Abstract

Within SAPHYRE three main steps to perform the evaluation of the proposed sharing solutions have been identified: reference topologies definition, performance metrics definition and description of the sharing models. At first physical layer sharing solutions have been generalized to a set of reference topologies. Next the proposed performance metrics were mapped onto system level KPIs. Finally, an overview of passive and active infrastructure sharing models is given.

Keywords

Backhaul sharing, gain evaluation, KPI, infrastructure sharing, performance metrics, QoS, relay sharing, scenarios, sharing models.
Executive Summary

The performance of the resource sharing schemes which will be developed within the SAPHYRE project has to be evaluated. The Deliverable D3.3a addresses this issue by providing an overview of the three main evaluation steps, we have identified: reference topologies definition, performance metrics definition and description of the sharing models.

In the first step the deliverable provides a detailed description of the four reference topologies, envisioned for SAPHYRE physical layer solutions. The first topology shows a typical interference channel model, which consists of nodes of different operators, sharing the same spectrum. The second topology enhances the scenario by collocating nodes, where for example joint coding, decoding or processing is possible. The last two topologies present scenario, which includes a relay node, thus enabling in the non-collocated case introduction of various different signal forwarding strategies, like: amplify-and-forward, decode-and-forward, among others. The relay collocated topology additionally allows for usage of joint signal processing techniques.

According to the second step of the proposed evaluation methodology, we propose a set of physical layer performance metrics, that could be used to describe the developed solutions. The representative metrics consist of single user rate, sum rate, outage probability, SINR or error rate. The metrics shall be used within the corresponding solutions to quantitatively describe the SAPHYRE gain. Where the SAPHYRE gain according to the deliverable can be understood as a system rate in the sharing scenario compared to the exclusive use of the spectrum and infrastructure by a single operator (typically TDMA) or as the fractional SAPHYRE gain which is rate of the user rate between the total utility received by users in the sharing scenario to the average utility received in single user scenarios. Furthermore we also provide a link between physical layer solutions and system level aspects such as Quality of Service. We achieve this through definition of Key Performance Indicators that are used to quantify operators Quality of Service levels. Binding the two aspects we propose a mapping between the performance metrics and Key Performance Indicators. Additionally we give a derivation of utility metrics to assess the users gain in the case of multiple antenna systems.

The last section of the deliverable deals with the description of the infrastructure sharing models. We give there a broad analysis of different aspects of the infrastructure sharing, starting from the top level division between active and passive sharing. Furthermore we expand and detail the division to five new categories: passive RAN sharing (sharing involves only passive elements such as site, masts), passive RAN
sharing with Access Transmission sharing (passive sharing with shared backhaul-
ing links), Active RAN sharing with MORAN (sharing of active resources with
static virtual resource division), Active RAN sharing MOCN and GWCN (sharing
of RAN active resources with dynamic resource assignment), and finally Roaming-
based sharing (full sharing based on inter-operator agreements). Among the listed
sharing paradigms, two elements of shared infrastructure are especially important,
as they might become system bottlenecks. The two elements are backhaul link
and relay node. Due to the availability of different technologies, we give a deep
analysis of the technologies behind backhaul link sharing. We put special attention
to the QoS provisioning problem, where we propose a solution based on IP QoS
service model and LAN virtualization using Ethernet Virtual Circuits. The analy-
sis of the relay node sharing presents initial assumptions on the interference relay
channel and the applicable forwarding scheme, such as amplify-and-forward, which
require the least signaling as well as minimum knowledge on the transmitted signal
(modulation, coding).

Eventually we conclude the deliverable providing a summary of research efforts
undertaken as well as highlighting the next steps to realize a complete vision of
performance measures for the resource sharing systems.
## Contents

Notation 7

1 Introduction 9

2 Description of Reference Topologies 13
   2.1 Motivation 13
   2.2 Scenario Classification 13
   2.3 Scenario Coordinates 13

3 Performance Metrics for Resource Sharing Schemes 17
   3.1 SAPHYRE Gain 17
   3.2 Definitions of Performance Metrics for Resource Sharing Schemes 18
   3.3 Quality of Service Aspects in SAPHYRE Solutions 21
   3.4 Performance Metrics in SAPHYRE Topologies/Solutions 25
   3.5 Utility Metrics Based on Multi-Antenna Channel Gains 25

4 Infrastructure sharing models 31
   4.1 Infrastructure sharing state-of-the-art 31
   4.2 Backhaul sharing model 38
   4.3 Relay sharing model 48

5 Conclusions 51

Bibliography 53
Notation

Mathematical Notation

C set of complex numbers
\( \mathbb{R} \) set of real numbers
Z set of integers

Some acronyms (please add your own acronyms here)

AGW access Gateway
AMC Adaptive Modulation and Coding
BER Bit Error Rate
CAPEX Capital Expenditures
CBS Commit Burst Size
CIR Commit Information Rate
CSI Channel State Information
DWDM Dense Wavelength Division Multiplex
EBS Excessive Burst Size
EIR Excessive Information Rate
eNodeB evolved NodeB
EPC Evolved Packet Core
EVC Ethernet Virtual Circuit
GTP GPRS Tunneling Protocol
IC Interference Channel
ICIC Inter-Cell Interference Coordination
IP Internet Protocol
IPSec Secured IP
IRC Interference Relay Channel
ITU-T International Telecommunications Union - Telecommunications
KPI Key Performance Indicator
KQI Key Quality Indicator
LAN Local Area Network
LMDS Local Multipoint Distribution System
LTE Long Term Evolution
MAC Media Access Control
MEF Metro Ethernet Forum
MIMO Multiple Input Multiple Output
MISO          Multiple Input Single Output
MME           Mobility Management Entity
MPLS-TP       Multi Protocol Label Switching - Transport Profile
NAS           Non-Access Stratum
NGMN          Next Generation Mobile Networks alliance
NMS           Network Management System
OAM           Operation, Administration and Maintenance
OPEX          Operational Expenditures
PBB-TE        Provider Backbone Bridging with Traffic Engineering
PHY           Physical Layer
PON           Passive Optical Networks
PWE3          PseudoWires
QCI           QoS Class Identifiers
QoS           Quality of Service
RAB           Radio Access Bearer
RAN           Radio Access Network
RED           Random Early Detection
RTT           Round Trip Time
SCTP          Stream Control Transmission Protocol
SDH           Synchronous Digital Hierarchy
SGW           Serving Gateway
SIMO          Single Input Multiple Output
SINR          Signal to Interference plus Noise Ratio
SLA           Service Level Agreement
TCO           Total Cost of Ownership
TCP           Transmission Control Protocol
TDM           Time Division Multiplex
TDMA          Time Division Multiple Access
UDP           User Datagram Protocol
VDSL          Very high bitrate Digital Subscriber Line
VLAN          Virtual LAN
WAN           Wide Area Network
1 Introduction

The vision of SAPHYRE is to demonstrate how the paradigm of exclusive resource allocation shifts towards cost, spectrum and energy efficient voluntary physical resource sharing which is realized through innovative use of radio spectrum and network infrastructure under economic and regulatory constraints. SAPHYRE realizes the vision by focusing on resource sharing aspects between wireless network operators, where by resource sharing one should understand passive or active pooling of available resources (spectrum, network elements, physical links or site hardware) for the joint purpose of cost savings, performance enhancement and overall greater efficiency. Thus, SAPHYRE intends to develop a number of techniques that enable introduction of resource sharing schemes that lead to increased utility. However, in order to ensure the feasibility of sharing schemes, in all objectivity, SAPHYRE solutions must be evaluated with common measures, which allow to emphasis the sharing gain.

The sharing gain, namely SAPHYRE gain, can be defined as the performance comparison in terms of various metrics (e.g. system sum-rate, achievable rate region). Each of the proposed resource sharing schemes shall clearly highlight the gain arising from implementation of the sharing scheme, in question. The metrics, which ought to show the SAPHYRE gain shall also be used to provide the notion of fairness and cover Quality of Service aspects. The notion of fairness between users is normally not explicitly built into the scheduling criterion, since the operator only seeks to maximize a capacity metric under possible QoS constraints. Therefore a new approach towards fairness in resource sharing schemes should be developed. Regardless from fairness aspects in the inter-operator domain still QoS levels must be maintained, therefore two way approach is necessary, where the QoS is divide into: guaranteed level (denotes minimum required service) and excessive level (denotes the maximum possible service, where notion of fairness is applied). It is worth pointing out that Quality of Service is especially important in the case of QoS-sensitive IP-based applications (e.g. VoIP), which are typically provided by 4G technologies and are envisioned also for future systems such as IMT-Advanced. On the business level QoS is ensured through contracts, called SLAs (Service Level Agreements), which oblige mutually operators to obey the policy rules. Typically policy rules describe the amount of network resources required to realize QoS services [33]. The quantitative measures for policy rules are given by KPIs (Key Performance Indicators). KPIs are set up to evaluate the system as well as monitor the current progress of the solutions in respect to set goals. In order to maintain the consistence bet However due to the general nature of KPIs, a mapping between link level and system level performance evaluation, KPIs need to be unambiguously
Figure 1.1: Two reference scenarios of SAPHYRE: a) spectrum sharing, b) infrastructure sharing.

