MATLAB in Business Critical Applications

Arvind Hosagrahara
Principal Technical Consultant
Arvind.Hosagrahara@mathworks.com
310-819-3970
Outline

- Problem Statement
- The Big Picture / Financial Applications
- Reference Architectures and integrations
  - TIBCO® Spotfire®
  - IBM® PureData™ for Analytics (powered by Netezza)
  - Microsoft® SQL Server®
  - Handling Big Data and out-of-memory problems (Hadoop, NoSQL databases)
- MATLAB in Business Critical Applications
- Conclusions
Problem Statement

- **Agility**: Is it possible to accelerate the development of data analytics?

- **Democratization**: Is it possible to make the results and insights from these analytics available to all stakeholders in an organization?

- **Production**: Is it possible to build data analytics algorithms in a flexible, scalable manner that is suitable for production usage and rigor?
Taking MATLAB into production

- Agility
- Lightweight processes
- Visualization
- Access to data
- Expressive high-level language
- Integration with best-in-class tooling

- Verification and validation
  - Correctness
  - Performance
- Automation of testing
- Test-driven development

- Performance
- Scalability
- Reliability
- Governance
- Auditability
- Maintenance
- Recoverability
- Support and Services
The Big Picture

- Reference architectures offer solutions to **finance industry applications** with an emphasis on selection of the right tool for the task leading to efficient workflows.
Finance Industry Applications

- Energy Trading
  - Load Forecasting
  - Model and Storage Assets
  - Price Energy Options
- Risk Modeling
- Portfolio Optimization
- ...

“The tools we developed with MATLAB are more reliable, scalable, and maintainable than our spreadsheet-based approach. We know the tools will work, we can add new capabilities, and we can update the production system without getting IT involved.” - Manuel Arancibia, Horizon Wind Energy

“MathWorks Consultants were well-qualified, professional, and fast. They understood not only the technical issues but also the business goals, which is essential when working on a core business system. We got more than we expected from MathWorks Consulting.” - Dr. Norbert Tönder, RWE
Reference Architecture

MATLAB Analytics with TIBCO Spotfire
MATLAB Analytics with TIBCO® Spotfire™
MATLAB Analytics with TIBCO Spotfire

- Rapidly create analytics for use in desktop, web, and mobile visualizations and dashboards
  - Run MATLAB scripts or programs within MATLAB sessions (small scale)
  - Access MATLAB analytics on MATLAB Production Server (large scale)
Reference Diagram
MATLAB Analytics Library

- MATLAB analytics available for use in Spotfire clients
Increasing Capacity and Redundancy
## Benefits

<table>
<thead>
<tr>
<th>Benefit</th>
<th>Tool</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rapid development and deployment of MATLAB analytics</td>
<td>MATLAB</td>
</tr>
<tr>
<td>Central management of locked down analytics</td>
<td>MATLAB Compiler SDK</td>
</tr>
<tr>
<td>Central management of locked down analytics</td>
<td>MATLAB Production Server</td>
</tr>
<tr>
<td>Scalable and reliable architecture</td>
<td></td>
</tr>
<tr>
<td>MATLAB Analytics available to all business users</td>
<td>TIBCO Spotfire Server</td>
</tr>
</tbody>
</table>

![Diagram showing MATLAB Production Server and MATLAB Analytics integration]
Reference Architecture

MATLAB Analytics with IBM® PureData™ / Netezza
MATLAB Analytics with IBM PureData

Off-Database
- Pure Data ODBC/JDBC Driver
  - ODBC
  - JDBC
  - Database Toolbox

On-Database
- IBM PureData (Netezza)

Low Performance “Processes” (UDAP or AE)
- Python
- Java
- MATLAB
- MDCS
- MPS
- MCR
- MATLAB Compiler

High-Performance “Function” (UDF / UDTF / UDA)
- C++
- Lua
- MATLAB Coder

Workflow
MATLAB Analytics with IBM PureData
Building and Running the MATLAB Analytics

