WINDOWS AZURE MARKETPLACE
DATAMARKET OVERVIEW

Christian Liensberger
Program Manager
cliens@microsoft.com
Why DataMarket?

Simple distribution

Information marketplace

Data providers

Easy discovery and sign up

Developers
IWIs
ISVs
Exploring DataMarket
DataMarket APIs

- OData-based APIs for access to the service
  - Query language over HTTP
  - Standardized metadata
  - Server-based filtering and sorting
  - Pagination for large result sets
  - Open standard (www.odata.org)
    - Support for a lot of languages, including C#, VB.NET, PHP, Objective C, etc.
    - ATOM feed as response

- Authentication using DataMarket account key
  - HTTP basic authentication
  - Secured through SSL
Consuming a DataMarket Dataset
DataMarket Architecture

DataMarket infrastructure

Web roles for exposed datasets (web services / public DataMarket APIs)

Web roles for marketplace website (ASP.NET)

Public APIs

Windows Azure compute infrastructure (VMs)

SQL Azure elastic databases

Windows Azure AppFabric Access Control System (ACS)
Processing user requests

Public API endpoints

Work dispatcher

Prioritized queue

Dataset worker

Dataset

Prioritized queue

Dataset worker

Dataset

Prioritized queue

Dataset worker

Dataset
Reaching out to content provider’s data

Data worker / Mapping layer

Data source connector

SQL Azure connector

Web service connector

3rd party cloud connector

SQL Caching layer

Data in database

Data as web service

Data in 3rd party cloud

Mapping store

CSDL

OData
Scaling out user databases

System data processing pipeline

Database access management

Partitioning function

Partition 2

System database Part1

System database Part2

System database Part3
Monitoring

- Discover issues before customer!
- Build monitoring as Windows Azure web role
  - Deployed to multiple data centers (leverage Azure infrastructure)
  - Executes sample queries on regular base
  - Stores result in Windows Azure Table
- Logs transferred on a regular basis to SQL Azure
  - Reporting services to create reports
  - Alerting on condition matching
Securing high value data

- **Problem:** how to secure 3rd party connection information?
- **Solution:** Certificates!
  - Windows Azure has dedicated certificate pipeline
  - Only small number of people need access to them
  - Connection information encrypted with final key during publication process
  - Certificates to decrypt connection information on the fly
  - Connecting to data sources through SSL
So why did we do this?
How do I buy it?

- Free and commercial datasets
- Content provider sets the price
- Two supported subscription models
  - Unlimited number of queries per month
  - Limited number of queries per month and minute
- Credit card payment
• You are in control of the
  • data
  • price
  • terms of use
• Easy publishing - we create the consumer facing APIs
  • Connect to SQL Azure
  • Wrap existing web services
• Receive periodic reports and payouts
Discover DataMarket datasets from Office
Questions?