FRONTHAUL AND BACKHAUL REQUIREMENTS OF FLEXIBLY CENTRALIZED RADIO ACCESS NETWORKS

JENS BARTELT, PETER ROST, DIRK WÜBBEN, JOHANNES LESSMANN, BRUNO MELIS, AND GERHARD FETTWEIS

Jens Bartelt and Gerhard Fettweis are with Technische Universität Dresden.

Peter Rost is with Nokia Networks.

Dirk Wübben is with the University of Bremen.

Johannes Lessmann is with NEC Laboratories Europe.

Bruno Melis is with Telecom Italia.

ABSTRACT

Cloud radio access networks promise considerable benefits compared to decentralized network architectures, but they also put challenging requirements on the fronthaul and backhaul network. Flexible centralization can relax these requirements by adaptively assigning different parts of the processing chain to either the centralized baseband processors or the base stations based on the load situation, user scenario, and availability of fronthaul links. In this article, we provide a comprehensive overview of different functional split options and analyze their specific requirements. We compare these requirements to available fronthaul technologies, and discuss the convergence of fronthaul and backhaul technologies. By evaluating the aggregated fronthaul traffic, we show the benefits of flexible centralization and give guidelines on how to set up the fronthaul network to avoid over- or under-dimensioning.

INTRODUCTION

Currently, LTE-based radio access networks (RANs) are implemented using two fundamentally different architectures, that is, a decentralized architecture where base stations (BSs) perform all RAN processing, and a centralized architecture where RAN processing is performed by central baseband units (BBUs). The latter is often identified as centralized RAN (C-RAN). In the distributed case, BSs are connected through backhaul (BH) links to the core network as well as to each other; in the centralized case, remote radio heads (RRHs) only convert the analog signal to the digital domain and forward these I/Q samples to central BBUs. The network connection between RRHs and BBUs is referred to as fronthaul as it imposes much more demanding latency and throughput requirements [1]. Both describe extreme cases of a mobile communication implementation; that is, in the case of a distributed implementation, BH carries the minimum necessary traffic toward the core-network, while in the centralized case FH entails the maximum traffic toward the central processing unit. Furthermore, full centralization facilitates cooperation between BSs, potentially reduces the operational costs of access points, and may exploit advanced data center technologies such as cloud-computing technology [2].

In future deployments, the transport network composed of BH and FH becomes a major cost component. As the number of BSs increases, the cost per BS decreases as they become smaller and transmit at lower power. Hence, the relative cost for BH and FH compared to each BS increases. For instance, deploying 1 m of optical fiber implies costs of up to $100 (in urban environments), while a microwave link with a range of a few tens of kilometers may be on the order of a few thousand dollars [3]. By contrast, a small-cell BS implies costs on the order of $5000–$10,000. In a fully centralized mobile network, the need for FH with its strong requirements in terms of capacity and latency would lead to even more expensive deployments. One option to reduce the overall network costs is to jointly optimize costs in the RAN and the FH network [4], and to utilize a heterogeneous set of FH technologies [5]. Furthermore, compression techniques have been proposed (e.g., in [6]) in order to reduce the required FH data rates. In this article we discuss another promising option: the flexible split ([7, 8]) of functionality between the BS and the central BBU, which reduces the FH data rate and yields additional gains through statistical multiplexing. Such flexibility, together with easier upgradability, is an important goal for 5G mobile networks, for both the radio access and the transport network. For FH and BH networks, this implies that flexible transport protocols must be deployed that are able to evolve with new RAN technologies, that BH and FH technologies need to converge to a unified, flexible transport network, and that automated control and management by means of software defined networking (SDN) must be supported. This includes the support of different functional splits of the RAN processing, also within the same network [2]. Each split may imply

