

# Case Study #2

Memilavi  
[www.memilavi.com](http://www.memilavi.com)

<https://t.me/learningnets>



*Muscat.auto*

**The Real Autonomous Car**

# *Muscar.auto*

---

- Manufactures autonomous systems for vehicles
- Has >10,000 vehicles on the roads right now
- Expects more than 200,000 vehicles by end of year
- Needs to reliably receive telemetry from cars and display data about them

<https://t.me/learningnets>



## Requirements

```
graph TD; Requirements[Requirements] --> Functional[Functional]; Requirements --> NonFunctional[Non-Functional];
```

### Functional

What the system should do

1. Web Based
2. Receive telemetry from cars (location, speed, breakdowns, etc)
3. Store telemetry in a persistent store
4. Display dashboards summarizing the data
5. Perform analysis on the data

### Non-Functional

What the system should deal with

## NFR - What We Know

1. Data intensive system
2. Not a lot of users
3. A lot of data
4. Performance is important

## NFR - What We Ask

1. *“How many expected concurrent users?”* 10
2. *“How many telemetry messages received per second?”* 7,000
3. *“What is the average size of message?”* 1KB
4. *“Is the message schema-less?”* Yes

NFR - What We Ask

5. *“Can we tolerate some message loss?”*

Sort of...

6. *“What is the desired SLA?”*

Highest Possible

## Data Volume

- 1 Message = 1KB
- 7,000 messages / sec = 7MB / sec
  - => ~25GB / hr
  - => ~605GB / day
  - => ~221TB / year ← That's a lot!

## Retention Period

Defines for how long records are kept in the database

What happens to them after the retention period?

- Deleted
- Moved to archive data store

## Retention Period

### Motivation:

- Keep database from exploding
- Improve query performance

AWS Config adds the ability to specify a **data retention** policy for your configuration items

## Retention Period

Muscar needs two types of data:

- Operational, near-real-time (location, speed, etc.)
- Aggregated and ready for analysis (BI – Business Intelligence)

## Retention Period

Data Type	Used for...	Retention Period
Operational	Monitor real time data from cars. Performance is critical	
Aggregated	Reports, BI. Not real time, can be slower.	

## Retention Period

Data Type	Used for...	Retention Period
Operational	Monitor real time data from cars. Performance is critical	1 week
Aggregated	Reports, BI. Not real time, can be slower.	Forever

## Data Volume

- 1 Message = 1KB
- 7,000 messages / sec = 7MB / sec
  - => ~25GB / hr
  - => ~605GB / day
  - => ~221TB / year

## Data Volume

- 1 Message = 1KB
- 7,000 messages / sec = 7MB / sec

=> ~25GB / hr

=> ~605GB / day

=> ~4TB / week

## Requirements

```
graph TD; Requirements[Requirements] --> Functional[Functional]; Requirements --> Non-Functional[Non-Functional];
```

### Functional

What the system should do

1. Web Based
2. Receive telemetry from cars (location, speed, breakdowns, etc)
3. Store telemetry in a persistent store
4. Display dashboards summarizing the data
5. Perform analysis on the data

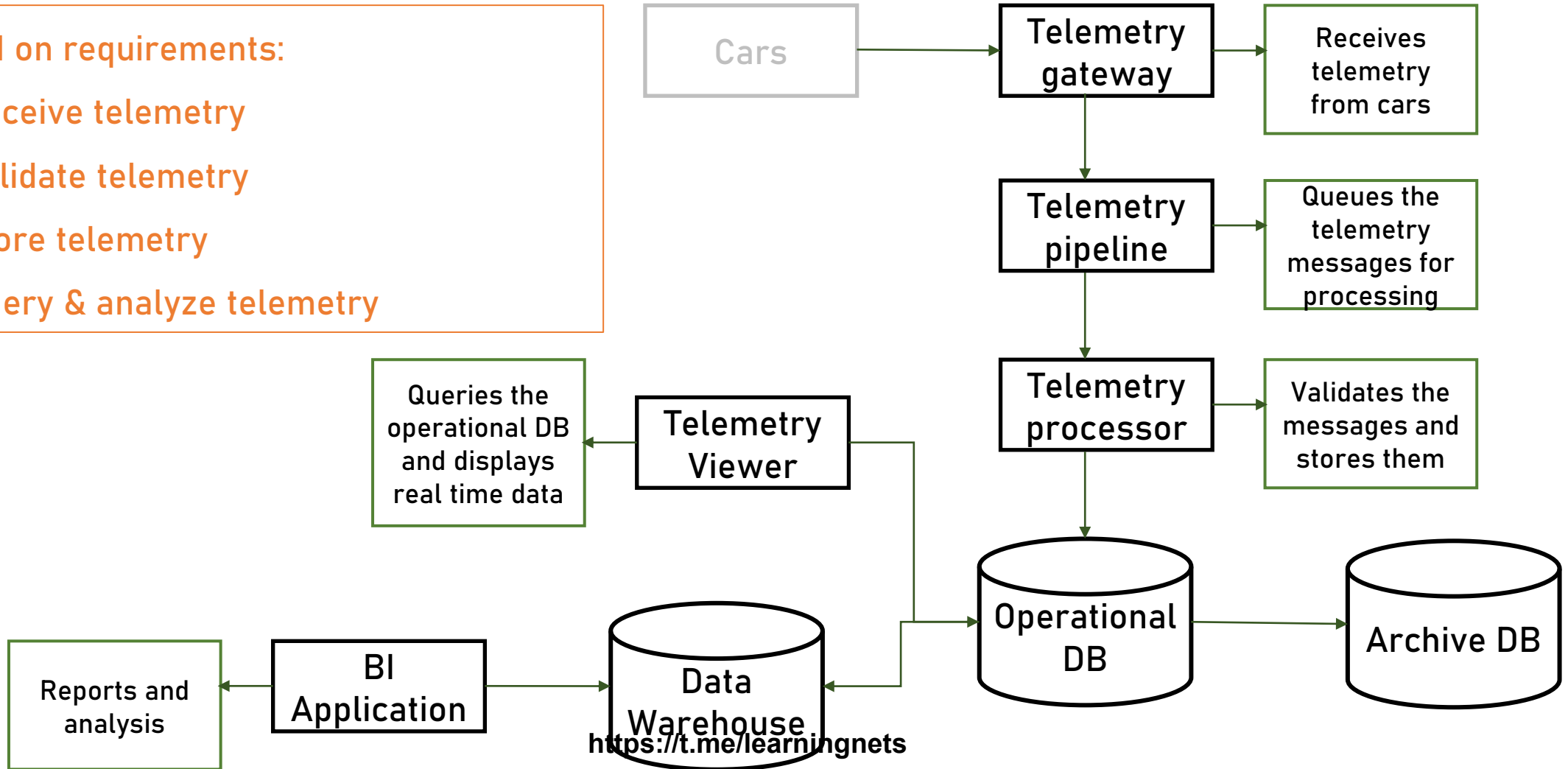
### Non-Functional

What the system should deal with

1. 10 Concurrent users
2. 7,000 msgs/sec
3. Max data in the operational DB: 4TB
4. Mission critical
5. Performance is critical

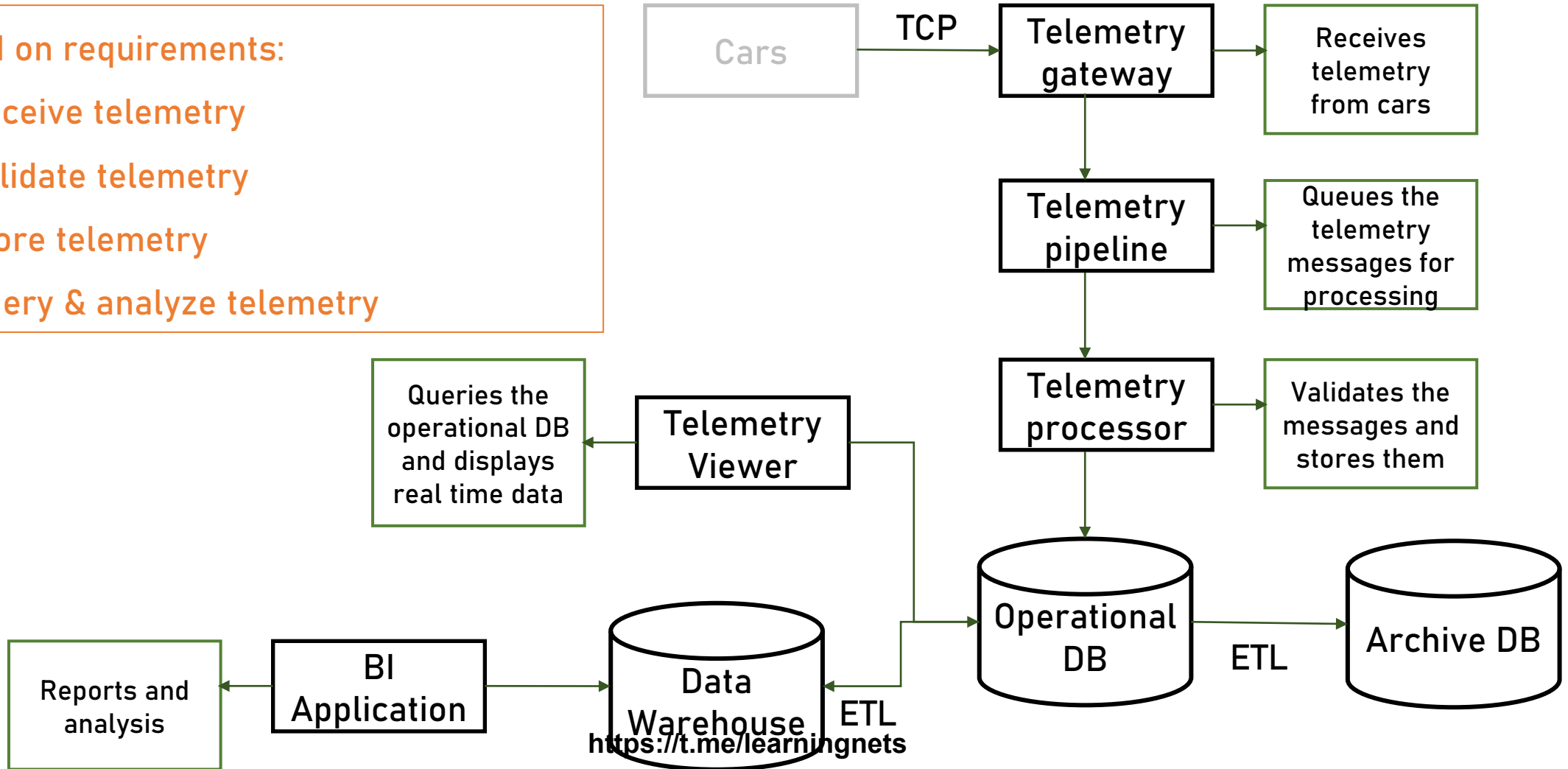
## Components

- Based on requirements:
1. Receive telemetry
  2. Validate telemetry
  3. Store telemetry
  4. Query & analyze telemetry

