DATA SATURDAYS #64 PARMA 2024

Non sottovalutare la potenza dello Storage Account!!

Massimo Bonanni

Technical Trainer @ Microsoft

CLABS

Sponsor & Org

iSolutions

engage LABS

DATA SKILLS
UNDERSTANDING THE WORLD

LIME ware

daniela malvisi COMMUNICATION

Lucient ITALIA

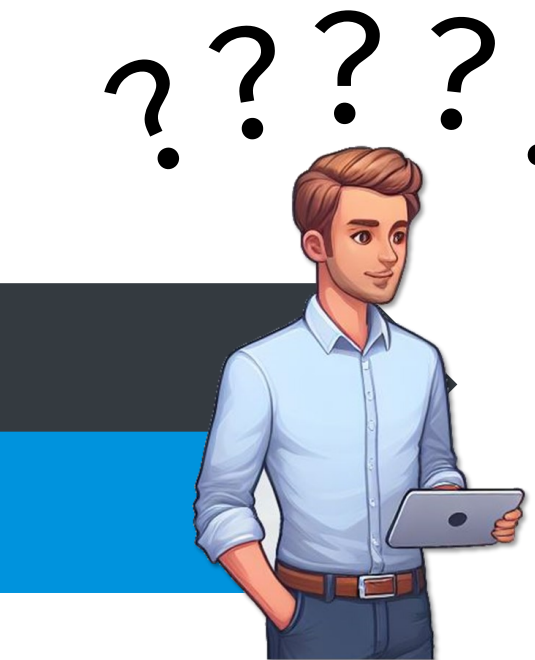
Bi Factory
DATA KNOWLEDGE ADVISOR

A day like many others, in a company like many others...

The requirements

- ✓ We produce an average of 100 Mb of new contents every hour in two different locations
- ✓ Contents are frequently accessed in the first 10 days, very rarely after a month, and can be deleted after 1 year
- ✓ The contents cannot be deleted for 10 days after publication
- ✓ The maximum downtime in a single month must be under 5 minutes
- ✓ The costs must be kept to a minimum
- ✓ We need daily statistics on the number of contents available, and on the space occupied

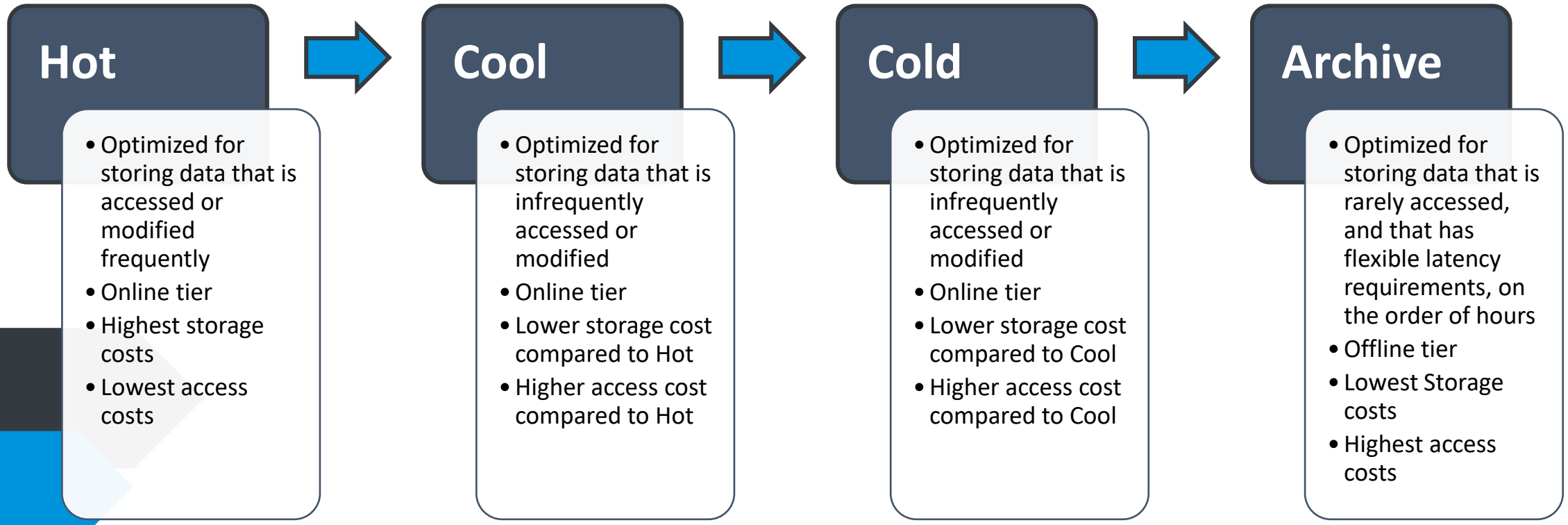
What if we used the Storage Account?



