



BRENT OZAR
UNLIMITED®

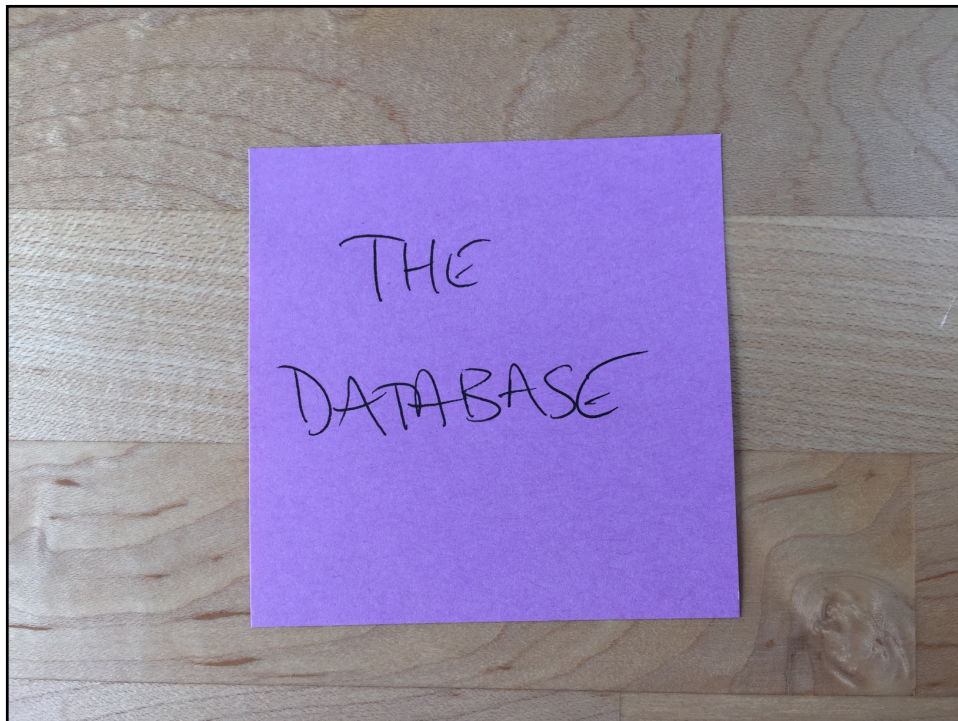
How to Choose the Right Scale-Out Tech

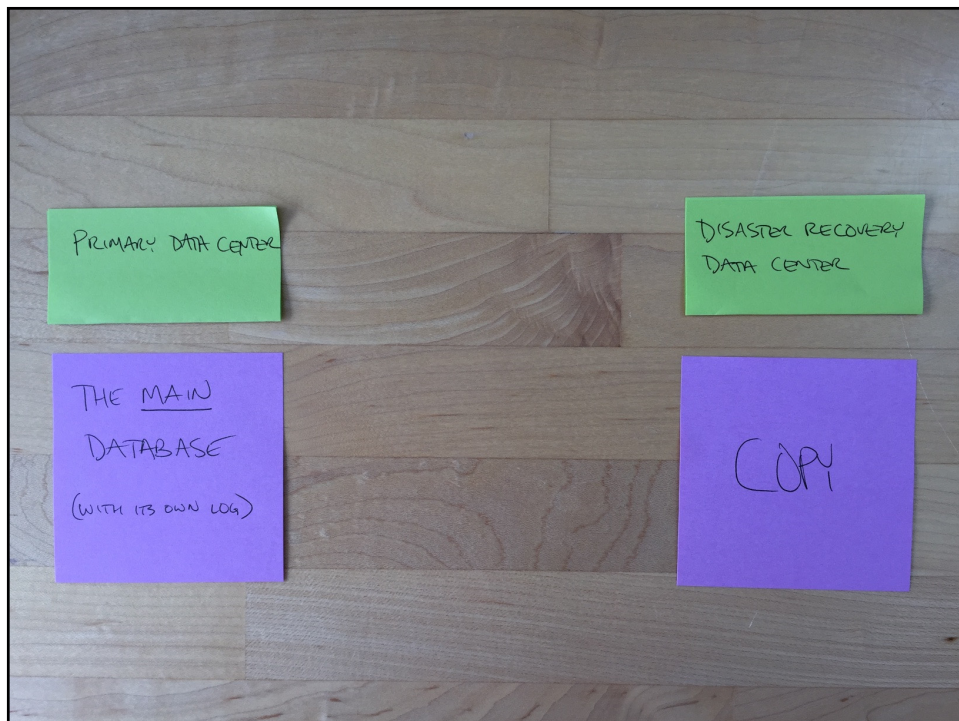
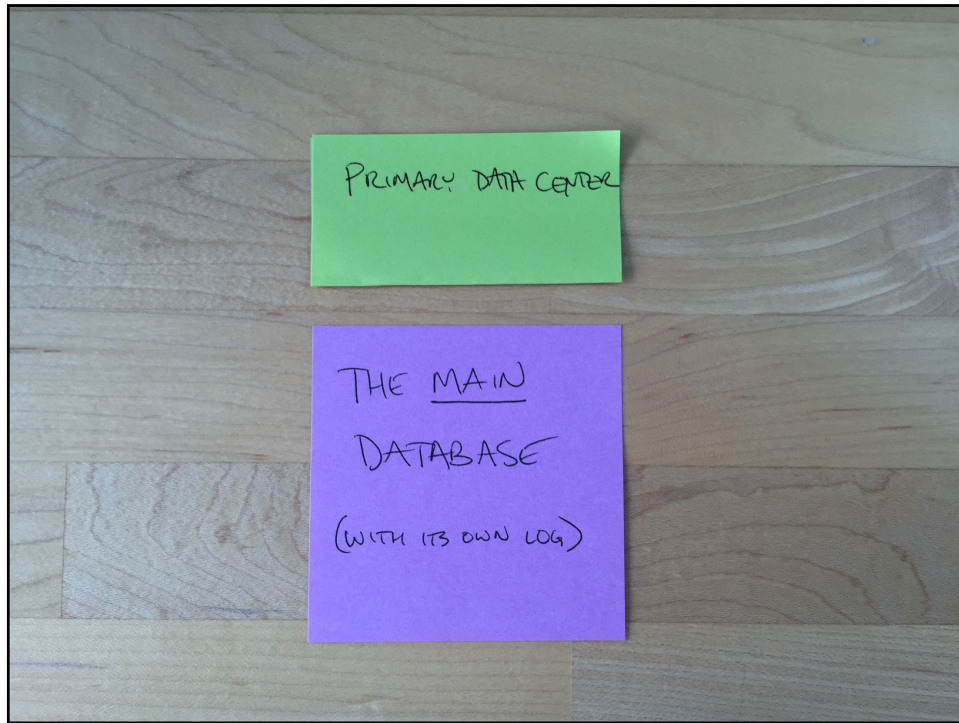
Matching up use cases to scale-out technology is hard.

