Project Number:	**FP7-257123**
Project Title:	**CONVERGENCE**
Deliverable Type:	**Report**
Dissemination Level	**Public**
Deliverable Number:	**D 8.3**
Contractual Date of Delivery to the CEC:	**30.06.2012**
Actual Date of Delivery to the CEC:	**06.09.2012**
Title of Deliverable:	**Report on trials and experimentation after cycle 2**
Workpackage contributing to the Deliverable:	**WP 8**
Editor:	**Richard Walker**
Author(s):	**Barbara Benincasa, Andrea Pede, Raffaele Di Fuccio, Richard Walker, Andrea Detti, Nicola Blefari Melazzi, Angelos-Christos-Anadiotis**
Abstract:	**In this report, we summarize results from the second phase of the CONVERGENCE user trials and the first experimental results from track 2. The results from the trial suggest that end-users perceive CONVERGENCE as potentially useful but still immature in terms of interaction design and layout. Developers are generally happy with the CONVERGENCE middleware. As far as concerns track 2, we report early evidence that deployment of the CONVERGENCE look up and routing architecture is feasible with current technologies and that the architecture is highly efficient in exploiting repetitions in user queries.**
Keyword List:	**Usability, user testing, expert review, feedback, focus groups, questionnaires, user scenarios, performance**

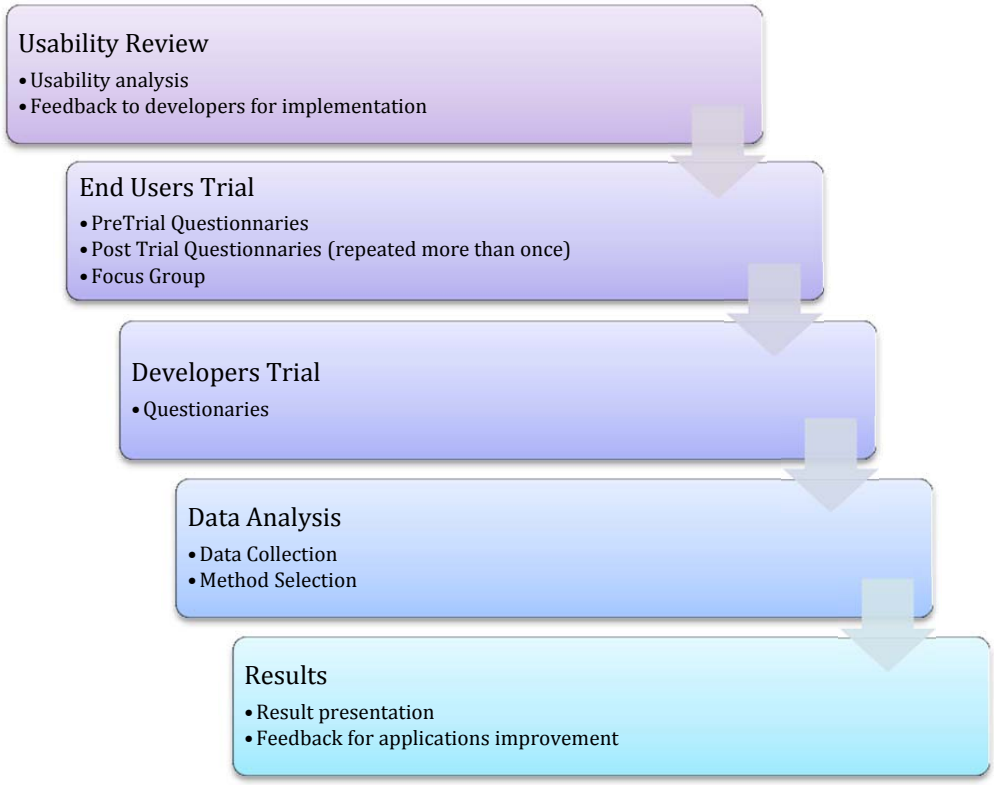
1 Executive Summary

In this document, we report results from the second CONVERGENCE end-user trial (track 1, phase 2) and the first network test (track 2, phase 1).

As described in the project proposal and in D8.1, the track 1 trial consisted of a pilot trial with friendly users in real-life conditions. The goal of the second trial, which exploited the results obtained in the first trial (see D8.2), was to improve the design of the CONVERGENCE applications for the third phase of the trials. In brief, it should be considered as a "formative evaluation". The trial involved five user scenarios defined in D2.2 and D8.1:

- Photos in the Cloud and down to Earth (Alinari)
- Videos in Cloud and Analysis on Earth (FMSH)
- Augmented Lecture Podcast (LMU)
- Smart Retailing (UTI)
- Smart Retailing (WIPRO).

Each trial involved a subset of CONVERGENCE applications. The third and last trial will test the full set. The evaluation involved five steps, as shown below.



The expert usability review was conducted by XIW in close collaboration with developers from LMU, ALI, FMSH, SIL, UTI, and WIPRO. The tests revealed a number of usability issues and software bugs. All the most serious were resolved prior to the start of the trials, as were the majority of the less serious problems.

The trials themselves involved 105 participants, including 99 end users and 6 developers. On average, each user used the software in three separate sessions, providing a large amount of feedback on their general impressions and about specific features of the CONVERGENCE applications.

In general, end-users believed that CONVERGENCE could be potentially very useful for their business. They appreciated improvements in speed and responsiveness with respect to the first trial but still had some difficult in navigating the functionalities provided by the applications and in appreciating the full potential of the system.

Developers were generally happy with the CONVERGENCE middleware, even if implementation is still in progress.

On the basis of the feedback received we have formulated the following recommendations for the next round.

1. *Functionality*. Even though CONVERGENE is not intended as an applications development project, it is essential that the last round of the CONVERGENCE trials shows functionality rich enough to effectively illustrate the advantages of the system. We repeat our suggestion from the previous round to create pop up text boxes providing additional information about what CONVERGENCE is doing when the user performs a particular action (e.g. creating a VDI, adding metadata to the VDI, defining a license for a VDI, publishing a VDI, subscribing to a VDI etc.). While this functionality would never appear in commercial implementations of the system it would be extremely useful for demonstrations.
2. *Usability and aesthetics.* Participants in the second round of trials again gave the CONVERGENCE poor marks for ease of use and layout, and many found it visually unattractive. We still believe that it would be useful to employ a graphics designer to improve this aspect of the applications. We also note a strong request from users to provide text tips for fields they have to fill in and better help functionality.
3. *Security.* Several participants expressed concern about the security of CONVERGENCE applications, (which was not shown in this phase of the trials). It is essential that the next phase should properly demonstrate these features.

The second part of the document describes the project's first network experiments and simulations (track 2, phase 1). We report the results of a test deployment on a laboratory network comprising a single CONET Subsystem and several end-nodes, serving-nodes and internal-nodes. Using this set-up we tested the CONVERGENCE API, CONVERGENCE networking, as well as the performance of the CONVERGENCE look up and cache architecture and the CONVERGENCE publish subscribe protocols. For each test the report describes the methodology used and the results. These provide initial evidence that it is feasible to implement the CONVERGENCE network using existing technologies and that our

approach provides effective load balancing in the presence of large numbers of repeated queries.

INDEX

2 Glossary

Term	Definition
Access Rights	Criteria defining who can access a VDI or its components under what conditions.
Advertise	Procedure used by a CoNet user to make a resource accessible to other CoNet users.
Application	Software, designed for a specific purpose that exploits the capabilities of the CONVERGENCE System.
Business Scenario	A scenario describing a way in which the CONVERGENCE System may be used by specific users in a specific context or, more narrowly, a scenario describing the products and services bought and sold, the actors concerned and, possibly, the associated flows of revenue in such a context.
CA	Central Authority
CCN	Content Centric Network
Cl_Auth_SC	Client Authentication with Smart Card (Challenge Response)
Cl_Auth_User_Pw	Client Authentication with Username and Password
Clean-slate architecture	The CONVERGENCE implementation of the Network Level, totally replacing existing IP functionality. See “Integration Architecture” and ”“Overlay Architecture” and “Parallel Architecture”.
CoApp	The CONVERGENCE Application Level.
CoApp Provider	A user providing Applications running on the CONVERGENCE Middleware Level (CoMid).
CoMid	The CONVERGENCE Middleware Level.
CoMid Provider	A user providing access to a single or an aggregation of CoMid services.

CoMid Resource	A virtual or physical object or service referenced by a VDI, e.g. media, Real World Objects, persons, internet services. It has the same meaning of "Resource" and it is used only to better specify the term "Resource" when there is a risk of a misunderstanding with the term "CoNet Resource".
Community Dictionary Service (CDS)	A CoMid Technology Engine that provides all the matching concepts in a user's subscription, search request and publication.
CoNet Provider	A user providing access to CoNet services, i.e. the equivalent of an Internet Service Provider.
CoNet Resource	A resource of the CoNet that can be identified by means of a name; resources may be either Named-data or a Named service access point.
Content-based resource discovery	A user request for resources, either through a subscription or a search request to the CONVERGENCE system (from literature). See "subscription" and "search".
Content-based Subscription	A subscription based on a specification of user's preferences or interests, (rather than a specific event or topic). The subscription is based on the actual content, which is not classified according to some predefined external criterion (e.g., topic name), but according to the properties of the content itself. See "Subscription" and "Publish-subscribe model".
Content-centric	A network paradigm in which the network directly provides users with content, and is aware of the content it transports, (unlike networks that limit themselves to providing communication channels between hosts).
CONVERGENCE Applications level (CoApp)	The level of the CONVERGENCE architecture that establishes the interaction with CONVERGENCE users. The Applications Level interacts with the other CONVERGENCE levels on behalf of the user.
CONVERGENCE Computing Platform level (CoComp)	The Computing Platform level provides content-centric networking (CoNet), secure handling (CoSec) of resources within CONVERGENCE and computing resources of peers and nodes.

CONVERGENCE Core Ontology (CCO)	A semantic representation of the CoReST taxonomy. See “CONVERGENCE Resource Semantic Type (CoReST)”
CONVERGENCE Device	A combination of hardware and software or a software instance that allows a user to access Convergence functionalities
CONVERGENCE Engine	A collection of technologies assembled to deliver specific functionality and made available to Applications and to other Engines via an API
CONVERGENCE Middleware level (CoMid)	The level of the CONVERGENCE architecture that provides the means to handle VDIs and their components.
CONVERGENCE Network (CoNet)	The Content Centric component of the CONVERGENCE Computing Platform level. The CoNet provides access to named-resources on a public or private network infrastructure.
CONVERGENCE node	A CONVERGENCE device that implements CoNet functionality and/or CoSec functionality.
CONVERGENCE peer	A CONVERGENCE device that implements CoApp, CoMid, and CoComp (CoNet and CoSec) functionality.
CONVERGENCE Resource Semantic Type (CoReST)	A list of concepts or terms that makes it possible to categorize a resource, establishing a connection with the resource’s semantic metadata.
CONVERGENCE Security element (CoSec)	A component of the CONVERGENCE Computing Platform level implementing basic security functionality such as storage of private keys, basic cryptography, etc.
CONVERGENCE System	A system consisting of a set of interconnected devices - peers and nodes - connected to each other built by using the technologies specified or adopted by the CONVERGENCE specification. See “Node” and “Peer”.
Dec_Key_Unwrap	Key Unwrapping and Content Decryption
DIDL	Digital Item Description Language
Digital forgetting	A CONVERGENCE system functionality ensuring that VDIs do not remain accessible for indefinite periods of time, when this is

	not the intention of the user.
Digital Item (DI)	A structured digital object with a standard representation, identification and metadata. A DI consists of resource, resource and context related metadata, and structure. The structure is given by a Digital Item Declaration (DID) that links resource and metadata.
Domain ontology	An ontology, dedicated to a specific domain of knowledge or application, e.g. the W3C Time Ontology and the GeoNames ontology.
Elementary Service (ES)	The most basic service functionality offered by the CoMid.
Enc_Key_Wrap	Encryption and Key Wrapping
Entity	An object, e.g. VDIs, resources, devices, events, group, licenses/contracts, services and users, that an Elementary Service can act upon or with which it can interact.
Expiry date	The last date on which a VDI is accessible by a user of the CONVERGENCE System.
Fractal	A semantically defined virtual cluster of CONVERGENCE peers.
Group_Sig	Group Signature
ICN	Information Centric Network
Identifier	A unique signifier assigned to a VDI or components of a VDI.
Integration Architecture	An implementation of CoNet designed to integrate CoNet functionality in the IP protocol by means of a novel IPv4 option or by means of an IPv6 extension header, making IP content-aware. See “Clean-state Architecture”, “Overlay Architecture”, “Parallel Architecture”
IP	Identity Provider
License	A machine-readable expression of Operations that may be executed by a Principal.
Local named resource	A named-resource made available to CONVERGENCE users through a local device, permanently connected to the network.

	Users have two options to make named-resources available to other users: 1) store the resource in a device, with a permanent connection to the network; 2) use a hosting service. In the event she chooses the former option, the resource is referred to as a local named-resource.
Metadata	Data describing a resource, including but not limited to provenance, classification, expiry date etc.
MPEG eXtensible Middleware (MXM)	A standard Middleware specifying a set of Application Programming Interfaces (APIs) so that MXM Applications executing on an MXM Device can access the standard multimedia technologies contained in the Middleware as MXM Engines.
MPEG-M	An emerging ISO/IEC standard that includes the previous MXM standard.
Multi-homing	In the context of IP networks, the configuration of multiple network interfaces or IP addresses on a single computer.
Named resource	A CoNet resource that can be identified by means of a name. Named-resources may be either data (in the following referred to as “named-data”) or service-access-points (“named-service-access-points”).
Named service access point	A kind of named-resource, consisting of a service access point identified by a name. A named-service-access-point is a network endpoint identified by its name rather than by the Internet port numbering mechanism.
Named-data	A named-resource consisting of data.
Network Identifier (NID)	An identifier identifying a named resource in the CONVERGENCE Network. If the named resource is a VDI or an identified VDI component, its NID may be derived from the Identifier (see “Identifier”).
Overlay architecture	An implementation of CoNet as an overlay over IP. See “Clean-state Architecture” and “Integration Architecture” and “Parallel Architecture”
Parallel architecture	An implementation of CoNet as a new networking layer that can

	be used in parallel to IP. See “Clean-state Architecture” and “Integration Architecture” and ”“Overlay Architecture”
PKI	Public Key Infrastructure
Policy routing	In the context of IP networks, a collection of tools for forwarding and routing data packets based on policies defined by network administrators.
Principal (CoNet)	The user who is granted the right to use a *CoNet Principal Identifier* for naming its named resources. For example, the principal could be the provider of a service, the publisher or the author of a book, the controller of a traffic lights infrastructure, or, in general, the publisher of a VDI. A Principal may have several Principal Identifiers in the CoNet.
Principal (Rights Expression Language)	The User to whom Permissions are Granted in a License.
Principal Identifier (CoNet)	The Principal identifier is a string that is used in the Network Identifiers (NID) of a CoNet resource, when the NID has the form: NID = <namespace ID, hash (Principal Identifier), hash (Label)> In this approach, hash (Principal Identifier) must be unique in the namespace ID, and Label is a string chosen by the principal in such a way that hash(Label) is unique for in the context of the Principal Identifier.
Publish	The act of informing an identified subset of users of the CONVERGENCE System that a VDI is available.
Publisher	A user of CONVERGENCE who performs the act of publishing.
Publish-subscribe model	CONVERGENCE uses a content-based approach for the publish-subscribe model, in which notifications about VDIs are delivered to a subscriber only if the metadata / content of those VDIs match constraints defined by the subscriber in his Subscription VDI.
Real World Object	A physical object that may be referenced by a VDI.
REL	Rights Expression Language

Resource	A virtual or physical object or service referenced by a VDI, e.g. media, Real World Objects, persons, internet services.
Scope (in the context of routing)	In the context of advertising and routing, the geographical or administrative domain on which a network function operates (e.g. a well-defined section of the network - a campus, a shopping mall, an airport -, or to a subset of nodes that receives advertisements from a service provider).
Search	The act through which a user requests a list of VDIs meeting a set of search criteria (e.g. specific key value pairs in the metadata, key words, free text etc.).
Serv_Auth	Server Authentication without Smart Card
Service Level Agreement (SLA)	An agreement between a service provider and another user or another service provider of CONVERGENCE to provide the latter with a service whose quality matches parameters defined in the agreement.
Sig	Signature
Smart_Card Role_Auth_SC	Role Authentication towards Smart Card
SP	Service Provider
Subscribe	The act whereby a user requests notification every time another user publishes or updates a VDI that satisfies the subscription criteria defined by the former user (key value pairs in the metadata, free text, key words etc.).
Subscriber	A user of CONVERGENCE who performs the act of subscribing.
Timestamp	A machine-readable representation of a date and time.
Tool	Software providing a specific functionality that can be re-used in several applications.
Trials	Organized tests of the CONVERGENCE System in specific business scenarios.
Un-named-data	A data resource with no NID.

Us_Reg_IP	User Registration to Identity Provider
Us_Reg_SP	User Registration to Service Provider
User	Any person or legal entity in a Value-Chain connecting (and including) Creator and End-User possibly via other Users.
User (in OSI sense)	In a layered architecture, the term is used to identify an entity exploiting the service provided by a layer (e.g. CoNet user).
User ontology	An ontology created by CONVERGENCE users when publishing or subscribing to a VDI.
User Profile	A description of the attributes and credentials of a user of the CONVERGENCE System.
Versatile Digital Item (VDI)	A structured, hierarchically organized, digital object containing one or more resources and metadata, including a declaration of the parts that make up the VDI and the links between them.

3 Goals and structure of this document

This document reports the results of the second end-user trial in the CONVERGENCE project (track 1, phase 2,) and the first results from network experiments and simulations (track 2, phase 1). Figure 3.1 shows the position of these trials in the overall planning of trials within the project.

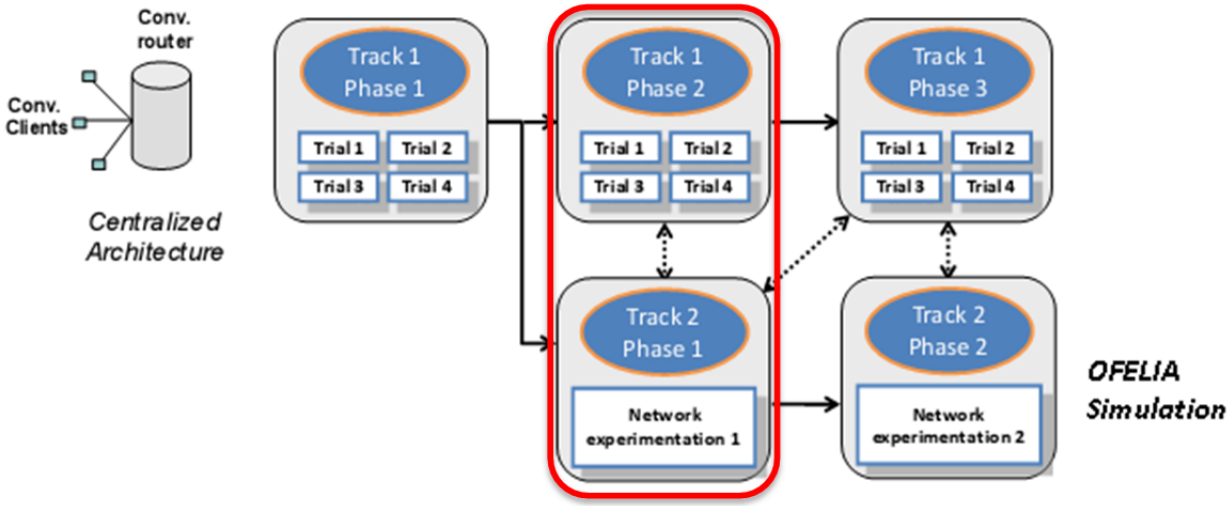


Figure 3.1: Overview of trials structures

The first part of this document is dedicated to the end-user trials and consists of three sections. Section one describes the evaluation methodology. Section two describes: i) the expert usability review that preceded the start of the trials and the measures taken to implement its recommendations; ii) feedback from participants in the trials (collected via questionnaires and an end-of-trial focus group); and iii) feedback from applications developers on the usability of the CONVERGENCE middleware. Section three summarizes our conclusions and formulates recommendations for subsequent phases of the trials.

The second part of the document focuses on network experiments and simulations and consists of two sections. Section one describes tests of the CONVERGENCE API, tests of the CONVERGENCE network and performance tests. Section two reports results from an evaluation of the CONVERGENCE publish describe protocol.

Detailed results from the trials are provided in an annex. The information provided includes tables of usability issues, questionnaires used to elicit feedback from participants, end-users replies and focus group scripts.

4 Track 1: Second Trial

4.1 Evaluation Methodology

The methodology used in the second round of trials was very similar to that used in the first round. This section therefore represents an updated version of the material already presented in our previous deliverable (D8.2). The description clearly identifies all cases, in which the two sets of trials were conducted differently.

While the first round of trials focussed on the individual functionalities provided by CONVERGENCE applications, the second round focussed on the functioning of CONVERGENCE *as a system*, paying particular attention to changes introduced as a result of the earlier trials and to user suggestions for improvement.

The trial tested the same five user scenarios as the first trial, namely

- Photos in the Cloud and down to Earth (Alinari)
- Videos in Cloud and Analysis on Earth (FMSH)
- Augmented Lecture Podcast (LMU)
- Smart Retailing (UTI)
- Smart Retailing (WIPRO)

However, the trials involved a larger number of users than the first phase (see table below) and a larger set of CONVERGENCE applications.

Scenario	Phase 1	Phase 2
Photos in the Cloud and down to Earth (ALI)	10	22
Videos in the Cloud and Analysis on Earth (FMSH)	9	17
Augmented Lecture Podcast (LMU)	17	27
Smart Retailing (UTI)	4	17
Smart Retailing (WIPRO)	9	17

The goal of the trials was to identify critical weaknesses in the second versions of the CONVERGENCE applications, to analyse the reactions of users who used the applications in

several different sessions and to formulate recommendations for improvement. More specifically the aim was to:

- Identify technical, performance and usability issues that could prevent users from making effective use of CONVERGENCE applications.
- Identify user attitude/expectations that could facilitate or hinder take-up of CONVERGENCE applications.
- Identify new applications functionality/modifications that could increase the value of applications to users.
- Identify critical issues in the deployment of CONVERGENCE hardware/software and in integration with legacy software.
- Identify ways in which CONVERGENCE applications could improve current workflows.
- Identify new middleware functionality that would facilitate the development of applications.

The evaluation involved five steps, as shown in Figure 4.1

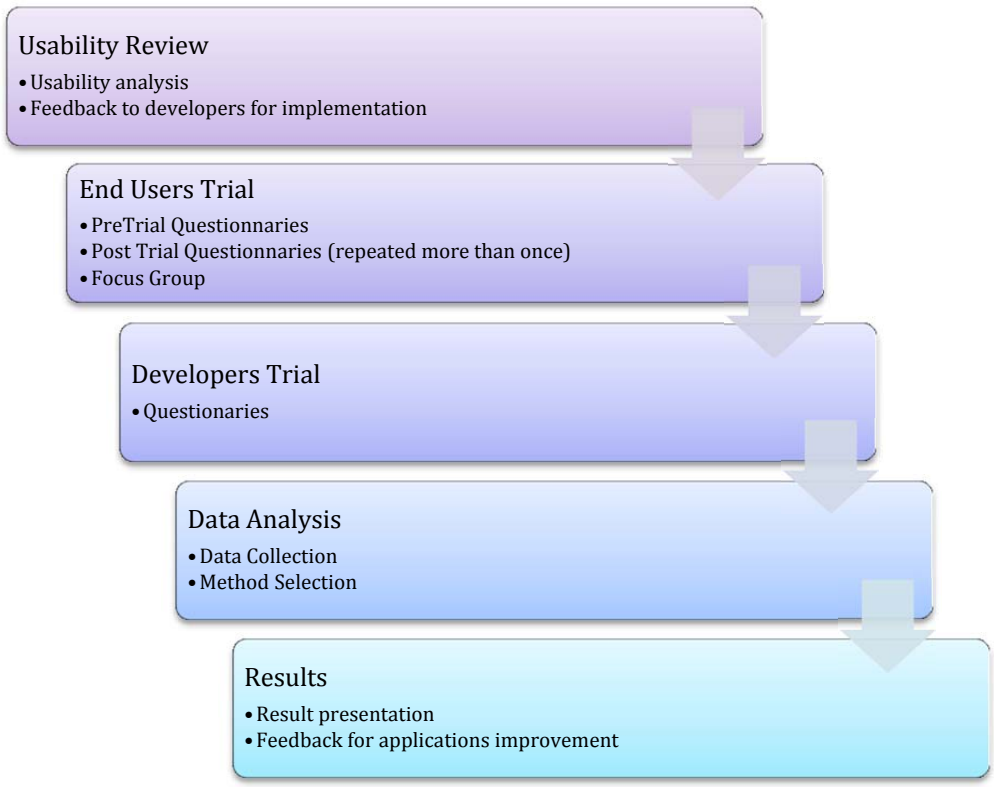


Figure 4.1: Evaluation Process

4.1.1 Usability Expert review

The aim of this task was to test and improve the usability of the user interface of the applications before their release for use in trials. The analysis, performed by a team from XIW, focused on four aspects of the user interface:

1. Interaction: user interaction and interface consistency
2. Efficiency: the speed with which users could perform key operations and the accuracy of error messages
3. Graphics: colour readability and consistency of icons
4. Contents: consistency of data, possible redundancy.

