Cloud and Big Data at Microsoft

Wenming Ye
Sr. Research Program Manager
Microsoft Research Connections
Twitter: @wenmingye

Agenda

- The Microsoft Cloud: Windows Azure!
- Windows Azure Storage
- HDInsight: Windows Azure + Hadoop
- Other Big Data Tools on Windows Azure

Cloud Computing

IaaS

Infrastructure-as-a-Service

host

PaaS

Platform-as-a-Service

build

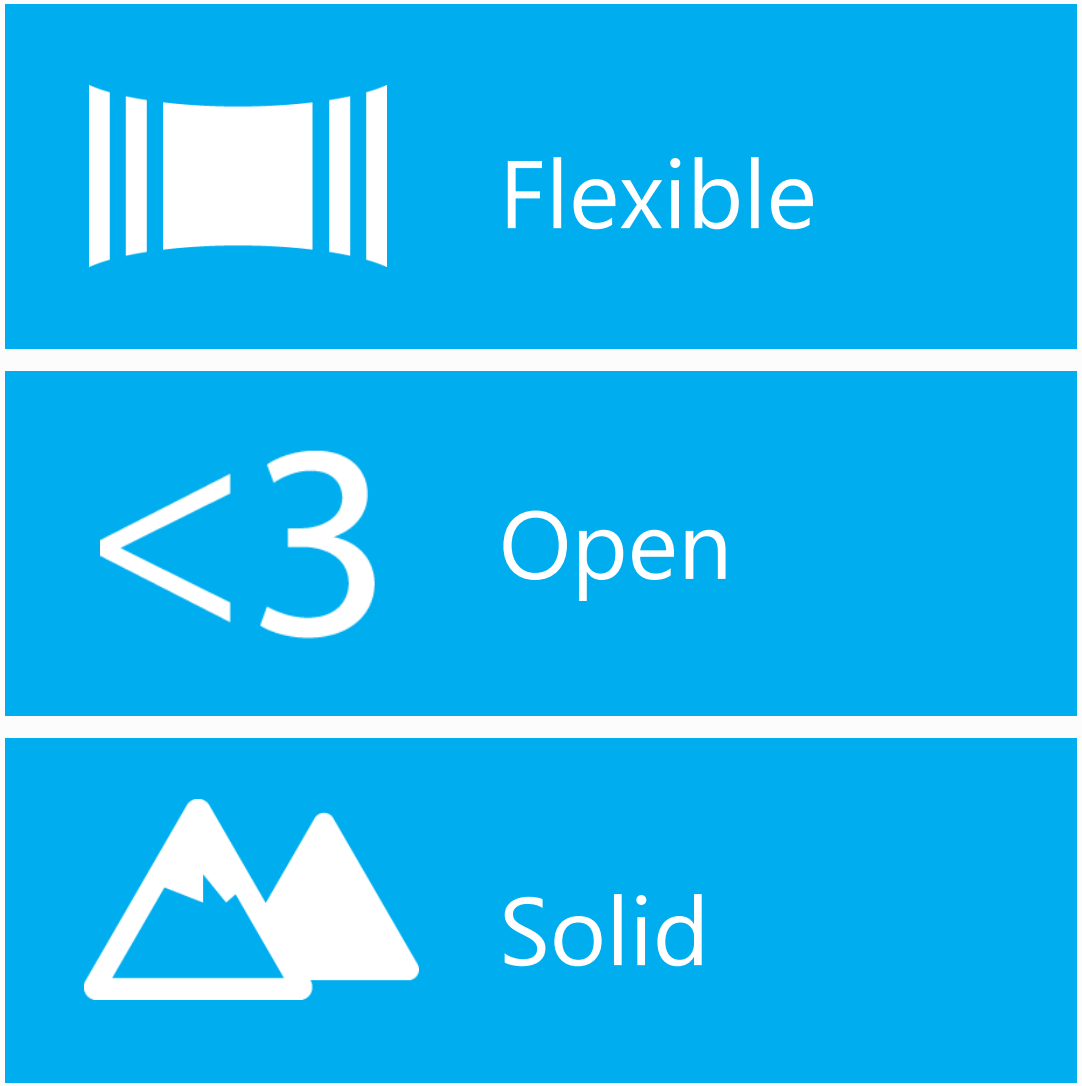
SaaS

Software-as-a-Service

consume

Windows Azure

Comprehensive set of services that enable you to quickly build, deploy and manage applications across a global network of Microsoft-managed datacenters



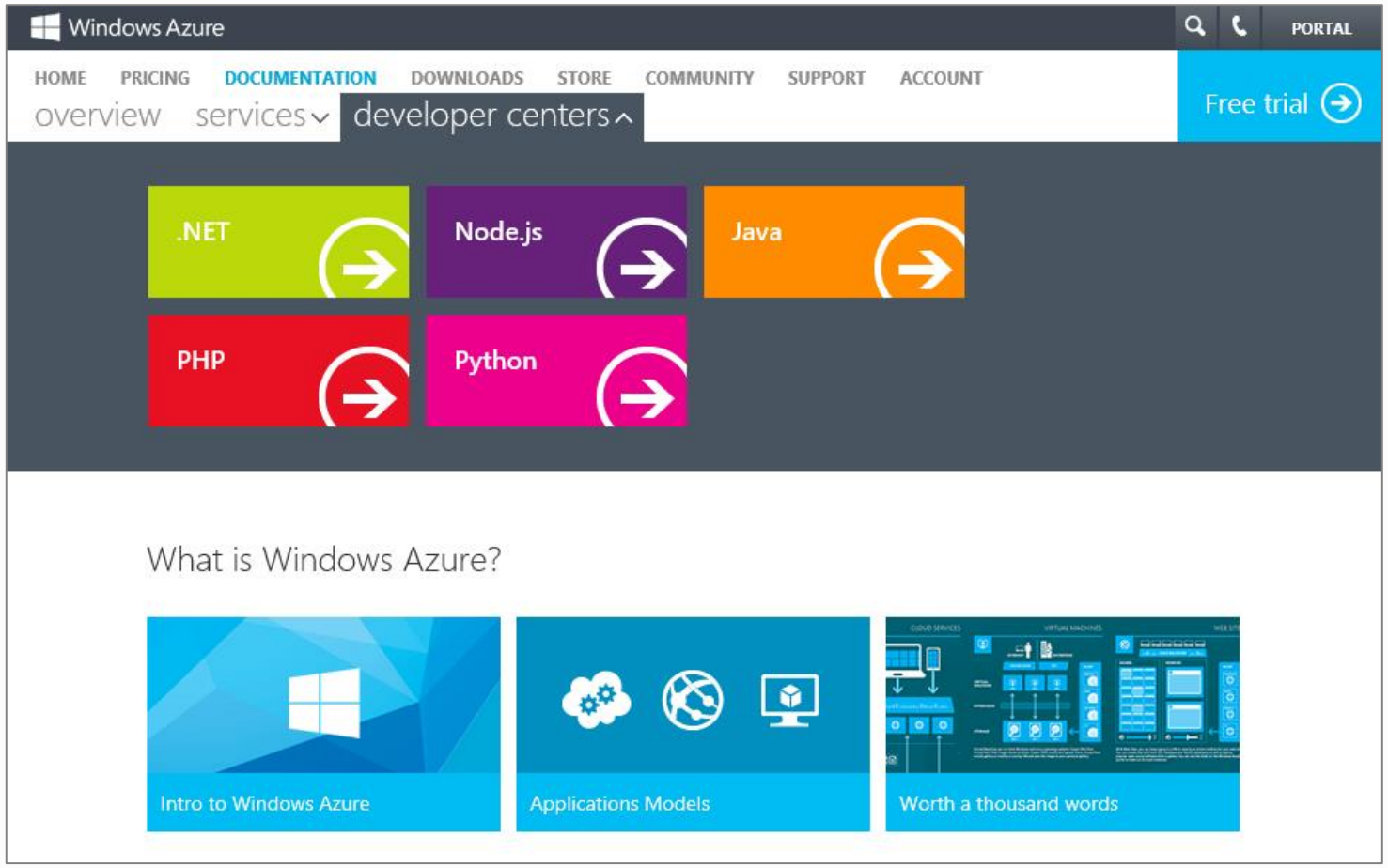
Global Footprint

Windows Azure

Virtual machines

Cloud services

Web sites



Multiple languages

→ http://WindowsAzure.com

Windows Azure

Virtual machines

Windows Server and Linux

Flexible Workload Support

Virtual Private Networking

Virtual machine operating system selection

- ALL
- PLATFORM IMAGES
- MY IMAGES
- MY DISKS

- Windows Server 2008 R2 SP1
- Windows Server 2012 Datacenter
- OpenLogic CentOS 6.3
- openSUSE 12.3
- RightScale Linux v13
- SUSE Linux Enterprise Server 11 SP2
- Ubuntu 12.04
- Ubuntu 12.10
- wenmingsaved
- whitehall
- boothdemo1-boothdemo1-0-2012

Microsoft SQL Server...

SQL Server 2012 SP1 Cumulative Update 2 Evaluation Edition (64-bit) on Windows Server 2008 R2 Service Pack 1. Virtual Machines created by using this SQL Server Evaluation Edition will expire on August 20, 2013. This image contains the full version of SQL Server. Some SQL Server 2012 components require additional setup and configuration before use. Medium is the minimum recommended virtual machine size for this image. To evaluate the advanced capabilities of SQL Server 2012, we recommend that you use a virtual machine size of Large or Extra Large.

PUBLISHER	Microsoft SQL Server Group
OS FAMILY	Windows
LOCATION	East Asia;Southeast Asia;North Europe;West Europe;East US;West US

2 3 4

CREATE A VIRTUAL MACHINE

Virtual machine configuration

VERSION RELEASE DATE

February 27, 2013

VIRTUAL MACHINE NAME

NEW USER NAME

azureuser

NEW PASSWORD

CONFIRM

SIZE

Extra Large (8 cores, 14 GB Memory)

UPLOAD SSH KEY FOR AUTHENTICATION

Ubuntu 12.10

Ubuntu Server 12.10 (amd64 20130227) for Windows Azure. Ubuntu Server is the world's most popular Linux for cloud environments. Updates and patches for Ubuntu 12.10 will be available until April 2014. Ubuntu Server is the perfect platform for all workloads from web applications to NoSQL databases and Hadoop. More information can be found at: http://www.ubuntu.com/business/server

PUBLISHER	Canonical
OS FAMILY	Linux
LOCATION	East Asia;Southeast Asia;North Europe;West Europe;East US;West US

NODEJS HOME

COMPUTE

DATA SERVICES

APP SERVICES

END-TO-END APPS

Forums

On the page (jump to):

Table of contents

What are the Windows Azure Command-Line Tools for Mac and Linux

How to install the Windows Azure Command-Line Tools for Mac and Linux

How to use the Windows Azure Command-Line Tools for Mac and Linux

This guide describes how to use the Windows Azure Command-Line Tools for Mac and Linux to create and manage services in Windows Azure. The scenarios covered include **installing the tools**, **importing your publishing settings**, **creating and managing Windows Azure Web Sites**, and **creating and managing Windows Azure Virtual Machines**. For comprehensive reference documentation, see Windows Azure command-line tool for Mac and Linux Documentation.

http://www.windowsazure.com/en-us/develop/nodejs/how-to-guides/command-line-tools/

Table of contents

WindowsAzure / azure-sdk-for-python

Unwatch | Star 60 | For

Code | Network | Pull Requests 0 | Issues 13 | Wiki | Graphs

Windows Azure SDK for Python — Read more

Clone in Windows | ZIP | HTTP | SSH | Git Read-Only | git@github.com:WindowsAzure/azure-sdk-for-pytl | Read-Only

branch: master | Files | Commits | Branches 2

zure-sdk-for-python /

43 c

Merge pull request #77 from WindowsAzure/dev ...

jeffwilcoxmsft authored 15 days ago — latest commit 738c79

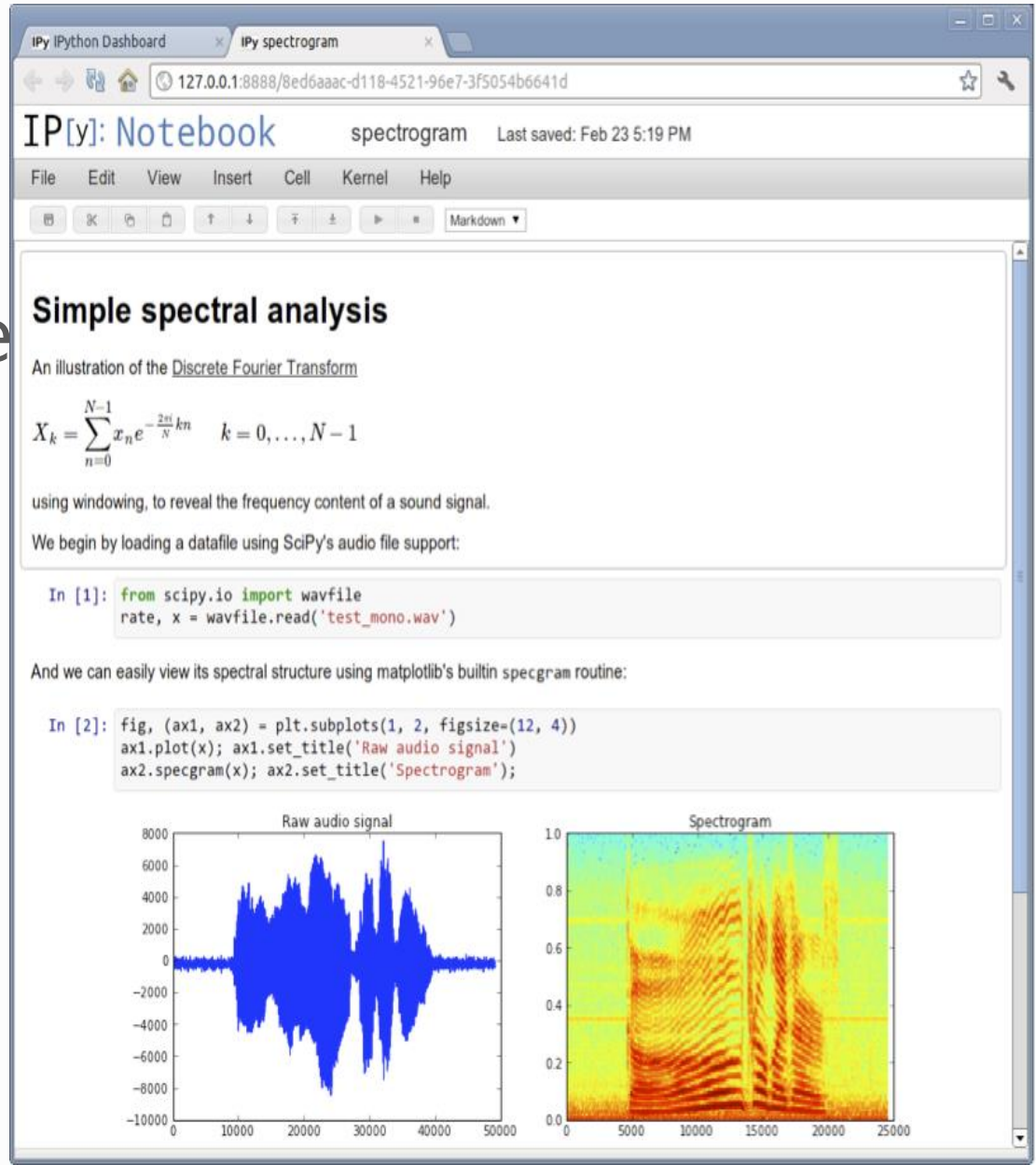
src	16 days ago	Updated version for next release [huguesv]
test	15 days ago	Split normal run and coverage run in separate batch files and added a... [huguesv]

Deploy & run Python on Virtual Machine

Demo

Python on Azure: IPython notebook

- Python IDE in browser
  - Any Browser
  - Any OS
- Backed by Python engine on Azure
  - Windows or Linux
- Key features
  - Intellisense, completion, …
  - Inline graphics
  - Markdown
- "Executable Document"
- IPython REPL also built-into PTVS



Windows Azure

Python Dashboard

https://dpedemo2.cloudapp.net

P[y]: Notebook

Logout

otebooks | Clusters

import a notebook, drag the file onto the listing below or **click here**.

ome / yew /

demo of K-Means clustering on the handwritten digits data

demo of structured Ward hierarchical clustering on Lena image

troML - Density Estimation

troML - Huber Loss Function

troML - SDSS Stripe 82 Moving Object Catalog

troML - SDSS imaging data and the SDSS Stripe 82 standard stars

Python Dashboard | IPy QuantLib-American Optio

https://dpedemo2.cloudapp.net/65b9e14e-78ab-4775-88c8-c89905b104a3

IP[y]: Notebook

Logo

QuantLib American Option Last saved: Apr 05 5:22 AM

File Edit View Insert Cell Kernel Help

Code Cell Toolbar: None

```
from QuantLib import *

# global data
todaysDate = Date(15,May,1998)
Settings.instance().evaluationDate = todaysDate
settlementDate = Date(17,May,1998)
riskFreeRate = FlatForward(settlementDate, 0.06, Actual365Fixed())

# option parameters
exercise = AmericanExercise(settlementDate, Date(17,May,1999))
payoff = PlainVanillaPayoff(Option.Put, 40.0)

# market data
underlying = SimpleQuote(36.0)
volatility = BlackConstantVol(todaysDate, TARGET(), 0.20, Actual365Fixed())
dividendYield = FlatForward(settlementDate, 0.00, Actual365Fixed())
```

IPy[y]: Notebook

QuantLib-American Option

Last saved: Apr 05 5:23 AM


```
process = BlackScholesMertonProcess(QuoteHandle(underlying),
                                    YieldTermStructureHandle(dividendYield),
                                    YieldTermStructureHandle(riskFreeRate),
                                    BlackVolTermStructureHandle(volatility))

option = VanillaOption(payoff, exercise)

refValue = 4.48667344
report('reference value',refValue)

# method: analytic

option.setPricingEngine(BaroneAdesiWhaleyEngine(process))
report('Barone-Adesi-Whaley',option.NPV())

option.setPricingEngine(BjerksundStenslandEngine(process))
report('Bjerksund-Stensland',option.NPV())

# method: finite differences
timeSteps = 801
gridPoints = 800

option.setPricingEngine(FDAmericanEngine(process,timeSteps,gridPoints))
report('finite differences',option.NPV())
```

Python Dashboard | IPy Visualizing the stock mark | IPy QuantLib-American Optic

https://dpedemo2.cloudapp.net/3515c455-c06f-4b68-b01e-78a0da7ab8d0

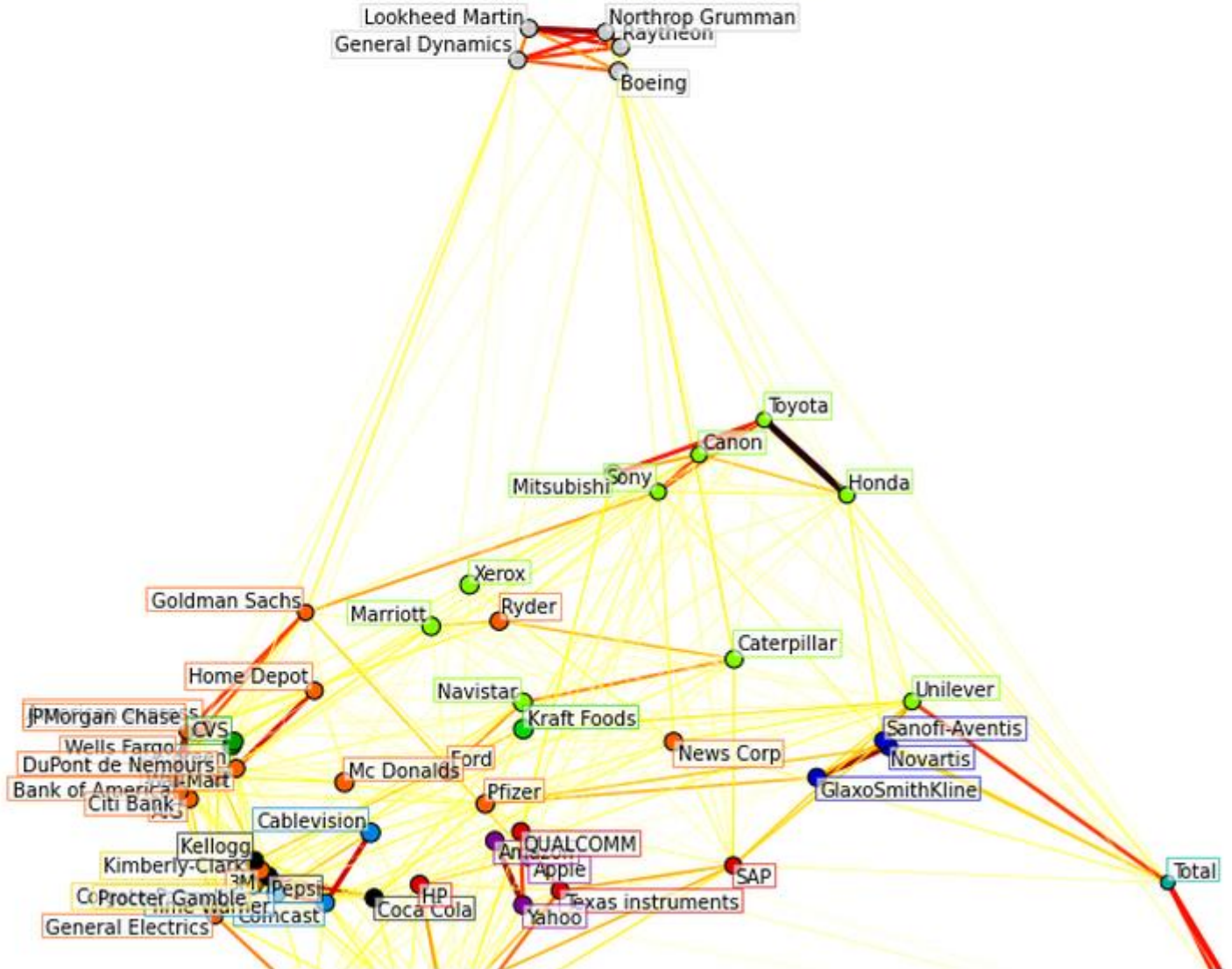
[y]: Notebook

Visualizing the stock market structure

Last saved: Apr 05 5:25 AM

Edit View Insert Cell Kernel Help

Code | Cell Toolbar: None



```
from azure import *
from azure.servicemanagement import *

sms = ServiceManagementService(subscription_id, certificate_path)

name = 'myvm'
location = 'West US'

# You can either set the location or an affinity_group
sms.create_hosted_service(service_name=name, label=name, location=location)

# Name of an os image as returned by list_os_images
image_name = 'OpenLogic__OpenLogic-CentOS-62-20120531-en-us-30GB.vhd'

# Destination storage account container/blob where the VM disk
# will be created
media_link = 'url_to_target_storage_blob_for_vm_hd'

# Linux VM configuration, you can use WindowsConfigurationSet
# for a Windows VM instead
linux_config = LinuxConfigurationSet('myhostname', 'myuser', 'mypassword', True)

os_hd = OSVirtualHardDisk(image_name, media_link)

sms.create_virtual_machine_deployment(service_name=name,
    deployment_name=name, deployment_slot='production', label=name,
    role_name=name, system_config=linux_config, os_virtual_hard_disk=os_hd, role_size='Small')
```

