How to Pick SQL Server Hardware

BRENT OZAR
UNLIMITED

The big picture

1. What SQL Server edition do you need?

2. Does your RPO/RTO dictate shared storage?

3. If you need shared storage, what's important?

4. No-brainer answers for Standard Edition

5. Enterprise Edition's socket choices

What SQL Server edition do you need?

Picking your SQL Server edition

	Standard	Enterprise
Cost per core	£2k	£5k
Cost for 2-socket, hex-core	£16k	£62k
Cost for 4-socket, hex-core	--	£125k
Max logical processors (cores)	16	OS max
Max memory used by SQL Server	128GB	OS max
Backup compression	Yes	Yes
Data compression	No	Yes
Transparent data encryption	No	Yes
Partitioned tables	No	Yes
In-memory OLTP (Hekaton), clustered column store indexes	Not Really	Yes

* These prices don't take into account your amazing negotiation skills.

Factor in the hardware costs

	SQL 2014 Standard	SQL 2014 Enterprise crippled	SQL 2014 Enterprise
Cores	12 (2x6)	12 (2x6)	12 (2x6)
Memory	128GB	128GB	768GB
Cool Stuff	No	Yes	Yes
Hardware w/TempDB SSDs	£9k	£9k	£21k
Licensing	£16k	£62k	£62k
Total Cost	£25k	£71k	£83k

* These prices don't take into account your amazing negotiation skills.

Got >500GB data? Forget Standard.

Merry-Go-Round Scans in Enterprise

tblSales – 1,000GB

My Query

Merry-Go-Round Scans in Enterprise

tblSales – 1,000GB

My Query →

Merry-Go-Round Scans in Enterprise

tblSales – 1,000GB

My Query

Your Query

Merry-Go-Round Scans in Enterprise

tblSales – 1,000GB

My Query

Your Query

Merry-Go-Round Scans in Enterprise

tblSales – 1,000GB

My Query

Your Query

Merry-Go-Round Scans in Enterprise

tblSales – 1,000GB

Continued

Your Query

Books Online "Reading Pages"

The storage engine uses prefetching to speed base table lookups from nonclustered indexes. The leaf rows of a nonclustered index contain pointers to the data rows that contain each specific key value. As the storage engine reads through the leaf pages of the nonclustered index, it also starts scheduling asynchronous reads for the data rows whose pointers have already been retrieved. This allows the storage engine to retrieve data rows from the underlying table before it has completed the scan of the nonclustered index. Prefetching is used regardless of whether the table has a clustered index. SQL Server Enterprise uses more prefetching than other editions of SQL Server, allowing more pages to be read ahead. The level of prefetching is not configurable in any edition. For more information about nonclustered indexes, see Nonclustered Index Structures.

◢ Advanced Scanning

In SQL Server Enterprise, the advanced scan feature allows multiple tasks to share full table scans. If the execution plan of a Transact-SQL statement requires a scan of the data pages in a table and the Database Engine detects that the table is already being scanned for another execution plan, the Database Engine joins the second scan to the first, at the current location of the second scan. The Database Engine reads each page one time and passes the rows from each page to both execution plans. This continues until the end of the table is reached.

Does your RPO/RTO dictate shared storage?

BrentOzar.com/go/fail

High Availability

For isolated server failures:

Windows crash, RAID controller failure, SQL or Windows patch fails, C drive full, bad memory chip, wrong box unplugged

RPO: Max acceptable data loss:

Biz Goal	IT Goal	Current	
			Zero
			1 minute
			1 hour
			1 day

RTO: Max acceptable downtime:

Biz Goal	IT Goal	Current	
			Zero
			1 minute
			1 hour
			1 day

Disaster Recovery

For widespread outages:

Data center power or network outage, domain controller failure, SAN failure, fire, quake, zombies in the data center

RPO: Max acceptable data loss:

Biz Goal	IT Goal	Current	
			Zero
			1 minute
			1 hour
			1 day

RTO: Max acceptable downtime:

Biz Goal	IT Goal	Current	
			Zero
			1 minute
			1 hour
			1 day

"Oops" Deletions

For human T-SQL error:

DELETE without a where clause, bug in stored procedure for updates, end user needing a restore due to human error

RPO: Max acceptable data loss:

Biz Goal	IT Goal	Current	
			Zero
			1 minute
			1 hour
			1 day

RTO: Max acceptable downtime:

