

Classified Hybrid Cloud Storage Based on Ceph RGW

UMCloud
Enming Zhang
2018.06.22



Outline

- Hybrid Cloud Storage Trend
- Current Solution and Its Limitations
- Our Solution
 - Storage Class
 - LC Transition
 - Auto-Generated Transition Policy
- Further Work

HYBRID CLOUD STORAGE TREND

Public Cloud Storage

- Unlimited Capacity
- Low Cost
- Performance not well
- Security and Controllability not Well

Hybrid Cloud Storage

- High Performance
- High Security and Controllability
- Unlimited Capacity
- Relatively Low Cost

Private Cloud Storage

- High Performance
- High Security and Controllability
- Limited Capacity
- High Cost

CURRENT SOLUTION AND ITS LIMITATIONS

- RGW Multisite Intro
- RGW Multisite-Based Cloud Sync
- Limitations of RGW Cloud Sync

CURRENT SOLUTION AND ITS LIMITATIONS

RGW Multisite Intro

- zone
a zone is *logical* grouping of one or more Ceph Object Gateway instances.

CURRENT SOLUTION AND ITS LIMITATIONS

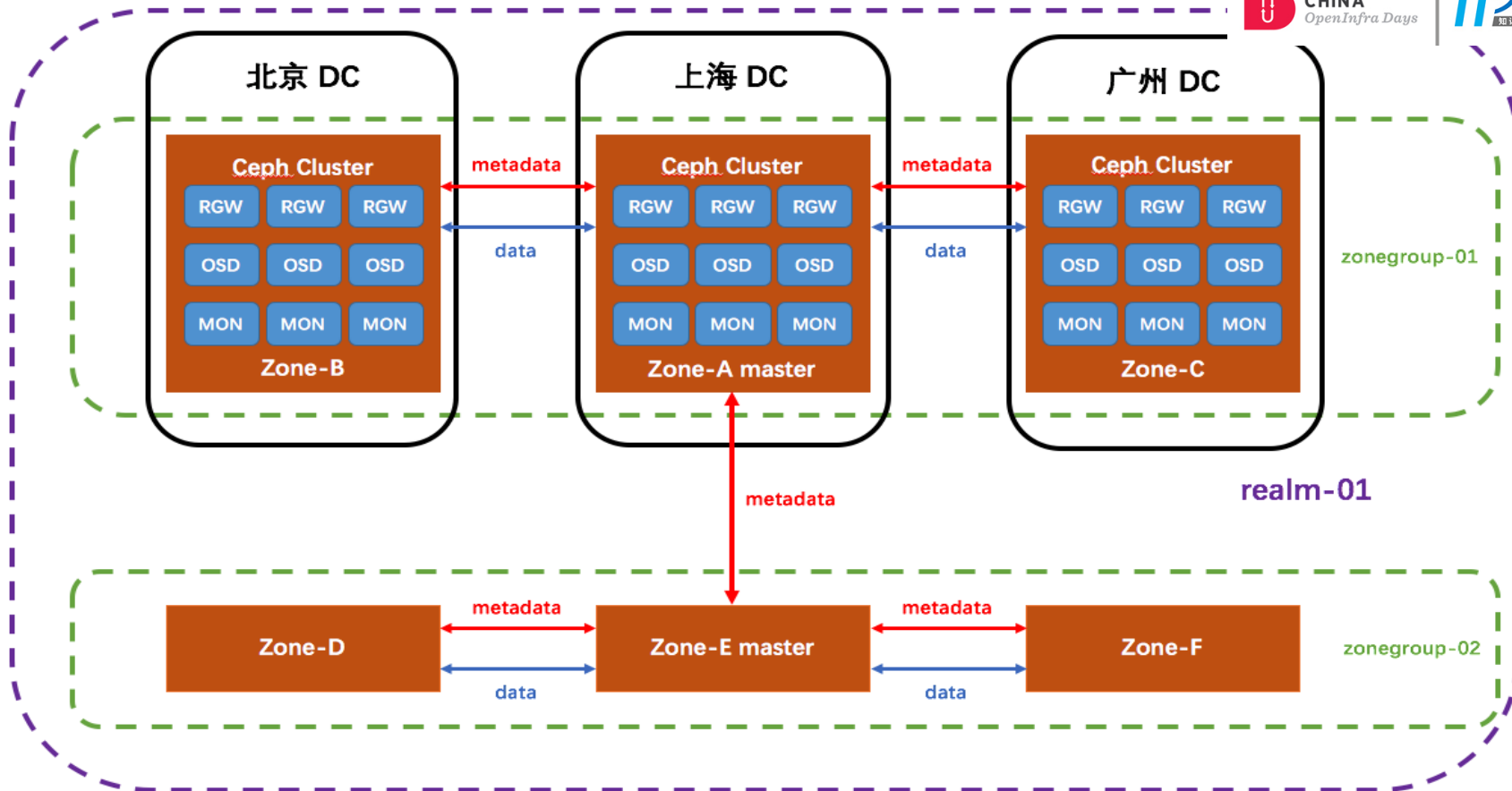
RGW Multisite Intro

- zonegroup
a zonegroup consists of multiple zones, metadata and data will be synced among zones in a same zonegroup.

CURRENT SOLUTION AND ITS LIMITATIONS

RGW Multisite Intro

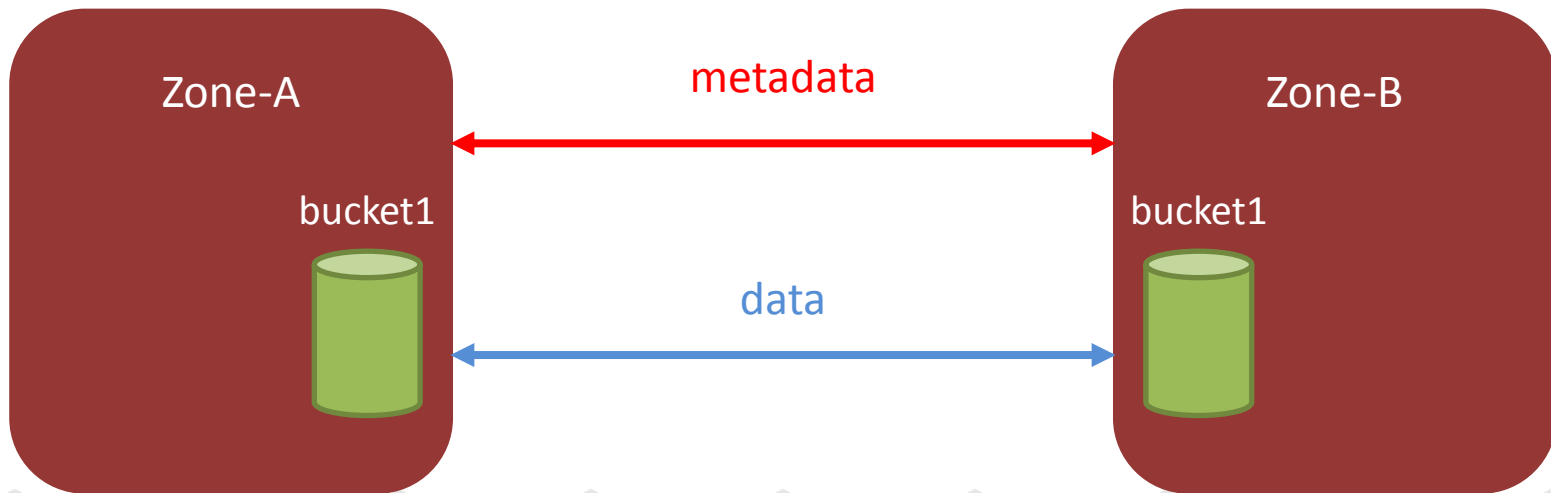
- realm
 - a realm is a container for zonegroups, metadata will be synced among zonegroups in a same realm.



