D 2.6

Architectural optimization for different geo-types

Dissemination Level:

- **Dissemination level:**

  PU = Public,
  RE = Restricted to a group specified by the consortium (including the Commission Services),
  PP = Restricted to other programme participants (including the Commission Services),
  CO = Confidential, only for members of the consortium (including the Commission Services)

Abstract:

The major objective of D ISCUS is to produce an end to end design for a future network architecture that can deliver very high speed broadband services to their users. The architectural design must meet this objective while remaining economically viable and scalable.

This deliverable reports on the architecture optimization, modeling of reference topologies using street and building data, modeling of LR-PONs for various geo-types, modeling of optical island core networks, and the corresponding combinatorial optimization results.

The main focus of this report is the end-to-end network. It aims at integrating access network optimization below the metro-core-node and core network optimization connecting metro-core-nodes into an end-to-end optimization process that can be used to evaluate the end-to-end DISCUS design and compare it with competitive solutions.

COPYRIGHT

The DISCUS Consortium consists of:

Participant Number	Participant organization name	Participant org. short name	Country
Coordinator			
1	Trinity College Dublin	TCD	Ireland
Other Beneficiaries			
2	Alcatel-Lucent Deutschland AG	ALUD	Germany
3	Coriant R&D GMBH	COR	Germany
4	Telefonica Investigacion Y Desarrollo SA	TID	Spain
5	Telecom Italia S.p.A	TI	Italy
6	Aston Universtity	ASTON	United Kingdom
7	Interuniversitair Micro-Electronica Centrum VZW	IMEC	Belgium
8	III V Lab GIE	III-V	France
9	University College Cork, National University of Ireland, Cork	Tyndall & UCC	Ireland
10	Polatis Ltd	POLATIS	United Kingdom
11	atesio GMBH	ATESIO	Germany
12	Kungliga Tekniska Hoegskolan	KTH	Sweden

Authors: (In alphabetic order)

Name	Affiliation
Alejandro Arbelaez	UCC
Norbert Ascheuer	ATESIO
Jiajia Chen	KTH
Deepak Mehta	UCC
Luis Quesada	UCC
Ata Sasmaz	UCC
Christian Raack	ATESIO
Marco Ruffini	TCD
Lena Wosinska	KTH

Internal reviewers:

Name	Affiliation
Harald Rhode	COR
Thomas Pfeiffer	ALUD

Due date: 31.10.2014

Contents

1 Introduction

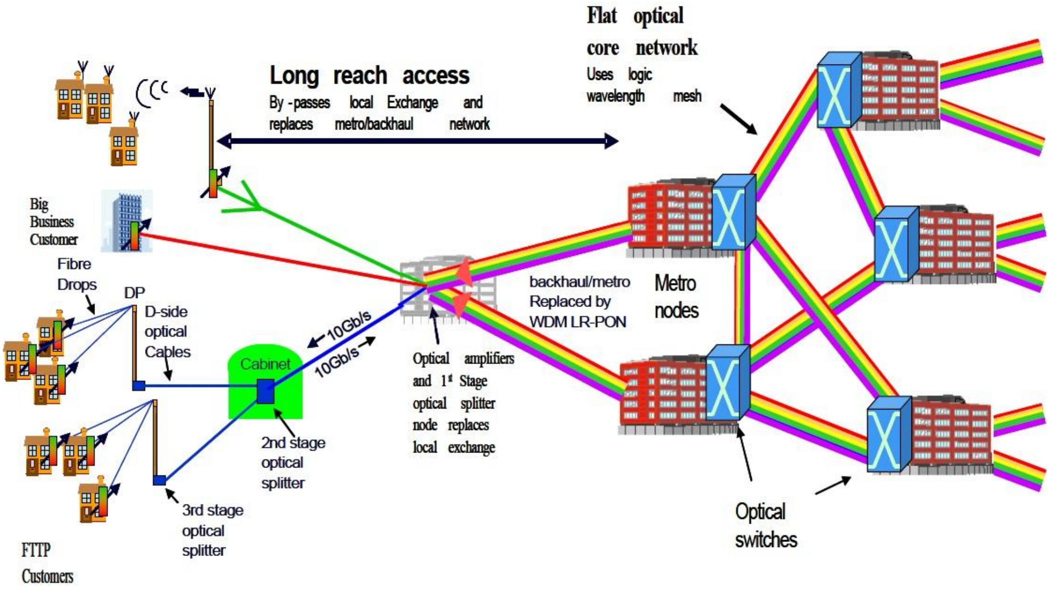


Figure 1: The end-to-end DISCUS architecture with LR-PON and flat optical core network

This deliverable provides the conceptual framework for optimizing and evaluating end-to-end solutions within the DISCUS project. In particular we show how the different optimization activities in WP4 (optimizing the LR-PON) and WP7 (optimizing the flat optical core) are combined into an end-to-end optimization process. We present the necessary data-models (e.g. reference networks, cost models, traffic demands) and methodology to compute end-to-end solutions and to come up with a global view on the DISCUS architecture. The end-to-end optimization process aims at providing evaluations of end-to-end solutions to compare alternatives of the DISCUS architecture but it also aims at comparing the DISCUS architecture with existing networks or current deployments, such as architectures based on GPON.

Any direct comparisons between DISCUS architectures or between architectures in general has to be done using the same reference data, that is, reference networks, reference demand scenarios, reference hardware technologies, and a reference cost-model. In Section 2 we provide the current status of the different reference models and show how they are used within the end-to-end process.

In Section 3, we describe the end-to-end optimization process in detail. We provide high-level views of the corresponding optimization problems tackled in WP4 (LR-PON deployments) and WP7 (core network optimization) and we show how global (nation-wide) solutions of optical islands and LE to MC assignments can be combined with local (regional) solutions of LR-PON deployments.

Providing a single optimized DISCUS solution based on the given data reference models is already a challenging task in itself, as we will see. Moreover, it is not a priori clear how to assess and evaluate the resulting end-to-end solutions. There are clearly different and often concurrent measures for deciding about the quality of an architecture solution. Besides cost (see Section 2.4) we will sketch how the availability of a network solution can be used as a performance indicator in Section 3.4.

Section 4 shows concrete examples for optimization both of the LR-PON and the flat optical core and Section 5 gives a preview on the planned activities towards the end of the project.

2 Reference data for end-to-end solutions

Below we introduce the data models that are used as a reference for the optimization activities within DISCUS. We provide reference networks, reference traffic demand scenarios, as wells as a tentative reference hardware and cost model.

2.1 Reference Networks for different geo-types

In this section we describe progress in Task 2.3 of the DISCUS project regarding the target to generate the reference networks necessary to perform techno-economic studies and optimization in the work packages WP2, WP4, and WP7.

Realistic reference topologies are critical for all techno-economic evaluations. This data is needed to make realistic comparisons of the variants of the DISCUS architecture but also to make comparisons with existing architectures such as GPON in the access and today's mix of Ethernet and IP in the aggregation and core networks. Clearly, if real network fiber topologies were available for all countries then there is no need to compute reference networks. However, such data is typically confidential and also network-operator-dependent. Below we provide networks that are as realistic as possible and as detailed as necessary. We used real-world data whenever available and closed data-gaps by using alternative and reasonable data sources.

We provide reference networks at different levels of detail:

- Section 2.1.1 reports on the progress regarding nation-wide networks on the basis of today's central office locations (CO). These topologies resemble national fiber topologies, and are the basis for LR-PON studies in the backhaul network and the consolidation of metro-core (MC) nodes, see, see Section 4.2. They are also used as a basis to compute core reference topologies.
- In Section 2.1.2 and 2.1.2 we introduce networks down from local exchange sites to buildings. These detailed reference networks provide a microscopic view and cover geographic areas at the size of smaller towns with several hundreds or thousands of customer premises. They are the basis for LR-PON studies in the optical distribution network, see Section 4.1.

2.1.1 Macroscopic: Nation-wide reference fiber topologies

In this section, we introduce nation-wide reference fiber topologies for the countries Italy, Spain, and the United Kingdom. For a microscopic nation-wide topology of Ireland see the next section. All of them contain local exchange sites (LEs) as nodes together with additional information such as the number of connected users. The distribution (location and number) of these LE sites reflects today's situation in the respective countries. This has been made possible by using (anonymized) data from the operators Telecom Italia (Italy), Telefonica (Spain), and British Telecom (United Kingdom).

However, the level of detail and the type of information we used differs strongly from country to country and will be explained below.

To compute the nation-wide reference networks, atesio established a process that combines location oriented data (geographic coordinates of LEs, connected customers, et cetera) with connection oriented data (fiber links). Since concrete fiber topologies are typically confidential, we used geo-referenced data from street networks to approximate these topologies. This approach is reasonable because in most cases laying fibers is done along streets. Street networks also reflect dense and sparse structures (towns and rural areas) and they reflect ‘forbidden’ fiber-routes as those across mountains or rivers. With open street maps (OSM) there exists extensive data sets that are open source and include fine-grained information down to residential roads and buildings.

That is, we combined some country-specific key-information obtained from the above mentioned network operators (location of COs, population/user statistics) with publically available data (streets, buildings).

To establish the mentioned process, we had to (i) extract the OSM-data from public sources for Italy, Spain, and the UK (openstreetmap.org) (ii) develop data-models to store the geo-referenced data (nodes with coordinates, trails/streets with their geographical representation, buildings, et cetera) (iii) implement tools that allow to process, simplify, and store the OSM-data (v) implement the eventual process that combines the geo- data with data from operators (coordinates, users) and computes the resulting nation-wide reference topologies.

In deliverable D2.2 [2], we already explained the main network generation framework resulting in a reference topology for Italy. However, in the meantime, we improved on our implementation which not only strongly accelerated the process but also strongly improved on the quality of the resulting networks. We will review the process below and comment on some of the improvements before we give more details about the individual networks for Italy, Spain, and United Kingdom.

Combining OSM data with location data

In the following, we will review the algorithmic framework to generate nation-wide reference topologies. However, for details we refer to Chapter 4 of Deliverable D2.2 [2].

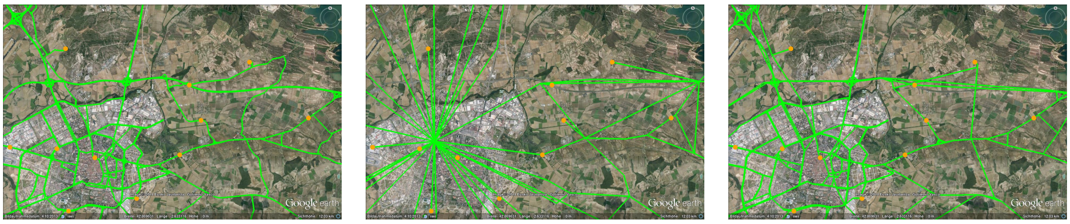
We first subdivide the countries into smaller regions. These regions are based on administration units such as the 20 regions (regioni) in Italy, the 50 provinces of Spain (provincias), and the 4 countries of the United Kingdom. This is done not only for performance reasons and to handle the massive amount of data. According to the network operators we talked to, fiber networks typically have a regional and inter-regional structure that is organized along the given administrational units, see below.

The process incorporates the geographic location (longitude and latitude) of central offices in the three countries. There are 10,708 COs for Italy, 5,578 COs for the UK, and 9,175 COs for Spain. For Italy and the United Kingdom we used anonymized coordinates from Telecom Italia and British Telecom, respectively. In these cases, the COs have been slightly changed and shifted from their original (real) location maintaining the nation-wide distribution characteristics. For Spain no such coordinates were available such that we aproximized the location of COs based on detailed population statistics using additional information on the distribution of COs in Spain obtained from Telefonica.

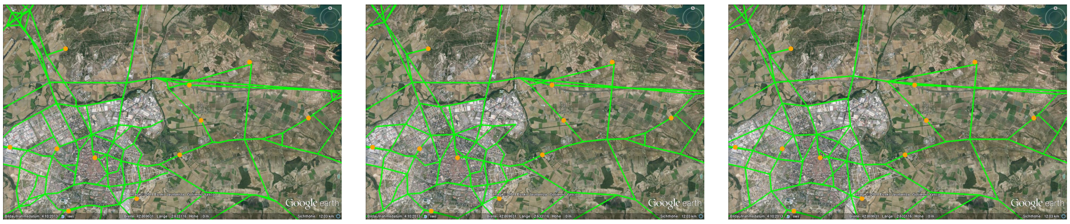
The main process, which is fully automated, has 5 major steps:

1. *Create a base-network for every individual region:* We extract the OSM-data mainly using the highest OSM-layers: highways, freeways, and primary roads. In addition to preprocessing and error correction we perform a first conservative reduction: (i) trails with the same end-nodes become one trail, (ii) nodes with degree 2 are removed but only if the main street structure is kept (turns, curves).

2. *Connect the given CO coordinates with the base-network in every region:* Connect all given COs to the (higher-layer) network by using paths in the OSM network. We first



(a) Unreduced network (b) Original reduction (c) Revised reduction, Iteration 1



(d) Revised reduction, Iteration 2 (e) Revised reduction, Iteration 3 (f) Revised reduction, Iteration 4

Figure 2: Reduction algorithm: The original node and link merging procedure from deliverable D2.2 (Figure 2(b)) led to some artifacts in urban areas in the resulting networks. We observe big star-like structures and nodes with large node-degree. This is mainly caused by too aggressive node-merging and by doing node-merging before link-merging. With the new iterative reduction algorithm (Figures 2(c)–2(f)) the main network structure from 2(a) is maintained.

shift the COs to close-by streets. Higher layer paths are preferred over lower-layer street-structures (secondary/tertiary roads, minor roads, residential roads). If possible we try to assure bi-connectivity (two disjoint paths) for each CO.

3. *Connect the individual regions to a nation-wide network:* To get a nation-wide network, the individual regional networks are connected. This is done using the highest OSM street layers, that is, by using inter-connecting highways and freeways.

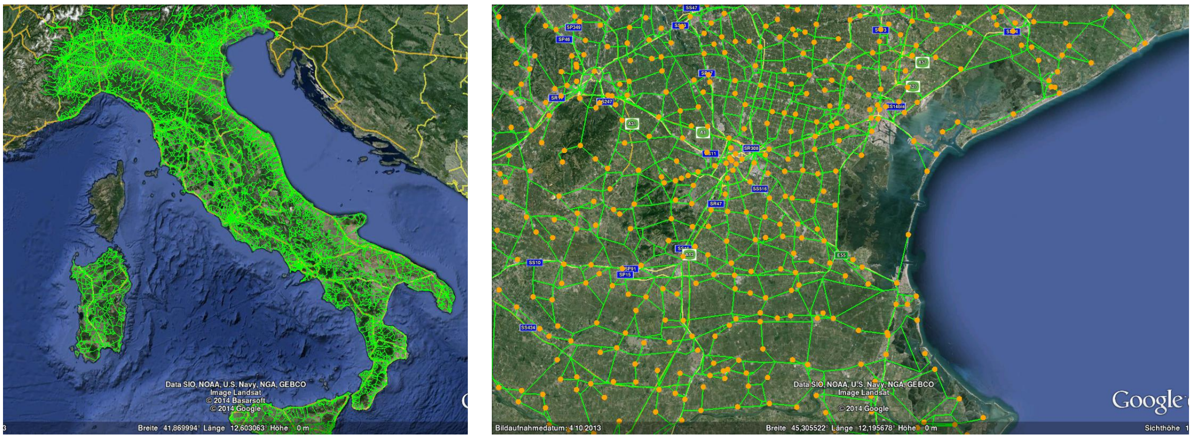
4. *Reduce and cleanup the network:* We reduce the street network to a network graph in the mathematical sense by performing two major reductions: (i) Merging links: Every node with a degree two that is not a central office location is removed. The two links are merged to one link, while we keep the information about the total length. Similarly, parallel links (having the same end-nodes) are merged. (ii) Merging nodes/crossings (not being central office locations): Nodes laying within a circle of certain radius get merged to one node. In addition we cleanup the network by removing degree-0 and degree-1 nodes.

5. *Cut the regional networks from the nation-wide network by using polygonal boundaries.*

The main improvements of our algorithm compared to the version introduced in D2.3 could be achieved by improving on Steps 2 and 4:

Instead of adding secondary roads from the beginning in Step 1, we ensure in Step 2 that every CO gets at least two different connections to the network. These connections might use lower layer structures. However, in the corresponding path computations we prefer higher layer over lower layer links. That is, the motivation for adding links to the network is now driven by the connectivity of the COs rather than by the OSM layers.

We also strongly improved on the network reduction. First of all we perform the reduction iteratively, that is, as long as the network changes we repeat the same reduction process. This leads to much cleaner networks with less nodes and links reflecting the same network structure.



(a) (b)

Figure 3: The Italian network: (a) the nation-wide topology, (b) the regional network of Veneto

However, within one reduction step we apply the node-merging in a more conservative way to maintain more fine-grained network structures within urban areas. This removed a couple of artifacts in the networks, in particular in cities, see Figure 2.

Recall that Step 1 and 2 are done for every individual region, while the regions are only connected with higher layer street structures. This is in line with existing network architectures, where regional networks are interconnected by fiber routes that follow large national highways or rail-way systems.

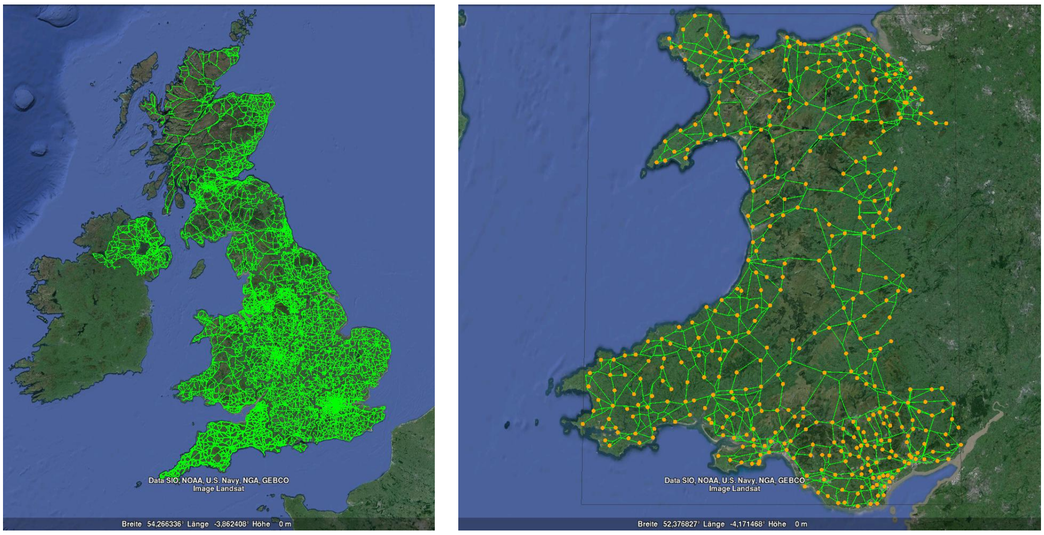
The Italian network

The generation of the nation-wide Italian network and the corresponding 20 regional networks is described in all detail in Chapter 4 of Deliverable D2.2 [2]. The Italian networks include 10,708 CO locations together with

- their anonymized coordinates (Latitude, Longitude),
- the number of connected cabinets,
- the number of connected residential customers,
- the number of connected business customers,
- the information about whether a CO is connected to the network by dual homing or not in the real topologies of Telecom Italia.

We mentioned above that we could improve on the quality of the networks. The improved Italian network became smaller with a total of now 23,689 nodes (before: 29,981) and 32,700 links (before: 39,615), see Figure 3, however, showing more details in urban areas and cities, see Figure 2 and Figure 3(b).

Recall that *nodes* in this context are mainly derived from street crossings in OSM. We may interpret them as potential splicing locations for cables and locations to house hardware. The set of COs is given as a proper subset of the set of nodes. The notion of *links* (derived from streets) refers to potential fiber (cable) paths. Their length is derived form the corresponding streets (highways, roads).



(a) (b)

Figure 4: The UK network: (a) the nation-wide topology, (b) the regional network of Wales

The British network

In cooperation with the IDEALIST project [13], we received anonymized coordinates of 5,578 CO locations in the United Kingdom from British Telecom. Similar to Italy every CO has a number of households attached (in total 29,374,911 households) such that the eventual process to generate the reference network is identical to the Italian case, see above. We subdivided the UK into the 4 regions corresponding to the 4 countries England, Northern Ireland, Scotland and Wales. The resulting nation-wide British network has 15,609 nodes and 23,025 links including the 5,578 COs as nodes.

The Spanish network

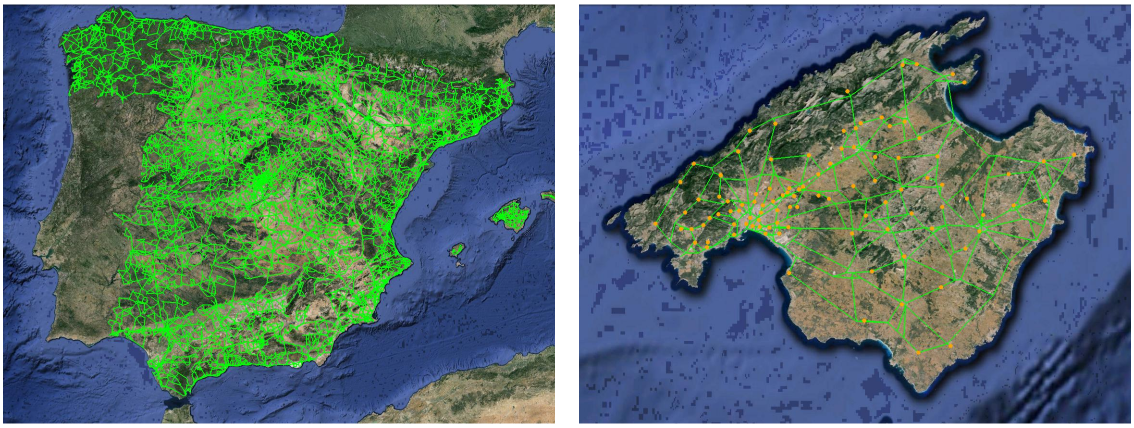
In case of Spain we had no access to coordinates of existing central office locations (anonymized or not). However, Telefonica provided us with detailed population and administrational statistics. At the core of this data set is a table of all 8,117 municipalities ('municipios') in Spain. This table has the following columns (among others), which we use in the process:

- a unique identifier,
- the municipality name,
- one coordinate (longitude, latitude) of a central location within the municipality region,
- the population (residential) of the municipality,
- the perimeter of the municipality region in meter, and
- the area of the municipality region in hectare.

