enhanced COntent distribUtion with Social INformation

www.ict-ecousin.eu

Deliverable D2.1

Initial Report on Use Cases and Requirements

Public deliverable, Version 1.1, 7 May 2013

Authors

FT	Patrick Truong, Bertand Mathieu, Luc Le Beller, Yannick Le Louédec
AL-BELL	Danny De Vleeschauwer
IMDEA	
TSP	Noël Crespi
ALUD	Ivica Rimac, Klaus Satzke
TUD	David Hausheer
TI	Claudio Venezia
UCAM	
UC3M	Rubén Cuevas Rumín

Reviewers Ángel Cuevas Rumín, Patrick Truong, Yannick Le Louédec

Abstract

The present deliverable D2.1 is the first technical deliverable of the eCOUSIN project. This document describes the use cases that allow to introduce a practical view of the objectives, the involved actors and the potential benefits obtained with the implementation of eCOUSIN project. In addition, starting from use cases requirements, we capture the functional and technical requirements in such a way that they will drive architectural decisions for the eCOUSIN system and they will be used to validate the final architecture. Finally, we briefly discuss privacy considerations.

Project funded by the European Union under the
Information and Communication Technologies FP7 Cooperation Programme
Grant Agreement number 318398

EXECUTIVE SUMMARY

The democratization of the Internet as a large repository of content, mainly driven by the emergence of voracious bandwidth-consuming services, induces new challenges for networking actors that need to continuously design innovative solutions for improving the efficiency of content delivery to end-users while optimizing their operational costs. Moreover, the recent uprising of Online Social Networks (OSNs) has considerably evolved the Internet usage to many-to-many distribution and sharing information with an increasing need for mobility, so that it is now well recognized that the social graph has an influence on content delivery. This means for example that social information can be used as a barometer to predict content popularity, and thus to enhance content delivery.

In the eCOUSIN project, we propose to measure the impact of OSNs on content distribution with a framework for a social-aware content delivery network architecture. To this end, we first lay out in this document the different use cases that cover the main objectives of eCOUSIN. These use cases allow us to express a complete set of requirements for the system that will then drive our architectural decisions. We organize requirements by adhering to the following layered scheme that will be helpful to find out technical solutions for designing and implementing a network architecture for enhancing content distribution with social information:

i. Social-Content Interdependencies layer, which integrates the social and content information extraction plane on the one hand, and the social and content information management plane on the other hand)
ii. Content Dissemination layer
iii. Network layer.

The resulting set of requirements will also be updated during the project development as we advance in maturity in our work.

TABLE OF CONTENTS

LIST OF FIGURES

INTRODUCTION

Global IP traffic has increased considerably over the past decade, and will continue to rapidly evolve over the next years, mainly due to new usage patterns on Internet, which are more and more content-centric. Nowadays, multimedia content is responsible for a major portion of the Internet traffic and its importance will grow in the coming years [1]. For instance Netflix [2], a video distribution platform in North America, already generates 33% of the downstream traffic in that continent as reported by the last Sandvine report from fall 2012 [3]. In addition, the same report shows that BitTorrent, the most important p2p application for file sharing, represents a large part, around 30%-40%, of the upstream traffic in North America and Europe. Aligned with these evidences the proliferation in the use of smartphones and tablets all over the world has led to an exponential increase in the access to multimedia content from mobile devices. Cisco envisions that by 2014 video content will account 66% of the global mobile data traffic.. This ever increasing traffic demand for multimedia content is mainly driven by the explosion of enhanced and simplified multimedia services, which takes advantage of the convergence and integration of networking technologies to create a unified digital world through a large set of heterogeneous connected terminals (PC, tablets, smartphones, connected TVs, set-top boxes, etc.) and high-rate access networks (ADSL, fiber, wireless networks, etc.). Network service providers are thus constantly coming up with engineering solutions to deliver content to end users with the best quality of service.

In the meantime, Online Social Networks (OSNs) have gained an enormous popularity in a very short time, shifting usage on Internet away from traditional information browsing to online consumption and sharing of any type of content, including user-generated content. OSNs do not only connect and make people interact together, but are also extensively used for searching and sharing information and multimedia content in real time. In addition, due to the fast viral spread of content sharing over OSNs, content popularity on the Internet is nowadays being dictated by the way OSNs' users can turn a piece of information into a buzz. Therefore, content distribution is highly influenced by the multiple social interactions happening among end users. This means that by exploiting information available in OSNs, we can propose novel algorithms, mechanisms and protocols to more efficiently distribute and deliver content through the Internet.

In the eCOUSIN project, we aim at designing a social-aware network architecture with built-in content dissemination functionalities that exploits the social-content interdependencies to improve the efficiency of content delivery. In particular, the overall goals are:

- Design of an on-net operational framework that tightly integrates network functionalities and content-related service functionalities
- Implementation of high performance distributed tools for collecting necessary data to study and model the social-content interdependencies
- Design of algorithms that exploit social information for placing and delivering contents in an optimized manner with a special focus on mobile environments

In order to sketch out the design of a network architecture satisfying the eCOUSIN objectives, one classical method is to formalize the requirements that the system must accommodate. Common inputs for deriving those requirements are use cases that should be as complete as possible to fulfil the main aspects of the project.

This document describes use cases relevant to the scope of eCOUSIN that are based on key usage scenarios. In addition, it captures functional and technical requirements from those use cases in such a way that they constitute a solid basis to drive architectural decisions for eCOUSIN and, moreover, can be used to fully validate the final architecture. Therefore, a representative set of use cases is described in the document to cover the different goals of the project.

The document is organized as follows. After introducing the scope of the use cases in Section 1, we describe them in Sections 2, 3 and 4. From these use cases, we extract, in Section 5, a set of global requirements for the social-aware content delivery ecosystem. In section 6, we present a brief discussion on privacy considerations that need to be carefully taken into account during the project implementation. We finalize the paper with some concluding remarks.

1. SCOPE OF THE USE CASES

The scope of this use case study is defined by the overall goals of the eCOUSIN project, as outlined in the Introduction. We have a strong emphasis on architectural design aspects. Security and privacy issues are only addressed in this document from a high-level perspective, since they will be covered in more detail in a separate technical report later in the project.

One of the most common practices to improve content delivery in current networks is using caching strategies, which allows reducing network traffic and access delay times. Caching is often used together with prefetching techniques that aim at predicting future content requests and provisioning those contents into the cache before next requests arrive. In order to further improve the content delivery efficiency, Content Distribution Networks (CDNs) have been designed to push and store content at the edges of the Internet. To this end, surrogate servers are placed on these locations to store a copy of the origin servers' content. The content placement problem, which consists in determining where to place surrogate servers and which surrogates will replicate which objects, is a key element in content delivery that needs to be optimized for getting better performance. In other words, the efficiency of content delivery depends on the right selection of the location where content to be delivered to end users is stored. Optimized content selection approaches can improve client downloading experience and, at the same time, reduce content providers (i.e. central servers) and network operators load.

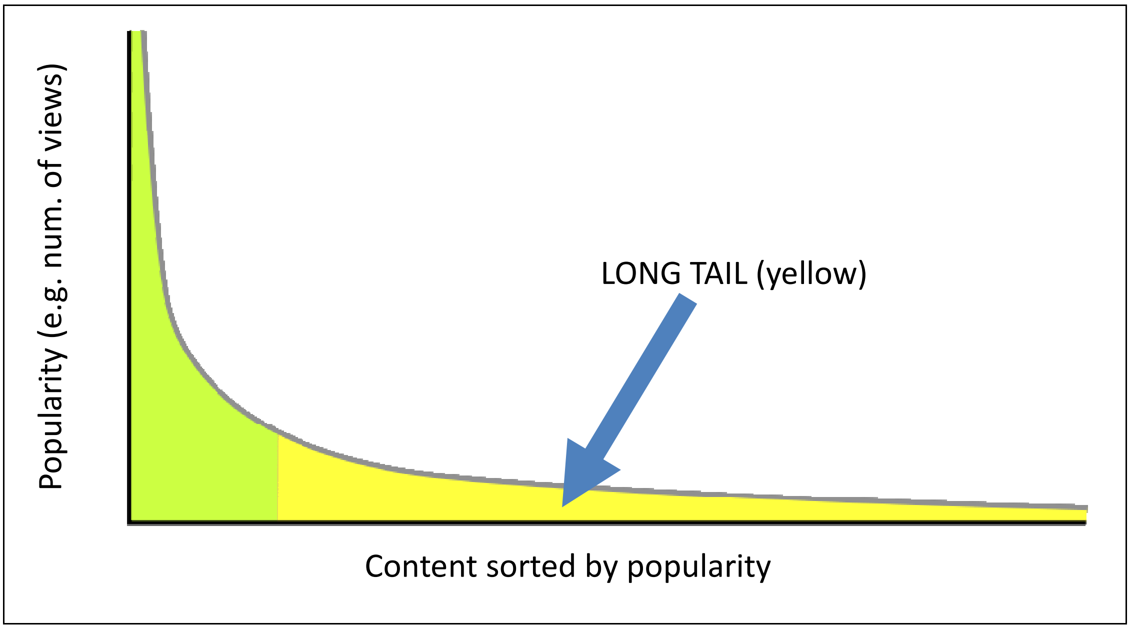


Figure 1: Long Tail Popularity Distribution

All the current techniques for content selection generally rely on content popularity statistics, which are based on recent historical request measurements and prediction methods. However, such statistics are often difficult to estimate for new content such as user generated content. Especially, this may introduce a bias when content consumption is based on the Long Tail model. Multimedia content ranges from very popular content (i.e. popular TV Shows, recent released movies, etc) to user generated content (UGC) whose audience is in many cases limited to the social environment of the content source. In the literature, Zipf has been shown to be a good approximation for the popularity distribution of multimedia content in systems like P2P file-sharing [4] or YouTube [5]. This means that a small portion of available content enjoys a high popularity while the majority is actually receiving few downloads/views, generating the so-called long tail distribution. Therefore, in the next years the proliferation of multimedia, and especially video, content will most likely increase the number of content in the long tail. In this context, there will be a large number of requests that are of interest to only small groups of users linked together by some social relationship such as the case in online social networks. Moreover, in case of Web content, some experiments (e.g. [6]) reveal that content replication based on clustering approaches, which determine groups of users based on correlation metrics (e.g. sharing similar usage patterns), are more appropriate as it allows to reduce client download time and the load on servers.

