



# Open Cloud Consortium: An Update

Stuart Bailey, CTO, Infoblox  
Open Cloud Consortium

OMG Cloud Interoperability Roadmap Session

December 10, 2009

# Part 1.

## Overview of the Open Cloud Consortium (OCC)



[www.opencloudconsortium.org](http://www.opencloudconsortium.org)

- 501(3)(c) Not-for-profit corporation
- Supports the development of standards and interoperability frameworks.
- Supports reference implementations for cloud computing.
- Manages testbeds: Open Cloud Testbed and Intercloud Testbed.
- Develops benchmarks.

# OCC Members

- Companies: Aerospace, Booz Allen Hamilton, Cisco, Infoblox, Open Data Group, Texeltek, Yahoo
- Universities: CalIT2, Johns Hopkins, MITLincoln Lab, Northwestern, University of Illinois at Chicago, University of Chicago
- Government agencies: NASA
- Organizations: Sector Project

# OCC Working Groups

1. Large Data Cloud Working Group
2. Open Cloud Testbed Working Group.
3. Intercloud Testbed Working Group
4. Open Science Data Cloud Working Group

# Part 2.

## Intercloud Testbed

# We have several cloud standards...

- Platform as a Service
  - Cloud Compute Services
  - Data/Table Cloud Services
  - Cloud Storage Services



Large Data Cloud  
Interoperability  
Framework

**SNIA**

SNIA Cloud Data  
Management  
Interface (CDMI)

- Infrastructure as a Service
  - Virtual Data Centers (VDC)
  - Virtual Networks (VN)
  - Virtual Machines (VM)



Open Virtualization  
Format (OVF)



Open Cloud Computing  
Interface (OCCI)

# Where are the Gaps?

- Naming entities in IaaS&PaaS
  - Bridging IaaS&DaaS
  - Services that span multiple VMs, ....
- 
- Platform as a Service
    - Cloud Compute Services
    - Data as a Service
  - Infrastructure as a Service
    - Virtual Data Centers (VDC)
    - Virtual Networks (VN)
    - Virtual Machines (VM)
    - Physical Resources



Large Data Cloud  
Interoperability  
Framework

**SNIA**

SNIA Cloud Data  
Management  
Interface (CDMI)



Open Virtualization  
Format (OVF)



Open Cloud Computing  
Interface (OCCI)

# Bridging the Gaps...A Small Step

- Platform as a Service

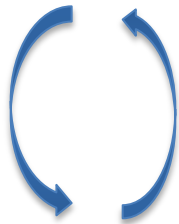
- Cloud Compute Services
- Data as a Service



Large Data Cloud  
Interoperability  
Framework

**SNIA**

SNIA Cloud Data  
Management  
Interface (CDMI)

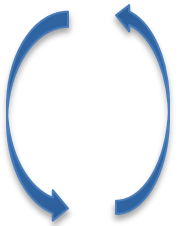


Dynamic infrastructure  
service linking IaaS and  
DaaS

- Infrastructure as a Service

- Virtual Data Centers (VDC)
- Virtual Networks (VN)
- Virtual Machines (VM)
- Physical Resources

Dynamic infrastructure  
service naming and linking  
entities in the IaaS layers