In addition we have the following information from Telefonica:

- Spain has a total of around 9,000 COs (all operators).
- Telefonica has a total of around 5,000 to 6,000 COs in Spain.
- A CO has currently around 1000 to 15000 connected (DSL) customers.

Although roughly fitting with respect to the total number of COs, the municipality locations cannot be directly used as CO locations. For instance, there is only one municipality for Barcelona and one for Madrid with several million inhabitants, respectively. On the other hand, more than half of the municipalities has less than 1,000 inhabitants and for more than 1,000 municipalities this number is below 100, see Figure 6.



(a) (b)

Figure 5: The Spanish network: (a) the nation-wide topology, (b) the regional network of the island Mallorca

To obtain a reasonable distribution of COs in Spain we applied a two-step procedure:

1. *Distributing:* For municipalities with a large number of inhabitants we randomly distributed COs in the corresponding region
2. *Merging:* We iteratively merged municipalities with a very small number of inhabitants to obtain COs covering several towns with lower population taking into account a copper reach distance constraint of 10km

The distributing operation starts from the given coordinate of a municipality with more than 15,000 customers and computes a radius around this location based on the given perimeter and area information of the municipality. It then randomly adds new COs within this radius with a random number of customers between 4,000 and 15,000. The distributing operation tries to avoid distances between two neighbored COs smaller than 2km if possible. The COs are also places centralized, that is, the probability to be placed decreases with the distance from the given central coordinate. Of course COs are not placed outside mainland Spain. The operation stops when less than 15,000 inhabitants are left and opens a CO at the given

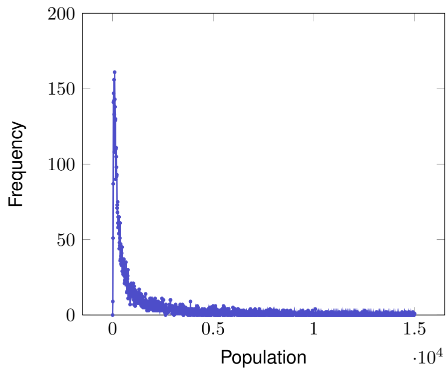
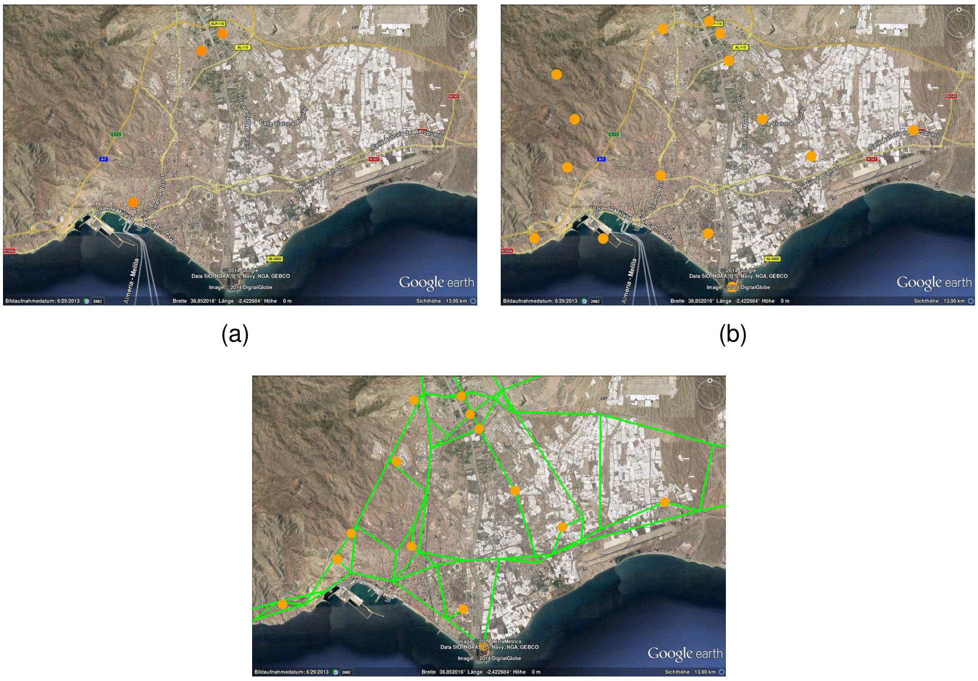


Figure 6: Population frequency of the Spanish municipalities.



(a) (b) (c)

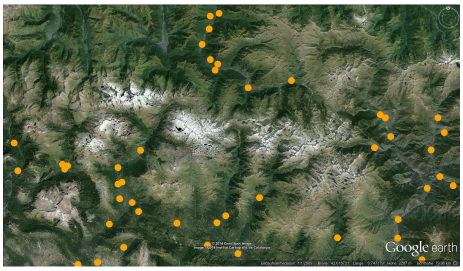
Figure 7: Distributing COs around the municipality of Almería: (a) The original coordinates (b) The distributed COs (c) The COs in the final network. Notice that the final network only contains those COs that could be connected to the (street) network.

municipality coordinate with this number of customers. See Figure 7. Recall that in the final network generation (Figure 7(c)) the COs are shifted or connected using street structures (Step 2 of network generation, see above).

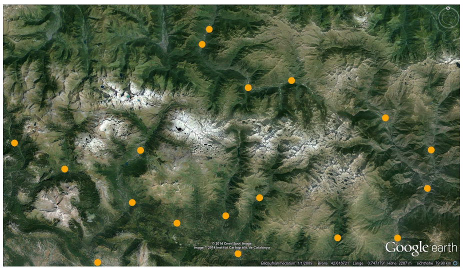
After the splitting operation we have a total of 11,259 (potential) COs. All COs have not more than 15,000 customers. However there are still many municipalities (potential CO coordinates) left with very few inhabitants (customers). In the merging operation we first increasingly sort the municipalities by their number of inhabitants. We then iteratively merge municipalities with less than 1,000 inhabitants with the closest neighbor municipality. The neighbor municipality inherits the inhabitants from the small municipality. However, the total number of customers should not exceed the number 15,000. As this process is repeated, a CO might get shifted several times until the resulting CO has more than 1,000 customers.

Since in today's copper networks the distance of a customer to its CO is typically not more 3-4 km and at most 10 km, we apply the merging in a very conservative way. We do not allow single shifts over more than 5 km. Moreover, customers should not get assigned to a new CO (municipality) more than 10 km away (The shifts should not sum up to more than 10 km), see Figure 8. We keep track of the shifts also to know about the maximum distance of a customer to its CO in the final network.

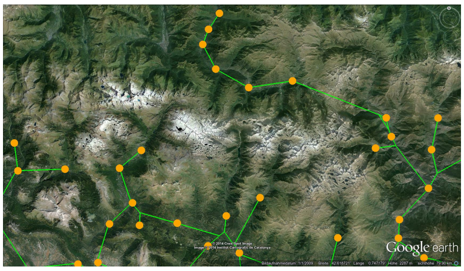
The merging returns with 8,897 CO locations. See Figure 9 for the frequency distribution of the CO customers across Spain. Notice that the resulting distribution is similar to the one of Italy and the one for the UK for which real data had been provided. These 8,897 locations with their coordinates enter the network generation algorithm described above. There are many COs that cannot be shifted to a (close) street structure since the coordinates are essentially randomly generated. Such COs are removed. The final Spanish reference network has 8,272 CO locations embedded into a topology of 18,819 nodes and 26,479 links, see Figure 5



(a)

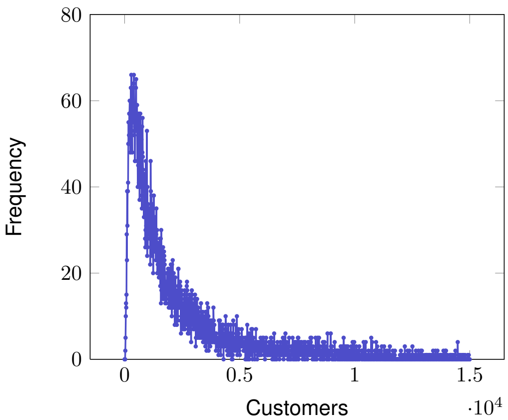


(b)

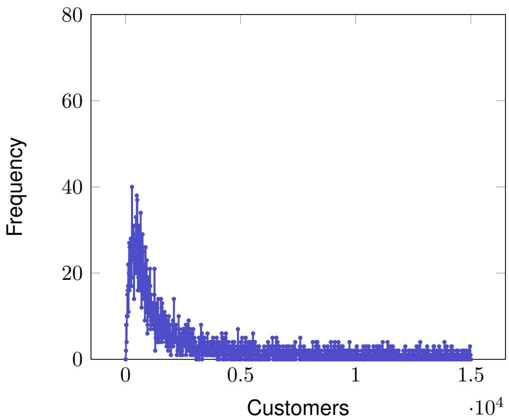


(c)

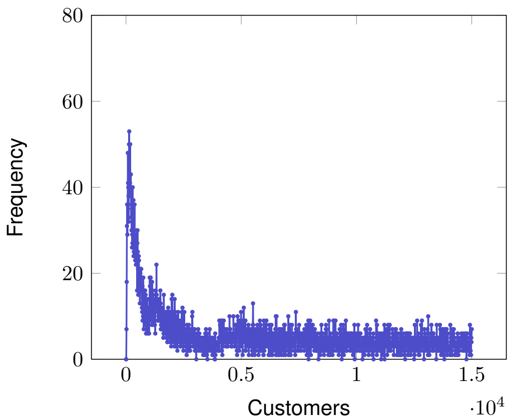
Figure 8: Merging municipalities somewhere in the Pyrenees: (a) The original coordinates (b) The COs after (distributing and) merging (c) The COs in the final network.



(a) Italy



(b) UK



(c) Spain

Figure 9: Frequency distribution of customers at COs in the reference topologies,

2.1.2 Microscopic: Reference networks for ODNs

In principle the approach described in the previous section can also be used to compute regional networks. What has to be changed is the set of anchor locations. In case of the nation-wide topologies we used geographical locations of central offices. In case of regional networks we will use coordinates of buildings. Moreover the level of detail has to be changed. While for the nation-wide topologies we used information from the highest OSM layers (highways, first level streets, etc.) we will have use more fine-grained structures such as residential roads to compute regional reference topologies. The quality of the resulting networks depends strongly on the quality of the data (even more than in the case of nation-wide coarse grained topologies). In the following first we show regional networks around Bologna and then focus on how to build the reference network for entire Ireland.

Regional networks around Bologna

We found that the OSM-data for Bologna is relatively high quality with many details regarding the location of address points along residential roads. In Figure 10 we show two preliminary reference networks for the center of Bologna in Italy and a rural area south of Bologna. We will use these networks to do studies on the deployment of optical distribution networks. Every address point will be considered being a point of fiber demand (a customer).

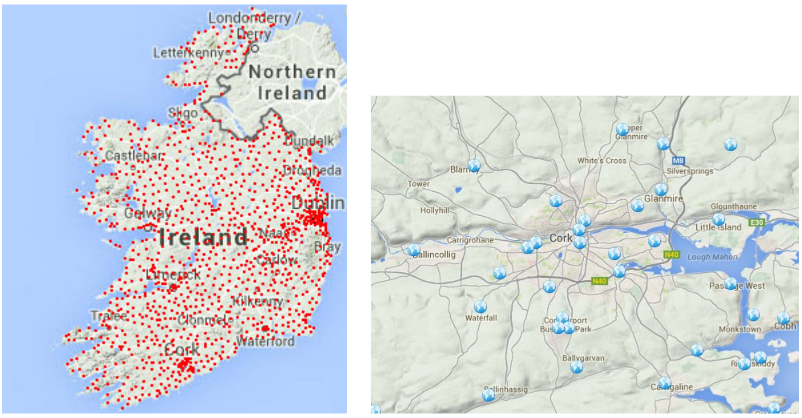
(a) (b)

Figure 10: A microscopic regional reference network: (a) Bologna urban, (b) Bologna rural; Each of the orange points corresponds to an address point (a potential customer). The green lines are residential streets or connections from the street to the building.

A Microscopic Network for Ireland

Also in the case of Ireland we combine OSM data with additional sources of information to come up with a network that contains nation-wide microscopic view down to the level of individual customers. However, the data sources differ from the approaches above. For Bologna street information and building data has been taken directly from the same OSM data set and exchange sites have been copied from the Italian reference network. For Ireland the road network data is still coming from OSM, similar to the above approaches. The customer (building) data, however, is coming from GeoDirectory Ireland (`https://www.geodirectory.ie`), and the locations of Exchange Sites data has been provided by Eircom (`http://www.eircomwholesale.ie/Our_Network/`).

The customer, or building data for Ireland is a simple data consists of latitude and longitude pairs. There are 2,189,121 customers from GeoDirectory. Figure 11 visualizes the exchange site data from Eircom on the map. Clearly, in rural areas exchange sites are more sparse. We remark that to handle the massive amount of different data we had to develop special techniques for storage and access. In particular, we made use of relational databases.



(a) (b)

Figure 11: LE sites (a) in Ireland and (b) in Cork

We now describe how to associate the location of buildings with OSM road network data and how to assign customers to LE sites within the network, see Figure 12 As they are from different sources, a relationship need to be defined between these separate data. First of all we associate buildings with closest road nodes as shown with blue lines in the Figure 12(a). This is identical to the approach taken for the Bologna networks, see Figure 10.

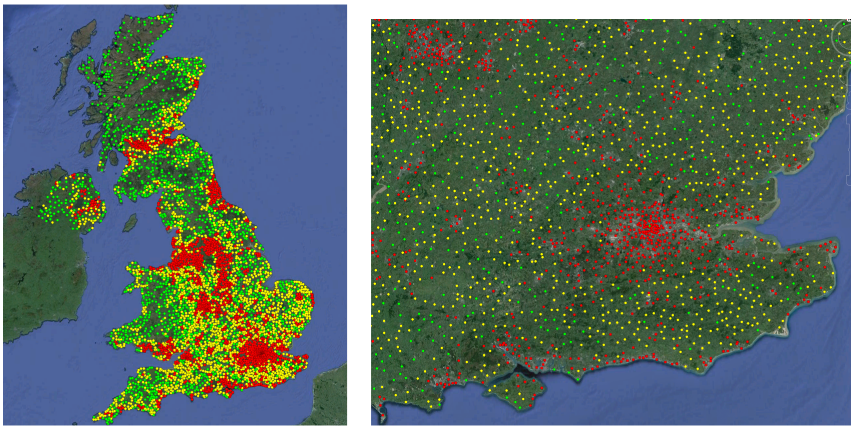
Once the customers and exchange-sites are put on the OSM map, we need to associate customers with their closest exchange-sites. We decompose this problem into smaller subproblem to avoid traversing the entire network numerous times. We create a small network for each exchange-site as follows:

1. Match each customer to their closest exchange sites by Euclidean distance and find the farthest customer per site.
2. Use this Euclidean radius to create a bounding box with latitude and longitude coordinates per site
3. For each exchange site extract the customer and road node data into separate text files within the bounding box
4. Compute the actual distances using Dijkstra's shortest-path-tree algorithm [11] for each exchange site

Doing these steps allowed us to divide the problem into smaller problems for parallel and distributed processing of the data in chunks. This clearly results in multiple matchings for some customers. We used the shortest actual distance for matching the customer to his LE.

(a) (b)

Figure 12: Reference network: (a) Buildings mapped to the network (b) An LE with associated buildings



(a) (b)

Figure 13: Classification of geotypes for British reference network - Overview and detailed view for south-east of England. *(red=urban, yellow=sub-urban, green=rural)*

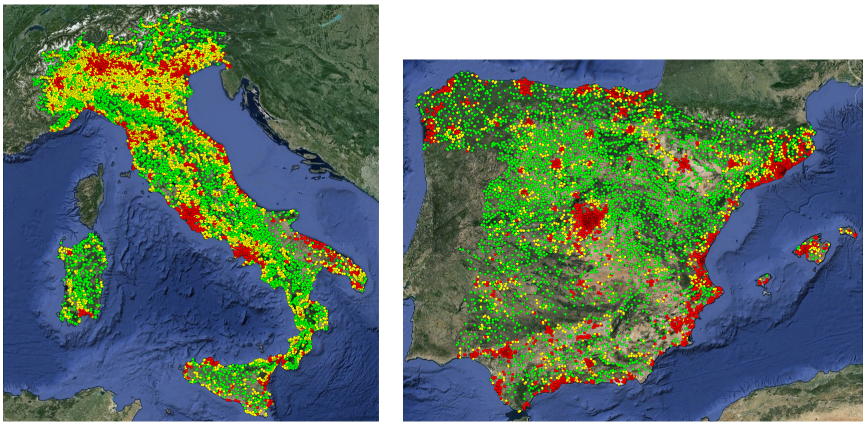
2.2 Geotypes by population densities

For choosing the appropriate LR-PON architecture at a given LE site it is necessary to have an understanding of the corresponding *geotype*. The geotype refers to the customer density in the area covered by the LE, that is, how many customers have to be connected and what is the potential extension of the optical distribution network (the customer furthest away) in this area. Based on the geotype we may decide about cascading LE topologies or ring structures, see for instance D2.1 [1].

Typically, geotypes are based on population densities, see [14]. We follow a similar approach here and introduce possible classifications of LE sites. As the geotypes are not provided as input data for our reference networks we will have to estimate them from available data. To this end, it is reasonable to classify all given LE locations into the three geotypes *rural*, *sub-urban* or *urban*. A more detailed classification, would be possible, but however, even for these three types it is not a priori clear how this classification has to be done and where the thresholds are.

The nation-wide topologies of Italy, Spain, and the UK as introduced above provide the LE locations plus the notion of distance in terms of possible fiber connections. Moreover, we know about the number of customers connected to each of the local exchanges. In contrast for Ireland we know the exact customer locations. Therefore we propose the following procedures to classify the geotype of the local exchanges L. Procedure (1) is applied to Italy, Spain, and the UK, whereas Procedure (2) is applied to Ireland.

This allows for the following classification of LE sites:



(a) (b)

Figure 14: Classification of geotypes for Italian and Spanish reference network. *(red=urban, yellow=sub-urban, green=rural)*

(1) Geotypes by network distances and LE customers

Given a local exchange LE $l \in L$ site, we collect all neighbor LE sites reachable within a distance d in km. In this case, distance refers to the length of the shortest path in the given reference topology (Italy, Spain, or UK). Let us denote this set of LE sites by $LE(d, l)$. Now we determine the total number $n(d, l)$ of customers connected to all LE in $LE(d, l)$. This can be done for every individual LE l. By introducing urban and suburban thresholds λ_1, λ_2 we may then classify the local exchanges into urban ($\lambda_1 \leq n(d, l)$), sub-urban ($\lambda_2 \leq n(d, l) < \lambda_1$), and rural ($n(d, l) < \lambda_2$). In addition we analyze the number of customers $cust_l$ served at the exchange site l.

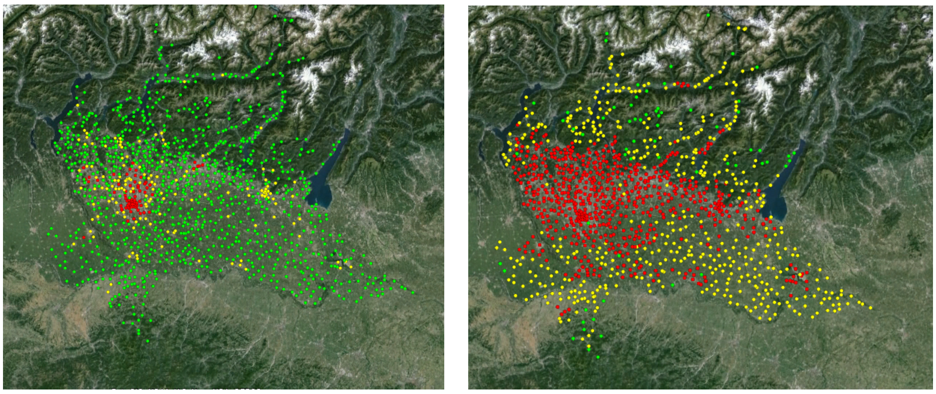
- **Urban:** LE serving a large number of customers ($cust_l \geq 18000$) and LE that have a large neighborhood ($n(5\ km, l) \geq 50000$ or $n(10\ km, l) \geq 100000$)
- **Sub-Urban:** LE l with (a)($cust_l \geq 2000$ and $n(10\ km, l) \geq 10000$) or (b)($n(10\ km, l) \geq 2500$)
- **Rural:** all other LE.

Figure 13 and Figure 14 shows the classification for Italy, Spain and the UK. In Figure 13 you see as well a more detailed zoom into the UK-network (south-east of England, containing the London area).

It should be noted that the classification is rather sensitive to the setting of the threshold values. Figure 15 shows possible classifications of LE sites for different thresholds using this approach.

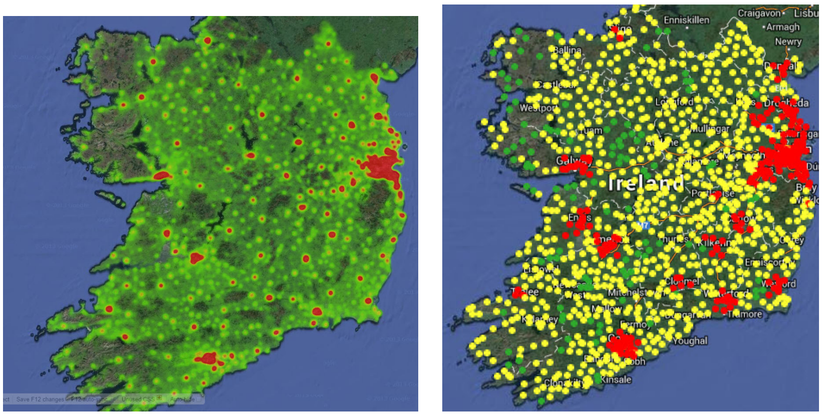
(2) Geotypes by grid cells and population density

We have generated the population density of Ireland by the data we have, as shown in Figure 16. Data helps to classify the density of the area easily. Our formula to classify if an



(a) (b)

Figure 15: Geotype classification in the region of Lombardia (Italy), green is rural, orange is sub-urban, red is urban. The classification depends on parametrization: (a) radius $d = 5$ km, thresholds $\lambda_1 = 13,000, \lambda_2 = 4000$. (b) radius $d = 10$ km, $\lambda_1 = 12,000, \lambda_2 = 1000$.