How do we minimize costs?

Access Tiers

Azure storage offers different access tiers so that you can store your blob data in the most cost-effective manner based on how it's being used.



Access Tiers

Contents will be stored in the **Hot** tier for the first **10 days**, then will be "moved" to the **Cool** tier for the next **30 days**, and then will be “moved” to the **Cold** tier. We can finally move them after 1 year to the **Archive** tier (or delete).

But... who “moves” the contents?

Blob Lifecycle Management

Azure Storage lifecycle management is a **rule-based policy** that allows you to “move” blob data to the appropriate access tiers or to expire data at the end of the data lifecycle.

A lifecycle management policy is comprised of one or more rules that define a set of actions to take based on a condition being met.

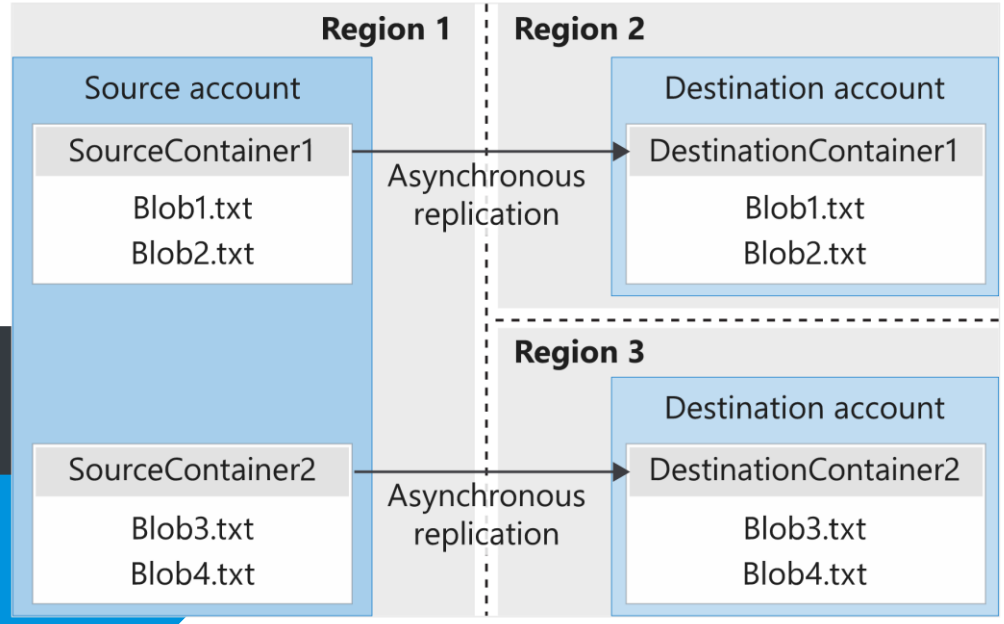
Blob Lifecycle Management

Let's create a rule that moves contents to the **Cool** tier **10 days** after their creation and moves them to the **Cold** tier after **30 days**...and deletes them after **1 year**.

But…we have **multiple storages in multiple regions** and all our users must see the same contents in all the regions.

Object Replication

Object replication asynchronously copies block blobs between a source storage account and a destination account.



- Minimizing latency
- Optimizing data distribution
- Optimizing costs

Object Replication

Contents produced in one region will be **replicated to the other regions**, so the web apps in each region will read all contents **locally.**

Cool…but…we need also to create a software that analyzes, **every day**, the contents in each storage to save **statistics**.

Storage Inventory

Blob inventory feature provides an overview of your containers, blobs, snapshots, and blob versions within a storage account managed by Azure.

