MIESIĘCZNIK STOWARZYSZENIA ELEKTRYKÓW POLSKICH

Przegląd Telekomunikacyjny

WIADOMOŚCI TELEKOMUNIKACYJNE

8–9 / 2012

WYDAWNICTWO SIGMA NOT

ISSN 1230-3496

Cena: 50,40 zł (w tym 5% VAT)

INSTYTUT ŁĄCZNOŚCI
PAŃSTWOWY INSTYTUT BADAWCZY

KTIT 2012

XXVIII Krajowe Sympozjum Telekomunikacji i Teleinformatyki

W kierunku Nowych Generacji Sieci

12-14 września 2012 r. Warszawa - Miedzeszyn

INSTYTUT ŁĄCZNOŚCI
PAŃSTWOWY INSTYTUT BADAWCZY

KSTiT 2012

Szanowni Państwo – Uczestnicy KSTiT,

XXVIII Krajowe Sympozjum Telekomunikacji i Teleinformatyki odbywa się w tym roku w Miedzeszynie i jest po raz pierwszy organizowane przez Instytut Łączności – Państwowy Instytut Badawczy. Tematem przewodnim tegorocznego Sympozjum są sieci nowej generacji – **NGN** (*Next Generation Networks*) oraz Internet Przyszłości (*Future Internet*). Tematyka ta stanowi jeden z ważnych nurtów w obszarze **ICT** (*Information and Communication Technologies*) w ramach 7. Programu Unii Europejskiej. Jej kontynuowanie przewidziano w następnej edycji projektów europejskich *Horizon 2020*.

Program techniczny konferencji obejmuje sesje tematyczne dotyczące tradycyjnych już zagadnień Sympozjum, to jest architektur i protokołów telekomunikacyjnych, pól komutacyjnych, systemów radiowych, sterowania i zarządzania sieciami, systemów transmisyjnych, bezpieczeństwa w sieci, przetwarzania sygnałów, jakości usług w sieciach, kompatybilności elektromagnetycznej, optoelektroniki oraz usług i aplikacji.

Ponadto w ramach Sympozjum zorganizowano sesje specjalne przeznaczone prezentacji rozwiązań realizowanych obecnie przez zespoły krajowe w ramach 7. Programu Ramowego, projektów POIG finansowanych ze środków strukturalnych oraz projektu POIG – *Inżynieria Internetu Przyszłości*, który to projekt jest wykonywany przez dziewięć organizacji i uczestniczących w tym przedsięwzięciu ok. 150 pracowników merytorycznych.

Programowi technicznemu towarzyszy wystawa, która w tym roku jest szczególnie interesująca, gdyż obejmuje pokaz rozwiązań wypracowanych podczas realizacji projektów.

Podsumowując ten krótki wstęp, trzeba podkreślić, że tegoroczne Sympozjum dobrze obrazuje stan wymienionych prac badawczych prowadzonych obecnie w kraju i będzie jak zwykle ożywionym forum dyskusji i wymiany informacji w tej dziedzinie.

Przewodniczący KSTiT
prof. dr hab. inż. Wojciech Burakowski

Andrzej Bęben, Jordi Mongay Batalla, Warsaw University of Technology
David Florez, Telefónica Investigación y Desarrollo
Wei Koong Chai, University College London
Spiros Spirou, Intracom Telecom
Ning Wang, University of Surrey
Michael Georgiades, PrimeTel PLC

The content mediator architecture for content-aware networks[1]

The paper proposes a novel content-oriented Internet architecture, which was designed and developed in the frame of EU FP7 COMET (COntent Mediator architecture for content-aware nETworks) project. The COMET architecture aims to simplify access to any type of content regardless of location and support efficient content delivery in a content-aware fashion. The key novelties of COMET system are: the global content naming scheme, the mediation mechanism supported by server and network awareness, and scalable content forwarding method. These mechanisms enable COMET system to select the best content source and content delivery path based on the knowledge about available content replicas, the status of content servers and available routing paths. The paper presents details of designed mechanisms and algorithms, outlines federated testbed environment and describes experiments that proved effectiveness of our approach.

1. Introduction

The Content Aware Networks (CAN), also known as Information Centric Networks (ICN), are designed to natively support access and delivery of multimedia content. The motivation for CAN comes from growing multimedia content (especially user-generated), abundance of content intermediary systems, like YouTube, Flickr, etc., and significant increase of multimedia traffic carried in today's Internet (even up to 80%, as reported in [1]). Although these are supported by the Internet today, there are a number of recognised limitations of the current host-centric Internet [2]. Among others the most significant are: (1) segmentation of content coming from the lack of global naming scheme, (2) inefficient content delivery due to network unawareness of transferred content and available replicas, (3) lack of anycast and point-to-multipoint connections, (4) dependency of content identifier on location, which leads to URL broken-link problem. The above problems were only partially solved by currently exploited content dissemination systems such as Content Delivery Networks (CDNs) [3] and peer-to-peer (P2P) networks [4] due to the fact that these solutions operate as overlay networks on today's Internet.

The aforementioned limitations motivate transition from the current Internet to a new "content–centric" network, which treats delivery of content as the key paradigm. Recently, several new concepts for improving content delivery have been investigated in [5], [6], [7], [8] and [9] that cover different approaches for routing by content name, location independent and self-certifiable content addressing schemes, scalable content resolution algorithms, publish/subscribe systems and in-network content caching [10], [11].