(a) (b)

Figure 16: Ireland population density (a) Google Maps heatmap feature without Exchange Sites (b) our algorithm indicating Exchange Sites

exchange-site is in rural, suburban or urban area is (i) Split the map of the country into X-cell grid (e.g. 10,000-cell grid); (ii) For each cell, assign the number of customers in that cell by latitude longitude; (iii) For each Exchange, calculate how many of the grid cells contain how many customers; (iv) The density of an exchange-site $l \in L$ is then defined by

$$\frac{\sum_{c \in C} TC(c,l) \times TCC(c)}{TCL(l)}$$

Here C is the set of all cells, E is the set of all exchanges, $TC(c,e)$ is the number of customers of exchange e in cell c, $TCC(c)$ is the number of customers in cell c, and $TCE(e)$ is the total number of customers of exchange e.

Our algorithm normalizes Exchange Site scores into an [0-1] interval. Then, in this case, we choose urban as [0.1, 1], suburban as [0.01, 0.1), and rural as [0, 0.01).

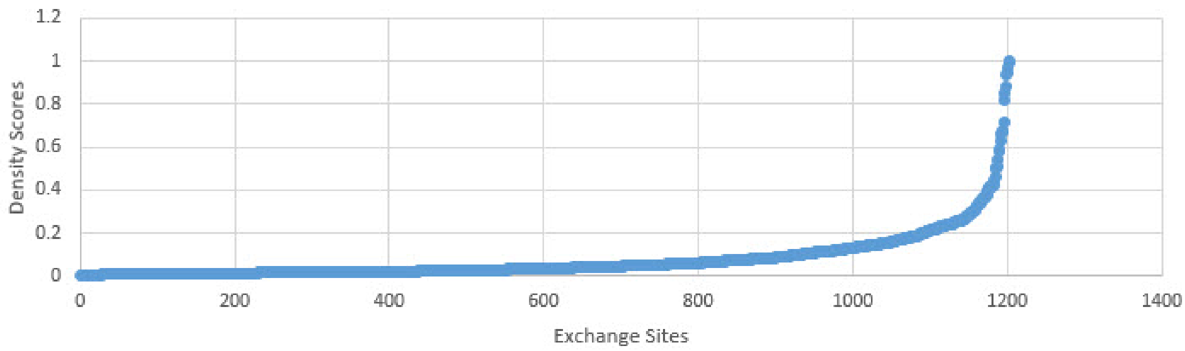


Figure 17: Ireland Exchange Sites vs Density Scores

2.3 Demand, traffic, and matrices

In this section, we wish to introduce the demand scenarios used for the optimization activities within DISCUS. Let as first clarify the terminology used within this section.

If we speak of *traffic* we refer to concrete (estimations of) data flows (typically in Mbps or Gbps). Traffic depends on the point in time, the user and service profiles, and also the corresponding network element (individual user traffic at the ONU or aggregated traffic at the MC node).

In contrast, if we speak of a *demand* we refer to a value against which network capacities are to be dimensioned. Typically, such a value refers to peak traffic (over time).

Node traffic (node demand) refers to traffic (demand) emanating at a particular network node (e.g. an MC node). Point-to-point traffic (demand), in turn, refers to traffic (demand) between two network nodes. A traffic matrix (demand matrix) states all pairwise point-to-point traffics (demands) between a set of network nodes.

Depending on the network layer and the network aggregation level there are different notions of demand.

In a multi-layer environment, there are demands at all the different network layers, e.g. IP demands in Mbps asking for optical channels, optical channel demands asking for fibers, or fiber demands asking for cables. Typically, higher-layer capacities create lower layer demands.

Regarding the optimization activities within DISCUS mainly the following demand types enter the optimization models:

Fiber demand

The DISCUS architecture assumes a fully transparent optical network between the customer and the MC node, that is, a network without optical-electrical conversion and without data-processing. We assume that every customer unit (a residential customer, a business unit) gets a single ONU with a single PON fiber connection. That is, the majority of customers has a fiber demand of 1. However, there is (a relatively small) number of larger business customers that would require dual parented protection form their premises. These customers would have more than one fiber termination at their premises and hence a fiber demand greater than 1. The ODN as well as the backhaul/metro network (towards the MC node) is dimensioned against all user fiber demands.

IP demand and demand matrices

Every user emanates individual user IP flows in Mbps (covered by the channel capacity transported over the single fiber mentioned above). These flows get groomed at the MC-node routers before entering the core network. The sum of all individual user traffic creates an aggregated MC node traffic. Because of the effect of statistical multiplexing, that is, user traffic profiles get merged at the MC-node routers, we do not have to assume that all user traffic profiles admit their peaks simultaneously. Instead we may sum up user traffic over longer time horizons (e.g. 1 hour) creating smooth traffic profiles (e.g. over a day). The MC node IP demand is then given by the peak traffic hour.

The network core has to be dimensioned against an IP demand matrix scenario. In the following we will explain how we set up MC node demand matrix scenarios based on individual user traffic scenarios.

2.3.1 Gravity models

Typically, point-to-point traffic statistics or forecasts are not available but a total traffic volume can be estimated. In this case, so-called gravity models can be used to compute point-to-point traffic. All gravity models are based on the assumption that the amount of traffic between two network nodes depends on the aggregated population at these nodes. Typically, the two values are multiplied. Very often, gravity models allow to include distance and service type information into the traffic computation, see [12].

We use a gravity model to compute inter-MC node traffic and demand that is not terminated at data-centers or peering-points, see also 2.3.2. In our model we ignore any distance dependency because we believe that most peer-to-peer services considered here are not (or only partially) distance dependent. The only exception to this rule is Voice over IP (VoIP), which we may ignore since it has relatively low bitrates, see below.

Let us in the following assume for every MC node i we are given the total (peer-to-peer) traffic demand d_i in Mbps that is generated at i. This value is obtained for instance from individual user demand multiplied with the number of customers at the respective MC node, see 2.3.2. Given individual values d_i we denote by $d(M) := \sum_{i \in M}$ the sum over the set of MC-nodes M.

First, we have to subtract from d_i the traffic that is reflected at the MC-node i, that is, traffic remaining below the MC-node not entering the core. The reflected (peer-to-peer) traffic at i increases with the relative size of the node given by the ratio $d_i/d(M)$. It is given by:

$$d_i^{ref} := d_i \frac{d_i}{d(M)}$$

such that for the core traffic generated at i it holds:

$$d_i^{core} := d_i - d_i \frac{d_i}{d(M)}$$

For every pair i, j of MC-nodes, we generate a non-negative point-to-point demand value $d_{ij} \in \mathbb{R}_+$ in Mbps as follows:

$$d_{ij} = d_i^{core} \times \frac{d_j^{core}}{\sum_{v \in M, v \neq i} d_v^{core}}.$$

Clearly, it holds that

$$\sum_{j \in M, j \neq i} d_{ij} = d_i^{core} = d_i - d_i^{ref},$$

that is the sum of the point-to-point demands corresponding to node i gives the node demand at i minus the reflected traffic.

2.3.2 User demand scenarios

Let us first review the current status of traffic modeling as described in the Deliverable D2.4 [6]. The traffic generator developed in D2.4 can be used to generate core traffic matrices reflecting traffic demands between given MC nodes, see next section. We essentially use functionality of the traffic generator to compute user demand scenarios.

The service profile

The traffic model assumes a given set of *services* S. Each service $s \in S$ has a name or type (e. g. "Internet surfing", "multi-media", "pure-data") and is defined by its source/target in the network and its (upstream/downstream) capacity in Mbps. Clearly, depending on the traffic flow direction (up or downstream) the notion of traffic source and traffic target changes. However, by convention, the end-user will always be the target of the traffic while the service source is given by one of the following:

- Internet exchanges (IX) also called peering points
- data centers (DC),
- or other end-users (P2P standing for peer-to-peer).

(See [6] for further details). In the following we will ignore the fact that up- and downstream-capacity might differ and always assume the maximum of the two as the given *service traffic* c_s in Mbps (which in most cases refers to the downstream). Notice that capacities (in the metro and core parts of the network) are typically installed bidirectional, while up and down-stream flows take the same network paths, so that we can ignore the smaller of the two values here.

Peering points as well as data centers are always located at MC nodes, see D2.3 [5].

The user profile and expected user demand

Once traffic source and target as well as service traffic are specified, one could in principal aggregate possible traffic flows across a given network topology. However, services are not used constantly at their maximum service traffic by all users simultaneously. The timing of individual users with respect to their usage of a particular service is called the *user profile*, see [6]. It specifies, for each service and different types of end-users (residential, business), a detailed statistical usage behavior (number of sessions, timing) over a typical day.

Based on the user profile for a particular service $s \in S$ we can compute its *expected service demand*. We compute this value as follows:

- We average user profile traffic over many users which gives an expected user profile.
- We average the expected user profile traffic over every hour of the day, which removes very local peaks
- We take the peak traffic hour and use it as the expected demand.

We will distinguish two different types of users: residential and business (with different expected user profiles). Let us denote by d_s^b the expected service demand for business users and by d_s^r the expected service demand for residential users, respectively. These expected demand values can be seen as averaged "mean"-demands over time and over many users. They are clearly only meaningful if many users are connected (and averaged) at the MC nodes (or LE nodes) assuming the effect of statistical multiplexing.

For simplicity we will also write d_{DC}^b, d_{DC}^r referring to the sum of all expected service demands with source DC (business and residential). Similarly, d_{IX}^b, d_{IX}^r denotes the expected IX demand per (business, residential) user and d_{P2P}^b, d_{P2P}^r the expected user P2P demand.

Notice that most of the complexity of the traffic generator lies in the user profile and in computing expected demand values based on these profiles.

Scenarios

A demand scenarios is simply a table that defines for residential and business users and the target types *DC*, *IX*, *P2P* the corresponding aggregated expected user demand:

Target	User	Value in Mbps
DC	residential	d^r_{DC}
IX	residential	d^r_{IX}
P2P	residential	d^r_{P2P}
DC	business	d^b_{DC}
IX	business	d^b_{IX}
P2P	business	d^b_{P2P}

Table 1: A demand scenario

As mentioned above, these values are obtained from the individual services, such that a demand scenario could also be stated by providing for every considered service:

- the service type/name ("Internet surfing", "pure data", "video")
- the service source ("DC", "IX", "P2P")
- the expected business user demand d^b_s
- the expected residential user demand d^b_s

Different demand scenarios are obtained by applying different assumptions on the service profile (up/down stream traffic), the difference between the services, and the user profiles (timing).

Notice that a demand scenario is completely independent of the network architecture or solution. Instead, we may compare different architectures, different solutions, and different resilience using the same demand scenario.

2.3.3 Computing a demand matrix based on a demand scenario

Next we show how to compute a demand matrix based on a demand scenario. A demand matrix is based on a demand scenario and a *network profile*. This is in line with the computation of traffic matrices in D2.4. However, the demand matrices here are based on expected user demands not on traffic. Moreover, we provide just one matrix while the traffic matrix generator in D2.4 states traffic matrices for each time slot (with different granularity's, e.g. 5-60 minutes slots).

We will explain how to compute an MC to MC demand matrix. However, the same approach can be used to compute matrices at different aggregation levels, e.g. an LE to LE matrix. We should take care with the number of users though. As mentioned above, expected user demands and expected node demands only make sense (can be used to dimension capacities) if the number of aggregated users is reasonably large.

The network profile

The network profile defines the number of MC nodes in the core network and the number of users (residential, business) connected to each of the MC nodes (primary MC). It also specifies which of the MC nodes is an internet exchange and which of the MC nodes acts as a data center. Here we use the same input as for the traffic matrix generator described in D2.4 [6].

Output and conventions

Demand matrices are evaluated before optimizing the core network every time a concrete network profile is available. Notice that this approach adds some flexibility. While service and user profile are relatively fixed, network profiles change with the considered country, architecture, resilience strategy, etc.

Based on service, user, and network profile, we compute an expected demand between any two specified MC nodes. Once a demand scenario and a network profile is given this computation is straight-forward. Let us assume the number of residential and business users at a particular MC node is given by $n^r > 0$ and $n^b > 0$, respectively. Then, the total DC demand d_{DC} of this MC node is given by

$$d_{DC} := n^r \cdot d^r_{DC} + n^b \cdot d^b_{DC}$$

Similarly,

$$d_{IX} := n^r \cdot d^r_{IX} + n^b \cdot d^b_{IX} \quad \text{and} \quad d_{P2P} := n^r \cdot d^r_{P2P} + n^b \cdot d^b_{P2P}$$

give the total IX and P2P node demand. To obtain dedicated point-to-point demands, in the computation, all traffic with source DC or IX is equally split among the given data centers and peering points, respectively. To compute peer-to-peer traffic demands between MC nodes, we use the gravity model defined above, see Section 2.3.1. That is, the total demand with source type P2P is split among the MC nodes based on the number of connected users.

In addition to the traffic that is derived from the different end-user services the traffic model in D2.4 assumes a certain amount of inter data center traffic. The model assumes a percentage of 41% of the traffic sourced at DC nodes as being inter data center traffic and adds it to the traffic matrix. We can do the same for demand matrices.

The traffic model for leased lines can be seen as a P2P service with a special user profile for business users, so it is included in the above modeling, that is, there is an expected service demand d^b_s for the leased line service s.

Resiliency

If we aggregate traffic or demand from users at their primary MC we assume the normal operating state. Depending on the resiliency and restoration concept we get a different traffic/demand matrices in every failure state. A failure state can be a failing MC node or a failing LE to MC connection. Depending on the resilience concept different numbers of users get mapped differently to different MC nodes. As the computation above solely depends on the number of users connected to an MC node we can handle any failure states if we know which user is sending its traffic to which MC node.

2.3.4 DC and IX scenarios

As mentioned above, we assume data centers and peering points to be co-located with metro-core nodes. However, there are many options for deploying data centers and peering points across a country. In fact there is a trade-off between the amount of core traffic and the number of specialized MC nodes with additional equipment. We will assume the following scenarios.

- *Single DC,IX*: There is only one DC and only one IX for every optical island. We will chose the largest MC node (based on the number of connected customers) as having DC as well as IX functionality.
- *All DC,IX*: Every MC node becomes a data-center and peering point.
- *Multiple DC,IX*: We assume that 5% and at least two of the MC nodes become IX locations and 30% of the MC nodes get data-center functionality. These nodes are chosen randomly favoring large MC nodes.

2.3.5 A demand scenario for 2015

in DISCUS a user demand scenario for 2015 was developed that is based on the framework described above and uses the traffic generator presented in Deliverable D2.4 [6]. It defines user profiles for the *moderate private user*, for the *versatile private user*, and for the *small business user* and it provides service profiles for the services *E-social*, *Video on demand* (VoD) either Standard Definition (SD) or High Definition (HD), *Voice over IP* (VoIP) , *Video Conferencing* (VC) either Standard Definition (SD) or Low Definition (LD), *File Sharing*, *Online Gaming*, and *Data Backup*. The service E-Social comprises online chat, internet surfing, email exchange, data exchange, news update, info sites (for example weather forecasts), internet banking as well as E-commerce services.

User	Service	Type	Mean # sessions	Mean duration in h	Traffic in Kbps
Moderate private	E-Social	IX	4.67	0.27	500
	VoD SD	DC	1.00	0.10	2,000
	VoIP	P2P	9.14	0.05	200
Versatile private	E-Social	IX	4.67	0.27	500
	VoD HD	DC	1.00	1.50	5,000
	VoD SD	DC	4.00	0.10	2,000
	VC LD	P2P	1.14	0.50	500
	VoIP	P2P	9.14	0.05	200
	File Sharing	P2P	0.47	0.50	4,000
	Online Gaming	P2P, DC	1.14	0.83	5,000
Small business	VC SD	P2P	1.14	0.75	1,000
	E-Social	IX	4.67	0.27	500
	Data Backup	DC	0.95	2.50	5,000
	VoIP	P2P	9.14	0.05	200

Table 2: User profiles for different users and services

The service characteristics for the different users are presented in Table 2 and 3. Table 2 provides the user profiles for the different services, that is, for every user profile it gives the mean number of sessions the service is used during a typical day, the mean duration time of such a session, and the average generated traffic. The latter is based on the downstream bitrate as it typically exceeds the upstream bitrate. Table 3 shows traffic during the peak hours of the day averaged over the services for different targets in the network.

Peak to Mean ratio

If we define the peak hour to be between 10 and 11 p.m., then the last column in Table 3 is used to compute the the user demand scenario for 2015.

As mentioned using the given averaged traffic demands in Kbps above to dimension network links and ports can only be done if the aggregation level is high, so many users are incorporated. If the number of users is small (e.g. at the LE or even closer to the customers) then

User	Target	2-3 a.m.	10-11 a.m.	9-10 p.m.	10-11 p.m.
		Traffic in Kbps			
Moderate private	P2P	0	22	12	10
	DC	0	0.6	65	56
	IX	0	0.15	44	44
Versatile private	P2P	176	86	1,300	1,700
	DC	176	6	2,700	3,300
	IX	0	39	31	31
Small business	P2P	0	185	37	36
	DC	4,400	0	0	0
	IX	0	677	5	2

Table 3: Averaged traffic in peak hours for different users and network targets

there might be a point in time within the chosen peak hour that sees much more traffic than the above demand multiplied with the number of users. This is because we have averaged the traffic over one hour. That's why network operators typically apply a peak-to-mean ratio to over-provision averaged traffic estimations or traffic measurements, see also [19]. A good factor clearly depends on the level of user aggregation.

For simplicity we will assume a peak-to-mean ratio of 2 for our user demand scenarios. That is, we multiply all values in Table 3 by two to cope with unexpected traffic peaks.

In case no population numbers are available for the different users we assume a distribution of 85% moderate users, 10% versatile users and 5% small business users, which gives the averaged user demand scenario in Table 4.

User	Target	Traffic in Kbps	%
Average	P2P	361	30.1
	DC	755	63.1
	IX	91	6.8
	Total	*1197*	*100.0*

Table 4: User demand scenario for 2015

2.4 Cost modeling

In this section we will introduce a tentative hardware and cost model that we will use for end-to-end optimization and cost evaluation. The cost model is based on the techno-economic modeling activities in WP4 and WP7 for the access and core part of the DISCUS architecture, respectively. The MC-node and core model is in large parts based on the capex-models developed in the STRONGEST and IDEALIST projects, see [17] and [4]. However, we adapted the model following the DISCUS architecture. Cost-tables can be found in Appendix A

For many reasons it is not feasible and also not desired to work with a highly-detailed equipment model in optimization. First of all the hardware structure in practice is typically changing quickly. Secondly, these models may complicate the optimization and shift the focus from the main cost-drivers. For example, from the optimization point of view it is not important to model any cost that is proportional to the number of customers and that is independent of the solution (e.g. ONUs, customer termination, etc.). This cost can be neglected although it might be interesting in cash-flow analysis or techno-economic modeling. Similarly, using a very fine-grained cable and fiber model (e.g. blown fiber tubes, blown fiber units, splicing) might just complicate if the network cost in the access is actually dominated by digging.

The cost model we provide below will be simplified even more depending on the optimization model and the focus of the optimization sub-problem. For example, if we decide about the location of MC nodes and the assignment of LEs to MCs, see Section 3.2 and 4.2, we will ignore the LE cost and might use a simple MC hardware model (e.g. a single setup cost value). For this particular optimization problem it is important to have a good understanding of the relation between the assignment cost (fibers and cabling) and the MC node cost that is independent from the number of connected customers.

In network optimization we typically work with node and link *designs*. Link designs are installed at network links and node designs are installed at network nodes. These designs refer to equipment blocks with certain properties that provide and consume resources (fibers, ports, slots). Every design has a certain cost that might depend on its size. For example a cable (which is a link design) provides a certain number of fibers and consumes duct space. Its cost depends on the number of provided fibers and the length. Similarly, a splitter (which is a node design) provides fiber ports and consumes housing space.

In many cases, in particular in case of a one-to-one correspondence, it is feasible to combine different network elements and work with designs that refer to blocks of equipment, e.g. we might include the cost for housing in the cost for a splitter.

The cost parameters to be used for optimization have been aligned with and derived from the cash flow models for which detailed cost models of network elements and subsystems are developed and utilized. Deliverable D2.8 will contain more details in this direction.

The analysis of cost and power consumption is an ongoing process within DISCUS. As individual cost values are still subject to potential change we decided to provide the tables in the appendix of this deliverable. Tables 15, 16, and 17 in Appendix A present the tentative cost and hardware for the LR-PON, the hardware at the MC node, and optical equipment, respectively. All cost values are given both in EUR and in the IDEALIST cost unit (ICU). For tables 16 and 17 we partially reused values from the cost model developed in IDEALIST (which in turn is based on the model from STRONGEST).

2.4.1 LR-PON cost

Table 15 in Appendix A summarizes the hardware model for the LR-PON. It shows the different equipment design elements and provides cost either per piece of equipment (splitters) or per km (in case of cables or ducts). It also states the area in the LR-PON where the equipment is installed. The corresponding areas can be found in Figure 15 which gives an overview of the LR-PON structure.

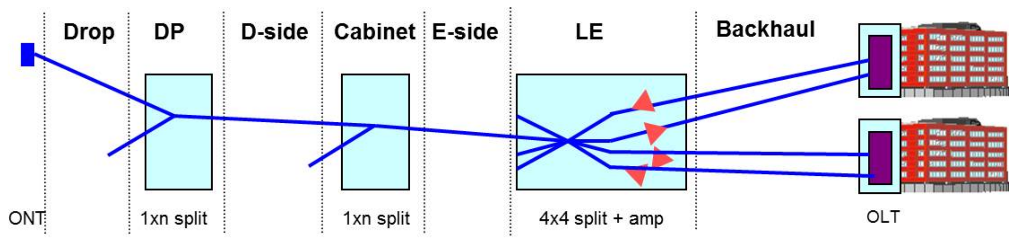


Figure 18: Abstract hardware model for the LR-PON

For the cabling we use cables of different sizes and distinguish cables used in the backhaul or between Cabinet and LE (E-side) and cables used between DP and cabinet (D-side) and in the Drop section. In the first case, the model assumes classical cables while in the latter case the model combines blown fiber tubes and blown fiber units and different sizes of BFU). The cost for splicing is assumed to be contained in the cable cost. Depending on the focus of the optimization problem it may be feasible to work with a simple fiber cost per kilometer instead of the detailed cable model, see Appendix A.

The duct cost per kilometer includes the cost for digging and duct deployment. Clearly, not in all areas of the network we will have to provide ducts. In optimization, we will use different duct deployment factors (percentage of links with empty ducts available), which will depend on the country (UK, Ireland, Italy, Spain), the LR-PON area (backhaul, E-side, D-side, Drop) as well as the geo-type, see Section 2.2.

All splitters include the cost for housing, optical splitter ports, splicing, and splicing trays In case of the DP and cabinet splitter we assumed a 12-fiber cable being attached. The 4x4 amplifier node includes the 4x4 splitter, the amplifier itself, and the amplifier housing.

