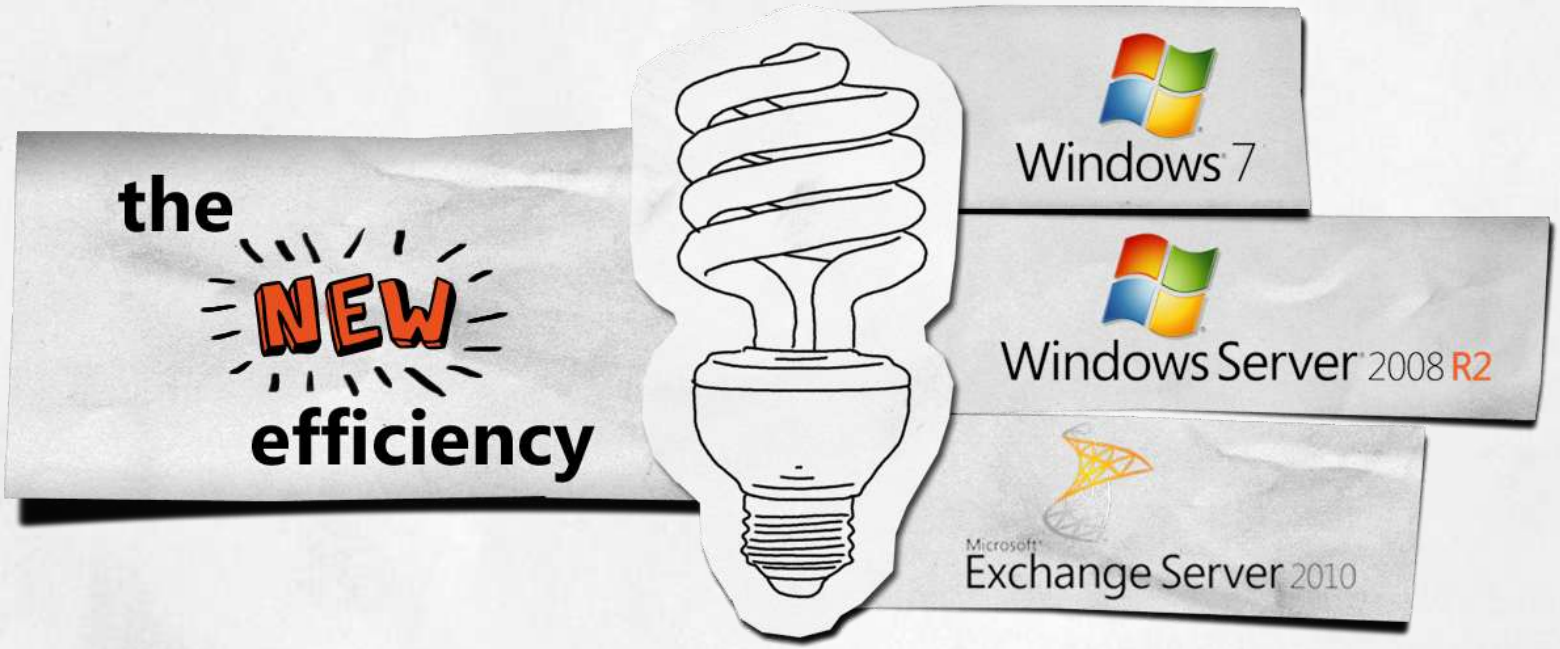


Exchange 2010 High Availability

ANGI LIVERMORE
PRINCIPAL TECHNOLOGY SPECIALIST
MICROSOFT CORPORATION

Why is E-mail Availability Important?

E-mail is the nucleus of business communication

As e-mail becomes more business critical:

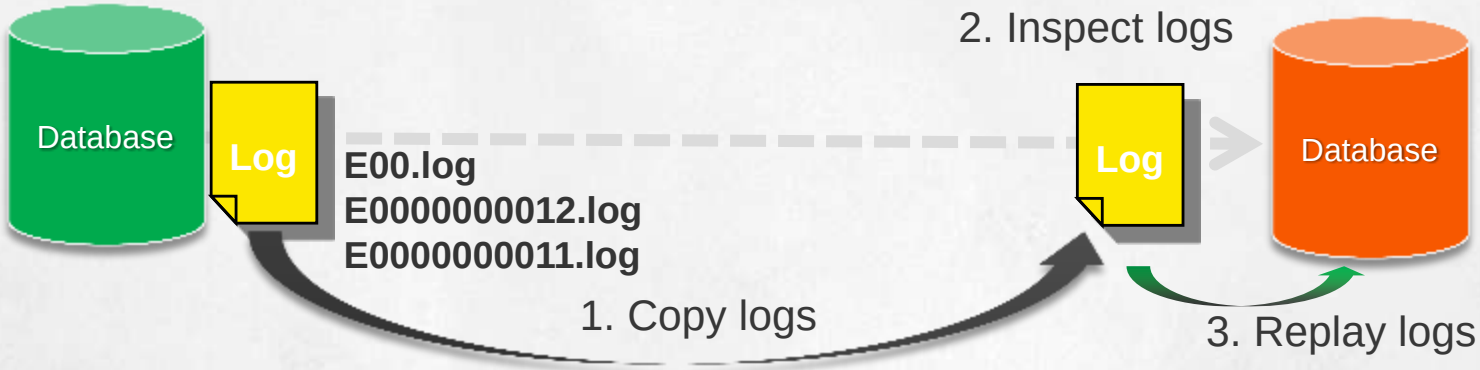
- Time loss after a failure is measured in seconds
- Data loss after a failure needs to be close to zero

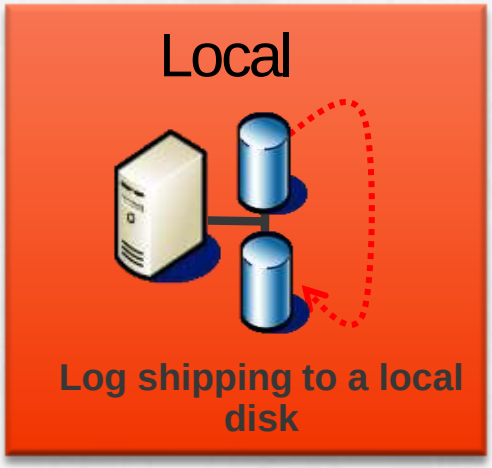
Business users report that they currently spend 19% of their work day, or close to 2 hours/day, on e-mail.

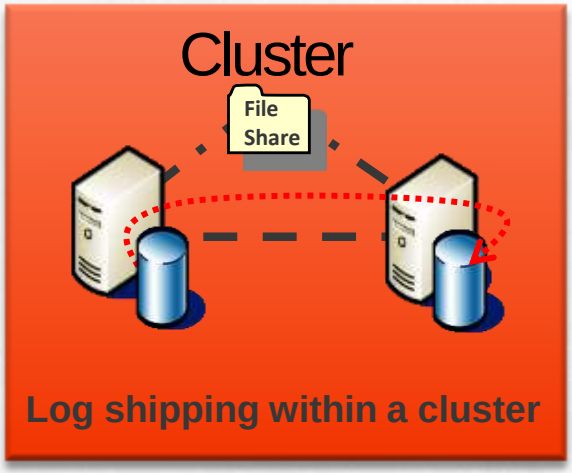
Radicati, 2007

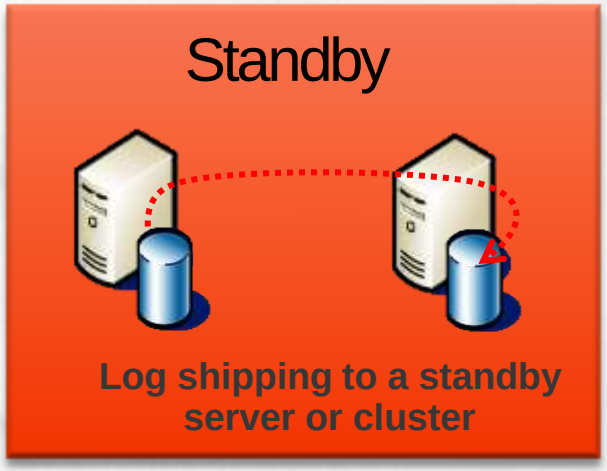
Exchange Server 2007

Continuous Replication

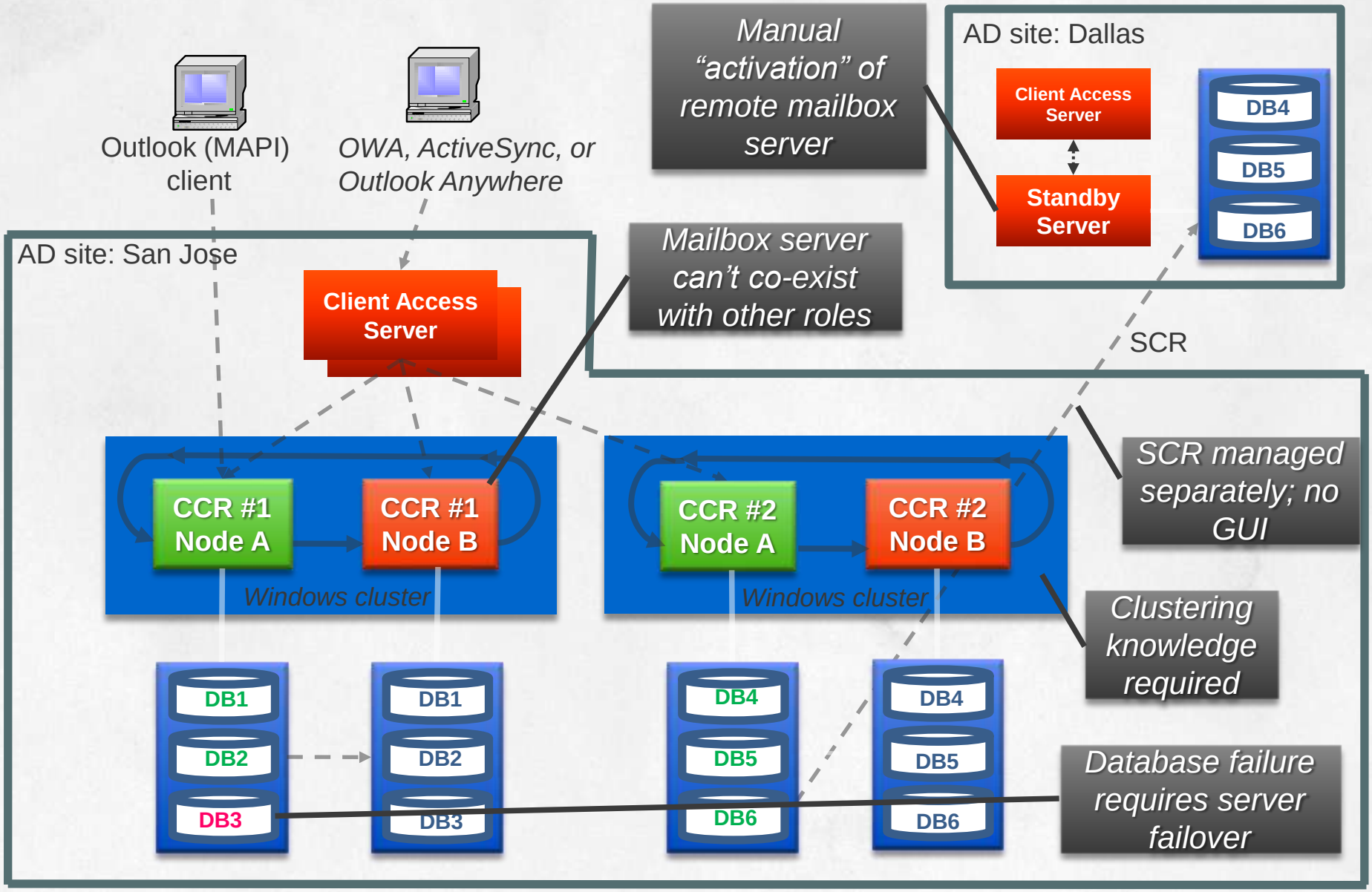




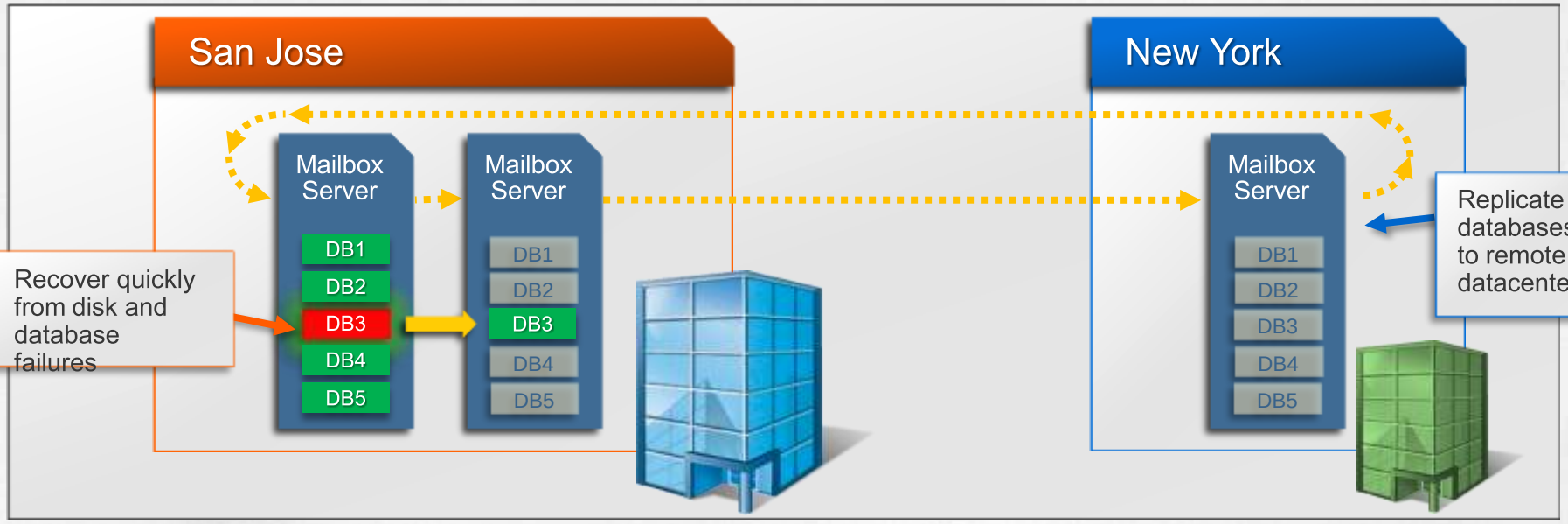




Exchange Server 2007 CCR + SCR



Simplified Mailbox High Availability, Disaster Recovery, and Backup



- New Unified Platform for High Availability, Disaster Recovery, and Backup
- Evolution of Continuous Replication technology
- Provides full redundancy of Exchange roles on as few as two servers
- Allows each database to have up to 16 copies

Exchange Server 2010 High Availability Goals

- Reduce complexity
- Reduce cost
- Native solution - no single point of failure
- Improve recovery times
- Support larger mailboxes
- Support large scale deployments

Make High Availability Exchange deployments mainstream!

Exchange Server Improvements

Improved mailbox uptime

- Improved failover granularity
- Simplified administration
- Incremental deployment
- Unification of CCR + SCR
- Easy stretching across sites
- Up to 16 replicated copies

More storage flexibility

- Further Input/Output (I/O) reductions
- RAID*-less/JBOD** support

Better end-to-end availability

- Online mailbox moves
- Improved transport resiliency

Key Benefits

- ✓ Easier and cheaper to deploy
- ✓ Easier and cheaper to manage
- ✓ Better Service Level Agreements (SLAs)

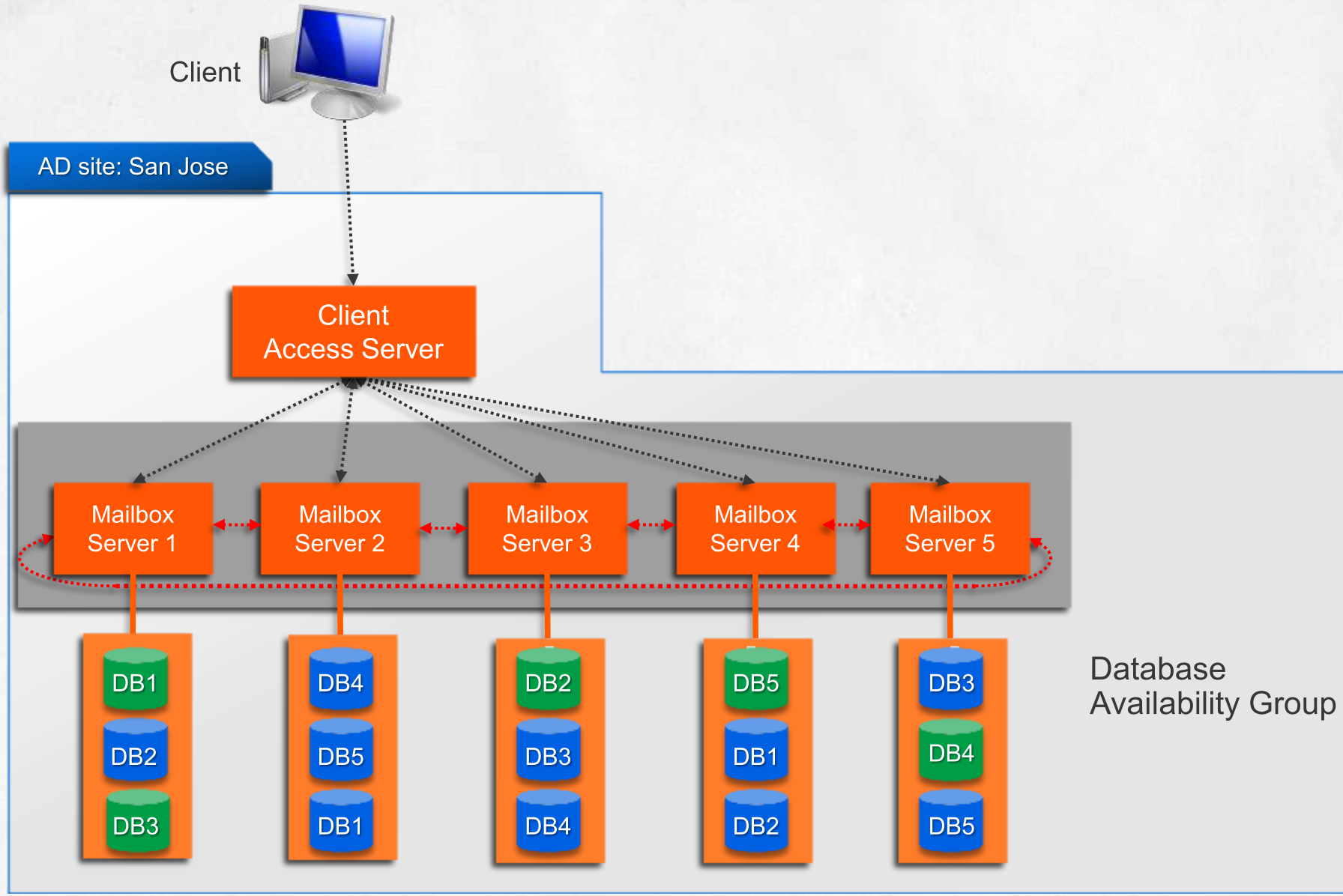
- ✓ Reduced storage costs
- ✓ Larger mailboxes

