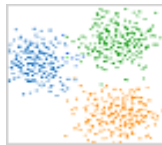
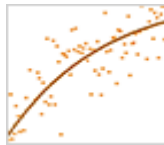


Scaling up MATLAB Analytics



Marta Wilczkowiak, PhD
Senior Applications Engineer
MathWorks

Agenda

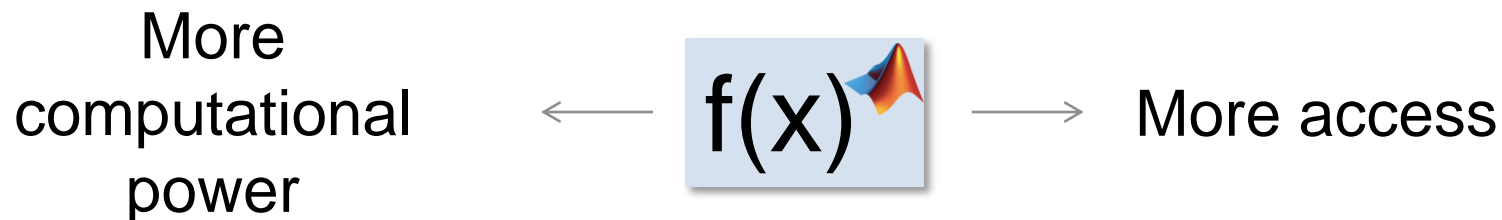
- Giving access to your analytics to more users
- Handling larger problems

When Scalability is a Challenge

- You have great functionality, yet people do not have access to it
- You have purchased more computers, yet you can not execute much more functions
- You hired more people, yet software gets developed at the same scale

Examples

- Increase number of Monte Carlo simulations
- Share portfolio, risk or trading apps with a few traders
- Deploy pricing functions into an enterprise pricing platform serving front, middle and back offices
- Update prices nightly, or intraday, in an enterprise database
- Add strategies to a trading platform

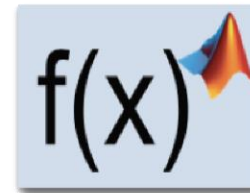


Multi-tier Architecture

Presentation
(Client)



Logic
(Algorithm)



Data

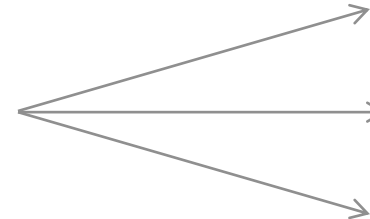
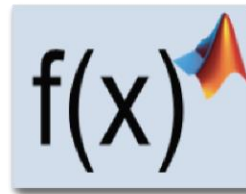


What We Are Looking at in This Session

More access

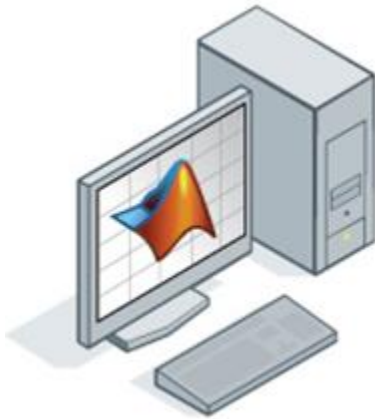


More
computational power

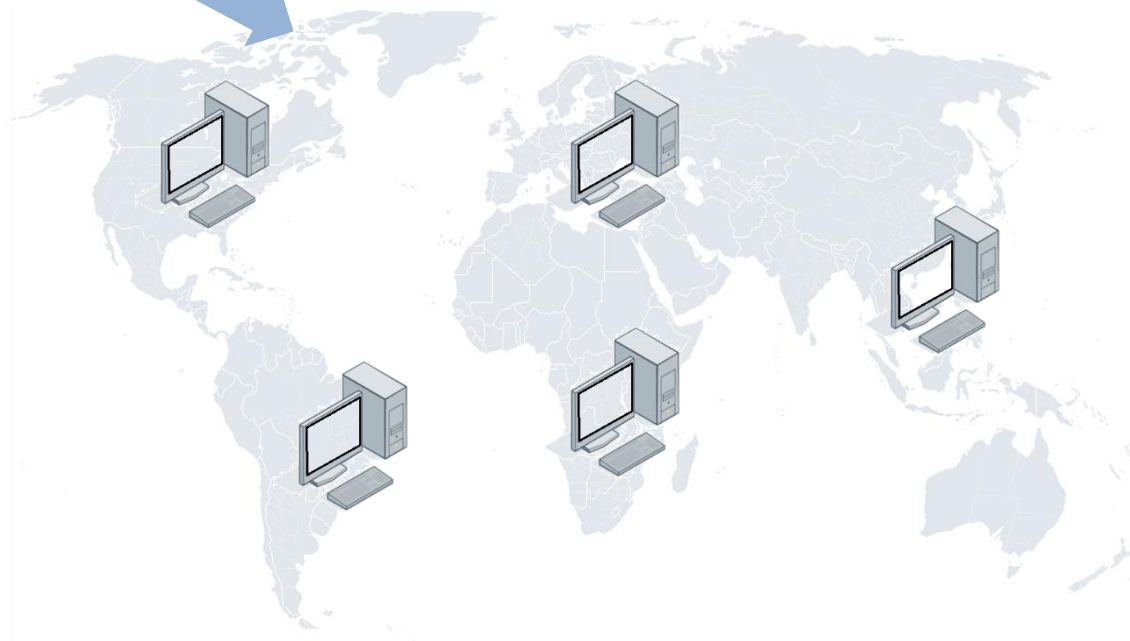
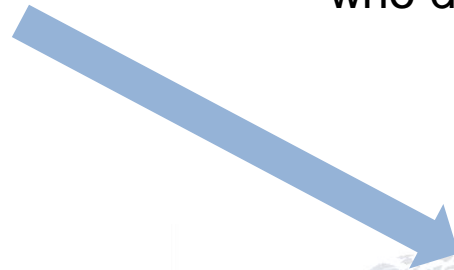


Enabling Access to MATLAB Algorithms (Deployment)

What Is Application Deployment?



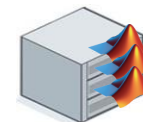
- Sharing MATLAB apps or native files with other MATLAB users
- Sharing MATLAB programs with people who do not have MATLAB



Tools

Deployment options

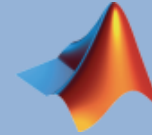
- Live MATLAB
 - Apps
- MATLAB Compiler
 - Standalone applications
 - Software modules
- MATLAB Production Server



Deployment options

- Live MATLAB

- Apps

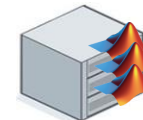


- MATLAB Compiler

- Standalone applications
 - Software modules



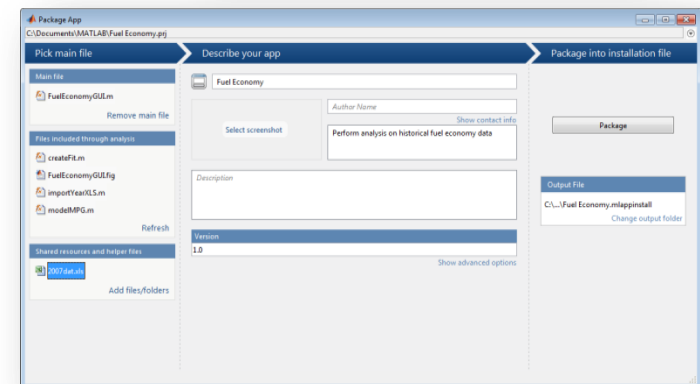
- MATLAB Production Server



MATLAB Apps

MATLAB Apps: Custom interactive applications running in MATLAB

- Easy to package, install and find
- Support full MATLAB language
- Require MATLAB to run



