ACROSS
Autonomous Control for a Reliable Internet of Services

ICT COST Action IC1304

Autonomous Control for a Reliable Internet of Services (ACROSS)

Task Force 4: Cloud Federations

Wojciech Burakowski, Attila Kertesz, Steven Latre , Patrick Gwydion Poullie and Maciej Sosnowski

cost
EUROPEAN COOPERATION IN SCIENCE AND TECHNOLOGY

ACROSS
Autonomous Control for a Reliable Internet of Services

TF4: Cloud Federations

Partners:

- Warsaw University of Technology, Poland
- University of Szeged and MTA SZTAKI, Hungary
- University of Antwerp-iMINDS, Belgium
- University of Zurich, Switzerland

and (maybe)

- Alexandru Ioan Cuza University of Iasi, Romania

ACROSS
Autonomous Control for a Reliable Internet of Services

TF4: Cloud Federations

Agenda:

- The objective of TF4
- Architectures for CFs
- Planned studied scenarios
- Models for CFs
  - On task flow level
  - On resource level
  - Model of resources
- Research topics
- Experiments possibilities
- Conclusions
-

ACROSS
Autonomous Control for a Reliable Internet of Services

TF4: Cloud Federations

The objectives:

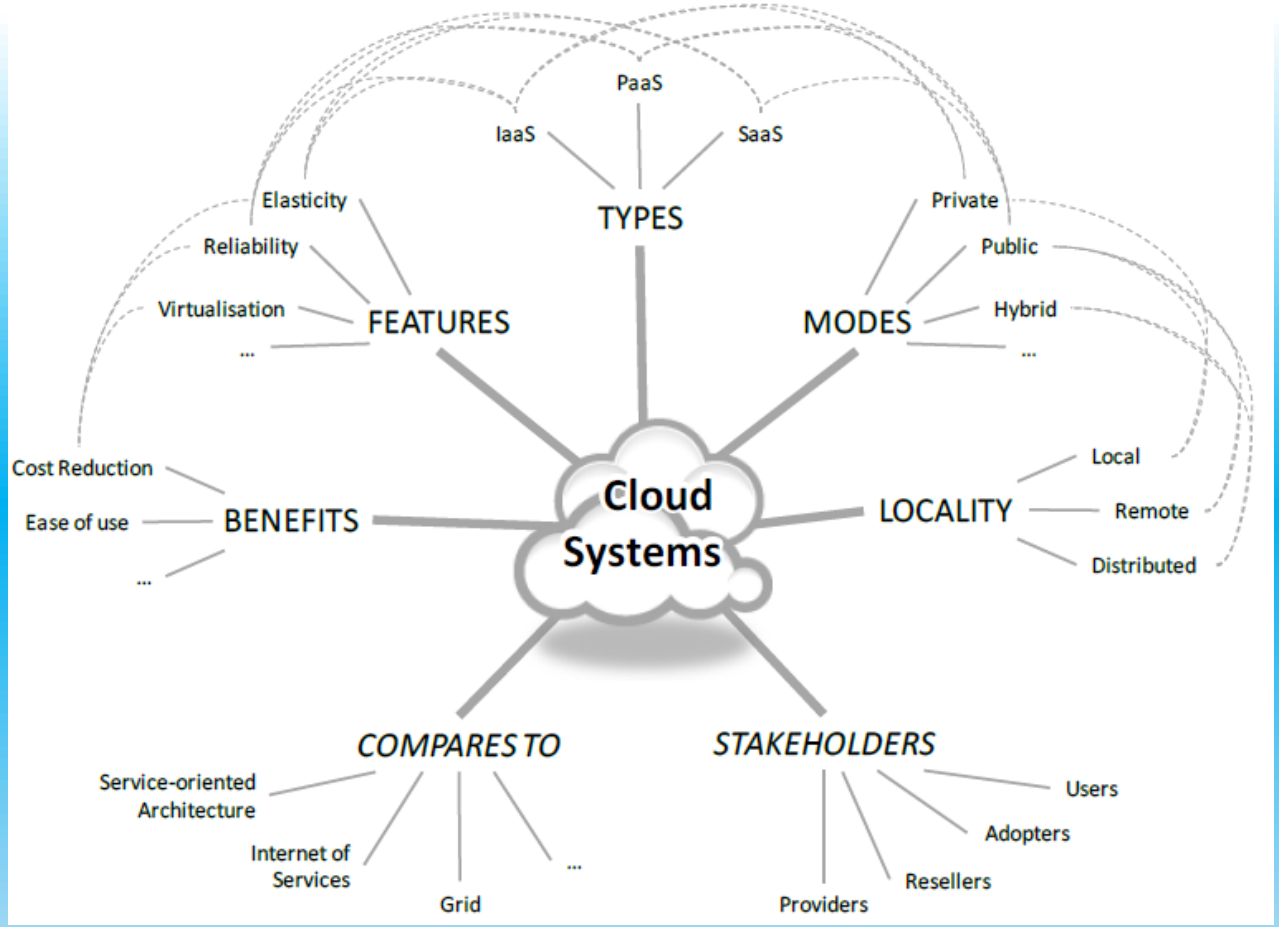
- To recognize important research topics corresponding to Cloud Federations
- To specify adequate models (for analysis, simulations and experiments) and traffic scenarios
- To present solutions
- To test solutions , if possible
- In addition
  - To enforce international cooperation (e.g.common papers, to establish european projects?)