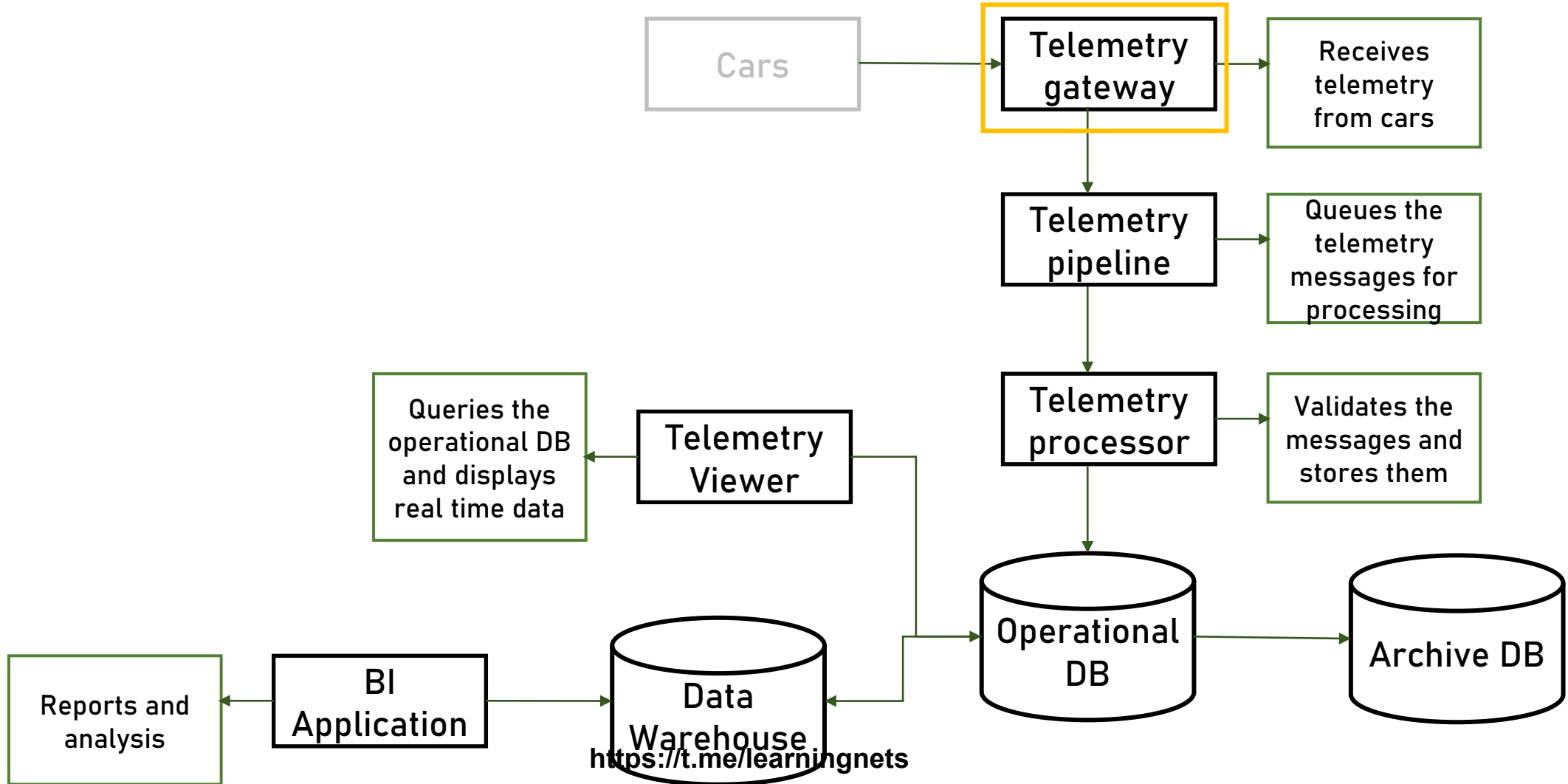


## Messaging

- Based on requirements:
1. Receive telemetry
  2. Validate telemetry
  3. Store telemetry
  4. Query & analyze telemetry



# Components





## Telemetry Gateway

What it does:

- Receives telemetry data from cars using TCP
- Pushes the telemetry data to the pipeline

## Application Type

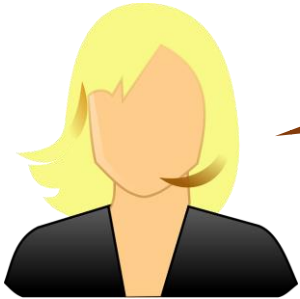
- Web App & Web API 
- Mobile App 
- Console 
- Service 
- Desktop App 

## Technology Stack

### Considerations:

- Load (7,000 msgs/sec)
- Performance
- Team's current knowledge
- Environment (OS, etc)

## Technology Stack



Our developers are familiar with Python, and are experts in JavaScript. In addition, we use only Linux servers.

**Python can't be used for the gateway**

**Too slow**

We look for a language with great performance, runs on linux, and leverages current skills (Python & JavaScript)



## Technology Stack



Great performance



Runs on Linux



Leverages JS skills

## Listeners in Azure



IoT Hub



Event Grid



App Service



Functions

**No raw TCP support**

## Listeners in Azure



- Not ideal
- Requires the most manual maintenance
- ...but allows most flexibility

## Scaling

Remember:

- Load (7,000 msgs/sec)
- Performance

# Scaling



Scaling



+



## Virtual Machine Scale Sets

REGION:  OPERATING SYSTEM:  TYPE:  TIER:

CATEGORY:  VM SERIES:  INSTANCE:

VIRTUAL MACHINES:

### Savings Options

Save up to 72% on pay-as-you-go prices with 1-year or 3-year Reserved Virtual Machine Instances. Reserved Instances are great for applications with steady-state usage and applications that require reserved capacity. [Learn more about Reserved VM Instances pricing.](#)

**Compute (D4s v4)**

- Pay as you go
- 1 year reserved (~41% discount)
- 3 year reserved (~62% discount)

COMPUTE PAYMENT OPTIONS:

\$127.60  
Average per month  
(\$0.00 charged upfront)

= \$127.60  
Average per month  
(\$0.00 charged upfront)

Upfront cost	\$0.00
Monthly cost	\$127.60

### Load Balancer

REGION:  TIER:

---

#### Load Balancer rules

Rules = \$18.25

---

#### NAT rules

ⓘ NAT rules are free. = \$0.00

---

#### Data processed

× \$0.005 Per GB = \$5.00

---

Upfront cost	\$0.00
Monthly cost	\$23.25

Architecture

Traditional:

User Interface /  
Service Interface

Business Logic

Data Access

Data Store

<https://t.me/learningnets>

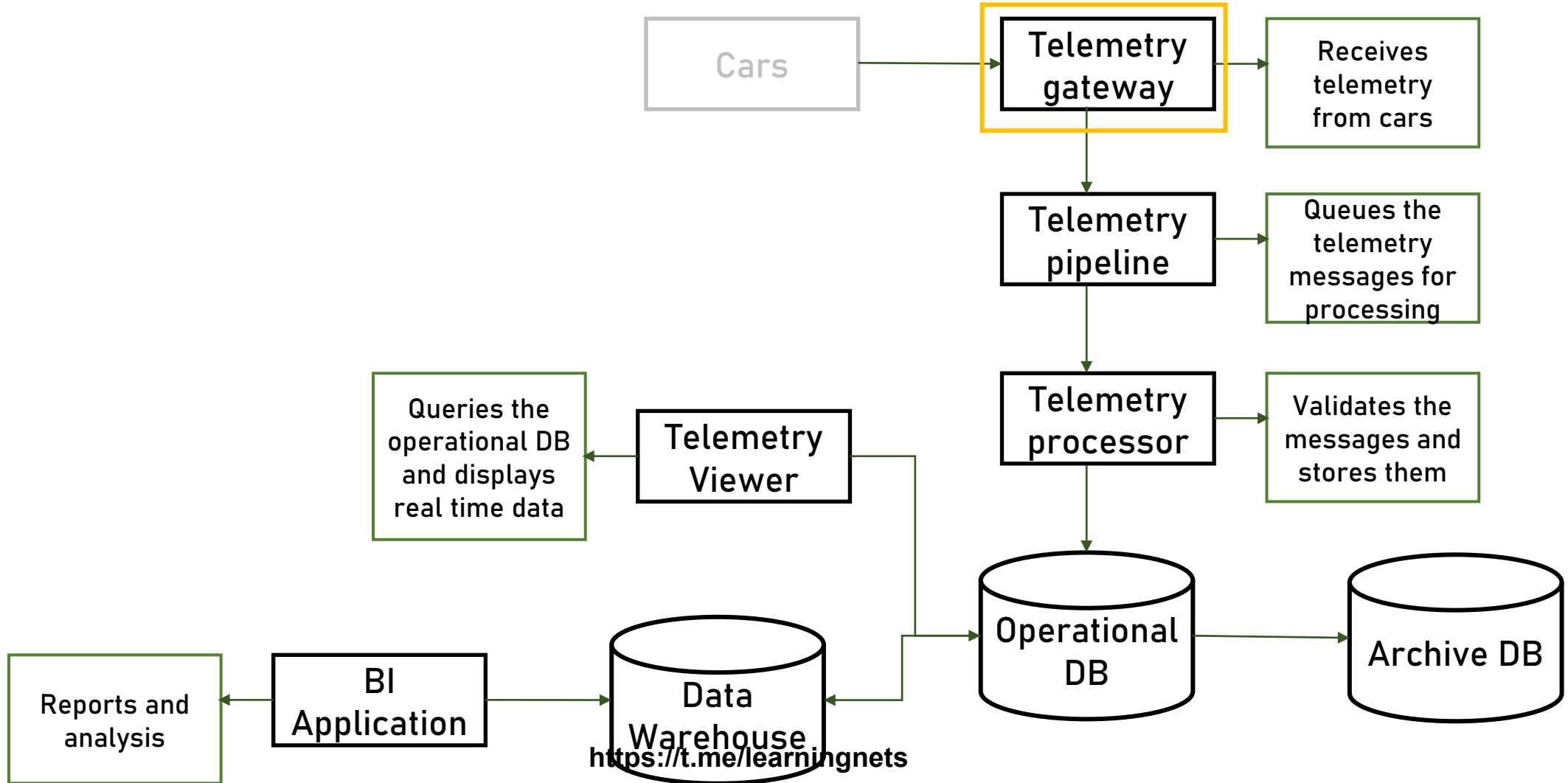
Architecture

In our case:

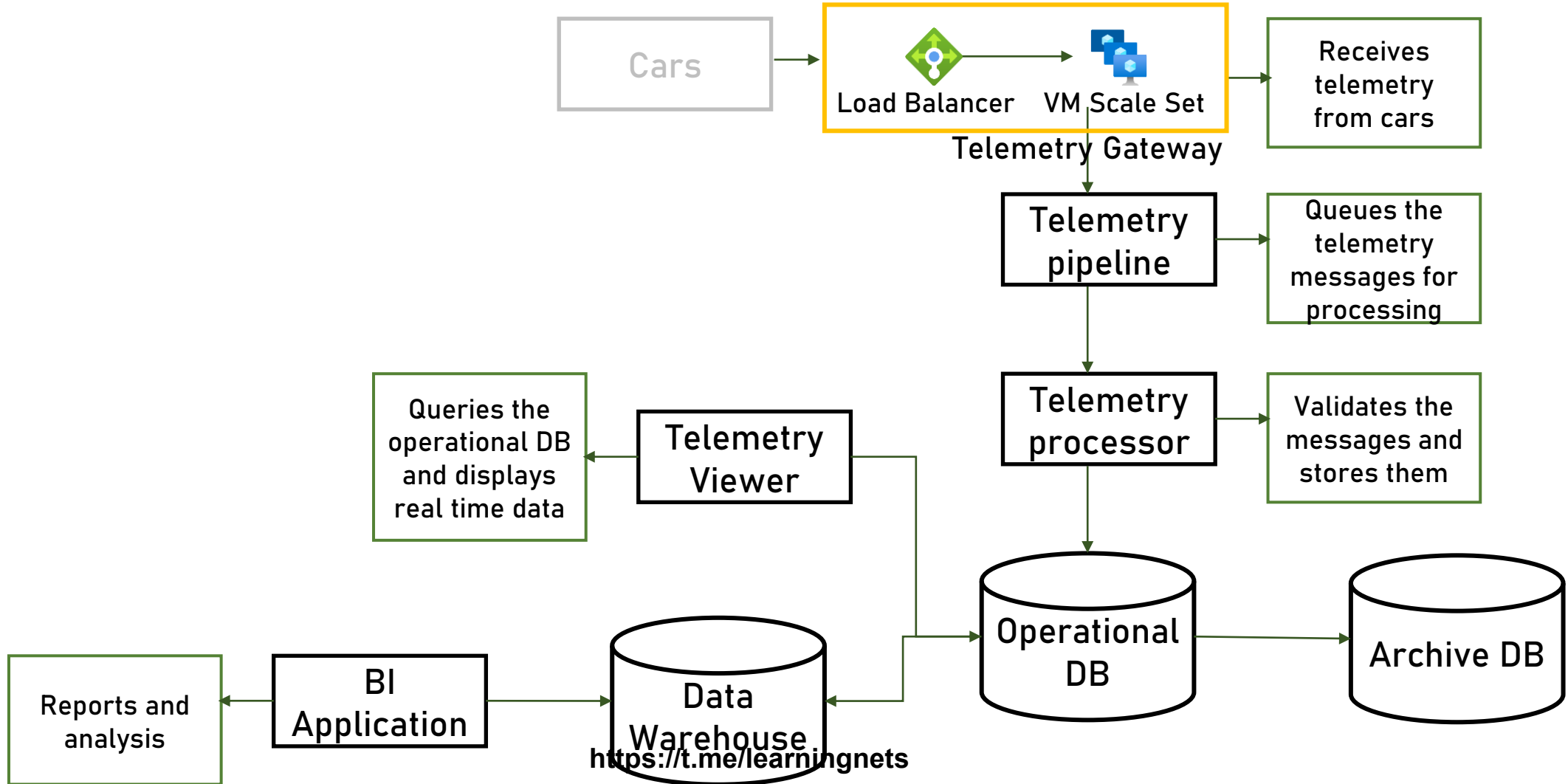
Service Interface

Pipeline

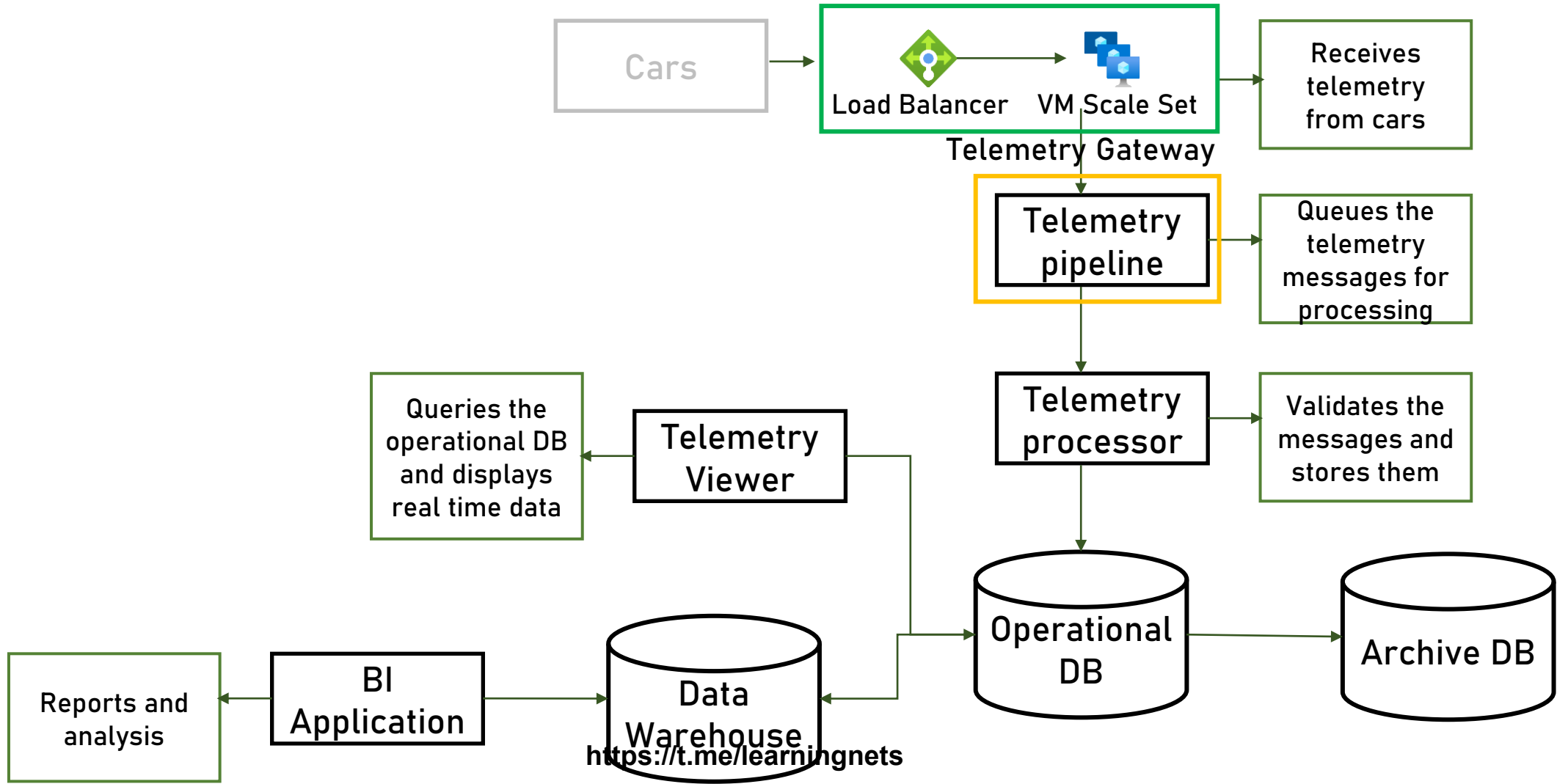
# Components



# Components



# Components







## Telemetry Pipeline

What it does:

- Gets the telemetry messages from the gateway
- Queues the telemetry for further processing
- Basically – a queue for streaming high volume data

## Messaging in Azure

Service	Used For...	Guarantees Order	Max Msg Size	And also...
Storage Queue 	Dead simple queueing	Yes	64KB	Extremely simple, no additional cost
Event Grid 	Event driven architectures	No	1MB	Great integration with other services
Service Bus 	Advanced queueing solutions	Yes	256KB	Advanced messaging features, durable
Event Hubs 	Big data streaming	Yes	1MB	Low latency, designed for heavy load

Each TU supports up to 1k msgs / sec

**Event Hubs**

REGION: West Europe

TIER: Standard

---

**Units**

Maximum throughput units: 20. Up to 1 MB per second of ingress events. Up to 2 MB per second of egress events.