Apart from KPIs, each of the proposed performance metrics need to be applied also to the specific reference topology, which would enable comparison and classification of the developed resource sharing schemes. In Chapter 2, the key topologies for sharing are presented. The enhancement in the sharing context, exposes novel physical layer processing techniques related to the sharing of a relay node. From the functional point of view reference topologies are used to provide an insight into the assumptions of the particular solution, for example duplex mode, element collocation, available demodulation and coding scheme.

Apart from the performance metrics and reference topologies typically available in the methodology, also infrastructure sharing models need to be provided. The infrastructure sharing models in SAPHYRE, can be divided into passive and active, which then can be further divided according to different technical solutions based on the type of the network element or site equipment that is being shared [17]. This creates a multidimensional sharing problem, therefore a detailed analysis of different degrees of infrastructure sharing will be given in the document. Furthermore the document will provide a deeper insight into the problems related to backhaul link sharing as well as relay node sharing, which will further be used in deliverable D3.1a to aid the development of resource sharing schemes.

The aim of this deliverable is to provide a general understanding of the assumptions underlying the performance evaluation and potential benefits of the physical layer techniques developed within SAPHYRE project. Therefore in the chapter ??, we start the document with unified reference topologies. In chapter 3 we propose a set
of performance metrics to describe SAPHYRE sharing solutions with enhancement of the metrics to system level and example detailed analysis of the utility achieved in multiple antenna scenario. Leading deeper with the discussions on sharing, in chapter 4, we provide analysis of infrastructure sharing models with a closer look into the backhaul link and relay node sharing.
2 Description of Reference Topologies

2.1 Motivation

Each WP, layer or viewpoint perspective (technical, business, etc.) requires a different level of abstraction when defining what is called by the term “scenario”. Different aspects of the “scenario” are defined here for classifying solutions in different contexts.

2.2 Scenario Classification

Here, a multi-dimensional “scenario coordinate map” is proposed.

- Each coordinate (dimension) introduces the classification relevant to a given layer/viewpoint. This creates required flexibility and avoids confusion around the term scenario later on.
- The scenario description is intentionally kept in a very generic form and defines only fundamental characteristics. All finer details (e.g. various quantitative parameters, finely defined subclasses) should be only specified as a parameter of that scenario.
- Note that we use the term “scenario” to express a classification of business related contexts. Technically-oriented such as WP2 and WP3 for instance will use the term e.g. “topology” in order to mark the difference.

2.3 Scenario Coordinates

2.3.1 Scenario (S) Coordinate Classes

The Scenario (S) coordinate describes operator’s viewpoint. This is the first (top level) coordinate.

Values:
- S1 = shared RAN & shared spectrum
- S2 = shared RAN only
- S3 = shared spectrum only
2 Description of Reference Topologies

The Topology (T) coordinate describes characteristics important mainly from the perspective of communication system design and algorithms. In particular it indicates

- The presence or absence of relays,
- The physical collocation of Base Station belonging to different operators.

The following key topologies are considered in SAPHYRE (see Fig. 2.1):

- TA = no relay, no base station collocation, spectrum sharing (interference channel)
- TB = no relay, base station collocation (with perhaps joint modulation/coding, joint demodulation/decoding at terminal)
- TC = relay present and shared, no base station collocation (all possible relaying strategies A&F, Joint D&F, C&F, HD&F, etc)
- TD = shared relay, share base station (all relaying strategies with joint modulation/coding, joint demodulation/decoding at terminal)

Notes:
- Presence of various link types (bidirectional/unidirectional, presence of direct link, presence of side-information link, etc.) is only reflected through the parameters of the topology. The same holds for all other attributes like SNR, channel type, synchronization assumptions, etc.

2.3.3 Other Coordinate Classes

A number of other classes, not relevant to WP3, are defined. Their particular description is in the respective WPs. Namely, there are

Notes:
- RAN sharing refers to sharing whatever type of HW (including e.g. the relay)
2.3 Scenario Coordinates

- Business (B) coordinate classes
- Regulatory (R) coordinate classes
3 Performance Metrics for Resource Sharing Schemes

3.1 SAPHYRE Gain

As was mentioned in D3.1a, in order to evaluate the benefits of novel signal processing algorithms to exploit the additional degrees of freedom brought by sharing in multi-user and multi-cellular environments, it is important to:

- define a performance metric,
- show the gain (loss) with respect to the chosen performance metric as compared to a non-sharing scenario,
- point out conditions when a significant gain can be achieved for the chosen scenario (topology), and
- illustrate the order of magnitude of this gain.

We define this sharing gain as the SAPHYRE gain. Formally it can be defined as the performance comparison in terms of various performance metrics (e.g., the system sum-rate, the achievable rate region, etc.). In this deliverable, we define two types of SAPHYRE gains in terms of system rate of the sharing scenario compared to the exclusive use of the spectrum and infrastructure by a single operator (time division case, in this case, the users are multiplexed via TDMA). The absolute SAPHYRE gain is defined as

$$\Xi_A = \sum_{k=1}^{K} U_k - \frac{1}{K} \sum_{k=1}^{K} U_{k}^{SU},$$

and the fractional SAPHYRE gain is defined as

$$\Xi_F = \frac{\sum_{k=1}^{K} U_k}{\frac{1}{K} \sum_{k=1}^{K} U_{k}^{SU}},$$

where $k \in \{1, 2, \cdots, K\}$ is the index of the users. The utility function of the $k$th user in the sharing scenario and the time division case are denoted by $U_k$ and $U_{k}^{SU}$, respectively.
3.1.1 Fairness Performance

An important quality of service element in wireless networks is the ability to provide a satisfactory communication performance to the user, regardless of where it is located, or at least doing so with a high probability level such as 90% of the time and coverage area. Since spectrum sharing results in interference, it is expected that users located in areas of strong overlap between the sharing operator’s coverage zones will suffer more from the sharing, compared with users which remain in areas dominated by a single provider and are naturally isolated from sources of interference (thanks to terrain, distance-based path loss, etc.).

As a consequence, it is important for sharing schemes to be evaluated not just in terms of the gain on the average data rates or spectrum efficiency but also in terms of the fairness level that can be maintained between the mobiles users, regardless of which provider they belong to. There is not a uniquely accepted definition of fairness but several options are available to measure this dimension of the system performance. Two key examples are:

- The Jain’s index: this index measures the discrepancy between the mean and the variance across the rates (or utilities) achieved at the users and is expressed by:

\[ J(U_1, U_2, \ldots, U_K) = \frac{\left( \sum_k U_k \right)^2}{K \sum_k U_k^2} \]  

In particular the Jain’s index is one for perfect fairness (equality among positive utilities) and closer to zero for unfair distributions of the utilities.

- The ratio of the mean across all rates or utilities and the mean across utilities for the \( x \%) worst users, where \( x \) is left to be defined (e.g. 10% worst users).

3.2 Definitions of Performance Metrics for Resource Sharing Schemes

3.2.1 Definition of Physical Layer Metrics

Saphyre results can only be assessed with use of appropriate performance metrics. The assessment will be done by means of QoS realization, which in highly loaded packet-switched networks needs to be preserved by all means [12]. The following physical layer metrics will influence the resource allocation process (done in Radio Resource Management layer), in a way that they will reflect the possible QoS levels that can be provided to a single users for specific service.

In order to properly assess the results of Saphyre WP2 and WP3 solutions we need to define the set of performance (utility) metrics for resource sharing schemes.
3.2 Definitions of Performance Metrics for Resource Sharing Schemes

proposed in the physical layer:

- **Sum-rate** - maximum system achievable throughput regardless of fairness in resource allocation between the users. The maximum system achievable throughput is in fact a maximum sum of rates taken over all rate vectors in the rate region [18]:

\[
C_{SR} = \max_{(R_1,...,R_k) \in C} \sum_{k=1}^{K} R_k
\]

- **Quality-of-Service related metrics:**
  - *Single user rate* (most likely presented in bits/s or bits/s/Hz) - the user rate can be split into two values, one is a guaranteed theoretical rate and the other is a maximum theoretical rate. The guaranteed value represents a rate that can be offered and supported to a single user of one operator/technology in case of spectrum/infrastructure sharing scenarios, typically when operator/technology owns an exclusive band or its traffic has higher priority over others (this rate will not be violated by unfavorable sharing conditions - high load of other operator(s) traffic). The maximum rate represents theoretical maximum capacity that can be achieved by a single user in spectrum/infrastructure sharing scenario, the maximum rate can not be anyhow guaranteed and it is dependent on the current load situation. The relation between the values is described by inequality: Guaranteed user rate \(<\) Maximum user rate

Single user rate can be calculated by means of either ergodic capacity or outage capacity. Example calculations of single user rate for time variant and invariant wireless channels with single and multiple antennas at the transmitter and receiver have been derived in [18].

