# Cloud Computing Challenges and Potential

Dennis Gannon
Cloud Computing Futures
eXtreme Computing Group
Microsoft Research

Microsoft

# Outline

- Cloud Concepts
  - Data Center Architecture
  - The cloud flavors: IaaS, PaaS, SaaS
- Our world of client devices plus the cloud
- Programming a Cloud Application
- Science in the Cloud

# The Cloud

- A model of computation and data storage based on “pay as you go” access to “unlimited” remote data center capabilities.
- A cloud infrastructure provides a framework to manage scalable, reliable, on-demand access to applications.
- Examples:
  - Search, email, social networks
  - File storage (Live Mesh, Mobile Me, Flicker, …)
  - Just about any large-scale web service is a cloud service.




Microsoft

# The Current Cloud Challenge

- The current driver: how do you
  - Support email for 375 million users?
  - Store and index 6.75 trillion photos?
  - Support 10 billion web search queries/month?
- And
  - deliver deliver a quality response in 0.15 seconds to millions of simultaneous users?
  - never go down.
- The future applications of the cloud go well beyond web search
  - The data explosion
  - The merger of the client (phone, laptop, your personal sensors) with the cloud.

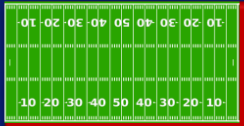
# The Physical Architecture of Clouds

Microsoft

# Clouds are built on Data Centers

- Range in size from “edge” facilities to megascale.
- Economies of scale
  - Approximate costs for a small size center (1000 servers) and a larger, 100K server center.

| Technology | Cost in small-sized Data Center | Cost in Large Data Center | Ratio |
|---|---|---|---|
| Network | $95 per Mbps/ month | $13 per Mbps/ month | 7.1 |
| Storage | $2.20 per GB/ month | $0.40 per GB/ month | 5.7 |
| Administration | ~140 servers/ Administrator | >1000 Servers/ Administrator | 7.1 |



Each data center is
**11.5 times**
the size of a football field

Microsoft

# The Challenge of Data Centers & Apps

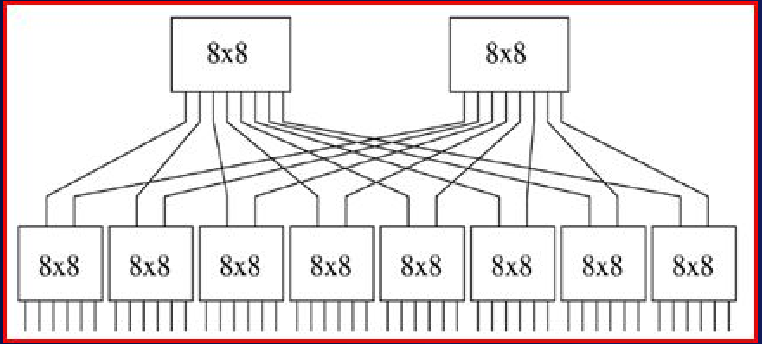
- The impact on the environment
  - In 2006 data centers used 61 *Tera*watt-hours of power
    - 1.5 to 3% of US electrical energy consumption today
  - Great advances are underway in power reduction
- With 100K+ servers and apps that must run 24x7 constant failure must be an axiom of hardware and software design.
  - Huge implication for the application design model.
  - How can hardware be designed to degrade gracefully?
- Two dimensions of parallelism
  - Scaling apps from 1 to 1,000,000 simultaneous users
  - Some apps require massive parallelism to satisfy a single request in less than a second.


Microsoft

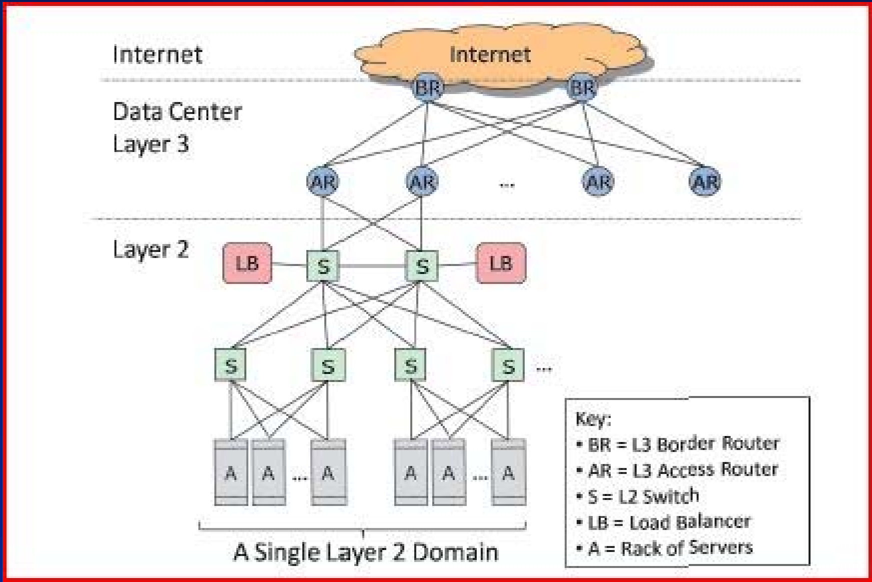
# Data Center vs Supercomputers

- Scale
  - Blue Waters = 40K 8-core "servers"
  - Road Runner = 13K cell + 6K AMD servers
  - MS Chicago Data Center = 50 containers = 100K 8-core servers.
- Network Architecture
  - Supercomputers: CLOS "Fat Tree" infiniband
    - Low latency – high bandwidth
    - protocols
  - Data Center: IP based
    - Optimized for Internet Access
- Data Storage
  - Supers: separate data farm
    - GPFS or other parallel file system
  - DCs: use disk on node + memcache

