Big Data Meets HPC:
Getting better acquainted

Dan Reed
Vice President for Research and Economic Development
University Computational Science and Bioinformatics Chair
Computer Science, Electrical Engineering & Computer Engineering, and Medicine

dan-reed@uiowa.edu

www.hpcdan.org

The University of Iowa

Quantum mechanics spacetime gravity

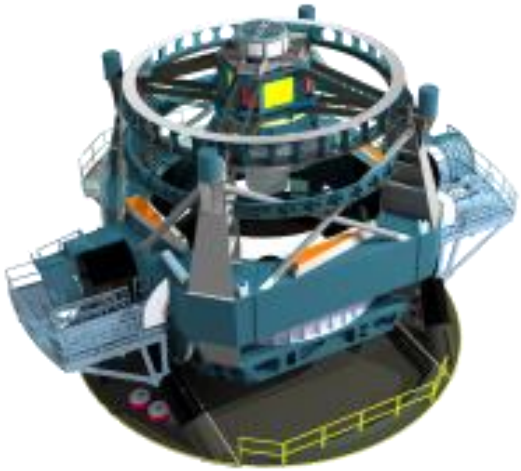
$$W = \int_{k<\Lambda} [Dg][DA][D\psi][D\Phi] \exp\left\{ i \int d^4x\sqrt{-g}\left[\frac{m_p^2}{2} R \right.\right.$$

$$\left.\left. -\frac{1}{4} F^a_{\mu\nu} F^{a\mu\nu} + i\bar{\psi}^i \gamma^\mu D_\mu \psi^i + \left(\bar{\psi}^i_L V_{ij} \Phi \psi^j_R + \text{h.c.} \right) - |D_\mu \Phi|^2 - V(\Phi) \right] \right\}$$

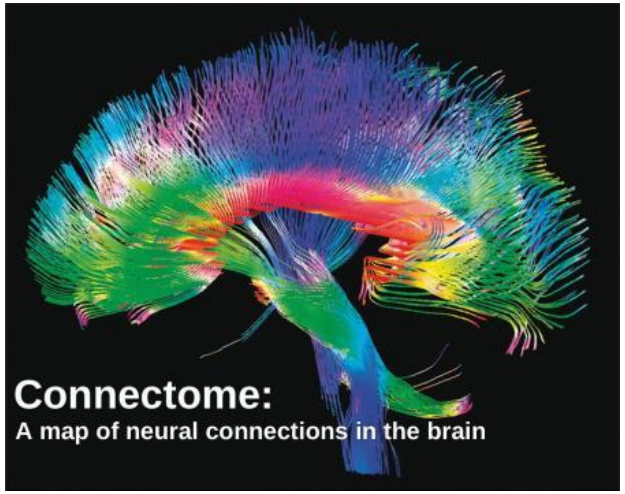
Other forces matter Higgs

Human Genome Functional Elements

http://www.preposterousuniverse.com



Large Synoptic Survey Telescope



Brain Connectome

Ocean Observing System

Square Kilometer Array

Dot.com boom 2.0

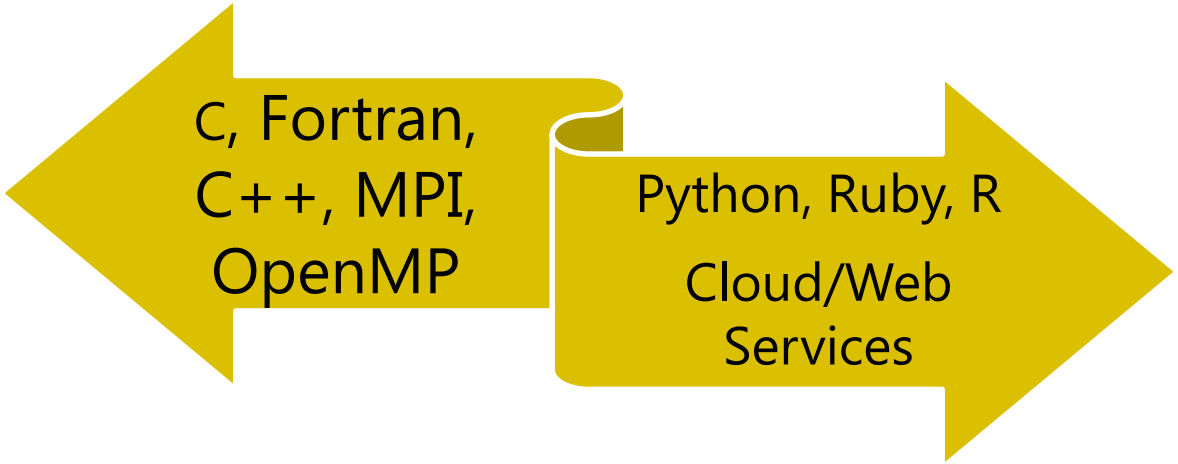
Diverging cultures

Technical application complexity is rising

- Multidisciplinary fusion
- Temporal and spatial adaptation
- Data assimilation and processing