Due to the exponential growth of connected mobile devices, a large portion of the Internet traffic goes now through mobile networks. This has led to a new paradigm for mobile data offloading from cellular networks to WLAN networks (such as Wi-Fi). For network operators, the main motivation for doing mobile data offloading is to ease cellular networks to avoid congestion. For end users, the purpose for offloading is twofold: (i) they reduce data service billing, and (ii) they enjoy a faster connection. As a consequence, if we are able to efficiently exploit time-based usage patterns (users usually connect on their usual Wi-Fi networks in specific periods of time), mobile data offloading can be used as a complementary method in addition to aforementioned practices for improving content delivery. As an example, a user whose interests in content are driven by social recommendations may choose to install a simple client on its smartphone that would proactively prefetch relevant recommended content when her/his mobile device is connected through a Wi-Fi access point.

Based on these observations and the project objectives, we have highlighted the following key concepts to define representative use cases in the context of eCOUSIN. In addition, these use cases allow us to extract general requirements that will serve as a basis for designing the eCOUSIN system, namely a social-aware content delivery network architecture that takes into account social relationship and content information to improve content dissemination within the Internet:

- Content Placement (Section 2): how to strategically place content based on users' social and location information and how to efficiently provision network resources for optimizing content delivery?
- Time-Unconstrained Content Delivery (Section 3): how to exploit mobile data offloading principle to leverage content delivery for getting an improved quality of experience for end users on mobility?
- User-Generated Content Delivery (Section 4): user generated content is a trending usage in today's networking applications such that it raises an important problem with the current content selection techniques which are not adapted for the Long Tail distribution.

2. CONTENT PLACEMENT USE CASES

An ever-growing fraction of Internet traffic is produced by content distribution applications. This has triggered the development of new content placement technologies, such as CDNs that aim to place content close to end users, leading to substantial benefit for different players in the Internet such as Access ISPs or Content Providers, which have reduced their operational costs (e.g., reducing their transit traffic), and end users who have improved their QoE (e.g. lower delivery delay). The recent irruption of Online Social Networks in the Internet has provided a new important source of information (e.g., end-users' relationships, information flow patterns, etc.) that can be leveraged to further improve existing content placement techniques. In this section we present two practical use cases that show how we can leverage social information in eCOUSIN to develop enhanced content placement solutions.

2.1 Enhanced Content Placement Using Users' Social and Coarse-Grain Location Information

CDNs provide efficient mechanisms to distribute popular content by placing this content close to the end user and thus reducing the network traffic. However, the offered performance for mid and low popular content is not good enough mostly due to the difficulty to predict where this type of content is going to be consumed. Fortunately, the recent irruption of Online Social Networks (OSNs) offers valuable information to predict where this mid and low popular content is going to be consumed. An important fraction of this content, in many cases User Generated Content (UGC), is nowadays distributed through Online Social Networks (OSNs). Therefore, we can leverage the information of a specific user's social relationships to predict where the content that (s)he shares is going to be consumed.

2.1.1 Description of the Use Case

2.1.1.1 Scenario-Based Illustration

Let us consider the example depicted in Figure 2:

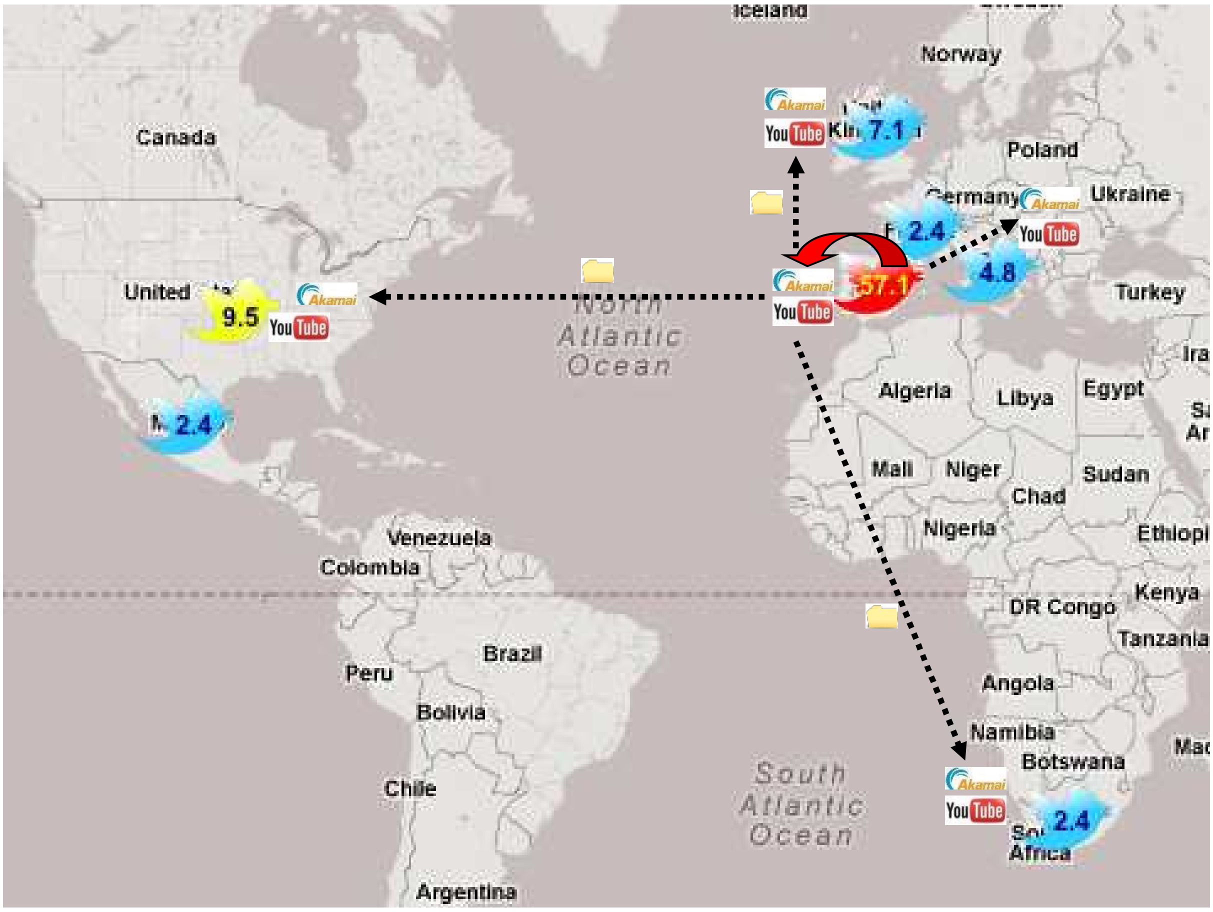


Figure 2: Enhanced Content Placement Using Users' Social and Location Information

The Figure 2 shows the geographical distribution of the followers for a user, Alice, in Twitter. Therefore, when Alice publishes a tweet including a link to a given content (e.g., a photo or a YouTube video), based on the distribution of her followers we know that the content is more likely to be consumed in Spain where 57% of Alice's followers are located. The second location in terms of followers is the US with 9.5% of the followers. Hence, information available in OSNs indicates where a given (typically unpopular) content is more likely to be consumed. Note that the more accurate the location information of OSN users is, the higher the granularity and accuracy of the predicted geographical pattern of consumption for Alice's published content will be.

The information regarding where content shared throughout an OSN is likely to be consumed (i.e., where are located the followers/friends of a given user) can be collected by the OSN provider itself, but also by external entities by means of crawling techniques. This information can be shared with Content Providers, CDN operators or ISPs so that they can place content in the appropriate location (e.g., datacenter, cache node, etc.). An important aspect to consider is the level of detail in the information that different players can provide. For instance, Bob may follow Alice but never consumes her shared content. An OSN provider has access to this type of information since it can log all the activities from its users. However, an external entity using monitoring techniques would be more limited. Furthermore, some thresholds need to be configured in order to distinguish users for which using this solution brings real benefits.

Finally, it is worth noting that the end user would benefit from this use case since (s)he will experience lower delays when accessing content and thus enjoy a better QoE.

2.1.1.2 Involved Actors

OSN Providers and/or external monitoring entities

The OSN Provider (or an external monitoring entity) is a key piece in the described scenario because it is the entity responsible for collecting the social relationships information in order to predict where a specific content is likely to be consumed. In particular OSN providers (and/or external monitoring entities) can benefit from the described scenario by offering a new business model (not implemented so far) that allows other Internet Players (e.g., ISPs or Content Providers) to improve their services and reduce their operational costs.

Content Providers

Content Providers are responsible for distributing a large number of content through the Internet. The presented use case would help content providers to place an important fraction of their content (specially unpopular or mid-popular content) close to the end user in caches provided by CDN operators or ISPs. This would lead to a reduction of their transit traffic costs.

CDN Operators

The business of CDN operators is storing and delivering the content generated by Content Providers close to the end users so that both the Content Provider and the access ISPs reduce their transit traffic. Furthermore, users perceive a better QoE due to the lower delay in the access to the content. The proposed use case would help CDN operators to enhance their current algorithms to decide which content to place in which cache/node by leveraging connectivity and relationships information from popular OSNs. As mentioned before, this information is especially relevant to enhance caching of low and mid-popular content.

Eyeball ISPs

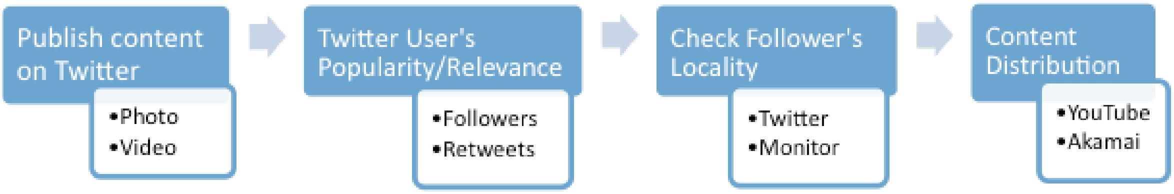
Most of the traffic in the current Internet is associated with content and a reinforcement of this trend is expected in the coming years. Therefore, Eyeball (or access) ISPs aim to store content consumed by their customers within their own infrastructure (e.g. Caches or CDN nodes) or to establish dedicated peering links with the entities providing content. This results in a twofold advantage: the ISP reduces the transit traffic costs because content is locally available and users enjoy a better QoE since the access to content takes advantage of lower delay. The proposed use case follows the described trend and helps ISPs (in collaboration with other entities such as OSN or Content Providers) to leverage social information so as to enhance the existing content placement techniques.