Fat tree network

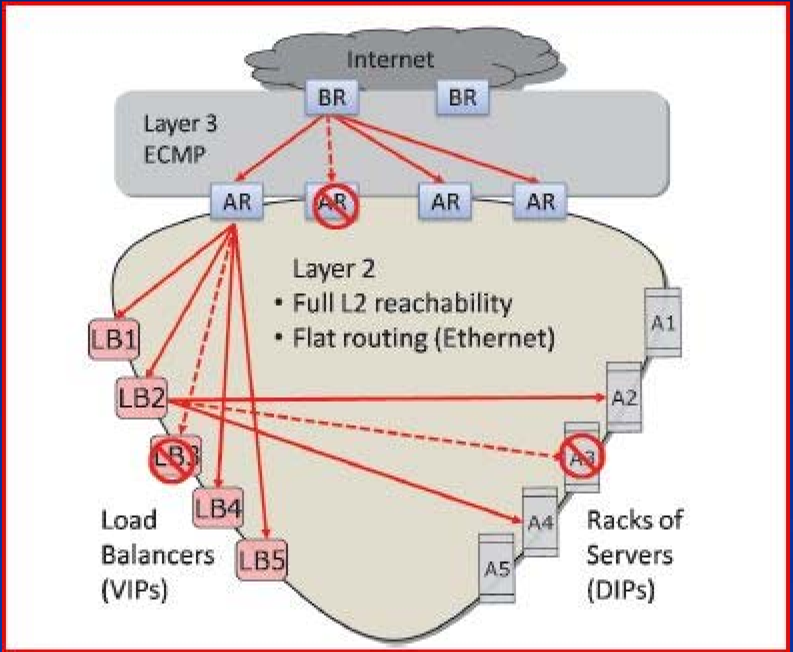


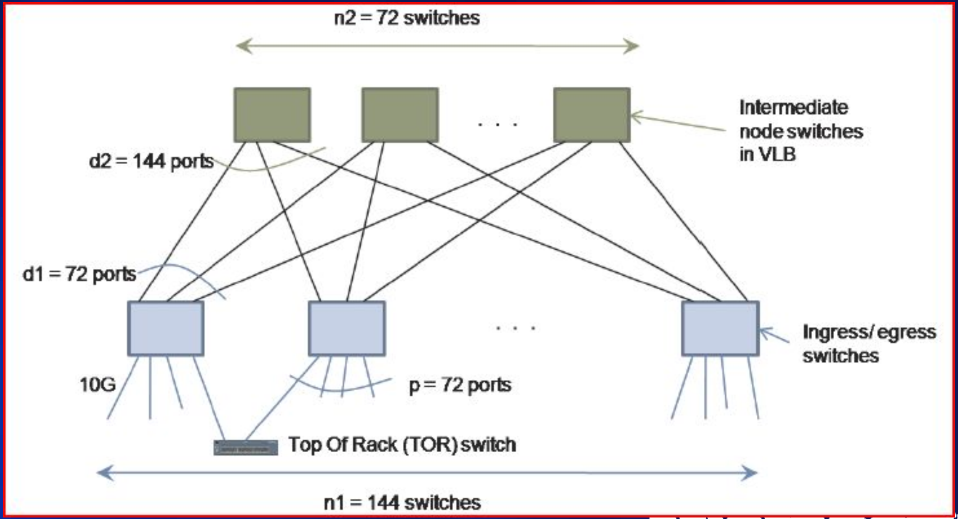

Standard Data Center Network



# Next Gen Data Center Networks

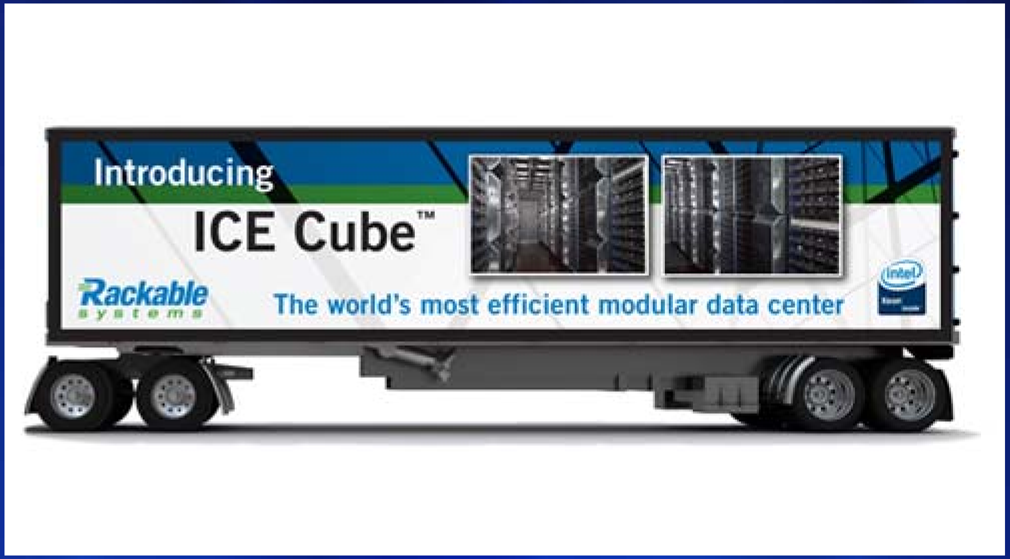
- Monsoon
  - Work by Albert Greenberg, Parantap Lahiri, David A. Maltz, Parveen Patel, Sudipta Sengupta.
  - Designed to scale to 100K+ server data centers.
  - Flat server address space instead of dozens of VLANS.
  - Valiant Load Balancing.
  - Allows a mix of apps and dynamic scaling.
  - Strong fault tolerance characteristics.



