SUPERFLUIDITY

a super-fluid, cloud-native, converged edge system

Research and Innovation Action GA 671566

Deliverable D2.1:

Use cases, technical and business requirements

Deliverable Type:	Report
Dissemination Level:	Public
Contractual Date of Delivery to the EU:	01/04/2016
Actual Date of Delivery to the EU:	01/04/2016
Workpackage Contributing to the Deliverable:	WP2

Editor(s): Philip Eardley (BT)

Author(s): Carlos Parada (ALB), Francisco Fontes (ALB), Isabel Borges(ALB), Omer Gurewitz (BGU), Philip Eardley (BT), Okung Ntofon (BT), Andy Reid (BT), Giuseppe Bianchi (CNIT), Nicola Blefari Melazzi (CNIT), Pierpaolo Loreti (CNIT), Claudio Pisa (CNIT), Stefano Salsano (CNIT), George Tsolis (CITRIX), Michael J. McGrath (Intel), Filipe Manco (NEC), Felipe Huici (NEC), Saverio Niccolini (NEC), Lionel Natarianni (NOKIA-FR), Bessem Sayadi (NOKIA-FR), Erez Biton: (NOKIA-IL), John Thomson (OnApp), Julian Chesterfield (OnApp), Michail Flouris (OnApp), Anastassios Nanos (OnApp), Haim Daniel (Red Hat), Livnat Peer (Red Hat), Pedro de la Cruz Ramos (Telcaria Ideas S.L.), Juan Manuel Sánchez Mateo (Telcaria Ideas S.L.), Raúl Álvarez Pinilla (Telcaria Ideas S.L.), Emilio Mino (Telefónica, I+D), Raquel García (Telefónica, I+D), Pedro A. Aranda (Telefónica, I+D), Ignacio Berberana (Telefónica, I+D), Dirk Griffioen (Unified Streaming), Rufael Mekuria (Unified Streaming), Costin Raiciu (UPB), Walner Saint-Aubin (EBlink)

Internal Reviewer(s): Carlos Parada (ALB), Stefano Salsano (CNIT), Pedro de la Cruz Ramos (Telcaria Ideas S.L.), Pedro A. Aranda (Telefónica, I+D)

Abstract: This document presents use cases and requirements for a “superfluid” network, which is one that will have the ability to instantiate services on-the-fly, run them anywhere in the network (core, aggregation, edge) and shift them transparently to different locations. It is an output of the Superfluidity collaborative research project.

The report includes a comprehensive set of 23 use cases, covering themes such as wireless access, mobile edge computing and on-the-fly monitoring. Each use case is described, along with its business and technical requirements.

The report takes into account our on-going architecture work, which will be described in a later document. At this stage we provide some initial thoughts on how the requirements, arising from the use cases, will impact on our technical work.

The intended readership for this document is quite general – other researchers into future networks, IT and their convergence; business leaders who will decide what use cases should be invested in; and engineers who will make the superfluid world a reality.

Keyword List: Superfluidity, use cases, requirements

Version Control Table			
Version N.	Purpose/Changes	Author	Date
1	First release	Superfluidity project	30 March 2016
1.1	Final revision	Nicola Blefari Melazzi	31 March 2016

Executive Summary

A “superfluid” network will have the ability to instantiate services on-the-fly, run them anywhere in the network (core, aggregation, edge) and shift them transparently to different locations. Such capabilities are a key part of the converged cloud-based 5G future - they will enable innovative use cases in the mobile edge, empower new business models and allow almost instant roll-out of new services, and reduce investment and operational costs.

Today’s networks suffer from several shortcomings, including a lack of service agility (to enable the flexible, rapid and tailored creation of new services), a lack of implementation agility (the reliance on rigid, cost-ineffective hardware devices with long provisioning times), and increasing complexity to cope with ever-growing scale and heterogeneity of the traffic, the services and the hardware technologies.

So there are a large number of potential use cases for a superfluid network, covering themes such as wireless access, mobile edge computing and on-the-fly monitoring. By studying these use cases we can better understand both business requirements, such as service agility and cost savings, and technical requirements, for example quality of experience and scalability.

Whilst use cases are interesting in themselves, they also help guide the technical work of the project. We have not attempted a strict ‘top down approach’, where the architecture would be derived in a formal process from the requirements of the various use cases. Instead we are using an iterative approach where the requirements analysis takes into account our on-going architecture work. A later document will describe the Superfluidity architecture – at this stage we provide some initial thoughts on how the requirements, arising from the use cases, impact our technical work. We include business requirements, such as service agility and cost savings, and technical requirements, for example quality of experience and scalability, as well as architectural concepts like reusable functional blocks.

INDEX

List of Figures

List of Abbreviations

ABR	Adaptive Bit Rate
API	Application Program Interface
AR	Augmented Reality
BBU	Baseband Unit
CDN	Content Distribution Network
CRUD	Create , Read , Update , Delete
DDoS / DoS	(Distributed) Denial of Service
DPI	Deep Packet Inspection
DRM	Digital rights management
ETSI ISG	European Telecommunications Standards Institute Industry Specification Group
GGSN / P-GW	Gateway GPRS Support Node / Packet Gateway
HARQ	Hybrid Automatic Repeat Request
IDPS	Intrusion Detection and Protection System
IETF	Internet Engineering Task Force
IoT	Internet of Things
ISA	Instruction Set Architecture
KPI	Key Performance Indicator
LCM	Life Cycle Management
LIPA-SIPTO	Local IP Access and Selected IP Traffic Offload
LTE	Long Term Evolution
LTM	Late TransMuxing
M2M	Machine to Machine
MEC	Mobile Edge Computing
MIMO	Multiple-Input Multiple-Output
NFV	Network Function Virtualization
NFVI	NFV Infrastructure
NDP	Neighbour Discovery Protocol
NIC	Network Interface Card

OSS	Operational Support Systems
OTT	Over-the-top
PDP/PDN	Packet Data Protocol/Packet Data Network
PoC	Proof of Concept
QoE	Quality of Experience
QoS	Quality of Service
RAN	Radio Access Network
RFB	Reusable Functional Block
RNC	Radio Network Controller
RRH	Remote Radio Head
RRM	Radio Resource Management
SDN	Software Defined Networking
SDO	Standard Defining Organization
SFC	Service Function Chaining
S/Gi	Reference point defined by 3GPP between the mobile packet core and PDN (Gi is between GGSN and PDN; SGi is between P-GW and PDN)
SLA	Service Level Agreement
SON	Self-Organising Network
TCP	Transmission Control Protocol
UE	User Equipment
UMTS	Universal Mobile Telecommunications System
vCS	virtual Convergent Services
vHGW	virtual Home GateWay
VM	Virtual Machine
VFN	Virtual Network Function
VNFC	Virtual Network Function Component
VNFM	Virtual Network Function Manager

1 Introduction

Today's networks suffer from several shortcomings, including: long provisioning times, with wasteful over-provisioning used to meet variable demand; reliance on rigid and cost-ineffective hardware devices; daunting complexity emerging from the heterogeneity of the traffic, the services and the hardware technologies; doubts about whether they can scale sufficiently to cope with the imminent Internet of Things; inflexibility from a single ownership business model; and a lack of agility to cope elegantly with disaster scenarios or to enable the flexible, rapid and tailored creation of services.

The Superfluidity project believes that the starting point to overcome these shortcomings is virtualisation: network function virtualisation, software defined networking and more generally 'softwarisation' of the network. We want to take this further in several new ways: the ability to extend virtualisation into domains such as radio access; the capability to instantiate on-demand chains of small re-usable functions in order to create the specific network service required; graceful inclusion of multiple vendors, whether they're supplying hardware, a network function or a service; naturally recursive solutions, both vertically (layering) and horizontally (domain-by-domain); a balance of both hardware and virtualised functions, optimised for the particular functionality required; and so on.

The goal of this document is to outline some use cases where such flexibility, which we call "Superfluidity", can help the future 5G world and to derive technical and business requirements for the Superfluidity architecture. The identified use cases fall into a number of overlapping areas:

- The wireless access part of the network – where for example we want to allow software-defined wireless fronthauling and radio resource management; and to support the Internet of things (with its difficult scaling and power requirements) as well as more traditional network services
- Monitoring – on-the-fly monitoring is needed in order to understand how to optimise the placement and performance of the various network functions, whilst minimising the overhead of regular monitoring
- Services delivered from the edge – for instance localised processing and storage; 'mobile edge computing' to improve content distribution, adaptation and personalisation; delivery of services or content from a local node, where this is possible; and enabling the network to perform services traditionally done on the end-user's device (such as ad removal)
- Security – there are several new threats arising from the extra flexibility, from the extra components (such as hypervisors and containers), and from virtualisation itself.

Whilst use cases are interesting in themselves, we use them to help guide the technical work of the project. We have not attempted a strict 'top down approach', where the architecture would be derived in a formal process from the requirements of the various use cases. Instead we are using an iterative approach where the requirements analysis takes into account our on-going architecture work. A later document will describe the Superfluidity architecture – at this stage we provide some initial thoughts on how the requirements, arising from the use cases, impact our technical work. We include business requirements, such as service agility and cost savings, and technical requirements, for example quality of experience and scalability, as well as architectural concepts like reusable functional blocks.

The document is structured as follows:

- Section 2 briefly describes the approach we followed to develop the use cases and requirements
- Section 3 details our 23 use cases and the business and technical requirements of each
- Section 4 outlines how we “clustered” the output of Section 3 into a more manageable number of categories
- Section 5 takes each of the ten “clusters” of requirements and discusses its potential impact on Superfluidity’s on-going technical work
- The Appendices provide some further details on state-of the-art use cases and requirements (Appendix A and B), and the approaches to clustering that we tried (Appendix C).

The intended readership for this document is quite general – other researchers into future networks, IT and their convergence; business leaders who will decide what use cases should be invested in; and engineers who will make the superfluid world a reality.

Overall Superfluidity aims to help achieve a converged cloud-based 5G future that will enable innovative use cases in the mobile edge, empower new business models and allow almost instant roll-out of new services, and reduce investment and operational costs.

2 Our approach

In this section we outline our approach.

First we reviewed use cases from the external world, as well as in the Superfluidity technical proposal:

- ETSI ISG on NFV (Network Function Virtualisation)
- ETSI ISG on MEC (Mobile Edge Cloud)
- 3GPP's document on "Feasibility study on new services and markets technology enablers"
- IETF working group on SFC (Service Function Chaining) – use cases for mobile networks
- SDx Central's use cases
- Cisco has a public report on the use cases it foresees
- The eight use case samples which we proposed earlier in Superfluidity's technical annex (i.e. the project proposal)

These are all detailed in the Appendix.

Based on this prior work, and enhanced and expanded by knowledge about what uses cases sparked the most interest within Superfluidity's Partner organisations, we generated a consolidated list of 23 use cases for further study. Section 3 describes each use case and lists some of the technical and business requirements associated with each one.

Since many of the use cases have significant features in common, it makes sense to "cluster" them in order to identify their key features. We looked at several potential ways of "clustering" these use cases and their requirements, as discussed in Section 4.

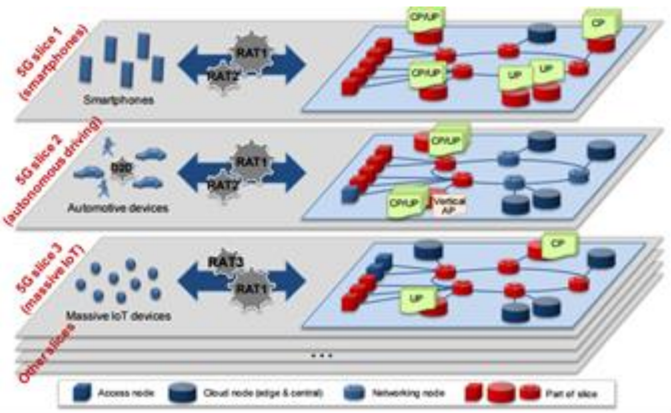
We next considered the impact of these "clusters" of requirements (associated with the Superfluidity use cases). In Section 5, for each "cluster" we discuss its likely impact on the Superfluidity architecture (which will be detailed in Deliverable D2.2), and more generally the technical work within the project.

3 Superfluidity use cases

In this section, we detail our consolidated list of use cases. For each, we provide a description and detail its technical and business requirements.

3.1 5G RAN network slices

3.1.1 Use case description

Rough Classification	Main: • Flexible RAN architecture Secondary classification labels: • Reprogrammable RAN • Containers and FPGA • Micro services • Elasticity
Source of the scenario	5G will be more than a new air interface, it will be a new and disruptive network architecture that will offer flexibility in the services and great scalability. Network Slices are a way to split the RAN into different portions, each having its own “size” of the resources. For example, one slice for multimedia, another for M2M and another for IOT. This requirement has been identified by the NGMN [https://www.ngmn.org/uploads/media/NGMN_5G_White_Paper_V1_0.pdf].  *Figure 1: Network slices*
Scenario description in a nutshell	We should be able to support different types of slices and dynamically change the size of each type of slice. More generally, we want the ability to split the 5G Access Network into chained micro services that can be

<table>
<tr><td colspan="2"></td><td>instantiated and scheduled dynamically over a combination of embedded systems and “cloudlets”.</td></tr>
<tr><td colspan="2">Extended description and examples</td><td>For example, we start with a slice for multimedia handsets and then introduce a slice for IOT. As the number of IOT devices grow, the second slice also grows. These two slices are “centralized” (i.e. located in a cloudlet via an Ethernet network). Then we add a new slice for M2M where latency is key. This slice is partially executed on an embedded system located “next to the antenna” for latency purposes.</td></tr>
<tr><td colspan="2">Architecture and/or networking and/or technical challenges</td><td>1) micro-services decomposition of the 5G RAN
2) instantiation in IT atomic entities (such as containers, see www.docker.com) and embedded atomic entities (such as partial reconfiguration HW accelerator, see http://www.xilinx.com/support/documentation/application_notes/xapp1159-partial-reconfig-hw-accelerator-zynq-7000.pdf) in a unified way
3) orchestration, scheduling, service chaining capabilities
4) live and dynamic re-composition</td></tr>
<tr><td colspan="2">Benefits and innovation</td><td>The high flexibility of the proposed use case will make it possible to deploy fully “programmable networks” able to support an infinite number of usages of the access network. Therefore, thanks to the flexibility, access network will be an excellent “substrate” for new innovative services.</td></tr>
<tr><td rowspan="2">ANALYSIS (rough)</td><td>Foreseen components, functions, or primitives</td><td>1) Elementary functions decomposing the RAN, both in the server and the front end sides
2) Middleware for supporting atomic entities on x86 and FPGA
3) Controllers for orchestration, service chaining, scheduling
4) Measurement tools for validating performance</td></tr>
<tr><td>Foreseen Requirements,
Performance issues</td><td>1) Speed of reconfiguration, on server side, on front end side, between both sides
2) Speed for scaling up and down</td></tr>
<tr><td colspan="2">Any other note or comment</td><td>This use case is challenging (!!) and will require strong interaction with other projects (FOSS or H2020) to delineate what is the “best” functional decomposition of the 5G RAN, deploy and test it.</td></tr>
</table>

3.1.2 Use case requirements

Description in a nutshell	The Network Slices is a way to split the RAN into different functions: multimedia, M2M, IOT, and so on, with each having its own “size” of the resources. Each 5G slice and functions is composed of chained micro-service that can be

	instantiated and scheduled dynamically over a combination of embedded systems and “cloudlets”.	
Technical requirements	Reconfigurable HW accelerators	Micro-service based software components need to be accelerated with hardware that brings offloading capabilities. There are systems on-chip that allow partial reconfiguration like programmable FPGAs.
	Microservice platform	Platform that allows management, connectivity and monitoring of distributed microservice in a cloud based or cluster based architectures.
	SDN controllers	This is a requirement for fully end-to-end technology agnostic SDN controllers to improve network capacity and flexibility by providing dynamic control.
	High level programming	Framework for writing programs that can execute on heterogeneous platforms and processors like OpenCL.
	Orchestrators	This is a requirement for more-granular and fine-grained orchestrator that can manage complex cross-domain processes and workflows for complex compositions of microservice (Choreography oriented). Orchestrators should be capable of adjustments from monitoring tools.
Business Requirements	Cost savings	It brings network optimization, scalability and resource isolation for software-defined functions. Network slices is deployed and scale using cheap Commodity hardware and off-the-shelf standards components
	Reduced time to market	Each layer is composed of functional blocks implemented as microservice. With microservices, software functionality is delivered as discrete and lightweight services Fast development, fast iteration, small blocks can be deployed independently enable patterns like continuous integration and delivery iteration
	QoS & QoE	QoS and QoE will be the main KPIs for future 5G users. By deploying slice layers with microservice based architecture, we can deliver fine-grained peace of code quickly more frequently with a better reliability. With continuous integration and delivery, it means much faster bug fixes, and provides better overall software quality, better QoS & QoE.

3.2 Wireless Software Defined fronthauling

3.2.1 Use case description

Rough Classification	Main: • Software Defined Wireless fronthauling Secondary classification labels: • Fronthaul network • Dynamic rates and topology • Network integration • Slicing
Source of the scenario	Fronthaul network is experiencing a strong "reboot" (see http://labs.chinamobile.com/cran/2015/07/13/the-1st-ngfi-next-generation-fronthaul-interface-workshop-2/). In this context we consider the fronthaul network to be an integral part of the 5G RAN and, as such, it has to be fully "reprogrammable". Due to the high data rates, only fibre and "wireless line of sight" can contend.
Scenario description in a nutshell	The wireless fronthaul connects the remote radio head (RRH) to a server. We propose to deploy a fully reprogrammable (via SDN) hybrid fixed and wireless fronthaul expected to be representative of a future 5G fronthaul system. In this scenario, we have a number of 10GbE switches and wireless relays, all controlled by an SDN controller (OpenDayLight). They connect RRH on one hand, server on the other hand. Due to some modification in the access network (e.g. slices update), the traffic conditions change and need to be reconfigured 'on the fly'. The wireless relay is connected by the Ethernet port.
Extended description and examples	There is a change in the slices, for instance after a 'multimedia slice' we add a 'low latency' slice (such as M2M) requiring to add a new flow in addition to existing "multimedia flows". According to the use case "5G RAN network slices", we need to add a "fast lane" for doing the new service chaining. A new flow is instantiated over wireless and wireline using OpenFlow. As the traffic pattern changes, the flows need to evolve also. At some point, the multimedia slice is recomposed to be more centralized (at least some of its components). This changes again the configuration. Parallel transmission over wireline and wireless may be needed to cope with capacity.
Architecture and/or networking and/or technical challenges	1) Plugins to control wireless relays via SDN 2) Coherent function service (re)chaining and flow (re)configuration in order to avoid service disruption 3) Additional failover mechanisms for the wireless link

		4) Unified wireless and wireline controller & applications 5) Parallel transmission over wireless and wireline 6) Integration of “server side components” and embedded systems
Benefits and innovation		5G will borrow its capacity from various technologies like massive MIMO, larger bandwidths and frequencies, more efficient phy processing (network MIMO) and denser network (frequency reuse). All of this puts a larger burden on the fronthaul – the link between the antenna and the cloudlet. Optical links may not be able to reach all sites and wireless relays can be a good way to complement and keep up with the increase, provided it is well integrated with the rest of the network (capacity to change dynamically, SDN integration, Ethernet interfaces…).
ANALYSIS (rough)	Foreseen components, functions, or primitives	- Ethernet port - SDN agent - Load balancing functions (parallel) - Interface and sync with micro services management
	Foreseen Requirements, Performance issues	- No loss, low latency, high throughput - Reconfigurability speed
Any other note or comment		This use case will enable live scenarios, provided we have 5G receivers The setup will be complex (wireless and wireline). It will be connected to the other use cases (for the service part).

3.2.2 Use case requirements

Description in a nutshell	The wireless Fronthauling technique consists of connecting a wireless Fronthaul system anywhere we can access multiple points of fibre links directly from the Access Nodes to bring capacity or coverage to the end users in urban or suburban area. 5G will borrow its capacity from various technologies like massive MIMO, larger bandwidths and frequencies, more efficient physical processing (network MIMO) and denser network (frequency reuse). All of this puts a larger burden on the fronthaul – the link between the antenna and the cloudlet. Optical links may not be able to reach all sites. Wireless fronthaul can be a good way to complement and keep up with the increasing capacity and coverage demand while it is well integrated with the rest of the network (capacity to change dynamically, SDN integration, Ethernet interfaces…).	
Technical requirements	Wireless Fronthaul programmability	In this context of 5G, flow data will vary depending on service and traffic demand. Traffic type (Fronthaul or Backhaul) can also lead to reconfiguration. Therefore, the wireless fronthaul

		system has to be fully "reprogrammable" via the SDN.
	Interoperability	The wireless fronthaul should be fully interoperable and integrated with the rest of the network (Backhaul/Fronthaul integration, SDN integration, Ethernet interfaces, etc.)
	SDN control agent	For monitoring and control management purpose, the wireless Fronthaul system will implement agents to be able to communicate with the SDN controller node and meets on-demand adaptation.
	Latency & Synchronization	Communication between the base band unit (BBU) and the Remote Radio Head (RRH) should be super fluid. Therefore, incoming and outgoing message flows should be transmitted and received in time, in order to avoid any discrepancies within the RAN. The Wireless Fronthaul must be compliant with 5G RAN latency and synchronization requirements.
	Throughput	Wireless Fronthaul interface must support total throughput aligned with the throughput 5G requirement for fronthaul. Millimetre wave frequency band will be implemented to meet this requirement.
	Supports both CPRI and new 5G protocols	Wireless Fronthaul should be flexible and able to support both legacy protocol (CPRI) and new 5G protocols (CPRI over Ethernet, I/Q over Ethernet, etc.).
Business Requirements	Cost savings	With the wireless fronthauling, RRHs can be installed in locations previously impractical or impossible due to the high cost of civil works or the unwillingness of a site owner for aesthetic reasons. The wireless Fronthauling brings a new value proposition to the site deployment of mobile networks as it greatly simplifies site acquisition, site design and construction, hence translating into cost savings.
	Reduced time to market	It minimizes the time spent on site acquisition and installation, and maximizes the flexibility of remote radio unit deployments and reduces time to market.

3.3 Dynamic MAC services allocation in Cloud RAN

3.3.1 Use case description

Rough Classification	Main: • Dynamic Radio Resource Management (RRM) services in Cloud RAN • Life Cycle Management (LCM) of RRM services in the access and aggregation network Secondary classification labels: • scheduling as a service • MIMO as a services • SON (Self Organizing Network) as a service
Source of the scenario	Wireless networks are highly dynamic with changes in load, positions of the mobile users, handovers, and changes in the interference paths, etc. Radio Resource management services should cope with such dynamicity of the wireless networks. This translates in instantiation, migration and termination of RRM services either in the access or aggregation networks. In addition to the NFV resource management and LCM it requires management of the networks and forwarding/load balancing boxes.
Scenario description in a nutshell	Migration or instantiation of RAN/RRM services is a complex task that is initiated by the RRM and involves management of NFV and SDN resources. The LCM of the RRM services is at the heart of this scenario.
Extended description and examples	Due to load changes or handovers there is a need to update and modify the collaborative MIMO schemes. This change may require migration of the scheduling and MIMO streaming from one access network to another. Once such a service is instantiated, the NFV management system should provide the service with the required performance in terms of bandwidth and delay guarantees.
Architecture and/or networking and/or technical challenges	1) Instantaneous creation and migration of RRM services 2) Description of complex LCM operation of the RRM services (application specific operation to be performed in a generic VNFM – VNF Manager) 3) QoS provision for the demanding RRM services
Benefits and innovation	Tackling and solving those challenges will pave the road for a true cloud RAN solution with the operational benefits of NFV accompanied by the performance gain from a joint scheduling and collaborative MIMO across multiple RRH.

ANALYSIS (rough)	Foreseen components, functions, or primitives	- VNFM for RRM services - SDN for RRM - QoS provisioning
	Foreseen Requirements, Performance issues	- High bandwidth and low latency networks between RRH and access and aggregation networks - Instantaneous instantiation of containers/virtual machines

3.3.2 Use case requirements

Description in a nutshell	Migration or instantiation of RAN/RRM services is a complex task that is initiated by the RRM and involves management of NFV and SDN resources. The LCM of the RRM services is at the heart of this scenario.	
Technical Requirements	Dynamic placement	The platform shall support dynamic placement of RRN functionalities at the RRH, edge or aggregation node.
	Connectivity	The network shall support the high bandwidth connectivity between the RRH and the RRM functions (located either at the edge cloud or the aggregation network)
	Latency	RRM functions should support the strict scheduling and HARQ (Hybrid Automatic Repeat Request) timing of 1ms or 3ms correspondingly
Business requirements	User QoE	The user quality of experience should not be sacrificed at the event of migration, scaling or healing
	Cell coverage and capacity	The cell coverage and capacity should be increased by the usage of cloud RAN with centralized multi RRH synchronization and scheduling
	Cost saving	The cost of implementing dumb RRH with edge and aggregation node for performing the signalling and scheduling processing should be lower than the cost of having smart base stations

3.4 S/Gi-LAN services on the (mobile) edge

3.4.1 Use case description

Rough Classification	Main: • Network/Application-Aware Transport/Content Optimization Secondary classification labels: • TCP Optimization • Adaptive Bitrate Video • Encrypted Content • RAN-awareness • Network Congestion
Source of the scenario	One of the key use cases of ETSI Mobile Edge Computing ISG (see slide 5 of http://www.etsi.org/images/files/technologies/MEC_Introduction_slides__SDN_World_Congress_15-10-14.pdf). There are early IETF standardization activities that are also relevant (see https://www.ietf.org/proceedings/93/slides/slides-93-iccrg-3.pdf), but the focus seems to be on delivering a stop-gap solution.
Scenario description in a nutshell	In today's mobile networks, services involving traffic management/DPI and transport/content optimization have been traditionally deployed on the Internet side of the GGSN/P-GW, i.e. in the S/Gi-LAN. Even though the industry recognizes the utility of these services, always in the context of the Mobile Data Tsunami and the desire of operators to differentiate from their competitors on the basis of QoE, deploying such solutions in a scalable fashion is becoming increasingly challenging/costly. Moreover, the lack of accurate visibility on RAN conditions makes it very difficult to deliver traffic management and transport/content optimization in a way that achieves balance between network efficiency and QoE. The flattening and "IP-fication" of the network, in the evolution from UMTS to LTE and, eventually, to 5G, provides opportunities of pushing these services into the RAN and towards the Edge of the network.
Extended description and examples	The proposition is, instead of trying to forward information using various channels from the eNodeB to the Traffic Manager located in the S/Gi-LAN, to migrate relevant services closer to the Edge of the mobile network. Examples of such services are deep packet inspection, content detection, TCP optimization and video optimization, the latter focusing primarily on Adaptive Bitrate Video formats (MPEG-DASH, Apple HTTP Live Streaming, MS Smooth Streaming, Adobe Dynamic Streaming, etc.).

		In their new place, and with the provision of appropriate interfaces/APIs, the above services will be able to take into account network behaviours on a per data session basis (cell handovers, RAT changes, RRC states, etc.). By having an accurate picture of localised network congestion, we will be able to make better decisions on when/how to effectively apply the above, and potentially other, traffic management and optimization services.
Architecture and/or networking and/or technical challenges		At least for the case of OTT content, deploying content optimization and content caching in front of the GGSN/P-GW has been always problematic, due to the billing/charging and lawful interception challenges. However, the introduction of faster networks, in conjunction with the increasing adoption of encryption by content providers, has moved focus on optimization schemes that are lighter-weight and largely transparent. This has increased the attractiveness/viability of the proposed use case.
Benefits and innovation		As mentioned above, the benefits to mobile operators will be immediate, both in terms of maximizing subscriber satisfaction, but also in terms of making the large-scale deployment of such services more affordable.
ANALYSIS (rough)	Foreseen components, functions, or primitives	- RAN conditions/congestion interface/API - example transport/content optimization services/applications
	Foreseen Requirements, Performance issues	- efficient virtualised implementation of the services of interest

3.4.2 Use case requirements

Description in a nutshell	In today's mobile networks, services involving traffic management/DPI and transport/content optimisation have traditionally been deployed on the Internet side of the GGSN/P-GW, i.e. in the S/Gi-LAN. The flattening and "IP-fication" of the network, in the evolution from UMTS to LTE and, eventually, to 5G, provides opportunities of pushing these services into the RAN and towards the Edge of the network.	
Technical requirements	User-Plane Access	The traffic inspection, management and optimization services will require access to user-plane traffic, i.e. the data flows enclosed in packet data protocol sessions.
	Content/Application Awareness	Traffic must be inspected (on the basis of data flows) to identify the type of content being transferred and the type

		of application that is consuming the content. Focus would be on identifying traffic that can be optimised (e.g. ABR video, even if it is encrypted).
	Mobile Network Awareness	The traffic management/optimization functions should be aware of the radio-access type, radio resource allocation, cell congestion level, cell location/topology and other such information necessary for adapting the delivery of services.
	Subscriber/Device Awareness	The services must have access to subscriber identifiers (MSISDN, or equivalent) and, optionally, device identifiers (IMEI, or equivalent) to apply traffic policies with the corresponding granularity (i.e. on a per user or UE basis).
	Mobile Edge APIs	Accessing the above information from the control-plane should not require implementing the relevant protocols, but must be exposed by APIs provided by the Mobile Edge infrastructure.
	Policy Engine	Combining the data-plane and control-plane inputs with the configuration/preferences of the operator to make decisions and apply actions requires a Policy Engine.
	End-to-End Transparency	Traffic manipulation actions must not break end-to-end transparency. Session state will still have to be tracked.
	Transport Optimization	Transport-layer (L4) optimization, initially focusing on TCP. Objective is to achieve high network speeds, efficient use of network resources, congestion control and low latency.
	Traffic Regulation	Regulation/pacing of traffic flows is a traffic management action that helps to conserve network resources. Useful for ABR video (nowadays and in the foreseeable future).
	High Speeds	Services must support per-user speeds aligned with the maximum bounds offered by the 5G mobile network.
	Low Latency	Injection/activation of these services must not increase end-to-end latency by more than a few milliseconds.
	Mobility Aware	Services must be robust to mobility effects, even in the presence of state that is maintained on a per flow basis.
	Dynamic Placement	Instantiation/tear-down and placement of the services must be dynamic, taking into account time-of-day, traffic levels, cell congestion levels, user experience metrics, etc.
	Throughput	Service instances must support total throughput aligned with the throughput of Mobile Edge Data Centres (DCs). Resource efficiency can be optimised internally, if the

		implementation uses lightweight instances, and traffic processing takes place in the user-plane (not kernel).
	Auto Scaling	Resources allocated to the service instances must be automatically/elastically scaled, on a per Mobile Edge DC basis, depending on load and to support transient peaks.
	Availability	Traffic processing services must offer 5 nines availability.
	Load Balancer	If traffic needs to be balanced between multiple service instances, load balancer must support stickiness on a per subscriber/UE (or underlying data flow) basis.
	Orchestration	Intelligent orchestration infrastructure that ensures that services are deployed on the most appropriate platform, taking into account the heterogeneity of resources and scaling them according to the business requirements.
	Service Chaining	Since these services will be probably deployed together with other traffic inspection or traffic processing services, the Mobile Edge infrastructure needs to provide a traffic steering and/or service chaining function.
	Service Monitoring	Estimates the QoE delivered by these service instances.
	Resource Monitoring	Monitors HW resources to drive orchestration decisions.
	Security	Injection of these services must preserve overall security
Business Requirements	User Satisfaction	In advanced networks, the primary business goal of traffic optimization services is to improve user satisfaction and reduce churn, via improving quality of experience (QoE). Migrating such services from the S/Gi-LAN to the Mobile Edge maximizes chances of achieving the above objective.
	Benchmark Testing	Another objective of traffic optimization services is to improve ranking in annual benchmarks (where operators of the same market compare with each other, in terms of network performance). Deploying these services in the Mobile Edge provides the required precision to succeed.
	Network Efficiency	Mobile operators would want to maximize the use/yield of their RAN network investment. Transport optimisation helps with accomplishing this. Migrating these services to the Mobile Edge will allow them to leverage information that is not available in their current placement (S/Gi LAN), which will further assist accomplish the specific goal.
	Congestion Management	Effective traffic management and optimization services need to be also conscious of the congestion levels in the radio network. For example, being too aggressive in transport optimization may actually make the situation

		worse in a congested network. One downside of moving these services to the Mobile Edge is that they will no longer be able to manage congestion in the mobile core.
	Cost Savings	Instantiating traffic optimization services only in areas that require them will make the costs of deploying them more sustainable, in both emerging and advanced Telcos.
	Quick Deployment	Assuming the pre-existence of the Mobile Network Edge and service orchestration infrastructure (see technical requirements), deployment of these services should be faster (shall take days-weeks rather than many months).
	Service/Performance Management	As in the current deployment model, the network operations and planning team will need to be able to manage the operation and configuration of the services, and monitor their performance to assess their benefits.
	Network Analytics	Traffic management and optimisation solutions have always been valuable data input sources for network analytics solutions. We envision that detailed KPI metrics generated by these services shall be integrated into the big data analytics infrastructure, supplementing or superseding the metrics generated by (virtual) probes. Based on the historical metrics, and using trending, forecasting and machine learning techniques, insights extracted can be fed back to the traffic handling services, in terms of policies, towards optimizing them further.
	Transparency	As all data plane services, they need to be adequately "low touch" to not impact transparency of important services, i.e. Lawful Interception, Billing & Charging, etc.