Gallery Images Available

Microsoft

Windows Server 2008 R2

SQL Server Eval 2012

Windows Server 2012

Biztalk Server 2013 Beta

Microsoft® SQL Server® 2012

Microsoft® BizTalk® Server

Open Source

OpenSUSE 12.2

CentOS 6.3

Ubuntu 12.04/12.10

SUSE Linux Enterprise Server 11 SP2

Windows Azure™

Virtual machine portability

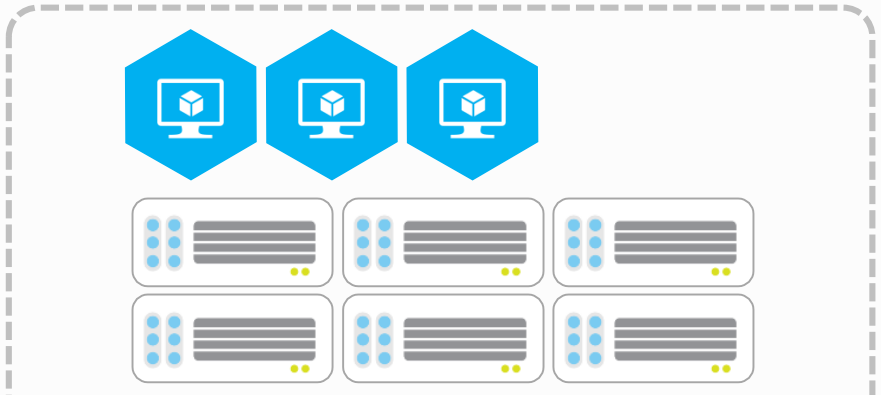
Windows Azure

Your Data Center

Windows Azure

Windows Azure

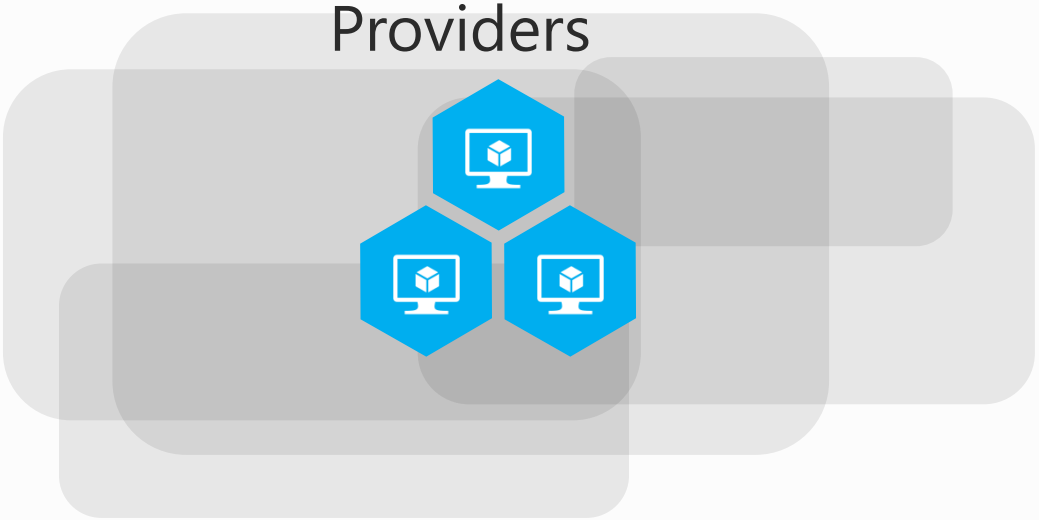
Your Data Center



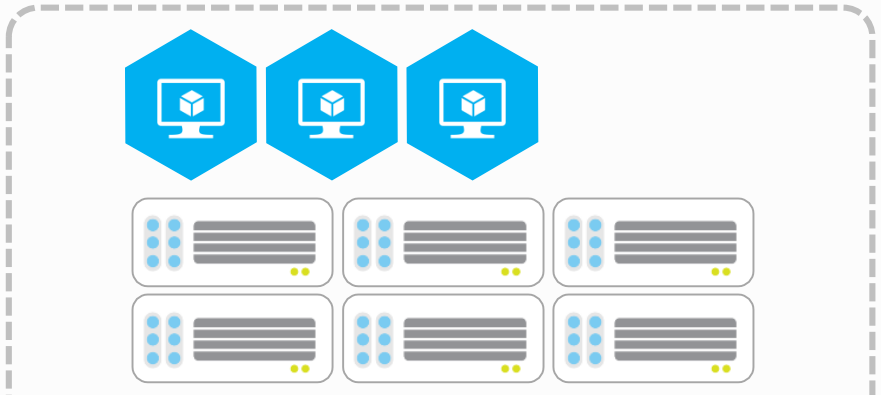
Windows Azure

Windows Azure

Other Service Providers



Your Data Center

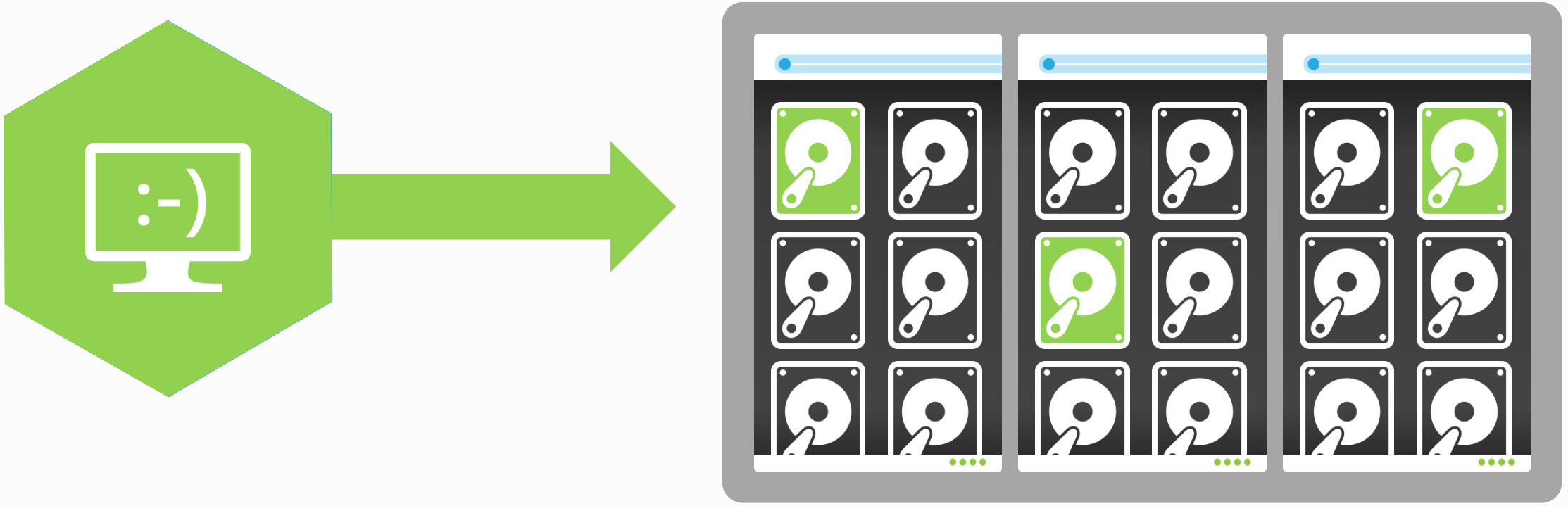


Windows Azure

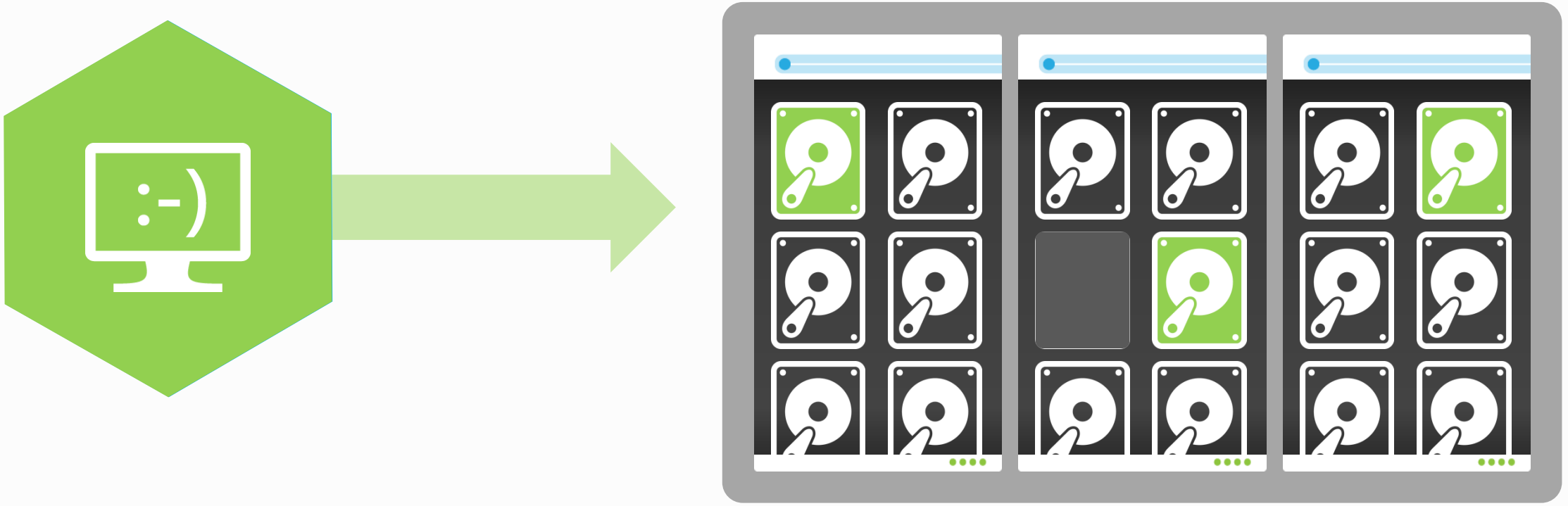
VM with persistent drive

Windows Azure Storage

VM with persistent drive



VM with persistent drive

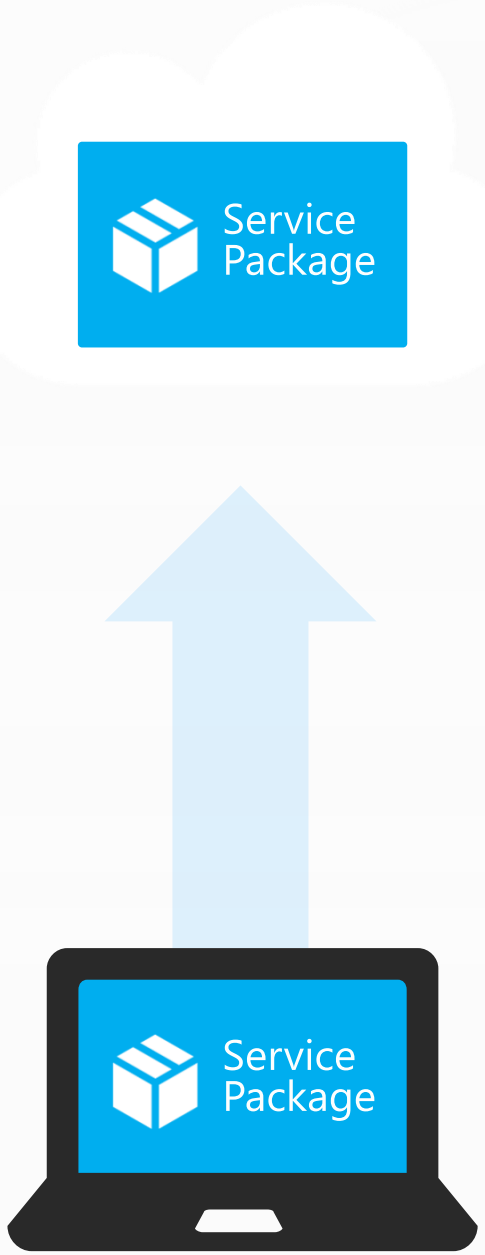


Cloud services

Build infinitely scalable apps and services

Support rich multi-tier architectures

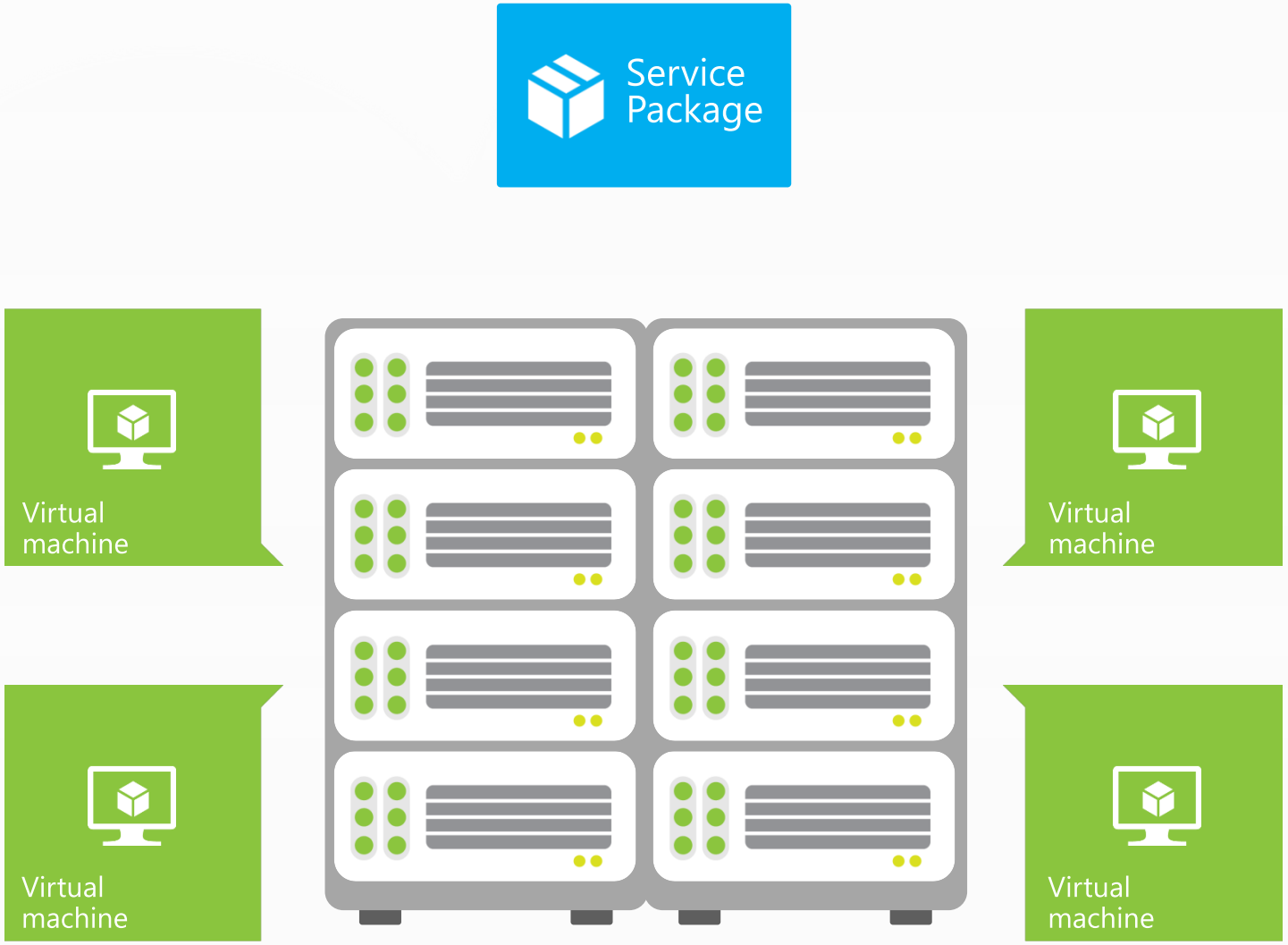
Automated application management


Windows Azure
Service Package
Service Package

Provision Role Instances

Deploy App Code

Configure Network

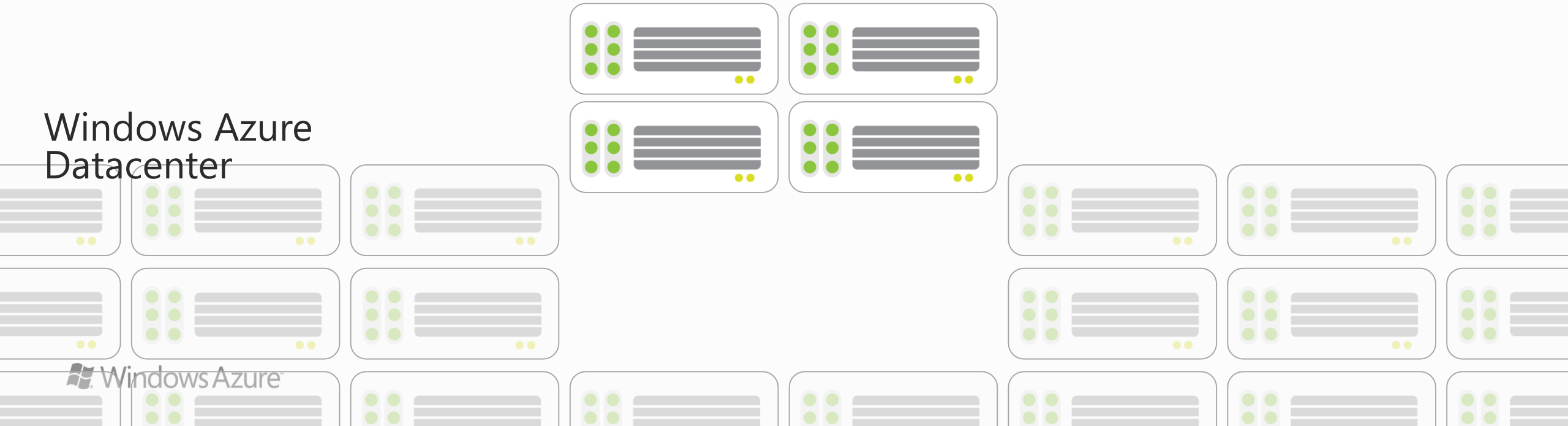


Windows Azure

Provision Role Instances

Deploy App Code

Configure Network

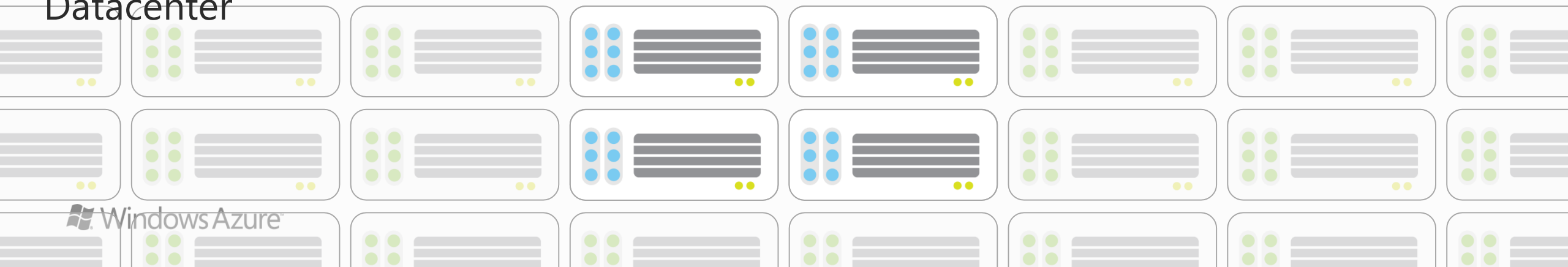


Provision Role Instances

Deploy App Code

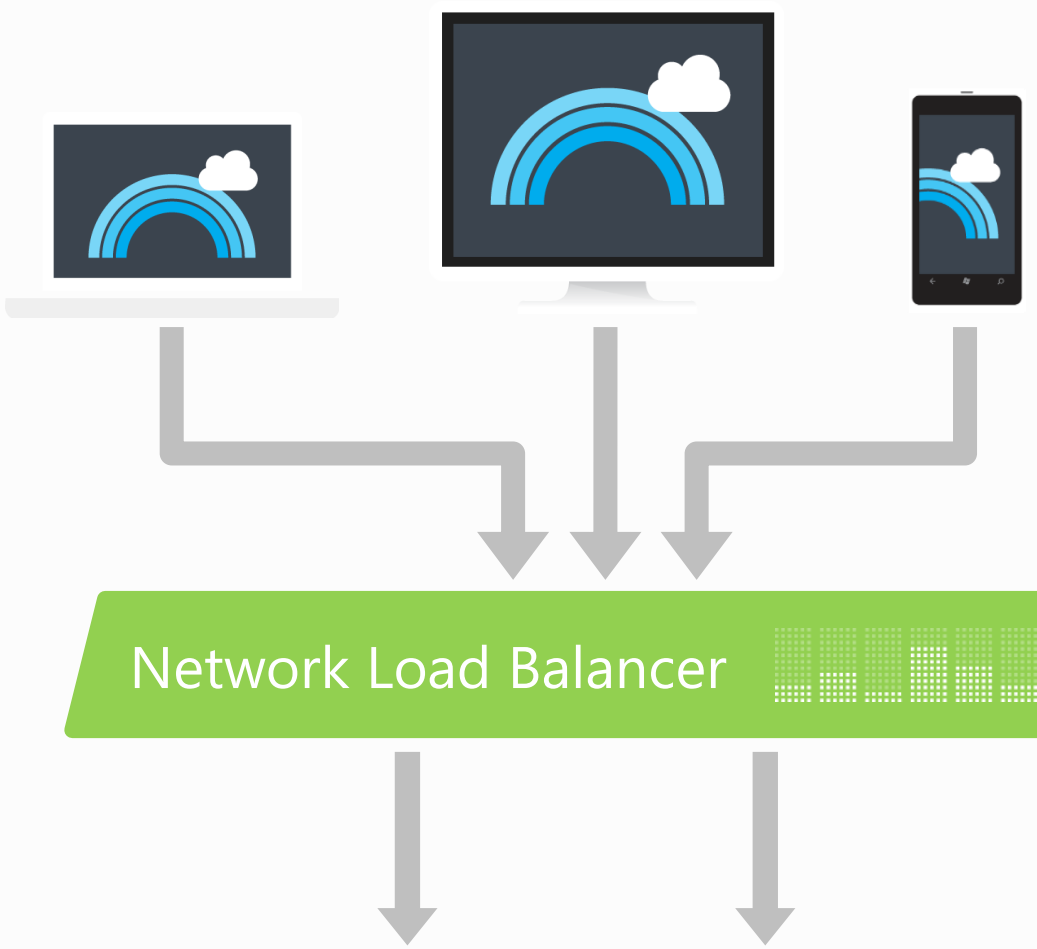
Configure Network

Windows Azure
Datacenter



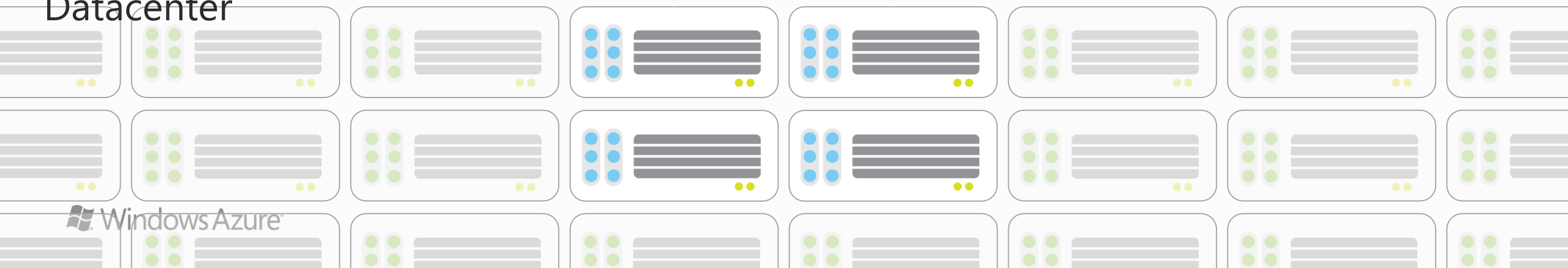
Windows Azure

Provision Role Instances
Deploy App Code
Configure Network



← Network load-balancer configured for traffic

Windows Azure Datacenter



Windows Azure

Cloud services: Development

