Storage Performance for Microsoft Server Applications

Dennis Martin

President, Demartek

February 2009

Agenda

- Quick Review of January 2009 Presentation
  - Disk Storage Comparison Review
  - Microsoft Server Application Best Practices and Recommendations
- Performance Data – February 2009

Introduction

- Storage is where the data resides
- Access to storage in Exchange Server, SQL Server and SharePoint Server is of critical importance
- Understanding storage technologies is important

Solid State Disk (SSD)

- Uses memory technology designed to appear as an online storage (disk) device
  - DRAM, NAND flash or combination
- Extremely fast
- Capacities vary from 8 GB to 1 TB+
- Expensive (although prices dropping)
- DRAM-based storage almost always includes battery-backup and disk-drive for safety

Disk Drives: Comparison Chart

Device	Enterprise	Desktop	Notebook	Consumer
Avg. seek time	3 – 5 ms	8 – 11 ms	10 – 15 ms	12 – 15 ms
Xfer rate (MB/s)*	70 – 170	60 – 120	30 – 80	6 – 40
RPM (K)	10, 15	5.4, 7.2, 10	4.2, 5.4, 7.2	3.6, 4.2
Capacities	Large	Very large	Medium	Small
Processors	2	1	1	1
Cmd. Queuing	TCQ or NCQ	NCQ	NCQ	-
Power need	High	Medium	Low	Very low
Warranty	5 – 7 years	3 – 5 years	1 – 5 years	1 – 3 years

* Maximum device transfer rate in megabytes per second from drive surface to buffer. Sustained rates are lower. This is not the same as the interface transfer rate.

Disk Drives: Recommendations

- Microsoft Server Applications, only consider enterprise and desktop drives
- Enterprise drives = Performance
  They will run out of capacity before they run out of performance
- Desktop drives = Capacity
  They will run out of performance before they run out of capacity

Storage Interface Comparison

	SATA	SAS	Fibre Channel	USB
Number of devices	1	16K	16M	127
Maximum distance	1 meter	10 meters	100+ KM	5 meters
Cable type	Copper	Copper	Fiber Optic	Copper, wireless
Interface type	Serial	Serial	Serial	Serial
Transfer speeds (MB/sec)	150, 300	300, 600	100, 200, 400, 800	0.15, 1.5, 50, ~500*

The updated version of this chart: http://www.demartek.com/Demartek_Interface_Comparison.html

MB/sec = Megabytes per second, which is generally calculated as megabits/second (Mbps) divided by 10 for planning purposes

* *SuperSpeed USB devices expected in 2010*

RAID

(**R**edundant **A**rray of **I**ndependent **D**isks)

- RAID 0: Interleaving or “striping” data across two or more disks
- RAID 1: Disk mirroring – same data written on two different disks (data can be rebuilt if drive fails)
- RAID 5: Data striping with parity across multiple disks (data can be rebuilt if drive fails)
- RAID 6: Data striping with double parity across multiple disks (data can be rebuilt if two drives fail)
- RAID 1+0 or RAID 10: combine RAID 1 and RAID 0

RAID Comparison Chart

	Transactional I/O Performance	Capacity Utilization	Disk Failure and Rebuild Performance
RAID 0	Good	Best	Poor
RAID 1	Best	Poor	Best
RAID 5	Good	Good	Moderate
RAID 6	Good	Moderate	Good
RAID 10	Best	Poor	Best

RAID Recommendations

- Choose RAID to spread data over multiple disks (“spindles”) to get better performance and reliability than using individual disks
- Best overall performance: generally RAID10
- Best capacity (with recoverability): generally RAID5

Recommendations – 1

- Configure database servers with performance and availability as design criteria
- Use more disks and faster disks for best performance
  - If you choose SATA disk drives, you're usually emphasizing capacity above performance (this choice may also reflect your budget)

Recommendations – 2

- Remember some things from last month's presentation:
  - Disk alignment
  - Format Allocation (cluster) size
  - Disk stripe size
  - Microsoft Best Practices TechNet references

