OpenGridForum
OPEN FORUM | OPEN STANDARDS

A Perspective on Scientific Cloud Computing

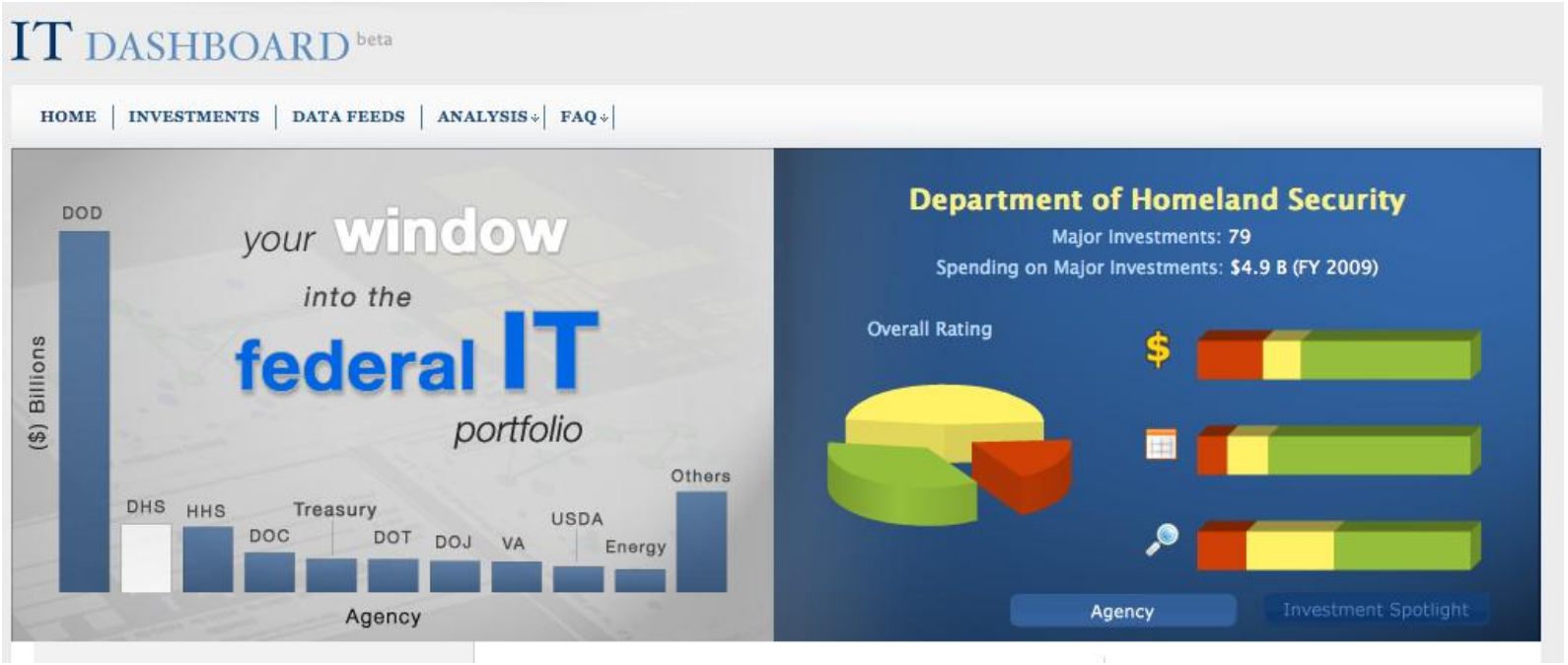
Science Cloud Workshop, June 21, 2010

Dr. Craig A. Lee, lee@aero.org
President, OGF, www.ogf.org
Senior Scientist, The Aerospace Corporation
A non-profit, federally funded R&D center

Introduction

- Cloud computing has many potential benefits
  - *As well as potential pitfalls*
- Different organizations -- industrial, scientific & governmental – will have different requirements and perceived risks for cloud computing
  - *How will these different requirements and risks drive their cloud adoption?*
- How will these different requirements drive cloud deployment trajectories?
  - *And what can we do about it?*

