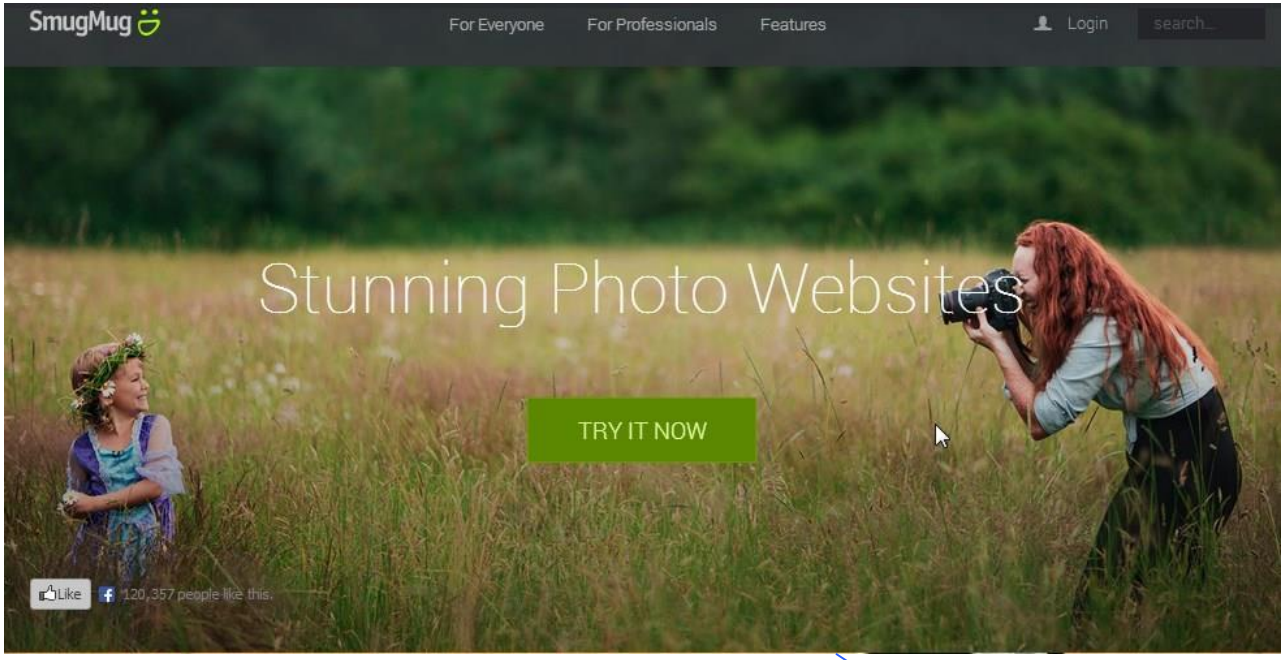
Service Level Agreements

CMT308

Omer Rana

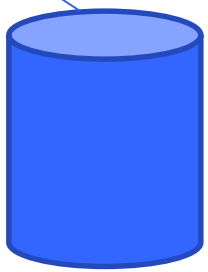
Service Level Agreements



Smugmug.com

Service Level Agreement

Increasingly, companies making use of 3rd party services



Amazon.com S3 service

SLAs

- Service Level Agreements (SLAs) and Quality of Service are closely related
- SLA are used to:
  - *Define pre-agreed quality terms between a service user and (one or more) provider(s)*
  - *Identify constraints on quality that a provider must observe*
  - *Identify requirements that a user/client has on the provider*
- SLAs are particular useful:
  - *In the context of service outsourcing*
  - *Important to be able to determine what is requested of a provider*
  - *Important to be able to decide whether a particular requested quality has been met*
- SLA from ITIL (IT Infrastructure Library for IT Service Manag.)
  - *Service-based*
  - *Customer-based*
  - *Hierarchical*

Amazon

- **simple storage service (s3)**
  - *Storage for the Internet*
  - *Simple Web Services interface to store and retrieve data*
  - *REST and SOAP Web Service interfaces*
- **elastic compute cloud (ec2)**
  - *Compute on demand*
  - *Virtualization*
  - *Integration with S3*

Amazon … SLA

- **simple storage service (s3)**
  - *http://aws.amazon.com/s3-sla/*
  - *Error Rate*
  - *Monthly Uptime Percentage*
  - *Service Credit*
- **elastic compute cloud (ec2)**
  - *http://aws.amazon.com/ec2-sla/*
  - *Monthly Uptime Percentage*
  - *Region Availability*
  - *Service Credit*

Salesforce… SLA

- http://trust.salesforce.com/

- Cloud provider based on a “multi-tenancy” environment
- Hosts capability from multiple customers on the same physical infrastructure
- Users can check system status:
  - *http://trust.salesforce.com/trust/status/*

PRODUCT

Health Cloud

INSTANCES | MAINTENANCES

10/15/19 - 12/31/20 • 1588 Items

Expand All | Collapse All

Quick Find

ID	DATE	START TIME	SUBJECT	INSTANCES	SERVICES	TYPE
PAST 33 DAYS (993)						
49496	Sep 14	1:00 pm BST	Database Maintenance	DB1	Marketing Cloud Core Service, Marketing Cloud Login, Marketing Cloud REST API, Marketing Cloud SOAP API	Scheduled Maintenance
49100	Sep 16	10:30 pm BST	Chat (formerly known as Live Agent) / Omni-Channel Winter '20 Release Preparation Notice	UM4	Live Agent	Release
49099	Sep 16	10:30 pm BST	Chat (formerly known as Live Agent) / Omni-Channel Winter '20 Release Preparation Notice	UM5	Live Agent	Release
49108	Sep 17	12:30 am BST	Chat (formerly known as Live Agent) / Omni-Channel Winter '20 Release Preparation Notice	EU19	Live Agent	Release
49104	Sep 17	12:30 am BST	Chat (formerly known as Live Agent) / Omni-Channel Winter '20 Release Preparation Notice	EU30	Live Agent	Release
49107	Sep 17	12:30 am BST	Chat (formerly known as Live Agent) / Omni-Channel Winter '20 Release Preparation Notice	EU25	Live Agent	Release

