UKTIN

Future Capability Paper

Core Technologies

TABLE OF CONTENTS

Acknowledgements

Chairperson: Neil McRae

Rafay Ansari
Saleem Bhatti
Jeff Land
David Neil
Joel Obstfeld
Andy Odgers
Andrews A. King
Dimitrios Pezaros
Louis “Sam” Samuel
Rick Taylor

We thank you for your time in reading our paper, as chair I would like to thank the UKTIN organisation, Digital Catapult and Bristol University for their assistance in this piece of work, and particularly the help of Nick Coombes from UKTIN, who has been a fantastic member of the team.

EXECUTIVE SUMMARY

Introduction
Core Network Technologies are defined explicitly as the critical networking technologies across the ISO layer model that enable networking. These are wide and varied, and the group made an early decision to separate the areas somewhat aligned with but not exactly representing the ISO model as a way of making the topics more manageable for the reader.

Core Network has many meanings; this paper is not about the 3GPP Mobile Enhanced Packet Core or the Core network of a transport/fixed network. Whilst many of the technologies referenced in the paper are part of those architectures, we want to make clear that his paper is much broader than just those two items.

Core network technologies create key enablers or capabilities in networking and are, for the most part, used in conjunction with other network technologies to form a use case. They are also wider than just this paper, wireless, optical and semi-conductors are key core networking technologies, these are covered in other papers.

Approach
In the paper, the EWG assume that security and network management are taken as implied. In every network, ensuring that risks are managed, and a security policy and best practices are in place is mandatory. We assume a full FCAPS architecture will enable successful network management.

The chair has been disciplined in ensuring that we stick to the focus of the UKTIN Future Capability request to highlight current activities specifically around commercial research and development of the scoped areas.

Core Technology standards are key enablers for network evolution and should be informed by working solutions. This represents a difference in approach between IETF and 3GPP and potentially between the Core Technology area and that of some of the technology areas of the other EWGs. We do not see any evidence that standards have affected the UK telecommunications industry's current situation.

The first section of the paper introduces the scope and the approach we have taken to “slice” this paper into a more manageable document and attempt to present guidelines around that scope. In the second section, we highlight the scope we have focused on and, importantly, what we have not focused on. While we have covered a wide range of topics, we have not drilled into those topics in detail; however, a detailed review of these technologies isn’t required to answer the future capability request. We have also specifically pulled out a section on a high-level overview of quantum networking. This exciting field is unique, given its infancy and wide range of directions and viewpoints.

It should be obvious that this is not an exhaustive review of all Core Network Technologies; there will no doubt be areas we have yet to cover, both intentionally and unintentionally, the latter mainly owing to a lack of knowledge or a view from the experts that this is not relevant. However, we strongly believe this doesn’t significantly impact our analysis of the UK’s future capability needs in core network technologies. In section 4, we illustrate how the core networking technologies come together to produce something of value, and we analyse the current core networking technologies system in the UK.

SUMMARY OF RECOMMENDATIONS

In terms of Core Network Technologies' commercial research and development, there is very little happening at significant scale in the UK today when compared to other countries such as Japan[1]. Whilst a lot is happening in academia, very little is making its way into the commercial sector.

To encourage development, we have made a series of recommendations:

- **Scale – Create more scale by incentivising UK telecoms suppliers in UK to grow and play on an international basis**
  - The UK Market needs to be more significant to drive enough scale to create the industrial research and development needed to create significant national leadership or sovereign capability. Core networking technologies are pieces of a larger ecosystem of standard solutions, and scale is vital.
  - Many successful smaller companies, such as Lumenisity, are acquired by large international organisations (in this case, Microsoft Inc.).
  - Smaller companies find it exceedingly difficult to export their products, making it more difficult to expand outside of the UK.

- **Stronger Co-ordination – Use regulation to maintain competition, create less fragmentation, reduce duplication (e.g. basic infrastructure duplication), and optimise investment.**
  - The UK market needs to be healthier as a basis of creating scale. The current regulatory policies are driving fragmentation of the market, which is reducing scale and, thus, the buying power of the UK. Over 140 companies are building fibre infrastructure in an un-coordinated way, of varying quality, many of which are unlikely to survive in the medium term. This further dilutes the ability of the UK to create scale. This will have a staggeringly negative impact on the UK market for decades and will not meet the government's objective of being a digital communications leader.
  - If we draw lessons from history, the reason for Marconi's failure was ultimately competitiveness; the UK operators needed lower-priced platforms, more responsive to customer needs and technology trends, and Marconi couldn't deliver them.
  - Academic research—There is a growing need for greater coordination across the UK academic community. While there are many great and globally recognised universities within the UK ecosystem, more needs to be done to ensure the work being done has a greater link to industrial research and development. There are pockets of success (Lumenisity, a spin out from Southampton University (for example) and we should use this as an exemplar to repeat across the academic community and enable them to take more commercial risk.

[1] Innovative Optical and Wireless Network

- **Better investment choices and more investment**
  - There is almost no industrial research in core network technologies in the UK.
  - The UK has run a series of competitions to create opportunities such as 5G testbeds and trial networks. This has not created enough value for the UK and we would advocate a move that re-directed this money to be used in a way that was part of a sustainable investment programme in telecommunications startups. (In addition, these trials had no significant impact on the deployment of 5G or fibre networks).
  - The UK should set an audacious goal for the nation's telecommunications infrastructure to enable innovation in services and potentially the creation of sovereign telecommunications capability, and importantly empower industry and academic research to work together on a valuable outcome - Quantum Networking has this potential. An example of this would be IOWN in Japan[2], a 100Bn Yen network vision to create the Internet of the future and the core network technologies to underpin it[3] - "the IOWN Global Forum targets approximately 100-fold improvements in power consumption, end-to-end latency, as well as transmission capacity levels compared to traditional networks". An area of focus for this work could be Quantum Networking where the UK is seen as a leader but is being overtaken by Japan and Switzerland.

- **UKTIN is a great first step in stronger coordination, we need to take this forward to build on the strong foundation created. The industry and academic coordination process that has been implemented so far should be formalised in "permanent" value in creating organisations and structured to further emphasise the high importance that the government has placed upon the telecommunications industry.**
  - As seen from this paper fragmentation is significant – creating more value in specific niche areas is likely to be more successful in the short to medium term.

- **Skills – Grow technology skills in the UK across and beyond the whole scope of Core Network Technologies**
  - The UK has some significant technology and computing successes, such as finance and video games industry, while there needs to be a stronger, capable, technology skills in telecommunications, specifically in infrastructure, computer science (including software engineering), semiconductors, and optical transmission technologies. The UK has some expertise in the integration of complex systems, and this is a strength to build upon, but it is still niche. This needs to be supported to foster, retain, and encourage skilled people for the UK communications sector, with collaboration between industrial and academic settings.
  - Core network technologies deployment requires the skills in the hardware integration as well as service creation through the integration of s/w applications and firmware. As an example, disaggregation in telecoms service providers needs a greater knowledge of cloud/compute infrastructure which many operators have struggled with, also the interaction of these different technologies has a wide scope and thus needs time to learn and build experience.

[2] Innovative Optical and Wireless Network
[3] New Medium-Term Management Strategy - New Value Creation & Sustainability 2027 Powered by IWON

- **Standards – Create solutions that create the standards.**
  - Many stakeholders across many standards bodies drive standards creation. However, at the heart of the standards process is the need to solve a problem either for network operators or to create a capability for network operators or vendors to monetise. Any single state, organisation, or individual can't control this collaborative environment, and any attempt to do so will likely drive standards in the opposite direction. Core Technology standards are key enablers for network evolution and should be informed by working solutions.

- **Are Core Network Technologies the right area for focus for R&D?**
  - Whilst the above recommendations apply to nearly all the EWG subjects, the question specific to core networking technologies is that, can the UK's relatively small presence in this area be increased? Even with the best execution of the above recommendations, it's incredibly hard to see how this could be turned into a major sustainable area of value beyond emerging niche areas e.g. Quantum networking. Thus, investment in this area would need to be compared to potentially better investment returns in other EWG subjects.

    - A better approach maybe to partner with other countries or organisations to develop until sufficient scale is possible.
    - Money is better spent on improving the exploitation of Core Technologies for specific objectives e.g. R&D developing better on customer-focussed systems integration or operationalisation tools and techniques and monetising this tooling.

Government policies must enable scale, stronger coordination, less duplication and waste, improved investment choices that create significant sustainable value, and reduced barriers enabling business outside of the UK if it wishes to create more research and development in the UK and, indeed, more sovereign capabilities in the core networking technologies area.

1/ INTRODUCTION

This paper summarised the initial discussion, findings, and recommendations from the UKTIN Core Network Technology Expert Working Group (EWG). In the sections that follow we discuss the:

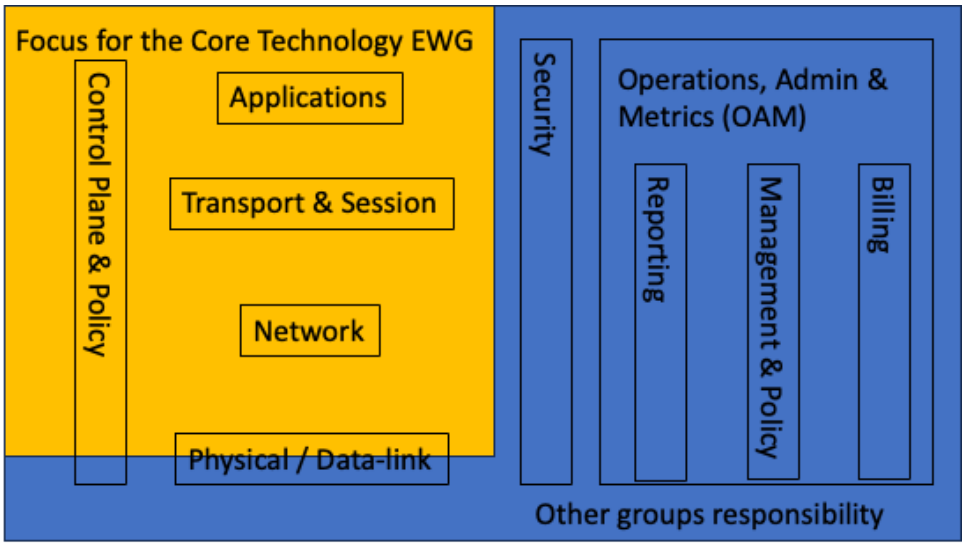
- Main technologies in scope
- Direction of evolution of those technologies
- Some examples of the assembly of those technologies into a network "system"
- Discussion of the Strengths, Weaknesses, Opportunities and Threats for the UK telecoms ecosystem for each of those technologies
- Initial Recommendations from the EWG including R&D investment focus

For the sake of clarity, the scope of the group is focussed on the key technologies forming a telecom network not the multitude of core network types and architecture themselves. Some examples of "systemisation" are provided to explore some of the significant business requirements driving the architectures.