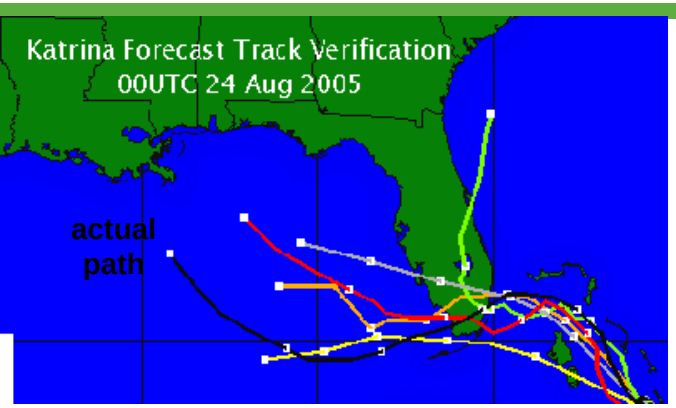
www.gpoaccess.gov/usbudget



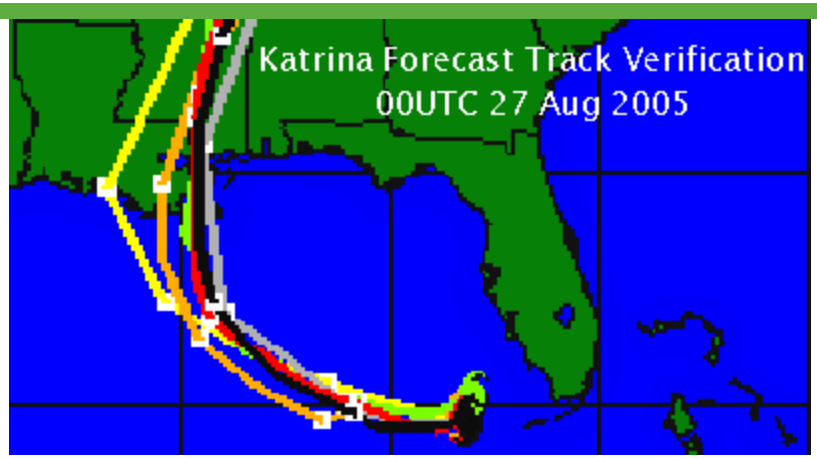
- FY 2010 US Federal IT Budget: $79B
  - ~70% spent on maintenance
- US Federal CIO pursuing cloud computing
  - Apps.gov, data.gov web sites stood up
  - NASA Ames *Nebula* cloud to be first “back-end”

A Scientific and Operational Grand Challenge

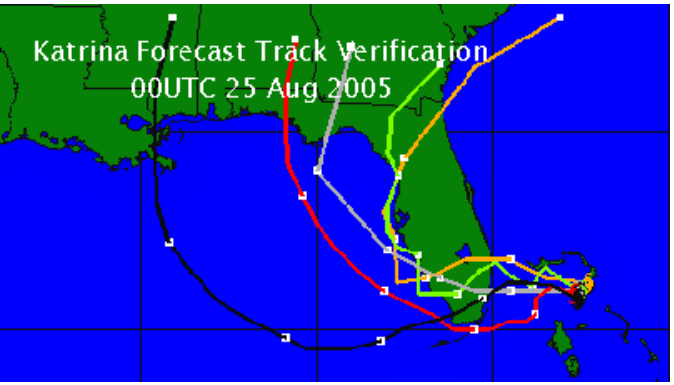
Basic Science, Computational Science, Data Access, On-Demand Resources



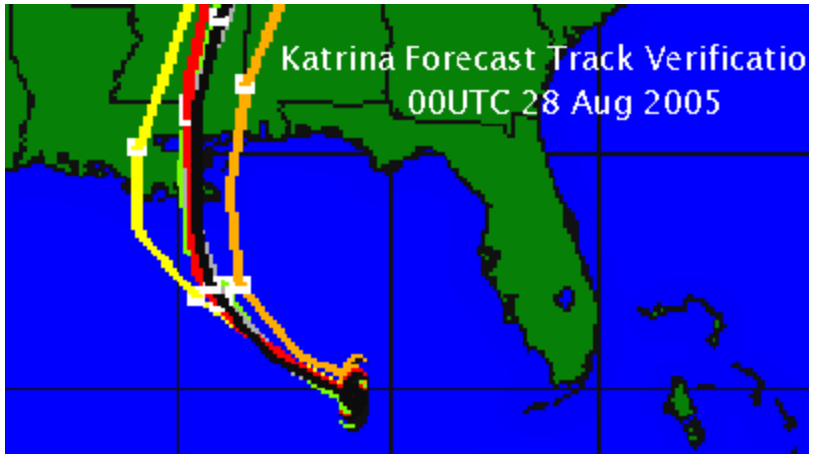
5 days out



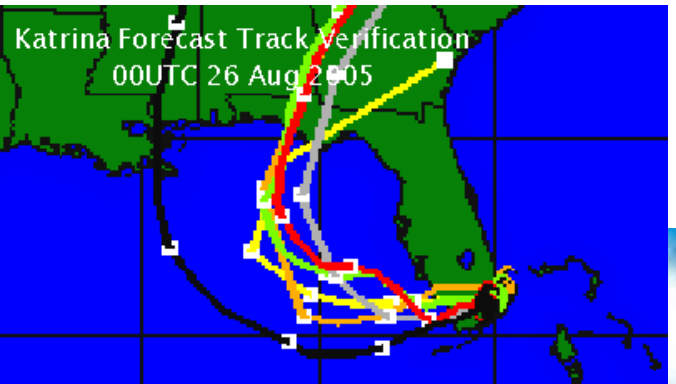
2



4



1



3

Philip Bogden, (Past) SCOOP Program Director
scoop.sura.org/documents/MTS_Journal_Article_Final.pdf

