



POST INCIDENT REPORT

Incident Summary	
Incident ID:	E2745
Date / Time of Incident:	01/10/2012 3:00PM ET
Duration of Incident:	3d 1h 12m
Scope of Incident:	Intermittent Mail Performance

Incident Details
Description of Incident
<p>Early morning 10.Jan.2012, the hard disk drives (HDDs) on mail77 were upgraded with brand new hard drives.</p> <p>Around 3PM 10.Jan.2012, the e-mail spool on mail77.safesecureweb.com reached approximately 20,000 messages. E-mail had been experiencing some delays prior but still sending. At the time no indication of a software or hardware issue had presented itself. The messages in the spool were moved to another folder and fed back into the spool manually. We'd seen some indications of high read/writes (disk I/O) to the mail server's hard drives. While not uncommon for a mail server to have significant quantities of data read and written to disk, the load was spiking in a way indicative of the mail server receiving a high amount of traffic akin to a ddos, though when pouring through the possibilities on an already slowly performing server (SMTP connections and statistics for mail sent, etc.) there was no evidence to point to a single issue. After ruling out as many potential issues as possible, we settled on I/O to the hard disks. We found windows indexing service running and stopped the indexing service as well as the nightly backup which was still in progress. This did alleviate some of the issues and mail began to send as it normally processed at the time.</p> <p>When the spooling issues appeared to continue in the evening, the new chassis was set live when it was ready at 9PM 10.Jan.2012. The spool was divided into multiple sub-spools then fed back into the main spool. During the migration of Mail77 to new hardware, all mail was delivered to the original server. Once the new server was brought online and functioning correctly, we cut over to it. Mail queued to send in the spool on mail77 was clear shortly afterward. At the time symptoms of performance issues had not reappeared and the server was performing well. Load was low and disk I/O was within normal boundaries.</p> <p>Load was high as normal business hours resumed 11.Jan.2012. The new hardware was experiencing a similar issue as the old hardware. With a pair of drives from the old server in the new hardware, the</p>

new server was allowed to rebuild its HDD RAID array overnight. Rebuilding of a HDD RAID array during off-peak hours with new hardware is generally completed in a few hours. When looking at the new hardware's RAID rebuild status in the morning on 11.Jan.2012, the percentage complete stayed at 0%. The drives were failing and the RAID card manager was not reporting the error.

After it was discovered the new hardware might have a set of bad drives, we set the old hardware live again and began a robocopy of data from the old server to the new hardware. We felt it better to have slow performance than none at all.

With increasing stability through the day we continued to see spooling into late afternoon on 11.Jan.2012. That said, the level was significantly reduced from the previous day but was still high and the server sending mail. We connected the old replacement hardware to the new replacement hardware server and initiated a robocopy. The robocopy would ensure the new hardware has a non-fragmented copy of data which would also contribute to decreased load on the hard drives.

In the evening of 11.Jan.2012, we replaced the RAM, updated the firmware on the RAID card and updated the BIOS on the live server just to be sure none of those was an issue. The versions of firmware and BIOS were very recent and do get updated regularly on our servers, this server just did not have the very latest version. Performance had once again increased on the old hardware after the updates and replacement of RAM. We were confident the server's issues were close to resolution but needed to wait for load to return to normal to know for sure.

After performing well and in a stable manner until approximately noon on 12.Jan.2012, the spool began to grow again. We found a bad drive in the new replacement server which had the same manufacture date as the two drives in the live server. In short, the RAID card was now throttling the drives so they would not lose data parity in the RAID array. Because the drives were not reporting errors in the live server there was no indication they were the cause of the performance issues two days prior. Thus when RAID was broken to bring the most recent copy of data from the live server to the old replacement hardware, we had no way to know at the time the bad drives had been chosen. After pulling the drives from that suspect batch out of the live server, the server began to perform significantly better but was running on half the hard drives. We placed a new drive in the first RAID array on the live server and it successfully rebuilt in three hours. We placed a new hard drive in the second RAID array on the live server and it rebuilt successfully during the evening.

The lingering issue during the day was the number of connections available on the live server and after switching traffic to a new network card, we saw POP3 connection issues to the server reduce significantly. With the issue finally known on 12.Jan.2012 and the live server stable, we began a repeating robocopy to synchronize data between the live server and new replacement hardware. Smartermail on the live server was set to be upgraded in a week. If the live hardware failed before the SmarterMail upgrade scheduled for the following week, we would be able to cut over to the new hardware quickly and with minimal downtime. The server stayed stable through the day of the SmarterMail upgrade and continues to be stable.

Root Cause of Incident

The hard disk drives (HDDs) on mail77 were upgraded early in the morning of 10.Jan.2012. The HDDs used in the drive upgrade that morning were from a batch of HDDs not known to have an issue. The RAID card in the live server or in the initial replacement chassis were not reporting issues or

degradation of the HDDs.

Incident Resolution

Once the hard drives were discovered to be the issue, they were replaced. This allowed RAID to be rebuilt on the server and disk performance to improve to acceptable levels.

Recommended or Corrective Action

Reporting the drives to the manufacturer including all relevant information. During the process we attempted to provide regular updates to customers and staff via internal communication and blog.hostmysite.com. The frequency proved to be insufficient. We are going to increase frequency of updates to the blog in future incidents.