

## Re: Large Number of Storage Groups

---

*Source:*

<http://www.tech-archive.net/Archive/Exchange/microsoft.public.exchange.admin/2007-08/msg02882.html>

---

- *From:* "Mike Abbaticchio" <[nospam@xxxxxxxxxxxxxxxxxx](mailto:nospam@xxxxxxxxxxxxxxxxxx)>
  - *Date:* Thu, 30 Aug 2007 23:32:58 -0400
- 

We will be non-clustered mailbox role only servers. They will be quad core with 16GB of RAM. I used Microsoft's general guidelines at Tecnet, assuming every user, a heavy user, to size the RAM. Design includes db replication on a different SAN frame than prod, and VSS based backups. I still have to meet with EMC about our backup options.

Speaking of IOPs, is there a good calculator spreadsheet for 2007? I have one for 2003.....

"John Fullbright" <[fjohn@donotspamenetappdotcom](mailto:fjohn@donotspamenetappdotcom)> wrote in message [news:OFPKE436HHA.3940@xxxxxxxxxxxxxxxxxxxxxxxxxxxx](mailto:news:OFPKE436HHA.3940@xxxxxxxxxxxxxxxxxxxxxxxxxxxx)

I've done designs in 30K – 120K seat range for Exchange 2007 on SAN (NetApp of course). 2800 users is a fairly low number for a single node. Make sure you get the latest MS calculator and use those numbers as the base. I tend to cross check the IO numbers generated by the log counts, message counts, and avg message size with permon data physical disk data (transactions/sec reads/sec and writes/sec) So far, it's always been +/- 10%. Having done 2003 to 2007 migrations, I'd have to say we really are seeing a 60% to 70% decrease in IO. In most of these, a major feature of the upgrade was also an increase in mailbox size. When you combine the reduced IO requirement with the increased space requirement, going with fewer bigger drives is a no brainer.

One thing to consider is "when and if performance degradation is acceptable". Going beyond normal operational states, you have to consider cluster failover, cold start, and log replay. When and if performance degradation is acceptable also impacts your backup strategy in CCR clusters.

You can backup from the passive node offloading the IO consumed by the backup process. On the surface, this seems attractive until you consider the restore scenario. If you need to restore any single DB, then you must fail the entire cluster over to the passive node. This is due to limitations of the Exchange 2007 VSS Writer; you can only restore to an active node. This means failover, log replay, and cold start impacting all users; the passive is now active. You have to take this into

## Re: Large Number of Storage Groups

consideration when sizing storage on the passive node.

If, on the other hand, you backup the active node then you can simply restore a single SG and only impact the users on the single database contained in the SG. There's no replay or cold start scenario, but you have to add the additional IO capacity to support the backups on the active node. If you backup both the active and passive nodes, you get single SG recoverability with minimal user impact regardless of the state of the cluster. The cost is the added IO capacity on both the active and passive for the backups.

Anyway, as to your specifics, you didn't mention mailbox size or IOPS/user or if you will be using clustering and if so what form. The clustering type determines the max recommended DB size, and the mailbox sizes and number of users get you to how many DBs. Using the recommended 1:1 mapping of DBs to SGs, you get the number of SGs. The IOPS/user gets you to the cache per user requirement which combined with RAM requirements for the number of SGs you have gives you the RAM you want in the host. It is very important to have enough RAM in the host. If you undersize the RAM, then IO to the storage increases due to increased cache database page faults. You'd need to provide more information to determine if the host is sized properly.

John

"Mike A" <nospam@xxxxxxxxxxxxxxxx> wrote in message  
[news:1188486000.797887.174660@xx](mailto:news:1188486000.797887.174660@xx)

We are aware of the SG design. At one time Microsoft's position on this was to minimize the number of SG's and that was what our design was based on. EMC had originally recommended RAID 5 configs for the db's. The disk read performance was really bad, and we ended up redoing the existing servers with RAID 10 configs for our existing servers and configuring all new servers this way, right off the bat.

For 2007, I am planning one DB per SG. I wasn't sure if Exchange could use mount points or not. That would definitely solve the problem. For our existing config, performance is stable at this time, and I am hoping it will remain that way until we start moving towards 2007.

Anyone know of organizations right now running large back end servers? I am wondering if I can get away with planning around 2800/ server. The hardware would be quad core with 16GB RAM.

On Aug 30, 6:30 am, "Mark Arnold [MVP]" <m...@xxxxxxx> wrote:

On Thu, 30 Aug 2007 06:00:31 -0400, "Mike Abbaticchio"

## Re: Large Number of Storage Groups

<nos...@xxxxxxxxxxxxxxxx> wrote:

My current layout is Exchange 2003 with 2 SG's / 8Db's all on the same same RAID 10 LUNs, with separate RAID 1 LUNs for xlogs. I went through all the IOP calculations with EMC and have acceptable performance, but the read times on the DB drive were never as good as I expected. Needless to say, I am being real careful with the 2007 plan.

Thanks for your response.... I was just wondering if there was a way out of the drive letter limitations. I will more than likely end up using less storage groups and end up with larger db sizes as a compromise.

Mike Abbaticchio

I'm never convinced what RAID10 gets you on platforms like EMC. Are you expecting drives to fail that often? The performance gains are negligible. Far better would be to go RAID6 and spread out on all those suddenly available spindles. Like the other posted says, Mount points are the way to go in terms of eliminating the drive letter challenge. You've no choice later in 2007 if you fill the full 50 stores anyway :-)

Your configuration is not in line with 2003 best practice though (I assume you are on W2K3?). The "fill a storage group with stores and then make another storage group" thing was correct in 2000 (on 2000 or 2003) but not in 2003 on 2003. You are affecting your log

## Re: Large Number of Storage Groups

disk write  
performance and potentially hitting buffer problems.

First job for me would be to make two more SGs and spread  
all the  
users across them. Then I'd look at how my logs were  
performing. Then  
I'd have a think about the read cache on the EMC (whatcha  
got? DMX,  
Clariion, Celera?)