

# SLA-Aware Resource Management

QUB, Intel, XLAB

SLA@SOI Consortium

SLA in Grids Workshop 2009, Banff, Canada.  
October 13, 2009

Speaker: Philipp Wieder (on behalf of QUB, Intel, XLAB)  
Organisation: Dortmund University of Technology, Germany

**Motivation & Goal**

**Core Innovation**

**Architecture**

**SLA Negotiation**

**Resource Provisioning & Re-Provisioning**

**Monitoring**

**Current Status & Future Work**

**Acknowledgements**

# Motivation & Goal

## Service Consumer

- dynamic demand for complex business solutions at low costs

Flexible usage  
Business  
Services

## Service Provider

- service economy requires dependable services

Automated SLA  
negotiation and  
management

## Infrastructure Provider

- virtualization technologies allow for adaptive SOIs

SLA  
enforcement  
via adaptive  
infrastructures

## Software Provider

- SOAs provide unprecedented flexibility

Engineering of  
predictable  
services

## Vision of SLA@SOI

A business-ready service-oriented infrastructure empowering the service economy in a flexible and dependable way

# Motivation & Goal

## Service Consumer

- dynamic demand for complex business solutions at low costs

Flexible usage  
Business  
Services

## Service Provider

- service economy requires dependable services

Automated  
negotiation and  
management

## Infrastructure Provider

- virtualization technologies allow for adaptive SOIs

SLA  
enforcement  
via adaptive  
infrastructures

## Software Provider

- SOAs provide unprecedented flexibility

Engineering of  
predictable  
services

## SLA-Aware Resource Management

Service-oriented  
infrastructure  
empowering the  
service economy  
in a flexible and  
dependable way

**Motivation & Goal**

**Core Innovation**

**Architecture**

**SLA Negotiation**

**Resource Provisioning & Re-Provisioning**

**Monitoring**

**Current Status & Future Work**

**Acknowledgements**

## Adaptive SLA-aware infrastructures

- Standardized interfaces for adaptive infrastructures with harmonized access to different virtualization technologies.
- Advanced technologies for SLA enforcement on infrastructure level.
- Efficient resource usage with reliable SLA enforcement at infrastructure level.