- *Outage probability* - the probability that the target bit error rate performance of the users can not be met, that is due to the fact that the power at the receiver is below the minimum reception threshold (which can be seen as outage in transmission). The metric can be used to represent the combination of transmission/reception methods (MIMO, MISO), transmission schemes (Network Coding), level of noise and interference on the reception of the signal. Additionally it can be also used to evaluate the system deployment of relay nodes in the scenarios with relays. The typical models for outage probability under combined path loss and shadowing, and and fading channel are given below [18]:

\[
p_{out}(P_{min}, d) = p(P_r(d) \leq P_{min})
\]

\[
P_{out}(\gamma \leq \gamma_0) = \int_0^{\gamma_0} p_{\gamma_\alpha}(\gamma) d\gamma
\]
where \( P_{\text{min}} \) denotes minimum required power, \( P_r(d) \) received power at given distance from the transmitter (which is log-normally distributed), \( \gamma_s \) received SNR, which is a random variable with distribution \( p_{\gamma_s}(\gamma) \), \( \gamma_0 \) specifies the minimum SNR required for acceptable performance.

- **Service latency/delay** - represents the time needed to transmit the information over the radio interface between the transmitter and receiver (sometimes can be expressed also as in term of RTT), factors which influence latency: channel type, number and complexity of processing operations (both at transmitter and receiver), size of the information burst, etc. The latency constraint can be used to describe the feasibility of the scenario as in respect to specific services offered by operators, e.g. low latency is crucial for operators providing real-time applications such as video streaming, voice. In the most general form can be expressed as combination of various delays (measured in the units of seconds):

\[
\text{Latency} = \text{propagation\_delay} + \text{transmit\_time} + \text{processing\_time}
\]

(3.6)

- **SINR** (Signal-to-interference-plus-noise-ratio) - describes channel quality (especially crucial in multiuser systems where it describes interference level) and therefore directly influences the rate achievable by each user. The idea is to keep it as constant as possible so to maintain channel quality for the user and provide stable QoS value. The metric can be used as a good descriptor of the interference level (important factor in spectrum sharing) produced in the system and QoS provisioning, e.g. if SINR is highly variable it is unlikely to provide high QoS to the users. The most general model for SINR at the receiver (measured in dB) is given by:

\[
\text{SINR} = \frac{P}{N_0 + I}
\]

(3.7)

where \( P \) denotes average received power of the signal, \( N_0 \) average received noise power and \( I \) average received interference power.

- **Error rates** (BER) - empirical metrics defined as number of erroneous bits to the total number of received bits, which can be used to represent the reliability of the received data information. Factors which influence the error rates: noise, interference, distortion, fading, etc. Generally error rate can be improved by providing better SINR (stronger transmitted signal power), lower modulation schemes, etc. The metric can be used to describe the reliability of the transmission in respect to the modulation and coding schemes proposed, if the schemes are less prone to interference then the error rates shall be smaller.
3.3 Quality of Service Aspects in SAPHYRE Solutions

Typically operators when deploying new networks seek to maximize their available capacity under specific QoS constraints. In particular QoS is an important measure that allows the operators to differentiate from each other [20]. It is important to remember about it when designing different sharing schemes for SAPHYRE. In principle resource sharing solutions for cellular networks shall provide QoS awareness. This imposes the same requirement on the design of performance metrics which will evaluate the sharing solutions, as in ideal case they should show also the compromise between sharing and individual QoS achievements.

3.3.1 Mapping of Performance Metrics on System Level KPIs

Key Performance Indicators are primary metrics to define the success rate of an enterprise. In principle they are set up to evaluate the system as well as monitor

\[
BER = \frac{\text{Number of errors}}{\text{Total number of bits transmitted}} \tag{3.8}
\]

- **Amount of additional side-information:**
  - *exchanged outside PHY* (inter/intra operator) - depending on the sharing scenario, operators (or technologies in case intra-operator scenario) may need to exchange information regarding the interference conditions, used (free) time slots, power allocations, CSI (Channel State Information). The metric describes the signaling overhead required to distribute the information, as well as additional interfaces and network components.
  - *required inside PHY* - typically Side Information the amount of information on other transmitted information streams (send to other destinations) in multi-user system, that is used to decode own information stream. Can be split into either full or partial side information knowledge, regarding the coding technique used. The side information can either be sensed from the environment or it can be supplied via pilot channel (see also cognitive pilot channel) or cable link. The metric can be used to represent limitations to the planning of the network and additional links that need to be taken into account when deploying the scenario.
  - *overhead penalization* (total rate reduction) - used as a weighting factor the utility function, the metric describes single user rate/sum-rate reduction due to additional overhead related with the exchange of side information such as supplying the CSI to the transmitter.
the current progress of the solutions in respect to set goals. According to 3GPP, KPIs describe strategic goals of enterprise, and cascade through the entire organization. KPIs are specified through definition and measurement of key parameters of input/output of internal network system and/or maintenance & operation progress of an enterprise [10]. The term strictly connected with KPIs is SLA, which can be seen as a contract which describes common understanding of the service as well responsibilities between parties involved and performance objectives. In order to provide information on how the SLA agreements are realized, KPIs (indication of service resource performance) and Key Quality Indicators (service element performance) are measured and compared towards objective targets included in the SLA. KPIs are proved by aggregation of network performance data from network elements [10].

3GPP has proposed such classification of KPIs for 2G and 3G systems, with definitions from ITU-T recommendation [3]:

- **Serveability** - The ability of a service to be obtained (with specific tolerance and other conditions) when requested by the user and continue to be provided without excessive impairment for a requested duration.
  - **Accessibility** - The ability of a service to be obtained (with specific tolerances and other given conditions) when requested by the user.
  - **Retainability** - The ability of a service (once obtained) to continue to be provided under given conditions for a requested duration.
  - **Integrity** - The degree to which a service (once obtained) is provided without excessive impairments.

- **Availability** - The ability of an item to be in a state to perform a required function at a given instant of time or at any instant of time within a given time interval, assuming that the external resources, if required, are provided.
  - **Reliability** - The ability of an item to perform a required function under given conditions for a given time interval.
  - **Maintainability** - The ability of an item under stated conditions of use, to be retained in or restored to, a state in which it can perform a required function, when maintenance is performed under given conditions and using stated procedures and resources.
  - **Utilization** - Indication of the network resource utilization, such as throughput on specific interface.
  - **Mobility** - The description of abilities to perform handovers.

**Saphyre Key Performance Indicators**

Based on the 3GPP proposed KPIs and their classification, Saphyre project will use its own set of KPIs to describe the SAPHYRE gain. The proposed KPIs are used for business level modeling as well as scenario benchmarking:
3.3 Quality of Service Aspects in SAPHYRE Solutions

- **KPI:**
  - *Accessibility* - see the definition from [3].
  - *Retainability* - see the definition from [3].

- **KQI:**
  - *Capacity* - The resources that can be provided to perform the service, the resources can be represented as link capacity, number of users, etc.
  - *Coverage* - The area, which is serviced by the proposed technical solutions.
  - *Latency* - The introduced time constraints to the performed service.
  - *Fairness* - The distribution of the user rates around the mean. The fairness can be measured by the Jain’s index which characterizes the variance of the rates with respect to the mean computed across all users.

What is more also an important issue is impact of PHY features on CAPEX (Capital Expenditure) and OPEX (Operational Expenditure):

- **Implementation cost/profit** - The cost/profit generated from the introduction of specific solution. This cost/profit occurs one time only at the installation or implementation phase.

- **Operational cost/profit** - The cost/profit generated periodically (e.g. each month) due to the availability/unavailability of specific solution. This cost/profit occurs due to rental fees, maintenance fees, person months, etc.