- ✓ Easier and cheaper to manage
- ✓ Better SLAs

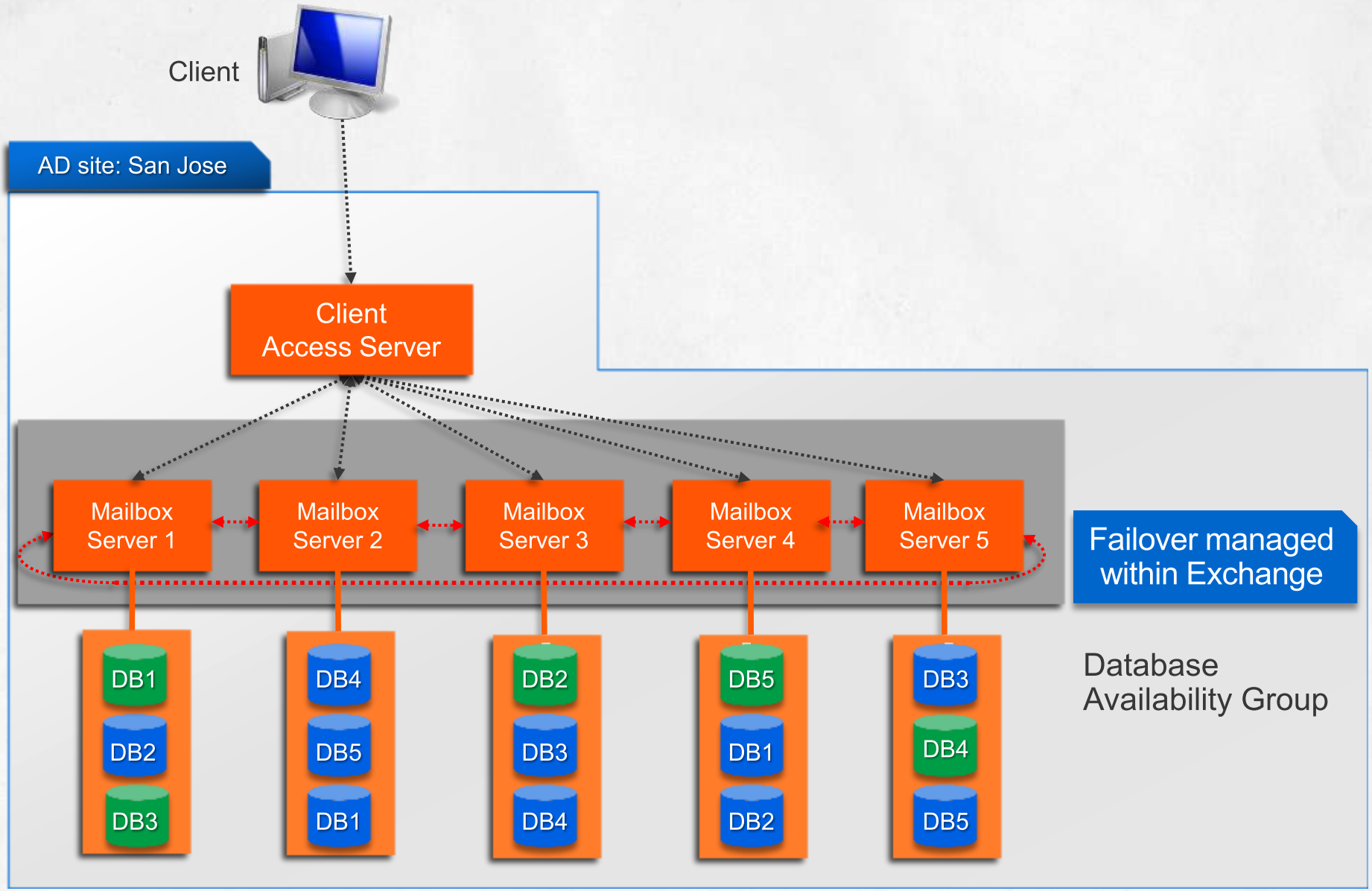
*Redundant Array of Independent Disks (RAID)

**Just a Bunch of Disks (JBOD)

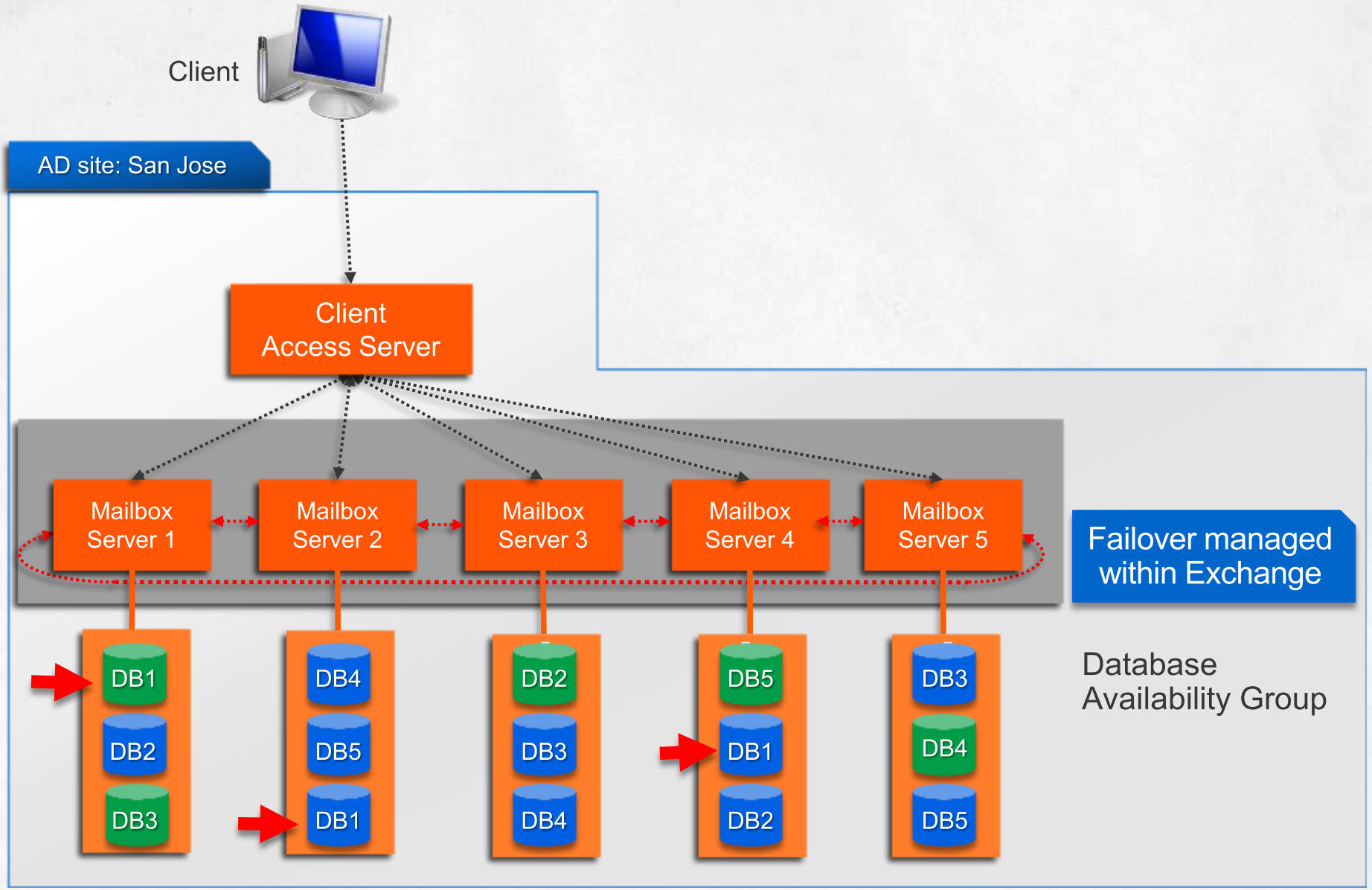
Exchange 2010 Mailbox Resiliency Overview



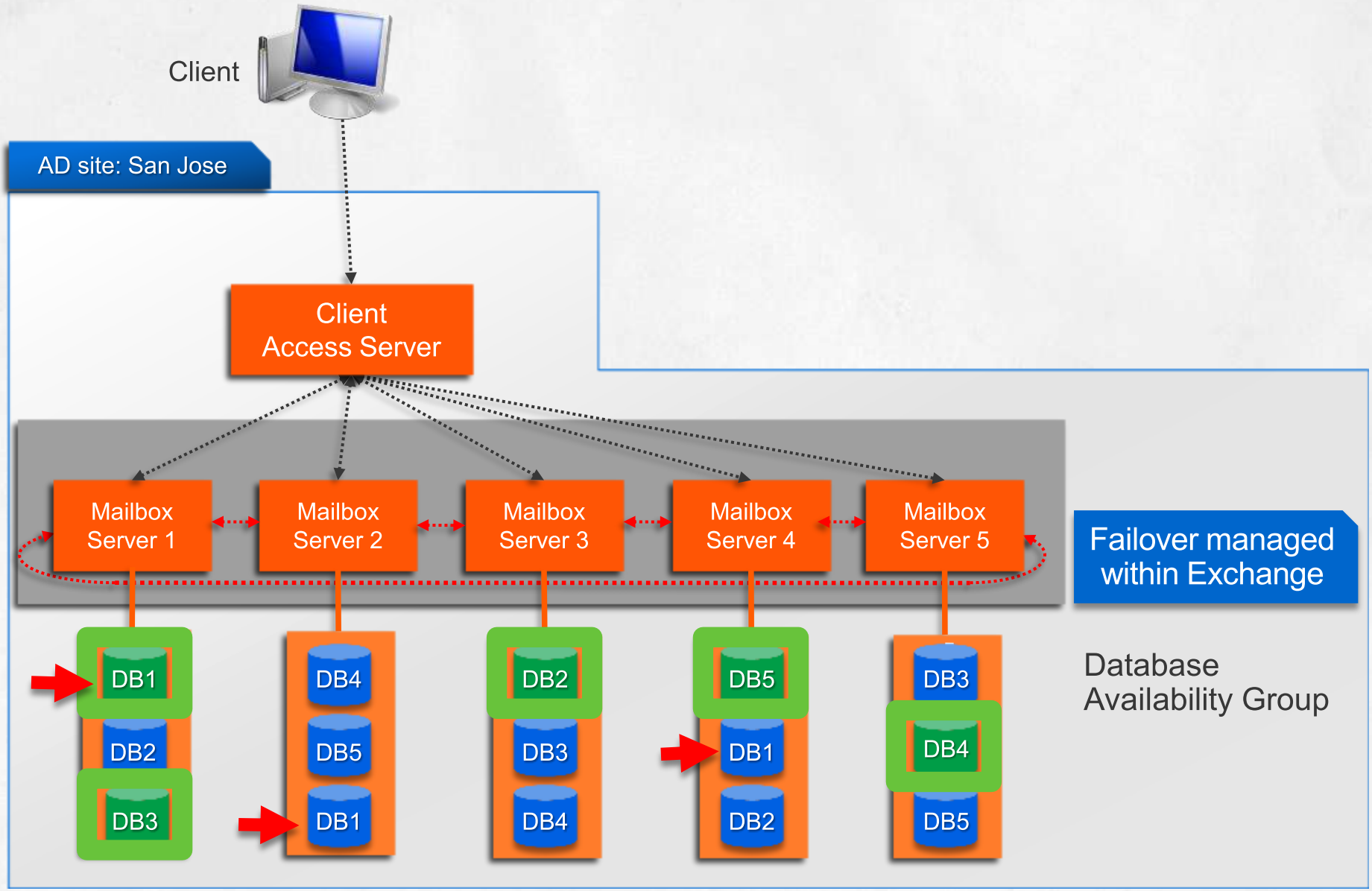
Exchange 2010 Mailbox Resiliency Overview



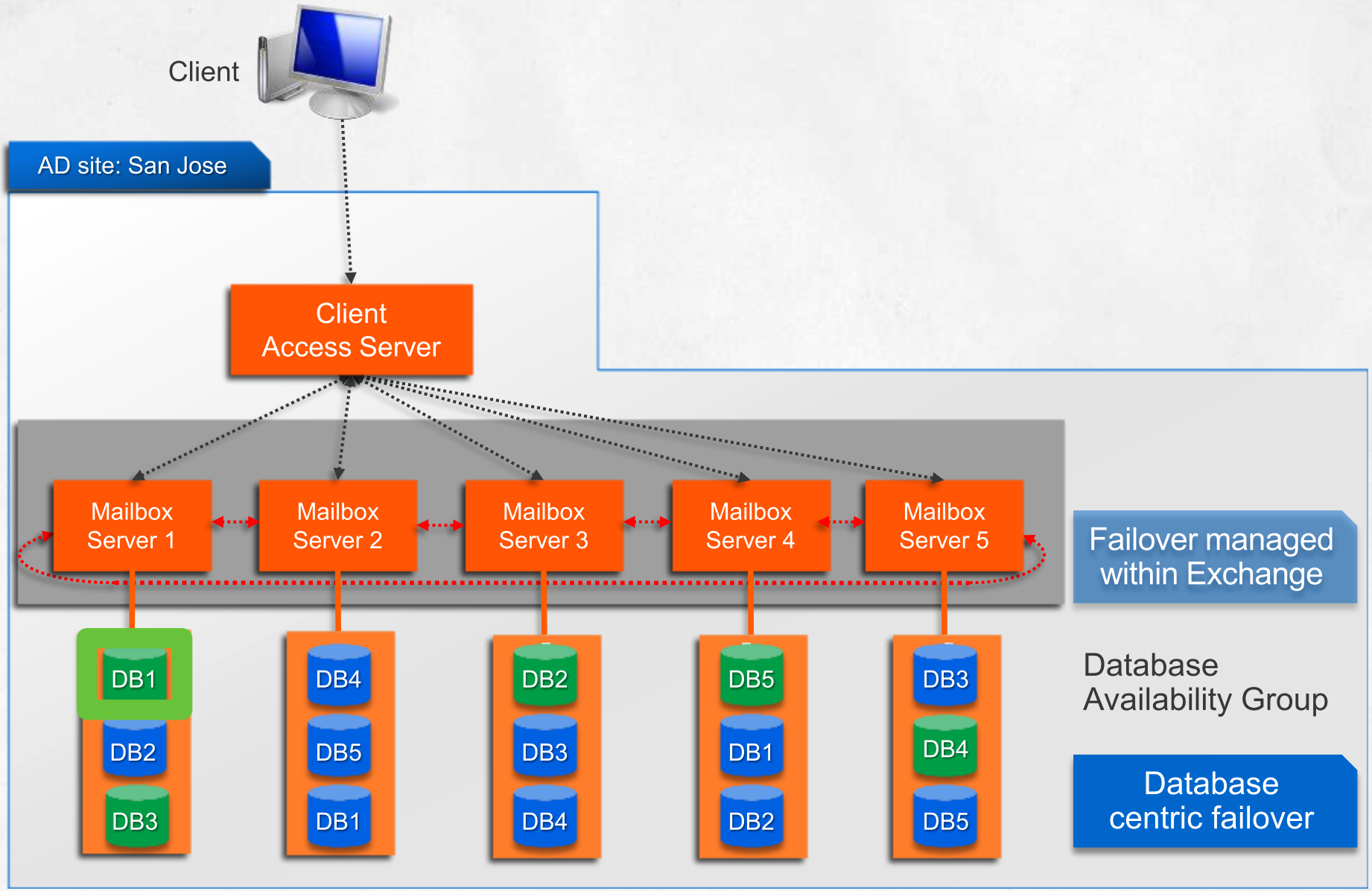
Exchange 2010 Mailbox Resiliency Overview



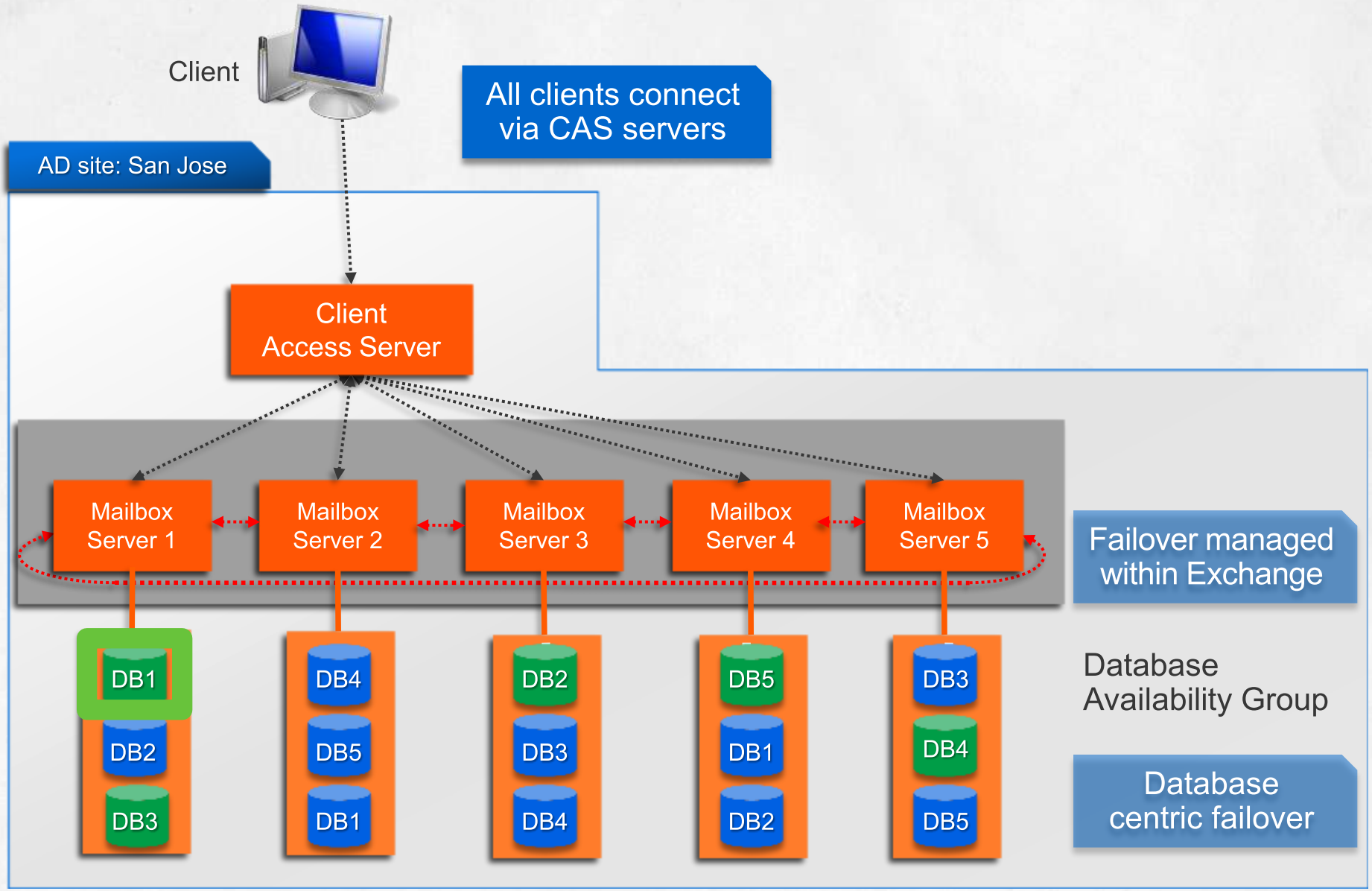
Exchange 2010 Mailbox Resiliency Overview



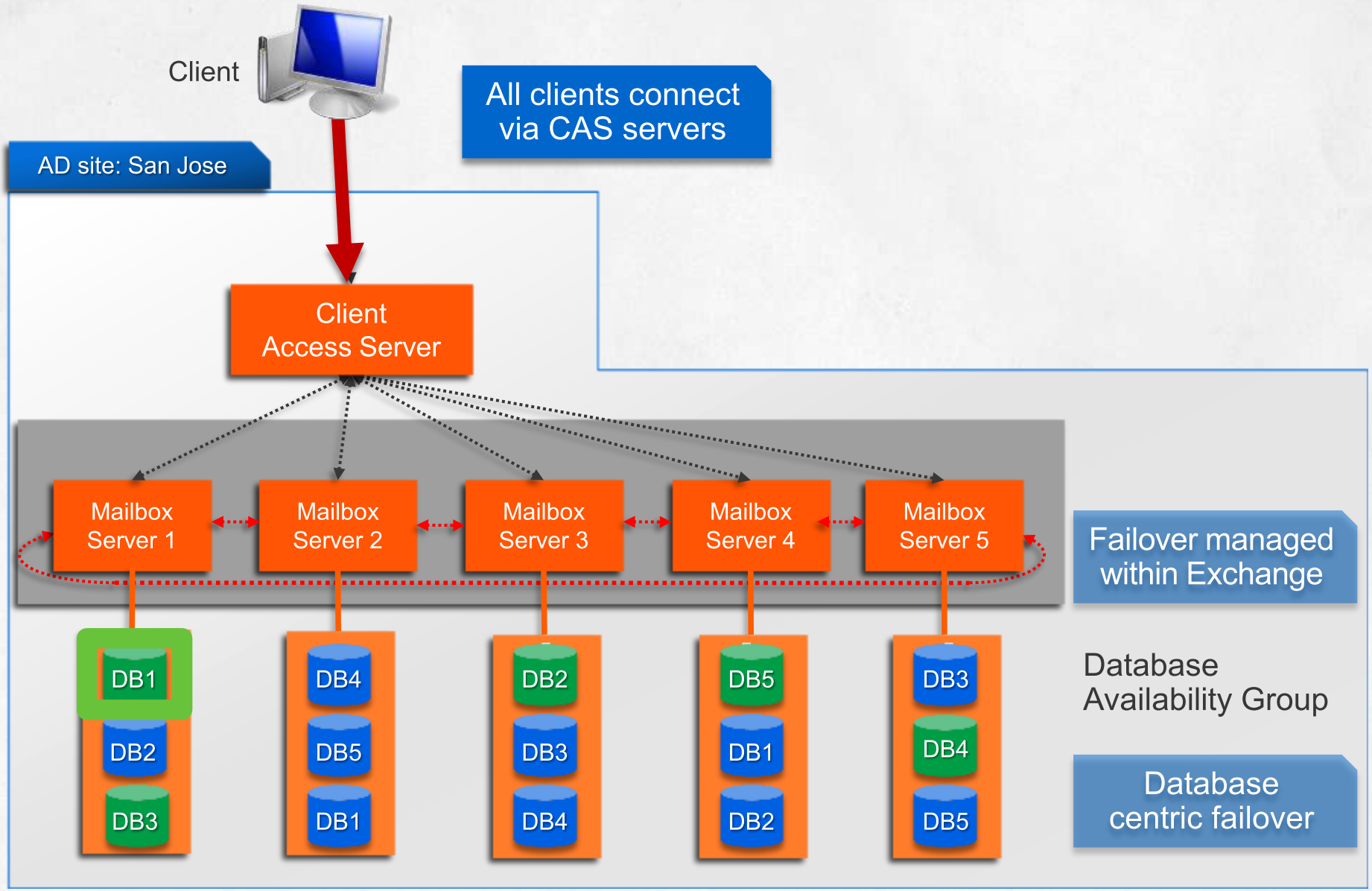
Exchange 2010 Mailbox Resiliency Overview