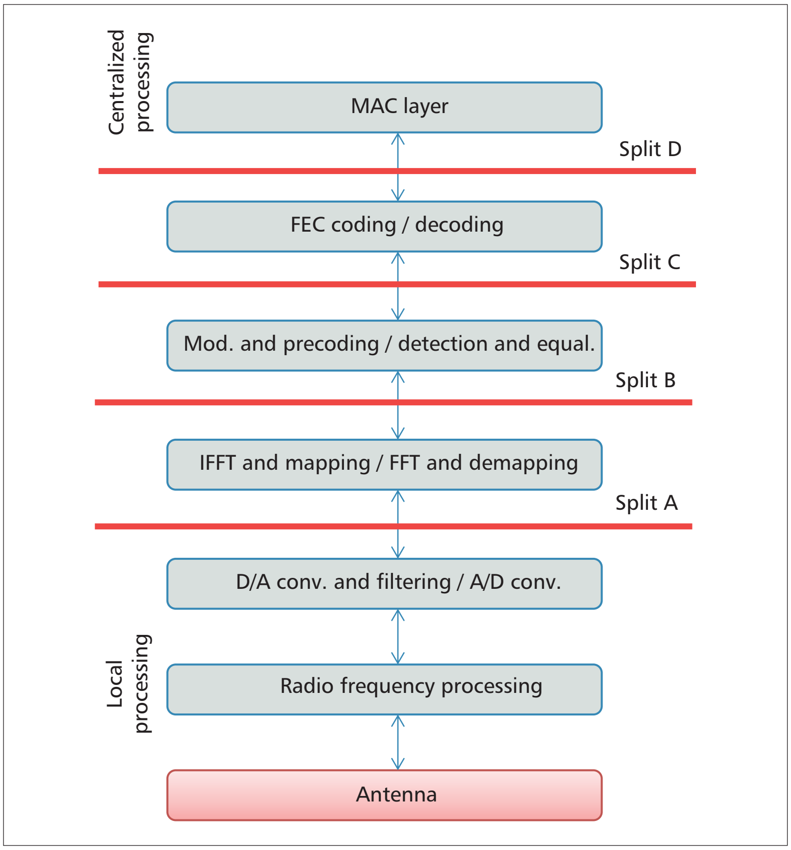

Figure 1. Functional split options.

different quality of service (QoS) requirements in terms of latency, throughput, and packet loss rate. This introduces a new level of complexity that must be handled by the transport network. In the following, we use FH in order to refer to the connectivity between a radio access point and a central processing unit if some or all of the RAN functionality is centralized.

In this article, we provide a comprehensive study of flexible centralized RAN and its impact on a converged FH and BH network in terms of required throughput and latency as well as protocol and infrastructure requirements. In particular, we derive and discuss per-user data rate requirements depending on different RAN functional splits, which extends previous works [7, 8] by including variable parameters such as the channel quality and traffic load. Using these per-user requirements, novel aggregated data rate requirements based on system-level results are discussed, and conclusive results on the feasibility of functional split under FH constraints are presented. We discuss the resulting transport network requirements to support flexible functional splits, compare them to available technologies, and describe the necessary convergence of FH and BH technologies.

Partially Centralized Radio Access Networks

The currently considered fully centralized and fully decentralized architectures are two extreme concepts, both with disadvantages. A decentralized network requires relatively low BH capacity but does not allow for joint processing, which may significantly improve network efficiency. A centralized RAN enables joint processing techniques such as multi-user detection and coordinated multi-point transmission, but it implies much more stringent requirements on the FH; that is, FH links usually require multi-gigabit-per-second capacity and latency on the order of a few microseconds. Furthermore, FH links are not compatible with existing transport network technologies, which increases capital expenditures (CAPEX) significantly. Hence, there is a need for technologies that offer a centralization gain while relaxing the requirements on the transport network. This can be achieved by splitting the RAN functionality into two parts, one executed locally at the BS, and one executed at a central processing unit. Depending on the chosen split, the FH requirements are reduced, and a different degree of centralization gain is achieved. There are many factors that determine the required data rate for each split, and there are numerous options to split the RAN processing. In this article, we review four major and representative options that show how the required FH capacity scales with the chosen functional split. Figure 1 illustrates the RAN signal processing chain of a Long Term Evolution (LTE) system as well as the four selected functional splits discussed in the following sections.

Split A

This option corresponds to full centralization, or the C-RAN approach. The Common Public Radio Interface (CPRI) [9] is the commonly used standard to exchange signals between the BSs and the central processing unit for this option. In the uplink, the received signals are down-converted to baseband and converted to the digital domain. As the complete baseband signal is forwarded, the required FH rate is static for a given system setup, that is, for fixed system bandwidth, number of receive antennas per sector, number of carriers, and number of sectors per site. Furthermore, the required bit rate scales linearly with the bit resolution of the analog-to-digital (A/D) conversion (and vice versa), which is usually chosen to be 15 b/symbol because of the high peak-to-average power ratio of the time domain signal, and to ensure precise channel measurements. The main drawback of Split A is the independence of FH data rate and actual user traffic; that is, even BSs that currently serve no user will require the full FH capacity.

Obviously, this split has no restrictions regarding the type of centralized processing that can be performed. Furthermore, apart from digital filtering and the FH protocol, no local processing in the BS is required. On the other hand, time and frequency synchronization as well as power control is performed centrally, which implies strong latency requirements for the FH. CPRI defines a round-trip time (RTT) of 5 μs plus propagation delay [9] on the FH, which sums up to a few hundred microseconds at typical distances of a few tens of kilometers.