SURA Coastal Ocean Observing and Prediction (SCOOP) Program

Expected Benefits, in General

OpenGridForum

Provider-oriented Benefits

- Improver server utilization
- Improved reliability
- Greener IT
- Clear business models

User-oriented Benefits

- Commodification of compute resources
- Managing surge requirements
- Ease of application deployment
- Virtual ownership of resources

Issues

General

- Security
  - *Information Assurance*
- Governance
  - *Regulatory, Legal -- SLAs*
- Portability & Interoperability
  - *Data and Applications*
- Licensing
  - *Cloud licenses must be hardware & location agnostic*
- Cost
  - *Where's the break-even point?*

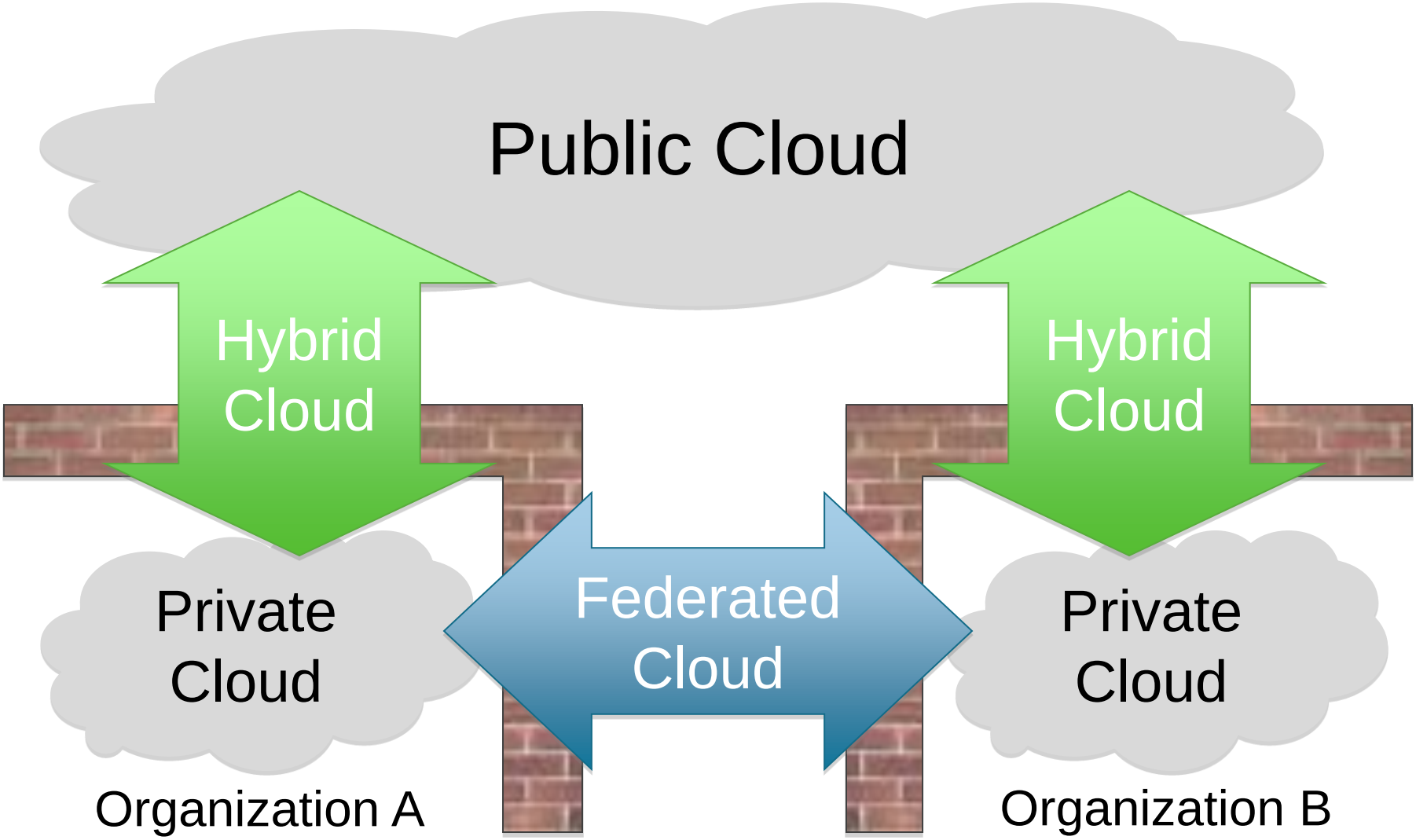
Scientific

- Performance Management
  - *Abstraction vs. Control -- SLAs*
- Portability & Interoperability
  - *Data and Applications*
- Execution Models
  - *Frameworks & SaaS*
- Security
  - *Information Assurance*
- Governance
  - *Regulatory, Legal -- SLAs*
- Cost
  - *Where's the break-even point?*

Major Cloud Deployers & Adopters?