Inventory reports for blobs and containers

Custom Schema

CSV and Apache Parquet output format

Manifest file and Azure Event Grid event per inventory report

Storage Inventory

Blob inventory feature provides an overview of your containers, blobs, snapshots, and blob versions within a storage account managed by Az

Using inventory, **every day**, we will have a **CSV file** with all the info we need to create our statistics.

We will use **Event Grid and Azure Functions** to extract data from Inventory file when it will be generated by Azure.

Manifest file and Azure Event Grid event per inventory re

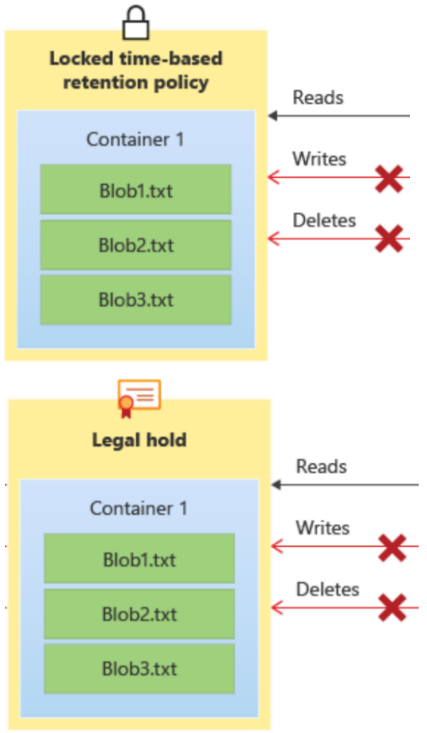
…and what about **avoid** contents **deletion** in the first **10 days** after their creation?

Immutable Storage

Immutable storage for Azure Blob Storage enables users to store business-critical data in a **WORM** (**Write Once, Read Many**) state.

While in a WORM state, data cannot be modified or deleted for a user-specified interval.

