

# 微软技术大会

## Microsoft Ignite China

# 使用虚拟磁盘和共享文件存储 来优化虚拟机

唐君

PM @ Azure Storage Team

# 课程内容

Azure 虚拟磁盘 + 深度了解

Azure 共享文件系统 + 深度了解

# Azure 全球架构



30 GA, 4 coming soon – Storage is available in every region

# 什么是 Azure 云存储

## Azure 基础架构

Azure Services: SQL Data Warehouse, HDInsight, Data Lake Store, Event Hubs, IoT Hubs...

Microsoft Services: Office 365, OneDrive, Xbox, Skype...

### 超大规模

> 120 Trillion Objects, > 19 million transactions per second

### 耐久性

Never lose your data. Multiple redundancy options. Automatic data checks

### 安全性

Encryption at Rest. Client side Encryption. Integration with KeyVault

### 高可用性

Fault tolerance to hardware/software issues. Automatic load balancing

### 开放性

REST API, Open sourced Client Libraries – .NET, Java, C++, Python, Node.js, iOS, Android, Xamarin...

### 混合云

Extensive partner ecosystem. Azure Stack for private/hosted clouds.

# Azure 云存储一览

## IaaS – 架构即服务



Storage



Virtual machines



Networking

## PaaS 平台即服务



Existing frameworks



Web and mobile



Microservices



Serverless Compute

### 虚拟磁盘 Disks

Persistent disks for Azure IaaS VMs

Premium Storage Disks option: SSD based, high IOPS, low latency

### 共享文件 Files

Fully Managed File Shares in the Cloud

SMB and REST access  
"Lift and shift" legacy apps

### 对象存储 Objects

Highly scalable, REST based cloud object store

Block Blobs: Sequential file I/O  
Cool Tier Available  
Page Blobs: Random-write pattern data  
Append Blobs

### NoSQL Tables

Massive auto-scaling NoSQL store

Dynamic scaling based on load

Scale to PBs of table data

Fast key/value lookups

### 消息队列 Queues

Reliable queues at scale for cloud services

Decouple and scale components

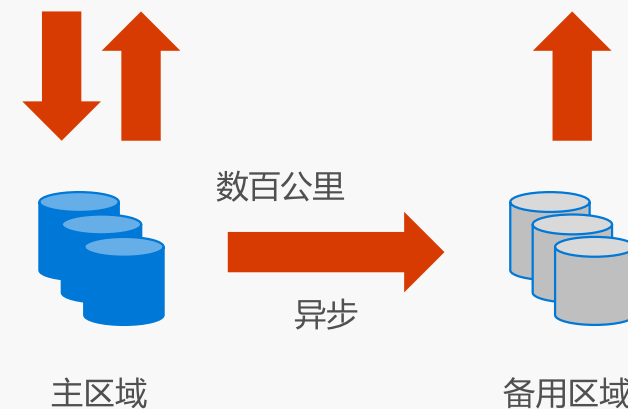
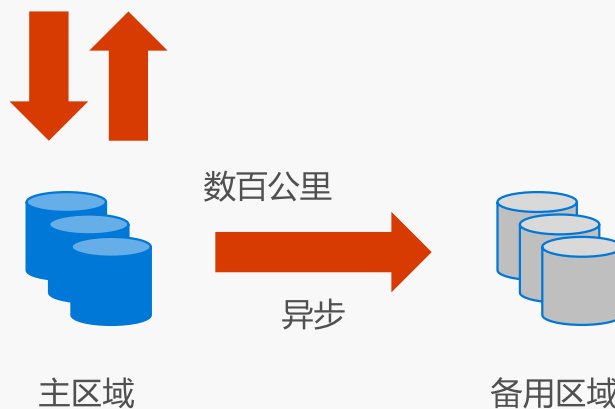
Message visibility timeout and update message to protect against unreliable

dequeueers

## Built on a unified Distributed Storage System

Durability, Encryption at Rest, Strongly Consistent Replication, Fault Tolerance, Auto Load-Balancing

# Azure 云存储 - 耐久性



## 本地冗余备份

### LRS

- 3 replicas, 1 region
- Protect against disk, node, rack failures
- Write is ack'd when all replicas are committed
- Superior to dual-parity RAID

## 地域冗余备份

### GRS

- 6 replicas, 2 regions (3/region)
- Protects against major regional disasters
- Asynchronous to secondary

## 地域冗余备份可读访问 - 推荐

### RA-GRS

- GRS + Read access to secondary
- Separate secondary endpoint
- RPO delay to secondary can be queried

Strong consistency, checksums/CRC and background scrubbing to preserve data integrity

# 内容一览

## Disks Overview

Features

Benefits

Roadmap

## Disks Deep Dive

Migration

Backup

Best Practices

## Files Overview

Scenarios

Benefits

Application  
Architecture

## Files Deep Dive

Snapshot

Best Practices



# Azure Disk - 虚拟磁盘

# 简介

## 什么是虚拟磁盘

Azure虚拟机作为系统盘和数据盘  
使用 Page Blob  
高级和标准: SSD & HDD

## 特性

高磁盘吞吐和低延迟  
支持磁盘加密(BitLocker)  
读写缓存  
可通过REST协议访问  
企业级耐久性服务  
快照及备份  
即时扩展

# 虚拟磁盘: 高级 vs. 标准

	Premium Storage Disks	Standard Storage Disks
Disks Types	SSD	HDD
Recommended Workloads	Mission critical applications which require high-performance, high-availability, and low latency for IO-intensive Enterprise workloads	Infrequently accessed data and Dev/Test workloads
Performance Expectations	5,000 Provisioned IOPS and 200 MB/sec throughput for 1 TB disk  80,000 Provisioned IOPS and 2,000 MB/sec throughout with GS5 (largest VM)  160,000 Provisioned IOPS and 1,600 MB/sec throughout with GS5 (largest VM)  Single digit millisecond latency  Consistency of provisioned performance	Up to 500 IOPS and 60 MB/sec for 1 TB disk  32,000 IOPS with G5 (largest VM)    Performance is not provisioned