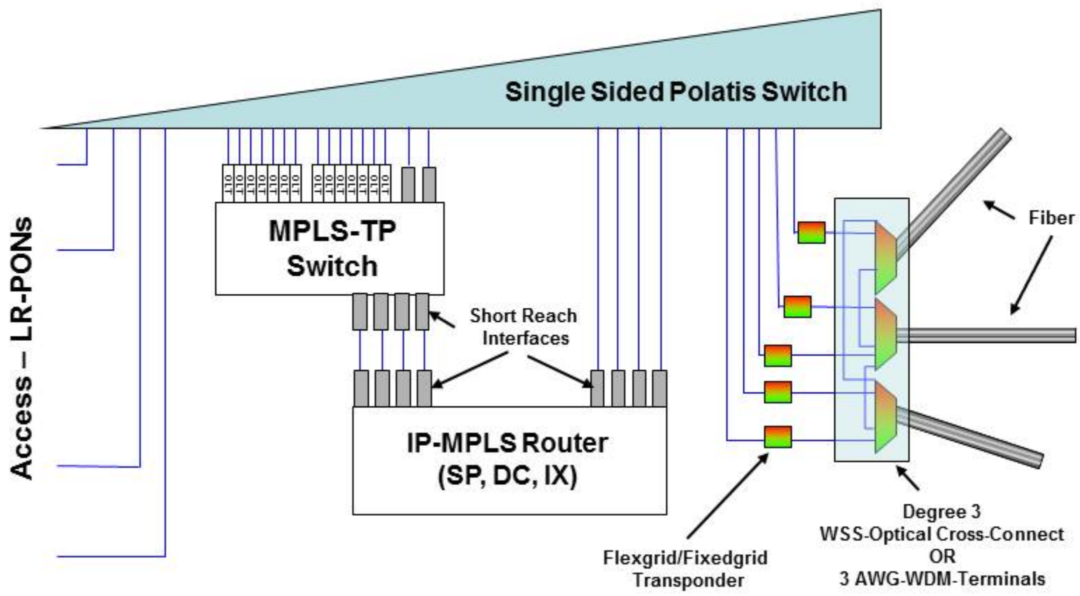


Figure 19: Simplified MC-node model

2.4.2 MC node switching cost

The MC node hardware model in Table 16 aims at providing cost values for the equipment at the MC node as defined in D2.3 [5], see Figure 19. As mentioned above, for most of the MC-node hardware we partially reused the models developed in the IDEALIST and STRONGEST projects and published in [4] and [17], respectively. In particular we copied the models for IP-MPLS (router, cards) and MPLS-TP (switches, cards), except for the OLT-cards. We only reused the 2014+ equipment and restricted ourselves to 40G, 100G, and 400G ports.

The node model assumes the L2, L3 routing and switching elements to be organized in three levels: chassis, cards, and transceivers, see Figure 20. The chassis is characterized by its capacity in terms of slots, the cards in terms of throughput and type and number of ports, and

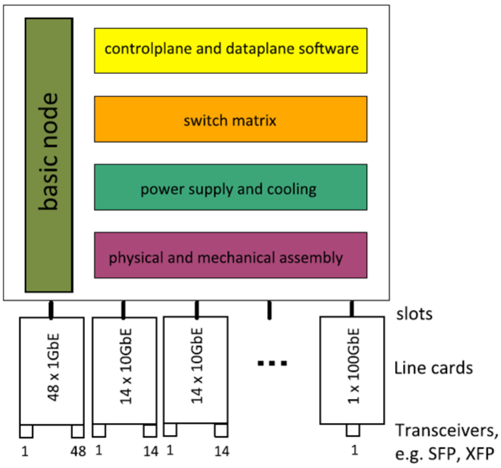


Figure 20: IP-MPLS or MPLS-TP switching element at the MC-node

the interfaces in terms of their client rate. Line-Cards are supposed to require one slot in the chassis. We assume 400G slots for IP-MPLS as well as MPLS-TP. All cards may be equipped with short reach transceivers which in turn can be connected to transponders. We do not assume colored long reach transceivers here for simplicity. In fact, we may assume a colored interface at the cost of a short reach interface plus a transponder. For the cost of transponders see the optical equipment below.

We will assume MPLS-TP switches at all MC nodes. For simplicity, in optimization we will not distinguish core-side and access-side switches. We introduce OLT-cards to terminate the access side LR-PONs at the MPLS-TP switch, see Table 17. These are not provided in the STRONGEST model [17]. We will assume IP-MPLS routers and equipment only at service nodes, that is, IP-nodes of service providers (SP), data-center (DC), or internet peering points (internet exchanges, IX), also see Section 2.3.

We remark that IP-routers with a slot count of more than 16 are multi-chassis routers. There is a significant cost increase from the single chassis router (16 slots) to the 2-chassis router with 32 slots. However the cost for larger routers then increases roughly linearly and can be computed with a simple formula based on the number of required slots, see [4] and [17].

2.4.3 Core optical equipment cost

Optical core equipment consists of the following element blocks:

- Polatis Switch: This switch is part of every DISCUS MC node. It switches access side and core-side optical fibers. Its cost depends on the number of fiber ports.
- Optical Line Amplifiers (OLA) have a fiber span of 80 kilometers. They need to be installed for core fibers with a length of more than 80 kilometers. All nodes already include line amplification.
- Digital Gain Equalizers (DGE) have a fiber span of 320 kilometers and are to be installed for all fibers accordingly. All nodes include DGE functionality.
- Transponders (Fixedgrid and Flexgrid) are used on both sides of every core optical channel in combination with a short-reach transceiver interface at the MPLS-TP switches. Depending on the signal, transponders have a certain signal-reach. The reach can be extended by regenerating the signal which is done using regenerators.

- AWG terminals are used to multiplex fixed-grid optical channels on a single fiber. The switching of fixedgrid signals is done using the Polatis switch.
- WSS-based optical cross-connects (including mux/demux) are used to switch flexgrid signals in the core.

Table 17 in Appendix A aims at providing the optical equipment cost. Again we reused values and formulas from IDEALIST and STRONGEST (Terminals, OXCs, OLAs, DGEs, Fixedgrid transponder). However, we introduce a model for the Polatis switch and the flexgrid transponders have been aligned according to the 6 flexgrid signal types with 37.5 GHz spacing defined inf D7.2 [3]. The single-sided Polatis switch is assumed to have a setup cost and a cost depending on the number of fiber ports, see Appendix A.

To model cabling and fiber cost in the core network we can make use of the cable model for the E-side and backhaul from Table 15 or a simple fiber cost per kilometer. However, typically dark fibers and cables can be assumed to be available such that cabling is not a cost factor for core deployment. The same applies to ducts.

3 An end-to-end optimization process

Optimizing an entire nation-wide DISCUS network down from the level of individual customers as a one giant monolithic network optimization problem is clearly beyond the focus of the optimization activities within DISCUS. It is not only challenging but has also the disadvantage of being inflexible. Using a global end-to-end cost function, one optimization would result in a single end-to-end network solution. In order to study alternatives and to understand the influence of different network parts with respect to the overall network cost and performance the same huge optimization problem had to be solved over and over again with different parametrizations of the objective cost function.

Moreover, fine grained end-to-end solutions on a nation-wide scale are impossible to compute simply because most of the data is missing or confidential (location of customers and buildings, nation-wide duct-systems, etc.). As described in Section 2 data to study the deployment of optical distribution networks is available only for smaller regions (or smaller countries such as Ireland) while nation-wide reference networks are available only starting from the level of a local exchange.

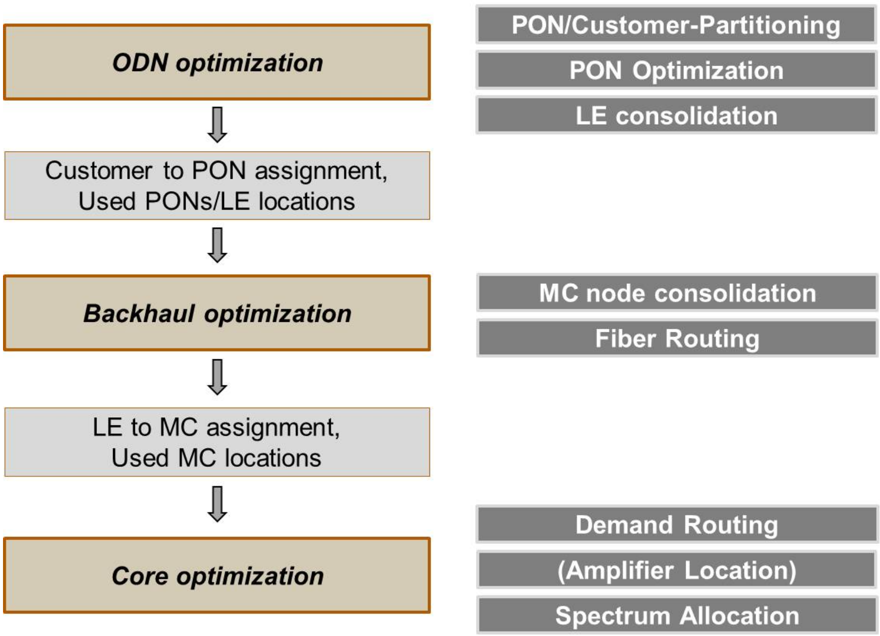


Figure 21: End-to-end optimization process: Global view

For these reasons we subdivide the optimization problem into three main sub-problems

- ODN optimization (for some regions): customers and their connections to local exchanges via sequences of optical splitters,
- Backhaul optimization: local exchanges, where the last amplifiers are placed and the first splitting takes place, and their connection to metro-core nodes via disjoint fiber routes
- Core optimization: metro-core nodes and their connection among each other

This increases the flexibility and allows to study local as well as global correlations. We clearly follow a bottom-up approach, see Figure 21, where first we optimize the optical distribution network reusing most of the existing infrastructure between customers and their exchange-sites. Based on its output we determine LE to MC assignments determining the number and location of the MC nodes and cable deployment optimally. Finally we focus on the dimensioning of the core network where we try to ensure that MC nodes are connected with each other so as to ensure the forwarding of estimated end-to-end traffic demands.

For each of these sub-problems we need to make several decisions by solving one or more combinatorial optimization problems. As each subproblem is solved independently decisions made in one network may impact decisions taken in another part of the network. Therefore is important to investigate different criteria and parametrizations. In the following sections we will introduce the corresponding optimization problems in more detail and present a high-level optimization framework based on combining the results of subproblem optimization

In this chapter, We do not present results but describe the process that will be followed throughout the rest of the DISCUS project. The end-to-end optimization process aims at providing evaluations of end-to-end solutions to compare alternatives of the DISCUS architecture. Furthermore, it also aims at comparing the DISCUS architecture with existing networks or current deployments, such as architectures based on GPON. Comparisons are done using the same nation-wide reference networks and the same demand scenarios, see Section 2.1 and Section 2.3. We also use a common component catalog, see Section 2.4.

Combining local with global solutions

Only in case of Ireland we have nation-wide data down to the level of customers. In case of Italy, Spain, and the UK we are able to optimize LR-PON (or GPON) deployments for a reasonable number of smaller reference regions, see Section 2.1.2 and Section 4.1. This can be done comparing alternatives of different LR-PON architectures (e.g. dual homing, versus ring structures, or cascaded local exchanges) depending on the customer density and customer to LE distances, that is, depending on the geotype at a given local exchange, see Section 2.2. We will classify the given smaller regions based on their customer density as rural, or "urban" and will optimize the LR-PON structure exploiting this geotype, see 4.1.

Studying the influence of architecture alternatives, we will come up with an understanding of the cost of the DISCUS network below a given LE as a function of the number of connected customers and the geotype. That is, we may derive an average cost per customer of the ODN network for rural, suburban, and urban LEs. Clearly, these values will be verified with the techno-economic modeling and cash-flow analysis in WP2 and deliverable D2.8.

Now given a list of LEs together with the number of connected customers, which we have for instance in case of Italy, Spain, the UK, and Ireland, and a geo-type classification of all these LEs into rural, suburban, and urban we are able to compute a nation-wide cost of the ODN network.

Based on the same set of LEs we will optimize the backhaul deployment and the core network, see Section 4 as well as deliverables D4.5 [7] and D7.4 [8]. Evaluation of the backhaul plus core cost together with the cost of the ODN we end up with an estimate of the cost for a nation-wide deployment of the given architecture. This can be done for LR-PON with flat optical core (DISCUS) but it can, in principal, also be done for a GPON based architecture.

Moreover, the presented high-level process allows to study alternative deployments, e.g. with smaller or larger numbers of MC nodes, or comparing the DISCUS architecture to a two-tier architecture where a larger number of MC nodes is connected to a second aggregation level of central core node.

Given a series of different optimized solutions we will compare performance indicators such as cost and availability, see Section 2.4 and Section 3.4.

3.1 Optical Distribution Network

The LR-PON uses optical amplification to support greater total split which means that it is possible to have multiple split stages to further increase infrastructure sharing and minimize cost per customer. In particular we will be focusing on minimizing the costs results from:

- Splitter placement
- Fiber, cable and duct deployment

The possible locations for splitting points will be the junction points based on the reference network, and existing distribution points (DP) close to the customer premises, primary cross connect (PCP) or cabinet locations, typically less than 1km from customers, and the local exchange or central office site where the optical amplifiers will also be located as this node has electrical power available.

We exploit the relation between the number of splits and the maximum fiber length from the local exchange-site to a customer for designing optical distribution network. This would allow to create PONs of different sizes for different geo-types. In general the length of fiber from exchange-site to a customer can go out as far as 10km. However in sparse rural areas the ODN reach may need to be extended.

The relevant decisions that can affect the cost are

- How many PONs are required for a set of customers associated with a given exchange-site? What should be the sizes of PONs in different scenarios (e.g., rural and urban)?
- Where to install splitters? What should be the size of each splitter? What splitter hierarchy should be applied?
- Which routes should be used for fiber/cable deployment?

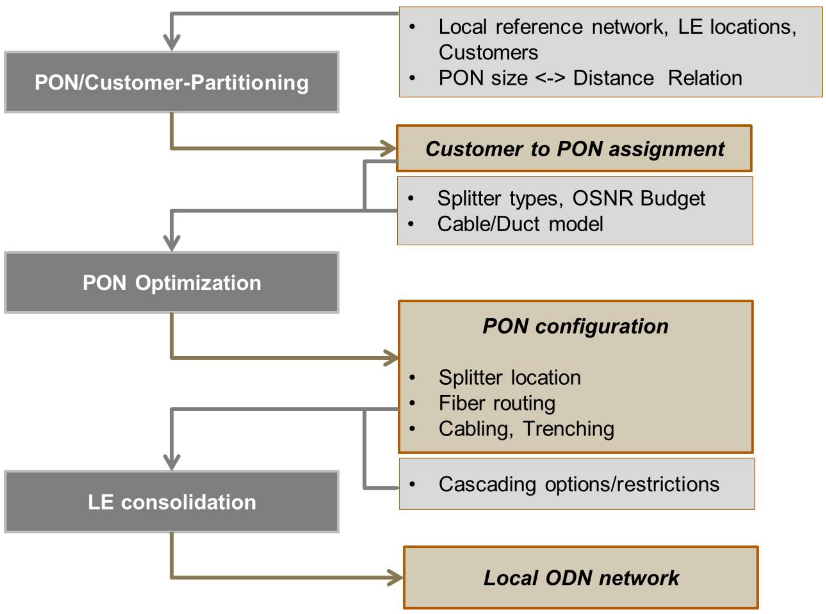


Figure 22: ODN optimization: Detailed view

Given a local exchange-site and the number of customers associated with it, the problem is to construct a set of optimal PONs. We decompose this problem in to three sub-problems as shown in Figure 22:

- **PON/Customer Partitioning.** The first subproblem is to estimate the number of PONs for each exchange-site by partitioning the set of customers such that each subset of customers will be associated with one PON. Notice that here we would like to take in to account different geo-types which would have an impact on the sizes (i.e. number of customers) of PONs.
  - Input: Reference network with customers and LE locations, relation between size of PON and the maximum LE to customer distance e.g.,((512,10),(256,20),etc.), and fill-factor to restrict the utilization of each PON.
  - Output: Number of PONs for each LE, Size of each PON, assignment of each customer to a PON.
- **PON Optimization.** The second subproblem is to find the locations of a set of splitters required to connect a given set of customers and their paths to the exchange-site. Here the assumption is that the first-level splitter in the splitter-hierarchy is always placed in the local exchange-site.
  - Input: Reference network restricted to the set of customers associated with a PON of a LE, Maximum distance to LE, possible splitter sizes, OSNR Budget
  - Output: Routes from Customers to exchanges, the locations of splitters, and the sizes of splitters.
- **Exchange-sites Consolidation.** In rural area, exchanges might be supporting only few customers and therefore it might be associated with very few PONs. In order to increase the sharing of expensive equipment and fibers it might be possible to consolidate these exchanges by connecting PONs of nearby exchanges in a cascaded manner. The result would be that not all exchanges would be directly connected to metro-core nodes. Consequently, the third subproblem is to consolidate some of these local exchange-sites, where possible, by moving the locations of the first-level splitters based on either cascaded PON configuration or open ring configuration as described in Section 4.1.
  - Input: All LEs, PONs of all LEs (optional), Specification related with open ring and cascaded PON configuration, Parameter to filter LEs which should not be considered
  - Output: Chains of PONs, Which LE require direct connection to MC cascaded) configuration,

3.2 Backhaul Network

Backhaul optimization (see Figure 23) aims at calculating the number of required MC-nodes, their locations and the assignment of the LE sites to the MC nodes, respecting the given resilience concepts, thereby optimizing the fiber-paths. This overall problem is currently decomposed into two sub-problems, namely the *MC node consolidation* and the *optimization of fiber routes*. (For a detailed discussion of the problems see as well Deliverables D4.5 [7] and Section 4.2.)

MC node consolidation

Given a network and a set of local exchange sites L, we calculate a subset of these to serve as MC-nodes. We calculate as well the assignment of the LE to a pair of MC and a possible maximally disjoint path-routing.

- Input: Reference network (Italy, Spain, GB) with LE locations, maximum LE to MC distance (e.g., 90 km, 100 km, 115 km)

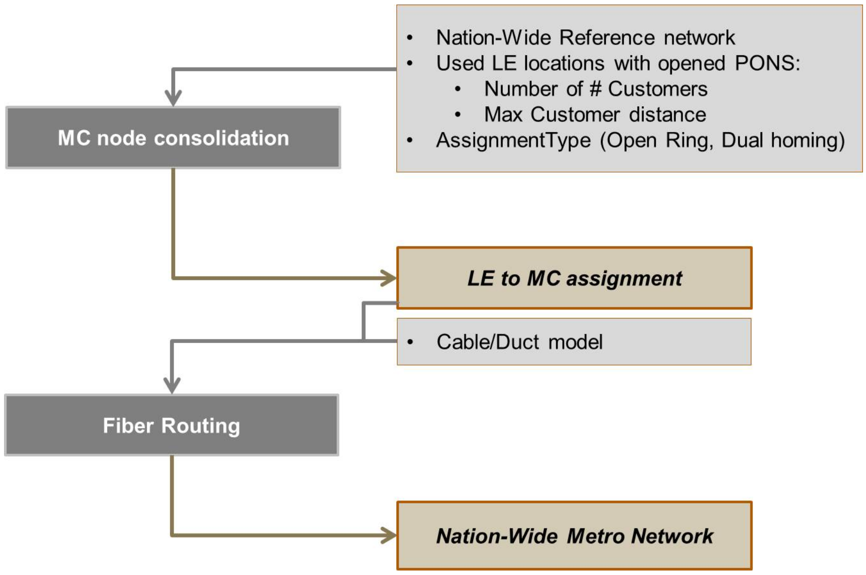


Figure 23: Backhaul optimization: Detailed view

- Output: Set of MC nodes in the network, Assignment of each LE to two MC by two (maximally) disjoint fiber routes

In this optimization-step we guarantee to choose nodes that are reasonable to the next step of core optimization (i.e., we restrict our self to nodes that have a certain connectivity level, etc.) In the description above we use a dual-homing strategy. It might be worth mentioning that other strategies like (single-homing, dual-homing with not necessarily disjoint paths, etc.) might be applied as well.

Fiber route optimization

The solution to the *MC node consolidation* calculates maximally disjoint path between the LE and its corresponding MC. The paths are calculated locally, i.e. for each (LE,MC,MC)-triple and does not take synergetic effects by sharing sub-paths into account. This will be done in *fiber route optimization*.

- Input: Reference network (Italy, Spain, GB) LE to MC assignment, maximum LE to MC distance (e.g., 90 km, 100 km, 115 km), cable and/or duct capacity/cost model
- Output: Set of fiber routes (two disjoint for each LE) such that cable sharing is maximized (cable cost is minimized)

MC node consolidation and fiber route optimization

In the remainder of the project we will investigate possibilities to solve both sub-problems in one.

- Input: Reference network (Italy, Spain, GB), maximum LE to MC distance (e.g., 90 km, 100 km, 115 km), cable and/or duct capacity/cost model
- Output: Set of MC nodes in the network, Assignment of each LE to two MC by two (maximally) disjoint fiber routes. Cable sharing considered

In fact, we can also look at the backhaul optimization being a single problem with different objective functions. The main task is to open MC nodes and to assign LE site to these opened

MC nodes. Now there are different concurrent criteria: (i) minimize the number of MC-nodes (ii) maximize disjointedness (iii) minimize distance based cost (by sharing cables), or (iv) minimize over-provision (see D4.5 [7]). Clearly, all these criteria or subsets of them can be combined into different multi-objective optimization problems.