INSTANCE

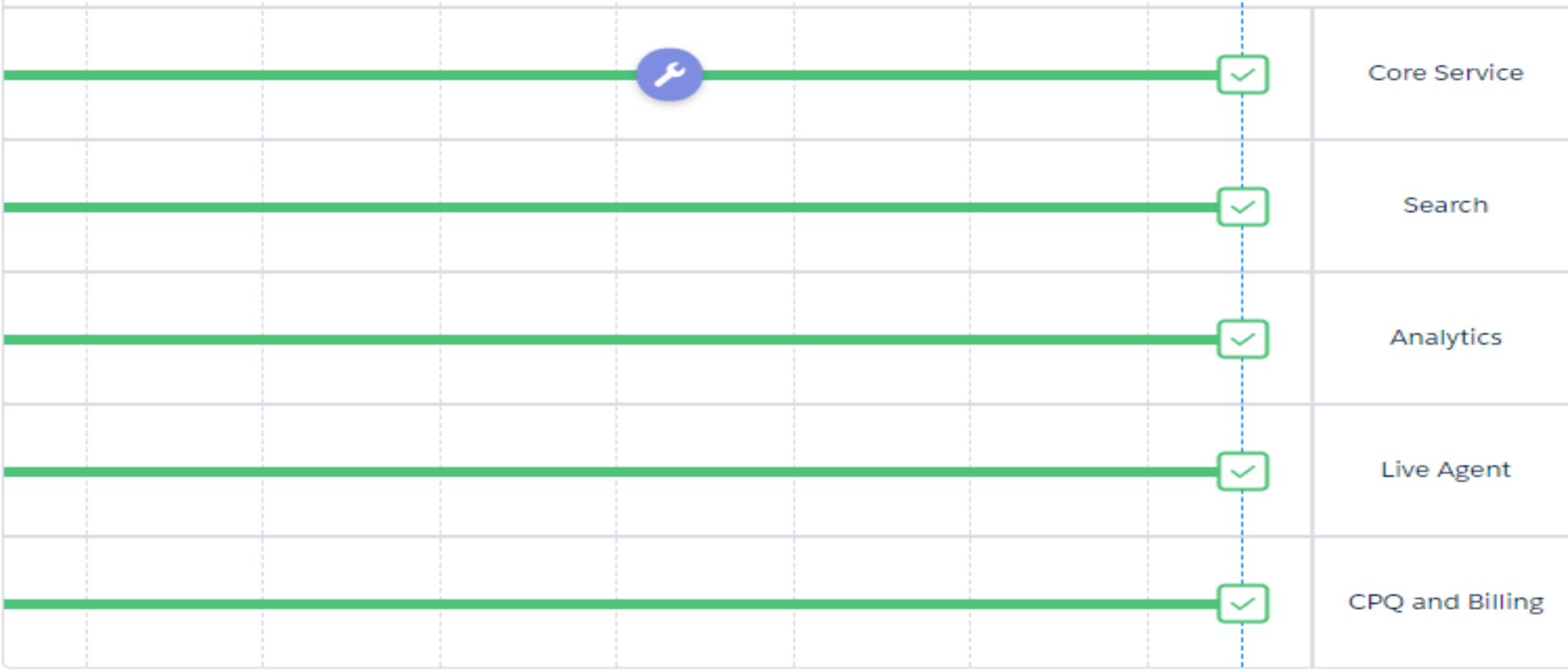
NA90

CURRENT STATUS | HISTORY | MAINTENANCES

7 Days | 3 Days | 24 Hours

Select a date | Now

9 | 10/10 | 10/11 | 10/12 | 10/13 | 10/14 | 10/15 | NOW



INSTANCE
NA90

RSS | Subscribe to Notifications

CURRENT STATUS | HISTORY | MAINTENANCES

10/15/19 - 12/31/20 • 8 Items

Expand All | Collapse All | Quick Find

ID	DATE	START TIME	SUBJECT	INSTANCES	SERVICES	TYPE
PAST 33 DAYS (6)						
49168	Sep 24	7:30 am BST	Chat (formerly known as Live Agent) / Omni-Channel Winter '20 Release Preparation Notice	NA90	Live Agent	Release
49286	Oct 12	6:30 am BST	CPQ Patch for Instanceless URL fixes for Production	NA90	CPQ and Billing	Scheduled Maintenance
48976	Oct 12	7:00 am BST	Financial Services Cloud - Winter '20 R2A Release	NA90		Release
28751	Oct 12	7:00 am BST	Winter '20 Major Release	NA90	Core Service	Release
48948	Oct 12	7:00 am BST	Health Cloud - Winter '20 R2A Release	NA90		Release
48119	Oct 12	7:30 am BST	QTC Winter '20 Major Release (222) - R2a	NA90	CPQ and Billing	Release
TODAY - OCTOBER (0)						

Instance Details

Version:
Winter '20 Patch 9.7

Region:
Americas

Products:
Sales Cloud, Service Cloud, Community Cloud, CPQ and Billing, Einstein Analytics, Financial Services Cloud, Health Cloud, Lightning Platform, LiveAgent / Omni-Channel

Winter '20 Major Release

Core Service | NA90

MAINTENANCE INFORMATION

ID# 28751

Status
Resolved

Impacted Instances
NA90

Impacted Services
Core Service

Type
Release

Availability
This instance will not be available during this maintenance window.

Planned Start Time:
7:00 am BST, Oct 12

End Time:
7:01 am BST, Oct 12
(Planned End Time: 7:05 am BST, Oct 12)

Duration
1 minute

MAINTENANCE HISTORY

FEATURES ENABLED

The upgrade activities are now complete and all major release features are available.

Posted 7:26 am BST, Oct 12

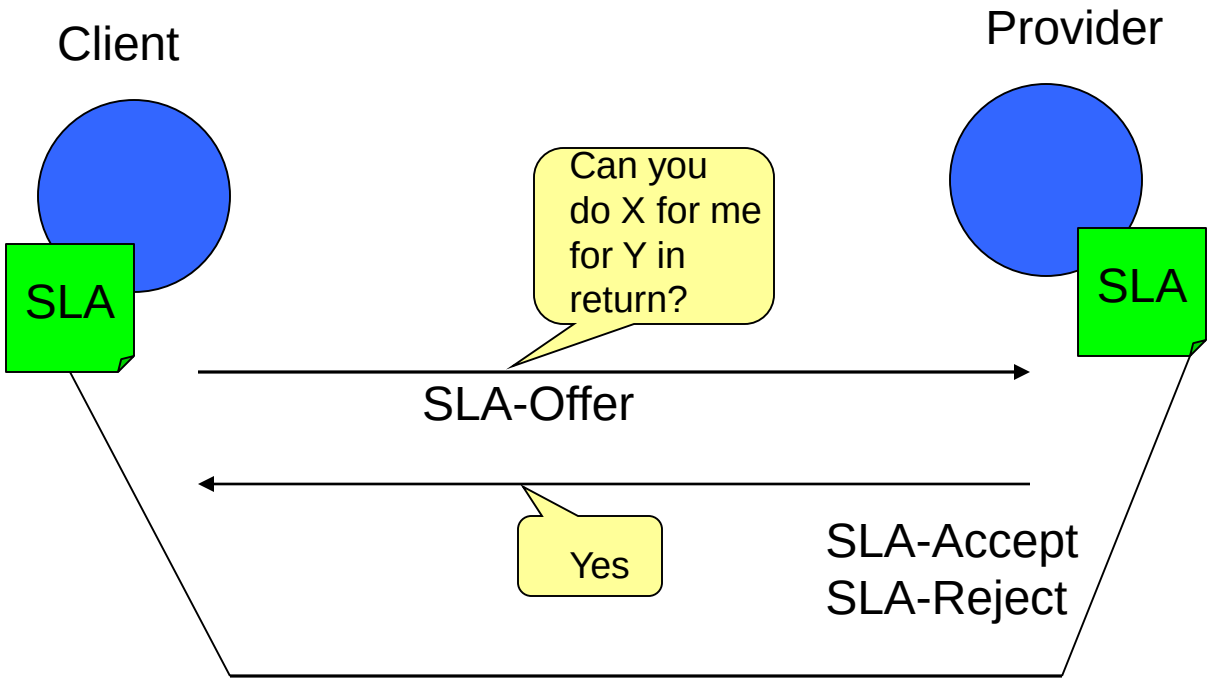
RELEASE IS LIVE

The release is now live. The instance should be generally available as we continue to perform upgrade activities including feature enablement, which typically completes within six hours and no later than 24 hours.

Posted 7:01 am BST, Oct 12

What is a Service Level Agreement (SLA) and how is it formed?

A relationship between a client and provider in the context of a particular capability (service) provision