3.5 On-the-fly network monitoring

3.5.1 Use case description

Rough Classification	Main: • Dynamic & optimal placement of network monitoring probe instances Secondary classification labels: • Usage monitoring • Service monitoring

	• Network planning information • Per user/per-line usage analysis
Source of the scenario	Network-wide deployment of traditional DPI solutions is very expensive for most operators. However, operators have a strong requirement to understand how customers are utilizing their products and services.
Scenario description in a nutshell	DPI solutions provide detailed views of network usage allowing operators to generate insights into capacity and service usage, and also to proactively plan for network evolution e.g. capacity uplifts, cache placement, peering and transit. However, the costs of achieving network-wide deployments of traditional DPI solutions are prohibitive for most operators. Therefore, alternative approaches such as the dynamic placement of monitoring probes to specific segments of the network are becoming increasing important. This is especially useful as different segments of the network potentially have customers with varying characteristics of the entire customer base.
Extended description and examples	A network operator wishing to observe different segments for usage monitoring, service monitoring, and network evolution. The operator is able to dynamically place monitoring probes in segments where the customer base is representative of the population for the metric being analysed. For usage monitoring, this involves observing the network bandwidth usage per application, per product For service monitoring, this relates to performance and quality metrics per application, and also per device type For network evolution e.g. CDN deployment, peering and transit, this involves being able to monitor bandwidth delivered from CDN caches on- and off-net, and also to identify services carried over peering and transit networks to ensure quality requirements are satisfied
Architecture and/or networking and/or technical challenges	1) identification of network primitives for this application 2) automated integration of monitoring probes into already existing service chains 3) algorithms for dynamic placement of monitoring probes 4) need to segment the network into appropriate sample spaces if entire network cannot be covered by monitoring probes
Benefits and innovation	Operators do not need to incur the high cost associated with network-wide deployment of DPI solutions to gain accurate insight of customer usage trends. The ability to dynamically instantiate such services in network segments of interest provides a cost effective approach. We also believe this will drive new business innovations in network monitoring such as alternatives to deep packet inspection. For example,

		applications that combine existing approaches such as NetFlow with network services such as DNS
ANALYSIS (rough)	Foreseen components, functions, or primitives	- function to determine optimal placement of monitoring probes - edge storage to capture data associated with usage and service monitoring - edge processing for monitoring logic - means to identify customer segment that is representative of population for metrics that should be monitored - function to collect data from distributed monitoring probes
	Foreseen Requirements, Performance issues	- privacy issues - requirement for extremely fast real time migration of probes - data corruption associated with migration of monitoring probes i.e. data is reported as being for user A after probe has been migrated to monitor user B

3.5.2 Use case requirements

Description in a nutshell	Implementation of Deep Packet Inspection (DPI) with virtual network functions (VNFs) for two scenarios (a) dynamic deployment of DPI to monitor selected network segments e.g. specific geographical regions, and (b) implementing multiple DPI deployments for disparate virtual customers sharing the same physical infrastructure	
Technical requirements	Dynamic placement	There is a requirement for an intelligent system that is able to automatically decide which network segments should be monitored. This could be based on time of day, geographically location, forecasted traffic levels
	Orchestration	This is a requirement for an orchestrator which has network-wide view and is able to dynamically instantiate all functions that compose a monitoring framework
	Service chaining	Most monitoring systems do not work in isolation so there is a requirement to support dynamic service chaining such as DPI → Analytics System → Traffic Shaper
Business Requirements	Cost savings	Continuous increase in bandwidth demands means network-wide deployment of DPI will become increasingly expensive and in most cases

		unsustainable. There is a requirement to enable relatively cheaper monitoring infrastructure
	Reduced time to market	For infrastructure providers/owners, there is a requirement to achieve reduced time to market especially when delivering services such as supporting multiple virtual customers sharing the same infrastructure with isolated DPI deployments
	User Experience	High quality of experience (QoE) for customers is a key requirement for Telco businesses. Enabling a dynamically configurable monitoring platform for observing customer usage of services with a view towards collecting inputs for capacity planning, cache placement, and infrastructure evolution/upgrades.

3.6 Transparent web service acceleration

3.6.1 Use case description

Rough Classification	Main: • Use a specialized accelerator VM instance to convert static content to use a CDN Secondary classification labels: • Transparent CDN • Re-routing of network traffic • Service chaining
Source of the scenario	Many hosting providers are happy to sell under-utilised resources, but buying has been a more difficult prospect. Federation of cloud resources is therefore a promising use case, see for example http://onapp.com/files/brochures/2015/onapp-accelerator-datasheet-170715-web.pdf
Scenario description in a nutshell	Benefit from CDN acceleration without explicitly re-writing web services and content.
Extended description and examples	Web developers may not have much experience with CDN platforms. To expose this, a transparent web content accelerator VM is placed in the same network as the VM instances and can instantly benefit from having multiple connected locations.

Architecture and/or networking and/or technical challenges		1) Analyse static content that is available from a web server 2) Compress the content so it is easy to mirror across CDN sites / within the Federation 3) Re-write local routings to utilize the CDN platform 4) Integrate with existing edge-services 5) Re-route traffic to utilize a closer edge provider
Benefits and innovation		For a large number of end-users, CDN platforms are a step beyond their technical knowledge and have a barrier to entry. To reduce this barrier and allow more people to benefit from CDN and Federation technology the complexity is being passed to the accelerator application. This allows the 'smarts' of the CDN logic to be kept from the web-developers and instead handled by the service provider which is a clear differentiator. Also as there are no changes to the original content, it can be combined with any other technologies that need the unmodified HTTP content.
ANALYSIS (rough)	Foreseen components, functions, or primitives	1) Accelerator VM to be created in the same network 2) Re-routing rules to work with the CDN platform. 3) Static content analysis 4) Compression system 5) CDN replication engine integration
	Foreseen Requirements, Performance issues	The accelerator will need to schedule the conversion of content for new web-sites. The information should be cached and when changes are made only the parts that are changed should be re-analysed to avoid over contention of the accelerator Difficult to transport secure 'session' based traffic e.g. https. Dynamic content will also need to be provided by the main origin until there can be a consistent view across any of the edge-providers.
Any other note or comment		This is quite high in the stack. It is relevant and has clear commercial benefits from a company perspective. It will likely be continued from a business perspective so is likely to have good exploitation and dissemination.

3.6.2 Use case requirements

Description in a nutshell	Re-writing of static content is performed to take advantage of CDN paths that is then synchronised with the Edge nodes.	
Technical requirements	Processing Latency	The latency is dependent on the size of the VM to be accelerated and the resources allocated to the accelerator

	Latency	Latency after the first re-direct should be that of the closest Edge-Node
	Availability of processing platform	The service should be stateless and be run as a VM on the host site. If it fails, this should be detected and the VM restarted and the admin notified.
	Availability	Edge nodes on the Federation will have their own availability values that will be captured in an SLA
	Scalability of processing platform	Currently only a single accelerator is allowed per network. Ideally this service will be scalable depending on the number of VMs that need to be accelerated
	Scalability	Many edge nodes / PoPs can be used. The load balancer with the re-direct should also be scalable
	Caching	As the content is all at the same site as the accelerator no caching is required
	Caching at the edge	Content caches should be large enough to fit all the content for which subscribers are paying for
	Security of processing system	The platform has access to all the VM resources that will be accelerated. Great care should be taken to ensure that content from users are isolated and cannot have an effect on other users of the platform
	Security of edge nodes	Given that the static content is replicated from the source, the integrity should be the first priority. A compromised edge node may try to perform a DoS attack but this can be mitigated.
	Platform	Requires a-priori knowledge of CDN edge sites that are compatible with the accelerator and a pre-established agreement with edge nodes for publishing content
	Acceleration	The level of performance required is expected to be low. If lower latency is required then acceleration may be required.
	Orchestration	The platform needs access to the VM content resources for accelerating. The platform also needs to be able to synchronise with edge-nodes in a consistent manner.
	Orchestration of CDN platform	Edge nodes need to be notified of content changes on the source. If a PoP goes down all routes through that PoP must be recalculated and performed on other Edge nodes.

	QoE	The latency should generally be that of the closest edge node. For dynamic content, delays are expected to retrieve content from the master.
Business Requirements	New additional service	This would normally be done by someone performing some manual operations to change the web-page content to use a CDN platform, or configure the web-development platform to use the CDN. This is a new value added service that helps those unfamiliar with CDN to take advantage of the platform
	The compression, HTML re-writing and the synchronisation require compute time, storage, network and memory	This takes resources away from other VMs on the platform and so should be taken into consideration. The resources though are taken from the same cloud operator as it is run on the operators' infrastructure.
	Possible extension – utilising an alternate CDN network	If another CDN network is used as the end-point rather than in-house CDN network then this CDN'ification of a web-site would need a different business model.
	User experience	The main business model is to keep the system very simple with few technical modifications and interventions required.
	Security	Confidentiality of web-site content must be assured as well as Integrity. For purposes of logging the changes and being able to assure users that only authenticated changes were made there should be non-repudiation

3.7 Virtual CDN for TV content distribution

3.7.1 Use case description

Rough Classification	Main: • Content distribution optimization Secondary classification labels: • Private Virtualised CDN • TV contents
Source of the scenario	Identified need for triple-play, multi-screen TV providers. The use case was identified by one of the project partners, providing such services.

Scenario description in a nutshell	This scenario addresses content caches deployment, close to network edge, running own rules, based on behaviour analysis and consumptions' forecasts. This may be associated to virtualised CDN, allowing several players to deploy their own CDN, according to their rules and needs. Current Internet contents distribution CDNs are based on traditional content caching algorithms, based on observed contents popularity and the distribution rules are similar for all contents and producers. However, other type of services require different rules to cache contents closer to users' location, like TV. Contents are stored in the edge repository but only become available after their scheduled transmission time. After that, they stay available in the edge for a defined period of time; after that, they become available only from a central data centre, freeing storage space for other contents. This may be associated to mobility, with contents following potential consumers, while they change attachment points to the network, creating a more demanding use case, closer to Superfluidity objectives.
Extended description and examples	Current CDNs are based on the distribution of contents, based on observed popularity. This is effective for Internet, where contents become increasingly seen as their popularity growths. This is not the case for TV contents, which follow a different behaviour. In this case, it is possible to forecast visualizations of certain contents based on the past observed popularity of similar ones (e.g. football matches and reality shows) or for periodically broadcasted programs (e.g. series and recorded news). In addition, in many of those scenarios, those contents remain popular for 2 or 3 days, being of no interest to maintain them at edge caches after that period. In that context, for instance, popular TV series may be cached some minutes/hours before scheduled emission time and be kept there for one or two days. In addition first seconds/minutes of all movies may also be cached in order to allow a faster visualization while users are browsing contents or to download to the edge the complete movie. In this context, it is of interest for operators to have access to CDN supporting infra-structures where contents' distribution follow defined own rules, according to content type. Even though CDNs are more related with content distribution, rather than processing, still require processing to select and deliver contents to consumers. This may also be associated with other functions like content adaptation and mobility. Work in Superfluidity should not focus in specific rules for TV contents distribution optimization but look into CDN in general in the context of 5G. Contributions to MEC may be foreseen.

Architecture and/or networking and/or technical challenges		1) Content caching at the edge 2) Content mobility 3) Content adaptation
Benefits and innovation		- Better storage resources usage. - Faster and better service delivery to customers. - Service scalability. - Service functions composition benefits.
ANALYSIS (rough)	Foreseen components, functions, or primitives	- Virtual storage at network edges and along the data path from central data centres - Decision algorithms, taking into consideration the existence or the need to deploy contents at and from network edges
	Foreseen Requirements, Performance issues	- Fast setup/migration/teardown of VM at the edge with high storage capacity - Capacity to rapidly move contents between edge and core, and between edge points

3.7.2 Use case requirements

Description in a nutshell	This scenario addresses content caches deployment, close to network edge, running own rules, based on behaviour analysis and consumptions' forecasts. This may be associated to virtualised CDN, allowing several players to deploy their own CDN, according to their rules and needs. Current Internet content distribution CDNs are based on traditional content caching algorithms, based on observed content popularity and the distribution rules are similar for all content and producers. However, other types of services require different rules to cache content closer to users' location, like TV. Content is stored in the edge repository but only becomes available after their scheduled transmission time. After that, content stays available in the edge for a defined period of time; after that, it becomes available only from a central data centre, freeing storage space for other content. This may be associated to mobility, with content following potential consumers, while they change attachment points to the network, creating a more demanding use case, closer to Superfluidity objectives.	
Technical requirements	Storage	Need to deploy storage capacity at the edge to store multimedia contents
	Analytics	Traffic and service monitoring to extract statistics on the viewed contents

	Connectivity	For fast content transfer between edges, following users
	Users tracking	Be aware of users' profiles served by a certain edge
	Content mobility	Be able to efficiently transfer content between edges
	Multicast/unicast	While contents are distributed and transferred using multicast between edges, efficiency in the radio interface can be improved with the adoption of multicast; a unicast/multicast adaptation function may exist
	Monitoring	Produce cache usage indicators
	Content adaptation	In order to better adapt to radio conditions, video adaptation at the edge can be provided.
Business Requirements	Accountability	Account for: - storage space used at the edge - transferred data volumes - transcoding operations - uni/multicast adaptations - served volume of data - served number of users - number of visualizations

3.8 Video streaming

3.8.1 Use case description

Rough Classification	Main: • Audio/Video streaming Secondary classification labels: • Cloud • Caching • Video delivery optimization
Source of the scenario	Optimizing video delivery and caching for CDN-like setups
Scenario description in a nutshell	Optimization of a complete ABR (adaptive bit-rate) video streaming workflow with state of art adaptive bit-rate protocols (Apple HLS, MPEG-DASH, Adobe HDS, Microsoft Smooth). An important part of such a

	workflow is content preparation of desired protocol specific media chunks. By distributing the preparation of protocol specific chunks throughout the network, the overall workflow can be optimized (reducing backhaul traffic, cache storage, CPU server load). In addition by taking advantage of Superfluidity location awareness and virtualization/provisioning better resource allocation will be possible. Overall, implementation of this use case in Superfluidity will introduce higher Quality of Experience to Users and Decrease Network and Computational Costs for relevant stakeholders in ABR Video Streaming
Extended description and examples	Video distribution online uses Adaptive Bit Rate Streaming (ABR). This involves complex content preparation (encoding, chunk generation and DRM). Doing all these operations at the core is not the most efficient solution. Therefore, we propose late segment generation at the edge (late transmuxing, LTM) for each of the protocols. This can reduce backhaul traffic and caching storage reducing operator cost. In addition, user and location based personalization (user specific content) with lower latency can be achieved at the edge resulting in a better user experience. On top of that backend storage caching to reduce the load on the origin server will be deployed, reducing the number of requests for original content made by the origin server. Example 1. Edge The CDN edge could use more of its CPU and less of its local storage when content is muxed as late as possible. User experience will benefit as start-up times are reduced. Example 2. Home Gateway A home gateway may be equipped with LTM to fetch only a mezzanine set of samples from and origin (or maybe provided with off-hours download in regions where there is low-bandwidth generally) to then mux the content on request to the device (iOS, Android – etc.) in the home.
Architecture and/or networking and/or technical challenges	The setup should be able to add and remove edges based on need, so cloud based autoscaling. Locations upstream origins (to fetch samples once for on-the-fly-conversion) should be configurable when an edge starts. A tiered layout allowing for multiple upstreams (local, remote, far) can be envisioned as part of the architecture.
Benefits and innovation	Creating a setup that combines caching and on-the-fly leads to a setup where videos can be muxed directly from the edge. Compared to other setups this lowers internal traffic and reduces load on the origin. All of which results in faster, cheaper and more efficient video streaming. Also, it will be easier to deploy large scale video streaming services in networks.

ANALYSIS (rough)	Foreseen components, functions, or primitives	Lightweight edge with muxing and caching capabilities. Origin providing mezzanine samples. Location service for edges to get configuration.
	Foreseen Requirements, Performance issues	Latency between origin and edge should be low enough (but cache priming or 'pre-fetch' may be deployed as well).

3.8.2 Use case requirements

Description in a Nutshell	Optimization of a complete ABR (adaptive bit-rate) video streaming workflow with state of art adaptive bit-rate protocols (Apple HLS, MPEG-DASH, Adobe HDS, Microsoft Smooth). An important part of such a workflow is content preparation of desired protocol specific media chunks. By distributing the preparation of protocol specific chunks throughout the network, the overall workflow can be optimized (reducing backhaul traffic, cache storage, CPU server load). In addition by taking advantage of Superfluidity location awareness and virtualization/provisioning better resource allocation will be possible. Overall, implementation of this use case in Superfluidity will introduce higher Quality of Experience to Users and Decrease Network and Computational Costs for relevant stakeholders in ABR Video Streaming.	
Technical Requirements		
	Scalability	The video streaming service shall provide content ranging from low quality to high quality and shall be deployable for a small number of users up to millions of users. Hence, scalability in both quality and number of users is required. This shall be achieved for both unpopular content (long tail) and highly popular content. To achieve this the Superfluidity platform should be able to allocate both network and computation resources when needed.
	Backhaul Traffic	Backhaul Traffic should be reduced. A way to do this is to move the content preparation of protocol specific media chunks to the edge of the network. This avoids duplicate backhaul traffic for chunks based on a common media source. We refer to this as late transmuxing (LTM). LTM can be

		deployed in the edge of the network using mobile edge computing (MEC).
	Origin Server Computational Resources	The resources used by the origin server that prepares protocol specific media chunks from a stored or live media source should be reduced. One way to this is through caching raw stored media in or near the Origin Server. Next, efficient caching of segments in the network edge should reduce the number of requests to the server, thus its resource usage. In other words, back-end and edge caching should save computational resources on the origin server.
	Storage	Storage in caches and the core should be reduced. This will be again done by distributing the protocol specific chunk generation to the edge using late transmuxing. This way intermediate formats relevant to each of the specific protocols can be stored, reducing the storage need. In addition, Superfluidity storage resource allocation mechanism will be deployed to further optimize allocation of storage throughout the workflow.
	Monitoring	It shall be possible to monitor the backhaul traffic, storage and resource utilization in real-time. This will be based on integral parts of the Superfluidity framework for monitoring.
	Provisioning	When the number of users increase resources (storage, computation) should be provisioned on the fly. This will be based on core functionality provided in the Superfluidity framework.
	Location Edge Computing	To enable the ABR workflow optimization localized edge computing near the user should be available.
Business Requirements	Analytics and Monitoring	Data on the resource usage and user analytics shall be available to stakeholders
	Security, DRM	Content should be accessed in a secure way, and Digital rights management (DRM) should be incorporated
	Personalization	Personalized presentations should be available, possibly including personalized location aware add content in the media presentation

3.9 Video orchestration and optimisation

3.9.1 Use case description

Rough Classification	Main: • Video Orchestration • Video Optimization Secondary classification labels: • Location Based Services • Advertisement
Source of the scenario	This use case addresses a very well-known concept of today, regarding the delivery of customized video contents to the users, considering their location (advertisements), as well as the optimization of the video according to the access conditions. This use-case is similar to others identified e.g. by ETSI MEC and fits to TV services providers and OTTs, in order to better provide advanced video features for their customers.
Scenario description in a nutshell	Video Orchestration intends to orchestrate the video contents sent to users, in advanced scenarios like multi-camera events (e.g. football match), where the user may choose the camera he wants to see. This may include advertisement videos. Video Optimization intends to increase the quality experienced by the user, considering the current access conditions. The quality of the video is adapted to the receivers.
Extended description and examples	The Video Orchestration and Optimizations intends to orchestrate the video to be delivered to the user with the most appropriate quality, according to the user's access conditions. In the case of content orchestration, it intends to perform the required orchestration of the video, according to the user selection, especially in situations when there are multiple sources, e.g. in multi-camera football matches, when the user can see the match from different angles. This orchestration may include advertisement videos, in case the operation and/or content provider business model includes that option. The video orchestration may be different from region to region, considering a local customization of contents, e.g. regional sport news. Caching mechanisms may be also involved. In the case of video optimization, the user must be provided with the more suitable video quality, depending on the access conditions at each moment. This means that if the user is crossing an area with lower coverage, the content must be reduced in quality to accommodate that limitation. The need for adaptation may be supported by local

		monitoring systems that provide metrics about the user's quality of the link.
Architecture and/or networking and/or technical challenges		In order to implement the Video Delivery and Optimization features, it is required an architecture that is able to deliver video to users on the edge of the network according to the MEC architecture. For that, the traffic must be offloaded in the edge, not going to the mobile core.
Benefits and innovation		This model brings multiple advantages to operators and users. • The video orchestration on the edge of the network benefits the operators, saving bandwidth on the backhaul, increasing the quality of experience and reducing costs. • The video orchestration on the edge of the network benefits the users, increasing the available bandwidth and reducing latency. • By using the context information (MEC), a more effective quality adaptation can be performed, delivering the most appropriate video quality. • By using the context information (MEC), a more focused advertisement can be performed, delivering sponsored videos e.g. based on the current location of the user (shopping mall). The main innovation of this use case is the operator capability to orchestrate and optimize video contents on the edge, taking advantage of context knowledge, high bandwidth and low latency.
ANALYSIS (rough)	Foreseen components, functions, or primitives	Some foreseen components are: • Video orchestration server • Video optimization server/gateway • Advertisement logic • Advertisement video database
	Foreseen Requirements, Performance issues	Some high level requirements: • Video orchestration capabilities • Video optimization capabilities • Video server capabilities • Advertisement logic capabilities • Monitoring QoS/QoE capabilities

3.9.2 Use case requirements

Description in a nutshell	Video Orchestration intends to orchestrate the video contents sent to users, in advanced scenarios like multi-camera events (e.g. football match), where the user may choose the camera he wants to see. This may include advertisement videos.

	Video Optimization intends to increase the quality experienced by the user, considering the current access conditions. The quality of the video is adapted to the receivers.	
Technical requirements	Connectivity	The Mobile edge platform shall provide a secure environment for providing and consuming services when necessary.
		The Mobile edge platform shall be able to select one or more applications to which the same traffic will be routed and assign priorities to them. The selection, prioritisation and routing during traffic redirection shall be based on re-direction rules defined per application
		The MEC platform needs to provide mechanisms to connect the local production devices (e.g. video cameras and sensors) to the video orchestration Mobile edge application as well as the UEs of the consumers that use the video orchestration service.
		The Mobile edge system shall be able to maintain connectivity between a UE and an application instance when the UE performs a handover to another small cell to another cell associated or not with the same Mobile edge server
		The Mobile edge platform may use available radio network information to adapt contents, guaranteeing service delivery.
		A video adaptation feature may be associated to the service, being provided by the edge platform
	Security	The Mobile edge platform shall provide a service to allow authorised applications to communicate with services provided by the platform
		The Mobile edge platform shall provide a secure environment for providing and consuming services when necessary.
	Location	The platform shall provide a service presenting the list of UEs per particular locations, with context information
	Monitoring	There shall be a service providing appropriate up-to-date radio network information including measurement and statistics information related to the user and radio network.
		There should be an application that will dynamically adapt user traffic requirements to radio network conditions improving QoE
	Routing	The Mobile edge platform shall provide a service to allow authorized applications to send/receive user plane traffic to/from UEs

		The Mobile edge platform shall provide a service to allow authorized applications to inspect and modify selected uplink and/or downlink user plane traffic
		The MEC management functionality shall allow the management of the traffic redirection rules
		The Mobile edge platform shall be able to select one or more applications to which the same traffic will be routed and assign priorities to them. The selection, prioritization and routing during traffic redirection shall be based on re-direction rules defined per application. (The prioritization is used to determine the routing order between the applications)
		The Mobile edge platform should provide data distribution taking advantage of multicast technology (e.g. unicast/multicast adaptation)
Business Requirements	Monitoring	The Mobile edge platform shall allow the collection of charging-related information (traffic usage, application instantiation, access, usage duration, resource usage, etc.,), log it in a secure way and make it available for business processing
	Stakeholders	Stakeholders, as for example OTT should have well defined rules to access MEC ecosystem
		The Mobile edge platform shall be able to present available services and the related APIs to applications for 3rd party players (OTTs, advertisement)
	Big data analytics	It should be gathered customer information to enhance QoE and perform better customization addressing new stakeholders offers
		Location based services can be used to embrace new business partners in value chain

3.10 Augmented reality

3.10.1 Use case description

Rough Classification	Main: • Augmented Reality Secondary classification labels:

	• Enhanced Platform Awareness • Monitoring and analytics • Low Latency /High Throughput Applications • High Bandwidth / High Processing Capacity
Source of the scenario	The 'video game' generation has become accustomed to immersive environments which provide high level of engagement and stimulation. Also given the increasing mobility of the global population whether for holidays or business travel there is a growing need to navigate and interaction with unfamiliar environments than present challenges such as language barriers, navigation either by foot, by cars/public transport, or local knowledge of services such as food and drink. Augmented reality may have potential to improve human safety for example when driving a car. Augmented reality is a logical evolution for users which high expectations of immersive entertainment or users that want informed information retrieval over what is currently afforded by applications such as Google maps.
Scenario description in a nutshell	The ability to provide rich and dynamic information at very low latencies to support applications such as augmented reality dashboards for drivers, augmented reality glasses tendered to a 5G mobile handset to provide ambient intelligence to users such as tourists visiting a new city or providing multisensory stimulation to an on-line gamer interacting with their peers.
Extended description and examples	There is a growing interesting in using augmented reality to enhance user experience for both entertainment and practical purposes such as ambient environmental interactions. Examples include: • Enhanced heads up display which overlay information on top of the windscreen to provide the drive information relating to safety such identifying objects in the dark which are obstructing a car's path, or providing the driver with navigation and road position information associated with a required road manoeuvre. • Augmented Reality (AR) applications running on 5G smart-phones or tablets can provide overlay augmented reality content onto objects viewed on the device camera. This has application such as Smart Cities to enhance the experience for visiting tourists. • Other scenarios include sporting event information, personalised advertisements etc.
Architecture and/or networking and/or technical challenges	1) Delivery of significant data volumes with very low latency (~ms). 2) Local caching of AR content to minimise round trip time and maximize throughput for optimum quality of experience. 3) High availability for safety enhancement related applications. 4) High reliability

		5) Scalable or elastic infrastructure to support transient peaks demands such as cultural or sporting events. 6) Capture of mobile device, application, infrastructure platform and network metrics and fusing them into and efficient and meaningful manner which can be used to quantify the user experience and to drive Orchestration actions which maintain a desired level of experience. 7) Successful and automatic differentiation of the landscape to support performant service deployment is necessary.
Benefits and innovation		• Enhanced user experience which creates 'stickiness' for a location or event. • Drive new business innovations such as location or context aware services.
ANALYSIS (rough)	Foreseen components, functions, or primitives	- Local VR Content Caches - Hardware Acceleration for content generation or rendering - Resource and platform aware Orchestration to intelligently manage VR applications to deliver the best available user experience.
	Foreseen Requirements, Performance issues	-How to include user experience in the loop. -High speed mobile cloud edge scale out or down in respond to spikes in utilisation and to minimise congestion. - Dynamic allocation of resources. - Dynamic network capacity -High reliability radio communications which support rapid location changes i.e. high mobility -Non-repudiation of information/content -indoor & outdoor connectivity - QoS for safety related applications
Any other note or comment		High potential use case given the number of possible devices that could utilise this form of technology.

3.10.2 Use case requirements

Description in a nutshell	The ability to provide rich and dynamic information at very low latencies to support applications such as augmented reality dashboards for drivers, augmented reality glasses associated with a 5G mobile handset to provide ambient intelligence to users such as tourists visiting a new city or providing multisensory stimulation to an on-line gamer interacting with their peers.

Technical requirements	Latency	<20msec for multimedia content or safety related data.
	Availability	>5 nines availability for safety related applications
	Scalability	Scalable or elastic infrastructure to support transient demand peaks.
	Caching	Local caching of content to improve. Linux platform that can support high-speed content delivery.
	Security	Ensure content is only available to authorised users. Content can only be added or modified by authorised users
	Platform	Highly distributed edge node platform to provide complete geographic coverage in an urban. For safety related applications such as intelligent cars urban and rural coverage is required.
	Acceleration	Hardware acceleration for content generation or rendering
	Orchestration	Need orchestration system which have a system wide view of resources and service needs and can reactive in dynamic fashion rather than in a predefined manner.
	QoS	Need to guarantees over the quality of the service a user receives
Business Requirements	User Experience	Need to include the user experience as integral element of the service. Need adjust the service dynamically to maintain the user experience. Service offerings with different costs models which are tied to the quality of the user experience.
	Dynamic Allocation of Resources	Need the ability to automatically re-provision resources in response to a geographic encapsulated demand spike.
	Non-repudiation of content/information	Need to ensure the content being received is trustworthy, particularly important for safety related applications. Opportunity for service providers to differentiate their offerings.
	Service Assurance	Reliable and consistent content delivery is required for a good user experience.

	Indoor and Outdoor connectivity	Pervasive connectivity is required to ensure services are available. Some services for example will take specific in-door use cases such as museum and gallery visits, personalised adverts in a shopping centre.
	New services creation	It should be easy to rapidly create new services, deploy them across the entire network and advertise their availability to consumers. BSS/OSS functions should be tightly coupled with service orchestration

3.11 Context-adapted data delivery

3.11.1 Use case description

Rough Classification	Main: • RAN-aware context-based content delivery and optimization Secondary classification labels: • Location based services • Monitoring and analytics • Adaptive media delivery
Source of the scenario	The scenario of adapting the delivery of information to users according to their context is not novel. Context awareness is a key feature of the so-called ubiquitous computing or pervasive computing concepts. Most often, context awareness is realized at application level with minimal or no interaction with networking infrastructures. The interest of the project in this use case is to identify generic and reusable functionality that can be provided by the networking infrastructure and taken out of the application itself.
Scenario description in a nutshell	Ability to adapt the delivery of content based on context information, including location but not restricting to it and including also an estimate of the user behaviour (e.g. user on the move, potential user interests, etc.)
Extended description and examples	A user walking in a mall subscribes to a context/location based service (concretely, an app), where content relevant to the 'mall' context is pushed to the user terminal. Context is built not only based on location, but also user behaviour, such as user standing in front of a shop opposed to user walking/driving. The system should make sure that the right type of content type is pushed to the user: a walking/driving user is arguably

		not expected to be interested in full quality video opposed to a sitting/standing user.
Architecture and/or networking and/or technical challenges		1) identification of basic (reusable) networking primitives underlying this application 2) need for improved localization and context sensing primitives, including information gathered by user terminals (e.g. accelerometer) 3) need for edge systems able to parse user information and match content with users 4) need for context-aware scheduling and resource allocation strategies
Benefits and innovation		Today, applications of this kind must be developed "from scratch" in every environment, meaning that each mall or museum or venue has to develop its own application. We see an advantage in providing basic (and hopefully standardized or at least systematically addressed) "ambience" primitives which are offered to the "venue operator" (e.g. the mall) so as to rapidly deploy, provision, and customize the desired service. The expected market is for 5G operators to become "ambience service providers", so as to provide the wireless delivery primitives and tuning knobs that the "venue provider" can leverage to implement its own application, just focusing on the design of the application business logic itself, rather than being forced to deal with the underlying content adaptation and delivery details.
ANALYSIS (rough)	Foreseen components, functions, or primitives	- means to monitor, collect and analyse context information - ability to gather sensing information from user terminals - ability to process sensed data on the edge and classify "context" - edge storage and processing of content for low latency delivery - flexible and time-varying resource allocation
	Foreseen Requirements, Performance issues	- privacy issues - resource allocation and scheduling must go beyond bandwidth/delay and include context information - massive peak content downloads (e.g. HD video for "still" users) - very low adaptation delay - how to include user in the loop (interactivity, usability!) - all the above mentioned primitives should be provided as "cloud-like" to the venue provider!