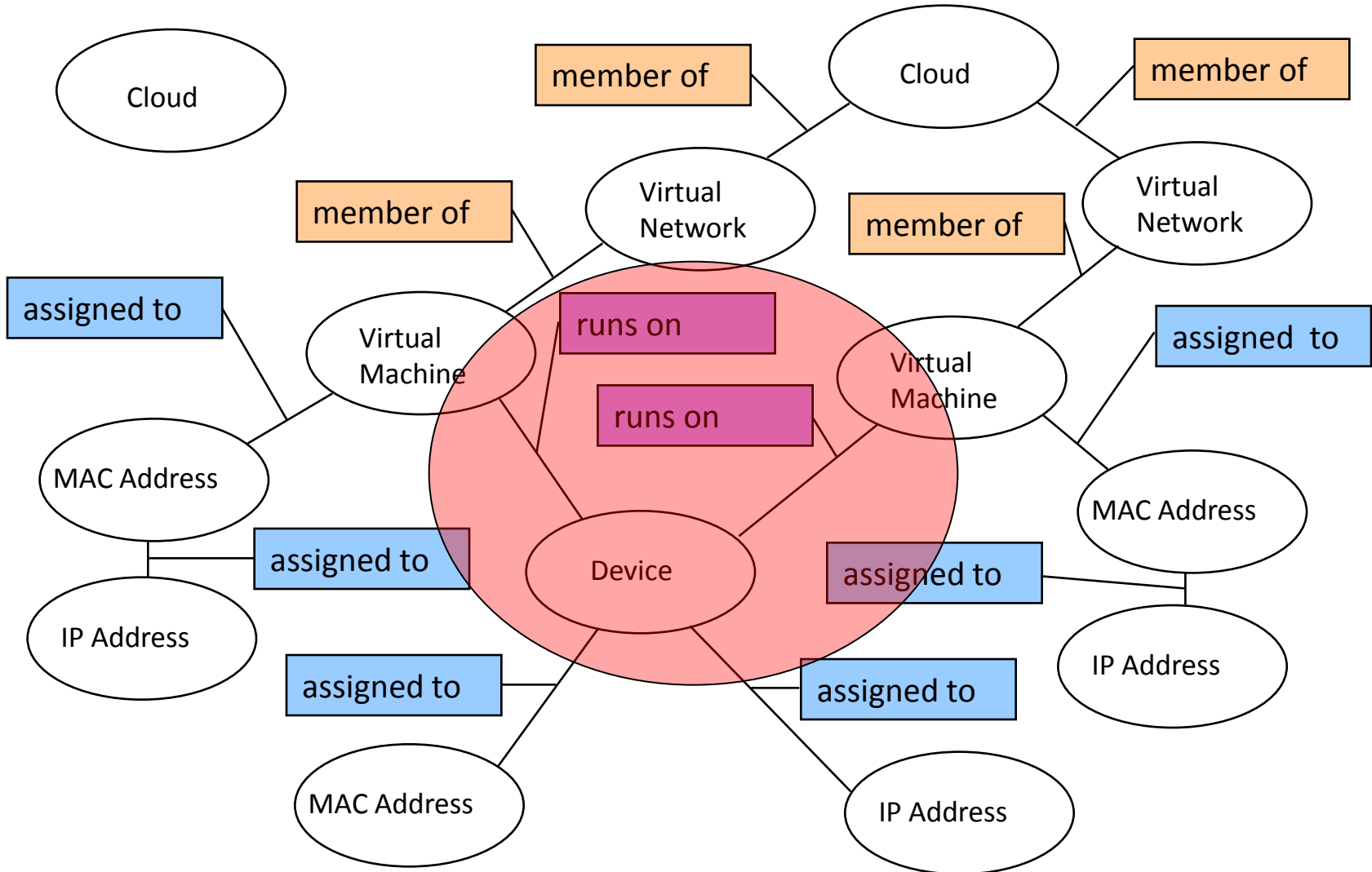


Open Virtualization  
Format (OVF)