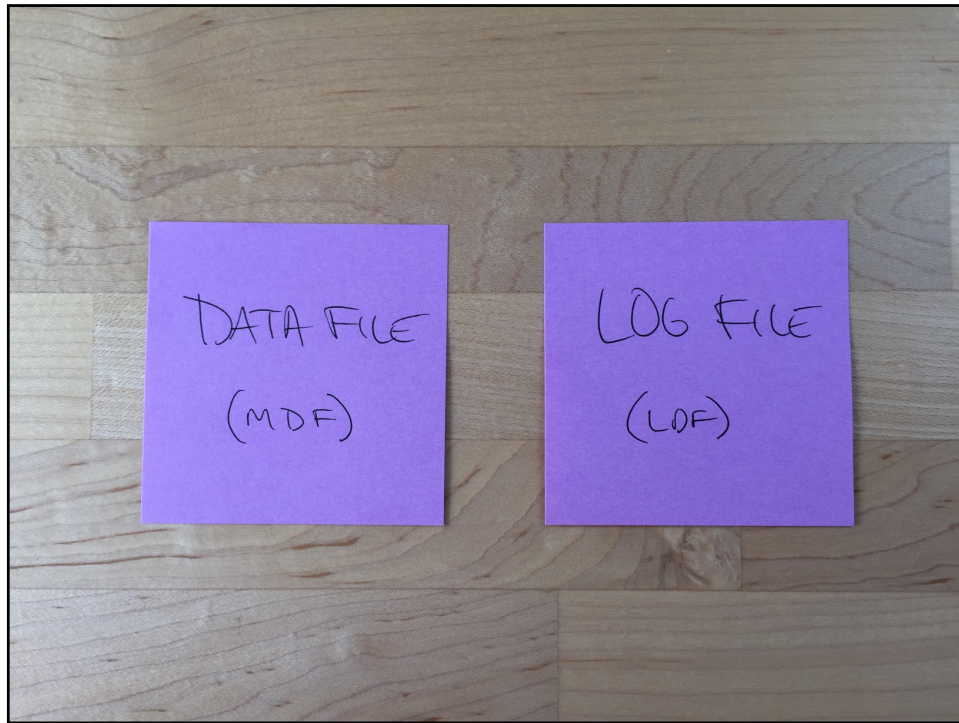
- Many concurrent readers of small data (product inventory): Hekaton
- Many concurrent writers, but writing to different groups of records: sharding bidirectional transactional replication, merge replication
- Many readers of data that is constantly updated, but readers can deal with data that's a few minutes old: AlwaysOn AG readable secondaries, transactional replication
- A few readers that write horrifically bad queries that tend to block others, and they only need a snapshot of the data as of 9AM this morning: log shipping, database mirroring snapshots

Session agenda

- Explain why scale-out starts with separation for HA/DR
- Lay out common data groupings with examples
- Designing your own data groupings
- Picking the right scale-out tech for your data groupings