7 Throughput units × 730 Hours × \$0.030 Per unit/hour = \$153.30

Enable Capture

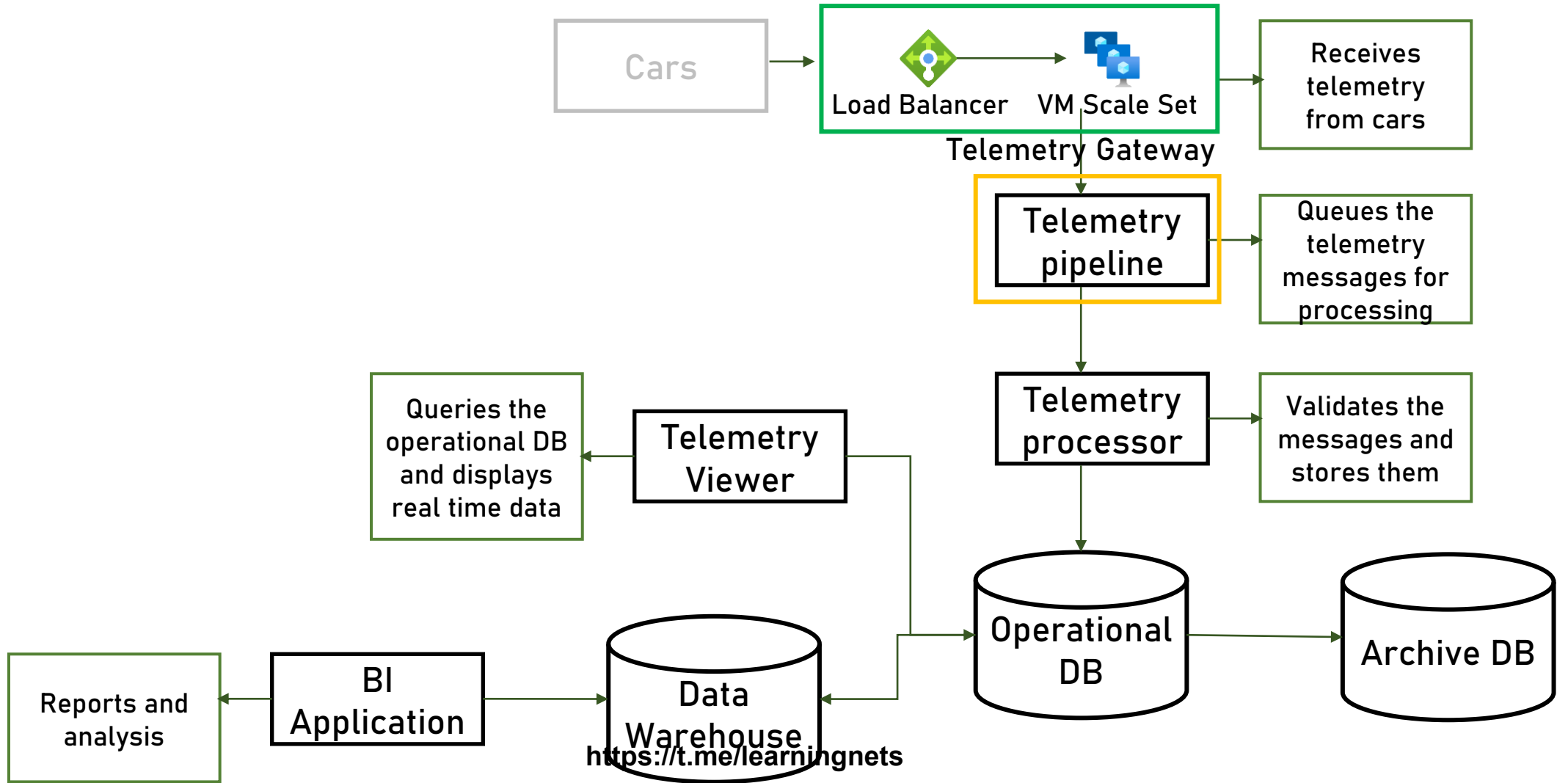
**Ingress**

10 Million Events per month × \$0.028 Per million Events / month = \$0.28

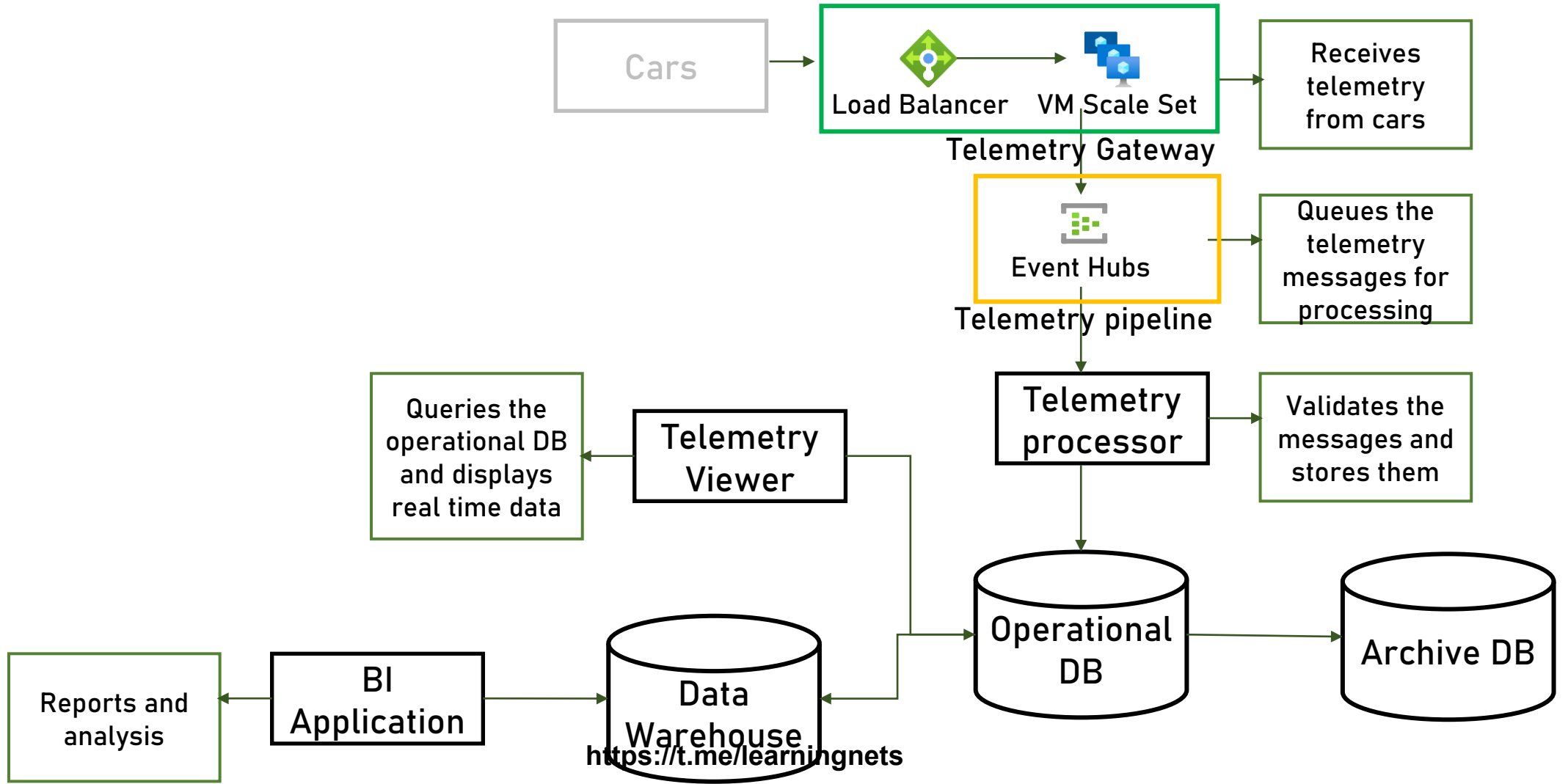
---

Upfront cost	\$0.00
Monthly cost	\$153.58

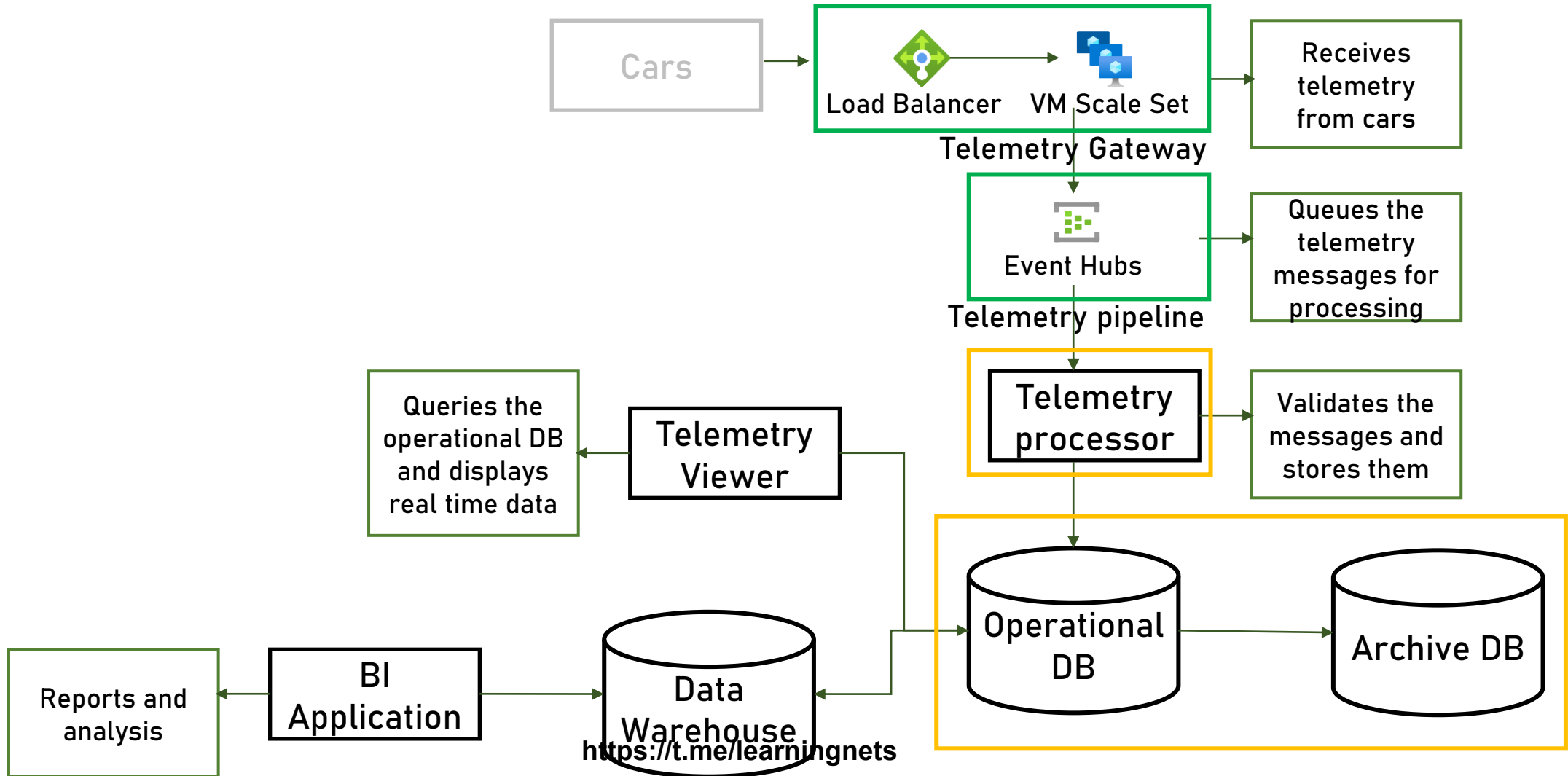
# Components



# Components



## Components



## Telemetry Processor

What it does:

- Receives the messages from the pipeline
- Processes the messages (mainly validation)
- Stores the messages in a data store

## Cloud Services

For:

- The processor
- The datastore

## Processor

### Function App

- Designed for lightweight operations
- Great, built-in integration with Event Hubs
- Cost effective
- Autoscaling

### Azure Functions

REGION:  TIER:

*The first 400,000 GB/s of execution and 1,000,000 executions are free.*

**Executions**

Memory size:  ×  ×  = \$0.00  
Execution time (in milliseconds)      Executions per month

**Requests**

= \$1.80  
Execution count

---

Upfront cost	\$0.00
Monthly cost	<input type="text" value="\$1.80"/>

## Data store

### What we're looking for:

- Schema-less message support
- Quick retrieval
- No complex queries

## Technology Stack



Cosmos DB

- Schema-less message support
- Quick retrieval
- No complex queries
- In addition:
  - Multi-region read / write
  - Multiple APIs
  - Great performance



Azure Cosmos DB

DATABASE OPERATIONS: Standard provisioned throughput (manual) [v]

WRITE REGIONS: Single Region Write (Single-Master) [v]

No need for 7000 RU/s, Event Hubs balances load

### Savings Options

Save up to 65% on pay-as-you-go prices with 1 year or 3 year Reserved Capacity options.

- Pay as you go
  - 1 year reserved capacity
  - 3 year reserved capacity
- \$23.36  
Average per month  
(\$0.00 charged upfront)

Request units per second (RU/s) ⓘ

400 RU/s × 730 Hours [v]

Write Region:

West Europe [v]

400 RU/s ×

730 Hours ×

\$0.008 Per 100 RU/s per hour

Enable Availability Zones

= \$23.36

### Storage

Transactional Storage

4000 GB

Analytical Storage ⓘ

Enable Analytical Storage

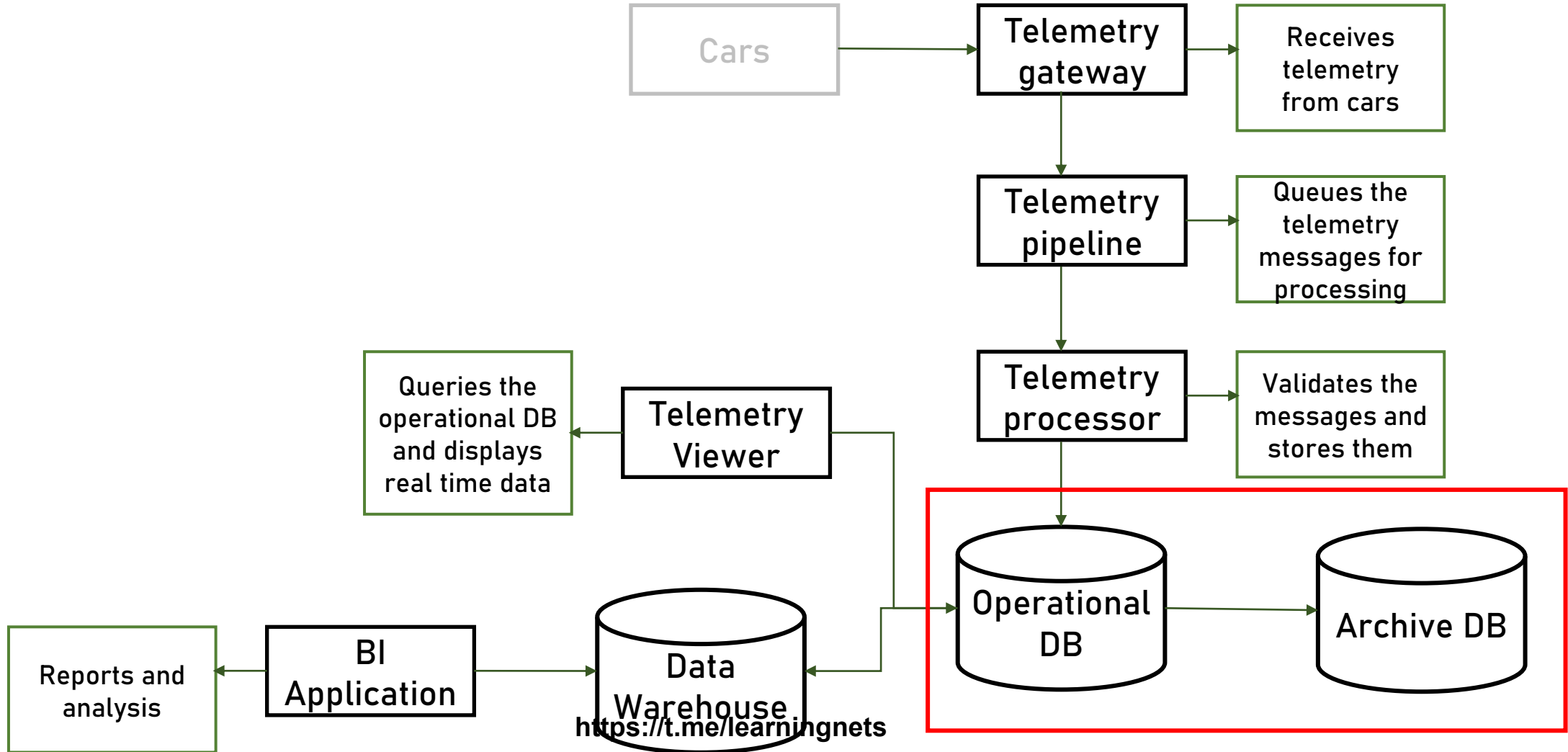
Regional Storage Costs

\$1,000.00

Upfront cost \$0.00

Monthly cost \$1,023.36

# Components



## Archive Database

### Archive– what we're looking for:

- Support for a huge amount of data (221TB / Year)
- Not accessed frequently
- No need for fast retrieval
- Save costs

## Archive Database



## Storage Account

- Huge amounts of data (221TB / Year)
- Not accessed frequently
- No need for fast retrieval
- Save costs



### Storage Accounts

REGION: West Europe | TYPE: Block Blob Storage | PERFORMANCE TIER: Standard | STORAGE ACCOUNT TYPE: General Purpose V2

ACCESS TIER: Archive | REDUNDANCY: LRS

Capacity: 221 TB

Early deletion fees may apply and are not included. [Learn more about early deletion fees.](#)

#### Savings Options

Save up to 38% on pay-as-you-go prices with 1-year or 3-year Azure Storage Reserved Capacity. [Learn more about Azure Storage Reserved Capacity pricing.](#)

Pay as you go  
 1 year reserved  
 3 year reserved

\$407.35  
Average per month  
(\$0.00 charged upfront) = \$407.35  
Average per month  
(\$0.00 charged upfront)

