

Content, Connectivity, and Cloud: Ingredients for the Network of the Future

Bengt Ahlgren¹, Pedro A. Aranda², Prosper Chemouil³, Luis M. Correia⁴
Holger Karl⁵, Sara Oueslati³, Michael Söllner⁶, Annikki Welin⁷

¹Swedish Institute of Computer Science, Stockholm, Sweden

²Telefónica, Investigación y Desarrollo, Madrid, Spain

³Orange Labs, Issy-les-Moulineaux, France

⁴IST/IT-Technical University of Lisbon, Lisbon, Portugal

⁵University of Paderborn, Paderborn, Germany

⁶Bell Labs / Alcatel-Lucent, Stuttgart, Germany

⁷Ericsson Research, Stockholm, Sweden

Abstract

A new network architecture for the Internet needs ingredients from three approaches: information-centric networking, cloud computing integrated with networking, and open connectivity. Information-centric networking considers pieces of information as first-class entities of a networking architecture, rather than only indirectly identifying and manipulating them via a node hosting that information; this way, information becomes independent from the devices they are stored in, enabling efficient and application-independent information caching in the network. Cloud networking offers a combination and integration of cloud computing and virtual networking. It is a solution that distributes the benefits of cloud computing more deeply into the network, and provides a tighter integration of virtualisation features at computing and networking levels. To support these concepts, open connectivity services need to provide advanced transport and networking mechanisms, making use of network and path diversity (even leveraging direct optical paths) and encoding techniques, and dealing with ubiquitous mobility of user, content and information objects in a unified way.

INDEX TERMS

Future Networks, Cloud Networks, Virtual Networking, Information-Centric Networking, Open Connectivity.

INTRODUCTION

The Internet's architectural model has sustained continuous development for the past four decades and provided an excellent substrate for a wide range of applications. The amount of mobile data has been growing exponentially, and one can expect this tremendous growth to continue. Despite Internet's uncontested successes, some challenges for this model are becoming apparent, like adding applications more complex than simple client/server or peer-to-peer ones (e.g., multi-tier), or deploying information-centric ones distributed over different providers; moreover, the range of so-far successful business models seems limited. Also, coordinating and integrating more diverse technologies, networks, and edge devices is getting overly expensive, and security issues are becoming real barriers to deployment and use.

Information itself has become more and more important in all aspects of communication and networking. Most of the traffic in today's Internet is related to content distribution, which includes file sharing, collaboration applications, and media streaming, among others. The interaction patterns of emerging applications no longer involve simply exchanging data end-to-end. These new patterns are centred on pieces of information, being accessed in a variety of ways. Instead of accessing and manipulating information only via an indirection of servers hosting them, putting named information objects themselves at the centre of networking is appealing, from the viewpoint of information flow and storage. This information-centric usage of the Internet raises various architectural challenges, many of them not being handled effectively by the current network architecture, which makes *Information-Centric Networking* an important research field. In this new paradigm, storage for caching information is part of the basic network infrastructure, a network service being defined in terms of named information objects (e.g., web pages, photos, movies, or text documents), independently of where and how they are stored or transported. This approach is believed to enable an efficient and application-independent large-scale information distribution.

Another problem is related to network applications, which can fluctuate rapidly in popularity and in terms of the amount of user interaction. This makes provisioning of both server and storage, as well as of networks, a difficult problem. On the server and storage side, cloud computing has successfully addressed many of these challenges, using virtualisation as a core technique. However, it is still unclear how to provide suitable network support for such highly variable applications when they run not just over the tightly controlled, custom-tailored, network

of a cloud computing operator, but rather inside more complex and diverse operator networks. In such a network, it might be possible to provide the computational resources, but it is not obvious how to dynamically provide the necessary networking support/capacity or the complex networking topology required. Furthermore, security both in networks and in cloud computing is a key challenge to success. One needs an integration of network resource management with cloud computing, an integration of provisioning distributed cloud resources with the network services to connect such distributed resources reliably, at a required quality. This combination is called *Cloud Networking*.

Transport of information is another matter that needs to be addressed. In the current Internet, transport relies on connectionless forwarding of small data packets that is not able to exploit the additional (semantic) information that is available in the end or edge systems; additionally, it is incapable of making use of the context information that defines and controls related flows throughout different network aggregation layers, leveraging the capabilities of heterogeneous transmission technologies. For example, it is practically impossible to exploit the diversity existing over different communication technologies between two endpoints (e.g., random variations in channel quality or structural differences in channel properties, like different delay/data-rate trade-offs), switching between technologies as the flow's required data rate changes. Similarly, efficient multi-path/protocol/layer optimisation is still unfeasible. In order to efficiently use such high-speed future network technologies, it is critical to implement cross-layer coordination with new inter-domain transport, switching, and routing protocols. Furthermore, the current Internet is a flat, service-neutral infrastructure; this is reflected in today's rigid peering agreements, which limit the type of business models and service agreements that can be applied at inter-provider interfaces. In today's cellular networks, the introduction of new services is a cumbersome process due to the complexity of setting up the necessary roaming agreements, since different networks may have different releases and features (besides the billing problem). *Open Connectivity* offers an approach to address these problems.

The aspects addressed above can be put into a perspective of joint planes for a new architecture, Figure 1. Three approaches are addressed in the current paper: information-centric networking, cloud networking, and open connectivity. The next sections present these concepts and discuss them in detail.

