Cloud Computing

Rob Fatland

Microsoft Research

For the MRC story: http://research.microsoft.com/azure

WRF two years and counting: http://weatherservice.cloudapp.net

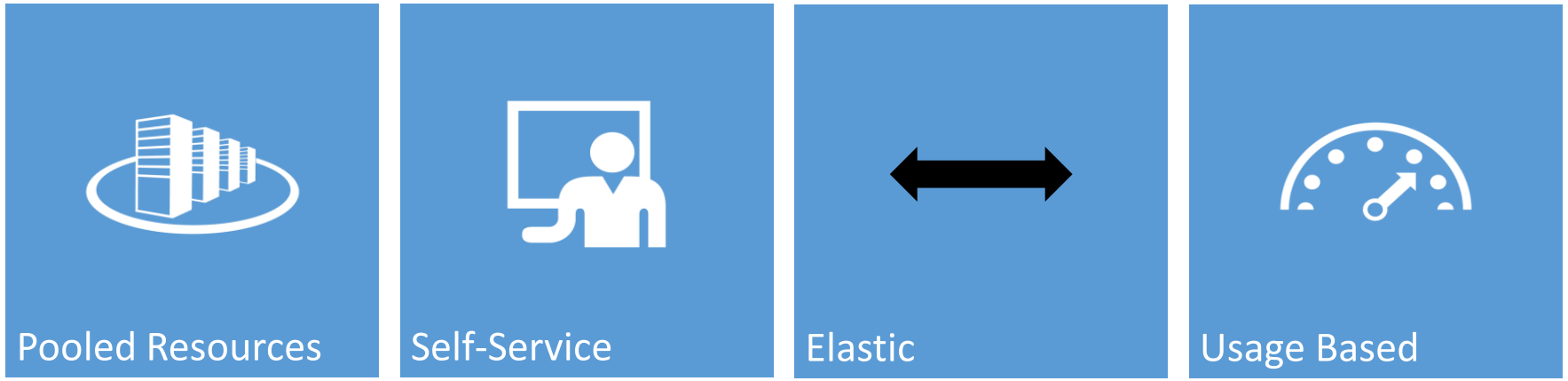
Pre-Talk Talk

- What does ESS look like as more of this CI arrives?

