MATLAB EXPO 2017

Scaling MATLAB

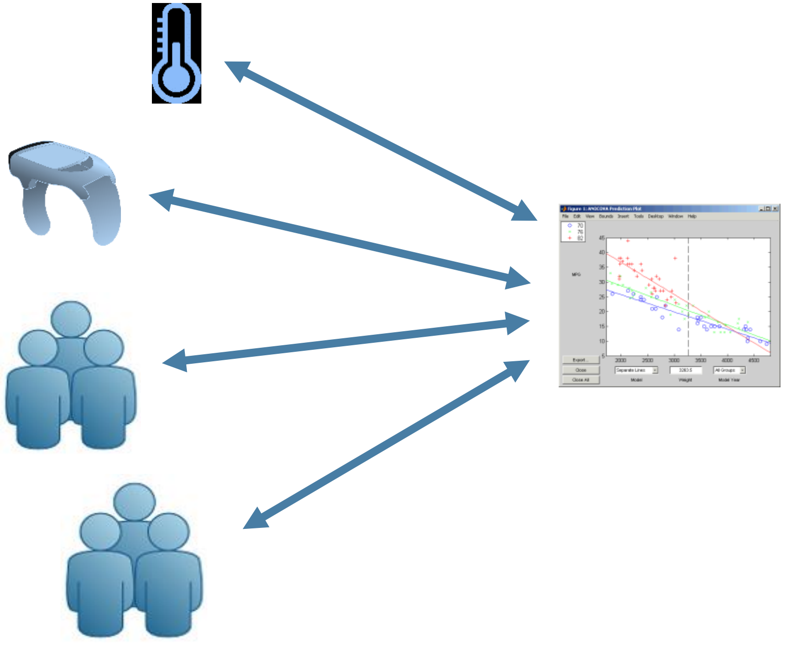
for Your Organisation and Beyond

Rory Adams

MATLAB at Scale

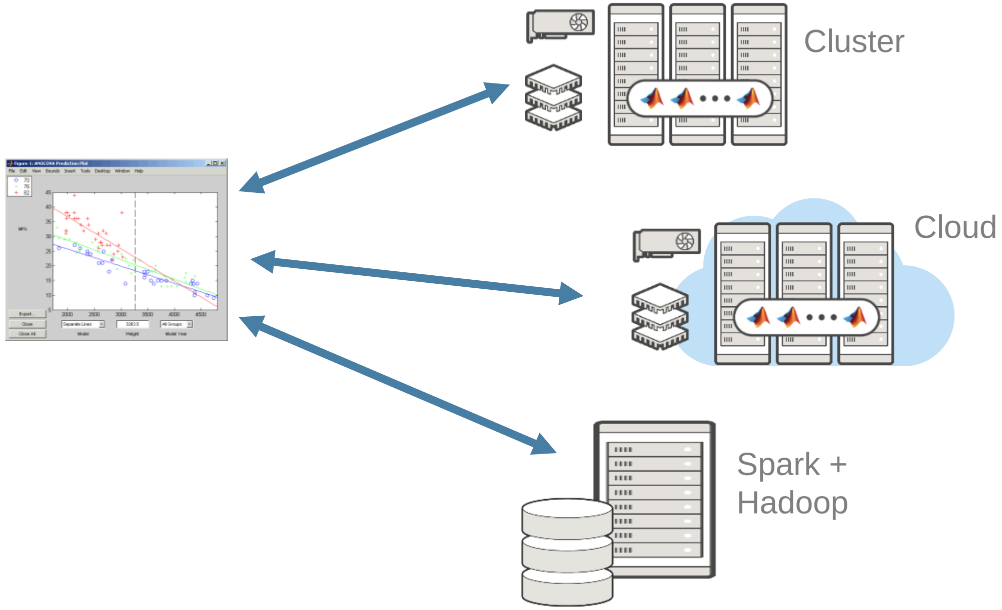
Front-end scaling

- Scale with increasing access requests



Back-end scaling

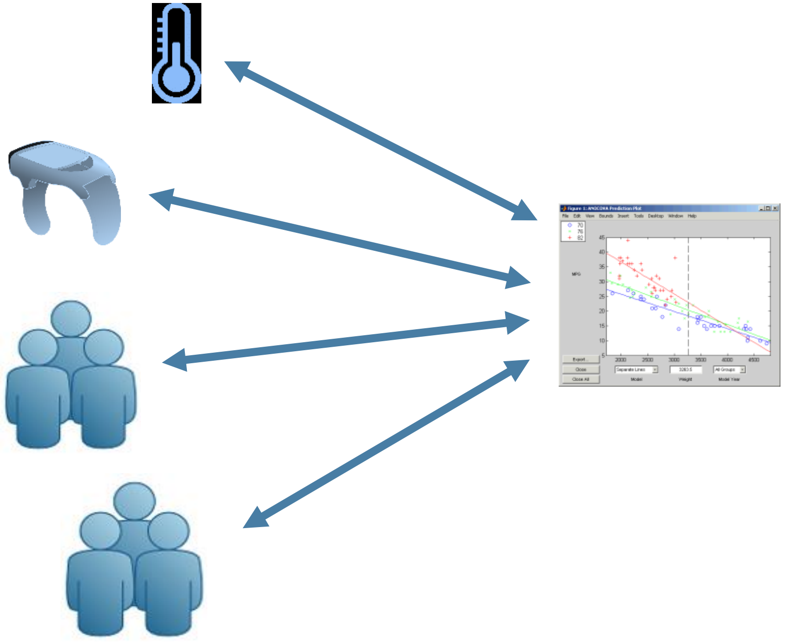
- Scale with increasing computational intensity
- Scale with increasing data volumes



MATLAB at Scale

Front-end scaling

- Scale with increasing access requests



Back-end scaling

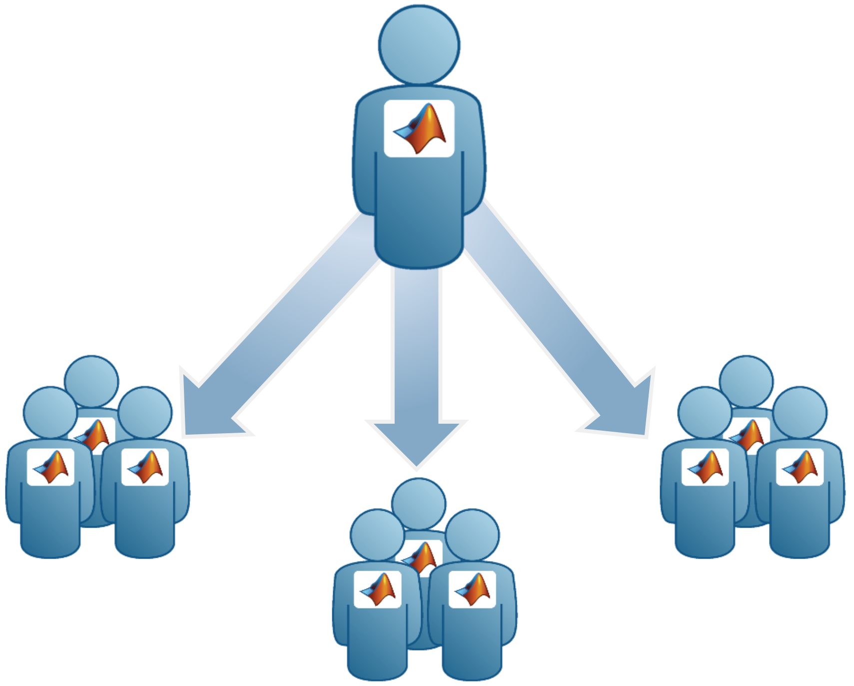
- Scale with increasing computational intensity
- Scale with increasing data volumes

Key Takeaways

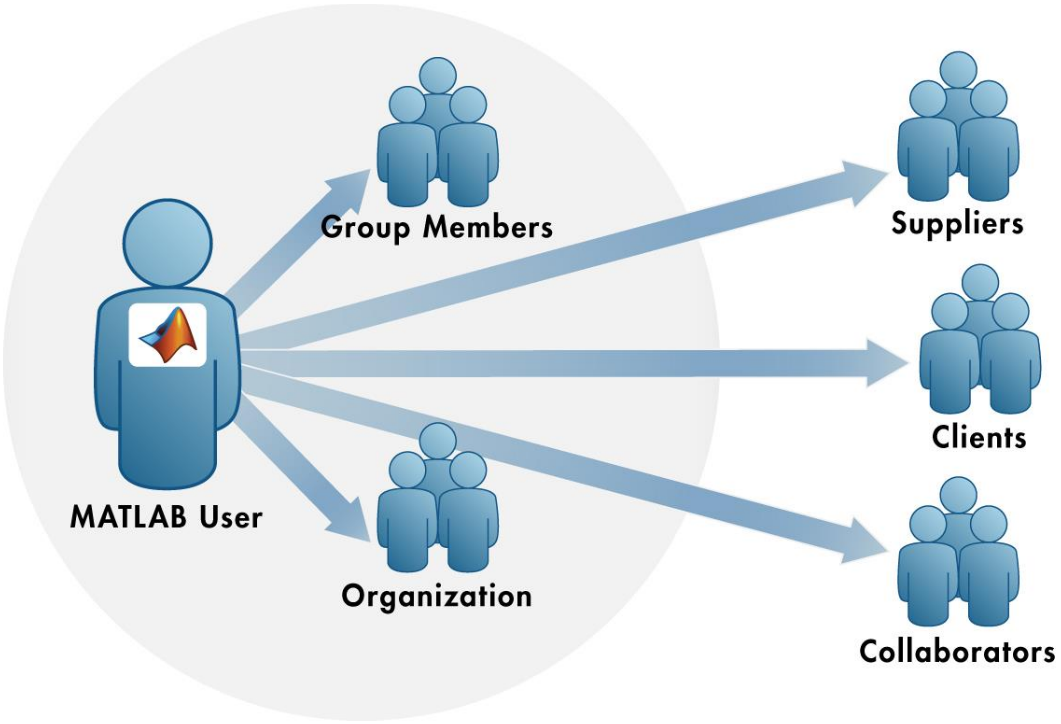
1. Share applications and algorithms with anyone
2. Integrate MATLAB functions into existing workflows and development platforms.
3. Deploy MATLAB applications to service simultaneous requests via web or cloud frameworks.