<table>
<thead>
<tr>
<th>Performance metric</th>
<th>KPI/KQI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Accessibility</td>
</tr>
<tr>
<td>Sum-rate</td>
<td>X</td>
</tr>
<tr>
<td>Single user rate</td>
<td>X</td>
</tr>
<tr>
<td>Outage probability</td>
<td>-</td>
</tr>
<tr>
<td>Latency/delay</td>
<td>-</td>
</tr>
<tr>
<td>SINR</td>
<td>X</td>
</tr>
<tr>
<td>Error-rate</td>
<td>-</td>
</tr>
<tr>
<td>Inf exchanged in PHY</td>
<td>X</td>
</tr>
<tr>
<td>Inf exchanged out PHY</td>
<td>X</td>
</tr>
<tr>
<td>Overhead penalization</td>
<td>X</td>
</tr>
</tbody>
</table>

Table 3.1: Map of PHY metrics on to system level Performance and Quality Indicators.
3.3.2 Examples of QoS Provisioning Agreements in Resource Sharing Scenarios

QoS and SLA in Shared Backhaul. The shared backhaul link transports multiple operator flows, which can be further divided into flows corresponding to different types of services. It is thus obvious that a QoS delivery shall be maintained on two levels: per operator flow and per service flow. All the constraints regarding the minimum and excessive parameters available for each of the flows are described in Service Level Agreements. The most popular solution that support provisioning of SLAs at the transport links is to use IP QoS service model [13]. The IP QoS model (DiffServ model) has been standardized to provide a variety of quality classes for various services traversing an IP network. The standard defines different service classes, QoS provisioning mechanisms along with architecture that can be applied to different network elements. 3GPP has also recognized this method as a standardized solution for traffic classification for IP Radio Access Bearer services [9]. The mapping between 3GPP QoS classes and DiffServ code points shall be defined by the operators. The packet classification is used by aggregation nodes to perform link scheduling (in fact rate limitation and packet marking), with respect to predefined thresholds derived from SLA [6]:

- Committed Information Rate (CIR) - guaranteed minimum throughput, this value is dedicated (it can not overlap with the throughput of other flows/operators) to specific flow. In principle the highest the priority of the flow, the highest the committed value.

- Excessive Information Rate (EIR) - an excessive throughput, which can be defined as maximum link rate that the service can be assigned. High value of EIR is typically assigned to 'best effort' services. The excessive throughput is taken from the common pool of resource for all the flows/operators. In case of leased lines, EIR is usually connected with additional fees.

- Committed Burst Size (CBS) - maximum size of Ethernet frame burst expressed in bytes with guarantees on performance.

- Excessive Burst Size (EBS) - maximum size of Ethernet frame burst with no guarantees on performance.

All the above mentioned mechanisms and limitations constitute for backhaul QoS support tools, which are used at aggregation nodes to actually serve the desired rates of traffic.
### 3.4 Performance Metrics in SAPHYRE Topologies/Solutions

#### Table 3.2: Usage of performance metrics in SAPHYRE solutions.

<table>
<thead>
<tr>
<th>Performance metric</th>
<th>SAPHYRE topologies</th>
</tr>
</thead>
<tbody>
<tr>
<td>...</td>
<td>Non-shared</td>
</tr>
<tr>
<td>Sum-rate (instantaneous, ergodic)</td>
<td>X</td>
</tr>
<tr>
<td>Single user rate</td>
<td>X</td>
</tr>
<tr>
<td>Outage probability (probability, rate)</td>
<td>X</td>
</tr>
<tr>
<td>Latency/delay</td>
<td>-</td>
</tr>
<tr>
<td>SINR</td>
<td>X</td>
</tr>
<tr>
<td>Error-rate</td>
<td>X</td>
</tr>
<tr>
<td>Inf exchanged in PHY</td>
<td>-</td>
</tr>
<tr>
<td>Inf exchanged out PHY</td>
<td>-</td>
</tr>
<tr>
<td>Overhead penalization</td>
<td>-</td>
</tr>
</tbody>
</table>

#### 3.5 Utility Metrics Based on Multi-Antenna Channel Gains

We consider $T$ transmitters and $K$ receivers sharing the same spectral band. Define the set of transmitters as $\mathcal{T} := \{1, \ldots, T\}$ and receivers as $\mathcal{K} := \{1, \ldots, K\}$. Each transmitter sends useful information to at least one receiver. For transmitter $k, k \in \mathcal{T}$, let $\overline{\mathcal{K}}(k) \subseteq \mathcal{K}$ denote the set of its intended receivers for which useful information is sent to, and let $\overline{\mathcal{K}}(k) = \mathcal{K} \setminus \overline{\mathcal{K}}(k)$ be the set of its unintended receivers. Each transmitter $k$ is equipped with $N_k$ antennas, and each receiver with a single antenna. The quasi-static block flat-fading instantaneous channel vector from transmitter $k, k \in \mathcal{T}$, to receiver $\ell, \ell \in \mathcal{K}$, is denoted by $h_{k\ell} \in \mathbb{C}^{N_k \times 1}$. The transmit covariance matrix of transmitter $k$ is given as $Q_k \in \mathbb{C}^{N_k \times N_k}$, $Q_k \succeq 0$. The basic model for the matched-filtered, symbol-sampled complex baseband data received at receiver $\ell$ is

$$y_\ell = \sum_{k=1}^{T} h_{k\ell}^H Q_k^{1/2} s_k + n_\ell,$$

(3.9)

where $s_k$ is the symbols vector transmitted by transmitter $k$ and $n_\ell$ are the noise terms which we model as independent and identically distributed (i.i.d.) complex Gaussian with zero mean and variance $\sigma^2$. Each transmitter has a total power constraint of $P := 1$ which leads to the constraint $\text{tr}(Q_k) \leq 1$, $k \in \mathcal{T}$. Throughout, we define the signal to noise ratio (SNR) as $1/\sigma^2$. The feasible set of covariance matrices for transmitter $k$ is defined as

1Differentiated Service model
\[ S_k := \{ Q_k \in \mathbb{C}^{N_k \times N_k} : Q_k \succeq 0, \text{tr}(Q_k) \leq 1 \}. \] (3.10)

The performance measure of a system in an interference network is usually described by a utility function. The utility function associated with a receiver depends on the power gains originating from the transmitters in the network. Define the power gain achieved by transmitter \( k \) at a receiver \( \ell \) as
\[
x_{k,\ell}(Q_k) = h_k^H Q_k h_{k\ell}, \tag{3.11}
\]
where \( x_{\ell}(Q_k) \in \mathbb{R}_+ \) since \( Q_k \) is positive semidefinite. The utility function associated with a receiver \( \ell \) is defined as
\[
u_{\ell} : \mathbb{R}_{+}^{T} \rightarrow \mathbb{R}_{+},
\]
where \( T \) is the number of transmitters in the network.

3.1 Assumption. The utility function \( u_{\ell}, \ell \in \mathcal{K} \), has the following properties:

A. If \( \ell \in \mathcal{K}(k) \), then \( u_{\ell} \) is monotonically increasing in the power gain from intended transmitter \( k \), i.e.,
\[
u_{\ell}(x_{1,\ell}(Q_1), \ldots, x_{T,\ell}(Q_T)) \leq u_{\ell}\left(x_{1,\ell}(Q_1), \ldots, x_{k,\ell}(\widehat{Q}_k), \ldots, x_{T,\ell}(Q_T)\right),
\]
for \( x_{k,\ell}(Q_k) \leq x_{k,\ell}(\widehat{Q}_k) \).

B. If \( \ell \in \mathcal{K}(k) \), then \( u_{\ell} \) is monotonically decreasing in the power gain from unintended transmitter \( k \), i.e.,
\[
u_{\ell}(x_{1,\ell}(Q_1), \ldots, x_{T,\ell}(Q_T)) \geq u_{\ell}\left(x_{1,\ell}(Q_1), \ldots, x_{k,\ell}(\widehat{Q}_k), \ldots, x_{T,\ell}(Q_T)\right),
\]
for \( x_{k,\ell}(Q_k) \leq x_{k,\ell}(\widehat{Q}_k) \). \( \square \)

Assumption 3.1 describes the settings where the performance measure at a receiver increases monotonically with increased power gain from intended transmitters and decreases monotonically with increased power gain from unintended transmitters. An example utility function which satisfies Assumption 3.1 is the signal to interference plus noise ratio (SINR).

The utility region is the set of all achievable utility tuples defined as:
\[
\mathcal{U} := \{(u_1(x_{1,1}(Q_1), \ldots, x_{T,1}(Q_T))), \ldots, u_K(x_{1,K}(Q_1), \ldots, x_{T,K}(Q_T))): Q_k \in S_k, k \in \mathcal{T} \} \subset \mathbb{R}_+^K. \tag{3.14}
\]

The efficient operating points in the utility region correspond to those in which it is impossible to improve the performance of one system without simultaneously degrading the performance of at least one other system. Such operating points are called Pareto optimal and are defined formally as follows.

