USE CLOUDSUMMIT APP!

SESSION FEEDBACK
BUSINESS NETWORKING
AGENDA
SPEAKERS AND SESSIONS

https://csmmt.eu/app

DID YOU KNOW?

YOU CAN USE THE APP TO SCAN OTHER ATTENDEES' BADGES, AND THEY WILL BECOME YOUR CONNECTIONS!

Microsoft

daenet – an ACP Digital company

Azure Best Practices

Architecting Complex Solutions with Heavy Loads

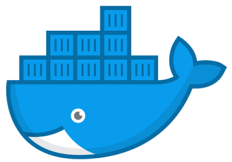
Damir Dobric

Agenda

- Running jobs with large RAM consumption
- Hosting Large Objects in Web
- Q&A

https://github.com/ddobric/azure-bestpractices

Services used in this talk



docker

acr

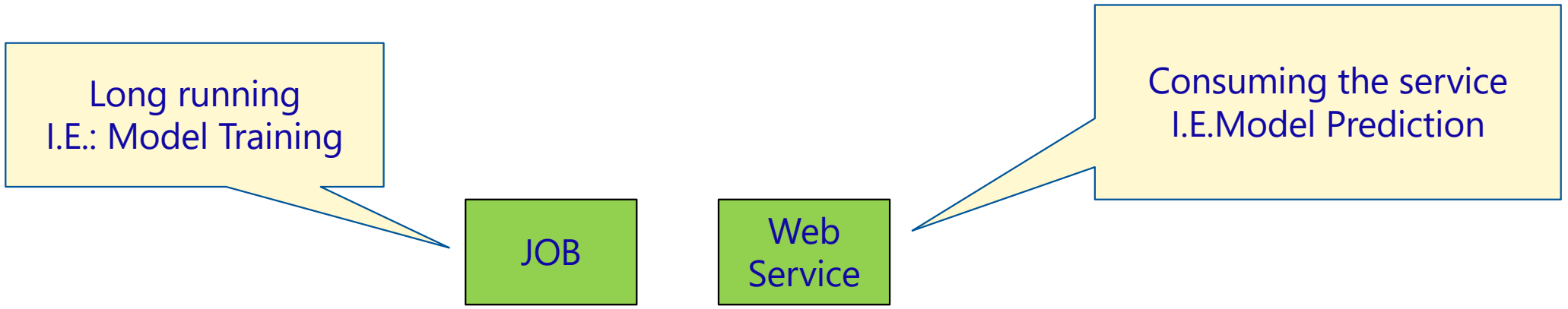
batch

appervice

aci

az

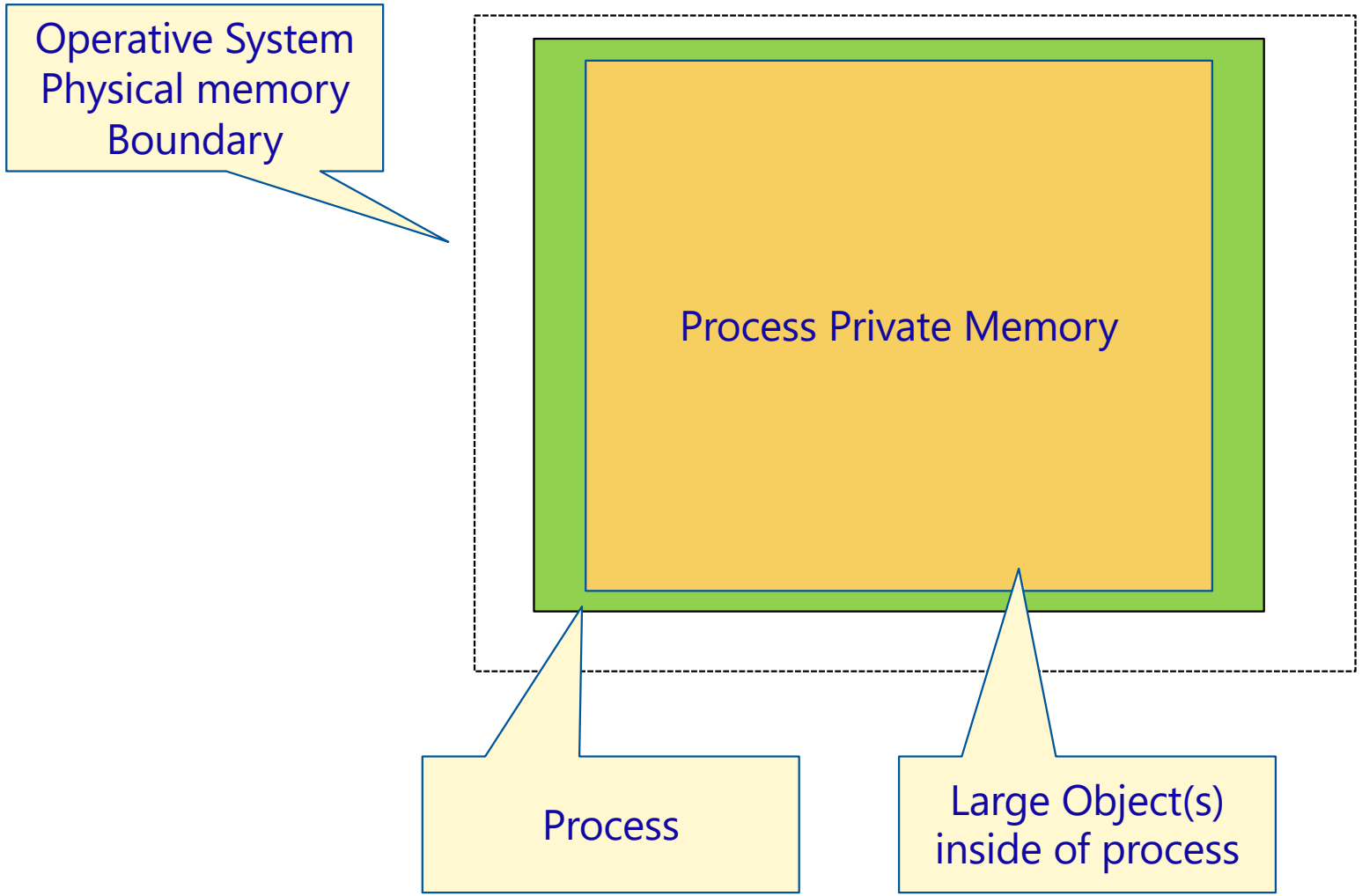
Typical ML Project



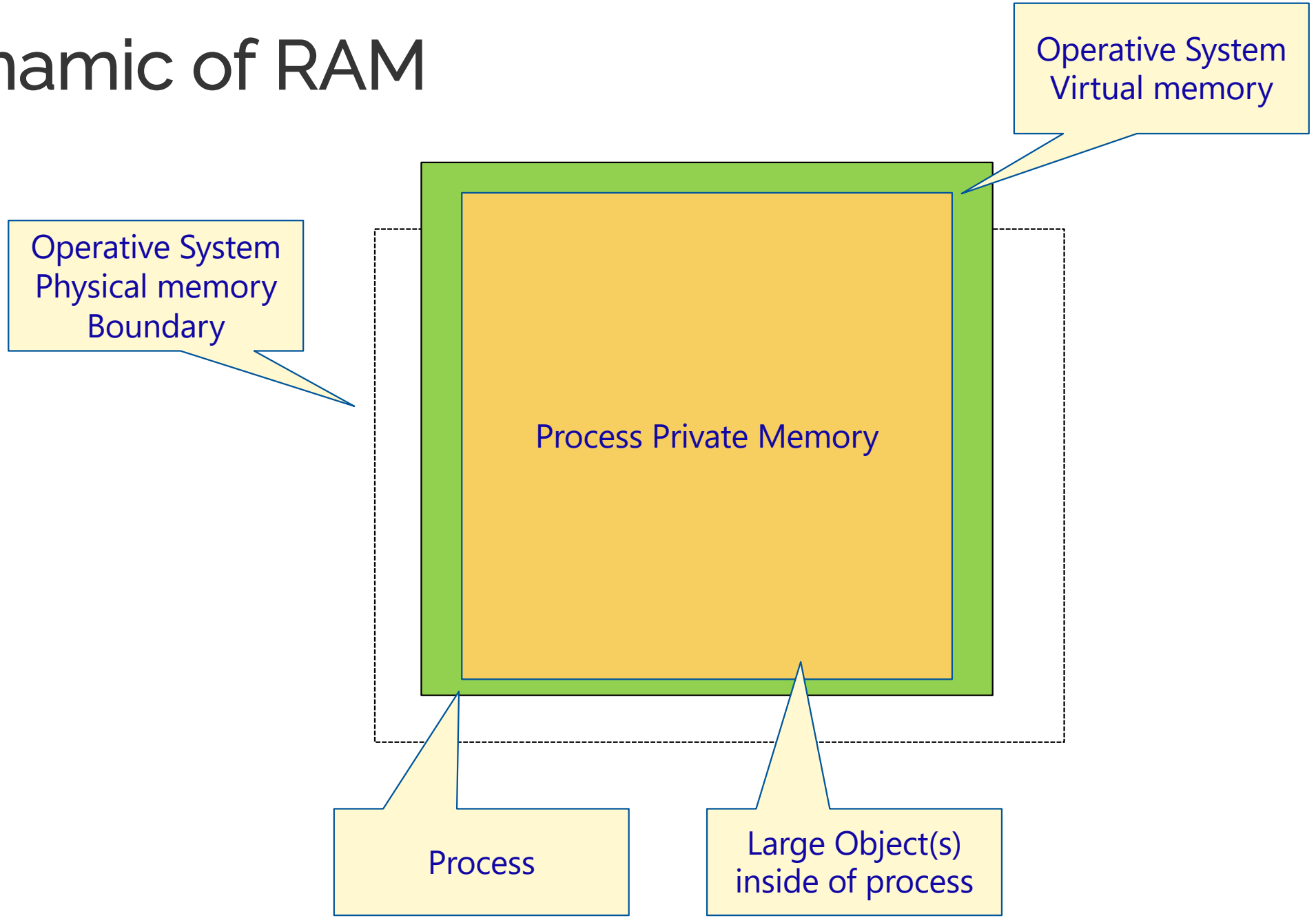
THE SECRET OF A DOCKER HYPE

RUNNING OBJECTS WITH A LARGE RAM CONSUMPTION

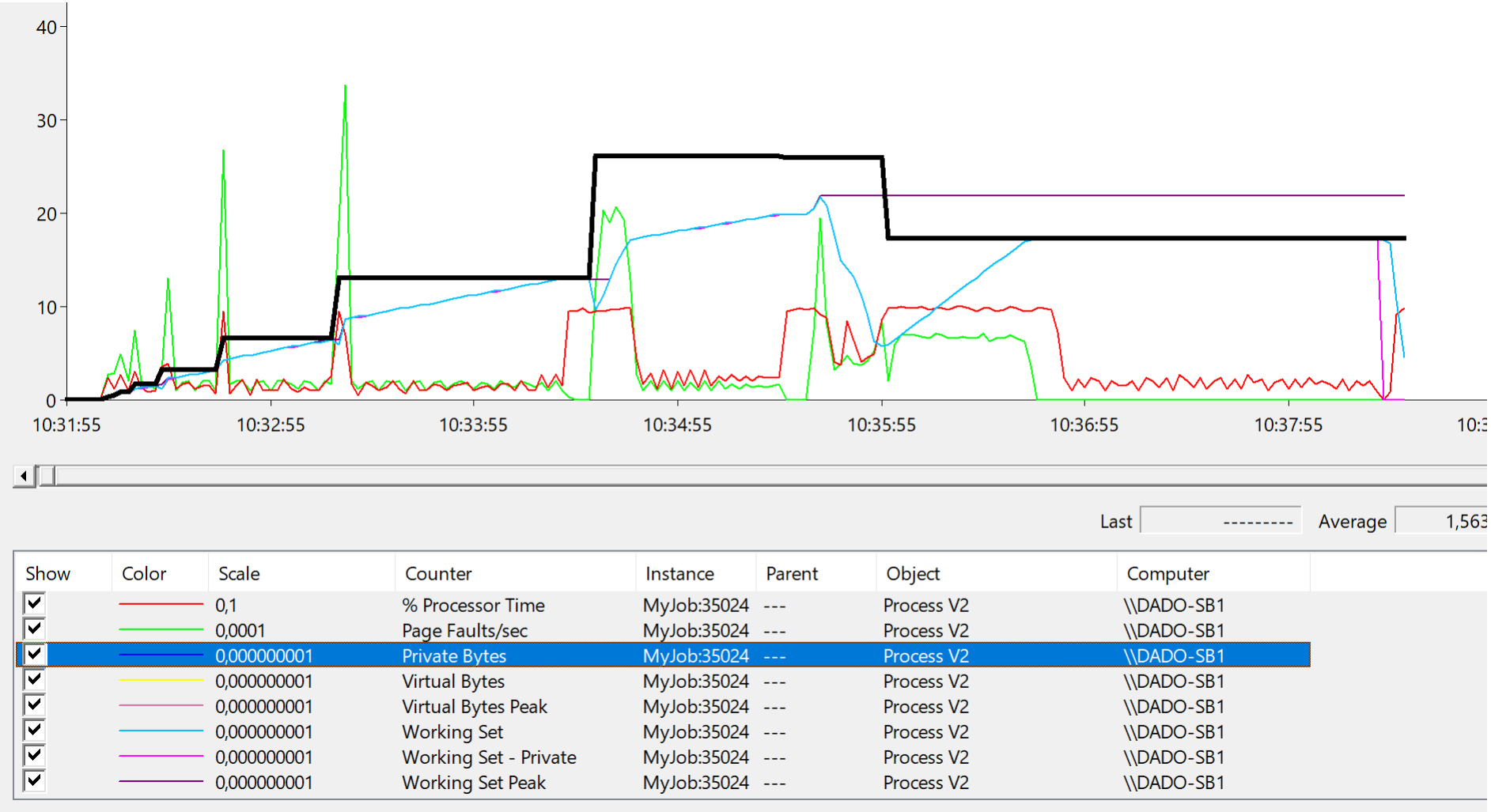
Dynamic of RAM



Dynamic of RAM



Run job locally



Show	Color	Scale	Counter	Instance	Parent	Object	Computer
☑		0,1	% Processor Time	MyJob:35024	---	Process V2	\\DADO-SB1
