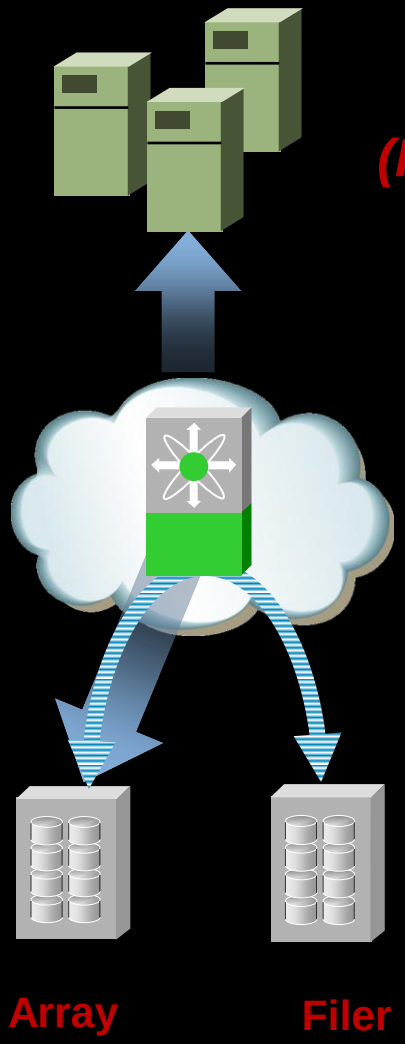
Data Compression and Deduplication

Data Redundancy Elimination Landscape

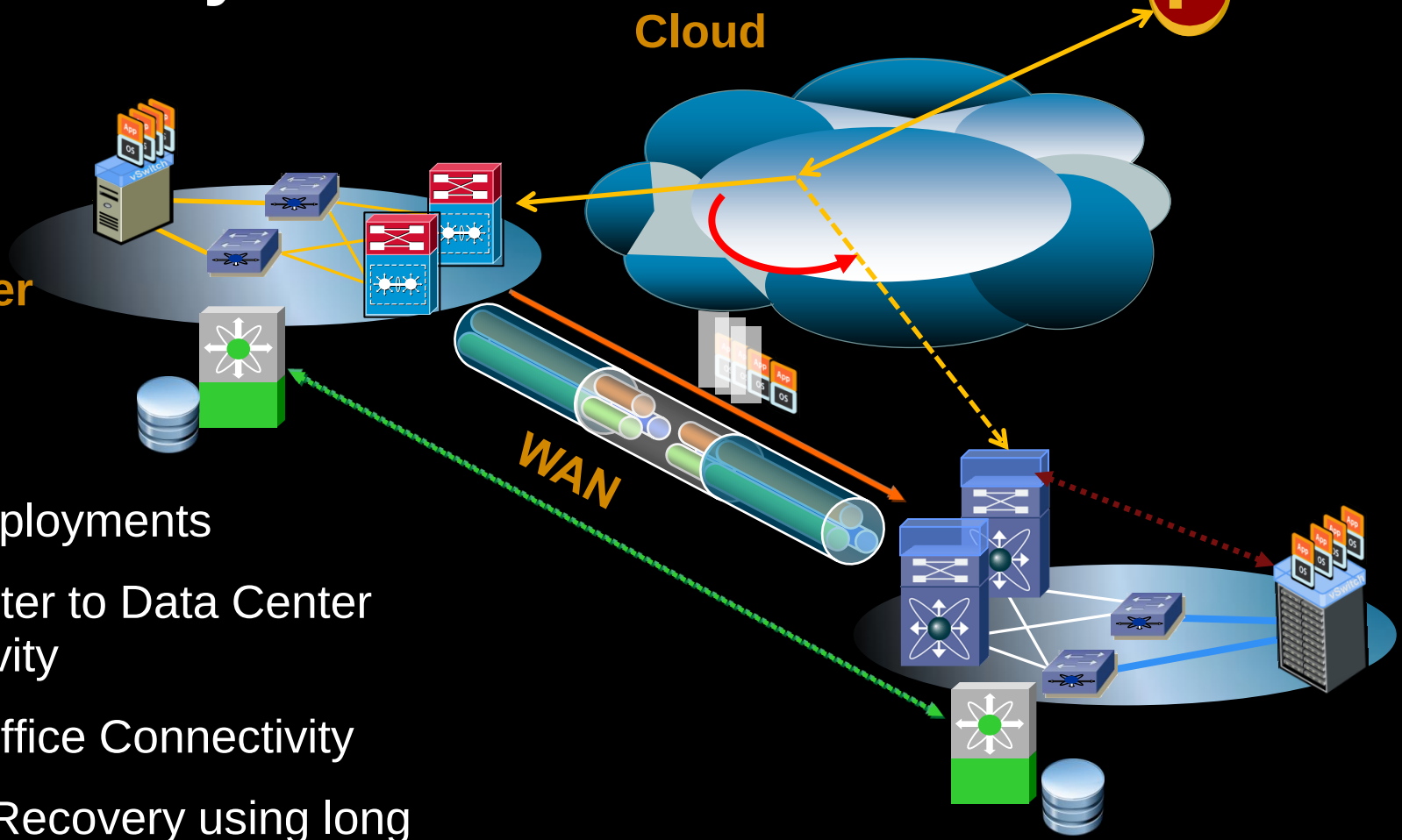
VMWARE DeDE
IBM DDE for Storage Tank
Solaris ZFS