**Motivation & Goal**

**Core Innovation**

**Architecture**

**SLA Negotiation**

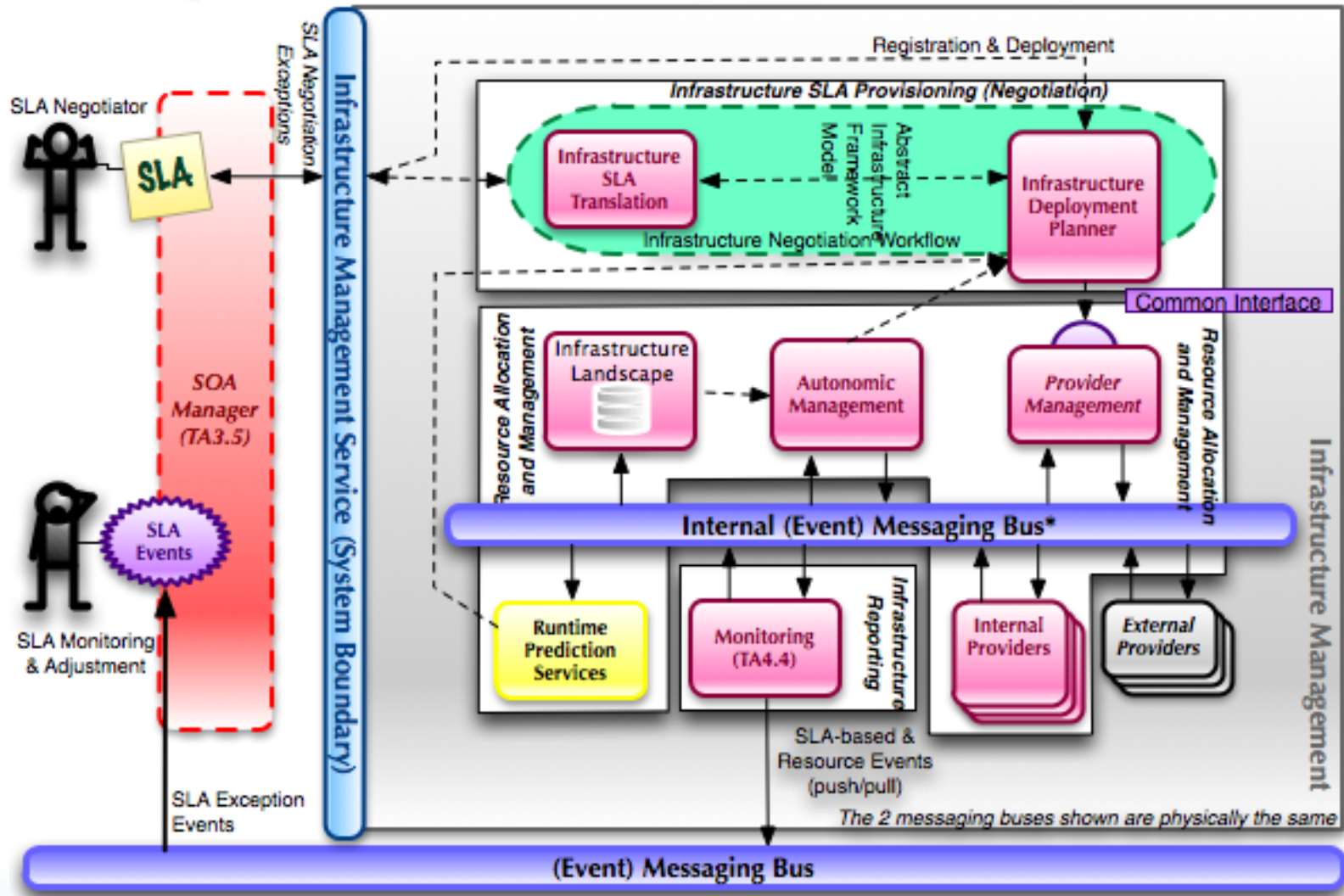
**Resource Provisioning & Re-Provisioning**