Demo

http://weatherservice.cloudapp.net/Gallery/List

List - Weather Forecasting ...

Storage - Windows Azure ... | Suggested Sites | Concur Hotel Reservation... | Windows Azure HDInsight

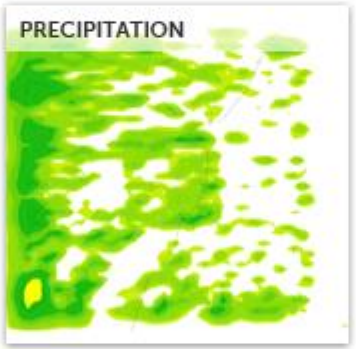
NORTHWIND WEATHER RESEARCH

HOME / GALLERY / NEW FORECAST / PROCESSING QUEUE

LIST | MAP | TIMELINE

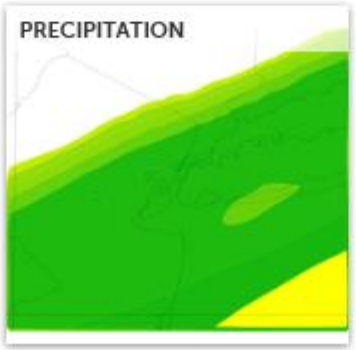
Sydney Harbour

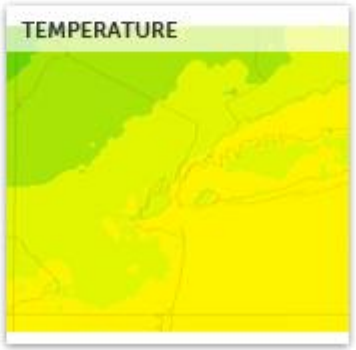
4/5/2013 12:00:00 AM

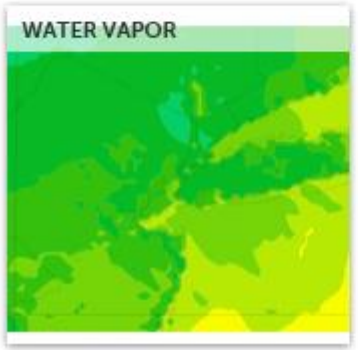


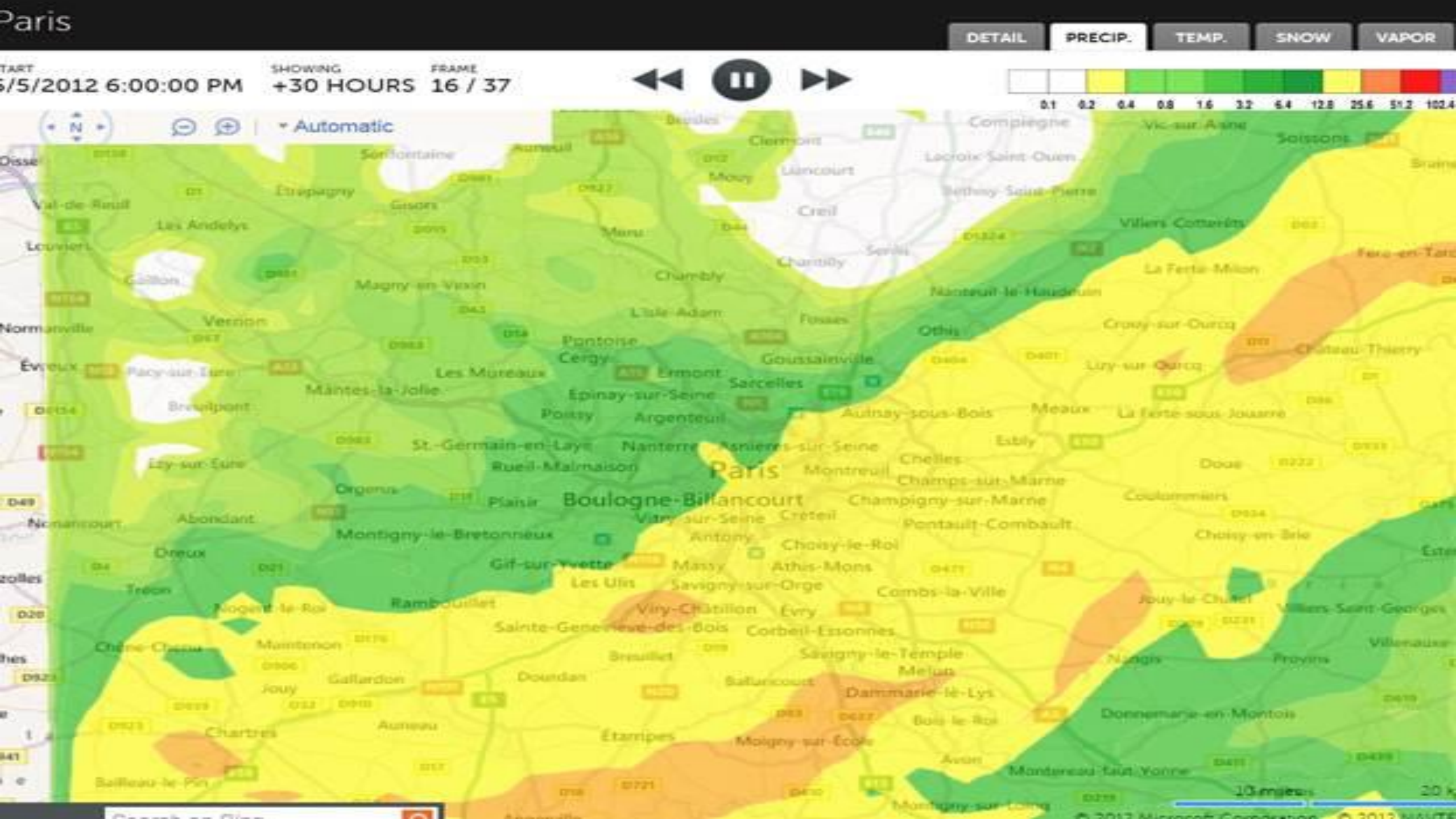
ny

4/4/2013 12:00:00 PM








Paris
DETAIL
PRECIP.
TEMP.
SNOW
VAPOR
START
5/5/2012 6:00:00 PM
SHOWING
+30 HOURS
FRAME
16 / 37
0.1
0.2
0.4
0.8
1.6
3.2
6.4
12.8
25.6
51.2
102.4
Automatic
Paris
Boulogne-Billancourt
Versailles
Saint-Germain-en-Laye
Nanterre
Argenteuil
Pontoise
Cergy
Meaux
Melun
Évreux
Dreux
Chartres
Étampes
Fontainebleau
Montereau-fault-Yonne
10 miles

weatherservice

DASHBOARD MONITOR CONFIGURE **SCALE** INSTANCES LINKED RESOURCES CERTIFICATES

PRODUCTION STAGING

roles

HEADNODE
MEDIUM VM

1 instance(s)

FRONTEND2
MEDIUM VM

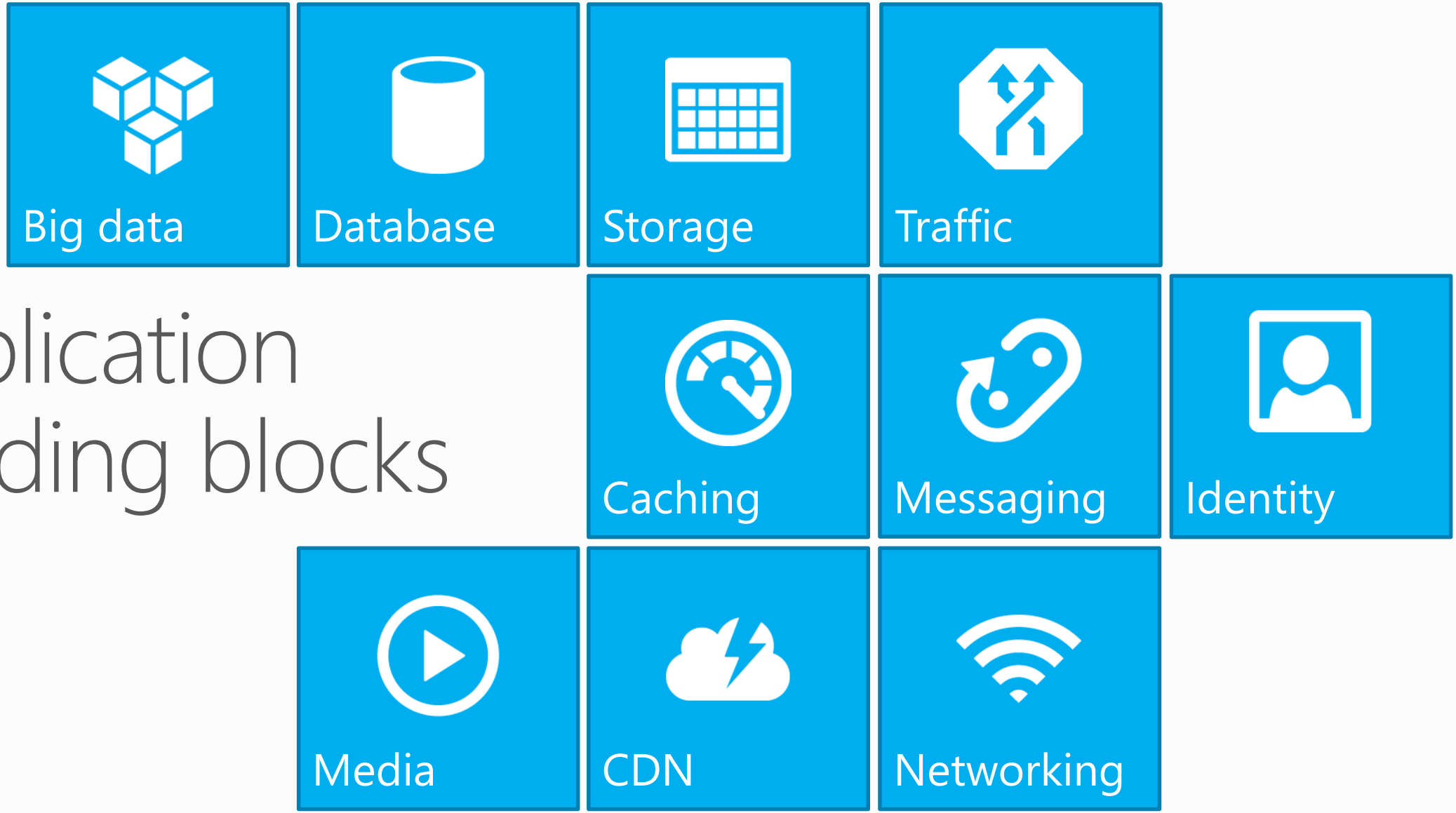
1 instance(s)

COMPUTENODE
EXTRA LARGE VM

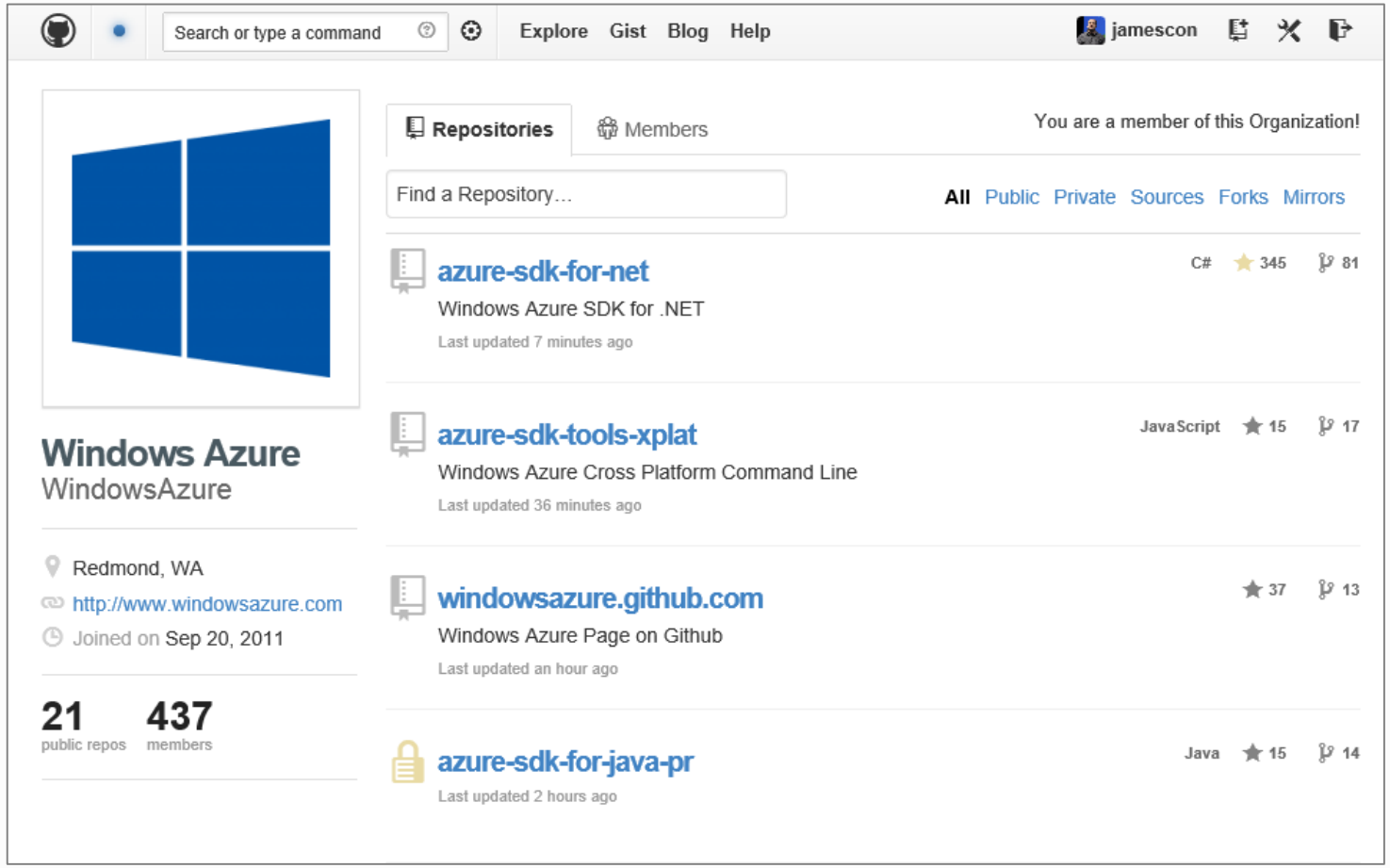
12 instance(s)

SAVE DISCARD

Application building blocks



Windows Azure



Open source

→ http://github.com/windowsazure

Windows Azure

Introduction to Windows Azure Storage

Windows Azure Storage

Cloud Storage - Anywhere and anytime access

- Blobs, Disks, Tables and Queues

Highly Durable, Available and Massively Scalable

- Easily build "internet scale" applications
- 8.5 trillion stored objects
- 900K request/sec on average (2.3+ trillion per month)

