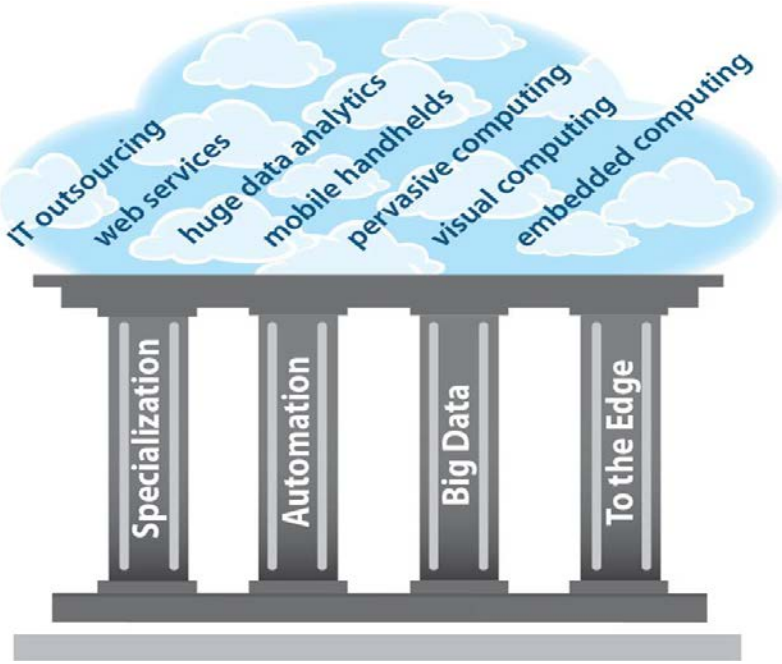
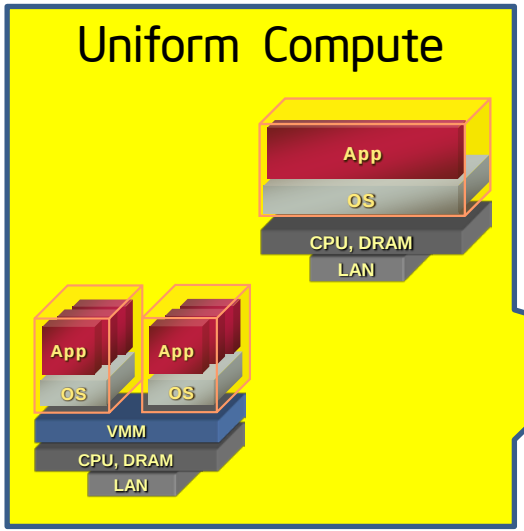
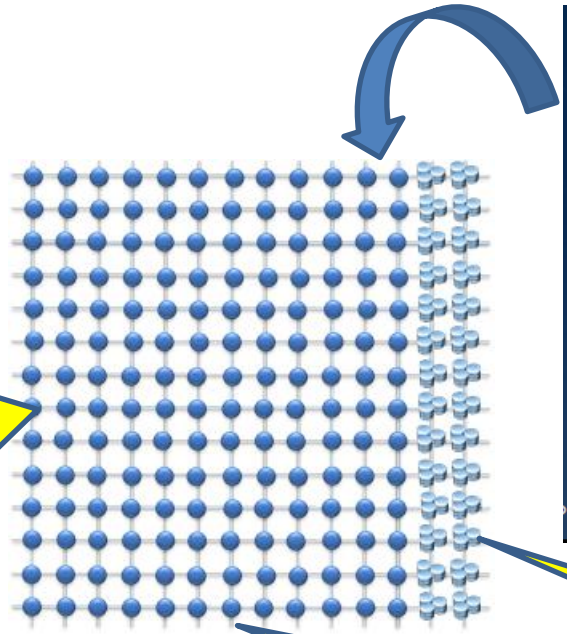