**Monitoring**

**Current Status & Future Work**

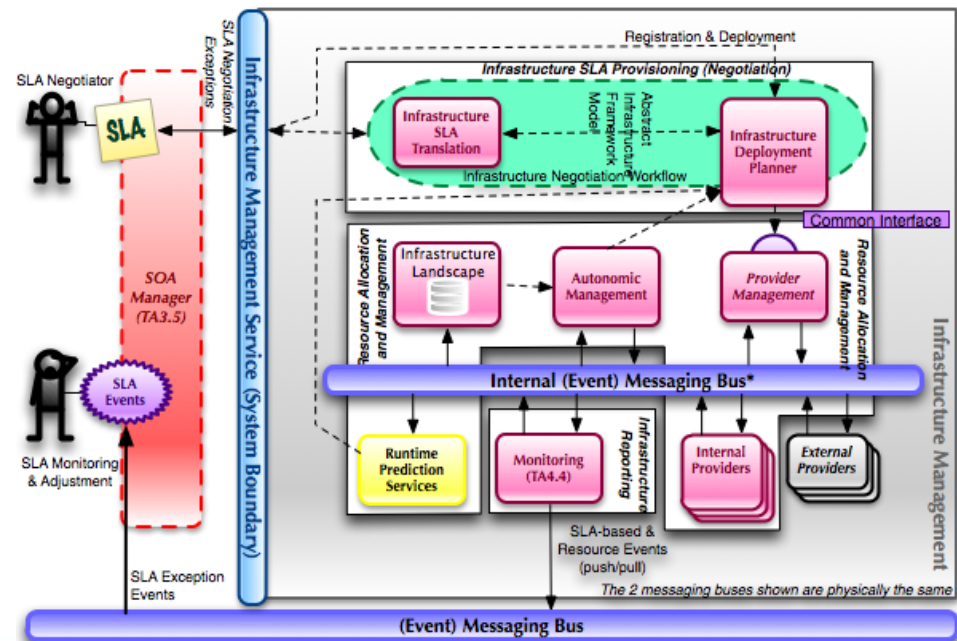
**Acknowledgements**

# An Overview of Architecture



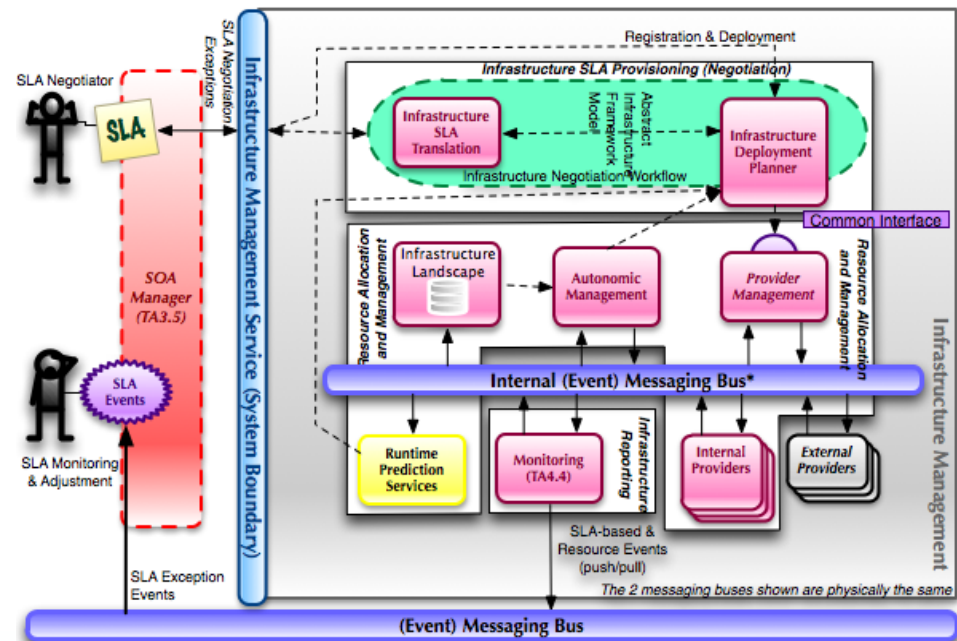
## Infrastructure Mgmt Service

- Customer facing interface for registration, provisioning, redeployment and management functionalities.



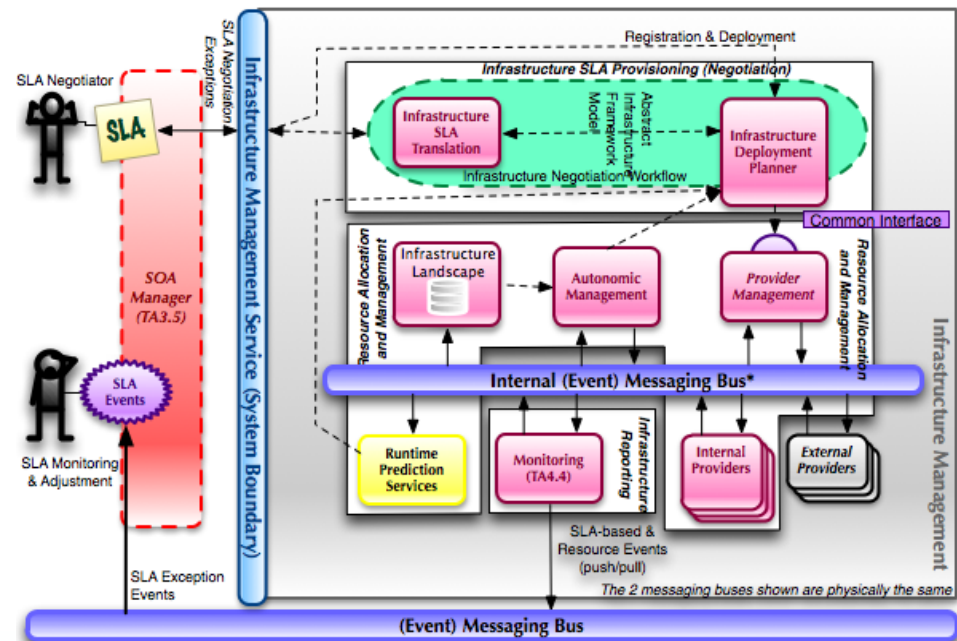
## Abstract Infrastructure Framework Model

- Describing required infrastructure resources.
- Accommodating infrastructure constraints defined in higher-level business SLAs.