The EU FP7 COMET project (COntent Mediator architecture for content-aware nETworks) [12] takes a unified approach to content location, access and distribution, irrespective of the intermediary used. COMET designed a global naming scheme along with mechanisms for optimizing both content source selection and distribution, by mapping content to appropriate network resources based on transmission requirements, user preferences, and network state. In this paper, we present a novel architecture for converged content handling services that simplifies access to content across multiple domains with network-aware distribution capability enabling end-to-end QoS. The aim is

[1] The research leading to these results has received funding from the EU FP7 COMET project - Content Mediator Architecture for Content-Aware Networks, under Grant Agreement 248784

to provide a unified interface for Internet-scale content discovery and access regardless of the temporal nature of the service and physical hosting servers. It is also capable of applying the most appropriate end-to-end transport strategy for different types of content services across multiple autonomous domains with heterogeneous QoS provisional policies. Our approach allows individual ISPs to have internal QoS provisioning methods, which are however, not exposed to third party ISPs outside the Federation. This is typically realized through dedicated content mediation brokers residing in each autonomous domain which collaborate with each other for supporting inter-domain QoS aware content access and delivery. Such a paradigm can readily be deployed over the current IP-based Internet without any radical deployment changes.

The paper is organised as follows. Section 2 presents the concept of content mediation which acts as a base for the COMET approach. Section 3 describes the COMET architecture focusing on basic processes, functions and entities. The proposed algorithms and mechanisms related to content naming, awareness processes, content resolution and delivery are presented in section 4. Section 5 discusses the COMET experimental environment and tests. The experiments and results illustrating the effectiveness of our approach are presented in section 6. Finally, section 7 summarises the paper and gives outline on further works.

2. The concept of the COMET system

In the COMET system, content are given unique content names that are solely used for the identification of the content. This is in contrast with the current IP address system whereby the address is overloaded both as an identifier (i.e., identification of the content) and a locator (i.e., the actual location of the content server). Exploiting the name-based content access paradigm, the COMET system proposes the concept of *mediation* where a novel mediation plane operating between the content ecosystem and the underlying network infrastructure is envisaged. It is designed as a tightly coupled overlay provisioned by a set of collaborative Internet Service Providers (ISPs) to mediate content access and delivery in a holistic manner. The conceptual depiction of this plane is shown in Fig. 1.

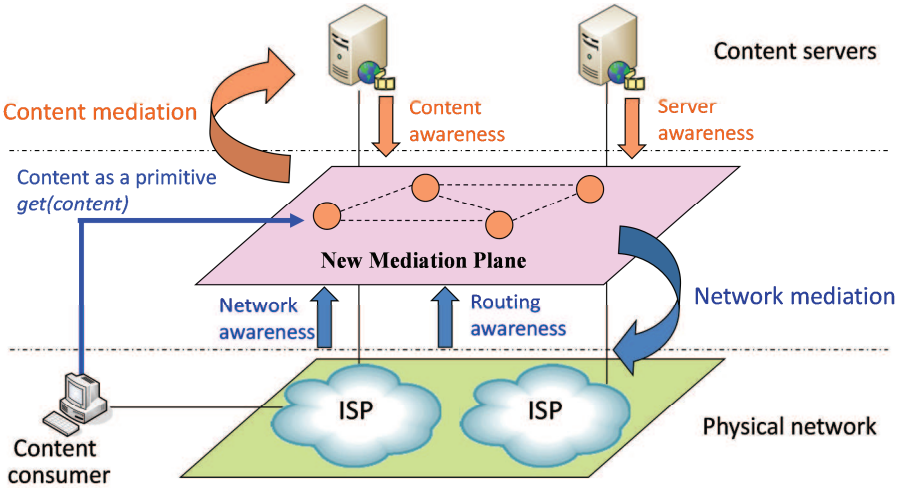

Fig. 1: The concept of COMET mediation plane.

Current network distribution platforms (e.g., content delivery networks (CDNs) and peer-to-peer (P2P) systems) operate as overlays that are oblivious of the network conditions and content characteristics. Recently, a closer collaboration between the overlays and the underlays has been investigated. For example, the IETF ALTO framework [13], [14] and the P4P [15] are proposed to provide network usage information to the overlays such that the overall content delivery performance can be optimized. The COMET system exploits the fact that an ISP-operated overlay has native access to these information and thus, is naturally suitable for assisting the selection of the best content source and delivery path.

The mediation operations rely on four types of awareness. For determining the most suitable content copy, the mediation plane will act as mediator for content publication, offering an interface for content publication and thus becoming aware of content characteristics, QoS requirements, etc. (content awareness), as well as the available content sources and their performance (server awareness). Since the mediation plane is provided by ISPs, it is also aware of the underlying network topology and the available routing paths between content servers and clients (routing awareness), as well as the network conditions (network awareness). These will enable the COMET system to effectively decide on the best content delivery option available.

Fundamentally, equipped with the awareness described above, there are two types of mediation operations in the COMET system.

- Content Mediation – this refers to the resolution of a content request to find the best available content copy. Primarily relying on the content and server awareness, if more than one copy of the same requested content is found, the mediation plane will be able to decide which copy is to be retrieved.
- Network Mediation – this refers to the delivery of content via the best transport strategy to both satisfy the user QoE and achieve high effective bandwidth utilization. The mediation plane allows ISPs to mediate content distribution by offering a common unified interface for content consumption. The ISPs are able to instruct the network on the best path(s) to use based on the network and routing awareness.