3.11.2 Use case requirements

Description in a nutshell	Ability to adapt the delivery of content based on context information, including location but not restricting to it and including also an estimate of the user behaviour (e.g. user on the move, potential user interests, etc.)	
Technical requirements	Capability to monitor, collect and analyse context information	Ability to provide basic (and hopefully standardized or at least systematically addressed) "ambience" primitives which are offered by the RAN operator to the application providers (e.g. a "venue operator") so as to rapidly deploy, provision, and customize the desired service
	Context processing in the edge	- ability to process sensed data on the edge and classify "context" - edge storage and processing of content for low latency delivery
	Context-aware scheduling and resource allocation strategies	The system should make sure that the right type of content type is pushed to the user: a walking/driving user is arguably not expected to be interested in full quality video opposed to a sitting/standing user
Business Requirements	Privacy issues	The information should be collected (and then exposed) with the user's consent
	Accounting	Depending on the business model it may be needed to account for the resource usage for context processing

3.12 Application-aware performance optimisation

3.12.1 Use case description

Rough Classification	Main: • Application Aware Performance optimization Secondary classification labels: • Enhanced Platform Awareness • Monitoring and analytics • Adaptive media delivery • Generation of content by the users
Source of the scenario	The 'social networking generation' who consume and generate multimedia content on a continuous basis expect to access and share their content in a seamless and real-time fashion. The use of cloud computing resources to support flexible deployment of multimedia resources to deliver a high quality user experience. The advent of

	heterogeneous cloud computing environments with specialized co-processors affords new opportunities to improve use experience and to enhance business for targeted consumer services.
Scenario description in a nutshell	The ability to adapt the delivery of content or to upload and process content based on context information through the use of edge clouds. Orchestration of the applications such as vCDN and video transcoding services is platform aware which ensures workloads can be scaled on computing platforms that have appropriate platform features and capabilities to deliver the required level of user experience.
Extended description and examples	At large outdoor events users are generating large volumes of multimedia content and viewing multimedia content generated by members of their social network who are also attending the same event. The use of distributed caching technology can provide backhaul and transport savings and improved QoE. Content caching has the potential to reduce backhaul capacity requirements by up to 35%. Secondly the QoE can be improved by ensure video processing or delivery applications which are capable of utilizing platform features such as co-processing cards e.g. multi-integrated cores can be scheduled to run on these resources when available. Thirdly the Orchestrator is context aware and can schedule additional cloud resources which are in closest proximity to the event to minimise video stalling and increase browsing throughput.
Architecture and/or networking and/or technical challenges	1) Efficient match of workloads to heterogeneous resources (compute, storage and network). What is optimal allocation of quantities and types for a given application context? 2) Capture of mobile device, application, infrastructure platform and network metrics and fusing them into and efficient and meaningful manner which can be used to drive system Orchestration. 3) Elastic formation of cloud services to address transient increase in resource utilization. 4) Intelligent Orchestration to ensure that services and deployed and scaled on the most appropriate platform and consider the heterogeneity resources in the decision making process. 5) Use of offline and real-time in an interleaved manner to develop insights in to the interplay between services and platform resources allocations on a longitudinal basis and in point in time to optimize service placement decisions. 6) Removing resource abstraction in cloud environments 7) Successful and automatic differentiation of the resource infrastructure landscape to support performant service deployment is necessary e.g. feature, topology, location of resources etc.

		8) Successful and automatic differentiation of the landscape to support performant service deployment is necessary.
Benefits and innovation		Heterogeneity in the clouds will continue to grow in the cloud. Platform features and capabilities have a significant impact on application performance such as multimedia and consequently user experience. By understanding the relationship between applications such as multimedia transcoding and the deployment infrastructure the user experience can optimized. Secondly, the adoption of multimedia applications at the network edge will improve user experience both from a generation and consumption perspective e.g. local content caching to reduce backhaul. Improved application orchestration by exposing platform features which can benefit either the initial application instance of subsequent application scale out.
ANALYSIS (rough)	Foreseen components, functions, or primitives	-Real-time and offline analytics -Real time telemetry -Mobile cloud edge for delivery/processing of user generated media and content. -Ability to gather application metrics from user mobile devices -Resource and service aware Orchestration
	Foreseen Requirements, Performance issues	-Resource allocation and scheduling that is both resource and state aware. -Modelling diverse data sources of varying quality and resources into an actionable representation of state and performance intent. -How to include user experience in the loop
Any other note or comment		Edge multimedia content use cases are already established however these are very generic and do not consider the platform and Orchestration aspects.

3.12.2 Use case requirements

Description in a nutshell	The ability to adapt the delivery of content or to upload and process content based on context information through the use of edge clouds. Orchestration of the applications such as vCDN and video transcoding services is platform aware which ensures workloads can be scaled on computing platforms that have appropriate platform features and capabilities to deliver the required level of user experience.

Technical Requirements	Metrics	Capture of mobile device, application, infrastructure platform and network metrics and fusing them into and efficient and meaningful manner which can be used to drive system Orchestration
	Orchestration	Intelligent Orchestration to ensure that services and deployed and scaled on the most appropriate platform and consider the heterogeneity of resources in the decision making process.
	Storage	High-speed storage at the network edge to store and deliver multimedia content. Storage provisioning needs to be location aware i.e. close to the point of generation or consumption.
	Scalability	Scalable or elastic infrastructure to support transient demand peaks
	Transcoding	Efficient and scalable video transcoding which is resource aware in order to leverage platform features to improve performance e.g. hardware acceleration AVX, GPU etc.
	QoS	Need to guarantees over the quality of the service a user receives. Need the ability to implement different levels of QoS depending on type of user or state of the network.
	Control	Users who are generating content need to have the ability to control how the content lifecycle and how it is exposed.
Business Requirements	Real-time and offline analytics	Real-time analytics are required are required to generate actionable insights at an orchestration level to react to dynamic changes in the network behaviour which could adversely affect service behaviours or performance resulting in poor user experience. Offline analytics are required for build intelligence on the behaviour of the network and the creating insights in the patterns of service consumption such as video content generation.
	Real-time telemetry	Required in order for orchestration processes in the network to react to changes in network state or events.
	Mobile device metrics collection	Is important to have metric from mobile devices to determine the quality of the user experience both from a content generation and consumption

		perspective. It is important to understanding the effect different device version or OS versions have on service performance.
	Pay for service models	Have the ability to offer different levels of service with different costs models. Support both short term (minutes, hours) and longer-term (monthly) contracts.

3.13 Virtual applications

3.13.1 Use case description

Rough Classification	Main: • To utilize an adapted version of the Integrated Storage platform on heterogeneous, low-power microservers Secondary classification labels: • Low resource overhead hypervisor • Massive consolidation of VMs • Communication and shared resources across hypervisor instances
Source of the scenario	A lean and highly-efficient Hypervisor platform, named MicroVisor, is of interest. It has been developed as part of Euroserver-FP7 Project by a project partner. On-going investigations include the algorithms for optimal resource placement along with orchestrators and improving the I/O performance of fluid VM instances across the platform.
Scenario description in a nutshell	Investigate how the new MicroVisor platform can be leveraged for fulfilling the objectives of Superfluidity on heterogeneous hardware architectures.
Extended description and examples	There is a sea change coming for x86 server systems. Current deployments rely on cache-coherence and this is becoming more limited with NUMA and some novel high-speed interconnects being used to extend the use of cache-coherent systems. Eventually there will be a time where cache coherency can no longer be maintained. This is the motivation behind the MicroVisor. In this project we will look at how to leverage liquidity of resources across heterogeneous hardware platforms including ARM and x86.
Architecture and/or networking and/or technical challenges	1) MicroVisor platform that is brought in as binary form (no source code).

		2) Investigate the interaction between ARM and x86 systems and see how performance can be improved when interacting between the platforms 3) Understand the workloads and look to move resources to best take advantage of x86 or ARM hardware in-line with global objectives and policies set by configurations and the orchestration system 4) Investigate how to maintain network communications and fluidity across heterogeneous resources
Benefits and innovation		The innovation comes from the generation of fluid VMs that can move across MicroVisor based platforms. For some workloads, utilizing a VM may not be the best way of migrating resources and maintaining fluidity. We will investigate when it might be appropriate to utilize virtualised resources and/or containers and also potentially look into VM / container / chroot exporting and importing. There are several potential business cases related to the MicroVisor, not limited to utilizing low power micro-servers that are coming to the market.
ANALYSIS (rough)	Foreseen components, functions, or primitives	1) Hypervisor platform 2) Storage, Network I/O resources 3) ARM / x86 systems 4) Interaction with SDN / networking resources
	Foreseen Requirements, Performance issues	We will need to continue the profiling. Beyond this we will need to look at improving the platform performance to allow for many 10s of 1000s of VMs to be instantiated that are light-weight, via MiniOS or other similar light-weight VM designs. Working on improving scheduling systems and see how it will integrate with orchestration systems such as OpenStack

3.13.2 Use case requirements

Description in a nutshell	Current VMs are quite large and heavy, requiring a full Linux O/S to be provisioned. We challenge this model to use much more light-weight instances that can be then used for a multitude of applications that are otherwise inaccessible. In particular we use our experience of storage systems to migrate content close to the end-user to improve the QoS

Technical requirements	Modification of the hypervisor platform to support light-weight instances	Xen / KVM and other hypervisor solutions currently use bloated VMs and so are efficient for loading them in the time-scales relevant to that type of VM. The scalability has been shown to be quite limited without modification. See NEC / ClickOS framework papers. Need to reduce the overhead in the VM handling in these platforms. Xen-Store was modified in previous work by NEC.
	Integration with hardware features	Take advantage of low-level hardware features for acceleration and attempt to couple the system as tightly possible for performance improvements.
	Scalability	Instead of the 100s of VMs that are currently possible on existing HV platforms the idea is to scale to many times this to take advantage of many-core platforms that are emerging from processor manufacturers.
	Availability	The availability of any one instance should be quite high but given that the instances are stateless, if an instance doesn't reply quickly enough, a new instance can be spawned. For high-performance storage systems the availability should approach 5 nines.
	Throughput	The throughput is the most important metric for the high-performance storage use-case. Each individual VM instance should take a fair share of the raw I/O performance. The overhead of the virtualised system's throughput performance should approach as close to 0 (e.g. raw performance) as possible.
	Security	Each storage instance should enforce isolation and there should be no confidentiality or integrity vulnerabilities
	Acceleration	The platform should take advantage of all NIC / RAID / CPU acceleration features to optimise the storage processing performance.
Business Requirements	Leverage benefits of unikernel	Light-weight VM instances do not offer multiple processes. The applications and services need to be re-constructed to be single process and stateless to be of maximum benefit.
	SFC	Multiple light weight services could be chained together, using the output of one Unikernels as the input of another. Good performance handling of VMs

		communicating on the same platform will lead to powerful combinations of services.
	Traditional VMs	The MicroVisor is more suited for light-weight instances. Traditional VMs if they are running on ARM would need to be compiled to make use of the ARM ISA (Instruction Set Architecture).
	ISA agnostic workloads	If a high-level DSL (Domain Specific Language) can be generated that describes the services, then the ISA that the service runs on may become of secondary concern as the workload could be transferred between different containers and run as required. Spare workload capacity could be auctioned off or sold, using a Superfluidity business model – to be worked on.

3.14 Virtual home gateways / virtual Convergent Services

3.14.1 Use case description

Rough Classification	Main: • Home Gateway Services • Set-Top-Box Services • Mobile Services Secondary classification labels: • Location Based Services • Caching Services • Functions Orchestration
Source of the scenario	The concept is similar to the vHGW (virtual Home GateWay), which is a well-known NFV/SDN use case for fixed networks. With the evolution of the vHGW to individually deal with the multiple persons inhabiting the home (e.g. different URL filtering rules), this model can be applied to mobile scenarios, building a convergent service, which can be used from different access technologies. Due to potential low-delay requirements, this use case may take advantage of edge deployments. Several operators are interested in the virtual home gateway and similar scenarios.
Scenario description in a nutshell	The vHGW use case intends to move traditional functions (e.g. firewall, parental control, NAT, etc.) residing on the customers' home to a virtual HGW (vHGW) in the cloud. In a convergent scenario, these services apply

	both to fixed and mobile environments, providing a convergent desired behaviour, either when the user is at home or using a mobile device. The vSTB component complements the use case, extending the usage of a virtual STB to a multi-screen scenario, simplifying the convergent environment, reducing operator investment and making easier the upgrade and deployment of new services, being accessible to the user from any terminal.
Extended description and examples	The vCS (virtual Convergent Services) use case is built on top of a generic platform composed by a chaining engine and a set of functions, which can be dynamically sequenced as desired per person/user. The traffic generated by the user will be driven according to defined chaining rules and this sequence will determine the treatment the traffic will receive, e.g. crossing firewalls, parental control, NAT, etc. In a convergent scenario, users take advantage of the operator capability to provide the same behaviour on fixed and mobile environments. This is the value-added that this vCS use case brings to users. With this model, operators can create a separation between the access network and services. For the users, service configurations apply, no matter the access network they are using. The deployment of vHGWs on the data centre, make operators free to deploy new services, or upgrade them without changing anything on customers equipment. This provides independence from smartphones/HGWs and flexibility to manage services, lowering operational costs (no intervention is needed with the customer). This kind of services largely benefits customers when deployed on the edge of the network, since they can get low delay and high bandwidth, to do, in the network, what today use to be done locally at customers' home or terminal. Taking advantage of this generic platform, operators can deploy other services than traditional functions. Good examples of this are CDN features, which increase the overall quality of experience of customers and reduce the required bandwidth on the operator's core and backhaul network.
Architecture and/or networking and/or technical challenges	To implement the vCS use case, the architectures of multiple SDOs must be combined, in particular: • TMForum (OSS/BSS issues) • ETSI NFV (NFV issues) • ONF (SDN issues) • IETF SFC (Chaining issues) The following picture intends to provide an overall architecture.

<table>
<tr>
<td></td>
<td>
Note: This combination is the main challenge of this use case.

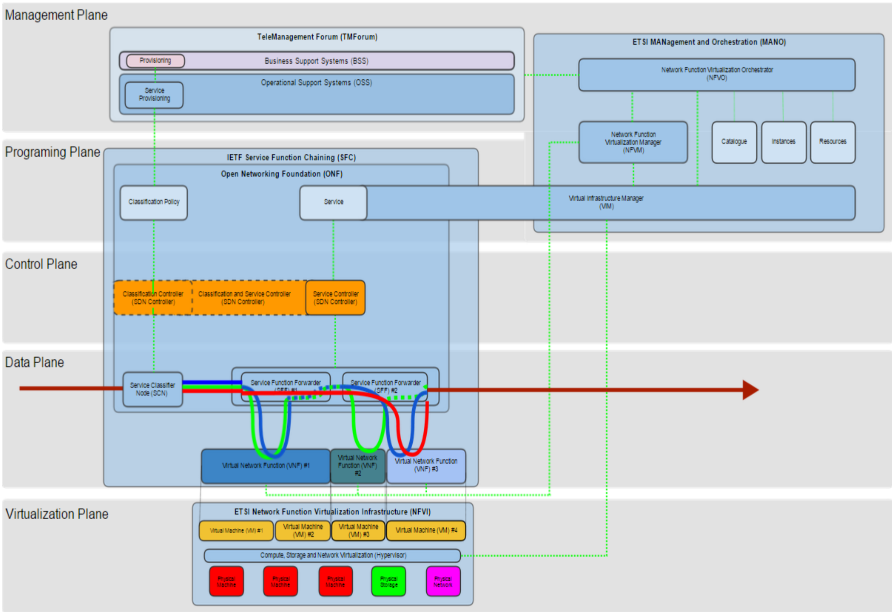


Figure 2: Overall architecture for virtual Convergent Services use case
</td>
</tr>
<tr>
<td>Benefits and innovation</td>
<td>
The migration of services from the customer home/terminal to the cloud has the following advantages

- On fixed networks, it allows operators to deploy simple and cheaper equipment on customers' home (L2 HGW); due to the high number of customers, this usually represents a significant cost saving for operators. On mobile networks, it allows customers to save computing power and battery on their terminals.
- It brings agility to operators to manage services (add, upgrade, swap), without requiring any change on the customers' home. Many times, new features require significant upgrades or even equipment replacement, which results in huge costs for operators. On the other hand, operators are no longer "on the hands" of the functions vendors.
- For fixed networks, it allows a more flexible management of functions, not requiring the displacement of a technician to the customers' home to upgrade/replace the HGW.
- Other functions can be added to the function chains, such as CDNs, improving the quality experienced by customers, making a more efficient use of the network resources and reducing costs.

In short, the main innovations have to do with the agility of functions management and the convergence between fixed and mobile functions available.
</td>
</tr>
</table>

ANALYSIS (rough)	Foreseen components, functions, or primitives	According to the architecture depicted above: • OSS/BSS Service Provisioning • VIM, NFVI, NFVM, NFVO, Catalogue DBs • MANO & Chaining features (Classification Policy, Service Provisioning, SDN Controller, Classification Controller, Service Classifier, Service Function Forwarder/OVS, etc.) • HGW/STB Functions and VNFs (Firewall, Parental control, NAT, DHCP, Broadcast Video Delivery, VoD Delivery.
	Foreseen Requirements, Performance issues	Some high level requirements. • VNF management and orchestration capabilities • Service chaining capabilities • Service provisioning capabilities

3.14.2 Use case requirements

Description in a nutshell	The vHGW use case intends to move traditional functions (e.g. firewall, parental control, NAT, etc.) residing on the customers' home to a virtual HGW (vHGW) in the cloud. In a convergent scenario, these services apply both to fixed and mobile environments, providing a converged desired behaviour, either when the user is at home or using a mobile device. The vSTB component complements the use case, extending the usage of a virtual STB to a multi-screen scenario, simplifying the convergent environment, reducing operator investment and making easier the upgrade and deployment of new services, being accessible to the user from any terminal.	
Business Requirements	Functional	The solution must be able to provide, from the cloud, convergent services to mobile and fixed users (services can be traditional HGW services, e.g. Parental Control, Video Streaming, or new services)
	Accountability	The solution must be able to account the details about the provided convergent services
	Scalability	The solution must be able to scale according to the required load
Technical requirements	User-Data Plane Access	The Edge System must be able to get access to user-data traffic (Uplink/Downlink)
	Traffic inspect and processing	The Edge System must be able to access user-data traffic to/from UEs, according to a set of traffic rules, redirecting it to the Services

	Latency	Latency should be unperceivable to users (in comparison with a real STB using remote control) and jitter should not preclude existent quality
	Service management	The system shall be able of perform the management of the service core control functional blocks (Firewall, Parental Control, NAT, etc.)
	Lifecycle management	The system shall be able to manage (Create/Activate/...) the lifecycle logical convergent services
	Services chaining	Multiple Converged Services can be chained in order to delivered combined services to the UE
	APIs	The Edge System functions are accessed through APIs
	Dynamic Services	The platform shall be able to dynamically change services to be made available to users
	Performance management	The platform shall collect and analyse data parameters (received packets, sent packets, lost packets, CPU usage and memory usage) to perform planning, provisioning, maintenance and quality measurements.
	Security	The system shall be able to guarantee service resources protection and isolation

3.15 Business communication services

3.15.1 Use case description

Rough Classification	Main: • Business communication services at the edge Secondary classification labels: • Small Cells • Mobile Edge Computing • Services platforms • Edge breakout • Enterprise/operator converged services • Community services

Source of the scenario	This use case presents a possible evolution of currently provided business services like ipCentrex. Several operators already have products (IMS/ipCentrex) on convergent communications areas and this use case is the natural evolution of it. It develops the use case is presented by MEC [ETSI GS MEC 002 V0.4.2(2015-07)].
Scenario description in a nutshell	Enterprises and operators with installed small cells may take advantage of traffic breakout at the edge, allowing an efficient communication at their premises without requiring the traffic to go to the core of the network. This is valid for data exchange but also for other type of communication services, like voice and Unified Communications [http://www.scf.io/en/documents/081_-_Enterprise_unified_communications_services_with_small_cells.php]. In particular, the applications to support the unified communications can be deployed at the edge, become even mobile, following the user when outside the enterprise, reducing the backhaul resources consumption and increasing the network responsiveness (low delay). Context: • Smartphones, tablets, etc. (mobile terminals), replace traditional fixed terminals at the enterprises • The 'office' moves to the terminal, using services provided by the operator, in both cases possibly hosted in a Cloud • Mobile networks replace private traditional enterprise networks • Small cells being installed at companies to increase capacity • The operator network is used as additional, distributed, enterprise resources
Extended description and examples	The Unified Communications concept puts in the mobile terminals the traditional enterprise services, serving as a single platform for the user to have unified access to all sort of messaging, voice, collaborative and other services, being them provided by the enterprise or by the operator. The deployment of Small Cells, besides improving QoE regarding traditional operator services, may also integrate with Wi-Fi accesses and may provide breakout features to access local enterprise services. The adoption of the virtualisation technology, also under analysis to be applicable in the small cells context [Small Cells Forum, "Virtualisation for small cells: Overview", document number SCF106]. It may represent a step forward in the evolution of both aspects, making possible to deploy dynamically, at the right place, the required functions, from network to service functions. MEC, as presented in the previously referred ETSI use case, can also be added, inclusively as a platform being present in future small cells'

	solutions. MEC will provide the platform for applications deployment at the edge that may be, in fact, inside enterprises environment. MEC can also be used to host the virtualised part of Small Cells. Thus, the merging of unified communications, small cells, virtualisation and mobile edge computing can be exploited to provide better, more flexible services to business customers. There are several advantages with this approach: (i) for the enterprise, reducing the cost of acquisition and maintenance of the enterprise network and services; (ii) for the users, making their life easier; (iii) for the operator, additional business. Edge computing resources can be used to deploy localised services to serve a community served by a geographically close set of 5G cells (videoconferencing, content caching, etc.). Service mobility is also improved, since when the user is away from its enterprise environment, the (unified communication) environment follows him. From his mobile terminal, the user runs traditional services hosted at the enterprise and/or at the operator's platforms. When the user is away, services may be run at their traditional servers (enterprise or operator) or at the closest edge. When the user terminal is under his company's small cell, some services traffic is exchanged directly with the enterprise's services platforms. The use case scope can be further extended to other type of communities, beyond business, which are geographically dense and require a specific communication service.
Architecture and/or networking and/or technical challenges	1) Service platforms have to be developed considering the possibility to be deployed at the enterprise, operator data centre or at the edge. 2) Move communication service platforms while communications are ongoing 3) Virtualisation of Small Cells 4) Usage of MEC inside the enterprise environment
Benefits and innovation	Operator - New business model - Optimal network resources usage (local breakout) - Reduced traffic on backhaul network Users - Reduced latency times - Increasing bandwidth - Network/Service follows the user

		Enterprises - CAPEX reduction on network/services acquisition - OPEX reduction to manage and maintain enterprise networks and services
ANALYSIS (rough)	Foreseen components, functions, or primitives	- Service components (voice, conferencing, storage, etc.) - Small Cell components - Traffic offload mechanisms - MEC platform
	Foreseen Requirements, Performance issues	- Service state being transferred between edge service platforms, on the run - Local breakout feature for unified traffic

3.15.2 Use case requirements

Description in a nutshell	Enterprises and operators with installed small cells may take advantage of traffic breakout at the edge, allowing an efficient communication at their premises without requiring the traffic to go to the core of the network. This is valid for data exchange but also for other type of communication services, like voice and Unified Communications. In particular, the applications to support the unified communications can be deployed at the edge, become even mobile, following the user when outside the enterprise, reducing the backhaul resources consumption and increasing the network responsiveness (low delay). Context: • Smartphones, tablets, etc. (mobile terminals), replace traditional fixed terminals at the enterprises • The ‘office’ moves to the terminal, using services provided by the operator, in both cases possibly hosted in a Cloud • Mobile networks replace private traditional enterprise networks • Small cells being installed at companies to increase capacity • The operator network is used as additional, distributed, enterprise resources	
Technical Requirements	Connectivity	The Mobile edge system shall be able to maintain connectivity between a UE and an application instance when the UE performs a handover to another small cell to another cell associated or not with the same Mobile edge server

		On demanding network should be made available to enterprise employees depending on the used business applications (HD videoconference, enterprise video messaging...)
		The Mobile edge platform shall be able to select one or more applications to which the same traffic will be routed and assign priorities to them. The selection, prioritisation and routing during traffic redirection shall be based on re-direction rules defined per application
		Small cells interconnection should be ensured by automatic setup
	Storage	There should be made available elastic and persistent content storage space
	Security	The Mobile edge platform shall provide a service to allow authorized applications to communicate with services provided by the platform
		User communications should be based on a single user identity providing easy access management service controlled by a MEC application
	Breakout	The Mobile edge platform shall provide a local breakout feature
		The LBO should filter and route the intended enterprise business traffic
		The Mobile edge platform shall provide a service to allow authorized applications to inspect selected uplink and/or downlink user plane traffic
		Redirection of traffic between UE and to/from Internet applications to Mobile edge platform according to configurable parameters.
	Monitoring	A MEC application should be designed to gather statistics concerning usage of available services (in order, for instance, of pre downloading and caching information that could be of interest to other employees)
		To implement a MEC application able to ensure user traffic dynamic adaptation (e.g. via contents recoding) to the current small cell capacity (bandwidth, latency and so on)

Business Requirements	Monitoring	The Mobile edge platform shall allow the collection of charging-related information (traffic usage, application instantiation, access, usage duration, resource usage, etc.), log it in a secure way and make it available for business processing
	Configuration	MEC applications configuration shall be easy in terms of offering new service Bundles with dynamic pricing offers
	Stakeholders	Stakeholders shall have well defined rules to access MEC ecosystem
	Big data analytics	Customers behaviour information shall be gathered to enhance QoE and perform better customisation addressing new stakeholders offers
	Management	Applications to converge enterprise and operator services should be scalable allowing the adding, modifying or deleting of users and services.
	Location	It shall be possible to deploy the Mobile Edge Server at various locations (at operator or at enterprise) according to chosen business model.

3.16 Local breakout

3.16.1 Use case description

Rough Classification	Main: • Network Forwarding Efficiency Secondary classification labels: • Local Network Connectivity • Corporate/Campus Connectivity • Big Events (Crowd) Connectivity • Tactile Communication Services • Moving Traffic to Local network avoiding going to the core
Source of the scenario	This use case addresses a very well-known concept on 3GPP networks, in order to avoid Internet traffic to go through the Home Network when a user is roaming out. In this case, the same concept is applied, but to a more restricted area (the edge). This use case will help to flatten the network and make it more efficient.

Scenario description in a nutshell	Local Breakout intends to avoid user traffic to be sent to the mobile network core, when communication parties are on the same edge network (e.g. eNB). On 3GPP networks, by default, all traffic is terminated on the mobile core (PDP/PDN). However, in many cases, knowing that users are attached on the same edge, communications can be shortcut, making the connectivity more efficient. A similar concept may also be applied to fixed networks.
Extended description and examples	The Local Breakout use case may apply to many situations where communication parties are attached to the same edge of the network. In this case, it is desirable that the traffic can flow between them directly, not going to the mobile core, which is today the default behaviour. Use cases that may take special advantage of that scenario are: some big events (sport, concerts, etc.) and corporate services, among others. In the case of big events, the operator can deploy services, like event video streams from multiple cameras, statistics, virtual reality, augmented reality (e.g. overlay player names), etc. on the edge, benefiting from the significant backhaul traffic reduction, increasing the bandwidth and reducing the latency available for users, resulting in an improved quality of experience. Corporate connectivity services can also benefit from the Local Breakout scenario. In this case, corporate and other big customers (like universities and schools) users, use to generate large amounts of traffic to access to local corporate services or other corporate users. Using Local Breakout, operators may save a lot of resources, increasing the overall quality experienced by the user (bandwidth, latency, etc.).
Architecture and/or networking and/or technical challenges	In order to implement the Local Breakout behaviour, the mobile network has to change the default forwarding behaviour, which is sending all traffic to the core. To achieve this behaviour on current deployments, the easiest way is to strip the GTP-U encapsulation (between eNB and Core) and access to the user IP traffic, delivering it (NATed) to the desired destination. On traffic return, a similar process is performed, this time encapsulating the user IP traffic into GTP-U. However, a cleaner solution would be ending the tunnel in the edge, bringing the (today's core) Gateways to the edge. This solution would not require any workaround on the GTP protocol, but this is only economically reasonable on green field operations. In the last Releases, 3GPP has been working on several technologies that can support at some extent the technical requirements needed for this use case, comprising a midway solution. For example, [TR23.829] describes multiple options of local/offloading communication that can fit the RNC/eNB breakout requirements (see Figures below).

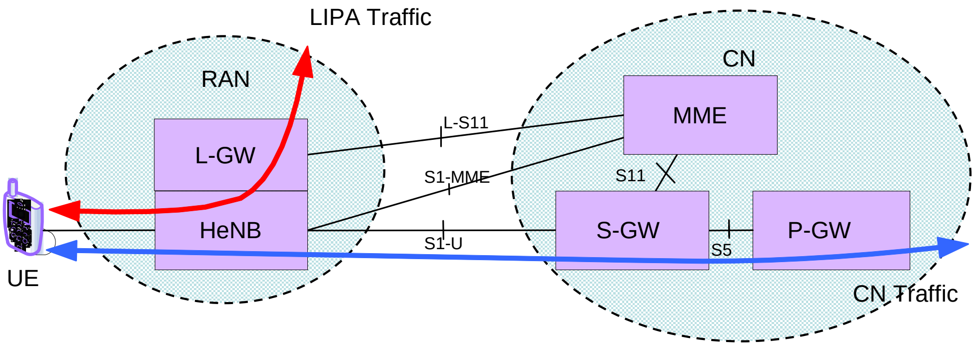
<table>
<tr><td colspan="2"></td><td>


Figure 3: LIPA solution for HeNodeB

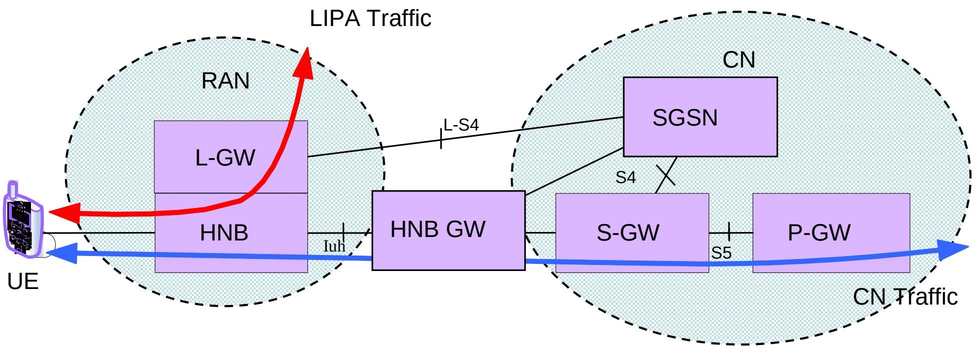


Figure 4: LIPA and SIPTO solution for Home Node B (HNB) with S4-SGSN
</td></tr>
<tr><td colspan="2">Benefits and innovation</td><td>
This model brings multiple advantages to operators and users.

- Performing local breakout, the operator can save resources on backhaul links, since part of the traffic is forwarded locally.
- As the traffic is not forwarded through the core, users have more bandwidth available for local traffic, e.g. corporate, event-related, etc.
- As the traffic is not forwarded through the core, users have lower latencies, benefiting especially some delay-sensitive applications, such as voice/video calls, gaming, virtual reality or augmented reality.
- This model allows operators to easily set up local services for temporary events, without a significant backhaul capacity, decreasing deployment times and reducing deployment costs. These costs reduction can also be reflected to the customer

The main innovation of this use case is the operator capability to change the forwarding model, decentralizing the routing process and making it more efficient, especially for some particular scenarios. Operators increase user satisfaction and reduce costs; users get a better quality of experience (tactile communications).