Performance Tests – Components

- Comparison of Microsoft application workloads using SSD, SAS and SATA disks
  - Server: Dual Intel Xeon E5320, 1.86 GHz (8 total cores), 4GB RAM, Windows Server 2003 R2 Enterprise x64
  - SSD: Fusion-IO ioDrive, 160GB, SLC NAND-flash, PCI-express 1.1 interface, no cache (quantity 1)
  - SAS: Seagate Cheetah 15K.5, 146GB, 15K RPM, 16MB cache (quantity 10)
  - SATA: Seagate Barracuda 7200.11, 500GB, 7200 RPM, 32MB cache (quantity 10)
  - Disk controller: Intel SRCSASJV, 512MB Cache, supports up to 240 SAS or SATA disk drives

Performance Tests

1. Microsoft SQLIOSim
   - Microsoft SQL Server(c) Simulator Stress Test Version 9.00.1399.05
   - Simulates SQL Server I/O workloads
2. Microsoft Exchange Jetstress
   - Microsoft Exchange Server Jetstress Version 08.02.0060.000
   - Simulates Exchange Server 2007 workloads

Performance Data – Scenario 1

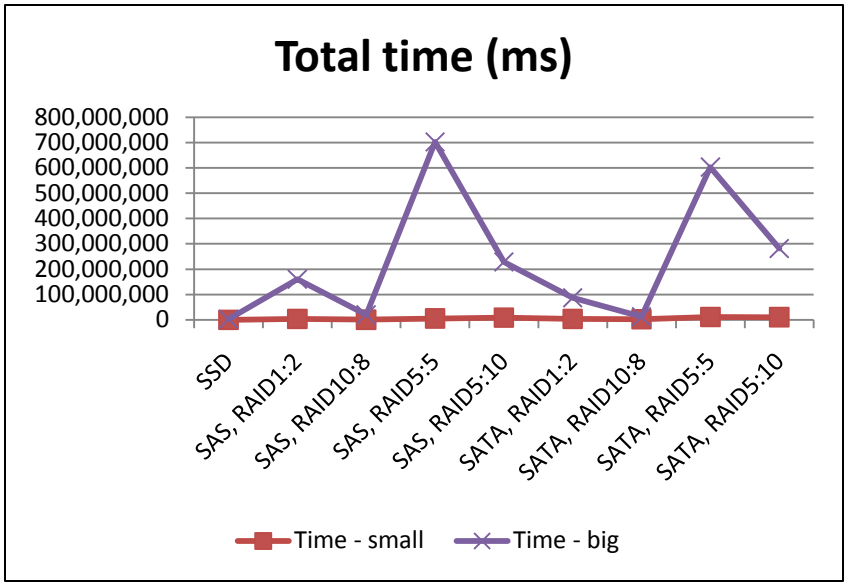
- Used Microsoft's SQLIOSim utility
  - Database and Logs on same volume (not best practice, but kept the tests equivalent)
  - Log: InitialSize = 50 MB, MaxSize = 50 MB, Increment = 0 MB, LogFile = Yes, Shrinkable = No, Sparse = No
- Two sets of tests with different database sizes
  1. InitialSize = 500 MB, MaxSize = 1000 MB, Increment = 50 MB, LogFile = No, Shrinkable = No, Sparse = No
  2. InitialSize = 5000 MB, MaxSize = 10000 MB, Increment = 500 MB, LogFile = No, Shrinkable = No, Sparse = No
- SQLIOSim detected no disk cache on SSD

