



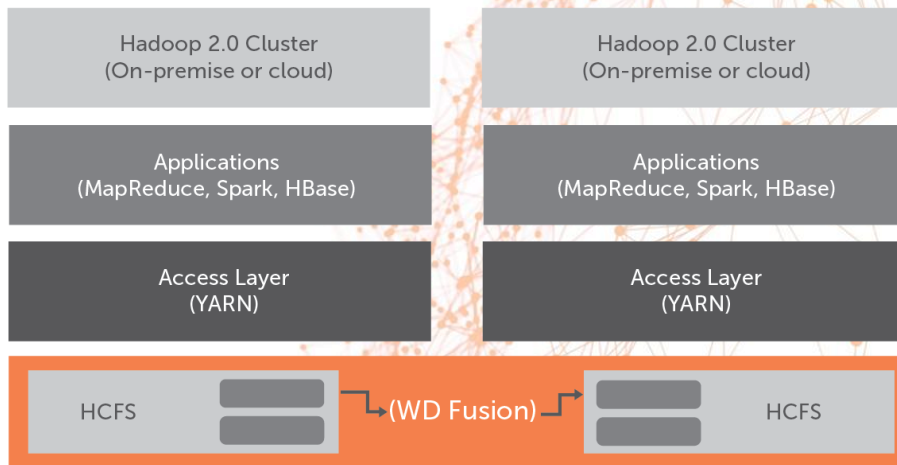
The **LIVE DATA** Company

WANdisco Fusion

Technical overview

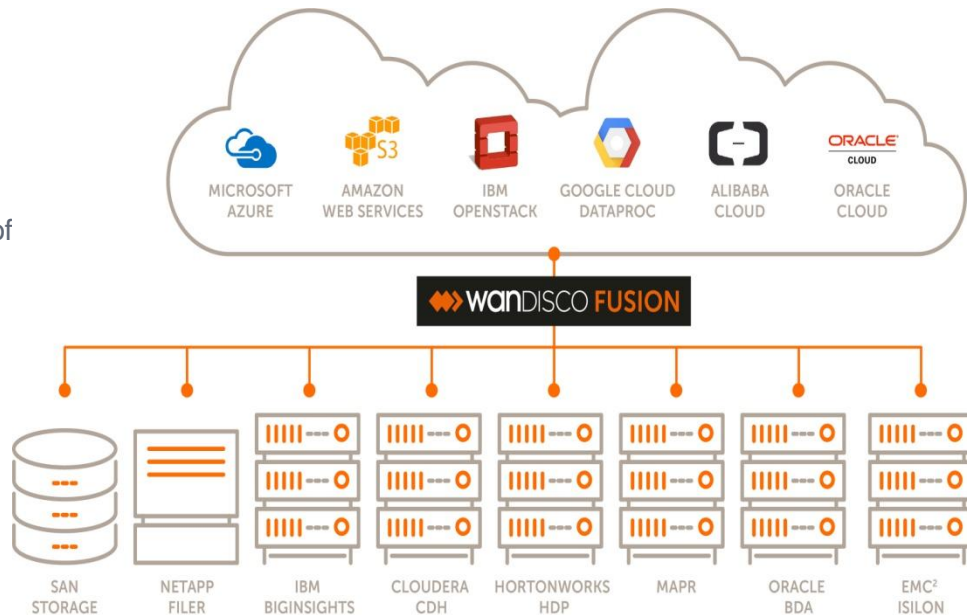
WHAT IS WANDISCO FUSION?

- Active-active replication of Hadoop data across multiple appliances, clusters, distributions, clouds.
- WAN-capable for geo-replication
 - Selective, opt-in, cross-distribution
- Enabler for extending Big Data solutions – on premise and cloud
 - Can replicate between on-premises and cloud environments or cloud to cloud
 - Enables bursting from on-premises hardware into Cloud Services
 - Enables adoption of Cloud services
- Supports all major Hadoop distributions and versions, operating as an application without change to the underlying cluster



WHAT DOES FUSION DO?

- ◆ Replicates data automatically with guaranteed consistency across Hadoop clusters running on any supported distribution, cloud object storage and local and NFS mounted file systems
- ◆ Patented distributed coordination engine enables:
 - Guaranteed data consistency across any number of sites at any distance
 - Minimized RTO/RPO
- ◆ Totally non-invasive
 - No modification to source code
 - Easy to turn on/off



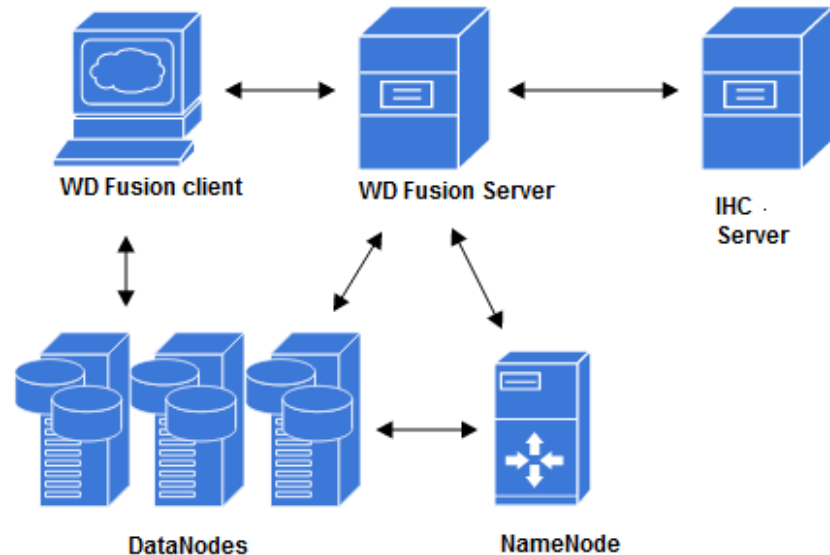


Fusion Components

Three Fusion components are deployed to each cluster, aka “zone”:

- Fusion Server
- IHC Server
- Fusion Client
- Fusion UI

Typical Data Center Configuration





Fusion Server

- Fusion Server is the core piece of WANDisco Fusion
- Proxies file write operations from clients
- Coordinates changes across clusters using DConE
- One Fusion server required per zone; generally 2 (or more) deployed for high availability and scalability, 3 preferred for quorum
- Should not be co-located on a Datanode. Data replicated to a destination, will fill Datanode storage first due to block placement policy.
- Typically the same system runs the Fusion UI, a web server for browser-based administration



IHC Server

- Inter Hadoop Communication (IHC) servers handle traffic between zones
- When data is being replicated, an IHC server reads it from the local zone and streams it to a Fusion server in the remote zone. That Fusion server writes to its local zone. This allows all reads and writes to be performed as “local” operation. No special cross-zone security is required.
- One IHC server required per zone; generally 2 (or more) deployed for high availability and scalability, 3 preferred for quorum.
- Automatically deployed on the same node as the Fusion Server.
- IHC is split out from the Fusion server to isolate the components that depend on HDFS distro; also increases scalability



Fusion Client

- Set of .jar files installed on each Hadoop client, such as mappers, reducers, and gateway nodes
 - Not a service or agent
 - Deployed with Stacks (Amabri) or Parcels (Cloudera Manager) or manual RPM installation on each node (management and data)
- Configurations tell apps to use Fusion classes for HDFS operations
- These classes redirect all write requests through the Fusion server in most cases. Replicated writes are coordinated.
 - Configurations determine whether, and how, clients contact the Fusion server



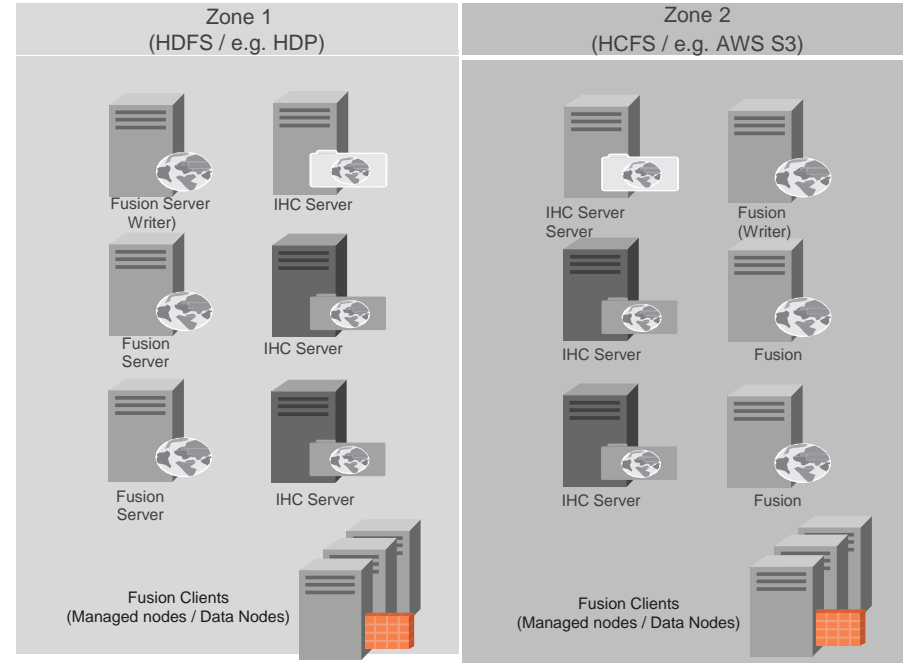
Fusion UI

- A separate server that provides administrators with a browser-based management console for each WANdisco Fusion server.
- Can be installed on the same machine as WANdisco Fusion's server or on a different machine within your data center.