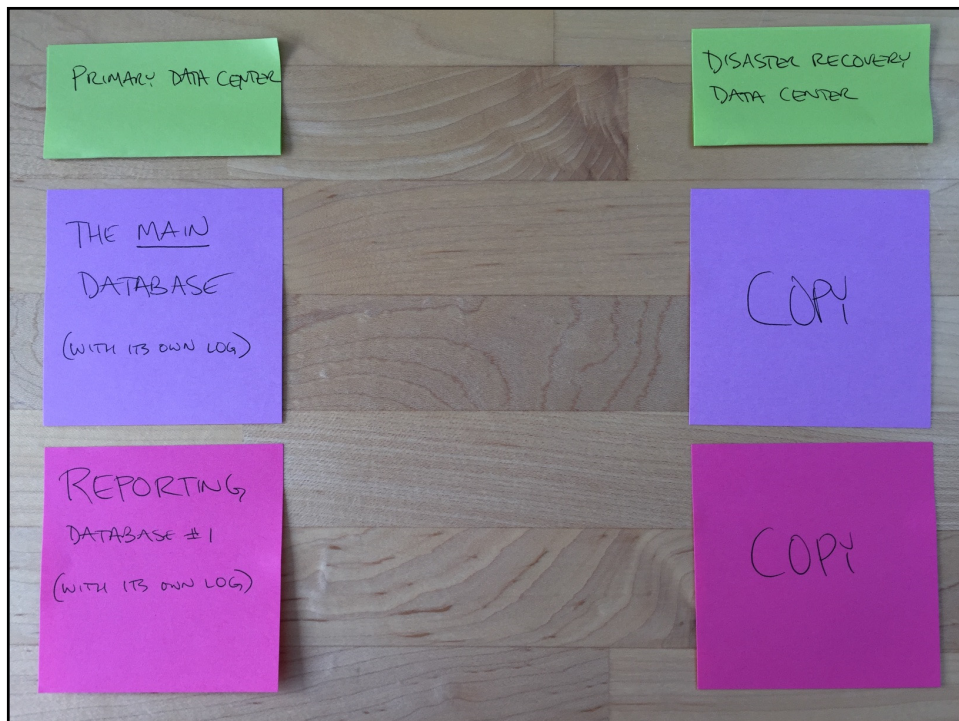
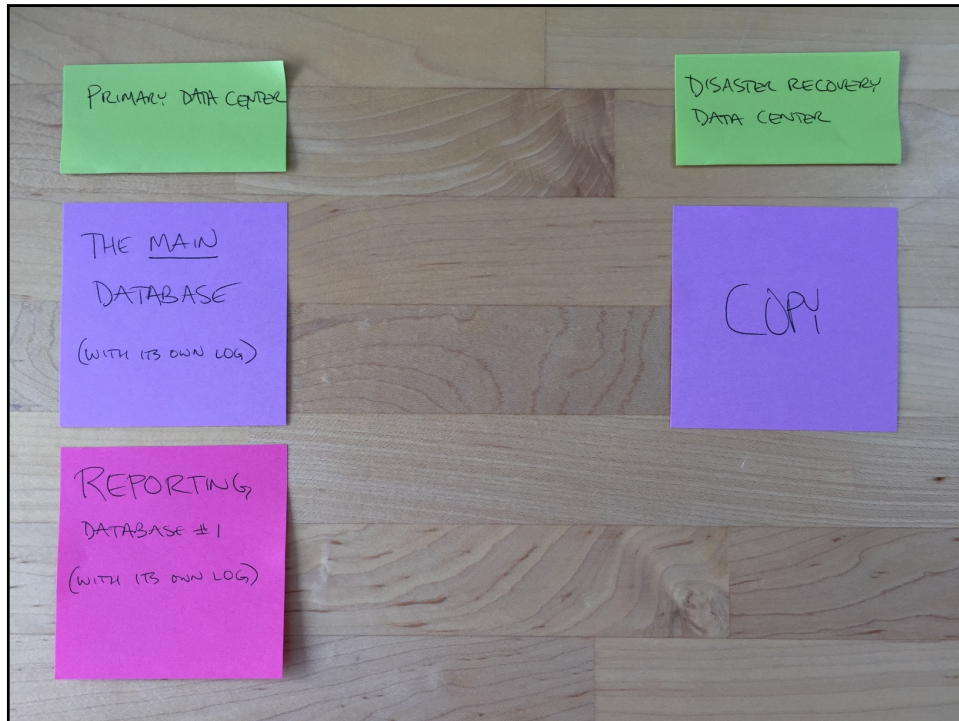