Deployment options

- Live MATLAB

- Apps

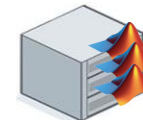


- MATLAB Compiler

- Standalone applications
- Software modules



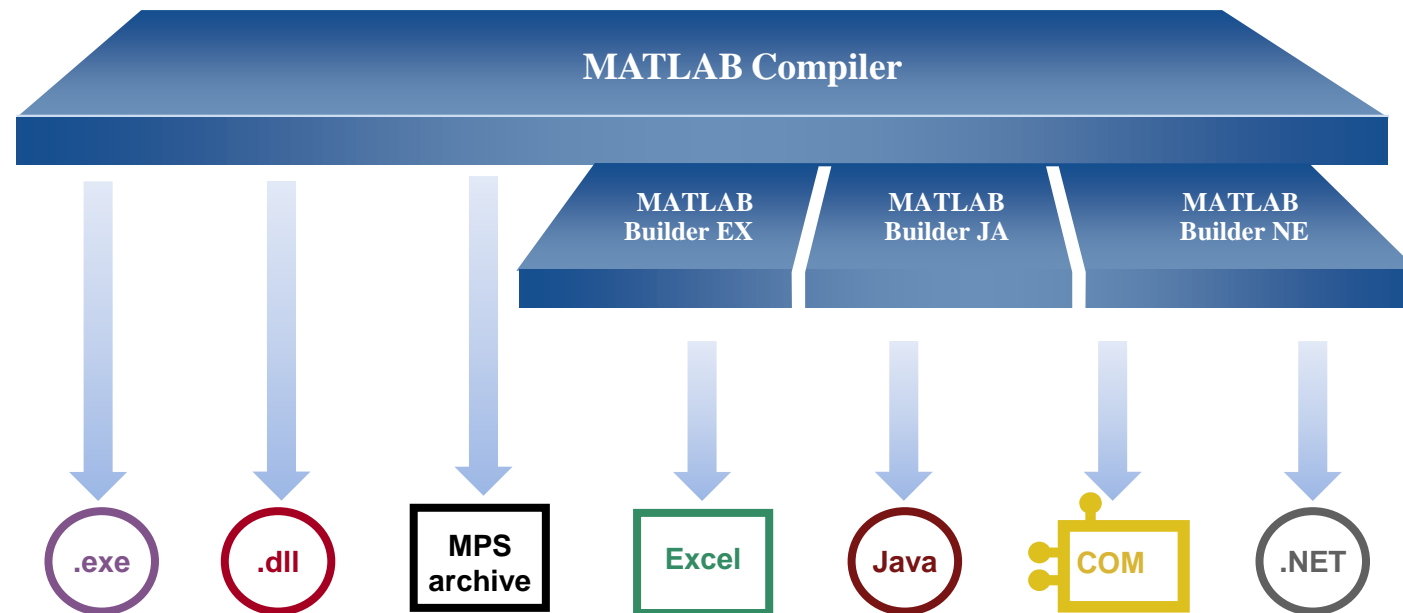
- MATLAB Production Server



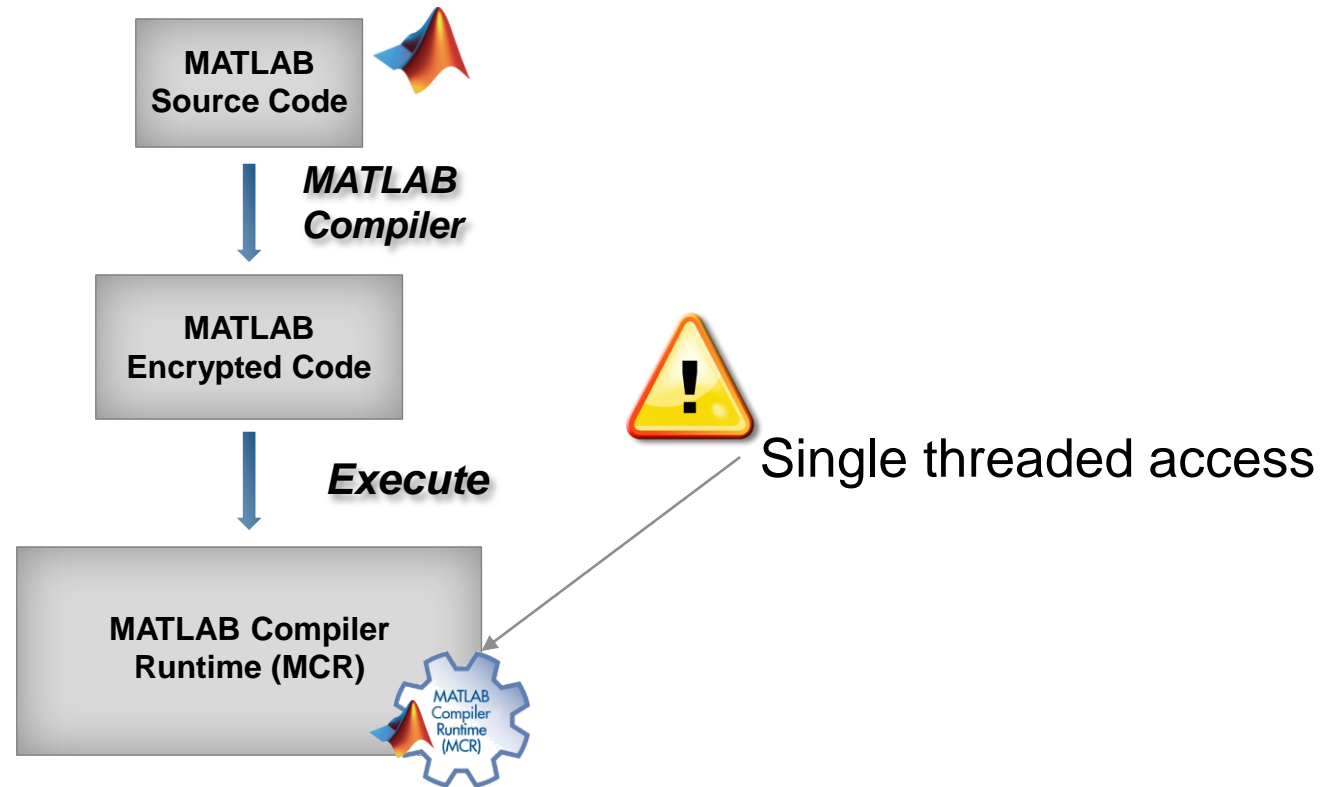
What is MATLAB Compiler?

Software that let's you to:

- Package your MATLAB code so that it does not require MATLAB license
- Encrypt your intellectual property



MATLAB Compiler Runtime

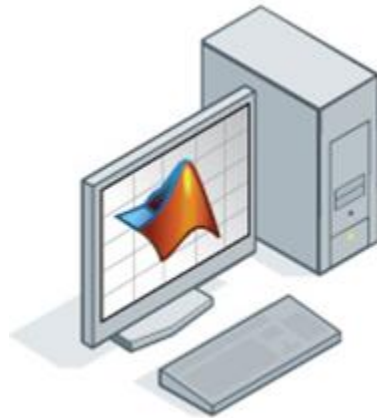


Typical deployment process for desktop applications

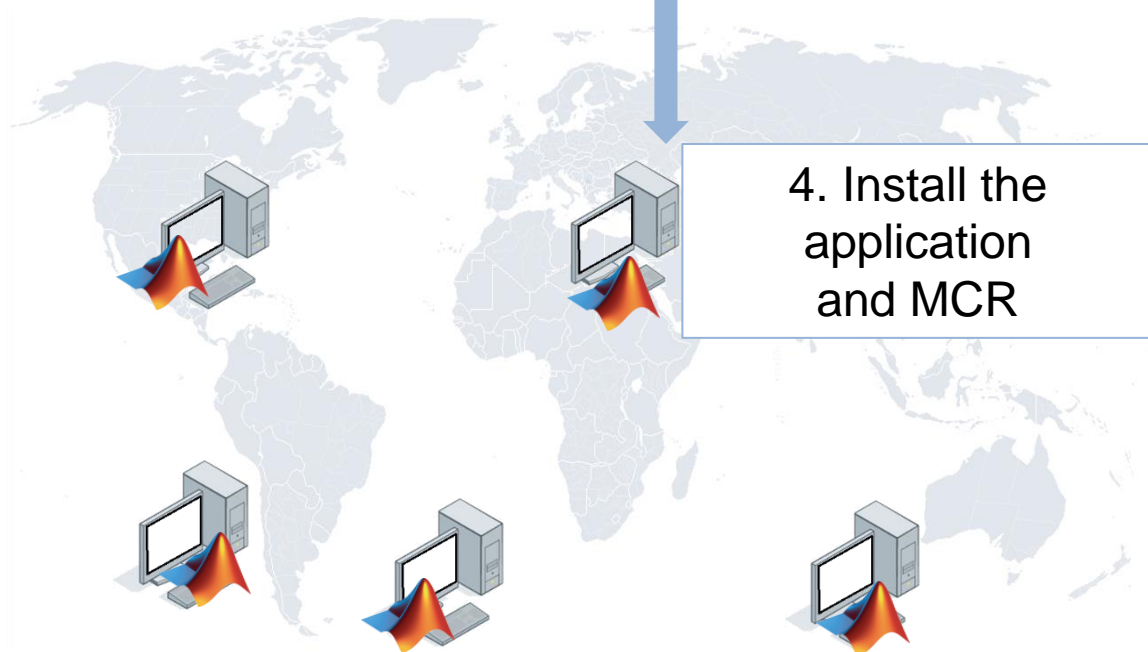
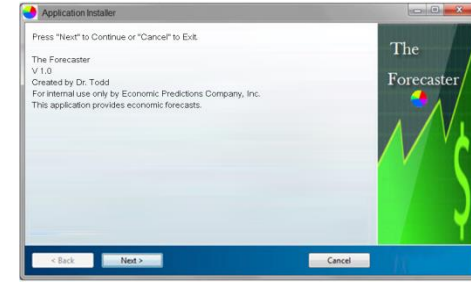
1. Create MATLAB code

