

Microsoft Azure Certified Data & Al Track

Vendée Globe Business Case

Table of Contents

Foreword

Information about Vendée Globe	3
Your Challenge	
Sailing Race	5
Getting Started	
How to create the Event Hub	7
How to get the Event Hub Namespace connecti	on string 11
Download and configure the Python sailing simu	ılator 13
Run the Python sailing simulator	16
Next Steps	18



Vendée Globe - Business Case

Foreword

The Vendée Globe is a solo non-stop round the world sailboat race founded by Philippe Jeantot in 1989. The race takes place every four years and is considered an extreme quest of individual endurance and the ultimate test in ocean racing.

The 9th edition of the Vendee Globe was held in 2020-2021 and was won by French sailor Yannick Bestaven who took slightly over 80 days to sail non-stop around the globe.

During the race, spectators could follow the action live on an online racing dashboard. The technology for tracking the boats live was provided by Nokia.



Vendée Globe - Business Case

Your Challenge

What will you need to do during this business case? Find out during this section of the guide!

Vendee Globe - Sailing Race

Learn more about your challenge during this business case.

The Challenge

In this business case, you will assume the role of Nokia. You will build a cloud-based Lambda Architecture to process the telemetry data from the sailing boats. Your architecture should run in Azure and contain a real-time path for processing sailing boat data in real-time, and a batch-processing path for collecting sailing boat data in batches and performing calculations on those batches.

Lambda Architecture

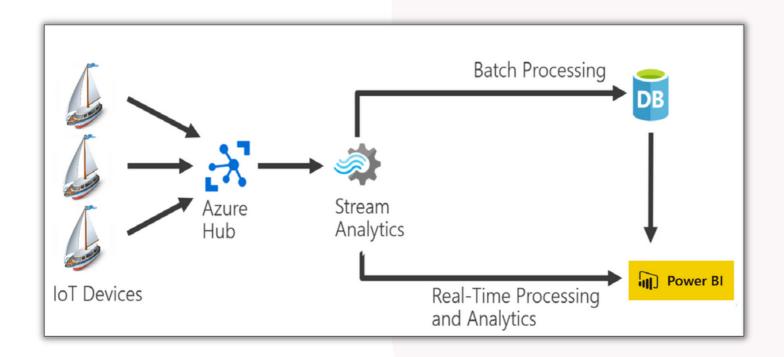
The Lambda Architecture should send all boat data to a PowerBI dashboard. The dashboard should display a world map with the current position of each boat, and a table with a ranking of racing teams. The ranking table should be sorted by who is currently in the lead.

We are currently between races, so unfortunately we cannot use the actual data from the boats participating in the race. Instead, you are going to use a Python application that will simulate boat telemetry data from a fleet of 10 race participants.

To complete the business case...

- Create a Lambda Architecture in Azure with a real-time path and a batch-processing path. Your architecture should include an Event Hub, a Stream Analytics Job, and an output to PowerBl. Your batch-processing path can use any data storage service you prefer: a Data Lake, a SQL Database, a Cosmos database, or a Synapse Analytics Workspace.
- Create a PowerBI dashboard that displays a world map with the current location of each racing team, and a table with the teams ranked by position in the race.
- Download the Python race simulation app to your local computer. Configure the app to send data to your EventHub.
- Start the Python app. Every 60 seconds, the telemetry of the simulated racing boats will be sent to your Azure cloud.

You will have completed the business case if your PowerBI dashboard correctly shows the position and ranking of each sailing team in the race.





