ViPR Distributed Storage System

Shashwat Srivastav, Sr Director Engg.

Kamal Srinivasan, Principal Prod Manager

Agenda

- ViPR Overview
- ViPR Architecture for scalability
- Geo distributed storage
- Demo
- Q&A

ViPR Overview

Storage Systems Today

Storage silos

- Impede development of applications
- Requires movement of data from one to another (e.g. File to HDFS)

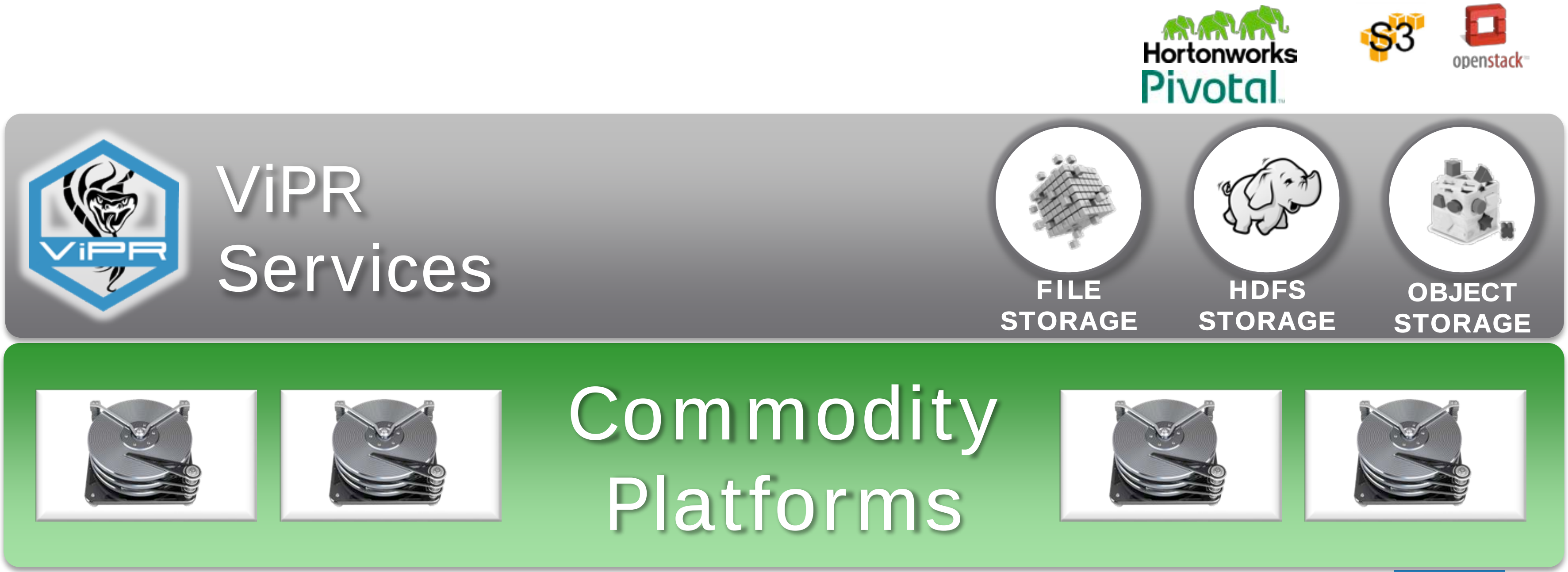
Enterprise scale

- Can’t economically scale for cloud
- Lack of elasticity

Not ready for modern apps

- Choice of API and HW
- Consistency semantics

Deliver Storage On Commodity Platforms with ViPR



ViPR Services

FILE STORAGE