1.1/ Scope Of Core Network Technologies

Core Network Technologies cover the wide range of IT and telecommunications technology elements assembled to create networks that ultimately form a telecoms network that provides services to users, namely consumers and private and public enterprises.

To simplify the coverage of this approach and make the scope manageable for the Expert Working Group (and reader) to handle, we have adopted the model, as shown below, which is adapted from the universally recognised OSI 7-layer reference model[4].



[4] CCITT standard X.200

Introduction

The network technology functions are broken down into layers, which may have a series of dependencies on the other layers through specific interfaces or APIs. The diagram also shows the areas in scope for the Core Technologies EWG and those which we considered to be within the scope of the other EWGs in the UKTIN project.

Core networks must transport large amounts of data efficiently and correctly (i.e. with required integrity between the right places at the required rate) on behalf of processes in the application layer (or 'apps'), fulfilling diverse purposes and being distributed as cooperating components across many locations. The number of such components, which can be physical sensors and actuators in a factory or a field, or large-scale computing processes, such as a 5G Core executing in multiple public, private and hybrid clouds, is growing rapidly. The amount of data that they generate is growing even more rapidly.

Delivery with required integrity at the right rate means that the core network components (clouds, edge computing systems, switches, routers, data-links and physical media) must be configured correctly. As such the Core Network Technologies EWG examines how physical media can be built into deployable networks making use of the data-links they provide, with off-the-shelf standard commercial, optical, copper or wireless solutions generally used. Policy must be applied consistently, and resources allocated across all devices involved for all users using the apps. This configuration requires control planes that can initiate sessions, create connectivity if necessary or discover a connectivity topology, and set parameters in the components. It also requires efficient transport layer protocols to carry the application data end-to-end between the applications' processes across a network layer that ensures delivery to the right place.

The UK has a significant R&D capability in the areas of applying the technology to future requirements e.g. systemisation of this technology. Any R&D project which builds a network will use technology components from all these layers, but few are carrying out fundamental research on the technologies themselves. Most of the fundamental lower-layer network R&D (physical layer and data-link layer) is done overseas, in large commercial companies e.g. hyperscalers and telecoms suppliers, and usually in-conjunction with international standardisation bodies such as IETF or ITU-T, partnering here with investment and resources could unlock this for the UK. To discuss the state of the art and future direction, In the following sections we have defined the key technologies in each of these layers in more detail.

1.2/ Application Layer

The application layer serves as a gateway to the network for end-users and is a focus for future research and development. Principle technologies in this layer include Hypertext Transfer Protocol Secure (HTTPS), as used for the interaction between functions in a mobile network, and protocols to manage the internet-wide IP address management such as Domain Name System (DNS) and Dynamic Host Configuration Protocol (DHCP), as examples along with many more protocols for data access.

The team has identified the focus for future research and development in this layer is around measures to increase network efficiency and scale enabling ubiquitous coverage or in quantum networking which is covered in outline in Section 2.1of this paper, however we must move at pace to ensure continued relevance.

1.3/ Transport & Session Layer

The transport layer protocols provide the interface between the basic, network layer packet transmission, and the communication services required for the communication endpoints within applications.

The purpose is the transmission control of information between end-points at the operating system's application layer – where the function is to control transmission, receipt and retransmission of information between the end-points in order to manage information flows, as an example, to minimise congestion.

Major technologies in this layer are the ubiquitous Transmission Control Protocol (TCP) used in most internet communications and network "boundary functions" such as end-user traffic shaping. Another protocol that is increasing in use is User Datagram Protocol (UDP), with its use in real-time audio and video.

Another technology in this layer, already standardised, is the development of a protocol named QUIC by IETF[5], which is also built on UDP, which better integrates functionality carried out by multiple protocols and functions today and provides encrypted flows in the process. Other enhancements are underway to support new applications such as Quantum Computing and AI.

1.4/ Physical / Data-link Layer

This is the physical medium over which information is passed between nodes.

This layer is typically under the control of a single administrative domain – There is the concept of addressing, covering both end-points and intermediary points. The physical/data-link layer provide reachability between end points (2 or more). An intermediary point may relay information towards an end-point or an intermediary point towards the end-point.

[5] IETF RFC9000 in 2021

The main technologies in this layer are Optical comms & components, Electro-magnetic communications (radio frequencies) and Electrical communications. Most of these technologies fall within the scope of the other UKTIN Expert Working Groups therefore we are including this here for completeness.

1.5/ Network Layer

The network layer is where the "systemisation" of the network components starts. There may be a potential technology/economic "control point" for the UK in the network technology ecosystem, where the network technology components start to come together e.g. enabling policy to be realised. The layers discussed so far form the pieces of the network technology jigsaw. Network and Control Plane are mechanism to bring these together to realise a user need. The Network Layer provides the glue between physical layer connectivity and the application traffic flows that assume and require seamless end-to-end transport and support for service differentiation. The network builds on top of physical and data-link infrastructure that is more tightly coupled with node-local hardware and scope and provides the logical abstraction of a network across different scopes (from local to Internet-wide) to network-wide and end-to-end services and applications. The main technologies included in this layer are routing protocols and functions such as OSPF or Software Defined Networks (SDNs).

As mentioned previously, standards are set by international bodies that have contribution from UK experts drawn from operators, suppliers, and academia, but there is little UK specific R&D on the fundamental technologies. The technologies are used to some extent, in every network that is built as part of the applied telecoms research where the UK has strength.

The major issues here are to evolve the network incrementally- – historically, big step changes have not been very successful. IPv6 for example, is still not fully adopted despite being released since 1998 as a solution for network address space limitation. Therefore, to increase the value of this technology to the UK a key question is - How do we create and sustain momentum for deployment of advances in networking technologies to happen faster? A strong economic driver is required, an example being legacy (copper-based) network replacement was very slow until electricity prices increased significantly, which in turn provided and incentive to speed up the deployment of optical networks which have lower operating costs.

Niche customer needs maybe also be a driver for example with the need for reliable network coverage everywhere to support mission critical customers. This could be architected as heterogeneous networks with connections back to the internet. There is some activity around alternative internet protocols as part of an ETSI initiative[6] but this is yet to deliver significant output.

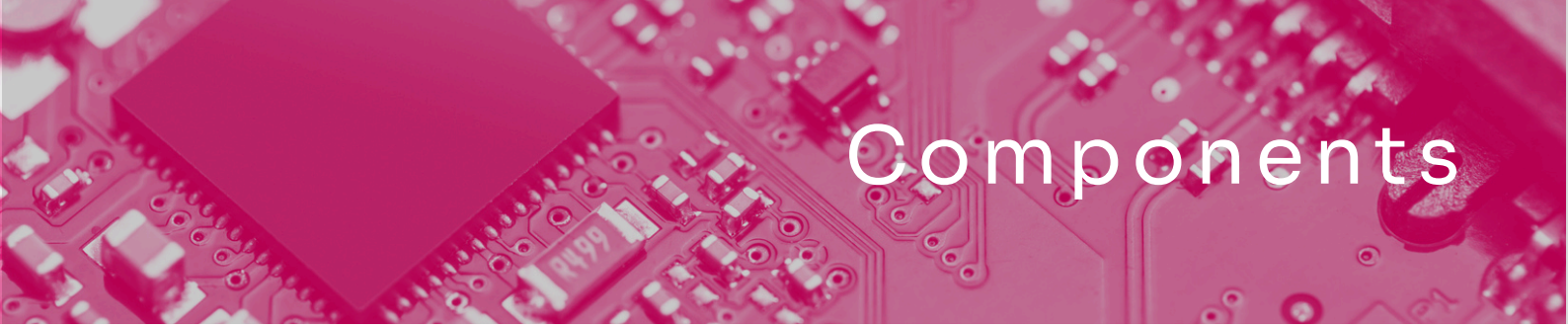
[6] ETSI Non-IP Networking

1.6/ Control Plane And Policy Layer

The Control Plane is concerned with the real-time management of traffic flows across the network and can interact with all layers of the networking stack to establish the topology and the dynamic configuration of the network used to carry user traffic within the constraints of the defined policies.

Main technologies in this layer are the many protocols and management functions used for this end-to-end traffic management. Examples being used for the setup of a voice call or a flow used for internet access or X-apps and R-Apps in ORAN deployments, if app affects networks in real-time and reacting to network information, SDN leveraging network capability to be setup in direct response to a workload requirement. Functions employ the dynamic configuration of routing and switching. To qualify for the layer, functions should never need human engagement.

Here again many of the protocols and architectures are defined by international standards bodies but the integration into a working network composed of various physical or logical functions, to support a specific set of customer services, is defined by the function of the systems integration function within an operator or R&D project team.



Components

2/ HOW TECHNOLOGY COMPONENTS COME TOGETHER

The Telecommunications industry is built on standards, with interoperability between components enabling the deployment of architectures to support customer outcomes.

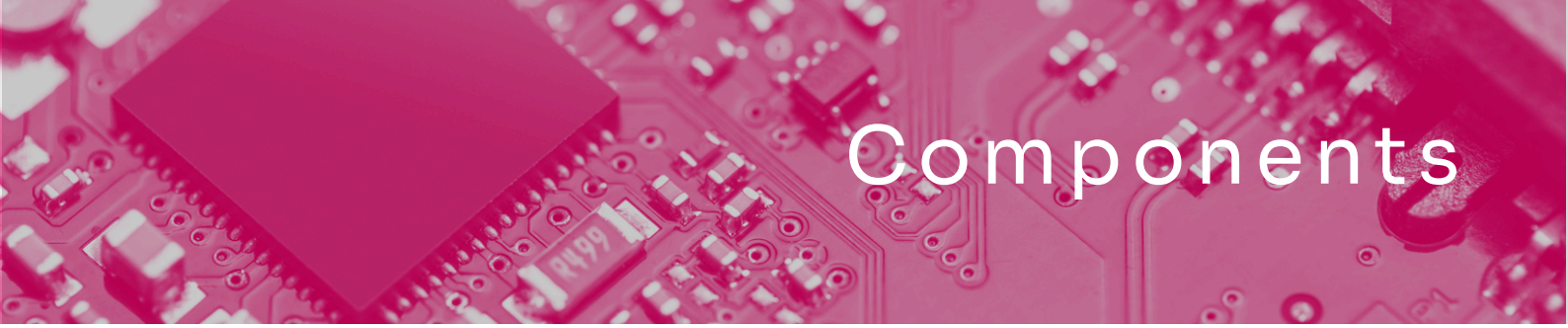
There are examples from recent history, of networks constructed from components provided by a small set of vendors. Some of these networks were deployed at national scale (e.g Minitel / France). The direction of travel has been away from vertically integrated systems and increasingly towards disaggregation. This can only be realised with cooperation and coordination between parties. (As a counter-example, international direct dialling telephone communications was premised on standards).