TF4: Cloud Federations

Architecture:

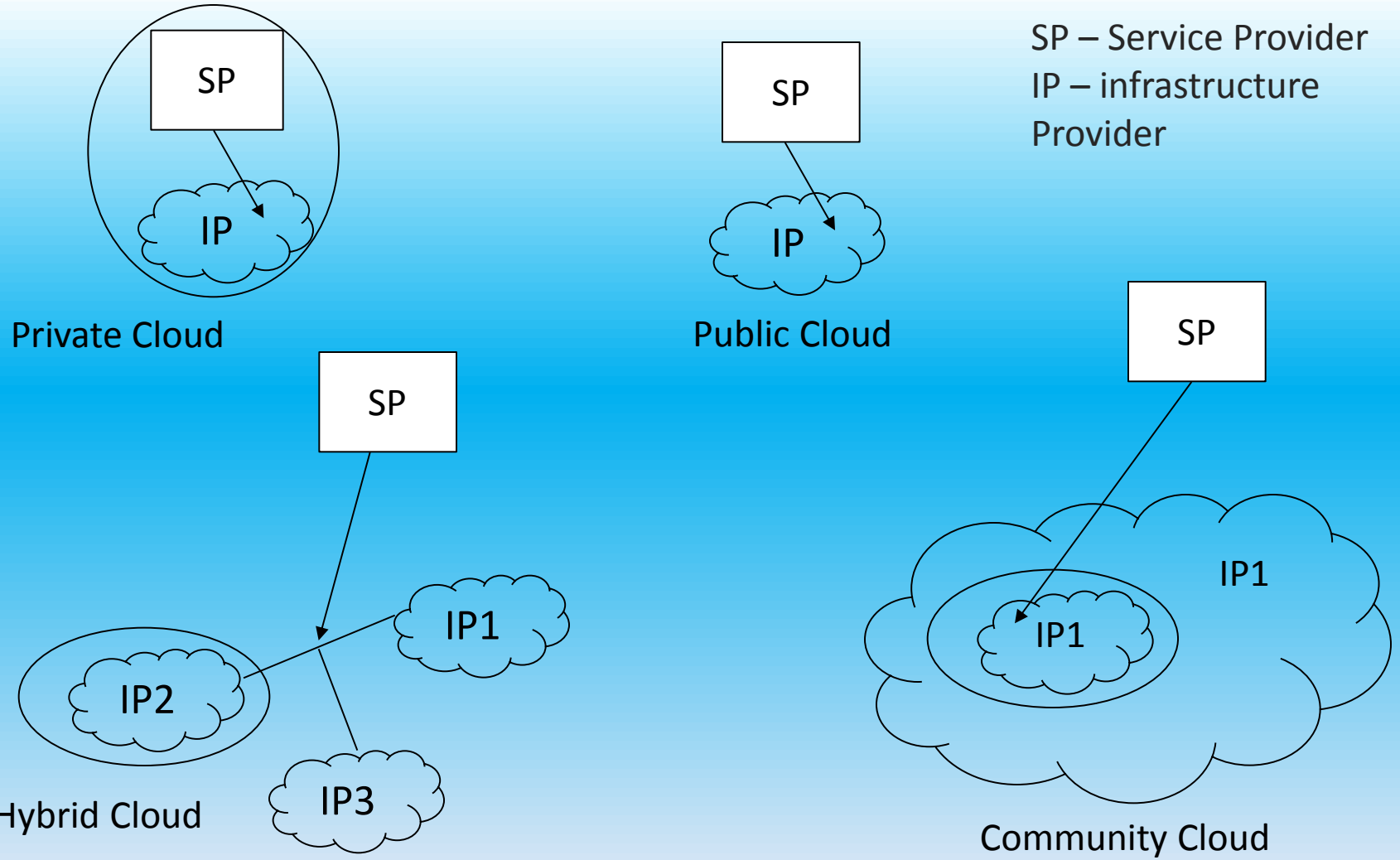
- Try to define a reference architecture we will assume for our studies

European Commission's definition of Clouds[2]