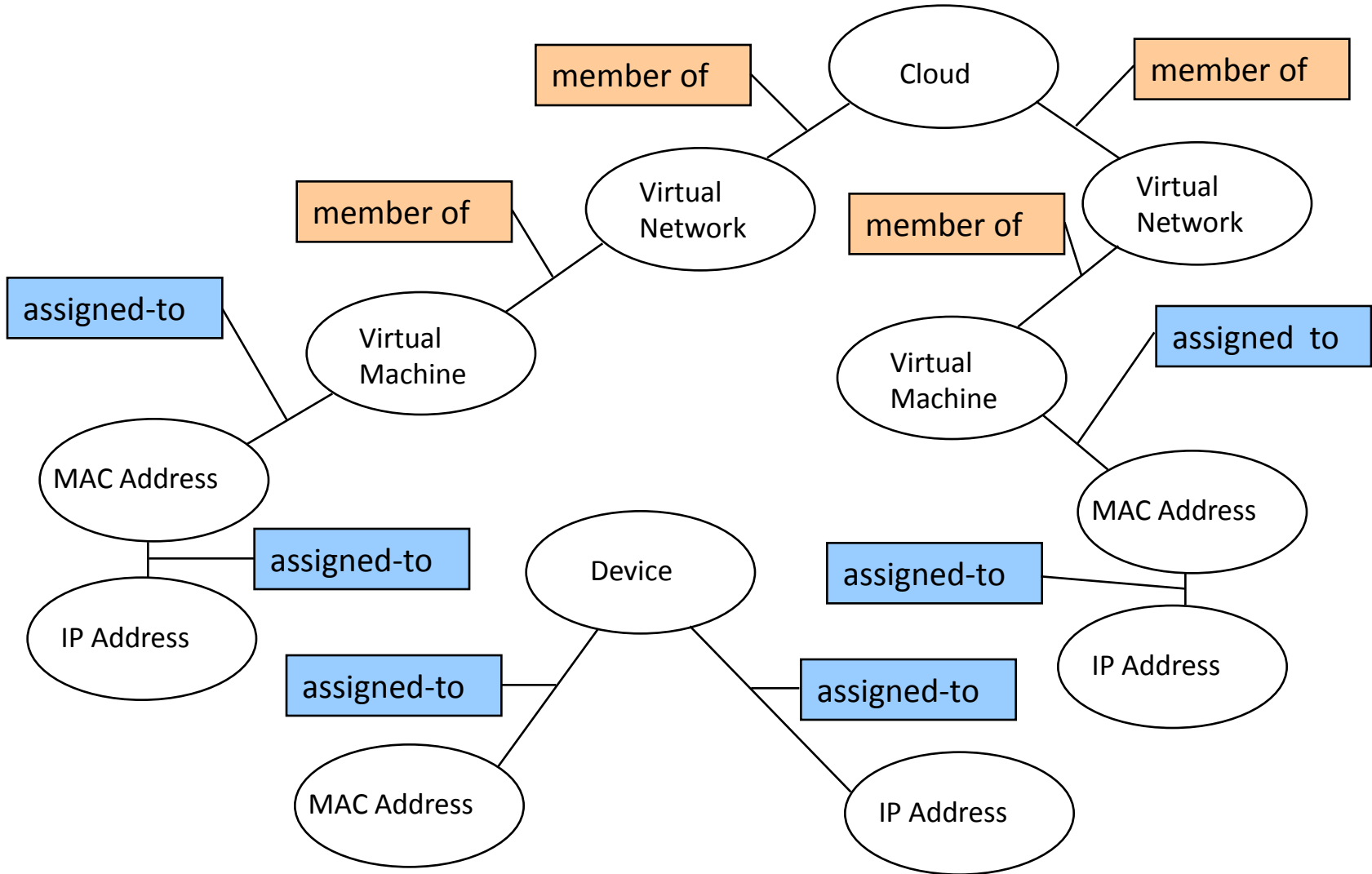


Open Cloud Computing  
Interface (OC CI)