Key technical advantages that can be achieved thanks to this mediation are:

- Unified access to the content whatever its nature and location.
- Content delivery with guaranteed QoS.
- Point-to-multipoint content delivery, reducing bandwidth needs for live contents.
- Graceful handover of the content delivery path, providing more resilience and flexibility for multi-homed users.
- Advanced publication mechanisms, allowing Content Providers to update content servers on-the-fly, while switching among different ways of distribution.

3. The COMET architecture

COMET follows a two-plane approach, namely the Content Mediation Plane (CMP) and the Content Forwarding Plane (CFP). The CMP is responsible for name and content resolution as well as the preparation of path used for content delivery while the CFP is responsible for the content delivery. This is done based on mediation performed by the CMP taking into account the information about server and network conditions. Fig. 2 depicts the overall COMET functional view.

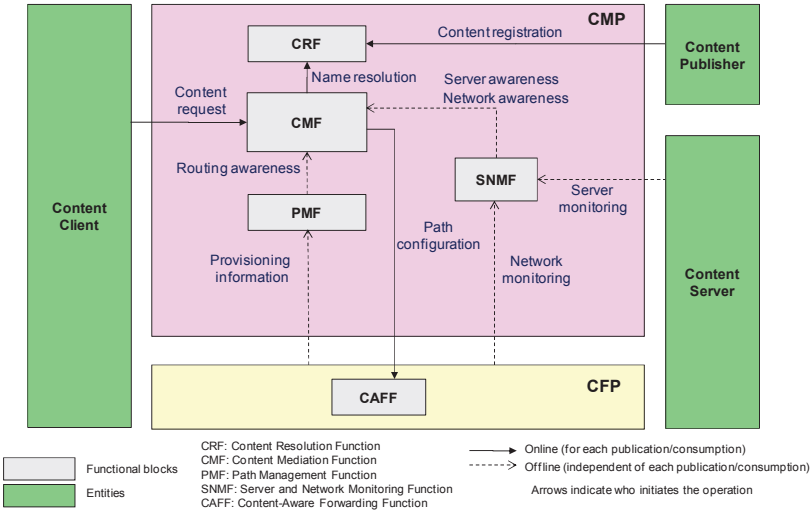

Fig. 2: The COMET system architecture.

The main functional blocks are described below.

- **Content Resolution Functional block (CRF)** – The primary tasks for this function is to maintain content records and to resolve content requests by resolving content names to content properties included in the content records. This function is invoked during the content publication process whereby a new content record is created and maintained for each new content published. The content record maps the content name to the physical content source(s) in the network. It also contains content metadata (e.g., aliases, media type etc.) and server load condition(s). The content resolution functions were implemented in Content Resolution Entity (CRE). A content request will trigger the CRE to resolve content names to content properties.
- **Path Management Functional block (PMF)** – It is responsible for the gathering of network reachability information (NRI) which relates mainly to the underlying network topology (i.e., inter-domain connections/links, link availability etc.) and the information related to QoS capabilities (e.g., QoS classes supported along the route). It is important to note that this information is long-term information. The path management function is performed by Routing Awareness Entity (RAE).
- **Server and Network Monitoring Functional block (SNMF)** – It is responsible for the collection of status data regarding: i) content servers, namely their availability and load and ii) the underlying network conditions, namely ingress and egress load on peering links, and load on access links (e.g., for admission control). This information is then fed to the CMF for it to make the appropriate decisions during the Content Consumption operation. The server and network monitoring functions were implemented in Server and Network Monitoring Entity (SNME).
- **Content Mediation Functional block (CMF)** – Effectively, the CMF is the "decision maker" in the CMP where it gains the necessary awareness from the CRF, SNMF and PMF and then mediates the content resolution and delivery. Specifically, it selects the best content copy based on the server and network conditions (from the SNMF) and decides the best delivery path using the information on the available paths (from the PMF). It interacts with the CFP to perform the necessary configuration on the delivery paths. The CMF function was implemented by Content Mediation Entity (CME).
- **Content-Aware Forwarding Functional block (CAFF)** – It is the main function in the CFP and is responsible for forwarding content via the paths determined by the CMF. The CAFF will have enhanced capabilities in order to provide the required QoS for end-to-end content delivery. These capabilities include traffic classification, point-to-multipoint forwarding, and masquerading functionality in order to hide the Content Server's IP address to the Content Client. The CAFF was implemented as Content Aware Forwarding Entity (CAFE), which is a specialised router enhanced with CAFF logics. Also, it is not necessary for all routers in a network to be collectively upgraded to handle COMET functions. It is sufficient at the early deployment stage that only border routers become the CAFE.

We briefly illustrate a typical flow of operations in the COMET mediation system. When a content publisher wants to make available a new content in the Internet, this content will have to be registered to the CRE that is responsible for the global maintenance of a distributed content repository. This repository contains the content records of all registered content items. The CRE may be realized with a set of content resolution servers arranged in a hierarchical structure similar to the current DNS system for scalability reasons. A content consumer requests a content using its content name via a unified interface provided by the CME. Upon reception of the content request, the CME will communicate with the CRE to find the content locations of the requested content. Based on the location information, the CME will check the actual server load conditions provided by the SNME and also the path reachability information provided by the RAE. The best server with a corresponding delivery path will be decided based on this information. Once the server and path are determined, the CME will configure the underlying content delivery path, i.e., interacting with the edge CAFE located close to the content server, for the actual content transmission.