The use case complements and fosters Edge Computing.
</td></tr>
<tr><td>ANALY SIS</td><td>Foreseen components,</td><td>
Some foreseen components are:

- Mobile Local/Offloading (LIPA/SIPTO features)
</td></tr>
</table>

	functions, or primitives	• Edge Forwarding Services
	Foreseen Requirements, Performance issues	Some requirements are: • 3GPP Local/Offloading (LIPA/SIPTO) capabilities • Edge Forwarding Capabilities

3.16.2 Use case requirements

Description in a nutshell	Local Breakout intends to avoid user traffic to be sent to the mobile network core, when communication parties are on the same edge network (e.g. eNB). On 3GPP networks, by default, all traffic is terminated on the mobile core (PDP/PDN). However, in many cases, knowing that users are attached on the same edge, communications can be circumvented, making the connectivity more efficient. A similar concept may also be applied to fixed networks.	
Business Requirements	Functional	The solution must be able to forward specific UE traffic to a Corporate Network and vice-versa
	Accountability	The solution must be able to account the details of the traffic forwarded/received from the Corporate Network
	Scalability	The solution must be able to scale according to the required load
Technical requirements	User-Data Plane Access	The Edge System must be able to get access to user-data traffic (Uplink/Downlink)
	Traffic Inspection	The Edge System must be able to inspect user-data traffic (Uplink/Downlink) considering a set of traffic rules
	Traffic Breakout	The Edge System must be able to break out user-data traffic coming from UEs, according to a set of traffic rules, and forward it to a Corporate Network
		The Edge System must be able to break out user-data traffic coming from a Corporate Network, according to a set of traffic rules, and forward it to some UEs
	Filtering Rules	Traffic rules must support up to layer 7 parameters
	APIs	All Edge System functions are accessed by using specific APIs provided by the Edge System

3.17 Mobile services offloading

3.17.1 Use case description

Rough Classification	Main: • Cloud based mobile services Secondary classification labels: • Location based services • User following services • Cloud offloading • Migrating services from mobile devices to mobile edge
Source of the scenario	The use case enables operators to sell new and innovative valued-added services to their customers.
Scenario description in a nutshell	Applications currently running in users' mobile devices (e.g. parental control, ad removal) can be offloaded to the edge of the network, for improved security and reduced battery usage.
Extended description and examples	Taking advantage of the Superfluid Cloud properties, ISPs can offload many of the services that currently run on users' mobile devices to the cloud, providing extra security and reduced battery usage, while keeping a near local user experience. Given the small footprint envisioned for the Superfluid Cloud guests, services can be deployed in a VM per user fashion, providing extreme personalization and unmatched security and privacy. The small footprint together with the mobility properties of the Superfluid Cloud, enable the processing units to follow the user around. Deployed in the edge, from the home gateway to the closest BTS, and therefore eliminating all sorts of overheads associated with centralized approaches, for a user experience that matches a local application. Examples: 1. Parental control By offloading parental control to the cloud, parents have extra assurance that their children cannot tamper with the service, but are still monitored independently of their location, being it at home through the Wi-Fi connection or outside, using 3G/LTE, and even independently of the device their using. 2. Ad removal Ad removal is an important part of web surfing these days that the users want available in all their devices. By using the Superfluid Cloud

	operators could provide personalized ad removal on the cloud, eliminating all of the browser extensions and mobile apps, and reaching new devices, like Smart TVs that start to have integrated browsers but no add removal functionality.
Architecture and/or networking and/or technical challenges	The use case depends directly on the properties provided by the Superfluid Cloud, and therefore depends on the development of such technology. The following are some of the research topics especially interesting in this context. 1. There are all sorts of new "lightweight virtualisation technologies" that can be used to provide small footprint guests for the Superfluid Cloud, ranging from containers to Unikernels or minimalistic Linux deployments. However, porting applications to these technologies with low effort is one of the biggest road blockers at this point in time and needs to be addressed. 2. A management framework for the Superfluid Cloud is another open issue. The proposed system quickly scales to thousands or hundreds of thousands of guests, deployed over a wide area and in constant movement. These characteristics put a special burden on a Superfluid Management Framework, that will need to address two particular requirements for this use case: • Extreme scalability: the framework should be able to manage hundreds of thousands of guests; • Low footprint: the software must run on a wide variety of devices, from powerful datacentre servers, to low power single board computers deployed at the edge. 3. For the successful deployment of such a system platform operators need to open the system to 3rd parties who will develop some of the applications. Therefore, the development of APIs through which 3rd parties can use the system is another essential topic. These APIs should also provide the means through which users will be able to request services. 4. Security and privacy of the users and users' data in such a system is essential. Users need to be provided with means enabling them to choose the type of data (network traffic) each of the applications have access to.
Benefits and innovation	Current services are either local (in terms of the device) or centralized (in the cloud). By moving cloud services closer to the user operators can provide the advantages of cloud services with the user experience of a local application. Benefits: • For the user: o Longer battery life on their mobile devices

		◦ Extra security ◦ Device independent services • For the operator: ◦ Provide value-added services
ANALYSIS (rough)	Foreseen components, functions, or primitives	• Superfluid Cloud porting tool • Virtualisation infrastructure • Management framework
	Foreseen Requirements, Performance issues	• Extremely fast times for instantiation, tear-down and migration of appliances • Extremely scalable management

3.17.2 Use case requirements

As an example of mobile services offloading, we consider the requirements for the ad removal sub use case.

Description in a nutshell	Given the proliferation of online advertisements and the fact that a large number of players base their revenue model on them, it is perhaps unsurprising that ad blockers have become commonplace. While certain useful, they do consume CPU cycles as they scan incoming traffic which results in reduced battery life when talking about mobile devices. In this use case we propose to provide an on-the-fly, virtualised ad blocker service that can be run in edge networks, essentially offloading this functionality from mobile devices in order to increase their battery life.	
Business Requirements	Functional	The solution must be able to remove online advertisements from incoming traffic
	Accountability	The solution must be able to log its usage for accounting purposes
	Scalability	The solution must be able to scale according to the required load
Technical requirements	User-Data Plane Access	The system must be able to get access to user-data traffic
	Traffic Inspection	The system must be able to inspect user-data traffic considering a set of traffic rules
	Ad removal	The system must be able remove online advertisements from incoming traffic.
	Throughput	The system should be able to remove ads at relatively high throughput rates (e.g., 1-10Gb/s) in

		order to be able to service a large number of devices concurrently. The device should be able to scale its performance to current demand.
	Service Instantiation	The service should be able to be instantiated on-the-fly, when needed (e.g., in timescales of 10s of milliseconds) and torn down when no longer needed.

3.18 Context-aware smart living

3.18.1 Use case description

Rough Classification	Main: • Intelligent Connected Devices to Support Smart Living Scenarios Secondary classification labels: • Context and Location based services • Autonomous intelligent services
Source of the scenario	Exponential growth in the number and types of devices (sensors, actuators, mobile devices) that have a wide range of characteristics and infrastructure related demands. Potential growth to tens or hundreds of discrete or integrated devices for individual users. Currently users already have a number of personal devices such as a smartphone, tablet, laptop, Bluetooth-enabled devices e.g. activity tracking devices. This trend will continue and will evolve to include ambient devices which do not require direct interaction by the user but provide information and support services that are visible and invisible to the consumer. 5G needs to be designed to accommodate such growth in device numbers.
Scenario description in a nutshell	A key focus for the proliferation of IOT devices is realisation of a connected world where devices and sensors are connected in a seamless manner to support humans in the daily activities of living. The combination of sensor, actuators, mobile devices, ubiquitous high speed wireless connectivity and cloud infrastructures will the support the realisation of a smart society where intelligence is embedded into all aspects of daily life such as smart transport, smart home, smart health and wellness and so forth.
Extended description and examples	The term smart is being used to describe a dramatic evolution in the way we live our daily lives. It has led to the advent of concepts such as smart cities, smart homes, smart living, smart society based on dense wireless sensor networks with ubiquitous connectivity and data accessibility and

	context aware intelligence. Distributed networks of sensors together with personal sensors/devices will be used to support our daily activities. In homes, various ambient sensor and actuators such as temperature sensors, security devices, heating controllers, meters and home appliances will be connected wirelessly. While the majority of these sensors will be low data rate devices, devices such as HD security cameras will be high data rate and will need to be accessible from any location outside the home including different countries by the home owner on personal mobile devices. The task for 5G will be to integrate the management of these diverse connected devices. Examples of smart living scenarios include: • Smart clothing to support health and activity monitoring. It is expected that the use of wearables consisting of multiple types of connected devices and sensors will become more pervasive. Smart clothing with embedded ultra-light, low power, waterproof sensors particularly for sports and leisure activities are starting to emerge. These sensors can measure various environmental and biometric parameters such as temperature, heart rate, blood pressure, body temperature, breathing rate, blood oxygen levels, skin moisture, etc. Information from ambient sensing such as pollution monitoring can be used to inform the user area will provide them with the best air quality when exercising outdoors. A key challenge for this use case is the overall management of the number of devices as well as the data and applications associated with these devices. • Wearables, smartphones, tablets, and other devices with sensors that are location and context aware will work together with apps and services that people use on a daily basis. For example a person is having a business meeting with a customer contact in their calendar. A few minutes prior to the meeting, their mobile device might share some data about that person by quickly sending content cached nearby in the cloud to help their preparation for the meeting while they are in transit to the meeting.
Architecture and/or networking and/or technical challenges	1) Smart aggregation devices which can support multiple radio protocols with different data bandwidths and QoS. 2) Improved localization and context sensing primitives combined with analytics to improve ambient intelligence and to provide user specific information and insights. 3) Data Aggregation and processing at edge systems. 4) Support for sensor network which are highly heterogeneous in nature from a wide variety of vendors.

		5) surge accommodation i.e. ability to handle scenario where a large number of devices attempt to access the network simultaneously
Benefits and innovation		• IOT is expected to be a key driver for 5G. To achieve the vision of IOT which is a smart and hyper-connected internet of everything world requires the high bandwidth, low latency and ubiquitous connectivity. 5G will support the massive deployment of connectivity devices (sensors and actuators) to support a wide variety of scenarios. The expected level of embedded and ambient sensing coupled with 5G connectivity will support the development of new services value add services by catalysing new businesses. For example wearable device could connect to other devices and this could lead to new kinds of experiences which can be monetised into value add services for consumers.
ANALYSIS (rough)	Foreseen components, functions, or primitives	• Cloud based radio access networks to support dynamic scalability • Ability to gather sensing information from user device e.g. smartphones, body worn sensors, smart clothing, ambient sensors etc. • Intelligent orchestration to support adaptable service delivery based on based on system wide context • means to monitor, collect and analyse context information • ability to gather sensing information from user terminals • ability to process sensed data on the edge and classify "context" • edge storage and processing of content for low latency delivery • flexible and time-varying resource allocation
	Foreseen Requirements, Performance issues	• Long life batteries 10+ years • Smooth mobility between cells, layers and radio access technologies needs to be assured • high capacity and low latency backhaul • network need to be programmable, software driven and managed in an integrated way • support for any-to-any communication • reliability • security • performance (latency, throughput) • Varying payload size and frequency of transmission
Any other note or comment		The IOT can be broken down into many specific use cases examples. Starting a general overarching scenario which may need to be broken down into more specific use case scenario's such as Smart Cities, Smart Home, Smart Health etc.

3.18.2 Use case requirements

Description in a nutshell	A key focus for the proliferation of IOT devices is realisation of a connected world where devices and sensors are connected in a seamless manner to support humans in the daily activities of living. The combination of sensor, actuators, mobile devices, ubiquitous high-speed wireless connectivity and cloud infrastructures will the support the realisation of a smart society where intelligence is embedded into all aspects of daily life such as smart transport, smart home, smart health and wellness and so forth.	
Technical Requirements	Aggregation	Base stations for IOT devices will need to be deployed at sufficient density to ensure that at least one wireless connection type is available at all times to the IOT device.
	Orchestration	Intelligent Orchestration will be required to manage the IOT aggregation nodes at large scales to enforce consistent device policies. An orchestrator will be capable of isolating one or more base stations from the network when an anomalous event occurs such as a security breach due to a malicious attack.
	Connectivity	Aggregators will need to support a variety of front side radios such as Bluetooth, Bluetooth Smart, Ant, Zigbee, Wi-Fi etc. for sensor or device communications. Software define radios maybe a requirement to eliminate the need to implementation multiple independent basebands.
	Flexibility	A base station will have the ability to support multiple use cases simultaneously or can be rapidly reconfigured via a remote firmware push to support either a new service to extent its existing service offerings
	Scalability	Scalable or elastic infrastructure to support transient demand peaks
	Security	Aggregators need to provide secure connection to sensors and actuators Aggregators need to provide secure transfer of data over the network to point of final aggregation e.g. Cloud DC.
	Reliability	Reliable communication should be available at times independent of location.

Business Requirements	Real-time and offline analytics	Real-time analytics are required are required to generate actionable insights at an orchestration level to react to dynamic changes in the behaviour of aggregators due to changes in the density or types of devices connecting to them. Offline analytics are required for build intelligence on the behaviour of the network and the creating insights in the patterns of data generation or information consumption.
	Context Information	Need the ability to monitor, collect and analyse context information in order to tailor service
	Outdoor and in-door connectivity	Device connectivity is required both in-doors and outdoors. Outdoor connectivity should be subject to minimal weather influence.
	Support device heterogeneity	A wide of variety devices need to be supported ranging from environmental sensors, body worn device, health devices, home sensors etc.
	Support for any-to-any communications	Support various uplink or downlink communications, device to device and node to node self-backhauling
	Ability to collect sensor data from user smartphones	Sensors on smart phones provide valuable data which can be fused with other data from the phone such as location to provide context.

3.19 Internet of things

3.19.1 Use case description

Rough Classification	Main: • IoT Platform Secondary classification labels: • Machine 2 Machine communication • IoT Virtual Network
Source of the scenario	IoT is a well-known scenario, experts estimate that the IoT will consist of almost 50 billion objects by 2020, the number of smart objects connected to the Internet will inflate the scale of the network up to two

	or three order of magnitudes and will bring a never seen pervasive interaction between humans, hosts, things in any kind of combination.
Scenario description in a nutshell	Superfluidity and interoperability in the IoT: the ability to instantiate IoT services on-the-fly (pub/sub brokers, gateway between different IoT networks), run them anywhere in the network edge (micro server) and shift them transparently to different locations, integrating vertical and/or proprietary platforms and devices. The ability to create on the fly a network (ad-hoc or/and infrastructure) composed by smart objects. The ability to use local communication (LTE Direct, BLE) instead of remote when needed: e.g. not all nodes can reach the internet, e.g. some nodes have to save energy.
Extended description and examples	The Internet of Things (IoT) is the “connection of physical things to the Internet”, which makes it possible to access remote sensor data and to control the physical world from a distance. As regards, IoT services and vertical applications examples are endless and include sectors such as healthcare, security/surveillance, transport, factory (Industry 4.0), energy/grid, agriculture. Most of this examples can be categorized in two main categories based on how the IoT devices connects to the internet: i) Using a specialized protocol, e.g. ZigBee, BLE, BACnet ii) Integrating the device into the IP-world, e.g. 6LoWPAN, CoAP, MQTT. In both cases the IoT devices connects to the internet using a kind of “gateway” that provides the entry to the network (i) or that translate between protocols (ii, e.g. 6LoWPAN to IPv6, CoAP to HTTP). The functions, performed by these gateways, could be virtualised using the Superfluidity platform (e.g. on a microserver). As a concrete example we can think to a sensors network used for environmental/structural monitoring. Sensors can be deployed in a random topology, or the topology can vary due to moving sensors or environmental changes. In this case having ability to virtualise/move network functions is a fundamental property. For example: it will be possible to keep the data aggregation functions near nodes that are actually producing more data or near nodes that needs to save energy, it will be easy to integrate new sensors and new protocols on the fly just upgrading a VM and to change routing depending on different metrics.
Architecture and/or networking and/or technical challenges	1) Scalability: support for dense crowds of devices, users 2) Interoperability: efficiently working with the inherently heterogeneous mix of communication protocols and media 3) Self-configuration and reconfiguration 4) Support different network topologies: nodes can be deployed in deterministic or random topologies; in both cases, the network must

<table>
<tr><td colspan="2"></td><td>adapt to the different topologies and adapt to changes (energy level of the nodes, unpredictable operating environment, different Data reporting model)
5) Support to mixed scenarios with ad-hoc and infrastructure network
6) Seamless combination of heterogeneous wireless accesses</td></tr>
<tr><td colspan="2">Benefits and innovation</td><td>Today the IoT scene is composed by a set of different solutions: different hardware platforms, different radio interfaces, different communication protocols and different development platforms. All this has created a myriad of different IoT solutions, in case of both open and proprietary solutions the end user (user or developer) must actively build and manage the communication infrastructure between the different connected objects and between the different IoT systems. The coming of 5G networks with the development of a platform that can easily integrate and manage existing IoT networks could be a key aspect in the creation of an IoT ecosystem where the user, or the developer, do not have to worry about creating and manage the network, but they can simply use the device / service.

The benefits of such platform are manifold:

- easily exploit new network
- partition/share an IoT network infrastructure
- classification of data produced by IoT networks
- easily aggregate and process data produced by IoT at the edge
- hybrid IoT network</td></tr>
<tr><td rowspan="2">ANALYSIS (rough)</td><td>Foreseen components, functions, or primitives</td><td>- layer that abstract the different network types (wireless, BLE, LTE Direct, LoRa…)
- dynamic configuration of connected objects
- algorithms that build the network topology and/or position the network functions using different metrics (battery consumption, location, meteorological conditions, time of day ....)
- service discovery
- - identify smart object and the conversation between them</td></tr>
<tr><td>Foreseen Requirements,

Performance issues</td><td>- handle a high number of connected devices
- low energy consumption (computation/communication)
- privacy & security</td></tr>
</table>

3.19.2 Use case requirements

Description in a nutshell	Superfluidity and interoperability in the IoT: the ability to instantiate IoT services on-the-fly (pub/sub brokers, gateway between different IoT networks), run them anywhere in the network edge (micro server) and shift them

	transparently to different locations, integrating vertical and/or proprietary platforms and devices. The ability to create on the fly a network (ad-hoc or/and infrastructure) composed by smart objects. The ability to use local communication (LTE Direct, BLE) instead of remote when needed: E.g. not all nodes can reach the internet, e.g. some nodes have to save energy.	
Technical requirements	Integration of IoT gateways in RAN	IoT devices connects to the internet: i) Using a specialized protocol, e.g. ZigBee, BLE, BACnet ii) Integrating the device into the IP-world, e.g. 6LoWPAN, CoAP, MQTT. In both cases the IoT devices connects to the internet using a kind of "gateway" that provides the entry to the network (i) or that translate between protocols (ii, e.g. 6LoWPAN to IPv6, CoAP to HTTP). The functions, performed by these gateways, could be virtualised using the Superfluidity platform (e.g. on a microserver).
	Integration with local communication (LTE Direct, BLE)	- Self-configuration and reconfiguration - Support different network topologies: nodes can be deployed in deterministic or random topologies; in both cases, the network must adapt to the different topologies and adapt to changes - Support of mixed scenarios combining ad-hoc networks and provisioning of network infrastructure - Seamless combination of heterogeneous wireless accesses
Business Requirements	Support of third party access and gateway services in RAN	The IoT gateways and services could be provided by third party entities that should be allowed to deploy their own virtualized entities in the RAN. Computing resources in the RAN would be shared by the RAN operator to the third party providers.

3.20 Third-party in-network processing

3.20.1 Use case description

Rough Classification	Main: • In network processing for third-parties

<table>
<tr><td colspan="2"></td><td>Secondary classification labels:
• Cloud architecture
• Network function virtualisation as a service</td></tr>
<tr><td colspan="2">Source of the scenario</td><td>Content providers are keen to place their content and processing close to users, and mobile apps want to leverage nearby processing for quick and cheap offloading. Until now, only major providers such as Google or Microsoft could afford to do this (e.g. the Google Global Cache).</td></tr>
<tr><td colspan="2">Scenario description in a nutshell</td><td>Mobile operators are deploying racks of machines already for NFV and this trend will intensify in the future with in 5G. Superfluidity will add processing capabilities to all parts of the network.

Such capabilities could be rented out to third parties that would run processing there. The key challenge is to ensure scalability and security in this content: third-party processing should not attack the operator’s network or other tenants.</td></tr>
<tr><td colspan="2">Extended description and examples</td><td>How can we enable the following functionality:
- A mobile app wants to batch incoming messages to save mobile device energy and to get them through the firewall of the operator.
- A client wishes to filter traffic by renting processing close to the source (e.g. DDoS filtering)
- A content provider wishes to quickly create a dynamic cache close to users in an operator’s network.</td></tr>
<tr><td colspan="2">Architecture and/or networking and/or technical challenges</td><td>1) How to ensure third party processing is safe for the operator in a cheap way.

2) How to ensure scalability to many tenants per box

3) How to ensure we can use different types of commodity hardware to ensure performance while giving isolation.</td></tr>
<tr><td colspan="2">Benefits and innovation</td><td>This use case can democratize in-network processing and allow anyone in the Internet to innovate, not just the major players. This will unleash creativity and will evolve the network in an organic way.</td></tr>
<tr><td rowspan="2">ANALYSIS (rough)</td><td>Foreseen components, functions, or primitives</td><td>- language that enables in-network processing that can be checked statically for security

- Implementation based on single-function VMs (e.g. ClickOS)

- Ability to run on different types of hardware.</td></tr>
<tr><td>Foreseen Requirements,

Performance issues</td><td>- trade-off between expressiveness of language and the ability to statically analyse it

- trade-off between using sandboxing (runtime cost) vs. static analysis (pre-processing cost)</td></tr>
</table>

3.20.2 Use case requirements

Scenario description in a nutshell	Mobile operators are deploying racks of machines already for NFV and this trend will intensify in the future with in 5G. Superfluidity will add processing capabilities to all parts of the network. Such capabilities could be rented out to third parties that would run processing there. The key challenge is to ensure scalability and security in this content: third-party processing should not attack the operator's network or other tenants.	
Technical Requirements	Latency	Static analysis must be performed at instantiation time of network processing, or interactively before deployment (e.g. in the Openstack GUI). It must therefore be able to run in well under a second for typical configurations.
	Throughput	The static analysis engine must be provisioned to cope with the expected verification load. This can be achieved by scaling it out.
	Controller	The static analysis tool will be connected to the Openstack and SDN controller of the provider in order to receive client configurations in real time and to check data plane configurations.
	Accuracy	Static analysis must not provide incorrect answers. In particular, for security questions such as "can this processing component send traffic to destination X?" a *no or yes* answer must be guaranteed to be correct (i.e. under no circumstance can the processing component send traffic to X, or it will send traffic to X always). When the answer cannot be determined with accuracy, it is preferable to provide a "don't know for sure" answer instead, such that the operator can, for instance, send the component traffic through a firewall.
Business Requirements	Hardware cost savings	Static analysis allows massive consolidation of user processing components onto the same platform without the need to perform costly traffic filtering. This will make it much cheaper for operators to offer third-party processing.
	Cheaper management	Static analysis allows operators to check configurations before deployment, and reduces the

		need to do costly data-plane deployment debugging with expert personnel.

3.21 Preventing NDP (Neighbour Discovery Protocol) spoofing

3.21.1 Use case description

Rough Classification	Main: • Implementation of virtual functions to prevent NDP Spoofing Secondary classification labels: • Security • Software Defined Networking • Network Functions Virtualisation • Hardware Accelerators
Source of the scenario	IPv6 NDP (Neighbour Discovery Protocol) Spoofing is analogous to the ARP one, but since IPv6 is not as extended as v4, some companies ignore this security threat due to the relation between solution cost and risk.
Scenario description in a nutshell	Implementation of a defence against NDP-Spoofing, for example SEND protocol, via Software.
Extended description and examples	The exchanges of NDP messages in datacentres is a matter of time, as IPv6 is destined to replace v4. For that reason, it is imperative to implement a solution to a potential spoofing in the Local Area Network, as it has been done with ARP. The inclusion of a defence against the mentioned threat, such as virtual implementation of the Secure Neighbour Discovery protocol in a superfluid network would be a great addition for the near future, furthermore, if hardware accelerators allow to speed up some of the functions of the protocol.
Architecture and/or networking and/or technical challenges	1) ensure security in LANs that use Neighbour Discovery Protocol 2) solution should have lower costs than current ones in order to achieve a better acceptation among companies
Benefits and innovation	Today, this kind of application is not considered a priority but as IPv6 continues to grow, anti NPD-Spoofing will become necessary. An early superfluid development offers a strategic opportunity for the distribution of a product that will become essential in the near future, while there is almost no competence.

ANALYSIS (rough)	Foreseen components, functions, or primitives	- virtualised NDP functions - anti spoofing mechanisms - potential hardware accelerators of the previous mechanisms
	Foreseen Requirements, Performance issues	- effective defence against spoofing and impersonation - lower costs than previous solutions - potential higher speed than previous solutions
Any other note or comment		The router vendors offer solutions to this particular problem since a few years ago, so costs are not as high as they used to be and will become lower as time passes. Despite these cons, 'Superfluidity' should be compatible with IPv6 security as it seems the future of the internet.

3.21.2 Use case requirements

Description in a nutshell	Implementation of a defence against NDP-Spoofing, for example the SEND protocol, by means of VNFs.	
Technical requirements	SEND protocol	Implementation of the SEND protocol
	Detection time	Seconds or minutes
	Identification time	Seconds or minutes
	Reaction time	Seconds or minutes
	Effectiveness	At least similar to current implementations Measured as: • Percentage of attacks detected • Percentage of attacks blocked • Percentage of false positives
	Building Blocks	Implementation of the following building blocks: • Idle: discards packets • Switch: sends packet stream to settable output • Hub: duplicates packets like a hub • Queue: stores packets in a FIFO queue • EtherSwitch: learning forwarding Ethernet switch • CheckIPHeader: checks IP header • IPFilter: filters IP packets by contents • SEND: SEND protocol implementation

Business Requirements	Protection against NDP Spoofing	Protection against NDP Spoofing
	Cost savings	Development and deployment costs less than current solutions
	Deployment time	Deployment time much shorter than current solutions
	Reaction time	Reaction time shorter than current solutions
	Effectiveness	Effectiveness at least equal to current solutions

3.22 Protection against DDoS (Distributed Denial of Service) attacks

3.22.1 Use case description

Rough Classification	Main: • Defence against DoS attacks Secondary classification labels: • Cyber-Security • Cloud architecture • Software defined networking • Network function virtualisation
Source of the scenario	Several small companies have suffered Denial of Service attacks in recent times. The attackers are hard to stop because of the distributed nature of the attack and the companies tend to succumb to blackmail. For this reason defences against DDoS attacks become more proficient over time and use every new useful technology, including parts of Superfluidity such as Cloud Architecture and Software Defined Networking.
Scenario description in a nutshell	Exploit of the elements that compose Superfluidity, to design defences against the increasing number of DDoS attacks.
Extended description and examples	Distributed Denial of Service attacks are currently hard to deal with. There are efficient defences against them but not infallible, nor universally extended. Technologies included in Superfluidity such as cloud architecture and software defined networking have been proposed as a defence against these attacks.

		Start-ups could use in their data centres, a defence already implemented in the network with a minor cost.
Architecture and/or networking and/or technical challenges		1) identification of possible defences against DDoS attacks 2) determine which are implantable in superfluid networks 3) propose the implementation of the most efficient of the studied defences
Benefits and innovation		Cybersecurity is a field in constant flow, as attacks evolve so must defences. Distributed Denial of Service attacks, although being relatively simple, are troublesome for start-ups with data centres but insufficient defence. A defence against DDoS attacks already implemented in the network means an added value to the Superfluidity project. A defence which implementation also includes other network functions could probably have an overall lesser cost than an isolated functions. This offers a business opportunity to sell the DDoS superfluid defence as low cost, especially among recent and small companies.
ANALYSIS (rough)	Foreseen components, functions, or primitives	- means to identify DoS attackers - drop traffic from attackers - determine if former attackers are no longer infected and restore traffic flow if they're not - potential use of hardware accelerators to speed up the performance of the defensive mechanisms
	Foreseen Requirements, Performance issues	- security issues - solution shouldn't slow the network - minimise the effect of the defence in traffic coming from non-attackers - economic issues - solution should be open for future improvements and updates
Any other note or comment		Security may not be a priority for the development of superfluid networks, but it eventually becomes an essential part of every network. For this reason, early attempts of implementing security in 'Superfluidity' should at least be taken into consideration. Link to another project dealing with security: ENSURE

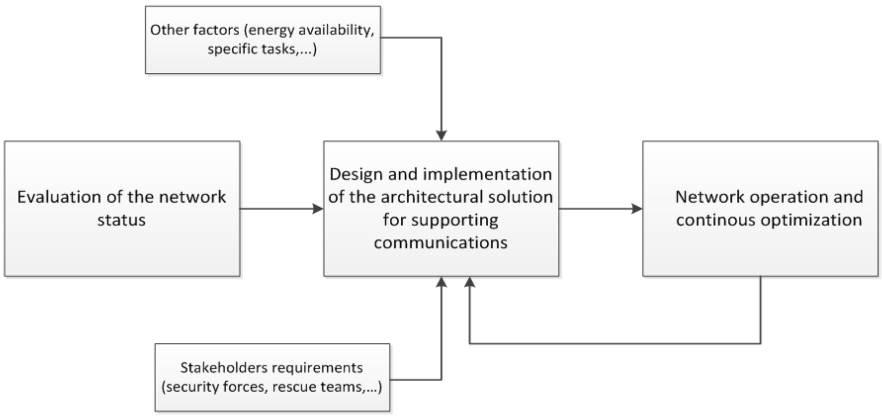
3.22.2 Use case requirements

<table>
<tr><td>Description in a nutshell</td><td colspan="2">Implementation of a defence against DDoS attacks, by means of VNFs.</td></tr>
<tr><td rowspan="7">Technical requirements</td><td>Detection time</td><td>Seconds or minutes</td></tr>
<tr><td>Identification time</td><td>Seconds or minutes</td></tr>
<tr><td>Reaction time</td><td>Seconds or minutes</td></tr>
<tr><td>Restoration time</td><td>Seconds or minutes</td></tr>
<tr><td>Tracking time</td><td>Seconds or minutes</td></tr>
<tr><td>Effectiveness</td><td>At least similar to current implementations
Measured as:
• Percentage of attacks detected
• Percentage of attacks blocked
• Percentage of false positives</td></tr>
<tr><td>Building Blocks</td><td>Implementation of the following building blocks:
• Idle: discards packets
• Switch: sends packet stream to settable output
• Counter: measures packet count and rate
• Queue: stores packets in a FIFO queue
• IPRouteTable: IP Routing Table superclass
• RadixIPLookup: IP lookup using a radix try
• CheckIPHeader: checks IP header
• IPFilter: filters IP packets by contents
• IPRateMonitor: measures incoming and outgoing IP traffic rates</td></tr>
<tr><td rowspan="5">Business Requirements</td><td>DDoS Protection</td><td>Protection against DDoS attacks</td></tr>
<tr><td>Cost savings</td><td>Development and deployment costs less than current solutions</td></tr>
<tr><td>Deployment time</td><td>Deployment time much shorter than current solutions</td></tr>
<tr><td>Reaction time</td><td>Reaction time shorter than current solutions</td></tr>
<tr><td>Effectiveness</td><td>Effectiveness at least equal to current solutions</td></tr>
</table>

3.23 Emergency communications support

3.23.1 Use case description

Rough Classification	Main: • Network auto-configuration Secondary classification labels: • Location based services • Monitoring and analytics • Adaptive media delivery • Content centric networking • Enhanced group communications • Intrinsic security mechanisms
Source of the scenario	Emergency communications is one of the classic use cases considered for 5G (e.g., NGMN, METIS). 3GPP is also working to improve for capabilities specifically aimed at critical communications requirements. It even has created a new specification group, SA6, Mission-critical applications, which is responsible for the definition, evolution and maintenance of technical specification(s) for application layer functional elements and interfaces supporting critical communications (e.g. Mission Critical Push To Talk). Other bodies are also considering this kind of services and the systems supporting them, like the TETRA Critical Communications Association (TCCA).
Scenario description in a nutshell	Capability of the network to repurpose itself after a disaster for providing specific services based on the remaining infrastructure available, including user devices and other networks. Support of specific services for specific groups (police, firefighters, rescue teams,…), including the capability to locate specific devices, broadcast messages and form device supported extensions to the network.
Extended description and examples	A natural disaster (earthquake, tsunami,…) or man produced one (terrorist attack, industrial accident,…) results in the loss of part of the communications network infrastructure. The existing operating infrastructure should reconfigure itself in order to provide a set of basic services • Support of emergency teams' communications. • Support of the broadcast of alert messages with enough flexibility in terms of selecting areas or user groups where location-specific alert information can be directed • Control of the operating mode of the devices connected (e.g., forcing reduced energy consumption operational modes, like half duplex)

	• Support of safety related functionalities, like the location of devices for rescue purposes. • Access control to the network in order to avoid it being saturated by lower priority traffic. The network should be able to integrate new elements once they become available. It should be able to integrate with broadcast and satellite systems.
Architecture and/or networking and/or technical challenges	1) The network should be able to evaluate its own status in order to determine which is the best configuration feasible for providing the functions expected 2) The network should be able to learn from its own operation in order to optimize the configuration and the use of resources  *Figure 5: Approaches for supporting Emergency Communications* 3) The network should be able to incorporate new elements and nodes as they become available. 4) The network should be able to prioritize the communication services for certain groups, like rescue teams. At the same time, it should be possible to block users to prevent that they make the network unusable 5) It should be possible to configure the network in order to support specific tasks, e.g., the location of victims in collapsed buildings. 6) It should be possible to seize resources (e.g., frequency channels) from other operators or systems to support the operation. 7) It should be possible for the network to activate the safety capabilities of devices without the users' intervention. 8) The network should be able to provide intrinsic security if the centralized security infrastructure (e.g., HSS/AuC, MME in LTE) is not accessible
Benefits and innovation	Today, specific communications networks (TETRA 1, TETRAPOL and P25) should be deployed for supporting emergency/mission critical

		communications. However, these services can be more effectively supported by commercial networks infrastructure. Two main sets of innovations are required for providing future mobile networks with the necessary capabilities to support this kind of communications: • The capacity to configure a new network from the infrastructure that is available, possibly using resources (spectrum, processing and networking resources) from other systems and networks. • The capacity for supporting specific services and operational modes required for safety procedures. These innovations may be reused for supporting other commercial mobile wireless services.
ANALYSIS (rough)	Foreseen components, functions, or primitives	- Means to monitor, collect and analyse context information. - Capability for deriving the topology of D2D supported subnetworks - Ability to gather sensing information from user terminals - Network elements supporting content centric networking capabilities
	Foreseen Requirements, Performance issues	Several KPIs can be defined, assuming a minimum infrastructure available: - Time needed by the network to set up the "emergency mode" and make the service available shall not exceed - Capacity to support a minimum number of simultaneous connections - Capacity to locate specific devices within a given time frame - Capacity for operating without energy supply for a certain period of time

3.23.2 Use case requirements

Description in a nutshell	Capability of the network to repurpose itself after a disaster for providing specific services based on the remaining infrastructure available, including user devices and other networks. Support of specific services for specific groups (police, firefighters, rescue teams,…), including the capability to locate specific devices, broadcast messages and form device supported extensions to the network.	
Technical requirements	Emergency detection	The network should be able to detect when an emergency situation has happened and, consequently, the emergency communications operational mode should be activated. This activation should happen

		with enough security guarantees to avoid false alarms that may disrupt the normal operation of the network.
	Self-evaluation	After an emergency situation has been detected, the network should be able to assess the resources that are available for the support of the services. For these purposes, the network should be able to carry out connectivity tests as well as being able to report the processing resources available at different locations. These self-evaluation procedures should encompass the infrastructure from different operators and tenants
	Self-configuration	Based on the results of the self-evaluation process, the network should be able to configure itself in order to provide emergency services. In this sense it should be able to assign roles and allocate functionalities to the resources available.
	Resources grabbing	The network should be able to take resources from other applications in order to support emergency communications. They may include, e.g., spectrum resources allocated to other services like broadcast services, or computing and storage resources in data centres
	Service prioritisation	The network should be able to prioritize the different service requests according to the policy that has been defined for emergency situations. This may include blocking end user services in order to avoid congestion situations.
	Broadcasting support	The network should support the broadcasting of specific messages, like alerts or instructions for the population.
	Energy saving	The network should be able to control the configuration of devices and network elements in order to save energy, especially if access to the power grid is not feasible or is not reliable
	Special services support	The network should be able to support special services, like group communications between specific groups (police, firefighters, rescue teams,…), location assistance services (e.g., for people under collapsed buildings), etc.
	Network expansion through D2D	The networks should be able to use UEs in order to extend the coverage area. The network should be able

		to discover the topology of the connections established among the devices.
	Name based addressing	The network may be able to support the access to specific contents based on their name, regardless of their collection.
	Network securing	The network should be able to provide basic security services even in the case that the security infrastructure is not available or reachable.
		The network should be able to prevent or minimize the effect of Denial of Service (DoS) attacks.
Business Requirements	High availability	For certain network resources, high availability should be provided in order to increase the network survivability in case of disaster.
	Reusability	A subset of the procedures defined for emergency communications should be reusable for dealing with other, less dramatic situations, like failures in different parts of the network, loss of connectivity, sabotages, etc.
	Public private partnership	Public sector should be involved in the definition of the procedures and policies to be used for emergency communications, as well as in the financing of the additional infrastructure that may be required to guarantee a certain degree of reliability and survivability.