CURRENT SOLUTION AND ITS LIMITATIONS

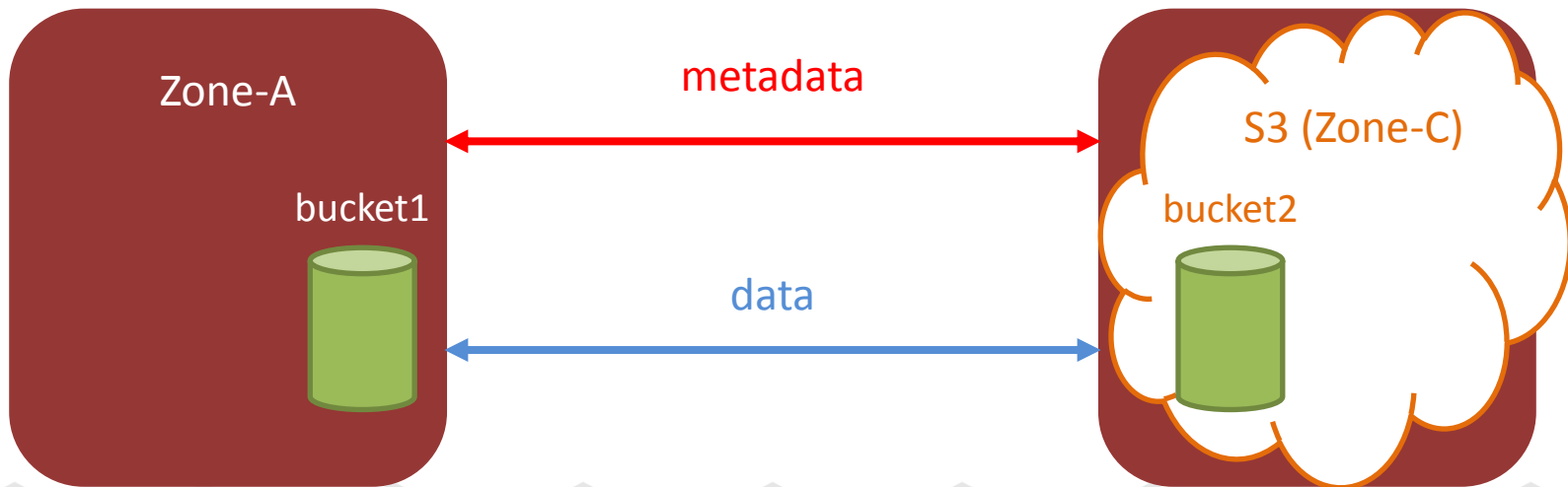
RGW Multisite Intro

bucket sync enable
bucket sync disable



CURRENT SOLUTION AND ITS LIMITATIONS

RGW Multisite-Based Cloud Sync



CURRENT SOLUTION AND ITS LIMITATIONS

Limitations of RGW Cloud Sync

- Sync Granularity (zone / bucket level -> **object level**)
- Schedule Flexibility (immediately -> **time-based**)

OUR SOLUTION PART I – STORAGE CLASS

- Why Classification ?
- RGW Data Placement
- Object Storage Class

OUR SOLUTION PART I – STORAGE CLASS

Why Classification

Storage Medium

- SSD
- HDD
- Blue Ray

Storage Policy

- 3x Replication
- 2x Replication
- Erasure Code

Storage Vendor

- UCloud
- AWS

OUR SOLUTION PART I – STORAGE CLASS

RGW Data Placement

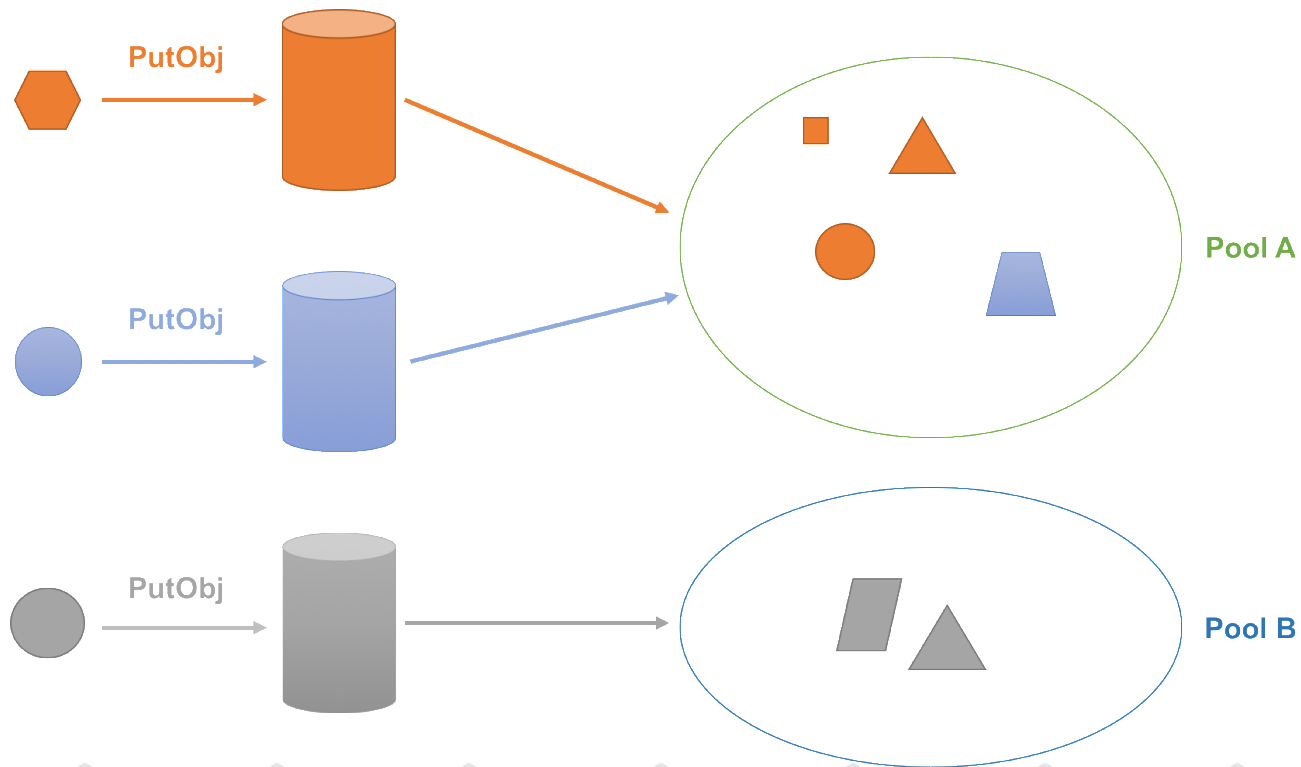
```
},  
"placement_pools": [  
  {  
    "key": "default-placement",  
    "val": {  
      "index_pool": "default.rgw.buckets.index",  
      "data_pool": "default.rgw.buckets.data",  
      "data_extra_pool": "default.rgw.buckets.non-ec",  
      "index_type": 0,  
      "compression": ""  
    }  
  }  
],  
"metadata_heap": "",  
"tier_config": [],  
"realm_id": ""  
}
```