While proprietary methods and solutions can and do exist within and between components, the layered architecture offers a key design element - clearly defined interfaces between the layers and components in the technology 'stack', enabling independent implementations within a component and advancement. For example, there have been significant advancements in optical transmission systems which reside in the physical/data-link layer of the stack. That advancement has not resulted in the requirement to change applications or transport/sessions layer elements, and vice-versa. Similarly, adherence to standards enables interoperability between elements within the layer.

One of the most broadly deployed networking technology stacks is that of the Internet protocol architecture. This stack has evolved to encompass a broad set of deployment architectures enabling both public and private uses, using the same fundamental components. The scale adoption has also driven down the costs of manufacturing.

Network operators (public or private) typically provide the physical / data-link layers and the network layer. They will implement a control-plane and policy to drive their network. Network users connecting to the network, will use systems (such as device operating systems) that provide the transport / session layer. The user will also determine which applications will use the network, via the transport and network layers.

For example, fixed-access and wireless-access network operators will (typically) deploy infrastructure forming the core of their network. This includes components such as optical transmission systems, packet switches and routers. As the network extends towards the customer, the access technology components will vary – from the systems required for cellular communications including the radio antenna and the packet gateway, to the systems required for fibre broadband including optical line systems and Ethernet switches.



Components

In private environments, the same networking technology stack is used, with hardware or software components being selected that are appropriate for the particular deployment scenario. A company campus may use optical transmission systems between buildings, using packet switches and routers to connect to a central computer data centre. They may also use a private cellular communications system to offer mobile communications coverage around the campus.

Little of this is of concern to the network user or the applications that they wish to use. Successfully engineered data communications networks can be highly performant and resilient to the point the 'network' disappears from user's minds – it just works. Like water, like electricity.

Data communications networks are now (almost) ubiquitous and should be considered as part of the UK's critical national infrastructure, since if they fail, there can be significant economic damage (and potential risk to life).

The project to upgrade the UK Emergency services communications network from the current radio network capable of supporting voice and low-bandwidth data applications, will see the use of a 4G-based cellular data network, enabling applications carrying voice, video and data for first responders. Those applications are (likely) to be premised on the use of Internet-based protocols – with the necessary security and access control protections.

The use of standards on the interface between the network layer and the transport / session layer means that a (near) common experience can be provided to network users. Applications are (nearly) unaware of the underlying network. Data, regardless of the application, can be passed over the underlying network.

For the UK, opportunities exist to advance the state of the art within and across the layers of the network. Cross-layer will be more complex given the increased number of stakeholders but with tenacity and influence, this can be achieved.

The benefit offered by a network (of any kind) must be greater than the sum of its parts. As with the advent of the telephone, the 'network effect' is the key. In the case of the networking industry, the benefit comes from hardware and software implementing the protocols and driving economies of scale through interoperability.

Change to the status-quo typically requires two key traits – the change needs to offer an improvement (cheaper, better, faster) and that improvement must be at least an order of magnitude in order to overcome inertia. And change takes time, so a long-game approach is necessary. About 20 years ago, the UK government supported initial explorations in Quantum computing. 20 years later, the UK is a world leader in this emerging technology.

2.1/ Quantum Networking

At the time of writing, the quantum networking area is far less mature than the general area of quantum computing.

Quantum networking refers to the field of study and technology that involves the transfer of quantum information between nodes in a network. Unlike classical networking, which relies on classical bits to encode and transmit information, quantum networking utilizes quantum bits, or qubits, to encode and transmit information.

The key principles of quantum mechanics, such as superposition and entanglement, are harnessed in quantum networking to enable secure communication and enhance computational capabilities. Quantum networks typically involve quantum computers, quantum cryptography, and quantum communication protocols.

One of the most promising applications of quantum networking is quantum key distribution (QKD), which enables the secure exchange of cryptographic keys using the principles of quantum mechanics. QKD ensures that any attempt to intercept the keys would disturb the quantum state, thereby alerting the legitimate users to potential eavesdropping.

Quantum networks hold the potential to revolutionize various fields, including cryptography, communication, and computation, by providing unprecedented levels of security and efficiency. However, building practical quantum networks still faces significant technological challenges, such as maintaining the coherence of qubits over long distances and integrating quantum devices with existing classical communication infrastructure

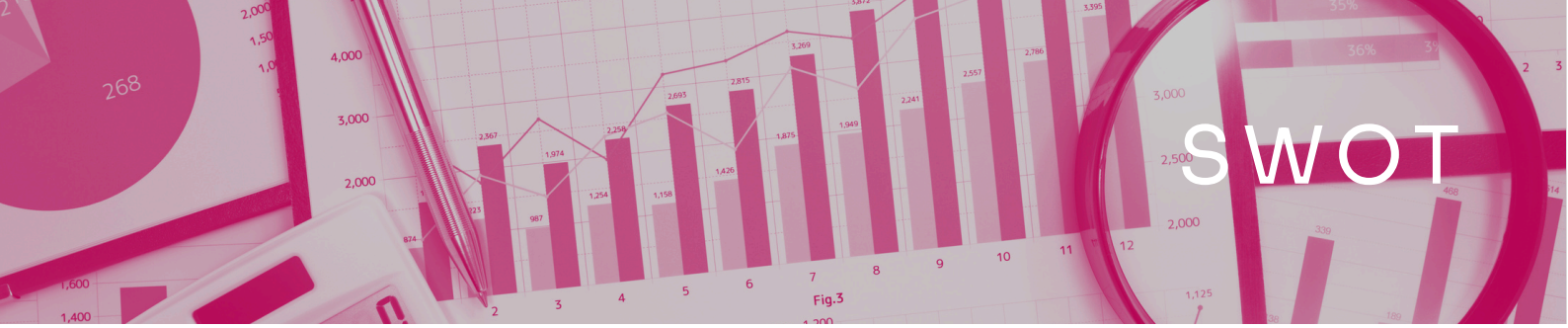
The IRTF Quantum Networking Research Group (QNRG https://www.irtf.org/qirg.html) has documented (non-standards track) architectural principles for a Quantum Internet (RFC9340(I)), the Abstract for which notes:

This is intended for general guidance and general interest. It is also intended to provide a foundation for discussion between physicists and network specialists.

There is significant global effort on Quantum Networking and given the UK's leadership in compute, physics and optical networking the UK could establish a leading position, with suitable investment and resourcing.

There is UK involvement in an EU project (Qurope[7]), but this is an area in which the UK needs to take greater interest and make more investment.

[7] Quantum Repeaters using On-demand Photonic Entanglement



3/ SWOT ANALYSIS

The EWG held a SWOT session considering the technologies previously identified as being in scope. We list the summary points below. These points were then used as a foundation for the discussion of challenges and recommendations made later in this paper.

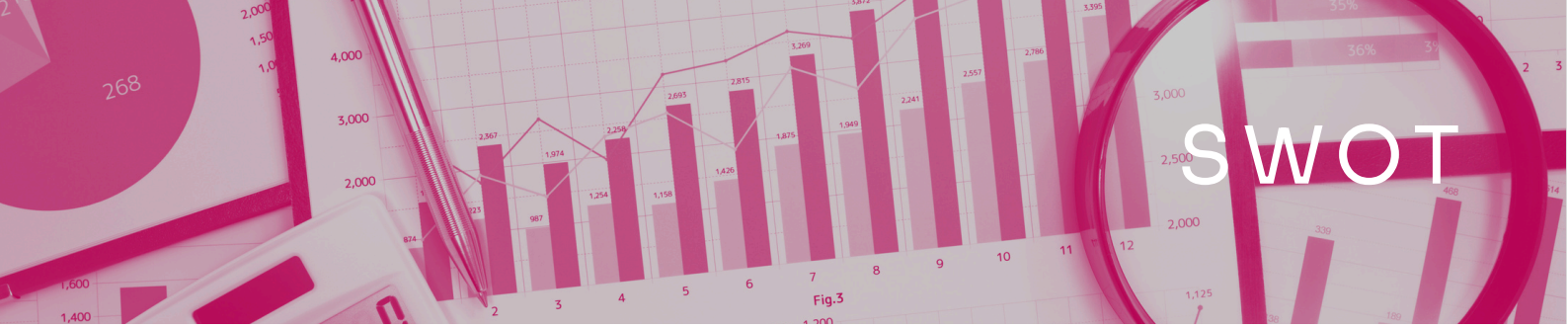
Strengths

- Strong academic contribution to telecoms standards and internet standardisation (Aberdeen, St Andrews, UCL, Glasgow etc).
- Efficient Open Spectrum Licensing process fostering competition.
- Wireless and optical companies well developed including NTN industry.

Weaknesses

No massive UK-based product companies (e.g. Nokia, Ericsson, Qualcomm, Huawei), not many start-ups in the core network technology scope.

- UK contribution to standards creation not matched by companies able to exploit that contribution in product development.
- Little business incentive to grow companies in UK rather than getting bought by larger international players.
- No organisations to assemble individual technology components into end-to-end system. Other countries e.g. India have strong systems integration companies.
- UK absent from activity in certain technology groups e.g. Overlay/underlay technology.
- Most UK presence is in optical and radio specific areas e.g. in various UK universities (e.g. Surrey, Bristol, UCL, Edinburgh).
- Companies in the US have easier access to funding, in the UK start-up and second stage funding is difficult.
- There is poor training in how to monetise good ideas into a business plan.
- The research community is fragmented – consortia are not strong compared to other countries e.g. Japan, South Korea.
- Uncertain outcomes of current academic telecommunications research
- The “Western” market for telecoms products (inc wireless) is notably smaller than the “Eastern” (China centric) market.
- UK market size is too small to give good ability to develop sovereign capability. A huge investment is needed & consequences of country only approach not positive in the medium/long term (e.g. Minitel France example).

