

Multicore, HPC and Clouds: It's All About Convergence

Dan Reed
reed@microsoft.com

Corporate Vice President
Extreme Computing Group (XCG)

Microsoft[®]

Our Decadal Research Changes

● Commodity clusters

- Proliferation of inexpensive hardware
 - “Attack of the Killer Micros”
- Race for MachoFLOPS
- Low level programming challenges



● Rise of data

- Scientific instruments and surveys
- Storage, management and provenance
- Data fusion and analysis

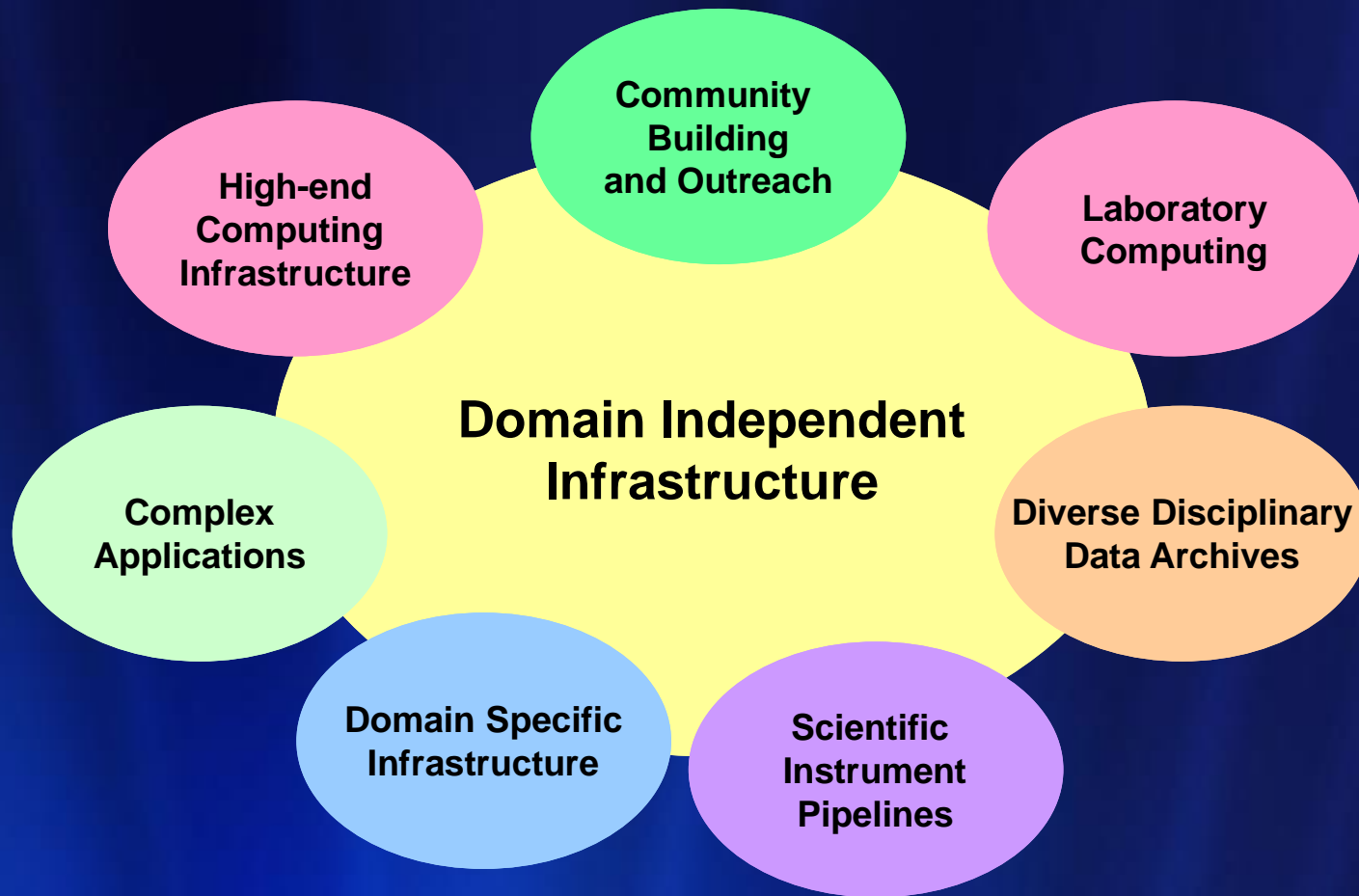


● Distributed services

- Multidisciplinary collaborations
- Interoperability and scalability
- Multi-organizational social engineering

Microsoft[®]

