

Workload Analysis of SPECmail2009

Hitoshi Oi and Sho Niboshi

Computer Architecture and Operating Systems Group,

The University of Aizu

Aizu-Wakamatsu, JAPAN

Email: {oi,nibo}@oslab.biz

Abstract—SPECmail2009 is a new benchmark suite for measuring the performance of corporate mail servers. It assumes the IMAP4 protocol for retrieving messages and a large fraction of transactions are for manipulating hierarchical messages and mail folders from the users connected via high-speed networks. In this paper, we present a performance analysis of SPECmail2009, including disk I/O behavior, quality of service metrics, and CPU time breakdown. The IMAP4's capability of directly manipulating the messages on the server makes SPECmail2009 a highly I/O-bound workload. Our experiments show that a combination of a small but fast access SCSI drive for indexing messages and larger but slower SATA drives for message files can be an option for building a cost-effective mail storage. Both of multiprocessing and multithreading are quite effective for SPECmail2009 especially in the user mode execution.

Keywords: Mail server, performance analysis, disk I/O.

I. INTRODUCTION

Email has been the most basic and crucial service on the communication networks. It was even available in the era of the time-sharing systems, on which users were connected to the host computers using the serial lines (such as RS-232C) with dumb terminals. Naturally, the data transfer rates were low and messages were in the plain-text format.

In late 90s, the Internet access became available from home through the Internet service providers (ISPs). People first connected to the Internet over the telephone lines, which limited the bandwidth to tens of Kbps. Later, so-called broadband technologies, such as accessing the Internet over the cable-television or ADSL, became available and they drastically increased the communication bandwidth. During this period, people started using attachment, the mail messages accompanied with picture, audio and movie files. The users downloaded their messages from ISP's servers to their client machines (PCs or mobile devices) using the POP3 protocol [1]. SPECmail2001 is the benchmark program from the Standard Performance Evaluation Corporation (SPEC) evaluating the ISP's mail servers used in this manner [2].

While the mail servers modeled in SPECmail2001 are still common these days, users in the enterprise environment are accessing messages using a different protocol, IMAP4 [3]. Unlike POP3, users can directly manipulate the messages on the server with IMAP4, such as browsing or deleting, in folders that have hierarchical structures. This capability of direct message manipulation, together with a higher message frequency and larger message sizes of the enterprise environment, places different workload on the mail servers than

those modeled in SPECmail2001. SPECmail2009 is a new benchmark suite from SPEC for measuring the performance of such corporate mail servers based on the IMAP4 protocol [4]. In this paper, we present a case study of SPECmail2009 performance analysis, including disk I/O behavior, quality of service metrics, and CPU time breakdown.

This paper is organized as follows. In the next section, the workload design and performance metrics of SPECmail2009 are explained. In Section III, the analysis of performance measurements, in terms of disk I/O behavior and CPU time breakdown are presented. Section IV introduces related work and the paper is concluded in Section V.

II. SPECMAIL2009

SPECmail2009 is designed to evaluate the performance of a mail server used in an enterprise environment. The primary scaling factor is the number of provisioned users (SPECmail_Ent2009 users or SPECmail_Ent2009Secure users), and the size of mail store and transaction rates are proportional to this value. The data collected at three companies, two universities and an out-sourced mail server are analyzed by the SPEC subcommittee and the following points are taken into consideration to make the workload realistic (summarized from Section 2.1 of [5]): **Mail store:** the numbers of folders and messages in folders. **Mail Store Folder Structures:** the numbers of hierarchical levels and sub-folders. **Message MIME Structures:** the numbers of attachments and their sizes. **Arrival Rate:** the number of requests received by the server per hour. **Peak Hour Simulation:** the workload reflects the operation in the busiest hour of the day. **Realistic Operation:** such as logging and exclusion of benchmark-specific optimization.