Each step in the usability analysis identified potential issues (menu structure, general user interface logic, naming, screen layouts etc.) making recommendations for improvement, which were passed back to development teams. Due to limited time not all of these issues were resolved in time for the trials; several solutions will only be introduced in the next releases of the CONVERGENCE software.

At the end of the trials, the XIW team analysed the feedback from end-users to see how far these matched the results of the expert review and to assign priorities for future implementation work.

4.1.2 End User Trial

For each user scenario, we recruited a group of representative end users who might be likely to use the future CONVERGENCE application in their work. Each participant was assigned a "role" in the scenario. Under the supervision of the leader of the work package, MORPHO, Xiwrite (XIW) prepared a series of questionnaires, focus group scripts and report templates, designed to identify difficulties encountered by users during their use of the CONVERGENCE applications and to collect suggestions for improvement.

In this phase XIW implemented new tools for data collection and data analysis in order to optimize the trial management. For further information see section 4.1.3. To facilitate data collection, the questionnaires were published online.

Each trial involved ten steps:

1. Initial presentation. The trial began with a meeting between the participants and at least two representative of the partner organizing the trial. In this meeting, the trial leader made a short PowerPoint presentation of CONVERGENCE (see D8.3 Annex), followed by a presentation of the trial itself (goals and organization).
2. Pre-trial questionnaire. Before the trial began, each participant was asked to fill in a brief online questionnaire describing his or her personal characteristics, capabilities and expectations.

3. First Trial Session. Each participant was asked to carry out a sequence of tasks following instructions from the trial leader. Tasks were personalized to the "role" of individual participants in the scenario. Table 4.1 below summarizes the tasks for each scenario. As participants performed the tasks, the second partner representative observed their performance, taking note of any difficulties they encountered.
4. Post trial questionnaire 1. At the end of the first test session, each participant filled in an online questionnaire providing information on:
    - His/her overall satisfaction with the applications tested in the trial
    - His/her satisfaction with specific functionality used in each task
    - His/her satisfaction with the overall usability and performance of the applications
    - The potential impact of the applications on his/her work
    - His/her opinion on the potential business impact
    - Technical, usability and performance issues encountered during the trial
    - Suggestions for new functionality/modifications in existing functionality
    - Other suggestions for improvement.
5. Other Trial Sessions. In the days following their first session with CONVERGENCE, each user retested the platform in two or three sessions over a period of 10 days. Each session involved the same tasks and the same test procedures.
6. Post trial questionnaire 2. After each repeated session, users filled in a new questionnaire providing information on:
    - His/her overall satisfaction with the applications tested in the trial
    - His/her satisfaction with the overall usability and performance of the applications/platform
    - Technical, usability and performance issues encountered during the sessions
    - The time taken to learn to use the applications (learning curve)
    - Differences between CONVERGENCE and systems users knew already
    - Suggestions for improvement.
7. Focus Group. At the end of the trial, participants were invited to take part in a focus group, designed to elicit specific suggestions for improvement.
8. Developers feedback. At the conclusion of the trial, developers were also asked to fill in a questionnaire, eliciting information on:
    - Technical/performance issues related to the applications
    - Suggestions for improvement.
9. Data Collection. At the conclusion of the trial the data collected was inserted into a database, which was used for the subsequent data analysis. Qualitative information – in particular recommendations for improvement - was saved in report templates filled in by the trial coordinator.
10. Data Analysis. Given the small size of our samples, sophisticated statistical methods would have been out of place. The data analysis in Section 2 outlines the results for

each scenario and for each specific target group. The final section draws some preliminary conclusions.

Task	Photos in the Cloud and down to Earth (ALINARI)	Videos in Cloud and Analysis on Earth (FMSH)	Augmented Lecture Podcast (LMU - Students)	Augmented Lecture Podcast (LMU - Lecturers)	Smart Retailing (WIPRO - Manufacturer)	Smart Retailing (WIPRO - Retailer)	Smart Retailing (UTI - Retailer)	Smart Retailing (UTI - Customer)
1	Register to the application	Register to the application	Register to the application	Register to the application	Register to the application	Register to the application	Register to the application	Register to the application
2	Upload a new photo	Manage a video	Search for lecture podcast	Create podcast components and lecture podcast	Create product VDI	Products creation	Navigate through products	Navigation through product
3	Publication of photo	Manage an analysis	Watch lecture podcast	Update and Revoke podcast components and lecture podcasts VDIs	View created product	Get a subscription match	Promotion of products over the Internet	Subscription to product
4	Subscription to a photo	Subscription to a video	Annotate lecture podcast		Publish product VDI		Customer Feedback Management	Digital information about products
5		Subscription to an analysis	Subscription to annotation					

Table 4.1: Tasks used in trial by scenario and partner role

4.1.2.1 Evaluation tools

4.1.2.1.1 End-User Questionnaires

The evaluation used three ad hoc questionnaires designed by XIW.

1. Pre-trial questionnaire (see D8.3 Annex). The goal of this questionnaire was to characterize the sample. The questionnaire, which was identical to the questionnaire used in the first trial was the same for all scenarios and all participants.
2. Post-trial questionnaire 1 (see D8.3 Annex). Post-trial questionnaire 1 was designed to measure user satisfaction with the functionality provided and was specific for each scenario and to each user "role".

3. Post-trial questionnaire 2 (see D8.3 Annex). Post-trial questionnaire 2 was designed to measure satisfaction and to elicit information about the ease with which users learned to use the application.

Each questionnaire contained three kinds of question.

- Multiple-choice questions with answers on a Likert scale from 1 (not at all, very little) to 5 (a lot, excellent). This kind of question was used to elicit information on user satisfaction with specific functionality provided by the application.
- Open questions. This kind of question was used to elicit suggestions for improvements to the applications.
- Ranking questions. This kind of question was used to identify user preferences with respect to specific features of the applications.

To facilitate data collection and the subsequent analysis, all the questionnaires were published online. Questionnaires for trial participants were anonymous. A numerical code made it possible to match pre- and post-trial questionnaires. Developer questionnaires were signed. Trial participants were instructed to fill in the questionnaire step by step as they performed the sequence of tasks defined in Table 4.1. The data was collected in a dedicated database.

4.1.2.1.2 Focus Groups

After the end of each trial, participants were invited to take part in a focus group, coordinated by the two partner representatives, and based on a pre-defined script. This script, designed by XIW (see below) was completely redesigned with respect to the script used in the first round of trials, and focussed on the following aspects of the system:

- Navigation
- Help support
- Layout/graphics:
- Performance
- Robustness/ Technical problems
- Privacy
- Security
- Ease of use and learning curve

Partner representatives took great care to ensure that the discussion focused on the key issues identified in the script. All focus groups were conducted in participants' mother tongues and recorded for later analysis. At the end of the focus group, partner representatives filled in a report template, designed by XIW. The reports were then analysed scenario by scenario.

4.1.2.1.3 Developers' Questionnaires

In the original plans for the evaluation (D8.1), we proposed to conduct individual interviews and/or focus groups with developers[1]. Already in round 1, however, it became clear that different interviewers would have interviewed different developers, making it hard to compare the results. We therefore decided that the final step in the evaluation would consist of a questionnaire for applications developers[2] . The goal of this questionnaire was to identify any problems developers had experienced while using the CONVERGENCE middleware, and to collect suggestions for improvement. The questionnaire consisted of two sections:

- A series of questions relating to common tools developed in WP7 present in most of scenarios.
- Open questions on issues related to the coding of the end-user applications.

4.1.3 Main differences with respect to the first trial

As already mentioned in the previous section, the second round of trials was marked by several changes in methodology and tools, with respect to the previous round. For the convenience of the reader we summarize these changes below.

1. Methodology:
   - The trials involved a higher number of users than the trials in round 1
   - Each user tested the CONVERGENCE applications in three separate test sessions over a period of ten days. In the first round of trials users tested the applications only once. This change in procedure made it possible to study the learning curves for the applications.
2. Tools:
   - Improvements in post-trial questionnaire 1
   - Creation of a new post-trial questionnaire for subsequent sessions
   - Redesign of the Focus Group scripts
   - Improvement in the questionnaire for developers
3. Data collection and analysis:
   - Deployment of LIME SURVEY (www.limesurvey.org), an open source tool that provided the team with a new platform to manage the questionnaires.

[1] For the purposes of the evaluation developers were defined as technical staff involved in the development of applications for the trial and/or interfaces to legacy software (see D8.1).

[2] In CONVERGENCE the design and implementation of the applications used in the trials was separate from the design and implementation of the CONVERGENCE middleware. Applications developers, while obviously not independent of the project, can thus be considered as *users* of the middleware.

4.2 Results of the trials

4.2.1 Usability Expert Review

The expert usability review was conducted by XIW in close collaboration with developers from LMU, ALI, FMSH, SIL, UTI, and WIPRO. For “Photos in the cloud and down to earth”, “Videos in the Cloud and Down to Earth” and “Smart Podcast” the tests were based on the web interfaces to the application. For “Smart Retailing” and “Smart Retailing with POS” they used remote terminal sessions.

The review revealed far fewer usability issues than the review that preceded the first round of trials. All the most serious problems found during the review were resolved prior to the start of the trials, as were the majority of the less serious problems. In particular the implementation teams resolved critical problems with user registration, introducing the possibility of logging on via Google and Facebook. This greatly increased the usability of the applications. The teams also resolved problems with date validation. In the case of the UTI “Smart Retailing” trial, the application was significantly rewritten. Major problems, which were not resolved in this stage, will be dealt with before the third phase of the trials. Solutions for some minor problems may be postponed until a later stage.

The detailed results of the usability analysis are shown in Annex 1. For each problem identified, we provide a description of the problem, an estimate of its severity, recommendations for a solution, and the current status of work to implement the recommendation. The main problems found were the following.

- *Navigation and menu items.* Applications often failed to provide users with a clear indication of the function they were using. Labels were placed in positions that made them hard for users to find. In some case menus are not well organized from a logical point of view.
- *Instructions and tips.* This problem, originally detected in trial 1 has been partially solved. However, the team believes it is still necessary to provide better text hints to users when they have to fill in text fields in the applications.

The study also detected several minor problems, with only a minor impact on usability. Examples include the following:

- *Data entry*. Data entry sometimes required the mouse; this made it less efficient than necessary.
- *Poor graphics design.* Applications still used multiple graphics and texts styles and make excessive use of “decorative” graphics with no functional purpose.

- *Data visualization.* Graphics objects were often displayed in inappropriate sizes and positions. In some cases visualizations contained very little data. In others, users were asked to handle lists of objects containing very large amounts of information.

Many of these issues were also mentioned by users in questionnaires and focus groups.

4.2.2 The trials

4.2.2.1 Photos in the Cloud and down to Earth (ALI)

The scenario for the trial was designed by ALINARI. The application used in the trial was implemented by SIL. The goal was to test whether CONVERGENCE can provide useful support for new business models in the professional photography business, making it easier for photographers to contribute and describe photos, improving access for users and generally facilitating the management of the services Alinari offers to its customers.

Participants in the trial used custom applications running on top of the CONVERGENCE framework and the CONVERGENCE network. These applications provided Alinari staff and freelance photographers with a user interface that made it easy for them to create, publish, un-publish, describe, update and subscribe to photos (represented by Resource VDIs and Publication VDIs). Alinari staff could also retouch the photos and add them to the Alinari catalogue. End-users could view and buy photos using a local application connected to the Alinari server. A new feature implemented in this trial allowed them to encrypt pictures, making them inaccessible to non-authorized users.

4.2.2.1.1 End-User Sample

The trial consisted of two separate sessions each involving 22 participants. 8 participants were between 31 and 40 years old; other users were older. 19/22 had a university degree (a first level degree or a master's degree). 18/22 considered themselves to be expert users of PCs, which they used for between 3 and 8 hours a day (13/22). 15 had used PCs for more than 8 years. 14/22 participants were also familiar with tablet and smartphone technology. 5 worked as photographers, 2 were teachers of photography. The other participants included 3 managers, 1 director, 1 exhibition supervisor, 3 experts in the photo management software, 2 indexing experts, 1 developer, 1 art historian, 2 administrative staff, and 1 trainer in photography. The sample did not include any students.

4.2.2.1.2 Results in a nutshell

The data collected during the ALINARI trial suggests that CONVERGENCE could be extremely valuable for professional users. Nonetheless participants raised many issues concerning the use of graphics and also identified other usability issues. In general, participants would like to see a system that is easier to use and to personalize, more attractive and professional. They also suggest improving the help system, the graphic layout and browser compatibility (especially with Apple Safari). Several criticized the difficulty of sharing resources with other photography professionals on the CONVERGENCE network.

Despite these criticisms, the majority of users were aware that CONVERGENCE is still in an early stage of development and believed that future, enhanced versions could be extremely useful for their work. They also thought that CONVERGENCE could be useful for other categories of user including content providers, museum curators, gallery managers, artists, etc. Many thought that with suitable improvements, especially with respect to graphics, CONVERGENCE could replace the systems they use at present. It appears that, at least for this target population, layout, graphics and the aesthetics play an essential role in purchasing decisions.

4.2.2.1.3 Results in detail

4.2.2.1.3.1 Post-Trial questionnaire 1

Task – 1: Register and login to the application

The user was asked to create a new account, filling in his/her name, username and password. Users were also offered the possibility of logging in via their Gmail or Facebook accounts.

In reality, users logged in either through their CONVERGENCE account (16/22) or through Gmail (6/22). None uses Facebook authentication. Users considered the implementation of this function to be slightly better than adequate (3.3/5.0), relatively easy to use (3.1) and good in terms of error handling (3.9/5). However, they gave the system a poor mark for layout (2.6). Only one participant experienced problems in logging in (a time out during the login process). Users suggested there should be more support functions (e.g. online help, account verification, information about content licensing options). Several suggested improvements in graphics and layout.

Task - 2: Upload a new photo (creation of R-VDI)

The user was instructed to select a photo, fill in the metadata, (title of photo, caption, date, location, tags); specify licensing conditions and upload the photo.

The 59% of participants (13/22) had already uploaded photos using open source tools and/or commercial services and considered CONVERGENCE as more effective than these tools.

The majority thought that this functionality could be very useful for their work (15/22). Users were relatively negative about layout (1.9/5) but gave good rating for responsiveness (4.1/5), speed (3.8/5) and overall quality (3.1/5). Suggestions for improvement included the introduction of new icons to manage specific features and improvements in button design.

Task -3: Publishing a photo (creation of a P-VDI)

The user was instructed to publish a photo, creating a P-VDI and to fill in the tags CONVERGENCE would use to match the publication VDI to subscription VDIs.

55% of participants reported that they had never previously published photos online. Those who had used a variety of tools include Adobe suite, Apple iPhoto, CMS, Flickr, Cyberduck and Kodak software. Their feedback suggested that CONVERGENCE was better than these tools (12/22). Nevertheless they still gave CONVERGENCE a relatively low mark (2.3/5). Although they believed that this functionality was very useful for their work (4/5), they gave CONVERGENCE low marks for ease of use (2.6/5), layout (1.7/5) and overall quality (2.6/5) and pointed to a number of specific problems (lack of online help or avatar-based support; poor interface design). They also suggested that the application should provide a broader range of tools and options.

The system received good scores for speed (3.9/5), responsiveness (4.2/5) and error handling (4.2/5).

Task -4: Subscribing to photos (creation of an S-VDI)

Users were instructed to create a subscription to photos meeting user-specified search criteria. CONVERGENCE then matched these criteria against photos present in the system and returned a list of results. Users could click on an item in the list and see the photo and related information.

5 participants did not test this functionality, 16 reported that subscribing to a photo could be useful or very useful for their work. However, they also reported less than adequate ease of use (2.5/5). Participants gave the application poor marks for layout (1.7/5) but better marks for error handling (3.3/5), speed (4/5) and responsiveness (4/5). Regarding overall quality users considered it to be adequate (3/5). Suggestions for improvement included requests to improve help and support and to provide video training materials. Users also suggested the introduction of new functionality to support the buying and selling of photos, and to provide advice.

Final notes

Even though participants considered that CONVERGENCE provided functionalities that would be potentially useful for their work (20/22), its overall score (2.1/5.0) was low.

Only four users reported making mistakes in pushing buttons when they used the application for the first time and three out of four said that these errors were rare. However the same users

reported that the error messages they had received were not clear. Only two complained about difficulties in using the software, lack of appeal or bugs (a problem during login, failure to provide feedback when uploading a photo). Half of the sample (11/22) would consider using CONVERGENCE instead of the application they use currently; 9/22 users would recommend CONVERGENCE to a colleague. Regarding photo encryption, the majority of users considered this tool important for their job. A small number did not understand how to use the tool. 64% considered the error messages useful. All the participants believed that it would be useful to have online help.

Supervisor's report

The supervisor did not report any significant problems.

4.2.2.1.3.2 Post trial questionnaire 2

The second post-trial questionnaire, administered after users' second and subsequent sessions, measured how far users' experience in the first session made it easier to use the application a second time.

Participants who filled in the questionnaire stated they had had no difficulty in remembering how the system worked. 77% found it easy to recover the photos they had posted in their last session, remembering how to perform this and other tasks. The majority (12/22) reported that they had improved their navigation skills. However, 18/22 gave CONVERGERNCE the same mark they had given it in the previous session, justifying their evaluation with the lack of improvements since the first session and the unattractive design. Users gave the application poor scores for layout (1.3/5), ease of use (2.4/5) and general quality (2.3/5). Several commented that it was not user friendly, citing difficulties in understanding the workflow and the poor interface. One reported that she had not been able to assign a licence. There were also reports of technical difficulties. Four users experienced problems in the subscription procedure (system errors and problems with tags), in uploading (upload failure, tag problems) and in publishing (slowness and problems with filling in the form). One participant found bugs in the login/registration procedure.

By contrast, participants strongly appreciated the speed and the responsiveness of the system, giving it a mark of more than 4. They also offered a positive assessment of the system's error handing capabilities (3.6/5) and commented positively on the new photo encryption feature, noting its importance for privacy and data protection. However, a minority had difficulty in understanding the function.

As far as concerns their overall assessment of the system, all participants considered the platform as potentially useful for their work and believed that it could improve their productivity. However, only eight were ready to begin using it immediately - a lower number than in the first session. These users would also be willing to recommend CONVERGENCE

to a colleague. Fourteen participants preferred using other systems. The main reasons they gave were that CONVERGENCE was too complicated to use (55%), that it had an unattractive design (45%) and that it was not sufficiently popular (31%). The mean score for overall quality was the same as in the first session: 1.9/5.

4.2.2.1.3.3 Focus Group

Navigation: Most participants did not appreciate the GUI or the navigation experience. Users reported that they had often failed to understand the workflow: which bottom to press, what to do, and where to go. Participants suggested the interface should use different colours for different actions (i.e. pink for subscription, green for publishing, blue for searching, yellow for editing). Another suggestion was to number actions in sequence (i.e.: 1,2,3,4,5,6), making it easier to move from one to the next.

Help: Several users reported that the system showed inappropriate error messages, without any specific tracking code or debugging function. Participants suggested an online help providing extensive explanations of each function. One user noted: *"Categories of Analysis Subscription are not easy to understand and to fill in."*

Layout/graphics: Participants considered the graphics and layout very poor for a potentially commercial application and suggested that designers should focus on the logic of the layout, on the basic template and on the way they used colour. Participants also suggested increasing the size of the fonts and the images previews.

Performance: Participants were satisfied with CONVERGENCE's speed. Some reported a time out an error message during the session. This was probably due to Internet congestion.

Robustness/ technical problems: When used with Apple Safari, the browser most commonly used by Italian photographers, the application crashed. Participants reported that it was stable when used with other browsers but added that compatibility with Safari was an important issue for them.

Privacy: Participants considered synchronization and expiry dates as very important features for privacy protection. Many would have liked to receive a technical description of the level of privacy the system was able to guarantee. Participant comments included the following: *"I am very concerned about security and like the project approach but want to learn more of what is happening in the real world, under the "hood": "The system should provide an indicator showing the strength of privacy protection provided and warnings about possible attacks from intruders"*.

Security: Many users expressed the view that the system provided very interesting security features and believed that that their content was better protected than in existing systems. One suggested that users should be able to view the license for a VDI at any time (*"I want to see the licenses that I'm subscribing to."*). Participants were also interested in VDI-level security

and possible scenarios for attacks (*"I want to see what level of security I am getting, what is happening during my navigation process, how effectively I am protected."* Participants suggested that one way of increasing the security of personal information would be for users to make a clean separation between content and personal data.

Ease of use and learning curve: Participants noted that it took them time to learn to use CONVERGENCE and that, especially at the beginning, the learning curve was quite steep. Many reported that they had felt lost, not knowing what to do. However, they also reported that using the application had been easier in the second session (*"After using it for a while I can use it quite easily").*

Best features: According to participants the best features of the application were the security and sharing mechanisms. They also believed that the application was useful for searching for, managing, and sharing photos. Participants considered CONVERGENCE as an innovative application, offering features that are not present in competing platforms. One participant stated: *"Today content sharing is very popular on Flickr, YouTube, Facebook and so on. However CONVERGENCE can provide unique and strong solutions to help professional photo agencies in their licensing effort". Another considered CONVERENCE as "a very innovative and unique solution for content providers"*. Users emphasized CONVERGENCE's great potentiality, and understood that much of this potentiality had yet to be realized. *"I want to use it as soon as it become mature"*; *"We hope to use it more extensively once the system will be fully tested and is in full operation).* In their opinion the system has great potential in the cultural heritage and photographic sector, as a tool for content providers, museum curators, gallery managers, artists, photographers etc.

Worst features: Participants commented negatively on the application's layout and the ease of use. As one typical comment put it: *"The graphic interface and whole system usability are quite poor", "They might be good for the R&D world, but they are not attractive at all (...) for the commercial and consumer sector*". Others complained that the application provided them with relatively little information and did not have advanced help functionality (*"Subscription information and the parameters for defining the information were not very easy to see and understand. The application should provide more clear info about the photographic display").*

Improving CONVERGENCE: Users asked for more information and online help, especially during search and subscription procedures. Despite their positive assessment of the system's privacy and system protection features, participants thought it necessary to make the system even more secure.

4.2.2.2 Videos in the Cloud and Analysis on Earth (FMSH)

The second scenario tested was "Videos on the Cloud and Analyses on the Earth", by FMSH. In this scenario, CONVERGENCE is used to improve the management of audio-visual archives and to exploit the potential of semantic techniques, when the same video resources are exploited several times in different contexts of use (analyses using different domain ontologies, posting on different video channels). In the scenario, Video Material Owners (VMO) provides videos to the archive; Video Distributors (VD) receives notifications of new videos, which they then download, decrypt and make available for streaming. Analysts download the videos from CoNet, decrypting and analysing the content. They then upload their analyses to CoNet and notify interested VCOs that they have done so. Video Channel Owners (VCO) manages video channels and content posted on the channels they own. Channel Holders (CH) receive notifications of new channels and new posts update the channels and make them available on the Internet. Video Channel Users (VCU) explores the content of Channels, subscribe to new posts and browse channels.

4.2.2.2.1 Sample

The trial took place in France (twelve participants) and in Peru (five participants). participants were asked to play the role, either of Video Material Owners (VMO) or of Video Channel Users (VCU). Seventeen participants completed the first session of the trial, twelve completed a second session and five a third. Sixteen users participated in the focus group, including one Communication Manager, one Engineer, one Film director, two anthropologists, two students, one musician, one analyst, five Researchers, one Scientific Communicator, one musician, two Web developers and one Dentist. Nine came from outside FMSH. These participants included experts in video production and management (researchers, students, artist, and independent video producers). Eight users came from the ESCoM laboratory at FMSH, where they specialized in media production and analysis. All except one had a university level education. All used computers regularly in their work and leisure activities; the majority used PCs between from three to eight hour per day. All considered themselves to be expert users of computers giving themselves an average mark of 4/5. Knowledge of smart phones and tablets was less widespread (2.4/5). Only seven out of seventeen considered themselves skilled in this area.

4.2.2.2.2 Results in the nutshell

The data from the trial shows that users were interested in CONVERGENCE and believed it could help them in their work. However, its lack of features and poor usability, combined with technical difficulties, meant that in their opinion it was not still an attractive and competitive solution for professional use.

Participants were aware that CONVERGENCE is still at an early stage in its development. The improvements they would most like to see include richer and more flexible ontologies, as well as an more attractive user interface, an increase in the number of file formats supported by the system and more details in the analysis functions. Many users wanted improvements in the help system. Their general impression was that CONVERGENCE does not fit the needs of the general user but that it could be extremely useful for target populations (for example in research, education or art) with a strong desire to control the content they produce. Some had difficulty understanding technical aspects of CONVERGENCE and the way it manages the data as VDIs.

As in the Alinari trial, users saw the system as very fast and responsive, relatively easy to use and potentially useful to their work. They were especially impressed by the new analysis features implemented in this second version (new functions to improve the traceability, control and sharing of content; improved security for intellectual property). They were also favourably impressed by other innovative features, such as licenses, downloads, subscriptions and expiry dates. Several participants suggested that the system should be extended to other domains. Suggested applications included e-learning, tourism, joint authoring of works of literature, other forms of collaborative work, managing of payments to artists, tracing the uses of research articles, and monitoring the way users reference and use web sites.

4.2.2.2.3 Results in detail

4.2.2.2.3.1 First session

Task 1: Register and login to the application

The user was instructed to create a new account, filling in her name, username, password, etc. Alternatively she could log in using her Gmail or Facebook account.