Split B

In the case of Split B, the received uplink signal

Split	Quant. resolution	Number of antennas	No. spatial layers	Bandwidth (FFT size)	Occupied resources	Modulation scheme	Code rate	Required FH rate
A	X	X		X				1.23 Gb/s
B	X	X			X			311 Mb/s
C	X		X		X	X		155 Mb/s
D	X		X		X	X	X	44 Mb/s

Table 1. Exemplary FH rates and dependence of functional splits on LTE system parameters (for 10 MHz LTE, normal CP, 33 kbit transport block, 64-QAM, two antennas, code rate 0.85, one spatial layer, and 25 percent FH overhead).

is transformed to the frequency domain so that it is possible to extract and separately process the various physical channels. Therefore, unutilized resources are not forwarded to the central processing unit, so the required FH capacity scales with the actually occupied physical resources, and guard carriers and cyclic prefix (CP) can be omitted. In the case that synchronization and equalization is performed at the BS, even reference symbols and synchronization signals can be omitted. In contrast to Split A, the required FH data rate scales with the actual data traffic, which increases with the radio access channel quality. This allows for exploiting the statistical multiplexing gain on the basis of occupied physical resources, as discussed later. Furthermore, as analyzed in [8], the frequency domain signals also require a lower bit resolution; that is, the amount of quantization bits can be reduced to 7–9 b/symbol (depending on uplink and downlink).

Similar to Split A, this functional split implies almost no restrictions on realizing centralization gains; user processing can be performed jointly at the central processing unit in order to enable advanced cooperative algorithms. A limitation for this split in terms of latency can originate from the channel coherence time (in downlink) and the hybrid automatic repeat request (HARQ) process (in uplink). Equalization and precoding are performed centrally, and therefore assume a high correlation of actual and estimated channel state that cannot be achieved in the case of high-latency FH. However, for slow-moving users, the channel coherence time will be large such that the latency constraints of the subsequent splits will be the critical ones. Furthermore, the HARQ process requires decoding acknowledgments within 4 ms, which is a tight requirement, but can be addressed using opportunistic HARQ [10].

Split C

In this split, equalization and demodulation/precoding and modulation are performed locally; forward error correction (FEC) decoding/coding as well as all functionality at higher layers is performed centrally. This implies that joint detection techniques cannot be implemented, while joint decoding is still feasible. In the uplink, the equalization and demodulation leads to soft information values with maximum 3 bits per coded bit, that is, 6 bits for quadrature phase shift keying (QPSK), 12 bits for 16-quadrature amplitude modulation (QAM), and 18 bits for 64-QAM. The input of the modulator in downlink is also tied to the modulation scheme as the coded bits per symbol need to be forwarded, that is, 2 bits for QPSK, 4 bits for 16-QAM and 6 bits for 64-QAM. Furthermore, multiple-input multiple-output (MIMO) processing is performed in the BS so that the spatial layers are mapped to antennas. Link adaptation and MIMO scheme selection are performed depending on the channel quality; hence, the required FH rate depends on the signal-to-interference-plus-noise ratio (SINR) and channel rank experienced by each user. The FH latency requirement is dominated by the link adaptation in the downlink and the HARQ process in the uplink. Link adaptation may be critical if the channel coherence time is lower than the FH latency. This would result in a higher probability of block errors and a violation of QoS constraints.

Split D

In the case of Split D, all physical layer processing is performed locally such that only layer 2 and 3 functionality is centralized. Hence, the main centralization gain follows from joint scheduling, interference coordination, and path management techniques, which may still have a significant impact on network performance. The required FH capacity is closely tied to the actual user throughput determined by the user channel quality. Hence, the main bottleneck is represented by the FH latency, which may not exceed a few tens of milliseconds. Otherwise, scheduling and link adaptation may perform suboptimally, leading to performance degradation.

Numerical Example

To illustrate how the different splits impact the required FH data rates, Table 1 gives an example based on the uplink of a 10 MHz LTE system. In the table, we also summarize the dependencies between the required data rate and basic system parameters.

Backhaul/Fronthaul Convergence

The functional splits discussed in the previous section represent classical FH links. BH links would imply a split on even higher layers, such as decentralized medium access control (MAC). As such, the requirements on the BH technology are comparable to or slightly more relaxed than those of Split D. However, it is highly desirable

