INCOSE International Symposium
Philadelphia, PA
June 24-27, 2013

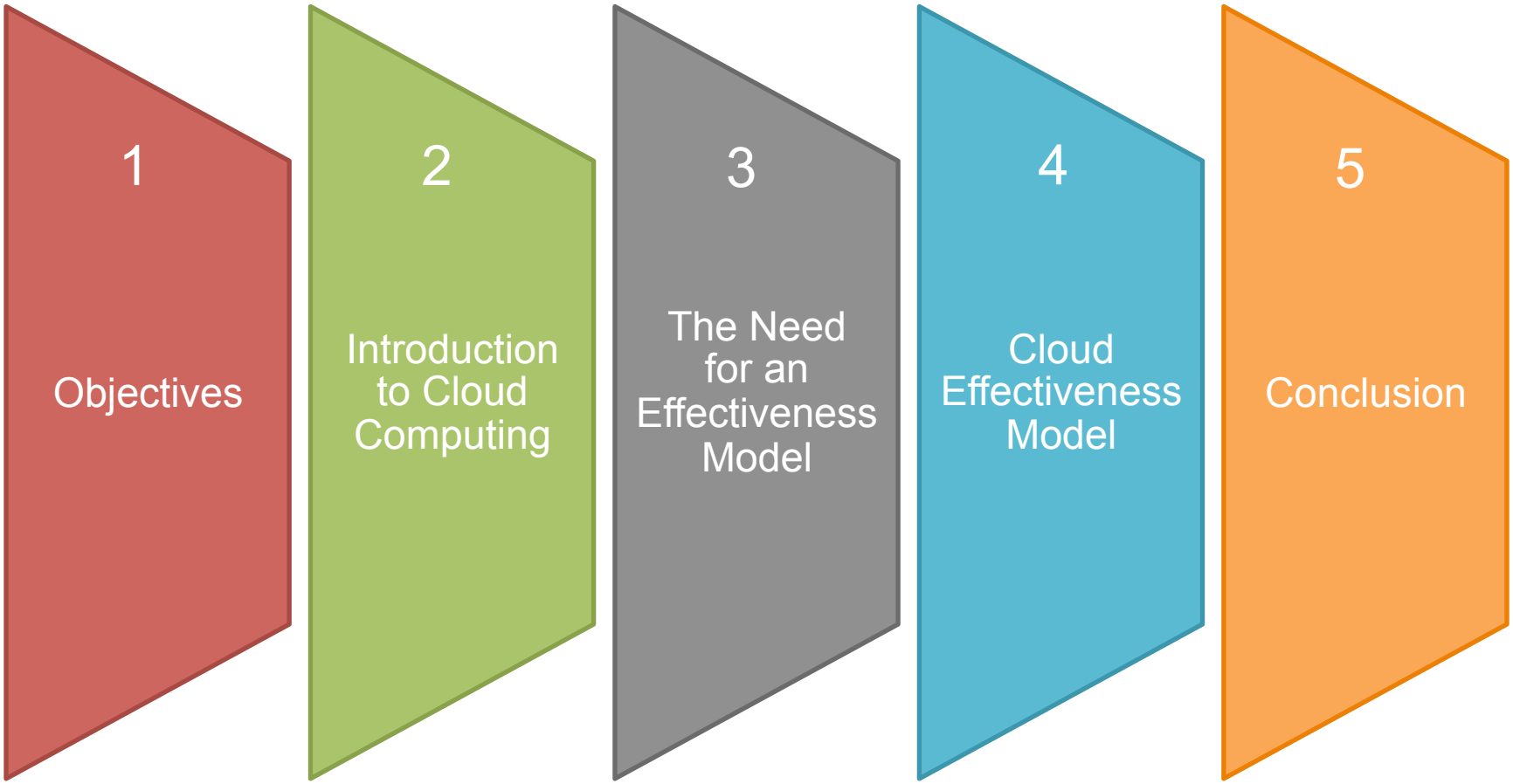
Cloud Effectiveness Model

Rich Feehs

The SI Organization, Inc.

rich.feehs@thesiorg.com

Agenda

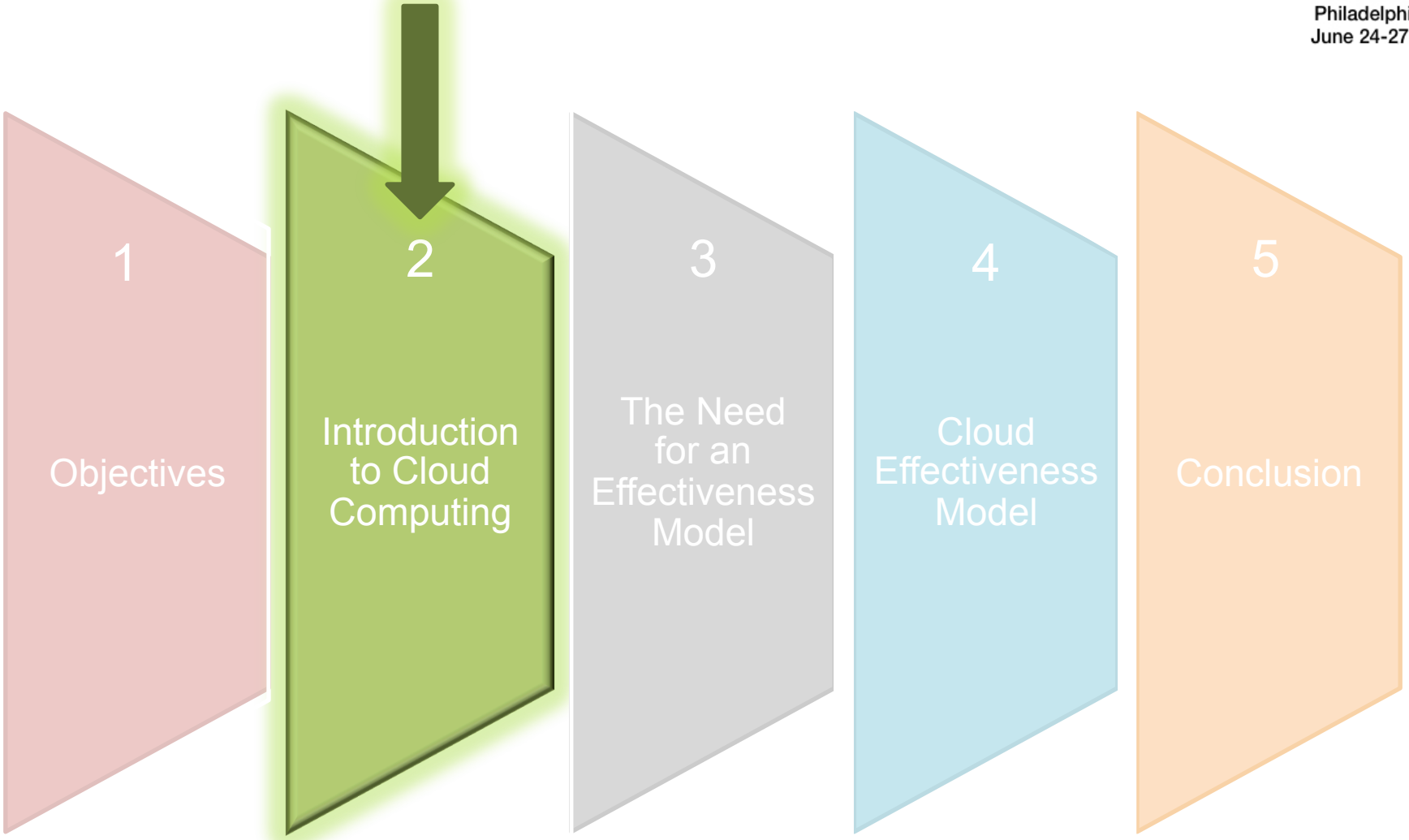

1
Objectives
2
Introduction to Cloud Computing
3
The Need for an Effectiveness Model
4
Cloud Effectiveness Model
5
Conclusion

Objectives

For the audience to understand:

1. **The NIST definition for Cloud Computing**
   including the "Essential Characteristics"
2. **Why an Effectiveness Model is needed**
3. **The Cloud Effectiveness Model**

Introduction to Cloud Computing



Introduction to Cloud Computing

Cloud Myths/Misconceptions:

- It's just a business model
  - It's a fad
    - It's good
      - Virtualization = Cloud
    - It's a new idea
  - It's commonly understood

In 1961, John McCarthy stated the following at the MIT Centennial:

"If computers of the kind I have advocated become the computers of the future, then computing may someday be organized as a public utility just as the telephone system is a public utility... The computer utility could become the basis of a new and important industry."

You may not realize it, but you've probably already gone "to the cloud."

Cloud Computing Definition

The interesting thing about cloud computing is that we've redefined cloud computing to include everything that we already do. I can't think of anything that isn't cloud computing … Maybe I'm an idiot, but I have no idea what anyone is talking about. What is it? It's complete gibberish. It's insane. When is this idiocy going to stop?

