

Foundations and applications of SLAs for Elastic Computing

Foundations and applications of SLAs for Elastic Computing



isa
INGENIERIA DEL SOFTWARE APLICADA

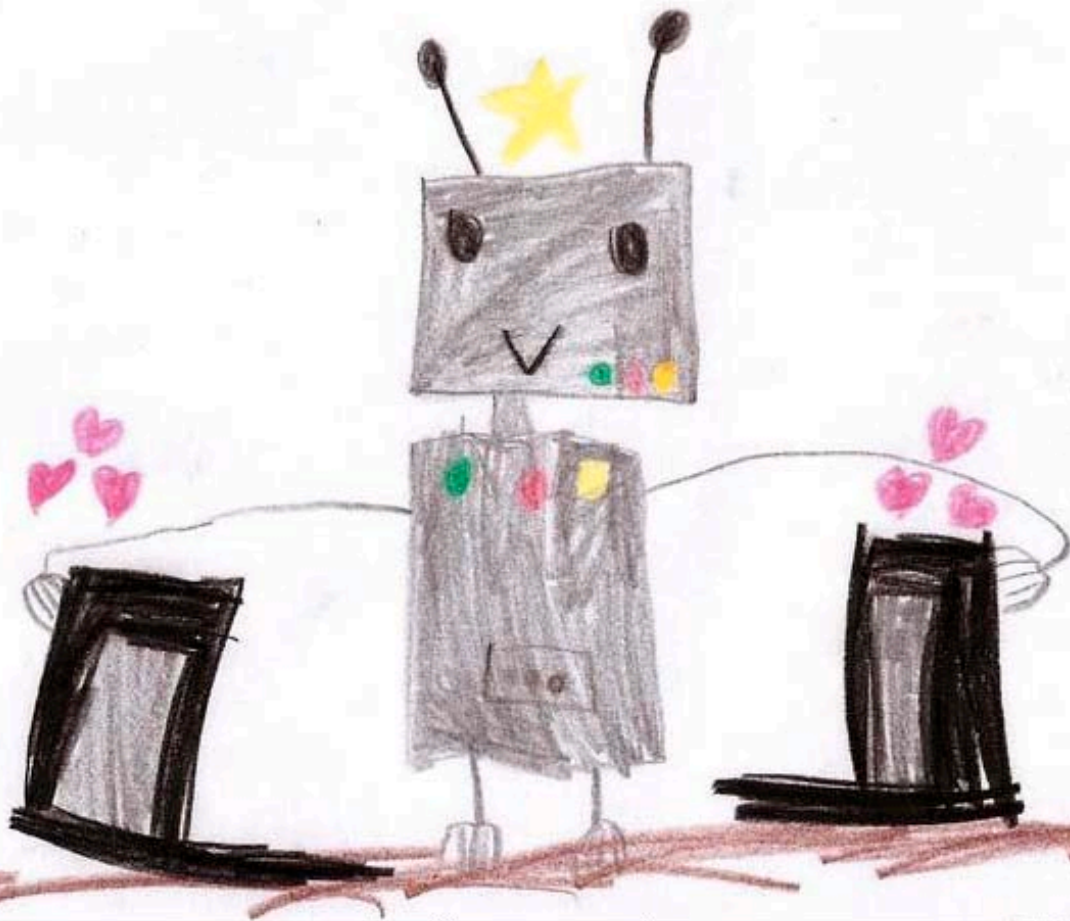


Prof. Dr. Antonio Ruiz-Cortés

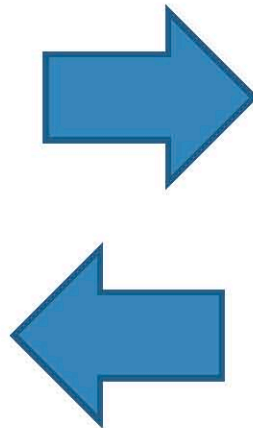








Laura



Provision

Compliance

Monitor

Workload optimization

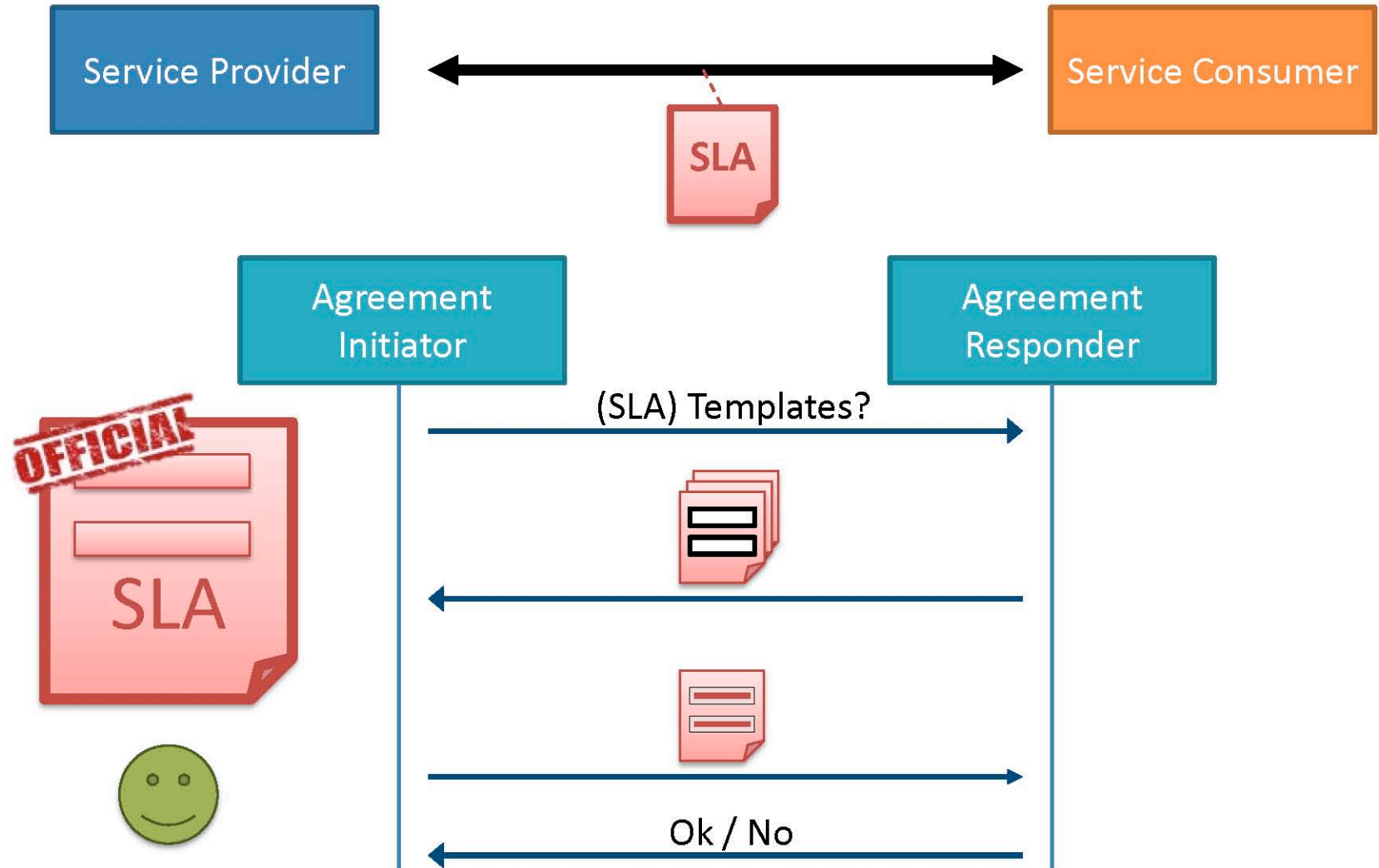
QoS

Service Level Objective

Pricing

Customizing

WS-Agreement



WS-Agreement documents structure

Identifier

Context

Parties / Roles (Initiator/Responder) / Expiration time / ...

Service Properties

- Var1 Id, Var1 Location, **Metric Lang1**
- VarN Id, VarN Location, **Metric LangN**

Terms -All (and)-

Any nested combination of:
{ All, ExactlyOne (xor),
OneOrMore (or), SDT, GT }

Service Description Term 1

SDT Lang 1

Service Description Term M

SDT Lang M

Service Reference 1

SR Lang 1

Service Reference P

SR Lang P

Guarantee Term 1

- Obligated: "Provider or Consumer"
- QualifCondition: **QC Lang 1**
- ServiceLevelObj: **SLO Lang 1**

Guarantee Term J

- Obligated: "Provider or Consumer"
- QualifCondition: **QC Lang J**
- ServiceLevelObj: **SLO Lang J**

Creation Constraints (for templates only)

Items

- Item1 Name, Item1 Location, **Item Constraint Lang 1**
- ItemK Name, ItemK Location, **Item Constraint Lang K**

Constraints

General Constraint Lang 1

General Constraint Lang L

WS-Agreement documents structure

Identifier

Context

Parties / Roles (Initiator/Responder) / Expiration time / ...

