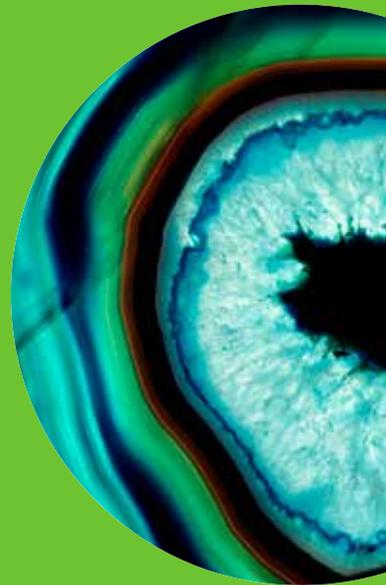


Review

ERICSSON
TECHNOLOGY



MULTIPATH TCP & THE LAST MILE



BOLSTERING

The last mile

WITH MULTIPATH TCP

The rapid uptake of bandwidth-consuming services such as video on demand and linear TV has many service providers struggling to keep pace with ever increasing bandwidth demands. The problem is particularly acute on the last mile: the segment of the network that delivers broadband services to users' homes and workplaces.

ROBERT SKOG,
DINAND ROELAND,
JAUME RIUS I RIU,
UWE HORN,
MICHAEL ERIKSSON

AS AN ALTERNATIVE to building out the physical communications infrastructure – which in some geographical areas may be too costly or time consuming – Ericsson proposes an access aggregation solution based on Multipath TCP. Our solution consists of a carrier-grade Multipath TCP proxy that allows the use of Multipath TCP across access networks without the need to introduce it in end devices or internet servers.

■ The last mile is the part of the telecommunications network that physically reaches user premises, either by wireless technology (cellular networks) or wireline technology such as cable, fiber or digital subscriber line (DSL). The achievable data rates for each of these access

technologies vary, but in many cases the bandwidth depends on the distance between the access termination point in the service provider network and the device in the user premises. This means that no matter how fast the service is up to the access termination point, the users who are farthest away from it will experience significantly slower service than the ones who are closer.

For example, although the most recently standardized DSL technologies allow bitrates of up to 1Gbps, most subscribers today are still getting less than 20Mbps. The reason for this is the dependency between the achievable bitrate and the length of the copper line connecting a household to the DSL access multiplexer (DSLAM). As *Figure 1* shows, if the distance between the user premises and the DSLAM

exceeds 2km, DSL speed falls quickly below 20Mbps. The obvious solution is to reduce the length of the last mile. If the copper line distance can be reduced to less than 250m, new technologies and standards such as vectoring and G.fast will allow bitrates of about 1Gbps. However, reducing the copper line distance is costly because it requires the deployment of more street cabinets connected by fiber lines to the backbone network. To get around this, some fixed broadband service providers have started to launch offerings that combine DSL with LTE as a cheaper way to boost the bitrate for DSL customers than deploying more fiber-connected DSLAM street cabinets.

Similarly, LTE/wi-fi aggregation is useful as a booster for mobile phones. Some operators have started deploying solutions that combine Wi-Fi and LTE accesses in areas such as shopping malls and big event venues as a means to increase user capacity while at the same time offloading their cellular network traffic to the fixed networks when possible.

Technologies for access aggregation

Many standardized aggregation technologies only support use cases in which links using the same access type are aggregated. This is known as bonding, and examples include the bonding of several Ethernet links, or of two DSL access links. Notable exceptions are IP Flow Mobility and multiple-access PDN connectivity – both defined by 3GPP – which are able to support aggregation of multiple access types [1]. However, these two technologies have gained little traction because their introduction on mobile devices would require a significant implementation effort, and even the apps running on them would require modifications.

Multipath TCP as specified by the IETF [2] can be deployed in existing networks more easily than other

MULTIPATH TCP CAN BE USED ACROSS ALL KINDS OF ACCESS NETWORKS, PROVIDING A RICH TOOLKIT THAT SUPPORTS ACCESS AGGREGATION FOR USE CASES SUCH AS BANDWIDTH AGGREGATION, RELIABILITY AND SEAMLESS CONNECTIVITY.

alternatives because it is an evolution of TCP [3] – the most widely used protocol in the internet today. This guarantees interoperability between equipment from different vendors. Like TCP, Multipath TCP works on top of IP. Since IP is the foundation of all internet protocols, Multipath TCP can be used across all kinds of access networks, providing a rich toolkit that supports access aggregation for use cases such as bandwidth aggregation, reliability and seamless connectivity. In addition, there is an open source reference implementation for Multipath TCP that is continuously developed and improved by a large community of developers [4].