Biz Goal	IT Goal	Current	
			Zero
			1 minute
			1 hour
			1 day

BrentOzar.com/go/fail

Recovery Time Objective (RTO) - Downtime

	Zero-1 Sec	1 Minute	1 Hour	1 Day
Zero-1 Sec	**>$1,000,000** Multiple active servers with bidirectional replication (usually requires code change)	**$100k-$500k** Clustering w/SAN, synch AlwaysOn Availability Groups (EE), synch database mirroring	**$100k-$250k** Synch SAN replication, synch VM replication	
1 Minute			**$50k-$250k** Async AlwaysOn Avail Groups (EE), async DB mirroring (EE)	**$5k-$100k** Log shipping, async SAN replication, async VM replication
1 Hour			**$5k-$100k** Log shipping, async SAN replication, async VM	
1 Day				

Recovery Point Objective (RPO) - Data Loss

Comparison of SQL Server Availability Features

	>$1mm	$100k-$500k			$100k-$250k		$50k-$250k		$5k-$100k		
	Replication	Failover Clustered Instances (FCI)	AlwaysOn Availability Groups - Synch	Database Mirroring - Synch	SAN Replication · Synch	VM Replication · Synch	AlwaysOn Availability Groups - Asynch	Database Mirroring - Asynch	Log Shipping	SAN Replication · Asynch	VM Replication · Asynch
Recovery Point Objective (Data Loss)											
Data loss possible	Yes	No*	No*	No*	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Typical RPO goal	Zero	Zero	Zero	Zero	Zero	Zero	1 Minute	1 Minute	1 Hour	1 Hour	1 Hour
Recovery Time Objective (Downtime)											
Failover automatic or manual	Automatic	Automatic	Automatic	Automatic	Optional	Optional	Manual	Manual	Manual	Manual	Manual
Can be failed over by the DBA alone	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	No	No
Easy to fail back and forth for patching	Some	Yes	Yes	Yes	No	No	Yes	Yes	Some	No	No
Multiple live writeable servers	Optional	No	No	No	No	No	No	No	No	No	No
Fails over Agent jobs, SQL logins	No	Yes	No	No	Optional	Yes	No	No	No	Optional	Yes
3rd party apps fail over easily	No	Yes	Yes	Some	Yes	Yes	Yes	Some	No	Yes	Yes
Typical RTO goal	Zero	1 Minute	1 Minute	1 Minute	1 Hour	1 Hour	1 Hour	1 Hour	1 Day	1 Day	1 Day
Groups of Databases											
Are the same point in time	No	Yes	No*	No	Yes	Yes	No	No	No	No	No
Fail over together	No	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Hardware & Software Requirements											
SQL Server minimum version supported	Apps vary	Any	2012	2005	Any	Any	2012	2005	Any	Any	Any
SQL Server minimum edition required	Apps vary	Any	Enterprise	Standard	Any	Any	Enterprise	Enterprise	Any	Any	Any
May require application changes	Yes	No	No	No	No	No	No	No	No	No	No
Req storage for multiple copies of DBs	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Requires shared storage (SAN)	No	Yes	No	No	Yes	Yes	No	No	No	Yes	Yes
Requires virtualization	No	No	No	No	No	Yes	No	No	No	No	Yes
Can use local solid state storage	Yes	No	Yes	Yes	No	No	Yes	Yes	Yes	No	No
Performance											
Has noticeable performance impact	Maybe	No	Yes	Yes	Maybe	Maybe	Minimal	Minimal	No	Maybe	Maybe
Can offload backups, DBCCs, reports	Yes	No	Yes	No	No	No	Yes	No	Some	No	No