The workload of SPECmail2009 consists of SMTP and IMAP transactions. The SMTP transaction rate is relatively simple. It is determined by the number of active users (78% of provisioned users), the average number of messages received by an active user during peak hour (5.06), and the average number of recipients per message (3.06). The distribution of message flow is local to local (56%), remote to local (31%) and local to remote (13%). The IMAP4 transaction model is defined as follows. First, five command sequences, each of which is a series of IMAP4 commands, are defined. Next, based on the command sequences issued, four client types are defined. Each client issues a command with various parameters. The combinations of commands and parameters

together with the transition of each client state are represented as a finite state machine with probabilities of transitions.

SPECmail2009 defines two performance metrics: SPECmail_Ent2009 and SPECmail_Ent2009Secure. These metrics represent the number of IMAP4 sessions per hour. The only difference between these two metrics is that in the system for the latter metric, secured transport channels are used in IMAP4 and SMTP connections. At the time of writing (April 2010), however, no SPECmail_Ent2009Secure result has been published by SPEC. For the validation of benchmark results, various Quality-of-Service (QoS) criteria are defined: 95% of response time of two SMTP and 13 IMAP4 transactions must be within 5 seconds. Also, 95% of messages to local users and remote users must be delivered within 60 seconds and within the measurement period, respectively. Finally, overall error rate must be less than 1%.

III. PERFORMANCE ANALYSIS

In this section, we show the results of measurements and their analysis. First the experimental environment is described. Next, the behaviors of disk I/O and QoS metrics against varying scaling factor of SPECmail2009 are analyzed. In Section III-D, CPU time breakdown and the effectiveness of multithreading and multiprocessing are presented.

A. Experimental Environment

Table I shows the specifications of the hardware and software components used in the experiments. The mail server has two Xeon CPUs with Hyper-Threading [7]. Until Section III-D, we use both CPUs with Hyper-Threading disabled. For each number of users, we created a mail storage and validated its structure using the benchmark manager included in the SPECmail2009 kit. The same mail storage is used for all the execution of the benchmark for the same number of users. We run the benchmark at least five times for each configuration and take the average among them. While we have followed the rules defined in [6], the results presented in this paper are not validated by SPEC. Therefore, our results are considered to be “Research and Academic Usage of SPECmail2009” defined in Section 3.2.2 of [6], and must not be compared with the results officially published from SPEC.

Component	Specification
Processor	Xeon 3GHz × 2
L1 Data Cache	8KB
L2 Cache	2MB
Memory	2GB
SATA HDD	7.2Krpm, 8MB disk cache × 2
SCSI HDD	15Krpm, 16MB disk cache, Ultra-320 SCSI
Operating System	Linux (CentOS) Kernel 2.6.18
IMAP4 Server	Dovecot (version 4.69)
SMTP Server	Exim (version 1.0.7)

TABLE I
BENCHMARKING ENVIRONMENTS

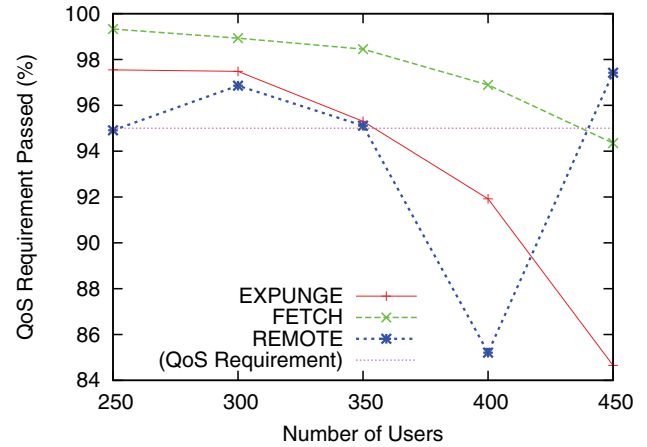


Fig. 1. Quality of Service Metrics (2 SATA Disks).

