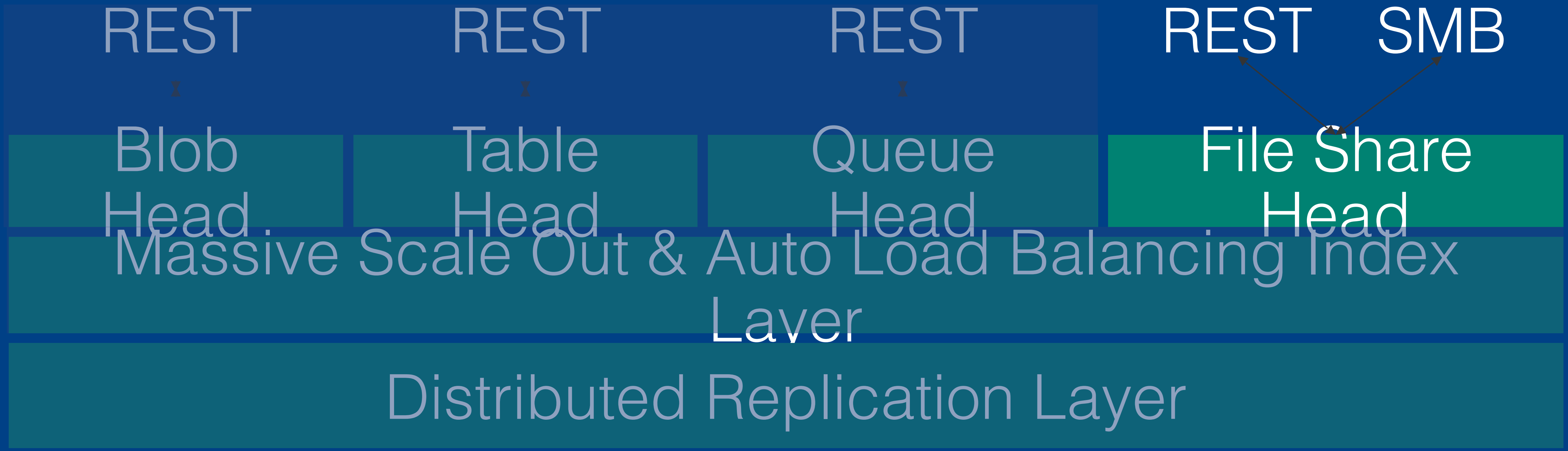
Microsoft Azure

Azure Data Storage

Microsoft

Azure Storage Architecture

Microsoft Azure


REST
REST
REST
REST
SMB
Blob Head
Table Head
Queue Head
File Share Head
Massive Scale Out & Auto Load Balancing Index Layer
Distributed Replication Layer

