Making MATLAB Data Analytics Accessible Across The Enterprise

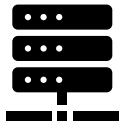
July 2nd 2020 | Online

MathWorks
AUTOMOTIVE
CONFERENCE 2020

MathWorks®

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

Compute  People 

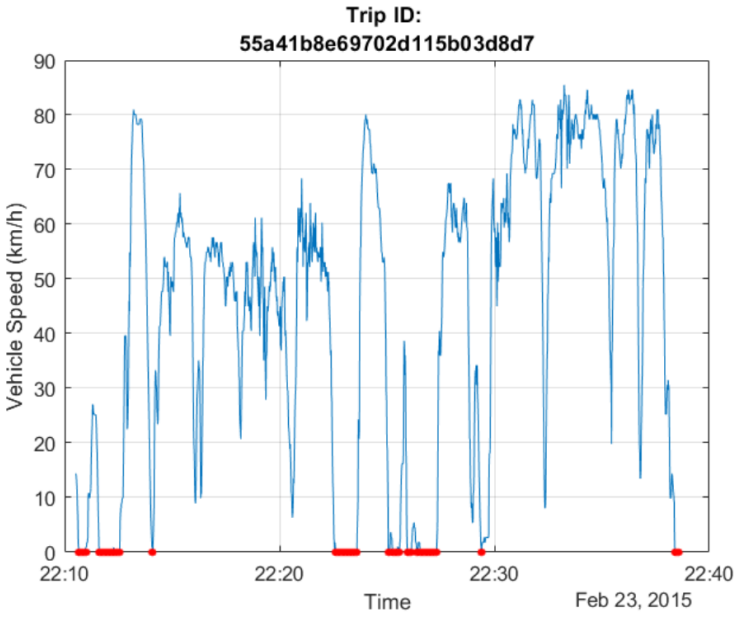
- The diversity of systems require interoperability of tools
  - MathWorks integrates with many different frameworks

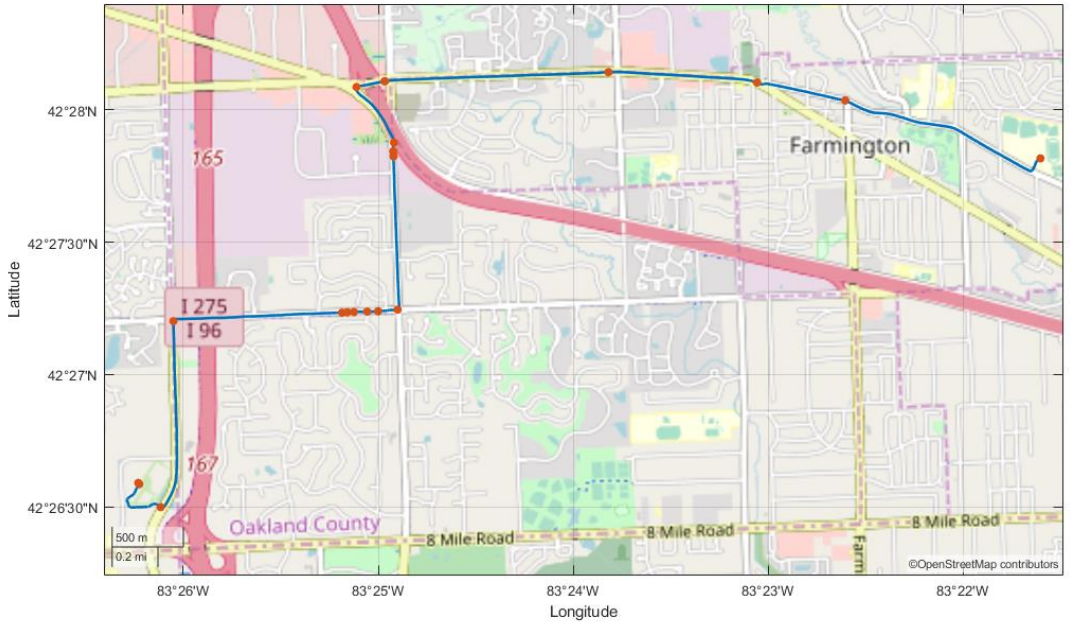
Automotive use case

Access and Explore Data → Preprocess Data → Analyze Data → Develop Algorithms & Models → Operationalize Models