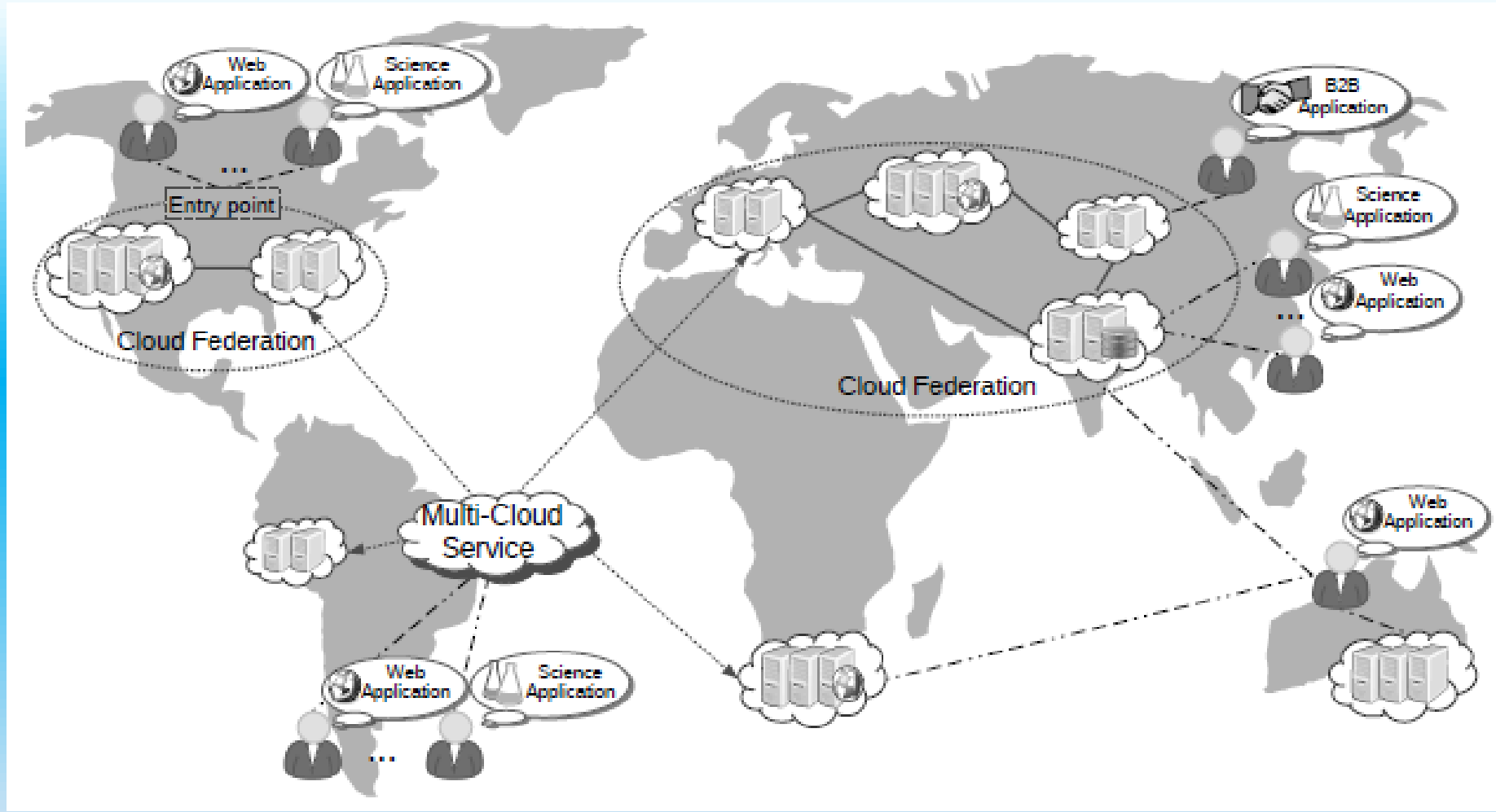
A 'cloud' is an elastic execution environment of resources involving multiple stakeholders and providing a metered service at multiple granularities for a specified level of quality.

Architectural view by the EC



2. K. Jeffery and B. Neidecker-Lutz, The Future of Cloud Computing, Opportunities for European Cloud Computing beyond 2010. Expert Group Report, January 2010.

Inter-Cloud overview[3]



3. Nikolay Grozev and Rajkumar Buyya, Inter-Cloud Architectures and Application Brokering: Taxonomy and Survey, Software: Practice and Experience,, Wiley Press, New York, USA, 2013.

Benefits of an Inter-Cloud

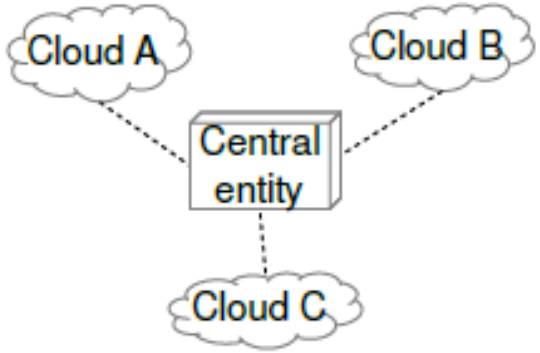
For *users*:

- Diverse geographical locations
  - Legislation compilant services: decide where data is stored
- Better application resilience
  - Avoid cloud service outages, multiple data centers for fault tolerance
- Avoidance of vendor lock in
  - Distribute workload, price-sensitive usage, easy migration

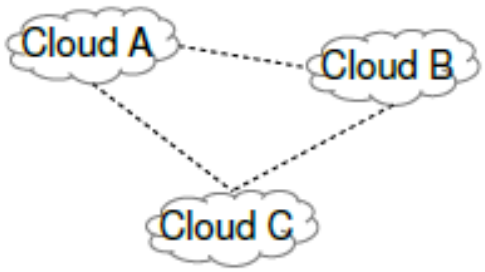
For *providers*:

- Expand on demand
  - Offload resource utilization to other providers
- Better SLA to customers
  - Support worst-case scenarios, survive outages

Federations

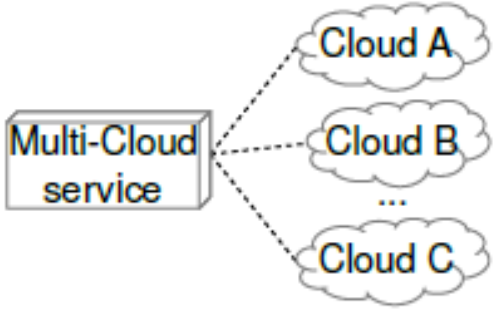


a) Centralised Inter-Cloud Federation. Clouds use a central entity to facilitate resource sharing.

