Making MATLAB Data Analytics Accessible Across The Enterprise

July 2\textsuperscript{nd} 2020 | Online

MathWorks AUTOMOTIVE CONFERENCE 2020
Key takeaways

- Data for decision making should be available throughout the Enterprise
  - Not only engineers need data – decision makers of all sorts need data

- Large volumes of data require new paradigms
  - Scaling out is necessary

- The diversity of systems require interoperability of tools
  - MathWorks integrates with many different frameworks
Automotive use case

Automotive

• Vehicles
• Engines
• Controllers

[Graph showing vehicle speed over time]

[Map showing a route]
Motivating factors for enterprise accessibility of Data Analytics

- Not all users of data are engineers
- Reuse among engineers
Processing data

- Exploratory work on a desktop
- Big Data cannot be processed on a single machine
Processing data

**Integrate** with a broad array of databases, data stores and streaming services

Data sources / applications

Streaming data

MATLAB Production Server

Worker processes

Request Broker

**Scale** to service hundreds of concurrent requests while **Securing** access

Access from a wide range of enterprise applications

Enterprise Application

Mobile / Web Application

Power BI

MATLAB

Integrate with a broad array of databases, data stores and streaming services

MATLAB Production Server

Worker processes

Request Broker

**Scale** to service hundreds of concurrent requests while **Securing** access

Access from a wide range of enterprise applications

Enterprise Application

Mobile / Web Application

Power BI

MATLAB
MathWorks provides a comprehensive end-to-end solution Data Analytics in the Enterprise.
Live Demonstration
MAC Europe 2020 - Demo - Making MATLAB Data Analytics Accessible Across Enterprise

This demo shows 3 different workflows with MATLAB and Databricks:

1. Communicating with Databricks
2. Data Exploration using Databricks-Connect and Spark
3. Data Engineering on Databricks

Part 1: Communicating with Databricks

We can easily work with data on Databricks directly from the MATLAB desktop

```matlab
db = databricks.DBSFS();
files = db.ls('/public/fleetdata/files')
```

We can directly download a file, to work with it in MATLAB

```matlab
db.download('/public/fleetdata/files/data_01.parquet')
```

and perform other filesystem tasks

```matlab
ans = 'C:\\E:\Platforms\databricks\Software\MATLAB\public\examples\MAC2020\data_01.parquet'
% Create a directory
db.mkdir('public/data')
% Remove a file
db.rm('/public/fleetdata/file_01.parquet')
% Upload a file
db.upload('trip_01.mat', '/public/fleetdata/mat')
```

Part 2: Exploration using Databricks-Connect and Spark

MATLAB users can use databricks-connect to connect to the Spark cluster to explore datasets. The tooling allows users to slice / dice and query data by leveraging the distributed compute cluster that works with the storage to expose the datasets via a Spark API.
Key takeaways

• Data for decision making should be available throughout the Enterprise
  • Not only engineers need data – decision makers of all sorts need data

• Large volumes of data require new paradigms
  • Scaling out is necessary

• The diversity of systems require interoperability of tools
  • MathWorks integrates with many different frameworks
Please contact us with questions

asolland@mathworks.com