

# A Measurement Study on the BitTorrent File Distribution System

Zhicheng Yang, [yangzhicheng@wustl.edu](mailto:yangzhicheng@wustl.edu) (A paper written under the guidance of [Prof. Raj Jain](#))



## Abstract:

The file distribution protocol BitTorrent(BT) is very popular nowadays, and people are used to sharing files with BT clients everyday. In this project, a two factor full factorial design on workloads and BT clients was proposed to test which effect is significant for BT downloading. The first part presents the background knowledge of Peer-to-peer(P2P) and BT and the introduction of several popular BT clients. Next, a two factor full factorial design was shown. The result indicated that "torrent" was the primary significant factor, while "BT client" was the secondary factor. Also, there was another experiment on out-of-date torrents to test which BT client was successful to download them with few seeder and peers. The result indicated the BT client " $\mu$ Torrent" was strongest in this task.

**Keywords:** Measurement Study, Peer-to-peer, Experimental Design, BitTorrent, BitTorrent Client, Torrent

## Table of Contents

- [1. Introduction](#)
  - [1.1 P2P](#)
  - [1.2 BT](#)
- [2. BT Clients](#)
  - [2.1.  \$\mu\$ Torrent](#)
  - [2.2. BitComet](#)
  - [2.3. Vuze](#)
  - [2.4. FlashGet](#)
  - [2.5. BitLord](#)
- [3. Workloads and Metric](#)
- [4. Experimental Design](#)
  - [4.1. The Two Factor Full Factorial Design](#)
    - [4.1.1. Data Processing](#)
    - [4.1.2. Computation of Effects](#)
    - [4.1.3. Analysis of Variance\(ANOVA\)](#)
    - [4.1.4. Visual Tests](#)
    - [4.1.5. Memory Usage](#)
  - [4.2. Experiment on the Out-of-date Torrents](#)
- [5. Summary](#)
  - [5.1. Significant Effects](#)
  - [5.2. Clients Choices](#)
  - [5.3. Reliability](#)
- [6. Future Works](#)
- [References](#)
- [Acronyms](#)

## 1. Introduction

BT is a very popular file distribution protocol based on P2P technique, and people use it to share files with BT clients everyday. In this section, the background knowledge of P2P and BT are shown.

### 1.1 P2P

P2P [[Kurose09](#), [Deaconescu09](#), [Peer-to-peer wiki](#), [Qiu04](#)] is a technique that means connecting user to user directly without a substantial server. It is different from the traditional client-server model in which clients only receive information sent from server. In the peer-to-peer model, peer is not only the client but also the server. A peer's partial resources, like network bandwidth, disk storage, are available to other peers in the same network directly. Today, many of popular applications are based on P2P architectures, such as BT and eMule for file sharing and Skype for Internet telephone.

## 1.2 BT

In the BT [[Kurose09](#), [BitTorrent wiki](#)] system, there are 3 elements: peer, tracker and seed. A peer is a user who is downloading files; a tracker returns a peer list for every peer, helping peers find each other; a seed has a complete file. The file is divided into many chunks and the information of those file chunks and tracker are saved in a "torrent" file, from which a peer can obtain file and tracker address information. When a peer is downloading the file chunks it is interested in, it is also uploading the chunks it has already downloaded. And a seed keeps uploading all the time. In the initial situation, a BT system need at least one seed. If a peer has finished downloading the file but doesn't leave the system, it will become a seed.

## 2. BT Clients

A BT client is a software which can download files from a torrent based on BT protocol. In this paper, five present popular BT clients were selected. The difference of those clients were regarded as a factor in the following experimental design.

### 2.1 $\mu$ Torrent

$\mu$ Torrent [ [\$\mu\$ Torrent wiki](#)] is a free BT client written in C++, whose platform are Windows and Mac OS X. The " $\mu$ " means "one-millionth" from the International System of Units(SI) prefix "micro-", indicating this program's small memory occupancy. Compared with BitComet and Vuze, it needs less system resources and its original installation file size is 532 KB.  $\mu$ Torrent has a good reputation in terms of feature set, stability and compatibility.  $\mu$ Torrent is also one of the most popular BT client worldwide. In this project, the version of  $\mu$ Torrent is 3.0.