3.2 Definition. A tuple \( (u_1, \ldots, u_K) \in \mathcal{U} \) is Pareto optimal if there is no other tuple \( (u'_1, \ldots, u'_K) \in \mathcal{U} \) such that \( (u'_1, \ldots, u'_K) \geq (u_1, \ldots, u_K) \), where the inequality is component-wise and strict for at least one component. The set of all Pareto optimal operating points constitutes the Pareto boundary \( \mathcal{PB} \) of \( \mathcal{U} \). \( \square \)
Next, we give an example setting where the systems’ utility functions satisfy Assumption 3.1.

**Example Setting**

Consider two transmitters each using three transmit antennas, and three single antenna receivers as depicted in Figure 3.1. The operation of the systems is as follows:

- **Broadcast Channel (BC):** Transmitter 1 transmits different useful data to receivers 1 and 2 simultaneously. We assume transmitter 1 chooses the transmit covariance matrices $Q_{11}$ with $\text{tr}(Q_{11}) = p_{11}$ for receiver 1 and $Q_{12}$ with $\text{tr}(Q_{12}) = p_{12}$ for receiver 2. Hence, transmitter 1 can be considered as two virtual transmitters, 11 and 12, coupled by the total power constraint, $p_{11} + p_{12} \leq 1$. The receivers are identified in the following receiver sets: $1 \in K(11), 1 \in K(12), 2 \in K(12), 2 \in K(11)$.

- **Multiple Access Channel (MAC):** Transmitters 12 and 2 send distinct useful information to receiver 2. Receiver 2 decodes the data from transmitter 12 and 2 successively. Thus, $2 \in K(12), 2 \in K(2)$.

- **Multicast:** Transmitter 2 sends common useful data in a multicast to receivers 2 and 3. The receivers are identified in the following receiver sets: $2 \in K(2), 3 \in K(2)$.

- **Interference Channel (IC):** Transmitter 2 induces interference on receiver 1, while transmitter 1 induces interference on receiver 3.

The receiver sets are summarized in Figure 3.1, and the solid and dashed arrows refer to useful and not useful signal directions, respectively. The achievable rate at receiver 1 is

$$u_1(x_{11,1}(Q_{11}), x_{12,1}(Q_{12}), x_{2,1}(Q_2)) = \log_2 \left( 1 + \frac{h_{11}^H Q_{11} h_{11}}{\sigma^2 + h_{11}^H Q_{12} h_{11} + h_{21}^H Q_{2} h_{21}} \right), (3.15)$$

which is monotonically increasing in $x_{11,1}(Q_{11})$ and monotonically decreasing in the power gains from transmitters 12 and 2. The utility at receiver 2 is its sum capacity,

$$u_2(x_{11,2}(Q_{11}), x_{12,2}(Q_{12}), x_{2,2}(Q_2)) = \log_2 \left( 1 + \frac{h_{12}^H Q_{12} h_{12} + h_{22}^H Q_{2} h_{22}}{\sigma^2 + h_{12}^H Q_{11} h_{12}} \right), (3.16)$$

which is monotonically increasing in $x_{12,2}(Q_{12})$ and $x_{2,2}(Q_2)$. The utility function at receiver 3 is the achievable rate,

$$u_3(x_{11,3}(Q_{11}), x_{12,3}(Q_{12}), x_{2,3}(Q_2)) = \log_2 \left( 1 + \frac{h_{23}^H Q_{2} h_{23}}{\sigma^2 + h_{43}^H Q_{11} h_{43} + h_{13}^H Q_{12} h_{13}} \right). (3.17)$$

which is monotonically increasing in $x_{2,3}(Q_2)$. Note that the transmission rate at transmitter 2 has to be chosen such that both receiver 2 and 3 can decode the data.
Figure 3.1: An example setting for the described system model. There exist two transmitters, each equipped with three antennas, and three single antenna receivers. The solid arrows refer to the intended receivers of a transmitter, while the dashed arrows refer to interference directions.

successfully. We do not consider this requirement in (3.16) and (3.17). These rates can be achieved using rateless coding. The utility functions in (3.15)-(3.17) satisfy properties A and B in Assumption 3.1.

In this example, it can be observed that the optimization of these three utility functions is in general a multi-criteria optimization problem. The corresponding utility regions are illustrated in Figure 3.2. The region is not convex and finding efficient operating points on the boundary of the region is difficult.
3.5 Utility Metrics Based on Multi-Antenna Channel Gains

Figure 3.2: Pareto boundary of the utility region of the setting described with SNR=15 dB and N = 3.
3 Performance Metrics for Resource Sharing Schemes
4 Infrastructure sharing models

4.1 Infrastructure sharing state-of-the-art

One of the most challenging tasks in nowadays radio network planning is to provide the solutions that would allow the operators to significantly reduce the ever-growing OPEX [28]. It is especially important in deployment phase of new wireless technologies, as although the new wireless technologies lead to the exponential traffic growth, the increased demand for service is not balanced with increased revenues for MNOs (Mobile Network Operators) (see also figure 4.1). The temporary solution to the problem is to be the first to introduce novel technologies (in the present tense it would be LTE and LTE-Advanced), which allows to benefit from the market monopoly (increased service pricing). However introduction of new technologies is limited by the cost of familiarizing with the technology, spectrum and site acquisition and pressures from regulatory bodies ([25]) to minimize number of sites in dense urban areas 1. All the factors lead to a shift in networking business model paradigms. Current interest of MNOs is moved from full ownership 2 to shared infrastructure networks. So far the network infrastructure sharing has also been implemented in different networking solutions, as co-location [25] or as full network sharing network [22]. Furthermore as standardized approach to provide new compatibility feature for MNOs [29]. Application of infrastructure sharing might be especially important (can provide high cost-efficiency) in any new roll-out, consolidation (old technology replacement) or in coverage-driven network expansions 3. In the case of capacity-driven network expansions 4 sharing might not be the most efficient solution, as capacity enhancements are usually connected with additional CAPEX related to software licenses or baseband card extension, which although providing non-negligible costs, can be compensated over time.

Network infrastructure sharing can be characterized in terms of [17]:

- **Business model** - describes parties involved in the sharing and relations between them.

---

1 Regulatory bodies would not only look for increased radio network capacity but also the introduction of sites to poor coverage rural, which are not attractive from business perspective, as a social responsibility of the MNOs.

2 MNO plans the network, acquires and builds the sites, eventually implements, operates and maintains the network (where some of the parts might be done by subcontractors). MNO owns also network equipment, software and spectrum licenses.

3 Rural area deployments.

4 Addition of new channel elements, driven by a growing number of users (typically dense urban environments).
4 Infrastructure sharing models

Figure 4.1: Growth in traffic data costs in comparison with operators revenues [15].

- **Geographic model** - describes physical footprint of each of the parties involved in infrastructure sharing, it involves ownership of the network elements and the depth of sharing solutions.

- **Technology model** - describes technological solutions that are implemented to provide sharing, e.g. aggregation nodes, schedulers.

The choice of a business model as well as a geographic model is highly dependent on the operator status (existing or greenfield operator), deployment purpose (network extension or new roll-out) and on valid regulatory policies (spectrum sharing might be forbidden by law regulations of the country). The issues concerning business and geographic models describing different types of network sharing proposed for SAPHYRE are discussed in details in the deliverables D5.2a/b and D5.1a/b respectively. The technological models are further discussed in this document.

4.1.1 Types of infrastructure sharing

Depending on the solution for infrastructure sharing the cost savings can vary and most of the time they are achieved by a sacrifice in the domain of standalone network control. This aspect is very crucial especially when active resources (such as baseband processing powers or backhaul capacity) are shared, as loss of standalone control might lead to QoS degradation and jeopardization of confidentiality of operator’s traffic patterns. In this case additional algorithms that support multi-operator aware and fair resource allocation are needed.

Different technical solutions for infrastructure sharing are created based on the type of the network element or site equipment that is being shared [17, 2]:

1. **Passive RAN**\(^5\) sharing - it is often referred to as site sharing or co-location.
   It has already become a solution for the operators to minimize capital (e.g. acquisition, civil work, mast) and operational expenditures (e.g. site rental fee, site maintenance fee) [25]. Possible solutions include also third party owners, which can be specialized rental companies that provide operation of

\(^5\)Radio Access Network
4.1 Infrastructure sharing state-of-the-art

Figure 4.2: Different types of infrastructure sharing in the domains of network control, sharing and cost savings.
e.g., telecommunication masts. Regulators are encouraging sharing of sites (what is also highly appreciated by the public) as it leads to the reduction of total number of sites. Site sharing may involve sharing of equipment such as site itself, mast, shelters, cabinets, electric power supply, air conditioning, diesel or biofuel generators, ducts and antennas\(^6\). The drawbacks of this type of infrastructure sharing are:

- Coordination of operational and planning aspects with sharing partners.
- In the case of antenna sharing, combined loss (combining) and no RX diversity (chaining).