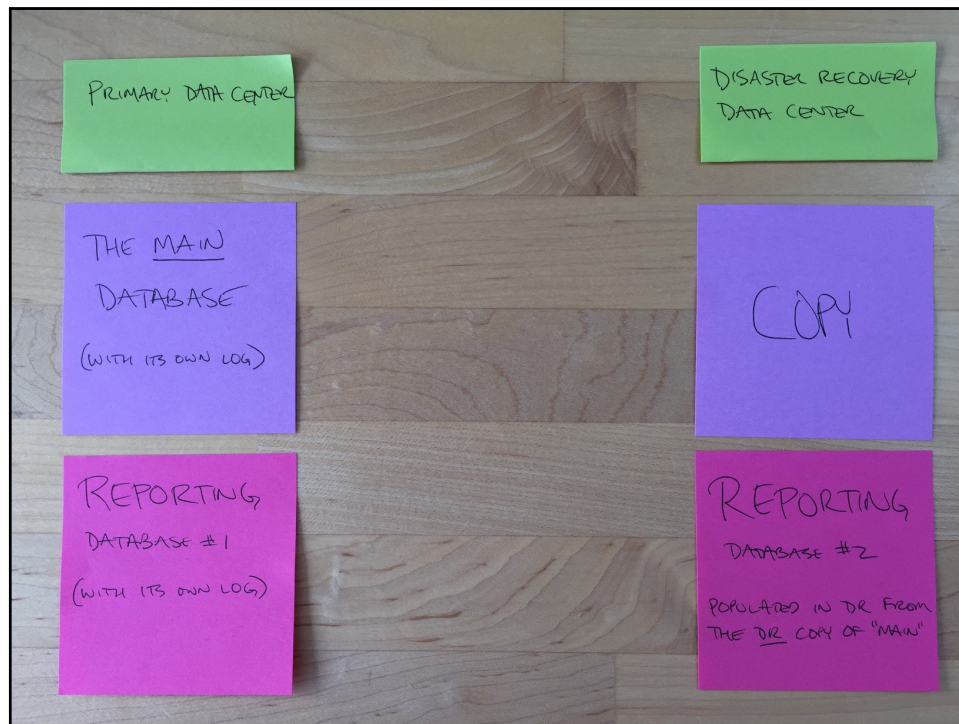




The problems with one log file

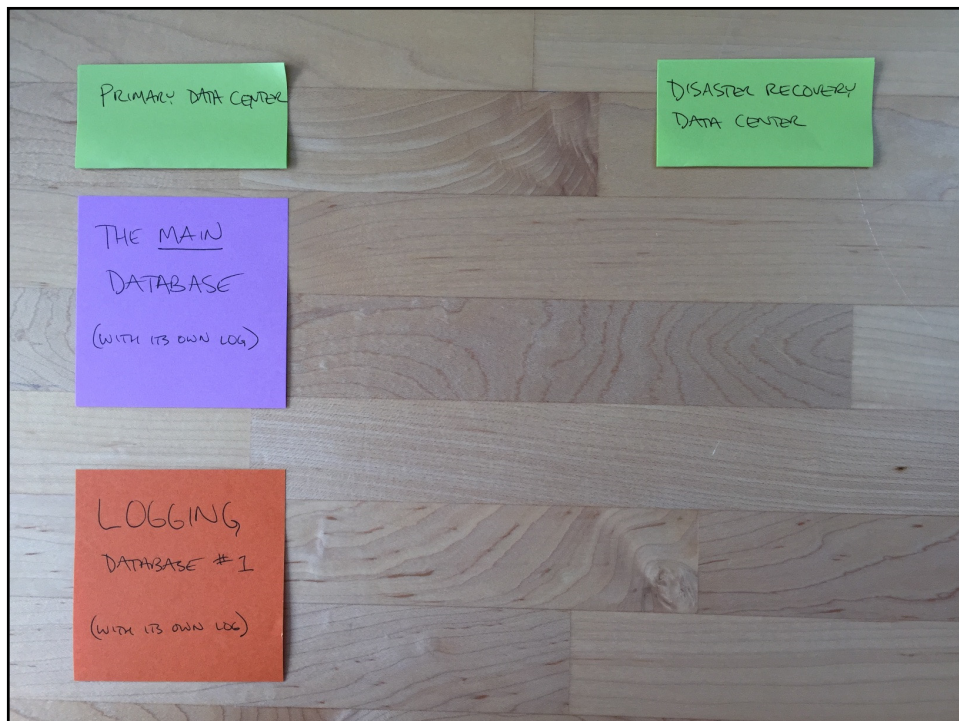
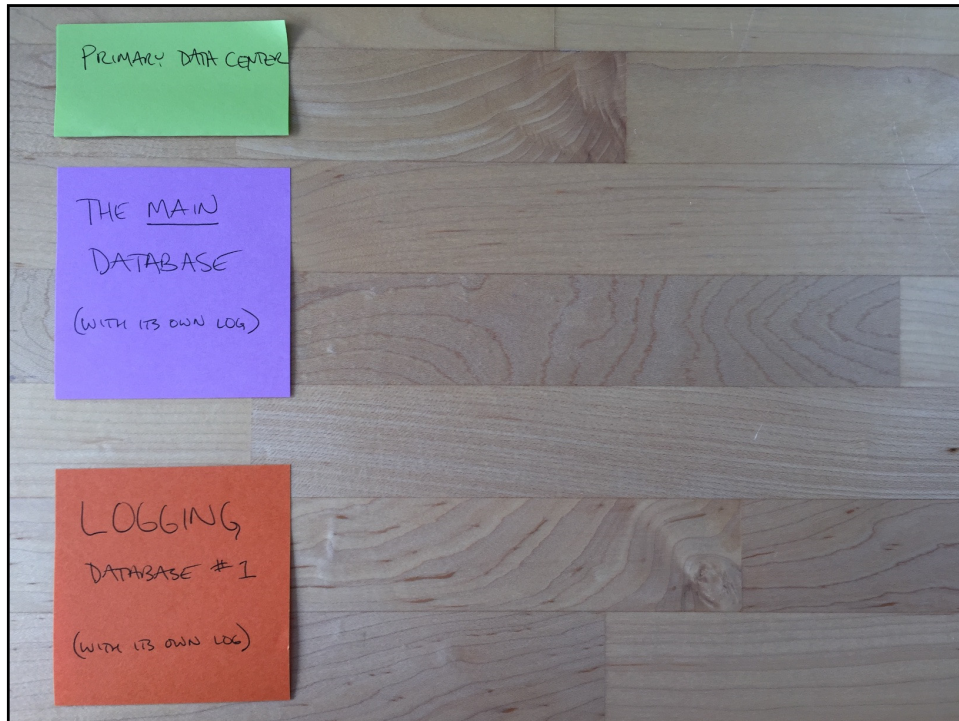
- We can't prioritize which data needs to get over to DR first
- Almost everything we do in the database gets logged (including maintenance tasks like rebuilding indexes)
- We can't set limits on how far behind log copies get: either we're synchronous, or we're asynchronous
- We can't really specify a low-value table to be non-logged*