2. Package the code

3. Share the installer



MATLAB Compiler



4. Install the application and MCR

Desktop Applications

Potential scalability challenges

- Distributing software updates to multiple users
- Installing application in uncontrolled IT setup (admin rights, different OS, etc.)
- Slow data transfer from a remote location
- Hardware resources on the client machines inadequate to application demands
- Distributing MCR updates to multiple users

Scaling Up: Demand for Server Analytics

Client

- Less need for processing power and memory



Server

- Simplified code management
- Optimal access to data and hardware resources
- Security

Challenge:
Giving access to MATLAB
to multiple requesters
(single threaded MCR)



Deployment options

- Live MATLAB

- Apps

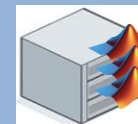


- MATLAB Compiler

- Standalone applications
 - Software modules

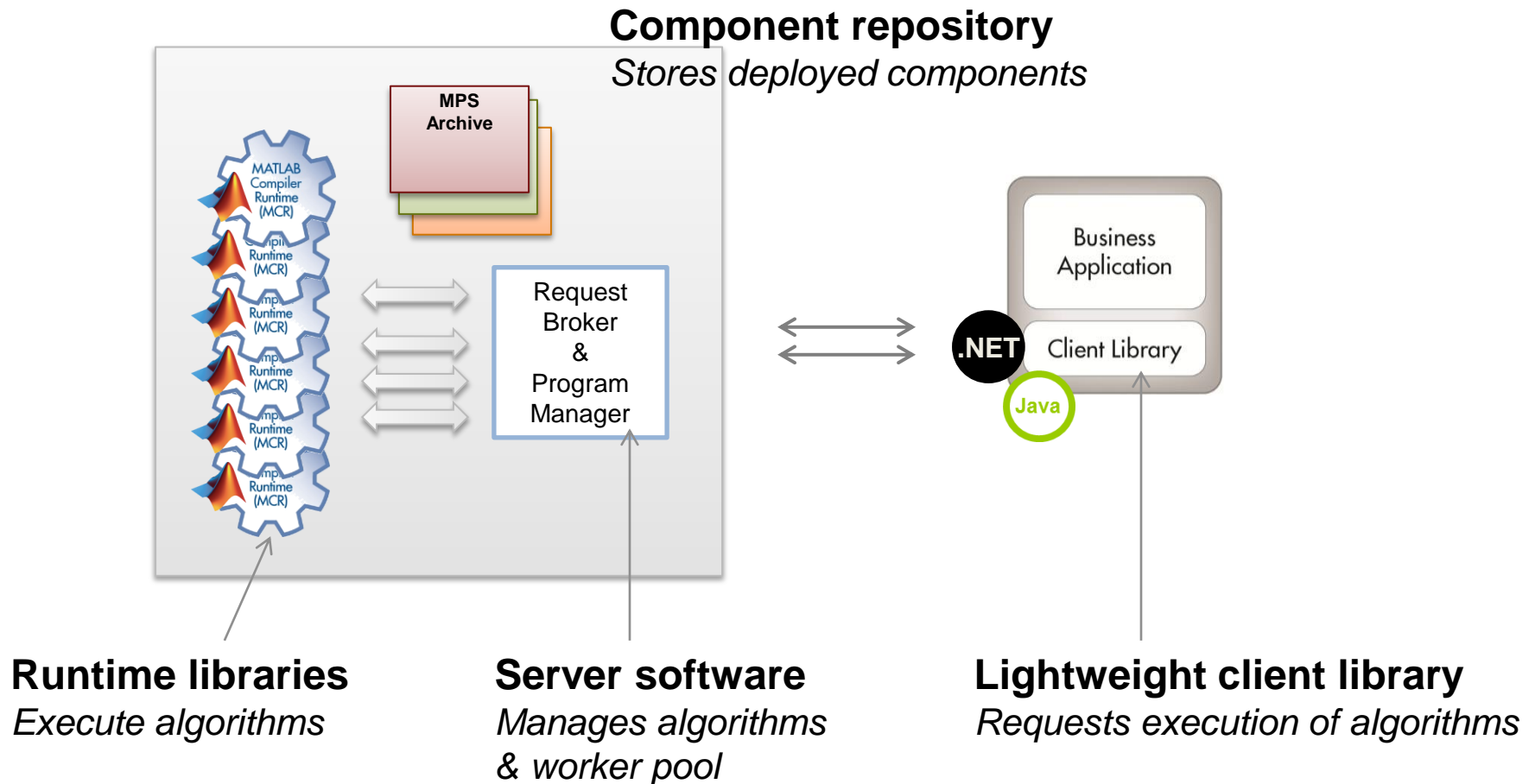


- MATLAB Production Server

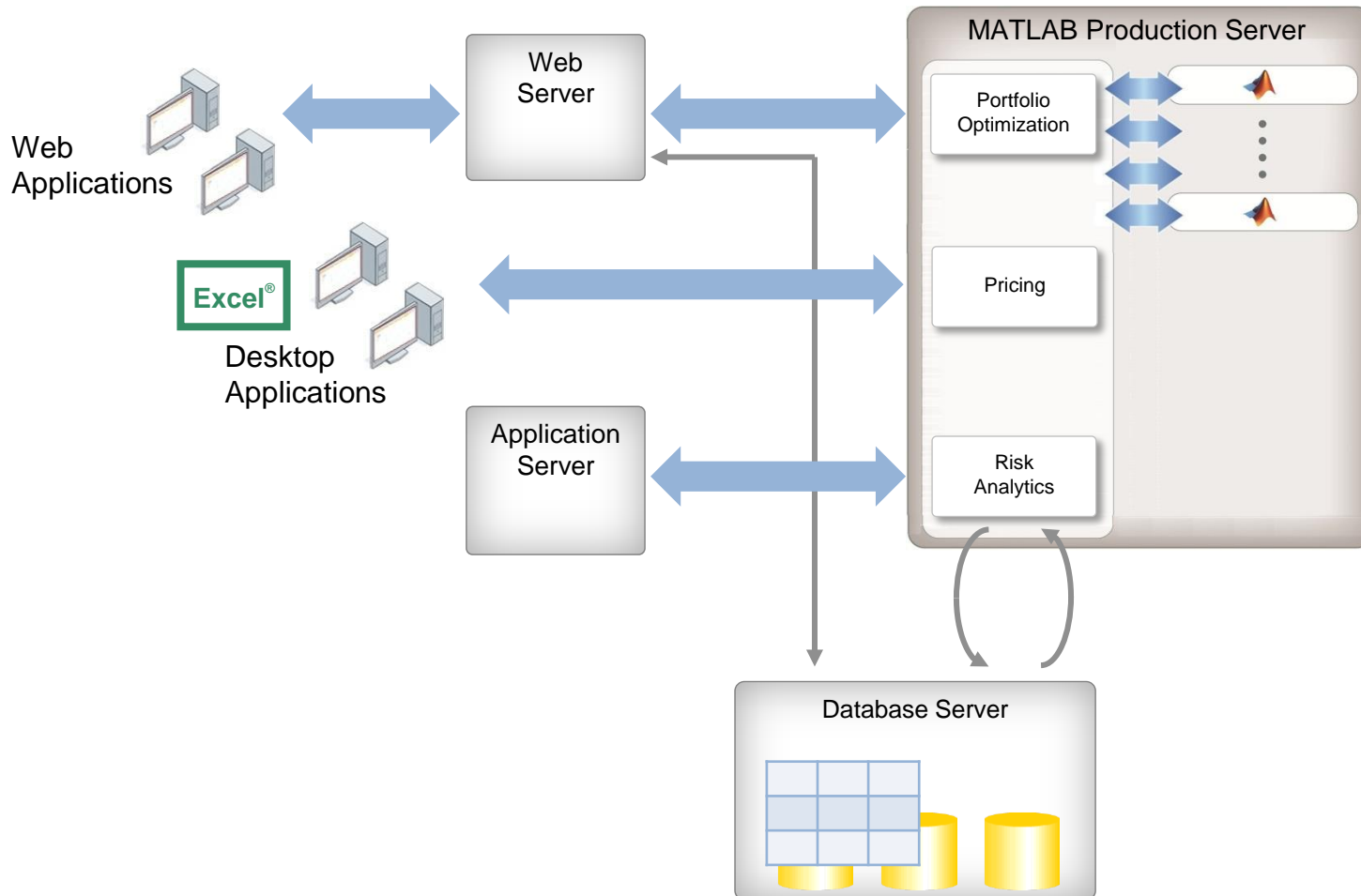


What is MATLAB Production Server?

Enterprise class framework for running packaged MATLAB programs

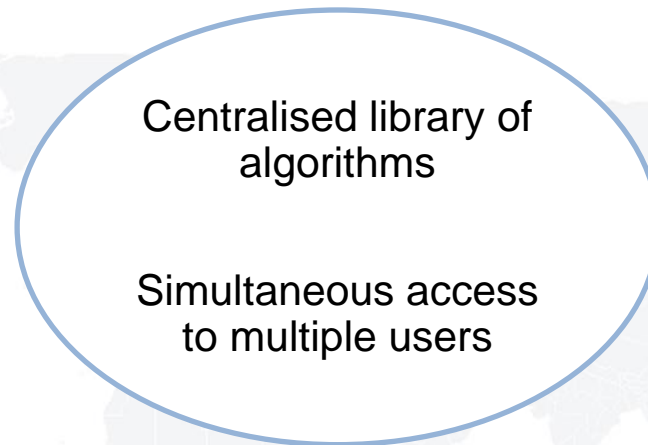


Integrating with IT Systems



MATLAB Production Server

Scalable and reliable way for giving access to MATLAB algorithms



Simple management

- Let MPS manage the worker pool, algorithms and MCRs
- Update the algorithms automatically
- Separate the development and execution of application and algorithmic code