- Science
  - Many operational grids introducing cloud functionality
- Industry
  - Beyond client-provider to *business-to-business*
- Government -- National Cloud Initiatives
  - US Cloud Storefront Concept
  - Japanese Kasumigaseki Cloud Concept
  - UK G-Cloud Concept

Cloud Deployment Modes



Key Observation

The distinction between private and public clouds is really a relative distinction between whether you "own" the resources or not -- whether the resources are inside or outside your security perimeter, i.e., your administrative domain.

- Can you enforce membership, usage, and security policies?
  - Do you know who your cloud tenants are and how they will behave?
- 451 Group identifies *Trust and Control* as the primary inhibitor of cloud technology, followed by *Interoperability, Portability and Licensing* ("It Is Cloud" report, June 2010)
- Cloud Security Alliance "Security Guidelines" document listing many security and regulatory issues

What is the Likely Trajectory?

- Top-Down deployment/adoption of “public cloud”?
  - Appealing from “end-goal” perspective, such as national cloud initiatives
  - *Challenge: Recruiting enough users whose security, reliability, and control requirements can be met*
- Bottoms-Up deployment of “organizational clouds”?
  - Realistic starting-point for many organizations -- much easier to control policy, security, risk, liability
  - *Challenge: Mitigating the risk of creating “cloud silos” that are non-interoperable -- cannot federate or hybridize*

My Opinion

- Commercial public clouds will not evolve fast enough to accommodate the spectrum of user requirements across industry, science and government
  - Many US gov agencies deploying their own private clouds
    - Informal US GSA survey identified at least 50 gov cloud projects in 2009
  - Science cloud users will want more control of their applications
- *Private clouds will be the predominant way that organizations adopt cloud computing for "serious work"*
  - *Allows in-house solutions to security, regulatory, performance issues*
- This premise produces a trajectory of deployment issues, which in turn, can be used to produce a deployment, develop and research roadmap

Progression of Issues & Concerns

OpenGridForum

	Private Everything within secure perimeter	Federated Known number of known tenants	Hybrid Unknown tenants but secure perimeter	Public No secure perimeter

Technical	• Job types & mixes
	• Data access/interop
	• Storage mgmt
	• Workload mgmt
	• Reliability
	• Energy mgmt
	• Governance
	  • VMs, VNs, VDCs
		• ID mgmt
		• VO mgmt
		• Distributed workload mgmt
		• Portability
		• Interoperability
			• SLAs
			• Full Security
			  • Privacy
			  • Data Leakage
			  • Denial of Service
			  • Eff. data deletion
				• Practical ways to operate on encrypted data (not!)
				• Virtual Private Clouds

Legal/Org.	• Costing models	• Agreement on joint operations • e.g., International Grid Trust Federation	• Outsourceable tasks • Cost Predictability • Liability • Reporting • Co-tenant reputation • Provider viability • Audits	• On-site inspections • Understand & test provider's operation

Landscape	• Avoid vendor lock-in
	• Harmonize/shake-out basic infrastructure standards
		• Harmonize/shake-out relevant standards
		  • Federation
		  • Distributed mgmt
			• Legal Precedent
			• Harmonize/shake-out relevant standards
				• Harmonize/shake-out relevant standards

Issues/Concerns Accumulate Left to Right

Draft Roadmap “Dartboard”

	Phase I	Phase II	Phase III	Phase IV
Deployment	• Existing IaaS processing and storage • ID mgmt	• VO mgmt • Dist Wrkld Mgmt • Data access & interop	• Hybrid infrastructure • SLAs • Auditing/Monitoring/Reporting	• VN, VDC, VPC • Secure operations • Organizations deprecate their own infrastructures
Development & Risk Mitigation	• Develop common, extensible data arch & semantics • Testbed Bake-off • Reliability practices • Workload mgmt • Costing models • Joint practices	• Energy Mgmt • SLAs at scale • PaaS capabilities • Auditing/Reporting mechanisms • Regulatory support • Set legal precedent	• Cloud DNS • VN, VDC, VPC demonstration • Various security tools	• Virtual Enterprise/Mission Support • Autonomic systems
Research	• SLAs & governance to support job mix • Event Notification at scale • Various security issues	• Naming – Cloud DNS • VN, VDC, VPC mgmt • Autonomic policy enforcement	• Virtual Missions • Virtual Enterprises • Encrypted operation	• Quantum computing