### 2.2 BitComet

BitComet [[BitComet wiki](#), [BitComet](#)] is a BT client written in C++, whose platform is Windows. It can also be used for "Hyper Text Transport Protocol(HTTP)"/"File Transfer Protocol(FTP)" download, and even eMule download if the eMule plugin is installed. BitComet supports 52 different languages. The most attractive feature of BitComet is that an Internet Explorer window is embedded to support users to search torrents. It also supports "Universal Plug and Play(UPnP)" gateway configuration, bandwidth scheduling and so on. The recent version has supported super-seeding [[Super-Seeding wiki](#), [Garbacki07](#)] and magnet links [[Magnet URI scheme wiki](#)]. BitComet provides a "preview download mode", which means that users can preview the content even though the file has not completely finished. This mode helps users to discard torrents in which the files are huge and different from the ones users want to download. BitComet also allows users to create their own torrents. In this project, the version of BitComet is BitComet Beta [20110325].

### 2.3 Vuze

Vuze [[Vuze wiki](#), [Vuze](#)] is a BT client written in Java. The former version is called Azureus. It support Windows, Mac OS X, Linux and Unix. It looks like not only a simple download tool but also a media center, whose interface is similar to that of iTunes. It is said to allow users to share the original DVD and HD quality materials [[Azureus' HD Vids Trump YouTube](#)]. In this project, the version of Vuze is 4.6.0.4.

### 2.4 FlashGet

FlashGet [[Flashget wiki](#), [Flashget](#)] is a free download tool, supporting almost all popular protocols, like HTTP, FTP, RSTP, BT protocols. The original developer is Yantang Hou, a Chinese Canadian. FlashGet is also one of the most popular download softwares in China. Since FlashGet has the English version, so it is selected in this project. In this project, the version of FlashGet is 3.7.

### 2.5 BitLord

BitLord [[BitLord](#)] is a free BT client derived from BitComet. It has some additional features, like chatting and seed searching, which are very convenient for users. The interface is similar to the one of BitComet In this project, the version of BitLord is 1.1.

### 3. Workloads and Metric

In this paper, the workload was torrent. Several torrents with different number of seeders, file size and file type were downloaded on 5 different BT clients. The following table showed the detailed information about them. All the torrents were from "the Pirate Bay"[[the Pirate Bay](#)].

**Table 1: Information of Torrent Samples**

Workload	Name	Number of Seeders	Torrent size	File size	Source URL	Added Date
I	Katy Perry - E.T Acoustic Cover By Christina [Youtube]	10	2.56KB	5.23MB	<a href="http://thepiratebay.org/torrent/6309184/Katy_Perry_-_E.T_Acoustic_Cover_By_Christina_[Youtube]">http://thepiratebay.org/torrent/6309184/Katy_Perry_-_E.T_Acoustic_Cover_By_Christina_[Youtube]</a>	Apr 10 2011
II	Ferrari Power Screensaver	27	8.25KB	23.83MB	<a href="http://thepiratebay.org/torrent/6309062/Ferrari_Power_Screensaver">http://thepiratebay.org/torrent/6309062/Ferrari_Power_Screensaver</a>	Apr 10 2011
III	Rio 2011 Hindi CAMRip x264 Sample Only NiCkkkDoN	30	2.11KB	4.7MB	<a href="http://thepiratebay.org/torrent/6308148/Rio_2011_Hindi_CAMRip_x264_Sample_Only_NiCkkkDoN">http://thepiratebay.org/torrent/6308148/Rio_2011_Hindi_CAMRip_x264_Sample_Only_NiCkkkDoN</a>	Apr 10 2011
IV	Typing Master 7.0	4	2.16KB	4.71MB	<a href="http://thepiratebay.org/torrent/6269409/Typing_Master_7.0">http://thepiratebay.org/torrent/6269409/Typing_Master_7.0</a>	Mar 25 2011
V	Mr. Sparkle Promotional Video ScreenSaver (Simpsons)	2	2.46KB	5.6MB	<a href="http://thepiratebay.org/torrent/6269004/Mr._Sparkle_Promotional_Video_ScreenSaver_(Simpsons)">http://thepiratebay.org/torrent/6269004/Mr._Sparkle_Promotional_Video_ScreenSaver_(Simpsons)</a>	Mar 24 2011

The metric was the download speed rate, which was a HB(High is Better) metric. And the actual extreme HTTP download speed was about 680 KB/s measured by the Internet Explorer downloader. The detailed information of the experiment platform was Table 2.

