INTERNATIONAL TELECOMMUNICATION UNION

TELECOMMUNICATION DEVELOPMENT BUREAU

Document NPM/4.1
28 February 2007
Original: English only

Telecom Network Planning for evolving Network Architectures

Reference Manual

Draft version 4.1

February 2007

ITU, Geneva, 2007

ITU-D

Telecom Network Planning for evolving Network Architectures

Reference Manual

Draft version 4.1

Disclaim:

These Guidelines have been prepared with the contribution of many volunteers from different Administrations and Companies coordinated by Riccardo Passerini, ITU- BDT.

The mention of specific Companies or products doesn't imply any endorsement or recommendation by ITU. Opinions expressed in this document are those of the contributors and do not engage ITU

PREFACE

These Guidelines have been prepared with the contribution of many volunteers from different Administrations and companies.

The mention of specific companies or products does not imply any endorsement or recommendation by ITU.

Revision Status :

Chapter	**Title**	**Revision Status**
1	Introduction	15 – January 2007
2	Overview of network planning	15 – January 2007
3	Service definition and forecasting	15 – January 2007
4	Traffic characterization	15 – January 2007
5	Economical modelling and business plans	15 – January 2007
6	Network architectures and technologies	15 – January 2007
7	Network design, dimensioning and optimization	15 – January 2007
8	Data gathering	15 – January 2007
Annex 1	Network planning tools	15 – January 2007
Annex 2	Case Studies	15 – January 2007
Annex 3	References	15 – January 2007

Reference Manual on the Telecom Network Planning for evolving Network Architectures

Table of Contents

Chapter 1 – Introduction

ITU Vision on Network Planning

Background

Telecommunication networks architectures are changing to meet new requirements for a number of services/applications (Broadband, IP, Multimedia, mobile, etc.).
New generation equipment (soft switches, databases, service controllers, new protocols and interfaces, etc.) and new call/mix traffic cases are going to be introduced in the networks.

Different solutions/network architectures can be taken into account for a smooth transition from existing network infrastructures (PSTN/PLMN) towards New Generation Network (NGN) as a result of the convergence process leading to different applications/services sharing network infrastructures.

The planning tool PLANITU, capable of dealing with some new traffic cases, can be considered a tool to introduce people to the Network Planning. However, any real Network Planning case should be dealt with using other powerful and modern tools available on the market.

Planning Strategy

Considering the different solutions/network architectures that exist, each Network Planning case has to be analysed and dealt with by using more than just one planning tool.
It means that maintaining and updating a unique tool is not the correct strategy to be applied for Network Planning.

The major concerned telecommunication Companies normally use different tools (or different packages integrated on a unique platform) for network Planning. They usually rely on the services of software companies who are in a position to provide quick updates as soon as required.

Therefore, countries' requests for assistance on Network Planning should be dealt with as follows:

a) First, to analyse the Network Planning case, taking into account the different technical aspects of the issue.

b) Second, after reaching the best solution in terms of cost and technical validity, to look for the appropriate partnership with whom to define a Project for the specific Network Planning case.

c) Implementation of the Project under the coordination and/or supervision of ITU-BDT.

This strategy has been endorsed by the World Telecommunication Development Conference - WTDC-02 (Istanbul, March 2002) in Program 2, point 1.3 (herewith attached) and reaffirmed during the last WTDC-06 (Doha March 2006).

WTDC-06 PROGRAMME 2: INFORMATION AND COMMUNICATION INFRASTRUCTURE AND TECHNOLOGY DEVELOPMENT

1.3 Network planning

The selection of new technology hinges on projected needs and consequent network development planning. In developing countries, the needs may be substantially different in urban and rural areas, and infrastructure and technology requirements will differ. In choosing technologies for a new or existing telecommunication network, a very wide range of factors needs to be considered.

The most difficult component of the network to build, and the least cost-effective to maintain, has proved to be the local access network. One of the main problems facing the developing countries is precisely the lack of access to broadband services, and low teledensity.

Adaptation of power-line communications and cable-television networks to provide telephony and internet services has converted them into broadband networks. The technology shall be of low cost, easy to maintain and adapted to the local environment.

The rural population will need to be connected to the information society. Choosing efficient and cost-effective and fast-deployment technologies such as wired and wireless networks will improve accessibility.

The architecture of the information and communication infrastructure is changing to accommodate the requirements of a growing number of ICT-enabled services/applications (broadband, IP, mobile, multimedia, streaming, multicasting, etc.) and evolving to next generation networks (NGN).

New-generation technology is being introduced in the networks, speeding up the convergence process, and obliging planners to apply different specialized up-to-date planning tools.

Network planning is a critical issue for network operators and network service providers in a time of globalization and intense competition. The current telecommunication market requires flexible and adaptive network planning methodologies for evolving network architectures to NGN. Practical guidelines, readily and easily applicable, should continue to be provided to be of use to operators and decision-makers. Moreover, there will be a need for powerful software tools to assist operators in developing their networks. ITU should continue entering into formal partnership agreements with outside partners, positioned to provide the Union with appropriate planning tools suitable for specific network planning requests. Taking into account the above considerations, and in order to contribute to bridging the digital divide, this programme should apply the following measures:

a) providing advice on the design, deployment and maximization of digital networks at an increased pace, including the roll-out of wireline broadband technologies such as, but not limited to, optical-fibre, xDSL, CATV, power-line and wireless broadband technologies, and the establishment of satellite earth stations;

b) facilitating the introduction of digital technology;

c) facilitating the design, production and availability of digital terminal equipment;

d) enhancing technical skills and management know-how;

e) promoting digitization of analogue networks and applying affordable wireline and wireless technologies to facilitate people's access to ICT, thereby also improving quality of service;

f) encouraging research on the information society, extensive networking, interoperability of ICT infrastructure, tools and services/applications to facilitate accessibility of ICTs for all;

g) optimizing connectivity among major information networks via regional ICT backbones in order to reduce interconnection costs and optimize the routing of traffic.

Who should use this Manual

To meet countries' requests for assistance in the field of Network Planning ITU has prepared this Reference Manual on the Telecom Network Planning for evolving Network Architectures.

The Reference Manual is intended for use by network planning experts from telecom operators, policy makers and regulators to facilitate the development of their respective strategies for evolution of the present network architectures and transition to the next generation networks - NGN.

The Reference Manual on the Telecom Network Planning for evolving Network Architectures intends to present an objective and technology neutral view of the issues to be addressed in the planning of the transition to NGN.

Content of the Manual

This reference Manual comprises 8 chapters and 3 annexes, each of which could be updated periodically, due to the rapid changes in the telecom networks.

Typical reason for revisions in the manual could be:

- introduction of innovative network technologies and corresponding planning methods
- appearance of new or improved planning tools on the market
- the need for better explanations in the presented material

Special emphasis in the Manual is given to the role of network planning today and the strong relation to the telecom business.

Chapter 1 provides the objectives and context of the manual as well as the content of the different chapters and relation to other ITU activities and documents.

Chapter 2 will review the aspects that a planner is confronted with when taking decisions on what to do in the network evolution, when to perform the changes, how to perform the corresponding actions and which processes to follow.

Chapter 3 addresses the needed modelling and characterization of services that is required for the planning activities.

Chapter 4 will give generic traffic characterization. Due to the overall modelling of the network for planning purposes, the needed traffic characterization is less detailed than the one needed for detailed system design.

Chapter 5 gives an overview on the economic modelling for planning and different evaluation procedures.

Chapter 6 describes different network architectures - existing telephony network architectures, data network architectures, data invasion of the telecommunication network, the future telecommunication network architectures. Special attention is drown on the next generation network (NGN) and the migration scenarios from the current TDM networks to this goal.

Chapter 7 presents an overview on the diverse models and methods used in the telecommunication network planning.

Chapter 8 lists the main input data needed for network planning. Network planning, especially performed with NP tools, requires collection of numerous data.

Annex 1 presents a portfolio selection of planning tools to support different planning activities. The selection criteria are: capability to model modern technologies, commercial availability and being well proven in the field.

Annex 2 provides selection of most frequent case studies (ie: Network extension, transmission, signalling, migration to NGN, mobile, etc.) in order to illustrate the application process.

Annex 3 contains list with references and glossary of the most frequently used terms and abbriviations.

Chapter 2 – Overview of network planning

Network planning activities evolve with the proper evolution of the network, the services, the technologies, the market and the regulatory environment. These evolutions imply a wider set of options to implement a network than in the past and as a consequence, the importance of careful planning and analysis for alternatives have larger impact on the network capabilities today in order to assure the needed capacities, the associated quality of service and the required investments.

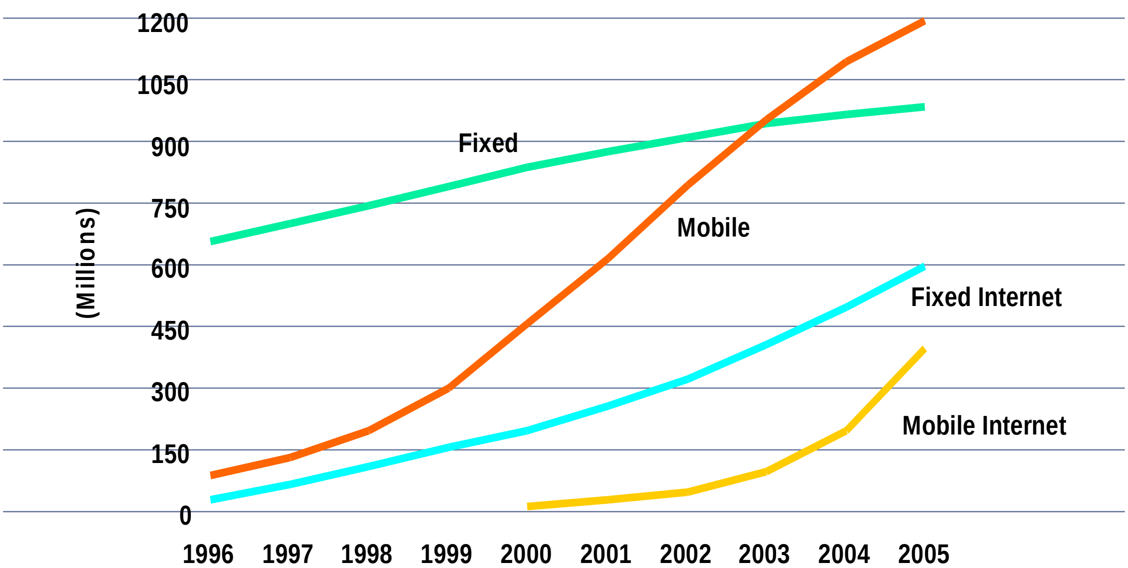
For general feasibility -- to economically justify the move towards the evolving architectures -- one should pay attention to planning of investments and services in a manner which makes sure there is no costly over-investment nor bad utilisation of already earlier made investments, and at the same time ensures fluent migration of the services for the large amount of existing subscribers.

This chapter will review the aspects that a planner is confronted with when taking decisions on what to do in the network evolution, when to perform the changes, how to perform the corresponding actions and which processes to follow.

2.1. Evolution of the Telecom context

- Services demand, associated traffic and revenues are evolving as indicated in the charts below:

Fig 2.1: Subscribers demand evolution in the period 96/05



- New network capabilities are due to the technologies (NGN, 2G to 3G. xDSL, FTTx, WDM, etc., new regulation and competition (market share, service promotion,, etc.) new services in the market (VoIP, VOD, UMS, MMS, etc.), service/platforms convergence through different technologies and pending communication coverage (Geo areas not covered, population not served, network expansion, etc.)

2.2. Requirements to the planners

Under the previous evolutionary context, the planner is confronted to a number of requirements in order to provide answers to the following needs:

- Business Oriented Needs
  - What are the best customer segments to address in multimedia?
  - Which new services have to be introduced through time?
  - What is the best service bundling per customer type?
  - How to increase market share?
  - How to maximize revenues?
  - How to reduce capital expenditure?
  - How to reduce operational expenditure?

- Network Oriented Needs
  - How to forecast multimedia services and related traffic demands?
  - How many nodes to install, especially for NGN?
  - What is best location for new systems and related communication media?
  - What is the best network architecture and routing in NGN?
  - Best balance between built and lease for infrastructure?
  - How to plan capacity evolution and solutions migration towards NGN and towards 3G
  - How to converge service applications and platforms through different access technologies?
  - How to ensure SLA and protection level?

- Operation Support Needs
  - How to evaluate alternatives for direct operation and outsourcing?
  - How to organize and engineer the new operation processes?

- Which IT applications ensure an efficient support to operation?
- How to train labor force on the new operational activities?