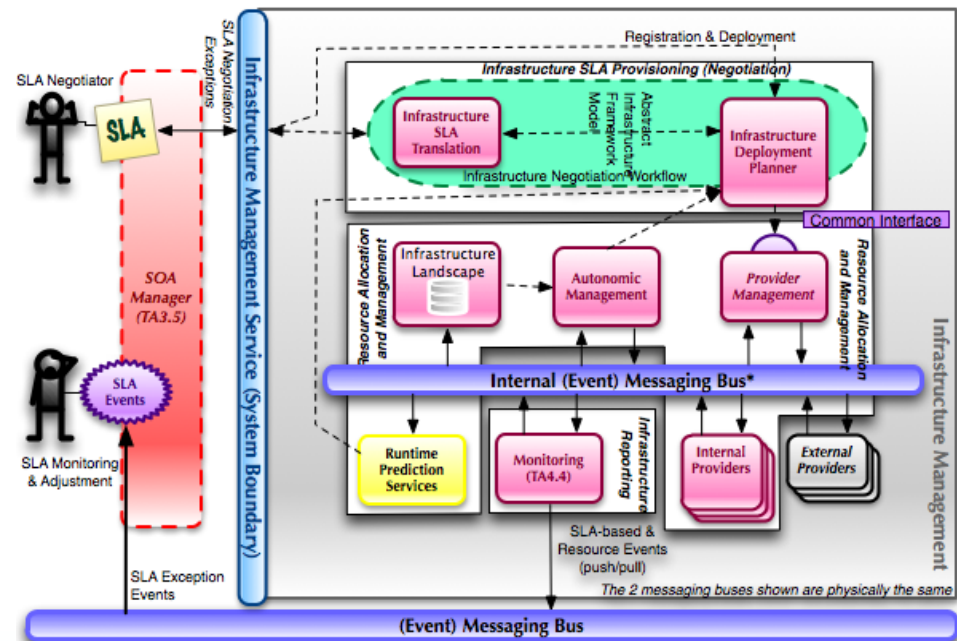
## Infrastructure SLA Translation

- Converts a provisioning request into an Abstract Infrastructure Framework representation.
- Potentially supports requests in multiple formats.



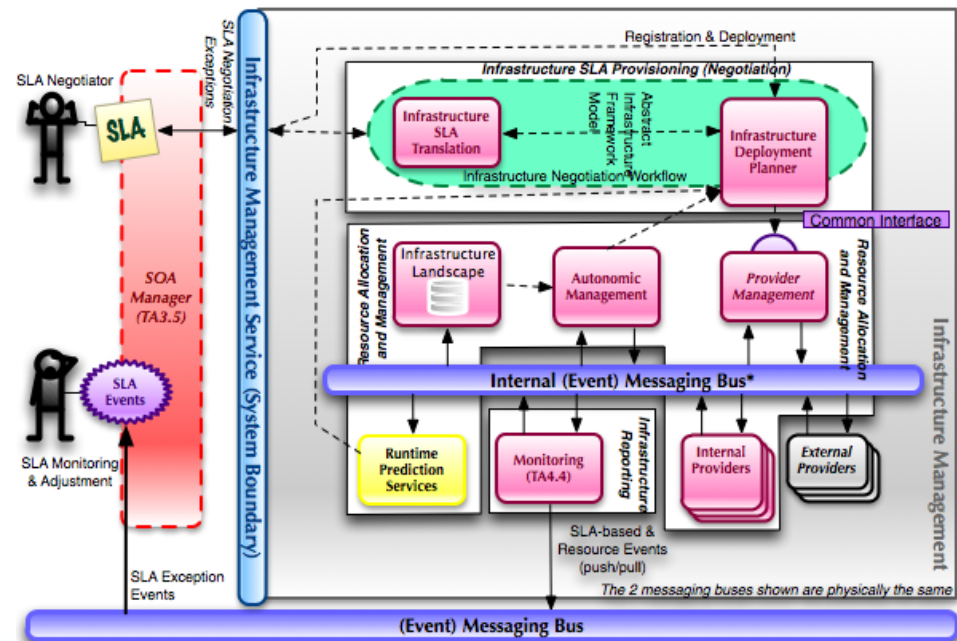
## Infrastructure Deployment Planner

- Analyses requests and converts them into individual virtual resources requirements and corresponding software images.
- Checks if resources can be provisioned.
- Reserves resources for short duration.