2. **Passive RAN sharing with Access Transmission Sharing** - additionally also transport links (backhauls) between base station and controller can be shared, e.g., in 2G between BTS and BSC (Base Station Controller), in 3G between NodeB and RNC (Radio Network Controller), in 4G between eNodeB and MME. It should be explained that typically backhaul sharing is realized by a physical separation of the two operators, where either physical links are separated but the aggregation/multiplexing equipment is common or operators use separate carriers to transmit their data. Both of the solutions provide traffic separation and moderate cost savings. Nonetheless backhaul can be shared also as an active element, where the scheduled resource is available capacity. Such type of sharing can be realized also via third party, which provides logical links with specific classes of service offered to different operators. Physically backhaul link can be realized via leased lines\(^7\), microwave communications (e.g. WiMAX), or fiber communications.

3. **Active RAN sharing Multi-Operator RAN (MORAN)** - the approach describes another degree of sharing, where also active (in the sense of changes in the equipments software) elements (e.g. base stations) of mobile networks are shared. In this approach even though shared element is active, operator maintains control over traffic flow as well as quality aspects (coverage, capacity, link parameters). Control is maintained by virtual (logical) and static division of shared network elements. Possible network elements that can be shared: Base Stations (baseband cards, power amplifiers), BSC, RNC or Relay Node\(^8\). Such virtual access networks are then connected to the respective operator core network. This technical solution allows operators to additionally reduce

---

\(^6\)Antenna systems can be shared in multiple ways: shared antenna radomes, shared antenna system by combining (combiners are used to connect antenna systems or receivers) or chaining (antennas are linked in chains)

\(^7\)The name is typically used to describe circuit-switched WANs, which allow permanent connection between two points set up by a telecommunications common carrier, referred to as private lines or dedicated lines [25].

\(^8\)Elements of mobile network infrastructure without a wired backhaul connection, that relay messages between the base station and mobile stations through multi-hop communication (to improve coverage, especially on the cell edges) [26].
Figure 4.3: Passive RAN sharing with shared backhaul link in 3G networks.
costs due to lower number of network elements\(^9\), maintain total independence in their roaming agreements and keep the sharing not visible to the users. Unfortunately Active RAN sharing solution has also many drawbacks [25]:

- operators have to use adjacent bands, e.g. due to technical limitations of power amplifiers,
- all the optional features of network elements have to be the same for both operators,
- capacity (CEs - channel elements) is pooled between the operators (one operator can exhaust available CEs),
- due to static division, elements may stay underutilized if operators have asymmetrical traffic volumes.
- O&M architecture to manage shared systems is very complex.

![Figure 4.4: A) Passive RAN sharing (co-located site and shared cabinet). B) Active RAN sharing (shared UTRAN).](image)

4. **Active RAN Sharing 3G Multi-Operator Core Network (MOCN) and (GWCN)) [29]** - multiple operators share UTRAN\(^{10}\), eUTRAN network elements and common frequency pool. GWCN approach introduces additionally partial sharing of the Core Network: in 3G MSC (Mobile Switching Center), SGSN (Serving GPRS Support Node) and in 4G MME (Mobility Management Entity). In case of shared accesses each cell in shared radio access network broadcasts (in the system information blocks) information on available core network operators PLMN-ids. This information is used by the users during attachment, handover and cell re-selection procedures. The available core network operators shall be the same for all cells of Location Area (3G) or

\(^9\)This leads to less power consumption, split of planning, optimization and maintenance costs [17].

\(^{10}\)UMTS Terrestrial Radio Access Network
Tracking Area (4G). In this solution network sharing is transparent to the users. As a result of this approach operators are loosing much of their control over traffic capacity and quality. Therefore this solution is mostly acceptable for low traffic rural areas. However with proper inter-operator agreements SLA (Service Level Agreements) and resource sharing algorithms QoS regime and fairness can be provided to satisfy network sharing operators. The saved costs are comparable to the ones achieved with MORAN approach.

5. **Roaming-based sharing and full network sharing** [25] - this type of sharing relies on the fact that each operator deploys its own network with a dedicated frequency band in specific region of the world/continent/country. In order to enable roaming-based sharing each of the operator’s networks need to support inter-PLMN mobility procedures. From legal perspective this means that operators require national roaming agreements, Service Level Agreements and agreement from the regulator to provide world-wide coverage. In fact up today this is a typical implementation for provisioning of international coverage (international roaming). National roaming can also be used by greenfield operators to provide nation-wide coverage before developing own network. In the full sharing case operators only retain portion of the core network sepa-

---

Figure 4.5: 3GPP vision of infrastructure sharing - Gateway Core Network sharing.
rated: HLR (Home Location Register), authentication (AUC and AAA server) and billing system (online and offline charging). Such a separation can be met in the deployment of MVNOs (Mobile Virtual Network Operators).

4.2 Backhaul sharing model

The broad introduction of 'all-IP' concept along with Long Term Evolution (LTE) networks [21], meant that transport networks solutions have to be definitely shifted from TDMA (Time Division Multiple Access) techniques towards Ethernet links, which are more flexible in handling IP datagrams. When deploying high capacity Ethernet links operators may initially experience high underutilization of the available link resources, i.e. users subscribe mainly for voice service or peak to average ratio of traffic volume is very high. In such a case operators are seeking the ways to reduce total cost of ownership (TCO) for backhaul links to the decrease the cost per bit of transmitted data. Ideal solution to maximize the utilization and reduce TCO is to share backhaul links with other operators. Backhaul sharing itself is one of the most interesting but very challenging aspects of Radio Access Network sharing. The biggest threat lies in the congestion (and therefore also dropping probability), that can occur when one or more operators traffic is exceeding its maximum share or backhaul capacity in general. This is highly unwanted situation against which a proper prevention mechanism shall be proposed. Furthermore it is not only about possible congestion but also fairness in the available resource distribution. In fact the key problem is to fairly distribute capacity resources with guaranteed Quality of Service levels for different services and still maintain high utilization of the available capacity. In principle there are at least three possible ways to deal with the problem:

1. **Backhaul dimensioning** based on maximum traffic ratio of each of the operators. The solution provides good support for different QoS levels, but it has many drawbacks that discard it as a potential candidate for backhaul sharing model:
   - high underutilization as the maximum rates usually occur only during small parts of the day\(^\text{11}\),
   - increased expenditures due to over-planning;
   - small flexibility towards future growth of demand and expansion of operator’s network.

2. **Load (Congestion) Control** algorithms, which are used to avoid bottleneck problem in the limited capacity links. The idea is to reduce the transmission rate of flowing packets. The possible methods for rate limitation include algorithms such as Random Early Detection (RED) [16] or token bucket class

\(^{11}\)In fact what is important here is the ratio between peak and average traffic volume.
4.2 Backhaul sharing model

of algorithms [32] where the incoming packets are dropped (blocked) depending on the link conditions, service priority and specific fairness rules. Such a solution can be applied to the backhaul link on different abstraction layers:

- transport layer where aggregation node or end router performs traffic classification and traffic scheduling to the link.

3. Flow Control (FC) algorithms, which are used to adapt the sending rate to the receiving rate at the final or intermediate node. The algorithm for flow control algorithms typically relies on end-to-end buffer state message exchange and RTT (Round Trip Time) measurements. Number of flow control algorithms has been proposed for Iub link between NodeB and RNC, as both entities can exchange their buffer state [23, 31]. The mechanism is however not fully applicable to LTE backhaul link as the communication with the Core Network may happen via external links, where routing paths might not be static, which could lead to high variations in RTT measured values. Nevertheless in [27] authors propose an algorithm for LTE radio link flow control that includes also the performance of the S1 interface.

There are also two additional aspects of backhaul sharing that are important in the process of shared backhaul link design\textsuperscript{12}:

- Underlying physical medium transport solution.
- Economical model.

Specifically physical layer solution for backhaul link need to address requirements on:

1. capacity supported at the served cell(s),
2. number of users,
3. traffic models (including day-night traffic patterns),
4. site location.