In practice, participants mainly used their CONVERGENCE account (9/17); four used Facebook account and another four used Gmail. Users expressed a positive evaluation of this functionality, giving it a score of 3.5 for layout, 3.8 for general quality, 4.5 for ease of use and 3.9 for error handling. In the areas of the questionnaire reserved for comments, users expressed their satisfaction about the opportunity to connect via Facebook or Google; and emphasized the tool's ease of use. One suggested the addition of a password recovery function.

Task 2: Manage a new video

Users were instructed to create a new video, fill in the appropriate metadata, (title, subtitle, type, producer of the video), define specific licenses for Video Channel Users, Distributors, etc., and upload the video.

Eleven out of seventeen users had previous experience with video management software. Other programs and services they had used included Daily Motion, ESCom, Facebook, GAAV, Streamlike, Vimeo, WimTV, and Youtube. All the participants believed that this was a very useful functionality for their work (4.1/5) and gave the CONVERGENCE implementation a high score for general quality (3.9). The users gave it high scores for ease of use (4.1), responsiveness (4.1), and speed (3.8). The scores for layout and error handling were lower (3.2 and 3.6 respectively). According to participants, the application needs several improvements before it can compete with programs already on the market. Suggestions included the introduction of new functionality for video sharing, to explain the difference between different types of license, to make it easier for users to recover from errors and to request help from technical support. There were also suggestions for improvements in the graphics and the layout. Most importantly of all, participants would like to manage video in different formats (for example .FLV) and in full screen mode. Some users, on the other hand expressed their satisfaction with this aspect of CONVERGENCE.

Task 3: Managing an analysis

Users were instructed to create an analysis VDI using a predefined format.

Eight out of seventeen had already used other software to analyse own videos on the Internet and considered this functionality useful for their work. The majority had used software developed by FMSH: the FMSH Video Analysis Tool and the Video Analysis Application. Other software they had used included YouTube and INA Studio. Users with previous experience reported that the functionality offered by CONVERGENCE matched that of other programs or was slightly better; the score assigned for this application was 4.0/5. However, the overall quality score was just 3.3/5. The lowest score was for layout (3.0/5.0). However, users gave good marks to other aspects of the application: ease of use (3.9/5), error messages (3.7/5), overall quality (3.6/5), speed (4.1/5) and responsiveness (4.1/5). One participant proposed the use of text tips to guide users during the analysis process. One suggested that the application should allow users to display more metadata. Another thought it should be possible to read an analysis without downloading it first. Other comments focused on the need to improve the aesthetics and the usability of the interface.

Task 4: Subscription to a video

Users were instructed to subscribe to a video. During the procedure, they could also browse (and delete) their previous subscriptions. CONVERGENCE used the metadata inputted during the subscription procedure to match the subscription to publications.

Some participants (7/17) had already used other services (Daily Motion, FMSH Interview Analysis, Google Alerts, RSS, YouTube,) to search for videos on the Internet and considered this functionality to be essential for their work. In general, they considered the CONVERGENCE service to be slightly better than other services they had used, giving it a

score of 4.4/5. A third of the participants had never searched for a video before and considered searching less important. Only three had problems with this functionality. Their main difficulty was in understanding tags. However, the majority gave a positive evaluation. Overall, ease of use received a score of 3.9, error handling 3.8, general quality 3.7 and layout 3.3. Users appreciated the speed and the responsiveness of the application, giving it a score of 4.3. Suggestions for improvement included the addition of new functionality allowing users to display the full set of subscription parameters, to modify or delete subscriptions, and to create new tags to facilitate search; another suggestion was to introduce new pre-defined tags. Other users suggested the introduction of new fields for the start date of subscriptions and the addition of new metadata to describe videos – in particular information about the location where the video was made. Further suggestions included improved systems for "tagging", mail notifications when a new publication matched a subscription, and improvements in online help.

Task 5: Subscriptions to an analysis

Users were instructed to create a subscription to an analysis VDI.

Only one participant had already used other software (the FMSH Interview Analysis tool) for this purpose. This user considered CONVERGENCE to be on a par with the other tool, giving it a score of 4.0. The majority of users consider this functionality to be useful for their work. Only two experienced problems with this task (difficulty in the criteria for selecting ontology). The highest scores were again for speed (4.3/5.0) and responsiveness (4.2/5.0). Ease of use, overall quality and error handling all received a score of 3.6. Layout obtained 3.2. User suggestions included the addition of new functionality, allowing them to add their own tags and keywords, improvements in the analysis search function, the introduction of additional information about location and date, better display of notifications, improvement in the interface design and a more advanced system of tag matching

Final notes

The majority of users (94%) considered CONVERGENCE useful for their work, and thought that it could improve their productivity. Only one disagreed, stating that she did not work online. Thirteen users would consider using CONVERGENCE instead of the application they use at present. Those who did not agree argued that CONVERGENCE had too few users, that the system was not user-friendly and that the platform was too similar to existing software. These users would not consider using CONVERGENCE even if it improved. Twelve out of seventeen would recommend CONVERGENCE to a colleague. Participants gave the application an overall score of 3.8. The most popular characteristic was its performance. The majority of participants (twelve out of seventeen) made no errors while using the application. However, nine said that the error messages they received were difficult to understand and thirteen though the application needed an online help function. Two reported technical problems.

Supervisor's report

The supervisor did not report any problems.

4.2.2.2.3.2 Post Trial questionnaire 2

As in the other trials, the second post-trial questionnaire, administered after users' second and later sessions with the software, measured how far users' experience in the first session made it easier to use the application a second time.

In general participants reported a positive learning experience, remarking on the ease with which they had recovered videos posted in the previous sessions. Sixteen out of seventeen had no difficulty in remembering how to execute specific tasks; fifteen thought they had improved their navigation skills. Participants again appreciated the speed (3.9) and the responsiveness (4.1) of the software but gave a negative evaluation of the help function (2.8). The sample gave the application moderately good scores for layout (3.1), ease of use (3.7), error handling (3.4) and general quality (3.6). In general, however, scores were slightly below those for the first session and the number of participants who would consider using CONVERGENCE instead of their current application fell from thirteen to eleven. Of these, ten would recommend CONVERGENCE to a colleague. The users who preferred other systems cited CONVERGENCE's lack of user friendliness. However, eight might change their opinion if the system improved.

Comparing the first and the second sessions, only five users changed their overall view of CONVERGENCE. Three gave it a lower rating than in the previous session, motivating the lower mark by difficulties in using the system, problems in licensing and subscription procedures and the impossibility of changing notifications. Two improved their mark.

User suggestions included improvements to the menu on the homepage, easier-to-use procedures for browsing video rights, and the display of more information on licenses. Only three users reported technical problems: two experienced difficulties when subscribing to an analysis; one reported a problem with the video subscription procedure.

The supervisor did not report any major technical problem. However, in the trial in Peru the Internet connection was poor. Therefore the trial had to be based on short videos rather than live use of the application. The main problem in the Peruvian trial was in preventing dropout by participants.

4.2.2.2.3.3 Focus Group

At the end of the trial, participants took part in a focus group organized along the same lines as in the Alinari trial. As in the Alinari case, the results of the focus group matched those from the questionnaires.

Navigation: Users familiar with computers did not experience major difficulties in navigating the application (*"The presentation and navigation are easy to understand."*). Participants with less experience sometimes found it difficult to find specific sections. The majority felt that the allocation of functions between the top and side menu was not logical. Some saw problems with the use of terminology. For example: *"When I click on Create Analysis, I expect a tool to open for staring an analysis, but nothing happens. Actually, I have to select an existing analysis file in order to upload it."* Some users complained about the license creation form, which used pop-up windows and never provided a summary of what licenses had been created. (*"Once the license is created, something that upsets me is the fact that a pop-up is opened, and then I have to close it without any helpful message. I cannot be sure that the license was successfully created."*). Several participants suggested that it would be useful to display a summary of recent events (videos and analyses created/published, notifications received, etc.) on the home page: One user commented: *"I think about a home page from where I could easily re-find the last activities I performed: my last videos, my last subscriptions, etc."*. Most gave positive feedback on the subscription parameters and the relevancy of the resulting matches (*"All the parameters were useful, it is easy to search and subscribe to a video and the matches are efficient.").*

Help: Many users were unsatisfied with the information provided about fields they had to fill in. This problem was particularly severe when they had to subscribe to an Analysis (*"Categories of Analysis Subscription are not easy to understand and to fill in."***).** Some would have liked more information about different kinds of licenses. More generally, participants believed that the application should have better help sections, provide better information about subscription parameters and licenses, and better error messages.

Layout/graphics: Many users suggested the need for a multi-language user interface and for tags in languages other than English. The majority did not like the quality of the layout, criticizing the choice of colours for fonts and the lack of icons. *("The colour of the fonts should be changed, especially the colour used for the message, which is the same used for error messages when I am browsing"*). Some suggested that participant should be able to personalize key aspects of the interface (language, colours, fonts, use of images etc.).

Performance: Participants in Paris found the application quick. Participants in Peru, whose maximum speed for uploading a video was 15 kb/s, found it slow. These users were unhappy about the performance and would like to see a progress bar inside the application (currently only available with Google Chrome).. Many users suggested that they should be able to fill in the Video Creation form while the file was uploading.

Technical problems: Almost all users experienced problems with the fields that relied on the CDS, with copy/paste operations and with the insertion of tags not included in the ontology. Participants reported that several of them had received incoherent matches to subscriptions

during the first session but that the problem had been resolved before the following sessions. One user experienced a technical problem when trying to browse a video. Some of them would have liked to have an enhanced ontology including relationships between tags. The majority did not like being restricted to tags in the ontology and some of them argued that the ontology did not cover their particular areas of interest. These users would have preferred to use free text, like Google (*"It is very important to have free categories for Tags and Type because the user should be free to give information about his own video. Sometimes, there is important data included in a rare video that you will never find e in a predefined list. So I think for a scientific application, it is really important to let users define their own videos with their own terms."*). Others argued that a default ontology could help users to focus their searches (*"It should be possible to have a related search option. For example, I write an expression, and even if it is not in the ontology, I would find a result because what I wrote would be related with the expression stored in the ontology."*). At the end of the discussion, participants agreed that the application should offer both options. According to another suggestion, the system should propose a preview of the results of a subscription before publishing it.

Privacy: Many users were concerned about privacy issues when logging into CONVERGENCE through their Google or Facebook account. These users thought that CONVERGENCE did not provide them with enough information about the way their information was used. Almost all users who logged in through Google or Facebook used their professional account or created a new private account exclusively for this purpose (*"I used my Google account because I only use it in a professional context. I did not use my Facebook account because I really did not want to communicate my personal data."*). The majority thought that synchronization and expiry dates provided valuable privacy protection (*"Licenses, expiry date and an easy way to delete private information are important improvements for the management of privacy.").*

Security: Many users considered that the system provided very interesting security features and that these were better than those provided by existing systems. They found that licenses were easy to manage, interesting and very useful. However, some would have liked to see the Event Reports (which were not shown in this trial). Some users also complained that the trial did not properly show the main innovation introduced by CONVERGENCE (interoperability of licences across different CONVERGENCE applications in different scenarios). Some of them thought that that a system in which everything converges (thanks to VDI technology) could easily be hacked and would never be safer than current technologies. In this case, VDIs could be a danger for data security. All participants would have liked to see the licenses for videos while they were browsing their subscriptions *("It would be more convenient if I could see the licenses when I'm subscribing. It would prevent me from receiving notifications about resources I cannot exploit."*). Many believed that subscriptions could include the characteristics of a resource's license of a resource as a parameter. Others suggested that a

user should be able to contact the VDI owner and ask for a license. Everyone agreed that it should be possible to add or remove a license for a VDI, even when it had already been published.

Ease of use and learning curve: Most users found the interface easy to use and to understand (*"In a general way, I liked to use the application; I found it intuitive and easy to use."*). All users found it easier to use it as they became more familiar with it. However, a couple still found it difficult after two sessions, even though the second session was easier than the first. Two users forgot their password after the first session, and criticized the lack of a password recovery facility. All participants found the registration process very easy. Some suggested the registration form should include more information, such as a photo, city or country of residence, themes of interests (with tags), a security question, etc.

Improving CONVERGENCE: Many participants saw the application as a tool for organizing, and managing video content that interested them. For this reason, they thought that it should be possible to add personal (and private) comments to resources and notifications. (*"What I missed in this application was the possibility to comment the possible uses I could have of the videos I found"*). Another suggestion was to add tools to organize notifications into private folders (i.e. the possibility of organizing the "Anthology" and "My Subscriptions" sections is the same way as they might be organized by a bookmark manager in an Internet browser) (*"When I make several subscriptions, I'd prefer to organize them by subject instead of having to deal with long lists."*). Many participants believed that the application had potential as a collaborative tool. They would also have liked to see support for a broader range of video formats. They suggested that the application should be enhanced so that users could contact each other to request licenses, or to ask questions or comment on an analysis. This kind of functionality could take the form of a forum or a private chat.

Best features: For most participants the best parts of the application were the subscription features. Most also believed that the application made it easy to find content and resources of interest (*"I can see in a glance my subjects of interest and their corresponding resources."*), and that the tools it provided for this purpose were more powerful than those available on the current Internet. Some believed that it could be a very powerful collaborative tool, for video analysis *("the CONVERGENCE system could help during the analysis workflow: a video producer publishes a video, then an analyst makes an analysis by segmenting the video, then another analyst proposes a domain –dedicated analysis of this segmentation, and another translates the analysis: every actor could be notified of what the others have done.")*. Many believed that CONVERGENCE could guarantee security and content quality when content is analysed collaboratively by a group of users. In their view, this is a real innovation with respect to existing market solutions. *("I'm very dissatisfied with existing solutions for videos. I can find different video material on YouTube, but in too low quality, with the wrong metadata. That's why I don't want to publish my videos on YouTube: they are research videos*

and it would make no sense to mix them with other kinds of content.") Another aspect appreciated by users was CONVERGENCE protection of intellectual property (*"CONVERGENCE licenses and protection for authors' rights and intellectual property are fundamental.").*

Worst features: For all the participants, the worst feature of CONVERGENCE was the lack of functionality to remove, modify or undo videos, analyses, licenses, subscription, etc. once they had been defined (*"If I change my mind and want to change my subscription parameters, I cannot easily go back, I am blocked. The system lacks flexibility."*). In their view, this is an essential feature for any user of any application. It was also the main reason given for not recommending the application to a colleague. Many users also pointed to the application's failure to provide them with adequate information when they were browsing content: for example, they found it impossible to re-find the licenses they had created for VDIs, or the parameters for their subscriptions. Users suggested that analyses should include additional metadata describing subjects, actors, themes, concepts, and videos (duration, format, size, etc.). **(***"What I really missed was knowing who made an analysis (...). I 'd always prefer to use an analysis published by (...) a person I trusted than someone I don't know.").*

4.2.2.3 Augmented Lecture Podcast (LMU)

This scenario is based on a web-based lecture podcast application that enables students to revise lectures by watching video podcasts with synchronized slides. CONVERGENCE provides a common basis for collaboration and information exchange, allowing the lecturer to update one file and keep all related files in sync, and ensuring that students always have access to the latest versions. Students who have downloaded files onto their local devices are notified of changes as soon as they connect to the CONVERGENCE network. The scenario involves the following categories of user:

a) Lecturers can create, store and publish Lecture Podcast VDI sequences, which are updated whenever videos or slides are modified, and retrieve statistical information about the use of Lecture Podcast VDIs (statistical information is collected by the service provider).
b) Students can search, subscribe to, download and watch Lecture Podcasts, receive notifications whenever the service makes available a new podcast or a new version of a podcast, create and publish annotations to specific portions of the podcast, receive notifications about published annotations, delete annotations.

4.2.2.3.1 End-User Sample

The LMU trial was the largest in the second round of trials. In total sixteen student and eleven lecturers completed all their assigned tasks. Thirteen students and eleven lecturers attended the focus groups. All of them used CONVERGENCE in at least two sessions, Twenty-two

tested the system three times and three used the software four times. Nine users had a first degree, a master's degree or a Ph.D., thirteen has a high school leaving certificate. All used computers regularly at work and in their leisure time, usually for many hours a day. They considered themselves to be highly skilled in using PCs and the Internet (4.6) as well as Smart phones and tablets (4.4).

Members of the two groups tested different aspects of the application. In what follows, we will report their results separately.

4.2.2.3.2 Results in a nutshell

Both groups found the interface acceptable, especially with respect to resizing windows, colours and layout but also reported a number of usability issues and bugs. The main suggestions were: improved compatibility with other browsers and third party software, a better user interface, support for new video formats, and better help.

In general, participants found the idea and concept of CONVERGENCE interesting. However, they also described it as difficult to understand, complicated, and inflexible. If participants were to describe CONVERGENCE to people without a computer science background, they would describe it as a new kind of Internet that aims to provide better handling of resources, ownership rights and security. However, they found it hard to find these features in the application they tested.

Participants made positive remarks about the application's ease of use and about its usefulness for their work. Students, in particular, considered CONVERGENCE to be slightly better than the systems they use at present.

4.2.2.3.3 Result in detail (Students)

The student sample consisted of sixteen participants aged between 21 and 30. All were undergraduates. All claimed a high level of skills in using computers (4.3/5.0) and the Internet (4.4/5.0), and reporting that they used these technologies intensively in their studies and during their leisure time. Self-reported skills in the use of tablets and smart phones were at the same level (4.3/5).

4.2.2.3.3.1 Post Trial questionnaire 1

Task 1: Register and login to application

Participants were instructed to create a new account, filling in their first name, last name, email, password, etc. Users did not experience problems during the registration/login procedure, and gave it a high mark for ease of use (4.6/5.0) but a lower rating for general quality (3.8), error handling (3.0) and layout (3.8). They suggested the system should provide better feedback when a registration is successfully completed. At the moment this feedback is

not clear. One user noted technical incompatibilities with some browsers (particularly Internet Explorer 9).

Task 2: Search for lecture podcast

Participants were instructed to search for a podcast, after which the system displayed the results from the search. Participants were instructed to select a podcast that interested them and download it.

Only two of thirteen participants reported that they had already used this kind of search function in their work (LMU suite or iTunes). For these users, CONVERGENCE was comparable to the other software they had used. They gave it a score of 4. Participants considered CONVERGENCE's functionality useful for their work (3.7/5). The application received very high marks for ease of use (4.7,) speed (4.7) and responsiveness (4.1). However, participants were less satisfied with error handling (3.1) and layout (3.5). The scored assigned for overall quality was 3.9. Two users would have liked to be able to download the podcast in other multimedia formats, in particular AVI, MOV and MPEG4. Another suggested improvement was to provide more filters for search results (name of author, expiry date, professor etc.) and to provide extra advanced search functionality. Participants also suggested adding an automatic completion feature for text fields.

Task 3: Watch lecture podcast

After searching for the podcast, users were instructed to watch it by clicking play. While watching, they could navigate by clicking on annotations or the progress bar or the slide and chapter numbers shown in the dropdown menus.

Four participants had already used this kind of application. The software they reported using included iTunes, lecturer' websites, LMU suite and QuickTime. Users reported that the CONVERGENCE application was slightly better than the services they had used previously. The mean score assigned to this CONVERGENCE functionality was 4.5. Most believed that this functionality was useful for their studies, giving it a score of 4.1. However 69% of participants were unimpressed by the video/audio quality, commenting that the resolution was too low and that the video was not sufficiently "fluent". Some students also noted problems with the audio. However, they gave the application good marks for ease of use (4.5), speed (4.5) and responsiveness (4.0). Marks for error handling (3.2), for general quality (3.8) and for layout (3.4) were lower. Participants would have liked to see videos and slides in full screen mode and to have had more control over the display (e.g. the possibility of resizing the display window). They would have also liked to choose between using the podcasts in audio or in video mode.

Task 4: Annotate lecture podcast

Users were instructed to annotate the podcast by clicking on a slide. The system would then respond by creating a dialog for the annotation, including a field for the expiry date. To publish the annotation, the user had to click on the SAVE button.

Only one of the participants had previously used a similar service, called Backstage, and considered CONVERGENCE slightly better than this software. For twelve participants, this functionality was completely new. In general, they thought it would be very useful for their work. Participants gave the application a good score for ease of use (4.2), speed (4.5) and responsiveness (4.0). Marks for layout (3.5) and error handling (3.2) were lower. Participants gave the application a good mark for overall quality (3.9). Several suggested the addition of new functionality allowing users to decide whether to display or hide annotations. One noted technical incompatibilities with certain browsers (particularly Internet Explorer 9).

Task 5: Subscribe to Annotation

Participants were instructed to subscribe to an annotation by clicking on the subscribe icon link next to the title. Participants would then be notified of replies to their annotation.

Only one of the participants had previously used a similar service. This user considered that, in this area, CONVERGENCE provided better functionality. In general, participants did not found this feature to be particularly useful for their work (3.1). However, they were satisfied by its speed (4.5) and found it relatively easy to use (4.5). The worst mark was for error handling (3.2). Users gave the application slightly better scores for layout (3.5) and for overall quality (3.7). Suggestions for improvement included requests for an “unsubscribe” button, and the visualization of a feedback message after a successful subscription.

Final notes

All participants believed that CONVERGENCE provided functionality that could be useful for their work. Twelve would be willing to use it and considered that CONVERGENCE could increase their productivity. Only one participant stated that CONVERGENCE was unattractive and that she would not use it, even if it improved. Overall, participants gave CONVERGENCE a score of 3.7/5. Only three experienced problems during the session (problems with login/registration; problems while subscribing to annotations). These users found the online help useful but not completely satisfying.

Supervisor’s report

The supervisor reported a server break-down (hardware failure) during the last day of the trial Participants who could not complete their session were asked to repeat the session on one of the following days.

4.2.2.3.3.2 Post Trial questionnaire 2

All participants tested CONVERGENCE in at least two sessions, filling in post-trial questionnaire 2, thirty-two times. Ten reported technical problems. The biggest problem was with the login procedure which denied users access, when they tried to re-enter the application after a logout. One user also noted problems with the watch functionality – and was not able to understand which button to click on. In general, however, the sample found the system easy to use: 85% recovered previously loaded data without error. None had any difficulty in remembering how to use it. 47% reported that they could now use the system better than in the first session. Participants gave the application high marks for layout (3.6), ease of use (4.1), overall quality (3.7), speed (4.2) and responsiveness (3.7) - results very similar to those for the first session. However, they gave it poor scores for the help system and for error handling (2.9). Only five users changed their assessment with respect to the first session (one improved her mark when she finally understood how to use it in her third session with the software). Ten would be willing to use the software, which they thought could improve their productivity. Five preferred other similar services. For these users the main problem with CONVERGENCE was poor layout and a lack of visible innovation. Five participants would consider using it in their work but only if it improved first. For these users the most important improvements required concerned Overall quality and Ease of Use. These responses matched those in the first session. Students' overall assessment of the application (3.6) was the same as in the first session.

4.2.2.3.3.3 Focus Group

All participants who filled in questionnaires also participated in the focus group.

Navigation: Participants cited several navigation problems. When they wanted to search for podcasts they were forwarded to the registration site and had to log in again to continue working. Some also reported that their username has not been recognized by the application at first and that they had to press the "back" button to log in. One participant reported that she did not read the information box at first and then could not figure out where to leave comments. (*"Sometimes it was a bit confusing").*

Help: Participants wanted more guidance while using the application. *(**"**I did not know what the expiry date was for when creating a comment. Some help or hint would have been useful"*).

Layout/graphics: One participant reported that some application buttons were hidden by the podcast video window and were not easily accessible. Another problem was that the video window was initially too small. After resizing, users had to scroll to see comments.

Usability problems: Participants reported several usability problems. Some had to spend time searching for the comments section, which they thought was poorly positioned. They

would also have liked to receive some kind of feedback after leaving a comment. One suggested that comments should be shown only if the user is interested in them (*"Not all comments were interesting for me. So maybe one should be able to fade them out.")* Another problem with the comments section was that users did not understand the meaning of the numbers in front of the comments (*"I did not know the meaning of the time in front of the comment. Was it the expiry date?")* One reported that all text fields in the registration form had to be filled out again if you forgot something when you filled it in the first time. Another participant complained that podcasts they had already found the previous day had to be searched for again the next day.

Performance: Most participants stated that the performance of the application was good and rather quick. One participant reported that the video had not run smoothly. However a more typical comment was: *"Everything went fast"*.

Robustness/ technical problems: One participant could not test the application properly because it would not work on Linux or on Windows XP (the operating systems she was using). On Windows 7 with the Chrome browser the users could only play a couple of minutes of the podcast video. However there were no problems with slides. Another participant reported that the video did not run smoothly and the sound did not work at all.

Privacy: Participants knew they were taking part in a study, trusted LMU and did not express any privacy concerns. Other users were invisible to them (*"I could not see other users. So it was ok; "I was not worried, that something would be public").*

Security: Participants did not give any feedback on this topic.

Ease of use and learning curve: In general, participants found the applications easier to use on the second session (*"On the second day I knew where I can leave comments.")*. However, they also reported that the instructions popped up every day, for no particular reason. One user reported that she never figured out how to use the subscription functionality. Participants tended to think that CONVERGENCE did not differ much from LMU's existing podcast application or from other applications for slide presentations (*"I thought I was testing the podcast application because of new features and not because of CONVERGENCE").*

Improving CONVERGENCE: Participants appreciated the fact that they could use CONVERGENCE without realizing it. (*"CONVERGENCE is an all-in-one tool).*

Best features: According to participants, the most attractive aspects of CONVERGENCE, as a system, were content synchronization and improved control over published content and expiry dates. The best features of the application were the possibility of fast-forwarding videos and adding comments by clicking on slides. Some participants also found the search function good.