2.3. Typical network planning tasks

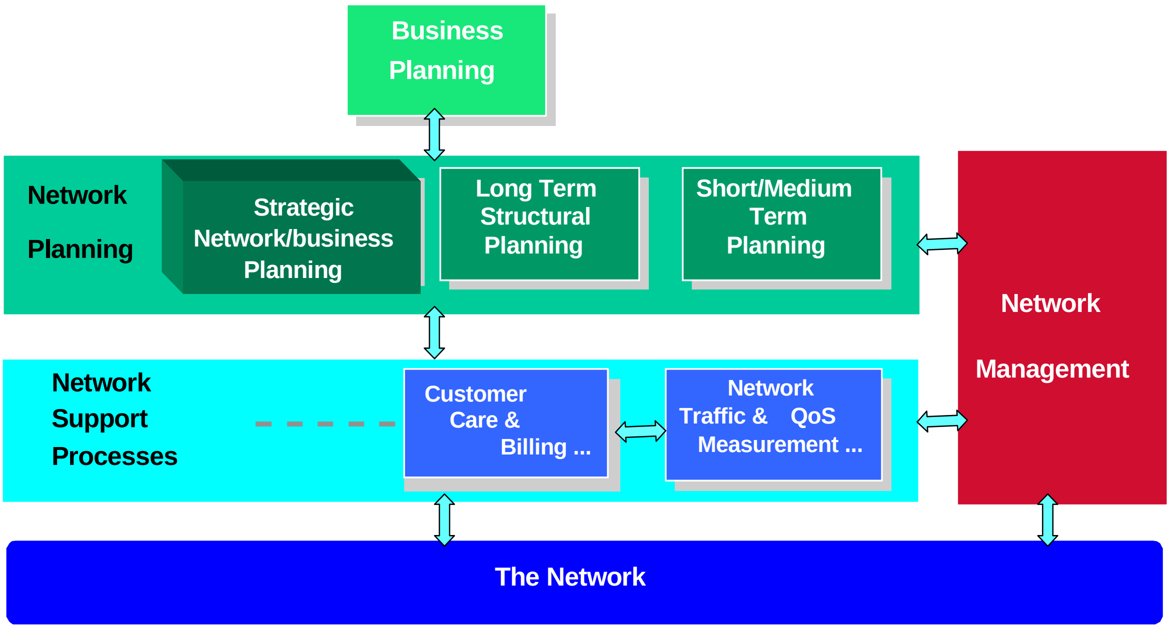
The most typical tasks that the planner has to perform to solve the complexity associated to the previous requirements are summarized as follows:

- Initial situation analysis for economy, customers, services and network
- Problem partitioning
- Data gathering
- Definition of alternatives per scenario
- Mapping solutions per scenario
- Design, dimensioning, location and costing
- Optimization
- Sensitivity analysis to uncertain variables
- Plan selection and consolidation
- Reporting

2.4. Network planning processes

- Due to the high speed of changes both on the environment and the technologies, the traditional planning activities that were performed in an separated way, today have to be strongly interrelated among themselves and to the other network related tasks. For that environment, the Strategic network planning, Business planning, Long term structural planning, Short/medium term planning have to be applied in iterative way with what-if analysis and also communicate with the related Network Management and Operation Support Processes like traffic measurement, performance measurement, etc. as illustrated in the figure:

Fig 2.2: Network Planning Processes and relation with other network activities



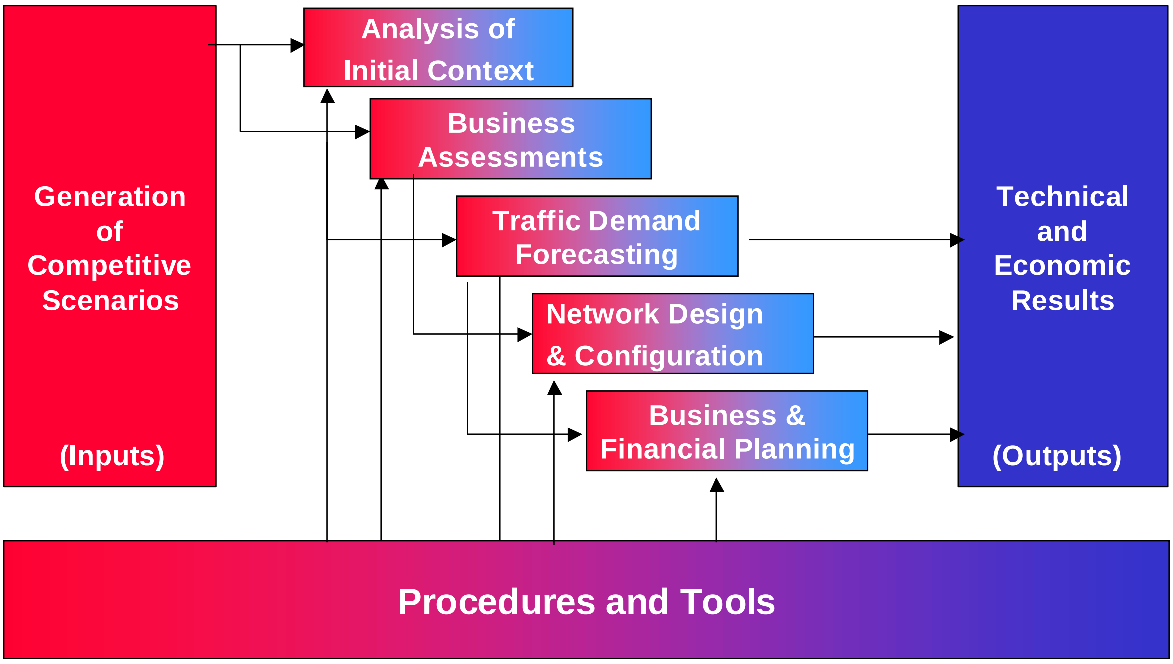
- Data on topologies, architectures, location, routing, etc from long term planning are transferred to the medium term and iteratively to short term activities

- Planning results are transferred to NM applications and vice versa, NM measurements and status are provided as inputs to the planning activities

- Operating System Processes also provide data to the short/medium term planning activities on the traffic demand, performance and Origin/destination flows

Due to the high number of scenarios possible in the competition, a special analysis of those scenarios is needed in order to derive which ones are feasible both from a technical and economical point of view. The following structured procedure is recommended to perform those analyses in an iterative manner:

Fig 2.3: Iterative Planning Sub-Processes for Competition Scenarios



-Telecom network scenarios are generated with the premises derived from realistic market share and competitive situation

-Final objective is to have a quantified design fulfilling the strategy for the operator the requirements of the society and being feasible from the business point of view

-Defined processes and tasks are needed for all solutions and technologies. Internal data and algorithms vary for each technology case

-Feedback among activities is needed to incorporate results of the optimization on the inputs and assumptions

-Business assessment is made at the start of the process to select feasible solutions and discard the ones not being realistic. A more detailed business plan is obtained at the end of the analysis for the selected solutions and providing the business and investment plans

2.4.1 Definition

Network planning addresses all the activities related to the definition of the network evolution in order to allow the transport of an expected amount of traffic demands, taking into account a set of requirements and constraints [2.1]. Depending on the timescale of the network evolution problem under study, three different planning activities can be performed:

- ***Long-term planning*** (LTP), whose objectives are to define and dimension the network parts which are characterised by a long lifetime and large investments for their deployment.

- ***Medium-term planning*** (MTP), whose framework should emphasise the behaviour and the relationships among the sets of entities (nodes, links, subnetworks) and the list of planning actions and procedures which are involved when planning a network to guarantee the convergence towards the established long term plans. Therefore, MTP should have as an objective the capacity upgrading of the network nodes and links; always, following the long-term (LT) deployment strategies of the optical network[1].

- ***Short Term Planning*** (STP), that determines the routes and the telecommunications systems that support a demand. That is, the network has to satisfy the current telecommunications demands with the already installed capacities without additional capital investments.

2.4.2 Long-term planning

The objective of the long-term planning (LTP) is to define and dimension the network aspects which are characterised by a long life time and large investments for their deployment; therefore mainly the topological and technological decisions and fibre cables capacity issues are addressed. LTP, then, elaborates a target network objective for the medium-term planning process; drawing to normally single-period processes.

Two different phases/approaches in LTP are generally considered (cf. Figure 2.4.1):

- the ***strategic planning***, which aims at defining the technology and architecture to be used in the network through the comparison of different options. It is generally based on a green-field approach and uses parametric models and typical values for the relevant network parameters.
- ***fundamental planning***, which uses as input the technology and network architecture selected by the strategic planning and defines the structure of the network[2]. The problems to be faced in the fundamental planning usually are the allocation of functions in the network nodes, the topology planning, the apportionment of functions between the optical and the client layer, the definition of an optimal network structure.

Dealing with LTP the focus of the project has been on the fundamental planning. Unless explicitly stated, LTP and fundamental planning are considered as synonymous in the following chapters.

Being more concrete, LTP defines the following aspects:

- Location and technological evolution of the network nodes.
- Partitioning into subnetworks (domain definition). In this aspect, the hub nodes for interconnecting the different domains should be identified. Additionally, the hierarchy between the different domains, if any, should be established.

[1] These strategies should be the outputs of the long-term planning process.
[2] Generally a green-field approach is used for the fundamental planning as well.

- Logical network structure for the considered network layer. Eventually a mapping of the telecommunications systems on the physical telecommunications infrastructures can be given.

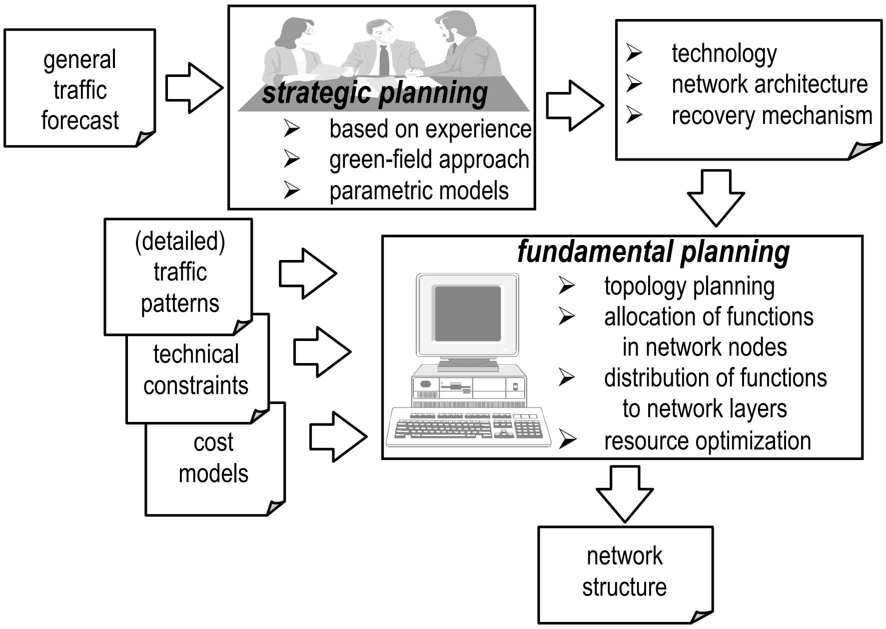


Figure 2.4.1 - Strategic and fundamental planning

The output of LTP is the dimensioned network structure. LTP uses as inputs the following data:

- Single-period long-term demand forecasts.
- Set of possible node locations. Even in the case of a new operator beginning the service in a greenfield zone, it is very usual that an initial set of possible locations is previously identified (own or allied-companies premises are frequently used as the initial set). Of course, this set may be as large as needed and even infinite (meaning that there is no constraint in the node location).
- Set of possible physical paths for the telecommunications infrastructures.
- Architecture to be used in each domain: ring, mesh. This aspect includes the protection/restoration schemes and the general routing/grooming criteria to be used.
- Component and telecommunications infrastructure costs. Normally, non-discounted costs of the target objective are used as minimisation function[3].

The cost elements used in the different cost calculations should have the same precision as the long-term demand forecasts. As these forecasts are normally not too much reliable, it is not worthy at all to use a very complex cost model and to perform very detailed cost calculations.

The timescale of the LTP is normally few years (from 3 to 5). In any case, the LTP exercises are performed to update the results, especially when the demand forecasts have significantly changed or when the NO has to implement the telecommunications installation plan (typically, each year). LTP is also performed whenever a rupture in technology is foreseen.

[3] That means that the cost evolution in time is neglected.

2.4.3 Medium-term planning

The objective of Medium-Term Planning (MTP) is the capacity upgrading of the network nodes and links following the long-term deployment strategies of the optical network. Then, the goal of the MTP is to determine the routing map and node capacities.

MTP is normally performed in a multi-period basis; setting of the different steps for moving from the installed plan (if any) to the long-term network objective (calculated by the LTP).

Being more concrete, MTP should generate the following results for each planning period:

- Detailed routing and grooming for each demand (traffic relation). It should not have conflicts with the defined LTP criteria.
- Telecommunications systems to be installed or uninstalled in all the periods. It should be done according to the MT forecasts and inside the set of nodes and telecommunications infrastructure supplied by LTP.
- Equipment to be installed, upgraded or uninstalled in all the periods. It should be done according to the MT forecasts and inside the set of nodes defined by the LTP.
- Scaling and possible delays in deploying/installing new network elements according to the budget constraints.

For producing these results, MTP should receive the following inputs:

- Network nodes (from LTP).
- Present and potential fibre routes (from LTP).
- Telecommunications systems in use.
- Installed equipment in each node.
- Forecasted demands for each planning period.
- Component costs. It should take into account, installation, upgrading and uninstallation costs of the different systems.

The discounted costs in each period are used[4]. MTP may take into account, as an additional constraint, ***budget restrictions***; that is a limitation of the available budget for the installation/upgrading/uninstallation of equipment in each period of time. This constraint may lead to possible delays in deploying/installing new network elements.

The MTP time scale should be equal to the one for LTP and is subdivided into several shorter periods (typically around one year each), as shown in Figure 2.4.2. In a first step, the LTP process is performed for getting the LTP target network (Figure 2.4.2a). This first step uses the demand forecasts and the installed plant. In a second step, the MTP process calculates the different steps for reaching the LTP target network (Figure 2.4.2b). This second process uses as inputs the MTP multi-period demand forecasts, the installed plant and the LTP plan (generated in the first step). Both steps should be repeated each time the demand forecasts change dramatically; in any case, it is very normal to repeat them in each planning period (T0, T1, …), typically each year. In case of strong variations of the demand forecasts, the LTP

[4] So the MTP cost for each network resource is a function of time and takes into account the depreciation due to diffusion or commercial/technical maturity of the resource.

target may change in each planning period. In this situation, the MTP plan (steps) calculated each year goes towards different targets; something like performing steps towards a "moving" target.