A Coordinated Short List

- Identity Management and Virtual Organizations
  - Role-based authorization
  - Supports cloud federation and business-to-business operations
  - Manage “ecosystem” of Cloud Trust Management
- Portability and Interoperability
  - Move workload between Cloud Provider A and B, while maintaining security and service levels
  - Move workload "out" and "back-in" to same cloud, while maintaining security and service levels
  - Common API semantics for managing cloud applications
- Control and Performance Management
  - Clouds are enabled by a simplified interface that "abstracts away" how the actual infrastructure works
  - What are the minimal abstractions that can be exposed – through Service Level Agreements -- to users to enable adequate control?

What Can We Do About This?

- *How can we drive cloud standards and adoption?*
- Coordinate technology roadmapping efforts
  - Leverage National Cloud Initiatives
- Coordinate existing groups & resources
  - Many standards organizations pursuing cloud standards
- Promote technology demonstrations
  - "Shake-out" standards and implementations
- Promote a more formal process

Cloud-Standards.org

- An informal group of Standards Development Organizations (SDOs) collaborating to coordinate and communicate standards for cloud computing, networks and storage
  - Wiki: cloud-standards.org
  - Mailing List: groups.google.com/group/CloudStandards

OGF DMTF distributed management task force, inc. SNIA CSA cloud security alliance OMG OBJECT MANAGEMENT GROUP tmforum OASIS Open Cloud Consortium CLOUDCOMPUTINGINTEROPERABILITYFORUM

- Different SDOs bring different but complementary technologies & capabilities
  - Storage, execution models, deployment models, service level agreements, security, authentication, privacy
- All interested, committed persons and organizations with relevant technical skills can participate