# Azure 虚拟磁盘路线图

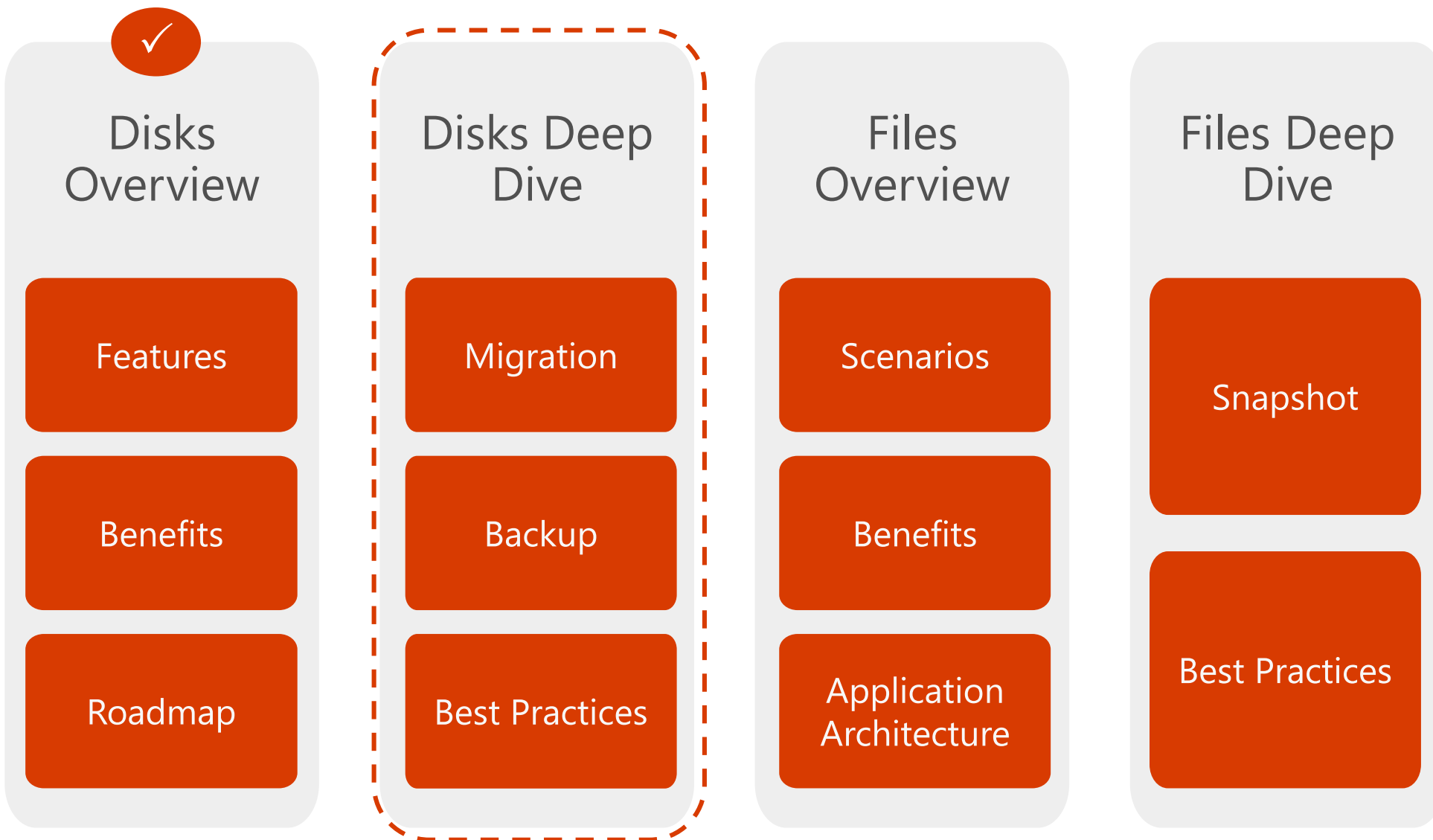
## Recently Released

- Azure Backup Support
- Encryption at Rest
- Azure Site Recovery Preview
- Incremental Snapshot Copy

## Future

- More disk sizes
- Disk analytics enhancement

# 内容一览



# 从标准磁盘到高级磁盘的迁移

## 两种方法

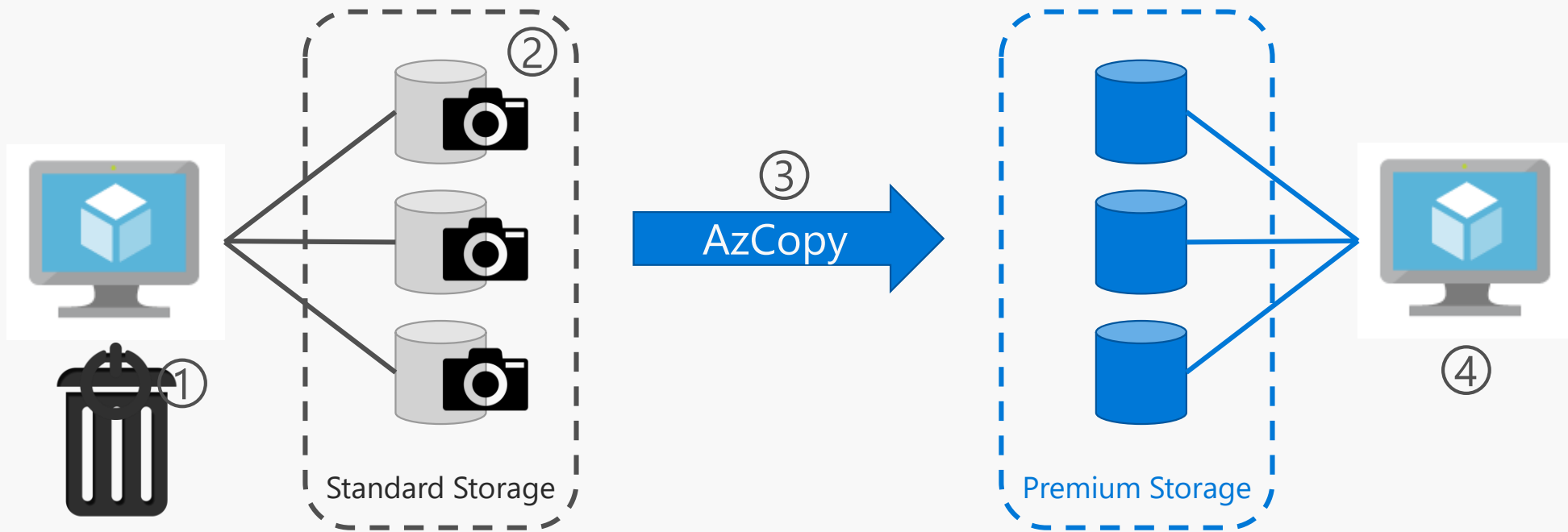
### 1. 离线迁移 - Offline Migration

Offline migration of the disks to Premium Storage using AzCopy  
Longer Downtime

### 2. 即时迁移 - Live Migration

Live migration of the VM/Disks using Azure Site Recovery or 3<sup>rd</sup> party solutions like Commvault  
Shorter Downtime

# 选项 1：使用AzCopy进行离线迁移



1. Create new VM with all the VM settings and the Standard Storage account
2. Copy data from the Standard Storage account to Premium Storage account  
Specify Availability Set, VNET and other VM settings during VM creation