Microsoft

Blobs

Microsoft Azure Storage Blob

Two Types of Blobs Under the Hood

Block Blob

Page Blob

Block Blob

Targeted at streaming workloads

Each blob consists of a sequence of blocks

Each block is identified by a Block ID

Size limit 200GB per blob

Optimistic Concurrency via Etags

Page Blob

Targeted at random read/write workloads

Each blob consists of an array of pages

Each page is identified by its offset from the start of the blob

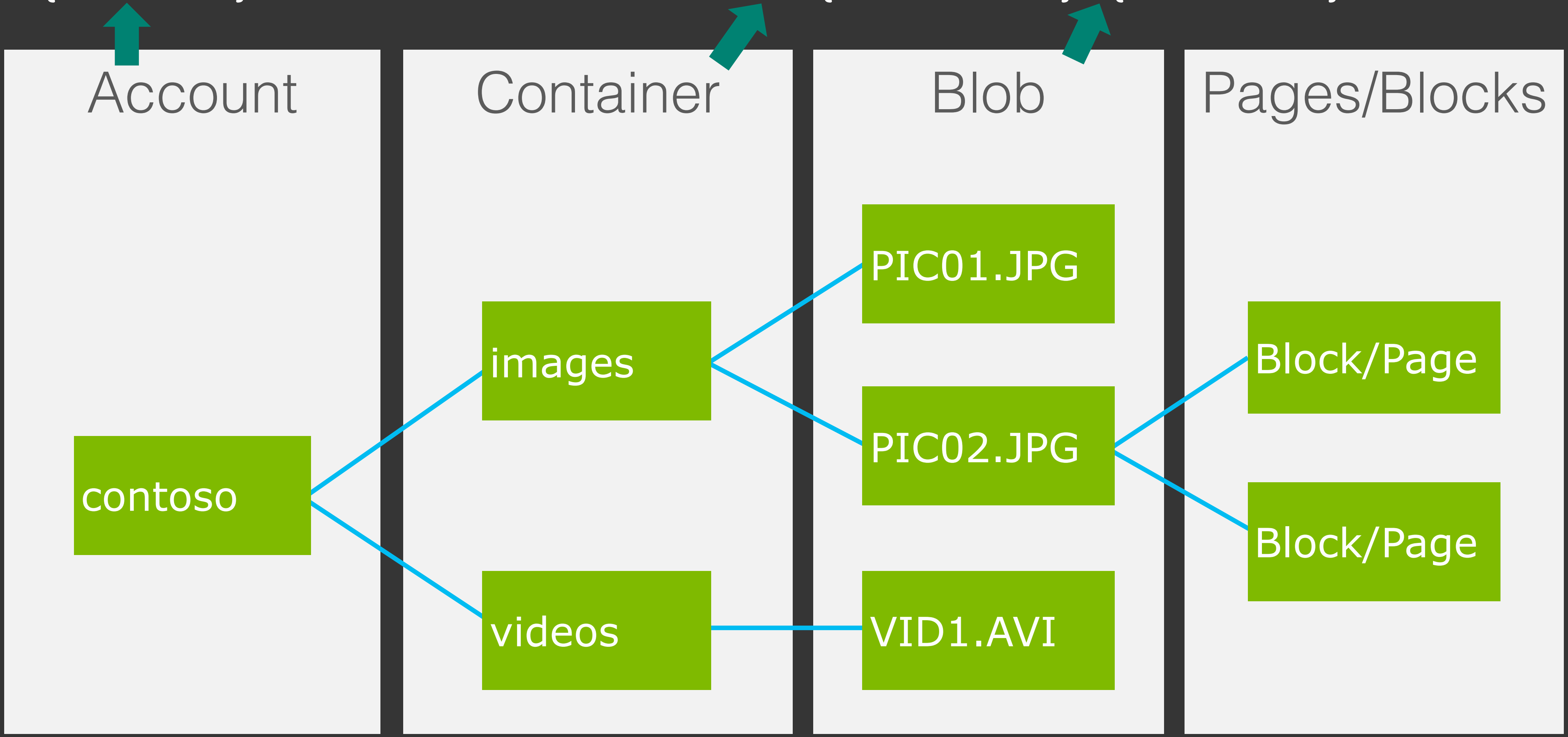
Size limit 1TB per blob

Optimistic or Pessimistic (locking) concurrency via leases

Blob Storage Concepts

Microsoft Azure

```
http://{account}.blob.core.windows.net/{container}/{blobname}
```



Microsoft

Microsoft Azure

Blob Details – Containers

- Multiple Containers per Account
- Special $root container

Microsoft

Microsoft Azure

Blob Details – Containers

- A container holds a set of blobs
- Set access policies at the container level
- Associate Metadata with Container
- List the blobs in a container
- Including Blob Metadata and MD5

  no search on metadata WHERE MetadataValue = ?

Microsoft

Blob Details – Throughput

- Effectively in Partition of 1
- Target of 60MB/s per Blob

PutBlob

GetBlob

DeleteBlob

CopyBlob

SnapshotBlob

LeaseBlob

Blob Details

Microsoft Azure

Associate metadata with blob

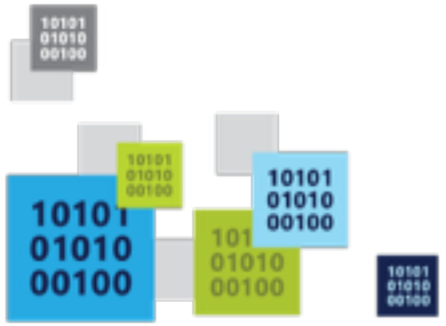
- Standard HTTP metadata/headers (Cache-Control, Content-Encoding, Content-Type, etc)
- Metadata is <name, value> pairs, up to 8KB per blob
- Either as part of PutBlob or independently

Microsoft

Blob Details – Blob always accessed by name

Can include '/' or other delimiter in name

$$\text{e.g. /<container>/}\underbrace{\text{myblobs/smurf.png}}_{\text{blob name}}$$



Blob Details

GET Blob operation takes parameters

Prefix

Delimiter

Include = (snapshots, metadata etc...)

Blob sample listing

http://
adventureworks.blob.core.windows.net/
Products/Bikes/
SuperDuperCycle.jpg
Products/Bikes/FastBike.jpg
Products/Canoes/Hybrid.jpg
Products/Canoes/Flatwater.jpg
Products/Canoes/Whitewater.jpg
Products/Tents/PalaceTent.jpg
Products/Tents/ShedTent.jpg

GET http://.../products?comp=list&prefix=Tents

```
<Blobs>
	<Blob><Name>Tents/PalaceTent.jpg</Name>[…]</
Blob>
	<Blob><Name>Tents/ShedTent.jpg</Name>[…]</
```

Blob sample listing full response

```
<Blobs>
	<Blob>
		<Name>Tents/PalaceTent.jpg</Name>
		<Url>https://readinesscloudcamp.blob.core.windows.net/products/Tents/PalaceTent.jpg</Url>
		<LastModified>Wed, 17 Dec 2014 09:00:26 GMT</LastModified>
		<Etag>0x8D1E7EF08F31520</Etag>
		<Size>150027</Size>
		<ContentType>image/jpeg</ContentType>
		<ContentEncoding />
		<ContentLanguage />
	</Blob>
	<Blob>
		<Name>Tents/ShedTent.jpg</Name>
		<Url>https://readinesscloudcamp.blob.core.windows.net/products/Tents/ShedTent.jpg</Url>
		<LastModified>Wed, 17 Dec 2014 09:00:26 GMT</LastModified>
		<Etag>0x8D1E7EF08EA6257</Etag>
		<Size>150027</Size>
		<ContentType>image/jpeg</ContentType>
		<ContentEncoding />
		<ContentLanguage />
	</Blob>
</Blobs>
```