3.3 Core Network

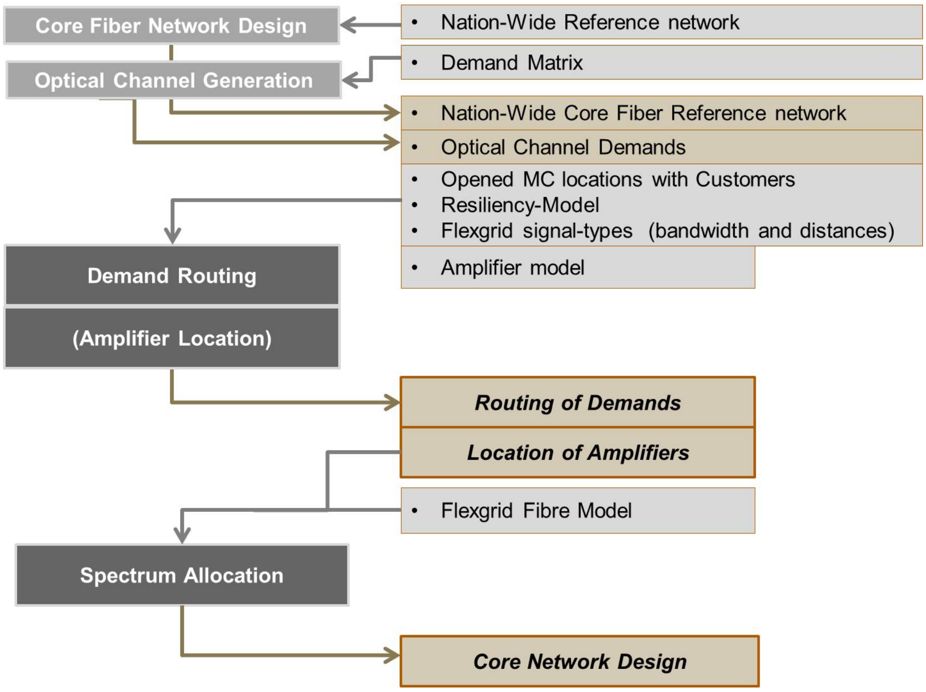


Figure 24: Core optimization: Detailed view

In this section, we specify the optimization problem associated with the interconnection of metro-core nodes in term of its input, constraints and expected output for a single transparent optical island, single translucent optical island and a resilient optical island.

In a transparent optical island it is desired that traffic stays in the optical domain between source and target of the traffic demand. Within a single optical island the structure of the virtual IP-layer is fixed, a full mesh. Also the virtual IP-routing is fixed, traffic is routed directly from source metro-core node to target metro-core node. It turns out that the overall problem for a single optical island reduces to a (single-layer) network design problem corresponding to the optical layer only. We briefly list out the outputs, constraints and objective involved in this problem. The following is the list of output for the core network:

- Determine the physical links for connecting all pairs of metro-core nodes
- Determine the number of optical channels and their types for each traffic request
- Routing path of each optical channel
- Slot and fiber allocation for each optical channel of each traffic-request

An optical island needs to respect the following constraints:

- Each pair of metro-core nodes must be connected and the length of the shortest path can not exceed the maximum signal reach.

- The traffic request in Gbits between all pairs of metro-core nodes must be satisfied. That is, the signals established between the demand end-nodes should have a sufficient bitrate capacity in Gbits.
- The signal type and the assigned spectrum (bandwidth interval) must be consistent over its entire fiber route.
- The limit on the distance that each signal can travel must be respected.
- The capacity constraints of the node hardware (e.g. maximal number of connected fibers, maximal number of switched channels) must be respected if any.

The main objective is to minimize the capital expenditure by minimizing the cost of physical links based on fibers/ducts/cables and amplifiers, and the cost for node hardware (OXC, regenerators, etc.). Another objective could be to design the core network such that energy consumption is minimized.

The main solution process is described in Figure 24. We are given the potential nation-wide network topology (Italy, Spain, UK, Ireland) as provided by Section 2.1 including a number of selected MC nodes and assumptions on the possible signal types as well as the resiliency. Moreover, we are given a demand matrix between these MC-nodes as computed in Section 2.3.

Before actually optimize the routing and spectrum allocation we have to consolidate the input. First we cannot start directly from the given fiber topology unless we believe that each of the links in this network may carry a core fiber, which is typically not the case. Instead we take the nation-wide street topology as input to a core fiber reference network design problem which results in a sub-network containing all potential core fiber links:

Core Fiber Network Design: In the network design subproblem we focus on minimizing the total lengths of physical links required to connect metro-core nodes such that the shortest paths between pairs of metro-core nodes is less than the maximum signal reach. The reachability constraint ensures that the traffic can be sent between pairs of metro-core nodes without electronic-optical-electronic conversion. The objective is to minimize the cost of the physical links. Another objective could be to minimize the sum of the lengths of the paths between all pairs of nodes, and optionally each path could be multiplied with the value of the traffic. This in some sense tries to minimize the total lengths of fibers.

- Input: Reference network with MC locations, maximum signal reach constraint
- Output: Pairs of MC nodes which are required to be physically connected in the optical island.

Secondly, the demand between any two MC nodes is typically larger than the capacity of a signal channel. To come up with a set of channel demands between any two MC nodes we solve a straight-forward Optical Channel Generation Problem:

Optical Channel Generation: Given the network and the traffic matrix the optical channel generation problem is to determine the partition of the traffic in to a set of optical channels. Each optical channel is associated with a signal type subject to the constraint that there exists a path between the pair of nodes whose length is not greater than the reach of the signal. In this problem one can either minimize the number of channels or minimize the number of slots required for these demands.

- Input: Core network containing MC locations, traffic matrix, flexgrid signal types

- Output: Partition of traffic in to a set of channels for each pair of MC nodes either based on minimizing number of slots or number of channels, or cost of the transponders required for each channel type.

At the center of the core optimization is a Routing and Spectrum Allocation problem including the location of Raman amplifiers:

- **Routing.** Given a network and a set of optical channels the problem of routing is to find a route for each optical channel in the network from its source metro-core node to target metro-core node such that the length of the path does not exceed the signal reach associated with the optical channel.
  - Input: Network containing MC locations and links between them, flexgrid signal types, optical channels associated with each traffic-request
  - Output: Route for each optical channel subject to signal reach.
- **Spectrum Allocation.** Given a network, a set of optical channels and a route for each channel, the spectrum allocation problem is to find assignment of slots to channels in a fiber subject to continuity and contiguous constraints. The objective is to minimize the cost related with the number of fibers. Among many optimal solutions, we might wish to prefer those that use the spectrum in the most efficient way. In a certain sense we want a solution that leaves enough consecutive bandwidth slots for future capacity expansions.
  - Input: Network containing MC locations, flexgrid optical channels, association between channels and routes
  - Output: Spectrum assignment for each channel.
- **Raman Amplifier Location.** A single transparent optical island might not be always be feasible for some large European countries because the distance between one or more pairs of metro-core nodes could be longer than the maximum optical signal reach. For such countries the goal would be to design a translucent optical island and use Raman amplifiers to increase the reach of the signals. Apart from the objective defined for the previous sub-problems for transparent optical island, for translucent island one would also like to minimize the number of Raman amplifiers. Notice that one Raman amplifier is used for one fiber and the reachability of all the signals passing through that fiber would be extended.
  - Input: Core network containing MC locations, traffic matrix, flexgrid signal types, association between channel types and routes, OSNR Tables, OSNR thresholds for different modulation techniques
  - Output: Locations of Raman amplifiers for optical channels whose routes are more than maximum signal reach and identifying fibers that require Raman amplification. The objective is to minimize the number of Raman amplifiers.
- **Resiliency.** Survivability to metro-core node or link failure is very important for optical island as a single failure could disrupt the communication between millions of users. Therefore it is important to design resilient optical island that can survive in case of a failure of a node or a link. When designing resilient optical island we want provide node disjoint paths between all pairs of metro-core nodes so that each pair of nodes are connected when a failure occurs. We would also like to make sure enough spectrum is available so that all demands affected in case of a failure can be allocated slots in their alternate paths for transmission.
  - Input: Reference network with MC locations, flexgrid signal types, traffic matrix
  - Output: Disjoint paths between each pair of MC-nodes and spectrum allocation for each traffic request. The spectrum will be allocated depending on the strategy

LR-PON	MTTR (hours)	MTBF (hours)
OLT Shelf / Backplane (18 tributary slots)	4	1,500,000
Power supply unit	4	500,000
Optical amplifier (EDFA)	4	1,000,000
OLT Downlink line card Hybrid-PON (10 x 10G)	4	400,000
Chassis of splitting point (not including splitter)	8	2,000,000
Arrayed Waveguide Grating (AWG)	8	4,000,000
Optical power splitter	8	7,500,000
Hybrid-PON ONU 10G tunable APD (hardware + software failures)	8	204,545
Polatis Optical Switch (192x192)	8	183,084
Fiber link (per km)		
Feeder Cable	24	1.350.000
Distribution Cable	24	1.500.000

Table 5: Input data for availability calculations

used for protection paths.

3.4 End-to-end availability

While network optimization is typically driven by cost, there are other measures to verify the quality of a network solution. One such performance indicator is the availability of connections.

In this section, we will first provide the definition of end-to-end *connection availability* and show how this measure can be computed, given a network solution and based on (un)availability times of network elements.

We will then focus on the access segment of the DISCUS architecture. Studying the proposed performance indicator, we will show that an alternative access architecture results in a better performance with respect to connection availability.

3.4.1 Definition of end-to-end connection availability for DISCUS architecture

Asymptotic *system availability* A_c is defined as the probability that a component c is operable at an arbitrary point of time, while asymptotic unavailability U_c is defined as the probability that the component is not operable. The system availability and unavailability can be expressed as in Equations (1) and (2), respectively:

$$A_c = \frac{MTTFF_c}{MTBF_c} \quad (1)$$

$$U_c = 1 - A_c = \frac{MTTR_c}{MTBF_c}, \quad (2)$$

where A_c, U_c give the system availability of component c, $MTTR_c$ stands for mean time to repair, $MTTFF_c$ denotes mean time to the first failure (represents the mean life time of the component/system) and $MTBF_c$ stands for mean time between failures (where $MTBF_c = MTTFF_c + MTTR_c$). Typically $MTTFF_c$ is much larger than $MTTR_c$.

See Table 5 for $MTTR$ and $MTBF$ figures for different components in the LR-PON of the DISCUS architecture.

End-to-end connection availability refers to the probability that a logical connection (of a user) is operable at an arbitrary point of time.

To compute end-to-end connection availabilities we have to compute the system availability of configurations of components. In general, there are two basic configurations for the connection availability calculation, namely series and parallel. The series configuration consists of two or more components (units) connected in series from the reliability point of view. It means that a series system fails if one or more components (units) fail. The parallel configuration consists of two or more components (units) connected in parallel from the reliability point of view. It means that a parallel system fails if all of the components (units) fail. Expressions for the system availability A_C for series or parallel configurations C are shown in Equations (3) and (4), respectively:

$$\text{Series configuration:} \quad A_C := \prod_c A_c \tag{3}$$

$$\text{Parallel configuration:} \quad A_C := 1 - U = 1 - \prod_c U_c \tag{4}$$

where A_c and U_c stand for availability and unavailability of a certain component c of the configuration, respectively. Typically, for the commercialized components, $1 > A_c >> U_c > 0$.

Clearly, based on the formulas (3) and (4) we may compute the system availability A_C of any component configuration and also parts of a given DISCUS solution.

DISCUS architecture builds on the concept of Long-Reach Passive Optical Network (LR-PON) in the access and a flat backbone partitioned into optical transparent islands. Thus, the end-to-end connection availability for DISCUS architecture should include both access and backbone segments. Accurately, one logical connection set up between two optical network units (ONUs) in DISCUS architecture should pass the LR-PON twice and the flat core once (see the orange solid line in Figure 25).

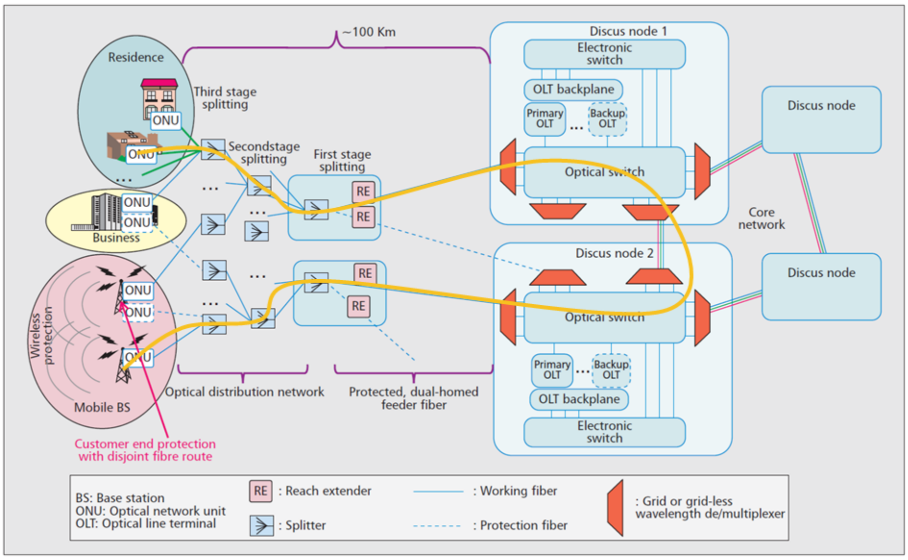
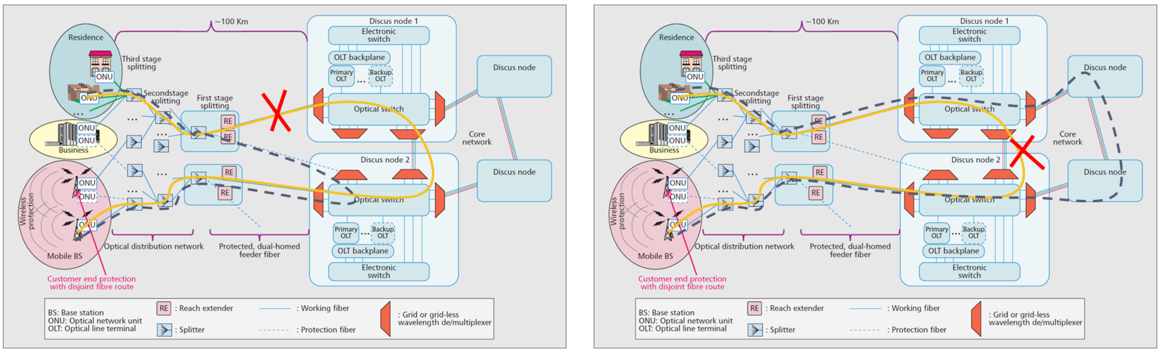


Figure 25: An example of a logical end-to-end connection in DISCUS architecture

In LR-PON of DISCUS architecture, dual-homed feeder fiber protection is always provided. Therefore, in case that the working feeder fiber is cut, the signal can be switched to the backup feeder fiber so that the connection between any pair of the ONUs will not be interrupted (see the example shown in Figure 26(a) where the dashed black line is the protection path keeping the connection up and running).



(a) (b)

Figure 26: Example of failures in the DISCUS architecture: (a) the failure occurs at feeder fiber of LR-PON (b) the failure accrues in the flat core of DISCUS architecture

The flat core of DISCUS architecture is based on mesh topology where survivability mechanisms are provided to deal with cable cuts and node failures in order to avoid connection interruptions due to the failures in the core segment of the network. Figure 26(b) shows an example where the connection is still on after a cable cut between two DISCUS metro/core nodes.

Given a DISCUS network solution S, let us now define the end-to-end connection availability $A(S)$ of this particular solution. We denote by O the set of all ONUs. Given two ONUs $i, j \in O$, the end-to-end connection availability for this connection can be expressed as follows:

$$A(i,j) := A^{i}_{odn} \cdot A_{feeder\_section+flat\_core} \cdot A^{j}_{odn}, \quad (5)$$

where A_{odn} represents the system availability of the unprotected parts of the two LR-PONs that the two considered ONUs i, j belong to, while $A_{feeder\_section+flat\_core}$ refers to the system availability of the part of DISCUS architecture where the resiliency schemes are implemented.

Now the expected end-to-end connection availability of a given solution S is clearly determined by the average over all possible connections:

$$A(S) := \sum_{i,j \in O} A(i,j) \quad (6)$$

This connection availability clearly depends on the concrete implementation of the discus architecture, that is, it depends on chosen fiber routes, distances, splitter hierarchies, etc.. And as shown above this measure is based on the individual system availability of components as in Table 5. In this respect, it is a solution measure and performance indicator similar to cost, which also depends on the concrete solution and cost values of individual components, see Section 2.4.

3.4.2 Analysis of Access Segment of DISCUS architecture (rural scenario)

It should be noted that for a connection that is not fully protected, the end-to-end connection availability is very much dependent on the availability of the unprotected part of the connection.

Therefore, the availability of the unprotected part can be seen as the upper bound of the end-to-end connection availability. Consequently, the resiliency provided in the core network doesn't need to offer much higher level of connection availability than the unprotected parts of the connections. For this reason, our further work on the survivable core network design will be dependent on connection availability limitations coming from the access part of the network. We will consider differentiated reliability scenario, where different services will be offered different levels of protection in the core segment, matching the availability level in provided in the access segment and according to the requirements specified in the SLAs.

As a first step we estimate the connection availability in the access segment.

We have already reported reliability analysis of LR-PON in ultra-dense and dense deployment scenarios in Deliverable 2.2. The conclusion was that the feeder section (including feeder fibers and OLT) should be protected in order to achieve sufficient connection availability (i.e., 99.99%) in the access segment.

In this deliverable, we focus on rural scenario and show the main conclusions in the next sub-section.

Here we consider connection availability in the access segment of DISCUS architecture located the basic rural scenario, i.e., we consider LR-PON between ONU and OLT located at DISCUS metro/core node. We take the basic rural deployment (shown in Figure 27) as case study. The MTTR and MTBF values for different components are collected among DISCUS partners and shown in Table 5.

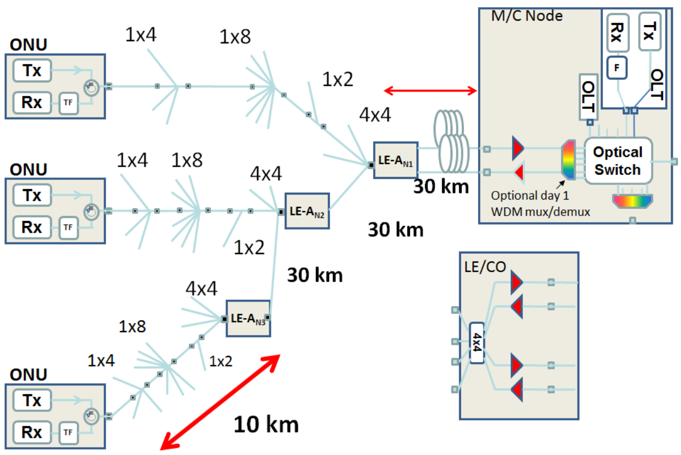
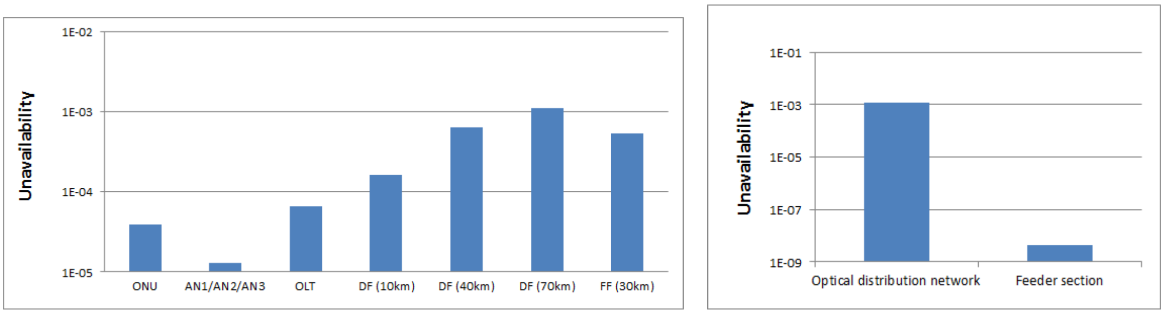


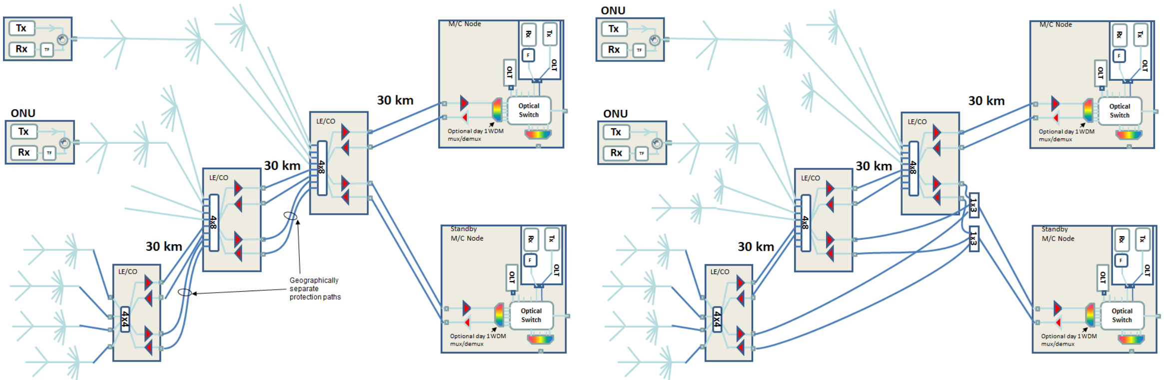
Figure 27: Distributed amplifier node solution for rural areas – basic concept

In the basic rural scenario (see Figure 27) dual-homed feeder section protection is taken into account. However, the total unprotected distribution fiber could have a length of up 70 km and several stages of amplifiers (active node) are needed in the field. Figure 28(a) shows the unavailability values obtained for network elements and fibers along the connection between ONU and OLT in the LR-PON. It can be seen that, the fiber links (including feeder fiber FF and distribution fiber DF) have the highest unavailability among all the network elements considered in the LR-PON segment. In Figure 28(b) the unavailability figures for the protected and unprotected parts of the access network are presented. In contrast to the feeder section where 1:1 protection is provided by dual homing, the optical distribution network in the basic rural scenario does not have any backup, leading to the low connection availability (lower than 99.9%, i.e., 3 nines) of the whole LR-PON segment, which is a consequence of the high unavailability of the 70km long unprotected fiber between the first amplifier node and the ONU. It should be also noted that with a segment having such a low availability the end-to-end connection availability cannot be better than 3 nines. However, some customers (in particular business users) may require a higher level of connection availability (which typically is included in service level agreement signed with the operators), in order to guarantee service continuation. In order to offer proper availability for LR-PON in rural areas, we would recommend to provide protection for DF, particularly for the DF between the first and second stage of amplifier nodes (i.e., A_{N1} and A_{N2} in Figure 27), which could affect up to more than 200 end users simultaneously. Figure 29 shows two possible options to provide protection to distribution fibers.