Automotive

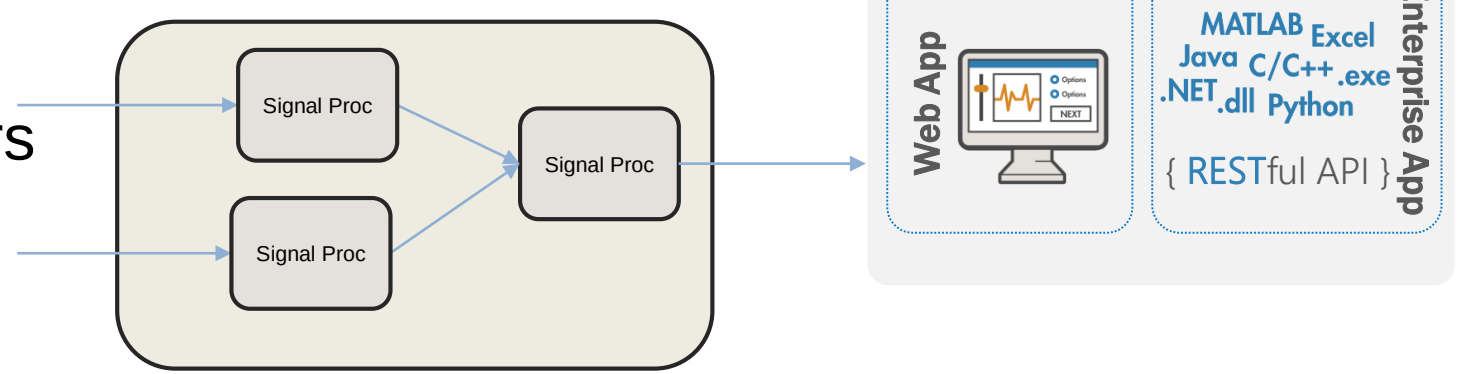
- Vehicles
- Engines
- Controllers



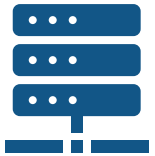


Motivating factors for enterprise accessibility of Data Analytics

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

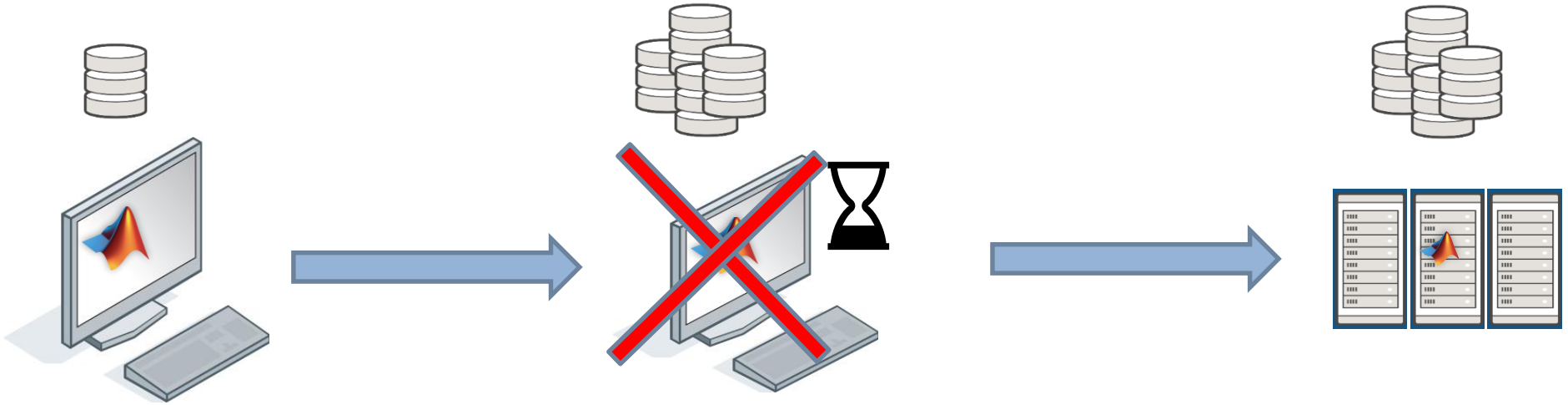


Processing data



Access and Explore Data | Preprocess Data | Analyze Data | Develop Algorithms & Models | Operationalize Models