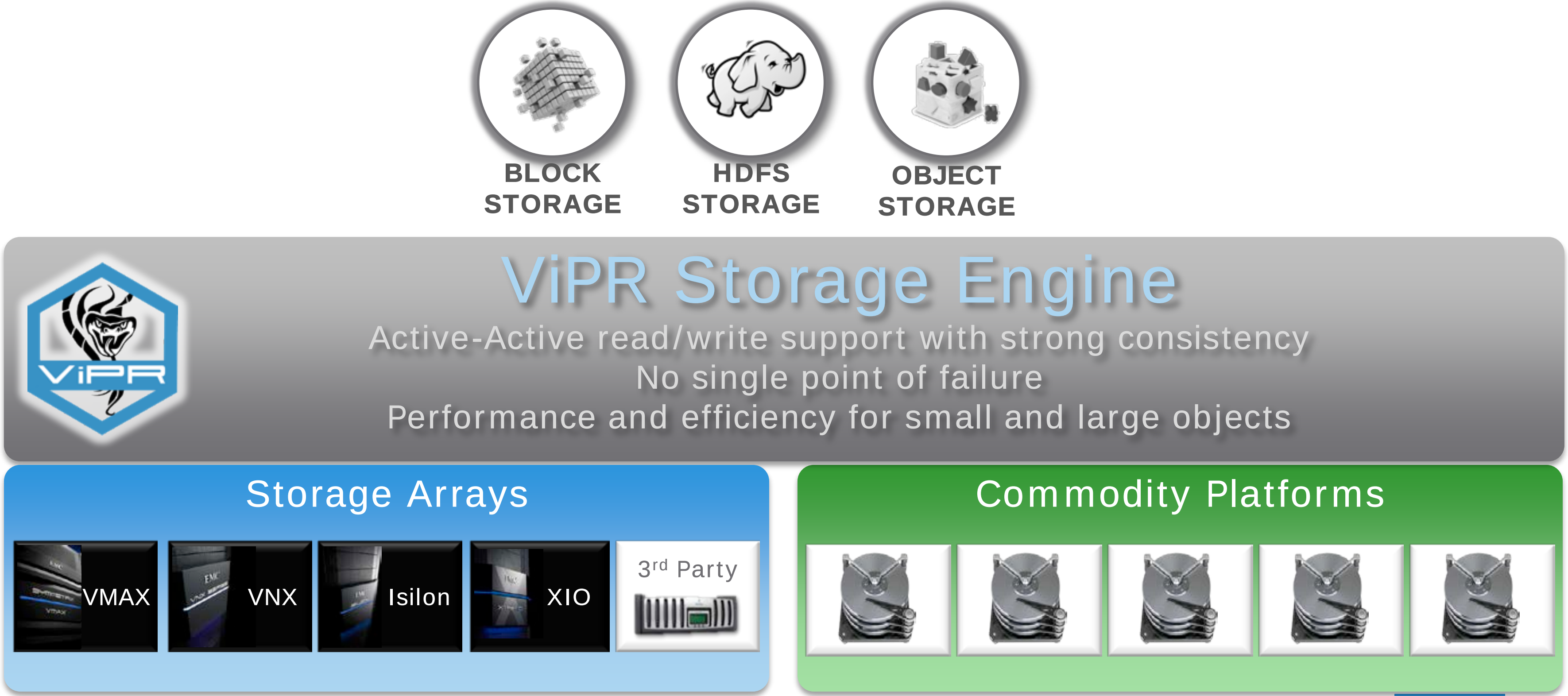
HDFS STORAGE

OBJECT STORAGE

Commodity Platforms

Scalable Architecture

ViPR Architecture



BLOCK STORAGE

HDFS STORAGE

OBJECT STORAGE

ViPR Storage Engine

Active-Active read/write support with strong consistency

No single point of failure

Performance and efficiency for small and large objects

Storage Arrays

VMAX

VNX

Isilon

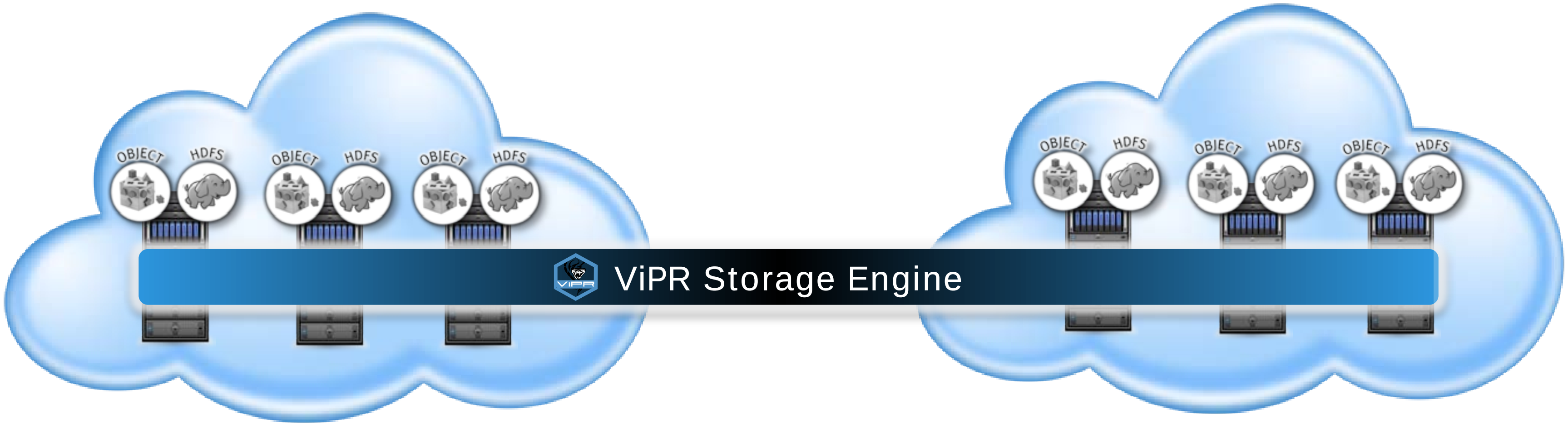
XIO

3rd Party

Commodity Platforms

Common Geo Functionality

GET https://account/bucket/object



ViPR Storage Engine

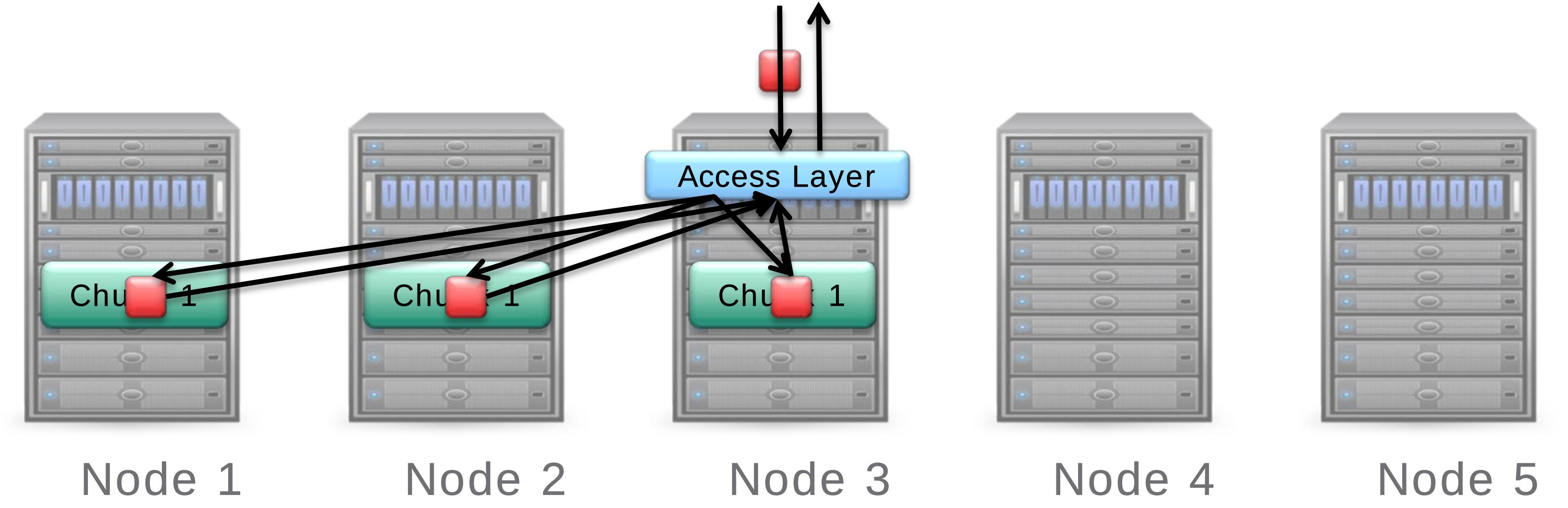
Chunks

- ViPR stores all types of data and index in “chunks”

