

Microsoft
Research



Microsoft Research Asia
Faculty Summit 2010



Cloud Computing Paradigm Shift

Enwei Xie

General Manager, Developer & Platform Evangelism

Microsoft China



What is Cloud Computing?

*“A style of computing where **SCALABLE** and **ELASTIC** IT-enabled capabilities are provided as a **service** to external customers using Internet technologies.”*

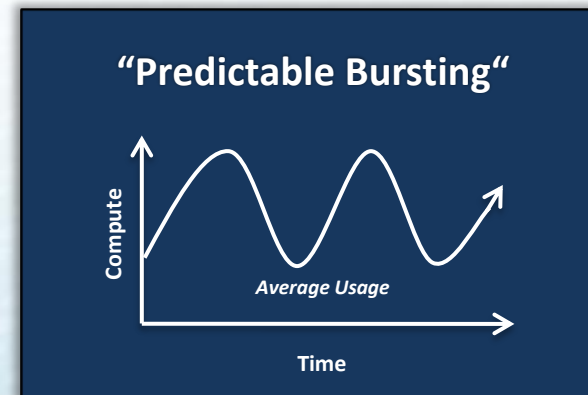
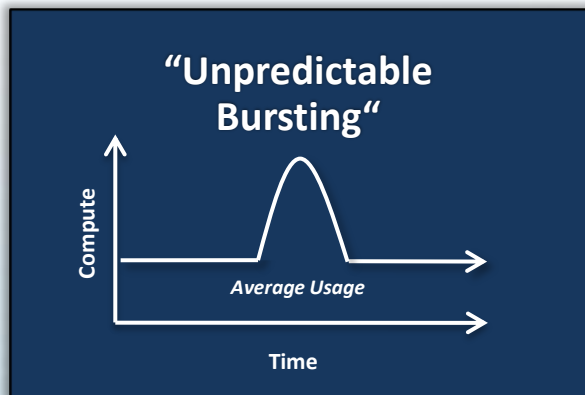
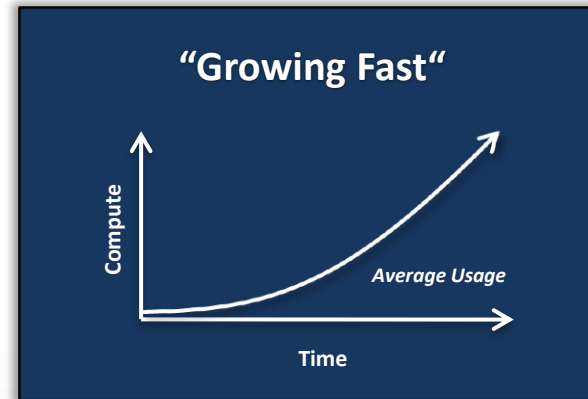
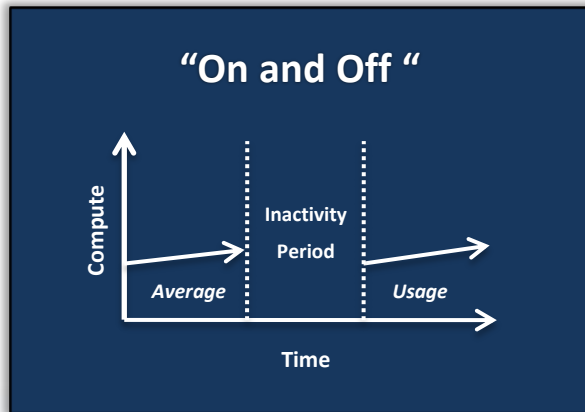
Gartner, Inc. “Hype Cycle for Cloud Computing,” 2009, by Daryl Plummer et al, July 2009.