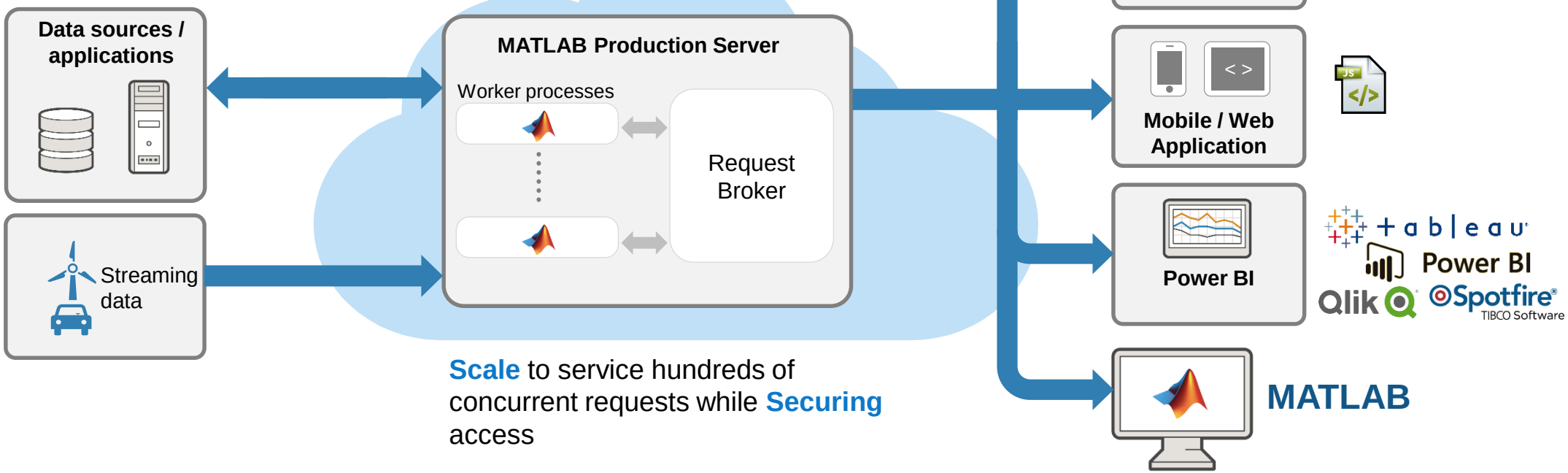
- Exploratory work on a desktop
- Big Data cannot be processed on a single machine



Processing data

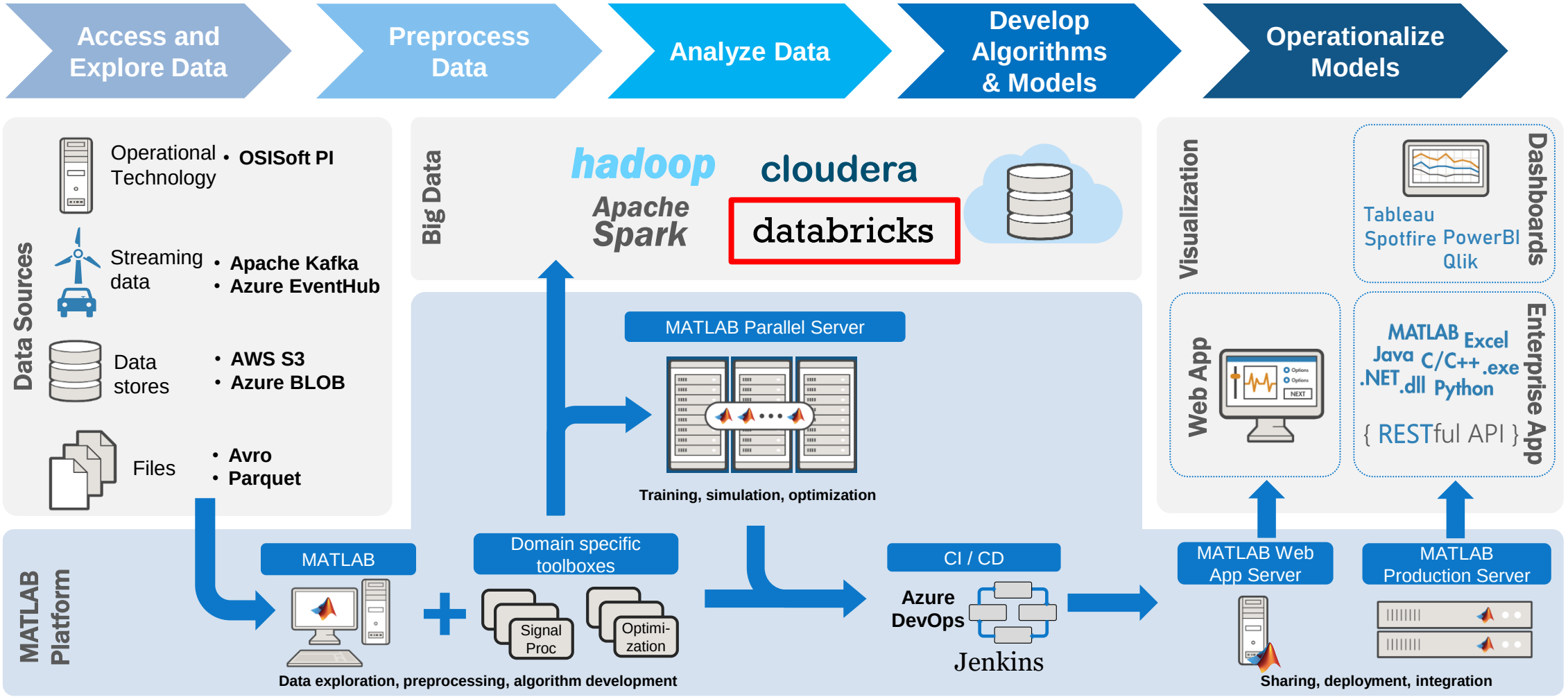
Integrate with a broad array of databases, data stores and streaming services