Comparison of SQL Server Availability Features

	>$1mm	$100k-$500k			$100k-$250k		$50k-$250k		$5k-$100k		
	Replication	Failover Clustered Instances (FCI)	AlwaysOn Availability Groups - Synch	Database Mirroring - Synch	SAN Replication Synch	VM Replication Synch	AlwaysOn Availability Groups - Asynch	Database Mirroring - Asynch	Log Shipping	SAN Replication Asynch	VM Replication Asynch
Recovery Point Objective (Data Loss)											
Data loss possible	Yes	No*	No*	No*	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Typical RPO goal	Zero	Zero	Zero	Zero	Zero	Zero	1 Minute	1 Minute	1 Hour	1 Hour	1 Hour
Recovery Time Objective (Downtime)											
Failover automatic or manual	Automatic	Automatic	Automatic	Automatic	Optional	Optional	Manual	Manual	Manual	Manual	Manual
Can be failed over by the DBA alone	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	No	No
Easy to fail back and forth for patching	Some	Yes	Yes	Yes	No	No	Yes	Yes	Some	No	No
Multiple live writeable servers	Optional	No	No	No	No	No	No	No	No	No	No
Fails over Agent jobs, SQL logins	No	Yes	No	No	Optional	Yes	No	No	No	Optional	Yes
3rd party apps fail over easily	No	Yes	Yes	Some	Yes	Yes	Yes	Some	No	Yes	Yes
Typical RTO goal	Zero	1 Minute	1 Minute	1 Minute	1 Hour	1 Hour	1 Hour	1 Hour	1 Day	1 Day	1 Day
Groups of Databases											
Are the same point in time	No	Yes	No*	No	Yes	Yes	No	No	No	No	No
Fail over together	No	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Hardware & Software Requirements											
SQL Server minimum version supported	Apps vary	Any	2012	2005	Any	Any	2012	2005	Any	Any	Any
SQL Server minimum edition required	Apps vary	Any	Enterprise	Standard	Any	Any	Enterprise	Enterprise	Any	Any	Any
May require application changes	Yes	No	No	No	No	No	No	No	No	No	No
Req storage for multiple copies of DBs	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Requires shared storage (SAN)	No	Yes	No	No	Yes	Yes	No	No	No	Yes	Yes
Requires virtualization	No	No	No	No	No	Yes	No	No	No	No	Yes
Can use local solid state storage	Yes	No	Yes	Yes	No	No	Yes	Yes	Yes	No	No
Performance											
Has noticeable performance impact	Maybe	No	Yes	Yes	Maybe	Maybe	Minimal	Minimal	No	Maybe	Maybe
Can offload backups, DBCCs, reports	Yes	No	Yes	No	No	No	Yes	No	Some	No	No

BRENT OZAR UNLIMITED

SQLServer minimum edition required	Apps vary	Any	Enterprise	Sta
May require application changes	Yes	No	No	
Req storage for multiple copies of DBs	Yes	No	Yes	Y
Requires shared storage (SAN)	No	Yes	No	
Requires virtualization	No	No	No	
Can use local solid state storage	Yes	No	Yes	Y

Your budget is limited.

Don't get SAN storage, switch gear, and a SAN admin unless you have to.

Or do you just want a SAN?

Already have a SAN that performs very well

Plan to move the database between hardware soon

Data too large to fit in internal drives
(which is pretty unlikely these days, except blades)