MATLAB Programs Can be Shared With Anyone

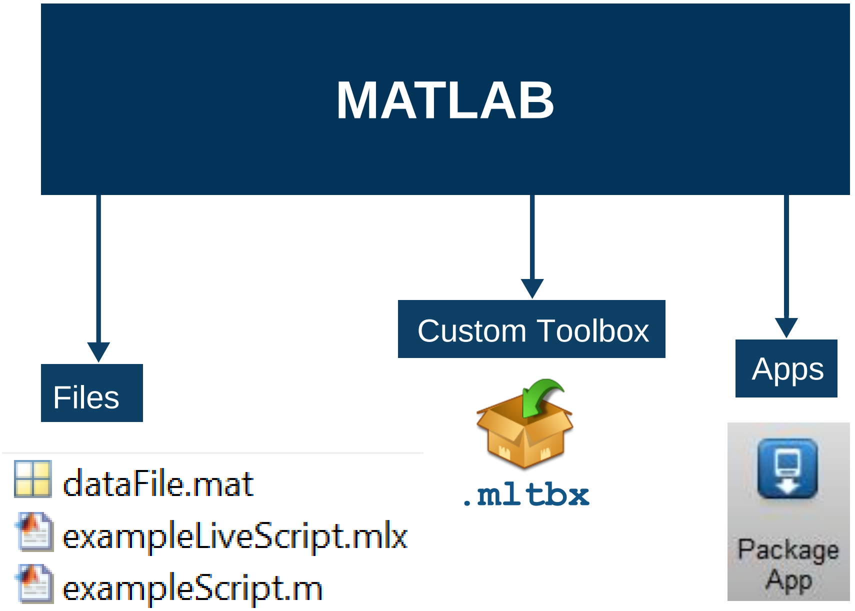
Share With Other MATLAB Users



Share With People Who do Not Have MATLAB

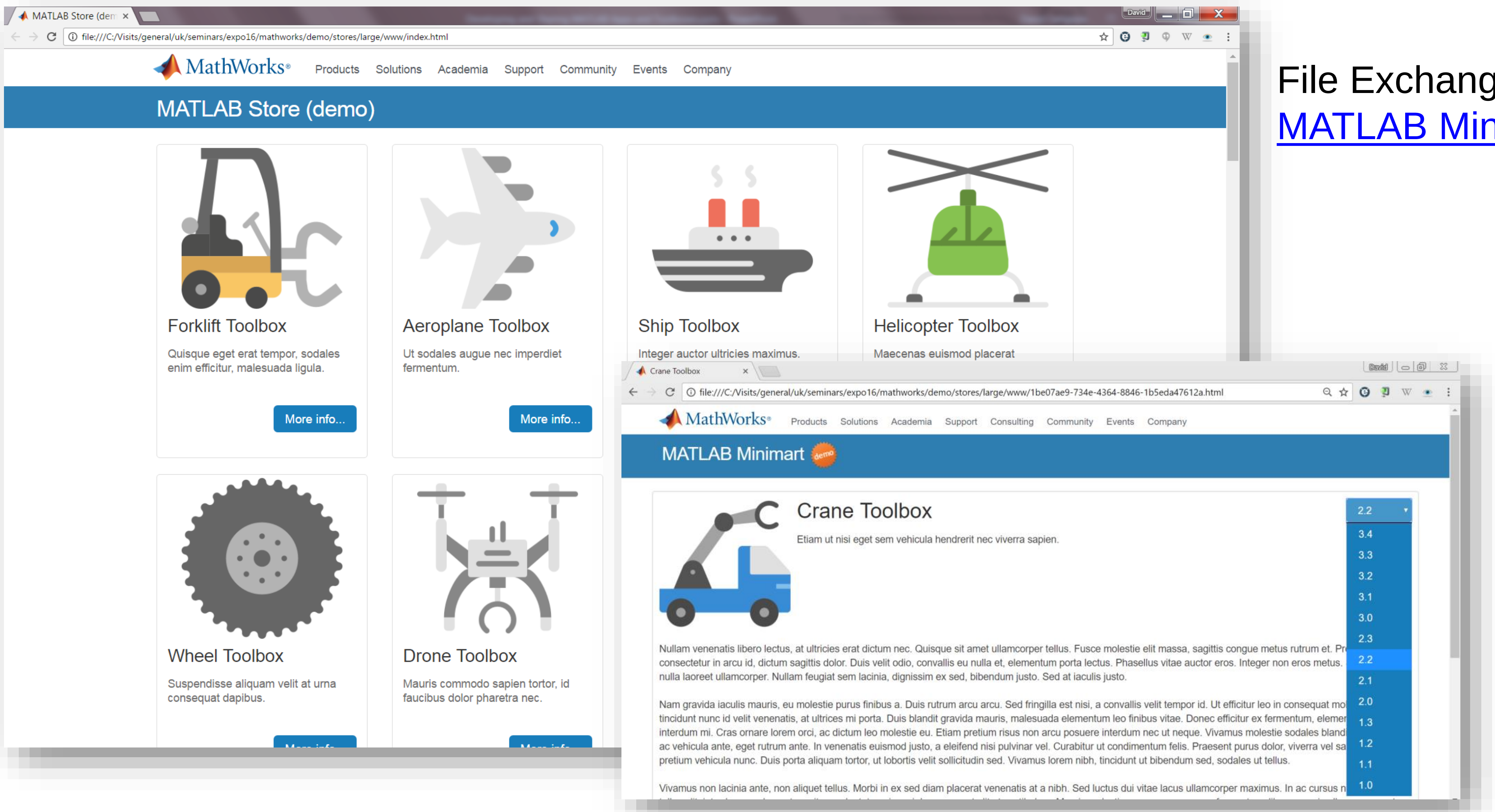


Share with MATLAB Users



- Directly share MATLAB files
- Package an App
- Package Entire Toolboxes

Scale Up Sharing with MATLAB Users



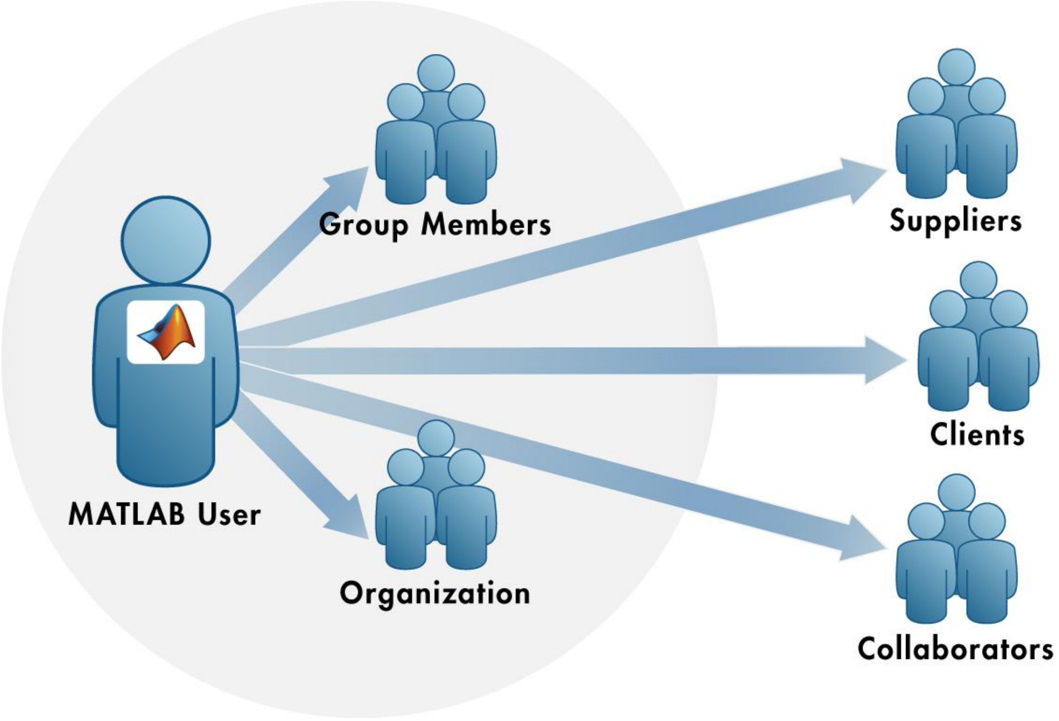
File Exchange:
[MATLAB Minimart]

Icons from [www.flaticon.com]

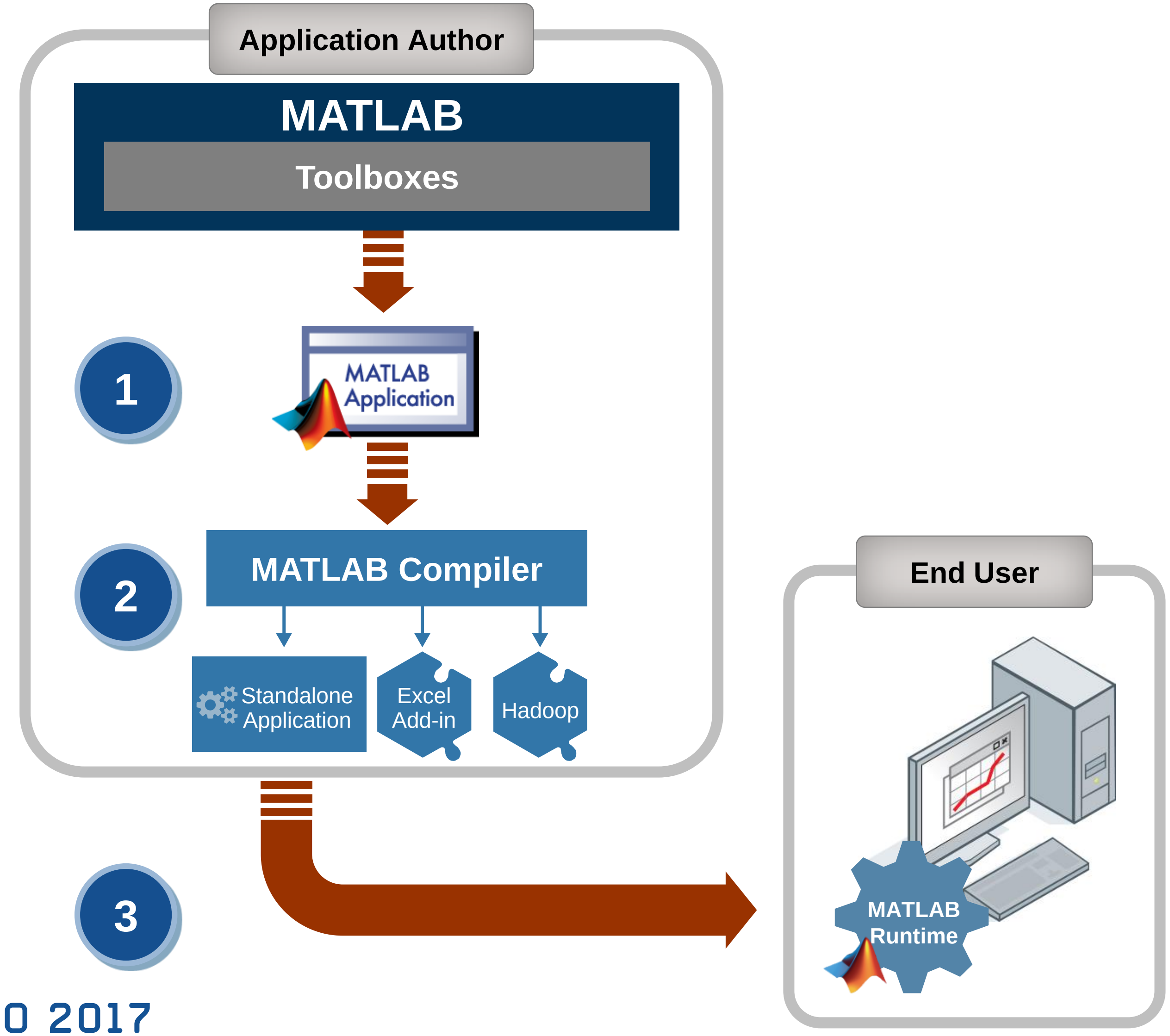
MATLAB Programs Can be Shared With Anyone