Vendée Globe - Business Case

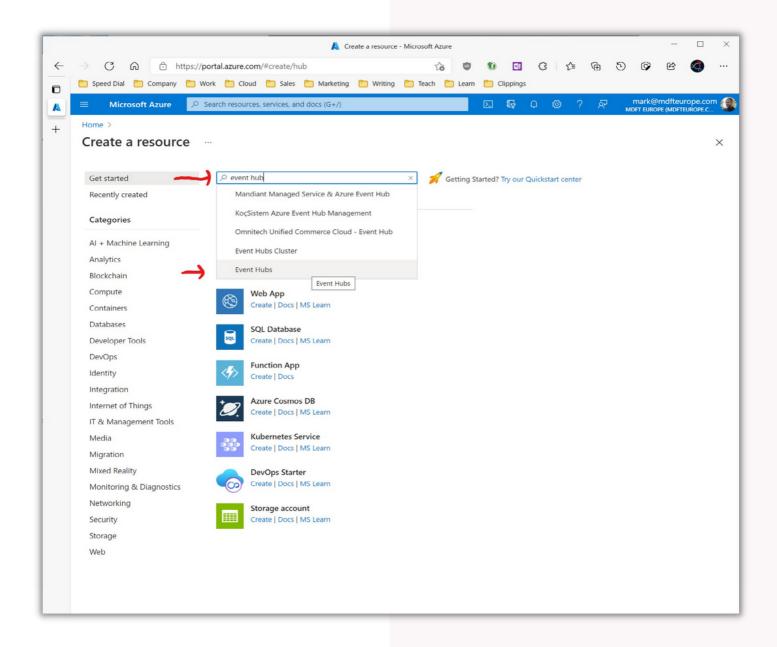
Getting Started

To help you get started, here are guided instructions to build the first piece of the Lambda Architecture and get the Python sailing simulator up and running.

Getting Started with the Vendée Globe Business Case

Step One:

The first thing you need to do is create an Event Hub Namespace in your Azure Tenant:



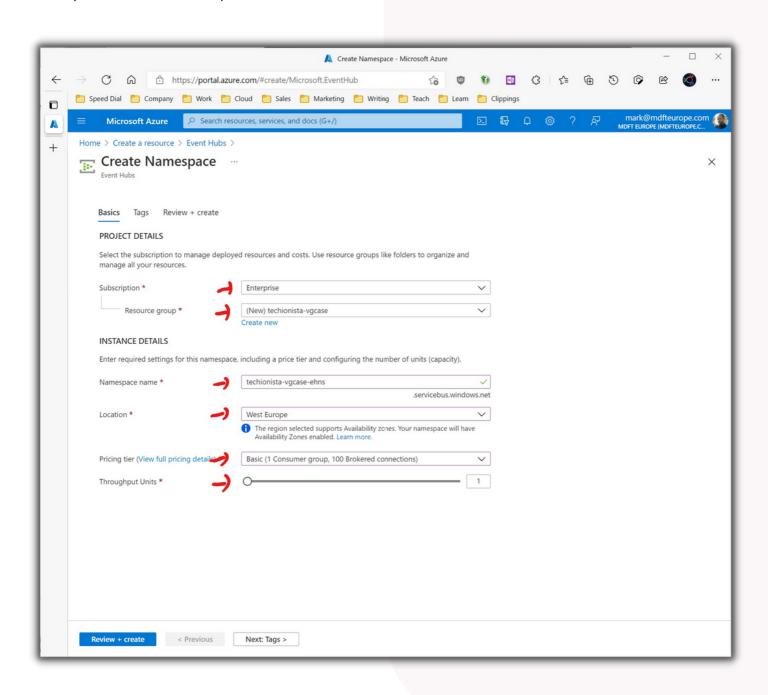
Getting Started with the Vendée Globe Business Case

Step Two:

Then configure the Event Hub Namespace as follows:

- Subscription: select your Azure Pass Sponsorship
- Resource group: create a new resource group to hold all the resources of your Lambda Architecture.
- Namespace name: choose a name for your namespace
- Location: select the region that's closest to your current location. This will be 'West Europe' if you're in the Netherlands.
- Pricing tier: set this to 'Basic'.
- Throughput Units: set this to 1.

Click the blue 'Review + Create' button at the bottom of the page, and then click the blue 'Create' button to create your Event Hub Namespace.