1. Generate Code and move to the IBM PureData Appliance

2. Compile the generated code into a shared object for the host and the SPU

3. Register the shared object as a library

4. Use the MATLAB analytic in the User-Defined Function / Table Function / Aggregate [ SELECT Histogram(*) ]
## Benefits

<table>
<thead>
<tr>
<th>• Rapid development and deployment of MATLAB analytics</th>
<th>MATLAB</th>
</tr>
</thead>
<tbody>
<tr>
<td>• High Performance Analytics written in MATLAB and generated as C++</td>
<td>MATLAB Coder and/or Simulink Coder with Embedded Coder</td>
</tr>
<tr>
<td>• MATLAB Analytics available to downstream users as a Stored Procedure/Function</td>
<td>IBM PureData for Analytics powered by Netezza or Microsoft Netezza</td>
</tr>
</tbody>
</table>
Reference Architecture

MATLAB Analytics with Microsoft SQL Server
MATLAB Analytics with Microsoft SQL Server

Off-Database

- ODBC/JDBC Driver
  - ODBC
  - JDBC
  - Database Toolbox

On-Database

- "Processes"
  - C# Common Language Runtime (CLR)
    - .NET Assembly
    - MATLAB Component Runtime
  - MPS Client Library
- "Transactions"
  - MATLAB Component Runtime
  - MATLAB Compiler SDK

MATLAB Production Server

MATLAB Analytics
MATLAB Analytics with Microsoft SQL Server

MATLAB Production Server

1. sp_configure 'clr enabled', 1;
   GO
   RECONFIGURE;
   GO

2. CREATE ASSEMBLY MATLABClient
   AUTHORIZATION db0
   FROM "C:\Program Files\MATLAB\MATLAB Production Server\Local\R2015a\client\dotnet\MathWorks.MATLAB.ProductionServer.Client.dll"
   WITH PERMISSION_SET = UNSAFE
   GO

3. using MathWorks.MATLAB.ProductionServer.Client;
   MKClient client = new MXHttpClient();
   try
   {
     // Attempt to connect to MPS and call the MATLAB functionality
     [messagingproxy("http://localhost:9910/RSDensity")]
     RS DensityDouble mSd = client.CreateProxy<RS DensityDouble>();
     mSd = mSd.RSDensityDemo(); // Call the method
     // Marshal and return the results
   }
## Benefits

<table>
<thead>
<tr>
<th>Benefits</th>
<th>Tools</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Rapid development and deployment of MATLAB analytics</td>
<td>MATLAB</td>
</tr>
<tr>
<td>• Analytics expressed in MATLAB with nearly all available toolbox</td>
<td>MATLAB Compiler SDK, MATLAB</td>
</tr>
<tr>
<td>functionality</td>
<td>Production Server</td>
</tr>
<tr>
<td>• MATLAB Analytics available to downstream users as a Stored Procedure/Function</td>
<td>Microsoft SQL Server</td>
</tr>
</tbody>
</table>
Big Data with MATLAB
Techniques for Big Data in MATLAB

Load, Analyze, Discard
parfor, datastore,

MapReduce

Distributed Memory
SPMD and distributed arrays

Embarrassingly Parallel

Non-Partitionable

out-of-memory
in-memory

Complexity
Big Data Capabilities in MATLAB

Memory and Data Access
- 64-bit processors
- Memory Mapped Variables
- Disk Variables
- Databases
- Datasets

Programming Constructs
- Streaming
- Block Processing
- Parallel-for loops
- GPU Arrays
- SPMD and Distributed Arrays
- MapReduce

Platforms
- Desktop (Multicore, GPU)
- Clusters
- Cloud Computing (MDCS on EC2)
- Hadoop
MATLAB in Business Critical Applications
Conclusions: MATLAB Analytics in Production

- Access and explore data from within MATLAB during prototype development
- Develop and integrate analytics with enterprise systems
  - Compact, precise language for expression of complex analytics
  - High-quality libraries and support
- Open and extensible platform for analytics
  - Data Visualization: TIBCO Spotfire,
  - Hadoop: Apache
  - Data warehouses: IBM PureData, Microsoft SQL Server
- MATLAB products and Services provide a single-stack solution when used with supporting technologies to address production data analytics demands.