Share With Other MATLAB Users

Share With People Who do Not Have MATLAB

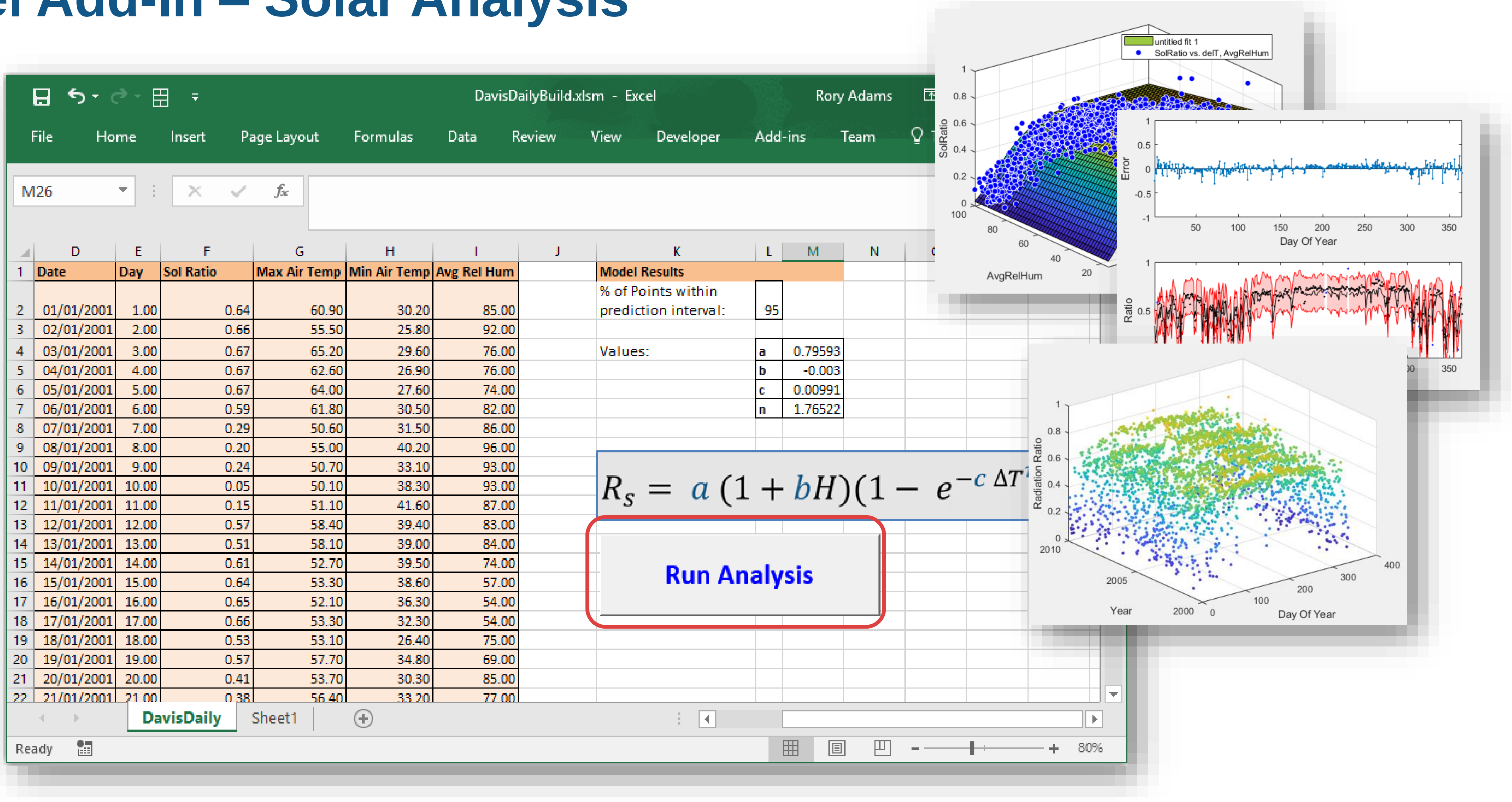


Share Applications Built Completely in MATLAB

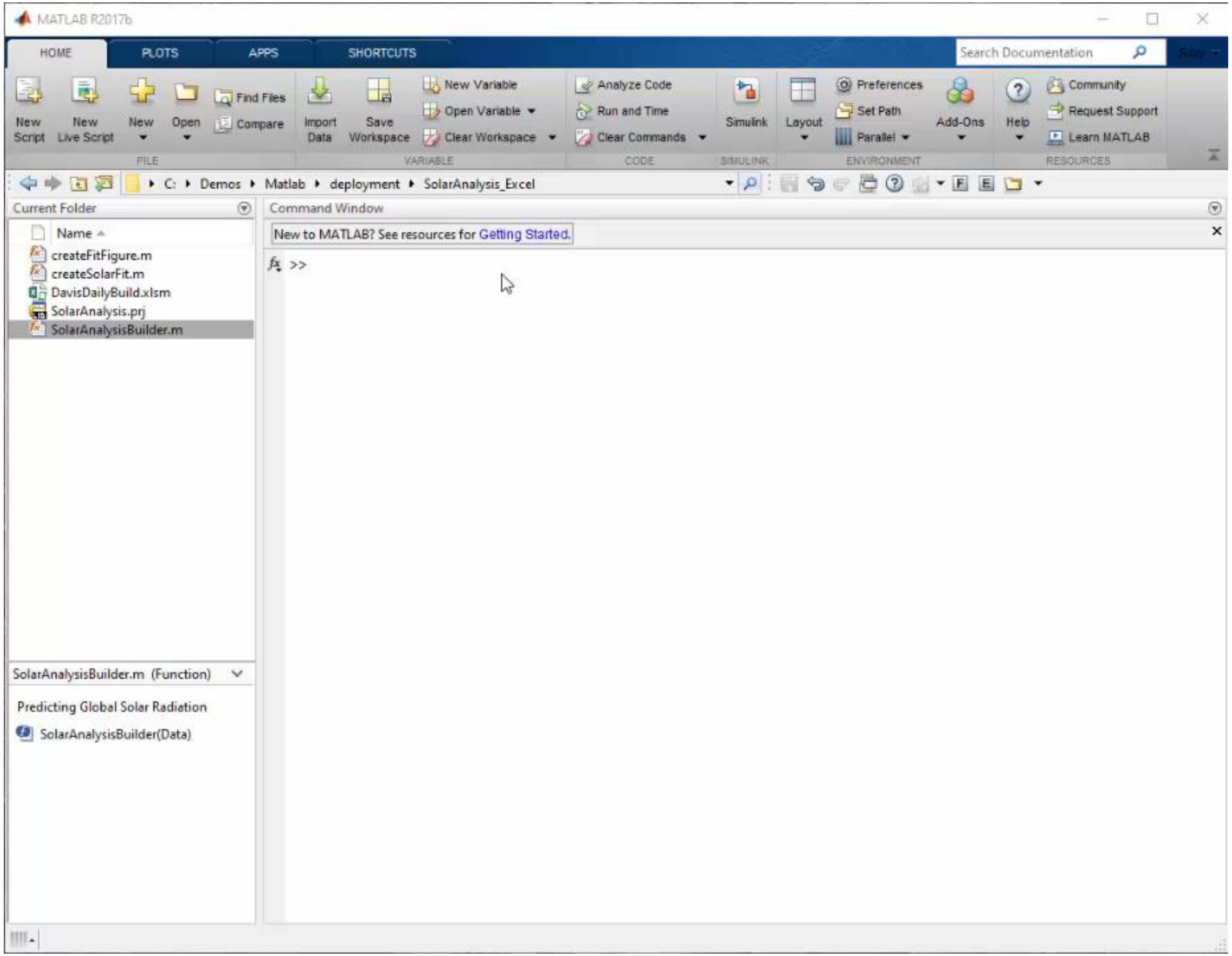


- Royalty-free Sharing
- IP Protection via Encryption

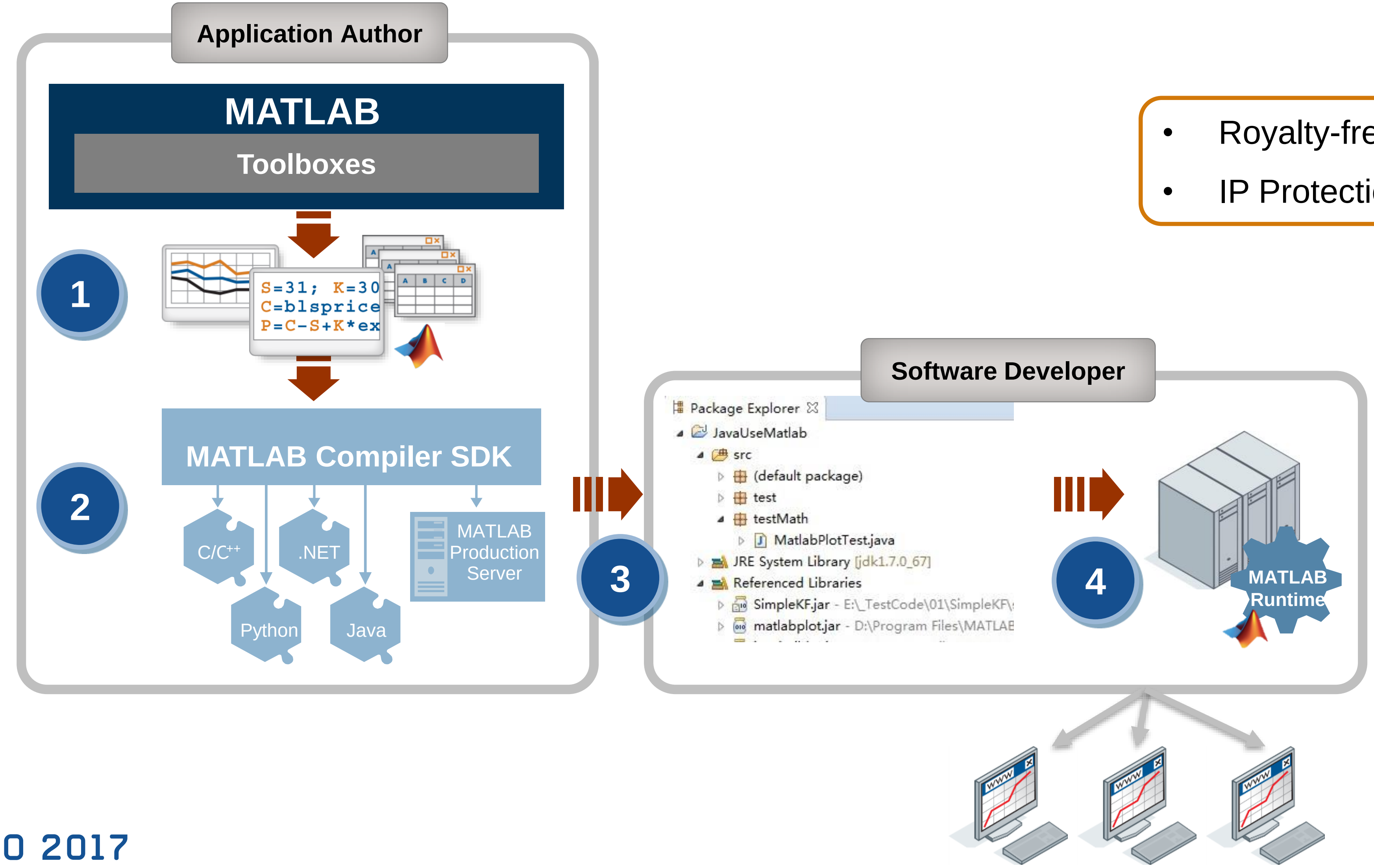
Excel Add-In – Solar Analysis



	D	E	F	G	H	I
1	Date	Day	Sol Ratio	Max Air Temp	Min Air Temp	Avg Rel Hum
2	01/01/2001	1.00	0.64	60.90	30.20	85.00
3	02/01/2001	2.00	0.66	55.50	25.80	92.00
4	03/01/2001	3.00	0.67	65.20	29.60	76.00
5	04/01/2001	4.00	0.67	62.60	26.90	76.00
6	05/01/2001	5.00	0.67	64.00	27.60	74.00
7	06/01/2001	6.00	0.59	61.80	30.50	82.00
8	07/01/2001	7.00	0.29	50.60	31.50	86.00
9	08/01/2001	8.00	0.20	55.00	40.20	96.00
10	09/01/2001	9.00	0.24	50.70	33.10	93.00
11	10/01/2001	10.00	0.05	50.10	38.30	93.00
12	11/01/2001	11.00	0.15	51.10	41.60	87.00
13	12/01/2001	12.00	0.57	58.40	39.40	83.00
14	13/01/2001	13.00	0.51	58.10	39.00	84.00
15	14/01/2001	14.00	0.61	52.70	39.50	74.00
16	15/01/2001	15.00	0.64	53.30	38.60	57.00
17	16/01/2001	16.00	0.65	52.10	36.30	54.00
18	17/01/2001	17.00	0.66	53.30	32.30	54.00
19	18/01/2001	18.00	0.53	53.10	26.40	75.00
20	19/01/2001	19.00	0.57	57.70	34.80	69.00
21	20/01/2001	20.00	0.41	53.70	30.30	85.00
22	21/01/2001	21.00	0.38	56.40	33.20	77.00

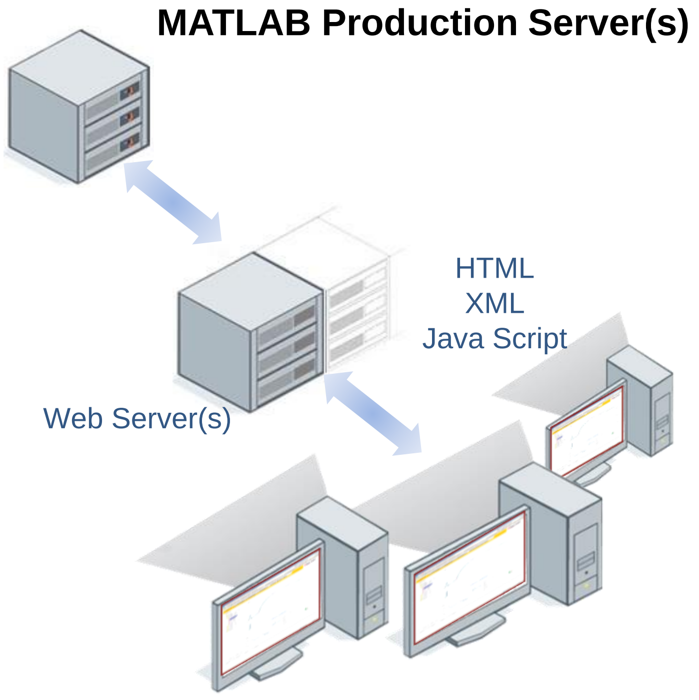

MATLAB R2017b
HOME
PLOTS
APPS
SHORTCUTS
Search Documentation
New Script
New Live Script
New
Open
Find Files
Compare
FILE
Import Data
Save Workspace
New Variable
Open Variable
Clear Workspace
VARIABLE
Analyze Code
Run and Time
Clear Commands
CODE
Simulink
SIMULINK
Layout
Preferences
Set Path
Parallel
ENVIRONMENT
Add-Ons
Help
Community
Request Support
Learn MATLAB
RESOURCES
C: Demos Matlab deployment SolarAnalysis_Excel
Current Folder
Name
createFitFigure.m
createSolarFit.m
DavisDailyBuild.xlsm
SolarAnalysis.prj
SolarAnalysisBuilder.m
SolarAnalysisBuilder.m (Function)
Predicting Global Solar Radiation
SolarAnalysisBuilder(Data)
Command Window
New to MATLAB? See resources for Getting Started.
>>

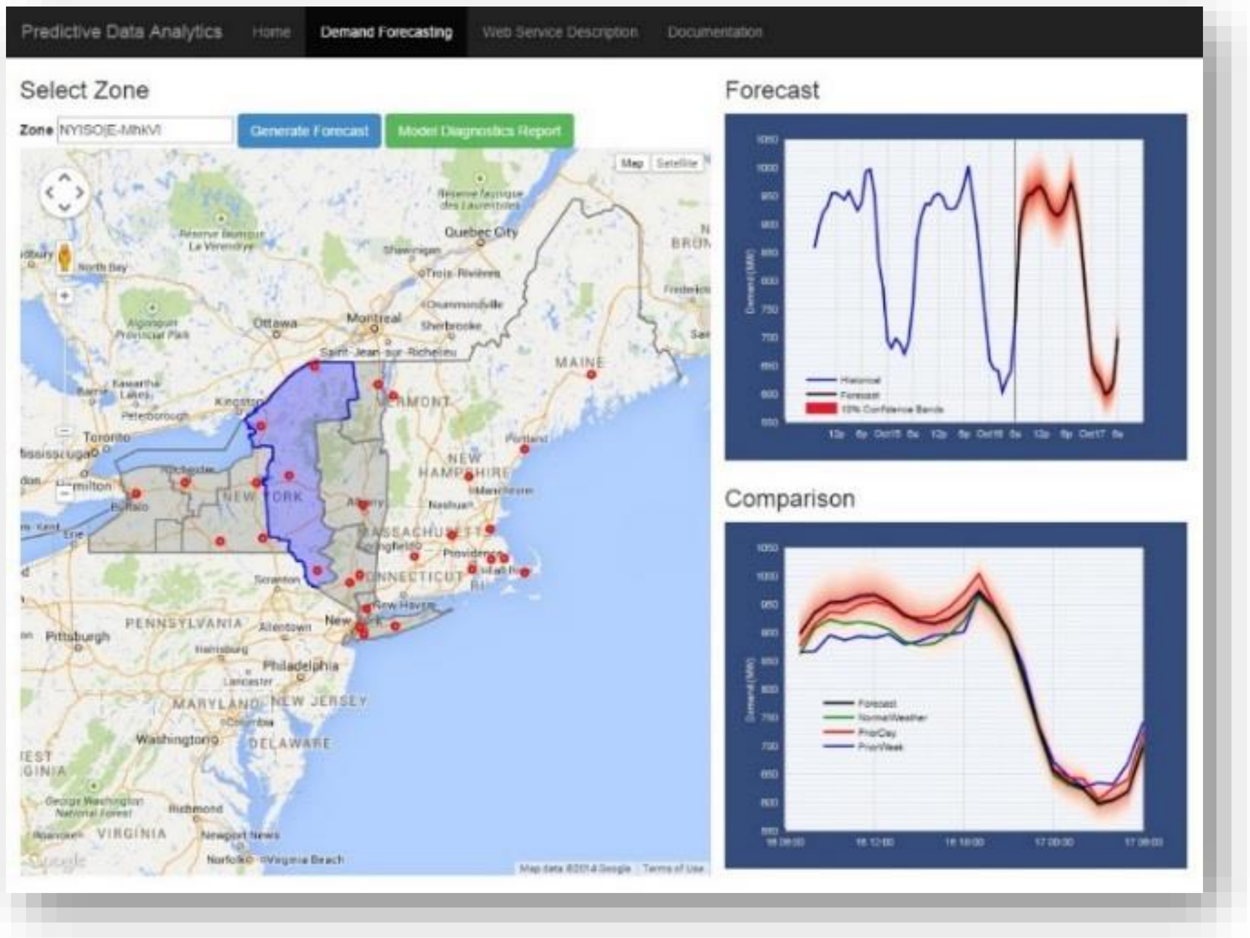
Integrate MATLAB-based Components With Your Own Software



- Royalty-free Sharing
- IP Protection via Encryption

Scaling up: Load Forecasting Demo





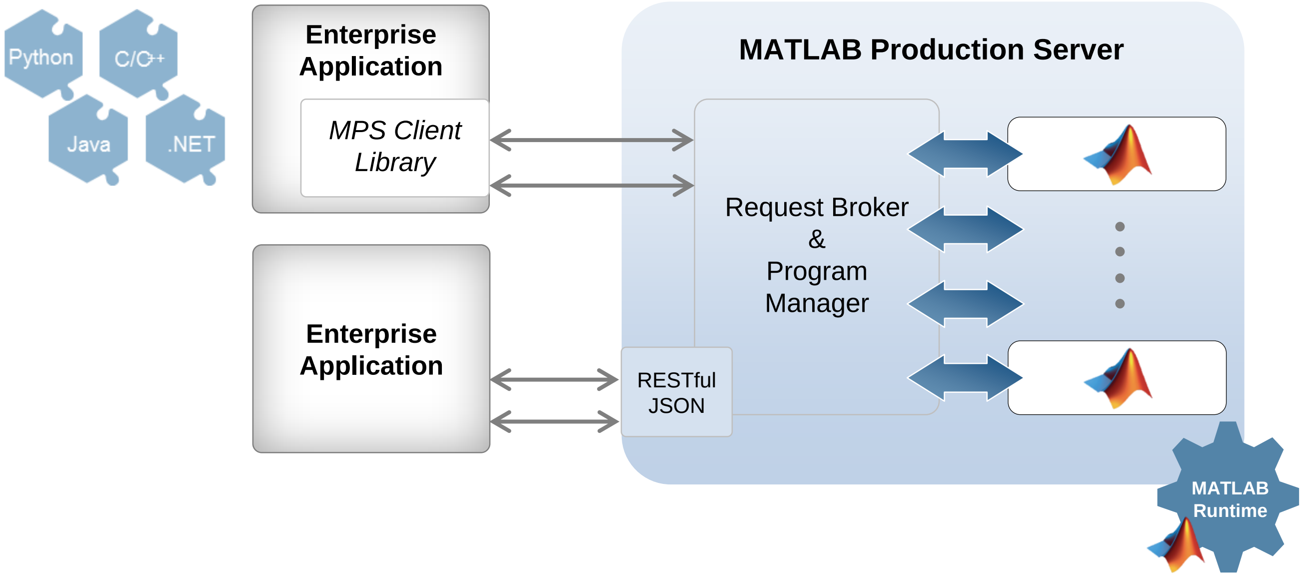
MATLAB and MATLAB Production Server

- The easiest and most productive environment to *take your enterprise analytics* or *IoT solution* from idea to a scalable production solution