# Advances in DC deployment

- Conquering complexity.
  - Building racks of servers & complex cooling systems all separately is not efficient.
  - Package and deploy into bigger units:



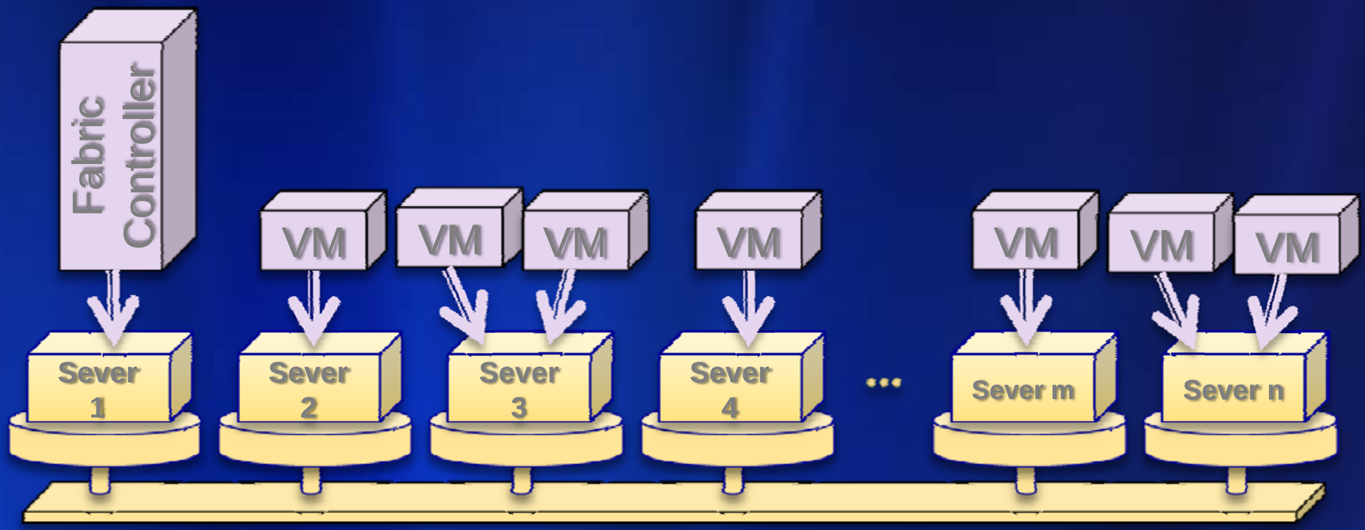

Generation 4 data center video

# Containers: Separating Concers


# Cloud Software Models

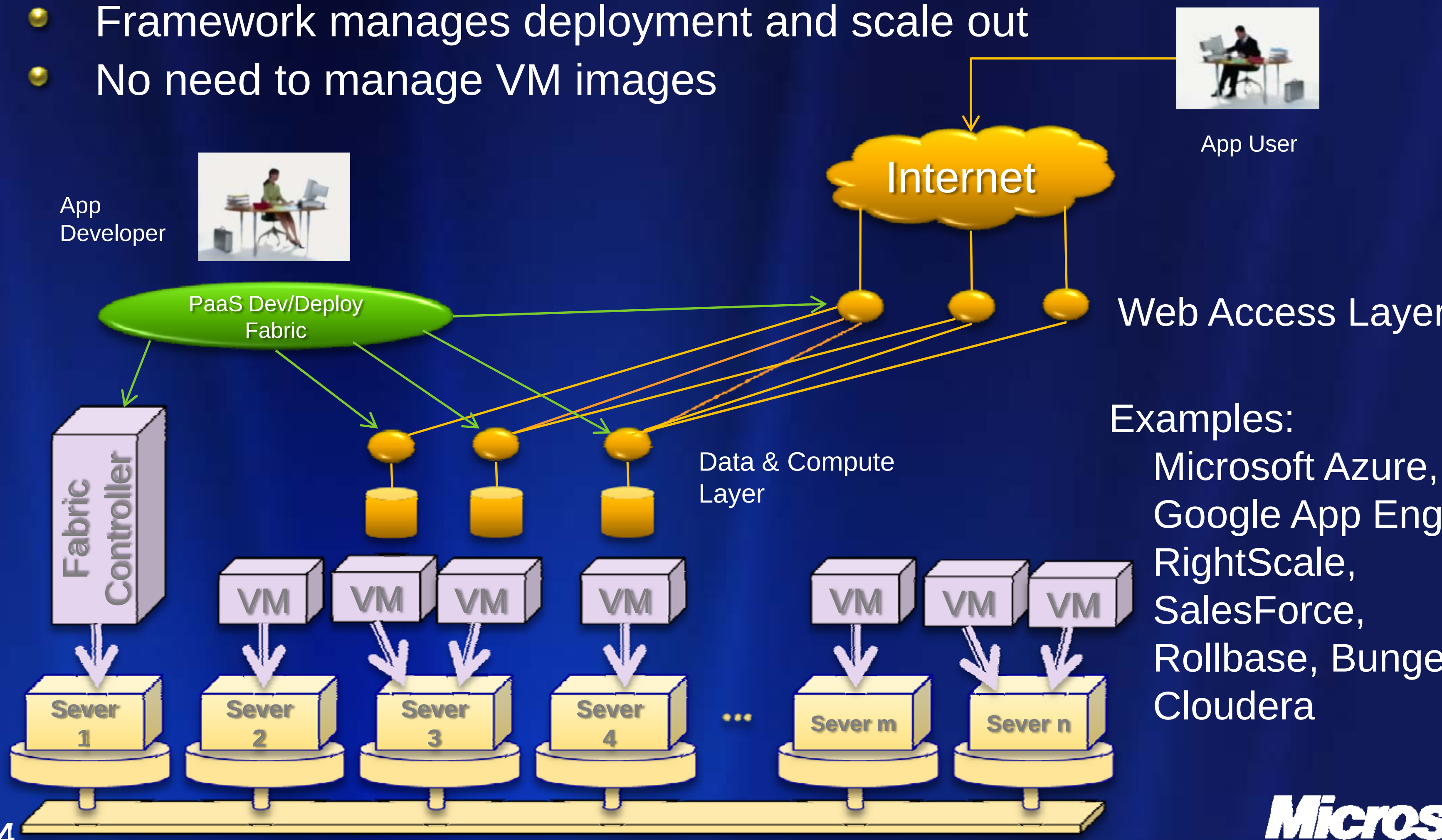
# Three Levels of Cloud Arcitecture

- Infrastructure as a Service (IaaS)
  - Provide App builders a way to configure a Virtual Machine and deploy one or more instances on the data center
  - The VM has an IP Address visible to the world
  - A Fabric controller manages VM instances
- Examples: Eucalyptus.com, Amazon EC2 + S3, Flexiscale, Rackspace, GoGrid, SliceHost, Nimbus



# Platform as a Service

- An application development, deployment and management fabric.
- User programs web service front end and computational & Data Services
- Framework manages deployment and scale out
- No need to manage VM images




Examples:
Microsoft Azure,
Google App Engine,
RightScale,
SalesForce,
Rollbase, Bungee,
Cloudera

# Software as a Service

- Online delivery of applications
- Via Browser
  - Microsoft Office Live Workspace
  - Google Docs, etc.
  - File synchronization in the cloud – Live Mesh, Mobile Me
  - Social Networks, Photo sharing, Facebook, wikipedia etc.
- Via Rich Apps
  - Science tools with cloud back-ends
    - Matlab, Mathematica
  - Mapping
    - MS Virtual Earth,  Google Earth
  - Much more to come.