Figure 2 shows two access aggregation scenarios enabled by Multipath TCP. The first scenario shows DSL/LTE aggregation, where an existing DSL connection is combined with LTE. If the DSL link provides 12Mbps and the LTE link provides 8Mbps, the aggregated

Terms and abbreviations

ACK – ACKnowledgment | **CCA** – Congestion Control Algorithm | **CPE** – customer premises equipment | **CPU** – central processing unit | **DPDK** – Data Plane Development Kit | **DSL** – digital subscriber line | **DSLAM** – DSL access multiplexer | **IETF** – Internet Engineering Task Force | **MFDN** – Media First Delivery Node | **RNA** – Radio Network Aware | **RTT** – round-trip time | **TCP RNA** – TCP Radio Network Aware | **VDSL** – Very high-speed DSL

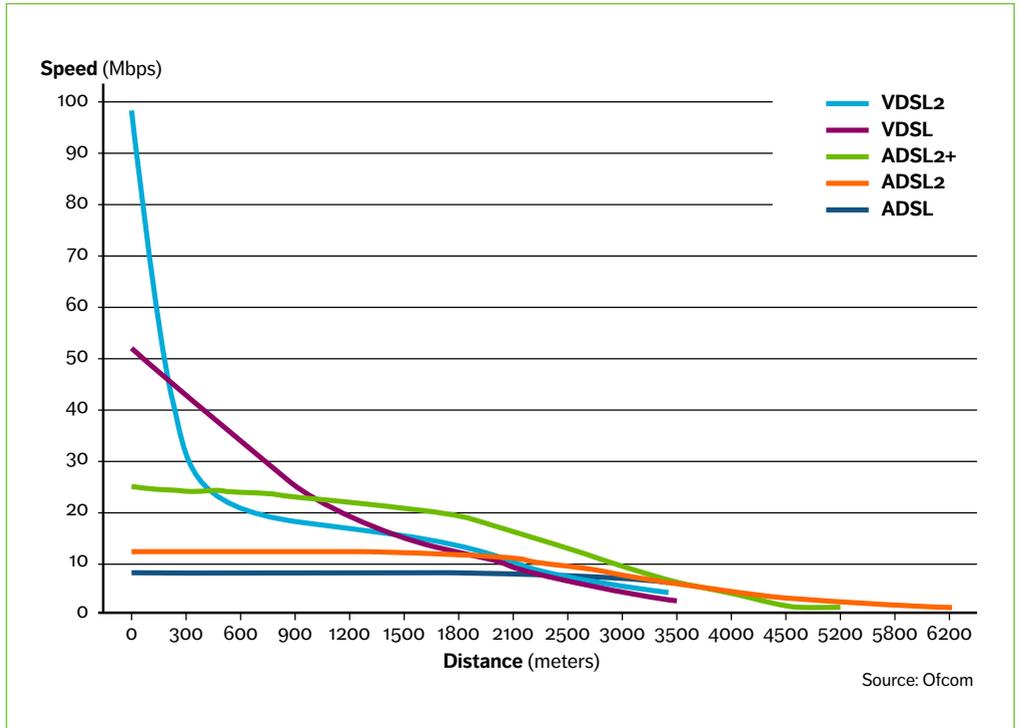


Figure 1 Speed versus copper line length between user premises and the DSLAM for the most widely deployed DSL technologies

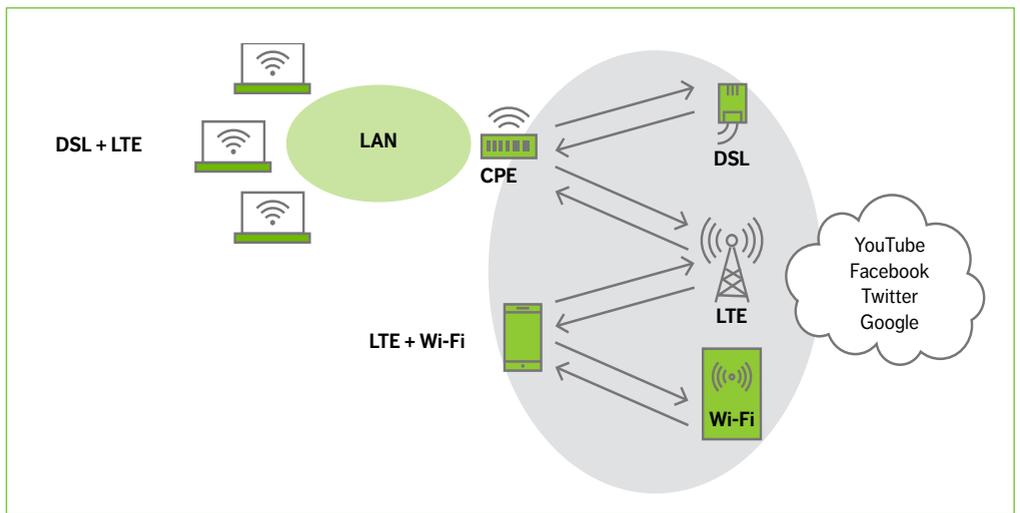


Figure 2 Examples of access aggregation enabled by Multipath TCP

bandwidth that can be obtained via Multipath TCP is roughly 20Mbps.

The second scenario shows LTE/Wi-Fi aggregation, which functions according to the same principle. Together with a mobile device manufacturer, Ericsson has performed successful field trials in public LTE and Wi-Fi networks using commercially available mobile devices. Only the firmware was modified to support Multipath TCP.

Although the benefits of Multipath TCP are often presented in the context of two different access networks, there is no limit in Multipath TCP that would prevent the use of three, four or more access networks. The access networks could even be operated by different service providers, which is an additional benefit for use cases aiming for improved resiliency.

Aggregating bandwidth

Bandwidth aggregation refers to the ability of Multipath TCP to combine the bandwidth of several links into one logical connection. *Figure 3* shows an example of how Multipath TCP adds together the bandwidth of DSL and LTE. This is equally valid for the LTE + Wi-Fi scenario depicted in the bottom part of *Figure 2*.