Worst features: Some participants found video quality to be unsatisfactory and others reported problems when subscribing to comments. According to one, the comments did not

display and this led to confusion. Several participants argued that allowing authors to control their posted content and set an expiry date could prevent other users from viewing content they might otherwise find useful. Complete control might also be a problem if a comment violated somebody else's rights or included untrue information about somebody else. Some participants suggested that if content were held in the cloud, security would need to be very strong.

4.2.2.3.4 Results in detail (Lecturers)

The lecturer sample consisted of research assistants at LMU and TU (Technical University of Munich) and included eleven participants with ages between 21 and 40 (three were older than 30). All had a university degree and all considered themselves as highly skilled with computers (4.7) and tablets, which they used intensively in their work and in their free time. Ten lecturers tested CONVERGENCE three times; one tested it in four sessions.

This group tested the CONVERGENCE Lecture Podcast Creator

4.2.2.3.4.1 Post Trial Questionnaire 1

Task 1: Register and login to application

As in the other trials, participants were instructed to create a new account, filling in their first name, last name, email, password, etc.

Participants were satisfied with this procedure, giving it a good mark for overall quality (3.9). The lowest score was for the error messages (3.2). Ease of use and layout also received positive marks (4.5 and 3.4 respectively). Suggestions for improvement included optimization of the procedures for specifying passwords, improvements in navigation functions, and a passport recovery function. Other suggestions included the addition of tools to check the strength of passwords and to confirm registrations via mail. One user suggested eliminating the default input values shown in the text fields.

Task 2: Create podcast components and lecture podcasts

Participants were instructed to create a slide or a video component. To do this, they had to select the location of the resource and add metadata (the title of the lecture, the name of the author and keywords) for each component and select an expiry date (optional tool). They were then instructed to view a list of slides and video components, select some of these and assemble the podcast, add a synchronization file (from their hard disk), add metadata for the podcast and publish it to the CONVERGENCE network.

64% of lecturers had previously created podcast components that used a range of different techniques (HTML, XML, text editors, FTP clients and self-developed software). In general, they considered CONVERGENCE to be better than the software they had used before, giving

it a score of 3.9. All participants reported that the application was useful for their work (3.7). Marks for error handling and layout were adequate but not high. The lecturers gave the application high marks for ease of use (4.2), speed (4.5) and responsiveness (4.4). Suggestions for improvement included better forms, better explanations for fields, more information on the file formats supported by the system and expiry dates, and previews of podcasts. One user noted compatibility problems with Firefox.

Task 3: Update and Revoke podcast components and lecture podcast VDIs

Participants were instructed to update or revoke a podcast. Users could modify files or change the metadata.

45% of lecturers had already updated or revoked podcasts using other software (Notepad++; XML, text editors, FTP clients, self-developed software). According to these users, CONVERGENCE was better than the software they had used previously. Overall, participants gave this function a score of 3.8. They all agreed that it could be useful for their work (4.1) and gave it a good score for ease of use (4.1), speed (4.5) and responsiveness (4.4). The scores for layout (3.6) and error handling (3.8) were relatively good. However, participants were unhappy with the error messages (2.6). Suggestions for improvement included feedback to confirm when a podcast had been created and new functionality to preview podcasts. Users were favourably impressed by CONVERGENCE expiry dates. The only lecturer who had met similar functionality in the past, had seen it in Moodle.

Final Notes

Eight out eleven lecturers believed that CONVERGENCE could be useful for their work; six considered that it had the potential to replace existing solutions; five would recommend it to a colleague, one might use CONVERGENCE but only if it improved. Only one user was seriously dissatisfied, complaining about too many bugs and inadequate error messages. The overall score was 3.4/5.

Five users experienced technical problems with podcast creation and with the update/revoke functionality; four considered that the error messages were not clear. The majority (8/11) believed that online help was very important.

4.2.2.3.4.2 Post Trial questionnaire 2

During the session, lecturers reported twelve distinct problems. The most serious involved the time required to send requests during the publication procedure, and the display of a list of keywords, even when they had not been defined. Four users reported an exception during the update task. The majority (62%) reported difficulties in finding data loaded in the previous session. However, they all remembered how to execute their tasks.

Nine lecturers reported finding it easier to use the software in the second than in the first session. However, only one changed the mark she gave the system. Eight considered that

CONVERGENCE was useful for their work, seven of whom would use the software immediately. Three disagreed, claiming that it was unattractive and had too many bugs. However, three could change their idea if the software was improved. Comments focused on features that were not yet implemented in this phase of development. There were also positive comments on the system's ease of use and good performances. The final score (3.3) was the same as in the first session. The supervisor observed that some users had experienced difficulties in updating their VDIs.

Supervisor's report

The supervisor did not note any major technical problems.

4.2.2.3.4.3 Focus Group

The lecturer focus group was separate from the student group.

Navigation: Participants stated that the navigation through the application had been good. Creating the podcasts had been easy. They commented: *"It was well designed."*

Help: Participants gave no feedback about this topic.

Layout/graphics: Participants considered the layout to be generally good, simple and pleasant. At the same time they noted that there were no major differences compared to LMU'S current podcast application. The only innovation introduced by CONVERGENCE was the possibility of setting an expiry date for podcasts. Another participant stated that even though the user interface was simple, the position and the purpose of the windows on the screen were not always clear.

Performance: The majority of the participants found the performance very good (*"I had no issues with the performance."*). They were also happy with the speed of uploads and updates, although one said that performance seemed slower on the last day. Just one participant reported that uploads had been slow throughout the trial.

Robustness/ technical problems: Some participants reported that they had received an error message while updating the podcasts on their second and third sessions. One participant who had used Firefox had not received the message. However, another participant reported she had not been able to use the application in Firefox at all. In Google Chrome, the application appears to have worked without major problems except for some unclear messages.

Usability problems: Participants reported problems while filling in keywords. The system proposed a long list of keywords, some from the author, and some from other users. This created some confusion (*"Suddenly, there were so many keywords.")* and participants viewed it negatively. Another usability problem was that participants had to select the files going into a podcast one by one instead of all together. There were also complaints that when they wanted to update a podcast they had to fill in the text fields all over again. Users considered

this as unnecessary work. Another problem was that the application did not give them enough information about who had uploaded the podcasts.

Privacy: Participants did not report any particular privacy concerns.

Security: Two participants said that they did not find the CONVERGENCE's application secure enough, or at least no more secure than other podcast applications. (*"It does not appear more secure than current solutions."*) One participant said that she could not estimate how secure it was in such a short trial.

Ease of use and learning curve: Participants reported that application became easier to use as they learned what to do and no longer needed to read the instructions. In these terms, the CONVERGENCE solution does not appear to be much different from other software *("I did not notice that there is a different system beneath the application.")*. Some participants said that they did not have much experience with podcast applications and therefore could not comment properly.

Additional features requested: Participants asked for better feedback when creating podcasts. They would also have liked to receive feedback on their progress while uploading podcasts (this was true only for participants, who could not see the progress indicator).

Best features: Participants had difficulties in naming the best feature of CONVERGENCE as a system, and felt that CONVERGENCE functionality was not explicitly visible. As far as concerned the podcast application, participants liked being able to add and delete podcasts, update copies of the content they that published, delete these copies and set expiry dates.

Worst features: For most participants, the worst feature was the need to upload the files required for a podcast, one at a time, rather than all at once. Some participants also said that the trial scenario did not allow them to properly test the expiry date feature, which they believed to be essential. They also stated that the podcast scenario was probably not the best choice for testing the features of CONVERGENCE.

4.2.2.4 Smart Retailing (UTI)

This scenario reproduces a smart retailing supply chain for electronic products. In the scenario, CONVERGENCE provides users with a wide range of services and operations. The users and beneficiaries are Retailers and Customers. Retailers subscribe to products from Manufacturers and advertise the products they sell through promotions, special offers, etc. Customers subscribe to, search for and compare products. Additional services improve Customers' shopping experience, allowing them to subscribe to information on specific products, receive notifications about the products and about sales events and immediately look up information on a product by scanning its barcode.

4.2.2.4.1 The Trial

The sample for the UTI trial consisted of seventeen users of which eight were aged between 21 and 30 and nine between 31 and 40. All participants had a university education. All reported that they had good computer skills and that they used computers regularly, mainly for work. In general, they were more familiar with traditional systems (Computer 3.8, Internet 4.1) than with smart phones and pads, which they nonetheless knew quite well (3.1). The scenario for the trial included fifteen (simulated) customers and two retailers. The roles of the fifteen customers were played by five engineers, one translator, three contract administrators, one project manager, one technical writer, one analyst, one tester, one Assistant Manager, and one software architect. The two retailer roles were played by two sales staff, external to the project. Fifteen participants tested CONVERGENCE twice; thirteen tested it four times.

4.2.2.4.2 Results in a nutshell

Customers and retailers both appreciated CONVERGENCE, even though they were aware that it is still in an early stage of development. All participants would be willing to use a later version of CONVERGENCE when it becomes available, making many suggestions about how to improve the application. On the customer side, the main demand was for improved search functionality, allowing users to personalize the way results were displayed. As far as concerned retailers, the most important suggestion was that the application should provide more information about customers and their subscriptions (customer preferences, previous purchases etc.). There were also requests to improve the layout of the results review and to provide better links to statistics.

All participants requested improvements in usability and a more attractive layout to help navigation. For many of them, CONVERGENCE still seemed too complex to use in their work, even if they appreciated its innovative aspects and technical quality.

4.2.2.4.3 Results in detail

4.2.2.4.3.1 Customers Post Trial Questionnaire 1

Task 1: Register and login to application

As in the other trials, participants were instructed to create a new account, filling in their first name, last name, email, password, etc.

Users were satisfied with all aspects of the procedure, giving it a mark of 4.3 for ease of use, 3.5 for error handling, 3.6 for layout and 3.7 for general quality. Nonetheless, two participants experienced technical problems and one customer had to reload the page and fill in a second time. User suggestions included the introduction of email confirmations for registration, a field where users could confirm their password, a tool to modify passwords, better error

messages, and better feedback after a successful login. One user asked for animation effects in the registration window. Another recommended the introduction of an advanced system to protect users' passwords.

Task 2: Navigation through products

Participants were instructed to browse products with specific features or characteristics.

Five participants had already used similar services provided by miniprix.ro, carturesti.ro, emag.ro, amazon.com, price.ro, asos.com and IBM Visual Store. In general, they believed that CONVERGENCE was slightly better than these services and thought this was a useful function; the majority (11/15) liked the CONVERGENCE browsing system considering it easy to use, accessible, and well designed. Five criticized the lack of logical structure in the list display and the small size of the photos. However, the majority (14/15) stated that the list of offers was easy to use. Participants gave positive scores to all aspects of the service, giving it a mark of 3.5 for error messages, 3.4 for layout, 3.7 for general quality, 3.8 for speed and 3.9 for responsiveness. Ease of use received a very strong mark of 4.1. The participants made some suggestions for improvement. These included the use of a tree structure to help users understand the level they had selected, a tool to show price trends and an online help function. One user found a technical problem: when the list of offers was too long, the application displayed an empty table.

Task 3: Subscribe to a product

Participants were instructed to subscribe to product type VDIs with specific features or characteristics.

Six reported that they had already used services providing this kind of functionality; expressing the view that CONVERGENCE was slightly better than other services they had used (flu.ro, asos.com, cel.ro). The majority (12/15) thought that this was a useful function, giving it very good scores for ease of use (4.3) and speed (4.1). Marks for error messages (3.5), layout (3.7), general quality (3.7) and responsiveness (3.8) were positive. Just one user complained about the time it took to complete the subscription process. Participants made several suggestions for improving the system, which they suggested should provide support for multiple subscriptions, display information about the total number of subscriptions in the system, display a warning (popup, text box, etc.) before completely deleting a subscription, and display more information in the list of subscription. Another user suggested that it would be useful to have a Subscription History Window where users could see their previous subscriptions.

Task4: Digital information about products promotions

Customers were instructed to browse product VDIs to find a detailed description of a particular product and relevant special offers.

Eight customers had already used similar systems (elefant.ro, kuponiada.ro, emag.ro, amazon.com, asos.com, hells.com, cel.ro) and believed that CONVERGENCE was slightly better than these services. Users strongly appreciated this feature, considering it to be useful (4.1). They also gave it high scores for ease of use (4.3), error messages (3.8), layout (3.8), general quality (3.9), speed (4.1) and responsiveness (4.1). Suggestions for improvement included the introduction of links between special offers and products and better email notifications.

Final notes

Thirteen participants thought that CONVERGENCE was very useful and that it could make a positive contribution to their productivity. The two users, who thought that CONVERGENCE was not useful, blamed slow performance and critical problems in displaying subscriptions. 80% of sample would consider using CONVERGENCE instead of the applications they use currently. Twelve would recommend it to a colleague. The overall mean score was 3.7. Nine users reported that they had sometimes chosen the wrong button during navigation, but only rarely. Six reported technical problems, three during login/registration.

4.2.2.4.3.2 Customers Post Trial Questionnaire 2

In later sessions, all users remembered how to use CONVERGENCE and 69% reported that they had improved their navigation skills. As in the first session, users appreciated the application's ease of use (4.2), speed (4) and responsiveness (4.0). Participants considered other feature adequate (Help: 3.5, Layout 3.4, Error Messages 3.5, Overall quality 3.7). Only 12% of users changed their overall mark for the application. User comments in the questionnaire suggested that their main reason for using the system was its technological quality and novel concepts. However, users would have liked to receive better information about the system. A majority (92%) preferred CONVERGENCE to the software they were using at the time of the study, considering it an important tool for their work. All these users would also recommend CONVERGENCE to a colleague. The minority was unsatisfied and would not use CONVERGENCE even if it improved. The final score (3.9) was similar to the score recorded in the first session. Just two participants reported time out problems and inadequate error messages during login and registration.

Supervisor's report

The supervisor did not note any major technical problems.

4.2.2.4.3.3 Retailers Post Trial 1

Task 1: Register and login to application

Like customers, retailers were instructed to create a new account, filling in their first names, last names, emails, passwords, etc.

Retailers were satisfied with the procedure giving it a mark of 4.0 out of 5.0. However, two suggested adding more text fields in the registration form. This would allow them to obtain more information about their customers.

Task 2: Navigate through products

Retailers were instructed to browse the Product VDIs in their database, searching for a product and displaying details about it. Finally they were asked to choose a category or product and make a subscription to it.

One had already used navigation systems on price comparison websites, ecommerce catalogues and coupon websites and thought this was a useful functionality (4.0). Both thought that CONVERGENCE could be very useful for their business, giving marks of 4.0 to most of the items in the questionnaire (ease of use, error messages, general quality, speed and responsiveness). The only item to receive a lower mark (3.0) was layout. One user thought there were too many categories and subcategories in the browsing system and that this made it difficult to navigate. None of the users had any difficulty in navigating the list of special offers. Retailers suggested that the system should be enhanced with a new advanced search function for categories and products and better options for offers.

Task 3: Promotion of products over the Internet

Retailers were instructed to import a Manufacturer Product VDI into their databases and to re-publish it to the CONVERGENCE network, as a way of advertising a product.

The retailers believed this function was essential for their work and considered that CONVERGENCE was slightly better than other services providing similar functionality (shopmania.ro, compare.ro, allshops.ro, cuponix.ro). The overall mark for this function was high (4.0/5.0). Ease of use, speed and responsiveness also received a mark of 4.0. Marks for error messages (3.5/5.0) and layout (3.0/5.0) were lower. Users suggested that one way of improving the system would be to allow retailers to customize the layout for special offers, and to provide more information (pictures, details, etc.).

Task 4: Customer Feedback Management

Retailers were instructed to monitor customer feedback on a specific product.

The retailers considered this function essential for their work and believed that CONVERGENCE was slightly better than other software they had used in the past, such as price.ro and shopmania.ro. However they also reported that illegible text in the images made the statistics impossible to understand. In general, they considered the implementation of this feature to be more than adequate in terms of ease of use, error messages, overall quality, speed and responsiveness, all of which received a score of 3.5. However, they were not happy with the layout, giving it a score of 2.0/5.0. One suggested a complete redesign of the list of special offers.

Final Notes

Retailers stated that they had made relatively few navigation errors. According to the comments, it is easy to push the back button without meaning to. One participant reported compatibility problems with Firefox. The two retailers believed that the error messages and online help functions were useful and did not report major technical problems. In general, they preferred CONVERGENCE to other products and expressed the belief that it could improve their productivity at work. They would also be willing to recommend it to colleagues. The final score was 4.5.

4.2.2.4.3.4 Retailers Post Trial Questionnaire 2

The two retailers repeated their test only once. Afterwards, they reported that they had always remembered how to perform specific tasks. Ease of use, speed and responsiveness all received a score of 4.0. Other features (help, layout and general quality) obtained a result of 3.0. Retailers reported that they now found it easier to navigate the system than in the first session and would consider using CONVERGENCE instead of the services they used at the moment. In their opinion the software could make a major contribution to their productivity and they would both be willing to recommend it to a colleague. The final score was the same as in the previous session: 4.5.

4.2.2.4.3.5 Focus Group

All participants took part in the focus group.

Navigation: Everybody said that the navigation tool was hard to use and not very useful. (“*The interface should point the path to the top before you click somewhere because you easily forget where you started*.”)**.** Users complained about the excessive number of products shown in the product’s list. One user said that the ontology mixed products and categories and needed reorganizing. *(“The products categories should be reorganized; some products are in inappropriate categories. The product categories need a better structure.”)*

Help: Many users requested a help manual or online help. *“It would be useful if every control was associated with contextual help”).* One person did not initially know what to do and searched for a user manual. *(“Better information about it (CONVERGENCE / the app) should be provided: for example by providing a help function or an information section on the site.”)*

Layout/graphics: Everybody said that the layout/graphics of the application need additional work. One user commented. “*The platform is useful; I just don't like the layout*.” According to another “*It’s not user friendly, the product details are far away from the navigation system and you always have to go around with the mouse*”. Many users thought that products and product offers should be more visible with bigger fonts, and larger, better quality pictures.

Performance: The majority of the users stated that the performance of the system was good. Apart from some remaining bugs the system responded to their demands quickly. (*"It's fast. You wait a bit at the beginning but after that everything loads quickly."*). However some reported that the application is too slow and sometimes seems to freeze. (*"The deletion of a subscription takes a very long time, around 10s."*).

Robustness/ technical problems: The review showed that the application needs additional work. The users remarked that CONVERGENCE still had some bugs and some functionality did not work properly in all browsers. One user reported: *"I can't enter in the application if I use IE8."*

Privacy: The sample considered the application's privacy functionality positively. "(Retailer*): I asked for the possibility to see the emails in the statistics, along with the interests of the user making the subscription, but the system doesn't allow this, as it would violate the user's privacy. As a conclusion, I think the application respects the privacy of every user."*). User subscriptions were individual, not public. This ensured that customers' interests remained private. (*"A user can't see the subscriptions of another user. That makes me feel that I have some privacy in this application*.").

Security: The users believed that the application was not very secure. Participants suggested a lot of improvements to make the system safer. Suggestions included *"Safer login e.g. captcha, email confirmation from inside the registration process"* and the addition of a field for password validation. One retailer suggested that the system should be enhanced with a validation system, allowing an administrator to validate the accounts of individual users.

Ease of use and learning curve: Everybody stated that the system was easy to use, and remembered how to use it, the second and third time they entered the application. **(***"Users who had used web applications understood how to use the application very quickly.") or* **(***"At first I understood everything I had to do because it's simple to get the idea, the only exception was the matching mechanism, which is a little trickier and I needed an explanation of what happens there and what I am supposed to expect.").*

Improving CONVERGENCE: Some participants proposed to create a customizable search function. Other users proposed an advanced search mechanism, which would allow the user to refine her search. Another idea was to add more subscription/offer attributes, beside price and expiry date.

Some participants suggested hiding the subscriptions list, showing it only when the user requests it by clicking a button. During the discussion, the idea emerged of an adaptive recommendation system, which could recommend products based on customers' tastes, and buying habits and the products' popularity. After a customer had subscribed to a product, the system would display information about other products often bought by customers when they buy the first product (e.g. the system could show shoes and some bags when a customer buys

a dress,). One person suggested that the application should provide more statistics, including statistics for customer product preferences and for how many times customers open promotional emails. Other suggestions included the possibility of sorting the list of special offers, showing product details in the subscriptions list, and grouping of subscriptions/offers by category. One user suggested SMS notifications for new matches.

Best features: All the users agreed that the subscription system and the matching mechanism were CONVERGENCE's best features. They appreciated the fact that the application shows the subscription list, avoiding the need for users to look up their favourite products every time they login. Users also liked the email notifications, saying they are very useful. They were impressed by the way the matching mechanism displayed only offers for products or categories, which they had shown an interest in (*"As long as the application notifies me only about what I'm interested in, I like it!"*). Participants said that receiving information about what they actually want is the first reason they would use an application like this. They also discussed the idea of having a single application that presents all possible products in the world. Opinions were divided. However, the majority of users was sceptical about the future of this functionality (*"The retailers who would benefit from this system would be those that don't have a monopoly. The big retailers (...) would reject it because it would not differentiate them from other retailers."*). In general, they thought that retailers would not be able to agree on standardized product descriptions. This is especially true for those who use many differentiators in the description of their products.

Worst features: Most of the participants said that the layout was not as appealing as that of other e-commerce sites. The interface needs some additional work, particularly as regards graphics and the use of colour. The strongest criticism focused on the products ontology tree. (*"The search functionality is more useful than the navigation tool but it's too simple."*). Users thought the tree showed too many products and that it was difficult to navigate (*"It's hard to use, you forget where you started and you don't have the big picture of the store."* or *"If there are too many products in the system you can't see anything."*) Most users preferred to use the search functionality and wanted more options to help them. One suggestion was to extend the "product details" section. Users would also like to be notified when a new product is added in the system.

Just one user found the subscription mechanism difficult to understand. One of the retailers complained about having to insert data about special offers manually. She reported that she usually had to insert a lot. The same user also asked for a check box system, allowing her to erase more than one offer at once. Users would appreciate the possibility of sorting offers by their expiry dates. This would make it much easier to delete those that had expired. One said that the email layout was too wide and that emails took too long to read.

4.2.2.5 Smart Retailing (WIPRO)

This scenario, similar in many respects to the scenario tested in UTI, involves the application of CONVERGENCE to a smart retailing supply chain for electronic products. The users and beneficiaries are Manufacturers and Retailers.

Manufacturers create, store, publish and certify Product Type VDIs containing information about their products (e.g. the name of a product, a general description, physical and technical characteristics of the product). Manufacturers also benefit from new flexible mechanisms for selling products to authorized Retailers.

Retailers subscribe to Product Type VDIs published by Manufacturers; create and publish Retailers' Product Type VDIs, augmenting Manufacturer Product Type VDIs with information about Retailer promotions, special offers, sales, etc., and create Product Instance VDIs including new information such as customer ID, serial number and warranty details. The use of Product Instance VDIs when selling to customers gives retailers better control over the products they sell and better knowledge of their Customers.

4.2.2.5.1 Trials

The group consisted of seventeen participants, including two Business Analysts, one Business Consultant, three Analysts, seven Engineers, two Managers, one Graphic Designer and one IT worker, all employed at Wipro. Eight participants acted as manufacturers and eight as retailers. All participants were external to the project and all participated in focus group. Almost all had a university or some other form of higher education. The group included ten participants in the 21-30 age group and six aged between 31 and 40. All had a good level of computer skills, using computers regularly, mainly for work. Like most participants in the trials, they reported very good computer skills (4.7/5.0) and good skills with tablets and smartphones (4.6). Each user used the system in two separate sessions.

4.2.2.5.2 Results in a nutshell

All of the participants were very attracted by the CONVERGENCE concept. Most of them described CONVERGENCE as a big step into the future, helping users to handle the huge amount of information now available on the Internet and speeding up the flow of information.

Everyone was also very attracted by concept of the VDI, as a digital data container. Participants also saw several other advantages in the CONVERGENCE offering, in particular the integration of a broad range of services and technologies into a single platform. Regarding the retail business area, everyone shared the opinion that CONVERGENCE brings a new approach to retailing and indeed to any transaction-centric businesses, not only by creating product VDIs with reliable information from a single source, linked to an easy-to-use mechanism for copyright management, but also by helping to catalogue products inside

organizations and to integrate catalogue information with core retail applications like RMS (Retail Merchandise System). From a consumer point of view, CONVERGENCE will offer an improved shopping experience and the optimization of post-sale services.

The main concern of the participants was CONVERGENCE security, and in particular the security of information within the VDIs when they are transferred between different entities and users. Participants affirmed that they would like to examine the way CONVERGENCE intended to use smart cards and to encrypt and decrypt data and the way these features would be integrated with the system applications.

Other interesting comments concerned the CONVERGENCE implementation of a product ontology and the way it was used in the application. The main requests were for better-structured ontologies, which could be shared between all actors in the value chain (manufacturers, retailers and customers), a more attractive human computer interface and fewer bugs.

4.2.2.5.3 Results

4.2.2.5.3.1 Manufacturer Post Trial Questionnaire 1

Task 1: Register and login to application

As in the other trials, participants were instructed to create a new account, filling in their first name, last name, email, password, etc.