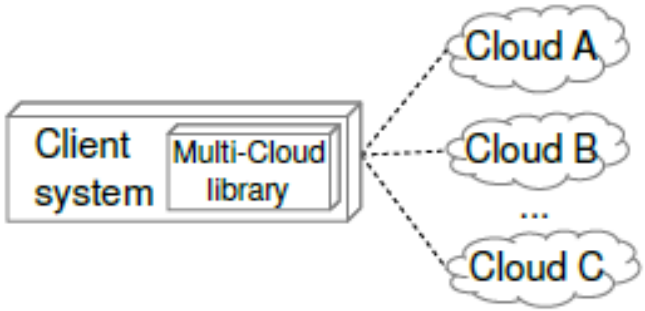


b) Peer-to-Peer Inter-Cloud Federation. Clouds collaborate directly with each other.

Multi-Clouds



c) Multi-Cloud Service. Clients access multiple clouds through a service.



d) Multi-Cloud Library. Clients develop their own brokers by using a unified cloud API in the form of a library.

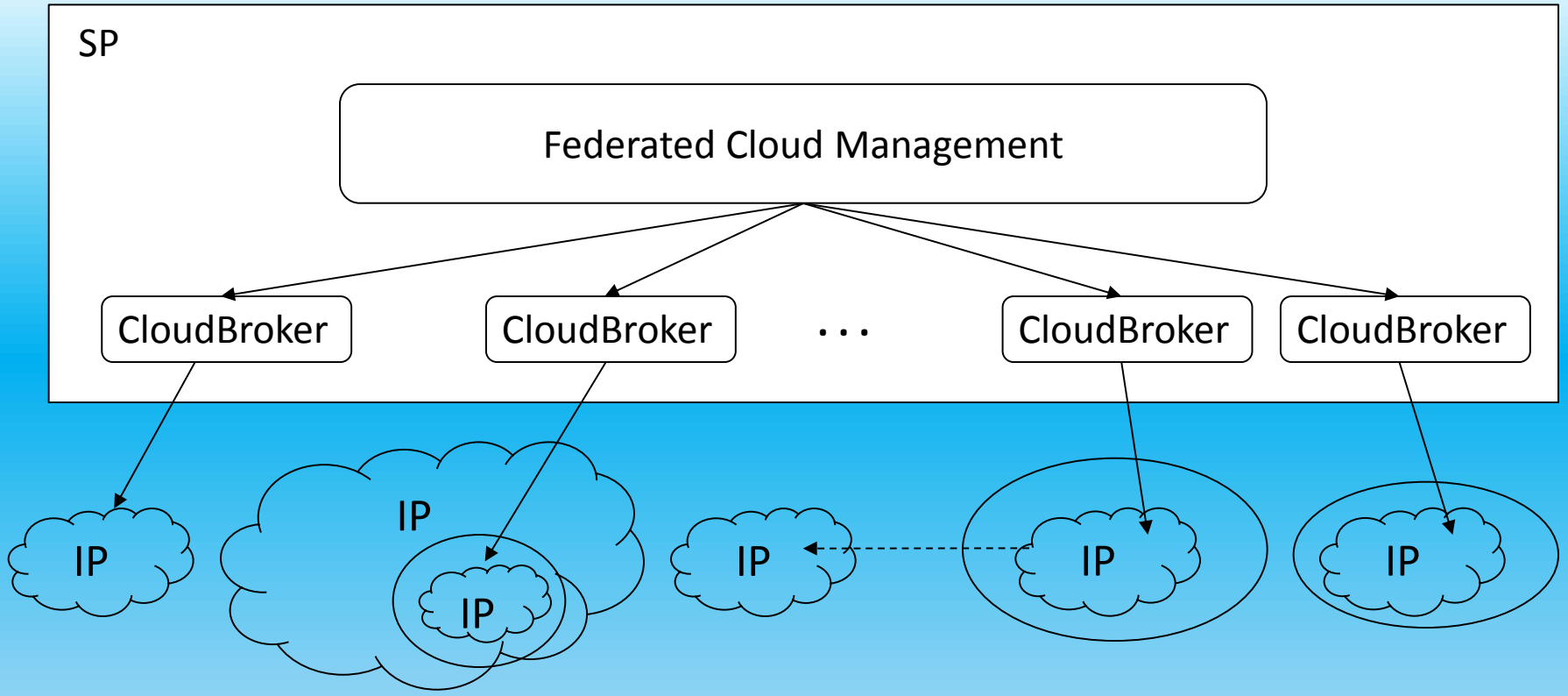
Inter-Cloud developments[3]

Inter-Cloud survey[3]

Project	Type, Organization	Architecture	Brokering Approach	Application Type	Awareness
InterCloud	Research project, University of Melbourne	Centralised federation	SLA based and Directly managed	Singular Jobs	Geo-location, Pricing
Contrail	Private and public European research organisations Funded by EU	Centralised federation and Independent Service	SLA based	Singular Jobs	Pricing
Dynamic Cloud Collaboration (DCC)	Academic research project supported by South Korean research funds.	Centralised federation	SLA based	Singular Jobs	Pricing
Federated Cloud Management (FCM)	Academic research project supported by EU funds.	Centralised federation	SLA based	Singular Jobs	Pricing
RESERVOIR	Private and public European research organisations Funded by EU	Peer-to-peer federation	SLA based and Trigger-Action	Singular Jobs	Pricing
Open Cirrus	Research testbed by academic and industry partners. Partially funded by US NSF.	Peer-to-peer federation	Directly managed	Singular Jobs	Data location
OPTIMIS	Private and public European research organisations Funded by EU	Peer-to-peer federation and Independent service	SLA based	Singular Jobs, Periodical Jobs, Compute and Data Intensive Interactive application	Pricing