Cisco MDS + WAAS

Data Domain
Netapp's A-SIS



Cisco's Focus Areas for Redundancy Elimination



- Cloud Deployments
- Data Center to Data Center Connectivity
- Branch Office Connectivity
- Disaster Recovery using long distance replication

Challenges in Deduplication

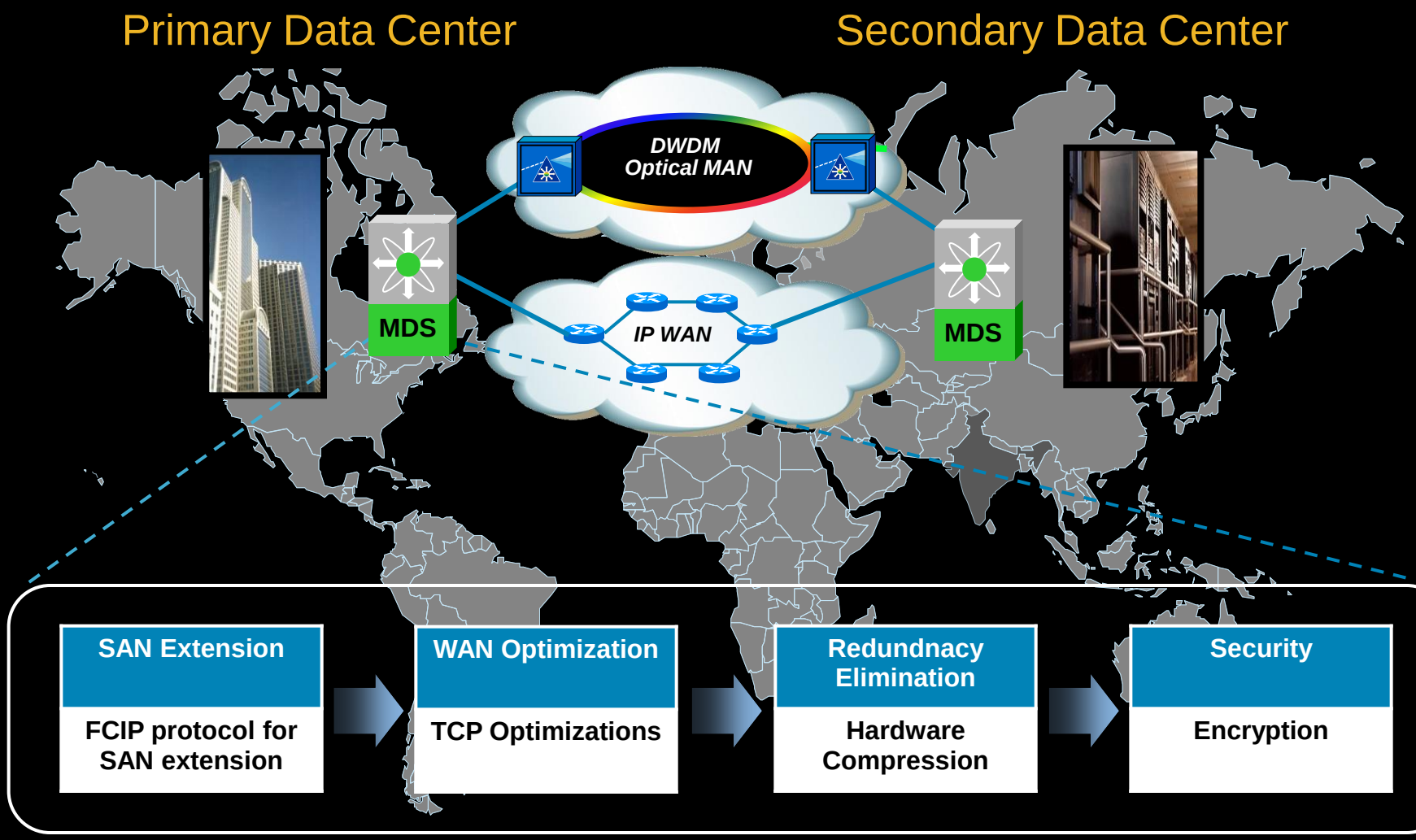
- CPU Intensive Chunking Operations
- CPU Intensive Fingerprint Calculation such as SHA-1
- Managing Large Indexes in Memory as well as in Persistent Storage
- Reducing Disk I/O for Index Lookup
- Load Balancing and Scalability in shared SAN environments
- Inter Host communication