Cyberinfrastructure Components

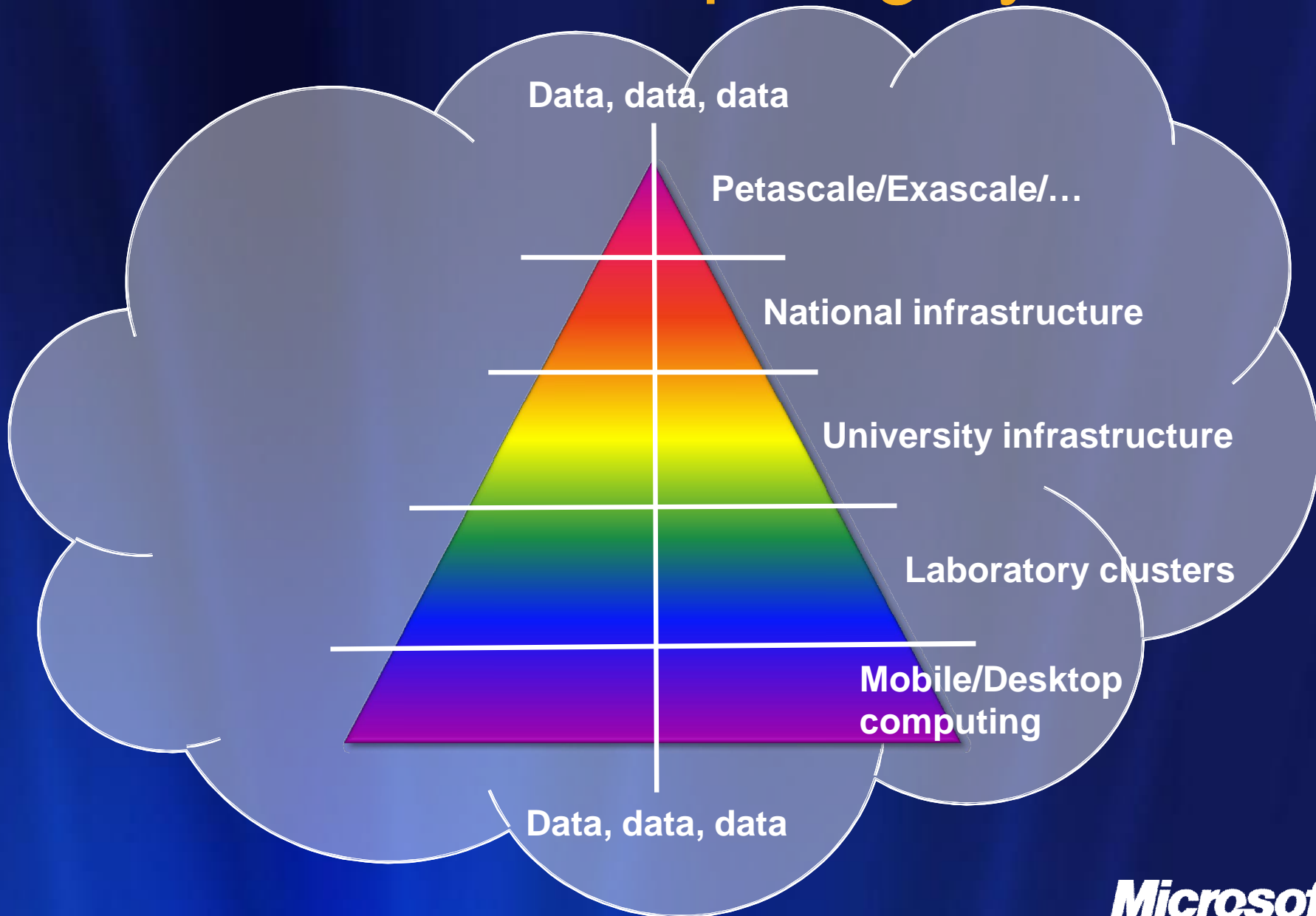


Research Infrastructure Challenges

- Insatiable demand
 - Cycles, storage, software, support
- Distributed acquisition/deployment
 - Sometimes, duplicative, non-shared infrastructure
- Distributed cost structures
 - Power, space, staff, staff, hardware
- Long-term sustainability
 - Decades rather than months/years
- The shape of the triangle
 - Apex versus mainstream users



The Technical Computing Pyramid



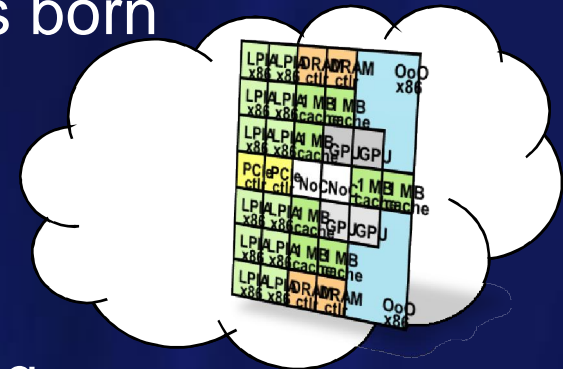
Today's Truisms (2009)



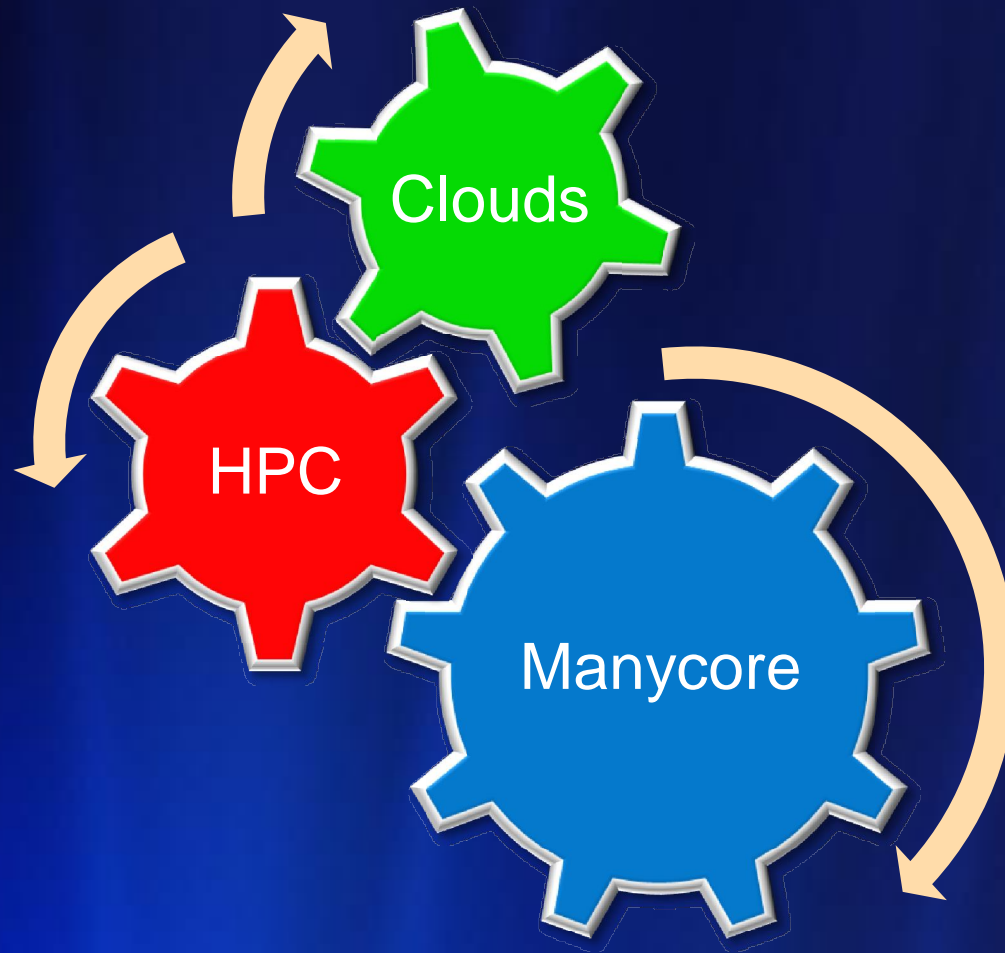
- Bulk computing is almost free
 - ... but applications and power are not
- Inexpensive sensors are ubiquitous
 - ... but data fusion remains difficult
- Moving lots of data is {still} hard
 - ... because we're missing trans-terabit/second networks
- People are really expensive!
 - ... and robust software remains extremely labor intensive
- Application challenges are increasingly complex
 - ... and social engineering is not our forte
- Our political/technical approaches must change
 - ... or we risk solving irrelevant problems

The Pull of Economics ...