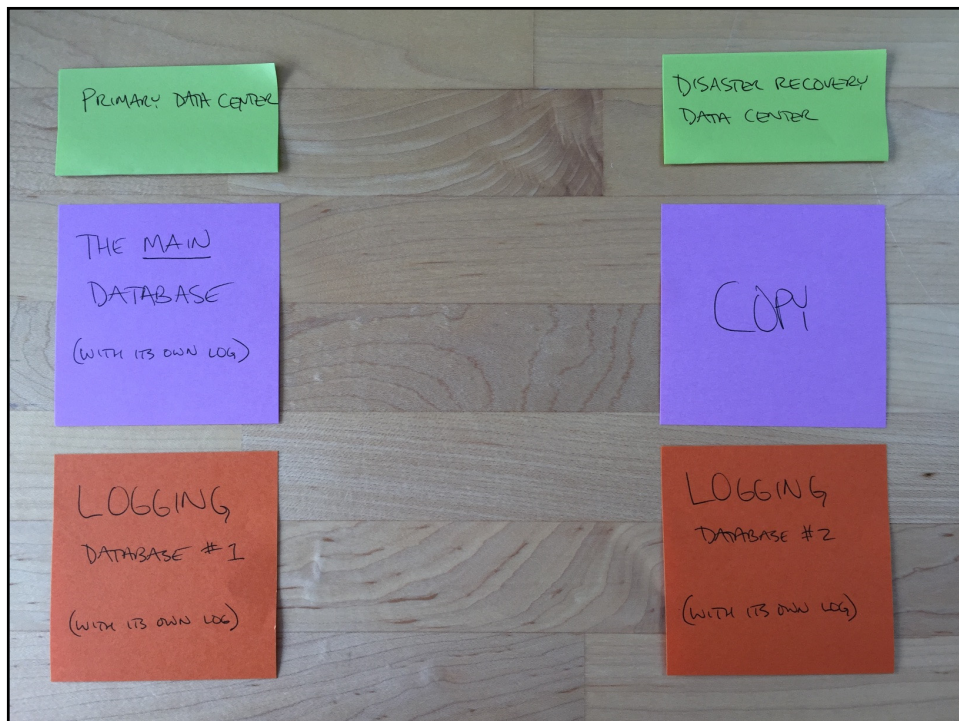
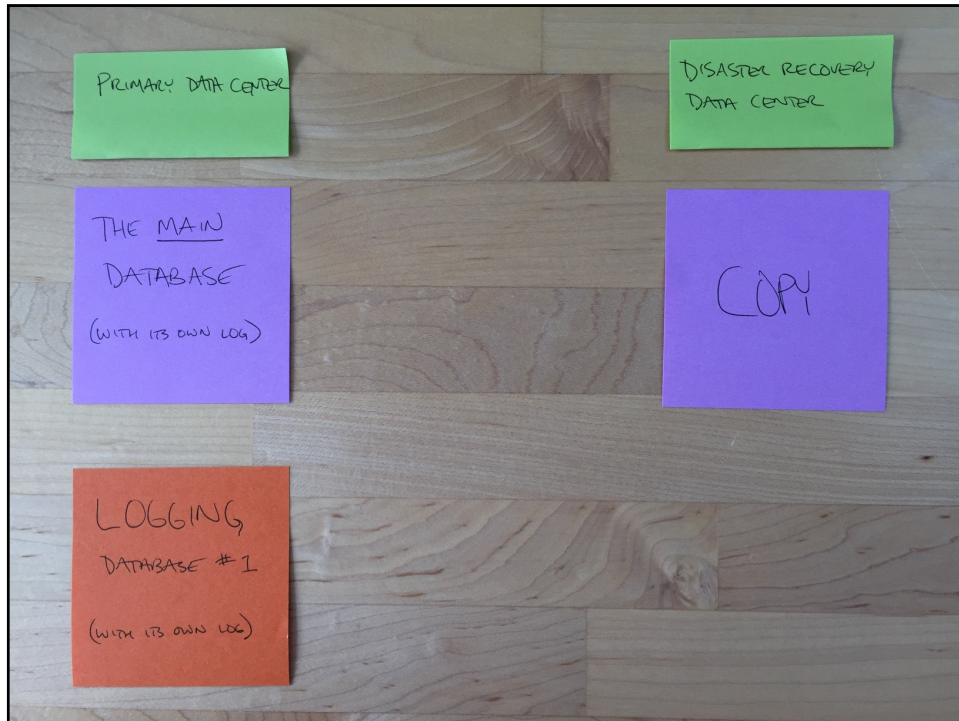




Considerations for splitting databases

- Could this table be easily recreated from other tables?
- Could the app be up without this table?
- Does this table have to be exactly in sync with other tables?
- If the table already exists, can I swap in a synonym or a view?