☑		0,0001	Page Faults/sec	MyJob:35024	---	Process V2	\\DADO-SB1
☑		0,000000001	Private Bytes	MyJob:35024	---	Process V2	\\DADO-SB1
☑		0,000000001	Virtual Bytes	MyJob:35024	---	Process V2	\\DADO-SB1
☑		0,000000001	Virtual Bytes Peak	MyJob:35024	---	Process V2	\\DADO-SB1
☑		0,000000001	Working Set	MyJob:35024	---	Process V2	\\DADO-SB1
☑		0,000000001	Working Set - Private	MyJob:35024	---	Process V2	\\DADO-SB1
☑		0,000000001	Working Set Peak	MyJob:35024	---	Process V2	\\DADO-SB1

Run a job as a Docker Container

```
> docker run -it --rm -m 8g myjob:latest
```

Run locally

Tag image before pushing to ACR

```
> docker tag myjob:latest damir.azurecr.io/myjob:latest
```

Push to ACR

```
> docker push damir.azurecr.io/myjob:latest
```

```
> az container create --resource-group 'my-container-apps' --name myjob --
image vcddevregistry.azurecr.io/myjob:latest --m 8  --cpu 4 --restart-policy
Never --os-type linux --registry-password RtMTAS+i2q1GkuYvFOIiRdRjpaxIAscF -
-registry-user vcddevregistry
```