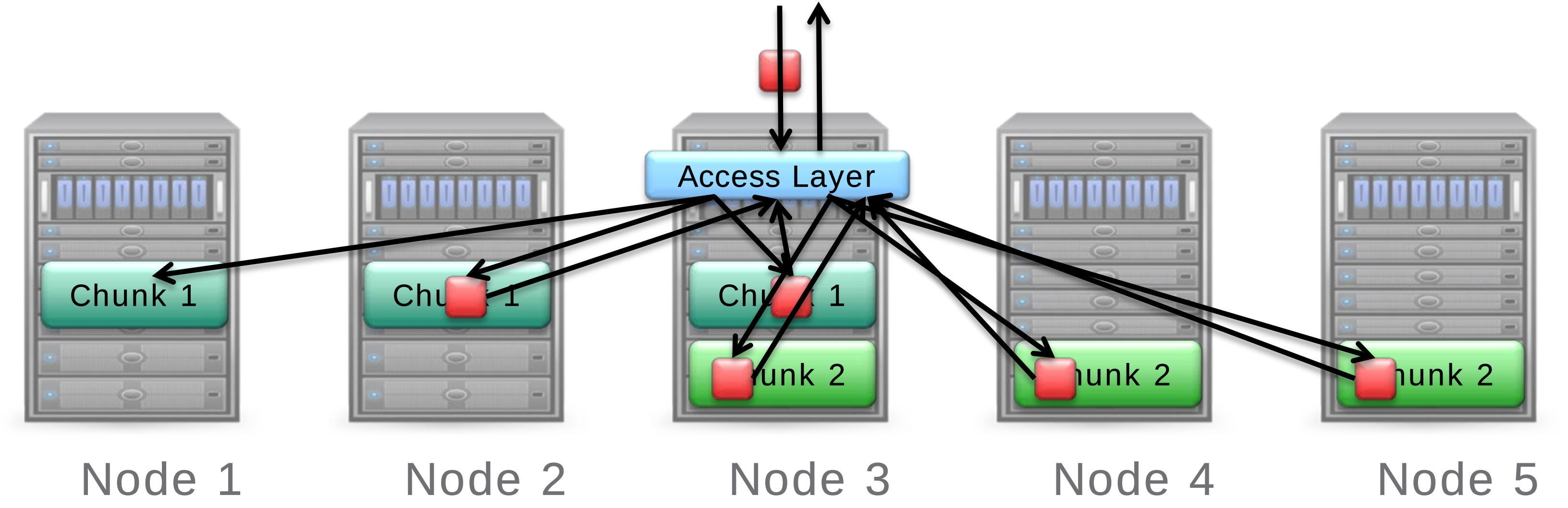
- Chunks are:
  - Logical containers of contiguous space (128MB)
  - Written in an append-only pattern