Blob sample listing with maxresults

http://
adventureworks.blob.core.windows.net/
Products/Bikes/
SuperDuperCycle.jpg
Products/Bikes/FastBike.jpg
Products/Canoes/Hybrid.jpg
Products/Canoes/Flatwater.jpg
Products/Canoes/Whitewater.jpg
Products/Tents/PalaceTent.jpg
Products/Tents/ShedTent.jpg

http://.../products?
comp=list&prefix=Canoes&maxresults=2

```
<Blob>Canoes/Hybrid.jpg</Blob>
<Blob>Canoes/Flatwater.jpg</Blob>
<NextMarker>1!28!Q2Fub2VzL1doaXRld2F0ZXIuanBn</
NextMarker>
```

Blob sample listing with maxresults

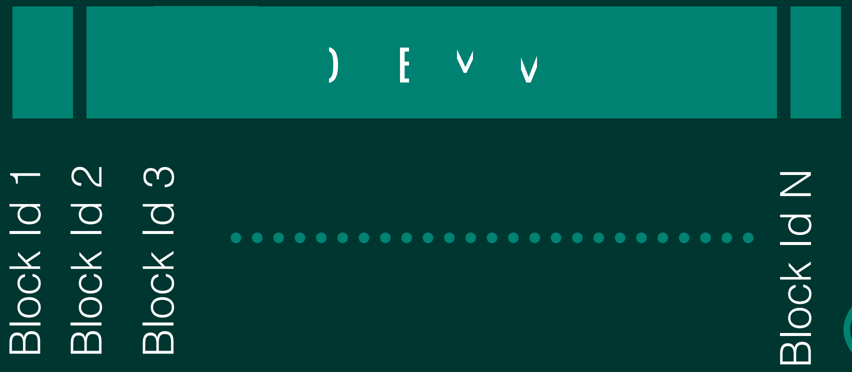
http://
adventureworks.blob.core.windows.net/
Products/Bikes/
SuperDuperCycle.jpg
Products/Bikes/FastBike.jpg
Products/Canoes/Hybrid.jpg
Products/Canoes/Flatwater.jpg
Products/Canoes/Whitewater.jpg
Products/Tents/PalaceTent.jpg
Products/Tents/ShedTent.jpg

http://.../products?
comp=list&prefix=Canoes&maxresults=2
&marker=1!28!Q2Fub2VzL1doaXRld2F0ZXIuanBn

<Blob>Canoes/Whitewater.jpg</
Blob>
</NextMarker>

Uploading a Block Blob

Uploading



THE BLOB

```
blobName = “TheBlob.wmv”;
PutBlock(blobName, blockId1,
block1Bits);
PutBlock(blobName, blockId2,
block2Bits);
…………
PutBlock(blobName, blockIdN,
blockNBits);
PutBlockList(blobName,
        blockId1,…,blockIdN);
```

TheBlob.wmv

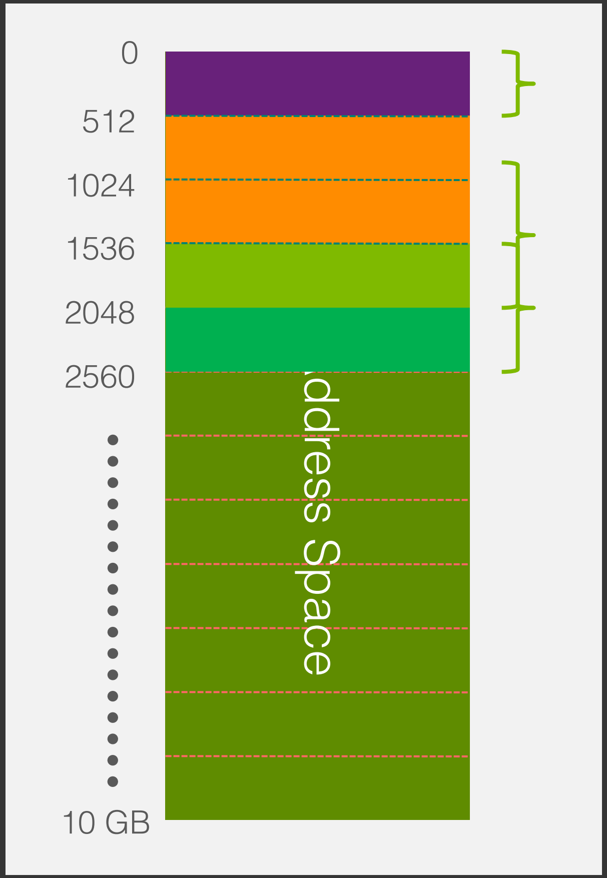
Blob block uploading benefits

Efficient continuation and retry

Parallel and out of order upload of blocks

Microsoft Azure

Page Blob – Random Read/Write



Create blob and specify Blob Size = 10 Gbytes

Fixed Page Size = 512 bytes

Random Access Operations:

PutPage[512, 2048)

PutPage[0, 1024)

ClearPage[512, 1536)

PutPage[2048,2560)

GetPageRange[0, 4096) returns valid data ranges:

[0,512) , [1536,2560)

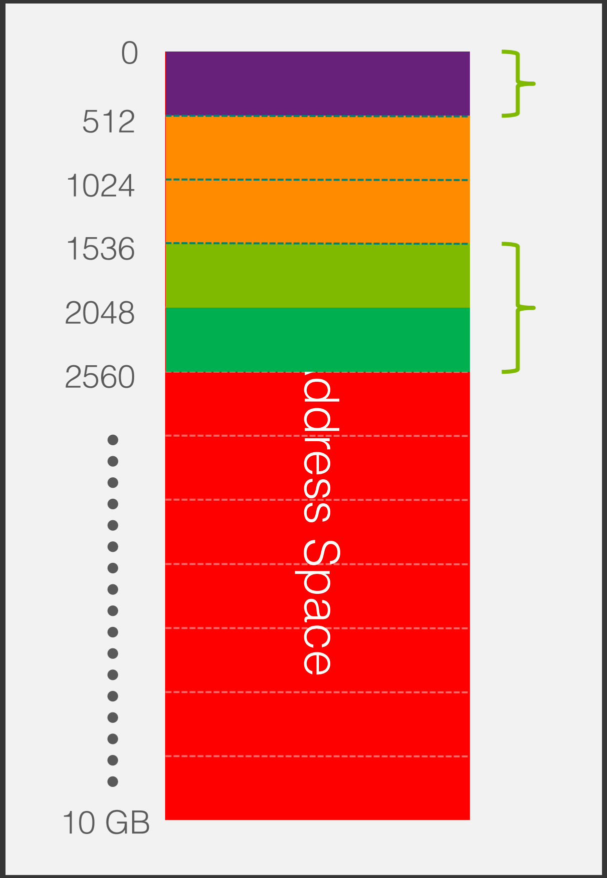
GetBlob[1000, 2048) returns:

All 0 for first 536 bytes

Next 512 bytes data stored in [1536,2048)

Microsoft

Page Blob – Random Read/Write



Sparse storage:
Only charged for pages with data stored in them

Shared Access Signatures

Fine grain access rights to blobs and containers

Sign URL with storage key – permit elevated rights

Microsoft Azure

Shared Access Signatures – Two broad approaches

Ad-hoc:
Stored Access Policy

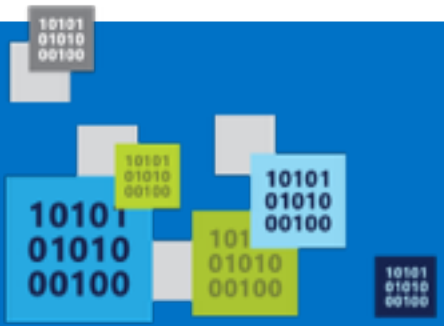
Policy based:
Shared Access Signature

Microsoft

Shared Access Signatures – Revocation

Use short time periods and re-issue

Use container level policy that can be deleted



Shared Access Signatures – Ad Hoc Signatures

Create Short Dated Shared Access Signature

- Signed resource Blob or Container
- AccessPolicy Start, Expiry and Permissions
- Signature HMAC-SHA256 of above fields

Shared Access Signatures – Ad Hoc Signatures

Use case

- Single use URLs
- E.g. Provide URL to mobile client to upload to container

Shared Access Signatures
Ad Hoc Signatures

Microsoft Azure

```
http://...blob.../pics/image.jpg?
sr=c&st=2009-02-09T08:20Z&se=2009-02-10T08:30Z&sp=w
&sig= dD80ihBh5jfNpymO5Hg1IdiJIEvHcJpCMiCMnN%2fRnbI%3d
```

Microsoft

Microsoft Azure

Store Access Policy – Policy Based Signatures

Create Container Level Policy

- Specify StartTime, ExpiryTime, Permissions

Microsoft

Store Access Policy – Policy Based Signatures

Create Shared Access Signature URL

- Signed resource Blob or Container
- Signed identifier Optional pointer to container policy
- Signature HMAC-SHA256 of above fields

Store Access Policy – Policy Based Signatures

Use case

Providing revocable permissions to certain users/groups

To revoke: Delete or update container policy

Store Access Policy
Policy Based Signatures

Microsoft Azure

```
http://...blob.../pics/image.jpg?
sr=c&si=MyUploadPolicyForUserID12345
&sig=dD80ihBh5jfNpymO5Hg1IdiJIEvHcJpCMiCMnN%2fRnbI%3d
```

Microsoft

Microsoft Azure

Files

Microsoft

Azure Files – Customer Quotes

Microsoft Azure

"I wish I could go to storage and provision a cloud drive, giving it a namespace, and that drive would then be UNC-addressable by the OSes."

"I need two VM's running with a shared drive. One will write to the drive, the other will read [it]."

"Hi, I have two VM's in Microsoft Azure. All I want to do is set up a file share between them. Is this possible?"

"Is it possible to share a secondary disk between different VM instances?"