*“A standardized IT capability, such as **SOFTWARE, APP PLATFORM, OR INFRASTRUCTURE**, delivered via Internet technologies in a **pay-per-use** and **self-service** way.”*

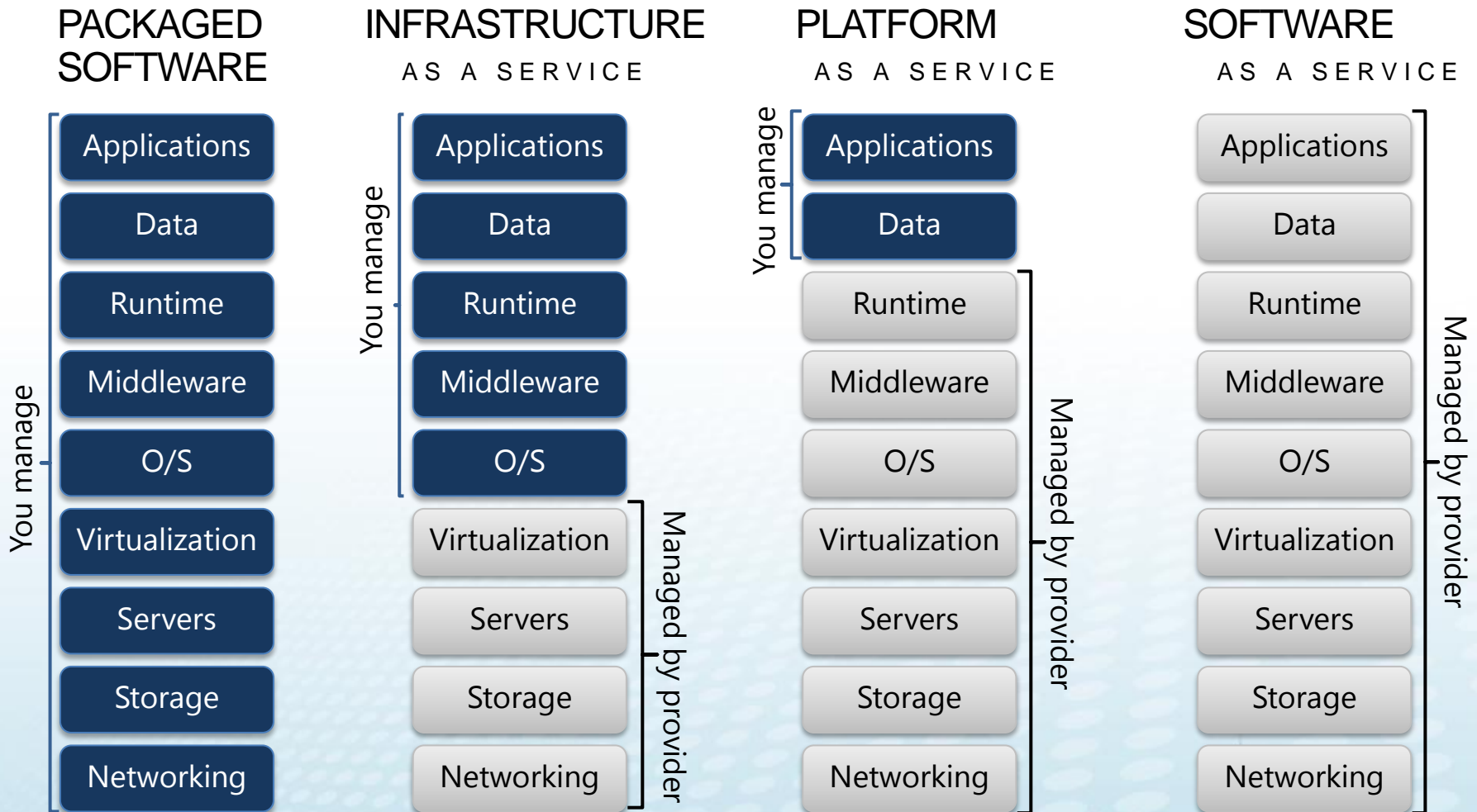
“How To Message “Cloud” Offerings And Not Get Lost In The Fog,” Forrester Research, Inc., July 2009.



Why Cloud Computing?