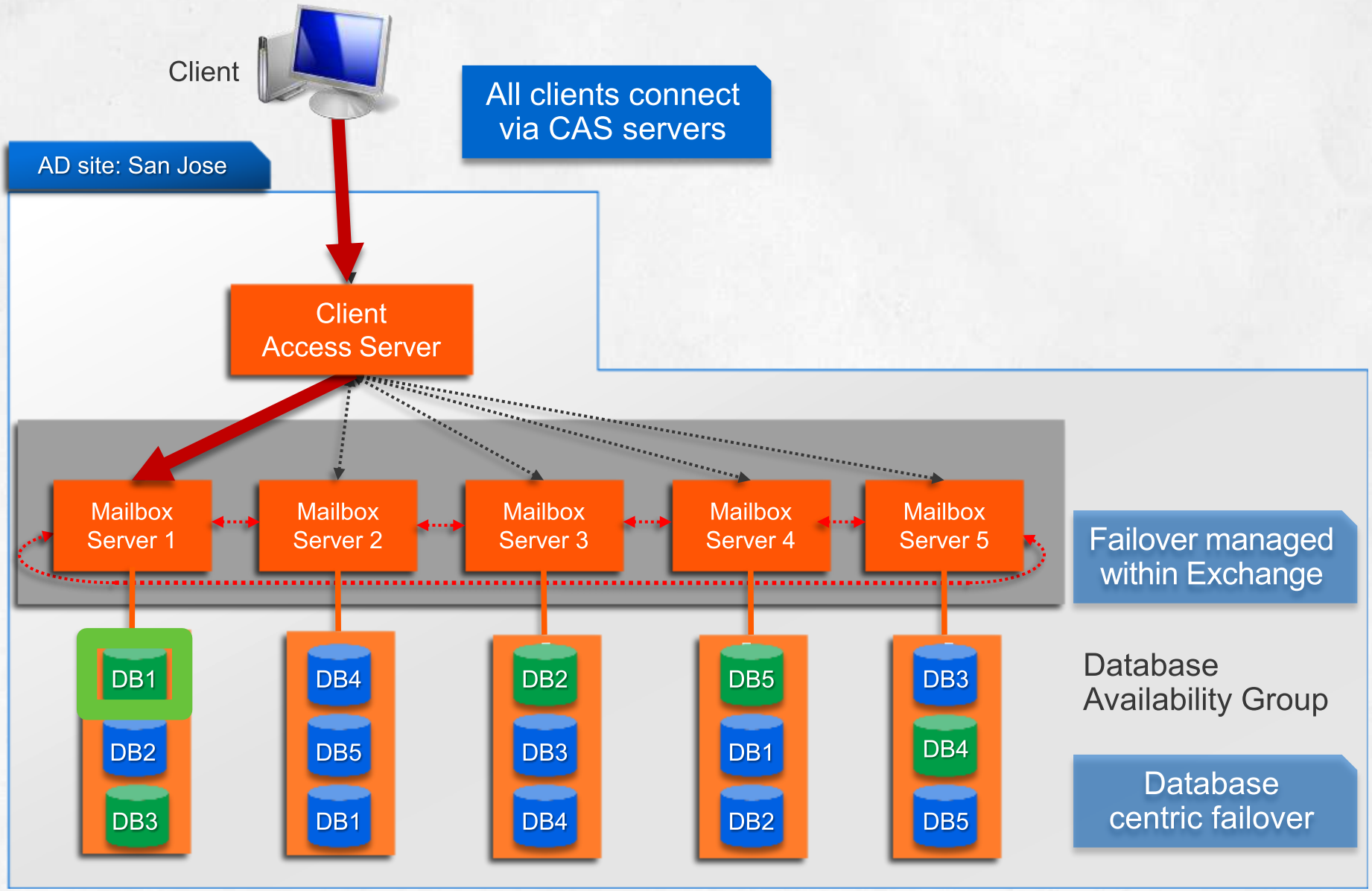
Exchange 2010 Mailbox Resiliency Overview



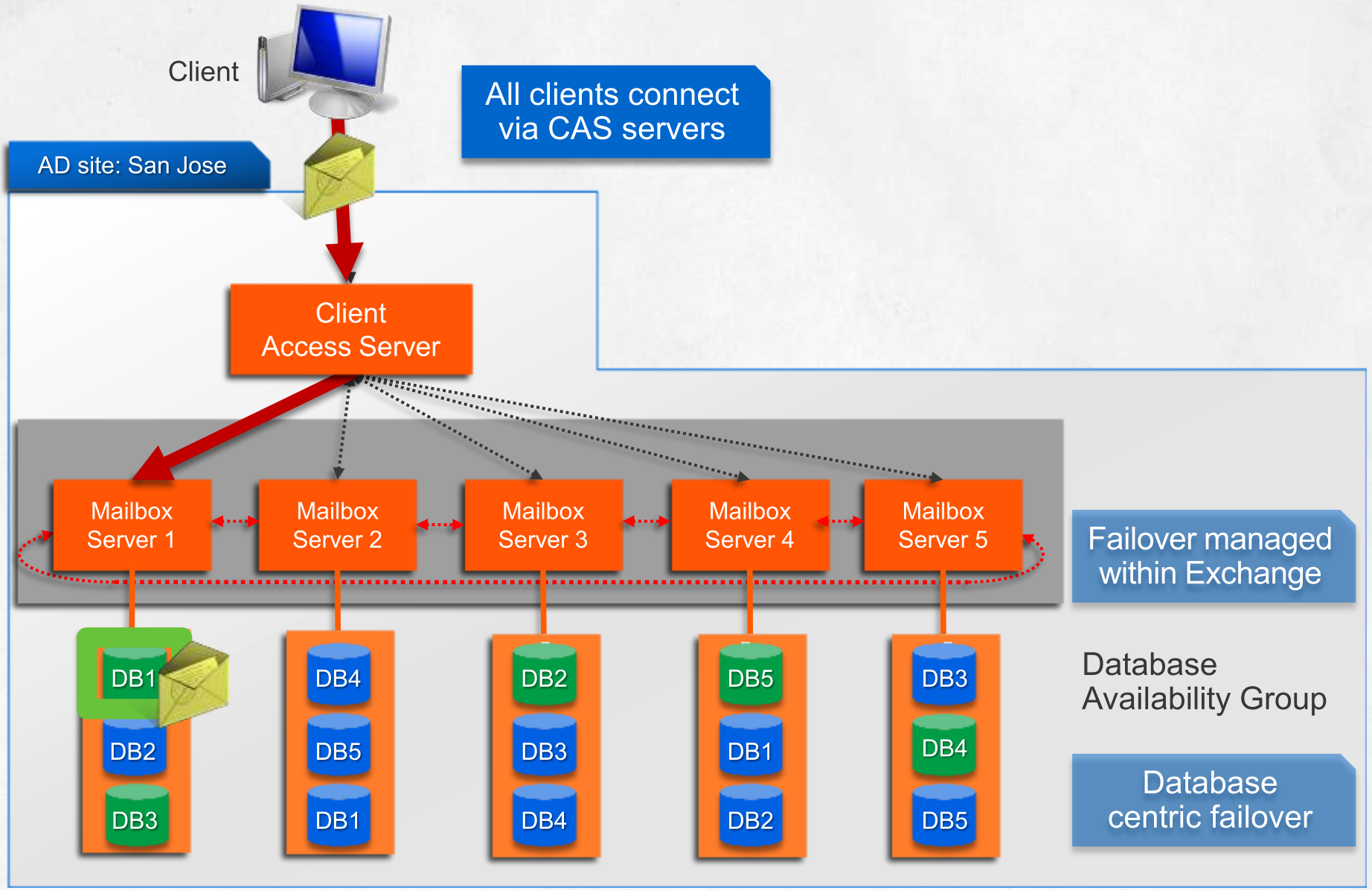
Exchange 2010 Mailbox Resiliency Overview



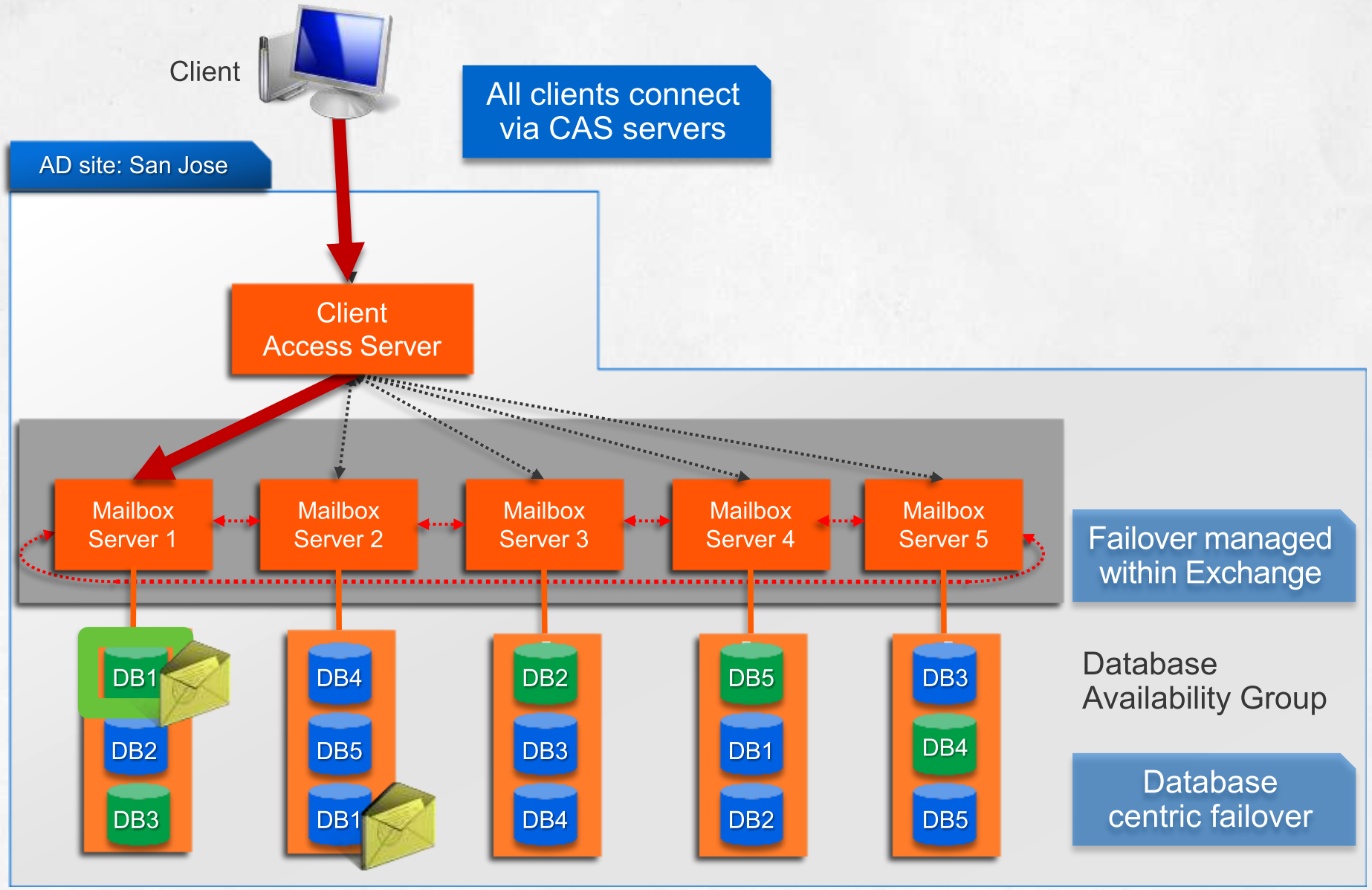
Exchange 2010 Mailbox Resiliency Overview



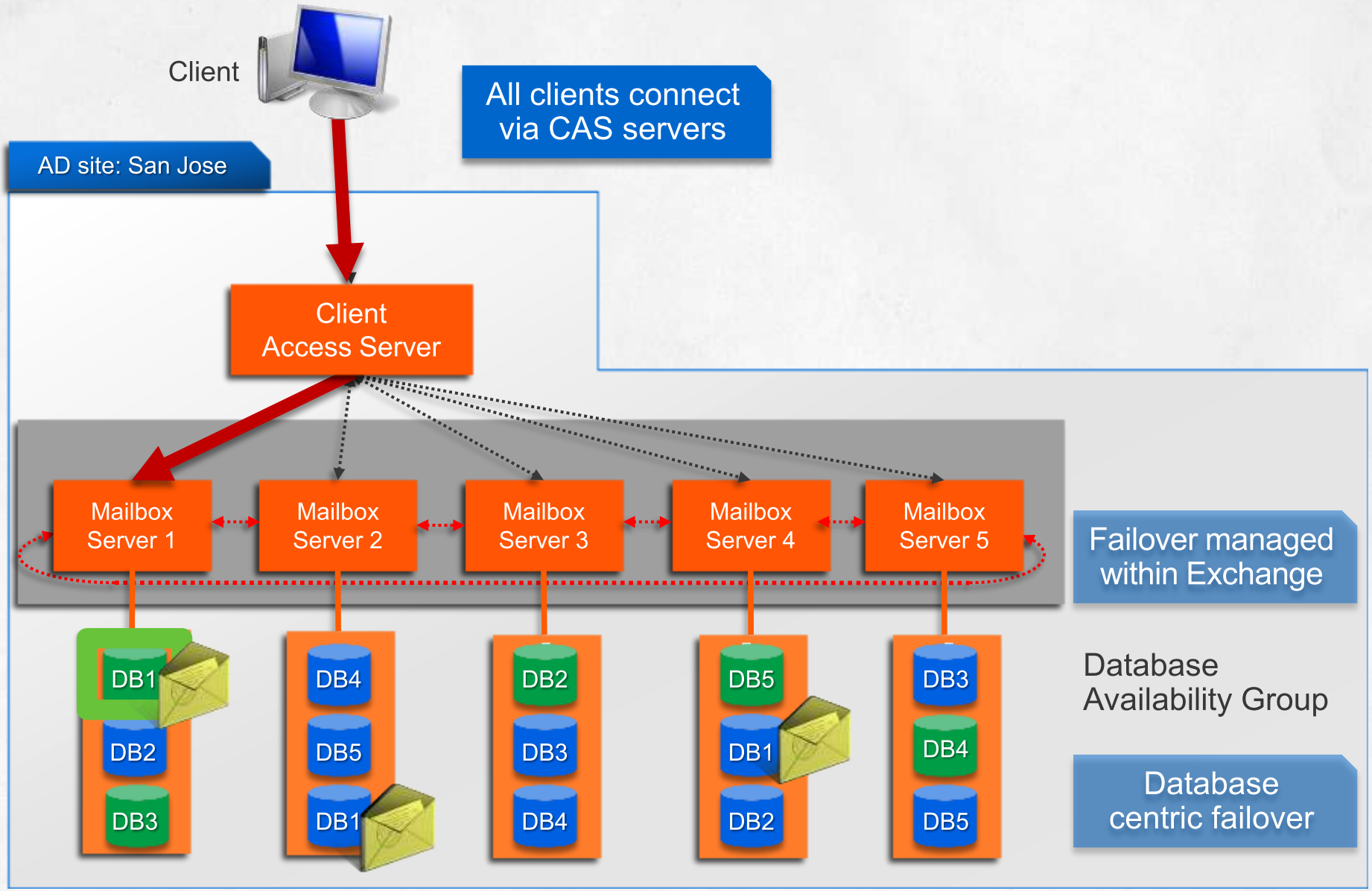
Exchange 2010 Mailbox Resiliency Overview



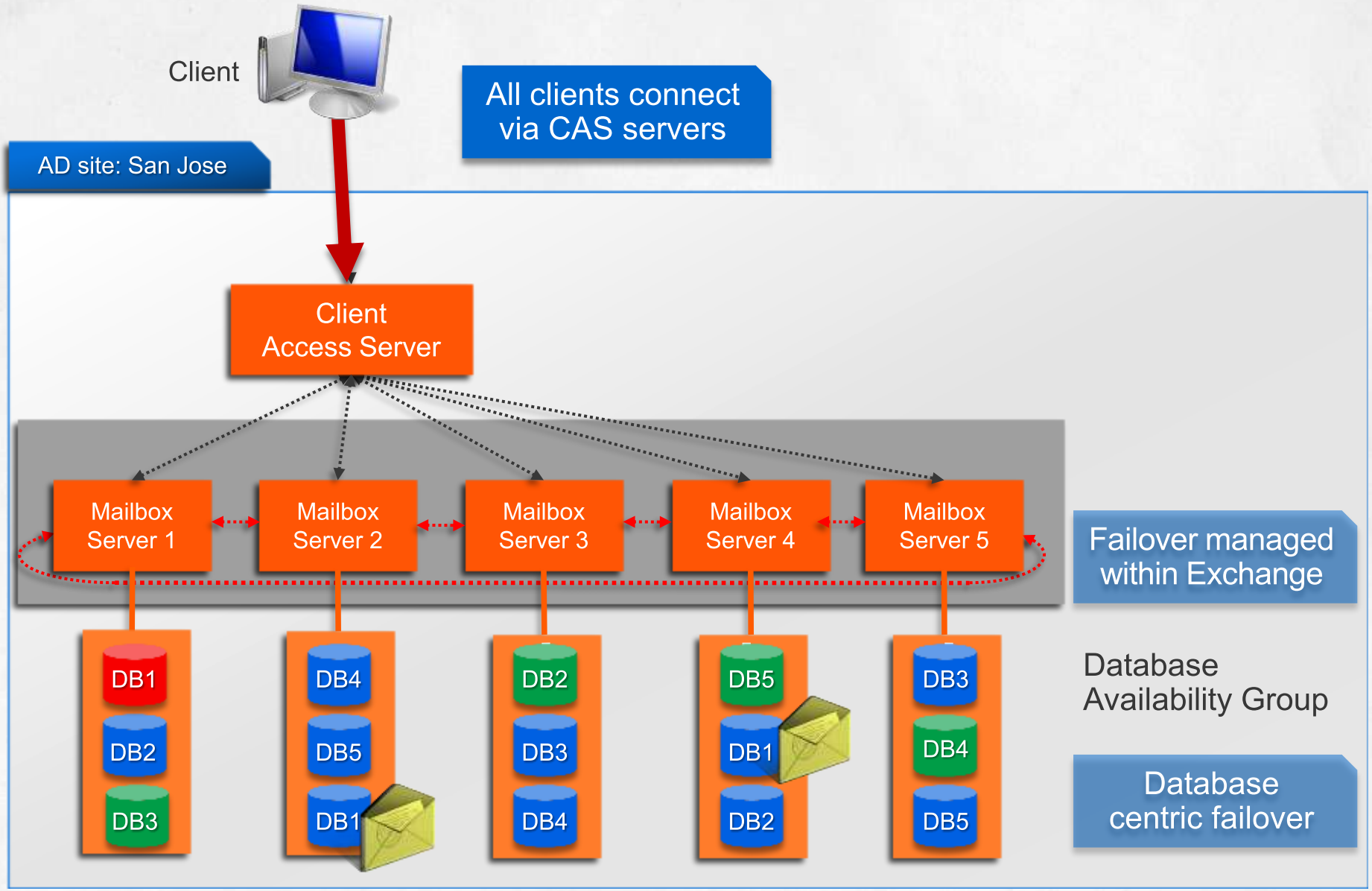
Exchange 2010 Mailbox Resiliency Overview