Common data groupings

Common data groupings

- **Main:** our transactional data. The app won't run without this.
- **Logging:** low-value, heavy-write, rarely-read. App would still work with an empty version of these tables.
- **Reporting:** aggregated data that's easily recreated from other transactional data. App could be down without it, or could live with a stale copy of this data.
- **Files/text/XML:** voluminous data. We do point lookups of individual rows or groups of rows based on the OLTP data, but we don't directly join this data to the above data.
- **Other examples:** session state, product inventory, staging, scratch

For each grouping:

- **Backup and recovery**
 - What's the acceptable data loss and downtime?
 - How close does this have to be to other data groups?
- **Change rate**
- **Geographic or legal restrictions**
- **Consumers – who's querying the data, from where?**
- **Type of data, natural home**
 - Is this data best served from SQL Server, or another platform
- **Performance bottleneck:**
 - Do you need to scale this group out?
 - What's the bottleneck requiring scale-out?



19

© SQLintersection. All rights reserved.
<http://www.SQLintersection.com>

Wikipedia*

** No, of course Wikipedia doesn't run on Microsoft SQL Server, but let's pretend.*

	Main OLTP (Users/Security)	Article Contents	Web Clicks	Web Click Analysis
RPO (Data Loss) for HA scenarios	0	1 day	1 day	(Unlimited, can reload from clicks)
RTO (Downtime) for HA scenarios	1 minute	1 minute	1 minute	1 day
Data producers/consume rs location	Web/app servers	Web/app servers	Web/app servers	SSIS servers
How it's queried	Small point queries, insert/update/del etes	99% reads	99% writes	Loaded nightly in batches, then read with reports
Performance bottleneck	Concurrent reads/writes	Lots of read connections	Transaction log writes	Storage reads b/c too large for RAM
Change rate	.01% per day	.01% per day	1% per day	1% per day
Type of data, natural home	Relational	Text, XML	Relational, cubes	Relational, cubes

Amazon*

** But first, a few words about "scale" and "real time data."*



amazon Prime Search All : bacon mints Go Your Account Prime Wish List

Departments Fire & Kindle For Brent Deals Gift Cards Help

Grocery & Gourmet Food Best Sellers Beverages & Coffee Snacks Breakfast Candy & Chocolate Gifts Baking Gluten Free Special Offers

Back to search results for "bacon mints"

Bacon Candy Sampler Gift Pack (4pc Set) - Bacon Mints, Jelly Beans, Gumballs & Lollipop
by Bacon Addicts
★★★★★ 2 customer reviews
Price: \$24.95 (\$6.24 / Item) & FREE SHIPPING
Only 6 left in stock.
Sold by Bacon Addicts.

OUR MOST POPULAR BACON NOVELTY CANDIES IN ONE PACK!
Bacon Addicts Bundle Pack Includes: Jumbo Bacon Swirl Lollipop, Bacon Jelly Beans, Bacon Flavored Mints & Bacon Flavored Gumballs. Each in collectible gift tin (except lollipop).
MAKES A GREAT GIFT All items in the pack are kosher & vegetarian safe!
FAST & FREE US SHIPPING!!

2 new from \$20.00

Qty: 1
\$24.95 + Free Shipping
In Stock. Sold by Bacon Addicts
Add to Cart
Turn on 1-Click ordering
Add to Wish List

Other Sellers on Amazon
\$20.00 (\$5.00 / Item)
+ \$4.99 shipping
Sold by: Meat Maniac
2 new from \$20.00
Have one to sell? Sell on Amazon

SQL intersection 23

© SQLIntersection. All rights reserved.
http://www.SQLIntersection.com

	Main OLTP (Orders)	Product Catalog	Inventory	Shopping carts
RPO (Data Loss) for HA scenarios	0	1 day (can regenerate from source, partners)	?	Unlimited
RTO (Downtime) for HA scenarios	0	0	1 hour (assume all items are in stock)	0
Data producers/consumers location	Web/app servers	Produced by ETL, consumed by web servers	Web/app servers	Web/app servers
How it's queried	Small point queries, inserts only	99.999% reads	Small point queries, readers/updaters	Small point queries, readers/inserts
Performance bottleneck	Transaction log writes	Many concurrent readers	Many concurrent readers	Transaction log writes
Change rate	1% per day	1% per day	100% per day	100% per day
Type of data, natural home	Relational	XML	Relational or XML	Relational or XML

Hotel Survey Company

Hotel survey company

- Hotel managers design the guest feedback questions to ask
- Hotels export their list of guests every night and we import it
- We email the guests a link to their customized survey
- Guest takes the survey
- We work on the comments (translation, sentiment analysis)
- We alert the hotel managers on truly bad comments that need help
- Hotel managers can analyze their survey feedback over time

	Main OLTP (The Database)	
RPO (Data Loss) for HA scenarios	1 minute	
RTO (Downtime) for HA scenarios	1 minute	
Data producers/consumers location	Web/app servers, ETL processes	
How it's queried	Small point queries, insert/update/deletes, plus analytical reports	
Performance bottleneck	Many: locking, reads, transaction log writes	
Change rate	1% per day	
Type of data, natural home	Everything!	