Pay for what you use

Exposed via easy and open REST APIs

Client libraries in .NET, Java, Node.js, Python, PHP, Ruby

Abstractions – Blobs and Disks

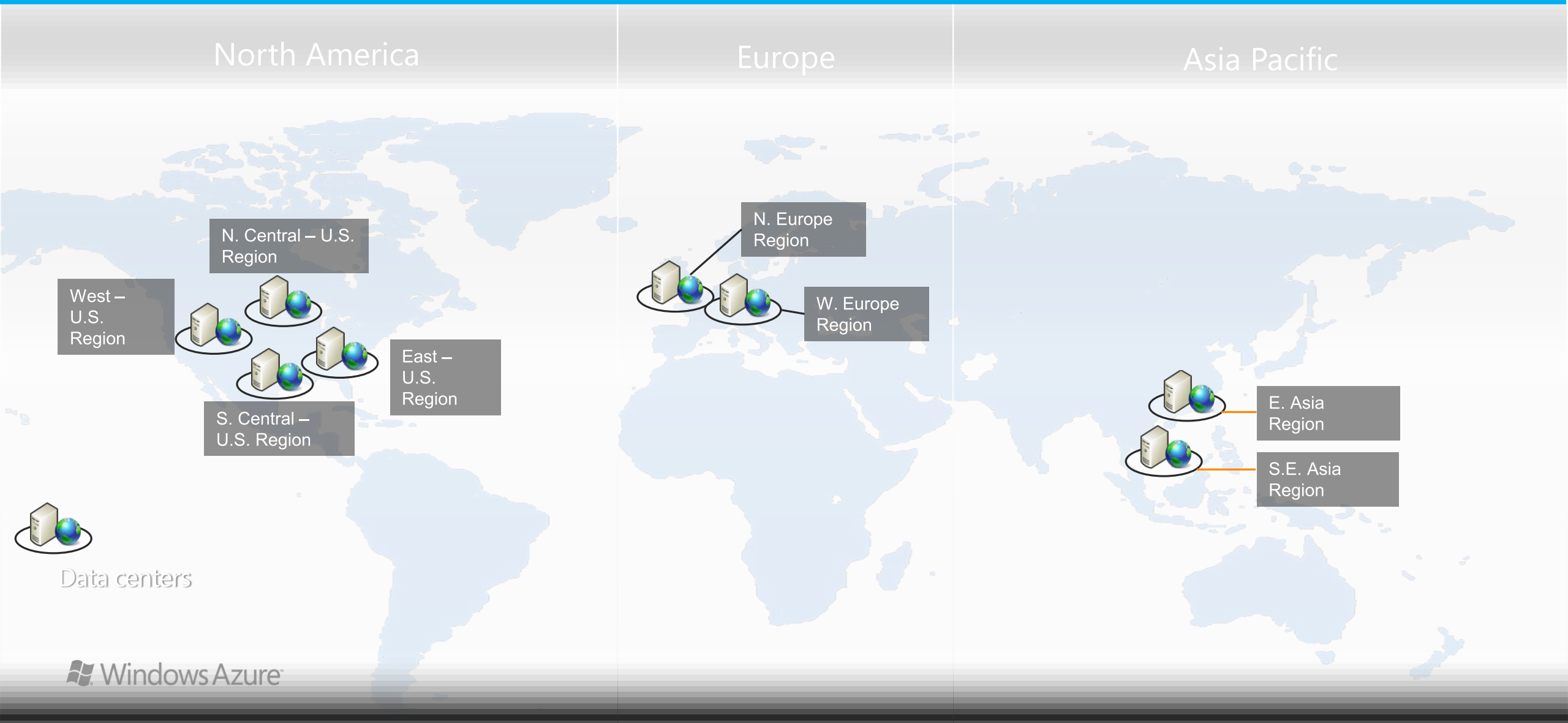
- Blobs – Simple interface to store and retrieve files in cloud
  - Data sharing – share documents, pictures, video, music, etc.
  - Big Data – store raw data/logs and compute/map reduce over data
  - Backups – data and device backups
- Disks – Network mounted durable disks for VMs in Azure
  - Mounted disks are VHDs stored in Azure Blobs
  - Move on-premise applications to cloud

```
# Create container
from azure.storage import BlobService
blob_service = BlobService(account_name, account_key)
blob_service.create_container('taskcontainer')

# Upload
from azure.storage import BlobService
blob_service = BlobService(account_name, account_key)
blob_service.put_blob('taskcontainer', 'task1',
file('task1-upload.txt').read(), 'BlockBlob')

#Download
from azure.storage import BlobService
blob_service = BlobService(account_name, account_key)
blob = blob_service.get_blob('taskcontainer', 'task1')
```

Abstractions – Tables and Queues

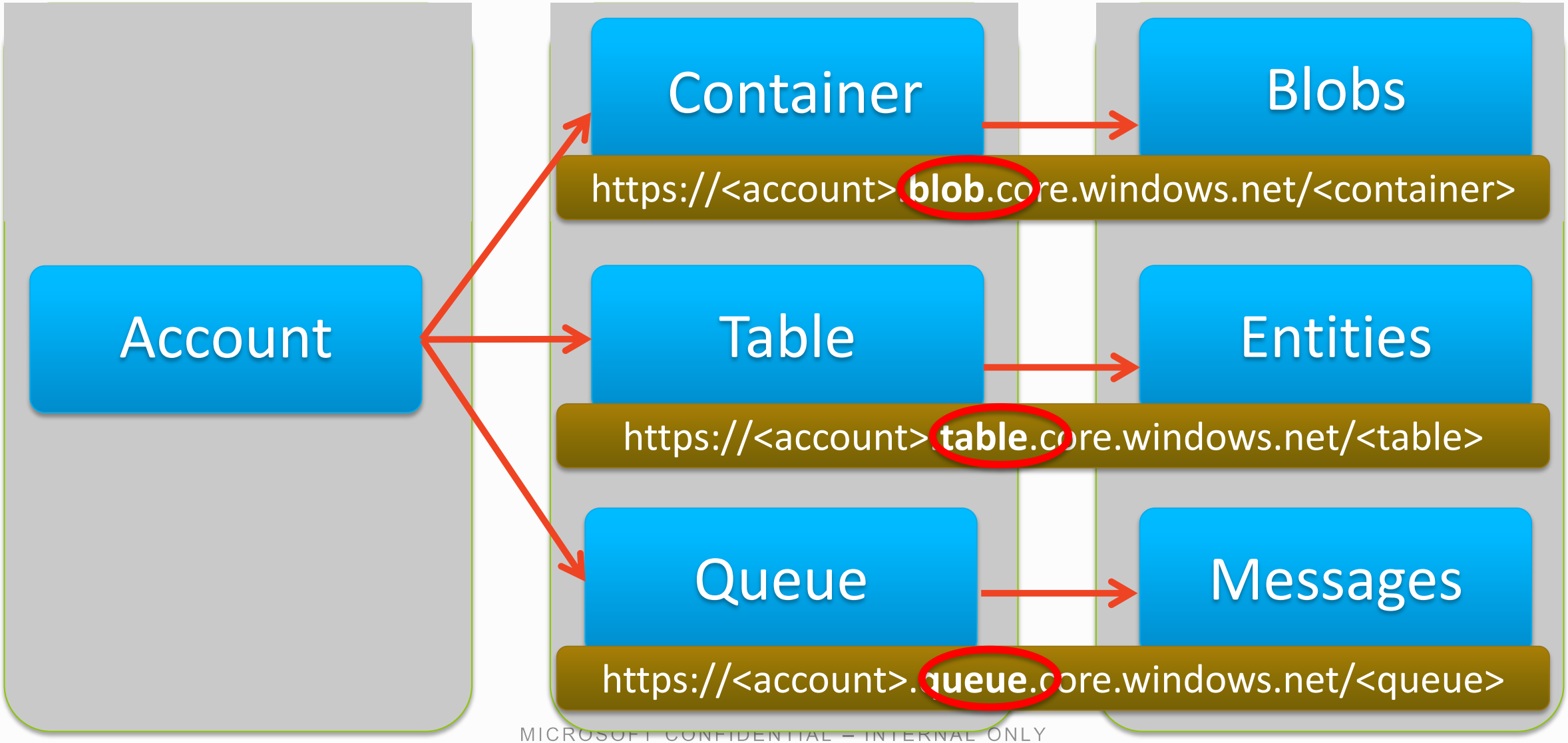
- Tables – Massively scalable and extremely easy to use NoSQL system that auto scales
  - Key-value lookups at scale
  - Store user information, device information, any type of metadata for your service
- Queues – Reliable messaging system
  - Decouple components/roles
    - Web role to worker role communication
    - Allows roles to scale independently
  - Implement scheduling of asynchronous tasks
  - Building process/work flows

Windows Azure Storage



Windows Azure

Windows Azure Data Storage Concepts



How is Azure Storage used by Microsoft?

Xbox: Uses Blobs, Tables & Queues for Cloud Game Saves, Halo 4, XBox Music, XBox Live, etc.

Skype: Uses Blobs, Tables and Queues for Skype video messages and to keep metadata to allow Skype clients to connect with each other

Bing: Uses Blobs, Tables and Queues to provide a near real-time ingestion engine that consumes Twitter and Facebook feeds, indexes them, which is then folded into Bing search

SkyDrive: Uses Blobs to store pictures, documents, videos, files, etc.

Internals

Design Goals

Highly Available with Strong Consistency

- Provide access to data in face of failures/partitioning

Durability

- Replicate data several times within and across regions

Scalability

- Need to scale to zettabytes
- Provide a global namespace to access data around the world
- Automatically scale out and load balance data to meet peak traffic demands

- Additional details can be found in the SOSP paper:
  - "Windows Azure Storage: A Highly Available Cloud Storage Service with Strong Consistency", *ACM Symposium on Operating System Principals (SOSP)*, Oct. 2011

Windows Azure Storage Stamps

Access blob storage via the URL: http://<account>.blob.core.windows.net/

Data access

Storage Location Service

LB

Front-Ends

Partition Layer

DFS Layer

Intra-stamp replication

Storage Stamp

Inter-stamp (Geo) replication

LB

Front-Ends

Partition Layer

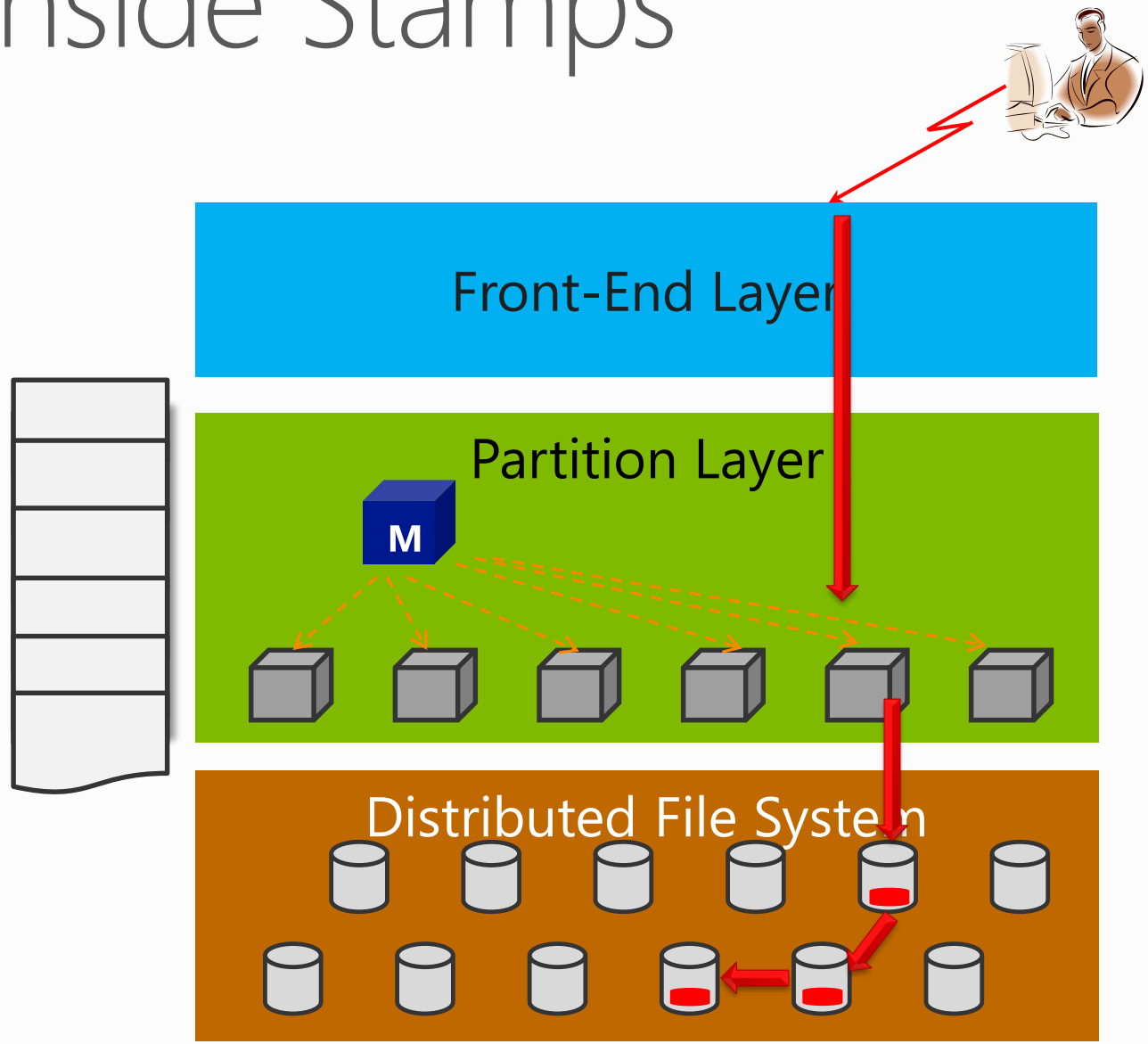
DFS Layer

Intra-stamp replication

Storage Stamp

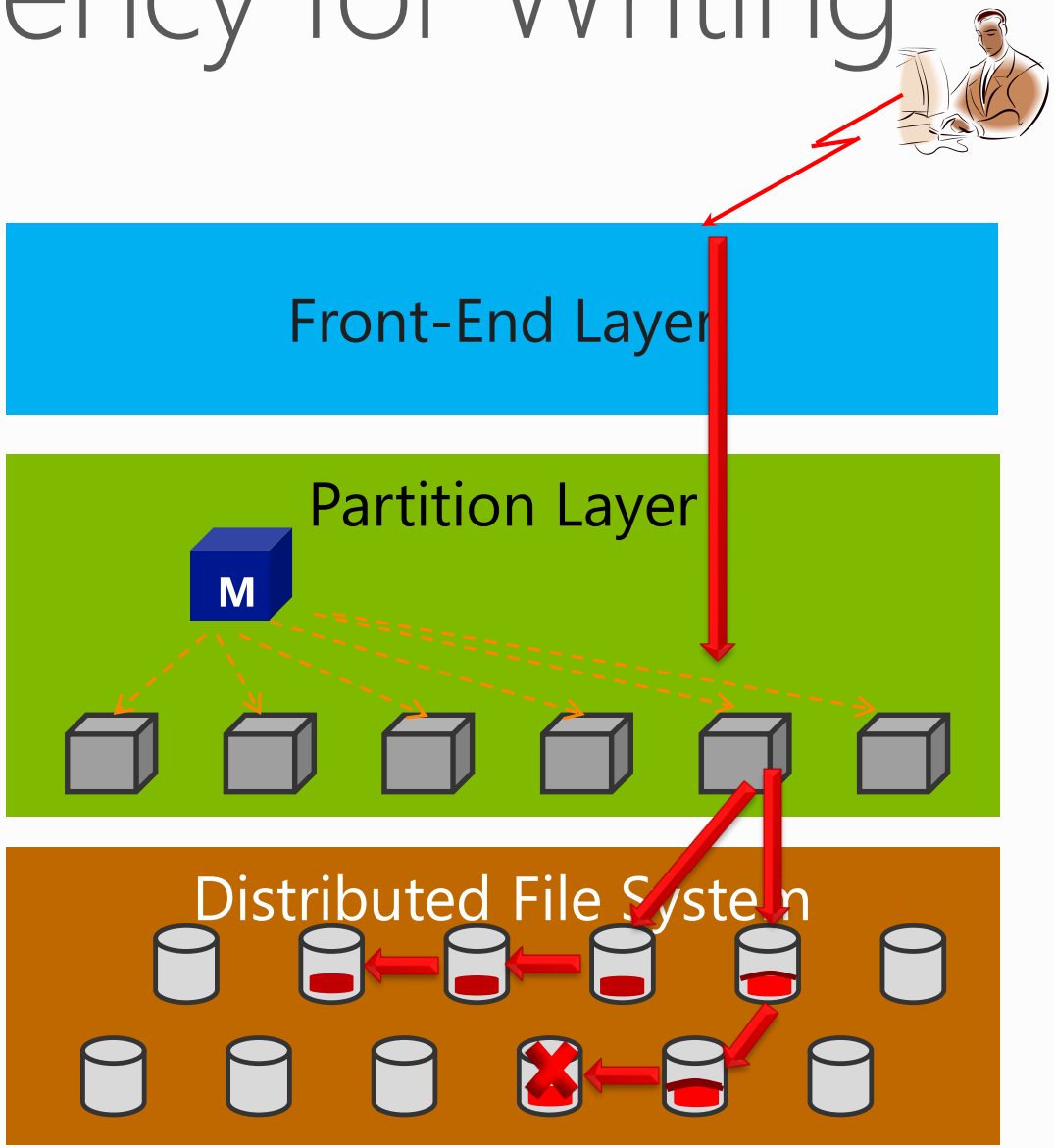
Architecture Layers inside Stamps

- Front-end Layer
  - REST front-end (blob, table, queue)
  - Authentication/authorization
  - Metrics/logging
- Partition Layer
  - Understands our data abstractions, and provides optimistic concurrency
  - Massively scalable index
  - Log Structured Merge Tree
    - Each log (stream) is a linked list of extents
- Distributed File System Layer
  - Data persistence and replication (JBOD)
  - Data is stored into a file called extent, which is replicated 3 times across different nodes (UDs/FDs)
  - Append-only file system



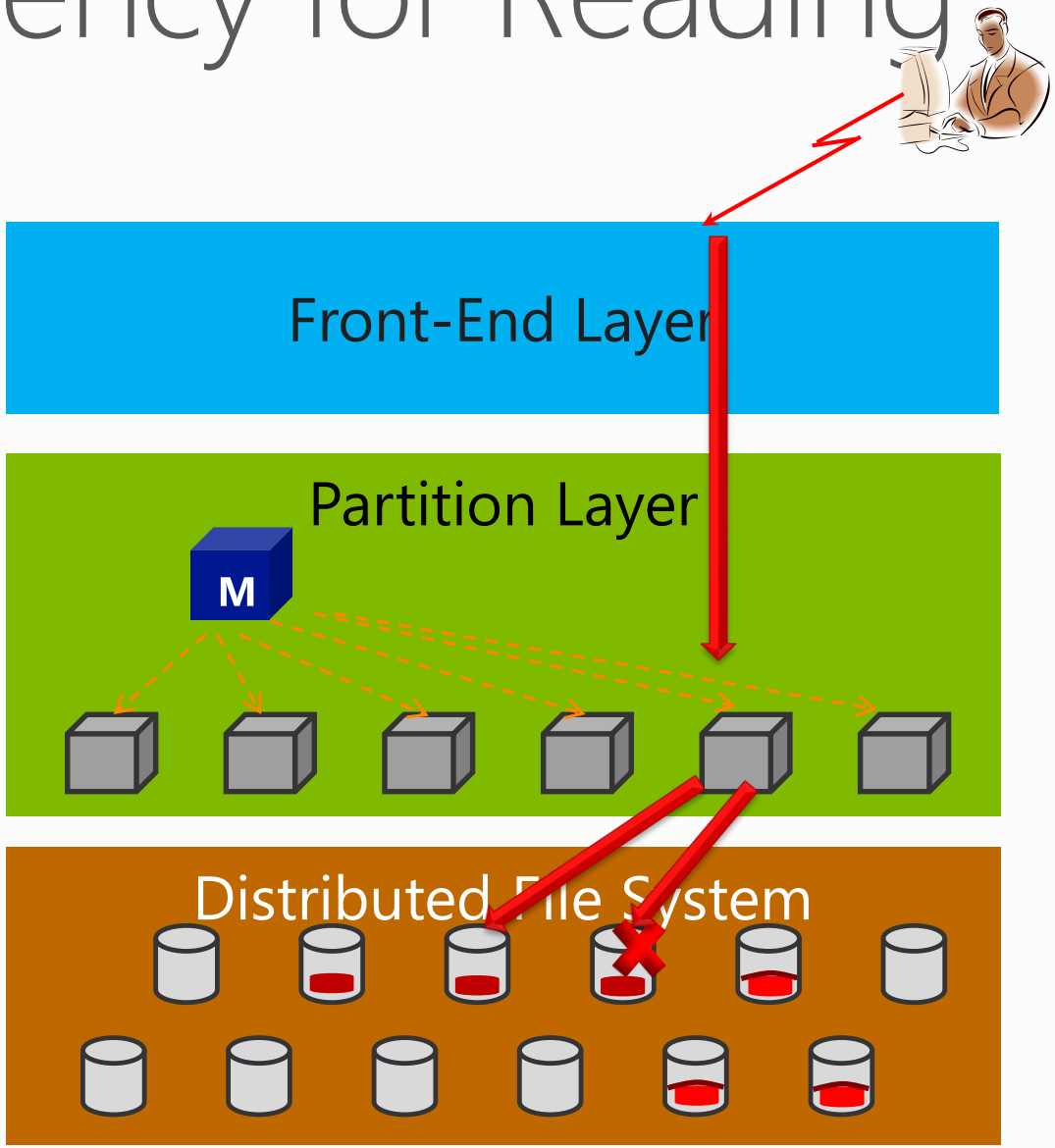
Availability with Consistency for Writing