– Oracle CEO Larry Ellison, Sept. 2008

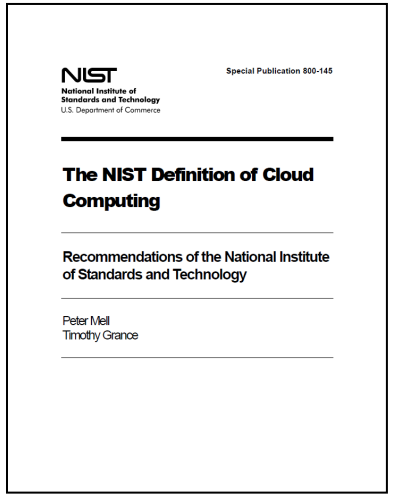
(The Problem)

Not having an agreed upon definition for "the cloud" is bad for consumers and good for vendors/providers

Cloud Computing Definition

Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources that can be rapidly provisioned and released with minimal management effort or service provider interaction.

(The Solution)

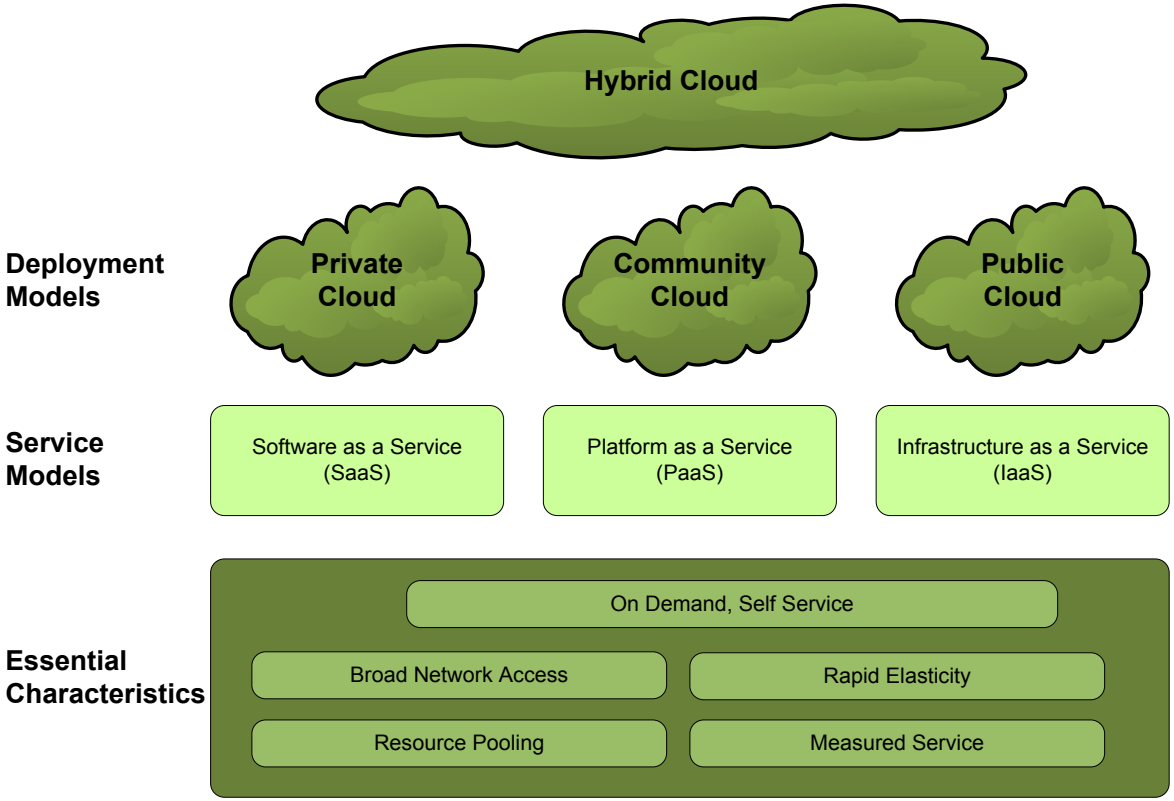

NIST
National Institute of Standards and Technology
U.S. Department of Commerce

Special Publication 800-145

The NIST Definition of Cloud Computing

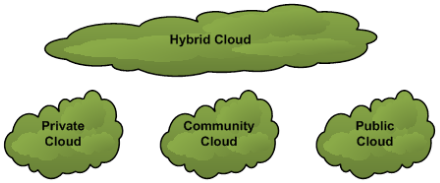
Recommendations of the National Institute of Standards and Technology

Peter Mell
Timothy Grance



The NIST definition for Cloud Computing is the de facto standard across the Federal Government and much of Industry

Deployment Model

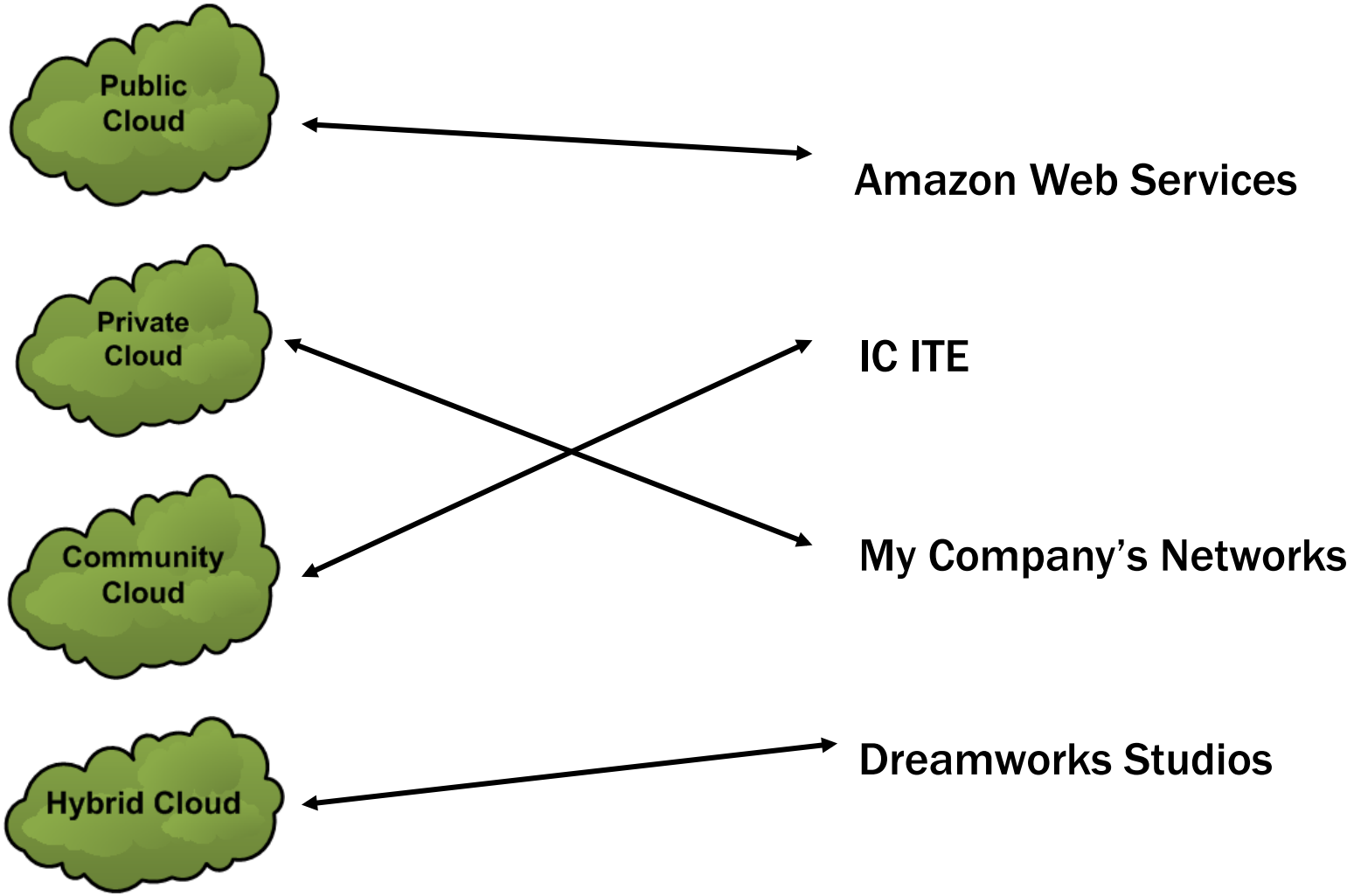


Model	Managed By	Owned By	Located	Accessible and Used By
Private	Organization 3rd Party	Organization 3rd Party	On-Premise or Off-Premise	Trusted
Public	3rd Party	3rd Party	Off-Premise	Untrusted
Community	Community Member 3rd Party	Community Member 3rd Party	On-Premise or Off-Premise	Trusted
Hybrid	A combination of the above			

The deployment model adopted by an organization is typically driven by cost and security