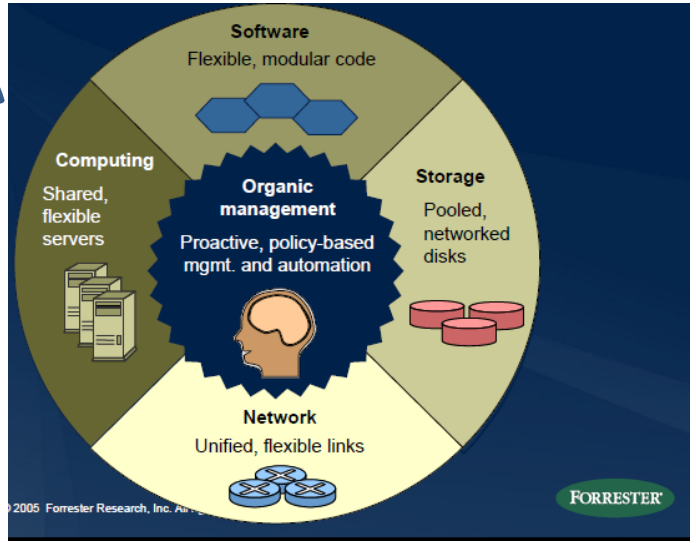
Clumpy Data Centers

Balint Fleischer
GM, Server Architecture
Intel

Idealized Grid Vision





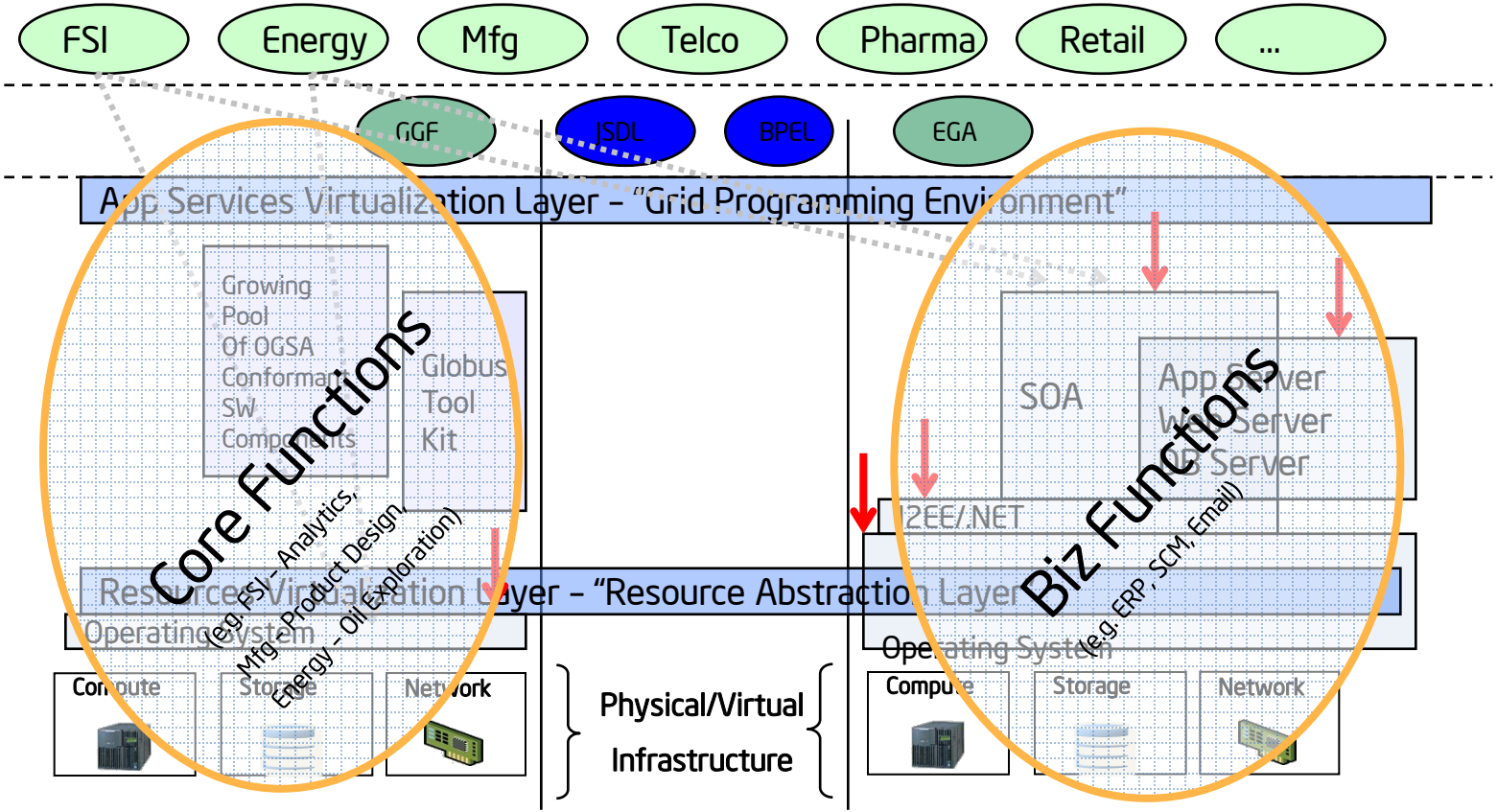


Automated WL and Resource Management

Uniform Fabric

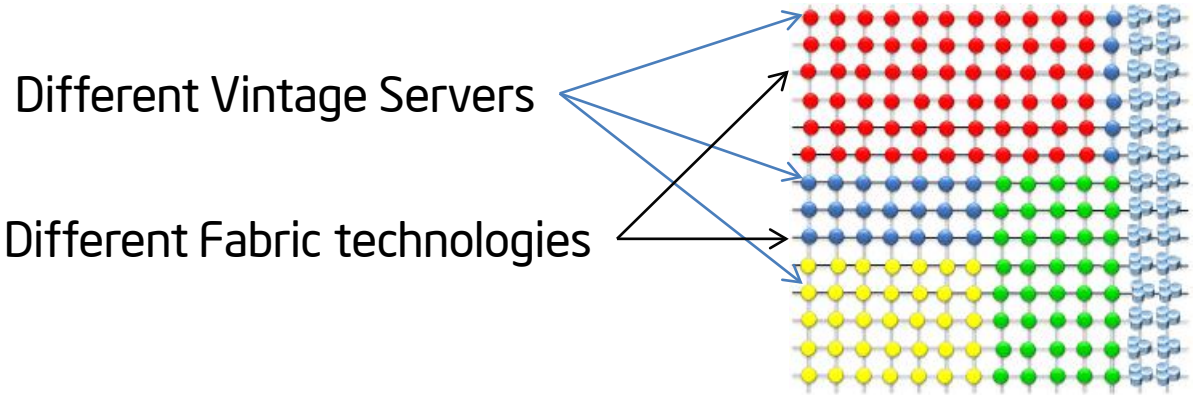
Shared Storage Pool

Enterprise Mapping onto Grid ~2005



Challenging uniformity