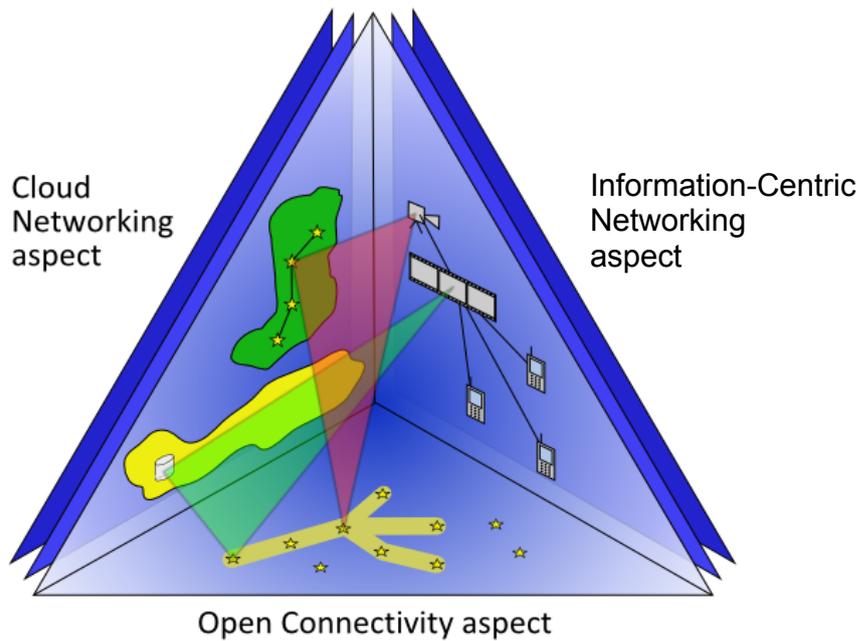


Figure 1 – Three aspects of a new network architecture.

A TARGETED SCENARIO

A scenario can help to put some of the previously mentioned aspects into perspective, and explicitly show an integrated approach to the problem. Obviously, a single example scenario cannot convey all aspects of a complex situation, but it is important to understand how the various approaches can fit together.

Consider a user, Alice, offering some piece of information (be it static or dynamic, e.g., streaming video), from her mobile handset to her content repository in the network. She shares this content, which becomes unusually popular, being viewed by many people, Alice's "Followers", most of whom use different network operators, thus, causing a large amount of relatively slow and quite expensive cross-operator, Figure 2 (a). This situation creates some incentive for the network operator to improve this content delivery situation (out of self-interest, but also to improve user-perceived quality).

A possible solution is the usage of the network-centric architecture, together with some open connectivity services, so that the increased load causes additional instances of Alice's content repository to be quickly spun up within some of these other operator's own networks, Figure 2 (b). The replication of the popular information to another location is facilitated by the information-centric caching mechanisms. If necessary, this infrastructure with the necessary

processing means, ensures that information is processed at and delivered from topologically advantageous places – unlike today’s cloud computing, where the processing can take place far away, with long round-trip delays. This allows for a reduction in the cross-operator traffic, since each of Alice’s “Followers” can now access information using only network operator-local traffic, and Alice’s video is only replicated once between her operator and each of the other operators.

However, this opens some transport problems, i.e.: by the time the additional nodes are operational, a substantial amount of video may already have been buffered at Alice’s (home network) node, which can cause problematic delays for new followers; existing followers that have been receiving cross-operator traffic will need to switch to their now-local instance of Alice’s node. These problems may be addressed by multi-path transport connectivity, which can handle the transport of the initial (previously buffered) video burst via higher bandwidth links for inter-operator traffic, before seamlessly falling back to the cheaper connectivity that is sufficient to keep up with Alice’s on-going video stream, Figure 2 (c). Hence, storing, processing, and transporting information turns into an integrated problem, while today only isolated solutions are available.

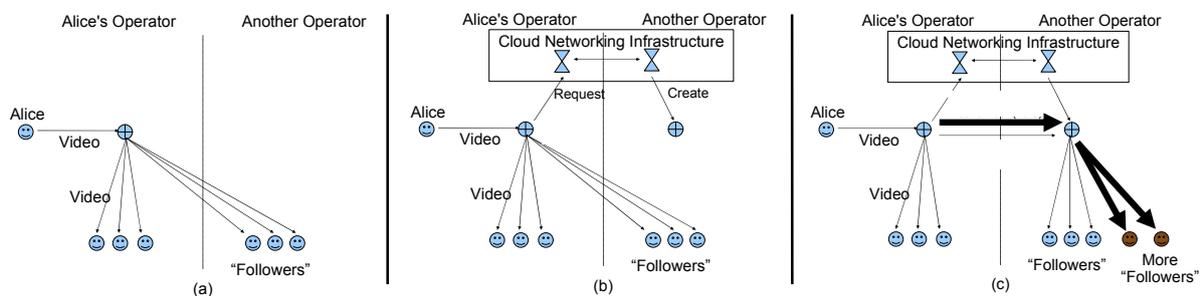


Figure 2 – Three steps in an advanced user-content-provisioning scenario.

INFORMATION-CENTRIC NETWORKING

The notion

The notion of Information-Centric Networking (ICN) has been proposed by several initiatives in the last few years. The core idea of most proposals is to consider pieces of information as main entities of a networking architecture, rather than only indirectly identifying

and manipulating them via a node hosting that information. Thus, information becomes independent from the devices they are stored in, enabling efficient and application-independent information caching in the network. This approach is believed to result in a network that is better adapted to information distribution and retrieval, which are the prevailing uses of current network technologies.