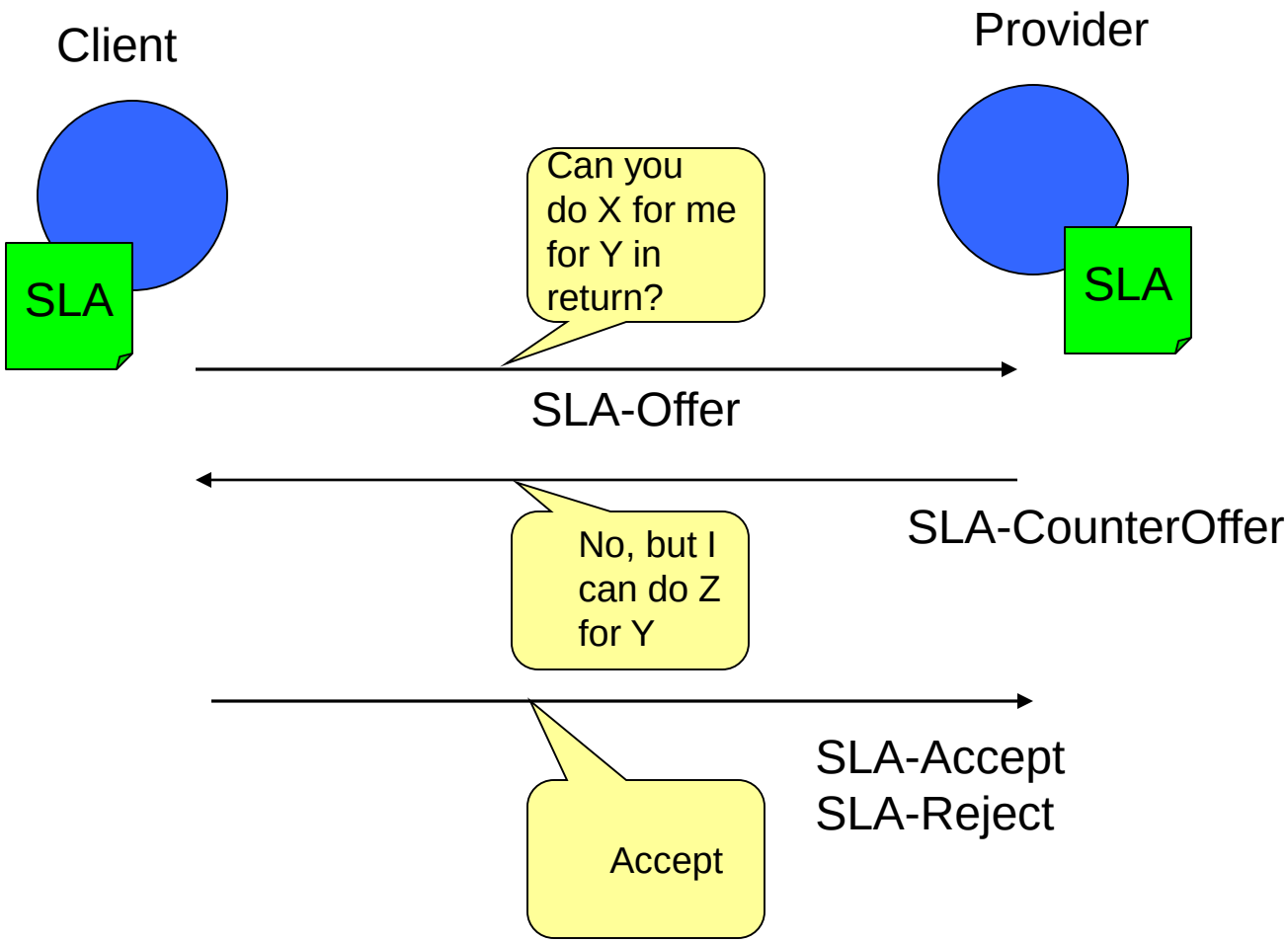


Distinguish between: **Discovery of suitable provider** → P2P Search, Directory Service

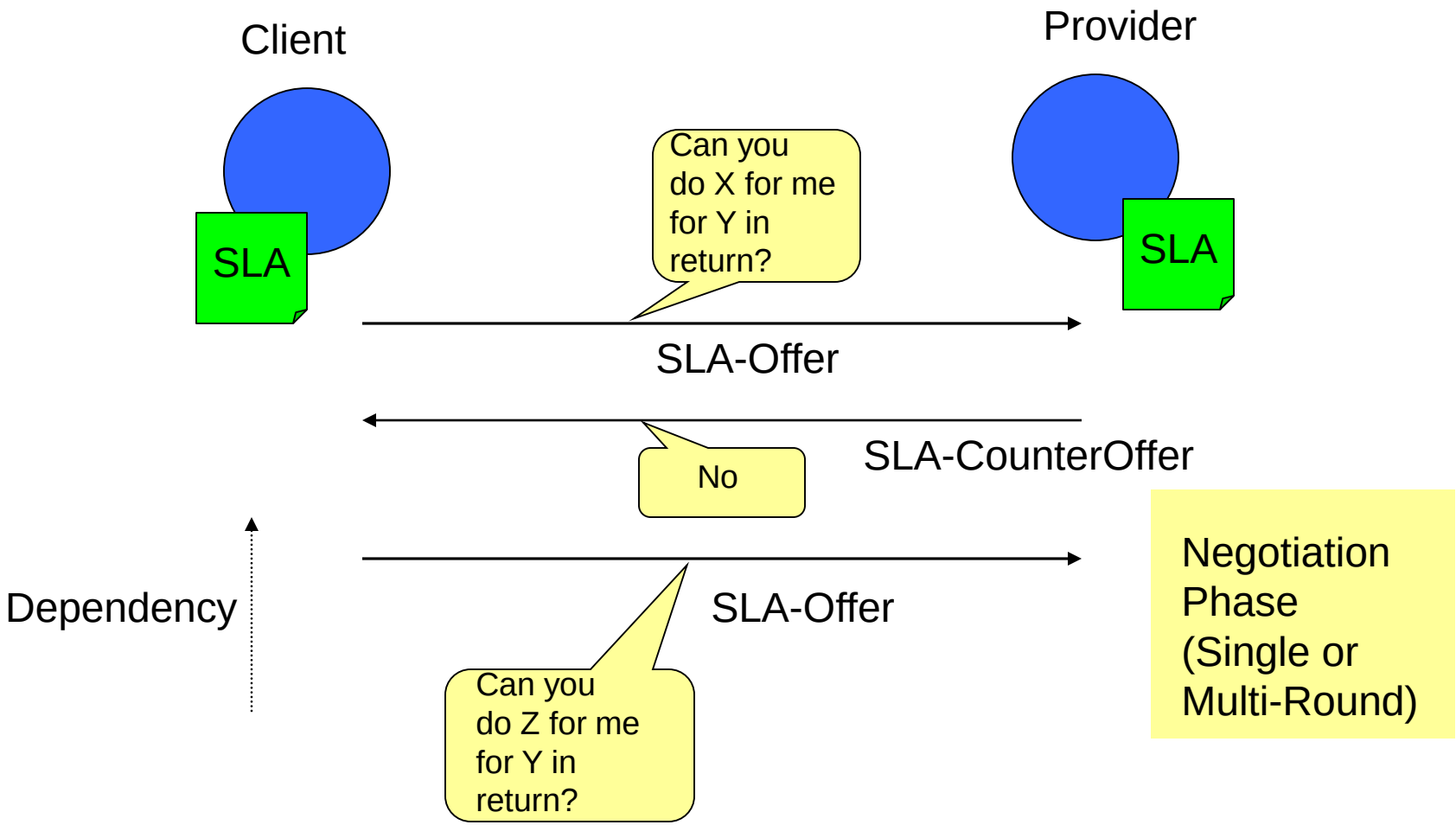
Establishment of an SLA

SLA as a basis to support adaptive behaviour

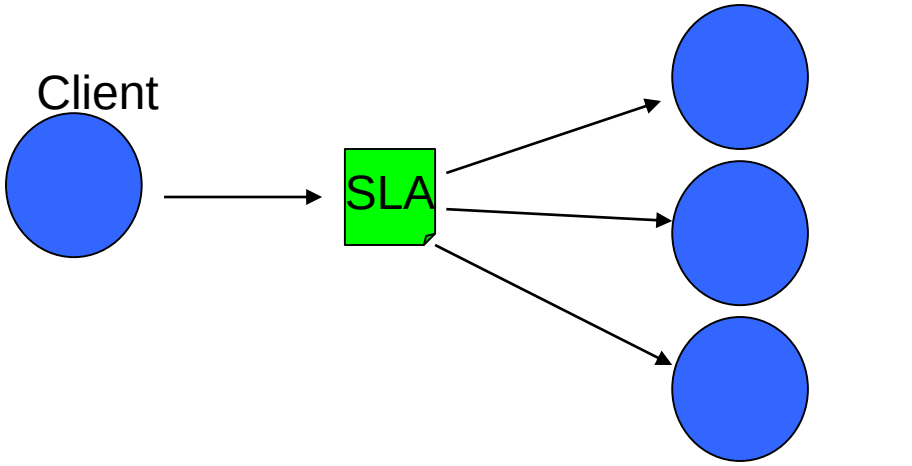
What is an SLA?

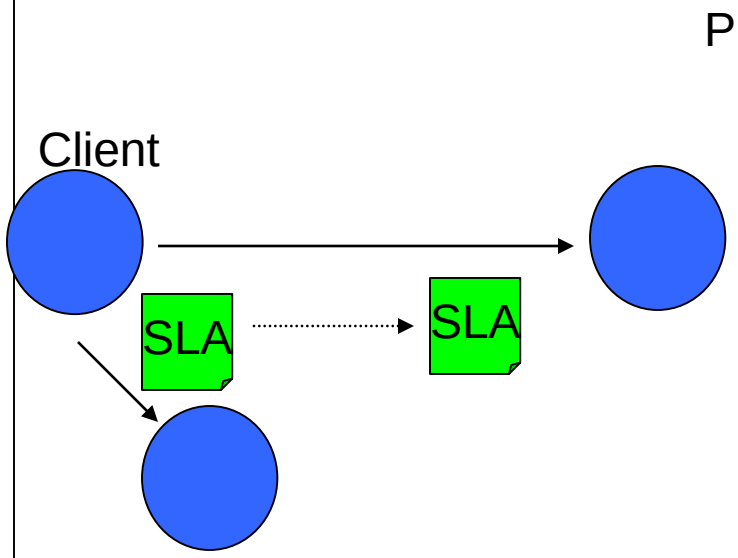


What is an SLA?



Variations





Multi-provider SLA

Single SLA is divided
across multiple providers
(e.g. workflow composition)

SLA dependencies

For an SLA to be valid, another
SLA has to be agreed
(e.g. co-allocation)

What is an SLA?