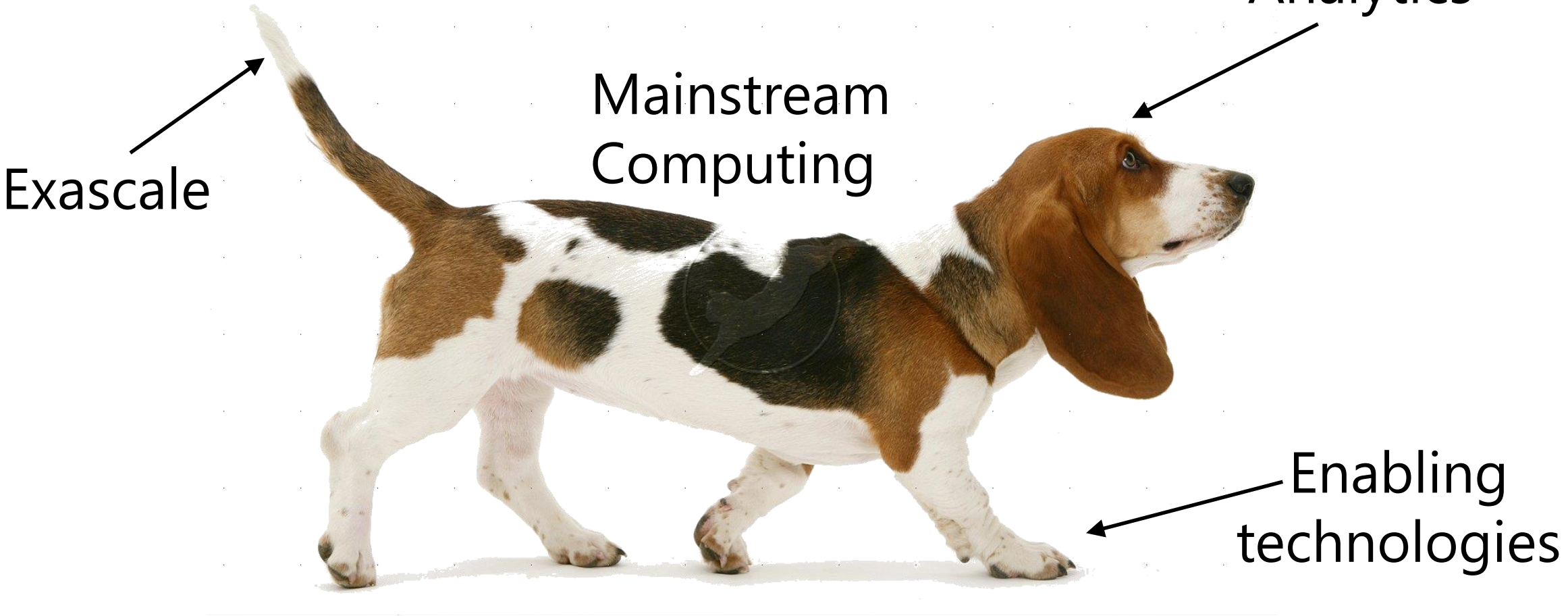
… along with multiple optimization axes

- Massive parallelism with heterogeneous cores
- Resilience/reliability at large scale
- Energy optimization for utility



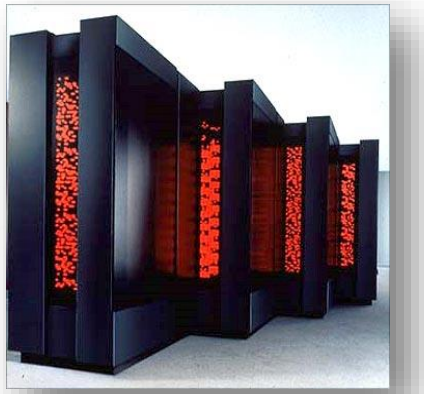
Technical and mainstream software development have diverged

Follow the money and the users …



… or the money and the users may not follow you

Computing transitions:
Riding the economic wave



Google

Windows Azure

amazon web services™

Mainframes | **Vectors** (1980s) | **MPPs** (1990s) | **Clusters & Grids** (2000s) | Clouds, Big Data and Devices

Relational DBMS | **Object-oriented DBMS** | **"No SQL" DBMS**

NVIDIA

The talent in data analytics have shifted from science to companies. We can't compete.