4 Clustering of use cases and their requirements

Section 3 describes 23 use cases, which in total produce quite a substantial list of requirements. It is clearly intractable to go through each individual use case, let alone each individual requirement, and consider how each could impact the Superfluidity architecture. Therefore we examined alternative ways for "clustering" both the use cases and the requirements. This work is described in Appendix C. At the end of this process, we came up with a requirements-based clustering. In the list below, the first five take into account business requirements, whilst the remaining seven are groupings of technical requirements:

Business requirements clusters:

- Service agility
- Cost savings
- User Experience
- Security

- Analytics

Technical requirements clusters

- Orchestration
- Quality of experience
- Scalability
- Platform
- Converged architecture
- Building blocks
- Security

4.1 Business requirements clusters

We examined the business requirements for the use cases (Section 3) and derived several business-related “clusters” of requirements.

The topic that clusters most requirements is service agility.

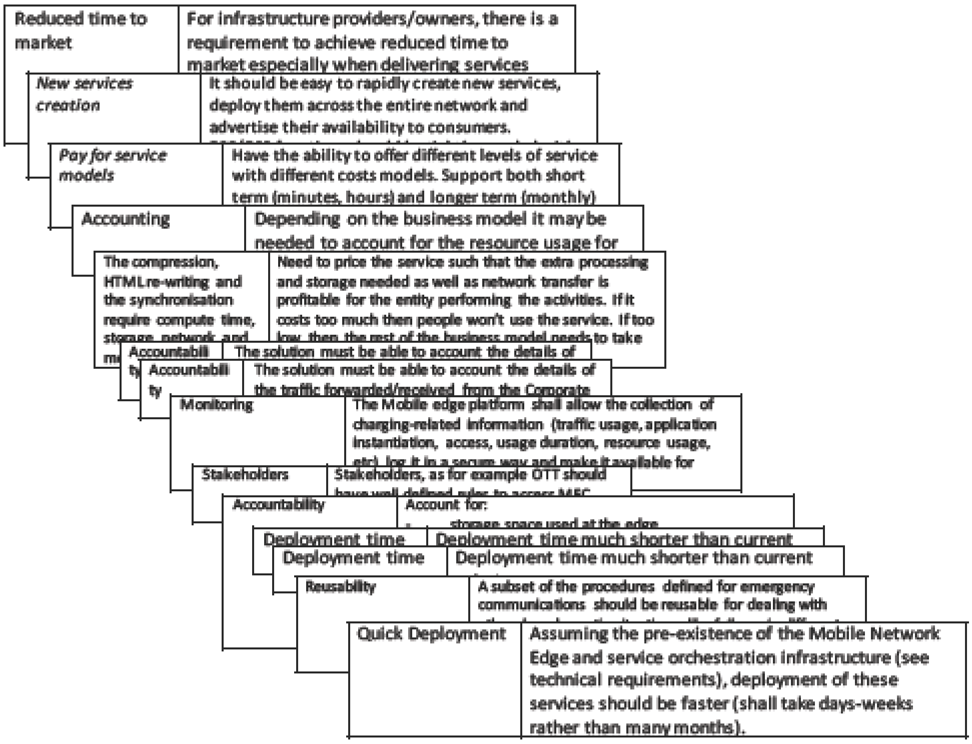


Figure 6: Business requirements: Service agility cluster

A significant number of requirements deal with cost savings:

Figure 7: Business requirements: Cost savings cluster

There are two clusters that reflect the centrality of the user, namely increased user experience and security concerns. After these four main clusters, another smaller cluster formed around the need for analytics embedded in the network as a means to extract intelligence from the network, to be fed back into the business processes.

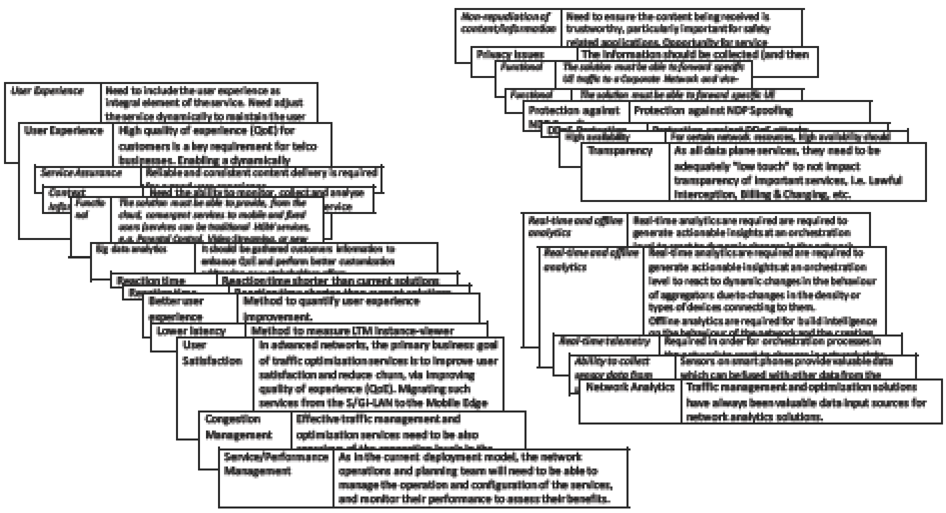


Figure 8: Business requirements: User experience, Security and Analytics clusters

Finally, there are other topics that form smaller clusters or cannot be included naturally in other clusters.

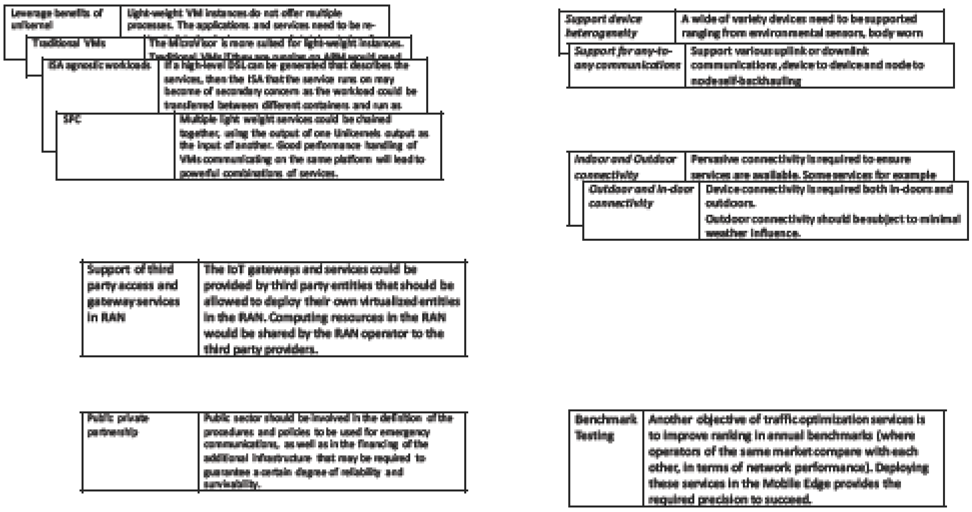


Figure 9: Business requirements: Minor clusters

The five main clusters reflect the increase in importance of these topics in 5G networks and how Superfluidity as a project ranks them. Figure 10 shows the interactions between the clusters in the Superfluidity use cases.

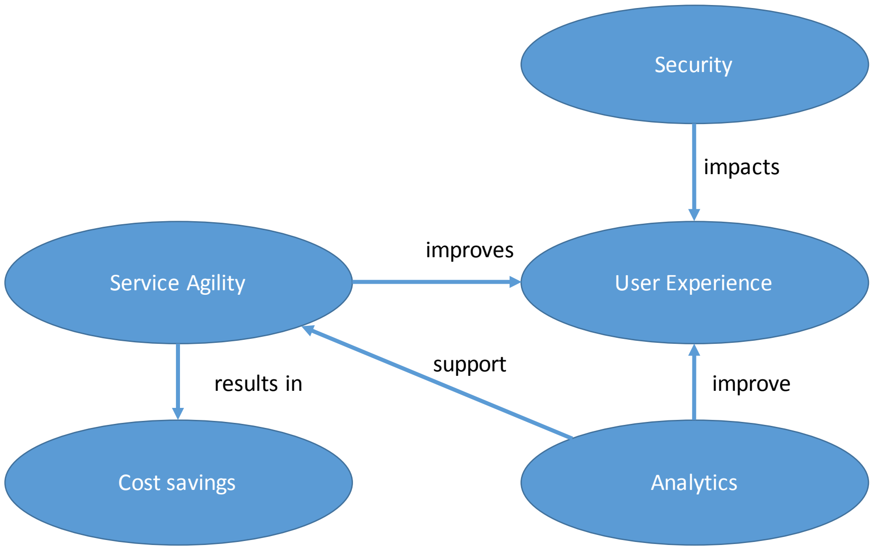


Figure 10: Interactions between the five main clusters of business requirements

4.2 Technical requirements clusters

We examined the technical requirements for the use cases (Section 3) and derived several technical-related “clusters” of requirements. The clustering was done in a more graphical way than for the business requirements, both because of the heterogeneity and because of the amount of requirements that were derived.

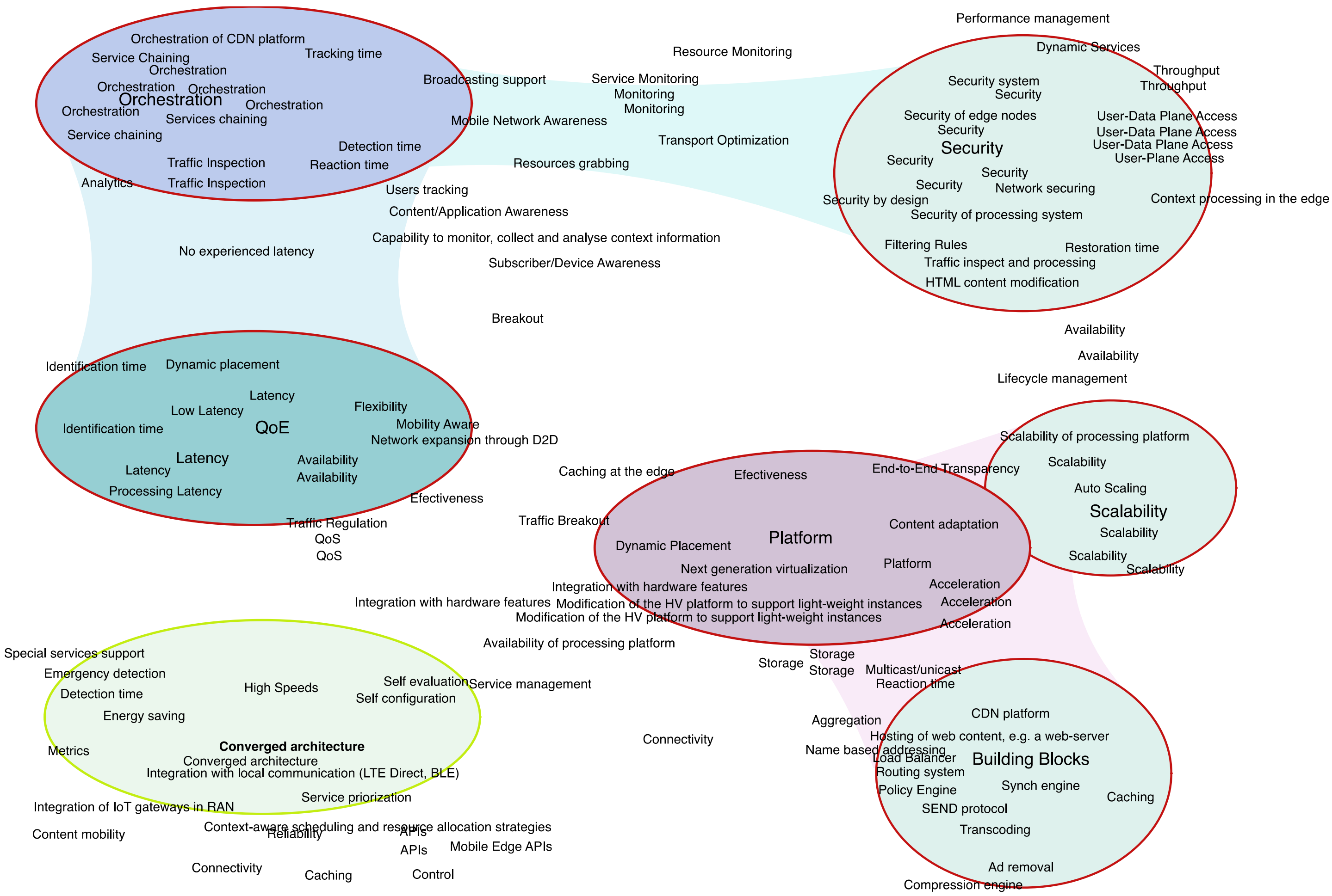


Figure 11: Overall clustering of the technical requirements

5 Impact of requirements on the Superfluidity architecture

In the previous section, we have identified 12 main clusters of system requirements, which can be classified as business requirements (service agility, cost savings, user experience, security and analytics) and technical requirements (orchestration, quality of experience, scalability, platform, converged architecture, building blocks and security).

In the following we will not differentiate between business level and technical level requirements and merge the two “security” clusters, as well as the “user experience” and “quality of experience” into a single cluster. Our final list includes the following clusters of requirements:

- Service agility
- Cost savings
- Security
- Analytics and Metrics
- Orchestration
- Quality of Experience
- Converged architecture
- Platform
- Scalability
- Building blocks

In this section we consider these clusters of requirements, in terms of their likely impact on the Superfluidity architecture and on the wider technical work in the project.

We are using an iterative approach (rather than a strict ‘top-down’ method where architecture would be derived from the requirements which had in turn been derived from the use cases). So in fact our requirements analysis takes into account the ongoing work on the Superfluidity architecture, which has itself been influenced by the initial work on use cases and requirements. Therefore, the analysis reported in the following subsection is not meant as an abstract identification of desirable functional and non-functional properties of a generic platform supporting 5G scenarios, but it is driven by some assumptions about the Superfluidity architecture under definition.

It is worth reporting some of these assumptions, in order to understand the context and clarify the terminology used in the next subsections.

The Superfluidity project aims at defining an agile, flexible and re-programmable architecture to support 5G scenarios relying upon existing architectural concepts and looking for their convergence. We build on the existing concepts and frameworks of NFV and SDN.

The Superfluidity architecture considers as a unified platform the set of resources provided by heterogeneous logical domains like Cloud-RANs, Mobile Edge Clouds, Cellular Core, and Traditional Clouds. Considering the physical deployment, these domains are mapped in heterogeneous physical data centres (Cell-site, Local, Regional, and Central Data Centres). In this respect, the Superfluidity architecture provides a *horizontal* integration of resources.

Another fundamental aspect of Superfluidity’s architecture is the decomposition of high-level monolithic functions into reusable components. This is based on the concept of the *Reusable Functional Block* (RFB). RFBs can be composed to realize services or to form more complex reusable components (details in sub-section below). The concept of RFB will be fully defined and detailed in

Deliverable D2.2 and Deliverable D3.1. The particularity of the approach is that it is recursively applied at different levels and in different contexts (*hosting environments*). Therefore, a Reusable Functional Block can correspond to a traditional VNF or VNFC implemented as a fully-fledged VM running in a hypervisor or in a container. However, a RFB can also be a tiny Unikernel VM running in a specialized hypervisor, or a module or component in a special purpose hosting environments, like a software router or radio signal processing chain. In this respect, the Superfluidity architecture provides a (conceptual) *vertical* integration of the different levels.

We have started to consider that there are three types of functionality: in-line functions, i.e. related to data plane and packet processing; in-line monitoring, i.e. the capability to measure, collect and analyse performance indicators at different levels in real time; and orchestration and management, i.e. the composition of RFBs, their mapping into the available resources offered by the execution environment, and their deployment over the infrastructure.

5.1 Service agility

5.1.1 Summary of the cluster of requirements

Service agility can be defined as the capacity of the operator to control and manage the user experience from the device through the entire network in order to create new services with short time to market (TTM), as well as allow for consumers and enterprises to personalize their services. For achieving these objectives, the network should provide the required agility in service creation, delivery and management.

It is recognized that network functions virtualization and software defined architectural frameworks may become significant enablers for service agility. Nowadays, it can take weeks or even months for operators to provision business services, if they can be supported with the existing infrastructure, but if these services require new supporting infrastructure, it can take more than six months of lab work for type approval and other certification processes. The new technologies will allow operators to offer context-aware network services, or the ability to ensure varying levels of service and performance according to application, user or location. With them, operators can define QoS/QoE goals for services according to users' specific needs and meet those goals at the network level. Additionally, they can integrate managed security and availability features into applications dynamically. Network operators will use this capability to differentiate their own higher-level services by tying them directly to network behaviour. Ultimately, providers will be able to charge accordingly for these granular services and performance assurances.

5.1.2 Likely extent of impact of requirements on Superfluidity

Service agility cluster is analysed here from a business perspective, and in that sense, almost all the requirements that are related with offering a better service, accordingly with customer needs at the proper time is included in this cluster.

The Table below shows the list of requirements and how important each one is regarding with what is envisioned as having the support of stakeholders such as SDOs, operators, academia and industry. The most important requirements are related to service deployment and creation time. The

Superfluidity architecture must deal with all the aspects that will impact on those requirements, at all levels, from infrastructure to applications, closely impacting with orchestration and management.

5.1.3 Likely impact of requirements on Superfluidity Architecture

Requirement	Impact to Architecture	KPIs
Reduced time to market New services creation Deployment time	• A framework for an easy and simplified service CRUD. • Compliant with SDOs, promoting quick alignment with industry • A set of built-in services to application developers	• Time frame to develop a new service, by using and re-using Reusable Functional Blocks (RFB)
SFC	• The ability of chaining services by making available a service catalogue. • The orchestrator must provide means to enable a dynamic the combination of services	
Dynamic Allocation of Resources	• Service/application continuity and mobility • Orchestration between involved entities • Context awareness	
Monitoring	A means for supporting monitoring: • xDRs enabling charging and billing; • quality related parameters (retrieved from network and user consumption)	• % of parameters values and scales monitored • Network metrics
Reaction time	A shorter reaction time to security threats and attacks	Benchmark with current measured values
Reusability	A prompt response to disaster recovery or other system downtime, by reusing and distributing modules following a redundancy strategy	
The compression, HTML re-writing and the synchronisation require compute time, storage, network and memory	The use of mechanisms to accelerate services and offload processing and storage, to ensure a better performance.	
Context Information	Access to MEC services (such as RNIS and LOC).	Capability of collecting and analysing context

		information in order to tailor service

5.2 Cost savings

5.2.1 Summary of the cluster of requirements

Cost savings much often begins with IT infrastructure. In the context of telecommunications networks it has always been a target for budget cost, as it is most certainly being near the top of an operator's expense list. There are several trends in IT, but few tend to impact the overall business as significantly as virtualization. Cost savings relate to technical requirements and several properties can be noted with regards to virtualization technology as one of the Superfluidity architecture's main drivers:

- Optimization of the use of existing resources.
- Simplification of Infrastructure and software administration, maintenance, and deployment.
- Reduction of hardware needs, resulting in less power consumption, less space required, and lower cooling costs.

5.2.2 Likely extent of impact of requirements on Superfluidity

Cost savings should be a straightforward outcome from the Superfluidity's work.

5.2.3 Likely impact of requirements on Superfluidity Architecture

The architecture aims to use COTS (Commercial-Off-The-Shelf) hardware. This allows reduced use of vendor-specific, higher-cost equipment. Furthermore, the heterogeneous property of the architecture lowers upgrade costs, since an operator can keep a large fraction of the existing equipment. An additional property of the architecture is it being automatic and self-driven, thus reducing the IT personnel expenses. With virtualization being Superfluidity's keystone - the ability to share the infrastructure resources and reuse the same physical hardware provides a further saving. Finally, simplicity is a major guideline to the ongoing architectural work, and simpler networks should have lower operational costs.

5.3 Analytics and metrics

5.3.1 Summary of the cluster of requirements

Optimising service/platform performance and operational efficiency requires intelligent automation through Orchestration. Telemetry and full-stack instrumentation, combined with analytics are critical inputs into an orchestration system to support automation of tasks such as scaling, migration, etc. Data from the different layers within the NFVI need to be synchronized, integrated, pre-processed, analysed and presented to users or functional entities such as an Orchestrator.

A number of key requirements for both analytics and metrics can be identified from the Superfluidity use cases. Both real-time and offline analytics are required to provide actionable insights at an

Orchestration level. For example, analytics are necessary to identify network behaviour which could adversely affect service behaviours or performance resulting in poor user experience. It is expected that a Superfluidity system has the potential to generate very high volumes of data, therefore a big data type solution for the aggregation, persistence and access provisioning to consuming analytics services is likely to be necessary.

Analytics requires data to analyse, and thus needs metrics to be measured. Metrics can be collected from all levels within the software and infrastructure stacks and with different temporal granularities, formats etc. It will be necessary to fuse the data into a common namespace which encodes the relationships between the metrics data. The availability of Telemetry data in real-time is also critical in order for an Orchestrator to react in an appropriate manner to changes in the context of a service such as approaching an SLA violation, need to scale or migrate etc. Metrics are also critical for quantifying service performance particularly from a user perspective. Metrics are necessary to determine the key performance indicators for services such as quality of service or quality of experience. In additional metrics within the Superfluidity system will be used for other purposes beyond Orchestration such as billing based on service or resource utilisation.

5.3.2 Likely extent of impact of requirements on Superfluidity

Metrics and analytics will play a key role in the Superfluidity system. The Superfluidity system must collect real-time metrics across the NFVI and the network. The metrics relate to the infrastructure (Compute, Storage and Network), the environment e.g. virtualization layer and the service layer (applications and NFV). The amount of metrics to be collected and processed will be of the orders of magnitude higher than in existing systems, given the scale of expected use cases.

Current approaches to metrics may not scale sufficiently. New approaches to metrics collection will likely be required. For example in steady state operations low resolution metrics are collected however when an abnormal event is detected the resolution of the metrics collection can change dynamically to provide higher resolution metrics to support diagnostics investigations.

5.3.3 Likely impact of requirements on Superfluidity Architecture

Requirement	Impact to Architecture	KPIs
Service Context Analytics	• Metrics collection required across the NFVI and network • Metrics collection subsystem must expose northbound interface to Orchestrator • Analytics subsystem required as part of Orchestrator to process NFVI and network metrics into actionable insights	Collection and processing of hundreds of megabits of metrics data per second.
Charging Metrics	• Required metrics need to be published by the E-NodeB to a MEC platform.	

	• MEC service need to be in place to process and expose required metrics to the MEC application layer	
Network Analytics	• Hierarchical network metrics collection required.	Standard Network Node Level KPI • Network Throughput • Latency • Jitter • Errors
Data Fusion	• Metrics aggregation required from distributed sources to form a common collection metrics for the required context	• GB processed per second
Real-Time Telemetry	• Existing metrics sources such as Ceilometer could be used in addition to new subsystems that could be developed specifically for Superfluidity architecture. • The telemetry system may need to support different virtualization environments e.g. OpenStack Cloud, Docker Containers, Unikernels. • Telemetry data may need to be related to infrastructure or deployment topologies to provide insights between service performance and the quantities, types and locations of resources allocated.	• Samples per second
Quality of Experience Metrics-Analytics	• Need metrics system which can capture the diversity of metrics from different sources, UE, Content sources, network etc. in order to measure QoE and identify problem points • QoE measurements may be exposed to the Orchestrator in order to trigger scaling actions in response to poor QoE.	• Std QoE apply for different classes for content. E.g. video – avg bit rate.

5.4 Quality of experience

5.4.1 Summary of the cluster of requirements

Quality of Experience (QoE) in the context of telecommunications networks has been defined as: "The degree of delight or annoyance of the user of an application or service. It results from the fulfilment of his or her expectations with respect to the utility and / or enjoyment of the application or service in the light of the user's personality and current state." In the context of the requirements gathering exercise described earlier in this Deliverable, QoE relates to technical requirements and user experience to business ones, and several aspects are apparent:

- Low latency is required for some applications, such as video and safety ones, perhaps <20msec latency
- Applications may need some guarantee about other aspects of the service performance, for example about the latency of the processing or the availability of the service (five nines say)
- There may need to be differentiation between the performance of different applications &/or users. This particularly applies if there are problems (bad congestion, failure of a server etc.) – for example, a favoured application will achieve better availability by being favoured in some way during the problem
- There may need to be some customisation in order to improve the QoE for a particular user /application
- The end user may move – their QoE needs to be maintained
- There needs to be the ability to measure what QoE is being delivered
- It is likely that the user (or perhaps the service provider, virtual network provider etc.) needs to pay extra for better QoE
- Emergency communications need, in extremis, to be able to grab resources from other applications

5.4.2 Likely extent of impact of requirements on Superfluidity

This cluster of requirements is likely to have quite a significant impact on the technical work of the project.

5.4.3 Likely impact of requirements on Superfluidity Architecture

We identify that QoE is likely to impact three parts of the architecture:

1. In-line functions
   - In-line functions: Transport and in-line processing should add only a small amount of latency – either to all applications or else to selected ones. Similarly for other performance parameters such as packet drops, consistent content delivery etc.
   - In-line functions: a function may need to be cognisant of potential nasty interactions with other functions (for example, radio optimisation and TCP congestion control fighting each other)

- In-line functions: some packets or processing requests may need to be prioritised over others, in order to meet the QoS requirements of the preferred applications. This is more likely to be required where the network or processing is slower – roughly speaking, at the edge of the network
- In-line functions: the chain of VNF and VM functions needs to be reliable. Most likely this requires redundancy mechanism(s), for instance fall-over to an alternative instantiation of the function, in a way that is invisible to the end user.

2. In-line monitoring
   - In-line monitoring: There needs to be the ability to measure the end-to-end performance, in order to understand what QoE is being delivered
   - In-line monitoring: monitoring may be needed for a particular segments of the network /service delivery, at a particular time, in a particular region, etc. – as controlled by Management functions. The information may be used to inform short-term network management and is likely to inform longer-term network planning
   - In-line monitoring: there may need to be special monitoring to help with Orchestration and Management activity
   - In-line monitoring: it is possible that monitoring may need to be end user aware, in some circumstances (not clear-cut)

3. Orchestration and management
   - Orchestration and management: there needs to be the ability to customise and adapt services. There are potentially lots of aspects to this, for example:
     - what service is delivered (no HD video when someone starts walking);
     - what functions are chained together (e.g. remove some security when on the corporate network);
     - where a function is carried out (e.g. at the edge or more centrally – which could be to improve the QoS for a specific user and/or to optimise the overall network, perhaps in order to improve the average QoE);
     - what optimisation function is added into the chain (e.g. a proxy to enhance performance or to regulate video quality, or a pre-cache to fetch predicted content)
   - Orchestration and management: the service offered to a user needs to be linked to AAAA (to see if the user is allowed to have the service, to bill them, etc.). These AAAA aspects may also need to depend on other factors, such as the type of device (/context), the virtual network operator, the application provider etc.
   - Orchestration and management: new instances of the VNFs and VMs need to be spawned, for example if a processor is overloaded, or if there is a fault
   - Orchestration and management: user mobility (movement of their device, movement of the user to a different device etc.) may require moving the functions that deliver the service.
   - Orchestration and management: there needs to be the ability to usurp resources (network, processing, storage) from other users in order to support emergency communications. This is only used in extreme circumstances

5.5 Building blocks

5.5.1 Summary of the cluster of requirements

The Superfluidity architecture envisages the concept of Reusable Functional Blocks (RFBs) as a way to decompose high-level monolithic functions into reusable components. The use of the RFB concept facilitates the Superfluidity goal to allocate resources (processing, networking, storage…) dynamically and efficiently. The concept of RFB will be fully defined and detailed in Deliverable D2.2 and Deliverable D3.1.

An RFB can be seen as a generalization of the concept of VNF (Virtual Network Function) and of VNFC (Virtual Network Function Component) defined in current ETSI NFV ISG. An RFB can be executed on what we call "hosting environment". For example, using the ETSI terminology, VNFC entities can be executed on a so-called "Virtualization Container", which constitutes its hosting environment.