B. Scalability

First, we use only two SATA disks for the mail storage and change the number of users from 250, which is the minimum number of users defined in SPECmail2009, until the system saturates. Fig. 1 shows the changes of three QoS metrics, IMAP EXPUNGE (EXPUNGE), IMAP UID_FETCH_RANGE_UID_BODYPEEK_HEADERFIELDS_FLAGS_RFC822SIZE (FETCH) and Remote Message Delivery (REMOTE), which we have found the most sensitive to the system scale-up in our experiments. The first two QoS metrics show the percentage of the corresponding transactions that finish in 5 seconds. “Remote Message Delivery” indicates the percentage of the message deliveries to remote users that finish within the measurement period. All these metrics must be 95% or higher. REMOTE changes randomly over the increasing number of users and the pass rate is the highest for the largest number of users (450), which is against our intuition. We consider this result as follows. The message destination (either remote or local) and size are determined by the probability distributions. When the load generator (client simulator) composes a large message to a remote domain near the end of the measurement period, the delivery of the message may not have been finished at the end of the measurement period. This message is placed in the queue for re-transmission and counted as a failure. Since the numbers of users we tested were relatively small, a small number of failures could lead to a lower percentage of the QoS metric. To determine whether or not the REMOTE is really a sensitive QoS metric in a small scale system, we need (for example) to run the measurement for a longer period than defined in [4] (1 hour). Other two metrics are monotonically decreasing over the number of users and from EXPUNGE, it is considered that the maximum performance of the system we used is obtained at around 350 users.

C. Disk Configuration

We use Dovecot as the IMAP4 server [8] that utilizes index files, which include the UID and other information of messages, for faster message retrieval. During the benchmarking shown in the previous subsection, we placed the maildir (the directory where each user's messages are stored), index files and mail queue (spool directory used by the SMTP server) in separate disk partitions and measured their I/O traffic using `iostat`.

Maildir	Index	Mail Queue
29.9	11.7	15.0

TABLE II

AVERAGE REQUEST SIZES FOR 300 USERS IN TWO SATA HDDs. SIZES ARE IN NUMBER OF 512B SECTORS. THESE SIZES ARE RELATIVELY CONSTANT OVER THE NUMBER OF USERS, BUT SLIGHTLY INCREASE FOR A LARGER NUMBER OF USERS. (Maildir is the most typical case where the size increased to 33.4 at 450 users)

Fig. 4 shows the summary of the disk I/O accesses for each of maildir, index and mail queue partitions. The average request sizes for 300 users are presented in Table II. For the numbers of users we tested (250 to 450), we only see the saturation of disk drives in write accesses to the maildir partition. For other partitions, the numbers of read and written sectors are almost proportional to the number of users. However, the average wait time of the access for each partition starts to grow rapidly beyond 300 users.

Another characteristic found in Fig. 4 is that the read request rates and throughputs of index and mail queue partitions are quite lower than those in the write mode access. In Dovecot, while indexes are written to files, they are also cached in the main memory. The lower read access rate of the index partition should imply that most accesses hit at the main memory. We consider that the reason of the lower read access rate than the write access in the mail queue partition is the same: while messages are written to the disk for possible re-transmission, they are discarded without being read as the first transmissions success.

By comparing the throughput and request rates of the maildir and index partitions together with the average request sizes in Table II, while the throughput (total of read and write) of the index partition is only 7 to 8% of the maildir partition, the request rate is around 21 to 23%. In other words, the index partition is accessed with (relatively) frequent smaller requests. Moreover, the storage space for the index is much smaller than that for the maildir: for 350 users, the former takes 715MB while the latter takes 49.5GB¹. These facts suggest us to use smaller but faster disk drives for the index partition and reduce the pressure on the maildir partition. This configuration should be a preferable solution for the users who cannot afford to use high-end HDDs for the entire storage space.

¹This storage space size on our system is translated to 145MB/user which is smaller than 160MB/user in [5].

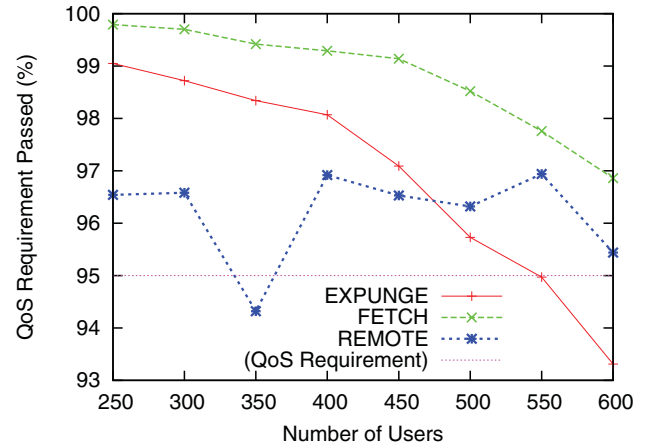


Fig. 2. Quality of Service Metrics (1 SCSI disk for index files).