- Moore's "Law" favored consumer commodities
 - Economics drove enormous improvements
 - Specialized processors and mainframes faltered
 - The commodity software industry was born
- Today's economics
 - Manycore processors/accelerators
 - Software as a service/cloud computing
 - Multidisciplinary data analysis and fusion
- They will drive change in technical computing
 - Just as did "killer micros" and inexpensive clusters

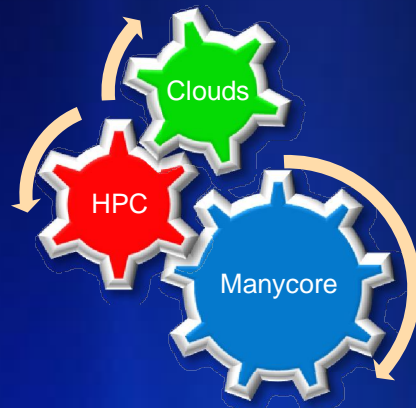


Convergence: It's Happening Now



Convergence Acceleration

- Very few users love technology itself
 - Clusters and parallel programming
 - Distributed services, grids or clouds
 - Data models and databases
- Successful technologies are *invisible*
 - They enable but are unobtrusive



- Desktop/mobile acceleration
 - Seamlessly accessible
 - Standard metaphors/tools

Client+Cloud

Computing
Everywhere



Continuous
Networking



Increasing
Abstraction
and Invisibility



New Data
Models



Natural User
Interface



Modeling
Replaces
Coding

Science In The Clouds



Emerging Cloud Opportunities

Ensure Service Continuity in the Cloud

- Protect against data loss and unauthorized access
- Address failure and disaster scenarios

Scale On-Demand and Cost Effectively

- Scale data throughput and store capacity
- Fuse and analyze multidisciplinary data

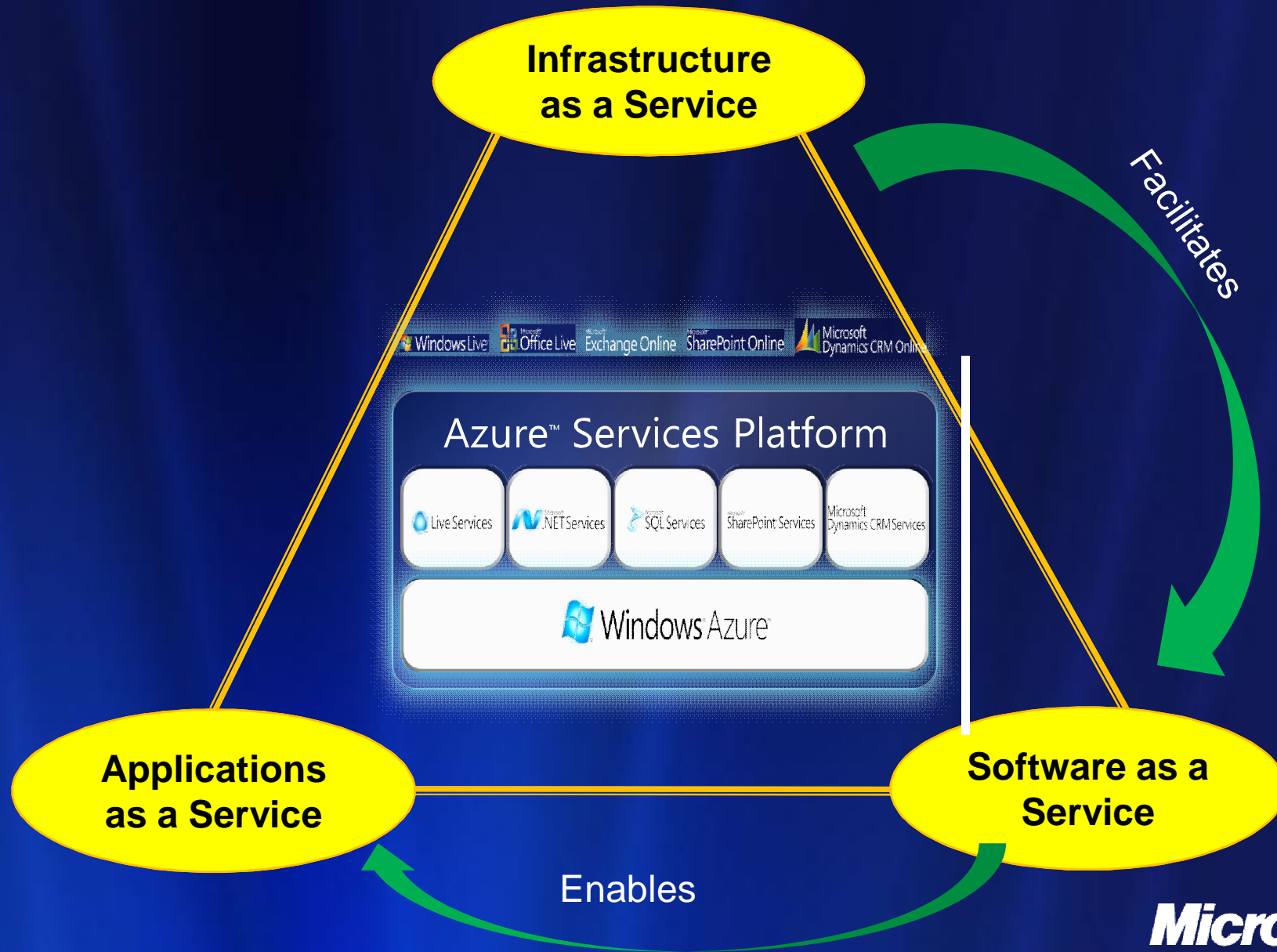
Support Emerging Applications Rapidly

- Enable rapid development of new applications/services
- Easy access to consume multiple data sources