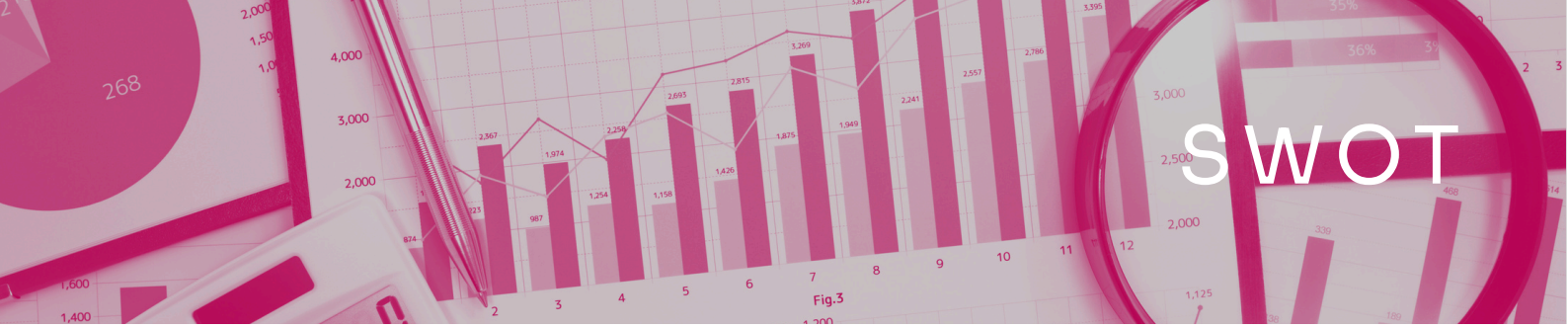


Opportunities

- Systems integration companies who have expertise in designing and assembling the multiple technologies into working E2E networks based on UK domestic and international operators.
- Private networking could be a disruptive market, not driven so strongly by international standards and subject to positive outcomes if the UK develops a successful technology.
- If the UK can control the “choke point” for core technologies to be commercially impactful. (Be ahead of the rest of the world to develop new technology niches)– potentially at the Network and Control Plane Layers. Be a “Big fish in a small pond” with selective technologies e.g. quantum networking
- Make the ecosystem more open – improve supplier diversity – but needs a big change in market dynamics and UK not well placed to contribute
- Quantum equivalent of OSI Layer 2, e.g. transport protocols or L2 Virtual Private Networks (VPN) though the IP for this may be owned overseas. This may be a longer-term opportunity, but the UK needs to be ready for when it goes mainstream.
- Solving the “Privacy vs Guaranteed QoS” dichotomy in network architectures.
- Develop niche networks, novel technologies catering to requirements of specific customers e.g. business critical networks.
- Exploit Network softwarisation to lower barriers to integration (and NaaS APIs), moving towards an “App ecosystem”.
- Need for a “Sovereign capability” in telecoms creates opportunities but also lack of focus of resources available.

Threats

- Increasing dominance of few extra-national players and hyperscalers. International suppliers are consolidating.
- Economic cutbacks on operator spend causing reduced investment in new technologies.
- Vendors continue to move development capability out of the UK.
- UK skills deficit. Future engineering talent focussed on AI aspects of Computer Science rather than network engineering. Lack of larger, stable companies for graduates to go to doesn’t help.
- Small UK companies find it easier to penetrate the network equipment market if sold to large international suppliers. Start-ups get acquired by large multinationals and IP is lost as a result. (Examples UK companies bought by Nokia/Ericsson and nearly every UK NM software company)



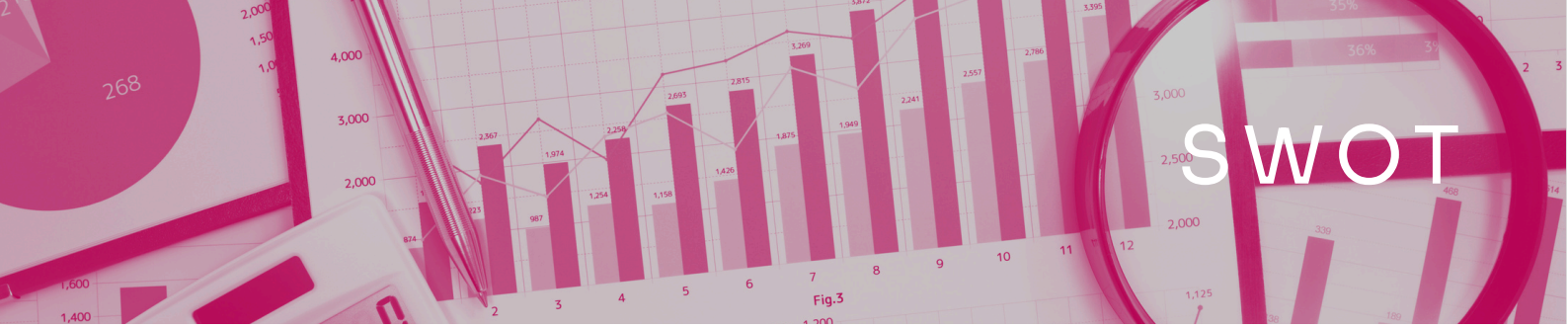
3.1/ Specific R&D Strengths & Prioritisation

While other EWGs focus on particular key topics with defined scope and objectives, even if they are transversal such as Security or AI, the topics covered within Core Network Technologies bring together many different issues. For example, this EWG is focusing on building the physical media into deployable networks and making use of the data-links they provide rather than their physics and engineering of the physical media. Off-the-shelf standard commercial, optical, copper or wireless solutions are generally used. R&D in these topics falls in the scope of the Wireless Networking Technologies, Optical and Semiconductor EWGs.

The definition and capability of a control plane can vary depending on the core network technologies and large-scale application platforms. In certain contexts, such as with 5G, global standardisation through bodies like 3GPP is essential for business continuity. However, in other scenarios, providers may opt for a proprietary control plane to distinguish their offerings in the market, or they may foster interoperability through industry alliances, agreeing upon protocols and their semantics. The O-RAN Alliance is a notable example of this collaborative approach.

The R&D landscape is however more dynamic at the network and transport layers, driven by the evolving demands placed on the Internet. As Internet technology is increasingly employed as a foundation for mass services, and critical infrastructure, it has become a synergy of the best of high-performance modern information, communications, and telecommunication technologies with the Internet's scalability at all levels, resilience and survivability. For example, where the original Internet was a flat collection of internetworked premises, campus, regional, national and transnational networks integrated by a global connectivity fabric, Core Network Technologies must now facilitate the creation and stable life of an unlimited number of Internets, distributed across the world as well as local or private, that may be isolated or interconnected, or overlaid or underlaid above/below one another as circumstances demand.

On this journey, original Internet design principles are being stretched, resulting sometimes in the loss of certain capabilities. Inefficiencies in some functions and protocols persist but are being addressed, as evidenced in the MP-TCP case study below. R&D and innovation in all aspects of Internet technologies and applications is global and extremely active, including towards standardisation in the IETF, both from academia and industry.

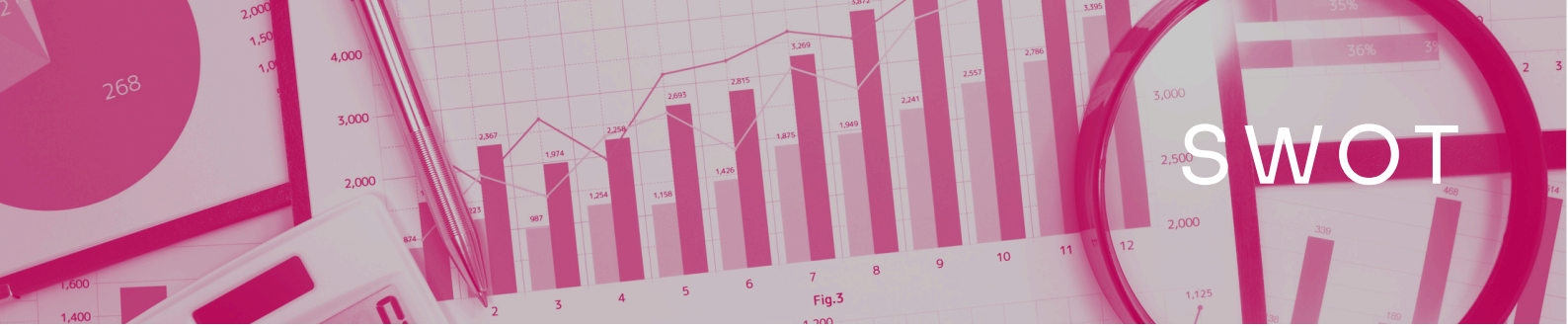


In response to these challenges, the UK is deploying a robust R&D effort. Alongside large-scale collaborative projects, such as the hubs described below, various universities including Aberdeen, St Andrews, UCL, Glasgow, Cambridge, QMUL, Imperial College London, and Sussex, among others, are actively engaged in research on telecoms networking, mostly focussed on "systemisation" of Core Network Technologies. Furthermore, many of these institutions play a role in contributing to telecoms standards and Internet standardisation efforts, ensuring the continued evolution and enhancement of core network technologies.

UKRI has awarded £6M to three consortia to establish long-lived communications hubs. The funded projects are:

- TITAN, a consortium of 17 universities leading research in critical segments of future communication networks and supported by four associate partners. The consortium aims to establish an open and productive platform for research collaboration and engagement across a large number of academic and industrial partners.
- The HASC hub brings together research teams from eight universities. They bring leading expertise in a wide range of wired and wireless technologies, in order to address the challenge of providing high-speed, low-latency access to internet services for future fixed and mobile users.
- CHEDDAR is a communications hub for empowering distributed cloud computing applications to drive research and networking across the UK academic community. The CHEDDAR hub aims to:
    - Inform the design of new communication surfaces that cater to emerging computing capabilities (neuromorphic, quantum, molecular), key infrastructures (energy grids and transport), and emerging end-user applications (swarm autonomy, air-service on demand).
    - Create integrated design of hierarchical connected human-machine systems that promote secure learning and knowledge distribution, resilient capabilities, sustainable operations, trust, and equality, diversity and inclusion-aware accessibility.

Yet, few UK companies are developing products in the core network domain, with limited activity in end-to-end integration or specialisation in specific technology groups like overlay/underlay technology. This is primarily due to uncertainty surrounding their return on investment, given the wide range of products that are already available on the market. Another reason is that the UK no longer has any large telecoms vendors to create demand for UK manufactured innovations and supply to UK operators with most Intellectual Property being held overseas. The limited industrial core networking research in the UK is bound to BT and the domestic presence of large off-shore multinationals, many a residue from historic acquisitions of UK innovator companies. However, system integrators, such as UK-based company Freshwave and AWTG, possess the potential to introduce significant innovations as most telecoms R&D projects involve system integration to a major degree.



We need to distinguish between "working on" from "working with" core network technologies. Examples of companies carrying out R&D “working-on” the in-scope areas are:

- Nokia (silicon in Cambridge)
- Luminosity (spin out from Southampton)
- BT and Vodafone
- Cambridge Quantum company and ORCA[8].

Key technical (and economic) challenges in the next 10 years which should prioritise R&D work include:

- The optimal distribution of infrastructure in hybrid, cloud-based network infrastructures
- Function distribution between CPE, edge and core
- Network instrumentation in increasingly complex, fragmented and encrypted networks e.g. with QUIC maybe also including "differential quality" approaches to performance management.
- Functional decomposition and exposure in “fine-grain” microservice-based network-as-a service architectures
- Approaches for sensor-fusion (+other 6G "wish list" stuff in network technology scope)
- The challenge of the “Gs” in mobile networks – where does fixed/mobile/NTN convergent networks play a part – a “road-mapping” challenge.
- New approaches to network quality/efficiency differentiation e.g. beyond L4S, TreeDN and MAUD
- Quantum impact on networks e.g. Quantum-safe, key distribution, quantum networking
- Approaches for Self-healing, Self-optimising and Self-monitoring networks (in conjunction with NM and AI)
- Systems integration and Systemisation tooling and practices focused on application & control planes in the model
- Architecture, application and optimisation of Quantum Key Distribution in telecoms networks
- Network reliability and availability; especially in the continued direction of generic compute and software functions and dis-aggregation.
- Symantec networking and Web 3.0

[8] OrcaComputing.com

4/ RECOMMENDATIONS

4.1/ Thought Leadership in the Evolution of Core Network Technologies

Whilst the recommendations set out below could apply to nearly all the EWG subjects, a challenge specific to core networking technologies is the ability to scale-up the size of current UK activity in this area. The global scale of 'the network' and advanced services running end-to-end mean that a small number of global suppliers have competitive advantage in a market traditionally hard to penetrate. However, developments such as the opening up of service interfaces and compute-communications technology convergence over the past decade create opportunities for new markets and players that need to be explored with a view to identify strategic investments that will create sustainable value for the future.
Potential "niche" areas for R&D focus and prioritisation were identified in the previous section, and these need further investigation and validation in the next phase of work.