The bandwidth aggregation features of Multipath TCP apply to both downlink and uplink directions. As a result, Multipath TCP also helps to improve uplink speeds, which are only a fraction of the downlink speed in existing (asymmetric) DSL consumer services. For instance, the uplink speed over a 6Mbps asymmetric

BANDWIDTH AGGREGATION REFERS TO THE ABILITY OF MULTIPATH TCP TO COMBINE THE BANDWIDTH OF SEVERAL LINKS INTO ONE LOGICAL CONNECTION.

DSL connection is usually below 1Mbps. Aggregating DSL and LTE makes it possible to boost the uplink speed to 10Mbps and more.

Examples of services that would benefit from the Multipath TCP bandwidth aggregation are:

- » A user watching HDTV (high definition TV) over a DSL access connection that is not capable of providing enough bandwidth – Multipath TCP can be used to schedule surplus traffic over LTE (particularly useful for the downlink).
- » A user uploading documents or photos to a server – when the DSL uplink capacity is exceeded, Multipath TCP can add LTE capacity for quicker upload.

Improving reliability

In the context of access aggregation, reliability refers to the ability to maintain data exchange within a session, even if one or several access links become unavailable. *Figure 4* compares the behavior of a

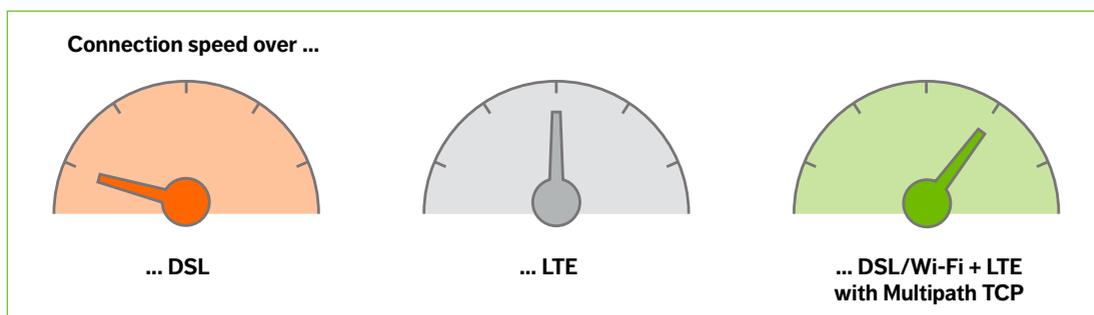


Figure 3
DSL and LTE
bandwidth
aggregation with
Multipath TCP

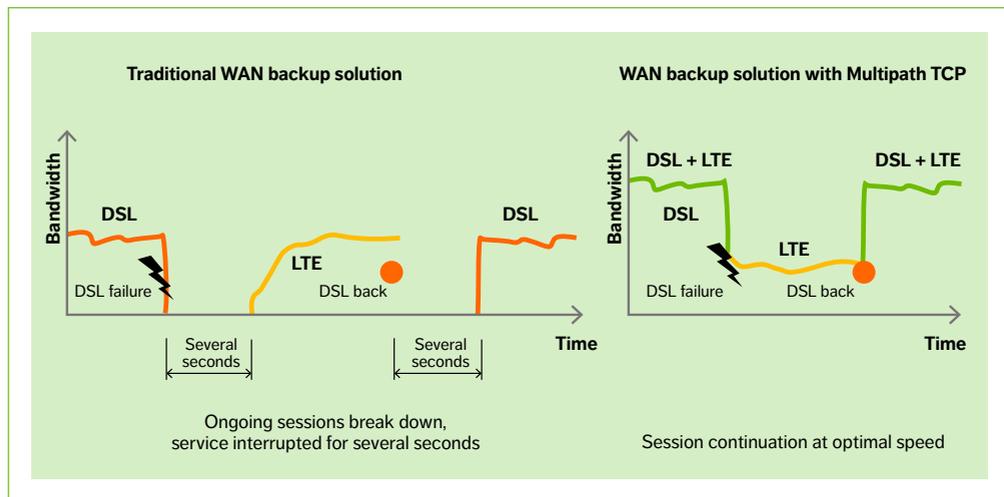


Figure 4 Improved connection resiliency with Multipath TCP

traditional WAN backup solution with that of a solution based on Multipath TCP. Traditional solutions cannot react quickly to the disappearance and reappearance of access links. Whenever a link disappears, sessions break and need to be reestablished, which can lead to data loss and the need for human intervention.

Multipath TCP is able to react more quickly to access links disappearing and reappearing. And as long as at least one access link is up and running, a Multipath TCP enabled session will continue without interruption – albeit at a lower bitrate. Likewise, if an access link reappears, the bitrate goes up. The connection always runs at an optimal speed in relation to the availability of the links involved.

Achieving seamless connectivity

The concept of seamless connectivity is related to reliability, referring more specifically to the ability of Multipath TCP to switch from one access to another without having any impact on the application. A typical use case would be a session started over Wi-Fi. If the mobile device leaves Wi-Fi coverage and enters mobile broadband coverage, the session will break and need to be reestablished. This can be quite

annoying and time consuming for the user, especially if two-factor authentication is involved. With Multipath TCP, the session does not get interrupted due to the change of access.