- Time-based retention policies
- Legal hold policies

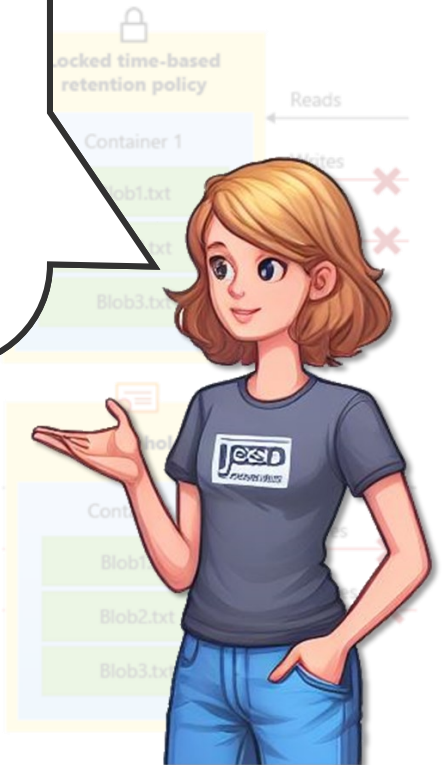


Immutable Storage

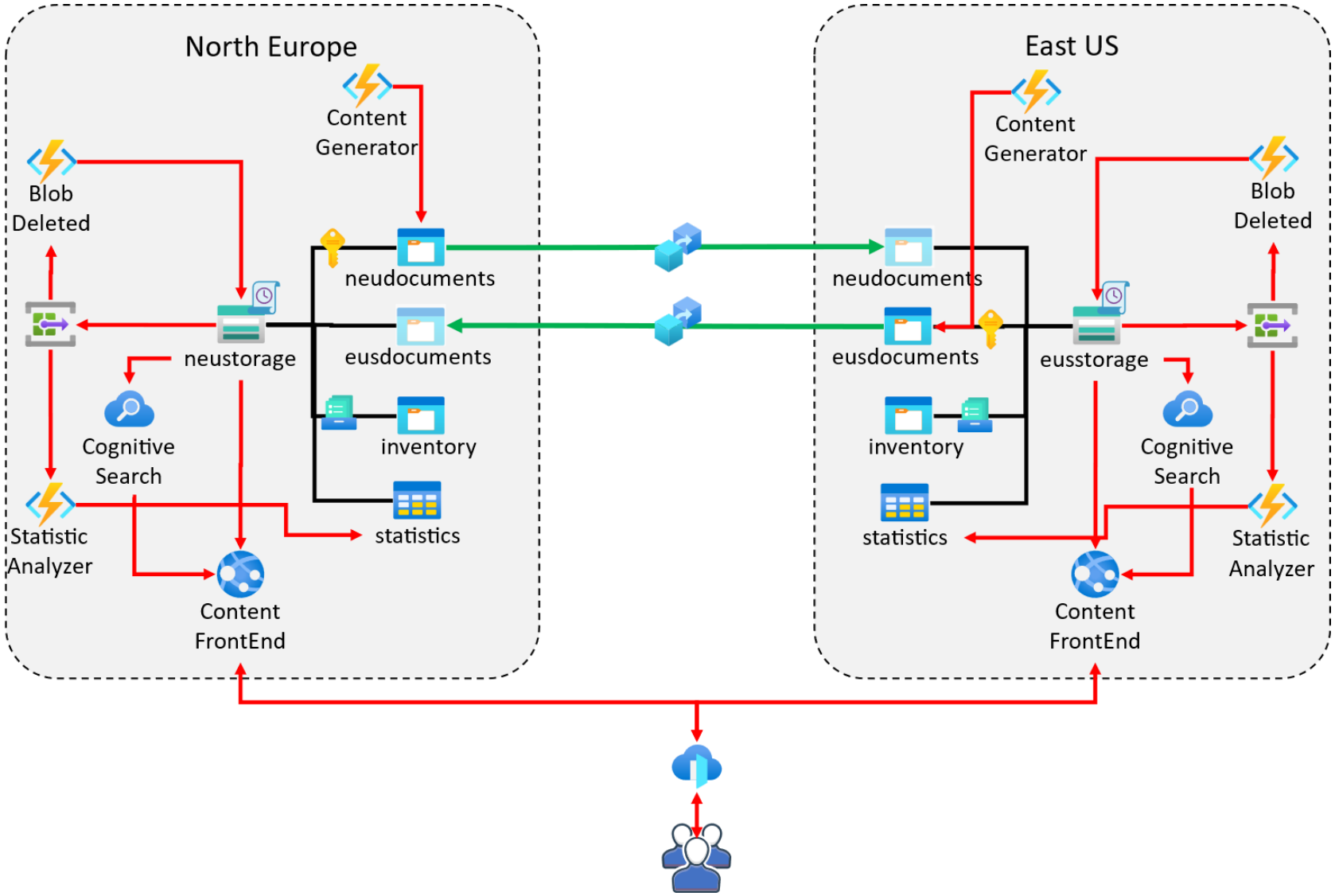
Immutable storage for Azure Blob Storage enables users to store business-critical data

We will set a **time-based policy** to avoid deletion in the first **10 days** after content creation.

Legal hold policies



The solution



DEMO

Storage Content Platform

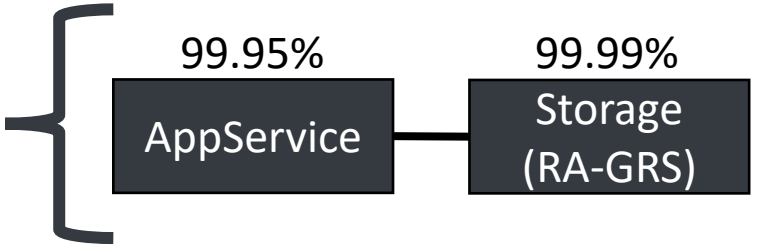
massimobonanni/StorageContentPlatform

…are you sure we can provide at most 5 minutes of downtime every month?