Notable examples are the work on Content-Centric Networking (CCN) [1], on publish/subscribe schemes [2], on directly embedding publish/subscribe schemes into the network fabric (PSIRP project) [3], the NetInf work by the 4WARD project [4], upon which our own on-going work is mostly based, or earlier the DONA project [5]. Similar ideas have also been considered in the context of wireless sensor networks (e.g., the idea to use predicate-based “interests” to identify which data shall be transported, with protocols like directed diffusion [6] realising that idea).

An example architecture: NetInf

Let us consider one of the approaches in more detail. The NetInf ICN architecture developed in the 4WARD project comprises three major components: a naming scheme for Information Objects (IOs), a name resolution and routing system, and in-network storage for caching. These components are illustrated at a high level in Figure 3, and described in the following paragraphs.

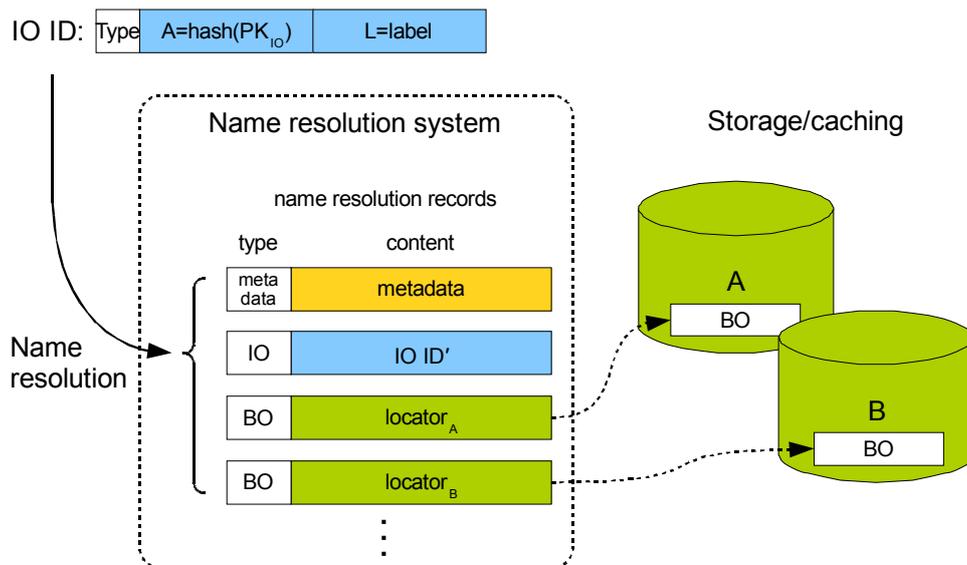


Figure 3 – Major components of the 4WARD NetInf Architecture.

The naming scheme is important for making information objects independent from the devices storing them. The hard part is not to make the names (information object identifier - IO ID in the figure) location-independent, but rather to fulfil the security requirements that result from the location independence. One cannot depend on host-based authentication of a delivering server, since one wants any node in the network, dedicated caches as well as end hosts, holding a copy to be able to share that with others. In order to be able to trust a copy of an information object coming from an untrusted device, the receiver must be able to independently verify the integrity of the object so that it becomes impossible to make forgeries. Therefore, the naming scheme has a cryptographic binding between the name itself (using field A of the ID) and the object, similar to DONA. Furthermore, the naming scheme supports dynamic objects, owner authentication and identification, changing owner of an object, and anonymous owners.

The purpose of the name resolution and routing system is to link the object names (IO IDs) to the actual information objects, so they can be queried and retrieved. It is a major challenge for all information-centric approaches to design this resolution system so that it scales to the global level. The system can be viewed as a variant of the Internet's Domain Name System (DNS). The system can also be viewed as a variant of IP (Internet Protocol) routing, where query packets are routed towards the location of the resolution records, or all the way to an actual copy of the information object. NetInf supports both of these models by allowing different name resolution protocols in different parts of the network. Protocols based on Distributed Hash Tables (DHTs) have been investigated as one suitable technology; multicast-based protocols have been investigated for implementing resolution in a local scope.

Storage for caching the actual bits of the information objects – the bit-level objects (BOs) in Figure 3 – is an integral part of the network service. Potentially, BOs are cached at all routers and end systems in the network. The goal is to deliver the requested object from the best cache(s) holding a copy to the client. A cached copy can be found either through the name resolution system or by a cache-aware transport protocol.

The NetInf Application Programming Interface (API) is inspired by publish/subscribe. A producer of information can *publish* an information object, creating a binding in the name resolution system, and *revoke* a publication, removing the binding. A consumer can *resolve* an information object name, returning the corresponding binding(s), and *retrieve* an object using the information in the binding(s).

Comparison of approaches

Other approaches make different design choices. Table 1 summarises the main conceptual differences for four exemplarily chosen, popular ICN variants, previously mentioned. The main aspects of difference are: (1) the choice of what to consider a piece of information with the corresponding naming model; (2) whether and how names are resolved into routable addresses of a simpler system (like IP) or whether name-based routing is used; (3) how transport and caching are integrated.

Table 1 – Comparison of different concepts.