Pop Quiz #1

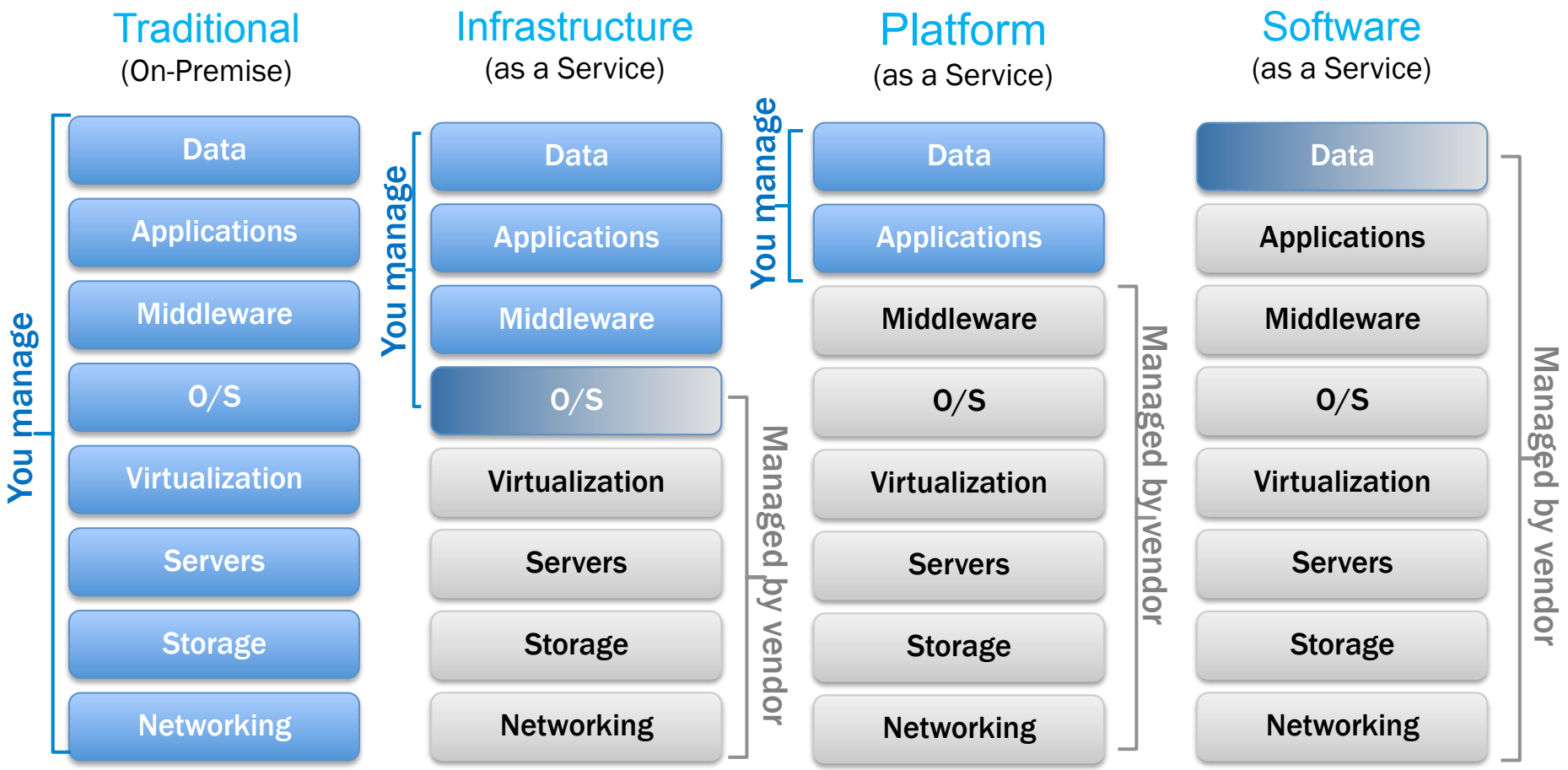
Match the following examples to their respective Deployment Model



Service Models

Software as a Service (SaaS) | Platform as a Service (PaaS) | Infrastructure as a Service (IaaS)

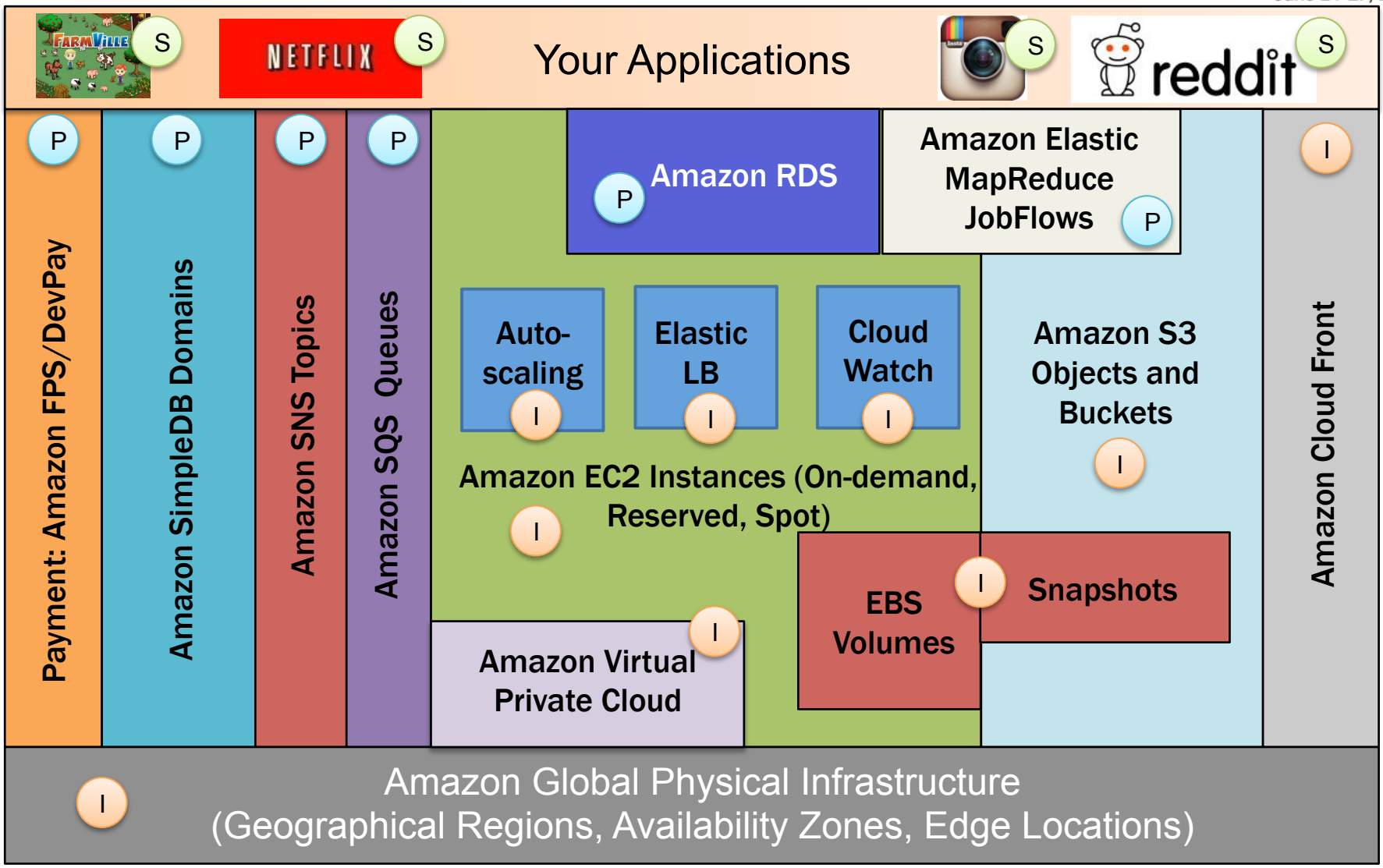
INCOSE International Symposium
Philadelphia, PA
June 24-27, 2013



Each Service Model incrementally increases the provider's control of higher layers of the stack and comes with a new set of APIs to enable automated access to these services

Pop Quiz #2

Software as a Service (SaaS) | Platform as a Service (PaaS) | Infrastructure as a Service (IaaS)



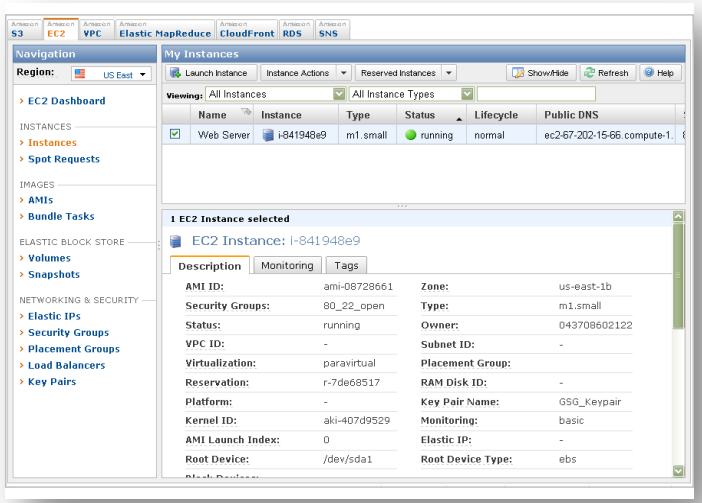
Which of these services are: I - IaaS P - PaaS S - SaaS

Essential Characteristics

INCOSE International Symposium
Philadelphia, PA
June 24-27, 2013

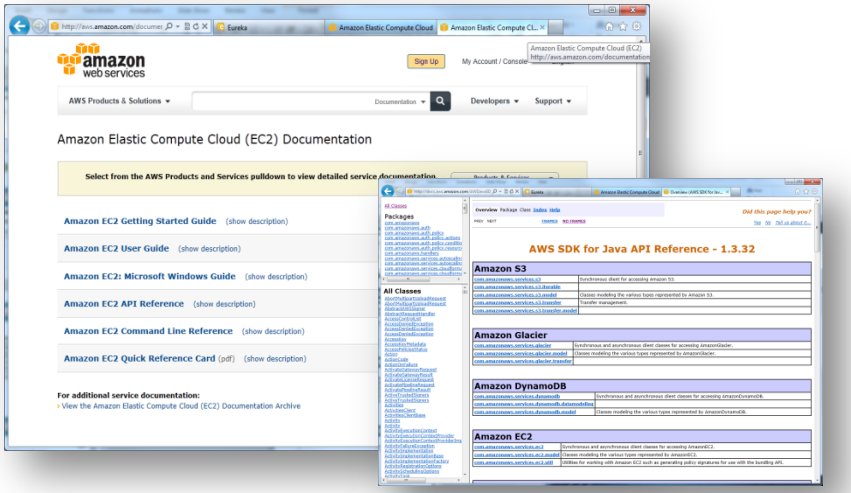
On-demand self service

A consumer can unilaterally provision computing capabilities, such as server time and network storage, as needed automatically without requiring human interaction with each service's provider.