- All writes are appends to the end of a log, which is an append to the last extent in the log
- Write Consistency across all replicas for an extent:
  - Appends are ordered the same across all 3 replicas for an extent (file)
  - Only return success if all 3 replica appends are committed to storage
  - When extent gets to a certain size or on write failure/LB, seal the extent's replica set and never append anymore data to it
- Write Availability: To handle failures during write
  - Seal extent's replica set
  - Append immediately to a new extent (replica set) on 3 other available nodes
  - Add this new extent to the end of the partition's log (stream)



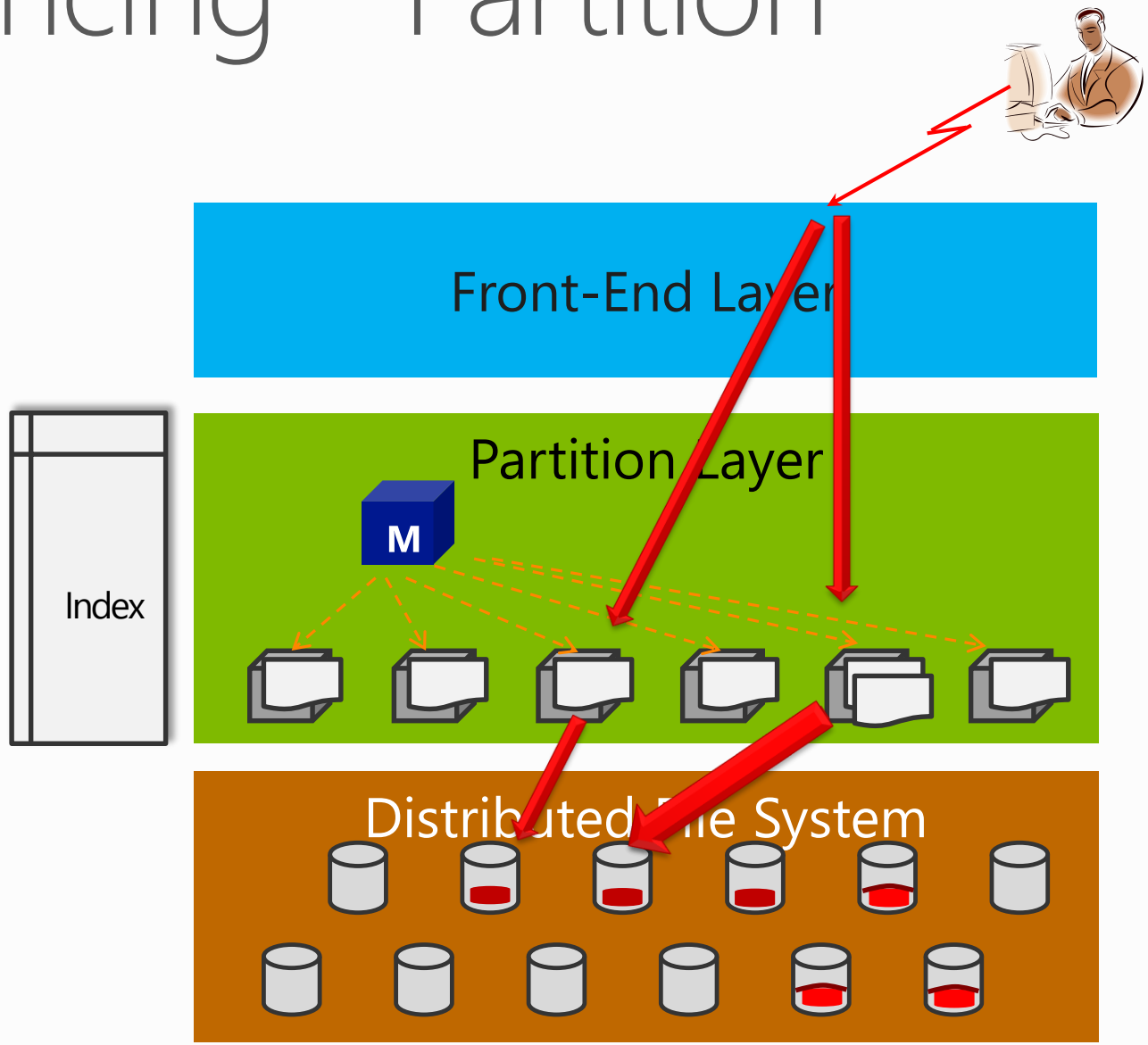
Availability with Consistency for Reading

- Read Consistency: Can read from any replica, since data in each replica for an extent is bit-wise identical
- Read Availability: Send out parallel read requests if first read is taking higher than 95% latency



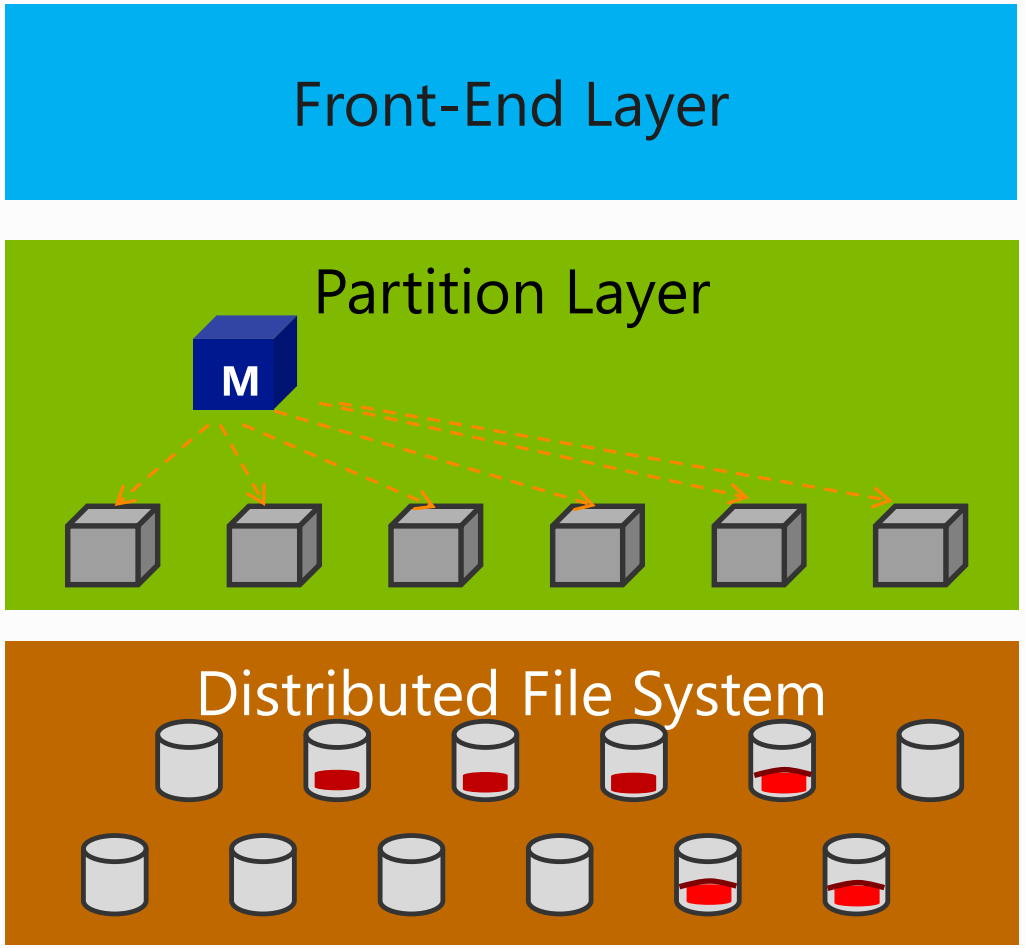
Dynamic Load Balancing – Partition Layer

- Spreads index/transaction processing across partition servers
  - Master monitors traffic load/resource utilization on partition servers
  - Dynamically load balance partitions across servers to achieve better performance/availability
- Does not move data around, only reassigns what part of the index a partition server is responsible for



Dynamic Load Balancing – DFS Layer

- DFS Read load balancing across replicas
  - Monitor latency/load on each node/replica; dynamically select what replica to read from and start additional reads in parallel based on 95% latency
- DFS write load balancing
  - Monitor latency/load on each node; seal the replica set with an overloaded node, and switch to a new extent on another set of nodes to append to
- DFS capacity load balancing
  - Lazily move replicas around to ensure the disks and nodes have equal amount of data on them
  - Important for avoiding hot nodes/disks



Architecture Summary

- **Durability:** All data stored with at least 3 replicas
- **Consistency:** All committed data across all 3 replicas are identical
- **Availability:** Can read from any 3 replicas; If any issues writing seal extent and continue appending to new extent
- **Performance/Scale:** Retry based on 95% latencies; Auto scale out and load balance based on load/capacity

- Additional details can be found in the SOSP paper:
  - "Windows Azure Storage: A Highly Available Cloud Storage Service with Strong Consistency", *ACM Symposium on Operating System Principals (SOSP)*, Oct. 2011

What's Coming

What's Coming by end of 2013

- Geo-Replication
  - Queue Geo-Replication
  - Secondary Read-Only Access
- Windows Azure Import/Export
- Real-Time Metrics for Blobs, Tables and Queues
- CORS for Azure Blobs, Tables and Queues
- JSON for Azure Tables
- New .NET 2.1 Library

Two Types of Durability Offered

Local Redundant Storage Accounts

- Maintain 3 copies of data within a given region
- ~ 2/3 price of Geo Redundant Storage

Geo Redundant Storage Accounts

- Maintain 6 copies of data spread over 2 regions at least 400 miles apart from each other (3 copies are kept at each region)

Geo Redundant Storage

Data geo-replicated across regions 400+ miles apart

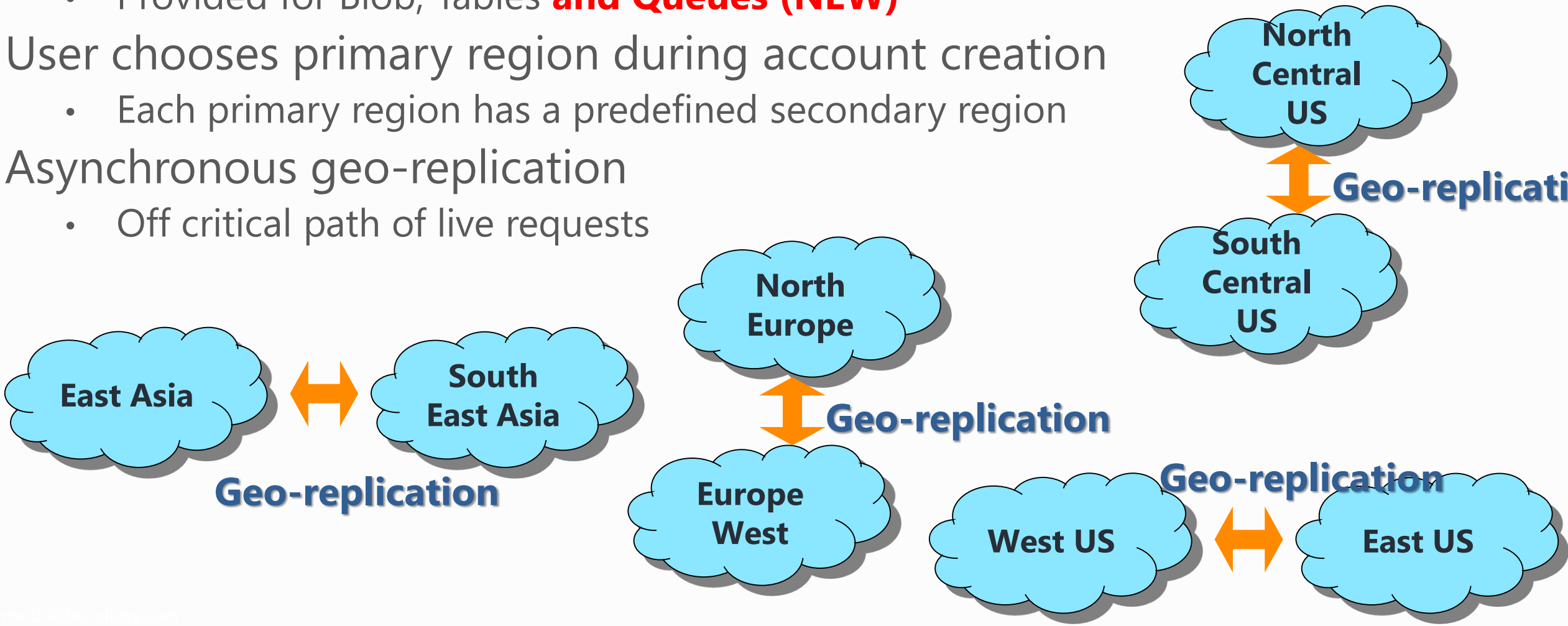
- Provide data durability in face of potential major regional disasters
- Provided for Blob, Tables **and Queues (NEW)**

User chooses primary region during account creation

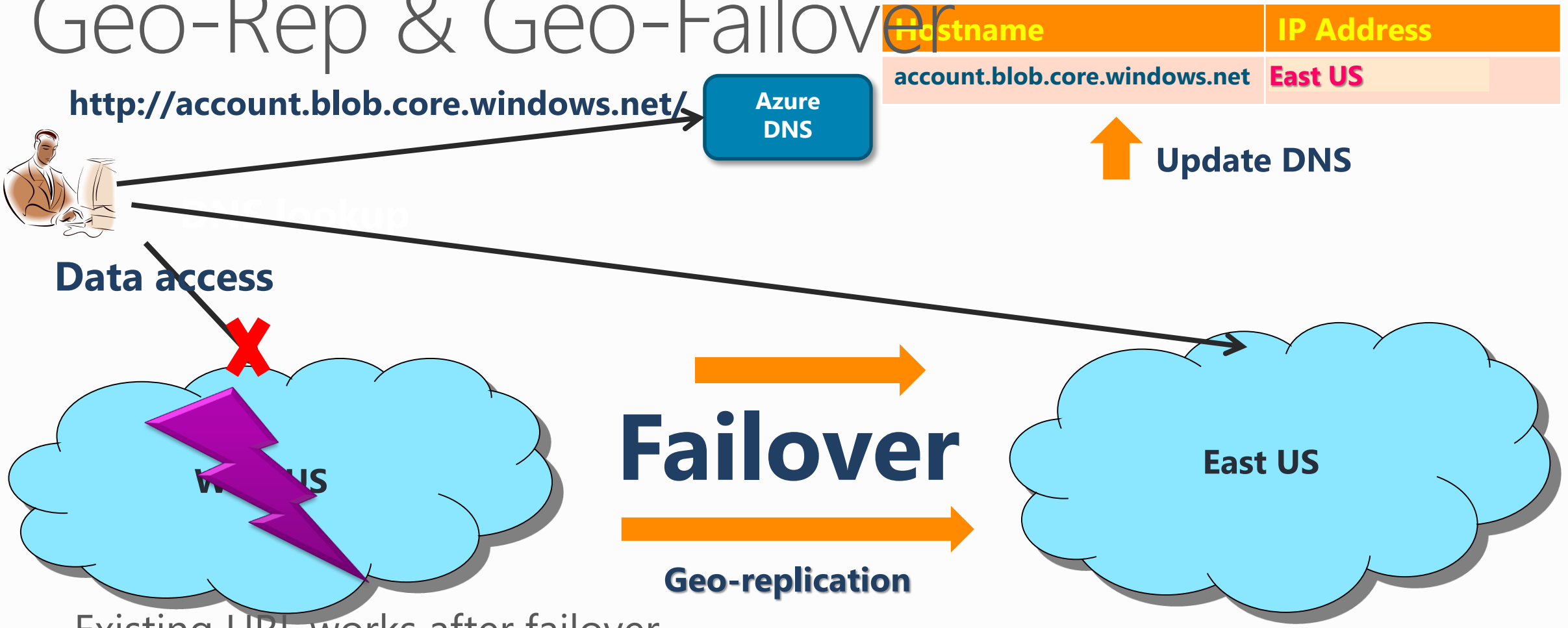
- Each primary region has a predefined secondary region

Asynchronous geo-replication

- Off critical path of live requests



Geo-Rep & Geo-Failover



- Existing URL works after failover
- Failover Trigger – failover would only be used by MS if primary could not be recovered
- Asynchronous Geo-replication – may lose recent updates during failover
- Typically geo-replicate data within minutes, though no SLA for how long it will take

Geo Redundant Storage Roadmap

- Customer Controlled Failover (Future)
  - Provide APIs to allow clients to switch the primary and secondary regions for a storage account
- Queue Geo-Replication (Done)
- Secondary Read Only Access (by end of CY13)

Secondary Read-Only Access – Scenarios

Read-only access to data even if primary is unavailable

- Access to an eventually consistent copy of the data in the other region

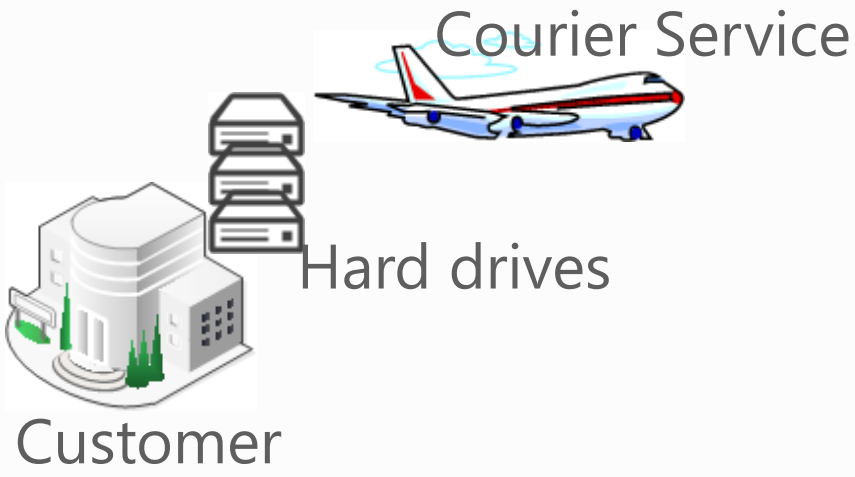
Provides another read source for geographically distributed applications/customers

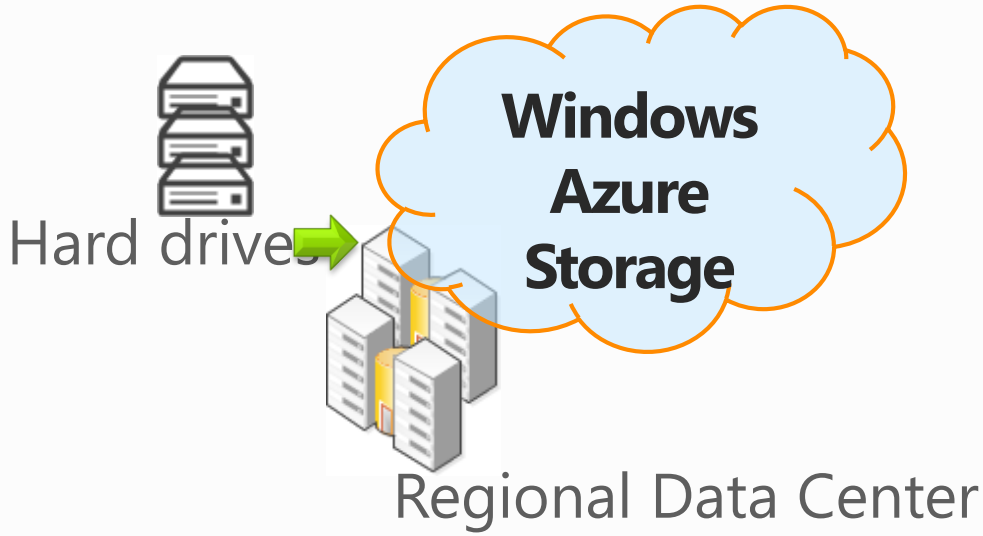
- Allows lower latency access to data in secondary region
- Have compute at both primary and secondary region and use the storage stored in that region