(a) (b)

Figure 28: Unavailability results: (a) each network element in access segment and (b) optical distribution network vs. feeder section



(a) (b)

Figure 29: Two options to provide protection for distribution fibers

4 Optimization and techno-economic studies for different geo-types

4.1 Optimizing LR-PONs, rural and urban solutions

In this section we focus on Optical Distribution Network of LR-PON consisting of local-exchange sites and customers. Figure 30 depicts the architecture for connecting a set of customers to the network with single fiber working in ODN and two fiber working in the Backhaul. Informally speaking, a passive optical network (PON) is a fiber tree network that uses point-to-multipoint fiber network in which unpowered optical splitters are used to enable a single optical fiber to serve multiple end-users. A PON consists of an optical line terminal (OLT) at the service provider's central office and a number of optical network units (ONUs) near end users. The starting point of a fiber in case of LR-PON is a metro-core node and the first split will usually occur in the exchange-site. An exchange-site can support tens of thousands of customers therefore it would also be the location where many first-level splitters for many PONs are going to be placed. To maintain dual homing each local-exchange site is connected to a primary and secondary metro-core nodes, and and therefore all PONs associated with the local-exchange site would be connected to the same pair of metro-core nodes. One of the problems is how to partition the customers of the exchange-site to create a set of PONs.

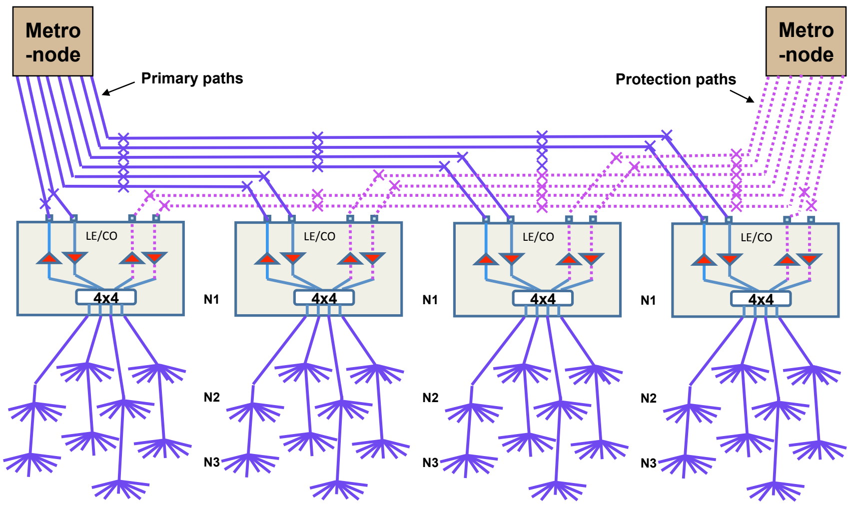
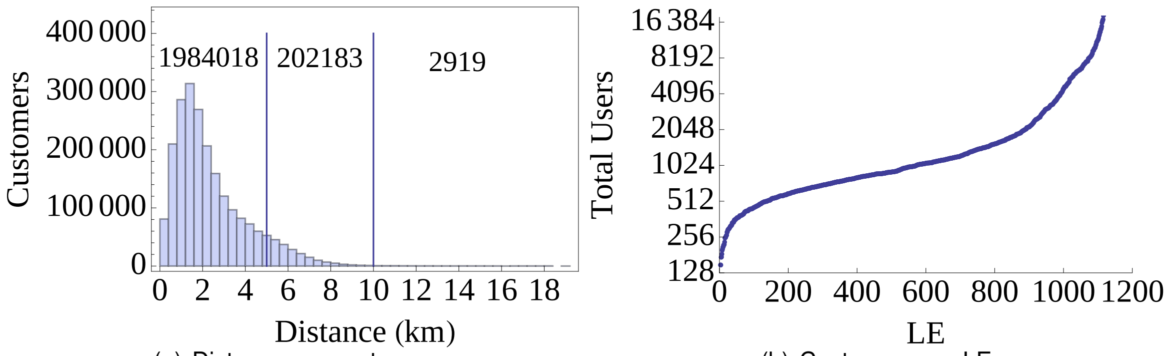


Figure 30: ODN based on cable tree model

The optical signal attenuation in a PON is due to the number of splits in the PON and the maximum length of the fiber between the exchange site and the customer. As signal attenuation is allowed up to some threshold, the upper-bound on the length of the optical fiber from the exchange site to any customer varies with respect to the maximum split size of PON (i.e., max. number of customers). In urban areas where an important number of customers are grouped within a few km ratio from the local-exchange site it is possible to design PONs of up to 512 customers. In rural areas local-exchanges might cover only few tens of customers in large and sparse areas.

Figure 31 depicts how customers are distributed in Ireland, the left figure (Figure 31(a)) shows the distance between local-exchange sites and customers, and the right figure (Figure 31(b)) shows the total number of users per local-exchange site. We use Euclidean distance of 2,189,120 customers and 1120 local-exchange sites. As expected the majority of the customers are grouped very close to the local-exchange site, indeed 1,984,018 customers are



(a) Distance per customer

(b) Customers per LE

Figure 31: Distribution of customers in Ireland

within 5 km radius, while on the other hand only 2919 customers are located more than 10 km far to the local-exchange. In addition, it can be observed that due to the geography of the country we observe urban areas with up to 26085 customers and rural areas with only 116 customers.

The problem is to optimally select a set of PONs by resolving the trade-off between the size of a PON and the maximum distance between a given exchange-site and its customer. For solving this problem we assume that we are given the following as an input:

- Network (containing locations of customers and local-exchange sites);
- customers associated with an exchange site;
- relation between split-size and maximum allowed distance;
- at least 20% of a given PON should be free for future customers.

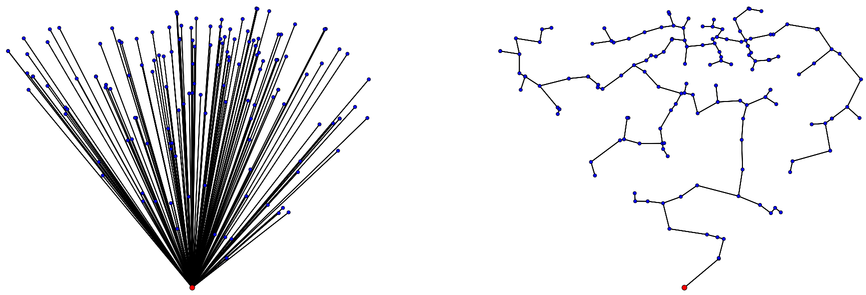
In order to solve this problem we partition the customers in to a set of trees

4.1.1 Cable Tree Model

A simple way of creating a set of PONs for a given exchange-site would to partition the set of customers by using the following steps:

1. create a sorted list of all the customers based on their decreasing distances to the exchange-site
2. select a PON size that is closest but greater than or equal to the distance of the farthest customer to the exchange-site
3. select a number customers from the front of the list that is equal to the number obtained by considering fill factor and the selected PON size
4. remove the previously selected customers from the list and go to step (2) if the list is not empty otherwise terminate

The problem with the above approach is that it may not consider locality of the customers. In the worst-case we may end up as shown in Figure 32(a) where each customer is connected to the exchange site without sharing any path. Certainly the option of connecting each customer directly to the exchange site leads to shorter connection paths. However, the drawback of connecting each customer directly is the total amount of fiber cable used. What we ideally want is to something like minimum spanning tree at the exchange-site as shown in Figure 32(b). Notice that this would maximize the sharing of paths. Certainly this option minimizes the total length of cable but the drawback is that we might be violating the maximum fiber length allowed between the exchange-site and any of its customers.



(a) (b)

Figure 32: (a) Customers are directly connected to an exchange-site. (b) Customers are connected to the exchange-site through a spanning tree.

We are interested in both restricting the length of the paths and the total amount of cable used. Keeping both requirements is known to be a hard problem [16]. There has been a significant amount of work on this bounded version of the spanning tree problem (see [15] for a short summary of the most relevant approaches). Constraint programming based techniques has also been suggested for tackling this problem [15, 10].

The Cable Tree Model (CTM) consists in building a distance bounded and capacity bounded tree networks such that each tree is rooted at a local-exchange site, each customer is present in exactly one tree and the total distance-based cost is minimized.

The number of trees denotes the number of optical fiber that would emanate from the local-exchange site towards the customers. The deterioration in the quality of the optical signal can not surpass a given threshold. Notice that the signal deterioration depends both on the maximum length of the fiber cable and the number of customers of a given tree. Therefore, when the former increases the latter decreases and vice-versa. The relationship between them is shown in Figure 33. In other words, in CTM we are given a set of customers, an exchange-site and a cost function, and the objective is to determine a number of fibers that start at the exchange-site and create a tree distribution network for each fiber such that each customer is included in exactly one tree, the total cost of all trees is minimized and both distance and capacity bounds are respected.

Constraint Optimization Formulation

In the following we present some notations, the formal definition of the problem and the constraint optimization model of the CTM:

- $G = (V, E)$ is a graph
- $V = \{v_0, v_1, \ldots, v_n\}$ is a set of vertices
- v_0 is the facility and $V \setminus \{v_0\}$ is a set of customers
- $E = \{(v_i, v_j) | v_i, v_j \in V; i \neq j\}$ is an edge set and an edge in G means that there exists a path in the ODN reference network that connects v_i and v_j without going through any v_k
- D is a matrix of distances where d_{ij} denotes the length of the path between customers v_i and v_j

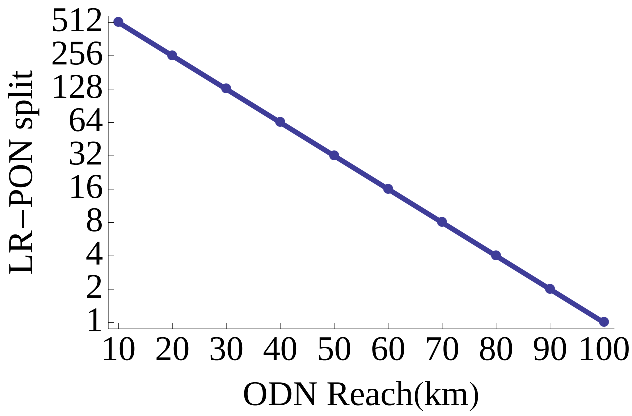


Figure 33: Relationship between the number of customers and the maximum fiber length

- C is a matrix of non-negative costs. Notice that C would be same as D if we are not given actual road distances.
- m is the number of fibers leaving exchange-sites towards customers
- T_i is the tree distribution network for fiber i

The Cable Tree Model consists in finding a set of m bounded trees of minimum total cost, starting at facility, such that every vertex in $V \backslash \{v_0\}$ is included exactly in one tree. A feasible solution is composed of a partition $R_1, \ldots, R_m$ of $V \backslash \{v_0\}$ and a tree $T_i = (R_i \cup \{v_0\}, L_i)$ with root v_0 and $L_i \subseteq E_{\downarrow R_i \cup \{v_0\}}$. A tree T_i is feasible if the cable visits a client exactly once, the distance from the facility to the client does not exceed a given bound and the size of the tree in terms of number clients does not exceed a given capacity limit.

Variables

- Let x_{ij} be a Boolean variable for each $\langle i, j \rangle \in V^2$ that denotes whether a cable between node i and node j exists in the tree.
- Let y_{ij} be a non-negative integer variable that denotes the number of customers in the sub-tree emanating from node j.
- Let z_i be a non-negative integer variable that denotes the length of the path from the facility to the client v_i.

Constraints

Each node (except root-node) must have one incoming cable:

$$\sum_{v_i \in V} x_{ij} = 1, \quad \forall_{v_j \in V \backslash \{v_0\}}$$

The root-node is connected to at least one another node:

$$\sum_{v_j \in V \backslash \{v_0\}} x_{0j} \geq 1$$

The total number of cables in the tree is equal to $|V \backslash \{v_0\}|$:

$$\sum_{v_i \in V} \sum_{v_j \in V \backslash \{v_0\}} x_{ij} = |V \backslash \{v_0\}|$$

If there is a cable from $v_i \in V$ to $v_j \in V$ then the length of the path from root-node to v_j is equal to the cost of the path from root-node to v_i plus the cost between nodes v_l and v_k:

$$x_{lk} = 1 \Rightarrow z_k = z_l + \kappa_{lk} \quad \forall_{\{v_l, v_k\} \subseteq V}$$

The number of customers relying on any cable node is more than the number of nodes in the sub-tree emanating from that node.

$$y_{ij} = \sum_{v_k \in V} y_{jk} + 1, \quad \forall_{\{v_i,v_j\} \subseteq V}$$

If there is no cable between nodes i and j then of course no customers are relying on this cable in the tree

$$\forall_{\{v_i,v_j\} \in V} : x_{ij} = 0 \implies y_{ij} = 0$$

The number of customers in any tree is dependent on the maximum distance between a customer and the facility. For example, if the maximum path length is greater than 10 then the cable tree network cannot contain more than 512 customers. Below we show constraints for 10, 20, 30 and 40 kms and the similar constraints for 40 to 100 in steps of 10 kms can be enforced.

$$z_i > 10 \implies y_{ij} < 512, \quad \forall_{\{v_i,v_j\} \subseteq V}$$

$$z_i > 20 \implies y_{ij} < 256, \quad \forall_{\{v_i,v_j\} \subseteq V}$$

$$z_i > 30 \implies y_{ij} < 128, \quad \forall_{\{v_i,v_j\} \subseteq V}$$

$$z_i > 40 \implies y_{ij} < 64, \quad \forall_{\{v_i,v_j\} \subseteq V}$$

Objective Function

The main objective is to minimize the total cost.

$$\min \sum_{v_l \in V} \sum_{v_k \in V} \kappa_{lk} \cdot x_{lk}$$

One might also be interested in minimizing the number of cable tree networks:

$$\min \sum_{v_j \in V \setminus \{v_0\}} x_{0j}$$

4.1.2 Open ring model

In addition to the previous cable tree model, in the DISCUS project we study an alternative cable model to connect PONs to their primary and secondary metro-core nodes. Unlike the cable tree model where each PON should be double covered by explicit fiber connections from the PON to the primary and secondary metro-core nodes (see Figure 30), in the chain model (see Figure 34) a set of PONs are chained through a fiber and the two metro-core nodes are connected to the two ends of the chain. However, in order to build chains a set of constraints, including distance between PONs and customers in the PON, must be fulfilled. And therefore, these chain models are mainly used in rural areas where a few hundred of customers are covered by a single local-exchange site. Similarly to the cable tree model, a set of customers are grouped into a PON forming a tree network.

Figure 35 depicts an example of the cable tree and open ring models. Let M_i denote a set of metro-core nodes, e_i denote a set of local-exchange sites, and P_i denote a set of PONs. In the right part of the (Figure 35(a)) we observe that each local-exchange is directly connected to the two metro-core nodes, while the open ring model (Figure 35(b)) allows a reduction of direct connections from metro-core nodes to local-exchange sites at a cost of adding cable links between pairs of local-exchange sites where PONs are being chained. In particular in the example, we observe two chains derived from c_1={P_1, P_2, P_3}, and c_2={P_4, P_5}, customers in

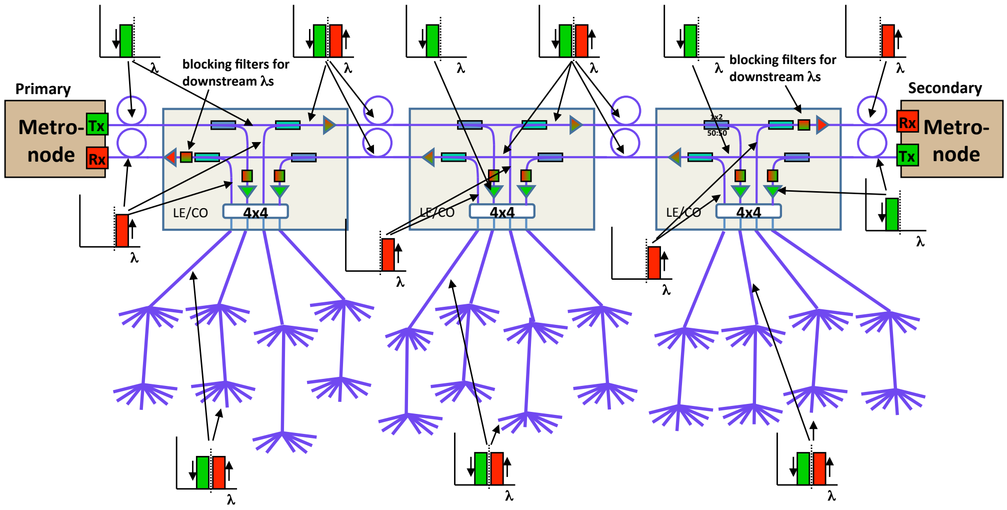
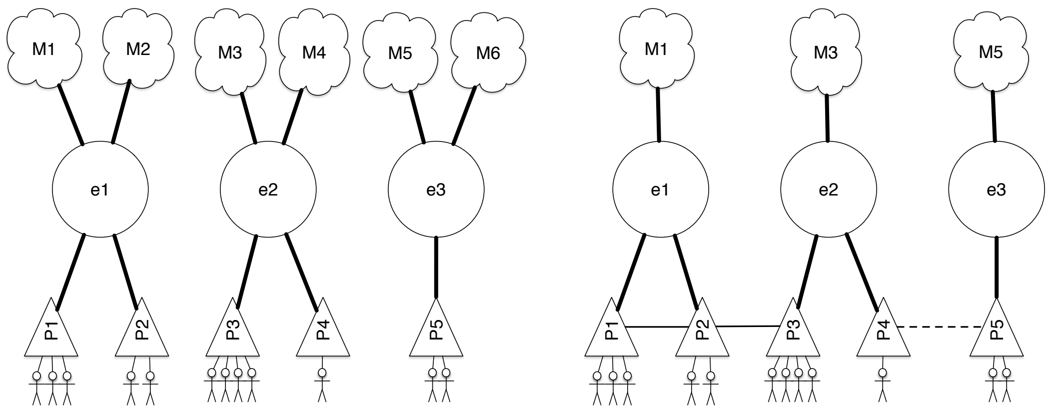


Figure 34: Open ring model

c_1 are covered by metro-core nodes M_1 and M_3, and customers in c_2 are covered by metro-core nodes M_3 and M_4.

Taking this into account, there is a trade-off between reducing cable link from local-exchange sites to metro-core nodes and adding cable links between pairs of local-exchange sites, the following optimization questions should be addressed:



(a) Cable tree model (b) Open ring model

Figure 35: Access network distribution examples

- What PONs should be chained?
- What PON should be starting/ending a chain?
- What metro-core nodes should be used as primary and secondary for chained PONs?

Furthermore, the following constraints must be fulfilled:

- There must be at most 5 PONs in a chain;
- the maximum length between any two PONs in a chain should be at most 50 km;
- the total distance from the metro-core node to any customer should be at most 125 km;
- chained PON should have at most 128 customers;

Split size	Distance (km)
128	10
64	20
32	30
16	40
8	50

Table 6: Maximum fiber length from the local-exchange site to any customer in the *open ring* model

- the maximum distance from the local-exchange site to the furthest customer is given in Table 6;

In the following we assume that the open ring model will be applied in rural areas with up to few hundred customers associated to the local-exchange site and with customers within a ratio of up to 10 km. The cable tree model will be applied in urban areas with up to tens of thousands customers associated to the local-exchange site, and also for rural areas where the customers are grouped within a ratio greater than 10 km. Notice that the 10 km limitation is given by the constraints of the model.

Objective

Due to the ring topology the *open ring* model might reduce direct connections from local-exchange sites to metro-core nodes, and therefore some local-exchange sites might require only one (or none) direct connections to the metro-core node. However, due to the constraints for building a chain, PONs are limited in size (number of customers covered). The objective function in the *open ring* model would be reducing the total cable length derived from connections between local-exchanges sites to metro-core nodes and pairs of local-exchanges.

Constraint Optimization Model

Constants

- $G = (V, E)$ is given a network;
- Let $p_i \in \mathcal{P}$ be the set of PONs;
- Let $l_i \in L$ be the set of local-exchange sites;
- Let $m_i \in M$ be the set of metro-core nodes;
- $V = \mathcal{P} \cup M\}$ is a set of vertices in the graph (either a PON or a metro-core node);
- $E = \{(v_i, v_j)|v_i, v_j \in V; i \neq j \wedge (v_i \neq M \vee v_j \neq M)\}$ is a set of edges in the graph;
- Let l_i be a constant denoting the local-exchange site of p_i;
- Let $lp_i \in lp$ be the set of PONs of a given local exchange l_i;
- Let M_i^1 (resp. M_i^2) be a constant denoting the primary (resp. secondary) metro-core node for l_i;
- Let κ_{lk} be a constant denoting the cost of connecting nodes l_l and l_k;
- Let κc_i be a constant denoting the cost of connecting local-exchange site l_i and the furthest customer in p_i;
- Let κm_i^1 (reap. κm_i^2) be a constant denoting the cost of connecting local-exchange site l_i to the primary (reps. secondary) metro-core node.

Variables

- Let x_{ij} be a Boolean variable that denotes whether a cable link between nodes n_i and n_j exits, where a node could be either a PON or a metro-core node;

- Let y_{ij} be a Boolean variable that denotes whether a cable link between local-exchange sites l_i and l_j exits;
- Let m_i^1 (reps. m_i^2) be a non-negative integer variable that denotes the primary (reps. secondary) metro-core node for p_i;
- Let mn_i^1 (reps. mn_i^2) be a Boolean variable indicating whether local-exchange site i is using M_i^1 as primary metro-core node (reps. M_i^2 as secondary metro-core node);
- Let f_i be a real number denoting the distance between L_i and the primary metro-core node starting the chain;
- Let b_i be a real number denoting the distance between L_i and the secondary metro-core node ending the chain;
- Let $I_j \in \{1..5\}$ be a non-negative variable denoting the position in the chain of p_i;

Constraints

- If there is a cable link from p_i to p_j then the two elements must have consecutive positions in the chain.

$$x_{ij} = 1 \quad \Rightarrow \quad I_j = I_i + 1, \quad \forall_{p_i, p_j \subseteq \mathcal{P}}$$

- Setting metro-core nodes that are starting and ending a chain

$$\begin{array}{lll} x_{ij} = 1 & \Rightarrow & m_i = M_j^1, \quad \forall_{m_i \in M, p_j \in \mathcal{P}} \\ x_{ij} = 1 & \Rightarrow & m_j = M_i^2, \quad \forall_{m_j \in M, p_i \in \mathcal{P}} \end{array}$$

- Two consecutive PONs must have the same primary and secondary metro-core nodes

$$x_{ij} = 1 \quad \Rightarrow \quad M_i^1 = M_j^1 \wedge M_i^2 = M_j^2, \quad \forall_{p_i, p_j \subseteq \mathcal{P}}$$

- The primary and secondary metro-core node for a PON must be different

$$M_i^1 \neq M_i^2, \quad \forall_{p_i, \subseteq \mathcal{P}}$$

- Every PON node in the graph must have a degree equal to two

$$\textstyle\sum_{p_i \in \mathcal{P} \setminus \{p_j\}} x_{ij} + x_{ji} = 2, \quad \forall_{p_j, \subseteq \mathcal{P}}$$

- A local-exchange site requires a connection to a metro-core node if at least one PON requires this cable link

$$x_{ij} = 1 \wedge (n_i = m_q^z \vee n_j \in m_q^z) \quad \Rightarrow mn_q^z = 1, \quad \forall_{p_i, p_j \subseteq \mathcal{P}, lp_q \in lp, z \in \{1,2\}}$$

links from the PON to the primary or secondary associated to the local exchange.