4.2/ Scale

Recommendation 1 - The UK Market needs to be more significant to drive enough scale to create the industrial research and development needed to create significant national leadership or sovereign capability. Core networking technologies are pieces of a larger ecosystem of standard solutions, and scale is vital.

Many successful smaller companies, such as Lumenisity, are acquired by large international organisations (in this case, Microsoft Inc.).

Smaller companies find it exceedingly difficult to export their products, making it more difficult to expand outside of the UK.

Measures should be devised and implemented to:

1. Including identifying the profitable sector targets within the scope of the Core Technology model discussed earlier in the report and choosing the next growing technologies.
2. And encourage the sector to grow to scale or to work better together, nationally or with friendly international partners, on targeted profitable technical niches.

Rationale

The Internet has developed under the auspices of bodies such as the Internet Engineering Task Force (IETF). The UK, both through commercial and academic representatives, have been involved in such bodies for many years. Contributions to various standards and implementations are an indication of the intellectual capability present in the UK, something which should be fostered and encouraged in the future. Examples include RFC9443 (QUIC), RFC8684 (multipath TCP) and RFCs 6740-6748 (ILNP).

Recommendations

However, the UK needs sufficient commercial scale in Core Network Technologies to be a serious player in the area which can compete with global suppliers such as Ericsson, Nokia or Huawei across a wide or niche scope. A more significant industrial presence will, in turn, make the sector more attractive for graduates contributing to alleviating the skills supply issue.

The team has discussed two possible strategies to achieving commercial scale:

1. Aim to create a major global telecoms player based in UK – a "UK Ericsson or Huawei" equivalent or a re-invention of a supplier akin to the old GPT/Marconi company that was lost at the end of the last century (acquisition investment required is large); or
2. Aim to dominate in smaller, growing, and profitable technology sectors (layers in our model), where the UK's medium size could achieve sector dominance. Achieved by "gearing of public sector investment".

The Network Layer could be one of the foci for R&D investment in analogy of the "thin waist" concept, where equipment manufacturers have been able to establish control of the ecosystem. A commercial choke point where a relatively small number of companies control the market. The UK could help make more this layer more open for supply chain diversification, but this requires a big change in marker dynamics to make an impact. Example: UK could become global leader in quantum networking as there is no present global market or the example of ARM in semiconductors.

The UK should expand on the "Systems aspect" of moving from technology to operational networks. Many companies in the UK are working on combining ethernet, security other tech into working networks. The UK could be "successful" having little presence in underlying technologies but be successful being network integrators and integration developers, making them work on E2E basis. This could either be to supply larger companies or create start-ups that become acquired (CISCO example).

The team considers other countries that are good in this area to be Israel, Singapore, and the US. The success is "patchy" across Europe where the attitude to risk being the main differentiator.

4.3/ Stronger Co-ordination

Recommendation 2 - Better investment choices

The UK has run a series of competitions to create opportunities such as 5G Testbeds but evidence of delivering significant sustainable value or accelerating 5G deployment in the UK has been sparse. If the UK wishes to have sovereign telecommunications capability, how can the government set an audacious goal to create something that will bring value? An example would be IOWN in Japan[9], a network vision to create the Internet of the future and the core network technologies to underpin it.

Initiatives should be pursued to:

1. Identify Future applications driving the future network growth
2. Manage Intellectual property to maximise UK growth

Rationale

The group sees key issue here is - how do we get organisations that can monetise the IP created for UK? which has been an historic weakness. There are three parts to this recommendation:

- Identify Future applications driving the network
- Making Deployment of New Technologies Happen Quicker and Better Return on Investment
  - linked to an external catalyst e.g. electricity price increase driving power saving enhancements
  - Niche networks beginning to appear - heterogeneous networks needed to drive differentiated user functionality e.g. electricity supply - should UK be focussed on these? - "big fish in small pond" approach.
  - "thin end of the wedge" approach to UK success where is the "choke point"? - Network and Control Plane as areas for UK focus. Joining technologies together to give end-to-end functionality is where UK could get maximum return for smaller investment going forward.
  - How change market dynamics?
  - Core network technologies suffer from an information deficit. As noted previously, new applications are being developed that work on the assumption that the network will ‘just work’. Once a problem is encountered, the network needs to be adapted to support that new application. Similarly, the network is slow to change since changes may disrupt existing applications. The evidence for this is the number of ‘adaptation technologies’ that exist, enabling applications to continue to operate while masking the changes. L2TP, VxLan, MPLS L2VPNs etc.

- Managing Intellectual property to maximise UK growth
  - How to stop IP being created in universities being exploited abroad?
  - How to stop companies being acquired by overseas giants? Is this necessary or should we aim to produce lots of start-ups for acquisition?

[9] IWON Japan

Recommendations

Recommendation 3 - Supporting Core Network and Services Research

Support core computer networking and future Internet research through appropriate state funding programmes for the development and demonstration of novel end-to-end services on top of connectivity.

There is a growing need for greater coordination across the UK academic community. While there are many great and globally recognised universities within the ecosystem, too few of these have had sufficient investment to allow creating sustainable value for the UK through, for example, sustained commercialisation of their research, or sustained engagement with the IETF and other standards / industry bodies.. This needs greater alignment with the industry and a greater focus on valuable outcomes. There are pockets of success (Lumenisity, for example), but not nearly enough.

Rationale

This document has hopefully demonstrated that the complexity of core networking technologies and mechanisms that need to be in place to support the delivery of evolving end-to-end services on top of connectivity can be overwhelming.

The scope and diversity of interconnected networks as well as the scale of the global Internet, make development of complex services on top of connectivity a significant challenge that requires sustained investment for research and experimentation. At the same time, commoditisation of connectivity makes network operators focus more on the offering of new services to create value add and increase revenue.

The diversity of expertise required to innovate in this space is such that requires significant investment to support research that would enable the development of innovative solutions and technologies that could then be licensed, standardised and/or otherwise commercially exploitable.

However, in the UK, we have seen a real terms reduction of funding in relevant areas of the research portfolio over the past twenty years that needs to be reversed drastically to support innovation over telecommunications networks as a national priority that requires building sovereign capability. This includes providing a sustainable line of expertise in these core areas that is based in the UK.

Recommendation 4 - Accelerate Transition from Research to Market

Develop structured funded programmes to provide academic researchers with the expertise, knowledge and training needed to convert their research into technologies, products and services in this key sector of the global economy, and create pipeline to move ideas out of the university lab and into the commercial market.

The UK market needs to be healthier. The current regulatory policies are driving fragmentation of the market, which is reducing scale and, thus, the buying power of the UK. Over 140 alternative networks are building fibre infrastructure in a totally un-coordinated way, of varying quality, many of which are unlikely to survive in the medium term. This further dilutes the ability of the UK to create scale but crucially waste investment money. This will have a staggeringly negative impact on the UK market for decades and will not meet the government's objective of being a digital communications leader. Crucially, the reason for Marconi's failure was ultimately competitiveness; the UK operators needed lower-priced platforms, and Marconi couldn't deliver them.

There needs to be flexibility and appetite to restructure the UK market and willingness to explore transformative even if controversial options such as the UK seeking to take a position as the 'testbed' where commercial concepts are tried out and then copied (sold) to other parts of the world.

Rationale

UK research has made significant and sustained contributions to the evolution of the Internet and many of the associated networking technologies and mechanisms, from developing protocol standards for the delivery of multimedia content over packet-switched best-effort networks[10] to high-speed end-to-end transport over redundant paths to harness advances in underlying bandwidth[11]. This is also evidenced by the fraction of IETF RFCs authored by authors with a UK affiliation[12]. At the same time, transitioning from academic or even industrial research to product development has not been adequately supported through structured mechanisms that could identify and accelerate pathways to commercialisation in an area that, by nature, would require substantially more start-up investment than, for example, mobile phone or Cloud software products.

In the cybersecurity domain, there has been such investment for commercialisation acceleration currently in its 8th year, funded by the UK Department for Science, Innovation and Technology and delivered by Innovate UK that has supported 11-month programmes offering highly effective interventions (including bootcamps, workshops, and mentoring) and giving participating teams essential insights into the key milestones necessary to commercialise their research (including in IP, New Product Development, Innovation Planning, Sales, Investment, Pitching, Communications).

[10] Colin Perkins csperkins.org
[11] TCP Extensions for Multipath Operation with Multiple Addresses
[12] Characterising the IETF Through the Lens of RFC Deployment

Recommendations

4.4/ Skills

Recommendation 5

Greater technology skills are required across the UK. The UK has some significant successes, such as banking and video games, but there needs to be stronger, capable technology skills in telecommunications, specifically in infrastructure, computer science (including software engineering), semiconductors, and optical technologies. The UK has some expertise in the integration of complex systems, and this is a strength to build upon, but it is still niche.

The EWG considers that the UK has a severe problem with respect to developing the technology skills in the scope of this EWG. Finding people who can understand how to build things is becoming harder and harder. We need to figure out how we sell the “mission”.

Initiatives should be sought to improve supply of trained engineers by making telecoms network engineering more attractive as a long-term career path, rather than opting for other STEM career areas. Note that the complexity and diversity of the national and international technology and systems landscape is such that future trained engineers could come from a variety of engineering, mathematical, and science backgrounds, and not just traditional electrical engineering backgrounds.

Rationale

The skilled workforce in the UK in the area in scope is, in the main working for foreign-owned companies with a local presence e.g. Cisco, Ciena and Nokia This due to lack of scale of UK owned companies in this area.

4.5/ Standards– punching above our weight

Recommendation 6

Many stakeholders across many standards bodies drive standards creation. However, at the heart of the standards process is the need to solve a problem either for network operators or to create a capability for network operators or vendors to monetise. Any state, organisation, or individual cannot control this collaborative environment, and any attempt to do so will likely drive standards in the opposite direction.

The UK needs to identify and support engagement in key network standards where there is a standardisation need that would aid R&D across UK industry and academia but at the same time be a strong and collaborative partner in any approach to working in standards.

Rationale

It is the teams view that Standards that get implemented are those where there is customer/market pressure more than vested interests or academic input. It is not standards controlling the destiny of the UK – it is the "market need”. This may be real or perceived.

Recommendations

Discussion on importance of customer and market drive to standards through larger suppliers e.g. Ericsson, Huawei etc. Vested interests in standards but most are driven by customer demand for feature. The market decides. Large suppliers build things customers want to buy and sometimes create new standards to realise.