Design aspect	CCN	NetInf	PSIRP	DONA
Naming and security of information objects	hierarchical, need to trust signing key to establish integrity	flat, self-certifying; support for versioning, and transfer of ownership	flat, self-certifying; notion of scope	flat, self-certifying
Name resolution and routing	name-based routing using longest prefix of hierarchical names	allows both resolution using, e.g., DHTs, and name-based routing	name resolution using a rendezvous function, within a specified scope	REGISTER and FIND primitives; hierarchical resolution handlers
Transport and caching	transport using name-based routing; finds cached objects through local search as well as on the path to the publisher	allows multiple transport protocols; finds cached objects through name resolution as well as cache-aware transport	transport routing and forwarding, using separate forwarding identifiers	caching in resolution handlers

Challenges in ICN

The ICN approach is still young, with many remaining research challenges, some of the most important ones being outlined in what follows.

Global scalability: An ICN needs to handle on the order of 10^{15} unique information objects at the global scale. Some solutions have been proposed, e.g., using DHTs, and calculations have been made suggesting that it is feasible to construct a global name resolution/routing system meeting this requirement. It still remains to be proven by experiments using real implementations.

Cache management: Resource management needs to go beyond considering link capacity, and has to address, in particular, cache storage. Some control of caching is needed to deliver a predictable service. Different cache replacement algorithms might be needed for different applications and usage patterns. Cache management protocols are needed for, e.g., collaboration between caches. Performance models are needed, accounting for distributed caching, statistical features of queried pieces of information (popularity, content size, usage patterns, correlations between objects), and the interplay between caching and data rate, notably for dimensioning.

Congestion control: ICN depart from today's Internet in two ways: they are receiver-oriented and they change the end-to-end principle. While the former implies that end-users may control the rate of information delivery, the latter creates an opportunity for implementing congestion control protocols between relevant nodes inside the network, through (chunk) query message pacing. This pacing mechanism may, e.g., be enforced between border routers of two different network providers in a consistent manner with respect to the charging model in use for information transport.

Deployment issues: To deploy information-centric schemes, there must be both incentives for users and operators, as well as the technical feasibility to introduce it. For operators, the appeal might lie in new business models (act as information host, cache provider) and in operational advantages (reduce inter-operator traffic, since information has to be exchanged only once between two operators). Incremental deployment is also a sine qua non condition; it is facilitated by schemes that can use existing routing and forwarding infrastructures (e.g., like NetInf can use different name resolution systems as plug-in and directly run on top of IP, as well as on lower layers).

CLLOUD NETWORKING

Clouds are restrictive

Provisioning of data processing and storage in *clouds* sitting at the edge of a network has proven to be extremely useful for a wide range of conventional applications; it is also a model that is well in accordance with today's network architecture. But when one considers either more demanding applications (e.g., with stringent latency requirements) or an advanced networking architecture, there are reasons to rethink the current cloud model.

Consider ICN as a case study: ICN requires storage and computing facilities distributed

into the network at a very fine granularity level, in particular, if it has to go beyond pure content distribution services and embrace active information objects. Leveraging current cloud computing solutions, based on server farms sitting at the edge of the network, provides an insufficient level of flexibility and performance, in particular, latency, to the end user. Serving ICN requests from the edge of the network will not result in acceptable performance, hence, while ICN will require cloud-like functionality, the notion of cloud computing has to be reconsidered. This implies the need to embed computation and storage deeply into the network to provide the required quality of experience. A cloud system serving an ICN architecture has to create ICN instances at various places in the network (and not just outside), and it has to provide these instances with a suitable and secure, possibly private, network. Hence, one needs to integrate cloud and (possibly virtual) networking services into cloud networking.

More generally, in a traditional cloud, massive amounts of data will be “sitting in the cloud”, waiting to be accessed by users anywhere and anytime. “Sitting in the cloud” also implies the need for a higher level of flexibility in the network: on the one hand, applications will reside in it, will be massively distributed (even over several cloud centres), and will be accessible to a massive number of users; on the other, the network itself will be composed of a vast range of different network infrastructures, which will be undoubtedly managed by different operators. These requirements are not new, and the TeleManagement Forum had been addressing them for some time in initiatives like IPSphere.

But today’s cloud solutions are based on concepts inherited from Grid Computing, and as such, they do foresee massive deployment of computing resources located at the edge of the network in general. Advanced solutions for distributed services were inspired by the grid (like Eucalyptus), but will not serve our purpose either. They implement the Infrastructure as a Service (IaaS) paradigm, being massively based on pushing computing and content to virtual machines localised at a few locations at the edge of the network, i.e., at large data centres. None of these approaches is suitable to act as an execution platform for ICN, where both storage and computing will be distributed, yet might still heavily interact with each other.

For a pervasive deployment of the kind of infrastructure one is aiming at, there is the need to provide the network with mechanisms to access and to *create* such computing resources at any place in the network they might be deployed at. More importantly, a tighter integration with virtualisation at all possible levels is necessary: applications in the cloud run in parallel, sharing the infrastructure, and need to be isolated from one another to provide predictable security and

performance guarantees at all levels, including the network plane. Additionally, current approaches to network virtualisation are too static: Virtual Private Networks (VPNs) at layers 2 and 3 are conceived as semi-static entities, which require often manual intervention when connections to end user locations are created or destroyed. Signalling, a possibility to make the operation of VPNs more dynamic, is currently based on protocols like BGP-4, which have been designed to minimise the oscillation probability in the infrastructure, therefore, not being too dynamic.