  Host based deduplication for Shared File Systems

Cisco's Network Centric Redundancy Elimination

- Fully Integrated System for lowering WAN OPEX
  - Deduplication
  - Compression
  - TCP optimization
  - Encryption
- Hardware Assistance for Compression
  - Scope for Hardware Assists in Deduplication
- Coordination less approach for Shared SAN file systems
- Application/Host/Array Agnostic

Interconnecting Data Centers using MDS 9000

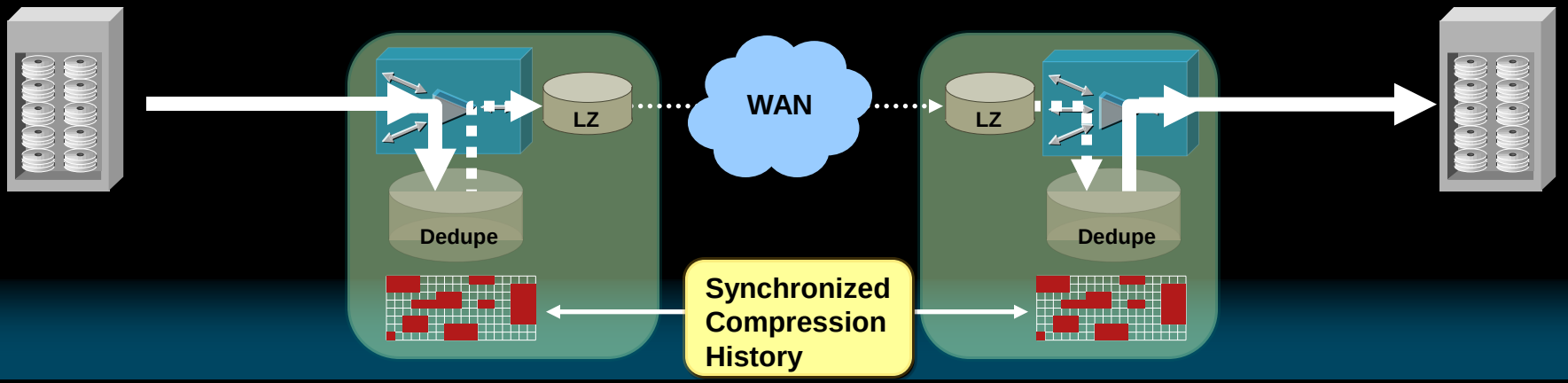


Cisco WAAS for Interconnecting Data Centers

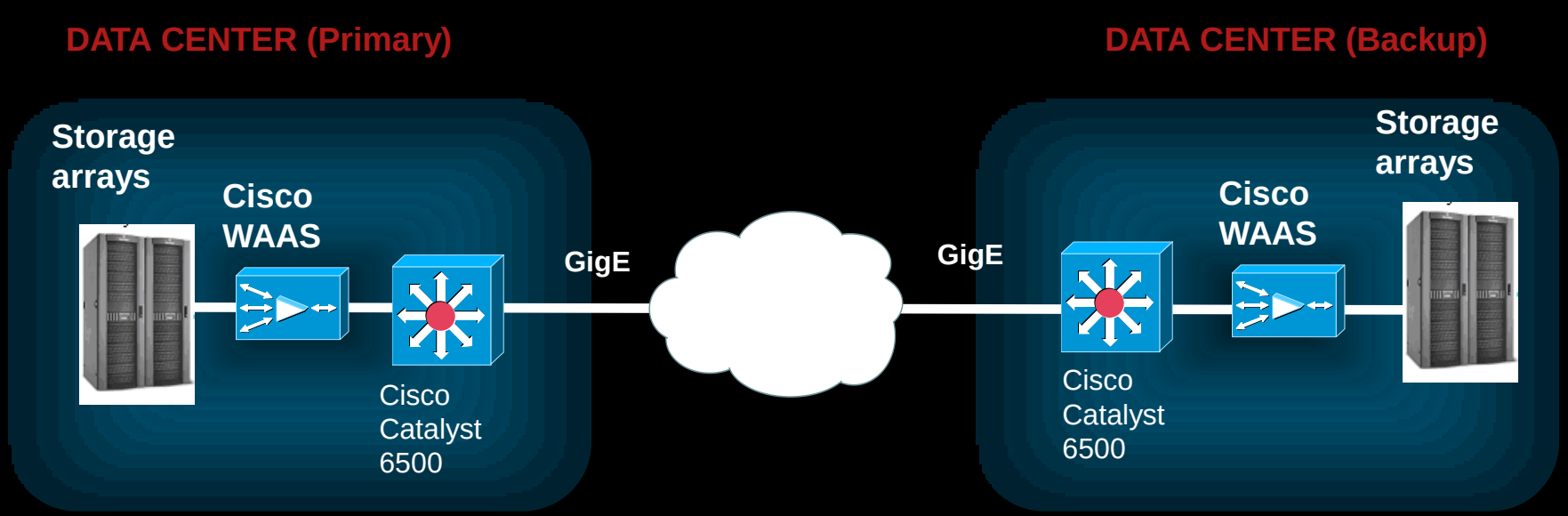
- Data Redundancy Elimination includes

  Data Deduplication: application-agnostic technique eliminates redundant data from TCP streams providing up to 100:1 reduction

  Persistent LZ Compression: session-based compression provides up to an additional 10:1 compression even for messages that have been optimized by deduplication