MATLAB Production Server

Enterprise Class Framework For Running Packaged MATLAB Programs

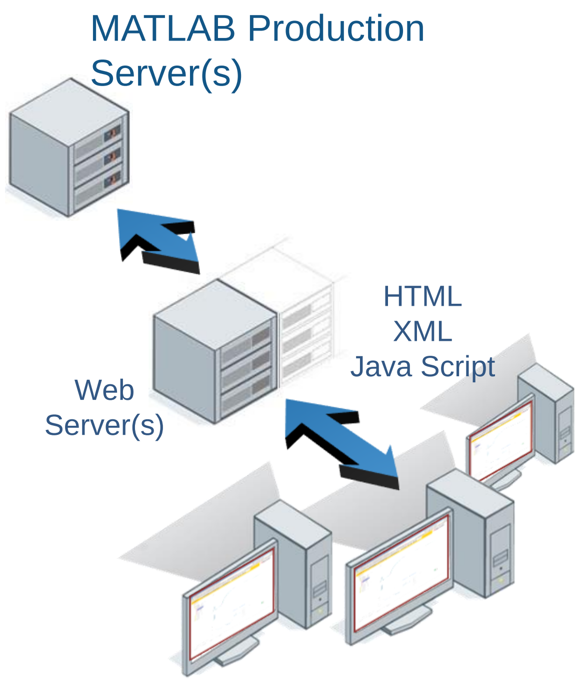
- Server software
  - Manages packaged MATLAB programs and worker pool
- MATLAB Runtime libraries
  - Single server can use runtimes from different releases
- RESTful JSON interface and lightweight client library
  - Isolates the MATLAB processing
  - Access using native data types



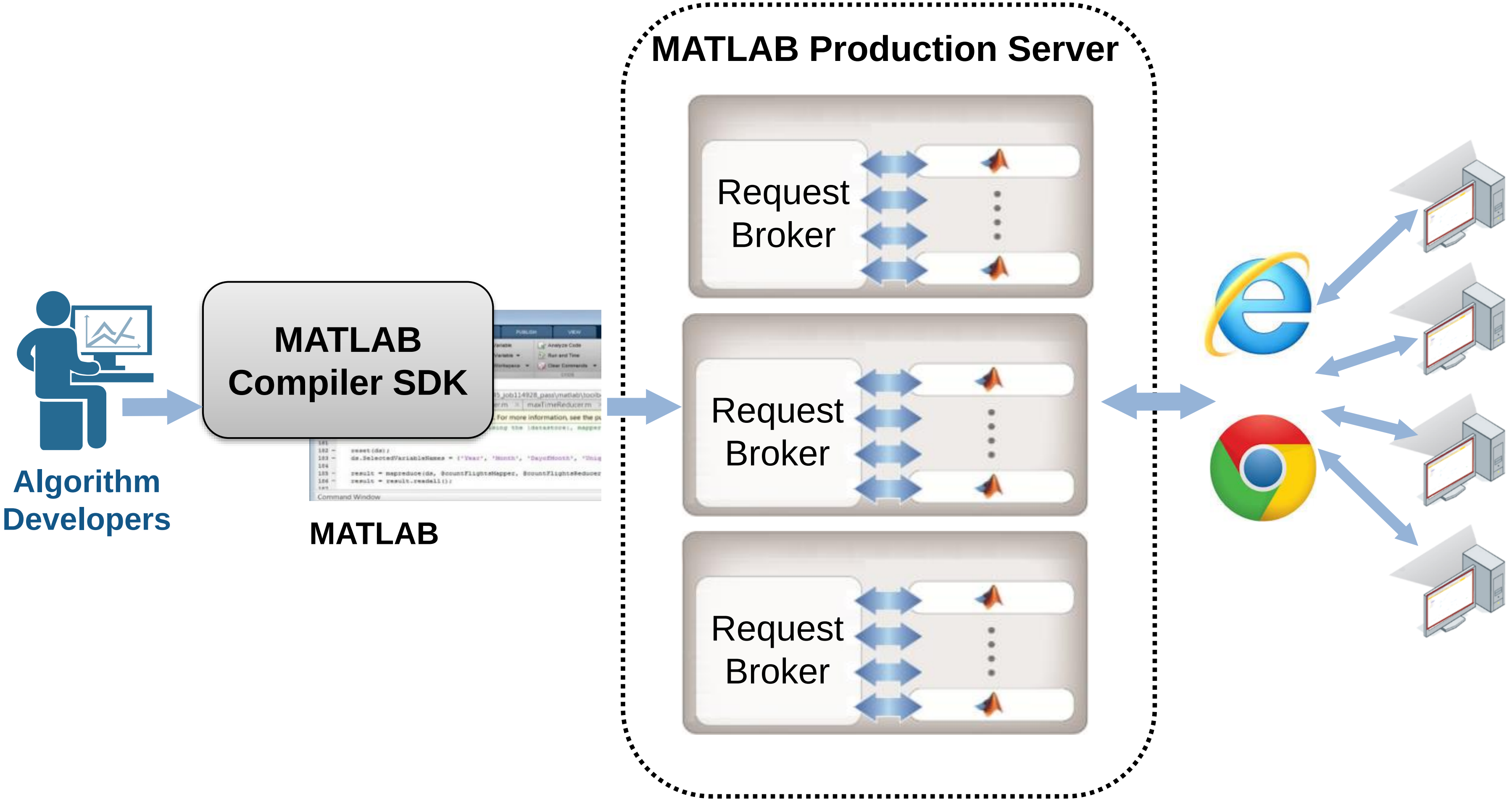
Scale Up with MATLAB Production Server™

- Scalable and reliable
  - Service large numbers of concurrent requests
  - Add capacity or redundancy with additional servers

- Directly deploy MATLAB programs into production
  - Automatically deploy updates without server restarts
  - Most efficient path for creating enterprise applications



Customer examples: Financial customer advisory service



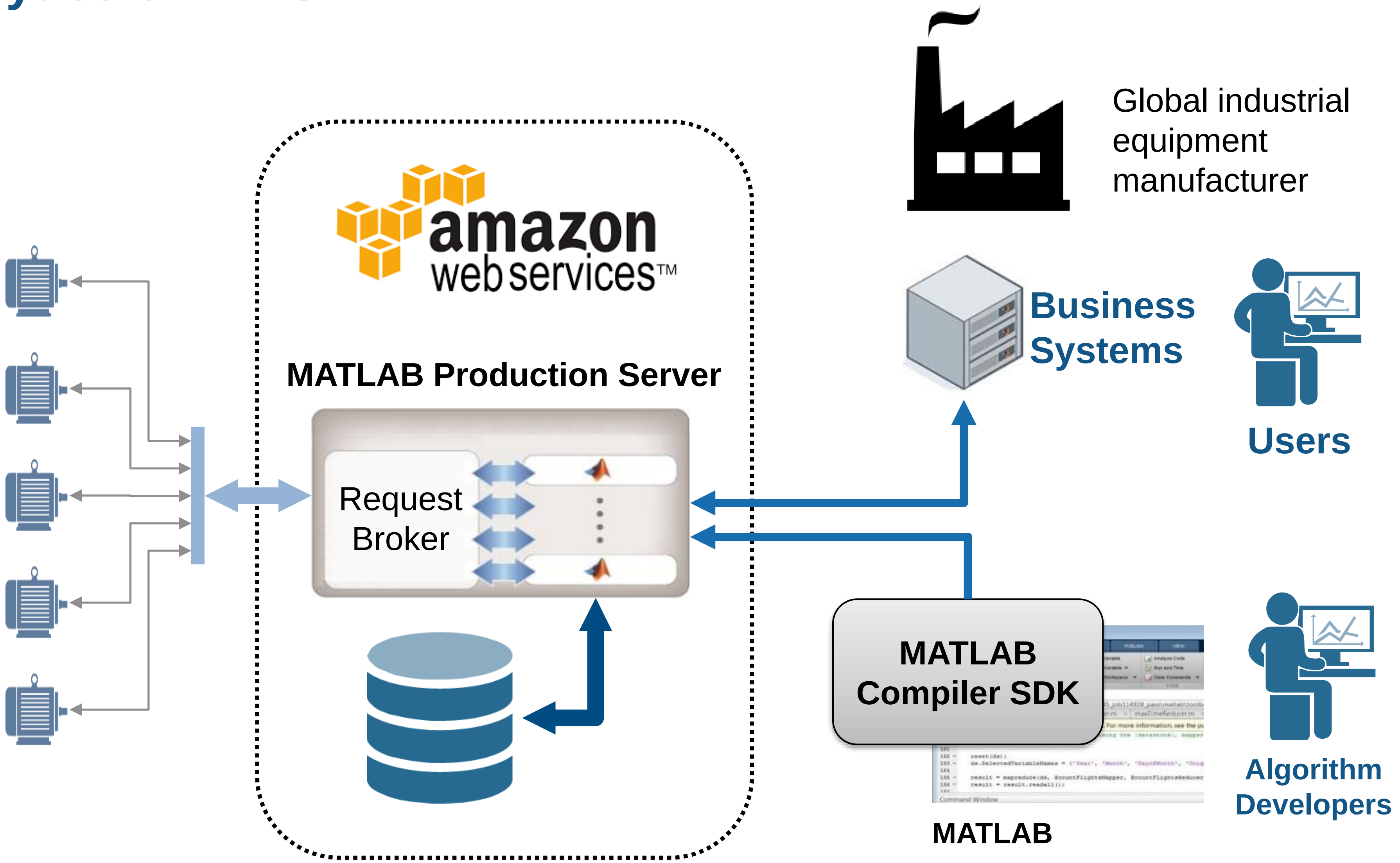
Global financial institution with European HQ

- Saved **€ 2 million annually** for an external system
- Quicker implementation of adjustments in source code by the quantitative analysts
- Knowledge + MATLAB = Build your own systems

Industrial IoT Analytics on AWS

Industrial Equipment

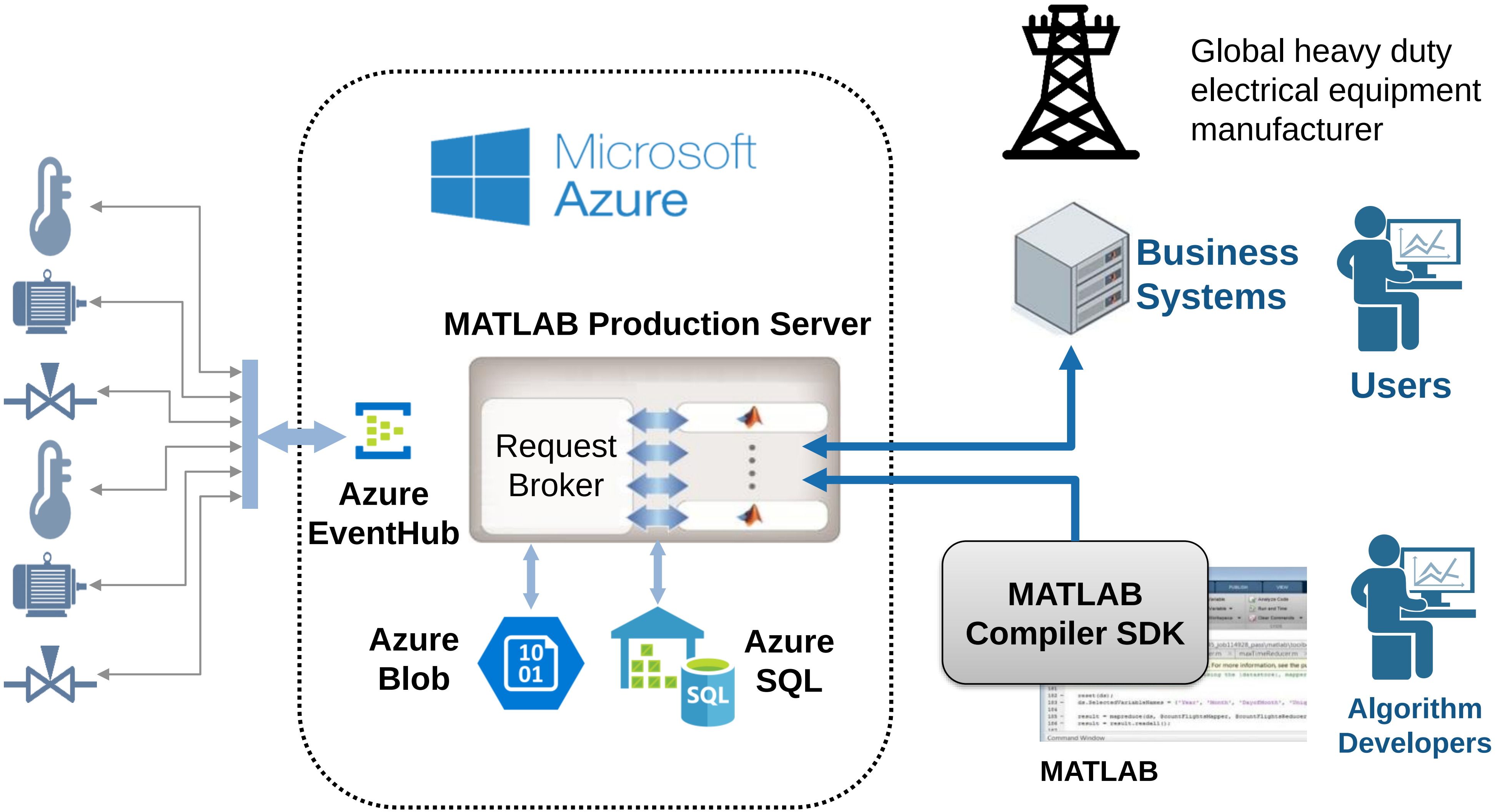
- Networked communication
- Embedded sensors
- Data reduction