Impact of Upgrade Cycles



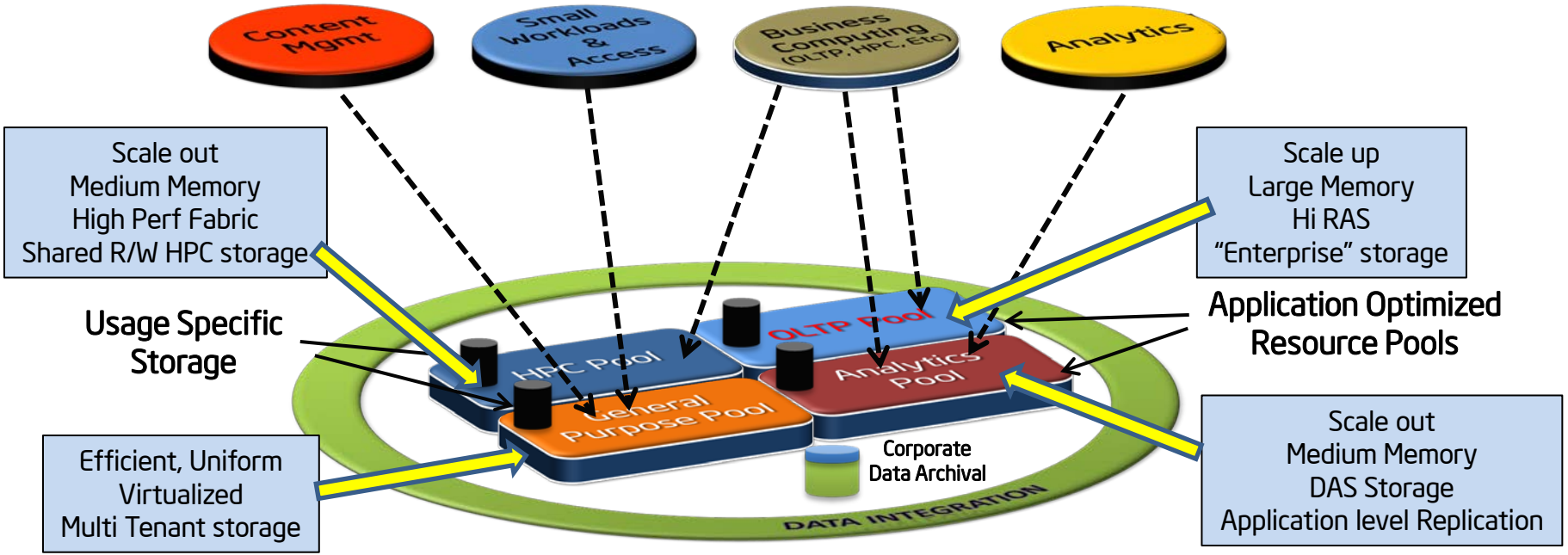
Large Compute capability variation across Data Center

Virtualization to std “Instances” does not abstract platform differences

Staged Fabric upgrades result in BW and latency Hotspots

Challenging Uniformity

New Service Classes Require new capabilities



Can we dynamically "assemble" these pools?