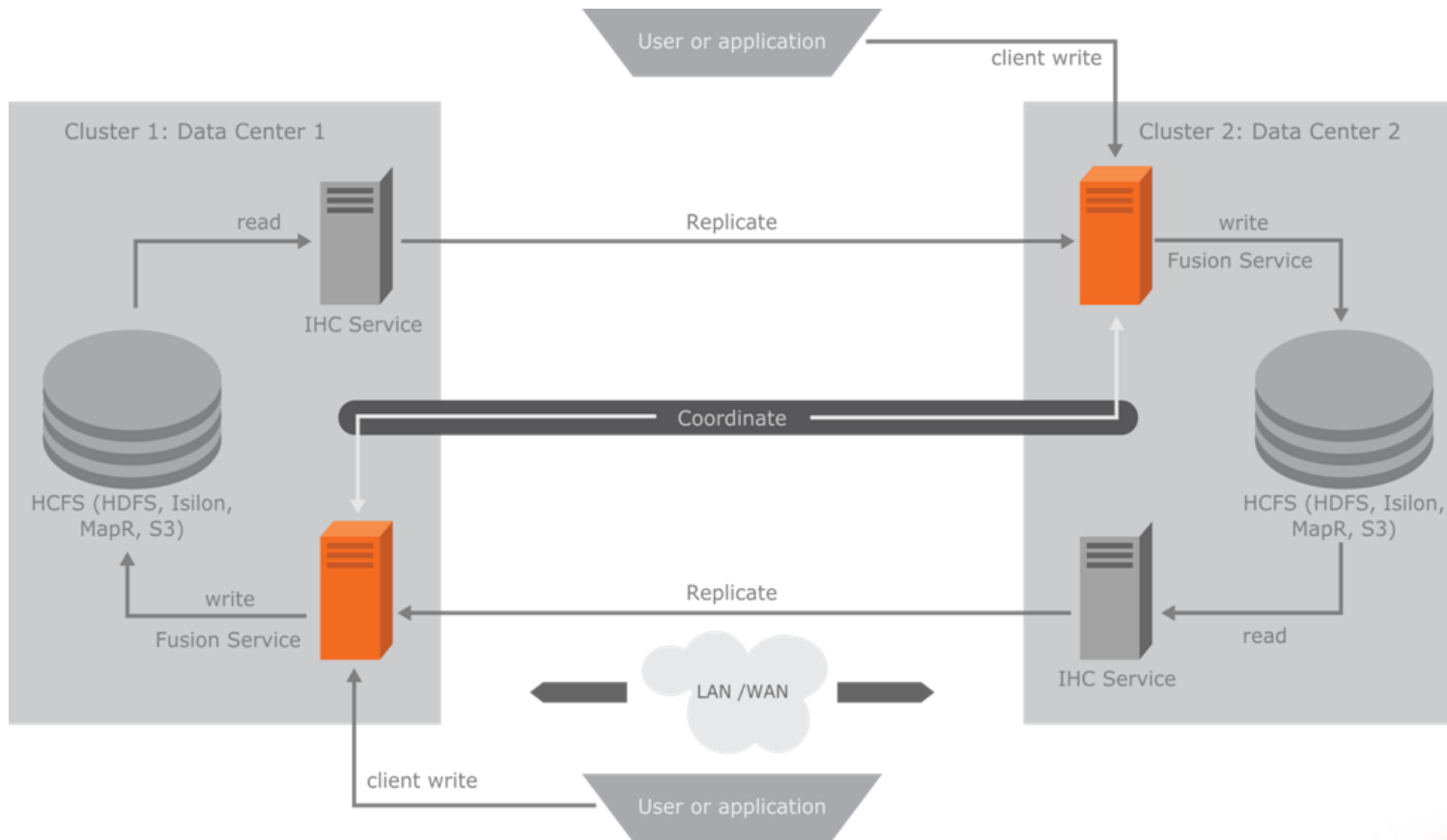


Architecture Components

- Each cluster is a “zone”
- Zones can have multiple Fusion servers and IHC servers
- A leader is elected for each replicated folder between zones (Priority Zone)
- A writer is elected for each replicated path per zone.
- Clients failover between Fusion servers for writes
- Fusion servers select between multiple IHCs for reading



Breakout Architecture





Architecture Principles

Strict consistency of metadata with fast data ingest

- Synchronous replication of metadata between data centers
 - Using Coordination Engine
 - Provides strict consistency of the namespace
- Asynchronous replication of data over the WAN
 - Data streamed in the background
 - Data can be transferred in file, block, or sub-block increments
 - Allows fast LAN-speed data creation and transfer without network spikes
 - Maximum amount of WAN bandwidth used for replication can also be configured
 - Small and large file transfers happen in parallel



Bypassing Replication

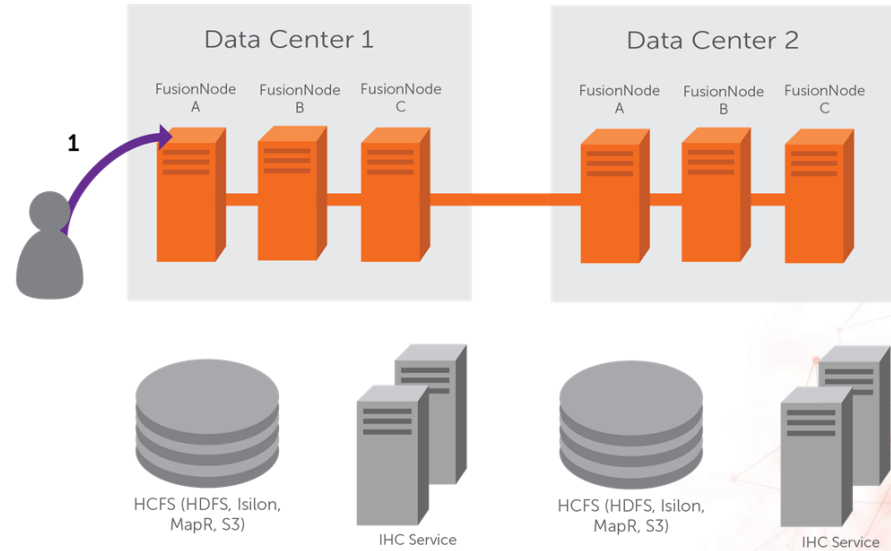
- Bypass is an operation that informs the client to skip the connection to the Fusion server if it is not available
- Eliminates reliance on Fusion server
- No single point of failure
- Disable HDFS client
 - Sets cluster to passive mode, removes writes from DConE process
 - Makes replication more manual
 - DR and migration use cases