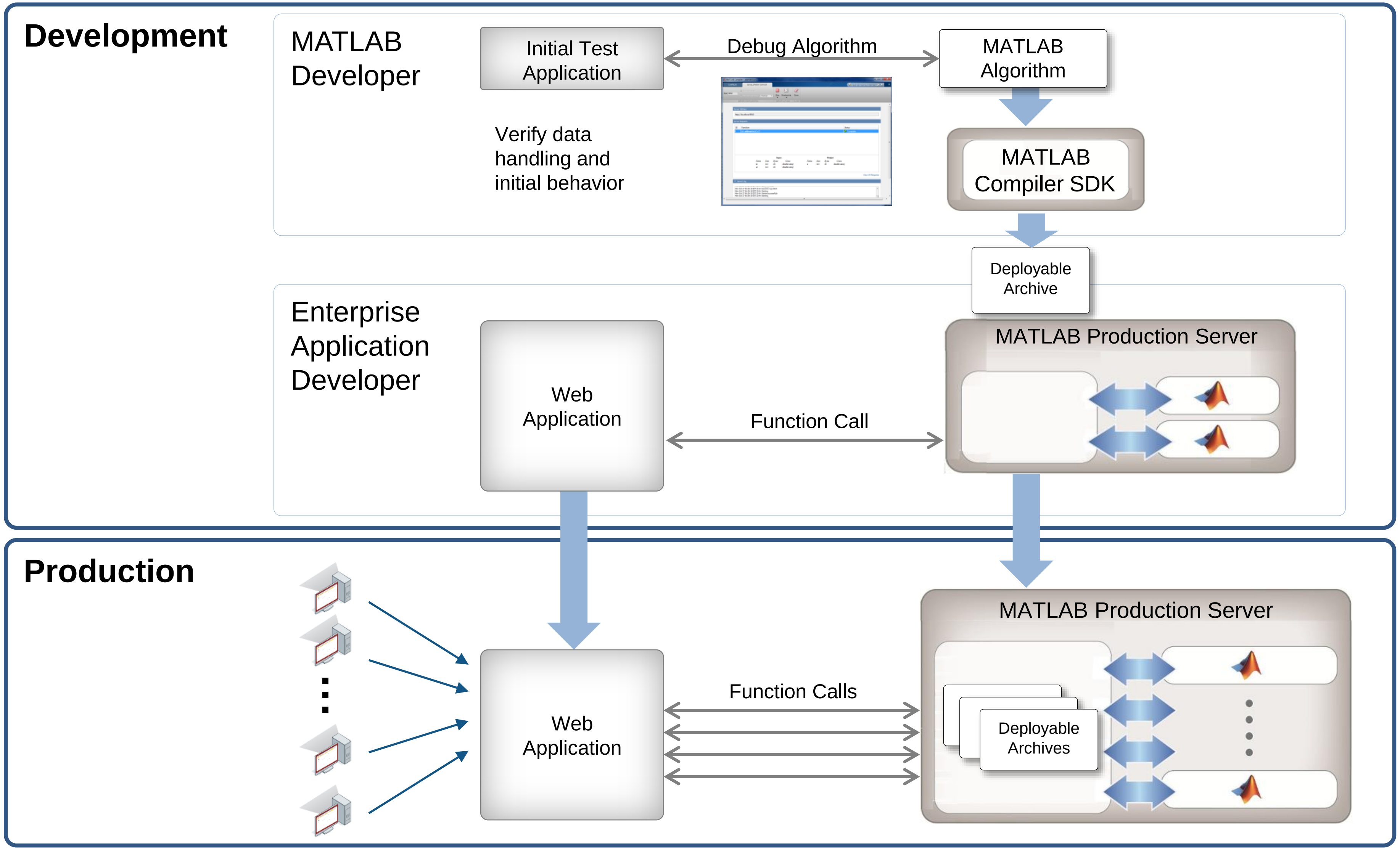
Building Automation IoT Analytics on Azure

Building/HVAC automation control system

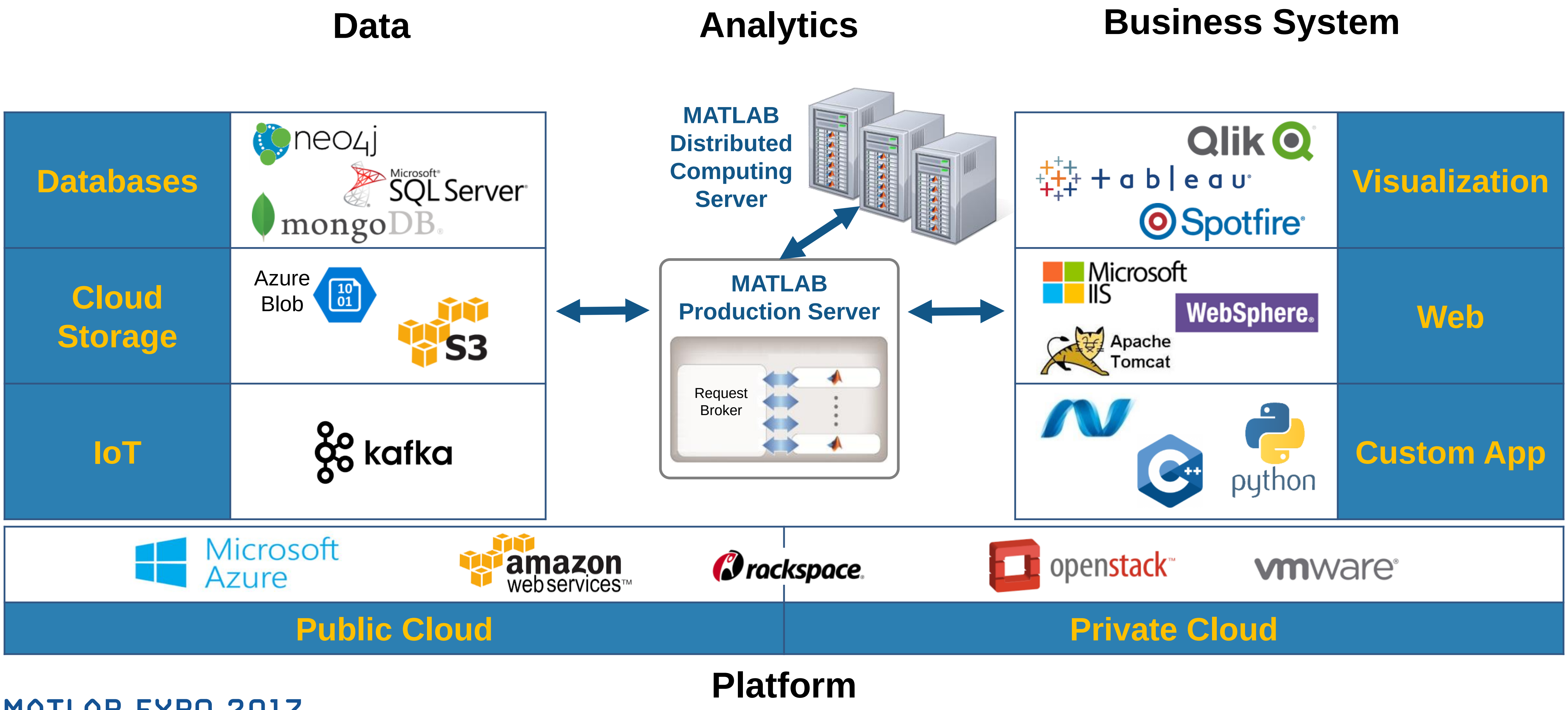
- Variety of sensors and controls
- Networked communication
- Data reduction



Production Deployment Workflow



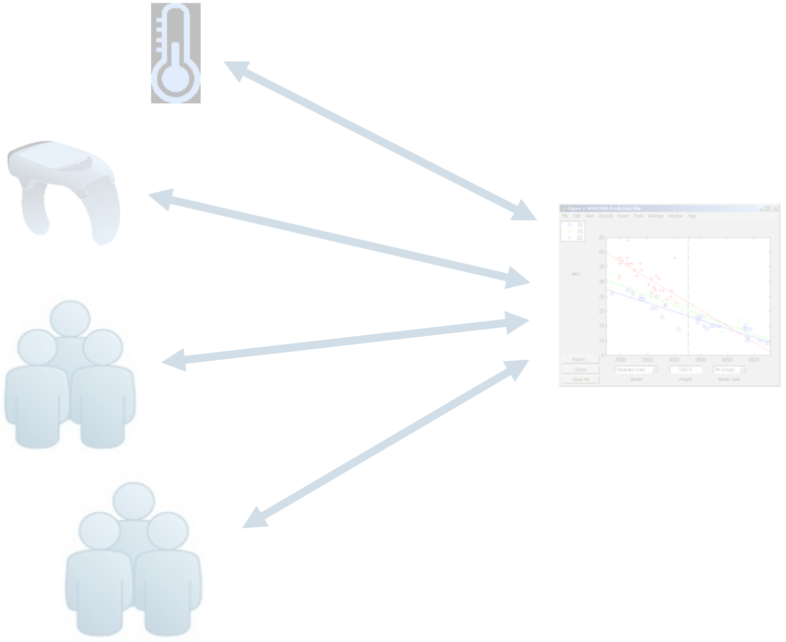
Technology Stack



MATLAB at Scale

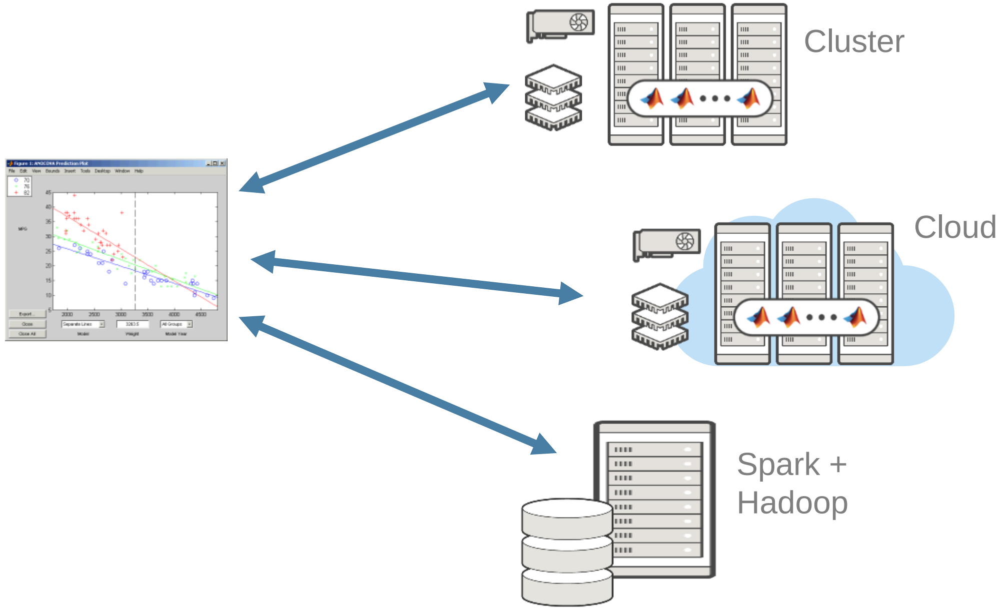
Front-end scaling

- Scale with increasing access requests



Back-end scaling

- Scale with increasing computational intensity
- Scale with increasing data volumes



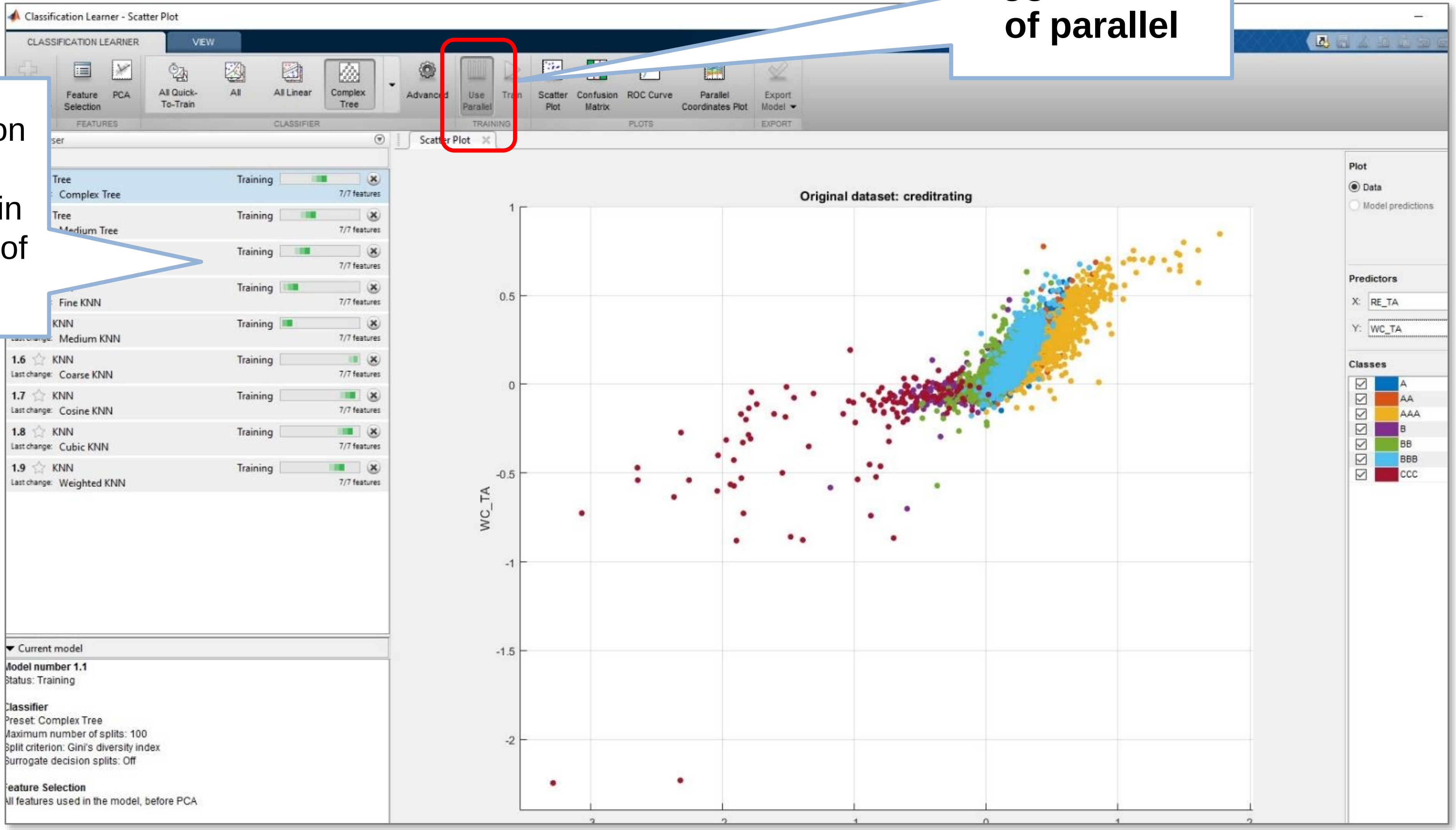
Key Takeaways

1. Leverage parallel computing
2. Handle big data
3. Seamlessly scale from your desktop to clusters or the cloud

Classification learner demo

One click to toggle the use of parallel

Run classification learner quick to train classifiers in parallel instead of one by one



Parallel-enabled Toolboxes

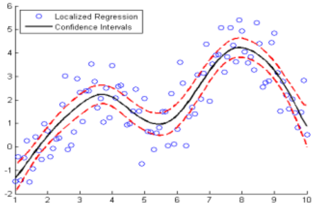
Enable acceleration by setting a flag or preference