Reduce Infrastructure and Management Costs

- Hardware and software independence
- Lower operational cost of managing data

Cloud Application Frameworks



Windows Live™

Microsoft Office Live

Microsoft Exchange Online

Microsoft SharePoint Online

Microsoft Dynamics CRM Online

Azure™ Services Platform

Live Services

Microsoft .NET Services

Microsoft SQL Services

Microsoft SharePoint Services

Microsoft Dynamics CRM Services



Windows Azure™

Compute

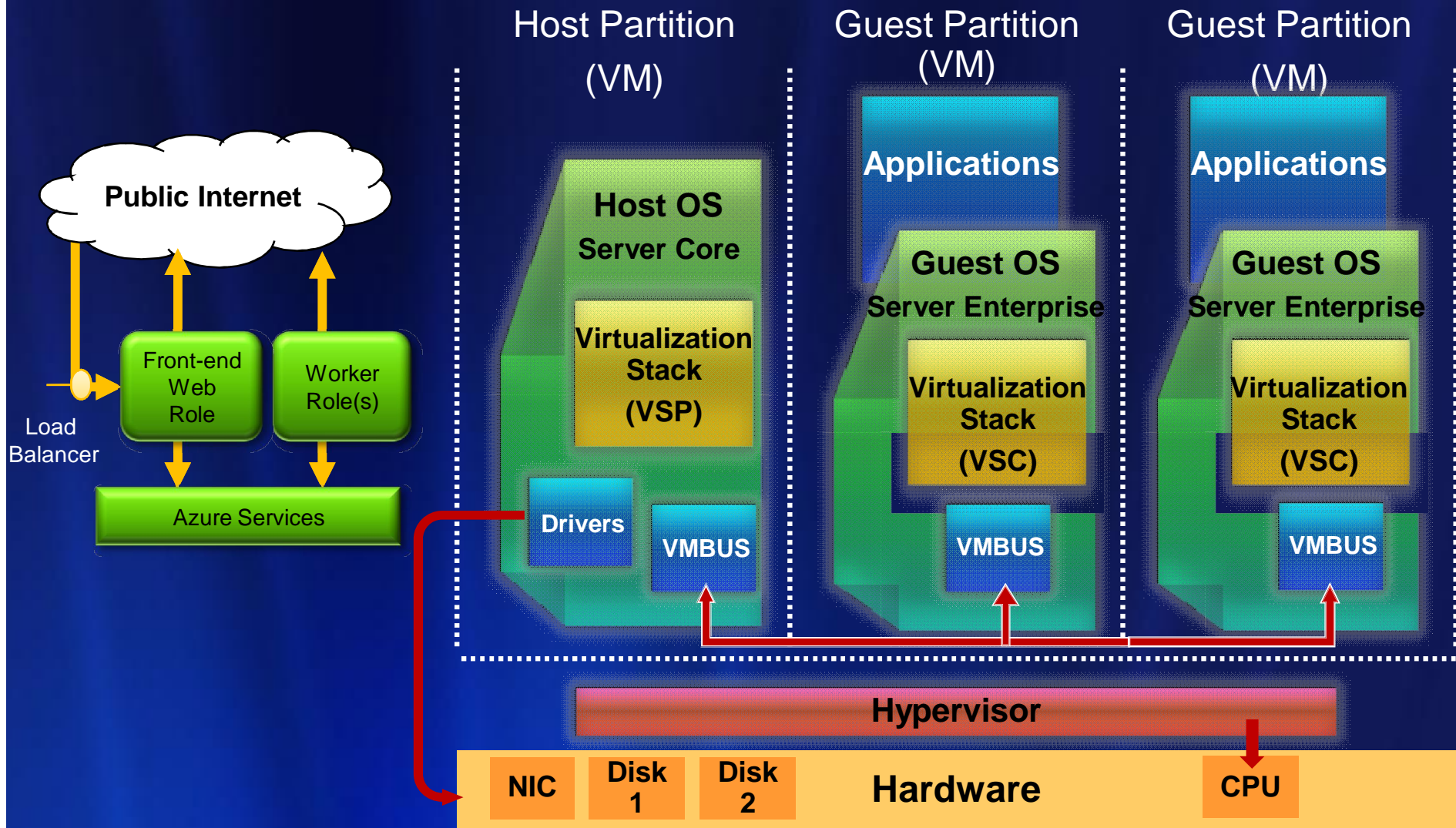