Changing from one access to another can also be triggered by service provider policies. For example, a service provider could have a policy to use LTE by default, but move some traffic to Wi-Fi when there is good coverage and available capacity. Or, alternatively, the service provider could set a policy where Wi-Fi is used by default and LTE is used to provide wide-area coverage. In all cases, the use of Multipath TCP prevents sessions from being interrupted if and when access systems change.

How Multipath TCP works

TCP [3] is one of the main protocols in the IP suite, providing a reliable means of communication between two endpoints. Once a TCP connection has been set up, both endpoints can send a data stream to each other. TCP is designed to cope with data that is damaged, lost, duplicated or delivered out of order. Furthermore, it provides a means to perform flow control. Upon receiving data, the receiver sends an acknowledgment

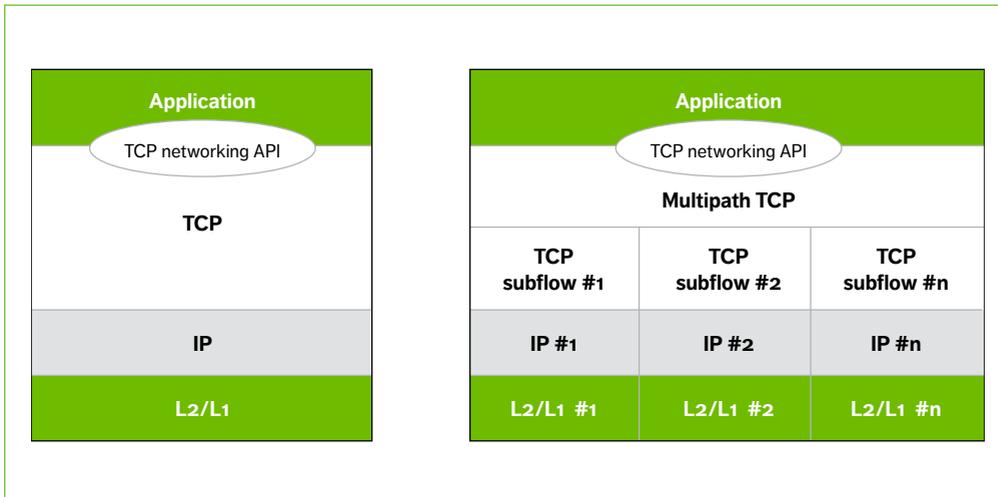


Figure 5 Protocol stack for TCP and Multipath TCP

(ACK) back to the sender. Such an ACK contains a “window,” which indicates the maximum number of bytes the sender is allowed to transmit before receiving further permission. This way, the receiver controls the amount of data transferred by the sender. Finally, the receipt or non-receipt of ACKs guides the TCP Congestion Control Algorithm (CCA) to determine the pace at which data may be sent.

Today, many endpoints have multiple data communication interfaces and therefore multiple IP addresses. For example, a laptop is often equipped with both a wired and a wireless interface, and a smartphone often has the capability to use multiple wireless communication technologies. Using regular TCP, these devices are capable of establishing multiple simultaneous TCP connections, with each connection tied to one specific IP interface. In other words, each TCP connection is bound to a single path defined by the IP addresses of the connection’s endpoints. Note, however, that a path is defined here in terms of endpoint identifiers; it is not the same as the route that individual packets take on their way from one endpoint to the other.

Multipath TCP [2] is a set of extensions to standard

TCP that allows connections to use multiple paths simultaneously. Multiple regular TCP connections, also known as subflows, are aggregated into a single Multipath TCP connection. *Figure 5* compares the protocols stack of regular TCP with that of Multipath TCP.

In regular TCP, an application initiates communication by opening a connection via an application programming interface (API) provided by the operating system. The TCP layer communicates in its turn with the IP layer. In Multipath TCP, the TCP layer has been extended. Upwards, the Multipath TCP layer exposes an interface that is perceived as regular TCP by the application. Downwards, the Multipath TCP layer may set up multiple regular TCP connections. These may be bound to different IP layers. In *Figure 5*, the host is equipped with multiple data communication interfaces. Each one is associated with its own IP address. The Multipath TCP layer aggregates the multiple TCP connections into a single Multipath TCP connection. The application does not need to be aware of which protocol stack is used.

Figure 6 shows an example of how a Multipath TCP connection can be established. It starts with the setup of a first subflow (steps 2-4). These steps consist

of a three-way handshake, similar to the process in regular TCP. The only difference for Multipath TCP is that an `MP_CAPABLE` option is used in the TCP header. With this option, the device indicates to its peer that it is Multipath TCP capable and wants to use it (step 2). If the peer is also able to use Multipath TCP, it replies with a similar capability indication (step 3). As part of the three-way handshake, the endpoints also exchange security keys. After setting up the first subflow, both endpoints can exchange data over the connection (steps 6–7).

Once a Multipath TCP connection has been established, each endpoint may initiate the setup of an additional subflow. In the example shown in [Figure 6](#), the device has two network interfaces. Each interface is associated with its own IP address. Here, the device takes the initiative to establish a second subflow via its second interface. Again, a three-way handshake is used to achieve this. But this time the option `MP_JOIN` is used to indicate that this is a new subflow that is to be joined to an existing Multipath TCP connection. A token (step 9), derived from the earlier received key (step 3), is used to correctly bind the two subflows. Additional authentication information is also exchanged to ensure the authenticity of both endpoints.

Once the new subflow has been established, both endpoints can use it to send and receive data. In our example, the device sends data to its peer (step 14). Note that the device needs to take an active decision regarding which subflow to use (step 13). How this decision is made is not defined in the standard, which gives the designer the freedom to implement the scheduling policy that is most appropriate for each case.