- Distance up to p_i from the primary and secondary metro-core nodes

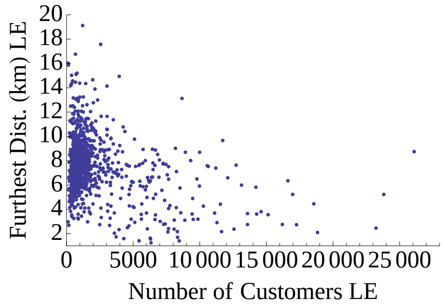
$$\begin{array}{ll} x_{lk} = 1 & \Rightarrow f_k \geq f_l + \kappa_{lk}, \quad \forall_{v_l, p_k \subseteq \mathcal{V}} \\ x_{lk} = 1 & \Rightarrow b_l \geq b_k + \kappa_{lk}, \quad \forall_{p_l, v_k \subseteq \mathcal{V}} \end{array}$$

- Maximum allowed distance from the primary (reps. secondary) metro-core node to the furthest customer

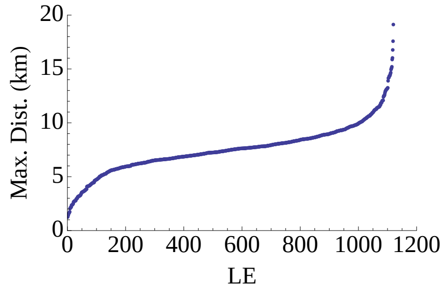
$$\begin{array}{lll} f_i + \kappa c_i & \leq & 125, \quad \forall_{p_i \in \mathcal{P}} \\ b_i + \kappa c_i & \leq & 125, \quad \forall_{p_i \in \mathcal{P}} \end{array}$$

- There is a cable link between two local-exchanges l_i and l_j if there is at least one fiber between any two PONs in l_i and l_j

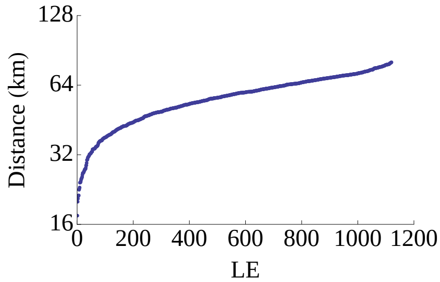
$$y_{ij} \geq x_{ij} \quad , \forall_{p_i, p_j \subseteq \mathcal{P}}$$



(a) Max. distance and users per LE



(b) Furthest customer per LE



(c) Distance to metro-core node $|M|$=18

Figure 36: Move operators

Objective

The objective is to minimize the total cost between any pair of local-exchange sites and from local-exchange sites to the primary and secondary metro-core nodes:

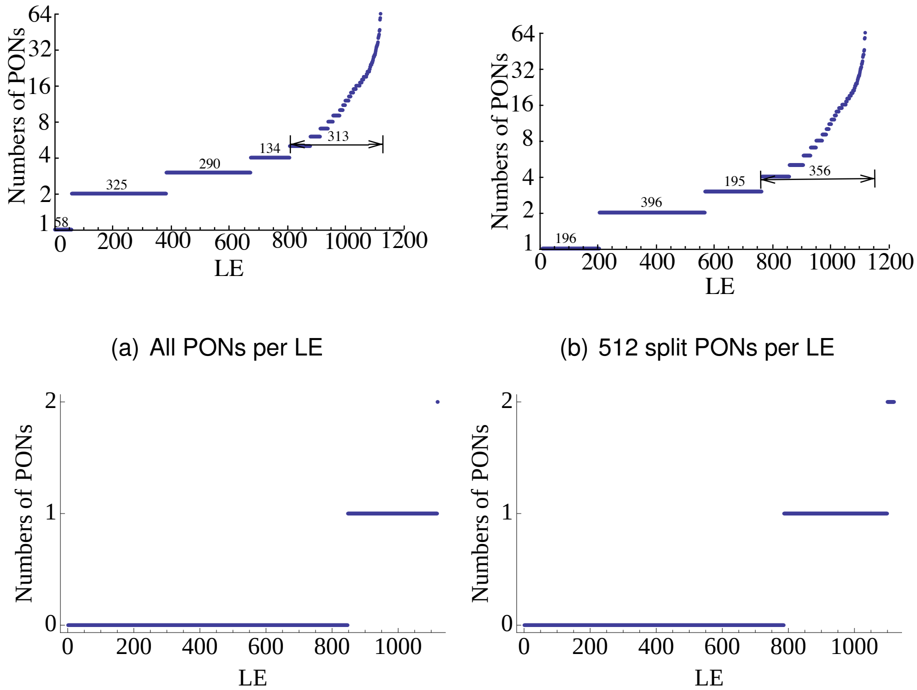
$$\sum_{l_l, l_k \subseteq L} y_{lk} \cdot \kappa_{lk} + \sum_{l_l \in L} \sum_{p \in 1,2} mn_l^p \cdot dm_l^p$$

Computational results

Figure 31 shows statistics of the input data from Ireland, Figure 36(a) shows for each point a relation between total number of users (x-axis) and the furthest customer in the local exchange (y-axis). Figure 36(b) details the distribution from the local-exchange site to the furthest customer for each local exchange, and Figure 36(c) depicts the distance from local-exchange sites to metro-core nodes.

Now we switch our attention to Table 7 where we detail for each max. split size (or PON type) the number of PONs, number of customers, minimum and maximum distance from the local-exchange site to its associated customers, and the utilization factor for each PON type. As expected the 512 PON type is covering the majority of the customers, about 97.2% of the customers in Ireland, mainly because we try to build PONs with as many consumers as possible. While PON types 256 and 128 are covering about 1.96% and 0.74% respectively of the entire population. In addition to the overall PON usage, Figure 37 depicts the number of PONs per local exchange. It is also worth pointing out that PON type 512 is close to the maximum allowed utilization 2.8% (i.e., 80%).

Figure 37 shows the distribution of PON types per local-exchange site. In the first figure we observe that only 58 local-exchange sites require only one PON, while 313 local-exchange require at least 5 PONs. In Figure 37(b) depicts the number of 512 split size PONs, and an important number of local-exchange sites are not using PONs with split size 256 or 128, that is because urban areas usually use 512 PON type.



(a) All PONs per LE
(b) 512 split PONs per LE
(c) 256 split PONs per LE
(d) 128 split PONs per LE

Figure 37: Number of PONs per local-exchange site

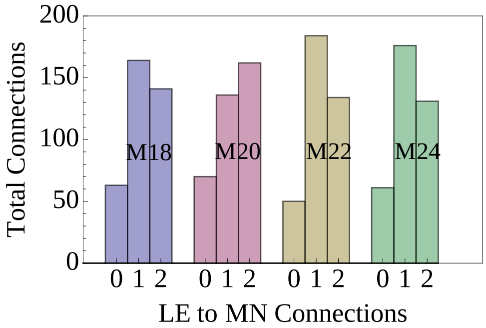
We now present experimental results for the open-ring model in rural areas, as pointed out above in order to apply the model a set of constraints regarding the distance between local-exchange sites and size of the PONs must be fulfilled. Therefore, we limit our attention to local-exchanges with at most 1024 customers and where the distance to the furthest customer is at most 10 km, furthermore only PONs of at most 128 split size are allowed.

Table 8 depicts preliminary results of the use of the open ring model to chain PONs in rural for a total of 368 local-exchanges with 2287 PONs. In this experiment we consider $|M| \in \{18, 20, 22, 24\}$ metro-core nodes. The cable tree column depicts the total cable distance in km of connecting local-exchange sites directly to their primary and secondary metro-core nodes, the open ring column depicts the total cable distance in km of each local-exchange site to reach its primary and secondary metro-core nodes, and the last column shows the percentage reduction of applying the model (i.e., cable tree / open ring). As it can be observed the open ring provides an interesting reduction between 65-67% against direct connections to metro-core nodes.

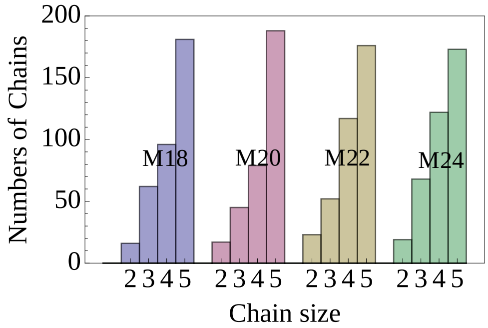
In Figure 38 we present statistics of the results of the open ring model. Figure 38(a) shows the total number of cable links from local-exchange sizes to metro-core nodes. As it can be observed after using the open ring model 141, 162, 134, and 131 local-exchanges for respectively $|M| \in$ {18, 20, 22, 24} will still require two direct connections to metro-core nodes; while

Split size	Total PONs	Customers	Min (km)	Max (km)	Utilization
512	5388	2129896	0.006	9.999	77.2%
256	275	43074	0.016	17.576	61.1%
128	355	16150	0.010	19.121	35.5%

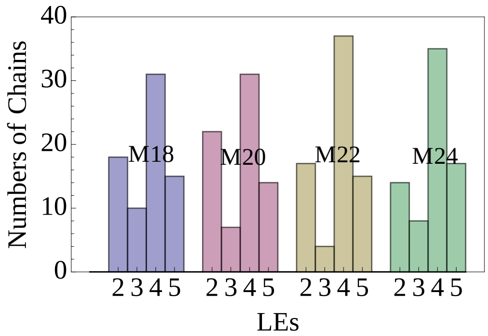
Table 7: PON usage statistics



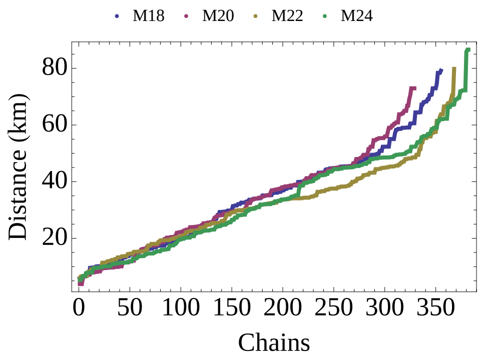
(a) LE to MC 5



(b) Chained PONs



(c) Chained LEs



(d) Chain distance

Figure 38: Open ring statistics

164, 136, 184, and 176 local-exchanges for $|M| \in \{18, 20, 22, 24\}$ will require one direct connection to a metro-core node; and 63, 70, 50, and 61 local-exchanges for $|M \in$ {18, 20, 22, 24} will require no direct connections to a metro-core node, that is, because these all PONs in this set of local-exchange sites are in the middle of a chain and will be using metro-core nodes connected to other local-exchange sites.

Figure 38(b) shows the distribution of the size of the chains for each experimented metro-node configuration. We observe similar distributions for each experiment, and the size of the chains ranges from 16 chains of size 2 for $|M|$=18 to 188 chains of size 5 for $|M|$=20. Figure 38(c) shows the total number of local-exchanges being linked by means of chains involving their PONs, interestingly here we observe slightly different distributions for different metro-core node settings, for instance $|M|$=18 has 10 chains that involves three local-exchanges, while M=22 has only 4 chains that involves three local-exchanges, in total there are 265, 259, 269, and 277 local-exchange sites involved in at least one chain respectively for $|M| \in \{18, 20, 22, 24\}$ out of 368 local-exchanges. Finally, we conclude with Figure 38(d) where we depict the distribution of the distances of the chains, we observe that the distance ranges from 3.8 km for $|M|$=20 to 86.5 km for $|M|$=24, it is worth noticing that all chains must fulfill a length constraint which indicates

$\|M\|$	Cable tree	Open ring	Reduction
18	34677	22999	66%
20	34099	22946	67%
22	32397	21699	66%
24	31178	20326	65%

Table 8: Total cable distance (km)

that the distance from the metro-core node to the furthest customer must be at most 125 km, so that, chains of larger distance will be unlikely as the distance from the local-exchange site to the metro-core node should be small.

4.2 Assigning Local Exchanges to Metro-Core-nodes

In this section, we present optimization studies that focus on the backhaul/metro design of the DISCUS architecture, see Figure 1. In particular we face the following questions:

- How many MC nodes are needed to connect all LE sites in a country respecting distance and capacity constraints and where should they be placed?
- How should the connection fibers between the LE sites and the MC nodes be routed such that cable cost is minimized and resiliency requirements are fulfilled?
- What is the impact of *dual-homing* in terms of the minimal number of feasible MC nodes and in terms of cost.

In deliverable D4.5 [7, Chapter 4.2, 4.6], we already gave partial answers to some of these questions. However, the corresponding computations were restricted to the LE assignment and MC placement, ignoring MC node capacities, fiber routes with cable sharing, and cable cost. Moreover, only the Italian network (see 2.1.1) was used as a reference topology.

In this section, we wish to extend the mentioned studies by answering all of the above questions. In particular, we compute fiber routes that share trails and hence exploit synergistic effects on cable and duct usage as well as on trenching costs. For this, we use models that go beyond the preliminary approaches presented in deliverable D4.5. Moreover, we will present numbers for all of the reference topologies, the Italian, Spanish, and UK network.

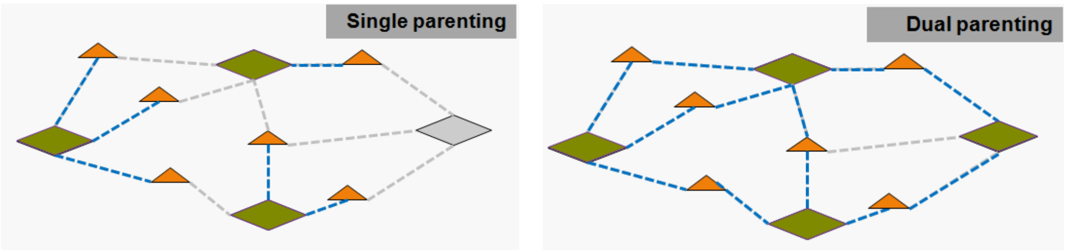
In the following, we will review some models already introduced in deliverable D4.5 [7, Chapter 4.2] emphasizing on improvements and extensions.

Recall that when optimizing the backhaul/metro design we assume the ODN architecture to be part of the input, that is, the connection of (residential, business) customers to LE sites is not subject to optimization. To this end we assume that the LE locations and the assignment of customers to these LE locations are fixed. For the LE sites we are given the number of connected customers/households and the splitting hierarchy towards the LE, which results in a certain PON and fiber demand at the LE site towards the MC node.

As already pointed out in D4.5, the optimization of the metro design can be decomposed as follows:

- *LE assignment optimization:* The task is to open MC nodes at potential locations in the reference network and to assign LE sites to opened MC nodes, see Figure 39. We have to respect distance constraints (maximal signal reach), resilience restrictions (single- or dual-homing) and capacity constraints at the MC node (e.g., maximum number of customers).
- *Fiber route optimization:* For every LE to MC assignment we have to decide about individual fiber routes, installing cables and using duct space in such a way that maximizes resource sharing (or minimizes the cost of fibers, cables, and ducts), see Figure 40.

We will first review models for the LE assignment before we provide alternative models tackling cable optimization and before we integrate the optimization of both problems.



(a) Single homing (b) Dual homing

Figure 39: *LE assignment:* Single vs. dual parenting (The triangles correspond to LE sites, the diamonds to potential MC nodes. The colored MC are actually opened, the blue dotted lines correspond to the LE-MC assignments)

4.2.1 MC-node selection and LE assignment

We use the following central facility location model as a basis to solve the assignment problems in the backhaul section of the DISCUS architecture (For details see [7, Chapter 4.2]):

$$\min \left(\sum_{l \in L} \sum_{k \in M^{(2)}} \kappa_{lk} x_{lk} + \sum_{m \in M} \kappa_m x_m \right)$$

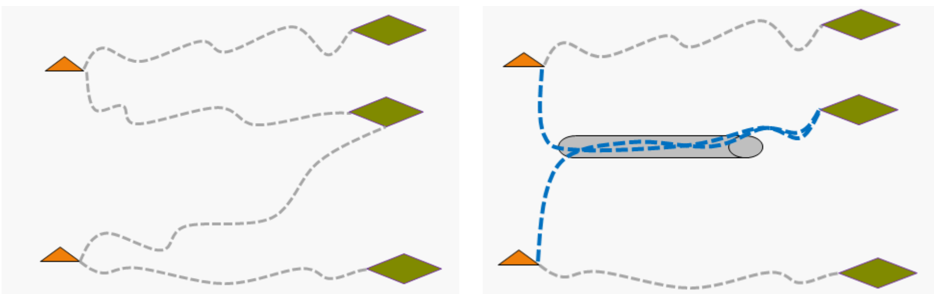
$$(DFL) \qquad \sum_{k \in M^l} x_{lk} = 1 \qquad \forall l \in L \tag{7}$$

$$x_{lk} \leq x_m \qquad \forall l \in L, k = \{m, n\} \in M^l \tag{8}$$

$$x_{lk} \leq x_n \qquad \forall l \in L, k = \{m, n\} \in M^l \tag{9}$$

This model opens individual MC nodes from the set of potential MC locations M and it assigns all local exchanges from the set L to a pair of opened MC nodes. The set of all pairs of MC locations is denoted $M^{(2)}$ while the set of MC pairs that is *feasible* for a single LE $l \in L$ is given by M^l. Clearly, $M^l \subseteq M^{(2)}$. We use the decision variables x_m to indicate whether or not an MC site m is opened and x_{lk} to decide whether or not an LE l is assigned to a pair k of MC sites with $k \in M^l$.

We can use the MC setup cost parameters κ_m and the assignment cost parameter κ_{lk} to model the cost objective. In the extreme case of just minimizing the number of used MC nodes we set $\kappa_{lk} := 0$ and $\kappa_m := 1$ for all $l \in L$, $k \in M^l$, and $m \in M$.



(a) No sharing (b) Path sharing

Figure 40: Optimizing fiber routes: (The triangles correspond to LE, the diamonds to MC)

Besides the cost function, the model (DFL), its optimal solutions, and the solvability of the problem strongly depend on a meaningful characterization of the set $M^{(2)}$, that is, the notion of feasibility for pairs k of MC nodes w.r.t a given LE site $l \in L$. In general, we only consider existing LE sites in the reference topology with a reasonable connectivity as being admissible to open MC nodes. Moreover, a pair of admissible MC nodes is feasible for a particular LE site if

- there exists (maximally) disjoint fiber routes between the LE and the two MC nodes and
- these fiber routes do not violate the distance constraint, that is, all customer behind the LE this pair is reachable within the maximum distance of 125km from both MC nodes.

This means that already setting up the sets M^l for LE sites $l \in L$ involves solving a series of optimization problems. For details on MC node feasibility, assignment policies, and how to compute disjoint fiber routes see D4.5 and [9].

Model (DFL) can be extended easily to model partial dual homing or to cover additional MC node capacity constraints such as a limitation on the number of customers connected to a single opened MC node, limitation on the number of LE sites or PONs connected to a single opened MC node, to model partial dual homing, (see Deliverable D4.5 for details)

Preliminary results on the MC-consolidation

region	$\|L\|$	60km	70km	80km	90km	100km	110km	120km	125km
abruzzo	337	–	17	14	10	9	8	6	6
basilicata	185	18	13	11	9	7	5	4	4
calabria	534	37	28	21	17	14	11	10	9
campania	660	26	19	13	11	8	7	6	6
emilia-romagna	928	35	24	20	15	11	11	8	8
friuli-venezia-giulia	258	15	14	8	7	6	5	4	4
lazio	735	26	19	15	11	9	7	6	6
liguria	318	13	10	9	8	6	6	5	4
lombardia	1192	–	28	22	17	14	11	9	12
marche	336	15	11	9	8	5	5	4	4
molise	126	9	7	6	6	4	3	3	2
piemonte	1168	–	28	22	16	13	10	9	9
puglia	358	–	23	18	14	10	9	8	8
sardegna	492	40	31	22	18	14	10	8	7
sicilia	586	46	36	23	17	13	11	9	9
toscana	949	38	26	20	15	13	10	9	8
trentino-altoadige	312	–	–	–	–	10	8	8	6
umbria	229	14	11	8	8	5	4	4	3
valle-d-aosta	72	–	4	4	3	3	2	2	2
veneto	845	27	20	16	11	11	10	7	6
Total (Region by region)	10620	359	369	281	221	185	153	129	123

Table 9: Computational results on the ***Italian*** reference networks

We present preliminary computational results on the MC-consolidation models using the three reference networks for Italy, Spain and Great-Britain. The optimization runs aim at minimizing the number of used MC nodes. The runs were performed with varying customer-MC-distance from 60 km to 125 km (dual homing with maximally disjoint paths).

See Table 9 for the results on the Italian reference network, Table 10 for the British network, resp. Table 11 for the Spanish reference network. In these tables the column `|L|` shows the number of LE, the columns `60km`...`125km` the number of selected MC-location for the given

region	$\|L\|$	60km	70km	80km	90km	100km	110km	120km	125km
england	3880	134	122	86	68	57	50	40	40
northern-ireland	186	17	13	11	9	6	6	5	5
scotland	948	–	–	56	44	38	32	26	24
wales	436	31	22	17	13	9	8	7	7
Total (Region by region)	5450	182	157	170	134	110	96	78	76

Table 10: Computational results on the ***British*** reference networks

region	$\|L\|$	60km	70km	80km	90km	100km	110km	120km	125km
Asturias	143	15	11	9	8	8	6	6	6
Valenciana	717	31	25	19	14	12	12	11	10
Navarra	171	13	11	8	7	7	6	4	4
Murcia	147	–	–	6	5	4	3	3	3
Extremadura	340	–	–	23	21	17	14	12	12
Pais Vasco	320	–	11	10	7	7	6	6	6
Baleares	135	9	7	7	6	6	6	6	6
Castilla-La Mancha	763	–	61	51	42	35	27	24	23
Cantabria	110	–	7	6	5	5	4	4	3
Galicia	423	–	29	26	22	20	16	15	14
Madrid	716	8	6	6	5	4	3	2	2
Cataluna	1147	–	–	–	22	18	15	13	12
Castilla y Leon	1375	108	79	63	48	42	37	30	27
La Rioja	94	8	7	5	5	4	4	3	3
Aragon	537	–	–	–	31	23	19	16	16
Andalucia	1128	–	79	63	51	46	41	33	32
Total (Region by region)	8266	249	374	332	299	258	219	188	179

Table 11: Computational results on the ***Spanish*** reference networks

(maximum) customer-to-MC-distance. It should be noted that some of the instances cannot be solved as w.r.t. to the given maximum distance no two disjoint paths to two MC-locations exist. These instances are marked with –.