4. The mechanism and algorithms

In this section we present key mechanisms and algorithms designed for both COMET planes. In the CMP, we defined new content identifier scheme and name resolution mechanism, the server and routing awareness processes as well as the decision algorithm, which selects the best content source and content delivery path based on the knowledge provided by awareness processes. On the other hand in the CFP, we defined stateless content forwarding method which allows dynamic selection of content delivery path for each content request.

4.1 Content Mediation Plane

Content identification in COMET consists of a content naming scheme, a content record specification and a name resolution mechanism. The naming scheme and the resolution mechanism are based on the Handle System [16]. The Handle System was selected because of its persistent and location-independent naming, extensible namespace, distributed architecture, security and efficient resolution. A content name in COMET is a Uniform Resource Identifier (URI) of the form:

```
<comet> :: <Naming Authority> "/" <Local Name>
```

The <Naming Authority> is a globally unique alphanumeric string - given by a registrar such as IANA - that identifies an organization or person who owns and wants to publish content. This organization freely administers the <Local Name> string to uniquely identify a content object under the <Naming Authority> namespace. Examples of content names are warnerbros.com/aMovie, for an organization, and bob@gmail.com/mySong, for a person. Content publication includes the creation of a content record, which points to the location of the content object and is indexed by the content name. Content type, QoS requirements and transport parameters are also part of the content record. Content publication stores the record in the name resolution system, which maps content names to content locations. The name resolution system is a distributed database resembling the DNS system. It is logically organized in two levels of hierarchy so as to reflect the <Naming Authority> and <Local Name> namespaces. Content mobility and replication are supported with the name resolution system by changing or adding location information in the content records.

The objective of the routing awareness process is to calculate content delivery paths in inter-domain network. The routing awareness is an off-line process performed in long time scale. It reacts to changes in inter-domain network reachability or re-provisioning of domains. The routing awareness is performed by Routing Awareness Entities (RAE). Comparing to BGP-4 inter-domain routing protocol, the RAE offers two new features: (1) multi-criteria (QoS) routing which allows RAE to build routing paths taking into account the requirements of COMET CoSs and (2) multipath routing as RAE builds a number of content delivery paths going towards a given prefix. These features provide COMET more information about the network and inter-domain routing paths, which enable COMET to deliver content with respect to QoS requirements as well as balance load in the network improving utilisation of network resources. The RAE provides information about discovered routes and their properties to the CME. This information is used by decision maker during content resolution process to select the best path for content consumption. Each RAE requires information about its domain, which among others covers: Autonomous System (AS) number, available network prefixes, IP addresses of peering RAEs, as well as CoS (Classes of Service) supported by COMET and values of their QoS parameters as maximum IP Packet Transfer Delay (IPTD), maximum IP Packet Loss Ratio (IPLR), maximum bandwidth (BW) for single content consumption assured between any ingress and egress points of its domain. Note, that values of QoS parameters should be valid in long term constituting an upper bound for actual values. This information should be provided by domain management system based on domain configuration, domain provisioning as well as SLA agreements between peering domains. Then, the RAE exchanges UPDATE or WITHDRAW messages with other RAEs located in peering domains to build content delivery routes. Each content delivery route is characterised by the list of AS numbers, supported COMET CoS and aggregated values of QoS parameters. The QoS parameters

are calculated from the egress CAFE of a source domain towards a given prefix. Moreover, RAE exchanges KEEP-ALIVE messages between peers in order to detect failures of peering RAEs. Whenever any change is detected, i.e. peer failure or domain re-provisioning, routing information is updated and propagated to the entire network.

Server Awareness is carried out in COMET by the SNME which will report information about server status to the CME for the servers in use inside the ISP managed by the CME. Typical metrics for assessing the server status are: available bandwidth, CPU and memory usage and disk occupation.

Once the CME gathers information about the content location, the available routing paths and the server status, it runs a decision algorithm, which is responsible for choosing the best server and path for delivery the content in the most efficient way. The CME takes decision for each content request independently. Each server has associated inherent information about the server's parameters as well as information related to the current state as e.g. server load. Each of the paths has attached information about path's characteristics obtained from RAE and information about the current state of the path provided by SNME. The information handled in the decision algorithm constitutes a set of complex parameters. This requires the use of multi-criteria algorithms. Multi-criteria algorithms are widely studied in the Multiple Criteria Decision Analysis (MCDA). MCDA considers both single and multiple objective optimization. In COMET, we considered multi-objective optimization, which assumes that the algorithm is able to supply a set of effective solutions. We leverage the multi-criteria decision algorithm presented in [15], which uses some a priori knowledge about the problem in order to select the effective solution. Without loss of generality, proposed decision algorithm uses three decision variables related to server load, path length and bandwidth. It could be easily extended to accommodate more decision variables. Our algorithm evaluates the impact of particular decision variable using two reference parameters, called reservation level and aspiration level. The reservation level is the hard upper limit for decision variable which should not be crossed by preferred solution. On the other hand, the aspiration level defines the lower bound for decision variable, beyond which preference of evaluated solutions is similar. The decision algorithm consists of three main steps:

Step 1: Decision Maker creates a decision space as the list of candidate solutions based on the information about servers and paths. Each candidate solution is a vector of decision variables, which has the following form:
Candidate solution [i]:

- *serverLoad* – this is numerical representation of server status,
- *pathLength* – the path length denotes the number of domains on the path. It is calculated based on the AS path parameter
- *bandwidth* – maximum supported bandwidth on the path.