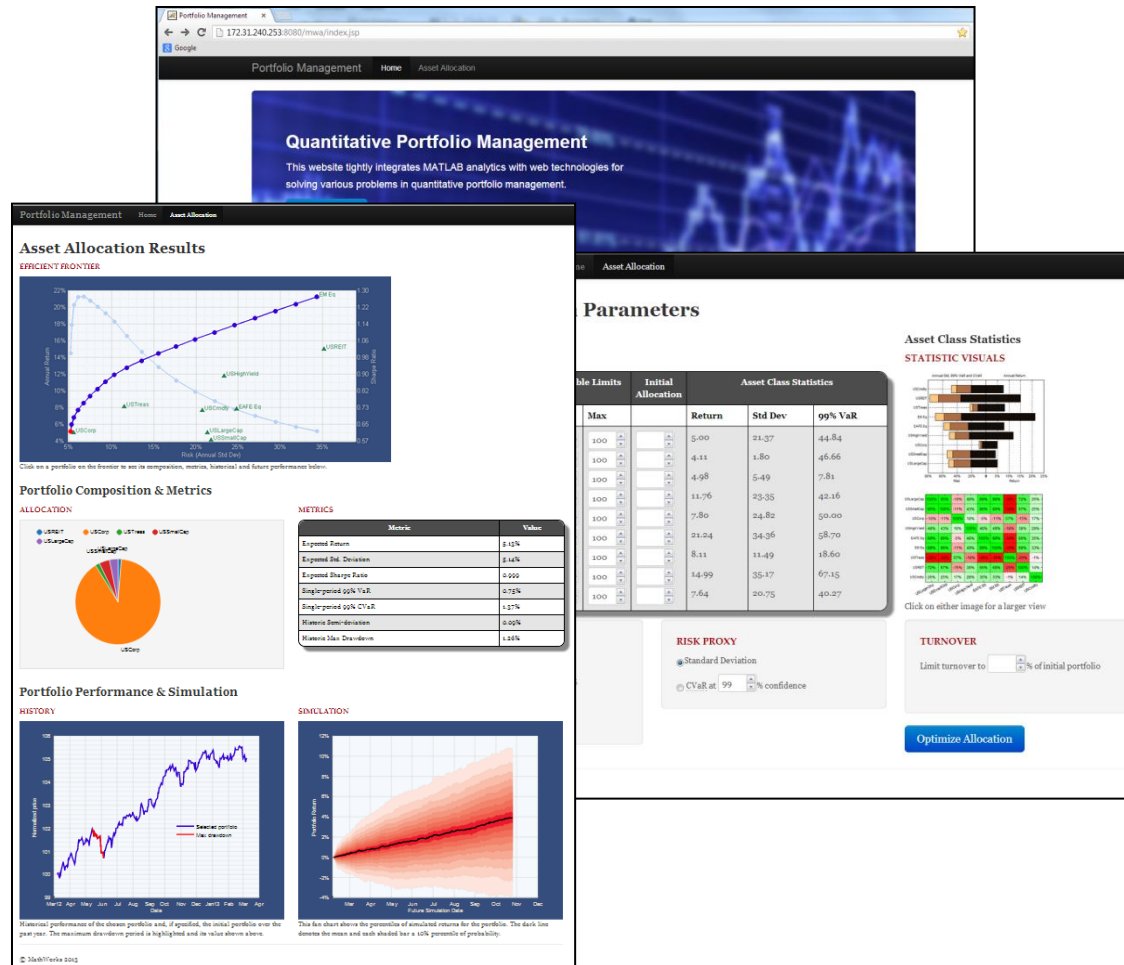
Execution efficiency

- Access immediately
- Execute with optimised hardware and access to the data

Ease of access

- Use with different clients: web, database & application servers
- Integrate easily with application software using the lightweight client library
- Call using native data types

Deploying Enterprise Applications



Deploying Your Code To a Server

Considerations

Server Deployment: Main Considerations

- Will Clients Call the Algorithm Simultaneously?
- What Interfaces the Client Can Use?
- What Should the Input / Output Be ?
- What are Other Dependencies of the Algorithm?
- How to Keep the Algorithm Available?
- What is the Response Time?

Client

$f(x)$

Data

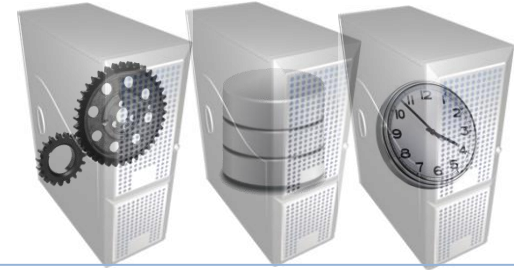
Will Clients Call the Algorithm Simultaneously?



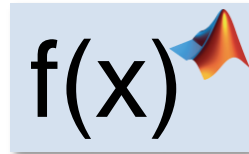
Desktop application



Web application



Database/web/application server



Providing access to multiple MCRs

- ✓ - Using MPS
- Using 3rd party or proprietary solutions (watch MCR start up and development time)

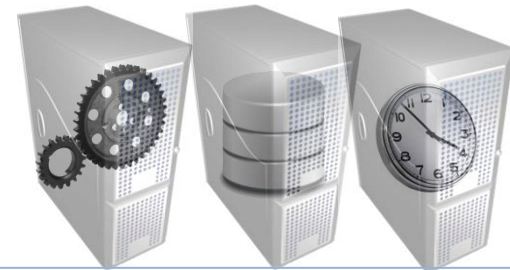
What Interfaces the Client Can Use



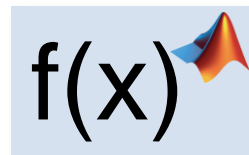
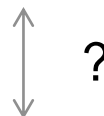
Desktop application



Web application



Database/web/application server

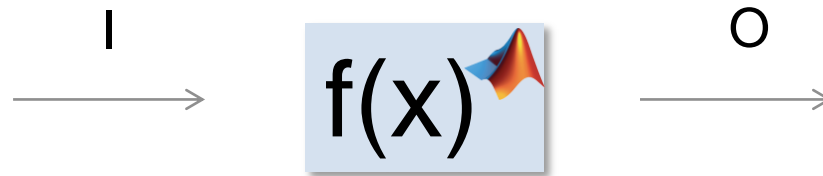


What “language” does the client speak?