The three participants playing the role of manufacturers gave this function an average mark of 4.1 for ease of use, 3.5 for error messages, 3.8 for layout, and 3.8 for general quality. As in other trials, users suggested the introduction of email validation during the registration process, the addition of a field in which users could confirm the password, as well as automatic detection of missing data and better explanations and support.

Task 2: Create product VDI

In this task, the user was instructed to fill in the create product form with product information, including the dimensions, weight, features, characteristics, barcode, of the product. During the procedure, the user also has to define the type of license offered to Retailers who subscribe to the product.

Five participants had already used services providing this kind of functionality (UBL Swinger and ORMS), which they found very useful for their work. In general they perceived these services to be slightly better than CONVERGENCE. Nonetheless, participants gave CONVERGENCE a good mark for layout (4.3) and reasonably good marks for ease of use (3.3), error messages (3.4), overall quality (3.5), speed (3.6) and responsiveness (3.6). One user suggested adding more descriptive text about the product, showing the date on which the

VDI was published. All thought that the list was easy to modify. One user suggested that the product description could be improved by implementing a price field and checking for links to other products in the list; another requested a feature to automatically detect missing fields.

Task 3: View product

Once the user had created the product VDI, she was instructed to display it.

Participants considered this CONVERGENCE's functionality more than adequate, giving it a score of 3.6. The manufacturers believed that this feature was essential for their job (4.4/5). However, the scores it received were only moderately good: 3.6 for ease of use and for error messages, 3.4 for layout, 3.6 for overall quality, 3.4 for speed and 3.4 for responsiveness. In the comments several users asked for a more attractive layout.

Task 4: Publish product VDI

The user was instructed to publish the VDI to the CONVERGENCE network. To do so, she had to select the product VDI and fill in a form containing: the name of the product and tags for retailers, (e.g.: Samsung LCD, HP Laptop). As part of the procedure she also had to define the license terms for retailers subscribing to the product. An additional option allowed her to define the type of user who would be notified when the VDI was published.

None of the participants had ever used a similar service before. However they all though it was useful (3.6) and gave it a very good score for ease of use (4.4), layout (4) and general quality (4). Scores for error messages (3.6), speed (3.6) and responsiveness (3.6) were lower.

Final notes

Seven out of eight participants would consider using CONVERGENCE as the main tool in their daily work. Only one– a manufacturer – rejected this option – claiming that there were too many problems at this stage in the implementation and that the system was not sufficiently popular. However, she would be willing to reconsider if the system improved. According to participants, CONVERGENCE could make a useful contribution to their productivity; a majority would thus be willing to recommend it to their colleagues (5/7).

All the participants though that the online help was useful and 50% thought the error messages were important. Only one user reported that she had made mistakes when using the buttons but said that such mistakes had been rare. Three users remarked technical problems in tasks two and four.

The final score (3.6/5) was relatively positive.

Supervisor's report

The supervisor did not observe any serious problems during the trial.

4.2.2.5.3.2 Manufacturer Post Trial Questionnaire 2

Eight manufacturers tested CONVERGENCE in a second session. One user could not access the list of information during the product creation process. Four found bugs during the publication of a product. One suggested that this could be due to the use of special characters. All users found it easy to remember how to perform their tasks. The manufacturers gave CONVERGENCE moderately good marks for the online Help (3.4), Layout (3.9), Error Messages (3.3), Overall quality (3.9), speed (3.3) and responsiveness (3.5) and an extremely good mark (4.4) for Ease of use. The majority (5/8) noted that their navigation skills had improved from the first to the second session. 63% changed opinion on the system during the second test. After the second session only three manufacturers still believed that CONVERGENCE was ready to replace current systems. However, five would use it, if it improved. The negative comments focused on lack of users and the unattractive design. Nonetheless participants considered that CONVERGENCE could contribute to their productivity at work. The final mark (3.4) was the same as for the first session.

4.2.2.5.3.3 Retailers Post Trial Questionnaire 1

Task 1: Register and login to application

This procedure was identical to the procedure for manufacturers.

Users gave it a good score for all items on the questionnaire (Ease of use: 3.9, layout: 3.5, overall quality: 3.9, ease of recovering from errors: 3.5). The only suggestion for improvement was to add a system for recovering lost passwords.

Task 2: Subscribe to a product

Participants were instructed to subscribe to product type VDIs with specific features or characteristics.

None had ever used this kind of service but they believed that it would be useful for business. Expiry dates were a novelty: however only four users considered them to be useful. There was general agreement that the quality of the service was fairly good (3.6). Ease of use received a very strong mark (4.4). However users gave a poor score to the error messages (2.6), layout (2.6), speed (2.4) and responsiveness (2.6). Suggested improvements included more information about text fields, and the addition of hints about keywords often used during the subscription process.

Task 3: Get a subscription match

When a customer buys a product, a notification appears in the retailer's list. The participants were instructed to browse it and check matches with their previous subscription.

None of the participants had ever used this kind of service before and none considered it useful for their business. The only aspect of the function that participants considered

positively was ease of use (3.4/5). On other items in the questionnaire, users returned scores between 2.3/5.0 and 2.8/5.0.

Final notes

All of the participants believed that CONVERGENCE could be useful for their work. Four preferred CONVERGENCE to the application they had used in the past, thought that it could increase productivity at work and were willing to recommend it to others. For the others, the main issues were frequent technical problems, slow performance, difficulties in using the software, and the lack of a user base. Nonetheless they would be willing to use CONVERGENCE, even as their main system, if it improved.

Just two retailers remarked that they had occasionally made mistakes when using buttons. 50% would have liked to see an online help. Three users reported technical problems (one for each task/functionality). The final score was 3.3/5.

4.2.2.5.3.4 Retailers Post Trial Questionnaire 2

Eight retailers tried CONVERGENCE in two sessions. Two reported technical problem during the subscription procedure. However, none of them had forgotten how to use the system and six reported that they had improved their navigation skills during the second test. As in the first session, retailers gave the application a good score for Ease of use, (4.1), but lower scores for other items on the questionnaire (Help: 3.6, Layout: 3.3, Error messages 3.3, Overall quality 3.8, Speed 3.5 and responsiveness 3.5). 63% of sample changed their assessment of CONVERGENCE, and many reported that they now had a better understanding of the CONVERGENCE concept. Half would be willing to use CONVERGENCE instead of their applications. The others thought that CONVERGENCE was too slow and the design too unattractive. Users considered the software to be potentially useful for increasing their productivity at work and would recommend CONVERGENCE to colleagues, mainly because it could save them time when searching for products. The overall score was 3.4, similar to the score in the first session.

Supervisor's report

The supervisor reported network problems, which made it hard for some participants to access and use the web application.

4.2.2.5.3.5 Focus Group

In this trial, unlike the trial in UTI, the focus group brought all the participants together in a single group.

Navigation: All of the participants in the trial classified the navigation functions as good, intuitive, easy to understand and user friendly (*"The navigation itself is very intuitive and*

easy to use." or *"Very clear and good navigation."*). They also stated they had become used to the application just by interacting with it a couple of times.

Help: Participants were also unanimously positive about the help provided by the application. (*"Besides being very intuitive the system also provides help messages in all of its screens.*). This was especially appreciated when they had to fill in forms *(There were very helpful tool tips in all the application's forms.")*

Layout/graphics: Although most of the participants thought that the application had a very user-friendly interface, and a layout that was sufficiently pleasant for normal use. (*"The layout is appropriate to an everyday, several times a day task ... the graphics play, in my opinion, a reduced part is this process and, are pleasant enough as they are."*) . Nonetheless some users thought that the graphics could be improved (*"It is a very user friendly interface, but probably a lit bit too simple and limited"*). The main suggestion was to improve the presentation of product information.

Performance: Most of the participants claimed that the tasks they had to perform were very slow and took a considerable amount of time to finish. The slowest functions were the creation and publication of VDIs *(It really takes a very long time to create a VDI").*

Robustness/technical problems: Some of the participants encountered technical problems during the trials, mostly due to problems with the network and with storage space on the server/database. All problems were identified and corrected during the trials. However, two participants observed data verification problems when creating product VDIs. The other participants stated that they had not had any technical problem during their interaction with the system. *("From the experience I had, there were no problems whatsoever").*

Privacy: Participants felt that CONVERGENCE offered the right level of privacy for this kind of application. However, they also thought that future versions would need more privacy options. Several emphasized the need for an abstraction layer between the manufacturer and the retailer, both as a protection for privacy and because of anti-trust considerations.

Security: In general, participants agreed that that the system seemed secure and safe and none raised any major concerns. Nonetheless one pointed out that the system's security would be improved if it used smart cards and VDI data was encrypted. Another stated that he could not answer questions about security without use cases and an in depth failsafe analysis of the system.

Ease of use and learning curve: All of the participants considered that the application was very easy to use and understand. However the group was not homogenous: some familiarized with the system right immediately (*"It was pretty intuitive on the first try, which is not very common among this kind of interfaces with retailer, for example."*) while a minority needed to use the system a couple more times before they reached the same level of confidence. In general, however, learning was rapid.

Improving CONVERGENCE: Most participants provided comments. As in the first trial, the topic that interested participants most was product ontologies. In general they felt the system would be enormously improved if it had a strong underlying ontology. Participants felt this need most when they had to interact with the subscription tool. Sometimes they failed to get all the matches they expected, because the keywords in the published VDIs were slightly different from the ones they had used in their subscriptions. Since the retail industry has no standard product ontology, the short-term solution would be to create a CONVERGENCE-specific ontology, at least for use in tests and trials. One participant proposed the development of an "ontology creation" tool, capable of creating and maintaining product ontologies for CONVERGENCE applications. Other users suggested improvements directly linked to their daily work. Several suggested the inclusion of some kind of an inventory management system, linked to and synchronized with product VDIs. Such a system, they suggested, could keep retailers informed about the status of their product inventories, even providing them with information about the location of their products. Another participant suggested the creation of a CONVERGENCE search engine. Even though she liked the publish and subscribe scheme and the way it worked, she would still like the ability to search for VDIs by keywords and receive all the results quickly, on a single screen, like the results page in a search engine.

A few participants also stated that they would have liked to see a lighter interface, and said they would have been more comfortable with a more appealing, minimalist approach.

Best features: For most participants, CONVERGENCE's best feature was the ability to create product VDIs. All participants mentioned VDIs as a new concept for a standard digital container. Participants also reported that the VDI creation application was very easy and intuitive to work with. For a smaller number of users, the best feature was the publish/subscribe system. Several users were really impressed by CONVERGENCE's ability to match the products created and published by manufacturers to retailers' subscriptions. Users were also very attracted by how easy it was to check the availability of a product – a feature that they thought could be very helpful for the retail business.

Worst features: Here, opinions among participants differed. For some, the worst aspect of CONVERGENCE was the impossibility of editing VDIs. For these participants, the idea that a VDI cannot be edited does not make sense, "*The worst is the fact that I can't edit VDI files as a manufacturer, because it obligates me to delete it and create it all from the beginning again if something is wrong.*" Others report that the application for browsing VDI information was poor and needed major improvements in terms of functionality and user interface. Regardless of their opinions on this matter, most participants thought that it was not possible to properly exploit CONVERGENCE publish subscribe functionality without a proper product ontology.

4.2.3 Developers' Feedback

The final stage of the evaluation involved the collection of feedback from the developers who had worked on the applications used in the trials. As in the previous trial, these were not the same programmers who developed the CONVERGENCE middleware and they came from different partner organizations. For obvious reasons, the number of participants was small (six). Nonetheless their responses provide interesting insights into the process of developing applications with the CONVERGENCE middleware.

Developers reported that most of the difficulties experienced in the previous round had been overcome thanks to a better exchange of information with the team responsible for the CONVERGENCE middleware. They also reported that there was now better documentation and more detailed code examples. However, some continued to experience problems due to the lack of unique methods to connect engines. Another general issue is the fact that the security infrastructures for the various methods provided by the middleware are not yet implemented. In what follows, we summarize developer feedback on specific aspects of the system:

USER REGISTRATION: Developers reported that this tool is still not completely implemented. In particular, it is still missing the connection with the orchestrator and has yet to be properly integrated in the applications. Developers said that at the moment it is still more efficient to access the APIs directly rather than use the tool, which is still immature, and which still has no sample application.

CONTENT REGISTRATION: The application is working and inadequacies in documentation and coding examples have been resolved. Bugs have mostly been solved but the tool still needs better security mechanisms. In general, developers feel that the functionality on offer is sufficient for their needs. However, they also report that the APIs are too complex to be truly developer- friendly.

PUBLISH VDI: This tool – essential for the publication of VDIs and the correct orchestration of relevant engines – works quite well. No relevant issues were noticed. The only bug encountered was solved in one day. The main problem is lack of clarity in the definition of parameters and the specification of the data needed by the tools. One possible step would be to introduce a single step generation and publication operation. For the subscription tool, the biggest need is for methods to parse subscription interests in S-VDIs.

UNPUBLISH VDI: This tool is still in development and is not yet integrated in the applications, due to the lack of integration with CONVERGENCE security mechanisms.

REVOKE VDI: This tool is still in development and is not yet integrated in the applications, due to the lack of integration with CONVERGENCE security mechanisms.

SUBSCRIPTION VDI: All developers used this tool in their application. The tool works quite well. The few bugs reported have been quickly solved. The documentation is complete and

there are no outstanding requests for improvement. A possible improvement would be the addition of methods to parse the subscription interest in S-VDIs.

BROWSE EVENT REPORT: The tool is working very well. The tool is stable, and offers a solid API. One possible improvement would be to add new functionality for creating Event Reports.

CREATE ANNOTATION REPORT: This tool will be not used in the project. LMU, the sole user has decided to access the CONVERGENCE APIs directly.

4.3 Recommendations from the end-user trial

As planned, the second stage of the CONVERGENCE trials involved 105 participants (99 end users and 6 developers). Many participants expressed the view that CONVERGENCE applications could be potentially valuable for their businesses, once they had reached a more mature stage in their implementation. This is a strong point in favour of the project.

It is interesting that the strong points in the applications were their responsiveness and speed - features that reflected the strengths of the underlying CONVERGENCE system. The weaker points were layout, ease of use and the help functions – issues that are vital for users but which are independent of the underlying middleware and network. Users also noted a number of technical problems – but did not consider them to be particularly important, given the current status of CONVERGENCE as a prototype system. In general they were strongly interested by the idea CONVERGENCE but had considerable difficulty in understanding the underlying concept.

Compared to the first round of trials many problems have now been resolved. Nonetheless some of our recommendations from the previous round have yet to be properly implemented. We therefore formulate three main recommendations for the next round

1. *Functionality*. Even though CONVERGENE is not intended as an applications development project, it is essential that the last round of the CONVERGENCE trials shows functionality rich enough to effectively illustrate the advantages of the system. We repeat our suggestion from the previous round to create pop up text boxes providing additional information about what CONVERGENCE is doing when the user performs a particular action (e.g. creating a VDI, adding metadata to the VDI, defining a license for a VDI, publishing a VDI, subscribing to a VDI etc.). While this functionality would never appear in commercial implementations of the system it would be extremely useful for demonstrations.
2. *Usability and aesthetics.* Participants in the second round of trials again gave the CONVERGENCE poor marks for ease of use and layout, and many found it visually unattractive. We still believe that it would be useful to employ a graphics designer to

improve this aspect of the applications. We also note a strong request from users to provide text tips for fields they have to fill in and better help functionality.

3. *Security.* Several participants expressed concern about the security of CONVERGENCE applications (which was not shown in this phase of the trials). It is essential that the next phase should properly demonstrate these features.

4.4 Concluding remarks on the end-user trial

There can be little doubt that the results of the end-user trial reported in the previous sections, provide extremely useful feedback for applications developers. However, it is clear that this trial cannot in itself provide a test of CONVERGENE as a system. CONVERGENCE is not an applications development project. Rather it touches on the basic architecture of the future Internet. This means, on the one hand that end-users cannot test (and, in some cases, even understand) the key characteristics and functionality of the system (just as they cannot test the key characteristics of TCP/IP); on the other hand user feedback issues such as ease of use, user interfaces etc., are of only limited use to the groups developing the CONVERGENCE architecture and middleware. This raises the question of how to best conduct the final round of trials.

5 Track 2: Network experiment and simulation

5.1 Testing CONVERGENCE Networking

According to the plan presented in Deliverable D8.1, the CONVERGENCE Network (CONET) was to be tested in two phases.

During phase 1 (January-June 2012), we planned to deploy and test a laboratory network consisting of a single CONET Subsystem (CSS), composed of end-nodes, serving-nodes and internal-nodes. This network is shown in Figure 5.1.

During phase 2 (July 2012 – February 2013), we planned to extend the laboratory test scenario by including multiple CSSs and using a distributed testing platform, like OFELIA or, as a contingency plan, OpenLab/PlanetLab.

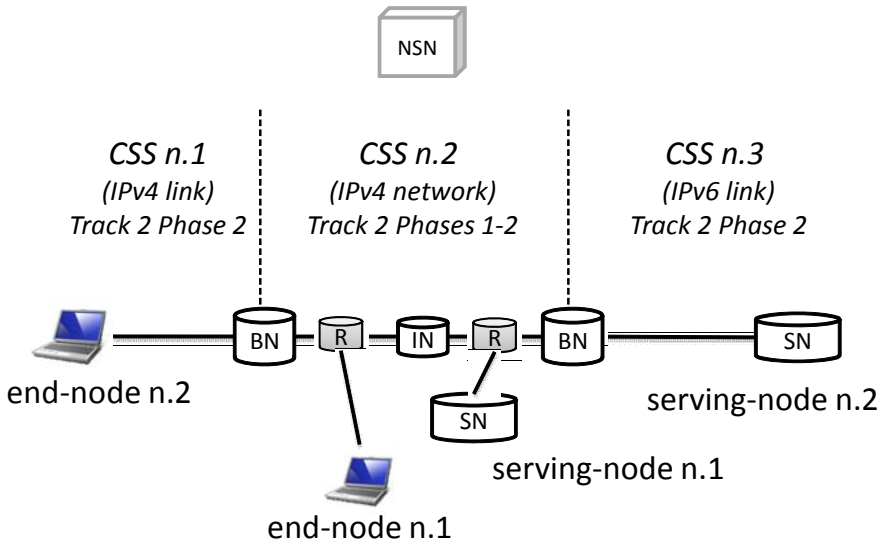


Figure 5.1: Laboratory scenario for CONET tests

This report comes at the end of phase 1 and includes most of the planned results for the phase 1 network test. It also anticipates some of our plans for phase 2. As we will report, we have already deployed a number of Border Nodes and carried out limited tests with multiple CONET Subsystems. This was necessary to verify the effectiveness of the CONET routing-by-name architecture, (Lookup-and-Cache) described in Deliverable D5.1 and [1]), and to test the scalability of the routing plane, a fundamental issue for information-centric networking. However it also meant that we had to delay some of our planned work in phase 1, which has now been shifted to phase 2. More specifically the deployment of Internal Nodes and the implementation and test of the SendToLocation and Route-by-path-info primitives (see

Deliverable D3.2) will now take place in phase 2. Table 5.1 reports the current status of the tests, planned in Deliverable D8.1.

Test	Phase planned in D8.1	Instrument	Status
API test	1	Laboratory network as in Figure 5.1 but including only IPv4 CSS n.2	Partially completed (test of SendToLocation and Route-by-path-info shifted to phase 2)
Networking test	1	Laboratory network as in Figure 5.1 but including only IPv4 CSS n.2	Partially completed (test of Internal Node shifted to phase 2)
Performance test	1	Laboratory network as in Figure 5.1 but including only IPv4 CSS n.2	Partially completed (test of Internal Node shifted to phase 2)
API test	2	Laboratory network as in Figure 5.1	Partially completed (only Get primitive tested)
Networking test	2	Laboratory network as in Figure 5.1	Partially completed (Missing test of distribution of routing information among NRS nodes)
Performance test (routing)	2	Laboratory network as in Figure 5.1	Almost completed
Performance test (routing)	2	Ofelia/OneLab Infrastructure	Test bed deployment and experiments on OneLab network in progress

Table 5.1: Test Plan and status for CONVERGENCE networking concepts

5.1.1 API test

5.1.1.1 Description of the test

The test set up consisted of a single IPv4 CSSs, in which an end-node communicates with a serving-node in the same CSS. The test verified the correct execution of the following API primitives offered by the CONET TE:

1) Store and advertise
2) Revoke
3) Get
4) Send2Name
5) advertiseSap

We also verified the correct execution of the `Get` operation with two CSSs, where the end-node of CSS n.1 (see Figure 5.1) is connected with the serving-node of CSS n.2.

5.1.2 Networking test

5.1.2.1 Description of the test

The aim of the Networking test is to verify the correct operation of CONVERGENCE network functionality supporting API services.

We created setups with a single IPv4 CSS in which an end-node communicates with a serving-node in the same CSS, and also with two CSSs where the end-node of CSS n.1 (see Figure 5.1) is connected to the serving-node of CSS n.2. With both setups, we verified correct execution of the following operations:

- Packetization of named-data
- Interactions with Name System Nodes to insert and lookup name-based routing information
- Lookup-and-cache
- Route-by-name
- Named-data caching

5.1.3 Performance test

CONET provides a set of extensions to CCNx that guarantee the scalability of the network, avoiding any deterioration in performance compared to the plain Internet and ensuring it can support a full range of services including conversational/interactive services.

As shown in Figure 5.2, CONET extends CCNx-based networks as follows:

Naming: For purposes of routing and scalability, CONET adopts a Principal/Label, naming scheme format.

***API*:** CONET extends the functionality offered by the CCNx API to provide better support for PUSH services, a feature that is especially useful for conversational and interactive services.

Routing-by-name: CONET implements a routing-by-name architecture, ("Lookup and Cache") that allows the use of classical CCNx routing and forwarding by name with very large number of name-prefixes.

Transport: CONET implements segmentation and transport level mechanisms that strongly improve the performance of CCNx transport.

Security: CONET extends CCNx's RSA-based security functionality with Identity Base cryptography and Elliptic Curve DSA, reducing the security overhead for individual data messages (named-data CIUs) transported over the network.

Caching: CONET extends CCNx with additional caching algorithms, (e.g. LRU) and will support secure caching, i.e. caching restricted to content with a verified digital signature.

In what follows, we evaluate the performance of the optimized segmentation and transport mechanism (Section 5.1.3.1) and of the Lookup and Cache architecture (Section 5.1.3.2).

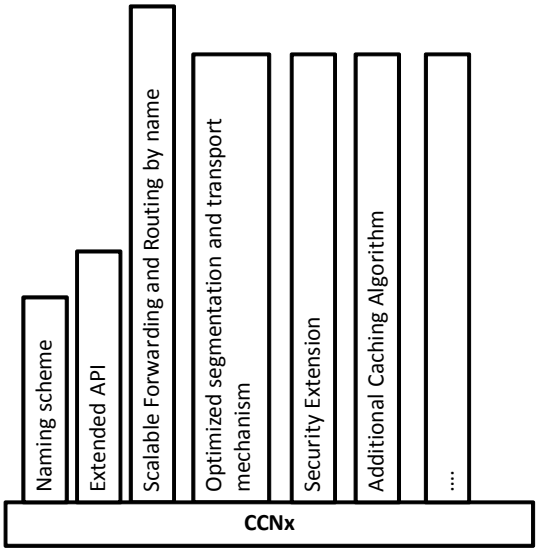


Figure 5.2: CONET extensions of CCNx

5.1.3.1 Throughput measurements

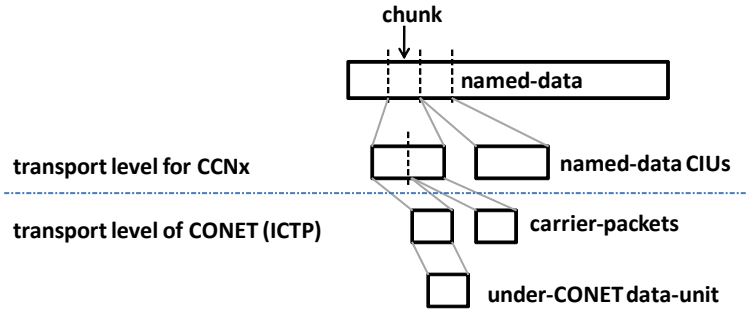
In this section, we report measurements of the throughput between an end-node and a serving-node directly connected by a 10 Mbit link. These measurements demonstrate the effectiveness of CONET segmentation and transport mechanisms, compared to the mechanisms provided by plain CCNx. For ease of reading we briefly summarize the CONET segmentation and transport mechanism.

5.1.3.1.1 Segmentation and Transport mechanism

It is well known that TCP endpoints exchange Segment/Ack sequences: the sender sends Segments and the receiver replies with Acks. This means that the rate at which Segments are sent is under the control of the sender, but also depends on the Acks (the feedback from the receiver).

CONET, by contrast, uses a receiver-driven protocol, which we can call Information-Centric Transport Protocol (ICTP). In the ICTP, endpoints exchange Interest-Data sequences and the rate at which they are exchanged is regulated by the receiver, following the principles described in [2] and [3]. Briefly, the receiver issues a sequence of Interest CIUs, each of which requests a *segment* of a named-data CIU, which is then transported in the payload of a

carrier-packet (Figure 5.3). By controlling the sending rate for these interest CIUs, we obtain a transport protocol that is TCP friendly and ensures fairness both among competing ICN flows and between ICN and TCP flows. The algorithms underlying the protocol are the same as for TCP (slow-start, congestion avoidance, fast retransmit and fast recovery).