- For these, the application semantics need to allow for eventually consistent reads

Windows Azure Import/Export

- Move TBs of data into and out of Windows Azure Blobs by shipping disks





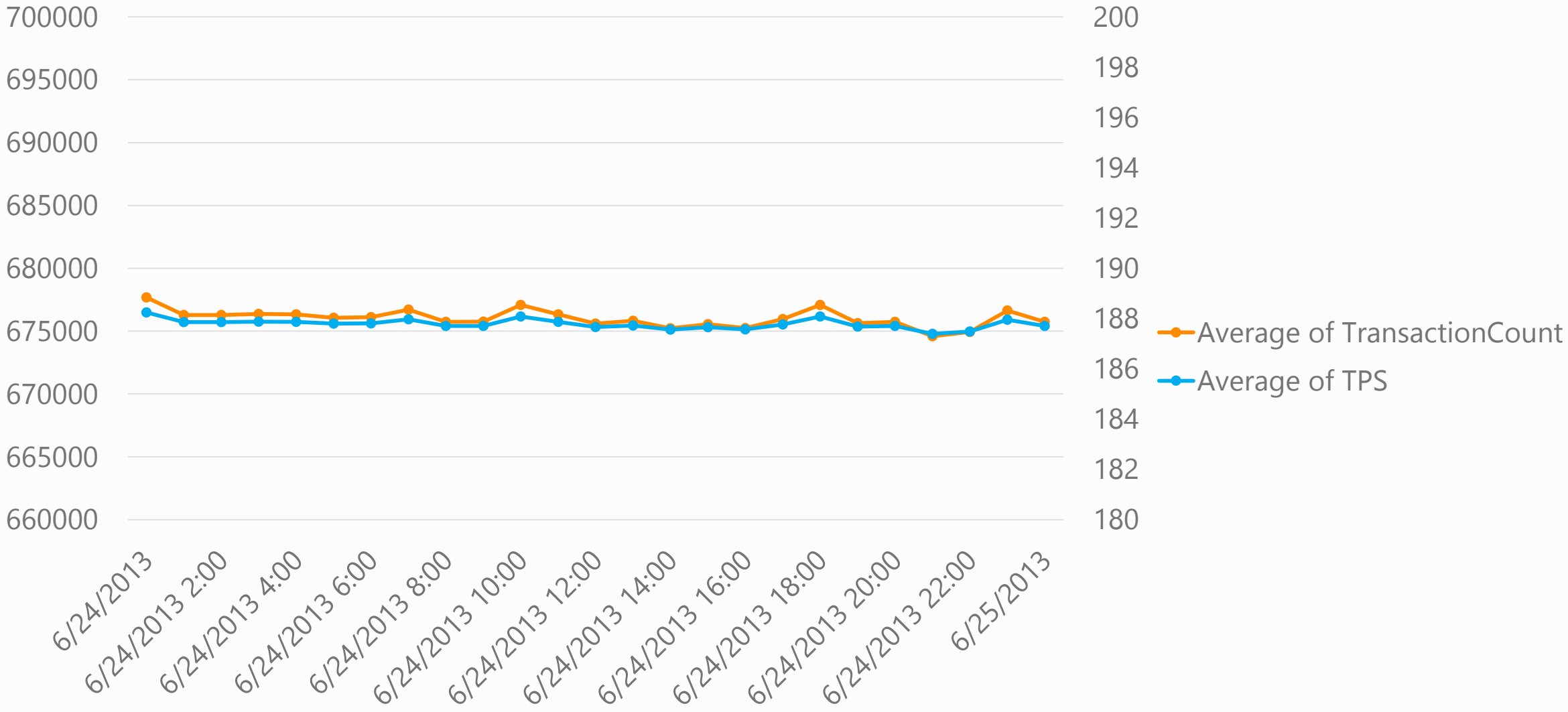
Import/Export Features

- Accessible via REST with Portal integration
- Each Job imports/exports data for a single storage account
  - Each Job can be up to 10 disks
- Support 3.5” SATA HDDs
- All Disks must be encrypted with BitLocker

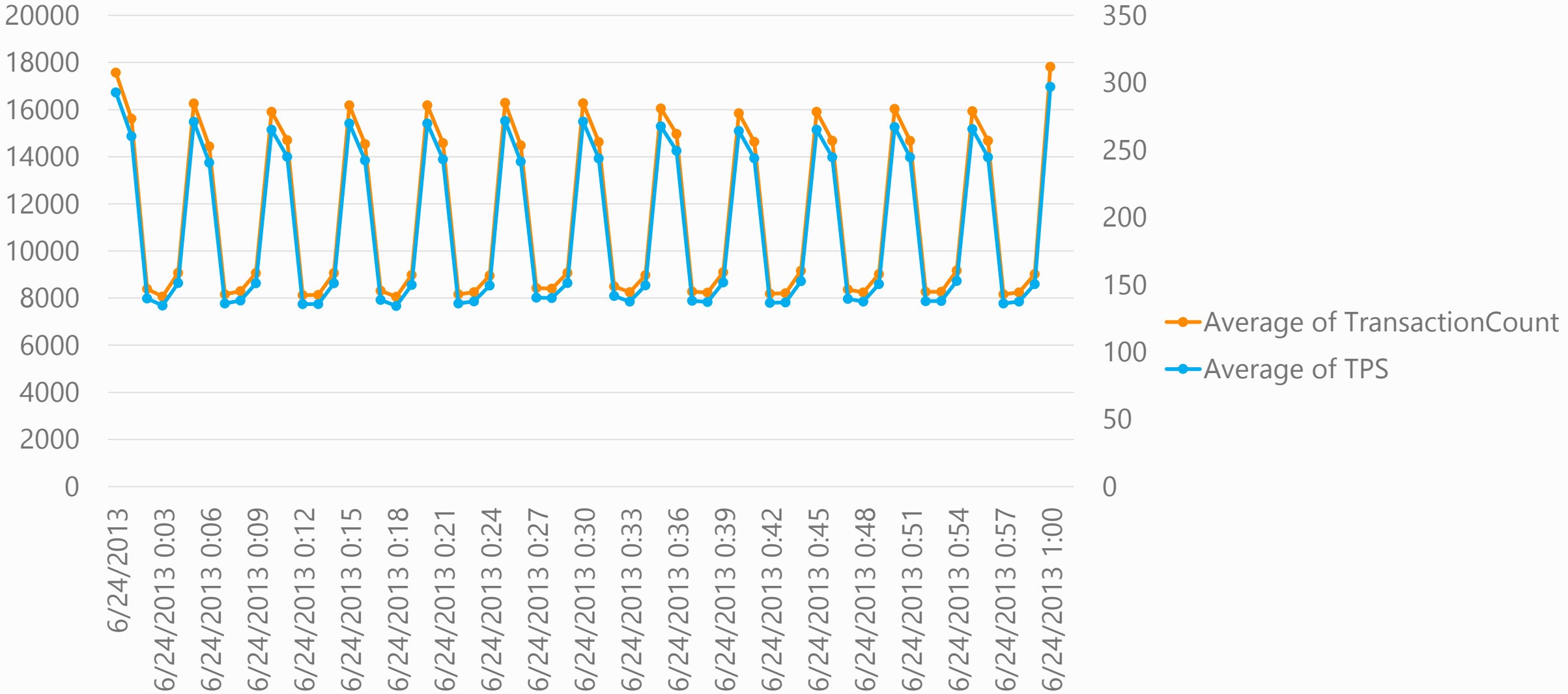
Realtime Metrics

- What?
  - Provide per minute aggregates at API and service level with up to 5 minute delay
  - Available for Blobs, Tables and Queues
  - Enable via SetServiceProperties
- This data is stored in tables under your storage account
  - $MetricsRealtimeTransactionsBlob, $MetricsRealtimeTransactionsTable and $MetricsRealtimeTransactionsQueue
    - Utilize this to monitor
  - Availability, Transaction rate, Ingress/Egress, Failure rate per category etc.
- Why?
  - Empower you to monitor/investigate problems in application in real-time
  - Hourly granularity smoothens spikes that may exist in traffic patterns

Hourly metrics - TPS



Realtime Metrics - TPS



JSON (JavaScript Object Notation)

- What?
  - A popular concise format for REST protocols
  - OData supports two formats
    - ATOMPub: We currently support this but is too verbose
    - JSON: OData has released multiple flavors of JSON
- Why?
  - Improves COGS for applications
    - Lower bandwidth consumption (approx. 70% savings), lower cpu utilization and hence better responsiveness
  - Many applications use JSON to represent object model
    - Efficient object data model to wire protocol

2.1 .NET Library

- New Features
  - Async Task methods with support for cancellation
  - Byte Array, Text, File upload / download APIs for blobs
  - IQueryable provider for Tables
  - Compiled Expressions for Table Entities
- Performance Improvements
  - Buffer Pooling
  - Multi-Buffer Memory Stream for consistent performance when buffering unknown length data
  - .NET MD5 now default (~20% faster than invoking native one)
- Available Soon @ http://www.nuget.org/packages/WindowsAzure.Storage

Best Practices – Account, Blobs, Tables and Queues

General Best Practices

- Locate Storage accounts close to compute/users
- Understand Account Scalability targets
  - Use multiple storage accounts to get more
  - Distribute your storage accounts across regions
- Cache critical data sets
  - As a Backup data set to fall back on
  - To get more request/sec than the account/partition targets
- Distribute load over many partitions and avoid spikes

General Best Practices (cont.)

- Use HTTPS
- Optimize what you send & receive
  - Blobs: Range reads, Metadata, Head Requests
  - Tables: Upsert, Merge, Projection, Point Queries
  - Queues: Update Message, Batch size
- Control Parallelism at the application layer
  - Unbounded Parallelism can lead to slow latencies and throttling

General Best Practices (cont.)

- Enable Logging & Metrics on each storage service
  - Can be done via REST, Client API, or Portal
  - Enables clients to self diagnose issues, including performance related ones
  - Data can be automatically GC'd according to a user specified retention interval
    - For example, have longer retention for hourly metrics and shorter retention for realtime metrics

Blob Best Practice

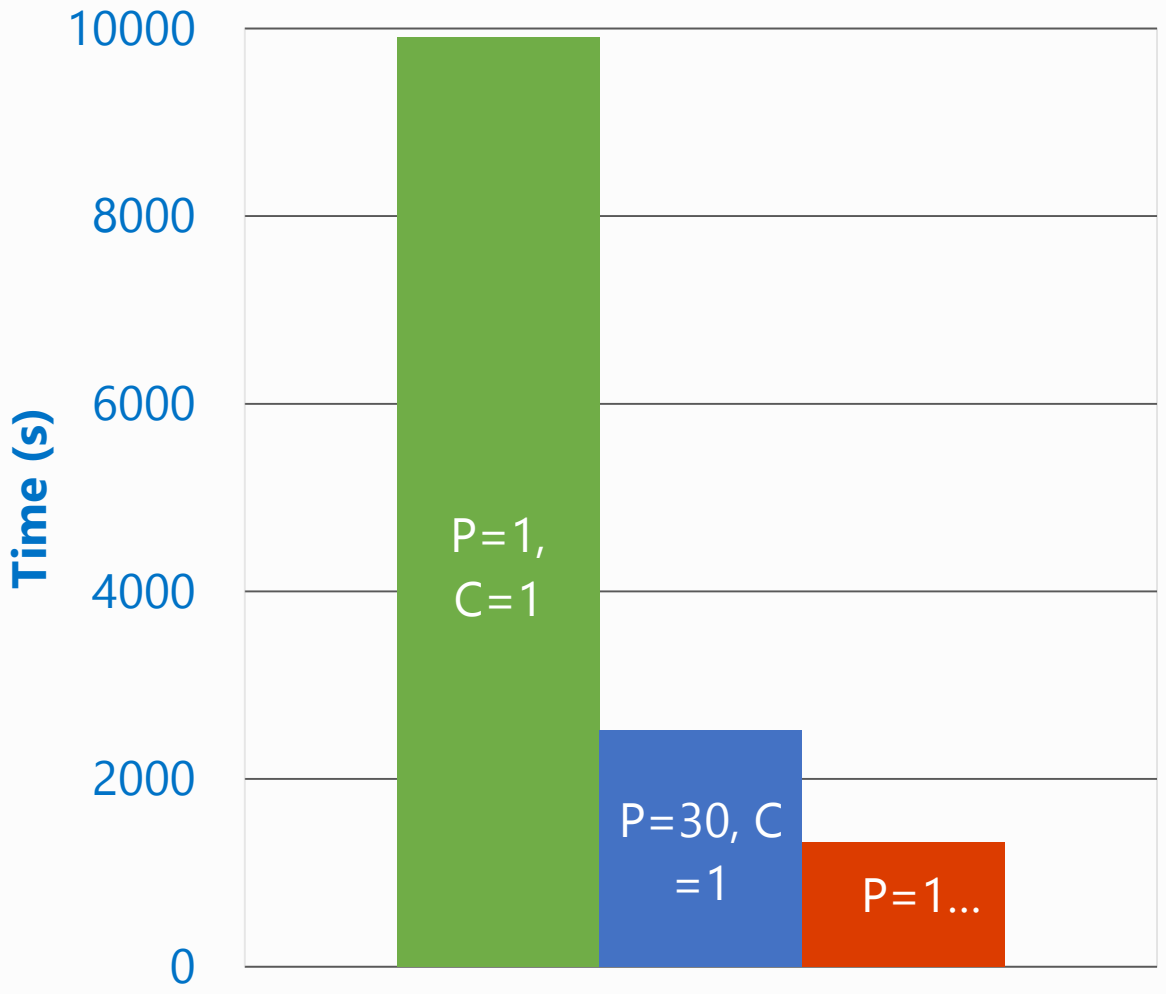
- Try to match your read size with your write size
  - Avoid reading small ranges on blobs with large blocks
  - CloudBlockBlob.StreamMinimumReadSizeInBytes/ StreamWriteSizeInBytes
- How do I upload a folder the fastest?
  - Upload multiple blobs simultaneously
- How do I upload a blob the fastest?
  - Use parallel block upload
- Concurrency (C)- Multiple workers upload different blobs
- Parallelism (P) – Multiple workers upload different blocks for same blob

Concurrency Vs. Blob Parallelism

XL VM Uploading 512, 256MB Blobs (Total upload size = 128GB)

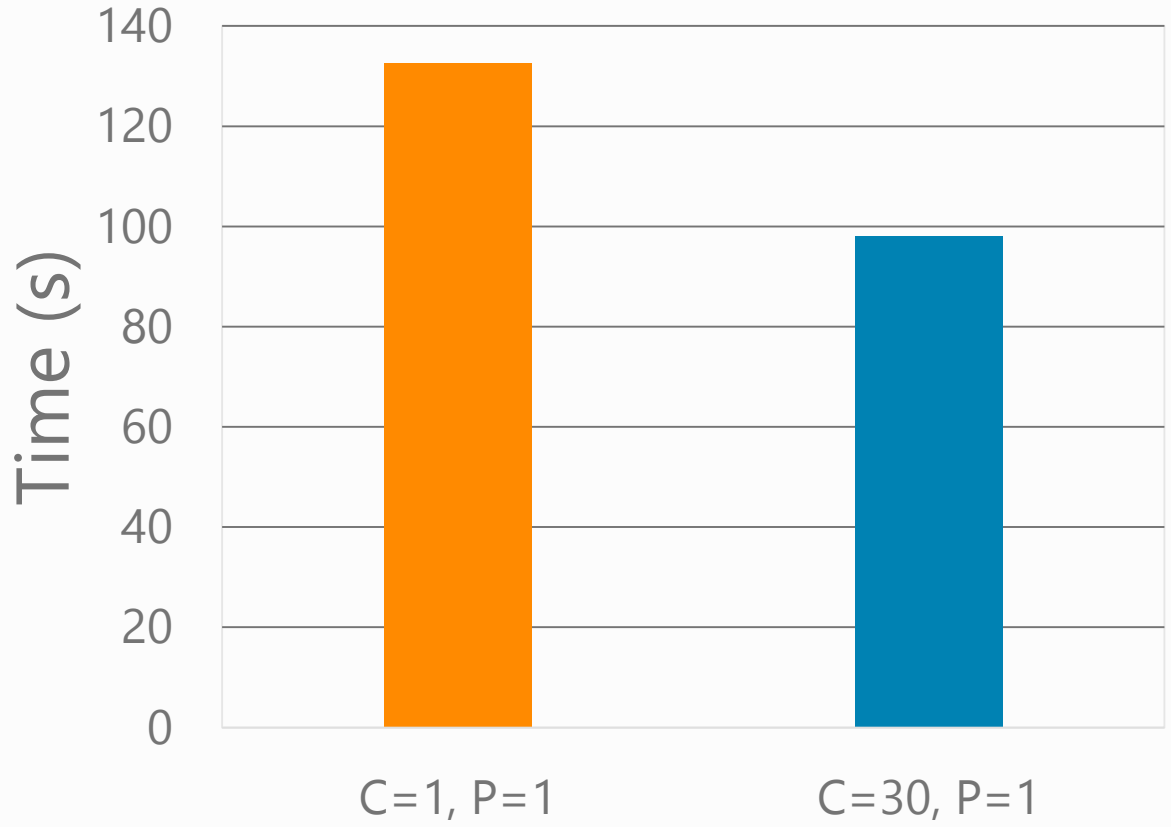
- C=1, P=1 => Averaged ~ 13. 2 MB/s
- C=1, P=30 => Averaged ~ 50.72 MB/s
- C=30, P=1 => Averaged ~ 96.64 MB/s

- Single TCP connection is bound by TCP rate control & RTT
- P=30 vs. C=30: Test completed almost twice as fast!
- Single Blob is bound by the limits of a single partition
- Accessing multiple blobs concurrently scales



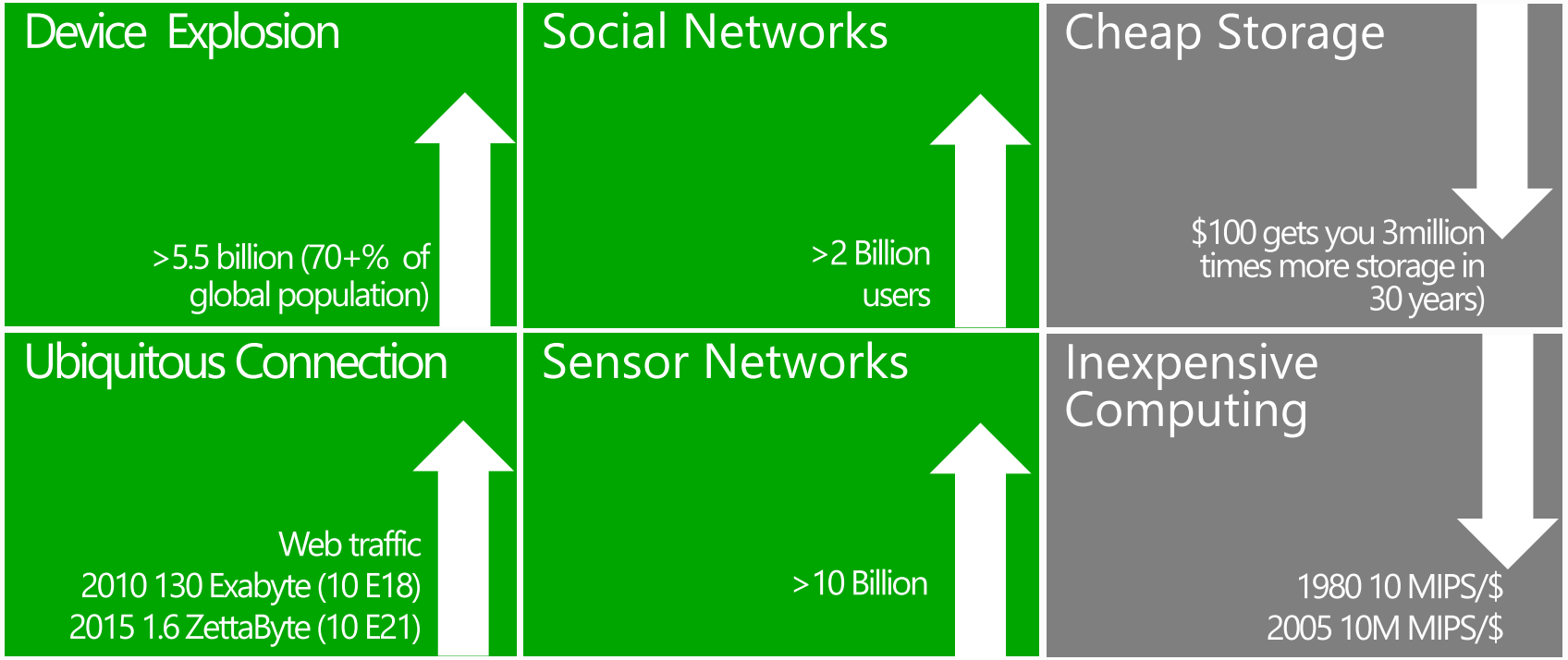
Blob Download

- XL VM Downloading 50, 256MB Blobs (Total download size = 12.5GB)
  - C=1, P=1 => Averaged ~ 96 MB/s
  - C=30, P=1 => Averaged ~ 130 MB/s