Bottom line: MATLAB components can be called from any application that has external interfaces



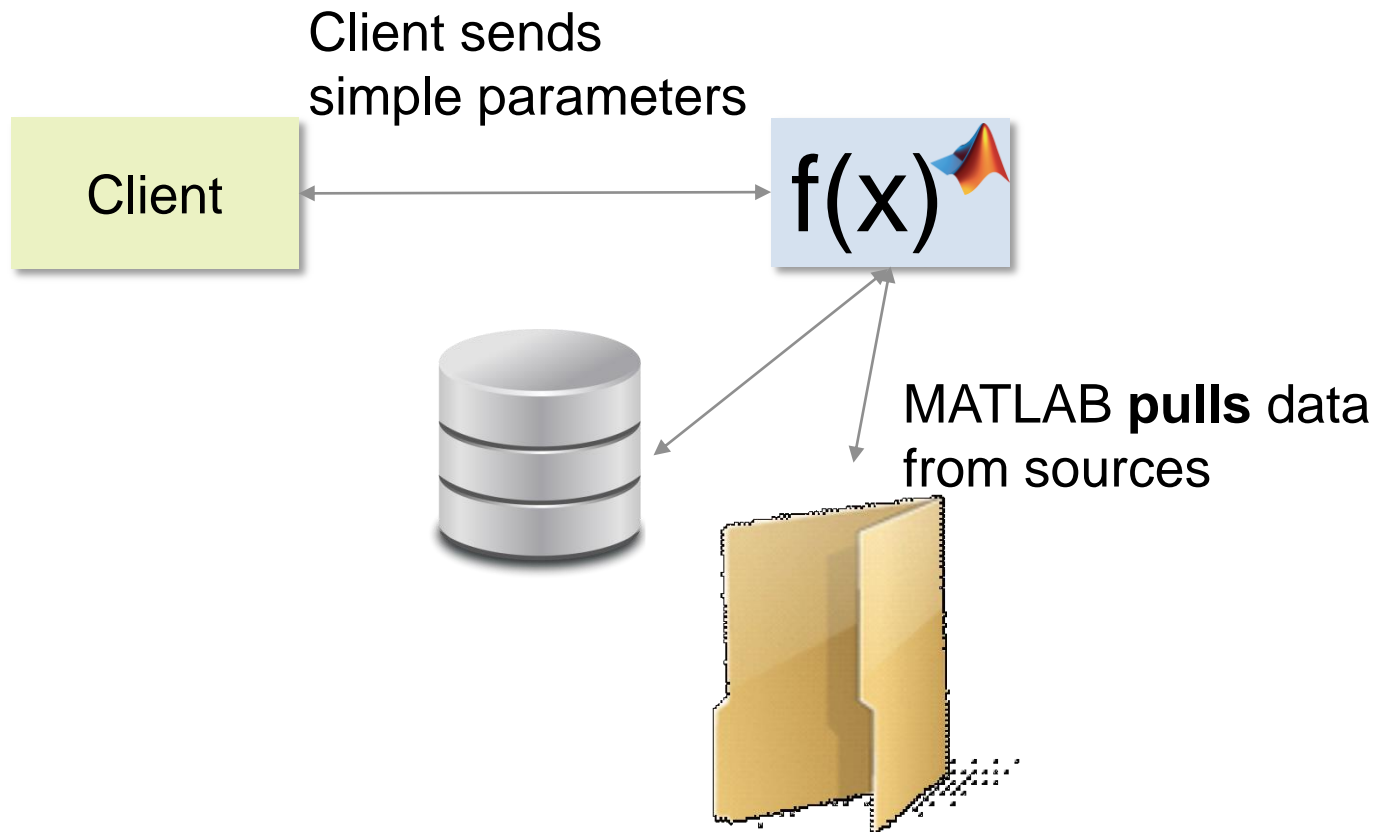
What Should Be the Input/Output?



What Should Be the Input/Output?

Example 1

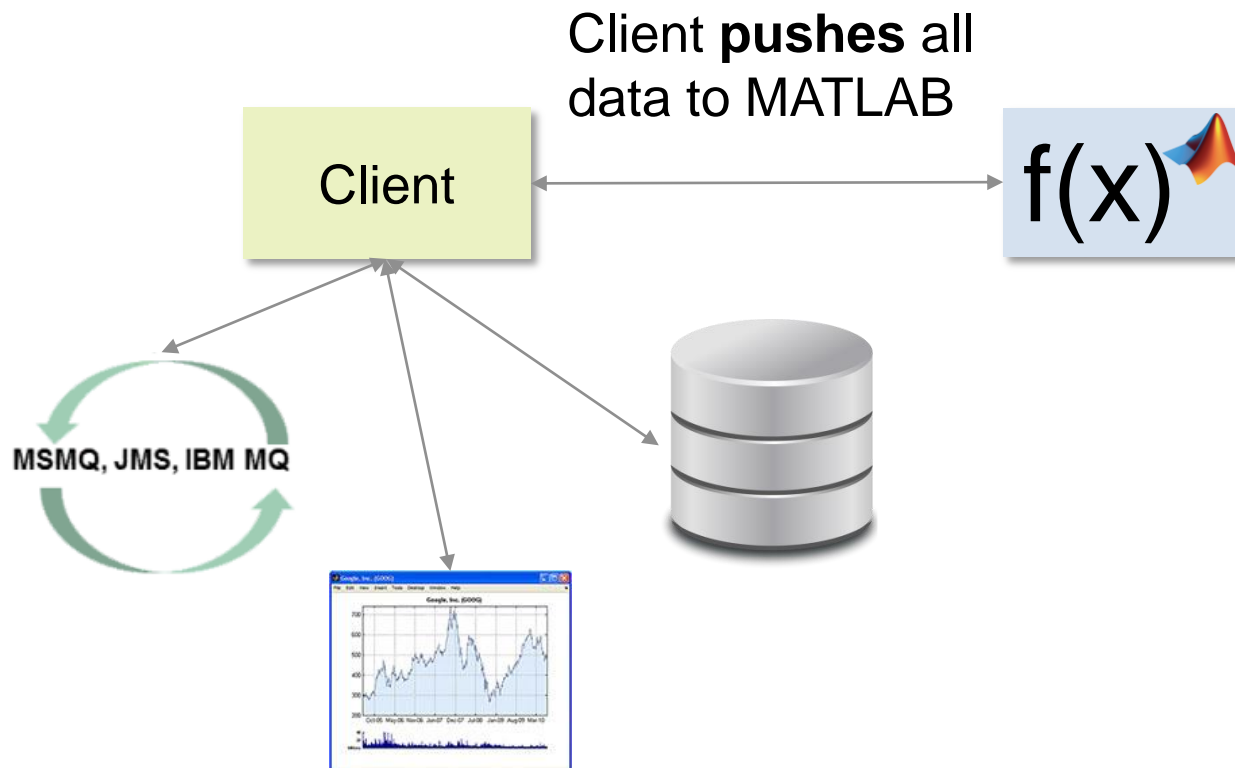
Deploy pricing functions into an enterprise pricing platform



What Should Be the Input/Output?