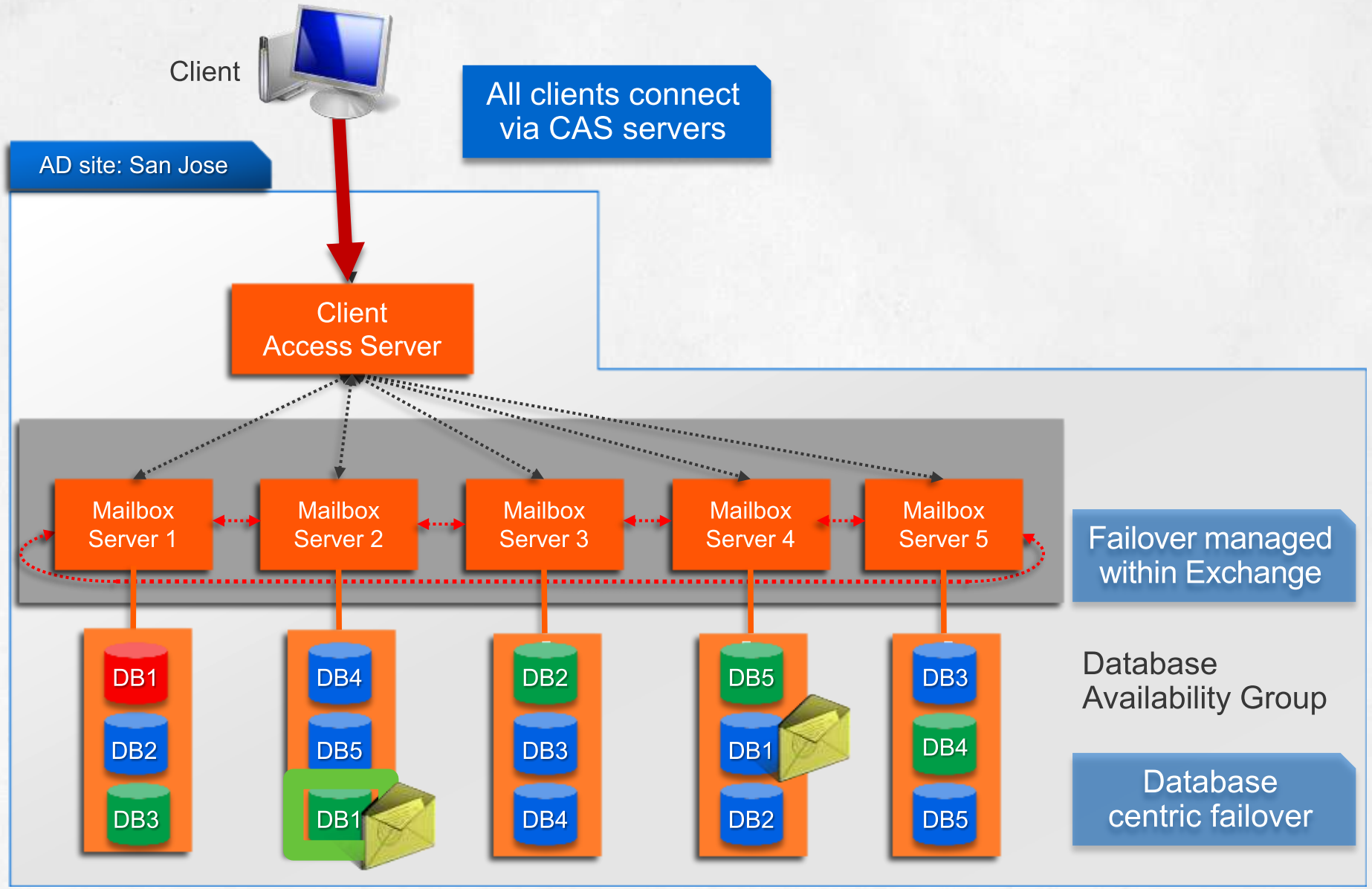
Exchange 2010 Mailbox Resiliency Overview



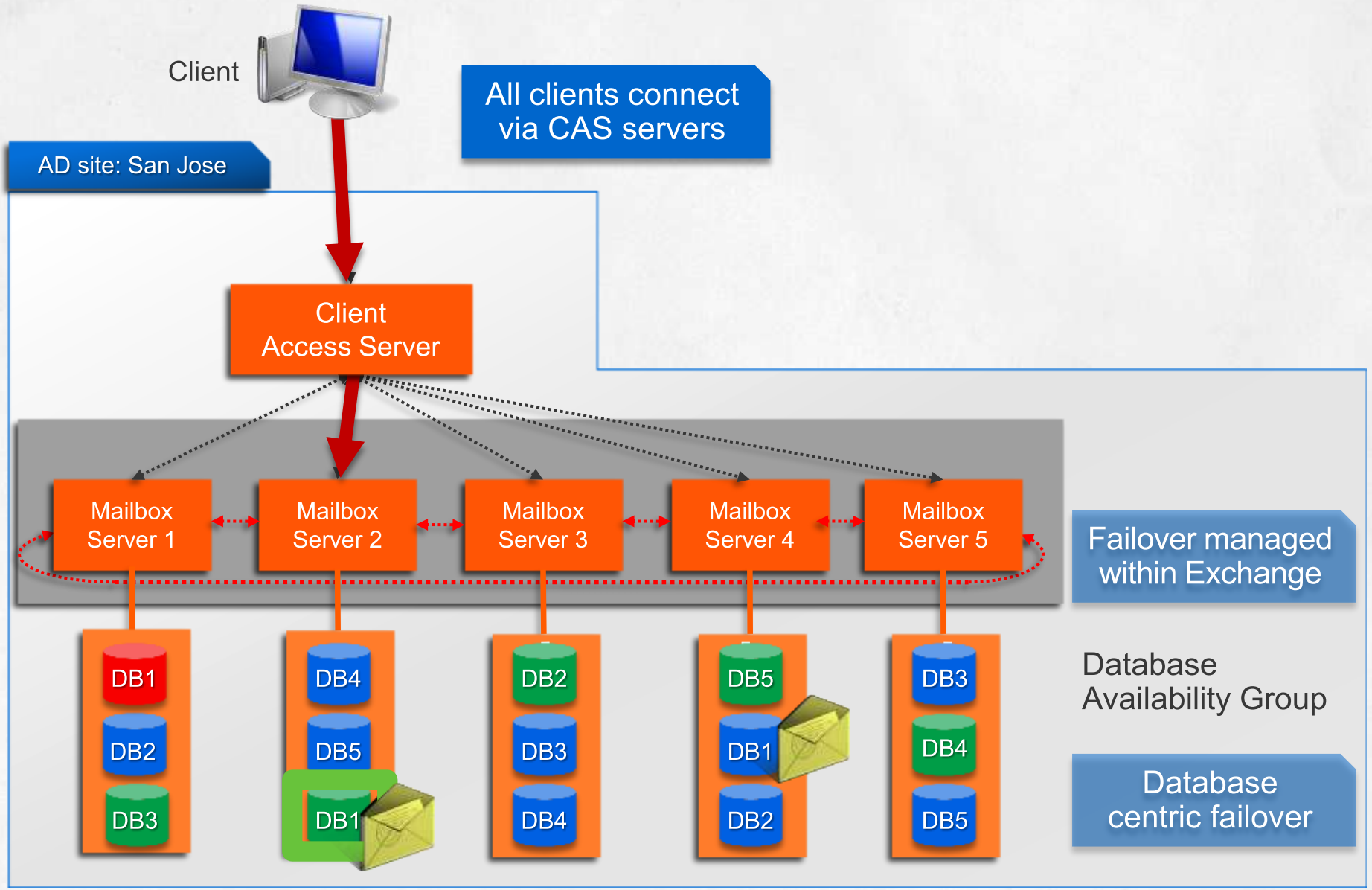
Exchange 2010 Mailbox Resiliency Overview



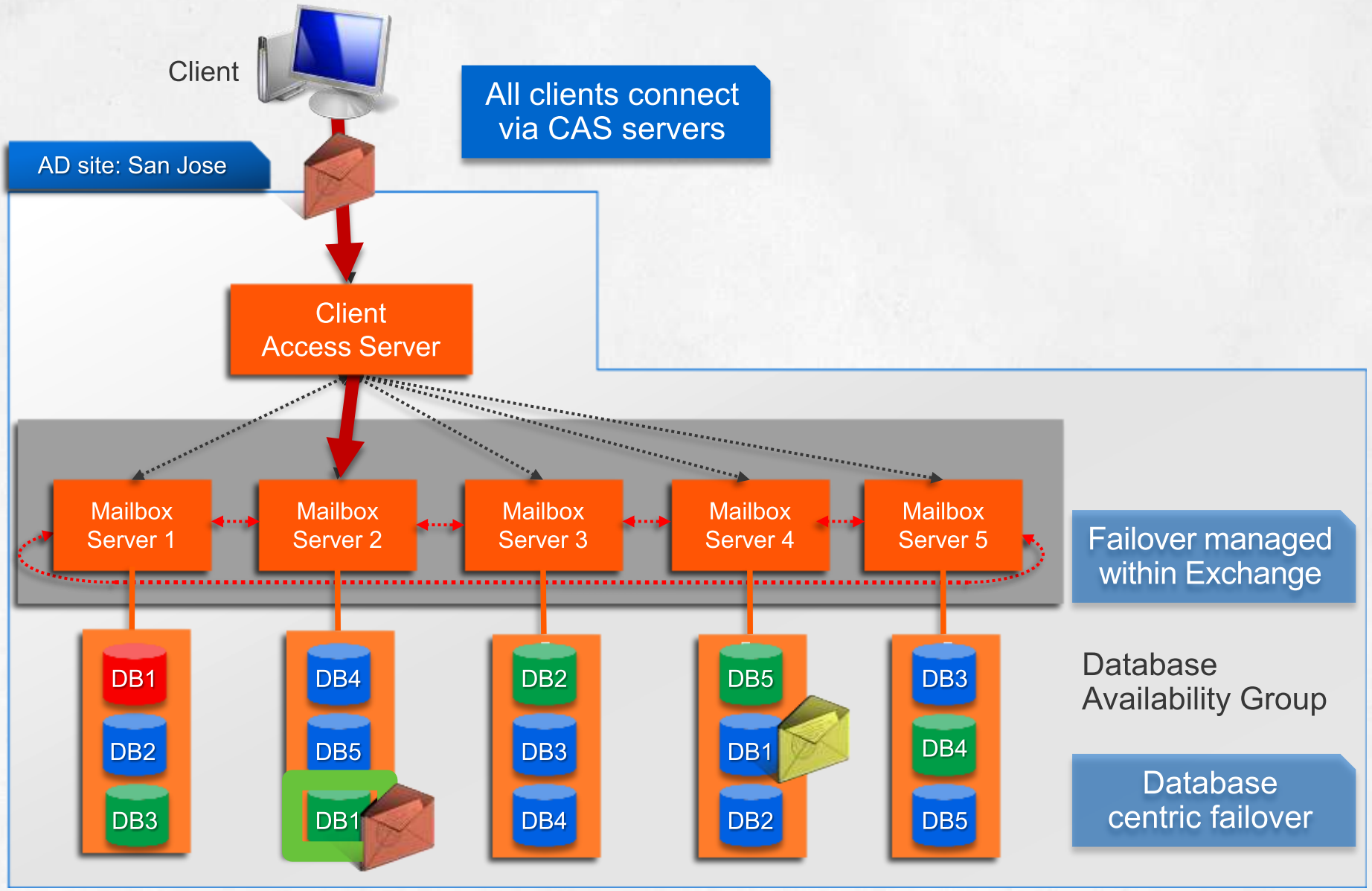
Exchange 2010 Mailbox Resiliency Overview



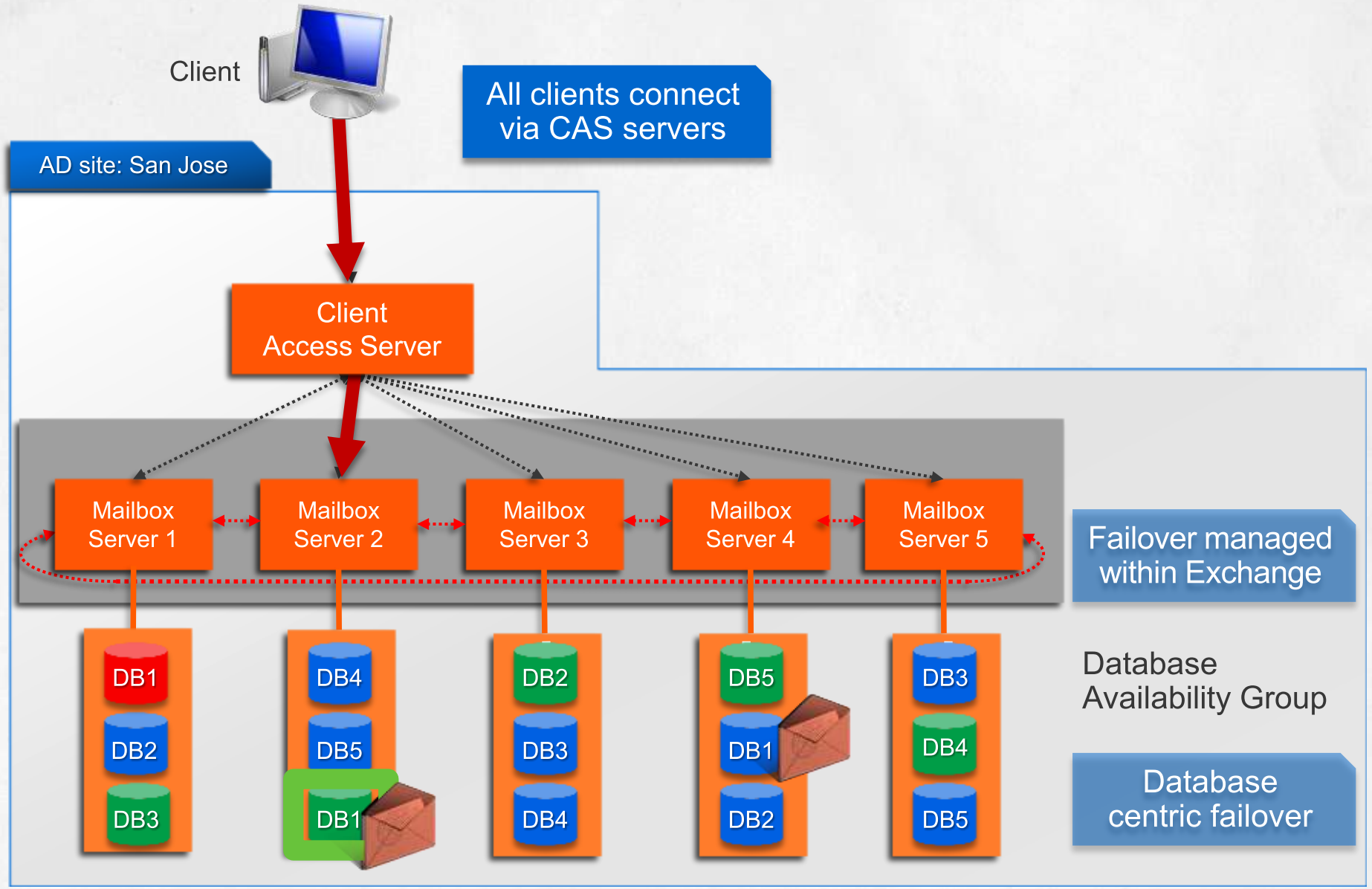
Exchange 2010 Mailbox Resiliency Overview



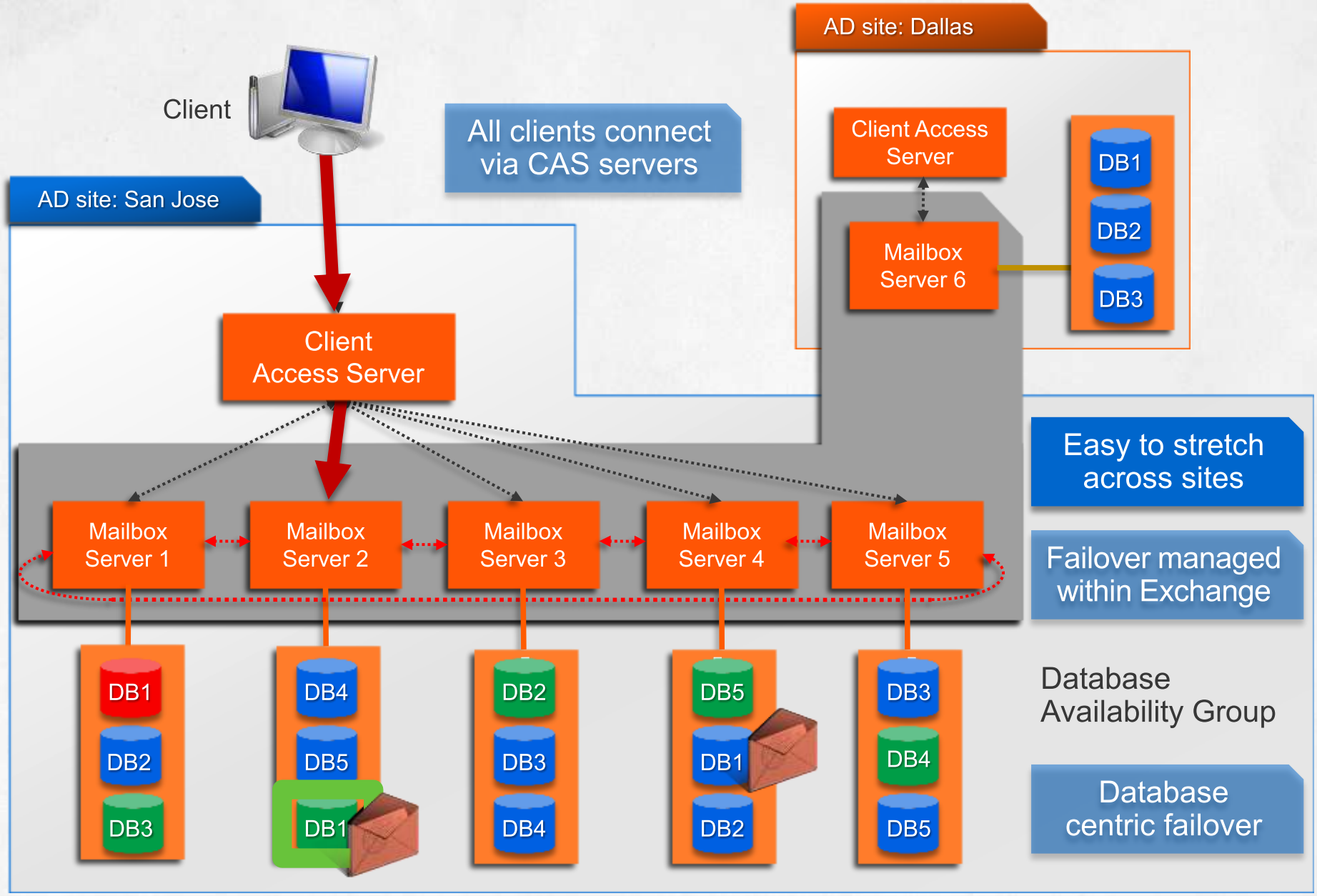
Exchange 2010 Mailbox Resiliency Overview