- A good example of the "perceived" market driven approach is the network slicing standardisation example. This may not driven by customer demand, but market hype becoming self-fulfilling.
- A good example of the "real market driven' approach is the Google example: selling proprietary network + handset (in US) allows more features.

In the end, customers are interested in service not underlying standards or how it works i.e. the customers are operating at the Application Layer of our model.

Anecdotally, bodies such as the IETF have become 'slower' as the number of participants has increased. More people involved, more opinions present, more politics. The alternative approach which has been used increasingly over the past decade, has been a 'build it and they will come' approach, outside of the IETF. The UK has the potential to be a leading proponent of the later approach – small enough to deploy new technologies as demonstrators, agile enough to deploy such technologies locally but with sufficient scale to show the power of what might be possible.

The remaining Annex in this paper describes in more detail the outcomes of discussion within the group on the definition of the scope of Core Network Technologies. We structured this Annex in line to the Scope model described in Section 1.1 above. We also have included detail on the "state of the art" and key developments in each of the model areas.

A.1/ Application

A.1.1/ Definition

Application layer serves as a gateway for end users to input information and access network services such as web browsing, email, directory services and file transfer through user friendly interfaces. The user requests are translated by application layer to perform network operations. It ensures that applications can communicate effectively with other applications running on a different network.

Application layer relies on transport layer protocols to establish information transfer between end points. It defines the standards and format for information exchange and specifies the protocols to allow the users to exchange information over the network. Moreover, application layer also allows multiple users to access shared resources over the network such as shared printers and files. In a nutshell, end users directly interact with the application layer to access network services and applications and aids in discoverability and visibility of users over a network.

There are several application layer protocols including Hypertext Transfer Protocol (HTTPS), Domain Name System (DNS), Border Gateway Protocol (BGP) and Dynamic Host Configuration Protocol (DHCP), to name a few. The purpose of this document is not to delve into details of each application layer protocol; however, HTTPS is presented as an example of how users interact with the network services using application layer. HTTPS is used to provide a secure communication over a network to the users using the web browser. HTTPS provides a secure channel for the information exchanged between the user's web browser and web server by ensuring confidentiality, integrity and authentication. However, it is pertinent to mention that since application layer is closest to the users in relation to user interaction with the network services this makes it a prime target for network security attacks due to the large threat surface. Common threat types at application layer include distributed denial-of-service (DDoS), HTTP Floods and Structured Query Language (SQL) injections.

A.1.2/ State of the Art

Application layer protocols and services are evolving with the advent of new technologies such as quantum computing. The current research and development related to application layer focuses on issues such as application layer protocol security, massive connectivity, integration of edge computing and support for artificial intelligence (AI)-powered applications.

The main aspects of application layer that are being explored currently and will support the future technologies include: scalability, to ensure that increasing workloads due to growing user traffic are handled efficiently; resource sharing, to ensure that network resources are available to the users; resource efficiency, to ensure that the multiple users can access shared resources leading to optimal utilisation; interoperability, to ensure that diverse applications can interact effectively; energy efficiency, to ensure that energy consumption is optimised while maintaining the required functionality and performance; low latency, to ensure that applications are able to provide real time response to user inputs and requests. Standardisation within this layer has also been of keen interest as it supports interoperability to ensure that applications from different vendors can interact effectively and securely.

Quantum computing, which leverage the principles of quantum mechanics unlike traditional computing that utilises classical physics, will bring in new challenges and opportunities in the context of networks. Quantum computing envisions quicker complex problem solving providing it an edge over traditional computing. It is expected to have considerable impact on AI applications because of its ability to perform calculation faster and effectively. Application layer bridges the interaction between quantum hardware and users, allowing users to benefit from the power of quantum computing. Therefore, quantum networking will bring in new opportunities and challenges which will also impact on how application layer technologies and protocols evolve.

A.2/ Transport & Session

A.2.1/ Definition

The purpose is the transmission control of information between endpoints at the operating system's application layer – where the function is to control transmission, receipt and retransmission of information between the end-points.

The transport layer protocols provide the interface between the basic, network layer packet transmission, and the communication services required for the communication endpoints within applications.

All transport protocols should provide multiplexing capabilities to delineate communication into separate "flows" for applications.

Transport protocols should provide the ability to react, adapt and respond to network capacity conditions, such as congestion, that may occur in the underlying network layers.

In principle, the lower layers of the network stack do not know about the transport and session layers, other than for its own purposes (its own applications such as routing protocols, which are a form of network control and management functions). [Case in point is MP-TCP, where the transport protocol wants to take advantage of multiple transmission paths between end-points but again the core network is unaware of this.]

One of the key functions of the transport layer is to compensate for the potential congestion and/or traffic loss occurring in the lower layers. If congestion occurs within a network, it typically results in packets traversing the network being discarded. There is no mechanism by which the lower layers indicate such congestion conditions towards the transport layer. Rather, the onus is on the transport layer to determine that congestion (or loss) is occurring and to signal to the higher layers (towards the application) to throttle its data transmission sending rates accordingly.

In today's networking stack, there is little meaningful interaction downwards, between the transport/session layer and the underlying network layer. Although the functionality exists in end-point operating systems and in networking equipment operating systems for the transport layer to 'mark' traffic in some way (providing a means of identification of traffic belonging to an arbitrary 'class' such as 'real-time' vs 'non-real time') in manner such that the classification may be used by the lower layers to provide some kind of treatment, the practical reality is that this may or may not be honoured by the network or across multiple administrative domains.

[Core networks do NOT typically perform any form of filtering of the traffic since this requires classification and interception which impose performance limitation in an area where speed and efficiency are key.

Network admission control, access control systems and other similar middleboxes tend not to be present in the core of the network nor are they core network devices. They tend to be edge devices, where control is being exerted over information entering the network or where they are being used to 'distribute' traffic towards a set of endpoints.]

A.2.2/ The importance of congestion control in the Core network

A network may look to protect itself against applications that can contribute to congestion, especially from "greedy" or "misbehaving" flows (e.g. from transport protocols that do not implement suitable congestion control, or are malicious – trying to consume as much network capacity as possible). Mechanisms can be implemented in the network layer to monitor and police traffic based on some pre-defined administrative policy, where traffic exceeding the policy is discarded, in turn causing the transport layer to react, adapt and respond.

Boundary Functions

Depending on the network connectivity, e.g. home network (using domestic broadband services) or enterprise network (using commercial fixed/leased connectivity) there are boundary functions that might require some cooperation with the Core network.

Such functions are typically based on traffic monitoring, and application of algorithmic policing and/or policy-based constraints, perhaps in accordance with a service-level agreement (SLA) between user and provider.

Functions could include traffic shaping; traffic filtering; traffic / path engineering, proxy/relay and application gateways (middleboxes). In a home network, these are typically applied independently within the ISP network, so from the home user's point of view, they are "Core network" functions. In an enterprise network, such functions might be distributed between site-border routers at the customer network and the ingress to the enterprise's service provider network: i.e. the boundary functions might rely on visibility of state above the network layer and cooperation with the site network.

Main technologies

At the transport layer, the technologies used matter in terms of how congestion control and boundary functions can be applied to them. There might be transport protocol functions or mechanisms that are in tension with network layer functions. The Core network can use Explicit Congestion Notification (ECN) in suitably instrumented transport protocols to provide cross-layer signalling for cooperative congestion control. Use of ECN has been defined for use in IP since 2001 (RFC3168(PS)), and is now considered Best Current Practice for Internet routers with Active Queue Management (AQM) (RFC7567(BCP)).

The only other mechanism available to the Core network to signal congestion (and to control congestion) is to pre-emptively drop packets with the aim that a transport protocol will respond to the the detection of packet loss by reducing transmission rate. This controlled packet-dropping is enabled through Active Queue Management (AQM), the use of which is now considered Best Current Practice in Internet routers (RFC7567(BCP)).

For both ECN and any traffic monitoring and policing with AQM, different vendors will have different algorithms that are used, and so behaviour might not be the same. However, as TCP is the prevalent transport-layer protocol, and has been for a number of decades, there is the expectation that there will be "TCP-like" or "TCP-friendly" behaviour.

TCP is still the most widely used transport protocol on the Internet. As well as being used for many popular applications for the user-data flows (HTTP/2, video streaming protocols, file transfer), it is also used in the control-plane for implementing signalling channels for other applications (e.g. BGP, conferencing applications, network management). Increasingly it is used with Transport Layer Security (TLS) so that TCP flows can be encrypted for security, and some authentication in initial communication set-up is possible.

UDP is increasing in use on the Internet, as it is used for the audio and video streams of real-time (interactive / conversational) applications (Teams, Zoom, some online-games such as RPG etc). UDP is not used directly for such services, but the Real-time Transport Protocol (RTP) (RFC3550(S)) is used, which is layered on top of UDP. As real-time applications increase in usage, so will the relative amount of UDP traffic. Unlike TCP, UDP and RTP do not implement any form of congestion control, rather, they reply the application layer to determine what do under congestion conditions.

Multipath-TCP (MP-TCP) (RFC8684(PS)) is an extension to TCP to use simultaneously mutliple network paths between end-points. MP-TCP does support congestion control, with each path being a "sub-flow" and using congestion control mechanisms as with single-path TCP. MP-TCP is not so widely used across the Core Internet (multiple addresses per host are required), but has notable use cases in the mobile domain and access network, e.g. inverse-mutliplexing across 4G and 802.11/WiFi.

The QUIC protocol (RFC9000(PS)) is increasing in usage. It is already well-supported by many commercial OS providers (Apple, Google, Microsoft). A key motivation for the use of QUIC is that it is the transport protocol defined for HTTP/3 (RFC9114(PS)). QUIC uses UDP, but adds additional functions, notably the ability to have multiple "streams" multiplexed within a QUIC flow with each stream controlled separately, and the use of end-to-end encryption as part of its definition, rather than relying on the use of TLS (as a separate protocol as is the case for TCP, e.g. when used for HTTP/2 and other protocols that need secure TCP communication). Congestion control based on packet loss detection and on the use of ECN are both defined for QUIC (RFC9002(PS)). MP-TCP (gracefully) degrades to single-path (normal) TCP usage for backwards compatibility.

A.2.3/ State of the Art

TCP is still the most widely used transport protocol:

- TCP has congestion control, and there are number of different congestion control algorithms in use.
- TCP responds to packet drops, and modern TCP implementations SHOULD support ECN by default (RFC9293(S)).
- Works directly on top of IP.
- There is good visibility of TCP header information and so middlebox functions can be applied based on IP and TCP header information using well-known techniques.
- Uses TLS for security.
- Usage likely to decrease relatively, displaced by QUIC as HTTP/3 increases in usage, and new applications appear opting for use of QUIC, but usage for control plane (such as for BGP) will remain.

UDP is also widely used, but is often treated not as a transport protocol in its own right, but as a way of getting access to IP with the bonus of flow multiplexing (port numbers) support within the OS.