**Table 2: Information of the Experiment Platform**

Computer Type	Laptop
Computer Brand	ThinkPad X300 (made in August 2008)
CPU	Intel(R) Core(TM)2 Duo CPU L7100
Memory Size	1GB
Disk Storage	80GB
Operation System	Windows XP SP3
Network Bandwidth	AT&T U-verse 6Mbps

### 4. Experimental design

In this section, a two factor full factorial design[[Jain91](#)] was presented. Another experiment focused on which BT client was strongest on downloading from out-of-date torrents.

#### 4.1 The Two Factor Full Factorial Design

A two factor design is used when the change of values of two parameters has impact on the performance. The first factor is workloads and the second factor is BT clients. A full factorial design means that we set the number of experiments based on the levels of the first factor and the second factor. For example, if Factor A has  $a$  levels and Factor B has  $b$  levels, then the number of experiments are  $a*b$ .

##### 4.1.1 Data Processing

First, the first three torrents were used in the two factor full factorial design. The first factor was BT client; there were 5 different BT clients. The second factor was torrent, including a 10-seeder music torrent, a 27-seeder software torrent and a 30-seeder video torrent. The data was the download rate observed only once and the unit was KB/s. Table 3 was the data.

**Table 3: Download Rates for the Different Torrents and Different BT Clients Experiment**

Clients	I	II	III
$\mu$ Torrent	31	394	289
BitComet	10	186	120
Vuze	12	186	115
FlashGet	27	400	320
BitLord	18	368	317

We noticed that the range of data was  $400-10=390$ , this value might be large, so we took the logarithmic method to reduce the range. Table 4 was the new data.

**Table 4: Logarithmic Download Rate for the Different Torrents and Different BT Clients Experiment**

Clients	I	II	III
$\mu$ Torrent	1.4914	2.5955	2.4609
BitComet	1.0000	2.2695	2.0792
Vuze	1.0792	2.2695	2.0607
FlashGet	1.4314	2.6021	2.5051
BitLord	1.2553	2.5658	2.5011

#### 4.1.2 Computation of Effects

**Table 5: Computation of Effects for the Different Torrents and Different BT Clients Experiment**

Clients	I	II	III	Row Sum	Row Mean	Row Effect
$\mu$ Torrent	1.4914	2.5955	2.4609	6.5478	2.1826	0.1715
BitComet	1.0000	2.2695	2.0792	5.3487	1.7829	-0.2282
Vuze	1.0792	2.2695	2.0607	5.4094	1.8031	-0.2080
FlashGet	1.4314	2.6021	2.5051	6.5386	2.1795	0.1684
BitLord	1.2553	2.5658	2.5011	6.3222	2.1074	0.0963
Column Sum	6.2572	12.3024	11.6070	30.1666		
Column Mean	1.2514	2.4605	2.3214		2.0111	
Column Effect	-0.7597	0.4494	0.3103			

The analysis of computation of effects was shown in Table 5. For each row and each column, we computed the sum and the mean of observations in that row or that column, overall sum and overall mean. The difference between a row or a column mean and overall mean indicated the row's or the column's effect.

Table 5 showed an average workload on an average BT client required  $10^{2.0111}=129.12$  KB/s to be finished. The average speed of downloading Workload I was  $10^{2.0111-0.7597}=17.84$  KB/s, and so on. The average speed of BT client  $\mu$ Torrent of downloading the three workloads was  $10^{2.0111+0.1715}=152.26$  KB/s, and so on. From Table 5, both workloads and BT clients had effects on download speed, because their effects were different from the overall mean.

#### 4.1.3 Analysis of Variance(ANOVA)

**Table 6: ANOVA Table for the Different Torrents and Different BT Clients Experiment**

Component	Sum of Squares	%Variation	DF	Mean Square	F-comp.	F-Table
y	65.5621		15			
y...	60.6682		1			
y-y...	4.8939	100.00	14			
Workloads	4.3766	89.43	2	2.1883	580.1890	3.1131
BT Clients	0.4871	9.95	4	0.1218	32.2878	2.8064
e	0.0302	0.62	8	0.0038		

ANOVA table is a way to analyse the significance of factors. Here, we used the ANOVA table for two factors without replications[Jain91]. From Table 6, the percentage of variation explained by the workloads was 89.43%, by the BT clients was 9.95% and by the errors was 0.62%. Table 6 showed the F-computation value 580.1890 of workloads was much more than the F-table value 3.1131 of them and the F-computation value 32.2878 of BT clients was more than the F-table value 2.8064, but not so much as the F-computation value of workloads.