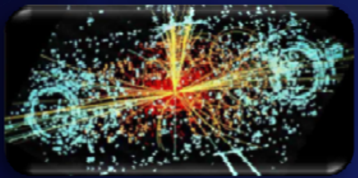
# The Clients+Cloud Platform

- At one time the “client” was a PC + browser.
- Now
  - The Phone
  - The laptop/tablet
  - The TV/Surface/Media wall
- And the future
  - The instrumented room
  - Aware and active surfaces
  - Voice and gesture recognition
  - Knowledge of where we are
  - Knowledge of our health


# The Future: an Explosion of Data

**Experiments**

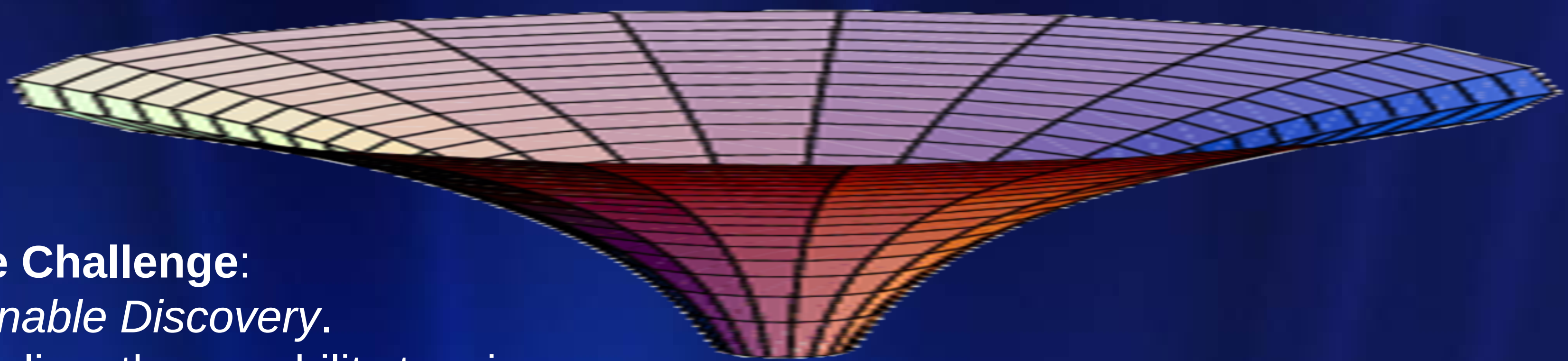
**Simulations**



**Archives**

**Literature**


**Instruments**



**The Challenge**:

*Enable Discovery*.
Deliver the capability to mine, search and analyze this data in near real time.

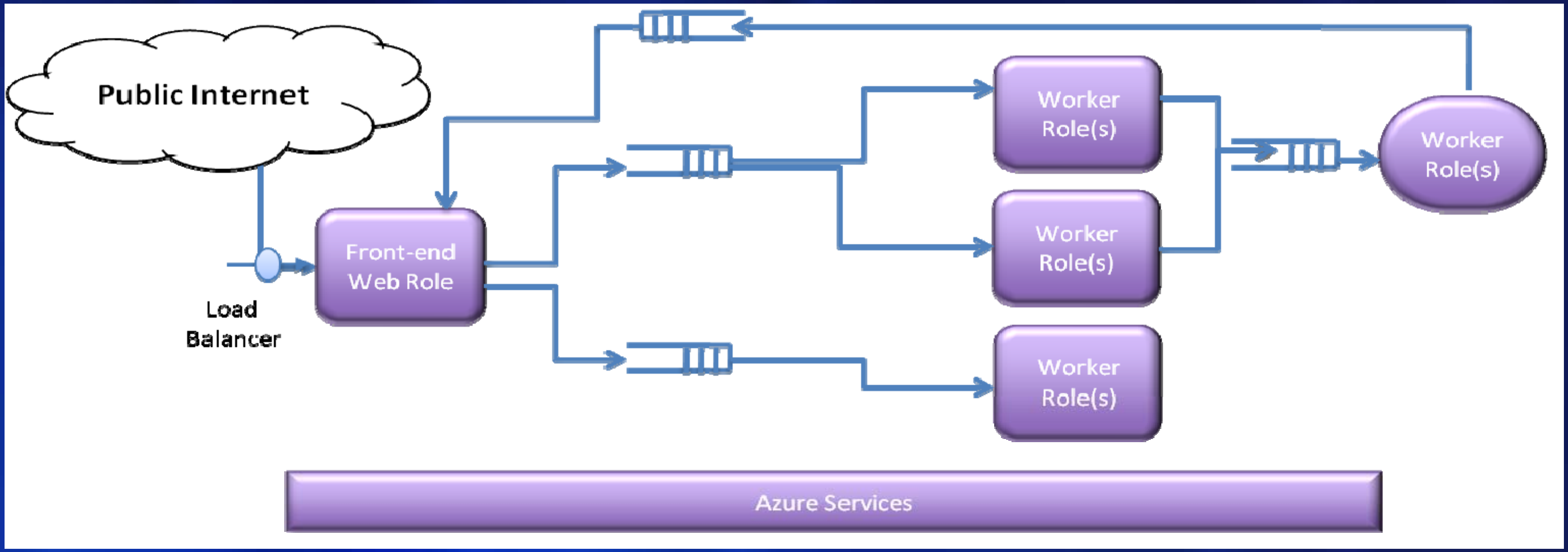
*Enhance our Lives*
Participate in our own heath care. Augment experience with deeper understanding.