HDFS ↔ HDFS

Fusion workflow

1. User makes a request to create/change a file

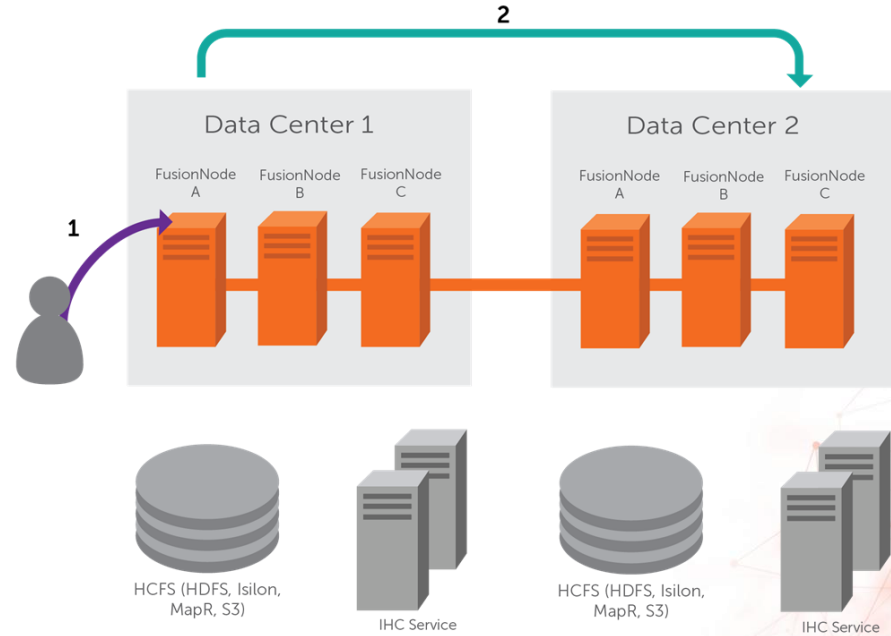




HDFS ↔ HDFS

Fusion workflow

1. User makes a request to create/change a file
2. Fusion coordinates File operation to other clusters involved

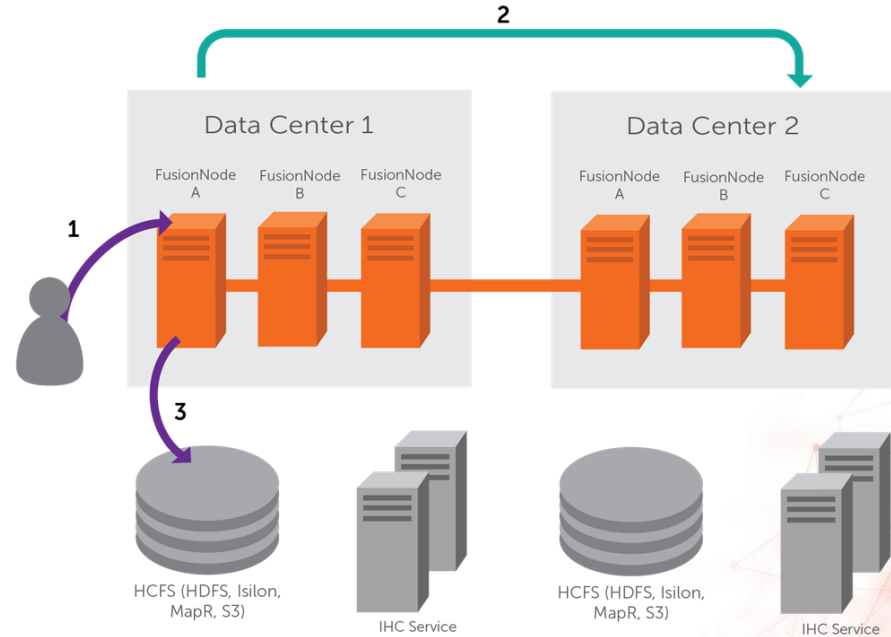




HDFS ↔ HDFS

Fusion workflow

1. User makes a request to create/change a file
2. Fusion coordinates File operation to other clusters involved
3. File is added to underlying storage