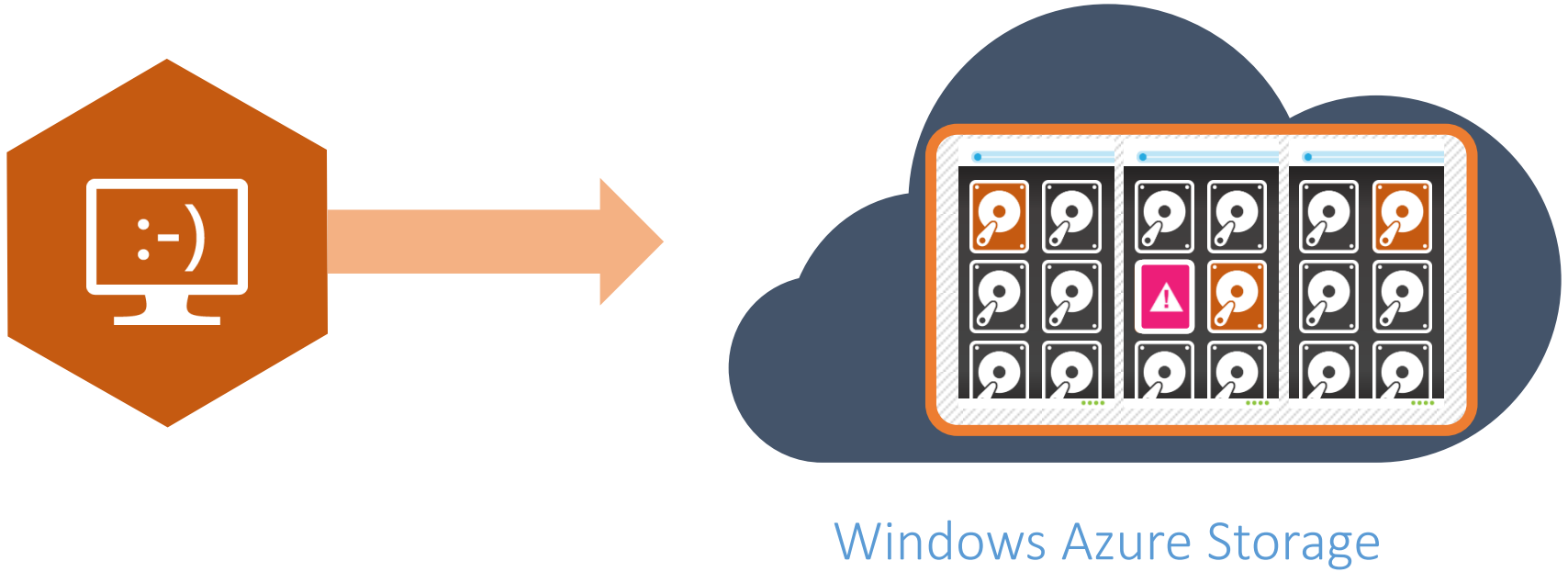
The Promise of Cloud Computing

- Solve the following problems without raising new ones
  - I don’t want to buy / fix / worry about / re-buy Servers
  - I want my legacy code to Just Work and scale
    - Modest SuperComputer On Demand at better than cost)
  - I want my Stuff to be available anywhere, always
  - …then there is something about Services… what are they and do I need them?
    - Database??
    - REST API?? Should I design one so you can talk to my Database??
    - Oh… You want me to USE an existing REST API.
      - …and so on down into the rabbit hole… let’s save this for later