Comparison of different resilience strategies

In the following we compare the impact of different resilience strategies on the number of required MC nodes. To this end we compare

- SINGLEHOMING: An LE is assigned to exactly one MC (model (FL) with n=1, see [7, Chapter 4.2])
- DUALHOMING: An LE is assigned to exactly two MC (model (FL) with n=2, see [7, Chapter 4.2])
- DUALHOMINGSTRICT: An LE is assigned to two MC using maximally disjoint paths (model (DFL))

In Table 12 (resp. Table 13) we show results for the British (resp. Spanish) networks, which gives the number of MC-locations using one of the strategies and a maximum customer-to-MC distance of 100 km or 125 km. It should be observed that the major increase in number of MC is when we switch from SINGLEHOMINGto DUALHOMING. There is only a slight increase when we require the disjointedness of the two paths.

region	SINGLEHOMING		DUALHOMING		DUALHOMINGSTRICT	
	100km	125km	100km	125km	100km	125km
england	23	17	49	33	57	40
northern-ireland	3	1	6	3	6	5
scotland	16	10	33	21	38	24
wales	4	3	8	5	9	7

Table 12: Comparison of resilience strategies on the ***British*** reference networks

region	SINGLEHOMING		DUALHOMING		DUALHOMINGSTRICT	
	100km	125km	100km	125km	100km	125km
Asturias	4	2	8	4	8	6
Valenciana	6	4	11	8	12	10
Navarra	2	1	4	2	7	4
Murcia	2	1	4	2	4	3
Extremadura	7	5	15	11	17	12
Pais Vasco	3	3	6	6	7	6
Baleares	3	3	6	6	6	6
Castilla-La Mancha	15	10	30	19	35	23
Cantabria	2	1	4	2	5	3
Galicia	7	4	15	8	20	14
Madrid	2	1	3	2	4	2
Cataluna	9	5	17	10	18	12
Castilla y Leon	17	11	36	23	42	27
La Rioja	2	1	4	2	4	3
Aragon	9	7	20	14	23	16
Andalucia	18	12	36	25	46	32
Total (Region by region)	108	71	219	144	258	179

Table 13: Comparison of resilience strategies on the ***Spanish*** reference networks

4.2.2 Fiber route optimization

In [7] we introduced optimization models that are based on the idea to generate a set possible feasible paths for each LE-MC-connection and choose the connection paths that share the most edges.

Path groups model

$$\min \left(\sum_{e \in E} \kappa_e \cdot u_e + \sum_{p \in \mathcal{P}} \kappa_p \cdot z_p \right) \tag{10}$$

$$\sum_{G \in \mathcal{PG}(l)} y_G = 1 \qquad \forall l \in L \tag{11}$$

$$z_p \geq y_G \qquad \forall G \in \mathcal{PG}, \forall p \in G \tag{12}$$

$$u_e \geq z_p \qquad \forall p \in \mathcal{P}, \forall e \in E_p \tag{13}$$

Conflicting path model

In addition to the model described in [7, Chapter 4.3] we introduce an additional model that seems to be computational promising as well.

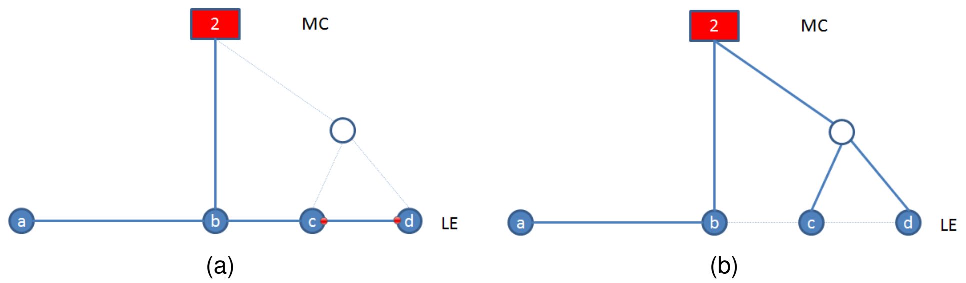


Figure 41: Steiner-Tree: (a) Iteration 1 (b) Iteration 2.

$$\min \left(\sum_{e \in E} \kappa_e \cdot u_e + \sum_{p \in \mathcal{P}} \kappa_p \cdot z_p \right) \tag{14}$$

$$\sum_{p \in \mathcal{P}(l,m)} z_p = 1 \qquad \forall l \in L, \forall m \in M(l) \tag{15}$$

$$z_{p_1} + z_{p_2} \leq 1 \qquad \forall (p_1, p_2) \in \mathcal{C} \tag{16}$$

$$u_e \geq z_p \qquad \forall p \in \mathcal{P}, \forall e \in E_p \tag{17}$$

Used notation:

$M(l)$: Set of MC to which the LE l has to be connected.
$\mathcal{P}(l,m)$: Set of paths connecting the LE l to the MC m.
$\mathcal{C}$: Set of all conflicting pair of paths (p_1, p_2), i.e., paths that cannot occur both in a solution.
$z_p \in \{0,1\}$ ($= 1$, if path p is used to connect an LE to an MC)
$u_e \in \{0,1\}$ ($= 1$, if edge e is used in any of the connection paths)

It will be part of further computational tests to evaluate which model will be computational superior.

To work with this model some aspects have to be discussed and resolved, namely

- How to generate the paths.
- How to determine conflicting paths.

Path generation Due to its exponential size it will be impossible to generate all possible paths as input to the fiber-opt-models. Therefore, one has to restrict oneself to promising generation techniques. Two options are:

- **Shortest Paths**: For each relevant (LE,MC) pair calculate the shortest path, add it to the set of paths, increase the weight of the path-edges, recalculate the shortest-path, etc. Repeat this for a limited number of iterations.
- **Steiner-Tree**: For each MC node m calculate a Steiner-Tree T connecting m and all l in $L(m)$. Test all paths p =(l,m) in T. If the path length d is $\leq l_{max}$ add path p. Update the edge-weights, i.e., increase weight for used edges and repeat this iteration n times. (see Figure 41)

From our point of view it will be necessary to apply *adaptive generation procedures*, i.e. start with a good subset of paths, analyze the optimization solution and generate additional paths that eliminate current solution deficits.

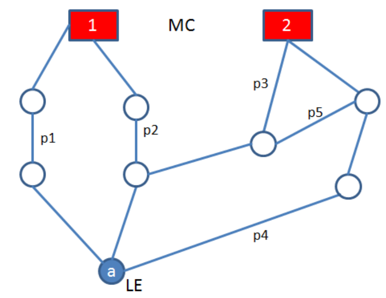


Figure 42: Generated (LE,MC)-paths $p_1, \ldots, p_5$: p_2 and p_3 are in conflict.

Conflicting paths. One central element in the *Conflicting Path Model* are the path conflicts $\mathcal{C}$, i.e. paths (p_1, p_2), that cannot occur both in a solution. (See Figure 42 for an example).

First, we have to guarantee that we select paths that represent the best possible choice with respect to the chosen resilience concept. To this end we introduce the following **disjointedness measure**. Given two path-groups $P_{ij} := (p_i, p_j)$ and $P_{jk} := (p_j, p_k)$ we say that P_{ij} dominates P_{jk} w.r.t. disjointedness, if it has (applied in that order):

1. less common used edge-length (used by both paths)
2. less common used nodes.

Among all paths we determine those two paths that have the best disjointedness measure. All pair of paths that have a worse disjointedness measure are said to be conflicting. Given two paths that are completely disjoint (except the LE) we can therefore guarantee that we accept only connections in the solution that are completely disjoint as well.

For further details see upcoming deliverables in WP4.

Preliminary computational results for Fiber-Opt-Optimization

In Table 4.2.2 we report on preliminary computational results that we performed on the British reference network. To this end we supplied for each regional network solutions (calculated with the MC-consolidation algorithms) that have a different number of MC-locations.

We used the locations calculated for a maximum customer-MC-distance of 80, 90 and 100 km. Each of these MC-locations-set was used to calculated fiber trails of 100 km.

It can be seen that for a higher number of MC-nodes, the more input-paths can be generated (column $|\mathcal{P}|$). This is due to the fact that for LE that have a distance of almost 100 km not many alternative paths can be found. As a consequence, in most cases the impact of common used trails is higher the more paths were generated (see column *Diff/Trail*, which gives the trail length in km that is reduced due to optimization). It is expected that while the trail-length is decreasing that the fiber-length will increase.

Instance			Optimization		Input (km)		Output (km)		Diff (km)	
Region	Dist	$\|M\|$	$\|\mathcal{P}\|$	$\|\mathcal{C}\|$	Trail	Fiber	Trail	Fiber	Trail	Fiber
england	80-km	109	97857	264828	61939	385337	57169	400691	4770	15354
england	90-km	93	91724	250006	58628	369796	54677	383351	3951	13555
england	100-km	57	66799	169267	45849	331806	43025	340651	2824	8845
northern-ireland	80-km	11	2814	6462	3109	11632	2740	12684	369	1052
northern-ireland	90-km	9	2989	7401	3146	13675	2843	14438	303	763
northern-ireland	100-km	6	3006	7644	3204	16400	2933	16913	271	513
scotland	80-km	56	12892	26272	13485	63897	12363	67000	1122	3103
scotland	90-km	44	12514	24038	13541	74975	12731	77425	810	2450
scotland	100-km	38	12288	23815	13709	82447	12870	84979	839	2532
wales	80-km	17	6930	16535	5680	28460	5069	30023	611	1563
wales	90-km	13	6463	14236	5724	33038	5182	34408	542	1370
wales	100-km	9	5982	11345	5710	40600	5386	41293	324	693

Dist : Maximum customer-MC distance (on input, output dist will always be 100 km)
$|M|$: Number of MC-locations
$|\mathcal{P}|$: Number of generated paths
$|\mathcal{C}|$: Number of generated path-conflicts
Input.Trail : Length of all trails (in km) used in the provided input solution
Input.Fiber : Length of all fibers (in km) used in the provided input solution
Output.Trail : Length of all trails (in km) used in the optimized solution
Output.Fiber : Length of all fibers (in km) used in the optimized solution
Diff.Trail : Reduction in trail-length between input and optimization (*Input.Trail-Output.Trail*)
Diff.Fiber : Increase in fiber-length between input and optimization (*Output.Fiber-Input.Fiber*)

Table 14: Fiber-Optimization (conflicting-path-model) on the British (regional) reference networks

5 Summary and Preview

In this deliverable we showed how end-to-end solutions can be computed in principle and how they should be evaluated resulting in an end-to-end optimization process. We defined the necessary input reference data (networks, costs, demands) as well as the all optimization problems in detail.

In Section 3 we already highlighted how we will tackle these challenging problems in the remainder of the project. Clearly, the main target is to verify that the DISCUS architecture outperforms other architectures in terms of cost, scalability, and flexibility, and to provide answers to questions such as: (i) What is the optimal number of MC nodes in a country like Italy, (ii) Should we build single optical islands, and if not possible (iii) how do we interconnect these islands optimally.

Our plan for the remainder of the project is to refine the tentative end-to-end cost model described in Section 2.4 such that it provides data for different scenarios of next generation broadband networks and also includes values for technologies such as GPON.

As a main end-to-end optimization study we would like vary the number MC nodes and compare the DISCUS architecture (single-tier, flat optical core) with an architecture that allows a second aggregation level, (a two-tier architecture).

We will start from different access technologies (i.e., considering different maximum reach) and calculate the number of network nodes required to cover a given country. We will then build a core network to interconnect such nodes. Finally based on the developed cost model we will compare the results cost-wise.

What we want to look at is how does the cost of access+core changes when we vary the reach of the access (20 km, 60 km, 100 km, 125 km). We expect that with a shorter reach there will be a large number of nodes (say around 2000). In that case the DISCUS flat core model will potentially not work, and the core network needs to have two tiers, one being a metro for packet aggregation and another a flat core for transport. The DISCUS one-tier model in this case will prove expensive due to the enormous number of wavelength channels. As we go towards longer reaches, the number of nodes will decrease up to a point where we expect that a one-tier model becomes cheaper than the two tier one.

We are interested in seeing if and where this transition point occurs for different countries, based on the reference topologies from Section 2.1: the UK, Italy, Spain, and Ireland. We are also interested to see what reach length will lead to the cheapest end-to-end solution.

These studies will be carried pout in cooperation with the IDEALIST project.

Type	Area	Design	Provides	Cost in €	Cost in ICU	Unit
Link	All	Duct	Duct-Space	58,000	1.1600	Km
Link	Drop and D-side	Cable	12 Fibers	390	0.0078	Km
Link	Drop and D-side	Cable	6 Fibers	313	0.0063	Km
Link	Drop and D-side	Cable	8 Fibers	270	0.0054	Km
Link	Drop and D-side	Cable	4 Fibers	225	0.0045	Km
Link	E-side and Backhaul	Cable	276 Fibers	5,750	0.1150	Km
Link	E-side and Backhaul	Cable	240 Fibers	5,125	0.1025	Km
Link	E-side and Backhaul	Cable	192 Fibers	4,250	0.0850	Km
Link	E-side and Backhaul	Cable	144 Fibers	3,375	0.0675	Km
Link	E-side and Backhaul	Cable	96 Fibers	2,500	0.0500	Km
Link	E-side and Backhaul	Cable	48 Fibers	1,625	0.0325	Km
Link	E-side and Backhaul	Cable	24 Fibers	1,250	0.0250	Km
Link	E-side and Backhaul	Cable	12 Fibers	1,000	0.0200	Km
Node	DP	Splitter	1 D-side, 32 Drop fiber ports	625	0.0125	Piece
Node	DP	Splitter	1 D-side, 16 Drop fiber ports	388	0.0078	Piece
Node	DP	Splitter	1 D-side, 8 Drop fiber ports	250	0.0050	Piece
Node	Cabinet	Splitter	1 E-side, 32 D-side fiber ports	1,203	0.0241	Piece
Node	Cabinet	Splitter	1 E-side, 16 D-side fiber ports	803	0.0161	Piece
Node	Cabinet	Splitter	1 E-side, 8 D-side fiber ports	603	0.0121	Piece
Node	Cabinet	Splitter	1 E-side, 4 D-side fiber ports	503	0.0101	Piece
Node	Cabinet	Splitter	1 E-side, 2 D-side fiber ports	453	0.0091	Piece
Node	Cabinet	Splitter	1 E-side, 1 D-side fiber port	428	0.0086	Piece
Node	LE	Amplifier	4 Backhaul, 4 E-side fiber ports	4,188	0.0838	Piece

Table 15: Tentative LR-PON hardware and cost model. The blown fiber model assumes 1 BFT at cost €100 and different sizes of BFU. The fiber cost per km is between €20.8 (276-fiber cable) and €26.0 (96-fiber cable) for bigger cables in the E-side. Ignoring smaller cables it may hence be feasible to assume an average fiber cost of €25.0 per km. Similarly, we can assume a fiber cost per kilometer of €40.0 for the D-side and drop

A Cost modeling

Tables 15, 16, and 17 present the tentative cost and hardware for for the LR-PON, hardware at the MC node, and optical equipment, respectively. All cost values are given in EUR. For tables 16 and 17 we reused the cost model from IDEALIST (which in turn is based on the model from STRONGEST). The latter models provide cost relative to the SCU (STRONGEST cost unit) and the ICU (IDEALIST cost unit). We use the following tentative conversion factors:

$$1 \text{ ICU} = 12.5 \text{ SCU} = €50{,}000. \quad (18)$$

Type		Provides	Cost in €	Cost in ICU
IP-MPLS router	16	400G Slots	215,000	4.3000
	32	400G Slots	1,143,500	22.8700
	...	...	...	...
	1152	400G Slots	416,451,000	8329.0200
IP-MPLS card	10	40 GE ports	128,000	2.5600
	4	100 GE ports	144,000	2.8800
	1	400 GE port	137,000	2.7400
MPLS-TP switch, 400G slots	16	400G Slots	192,000	3.8400
	32	400G Slots	384,000	7.6800
	...	...	...	...
	112	400G Slots	1,344,000	26.8800
MPLS-TP line-card	10	40G ports	43,360	0.8672
	4	100G ports	54,200	1.0840
	1	400G port	60,680	1.2136
MPLS-TP OLT-card	40	10G ports	t.b.d.	t.b.d.
	10	40G ports	t.b.d.	t.b.d.
Transceiver, Grey, Short Reach	1	40G port	400	0.0080
	1	100G port	1,600	0.0320
	1	400G port	4,000	0.0800

Table 16: Tentative MC-node hardware and cost model based on the values defined in [17] and [4]. We added OLT cards for MPLS-TP switches with a cost value still to be defined.

Type	Provides		Reach	Cost in €	Cost in ICU
Polatis Switch	100	Fiber ports		17,063	0.3413
	...	...		...	...
	400	Fiber ports		34,650	0.6930
Optical Line Amplifier (OLA)			80 km	15,000	0.3000
Digital Gain Equalizer (DGE)			320 km	8,000	0.1600
Fixed Grid only					
Transponder	1	40G port	2500 km	24,000	0.4800
	1	100G port	2000 km	60,000	1.2000
	1	400G port	150 km	68,000	1.3600
WDM Terminal (AWG)	1	Fiber port		24,000	0.4800
Flexible Grid only					
Transponder	1	40G port	2430 km	72,000	1.4400
	1	100G port	2430 km	144,000	2.8800
	1	100G port	1170 km	72,000	1.4400
	1	100G port	500 km	72,000	1.4400
	1	400G port	1170 km	180,000	3.6000
	1	400G port	500 km	180,000	3.6000
WDM Terminal (WSS)	1	Fiber port		360,000	5.76
Optical Cross Connect (WSS)	2	Fiber ports		576,000	11.52
	...	...		...	...
	20	Fiber ports		5,760,000	115.2

Table 17: Tentative hardware and cost model for optical equipment: The single-sided Polatis switch has cost $11,200 + n \cdot 59$ EUR, where n is the number of fiber ports with $100 \leq n \leq 400$. For the AWG WDM terminal we reused the formula in [17, Table X1V] with AWG and Amp values from [4]. The formula for WSS-based optical cross-connects (OXC) with a fiber degree of $2 \leq n \leq 20$ is based on [4, page 52] and the assumption that there is a 20% cost increase for flexgrid. The OXCs are assumed to be fully flexible (colorless, directionless, and contention-less). Fixing the add-drop percentage to 40% and ignoring the (non-significant) cost for amplification we used a simpler formula in Table 17: 1 OXC $= 12 \cdot WSS$ with $WSS = 0.48 ICU$. A WDM-Terminal for flexgrid is just an OXC with fiber degree of 1. We will assume regenerators to have a cost of 1.6 times the cost of the corresponding transponder.

References

[1] DISCUS, Deliverable 2.1, Report on the initial DISCUS End to End Architecture. Technical report, The DISCUS Project (FP7 Grant 318137), 2013.

[2] DISCUS, Deliverable 2.2, First WP2 Progress Report. Technical report, The DISCUS Project (FP7 Grant 318137), 2013.

[3] DISCUS, Deliverable 7.2, Optical Island Description. Technical report, The DISCUS Project (FP7 Grant 318137), 2013.

[4] IDEALIST, Deliverable 1.1, Elastic Optical Network Architecture: reference scenario, cost and planning. Technical report, IST IP IDEALIST: Industry-Driven Elastic and Adaptive Lambda Infrastructure for Service and Transport Network, 2013.

[5] DISCUS, Deliverable 2.3, Updated reference architecture. Technical report, The DISCUS Project (FP7 Grant 318137), 2014.

[6] DISCUS, Deliverable 2.4, Progress report on traffic and service modelling. Technical report, The DISCUS Project (FP7 Grant 318137), 2014.

[7] DISCUS, Deliverable 4.5, Optimization models for long-reach access networks . Technical report, The DISCUS Project (FP7 Grant 318137), 2014.

[8] DISCUS, Deliverable 7.2, Preliminary quantitative results for flat optical network. Technical report, The DISCUS Project (FP7 Grant 318137), 2014, in preparation.

[9] Norbert Ascheuer, Christian Raack, and Roland Wessäly. A nation-wide optimization study on the consolidation of local exchanges using long-reach passive optical networks. In *Proceedings of ICTON 2014 – 16th International Conference on Transparent Optical Networks*, Graz, Austria, 2014.

[10] Nicolas Beldiceanu, Pierre Flener, and Xavier Lorca. Combining Tree Partitioning, Precedence, and Incomparability Constraints. *Constraints*, 13(4):459–489, 2008.

[11] Edsger W Dijkstra. A note on two problems in connexion with graphs. *Numerische Mathematik*, 1(1):269–271, 1959.

[12] Anurag Dwivedi and Richard E. Wagner. Traffic model for USA long-distance optical network. In *Proceedings of the Optical Fiber Communication Conference (OFC)*, volume TuK1, pages 156–158. Optical Society of America, 2000.

[13] IDEALIST. IDEALIST: Industry-Driven Elastic and Adaptive Lambda Infrastructure for Service and Transport Networks , 2014.

[14] Analysys Mason. Final report for ofcom. telecoms infrastructure access. sample survey of duct access. 2009.

[15] Thiago F. Noronha, Andréa C. Santos, and Celso C. Ribeiro. Constraint Programming for the Diameter Constrained Minimum Spanning Tree Problem. *Electronic Notes in Discrete Mathematics*, 30:93–98, 2008.

[16] Jaewon Oh, Iksoo Pyo, and Massoud Pedram. Constructing minimal spanning/Steiner trees with bounded path length. *Integration*, 22(1–2):137–163, 1997.

[17] Franz Rambach, Beate Konrad, Lars Dembeck, Ulrich Gebhard, Matthias Gunkel, Marco Quagliotti, Laura Serra, and Víctor López. A Multilayer Cost Model for Metro/Core Networks. *J. OPT. COMMUN. NETW.*, 5(3):210–225, 2013.

[18] Helmut Simonis. A Hybrid Constraint Model for the Routing and Wavelength Assignment Problem. In Ian P. Gent, editor, *CP*, volume 5732 of *Lecture Notes in Computer Science*, pages 104–118. Springer, 2009.

[19] R. van de Meent, A. Pras, M. R. H. Mandjes, J. L. van den Berg, F. Roijers, L. J. M. Nieuwenhuis, and P. H. A. Venemans. Burstiness predictions based on rough network traffic measurements. In *Proceedings of the 19th World Telecommunications Congress (WTC/ISS 2004), Seoul, Korea*, page 6, Seoul, Korea, September 2004. WTC/ISS.