Astronomy researcher

The University of Iowa

Exaflops and big data: twins separated at birth

Mostly similar technology issues

- Node and system architectures
- Communication fabrics
- Storage systems and analytics
- Physical plant and operations
- Reliability and resilience

With substantial differences

- SAN/local storage models
- Virtualization and scheduling strategies
- Software development tools
- Culture and expectations

Cloud computing observations

Microsoft Supercharges Bing Search With Programmable Chips

BY ROBERT MCMILLAN 06.16.14 | 6:30 AM | PERMALINK

Programming efficiency
- Rich toolkits and expression

Network optimization
- Flatter networks
- Software virtualization and flow

Supply chain optimization
- The advantage of scale

Generic server design
- Workload-specific optimization
  - Functional accelerators
- ODM, not OEM partnerships

Energy optimization
- Substations and generation
- Switchgear control

Systemic resilience
- Failure management, not avoidance

Have you ever …

Requested 200 nodes and 2 PB for *four years*?

Logged onto a node and killed processes just to see what would happen?

Wished you could load containers rather than just applications?

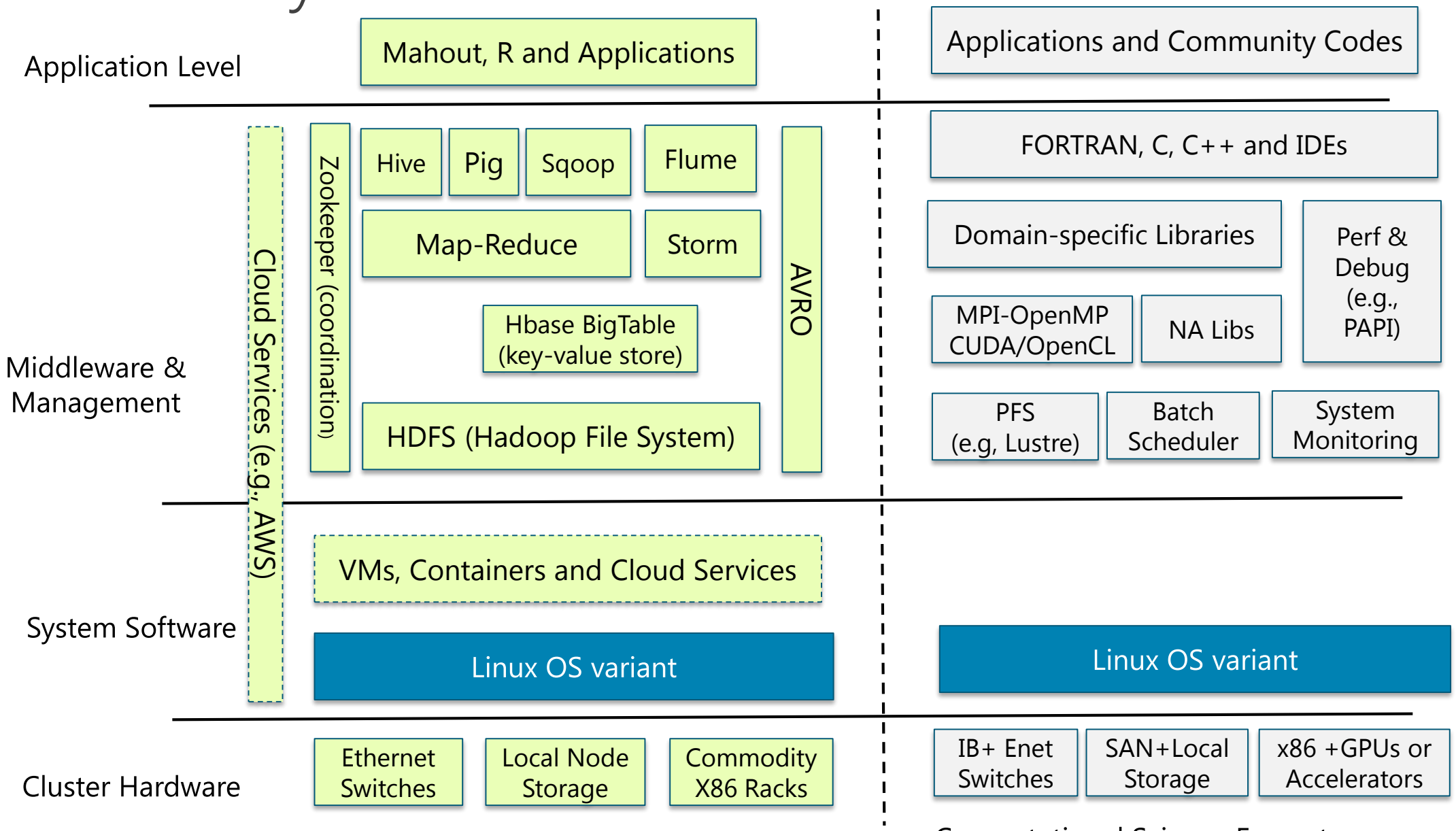
Found your code performance limited by the I/O bandwidth of a Raspberry Pi?