Under conditions of high uncertainty a NO could adopt a different MTP approach (cf. Figure 2c), having its medium-term plans (MTPs) partially disjoint from its long-term plans (LTPs). In this case the results of the LTPs are considered like a set of valuable constraints, rather than an absolute target to be reached. The most important reasons driving this option seem to be:

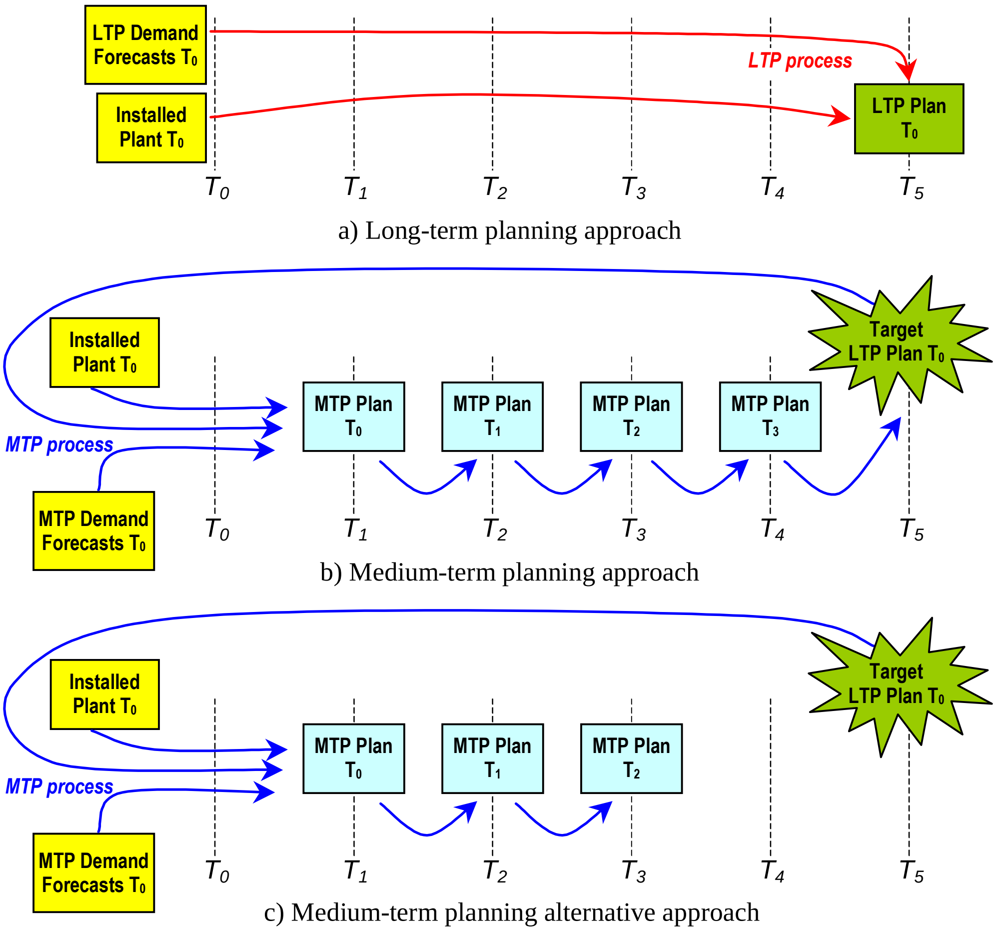


Figure 2.4.2 - LTP and MTP processes

- the Operator considers as useless to plan periods far away in the time since there is the highest probability to have unreliable forecasting leading to unreliable results;
- the optimised results attainable in a static LTP are due to the huge advantage to be able to use network resources in a long period of time selecting the best fitting with the incremental traffic. Unfortunately traffic demands are subject to time constraints (you are not allowed to delay the provisioning of a circuit in order to optimise the network filling) and the network resources' deployment is subject to budget constraints. Consequently through the months and the periods the network grows un-optimised compared to the LTP perspective and it will be impossible to stick to the LTP programs even if your MTP planning algorithm is the best possible one.

2.4.4 The breakdown approach for LTP and MTP solving

Dividing a problem into simpler sub-problems is recognised as an effective solution for very complex problems like the telecommunications network planning. The resulting planning approach, called breakdown approach in this document, is described in this section.

2.4.4.1 Breakdown approach in LTP

The ***LTP for telecommunications networks is a very complex problem*** due to the size and complexity of the realistic network planning tasks. There are several limiting factors that makes the solution of the planning problems difficult, such as the available computing resources and the limited practical applicability of general and unified formalisation of the optimisation problems.

The ***division of the overall planning problem in smaller sub-problems*** (called ***breakdown approach*** in the following) decreases the complexity of the planning activity and it has many positive consequences like simpler solution algorithms, shorter development periods, software re-usability, etc.

The main drawback of the breakdown approach is that it becomes more and more difficult taking under control the global optimisation of the planning problem when the number of sub-problems grows. That is because in this case the optimisation does not only depend on the efficiency of the algorithms that are used to solve the sub-problems, but also on the harmonisation of the sub-problems in a global process. In fact, as outputs of a sub-problem become inputs for another one, an order in the solution of the sub-problems should be identified.

However, realistic network planning problems are so complex that the breakdown approach is unavoidable (in spite of its disadvantages). That is why this approach is widely adopted for the telecommunications network planning problems.

The identification of sub-problems in the planning process is a cumbersome matter. Generally NOs adopt ad-hoc solution for each planning problem, with the aim of taking all the possible advantages in terms of simplification.

2.4.4.2 Breakdown approach in MTP

MTP for telecommunications networks is even more complicated than LTP. First of all that is due to the ***additional outputs*** required to MTP (cf. section 2.4.3). But other aspects add complexity to MTP. A single MTP formulation is: how to maximise flow minimising total cost; solving this problem, the MTP adopts a ***temporised perspective***, in which the time-scale is divided into time-slots, demand matrices has to be carried in each time slot, and the network costs are time related. As in the LTP there are technical constraints to the problem, but ***additional constraints*** can appear as a maximum budget (and the question about how to maximise its utility), and the duty of using previously installed resources but not paid off equipment. On the other hand, MTP decisions (annual periods) will often condition the future network profitability. As a result, planner should also ***consider LTP in his medium-term planning decisions***. If ***technological breakthroughs*** are considered, additional difficulties arise, as the unlimited options of upgrading a SDH network. Because there is a temporary cost evolution and opportunity capital cost, different alternatives appear: when to change to the newest technology (which period), total change instead of partial ones, etc.

The additional difficulties of MTP can be taken under control through a breakdown approach again. So a general methodology to solve a planning problem in the MTP perspective consists in ***dividing MTP into*** two separated and related parts: ***single-period and multi-period***.

Single-period planning process objective is to determine quantity and cost of network resources to meet a single-period incremental demand forecast. It is schematically shown in Figure 2.4.3. The process can be ***further divided into sub-problems***, like in the LTP case. However the problem is generally more complex than the single-period planning applied in LTP, both because more constraints, existing resources and their usage and available unused resources should be taken into account and because more detailed output are necessary.

Multi-period planning process and its relationships with single-period planning process can be viewed in Figure 2.4.4. In order to minimise the total network cost in all the considered periods, a relationship between single-period and multi-period planning is established, while an overall optimisation objective is in target. Results provided by single-period network planning process are the required network resources in the period. There is a relationship between one period and the next one. Multi-period planning process requires information about the total amount of resources of technology **p** and type **i**, purchased in period **t** and disposed in period **j** (new acquisitions cannot be available since purchase time), because these resources are inputs in the following period.

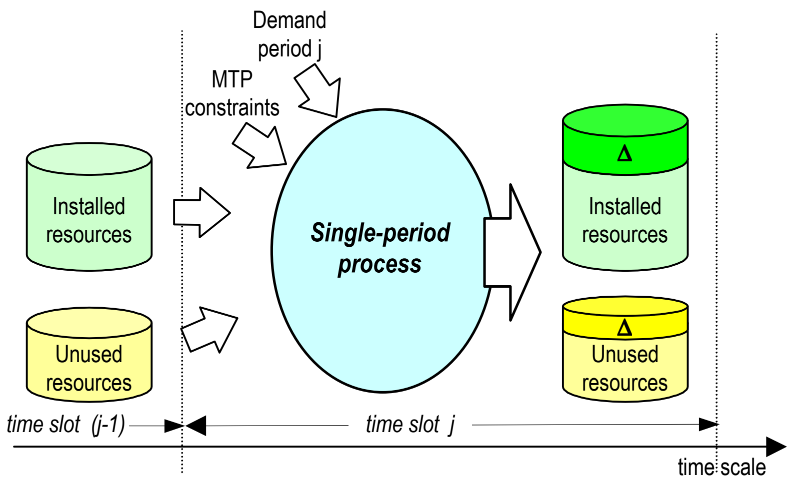


Figure 2.4.3 - Single-period process in MTP

Appropriate network models should be formulated in each period. Furthermore, it is necessary to take into account the different costs involved in the adopted solution: usage of installed and unused equipment (de-installation and new installation costs), usage of uninstalled equipment but purchased for taking advantage of scale economies (installation costs) and usage of purchased equipment (acquisition costs including installation). It is necessary to remark that the equipment cost in each period includes its temporary evolution, and total investments in each period are correctly discounted.

This kind of approach can be viewed as a ***time-scaled decision process***. Each step requires taking a decision among the available alternatives, each taken decision affects the future decision and the overall solution. As the tree of the solutions grows very fast in number of possible branches, different ways of ***reducing the decision tree*** (composed of the whole of solutions) are looked for. Typical levers to prune the decision tree are application of network development strategies, consideration of techno-economical constraints and reduction of the number of time-slots (i.e. MTP periods) taken into account.

As a result of this process, ***optimal network solutions*** are obtained. Three overall ***optimisation goals*** are possible, leading to different network results:

1. optimise at the same time the network cost in each period;
2. optimise the discounted sum of the investments from the beginning to the considered period;
3. optimise the next MTP period, only considering the structural part (network architecture, network structure) of the LTP results as a weak constraint.

These three goals answer to three different MTP interpretations, being the first well suited to the MTP process described in Figure 2.4.2b and the third to the MTP process described in Figure 2.4.2c. The second option can be adapted to both the interpretations of MTP. In each case, the costs taken into account in each period are:

- the cost of the acquired resources up to the considered period,
- the maintenance cost of the resources up to the considered period,
- network operation costs in the considered period,
- net saving costs from disposal of unused resources up to the considered period,
- net saving costs from disposal of used resources before the considered period.

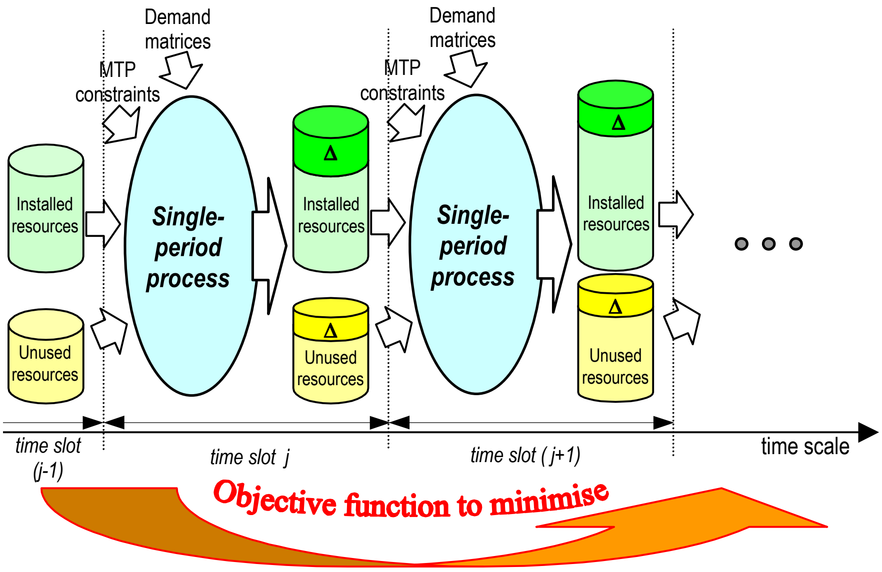


Figure 2.4.4 - Multi-period process in MTP

It is then possible to identify ***difficulties*** that arise in single and multi-period planning phases of MTP, when an optimal solution is looked for. Particularly, in single-period planning

- previously installed and not paid off equipment has to be used;
- limited budget difficult to use. Criteria for establishing network element priorities are needed;
- medium-term planning (MTP) needs to be agreed with long-term planning (LTP);

- criteria for sub-network definition;
- optimisation intrinsic problems;

while in multi-period planning

- temporary cost evolution of network elements has to be defined;
- demand uncertainty exists. Moreover, demand variance increases as temporary horizon does. Planner should consider it when a network solution is selected;
- technological breakthroughs have to be considered. Particularly, upgrading existing networks to NGN ones is needed (minimising cost and risk);
- single-period planning problem has to be resolved in each step;
- different alternatives have to be compared. As required investments are not simultaneous, a financial assessment rule is needed. NPV (Net Present Value) criterion may be used.

Summarising MTP is generally solved applying twice the breakdown approach:

- first a ***time breakdown*** allows to divide the single multi-period problem into several single-period problems;

then a ***LTP-like breakdown*** is applied to divide each single-period problem into simpler, solvable sub-problems.

2.5. Overall plans per network layer and technology

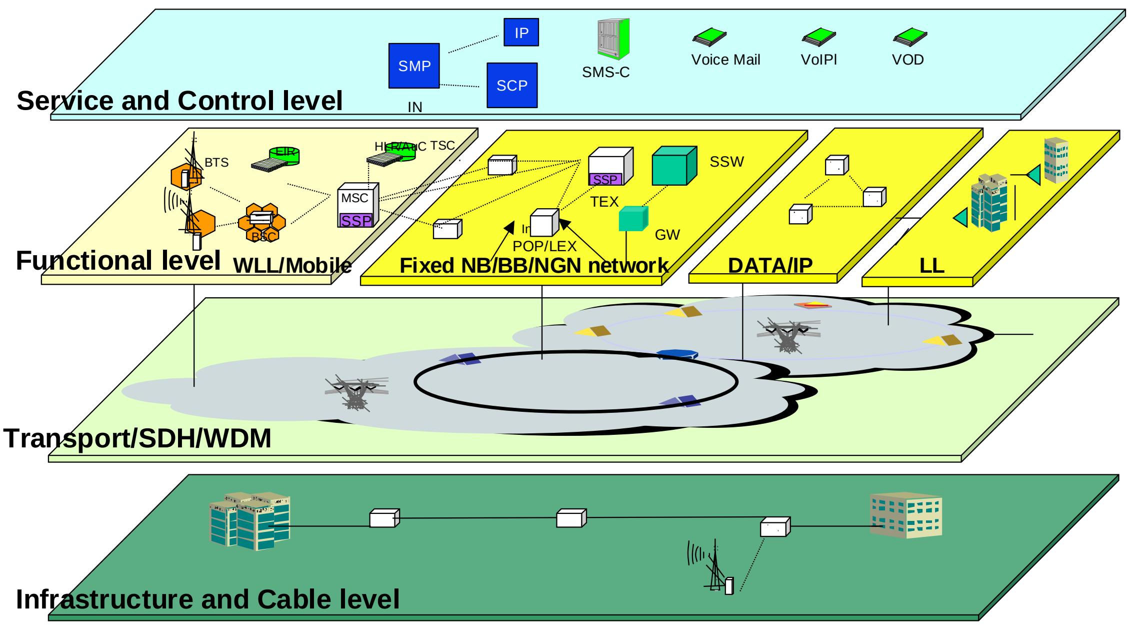
- o The inherent layering structure of the network and related technologies together with the complexity of the overall network implies that the network planning has to be performed also by layers, subnetworks and technologies:

- By Layers in a vertical dimension following the client-server relation (one layer is supported in the layer below and provides resources for the layer up) as indicated: Physical, Transmission, Routing/Switching, and Applications/Services/Control.