Example 2

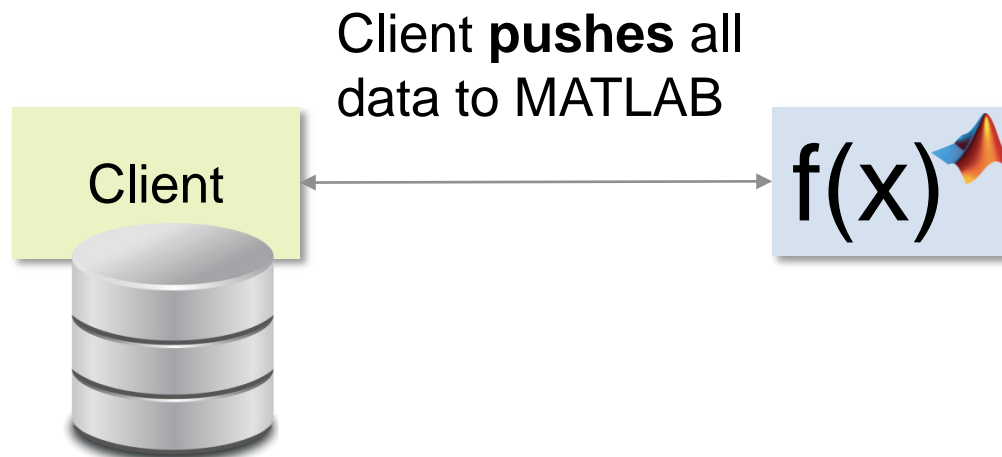
Add algorithms to a trading platform



What Should Be the Input/Output?

Example 3

Update prices nightly, or intraday, in an enterprise database



Bring algorithm to data
Examples



TERADATA

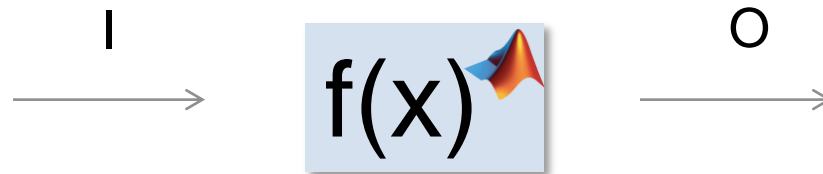
NETEZZA



hadoop

What Should Be the Input/Output?

Summary

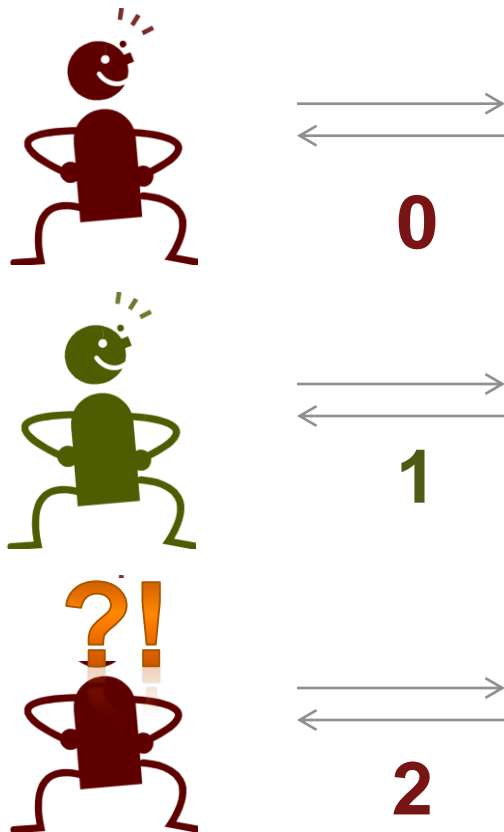


Main considerations

- What is needed
- Data types
- Is data on the client anyway?
- Need/want to use of existing enterprise I/O functions
- Marshalling time

What are Other Dependencies of the Algorithm?

State/persistent data



```
function out = getNextNumber
persistent n;
if isempty(n)
    n = 0;
end
n = n+1;
out = n;
```

What are Other Dependencies of the Algorithm?

```
function out = getNextNumber
persistent n;
if isempty(n)
    n = 0;
end
n = n+1;
out = n;
```

```
function out = getNextNumber(n)
out = n+1;
```

Removing state/persistent data

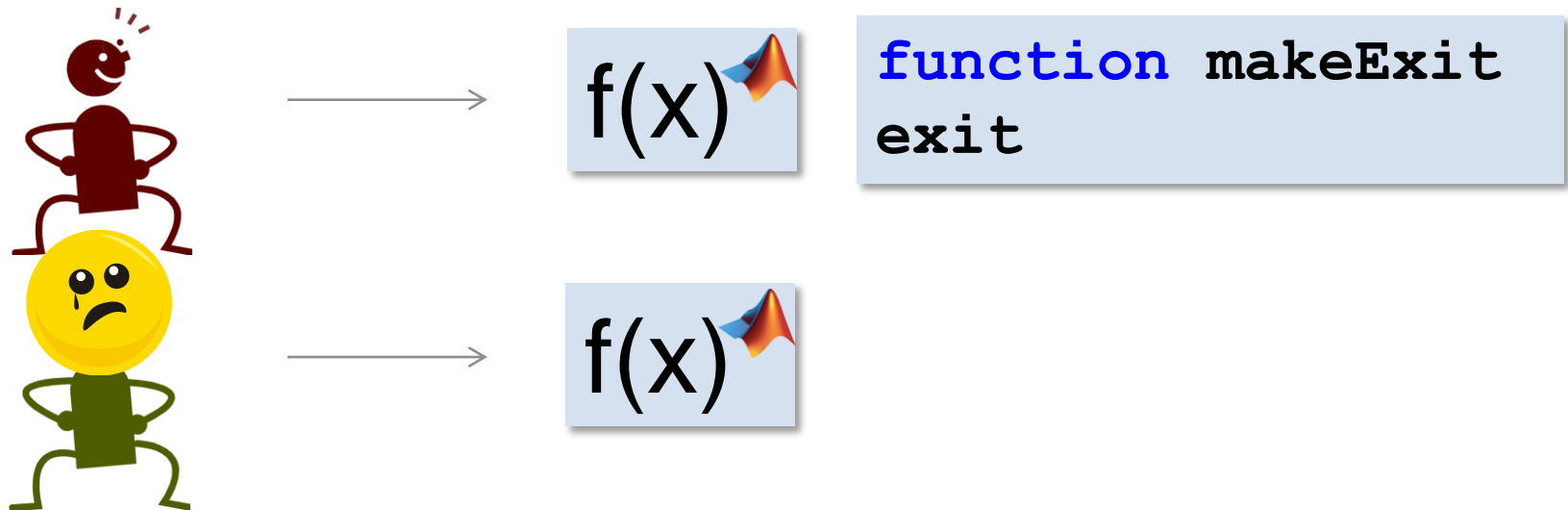
- Refactor
- Store state in
 - Database
 - MAT files
 - In-memory storage

Config files

- Make sure they will be present in the production environment

How to Keep the Algorithm Available

Imagine



Monitor and maintain the worker(s)

What is the Response Time

1 min

1min is fast for daily risk report

$f(x)$

1min is slow for interactive pricing

If the algorithm is not fast enough for the client

- Accelerate your code
- Warn the client
- Minimize call overhead
 - Reduce data marshalling
 - MCR start-up time
 - Algorithm closer to data

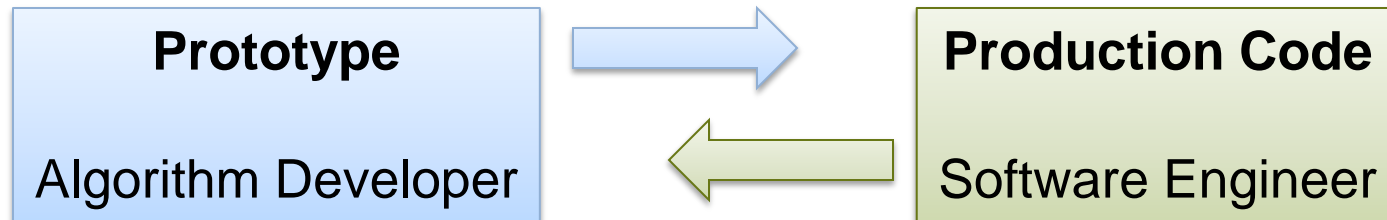


Off the Shelf Product Features for Server-Side Deployment

	MATLAB Compiler +Builders	MATLAB Production Server
Code locked	✓	✓
Logic out of process		✓
Reduced MCR start up time		✓
Simultaneous access		✓
Monitor and maintain the worker(s)		✓

✓ : Built in

Creating Production Ready MATLAB Algorithms



Internal collaboration
MathWorks Consulting Services

Using More Computational Power (Parallel Computing)

Parallel Computing Tools Address...