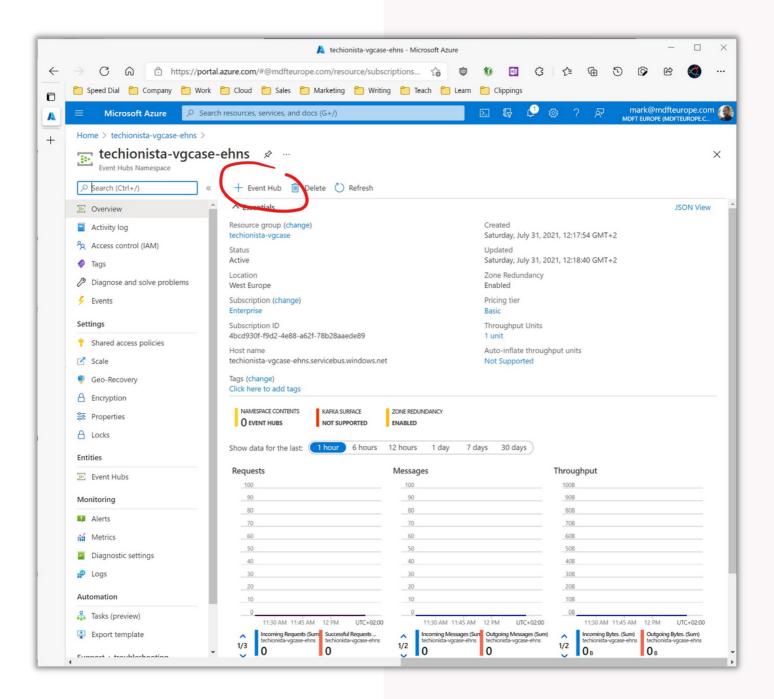


Getting Started with the Vendée Globe Business Case

Step Three:

Once the Event Hub Namespace has been created, click on the blue 'Go to resource' button to access the Event Hub Namespace overview page.

From there, click on the '+ Event Hub' button to add a new Event Hub to the Event Hub Namespace:



Getting Started with the Vendée Globe Business Case

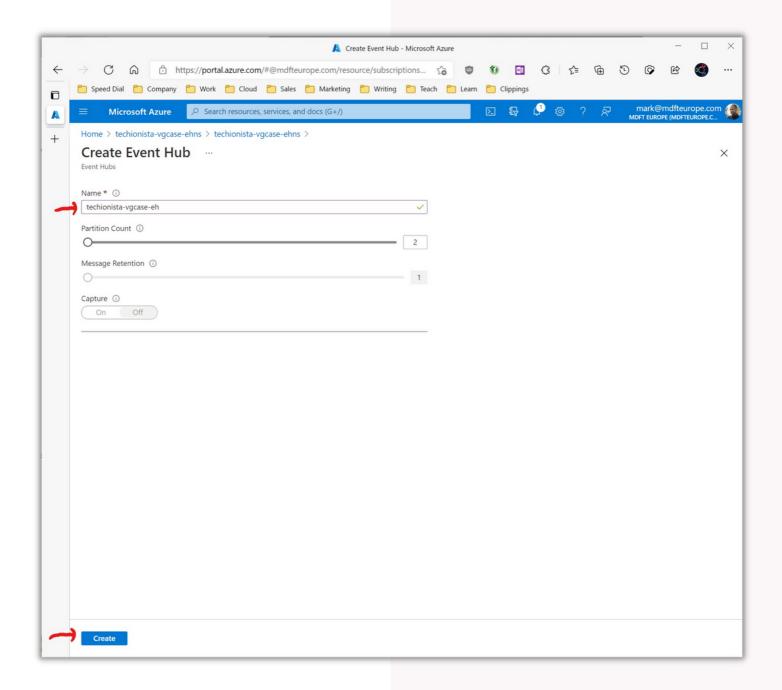
Step Four:

Configure the Event Hub as follows:

• Name: choose a name for your Event Hub.

Store the Event Hub name in notepad. You will need it in an upcoming installation step.

You can leave all other fields at their default value. Click the blue 'Create' button at the bottom of the page to create the Event Hub.