Run in ACI

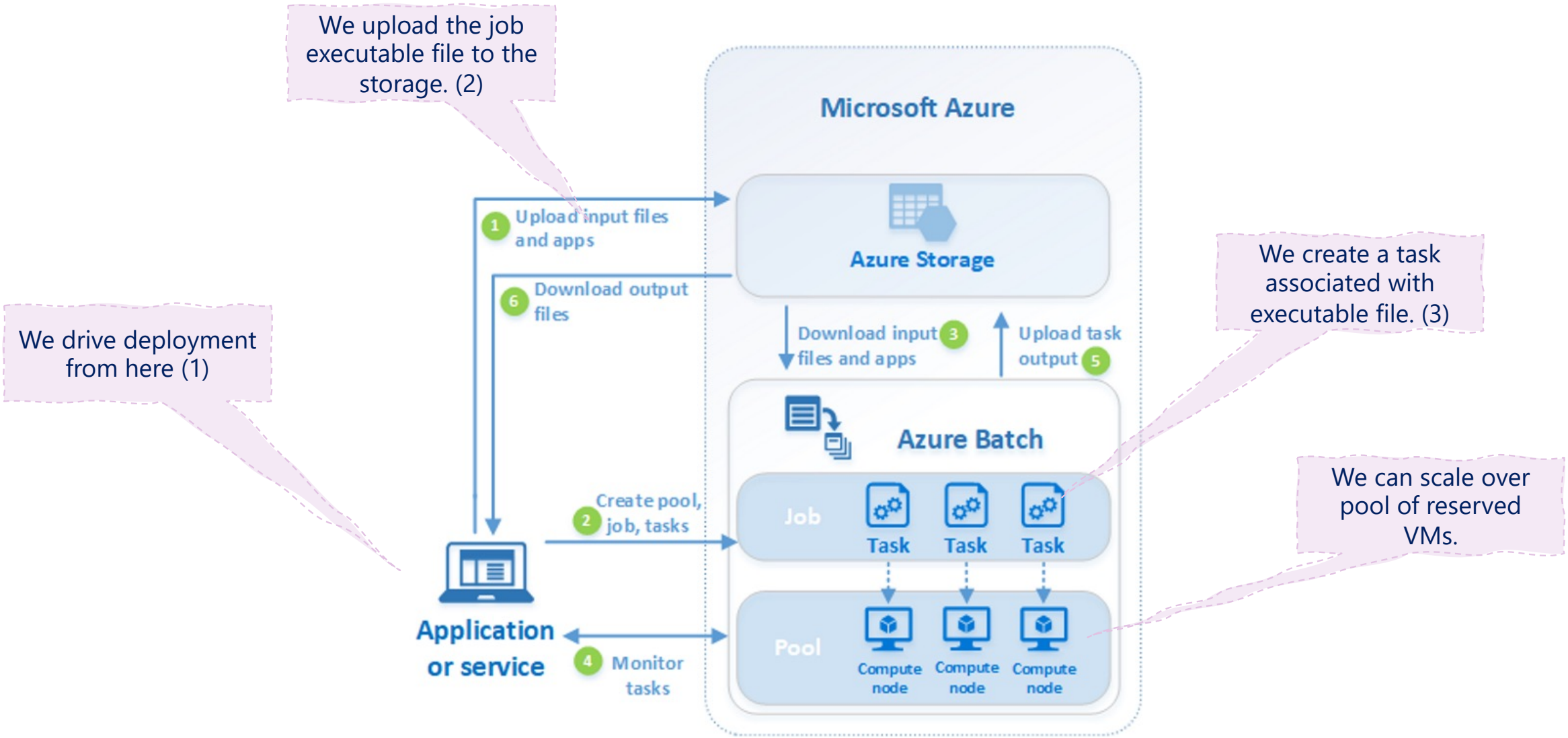
Killing my Docker softly

- Exit Code 134
  - Abnormal termination (SIGABRT)
  - Container Abnormally Terminated itself
  - Status *Out of Memory*
- Exit Code 137
  - Immediate Termination (SIGKILL)
  - Docker Kill
  - Status *OOMKilled*

Inspect the container

```
PS /home/damir> az container show -g RG-docker --name myservice4
{
  "containers": [
    {
      "command": null,
      "environmentVariables": [],
      "image": "cloudlesson.azurecr.io/myjob:latest",
      "instanceView": {
        "currentState": {
          "detailStatus": "Error",
          "exitCode": 134,
          "finishTime": "2022-05-02T08:41:13.060000+00:00",
          "startTime": "2022-05-02T08:41:01.809000+00:00",
          "state": "Terminated"
        },
```

Azure Batch Service



Docker vs. AzBatch with --memory 8G

Docker in ACI and local

```
24
PrivateMemorySize64:     4.44
25
PrivateMemorySize64:     4.44
26
Out of memory.
2022-07-02T13:38:09.9683023Z stdout F
```

stdout.txt

training_task

Download Delete Refresh

```
PrivateMemorySize64:     6.06
49
PrivateMemorySize64:     6.06
50
PrivateMemorySize64:     6.06
51
PrivateMemorySize64:     6.06
52
PrivateMemorySize64:     6.06
53
```

Azure Batch

HIDDEN PAINS OF CLOUD

HOSTING LARGE OBJECTS IN WEB APPLICATIONS

API with large RAM consumption

```
var api = new MyApi(arg);

var res = await api.Run();
```

Locking for solution

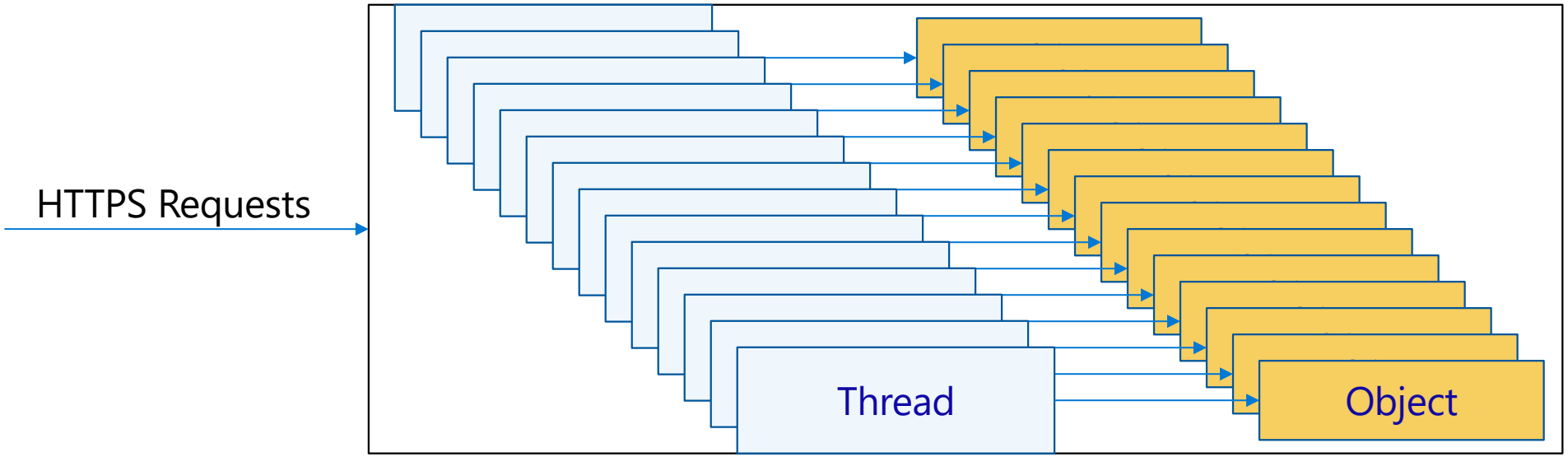
- Facts
  - Large APIs consume a lot of RAM
  - Not designed to run in the cloud
  - You want to run it as a service = REST service
  - Few API instances cannot serve many users

```
[HttpGet("{arg}")]
public async Task<string> Run(int arg)
{
    var api = new MyApi(arg);

    var res = await api.Run();

    return res;
}
```

- What to do?
  - Increase Vertically/Horizontally means an increase in costs
  - It is not easy to choose the right configuration