Image Processing

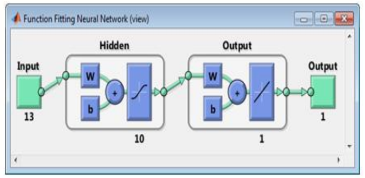
Original Image of Peppers

Recolored Image of Peppers

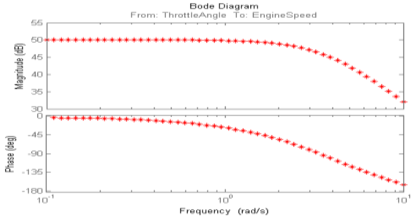
Statistics and Machine Learning



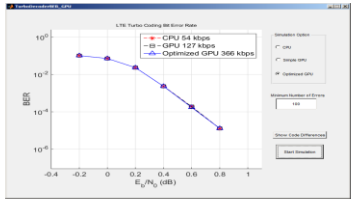
Neural Networks



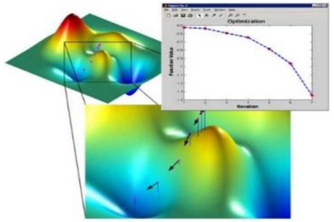
Simulink Control Design



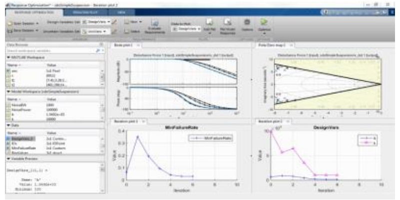
Signal Processing and Communications



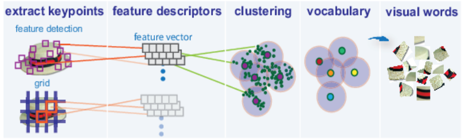
Optimization



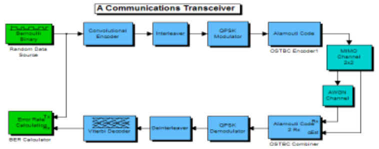
Simulink Design Optimization



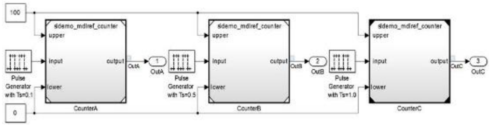
Computer Vision



Communication Systems Toolbox



Simulink/Embedded Coder

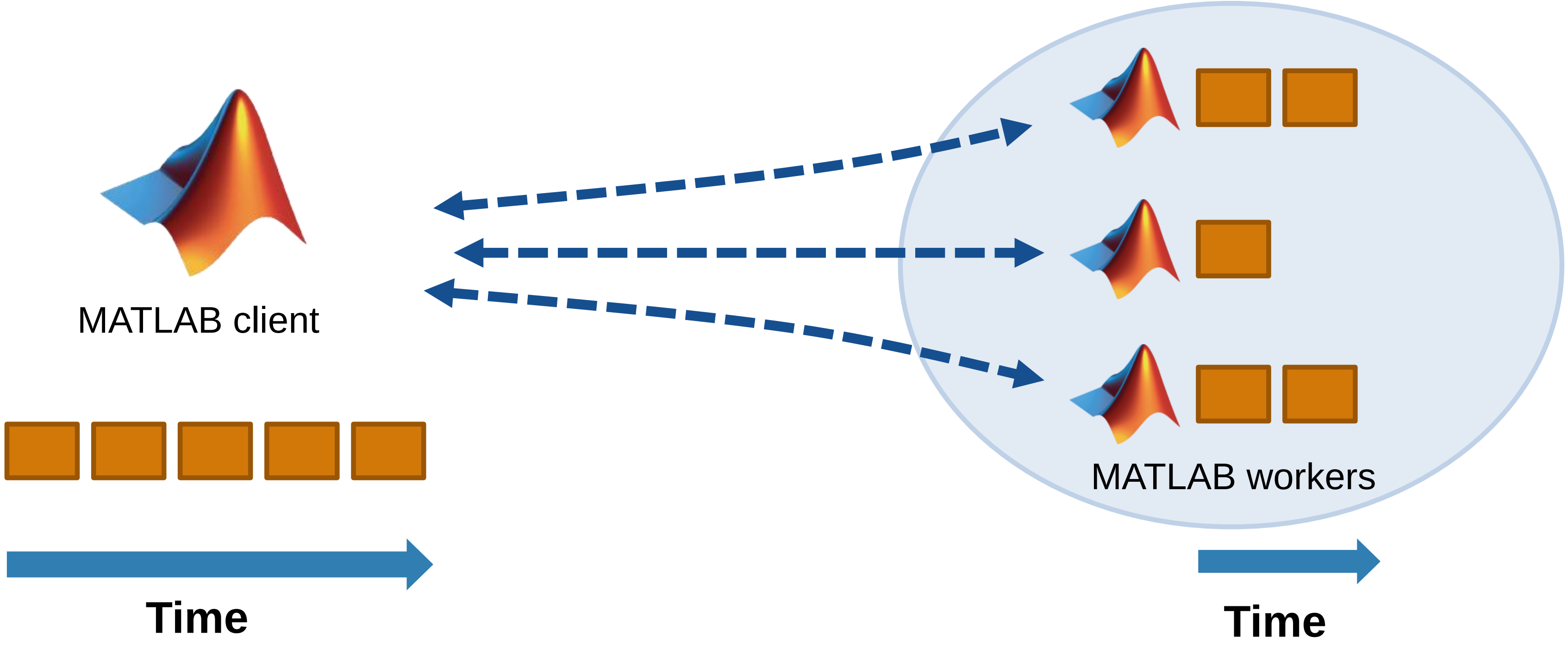


[Other Parallel-enabled Toolboxes]

Independent Tasks or Iterations

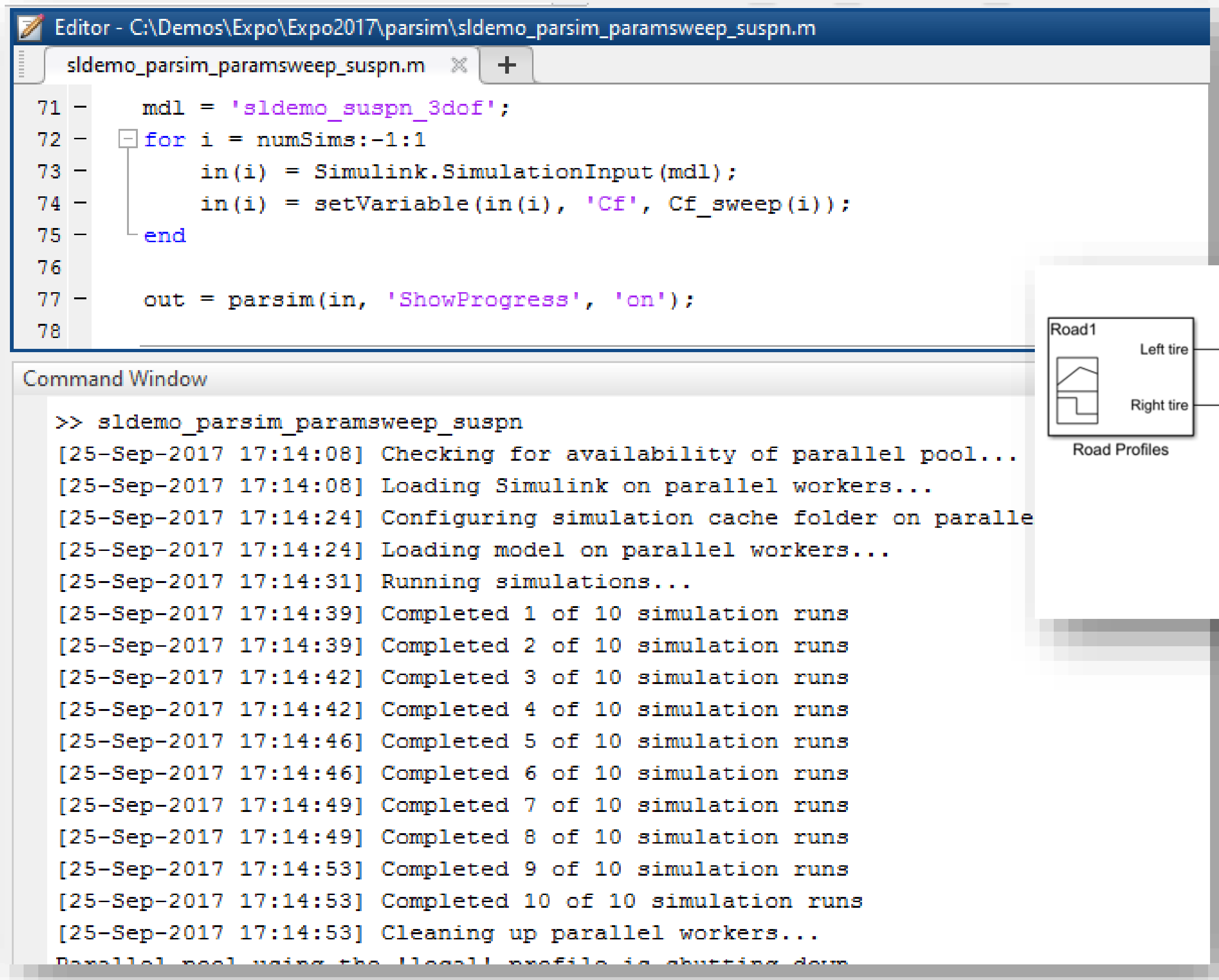
Simple programming constructs: `parfor, parfeval`

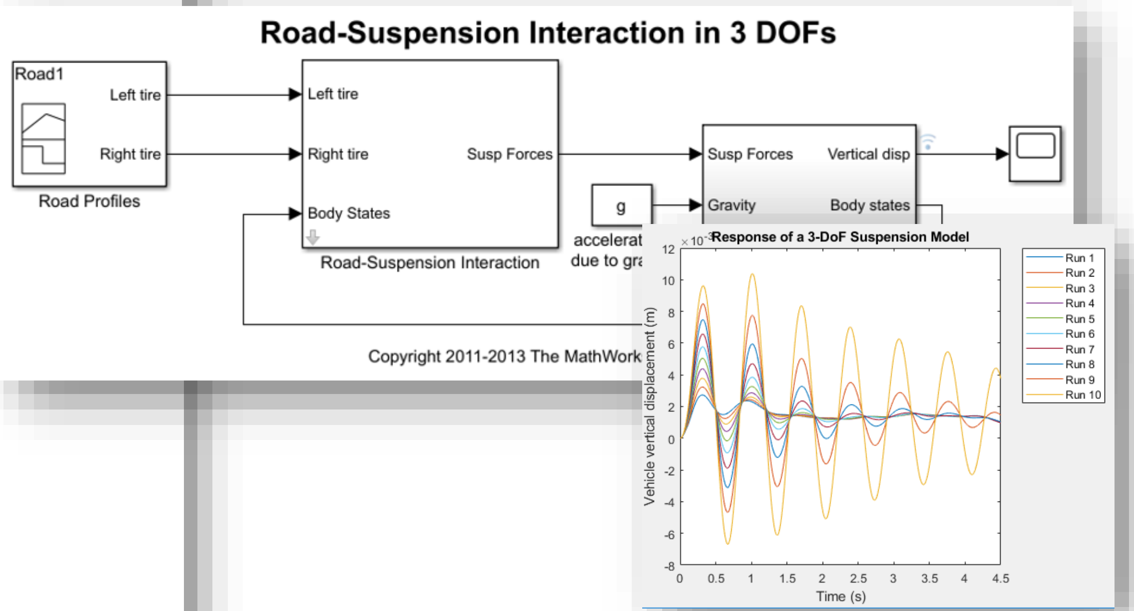
- Examples: parameter sweeps, Monte Carlo simulations
- No dependencies or communications between tasks



Run multiple parallel simulations from the `parsim` command

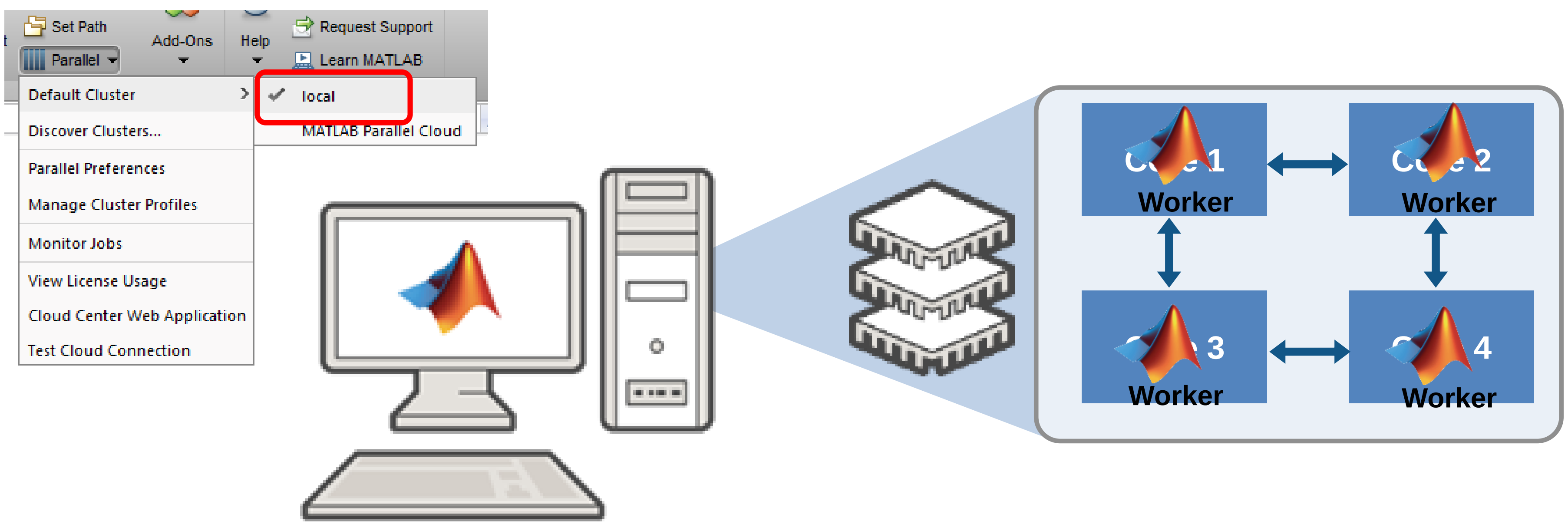
Run Simulink multiple simulations in parallel with simplified workflow





Parallel Computing

Multicore Desktops



Carnegie Wave Energy

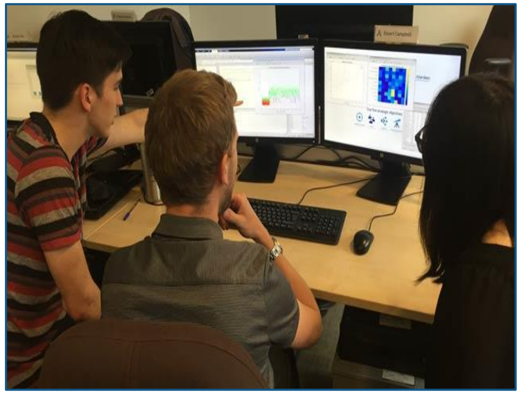
A CETO unit ready for deployment in the wave farm.