27

© SQLIntersection. All rights reserved.
<http://www.SQLIntersection.com>

	Main OLTP (Hotels, Guests, Surveys)	Survey Servers (Emails, Questions)	Survey Answers	Hotel Reports
RPO (Data Loss) for HA scenarios	1 minute	(Unlimited, can reload from OLTP)	1 minute	(Unlimited, can reload from OLTP answers)
RTO (Downtime) for HA scenarios	1 day (as long as people can still take surveys)	1 minute	1 minute	1 day
Data producers/consumers location	ETL process: hotels sent text files in, plus answers	Web/app servers around the world	Web/app servers around the world	SSIS servers
How it's queried	Small point queries, insert/update/deletes	99% reads	99% writes, plus reads to push data to OLTP	Loaded nightly in batches, then read with reports
Performance bottleneck	None	Lots of read connections in bursts	Transaction log writes	Storage reads b/c too large for RAM
Change rate	1% per day	100% per day	100% per day	10% per day
Type of data, natural home	Relational	Relational	Relational and text comments	Relational, cubes

Hotel survey company: scaling out = separating.

- For us, after simply splitting out the data and building some ETL processes, we didn't need database server scale-out technologies.
- But how would you do this today with AWS and Azure services?
- Sending a survey notification, answering a survey, or translating a survey could really be done off a queue service instead.



29

© SQLintersection. All rights reserved.
<http://www.SQLintersection.com>

Typical Multi-Purpose Web App with Reports

	Main OLTP	Logging	Reporting (DW)	Staging
RPO (Data Loss) for HA scenarios	0	1 day	(Unlimited, can regenerate from source)	(Unlimited, can restart the nightly process)
RTO (Downtime) for HA scenarios	1 minute	1 minute	1 hour	1 minute
Data producers/consumers location	Web/app servers	Web/app servers	End users nationwide w/Excel, Access	SSIS servers
How it's queried	Small point queries, insert/update/deletes	99% inserts, rarely queried	Loaded in nightly batches, then queried all day	Huge loads of writes at night
Performance bottleneck	Transaction log writes	Transaction log writes	Reading data from disk for reports	Storage writes/reads, CPU
Change rate	1% per day	1% per day	10% per day	100% per day
Type of data, natural home	Relational	XML	Relational, cubes	Relational

How this impacts tuning and scale-out technologies

- Many concurrent readers of small data (product inventory): Hekaton
- Many concurrent writers, but writing to different groups of records: sharding bidirectional transactional replication, merge replication
- Many readers of data that is constantly updated, but readers can deal with data that's a few minutes old: AlwaysOn AG readable secondaries, transactional replication
- A few readers that write horrifically bad queries that tend to block others, and they only need a snapshot of the data as of 9AM this morning: log shipping, database mirroring snapshots

Separate before you scale.

**Scaling out is seriously hard.
Don't scale out more data than you have to.**



33

© SQLIntersection. All rights reserved.
<http://www.SQLIntersection.com>

It's your turn.

	Main OLTP (Your current DB)			
RPO (Data Loss) for HA scenarios				
RTO (Downtime) for HA scenarios				
Data producers/consumers location				
How it's queried				
Performance bottleneck	<i>I don't expect you to have these answers yet, and as you separate the data, the answers may change.</i>			
Change rate				
Type of data, natural home				

	Main OLTP (Your current DB)			
RPO (Data Loss) for HA scenarios				
RTO (Downtime) for HA scenarios				
Data producers/consumers location				
How it's queried				
Performance bottleneck				
Change rate				
Type of data, natural home				

Choosing the right scale-out tech for a data group

Wikipedia*

* No, of course Wikipedia doesn't run on Microsoft SQL Server, but let's pretend.



	Main OLTP (Users/Security)	Article Contents	Web Clicks	Web Click Analysis
RPO (Data Loss) for HA scenarios	0	1 day	1 day	(Unlimited, can reload from clicks)
RTO (Downtime) for HA scenarios	1 minute	1 minute	1 minute	1 day
Data producers/consume rs location	Web/app servers	Web/app servers	Web/app servers	SSIS servers
How it's queried	Small point queries, insert/update/del etes	99% reads	99% writes	Loaded nightly in batches, then read with reports
Performance bottleneck	Concurrent reads/writes	Lots of read connections	Transaction log writes	Storage reads b/c too large for RAM
Change rate	.01% per day	.01% per day	1% per day	1% per day
Type of data, natural home	Relational	Text, XML	Relational, cubes	Relational, cubes

	Main OLTP (Users/Security)	Users/Security details - Query volume isn't all that high, but we're running into locking/blocking - Writes alone could easily be handled by 1 database server, and reads could be handled by 2-3 - We just need to split out writers and readers (and we already tried RCSI)
RPO (Data Loss) for HA scenarios	0	
RTO (Downtime) for HA scenarios	1 minute	
Data producers/consume rs location	Web/app servers	
How it's queried	Small point queries, insert/update/del etes	
Performance bottleneck	Concurrent reads/writes	
Change rate	.01% per day	
Type of data, natural home	Relational	