- By Segments or splitting of the end to end communication into sub areas as customer premises, access, core national, core international

- By Technologies or underlying technique as FO, WDM, PDH, SDH, PSTN, ATM, IP, NGN, GSM, 3G, etc.....

Fig 2.4: Network Layer Modeling for Planning and Design



- The planning process starts with a first phase for the services and traffic demand projection both at user interfaces and origin to destination interest

-A second phase considers the design for the functional level for the involved functions and technologies like: switching, routing, mobile, data, etc.

- Intermediate results are given as inputs for Transmission and

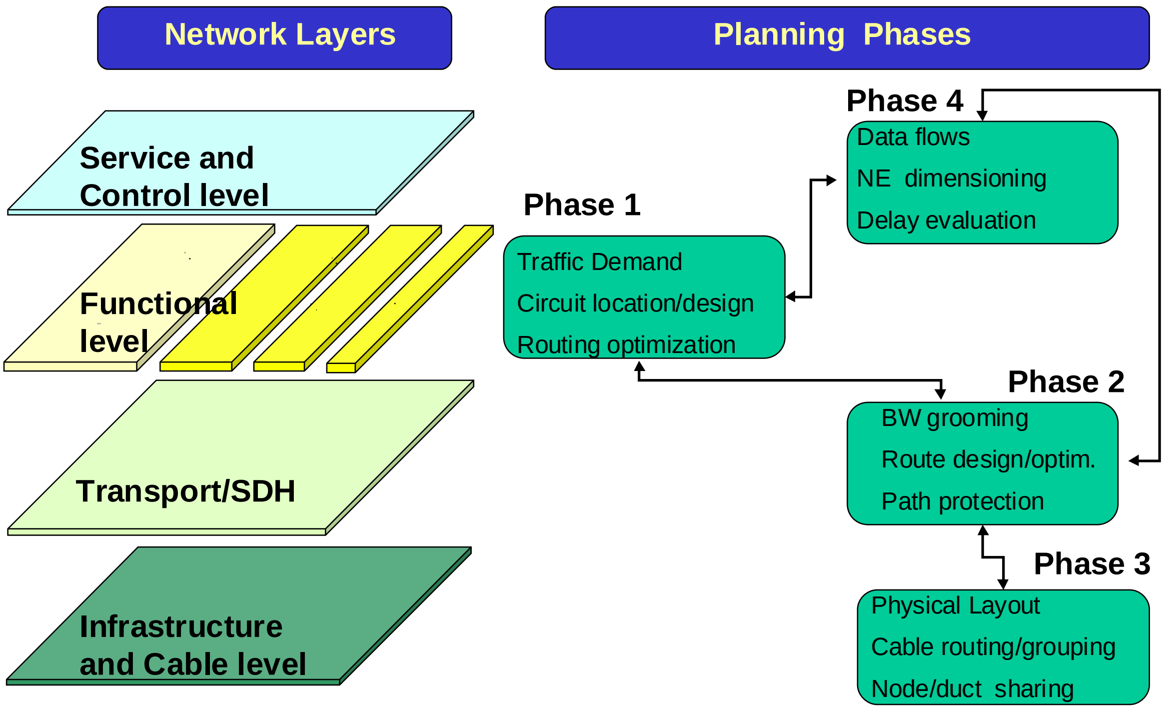
control layers

- In a third phase, the transmission design and planning is performed and the results are provided as inputs to the Physical layer

- Fourth phase contains the planning for the physical elements as ducts, buildings, cables, FO, etc.

- Iteration is made among layers for consolidation, being the functional layer the one that may require more what-if analysis for the central role played among all the other layers and the services/customers.

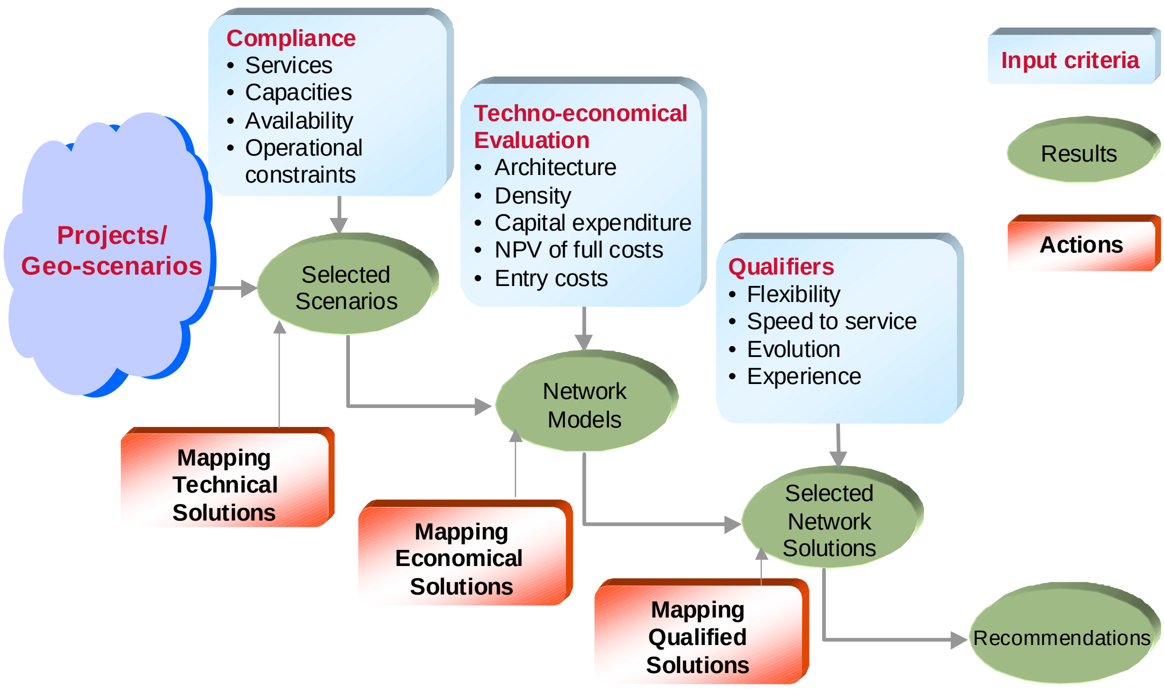
Fig 2.5: Phases of The planning process related to Network Layers



2.6. Solution mapping per scenario

- Due to the large variety of geo-scenarios defined by the combination of customers, services, geo-models, density, consumption, available solutions, etc. the planner has to analyze and decide which solution is going to be planned in more detail per scenario type. The recommended methodology is structured in the diagram with a first selection by technical compliance followed by an economical evaluation of the Cost Of Ownership that is self-explanatory.

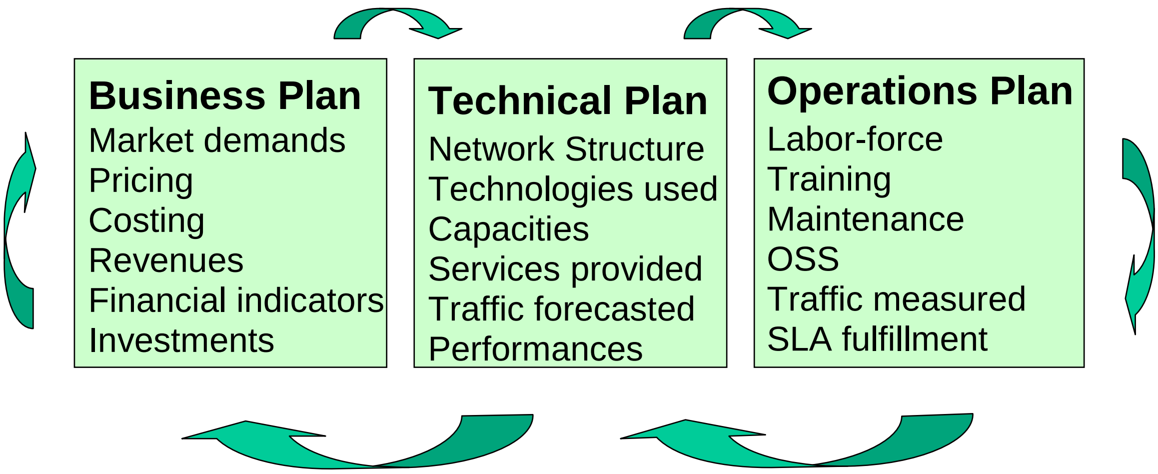
Fig 2.6: Methodology to obtain best techno-economical solutions per scenario



2.7. Relation among technical, business and operational plans

- The number of scenarios and high interrelation among decisions at each level of the organization: Financial, Technical and Operation requires implementing carefully an integrated processing for the information at each stage which is summarized in the following diagrams. The large ranges of variation in many cases and the need to optimize synergies in competition obliges to interchange results between the processes and have an information System across the organization based on Operational Support Systems (OSS) to facilitate consistency and speed of application

Fig 2.7: Rerlation between technical, business and operational plans



Iteration to consolidate all plans

2.8 Planning issues and trends when reaching NGN

Once the migration to NGN enters in the final evolution steps after network topology consolidadtion and access capacity increase to broadband, the specific NGN arquitecture and systems at transit and local segments have to be designed and optimized. Some of the key planning issues to be solved and related activities are summarized:

2.8.1. End to end multiservice traffic demand: Processes for services and traffic flows aggregation.

In a full NGN network, all the service flows need to be modelled with the IP traffic parameters at the five levels of: 1) Calls, 2) Sessions, 3) Bursts, 4) Packets and 5) Bits for each service class at the user origin. Due to the heterogeneous service types, they have to be aggregated by affinity of demand types first (like voice, audio streaming, video streaming, file transfer, etc.) in order to know the demand per user at the network origin points. In a second step, the service types have to be aggregated by Quality of Service category like a) constant speed, b) variable streaming and c) elastic category in order to be able to dimension network resources according to each grade of service and Service Level Agreement per category. A well defined Sustained Bit Rate (SBR) common unit and measurement period of reference (i.e.: 5 minutes) has to be used in order to maintain consistency in the statistical aggregation. This process is illustrated in the following diagram:

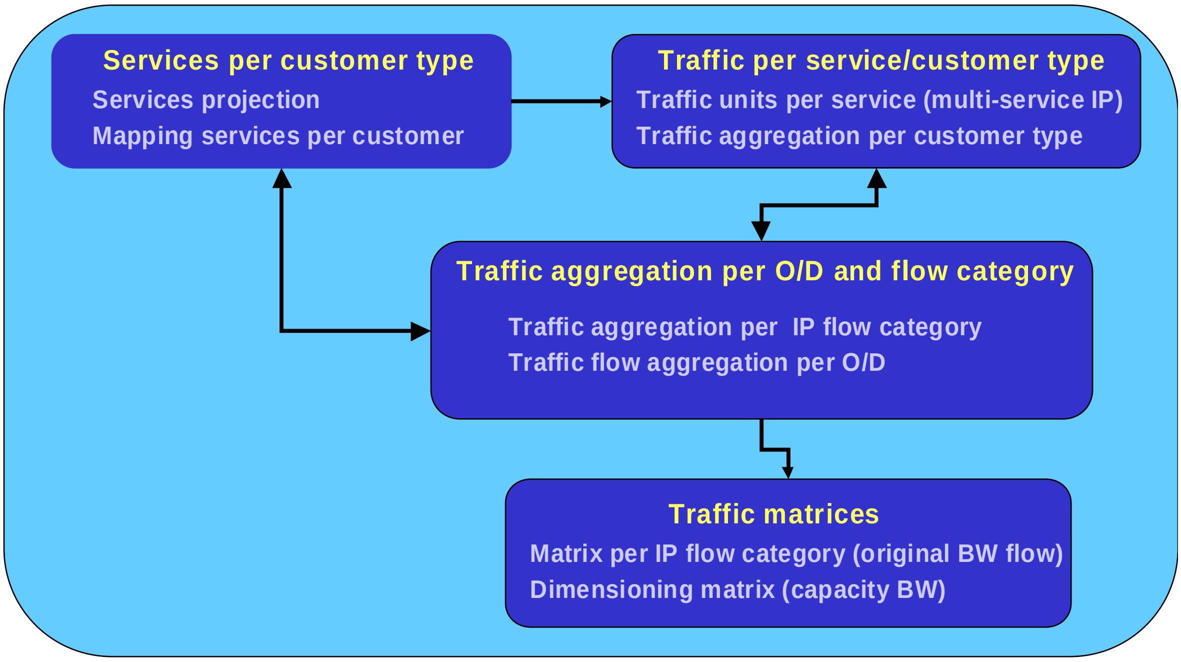


Fig: 2.8.1.- Multiclass Traffic evaluation process at network level

Once that all the matrices for that categories are well defined, the dimensioning of network resources is to be performed, according to the used routing procedures with the corresponding algorithm for each category. Most frequent algorithms are proposed and discussed within the contributions at the International Teletraffic Congress series.

