Leveraging the Cloud to Architect Digital Solutions
Agenda

- State of the art technology for IoT
- Table Storage Demo
- ML Clustering and Classification prototype
- Cortana Analytics
- Architecture for building today
- Patterns and anti-patterns
- Open Discussion
What is Digital?

Digital is more than technology.

It involves connecting that technology with the right data science, devices, design and business strategy.

It is putting a customer, device, organization or business process at the center of real change in how businesses do things and how customers experience them. In how we engage, invent, build and buy everything.

It means creating value by uniting the physical world—seamlessly, efficiently, meaningfully—to the virtual one we’re building.
# Microsoft Azure IoT services

<table>
<thead>
<tr>
<th>Devices</th>
<th>Device Connectivity</th>
<th>Storage</th>
<th>Analytics</th>
<th>Presentation &amp; Action</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Event Hubs</td>
<td>SQL Database</td>
<td>Machine Learning</td>
<td>App Service</td>
</tr>
<tr>
<td></td>
<td>Service Bus</td>
<td>Table/Blob Storage</td>
<td>Stream Analytics</td>
<td>Power BI</td>
</tr>
<tr>
<td>External Data Sources</td>
<td>External Data Sources</td>
<td>DocumentDB</td>
<td>HDInsight</td>
<td>Notification Hubs</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Data Factory</td>
<td>Mobile Services</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BizTalk Services</td>
</tr>
</tbody>
</table>

**Microsoft Azure IoT services**

- **Event Hubs**: Connect devices and applications through communication services.
- **SQL Database**: Store data from devices and applications.
- **Machine Learning**: Analyze and predict insights from IoT data.
- **App Service**: Offer a mobile and web experience.
- **Service Bus**: Message and data transport between IoT devices and applications.
- **Table/Blob Storage**: Store and retrieve data from IoT devices.
- **Stream Analytics**: Analyze live data streams.
- **Power BI**: Create interactive dashboards and reports.
- **External Data Sources**: Integrate data from non-IoT sources.
- **DocumentDB**: NoSQL database for IoT data.
- **HDInsight**: Big data analytics service.
- **Notification Hubs**: Push notifications to applications.
- **Data Factory**: Data integration and preparation service.
- **Mobile Services**: Offer mobile and web applications.
- **BizTalk Services**: Integration and orchestration services.
<table>
<thead>
<tr>
<th>&gt;90,000</th>
<th>&gt;1.5 Million</th>
<th>&gt;500 Million</th>
</tr>
</thead>
<tbody>
<tr>
<td>New Azure customer subscriptions/month</td>
<td>SQL Databases running on Azure</td>
<td>Users in Azure Active Directory</td>
</tr>
</tbody>
</table>

Cloud Momentum

<table>
<thead>
<tr>
<th>1.5 Trillion</th>
<th>777 Trillion</th>
<th>&gt;40%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Messages per month processed by Azure IoT</td>
<td>Storage Transactions per day</td>
<td>Revenue from Start-ups and ISVs</td>
</tr>
</tbody>
</table>
Managing Big Data

Azure IoT Suite

Device Connectivity and Security
Data Ingestion and Command & Control
Stream Processing & Predictive Analytics
Workflow Automation
Dashboards and Visualization
Preconfigured Solutions
Is IoT even a new thing?

Depending on who you ask, IoT is either:

1. Nothing new
   "We’ve been doing this for 40 years"

2. A unicorn
   Magic, and will soon change everything.

Legacy applications incorporate some ideas of a modern IoT application

But falling hardware costs, cloud services and relatively ubiquitous communications do enable new approaches
Emerging Challenges for IT

• **Scale**
  • # devices >> # users, and growing fast
  • Volume of data (and network traffic)

• **Pace**
  • Innovation pressure: analysis, command and control, cost
  • Skill pressure: data science, new platforms

• **Environment**
  • IT/OT collaboration
  • Security and privacy threats
  • Emerging standards
  • New competitors
Azure Machine Learning Conceptual Model

Azure Machine Learning: Basic workflow
Build models from data and operationalize a machine learning solution

Data collection & management
- Blobs & tables (Azure Storage)
- Hadoop (Azure HDInsight)
- Relational data (Azure SQL Database)
- Massive data stores (Azure Data Lake)

Machine Learning service
- ML Studio web app: Create machine learning models
- Output a web service that can be run on a scheduled basis and connected to a database.

Embedded ML model
- Add intelligence to apps or websites, or provide insights in BI tools
Data scientist
A highly educated and skilled person who can solve complex data problems by employing deep expertise in scientific disciplines (mathematics, statistics or computer science)

Data professional
A skilled person who creates or maintains data systems, data solutions, or implements predictive modelling
Roles: Database Administrator, Database Developer, or BI Developer

Software developer
A skilled person who designs and develops programming logic, and can apply machine learning to integrate predictive functionality into applications
Pattern: Think big. Start small

- Think big, but start small.
- Experiment, learn and refine.