Storage

Management

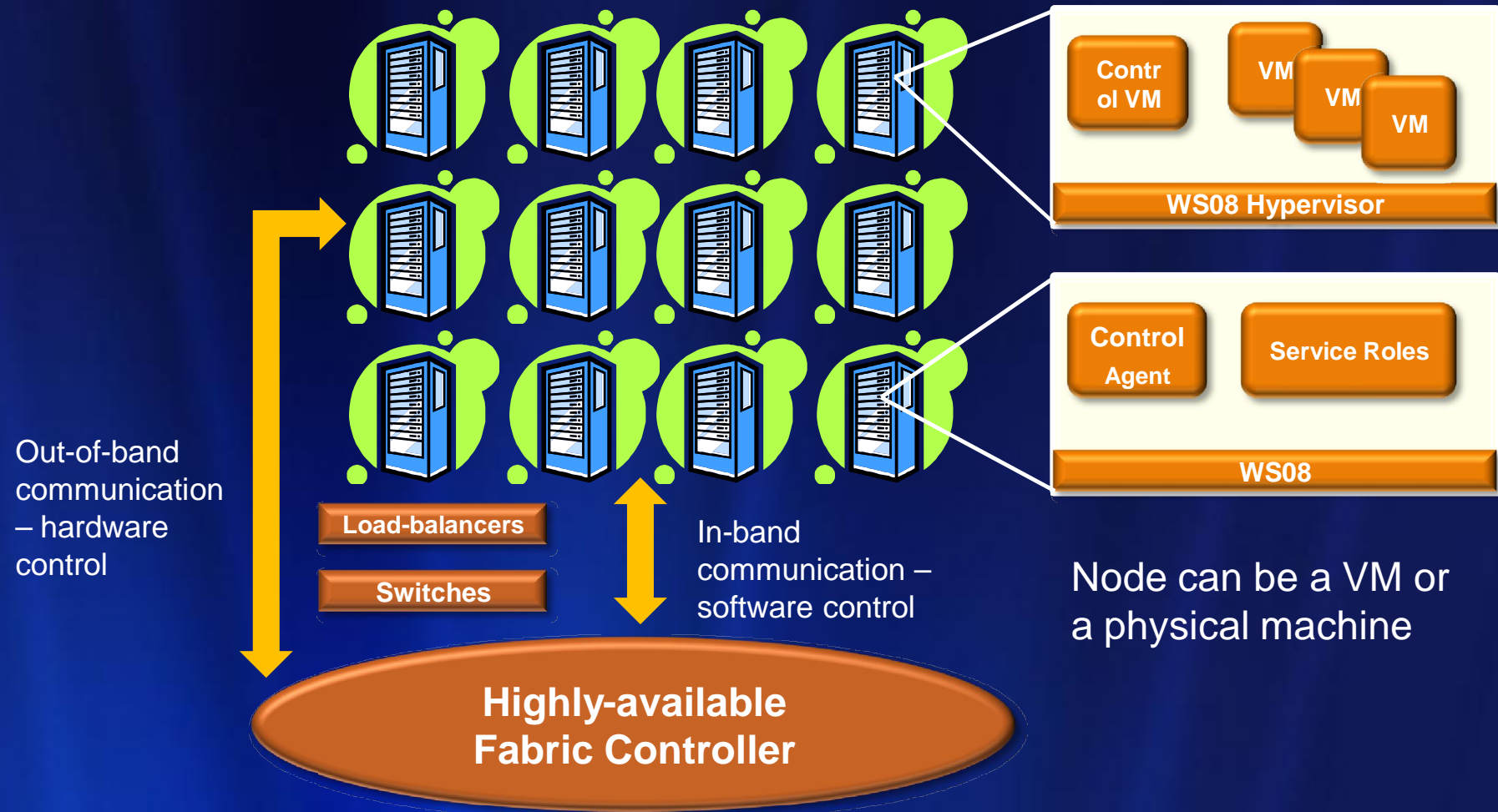
www.azure.com

Microsoft®

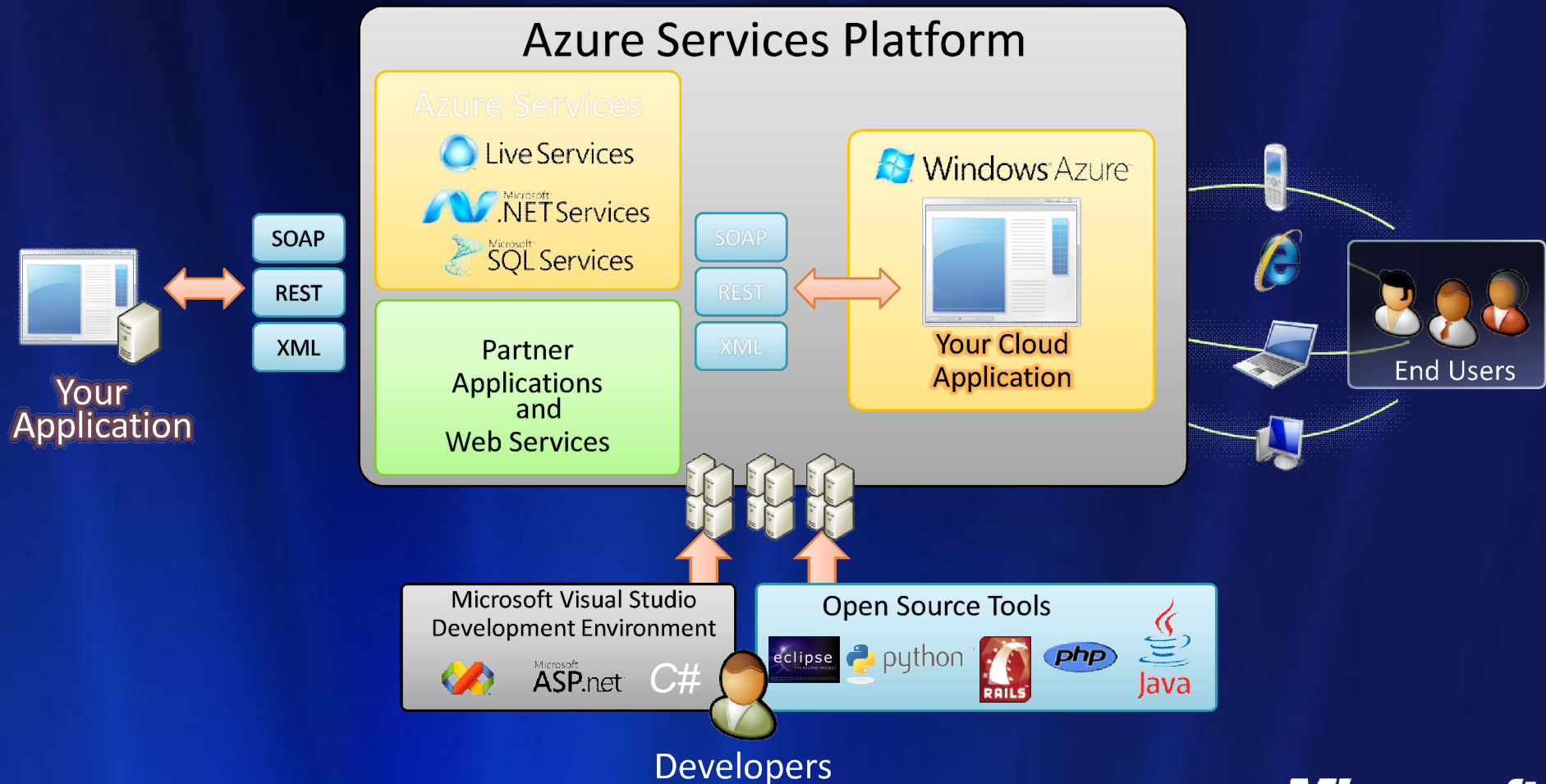
Azure Virtualization Architecture



Windows Azure Fabric Controller



Development Interoperability

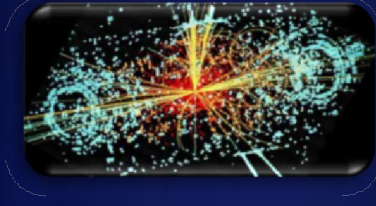


The Data Explosion

Experiments



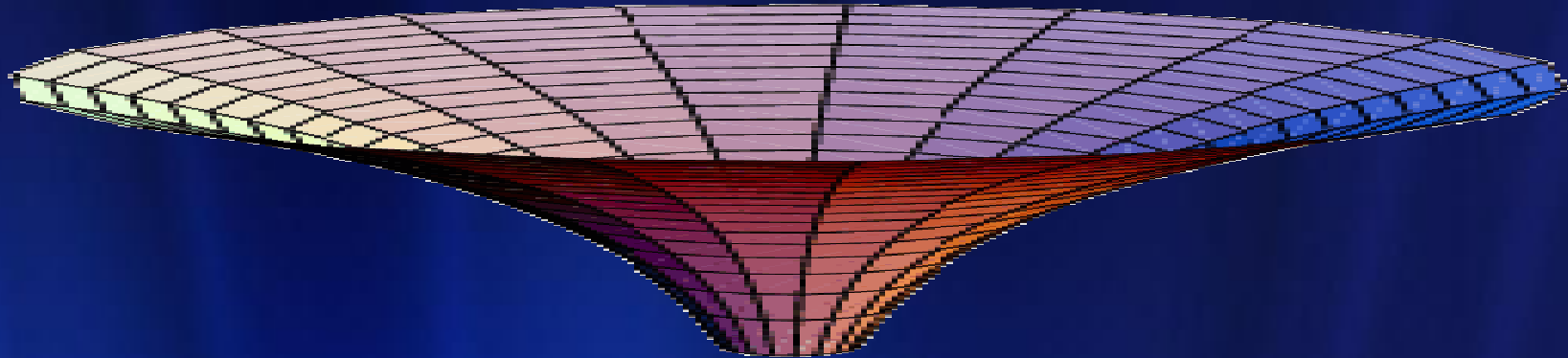
Simulations



Archives



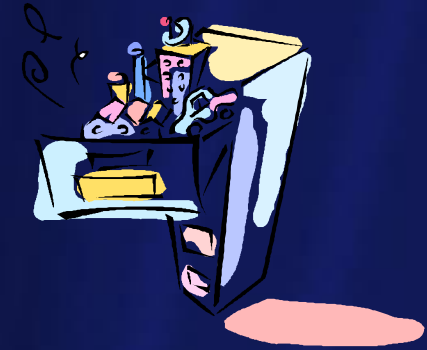
Literature



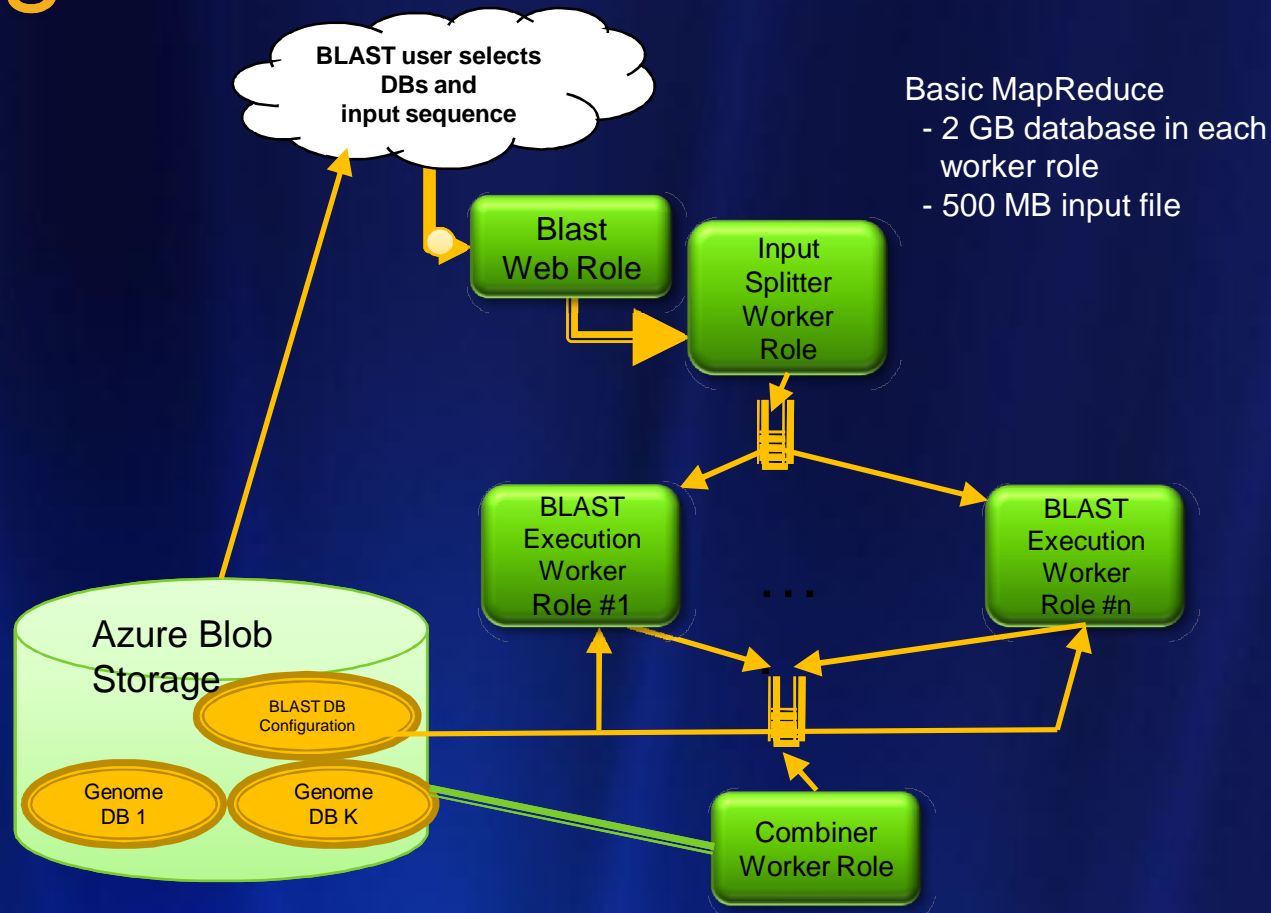
Many Petabytes
Doubling every
2 years

Social Implications of the Data Deluge

- Hypothesis-driven
 - “I have an idea, let me verify it.”
- Exploratory
 - “What correlations can I glean from everyone’s data?”
- Different tools and techniques
 - Exploratory analysis relies on deep data mining
 - supervised and unsupervised learning
 - “grep” is not a data mining tool
 - ... but an RDBMS really isn't either
- **Massive, multidisciplinary data**
 - **Rising rapidly and at unprecedented scale**