2.8.2. Functionality and location for SSWs.

Up to now most published information describes the NGN network nodes like Softswitches (SSW), Gateways (GW), etc. at a functional level. As soon as a design has to be made and optimized for a mature and large network, a number of new issues appear as:

- Decision on SSW multifunctional versus specific per type of control and application (Fixed network calls, mobile network calls, HLR, NM functions, OSS functions, etc.)

- Number of SSW by functionality, capacity and security level

- SSW locations as a function of all previous constraints and survivability required

- Number and location of GWs as a function of capacities and optimum design either at transit level, local level or hybrid assignment.

These and other more detailed issues are being analyzed today on a per case basis and methodologies are in phase of consolidation for a secure and optimum network evolution in a near time frame.

2.8.3. Design for security at network and information levels

Number of processing nodes at a mature NGN is much lower than in a traditional PSTN and is one of the causes for savings in CAPEX and OPEX. Nevertheless, in order to maintain a proper level of survivability to the network and services, the design criteria cannot just be extrapolated from the current networks and very robust methods have to be applied at the following domains:

- High physical security at topological level with higher connectivity ratios and diversity paths for high capacity and wide influence network nodes.

- High protection level for the energy supply in all key nodes with duplicated or triplicated sources of energy and diverse physical energy paths

- Design of large capacity routes and logical paths with high security criteria

- Design of high security and protected buildings for all involved elements and servers associated to key services

- High level of protection for intrusion, hacking and security for accessibility to SSWs and key NGN resources considering that within IP all network interfaces are potential gates candidates for access to the control.

In most of previous issues, basic mathematical procedures are available for a design, but application criteria need to be much more stringent and combination with new capacities will derive in innovative network designs and configurations. Special attention has to be given to the last one on accessibility that is new for a large, all IP, high quality network and really imply novel concepts and solutions.

2.8.4. Trends towards convergence at different network dimensions

Once that an NGN is implemented at transit, local and access network segments, the convergence possibilities are extended to more domains that the conventional fixed and mobile services. The expected trends in convergence have the following dimensions:

- Convergence at *Network Technology* level in wich synergies will be applied for all the network levels, hierarchies and geographical locations.

- Convergence at *User Terminals* or devices like mobiles, PDAs, GPS, etc. for all functionalities on communication, frequencies, protocols, control positioning, agenda, entertainement, etc.

- Convergence at *User Services* domains with the same functionalities across diffrent network types like fixed, mobile, xLAN, satellite, etc.

- Covergence at *OSS* for all functions on SLA management, Measurements, Service activation, Service management, Quality/performance mgt, Invoicing, Billing, Customer care, Provisioning, Inventory, Application monitoring, etc.

- Convergence at *IT platforms*, Databases and enablers for SSWs, NM and OSS,

Economies of scale for higher customer density, purchasing volume, traffic grouping, system sizes and technology scalation are the main business drivers for the implementation of convergence at the previous identified dimensions.

2.8.5. Planning inter-working and interoperability among domains

When multiple networks reach the NGN maturity stage, a number of inter-working principles have to be planed and designed to ensure the correct end to end operation. The operation of different networks either belonging to the same operator or to different operators is organized in management domains or set of network resources controlled by one management entity. Inter-working and interoperability apply to a given country, a region, an operator or a sub network with a given technological solution.

In order to ensure interoperability between NGN areas and administrative domains, a set of network capabilities have to be planned. Such network capabilities include:

- Converting and trans-coding the media traffic

- Static and dynamic routing configuration, policies and algorithms
- Conversion of name, number or address
- Signalling inter-working
- Exchanging charging and billing information
- Exchanging user and terminal profiles
- Security policy and authentication

The planner has a new set of tasks to specify, locate, design, dimension, cost and optimize the following network interfaces, points and functionalities:

- Network inter-working points, points of presence, peering points, provider edges that have to be deployed at the networks edges with the corresponding functionality, location and dimensioning

- Admission control procedures for the traffic flow acceptance on the base of flow priority, demanded sustained bit rate, Quality of Service, available capacities, network routing algorithms and coordination between the origin based and destination based acceptance criteria.

- Management and filtering functionality across networks for the sensitive control and management information like security level, authentication, authorization, user profiling, non-repudiation, data confidentiality, communication security, data Integrity, availability, privacy, etc.

- Protocol inter-working or adaptation for the different types of traffic flows and the information required to be interchanged for services across domains. Support multiple transport stratum address inter-working scenarios, i.e. inter-working scenarios among different address domains, such as IPv4 and IPv6 address domains, public and private address domains.

- Charging information required for the multimedia services, either based on calls, number of events, information volume or sustained bit rate with a given quality.

- PSTN emulation and simulation functionality to complete calls with origin or destination in existing PSTN networks while maintaining the corresponding characteristics of end to end flows service capabilities and interfaces as well as to ensure service continuity respecting the end-user experience unchanged irrespective of the changing of the core network or the crossing among different network types.

- SLA and e2e QoS management functionality with all procedures to measure and control parameters defined at the SLA such as performance ratios, throughput, delays, packet loss probability, path availability, etc. that have to be coordinated among multiple domains in order to ensure the properties signed with the customers.

The following diagram from the WG3 at the ITU Focused Group for NGN, provides a good reference configuration to illustrate the interconnection and interoperability points that have to be planned

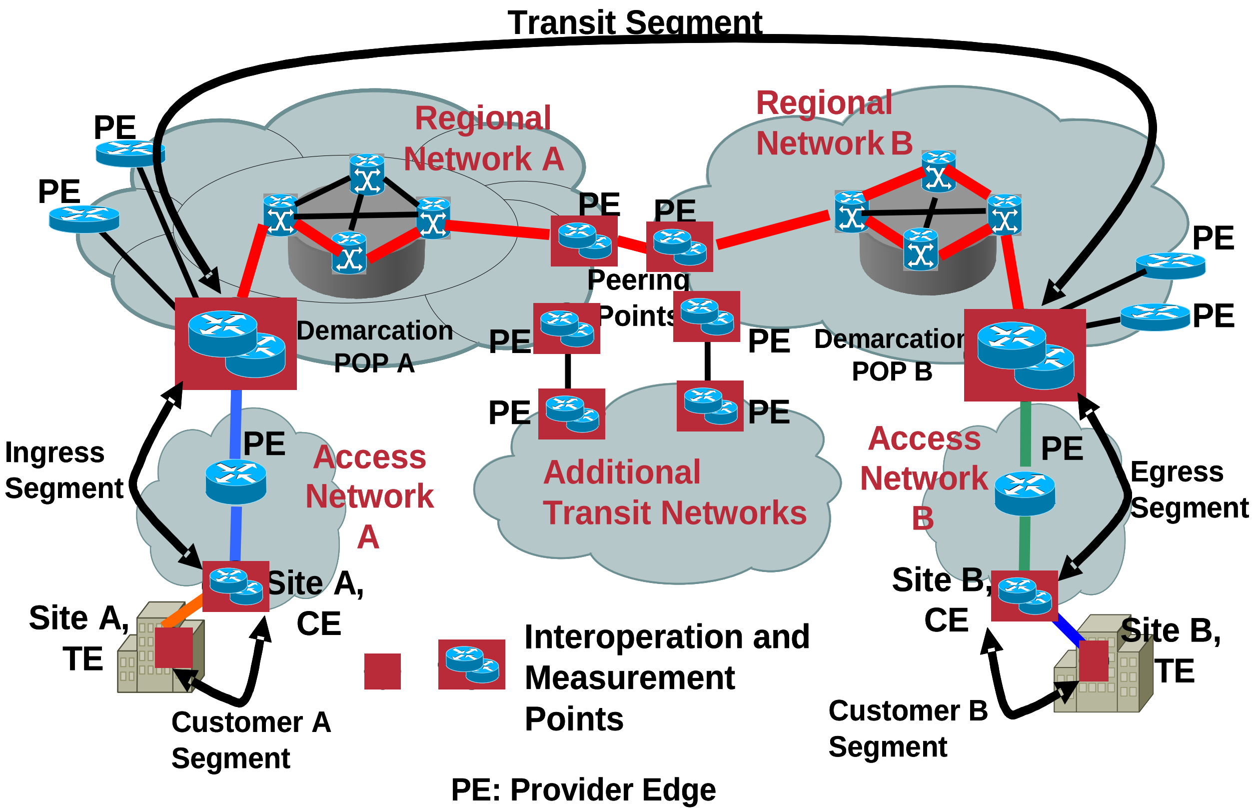


Fig: 2.8.2 Illustration of reference interoperation structure for NGN design by the ITU-FGNGN

2.8.6. Quality of Service considerations

Quality of Service was a concept very well defined, modelled, quantified and measured for classical Telecom networks at ITU both at end to end user service level and at a network and system levels. When networks migrate towards multiservice multimedia services on IP mode, the complexity of quality description enlarges to more domains, parameters and concepts implying an increase of difficulty for definition, measurement and standardization. In addition several entities conceive the quality with different perspectives, as in ITU, ISO, IETF, ETSI, ETNO, etc.

For the point of view of a planner, it is not required to address all operational details but it is needed to focus more on the macroscopic parameters and values that impact on the network dimensioning and costing as those aspects are the ones that have to be quantified with anticipation for the decision making on architectures and business planning.

The variety of different definitions demonstrates the difficulties in assessing all aspects related to the term QoS either focussed on the network provider view or the customer perspective. Basically ITU-T is oriented towards an overall QoS description for the different services with two perspectives:

- Phases of the service life cycle to analyze like: service provision, service enhancement, service support, service connection, service billing, service management, etc.
- Criteria for the quality observation like: availability, accuracy, speed, security, reliability, etc.

It is important to understand that QoS differs from network performance. QoS is the outcome of the user's experience/perception in a global manner, while the network performance is determined by the performances of network elements one-by-one, or by the performance of the network as a whole. This means that the network performance may be used or not on an end-to-end basis. For example, access performance is usually separated from the core network performance in the operations of a single IP network, while Internet performance often reflects the combined NP of several autonomous networks.

Thus QoS is not only defined or determined by measures that can be expressed in technical terms (network performance parameters), but also by a subjective measure which is the user-perceived quality and his quality expectations. Then QoS has to take into account both:

- Customer view: QoS requirements and perception
- Service provider view: QoS offering and achievement

The combination of both views and their relationship forms the basis of a practical and effective management of service quality including the convergence of those perspectives. The views and definitions by ITU-T are taken into account in following sections as a framework for the needed considerations on quality. It has to be emphasized that standardization for quality in NGN context is in progress and a more complete vision will be available at the completion of current Working Groups.

2.8.6.1 QoS parameter types

Quality of Service parameters characterize the quality level of a certain aspect of a service being offered, and ultimately the customer satisfaction. QoS parameters represent subjective and abstract user-perceived "quality" in terms of quantified values.
QoS parameters can be used by service providers to manage and improve how they offer their services, as well as by the customers (end users or partner providers) to ensure that they are getting the level of quality that they are paying for. They have now been used to support commercial contracts such as SLA (Service Level Agreement) formulation and verification. They are also used in call-minute trading, where price is determined by volume and quality grade.

<u>- Objective and Subjective measurements</u>

QoS parameters can be obtained from objective or subjective measurement methods. Objective QoS parameters are obtained from measurement of physical attributes of circuits, networks and signals. They are normally used as internal indicators for service quality characterization and improvement. The subjective QoS parameters are obtained by actually conducting well-designed customer opinion surveys. They are normally used as an external indicator, e.g. for customer relationship management.

<u>-Primary and derived QoS</u>

QoS metrics can be primary parameters that are determined by direct measurement of call characteristics or events, such as circuit noise, echo path loss, or signalling release cause. Alternatively, QoS metrics can be derived from a collection of primary parameters like Statistical calculation, opinion modelling based on measured parameters, opinion and equipment impairment factors, etc.

2.8.6.2 Survey of standardized QoS parameters

Conditions for a parameter to be effectively used as reference for QoS management are: the existence of QoS clear metrics, simplicity of use, proven accurate representations of customer perception, and commonly accepted as standards. This section provides a survey of existing QoS parameters/metrics and QoS class definitions.

A – *Call/session connection succes*

This metric relates to the issue of how successful the called party is reached for the requested session and provide definitions of the commonly used ASR (Answer-to-Seizure Ratio, the ratio of number of answered calls/sessions to number of seizures), and NER (Network Effectiveness Ratio) either for conventional networks or generalized for NGN IP based networks.

B - *Call/session connection delay*

This metric relates to the issue of how long is the waiting time to get to the called party or the call/session after the initial set-up request both for conventional circuit-switched networks and generalized for NGN IP based networks.

C - Conversation and voice quality

This metric relates to the issue of how satisfactory the conversation quality or voice quality is during the call. Conversation or voice quality can be affected by parameters such as noise, echo, speaking volume, transmission delay, and impairments due to voice compression, packet loss, and jitter. The following models are being used:

- *Subjective evaluation*
The most direct way to assess voice quality is via subjective evaluation using human subjects. ITU-T Recommendations provide specifications on a 5-point Mean Opinion Score (MOS) for voice quality assessment (1 = bad, 2 = poor, 3 = fair, 4 = good, 5 = excellent). While subjective evaluation produces results reflecting user perception, it is however costly, timing-consuming and difficult to carry out, particularly in operations. Objective evaluation techniques are therefore employed to estimate user-perceived MOS using signal-based or parameter-based psycho-acoustic models.

- *INMD measurement*
In-service Non-intrusive Measurement Device Voice Service Measurement defines the scope of measurement and accuracy objective of voice grade parameters such as speech level, noise level, echo path loss and echo path delay, based on non-intrusive monitoring of live calls.

- *The E-model*
The E-model, A Computational Model for Use in Transmission Planning gives the algorithm for the so-called E-model as the common ITU-T Transmission Rating Mode for assessing the combined effects of variations in several transmission parameters that affect conversational quality of 3.1 kHz handset telephony. The primary output of the model is a scalar rating of transmission quality but this can be transformed to give estimates of customer opinion. A major feature of this model is the use of transmission impairment factors that reflect the effects of modern signal processing devices. The E-model requires the knowledge of the end-to-end configuration, i.e. networks and terminals, and is intended for network planning purposes.
The transmission quality is calculated taking into account the basic signal-to-noise ratio, including noise sources such as circuit noise and room noise, the impairments caused by delay and the effective equipment impairment factor representing impairments caused by low bit rate codecs. It also includes impairment due to packet-losses of random distribution.