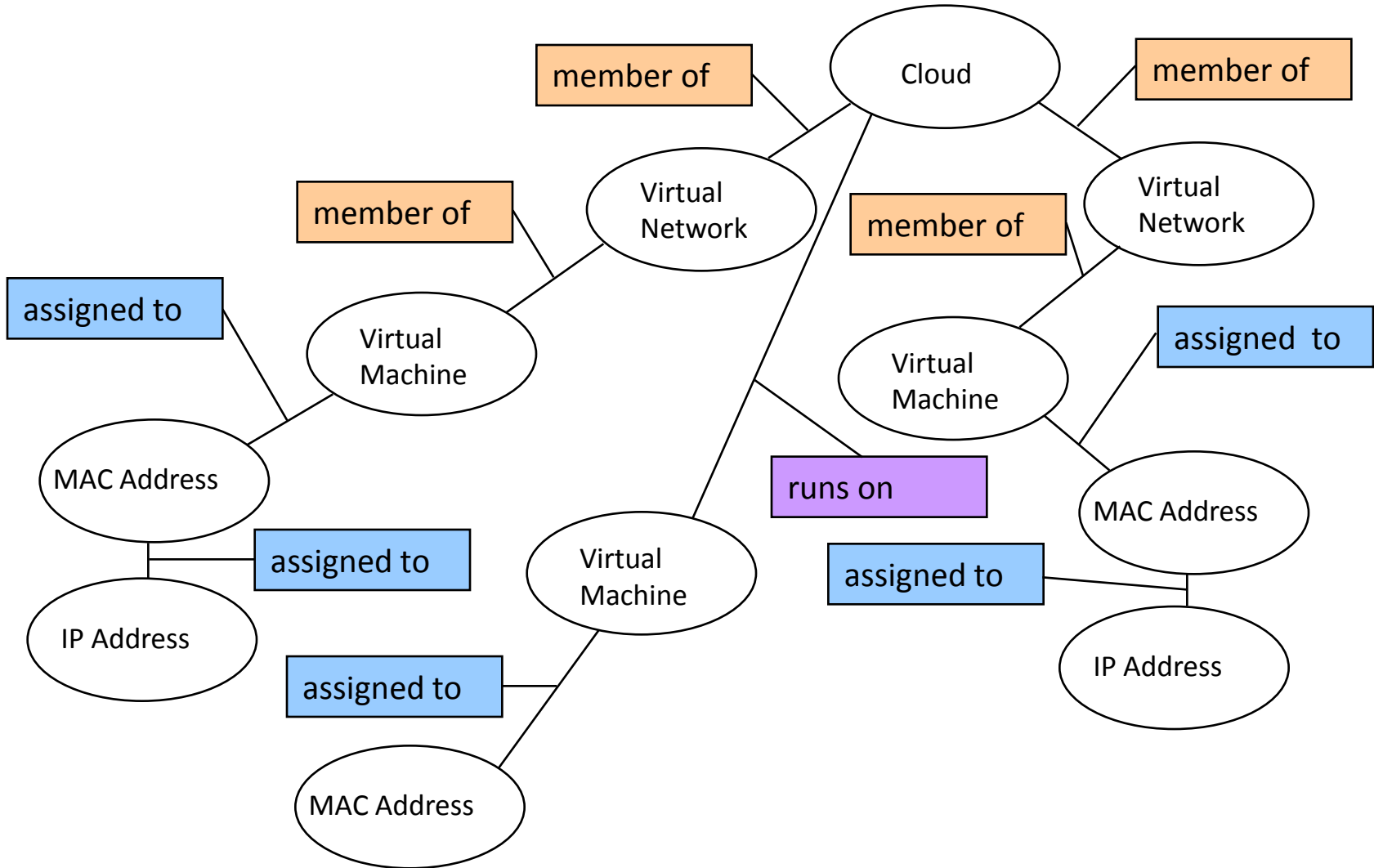
# Complex Binding Patterns May Emerge



# Binding Patterns Evolve



# Binding Patterns Evolve



# IF-MAP Operations

Publish: Tell others that... <metadata...>

- Clients store metadata into MAP for others to see
  - Incorporates create, modify and delete functionality

Search: Tell me if... *match*(metadata pattern)*j*

- Clients retrieve published metadata associated with a particular identifier and linked identifiers
  - Constrained by link-match and result-filter criteria
  - Constrained by maximum depth and size criteria