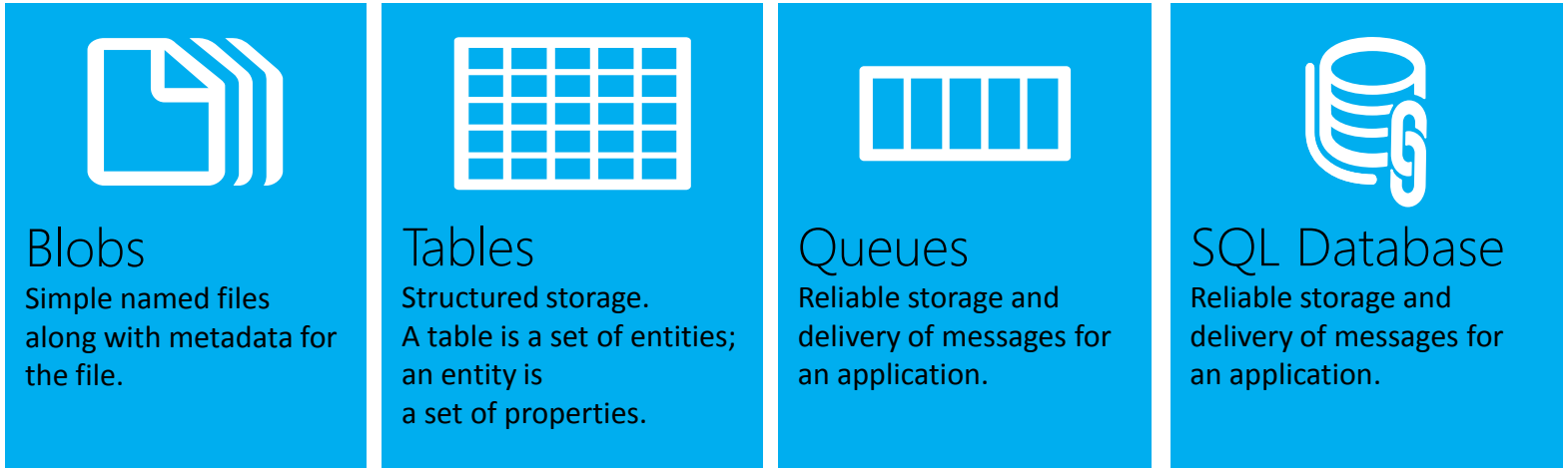
Today’s Transformation – Cloud



Persistent Virtual Machines with Geo-Replication

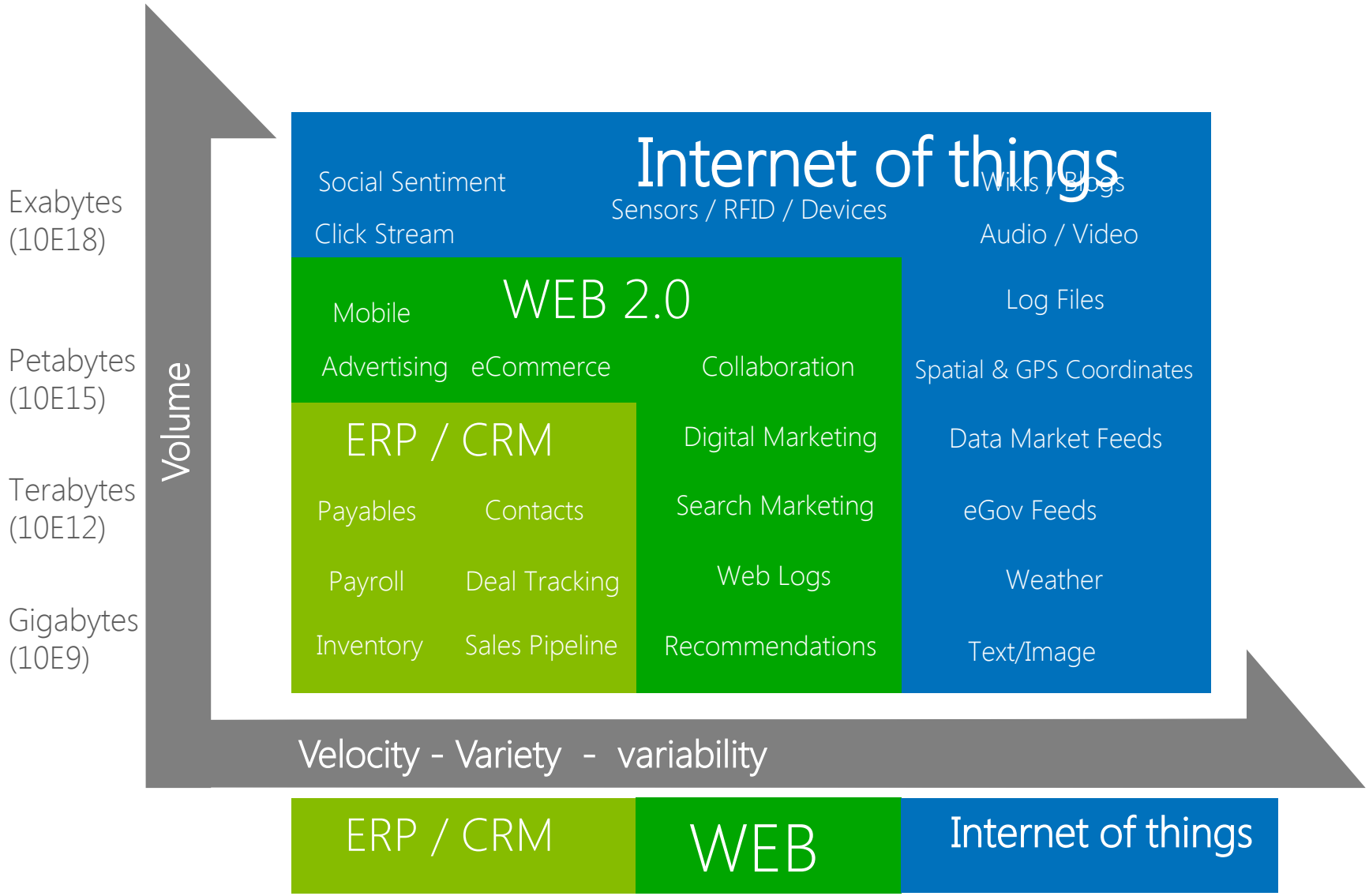


Cloud Storage

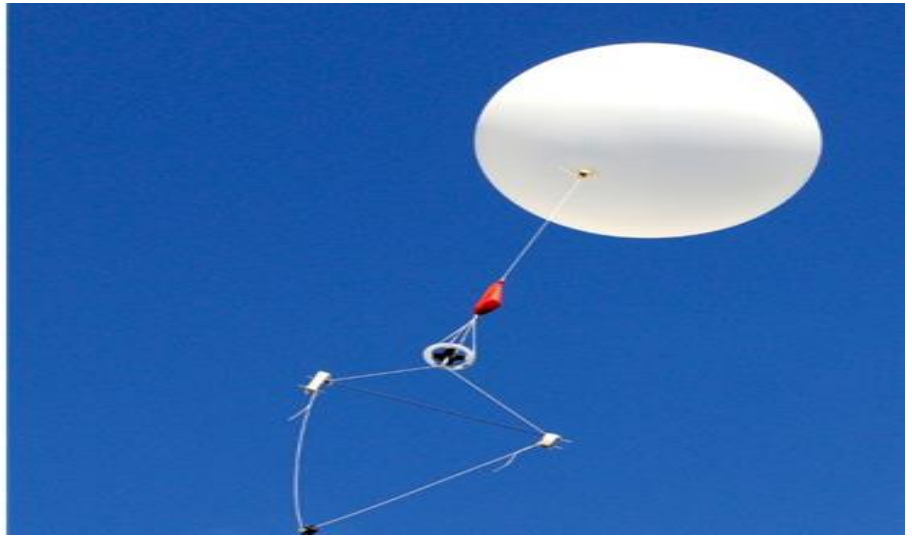


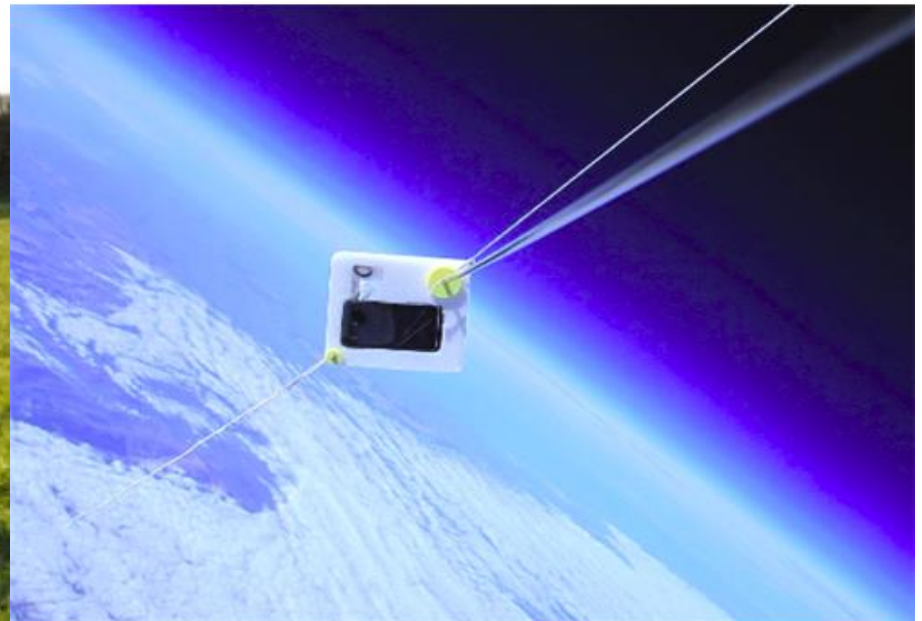
Windows Azure

Big Data


Exabytes (10E18)
Petabytes (10E15)
Terabytes (10E12)
Gigabytes (10E9)
Volume
Internet of things
Social Sentiment
Click Stream
Sensors / RFID / Devices
Wikis / Blogs
Audio / Video
WEB 2.0
Mobile
Advertising
eCommerce
Collaboration
Log Files
Spatial & GPS Coordinates
Data Market Feeds
eGov Feeds
Weather
Text/Image
ERP / CRM
Payables
Contacts
Payroll
Deal Tracking
Inventory
Sales Pipeline
Digital Marketing
Search Marketing
Web Logs
Recommendations
Velocity - Variety - variability
ERP / CRM
WEB
Internet of things
Windows Azure
Storage/GB
1980
190,000$
1990
9,000$
2000
15$
2010
0.07$