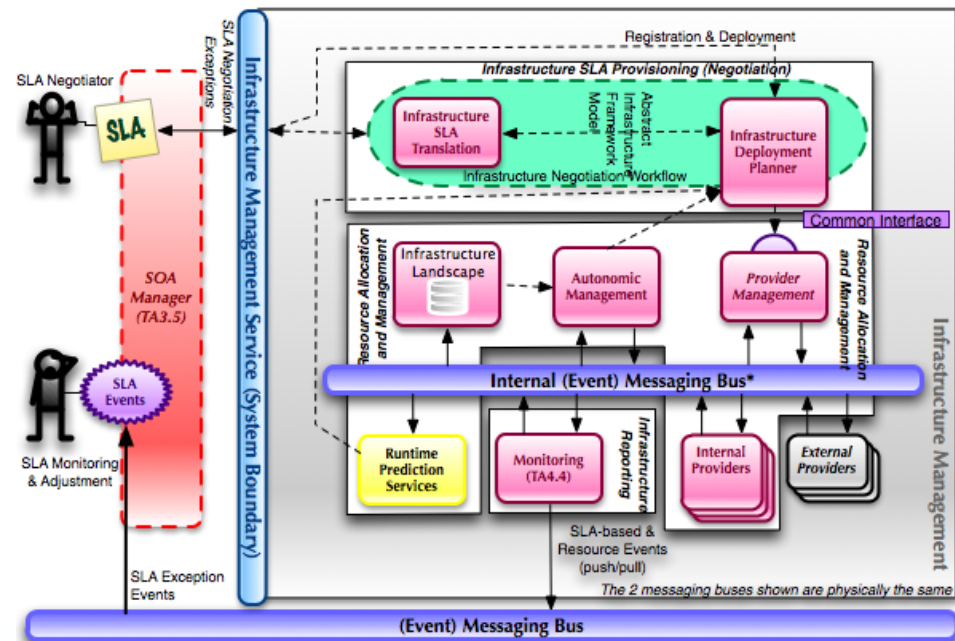
## Infrastructure Negotiation Workflow

- Uses the Infrastructure SLA Translation and Infrastructure Deployment Planner to see if resources can be provisioned.
- The customer may or may not decide to proceed with the provisioning request.