Fig. 2 shows the QoS metrics for the system using a SCSI HDD for the index and mail queue partitions. Similar to Fig. 1, REMOTE draws an irregular curve (but at a smaller degree than the SATA-only configuration). It drops to 94.3% at 350 users but rises to 96.9% at 400 users. Again, EXPUNGE and FETCH are monotonically decreasing. EXPUNGE drops slightly below 95% (94.97%) at 550 users which seems to be the largest number of users that our system with SCSI and SATA disks can handle. Fig. 5 shows the summary of the disk I/O access for the SATA+SCSI configuration. By comparing at 450 users, the utilization for the maildir is reduced from 67.2% to 41.8%. For larger numbers of users, while the utilization of the disks for the maildir increases to 65% (at 650 users), their wait time is 51ms, which is lower than the wait time in the SATA-only configuration at 450 users (88ms). While the pressure on the SATA disks has been reduced, its utilization is still higher than that of the SCSI disk (12% for index and 4% for mail queue at 450 users). This result implies that the ratio between the numbers of SCSI and SATA disks can be increased (say, 1 to 3 or more) for either scaling up the system or load distribution of the maildir storage.

D. CPU Time Breakdown and Effectiveness of Multiprocessing and Multithreading

Fig. 3 shows the breakdown of the CPU time utilization for the configuration that uses two SATA and a SCSI disks. First, only a small fraction of CPU time is spent for the 'meaningful tasks' and this confirms the I/O-boundness of the IMAP4 mail server. For example, at 550 users, the sum of user and system mode CPU times is only 6% while the I/O wait time is 31%. Second, while the user and system mode CPU times are proportional to the number of users, the I/O wait time grows super-linearly. This should be the result of steeply rising I/O wait time of the disk access (especially, for the maildir drives).

Tables. III and IV show the CPU utilization in the user mode and system mode for the combinations of single and dual CPUs and Hyper-Threading on and off at 550 users. MP-

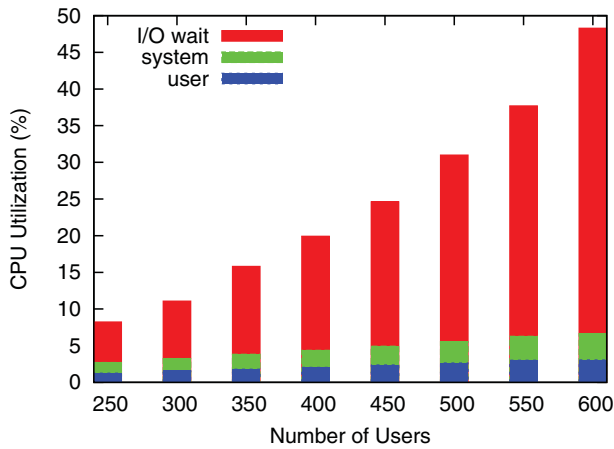


Fig. 3. CPU Time Breakdown

ratio (HT-ratio) indicates the 2CPU case divided by the 1CPU case (the Hyper-Threading=on case divided by the off case). The ratio of 0.5 means that the CPU time required for the same task is halved by adding either a physical (2CPU case) or a logical (HT=on case) CPU.

	HT=off	HT=on	HT-Ratio
1CPU	6.22	3.49	0.56
2CPU	3.05	1.71	0.56
MP-Ratio	0.49	0.47	

TABLE III
USER MODE CPU TIME FOR 550 USERS.

The MP-Ratios in Table III indicate that 2CPU achieves more than $2\times$ speed-up in the user mode. A possible reason of this super-linear speed-up is the reduction of the number of cache misses since each CPU has own L1 and L2 caches and they store the data structures necessary for handling the transactions running of the CPU. On the other hand, the MP-Ratios in Table IV are higher than those in Table III. In other words, dual-CPU is not as effective in the system mode as in the user mode. This is probably because the system tasks are likely to lock data structures for mutual exclusions which result in reduction of the parallelism between processes.