Otherwise, local solid state is a shortcut to speed,
so your hardware choice should have local drive bays.
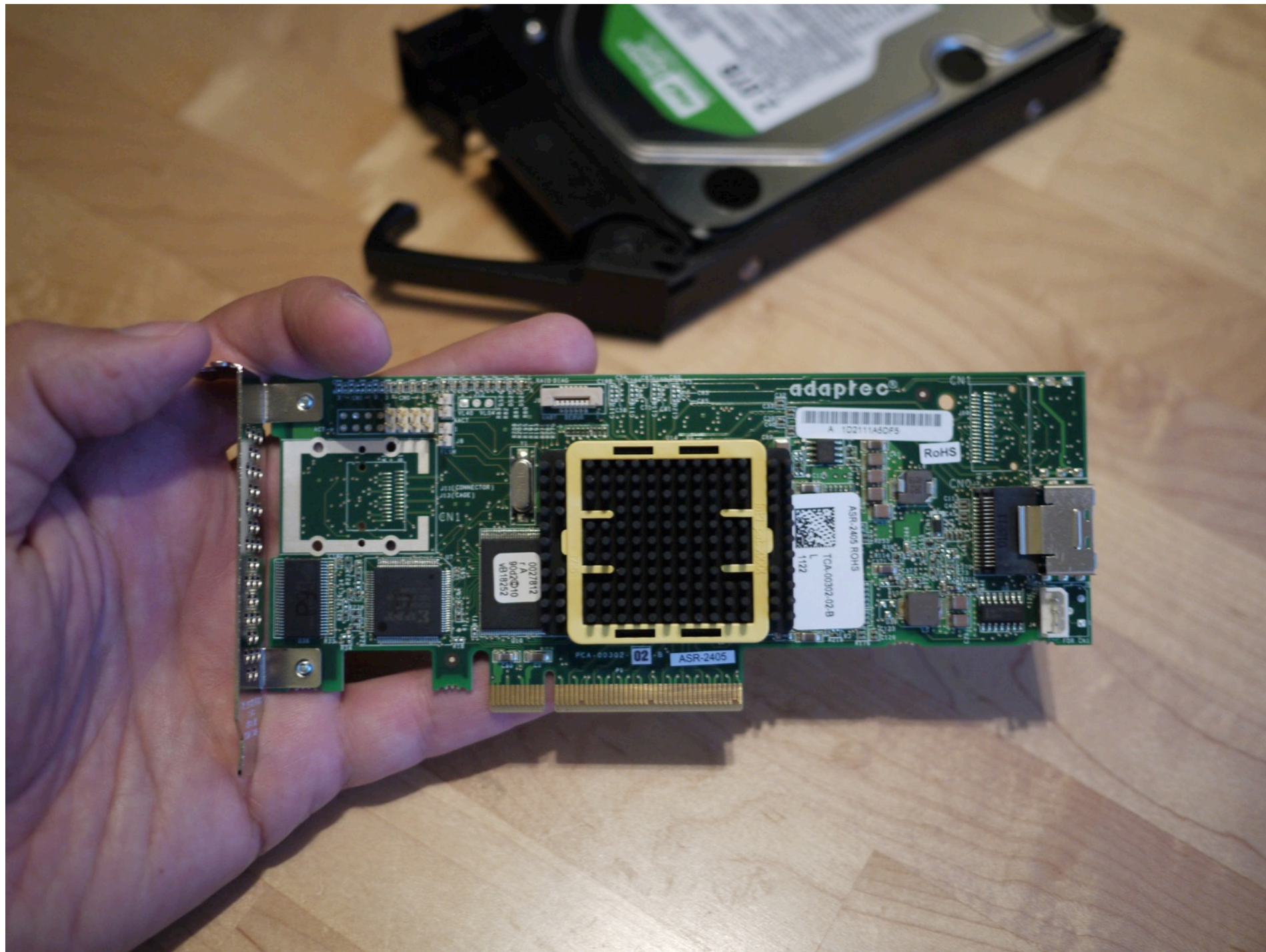
If you need a SAN, what's important?