Step 2: Decision maker calculates the rank value R_i for each candidate using objective function with reservation and aspiration levels specific for each decision variable.

$$R_i(.) = \min_{k=1,2,3}\left[\frac{r_k - q_k}{r_k - a_k}\right] \tag{1}$$

where: i is the number of candidate, k is the number of decision variable, q_k is current value of decision variable, r_k is a *reservation level* for decision variable k, while a_k is an *aspiration level* for decision variable k, which is determined as a multiplication of reservation level by aspiration coefficient α_k, $a_k=\alpha_k \times r_k$; and q_k is value of decision variable.

Step 3: Decision Maker selects the candidate with maximum rank as the best solution. Note that considered aggregate objective functions may have more than one effective solution into the Pareto optimal set. Thus, some tie-breaking rules (e.g., lower server load is preferred or just random selection) are required to ensure only one solution is selected.

As a result of the decision algorithm, the best server and path for content delivery are selected. The following step is configuration of edge CAFE located close to the content server to intercept packets streamed by the server and encapsulate them into content packet. For this process, the CAFE configuration process is invoked, which directly contact the edge CAFE and configures packet filtering rules based on source and destination addresses, ports numbers and protocol type. This information is provided by content owner during publication process.

4.2 Content Forwarding Plane

The key element of Content Forwarding Plane is Content Aware Forwarding Entity (CAFE), which delivers content from the content server to the consumer through selected content delivery path. We proposed new content forwarding approach where CAFEs maintain only the neighbourhood (local) information, i.e., how to forward content packet to the peering CAFE. All information about content delivery path is stored in the header of content packet. This header includes information of the end-to-end path, which allows CAFEs to forward packets along the selected path. The content packet header is attached and removed by edge CAFEs located close to the content server and the client, respectively. The path selected during resolution process is configured in the edge CAFE. Our approach follows the source routing principle at the domain level, which allows for flexible selection of content delivery path for each content request. Furthermore, the proposed approach is technology-agnostic because it enables domains to use different packet forwarding technologies between peering CAFEs. In COMET approach, the end-to-end path is made up by a chain of content delivery path segments between neighbouring CAFEs.

Fig. 3 presents COMET header and its processing in a CAFE node. The COMET header composes of three main fields: (1) *Length* (1 byte), which determines the length of the list of forwarding key, (2) *Index* (1 byte) indicates the active forwarding key and (3) *the list of forwarding keys*. Note that the list of forwarding keys is correlated with the length of content delivery paths. For typical inter-domain paths, the list of forwarding key has length about 6 ÷ 10 bytes (assuming 2 bytes per domain).

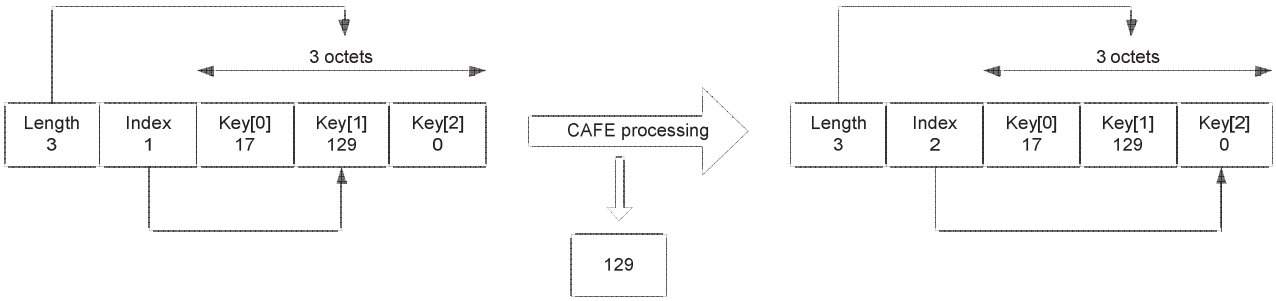

Fig. 3: COMET packet processing.

Each CAFE consumes 1 key from the list of forwarding keys included in the COMET header, i.e., {129}. It uses this key as index to lookup forwarding table, where forwarding behaviour is described, e.g., the output interface and a set of parameters required for encapsulation. Finally, the *Index* field must be increased by 1 to point out on the next key on the list.

The proposed approach allows for flexible selection of content delivery paths for each content request. Moreover, it allows applying specific packet processing within CAFE. In particular, each domain may define a set of locally specified CoS and use specific forwarding keys to instruct CAFE how assign packets into them. Although content forwarding is based on the stateless principle, the edge CAFE located at the server side has to classify packets belonging to given content flow and assign them appropriate values of the content header. This information is provided by the CAFE configuration component during the resolution process.

5. Federated testbed and validation tests

A federated testbed have been set up to carry out functional tests of the COMET approach. As presented on Fig. 4, the federated testbed comprises three individual testbed sites, located at the premises of three partners in three different countries (Cyprus, Poland and Spain) and connected by means of OpenVPN servers/clients.