Service Properties

- Var1 Id, Var1 Location, **XSD-Schema**
- VarN Id, VarN Location, **XSD-Schema**

Terms -All (and)-

Any nested combination of:
{ All, ExactlyOne (xor),
OneOrMore (or), SDT, GT }

Service Description Term 1

JSDL

Service Description Term M

JSDL

Service Reference 1

XSD-Schema

Service Reference P

XSD-Schema

Guarantee Term 1

- Obligated: "Provider or Consumer"
- QualifCondition: **JEXL**
- ServiceLevelObj: **JEXL**

Guarantee Term J

- Obligated: "Provider or Consumer"
- QualifCondition: **JEXL**
- ServiceLevelObj: **JEXL**

Creation Constraints (for templates only)

Items

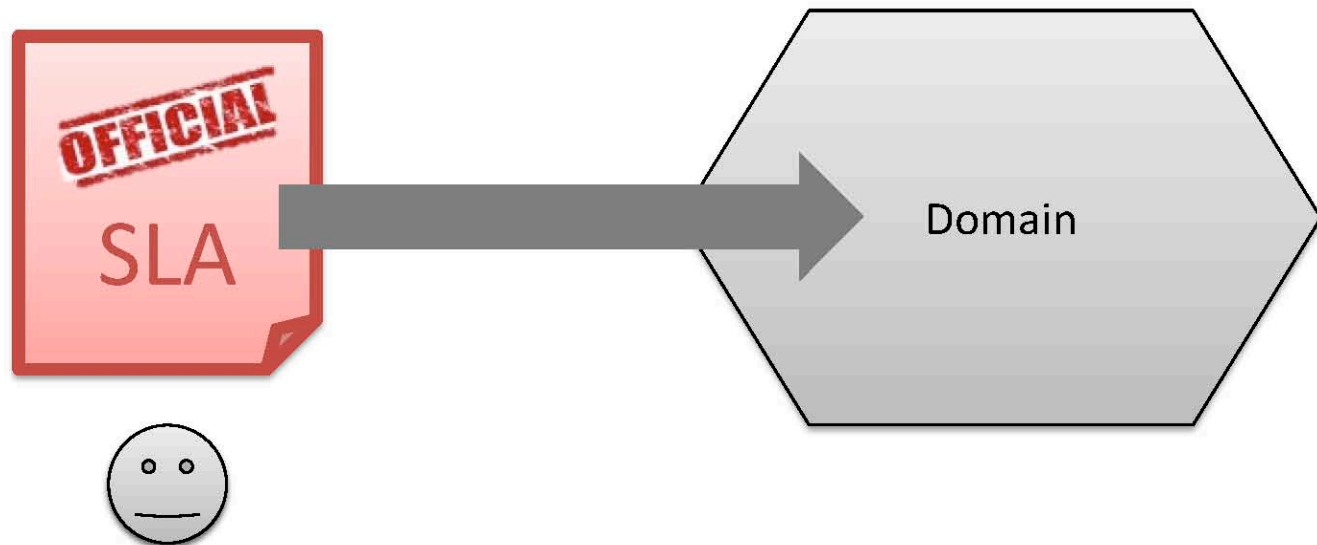
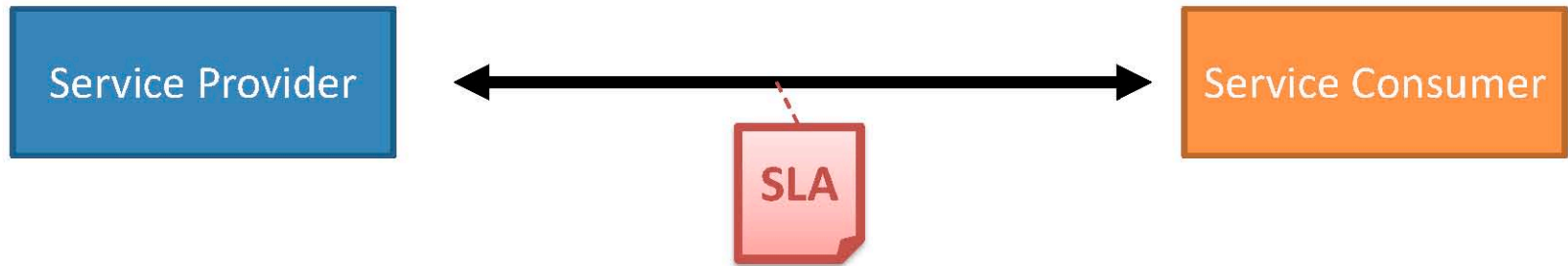
- Item1 Name, Item1 Location, **XSD-Restrictions**
- ItemK Name, ItemK Location, **XSD-Restrictions**

Constraints

JEXL

JEXL

WS-Agreement



WSAG4J Template

Identifier, Context, ...

Terms -All (and)-

Service Description Term

```
<jSDL:JobDefinition>
```

```
< ApplicationName = "WISDOM_PACK" />
```

```
< ApplicationVersion = "1.0" />
```

```
< POSIXApplication.InputFilename = />
```

```
< POSIXApplication.OutputFilename = />
```

```
< Resources.TotalCPUCount = />
```

```
</jSDL:JobDefinition>
```

Creation Constraints (for templates only)

WSAG4J Template

Identifier, Context, ...

Terms -All (and)-

Service Description Term

<jsdl:JobDefinition>

```
< ApplicationName = "WISDOM_PACK" />
< ApplicationVersion = "1.0" />
< POSIXApplication.InputFilename = />
< POSIXApplication.OutputFilename = />
< Resources.TotalCPUCount = />
```