Benefits of an integrated approach

One needs a solution that distributes the cloud (and its main benefits, on-demand availability of computing, and storage with massive benefits of scale) more deeply into the network, and disperses the cloud closer to the end user to reduce latency: one might talk about *mist computing* instead of cloud computing, Figure 4. Moreover, these “misty” resources need to be flexibly networked across a backbone network, with isolation and security in place, the allocation of storage, computation, and networking connectivity between them becoming an integrated problem – applications can only be mapped onto a part of a cloud, when the required networking resources are in place, both to other parts of a cloud and to the end-user population the cloud part is intended to serve. The approach proposed here intrinsically takes the finer level of granularity needed in order to implement an infrastructure into account, which is highly responsive and provides a tighter integration of virtualisation features at computing and networking levels, possibly trading off computing and networking against each other (e.g., use slower computing nearby vs. fast computing far away).

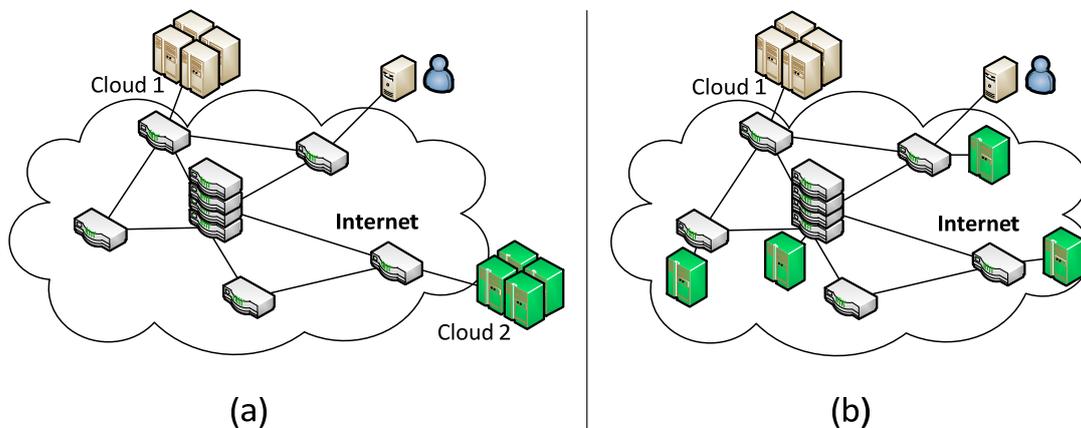


Figure 4 – From cloud (a) to mist (b) computing, supported by cloud networking: resources of Cloud 2 (shown in green) are spread much finer and deeper into the network, close to the actual point of usage.

The levels of envisioned adaptability provide better adaptation to network conditions and higher robustness to flash crowd effects. The network will adapt the amount of resources to the traffic needs, and move computational resources nearer to the physical locations where they are needed, creating the required network connectivity on demand. Additionally, it will support a rich ecosystem of middleware, which can be run concurrently and isolated in different virtual infrastructures; cloud networking is not meant as an exclusive runtime environment for ICN alone, but as a generic service accessible for many different kinds of applications that need to run in the network with similar levels of adaptability and scale. For example, the Software as a Service (SaaS) paradigm should also benefit from cloud networking, moving the provided software away from today's centralised and remote data centres closer to the customer. Moving closer to the customer is in the interest of resource efficiency too: usage dictates the network portions that are activated for a specific service.

Challenges

Some lines of current research in cloud networking focus on optimising networking inside big cloud data centres [7], since measurements in such environments show that most of the traffic stays within the data centre [8]. The envisioned architecture does not geographically confine traffic in this way. The impact of traffic patterns associated with cloud networking applications [8] needs to be studied in highly distributed scenarios as considered here.

Another challenge that arises in massively distributed environments is failure protection: an application at a given network location might not work as expected. This situation needs to be detected and corrected. Approaches like the one presented in [9] need to be explored.

OPEN CONNECTIVITY

Challenges of Internet transport and connectivity architectures

So far, the current Internet transport paradigm focused to a large extent on the provisioning of a transparent TCP/IP based point-to-point connectivity between addressable hosts irrespective of the underlying transport technologies. However, there is a tremendous increase in capacity in the lower level network technologies (fibre, copper, and wireless technologies), but the usable network capacity increasingly lags behind the demands of emerging resource-hungry networked

applications (created by, e.g., content distribution, cloud computing, or social networking). In addition, the heterogeneity of deployed network technologies makes it hard to exploit the particular network resources and features on an end-to-end, or even edge-to-edge, basis for the sake of new evolutions, like ICN or cloud networking.

Therefore, there is also a need for an advanced open connectivity service framework that addresses the issues in the *transport mechanisms* of a Future Internet. It aims at leveraging advanced features (e.g., multipoint, multipath, dynamic switching, and extended bandwidth) of link technologies, especially of optical transport networks, making use of network (and path) diversity and advanced encoding techniques, and at dealing with ubiquitous mobility of user, content and information objects in a unified way. Access to these mechanisms should be provided through new open and extensible interfaces, between client (user) and network, as well as between networks. Figure 5 presents multi-layer transport architecture and interfaces for open connectivity services.