**<u>Petabytes</u>**
**Doubling every 2 years**

# The Architecture of an Azure App

Roles are a mostly stateless process running on a core.

- *Web Roles provide web service access to the app by the users. Web roles generate tasks for worker roles*
- *Worker Roles do "heavy lifting" and manage data in tables/blobs*
- *Communication is through queues.*
- *The number of role instances should dynamically scale with load.*



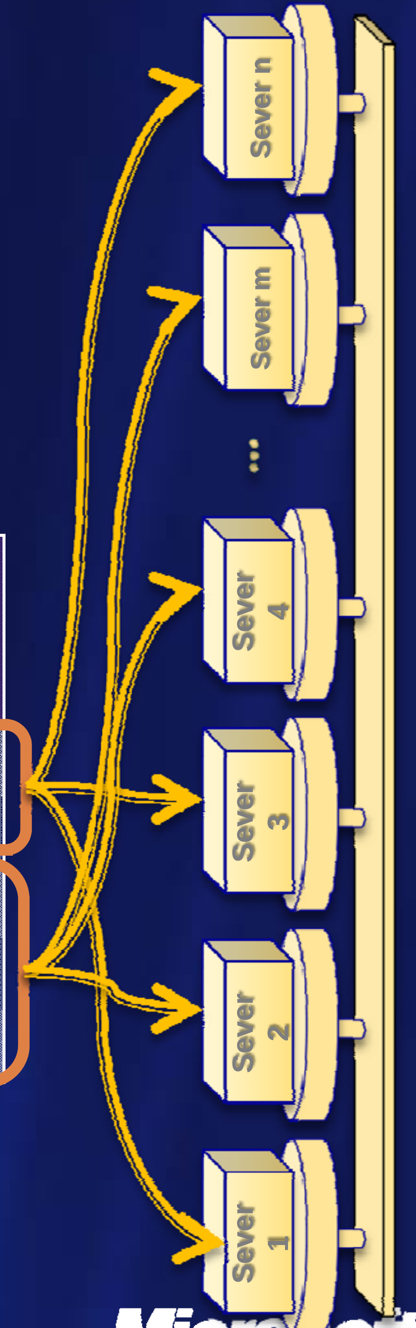



Microsoft

# Data Architecture

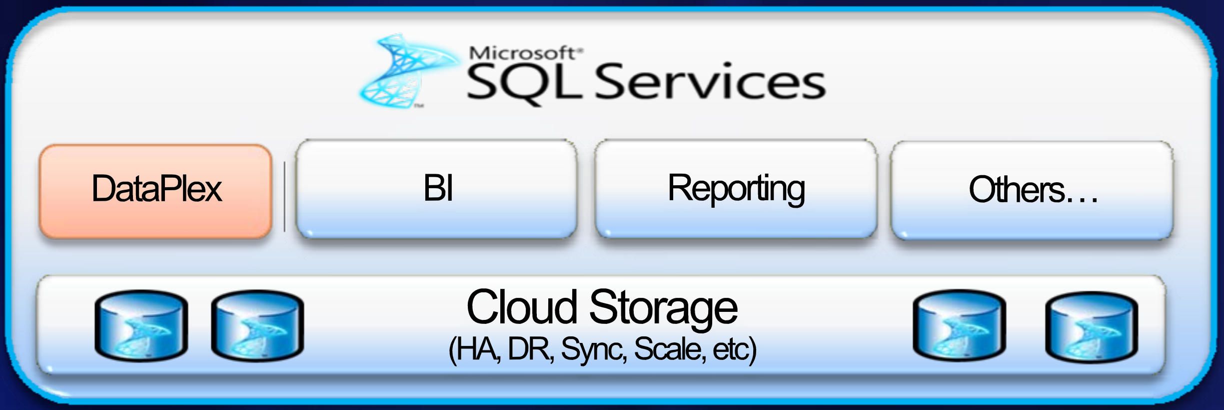
- Replicated, distributed file objects (blobs)
- Massive table storage (replicated, distributed)

| Partition Key Document Name | Row Key Version | Property 3 Modification Time | ….. | Property N Description |
|---|---|---|---|---|
| Examples Doc | V1.0 | 8/2/2007 | ….. | Committed version |
| Examples Doc | V2.0.1 | 9/28/2007 | | Alice's working version |
| FAQ Doc | V1.0 | 5/2/2007 | | Committed version |
| FAQ Doc | V1.0.1 | 7/6/2007 | | Alice's working version |
| FAQ Doc | V1.0.2 | 8/1/2007 | | Sally's working version |



