# GNITED MINDS
## Journals

# HANDLING INFORMATION OVERLOAD ON USENET

# Handling Information Overload on Usenet

**Kalpik Patel[1] Prof. Dr. Dhaval Kathiriya[2]**

[1]Research Scholar, Mewar University, Chittorgarh, Rajasthan, India

[2](Director) - Information Technology (Professor) - Computer Science, Anand Agricultural University

***Abstract** – **Usenet is the name of a worldwide network of servers for group communication between people. From 1979 and onwards, it has seen a near exponential growth in the amount of data transported, which has been a strain on bandwidth and storage. There has been a wide range of academic research with focus on the WWW, but Usenet has been neglected. Instead, Usenet's evolution has been dominated by practical solutions*

*Keyword: Communication, Information Overload.*

--------------------------◆---------------------------

## 1. INTRODUCTION

Usenet is the name of a worldwide network of servers for group communication between people. Since Usenet was created in 1979, it has seen an impressive growth from a small academic community to a network used by millions of people from a wide variety of backgrounds all over the world. The total size of the data flowing through Usenet has been more than tripling every year between 1993 and 2001.

This growth has not been without problems, and has raised significant challenges in how to handle the ever increasing volume of Usenet data flow. Very few are able to handle all of Usenet, and as the amount of users and data they produce increase, as do the challenges with having enough network bandwidth and storage capacity. Spending great sums of money on hardware components relieves the situation, but it does not solve it.

My motivation for this thesis was to find a way to reduce the problems we see today. I have introduced the idea of advanced caching methods as a general improvement for parts of the Usenet distribution network, as well as discussed other work that has been done to relieve network bandwidth and storage capacity. I also introduce methods for analyzing and evaluating caching strategies based on statistical data from news servers. Advanced caching will be an improvement for those news servers with users that do not read every available news article, which goes for most if not all news servers with users. However, caching does not solve the problem of exponential growth. When the available technology no longer can support enough network bandwidth and storage capacity, this will limit itself.

Usenet is the set of people who exchange articles tagged with one or more universally-recognized labels, called "newsgroups" (or "groups" for short). There is often confusion about the precise set of newsgroups that constitute Usenet; one commonly accepted definition is that it consists of newsgroups listed in the periodic "List of Active Newsgroups" postings which appear regularly in news, Lists, Misc., and other newsgroups. A broader definition of Usenet would include the newsgroups listed in the article "Alternative Newsgroup Hierarchies" (frequently posted to news, Lists, Misc.,). An even broader definition includes even newsgroups that are restricted to specific geographic regions or organizations. Each Usenet site makes its own decisions about the set of groups available to its users; this set differs from site to site.

Communication is in the form of articles. An article is simply a single text message, authored by one or more entities, usually humans. Some mailing lists are automatically reposted to Usenet. Other automated postings are common.

In the case of human users, the author writes the article in his favorite text editor, and then sends it to a newsgroup (also simply called a group) on a news server (called the injecting server) using a specialized program for news, a newsreader. This process is called posting an article.

The text editor can be internal to the newsreader or vice versa, for instance MS Outlook Express and Gnus respectively. The newsreader program is also referred to as a user agent (UA) for consistency with similar e-mail and WWW terminology). To control the creation, modification, and removal of newsgroups, and to give users a chance of withdrawing articles after they have been posted, there is a special kind of

articles known as control messages. Using a very basic authentication scheme through the **Control** header, one can send newsgroup, check groups, regroups, cancel and supersedes messages, which the news servers of the world may choose to honor or ignore, depending on the administrator's policy.

## 2.    REVIEW OF LITERATURES:

The Usenet News model has the following major aspects to consider: Message format Message distribution Message storage. The main flow of Usenet is commonly through the Internet, using the Network News Transfer Protocol (NNTP) [Kantor and Lapsley, 1986], a TCP based protocol for transmission. Most Internet standards are described in RFCs, and the IETF is working on several new standards. Usenet's standards are described in RFCs, but there are de facto Usenet standards not included in the RFCs, although the IETF is working on standardizing these enhancements.

The core of Usenet's organization is eight top level topically organized hierarchies frequently referred to as the Big 8. The alternate hierarchy alt, which is more "free" in how groups are created and organized, is also considered part of the core by many users and news administrators.

Information overload is closely linked to high cognitive load. According to Kelsey and St. Amant, "Cognitive load is the burden on working memory during information processing. It can be augmented by the individual's characteristics, but also by the content of the form and structure of the message. For example, a complex integration of modalities can elicit a cognitive load on the working memory so that the information is less remembered." A decision maker faces a cognitive load more when processing information for declarative knowledge than for procedural knowledge. An engineer or technician, for instance, might convert what he or she reads into an immediate action, requiring relatively little cognitive load. If the same information has to be stored in the brain to use for later application, then the cognitive load will be heavier. Therefore, reliance on memory can lead to the perception of information overload.

It is important for producers of information to keep in mind the time constraints that users face and create communication products that provide the right balance of information to prevent information overload—as well as information underload. At the same time, producers of information also need to package communication products according to the needs of the users, so as to prevent the perception of information overload.

**Article Level Control** The cancel and supersedes control message types are there to control actual articles. Sending a cancel message should — if it is honored— result in the removal of the article you ask to cancel out. A supersedes message has its own header, **Supersedes**, but is effectually just a combined cancellation of an old article, with a new article attached; effectively, it replaces the old one. The header can also be used for repositioning the new article within the thread it belonged to.