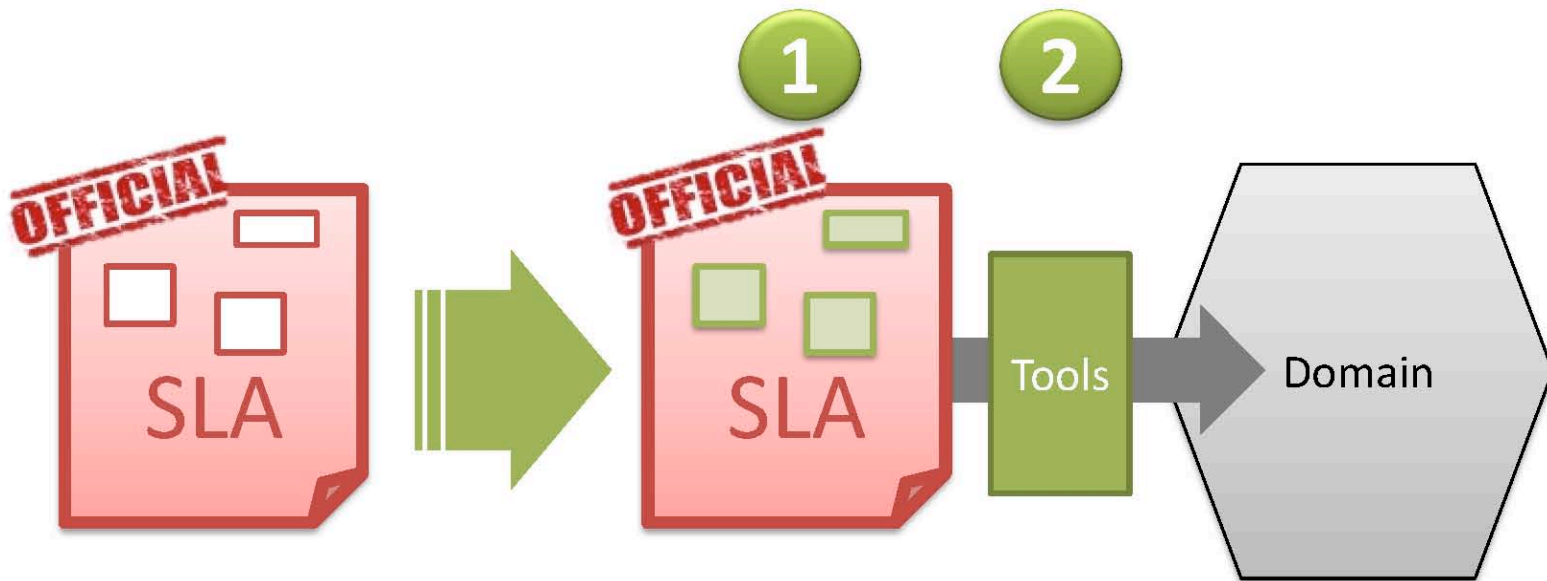
</jsdl:JobDefinition>

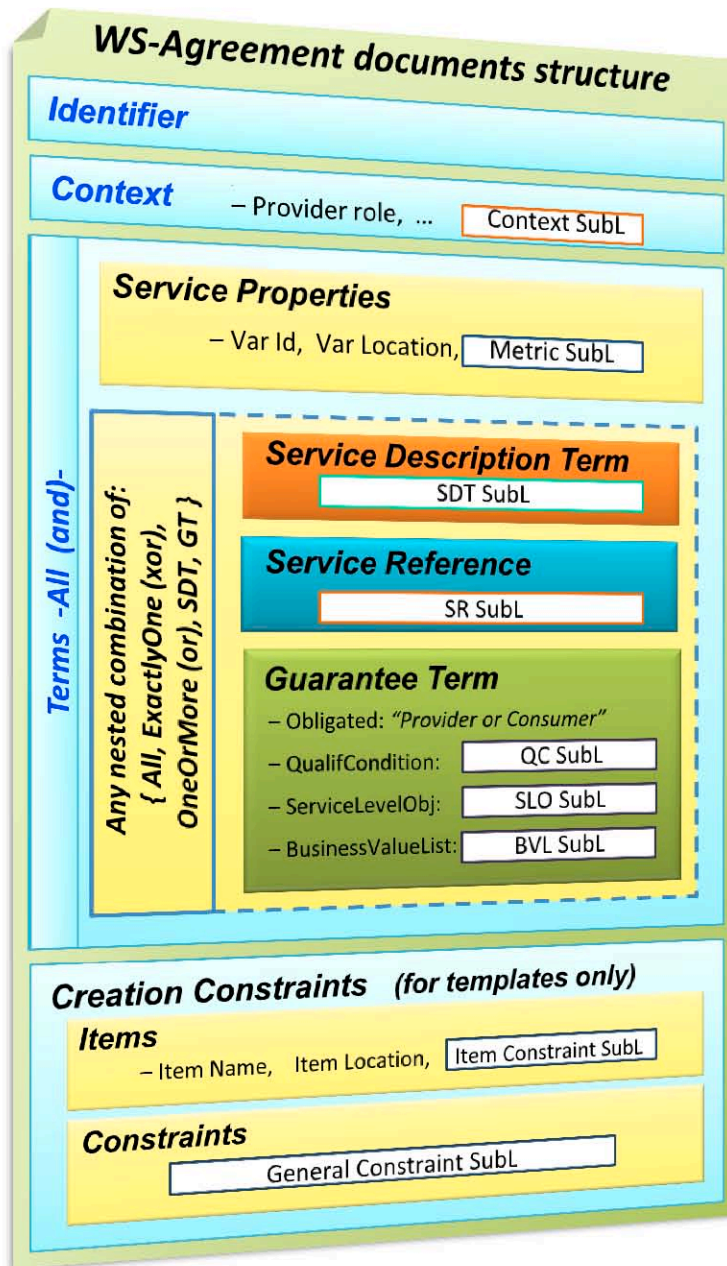
Creation Constraints (for templates only)

<xs:sequence>

```
< ApplicationName: xs:string = "WISDOM_PACK" />
< ApplicationVersion: xs:string = "1.0" />
< POSIXApplication.InputFilename: xs:string minOccurs=1 />
< POSIXApplication.OutputFilename: xs:string minOccurs=1 />
< Resources.TotalCPUCount: xs:decimal minOccurs=1 />
```

</xs:sequence>





WS-Agreement

Template AmazonS3 version 1.3
Provider Amazon as Responder;
Consumer Me;
Global Period: during 01/01/2012..12/31/2013
 except from 23:00 to 00:00 on SUN;
Defined Period:
 FreeTrialQuota: during 01/01/2012..12/31/2012;
Metrics [iAgree.generalMetrics, Amazon.metrics]

AgreementTerms
Service AWS-S3 available at. aws.amazon.com/s3;

Global Description:

Region;
 Interface;
 StorageSize;
 FirstProject;
 TotalPrice;

Monitorable Properties

global:
 MUP;
 ResponseTime;

Guarantee Terms

Exactly One between:

G1: Provider guarantees ResponseTime < 1000;
 on GlobalPeriod;
 onlyIf (Interface = SOAP);
 relative importance 0.2;

...
Creation Constraints:

C1: Region belongs {US, EU_Ireland};
 C2: TotalPrice = 0
 on FreeTrialQuota
 onlyIf (FirstProject and StorageSize <= 5)

iAgree

Template AmazonS3 version 1.3

Global Period: during 01/01/2013..12/31/2014
except from 23:00 to 00:00 on SUN;

Defined Period:

FreeTrialQuota: during 01/01/2013..12/31/2013;

...

AgreementTerms

...

Guarantee Terms:

G1: Provider guarantees MUP >= 99.9;

G2: Consumer guarantees TransferredGb < StorageSize * 100
AND ReadRequests >

WriteRequests;

on **FreeTrialQuota**;

G3: Provider guarantees ResponseTime < 1000;

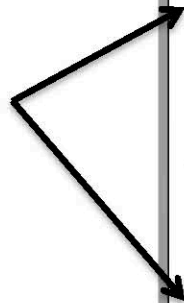
G4: Provider guarantees ResponseTime < 700;

on **from 00:00..12:00 during 01/01/2013..12/31/2013**;

G5: Provider guarantees MUP > 99.999;

on **GlobalPeriod** except during 01/01/2013..12/31/2013;

GlobalPeriod by default



Template AmazonS3 version 1.3

...
AgreementTerms

...
Guarantee Terms:

G1: Provider guarantees $MUP \geq 99.9$;

Only If $TransferredGb < StorageSize * 100$;

G2: Consumer guarantees $TransferredGb < StorageSize * 100$
AND ReadRequests >

WriteRequests;

Creation Constraints

C1: $MUP = 100 - (AmazonServerErrors / RequestsNumber)$



Analysis Operations Catalogue

ICore public ops.

existInconsistencies
 whyExistInconsistencies
 areCompliant(O.Ts,T.GTs)
 areCompliant(O.Ts,T.CCs)
 whyAreNotCompliant(O.Ts,T.GTs)
 whyAreNotCompliant(O.Ts,T.CCs)

ISolve private ops.

solve

 explain

implies
 whyNotImplies

IDocManager public ops.

createTemplate

 create...
 exportToWSAG4JOffer
 import...
 addCCs

 add...

getCompliantOffer
 getAgreement
 getAgVariants

Legend: Ts: terms, O:Offer, T:Term.

IChecker public ops.

existConflicts
 existDeadTerms
 existCondInconsisTerms
 areCompliant(O,T)
 whyExistConflicts
 whyExistDeadTerms

 whyExistCondInconsisTerms

whyAreNotCompliant(O,T)

INormalisation & IMap private ops.

getVariants
 temporalNormalisation

map

IFulfillment public ops.

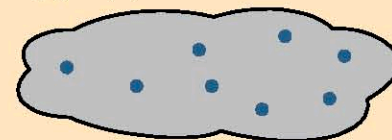
isFulfilled
 whyIsNotFulfilled
 generateMMD
 mergeSLAMMD

IValidator public ops.

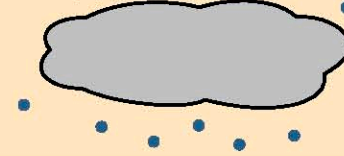
isValid
 whyIsNotValid(T)
 whyIsNotValid(O,T)