- *PESQ*
Perceptual evaluation of speech quality provides a standardized signal-based psycho-acoustic model (PESQ) to obtain predictions of user-perceived speech quality using an intrusive test-call approach. PESQ, which is a one-way listening model, attempts to generate a prediction of user-perceived MOS by comparing the transmitted reference

speech signal and the received degraded signal. The model takes into account impairment effects due to voice compression and IP network parameters (e.g. jitter and packet loss), in addition to conventional circuit-switched network impairments such as noise and echo. As it is a one-way listening model, the effect of absolute transmission delay is not included.

D - Video transmission quality

Video quality has additional complexity due to the image and visual effects that are treated as subjective evaluation in a 5-grade scale (1 to 5) for both quality-based (bad to excellent) and impairment-based (imperceptible to very-annoying) assessment of television signal quality.

E - IP network performance parameters

ITU-T Rec. Y.1540 defines network parameters that may be used in specifying and assessing IP network performance. They are applicable to network segments or end-to-end connections. The most common defined metrics include the commonly used parameters:

- IPTD (IP Packet Transfer Delay),
- IPDV (IP Packet Delay Variation, or jitter),
- IPLR (IP Packet Loss Ratio), and
- IPER (IP Packet Error Ratio).

These network performance parameters together with the associated target values for different QoS classes are useful for supporting SLA management at the wholesale level as well as at the end-user level.

2.8.6.3 QoS classes and performance objectives

In order to facilitate QoS management for service and business applications, different classes of QoS have been defined either for different service types, or the same service type but different price brackets. Performance targets can be specified for each QoS class in terms of the value ranges of pertinent QoS metrics. The following are examples of QoS class definitions provided by standardization organizations.

- VoIP QoS classes

VoIP QoS classes are defined by levels according overall parameters such as transmission rating R-factor, speech quality (equivalents of known voice-codec quality) and end-to-end delay. The result is a 4 level class as (4 = high, 3 = medium, 2 = acceptable, and 1 = best-effort/no-guarantee) that are used for the negotiation of the SLAs and the network capacity dimensioning

- Guidance for IP QoS classes

ITU-T defines six IP QoS classes based on applications, node mechanisms and network techniques:

Table 2.8.6.1 – Guidance for IP QoS classes

QoS class	Applications (examples)	Node mechanisms	Network techniques
0	Real-time, jitter sensitive, high interaction (VoIP, VTC)	Separate queue with preferential servicing, Traffic grooming	Constrained routing and distances
1	Real-time, jitter sensitive, interactive (VoIP, VTC).		Less constrained routing and distances
2	Transaction data, highly interactive (signalling)	Separate queue, drop priority	Constrained routing and distances
3	Transaction data, interactive		Less constrained routing and distances
4	Low loss only (short transactions, bulk data, video streaming)	Long queue, drop priority	Any route/path
5	Traditional applications of default IP networks	Separate queue, lowest priority	Any route/path

For each QoS class, IP network-performance objectives are defined in terms of value ranges (upper bound) of measured IP network parameters: IPTD, IPDV, IPLR and IPER. Because this guidance is specified from the network perspective, it is particularly useful for SLA support at the wholesale level (between service providers), where end-users' perception may not be directly measurable.

- End-user multimedia QoS categories

ITU-T specifies different multimedia QoS categories from the end-user's perspective. Performance considerations are addressed in terms of three parameters (delay, delay variation, and information loss) for different service applications, including:

- **Audio:** Conversational voice, voice messaging, high-quality streaming audio
- **Video:** Videophone, one-way video
- **Data:** Web-browsing (HTML), bulk data transfer/retrieval, transaction (e-commerce), command/control, still image, interactive games, Telnet, e-mail (server access), e-mail (server-to-server transfer), data low-priority transactions, etc.

- Speech transmission quality

ITU-T defines five categories of speech transmission quality that can be used as guidance in establishing different speech transmission quality levels in telecommunications networks. The definitions provided are independent of any specific technology that may be used in different types of network scenarios under consideration.

Based on the primary output of the E-model, the Transmission Rating Factor, R provides the following definitions of the categories of speech transmission quality in terms of ranges of Transmission Rating Factor R. Also provided are descriptions of "User satisfaction" for each category.
There are procedures to relate R values with other quality parameters like MOS and values lower than 50 are not recommended for any case.

Table 2.8.6.2 – Definition of categories of speech transmission quality

R-value range	Speech transmission quality category	User satisfaction
$90 \leq R < 100$	Best	Very satisfied
$80 \leq R < 90$	High	Satisfied
$70 \leq R < 80$	Medium	Some users dissatisfied
$60 \leq R < 70$	Low	Many users dissatisfied
$50 \leq R < 60$	Poor	Nearly all users dissatisfied

The R-value is a measure of a quality perception to be expected by the average user when communicating via the connection under consideration: quality is a subjective judgment such that assignments cannot be made to an exact boundary between different ranges of the whole quality scale. Rather, the quantitative terms should be viewed as a continuum of perceived speech transmission quality varying from high quality through medium values to a low quality as illustrated in the following Figure.

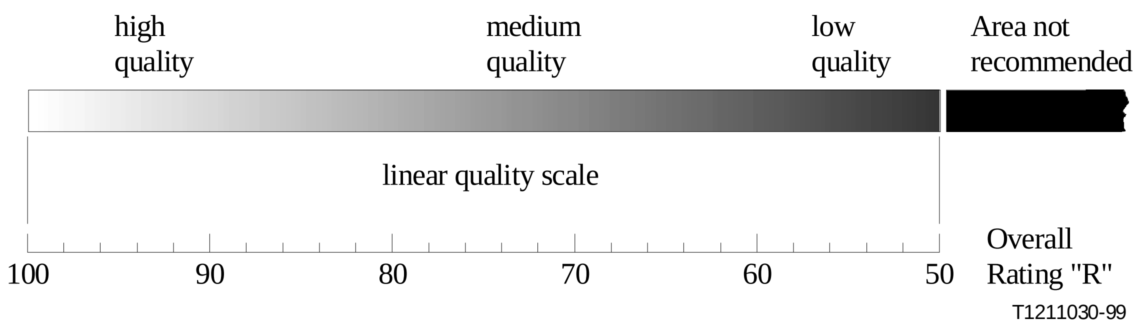


Figure 2.8.6.1 – Judgment of a connection on a linear quality scale

2.8.6.4 Service Level Agreement (SLA)

Due to the higher complexity for quality agreements among related entities motivated by the new parameters, the lack of history of new services and the provisional status of the ongoing standardization process, it becomes essential to exploit the capabilities of the SLAs in order to ensure the appropriate end to end quality across the different service or content providers involved in the provision of a given service.

A *Service Level Agreement* is a formal agreement between two or more entities that is reached after negotiation aiming to define service characteristics, responsibilities and priorities of the involved parties. An SLA may include statements about performance, billing, service delivery but also legal and economic issues.

The part of the SLA which refers to QoS is called a *QoS Agreement* and includes a formal programme mutually agreed by the two entities for choosing, measuring and monitoring QoS parameters. The goal is to reach the QoS agreed upon with the end-user and thus obtain the end-user's satisfaction. Although the definition of a SLA is a bilateral negotiation between the signing entities, from the QoS point of view it should consider at least the description of subsystem or interface for observation, the characterization of traffic flows, the selected QoS parameters with related objectives, the measurement procedures with observation time periods and the related corrective actions when deviations are detected.

Chapter 3 – Service definition and forecasting

A new pleyade of services is being incorporated to the traditional ones in the domain of voice, data and video due to the capabilities offered by the new technologies and the users demand in the information society.

In addition to the multimedia type of new services, it makes sense at the same time to specifically see how the availability of basic telecommunication services can be secured and means for extending their reach cost-efficiently, as it is seen that telecommunication services in the coming ears increasingly become available also to those (numerous) people who have so far lived outside service coverage, or have not afforded such services, and are now becoming the next billion of cost-conscious telecom services users.

The chapter addresses the needed modelling and characterization of services that is required for the planning activities.

3.1. Customer segments

3.1.1. Per socio-economical category: LE, SME, SOHO, Business, High-end residential, Low-end residential, etc.

3.1.2. Per consumption level: stratified per consumption unit (time, events, information volume)

3.1.3. Per type of end user class (innovators, followers, lazars, addicts, etc.)

3.2. Services definition and characterization. Categories

3.2.1. Service definition as voice, data, video, etc.

Service requirements:

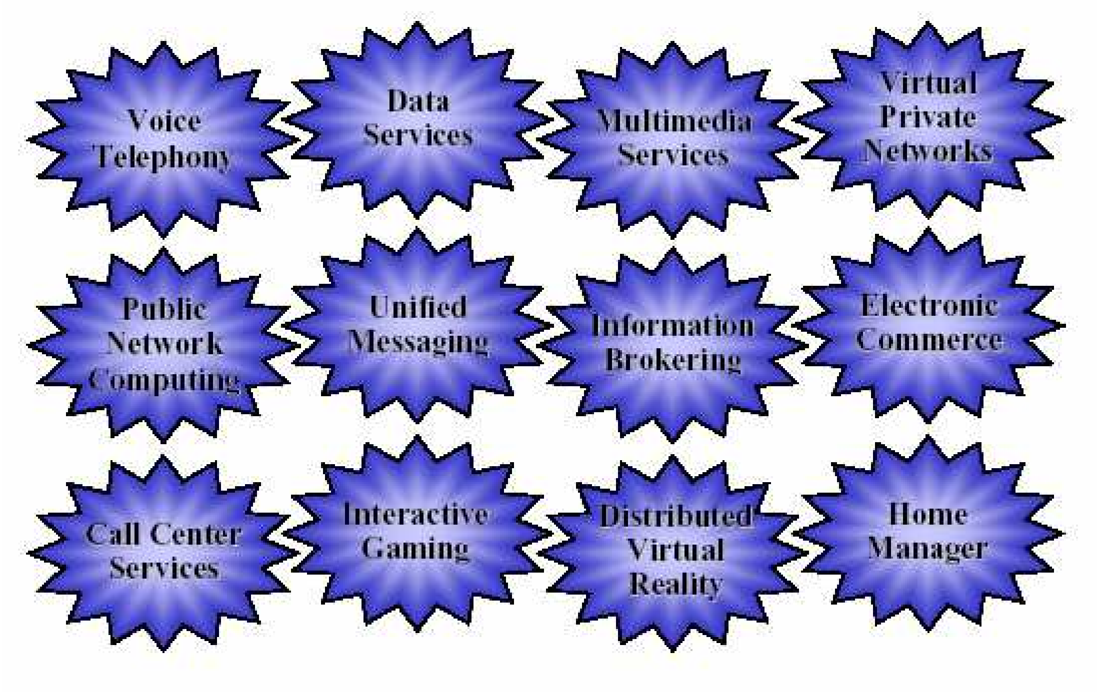
- bring services to customers in a way that is
  - in accordance with the trend to separate the roles of Service Providers, Network Providers, Content Providers
  - future-proof (easy incorporation of new services and network technologies)

- support levels of QoS in terms of delay, jitter, loss, reliability, availability

- support security

- faster access - where is the bottleneck? … and is the problem really speed or prioritization?
- be simpler/cheaper to operate/maintain/manage

Next Generation Service Architecture will support a wide variety of servicesIntroduction of a variety of new services and applications will be possible because of the open interfaces that are typical for NGN.



Voice Telephony **–** NGN will likely need to support various existing voice telephony services (e.g., Call Waiting, Call Forwarding, 3-Way Calling). Rather, it will initially focus on the most marketable voice telephony features.

Data (Connectivity) Services **–** Allows for the real-time establishment of connectivity between endpoints, along with various value-added features (e.g., bandwidth-on-demand, bandwidth management/call admission control).

Multimedia Services **–** Allows multiple parties to interact using voice, video, and/or data. This allows customers to converse with each other while displaying visual information.

Virtual Private Networks (VPNs) **–**allowing large, geographically dispersed organizations to combine their existing private networks with portions of the PSTN, thus providing subscribers with uniform dialing capabilities. Data VPNs provide added security and networking features that allow customers to use a shared IP network as a VPN.

Public Network Computing (PNC) **–** Provides public network-based computing services for businesses and consumers. For example, the public network provider could provide generic processing and storage capabilities (e.g., to host a web page, store/maintain/backup data files, or run a computing application).

***Unified Messaging* –** Supports the delivery of voice mail, email, fax mail, and pages through common interfaces, independent of the
means of access (i.e., wireline or mobile phone, computer, or wireless data device).

***Information Brokering* –** Involves advertising, finding, and providing information to match consumers with providers. For example, consumers could receive information based on pre-specified criteria or based on personal preferences and behaviour patterns.

***E-Commerce* –** Allows consumers to purchase goods and services electronically over the network. Home banking and home shopping fall into this category of services.

***Call Centre Services* –** A subscriber could place a call to a call centre agent by clicking on a Web page (the agent could be located anywhere, even at home - virtual call centres). Agents would have electronic access to customer, catalogue, stock, and ordering information, which could be transmitted back and forth between the customer and the agent.

***Interactive gaming* –** Offers consumers a way to meet online and establish interactive gaming sessions (e.g., video games).

***Distributed Virtual Reality* –** Refers to technologically generated representations of real world events, people, places, experiences, etc., in which the participants in and providers of the virtual experience are physically distributed.

***Home Manager* –** With the advent of in-home networking and intelligent appliances, these services could monitor and control home security systems, energy systems, home entertainment systems, and other home appliances.
The classification on the 12 categories covers most of the services spectrum
and could be subdivide into specific services if needed more detail.

VIDEO distribution services could form a category by itself due to the high importance in demend, traffic, revenues, etc or alternativelly mentioned in an explicit way to know in which category they are.