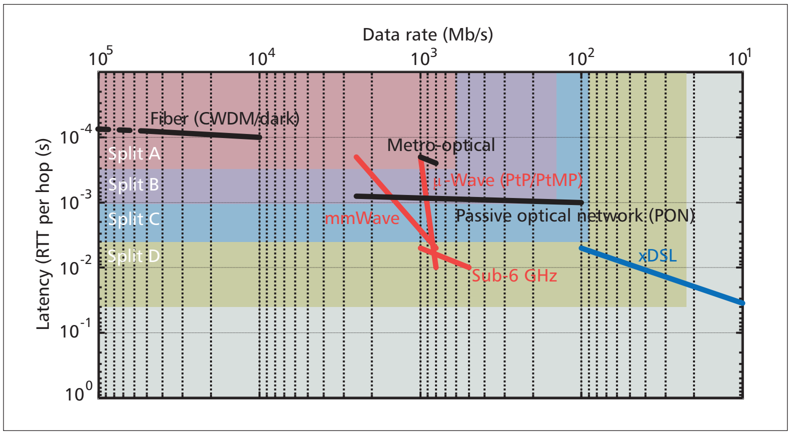

Figure 2. Latency and data rate requirements for different functional splits vs. available FH technologies.

that future transport networks are capable of supporting legacy technologies (full de/centralization) as well as novel implementation (partial centralization). The central processing unit could host, at the same time, core functionalities (e.g., gateways) serving classical decentralized BSs using BH links, as well as baseband functionality to serve (partially) centralized BSs via FH links. Hence, a convergence of FH and BH becomes necessary to avoid the need to deploy two separate transport networks. Figure 2 gives an overview of the latency and rate properties of available transport technologies that are able to support the previously introduced functional splits, and in the following, we discuss FH/BH convergence aspects across several network layers.

Layer 1 Technologies

Let us first consider physical layer technologies. Optical fiber is an intuitive solution for a converged FH and BH transport network as it provides future-proof capacity and latency. However, economics will forbid fiber to be deployed to every cell site, particularly in the case of small cells. While some operators may own a well developed fiber network, other operators may have to rely on leasing fiber or wavelengths from third parties, which is usually very expensive.

When fiber is not available or too costly, microwave connections become an alternative. Current state-of-the-art microwave solutions provide about 1 Gb/s capacity, which is not enough for lower-layer RAN centralization (Splits A and B). However, according to the roadmaps of most vendors, next-generation microwave technology will support multi-gigabit-per-second capacities, which would at least suffice for the last mile.

Another option is millimeter-wave radio technology. Depending on the desired channel bandwidth, such radios easily provide multi-gigabit-per-second speeds even today, and can be expected to scale to more than 10 Gb/s if needed. Multiple vendors offer these E-Band (60–90 GHz) devices already with built-in CPRI support. As such, they act as pure layer 1 pipes with per-device latencies as low as 10 ns, which allows chaining them back-to-back to enable multi-kilometer CPRI transmission. As layer 1 devices, they are oblivious to any layer 2 payload, and therefore in principle suitable not only for full (CPRI-based) RAN centralization but also for flexible functional RAN splits. Depending on the type of split and the required total air interface capacity, E-Band radio can be an option for the last mile as well as the first level of aggregation.

Finally, free space optics (FSO) can be a viable technology candidate for future transport solutions. The advantages are comparable to E-Band radio, but the lack of flexible equipment and economies of scale has so far prevented more widespread adoption.

Layer 2 Data Plane

The support for a flexible RAN split has a disruptive impact on data plane. Currently, the two extremes of fully centralized and fully decentralized RAN come with two completely incompatible transport technologies, CPRI and Ethernet. In cases where C-RAN has already been deployed, these are two fully separated network segments in terms of hardware interfaces and, more importantly, frame formats.

To increase cost efficiency, but also reduce network management overhead and spare part storage, a converged transport solution should have a single unified data plane interface regardless of the functional split. Such an interface with a versatile but unified frame format is a key enabler of future transport network architectures. It allows for distribution of cloud nodes in the periphery of the network rather than being limited to constrained locations in central core parts of the topology. These cloud nodes (i.e., very small data centers) in turn will be integral parts of the network and can be used to host functions such as caches, in-network data processing, network address translation (NAT), deep packet inspection (DPI), virtual border gateway protocol (BGP), but, most importantly in this context, centralized RAN functions.

If each functional RAN split came with its own hardware interface, frame format, and set of transport protocols, the set of cloud nodes where RAN functions could be centralized would be completely dictated by the setup of the deployed transport network infrastructure. However, once a unified solution is available, RAN functions can be freely placed and migrated to meet desired RAN and service-level key performance indicators (KPIs) in the best and most cost-efficient way.

The challenge is therefore to design this unified data plane despite the diversity of information that needs to be exchanged depending on the functional RAN split. Ideally, a widely deployed low-cost technology such as Ethernet could be reused and extended for this purpose. Work toward CPRI over Ethernet (the most difficult mapping) has already started [11], but other functional splits need to be taken into consideration as well. While Ethernet-like framing of digitized radio signals is certainly possible, it remains to be seen whether Ethernet can cope with the strict synchronization and jitter requirements of CPRI.