End Users

An important fraction of the current Internet traffic is associated with (entertainment) content distributed to end users (e.g. homes). Therefore, it is important to develop techniques that guarantee a good QoE to users consuming content in the future Internet. The proposed use case pursues this goal by providing techniques that are able to place content close to the end user regardless of the popularity of the specific content. In particular, this would lead to reduce the access delay to the content and thus a better QoE.

2.1.1.3 Use Case Flow-of-Interactions

The next scheme summarizes the different functional steps of the proposed use case:



1. A Twitter user publishes a piece of content (e.g., a YouTube video).

2. The OSN provider (or the external monitoring entity) decides, based on predefined thresholds, whether the user is representative enough (i.e., whether the content shared by the user is usually consumed by a relevant number of followers/friends). If so, we pass to the next steps.

3. The OSN provider (or the external monitoring entity) identifies the location of the user's followers (or friends) and the locations where the content is more likely to be consumed. Note that this task can be executed either online or offline.

4. The OSN provider (or the external monitoring entity) passes this information to the Content Provider (e.g., YouTube) or the CDN operator (e.g., Akamai) that places the content in the appropriate nodes (e.g. data centers, cache servers)

2.1.1.4 Measurable Results of Value for eCOUSIN

eCOUSIN will perform the following tasks associated with the described scenario:

1. eCOUSIN will implement a large-scale measurement infrastructure to collect public information regarding user locations and social relationships in Twitter. This would be a prototype of the aforementioned "External Monitoring Entity".

2. eCOUSIN will define the signalling interfaces between the involved actors.

3. We will use simulations (and may use a real testbed) in which we will emulate the complete scenario in order to evaluate the gains of the proposed solution.

2.1.2 Derived Requirements

This use case requires having a measurement infrastructure able to collect information associated with OSNs users. The information items that need to be collected to efficiently enhance the multimedia content distribution by means of OSNs data exploitation are: 1) the social relationships between users (friendship, friend-follower relationship); 2) the users' location (high granularity is desirable); 3) content published by users in the OSNs.

OSN providers are the most suitable players to play the role of collecting the information. However, external entities can run this service, offering obviously less scalability and accuracy.

In the case where an external entity is the one collecting the information the proposed use case requires to get access to users' information (e.g. location, contact list, etc.). This can be obtained either because users make that information public to everybody, or because the monitoring tool gets the user's approval to access its private information.

Finally this use case requires signalling interfaces between main players in the multimedia content distribution chain. These players are: OSN Operators/External monitoring entities, CDN Operators, Content Providers and Eyeball ISPs. This signalling infrastructure will allow each player to properly use information available in the remaining ones to achieve both a global but also a particular benefit.

2.2 Optimization for the Delivery of Premium Content

It is expected that in the near future the growth of traffic for content delivery on Internet will be more and more due to premium contents. As an illustration Netflix use accounts today for 33% of all downstream traffic in North America during the peak hours between 9 p.m. and 12 a.m. If we consider that premium live contents like concerts and sport games will also soon be delivered on Internet, the question of optimizing the delivery is becoming very complex due to a lot of uncertainties about the number of eyeballs, their location, the size of their screen, ... OSNs can be a powerful source of information to uncover those uncertainties and thus to bring the adequate network resources required to satisfy eyeballs and service providers.

2.2.1 Description of the Use Case

2.2.1.1 Scenario-Base Illustration

Bob and Alice are discussing on an OSN about the forthcoming show of Rihonno, the famous star, that will be broadcasted in Europe on Internet by EU-BroadNet. Alice says that she prefers the low cost offer (1€), Bob the premium offer (3€) with enhanced audio and video quality because he will follow the show from his girlfriend's flat on a terrific new home cinema.

Several OSNs, including Alice and Bob ones, have concluded a deal with EU-BroadNet: all statistical information related to Rihonno fans expected behaviour about the show (number of eyeballs, locations, size of their screens...) are extracted and communicated to EU-Broadnet against payment.

EU-BroadNet will be able to compute all OSN information and to derive the appropriate content storage and delivery capacity required to satisfy the demand of its clients. As we assume that an online European storage market exists thanks to several network operators that moved to SDN architecture, EU-BroadNet can reserve the storage and delivery capacities whenever it wants (but no later than the beginning of the show). A trade-off needs to be found between the progressing accuracy of the data extracted from OSNs as time goes by, and the progressing cost of the storage and delivery capacities for a last moment booking.

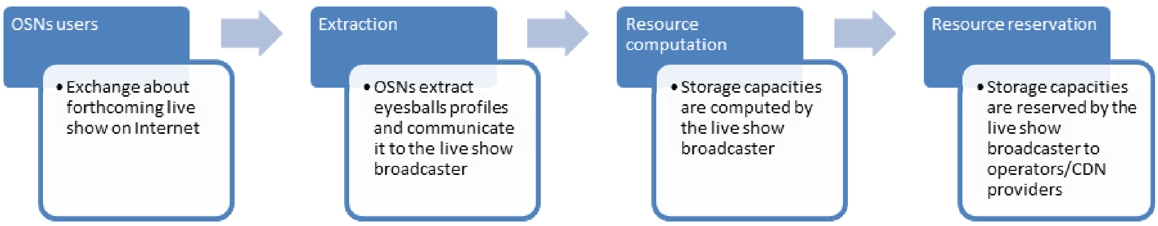
2.2.1.2 Involved Actors

The actors involved in the use case are:

- **Eyeballs** (whatever on fixed or mobile access)
- **Content publisher:** the company in charge of broadcasting live events on Internet
- **OSNs**
- **Network operators/CDN operators**

2.2.1.3 Use Case Flow-of-Interactions

The next scheme summarizes the different functional steps of the proposed use case:



1. Fans discuss on OSNs about how they will follow the show (location, device)

2. The OSN owner (or an external entity) extracts the information related to fans behaviour.

3. Based on one or several OSNs extraction, the broadcaster computes where storage and delivery capacities need to be allocated to fulfil eyeball expectations in terms of QoE.

4. The broadcaster reserves to network operators/CDN providers the storage and delivery capacities:

- either well ahead in time before the show at probably lower price but with less accuracy regarding eyes balls behaviour

- or very close to the show at probably higher storage cost but with the most achieved accuracy

2.2.1.4 Measurable Results of Value for eCOUSIN

To demonstrate the feasibility of this use case, eCOUSIN will provide:

1. OSNs data extraction mechanisms giving eyes balls location and type of screen used to watch the show
2. A data model used to communicate this information to a third party (broadcaster)
3. Resource computation programs giving size and location of storage and delivery capacities at the lower cost, based on a map of storage and delivery capacities with associated cost.

2.2.2 Derived Requirements

The following requirements can be derived from the proposed use case:

- Approximate number of eyeballs for the show can be extracted from OSNs
- Approximate geographic information (city level) of eyeballs can be extracted from OSNs
- Approximate expected QoE of eyeballs can be extracted from OSNs
- All this information can be computed to derive the most appropriate content delivery architecture

3. TIME-UNCONSTRAINED CONTENT DELIVERY USE CASES

This section defines two use cases related to Time-Unconstrained Content Delivery. While the first use case described in Section 3.1 relates mainly to the *download* of content, the second use case described in Section 3.2 relates to the *upload* of content from mobile devices.

3.1 Social-Assisted Time-Unconstrained Content Delivery

This use case focuses on the possibility to reduce costs and energy consumption at the side of mobile end user for the access to data-intensive content, such as videos that the user is interested in, thanks to social recommendations derived from the user's participation in online social networks. The main idea is twofold:

- 1) to proactively prefetch relevant data at the user device in time when Wi-Fi connections are available
- 2) to retrieve content from other users in proximity using device-to-device techniques, such as Wi-Fi ad hoc.

This way, the use of energy and monetary costly mobile connections can be reduced, while providing the user with the same or even better experience.

3.1.1 Description of the Use Case

The follow sub-sections provide further details about the use case about Social-Assisted Time-Unconstrained Content Delivery.

3.1.1.1 Scenario-Based Illustration

A typical user accesses content using a number of different devices, including traditional PCs and laptops but, nowadays, also to a large extent using mobile devices such as smartphones and tablets.

Throughout the day, these mobile devices are connected to the Internet using different networks and connection types. At home, they may be connected to a Wi-Fi network that provides DSL- or cable-based access to the public Internet. On the commute between home and work place, the devices use the cellular infrastructure of a mobile provider. At work or in public places they might be, again, connected to a Wi-Fi network. In comparison to the connection via Wi-Fi networks where flat rates are common subscription models, volume-based subscription models are predominant for cellular networks. While the former one allows to consume an unlimited amount of data (i.e., hundreds of GBs) per month, the latter one usually includes only a few hundred megabytes of data volume at full speed per month and connection speed is throttled to some minimum value once this volume is exceeded, or excess data volume has to be paid extra. In the context of this use case, it is assumed that users have such a volume-based data plan for their mobile devices.

The rather small amount of included data volume as well as the high-energy consumption of the cellular network interface renders the access to data-intensive content (e.g. videos) unattractive. If, in addition, the user moves while accessing the content (e.g. when using public transportation on the way to work), the situation gets even worse as the quality of the cellular connection is usually subject to high fluctuation and connection gaps.

Using a simple but intelligent caching/prefetching (c/p) app for the mobile device, this problem could be mitigated. It could detect when a user is connected to a Wi-Fi network and proactively prefetch data-intensive content (e.g. videos) that the user is most likely going to access in the next hours.

Therefore, also mobility patterns of the user could be used to initiate the prefetching, e.g., just before a user leaves his home towards his working place. The content to be prefetched would be determined based on the user interests that are highly driven by social recommendations as published by the friends and other contacts in Online Social Networks (OSNs).