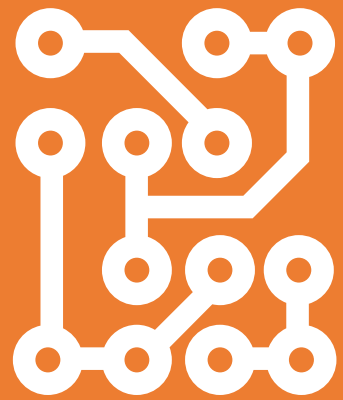
Collecting data above the cloud





Linux Offering

- Linux as a first class citizen in Azure
- Open Sourcing critical components
- Documenting API
- We will offer both Community
- and Commercial Distributions
- You will be able to buy support for the commercial distributions



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

Weather Forecast Computation as a Service

Windows Azure™

http://weatherservice.cloudapp.net

Benefits

Overview

- Acutal National Weather service modeling Code containing 15 groups of complex Tools
- Utilizes a Platform as a Service (PaaS)
- Initialization Data from NOAA
- Uses Bing Maps for visualization

Detailed Forecast

- Anyone will have the capability run a forecast
- The easiest & affordable way to deploy an HPC Cluster
- Conveniently Sharing simulation Data in the cloud

Details:

- 1500 simulations ran over 1 year
- 1 city 180km x 180km takes 7 hours on 1 Extra-large instance for a 3 day forecast
- 100 cities on 100 Extra-large instance still takes 7 hours
- http://aka.ms/oljnt2

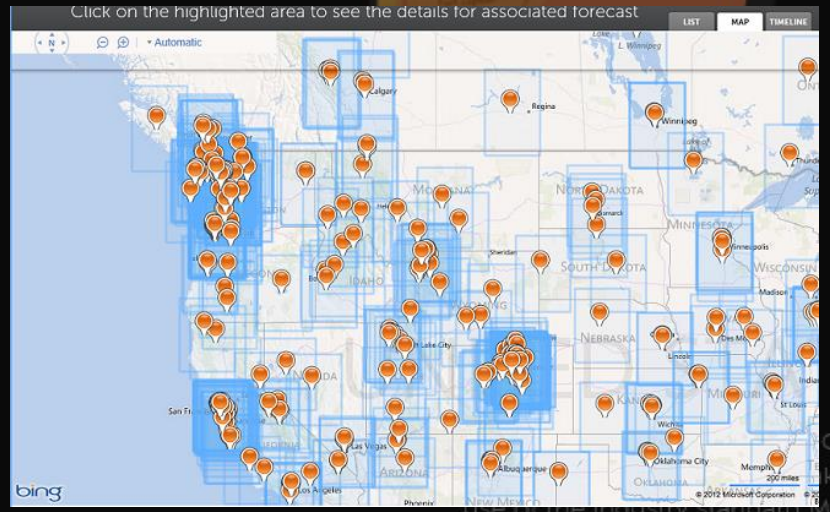
http://weatherservice.cloudapp.net/Home/Index

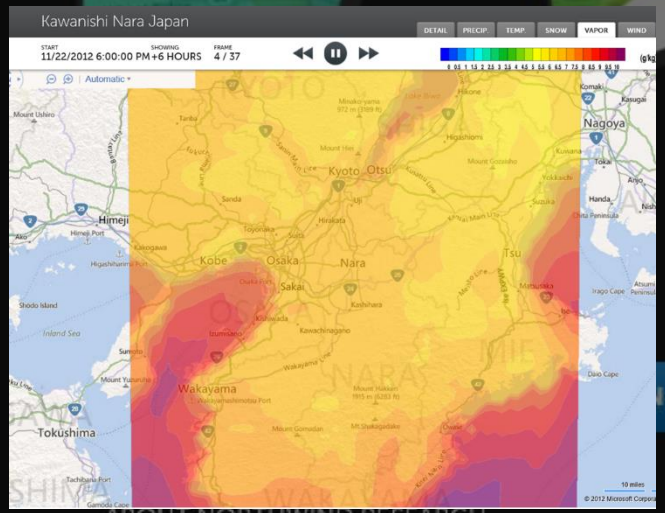
Cloud services - Windows Azure | Home - Weather Forecastin...