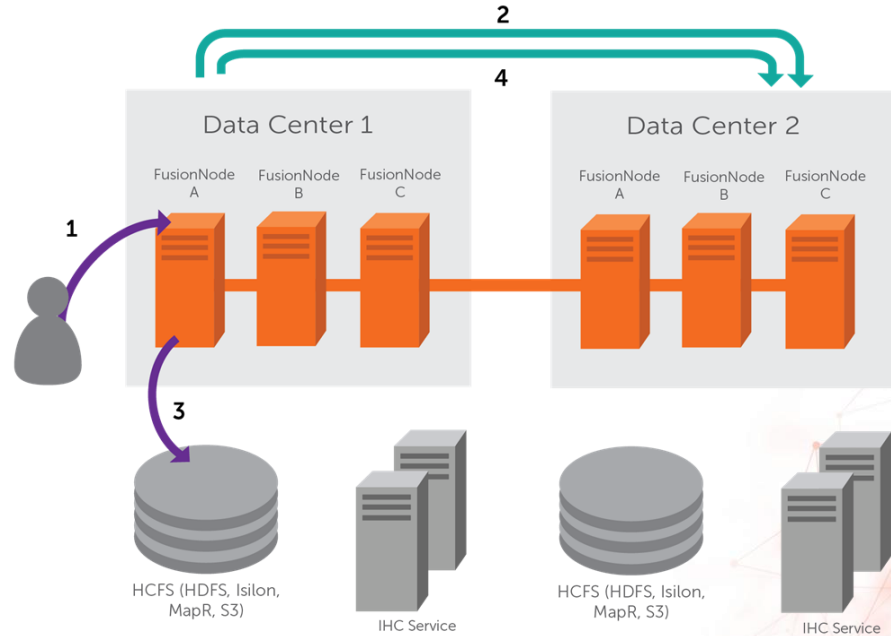




HDFS ↔ HDFS

Fusion workflow

1. User makes a request to create/change a file
2. Fusion coordinates File operation to other clusters involved
3. File is added to underlying storage
4. Fusion coordinates at configurable write increments and File Close with other clusters

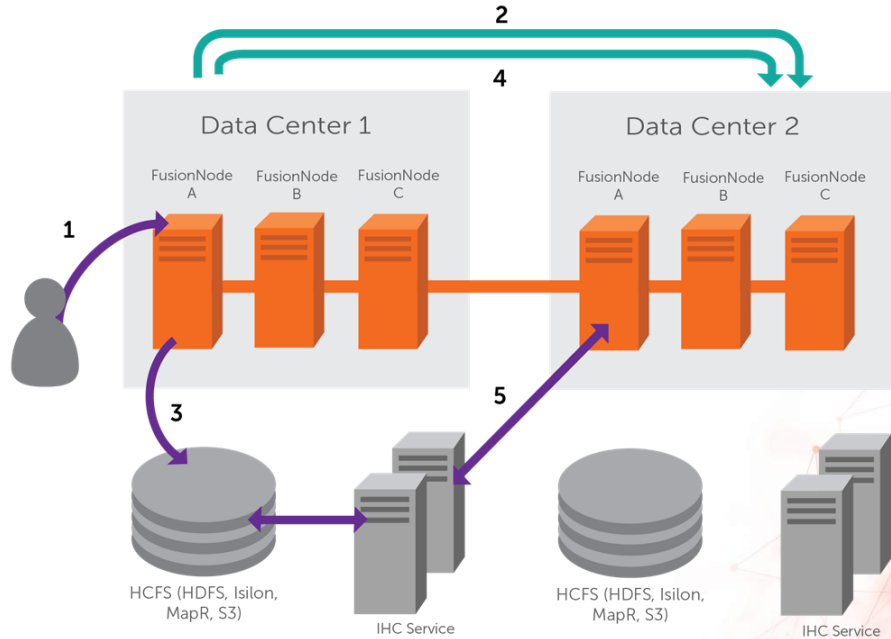




HDFS ↔ HDFS

Fusion workflow

1. User makes a request to create/change a file
2. Fusion coordinates File operation to other clusters involved
3. File is added to underlying storage
4. Fusion coordinates at configurable write increments and File Close with other clusters
5. Fusion server at remote cluster pulls data from IHC server on source cluster