- Dynamically established and managed relationship between two parties
- Objective is “delivery of a service” by one of the parties in the context of the agreement
- Delivery involves:
  - *Functional and non-functional properties of service*
- Management of delivery:
  - *Roles, rights and obligations of parties involved*

Forming the Agreement

- Distinguish between:
  - *Agreement itself*
  - *Mechanisms that lead to the formation of the agreement*
- Mechanisms that lead to agreement:
  - *Negotiation (single or multi-shot)*
  - *One-shot creation*
  - *Policy-based creation of agreements, etc.*

SLA Life Cycle

- Identify Provider
  - *On completion of a discovery phase*
- Define SLA
  - *Define what is being requested*
- Agree on SLA terms
  - *Agree on Service Level Objectives*
- Monitor SLA Violation
  - *Confirm whether SLO's are being violated*
- Destroy SLA
  - *Expire SLA*
- Penalty for SLA Violation

WS-Agreement

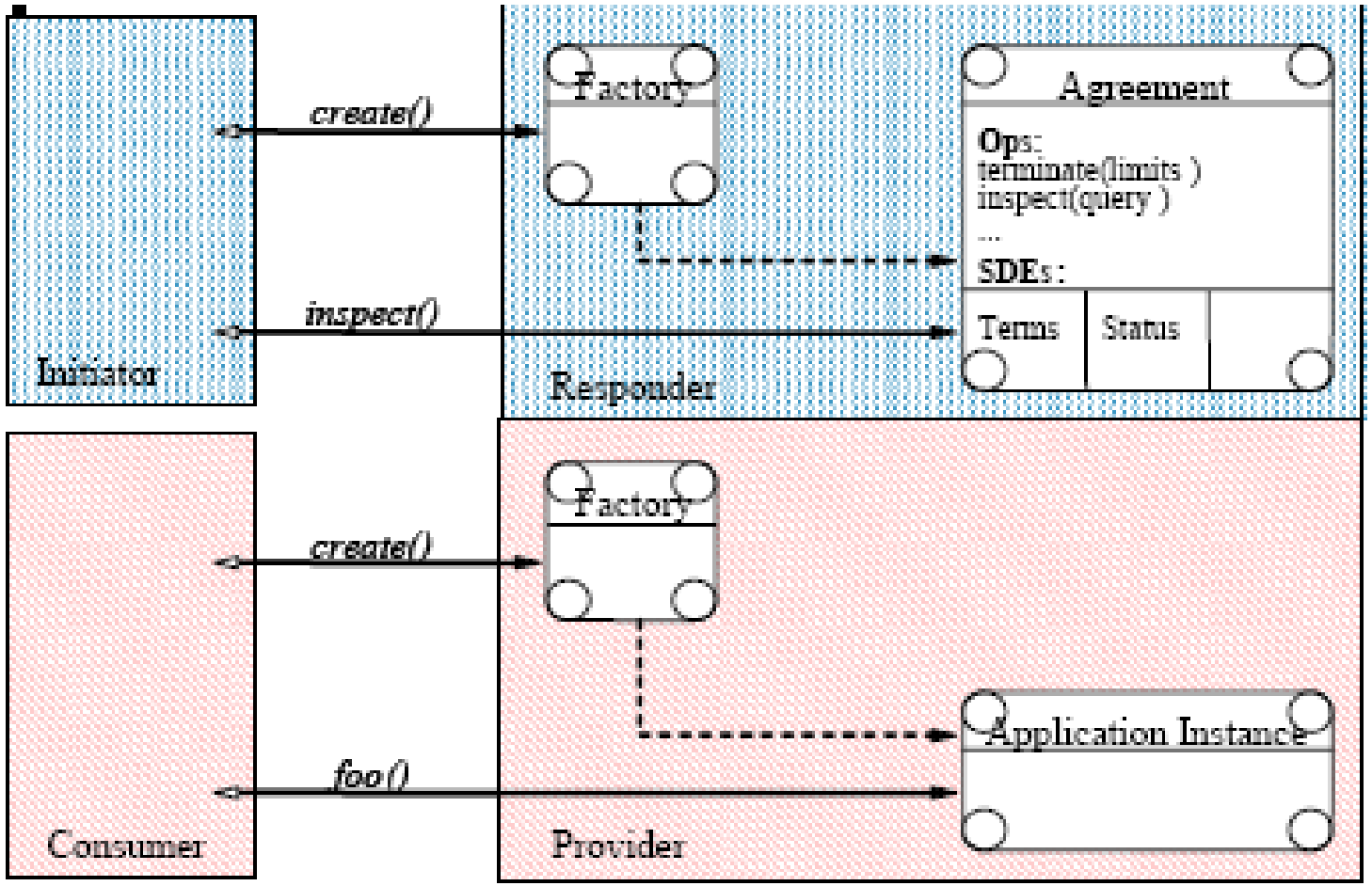
- Specification for Service Level Agreements
  - *Developed through GRAAP WG at the Open Grid Forum*
  - *WSLA (from IBM) – previous efforts*
- Provides:
  - *Schema for agreement terms*
  - *A very simple protocol (two stage)*
  - *A state sequence*
  - *Support penalty clauses*
- No support for negotiation

WS-Agreement Specification Document (GFD.107)

WS-Agreement

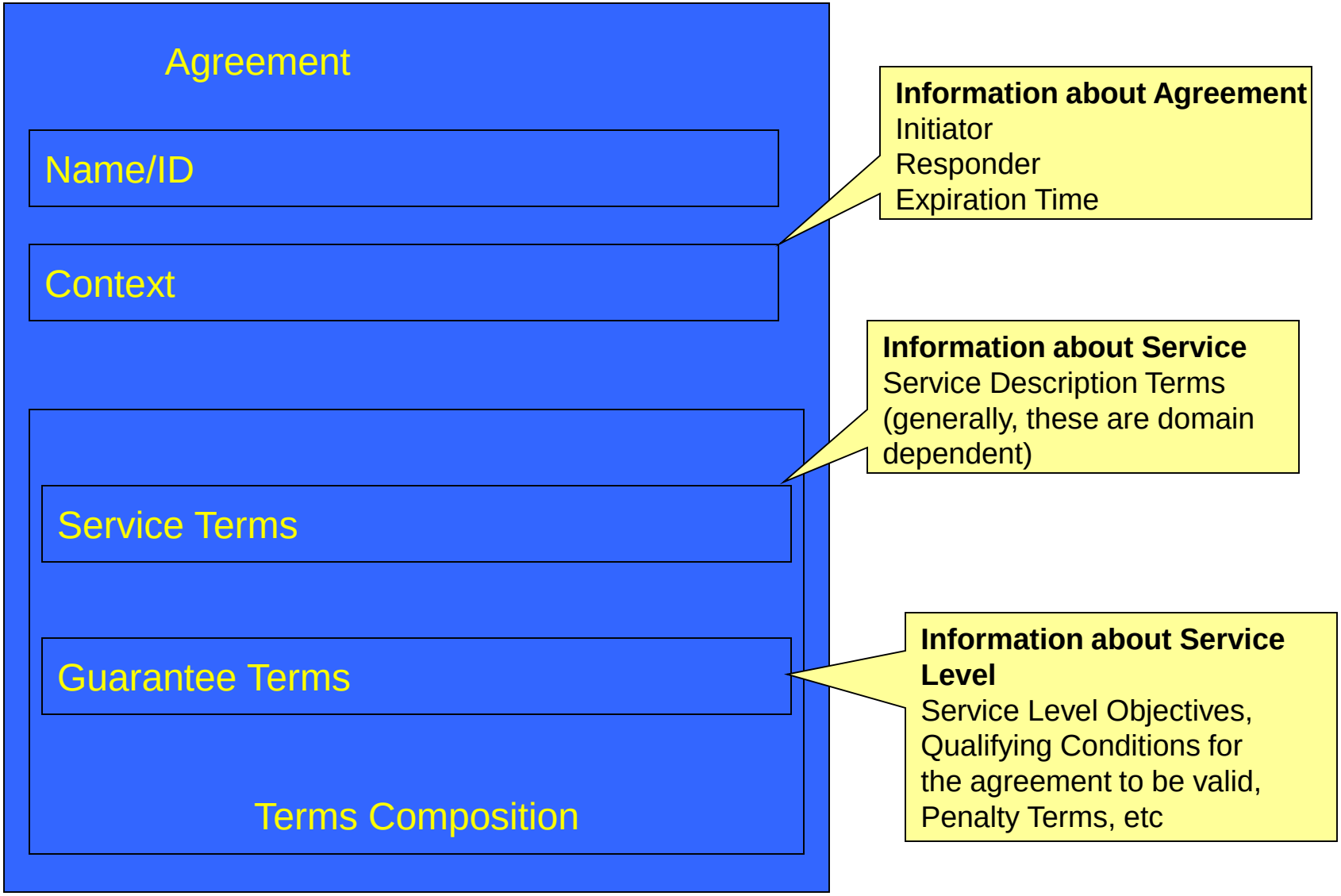
- Service Client/Provider does not need to be a Web Service
  - *Approach is quite general in scope*
- Provides a two layered model:
  - *Agreement layer: Web Service-based interface to create, represent and monitor agreements*
  - *Service layer: Application specific-layer of service being provided*