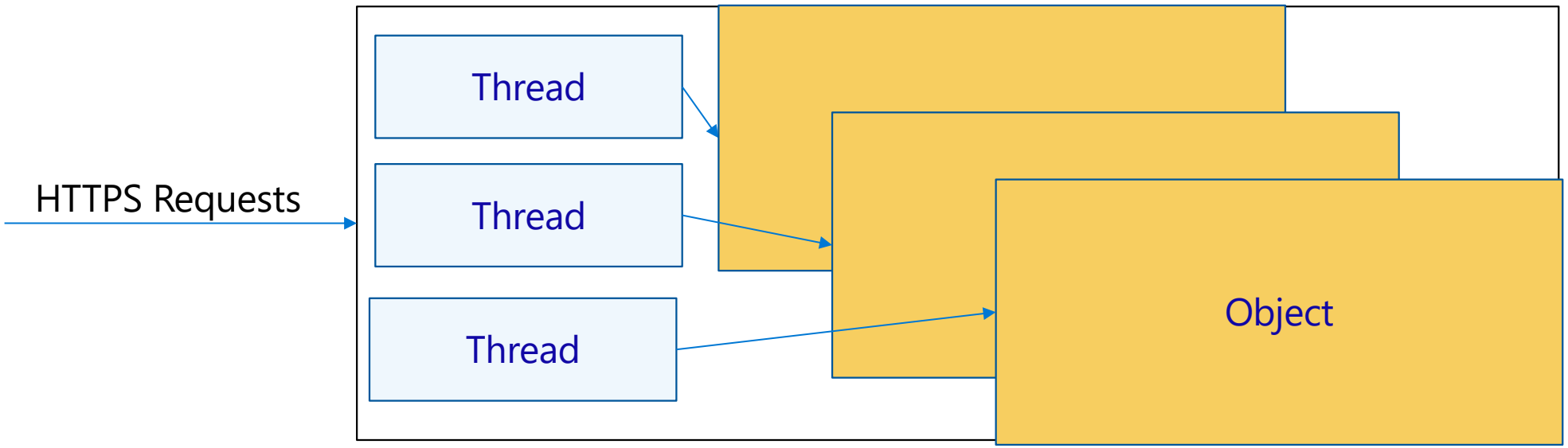
https://github.com/ddobric/azure-bestpractices

Processing of requests in ASP.NET WebApi




```
[HttpGet("{arg}")]
public async Task<string> Run(int arg)
{
    var obj = new MyApi(arg);

    var res = await obj.Run();

    return res;
}
```

Processing Large objects



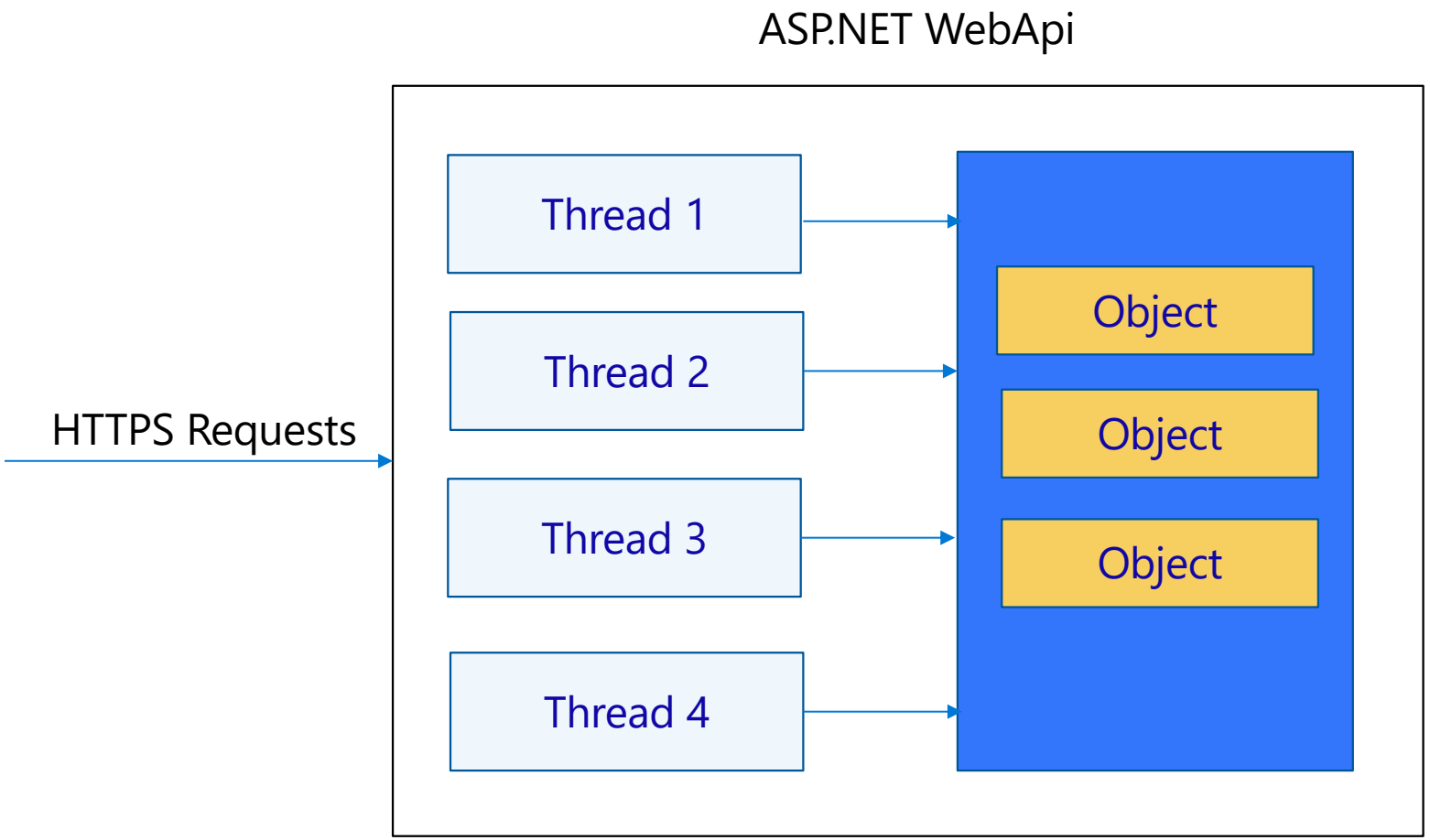
An example

- Your API consumes (just) 1 GB RAM.
- You have 100+ concurrent users.
  - 2 instances in 3.5GB
  - 100/2*144 EUR/month=7200 EUR/Month

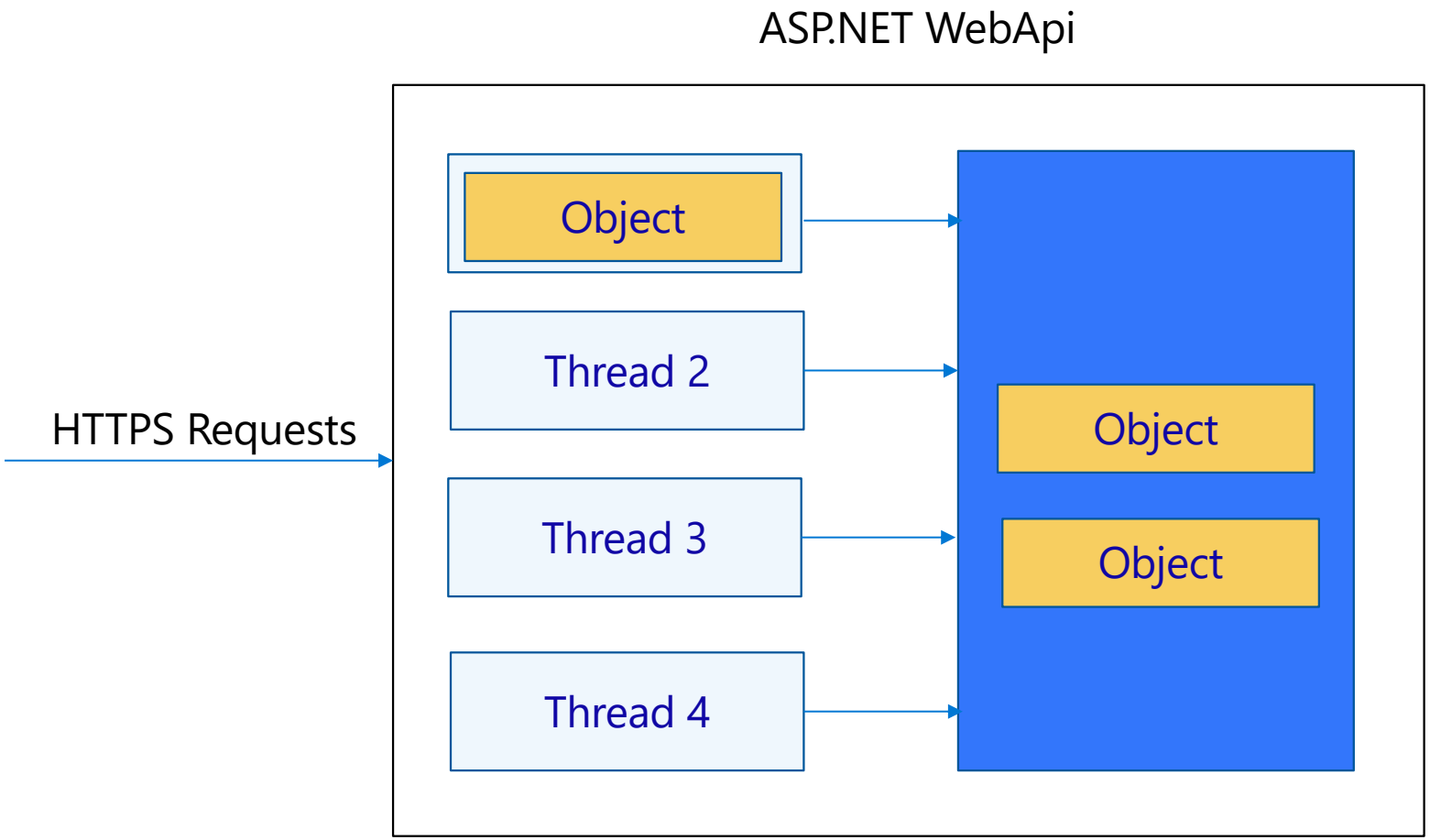
Instance	Cores	Ram	Storage	Pay as you go
S1	1	1.75 GB	50 GB	**$0.10**/hour (072)
S2	2	3.50 GB	50 GB	**$0.20**/hour (144)
S3	4	7 GB	50 GB	**$0.40**/hour (288)