IT as a Service





Server & Services Platforms

SERVICES

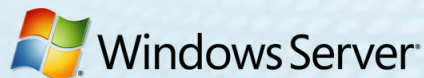
PLATFORM



STANDARDIZED SERVICE
LOWEST OPERATIONS COST
UPDATED BY MICROSOFT

SERVER

PLATFORM



CUSTOMIZABLE PRODUCT
SUPPORTS ALL EXISTING APPS
LOW OPERATIONS COST
UPDATED BY CUSTOMER

Windows Azure Platform



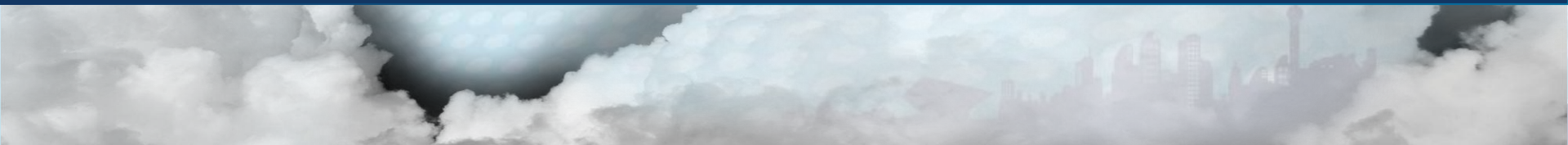
- Scalable compute and storage
- Automated service management
- Familiar tools, technologies, languages



- Relational storage for the cloud
- Consistent development model
- Automated database management



- Connect existing apps to the cloud
- Connect through network boundaries
- Easily control authorization to apps



Codename “Dallas”

Information Discovery

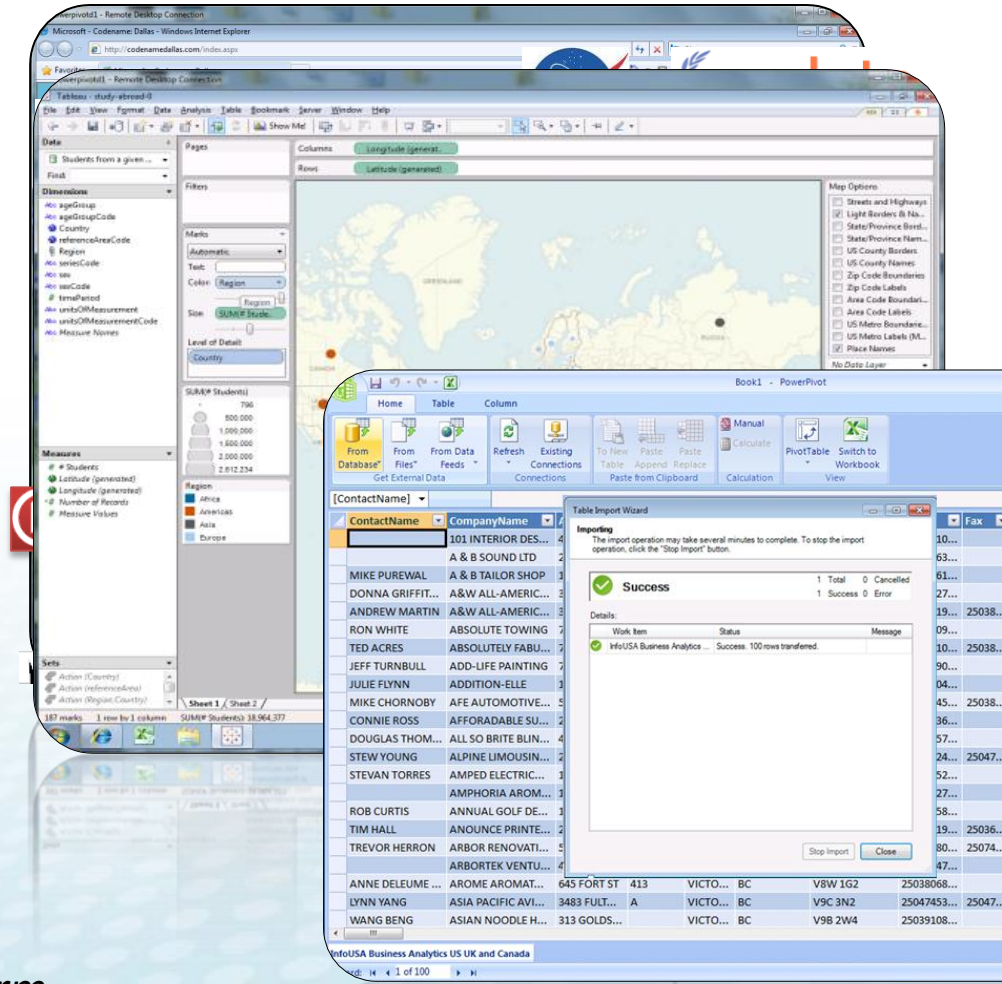
Discover, acquire, and consume structured and blob datasets to power any application – on any platform and any screen size.