The Hyper-Threading is less effective in the 1CPU system mode than all other cases. Similar to the difference in the MP-Ratios, the synchronization overhead should also be the cause of this difference. The effectiveness of the Hyper-Threading is equal for the 1CPU and 2CPU cases in the user mode (0.56), while it is different in the system mode (0.65 vs 0.56). The sharing of the caches by two logical CPUs on a single physical CPU might be the reason but further investigations are required.

IV. RELATED WORK

In 2008, SPEC published SPECmail2008 which also modeled the enterprise mail servers [9]. It was retired in July 2009 and its workload was refined to SPECmail2009. The

	HT=off	HT=on	HT-Ratio
1CPU	5.46	3.56	0.65
2CPU	3.24	1.82	0.56
MP-Ratio	0.59	0.51	

TABLE IV
SYSTEM MODE CPU TIME FOR 550 USERS.

analysis of mail server workload that resulted in the design on SPECmail2009 is found in [10]. The mail store structure in SPECmail2008 was modeled after the data from Sun Microsystems, while that of SPECmail2009 is modeled after the data from Apple. VMmark from VMware is a proprietary benchmark for a consolidated server [11]. It assumes a mail server running MS Exchange and LoadSim is used as the workload driver [12].

In [13], the authors described the standard Internet mail protocols (POP3, STMP and IMAP4) and presented workload characterizations of these protocols based on the data collected at a small ISP. They also presented the results of IMAP4 benchmarking but detail was not provided. The paper was published in 2004 and the assumptions (both in the system configuration and in the workload) are quite different from those in SPECmail2009. [14] is a mail server workload analysis based on SPECmail2001. They presented break-downs of SMTP and POP3 (with and without download) session latencies. [15] presents an analysis of the data collected at the OpenMail file system which was used by the HP Lab with 3000 users. They provided I/O request rates, request sizes, bandwidths and queue lengths, separated by read and write. They tried to find invariances (self-similarities) of these statistical data.

V. CONCLUSION

In this paper, we presented a case study of SPECmail2009, which was a benchmark program for evaluating a mail server in an enterprise environment. Unlike a POP3-based mail server which is the assumption of SPECmail2001, SPECmail2009 assumes the IMAP4 protocol by which users can directly access the messages on the server. This feature of the IMAP4 protocol makes SPECmail2009 a highly I/O-bound workload.

As an option of building a cost-effective mail storage, we evaluated a combination of SCSI and SATA disks. The Dovecot IMAP4 server utilizes message indexing for a faster access to the message files. We demonstrated that a SCSI disk for storing index files could effectively reduce the load on the SATA disks that were used for message files.

Multiprocessing and multithreading were quite effective in speeding up this typical server workload, especially in the user mode CPU time. Further investigation is planned to understand the less effectiveness of these techniques in the system CPU mode.

REFERENCES

- [1] "Post Office Protocol - Version 3," <http://www.ietf.org/rfc/rfc1939.txt>
- [2] SPECmail2001 <http://www.spec.org/mail2001/>
- [3] , "INTERNET MESSAGE ACCESS PROTOCOL - VERSION 4rev1," <http://tools.ietf.org/html/rfc3501>
- [4] SPECmail2009 <http://www.spec.org/mail2009/>
- [5] "SPECmail2009 Benchmark Architecture White Paper," SPEC OSG Mailserver Subcommittee, 22 December 2008.
- [6] "SPECmail2009 Run and Reporting Rules," <http://www.spec.org/mail2009/docs/runrules.html> .
- [7] Deborah T. Marr, et. al, "Hyper-Threading Technology Architecture and Microarchitecture," Intel Technology Journal, Vol. 6 Issue 1, pp1-12, February 2002.
- [8] Dovecot, <http://www.dovecot.org/>
- [9] SPECmail2008, <http://www.spec.org/mail2008/>
- [10] Mike Abbott and Yun-seng Chao, "SPEC OSG SPECmail2009 Benchmark, Workload Characterization for SPECmail_Ent2009 Metric," Standard Performance Evaluation Corporation.
- [11] "VMware Virtualization Benchmark," <http://www.vmware.com/products/vmmark/>
- [12] "Microsoft Exchange Server 2003 Load Simulator (LoadSim)," available from <http://www.microsoft.com/downloads/>
- [13] M. Calzarossa, "Performance Evaluation of Mail Systems," in *Performance Tools and Applications to Networked Systems*, vol. 2965 of Lecture Notes in Computer Science, pp51-67, Springer, 2004.
- [14] Jun Wang and Yiming Hu, "A Performance Study on Internet Server Provider Mail Servers," in *Proceedings of Ninth IEEE Symposium on Computers and Communications 2004 (ISCC'04)*, vol. 1, pp56-61, 2004.
- [15] Qiang Zou and Dan Feng, "OpenMail File System Workloads Analysis and Characterization," in *Proceedings of The 9th International Conference for Young Computer Scientists*, pp71-76, 2008.