- All data protection operations are done on chunks

Chunk Write

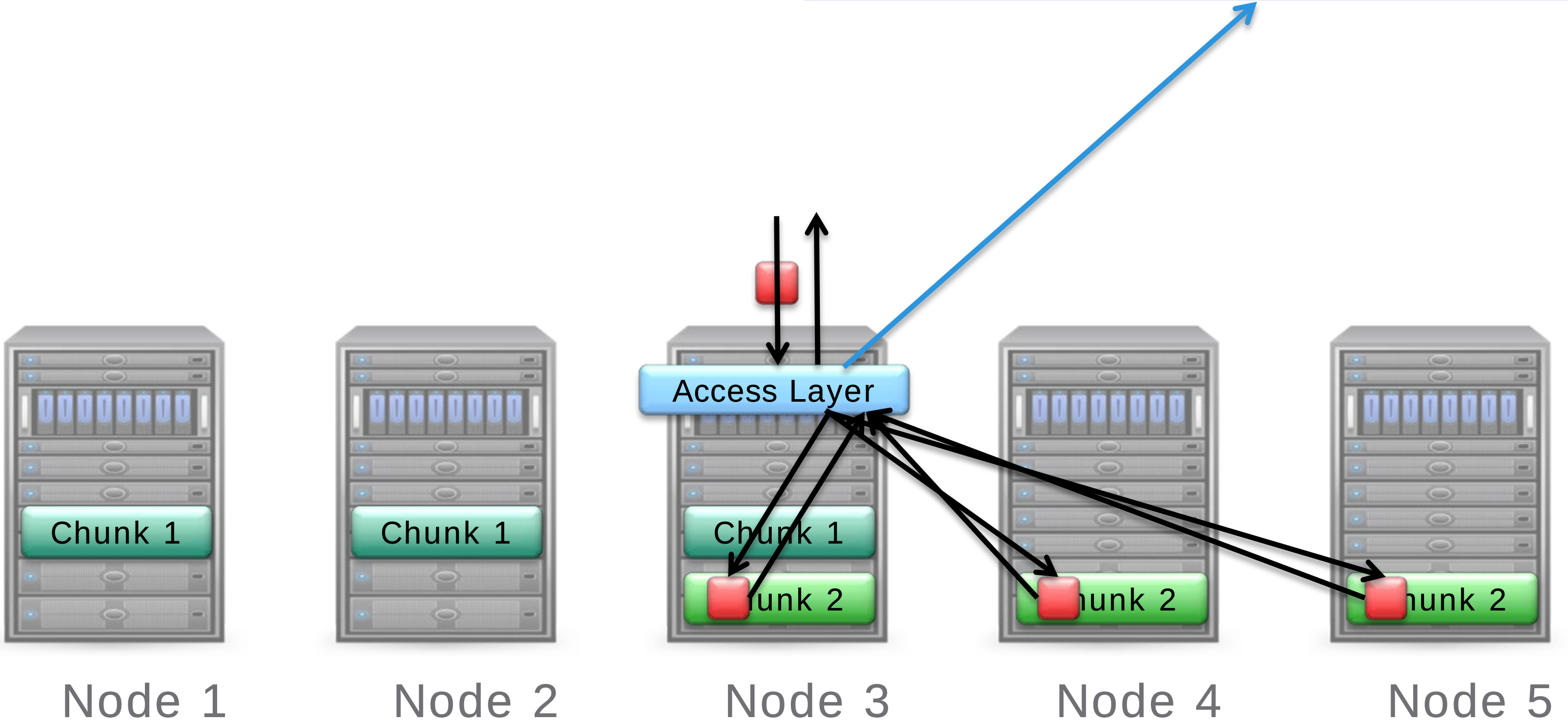


Chunk Write

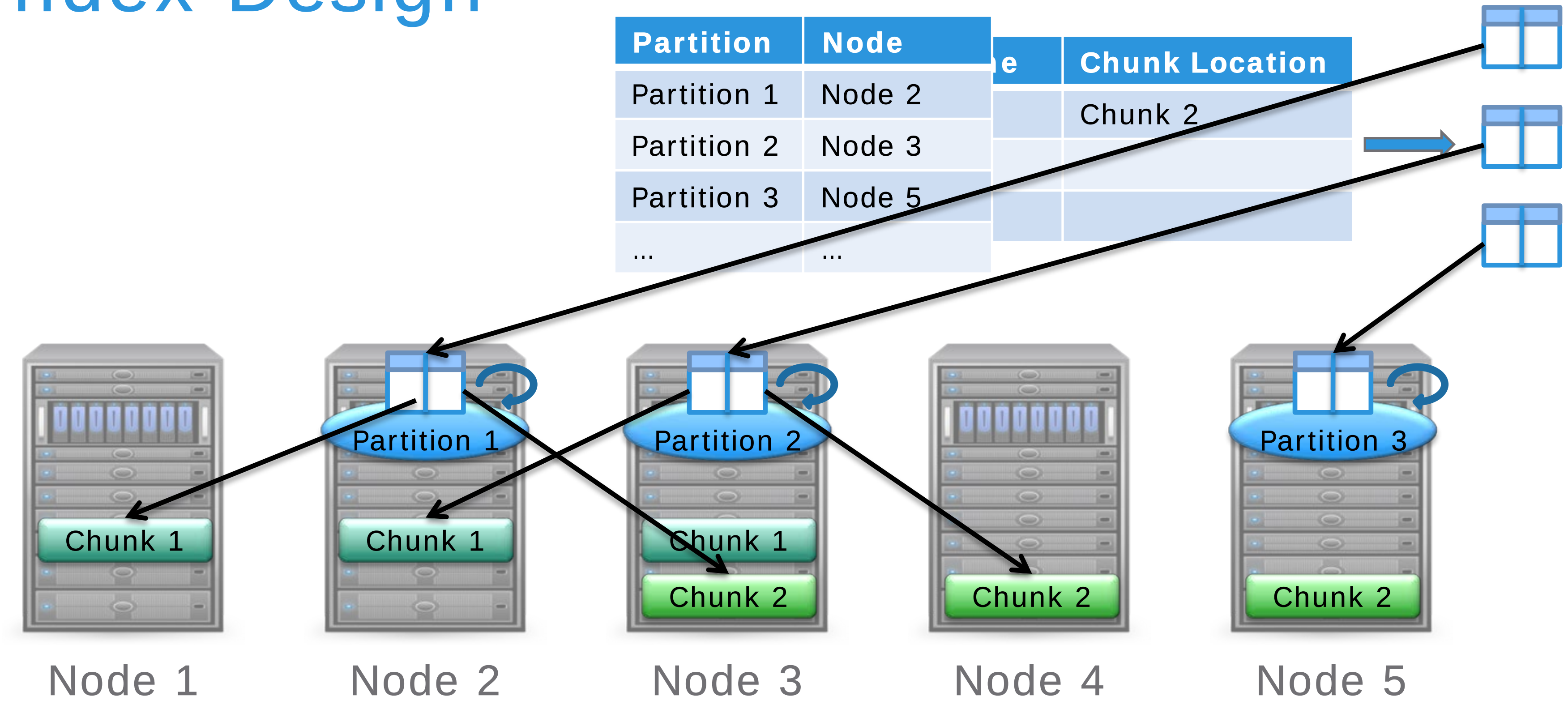

Access Layer
Chunk 1
Chu 1
Chu 1
unk 2
hunk 2
hunk 2
Node 1
Node 2
Node 3
Node 4
Node 5

Index

Content Name	Chunk Location
Image1.jpg	Chunk 2 (offset:100)



Index Design

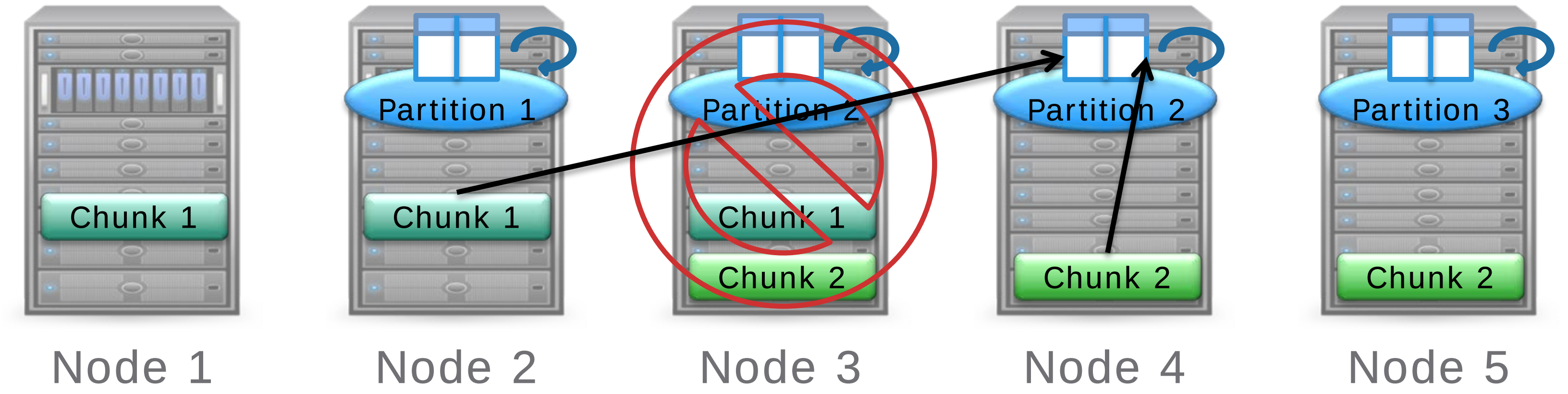


Partition	Node
Partition 1	Node 2
Partition 2	Node 3
Partition 3	Node 5
...	...