In the current model, VNFs and VNFCs correspond to Virtual Machines running in Hypervisors. The Superfluidity architecture extends this state of the art in two directions. On one hand, it includes the concept of very lightweight Virtual Machines supporting RFBs with fine granularity. In this case, the hosting environment is a hypervisor specialized in supporting tiny Virtual Machines. On the other hand, the Superfluidity architecture will support heterogeneous hosting environments that can be used to support the RFBs. Examples of such environments are the like the click modular router, or extended finite state machines based on OpenFlow, of modular Software Radio processing platforms.

5.5.2 Likely extent of impact of requirements on Superfluidity

These requirements will have a significant impact on the technical work of the project.

5.5.3 Likely impact of requirements on Superfluidity Architecture

The key challenge and requirement for the architectural modelling is trying to define the RFB concept in a general way, applicable to different hosting environments. The model should be then specialized for these different hosting environments, because each one will be characterized by a different specific set of features.

The following requirements for the modelling of Reusable Functional Blocks in the Superfluidity architecture are identified:

1. Functional decomposition: it should be possible to decompose Complex services/ functions into simpler, reusable, functional blocs (RFBs)
2. Functional composition: it should be possible to compose Complex services/ functions by programmatic composition of simpler RFBs
3. Definition of models for expressing the functional and non-functional properties of the RFBs. The models should support description of RFBs at different levels. For example, some RFBs will have a coarse granularity and their functionality will be informally described at a high level, while other RFBs will have a fine granularity and their functionality will be formally described with some proper language.

Definition of models and languages for expressing the composition of RFBs. Tools are required for automatic execution of the composed RFBs and/or validation of some properties of the composed

RFBs. In general, this cannot be applied to every hosting environment. Such tools will be considered for RFBs operating at fine granularity in specific hosting environments.

For in-line (data plane) functions, the Superfluidity project aims at promoting a novel network processing architecture. This data plane processing architecture should: (i) support a programming abstraction which retains the platform independent features of the original "match/action'" OpenFlow abstraction; (ii) allow to execute directly on the fast path stateful flow processing tasks, formalized in a vendor-agnostic manner via extended Finite State Machines.

5.6 Orchestration

5.6.1 Summary of the cluster of requirements

Orchestration is the capability that provides automated arrangement, coordination, and management of workloads and infrastructure resources (compute/storage/network). Orchestration is a fundamental building block for any complex system such as Superfluidity. Analysis of the use cases identified significant Orchestration related requirements which have an impact on the high level functional architecture of Superfluidity.

Common across many of the use cases was an overarching requirement for the orchestrator to have a "system and network wide and heterogeneous view". Different network and service domains must be visible and orchestrated in a coordinated way, including having real-time access to those elements' dynamic state. This guarantees fulfilment of other requirements categories such as "Intelligent service placement and management", "Dynamic allocation of resources", "Intelligent service placement and management", as well as "Service composition" and "Service Chaining".

Orchestration is also seen as a tool for network healing, considering the requirements on category "Exceptional event handling".

Other requirements such as "dynamic allocation of resources", requires the infrastructure and the deployed functions/applications to support migration without losing of state.

Finally, orchestration must guarantee appropriate network and services scalability, and its own scalability.

5.6.2 Likely extent of impact of requirements on Superfluidity

Orchestration is a key functionality of the Superfluidity system. Therefore requirements relating to Orchestration must be carefully considered and addressed in the overall architecture. Dealing with RFBs will pose new challenges, increasing the number of orchestrated elements and implicitly requiring services composition to achieve the required 'macro' functions. RFB descriptors have to include additional information to take into consideration specific compositions that will answer specific requirements (e.g. low latency, high scalability, high throughput).

5.6.3 Likely impact of requirements on Superfluidity Architecture

Requirement	Impact to Architecture	KPI's

System and Network wide view	• Orchestrator needs to have an integrated view of NFVI, C-RAN, MEC and Core Network. • Multi- layer Orchestrator maybe required. • Needs access to dynamic system and platform information (NFVI and network).	
Service Composition/ Service Chaining	• Orchestration needs to have appropriate intelligence to compose RFBs, achieving the most efficient implementation of an NFV and a Service comprised of NFVs.	
Dynamic Allocation of Resources	• Orchestration needs up to date information about the current state of the NFVI to select the appropriate resources (type/location) to allocate to a service. • Orchestration needs up to date information about the current state of the NFVI to move a service to a new location.	
Scalability	• Orchestration needs access to real-time metrics • Orchestration needs to manage tens of thousands of RFBs • Orchestrate needs to execute scalability actions in the order of millisecond	• Handle > xx events/sec • Manage RFB $>10^5$ (or e.g. 10x compared to NFV) • Service scalability in milliseconds
Metrics	• Orchestrator needs real-time collection of metrics from NFVI exposed to Orchestrator. • Pre-processing and filtering of metrics at NFVI layer maybe required for scalability. • Collection and integration of RFB metrics maybe required. • Collection and integration of application metrics will be required • Sampling rate configuration at deployment time. • Ability to deal with heterogeneous metrics sources beyond services, applications and NFVI e.g. UE metrics, network statistics etc.	
Rapid Deployment	• Local catalogues of RFB images maybe required • Distributed or hierarchical orchestration by required	• < 20 ms

Intelligent Service Placement/ Heterogeneous view	• Need to expose all platform features and capabilities to Orchestrator e.g. accelerators • Need view on current state of NFVI e.g. saturation, utilisation etc.	
Service Migration	• Need view on current state of NFVI e.g. saturation, utilisation etc. • Need to capture and migrate state. • Rapid re-instantiation of the service (RFB chain)	• 100s of ms

5.7 Converged architecture

5.7.1 Summary of the cluster of requirements

Superfluidity identified convergence as one of its pillar objectives. The converged architecture aims at handling the heterogeneity in different levels. Specifically, three levels of heterogeneity have been identified, namely, (i) Heterogeneity in data traffic and end-points, (ii) Heterogeneity in services and processing needs, and (iii) Heterogeneity in access technologies and their scale. The following requirements should be addressed.

- Orchestration - Intelligent Orchestration to ensure that services are deployed and scaled on the most appropriate platform and consider the heterogeneity of resources in the decision making process. The platform needs to be able to synchronise with edge-nodes in a consistent manner.
- Dynamic placement and allocation of resources – to support the heterogeneity in end points, processing need and access technology it is required to select the network segment to place functions (e.g., cloud-RAN services) and monitoring elements. Furthermore, there is a need to automatically re-provision resources in response to a geographic encapsulated demand spike.
- Aggregation - Base stations for IOT devices will need to be deployed at sufficient density to ensure that at least one wireless connection type is available at all times to the IOT device. Aggregators will need to support a variety of front side radios such as Bluetooth, Bluetooth Smart, Ant, Zigbee, Wi-Fi etc. A wide of variety devices need to be supported, ranging from environmental sensors, body worn device, health devices, home sensors etc. IoT devices connects to the internet: using a specialized protocol (e.g., ZigBee , BLE, BACnet), or integrating the device into the IP-world (e.g. 6LoWPAN, CoAP, MQTT). In both cases the IoT devices connects to the internet using a kind of “gateway” that provides the entry to the network or that translate between protocols (e.g. 6LoWPAN to IPv6, CoAP to HTTP).
- Connectivity (high bandwidth, low latency) - The Mobile edge system shall be able to maintain connectivity between a UE and an application instance when the UE performs a handover to another small cell to another cell associated or not with the same Mobile edge

server. The network shall support the high bandwidth & low latency connectivity between the RRH and the RRM functions (located either at the edge cloud or the aggregation network).

5.7.2 Likely extent of impact of requirements on Superfluidity

Convergence is a key objective in Superfluidity, accordingly the above functional requirements have a substantial impact on the Superfluidity architecture.

5.7.3 Likely impact of requirements on Superfluidity Architecture

1. In-line functions
   - Support a variety of front side radios such as Bluetooth, Bluetooth Smart, Ant, Zigbee, Wi-Fi etc.
   - The Mobile edge system shall be able to maintain connectivity between a UE and an application instance
   - The access network shall support the high bandwidth & low latency connectivity between the RRH and the RRM functions

2. In-line monitoring
   - Monitoring elements should be placed across the edge, aggregation and core network

3. Orchestration and management
   - Ensure that services are deployed and scaled on the most appropriate platform and consider the heterogeneity of resources in the decision making process
   - Synchronise core network with edge-nodes in a consistent manner
   - Optimized selection of the network segment to place functions (e.g., cloud-RAN services)
   - automatically re-provision resources in response to a geographic encapsulated demand spike

5.8 Platform

5.8.1 Summary of the cluster of requirements

The Superfluidity project aims to open new business and technology perspective by designing a next generation of heterogeneous dynamic cloud architecture for future 5G cloud native telecom services.

To address these challenges, the platform cluster highlights a set of business requirements and KPIs that Superfluidity should fulfil to manage and orchestrate Network Function Virtualization (NFV) by providing capabilities such as extreme service agility as well as computation and network efficiency.

5.8.2 Likely extent of impact of requirements on Superfluidity

Some specific requirements such as scalability, geographic reach, network and computation acceleration, advances cloud management, NFV orchestration, low latency and monitoring should highly influence the work on Superfluidity.

Related platform should be adaptable and adopt an extensive approach. It should to be decentralized and composed of heterogeneous location dependent components to address low latency and high throughput challenges on-demand from any part of the topology.

5.8.3 Likely impact of requirements on Superfluidity Architecture

Building lightweight, composable and functional blocks and using specific hypervisors and accelerated components able to instantiate them in an efficient manner.

Leveraging the fact that the telco owns the last mile in order to provide low latency and caching services to the edge.

Compensating for compute and network limitations with programmable acceleration technologies such as FPGAs, GPUs, SRIOV and DPDK.

Implementing NFVs as a set Functional Blocks with microservice on light VMs or containers.

Providing NFVs on a managed platform for dynamic services composition and orchestration.

Enabling real-time decision making, service delivery and KPI validation with monitoring systems.

5.9 Scalability

5.9.1 Summary of the cluster of requirements

"[Scalability] connotes the ability of a system to accommodate an increasing number of elements or objects, to process growing volumes of work gracefully, and/or to be susceptible to enlargement." – (Bondi 2000).

We use the definitions of scale-out/in and scale-up/down that are defined by ETSI / NFV (ETSI NFV 2014), namely:

Scale-out/in: ability to scale by add/remove resource instances (e.g. VM)

Scale-up/down: ability to scale by changing allocated resource, e.g. increase/decrease memory, CPU capacity or storage size

The main perspectives for the requirements:

SWOP – {Service, Workload, Orchestration, Platform} – scalability perspectives.

A (business) service can be empowered by multiple workloads (e.g. a web-hosting service may rely on Apache, MySQL and PHP). To enable continuation of a service (SLA/QoE etc.) there will usually be a load balancer (platform scalability) that can redirect traffic to a different provider (e.g. a scaled-out workload). The way that workloads are combined together and the order that they are performed is part of the orchestration framework.

To scale-up a service to meet higher levels of demand, workloads are scaled-out, which may lead to a scale-out on the platform. A workload can be scaled-up temporarily but most of the time a workload is scaled-out to meet demand. The orchestration framework is orthogonal to the virtualised resource placement and needs to be able to cope with changes in demands and infrastructure, as well as changes in policies and SLAs.

Notes on scalability

Physical resources tend not to scale-up/down. The exception to this is CPU frequency bursting and bursting in general, where a cap is de-limited to allow for greater performance. RAM, network, disk storage etc., can usually not be changed other than in a step-wise increase when a new piece of hardware is added to the device.

Workloads though can scale-up and scale-out. Scale-up is where the workload uses more logical resources available up until the cap allowed by a Virtual Machine. In some cases the platform may allow for a VM to temporarily have access to more CPU, disk, RAM, IO, N/W resources for a defined period of time although this is not usually the case. Scale-out is where a new instance of the workload processor can be created that can then handle more requests. This can be used to get around the limitations of having a single share of resources on a shared tenancy system. If the resources of a single device are insufficient then the workload may be scaled-out onto another device.

- Individual server has workload scalability requirements; Scale-up and scale-out.
    - This leads to scalability requirements for the hypervisor
- Scalability of resources. All end-user devices will have increased throughput, hence greater requirements for bandwidth (and storage and compute) throughout the Superfluidity spectrum.
- Scalability of communications. More communication with a greater number of devices. M2M / D2D
    - Leads to a scalability requirement in the network to provide sufficient communication channels (latency and bandwidth) between nodes
- The most common way of addressing scalability throughout use-cases is to have a load-balancer and to then spawn up a 'new service handling instance' based on QoE / QoS / SLA requirements.
- For a business to maximise revenue, it is important that the orchestration system receives feedback about the dynamic state of the nodes so that more accurate decisions can be made
    - Leads to a requirement of scalable telemetry and feedback systems
- Orchestration scalability
    - Needs to be able to respond (promptly) to rapid increases/decreases in demand.
- Platform scalability - More and denser workloads are expected with faster provisioning times.

5.9.2 Likely extent of impact of requirements on Superfluidity

To support more services and workloads in a given area with existing infrastructure, Superfluidity will need to increase the density of workloads on the existing hardware. Given the large increase in services and workloads, increasing density alone will likely be insufficient and there will need to be more edge devices to support the number of users. The orchestration system and platform itself will need to be able to manage many more nodes connecting into the system and also be able to cope with the greater bandwidth of data and instructions (control) needed. The controllers (in SDN parlance) will need to be able to cope with many more devices and greater communication between devices for traffic, as well as different routes.

5.9.3 Likely impact of requirements on Superfluidity Architecture

Requirement	Impact to Architecture	KPIs
Load balancer	• How many and where load-balancers should be placed • Distributed or centralised • Determining the important properties of load balancers (inc. consistency)	Number of nodes a load balancer can handle Number of nodes the Superfluidity network can handle
Ability to scale-up/out at any point in the network topology	• All nodes in the Superfluidity network need to be able to run workloads • How to migrate/translate workloads between nodes of different architectures • Heterogeneous resource types that may be limited in running certain applications	Translating workloads to work on different architectures Migration (cold/live) between nodes Communication fabric between nodes and replication/migration of data
Monitoring / feedback system. (Dynamic load of a system)	• Orchestration can only act on the information it has • Conversely, the information needed should be only information that the orchestration can act on • Requires agents on each node and/or a way of polling/collecting data • Black/grey/whitebox analysis of workloads on a node	• Sensors and information that is sent to the orchestration platform
Scaling up of resources / virtualisation (single node) Creation and running multiple services on an edge-site in a scalable manner	• Determines the number of workloads that can be run • Overhead removes from the ability to run workloads • Determines performance and throughput	• Number of nodes that can be provisioned • CRUD + cool down/ warm up time of virtual workload executors • Performance

Scale up of a service at an edge-site (micro data centre)	• Links business to requirements to the IT platform • Determines the ingress to the problem-space	• Number of users/services that can be run in the Superfluidity network (when combined with physical infrastructure)
Manageability of the whole platform - Communication in the platform should scale to allow for all resources to be used by a single service	• Determines the number of messages that need to be sent between orchestrators • The heuristics and algorithms for resource placement should scale with the number of nodes and the workloads. The decisions will need to be taken under certain time constraints	• Frequency of calculation of resource placement will determine timeout for heuristics/algorithms
Infrastructure scalability (physical)	• Superfluidity network should be able to periodically update itself with current state of nodes and ensure information is 'relatively' fresh • New nodes should be able to join the network • As resources are changed on individual nodes this information should be relayed to the orchestration platform	• Number of nodes supported by a Superfluidity system

5.10 Security

5.10.1 Summary of the cluster of requirements

The Superfluidity architecture and platform shall ensure the following properties of information*:

- Authenticity: Only well-identified users* may generate, access, transmit, or act upon the information*. Actions, once performed, cannot be repudiated.
- Confidentiality: Only authorized users* may access the information*.
- Integrity: Only authorized users* can modify the information, and only allowed actions can be performed upon it.
- Availability: The information* shall be accessible for use when required.
- Traceability: Actions upon the information* may be traced back to the users* that performed them.
- Isolation: Failure, malfunction, misbehaviour or overload of a resource* shall not affect other resources*. Network slicing shall be supported.

* NOTES:

- ‘information’ includes the information itself, and the services and systems that manage it.
- ‘users’ include human users, services and systems.
- ‘resources’ include building blocks, functions, services, devices, systems and network slices.

5.10.2 Likely extent of impact of requirements on Superfluidity

Security requirements are expected to have a significant impact on the technical work of the project. In fact, Security by Design* is one of the key aspects of Superfluidity's approach and architecture.

* Security by Design: Reusable Functional Blocks (RFBs) shall be verified to be secure before deploying them.

5.10.3 Likely impact of requirements on Superfluidity Architecture

Requirement	Impact to Architecture	KPI’s
Security (general)	• Security Subsystem within the Superfluidity platform • The Security Subsystem shall include some kind of Intrusion Detection and Protection System (**IDPS**) • RFBs, Services, and APIs, shall be classified as either critical or non-critical • The following requirements apply at least to critical RFBs, Services and APIs	For IDPS: • Detection Time (something abnormal is happening) • Identification Time (type of attack, misbehaviour, error, etc.) • Reaction Time (mitigation/solution actions start to be applied) • Restoration Time (normal behaviour/operation re-established) • Tracking Time (root cause identification) • Effectiveness: o Percentage of attacks detected o Percentage of attacks blocked o Percentage of false positives
Authenticity	• RFBs shall be authenticated • Users shall be authenticated • Platform services and APIs shall support authentication	• See KPIs for IDPS

	• Authentication Service/API in the SF platform	
Confidentiality	• Information transmitted from, to, or between RFBs and/or platform services/APIs shall be encrypted or otherwise protected • Platform services and APIs shall support encryption • Encryption Service/API in the SF platform	• See KPIs for IDPS
Integrity	• Information transmitted from, to, or between RFBs and/or platform services/APIs shall be digitally signed or checksum/CRC protected • Platform services and APIs shall support digital signature or checksum/CRC protection • Signature/Checksum/CRC Services/APIs in the SF platform	• See KPIs for IDPS
Availability	• High Availability (HA) shall be provided for critical RFBs, Services and APIs (e.g. by replication, hot standby, cold standby, etc.) • HA Services/APIs in the SF platform	• Percent of unsuccessful access attempts • Percent of interrupted accesses/connections
Traceability	• Accesses to RFBs, Services and APIs shall be logged, and log information securely stored • Platform services and APIs shall support access logging and tracing • Logging/Tracing Services/APIs in the SF platform	• Percent of events that cannot be traced
Isolation	• RFBs shall be executed in a sandboxed environment • The SF platform shall support resource and network slicing	• Percent of anomalies/ failures that propagate
Security by Design	• Static analyser for verification of RFBs prior to deployment • Dynamic analyser for behaviour verification, anomaly detection, etc.	• Percent of unsafe RFBs detected • Percent of anomalies detected before fault • Percent of false positives

6 Conclusions

In this document we have described a representative collection of Superfluidity use cases and scenarios and their induced requirements on the Superfluidity architecture. Specifically, we have presented a comprehensive set composed from an inclusive set of use cases defined by Superfluidity Partners (in this document we presented only a representative consolidated set of 23 use cases), as well as a representative set of use cases which are collected from the "external world", i.e., ones which are defined outside the Superfluidity project, such as those from the ETSI ISG on NFV, ETSI ISG on MEC, 3GPP's document on "Feasibility study on new services and markets technology enablers", vendor reports, etc.

The heterogeneous set of use cases introduces a substantial list of requirements. We have identified key features and clustered the requirements according to these features. Specifically, we have identified 12 main clusters of system requirements, which can be classified as business requirements (service agility, cost savings, user experience, security and analytics) and technical requirements (orchestration, quality of experience, scalability, platform, converged architecture, building blocks and security).

Finally, we considered the impact of these "clusters" of requirements on the Superfluidity architecture and on the broader technical work within the project. We did not differentiate between business level and technical level requirements. So we merged the two "security" clusters, as well as the "user experience" and "quality of experience" into a single cluster. Our final list includes ten requirements clusters: service agility, cost savings, analytics and metrics, quality of experience, building blocks, orchestration, converged architecture, platform, scalability, and security. Within each cluster, we identified a set of requirements analysing their impact on the architecture and identifying associated KPIs where possible.

This document provides a baseline for Deliverable D2.2, which will provide a detailed description of the consolidated Functional-Analysis of the Superfluidity system, identifying the monolithic functions and their decomposition into reusable components.

7 Appendix A - Use cases from the state-of-the-art

7.1 ETSI NFV ISG

The activity of ETSI NFV ISG has been divided in a number of phases. The document “Network Functions Virtualisation (NFV); Use Cases” (ETSI) [ref-1] has provided the following list of use cases for phase 1.

- Network Functions Virtualisation Infrastructure as a Service
- Virtual Network Function as a Service (VNFaaS)
- Virtual Network Platform as a Service (VNPaaS)
- VNF Forwarding Graphs
- Virtualisation of Mobile Core Network and IMS
- Virtualisation of Mobile base station
- Virtualisation of the Home Environment
- Virtualisation of CDNs (vCDN)
- Fixed Access Network Functions Virtualisation

Figure 12: ETSI NFV ISG phase 1 use cases

Phase 2 is still ongoing and in general the documents are not finalised nor publicly available. The document “Network Functions Virtualisation (NFV); Acceleration Technologies; Report on Acceleration Technologies & Use Cases;” [ref-2] is a phase 2 document published as draft. It reports the following set of use cases related to acceleration:

Compute Acceleration

- IPSec tunnels termination VNFC
- Next Generation Fire Wall (NGFW) Acceleration
- Virtual Base Station (VBS) L1 Acceleration
- Virtual Acceleration Interface for VNFs
- Transcoding
- Deep Packet Inspection

Network Acceleration

- Load Balancing and NAT
- NFVI Virtual Networking Offload
- NFVI Secure Overlay Offload
- Dynamic Optimization of Packet Flow Routing

STORAGE ACCELERATION

- NVMe Over Fabric Enabled Acceleration
- High Performance Persistent Memory on Compute Node

Figure 13: ETSI NFV ISG phase 2 doc with use cases related to Acceleration Technologies

7.2 ETSI MEC ISG

The ETSI MEC (Mobile Edge Cloud) ISG is working on a detailed document ("Mobile-Edge Computing (MEC); Technical Requirements") with 23 use cases, categorised according to three sets: "Consumer-oriented services", "Operator and third party services" and "Network performance and QoE improvements". This document, as its name indicates, extracts requirements and identifies features ("A feature is defined as a group of related requirements"). The document is not yet publicly available; hereafter we have collected information coming from public documents and presentations.

A public technical white paper (Mobile-Edge Computing – Introductory Technical White Paper [ref-3]) contains the set of use cases listed below:

- Active Device Location Tracking
- Augmented Reality Content Delivery
- Video Analytics
- RAN-aware Content Optimization
- Distributed Content and DNS Caching
- Application-aware Performance Optimization

Figure 14: Use cases from the Introductory Technical White paper of ETSI MEC ISG

A presentation about MEC goals [ref-4] includes the use cases listed below:

Network-Centric Applications

- Distributed Content and DNS Caching
- RAN-aware & Application-aware Content Optimization

Enterprise and Vertical Applications

- Active Device Location Tracking
- Intelligent Video Analytics

Efficient Delivery of Local Content

- Augmented Reality Content Delivery

Figure 15: Use cases from a presentation about MEC goals

The ETSI White Paper No. 11 [ref-5] considers the following use cases:

- Augmented Reality
- Intelligent Video Acceleration
- Connected Cars
- Internet of Things Gateway

Figure 16: Use cases from the ETSI White Paper No. 11

7.3 3GPP

3GPPP's Technical Specification Group "Services and System Aspects" has produced the document "Feasibility Study on New Services and Markets Technology Enablers;" [ref-6]. It has similarities to MEC's approach but more broad in scope since it "aims to identify the market segments and verticals whose needs 3GPP should focus on meeting, and to identify groups of related use cases and requirements that the 3GPP eco-system would need to support in the future". This is a very broad and wide-ranging endeavour, not specifically referred to NFV. The complete list of 59 use cases are outlined below:

1	Ultra-reliable communications	31	Temporary Service for Users of Other Operators in Emergency Case
2	Network Slicing	32	Improvement of network capabilities for vehicular case
3	Lifeline communications / natural disaster	33	Connected vehicles
4	Migration of Services from earlier generations	34	Mobility on demand
5	Mobile broadband for indoor scenario	35	Context Awareness to support network elasticity
6	Mobile broadband for hotspots scenario	36	In-network caching
7	On-demand Networking	37	Routing path optimization when server changes
8	Flexible application traffic routing	38	ICN Based Content Retrieval

9	Flexibility and scalability	39	Wireless Briefcase
10	Mobile broadband services with seamless wide-area coverage	40	Devices with variable data
11	Virtual presence	41	Domestic Home Monitoring
12	Connectivity for drones	42	Low mobility devices
13	Industrial Control	43	Materials and inventory management and location tracking
14	Tactile Internet	44	Cloud Robotics
15	Localised real-time control	45	Industrial Factory Automation
16	Coexistence with legacy systems	46	Industrial Process Automation
17	Extreme real-time communications and the tactile internet	47	SMARTER Service Continuity
18	Remote Control	48	Provision of essential services for very low-ARPU areas
19	Light weight device configuration	49	Network capability exposure
20	Wide area sensor monitoring and event driven alarms	50	Low-delay speech coding
21	IoT Device Initialization	51	Network enhancements to support scalability and automation
22	Subscription security credentials update	52	Wireless Self-Backhauling
23.	Access from less trusted networks	53	Vehicular Internet & Infotainment
24	Bio-connectivity	54	Local UAV Collaboration
25	Wearable Device Communication	55	High Accuracy Enhanced Positioning (ePositioning)
26	Best Connection per Traffic Type	56	Broadcasting Support
27	Multi Access network integration	57	Ad-Hoc Broadcasting
28	Multiple RAT connectivity and RAT selection	58	Use Case for Green Radio
29	Higher User Mobility	59	Massive Internet of Things M2M and device identification
30	Connectivity Everywhere		

Figure 17: Use cases list from the 3GPP's document on New Services and Markets Technology Enablers

7.4 IETF SFC Work Group

The Internet Engineering Task Force (IETF) has a group that is defining Service Function Chaining (SFC) [ref-7]. Of specific relevance to Superfluidity is the work on use cases and more concretely the draft on Service Function Chaining Use Cases in Mobile Networks [ref-8], from which we have extracted this list:

- Service chain model for Internet HTTP services
- Service chain for TCP optimization
- HTTP header enrichment in mobile networks

Figure 18: Use cases list from IETF SFC's draft on Mobile Networks

7.5 Use cases from non-SDOs

7.5.1 Use cases from SDx Central

SDx Central [ref-9] considers the following list of use cases [ref-10]:

Network Access Control - Set appropriate privileges for users or devices accessing the networks, including access control limits, incorporation of service chains as well as appropriate quality of service. Generally follows the user/device as they connect from different parts of the network.

1. Campus NAC
2. Remote Office / Branch NAC
3. M2M NAC
4. Unified Communications Optimization

Network Virtualisation - Creates an abstracted virtual network on top of a physical network, allowing a large number of multi-tenant networks to run over a physical network, spanning multiple racks in the datacentre or locations if necessary, including fine-grained controls and isolation as well as insertion of acceleration or security services

1. Data Center Virtual Networks
2. Campus / Branch Virtual Networks
3. Data Center Micro Segmentation
4. Network Functions as a Service

Virtual Customer Edge - Virtualising the customer edge either through creation of a virtualised platform on customer premises or by pulling in the functions closer to the core on a virtualised multi-tenant platform hosted either in a carrier point-of-presence, regional datacentre, central datacentre (enterprise, telco or over-the-top cloud SP)

1. On-premises vCPE
2. On-premises vCPE (OTT)
3. vCE (Telco)
4. vCE (OTT) (aka: Cloud CPE)

Dynamic Interconnects - Creation of dynamic links between locations, including between DCs, enterprise and DCs, and other enterprise locations, as well as dynamically applying appropriate QoS and BW allocation to those links.

1. BWoD
2. Virtual Private Interconnects / Cloud Bursting
3. Dynamic Enterprise VPN
4. Cross Domain Interconnect
5. Multi-Layer Optimization

Virtual Core and Aggregation - Virtualised core systems for service providers including support infrastructure such as vIMS, vEPC, as well as dynamic mobile backhaul, virtual PE and NFV GiLAN infrastructure

1. vEPC & vIMS
2. vPE
3. GiLAN
4. Mobile Network Virtualisation

Datacentre Optimization - Using SDN and NFV, optimizing networks to improve application performance by detecting and taking into account affinities, orchestrating workloads with networking configuration (mice/elephant flows)

1. Big Data Optimization
2. Mice/Elephant Flow Optimization

Figure 19: Use cases from SDX central

7.5.2 Use cases from a vendor's perspective

[Ref-11] provides a vendor's view of use cases:

1. Mobility Virtualisation

   Virtualisation of the packet core and GiLAN

2. Virtual CPE and Service Chaining

   Deliver CPE services from the cloud. SDN to automate the processes for creating new services

3. NFV and Service Orchestration

   Operating in a hybrid, physical/virtual environment. orchestration sits at the centre of this

4. WAN Optimisation & Innovation

   Better traffic engineering. holistic view across the different transport and IP network layers The innovation of SDN, in association with NFV, gives us the possibility to create simpler network designs, making the "delayering" of networks achievable

5. Policy Driven Application Provisioning & Delivery

Delivery of applications from a data centre. Turning the manual process of provisioning and delivering applications into a simplified, automated, policy-driven process

Figure 20: Use cases list from a CISCO web page

7.6 Use cases from Superfluidity technical annex

For completeness we describe the use case 'samples' that we had in mind for the technical annex.

Minimum-Delay Cloud storage (TA-1) Cloud storage has the potential for finally allowing people to throw out all of the clumsy hard drives, memory cards and USB sticks cluttering homes and travel bags. Unfortunately, the high data volumes and relatively low throughputs and high delays to the core/data centre mean that there is still a difference in the experience between local and cloud storage; deploying cloud storage services at the network edges would finally close this gap.
RAN As A Service (TA-2) Individual functions constituting a Cloud RAN would be readily deployed, following a dynamic life cycle (creation, attachment to core network and antenna site / RRH, hot upgrade, etc.) involving optimized placement decisions about CPU, NIC, memory, hardware-acceleration capabilities; Moreover, a RAN could be flexibly build and adapted to the context using different types of schedulers, different physical layer blocks pointing to various waveforms, etc.
Localised services (TA-3) Many services that require some sort of mixing server (e.g., video conferencing, online gaming, to name a few) often end up using a distant one in terms of delay. Instead, such a virtualised server could be deployed on-the-fly at the edge, drastically reducing delays and improving user experience.
Pooling (TA-4) User-specific functions attached to various cells (and even baseband computation units [Wer13]) can be pooled in a same host so as to maximize the host load / minimise the required number of hosts; intra-cluster live migration of functions would optimize system KPIs (pooling gain, total radio capacity, energy efficiency, etc.), and would comply with RRM handover requirements (e.g. the current intra LTE handover < 50ms would be readily attained by our technology);
Edge offloading (TA-5) One of the drawbacks of mobile devices is their short battery life. Many services (e.g., firewalling, anti-virus software, ad blockers) could be offloaded to the edge to reduce battery consumption.
Portable signal processing (TA-6) platform independence would permit portability of signal processing tasks between the edge cluster and the antenna site so as to minimise fronthauling requirements and maximize radio

capacity (as fronthauling requirements increase for larger radio bandwidth using carrier aggregation, more massive MIMO, more network MIMO…).
On-the-fly Monitoring (TA-7) On-the-fly Monitoring: The owner of the infrastructure could deploy a monitoring service in order to track usage of its tenants' services, or to for instance instantiate a DPI service on a particular suspicious flow.
Virtualised CDN operators (TA-8) Virtualised CDN operators: It is well known that the performance of CDNs improves the closer that content is from users. This, however, is an expensive proposition, and so restricts all but the biggest players from the market. Instead, newcomers could deploy (virtualised) content caches at network edges, effectively renting out infrastructure and growing it as their business grows.