Exchange 2010 Mailbox Resiliency Overview

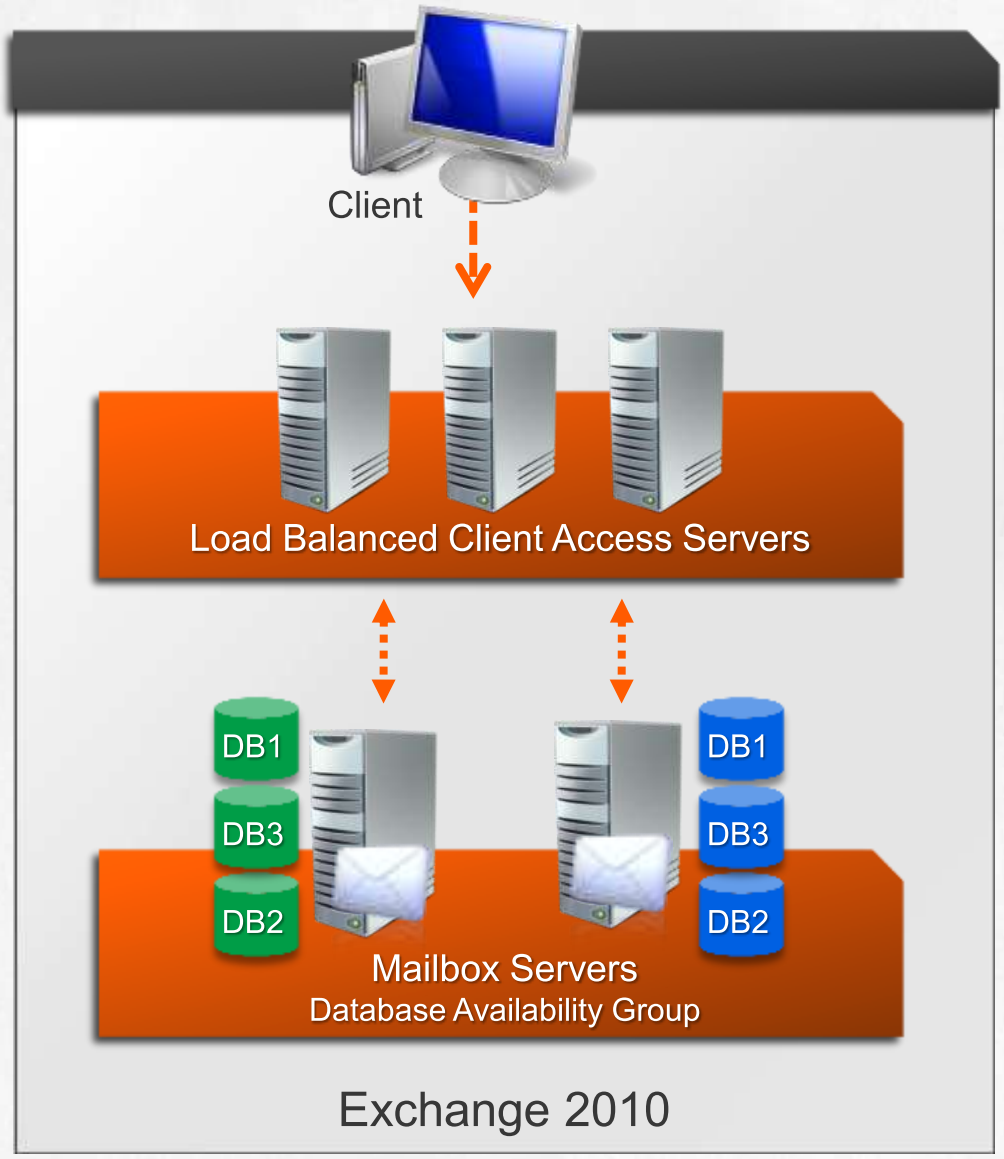


Exchange 2010 Mailbox Resiliency Overview



Improved Availability During Failures

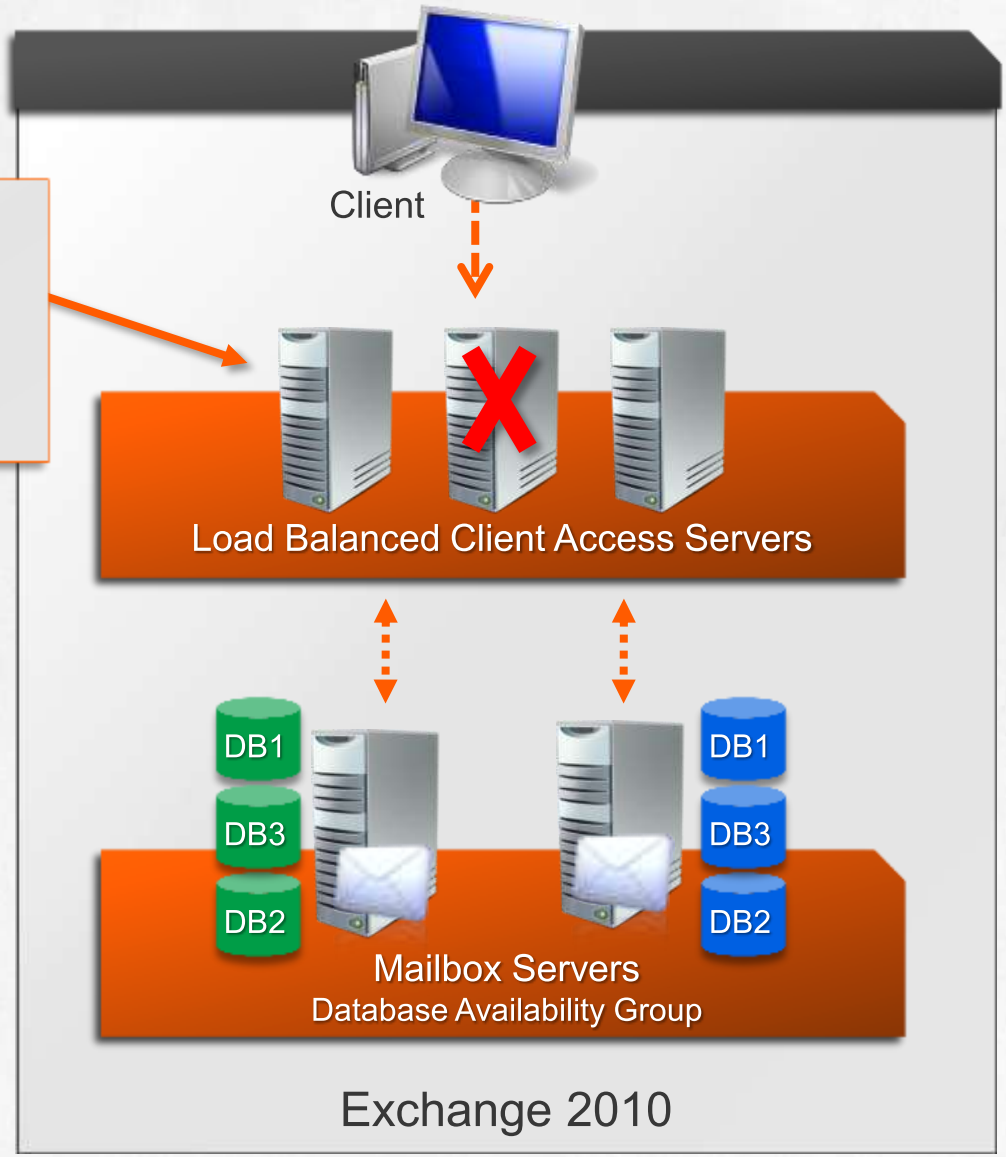
Keeping Users Connected



Improved Availability During Failures

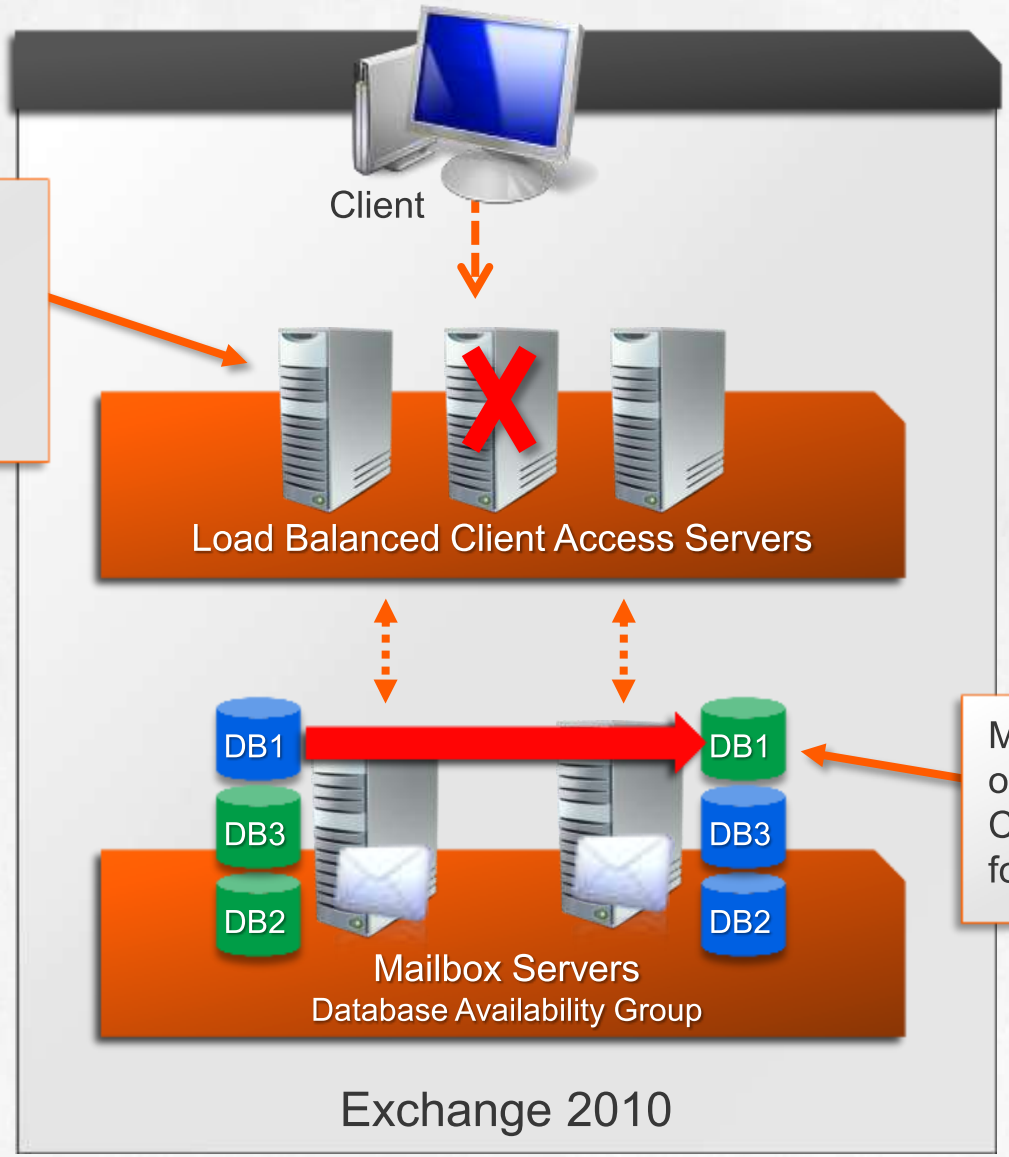
Keeping Users Connected

Client Access Server failure … Client reconnects through another Client Access Server



Improved Availability During Failures

Keeping Users Connected



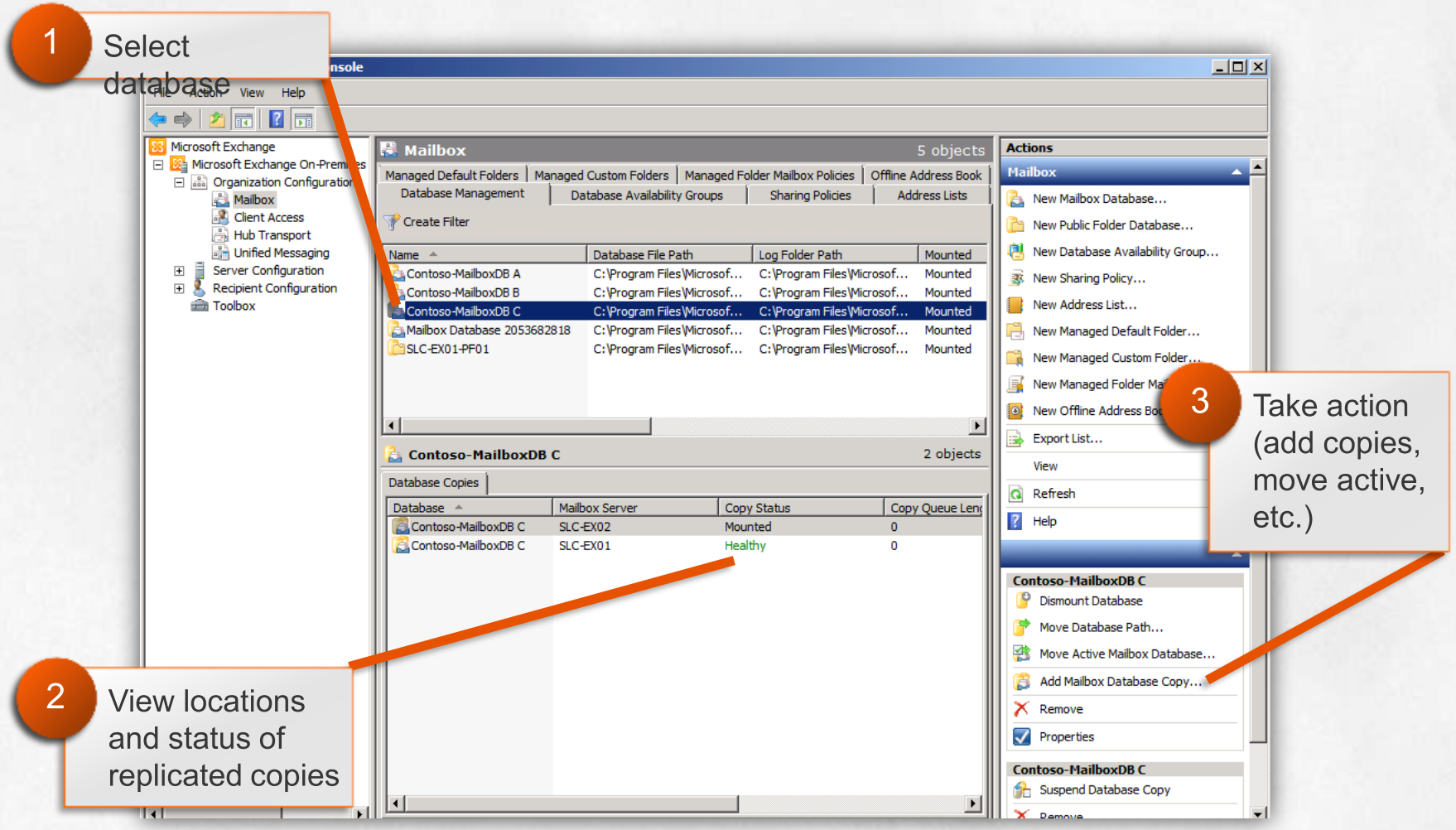
Client Access Server failure … Client reconnects through another Client Access Server

Mailbox Database or Server failure… Client disconnected for 30 seconds

Simplified Administration Reduces Cost and Complexity

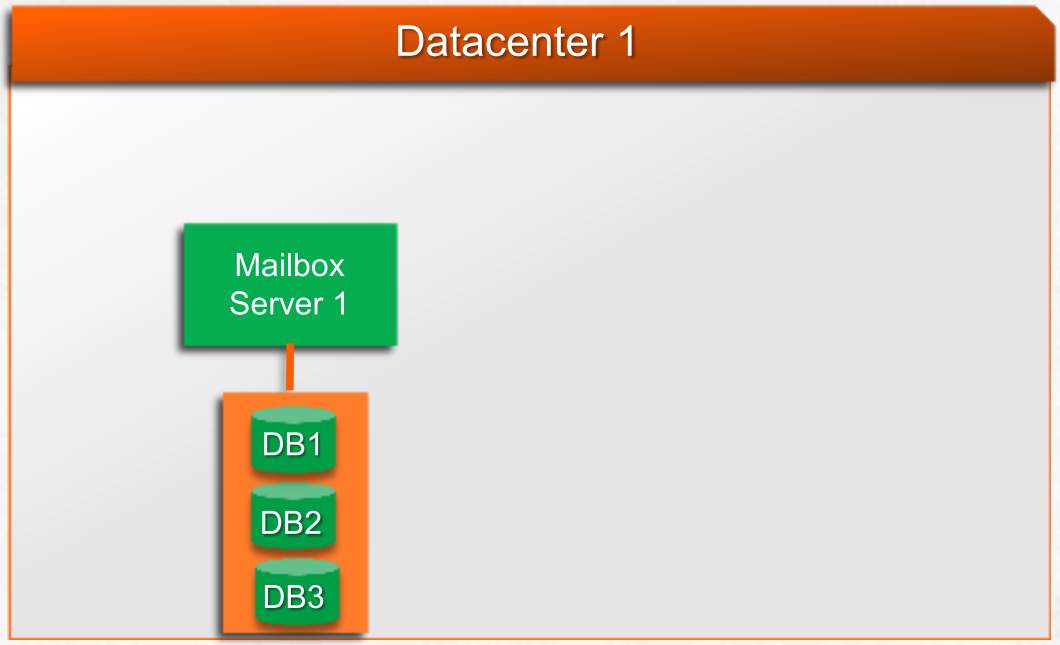
- High Availability administration all within Exchange
- Same automated database failover process used for a range for failures – Disk, Server, Network
- Simplified activation of Exchange services in a standby datacenter

Managing Availability in the Exchange Management Console



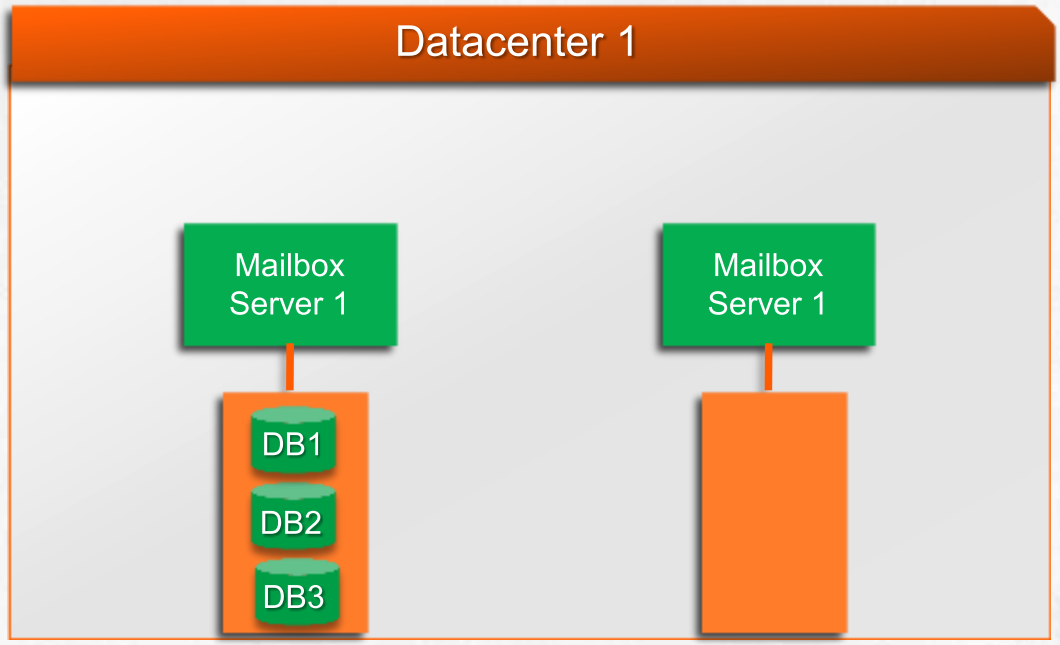
Incremental Deployment Reduces Cost & Complexity

- Easy to add High Availability to existing deployment
- High Availability configuration is post-setup
- High Availability mailbox servers can host other Server Roles



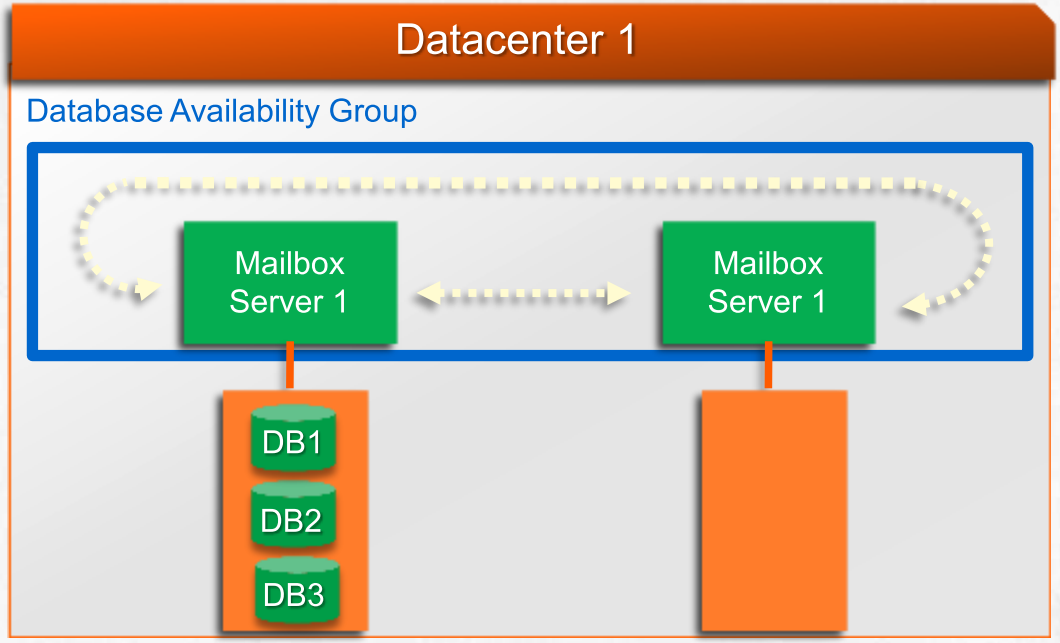
Incremental Deployment Reduces Cost & Complexity

- Easy to add High Availability to existing deployment
- High Availability configuration is post-setup
- High Availability mailbox servers can host other Server Roles



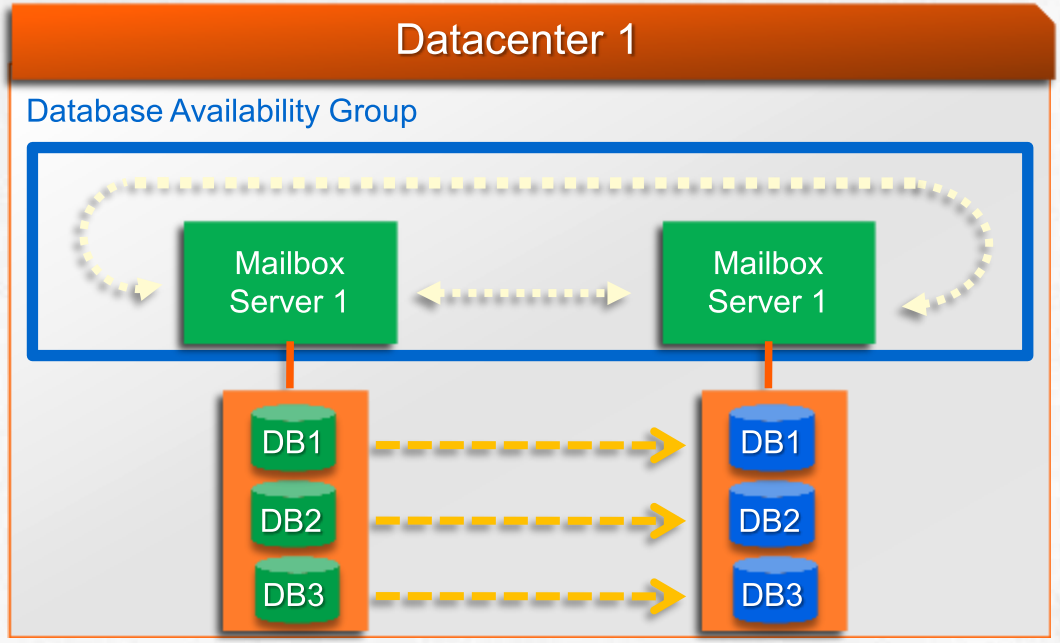
Incremental Deployment Reduces Cost & Complexity