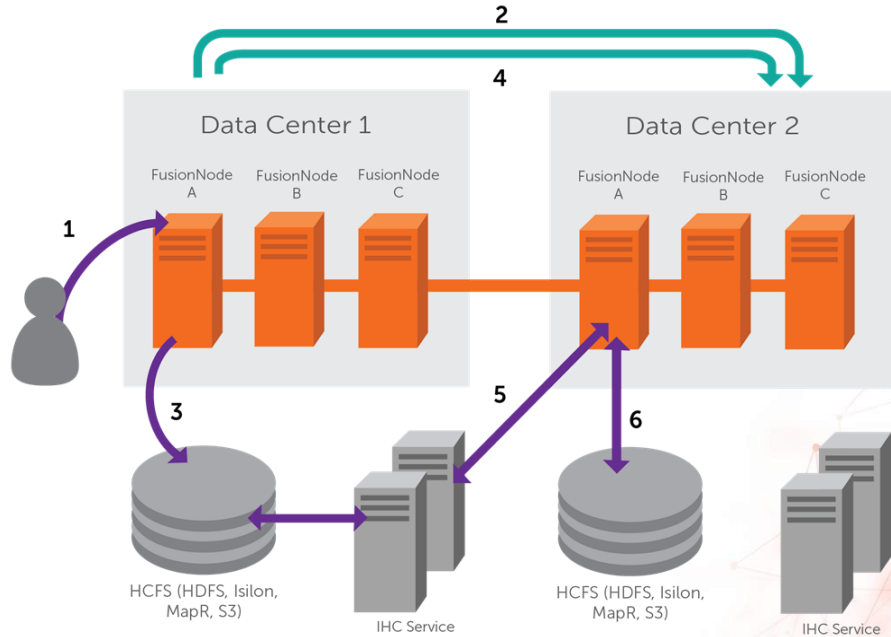




HDFS ↔ HDFS

Fusion workflow

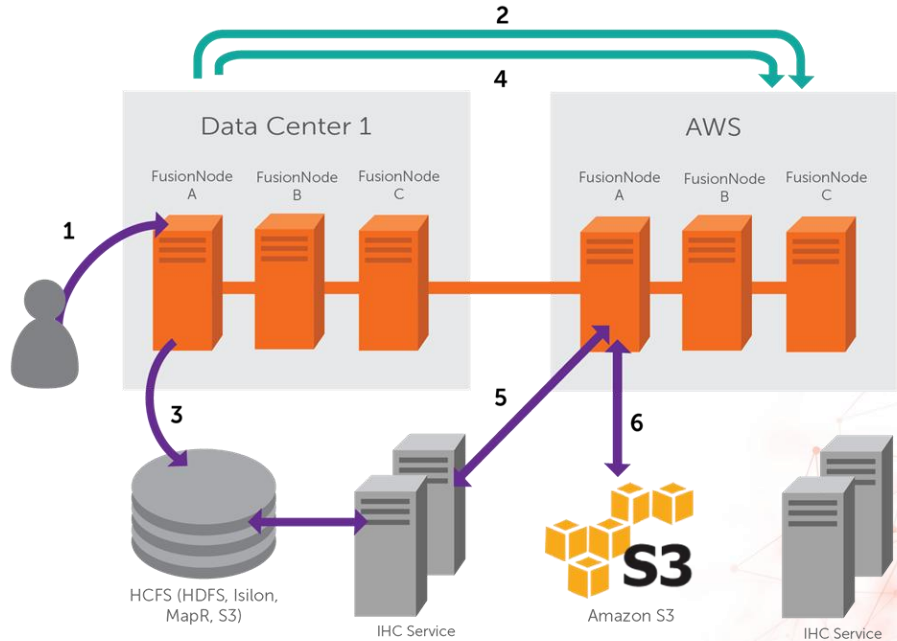
1. User makes a request to create/change a file
2. Fusion coordinates File operation to other clusters involved
3. File is added to underlying storage
4. Fusion coordinates at configurable write increments and File Close with other clusters
5. Fusion server at remote cluster pulls data from IHC server on source cluster
6. Fusion server at remote site writes data to its local cluster





HDFS ↔ HCFS (eg: S3)

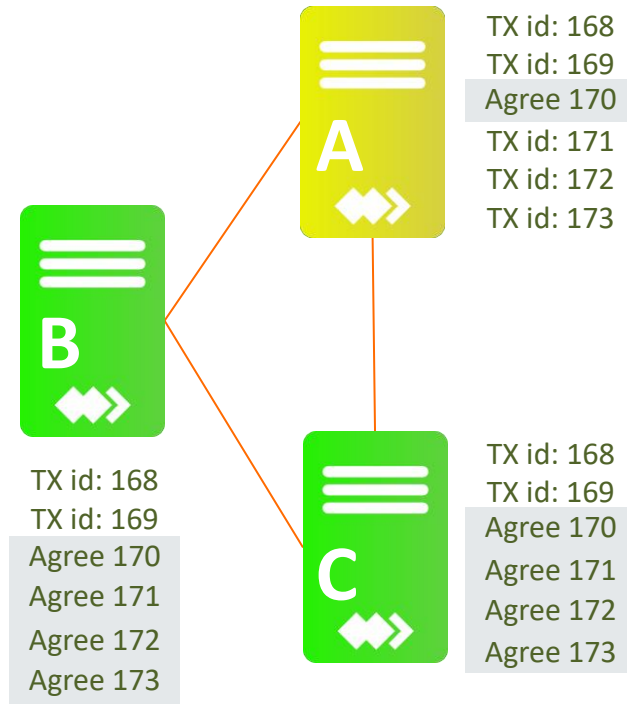
- Similar architecture to HDFS to HDFS
- Fusion server(s) run in EC2
- Works directly with S3 for migration
- EMR cluster is optional
- Customer does not pay for compute
- IHC handles translation between HDFS and S3
- Works with EMR for compute use cases
- Supports EMR's ability to spin up and shut down dynamically
- Distributes Fusion client to every node at EMR spins up
- Similar architecture with other cloud vendors