Thought SAN was just a typo in a message meant for Sam?

Asked your system for recommendations?

Wondered why R came after S and C doesn't matter?

Two ecosystems



File systems, SDNs and CDNs

SAN vs. node-local temporary data

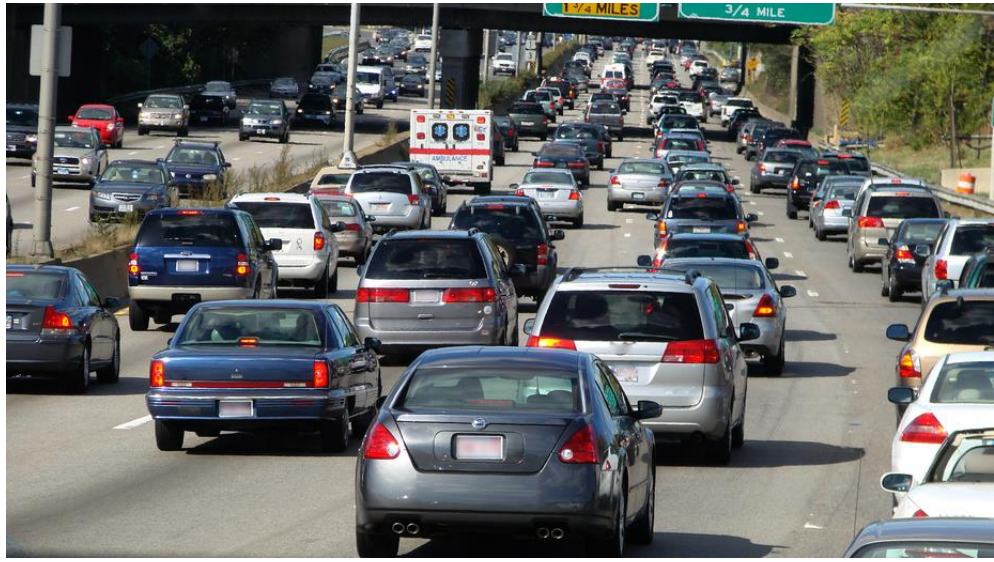
- Local disk store is a cache of SAN data subsets
- Local disks store all data

Software-defined networks (SDNs)

- Separating control and data planes
- Traffic management and prioritization

Content distribution networks (CDNs)

- Data caching and quality of service (QoS)



Resource allocation and policies

User allocations

- Normalized SUs is a univariate metric
- Jobs, particularly data intensive ones, use diverse resources

Software models

- Use a community code
- Bring your own application
- Bring your own data
- Bring your own data and software stack

Job scheduling

- Sensor data analysis has soft real-time constraints
- Run “forever” versus aperiodic intervals

Netflix Simian Army

Chaos Monkey
- *Random service termination to ensure other services continue operation*

Latency Monkey
- *Simulates service degradation and ensures services react*

Janitor Monkey
- Searches for and turns off unused resources

Conformity Monkey
- Ensures virtual machines meet specified standards

Doctor Monkey
- Monitors the “health” of various virtual machines

Security Monkey
- Monitors and analyzes system security

Blurring the system/application boundary

Virtual machines (VMs)

- Application isolation via operating system duplication
- Hypervisor implementation – hardware virtualization
- Lower limit on viability based on operating system needs