How to get the Event Hub Namespace Connection String

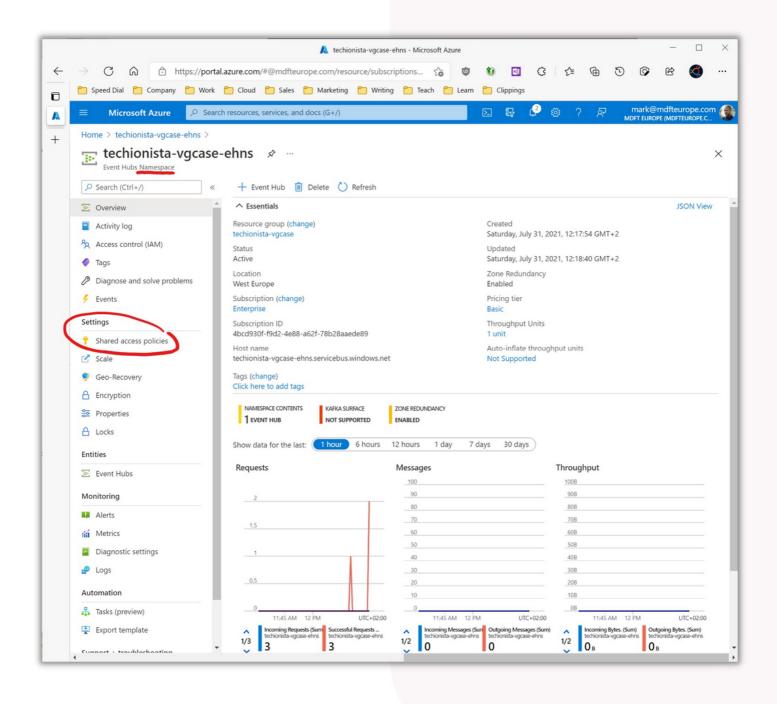
Getting Started with the Vendée Globe Business Case

Step Five:

Next, you will need to create a Shared Access Policy that will allow the Python sailing simulator to send data to this Event Hub Namespace.

The Shared Access Policy will contain a connection string that the Python code will use to connect to the Event Hub and upload data to the Azure cloud.

First, make sure that you are on the Event Hub Namespace overview page, and then click the Shared Access Policies menu item.



How to get the Event Hub Namespace Connection String

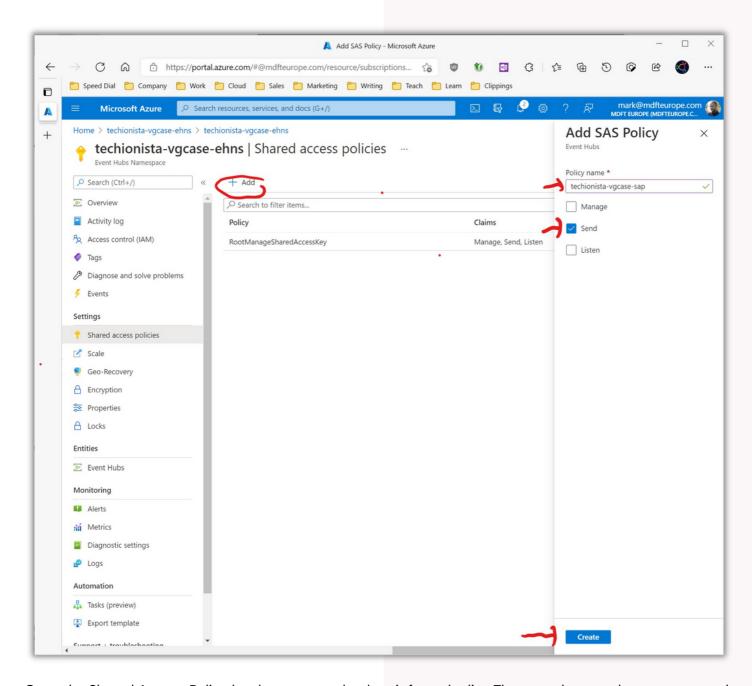
Getting Started with the Vendée Globe Business Case

Step Six:

Now click the 'Add' button to add a new Shared Access Policy and provide the following information:

- Name: choose a name for your Shared Access Policy
- Send: check this box to allow sending new data to the Event Hub Namespace.

Leave the other options unchecked, and click the blue 'Create' button at the bottom of the page.



Once the Shared Access Policy has been created, select it from the list. Then use the copy button to copy the 'Connection String-Primary Key' value to your clipboard.

Store the connection string in notepad. You will need it in an upcoming installation step.

Download and configure the Python Sailing Simulator

Getting Started with the Vendée Globe Business Case

Step Seven:

We do not have access to the real boat telemetry data from the Vendée Globe participants, so instead, we are going to use a Python application that will simulate boat telemetry data from a fleet of 10 race participants.

First, make sure that <u>Python</u> is installed on your computer. You can use either Windows, OS/X or Linux to run the sailing simulator.