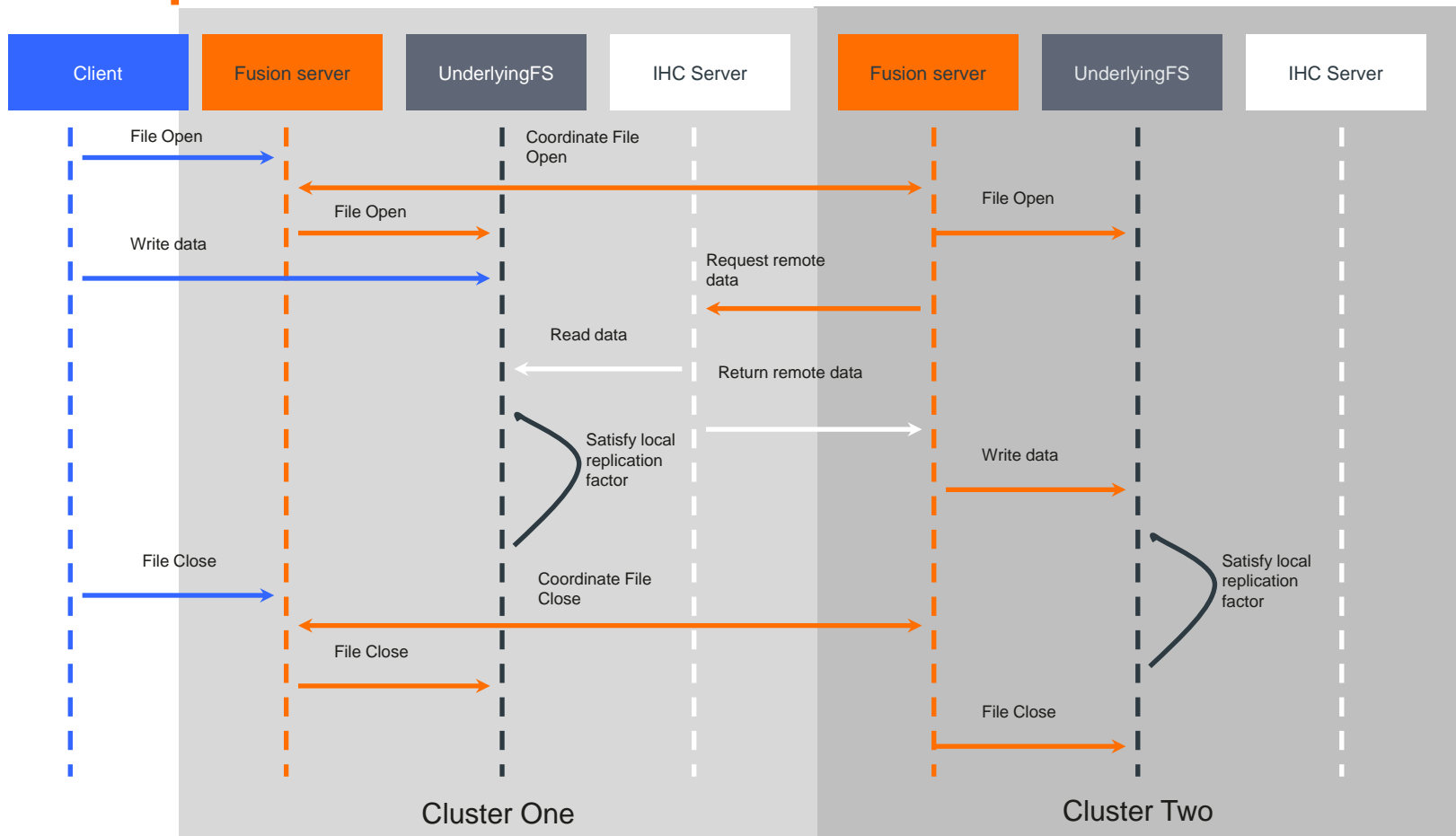
How DConE Works

WANdisco Active/Active Replication

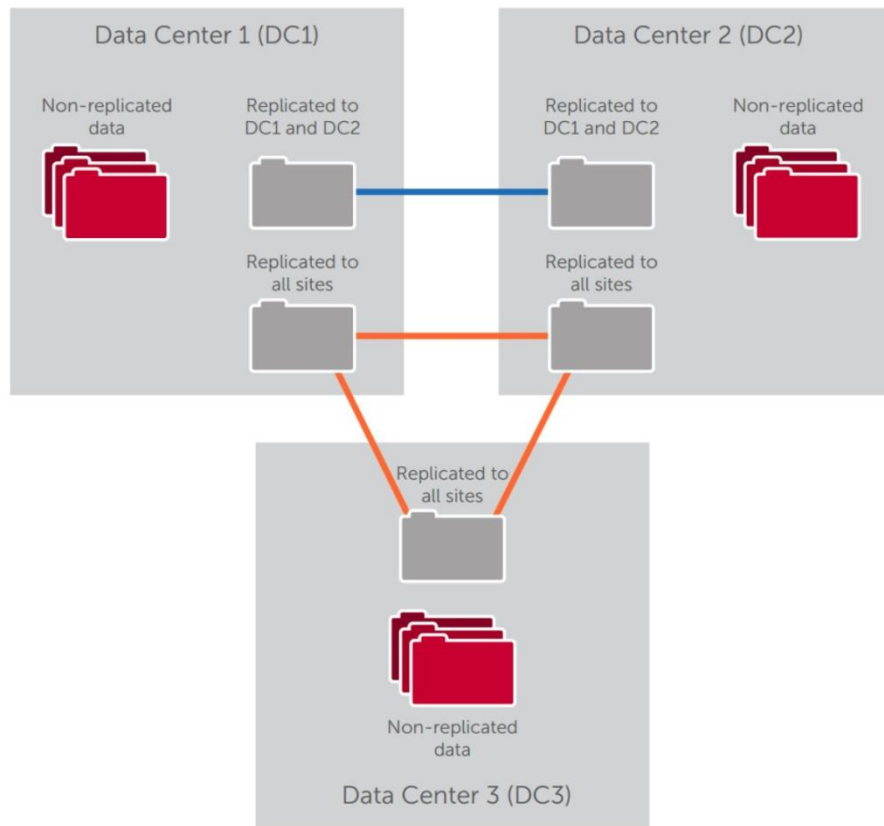
- ◆ Majority Quorum
 - A fixed number of participants
 - The Majority must agree for change
- ◆ Failure
 - Failed nodes are unavailable
 - Normal operation continue on nodes with quorum
- ◆ Recovery / Self Healing
 - Nodes that rejoin stay in safe mode until they are caught up
- ◆ Disaster Recovery
 - A complete loss can be brought back from another replica