Task-Parallel

Long computations

- Multiple independent iterations

```
parfor i = 1 : n
    % do something with i
end
```

- Series of tasks

Task 1

Task 2

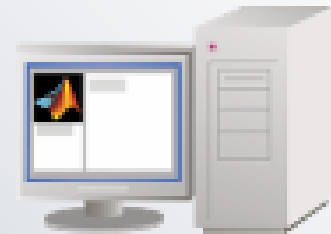
Task 3

Task 4

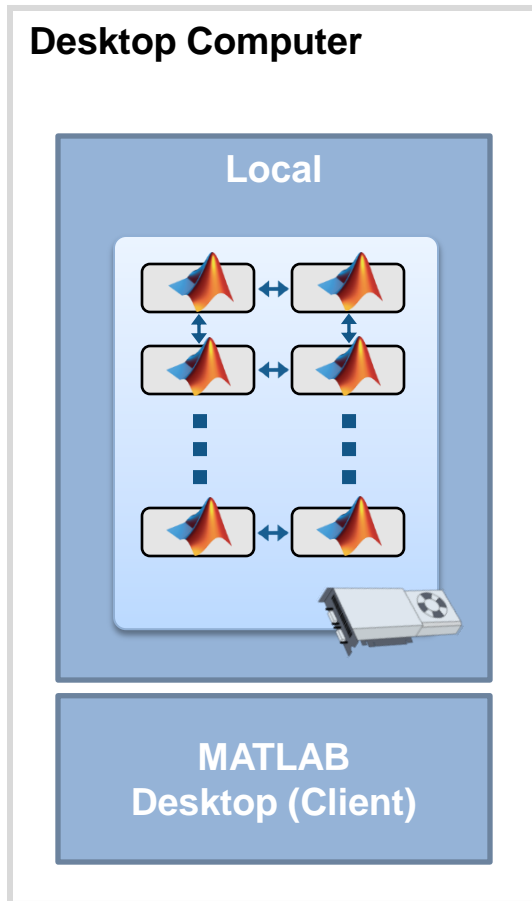
Data-Parallel

Large data problems

11	26	41
12	27	42
13	28	43
14	29	44
15	30	45
16	31	46
17	32	47
17	33	48
19	34	49
20	35	50
21	36	51
22	37	52

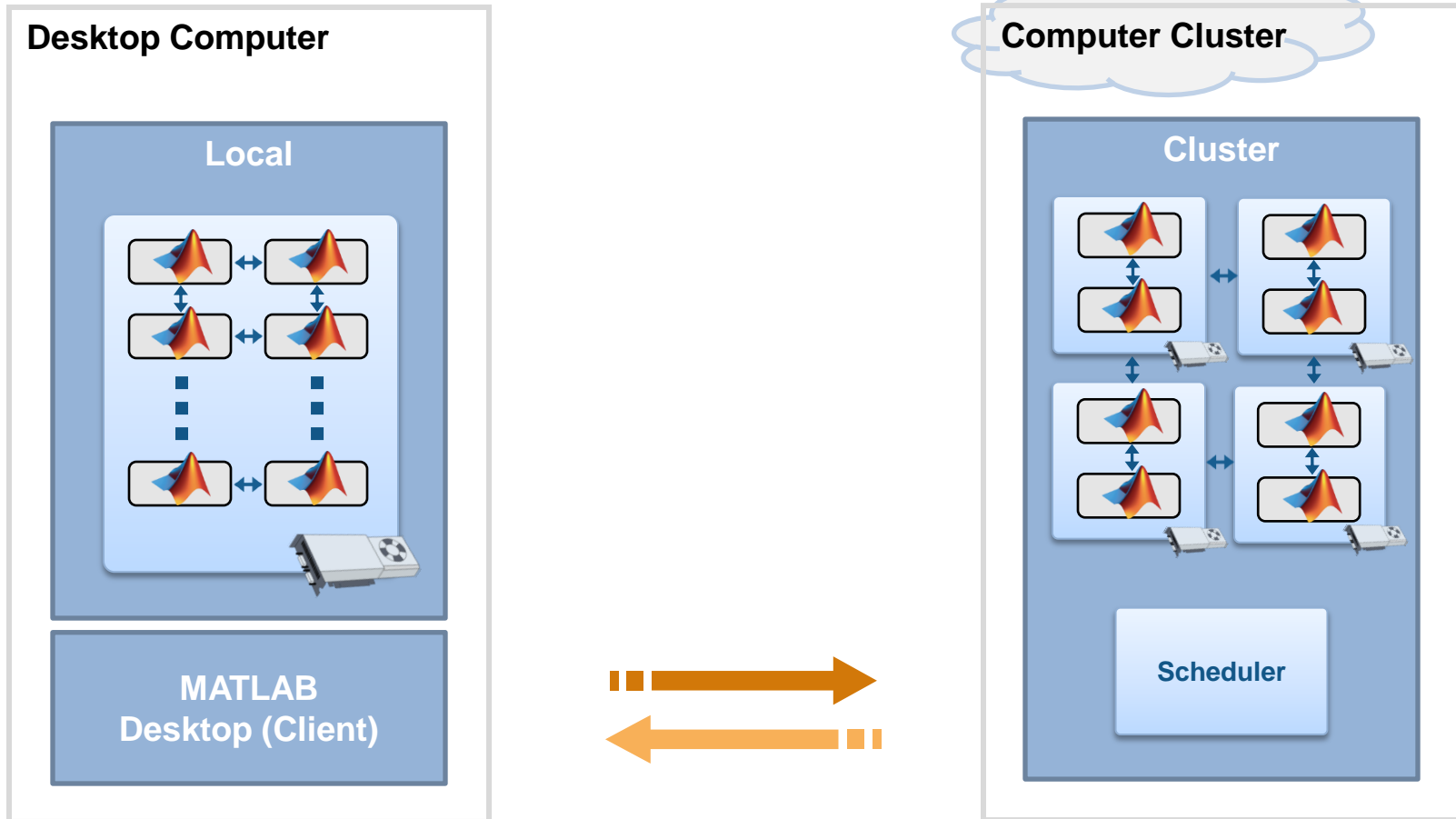


Parallel Computing Toolbox for the Desktop



- Speed up parallel applications
- Take advantage of GPUs
- Prototype code for your cluster