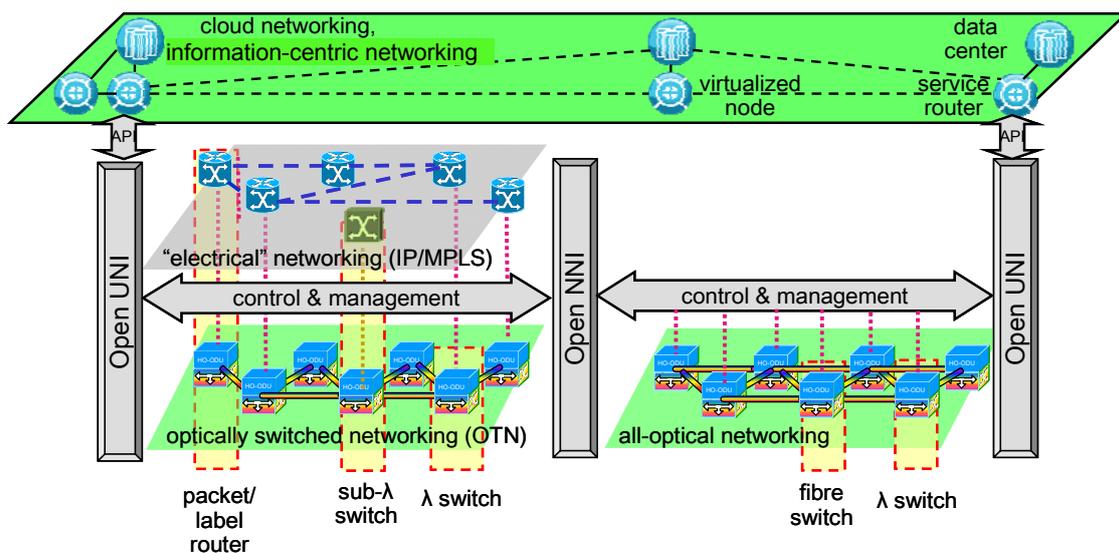


Figure 5 – Open Connectivity Services: multi-layer transport architecture and interfaces (UNI: User-to-Network interface, NNI: Network-to-Network interface).

While the physical networks offer an ever-growing optical bandwidth, and tend to aggregate links and switching/routing capabilities as much as possible for efficiency purposes in the core network, connectivity for the ICN approach will require high-performance distribution, referencing and managing a large number of inter-linked, but relatively small, information chunks located all over the world, preferably at the edge of the network. Today, this seems like

diverging interests, a challenge that needs to be addressed in an evolved transport system.

Therefore, the use cases for open connectivity will include the special needs for Wide Area Networks interconnectivity of new players, like distributed service centres and large enterprises (acting as information-centric nodes or cloud service providers), providing them with advanced and easy-to-use open APIs, to setup and efficiently control their private “virtual cloud networks” across multiple transport technologies and domains. Data centres in such a context will also comprise “mobile operation centres”, such as traditional IMS (IP Multimedia Subsystem) network functionalities running in a specific “mobility cloud”.

Leveraging lower layer transport efficiency

The current Internet is not capable of making use of context information that defines and controls related flows throughout different network aggregation layers, leveraging the capabilities of heterogeneous transmission technologies, including IP/MPLS (MultiProtocol Label Switching), WiFi, 3G/LTE (Third Generation / Long Term Evolution), Ethernet, and optical networks. For example, the TCP (Transmission Control Protocol) end-to-end transport with error protection, flow control, and congestion control is completely decoupled from routing and forwarding aspects of interconnected networks.

This architecture does also not allow leveraging advanced features of upcoming global network technologies, such as carrier-grade Ethernet or advanced optical switching techniques, e.g., concerning path management, resilience, or QoS (Quality of Service) mechanisms. For efficiently utilising such high-speed future network technologies, it is critical that there is cross-layer coordination with new inter-domain transport, switching, and routing protocols [10].

The evolution of transport networks is mainly driven by advances in optical transmission technologies, increasing the usable transmission bandwidth in optical fibres, as well as by the evolution of photonic integrated circuit technologies and the electrical processing in silicon, more and more used for switching of sub-lambda, single wavelengths and wavebands in a dynamical way [11]. The ability of using direct lightpaths and optical technologies for traffic off-loading the Internet core, and reducing (electrical) processing in intermediate hops, will have a beneficial impact on the energy budget of the global Internet overall, a problem being recognised only in recent years in the context of so-called “green ICT” [12]. However, this requires a new modified addressing and routing architecture, if a packet should be processed in less electronic steps, and then put into an optical path that ends up near the destination, in order not to run into routing

table explosion and scaling problems caused by the extensive use of multilayer techniques across multiple domains [13].

Related Work

The design of new transport architecture for the Future Internet has partly been addressed in previous projects. Most notably, the European project 4WARD (<http://www.4ward-project.eu>) has developed a functional architecture for generic paths that enables new forms of in-network processing. The 4WARD Generic Path concept encapsulates both interfaces and functionality central to data transport in an object-oriented fashion, thereby enabling access to routes, paths and transport functionalities, allowing to stack and connect them dynamically, e.g., to implement distributed mobility management. Such principles can now be extended to develop a lightweight, flow-aware concept of routing across and managing of rich communication paths (i.e., multi-point, multi-protocol, multi-path). This should allow for the development of open connectivity services that satisfy the needs of flash network slices (i.e., capable of being virtualised) and ICN (i.e., connecting content and data centres).