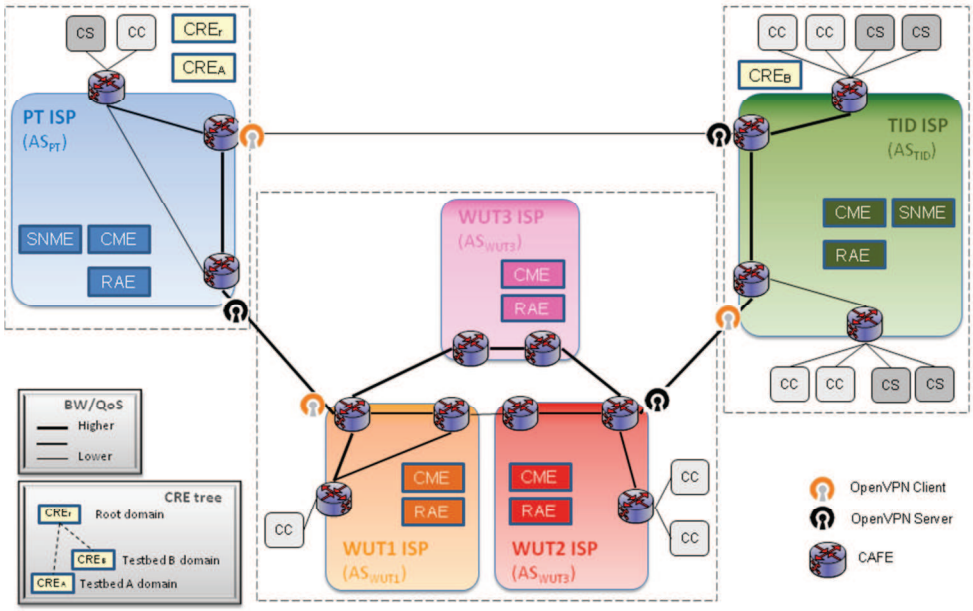

Fig. 4: Layout of the federated testbed.

The main features of the federated testbed are:

- Five ISPs are defined to emulate a testing ground as similar as possible to the current internet, including end ISPs and transition ones.
- Paths linking given ISPs are characterised in terms of QoS and bandwidth (BW).
- The RAE in each ISP gathers information about available paths to reach other ISPs, qualified with QoS and BW information.
- A number of CAFEs are in charge of packet forwarding in each ISP. Two different sorts of CAFEs are deployed, the edge CAFEs intercept packets for client and server and the border CAFEs carry inter-ISP traffic.
- CRE translates Content Names into Content Server characteristics. Two types of CREs are required: Authoritative CREs for storing Content Records in one or more domains, and the root CRE for identifying the Authoritative CREs.. A one-to-one mapping between ISP and Authoritative CREs is not required.
- A SNME in each ISP stores the status of the servers deployed there.
- CME finds an optimal server/path solution for retrieving the content identified by a Content Name and configures the selected path for download.
- CCs (Content Client) and CSs (Content Servers) emulate content consumers and producers, respectively. The CCs are assigned a preferred CoS, while the CSs are assigned a minimum CoS and BW/QoS parameters for the content being served.

Over this testing ground a variety of functional tests are performed:

- A number of CSs are deployed in different locations of the federated testbed to distribute a content identified by a unique Content Name. These CSs can be assigned different CoS and BW/QoS requirements, enabling that the CME mechanism will select optimal solutions (CSs/path) according to the CoS match, the CS loads, the location of the CSs with respect to

the CC in the federated testbed and the characteristics of the paths leading from the CSs to the CC.

- The composition of the CSs can be modified in the fly for a given Content Name, both in terms of amount, location and type (i.e. from Streaming Servers to VoD Servers) and the system will adapt automatically in order to provide an optimal solution (CS/path) for any new CC requesting the Content Name.
- Finally, a Content Name can be configured to create an offload hierarchy, enabling that if a set of CSs is overloaded (i.e, the high quality ones), the content can still be provided by redirecting the query to other set of servers (typically the low quality ones).

To ensure overall validation of the proposed architecture a number of performance tests have been also proposed for the federated testbed to complement functional testing and ensure that key operational aspects are met. Different characterisations have been firstly defined for monitoring purposes. These include defining the authoritative CRE performance in terms of retrieval time according to the number of CRs stored in its database for sporadic queries. Characterisation of retrieval times (minimum/ mean/ maximum and 95% percentile) vs. CRE occupation. The setup of query rates in accordance to the Poisson arrivals. Setting the SNME response time (minimum/ mean/ maximum and 95% percentile) vs. query rates according to load for each server.

A number of metrics and quantitative reference values have also been defined to ensure the achievement of operational expectations:

- The maximum number of CRs that can be stored in one single authoritative CRE should be of the order of billions (10^9) of CRs in the overall CRE hierarchy.
- The RAE should support up to 300 000 prefixes.
- The routing convergence time should be should be on the order of BGP-4 (expected few minutes)
- The edge CAFE should transfer packets with 1 Gbps throughput.
- The edge CAFE should be able to serve simultaneously at least 10,000 flows without packet losses.
- The configuration time configuration of edge CAFE should be in the order of a fraction of a second.
- The content resolution time (CRT) (95% percentile) should not exceed 2.5 seconds with success ratio less than 99.9% for sporadic queries (queries are sent in intervals larger than expected tolerable values for CRT when CCs and CSs are in the same ISP.

6. Performance evaluation

Besides the aforementioned federated testbed prototype and associated functional and operational tests, we evaluated COMET performance and its improvements in terms of content delivery as compared to current CDNs. One of the main features of COMET approach is that the content delivery exploits knowledge from both the underlying network and the content server conditions to improve delivery of the content packets. We evaluate this via simulations based on real domain-level Internet topology [18], [19] with ~36,000 domains and ~103,000 links, while the server distribution and server characteristics are modelled according to the Akamai infrastructure [20]. The content features such as duration, download bandwidth and content arrival and popularity are modelled based on existing works such as [21] to resemble current content delivery systems. The objective is to provide a comparison study on finding the optimised content access and delivery in terms of different objective function(s).