Simple SAN path
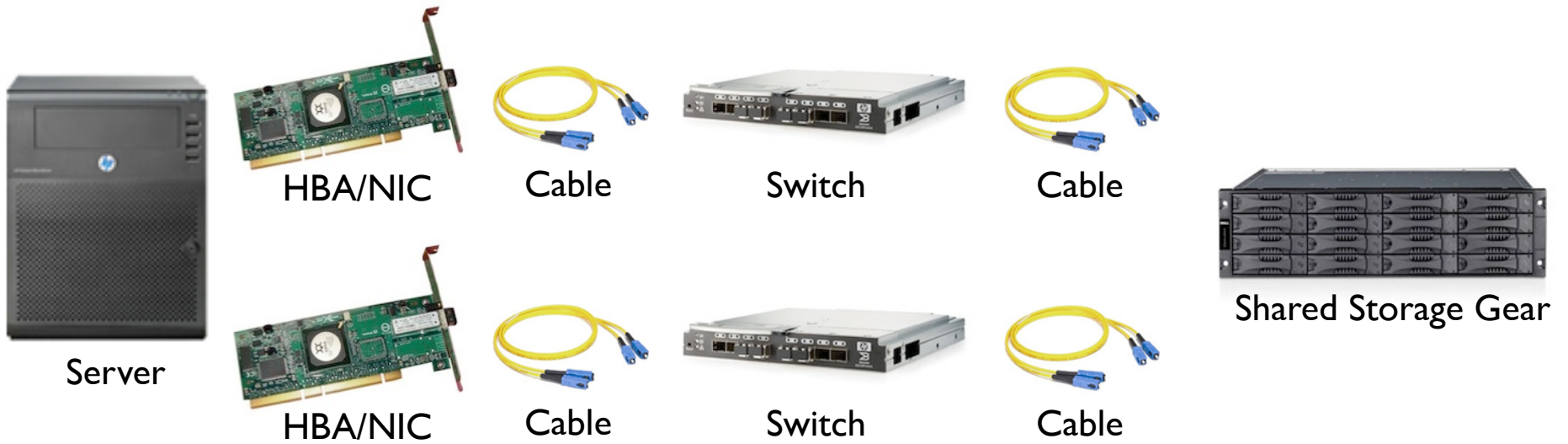

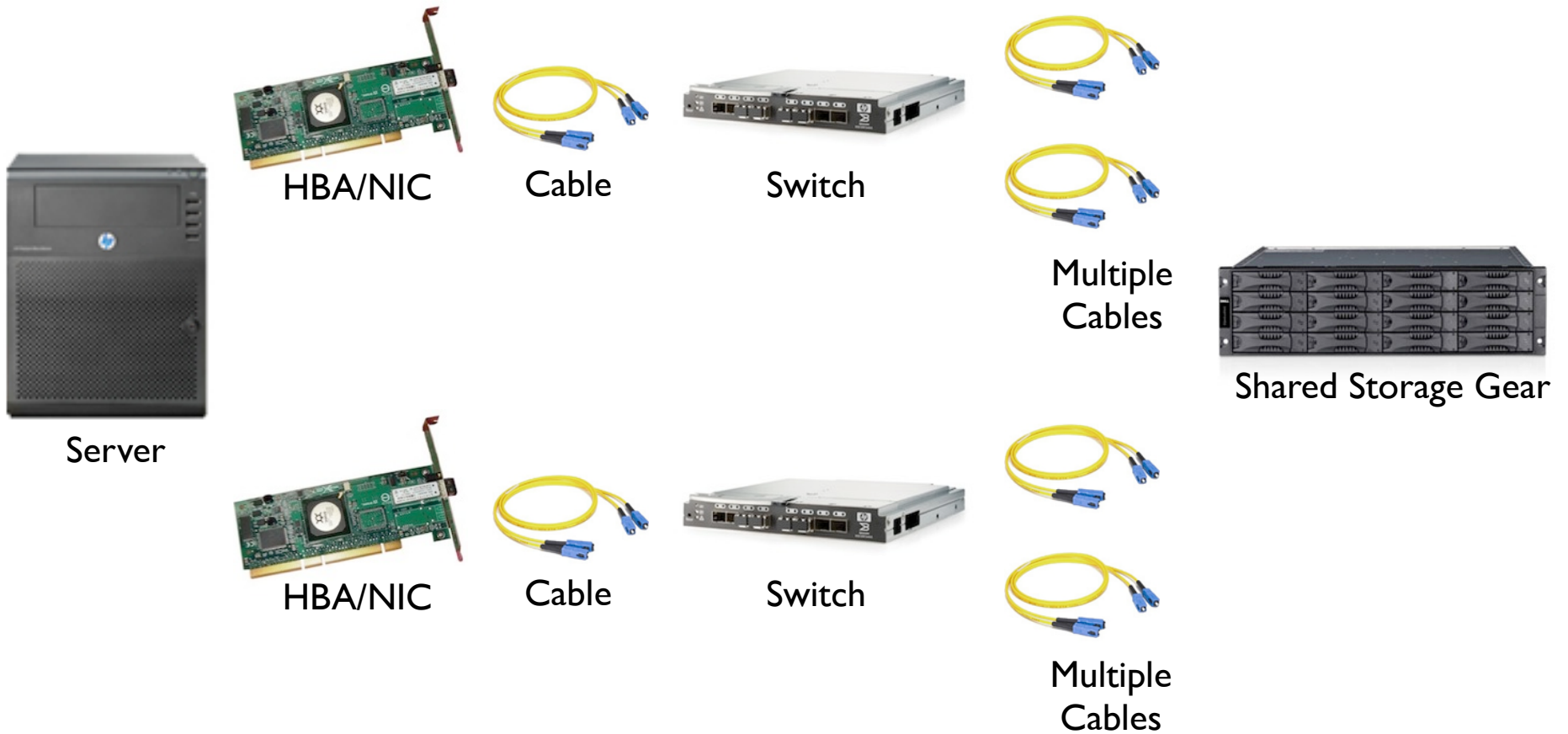Server   HBA/NIC   Cable   Switch   Cable   Shared Storage Gear

Two storage fabrics

Server — HBA/NIC — Cable — Switch — Cable — Shared Storage Gear

HBA/NIC — Cable — Switch — Cable

More paths, more throughput?

Server

HBA/NIC

Cable

Switch

Multiple Cables

Shared Storage Gear

HBA/NIC

Cable

Switch

Multiple Cables

Active/Active Clustering

Two servers, one SAN

Each server runs a separate SQL Server instance

Both instances are live at any given time

Higher hardware utilization

*But one database is only served
by one server at a time*

Active/Active Multipathing

Multiple active paths are actively carrying data,
but they may only be carrying data for one drive letter,
or they may only be active for writes or for reads.