Containers

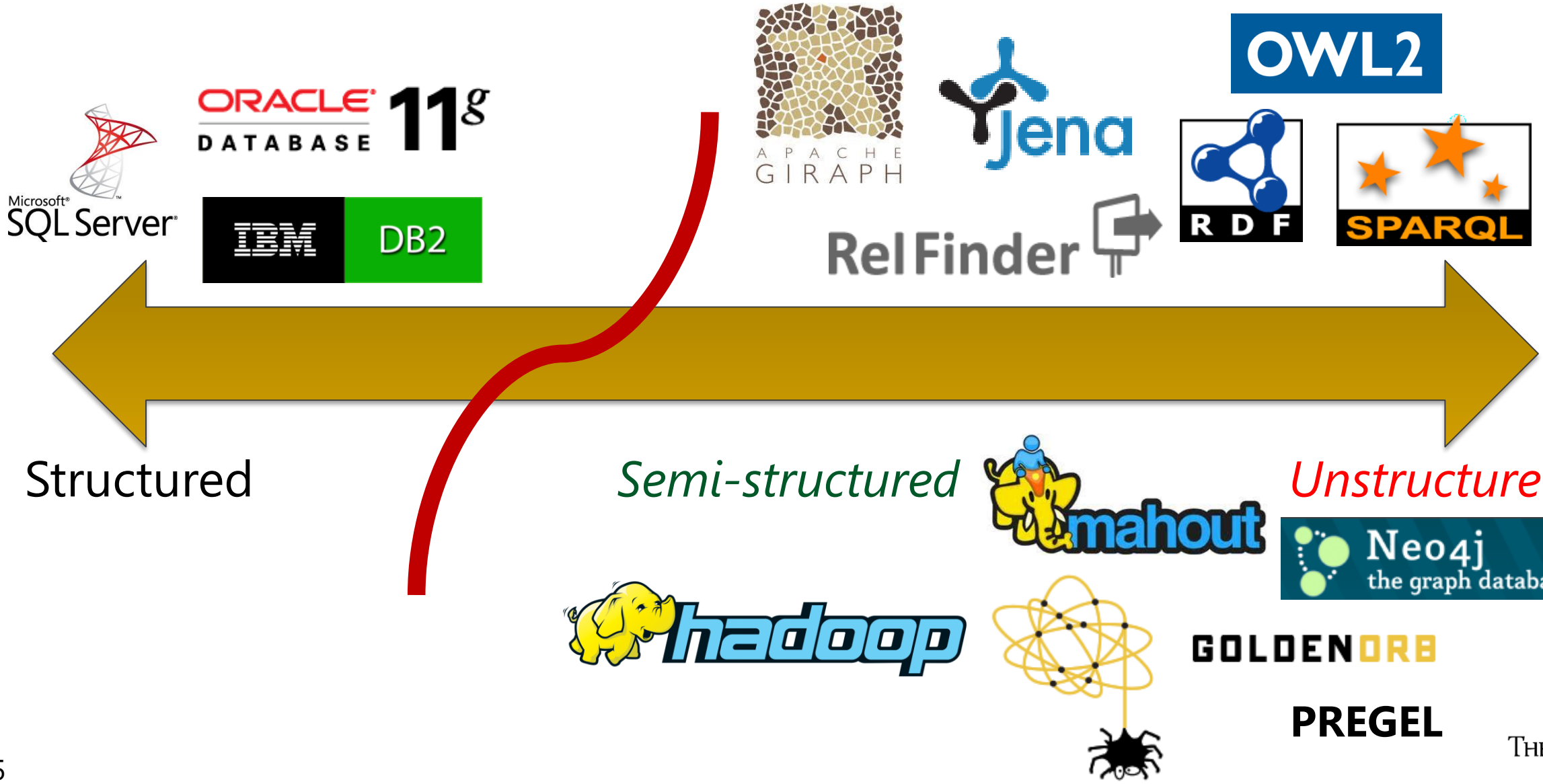
- Shared operating system
- (Potentially) separate libraries for each application
- Operating system virtualization
- Support via LXC and Docker
  - Container IDs for system calls and resource use
- Near native performance for many applications

Implications

- Users bring their own stacks
- Operators and systems provide minimal services

The big data continuum …

Next-generation architecture
(mostly open source and cloud based)



Knowns, unknowns and sociology

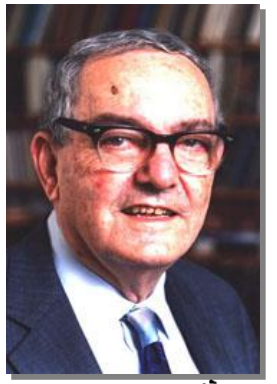
Known questions, the traditional approach

- I know the question, but not the answer
- I'll capture the data and I know the answer is "in there"

Unknown questions, the big data approach

- We have data, but I do not know what is could tell me
- I want to explore and learn

www.infoq.com/resource/presentations/big-data-persistence-nosql



What information consumes is rather obvious: it consumes the attention of its recipients. Hence *a wealth of information creates a poverty of attention*, and a need to allocate that attention efficiently among the overabundance of information sources that might consume it.