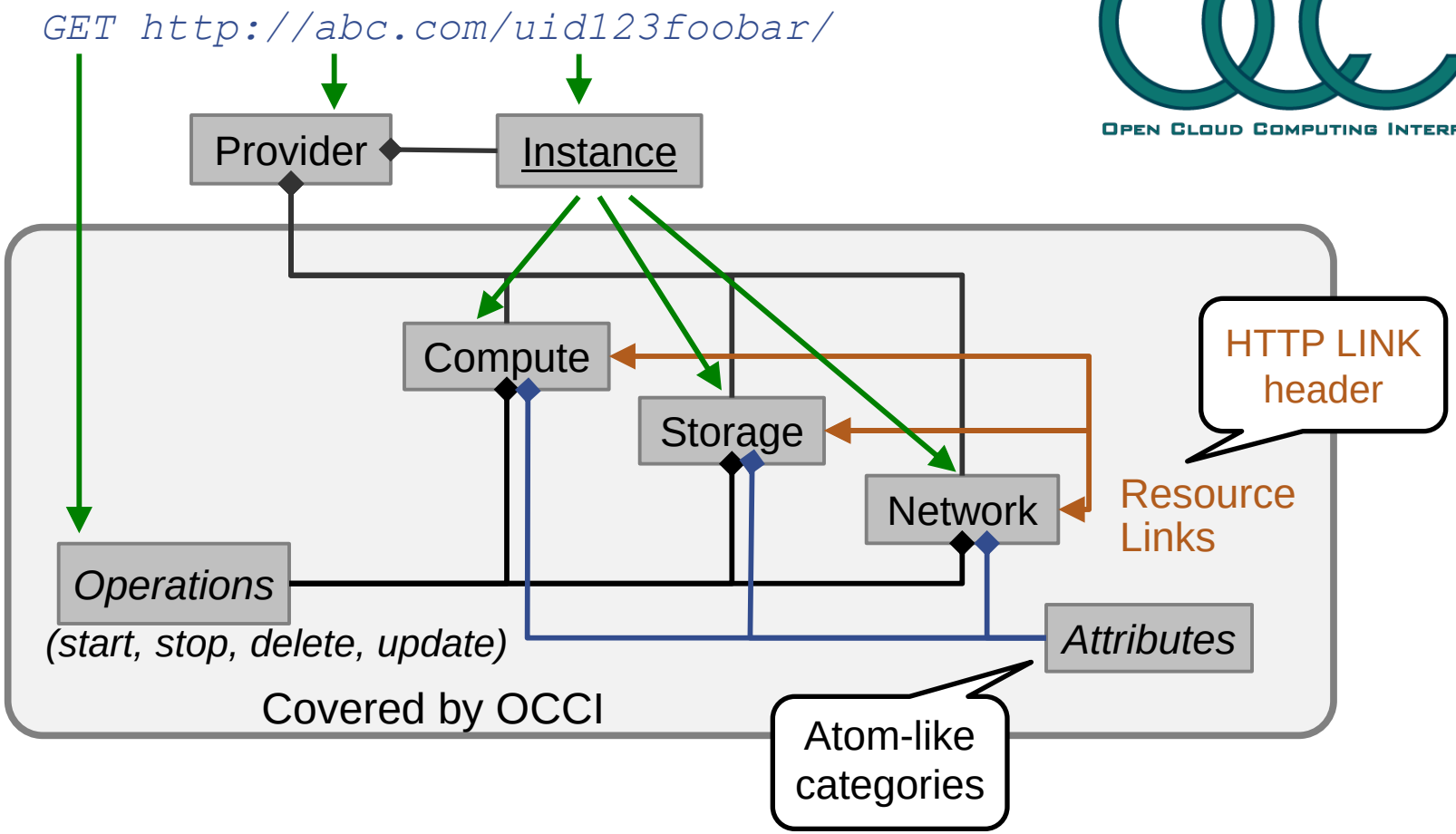
Emerging Cloud Standards

- OGF Open Cloud Computing Interface
- DMTF Open Virtualization Format
- SNIA Cloud Data Management Interface
- *Together these represent the basis for standard IaaS*

Open Cloud Computing Interface

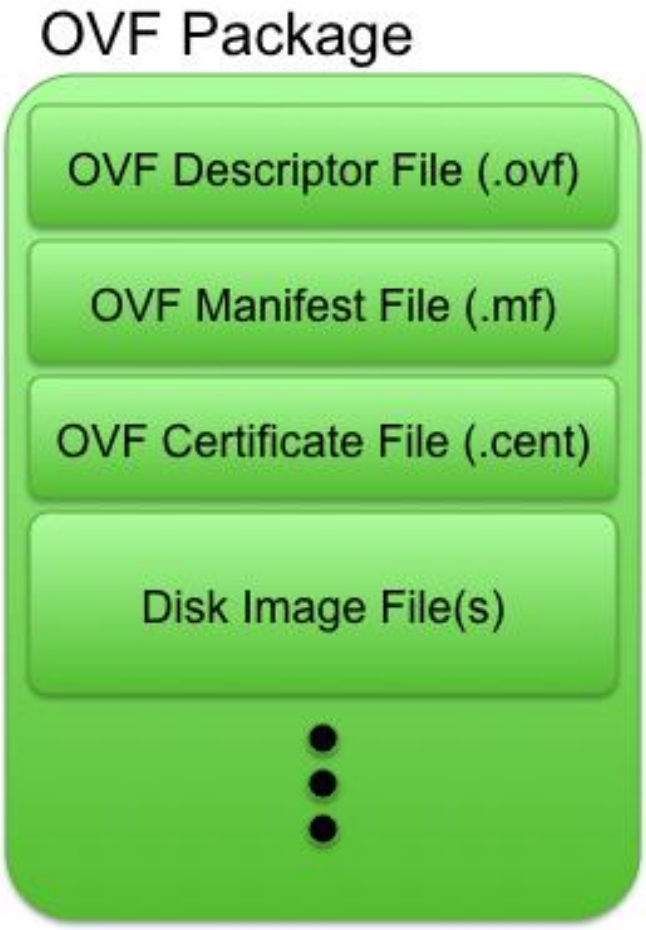
- Focus on interoperable IaaS Cloud API: *www.occi-wg.org*
- Goal: Creation of a simple and RESTful API
  - Simple and very extensible: ~15 commands
- Solid community interest
  - 160 members on mailing list across industry & academia
- Four Main Documents Available
  - The OCCI Core & Models
  - The OCCI Infrastructure Models
  - OCCI XHTML5 Rendering
  - OCCI HTTP Header Rendering
- Four implementations
  - OpenNebula (existing)
  - Istituto Nazionale di Fisica Nucleare (INFN) (existing)
  - SLA@SOI (planned)
  - anonymous

OCCI Overview



DMTF Open Virtualization Format

- A multi-vendor format enabling interoperability



Exactly one XML document defining the content and requirements of the virtual appliance

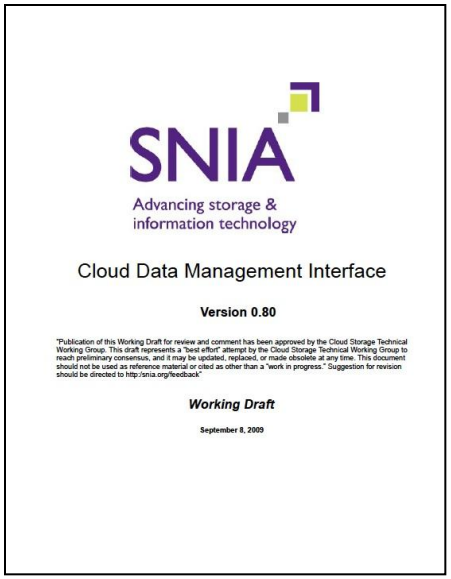
Optional SHA-1 digest for package data integrity