THE OBJECT POOL PATTERN

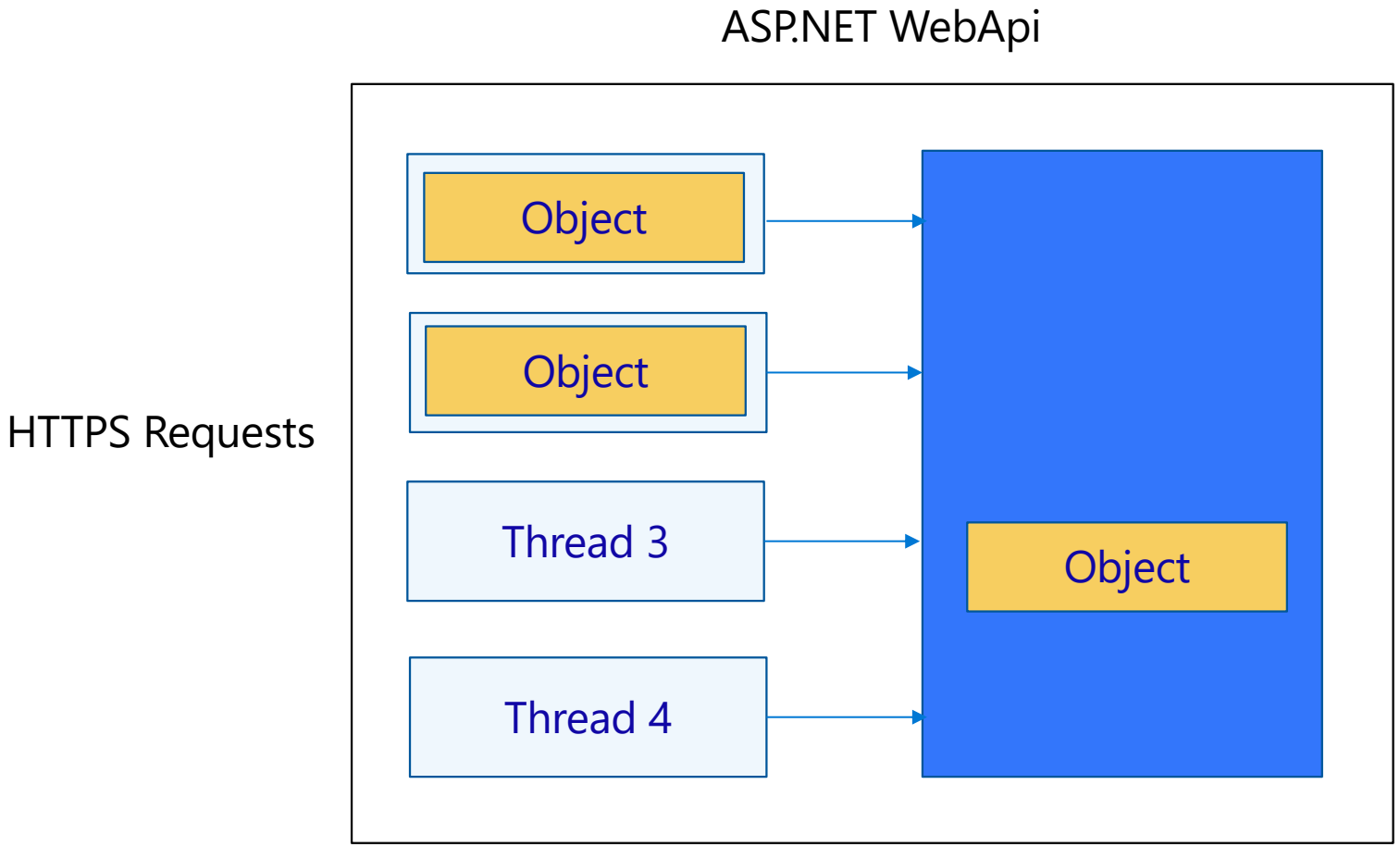
Object Pool Pattern



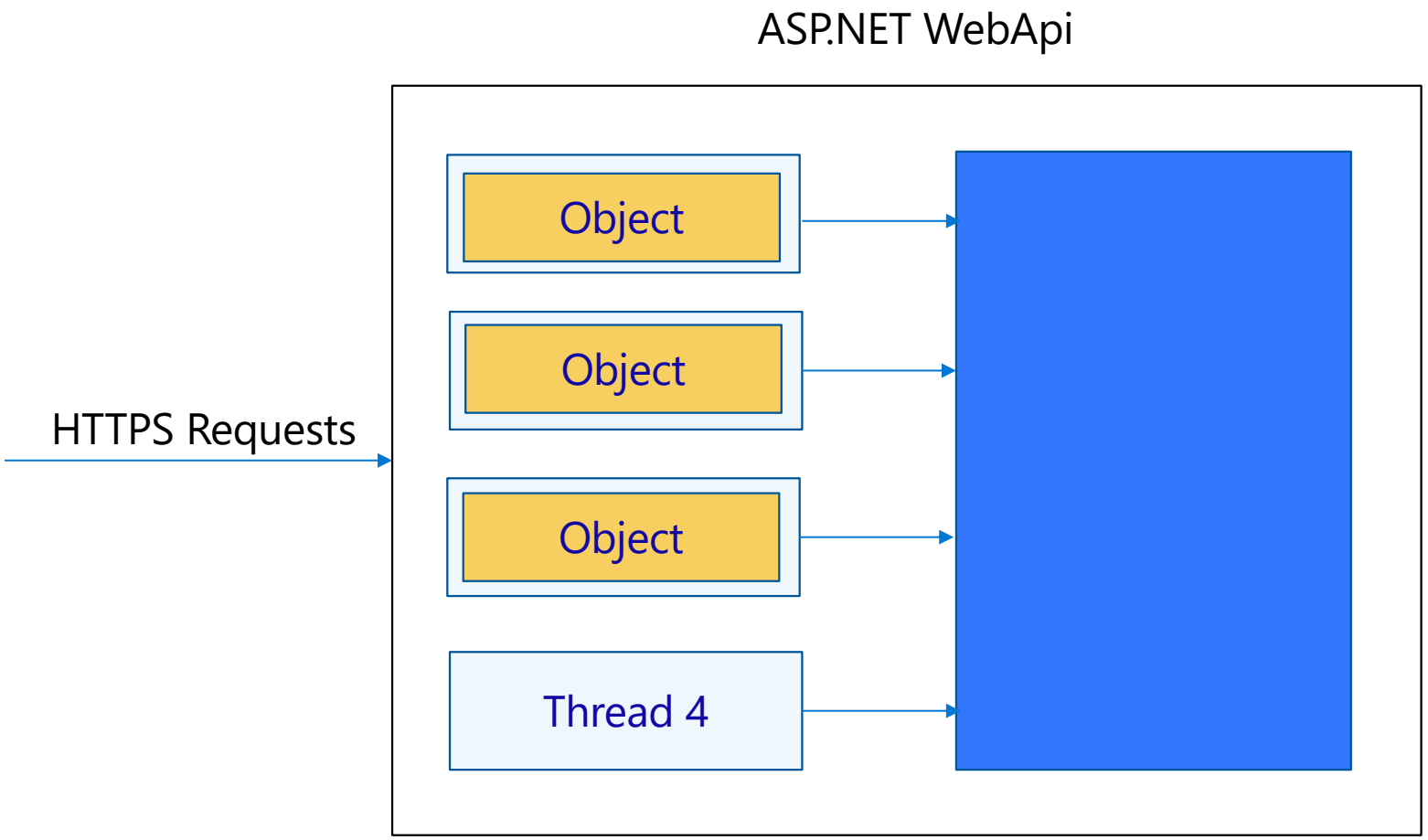
Object Pool Pattern



Object Pool Pattern

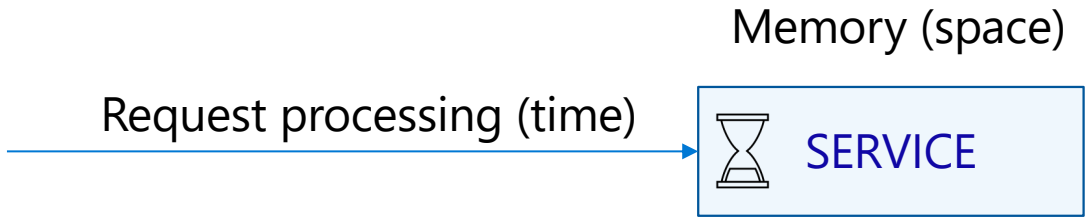