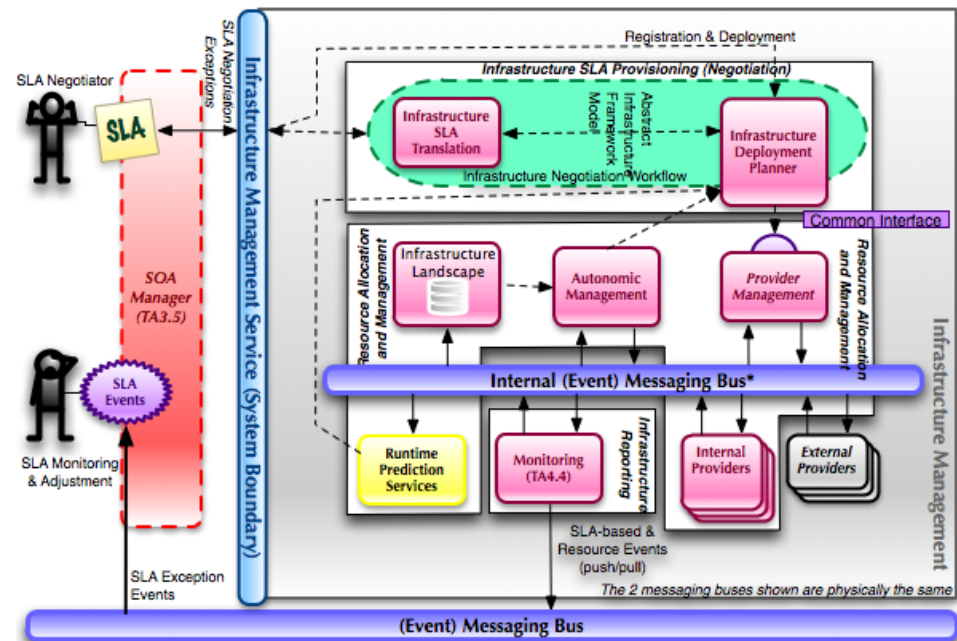
## Provider Mgmt

- A plug-in management system for communicating and controlling resource providers using a consistent abstracted interface.
- Performs provisioning and re-provisioning as required.



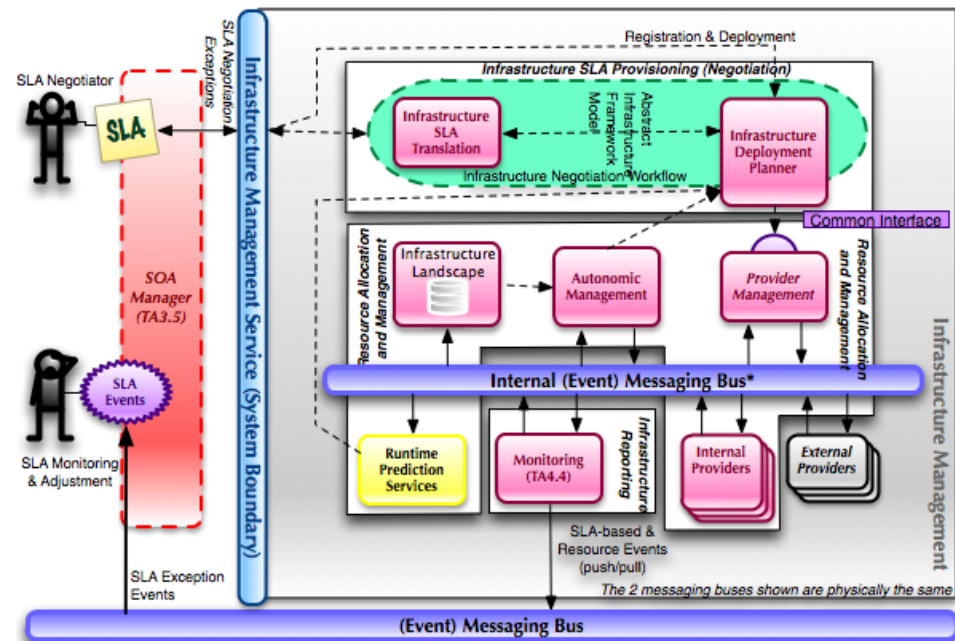
## Autonomic Mgmt (Optimisation)

- Sends a request to Deployment Planner to perform redeployment pre-emptively based on potential SLA violations identified by the Monitoring component.