WS-Agreement in general

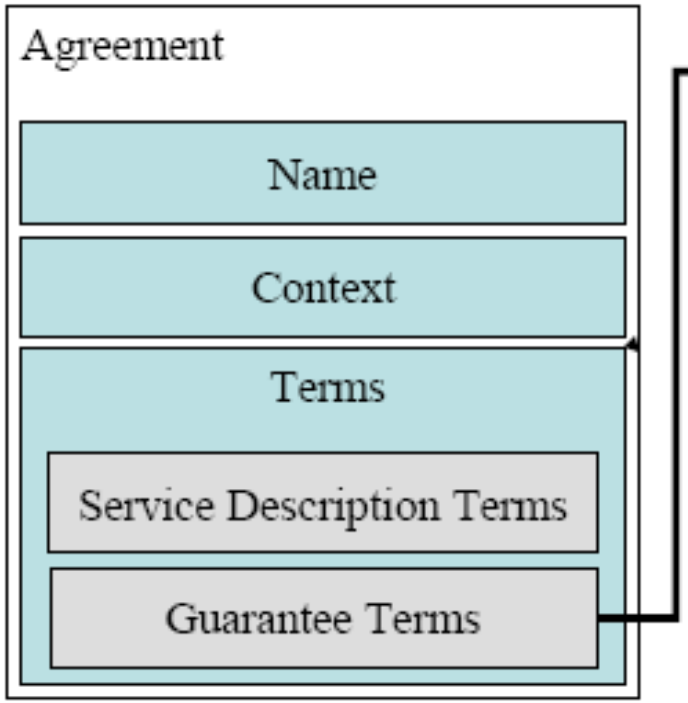


Agreement Initiator may be Service Consumer or Service Provider

WS-Agreement



WS-Agreement Terms



Service Level Objectives Business Value List
Service Level Indicators

SLO	SLI	BVL
CPUs	8	5
Start time >	9:00	10
Finish time <	23:00	4
Job Duration <	5	1
Job done:		59
Price per time unit:		8
Penalty for delays:		1

From: Viktor Yarmolenko (U Manchester)

Service Level Objectives

- Must be measurable items
  - *To some degree of accuracy*
  - *Who measures? (service clnt/user or service provider?)*
  - *Use of a trusted third party?*
- Must be mutually agreed terms
  - *Term semantics need to be resolved*
  - *Measurement interval and measurement value must be pre-determined*
- Control infrastructure
  - *Service provider must be able to control/manage the parameter of interest*

Service Level Objectives … examples

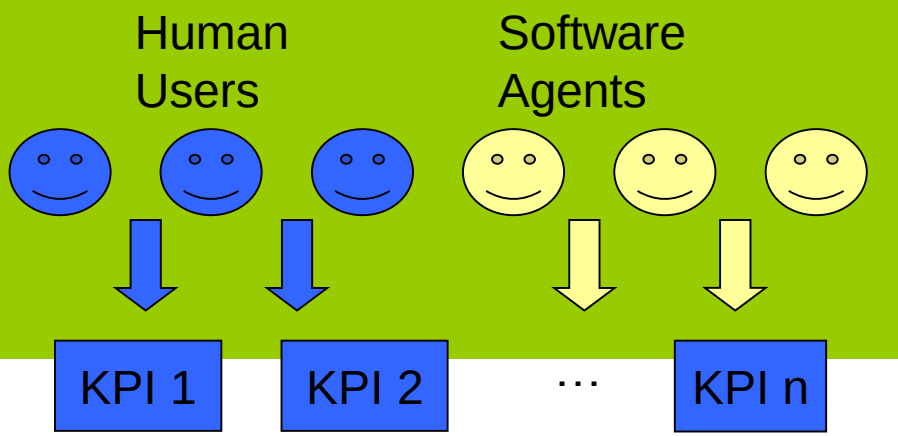
- Latency
  - *Network*
  - *Application* → **Response Time**
- Throughput
  - *Per time unit*
- Operational environment
  - *Can include static or dynamic parameters (such as OS, memory, number and type of CPU(s), etc)*
  - *Up time, Availability, Accessibility*
- Preferred completion time + start/end period (validity period)
- Owner + creator + user (identify individuals/systems responsible for this)
- Application specific terms
  - *Number of molecules, number of processing units, etc*

Key Performance Indicators to SLOs

- KPI: Indicators used to assess business efficiency
  - *May be directly measurable or derived*
  - *Based on business benefit*
- KPIs need to be mapped to SLOs
  - *A given KPI may map to multiple SLOs*
  - *Mapping may be influenced by a "business policy"*
  - *Indicates priorities for a business over a particular time frame*
- KPIs may be used to evaluate the success of an SLA
  - *Relationship between KPI and SLOs is important*

Key Performance Indicators (KPI)

KPI are specific to particular Application domains:
Business: “Reduce Employee Turnover” or “Increase Sales Volume”

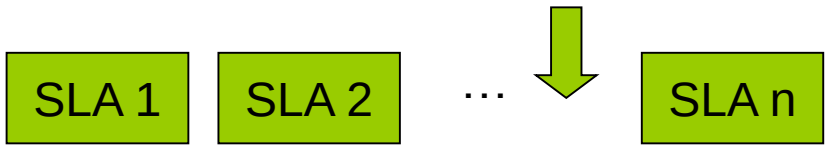


KPI Aggregation

Mapping to Operational Requirements

Operational Service Level Agreements

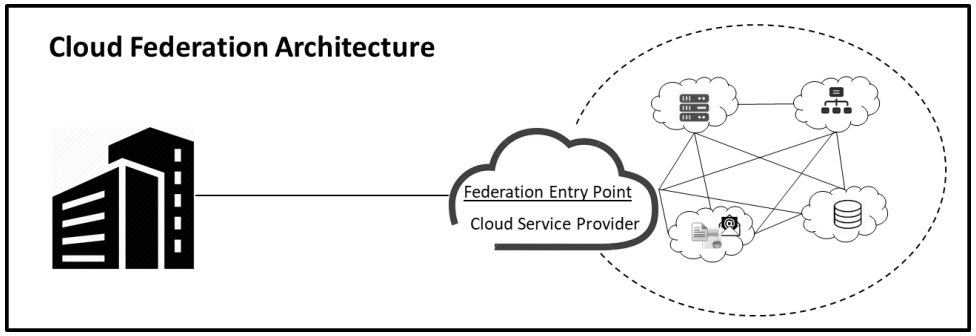
Service Level Agreements (SLAs) Specify operational constraints on the underlying resources – these would be compute, data and network resources in the context of a data center, for instance



Operational Resources

Federation Model

- Relying on single cloud – Not sufficient
  - Vendor lock-in
  - Geographical and legislative dependencies
  - Inefficient global responsiveness
  - Dealing with application requirements (Big data, HPC etc)
- Use '*Cloud of Clouds'* i.e. the cloud federation
  - Multiple providers combine their resources to provide dynamic coordination and distribution of load among a set of cloud data centers.
- Benefits include on-demand expansion, Low Capital Expense (CAPEX) and Operational Expenditure (OPEX) and better application resilience etc.