#### All other operations

1 Operations	×	\$0.004 Per 10,000 operations	= \$0.01
--------------	---	----------------------------------	----------

#### Data Retrieval

1000 GB	×	\$0.024 Per GB	= \$24.00
---------	---	-------------------	-----------

#### Archive high priority retrieval

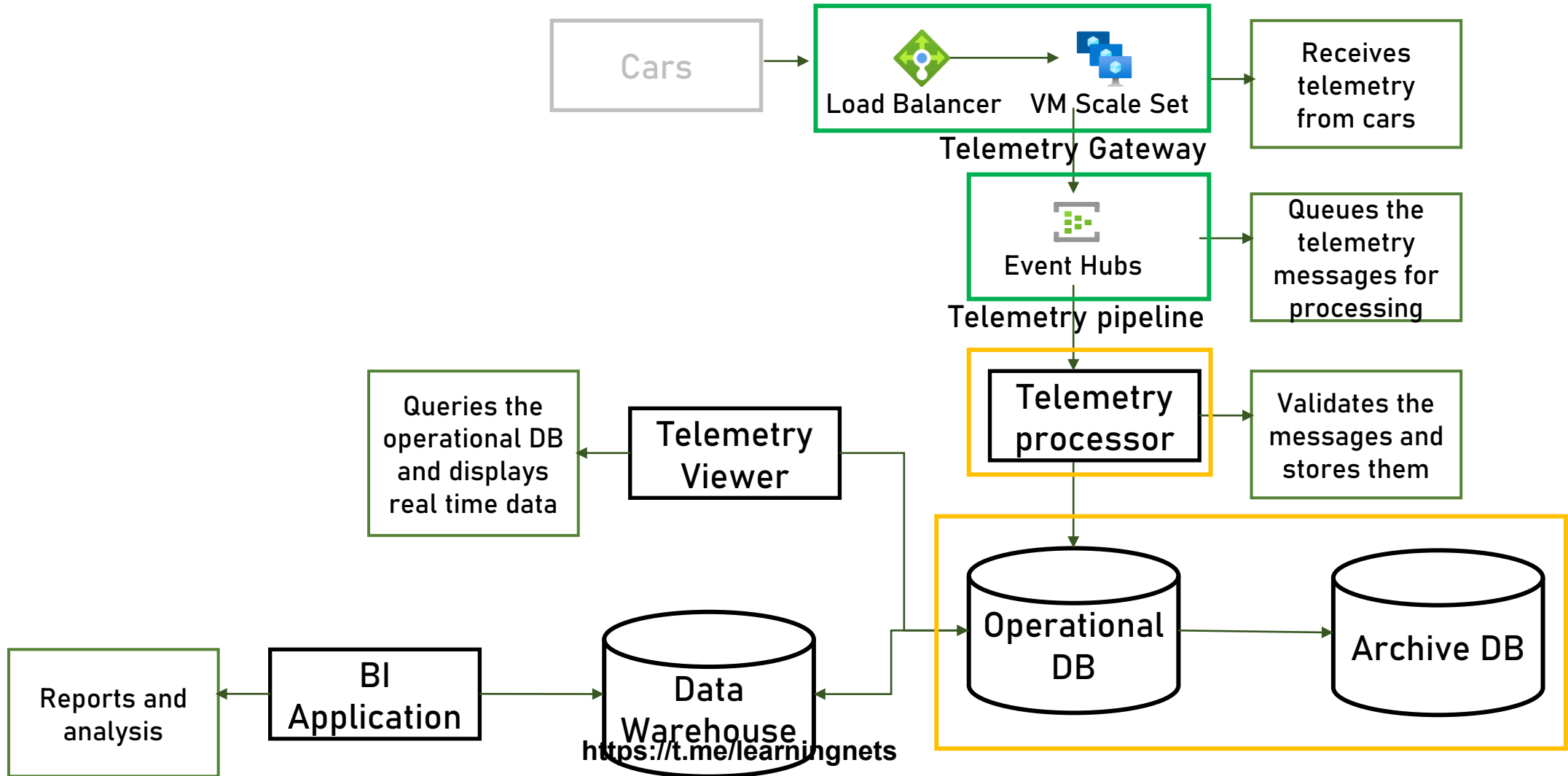
1 GB	×	\$0.130 Per GB	= \$0.13
------	---	-------------------	----------

#### Data write

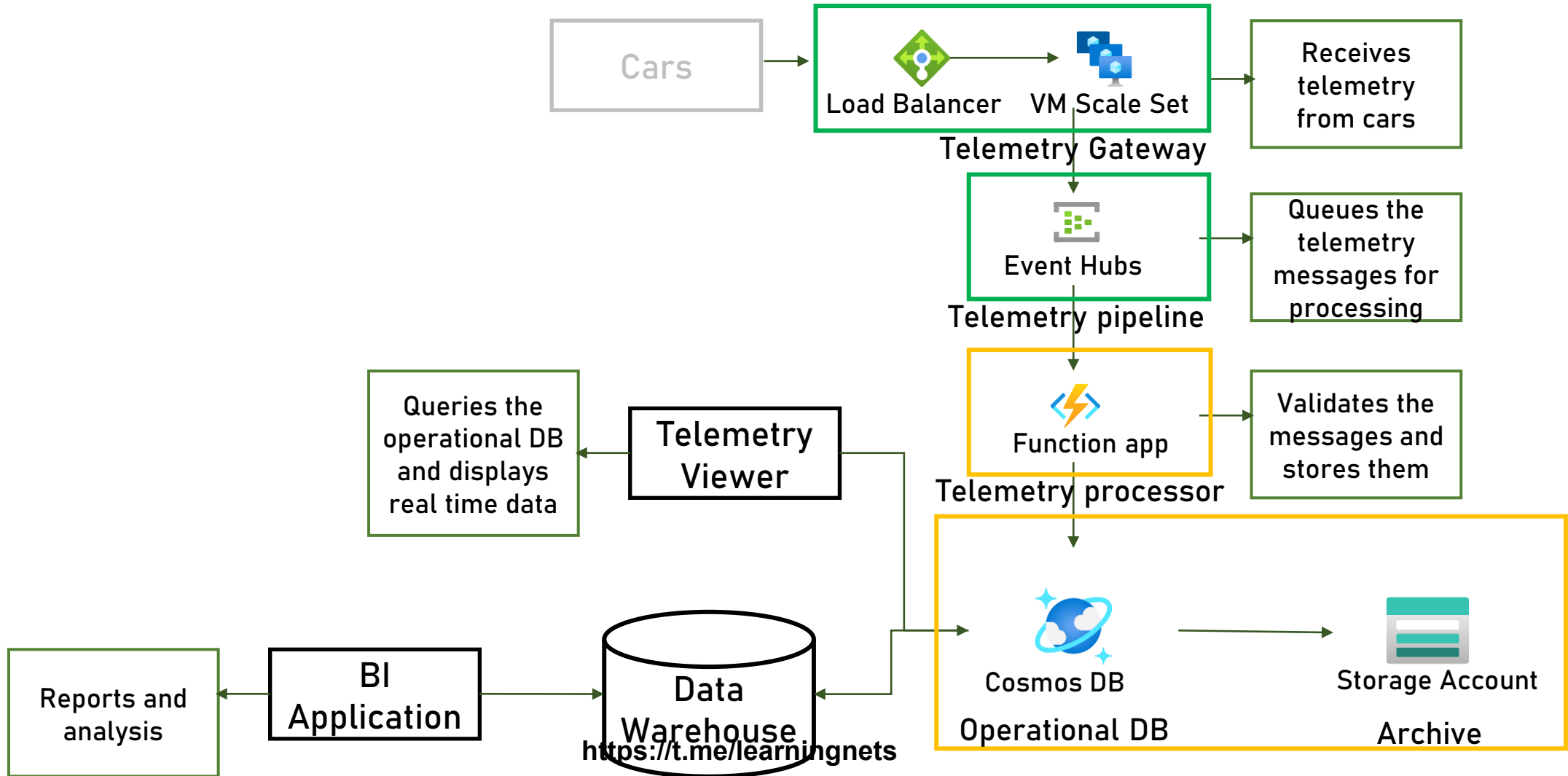
1000 GB	×	\$0.000 Per GB	= \$0.00
---------	---	-------------------	----------

Upfront cost	\$0.00
Monthly cost	\$433.22

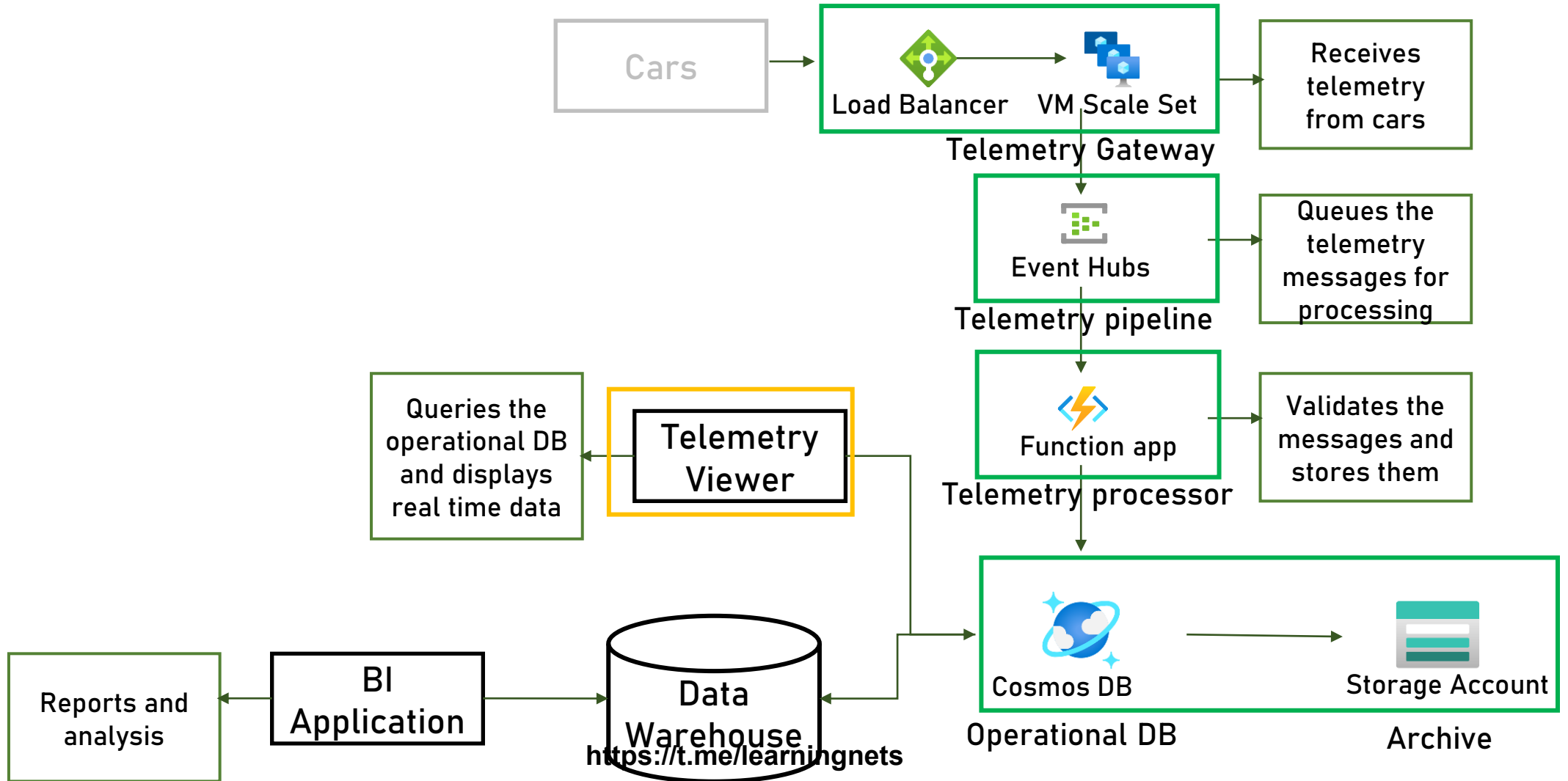
# Components



# Components



# Components



## Telemetry Viewer

What it does:

- Allows end users to query telemetry data
- Displays real time data

What it doesn't:

- Analyzes the data

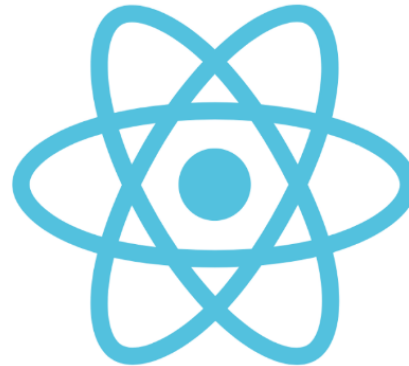
## Application Type

- Web App & Web API ✓
- Mobile App ✗
- Console ✗
- Service ✗
- Desktop App ✗

## Technology Stack

Back End

Front End



**React JS**

## Azure Web API



### App Service

- Fully managed web app & API
- Supports many platforms
- Autoscale
- Support for WebJobs

## Azure Web API

**App Service**

REGION: West Europe    OPERATING SYSTEM: Windows    **TIER: Standard**

**Standard**

**INSTANCE: S1: 1 Cores(s), 1.75 GB RAM, 50 GB Storage, \$0.100**

1 Instances × 730 Hours = \$73.00

SSL Connections

Upfront cost	\$0.00
Monthly cost	<b>\$73.00</b>

## Architecture

Service Interface

Business Logic

Data Access

Data Store

<https://t.me/learningnets>

## API


- Get latest errors for all cars
- Get latest telemetry for specific car
- Get latest errors for specific car


## API

Functionality	Path	Return Codes
Get latest errors for all cars	GET /api/v1/telemetry/errors	200 OK
Get latest telemetry for specific car	GET /api/v1/telemetry/{carId}	200 OK 404 Not Found
Get latest errors for specific car	GET /api/v1/telemetry/errors/{carId}	200 Ok 404 Not Found


## Telemetry Viewer Redundancy

### App service auto scale

**Default\*** Auto created scale condition 

Delete warning  The very last or default recurrence rule cannot be deleted. Instead, you can disable autoscale to turn off autoscale.

Scale mode  Scale based on a metric  Scale to a specific instance count







Rules  It is recommended to have at least one scale in rule. To create new rules, click [Add a rule](#).

Scale out

When	Default1	(Average) CpuPercentage > 70	Increase count by 1
Or	Default1	(Average) HttpQueueLength > ...	Increase count by 1

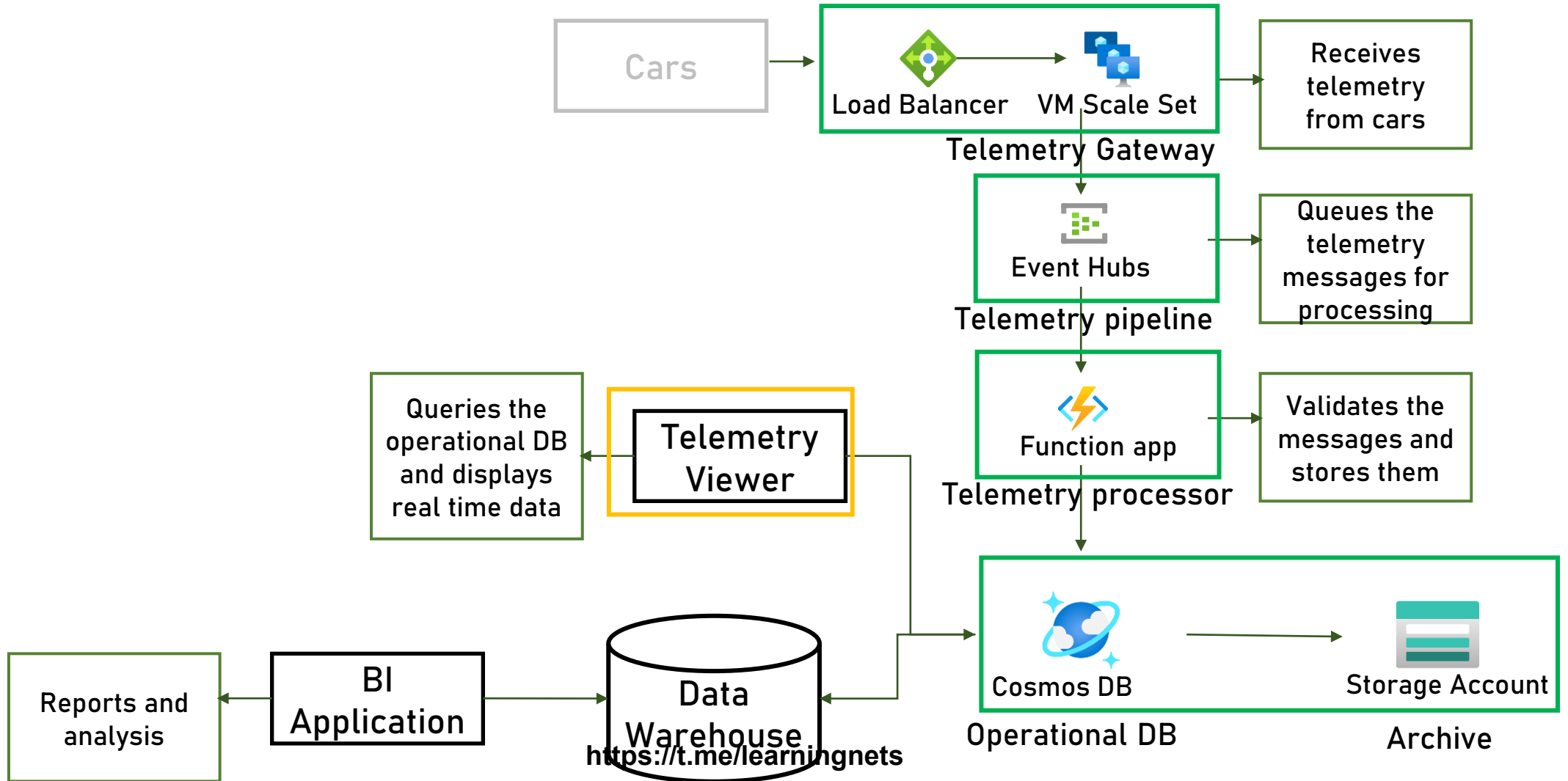
[+ Add a rule](#)

Instance limits

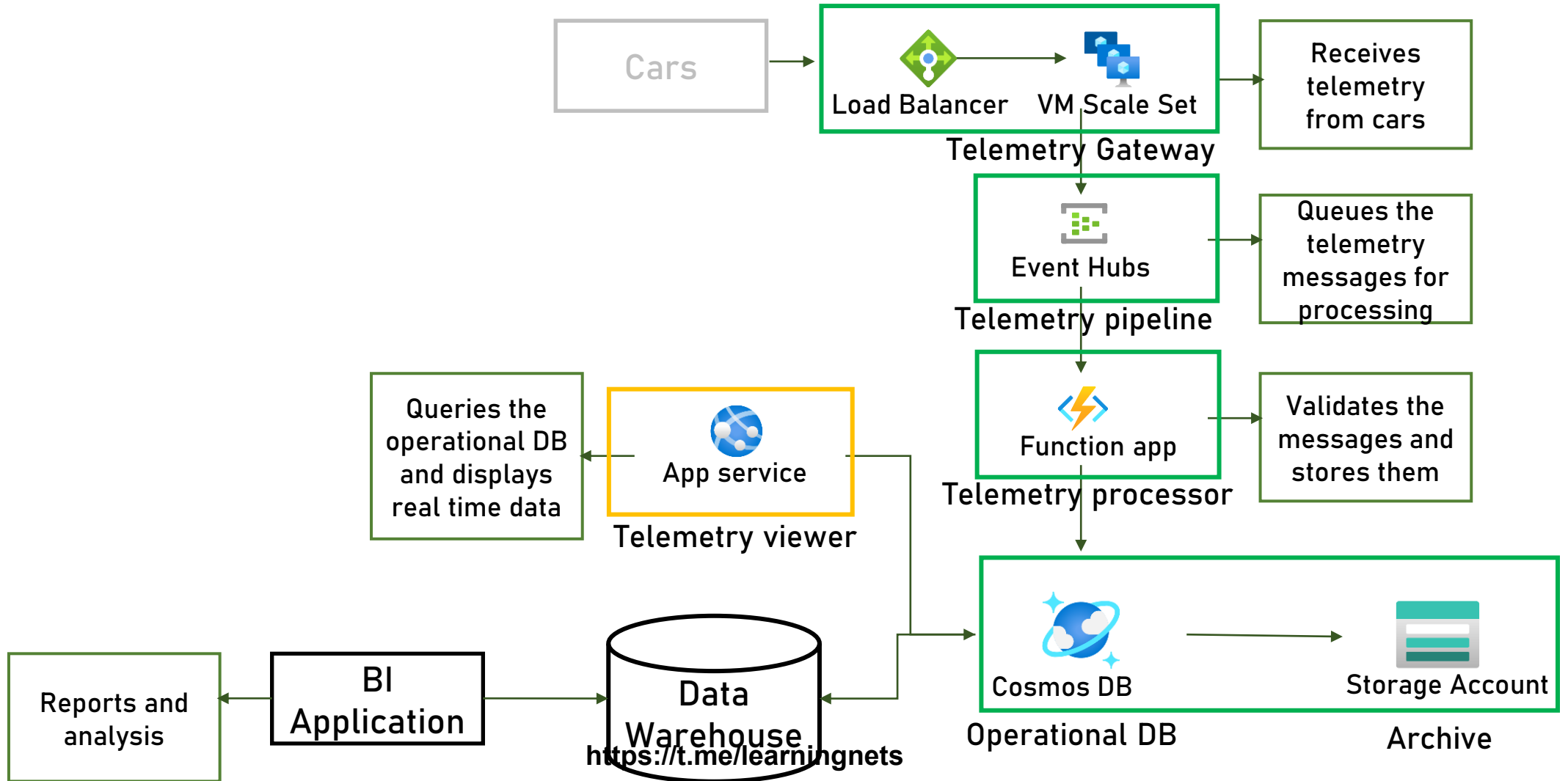
Minimum 	Maximum 	Default 
<input type="text" value="1"/> 	<input type="text" value="3"/> 	<input type="text" value="1"/> 