Subflows may come and go for various reasons, such as connectivity problems. To ensure reliable, in-order

delivery to the application, Multipath TCP uses a data sequence number that is carried in a Data Sequence Signal option (steps 6-7 and 14-15). Aside from ensuring in-order delivery, this number can be used in combination with the sequence numbers used by regular TCP at subflow level to execute retransmissions on different subflows, if needed. Multipath TCP can also synchronize congestion control over subflows in order to avoid unfairness to single-path users [5].

An additional benefit of Multipath TCP is that it can be introduced incrementally. In particular, if the receiver of the first subflow's TCP SYN does not support Multipath TCP, it will simply discard the capability option. It will reply with a TCP SYN ACK, but without adding the `MP_CAPABLE` option, and the connection will be made with standard TCP.

The proxy-based approach to Multipath TCP access aggregation

Proxies make it possible to achieve the benefits of Multipath TCP for access aggregation without requiring Multipath TCP support in all end devices and internet servers. An additional benefit of proxies is that they give the service provider control over the scheduling of the traffic. In this way, service providers can ensure that the available access alternatives are used in the most efficient and cost-effective way. The use of proxies has already been recognized by the industry, and work has been done and published by the Broadband Forum defining the architecture [6]. Ericsson is contributing actively to this work.

[Figure 7](#) provides a high-level overview of the proxy-based approach to Multipath TCP access aggregation. There are two proxies involved: a network proxy and a customer premises equipment (CPE) proxy. The

User space

In computer design, a distinction is made between kernel space and user space. Kernel space is where the operating system code runs – hardware device drivers, memory management and protocol stacks, for example. User space is where ordinary programs run. In designing our Multipath TCP solution, we chose to place a protocol stack (MPTCP) in user space rather than in kernel space. This results in faster packet processing, because packets don't need to travel from kernel space to user space. Instead, they go directly from the hardware interface to user space.