Arjuna Agility	Commercially owned	Peer-to-peer federation	Trigger-Action	Singular Jobs, Periodical Jobs, Compute and Data Intensive Interactive application	Local resources
Global Inter-Cloud by Bernstein et al.	Publications are by people from miscellaneous companies - CISCO, Huawei Technologies, EMC Corporation	Peer-to-peer federation	SLA based	Singular Jobs, Periodical Jobs, Compute and Data Intensive Interactive application	Data location, Local resources
mOSAIC	Private and public European research organisations Funded by EU	Independent service	SLA based	Singular Jobs, Periodical Jobs, Compute and Data Intensive Interactive application	Pricing
STRATOS	York University. Supported by Canada's NSERC funds, Amazon and CA Inc.	Independent service	SLA based	Singular Jobs, Periodical Jobs, Compute and Data Intensive Interactive application	Geo-location, Pricing, Legislation/ Policy and Local resources
Commercial Cloud Management Systems (RightScale, EnStratus, Scalr, Kaavo)	Commercially owned	Independent service	Trigger-Action	Singular Jobs	Geo-location, Data location, Pricing, Legislation/ Policy and Local resources
Libraries (JClouds, LibCloud, DeltaCloud, SimpleCloud, Apache Nuvem)	Open source projects	Multi-Cloud libraries	Directly managed	Singular Jobs, Periodical Jobs, Compute and Data Intensive Interactive application	Geo-location, Data location, Pricing, Legislation/ Policy and Local resources

Generalized FCM (covering all presented scenarios)



Two levels of brokering:

- at the federated level
- at single cloud level

Title of Activity or Paper

Planned scenarios for using Cloud Federations:

- **Scenario 1**: each Cloud belonging to Cloud Federation has its own customers, which regard it as Home Cloud. So, such Cloud at the first wants to provide service to these customers using its own resources.
  - If this cloud is saturated then it looks for resources from other clouds from the Federation.
  - If this cloud has spare resouces, then it want to rent them
- **Scenario 2**: The request for service is generated by a customer to Cloud Federation. Cloud federation treats each customer in the same way. It means that such Cloud Federation manages allocation of resources to a given request in a “fairness” way taking into account all Clouds

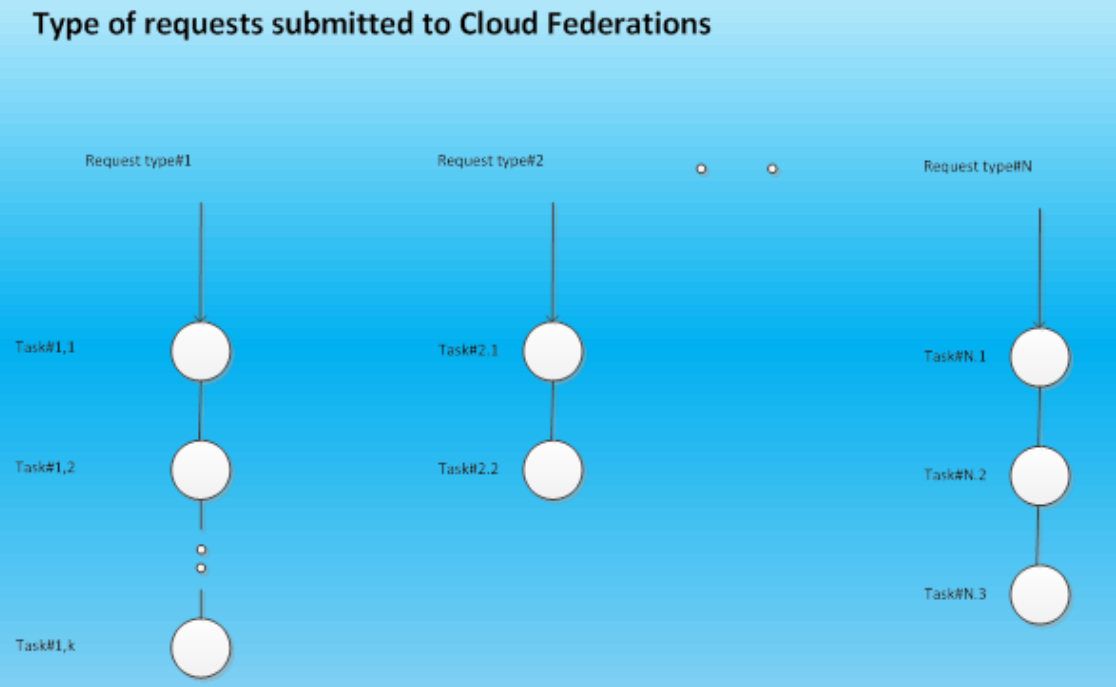
TF4: Cloud Federations

Models on task flow level:

- Two types of requests for services:
  - Type#1: Requests on demand for existing services (on-line)
  - Type#2: Requests for provisioning new services (off-line) I