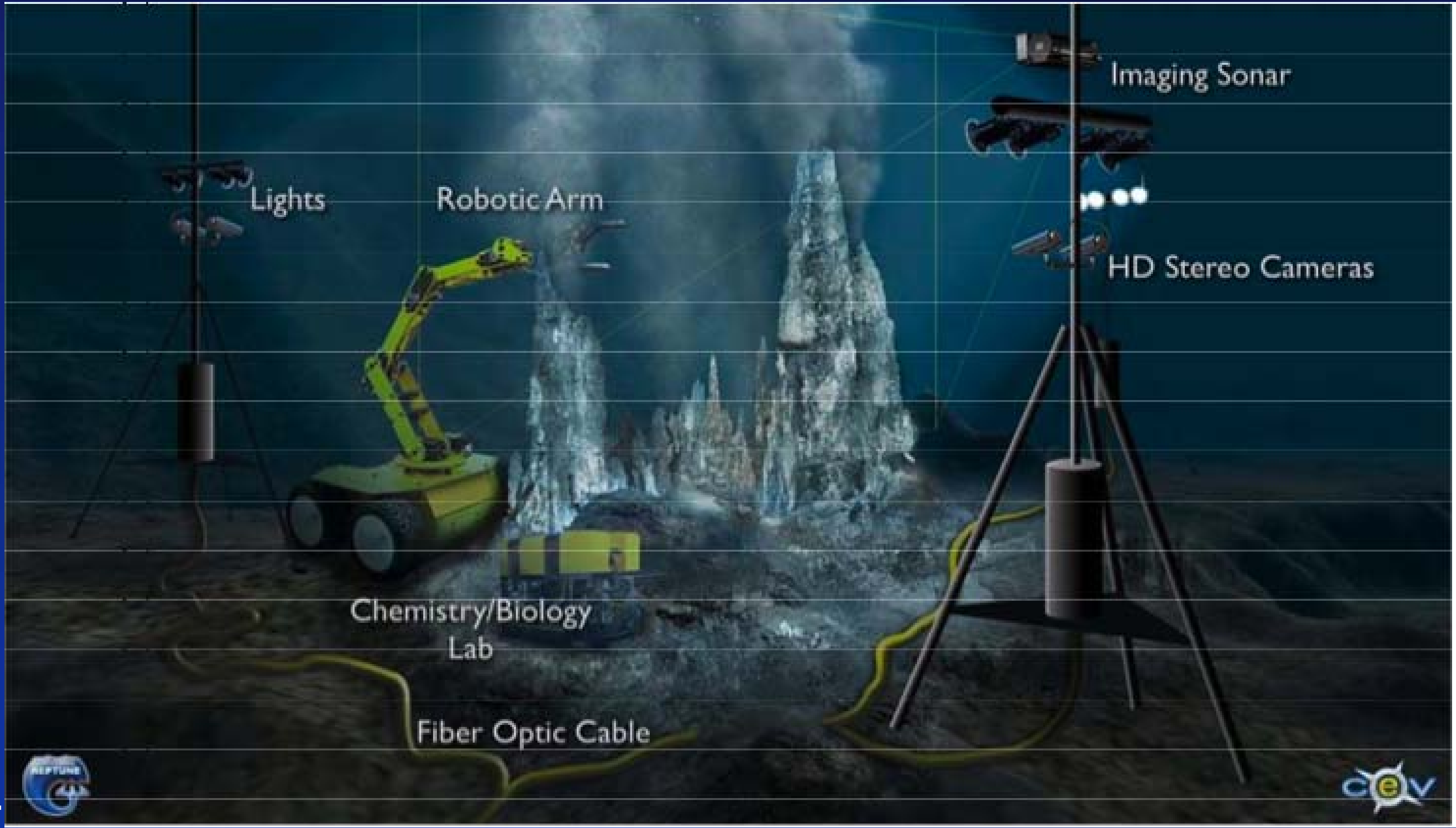

Microsoft

# Windows Azure and SQL Services




# Smart Sensors and Data Fusion

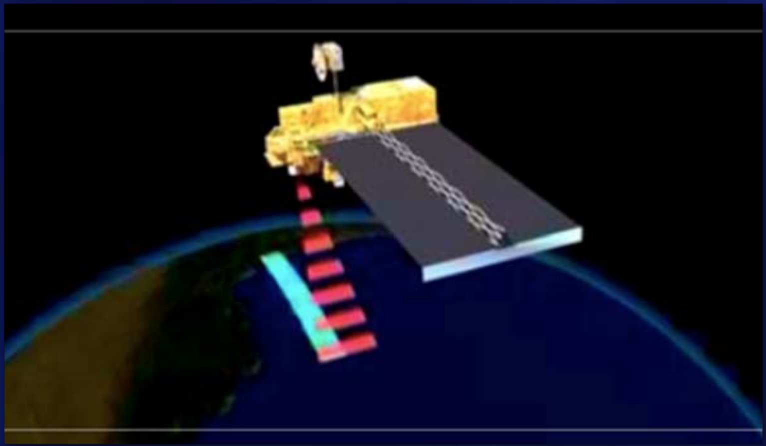
- The NSF Ocean Observing Initiative
  - Hundreds of cabled sensors and robots exploring the sea floor
  - Data to be collected, curated, mined




Conceptual representation of a future seafloor laboratory on the Regional Cabled Observatory network. *Credit: the NEPTUNE Project* www.neptune.washington.edu

# MODIS data analysis

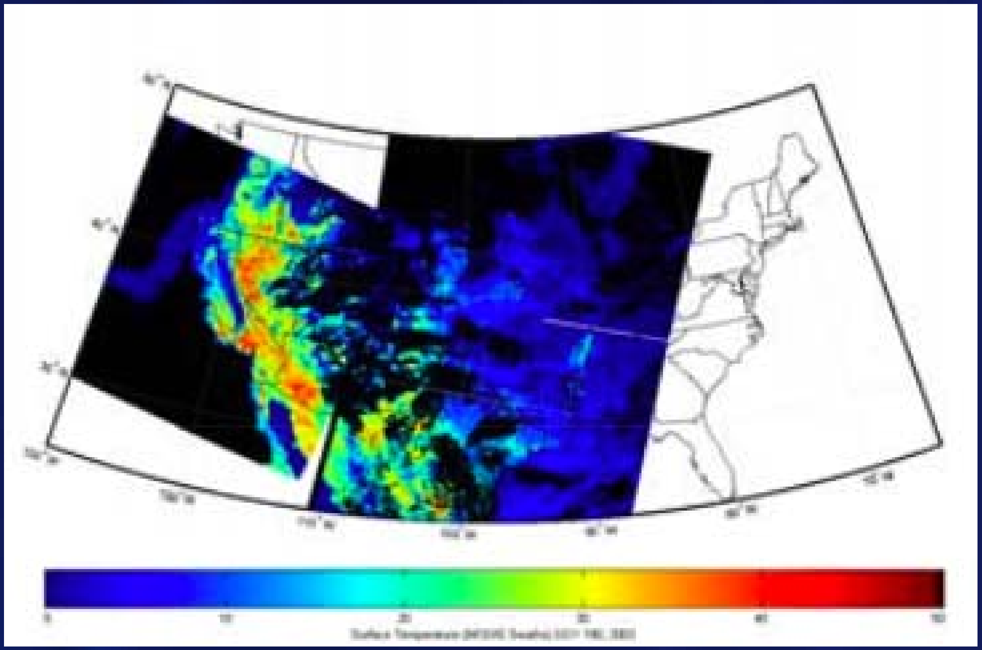
- Satellite image land use analysis
  - Two MODIS satellites
    - Terra, launched 12/1999
    - Aqua, launched 05/2002 •
    - Near polar orbits
    - Global coverage every one to two days
    - Sensitive in 36 spectral bands ranging in wavelength from 0.4 µm to 14.4 µm