	Article Contents	Article Contents details ▪ When querying the data for display on web pages, the data can be stale (a few minutes old, even hours) ▪ Even when the primary copy of the data fails, the readable copies need to stay online for read-only web pages ▪ A couple readable servers won't be enough: we may need dozens to handle reads
RPO (Data Loss) for HA scenarios	1 day	
RTO (Downtime) for HA scenarios	1 minute	
Data producers/consumers location	Web/app servers	
How it's queried	99% reads	
Performance bottleneck	Lots of read connections	
Change rate	.01% per day	
Type of data, natural home	Text, XML	

44

© SQLIntersection. All rights reserved.
<http://www.SQLIntersection.com>

	Web Clicks	Web Clicks details ▪ Clicks: mostly insert-only ▪ Also read nightly by ETL process that builds analytics cubes for the Web Click Analysis column ▪ One server alone can't handle the write workloads
RPO (Data Loss) for HA scenarios	1 day	
RTO (Downtime) for HA scenarios	1 minute	
Data producers/consumers location	Web/app servers	
How it's queried	99% writes	
Performance bottleneck	Transaction log writes	
Change rate	1% per day	
Type of data, natural home	Relational, cubes	

47

© SQLIntersection. All rights reserved.
<http://www.SQLIntersection.com>

Wait – why are we leaving SQL Server?

When you've separated your data out into groups, the easy fix is isolating the problem group onto its own database server.

You can isolate it with minimal changes to code or connection libraries.

But when one scaled-up server still isn't enough for it, and you split the data across multiple servers, you may need to:

- **Re-combine the data for reporting**
- **Build conflict resolution processes**
- **Change your application to connect to the right server**

And if you're doing all that, it's a great time to think about the data's home.



49

© SQLintersection. All rights reserved.
<http://www.SQLintersection.com>

	Web Click Analysis	Analysis details ▪ Analytical data can be independent from the source data format ▪ Well-known reporting front ends: Excel, SSRS, BusinessObjects ▪ Great fit for relational or cubes (but we still may need to scale out)
RPO (Data Loss) for HA scenarios	(Unlimited, can reload from clicks)	
RTO (Downtime) for HA scenarios	1 day	
Data producers/consumers location	SSIS servers	
How it's queried	Loaded nightly in batches, then read with reports	
Performance bottleneck	Storage reads b/c too large for RAM	
Change rate	1% per day	
Type of data, natural home	Relational, cubes	

51

© SQLintersection. All rights reserved.
<http://www.SQLintersection.com>

Amazon*

** But first, a few words about “scale” and “real time data.”*

	Main OLTP (Orders)	Product Catalog	Inventory	Shopping carts
RPO (Data Loss) for HA scenarios	0	1 day (can regenerate from source, partners)	?	Unlimited
RTO (Downtime) for HA scenarios	0	0	1 hour (assume all items are in stock)	0
Data producers/consumers location	Web/app servers	Produced by ETL, consumed by web servers	Web/app servers	Web/app servers
How it's queried	Small point queries, inserts only	99.999% reads	Small point queries, readers/updaters	Small point queries, readers/inserts
Performance bottleneck	Transaction log writes	Many concurrent readers	Many concurrent readers	Transaction log writes
Change rate	1% per day	1% per day	100% per day	100% per day
Type of data, natural home	Relational	XML	Relational or XML	Relational or XML



	Shopping carts	
RPO (Data Loss) for HA scenarios	Unlimited	Shopping Cart details ▪ dbo.Cart with: □ CustomerID INT □ CartContents XML ▪ <cart> □ <customerid>12346</customerid> □ <item> □ <productid>7891350</productid> □ <quantity>1</quantity> □ </item> □ <item> □ <productid>735616</productid> □ <quantity>10</quantity> □ </item> ▪ </cart>
RTO (Downtime) for HA scenarios	0	
Data producers/consumers location	Web/app servers	
How it's queried	Small point queries, readers/inserts	
Performance bottleneck	Transaction log writes	
Change rate	100% per day	
Type of data, natural home	Relational or XML	

56

© SQLintersection. All rights reserved.
<http://www.SQLintersection.com>

	Inventory	
RPO (Data Loss) for HA scenarios	?	Inventory: let's pretend. ▪ All our order data resides on a single SQL Server, not sharded out. ▪ We don't need pinpoint accuracy: if we're off by a few items, that's OK. ▪ Our developers aren't smart enough to implement caching, and we have to query the database every time.
RTO (Downtime) for HA scenarios	1 hour (assume all items are in stock)	
Data producers/consumers location	Web/app servers	
How it's queried	Small point queries, readers/updaters	
Performance bottleneck	Many concurrent readers	
Change rate	100% per day	
Type of data, natural home	Relational or XML	

60

© SQLintersection. All rights reserved.
<http://www.SQLintersection.com>



Recapping how to decide

Session agenda

- Lay out your app's data groupings (ideally before you run into scaling problems, because these changes can take time)
- Set RPO/RTO goals for each data grouping
- Define the performance bottleneck for the grouping
- Separate and scale up first: don't scale out anything you don't have to
- Pick the right home and scale-out method for an individual group, not a database full of unrelated objects