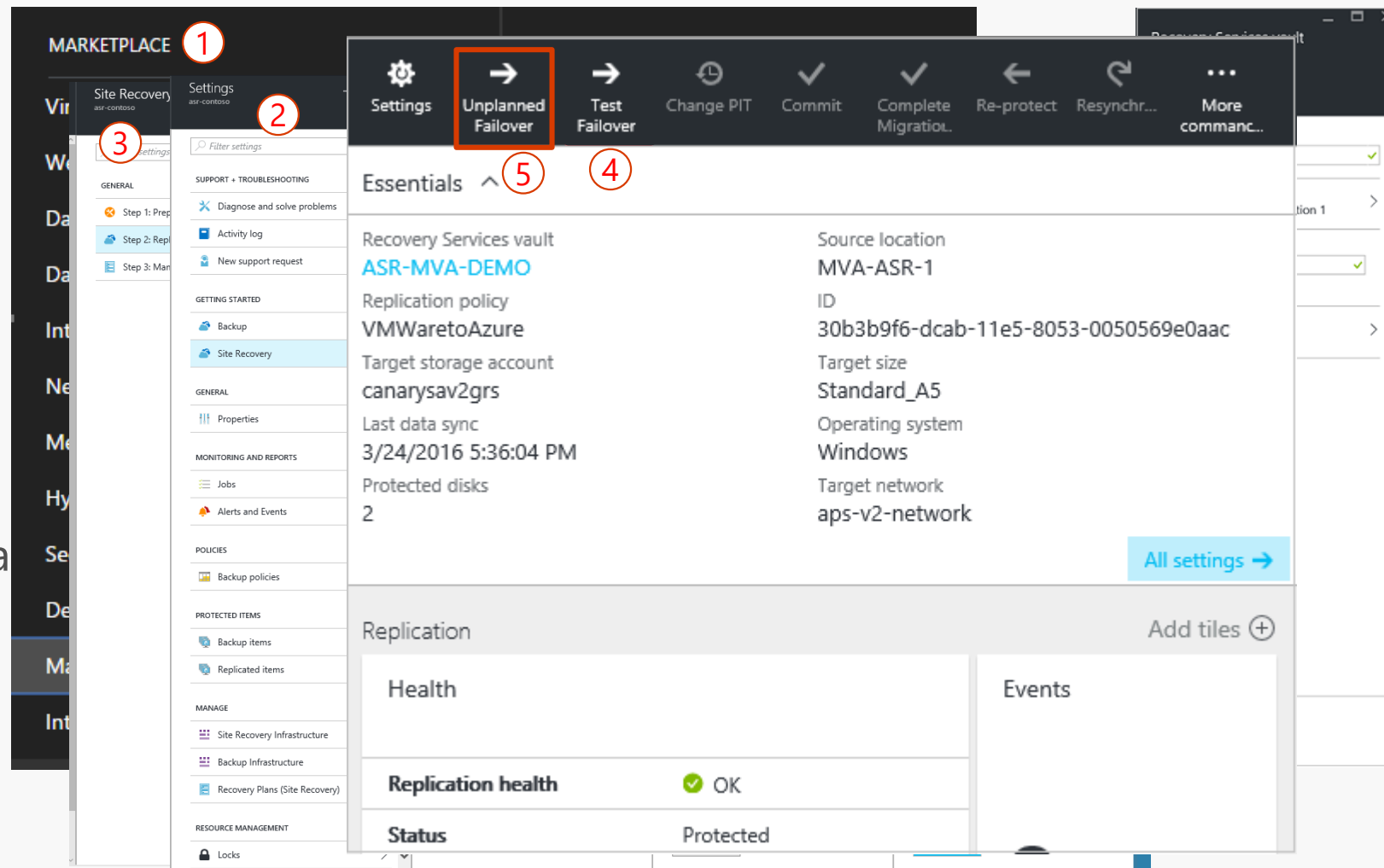
# 选项 2: 使用 ASR 进行即时迁移

## Benefits

Migrate using ASR service  
Shorter downtime

## Migration Steps

- ① Create Recovery Services Vault
- ② Setup Infrastructure for Migration
- ③ Replicate the Application
- ④ Test Failover (no downtime) and validate
- ⑤ Failover (Migrate)





# 磁盘备份策略

## 企业容灾需要可靠的备份来恢复应用及数据

Backup must use "consistent" snapshots

Creating "consistent" snapshots of running VMs require additional steps

---

### 1. 使用 Azure 备份服务 (推荐)

Azure Backup or a 3<sup>rd</sup> Party Backup solution, like Commvault

### 2. 创建新的备份方案

Build a custom solution for your needs

### 3. 停机备份

Alternative to create consistent snapshots with a short downtime

# 选项 1:使用 Azure 备份服务

## 备份虚拟机

## 推荐 适用于高级磁盘和标准磁盘的灾备方案

Creates consistent backups for running VMs

Supports Windows and Linux VMs

## 特性

Application consistent backups for Windows VMs

File System consistent backups for Linux VMs

Local and Geo-replicated Backup options

Flexible backup frequencies and retention periods

Encrypted backup storage

# 选项 2: 创建新的备份方案

## 创建一致性的备份

Freeze the disks, flush pending IOs and then create blob snapshots

For Windows, use WSB/VolumeShadow (VSS) to create Application consistent snapshots