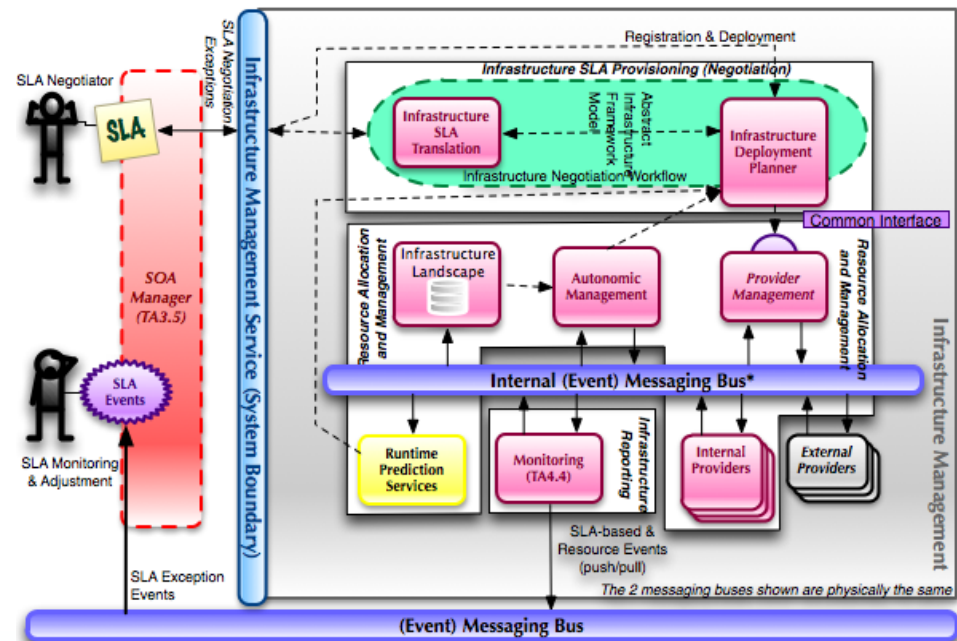
## Monitoring

- Receives events from internal or external providers.
- Standardises events and stores in a historical repository.
- Reviews the historical repository, correlates raw events, identifies escalations including potential and actual SLA violations



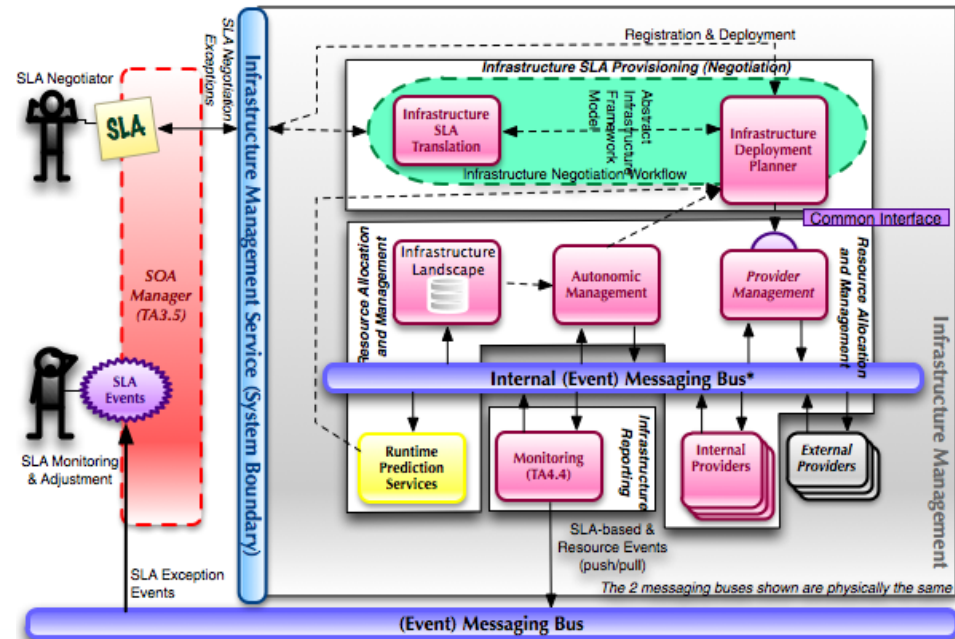
## Prediction Services

- Predicts actual resources required based on historical and any other information available.



## Infrastructure Landscape

- A representation of all currently running physical and virtual infrastructure resources which are under the control of the Infrastructure Provider
- Physical and Virtual infrastructure resources must be registered here upon activation.



**Motivation & Goal**

**Core Innovation**

**Architecture**

**SLA Negotiation**

**Resource Provisioning & Re-Provisioning**

**Monitoring**

**Current Status & Future Work**

**Acknowledgements**

## SLA Negotiation

- Based on WS-Agreement specification
- Infrastructure SLA (iSLA) {Terms, SLOs, Rules}
  - ◇ Terms
    - Resource specification, e.g. Virtual Machine
    - Includes functional and non-functional attributes
  - ◇ SLOs
    - Metrics to be monitored by Monitoring component.
  - ◇ Rules
    - Conditional actions to be taken upon stage change of one or more SLOs.
- iSLAs can be combined to form a request with multiple resource provisioning per agreement.
- Multi-round negotiation
  - ◇ Counter-offers.