Data Brokerage

Partner driven ecosystem and global reach to deliver data and functionality to developers and information workers.

Analytics and Reporting

Augment private data with premium commercial and public domain data: on premises or in the Cloud.



The screenshot displays the Microsoft Codename Dallas interface. On the left, a sidebar shows a 'Data' pane with a tree view containing fields like 'Students from a given...', 'Country', and 'Region'. Below it is a 'Measures' pane with '# Students' and 'SUM(SUM Students)'. A 'Map' pane shows a world map with a red dot in the US. A 'Table Import Wizard' dialog box is open in the foreground, showing a 'Success' message: 'InfoUSA Business Analytics - Success: 100 rows transferred.' Below the dialog, a data table is visible with columns for 'ContactName', 'CompanyName', and other fields.

ContactName	CompanyName	Address	City	Region	PostalCode	Country
MIKE PUREWAL	A & B TAILOR SHOP	101 INTERIOR DES...				
DONNA GRIFFIT...	A&W ALL-AMERIC...					
ANDREW MARTIN	A&W ALL-AMERIC...					
RON WHITE	ABSOLUTE TOWING					
TED ACRES	ABSOLUTELY FABU...					
JEFF TURNBULL	ADD-LIFE PAINTING					
JULIE FLYNN	ADDITION-ELLE					
MIKE CHORNOBY	AFE AUTOMOTIVE...					
CONNIE ROSS	AFFORDABLE SU...					
DOUGLAS THOM...	ALL SO BRITE BLU...					
STEVAN TORRES	ALPINE LIMOUSIN...					
ROB CURTIS	AMPED ELECTRIC...					
TIM HALL	ANNUAL GOLF DE...					
TREVOR HERRON	ARBOR RENOVATI...					
ANNE DELEUME ...	AROME AROMAT...	645 FORT ST	413	VICTO...	BC	V8W 1G2
LYNN YANG	ASIA PACIFIC AVI...	3483 FULT...	A	VICTO...	BC	V9C 3N2
WANG BENG	ASIAN NOODLE H...	313 GOLDS...		VICTO...	BC	V9B 2W4