- Easy to add High Availability to existing deployment
- High Availability configuration is post-setup
- High Availability mailbox servers can host other Server Roles



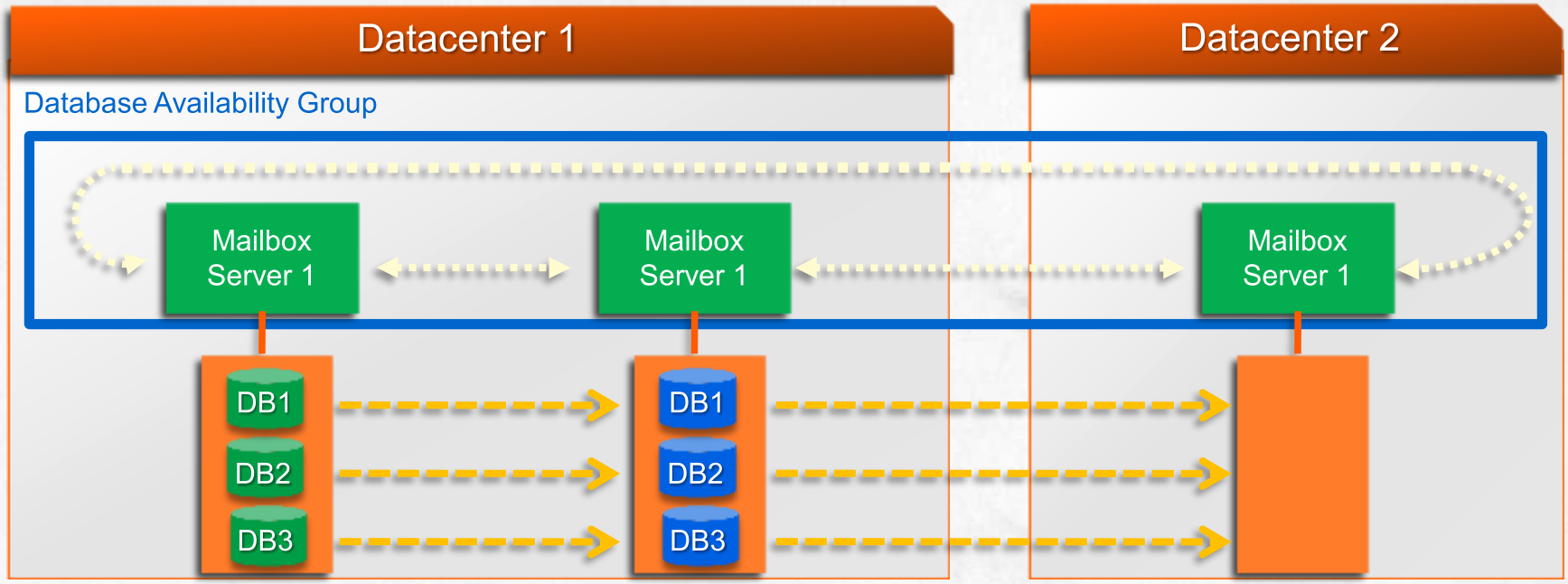
Incremental Deployment Reduces Cost & Complexity

- Easy to add High Availability to existing deployment
- High Availability configuration is post-setup
- High Availability mailbox servers can host other Server Roles



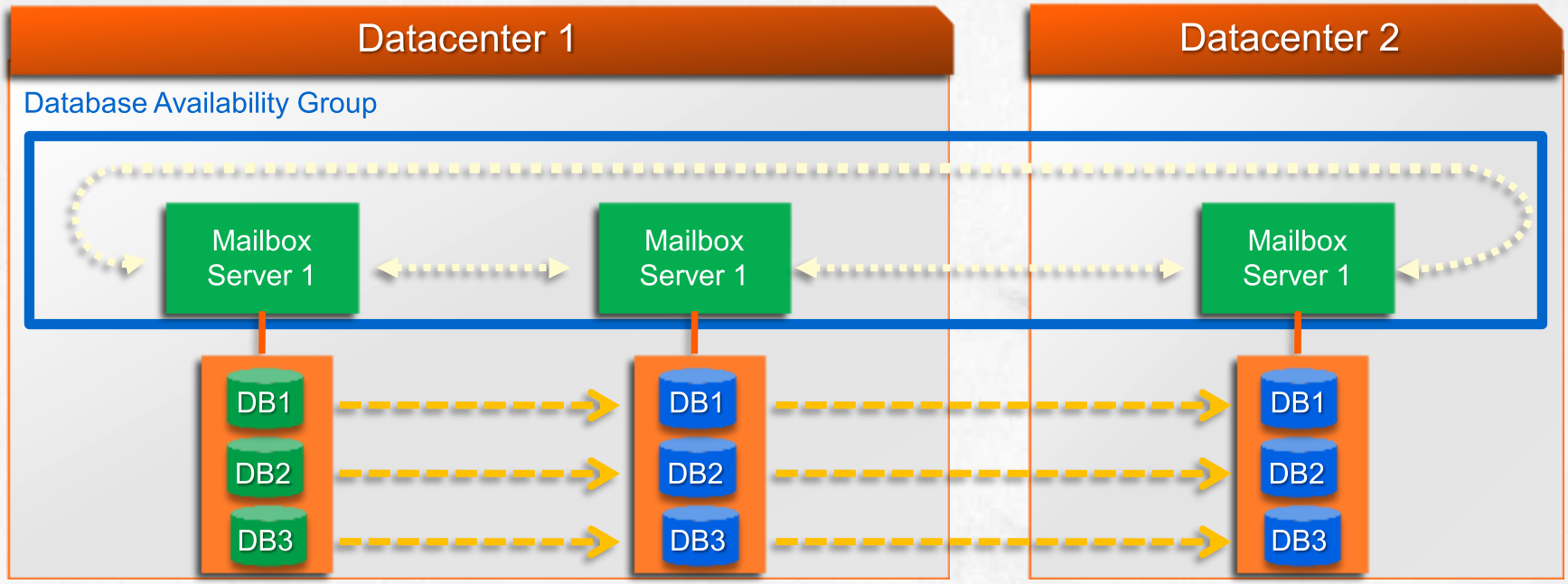
Incremental Deployment Reduces Cost & Complexity

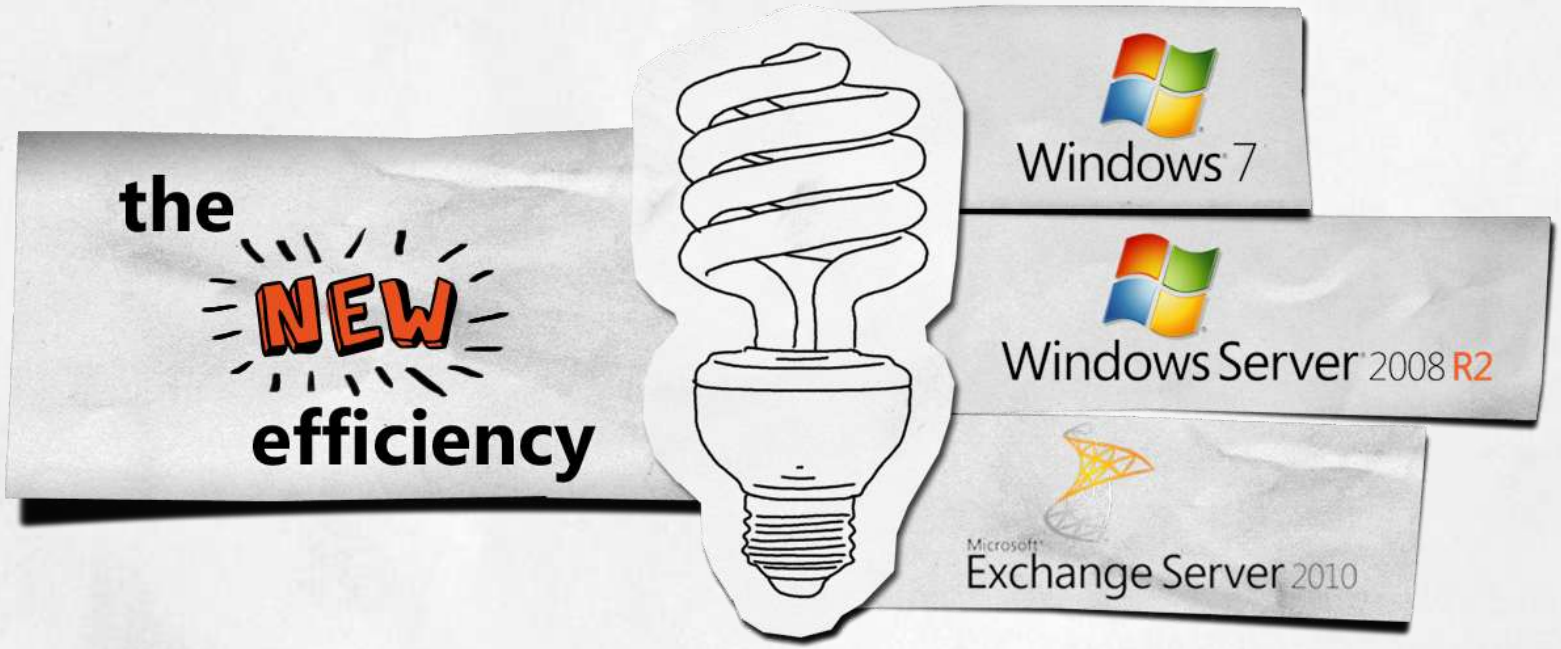
- Easy to add High Availability to existing deployment
- High Availability configuration is post-setup
- High Availability mailbox servers can host other Server Roles



Incremental Deployment Reduces Cost & Complexity

- Easy to add High Availability to existing deployment
- High Availability configuration is post-setup
- High Availability mailbox servers can host other Server Roles



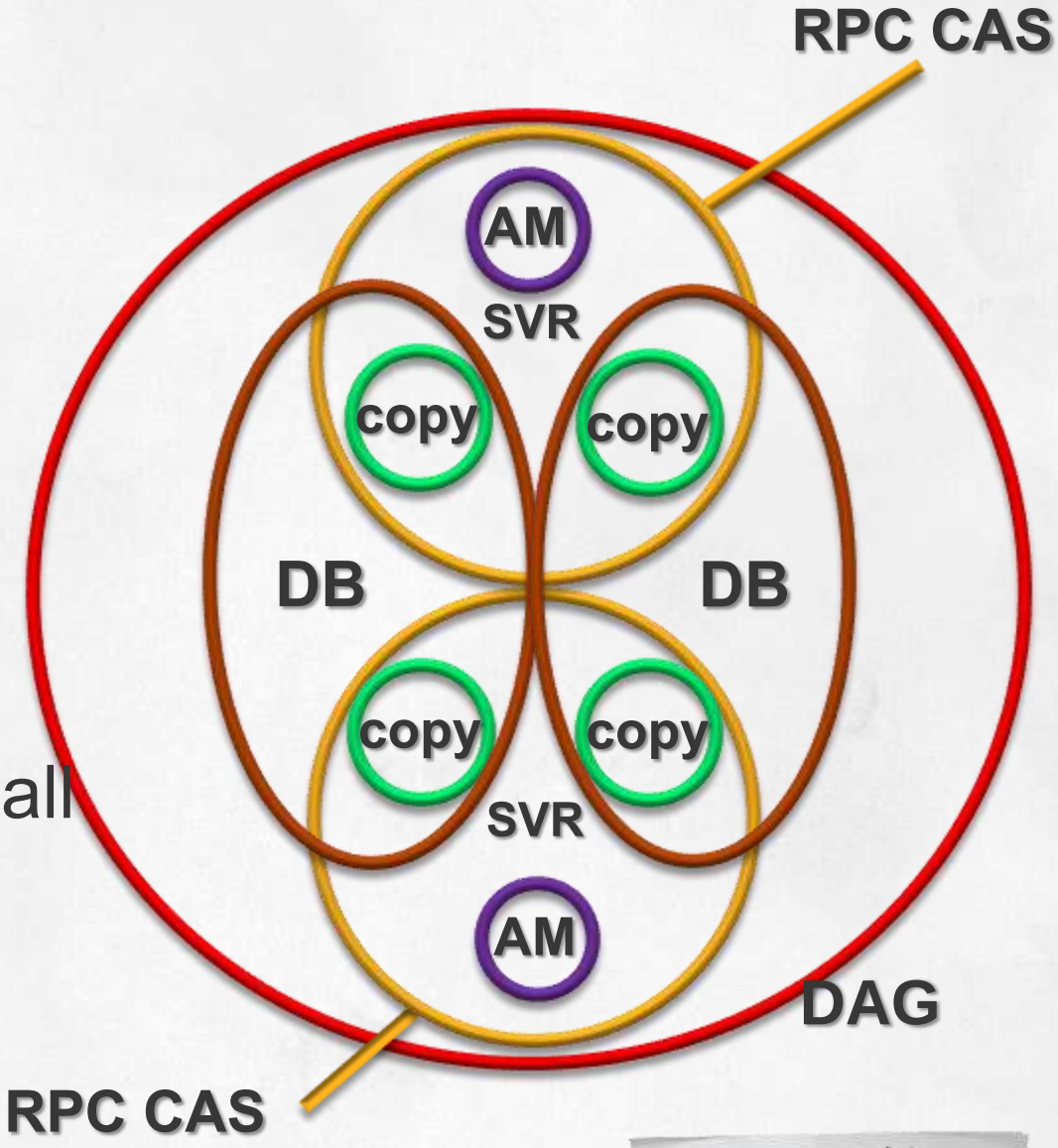


EXCHANGE SERVER 2010 HIGH AVAILABILITY FUNDAMENTALS

Because it's everybody's business

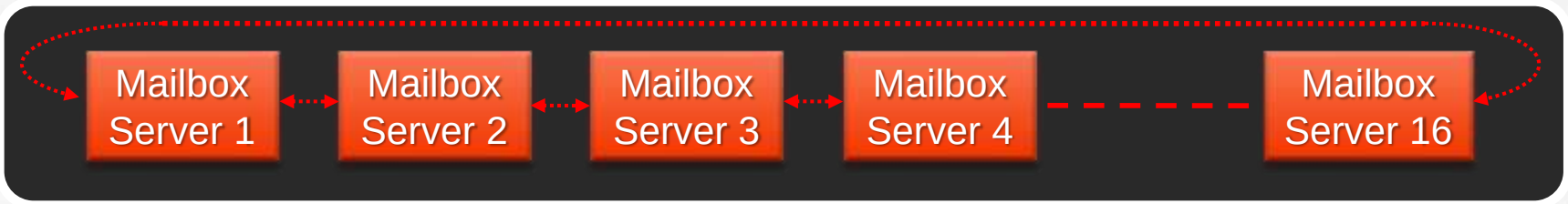
Exchange Server 2010 High Availability Fundamentals

- Database Availability Group
- Server
- Database
- Database Copy
- Active Manager (AM)
- Remote Procedure Call (RPC) Client Access service



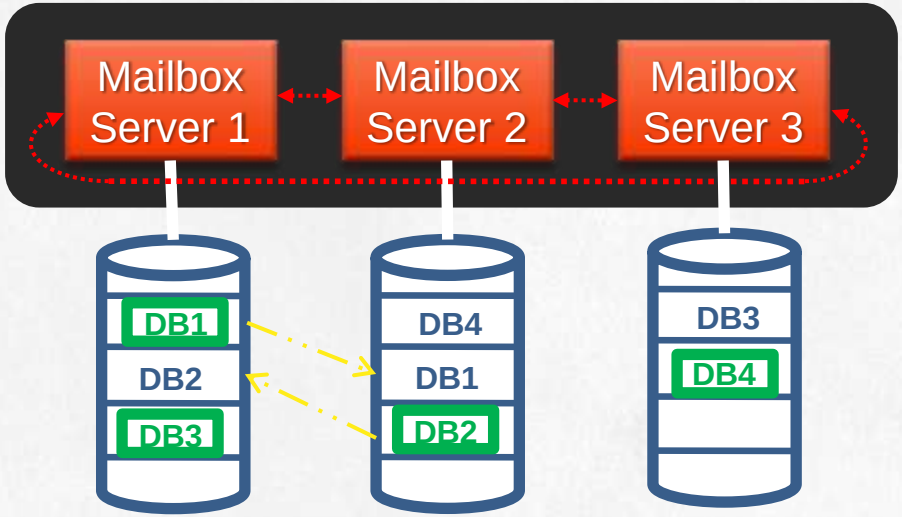
Exchange Server 2010 High Availability Fundamentals
Database Availability Group

- A group of up to 16 servers hosting a set of replicated databases
- Wraps a Windows Failover Cluster
  - Manages servers' membership in the group
  - Heartbeats servers, quorum, cluster database
- Defines the boundary of database replication
- Defines the boundary of failover/switchover (*over)
- Defines boundary for DAG's Active Manager



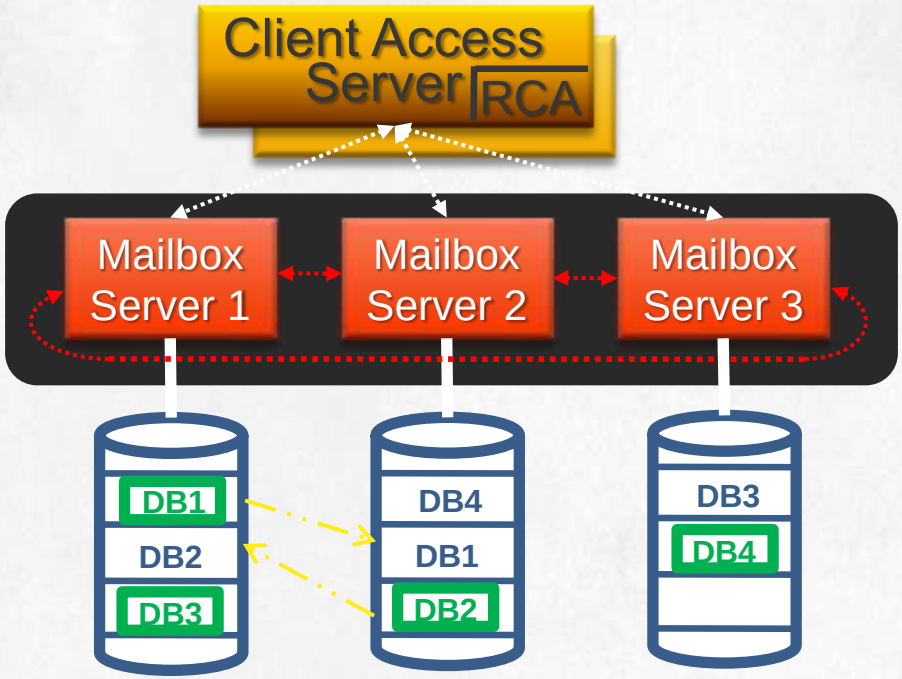
Exchange Server 2010 High Availability Fundamentals - Server

- Unit of membership for a DAG
- Hosts the active and passive copies of multiple mailbox databases
- Executes Information Store, CI, Assistants, etc., services on active mailbox database copies
- Executes replication services on passive mailbox database copies



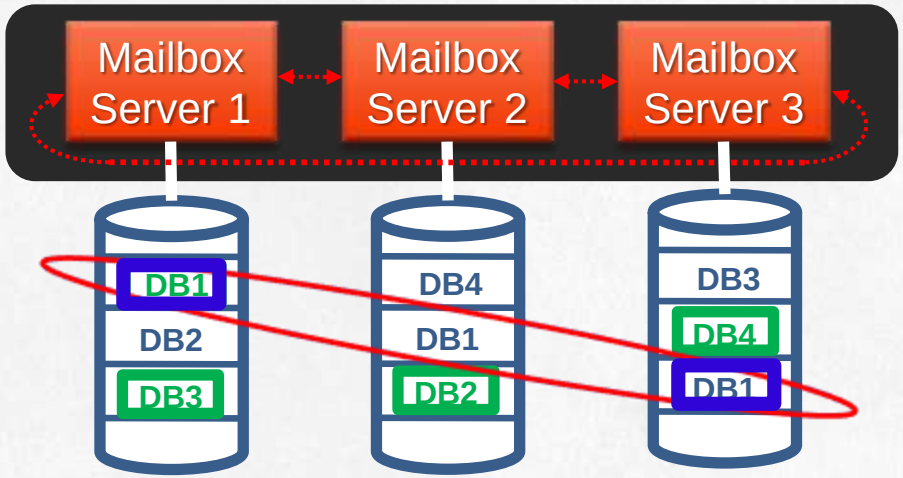
Exchange Server 2010 High Availability Fundamentals - Server (Continued)

- Provides connection point between Information Store and RPC Client Access
- Very few server-level properties relevant to High Availability (HA)
  - Server's Database Availability Group
  - Server's Activation Policy



Exchange Server 2010 High Availability Fundamentals - Mailbox Database

- Unit of *over
- A database has 1 active copy – active copy can be mounted or dismounted
- Maximum # of passive copies == # servers in DAG – 1