Schedule **This scale condition is executed when none of the other scale condition(s) match**

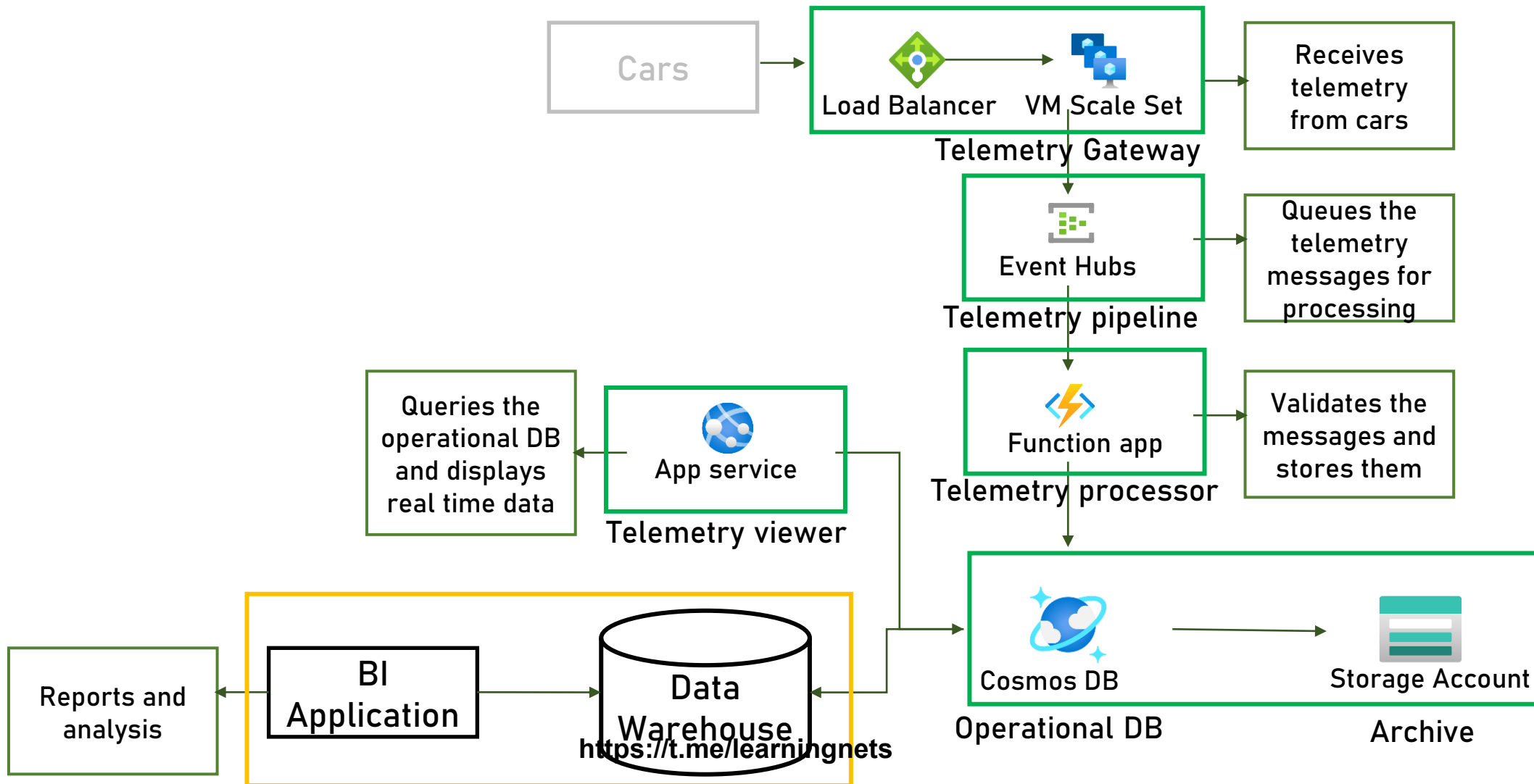
# Components

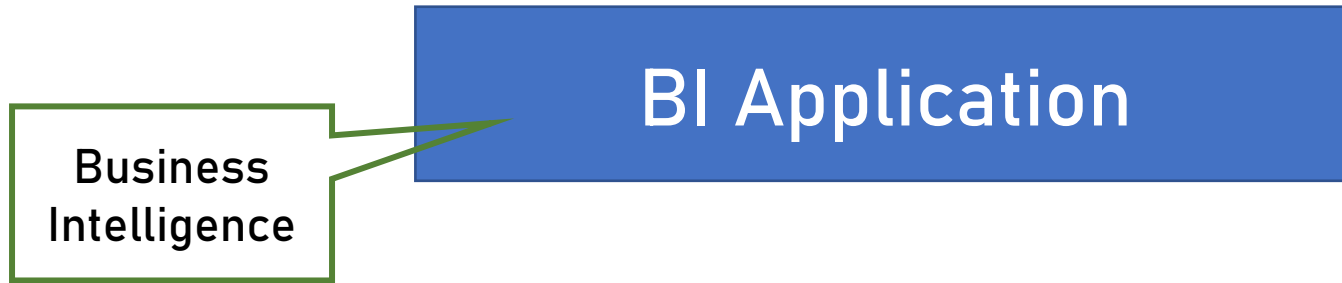


# Components



## Components





What it does:

- Analyzes telemetry data
- Displays custom reports about the data, trends, forecasts etc.
  - How many cars did break during the last month?
  - What is the total distance the cars drove?

## Application Type

- Doesn't matter
- BI Application is ALWAYS based on an existing tool

BI Tools



# BI Tools

Figure 1. Magic Quadrant for Analytics and Business Intelligence Platforms



## BI Tools

- An important lesson:
  - Designing BI solution is NOT part of the architect's job
  - ALWAYS use BI expert for this task

## Security

- Pay attention to:
  - Public accessible databases
  - Unprotected access to App Service

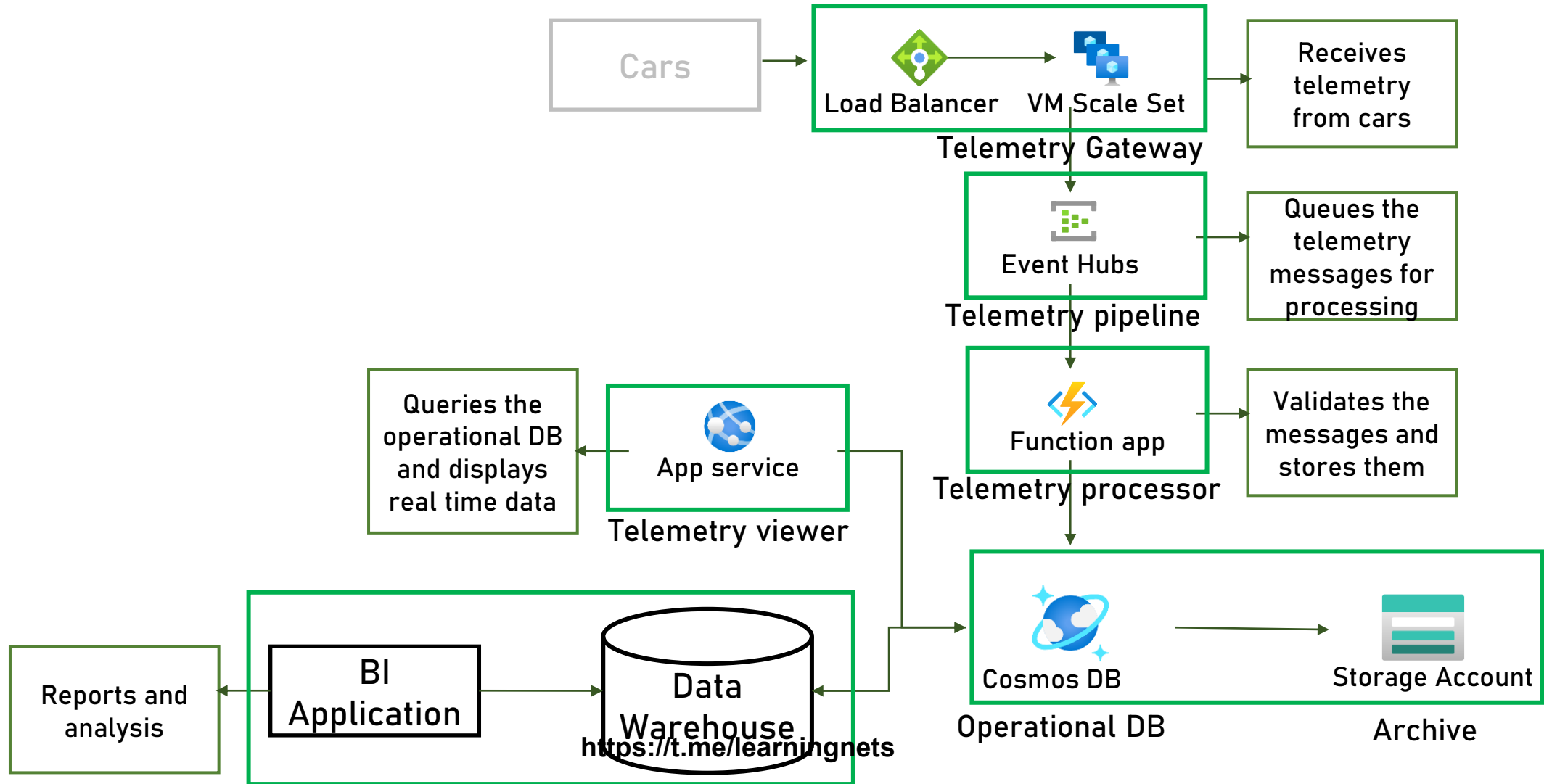
## Security

- To-Do:
  - Block access to databases from unauthorized IP addresses

## Security

- What about the App Service?
  - The client decided not to place WAF in front the App Service
    - Small service
    - Read-only operations
    - Save costs

# Architecture Diagram



Cost

Estimated upfront cost	\$0.00
Estimated monthly cost	\$1,835.82

Download detailed cost estimation  
from the lecture's resources