Server

HBA/NIC   Cable   Switch

Multiple Cables

Shared Storage Gear

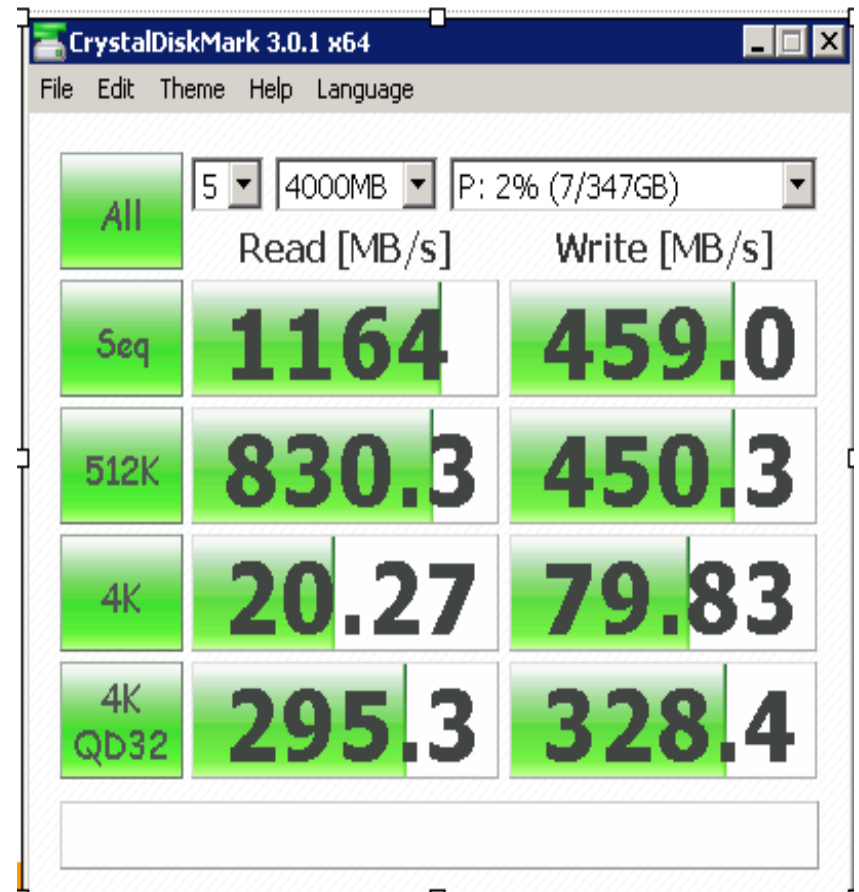HBA/NIC   Cable   Switch

Multiple Cables
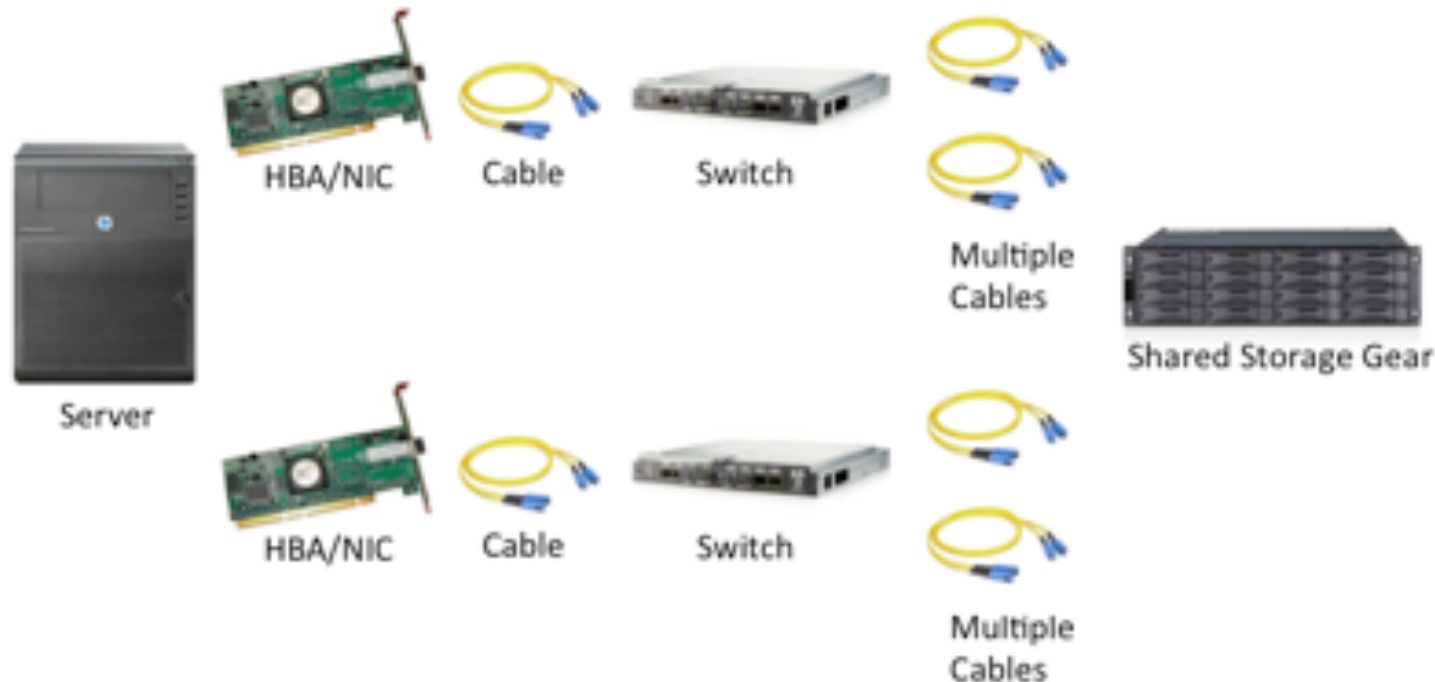
CrystalDiskMark

Free tool with instructions:
BrentOzar.com/go/cdm

Pick 5 tests, 4000MB test file, drive letter

Sequential:
roughly akin to backups, large table scans

4K QD32:
vaguely similar to active OLTP server or TempDB

Bandwidth Reference

	Fastest Time to Transfer 1GB	Theoretical Max Bandwidth	
USB 2.0	17 Seconds	60 MB/Sec	
1 Gb iSCSI	8 Seconds	125 MB/Sec	
4 Gb Fibre Channel	2.4 Seconds	425 MB/Sec	
USB 3.0	1.7 Seconds	600 MB/Sec	
SATA Revision 3.0 (6Gb)	1.3 Seconds	750 MB/Sec	
8 Gb Fibre Channel	1.2 Seconds	850 MB/Sec	
10 Gb iSCSI	800 Milliseconds	1.25 GB/Sec	

From BrentOzar.com/posters

Your goals

Aim for 200MB per second of read throughput,
per core of SQL Server

Example: 8 cores need at least 1,600MB/sec read
throughput

If you quickly and consistently hit a bottleneck lower
than that, check your SAN fabric's throughput

Try using multiple paths, active/active multipathing

Easier way: cache the data in RAM, or ...

Storing the data locally on SSDs

"But SSDs aren't reliable!"