Characteristics of Federation

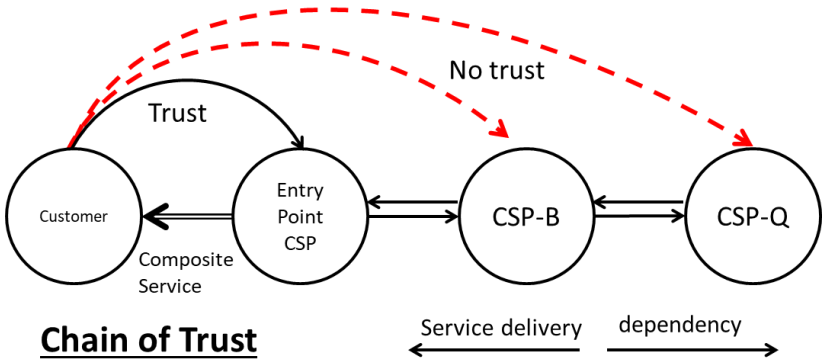
Attribute	Value
Service Providers	Multiple
Portfolio	Collaborative
Provider Interconnection	Dynamic (Short term)
Infrastructure	Heterogeneous
Service Composition	Multilevel
Service Junction	Federation entry point
User Service Level Agreement	Single
User-Provider Relationship	Static
Provisioning	Federation entry point
Scheduling	Transparent to user

Security Challenges of Federated Clouds

- Security in cloud computing infrastructure – *"old security transported to cloud"* [1]
  - mostly investigated and addressed in the traditional system and network security context e.g mechanisms for data protection, access control, mitigation of DDoS attacks, code verification etc.
- Lack of trust in cloud computing – already a show-stopper [2]
  - Loss of control over physical and logical aspects of data
  - Users have doubt regarding providers' intentions rather their performance.
  - Cloud adoption demands an increased level of confidence in the provider
- Added with the nature of federation – things are going to get worse
  - Newer trust challenges in federation

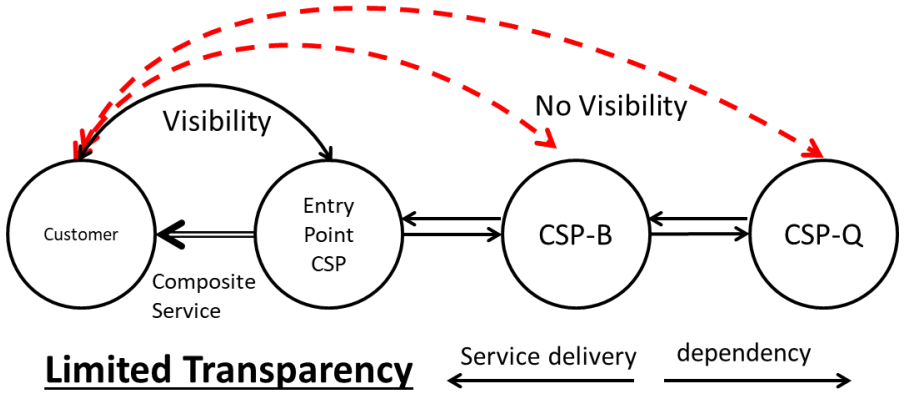
Trust Challenges – *chain of trust*

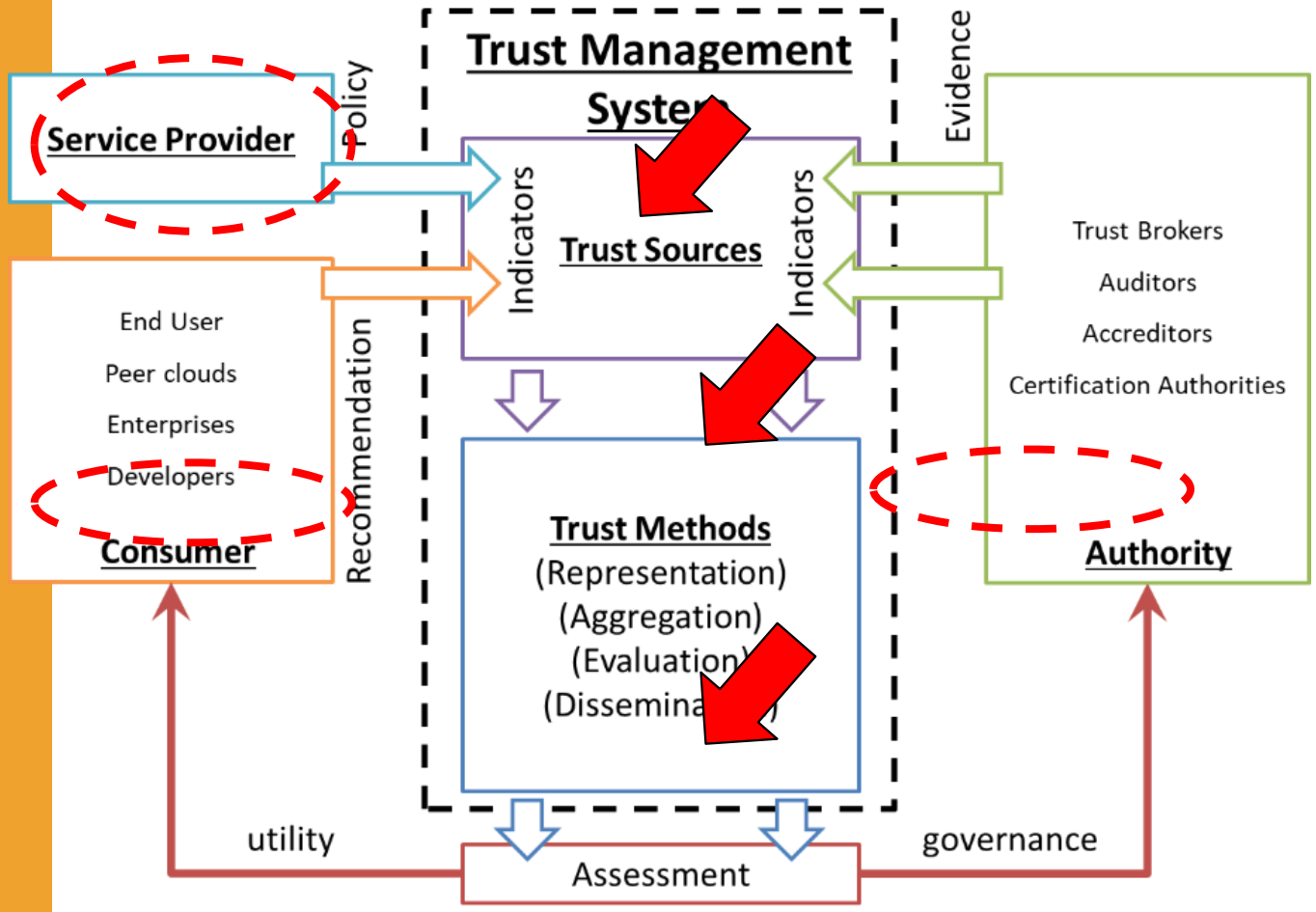
- Trust in cloud federation with chain of providers
  - Explicit interaction with only immediate neighbor
  - Enormous loss of control over data and application
  - Near to none user confidence
- Who is at stake?
  - User – owner of data and application
  - Entry point CSP – Signed a contract with the user
- Can the entry point CSP convince the user that everything is good within federation?
  - First it need to be confident itself



Trust Challenges – *Limited Transparency and Auditability*

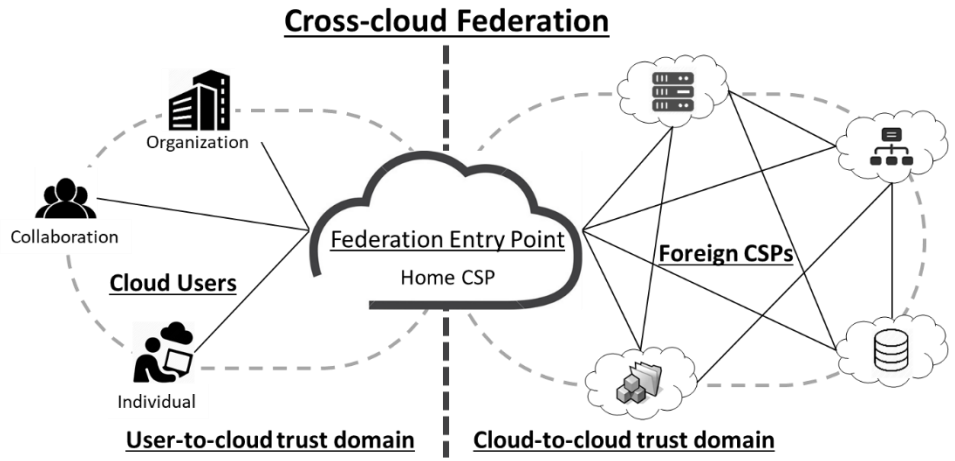
- A problem for cloud computing in general and for federated clouds in particular
  - Works two ways in federated clouds
- Non-transparent providers
  - No way to pin point malfunctioning service component
- ***User perspective*** - Where is my data and application currently residing?
- ***Provider perspective*** - Whose data (and what kind of data) is placed in its premises?