Due to the prefetching functionality, the mobile device can hold a number of potentially interesting content objects in its cache that the user is able to access during the time he or she is connected via cellular networks or even in cases he or she does not have any connection at all. Using the existing cellular connection, the user, in addition, is able to use functionalities such as rating, commenting, or sharing of the content with his contacts. In summary, the user would experience an enhanced performance when accessing the content, e.g., under the form of better video quality and playback experience, while reducing the monetary costs for the usage of the cellular infrastructure as well as the costs in terms of energy consumption for the transmissions over the cellular network interface of the mobile device. The latter improves the battery lifetime.

In an enhanced version of the scenario, the client application would also leverage the possibility of connecting to other devices using Wi-Fi Ad-hoc techniques to directly transfer prefetched data. In this case, locally popular content or available content from social contacts (i.e. from nearby friends, family members, or other users) would be proactively transferred and added to the cache as well as recommended to the user. This is especially interesting to access recently shared content in situations where there is only a cellular connection or no other connection at all available but content is available from other users' devices located nearby. For this part of the use case, it is assumed that the interest in content of nearby users is sufficiently correlated and the number of devices in proximity is sufficiently high, such that there is a non-negligible probability that the requested content is cached in a nearby device.

Figure 3 illustrates the key players and aspects of the overall use case.

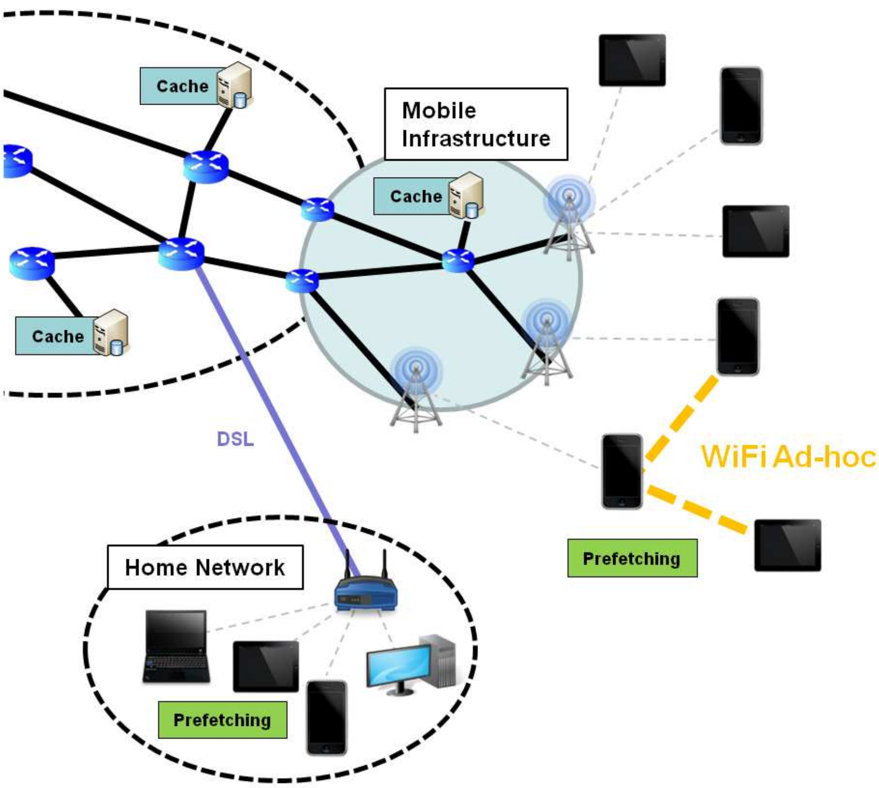


Figure 3: Social-Assisted Time-Unconstrained Content Delivery

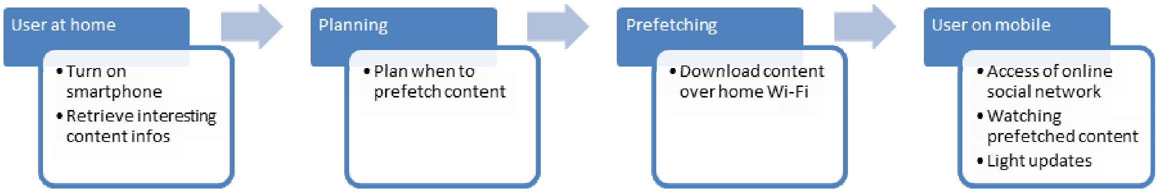
3.1.1.2 Involved Actors

The actors involving in the use case are:

- **End users** (and their mobile devices)
  - Benefit: end users will be able to enjoy content at improved performance and lower costs (in particular costs for mobile connection and battery power)
- **Content provider/OSNs**
  - Benefit: content providers will be able to increase the diffusion of content, and provide their users with a better content experience
- **Network provider** (could allow for light-weight solutions to discover other nearby clients for potential ad-hoc communication)
  - Benefit: They would benefit from an offloaded mobile network

3.1.1.3 Use Case Flow-of-Interactions

The next scheme summarizes the different functional steps of the proposed use case:



1. A user is at home and turns on its smartphone in the morning.
   a. The c/p app installed on the smartphone recognizes the current context of the user, retrieves information on potential content to be prefetched based on information from the online social graph of the user or some external service that provides potentially interesting content.
2. The app plans when to prefetch potentially interesting content before the user leaves the Wi-Fi network.
3. The app prefetches/downloads the content using the user's home Wi-Fi network.
4. The user leaves his home, forcing him to switch to infrastructure-based mobile connectivity (3G/4G).
   a. The user accesses its favourite online social network and watches the already prefetched video content, while more current, lightweight information is retrieved via the mobile network.
   b. For the extended scenario, the user would also try to establish ad-hoc connections to other nearby users to exchange prefetched data that might be relevant for the users based on social properties or locality.
5. Arriving at work, the Wi-Fi network is used again to decide on new prefetching actions.

3.1.1.4 Measurable Results of Value for eCOUSIN

eCOUSIN will perform the following tasks associated with the described scenario:

- eCOUSIN will develop mechanisms that provide enhanced performance for access to data-intensive content and access to content even when no connection to the Internet exists.

- eCOUSIN will reduce the costs due to reduced usage of volume-based subscriptions and reduce energy consumption for mobile devices and cellular network (by avoiding unnecessary costly transmissions), which will increase battery lifetime.
- eCOUSIN will offload the mobile network (if overloaded), so that more capacity and, thereby, better service quality can be offered to all connected users.
- eCOUSIN will provide mechanisms for prefetching at off-peak hours, reducing peak traffic and thus traffic costs for operators/service providers.
- eCOUSIN will use OSN crawled data for particular user and popularity prediction to decide on relevant data to be prefetched.

3.1.2 Derived Requirements

The requirements that are derived from this use case can be classified as follows:

- Social/User Layer (including end-user devices):
    - Geographic information of end users is known (GPS functionality of smartphones, known locations of access points and base stations the user is connected to, …).
    - Users have volume-based data plans for their mobile devices.
    - User interest is driven by social recommendations and users are interested in consuming data-intensive content (e.g. videos) in times when they do not have a Wi-Fi connection.
    - Mobility patterns of users and social properties of content need to be used to infer when to prefetch which content objects.
    - c/p app has access to information on content relevance according to its social properties (access to OSN information/prediction service).
    - The user device has enough storage to keep a number of prefetched content objects (rough estimation: several hundred megabytes).
    - For the enhanced scenario: The mobile devices are able to efficiently communicate in an ad-hoc manner (support by hardware and software) and interest in content is sufficiently correlated and the number of devices is sufficiently high, such that there is a non-negligible probability that the requested content is cached in a nearby device.
- (Access) Network Layer:
    - The user is able to access the Internet at his/her home or other places via Wi-Fi with a certain minimum connection speed to allow for an efficient and timely prefetching of content.
    - Nearby clients to potentially exchange prefetched data using Wi-Fi Ad-hoc communication need to be detected.
    - A user’s regularly used Wi-Fi networks (e.g., home, office) need to be identified.
- Content Delivery:
    - Content (e.g. videos) is available in a quality suitable for the mobile device (this requirement also correlates with the first one).

- Access to content can be adapted to make use of pre-fetched or cached copies. This may require cooperation of the content provider (e.g., in case of an SSL encrypted connection directly to the content provider, where an operator could not intercept and redirect traffic).

3.2 Mobile Content Uploading

The interaction between OSNs and content consumption should not only be focused on the delivery; ingestion of contents must be also addressed as proposed in this use case that is complementary to the one proposed in Section 3.1.

3.2.1 Description of the Use Case

3.2.1.1 Scenario-Based Illustration

Bob is living in Southampton; he is spending his holidays in Vancouver. As he used to do during holidays, he will shoot a lot of videos from his smartphone that he wants to share with his contacts in his various social communities (friends/family, sailing club, charitable association). Bob is using an app from his cloud service provider and feeds it with all the videos he made and wanted to share. Then he shares it as usual on his own OSNs/portals but the uploading is controlled by the app. The cloud service provider will:

- (MANDATORY) upload each video at the most appropriate time according to bandwidth radio capabilities (cellular + Wi-Fi), cost, smartphone battery state, Bob's OSNs contacts location (local time) and habits (time/frequencies of connectivity on the OSN), etc.
- (OPTIONAL) convert video on all required formats according to Bob's OSNs contacts request,
- (OPTIONAL) store each content at the most appropriate place according to Bob's OSNs contacts' location.

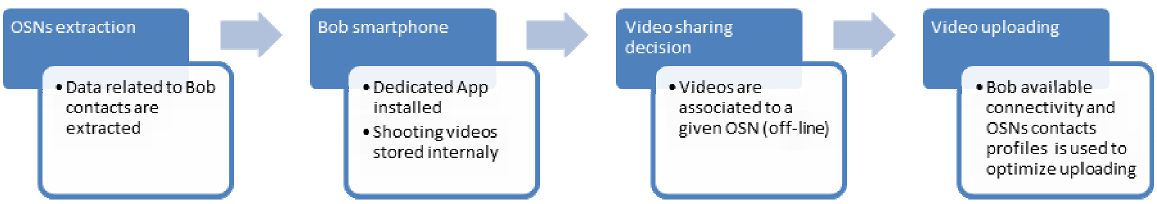
3.2.1.2 Involved Actors

The actors involving in the use case are:

- **End user** with his smartphone
- **End user cloud service provider:** the company in charge of storing user contents in the cloud
- **End user OSNs**
- **End user mobile operator** (and its roaming partners)

3.2.1.3 Use Case Flow-of-Interactions

The next scheme summarizes the different functional steps of the proposed use case:



1. The OSN owner (or an external entity) has extracted the information related to Bob's contacts behaviour (location, time/frequencies of connectivity on the OSN)

2. Bob has installed on his smartphone the dedicated App of his cloud service provider; the videos he shot are stored internally.