Microsoft

Microsoft Azure

Sharing Files – The old way

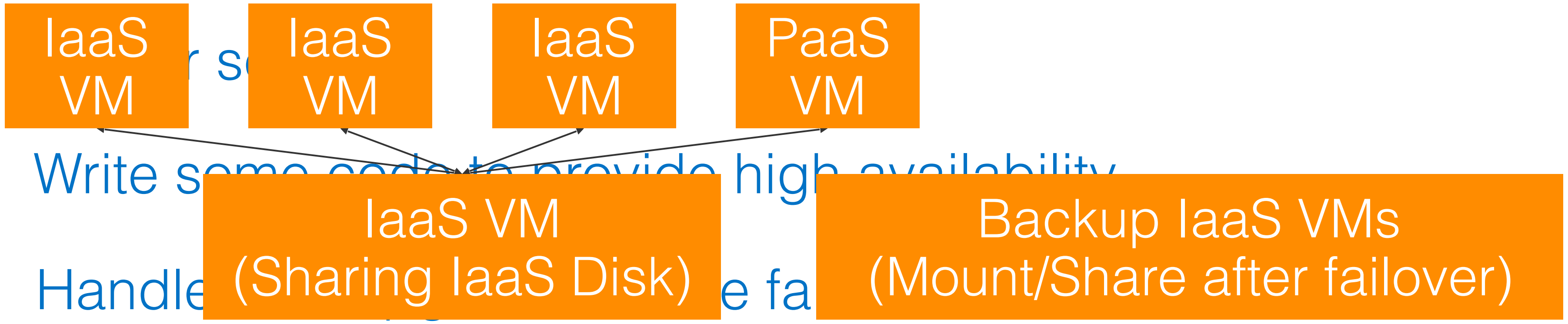
Setup an IaaS VM to host a File Share backed by an IaaS Disk

Write code to find the IaaS File Share from the rest of the VMs

r s

Write some code to provide high availability

Handle e fa

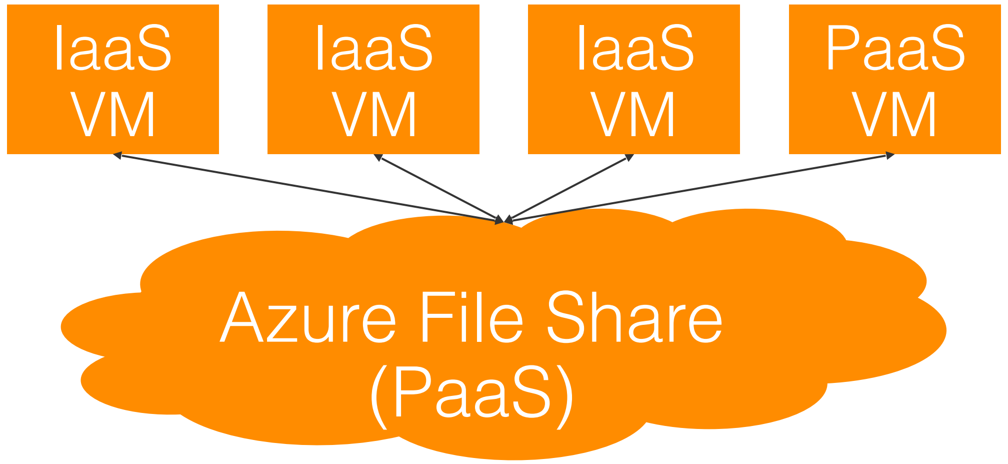

IaaS VM
IaaS VM
IaaS VM
PaaS VM
IaaS VM (Sharing IaaS Disk)
Backup IaaS VMs (Mount/Share after failover)

You can only access the File Share from other VMs

Microsoft

Azure Files

- Shared Network File Storage for Azure
- Availability, durability, scalability are managed automatically
- Supports two interfaces: SMB and REST



Azure Files – Usage

Microsoft Azure

- Share data across VMs and applications
- Share settings throughout services
- Dev/Test/Debug

Microsoft

Queues

Why use a Queue?

Microsoft Azure

Queue length reflects how well the backend processing nodes are doing.

Decouples the application.

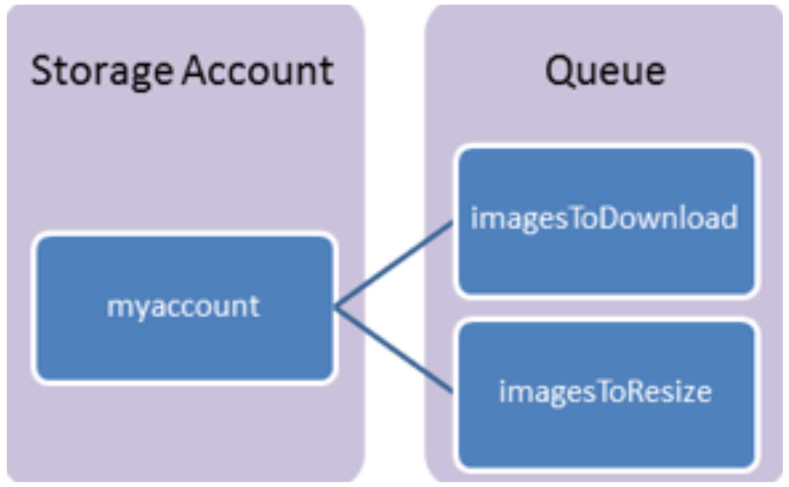
Flexibility of efficient resource usage within an application.