TF4: Cloud Federations

Requests of type#1 (on-line): for offered services

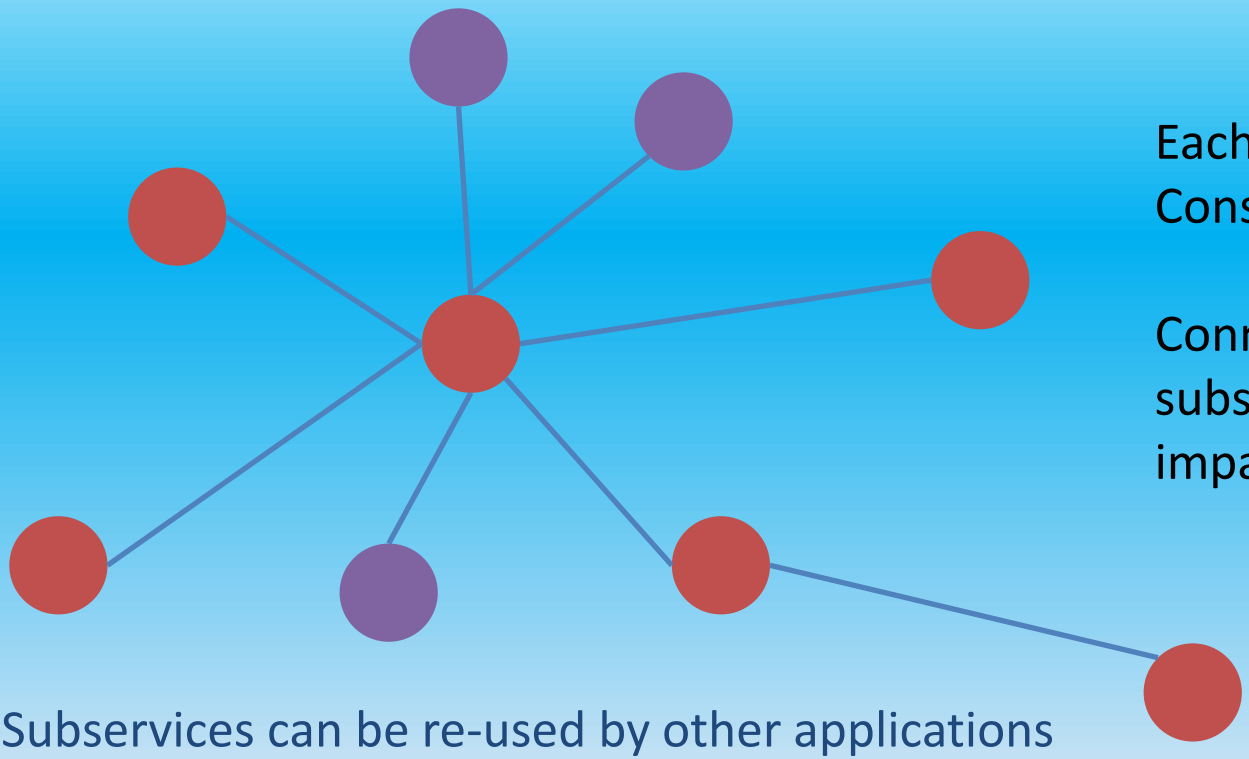


- Seuence of tasks (a bit simplified description)
- Request rate per task sequence
- Description of each task: sequence number, type of resource the task can be executed

TF4: Cloud Federations

Requests of type#2: for provisioning new services (offline)

- Requests for virtual machine configurations
- Application can be decomposed into multiple subservices



Each subservice will Consume resources

Connections between subservices also have an impact on the resources

Subservices can be re-used by other applications

TF4: Cloud Federations

Models of cloud resources:

- Type of resources:
  - Compute resources

Characterization of resources offered by Clouds

Compute resources are packed in racks and organized in clusters

- Rely heavily on parallelization of resources
  Need high-bandwidth network access
  - Gigabit Ethernet
  - Infiniband
- Two main aspects
  - Network topology
  - QoS

Metrics

- Throughput (Mbps)
- Latency
- Packet loss

Characterization of resources offered by Clouds

Compute resources are often connected to shared storage infrastructure

- Type of database system
  - ACID with limited scale-out (SQL)
  - BASE with unlimited scale-out (NoSQL)
- Has an impact on
  - Database size
  - Elasticity
  - Consistency

Metrics

- Consistency speed
- Scale-out factor
- Size

Characterization of resources offered by Clouds

Data centers consume energy

- Almost 50% of costs relate somehow to energy
- Prices may vary over time
- Prices may vary across locations
- Often less explicitly defined.

Metrics

- KWh
- Share of renuwable energy

ACROSS
Autonomous Control for a Reliable Internet of Services

TF4: Cloud Federations

Research topics:

- Task allocation
- Resource sharing
- Firness
- Virtual Network Infrastructure for Cloud Federations

TF4: Cloud Federations

Task allocation:

- Task allocation(when we perform a service):
  One of the essential problem in Cloud Federations is to make optimal task allocation to the resources in order to meet some requirements for service quality and resource usage.
- We need to specify the rules for task allocation to resources

Planned research topics: Task allocation (when we provision service)