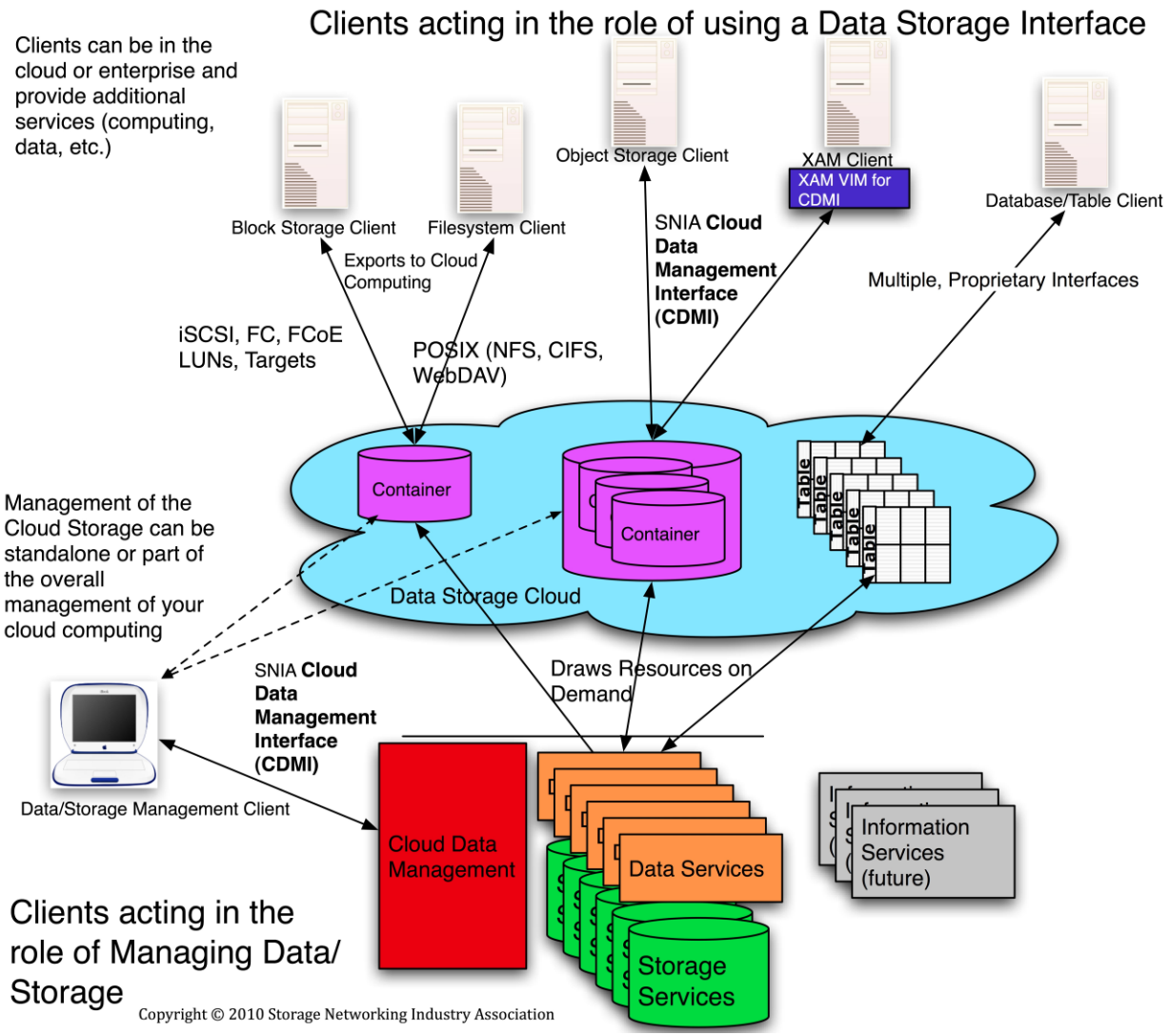
Optional certificate for package authenticity

Zero or more disk image files representing virtual disks for the virtual appliance

SNIA Cloud Data Mgmt Interface

Manages the provisioning of block-oriented, file-oriented & object-oriented storage