Packet Switch Evolution

Clearly, the tight transport-level performance requirements imply changes to packet switches as well. Significant advances have been made in ultra-low-latency switching, pushing down layer 2 per-node switching delays from some 10 μs to about 100 ns. However, a major bottleneck for further reducing latency comes from packet queuing. Current packet switches have a probabilistic scheduling paradigm that works based on relative packet priorities, typically given by QoS class identifiers. The exact scheduling discipline is configurable, but will generally favor the delivery of high-priority packets over low-priority ones. In order to support FH requirements, such probabilistic approaches might not always be satisfactory. Hence, future packet switches may need to support a more deterministic model, that is, guarantees on upper bounds for latency, jitter and packet loss rate. This again needs to be complemented with novel queue management disciplines, proper resource management, and access control, which prevents switch queues from becoming overloaded.

In terms of QoS support, FH performance requirements will in fact dominate application-level QoS requirements in a (partly) centralized RAN architecture for most user applications. Even relatively delay-sensitive traffic such as voice over IP (VoIP) needs end-to-end latencies on the order of 100 ms. In contrast, all functional RAN Splits A–D require much tighter latencies of a few milliseconds. Although the VoIP latency budget cannot be entirely spent within the RAN, RAN requirements are the dominating latency challenge for the transport network. Hence, application-level QoS differentiation becomes much less important, at least from a latency and jitter perspective. It can still be important from an overall capacity dimensioning perspective, as QoS information can be used to selectively drop packets in periods of congestion.

Control and Management Plane

Finally, a control and management plane is needed that can adapt switch configurations such as scheduling and QoS policing disciplines to the changing demands. Demands do change over time because of varying traffic load or dynamic migrations of functions between different cloud nodes in the network. Note that "migration" is a general term here which does not necessarily imply that a specific RAN function instance is physically transferred to another cloud node. Rather, RAN state could be externalized, for example, to cloud storage systems, and then be accessed simply from a new function instance at another node.

Generally, the control and management planes need to enable joint optimization of traffic steering in the transport network and RAN function placement across cloud nodes. From a provisioning perspective, this will be done using potentially extended SDN [12] and network functions virtualization (NFV) [13] technologies such as OpenStack and OpenFlow. In this spirit, there will be a logically centralized management entity with global RAN, and transport state knowledge and intelligence.

Besides centralized control, there is also a need to revisit localized organization, administration, and management (OAM) functionality.

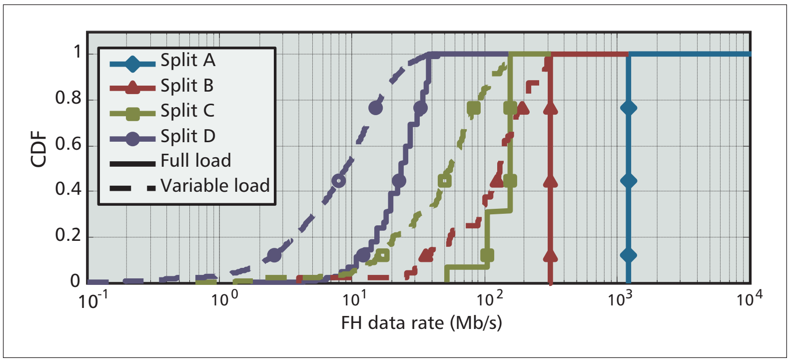

Figure 3. Cumulative distribution function (CDF) of required FH data rates of different functional splits considering full load in all BSs (solid lines) and considering a varying load (dashed lines).

For instance, in currently deployed BH networks, usually a failover latency of 50 ms after node or link failures is guaranteed. In the context of a partially centralized RAN, this latency is by far too high. It is open whether anything beyond 1+1 protection schemes is possible, where traffic is sent along a working and a protection path.

Quantitative Fronthaul Requirements

The high costs for BH and FH networks requires well dimensioned transport networks that are not over-dimensioned in order to minimize the deployment costs, and not under-dimensioned to avoid outage and satisfy user expectations. In order to provide a quantitative assessment of the required FH data rates, we utilize the four functional splits introduced earlier and combine them with system-level results that were obtained from a calibrated Third Generation Partnership Project (3GPP) LTE system-level simulator. In particular, these simulations assumed 500 m inter-site distance, a hexagonal urban macro setup with one small cell per macrocell, two antennas per small cell, and the urban macro propagation channel model. Based on this setup, we obtained the per-user SINR distribution. This distribution and the data rate requirements per functional split are used to derive the *aggregated* data rate requirements per functional split including the statistical multiplexing gain. We compare these requirements with the transport network technologies described earlier in order to assess the number of supported BSs for each functional split.