Metagenomics On Azure



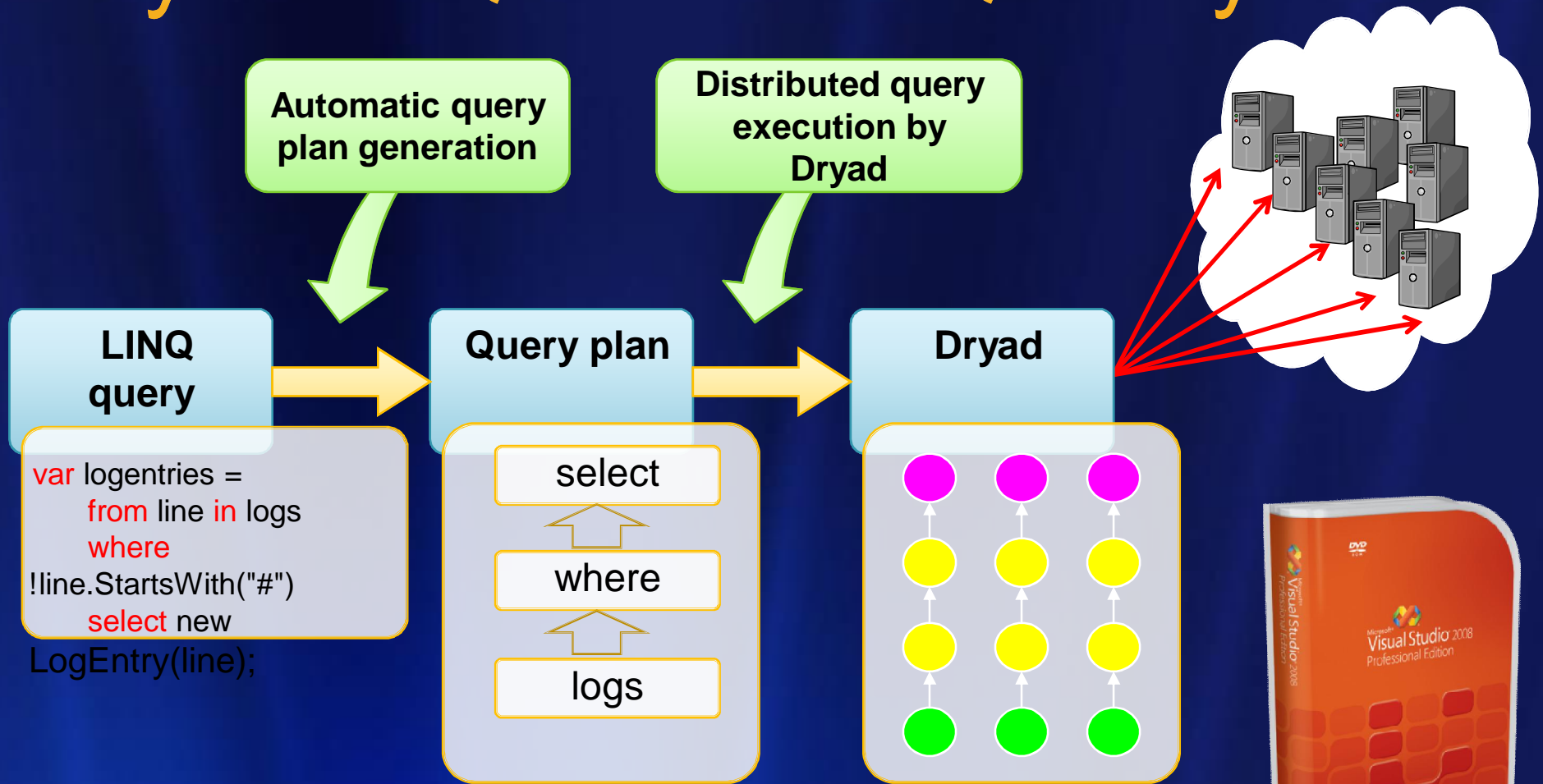
- Metagenomics sample

- 50 roles, speedup 45
- 100 roles, speedup 94

- Map reduce-style

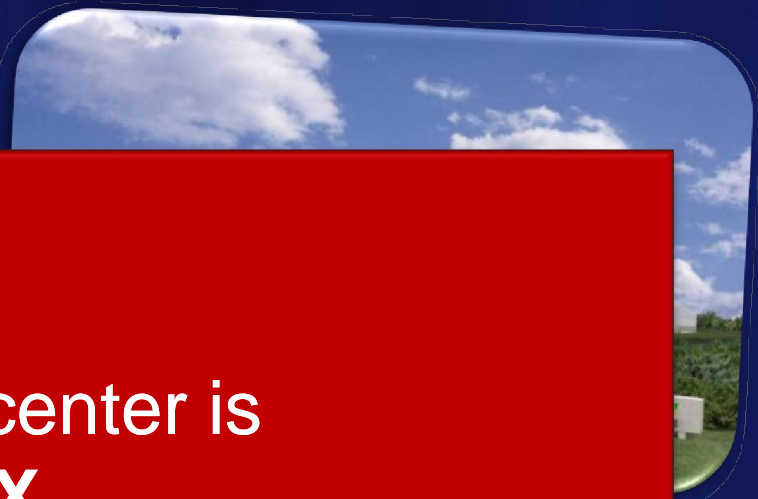
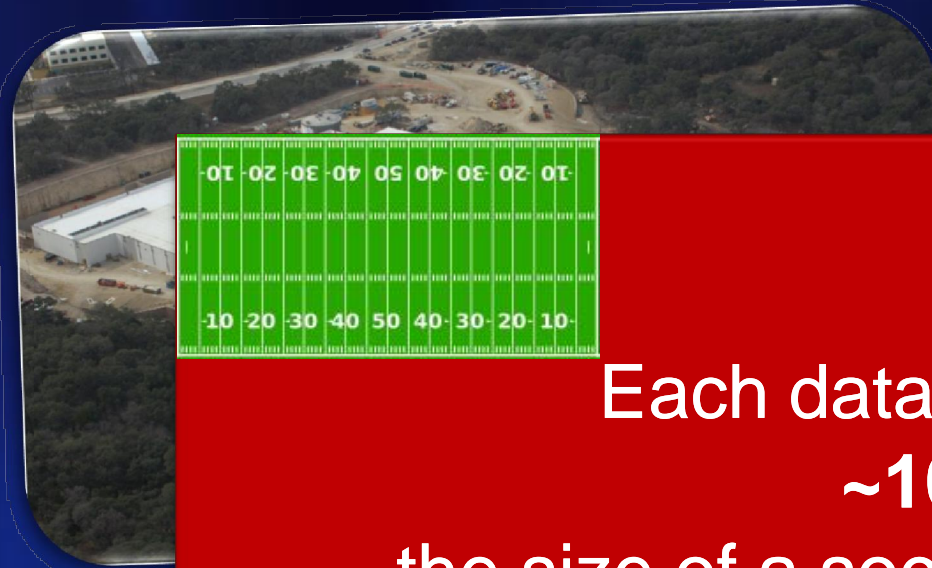
- Parallel BLAST
- DNA matching