Placement Rule

- index pool
- data pool
- data extra pool

A field in zone configuration

OUR SOLUTION PART I – STORAGE CLASS



OUR SOLUTION PART I – STORAGE CLASS

Object Storage Class

AWS S3 Storage Class

- STANDARD
- STANDARD_IA
- ONEZONE_IA
- REDUCED_REDUNDANCY
- GLACIER

OUR SOLUTION PART I – STORAGE CLASS

Object Storage Class

Storage
Class



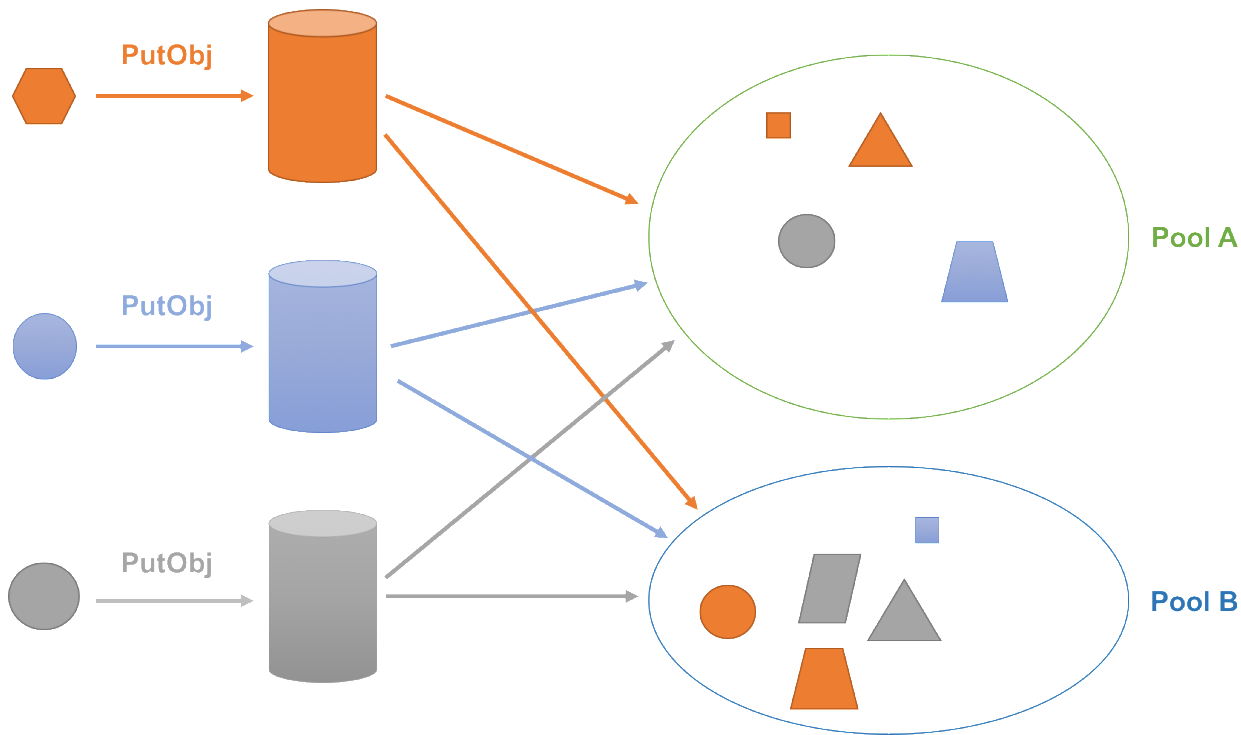
Pool

- Storage Medium e.g. SSD/HDD
- Storage Policy e.g. 2x/3x/EC
- **Storage Vendor** e.g. **AWS S3 / UCloud UFile**

Placement Rule

- Bucket
- **Object**

OUR SOLUTION PART I – STORAGE CLASS



OUR SOLUTION PART II – LC TRANSITION

Object Lifecycle Management

AWS S3

- Expiration actions
- Transition actions

Ceph RGW

- Expiration actions
- **Transition actions**

OUR SOLUTION PART II – LC TRANSITION

Object Lifecycle Management

LC Transition **VS** Cloud Sync

- Sync Granularity (zone / bucket level/**object level**)
- Schedule Flexibility (**time-based**)