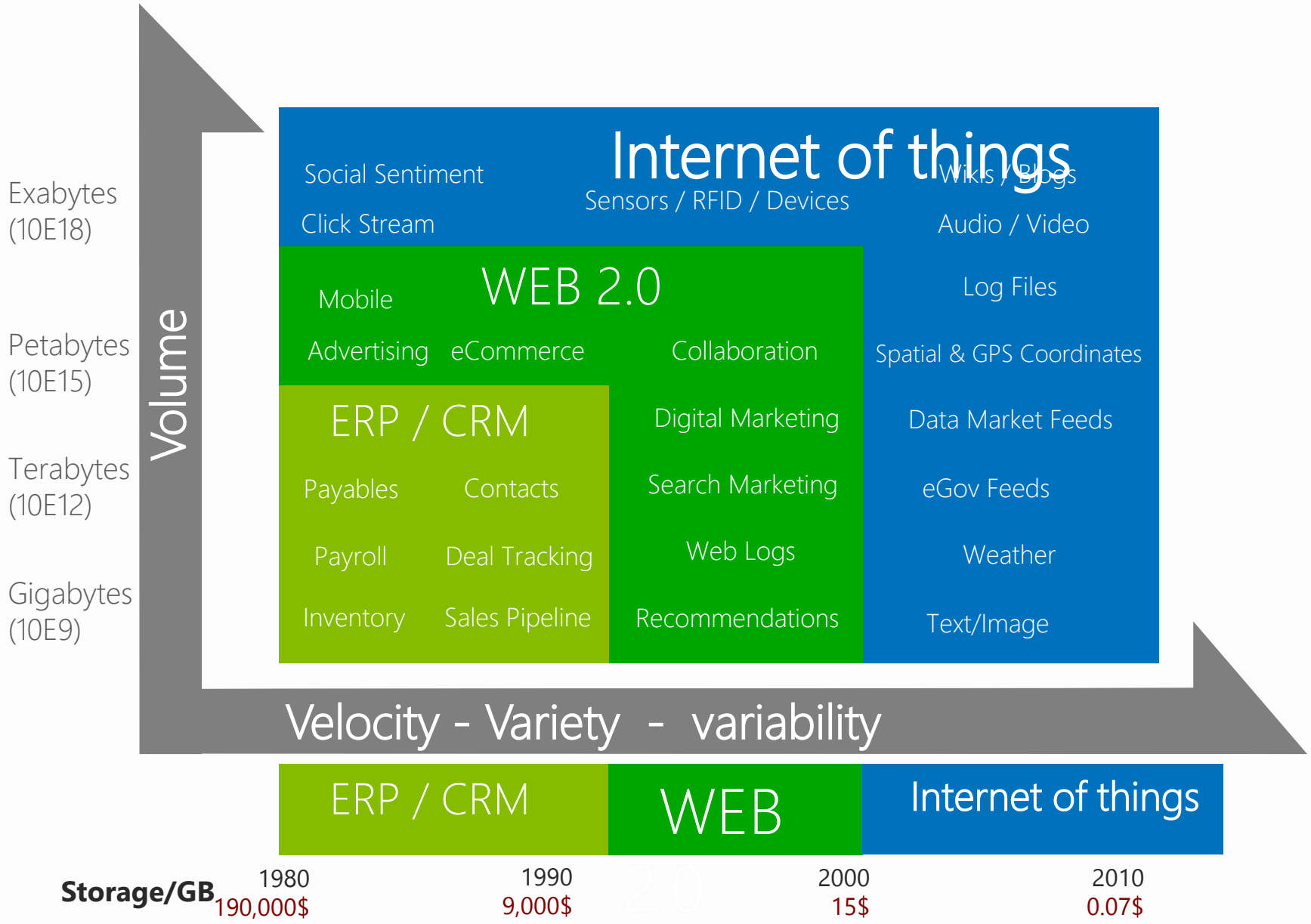


Understanding Big Data (Most of you already do)

KEY TRENDS


Device Explosion
>5.5 billion (70+% of global population)
Social Networks
>2 Billion users
Cheap Storage
$100 gets you 3million times more storage in 30 years)
Ubiquitous Connection
Web traffic
2010 130 Exabyte (10 E18)
2015 1.6 ZettaByte (10 E21)
Sensor Networks
>10 Billion
Inexpensive Computing
1980 10 MIPS/$
2005 10M MIPS/$

What is Big Data?



Big Data, BIG OPPORTUNITY

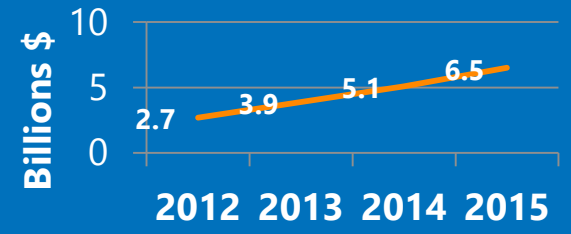
Big Data is a top priority for institutions

49% CEOs and CIOs are planning big data projects

Software Growth

34% compound annual growth rate[2]

Services Growth

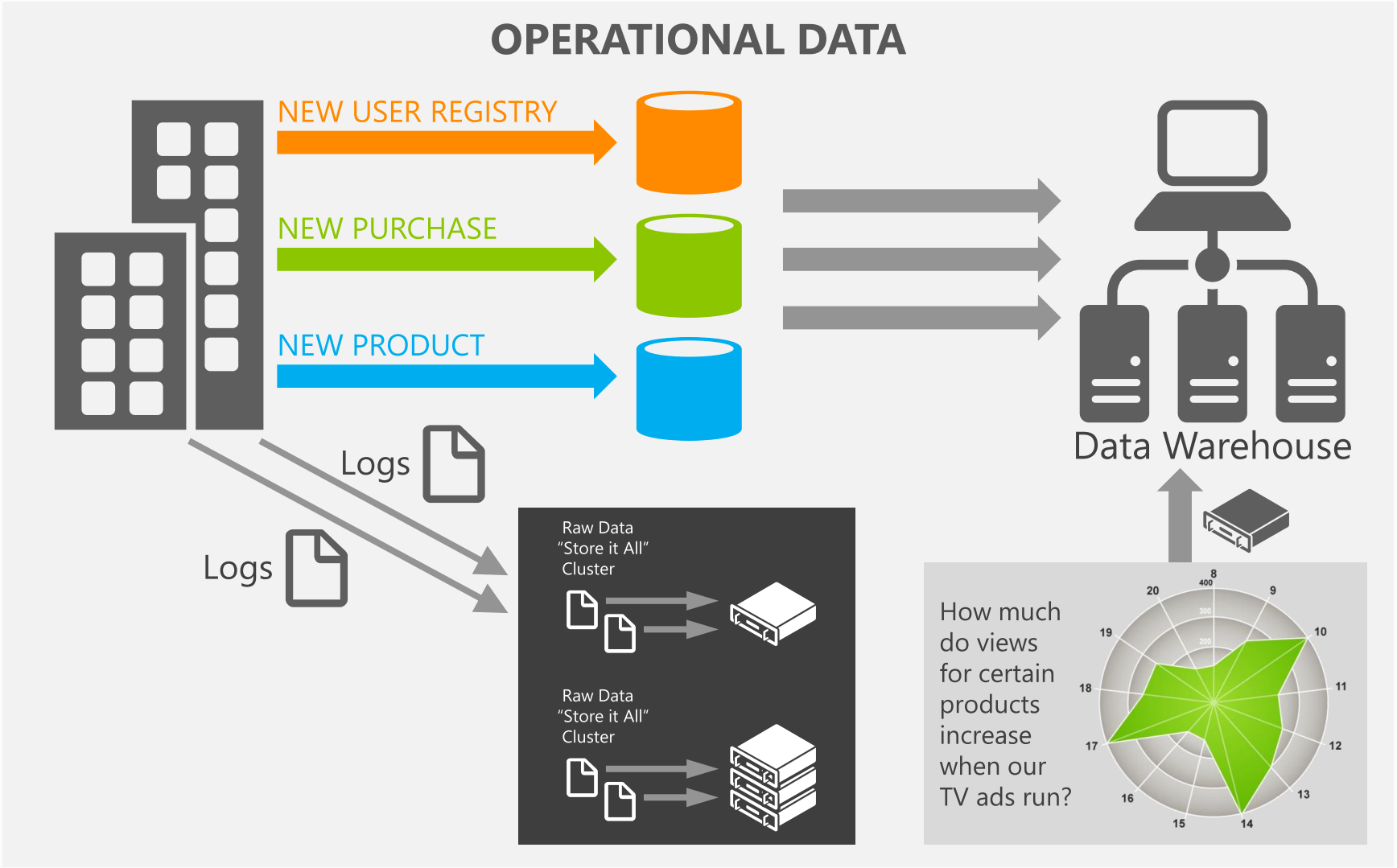


39% compound annual growth rate[2]

1. McKinsey&Company, McKinsey Global Survey Results, Minding Your Digital Business, 2012
2. IDC Market Analysis, Worldwide Big Data Technology and Services 2012–2015 Forecast , 2012

Big Data Scenarios

New workflow in Data Warehousing



Devices: Internet and Internet of things

Internet of things

Invisible devices

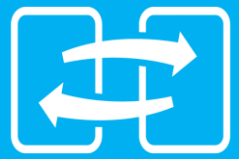
Trillions of computer-enabled devices which are part of the IoT

Trillions of networked nodes

100kBit/sec

Low bandwidth last-mile connection

Mostly addressed by local schemes



Machine-centric

Sensing-focus

Laptops / tablets / smartphones

6+billion people
1.5 billion use net
US: 4.3 devices per adult

Billions of networked devices

Cable: 10Mbs+
Fiber: 50-100Mbs

High-bandwidth access

Global addressing

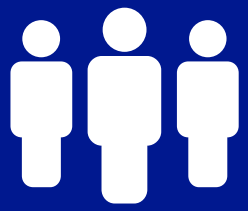
User-centric

Communication-focus

Internet

Collective Intelligence and Predictive analysis

What's the social sentiment of my product?

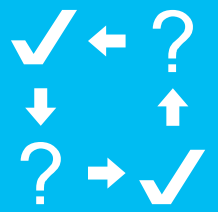


Social Analytics

Live Data Feed, Search

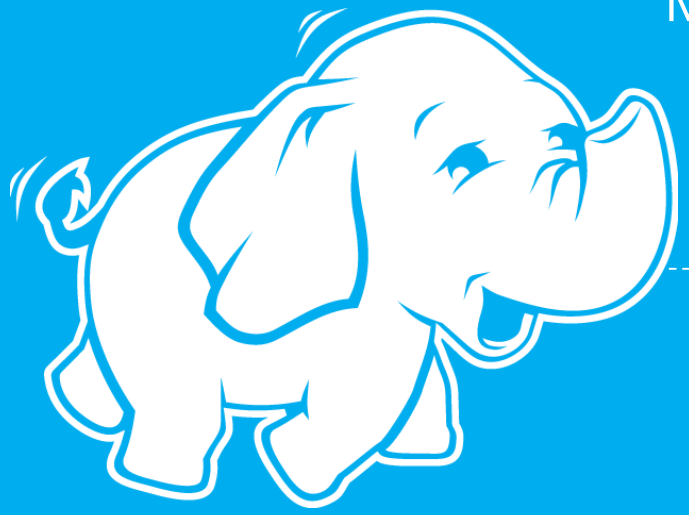
How do I optimize my services based on patterns of weather, traffic. How do I build a recommendation engine?

How do I better predict future outcomes?



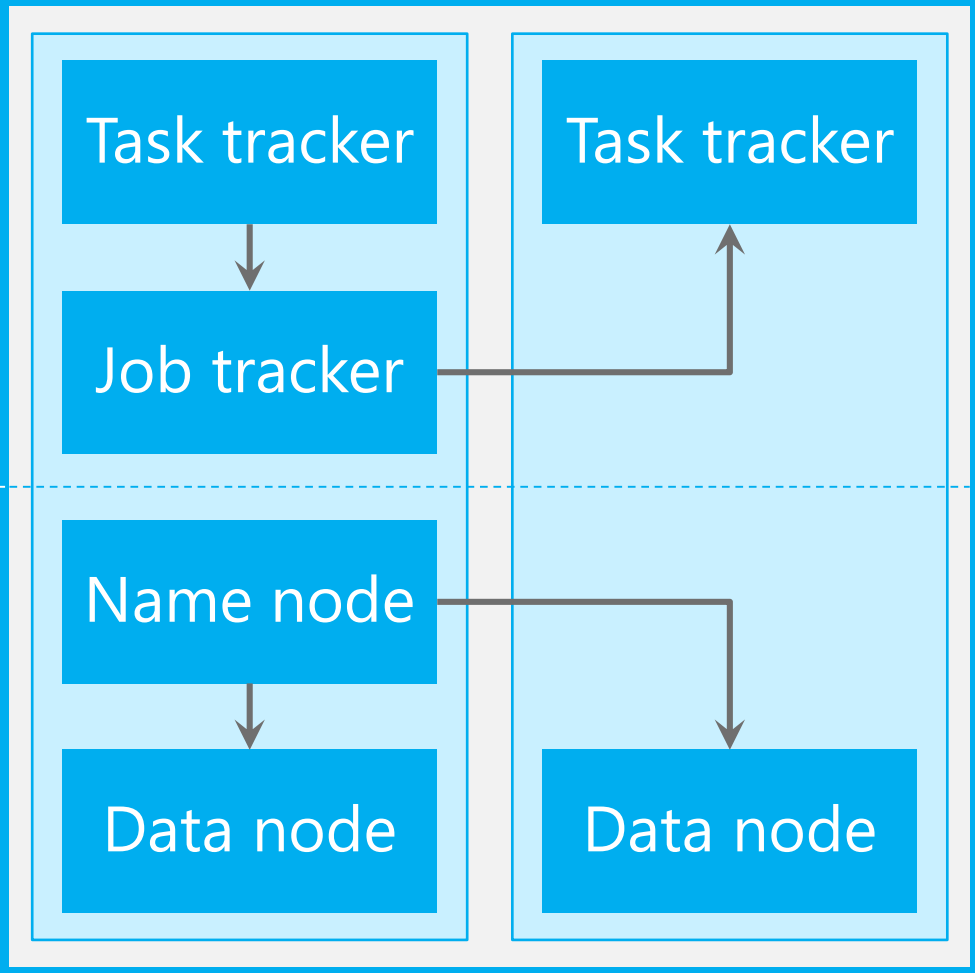
Advanced Analytics

Hadoop Distributed Architecture



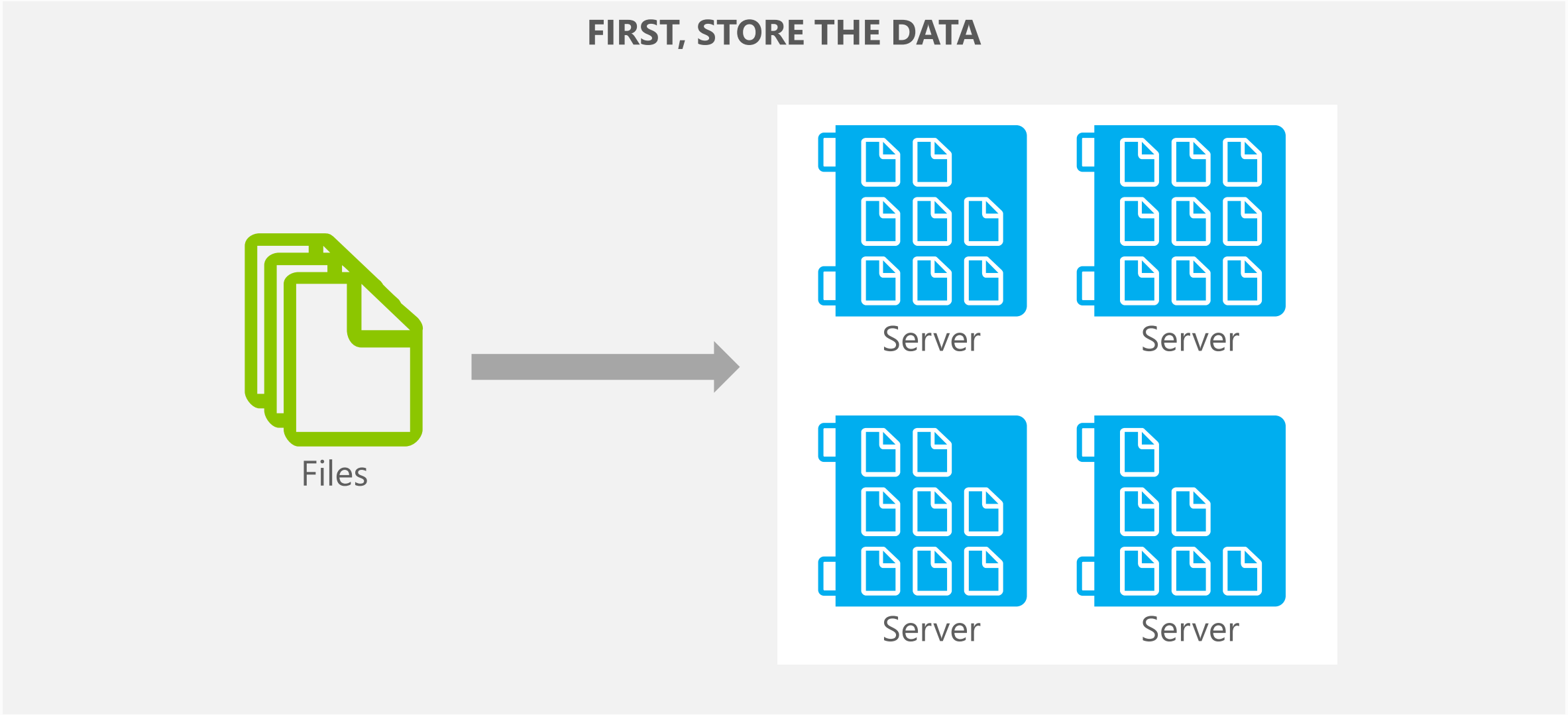
MapReduce Layer

HDFS Layer



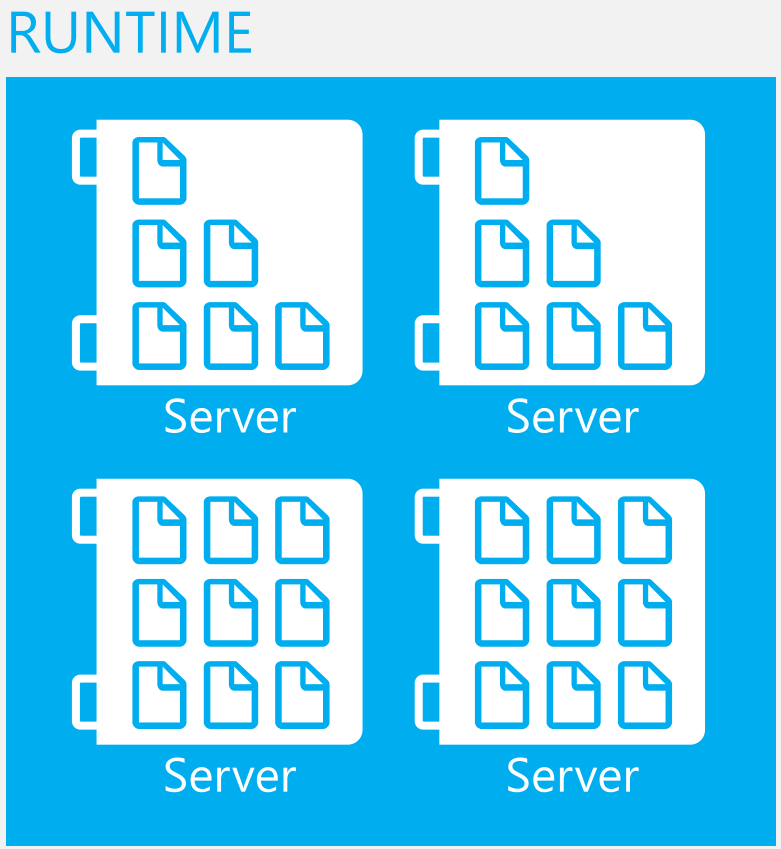
Reference: http://en.wikipedia.org/wiki/File:Hadoop_1.png

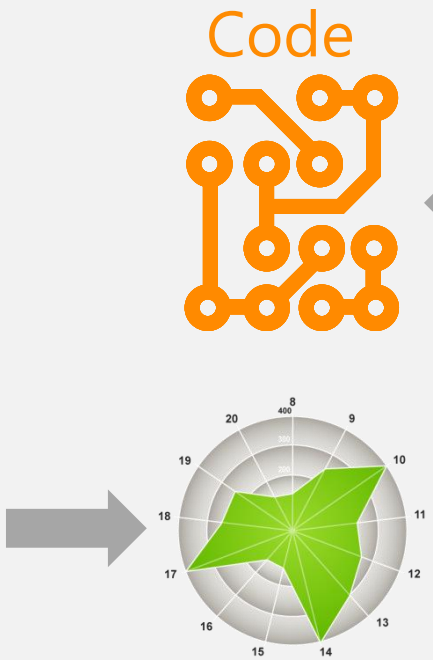
MapReduce: Move Code to the Data



So How Does It Work?