We then compute the server and path selection based on the multi-criteria optimization problem described earlier with the following optimization objective(s):

1. Random server with shortest AS (autonomous system) level path: this reflects what is happening in the current Internet;

2. Closest server with shortest AS-level path: this reflects the locality strategy adopted by most CDNs nowadays;
3. Best server with shortest AS-level path: the best server refers to the least loaded server hosting the requested content;
4. Best server with best AS-level path (the COMET approach): the best path refers to the least loaded path which may not necessarily be the shortest path. It is worth mentioning that we only consider the load on the inter-domain links along the considered AS path which can be captured by local SNMEs.

As will be indicated later, this set of objectives reflects the trade-off between system efficiency and simplicity (e.g., with or without SNMF) and offers flexibility by applying alternative algorithms into the proposed mediation system in a plug-and-play fashion.

Fig. 5 presents the content hit success rate when varying the incoming content request intensity. The considered thresholds for the maximum server load and inter-domain link capacity for content traffic are 180 concurrent connections and 1 Gbps respectively, and these values are in the range of operation of commercial content servers and consider the bandwidth by which large operators connect users [22] These thresholds are specially determined for observing both the effects: server overload and link overload in the same range of request arrival rates. The result confirms the intuition that the system gains in efficiency when the content resolution and delivery processes are optimized based on more information about server and network conditions obtained through monitoring functions. On the other hand, although the scheme of best server with best paths achieves the best overall performance, it is noted that such an approach has higher monitoring complexity on both server and delivery path conditions. The next best approach is best server with shortest AS path, and it can be easily inferred that no measurements on delivery paths (specifically on inter-domain links) are needed, but instead the selection is determined by the AS-path length, directly obtained from the underlying BGP routing information.

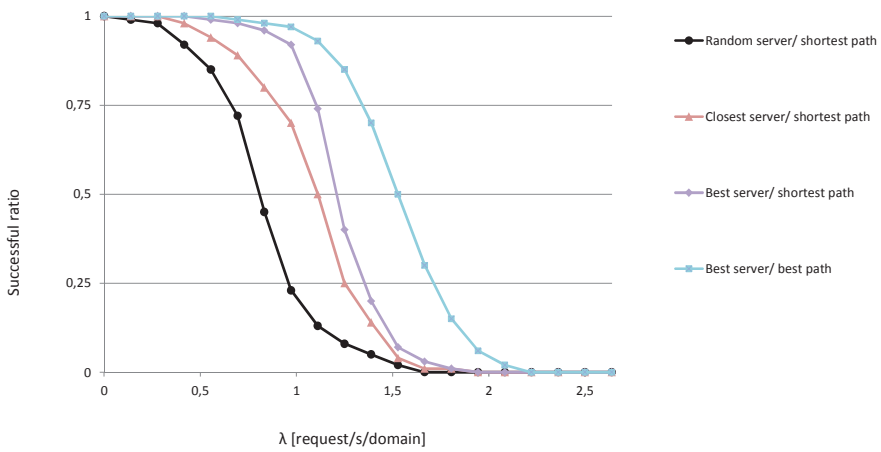

Fig. 5: Success ratio for server threshold 180 concurrent connections and link threshold, 1 Gbps.

In fact, more sophistication can be added in the CMF to further enhance the system's performance. Here we introduce an admission control mechanism on excessive incoming content requests. In the previous scenario, in case of insufficient server or path availability, blind acceptance of arbitrary numbers of content requests may disrupt ongoing content sessions due to congestion at the server or network side. To prevent such undesired incidents, a simple admission mechanism is implemented at the CME for actively blocking excessive requests according to the monitored server/network availability. In Fig. 6 we present the results where the system is equipped with the admission control functionality for avoiding service disruptions caused by uncontrolled requests acceptance. The performance gain in terms of successful ratio substantially increases, which demonstrates the advisability of implementing admission control functionality in the decision process algorithm.

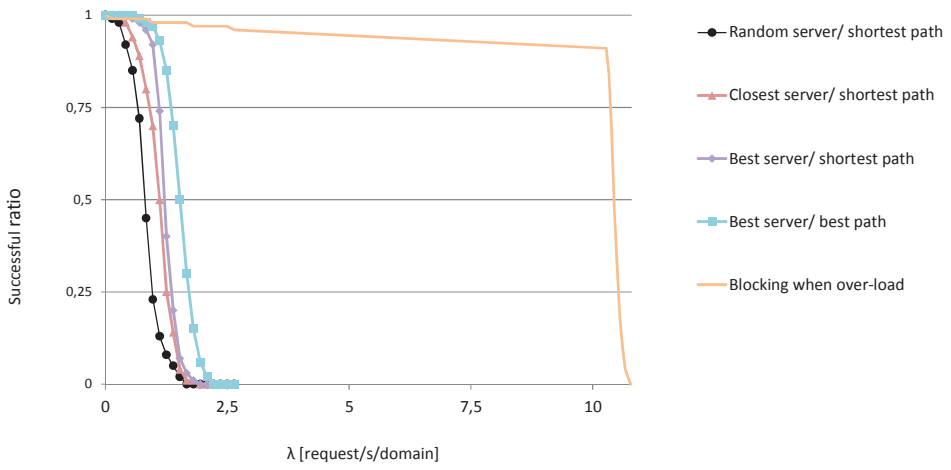

Fig. 6: Success ratio vs. Content requests intensity (with admission control at CME).