#### 4.1.4 Visual Tests

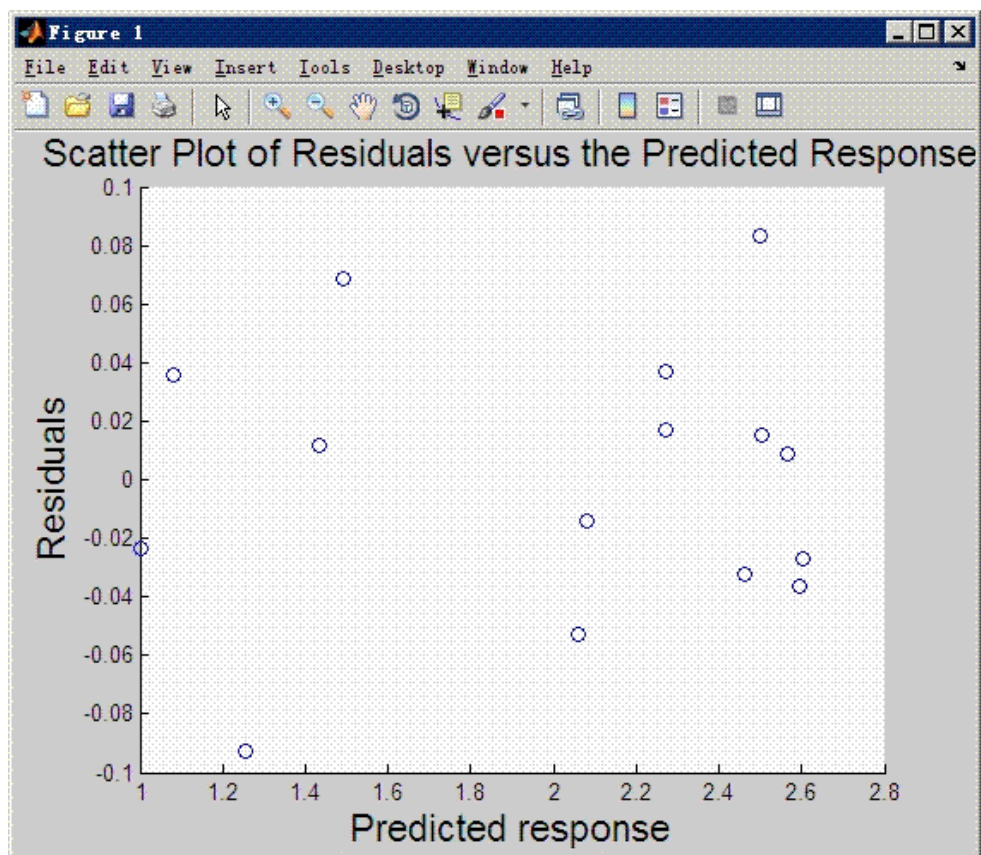
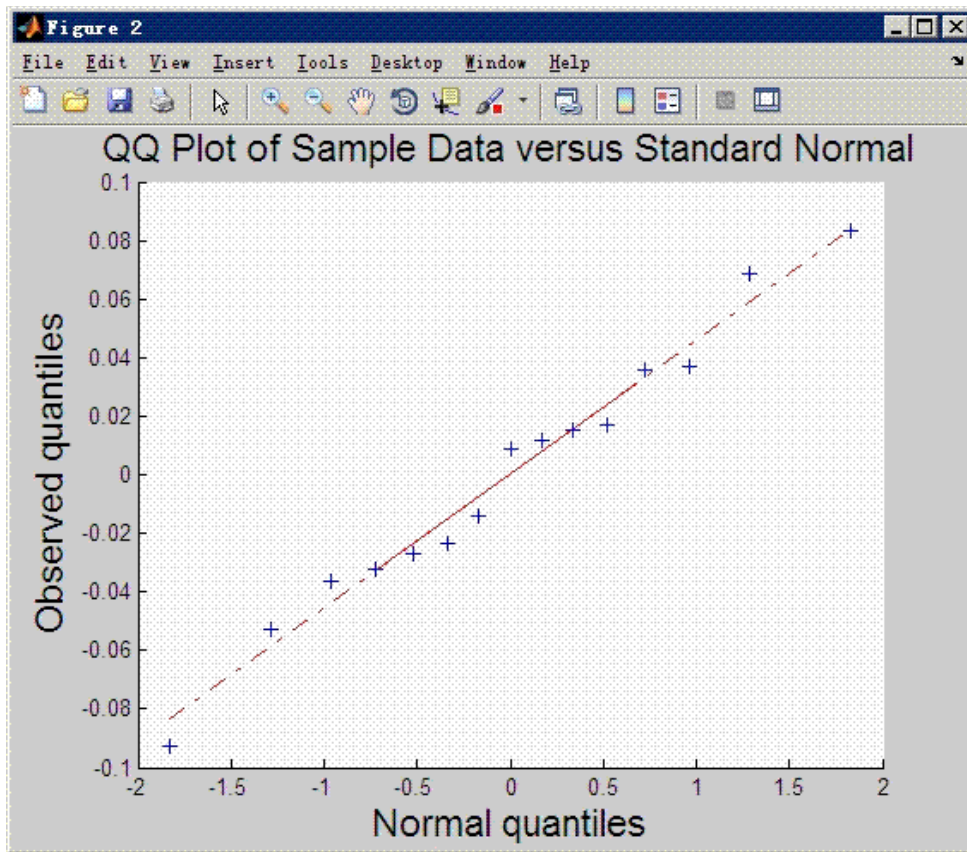