For Linux, use similar tools to create File System consistent snapshots

## 将备份复制到另一个存储账户

Target a Standard RA-GRS account for geo redundant backup

Implement incremental copying for efficiency

Use the new Differential Page Ranges API for incremental copying

# Differential API for Copying Snapshots

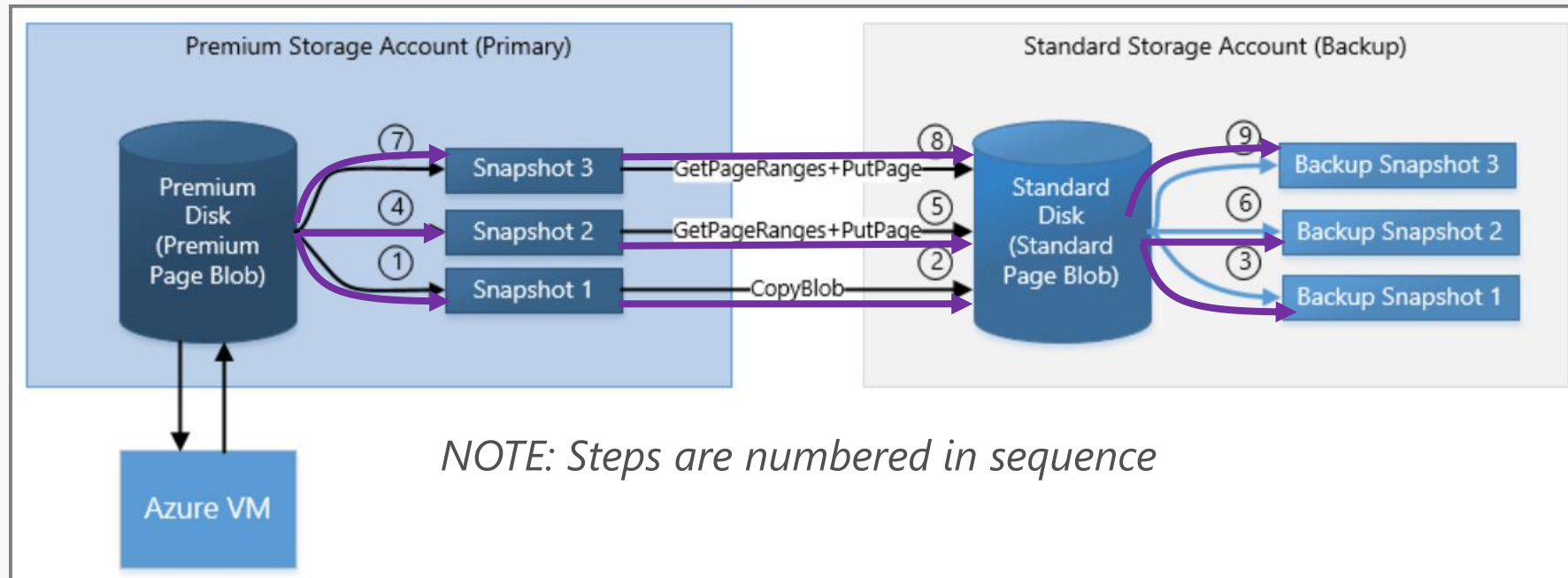
## Enhanced Get Page Ranges API

Use Storage REST version 2015-07-08 or newer

Support for snapshots or blobs created on or after January 1, 2016

Specify "snapshot" and "prevsnapshot" to get the list of changes between snapshots

<https://myaccount.blob.core.windows.net/mycontainer/myblob?comp=pagelist&snapshot=<DateTime>&prevsnapshot=<DateTime>>



# 选项 3: 停机备份

## 创建快照



## 将备份复制到另一个存储账户

Target a Standard RA-GRS account for geo redundant backup

# 从备份中恢复

## 1. 使用已有的备份方案

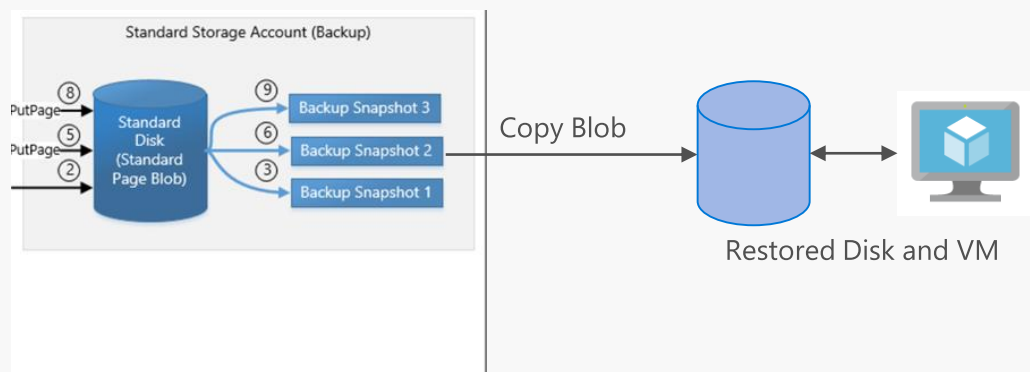
Restore mechanisms as provided by the backup solution

## 2. 使用新建的备份方案

Copy Blob the snapshot backups of all the disks

Create new VM using the copied disks

Validate the new VM, and then delete the prior VM / disks



Alternative is to promote the snapshot. VM must be deleted prior to doing that.

# 最佳实践: 性能

See [BRK4450](#) (Ignite 2015) for more

## IO 和磁盘吞吐

Are separate limits, enforced independently

Throttling kicks in when either of the limits are reached

Limits	IOPS	Throughput
P30 (1 TiB) Disk	5,000	200 MB/sec
DS3 VM	12,800	128 MB/sec

## 虚拟机及磁盘选择

Enforce independent limits on VM and Disk

Chose the right configuration based on workload needs

## 优化建议

Larger IO sizes for higher Throughput. Smaller IO sizes for higher IOPS.

Use multiple parallel operations / threads to push more traffic

Avoid heavy bursts by reducing the IO buffer to <5% of the IO limits

# 最佳实践: 可用性

## 生产环境使用高级磁盘

Provisioned disk performance for consistency  
Required for IO-intensive enterprise workloads

## 存储账户管理

Use random pre-fix for Standard Account Naming  
Limit to <40 disks per Standard Storage Account

## 冗余

Use VM availability Set with multiple VMs for redundancy  
Use separate storage accounts for VMs in the same Availability Set



# 最佳实践: 虚拟机及应用配置

## 缓存策略

Use "ReadOnly" cache for Data disks (Portal Default), ideal for Read-heavy and mixed workloads  
Use "None" cache for Write-only / Write-heavy Data disks (e.g. logs)  
Consider "ReadWrite" caching for Data Disks if the application handles IO flushing / durability  
Keep "ReadWrite" caching for OS Disks (Default)

## SQL Server

Separate disks for data and log files. And, use Premium Storage.  
Use the VM's temporary resource disk for Temp DB

## 其他建议

Use multiple instances for increasing throughput for single-threaded applications  
Example: setup rsync per folder  
Refer to Application best practices documents for details