- No congestion control.
- Works directly on top of IP.
- The UDP header has little information, so not much middleboxes can do with it.
- Can use Daragram TLS (DTLS) (RFC9147(PS)) for security, but DTLS does not provide exactly the same functionality as TLS for TCP.
- Usage will increase (real-time applications, and QUIC), but DNS usage already decreasing (due to centralised DNS services over TCP).

RTP is widely used for real-time and/or interactive applications.

- No congestion control in the base protocol, but extensions exist for enabling congestion control (RFC8888(PS)).
- Uses UDP.
- Additional header to UDP.
- If the RTP header is visible, middleboxes can use that information.
- No security in base protocol, but options / extensions available (RFC9335(PS), RFC6904(PS), RFC3711(PS)).
- Usage likely to increase as with the use of real-time applications, and RTP is also being defined for use over QUIC (IETF work in progress).

MP-TCP has specific uses in specialised domains.

- Has similar congestion control to TCP.
- Works directly on top of IP.
- MP-TCP-aware middleboxes are possible, but do not appear to be widely deployed.
- Has some built in security for its control plane, not clear it can use TLS.
- Existing usage likely to remain stable, usage not likely to increase in new applications as multipath for QUIC matures (IETF work in progress).

QUIC is growing in usage, and HTTP/3 will only work over QUIC.

- Has similar congestion control to TCP (loss-based and ECN-based).

QUIC is growing in usage, and HTTP/3 will only work over QUIC.

- Has similar congestion control to TCP (loss-based and ECN-based).
- Works on UDP.
- Middleboxes see very little of the QUIC header to the use of end-toe-end encryption.
- Has built-in security equivalent to TLS v1.3 (RFC9001(PS)).
- Usage likely to increase, and displace some usage of TCP for web applications, as well as displacing some usage of UDP for real-time applications. However, from the point of view of the core network, this will look like UDP, though with better behaviour with respect to congestion control.

These are not so widely used, or are for specialised application domains:
* SCTP (RFC9260(PS)), used for signalling (including SS7 and mobile).
Application-layer requirements have tended to drive functions within the transport and session layers. QUIC is an exemplar of this. Attention should therefore be paid to 'new' forms of application that bring them, their own sets of requirements for performant operation. One such application is the 'training' of AI models. Model training is a form of an application, where there very particular requirements in respect to traffic distribution, flows periods and such like. Other emerging applications include Quantum computing applications and large-scale AR/VR environments – while nascent today, it is likely that these will have an impact on the evolution of transport and session protocols. It must be acknowledged that whereas AI training and Quantum computing are likely to present predominantly in data-centre environments, and that these newer applications are likely to yield research efforts intended on providing improved performance within the transport and session layers, TCP and UDP (RTP) will remain in common use across most public networks.

Use of transport protocols in the Core network.

Applications supporting control plane and management plane activities for administration and management will be used in the core network. So, transport protocols will be in use in the core network.

For example: (i) the Border Gateway Protocol (BGP) (RFC4271(DS), RFC4545(PS)) uses TCP for communication with BGP peers; and (ii) some ASs use SNMP (RFC3417(I)), which was originally designed to work over UDP, but can also work over TCP (RFC3430(PS)).

The Stream Control Transport Protocol (SCTP) (RFC9260(PS)) is used in mobile networks especially, to allow interfacing with SS7 and SIGTRAN for such networks. It has similarities to TCP, but some notable and important differences: (i) it is message-oriented rather than stream-oriented; (ii) it support multihoming directly by the use of IP address values directly. This makes is useful for network management applications working across network boundaries.

The use of transport protocols in such cases for the core network is unlikely to change soon or quickly. There is no requirement or incentive for mature control and management systems using these mature protocols to change, e.g. 5G systems use SCTP for control plane and management plane functions.

A.3/ Physical / Data-link

A.3.1/ Definition

The physical medium over which information is passed between nodes.
This layer is typically under the control of a single administrative domain – There is the concept of addressing, covering both end-points and intermediary points. The physical/data-link layer provide reachability between end points (2 or more). An intermediary point may relay information towards an end-point or an intermediary point towards the end-point. The data-link layer imposes order over the structure of information placed on the link as well as methods by which information can be 'addressed' towards another node. Key is that the data-link layer address is not changed when passing between end-points or across an intermediary point. Between administrative domains – the physical/data-link layer can connect endpoints and intermediary points between a pair of administrative domains.

The goal is to provide a reliable communications medium.

A key concept is that between a pair of endpoints is a shared physical layer – the same medium type being used by both points. The two physical endpoints are homogeneous – if the entry point is wireless, the exit point is wireless. If the entry point is optical, the exit point will be optical, too. If information needs to be physically transformed, for example, optical encoding to electrical encoding on metal, then a device is required to perform the transformation.

Different physical layer technologies have differing capabilities and efficient deployment scenarios depending on the required 'reach', e.g. local-area deployment vs wide-area. In some cases, variants of one technology can be used in order to address a different deployment scenario, e.g. local area ethernet on copper to wide-area ethernet on fibre optics. While the encoding on the physical layer changes, the data-link format will still the same between any two nodes (for the most part).

The addressing mechanism used at the data-link layer is local to that link and all nodes connected to it. It is not administrative-wide. All devices on a shared medium require common timing and the ability to synchronise.

What technologies –

- Optical comms & components
- Electro-magnetic communications (radio frequencies)
- Electrical communications
- Encoding/decoding components (MAC/PHY)
- Device that are capable to passing photons or electrons from one 'interface' to another. The act of passing photons or electrons can only be performed by a physical device. The control mechanism that determines how and to which interface a photon or electron is directed could be provided in either hardware or software.
- A notional separation of control-plane function and data-plane function.

A.3.2/ State of the Art

Attention needs to be paid to adaption technologies – solutions that provide a physical/data-link layer on top of higher layers of the network. These can be considered as a form of 'overlay'. Examples include various Layer 2 tunnelling protocols such as L2TP, MPLS L2VPN and VxLan. Such tunnelling mechanisms are typically employed when a change occurs in network implementation that are disruptive to the data-link layer. The tunnelling mechanism provides a means by which the data-link layer functionality can be maintained.

Particular applications have emerged which concentrate in the physical/data-link layer without extending up into the network layer. These can be considered as a form of 'underlay'. Prime example of this include high-performance computing (HPC) and AI model training (GenAI being just one example of this). These applications, which typically occur within a single data-centre estate, drive for particular performance characteristics. This has resulted in data-link technologies that aim to provide for these high performance scenarios, such as the connection of CPUs and memories for direct data transfer using mechanisms such as Remote Direct Memory Access (RDMA). Technologies include Infiniband and RDMA over converged Ethernet (RCOE), providing a high-speed interconnect for computing and memory components which do not necessarily require high layers (such as the application layer) to be involved.

While these technologies are deployed within a single estate, like previous underlay technologies, it is likely that at some point in the future, overlay solutions will be required to enable to connection of computing resources across estates.

A.4/ Network

A.4.1/ Summary

The Network Layer provides the glue between physical layer connectivity and the application traffic flows that assume and require seamless end-to-end transport and support for service differentiation. The network builds on top of physical and data-link infrastructure that is more tightly coupled with node-local hardware and scope and provides the logical abstraction of a network across different scopes (from local to Internet-wide) to network-wide and end-to-end services and applications. The main technologies included in this layer are routing protocols and functions such as OSPF or Software Defined Networks (SDNs).

Again, standards are set by international bodies that have contribution from UK experts drawn from operators, suppliers, and academia, but there is little UK specific R&D on the fundamental technologies. The technologies are however used, to some extent, in every network that is built for applied telecoms research where the has strength.

The major issues here are how do we incrementally evolve the network? – historically big step changes have not too successful. IPv6 for example is still not fully adopted despite being in developed since 1998 as a solution for network address space limitation.

Therefore, to increase the value of this technology to the UK a key question is - How do you make momentum for deployment happen faster? A strong economic driver is required, an example being legacy network replacement was very slow until electricity prices increased a lot.

Niche customer needs maybe a driver - Example: reliable network coverage everywhere, mission critical customers. Heterogeneous networks with connections back to the internet.

Niche customer needs maybe also be a driver in the example of reliable network coverage everywhere, mission critical customers. This could be architected as heterogeneous networks with connections back to the internet.

An alternative network protocol is the Identifier Locator Network Protocol (ILNP), RFC6740-6748, “Experimental” status protocol from the IETF, with IANA codepoints, engineered as a superset of IPv6, some global connectivity tests completed successfully, full deployment capability not yet full assessed.

https://ilnp.cs.st-andrews.ac.uk/

A.4.2/ Definition

The network comprises the technologies, architectures, and protocols that enable the development of advanced services between endpoints on top of physical point-to-point connectivity. These include addressing architectures to support end-to-end, bidirectional interconnection of myriad heterogeneous devices; switching and routing within and across administrative domain boundaries to enable performance and policy-based traffic flow; and mechanisms to support the deployment of multiple services for diverse traffic flows, from traffic engineering to network-compute integration through network (dataplane) programmability and function virtualisation, to slicing and isolation.

The network provides the glue between physical layer connectivity and the application traffic flows that assume and require seamless end-to-end transport and support for service differentiation. The network builds on top of physical and data-link infrastructure that is more tightly coupled with node-local hardware and scope and provides the logical abstraction of a network across different scopes (from local to Internet-wide) to network-wide and end-to-end services and applications.

A fundamental property of mechanisms operating at this level is the ability to scale with global size and increasing traffic rates to 100s of gigabits and beyond. This is particularly important to consider in relation to programmable silicon and network (dataplane) programmability architectures that aim to accelerate innovation at this level.

The network also provides integration between mechanisms that operate as part of the data, control, and management planes, and interfaces with the Security and OAM (Operations – Admin – Metrics) verticals to support services provided by transport (e.g., firewall and intrusion detection middleboxes), and to expose interfaces for measurement and troubleshooting.

A.4.3/ State of the Art

Arguably the primary role of the Network is to provide routable paths between its ingress and egress, and unique addressing and naming within its scope, from local area to metropolitan, to backbone, to end-to-end over the Internet.

Routing protocols within the boundaries of an Autonomous System (AS) of single administrative ownership typically try to establish shortest paths between ingress and egress using number of hops, inverse of bandwidth or other relatively slowly changing metrics to represent weights of each of the network-level links comprising a network path. Such Interior Gateway Protocols aim to optimise performance of routing, and calculate network-wide topology based on implementing distance-vector (e.g., Routing Information Protocol – RIP – RFCs 1058, 1388, 1723) or link-sate (e.g., Open Shortest Path First – OSPF – RFCs 1131, 1247, 1583, 2178, 2328, 3101, 5709, 6549, 6845) algorithms to construct ingress-to-egress paths either based on propagating global state to adjacent links or adjacency state globally, respectively.