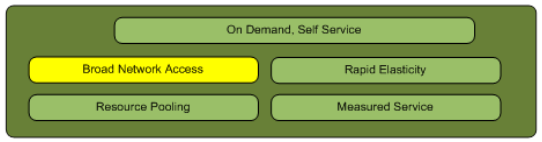
Web-based management console is provided to allow the consumer to request and manage Standard services

Web-Service interfaces and SDKs are provided to allow applications to request Standard services at runtime



Automated Provisioning is the enabler for "Self Service" and "Elasticity"

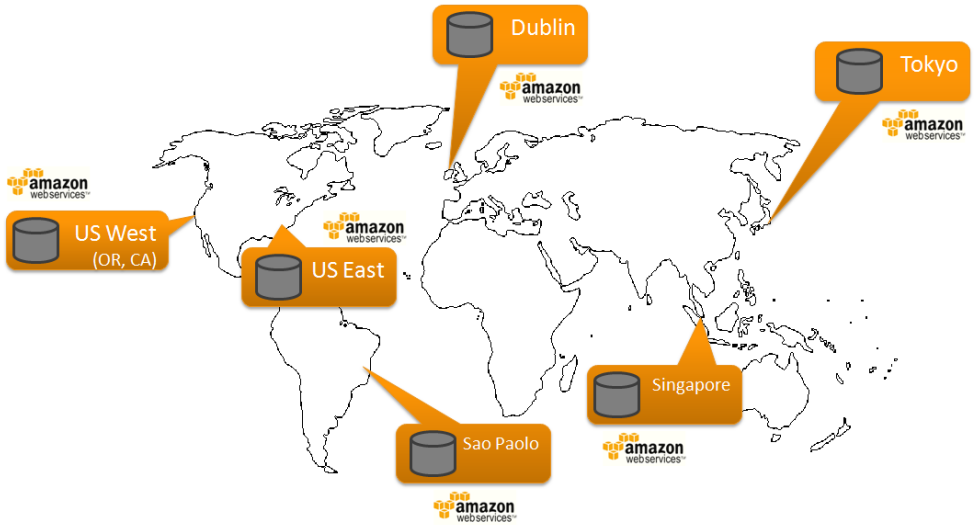
Essential Characteristics



Broad network access

Capabilities are available over the network and accessed through standard mechanisms that promote use by heterogeneous thin or thick client platforms (e.g., mobile phones, laptops, and PDAs).

Several techniques are used to ensure customers have access to data and services:



- **Multiple, regional data centers**
- **Multiple "availability zones" in a given region**
- **Network provider redundancy and diversity**
- **Forward caching**

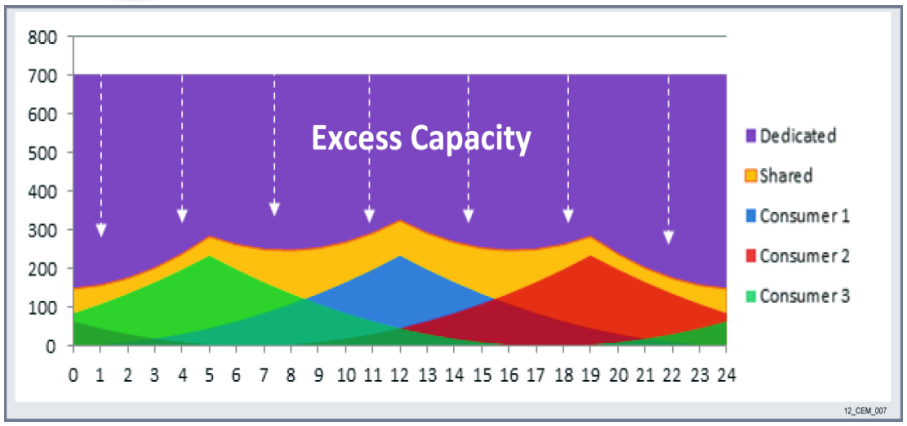
If your customer can't reach your cloud, you won't have a customer for long

Essential Characteristics

INCOSE International Symposium
Philadelphia, PA
June 24-27, 2013

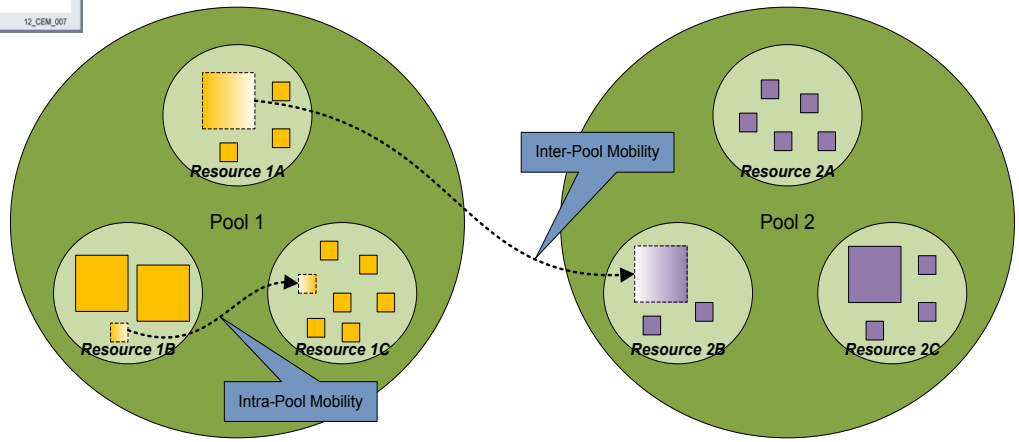
Resource Pooling

The provider's computing resources are pooled to serve multiple consumers using a multi-tenant model, with different physical and virtual resources dynamically assigned and reassigned according to consumer demand.



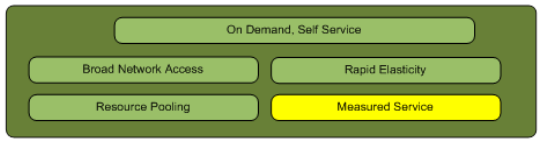
Resource pooling allows for a reduction in the total resources required to support many customers (when compared to dedicated resources)

Resource pooling enables a simpler approach to High Assurance (HA) and Disaster Recovery (DR)



Resource Pooling is key to most provider's business model

Essential Characteristics



Measured Service

Cloud systems automatically control and optimize resource use by leveraging a metering capability at some level of abstraction appropriate to the type of service. Resource usage can be monitored, controlled, and reported.

Why is metering needed?

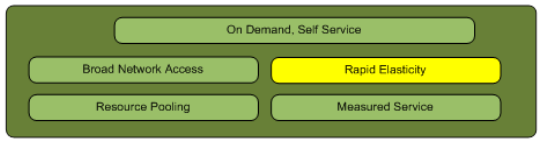
- Providers occasionally need it to charge customers
- Providers need it to manage their capacity & associated costs
- Consumers need it so that they know what they are paying for

Metering Approaches

1. Fixed Rate (e.g., Netflix)
2. Time Provisioned (e.g., AWS EC2)
3. Actual usage (e.g., Sun Cloud)

The "IDEAL" … plug in and pay for only what you use (similar to other utilities)