# 最佳实践: 迁移

## 从客户数据中心到 Azure

Lift and Shift demanding enterprise workloads to the cloud

Optimize VM sizes based on application needs

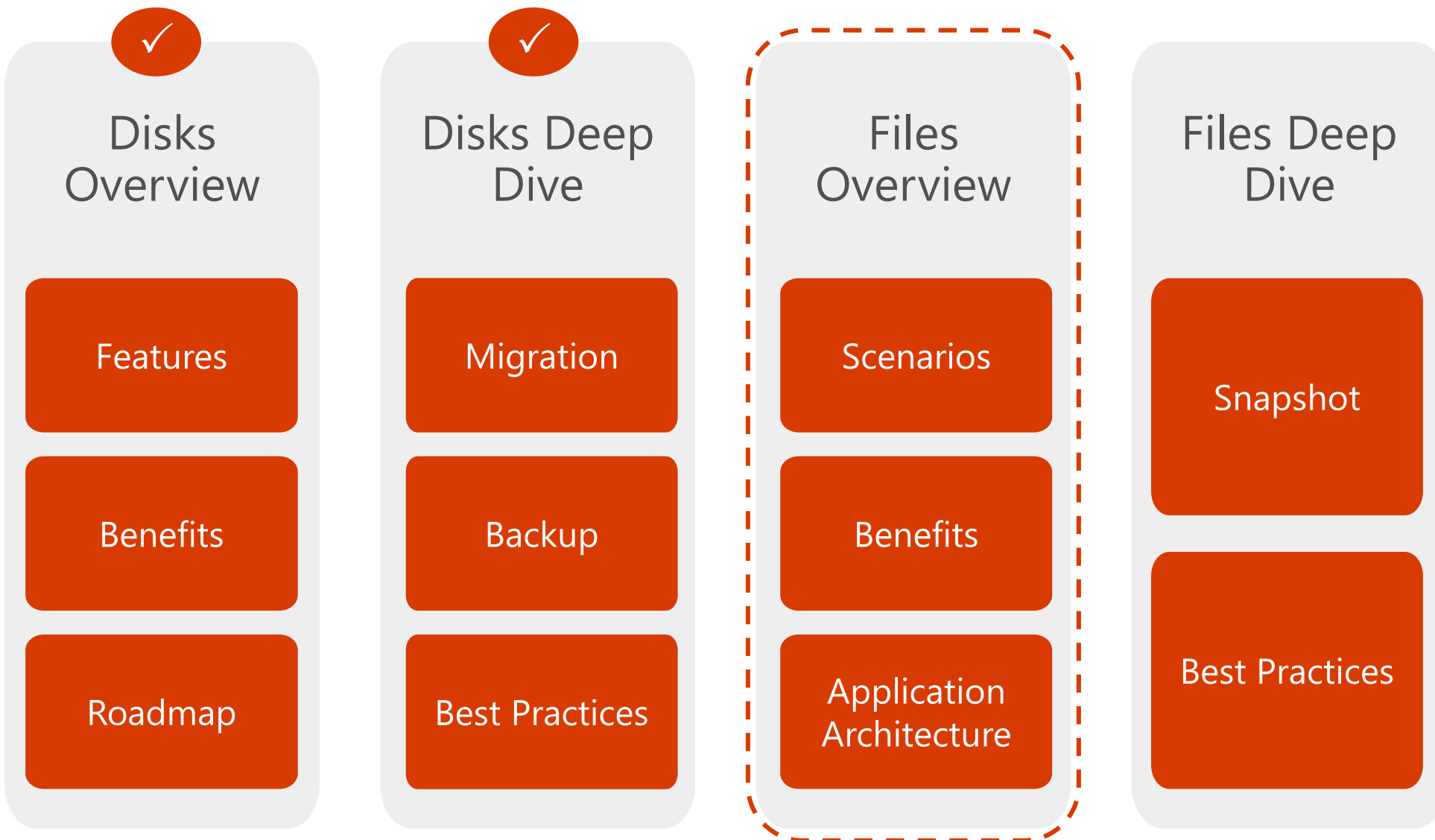
Optimize disks cost based on GB, IOPS and Throughput needs

## 迁移

Use AzCopy to migrate VHDs from On-Premise to Azure

Use ASR to shorten the downtime with live migration

# 内容一览



# 共享文件系统 Azure Files

# 简介

## 什么是 Azure Files

Fully Managed Cloud File Storage for use with IaaS and On Premise instances

## 应用场景

Lift and Shift

Host high availability workload data

Enable backup and disaster recovery

# 特性

## 协议及操作系统支持

SMB 2.1, 3.0, and REST

Mount from both Windows & Linux

## 易于访问

Accessible from both On Premise and IaaS instances

Available in all Azure regions including Azure China

## 高可用性

Replication for high availability (99.9% read/write SLA for LRS, GRS support)

# Azure Files 路线图

## Current

- Onboarding production workloads to Azure Files
- SMB and REST
- Standard Storage

## End of CY2016

- 1<sup>st</sup> party snapshots

## Future

- AD Integration
- Increase scale limits
- Larger share size
- Encryption at rest
- Backup integration

# Azure Files vs. Disks vs. Blobs

	Azure Files	Azure Disks	Azure Blobs
Scenario	Life & Shift applications which leverage native file system	Persistent disks to Azure Virtual Machines	Massively scale out object storage
Protocol	SMB 2.1/3.0, REST	VHD, REST	REST
Accessibility	SMB – Worldwide (requires Port 445) REST – Worldwide	VHD – Azure Data Center REST (Page Blob) – Worldwide	REST –Worldwide
Durability	LRS, GRS	LRS, GRS, RA-GRS (for Page Blob)	LRS, GRS, RA-GRS
Object Size	Up to 1 TB file	Up to 1 TB Disks (Can stripe up to 64 disks on G VM)	Up to 1 TB/Blob
Max IOPS (8K)	1000	5000 (Premium) 500 (Standard)	500 request/sec
Throughput	Up to 60 MB/s per share	Up to 200 MB/s per disks (Premium) Up to 60 MB/s per disk (Standard)	Up to 60 MB/sec per blob



# 客户实例

Customer Success: Global Collaboration  
with Talon CloudFAST™ and Azure Files



# Robert Bird Group

Global Specialists in Structural, Civil & Construction Engineering

Stephen Cottham

Chief Technology Officer

[www.robertbird.com](http://www.robertbird.com)

- 10 Offices Worldwide
- 600+ Employees Globally
- Offices in the Northern, Central, Southern and Western region's of Australia, United Kingdom and Middle East



# Introduction

## Talon – Microsoft Azure ISV Partner

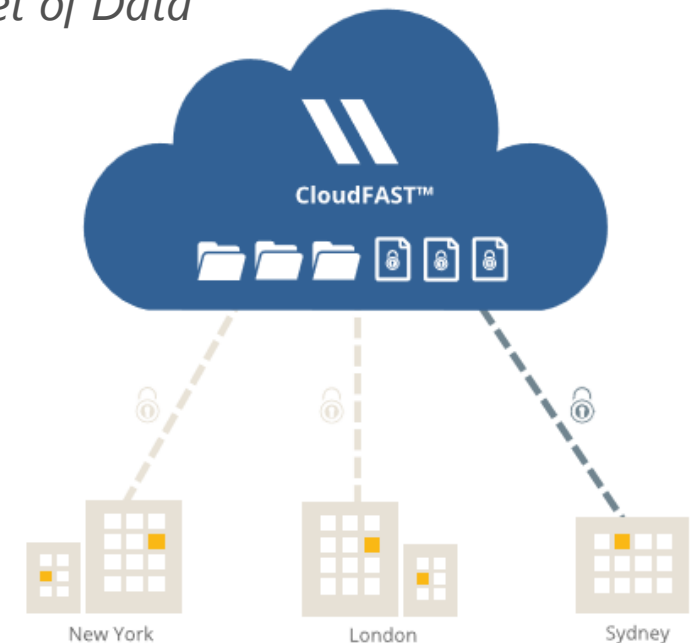
Jaap van Duijvenbode

Product Director

[www.talonstorage.com](http://www.talonstorage.com)



- **Software solution** that helps enterprises centralize file storage into a *"Single Set of Data"*
- **Centralizes your data** set into the Microsoft Azure Cloud
- Delivers instant **Enterprise Global File Sharing with Global File Locking**
- **High performance experience** at all geographically distributed locations
- **Eliminates** complexity, storage, infrastructure and backups at the branch
- **Cut storage costs up to 70% using Microsoft Azure**





# One Central Park

Architects: Jean Nouvel

Client: Frasers Property Group

**Winner of CTBUH Best Tall**

**Building Worldwide**

Images courtesy of Frasers Property Australia and Sekisui House Australia. Image © Simon Wood



# Robert Bird Group - Business and IT Needs



## ENABLE GLOBAL COLLABORATION

Need to be able to work in real-time with offices in Australia, Middle-East and UK