7.7 References for Appendix A

[1] "Network Functions Virtualisation (NFV); Use Cases" (ETSI) http://www.etsi.org/deliver/etsi_gs/nfv/001_099/001/01.01.01_60/gs_nfv001v010101p.pdfReference 1

[2] "Network Functions Virtualisation (NFV); Acceleration Technologies; Report on Acceleration Technologies & Use Cases;" ETSI GS IFA 001 V0.7.0 (2015-09) https://docbox.etsi.org/ISG/NFV/Open/Drafts/IFA001_Acceleration_-_UCs_report/NFV-IFA001v070.zip

[3] "Mobile-Edge Computing – Introductory Technical White Paper", https://portal.etsi.org/Portals/0/TBpages/MEC/Docs/Mobile-edge_Computing_-_Introductory_Technical_White_Paper_V1%2018-09-14.pdf

[4] "Boosting User Experience by Innovating at the Mobile Edge - An Introduction to the new ETSI Industry Specification Group (ISG) for Mobile Edge Computing (MEC)", presentation by Dr. Rolf Schuster, Vodafone http://www.etsi.org/images/files/technologies/MEC_Introduction_slides__SDN_World_Congress_15-10-14.pdf

[5] Yun Chao Hu, Milan Patel, Dario Sabella, Nurit Sprecher and Valerie Young , "Mobile Edge Computing - a key technology towards 5G", ETSI White Paper No. 11 http://www.etsi.org/images/files/ETSIWhitePapers/etsi_wp11_mec_a_key_technology_towards_5g.pdf

[6] "Feasibility Study on New Services and Markets Technology Enablers;", 3GPPP's Technical Specification Group "Services and System Aspects", 3GPP TR 22.891 V1.0.0 (2015-09)

[7] IETF Service Function Chaining WG charter - https://tools.ietf.org/wg/sfc/charters

[8] Haeffner et.al. "Service Function Chaining Use Cases in Mobile Networks" https://tools.ietf.org/html/draft-ietf-sfc-use-case-mobility-04

[9] sdx central - https://www.sdxcentral.com/

[10] “Network Functions Virtualisation and Software Defined Networking; SDN & NFV Use Cases Defined”, SDx central - https://www.sdxcentral.com/sdn-nfv-use-cases/

[11] “Top 5 Service Provider Use Cases for SDN and NFV in EMEAR”, CISCO http://emear.thecisconetwork.com/site/content/lang/en/id/1433

8 Appendix B – Requirements from the state-of-the-art

8.1 Introduction

This section reviews the challenges identified by the most relevant standardization bodies and the 5G-PPP initiative, and how they have been translated to technical requirements. Our study of the state-of-the-art was a crucial pillar when we derived the requirements for our use cases.

8.1.1 The 5G landscape, challenges and opportunities

As identified in [ref-1], the main challenges for 5G are to create "interoperable, ubiquitous and dynamic" 5G communication systems and applications. These emerge from the need to meet user expectations with regard to mobility and quality of service, while coping with an ever-growing data volume. Wireless communication services will expand into new market segments, such as smart grids and smart cities, e-health, Intelligent Transportation Systems, traffic control and safety, just to name a few. This expansion is required to realise the vision for an intelligent, mobile and connected society based on social, political and technological drivers [ref-2]. The key aspects of this vision are highlighted below:

- Final consumer – Mobile devices are ubiquitous in modern everyday life for a large and exponentially increasing set of users, accommodating new applications requiring very high data rates, low latency, security and reliability. People want to experience always-on mobile connectivity, with high quality, quick response and high reliability communication to interact seamlessly with the world around them. It is estimated that the global mobile data traffic will grow more than 200 fold from 2010 to 2020, and by nearly 20,000 times from 2010 to 2030.
- Machine to machine – Future wireless systems need to be designed with support for real-time machine-to-machine (M2M) communication in mind from the bottom-up. Some examples are driverless cars, real-time traffic control optimisation, emergency and disaster response, or efficient industrial communications, among others.
- High user density – Today, we are experiencing an ever-increasing mobile devices density per unit of area. Examples include large buildings with many employees, shopping centres, music festivals, sports events, etc.
- Enhanced multimedia services – User devices will have Ultra-High Definition (UHD) displays, including multi-view, mobile 3D projections, immersive video conferencing, augmented reality and mixed reality display, and interfaces enabling the provisioning of e-health, education and entertainment (as for instance gaming) applications. This will require considerably higher data rates.
- Internet of Things (IoT) – Every object can benefit from being connected through Internet technologies. These connected "things" can be smart phones, sensors, actuators, cameras, vehicles, etc., ranging from simple devices to highly complex and advanced devices. These connected devices are bound to have varying levels of energy consumption, transmission power, latency requirements, cost, etc. Smart cities, smart rural places will benefit from this level of ubiquitous connectivity. Examples include smart energy distribution grids, agriculture, healthcare, or vehicle-to-vehicle and vehicle-to-road infrastructure communications.

These drivers and other trends anticipate that global mobile traffic will grow and the technological systems will need to continuously evolve at an even faster rate to meet future demands.

Data forecasts

The following picture, taken from Cisco forecasts [ref-3], provides predictions for several global population evolution indicators.

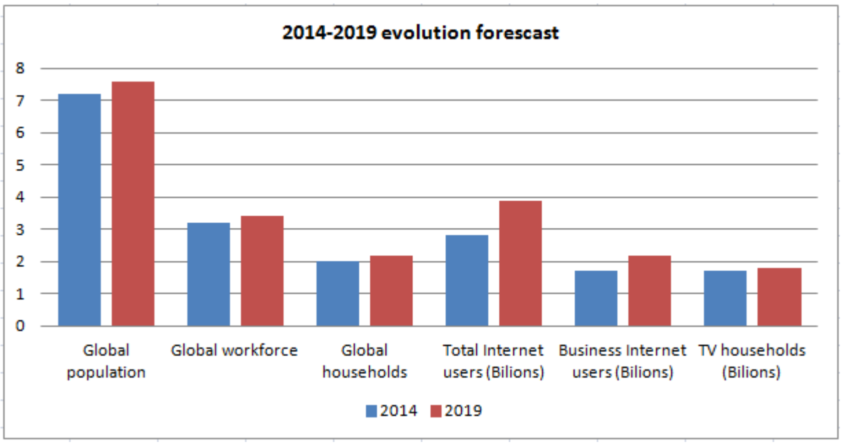


Figure 21: 2014 to 2019 evolution forecast

It can be seen from the figure that the total number of Internet users will grow at a faster rate than the global population, while TV households will maintain relatively static. According to Cisco, mobile data traffic will:

- Grow 10-fold from 2014 to 2019, at a compound annual growth rate of 57%
- Reach 24.3 Exabytes per month by 2019, up from 2.5 Exabytes per month in 2014
- Reach an annual run rate of 291.8 Exabytes by 2019, up from 30.3 Exabytes in 2014
- Grow 3 times faster than global fixed IP traffic from 2014 to 2019
- Account for 15% of global fixed and mobile data traffic by 2019, up from 4% in 2014
- Be equivalent, in 2019, to 266x the volume of global mobile traffic 10 years earlier (in 2009)
- Be equivalent, in 2019, to 12x the volume of the entire Global Internet in 2005
- Globally 47% of all networked devices will be mobile-connected in 2019

The next two figures provide additional insights on the evolution of mobile traffic and of smartphone devices, and about key indicators on the projected growth of connected devices between 2014 and 2019.

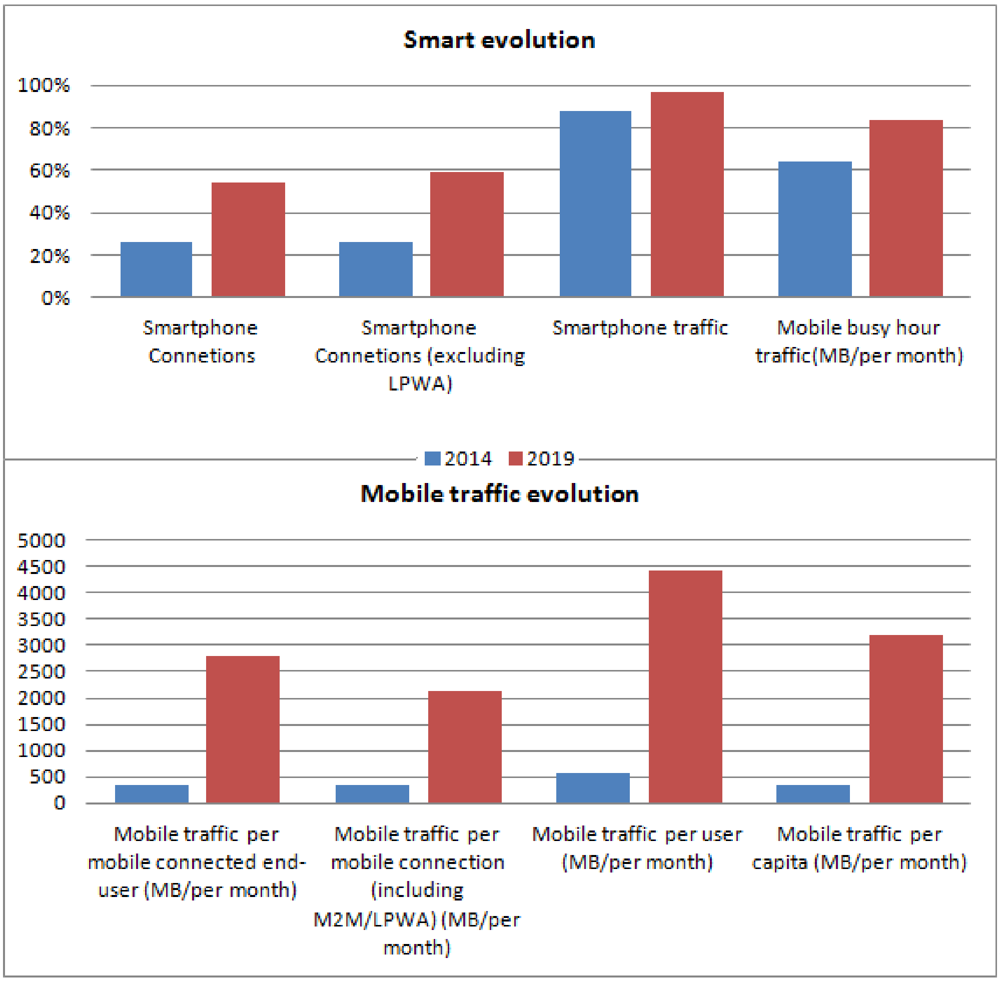


Figure 22: Mobile traffic and smartphone evolution

Devices evolution	2014	2019
Networked devices (billion)	14,2	24,2
Networked devices per capita	2	3,2
M2M modules	24% (3,3 billions)	43% (10,5 billions)
PCs	11% (1,5 billions)	6% (1,5 billions)
Tablets	4% (922,4 millions)	3% (419,2 millions)
Smartphones	19% (4,6 billions)	15% (2,1 billions)
Connected TVs	11% (1,5 billions)	12% (3,1 billions)
Non-smartphones	12,7% (3,1 billions)	32% (4,5 billions)
Other portables	4% (875,9 millions)	5% (693,1 millions)

Global IP Traffic from non-PC devices	39%	67%
IP Traffic from PCs	61%	33%
IP Traffic from TVs	29%	22%
IP Traffic from Portable devices (as smartphones and tablets)	10%	42%
IP Traffic from M2M	0,50%	2,70%
Consumer Internet Traffic from PCs	75%	33%
Consumer Internet Traffic from TVs	9%	10%
Total Internet Traffic from TVs	7%	8%

Figure 23: Projected growth of connected devices 2014 to 2019

According to [ref-4], the expected timeline for 5G is based on the assumption that the standards in 3GPP Rel-14/15/16, the Spectrum Allocation in WRC-19 and ITU Approval will occur by 2020, as planned.

8.1.2 Financial Trends

While there are some key data enabling some operational projections, there are no detailed reports with potential revenue predictions currently available with the exception of Juniper Research [ref-5]. This report forecasts the adoption for 5G, with service revenues set to exceed $65 billion by 2025, compared to just $100 million during its first year of commercial services in 2020. This figure would represent approximately 7% of all operator-billed service revenues.

8.1.3 5G Standardisation Activities

Currently, the ITU-T has established a focus group that is identifying the standardisation requirements for 5G networks, known as the International Mobile Telecommunications for 2020 and beyond (IMT-2020) study group. This group is focusing on the requirements of wired elements and interfaces in 5G networks. In parallel, the ITU-R has established an equivalent study group (that even shares the name with the ITU-T study group) to study the challenges and requirements for 5G system in the radio interface. This study group has produced a first report [ref-2], which provides an analysis of trends and envisaged usage scenarios for IMT-2020.

The ETSI NFV ISG is in the midst of a change from stage 2, which focuses on the information model, to stage 3, that includes data models and interoperability tests. Being late, the ETSI NFV is looking for accelerators that would allow it to cope with open-source projects that have already implemented parts of its architecture, as well as with other standardisation bodies that have defined the data models for some of the ETSI NFV APIs. Specifically, in a recent poll among the operators, the majority voted in favour of adopting OpenStack APIs as the de-facto standard for the interfaces between VIM-VNFM and VIM-NFVO, which corresponds to the ETSI NFV IFA005 and IFA006, respectively.

There is still a need to adapt the OpenStack to comply with the requirements of ETSI NFV stage 2, or possibly, update the ETSI NFV in accordance to OpenStack's implementations. This gap analysis and the way to bridge is currently a discussion point for the ETSI NFV.

For orchestration, the TMForum ZOOM (Zero-time Orchestration, Operations and Management) project aims to define a vision of new virtualised operations with an architecture based on the seamless interaction between physical and virtual solutions. To that end, the TMForum published a framework technical report on information framework enhancements to support ZOOM [ref-10]. This report defines 4 fundamental concepts for modelling NFV-based systems, namely, Virtual Resource, Network Function, Network Service, and Graph, as well as 2 general-purpose concepts, Catalogue and Event, which are also critical for realising SDN and NFV systems. Having already defined a set of APIs to OSSs, the TMForum is the leading candidate to be delegated by the ETSI NFV to define the data model for IFA012 and IFA013 (OSS-NFVO interfaces).

The 3GPP SA WG3 aims at specifying the requirements, architecture and solutions for provisioning and management of the network (RAN, CN, IMS) and its services. As such, it has requirements related to several ETSI NFV interfaces and architectural specifications, including network service life cycle management, VNF package management, virtual resource fault information, performance management information, and specifications of the Ve-Vnfm-em, Ve-Vnfm-vnf and Os-Ma-nfvo Reference Points with respect to 3GPP applications. The ETSI NFV IFA group cooperates in these areas with the 3GPP SA5 to align the specifications.

8.1.4 5G-PPP Initiative

The “Public Private Partnership on 5G” (5G-PPP) is the EU flagship initiative to accelerate research developments in 5G technology. Formally, it is an agreement between the EU commission and the “5G Infrastructure Partnership”, an industry association comprising the key European players in ICT. The European Commission will provide a public funding of €700 million to support this activity, in the context of the Horizon 2020 Programme. [Ref-6] is the “vision” document produced by the 5G-Infrastructure PPP.

8.2 General requirements from standards bodies

8.2.1 ITU-R

ITU-R has identified a set of ‘usage scenarios’ for ‘IMT for 2020 and beyond’ [ref-2]. These were organised along three use cases’ categories, as shown in the figure, reflecting the supported requirements along those three axis:

- Enhanced Mobile Broadband (eMBB)
- Ultra-reliable and low latency communications, also known as critical MTC (cMTC)
- Massive machine type communications (mMTC)

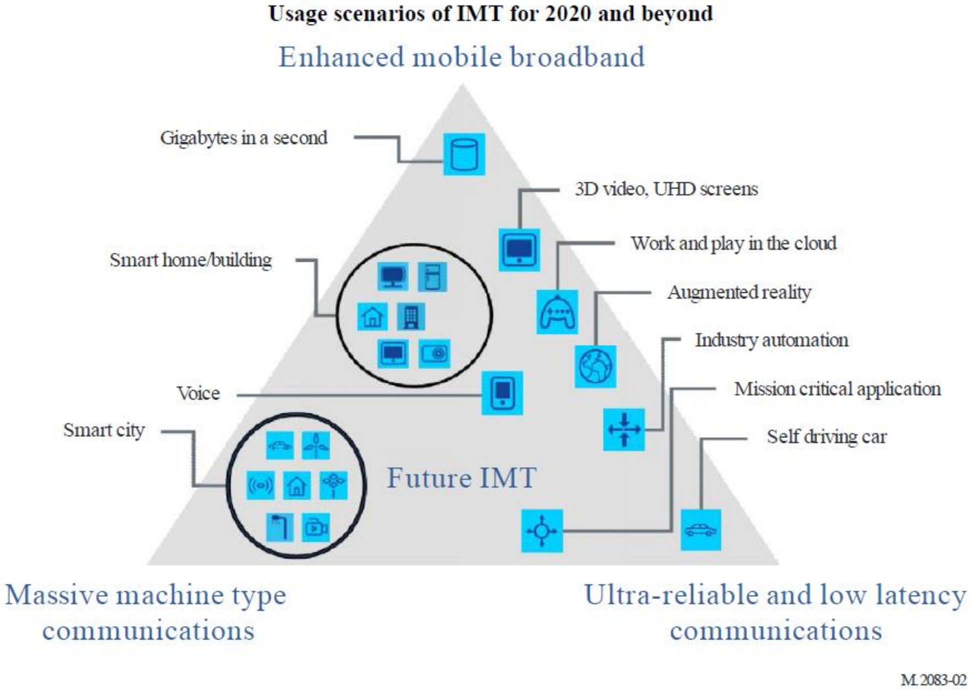


Figure 24: IMT 2020 and beyond usage scenarios

The target performance enhancements required to address those scenarios, are summarised by ITU-R into eight 'key capabilities':

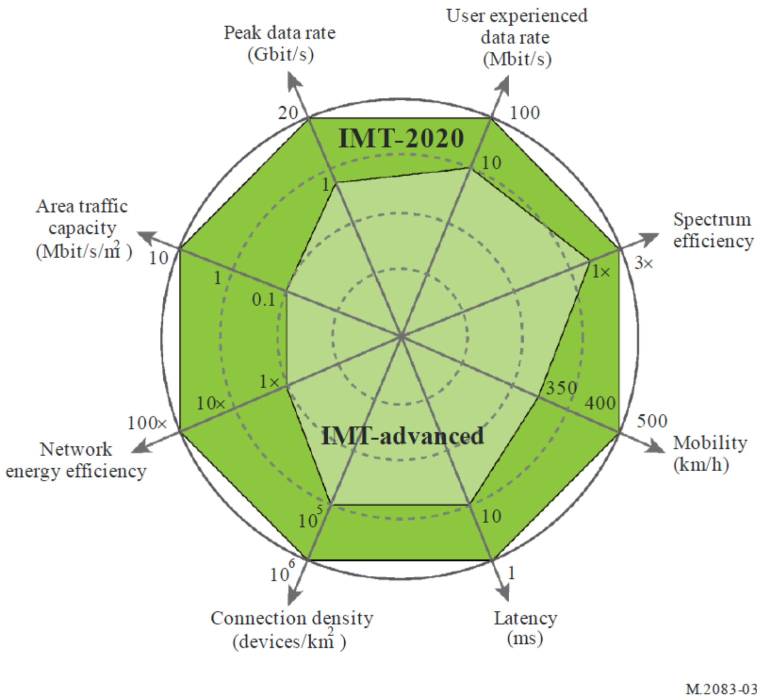


Figure 25: Enhancement of key capabilities from IMT-Advanced to IMT-2020

These eight capabilities can be grouped into the following five categories:

1. Capacity (global area traffic and connections)
2. Throughput (peak and experienced data rates)

3. Latency
4. Mobility
5. Efficiency (Spectrum and energy)

The established target values for these KPIs face significant challenges, to be solved over the coming years, with some requiring enhancements with factors of up to 100x.

Other organisations and vendors present slightly different numbers and add other KPIs, such as:

- Reduction to 20% in network management OPEX (5G-PPP)
- Services' deployment time below 90 minutes (5G-PPP)
- Cell spectral efficiency of 30 bps/Hz (Samsung)
- Cell edge data rate of 1 Gbps (Samsung)

The Figure below, also from ITU-R, identifies the relative importance ('high/medium/low') of the key capabilities relevant to the usage scenarios.

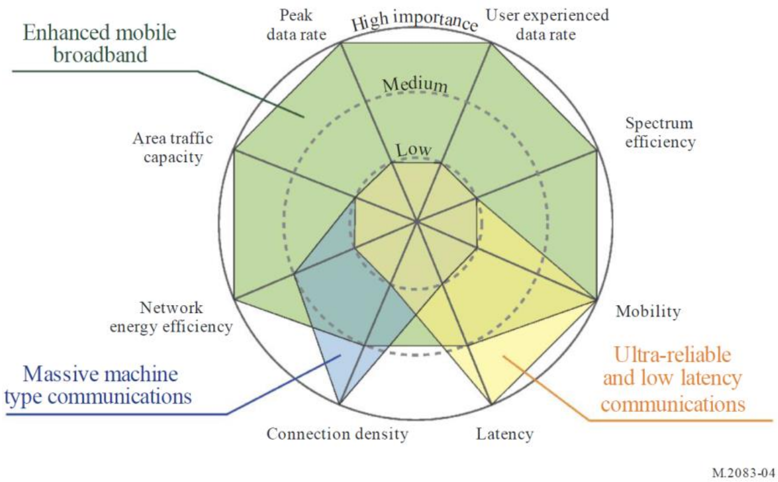


Figure 26: Importance of key capabilities in different usage scenarios

It is clear that the *enhanced mobile broadband* (eMBB) usage scenario represents the widest range of requirements, even if the focus is only on capacity, efficiency and mobility. *Massive machine type communications* (mMTC) and *ultra-reliable and low latency communications* have very specific requirements, the former focusing in efficiency, to cope with finite battery duration, and connection density (large number of devices per area unit), and the later focusing on mobility and latency.

8.2.2 3GPP

3GPP, via its SMARTER (*Study on New Services and Markets Technology Enablers*) Study Item [ref-11] is elaborating the TR 22.891, currently encompassing seventy-two 'use cases', extracting functional requirements to be realised in future 3GPP Releases' features. This work in progress, is complemented by results coming from *NexGen* (SA2 Architecture) and RAN (Radio Access Network) activities. From these and among many other requirements identified, we can extract the following as being directly relevant for Superfluidity:

- Ultra-reliable communications
- High availability
- Very low data error rate
- Network slicing
- On-demand networking
- Mobility between 3GPP and non-3GPP RATs
- Multiple simultaneous connections to several RATs
- Content caching capabilities at the edge of the network
- Services mobility to follow the user, with routing optimization
- Network capabilities exposure
- Broadcast/multicast services in a dedicated bearer

8.2.3 ETSI MEC WG

Mobile Edge Computing (MEC) is recognised by the 5G-PPP as one of the key emerging technologies for 5G systems (along with NFV and SDN). It will offer application developers and content providers cloud computing and an IT service environment at the edge of the mobile network (eventually on fixed environments in the future). The MEC environment will be characterised by:

- Proximity
- Ultra-Low Latency
- High Bandwidth
- Real-time access to network information
- Location Awareness

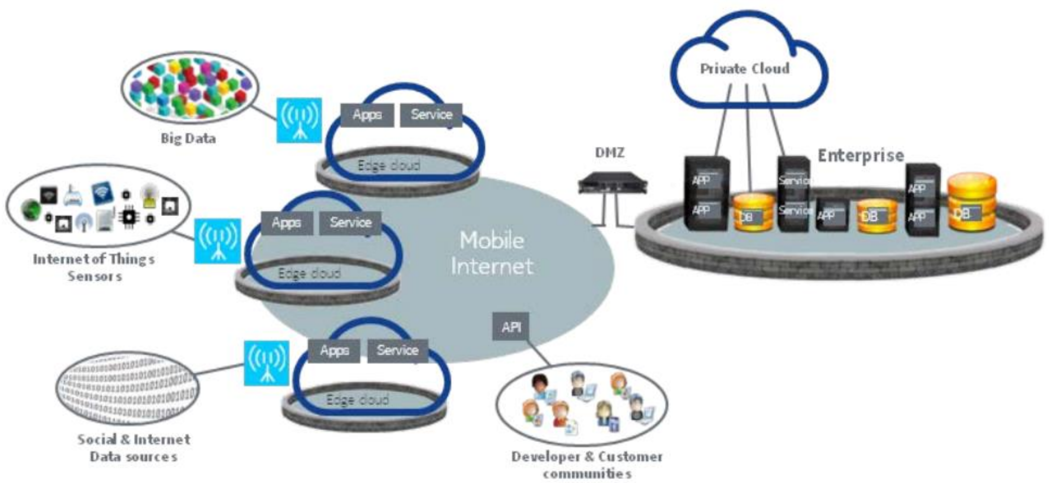


Figure 27: MEC provides IT and Telecoms Networking Convergence

The ETSI Industry Specification Group (ISG) on Mobile Edge Computing (MEC) produces normative Group Specifications that will enable the hosting of third-party applications in a multi-vendor MEC environment. Launched in December 2014, the group plans to deliver the first set of specifications within 2 years.

The initial scope of the ISG MEC focuses on use cases, requirements and the reference architecture, including the components and functional elements that are the key enablers for MEC solutions.

When the first documents reach the required maturity level, work on platform services, APIs and interfaces will commence. The ETSI whitepaper on MEC [ref-8] provides an initial overview of the MEC platform as shown in the figure above. The MEC platform is described as a hosting infrastructure, comprising hardware resources and a virtualization layer. The details of the actual implementation of the MEC hosting infrastructure (including the actual hardware components) are abstracted from the applications hosted on the platform. The MEC platform API [ref-9] will be application-agnostic and will foster the smooth portability of value-creating applications on every mobile-edge server with guaranteed SLA:

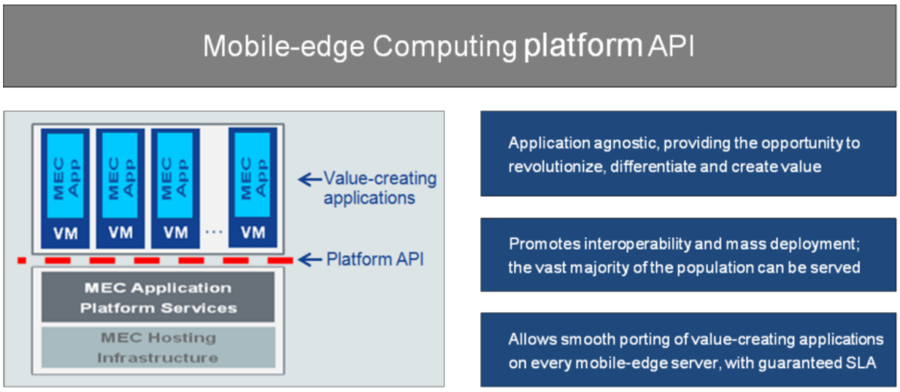


Figure 28: MEC Platform API

According to the work-in-progress Group Specification "ETSI GS MEC 002: Mobile-Edge Computing (MEC); Technical Requirements", MEC technical requirements are organised/outlined as follows:

I. Generic principles	III. Services	
1 NFV alignment 2 Mobility support 3 Deployment independence 4 Simple and controllable APIs 5 Smart application location 6 Representation of features	Platform service requirements 1 Services registry 2 Connectivity 3 Storage 4 Traffic routing 5 DNS support 6 Timing	Features 1 Feature *UserApps* 2 Feature *SmartRelocation* 3 Feature *RadioNetworkInformation* 4 Feature *LocationService* 5 Feature *BandwidthManager* 6 Feature *UEIdentity*
II. Generic requirements	IV. Operation and management requirements	V. Security, regulation, charging requirements
1 Requirements on the framework		

2	Application lifecycle
3	Applications environment
4	Support of mobility

Figure 29: ETSI MEC Technical Requirements

8.2.4 GSMA

The Figure below, taken from GSMA Intelligence [ref-12], identifies some other use cases, which are slightly different, but complementary to the ITU-R use cases. The GSMA goes into more detail and shows the dependency between the identified use cases and delay vs. bandwidth parameters.

The most demanding use cases are related to augmented/virtual reality and Tactile Internet, requiring simultaneously data rates above 1Gbps and delays in the order of a few milliseconds, which are in alignment with the ITU-R capabilities. In fact, those are the ones identified as not being enabled by today's networks (grey area).

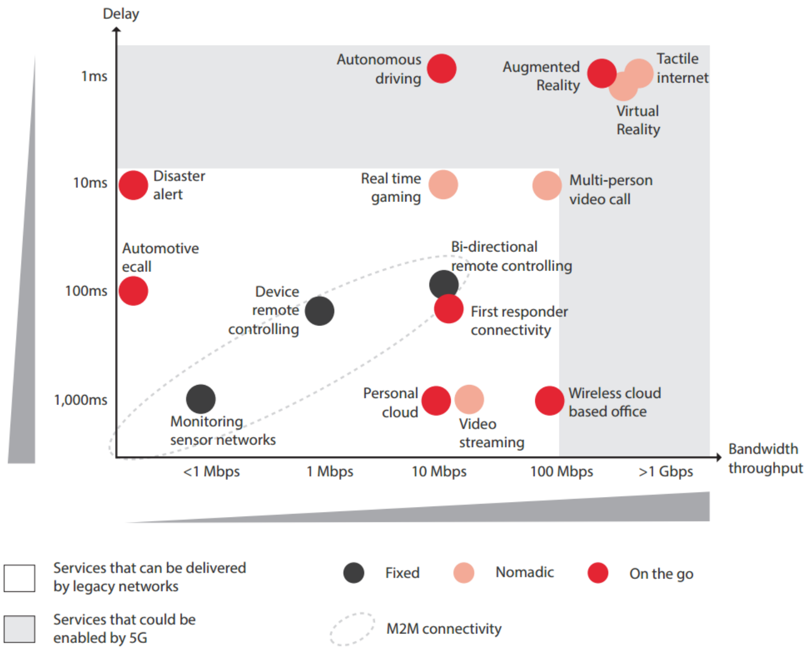


Figure 30: Bandwidth and latency requirements of identified 5G use cases

8.2.5 NGMN

The NGMN (*Next Generation Mobile Networks*) Alliance comprises Mobile Operators, Technology Vendors and Universities/Research Institutes, with the main objective of producing requirements, solutions and guidelines published in the form of as White Papers and Liaison Statements.