Figure 1: Plot of the residuals versus predicted response for the Different Torrents and Different BT Clients Experiment



**Figure 2: Normal quantile-quantile plot(QQ plot) for the residuals of the Different Torrents and Different BT Clients Experiment**

Visual tests was used to test the errors' distribution. Figure 1 was scatter plot, which is a technique to test if the values are independent or not. If the values have no trend, they are independent. In this paper, Figure 1 showed the errors between observations and the overall mean. There was no trend in this figure, that is, the errors were independent. Figure 2 was the QQ plot, which was a technique to plot the observed quantiles versus the theoretical quantile for small samples. It showed the errors appeared to be normally distributed.

**4.1.5 Memory Usage**

Table 7 was the average of memory usage monitored in the Windows Task Manager. Each value was obtained instantly after finishing downloading tasks in each BT clients. The memory usage of  $\mu$ Torrent was the lowest.

**Table 7: Memory Usage for the Different Torrents and Different BT Clients Experiment**

Clients	Memory Usage
$\mu$ Torrent	13,796 K
BitComet	60,578 K
Vuze	112,308 K
FlashGet	104,128 K
BitLord	37.984 K

**4.2 Experiment on the Out-of-date Torrents**

Second, the last two torrents were used in the second experiment showing which client could download those out-of-date torrents successfully. We noticed that these torrents were added half a month ago. The number of seeders was only 4 and 2. The "NA" was defined that during 30 minutes, the download process bar did not go forward any more. The timer began as soon as the download task was added.

**Table 8: Staring Time and Finishing Time for the Different BT Clients on Out-of-date Torrents Experiment**



Clients	IV		V	
	Starting Time	Finishing Time	Starting Time	Finishing Time
$\mu$ Torrent	1m05s	5m25s	1m42s	7m01s
BitComet	NA	NA	NA	NA
Vuze	3m08s	10m32s	4m43s	NA
FlashGet	36s	6m50s	59s	8m57s
BitLord	NA	NA	NA	NA

## 5. Summary

In this section, we summarized that the significant effects in the first experiment, the clients choices and the reliability of those 5 BT clients.

### 5.1 Significant Effects

From Table 6, because the F-comp values were bigger than F-table values, both of workloads and BT Clients were significant. However, the dominant significant effect was the workloads. It indicated the dominant element influencing the download rate was the torrent selection. From Table 1 and Table 3, in general, the higher the number of seeders was, the more download speed BT clients had. Therefore, for BT downloading, the most important thing was peers should start to download the interesting files from the networks as soon as possible. In the visual test graph, the errors didn't have obvious trend in the scatter figure, and the QQ plot showed errors were normally distributed. Thus this design model was acceptable.

### 5.2 Clients Choices

When we just focused on the comparison of these clients,  $\mu$ Torrent was suggested as the prior choice on BT tasks, based on the application size, speed and memory usage. FlashGet enjoyed a highest download speed in this experiment, but its memory usage was huge. BitLord had its own merit of seeder-rich torrent, though its prototype was BitComet, the performance of the former was much better than that of the latter. Vuze's download speed was just moderate and the memory usage was the highest compared with other BT clients. This issue was common for Vuze users and the official site of Vuze had a method trying to reduce the memory usage [[Reduce Memory Size](#)].