7. Summary

The paper presented novel content-oriented architecture, which was designed, developed and tested in the COMET project. The COMET architecture provides unified access to any type of content regardless of location and supports effective content delivery in a content-aware fashion. The proposed approach relies on the concept of mediation, which lies in the provision by Internet Service Providers of an intermediate plane between the world of contents and the world of network. The COMET architecture consists of two planes. The Content Mediation Plane (CMP) becomes aware of content location, server and network conditions, and available routing paths, and, based on that info, decides the best source and delivery path for each content request. On the other hand, the Content Forwarding Plane (CFP) is responsible for content delivery based on novel stateless content forwarding method which allows dynamic selection of content delivery path for each content request. In this paper we presented key mechanisms and algorithms that were designed for both planes. They cover: (1) new content naming and resolution scheme, which is based on the Handle System, (2) the routing awareness process, which enables multi-criteria (QoS) and multi-path routing, (3) the multicriteria decision algorithm, which optimises content delivery and resource utilization and (4) stateless content forwarding method, which exploits the source routing principle at the domain level.

The performance of proposed solution was evaluated by simulation experiments performed in the Internet-scale network. The obtained results confirm that COMET decision strategy outperforms currently used content delivery systems, i.e. CDNs, thanks to awareness processes, mediation and flexible content delivery method.

Our further works will focus on evaluation of the COMET prototype in the federated testbed. The tests will focus on verification of COMET functionalities, performance evaluation of developed COMET entities as well as the overall COMET system performance. Moreover, we will focus on evaluation of the COMET system scalability.

References

1. C. Labovitz, et al. ATLAS Internet Observatory 2009 Annual Report.
2. EC FIArch Group, Fundamental Limitations of Current Internet and the path to Future Internet, report ver. 0.9, 2010.
3. M. Pathan and R. Buyya, “A Taxonomy and Survey of Content Delivery Networks”, Technical Report, GRIDS-TR-2007-4, Grid Computing and Distributed Systems Laboratory, The University of Melbourne, Australia, February, 2007.

4. Eng Keong Lua, et al., “A survey and comparison of peer-to-peer overlay network schemes”, IEEE Communications Surveys & Tutorials, vol.7, no.2, pp. 72- 93, 2005.
5. V. Jacobson, et al., “Networking Named Content”, CoNEXT 2009, Italy, December 2009, pp. 1–12.
6. T. Koponen, et al., “A Data-oriented (and Beyond) Network Architecture”, Proc. ACM SIGCOMM ’07, Kyoto, Japan, August 2007, pp. 181–92.
7. M. D. Ambrosio, et al., “Second NetInf architecture description”, Deliverable 6.2 of project 4WARD, FP7-ICT-2007-1-216041-4WARD/D-6.2, January 2010.
8. P. Jokela et al., “LIPSIN: Line Speed Publish/Subscribe Inter-networking”, Proc. ACM SIGCOMM ‘09, Barcelona, Spain, August 2009.
9. I. Stoica et al., “Internet Indirection Infrastructure,” Proc. ACM SIGCOMM’02, Pittsburgh, USA, August 2002, pp. 73–86.
10. W. K. Chai, D. He, I. Psaras and G. Pavlou, “Cache ‘Less for More’ In Information-Centric Networks”, IFIP NETWORKING 2012, Prague, Czech Republic, May 2012.
11. I. Psaras, W. K. Chai, George Pavlou, “Probabilistic In-Network Caching for Information-Centric Networks”, ACM SIGCOMM Workshop on Information-Centric Networking, ICN’12, August, 2012, Helsinki, Finland.
12. W. K. Chai, et al., “Final specification of mechanisms, protocols and algorithms for the content mediation system”, COMET Deliverable D3.2, January 2012.
13. J. Seedorf and E. Burger, “Application-Layer Traffic Optimization (ALTO) Problem Statement”, IETF RFC 5693, October 2009.
14. The IETF Application Layer Traffic Optimization (ALTO) Working Group, Available http://datatracker.ietf.org/wg/alto/charter/.
15. H. Xie, Y. Yang, A. Krishnamurthy, Y. Liu and A. Silberschatz, “P4P: Provider Portal for Applications”, Proc. of ACM SIGCOMM 08, Seattle, WA, USA, August 2008.
16. Sam Sun, Larry Lannom, Brian Boesch, "Handle System Overview". Internet Engineering Task Force (IETF) Request for Comments (RFC), RFC 3650, November 2003. Available: http://www.ietf.org/rfc/rfc3650.txt.
17. A. Wierzbicki, “The use of reference objectives in multiobjective optimization”. Lecture Notes in Economics and Mathematical Systems, vol. 177. Springer-Verlag, pp. 468–484
18. David Meyer, “Route Views Archive Project”, University of Oregon, Available http://archive.routeviews.org/.
19. RIPE Network Coordination Centre, “Routing Information Service”, Available: http://www.ripe.net/data-tools/stats/ris/ris-peering-policy.
20. Akamai Technologies, Inc., “Facts & Figures”, http://www.akamai.com/html/about/facts_figures.html.
21. H. Yu, D. Zheng, B. Zhao, and W. Zheng, “Understanding user behavior in large-scale video-on-demand systems”. SIGOPS Oper. Syst. Rev. 40, 4 (Oct. 2006), pp. 333-344.
22. K. Florance, Netflix Content Delivery, “Netflix Performance on Top ISP Networks”, January 2011, http://techblog.netflix.com/2011/01/netflix-performance-on-top-isp-networks.html.