e	Chunk Location
	Chunk 2

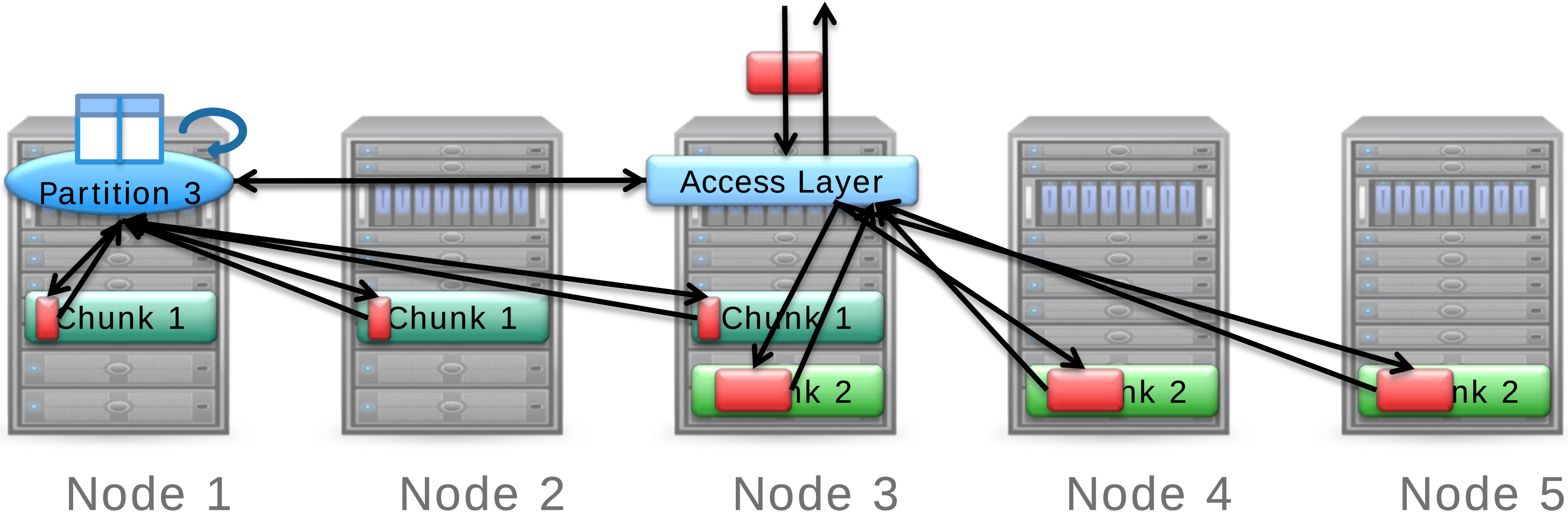
Index Design

Partition	Node
Partition 1	Node 2
Partition 2	Node 4
Partition 3	Node 5
…	…

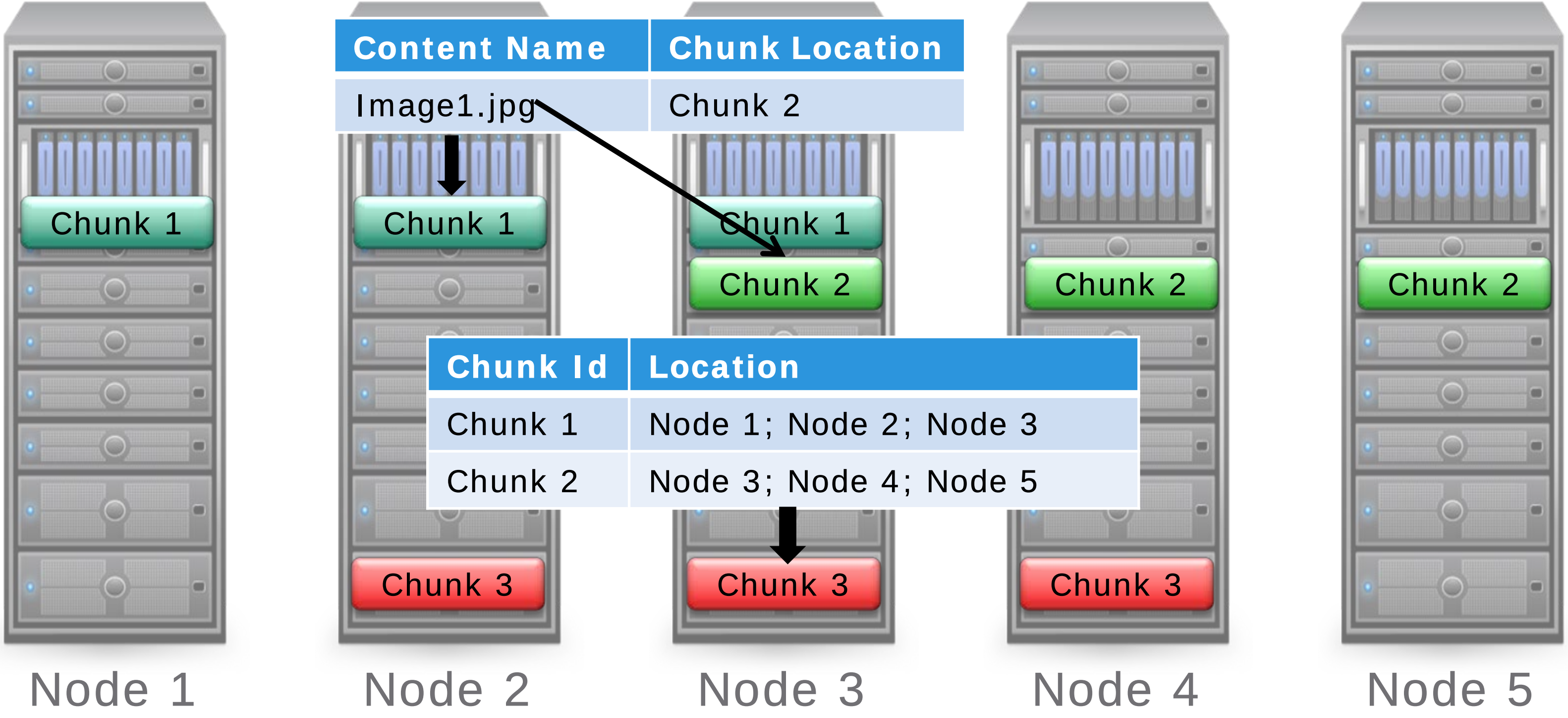


Transaction

Content Name	Chunk Location
Image1.jpg	Chunk 2 (offset:100)



Chunk Info

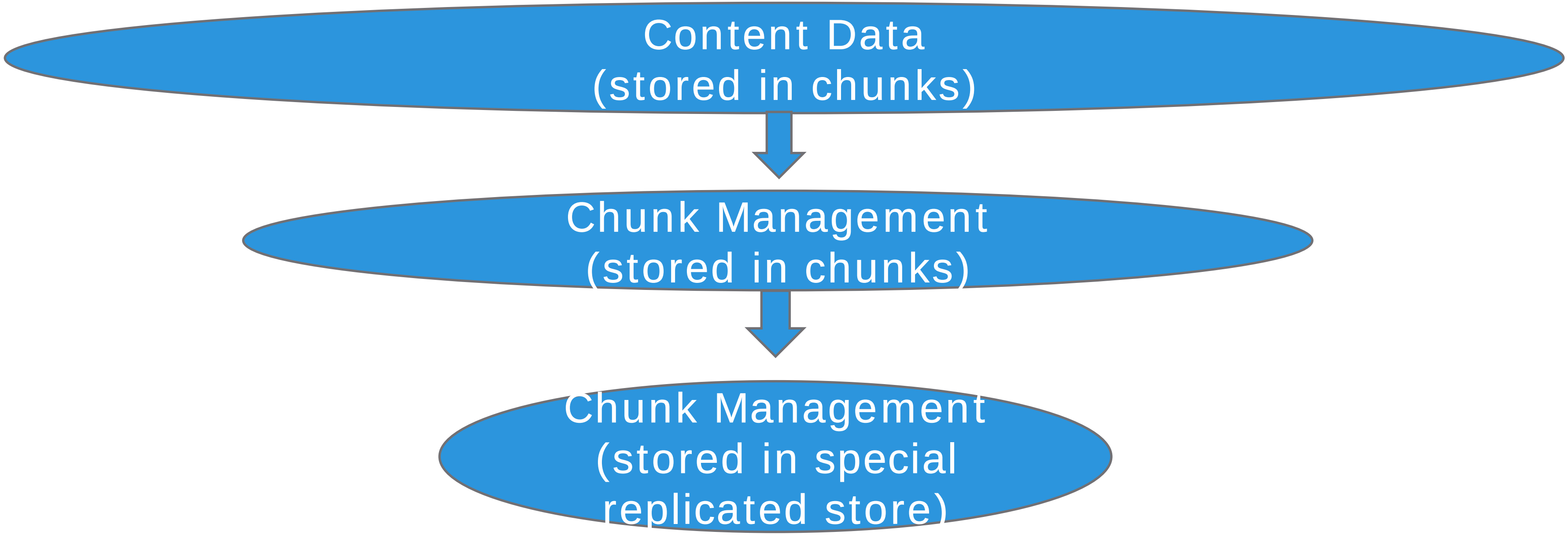


Content Name	Chunk Location
Image1.jpg	Chunk 2

Chunk Id	Location
Chunk 1	Node 1; Node 2; Node 3
Chunk 2	Node 3; Node 4; Node 5

Chunk Id	Location
Chunk 3	Node 2; Node 3; Node 4

High Scalability Technique



Content Data
(stored in chunks)