The SLA of the solution

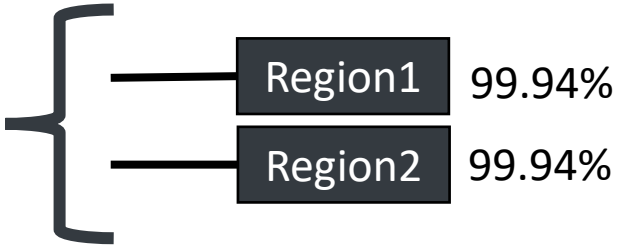
5 minutes of downtime every month → **SLA 99.9885 %**

SLA 1 region



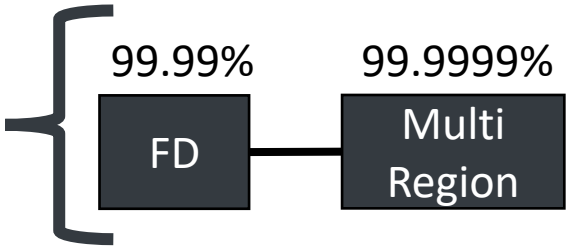
$$SLA_{sr} = 99.95\% * 99.99\% = 99.94\%$$

SLA 2 regions



$$SLA_{mr} = 100\% - (0.06\% * 0.06\%) \cong 99.9999\%$$

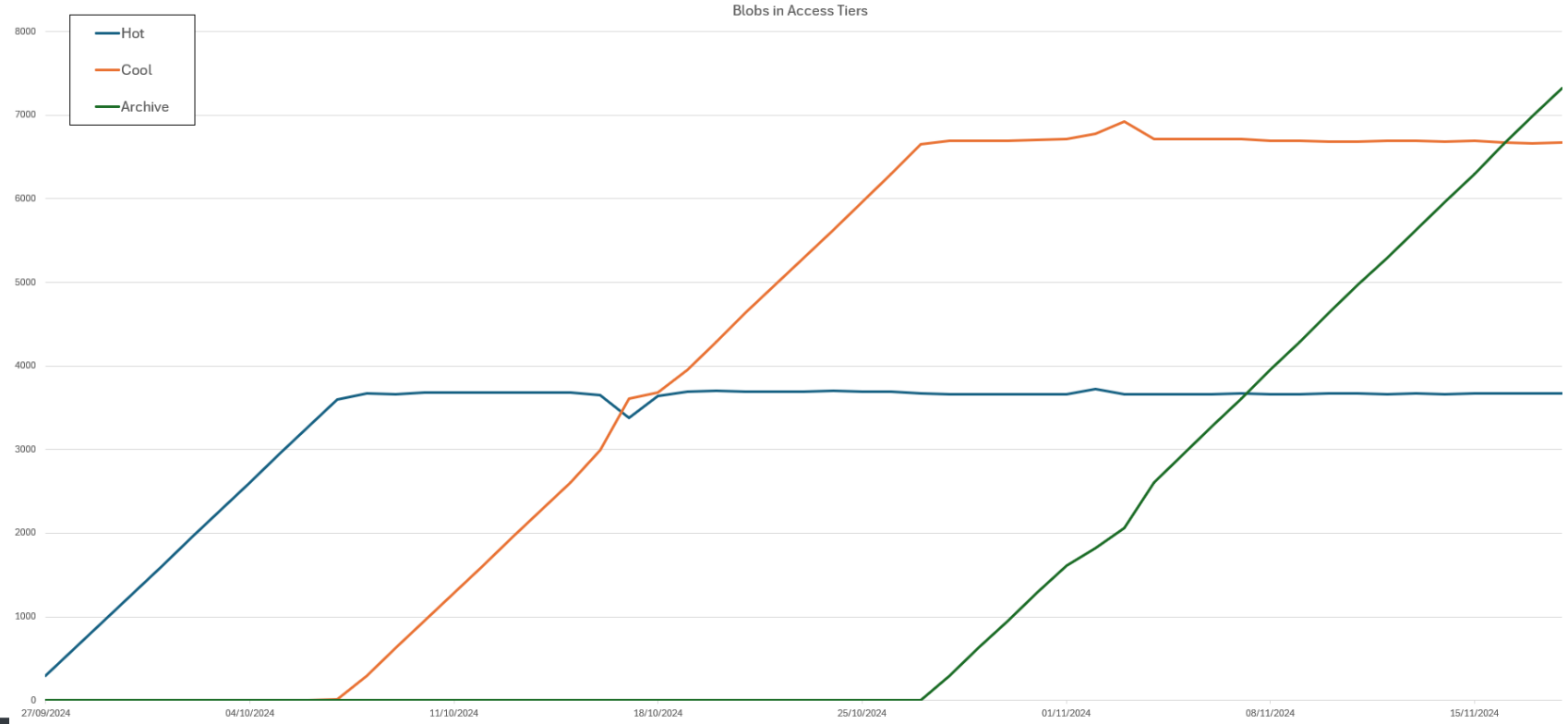
SLA full solution

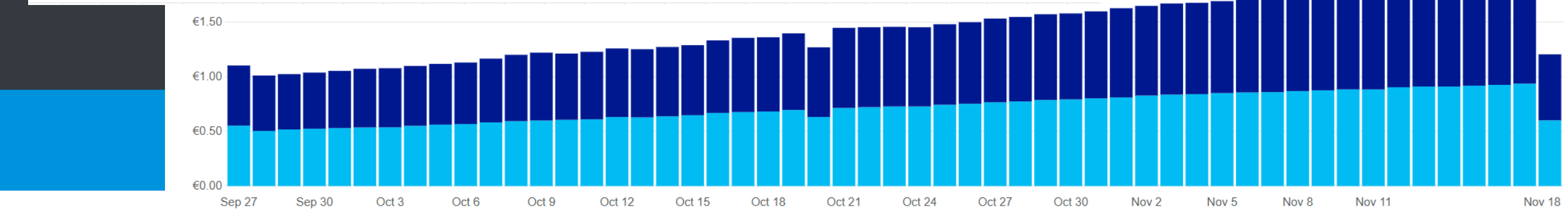


$$SLA_{tot} = 99.99\% * SLA_{mr} \cong \mathbf{99.9899\%}$$

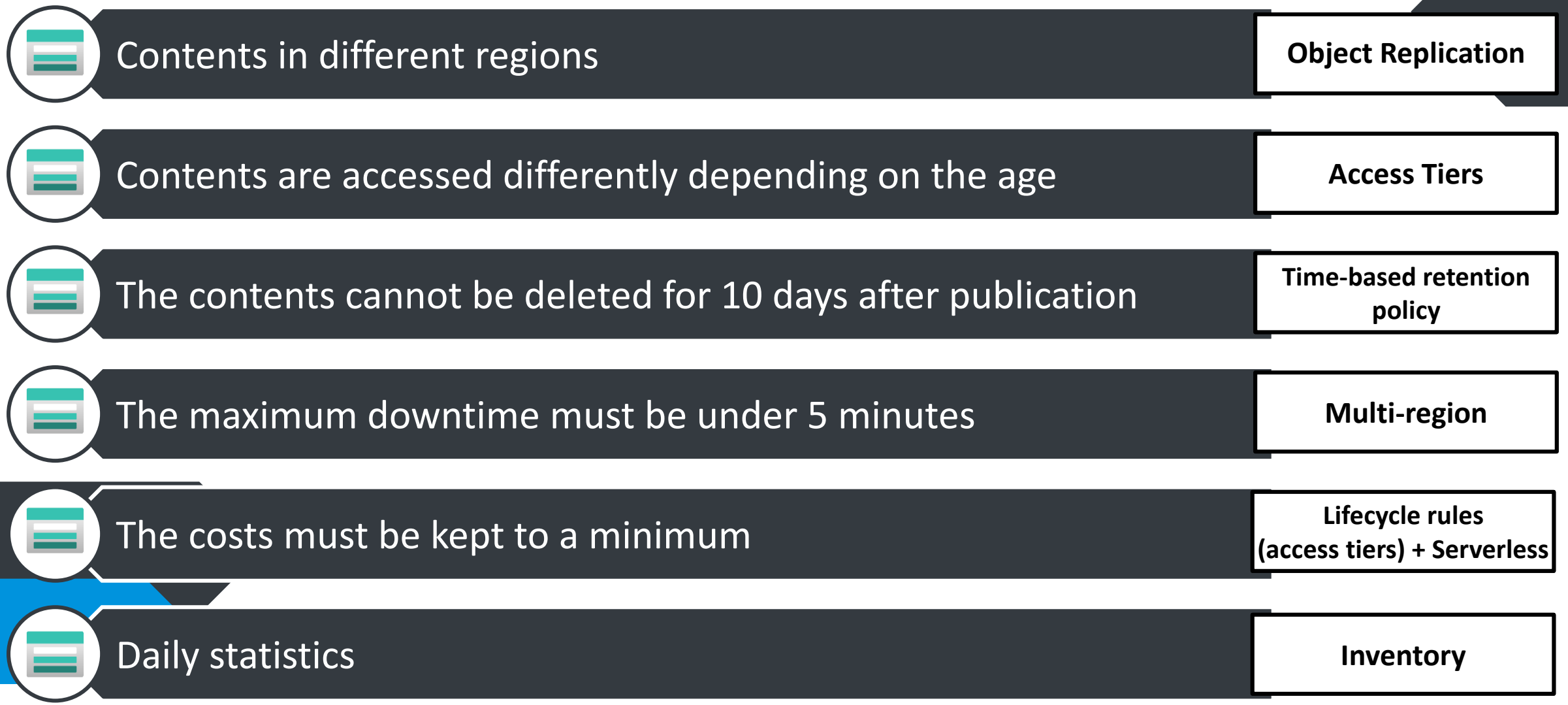
…and how will cost it?