The European project Trilogy (<http://www.trilogy-project.org>) proposes a unified architecture for control, divided into a delivery service and transport services. The delivery service is composed of a reachability plane, responsible for the outgoing link selection, enabling network-wide reachability, and a resource plane, responsible for sharing the transmission resource between packets. The transport services provide functions for reliability, flow control and message framing.

Likewise, alternative control plane architectures arise out of the on-going Future Internet research activities, e.g., OpenFlow (<http://www.openflowswitch.org>) provides a new interface to Ethernet switches, which enables experiments with novel protocols at flow level on the one hand, and allows for programmable networking on the other.

Further related activities exist in other on-going European projects working towards the Future Internet, such as ETICS, GEYSERS and STRONGEST; generally, all these projects deal with variants of multi-layer network architectures and their multi-domain interconnection. To our knowledge, only the SAIL project (<http://www.sail-project.eu>) focuses on providing the means for user/application controlled access to establish connectivity tailored to the needs of future internet stakeholders, such as cloud networking or ICN.

Connectivity as an open service

The novel proposed open connectivity approach will extend the current point-to-point related connectivity towards multi-p* (i.e., multi-path/point/protocol) transport and routing, investigating the interactions between multi-p* transport, path-selection and routing, and having an end-to-end, cross-layer and cross-domain approach for multi-p* management.

The proposed solution is based on a multi-domain architecture, allowing for open and extensible communication and cooperation between the control planes of different network domains (user-to-network, and network-to-network) in a unified way. It enables the generic exchange of resource information for data flows across technology boundaries, in order to support content and information delivery in ICN, and provide appropriate dynamic and virtualised connectivity for cloud networking. That will also allow end-to-end optimisation concerning transport energy efficiency or intelligent sharing of network resources, caches and data processing.

As an application, one expects the rise of specific cloud services provided by distributed and interconnected data centres, e.g., mobile access to community and social networks running in a specific “mobility cloud”. Networking of such mobility centres might require new forms of mobility management that go beyond serving mobile users in the access networks, and include mobility of content and information within the interconnected distributed operation centres. A significant efficiency benefit can be expected by making use of path diversity, both in advanced optical and wireless network technologies. A promising alternative to state-of-the-art mobility management, with its single centralised anchor point, is a dynamic distributed mobility management [14]. In the cloud, the network dynamically chooses the optimal location of mobility service anchor points on a per-user/per-device or even per-flow basis. Open connectivity services will enable the cloud to make use and manage the multi-flow and multi-path routing capabilities provided edge-to-edge across the networks.

CONCLUSIONS

Internet has been based up to now on an architectural model that has coped with a sustained continuous development and provided a good environment for a wide range of applications. Nevertheless, challenges for this model became apparent, namely at the applications level, not only from the technical viewpoint but also from the business one. This paper addresses aspects

of a new architecture, from three approaches: information-centric networking, cloud networking, and open connectivity services.

Information-centric networking considers pieces of information as main entities of a networking architecture, rather than only indirectly identifying and manipulating them via a node hosting that information; this way, information becomes independent from the devices they are stored in, enabling efficient and application-independent information caching in the network. Major challenges include global scalability, cache management, congestion control, and deployment issues.

Cloud networking offers a combination and integration of cloud computing and virtual networking. It is a solution that distributes the benefits of cloud computing more deeply into the network, and provides a tighter integration of virtualisation features at computing and networking levels. Current challenges encompass the optimisation of networking inside cloud data centres, the study of the impact of traffic patterns associated with cloud networking applications in highly distributed scenarios, and failure protection in massively distributed environments.

Open connectivity services address transport mechanisms issues, aiming at leveraging advanced features of link technologies, namely in optical networks, making use of network (and path) diversity and advanced encoding techniques, and at dealing with ubiquitous mobility of user, content and information objects in a unified way. Challenges address, among others, the development of a lightweight, flow-aware concept of routing across and managing of multi-point/protocol/path communications, satisfying the needs of flash network slices, supporting content and information delivery in information-centric networks, and providing appropriate dynamic and virtualised connectivity for cloud networking.

ACKNOWLEDGMENTS

The authors would like to thank all their colleagues from the SAIL project team. Without their contributions, this paper and the insights behind it would not have happened. This work has been partially funded by the European Commission, under grant FP7-ICT-2009-5-257448-SAIL.