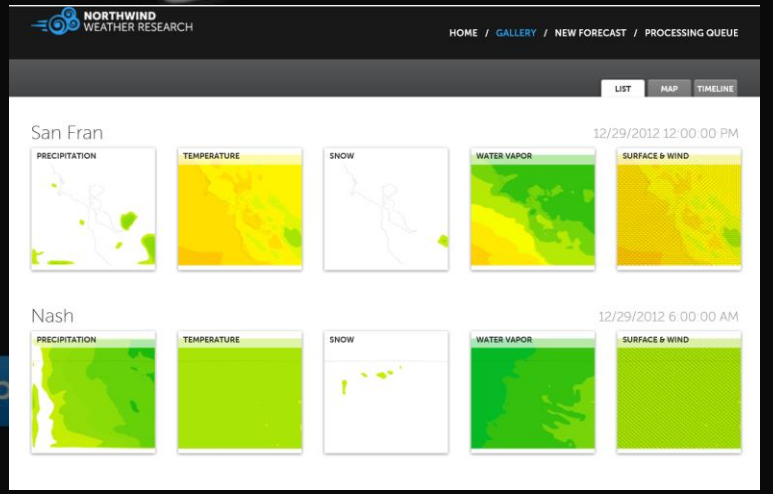
NORTHWIND WEATHER RESEARCH

HOME / GALLERY / NEW FORECAST / PROCESSING QUEUE

Custom weather forecasts, computed in the Cloud.







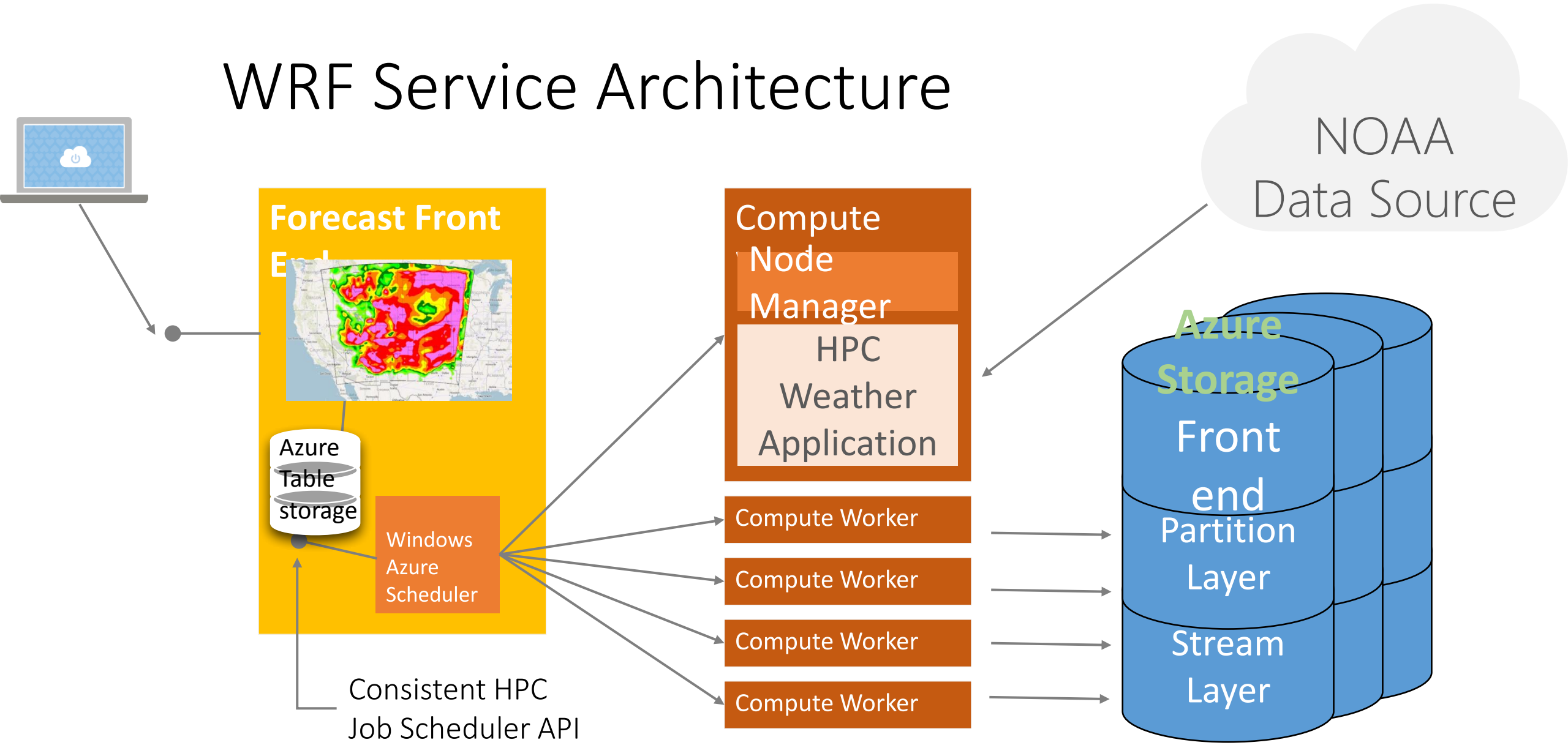
library to distribute the computation across processors, and it utilizes the new Scheduler to manage distribution of executing forecast