OUR SOLUTION PART III – AUTO-GENERATED TRANSITION POLICY

- Gathering Metrics - Bucket Logging
- Auto-Generated Transition Policy

OUR SOLUTION PART III – AUTO-GENERATED TRANSITION POLICY

Gathering Metrics - Bucket Logging

```

79a59df900b949e55d96a1e698fbacedfd6e09d98eac8f8d5218e7cd47ef2be mybucket [06/Feb/2014:00:00:38 +0000] 192.0.2.3 79a59df900b949e55d96a1e698fbacedfd6e09d98eac8f8d5218e7cd47ef2be 3E57427F3EXAMPLE REST.GET.
VERSIONING - "GET /mybucket?versioning HTTP/1.1" 200 - 113 - 7 - "-" "S3Console/0.4" -

79a59df900b949e55d96a1e698fbacedfd6e09d98eac8f8d5218e7cd47ef2be mybucket [06/Feb/2014:00:00:38 +0000] 192.0.2.3 79a59df900b949e55d96a1e698fbacedfd6e09d98eac8f8d5218e7cd47ef2be 891CE47D2EXAMPLE REST.GET.
LOGGING_STATUS - "GET /mybucket?logging HTTP/1.1" 200 - 242 - 11 - "-" "S3Console/0.4" -

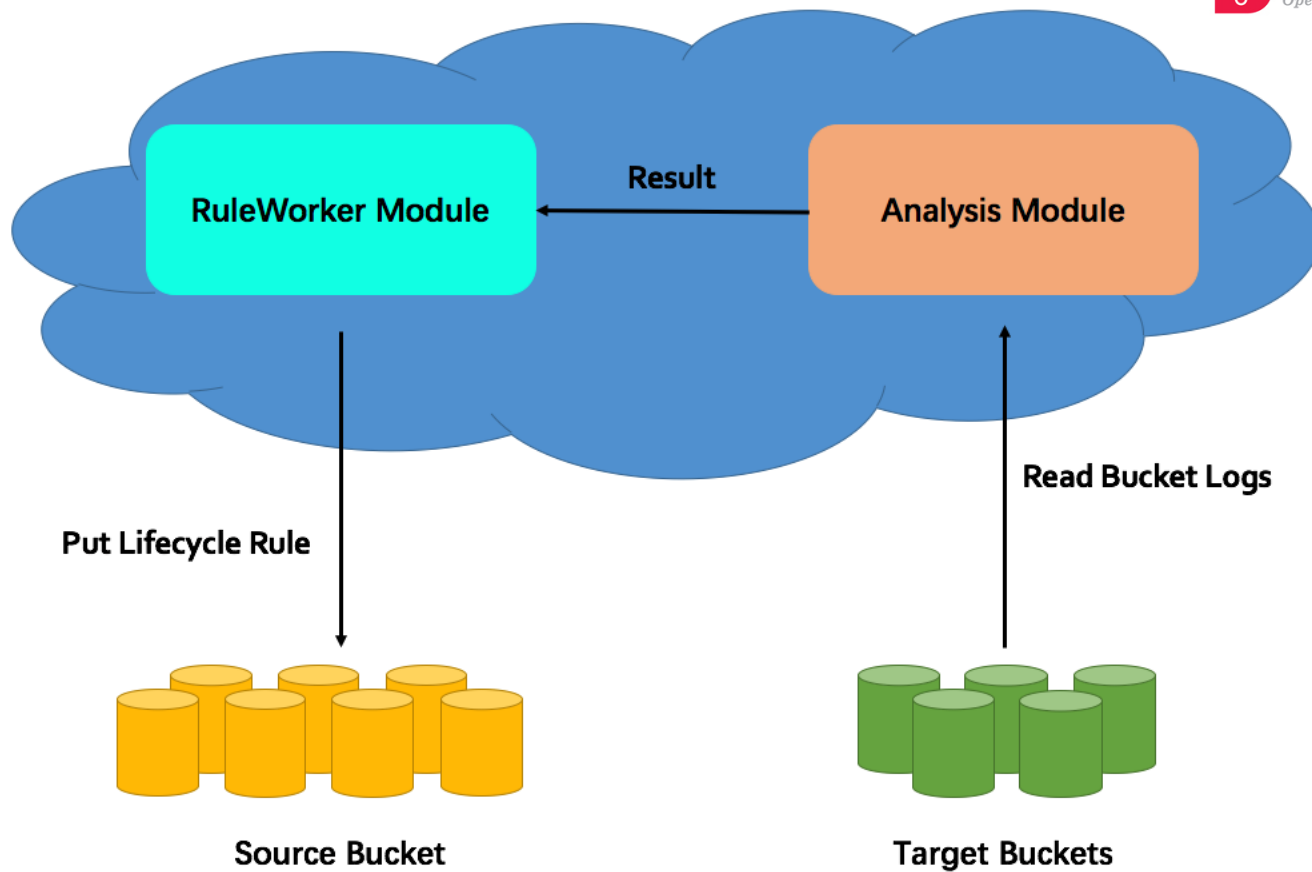
79a59df900b949e55d96a1e698fbacedfd6e09d98eac8f8d5218e7cd47ef2be mybucket [06/Feb/2014:00:00:38 +0000] 192.0.2.3 79a59df900b949e55d96a1e698fbacedfd6e09d98eac8f8d5218e7cd47ef2be A1206F460EXAMPLE REST.GET.