Develop unique technology for generating electric power from ocean waves

*"…we can run **simulations in parallel**, and with a twelve-core computer we see an almost **twelvefold increase in speed**."*

Jonathan Fiévez
Carnegie Wave Energy

Aberdeen Asset Management



Improve asset allocation strategies with machine learning techniques

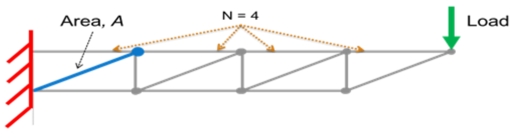
*"… can **develop prototypes** to test machine learning techniques **quickly**… **get rapid, reliable results** by running the algorithms with large financial data sets **on a distributed computing cluster**."*

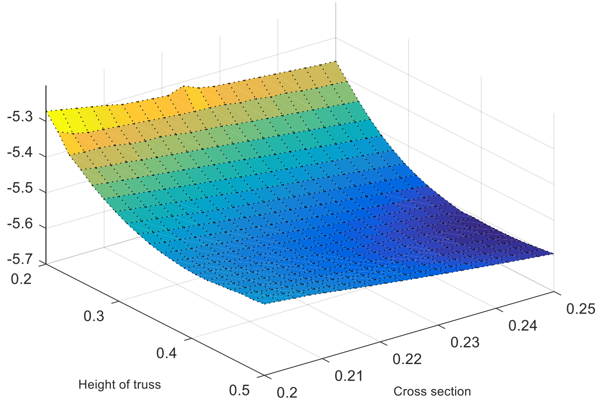
Emilio Llorente-Cano
Aberdeen Asset Management

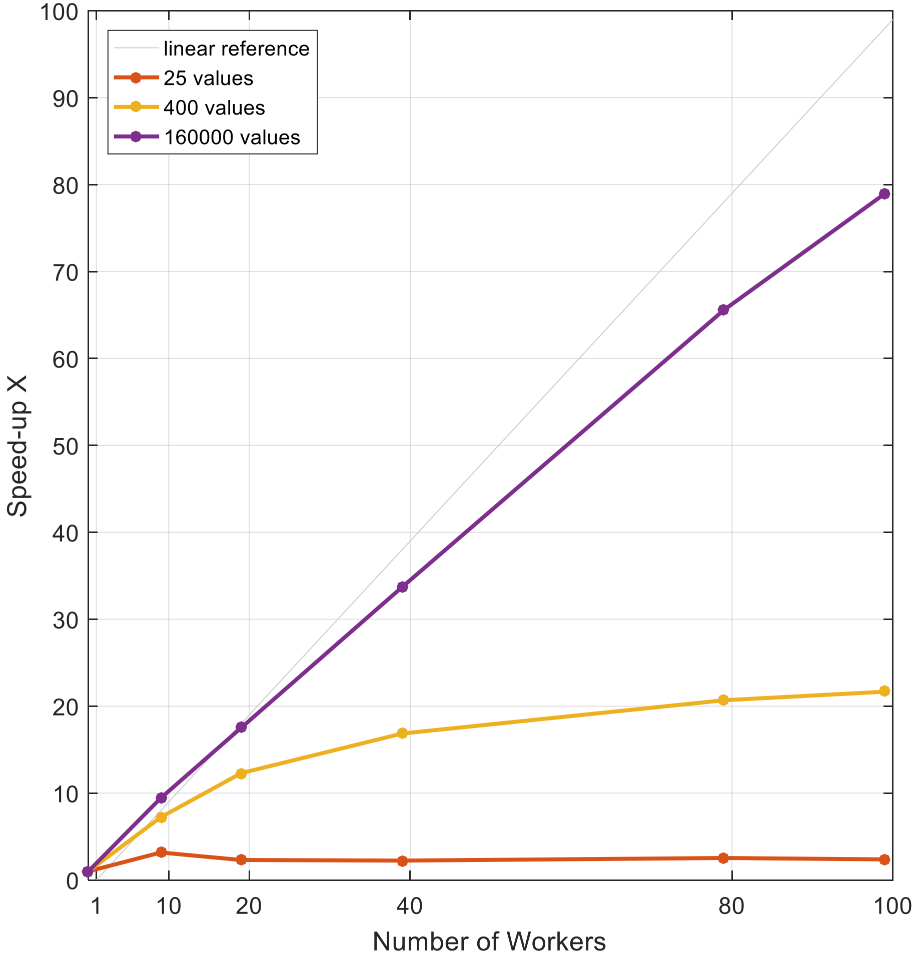
Why parallel computing matters

Scaling case study with a compute cluster

$$M\ddot{x} + C\dot{x} + Kx = F$$





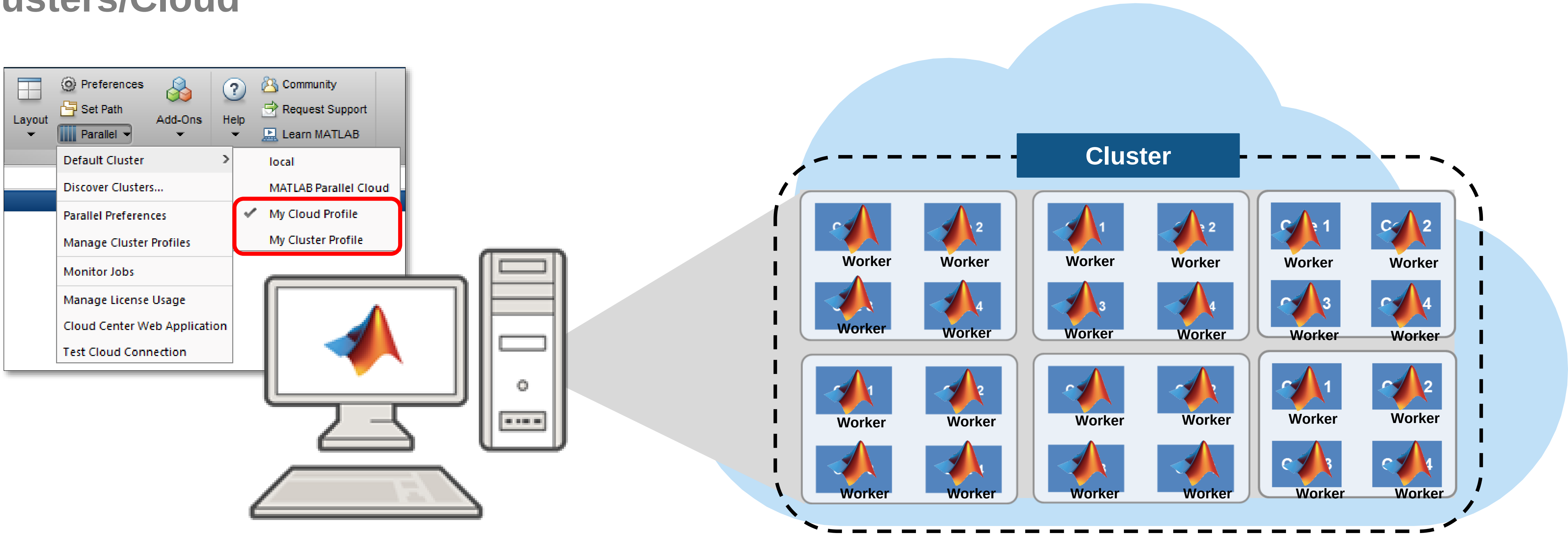


Workers in pool	Compute time (minutes)		
	160e3 values	400 values	25 values
1	140	0.38	0.03
10	15	0.05	0.01
20	8.0	0.03	0.01
40	4.2	0.02	0.01
80	2.1	0.02	0.01
100	1.8	0.02	0.01

Processor: Intel Xeon E5-class v2
16 physical cores per node
MATLAB R2016a

Parallel Computing – Scaling Up

Clusters/Cloud



Parallel Computing Toolbox

MATLAB Distributed Computing Toolbox

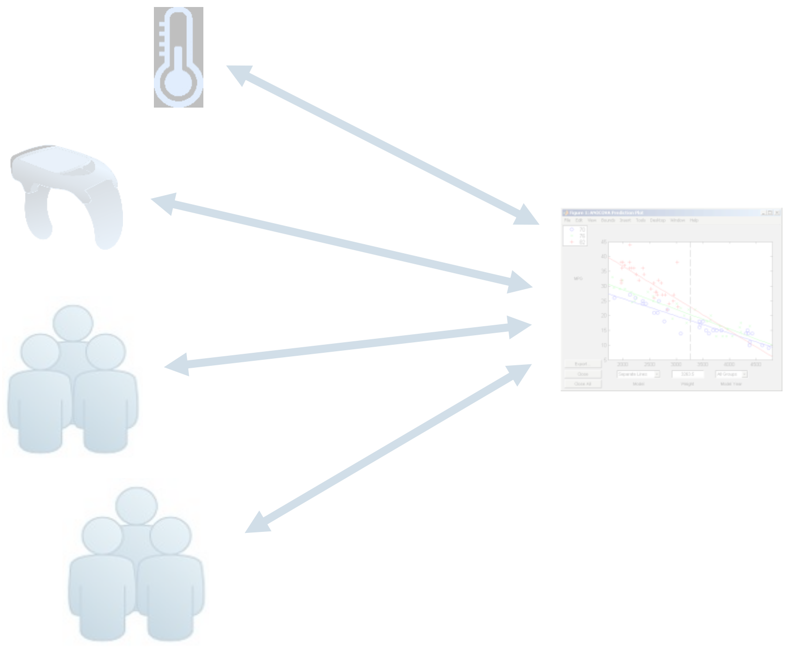
Considerations When Scaling to Clusters

- Workers need access to your code
- Workers need access to the data
- Operating system independent file path management
  - *fileparts, fullfile, filesep*

MATLAB at Scale

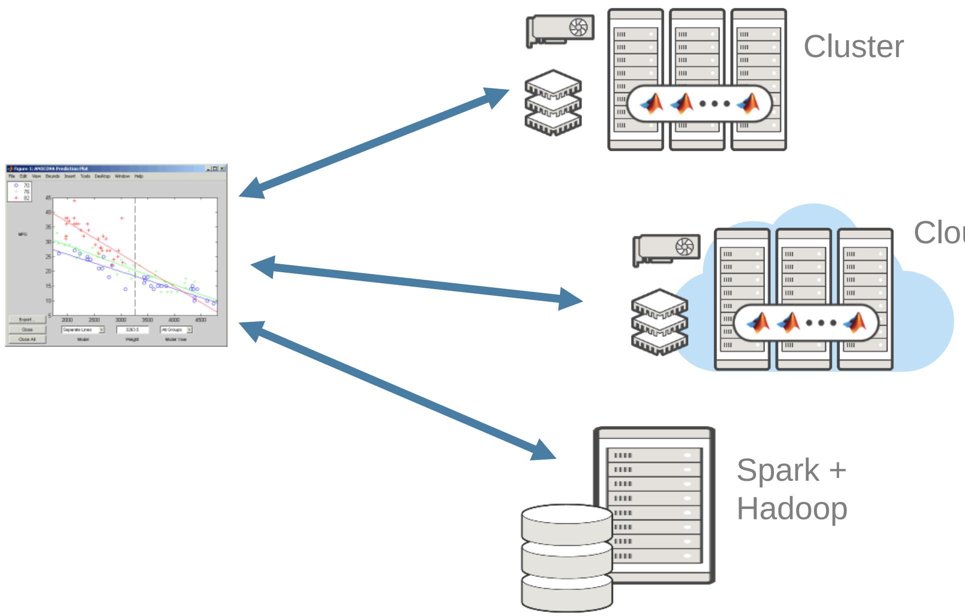
Front-end scaling

- Scale with increasing access requests



Back-end scaling

- Scale with increasing computational intensity
- Scale with increasing data volumes



Large Data Options

Data fits in memory of pool

- Distributed arrays
  - Look like normal MATLAB variables

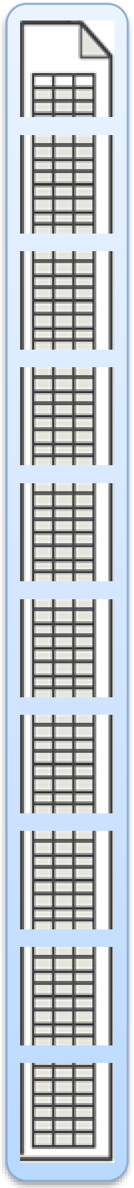
Data does not fit in memory (Big Data)

- Tall arrays
  - Looks like normal MATLAB variables
- Custom map-reduce functions
  - Can be painful to learn

`tall` arrays

- Data doesn't fit into memory
- Lots of observations - "tall"
- Looks like a normal MATLAB array
  - Numeric types, tables, datetimes, strings, etc…
  - Basic math, stats, indexing, etc.
  - Statistics and Machine Learning Toolbox (clustering, classification, etc.)