News flash: neither are hard drives.

Mitigations:

- Backups

- DBCCs

- 2 hot spare drives (hey, we've got all those bays in the server chassis, might as well use them)

- Xanax

Dell PowerEdge 720XD

Top version: holds 26 2.5" drives (2 are on the back)

Bottom version: old 3.5" drives, holds 12

HP DL380e G8

Similar, has models for different drive sizes.

You want 2.5" drive bays.

Lots of drive bays for SSDs

Option 1: vendor-approved, enterprise-class drives.

- Vendor supports them, replaces when they break

- Well-tested – but this also means old and slow, because vendors don't test the latest and greatest

- Redonkulously expensive (800GB ~£2,000 each)

Option B: off-the-shelf consumer drives

- You support them when they break
  (but they're dirt cheap, so buy spares)

- Redonkulously fast, cheap (1TB ~£350 each)

Your choice.

	Vendor-supplied 800GB drives	Off-the-shelf 1TB drives
Total cost	£33k	£6k
Usable space, RAID 10	6TB	8TB
Usable space, RAID 5	11TB	14TB
Speed	No specs given	Crazy fast

Controllers count too

2U servers will either have 1 RAID controller or 2

More controllers, more ~~problems~~ throughput

Dell PowerEdge 720: 16 bays, 2 controllers.
Less space, more speed.