Essential Characteristics

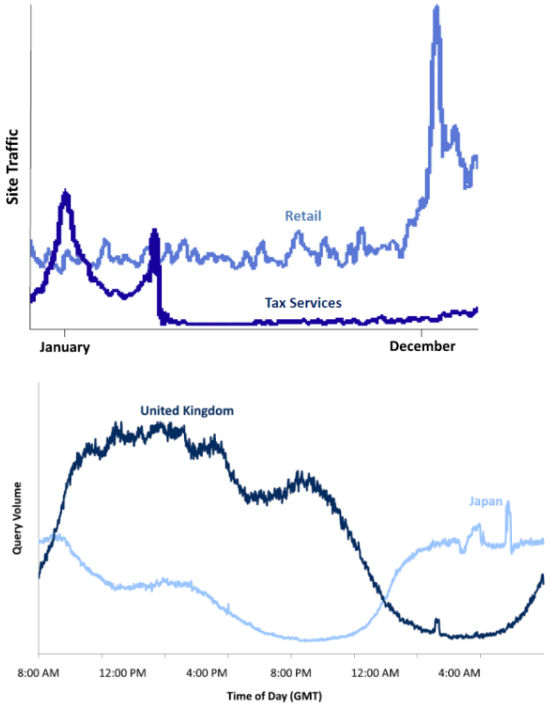


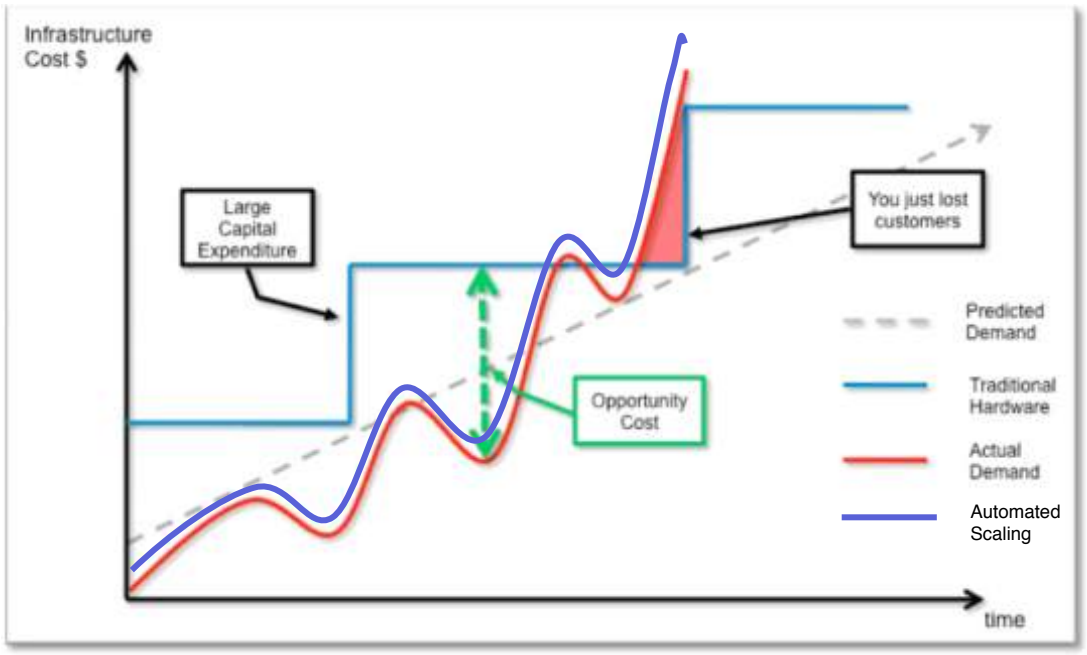
INCOSE International Symposium
Philadelphia, PA
June 24-27, 2013

Rapid Elasticity

Capabilities can be rapidly and elastically provisioned, in some cases automatically, to quickly scale out and rapidly released to quickly scale in. To the consumer, the capabilities available for provisioning often appear to be unlimited and can be purchased in any quantity at any time.

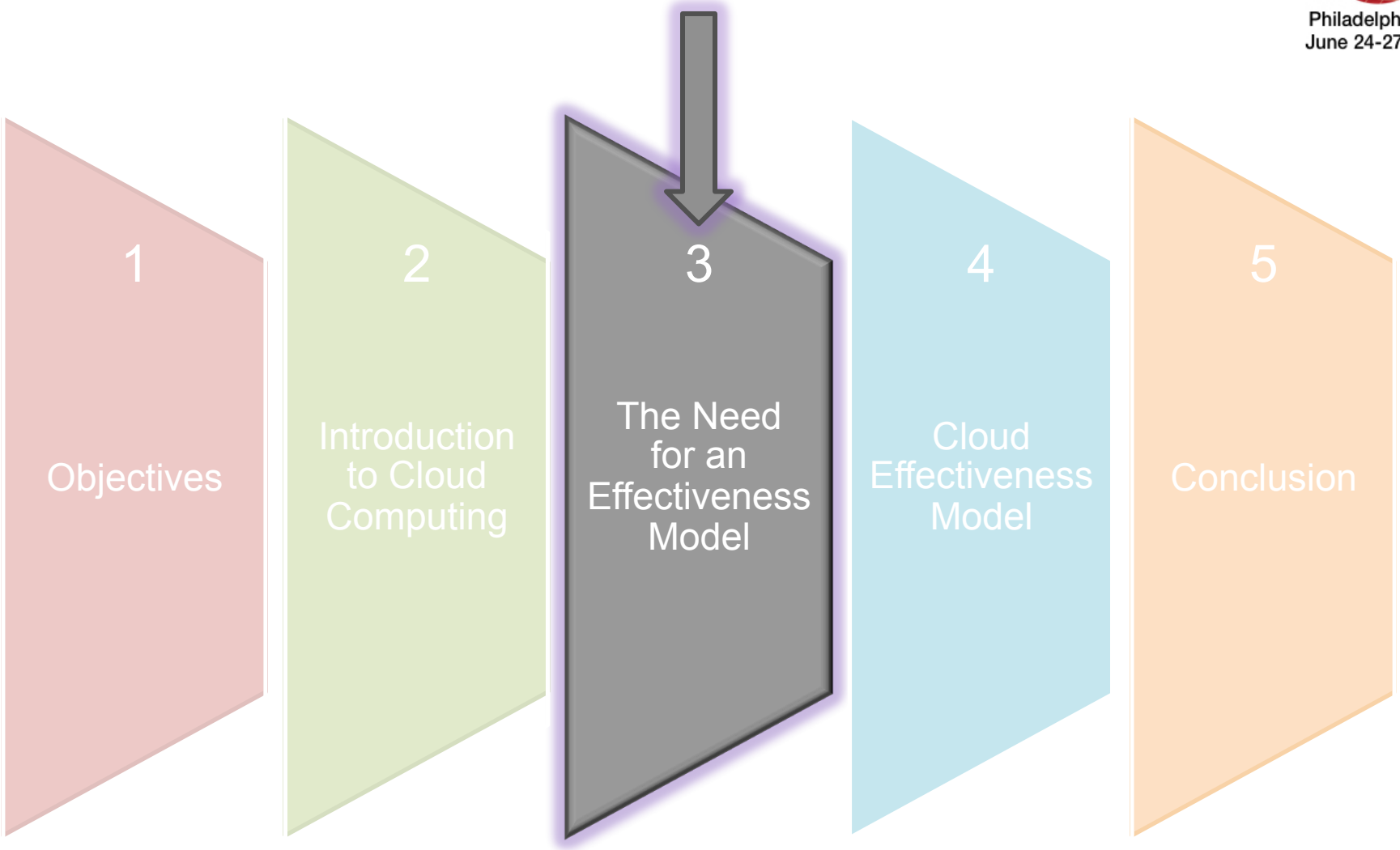
Why is elasticity needed, and how does it help?



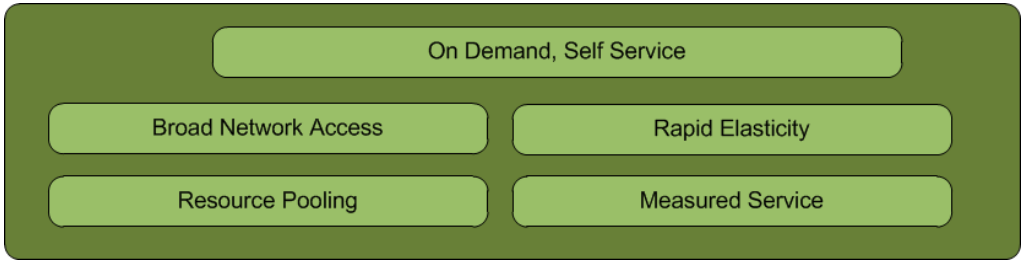


Without elasticity, you must plan/provision/pay for your peak (all the time)

The Need for an Effectiveness Model



The Need for an Effectiveness Model



The Essential Characteristics are what make Cloud Computing different.

---

es·sen·tial | iˈsenSHəl |

noun
something absolutely necessary.