Technology Trends

Projection based On Industry Data

Rack 2012

~600 Cores
~2+ TB Memory
~1/2 PB Storage
~1.4 Tbps Disk IO

3x
15x
8x
3x

Rack 2018

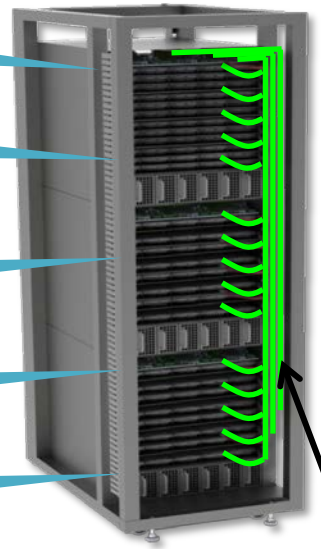
>2000 Cores
>40 TB Memory
~4 PB Storage
>5 Tbps Disk IO
Persistent Memory

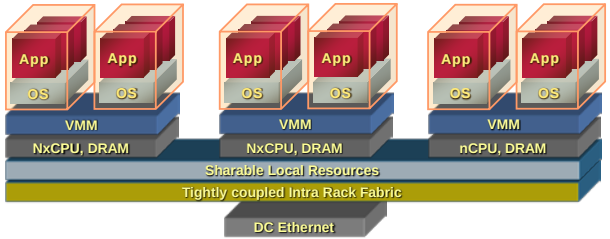
The growing Rack Density
Will certain class of problems fit within a Rack?

Opportunity to *Architect* the Rack

Some Architecture "knobs"

- Very high internal BW?
- Configurable Topologies?
- Low intra Rack Latency for Tight coupling?
- Optimized Local Protocols?
- Shared NVM?

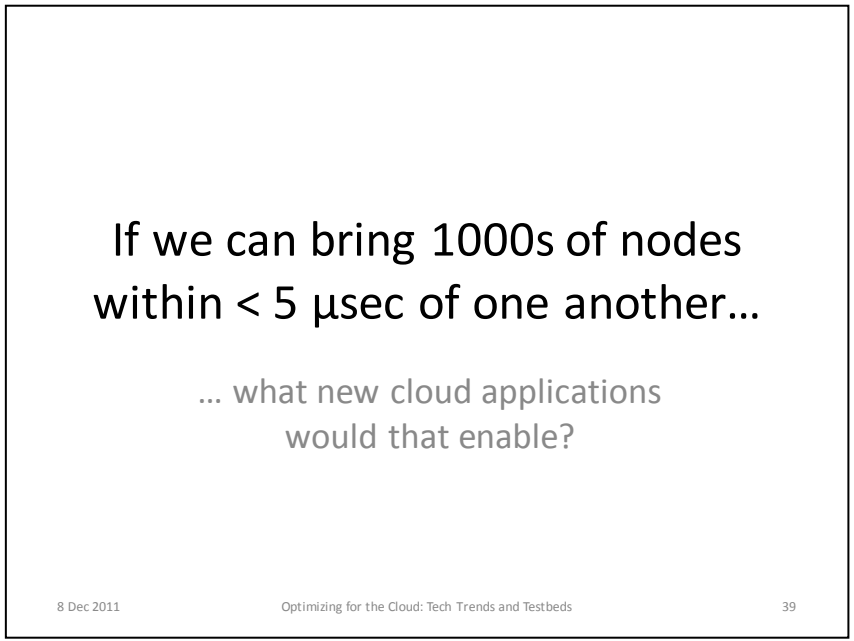




Intra Rack "wiring"

"RackScale Architecture"

From Rich Uhlig's 2011 Keynote



To make the challenge harder

If we can bring 1000s of ~~nodes~~ within < ~~5~~ usec of one another…

… what new cloud applications would that enable?