-

Figure 5.3: CONET and CCNx transport levels

The main differences between ICTP and the transport protocol provided by the CCNx tool [4] can be summarized as follows.

1) *Use of carrier-packets*. As shown in Figure 5.3 the data-units in the CCNx transport protocol are named-data CIUs (aka Data messages) each of which is relatively large, (e.g. 8 kB). This means they have to be segmented at the IP level. The use of such a large data-unit decreases the efficiency and promptness of congestion control. To overcome this problem, ICTP divides named-data CIUs into segments, each of which is transported by a carrier-packet, whose size is close to that of a TCP segment. In other words, the carrier-packet is the ICTP data-unit and no longer the whole named-data CIU. This suggests that the performance of the transport layer will be similar to that of the transport layer in TCP.
2) *Mimicking TCP congestion control.* CNNx implements a simple congestion control algorithm that resembles the traditional Selective-Repeat ARQ with fixed window size. Conversely, ICTP thoroughly mimics the TCP Reno congestion control algorithms.

5.1.3.1.2 Performance Evaluation

We evaluated the performance of ICTP in a laboratory scenario with a single IPv4 CSS. The experiment setup consisted of two Linux PCs connected by a 100 Mbit/sec Ethernet. The first PC acts as a data source (the CONET serving-node) and the other as a data-sink (i.e., the CONET end-node). On both machines, we used the Linux traffic control (tc) module, to limit the outgoing data rate to 10Mb/s at the IP level. The size of the headers for named-data CIU was set to 650 bytes; the size of carrier-packets was 1460 bytes.

In the first experiment, we measured ICTP goodput (application level throughput) during the transfer of a long file (10 Mbyte). To show the advantages of working with small data-units at the transport layer, we compared the performance of the "multi-segments" approach (the default in ICTP) with a "single segment approach" in which each transport layer data-unit transported a complete named-data CIU, which was then fitted into an Ethernet MTU, using IP fragmentation. In both cases, we used the ICTP congestion control algorithm, ensuring that the comparison was independent of the congestion control scheme.

Performance was tested for different loss rates (shown on the x axis) and for different sizes of named-data item (4 kB, 8 kB, 16 kB and 32 kB). For the sake of clarity, Figure 5.4 shows the results only for items of size 4 kB and 32 kB. Performance for items of size 8 kB and 16 kB lay between the values for 4 kB and 32 kB. For each case we measured performance for at least 10 runs and computed 95% confidence intervals. Note that the graph shows the lossless case (0% loss) as a loss of 10^{-4}, slightly abusing the logarithmic scale of the x-axis.

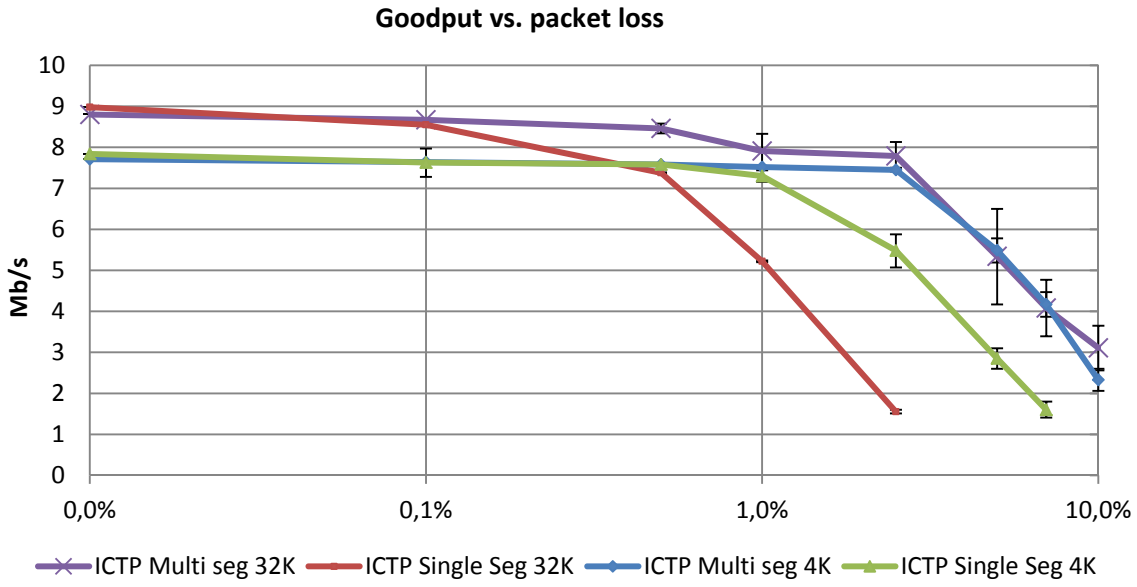


Figure 5.4: Multi segment vs. single segment goodput

In the lossless case, goodput with the single segment was slightly higher than for the multi-segment solution. This was true both for 4 kB and for 32 kB named data items. This is because ICTP carrier-packets (a single carrier packet per named-data CIU) had lower overhead. The named-data CIU header overhead (650 bytes per named-data CIU, mostly due to the signature) had a lower impact in the 32 kB case than in the 4 kB case. As a result goodput was higher in the former case. However, when the loss rate increased, the performance of the single segment case fell very sharply. This is because, in this case, the loss of a single IP packet implied the loss of a whole named-data CIU. In the single segment case, the loss was higher at the congestion control level than at the IP level, again because the loss of a single IP fragment caused the loss of the whole named-data CIU. This effect increased

with the size of the named-data CIU; thus, the goodput with named-data CIUs of 32 kB was smaller than the goodput with 4 kB.

We also compared the performance of ICTP against the performances of congestion control as currently implemented in CCNx. Figure 5.5 plots application goodput for CCN and for ICTP for different probabilities of IP packet loss and for two different sizes of the named-data CIU (4 K and 32 K). In the ideal lossless case, the current implementation of CCNx transport saves the overhead of the carrier packet and thus performs slightly better than ICTP. However, as soon as there is some loss of packets, the relative performance of ICTP improves sharply. This means that if CCNx flows and ICTP flows competed on the same bottleneck, ICTP flows would be more aggressive and would take an unfair share of resources.

In an evolutionary deployment of ICN, CCN or ICTP flows would coexist with TCP flows – a possibility we have analysed in [5]. We began with a more abstract experiment that did not use the current implementation of the ICTP protocol. Rather, we used a plain source-driven TCP Reno in which we increased the data-units length from 2K to 8K, mimicking the size of the named-data CIU in ITCP. We then measured the goodput of a CCN flow coexisting with ten regular TCP/IP flows. When we used small named-data CIU, the IP good puts were comparable: TCP/IP and CCN flows fairly shared the link capacity. Conversely, when we used large named-data CIUs, the CCN lost packets and the TCP/IP flows began to starve the CCN flows. This implies that ICN flows have to use carrier-packets if they are to coexist with TCP/IP flows.

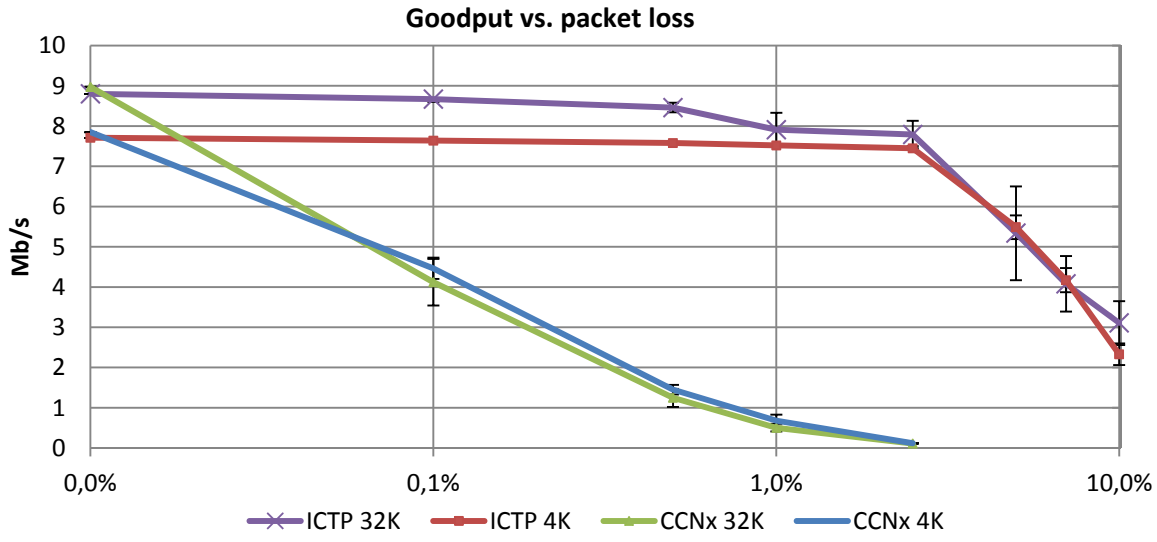


Figure 5.5: ICTP vs. existing CCNx transport

As reported in [5], we also investigated issues related to the coexistence of several CCN flows. We found that the large congestion control data units stretched the timescale for congestion control leading to significant oscillations in goodput. This could be a problem In

some circumstances – for instance in the case of non-real time video streaming services, which would require a large de-jitter buffer, penalizing quality of experience.

Measurements of coexisting CCN, ICTP and TCP flows on the real test bed are on-going and will be reported in [5].

5.1.3.2 Routing-by-name measurements

In this section, we evaluate the feasibility and the performance of the data plane of the CONVERGENCE Lookup and Cache architecture, as described in Deliverable D5.2 (for a more extensive analysis see [1]).

For ease of reading, we briefly describe the CONET Lookup and Cache mechanism.

The Lookup-and-Cache routing architecture uses the FIB of a Forwarding Engine as a route cache and deploys a centralized routing engine to serve all the nodes of a sub-system. Figure 5.6 shows an example of Lookup-and-Cache operations.

Node N receives an Interest message for a named-data item, with Network Identifier "ccn.com/text1.txt/chunk1". Since the FIB does not have a route for the item, the node temporarily queues the Interest message, lookups the route in a remote RIB, gets the routing information and stores it in the FIB, after which it forwards the Interest message.

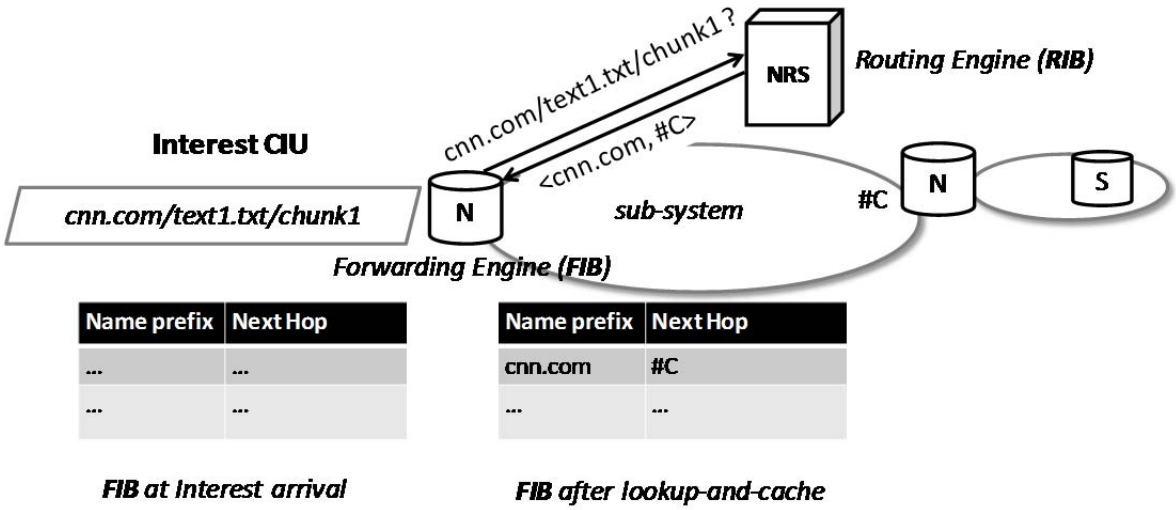


Figure 5.6: Lookup-and-Cache concept

5.1.3.2.1 Feasibility check

In this section we analyse the feasibility of deploying the Lookup-and-Cache architecture, by using currently available technology.

5.1.3.2.1.1 Reference scenario

We envisage a scenario in which CONET is used to retrieve content from the WEB. In this scenario, WEB clients become CONET end-nodes, WEB servers become CONET serving-nodes and WEB pages become CONET named-data items.

As far as concerns naming schemes, we assume that URLs are mapped to CONET Network Identifiers (NID), as discussed in Deliverable D5.1, section 5.4.2. The NID has a Principal/Label structure in which the Principal is the DNS second-level domain of a URL. For instance, for the URL "www.ict-CONVERGENCE.eu/documents/doc1.txt", the Principal is "www.ict-CONVERGENCE.eu" and the Label is "documents/doc1.txt". Today there are approximately 2 10^8 second levels domains. Allowing a little overestimation, we would thus have about 10^9 Principals. We observe that with this naming scheme, a Principal identifies a serving-node that contains all named-data items who's NID has the Principal string as a prefix. For instance, the serving-node of the Principal "www.ict-CONVERGENCE.eu" contains all named-data items, whose NID starts with "www.ict-CONVERGENCE.eu". This means that the CONET routing plane could operate using just Principal strings. A routing entry in the forwarding table used by a CONET node to route incoming Interest CIUs (an *ICN* route) would thus take the form:

<Principal, next-hop>

Where next-hop is the IP address of the next CONET node along the routing path. In this case, the only ICN route needed to route all Interest CIUs related to the named-data items published by the serving-node would be:

< www.ict-CONVERGENCE.eu, next-hop>

Since the number of Principals is in the order of 10^9, CONET nodes would also need to handle approximately 10^9 routes.

5.1.3.2.1.2 Scope of the feasibility check

In our study, we considered feasibility for carrier-grade CONET nodes located in the *default-free zone* of the network. These are nodes that do not have a default route and need to forward packets at very high speed. For obvious reasons, these nodes have higher requirements than other nodes in terms of processing speed (lookup rates) and storage (size of the FIB). Figure 5.7 shows the modules involved in the Interest CIU forwarding procedure for such a node. The architecture is directly derived from the architecture for an IP router [6][7][8], the main

difference being that the Routing Information Base (RIB) is *pulled out* by the node and located in a remote Name Routing System (NRS) node. A NRS node logically serves all CONET nodes belonging to the same Autonomous System.

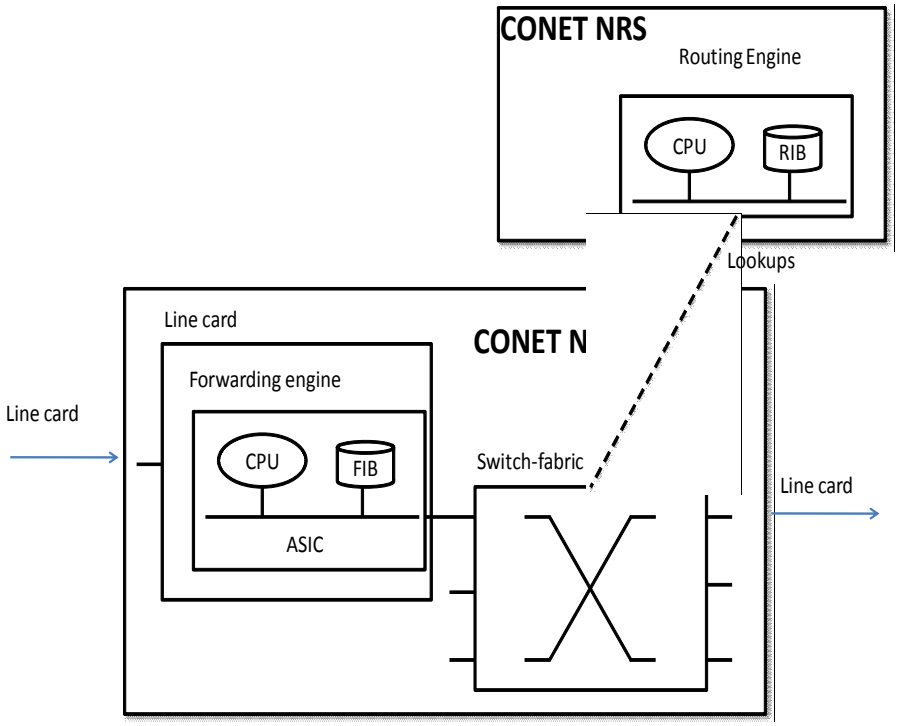


Figure 5.7: Forwarding modules of CONET carrier-grade node

A CONET node is composed of two major forwarding components: line cards that host a forwarding engine and a switch fabric. The forwarding engine of a line card receives incoming Interest CIUs and selects the output line card by looking up an on-board routing table, the Forwarding Information Base (FIB) [9]. The FIB contains one route per destination. To support forwarding at line rate, forwarding is performed by dedicated ASIC chips and the FIB is implemented with fast memories, such as SRAM or TCAM. These memories are costly, consume a lot of power, and do not follow Moore's Law [10]. After selecting the output interface, the forwarding engine injects the packet into the switching fabric: an NxN non-blocking crossbar, where N is the number of line cards (including the routing engine).

Each CONET NRS contains a routing engine, which handles the routing protocols and stores the routes in a routing table, called the Routing Information Base (RIB). In general, the RIB contains several routes to the same destination and is implemented using cheap, slow memory such as DRAM.

When a CONET node needs to forward an Interest CIU and there is no related ICN route in the FIB, it looks up the route in the remote RIB and caches it in the FIB. This approach is motivated by the fact that the current maximum size of a SRAM chip is 32 MByte [11].

Assuming an ICN route of 45 bytes[3] [12], a FIB could store some 10^6 routes, i.e. 32MB/45B. This is considerably lower than the 10^9 ICN routes required in our scenario. We assume, however, that in a given short period of time, the FIB will only need to use a restricted set of ICN routes and that this restricted set could easily be stored on the FIB.

The aim of our study was to verify these assumptions. More specifically the study sought to verify that:

i) the capacity (10^6 routes) provided by current FIB technology is enough to store the expected number of *active-routes*
ii) current database technology can support the desired route lookup rate in the NRS node providing the RIB;
iii) the invalidation rate would be compatible with current processing technology and would not generate too much traffic over the RIB-FIB path[4].

An ICN route on a given node at a given time is "active" if at least one flow of Interest messages uses the route, as shown in Figure 5.8, where there are three flows of Interest messages toward "cnn.com". The route toward "cnn.com" becomes active at the start of the first flow and becomes inactive at the end of the last flow. Figure 5.8 also shows a single flow of Interests for "bbc.com". This route will be active so long as this flow is active.

5.1.3.2.1.3 Feasibility Results

In the current Internet, a client sends a TCP ACK and receives TCP segments from the Web server. In CONET, a client sends Interest CIUs and receives named-data CIUs from the ICN server, or from an en-route cache. If a client uses CONET to download Web contents, traditional flows of TCP ACK messages are replaced by a flow of Interest messages and the couple < IP destination address, destination Port> in TCP ACK messages is replaced by a NID containing the Principal prefix advertised by the Web server.

Consider a scenario in which a host on the current Internet sends an HTTP request towards the domain name "cnn.com". The DNS translates the domain name "cnn.com" into an IP address, e.g. 157.166.226.25 and returns it to the host. The host then sends a request to 157.166.226.25:80 which directs a flow of data back to the requesting host which responds with a flow of TCP ACKs. In the CONET scenario, we are considering here, the flow of TCP ACKs is replaced by a flow of Interest CIUs for named-data items, whose NIDs contain the "cnn.com" Principal prefix.

[3] 40 bytes are for the name-prefix, 1 byte for the output port identifier and 4 bytes for the next hop info.

[4] The RIB uses invalidation to remove stale ICN routes cached in the FIB, as described in Deliverable D5.3 and in [5].

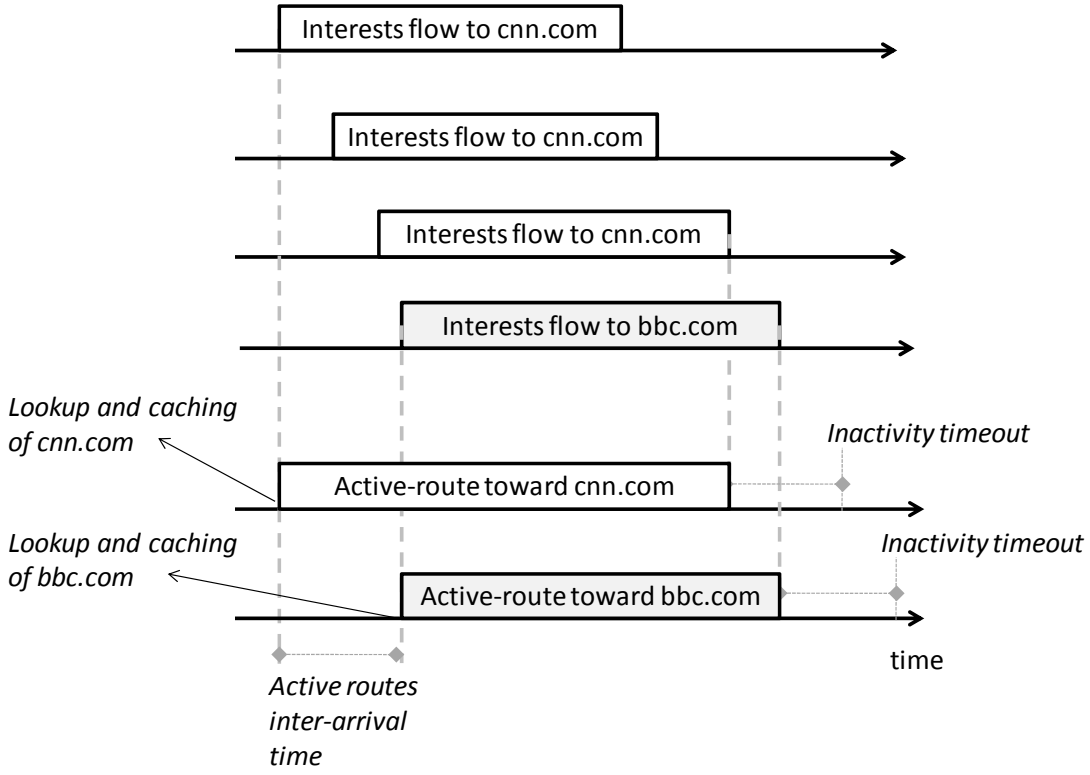


Figure 5.8: Flows and active-routes

Assuming a *one-to-one* relationship between an IP address and a domain-name[5], we can remap a trace of ACKs to an equivalent trace of CONET Interest CIUs. In what follows, we apply this *remapping* to real Internet traces, estimating the number of active-routes for a hypothetical CONET. We proceed as follows.

First, we extract ACKs addressed to port 80. We then consider each ACK as if it were an Interest CIU for a named-data item available at the destination Web server (the CONET serving-node)[6]. Following our naming scheme, the Principal part of the NID for the item is the domain name of the Web server[7]. Given that, routing-by-name operations consider only

[5] In cases where several domain-names share the same IP address(as in some web hosting services) the assumption of *one-to-one* relationship between an IP address and a domain-name may lead to an underestimation of the number of active-routes and the required lookup rate. Appendix II of [1] presents a simulation model that takes Web Hosting services into account. Simulation results show that Web hosting services increase the number of active-routes and the lookup rate with respect to the values computed with the *one-to-one* assumption; nevertheless the two assumptions yield estimates of the same order of magnitude. Thus our conclusions on feasibility remain valid.

[6] Interest messages refer to named-data items, which are typically greater than TCP segments. The number of Interest CIUS would be lower than the number of ACKs. However, the one-to-one correspondence between ACK and Interest allows tracking the activity of flows and routes, which is exactly what we need.

[7] We have only anonymized traces, from which we cannot derive the actual domain name of the server; so we used a fake domain name equal to the anonymized IP address of the server, considered as a flat string.

the Principal part (e.g., “cnn.com”)[8]. we do not need to know the other parts of the NID of the named-data item.

The remapped trace of Interest CIUs feeds a simple algorithm, which infers whether a route is active by using an inactivity timeout of 1s. A route is deemed to be active from the time of receipt of the first Interest message and inactive when there have been no related Interest messages in the last second; the end-time is thus the time of receipt of the last Interest message.

Table 5.2 summarizes our results. The Equinix-sanjose-* and Equinix-chicago-* traces are captured from the10 GigE interfaces of a tier-1 ISP. The Mawi-* traces are captured from a trans-Pacific line operating at 150 Mbit/sec. The Rome-Tor-Vergata trace is captured from the 1 GigE interface on our University’s router gateway for our University, a tier-3 network.

Trace id	Date (m-g-y, h:m)	Trace duration	Average n. of active-routes	Max n. of active-routes	Average active-routes inter-arrival	Ref.
Equinix-sanjose-dirA	11-18-2010, 13:50	60 s	2582	2681	1.2 ms	[13]
Equinix-sanjose-dirB	11-18-2010, 13:50	60 s	1293	1353	2.8 ms	[13]
Equinix-chicago-dirB	10-29-2010, 13:06	60 s	1197	1283	1.7 ms	[13]
Mawi-1	11-28-2011, 14:00	15 min	236	289	5.9 ms	[14]
Mawi-2	11-29-2011, 1400	15 min	246	303	4.1 ms	[14]
Rome-Tor-Vergata	03-04-2011, 15:06	9 min	176	227	7.3 ms	[15]

Table 5.2: Internet trace analysis report

[8] If we needed to know the full chunk-names, we should have performed an HTTP level analysis of the traces, rather than limiting ourselves to the analysis at TCP level.