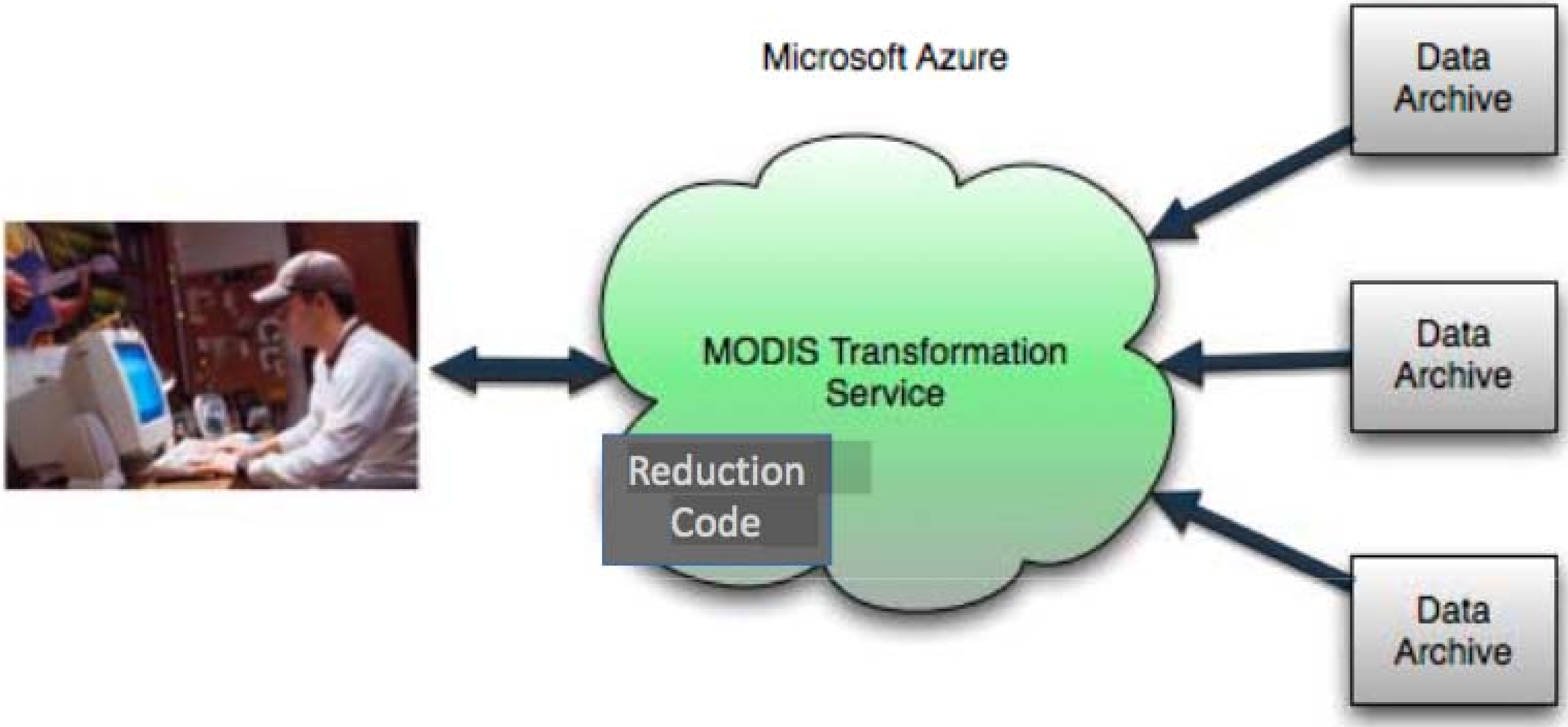
Microsoft

# Modis Cutter



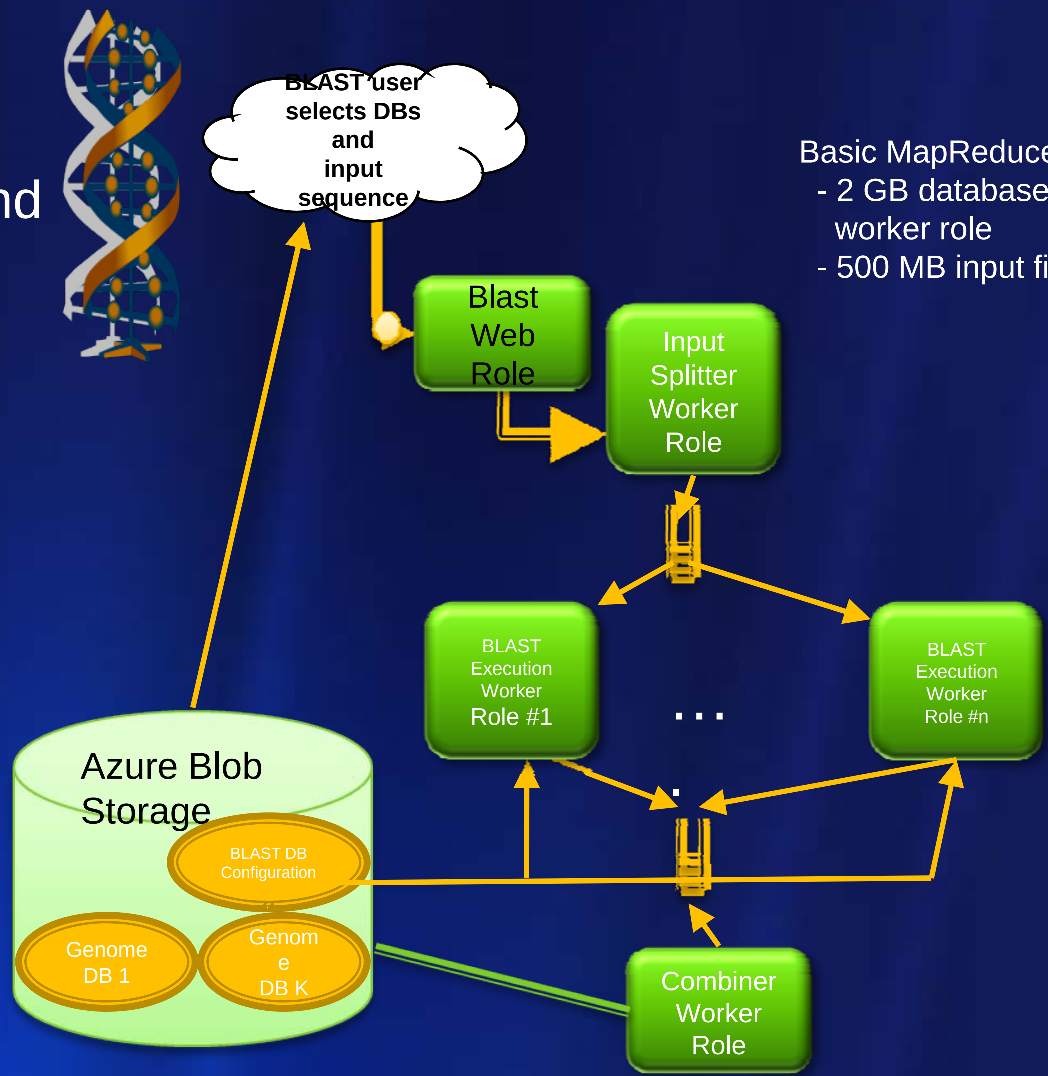
- Work by 3 Stanford student in a class project, Catherine Van Ingen, and Keith Jackson.
- Data Integration Problem
  - ~35 different science data products
  - Atmospheric and land products are in different projections
  - Need to reproject one to work with both
  - Different products are in different:
    - Spatial resolution – Temporal resolution •
  - Must integrate data from different swaths, different days
  - Data volume and processing requirements exceed desktop capacity