Chunk Management
(stored in chunks)

Chunk Management
(stored in special
replicated store)

Optimized Geo Protection

Key Points

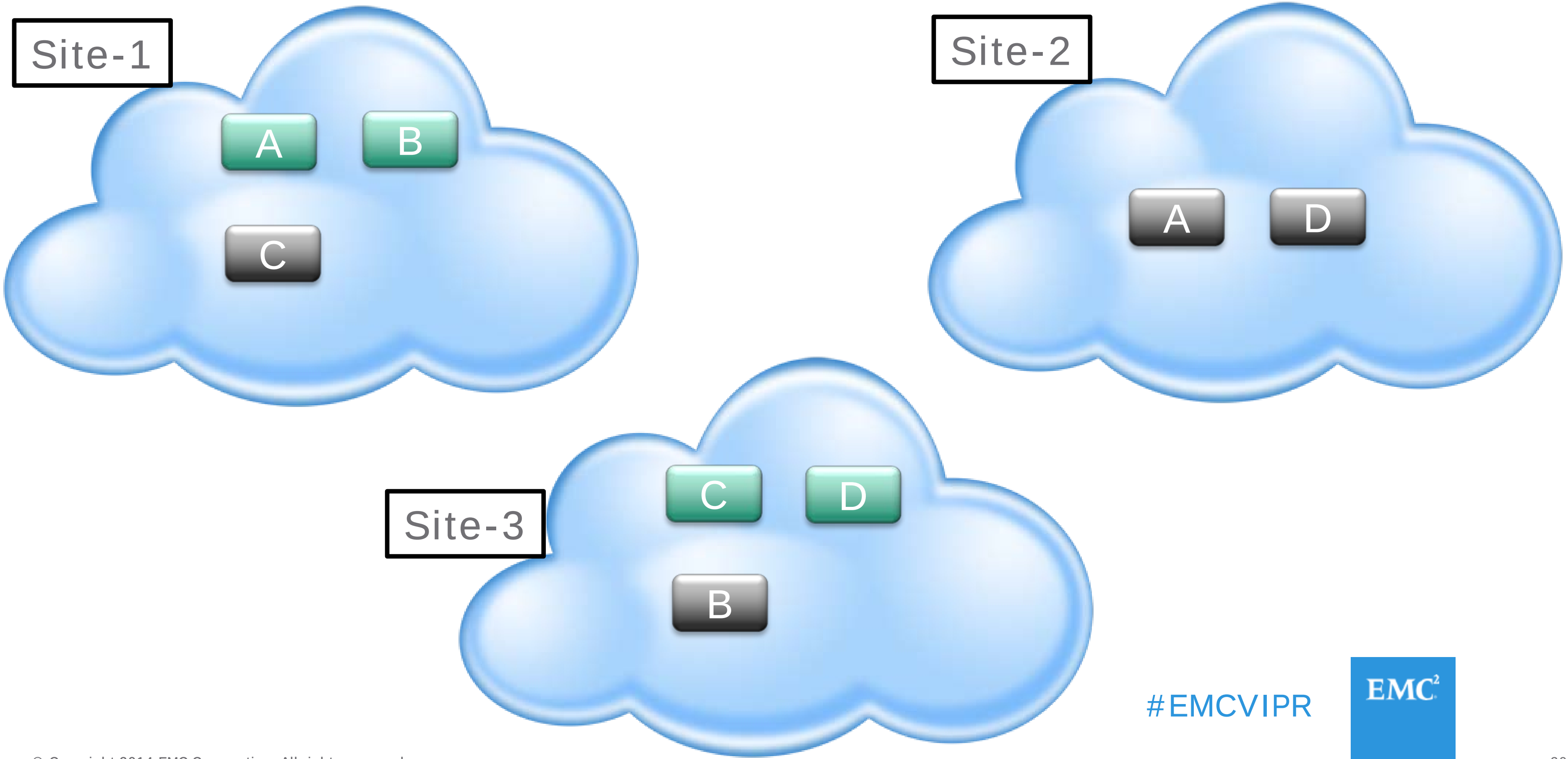
- The scheme can tolerate **one site disaster** along with up to **2 node failures** in all the rest of the sites.

- The node failures are repaired using fragments from **local site** without WAN traffic.

- Achieves **~1.8 copies across 4 sites** without having to reconstruct data across the WAN