Dynamic environment which is constantly changing

- New applications coming
- Existing applications being removed
- Existing applications change

application

demands

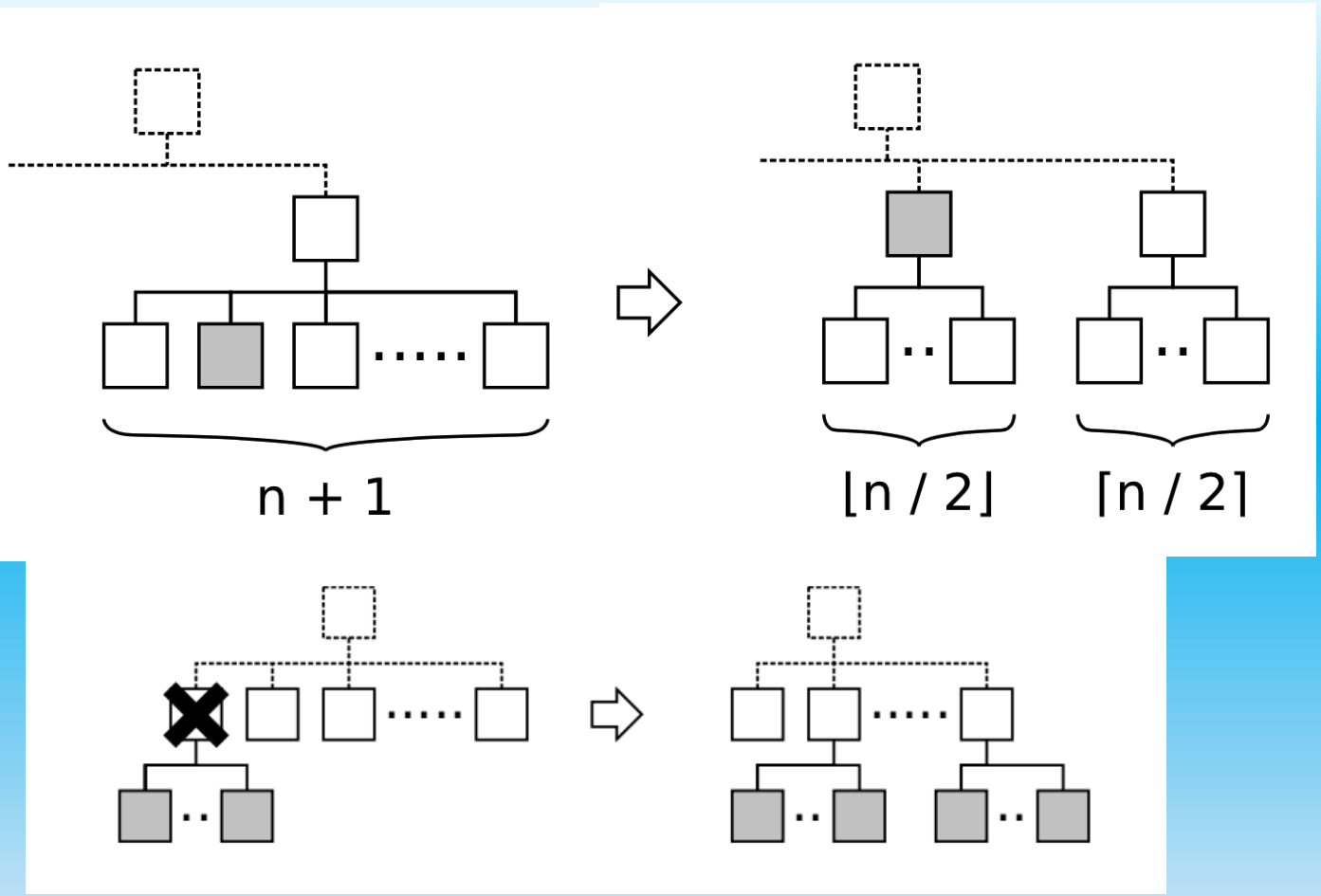
server infrastructure

capacity

Task allocation (when we provision service) : previous work

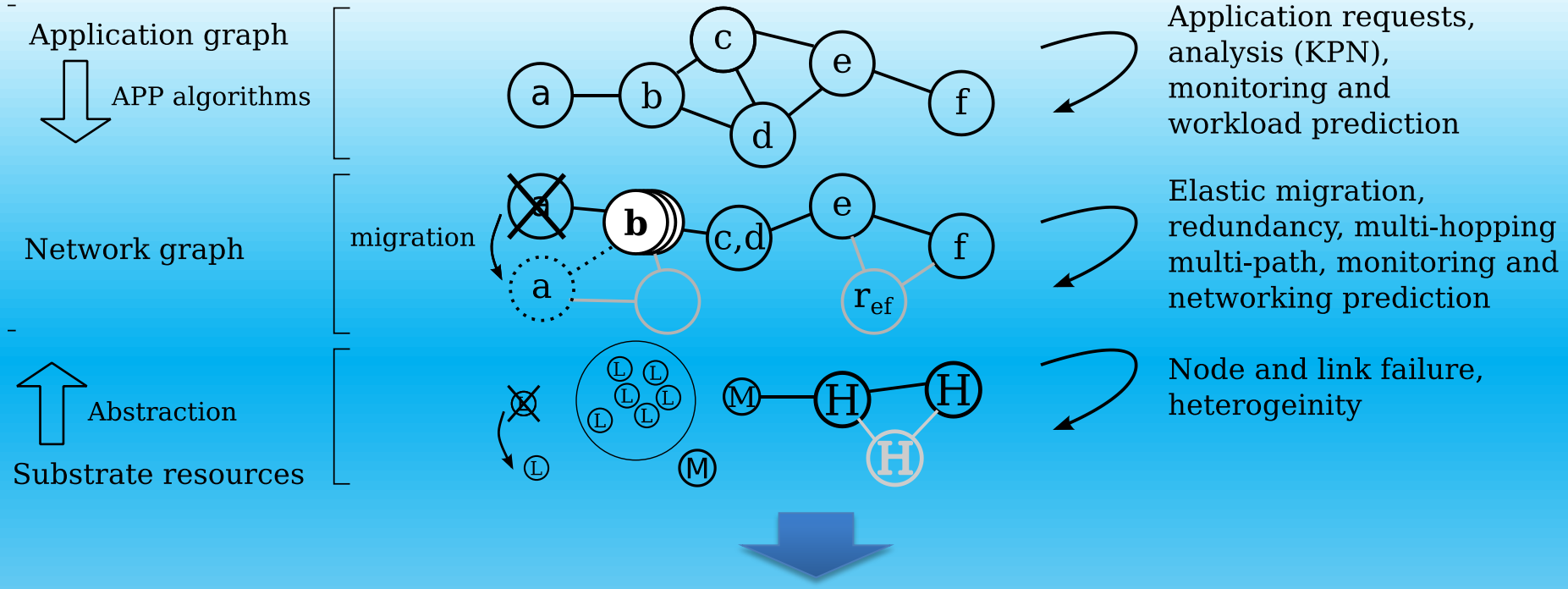
Maintaining a tree of peers that manage the application placement