**Motivation & Goal**

**Core Innovation**

**Architecture**

**SLA Negotiation**

**Resource Provisioning & Re-Provisioning**

**Monitoring**

**Current Status & Future Work**

**Acknowledgements**

## Provisioning

- Interact with arbitrary sources of infrastructure resources
  - ◇ Internal Provider, e.g. internal physical resources.
  - ◇ Remote (External) Provider, e.g. EC2, Flexiscale, etc.
- Harmonized and Abstract interface
  - ◇ Technology neutral interface to virtualization technologies.
  - ◇ Enables high level services to remain separate from details of underlying technologies that are used in physical infrastructure.
- Provider Management interface
  - ◇ To **find** suitable physical resource that satisfy the virtual resource SLA requirements.
  - ◇ To **reserve** a virtual resource on a physical resource.
  - ◇ To **instantiate** and start the virtual resource.
  - ◇ To **stop** the virtual resources.
  - ◇ To **re-adjust** or **re-provision** the virtual resources according to the SLA constraints.

## Live Migration

- A form of re-provisioning
  - ◇ Consolidation of services without affecting customer SLAs.
  - ◇ Relocation of virtual machines.
- Research experiments
  - ◇ Live migration takes as long as it takes for the memory of the running virtual machine to be transferred from one system to the other.
    - e.g. A virtual Machine with 512MB memory was migrated in approx. 49 seconds over a 100 Mbps network with almost no noticeable impact on the running services inside the virtual machine.

**Motivation & Goal**

**Core Innovation**

**Architecture**

**SLA Negotiation**

**Resource Provisioning & Re-Provisioning**

**Monitoring**

**Current Status & Future Work**

**Acknowledgements**

## Multi-layer monitoring architecture

- Layer 0 (Data Collection Layer)
  - ◇ Collection of raw input data.
  - ◇ Basic filtering and preprocessing of collected information.
- Layer 1 (Event Evaluation Layer)
  - ◇ Integration of monitors into a cascade of increasingly more complex monitors, ranging from simple metric checks to composed monitors.
- Layer 2 (Service Layer)
  - ◇ Collection of conceptually similar functions of L1 that provides the services used by any service dealing with the infrastructure.
  - ◇ Receives inputs from layers below it.
  - ◇ Configures and defines the meaning of monitoring events generated in lower layers.
  - ◇ Prevents top-level monitors from connecting to L0 and bypassing the L1.

**Motivation & Goal**

**Core Innovation**

**Architecture**

**SLA Negotiation**

**Resource Provisioning & Re-Provisioning**

**Monitoring**

**Current Status & Future Work**

**Acknowledgements**

## Current Status

- Preliminary proof of concept prototype was developed.
- It was realised that the core functionality concerning the SLAs of the infrastructure layer could be addressed independently of the core infrastructure of the provisioning layer.
- Decoupling allows high-level SLA modelling, management and negotiation concerns to be processed independently of the low-level infrastructure management components.

## Future Plans

- Separation of SLA management from resource provisioning system. → SLA management system can be used with any sort of provisioning system. e.g. Reservoir/OpenNebula, EC2, etc.
- Supporting storage and networking resources.
- Arbitrary customer groupings.
- Introducing software and service concepts in the infrastructure models.
- Identifying and submitting enhancements to OVF.
- Implementing OCCI interface into prototype provisioning layer.
- Providing SLA management for OCCI providers.

**Motivation & Goal**

**Core Innovation**

**Architecture**

**SLA Negotiation**

**Resource Provisioning & Re-Provisioning**

**Monitoring**

**Current Status & Future Work**

**Acknowledgements**

## Acknowledgements

- Supported by the SLA@SOI project, an European Commission's Seventh Framework Programme (FP7) under grant agreement FP7-216556 addressing Objective 1.2 SSAI "Service and Software Architectures, Infrastructures and Engineering".
- <http://www.sla-at-soi.eu>

**Thank  
you!**