Figure 3 shows the distribution of required data rates for one BS and the four functional RAN splits. These data rates may vary quantitatively in different scenarios and depend on the user density, small-cell density, and propagation environment, among other factors. However, the qualitative comparison and conclusions remain unchanged. In addition, we differentiate two models: fully loaded buffer, which implies that at any time each resource is used, and variable load, which reflects a time-variant traffic profile adapted from the load distribution in [14]. Evidently, the data rate of Split A is constant and does not change with the actual traffic demand. Split B varies with the actual traffic load as

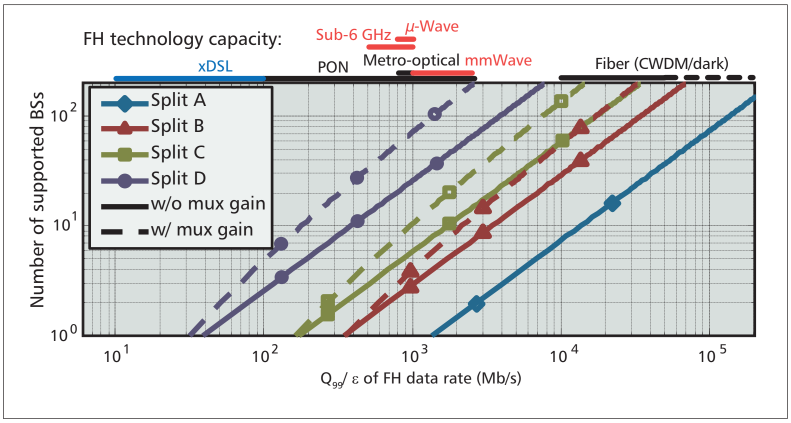

Figure 4. 99th percentile of the aggregated FH data rate (including safety factor) vs. the number of aggregated BSs.

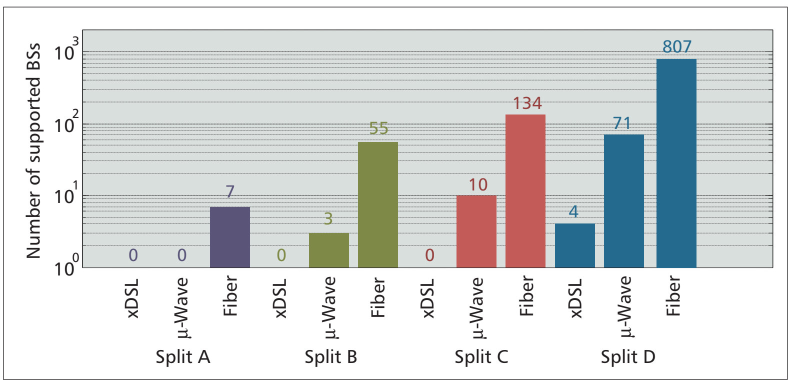

Figure 5. Number of BSs supported by different FH technologies: xDSL@100 Mb/s, μ-Wave@1 Gb/s, fiber@10 Gb/s.

reflected by the curve for time-variant traffic. Split C additionally varies with the three modulation schemes (4-QAM, 16-QAM, 64-QAM), while Split D also varies depending on the actual information rate due to the applied modulation and coding schemes (MCSs).

A widely accepted strategy for dimensioning the transport network capacity is to guarantee a certain outage percentile Q_x (per BS) and to divide it by a safety factor ε [15]. For example, Q_{99} implies a packet loss rate of less than 1 percent, and with $\varepsilon = 0.9$ it is guaranteed that the actual load does not exceed 90 percent of the transport network capacity with 99 percent probability. However, the FH traffic of multiple BSs is aggregated (e.g., at switches or routers) before it is processed at the central entity. At each aggregation point, we can exploit a statistical multiplexing gain, which implies that the peak demand of the aggregated traffic leaving the aggregation point is less than the sum of the peak demand of individual traffic arriving at the aggregation point. The multiplexing gain can be exploited in any scenario with non-static traffic and can be well determined using the central limit theorem. Figure 4 illustrates the scaling of the required FH data rates within the number of aggregated BSs. The multiplexing gain is the gap between the solid and dashed curves.

Furthermore, Fig. 4 compares the data rates to the physical layer FH technologies explained above. Based on this comparison, Fig. 5 shows the number of BSs that can be supported by three representative technologies and employing different functional splits. The number of supported BSs has a direct impact on deployment cost: the more BSs can be supported by a certain aggregation link, the fewer links need to be deployed.

Figure 5 further illustrates three important consequences for transport networks. First, partial centralization, particularly Split B and above, reduces the required network capacity and centralization overhead significantly. Second, higher functional splits allow the statistical multiplexing gain to be exploited, which has been up to a factor of three in our scenarios. Third, the multiplexing gain scales with the amount of traffic fluctuation and therefore increases with higher functional splits. For Split A, which is invariant in its data rate, there is no multiplexing gain to be achieved. In contrast, Split D, which minimizes overhead and exploits full traffic variance, achieves the highest multiplexing gain.