Self-evidently, possible solutions will vary depending on the value of the above mentioned parameters. The typical solutions recognized for backhaul links are [25, 21]:

- Leased lines (dedicated lines) - circuit-switched WANs (Wide Area Networks), which allow permanent connection between two end points of communication system via common carrier. Dedicated lines are typically used to connect geographically separated locations or to provide high capacity connection with the Internet. There are two basic types of leased lines realizations: p2p (point-to-point) and multipoint LMDS\textsuperscript{13}. p2p lines are used to connect

\textsuperscript{12}Which are however not stressed in the document as they refer to the business and financial part.
\textsuperscript{13}Local Multipoint Distribution System.
two locations directly for full-time and full capacity communication. LMDS lines are used to connect multiple locations to central facility over number of common transmission channels. Technologically leased lines can be realized as Very-high-bitrate Digital Subscriber Line (VDSL), NG SDH/SONET (Next Generation Synchronous Digital Hierarchy/Synchronous Optical Networking), Ethernet over fiber (Dense Wavelength Division Multiplex - DWDM) or PON (Passive Optical Networking).

- **Radio links** - p2p or multipoint connections realized over wireless media, where possible solutions include: licensed spectrum transmission paths (e.g. WiMAX), unlicensed spectrum transmission (e.g. WiFi) or even free space optics systems.

- **Self-backhauling** [19] - it is an approach proposed for LTE-Advanced to use eNodeB (which becomes anchor point) radio links as a backhaul for another eNodeB to communicate with the Core Network. It is IP layer multi-hop solution that reuses existing base stations and requires minimum system level adaptations.

The economical model in the state of the art solutions [7] for Ethernet link sharing can be realized in one of the two ways:

1. By rental of the links from fixed-line providers - where backhaul is realized as a combination of links: from base stations to the edge routers, between edge routers and operator’s PLMNs. This may in fact mean that the link can be realized through the Internet.

2. By consolidation of the transport network infrastructure with other operators - where traffic separation and link scheduling is realized by traffic aggregation nodes.

The following section provides a complete model of a shared backhaul solution for SAPHYRE. The description of the model includes: LTE backhaul link architecture, transport mechanisms as well as possible solutions for flow separation between the operators.

### 4.2.1 LTE backhaul link architecture

3GPP LTE offers significantly higher data rates in comparison to 2G and 3G systems. LTE is most likely to offer 100/50 Mbps in uplink/downlink (in 1 sector, with 20 MHz bandwidth) and even up to 1 Gbps for the whole site that consists of three sectors where downlink/uplink data rates are of 300/150 Mbps [30]. This in comparison with 3G peak data rates of 30 Mbps means that existing backhaul capacities need to be increased significantly by at least 10x the same capacity\(^\text{14}\).

\(^\text{14}\)It is expected that due to enhanced bandwidth and extra features proposed, LTE-Advanced will have even up to 1 Gpbs in downlink [4].
4.2 Backhaul sharing model

Another important factor that greatly differentiates 4G backhaul links from legacy systems, is the shift in architectural paradigm from hierarchical TDM (Time Division Multiplex) based architecture to flat IP-based. Flat architecture consists of lesser number of elements, which in turns forces mesh topology, and expects purely IP packet transmission between network entities. The figure 4.6 presents a flat architecture of backhaul link connectivity, where RANs are connected with different Core Network entities as well as with each other, constituting for meshed topology.

Figure 4.6: Flat architecture backhaul connectivity with inter-eNodeB communication.

Due to mesh network topology (inter-eNodeB connectivity), radio network controllers were replaced with evolved base stations (eNodeB), which perform dynamic resource allocation, radio admission control, connection mobility control, measurement configuration and intercell Radio Resource Management [21]. The introduction of new functionalities directly in Radio Access Networks lead to changes in the system architecture and appearance of new logical interfaces which constitute for evolved backhaul links (see also figure 4.7):

**S1-MME.** - which is a point-to-point\(^{15}\) link that carries control plane data between eNodeB and MME. Functions of the interface include: handling of RAB

\(^{15}\)Although it is possible to have connection towards multiple MMEs (Mobility Management Entity), terminal can be associated with only one MME at a time.
(Radio Access Bearer) procedures, handover procedures, NAS (Non Access Stratum) signaling and paging. The interface is required to provide high level of reliability in order to avoid message retransmissions and unnecessary delay in control plane procedure executions [21]. Due these requirements the interface uses reliable end-to-end transport layer communication via SCTP (Stream Control Transmission Protocol).\footnote{SCTP implements path selection and monitoring, flow control, validation and selective acknowledgements and order preservation [21].}

**S1-U.** - which is either point-to-point or point-to-multipoint (S1-flex\footnote{S1-flex means that eNodeB can be connected to multiple Serving GWs and MMEs. Where each logical control connection between UE and MME is marked with using different S1-AP Id [21]. This solution gives robustness towards Core Network node failures, more flexibility in network architecture and limitation of inter-Core Network handover procedures.}) link, that connects eNodeB with Serving Gateway(s) to transport user data packets. There is no need for flow control nor error control, nor any mechanism to guarantee data delivery over the S1-U interface, therefore in transport layer GTP (GPRS Tunneling Protocol) protocol over UDP (User Datagram Protocol) is used, which provides only data encapsulation [21].

**X2.** - which is a point-to-multipoint interface inter-connecting eNodeBs. It is used to transport:

- User data packets (via GTP protocol), user context information and signaling (via SCTP) in the case of handover procedures.
- Load indicators - to support load balancing management and to optimize handover decisions.
- Intercell interference coordination (ICIC) - information required to support ICIC, e.g.: allocated carriers, specific information is yet to be specified.

Altogether the three logical interfaces bring an astonishing increase in the amount of transported data and signaling, traffic prioritization and new connectivity solutions, which poses a number of requirements on the backhaul link transport architecture.

The detailed requirements for next generation backhaul transport networks have been specified by NGMN (Next Generation Mobile Networks) [24]:

- **High bandwidth** - current radio solutions require at least 450/150 Mbps in DL/UL. The demand for bandwidth (especially with introduction of LTE-Advanced) will increase over time, therefore transport network solution should be scalable to fit requirements of new radio interface solutions, different user environments (rural and urban sites) and rising interest in bandwidth consuming applications.
4.2 Backhaul sharing model

Figure 4.7: Evolved Packet Core backhaul link architecture.

- **Flat architecture** - eNodeBs and Access Gateways (aGWs) shall be connected in a mesh topology type to provide many-to-many connectivity.

- **Support for QoS mechanism** - radio QoS Class Identifiers (QCI) shall be mapped on transport QoS markings, so that packet classification can be performed in the transport layer. Transport equipment is required to implement packet scheduling algorithms to guarantee requested QoS over backhaul link (in case of congestion high priority packets are sent first).

- **Low latency** - expected two-way delay shall not be higher than 10ms and it shall be possible for operators to achieve even 5ms if required.

- **Synchronization** - NGMN backhaul solutions shall support clock reference distribution over packet network. The synchronization mechanism shall support distribution of frequency, phase and time source to enable alignment of eNodeBs.

- **Link availability and fault restoration** - the availability of backhaul shall be tunable according to operators needs (e.g. in case of microwave links it can be done via Adaptive Modulation and Coding). The expected availability shall be 99.99% of time and expected link outages (before path resiliency works) shall be in the range of 50ms - 250ms.

- **Fault Management** - backhaul network elements shall have OAM (Operation, Administration and Maintenance) protocols to reactively and pro-actively respond to link failures to support required end-user experience and Quality of Service. The OAM protocols shall enable backhaul link management (e.g. paths upgrade) via NMS (Network Management Systems).

- **Service continuity** - most of the already deployed sites in 2G and 3G technologies will still need to be maintained. It is highly likely that new LTE sites will
be co-located with the legacy ones, therefore there is a need for next generation backhaul technologies to support also emulation of TDM services over Ethernet links or hybrid architecture were both technologies are supported using different carriers or separate physical links.

**LTE backhaul transport protocols**

The protocols used in LTE transport network need to answer the NGMN requirements, and enable the implementation of aggregation nodes in the Layer 2, which allows for faster and less expensive traffic switching and aggregation. The figure 4.8 presents an abstraction of possible LTE backhaul protocol stack for the consolidation backhaul link, where the main focus is on transport provisioning, realized via Ethernet protocol (possibly also Ethernet over SDH), which is further incorporated with PBB-TE (Ethernet tunneling - Provider Backbone Bridging with Traffic Engineering) [1] or MPLS-TP (Multilabel Path Switching Transport Profile) [11] labeling to enable Virtual LAN creation. The specified protocols are used to transport either IP datagrams or legacy TDM services using circuit emulation (PWE3).

![Protocol stack for consolidated backhaul link](image)

In order to efficiently realize logical connectivity in LTE backhaul networks each interface needs to be mapped to an appropriate Ethernet service\(^\text{18}\) configuration and Ethernet Virtual Connections (EVCs), an example mapping [7]:

- S1 interface can be realized as either E-Line (number of statistically multiplexed point-to-point connections) or E-Tree (point-to-multipoint connections between leaves and roots\(^\text{19}\) as well as roots to a number of leaves) service.
- X2 interface can be realized as E-LAN service to facilitate direct point-to-multipoint interface towards other eNodeBs.