All powered by the Windows Azure platform

Windows Azure Platform Data Centers

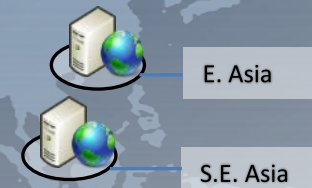
North America Region



Europe Region



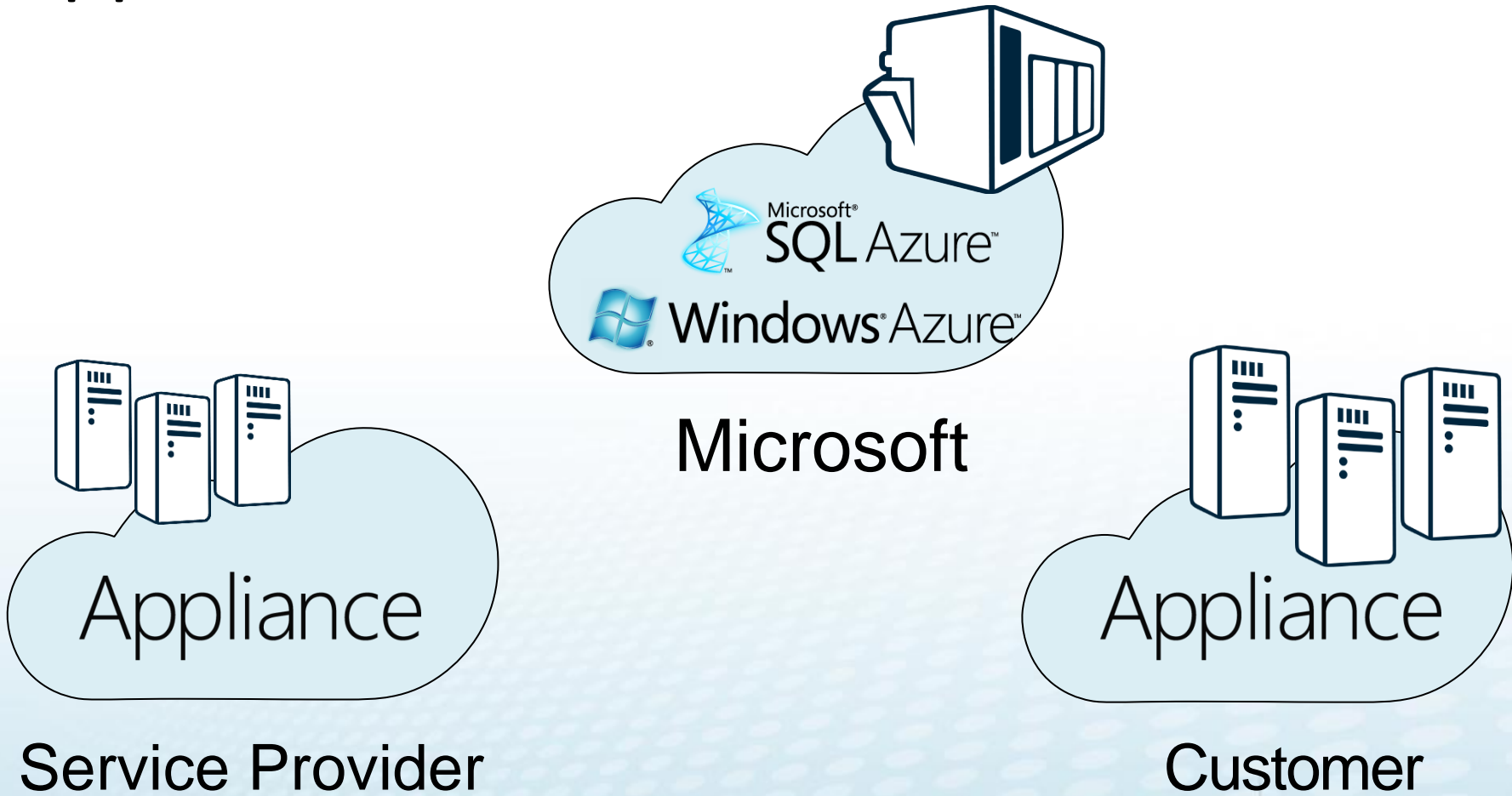
Asia Pacific Region



6 datacenters across 3 continents

Simply select your data center of choice when deploying an application

Windows Azure Platform Appliance



Scalability, Elasticity & Multi-tenancy

PARADIGM SHIFT



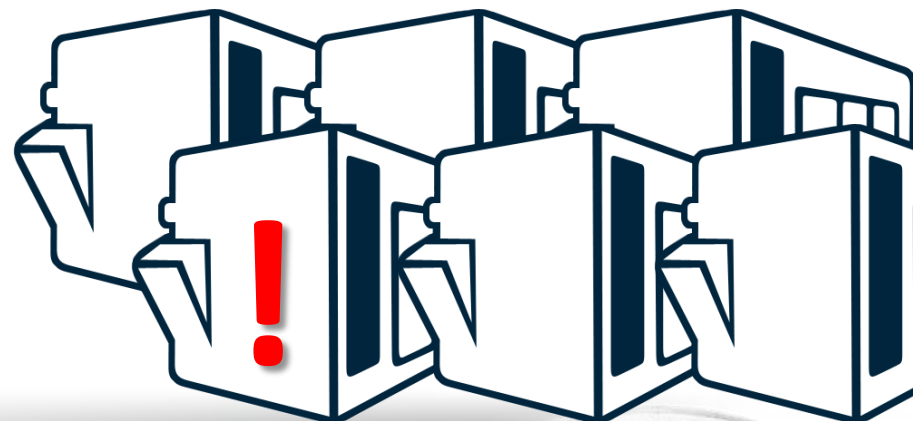
Scalability & Elasticity

Manage thousands of servers per admin

Application and hardware failures are expected

Constant feedback about usage and performance

Globally distributed workloads



Paradigm Shift

Scalability & Elasticity

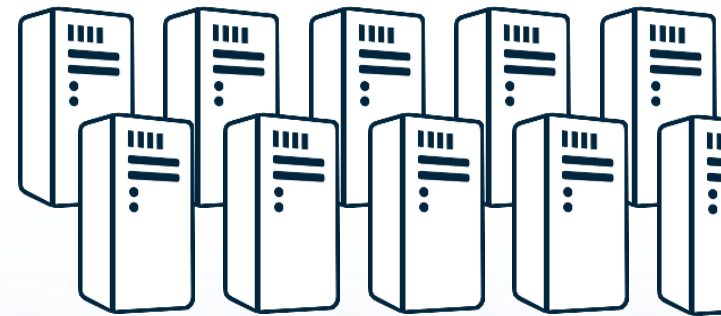
Quickly increase and decrease resources

Special considerations for instance startup and synchronization

Manage state across instances

Queues provide the glue for work coordination

```
<?xml version="1.0"?>  
...  
<Instances count="20"/>  
...  
</ServiceConfiguration>
```



```
CREATE DATABASE foo ...;
```

```
CREATE DATABASE foo_copy1 AS COPY OF foo;  
CREATE DATABASE foo_copy2 AS COPY OF foo;  
CREATE DATABASE foo_copy3 AS COPY OF foo;  
CREATE DATABASE foo_copy4 AS COPY OF foo;
```



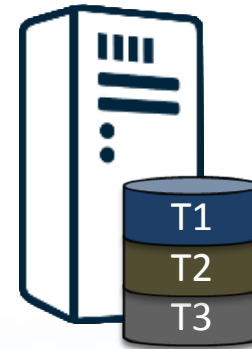
Paradigm Shift

Multi-tenancy

Oftentimes the key for cost reduction