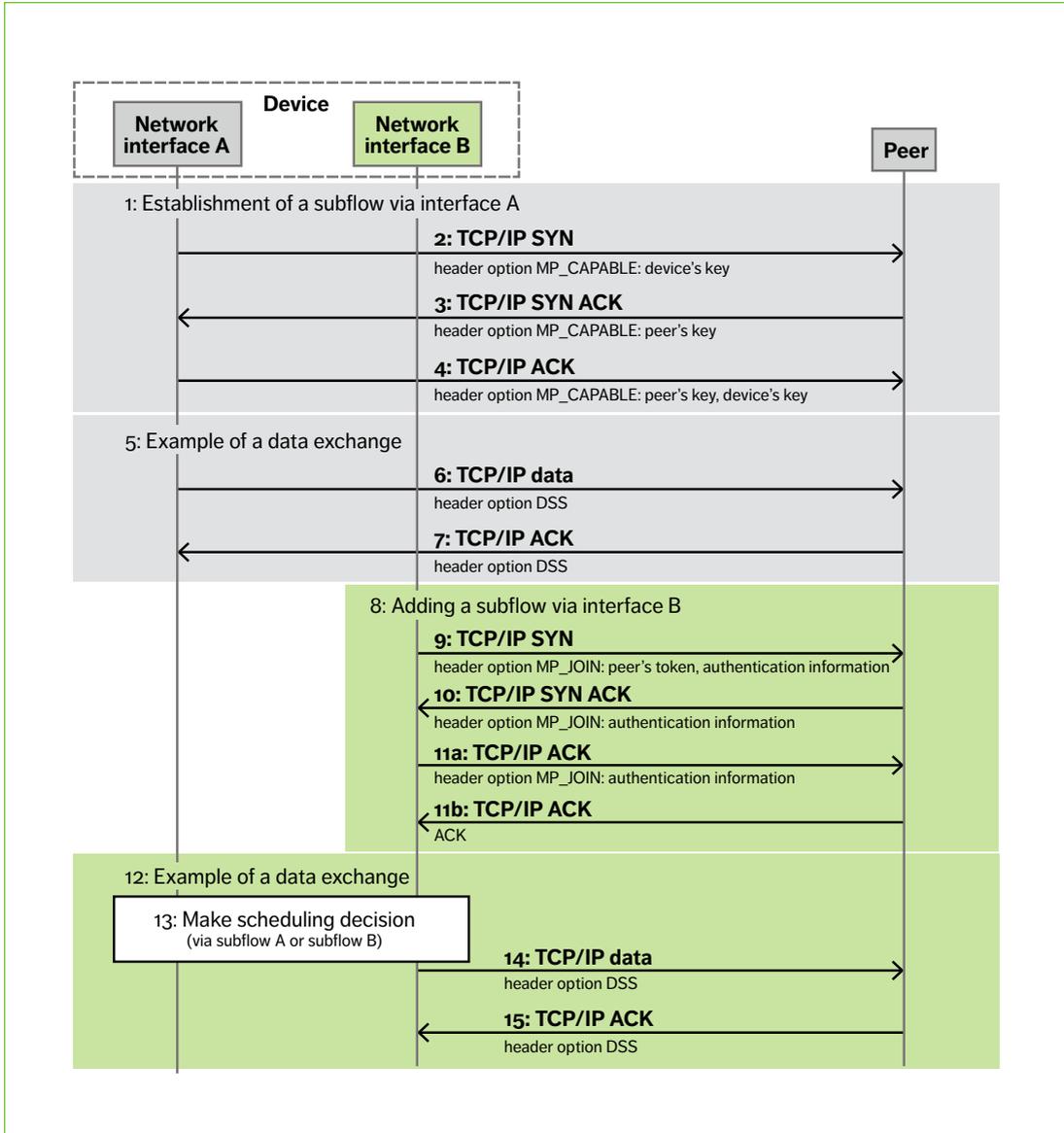


Figure 6 Establishment of a Multipath TCP connection

ERICSSON IS PARTNERING WITH CPE VENDORS AND CHIPSET MANUFACTURERS SUCH AS INTEL TO ENSURE EFFICIENT IMPLEMENTATION OF THE MULTIPATH TCP CPE PROXY.

network proxy is located in the service provider's network and converts TCP sessions from internet servers into Multipath TCP sessions that operate across multiple access networks. Similarly, the CPE proxy converts a Multipath TCP session with the network proxy back into a TCP session.

End devices with built-in Multipath TCP support could also connect directly to the network proxy. There are already some smartphones on the market with built-in Multipath TCP support that can be used to aggregate LTE and Wi-Fi. Ericsson has run tests that prove the feasibility of this setup in public LTE and Wi-Fi networks.

The proxies can be used to enhance standard Multipath TCP via additional traffic-steering capabilities that are optimized for the specific application scenario. For instance, a service provider might want to ensure that the DSL pipe is filled first before using the scarcer LTE bandwidth. This traffic-steering approach is often referred to as a cheapest-link-first policy. Service providers might also want to define policies to prevent or allow the use of heterogeneous access for specific services, or to force selected services to use only one of the available access links. All of this is possible with Multipath TCP, as the IETF standard does not prescribe a specific traffic-steering method.

In an implementation, the optional CPE proxy will be integrated in a CPE such as a home or office router. This setup can be used in a residential or enterprise setting, and when it is in place, all devices connecting to the

router will receive a faster and more reliable internet connection. Traffic steering can also be applied at the CPE proxy level to control the traffic in the uplink direction.

Ericsson is partnering with CPE vendors and chipset manufacturers such as Intel to ensure efficient implementation of the Multipath TCP CPE proxy. We also offer a reference design and a test lab environment for CPE vendors.

Carrier-grade Multipath TCP proxy implementation

One important requirement for a Multipath TCP proxy in the service provider network is the ability to support a high-performance, carrier-grade IP solution for traffic aggregation. *Figure 8* illustrates how Ericsson's solution can be used as a Multipath TCP network proxy, which can be deployed in either a virtualized or non-virtualized environment.

All components – including Multipath TCP functionality – are implemented in user space [7] to meet the capacity requirements. The TCP traffic can be accessed directly from hardware using a Data Plane Development Kit (DPDK) [8]. The packet distribution function is responsible for sending traffic to the Multipath TCP protocol stack, located in the user space on one or several central processing unit (CPU) cores.

The Ericsson solution implements Multipath TCP functionality as specified by the IETF [2], combined with a specifically designed TCP CCA called TCP RNA (Radio Network Aware). TCP RNA is designed to utilize the mobile ran in an optimal way, and solves the equations for the correct congestion window by using measurements of the speed of the arriving TCP ACKs in conjunction with reactions of lost TCP segments. The benefits of TCP RNA are:

- » maximum utilization of available bandwidth for both uplink and downlink
- » reduced retransmissions using traffic shaping
- » controllable latency
- » avoiding bufferbloat.

This solution is highly configurable and can be tailored to support multiple Multipath TCP use cases per access network. The traffic-steering settings are policy driven.

