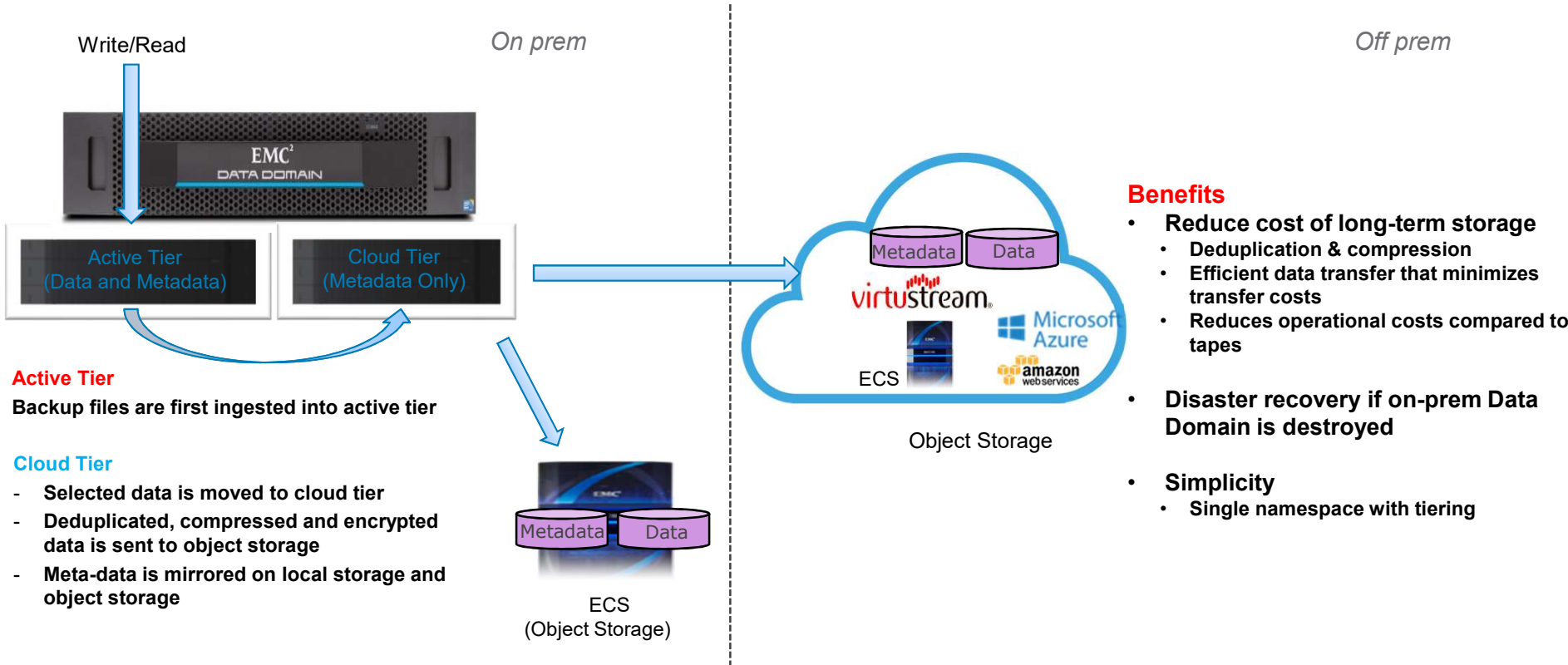


Data Domain Cloud Tier: Backup here, backup there, deduplication everywhere!

**Abhinav Duggal, Fani Jenkins, Philip
Shilane, Ramprasad Chinthekindi, Ritesh
Shah, Mahesh Kamat**

Dell EMC

Data Domain Cloud Tier



Benefits

- **Reduce cost of long-term storage**
 - Deduplication & compression
 - Efficient data transfer that minimizes transfer costs
 - Reduces operational costs compared to tapes
- **Disaster recovery if on-prem Data Domain is destroyed**
- **Simplicity**
 - Single namespace with tiering

Key highlights of our architecture

- Meta-data cached locally to assist deduplication and garbage collection
- Move incremental data with deduplication to the object storage efficiently
- Uses perfect hashing and breadth-first Merkle tree traversal
 - Find out how much space will be freed by moving data to cloud tier
 - Seed large amount of data efficiently to the object storage
 - Clean cloud tier (garbage collection)
- Provides efficient cleaning for private and public cloud providers using API integration and microservices

Summary of field evaluation results

- Analyzed field results from hundreds of sampled customers running cloud tier (first released in April 2016)
- Some of our customers have moved 20PB of data to the cloud tier saving millions of dollars due to benefits of deduplication and compression
- Total compression (deduplication * local compression) for our sampled cloud tier customers ranges from 4X to 100X

USENIX ATC '19, 04:10 PM, Track II on July 11th