BUCKETPOLICY - "GET /mybucket?policy HTTP/1.1" 404 NoSuchBucketPolicy 297 - 38 - "-" "S3Console/0.4" -

79a59df900b949e55d96a1e698fbacedfd6e09d98eac8f8d5218e7cd47ef2be mybucket [06/Feb/2014:00:01:00 +0000] 192.0.2.3 79a59df900b949e55d96a1e698fbacedfd6e09d98eac8f8d5218e7cd47ef2be 7B4A0FABBEXAMPLE REST.GET.
VERSIONING - "GET /mybucket?versioning HTTP/1.1" 200 - 113 - 33 - "-" "S3Console/0.4" -

79a59df900b949e55d96a1e698fbacedfd6e09d98eac8f8d5218e7cd47ef2be mybucket [06/Feb/2014:00:01:57 +0000] 192.0.2.3 79a59df900b949e55d96a1e698fbacedfd6e09d98eac8f8d5218e7cd47ef2be DD6CC733AEXAMPLE REST.PUT.OBJECT
s3-dg.pdf "PUT /mybucket/s3-dg.pdf HTTP/1.1" 200 - - 4406583 41754 28 "-" "S3Console/0.4" -

79a59df900b949e55d96a1e698fbacedfd6e09d98eac8f8d5218e7cd47ef2be mybucket [06/Feb/2014:00:03:21 +0000] 192.0.2.3 79a59df900b949e55d96a1e698fbacedfd6e09d98eac8f8d5218e7cd47ef2be BC3C074D0EXAMPLE REST.GET.
VERSIONING - "GET /mybucket?versioning HTTP/1.1" 200 - 113 - 28 - "-" "S3Console/0.4" -
  
```

Bias For Data Temperature Analysis



FURTHER WORK

- Support bi-direction sync
- Support proxy read and write