Trust Management System
Service Provider
Policy
Evidence
Indicators
Trust Sources
Indicators
Trust Brokers
Auditors
Accreditors
Certification Authorities
End User
Peer clouds
Enterprises
Developers
Consumer
Recommendation
Trust Methods
(Representation)
(Aggregation)
(Evaluation)
(Dissemina
Authority
utility
governance
Assessment

Trust Management for Federation

- First thing first – reduce the chain of trust
  - If somehow the entry point CSP has irrefutable evidence of trust in its peers that can be presented to the user and vice versa
  - **User-to-cloud** and **cloud-to-cloud** trust

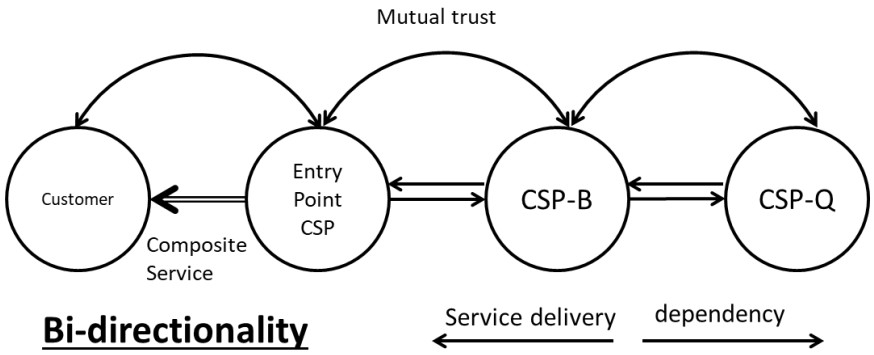


- Trivial (performance / recommendation based)
- Lots of literature already available

- Challenging
- Participants are competitors
- Multilevel services
- Needs to be transparent and presentable to user

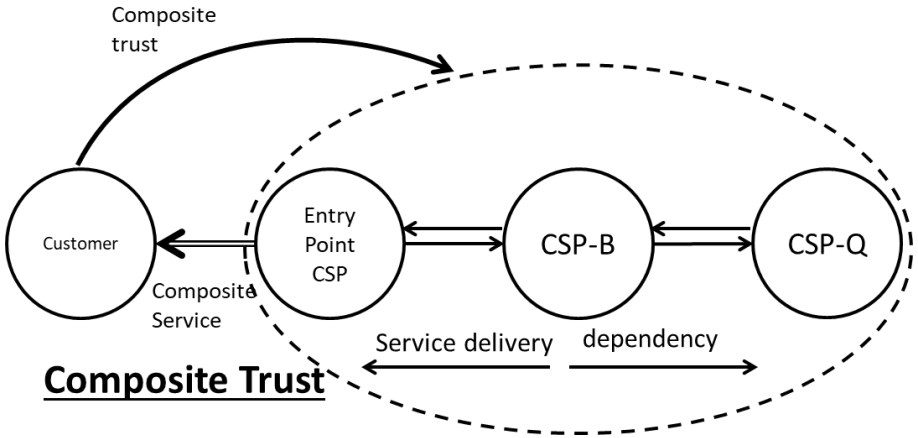
Cloud-to-cloud trust – *Bi-directionality*

- Conventional Cloud systems
  - Service requester needs to trust service provider
- Federation
  - Both service requester and service provider must mutually trust each other
- Already available trust models can be instantiated in both directions (easy said than done)



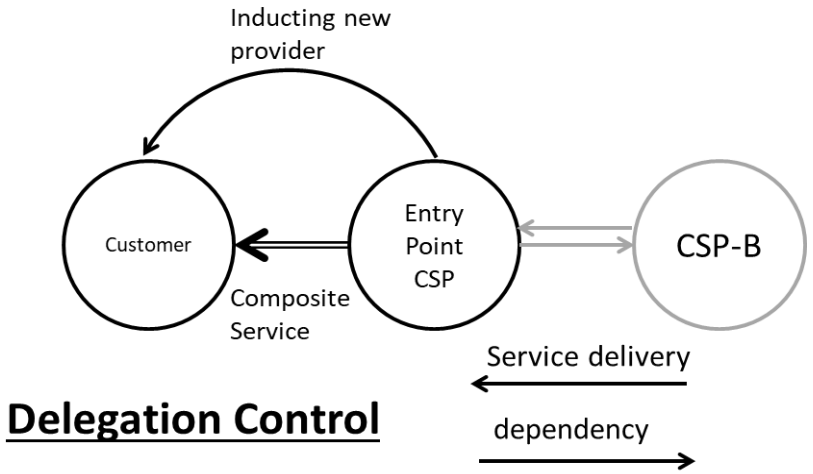
Cloud-to-cloud trust – *Compositional Trust*

- Trust the service in entirety
  - Trust should be composed in a similar manner as the service has been composed
  - Trust level of every individual provider joining a service should be reflected to the overall global trust value



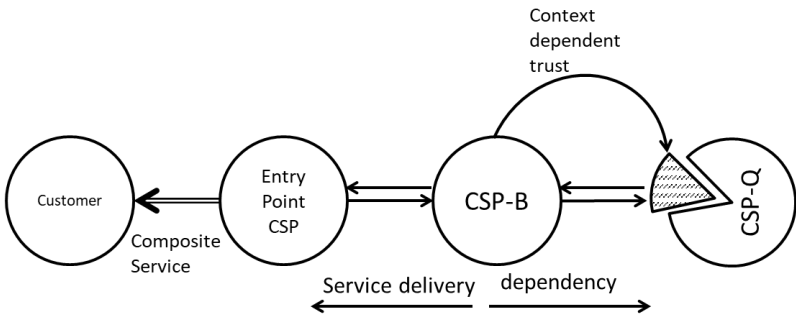
Cloud-to-cloud trust – *Delegation Control*

- In a collaborative scenario trust should never be unilaterally delegated
- User data/application entrusted to one provider should not be delegated to the other without owner's knowledge
- A third party may overlook the federation. But still !!!



Cloud-to-cloud trust – *Context Awareness*

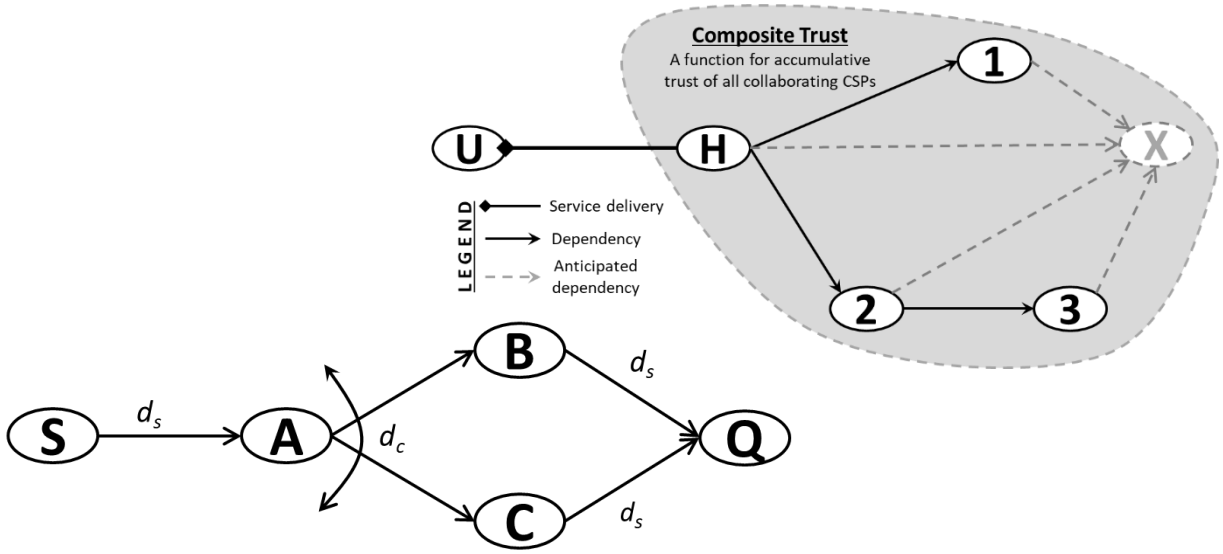
- A cloud provider may lease different type of resources in different contexts with different settings e.g.
  - Storage, compute etc.
  - Data back-up and recovery
  - Computational spikes, Data bursts
  - Identity verification etc.
- Global assessment of a CSP is not recommended for context dependent evaluations
  - A provider demonstrating sufficient overall trust may lack in something the others are only interested in.