### 5.3 Reliability

From Table 8,  $\mu$ Torrent was the strongest among those BT clients. During the experiment, it used shortest time to finish the two tasks, and FlashGet was the second shortest. Vuze was not so strong as  $\mu$ Torrent but it succeeded to download Workload IV. BitComet and BitLord were disappointing in this experiment. In conclusion,  $\mu$ Torrent was still recommended to be the first choice for downloading old torrents.

## 6. Future Works

Because of the limitation of time to finish huge file downloading on time, the file size in those torrents were not big. As Table 1 shows, the video file was only about 4 MB. That might cause the speed couldn't reach the limited speed, because before clients went higher speed, the task had already been finished. Also, some designs including more factors should be used to test more comprehensive result.

## References

- [Kurose09] James F.Kurose, Keith W.Ross, "Computer Networking," Pearson, 2009, ISBN-10: 0131365487; ISBN-13: 9780131365483.
- [Jain91] Raj Jain, "The Art of Computer Systems Performance Analysis," John Wiley & Sons, INC, 1991, ISBN-10: 0471503363.
- [Deaconescu09] Deaconescu, R., R. Rughinis, N. Tapus. "A BitTorrent Performance Evaluation Framework," In: Proceedings of ICNS 2009.
- [Peer-to-peer wiki] "Peer-to-peer"; <http://en.wikipedia.org/wiki/Peer-to-peer>.
- [BitTorrent wiki] "BitTorrent(protocol)"; [http://en.wikipedia.org/wiki/BitTorrent\\_\(protocol\)](http://en.wikipedia.org/wiki/BitTorrent_(protocol)).

6. [ $\mu$ Torrent wiki] " $\mu$ Torrent"; <http://en.wikipedia.org/wiki/MTorrent>.
7. [BitComet wiki] "BitComet"; <http://en.wikipedia.org/wiki/BitComet>.
8. [BitComet] "BitComet"; <http://www.bitcomet.com/index.php>.
9. [Vuze wiki] "Vuze"; <http://en.wikipedia.org/wiki/Vuze>.
10. [Vuze] "Vuze"; <https://www.vuze.com/>.
11. [FlashGet wiki] "FlashGet"; <http://en.wikipedia.org/wiki/FlashGet>.
12. [FlashGet] "FlashGet"; [http://www.flashget.com/index\\_en.htm](http://www.flashget.com/index_en.htm).
13. [BitLord] "BitLord" [BitLord official site]; <http://www.bitlord.com/>.
14. [Qiu04] Dongyu Qiu, R. Srikant, "Modeling and performance analysis of BitTorrent-like peer-to-peer networks," SIGCOMM, 2004, 34(4).
15. [Super-Seeding wiki] "Super-Seeding"; [http://en.wikipedia.org/wiki/Super\\_seeding](http://en.wikipedia.org/wiki/Super_seeding).
16. [Garbacki07] P. Garbacki, D. H. J. Epema, and M. van Steen. "Optimizing peer relationships in a super-peer network." In ICDCS, page 31, 2007. 6.
17. [Magnet URI scheme wiki] "Magnet URI scheme"; [http://en.wikipedia.org/wiki/Super\\_seeding](http://en.wikipedia.org/wiki/Super_seeding).
18. [Azureus' HD Vids Trump YouTube] "Azureus' HD Vids Trump YouTube"; <http://www.wired.com/software/coolapps/news/2006/12/72223>.
19. [the Pirate Bay] "the Pirate Bay"; <http://thepiratebay.org>.
20. [Reduce Memory Size] "Reduce Memory Size"; [http://wiki.vuze.com/w/Reduce\\_memory\\_usage](http://wiki.vuze.com/w/Reduce_memory_usage).

## Acronyms

ANOVA = Analysis of Variance

BT = BitTorrent

DF = Degree of Freedom

FTP = File Transfer Protocol

HTTP = Hyper Text Transport Protocol

P2P = Peer-to-peer

QQ plot = Quantile-Quantile plot

SI = International System of Units

UPnP = Universal Plug and Play

URL = Uniform Resource Locator

---

Last modified on April 24, 2011

This and other papers on latest advances in performance analysis are available on line at <http://www1.cse.wustl.edu/~jain/cse567-11/index.html>

[Back to Raj Jain's Home Page](#)