Costs

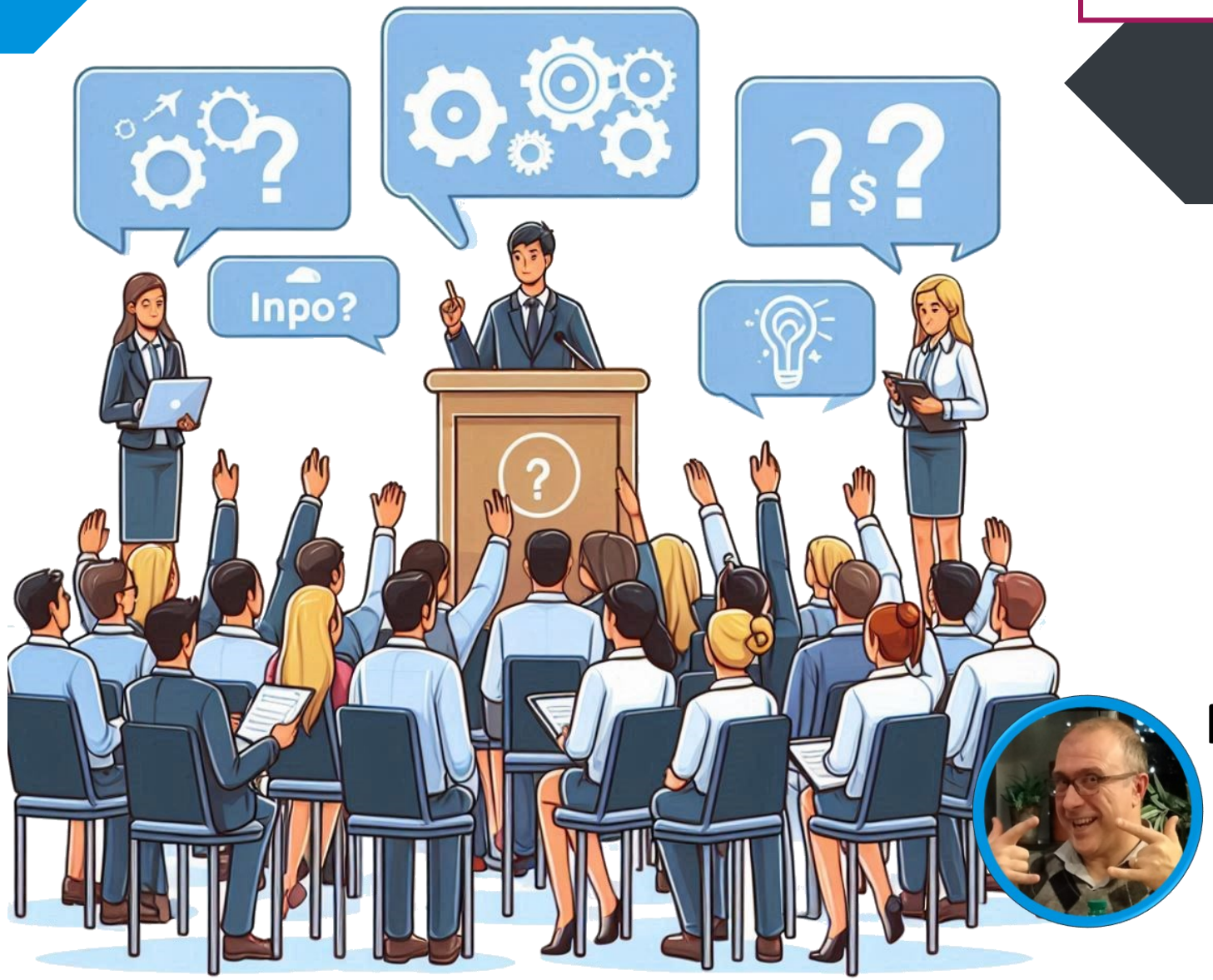

Blobs in Access Tiers
Hot
Cool
Archive
8000
7000
6000
5000
4000
3000
2000
1000
0
27/09/2024
04/10/2024
11/10/2024
18/10/2024
25/10/2024
01/11/2024
08/11/2024
15/11/2024


€1.50
€1.00
€0.50
€0.00
Sep 27
Sep 30
Oct 3
Oct 6
Oct 9
Oct 12
Oct 15
Oct 18
Oct 21
Oct 24
Oct 27