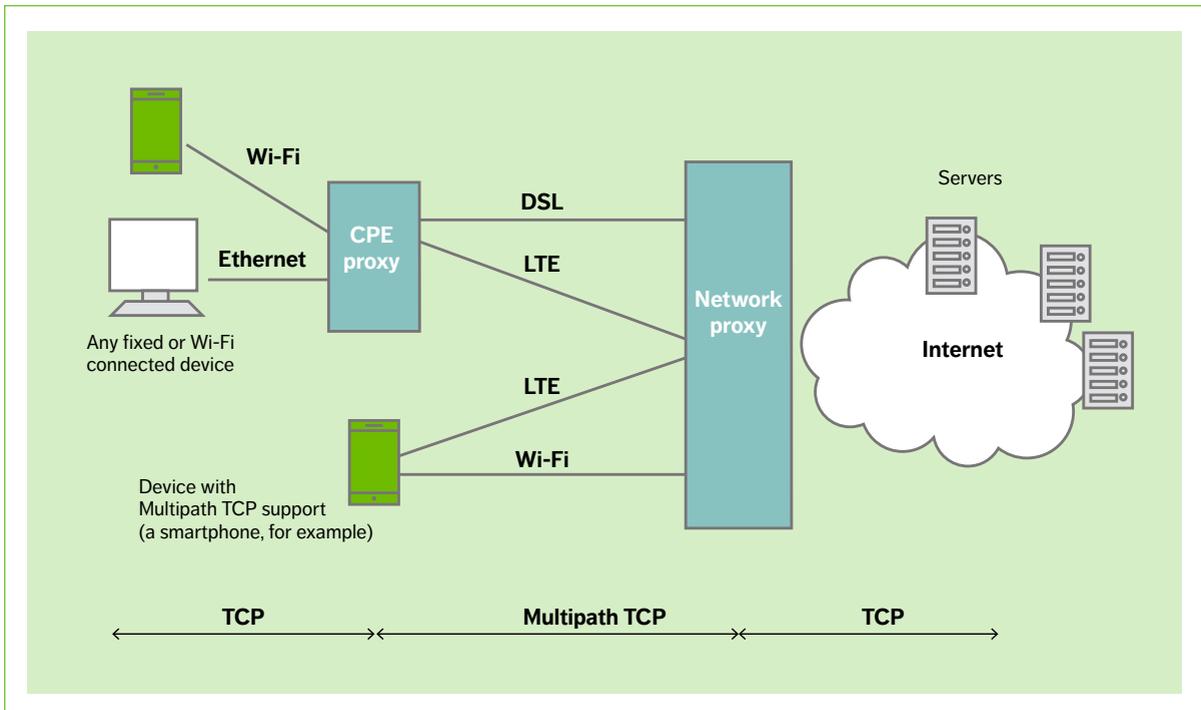


Figure 7 Proxy-based approach for Multipath TCP access aggregation

One configuration example is to send Multipath TCP traffic on one preferred subflow, such as the DSL link. When the DSL link has reached its limit, any surplus Multipath TCP traffic will be sent on another subflow – most commonly the LTE link.

Another configuration example aims to optimize radio usage on a system-wide level. If Multipath TCP traffic is sharing radio spectrum with other non-Multipath TCP traffic – from LTE-only mobile phones, for example – it might be preferable to avoid excessive use of the LTE link from Multipath TCP traffic. This can be achieved by configuring the TCP RNA for the LTE link to behave like background delivery. The result is that Multipath TCP traffic will back off when TCP RNA detects that the cell is congested, in favor of LTE-only traffic.

At times, it might be desirable to configure

Multipath TCP for maximum throughput – when combining LTE with Wi-Fi access for fast file download, for example. In such a scenario, the solution can be configured to use round-trip-time-based (RTT-based) traffic steering. Such traffic steering is achieved by sending data over the subflow with the lowest RTT. If that link reaches its capacity limit and there is more data to send, the rest of the data is sent over the other subflow. If one subflow can handle all the data, only the link with the lowest RTT will be used.

Conclusions

Access aggregation is a viable option for service providers to boost bandwidth across the last mile in areas where it is too costly to increase the capacity of legacy access. Typical access aggregation scenarios are the combination of DSL with LTE or the combination