3.2.2. Service characterization by traffic, bandwidth, etc.

3.3. Services mapping to customer segment

3.4. Service forecasting per segment

Earlier, fixed lines were the way to build telecom services.. Today, the mobile has become also choice in increasing welfare by telecommunication services.

Voice and simple messaging services have become a key element in increasing the welfare of both society and the individual. However, these basic telecommunications services have so far been too expensive to afford, or have not been available at all for those most in need. Moreover, the current business models and cost levels of telecommunication operators are not sufficient to support extending the availability of basic communication services as widely as demanded, but this can be changed..

Affordability and cost of service are clear drivers for the new mobile end-user segments. This requires new thinking in business practices and initial network planning, right through to network operations and maintenance. Many operators are already active in the new user segments. Also, other industry players are recognising the opportunity.

Regulators — both on trade and technology — are in a focal role to define if the telecommunication services are to be made available for wider part of the population than today, contributing to regional development.

For the operators there is no single set of applicable rules for cost reduction. The key areas requiring attention naturally are how to reduce operational expenditure (OPEX) and capital expenditure (CAPEX), minimise average cost per user (ACPU), and enable profitable business from segments with low average revenue per user (ARPU).

The industry believes that lowering the total cost of ownership for consumers — for the benefit of also entry-level segment — will create growth opportunities in low mobile penetration markets.

Basic telecommunications as catalyst for improving welfare

Basic telecommunication services

Basic telecommunication services are defined primarily as voice and simple messaging to other users of telecommunication services on a national scope. They can also include basic data communication services enabling the use of e-mail and access to the Internet. The availability of these services is a significant contributor to the development of local and national economies, including the health of people, education, social contacts, and supporting the government in their effort to serve the nation in the best possible manner.

Focus has shifted to mobile telecommunications

Until the past decade, the implementation of such services was dependent on the availability of the fixed telephony infrastructure, with limited ability to expand these services to previously unanticipated volumes of new subscribers.

Since then, the provision of basic telecommunication services through mobile telephony has changed both the affordability and expandability of the service. What used to be considered as luxury has become a justified commodity in a majority of countries.
However, as the deployment of the mobile telephony infrastructure has taken place through commercial implementation of the services, the service has initially only been made available where the local domestic economy results in individuals with sufficient wealth for these services, i.e. primarily in cities and urban areas.

Lack of telecommunications facilities both in urban and rural areas

The resulting lack of service coverage has so far ensured basic telephony services are unavailable to a significant number of people in some of the more rural areas, where the common challenges of life would favour a wireless telecommunications solution the most.

In addition to the rural areas, basic telephony services are still considered to be inaccessible for large numbers of people living in cities, or around them, either due to a lack of fixed telephone lines, because significant expansions of the existing wireline infrastructure are laborious and time-taking, or because they are not able to afford subscribing to telephone services.

MAJORITY OF NEW TELECOMMUNICATION SERVICE USERS IN THE COMING TIMES WILL COME FROM THE ENTRY SEGMENT

For the 6.4 billion people in the world, there are currently 2.2 billion telephone lines (fixed lines and mobile subscriptions altogether).

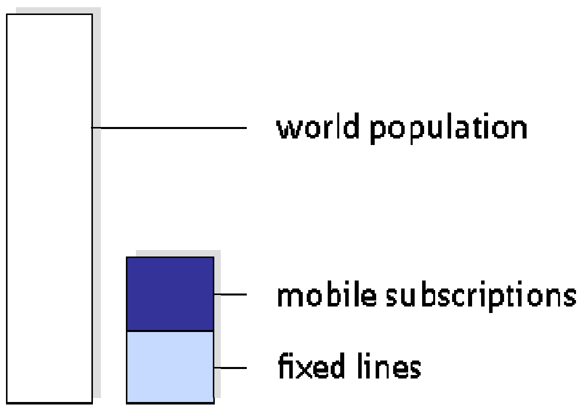


Figure 3.1. World population and number of telephone lines

Mobile lines have been estimated to have surpassed the number of fixed lines in the year 2002, as the more flexible and economic way of building new telephony services.

It is characteristic of the telephone services that their availability is unevenly distributed globally; in some countries telephone penetration already exceeds 100%, whereas in some countries the lack of any basic communication services is severe.

Table 3.1. Population and telephone line statistics examples (from 1998-2002)

Country	Population (millions)	Fixed + cellular lines (millions)	Population minus number of telecom lines (millions)
India	1046	36	1010
China	1284	327	957
Indonesia	231	15	216
Pakistan	148	4	143
Bangladesh	133	1	132
Nigeria	130	2	128
Brazil	176	48	128
Russia	145	44	101

From a service provider's view, the very high number of such potential subscribers can compensate for the limited revenue potential of these new subscribers. By improving their internal efficiency and business support processes, several operators, e.g. in India and the Philippines, are creating a profitable business case out of this new subscriber potential through solutions such as prepaid and short messaging.
As global telecomms penetration is expected to double during this decade, there will no longer be such strong growth in many of the countries with established telecommunications services, the subscriptions in most of the new growth market countries will increase many-fold, creating a challenge for the telecommunication operators in those countries.
While limited availability of service is a key limitation to growth, the clear driver for wide adoption of telecommunication services beyond the current subscribers is the cost level seen by the end-user, including acquisition of a suitable phone, the subscription, and the cost of actual usage. The recipe for increasing affordability of the offering is formed as a sum of multiple contributors:

- operators are contributing by extending their reach in distribution and lowering their overall cost per subscriber, thus enabling themselves to provide more affordable services to end-users on a profitable basis

- telecommunication vendors are contributing by trying to find ways to produce more cost-effective products, with minimized logistics and distribution costs

- in general, governments have contributed by enabling competition on the market (to drive the cost levels down), and re-grading telecommunication products and services into basic rather than luxury items in taxation and duties.

3.4.1. Forecasting methods

The forecasting methods could be divided in the following generalized groups:

- Time trend forecasting methods – it is assumed that development will follow a curve which has been fitted to existing historical data
- Explicit relationships between demand and various determining factors – basic assumption is that these will remain the same in the future
- Comparing various steps of telecommunication development – it is assumed that the less-developed country (or area) will develop to the level of the more developed one
- Personal (subjective) judgment in the forecast – the future will resemble the person's previous knowledge and experience of past developments

As one example the Logistic model from the time trend forecasting methods is presented.

In the Logistic model (Fig. 3.4.1) the development is supposed to follow a curve which first accelerates, then passes a point of inflection, and finally the development slows down and approaches an asymptote, the "saturation level", or "the maximum density".

That model fits very well with change in time of the density of group of customers from a customer class, populated place, region, etc.,

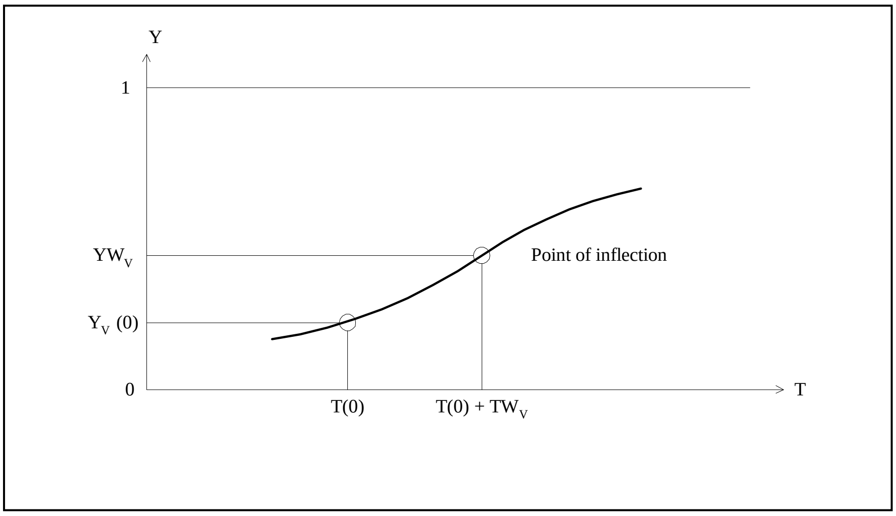


Fig 3.4.1. Forecasting methods- Logistic model

Mathematical expression for the Logistic model function $\mathbf{Y_V}$ and corresponding density calculation $\mathbf{D_V}$ is:

$$D_V = Y_V \cdot DMAX_V$$

$$Y_V = \frac{1}{\left(1+e^{-C_V(T-T_0)}\right)^{1/M_V}}$$

To define the Logistic function is sufficient to know two points from the curve and the saturation value. The two points could be present number of customers and number of customers in some past moment, e.g. one year ago.

Practically is better to perform forecast with the Logistic model on the customer's density, as far as saturation on the density is a clear parameter (e.g. one access point per household), whilst saturation on the customers varies with the time (e.g. population changes with the time).

3.4.2 Demand forecasting per site and per area

A **site** presents group of customers/subscribers, concentrated in one point (one town, village, group of houses, etc.) or location of large business, which will be connected by one link (e.g. business center connected by optical fiber).
Site is typical model for customers from villages and small towns in rural regions or for large business locations in cities.
If site model is used customer densities are defined per site or the site presents one access pint.
Also each site is described with a specified mix between different categories of customers.
The site model could be related to Graph model (Fig. 3.4.2) with customers in the nodes of the graph and arcs of the graph representing geographical distances.

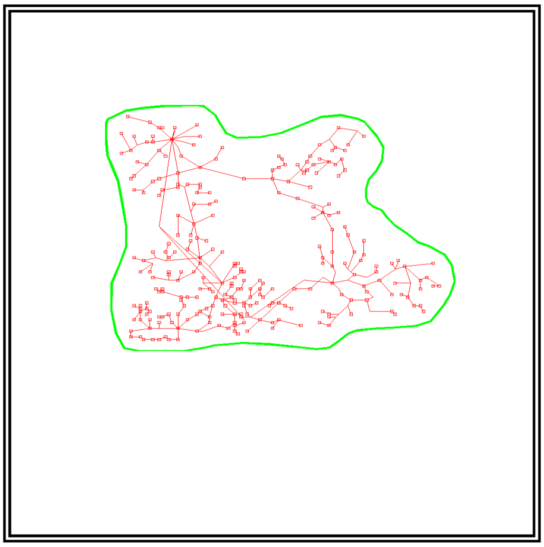


Fig 3.4.2. Forecasting per site

An **area** presents group of customers/subscribers, homogeneously distributed in a geographical area (group of buildings, houses, etc.).
They can be just a few or a substantial number depending upon the studied region and the desired precision.
It is typical model for metropolitan zones (Fig 3.4.3), where in the suburbs areas could be quite big (e.g. one residential district) but in the center they are much smaller (e.g. just one administrative building).
If area model is used customer densities are defined per area, e.g. per square kilometre.
Also each area is described with a specific mix between different categories of customers.

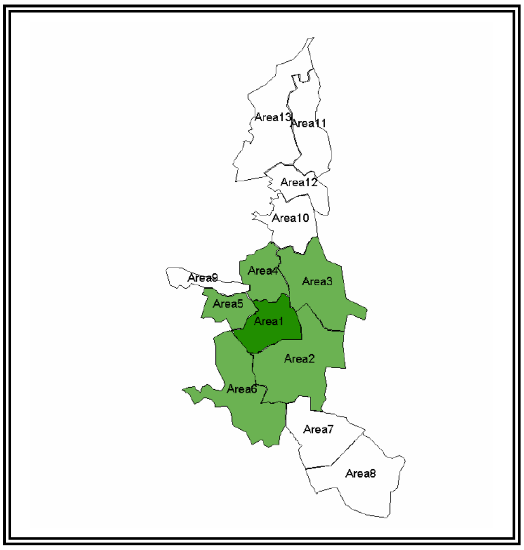


Fig 3.4.3. Forecastung per area

3.5. Service bundling

Service bundling is the packaging of a number of services together in such way that the price of the bundle is less than the price of the individual services or smaller bundled packages of those services.

3.6. Service security

Chapter 4 – Traffic characterization

Traditionally demand is expressed as calls generated by the users and their traffic characteristics. New parameters are Sessions, Information or requests generated at users interface, packets handled at a given resource through the network, Mbits transported through a given network link or path, etc. Due to the overall modelling of the network for planning purposes, the needed traffic characterization is less detailed than the one needed for detailed system design. That generic traffic characterization is given here.

Network interconnections management is a significant area, both from technical perspective (to ensure that services can be provided across network boundaries) and from financial perspective (as the operators want to have visibility-control-charging for use of the network resources).

4.1. Traffic units for service characterization

In teletraffic theory the word *traffic* is used to denote the traffic intensity, i.e. traffic per time unit [4.1].

4.1.1. Traffic in Erlang

Definition of Traffic Intensity: The instantaneous traffic intensity in a pool of resources is the number of busy resources at a given instant of time.

The pool of resources may be a group of servers, e.g. trunk lines. The statistical moments of the traffic intensity may be calculated for a given period of time *T*. For the mean traffic intensity we get:

$$\boxed{Y(T) = \frac{1}{T} \cdot \int_0^T n(t)dt.}$$

where $n(t)$ denotes the number of occupied devices at the time t.

Carried traffic $Y = Ac$: This is the traffic carried by the group of servers during the time interval *T*. In applications, the term traffic intensity usually has the meaning of average traffic intensity. The unit usually used for traffic intensity is *erlang* (symbol *E*).

4.1.2. Bit rate – Mean rate, Pick rate

In a bit stream, the number of bits occurring per unit time, usually expressed in bits per second is called bit rate (BR).

Fig. 4.1.1. shows three different cases of bit streams with different variations of the bitrate.