Oct 30
Nov 2
Nov 5
Nov 8
Nov 11
Nov 18

Feature-to-Requirement mapping

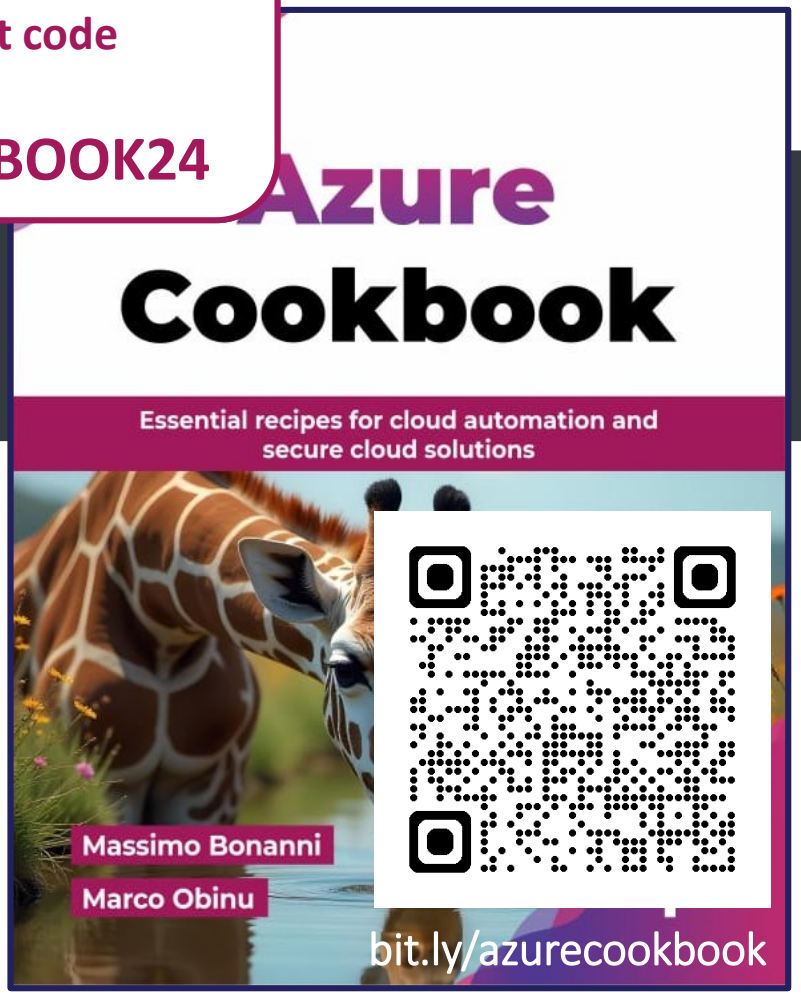


Why don’t we use Storage Account?



20% discount code

AZURECOOKBOOK24



Massimo Bonanni

Technical Trainer @ Microsoft

massimo.bonanni@microsoft.com
@massimobonanni

aka.ms/maxlinkedin

Thanks!!!