Object Pool Pattern



RELATIVITY THEORY OF THE LOAD BALANCING

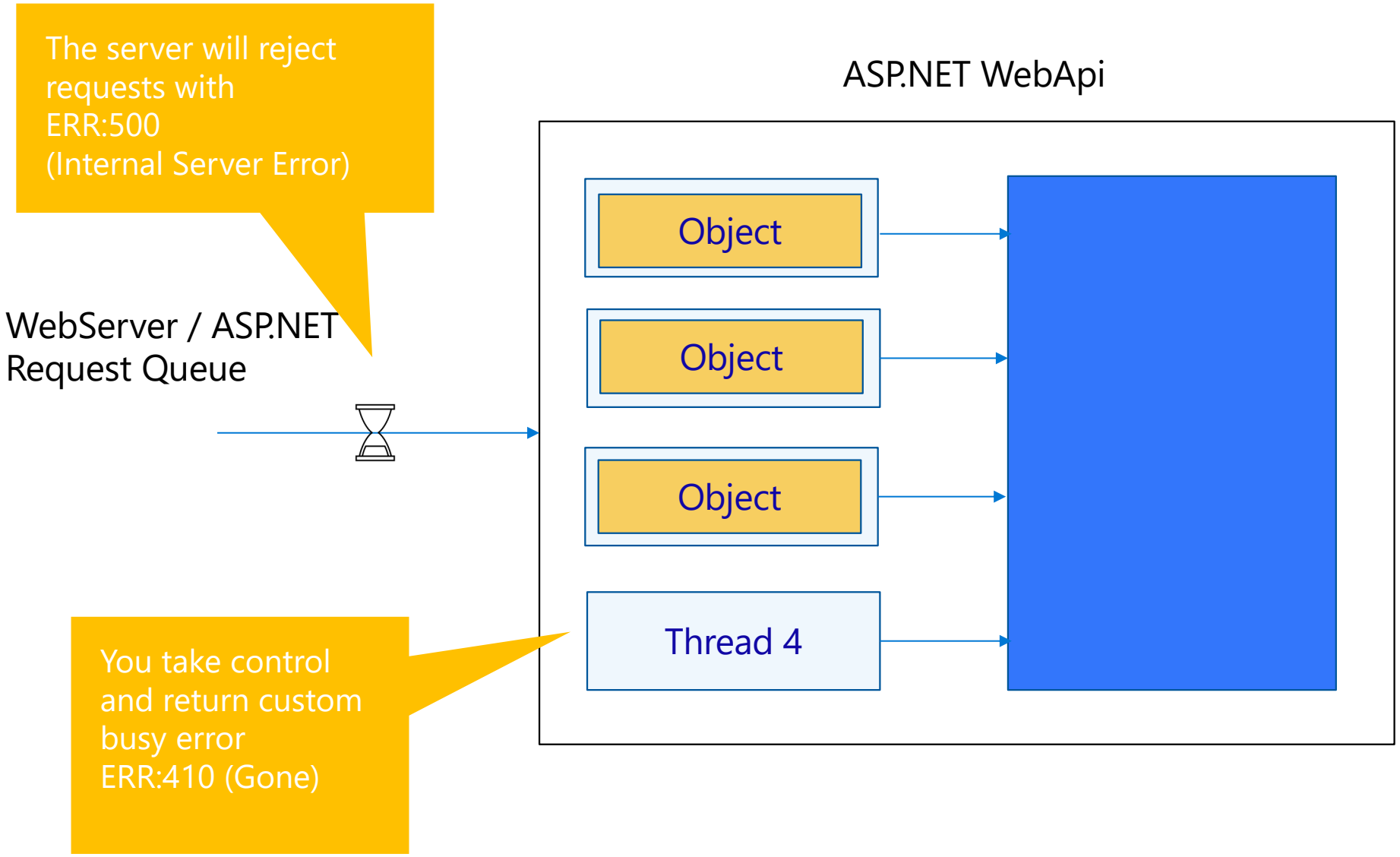
Cost Saving by prolonging request time



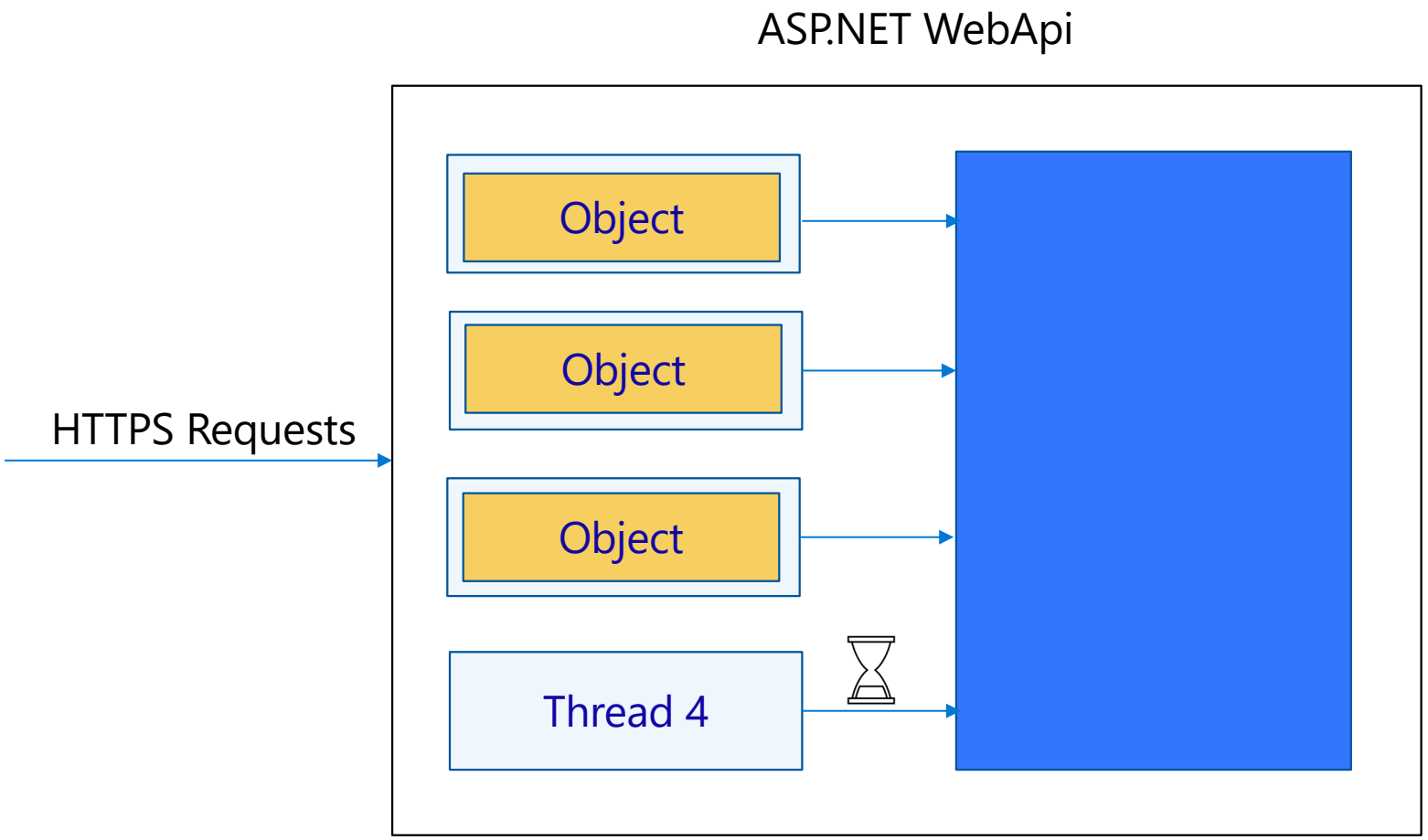