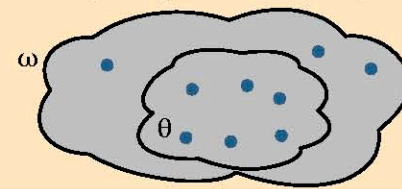
$solve(\delta) \neq \{\emptyset\}$



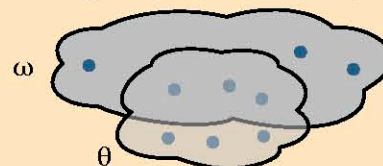
$solve(\delta) = \{\emptyset\}$



$implies(\theta. GTs, \omega. Terms)$



$\neg implies(\theta. GTs, \omega. Terms)$



Valid



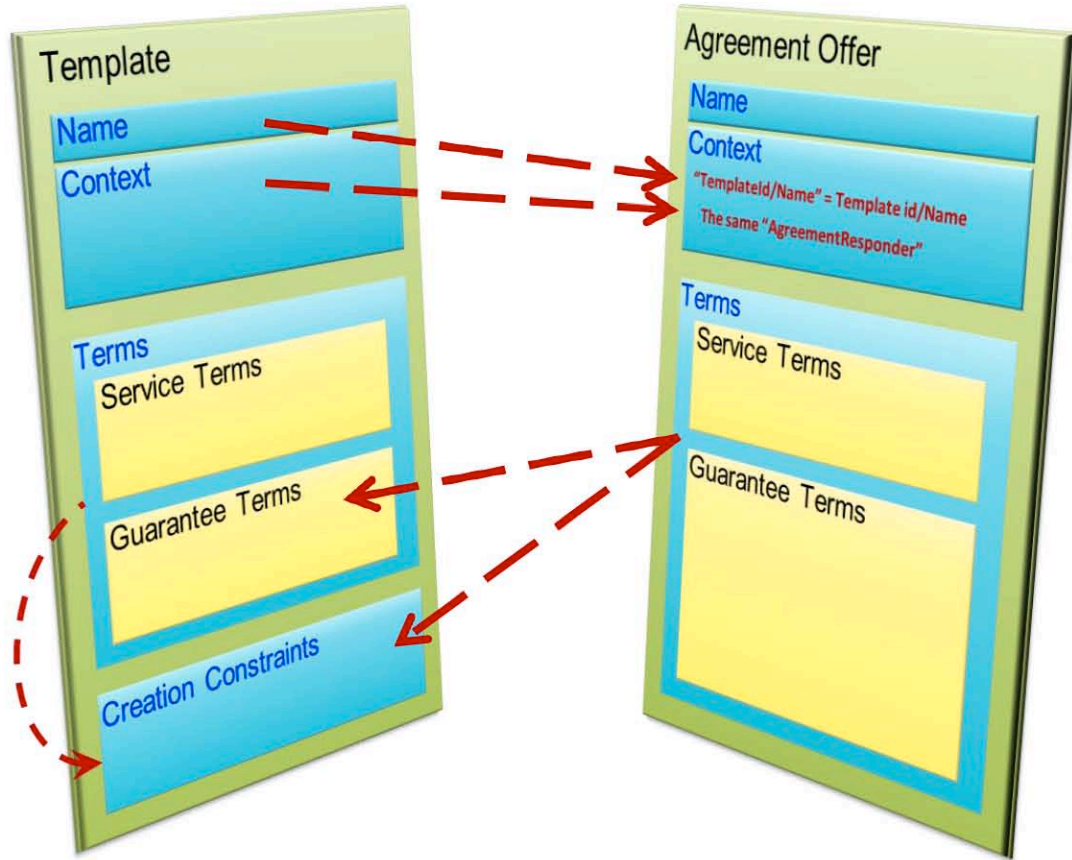
No conflicts