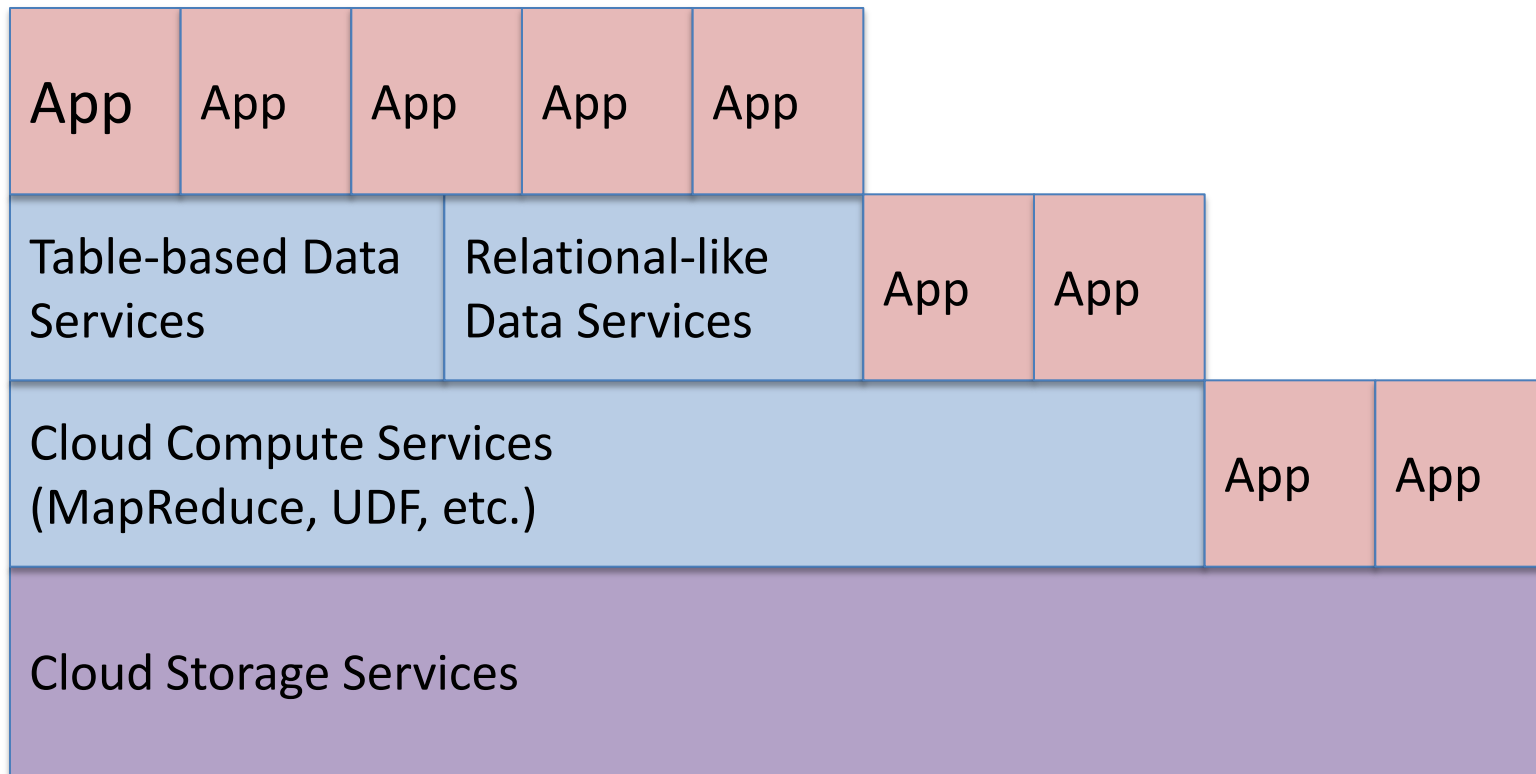
Subscribe: Tell me when... *match*(metadata pattern)

- Clients request asynchronous results for searches that match when others publish new metadata
  - A client's subscription consists of a list of one or more searches
  - Client names its searches so that asynchronous results are unambiguous

# Part 3. Large Data Cloud Working Group

\*

# Focus of Working Group



- Developing APIs for this framework.