Time Dilation

You can save the space if you slow dwn the processing time

WebServer will reject requests on high load

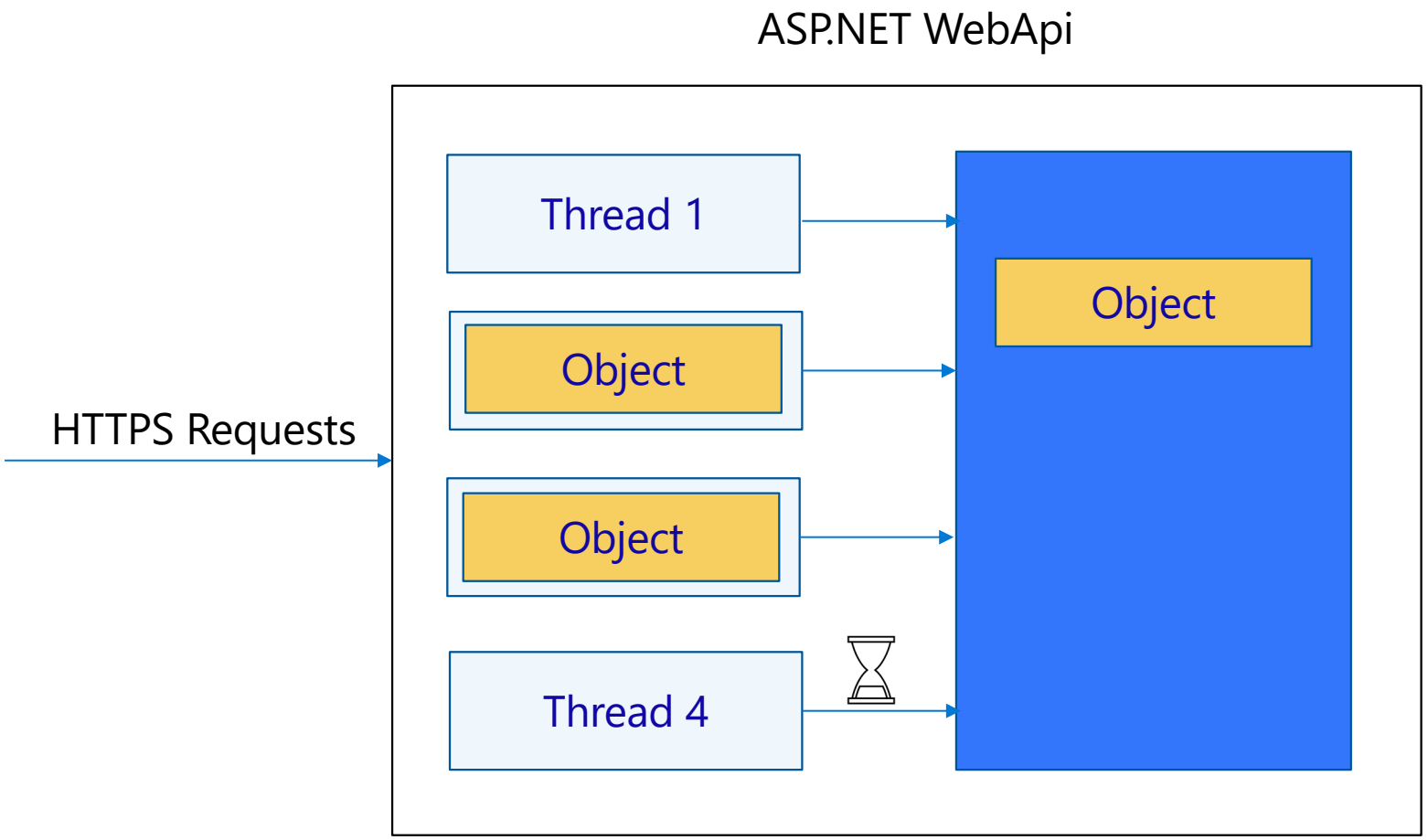


Object Pool Pattern with the “request queue”