DryadLINQ: From LINQ to Dryad

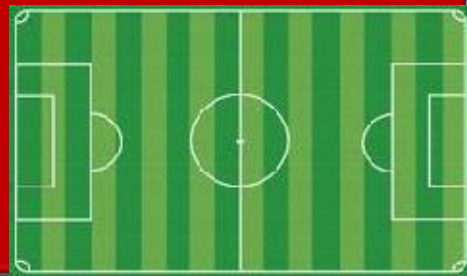


- LINQ: .NET Language Integrated Query
 - Declarative SQL-like programming with C# and Visual Studio
 - Easy expression of data parallelism
 - Elegant and unified data model

A Bit of Scale Calibration ...



Each data center is
~10X
the size of a soccer/football field

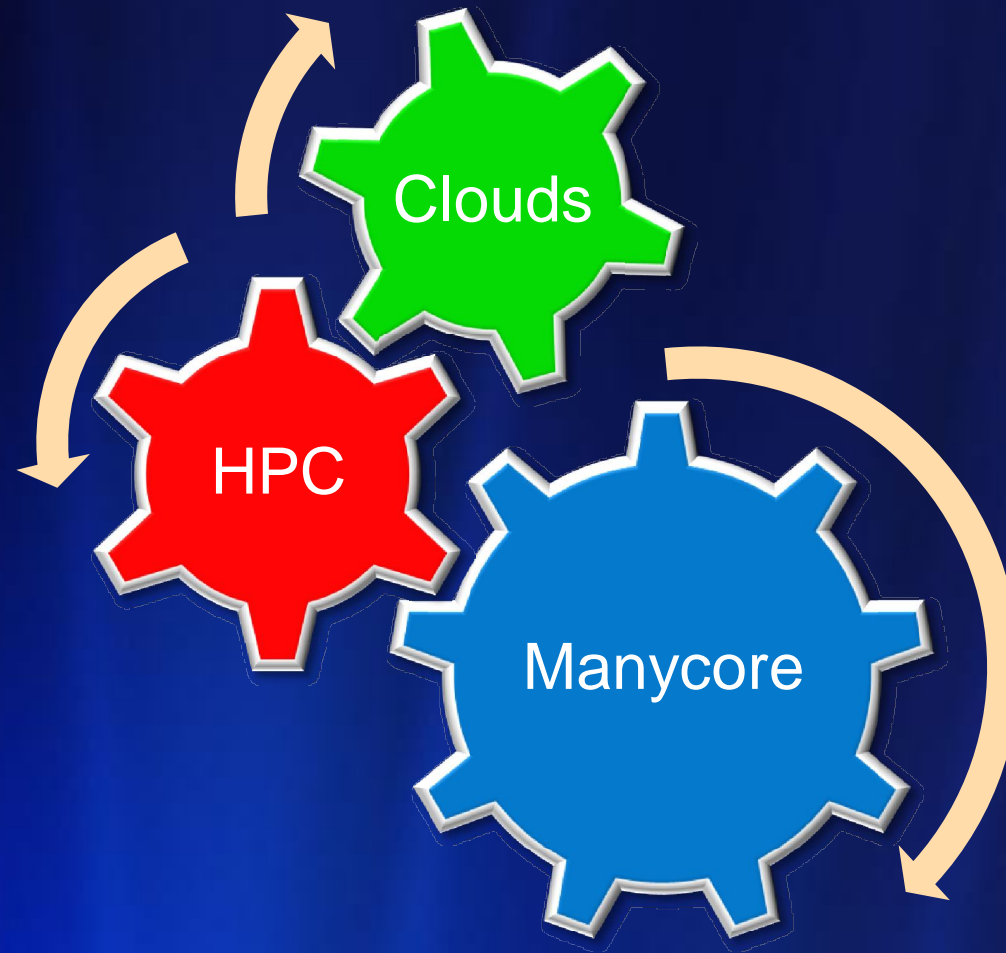


Clouds/Exascale Technical Issues

- Cooling technologies
 - **Operating points**, heat dissipation, ...
- New packaging technologies
 - Optoelectronics, **memory stacking**, ...
- New storage models/algorithms
 - **Solid state** storage
- Locality-aware algorithms
 - The speed of light is pretty **slow**
- Programming models
 - Effective **scale-invariant** abstractions
- Intelligent power management
 - Adaptation and **power down**
- System adaptation and integration
 - **Reliability** and power as first class objects



Convergence: It's Happening Now



Microsoft[®]

Your potential. Our passion.[™]

© 2009 Microsoft Corporation. All rights reserved. Microsoft, Windows, Windows Vista and other product names are or may be registered trademarks and/or trademarks in the U.S. and/or other countries. The information herein is for informational purposes only and represents the current view of Microsoft Corporation as of the date of this presentation. Because Microsoft must respond to changing market conditions, it should not be interpreted to be a commitment on the part of Microsoft, and Microsoft cannot guarantee the accuracy of any information provided after the date of this presentation.
MICROSOFT MAKES NO WARRANTIES, EXPRESS, IMPLIED OR STATUTORY, AS TO THE INFORMATION IN THIS PRESENTATION.

Microsoft[®]