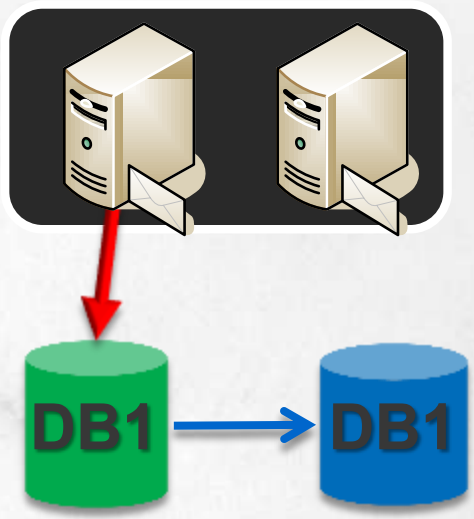
Exchange Server 2010 High Availability Fundamentals

Mailbox Database (Continued)

- ~30 seconds database *overs
- Server failover/switchover involves moving all active databases to one or more other servers
- Database names are unique across a forest
- Defines properties relevant at the database level
  - Globally Unique Identifier (GUID): a Database's unique ID
  - EdbFilePath: path at which copies are located
  - Servers: list of servers hosting copies

Exchange Server 2010 High Availability Fundamentals
Active/Passive vs. Source/Target

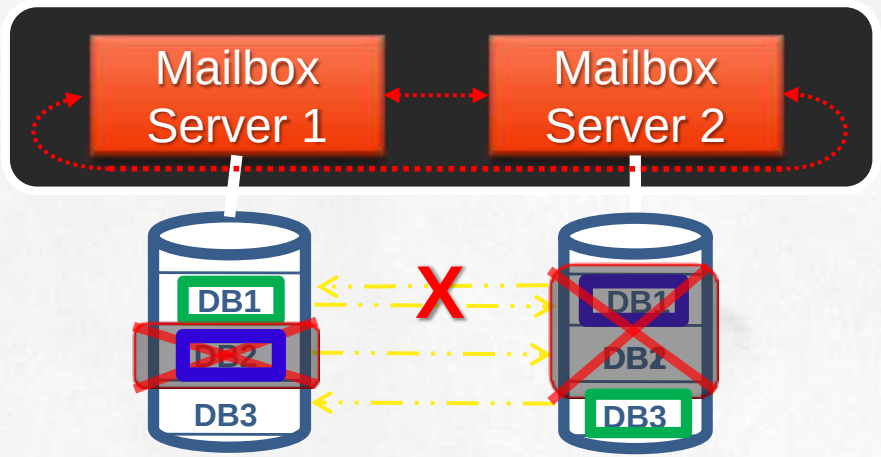
- Availability Terms
  - Active: Selected to provide email services to clients
  - Passive: Available to provide email services to clients if active fails
- Replication Terms
  - Source: Provides data for copying to a separate location
  - Target: Receives data from the source



Exchange Server 2010 High Availability Fundamentals

Mailbox Database Copy

- Scope of replication
- A copy is either source or target of replication at any given time
- A copy is either active or passive at any given time
- Only 1 copy of each database in a DAG is active at a time
- A server may not host >1 copy of a any database

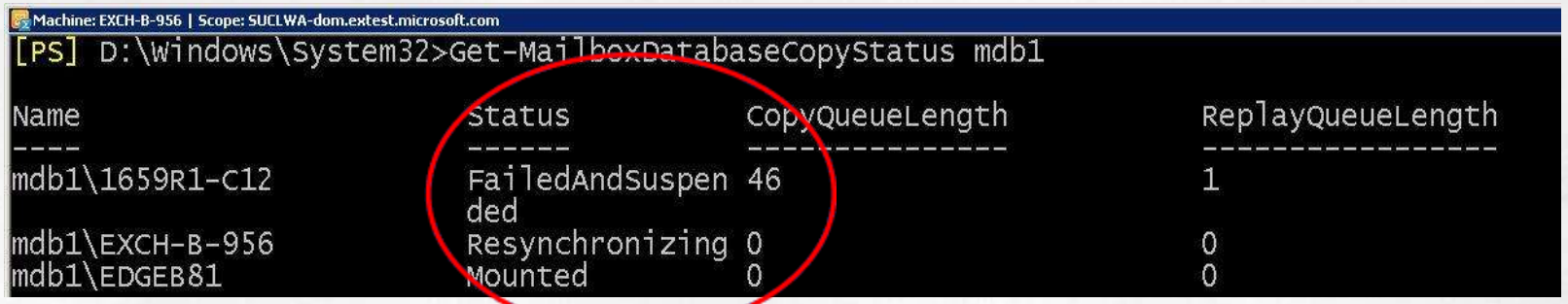


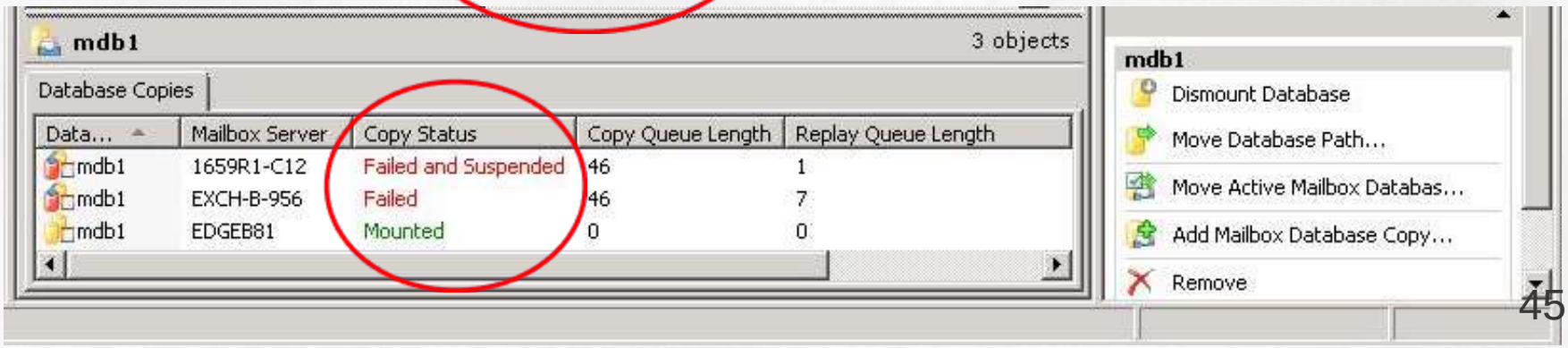
Exchange Server 2010 High Availability Fundamentals

Mailbox Database Copy

Defines properties applicable to an individual database copy

- Copy status: Healthy, Initializing, Failed, Mounted, Dismounted, Disconnected, Suspended, FailedandSuspended, Resynchronizing, Seeding
- CopyQueueLength
- ReplayQueueLength
- ActiveCopy
- ActivationSuspended





Exchange Server 2010 High Availability Fundamentals - Active Manager

- Exchange-aware resource manager (high availability's brain)
  - Runs on every server in the DAG
  - Manages which copies should be active and which should be passive
  - Definitive source of information on where a database is active or mounted
    - Provides this information to other Exchange components (e.g., RPC Client Access and Hub Transport)
    - Information stored in cluster database

Exchange Server 2010 High Availability Fundamentals - Active Manager

- Active Directory is still primary source for configuration info
- Active Manager is primary source for changeable state information (such as active and mounted)
- Replication service monitors health of all mounted databases, and monitors Extensible Storage Engine (ESE) for I/O errors or failure

Active Manager

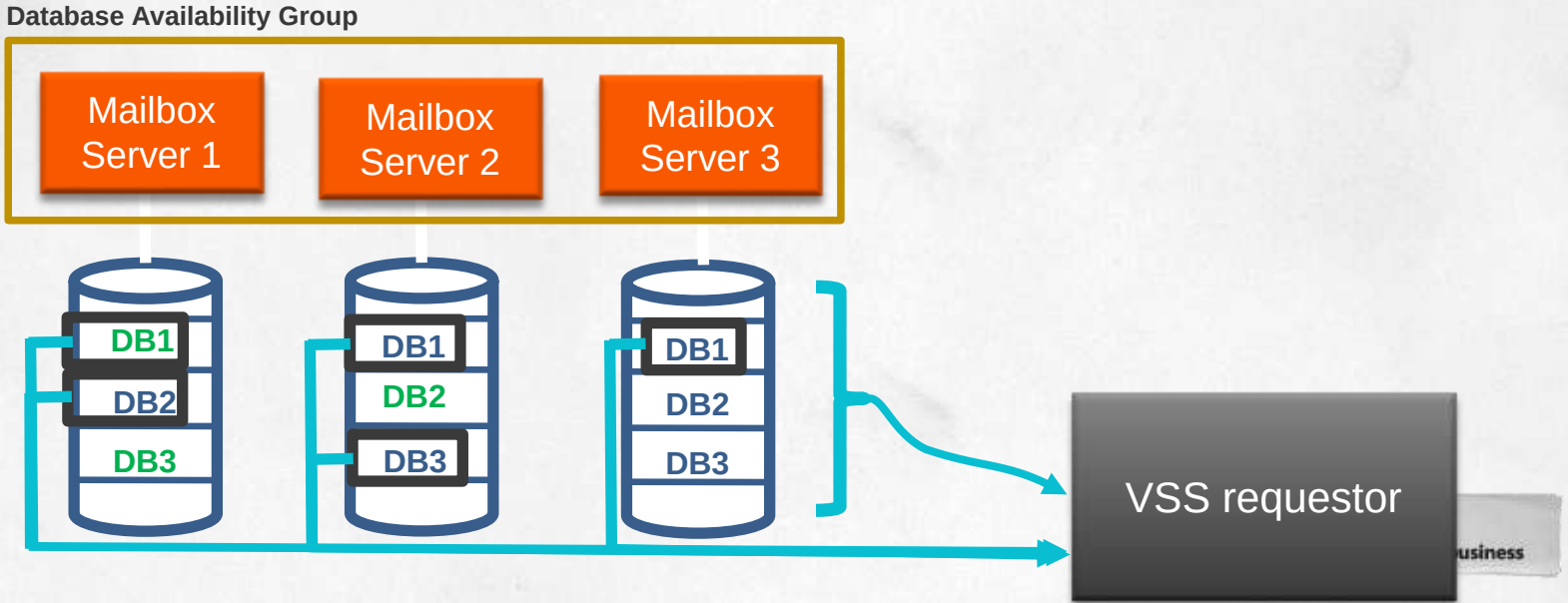
- Primary Active Manager (PAM)
  - Runs on the node that owns the default cluster group (quorum resource)
  - Gets topology change notifications
  - Reacts to server failures
  - Selects the best database copy on *overs
- Standby Active Manager (SAM)
  - Runs on every other node in the DAG
  - Responds to queries from other Exchange components for which server hosts the active copy of the mailbox database

Exchange Server 2010 High Availability Fundamentals
Continuous Replication

- Continuous replication has the following basic steps:
  - Database copy seeding of target
  - Log copying from source to target
  - Log inspection at target
  - Log replay into database copy

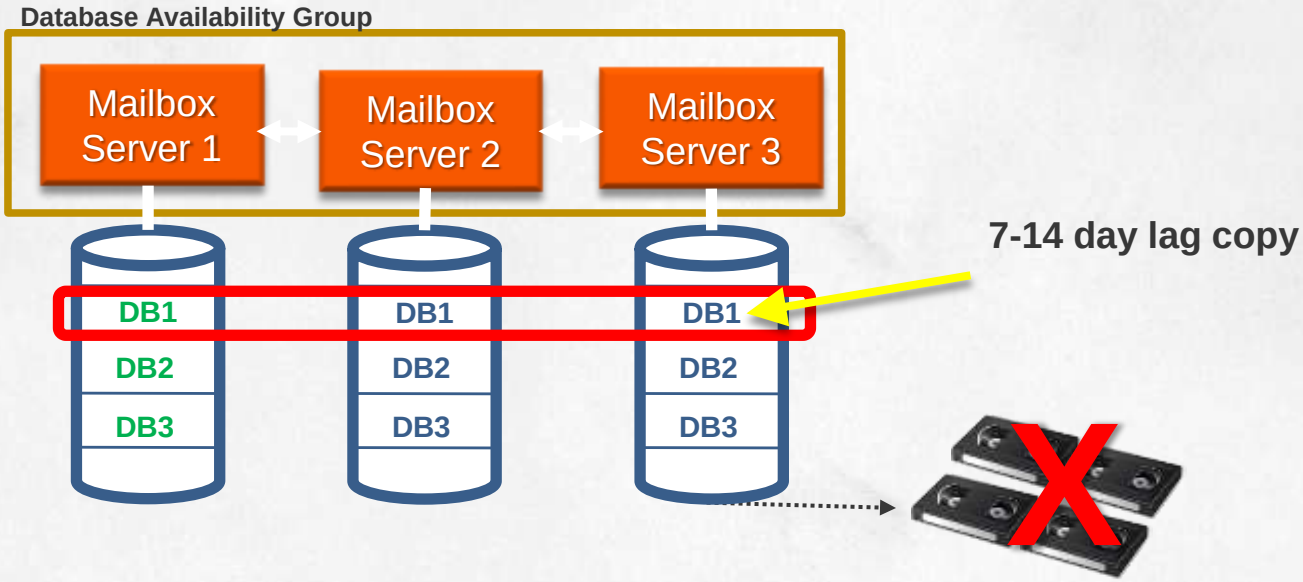
Exchange Server 2010 High Availability Fundamentals - Backups

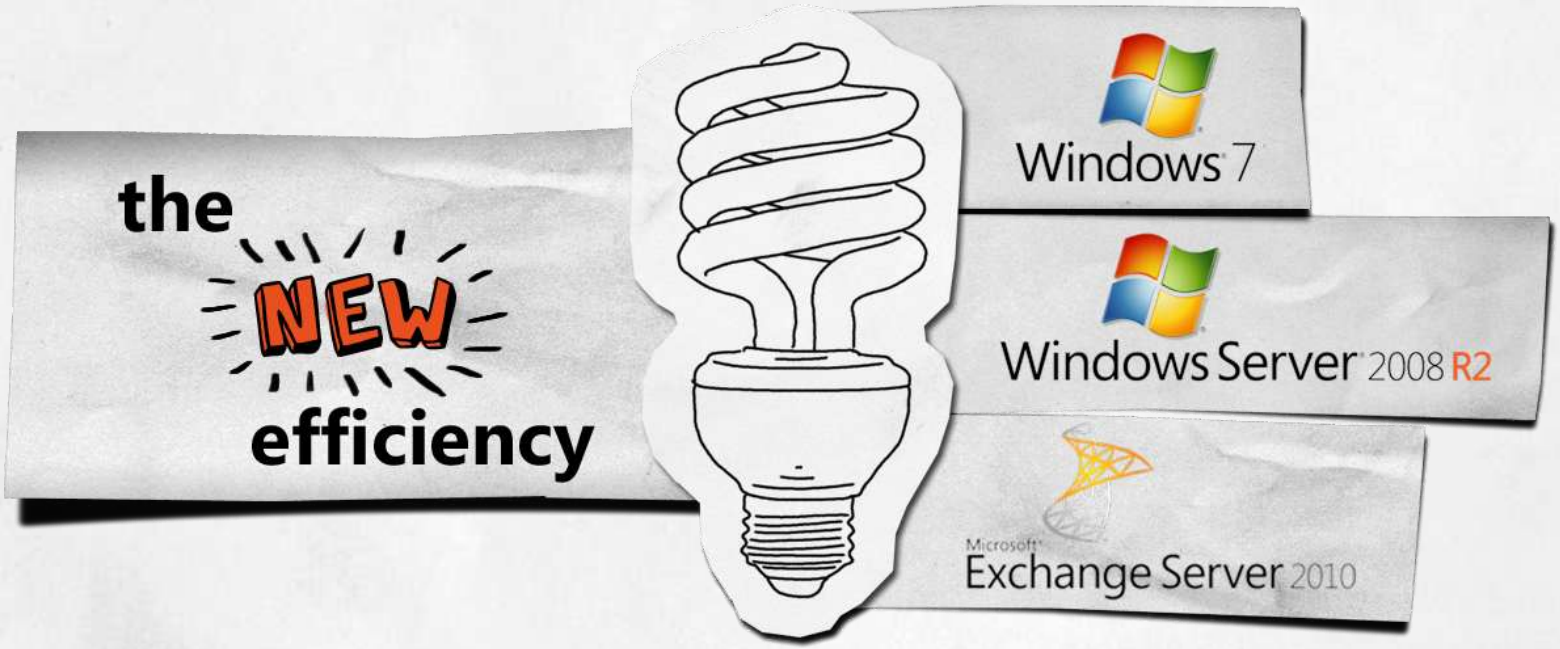
- Streaming backup APIs for public use have been cut, must use Volume Shadow Copy Service (VSS) for backups
  - Backup from any copy of the database/logs
  - Always choose Passive (or Active) copy
  - Backup an entire server
  - Designate a dedicated backup server for a given database
- Restore from any of these backups scenarios



Multiple Database Copies Enable Backupless Configurations

- Site/server/disk failure → Exchange Server 2010 HA
- Archiving/compliance → E-mail archive
- Recover deleted items → Extended/protected dumpster retention