## LARGE FILE SHARING

Complex engineering models = large file sets, including application data like Autodesk, Microsoft Office and image files



## IMPROVE PRODUCTIVITY AND PERFORMANCE

Improve application performance and increase productivity for teams globally



## REDUCE STORAGE COST AND COMPLEXITY

Find better ways to reduce the storage footprint, eliminate 'doubling up' on storage space



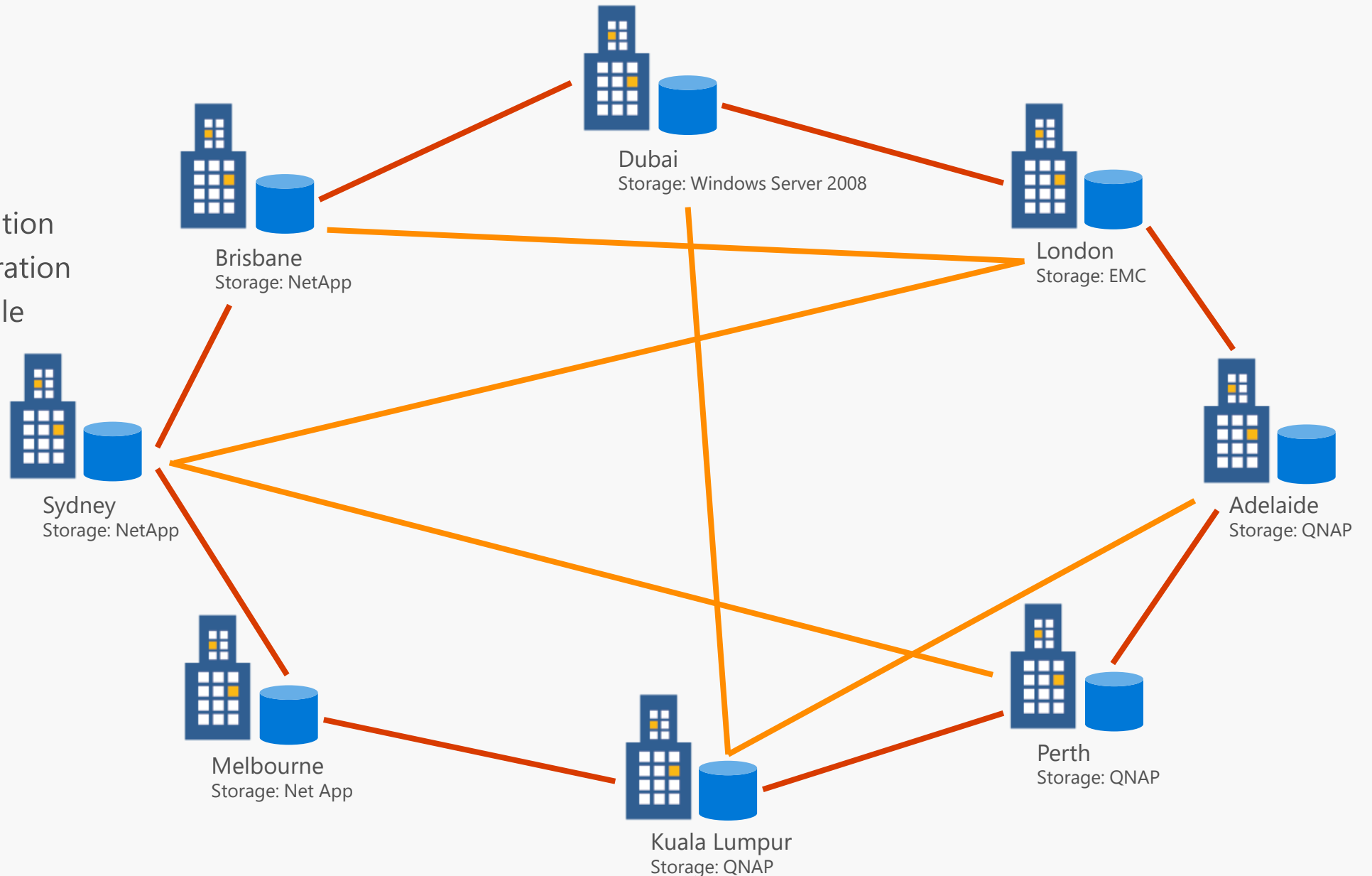
## SHARE SPECIALIST KNOWLEDGE AND WORK BETWEEN OFFICES

Access to data, anytime, anywhere for knowledge workers, meet project deadlines

# Before: De-Centralized = Collaboration Painful

## CHALLENGES

- Local storage in each site
- Local backups
- Trouble managing replication
- High latency / WAN saturation
- Collaboration is impossible





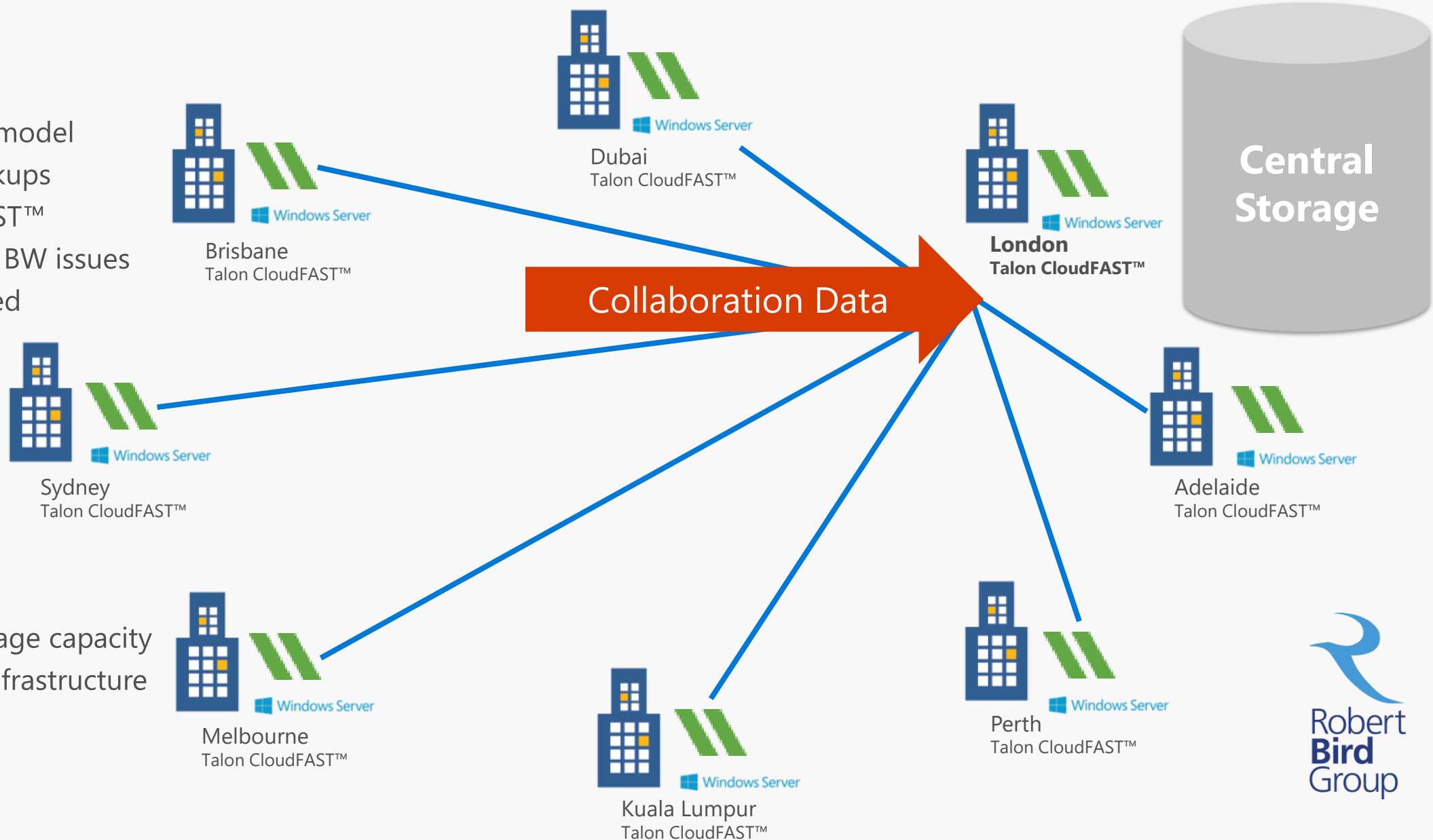
# Step 1: Centralize Data to 'on-premise'