REFERENCES

- [1] V. Jacobson, D.K. Smetters, J.D. Thornton, M.F. Plass, N.H. Briggs and R.L. Braynard, "Networking named content", in *Proc. of CoNEXT'09 - 5th International Conference on Emerging Networking Experiments and Technologies*, Rome, Italy, Dec. 2009.
- [2] P.T. Eugster, P.A. Felber, R. Guerraoui and A.-M. Kermarrec, "The many faces of publish/subscribe", *ACM Computing Surveys*, Vol. 35, No. 2, June 2003, pp. 114-131.
- [3] A. Zahemszky, A. Csaszar, P. Nikander and C.E. Rothenberg, "Exploring the Pub/Sub Routing & Forwarding Space", in *Proc. of IEEE International Conference on Communications Workshop on the Networks of the Future*, Dresden, Germany, June 2009.
- [4] B. Tarnauca and S. Nechifor (eds.), *Netinf evaluation*, EC FP7-ICT-4WARD Project, Deliverable D-6.3, June 2010 (<http://www.4ward-project.eu>).
- [5] T. Koponen, M. Chawla, B.-G. Chun, A. Ermolinsky, K.H. Kim, S. Shenker and I. Stoica, "A Data-Oriented (and Beyond) Network Architecture", in *Proc. of ACM SIGCOMM'07*, Kyoto, Japan, Aug. 2007.
- [6] C. Intanagonwiwat, R. Govindan, D. Estrin, J. Heidemann and F. Silva, "Directed diffusion for wireless sensor networking", *IEEE/ACM Transactions on Networking*, Vol. 11, No. 1, Feb. 2003, pp. 2-16.
- [7] M. Al-Fares, A. Loukissas and A. Vahdat, "A Scalable, Commodity Data Center Network Architecture", in *Proc. of ACM SIGCOMM'08*, Seattle, WA, USA, Aug. 2008.
- [8] Arista Networks, *Switching architectures for cloud network designs* (http://www.aristanetworks.com/media/system/pdf/SwitchingArchitecture_wp.pdf), Apr. 2010, *Architecting low latency cloud networks* (<http://www.aristanetworks.com/media/system/pdf/CloudNetworkLatency.pdf>), May 2009, Menlo Park, CA, USA.
- [9] A. Carzaniga, A. Gorla and M. Pezze, "Healing web applications through automatic workarounds", *International Journal on Software Tools for Technology Transfer*, Vol. 10, No. 6, Oct. 2008, pp. 493-502.
- [10] K. Sato and H. Hasegawa, "Optical Networking Technologies That Will Create Future Bandwidth-Abundant Networks", *IEEE/OSA Journal of Optical Communications and Networking*, Vol. 1, No. 2, July 2009, pp. A81-A93.
- [11] D.T. Neilson, "Photonics for switching and routing", *IEEE Journal of Selected Topics in Quantum Electronics*, Vol. 12, No. 4, July-Aug. 2006, pp. 669-678.
- [12] J. Baliga, R. Ayre, K. Hinton and R.S. Tucker, "Green Cloud Computing: Balancing Energy in Processing, Storage and Transport", *Proceedings of the IEEE*, Vol. 99, No. 1, Jan. 2011, pp. 149-167.
- [13] G.J. Eilenberger, S. Bunse, L. Dembeck, U. Gebhard, F. Ilchmann, W. Lautenschlaeger and J. Milbrandt, "Energy-Efficient Transport for the Future Internet", *Bell Labs Technical Journal*, Vol. 15, Issue 2, Sep. 2010, pp. 147-167.
- [14] F. Bertin (ed.), *Description of generic path mechanism based on resource sharing and mobility management*, EC FP7-ICT-4WARD Project, Deliverable D-5.2.1, Dec. 2009 (<http://www.4ward-project.eu>).

BIOGRAPHIES

BENGT AHLGREN received his Ph.D. in computer systems in 1998 from Uppsala University, Sweden. He conducts research in the area of computer networking including the protocols and mechanisms of the Internet infrastructure. His main interest is the evolution of the Internet architecture, especially issues with naming and addressing on a global scale. Lately his research focus is on designing networks based on an information-centric paradigm.

PEDRO A. ARANDA obtained his Telecommunications Engineer title at the UPM's Telecommunications school in Madrid, Spain. He joined Telefónica I+D in 1991 and is currently a Technology Specialist, conducting research in the areas of the Future of the Internet and Service Agnostic Networks. His main research interests are the design of Internet grade architectures and the behaviour of BGP-4. Lately he has been working on the evolution of the Internet, especially all the issues related with inter-provider and inter-domain relationships.

PROSPER CHEMOUIL [F'03] received his PhD in control theory in 1978 from Nantes University. In 1980, he joined Orange Labs (then CNET), France Telecom's R&D Centre, where he is currently Director of a Research Program concerned with the design and management of Future Networks.

LUIS M. CORREIA [SM'03] received his Ph.D. in Electrical and Computer Engineering from IST-TUL in 1991, where he is currently a professor in telecommunications, with his work focused on wireless/mobile communications. He has been active in various projects within

European frameworks. He was part of the COST Domain Committee on ICT and has been involved in Net!Works activities.

HOLGER KARL received his PhD in 1999 from Humboldt University Berlin; afterwards he joined Technical University Berlin. Since 2004, he is Professor for Computer Networks at University Paderborn. He is also responsible for the Paderborn Centre for Parallel Computing and has been involved in various European and national research projects.

SARA OUESLATI received her Ph.D. degree in Computer Science & Networks from École Nationale Supérieure des Télécommunications, Paris, in 2000. She next joined France Telecom R&D as a research engineer in the field of performance evaluation and design of traffic controls for multiservice networks, and is leading the “Traffic and Resource Management” research team since 2005.

MICHAEL SÖLLNER is a technical manager at Alcatel-Lucent Bell Labs in Stuttgart, Germany. After he had received a PhD degree in applied mathematics (1984), he held various positions in communication industry where he focused on systems engineering and research for network architectures and protocols, now for mobile systems beyond 3G and the future internet. He has been involved in various European cross-industry research projects.

ANNIKKI WELIN is a Senior Researcher at Ericsson Research, department Packet Transport and routing. She joined Ericsson in 1998. Her research interests include packet transport and overlay networks. She has co-authored more than 20 papers and over 20 patents. She has been active in various projects within European frameworks.