3. Bob decides for each video on which OSN it will be published and how it will be published (comments, priority); this step can be done off-line

4. Once the smartphone has connectivity (cellular or Wi-Fi) the App determines if a video needs to be uploaded; decision is based on contacts profiles behaviour, cellular cost, video priority...

3.2.1.4 Measurable Results of Value for eCOUSIN

To demonstrate the feasibility of this use case, eCOUSIN will provide:

1. OSNs data extraction mechanisms giving Bob's contacts location and habits.
2. A data model used to communicate that information to a third party (cloud service provider).
3. Optimization programs aiming to reduce the cost of uploading.

3.2.2 Derived Requirements

The following requirements can be derived from the proposed use case:

- Extract from OSNs the location and behavior (time/frequencies of connectivity on the OSNs) of the contacts of a given user,
- Extract user videos classification (targeted OSNs, priority…),
- Extract user information related to its mobile data plan (including roaming conditions), Wi-Fi availability, battery state...,
- Derive for each video the most appropriate access points and moments to upload it.

4. USER-GENERATED CONTENT DELIVERY USE CASES

As we have stated in the introduction of this document, the proliferation of user-generated content will very likely affect the multimedia content traffic in the Internet. Most of those user-generated content belongs to the long-tail within the content popularity distribution, thus being expensive to be distributed around the world given the low number of users that will access it. Therefore, in this section we present two use cases that address the distribution of user-generated content from two different perspectives. The first use case proposes the use of personal sharing clouds, while the second one exploits the concept of Information-Centric Networking (ICN) to improve the distribution of user-generated content.

4.1 Personal Content Sharing Clouds

Several network operators, including Telecom Italia, have been trying to market Internet-based multimedia content distribution platforms. Typically, the main objective has been to bridge the television world with the Internet one, towards novel Social TV paradigms. Managed network approaches, based on infrastructure-side components hosting and delivering multimedia content, have come up against major issues in trying to be commercially successful and to catch final users' expectations to both consume professional multimedia content and share user-generated one. Based on such experience, players like Telecom Italia have been investigating alternative approaches. For example Telecom Italia aims at evolving Cubovision, the Telecom Italia IPTV offer, to support not only infrastructure-based content delivery, but also to locally store private content and share it over the Internet with remote people linked by social relationships.

The integration of social networking with decentralized communication, such as Peer-to-Peer (P2P), has the threefold objective of introducing a novel paradigm of Social Content "Crowdsourced" distribution chain, providing higher degree of data ownership to final users while incentivizing the adoption of large bandwidth fixed connectivity provided by ISPs. The proposed solution also integrates widely spread off-the-shelf Universal Plug and Play (UPnP) devices/services, taking advantage of Digital Living Network Alliance (DLNA) AV Media Controllers and Servers already available in users' home networks.

4.1.1 Description of the Use Case

To better understand the target use case about Personal Content Sharing Clouds and how it achieves these goals, we consider the following scenario (see Figure 4).

4.1.1.1 Scenario-Based Illustration

A. Alice has a DLNA-enabled Smart TV she usually exploits to browse the Internet, e.g., to read news and access Facebook. Bob has a Telecom Italia's Cubovision (or any other similar set-top box) where he stores his own pictures and music files, to be available on other remote nodes via DLNA AV Media Server. Smart TV and Telecom Italia's Cubovision devices are deployed in Alice and Bob home networks respectively;

B. Alice and Bob meet on their summer holiday; when they come back home, they become friends on their preferred Online Social Networks (OSNs), eventually exploiting the Federated Social Networking Standard Implementations (FSN). The creation of a new social relationship triggers the automatic generation of a communication link among Alice's and Bob's gateways, already deployed in their home networks to provide Internet connectivity;

C. Once friends, Alice uses her Smart TV to access via DLNA the pictures, videos and music that Bob stores on his Telecom Italia's Cubovision and has tagged as "shared with friends". Alice accesses

Bob's Telecom Italia's Cubovision without any specific goal, just for her own curiosity to see pictures and videos of her new friend, e.g., to view photos of his previous vacations or check his music selection;

D. After a couple of days, Bob saves his holiday pictures in the "Summer 2012" directory on the Telecom Italia's Cubovision and then notifies via FSN his friends (including Alice) of the new photo album. Alice clicks on the FSN post exploiting her Smart TV and accesses via HTTP the whole directory content;

E. Bob takes a new video with his smartphone. The FSN client automatically uploads the video on his Telecom Italia's Cubovision at home and notifies his friends of the new video location. Alice can access that specific video stored in the Bob's home private network simply clicking on the new post.

F. Finally, Bob accesses a YouTube video from his mobile and decides to share it in using one or more OSNs. Automatically, this video is downloaded and stored in his Telecom Italia's Cubovision system. While Alice is browsing her OSNs contacts she clicks on the YouTube video Bob shared, and the Content provider (i.e. YouTube) redirects Alice to Bob's Cubovision that serves the video to Alice.

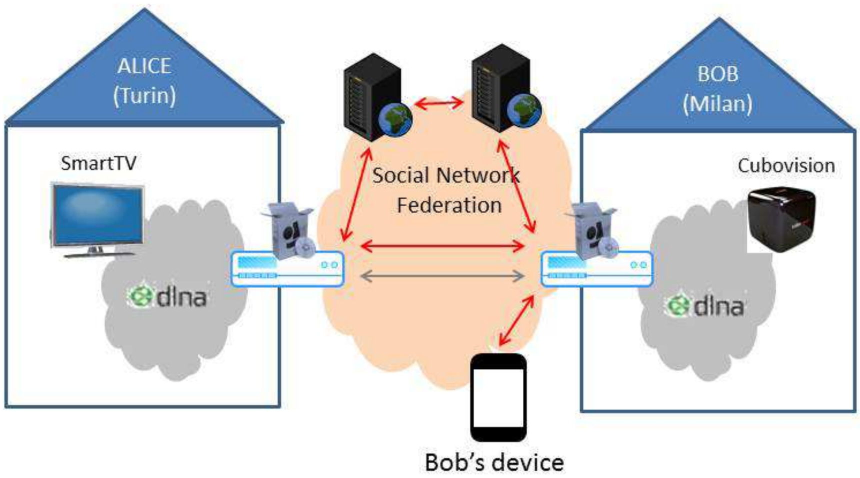


Figure 4: Personal Content Sharing Clouds

It is worth noting that Bob shares his resources while preserving data ownership. In fact, while Bob notifies his friends of new pictures and videos via FSN posts, actual resources are stored in his own Telecom Italia's Cubovision. If Bob desires to completely stop sharing his resources, he can simply turn off his Telecom Italia's Cubovision (or disable the DLNA AV Media Server).

4.1.1.2 Involved Actors

End users experience of social mediated multimedia content sharing is critically impacted by this use case. Users are permitted to share their multimedia contents in a P2P manner without having to upload them to an online social network server. We believe this capability will improve their willingness to share more contents given that they would feel like having them under closer/concrete control. Likely they will be willing to produce more user-generated contents in order to share them by means of several devices such as PDAs and smart phones besides set top boxes. Furthermore the system will be able to compute their social relationship and their consent to share even if two users are not subscribed to the same Social Network.

Telco Operators are impacted by the scenario because likely end users will be demanding for higher bandwidth capabilities in order to be able to directly serve their friends with their content stored locally. Telcos have already deployed an enhanced network infrastructure which is already capable of providing an improved connectivity to end users. eCOUSIN's scenario is definitely targeting a bandwidth hungry application case which will facilitate the exploitation of these improved network capabilities.

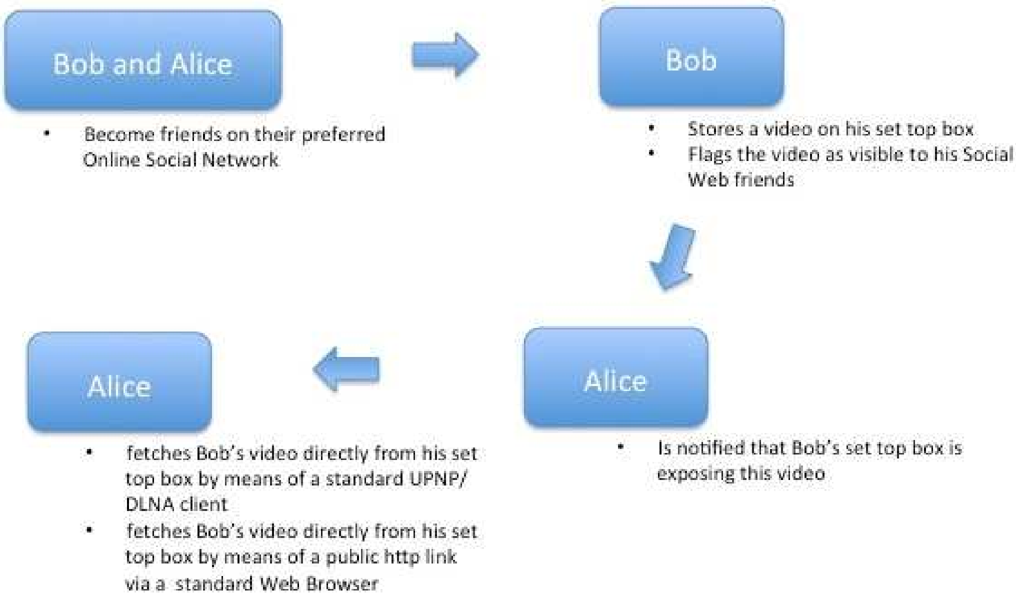
Content providers are also involved in the scenario, which will facilitate a novel crowd-sourced content distribution chain. We believe this would also encompass professional contents, subject to copyright, therefore content providers should be able to figure out novel mechanisms to manage content copyright for those cases in which contents are proxied from one friend to the other.