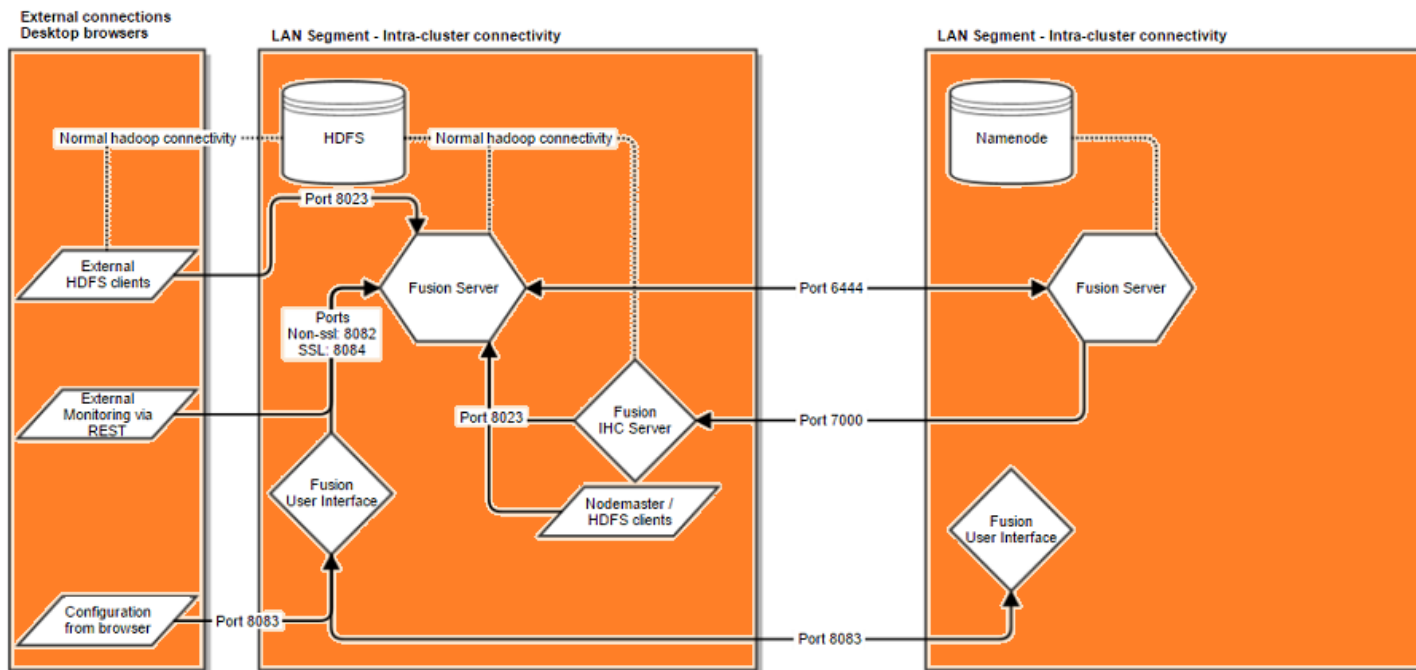
Write path



Fusion Replicate Data on Folder Levels



Network Ports Used By Fusion





Installation

- Per Platform
 - Debian
 - RedHat and derivatives (eg; CentOS)
- Per Storage Type or Distribution
 - Cloudera
 - Hortonworks
 - Cloud Storage
 - OpenStack/IBM, Google, Microsoft, Amazon, Alicloud
 - LocalFS (Mounted NFS/NAS/CIFS/etc or Local)
- Hive or Standard
 - Only applies to Hadoop distros
- Example – Fusion on RedHat for Cloudera with Hive replication.
- Example – Fusion on Ubuntu for Azure WASB without replication.
- Example – Fusion on LocalFS without Hive Replication



System Requirements

- Varies by scale
- High Availability – Requires 2 or more servers, numbers can vary by location
- Can be Virtual or Physical
- Hadoop – Should be part of the cluster, client role only
- Isolated machines “Edge Nodes” are not required – However, highly recommended.
 - Isolated machines allow Fusion to use its own resources for replication and not contend with existing services in the cluster
 - Engineered systems (eg; Oracle BDA) with no Edge Nodes.
- Software Reqs – Falls in line with Hadoop requirements (Java 1.7/1.8, Perl)
- Hardware Reqs – These are average system specs
 - 8 CPU
 - 32 to 64GB RAM
 - 200 to 1TB Disk - The larger the disk the longer Fusion can ‘cache’ transactions for self healing (currently with equal heap requirements)
 - See docs for specific details and updates -
https://docs.wandisco.com/bigdata/wdfusion/2.11/#_prerequisites_checklist



Hardware Sizing Guidelines

- Low End Scale (under 100TB/year replication)
 - 4CPU, 32GB RAM
- Large Scale (~100TB/year replication)
 - 8CPU, 64GB-128GB RAM
- Billion Object / Petabytes Scale
 - 16CPU, 128GB-256GB RAM



The **LIVE DATA** Company

ANY QUESTIONS?



The **LIVE DATA** Company

THANK YOU