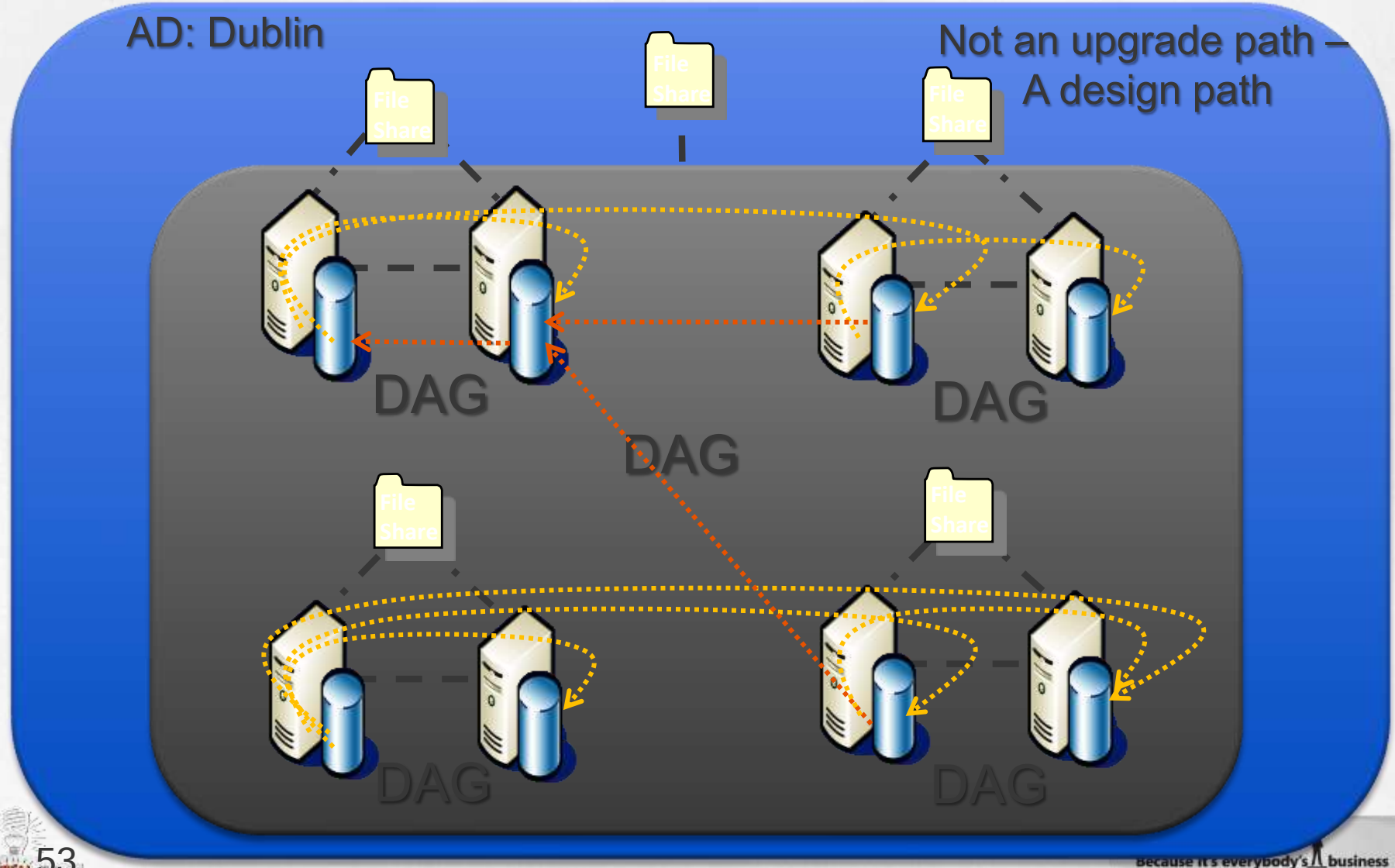




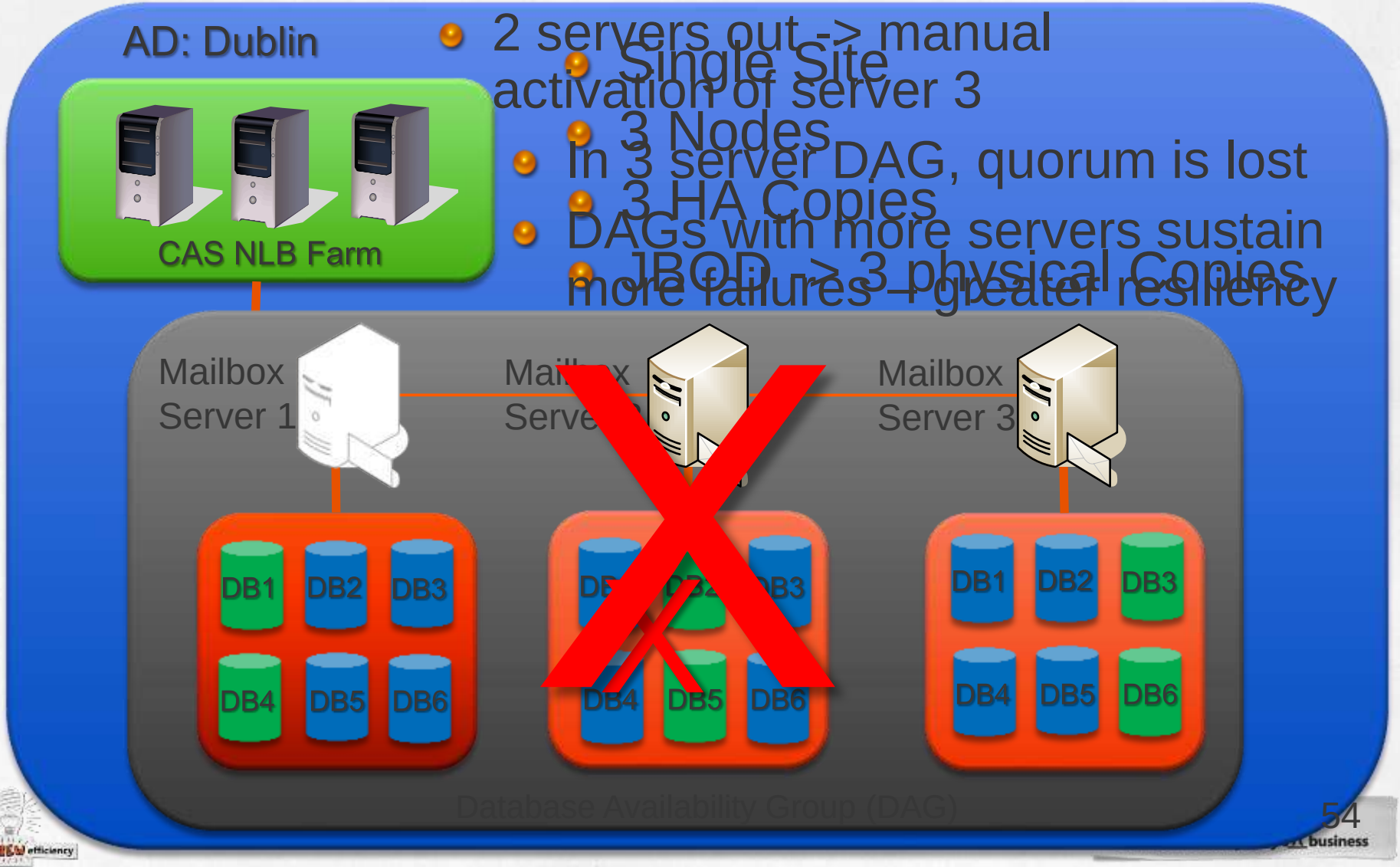
EXCHANGE SERVER 2010 HIGH AVAILABILITY DESIGN EXAMPLES

Because it's everybody's business

High Availability Design Example
CCR Design -> DAG Design



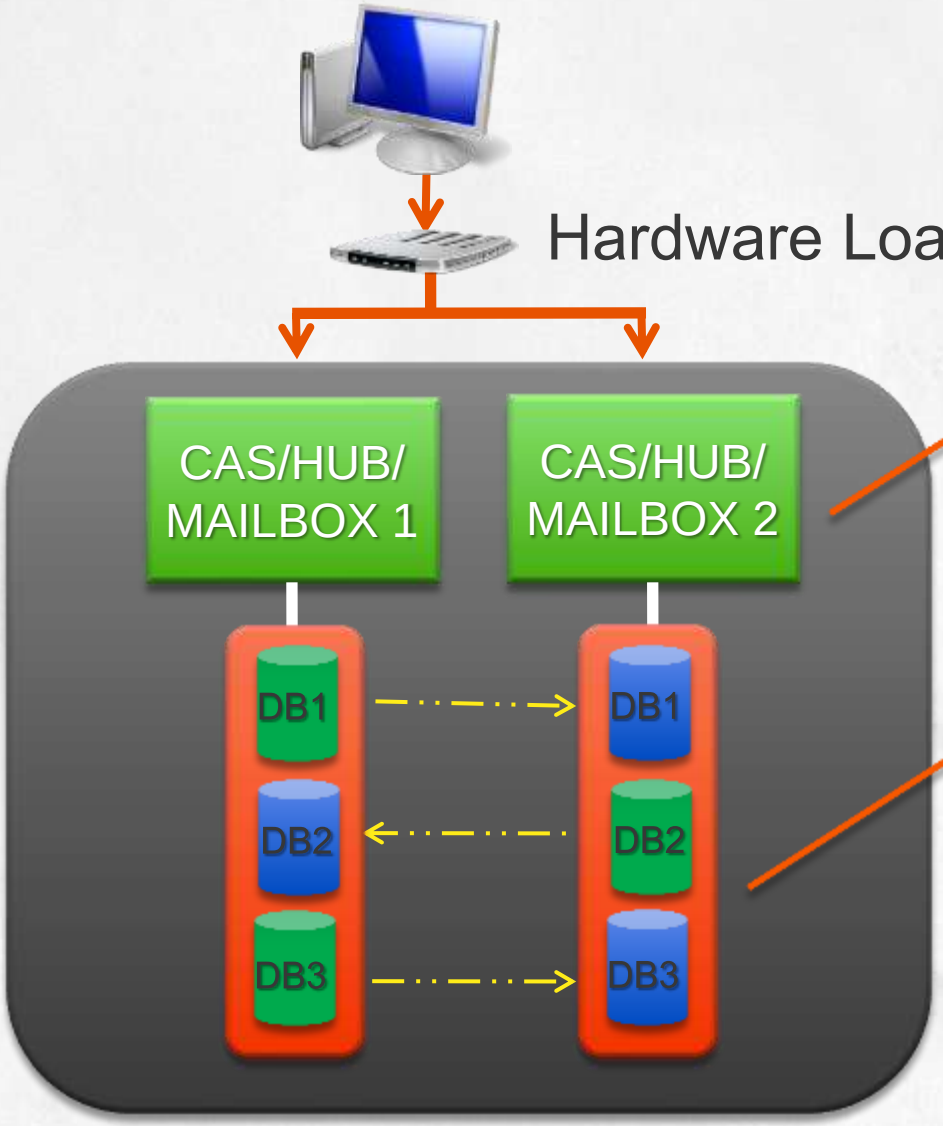
High Availability Design Example
Double Resilience – Maintenance + DB Failure



- 2 servers out -> manual activation of server 3
- In 3 server DAG, quorum is lost
- DAGs with more servers sustain more failures – greater resiliency

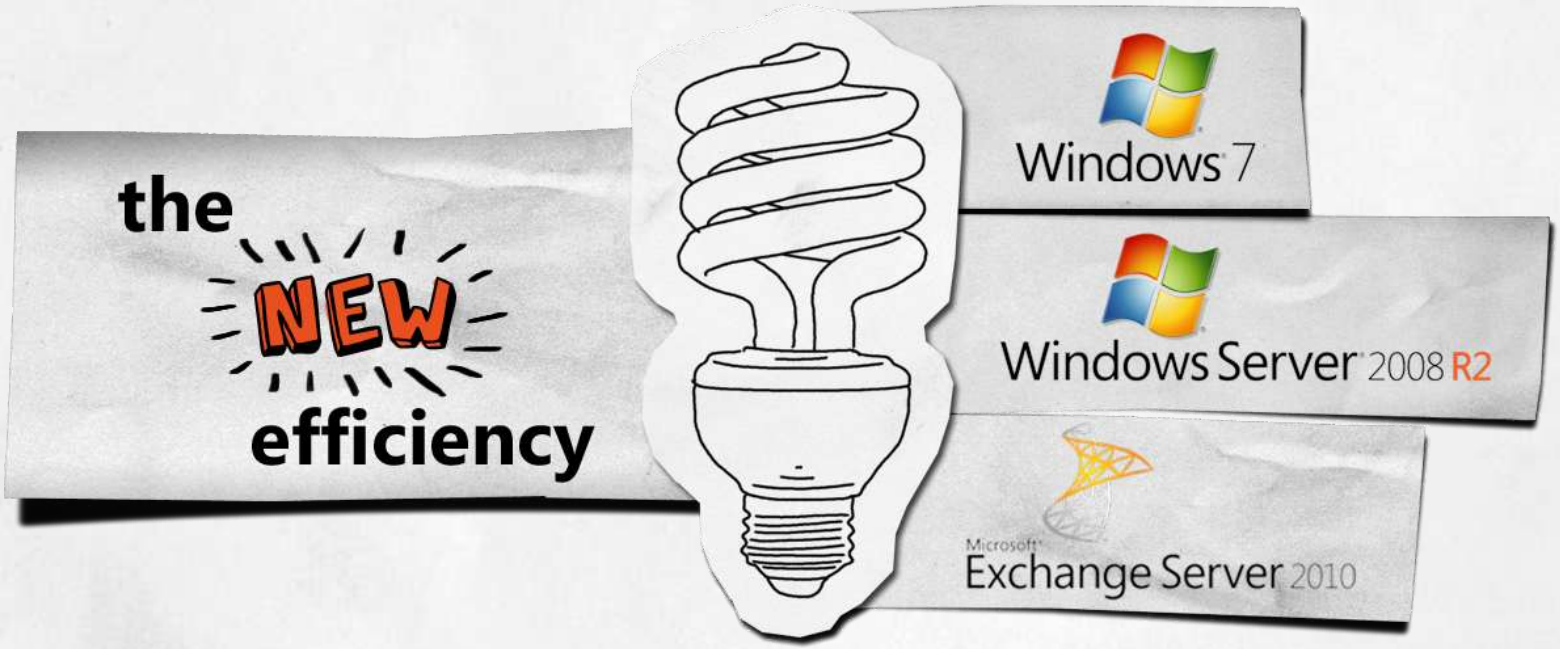
- Single Site
- 3 Nodes
- 3 HA Copies
- JBOD -> 3 physical Copies

High Availability Design Example
Branch Office or Smaller Deployment



Member servers of DAG can host other server roles

2 server DAGs, with server roles combined or not, should use RAID



EXCHANGE SERVER 2010 SITE RESILIENCE

Because it's everybody's business

Exchange Server 2010 *Over Cases

- Within a datacenter
  - Database *over
  - Server *over
- Between datacenters
  - Single database *over
  - Server *over
- Datacenter failover (which is really a switchover)

Single Database Cross-Datacenter *Over

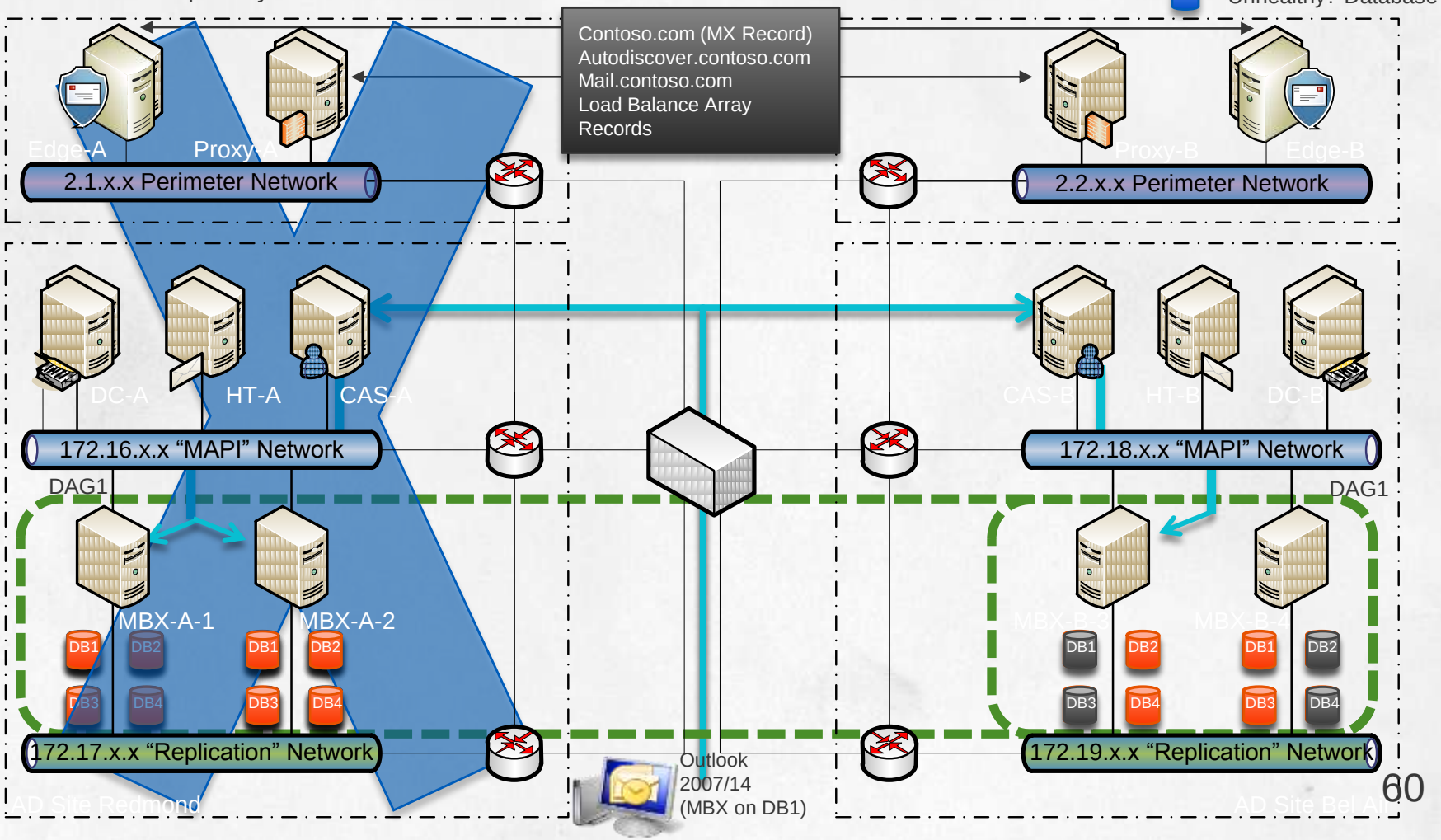
- Database mounted in another datacenter and another Active Directory site
- Serviced by “new” Hub Transport servers
- “Different OwningServer” – for routing
  - Transport dumpster re-delivery now from both Active Directory sites
- Serviced by “new” CAS
  - “Different CAS URL” – for protocol access
  - Outlook Web App now re-directs connection to second CAS farm
  - Other protocols proxy or redirect (varies)

Datacenter Failover

- Customers can evolve to site resilience
- Standalone → local redundancy → site resilience
- Consider name space design at first deployment
- Keep extending the DAG!
- Monitoring and many other concepts/skills just re-applied
- Normal administration remains unchanged
- Disaster recovery not High Availability (HA) event

Failure Scenario – Primary Data Center Failure

1. Verify primary data center is capable of hosting service
2. Activate standby WSFC nodes in primary data center
3. Restore DAG quorum – Stop-DatabaseAvailabilityGroup DAG1 –ActiveDirectorySite Redmond –ConfigurationOnly (in both data centers)
4. Stop the cluster service on the nodes in primary data center
5. Restore database availability group – Restore-DatabaseAvailabilityGroup DAG1 –ActiveDirectorySite "Bel Air" –AlternateWitnessDirectory c:\DAG1 –AlternateWitnessServer HT-B
6. Change MX records and HTTP access to primary data center
7. Mount databases in primary data center
8. Move databases back to primary data center (no activation block scenario)



Takeaways

- With each release, our goals are to make Exchange high availability:
  - Easier and cheaper to deploy
  - Easier and cheaper to manage
  - Support better SLAs with faster and more granular recoveries
  - Improve site resiliency support
- Our other goal is for highly available deployments to be mainstream!

Exchange 2010 High Availability

One Technology for High Availability, Disaster Recovery and Backup

- Incremental Deployment
- Simplified Administration
- Granular failover & recovery
- Improved user uptime

Exchange Server 2010 High Availability Fundamentals
Database Seeding

- There are three ways to seed the target instance:
  - Automatic Seeding
    - Requires 1st log file containing CreateDB record
  - Update-MailboxDatabaseCopy cmdlet
    - Can be performed from active or passive copies
  - Manually copy the database

Exchange Server 2010 High Availability Fundamentals
Log Shipping

- Log shipping in Exchange Server 2010 leverages Transmission Control Protocol (TCP) sockets
  - Supports encryption and compression
  - Administrator can set TCP port to be used
- Replication service on target notifies the active instance the next log file it expects
  - Based on last log file which it inspected
- Replication service on source responds by sending the required log file(s)
- Copied log files are placed in the target's Inspector directory

Exchange Server 2010 High Availability Fundamentals
Log Inspection

- The following actions are performed to verify the log file before replay:
  - Physical integrity inspection
  - Header inspection
  - Move any Exx.log files to ExxOutofDate folder that exist on target if it was previously a source
- If inspection fails, the file will be recopied and inspected (up to 3 times)
- If the log file passes inspection it is moved into the database copy's log directory

Exchange Server 2010 High Availability Fundamentals - Log Replay

- Log replay has moved to Information Store
- The following validation tests are performed prior to log replay:
  - Recalculate the required log generations by inspecting the database header
  - Determine the highest generation that is present in the log directory to ensure that a log file exists
  - Compare the highest log generation that is present in the directory to the highest log file that is required
  - Make sure the logs form the correct sequence
  - Query the checkpoint file, if one exists
- Replay the log file using a special recovery mode (undo phase is skipped)

Exchange Server 2010 High Availability Fundamentals - Lossy Failure Process

- In the event of failure, the following steps will occur for the failed database:
  - Active Manager will determine the best copy to activate
  - The Replication service on the target server will attempt to copy missing log files from the source - ACLL
    - If successful, then the database will mount with zero data loss
    - If unsuccessful (lossy failure), then the database will mount based on the AutoDatabaseMountDial setting
  - The mounted database will generate new log files (using the same log generation sequence)
  - Transport Dumpster requests will be initiated for the mounted database to recover lost messages
  - When original server or database recovers, it will run through divergence detection and perform an incremental reseed or require a full reseed

Exchange Server 2010 High Availability Fundamentals

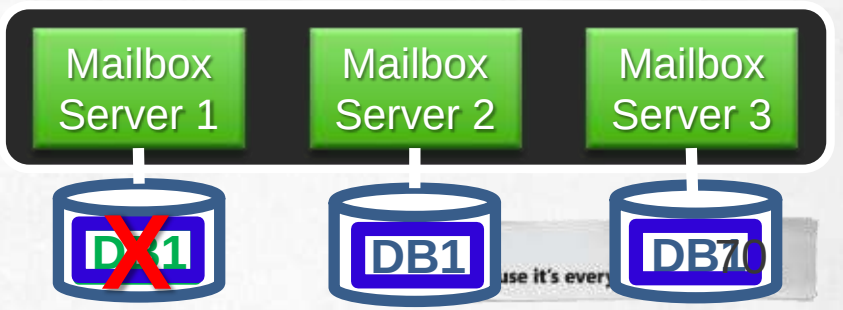
Active Manager Selection of Active Database Copy

- Active Manager selects the “best” copy to become active when existing active fails

Catalog	Healthy
Copy status	Healthy, DisconnectedAndHealthy, DisconnectedAndResynchronizing,
or	
	SeedingSource
ReplayQueueLength	< 10 < 50
ReplayQueueLength	< 50

Exchange Server 2010 High Availability Fundamentals - Incremental Resync

- Incremental reseed scenario
  - Active DB1 on server1 fails
  - Passive DB1 on server3 takes over service
  - Sometime later, failed DB1 on server1 comes back as passive – contains inconsistent data
  - Make DB1 on server1 consistent with new active
- Transaction logs of active and failed copy are compared to find divergence point
- Determines from logs the database pages that changed after divergent point
- Copies database pages from active to failed copy, then play new logs, until in-sync
- Replaces Exchange Server 2007's Lost Log Resilience (LLR)
  - LLR is set to 1



Split Brain Management

- Two datacenter *overs have a risk of split brain
- Primary datacenter power outage is classic example
- Exchange Server 2010 datacenter failovers maintain DAG membership but shrink cluster membership to create a new, “available topology” in the standby datacenter
- Exchange Server 2010 provides a safe answer with datacenter activation coordination (DAC) mode
  - Requires a DAG with three nodes
  - Requires activation in partial datacenter failure cases is “done right”
    - Mailbox servers must be “stopped” or powered off
  - Implements a “Mommy may I protocol” before active manager mounts databases

Split Brain Management (Cont'd)

- If DAC is not enabled, the DAG will not restart and mount databases until a majority of servers are restored
- If DAC is enabled, the "Mommy May I Protocol" is used to coordinate with Active Managers in DAG to determine state and recoverability
- There are several requirements that must be satisfied to prevent split brain between datacenters after datacenter failover