Access from a wide range of enterprise applications

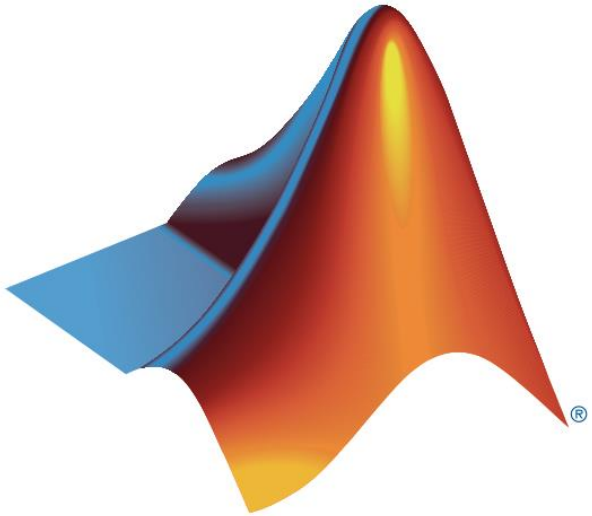


Scale to service hundreds of concurrent requests while **Securing** access

MathWorks provides a comprehensive end-to-end solution Data Analytics in the Enterprise



Live Demonstration



databricks

MATLAB R2020a

HOME PLOTS APPS LIVE EDITOR INSERT VIEW

Search Documentation Anders

New Open Save Find Files Compare Print FILE Go To Find NAVIGATE Text Normal B I U M TEXT Code Task Control Refactor CODE Section Break Run Section Run and Advance Run to End SECTION Run Step Stop RUN

C: ▸ EI ▸ Platforms ▸ databricks ▸ Software ▸ MATLAB ▸ public ▸ examples ▸ MAC2020

Live Editor - C:\EI\Platforms\databricks\Software\MATLAB\public\examples\MAC2020\MAC2020_Demo.mlx

MAC2020_Demo.mlx

MAC Europe 2020 - Demo - Making MATLAB Data Analytics Accessible Across Enterprise

This demo show 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

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

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

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

```
.........
ans = 'C:\EI\Platforms\databricks\Software\MATLAB\public\examples\MAC2020\data_01.parquet'
```

and perform other filesystem tasks

```
% 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.

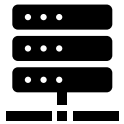
```
inputLocation = '/public/fleetdata/files/*.parquet';
```

UTF-8 script

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

Compute  People 

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

Q&A

Please contact us with questions

asolland@mathworks.com

MathWorks®