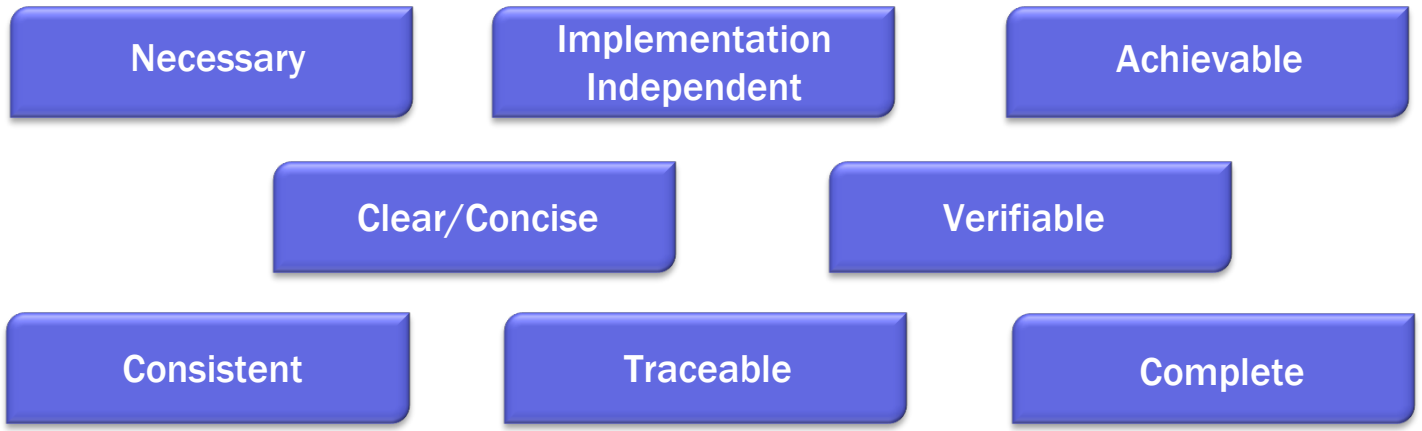
adjective
absolutely necessary; extremely important

synonyms
fundamental. substantial. basic. main. material.

System Engineers use the words “required” or “requirement”

The Need for an Effectiveness Model

But what are the characteristics of a "good" requirement?

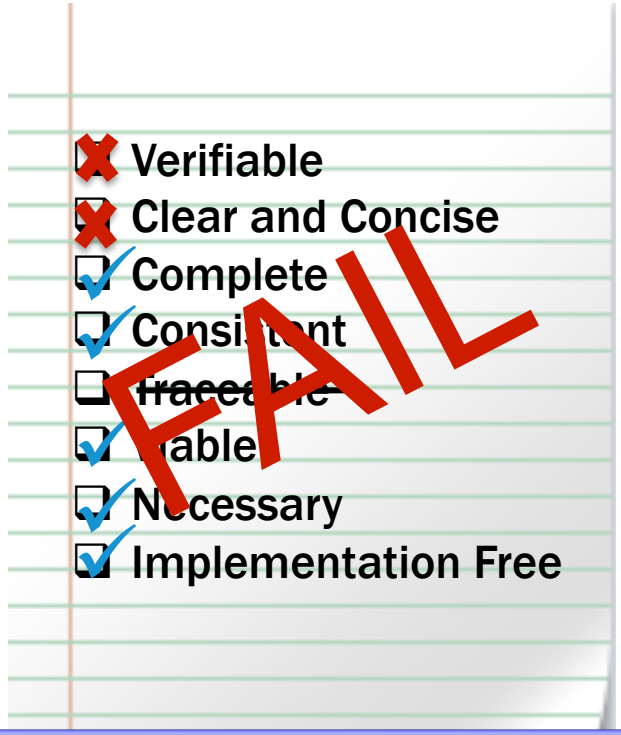


Pop Quiz #3

Rapid Elasticity

Capabilities can be rapidly and elastically provisioned, in some cases automatically, to quickly scale out and rapidly released to quickly scale in. To the consumer, the capabilities available for provisioning often appear to be unlimited and can be purchased in any quantity, at any time.

Is this…



REASON #1: We cannot confidently use what we can't measure. "Essential" isn't measurable.

The Need for an Effectiveness Model

INCOSE International Symposium
Philadelphia, PA
June 24-27, 2013

As System Engineers,

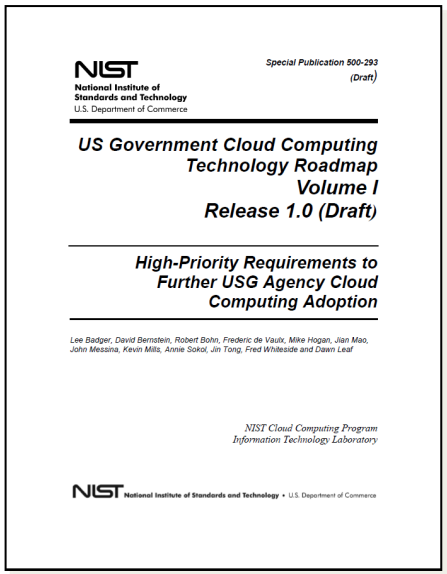
we (and/or our customers) may be asking…

- How important are each of the essential characteristics to me? Which ones really matter?
- Do I need to migrate to a cloud architecture? If so, are these the right requirements for my mission objective?
- Where should I focus my resources to improve my cloud offering?
- What improvements would my customers value?
- Have I reached the desired end-state?

REASON #2: They may be "essential", but they are not all valued equally (by providers or customers).

The Need for an Effectiveness Model

The NIST definition is necessary (and good), but insufficient



NIST
National Institute of Standards and Technology
U.S. Department of Commerce

Special Publication 500-293 (Draft)

US Government Cloud Computing Technology Roadmap Volume I Release 1.0 (Draft)

High-Priority Requirements to Further USG Agency Cloud Computing Adoption

Lee Badger, David Bernstein, Robert Bohn, Frederic de Vaulx, Mike Hogan, Jian Mao, John Messina, Kevin Mills, Annie Sokol, Jin Tong, Fred Whiteside and Dawn Leaf

NIST Cloud Computing Program
Information Technology Laboratory

NIST National Institute of Standards and Technology • U.S. Department of Commerce

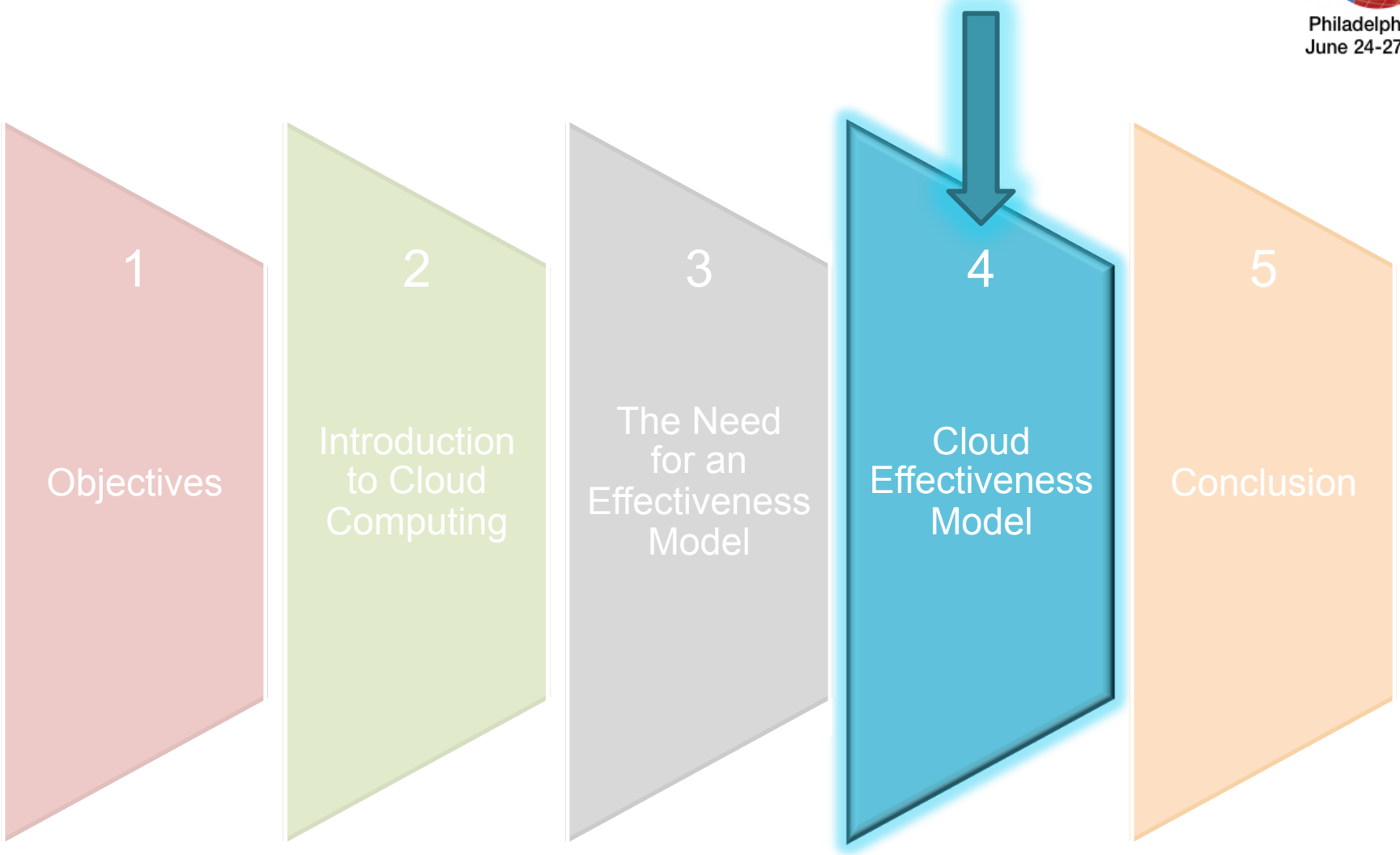
Requirement 10: Defined & implemented Cloud Service Metrics
Industry needs to establish Cloud Service Metrics, including Standardized Units of Measurement for Cloud Resources.

