of other people. Security is also a related problem with electronic data

dissemination. Every year, millions of dollars are lost to electronic theft and fraud. Every time a solution is found for a security problem, another malicious programs spring up somewhere in the globe.

Many companies set up precautionary measures against security invasion in their information systems. Some set up user accounts with varying privileges to data access. Many set up internet firewalls and antivirus software on their computers to prevent intrusions.

Data dissemination is a very substantial aspect of business operation. Most of today's businesses are data driven. It is a common scenario where business organizations invest millions for data warehouses including hardware, software and manpower costs, to make data dissemination fast, accurate and timely. Information gathered from disseminated data form as basis for spotting industry trends and patterns and decision making in companies.

Data dissemination in asymmetrical communication environment, where the downlink communication capacity is much greater than the uplink communication capacity, is best suited for mobile environment. In this architecture there will be a stationary server continuously broadcasting different data items over the air. The mobile clients continuously listen to the channel and access the data of their interest whenever it appears on the channel and download the same. The typical applications of such architecture are stock market information, weather information, traffic information etc. The important issue that is to be addressed in this type of data dissemination is – how quickly the mobile clients access the data item of their interest i.e. minimum access time so that the mobile clients save the precious battery power while they are on mobile. There are two fundamental information delivery methods for wireless data applications: Point-to-Point access and Broadcast. Compared with Point-to-Point access, broadcast is a more attractive method. A single broadcast of a data item can satisfy all the outstanding requests for that item simultaneously. As such, broadcast can scale up to an arbitrary number of users.

There are three kinds of broadcast models, namely push-based broadcast, On-demand (or pull-based) broadcast, and hybrid broadcast. In push based broadcast, the server disseminates information using a periodic/aperiodic broadcast program (generally without any intervention of clients). In on demand broadcast, the server disseminates information based on the outstanding requests submitted by clients; In hybrid broadcast, push based broadcast and on demand data deliveries are combined to complement each other. Consequently, there are three kinds of data scheduling methods (i.e., push based scheduling, on demand scheduling, and hybrid scheduling) corresponding to these three data broadcast models.

1. Push based data scheduling

In push based data broadcast, the server broadcasts data proactively to all clients according to the broadcast program generated by the data scheduling algorithm. The broadcast program essentially determines the order and frequencies that the data items are broadcast in. The scheduling algorithm may make use of precompiled access profiles in determining the broadcast program. In the following, two typical methods for push based data scheduling are described, namely flat broadcast and broadcast disks. Flat Broadcast The simplest scheme for data

scheduling is flat broadcast. With a flat broadcast program, all data items are broadcast in a round robin manner. The access time for every data item is the same, i.e., half of the broadcast cycle. This scheme is simple, but its performance is poor in terms of average access time when data access probabilities are skewed. Broadcast Disks Hierarchical dissemination architecture, called Broadcast Disk (Bdisk), was introduced in. Data items are assigned to different logical disks so that data items in the same range of access probabilities are grouped on the same disk.
Data items are then selected from the disks for broadcast according to the relative broadcast frequencies assigned to the disks.

This is achieved by further dividing each disk into smaller, equal size units called chunks, broadcasting a chunk from each disk each time, and cycling through all the chunks sequentially over all the disks. A minor cycle is defined as a sub cycle consisting of one chunk from each disk. Consequently, data items in a minor cycle are repeated only once. The number of minor cycles in a broadcast cycle equals the Least Common Multiple (LCM) of the relative broadcast frequencies of the disks. Conceptually, the disks can be conceived as real physical disks spinning at different speeds, with the faster disks placing more instances of their data items on the broadcast channel. However, if the number of minor cycles in a broadcast cycle is not equal the Least Common Multiple (LCM) of the relative broadcast frequencies of the disks, dividing precisely the desired number of chunks, is a problem. addressed this problem by suggesting to fill up the disk with other information and making it divisible so that the number of minor cycles is equal to the LCM of relative broadcast frequencies.

2. On-demand data scheduling

Push based wireless data broadcasts are not tailored to a particular user's needs but rather satisfy the needs of the majority. Further, push-based broadcasts are not scalable to a large database size and react slowly to workload changes. To alleviate these problems, many recent research studies on wireless data dissemination have proposed using on-demand data broadcast. A wireless on demand broadcast system supports both broadcast and on demand services through a broadcast channel and a low bandwidth uplink channel. The uplink channel can be a wired or a wireless link. When a client needs a data item, it sends to the server an on demand request for the item through the uplink. Client requests are queued up (if necessary) at the server upon arrival.
The server repeatedly chooses an item from among the outstanding requests, broadcasts it over the broadcast channel, and removes the associated request(s) from the queue. The clients monitor the broadcast channel and retrieve the item(s) they require. The data-scheduling algorithm in on demand broadcast determines which request to service from its queue of waiting requests at every broadcast instance.

3. Hybrid data scheduling

Push-based data broadcast cannot adapt well to a large database and a dynamic environment. On-demand data broadcast can overcome these problems. However, it has two main disadvantages: i) more uplink messages are issued by mobile clients, thereby adding demand on the scarce uplink bandwidth and consuming more battery power on mobile clients

ii) if the uplink channel is congested, the access latency will become extremely high. A promising approach, called hybrid broadcast, is to combine push-based and on-demand techniques so that they can complement each other.

In the design of a hybrid system, one of the main issues is the assignment of a data item to push-based broadcast, on-demand broadcast or both. Concerning this issue, there are different proposals for hybrid broadcast in the literature. In the following, we introduce the techniques for balancing push and pull and adaptive hybrid broadcast. Balancing Push and Pull:

Hybrid architecture was first investigated in. In that model, items are classified as either frequently requested (frequest) or infrequently requested (irequest). It is assumed that clients know which items are frequests and which are irequests. The model services frequent using a broadcast cycle, and irequests using on-demand. In the downlink scheduling, the server makes consecutive transmissions of frequented items (according to a broadcast program), followed by the transmission of the first item in the irequest queue (if at least one such request is waiting).

Data allocation over multiple broadcast channels

Multiple physical channels have capabilities and applications that cannot be mapped on to single channels. As stated in some example advantages include better fault tolerance, configurability and scalability. By having access to multiple physical channels fault tolerance is improved. For example if a server broadcasting on a certain frequency crashes, its work must be migrated to another server. If this server is already broadcasting on another frequency it can only accept the additional work if it has the ability to access multiple channels. More flexibility is allowed in configuring broadcast servers. Assume that there are two contiguous cells, which contain broadcast servers that transmit at different channels. A single server that wishes to take over the responsibility of transmitting in both cells can only do so if it can transmit over multiple channels. Finally, being able to transmit over multiple channels has scalability benefits. A broadcasting system must be able to handle both high powered and low powered clients. In order to do so, multiple channels can be used and clients can monitor a number of channels commensurate to their capacities and data needs.