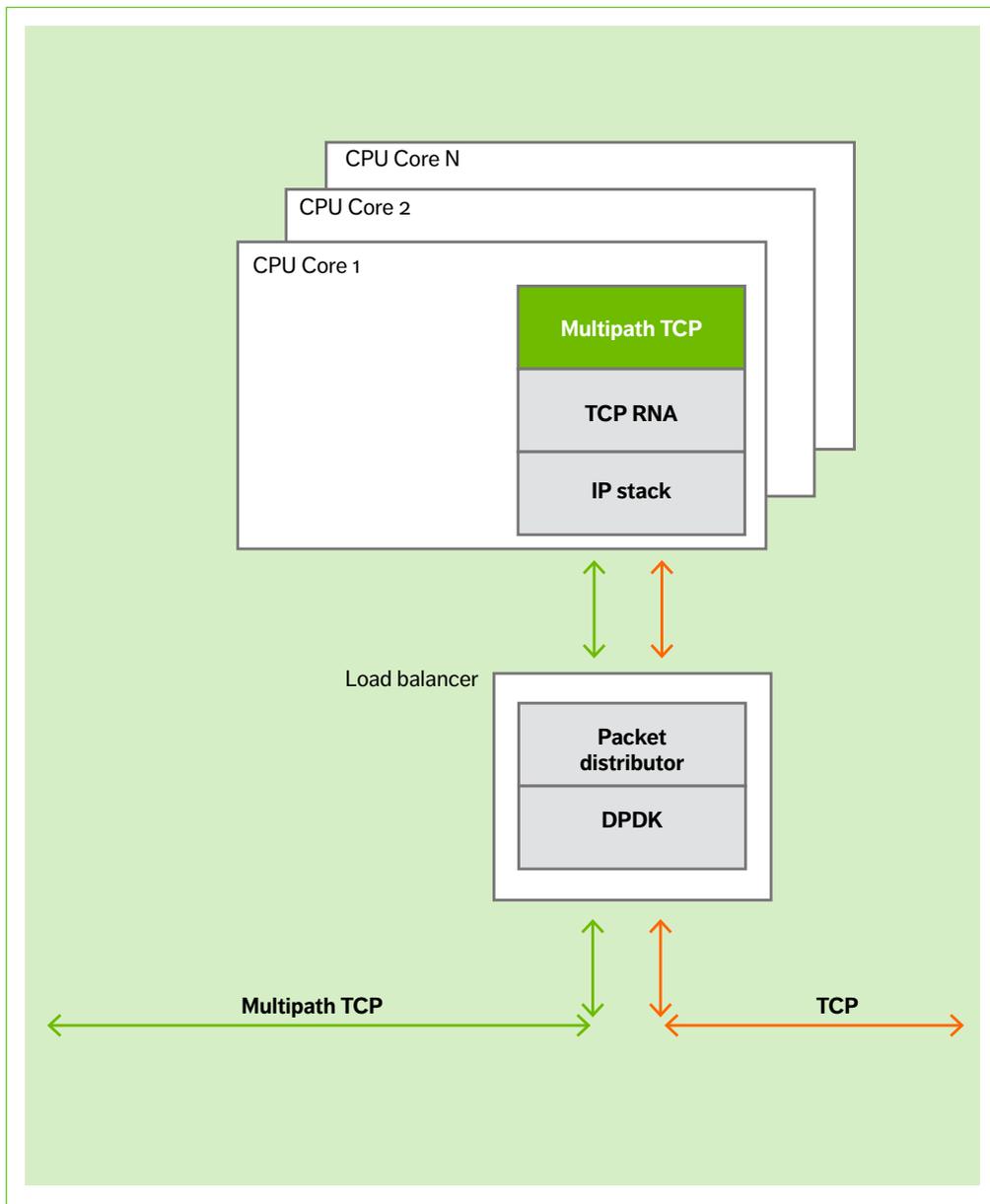


Figure 8 Multipath TCP network proxy

of LTE with Wi-Fi. Multipath TCP, as specified by the IETF, is ideal for access aggregation in the last mile, as it is able to boost bandwidth significantly, while simultaneously increasing reliability and ensuring seamless connectivity.

Multipath TCP comes as a set of extensions to standard TCP. It leverages all of the benefits of TCP such as fairness, flow control and reliability, as well as allowing the use of multiple paths through a network simultaneously. Multipath TCP proxies allow service

providers to use Multipath TCP for access aggregation without the need for end devices and internet servers to be aware of it.

Ericsson has created a Multipath TCP proxy that is tailored to the specific needs of service providers. It is carrier-grade, optimized for high traffic throughput and allows service providers to implement traffic-steering policies for the use of available access networks in the most cost-effective and efficient way. *

References:

1. 3GPP TS 23.402, Architecture enhancements for non-3GPP accesses, available at: www.3gpp.org/DynaReport/23402.htm
2. IETF RFC 6824, TCP Extensions for Multipath Operation with Multiple Addresses, available at: <https://tools.ietf.org/html/rfc6824>
3. IETF RFC 793, Transmission Control Protocol, available at: <https://tools.ietf.org/html/rfc793>
4. Linux Kernel Multipath TCP Project, available at: <http://www.multipath-tcp.org/>
5. IETF RFC 6356, Coupled Congestion Control for Multipath Transport Protocols, available at: <https://tools.ietf.org/html/rfc6356>
6. Broadband Forum, Hybrid Access Broadband Network Architecture (TR-348), available at: <https://www.broadband-forum.org/technical/download/TR-348.pdf>
7. Jonathan Corbet, Alessandro Rubini, Greg Kroah-Hartman, Linux Device Drivers, 3rd Edition. Nutshell Handbooks, 2005.
8. DPDK – Data Plane Development Kit, available at: https://en.wikipedia.org/wiki/Data_Plane_Development_Kit

THE AUTHORS

Robert Skog

◆ is a senior expert in the field of media delivery. After earning an M.Sc. in electrical engineering from KTH Royal Institute of Technology in Stockholm in 1989, he joined Ericsson's two-year trainee program for system engineers. Since



then, he has mainly worked in the service layer and media delivery areas, with everything from the first WAP solutions to today's advanced media delivery solutions. In 2005, Skog won Ericsson's prestigious Inventor of the Year Award.

Dinand Roeland

◆ joined Ericsson in 2000 as a systems manager for core network products. At Ericsson Research since 2007, he is currently a senior specialist in core network architectures and features. He has been a key contributor to the



standardization of multi-access support in the 3GPP EPC architecture, especially in Wi-Fi. Roeland holds an M.Sc. cum laude in computer architecture from the University of Groningen, the Netherlands.

Jaume Rius i Riu

◆ joined Ericsson in 2004 and has been principal



researcher in connectivity architectures at Ericsson Research since 2014. His work focuses mainly on the standardization of fixed-mobile convergence networking technologies (hybrid access, mobile transport, converged policy control, IPv6,

network controlled Wi-Fi and so on) and proof of concept development. Rius i Riu holds an M.Sc. in physics from UAB Autonomic University of Barcelona and a Ph.D. in experimental physics from KTH Royal Institute of Technology in Stockholm.

Uwe Horn

◆ is a solutions director within Ericsson's Global Customer Unit Vodafone. He has worked in the



telecommunications industry for more than 15 years and held various positions in R&D, consulting, marketing and sales. For the past 10 years, he has worked closely with Tier-1 service providers to develop new solutions based on the latest technologies. Horn holds a Ph.D. in telecommunication engineering from Friedrich-Alexander-Universität Erlangen-Nürnberg in

Germany and a diploma in computer science from the University of Bonn, Germany.

Michael Eriksson

◆ is a senior researcher at Ericsson Research. During most of his more than 20



years with Ericsson, his research has focused on the areas of computer science and networking. His current focus is on the design and implementation of advanced networking prototypes. Eriksson holds an M.Sc. in electrical engineering from KTH Royal Institute of Technology in Stockholm.



ISSN 0014-0171
284 23-3292 | Uen

© Ericsson AB 2016
Ericsson
SE-164 83 Stockholm, Sweden
Phone: +46 10 719 0000