Since the late-2000s, significant research and development efforts focused on the design of network architectures and primarily intradomain routing protocols for datacentre networks mainly to support large partition-aggregate workloads that underpin many Cloud Computing services. Such network architectures included Clos, fat-tree and server-centric (e.g., DCell) architectures that used high network path redundancy and mechanisms to balance load and maximise bisection bandwidth in short timescales. In this context, Equal-Cost Multi-Path (ECMP) routing has been employed to balance load over redundant paths based on hashing flow-related data in the packet header to avoid packet-reordering within flows (RFC 2992).

Further research and protocol development has focused on routing algorithms for better or explicit support for load balancing and Traffic Engineering (TE) across network paths such as Segment Routing (RFC 8402) that allows a sender of a data packet to partially or completely specify the route the packet takes through the network, and OSPF-TE extensions (RFC 3630) to add traffic engineering capabilities to intradomain link-sate routing and interoperate with data link layer TE protocols such Multi-Protocol Label Switching (MPLS – RFC 3031). In the last 10 years, use of Segment Routing (SR) (RFC8402(PS), RFC9256(PS)) has started to displace original MPLS usage, e.g. use of Segment Routing with MPLS (SR-MPLS RFC8660(PS), RFC8663(PS)).

Across AS boundaries, routing is based on enforcing policies that reflect business relationships and agreements between interconnecting network operators. The Border Gateway Protocol (BGPv4 – RFC 4271) is the de facto Exterior Gateway Protocol used over the Internet today that defines global connectivity based on a distance-vector routing algorithm that uses weights to reflect peering and transit policies between providers rather than performance metrics directly.

Naming and addressing architectures ensure that every node connecting to a network has a unique identifier through which it can exchange data within the network scope. Over the Internet, unique identification of any system connected to the global medium is required. The Internet Protocol (IPv4 – RFC 791) provides globally routable addressing end-to-end over the Internet with the main challenges being the scale of unique addresses that need to be accommodated and the scale of global routing tables that need to be maintained in Internet routers (in conjunction with executing the BGP algorithm above). The original IPv4 still in use as Internet's main addressing architecture today can accommodate up to a theoretical maximum of ca 4bn addresses since these are captured within a 32-bit integer in the relevant protocol fields. The tremendous growth of the Internet over the past decades has been supported mainly through mechanisms such as Network Address Translation (NAT – RFC 2663) that virtually expand the globally routable IP address space by using a one-to-many mapping between public and private (local network scope) IP addresses. IP version 6 (IPv6 –- RFC8200(S)) has been standardised and implemented widely, mainly to address the IPv4 address shortage and able to support ca 340 undecillion addresses.

IPv6 also offers other advantages over its predecessor including the selective processing of optional packet header data and Segment Routing and, although we are now over 10 years past the World IPv6 Launch (6 June 2012) the majority of traffic is still carried over IPv4 for most parts especially of the western world[13] – demonstrating the challenges of adoption of technology at Internet scale.

The Internet addressing architecture implicitly ties node identification and network location due to way blocks of IP addresses are assigned by the Internet registries. This hinders the seamless development of IP(v4/v6) mobility, multi-homing, IP security, etc. To address these issues and improve future-proofing of the Internet Protocol, the Identifier-Locator Network Protocol (ILNP – RFC 6740) has been developed to divide the two functions of network addresses (topological information form network identity) by replacing the concept of an IP address with separate Locator and Identifier names. ILNP is backwards-compatible with existing Internet Protocol v6 functions, and is incrementally deployable.

Naming is another fundamental aspect of the network as it facilitates the translation between unique machine identifiers (addresses) and humanly memorisable names for different types of content. The Domain Name System (DNS – RFCs 1034 1035) is the hierarchical and distributed naming system for computers, services, and other resources across the Internet. Furthermore, over the past fifteen years, research and development efforts have explored Information-Centric Networking (ICN) as the evolution of the Internet infrastructure from the existing host-centric design to a data-centric architecture altogether, where data by name becomes the essential and addressable network primitive independent of data location, hence enabling native multicast delivery, ubiquitous in-network caching and replication of data objects. Named Data Networking (NDN) and Content-Centric Networking (CCN) are two realisations of the ICN architecture (ICN – RFC 7927) that implement a lookup service through defining two packet formats: Interest packets that request content by name, and Data packets that carry the requested content. ICN nodes then fulfil the role of a data producer, a data consumer, and/or a forwarder for Interest and Data packets, while the forwarding plane in an ICN network relies on three data structures that capture a node's state: a Forwarding Interest Table (FIB), a Pending Interest Table (PIT), and a Content Store (CS). Even though ICN efforts rapidly gained momentum, it is now less clear that the benefits of deploying such a radically different architecture at Internet scale outweigh the cost, while associated challenges of adoption might prove insurmountable.

[[12] IPv6 10 Years Out: An Analysis in Users, Tables, and Traffic

Time-Sensitive Networking (TSN) is another area of work to enable delivery of data with strict timing guarantees and hence able to support applications with deterministic delivery requirements such as audio and video, industrial control networks, safety-critical and mixed-media networks (e.g., in vehicle). TSN requires three key component categories to realise its full potential, sub-microsecond precision time synchronisation across all participating nodes and end-devices achieved through the network itself, scheduling and traffic shaping to allow for the coexistence of different traffic classes with different priorities on the same network, and path selection, reservation and fault tolerance to ensure sufficient resources are in place to guarantee delivery of certain (and multiple) traffic profiles.

Software-Defined Networking (SDN) has been a major development over the past fifteen years in an effort to horizontalize a traditionally vertically integrated industry where vendors had full control of their products and services they offered. SDN defined a protocol and a paradigm that allowed the separation of the network control plane from the routers which can now configure packet forwarding policies through a well-defined API that can be programmed by network administrators and also application developers. SDN strove to enable innovation in a so far closed ecosystem, reduce complexity of the devices themselves (through only encoding the required functions and algorithms at any given time), and facilitating network telemetry and increase visibility into the network's operation.

Even though SDN let network operators take charge (from vendors) of how their networks are controlled, it had various shortcomings including the inability to perform line rate packet processing due to the frequent crossing of the control-data plane interface. Network (dataplane) programmability has been introduced over the past decade to now give the operators control of how their packets are being processed. Advances in (and cost reduction of) programmable silicon meant switches can now support arbitrary packet matching and programmable pipelines in hardware, and support the development of real-time per-packet services including firewalling, load-balancing, and access control, all the way up to executing machine learning-based inference. P4 has been a dominant dataplane language to program reconfigurable match-action table (RMT) router architectures, it has been supported in hardware (by Barefoot networks and later Intel) and demonstrated widely to develop diverse dataplane functionality at high speeds and high switch port densities. At the same time, supporting stateful operations in the RMT architecture is challenging and costly (due to recirculation) while at the same time maintaining line-rate bound performance. One of the main use-cases for dataplane programmability has been In-band Network Telemetry (INT), where programmable switches can add custom monitoring and path information to packet headers to then be consumed transparently at remote nodes in the network.

Both SDN and dataplane programmability have been mainly showcased over datacentre networks where an operator has full control of the hardware-software infrastructure and less so over backbone and telecommunication networks over the Internet. More recently, the need to support low-latency services at the Edge has led to the advent of application offload and network stack bypass frameworks, including the extensive support for the eBPF instruction set by the Linux kernel and Intel's Data Plane Development Kit (DPDK). Exploiting such device driver and OS support (e.g., Linux's eXpress Data Path – XDP) for in-network application acceleration has been shown to offer orders of magnitude performance increase in application code while being able to support more (or complete) stateful processing.

Both Linux and FreeBSD have opensource routing platforms that could be used for exploration and innovation within this space. Although Linux is more widely used as base for products commercially, that FreeBSD, the latter has a less restrictive licence which is attractive for some commercial actors. Examples of opensource platforms that could be (and have been) used by industry include FRRouting (https://frrouting.org/) based on Linux, and OPNsense (https://opnsense.org/) based on FreeBSD.

A.4.4/ Challenges

The requirement for the network and mechanisms operating at this layer to scale in space and time potentially to Internet-scale poses significant challenges to the adoption of new technology. At the same time, user demands from the network are increasing, especially since connectivity is increasingly considered a commodity and users will pay a premium for advanced services on top of connectivity. User requirements such as privacy are increasingly taken into account when designing new services and legislation, yet requirements around higher availability and predictable or guaranteed performance receive less attention not least because of the complexity of implementing and the amount of state required to support them especially across administrative boundaries.

The nearly thirty years that have passed since the specification of IPv6 and the challenges that still prevent its more widespread deployment and eventually succession of IPv4 as the main addressing and routing architecture over the Internet demonstrates the need to develop mechanism and technologies that can be adopted faster and in a more incremental manner, and the need to make momentum for deployment happen faster. This is particularly crucial at the network layer since this is where the systemisation of the network components starts.

A.5/ Control Plane & Policy

A.5.1/ Definition

Policies are defined by the operators of a network to provide guidance on how the network should behave. e.g. What traffic should be prioritised or what quality of service should be offered to subscribers. The control plane is then responsible for establishing the topology and the dynamic configuration of the network used to carry user traffic within the constraints of the defined policies. The control plane can interact with the user plane at all layers of the networking stack (Session, Network, Physical); in order to establish the required network topology and configuration. The control plane and the user plane are distinguished by the typical speed of the processes involved; the control plane runs protocols that may over time lead to decisions to change the network topology or configuration; the user plane processing is much faster and more time critical and substantial effort is put into increasing the performance.

Examples of things that live in the control plane:

- The dynamic configuration of routing and switching
- The dynamic configuration of quality of service
- The dynamic configuration of network slicing
- The scaling up and down of the network to dynamically increase/decrease capacity
- Supporting troubleshooting initiated by the management plane
- The gathering of measurement needed by the management plane
- Device authentication controlling which devices can participate in the network

Access control defining who is allowed to update the network. Examples of things that are excluded from the control plane.

- Subscriber management - this lives in the management plane and is being handled by another working group
- Subscriber authentication - this lives in the security plane and is being handled by another working group
- Static management of network components - this lives in the management plane and is being handled by another working group

A.5.2/ State of the Art

There are a vast number of protocols used as part of the control plane and policy. Some the more significant ones are:

ICMP – Used alongside IP; typically used for diagnostic or control purposes or generated in response to errors in IP operations.

RRC – Protocol used between a UE and the 3GPP NodeB in order to setup / modify / teardown data sessions.

NAS – Protocol used between a UE and the 3GPP core network in order to setup / modify / teardown data sessions.

RTCP – Control protocol used alongside RTP to monitor the performance of the RTP stream.

OSPF – a routing protocol for IP networks.

BGP - a standardized exterior gateway protocol designed to exchange routing and reachability information on the Internet.

MPLS – Can be used alongside IP routing for traffic engineering providing performance gains for packet-forwarding.

SR – segment routing has a similar goal to MPLS, and, indeed, it is possible to use MPLS as a way of provisioning SR capability.

X-Apps / R-Apps - monitor the performance of the 3GPP radio network and adapt the network behaviour based on the observed performance.