\(^{18}\)For more information on Ethernet services, see [5].
\(^{19}\)In case of S1-flex multi-rooted E-Tree can be used.
4.2 Backhaul sharing model

In order to enable coexistence of multiple operators among same transport links at least three aspects need to be covered:

1. **Quality of Service delivery.** It shall be possible to classify packets based on the service they are carrying. The classification can be used then by traffic marking and limiting mechanisms to provide service differentiation and to guarantee minimum QoS classes described in Service Level Agreements.

2. **Traffic separation.** The second important aspect in designing shared backhaul solution is the traffic separation between the operators. Appropriate solution should enable application of different QoS policies to different traffic flows independently for each operator. The solution that enables logical separation of physical links is VLAN marking [8]. In VLAN approach flows of different operators are assigned to different VLANs and packets of each flow are marked with corresponding VLAN id. Each operator has at least one VLAN id, a number of VLANs can be used by one operator to separate also different logical interfaces.

3. **Synchronization.** Complete solution for shared backhaul shall address also the problem of synchronization provisioning (the operators typically prefer to maintain their own source of synchronization [7]). The distribution of synchronization signals (time, phase, frequency) is a challenging task in natively asynchronous packet networks. Due to the existence of multiple operators on one link, an end-to-end packet-based methods are required to provide proper separation of operators, instead of typical incorporation of reference signal into the physical layer, e.g. SDH or Sync-Ethernet.
Operator’s flows separation

Traffic separation in a backhaul operated by multiple operators is realized via division of the available physical links into separate network domains, called Virtual LANs. Based on the general idea, there are two possible solutions to divide the resources [8]:

- Each operator maintains one VLAN network to transport its traffic.
- Each operator receives a pool of available VLAN networks to separate its traffic from other operators and to separate also different logical interfaces (S1 and X2). It is possible also to separate S1-U and S1-MME interfaces from each other, such solution however is not necessary as Core Network edge entity (aGW) routes the C-plane and U-plane traffic to destined entity.

![Figure 4.10: Solution for shared backhaul with EVCs for different interfaces.](image)

The figure 4.10 shows VLANs separation between different interfaces. Each packet of a flow assigned to specific VLAN is marked with a label called Virtual Path Label Switched (VPLS)\(^{20}\). Typically the label is assigned by a customer edge router\(^{21}\). In the aggregation network all packets marked with the corresponding VLAN are assigned so called tunnel label that allows routing of the packets between aggregation nodes. Such solution is highly appreciated by operators as usually routers operating on lower layers are much cheaper in comparison to the ones operating on IP layer.

\(^{20}\)MPLS or PBB-TE multiprotocol encapsulation of Ethernet frames.

\(^{21}\)In the case of shared backhaul links labels can be assigned also by the base station’s Transport Module.
Another degree of traffic separation can be achieved from usage of security protocols (i.e. S1-MME can be encapsulated with IPSec protocol) to minimize the possibility of jeopardizing the operators confidential data.

Each VLAN can be abstracted as Ethernet service that needs to be served with the appropriate QoS class. The provisioning of QoS agreements is done by the aggregation node. The QoS provisioning is a two-level structure that comprises both demands of services and particular operators. The realization is done through the following mechanisms:\footnote{\textsuperscript{22}The presented solution is a generalization of the solution for Carrier Ethernet transport networks presented in [6].}

1. **Packet classification mechanism** - which is responsible for mapping of DiffServ QoS classes marked in IP header to corresponding CoS (Class of Service) provided by the transport layer protocol.

2. **Per service flow metering (rate limitation) and scheduling** - which is responsible for assignment of traffic rates (according to SLA) and pre-allocation of transport resources. The typical solution for rate limitation (and congestion control) in transport network is to use token bucket algorithms to mark the packets with specific color. Marked packets are sent to the queue of corresponding color and afterwards scheduled to the link. The queues are managed by Random Early Detection (RED) algorithm \cite{16} to identify a priori possible congestion situations and provide early dropping of packets.

3. **Per operator flow rate limitation and scheduling** - once the requirements for services are fulfilled it is time for the aggregation nodes to confront the incoming rates against the link capacity resources and apply inter-operator resource sharing agreements. The SLA adaptation can be done again via rate limitation algorithm (this time without coloring), however it is important to note, that the proposed algorithm need to maintain minimum required rate that needs to be available to the operator’s traffic. The traffic exceeding link capacity is either dropped or blocked depending on the queue management solution.

Such a three step mechanism allows proper application of strict QoS parameters (e.g. throughput) to most demanding services and controlled distribution of resources among operators.

**Summary**

Backhaul link sharing is an important aspect of a shared infrastructure, especially due to the fact that the backhaul performance can affect the call admission process at the radio link. Therefore in the case of shared backhauls it is very important to start the deployment with careful dimensioning, then providing an appropriate transport layer mechanisms for flow separation and service differentiation between
the operators, and finally design of aggregation nodes which implement QoS provisioning (via rate limitation and scheduling).

4.3 Relay sharing model

Recently, relays have received an increased interest due to their potential abilities of reducing the deployment cost, enhancing the network capacity, mitigating shadowing effects, and so on. Thus, it is interesting to exploit the benefits of the relay sharing in wireless networks. In this section, we will introduce two kinds of relay sharing models that have been investigated under the scope of SAPHYRE. The application of these two models are also described.

One relay sharing model is called multiple operator two-way relaying, which corresponds to the metropolitan scenario as shown in Figure 4.12(a). In this case, strong shadowing effects will cause many coverage holes and thereby dense networks are...
required to guarantee the QoS at the user terminals (UT). Moreover, considering the geometric constraints and network deployment costs, relays would be more suitable. Also taking into account that more than one operator or service provider would operate in the same area, if we share the relays as well as the spectrum, at the first glance it means lower capital expenditures and operating expenditures for all the operators. This model can also be applied to a disaster scenario where the base station cannot provide services any more. Then the relays can be deployed to temporarily maintain the communication among the local residents.

The other model incorporates BSs into the system, shown in Figure 4.12(b). In order to guarantee the QoS of the cell-edge users, a relay is usually employed to assist the BS in addition to a direct transmission (might be quite weak), which is modeled as the relay channel (RC). We are motivated to extend this model to the two transceiver pair case belonging to two different operators to further improve the spectral efficiency, where the relay is shared and accessed by both BSs at the same time instead of an exclusive use of the relay for each operator in a TDMA mode. We refer to this model as an interference relay channel (IRC). More specifically, it describes the channel model where two independent transceiver pairs with multiple antennas communicate with the assistance of one relay, which operates in half-duplex mode and employs an amplify-and-forward (AF), a decode-and-forward (DF) or a compress-and-forward (CF) strategy.

In both models, the AF relays are highly preferable not only because they significantly reduce the delay and the complexity but also because they avoid complex signaling (a relay does not need to acquire signaling knowledge of different operators). Moreover, since each operator does not need to know the modulation and coding schemes of the other operators, they do not need to share their data which leads to more independences to the operators. Furthermore, since the relay can even belong to a third-party, this kind of relay sharing will not harm the competitiveness of operators.

To further exploit these new relay deployment scenarios, we need to clarify the traffic loads of different operators. For notational simplicity, it is assumed that all the operators have infinite buffers and the same amount of traffic loads in our work.

Traditionally the transmit strategy is to assign the physical resources to all operators in an orthogonal manner, e.g. different time slots. When the physical resources (the relay and the spectrum) are shared between operators, two important questions arise:

- What spectral efficiency can we achieve?
- Compared to the exclusive approach, how much sharing gain in terms of sum rate can we obtain?

These are the questions that we will answer in the SAPHYRE deliverable D3.1a.
4 Infrastructure sharing models
5 Conclusions

SAPHYRE focuses on two forms of resource sharing, namely spectrum sharing and infrastructure sharing. The spectrum sharing introduces interference which can be resolved by the use of MIMO (multi-antenna) algorithms. Infrastructure sharing is a complex multi-level concept which comprises a number of different scenarios among which one distinguish especially the concepts of backhaul sharing and fixed wireless relay sharing. Several backhaul sharing models are provided and methodologies used to evaluate the sharing performance are outlined. A SAPHYRE gain is defined aimed at grasping the benefits of the proposed methods in a single performance number.
Bibliography


[7] Lte backhaul: Meeting operator requirements, how transport providers can realize the business opportunity of lte with sla assurance. RAD Data Communications, December 2009.