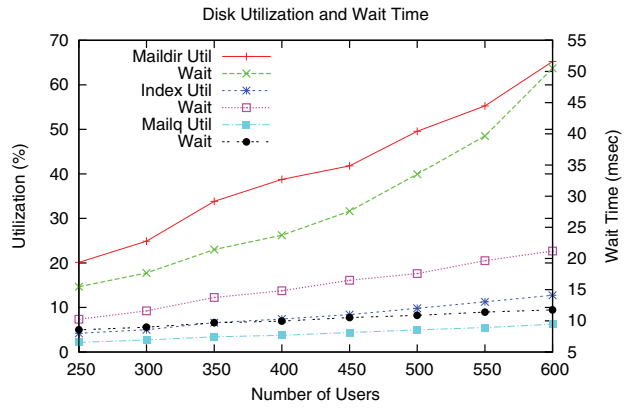
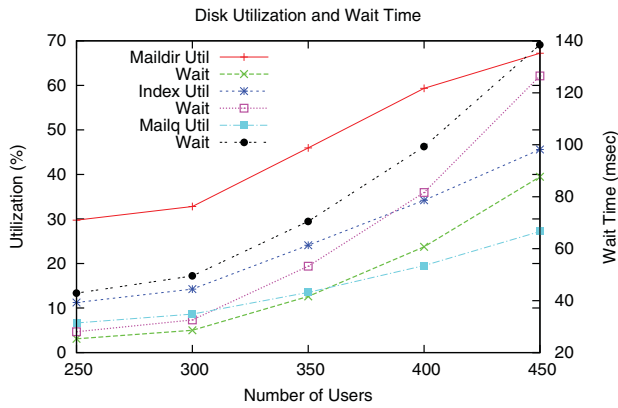
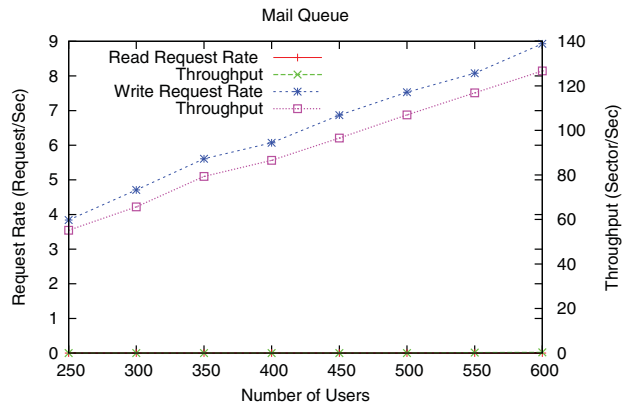
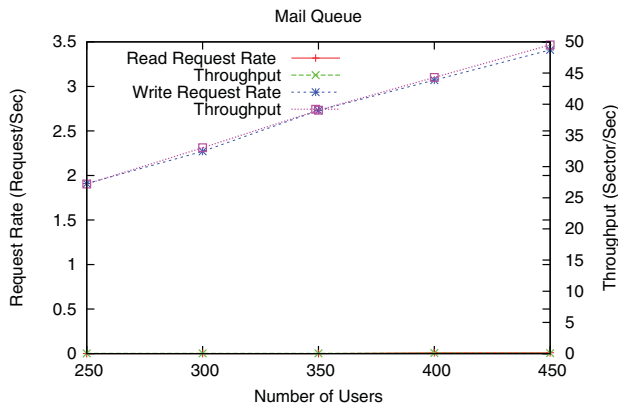
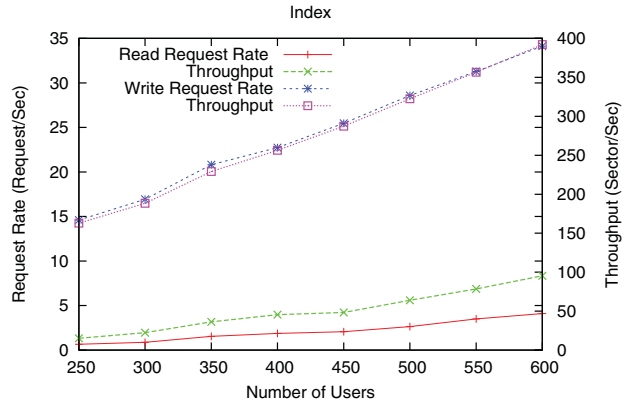