## BENEFITS

- Centralized storage model
- Eliminated local backups
- Introduced CloudFAST™
- Eliminated Latency / BW issues
- Collaboration enabled

## CHALLENGES

- Increase central storage capacity
- Expensive storage/infrastructure
- MPLS connectivity

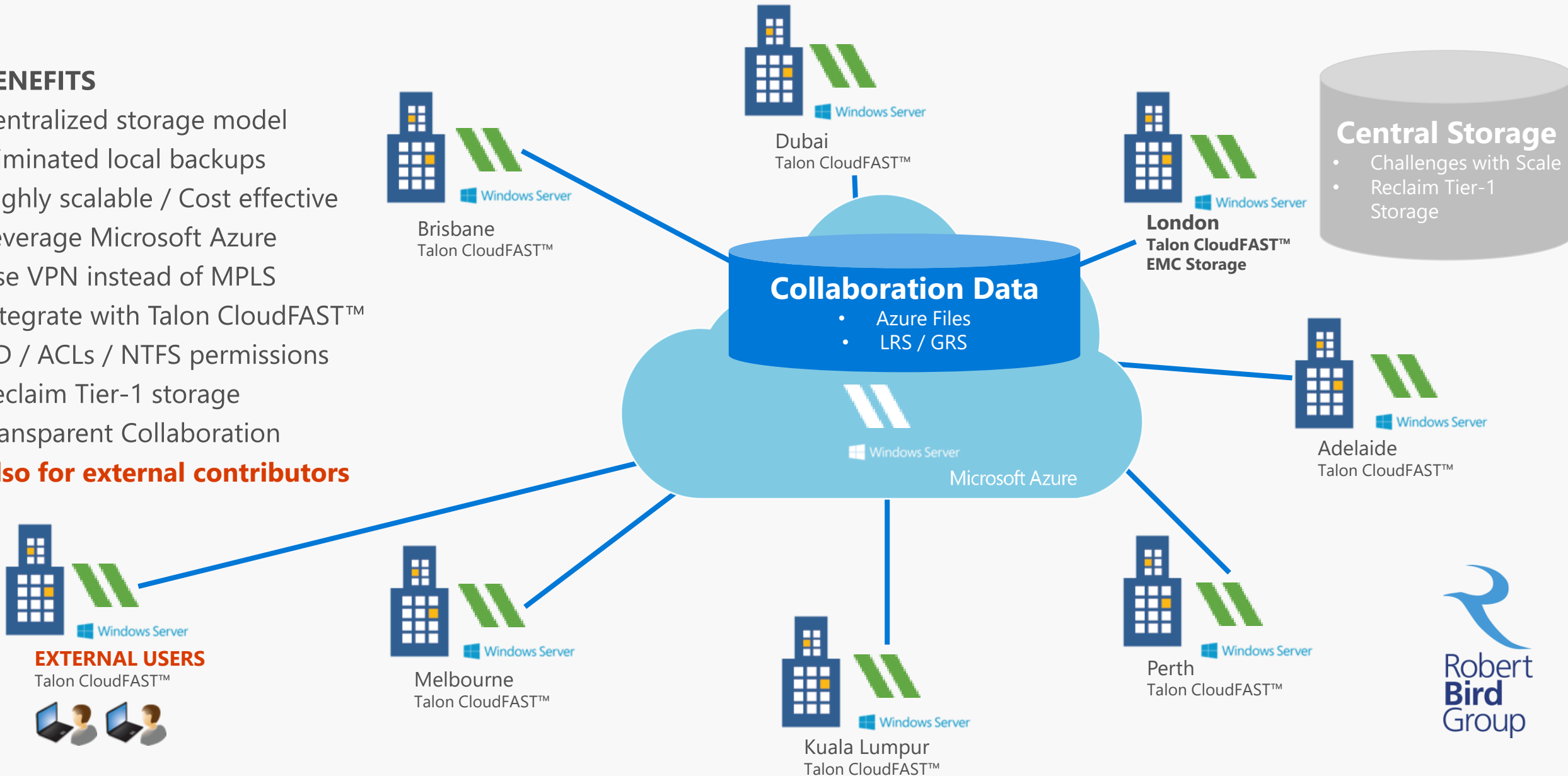


# Step 2: Leverage Microsoft Azure Files

## BENEFITS

- Centralized storage model
- Eliminated local backups
- Highly scalable / Cost effective
- Leverage Microsoft Azure
- Use VPN instead of MPLS
- Integrate with Talon CloudFAST™
- AD / ACLs / NTFS permissions
- Reclaim Tier-1 storage
- Transparent Collaboration

**Also for external contributors**





# Next steps for Robert Bird Group

Move from a traditional 'on-premise' environment

How to leverage Azure IaaS optimally

- 100's of different Azure Solutions available today, growing every month
- Azure IaaS as an alternative to traditional infrastructure
- OMS / LogAnalytics, Azure Automation, etc.


Improved Data Management

- Moving active and archive unstructured data from 'on-premise' to Azure
- NTFS Permissions and ACLs for internal and external collaboration

Backups & Snapshots?