From these observations, important conclusions can be drawn for network operators: a flexible functional split can reduce the required FH capacity drastically. It should therefore be carefully chosen based on the type of centralized processing that actually yields benefits in a certain scenario. If the transport network is temporally overloaded, higher functional splits could be used as a fallback. This decreases the amount of centralized processing at the cost of lower user throughput, but avoids critical FH overload. Multiplexing of BSs at aggregation points yields another large reduction in required FH capacity and must not be neglected to avoid over-dimensioning the FH network. To exploit the multiplexing gain as far as possible, operators should aim to aggregate cells with a great variance of traffic, potentially combining areas with different daily traffic profiles (e.g., residential and commercial areas) and SINR distributions with sufficient variance. The flexible functional split can also increase the variance of FH traffic if it is adapted to the current network load.

These findings underline both the importance of a more flexible functional split as well as the benefits of the multiplexing gain in cloud-based networks.

Conclusions

This article provides a comprehensive overview of requirements for a converged FH/BH architecture of cloud radio access networks implementing a flexible functional split. The flexible centralization will considerably reduce the requirements on the FH network in both data rate and latency, enabling the utilization of technologies other than the currently favored but expensive fiber-only FH. To simplify deployments of such a heterogeneous and flexible architecture, a convergence of FH and BH must be considered. By adapting the functional split to the availability of the converged FH/BH network as well as actual user requirements, multiplexing gains can be exploited that further reduce the overall requirements. With this approach, future access networks could become truly virtualized, while keeping the deployment cost at a reasonable level.

Acknowledgment

The research leading to these results has received funding from the European Community's Seventh Framework Programme FP7/2007-2013 under grant agreement no. 317941 — project iJOIN, http://www.ict-ijoin.eu. The European Union and its agencies are not liable or otherwise responsible for the contents of this document; its content reflects the view of its authors only. We gratefully recognize the great contributions of many colleagues from iJOIN, who, in fruitful cooperation, contributed with valuable insight, surveys, and vision.

References

[1] P. Chanclou, *et al.*, "Optical Fiber Solution for Mobile Fronthaul to Achieve Cloud Radio Access Network," *Proc. Future Network and Mobile Summit*, Lisboa, Portugal, July 2013.
[2] P. Rost *et al.*, "Cloud Technologies for Flexible 5G Radio Access Networks," *IEEE Commun. Mag.*, vol. 52, no. 5, May 2014, pp. 68–76.
[3] T. Naveh, "Mobile Backhaul: Fiber vs. Microwave," Ceragon Networks Ltd., white paper, Oct. 2009; http://www.ceragon.com/images/Reasource_Center/White_Papers/Mobile_Backhaul_Fiber_Microwave-White_Paper.pdf. Accessed July 28, 2015.
[4] V. Suryaprakash, P. Rost, and G. Fettweis, "Are Heterogeneous Cloud-Based Radio Access Networks Cost Effective?" to be published, *IEEE JSAC*.
[5] J. Bartelt *et al.*, "Heterogeneous Backhaul for Cloud-Based Mobile Networks," *Proc. IEEE VTC-Fall 2013*, Las Vegas, NV, Sept. 2013.
[6] B. Guo *et al.*, "LTE/LTE-A Signal Compression on the CPRI Interface," *Bell Labs Tech. J.*, vol. 18, no. 2, Aug. 2013, pp. 117–33.
[7] D. Wübben *et al.*, "Benefits and Impact of Cloud Computing on 5G Signal Processing," *IEEE Signal Processing Mag.*, vol. 31, no. 6, Nov. 2014, pp. 35–44.
[8] U. Dötsch *et al.*, "Quantitative Analysis of Split Base Station Processing and Determination of Advantageous Architectures for LTE," *Bell Labs Tech. J.*, vol. 18, no. 1, May 2013, pp. 105–28.
[9] CPRI Spec. V6.0, 2013; http://www.cpri.info/; accessed July 28, 2015.
[10] P. Rost and A. Prasad, "Opportunistic Hybrid ARQ Enabler of Cloud-RAN over Non-Ideal Backhaul," *IEEE Wireless Commun. Lett.*, vol. 3, no. 5, July 2014, pp. 481–84.
[11] IEEE Std P1904.3, "Standard for Radio over Ethernet Encapsulations and Mappings," 2014; http://standards.ieee.org/develop/project/1904.3.html; accessed July 28, 2015.
[12] Open Networking Foundation, "Software-Defined Networking: The New Norm for Networks," white paper, Apr. 2012; https://www.opennetworking.org/images/stories/downloads/white-papers/wp-sdn-newnorm.pdf. Accessed Aug. 2013.
[13] ETSI Industry Specification Group for Network Functions Virtualisation; http://www.etsi.org/technologies-clusters/technologies/nfv; accessed July 28, 2015.
[14] H. Klessig, M. Gnzel, and G. Fettweis, "Increasing the Capacity of Large-Scale HetNets through Centralized Dynamic Data Offloading," *Proc. IEEE VTF-Fall 2014*, Vancouver, Canada, Sept. 2014.
[15] Y. d'Halluin, P. A. Forsyth, and K. R. Vetzal, "Wireless Network Capacity Management: A Real Options Approach," *Euro. J. Oper. Res.*, vol. 176, no. 1, 2007, pp. 584–609.