Herbert Simon

Recommender techniques

Item hierarchy (Amazon)

- You bought a Kindle™, you'll want a cover

Attributes (Pandora)

- You like 70's pop, you'll like *Simon and Garfunkel*

Item similarity (Netflix)

- You liked *Batman*, you'll like *Spiderman*

User similarity (Walmart)

- People who buy beer generally buy chips

Social network (Linkedin)

- Your connections liked this job posting, so will you

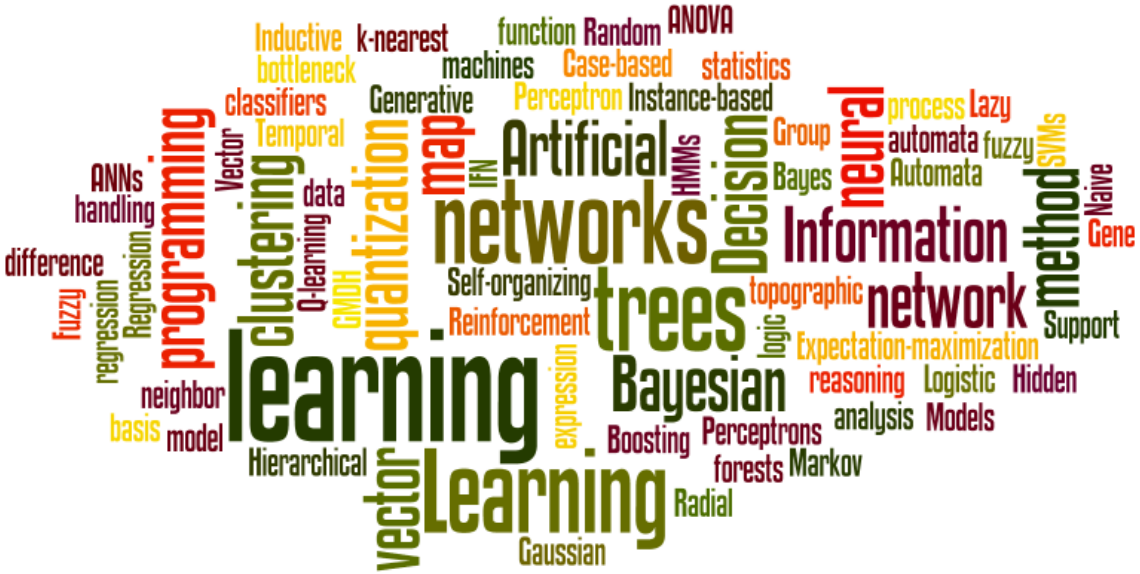
Model based (HPC challenges and needs)

- ***Training, singular value decomposition (SVD), support vector machines (SVM), …***

Deep learning and algorithms

Clustering (grouping similar items)

- Partitioning (K-means) or hierarchical methods



Hidden Markov models (HMMs)

- Markov process with unobserved (hidden) states
- Goal is to infer a process from observations
- Often used for temporal pattern recognition
  - speech, handwriting and gesture recognition