As expected, the interface with the highest number of active routes is the interface for the tier-1 ISP and the interface with the lowest number of routes is our own university interface. Even in the worst case of the Equinix-sanjose-dirA trace, the maximum number of active-routes is of the order of 10^3, approximately 10^3 times lower than the maximum capacity of an off-the-shelf SRAM based FIB, i.e. 10^6 ICN routes.

Figure 5.9 plots the number of active-routes versus time for the Equinix-sanjose-dirA trace. The number of active-routes varies in a relatively tight range around the mean value of 2582. This simplifies the dimensioning of the FIB, which can be set close to the observed mean, without requiring a large safety margin [16].

A flow-level analysis [9] reveals that 2.6 million Interest flows, i.e. content downloads in the flow use just 11,000 routes[10] and that roughly one tenth of these routes account for about 96% of flows. These results are consistent with previous work on IP [16]. Figure 5.11 shows the cumulative density function (CdF) for *route* popularity versus the route ranking (*k*), where CdF(*k*) is the probability that a flow uses a route that belongs to the first *k* mostly used routes. We compare this function to the CdF predicted by the Zipf distribution, which is known to provide a good representation of *content* popularity. The figure shows that route popularity is even more skewed than predicted by the Zipf function. The explanation is that ICN routes do not address specific items of content but servers, and that some very popular servers (e.g., YouTube) provide access to many items of content.

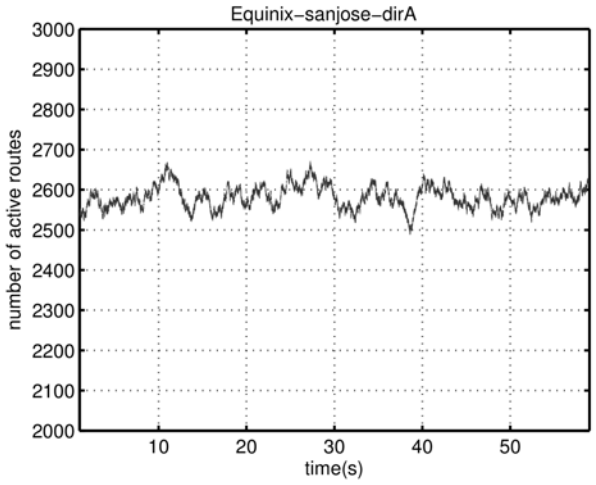


Figure 5.9: Number of active-routes for the Equinix-

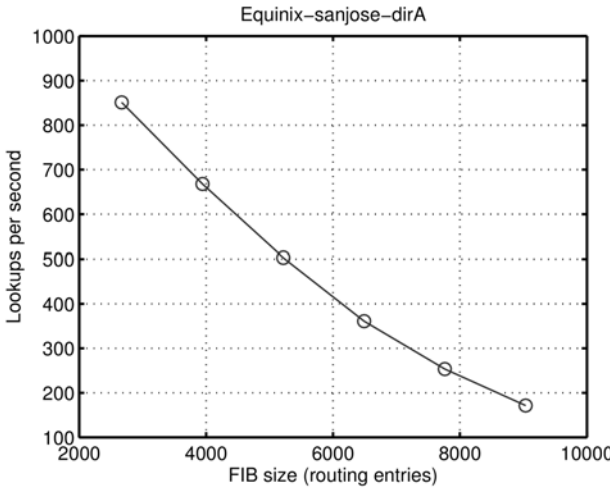


Figure 5.10: Lookup rate versus FIB size for the

[9] We measured the number of active-flows using the same methodology used for the number of active-routes, i.e. based on an inactivity time out of 1 sec. An active-flow is identified with the triple <client-address, server-address, client port>.

[10] The average number of flows active at a given time *t*, i.e. the number of *active-flows*, is equal to 27865 versus a number of active-routes equal to 2582.

sanjose-dirA trace

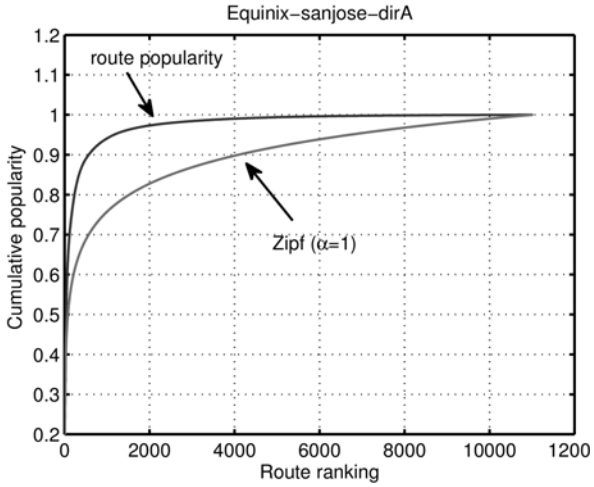


Figure 5.11: Cumulative route popularity for the Equinix-sanjose-dirA trace

Equinix-sanjose-dirA trace

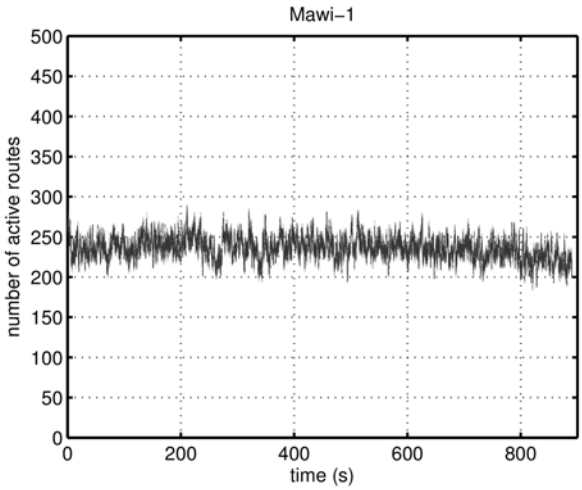


Figure 5.12: Number of active-routes for the Mawi-1 trace

In the next stage of our study, we investigated whether current database technology is sufficient to support the required lookup rate for the RIB. Table 5.2 shows that in the traces we considered, the average time between the start of two consecutive active-routes was of the order of milliseconds. Assuming that FIB memory is dimensioned to contain all active-routes, the inverse of the inter-arrival time for active routes provides an upper bound on the required lookup rate[11]. Therefore, an average active-route inter-arrival time of a few ms would require a lookup rate in the order of 1000 lookups per second, in the worst case. This level of performance can easily be achieved with current database technology. For instance, we have implemented an NRS node with a Bind9 server, running on an old Linux laptop with an Intel Pentium Processor M at 1.4 Ghz. This set-up offered a sustainable rate of about 15 000 lookups per second.

We also investigated the possibility of *over-provisioning* the FIB, so as to reduce the required lookup rate. A FIB is said to be over-provisioned, when it has a capacity greater than the maximum number of expected active-routes. For this analysis, we used an ideal route replacement policy that randomly replaces inactive-routes. Figure 5.10 shows required lookup rates vs. FIB size for the Equinix-sanjose-dirA trace. From the data, we see that as the FIB size increases, the required lookup rate decreases. This is because individual routes switch on and off very frequently: in the trace we observe about 54000 route activations/deactivations for just 11000 unique routes. These results imply that the FIB size has a significant impact on the cache hit-ratio and, hence, on the required lookup rate.

[11] In reality, lookups are only required if the route has not already been cached in the FIB.

We also note that with cached routes (as in our scenario) the relationship between hit-ratios and cache size is stronger than with cached content. In this latter scenario, the hit-ratio grows with the log of cache size [17]. Figure 5.12, Figure 5.13 and Figure 5.14 show equivalent data for the Mawi-1 trace. This analysis (and other analyses not reported here) confirms our previous conclusions.

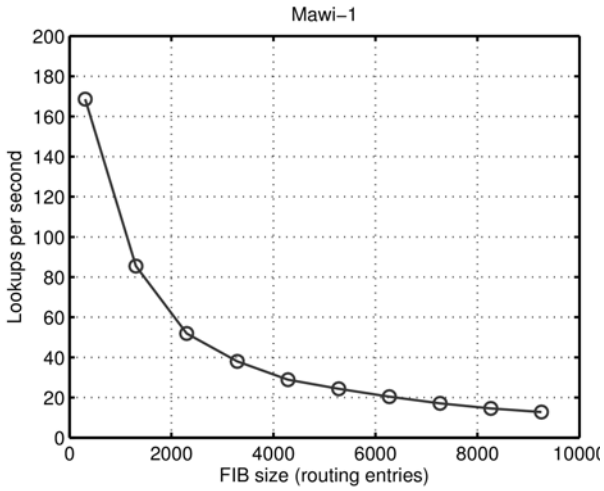


Figure 5.13: Lookup rate versus FIB size for the Mawi-1 trace

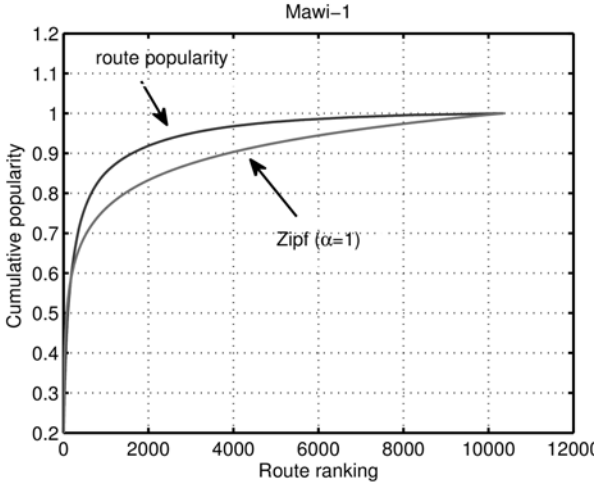


Figure 5.14: Cumulative route popularity for the Mawi-1 trace

The next stage in our study was to investigate the feasibility of the invalidation approach described in Deliverable D5.3 and in [1]. Briefly, the NRS node contains the status of the FIB on the managed CONET node. When a RIB entry becomes stale, it sends an invalidation message to the FIB. We began by measuring what happens today in the IP routing-plane. We then used the results to derive conclusions that were valid for CONET.

BGP statistics reported in [18] show that the sum of the average number of IP prefix updates and the average number of withdrawals is close to three per second. This figure provides an upper bound on the required rate of change for an IP FIB: in reality, only a subset of BGP updates require a modification of the IP FIB, depending on the topological location of the router. For instance, the FIB update rate for the router AS65000 [18] is about one per second.

Today there are approximately 2 10^8 Internet domain-names and about 400k IP prefixes. Thus, each IP prefix serves, on average, 500 domain-names. In our scenario, Principal strings are domain-names. Therefore, the number of route updates and withdraws that the ICN routing-plane would have to handle would be of the order of 1500 (i.e., 3*500).

A route update triggers an invalidation procedure, only if the ICN route is contained in the FIB. Therefore, the invalidation rate is equal to the route update rate multiplied by the probability that an ICN route is contained in the FIB. This probability is roughly equal to the ratio between the size of the FIB and the number of domain-names. In the deployment scenario we are considering, the FIB size contains of the order of 10^6 entries. Thus the

invalidation rate is of the order of 7.5 invalidations per second (i.e. $1500 * 10^6 / (2 * 10^8)$). This rate is easily supported by current technologies.

5.1.3.2.2 Test bed

In the next stage of our study we used a test bed setup to analyse the performance and limitations of the CONET Lookup-and Cache routing architecture.

5.1.3.2.2.1 Hardware setup

Figure 5.15 shows the network architecture for the tests. The setup consisted of two CONET sub-systems in which each sub-system was an IP network connected by a 100 Mbit/s Ethernet switch. Sub-system *A* contained three clients (i.e. CONET end-nodes). Sub-system *B* contained a server and an NRS node. The sub-systems were interconnected by a border node *N*, equipped with two network interfaces.

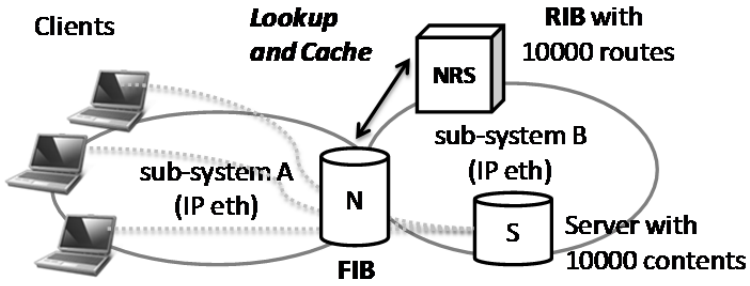


Figure 5.15: Test bed setup

5.1.3.2.2.2 Software setup

All devices ran Linux OS. Clients, node *N* and server ran an implementation of CONET based on a modified version of CCNx release 0.5.0 [4], in which we implemented our own Lookup-and-Cache routing scheme, using the Least Recently Used (LRU) and the Inactivity Timeout (ITO) algorithms, described in Deliverable D5.3 and in [1] (see also footnote 12).

Client FIBs had a single default route toward node *N*. Hence, clients did not perform Lookup-and-Cache procedures. Node N did not have a default route and used the Lookup-and-Cache routing-by-name mechanism to feed its FIB, up to a fixed size of 100 routing entries.

---

[12] Briefly, the ITO algorithm is based on estimated route activity. If a route contained in the FIB is active, than it cannot be not removed from the FIB, unless there are new routes that have been waiting to be inserted in the FIB for at least *MPT* seconds (Maximum Pressuring Time). In this sense, we could classify the ITO algorithm as *partially non pre-emptive*. An active route in the FIB cannot be interrupted, except in the case of an MPT timeout. If the FIB is full when a new Interest CIUs arrives, the required route is not present in the FIB and the route cannot be inserted in the FIB since all routes are active, and the incoming Interest CIU is discarded. It will then be retransmitted by transport layer mechanisms.

The server had a repository containing 10 000 content items; content size followed a Pareto distribution (k=133, α = 1.1) [21]; each content item was divided into 4 kByte chunks [13].

The NRS node had a RIB indexing all 10 000 contents. The next-hop value for each RIB entry was the IP address of the server. The RIB was implemented using the Bind-9 DNS server; therefore, the exchange of Lookup-and-Cache messages complied with the DNS protocol. To emulate delay on a wide area network, we artificially set up a two-ways delay of 100 ms between node *N* and the NRS node. *Local* lookup-and-cache procedures, such as inserting the route in the FIB, de-queuing waiting Interest messages, etc. added an additional 60 ms delay. Thus each lookup-and-cache operation added an average delay of 160 ms to an incoming Interest message with no route in the FIB.

The size of the "content" (i.e. named-data) cache was set to zero for all nodes, ensuring that *content* caching had no influence on performance. The CCNx InterestLifeTime parameter was set to 1 sec.

5.1.3.2.2.3 Persistent workload

A crucial variable for studies of Lookup-and-Cache routing is the number of *active-routes* – a variable that can also be used to compare different route replacement strategies. We therefore set up a *persistent* workload model that generated a constant number of active-routes. The workload guaranteed that the number of concurrent downloads performed by clients was constant. Each download fetches a content that had never been downloaded before. This meant that the number of active routes was equal to the number of downloads. Each time a download ended, a new download started 50 ms later. If the first Interest message was dropped, the message was resent 2 seconds later. To download an item of content, clients used the ccncatchunks2 utility [4], configured with a maximum congestion window of four chunks. Thus, to implement a scenario with 30 active-routes we ran ten instances ccncatchunks2 on three clients, with each ccncatchunks2 instance downloading different content, continuing this scheme for the duration of the test (10000 downloads).

5.1.3.2.2.4 Test bed results

We compared Lookup-and-Cache performance with a FIB limited to 100 entries, to performance with an *unlimited* FIB, in which *all* routes were preloaded. We repeated each test five times, measuring average performance and 95% confidence intervals.

[13] We also analysed cases in which all items of content were of the same size, repeating the analysis for different sizes (500 kB, 1MB, 10MB). The resulting behaviour was very similar to the behaviour reported here.

5.1.3.2.2.5 Analysis of delays and lookup rates

Figure 5.16 plots average download time against number of active-routes. Figure 5.17 shows the average number of lookups per download, which is equal to the ratio between the number of lookups performed by node *N* and the overall number of downloads. In what follows, we begin by analysing the performance of an unloaded FIB (number of active-routes lower than 100); we then considered an overloaded FIB.

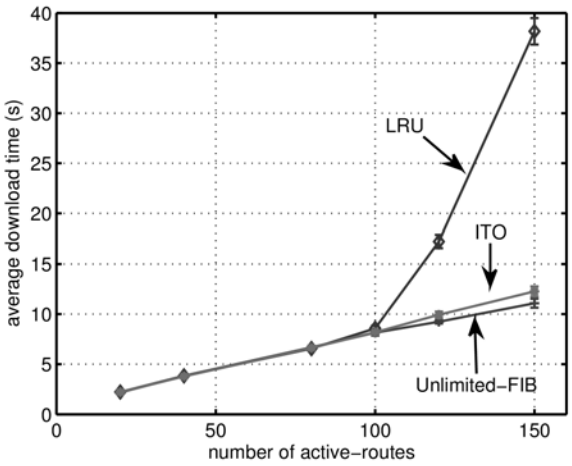


Figure 5.16: Average download time versus number of active-routes

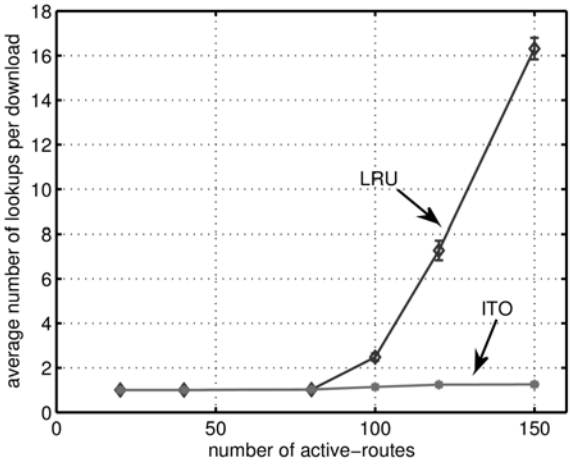


Figure 5.17: Lookup rate versus number of active-routes

5.1.3.2.2.6 Unloaded FIB

As expected, Lookup-and-Cache works well when the number of active-routes is lower than the size of the FIB (100 routes). Download times are comparable with times for the case for an unlimited-FIB case; the number of lookups per download is close to one. This means that after a first lookup-and-cache cycle, the route is held in the FIB for the duration of the download time. The download time is given by the sum of the initial lookup delay (160 ms, if the route is not already in the cache) plus the content download time. This latter time depends on available link capacity and thus increases linearly with the number of competing downloads (active routes). In unloaded conditions ITO and LRU perform almost in the same way.

5.1.3.2.2.7 Overloaded FIB

When the number of active-routes is equal or greater than the FIB size, performances start to degrade. The greater the overloading, the greater the degradation, with respect to the case of the unlimited-FIB. In these conditions, it is easy to detect the difference between different route replacement algorithms. With ITO, the download delay is a bit greater than in the unlimited-FIB case and the number of lookups per download remains quite limited. This is a

consequence of the non-preemptive behaviour of the ITO algorithm. With a FIB of 100 entries, only 100 downloads actually transfer data; other downloads wait for their routes to be cached in the FIB, when one of the 100 downloads ends or when the MPT timeout of 1 sec expires – something that happened only very rarely in our experiment. Thus, ITO schedules the use of transmission resource for 100 downloads at a time. With an unlimited-FIB, in contrast, all downloads transfer data concurrently. Therefore, the main difference between the two cases is in the way they schedule the use of transmission resources. This has only a minor impact on average delay.

Figure 5.17 shows that in both unloaded and overloaded conditions, ITO limits the number of lookups per download to 1.2. The rise of the lookup rate is due to transport level time outs, when route timeouts set by ITO elapse. In these cases, the route is removed from the FIB, and then briefly reinserted, increasing the number of lookup per item of content.

With LRU, the number of lookups per download and the download times increase significantly. When the number of active routes is equal or greater than the size of the FIB, the LRU algorithm pre-emptively replaces the FIB entries associated with active routes. The result is in/out flapping of routes in the FIB, leading to an increase in download delay and in the number of lookups per item of content. Although we have not tested this, we expect that if we further increased the overload, the number of lookups per contents would approach the upper bound of one lookup per Interest message - as in the limit case in which each route hold in the FIB is used to forward a single Interest message.

5.1.3.2.2.8 Fairness analysis

In the study reported here, we analysed the fairness performance of our architecture, evaluating the download bit-rate (goodput) as a function of content size, and checking if large contents (e.g. 10 Mbytes) were downloaded with a different goodput with respect to small contents (e.g. 200 Kbytes).

We organized contents into different *classes*, where each class included contents whose size was within a given range, (e.g., 0-500 Kbytes). For each class, we evaluated the average goodput achieved by downloading contents of that class. Then, we computed Jain's fairness index for class goodputs, defined [22] as follows:

$$F_J = \frac{(\sum_{i=1}^{N} g_i)^2}{N \sum_{i=1}^{N} {g_i}^2}$$

where g_i is the average goodput obtained by downloading contents of the i-th class and N is the number of classes. Perfect fairness is achieved when $F_J = 1$ and absolute unfairness is achieved when $F_J = 1/N$.

We first discuss the case of an unloaded, then that of an overloaded FIB.

5.1.3.2.2.9 Unloaded FIB

Figure 5.18 plots the Jain's fairness index against the number of active-routes for our two route replacement strategies (ITO and LRU). To perform the necessary computations, we arranged contents in classes in which the first class included content items with sizes in the range {0, 500 kbytes}, the second class contained items with sizes in the range {500 Kbytes, 1Mbytes}, and so forth. For numbers of active-routes lower than the size of the FIB (i.e. 100 routes), the performance of the two algorithms was similar, and was close to the performance achieved with an unlimited-FIB. We conclude that, in unloaded conditions, Lookup-and-Cache provides good fairness.

When the number of active–routes increased from 20 to 40, the fairness index improved, as seen in Figure 5.19 and Figure 5.20.

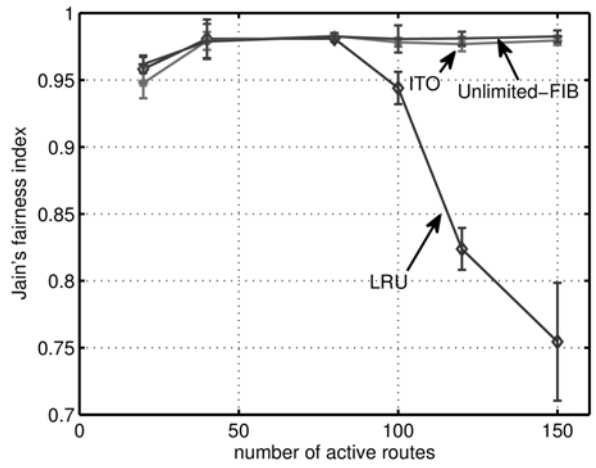


Figure 5.18: Jain's fairness index versus number of active-routes

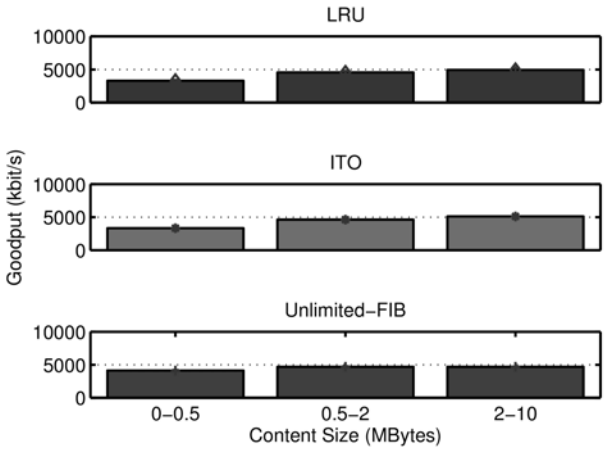


Figure 5.19: Goodput versus content size in case of 20 active-routes

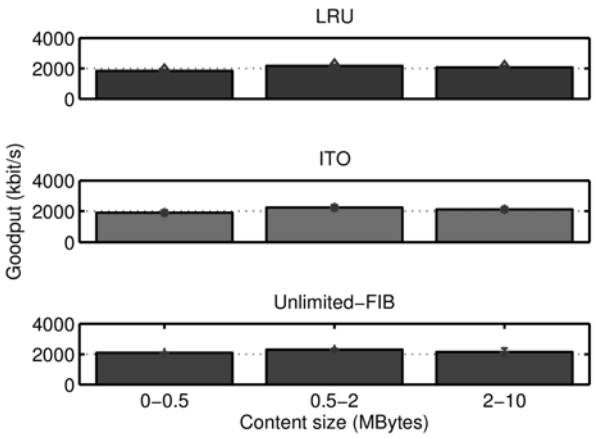


Figure 5.20: Goodput versus content size in case of 40 active-routes

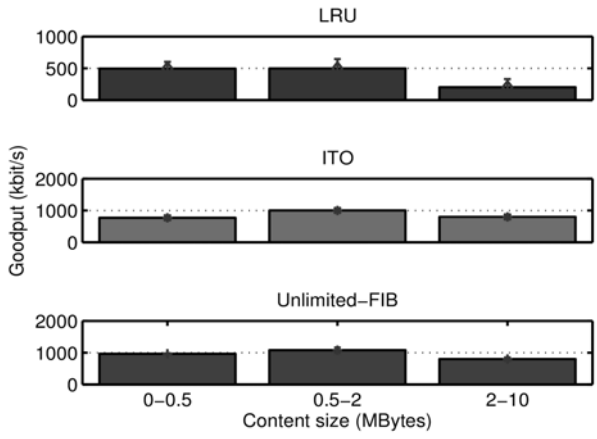


Figure 5.21: Goodput versus content size in case of 150 active-routes

With 20 active-routes, the download of small content items was slightly penalized, probably due to the time needed to execute Lookup and Cache procedures. In these conditions, therefore, Lookup and Cache was less fair than an unlimited FIB. This suggests that network paths connecting the NRS node to other network nodes should be properly dimensioned to make the round trip time as short as possible.