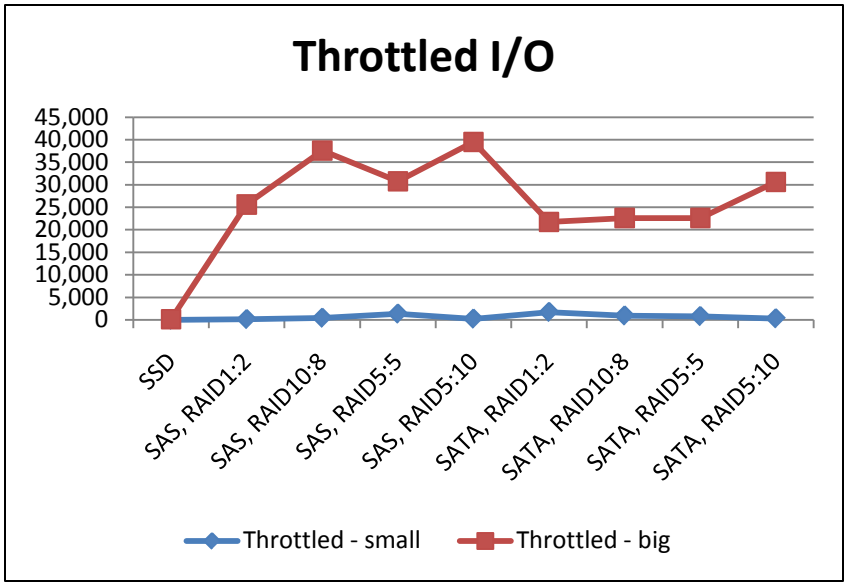
Performance Configuration 1

Device	Quantity	Capacity	RPM	RAID
SSD	1	160GB	-	None
SAS	2	146GB	15000	RAID1
SAS	8	146GB	15000	RAID10
SAS	5	146GB	15000	RAID5
SAS	10	146GB	15000	RAID5
SATA	2	500GB	7200	RAID1
SATA	8	500GB	7200	RAID10
SATA	5	500GB	7200	RAID5
SATA	10	500GB	7200	RAID5

All SQLIOSim tests used a single 100GB partition, aligned at 64K, format allocation unit (cluster size) = 64K

Performance Results – 1





	Small		Big	
Device, RAID:Qty	Throttled	Total time (ms)	Throttled	Total time (ms)
SSD, noRAID:1	0	87,966	151	2,256,668
SAS, RAID1:2	141	3,753,640	25,613	159,895,463
SAS, RAID10:8	432	640,560	37,568	20,112,550
SAS, RAID5:5	1,348	5,305,152	30,781	702,111,985
SAS, RAID5:10	235	8,399,242	39,485	228,298,583
SATA, RAID1:2	1,720	3,722,299	21,740	87,231,755
SATA, RAID10:8	931	2,364,109	22,614	11,981,468
SATA, RAID5:5	806	11,047,995	22,622	603,086,114
SATA, RAID5:10	307	10,421,506	30,611	281,417,172

Small = 0.5GB – 1GB database
Big = 5GB – 10GB database

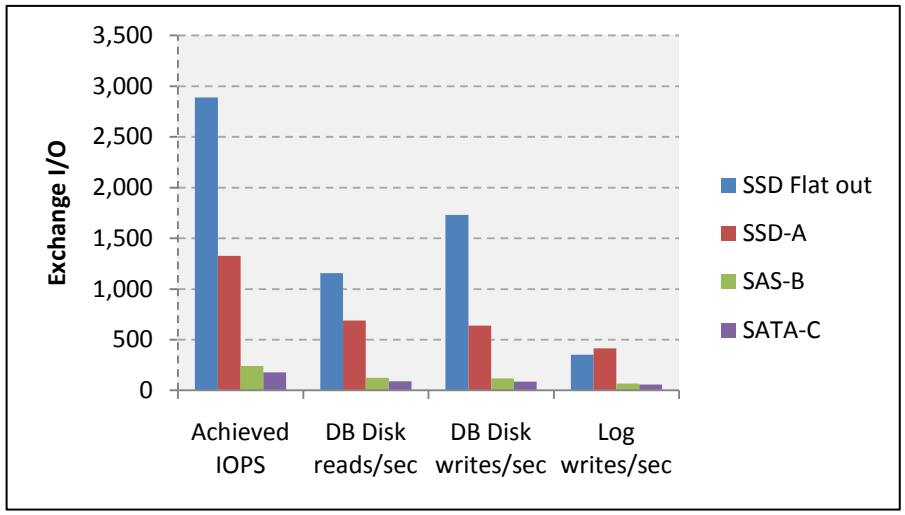
For "Big" databases, SSD and RAID10 (8 drive) configurations had no delayed I/O. All other configurations of "Big" databases had delayed I/O of at least 15 seconds.