Biographies

Jens Bartelt (jens.bartelt@tu-dresden.de) received his Dipl.-Ing. (M.S.E.E.) from TU Dresden, Germany, in 2012. In 2011–2012 he worked as an intern for Rohde & Schwarz in Munich. Since 2013, he has been a research associate at the Vodafone Chair Mobile Communications Systems at TU Dresden, Germany, working toward his Ph.D. His research interests include cloud-based mobile networks, millimeter-wave communication, and physical-layer signal processing.

Peter Rost (peter.rost@nokia.com) received his Ph.D. degree from Technische Universität Dresden, Germany, in 2009 (supervised by Prof. G. Fettweis) and his M.Sc. degree from the University of Stuttgart, Germany, in 2005. Since May 2015, he has been a member of the Radio Systems research group at Nokia Networks, Munich, Germany, where he contributes to the European H2020 projects 5G-NORMA and METIS-II, and works in business unit projects on 5G architecture. He has been involved in several EU projects (e.g., FP7 iJOIN as Technical Manager) and standardization (e.g., 3GPP RAN2). Currently, he serves as a member of IEEE ComSoc GITC, IEEE Online GreenComm Steering Committee, and VDE ITG Expert Committee Information and Communication Theory. He is an Executive Editor of *IEEE Transactions on Wireless Communications*.

Dirk Wübben (wuebben@ant.uni-bremen.de) received a Dipl.-Ing. (FH) degree in electrical engineering from the University of Applied Science Münster, Germany, in 1998, and Dipl.-Ing. (Uni) and Dr.-Ing. degrees in electrical engineering from the University of Bremen, Germany, in 2000 and 2005, respectively. In 2001, he joined the Department of Communications Engineering, University of Bremen, Germany, where he is currently a senior researcher and lecturer. His research interests include wireless communications, signal processing, cooperative communication systems, and channel coding.

Johannes Lessmann (johannes.lessmann@neclab.eu) is a senior researcher at NEC Labs Europe, Germany, where he manages various R&D projects in the domain of mobile backhaul/fronthaul networks. He has also acted as Project Manager for operator collaborations and trials, Technical Manager of EU R&D projects, Product Manager for NEC's transport network optimization tool suite, and Chair of an IEEE 802.21 standards group on backhaul networks. Prior to NEC, he worked for Siemens on related topics. He is the recipient of multiple awards and a frequent speaker at key industry events on transport networks.

Bruno Melis (bruno1.melis@telecomitalia.it) graduated in electronic engineering from the Polytechnic of Turin in 1995. In the same year he joined Telecom Italia where at present he operates in the Wireless Innovation department. He has gained significant experience in system definition and design of signal processing algorithms for wireless communication systems including GSM, UMTS/HSDPA, and LTE. His research interests include wireless communications, physical layer signal processing, MIMO, and channel coding.

Gerhard P. Fettweis (gerhard.fettweis@vodafone-chair.com) earned his Dipl.-Ing. and Ph.D. degrees from Aachen University of Technology (RWTH), Germany. From 1990 to 1991, he was a visiting scientist at the IBM Almaden Research Center in San Jose, California, working on signal processing for disk drives. From 1991 to 1994, he was with TCSI Inc., Berkeley, California, responsible for signal processor developments. Since September 1994 he has held the Vodafone Chair at Technische Universität Dresden. In 2012, he received an honorary doctorate from Tampere University. He is also a well-known serial entrepreneur who has co-founded 11 startups to date. He has been an elected member of the IEEE Solid State Circuits Society's Board (Administrative Committee) since 1999, and he is also an elected member of the IEEE Fellow Committee. He coordinates two DFG centers at TU Dresden: Highly Adaptive Energy-Efficient Computing (HAEC) and the Center for Advancing Electronics Dresden (CfAED).

By adapting the functional split to the availability of the converged FH/BH network as well as actual user requirements, multiplexing gains can be exploited that further reduce the overall requirements. With this approach, future access networks could become truly virtualized, while keeping the deployment cost at a reasonable level.