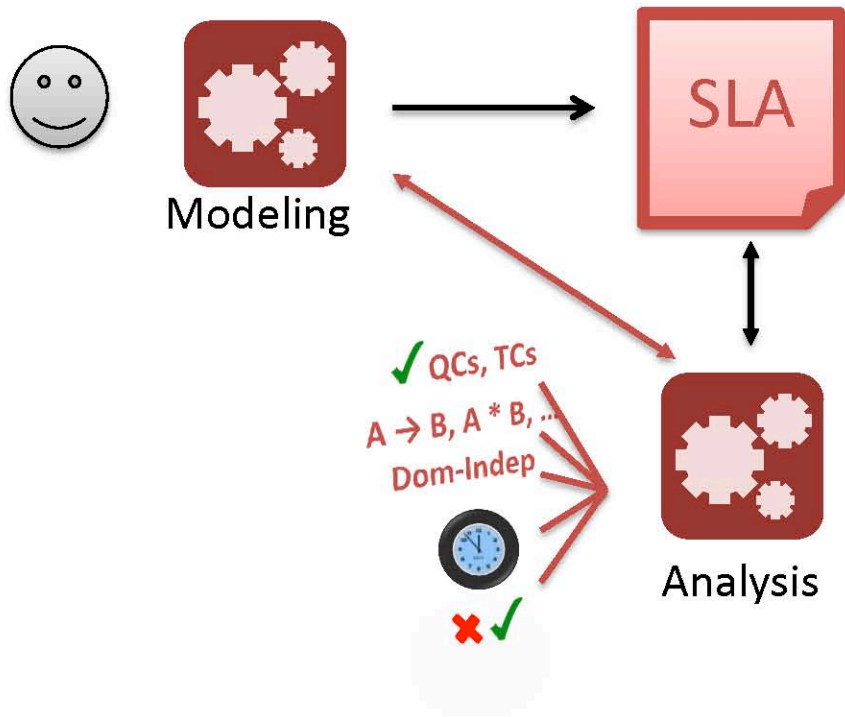


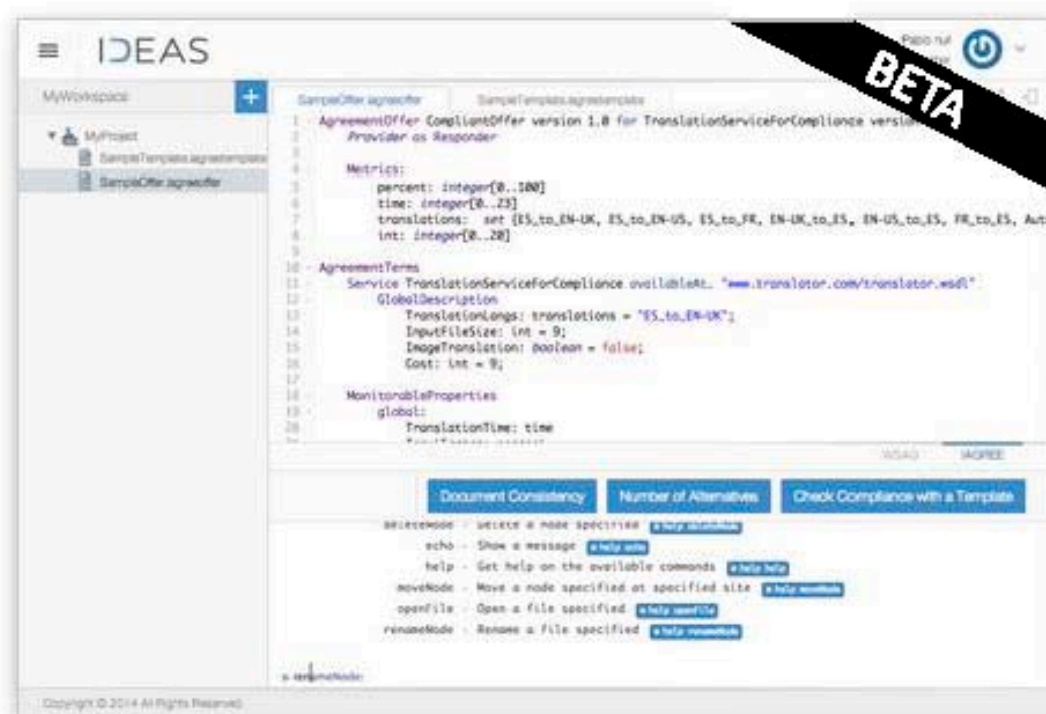
- Inconsistencies,
- Dead terms



Compliance

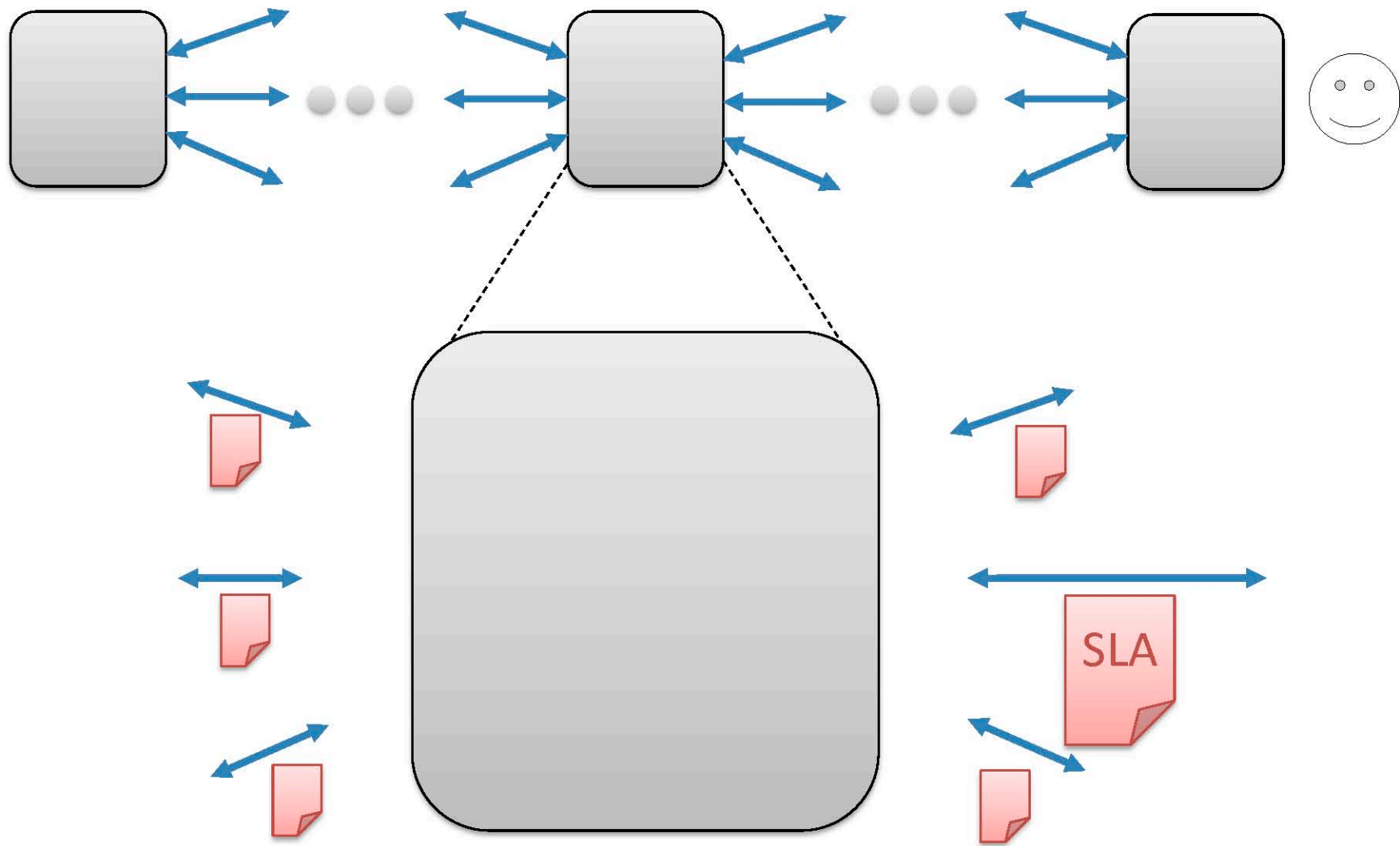


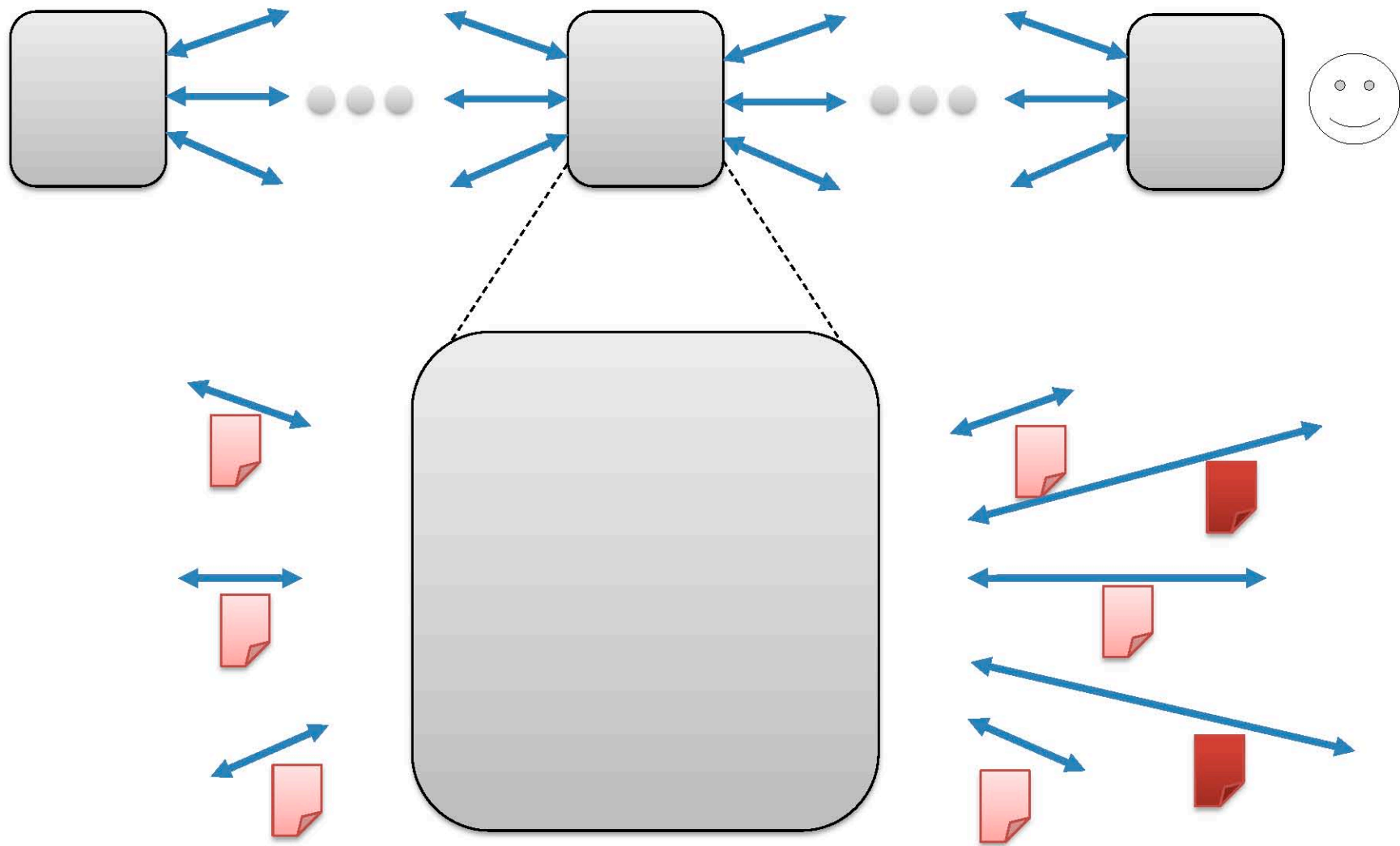


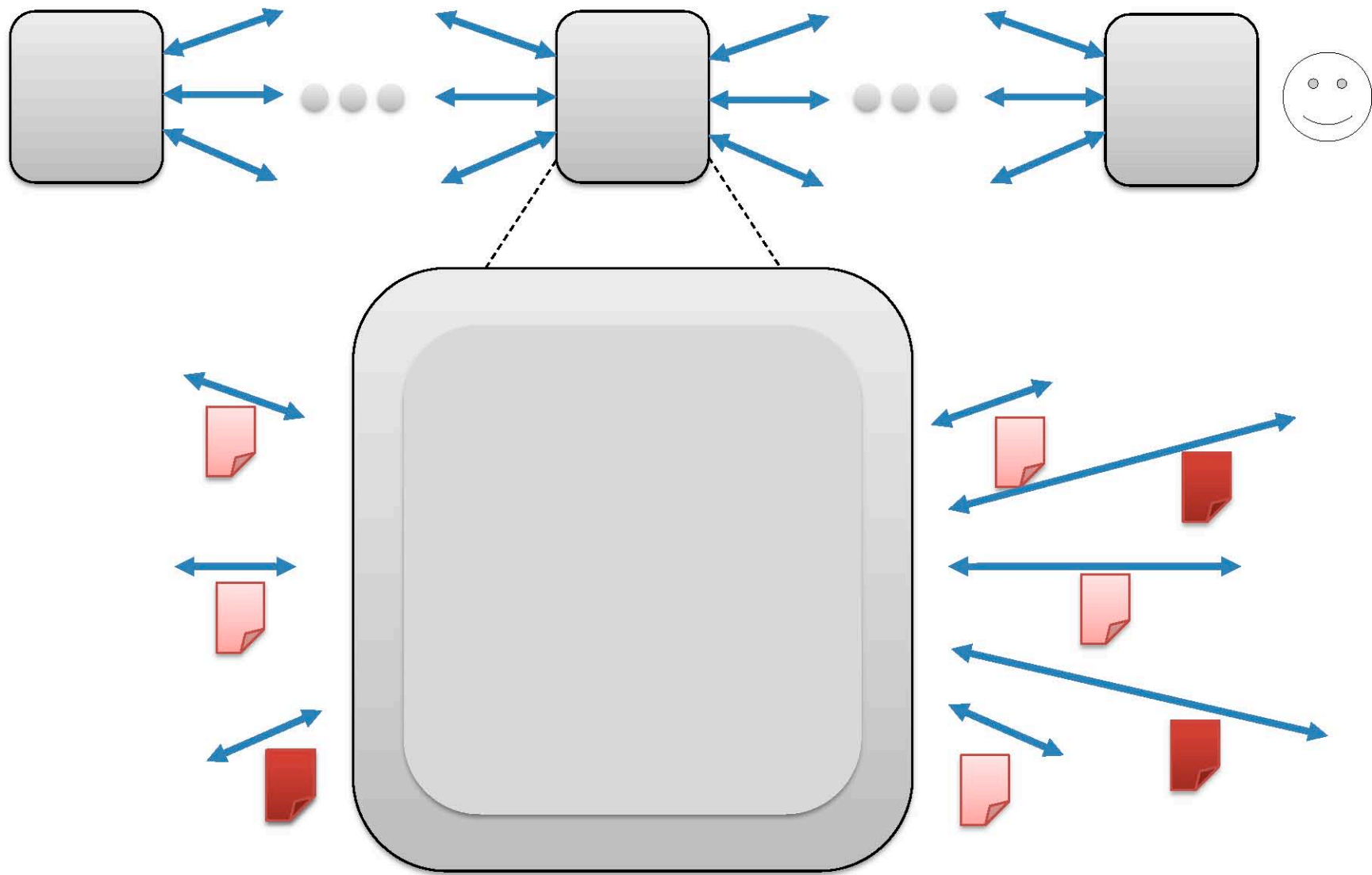


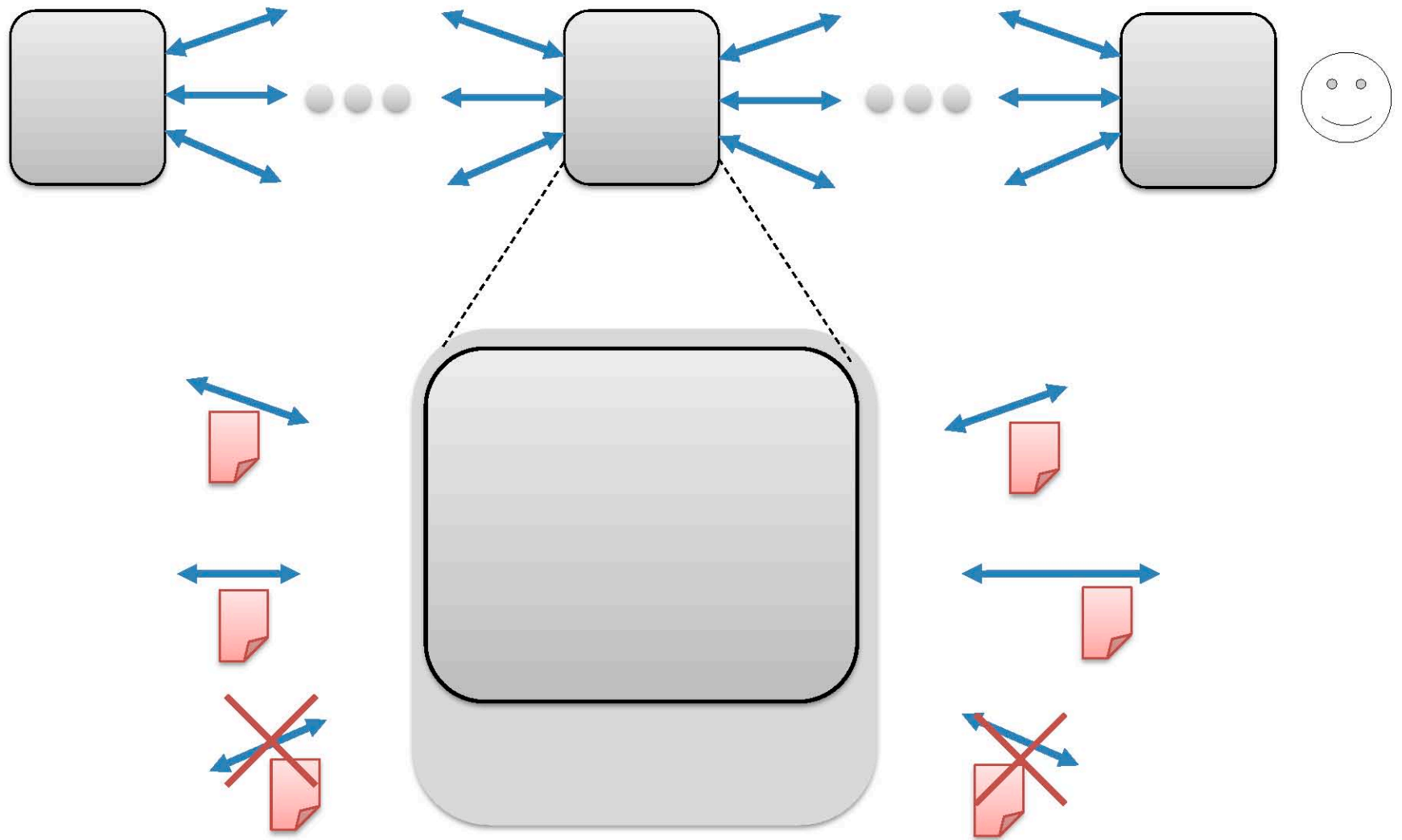