"In contrast to the precision with which we categorize units of measurement in electricity, light, or fuels, cloud computing measurements are relatively imprecise… consumers cannot determine and request cloud services as a utility with a high degree of predictability, and cannot achieve maximum cost-effectiveness in cloud computing service application."[1]

REASON #3: The definitions aren't precise and metrics are

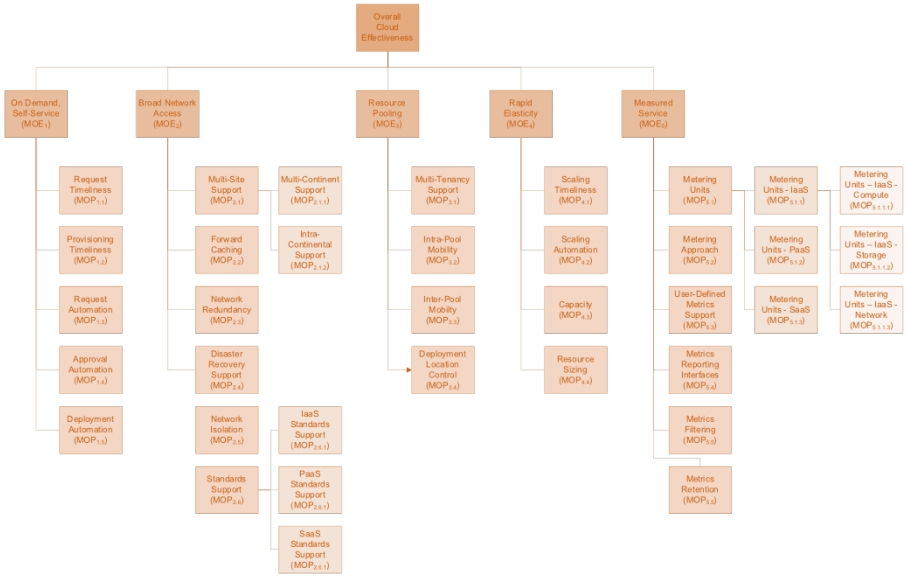
1 - NIST Special Publication 500-293, US Government Cloud Computing Technology Roadmap, Volume I, Release 1.0 (Draft) November 2011

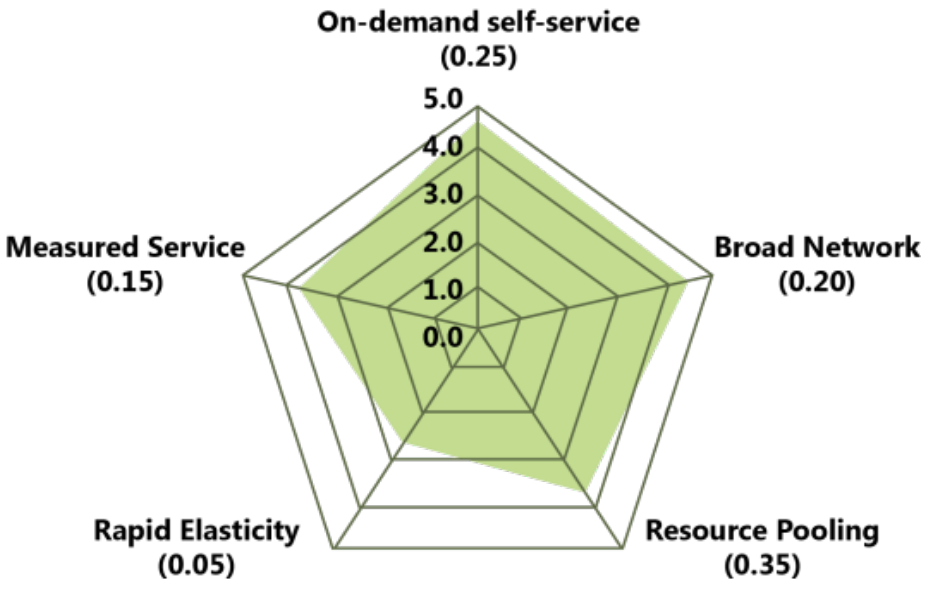
Cloud Effectiveness Model



What is the Effectiveness Model?

INCOSE International Symposium
Philadelphia, PA
June 24-27, 2013





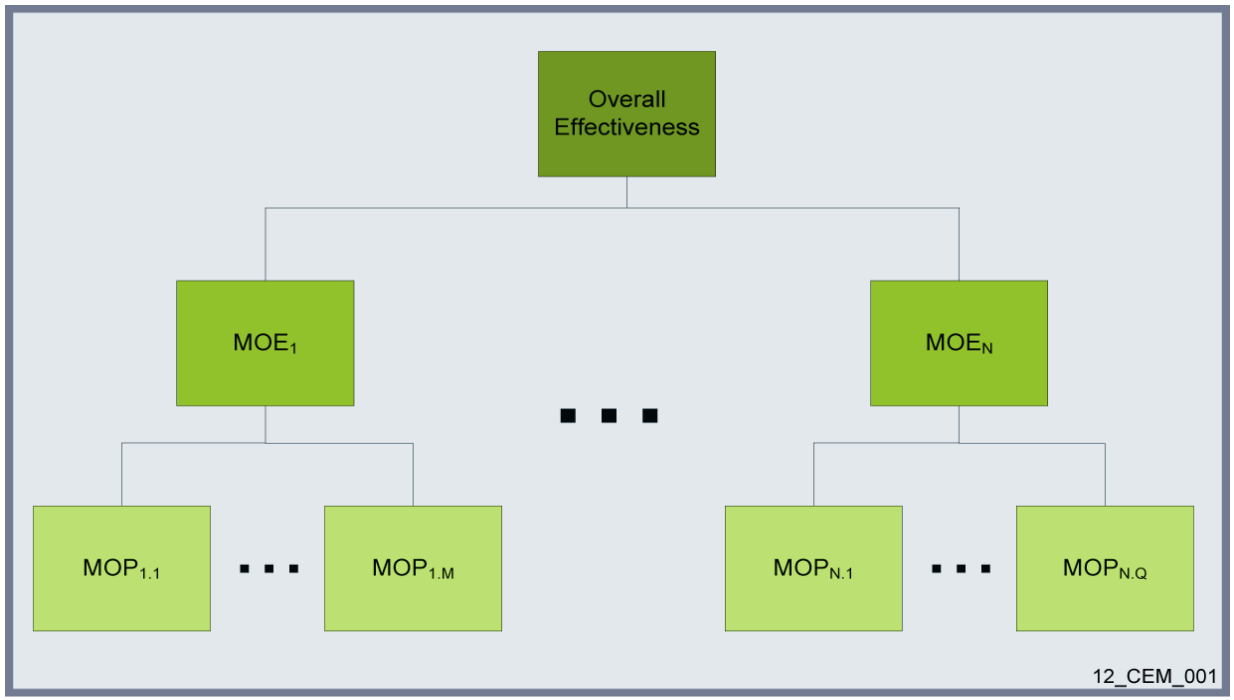
The Cloud Effectiveness Model

- A set of Measures of Effectiveness (MOEs) and associated Measures of Performance (MOPs)
- Adopts the NIST “Essential Characteristics of cloud computing” as the top level MOEs because they reflect the “operational” measures of success
- Tailored to reflect a customer’s priorities and summarized in an intuitive radar chart

What is the Effectiveness Model?

The Framework (i.e., the math)

$$Overall\ Effectivness = \sum_{x=1}^{N} (\alpha_x * MOE_x)$$



$$MOE_x = \sum_{y=1}^{N} (\alpha_{x.y} * MOP_{x.y})$$

$$\sum_{y=1}^{N} \alpha_{x,y} + \alpha_{x,y+1} + \ldots + \alpha_{x,}$$

where

The weights are tailored based on the customer's priorities

What is the Effectiveness Model?

INCOSE International Symposium
Philadelphia, PA
June 24-27, 2013

Process to develop the MOPs

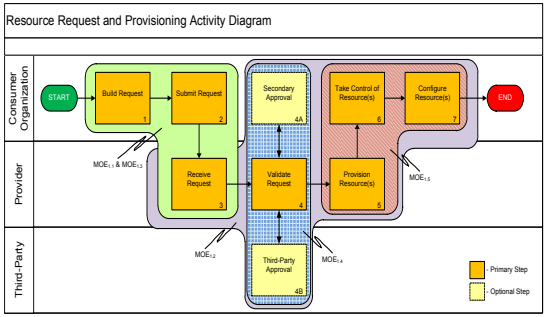
Step 1: Extract the important objectives from the definition

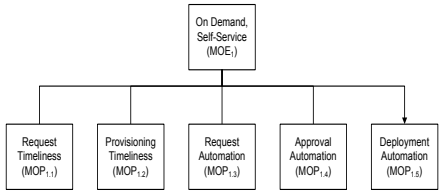
Step 2: Think of quantifiable measures that relate to each objective. They become MOPs

Step 3: Determine the scale for characterizing the performance of each measure

On-Demand, Self-Service

A consumer can unilaterally provision computing capabilities, such as server time and network storage, as needed automatically without requiring human interaction with each service's provider.