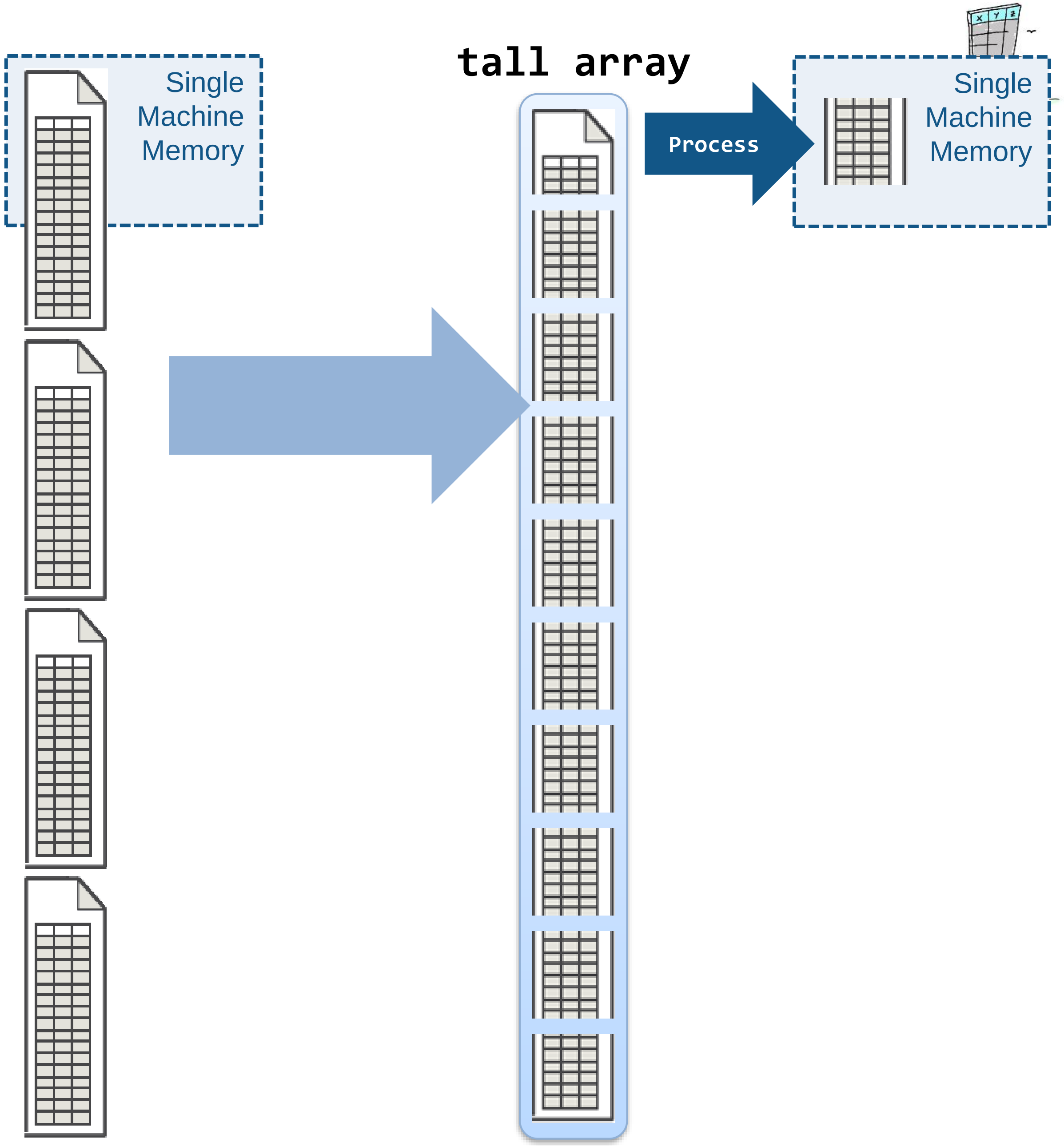
`tall array`



R2016b

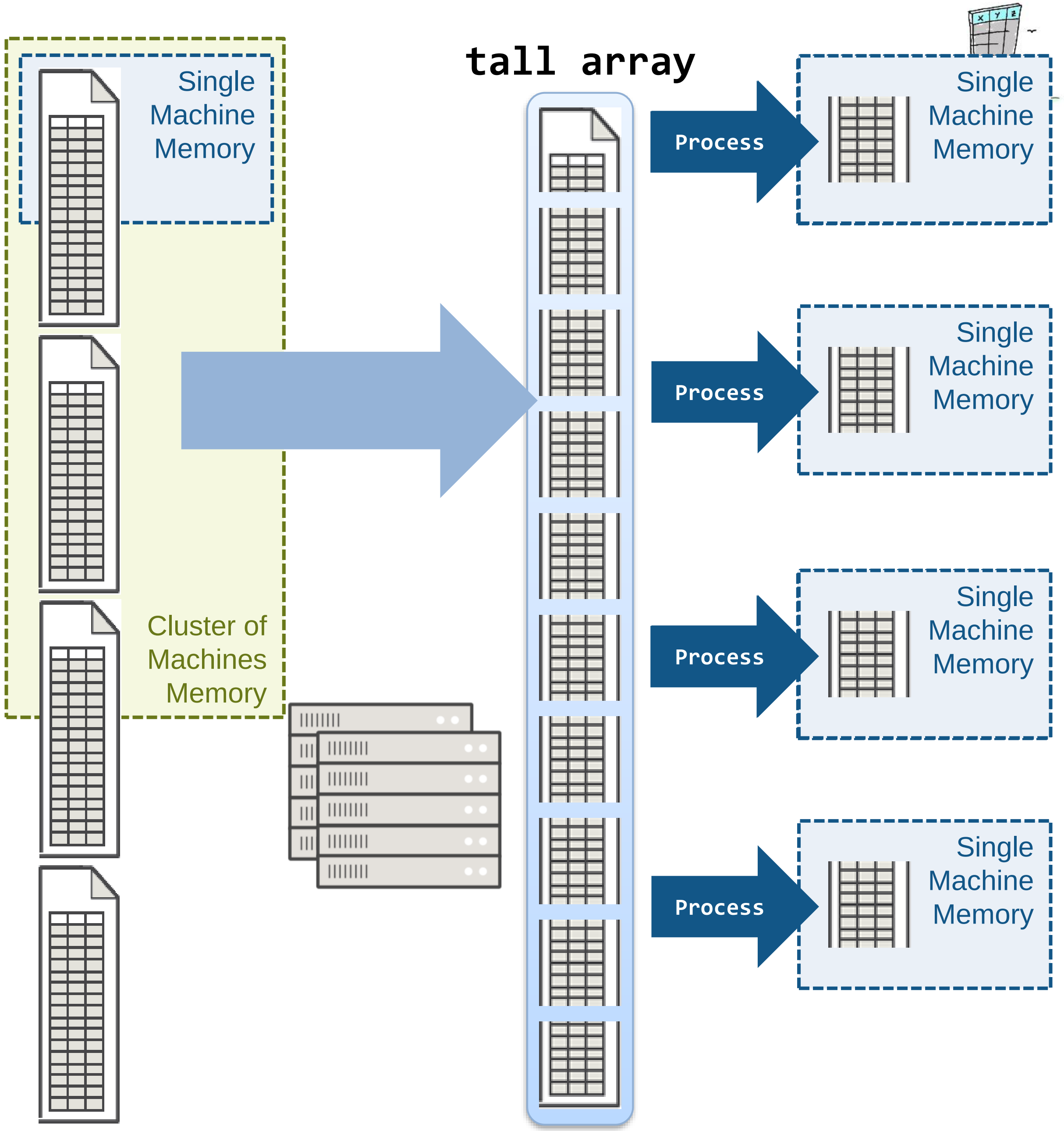
tall arrays

- Automatically breaks data up into small “chunks” that fit in memory
- “Chunk” processing is handled automatically
- Processing code for tall arrays is the same as ordinary arrays



tall arrays - Scaling

- Process several “chunks” at once
- Scale up to clusters



Big Data workflow

ACCESS

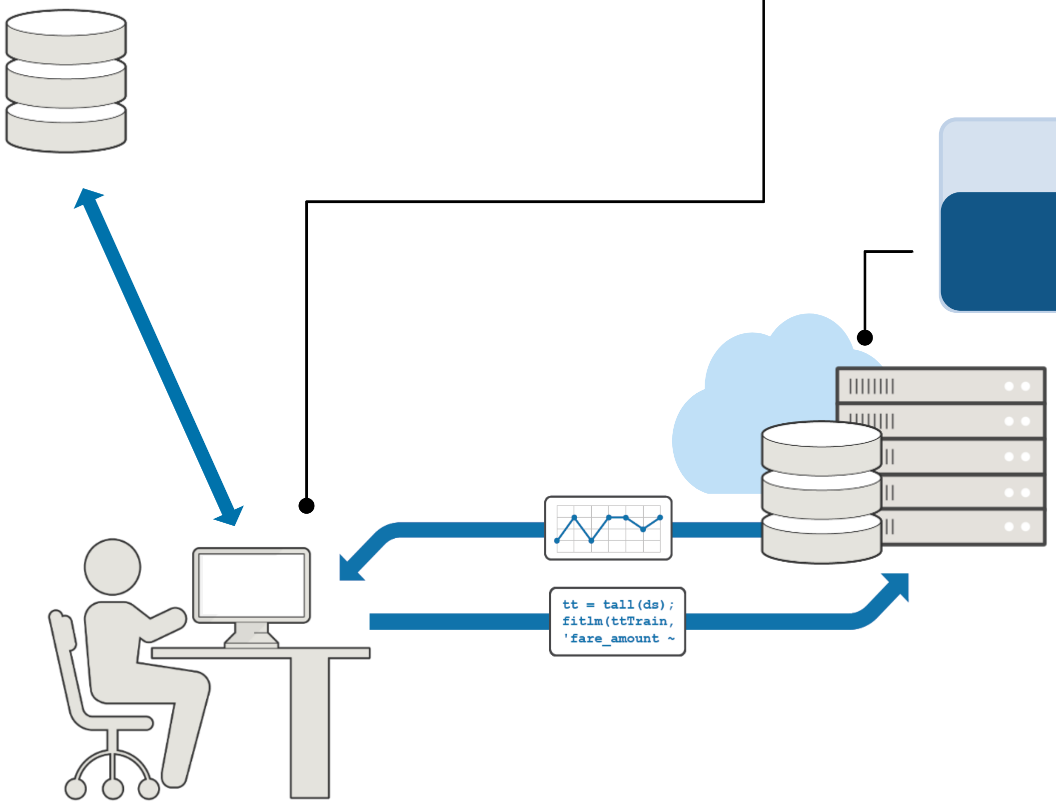
More data and collections of files than fit in memory

PROCESS AND ANALYZE

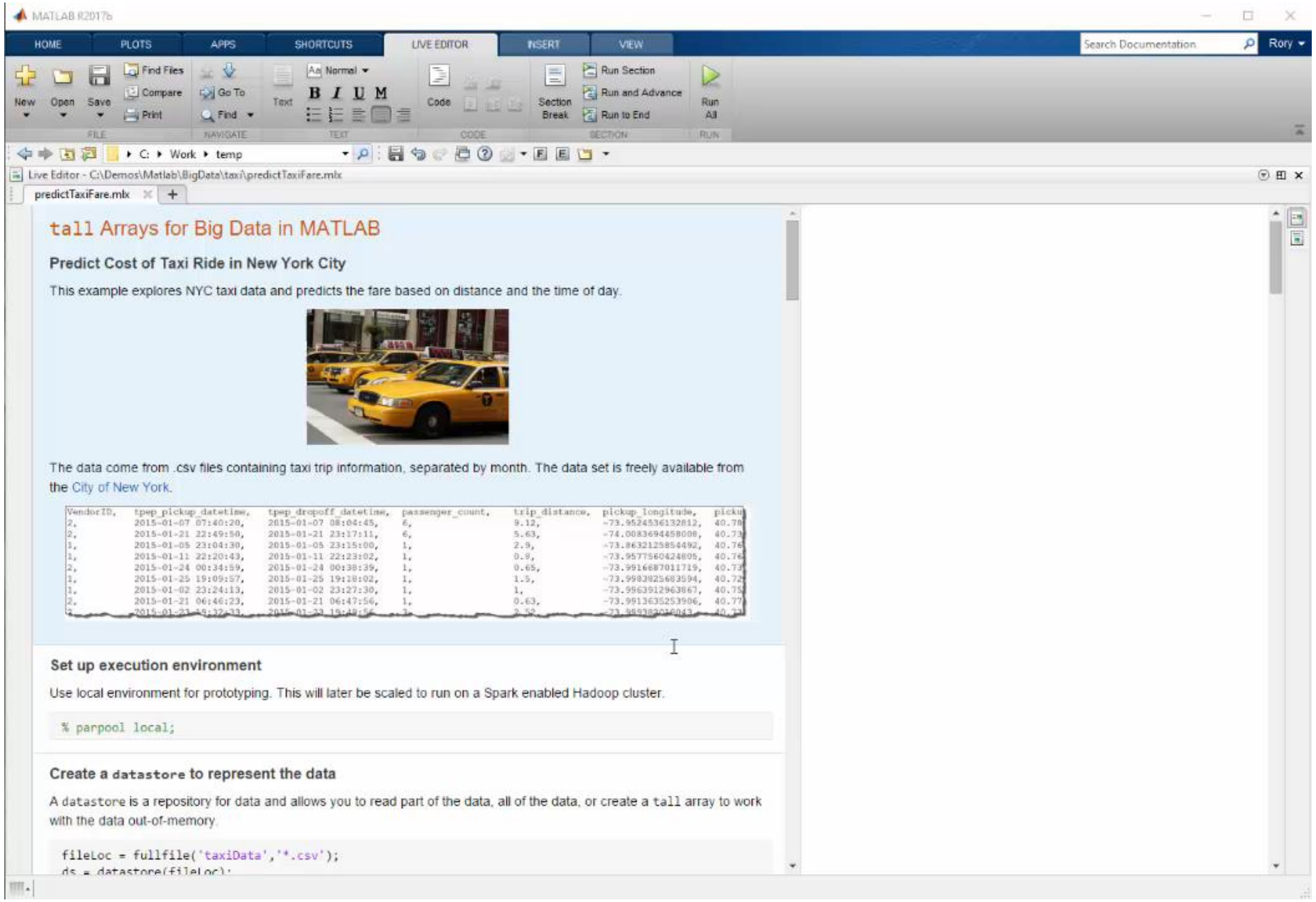
Adapt traditional processing tools or learn new tools to work with Big Data

SCALE

To Big Data systems like Hadoop

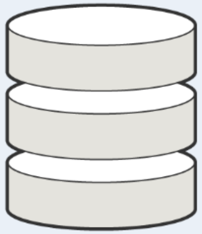


Example: Scaling up to Spark and Hadoop



Using Tall Arrays

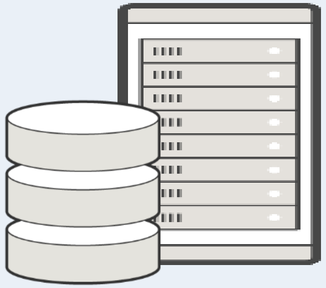
Local disk
Shared folders
Databases
HDFS



- **Tall arrays**
  MATLAB
- **100's of functions supported**
  MATLAB
  Statistics and Machine Learning Toolbox
- **Run in parallel**
  Parallel Computing Toolbox

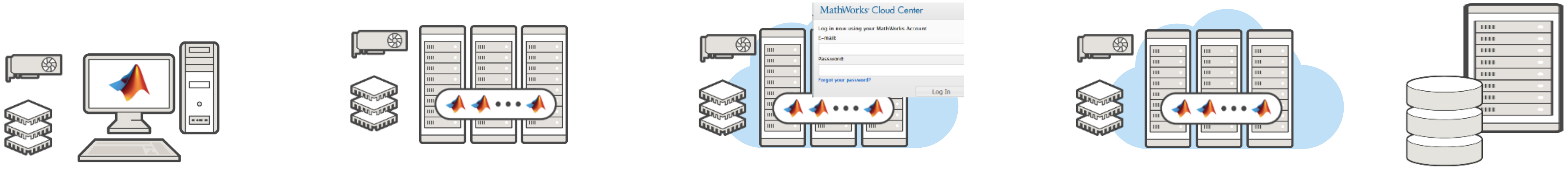
- **Run in parallel on compute clusters**
  MATLAB Distributed Computing Server

Spark + Hadoop



- **Run in parallel on Spark clusters**
  MATLAB Distributed Computing Server
- **Deploy MATLAB applications as standalone applications on Spark clusters**
  MATLAB Compiler

Summary - Scale your applications beyond the desktop



Option	Parallel Computing Toolbox	MATLAB Distributed Computing Server	MATLAB Distributed Computing Server *for Amazon EC2*	MATLAB Distributed Computing Server *for Custom Cloud*	MATLAB Distributed Computing Server *for Hadoop + Spark*
Description	Explicit desktop scaling	Scale to clusters	Scale to EC2 with some customization	Scale to custom cloud	Scale to custom cloud
Maximum workers	No limit	No limit	256	No limit	No limit
Hardware	Desktop	Any	Amazon EC2	Amazon EC2, Microsoft Azure,Others	Hadoop + Spark
Availability	Worldwide	Worldwide	United States, Canada and other select countries in Europe	Worldwide	Worldwide

Learn More: Parallel Computing on the Cloud

MATLAB at Scale

Front-end scaling

- Scale with increasing access requests

Back-end scaling

- Scale with increasing computational intensity
- Scale with increasing data volumes