Data partitioning and synchronization becomes a key architecture concern

Tenant isolation is critical for security and minimizing side affects



Biological Sequence Comparison in the Cloud

AZUREBLAST

In 15 years we'll have all the sequence, a list of the genes everyone has in common and those that differ among people. We know only something like a tenth of 1 percent of the sequence at the moment...

[Walter Gilbert](#)

All-Against-All Experiment

Discover the interrelationships of known protein sequences
Theoretically, 100 billion sequence comparisons, estimated 6.1 years on desktop

Allocated a total of ~4000 instances

475 extra-large VMs (8 cores per VM) across four datacenters worldwide

8 deployments of AzureBLAST w/co-located storage service

Divide 10 million sequences into multiple segments

Output result is ~230GB with 1,764,579,487 total hits



Large-scale Parallel Machine Learning Toolkit

SIGMA



Sigma Architecture

Sigma Azure UI
<http://msrasigma.cloudapp.net/>

Machine Learning algorithms

Sigma Algorithm Framework
for Windows Azure

Sigma Azure
Job System

Sigma Azure
Data System



Windows® Azure™

LET'S MOVE FORWARD TOGETHER



Call To Action

Engage with Microsoft to help us solve the hard problems

Teach distributed and scale out programming

Build your own Cloud or partner with Cloud Providers for hands-on experience



Microsoft[®]