In its '5G White Paper' [ref-7], NGMN identifies a large number of requirements, organised according to the 'dimensions' as depicted in the figure:

Figure 31: NGMN Requirements Dimensions

In detail, the identified requirements are as follows:

User Experience	System Performance	Device Requirements
1 Consistent User Experience 2 User Experienced Data Rate 3 Latency 4 Mobility	1 Connection Density 2 Traffic Density 3 Spectrum Efficiency 4 Coverage 5 Resource and Signalling Efficiency	1 Operator Control Capabilities on Devices 2 Multi-Band-Multi-Mode Support in Devices 3 Device Power Efficiency 4 Resource and Signalling Efficiency
Enhanced Services	New Business Models	Network Deployment, Operation and Management
1 Connectivity Transparency 2 Location 3 Security 4 Resilience and High Availability 5 Reliability	1 Connectivity Providers 2 Partner Service Provider and XaaS Asset Provider 3 Network Sharing Model	1 Cost Efficiency 2 Energy Efficiency 3 Ease of Innovation and Upgrade 4 Ease of Deployment 5 Flexibility and Scalability 6 Fixed-Mobile Convergence 7 Operations Awareness 8 Operation Efficiency

9	Ultra-Low-cost Networks for Very Low-ARPU Areas
10	Ultra-Low-Cost Networks for Very Low-ARPU MTC Services

Figure 32: NGMN Requirements

NGMN also provides a number of use cases, organised in families and categories, giving an overview on the relationships between requirements and use cases:

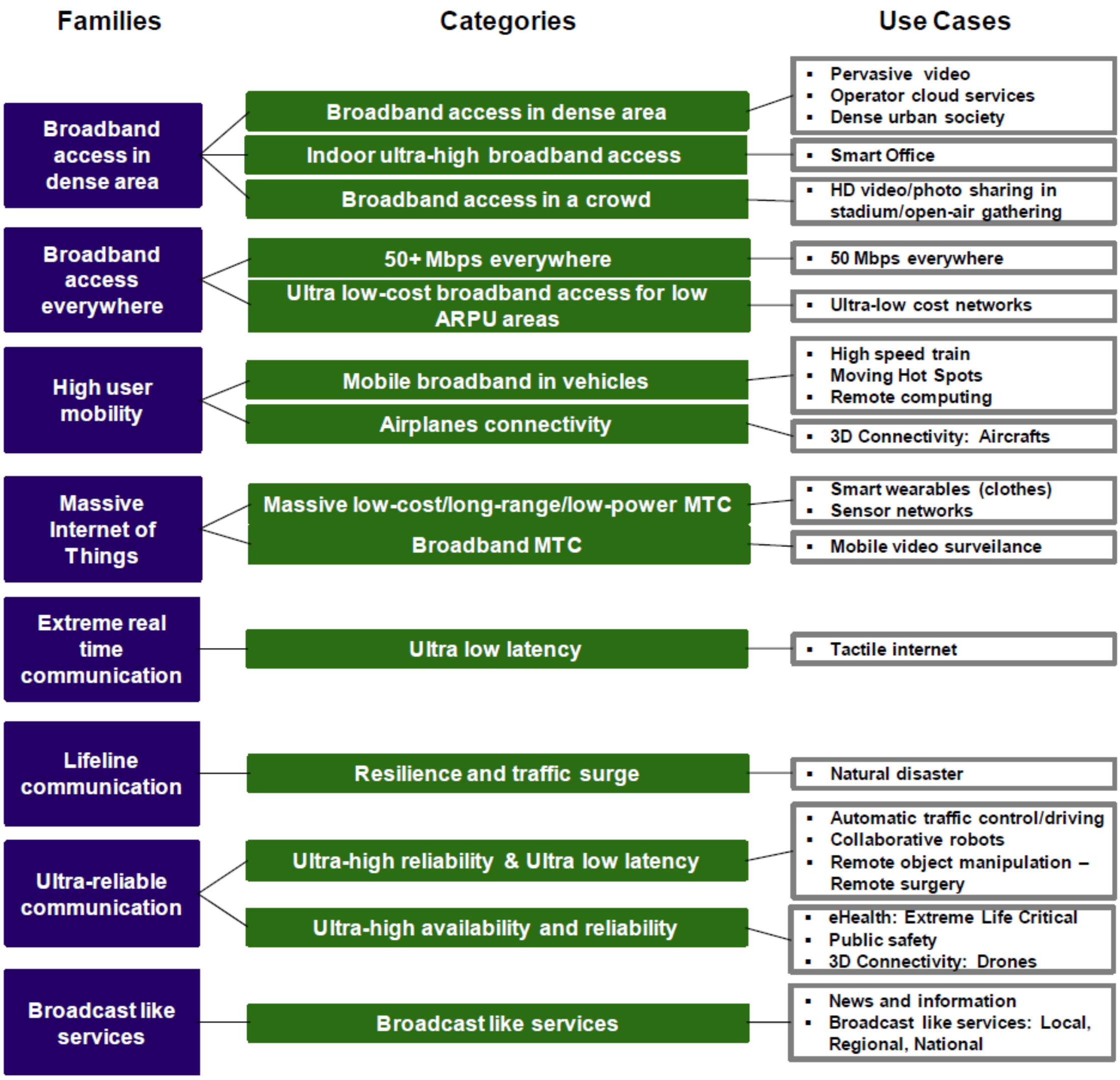


Figure 33: NGMN Use Cases

8.2.6 4G Americas

4G America [ref-13] divides the identified requirements in user and network driven requirements:

User-driven	Network-driven
1. Battery Life 2. Per-User Data Rate and Latency 3. Robustness and Resiliency 4. Mobility 5. Seamless User Experience 6. Context-Aware Network	1. Scalability 2. Network Capacity 3. Cost Efficiency 4. Automated System Management and Configuration

Figure 34: 4G Americas' Recommendations on 5G Requirements

8.3 General Requirements from the 5G-PPP Community

The 5G-PPP community is considering a set of Key Performance Indicators (KPIs), classified as Performance, Societal and Business KPIs. As part of the 5G-PPP community, Superfluidity will also contribute to the achievements of a subset of these KPIs. In the following tables, the 5G-PPP KPIs are listed and the relevance of the KPIs with respect to the scope of Superfluidity is analysed.

Performance KPIs

KPI		Relevance (High/Medium/ Low / N.A.)	Details on planned project contribution towards achieving the KPI
P1	Providing 1000 times higher wireless area capacity and more varied service capabilities compared to 2010.	High	Solutions to lower deployment and maintenance costs (e.g. BBU aggregation on the edge) through the use of standard high volume servers. Increased service diversity thanks to our slicing concept
P2	Reducing the average service creation time cycle from 90 hours to 90 minutes.	High	Automatic verification of service deployments before instantiation via static analysis to reduce the need for manual checks. Symbolic execution is a key enabler for this. Adoption of light weight virtualisation approaches such as unikernels or container-based approaches to significantly reduce service instantiation times Simplicity in network function deployment thanks to platform agnostic configuration/programming interfaces

P3	Facilitating very dense deployments of wireless communication links to connect over 7 trillion wireless devices serving over 7 billion people.	Medium	Cloud RAN concept particularly effective for dense environments Synchronization and collaboration between cells: Cloud-RAN, HW selection and acceleration, tackle this issue with a centralized BBU. Work on mobility at layer 4, based on Multipath TCP, to allow seamless handovers across a range of communication technologies. Connect to multiple technologies at once, favouring slow handover while connected via many links to fast handover of a single link, as done today
P4	Creating a secure, reliable and dependable Internet with a "zero perceived" downtime for services provision.	Medium	Almost instantaneous service instantiation. Use of symbolic execution to ensure robustness of the network, to check configurations and services before they are applied, and to ensure security by design, without sandboxing.

Societal KPIs

KPI		**Relevance** **(High / Medium / Low / N.A.)**	**Details on planned project contribution towards achieving the KPI**
S1	Enabling advanced User controlled privacy;	N.A.	
S2	Reduction of energy consumption per service up to 90% (as compared to 2010);	Medium	Reducing the Energy consumption in NFV data centres by improving infrastructure utilisation through more intelligent placement of workloads and continuous optimisation of placement decisions. Fast Service migration and shift or processing from one hardware to the other enables energy efficient scheduling (and having most of the platform in low energy consumption mode). More in general, cloud networking helps to adapt resources allocation, increasing efficiency by matching the workload

			characterisation to optimised allocation of resources through resource aware intelligent orchestration.
S3	European availability of a competitive industrial offer for 5G systems and technologies;	High	Consortium partners aiming at strategically placing themselves as a driving force in the area of converged 5G service and network architectures by becoming early adopters of Superfluidity's system Interface definition and contribution to standards Contributions to industry initiatives such as OPNFV to accelerate the adoption and deployment of open platform solutions to drive the adoption of NFV. Creation of high fidelity prototypes to demonstrate feasibility of the approach developed by Superfluidity
S4	Stimulation of new economically-viable services of high societal value like U-HDTV and M2M applications;	Medium	Edge cloud concept facilitates delay-sensitive services and crowd-targeted video services (e.g. multi-camera video streaming apps of events) Easing the development of M2M markets via virtualized infrastructure
S5	Establishment and availability of 5G skills development curricula (in partnership with the EIT).	N.A.	

Business-related KPIs

KPI		**Relevance** **(High / Medium / Low / N.A.)**	**Details on planned project contribution towards achieving the KPI**
B1	Leverage effect of EU research and innovation funding in terms of private investment in R&D for 5G systems in the order of 5 to 10 times;	Medium	Exploitation of Superfluidity results and prototypes by industrial partners into internal R&D programs Contribution to the open source community.

B2	Target SME participation under this initiative commensurate with an allocation of 20% of the total public funding;	Medium	Budget allocated to SMEs=€ 1.206.375, corresponding to 15,28% of the total public funding
B3	Reach a global market share for 5G equipment & services delivered by European headquartered ICT companies at, or above, the reported 2011 level of 43% global market share in communication infrastructure.	High	Use of Superfluidity collateral with customers to demonstrate feasibility of 5G use cases and solution paths to major 5G technical challenges. Investments by industrial partners.

8.4 References for Appendix B

[1] Ramjee Prasad and Albena Mihovska, "Challenges to 5G standardisation", https://itunews.itu.int/En/4619-Challenges-to-5G-standardization.note.aspx

[2] Recommendation ITU-R M.2083-0 (09/2015) IMT Vision – Framework and overall objectives of the future development of IMT for 2020 and beyond. http://www.itu.int/dms_pubrec/itu-r/rec/m/R-REC-M.2083-0-201509-I!!PDF-E.pdf

[3] Cisco Visual Networking Index: Forecast and Methodology, 2014-2019 White Paper, http://www.cisco.com/c/en/us/solutions/collateral/service-provider/ip-ngn-ip-next-generation-network/white_paper_c11-481360.html

[4] Wonil Roh, "5G Mobile Communications for 2020 and Beyond - Vision and Key Enabling Technologies", Samsung Electronics Corp., EUCNC 2014, Bologna, June 2014.

[5] Juniper Research, 5G Market Strategies, http://www.juniperresearch.com/researchstore/enabling-technologies/5g-market-strategies/4g-lte-evolution-spectrum-analysis

[6] "5G Vision - The 5G Infrastructure Public Private Partnership: the next generation of communication networks and services", February 2015, Available: https://5g-ppp.eu/wp-content/uploads/2015/02/5G-Vision-Brochure-v1.pdf

[7] NGMN 5G Whitepaper, https://www.ngmn.org/uploads/media/NGMN_5G_White_Paper_V1_0.pdf

[8] Mobile-Edge Computing – Introductory Technical White Paper, https://portal.etsi.org/Portals/0/TBpages/MEC/Docs/Mobile-edge_Computing_-_Introductory_Technical_White_Paper_V1%2018-09-14.pdf

[9] Mobile Edge Computing - A key technology towards 5G, http://www.etsi.org/images/files/ETSIWhitePapers/etsi_wp11_mec_a_key_technology_towards_5g.pdf

[10] TR244 TM Forum Information Framework Enhancements to Support ZOOM R15.0.1 Exploratory Report, https://www.tmforum.org/resources/technical-report/tr244-tm-forum-information-framework-enhancements-to-support-zoom-r15-0-0/

[11] Study on New Services and Markets Technology Enablers, 3GPP TR 22.891 SMARTER Study item, http://www.3gpp.org/DynaReport/22891.htm

[12] https://gsmaintelligence.com

[13] 4G Americas' Recommendations on 5G Requirements and Solutions http://www.4gamericas.org/files/2714/1471/2645/4G_Americas_Recommendations_on_5G_Requirements_and_Solutions_10_14_2014-FINALx.pdf

9 Appendix C – Clustering of use cases

In this Appendix we record information about our different attempts at grouping use cases. Sections 8.1 to 8.5 details different approaches we explored to classify use cases, and Section 8.6 contains some results from the different classifications. We ended up with a clustering based on requirements, as described in Section 4.

9.1 Categorisation by primary client actor

This categorization takes into account the primary actor or beneficiary of the considered use cases. We took into account three proposals for the classification based on the set of actors, than we converged on a fourth proposal. For the record, we first report the three proposals which have been discussed, and then the one that we have currently used for this categorization.

Proposal 1:

a) Residential end user
b) Business end user
c) “a thing” end user (i.e. IoT)
d) Internal user within the Superfluidity system: the provider that operates the Superfluidity system
e) Internal user outside the Superfluidity system e.g. a (virtual) Network Provider (carrier), Network operator, or the like

Proposal 2, a variant of proposal 1 with a different definition of d) and e) actors.

a) Residential end user
b) Business end user
c) “a thing” end user (i.e. IoT)
d) A Superfluidity component
e) Network/Service operator using the Superfluidity system

Proposal 3 was derived from a similar classification of use cases proposed by ETSI MEC

a) "Consumer-oriented services",
b) "Operator and third party services"
c) "Network performance and QoE improvements"

We tried to apply the different proposal to the classification of the use cases, however it was difficult to differentiate cases d) and e) in proposals 1) and 2) and therefore the decision was made to merge them in a single one. It was found that differentiating between residential and business end users was useful, so we maintained it throughout the process, and the network performance category from proposal is also included 3), the final categories are listed hereafter.

a) Residential end user
b) Business end user
c) “a thing” end user (i.e. IoT)
d) Network operator / Service provider (also virtual)
e) Network performance and QoE improvements

9.2 Classification by network section

The following categories were identified with respect to the affected network sections:

a) Wireless access / RAN /fronthaul
b) MEC
c) Backhaul (anything between the radio access and core)
d) Core (includes everything behind the packet GW)

The "MEC" section refers to the elements that are capable of providing processing or storage resources close to the RAN elements.

9.3 Classification based on a harmonized set of tags

In the use cases provided by the partners, each partner has independently proposed a set of tags, indicated as "secondary classification labels". We have selected a subset of these labels and tried to reclassify the use cases based on it. We report hereafter the set of tags that have been selected and then the tags that have been "discarded". This classification is somehow redundant with respect to the one described in the next section, based on "classes" or use cases. Anyway, it has been a useful exercise to understand the different aspects of the uses cases and to identify a good set of use case categories.

These are the tags that have been considered:

- Reprogrammable RAN / RAN slicing
- Layer 1 services / Mac services
- Dynamic rates and topology
- Usage/Service/platform monitoring and analytics
- Location Based Services / Mobile edge computing
- Local breakout (Corporate/Campus - Big Events/Crowd)
- Caching Services
- TCP/http Optimization / Adaptive media delivery
- Content encryption / content processing
- M2M communication
- IoT /IoT virtual network
- Low Latency /High Throughput Applications
- Tactile Communication Services (*)
- Functions Orchestration / Service chaining
- Security

(*) The Tactile Internet will require extremely low latency combined with high availability, reliability and security.

The tags that have been left out are the following:

- Context and Location based services
- Autonomous intelligent services
- Enhanced Platform Awareness
- Scheduling as a service
- MIMO as a services
- SON as a service
- User following services
- Cloud offloading
- Transparent CDN
- Re-routing of network traffic
- Low resource overhead hypervisor
- Massive consolidation of VMs
- Communication and shared resources across hypervisor instances
- Advertisement
- Private Virtualised CDN
- TV contents
- Small Cells
- Services platforms
- Enterprise/operator converged services
- Community services
- Hardware Accelerators
- Enhanced group communications
- Intrinsic security mechanisms
- Advertisement
- Targeting
- Origin caching

9.4 Classification based on classes of use cases

Using the 8 use case samples provided in the technical annex (see Section 7.6), we classified the 23 use cases. We found that this classification was not satisfying, because some contributed use cases were not fitting in the sample use cases, and some sample use cases were duplicated, because they were matching the same set of contributed use cases. We considered a different set of classes, as listed hereafter:

C1	RAN, wireless access (TA-2) (TA-6)
C2	On-the-fly Monitoring (TA-7)
C3	MEC / Localised services – processing (TA-3)
C4	MEC / Localised services – storage (TA-3, TA-1, TA-8)

C5 MEC / Localised service – local breakout (TA-3)

C6 Edge offloading (TA-5)

C7 Emergency communications

C8 Security

In section 9.6 we show how the new classes are able to provide a better clustering of the contributed use cases.

9.5 Categorisation by layer in dynamic design lifecycle

Another approach that is to be proposed for use-case categorization is the application of a dynamic design lifecycle. With this approach, the resource/service stack is divided into layers such that there is a recursively layered relationship between the runtime and design phases required for service delivery. This principle is depicted below, where the dynamic runtime actions in one layer, i.e. host layer, create service entities that are used for the design of another layer, i.e. the client layer. In contrast with traditional approaches where the design phases are always assumed to be static, this approach will enable fluidity, flexibility, and dynamic reconfiguration.

	Static Elements	Dynamic Operations	Service Entities	Abstracted Dynamics
Network Service Layer (Service Chains)	▪ Virtual Network Functions (VNFs)	▪ Create VNFs ▪ Configure VNFs ▪ Connect VNFs	▪ Interconnected Network Services	▪ Instantiate Components ▪ Connect Components ▪ Configure Components in Context
	Client		Host	
Virtual Network Function Layer	▪ Virtual Network Functions Component Instances (VNFCIs)	▪ Create VNFCIs ▪ Configure VNFCIs ▪ Connect VNFCIs	▪ Interconnected VNFs	▪ Instantiate Components ▪ Connect Components ▪ Configure Components in Context
	Client		Host	
Network Function Virtualization Infrastructure Layer	▪ Virtual Machines (VMs) ▪ Virtual Networks (VNs)	▪ Create VMs & VNs ▪ Configure VMs & VNs ▪ Connect VNs to VMs	▪ Interconnected VNFCI hosts	▪ Instantiate Components ▪ Connect Components ▪ Configure Components in Context
	Client		Host	
Infrastructure Isolation Layer	▪ Hypervisor ▪ Operating System (OS) ▪ Network Transport Paths	▪ Install OS & Hypervisor ▪ Configure Hypervisor ▪ Create vSwitches ▪ Create Transport Paths ▪ Join Transport Paths to vSwitches	▪ VM Host ▪ VN Host	▪ Instantiate Components ▪ Connect Components ▪ Configure Components in Context
	Client		Host	
Physical Infrastructure Layer	▪ Servers ▪ Storage ▪ Fibre, Cable ▪ Switches etc.	▪ Install Hardware Devices ▪ Plug Interface Connectors ▪ Boot & Configure Hardware Devices	▪ Compute Machine ▪ Network Transport Infrastructure	▪ Instantiate Components ▪ Connect Components ▪ Configure Components in Context

Figure 35: Recursively layered design and runtime layers for service delivery

9.6 Results from classification of the use cases

The results from the mapping the 23 contributed use cases into the original set of classes provided in the technical annex is shown below. An "X" in the intersection between a use case and a use case class indicates a strong matching, a "/" indicate a loose matching.

	Minimum-Delay Cloud storage (TA-1)	RAN As A Service (TA-2)	Localized services (TA-3)	Pooling (TA-4)	Edge offloading (TA-5)	Portable signal processing (TA-6)	On-the-fly Monitoring (TA-7)	Virtualized CDN operators (TA-8)
Wireless Software Defined fronthauling (ALUBL-2)		X		X		/		
On-the-fly monitoring (BT-1)							X	
S/Gi-LAN Services on the Edge (CITRIX-1)			X		/			X
Dynamic MAC services allocation in Cloud RAN (ALUIL BGU-1)		X		X				
Internet of Things (IoT) & SUPERFUIDITY Platform Scenario (CNIT-1)			X					
Context-adapted data delivery (CNIT-2)			X					X
Mobile Based Augmented Reality for User Experience Enhancement (Intel-1)			X		/			
Performance Optimization for Distributed Multimedia Content Delivery (Intel-2)			X					X
Context Aware Smart Living (Intel-3)			X					
Mobile services offloading (NEC-1)	X				X			
Transparent web service acceleration (ONAPP-1)			X					
Rapid and massiv.-scalable instant. of high perf (virt.) applic. instances (ONAPP-2)			X					
Local Breakout (LBO) (PTIN-1)	X							
virtual Convergent Services (vCS) (PTIN-2)								
Video Orchestration and Optimization (PTIN-3)			X					
Virtual CDN for TV contents distribution (PTIN-4)	X		X					X
Business Communication Services (TPIN-5)			X					
Anti NDP Spoofing software implementation (Telcaria-1)								
Protection against DDoS (Telcaria-2)								
Emergency communications (TID-1)		/	/					
Late transmuxing (USTR-1)			X		/			
Remix (USTR-2)			X					
Backend Storage Caching (USTR-3)	?		/					/

Figure 36: Mapping between the contributed use cases and the use case classes from the technical annex

The next figure shows the relations between the contributed use cases. The use cases are not properly ordered, because the matrix appears almost random.

	5G RAN "network slices" (ALUBL-1)	Wireless Software Defined fronthauling (ALUBL-2)	On-the-fly monitoring (BT-1)	S/Gi-LAN Services on the Edge (CITRIX-1)	Dynamic MAC service allocation in Cloud RAN (ALUIL BGU-1)	Internet of Things (IoT) & SUPERFLUIDITY Platform Scenario (CNIT-1)	Context-adapted data delivery (CNIT-2)	Mobile Based Augmented Reality for User Experience Enhancement (Intel-1)	Performance Optimization for Distributed Multimedia Content Delivery (Intel-2)	Context Aware Smart Living (Intel-3)	Mobile services offloading (NEC-1)	Transparent web service acceleration (ONAPP-1)	Rapid and massively-scalable instant. of high perf (virt.) applic. instances (ONAPP-2)	Local Breakout (LBO) (PTIN-1)	Virtual Convergent Services (VCS) (PTIN-2)	Video Orchestration and Optimization (PTIN-3)	Virtual CDN for TV content distribution (PTIN-4)	Business Communication Services (TPIN-5)	Anti NDP Spoofing software implementation (Telcaria-1)	Protection against DDoS (Telcaria-2)	Emergency communications (TID-1)	Late transmuxing (USTR-1)	Remix (USTR-2)	Backend Storage Caching (USTR-3)
5G RAN "network slices" (ALUBL-1)		/			/																/			
Wireless Software Defined fronthauling (ALUBL-2)	/				/									/										
On-the-fly monitoring (BT-1)																								
S/Gi-LAN Services on the Edge (CITRIX-1)						/	X	X	X	/		X	X		/	X	X	X				X	X	X
Dynamic MAC services allocation in Cloud RAN (ALUIL BGU-1)	/	/																						
Internet of Things (IoT) & SUPERFLUIDITY Platform Scenario (CNIT-1)				/						X			/	X										
Context-adapted data delivery (CNIT-2)				X					X			X				X		X				X	X	/
Mobile Based Augmented Reality for User Experience Enhancement (Intel-1)				X								X	X			X	X					X	X	X
Performance Optimization for Distributed Multimedia Content Delivery (Intel-2)				X			X					X	X			X	X	X				X	X	X
Context Aware Smart Living (Intel-3)				/		X								X										
Mobile services offloading (NEC-1)														/								/		/
Transparent web service acceleration (ONAPP-1)				X			X	X	X				X			X	/	X				X	X	/
Rapid and massiv.-scalable instant. of high perf (virt.) applic. instances (ONAPP-2)				X		/		X	X			X				X	/	X				X	X	/
Local Breakout (LBO) (PTIN-1)		/				X				X	/							X			X			/
Virtual Convergent Services (VCS) (PTIN-2)				/												X	/	X				X	X	
Video Orchestration and Optimization (PTIN-3)				X			X	X	X			X	X		X		/	X				X	X	
Virtual CDN for TV contents distribution (PTIN-4)				X				X	X			/	/		/	/						/	/	X
Business Communication Services (TPIN-5)				X			X		X			X	X	X	X	X								
Anti NDP Spoofing software implementation (Telcaria-1)																				X				
Protection against DDoS (Telcaria-2)																			X					
Emergency communications (TID-1)	/													X										
Late transmuxing (USTR-1)				X			X	X	X		/	X	X		X	X	/						X	/
Remix (USTR-2)				X			X	X	X			X	X		X	X	/					X		/
Backend Storage Caching (USTR-3)				X			/	X	X		/	/	/	/			X					/	/	

Figure 37: Correlation between the contributed use cases, in the original order.

With a proper reordering of the use cases, which was based on the analysis of the correlation among the use cases, as well as on the classifications using the network section and the main actor criteria, we obtained the correlation matrix shown in the following figure. The matrix is almost block-wise diagonal, showing that a satisfactory ordering of the use cases has been reached.

	5G RAN "network slices" (ALUBL-1)	Wireless Software Defined fronthauling (ALUBL-2)	Dynamic MAC service allocation in Cloud RAN (ALUIL BGU-1)	On-the-fly monitoring (BT-1)	S/Gi-LAN Service on the Edge (CITRIX-1)	Context-adapted data delivery (CNIT-2)	Performance Optimization for Distributed Multimedia Content Delivery (Intel-2)	Transparent web service acceleration (ONAPP-1)	Rapid and massiv-scalable instant. of high perf (virt.) applic. instances (ONAPP-2)	virtual Convergent Services (vCS) (PTIN-2)	Video Orchestration and Optimization (PTIN-3)	Mobile Based Augmented Reality for User Experience Enhancement (Intel-1)	Virtual CDN for TV content distribution (PTIN-4)	Business Communication Services (TPIN-5)	Late transmuxing (USTR-1)	Remix (USTR-2)	Backend Storage Caching (USTR-3)	Local Breakout (LBO) (PTIN-1)	Context Aware Smart Living (Intel-3)	Internet of Things (IoT) & SUPERFLUIDITY Platform Scenario (CNIT-1)	Mobile services offloading (NEC-1)	Anti NDP Spoofing software implementation (Telcaria-1)	Protection against DDoS (Telcaria-2)	Emergency communications (TID-1)
5G RAN "network slices" (ALUBL-1)		/	/																					/
Wireless Software Defined fronthauling (ALUBL-2)	/		/															/						
Dynamic MAC services allocation in Cloud RAN (ALUIL BGU-1)	/	/																						
On-the-fly monitoring (BT-1)																								
S/Gi-LAN Services on the Edge (CITRIX-1)						X	X	X	X	/	X	X	X	X	X	X	X		/	/				
Context-adapted data delivery (CNIT-2)					X		X	X			X			X	X	X	/							
Performance Optimization for Distributed Multimedia Content Delivery (Intel-2)					X	X		X	X		X		X	X	X	X	X							
Transparent web service acceleration (ONAPP-1)					X	X	X		X		X	X	/	X	X	X	/							
Rapid and massiv.-scalable instant. of high perf (virt.) applic. instances (ONAPP-2)					X		X	X			X	X	/	X	X	X	/			/				
viltual Convergent Services (vCS) (PTIN-2)					/						X		/	X	X	X								
Video Orchestration and Optimization (PTIN-3)					X	X	X	X	X	X		X	/	X	X	X								
Mobile Based Augmented Reality for User Experience Enhancement (Intel-1)					X			X	X		X		X		X	X	X							
Virtual CDN for TV contents distribution (PTIN-4)					X		X	/	/	/	/	X			/	/	X							
Business Communication Services (TPIN-5)					X	X	X	X	X	X	X							X						
Late transmuxing (USTR-1)					X	X	X	X	X	X	X	X	/			X	/				/			
Remix (USTR-2)					X	X	X	X	X	X	X	X	/		X		/							
Backend Storage Caching (USTR-3)					X	/	X	/	/			X	X		/	/		/			/			
Local Breakout (LBO) (PTIN-1)		/												X			/		X	X	/			X
Context Aware Smart Living (Intel-3)					/													X		X				
Internet of Things (IoT) & SUPERFUIDITY Platform Scenario (CNIT-1)					/				/									X	X					
Mobile services offloading (NEC-1)															/		/	/						
Anti NDP Spoofing software implementation (Telcaria-1)																							X	
Protection against DDoS (Telcaria-2)																						X		
Emergency communications (TID-1)	/																	X						

Figure 38: Correlation between the contributed use cases, properly reordered

The mapping of the uses case into the newly proposed classes is shown next. Also in this case the matrix is more regular than the one presented earlier and allows identifying clusters of correlated use cases.

	RAN, wireless access	On-the-fly Monitoring (TA-7)	Localized services, processing	Localized services, storage	Localize services, local breakout	Edge offloading (TA-5)	Security	Emergency communications
5G RAN "network slices" (ALUBL-1)	X							
Wireless Software Defined fronthauling (ALUBL-2)	X							
Dynamic MAC services allocation in Cloud RAN (ALUIL BGU-1)	X							
On-the-fly monitoring (BT-1)		X						
S/Gi-LAN Services on the Edge (CITRIX-1)			X	X		/		
Context-adapted data delivery (CNIT-2)			X	X				
Performance Optimization for Distributed Multimedia Content Delivery (Intel-2)			X	X				
Transparent web service acceleration (ONAPP-1)			X					
Rapid and massiv.-scalable instant. of high perf (virt.) applic. instances (ONAPP-2)			X					
virtual Convergent Services (vCS) (PTIN-2)								
Video Orchestration and Optimization (PTIN-3)			X					
Mobile Based Augmented Reality for User Experience Enhancement (Intel-1)			X			/		
Virtual CDN for TV contents distribution (PTIN-4)			X	X				
Business Communication Services (TPIN-5)			X					
Late transmuxing (USTR-1)			X			/		
Remix (USTR-2)			X					
Backend Storage Caching (USTR-3)			/	X				
Local Breakout (LBO) (PTIN-1)				X	X			
Context Aware Smart Living (Intel-3)			X					
Internet of Things (IoT) & SUPERFUIDITY Platform Scenario (CNIT-1)			X					
Mobile services offloading (NEC-1)			/	/		X		
Anti NDP Spoofing software implementation (Telcaria-1)							X	
Protection against DDoS (Telcaria-2)							X	
Emergency communications (TID-1)	/		/					X

Figure 39: Mapping between the (reordered) contributed use cases and the new classes

The final two figures provide a classification based on the tags, and according to the main actor and the affected network section.

	5G RAN "network slices" (ALUBL-1)	Wireless Software Defined fronthauling (ALUBL-2)	Dynamic MAC services allocation in Cloud RAN (ALUILBGU-1)	On-the-fly monitoring (BT-1)	S/Gi-LAN Services on the Edge (CITRIX-1)	Context-adapted data delivery (CNIT-2)	Performance Optimization for Distributed Multimedia Content Delivery (Intel-2)	Transparent web service acceleration (ONAPP-1)	Rapid and massive.-scalable instant. of high perf (virt.) applic. instances (ONAPP-2)	virtual Convergent Services (vCS) (PTIN-2)	Video Orchestration and Optimization (PTIN-3)	Mobile Based Augmented Reality for User Experience Enhancement (Intel-1)	Virtual CDN for TV contents distribution (PTIN-4)	Business Communication Services (TPIN-5)	Late transmuxing (USTR-1)	Remix (USTR-2)	Backend Storage Caching (USTR-3)	Local Breakout (LBO) (PTIN-1)	Context Aware Smart Living (Intel-3)	Internet of Things (IoT) & SUPERFUIDITY Platform Scenario (CNIT-1)	Mobile services offloading (NEC-1)	Anti NDP Spoofing software implementation (Telcaria-1)	Protection against DDoS (Telcaria-2)	Emergency communications (TID-1)
Reprogrammable RAN / RAN slicing	X	X																						
Layer 1 services / Mac services	X	X	X																					
Dynamic rates and topology		X																						
Usage/Service/platform monitoring and analytics				X			X																	X
Location Based Services / Mobile edge computing					X	X			X	X				X	X	X	X				X			X
Local breakout (Corporate/Campus - Big Events/Crowd)										X				X				X						
Caching Services					X	X					X		X		X		X				X			X
TCP/http Optimization / Adaptive media delivery					X	X		X			x		X		X	X								
Content encryption / content processing						X								X										
Low Latency /High Throughput Applications					X		X					X									X			
Tactile Communication Services(*)					X				X			X												
Functions Orchestration / Service chaining					X		X		X			X												
M2M communication																			X	X				
IoT /IoT virtual network																			X	X				
Security																						X	X	X

Figure 40: Classification of the contributed use cases, based on the selected tags

	a) residential end user	b) business end user	c) a thing end user (ie IoT)	d) Network operator / Service provider (also virtual)	e) "Network performance and QoE improvements"	a) Wireless access / RAN / fronthaul	b) MEC	c) Backhaul	d) Core
5G RAN "network slices" (ALUBL-1)				X		X		X	
Wireless Software Defined fronthauling (ALUBL-2)				X		X			
Dynamic MAC services allocation in Cloud RAN (ALUIL BGU-1)				X		X			
On-the-fly monitoring (BT-1)				X	X	X	X	X	X
S/Gi-LAN Services on the Edge (CITRIX-1)	/	/	/	X			X		X
Context-adapted data delivery (CNIT-2)	X	X			X		X		
Performance Optimization for Distributed Multimedia Content Delivery (Intel-2)	X	X			X		X		
Transparent web service acceleration (ONAPP-1)	X	X					X		
Rapid and massiv.-scalable instant. of high perf (virt.) applic. instances (ONAPP-2)	X	X	/				X		
virtual Convergent Services (vCS) (PTIN-2)	X	/					X		
Video Orchestration and Optimization (PTIN-3)	X	/					X		
Mobile Based Augmented Reality for User Experience Enhancement (Intel-1)	X	X			X		X		
Virtual CDN for TV contents distribution (PTIN-4)				X			X		
Business Communication Services (TPIN-5)		X					X		
Late transmuxing (USTR-1)	X	/					X		
Remix (USTR-2)	X	/					X		
Backend Storage Caching (USTR-3)	X	/					X		
Local Breakout (LBO) (PTIN-1)				X	X	X		X	/
Context Aware Smart Living (Intel-3)			X				X		
Internet of Things (IoT) & SUPERFUIDITY Platform Scenario (CNIT-1)			X				X		
Mobile services offloading (NEC-1)	X	X			X		X		
Anti NDP Spoofing software implementation (Telcaria-1)				X	X	X	X		
Protection against DDoS (Telcaria-2)				X	X	X	X		
Emergency communications (TID-1)				X	X	X	X	X	X

Figure 41: Classification of the (reordered) contributed use cases with respect to: main actor (on the left), network section (on the right)