Scale at large > 100TB's

# 内容一览




Disks  
Overview

Features

Benefits

Roadmap




Disks Deep  
Dive

Migration

Backup

Best Practices



Files  
Overview

Scenarios

Benefits

Application  
Architecture

Files Deep  
Dive

Snapshot

Best Practices

# Azure Files 深度了解: 文件快照

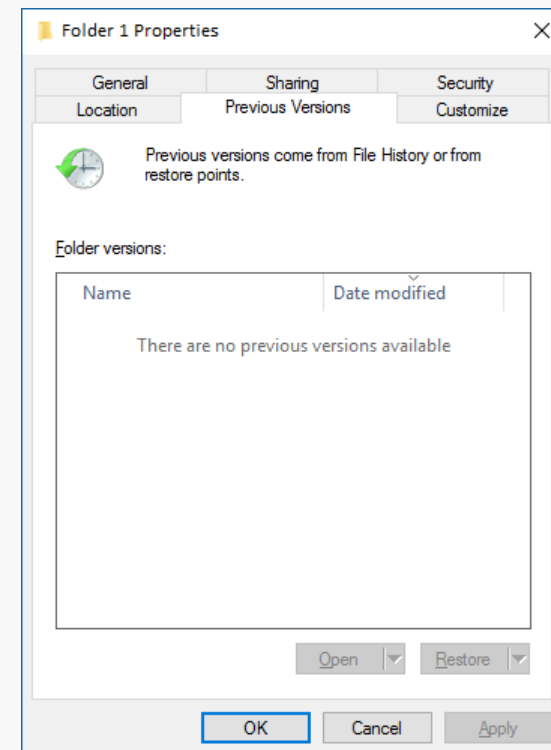
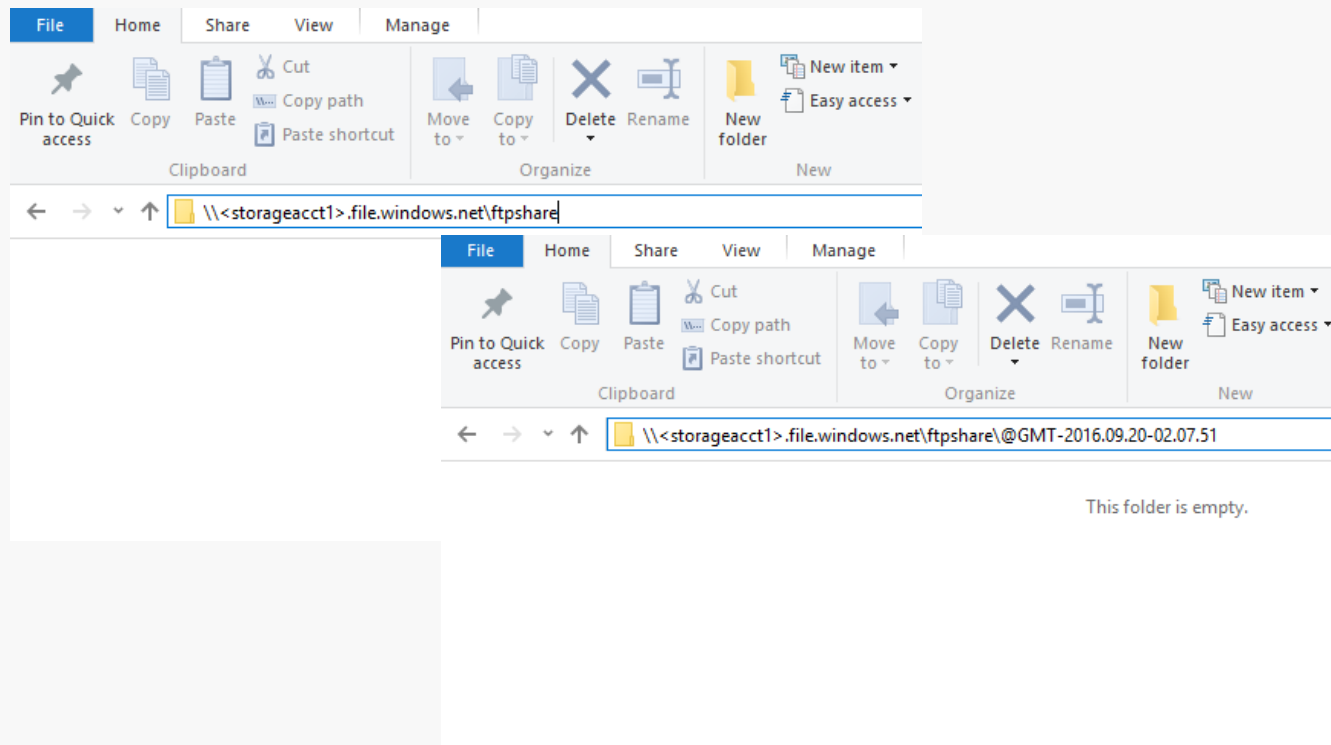
## Protect and recover your data using share snapshots

Create up to 100 per share, at most every minute

REST API, PowerShell and CLI for managing snapshots

Previous versions integration for viewing snapshot content

Application consistency



# Azure Files 深度了解: 最佳实践

## 性能

Use **1MB IO Size** for best throughput

Ensure **KB3114025** applied on Windows 8.1 or Server 2012 R2

## 安全

Use **Win 8.x or Server 2012** as Windows clients

Use **SMB 2.1 in the same DC** for Linux clients

Mount file for **correct user context**

Enable **Port 445 outbound** to access Azure Files from On Premise instances

## 易用性

Write to a file periodically to **keep connection open in Linux**, until [kernel fix](#) is backported

Mount from **logged in user context** to see drive in **Windows Explorer**

# 设计模式 – 云端的FTP服务器

ftp://ftp.domain-one.com/pub/

Azure



Load Balancer

App/FTP  
Server  
VM 1

App/FTP  
Server  
VM 2

App/FTP  
Server  
VM 3

\\<storageacct1>.file.windows.net\ftpshare

Azure files

Folder 1

Share 1

Folder 2

Sub Folder 1

Azure

Scenario 1

# 演示

Azure File Share

# Learn More

## Azure Disks Resources

[Premium Storage: High-Performance Storage for Azure Virtual Machine Workloads](#)

[Design for High Performance](#)

[Migrating to Azure Premium Storage](#)

[Blob Service REST Operations with Azure Premium Storage](#)

[Backup with Incremental Snapshot copying](#)

[Azure VM specifications](#)

[Mark's Channel 9 video on Premium Storage](#)

## Azure Files Resources

Getting started blog with many useful links: <http://blogs.msdn.com/b/windowsazurestorage/archive/2014/05/12/introducing-microsoft-azure-file-service.aspx>

SMB features currently not supported: <https://msdn.microsoft.com/en-us/library/azure/dn744326.aspx>

Naming restrictions for REST compatibility: <https://msdn.microsoft.com/library/azure/dn167011.aspx>

Scalability targets: <https://azure.microsoft.com/en-us/documentation/articles/storage-scalability-targets/>

Various ISPs and Port 445 status: <http://social.technet.microsoft.com/wiki/contents/articles/32346.azure-summary-of-isps-that-allow-disallow-access-from-port-445.aspx>