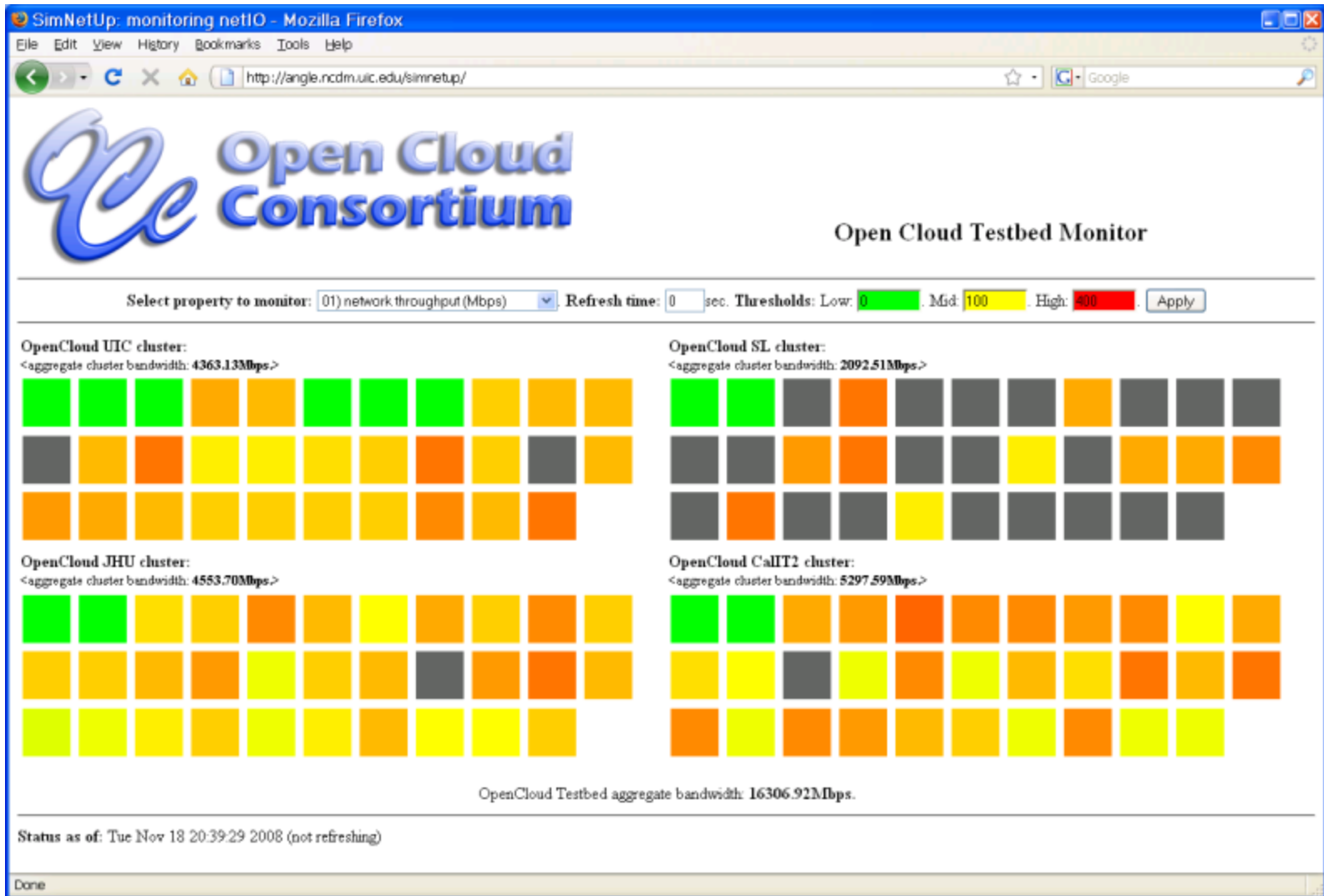
# Benchmarks for Large Data Clouds

- Until recently, the only benchmark used was Terasort (sorting 10 billion 100 byte records)
- Replaced by Gray Sort and Minute Sort
- Gray Sort tries to maximize TB / min sorted on 100 TB or more of data.
- Hadoop holds the current Gray Sort and Minute Sort records.
- Problem: sort is just one of the types of work load for analytic applications

# MalStone

- MalGen – generates synthetic data with realistic distributions.
- MalStone A & B – “stylized” computations that can be used as benchmarks for architectures, software and systems for large data clouds.
- Open source and available at [code.google.com/p/malgen](https://code.google.com/p/malgen)