- Build to an architecture that will scale, but start prototyping with a small number of devices.
- It’s hard to predict what data provides value -- which impacts which sensors and devices are necessary -- until you build something.
- The options can be overwhelming: set crisp goals up front and use those to define and refine.
- It’s much easier to work through device identity, management/update and security at small scale.
Pattern: Telemetry first

Start with telemetry.

The important data may not be what you expected.

Address privacy, management and security before command & control.

- It is very hard to predict in advance what data will be useful.
- It is tempting, but likely inefficient to try for business transformation in the first step.
- Think about not only device telemetry but also diagnostic telemetry.
- Privacy and security implications of telemetry are generally lesser than for command and control.
Telemetry today

- High scale data ingestion stream processing
- Storage for cold-path analytics
- Processing for hot-path analytics
Azure Table Storage
Storage

**TABLES:**
Provide structured storage. A table is a set of entities which contain a set of properties.

**BLOBS:**
Provide a simple interface for storing named files along with metadata for the file.

**QUEUES:**
Provide reliable storage and delivery of messages for an application.

**SQL Azure:**
A relational database hosted in a MS data center, triply replicated.
# Traditional RDBMS vs. Table Storage

<table>
<thead>
<tr>
<th>Data Size</th>
<th>TRADITIONAL RDBMS</th>
<th>TABLE STORAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Structure</td>
<td>Gigabytes (Terabytes)</td>
<td>Petabytes (Hexabytes)</td>
</tr>
<tr>
<td>Integrity</td>
<td>High (ACID)</td>
<td>Low</td>
</tr>
<tr>
<td>Scaling</td>
<td>Nonlinear</td>
<td>Linear</td>
</tr>
<tr>
<td>DBA Ratio</td>
<td>1:40</td>
<td>1:3000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Value</th>
<th>Metric</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td>kB</td>
</tr>
<tr>
<td>1000²</td>
<td>MB</td>
</tr>
<tr>
<td>1000³</td>
<td>GB</td>
</tr>
<tr>
<td>1000⁴</td>
<td>TB</td>
</tr>
<tr>
<td>1000⁵</td>
<td>PB</td>
</tr>
<tr>
<td>1000⁶</td>
<td>EB</td>
</tr>
<tr>
<td>1000⁷</td>
<td>ZB</td>
</tr>
<tr>
<td>1000⁸</td>
<td>YB</td>
</tr>
</tbody>
</table>
Entity Properties

• Entity can have up to 255 properties
  – Up to 1MB per entity

• Mandatory Properties for every entity
  – PartitionKey & RowKey (only indexed properties)
    – Uniquely identifies an entity
    – Defines the sort order
  – Timestamp
    – Optimistic Concurrency
    – Exposed as an HTTP Etag

• No fixed schema for other properties
  – Each property is stored as a <name, typed value> pair
  – No schema stored for a table
  – Properties can be the standard .NET types
  – String, binary, bool, DateTime, GUID, int, int64, and double
Purpose of the PartitionKey

- **Entity Locality**
  - Entities in the same partition will be stored together
  - Efficient querying and cache locality
  - Endeavour to include partition key in all queries

- **Entity Group Transactions**
  - Atomic multiple Insert/Update/Delete in same partition in a single transaction

- **Table Scalability**
  - Target throughput – 500 tps/partition, several thousand tps/account
  - Windows Azure monitors the usage patterns of partitions
  - Automatically load balance partitions
    - Each partition can be served by a different storage node
    - Scale to meet the traffic needs of your table
Azure Machine Learning: Basic workflow

Build models from data and operationalize a machine learning solution

Data collection & management
- Blobs & tables (Azure Storage)
- Hadoop (Azure HDInsight)
- Relational data (Azure SQL Database)
- Massive data stores (Azure Data Lake)

Machine Learning service
- ML Studio web app: Create machine learning models
- Output a web service that can be run on a scheduled basis and connected to a database.

Embedded ML model
- Add intelligence to apps or websites, or provide insights in BI tools
Azure IoT Reference Architecture

1. **Devices and Data Sources**
   - IP capable devices
   - Existing IoT devices
   - Low power devices

2. **Data Transport**
   - Agent Libs

3. **Device and Event Processing**
   - Provisioning API
   - Solution Portal
   - Identity & Registry Stores
   - Device State Store
   - Stream Event Processor
   - Storage
   - Analytics/Machine Learning
   - Control System Worker Role

4. **Presentation**
   - Data Visualization & Presentation
Demo: ML Clustering and Classification of Data

Premise:
Wireless carrier develops a prototype of a public cloud service that can track IP connected phones, tablets, and smart watches.