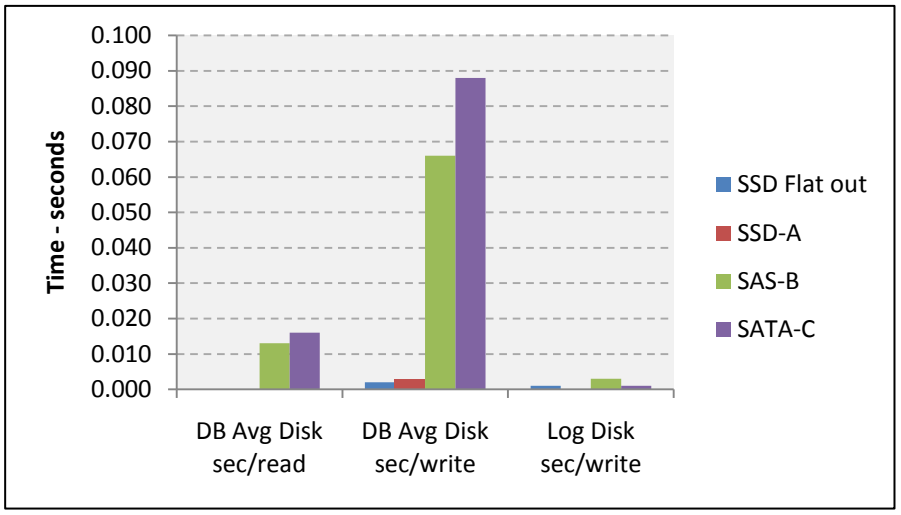
Performance Data – Scenario 2

- Microsoft Exchange 2007 Jetstress
  - 500 mailboxes (550 for SATA)
  - Mailbox size=250MB

Config.	Device	Exchange volumes	Exchange IOPS	Qty.	Capacity	RPM	RAID
SSD-A	SSD	DB + Log	0.5 (heavy)	1	160GB	-	None
SAS-B	SAS	DB	0.5 (heavy)	2	146GB	15000	RAID1
	SAS	Log	0.5 (heavy)	2	146GB	15000	RAID1
SATA-C	SATA	DB	0.3 (light)	2	500GB	7200	RAID1
	SATA	Log	0.3 (light)	2	500GB	7200	RAID1

Performance Results – 2





	Achieved IOPS	DB Disk reads/sec	DB Disk writes/sec	Log writes/sec
SSD Flat out	2889.190	1157.621	1731.569	350.849
SSD-A	1327.780	688.065	639.715	414.962
SAS-B	242.249	123.395	118.854	66.224
SATA-C	177.069	89.427	87.642	59.254

	DB Avg Disk sec/read	DB Avg Disk sec/write	Log Avg Disk sec/write
SSD Flat out	0.000	0.004	0.001
SSD-A	0.000	0.005	0.000
SAS-B	0.013	0.066	0.003
SATA-C	0.016	0.088	0.001

Performance – Addendum

- Interesting Jetstress comparison with same SSD on different servers and operating system versions
  - Server 1: as described previously (4GB RAM)
  - Server 2: Dell PowerEdge 2900, dual Intel Xeon E5345 (2.33 GHz, 8 cores), 32GB RAM, Windows Server 2008 x64
  - Exchange Server 2007 changes the I/O mix with increased system RAM to be more efficient

	Achieved IOPS	DB Disk reads/sec	DB Disk writes/sec	Log writes/sec
1: SSD Flat out	2889.190	1157.621	1731.569	350.849
1: SSD-A	1327.780	688.065	639.715	414.962
2: SSD Flat out	6121.350	3397.941	2723.409	757.635
2: SSD-A	2303.668	1236.723	1066.945	699.417

	DB Avg Disk sec/read	DB Avg Disk sec/write	Log Avg Disk sec/write
1: SSD Flat out	0.000	0.004	0.001
1: SSD-A	0.000	0.005	0.000
2: SSD Flat out	0.000	0.002	0.001
2: SSD-A	0.000	0.003	0.000

Contact me

Dennis Martin, President
Demartek
(303) 940-7575
dennis@demartek.com

www.linkedin.com/in/dennismartin