Dell PowerEdge 720XD: 26 bays, 1 controller.
More space, less speed.

More details: http://www.brentozar.com/archive/
2013/08/load-testing-solid-state-drives-raid/

The right hardware for Standard Edition

The Standard Standard

2-socket, 4-6 cores (get fastest cores possible)

128GB-192GB memory (SQL will use 64GB-128GB)

2 local SSDs in a mirrored pair for TempDB

Same spec kit whether you're doing cluster, non-clustered with a SAN, or non-clustered with local disks

Example: Dell PowerEdge 720 with 3.5GHz processors and 128GB memory: £9k

On Enterprise, how many CPU sockets do you need?

Do you even need sockets?

	Standard Edition	Enterprise Edition
Physical	2 quad-core processors Licensing: $14,344 Hardware: $10,775 Total: $25,119	2 quad-core processors Licensing: $54,992 Hardware: $23,206 Total: $78,198
Virtual	Guest with 4 cores Licensing: $7,172 Hardware: FREE* Total: Unknown	Host with 2 quad-core processors Licensing: $54,992 Hardware: $23,206 Total: $78,198

Source: BrentOzar.com/go/smallest

Physical servers

Summary			
Sockets	2	4	8
Memory	768GB	3TB	4TB
PCI Express slots	6	9	16
Drive bays	26	10	8
Fastest low-core, low-licensing			
Total speed, GHz	28	82	89
Total cost	$ 80,489	$ 335,430	$ 688,939
Most cores available			
Total speed, GHz	65	168	192
Total cost	$ 192,613	$ 569,547	$ 965,267

# of CPU Sockets	2	4	8
Models	Dell R720		
	HP DL380 G8	HP DL580 G8	HP DL980 G7
Memory	768GB	3TB	4TB
PCI Express slots	6	9	16
Drive bays	26 Dell, 25 HP	10	8
Rack space required	2U	4U	8U
Fastest low-core, low-licensing			
Processor speed, GHz	3.5	3.4	1.86
# of cores	4	6	6
Total speed, GHz	28	82	89
Boost, theoretical		291%	319%
Cache, MB	15	38	18
Cache per core	3.8	6.3	3.0
Model	E5-2637 v2	E7-8893 v2	E7-4807
Hardware cost	$ 25,497	$ 170,454	$ 358,987
SQL Server licensing	$ 54,992	$ 164,976	$ 329,952
Total cost	$ 80,489	$ 335,430	$ 688,939

# of CPU Sockets	2	4	8
Models	Dell R720		
	HP DL380 G8	HP DL580 G8	HP DL980 G7
Memory	768GB	3TB	4TB
PCI Express slots	6	9	16
Drive bays	26 Dell, 25 HP	10	8
Rack space required	2U	4U	8U
Most cores available			
Processor speed, GHz	2.7	2.8	2.4
# of cores	12	15	10
Total speed, GHz	65	168	192
Boost, theoretical		259%	296%
Cache, MB	30	38	30
Cache per core	2.5	2.5	3.0
Model	E5-2697 v2	E7-4890 v2	E7-4870
Hardware cost	$ 27,637	$ 157,107	$ 415,347
SQL Server licensing	$ 164,976	$ 412,440	$ 549,920
Total cost	$ 192,613	$ 569,547	$ 965,267

Physical servers

Summary			
Sockets	2	4	8
Memory	768GB	3TB	4TB
PCI Express slots	6	9	16
Drive bays	26	10	8
Fastest low-core, low-licensing			
Total speed, GHz	28	82	89
Total cost	$ 80,489	$ 335,430	$ 688,939
Most cores available			
Total speed, GHz	65	168	192
Total cost	$ 192,613	$ 569,547	$ 965,267

Boiling it down

Need <768GB RAM and <82GHz CPU? 2-socket.
The 25-26 local drive bays give you great options.

Need 1-3TB RAM and 8-10 PCI slots? 4-socket.
The March CPU updates beat 8-socket boxes.

Need over 3TB RAM or over 10 PCI slots?
8-socket, but start making plans to split this monster.

Recap

The top 5 things we covered

1. For Standard Edition, think virtual first.

2. If your RPO/RTO doesn't require shared storage, try local SSDs.

3. If you have to have a SAN, pathing is incredibly important. Get lots of fast paths.

4. CrystalDiskMark test for 200MB/sec/core.

5. In mid-2014, 2-sockets are an amazing deal.

Learn more: BrentOzar.com/go/faster