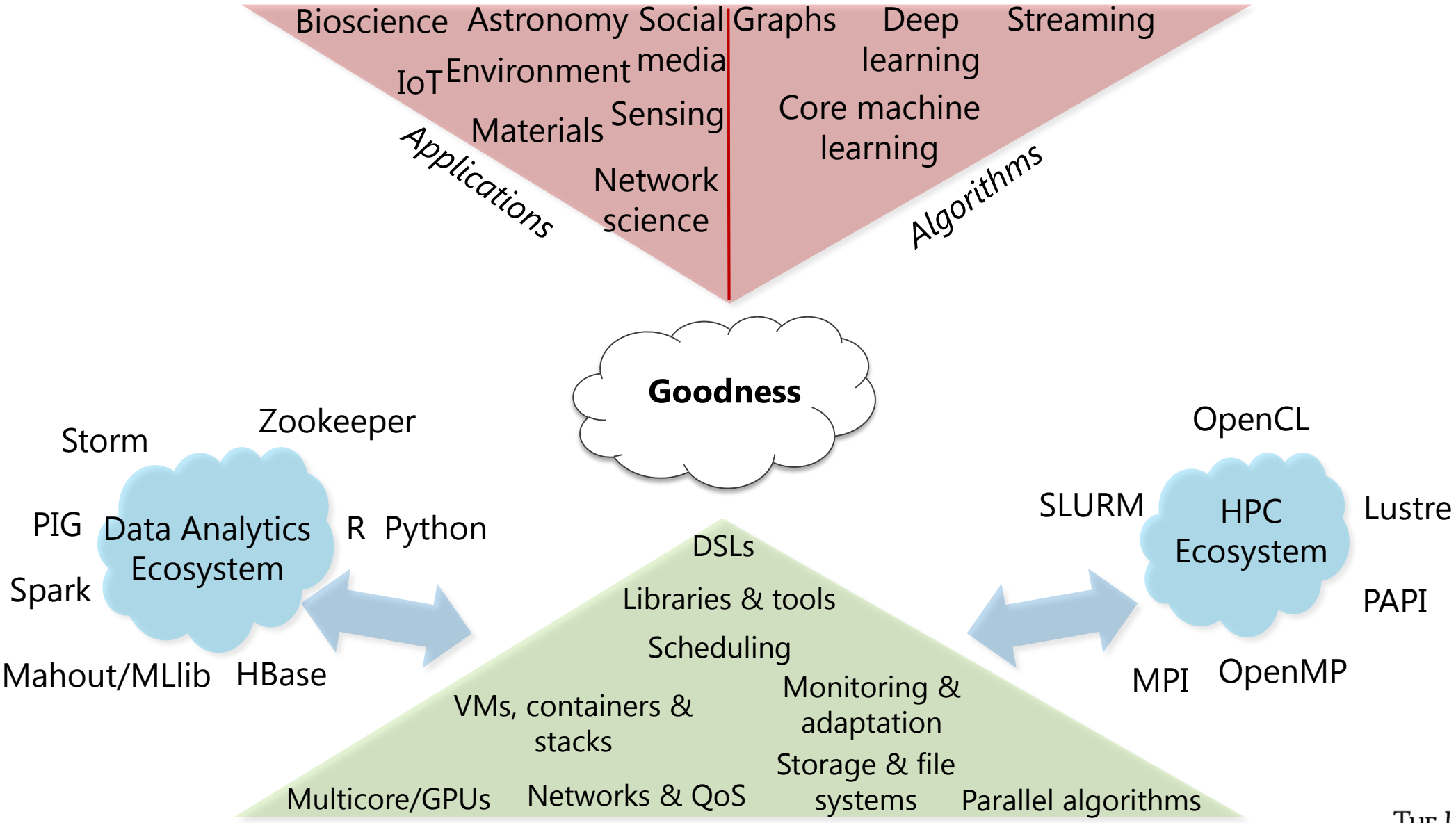
Blind signal separation/feature extraction for dimensionality reduction

- Principal component analysis (PCA), singular value decomposition (SVD)

Artificial neural networks (ANNs)

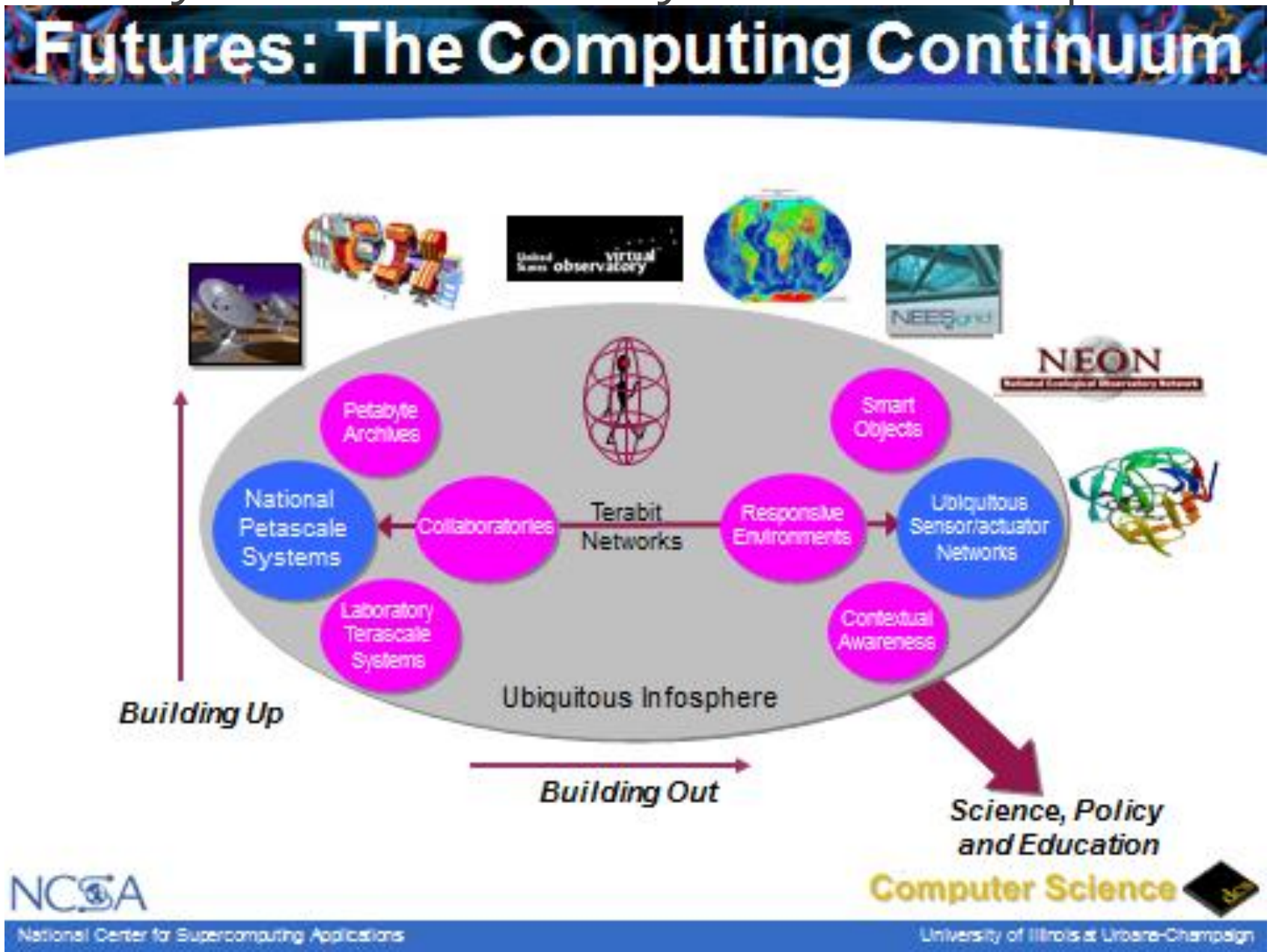
- Training based on inputs/outputs
- Extract linear combination from input, output nonlinear function of combinations