not suited for federated environments



Moens, Hendrik, et al. "Design and evaluation of a hierarchical application placement algorithm in large scale clouds." *Integrated Network Management (IM), 2011 IFIP/IEEE International Symposium on*. IEEE, 2011.

Task allocation in federations



Substrate resources on different sites

Different sites have

- Important differences in capabilities
- Important differences in reliability
- Important differences in connectivity

TF4: Cloud Federations

Resource sharing:

- Interdependency of host resources, e.g., a VM utilizes RAM slower, when it receives less CPU
- Which dependencies exist and how to describe them mathematically?
- Hitherto no consideration for dependencies or overly simplified assumptions e.g., Leontief utility functions
- Investigation by executing workloads on different VM configurations
- Combined with different stress tests
- Monitoring resource consumption over time

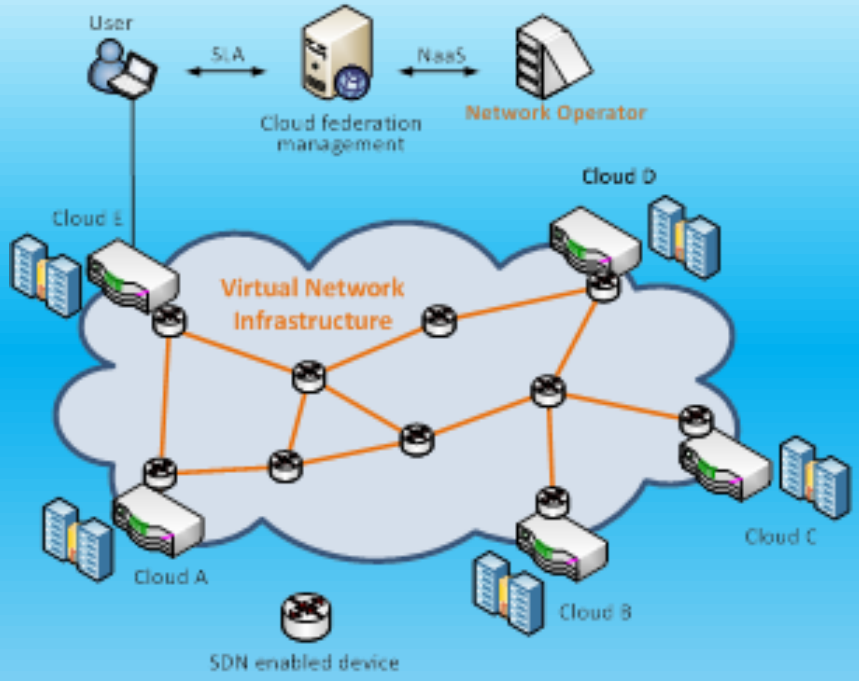
TF4: Cloud Federations

Fairness:

- Straight forward definition for single resource: Max-Min fairness Clouds
- Heterogenous resources on two levels
  - Different types of physical machines
  - Different types of host resources, e.g., CPU, RAM
- Heterogenous resources make share values subjective
- No access to utility functions
- Envy-freeness (economic concept) not applicable
- How to define fairness in clouds?
- Relevant for private clouds, where resource allocation is not prescribed by SLAs

TF4: Cloud Federations

Virtual Network Infrastructure for Cloud Federations :



- What traffic is carried by the network
- Which Classes of Services are needed, if needed
- To find algorthm to network dimension and adaptation to current needs

TF4: Cloud Federations

Testing capabilities:

- CloudSim for simulations
- European test-beds for experiments
  - Fed4Fire (more detailed information during next meeting)
  - OPEN CIRRUS - a federated testbed of distributed clusters for systems and applications research

ACROSS
Autonomous Control for a Reliable Internet of Services

TF4: Cloud Federations

Conclusion:

- The TF4 just started
- We plan to finish „white paper” (next week) and to upload it on the server
- We still discuss research topics
- We are open for new partners and co-operation with other TFs: e.g. with TF3 for specification of task sequences and rules for tasks allocation at the provisioning and service execution levels.

Any suggestions are welcome