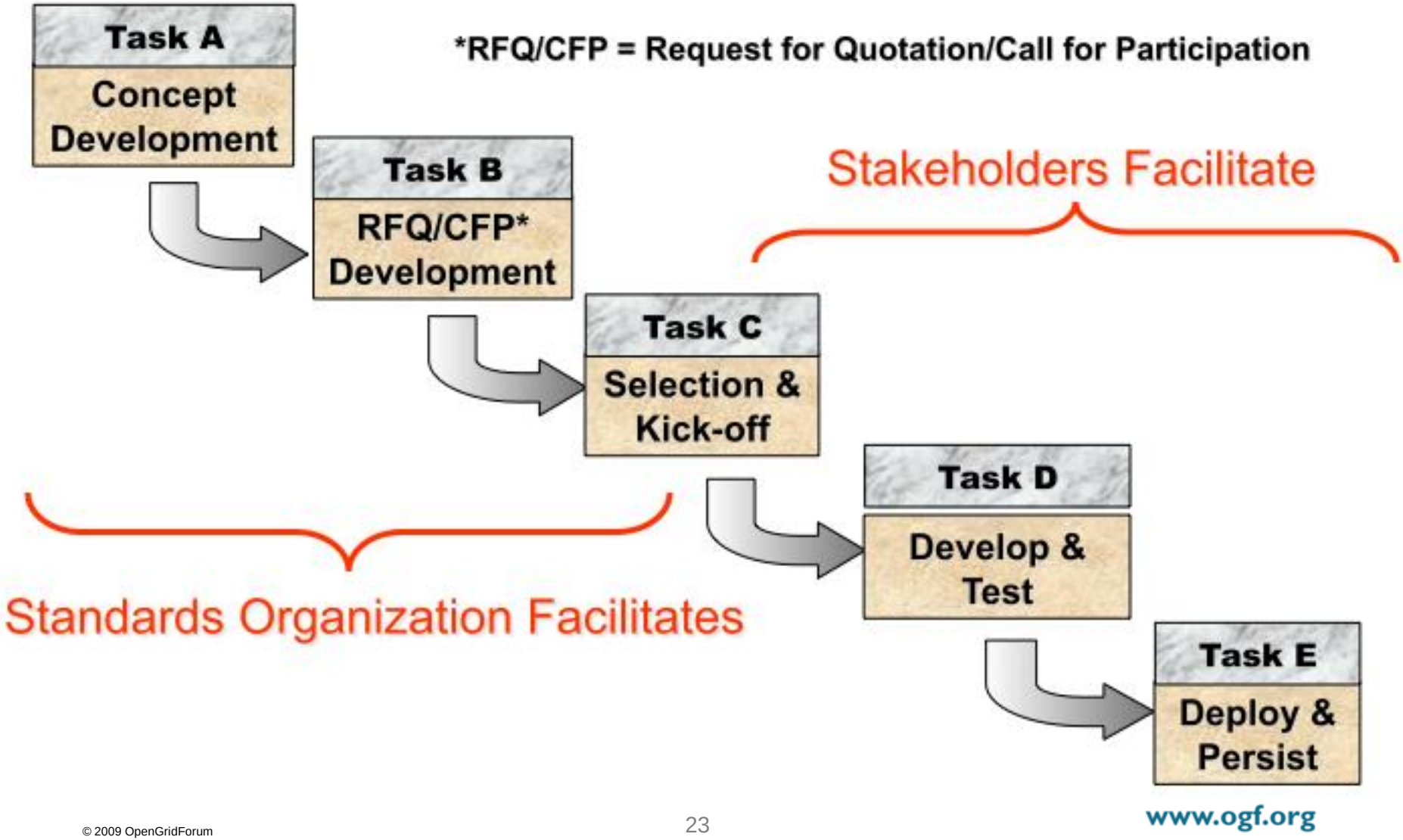


Used by permission.

A More Formal Process

- Drive ***incremental progress*** through near-term projects
- Build Critical Mass of Key Stakeholders
  - Continual polling and coordination across the community
- Forge agreement on:
  - Clear Goals
  - Clear Schedule (“time-box” the process)
  - Clear Responsibilities
  - Proper Provisioning of the Effort
- Major Stakeholders Must Contribute:
  - Time, Money & ***People***
  - Technical staff must engage to do the real technical work
  - The more people "invested", the more your agenda addressed!
- Deliver ROI to the Stakeholders
  - ***Targeted Projects on Key Issues***

One Iteration of this Process



Summary & Take-Home Message

- Argued that bottoms-up approach of evolving from private to public clouds acknowledges and leverages how the real-world works
- Draft roadmap “dartboard” presented
  - Deployment, Development /Risk Mitigation, Research
  - Coordinated Short List of Capabilities
- Reviewed status of OCCI, OVF, CDMI
  - Series of demonstrations possible to harmonize/shake-out existing/developing standards
- Discussed ways to *drive progress*
  - *Many roadmapping efforts: NIST, SIENA, GICTF*
  - *We must coordinate and collaborate*

Thank you
Questions?

lee@aero.org

Upcoming Event: OGF-30

- Brussels, October 25-29, 2010
- SIENA Roadmap Meeting
- ACM/IEEE Grid 2010 conference

www.ogf.org