Available soon at
www.isa.us.es/IDEAS !

Foundations and applications of **SLAs** for Elastic Computing

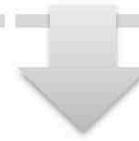
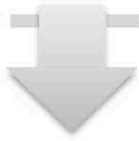








ELASTICITY



What is my risk?

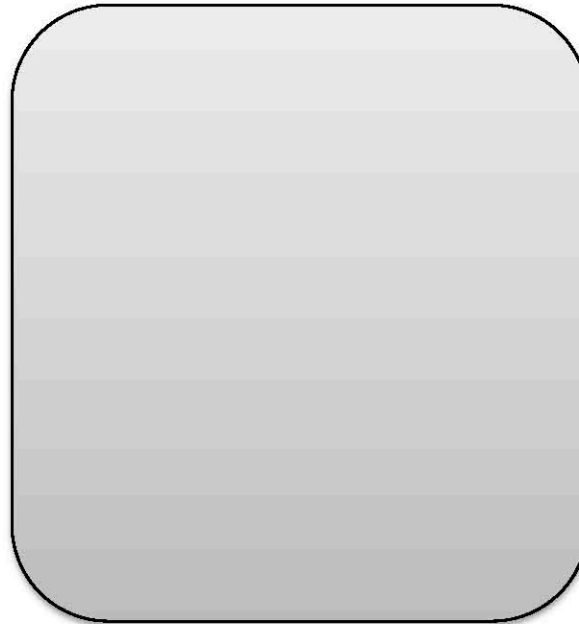
Can I be more competitive?