ISPs are also actors and impacted by this scenario. A critical requirement for this scenario is that each node, pertaining to a user, must be assigned with a public IP address for the system to properly scale. Some ISPs are already using NAT technologies due to the lack of availability of IPv4 addresses, probably some NAT transversal mechanisms will have to be taken into account.

Online Social Network providers, in order to be able to offer the possibility to their customers, to join this novel content distribution chain should take into account to comply as much as possible to existing Web Identity Federation recommendations.

4.1.1.3 Use Case Flow-of-Interactions

The following figure summarizes a basic flow of interactions, which should be supported by eCOUSIN.



4.1.1.4 Measurable Results of Value for eCOUSIN

The envisioned scenario greatly increases egress Internet traffic, since access to shared resources triggers large file transfers from Bob's home network to his friends' ones. Thus, there is the need for large bandwidth (eventually symmetric) ingress and egress Internet connectivity, providing high data throughput not only from the Internet to home networks but also from home networks to the Internet. Moreover, the use of peer-assisted video delivery will help content providers in serving unpopular content usually located in the long-tail within the content popularity distribution. Serving this unpopular content is usually costly for content providers since they need to spend their resources for just some few views/downloads of a particular content Finally, users can fruitfully exploit already deployed legacy UPnP devices/services with new and enhanced features, enabled by our novel solution: Alice and Bob can exploit their Telecom Italia's Cubovision and Smart TV in a novel and powerful way, thus also increasing the perceived value of their already available devices.

The positive aspects of the scenarios will be the findings in the areas that are going to be addressed for this to become deployable, in fact the main challenges to be mastered are:

1. Federation of identities on the Web;
2. Social federation of Smart Environments;
3. Security and privacy;
4. Usability and configurability;
5. Peer assisted protocol design.

With respect to 1, eCOUSIN will provide APIs retrieving and merging information coming from different Social Networks achieving a twofold goal. On the one hand, Cubovision users are not bound to subscribe to the same Online Social Network provider to exchange multimedia content; on the other hand, it avoids to further increase fragmentation creating yet another community of Cubovision users. This set of APIs will also ensure that Cubovision users' micro-communities will still be able to interact with one another while easily sharing multimedia content also with traditional Online Social Network users.

With respect to 2, eCOUSIN will identify protocols and mechanisms supporting content sharing on top of overlay networks. We expect our solution to be able to create overlay networks composed of a) isolated Smart Environments related to different private spheres and b) dynamically federated Smart Environments driven by inter-user social relationships.

With respect to 3, eCOUSIN will provide mechanisms to ensure encryption, traceability and certainty of content origin and will consider users' willingness, hardware capabilities and bandwidth availability to adopt the right trade-off between quality of experience and security.

With respect to 4, eCOUSIN will ensure cross device (Smart TV, smartphone, tablet) user experience by means of intuitive user interfaces and smart content filtering based on users' preferences and/or profiles. On the one hand, we will support traditional multimedia sharing protocols end users are already used to, such as HTTP and UPnP/DLNA, also in the challenging scenario of overlay networks composed of federated Smarts Environments

With respect to 5, eCOUSIN will provide algorithms to decide when a user can retrieve a particular content from an OSN friend that shared that content in its OSN wall. In addition, eCOUSIN will define a suitable framework where OSNs operators, content providers and ISPs collaborate together so that all of them will take advantage of the peer-assisted video delivery.

4.1.2 Derived Requirements

This use case establishes a set of requirements for its successful development. First of all, this scenario mainly relies on commercial DLNA enabled devices such as smart TVs, smartphones, tablets, PDAs or commercial Internet enabled multimedia servers (i.e. Telecom Italia's Cubovision) that are also known as set-top boxes.

UPnP and DLNA have to be accounted as reference protocols because they are widely spread, standard and supported by most of the commercial multimedia devices. Of course these protocols have been specified to enable zero configuration and work properly in Local Area Networks. eCOUSIN will have to study and implement the necessary extensions and enhancements for them to be able to be bridged over the Internet.

Furthermore, since the set-top boxes will deliver multimedia content to third parties they will require high-speed uplink connections so that users retrieving content from other set-top boxes do not experience any performance downgrade as compared to retrieve the same content (if applicable) from CDNs, cloud services or content providers' servers.

We also mentioned in the above paragraph the critical requirement to have a public IP address for each sharing node. In case some of these nodes will be placed behind a NAT, eCOUSIN should be able to put in place some NAT transversal functionalities for them to be able to join the ecosystem properly. Of course we will foresee contexts in which a single user will be sharing contents distributed over different devices e.g. set top box and smartphones, in this case the system will require at least one of them, e.g., set top box to be assigned with a public IP address.

In order to perform peer-assisted video delivery content providers, ISPs and OSNs operators may need to share some information to efficiently redirect end users to retrieve the video from a suitable set-top box.

The proposed scenario requires that set-top boxes will remain always on as they will become the main storage points for multimedia content. However, this requirement is quite reasonable since end users use to keep this equipment switched on most of the time.

As said a further requirement is that eCOUSIN implements all the available recommendations in terms of APIs and protocols to fully support Input/Output interoperability with Online Social Networking initiatives. It will therefore be necessary that at least a minumum set of information related to identity and profiles will be accessible from the system regardless of the Online Social Networking server which actually stores them. On the contrary it will not be mandatory to have an account of a OSN to be part of the ecosystem.

4.2 Information-Centric and Social-Driven Content Delivery

In response to the needs of emergent bandwidth-consuming applications, mainly driven by the popularity of video content and social networks, several projects and initiatives for the Future Internet have started over the last years to envision a clean slate foundation of the Internet by moving away from the current endpoint-oriented approach to promote information at the center of networking design considerations. As a result, several Information-Centric Networking architectures have been proposed [7] targeting a well-defined set of architecture invariants, which are persistent/unique naming for content objects independently on location, efficient content distribution and discovery through name-based routing, in-network caching, security and mobility.

As the ICN concept redefines interdependencies between network entities to make them depending more on naming and semantics of pieces of content, in eCOUSIN we aim at exploring the benefits of using ICN as a network layer for content delivery in OSNs while also taking into account the relevant interactions inside the social graph resulting from OSNs. With a use case based on user-generated content sharing, we show in the following sections the requirements that need to be adhered to design an ICN architecture to improve content exchanges between users in an OSN.

4.2.1 Description of the Use Case

This section describes the use case “Information-Centric Networking for social-driven content delivery”, based on the following scenario.

4.2.1.1 Scenario-Based Illustration

More and more OSNs propose now to share your music with your friends: whatever you are listening on your computer or smartphone, your friends can listen with you at the same time. Typical examples of OSNs offering this functionality are Facebook or Russian VK application. We propose to go further with an extension towards video live streaming over OSNs, with an optimized delivery of content thanks to the ICN features, linked with the OSN ones. The proposed use-case is related to the sharing of content by one end user with his friends. Bob has just come back from his holidays in Japan, and wants to tell the trip to his friends on his favourite OSN. What can then be more pleasant than inviting his friends for a special night on the OSN to tell his wonderful tribulations in Japan to them? Instead of just sharing his photos with his friends, which is somewhat annoying, Bob broadcasts on the OSN the videos he took in Japan, so that his friends can watch together with him these videos in live streaming. Sometimes, to give some details, Bob can also comment some videos either by chat or by a voice or video conversation. Bob can also show some souvenirs he bought in relation with the live video streamed on the OSN. Latecomers or friends who cannot attend the great show can always watch the videos later on demand, whenever they have time.

The use-case can be detailed with the following steps:

A - Just back from his holidays in Japan, Bob sends an invitation for day D and time T, to his friends on his favourite OSN, for telling them his trip.

B – Receiving the invitation on the OSN, his friends are aware of this event and can save the date.

C – The friends being active on the OSN, the application can remind them about this event a few minutes before (otherwise the friends should have it in mind or should rely on another reminder application and connect to the OSN).

D – The interested friends declare their wish to attend this event, to receive the contents and possibly contribute, ask questions, send comments.

E – At day D and Time T, Bob starts the streaming of his video and can add some other information in the chat box or others.

F – All interested friends receive the stream and enjoy the story told by Bob. Some can contribute by exchanging information, souvenirs, and questions.

G – While being transmitted on the OSN, the stream is recorded in order to be played later (as a VoD content) for absent friend.

H – At the end of the event, Bob stops the streaming (the recording of the content is then stopped and saved by the Service Provider) and everyone can leave the virtual room.

4.2.1.2 Involved Actors

The actors involving in the use case are:

- **End Users** share their personal content (photos, audios and videos) on the OSN. User-generated contents are directly streamed towards any user interested in (this means that users are also content providers), and while sharing their contents, users have the possibility to have these contents stored by content service providers that are then in charge of broadcasting later at user's request (similarly to YouTube, replay TV services or VoD). As a consequence, end users are also consumers, requestors and receivers of content offered by other users.
- **Content Service Providers, CSPs,** (hosting user generated content among others) offer media streaming services, such as live or on-demand services, based on content provided by end users. Multiple types of content services providers may be distinguished depending on the business philosophy, such as Internet TV broadcasters (offering personal TV to users), ISPs offering quad-play services or content aggregators e.g. YouTube, Netflix.
- **ICN Service Providers, ICNSPs,** deploy ICN infrastructures (including cache or replica servers) to provision connectivity between Content Providers and users. In particular, CSPs use ICNSP services for distributing their content (including user generated content) to End-Users.
- **Transit Internet Service Providers (ISPs), or other Network Service Providers (NSPs),** are connected to ICN service providers for peering or transit services for large-scale Internet connectivity.

4.2.1.3 Use Case Flow-of-Interactions

The following Figure5 depicts the high-level architecture for the use case. Content objects (which represent pieces of user generated content) are named by following a naming scheme, which also integrates semantic information about content (type of content, author, title, etc.) as well as social information to reflect possible social-content interdependencies (e.g. this content is shared with which friends or with any user, or the content is only for private use).