1) Table Storage
2) Clustering Model
3) Classification Model

http://tableclientdemo.azurewebsites.net/
Table Storage Sharding
Cortana Analytics Suite
Azure Machine Learning Conceptual Model

Azure Machine Learning: Basic workflow

Build models from data and operationalize a machine learning solution

Data collection & management
- Blobs & tables (Azure Storage)
- Hadoop (Azure HDInsight)
- Relational data (Azure SQL Database)
- Massive data stores (Azure Data Lake)

Machine Learning service
- ML Studio web app:
  - Create machine learning models
- Web service:
  - Output a web service that can be run on a scheduled basis and connected to a database.

Embedded ML model
- Add intelligence to apps or websites, or provide insights in BI tools
### Data, Information, Knowledge, Wisdom Hierarchy

<table>
<thead>
<tr>
<th>Look to the Past</th>
<th>Learning</th>
<th>Look to the Future</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gathering</strong></td>
<td><strong>Learning</strong></td>
<td><strong>Doing</strong></td>
</tr>
<tr>
<td>Know-Nothing</td>
<td>Know-What</td>
<td>Know-Why</td>
</tr>
<tr>
<td>Data (numbers, symbols, facts)</td>
<td>Theory (a framework for explaining behavior)</td>
<td>Understanding (assign meaning, explain why, apply to analysis)</td>
</tr>
<tr>
<td></td>
<td>Information (data processed and organized to describe who, what, where and when)</td>
<td>Knowledge (collected information, instructs how)</td>
</tr>
<tr>
<td></td>
<td>Experience (knowledge gained through doing)</td>
<td></td>
</tr>
<tr>
<td><strong>Extrinsic</strong></td>
<td><strong>Intrinsic</strong></td>
<td><strong>Executing</strong></td>
</tr>
<tr>
<td>Analyzing</td>
<td>Deciding</td>
<td>Executing</td>
</tr>
<tr>
<td>Control and Efficiency</td>
<td>Value and Effectiveness</td>
<td></td>
</tr>
<tr>
<td>Doing things right</td>
<td>Doing the right things</td>
<td></td>
</tr>
</tbody>
</table>
Cortana Analytics Suite
Transform data into intelligent action

- Information Management
  - Azure Data Factory
  - Azure Data Catalog
  - Azure Event Hub

- Big Data Stores
  - Azure Data Lake
  - Azure SQL Data Warehouse

- Machine Learning and Analytics
  - Azure Machine Learning
  - Azure HDInsight (Hadoop)
  - Azure Stream Analytics

- Dashboards and Visualizations
  - Power BI

- Personal Digital Assistant
  - Cortana

- Perceptual Intelligence
  - Face, vision
  - Speech, text

- Business Scenarios
  - Recommendations, customer churn, forecasting, etc.

DATA → INTELLIGENCE → ACTION
Azure Machine Learning Conceptual Model

Azure Machine Learning: Basic workflow

Build models from data and operationalize a machine learning solution

Data collection & management
- Blobs & tables (Azure Storage)
- Hadoop (Azure HDInsight)
- Relational data (Azure SQL Database)
- Massive data stores (Azure Data Lake)

Machine Learning service
- ML Studio web app: Create machine learning models
- Output a web service that can be run on a scheduled basis and connected to a database.

Embedded ML model
- Add intelligence to apps or websites, or provide insights in BI tools
Visualizing ML Results in Power BI
The need to know what could be…
Deriving Business Value from Big Data
Machine learning & predictive analytics are core capabilities that can help create added business value.

Using past data to predict the future

- Churn analysis
- Social network analysis
- Recommendation engines
- Location-based tracking and services
- Vision Analytics
- Weather forecasting for business planning
- Legal discovery and document archiving
- Equipment monitoring
- Advertising analysis
- Pricing analysis
- Fraud detection
- Personalized Insurance

Machine learning & predictive analytics are core capabilities that can help create added business value.
Summary

- Think big (architecture), but start small (experiment, learn and refine).
- Start with telemetry. Address privacy, security and manageability before moving to command and control.
- Don’t interrupt the fast path and create processing bottlenecks.
- Build to the reference architecture to ease the move to IoT Suite.
Future Topics

• Cortana Analytics Suite
• Azure Portal
• Scalable Cloud Applications
  • PaaS development
• Event Hub
• Stream Analytics
• Machine Learning Studio
...rabbits eat grass...
grass grows in soil...
rain falls from clouds...
wolves eat rabbits...

etc... etc...

if I kill off all the wolves, the rabbits will eat up the grass, & the soil will all wash away!

INFORMATION...

KNOWLEDGE...

...WISDOM!