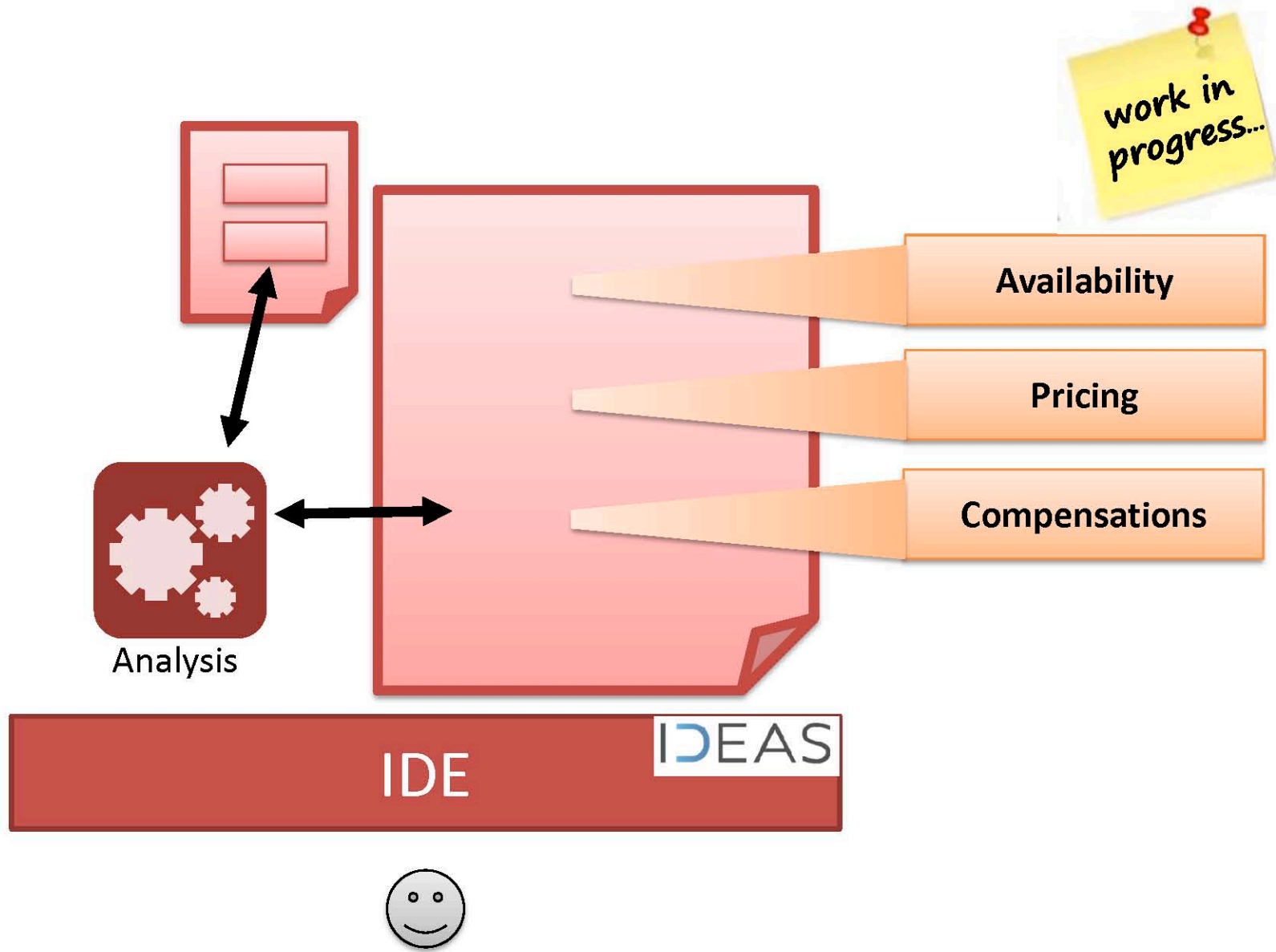
Should I change my provider?

Should I break my guarantees?

Am I using all the resources
that I paid for?

Can I accept more customers?





Availability

Amazon S3 SLA

Effective Date: June 1, 2013

This Amazon S3 Service Level Agreement ("SLA") is a policy governing the use of Amazon Simple Storage Service ("Amazon S3") under the terms of the Amazon Web Services Customer Agreement (the "AWS Agreement") between Amazon Web Services, Inc. ("AWS", "us" or "we") and users of AWS' services ("you"). This SLA applies separately to each account using Amazon S3. Unless otherwise provided herein, this SLA is subject to the terms of the AWS Agreement and capitalized terms will have the meaning specified in the AWS Agreement. We reserve the right to change the terms of this SLA in accordance with the AWS Agreement.

Service Commitment

AWS will use commercially reasonable efforts to make Amazon S3 available with a Monthly Uptime Percentage (defined below) of at least 99.9% during any monthly billing cycle (the "Service Commitment"). In the event Amazon S3 does not meet the Service Commitment, you will be eligible to receive a Service Credit as described below.

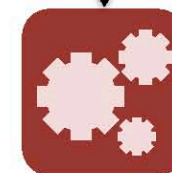
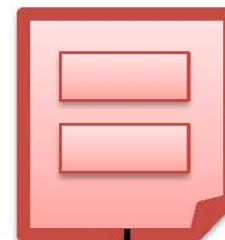
Definitions

- "Error Rate" means: (i) the total number of internal server errors returned by Amazon S3 as error status "InternalError" or "ServiceUnavailable" divided by (ii) the total number of requests during that five minute period. We will calculate the Error Rate for each Amazon S3 account as a percentage for each five minute period in the monthly billing cycle. The calculation of the number of internal server errors will not include errors that arise directly or indirectly as a result of any of the Amazon S3 SLA Exclusions (as defined below).
- "Monthly Uptime Percentage" is calculated by subtracting from 100% the average of the Error Rates from each five minute period in the monthly billing cycle.
- A "Service Credit" is a dollar credit, calculated as set forth below, that we may credit back to an eligible Amazon S3 account.

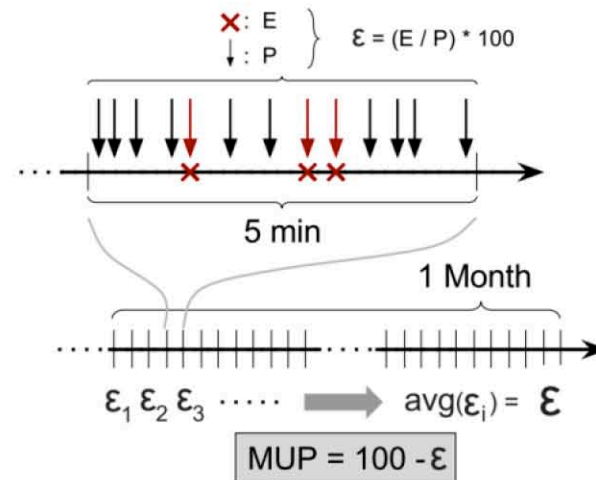
Service Credits

Service Credits are calculated as a percentage of the total charges paid by you for Amazon S3 for the billing cycle in which the error occurred in accordance with the schedule below.

Monthly Uptime Percentage	Service Credit Percentage
Equal to or greater than 99% but less than 99.9%	10%
Less than 99%	25%



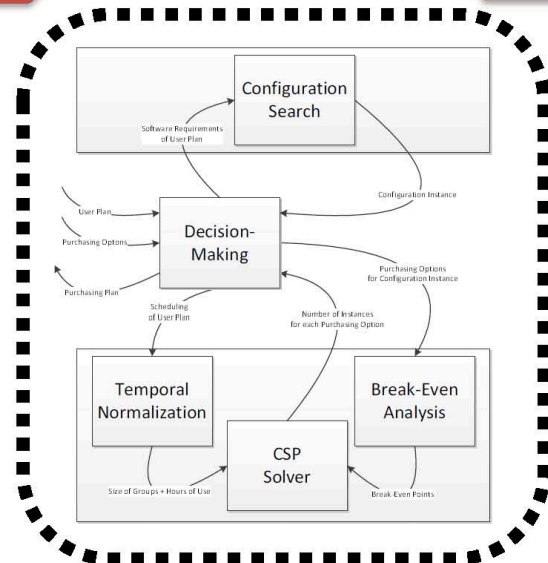
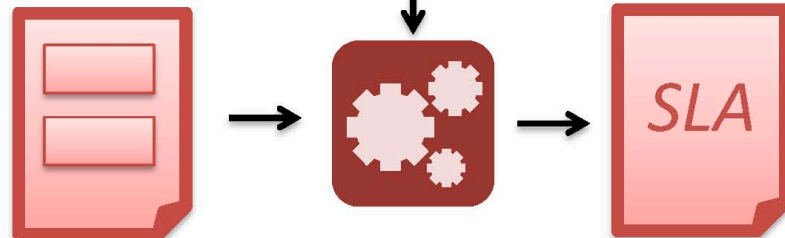
What if...?



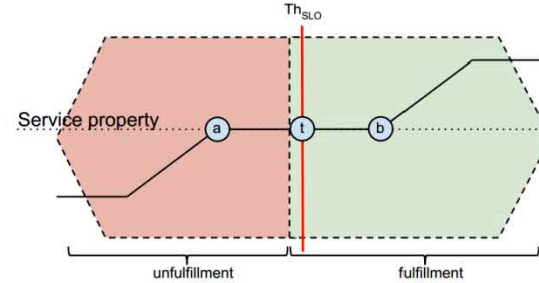
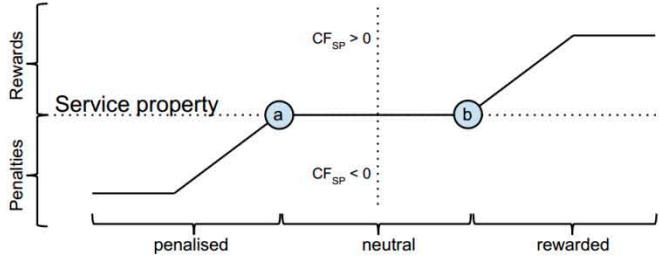
Pricing