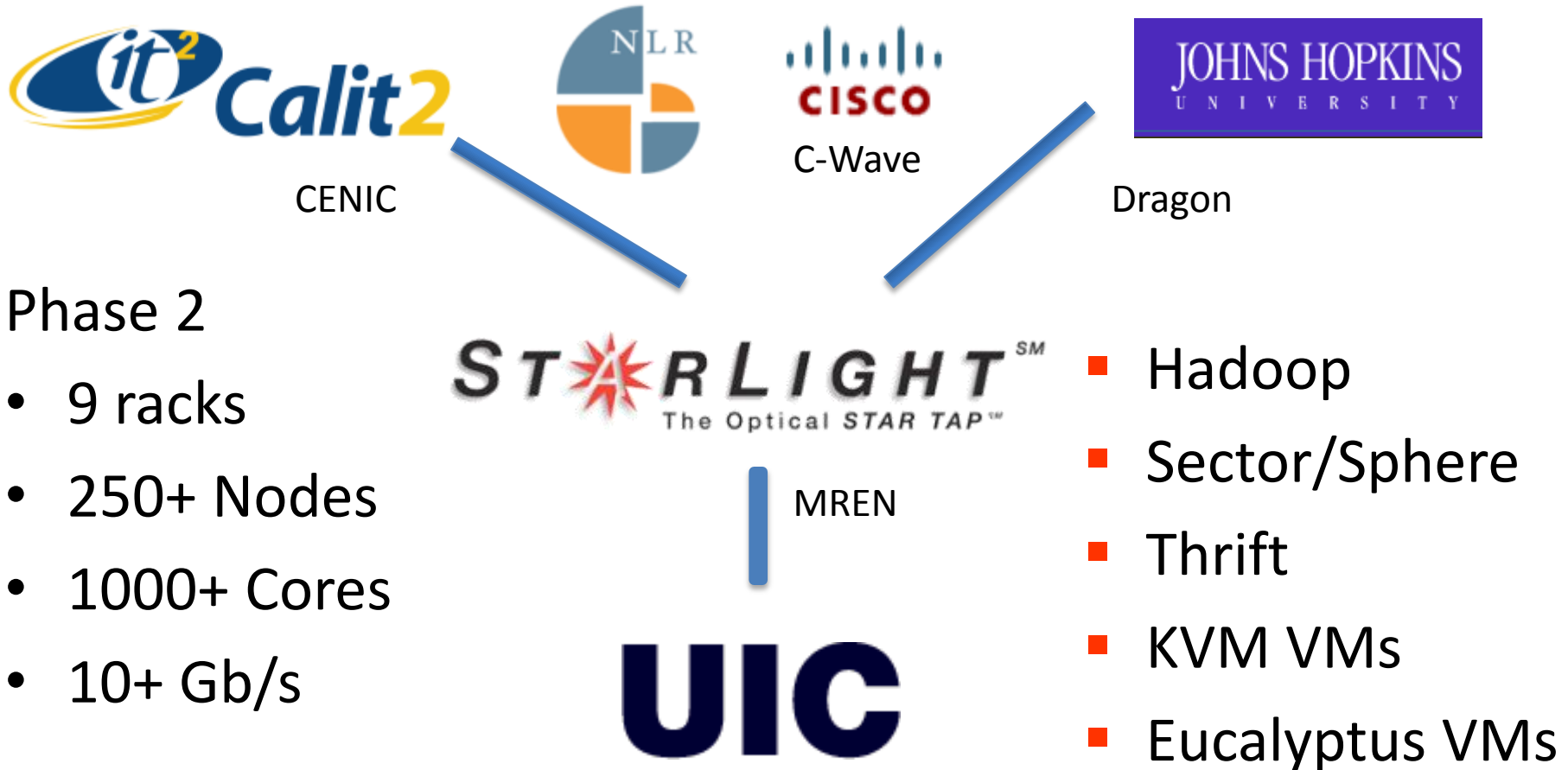
# Part 4. OCC Testbeds & Clouds



# Condominium Clouds

- In a condominium cloud, you buy your own rack or bunch of racks.
- The racks are managed and operated by the condominium association, in this case the OCC.
- If your rack is 120 TB, you get the rights to c. 40 TB of storage in the cloud. The rest is a shared resource.
- The Open Cloud Testbed is a condo cloud managed by the OCC.

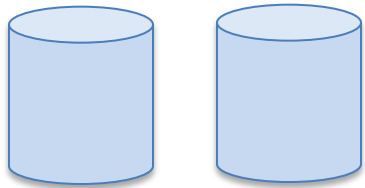
# Open Cloud Testbed



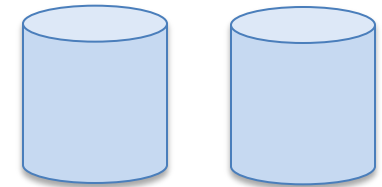
## Phase 2

- 9 racks
- 250+ Nodes
- 1000+ Cores
- 10+ Gb/s

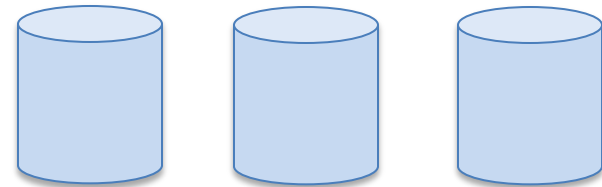
# Open Science Data Cloud



network  
cyber cloud



sky cloud



biocloud

Planned Launch Dec 16, 2009

# For More Information



[info@opencloudconsortium.org](mailto:info@opencloudconsortium.org)

[www.opencloudconsortium.org](http://www.opencloudconsortium.org)