Context Awareness

Proposed Research- *Compositional Trust*

- Compositional trust
  - Represent federation as graph based structure
  - Service dependency evaluation i.e. singular or concurrent
  - Trust evaluation for service dependency extending Dempster-Shafer model for trust evaluation based on Cloud Security Alliance STAR certification dataset



CSA STAR Program

- Cloud Security Alliance (CSA) offers a Security, Trust and Assurance Registry (STAR) program helping to incorporate
  - transparency
  - rigorous auditing and
  - harmonization of standards
- Indicates best practices for validation of security posture of cloud offerings
- The program is based upon:
  - Cloud Control Matrix (CCM) framework
  - Consensus Assessment Initiative Questionnaire (CAIQ)
- STAR Database
  - Publicly accessible registry database of security and privacy controls provided by popular cloud computing offerings

CSA STAR Program

- STAR consists of three levels of assurance
  - Self Assessment
  - 3rd party certification and
  - Continuous auditing

Cloud Control Matrix

- Offers a framework of cloud-specific security controls
- Provide organizations with the needed structure, detail and clarity relating to information security tailored to cloud computing
- Contains 16 Control Domains, each having a detailed specification and architectural relevance to cloud
- Each control specification is based on the its relationship with other industry standards, regulations, and frameworks like
  - ISO 27001/27002, ISACA COBIT, PCI, NIST, Jericho Forum and NERC CIP, HIPAA

CCMv3.0.1 CLOUD CONTROLS MATRIX VERSION 3.0.1

Control Domain	CCM V3.0 Control ID	Updated Control Specification	Architectural Relevance					
			Phys	Network	Compute	Storage	App	Data
Business Continuity Management & Operational Resilience Documentation	BCR-04	Information system documentation (e.g., administrator and user guides, and architecture diagrams) shall be made available to authorized personnel to ensure the following: • Configuring, installing, and operating the information system • Effectively using the system's security features		X	X	X	X	X
Business Continuity Management & Operational Resilience Environmental Risks	BCR-05	Physical protection against damage from natural causes and disasters, as well as deliberate attacks, including fire, flood, atmospheric electrical discharge, solar induced geomagnetic storm, wind, earthquake, tsunami, explosion, nuclear accident, volcanic activity, biological hazard, civil unrest, mudslide, tectonic activity, and other forms of natural or man-made disaster shall be anticipated,	X					

Consensus Assessment Initiative Questionnaire (CAIQ)

- Expands on CCM using the terms and descriptions considered to be a best practice by the CSA
- A set of Yes/No assertions that a cloud consumer and cloud auditor may wish to ask of a cloud provider to ascertain their compliance to the CCM
- standardized way for cloud providers to showcase their detailed capabilities and competencies

CCM				CAIQ		
Control Domain	Control ID	Description	Relevance	Question ID	Question	Options
Application & Interface Security *Customer Access Requirements*	AIS-02	Prior to granting customers access to data, assets, and information systems, identified security, contractual, and regulatory requirements for customer access shall be addressed.	Phy, N, C, S, App, D	AIS-02.1	Are all identified security, contractual, and regulatory requirements for customer access contractually addressed and remediated prior to granting customers access to data, assets, and information systems?	1- Yes 2- No 3- NA 4- unanswered
				AIS-02.2	Are all requirements and trust levels for customers' access defined and documented?	

STAR Adoption

- More than 200 participants in STAR registry including

- CSA Certified Auditors and Attesters

Cloud Security Alliance (Business Continuity)

- Determine the best acceptable RTO/RPO (Recovery Time Objective/Recovery Point Objective) for all systems whether Cloud or Non-cloud Apps/Data/Systems.
- Integrate Cloud versus Non-cloud Services into SLAs.
- Establish Performance Requirements.
- Cite Bandwidth Requirements for things like replication to the CSP and client access to hosted services.
- Detail System Requirements for Recovery Workstations.
- Provide a Declaration of DR Procedures and Workflow.
- Specify how failover and failback are handled.

Cloud Security Alliance (Business Continuity)

- What is the value of an entity's data to the business and how much will it cost to replace it, if it is lost or stolen?
- After conducting a Risk Assessment/Risk Analysis, what data should go into the cloud?
- How effective is the Cloud Service Providers own BC/DR Planning?
- How is access to the BC/DR systems and access from the BC/DR systems to third parties managed? If there is an established relationship between the CSP and the BC/DR provider, procedures already may be in place. Do they mesh well with the consumer's needs?

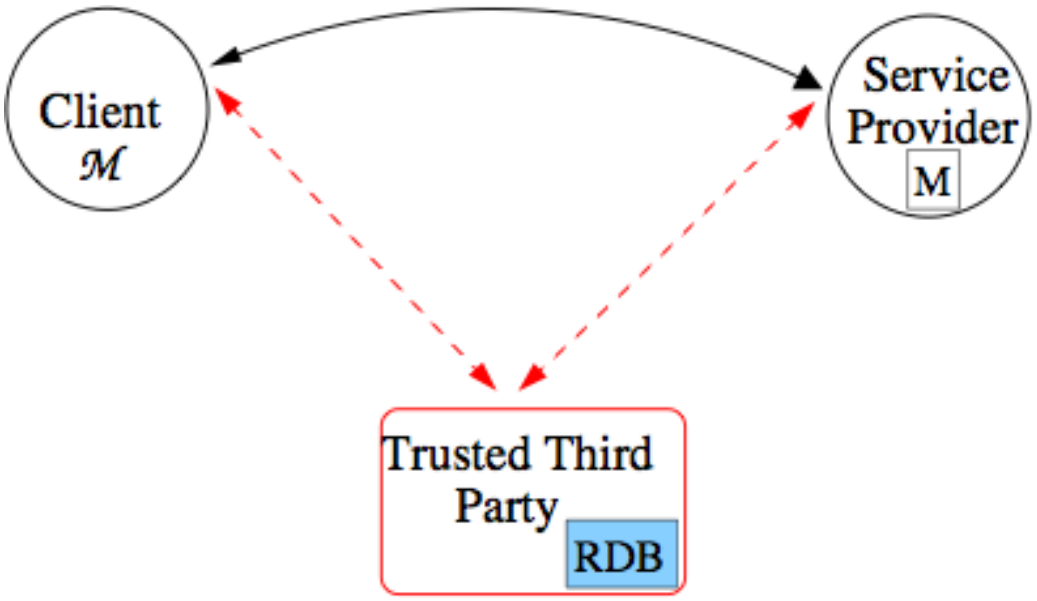
Cloud Security Alliance (Business Continuity)

- **Data Replication** is NOT the same as data archiving or backing up data. Unless snapshots and journaling are used, replication will only replicate whatever is "live" so deletions/errors/corruptions will be replicated to the local Highly Available (HA) instance and the remote BC/DR instance(s).
- To provide solid DR, backups, whether as snapshots to disk or more traditional backups to tape, also should be implemented.
- **Elasticity of the Cloud Provider** – Can they provide all the resources if BC/DR is invoked?
- **How is DR testing achieved**? Does the CSP support DR testing?
- **How are physical and virtual workloads accommodated** when using cloud for BC/DR?
- How are the DR services accessed if invoked?

SLA Violations

- Significant existing body of work on specifying and monitoring SLOs
- Limited studies on *how* SLOs may be impacted by specifying penalty clauses
  - *WS-Agreement contains "penalty" and "reward" terms – but these are not very clearly defined*
  - *Penalty and Rewards are central to business use of SLAs*
- SLAs containing penalty clauses better reflect 'real world' contracts
- **Business scenarios**: violations are important to support financial sanctions
- **Scientific scenarios**: violations used to determine choice of resources

SLA Monitoring & Violations



- ‘All or nothing’ provision
- ‘Partial’ provision
- ‘Weighted partial’ provision

Types of Violations

Monitoring plays a crucial role in discovering violations:

- Trusted Third Party (TTP)
- Trusted Module on server site (M)
- Model at client site (M)

A prerequisite for contract (SLA) enforcement

Example penalty clauses … from a providers perspective

If 90% of the number of requested CPUs, and 90% of requested memory have been delivered, then these SLOs have not been violated.

For provisioning below 90% of CPU and memory, and for each percent the provider must incur a penalty of α monetary units.

If 90% of the execution times are not in the 2 second range,

then for each deviation from the 98% of between 2 and 5 seconds, the penalty to the provider is β monetary units,

and for each percent of the 98% of execution times more than 5 seconds, the penalty is γ,

and for other percents that are more than 5 seconds, the penalty is α monetary units.