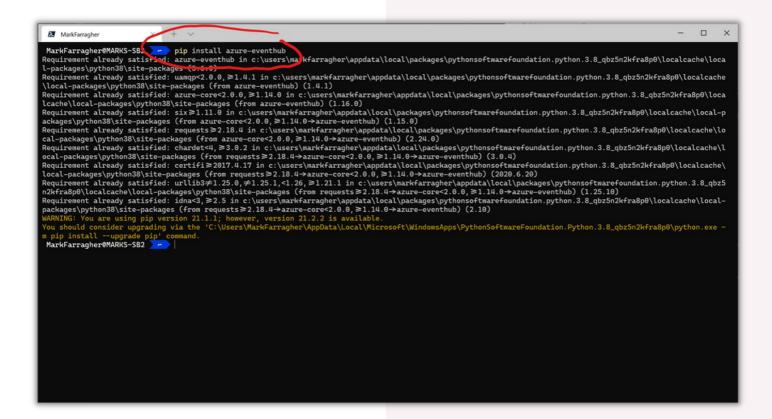
Now download the <u>Python Sailing Simulator</u> and store it somewhere on your computer. We suggest you save the file in your Documents folder. Make sure you save the file as *race_simulator.py*.

Next, open a Powershell window (or a Terminal window if you're using OS/X or Linux) and run the following command:

cd ~/Documents

This will change the current directory to the Documents folder where you have saved the racing simulator. Now you need to install the Python library for Azure Event Hubs. Run the following command: pip install azure-eventhub

This will install the Azure Event Hub library for Python on your computer.



The racing simulator is just a single file named *race_simulator.py*. Open this file with your favorite text editor and take a look at lines 33 and 34.

Download and configure the Python Sailing Simulator

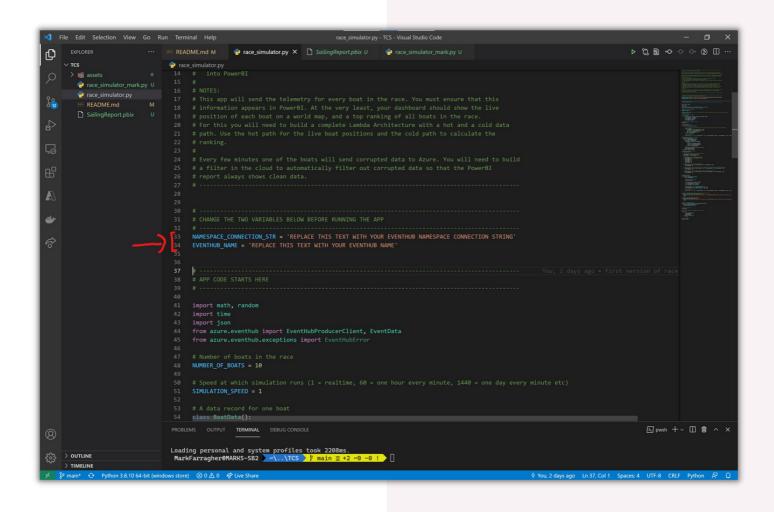
Getting Started with the Vendée Globe Business Case

Step Eight:

You must do the following:

- Replace the value of line 33 with the Connection String-Primary Key of your Event Hub Namespace. You saved this value in notepad earlier.
- Replace the value of line 34 with the name of your Event Hub. You saved this value in notepad earlier.

After you have made these replacements, your file should now look like this (keep in mind that the screenshot shows my specific configuration values, your values will be different):

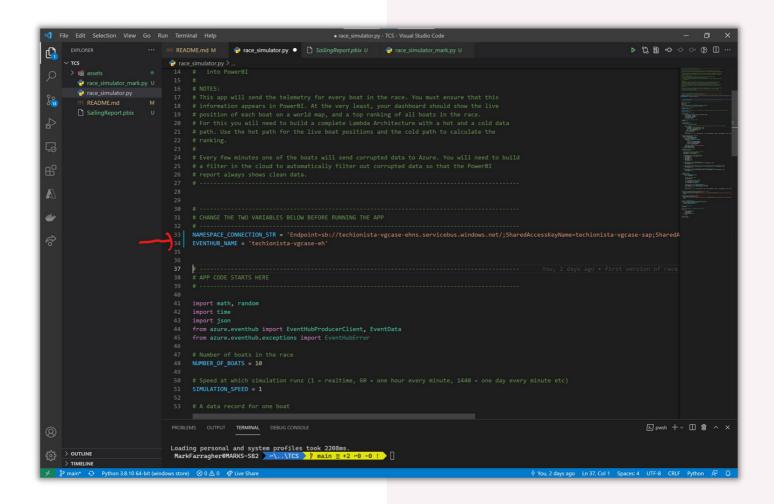


Download and configure the Python Sailing Simulator

Getting Started with the Vendée Globe Business Case

Step Nine:

Save the file. You are now ready to run the simulator.