Absorb traffic bursts and reduce the impact of individual component failures.

Microsoft

Queue Components

- Storage Account: All access to Azure Storage is done through a storage account.
- Queue: A queue contains a set of messages.
- Message: A message, in any format, of up to 64KB.



Microsoft Azure

Queue URL format

Queues are addressable using the following URL format:

http://{storage-account}.queue.core.windows.net/{queue}

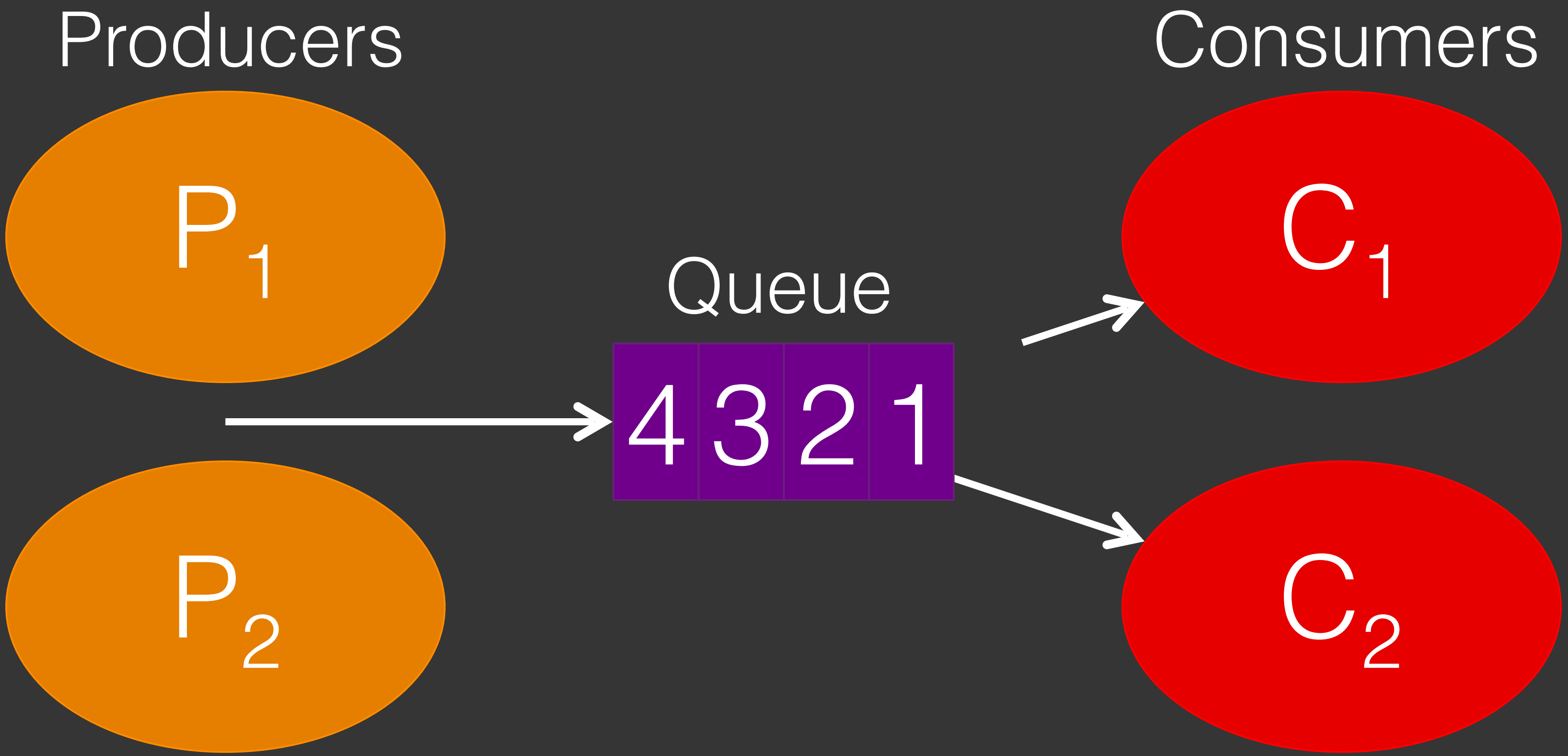
Microsoft

Queue URL format

Example:

http://myaccount.queue.core.windows.net/imagesToDownload

Queue-based Load Levelling Pattern



Queue Considerations

Messages are ordered but not guaranteed FIFO.

Message will be processed at least once.

Message may be processed more than once.

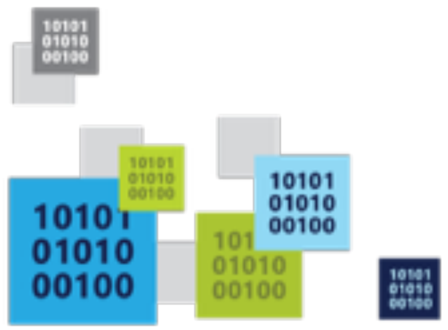
.DequeueCount increases every time.