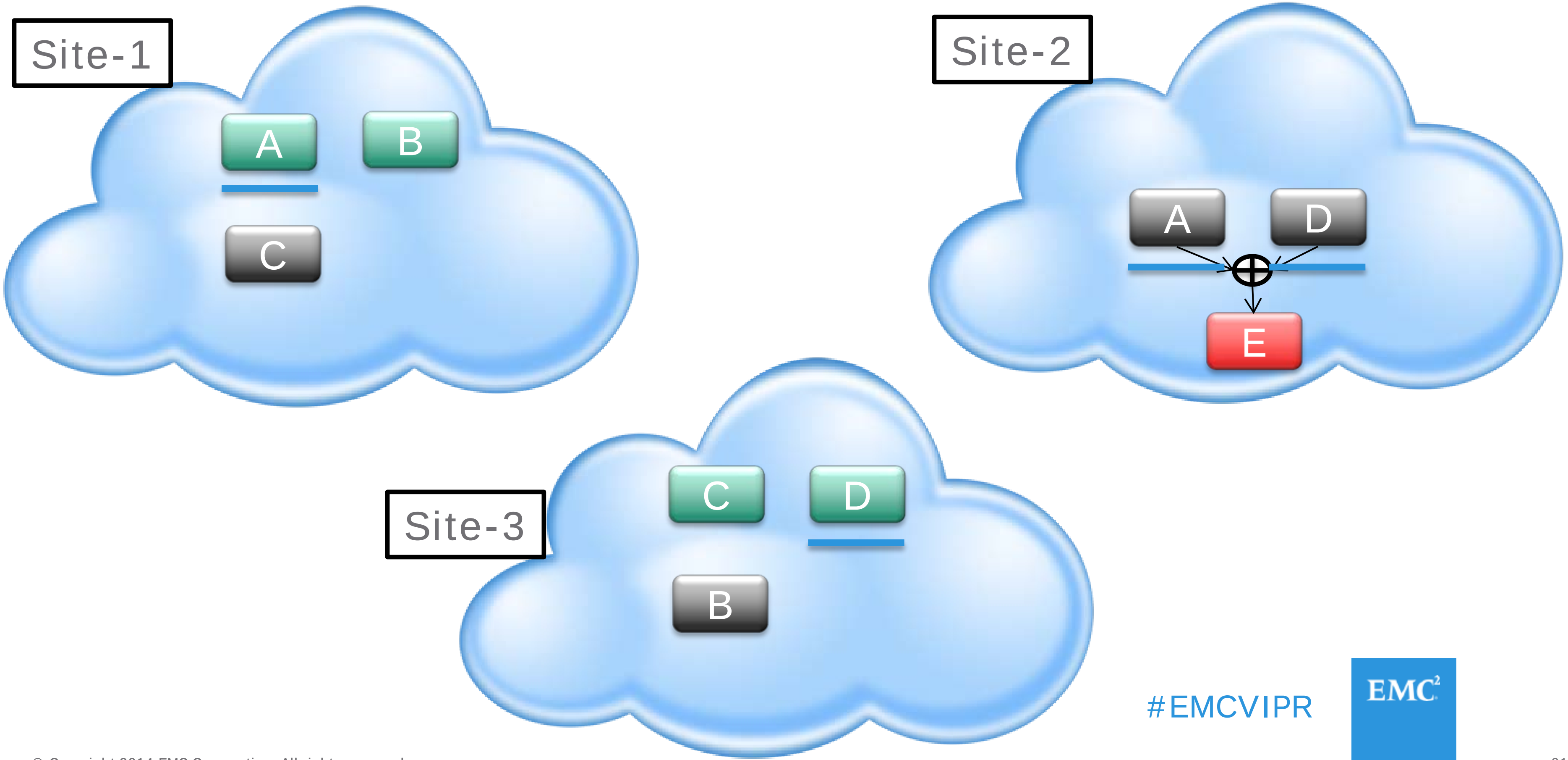
#EMCVIPR

Chunk Backup



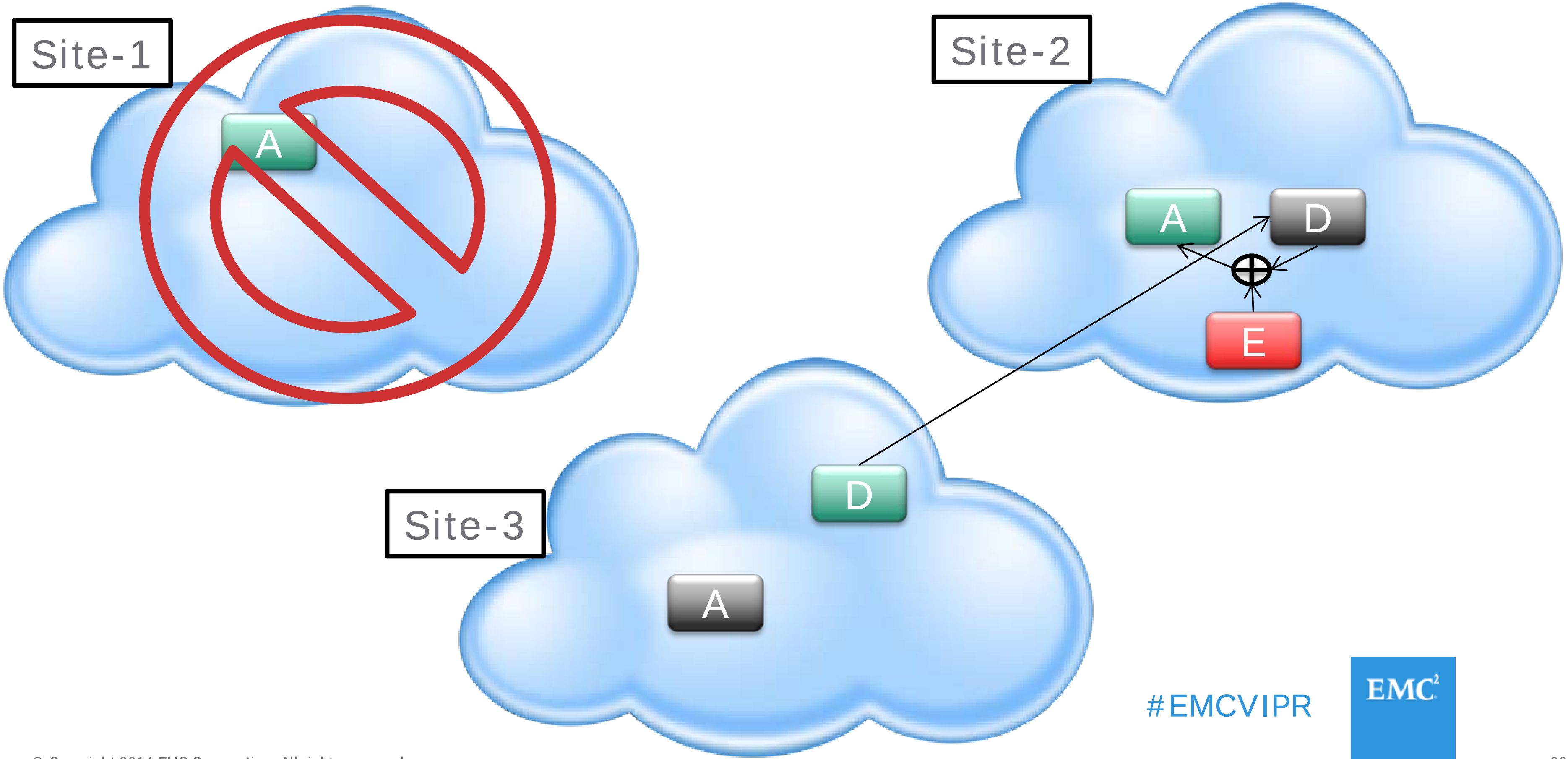
#EMCVIPR

Backup Compaction



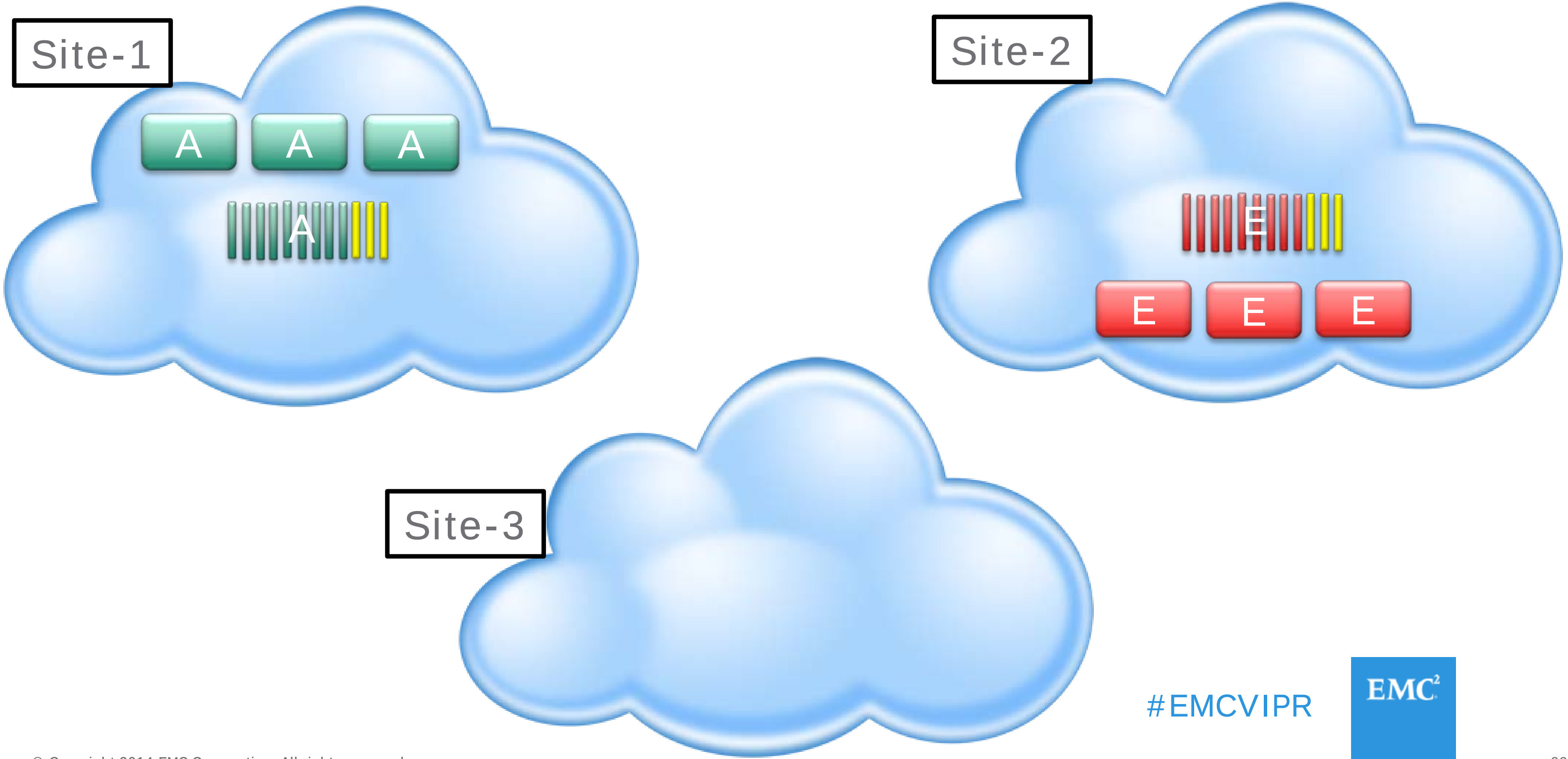
#EMCVIPR

Chunk Recovery



#EMCVIPR

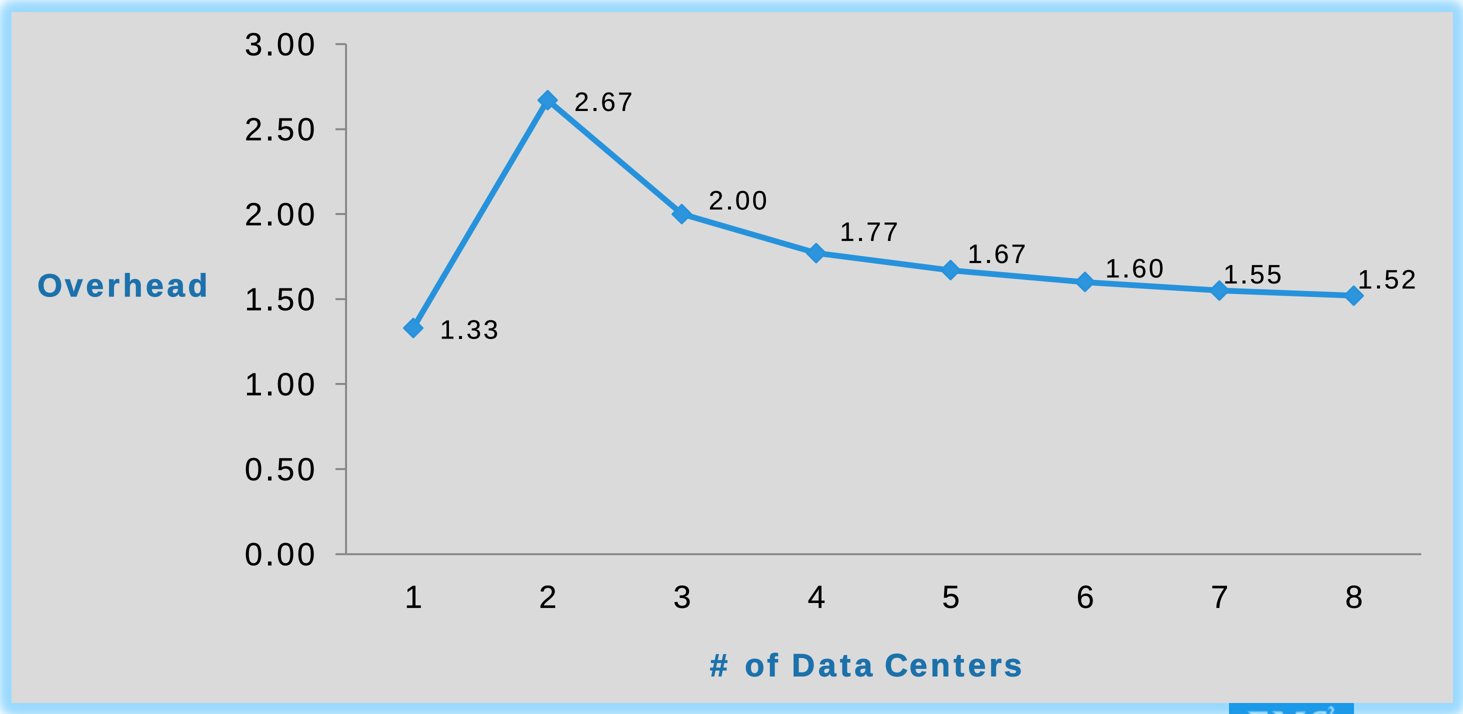
Local Protection



#EMCVIPR

Storage Overhead

Optimized data access, protection and efficiency



Location Agnostic Access With Strong Consistency

Industry Solutions

- Geo location partitioned namespace
- Eventual consistency across geo locations
- Sync write all transaction across geo locations

ViPR Solution

Scalable Geo protection

- Each bucket, object, directory, and file is represented as an entity in the index

Traffic heuristics

- Sense traffic pattern individually for each entity.

Strong consistency

- Different techniques to avoid WAN round trip

EMC^2®