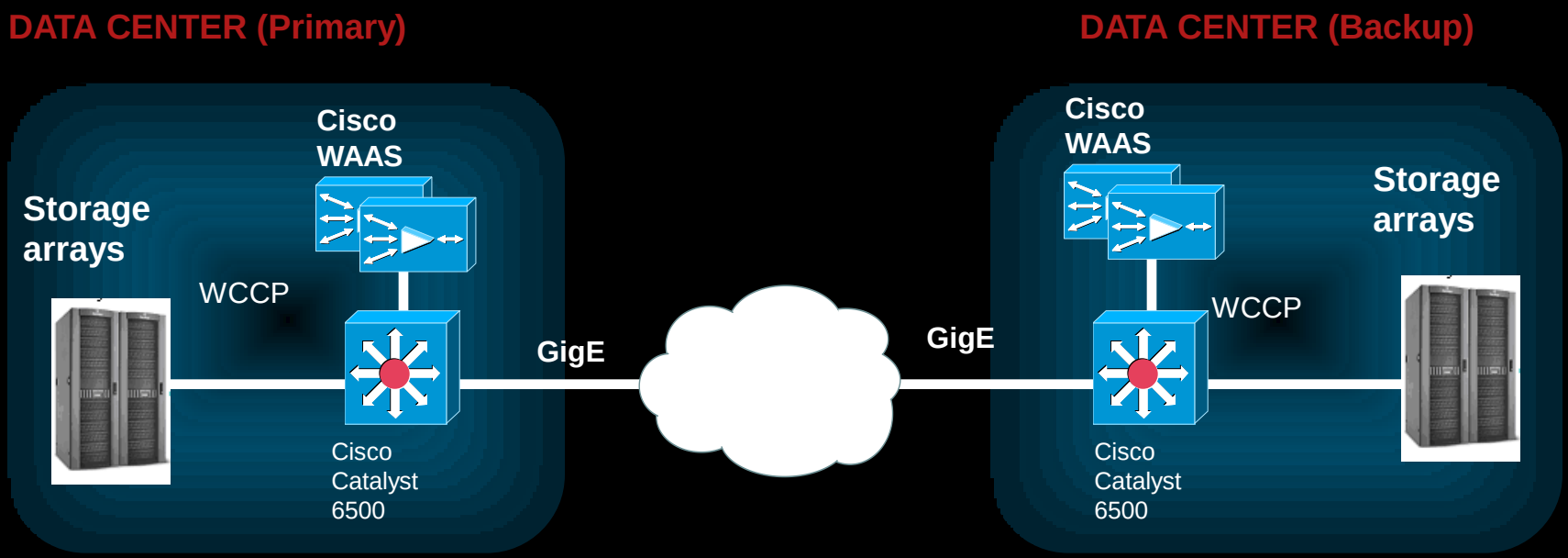


DC-DC Replication Using Cisco WAAS



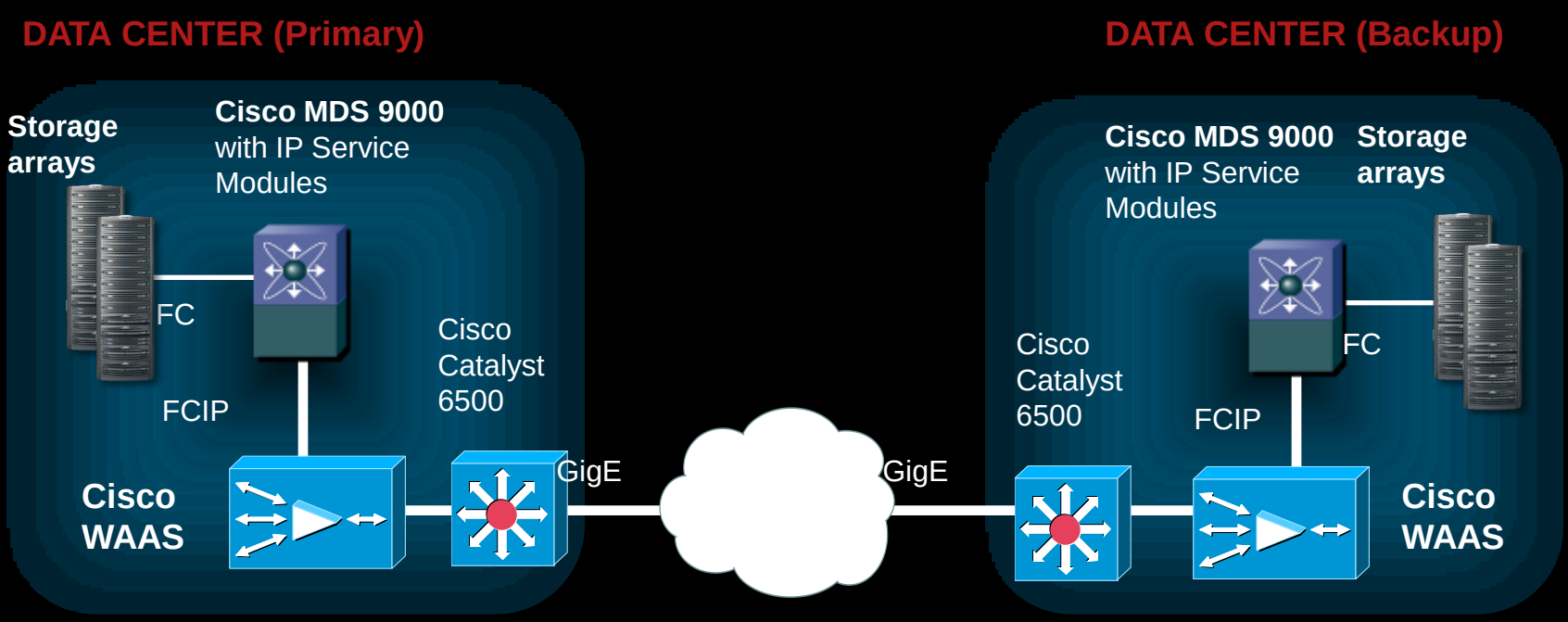
- **Ease of deployment**
- **Integration with existing network topologies**

Storage Replication in WCCP Mode



Higher availability and scalability through N+1 clustering

Storage Replication using Cisco MDS and WAAS



- **Improve end-to-end network resilience (against port failure)**
- **Optimize data replication for legacy FC storage**
- **Support multiple heterogeneous FC storage arrays**

Trends in Deduplication

- Use of Solid State Disks to overcome Index Lookup and retrieval latency
  - Price Performance Trade off
- Design Choices to overcome poor random write throughput of SSDs
  - Random write IOPS is still about 350
- Adaptive deduplication that combines Inline and Offline methods
- Data Agnostic Deduplication
  - Integrated approach for File and Block Deduplication