On-Demand Instance Prices

Linux						
RHEL SLES Windows Windows with SQL Standard Windows with SQL Web						
Region:	US East (N. Virginia) US East (N. Virginia) US West (Oregon) US West (Northern California) EU (Ireland) Asia Pacific (Singapore) Asia Pacific (Tokyo) Asia Pacific (Sydney) South America (Sao Paulo)					
General Purpose	Memory (GiB)	Instance Storage (GB)	Linux/UNIX Usage			
m3.medium	3.75	1 x 4 SSD	\$0.070 per Hour			
m3.large	2	6.5	7.5	1 x 32 SSD	\$0.140 per Hour	
m3.xlarge	4	13	15	2 x 40 SSD	\$0.280 per Hour	
m3.2xlarge	8	26	30	2 x 80 SSD	\$0.560 per Hour	
Compute Optimized - Current Generation						
c3.large	2	7	3.75	2 x 16 SSD	\$0.105 per Hour	
c3.xlarge	4	14	7.5	2 x 40 SSD	\$0.210 per Hour	
c3.2xlarge	8	28	15	2 x 80 SSD	\$0.420 per Hour	
c3.4xlarge	16	55	30	2 x 160 SSD	\$0.840 per Hour	
c3.8xlarge	32	108	60	2 x 320 SSD	\$1.680 per Hour	
GPU Instances - Current Generation						
g2.2xlarge	8	26	15	60 SSD	\$0.650 per Hour	
Memory Optimized - Current Generation						
r3.large	2	6.5	15	1 x 32 SSD	\$0.175 per Hour	
r3.xlarge	4	13	30.5	1 x 80 SSD	\$0.350 per Hour	
r3.2xlarge	8	26	61	1 x 160 SSD	\$0.700 per Hour	
r3.4xlarge	16	52	122	1 x 320 SSD	\$1.400 per Hour	
r3.8xlarge	32	104	244	2 x 320 SSD	\$2.800 per Hour	
Storage Optimized - Current Generation						
i2.xlarge	4	14	30.5	1 x 800 SSD	\$0.853 per Hour	
i2.2xlarge	8	27	61	2 x 800 SSD	\$1.705 per Hour	
i2.4xlarge	16	53	122	4 x 800 SSD	\$3.410 per Hour	
i2.8xlarge	32	104	244	8 x 800 SSD	\$6.820 per Hour	
hs1.8xlarge	16	35	117	24 x 2048	\$4.600 per Hour	



Compensations



Type	Measurement	Penalty
Quarterly Status Report	Delivered at quarterly intervals and not less than five business days before scheduled review meeting	5% of monthly invoice

Sample 1 (page 5 of the template)

Severity Code	Initial Response	Estimation Response	Subsequent Responses	Resolution
1	15 minutes	2 hours	Every 30 min.	4 hours

Type	Measurement	Reward	Penalty
Severity 1 Resolution	All Severity 1 problems are resolved in less than 2 hours.	10% of monthly fees	NA
	One or more Severity 1 problems are resolved in over 4 hours.	NA	10% of monthly fees

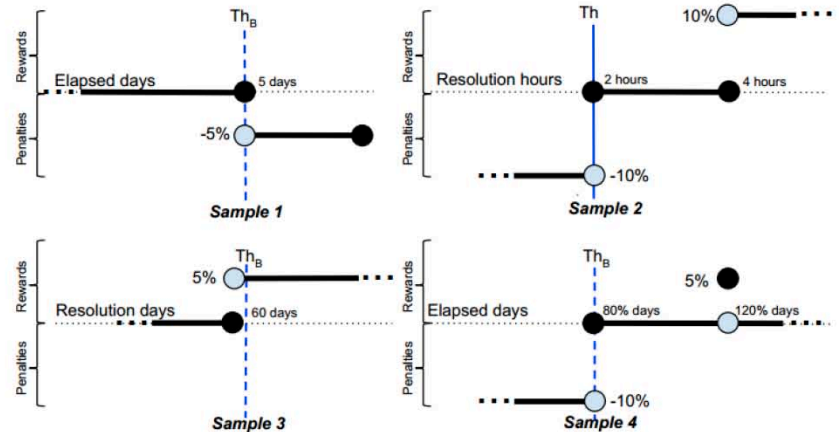
Sample 2 (page 7 of the template)

Type	Measurement	Reward	Penalty
Maximum Problem Aging	No problem is older than 60 days.	5% of monthly fees	NA

Sample 3 (page 8 of the template)

Type	Measurement	Reward*	Penalty
Project Delivery	Total elapsed days until delivery is more than 20% greater than planned.	NA	10% of the amount invoiced for the project.
	Total elapsed days until delivery is 20% less than planned.	5% of the amount invoiced for the project.	NA

Sample 4 (page 9 of the template)



Template/AgreementOffer/Agreement DocumentName version...

...
AgreementTerms

...
Guarantee Terms

```

guaranteeId: StakeHolder guarantees SLO; //GT id, obligated, and SLO
[upon scopedServiceOperations;           //serv. operation(s) affected
[onlyIf qualifyingCondition;              //qualifying condition
[with Interval CompensationType
[of amount if valueExpression;]*] * //penalties/rewards
    
```

ELASTICITY

What is my risk?

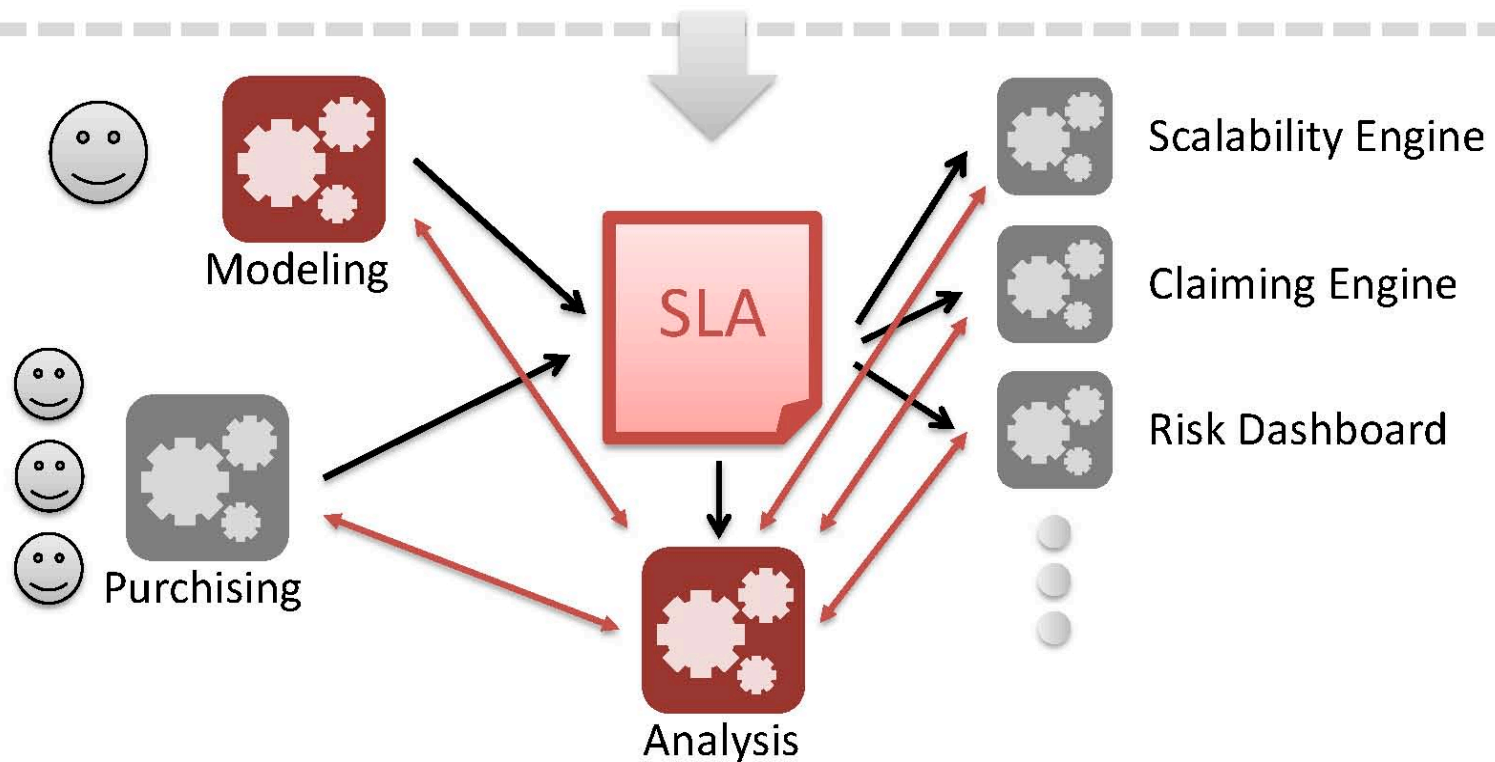
Can I be more competitive?

Should I change my provider?

Should I break my guarantees?

Am I using all the resources
that I paid for?

Can I accept more customers?



Ευχαριστώ!!
Thank you!!