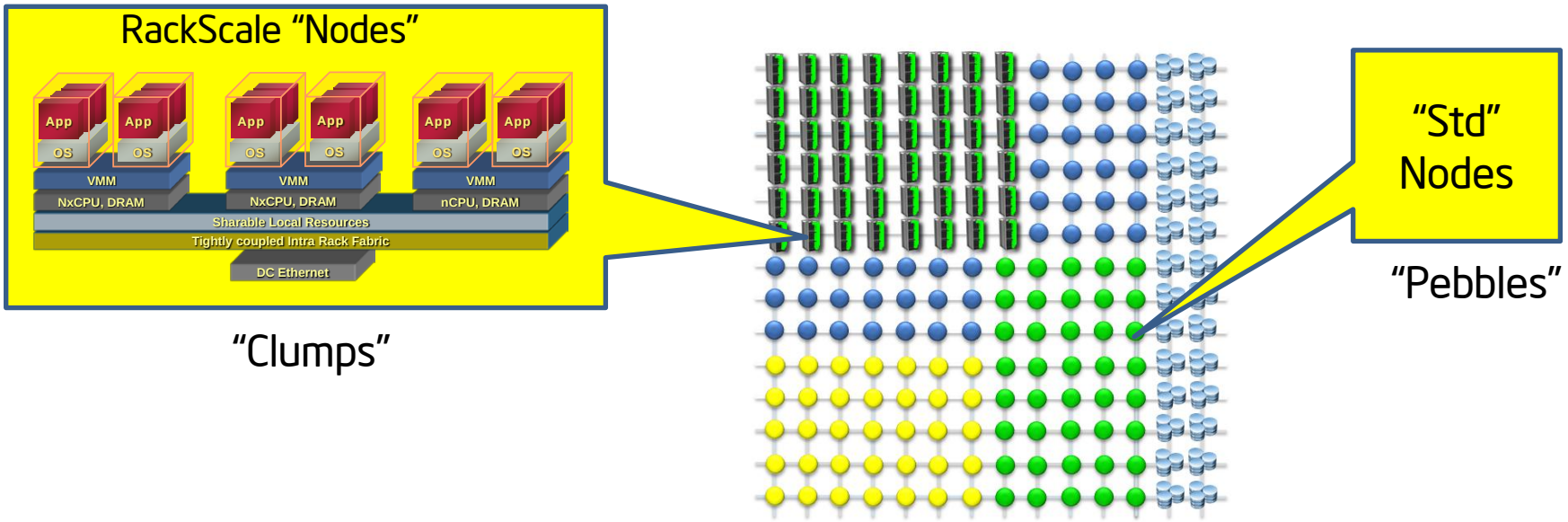
If we can bring 1000s of cores, PB's of Storage AND 10's of TB of memory Within 0.5 usec of one another…..

…. What will happen to "big" systems?
……What will happen to applications?
……What are the usages?

The emergence of the New Mainframe or the New MPP systems?

The impact on Data Centers

Clumpy Data Center



How do we manage this?

How do we integrate "Clumps" and "Pebbles"?

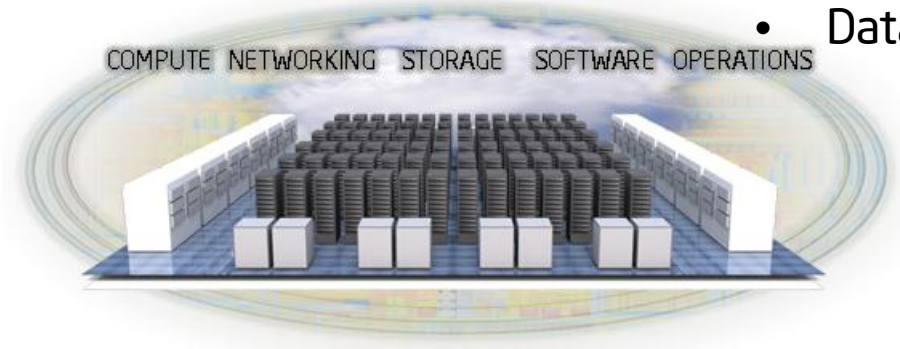
Emergence of Locality

- Neighborhood (Rack)

Xtreme low Latency
New Semantics?
Resource sharing?
>100 Gbps
Traditional Semantics

- Cluster

Low Latency
HPC style Semantics in Cloud?
100 Gbps
Traditional Semantics



- Data Center

Moderate Latency
>100 Gbps, EW Traffic
SDN
Traditional Semantics

Locality based Application architecture and Scheduling

Wish List

Brainstorm using RackScale

Rethink clustering APIs and clustered applications

Rethink small radix Networks and protocols

What is shared memory in a Rack?

Composable Architectures