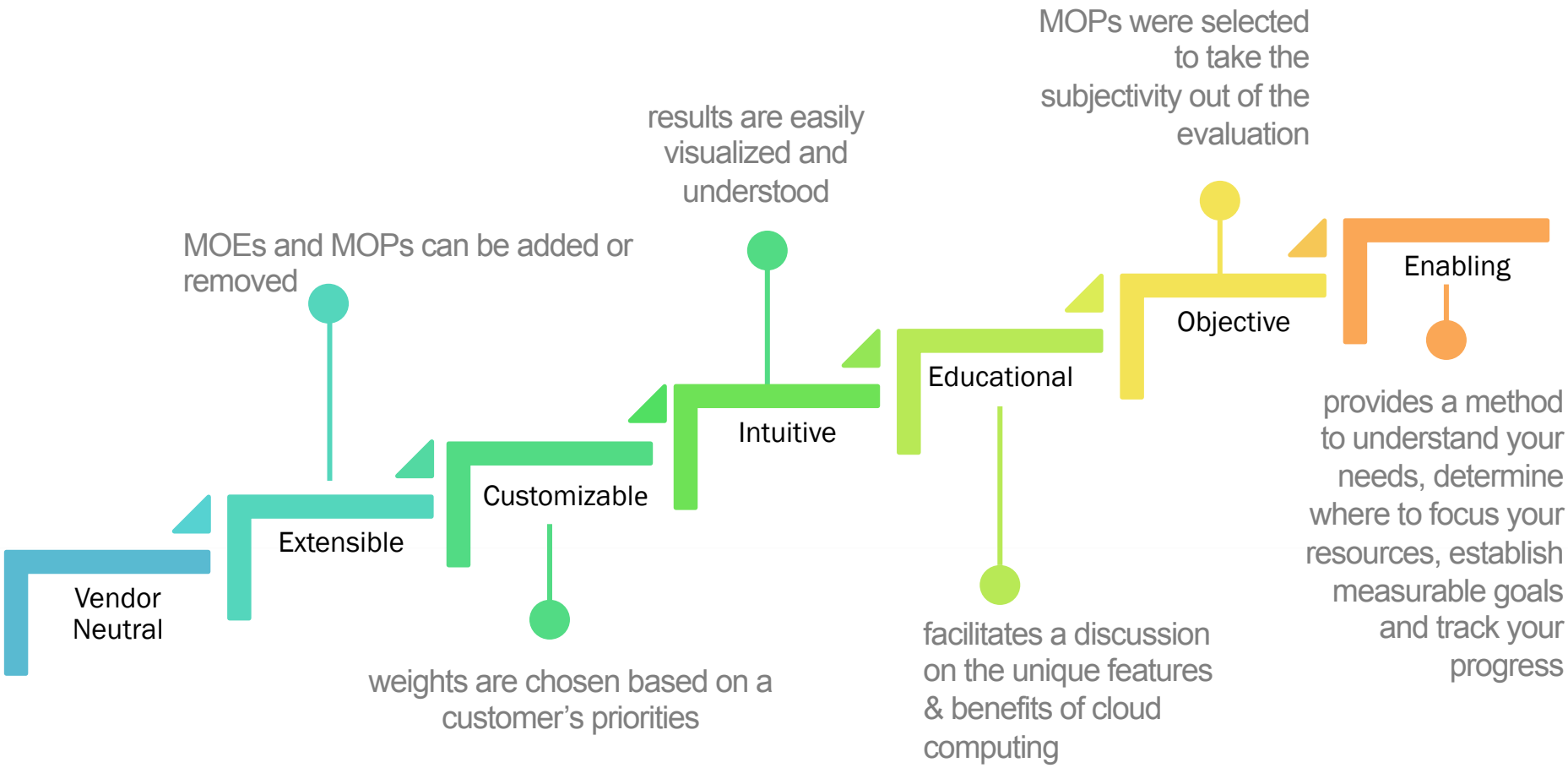


MOP	Weight	Score	Performance Characteristics
Request Timeliness $MOP_{1,1}$	$\alpha_{1,1}$	5	Minutes
		4	Hours
		3	Days
		2	Weeks
		1	Months
Provisioning Timeliness $MOP_{1,2}$	$\alpha_{1,2}$	5	Minutes
		4	Hours
		3	Days
		2	Weeks
		1	Months
Deployment Automation $MOP_{1,5}$	$\alpha_{1,5}$	5	Automated provisioning, no manual configuration required
		4	Automated provisioning, manual configuration by user
		3	Automated provisioning, manual configuration by provider
		2	Manual provisioning, manual configuration by user
		1	Manual provisioning, manual configuration by provider

What is the Effectiveness Model?

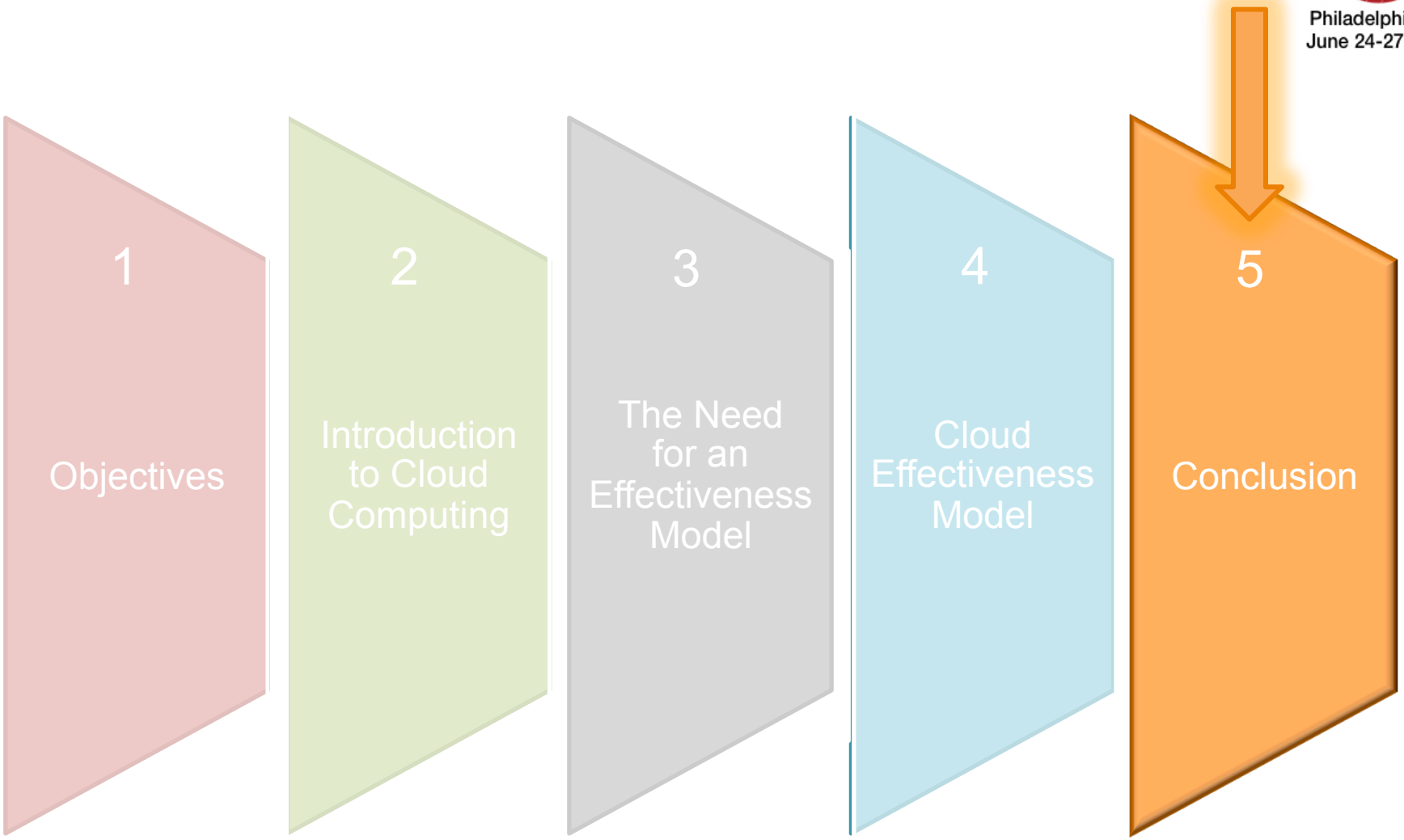
INCOSE International Symposium
Philadelphia, PA
June 24-27, 2013

The Cloud Effectiveness Model provides a framework that is:



The Cloud Effectiveness Model can help you know when and why you need the cloud

Conclusion


1
Objectives
2
Introduction to Cloud Computing
3
The Need for an Effectiveness Model
4
Cloud Effectiveness Model
5
Conclusion

Conclusion

You should understand:

1. **The NIST definition for Cloud Computing**
   Including the "Essential Characteristics"
2. **Why an Effectiveness Model is needed**
3. **The Cloud Effectiveness Model**

next steps:

1. **continue to** refine the model (**MOEs and MOPs**)
2. engage cloud providers **to characterize their "effectiveness"**

Questions?

INCOSE International Symposium
Philadelphia, PA
June 24-27, 2013

cloudexperts@thesiorg.com