SECOND, TAKE THE PROCESSING TO THE DATA






```
// Map Reduce function in JavaScript

var map = function (key, value, context) {
var words = value.split(/[^a-zA-Z]/);
for (var i = 0; i < words.length; i++) {
          if (words[i] !== "")
context.write(words[i].toLowerCase(),
1);}
}};

var reduce = function (key, values, context)
{
var sum = 0;
while (values.hasNext()) {
sum += parseInt(values.next());
     }
context.write(key, sum);
};
```

MapReduce – Workflow

Data Acquisition & Modeling

Collaboration & Visualization

Analysis & Data Mining

Dissemination, Sharing, Preservation

"It *takes more time to hand a project from the seismic guys to me to the engineers in production than it does to figure out the oil field plays.*"

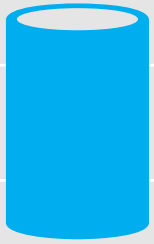
Geologist,
Major oil and gas company

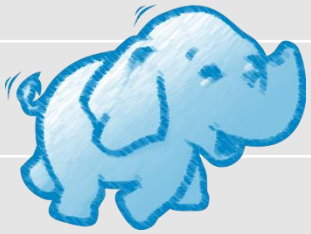
"*Our weather model and resulting data sets should be accessible to universities and other institutions.*"

Aerospace Development Manager,
U.S. Federal Government

Traditional RDBMS vs. NoSQL

	TRADITIONAL RDBMS	HADOOP
Data Size	Gigabytes *(Terabytes)*	Petabytes *(Hexabytes)*
Access	Interactive and Batch	Batch
Updates	Read / Write many times	Write once, Read many times
Structure	Static Schema	Dynamic Schema
Integrity	High (ACID)	Low
Scaling	Nonlinear	Linear
DBA Ratio	*1:40*	*1:3000*

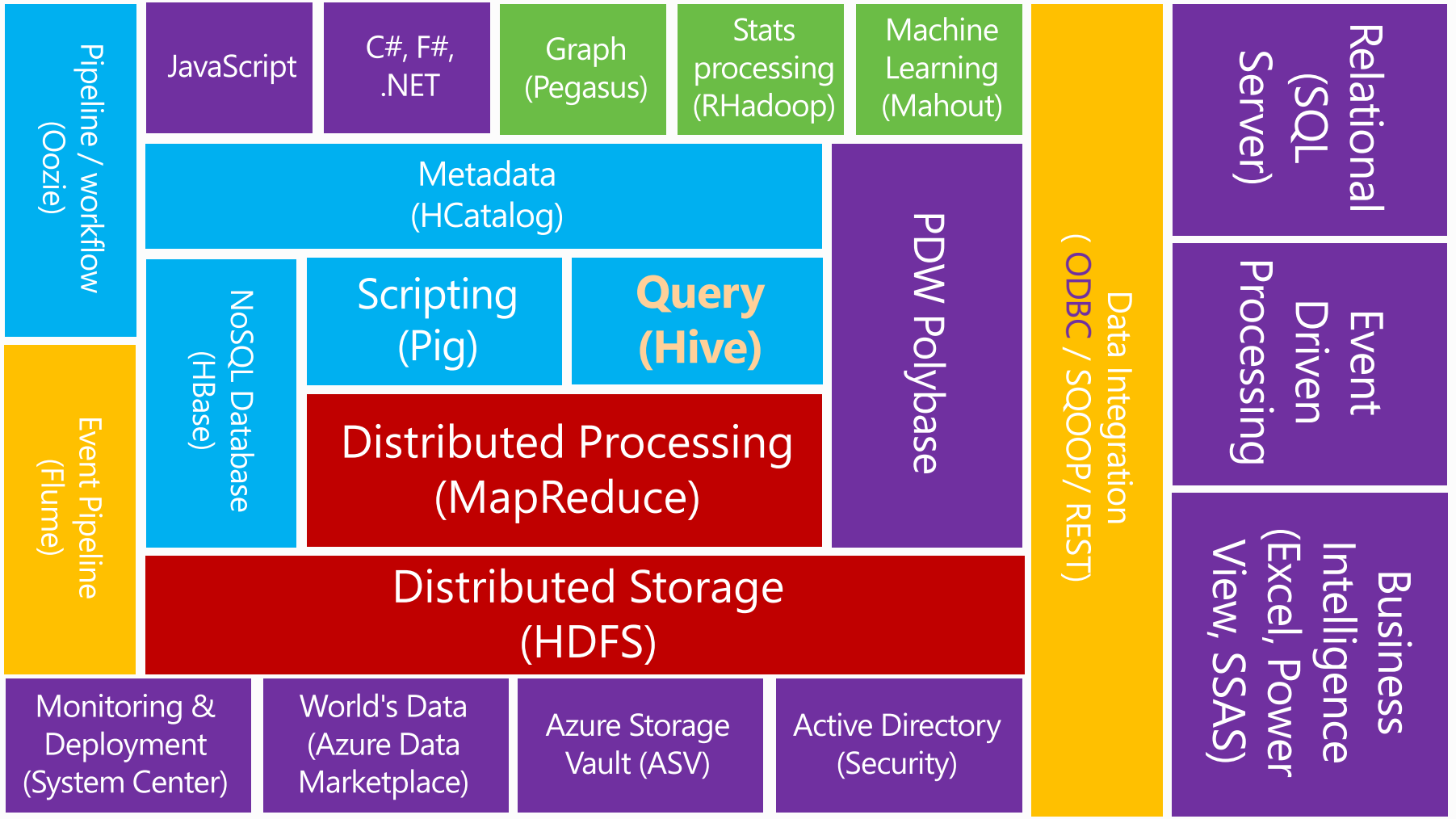




Reference: Tom White's Hadoop: The Definitive Guide

Windows Azure HDInsight Service

HDINSIGHT / HADOOP Eco-System



Legend

Red = Core Hadoop

Blue = Data processing

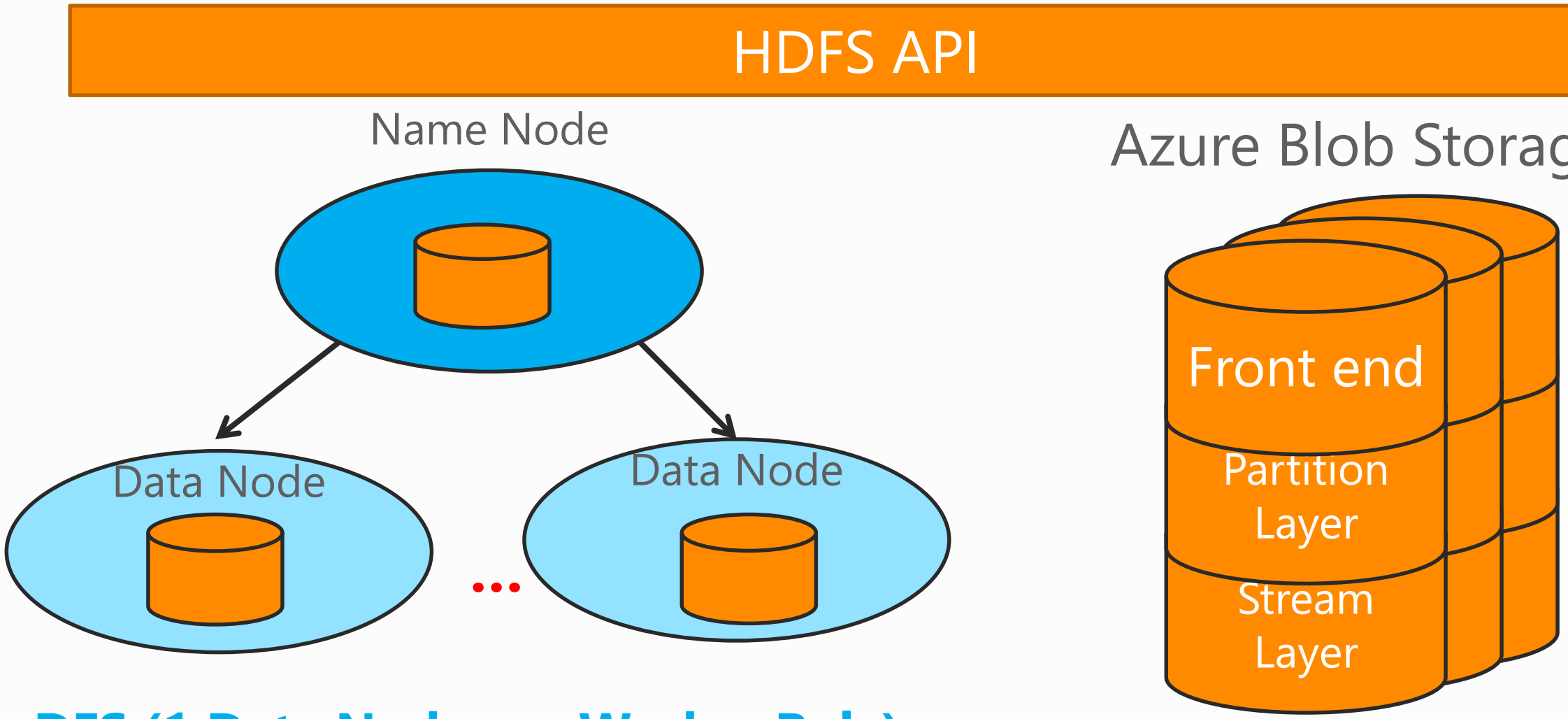
Purple = Microsoft integration points and value adds

Orange = Data Movement

Green = Packages

Storing Data with HDInsight

HDFS on Azure: Tale of two File Systems



DFS (1 Data Node per Worker Role) and Compute Cluster

Azure Storage (ASV)

Azure Storage (ASV)

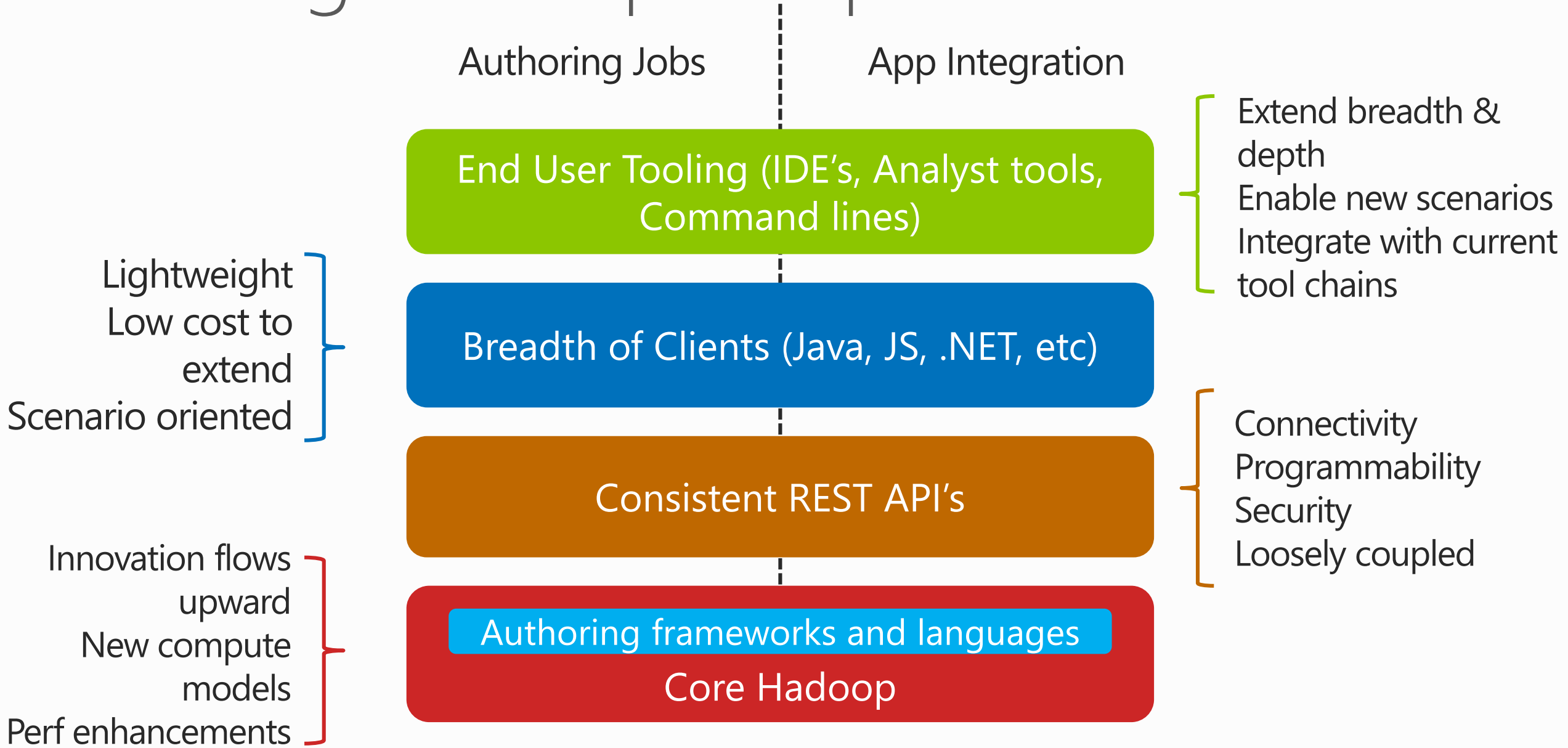
- Default file system for HDInsight Service
- Provides sharable, persistent, highly-scalable Storage with high availability (Azure Blob Store)
- Azure storage itself does not provide compute
- Fast access from compute nodes to data in same data center
- Several file systems, addressable via:
  `asv[s]:<container>@<account>.blob.core.windows.net/<path>`
- Requires storage key in core-site.xml:

```
<property>
  <name>fs.azure.account.key.accountname</name>
  <value>enterthekeyvaluehere</value>
</property>
```

Programming HDInsight

Existing Ecosystem	Hive, Pig, Mahout, Cascading, Scalding, Scoobi, Pegasus…
.NET	C#, F# Map/Reduce, LINQ to Hive, .NET management clients
JavaScript	JavaScript Map/Reduce, Browser hosted console, Node.js management clients
DevOps / IT Pros	PowerShell, Cross Platform CLI tools

Building Developer Experiences



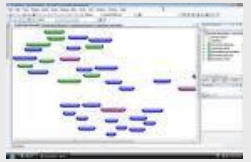
Microsoft Big Data Solution

FAMILIAR END USER TOOLS

Power View

Excel with PowerPivot

Predictive Analytics



Embedded BI

BI PLATFORM

SSAS

SSRS

HDInsight Service

Hadoop On Windows Server

Connectors

Microsoft PDW

UNSTRUCTURED & STRUCTURED DATA

Sensors

Devices

Bots

Crawlers

ERP

CRM

LOB

APPs

Deploying and Interacting With HDInsight Service

demo

Microsoft Hadoop Vision

Insights to all users by activating new types of data

INSIGHTS	Integrate with Microsoft Business Intelligence
ENTERPRISE READY	Choice of deployment on Windows Server + Windows Azure Integrate with Windows Components (AD, Systems Center)
BROADER ACCESS	Easy installation and configuration of Hadoop on Windows Simplified programming with . Net & Javascript integration Integrate with SQL Server Data Warehousing

Contributions proposed back to community distribution

Big Data Processing

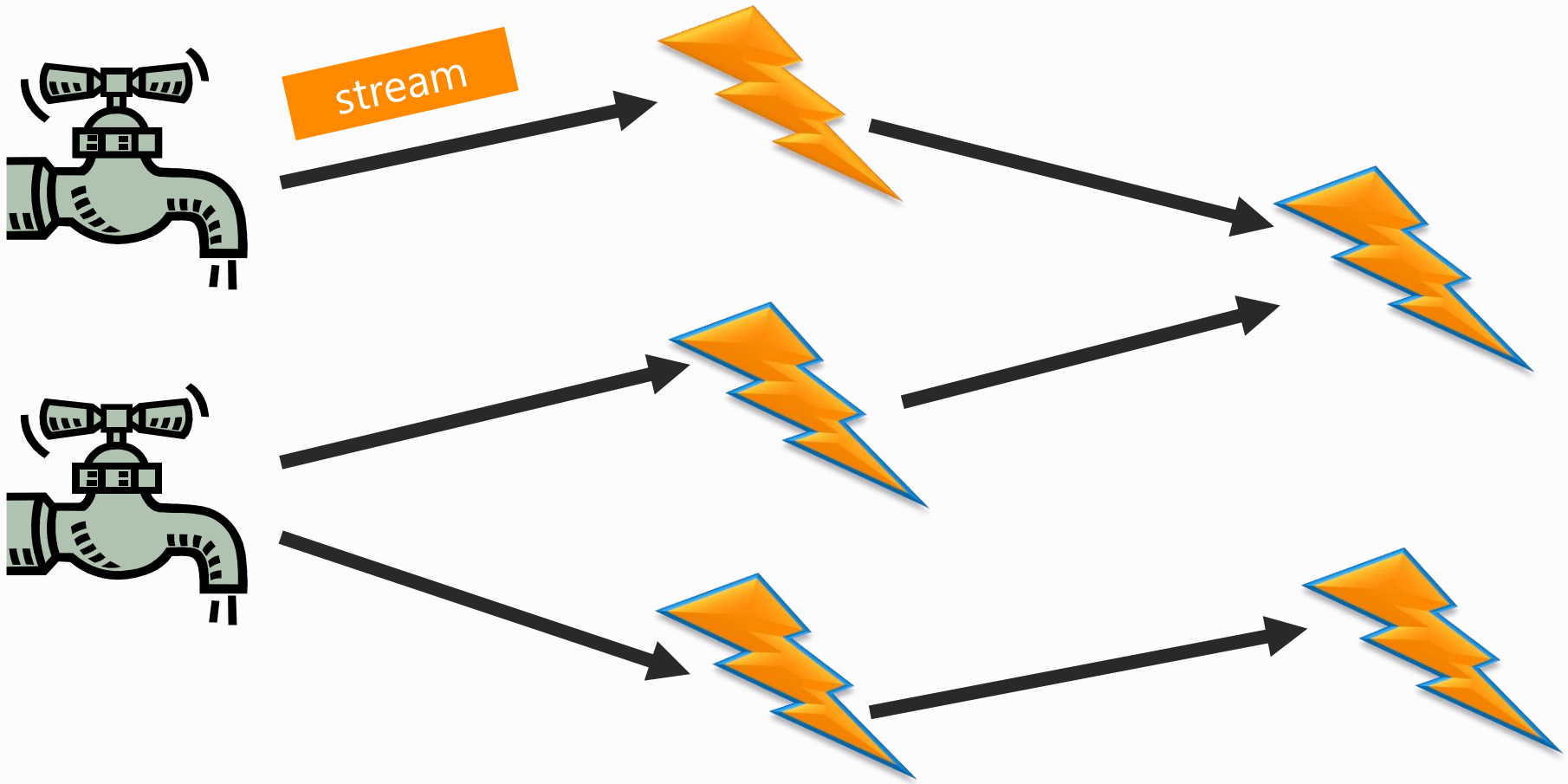
	Batch Processing	Interactive analysis	Stream processing
Query runtime	Minutes to hours	Milliseconds to minutes	Never-ending
Data volume	TBs to PBs	GBs to PBs	Continuous stream
Programming model	MapReduce	Queries	DAG
Users	Developers	Analysts and developers	Developers
Originating project	Google MapReduce	Google Dremel	Twitter Storm
Open source project	Hadoop / Spark	Drill / Shark /Impala Hbase / Cassandra	Storm / Apache S4 /Kafka

STORM Concepts

- Streams

- Spouts: Source of streams
- Bolts: Functions, Filters, Aggregation, Joins, DB R/W
- Topologies: Grouping of Spouts and Bolts

Topology: Network of Spouts and Bolts



Thank you!

Resources

- Windows Azure Developer Website
  - http://www.windowsazure.com/en-us/develop/net/
- Windows Azure Storage Blog
  - http://blogs.msdn.com/b/windowsazurestorage/
- SOSP Paper/Talk
  - http://blogs.msdn.com/b/windowsazurestorage/archive/2011/11/20/windows-azure-storage-a-highly-available-cloud-storage-service-with-strong-consistency.aspx

Resources

- Windows Azure Python SDK
- Windows Azure
- **How to use Service Management from Python**
- **http://www.windowsazure.com/en-us/manage/linux/other-resources/command-line-tools/**
- **http://research.microsoft.com/en-us/projects/azure/**

- Follow @wenmingye for Questions and latest info.