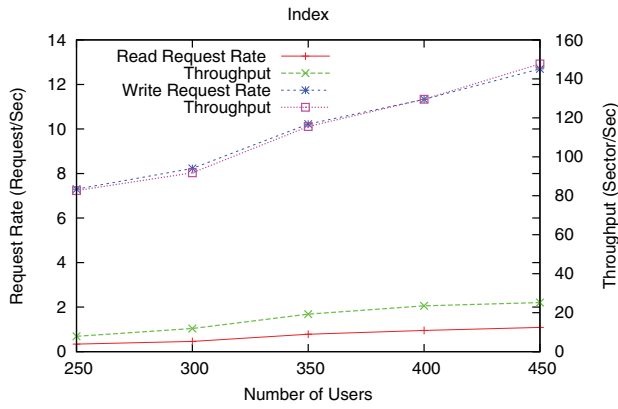
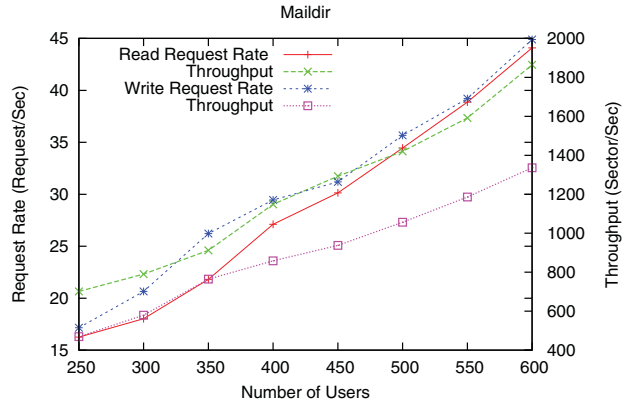
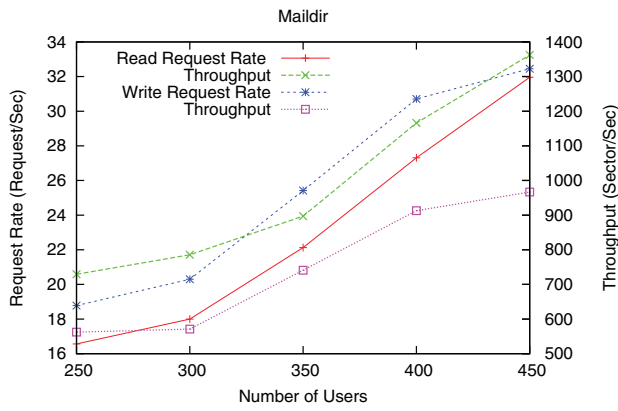


Fig. 4. I/O Behavior for Two SATA Disks

Fig. 5. I/O Behavior for 2 SATA (Maildir) and 1 SCSI (Index and Mail Queue)