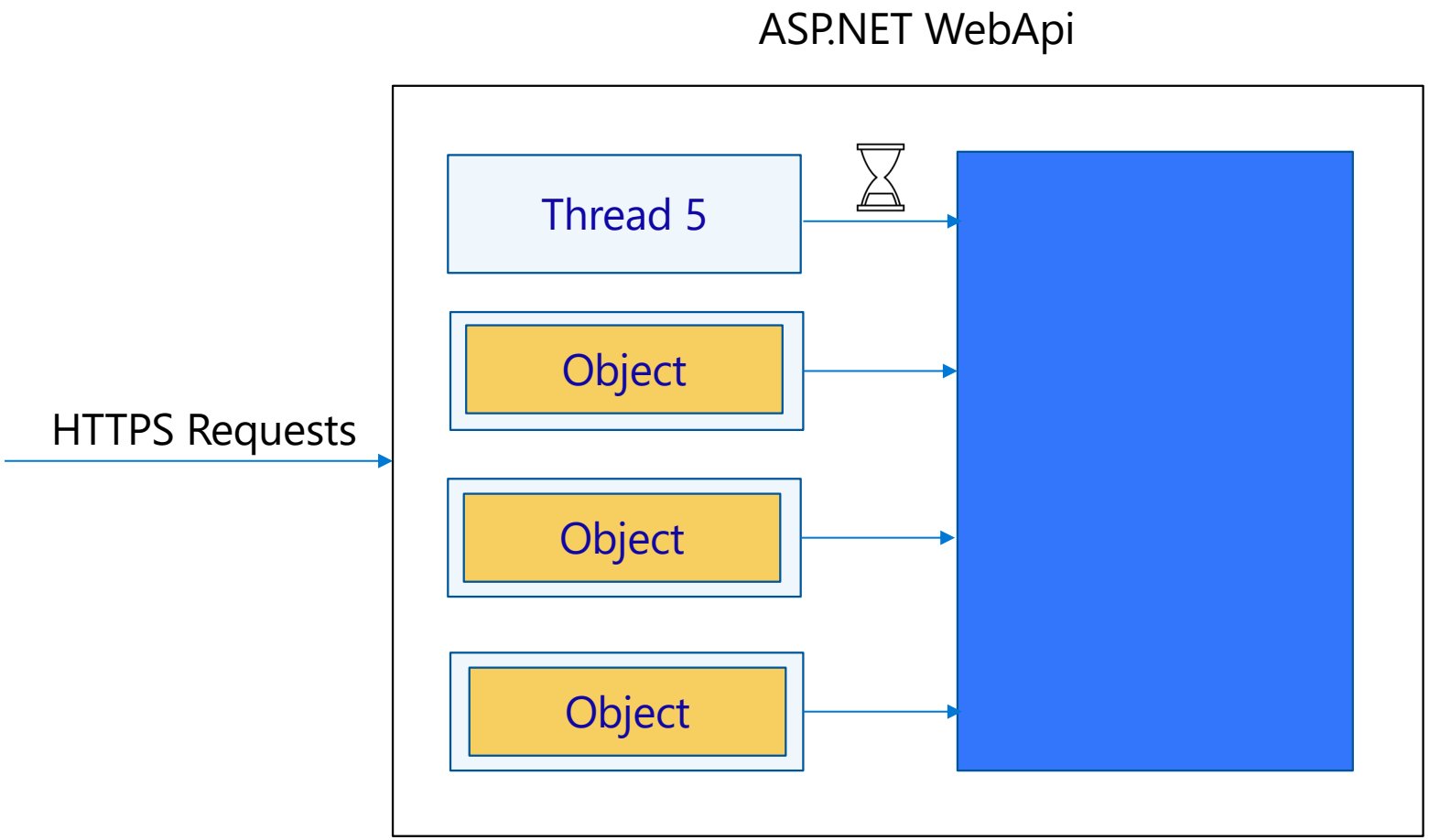


machinelearning/ObjectPool.cs at 510f0112d4fbb4d3ee233b9ca95c83fae1f9da91 · dotnet/machinelearning (github.com)

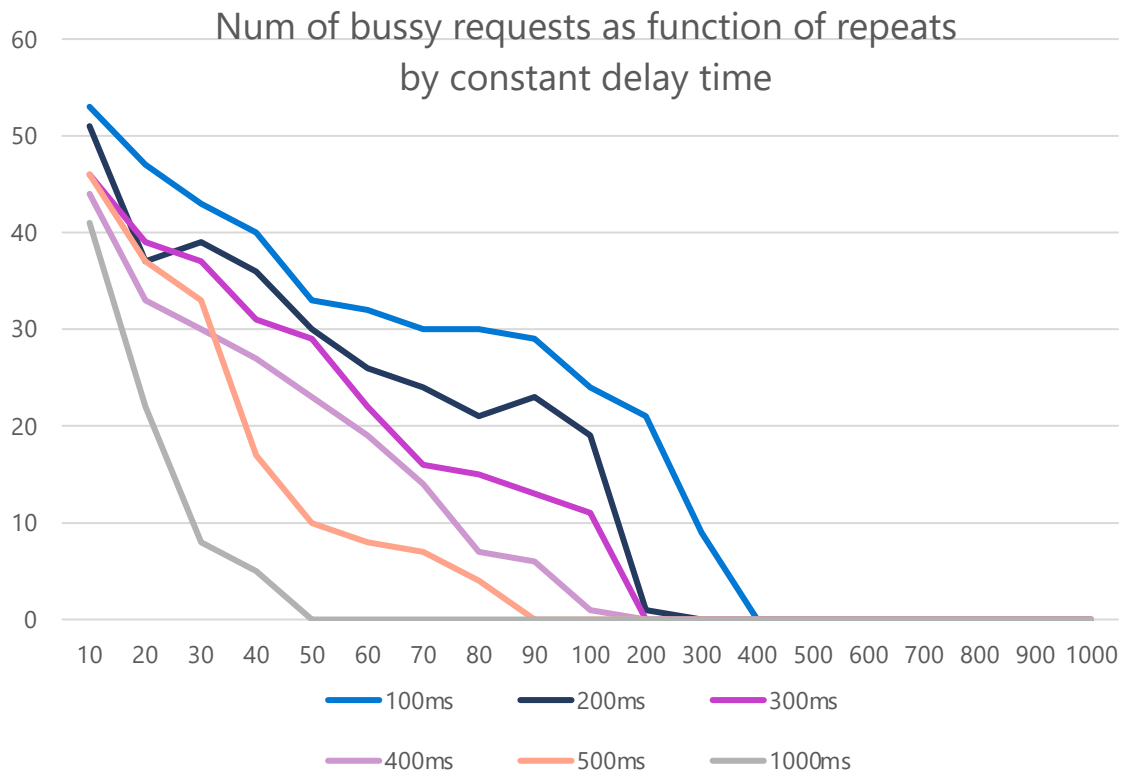
Object Pool Pattern with “request queue”

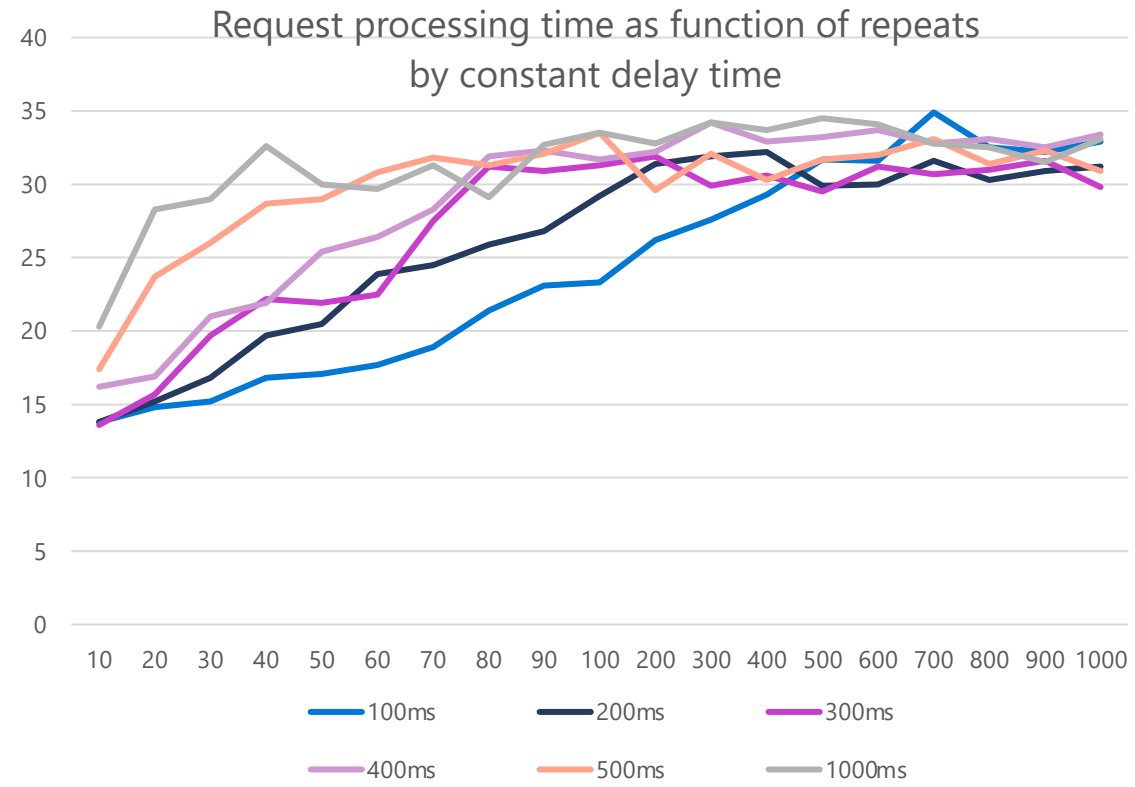


Object Pool Pattern with “request queue”



This really works ☺





Summary

- Large objects consume an unpredictable amount of RAM.
- Docker Host terminates your code with Err 134 and 137.
- Batch Service and VM are more tolerant
- You can implement Object Pool = Large Objects as a Service
- You can keep the size and costs, but you must slow down processing (time dilation).

https://github.com/ddobric/azure-bestpractices

Vielen Dank!

- Vielen Dank für deine Teilnahme!
- Dein Feedback ist uns wichtig!.

Damir Dobric
daenet GmbH – ACP Digital

Lead Software Architect
Microsoft Regional Director
Azure MVP

https://damirdobric.me

daenet – an ACP Digital company

thank you

questions?

TWITTER HANDLE

BLOG URL