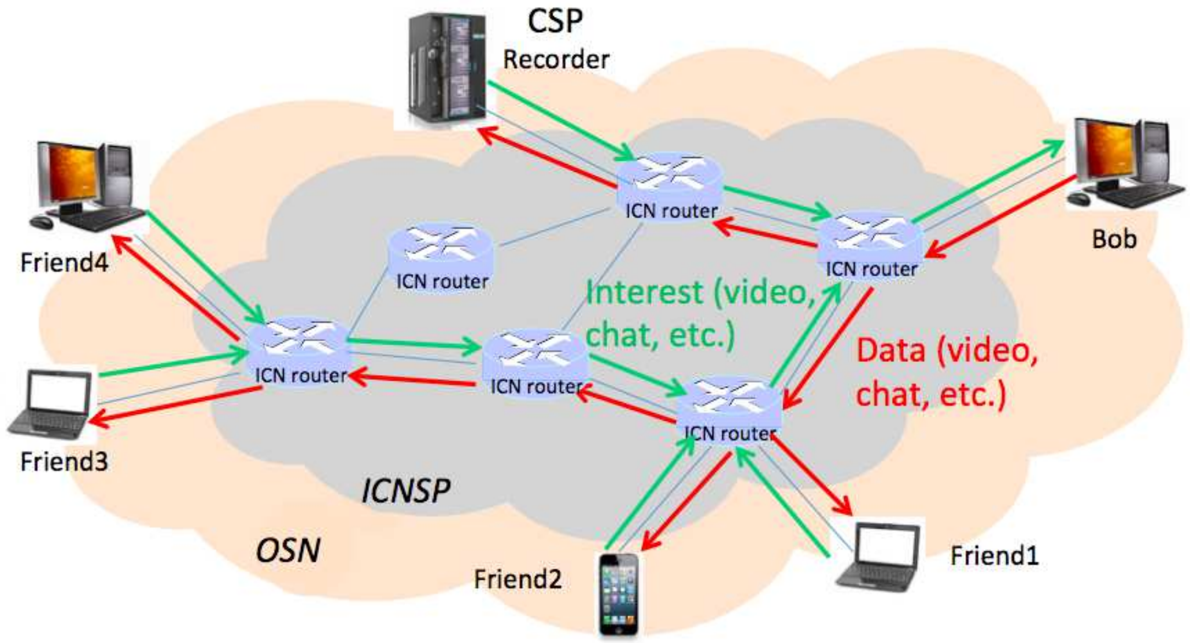


Figure 5: Information-Centric and Social-Driven Content Delivery

End-users (or end-user applications) find media services of their interest on the OSN through a Name Resolution Service. This service offers two capabilities. The first one is mechanisms for searching content based on human-friendly meta-descriptions, and the second one allows to resolve these descriptions into appropriate names used by the ICN forwarding plane to route/find content objects.

The end-user request, expressed as an **Interest** message for a content object, is routed to the origin Content Service Provider server, unless a copy is found somewhere in a cache along the path. A **Data** message carrying the queried content object is then forwarded back to the user along the reverse path.

The ICN Service Provider is responsible for the reliable, timely and secure delivery of the requested media streams, based on its ICN infrastructure, which integrates in-network caching.

Finally, the Content Service Provider can also host user-generated content (such as Bob's videos in our scenario). Interest messages are then issued from the CSP recorder server to get UGC to be stored.

4.2.1.4 Measurable Results of Value for eCOUSIN

With this use-case, we propose to highlight the benefits of an ICN architecture for improving overall performance of user generated content (UGC) dissemination and delivery, for users as well as for network or content providers. This new ICN architecture, with its native in-network functions (such as multicasting, caching or chunking), has many advantages for live content, and the social links derived from the OSN have a direct influence on the network delivery efficiency. ICN nodes in the delivery path can cache and forward content to subsequent end users (the friends watching the videos) requesting the same contents. All the content the user can publish will be disseminated in different caching locations in the network as soon as her/his followers start to update their timeline. Thus for socially driven contents which are likely to be more geographically local and a short tail distribution, end users in the same community can request the same content from the closest locations, enabling the network load to be greatly reduced, starting from the operator network. The

ICN node will receive the content only once from the upstream node/server and deliver it downstream to all interested entities. There is no need for all end users to fetch it from the original content server.

Content Providers get thus benefits for improving delivery performance to users and resource consumption since they can base on the ICN infrastructure to provide native pseudo-multicasting and improved caching mechanisms not only for popular but also for mid-popular content (socially driven content within a community). In return, end users can benefit from enhanced Quality of Experience (QoE) when consuming content originated within social network communities, since the content is disseminated in various caches closer to end users and can be delivered more rapidly. Regarding ISP, transit traffic is consequently reduced with copies of content near users' location.

The positive results of an evaluation reported in [8] for deploying Twitter over an ICN infrastructure, compared to the current classical IP-based delivery or the CDN-based solution, back up our analysis about the aforementioned measurable results of value of our social-aware and content-centric delivery use case.

4.2.2 Derived Requirements

We first present the functional requirements that are derived from the use case.

- As the technological and cost barriers of content production continue to get lower, User Generated Content (UGC) based applications grow rapidly, End-Users can therefore become also CPs.
- As End-Users and CSPs are customers of ICN Service Providers, their interactions should be based on respective SLAs, which specify various aspects related to large-scale content distribution, such as content scope (e.g. limiting content access to only authorized users) and geo-restrictions, content integrity and availability, performance (delay, network load, QoE for end users, etc.), mobility and security.
- End-Users should all be reachable, being clients of the same ICNSP or different ICNSPs. They should then be gateway or transit network to make the connection between different ICNSPs.

We list now the technical requirements from the use case, which address several aspects that will be studied in the eCOUSIN project, related to the ICN-based delivery of OSN contents.

- Naming scheme for addressing content objects generated in the OSN. According to the ICN principles, it is important to define a coherent and efficient naming strategy for contents shared in the OSN. The naming can take into account social links within OSN for more efficient delivery.
- Name-based forwarding plane with an in-path resolution of content names into locators (hosts storing content) and a native integration of content caching functions. This means that the request for a content object (or chunk), using an Interest message issued by the end user, is forwarded towards the origin content provider, unless a copy of this chunk is cached somewhere along the path, and then a Data message containing the queried chunk is transferred back to the requestor along the reverse path. Moreover, the built-in caching capability should allow each forwarding node to store copies of chunks in forwarded Data messages for future use based on a defined policy.
- Media streaming transport protocol over ICN network layer. The ICN transport layer being different than the IP one, the streaming mechanism should be adapted to fit with ICN messages and way of working. Typically, the split of content into chunks, naming of chunks and exchange (and possible caching) of them change the way to stream contents.

- Traffic engineering for efficiently delivering OSN content objects, with routing management and cache strategies to meet the time requirement of live content streaming. It can exploit social information related to users' relationship, users' location, and users' interest for a better efficiency of content delivery.
- Adaptation of the OSN application for the ICN protocol. The network transport being ICN, the OSN application should be adapted to "speak" the ICN protocol and mainly include Interest and Data messages, as used by ICN. This would also allow to evaluate the complexity of having ICN-compliant applications.

5. REQUIREMENT SPECIFICATION FOR ECOUSIN

This section defines the requirements that are necessary to successfully develop eCOUSIN. We could follow either a top-down approach or a bottom-up approach. In the first one, we should have defined abstract requirements that later could be narrowed down to match the requirements established by the different use cases introduced in this document. The main problem of this approach is that usually abstract requirement definition is too far away from practical issues, so that you need to later find uncomfortable ways to match those abstract requirements into the use cases that are actually addressing real and practical scenarios. In contrast, by choosing a bottom-up approach, we rely on the use cases to extract requirements imposed by practical considerations included in those use cases. Those requirements allow us to extrapolate more abstract concepts that can be extended to a number of other practical scenarios, but at the same time keep a practical consistency. Therefore, based on these arguments, we decided to implement a bottom-up approach to define initial requirements for eCOUSIN.

Moreover, we organize requirements by following a layered scheme that will help us to, later in the project, derive technical solutions for designing and implementing a network architecture for enhancing content distribution with social information. We envision having three main layers in our architecture, which orientate our requirement specification and classification:

i. Social-Content Interdependencies layer, which integrates the social and content information extraction plane on the one hand, and the social and content information management plane on the other hand)
ii. Content Dissemination layer
iii. Network layer.

5.1 Requirements for the Social and Content Information Extraction Plane

To be able to support the use cases described in the previous sections, some generic functionalities are required when collecting social and content information within OSNs to uncover the interactions among users and the way content is exchanged or shared as a result of interdependencies within the social applications.

First, most use cases require that the "relationship graph" between friends can be retrieved either in one OSN or aggregated over several OSNs (via FSN techniques). In other words the "friends" list for each person needs to be identified.

Second, for most uses cases, "individual user interests" of OSN users need to be obtained. These user interests have to be seen in a broad sense. For instance a preliminary list of interests could include , but is not necessarily limited to, preferences for types of content (e.g. YouTube videos, friends pictures, etc) and preferences for consuming content at specific hours in the day or at specific locations. This type of information can be implicitly obtained (via user profiling engines) or by a user explicitly stating his or her interest in some piece of content that this user makes publicly available.

Third, the "user location" and/or "mobility pattern" (i.e., at what moment in time the user is likely to be connected to which network (e.g. fixed, Wi-Fi, cellular network) together with the "congestion state" of the network need to be acquired.

Fourth, some use cases require to know in advance the context of content (also referred as the content scope) to define premium content or usage restrictions such as non-cacheable, authorization access only for a specific circle of users/friends, etc.

All the use cases discussed above combine these four functions in their own way. The most common way (see use cases 2.1, 2.2, 3.1, 4.2 and partly use case 4.1) is to combine individual user interests and location with relationships graphs to predict where or when content will very likely be consumed. In this way, content can be pre-loaded to an optimal location nearby the users or be pre-fetched by an end-user device. Some use cases (3.1, 3.2) take the actual or predicted network state into account: the advanced version of use case 3.1 relies on the (predicted) proximity of users, while use case 4.2 takes network congestion into account. Finally, some use cases prevent some parts of the content to be locally stored either for privacy reasons (see use case 4.1) or geo-restrictions (see use case 4.2).