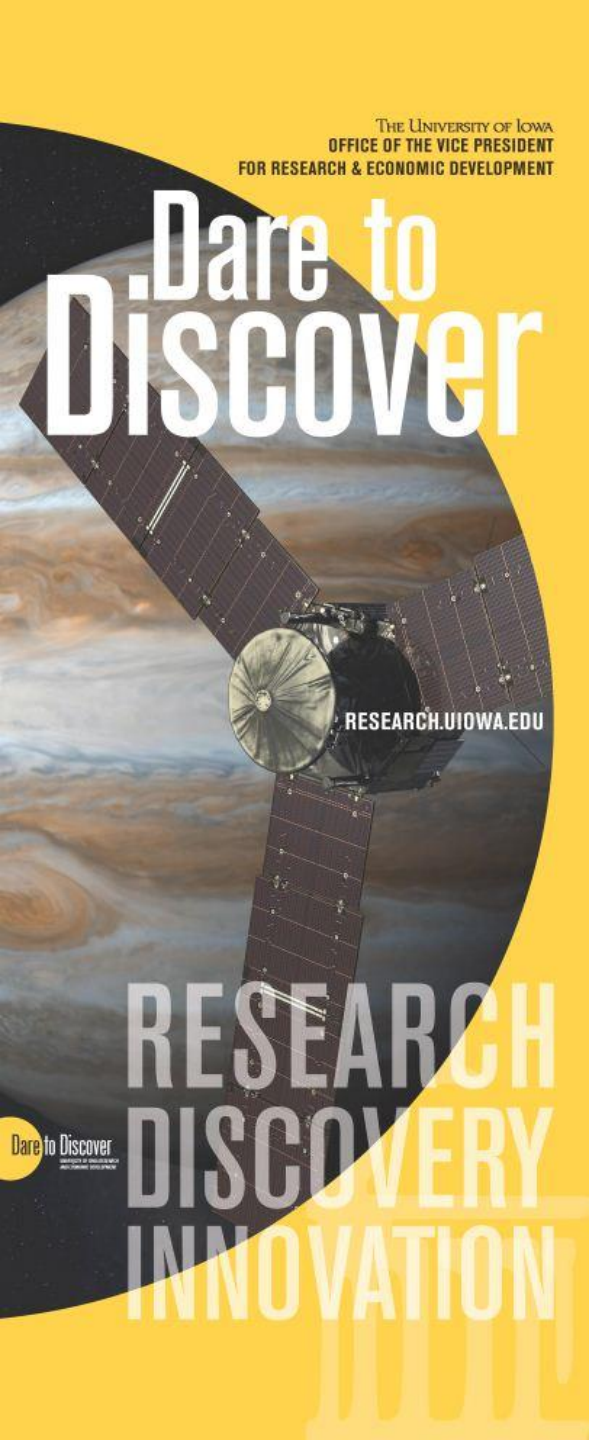
Scalable Ecosystems for Data Science (SEDS)



Reed, Beckman, Choudhary, Dongarra, Fox, Gropp

Not yet fully realized … my 2002 NSF presentation





Discussion

The University of Iowa