When the number of active-routes increased, download rates decreased. In these conditions, processing delays had a lower impact on the goodput of small contents and fairness improved, as can be seen in Figure 5.20, which shows performance with 40 active-routes.

5.1.3.2.2.10 Overloaded FIB

When the number of active-routes was equal to or greater than the FIB size, the fairness performance of ITO was very close to that of the unlimited-FIB case. Conversely, LRU performed less well with higher numbers of active-routes. This deterioration was primarily due to in/out route flapping events that penalized the download of large items of content. Figure 5.21 shows goodput for different classes of content with 150 active-routes. With LRU, the greater the content size, the lower the goodput.

Finally, we considered the effectiveness of the ITO MPT timeout as a means of avoiding traffic starvation in the presence of routes that remain active for a long time. In this analysis, we consider a FIB with three entries, and four downloads involving content of *unlimited* size, starting at times 0, 0.2, 0.4, and 0.6 s. For the ITO algorithm, we set the MPT to 1 second. Figure 5.22 plots the presence (in) and the absence (out) of routes in the FIB against time. We observe that when the route of a download is absent in the FIB, this condition persists for about 1 second, i.e. the value of the MPT timeout.

5.1.3.2.2.11 QoS analysis

In this section, we evaluate the priority-ITO algorithm described in Deliverable D5.3 and in [1], where different classes of traffic have different priorities in using the FIB space. We consider three classes of traffic. With reference to the scenario in Figure 5.15, client n.1 generates traffic of class 1 (lower priority), client n.2 generates traffic of class 2 (medium priority) and client n.3 generates traffic of class 3 (highest priority). At the start of the test, all clients request 50 different items of content, each of size 10 Mbytes. The size of the FIB is set to 50 entries, meaning it is limited to supporting 50 concurrent downloads. Figure 5.23 plots average download time measured for different classes of content. As expected, the greater the class priority, the lower the delay.

During the test we observed three phases. First, the 50 downloads in class 3 filled the FIB; in this phase downloads were limited to content in class 3 and downloads of other classes of content had to wait for space in the FIB. In the second phase, downloads of class 3 content

came to an end and the FIB was filled with 50 downloads of content belonging to class 2. Now downloads were limited to content in class 2, and downloads of class 1 had to wait. In the third phase, downloads of class 2 came to an end, allowing content belonging to class 1 to start downloading.

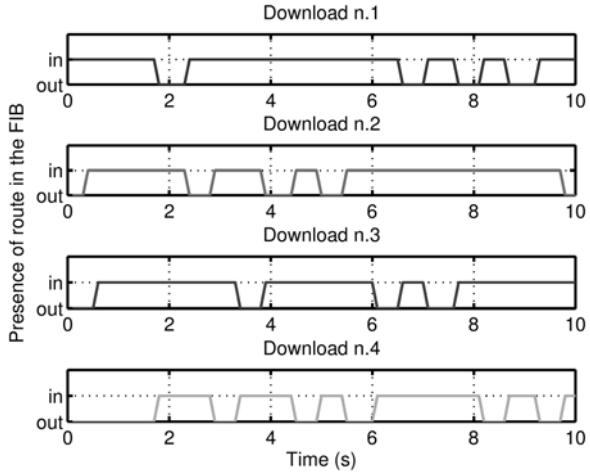


Figure 5.22: Presence and absence of routes in the FIB

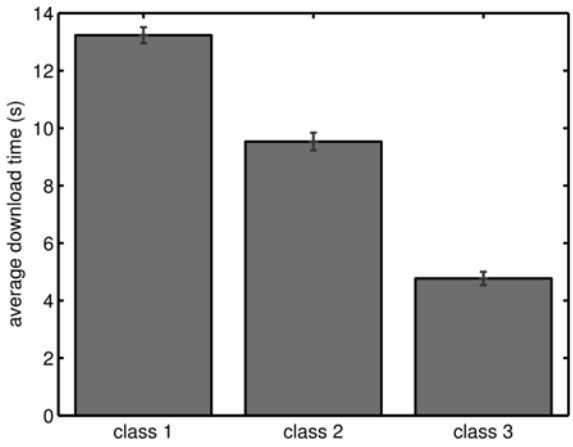


Figure 5.23: Goodput versus class of traffic, in case of priority-ITO

5.2 Evaluation of the CONVERGENCE publish/subscribe infrastructure

5.2.1 Protocol Overview

CONVERGENCE has introduced ANUBIS, a content based publish subscribe protocol that exploits users' repetitive behaviour when they search for content to introduce a distributed cache for storing queries and information related to search results. ANUBIS is a two-layer protocol. The first layer uses a gossip protocol to route content to any possible destination[14]. The second layer exploits the fact that users often make the same queries to provide a distributed cache [24]. The caching scheme uses a DHT in which the peers are the same as in the first layer (i.e. the publisher only). Each peer caches a metafile which uses the query hash as the key together with meta-information about subscribers who have downloaded the content referenced by that key. Subscribers then use a peer to peer protocol to share content that corresponds to the same (repeated) query stored in the distributor caches. In this way, the

[14] In the literature, content and keyword based- routing can be performed either with random walks, which do not have a high success rate, or with flooding protocols [23], such as gossip.

query is performed only once for all the publishers, and ANUBIS achieves both reduction of lookup messages and better load balancing than would be possible in the plain flooding protocols used in the literature.

5.2.2 System Setup

The authors of [24] suggest that 80% of queries in a distributed system are duplicates. To measure the performance of ANUBIS, we considered a scenario in which 1500 subscribers were performing queries. First we analysed a relatively small system with 1500 publishers, in which subscribers made 1500, 3000, 4500 and 6000 queries with repetition probabilities ranging from 0.0 to 0.8/ query with step 0.2. When this system performed as expected, we went on to consider systems with 4800, 9600 and 14400 publishers, in which 1500 subscribers performed 15000, 30000, 45000 and 60000 queries, again with repetition probabilities ranging from 0.0 to 0.8/ query with step 0.2. In both cases, we defined queries in such a way that each query matched ten publications uniformly distributed among publishers.

5.2.3 Experimental Results

ANUBIS aims, on the one hand, to reduce the overhead involved in discovering publications matching a query and, on the other, to balance loads when transferring results (publications) to subscribers. Therefore, the goal of our simulations was to measure the overhead (number of lookup messages) necessary to discover the publications matching a given subscription and, then, to measure the number of result (publication) messages sent by publishers and by subscribers.

In our initial network with 1500 peers, hit rates were higher than 98%, as shown in Figure 5.24. As the number of queries grew from 1500 to 6000, the hit rate also increased, reaching close to 99.5%. This is evidence that the protocol behaves best when demand is high. The graph also shows that even with different repetition probabilities, performance converges for high numbers of queries.

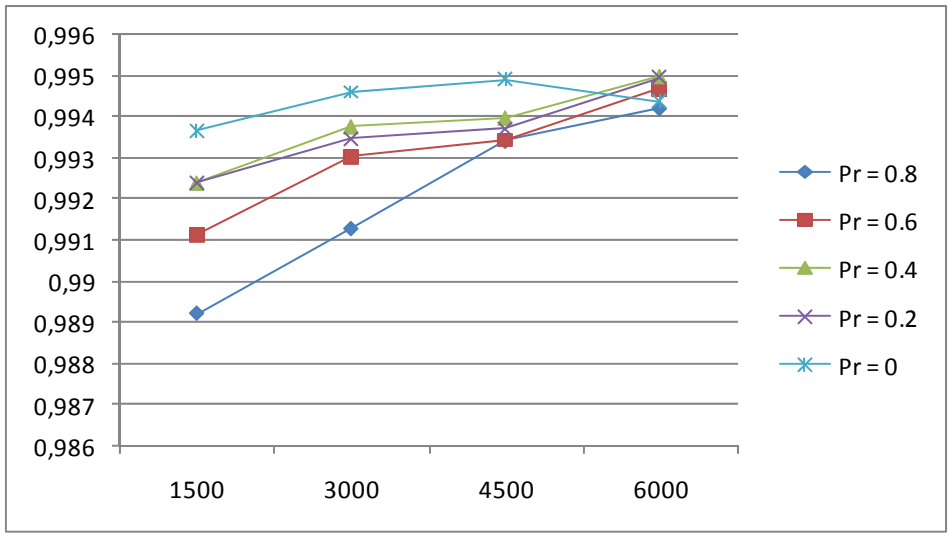


Figure 5.24: Hit Rate with 1500 publishers

Figure 5.25 shows the number of messages sent directly by publishers to subscribers (overhead) i.e. the number of occasions in which subscribers could not retrieve content from other subscribers. The graph clearly shows how ANUBIS exploits query repetition; the rate of growth of messages with respect to the number of queries is much lower for higher repetition probabilities.

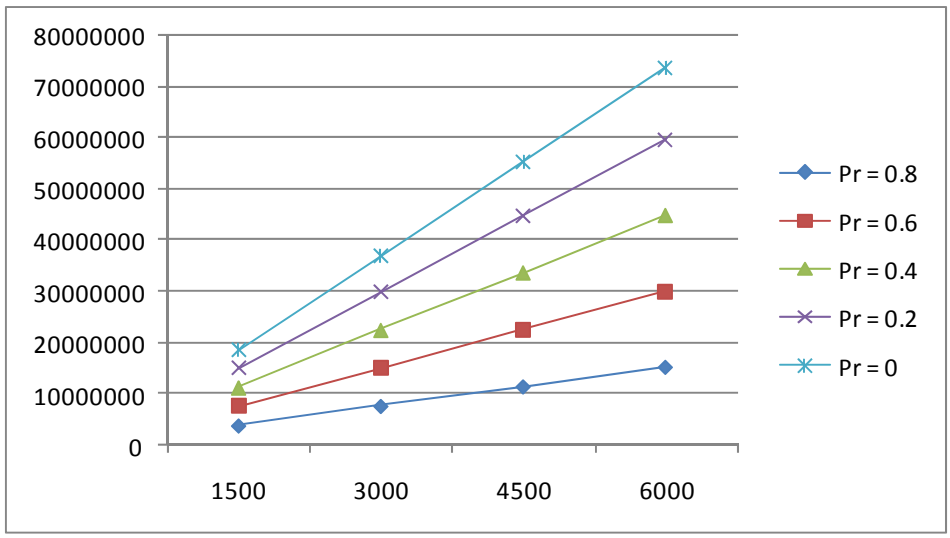


Figure 5.25: Overhead for 1500 publishers

Figure 5.26 compares the number of content (match) messages coming from publishers and from subscribers. The results show clearly that higher repetition probabilities lead to better-balanced loads.

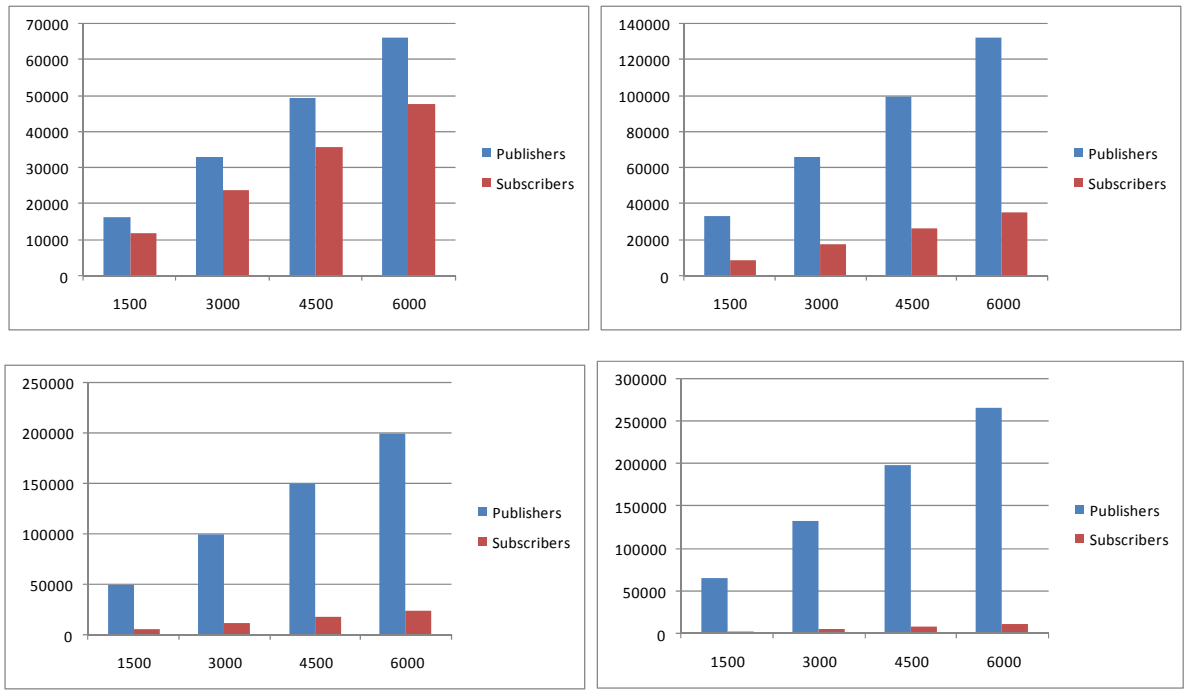


Figure 5.26: Load Balancing for Matches

Figures Figure 5.27, Figure 5.28, Figure 5.29 and Figure 5.30 plot the number of publisher messages for 15000, 30000, 45000 and 60000 queries respectively. Even though this grows linearly, a consequence of the use of the gossip protocol for the initial discovery, it is clear that ANUBIS effectively exploits query repetition.

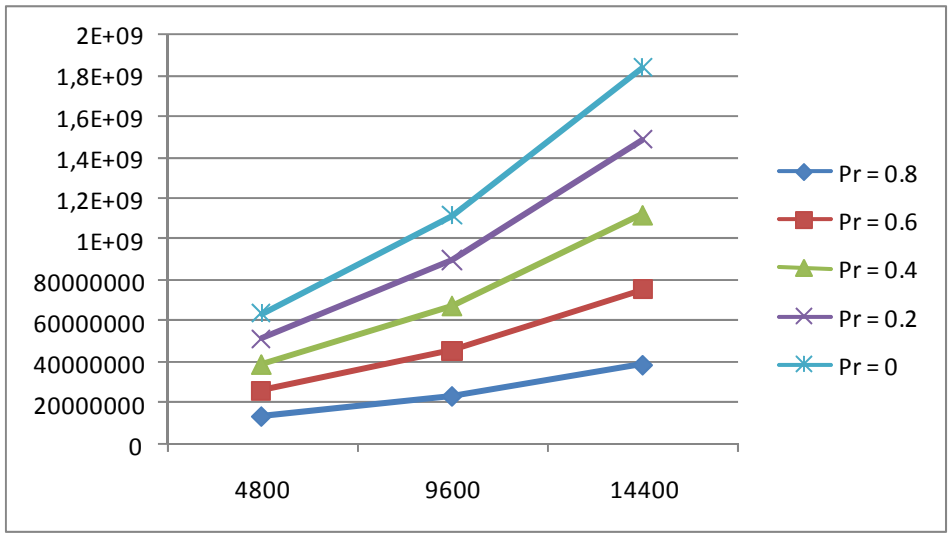


Figure 5.27: Publishers lookup messages for 15000 queries

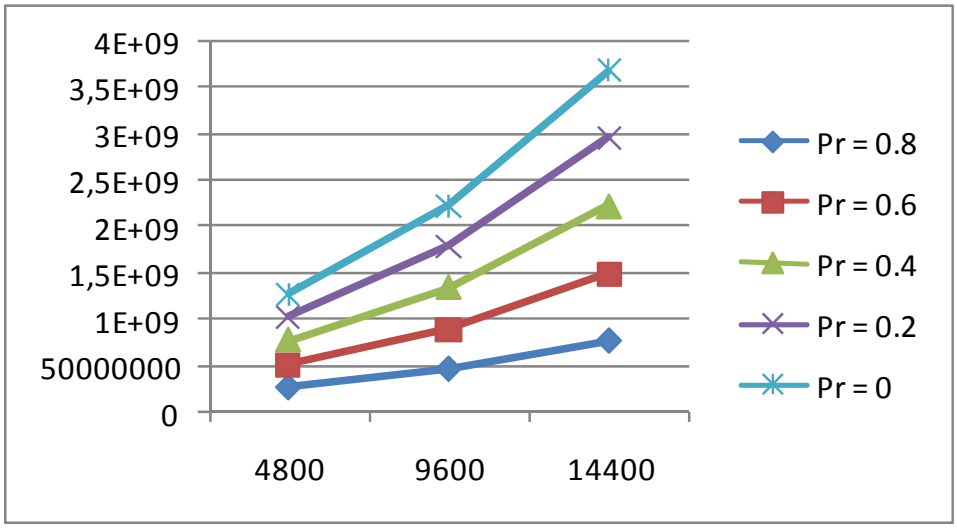


Figure 5.28: Publishers lookup messages for 30000 queries

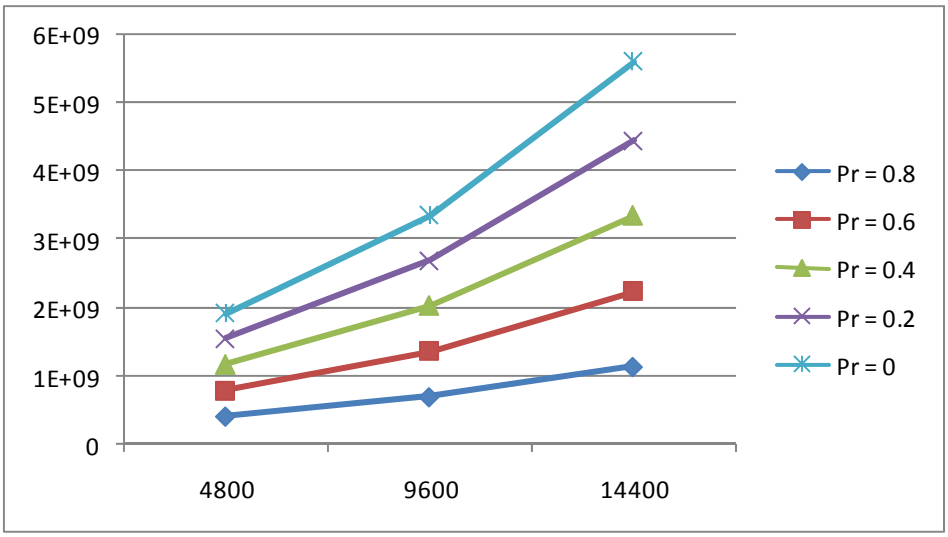


Figure 5.29: Publishers lookup messages for 45000 queries

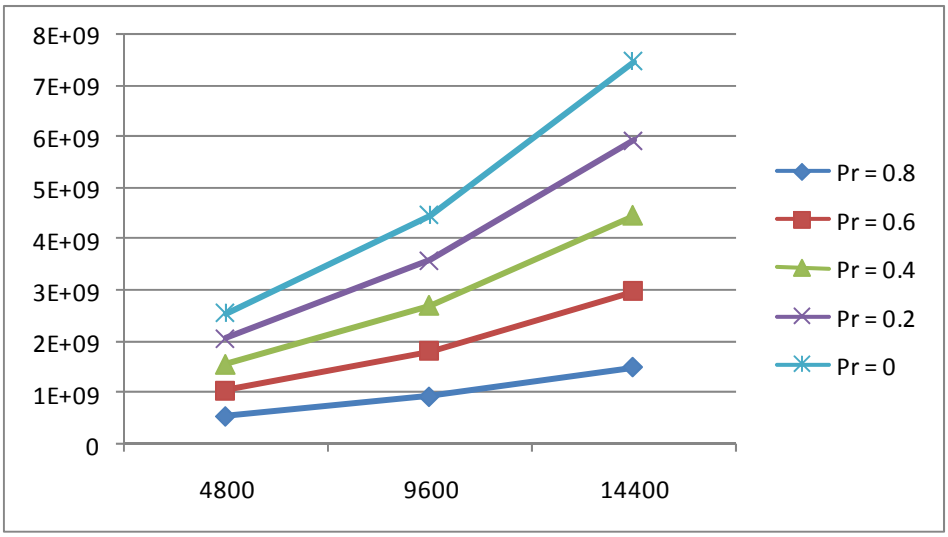


Figure 5.30: Publishers lookup messages for 60000 queries

5.2.4 Conclusions and next steps

To achieve high hit rates ANUBIS uses the gossip protocol. Necessarily, therefore, the behaviour of the protocol is linear. Nonetheless, the data we have presented demonstrates that the protocol is highly effective in exploiting repetition probability and that higher probabilities are associated with lower overhead.

The model we have used so far (limited query repetition) is rather conservative. In the next stages of our work, we will test other models that have recently emerged in the literature. These will include the model presented in [25], which suggests that query repetition is much more frequent than previously thought and that it follows a lognormal distribution. In these conditions we expect ANUBIS to yield an even better performance, as the number of repeated queries can reach much greater magnitudes than the authors of [24] are suggesting and, therefore, the cache is expected to be used even more often leading to a further reduction in the lookup and content messages.

6 References

[1] A. Detti. M. Pomposini, N. Blefari-Melazzi, S. Salsano, “Supporting the Web with an Information Centric Network that Routes by Name – Extended version”, available at http://netgroup.uniroma2.it/Andrea_Detti/Lookup-and-Cache/comnet-web-ICN-ext.pdf

[2] V. Jacobson, et al., ”Networking named content”, in Proc. of ACM CoNEXT 2009

[3] A. Kuzmanovic, E.W. Knightly “Receiver-Centric Congestion Control with a Misbehaving Receiver: Vulnerabilities and End-point Solutions”, Elsevier Comput. Network. 2007, 51, 2717–2737

[4] CCNx project web site: www.ccnx.org

[5] N. Blefari Melazzi, M. Cancellieri, A. Detti, M. Pomposini, S. Salsano: “The CONET solution for Information Centric Networking”, Tech. report, available at http://netgroup.uniroma2.it/CONET.

[6] Xiaoliang Zhao, Member, IEEE, Dante J. Pacella, and Jason Schiller, “Routing Scalability: An Operator’s View”, IEEE Journal on Selected Areas in communications, vol. 28, no. 8, October 2010

[7] “Cisco Carrier Routing System”, available at http://www.cisco.com/en/US/prod/collateral/routers/ps5763/prod_brochure0900aecd800f8118.pdf

[8] “Juniper T Series Core Routers” available at http://www.juniper.net/elqNow/elqRedir.htm?ref=http://www.juniper.net/us/en/local/pdf/datasheets/1000051-en.pdf

[9] G. Trotter , “Terminology for Forwarding Information Base (FIB) based Router Performance”, IETF RFC 3222

[10] D. Meyer , L. Zhang , K. Fall , “Report from the IAB Workshop on Routing and Addressing”, IETF RFC 4984

[11] D. Perino, M. Varvello, “A Reality Check for Content Centric Networking”, ACM SIGCOMM 2011, Workshop on Information-Centric Networking

[12] T. Koponen, M. Chawla, B.G. Chun, et al.: “A data-oriented (and beyond) network architecture”, ACM SIGCOMM 2007

[13] CAIDA Internet Trace Storage, https://data.caida.org/datasets/passive-2010/

[14] MAWI Working Group Traffic Archive, http://mawi.wide.ad.jp/mawi/samplepoint-F/2011/

[15] http://netgroup.uniroma2.it/Andrea_Detti/Lookup-and-Cache/

[16] Changhoon Kim, Matthew Caesar, Alexandre Gerber, and Jennifer Rexford, "Revisiting route caching: The world should be flat," in Proc. Passive and Active Measurement Conference, April 2009.

[17] L. Breslau et al., “Web Caching and zipf-like Distribution: Evidence and Implications”, in Proc. IEEE INFOCOM, 1999

[18] “BGP Routing Table Analysis Report”, available at http://bgp.potaroo.net

[19] “Daily estimated size of World Wide Web”, http://www.worldwidewebsize.com/

[20] “BGP Routing Table Analysis - DIX-IE Data”, http://thyme.apnic.net/current/

[21] P. Barford and M. Crovella, “Generating Representative Web Workloads for Network and Server Performance Evaluation”, In Proc. of ACM SIGMETRICS ’98

[22] R. Jain. The Art of Computer Systems Performance Analysis. John Wiley & Sons, 1991

[23] R. Ahmed, R. Boutaba, A survey of distributed search techniques in large scale distributed systems, Communications Surveys Tutorials, IEEE 13 (2) (2011) 150-167.

[24] B. Smyth, E. Balfe, J. Freyne, P. Briggs, M. Coyle, O. Boydell, Exploiting query repetition and regularity in an adaptive community-based web search engine, User Modeling and User-Adapted Interaction 14 (5) (2005) 383–423.

[25] C. Bauckhage, Insights Into Internet Memes, in Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media, 2011.