5.2 Requirements for the Social and Content Data Management Plane

The social and content data management plane is responsible for gathering, analyzing and statistically correlating both social and content data collected from OSNs and content providers, in order to derive relevant parameters that can be later exploited in the content dissemination layer. For instance, as described in the aforementioned section, it is important to exploit social-driven content popularity patterns (used in the use case 2.2 on the premium content delivery) and the content geo-popularity metrics that help to predict future content diffusion patterns ahead of time (as relevant to the use case 2.1).

Note that the expected amount of data to analyze is huge and in some cases will not be computationally affordable. This justifies then the use of aggregation and abstraction methods (e.g. based on data mining techniques) for processing with a sub-linear cost in terms of computational resources and the implementation of a federation plane between heterogeneous online social networks, as required in the use case 4.1, to obtain aggregated information from different sources.

Finally, for name-base content delivery, as proposed in the use case 4.2, social information extracted in OSNs and content applications will also be investigated to define naming schemes that integrate social interactions between content objects. Social-enhanced naming allows then asserting routing hints that can improve content delivery in an information-centric networking architecture. This means in particular that the forwarding nodes can exploit a set of routing meta-information (hints) to discover the path to the destination hop-by-hop and on the fly.

5.3 Requirements for the Content Dissemination Layer

The content dissemination requirements provide the framework to efficiently distribute contents with intelligent content pre-fetching and caching systems. The algorithms for content placement (including content selection, replica server placement and content outsourcing) and cache management (i.e. appropriate in-network cache distribution and caching techniques) will then be based on the exploitation of information coming from the social and content data management plane (as outlined in Section 5.2). The objective is to reduce network load, operational and energy costs and at the same time improve users' experience when consuming content.

In order to enable content prefetching based on social recommendations or other policies (as in the use case 3.1) or to let users the possibility to search for content matching their interest (as relevant to the use case 4.2), a content discovery service needs to be implemented. This means that some of the network nodes will further act as information processing nodes at the higher level of the social and content data management plane, holding databases of aggregate information and supporting sophisticated search functionality over information stored at this level of the plane. Each search node can receive data in real-time from associated network nodes based on subscribed interest, which is subsequently mapped onto an information model and stored locally. Search functions, invoked from the management applications through query processing, will be executed as distributed algorithms on the graph of search nodes, and appropriate models for describing content and related services need also to specified.

For the information-centric networking use case (Section 4.2), the content discovery service also provides capabilities first for searching content based on user-friendly meta descriptions, and secondly for resolving these descriptions into appropriate names used by the underlying ICN network to route content.

5.4 Requirements for the Network Layer

Considering the proposed use cases, the network layer will integrate a set of network tools capable of:

- Automatic topology discovery taking into account a large number of networking solutions (based on CDN, ICN, wireless networking environments including 3G/LTE, Wi-Fi, Ad-hoc Wi-Fi, etc.). Special considerations relevant to mobile data offloading and ad-hoc communications will also be investigated to enable time-unconstrained content delivery that allows prefetching or uploading content in the case when high bandwidth and low cost access networks such as Wi-Fi hotspots are available (cf. use cases in Section 0);
- Traffic monitoring for collecting data related to social and control information extraction plane (in line with the requirements of Section 5.1), as well as for analyzing network status to reveal changes in network conditions (available bandwidth, delay time, flash crowd, etc.);
- Efficient mechanisms for identifying and orchestrating the required in-network processing functions for matching content with users' interest, content search, resolving content names into physical host identifiers for ICN routing, etc.

Note that Information-centric networking solutions will be taken into account as further evolutionary development of the network layer to enhance content-aware routing, mobility and caching in content-aware routers with social and content information. Relative to the use case 4.2, the ICN transport protocol should be efficiently designed to initiate and control users' requested media streams from the content provider and the different network caches, while taking into account the possible issues of the underlying network such as detecting and adapting the underlay congested path to rapidly maintain an acceptable quality of service for live streaming services.

For efficiently delivering content in collaboration with the requirements for content dissemination as described in Section 5.3, Traffic Engineering (TE), which consists of the cache and route management and control plane, is also an important part of the network layer to provide the necessary intelligence for collaboratively:

- Determining and dynamically managing, based on social information and actual network usage and load conditions, the placement of caches in the network, the policies that caches use for content selection, and the placement of monitoring and/or processing functions in the network;
- Optimizing dynamic cache-aware routing, e.g. to both minimize delays for content delivery and reduce transmission bandwidth by properly routing requests to the best possible cache, rather than to the original server storing the content.

Specific solutions for the optimization of cellular mobile wireless networks will finally be addressed to address issues related to temporal service disruptions and degradations (e.g., such as low data rates and high latencies) by integrating social information with planning and management decisions.

5.5 Synthesis of the requirements

As an overview, the requirements extracted from all the uses as outlined in Sections 2, 3 and 4 can be summarized and organised as described in the following table:

Architecture Layer	Requirements
Social and Content Information Extraction Plane	- Collecting information associated with OSNs' users (social relationship, user interests, user usage patterns, user location, mobility pattern, etc.) - Collecting information about content exchanged/shared between users (content properties, content scoping, etc.)
Social and Content Data Management Plane	- Data aggregation and abstraction -Data mining to highlight useful information from collected social and content data for content dissemination and delivery -Federation between heterogeneous social networks - Understanding the naming scheme in ICN while taking into account the social and content information collected in the OSN
Content Dissemination Layer	- Social-enhanced cache strategies - Social-enhanced content placement strategies -Content Search algorithms based on users' interest
Network Layer	- Network topology detection - Social-enhanced content delivery specific to the network architecture (CDN, ICN, mobile networks, adhoc Wi-Fi, etc.) - Time unconstrained content delivery algorithms - Adhoc Wi-Fi content exchange - Network availability detection - Load and congestion monitoring - Mobile device proximity detection

And last but not least, an important requirement that we have to consider from now is the necessity that all architectural solutions that we will design need to have backward compatibility with current services relying on content distribution.

6. PRIVACY CONSIDERATIONS

We wanted to finish this document by briefly including some privacy considerations that need to be taken from the very beginning of this project. First of all, this consortium is in contact with the different national authorities that are in charge of controlling the implementation of good practices for those cases where Internet users' information is being used and analysed. For instance CNIL, the French authority, is aware of the data collection processes that will be performed during eCOUSIN project. Furthermore, we have crosschecked that eCOUSIN data collection techniques are aligned to German regulation. Therefore, from the very beginning our intention is to be very sensitive to potential privacy issues, and our goal is to find the right balance to go as far as possible in our research to achieve the established goals, and at the same time being compliant with privacy regulations.

CONCLUSION

The present deliverable D2.1 is the first technical deliverable of the eCOUSIN project. This document describes six use cases that constitute a comprehensive effort that fulfils the main aspects of the project. This document also captures the functional and technical requirements from those use cases in such a way that they will drive architectural decisions for the eCOUSIN system and they will be used to validate the final architecture.

The requirements are defined by applying a layer scheme that will help the project to derive technical solutions for designing and implementing the eCOUSIN architecture. The deliverable D2.1 introduces the four following layers along with their companion requirements:

- The **Social-Content Interdependencies layer**, which integrates the **social and content information extraction plane** on the one hand, and the **social and content information management plane** on the other hand,
- The **Content Dissemination layer**,
- The **Network layer**.

It must be noted that this document establishes an initial set of requirements that will be rigorously updated during the project development. For instance, it is very likely that while discussing and defining the eCOUSIN architecture, we will realize that new requirements have not been initially defined. In such a case, we will update the present document to include them. The architectural process is thus meant to be an iterative and incremental approach. Our first candidate architecture will be a high-level design that we can test against use cases, requirements, known constraints or issues. As we refine our candidate architecture, we could learn more details about the design and could be able to further expand use case scenarios with new resulting requirements and improve our approach to address emerging issues.

REFERENCES

[1] Cisco, "Visual Networking Index: Forecast and Methodology, 2011- 2016," May 2012.

[2] http://www.netflix.com

[3] Sandivine, "Sandvine global internet phenomena complete'," Spring 2011, Fall 2011, Spring 2012, Fall 2012.

[4] K. Gummadi, R. Dunn, S. Saroiu, S. Gribble, H. Levy and J. Zahorjan. "Measurement, modeling, and analysis of a peer-to-peer file-sharing workload" Proc. of the 19th ACM Symposium on Operating Systems Principles (SOSP), 2003.

[5] Meeyoung Cha, Haewoon Kwak, Pablo Rodriguez, Yong-Yeol Ahn, and Sue Moon. 2009. Analyzing the video popularity characteristics of large-scale user generated content systems. IEEE/ACM Trans. Netw. 17, 5 (October 2009), 1357-1370.

[6] A. Sidiropoulos, G. Pallis, D. Katsaros, K. Stamos, A. Vakali, Y. Manolopoulos, "Prefetching in Content Distribution Networks via Web Communities Identification and Outsourcing", Journal Worl Wide Web, Volume 1 / 1998 - Volume 16 / 2013

[7] B. Mathieu, P. Truong, J.F. Peltier, W. You, G. Simon, Book chapter "Information-Centric Networks: Communication Model and Architectures", in book "Media Networks; Architectures, Applications, and Standards", CRC Press (Taylor & Francis Group), ISBN 9781439877289, May 2012.

[8] B. Mathieu, P. Truong, W. You, J.F. Peltier, "Information Centric Networking : A Natural Design for Social Network Applications", IEEE Communications Magazine, special issue « Information-Centric Networking », July 2012

ACRONYMS

API	Application Programming Interface
CDN	Content Delivery Network
C/P	Caching/Prefetching
CSP	Content Service Provider
DLNA	Digital Living Network Alliance
eCOUSIN	Enhanced Content distribution with Social Information
FSN	Federated Social Networking Standard
ICN	Information Centric Network
ICNSP	Information Centric Network Service Provider
ISP	Internet Service Provider
NAT	Network Address Translation
NSP	Network Service Provider
OSN	Online Social Network
P2P	Peer-to-Peer
PDA	Personal Digital Assistant
QoE	Quality of Experience
QoS	Quality of Service
SDN	Software Defined Network
SLA	Service Level Agreement
SSL	Secure Sockets Layer
TE	Traffic Engineering
UGC	User-Generated Content
UPnP	Universal Plug and Play