ABOUT NORTHWIND RESEARCH

Northwind is an advanced science and technology company. Built by ...wind combines the science of ... to bring you weather forecast on demand

http://weatherservice.cloudapp.net

WRF Service Architecture



To Summarize

- Cloud computing gets me hassle free cheap compute power, it frees me from the Server game, and gives me universal access…

- …and it offers to help with XaaS: Services and Databases and Infrastructure…

Some

- Actually the “my code just works” part is coming along pretty well
  - MATLAB in Azure, for example, is moving along (Dozier)
- People can easily exaggerate compute power in Cloud scenarios
  - This isn’t the latest CRAY; it’s more like having 50 PCs tied together by an Internet cable. You still don’t get massive parallelism for free: Communication is still a hurdle.
- Cloud service providers will be a marketplace driven by supply and demand; so they are heavily motivated to price competitively against “buy your own servers”. Without that they got no business.
- The ecosystem that grows up around this already features One Click Installs: Wikis, Word Press, Virtual Machines
  - For example Azure gives you Linux VMs

Part 2: Services, Databases and Cyberinfrastructure

- I need 50 years of climate data
- Ok: Get 50 nodes, each one gets a year
- Node 51 scoops them all up for you.

Code to System

- I send my code to run in your corner of the cloud to operate on data there
- Result << data

- Post-Talk Talk
  - Idea: Perhaps democratization of ESS research is possible
  - Idea: self-organizing taxonomy of research problems
  - Idea: Intentional new approach to trans-disciplinary + CS graduate students

End of Talk

Workshop Goal

The overall goal of the workshop is to identify similarities and differences in how cyberinfrastructure programs serve polar sciences versus other disciplines. The workshop and the report will address engagement and connections between computer and polar sciences concerning what can be accomplished in the short-term (1-5 years) and long-term (5-10+ years). The outcomes of this workshop will inform the Polar Cyberinfrastructure Program at the National Science Foundation concerning the past and current polar cyberinfrastrucutre activities and will provide support for a community-driven design and architecture development of a polar science cyberinfrastructure that is aligned with the following end-users' needs: (1) long-term sustainable curatorship, standardization, management and discovery of data and metadata; visualization, manipulation, and analysis; (2) use of high performance computing (HPC) for direct and sustainable advances in current polar research; (3) big data and data access; (4) interoperability with data from other domains; (5) e-learning and educational tools based on cyberinfrastructure components; and (6) virtual organizations.