# Modis Cloud Image Transformation and Reduction Service



# Life Science in the Cloud

- Map Reduce-style
  - Parallel Blast
  - Take DNA samples and search for matches
- Full Metagenomics sample
  - 363,876 records
  - 50 roles 94,320 sec.
  - Speedup = 45.
  - 100 roles 45,000 sec.
  - Speedup = 94.
- Next Step
  - 1000 roles
  - 20 GB input sample



Basic MapReduce
- 2 GB database in each worker role
- 500 MB input file.

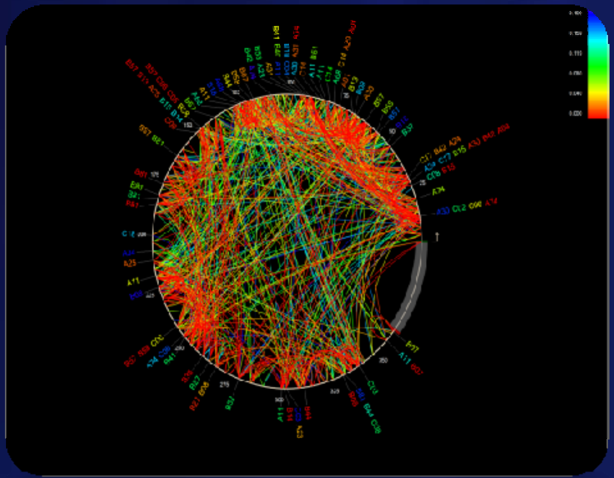




Microsoft

# PhyloD as an Azure Service

- Statistical tool used to analyze DNA of HIV from large studies of infected patients
- PhyloD was developed by Microsoft Research and has been highly impactful
- Small but important group of researchers
  - *100's of HIV and HepC researchers actively use it*
  - *1000's of research communities rely on results*



**Cover of PLoS Biology November 2008**

Typical job, 10 – 20 CPU hours, extreme jobs require 1K – 2K CPU hours

- Very CPU efficient
- Requires a large number of test runs for a given job (1 – 10M tests)
- Highly compressed data per job ( ~100 KB per job)



Microsoft

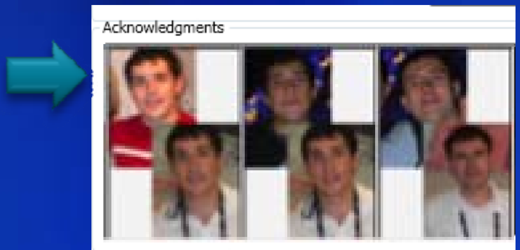
# Challenges for Science

- There is no effective Supercomputer Cloud
  - Supers are about peak performance at the expense of reliability. Batch mode operation. Also poor data access. Virtualization considered bad.
  - Clouds are about scalable, on-demand reliable access by millions of simultaneous users. Optimal for large scale data analysis. Heavy use of virtualization
- Projects like LEAD need both HPC & cloud.
  - Want to run hundreds of copies of WRF on-demand. Resource needs to scale out dynamically. Need rapid access to data streams and archival data. Complex workflows.
- Possible solution
  - Cloud servers composed of massive many-core processors – run as separate cores or ganged.

# CCF Applications

- The Goal: to identify and build applications that
  - Explore exciting future scenarios that are enabled by advanced data center architectures
  - Show deep integration of the client with the cloud
  - Demonstrate and test the Orleans programming model
- Examples

- Intelligent Memory Assistant
  - From phone to datacenter face recognition application
- Adaptive code tier splitting
  - Depending on environment Marlowe moves parts of code execution from phone to data center at runtime


- Virtually Real Worlds
  - Merge 2nd life with Photosynth and telepresence
- Scale real-time VR interaction from a few dozen simultaneous users/avatars to millions.
  - Total stress on data center network

# Conclusion

- Cloud technology transforming the service space.
  - Pay-as-you-go scalability
  - Economics favor massive commercial deployment
- There is a strong chance we will change the research model in many disciplines.
  - The clients + the cloud will be a game changer driven by the shift to data driven science.
  - Can we build the tools to manage and mine streams of data?
  - Can we merge HPC and the cloud effectively?
- The government challenges
  - Changing the mindset in the federal government to allow for grants to shift capex (buying computers) to opex (pay-as-you-go service).

# Windows Azure Data Storage Concepts