Run the Python Sailing Simulator

Getting Started with the Vendée Globe Business Case

Step Ten:

Open a command-line window and run the following command: python.\race_simulator.py

This will run the Python sailing simulator. The app will simulate a fleet of 10 sailing boats that start the race near the port of Cascais in Portugal and proceed south along the coast of Africa and then east towards Australia.

While the app is running, it will display simluated boat telemetry data every 60 seconds:

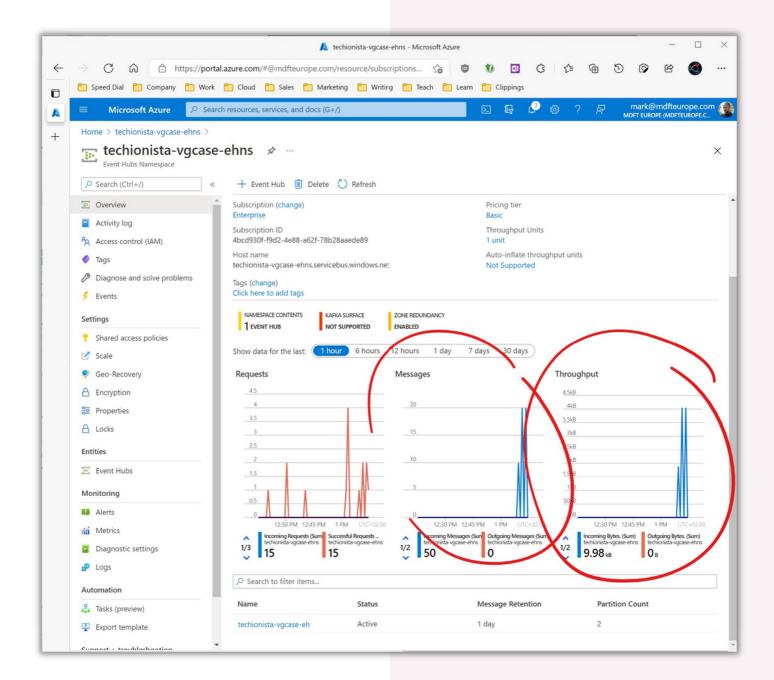
The race will continue for as long as you keep the app running. If you stop and restart the app, all sailing boats will reset back to their starting position in Portugal.

Run the Python Sailing Simulator

Getting Started with the Vendée Globe Business Case

Step Eleven:

Keep the app running and check your Event Hub Namespace overview page. After approximately 5 minutes, the first data should appear in the monitoring graphs:



Congratulations! Your app is now working correctly and sending simulated race data to your Azure cloud.

Next Steps

Getting Started with the Vendée Globe Business Case

Next Steps

To complete the business case, here's what you need to add:

- The remaining pieces of the Lambda Architecture: a Stream Analytics Job, an output to PowerBl, and a second output to a data storage service of your choice.
- A data storage service of your choice to store batch data: a Data Lake, a SQL Database, a Cosmos database, or a Synapse Analytics Workspace.
- A view or stored procedure or Synapse notebook to calculate the table of teams ranked by position in the race.
- A PowerBI dashboard that displays a world map with the current location of each racing team, and a table with the teams ranked by position in the race.

You will have completed the business case when your PowerBI dashboard correctly shows the position and ranking of each sailing team in the race.

Challenges

During the business case, you will need to address several challenges:

- The Python app occasionally produces garbled data. How are you going to ensure that only clean data arrives in the PowerBI dashboard?
- How are you going to calculate the ranked list of sailing teams? How will you calculate who is ahead in the race?
- Which data service are you going to use for the batch-processing path in the Lambda Architecture?
- How will you present your data in the PowerBl dashboard?

There are multiple solutions for each challenge.