-> Processing must be idempotent.

Microsoft Azure

Queue Considerations

Messages are stored up to 7 days

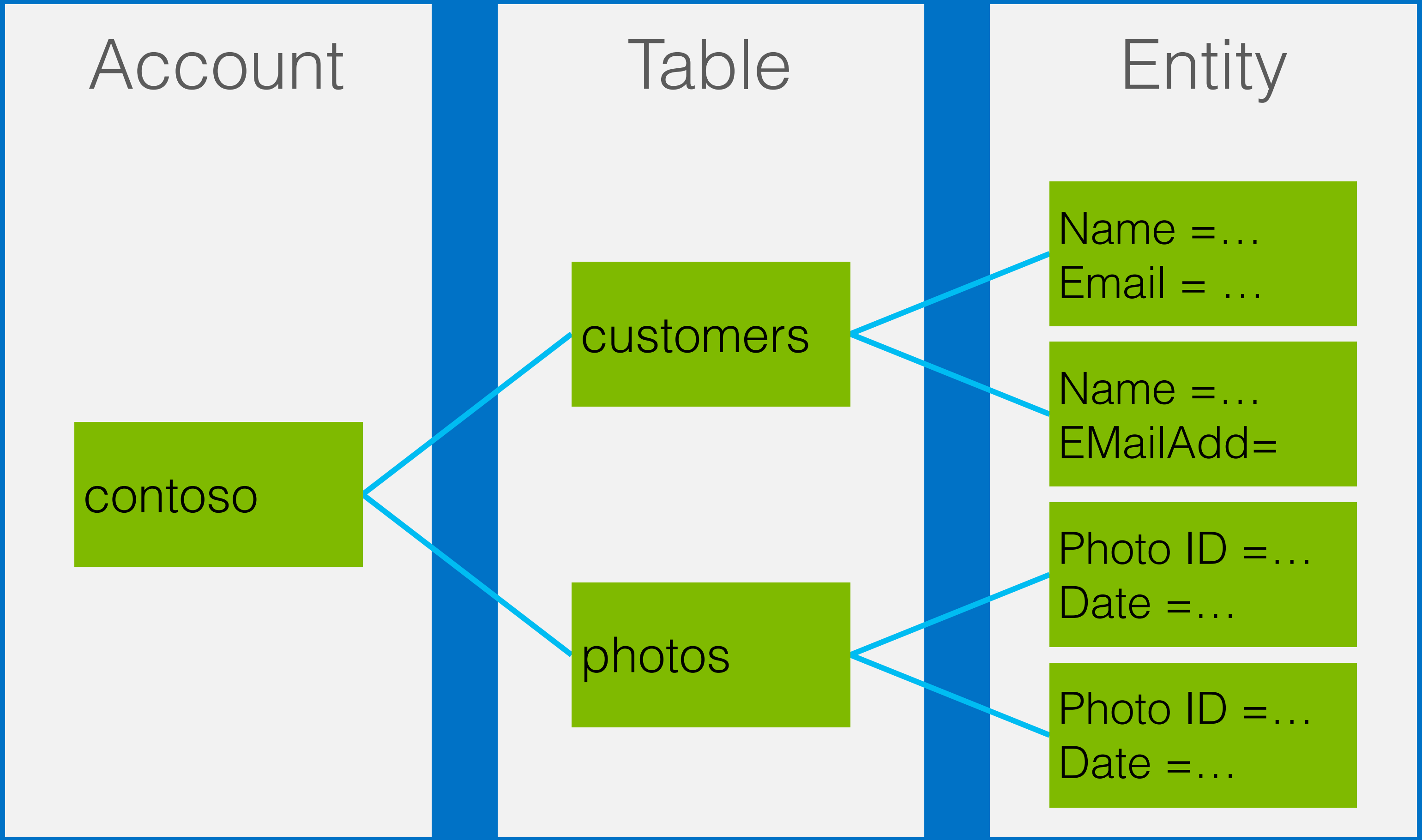


Microsoft

Tables

Table Storage Concepts

Microsoft Azure



Microsoft

Table Storage Details

Not an RDBMS Table!

The mental picture is ‘Entities’

Microsoft Azure

Table Storage Details

Entity can have up to 255 properties

Up to 1MB per entity

Microsoft

Table Storage Details

Microsoft Azure

Entity Properties

PartitionKey & RowKey are mandatory properties

Composite key which uniquely identifies an entity

They are the only indexed properties

Defines the sort order

Microsoft

Table Storage Details

Microsoft Azure

Purpose of the PartitionKey

Entity Locality

Entity Group Transactions

Table Scalability

Microsoft

Microsoft Azure

Table Storage Details

Purpose of the PartitionKey

Entity Locality

- Entities in the same partition will be stored together
- Efficient querying and cache locality
- Endeavour to include partition key in all queries