A moderated newsgroup is no different from an ordinary newsgroup, with a few exceptions. All articles posted to such a newsgroup are automatically forwarded via e-mail to a moderator. That moderator can either be a program or a person, who checks the articles according to some criteria defined for that particular newsgroup. Articles that fulfill those criteria are posted to the newsgroup through the use of the **Approved** header. Other articles are discarded, and sometimes returned to their originating author.

Decision makers of all kinds are often overwhelmed with information and lack the time management techniques to cope with the problem. The definitions of information overload given earlier provide a good start for studying the subject, but howdo we deal with the elements that make information overload such an international and intercultural challenge for professional communicators in the engineering, scientific, technical, and business fields? To conceptualize the challenge, we consider information overload in terms of information and time management. How can professional communicators assist? In this book, we argue that communicators need to know how their products are processed by their clients to help them minimize and/or avoid information overload. To gain this understanding, communicators must be intimately familiar with the supplier/producer/ writer and client/user/reader perspective. In addition, the intercultural perspective has to be taken into account. Finally, a focus on innovation needs to be maintained to try and manage information overload.

## 3.    THE HISTORY AND DEVELOPMENT OF USENET

In 1999, Usenet News turned 20 years. In those 20 years, many things have changed, but some underlying principles have remained.

When BBSes (Bulletin Board Systems) were very popular, many people expressed that Usenet was just another BBS. Where BBSes (with few exceptions) were limited to single computers and people connected with their modems (or whatever means they had) to post their messages and discuss with others of like or different mind, Usenet was from the beginning a distributed system, where messages were transmitted between different computers to be available from more servers. Usenet was probably best compared with a network of BBSes, each carrying the same discussions.

## 4.    TRAFFIC GROWTH

Usenet traffic — meaning the number and size of the daily accepted flow of articles for a site that attempts

Kalpik Patel[1] Prof. Dr. Dhaval Kathiriya[2]

to get "all" that is posted to Usenet —has never increased slowly, at least not volume measured in bytes. According to Hardy ([Hardy, 1993], citing Gene Spafford), the rate of growth was fairly constant from 1979 to 1988 (see table 1.7 on the facing page, based on a similar table in [Spafford, 1990]).

About 1993, the use of services available via the Internet (such as Usenet) began increasing dramatically with the introduction of the World Wide Web (WWW), and the following success stories of the commercial ISPs.

## CONCLUSION

I have presented the history of Usenet from a growth perspective, and shown that there are technical problems with the its continued growth. Smaller sites cannot afford to offer their users all the newsgroups they might want to read, and the problem seems to be growing. While other solutions than caching — such as filtering — greatly reduce the size of a full newsfeed, they are rigid and do not adapt the incoming flow depending on usage, as caching will. The World Wide Web has used various caching methods for years, and a lot of work and research has been done to optimize caching for the web. However, nobody has worked with solutions for news. My proposed advanced caching methods for Usenet will help the smaller sites to appear to offer a greater amount of newsgroups and articles, but does not address the problem of the seemingly exponential growth. However, even a linear reduction in newsfeed size will buy the news administrators time to postpone the next hardware upgrade, which means they will save money. One small weakness is that I do nothing to help the backbone Usenet sites, which are the ones who carry the bulk of Usenet traffic today.

## REFERENCES:

[Kantor and Lapsley, 1986] Kantor, B. and Lapsley, P. (1986). RFC 977: Network News Transfer Protocol — A Proposed standard for the Stream-Based Transmission of News. RFC.

[Spafford, 1990] Spafford, G. (1990). Re: The List again :-). http://communication.ucsd.edu/bjones/Usenet. Hist/Nethist/0014.html.

[Hardy, 1993] Hardy, H. E. (1993). The Usenet System. *ITCA Yearbook*.

Jan Ingvoldstad "Handling Information Overload on Usenet Advanced Caching Methods for News" [4-aug-2001]

Emil Sit, Robert Morris, and M. Frans Kaashoek UsenetDHT: A low-overhead design for Usenet [MIT CSAIL]

Deepika Saxena, Monika Saxena "Methods Used to Handle Overloading of Information in Usenet" [International Journal of Scientific & Engineering Research, Volume 2, Issue 1,

January-2011 1 ISSN 2229-5518] Emil Sit, Frank Dabek and James Robertson MIT Computer

Science and Artificial Intelligence Laboratory "UsenetDHT: A Low Overhead Usenet Server"

Rich Salz – Open Software Foundation InterNetNews: Usenet transport for Internet sites

[Bumgarner, 1995] Bumgarner, L. S. (1995). USENET — The Great Renaming — 1985–1988. http://www.vrx.net/usenet/history/rename.html.

[Cidera Inc., 2001] Cidera Inc. (2001). Cidera usenet news service. http://www.cidera.com/services/usenet_news/index.shtml.

[Collyer, 1992] Collyer, G. (1992). newsoverview - netnews overview files. newsoverview(5) man page.

[Crocker, 1982] Crocker, D. H. (1982). RFC 822: Standard for the Format of ARPA Internet Text Messages. RFC.

**Kalpik Patel[1] Prof. Dr. Dhaval Kathiriya[2]**