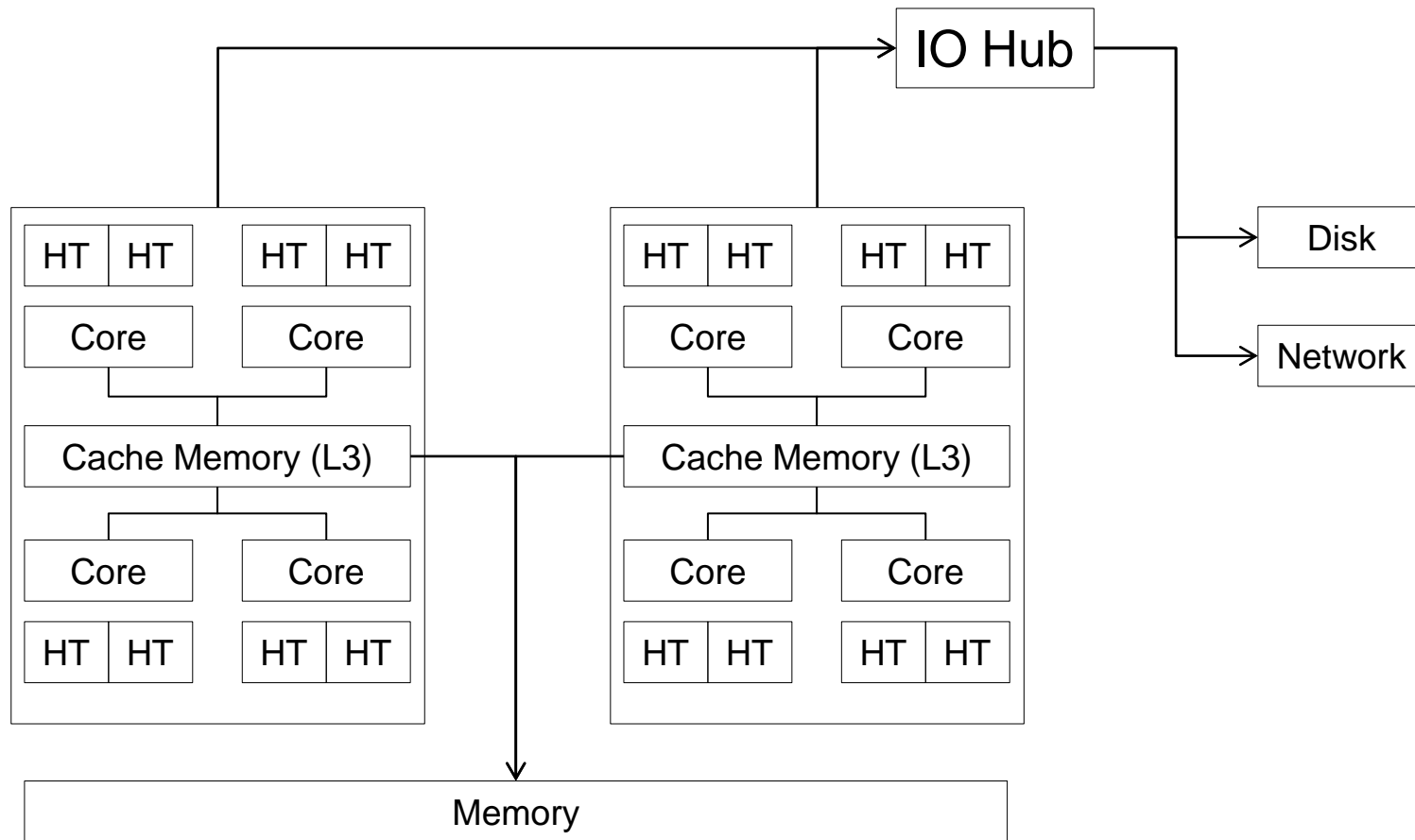
Scale Up to Clusters and Clouds



Main Considerations for Scalability

- Data transfer
- Resource contention
- Other overheads

Resource Contention



Speedup vs. num. Concurrent Processes

Performance

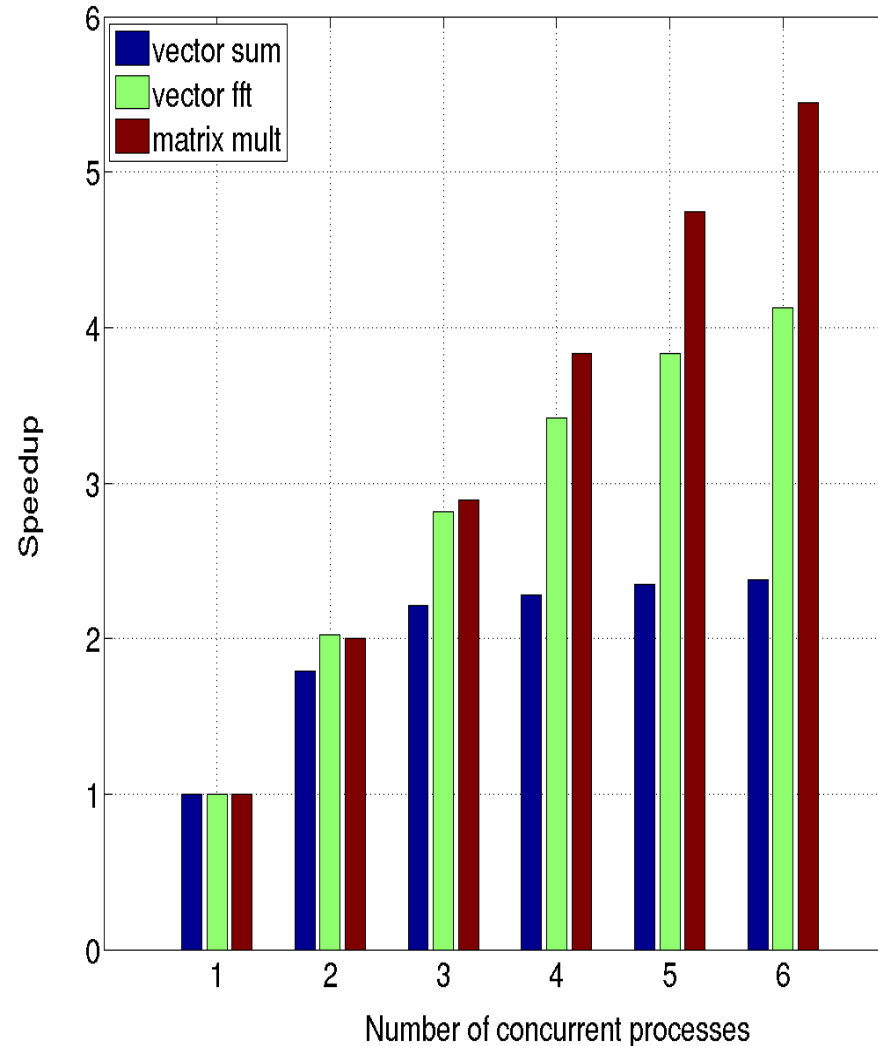
```
a = bigMatrix
```

```
a*a
```

```
fft(a)
```

```
sum(a)
```

Effect of number of concurrent processes on resource contention and speedup



Speedup vs. num. Concurrent Processes

Performance

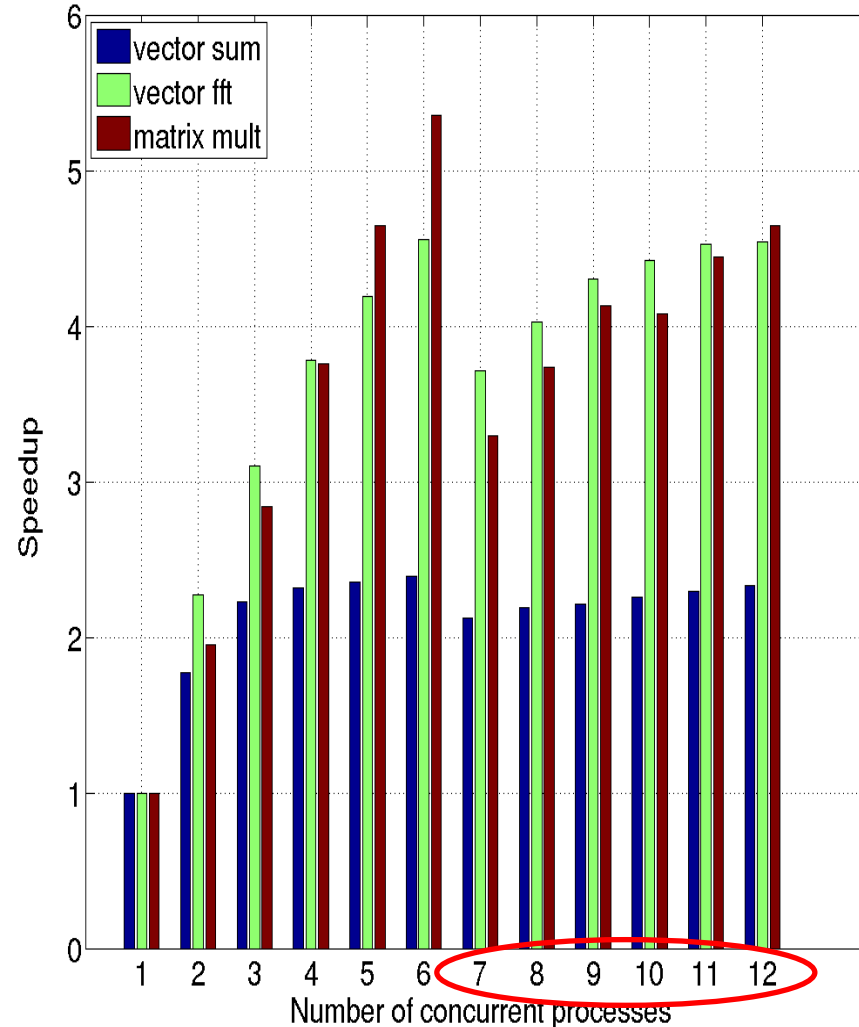
```
a = bigMatrix
```

```
a*a
```

```
fft(a)
```

```
sum(a)
```

Effect of number of concurrent processes on resource contention and speedup



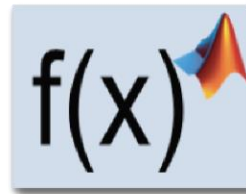
Hyperthreaded
Cores

What We Have Seen

More access



More
computational power



More Data



Scaling Up MATLAB Analytics

- Rapid prototyping of ideas
- Scalable algorithms
 - using more computational power
 - with wide access