Microsoft

Table Storage Details

Microsoft Azure

Purpose of the PartitionKey

Entity Group Transactions

Atomic multiple CRUD in same partition in a single transaction

Microsoft

Table Storage Details

Microsoft Azure

Purpose of the PartitionKey

Table Scalability

Target throughput – 500 tps/partition, several thousand tps/account

Microsoft Azure monitors the usage patterns of partitions

Microsoft

Microsoft Azure

Table Storage Details

Purpose of the PartitionKey

Table Scalability

- Automatically load balance partitions
- Each partition can be served by a different storage node
- Scale to meet the traffic needs of your table

Microsoft

Table Storage Details

Entity Properties

Timestamp property

Optimistic Concurrency

Exposed as an HTTP Etag

Microsoft Azure

Table Storage Details

Entity Properties

No fixed schema for other properties

Each property is stored as: <name, typed value>

Properties can be the standard .NET types:

string, binary, bool, DateTime, GUID, int, int64, double

Microsoft

Table Storage Details

Supports full manipulation (CRUD)

Including Upsert and Entity Group Transactions

Tables can have metadata

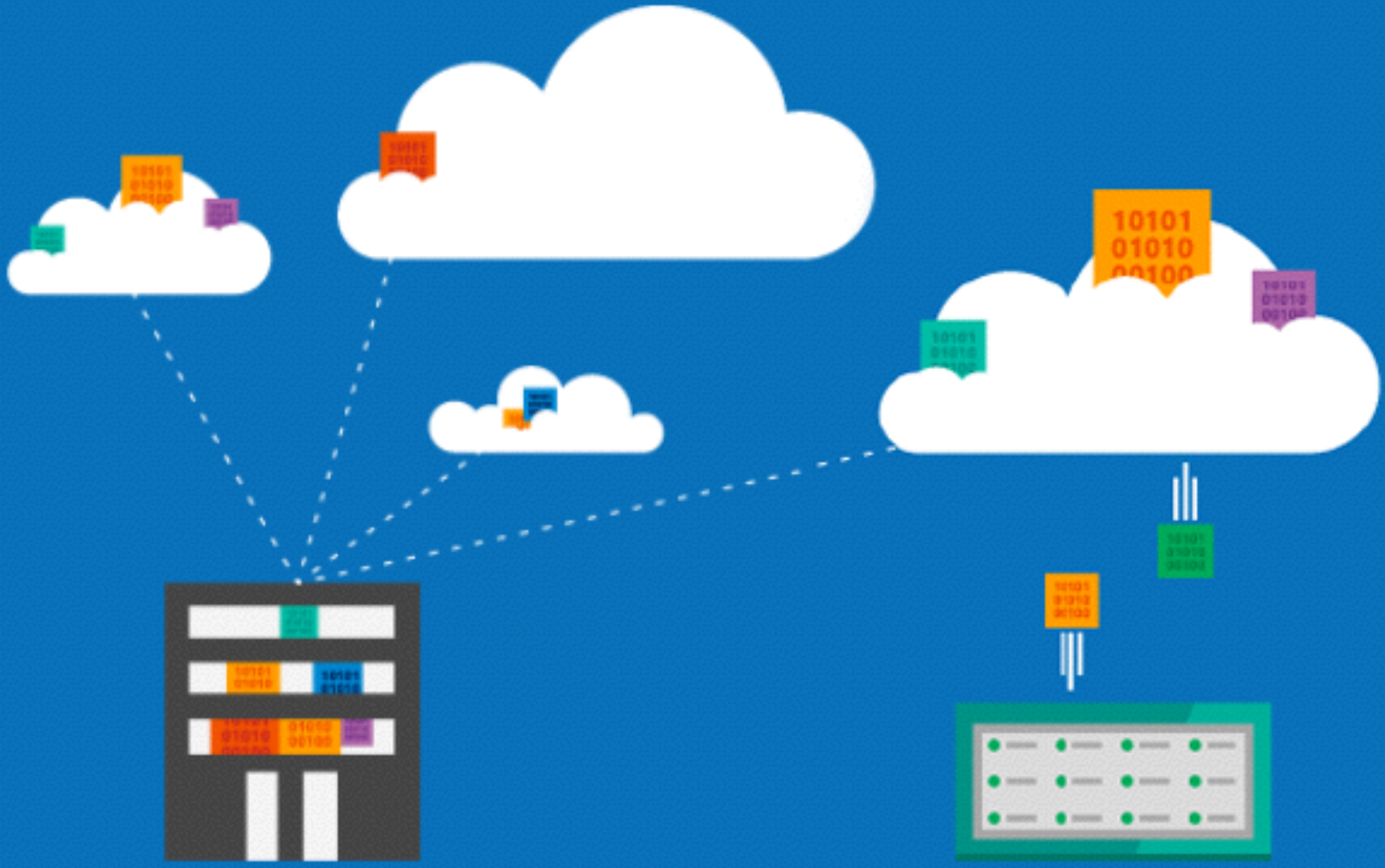
Microsoft Azure

StorSimple

Microsoft

Microsoft Azure

StorSimple + Microsoft Azure = Hybrid Cloud Storage



Microsoft

Microsoft Azure

StorSimple

Designed to:

Reduce storage costs

Simplify storage management

Improve disaster recovery capability and efficiency

Provide data mobility.

Microsoft