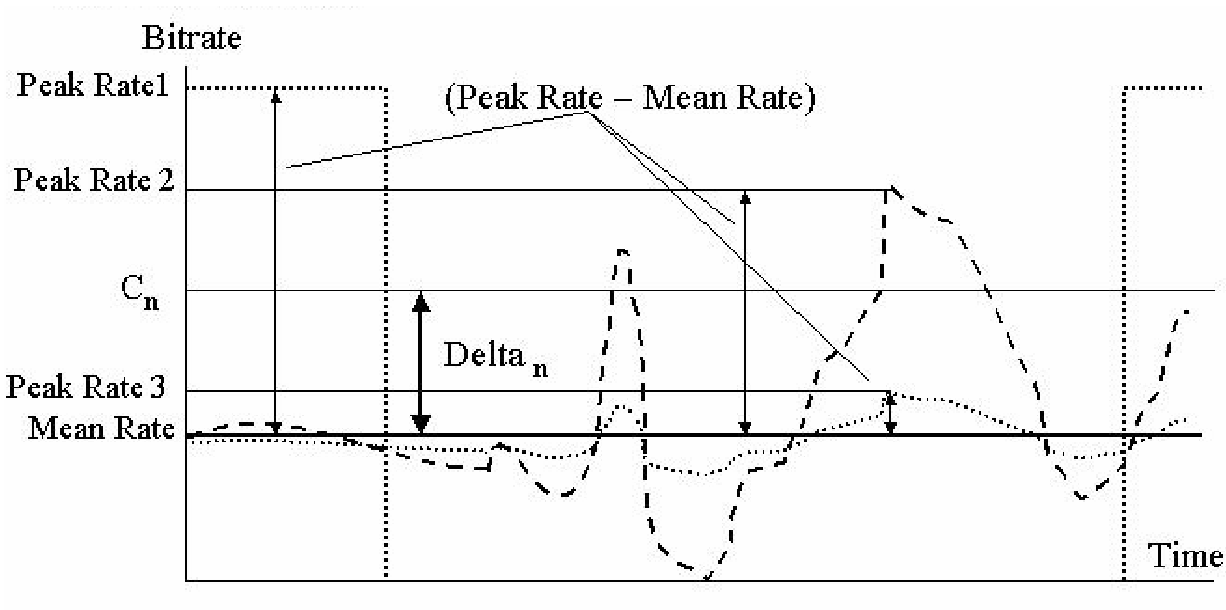


Fig. 4.1.1 Cases of bit streams with different variations

The mean rate is for all 3 cases the same. The peak rates and thus the terms (Peak Rate – Mean Rate) are significantly different.

The effective bit rates or capacities C_n (C_1, C_2, C_3), which must be allocated is the Mean Rate plus a Delta, depending on (Peak Rate – Mean Rate), H (Hurst) and the Buffer Size B.

4.1.3. Total traffic, present of service

The total originating and terminating traffic is relatively easy to be calculated and forecasted, as far as it is proportional to the number of customers/subscribers and the average calling rate/traffic per subscriber.

Usually the originating and terminating traffic per subscriber is measured and known. Also the percentage of the outgoing/incoming long-distance, national or international traffic may be known.

A more presize traffic study will include not only the traffic per subsctiber, but also traffic per each service used by a subscriber from a customer group.

For example, if VoIP service is presented, the required bit rate is specified and a VoIP matrix is created. The VoIP matrix contains the numbers of subscribers using this service simultaneously.

4.1.4. Service and degree of usage

4.2. Reference periods for dimensioning

Busy Hour: The highest traffic does not occur at same time every day. We define the concept *time consistent busy hour, TCBH* as those 60 minutes (determined with an accuracy of, e.g. 15 minutes) which during a long period on the average has the highest traffic.

It may therefore some days happen that the traffic during the *busiest hour* is larger than the time consistent busy hour, but on the average over several days, the busy hour traffic will be the largest.

We also distinguish between busy hour for the total telecommunication system, for one node, e.g. exchange or router, and for a single group of servers (link), e.g. a trunk group. Certain links may have a busy hour outside the busy hour for the node (e.g. trunk groups for voice calls to the USA).

In practice, for measurements of traffic, dimensioning, and other aspects it is an advantage to have a predetermined well-defined busy hour.

4.3. Traffic aggregation process

Traffic aggregation process assumes sub summation of different flows sharing the same identifier (may be additional) across a common path in the network. Thus if only this identifier is used to switch traffic through the network, the flows inside an aggregate are not distinguishable any more.

Some well-known advantages of aggregation – especially in the case of architectures that keep flow states (like ATM and MPLS) – are reduced time and space requirements in core nodes and multiplexing gain for bandwidth and buffer.

4.4. Traffic profiles

The teletraffic varies according to the activity in the society. The teletraffic is generated by single sources, subscribers, who normally make telephone calls independently of each other. A investigation of the traffic variations shows that it is partly of a stochastic nature partly of a deterministic nature. Fig. 4.4.1 shows the variation in the number of calls on a Monday morning. By comparing several days we can recognize a deterministic curve with superposed stochastic variations.

During a 24 hours period the traffic typically looks as shown in Fig. 4.4.2. The first peak is caused by business subscribers at the beginning of the working hours in the morning, possibly calls postponed from the day before. Around 12 o'clock it is lunch, and in the afternoon there is a certain activity again.
Around 19 o'clock there is a new peak caused by private calls and a possible reduction in rates after 19.30. The mutual size of the peaks depends among other thing upon whether the exchange is located in a typical residential area or in a business area. They also depend upon

which type of traffic we look at. If we consider the traffic between Europe and e.g. USA most calls takes place in the late afternoon because of the time difference.

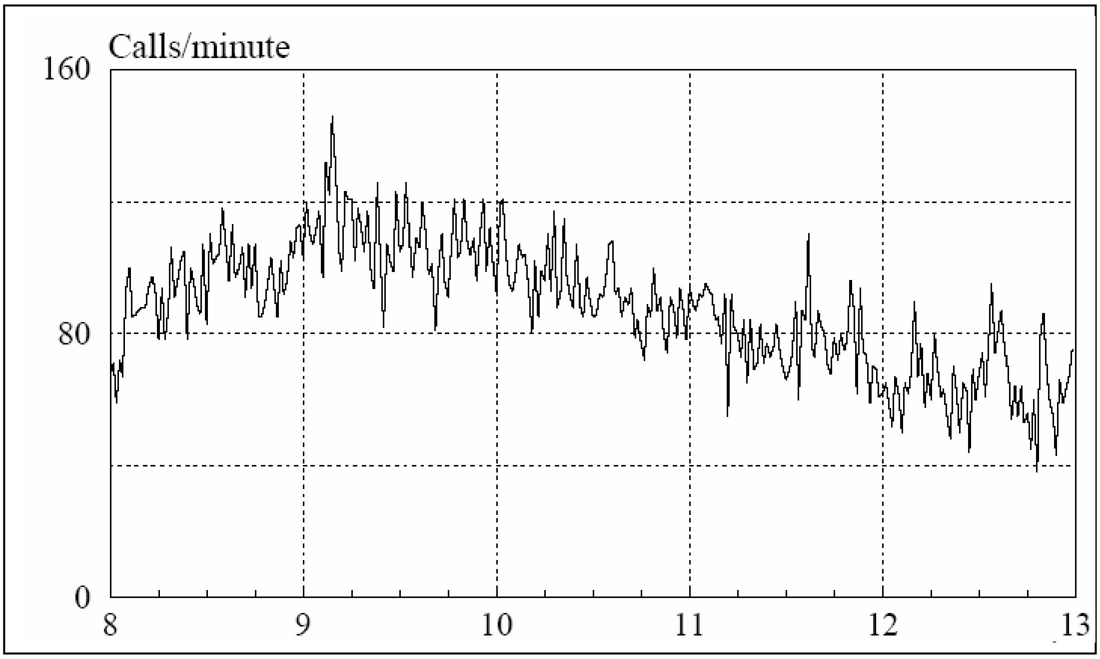


Fig. 4.4.1 Variation in the number of calls on a Monday morning

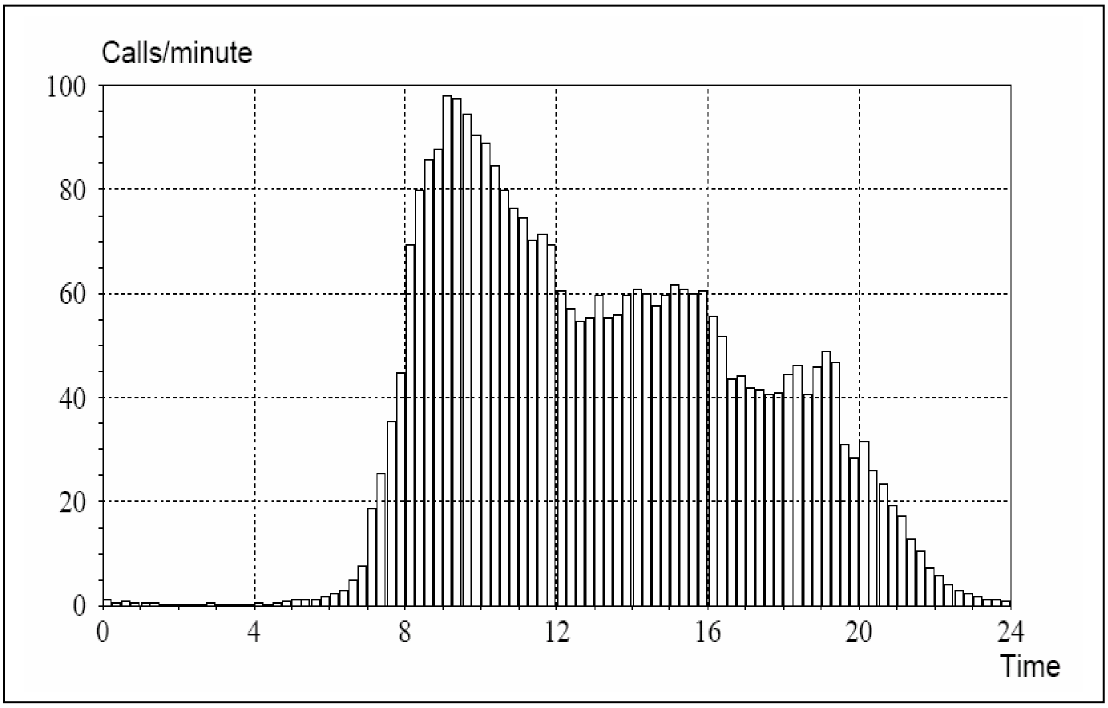


Fig. 4.4.2 Typical traffic profile for 24 hours period

The variations can further be split up into variation in call intensity and variation in service time.

4.5. Origin/destination of the traffic flows in Local, Metropolitan, Regional, National, Continental and Intercontinental networks

The bases for effective network planning are the traffic data between each two nodes of the network.

Such traffic values are typically shown in an origin-destination traffic matrix, based on the origin/destination of the traffic flows in local, metropolitan, regional, national, continental and intercontinental networks.

The traffic matrix presents point-to-point traffic between nodes of local, metropolitan, regional, national, continental and intercontinental networks.

On Fig 4.5.1 is illustrated a set of traffic matrices consisting of one traffic matrix for each service.

i \ j	1	2	- - -	Σ	LD	Σ
1						
2				$A_{iD}(T)$		
- - -						
Σ		$A_{Dj}(T)$		$A_{DD}(T)$		
LD					0	- - -
Σ					- - -	

Fig. 4.5.1 Set of traffic matrices - one traffic matrix for each service

4.6. Interest factors, i.e. attraction coefficients between areas or cities

Normally the total originating and terminating traffic is known and has to be distributed in the traffic matrix.

Also the percentage of the outgoing/incoming long-distance, national or international traffic may be known.
The distribution of point-to-point traffic could be done:

- Based on measured traffic matrix

- With fixed percentage of internal traffic
- With Interest factor or Destination factor method

One well known and used in the practice method is the Kruithof double factor method.

The traffic values in the traffic matrix, at present, are assumed to be known and so is the future total originating and terminating traffic, i.e. the row and column sums.

The procedure is to adjust the individual traffic ***A(i, j)*** so as to agree with the new row and column sums:

A(i, j) is changed to

$$A(i,j)\frac{S_1}{S_o}$$

Where, S_o is the present sum and S_1 the new sum for the individual row or column.

4.7. Traffic evolution

4.8. *Traffic models*

In this subchapter we shortly review the classical teletraffic theory as background for a very simple and general model which is applicable to performance evaluation of both circuit switched multi-rate networks and packet switched networks (IP-based networks).

4.8.1. Introduction – traffic engineering

Teletraffic theory is the use of mathematical, numerical and simulation models for design and resource allocation in telecommunication networks. The development of teletraffic theory started about 100 years ago. The pioneering works in this field was that of the Dane Agnar Krarup Erlang, whose works were published between 1909 and 1928 [4.2].
When we model a communication network we have to include the following elements: the traffic (e.g. subscriber behavior), the system (e.g. topology, link capacity) and the strategy (e.g. routing strategy, priority, accessibility). We want to find the network performance (e.g. loss probability, mean waiting time) when we know the above elements. By this we may design and dimension for an optimal system. First we look at circuit switched systems, then on packet switched systems, and finally we present an integrated model which is independent of time distributions of the process, only the mean value being of importance.

4.8.2. Traffic concepts

The term traffic means traffic intensity, i.e. traffic per time unit, and we have different traffic concepts.

Carried traffic Y *:* the traffic carried in a group of channels is equal to the average number of busy channels. The carried traffic is obtained from measurements during, typically 15 minutes or one hour.

Offered traffic A*:* in mathematical models we operate with the concept offered traffic, which is defined as the traffic carried when there always is a sufficient number of channels. The offered traffic may also be defined as the average number of calls offered per mean service times. If the average number of calls (arrival intensity) is λ and the mean holding time is s, then we have
$A = \lambda . s$

Lost traffic $A - Y$ *:* when the number of channels n is limited, then some calls may be lost. The lost traffic is the difference between the offered traffic and the carried traffic.

Above we have implicitly assumed that all calls use one channel as in the plain old telephone systems. In digital systems we may have calls with individual bandwidth requirements (slot-size).

Then we have to specify whether the traffic is measured in connections or in channels.

For data communication networks we often measure the traffic in bits or bytes per second. The traffic offered is related to the capacity of the link and we consider the utilization ρ, which is the proportion of capacity used.

4.8.3. Traffic variations

The actual traffic observed is varying during the day, week, month, and year. In Fig. 4.8.1 we have typical variations for conventional telephone traffic during the day. Other services and traffic types have other patterns of variation.

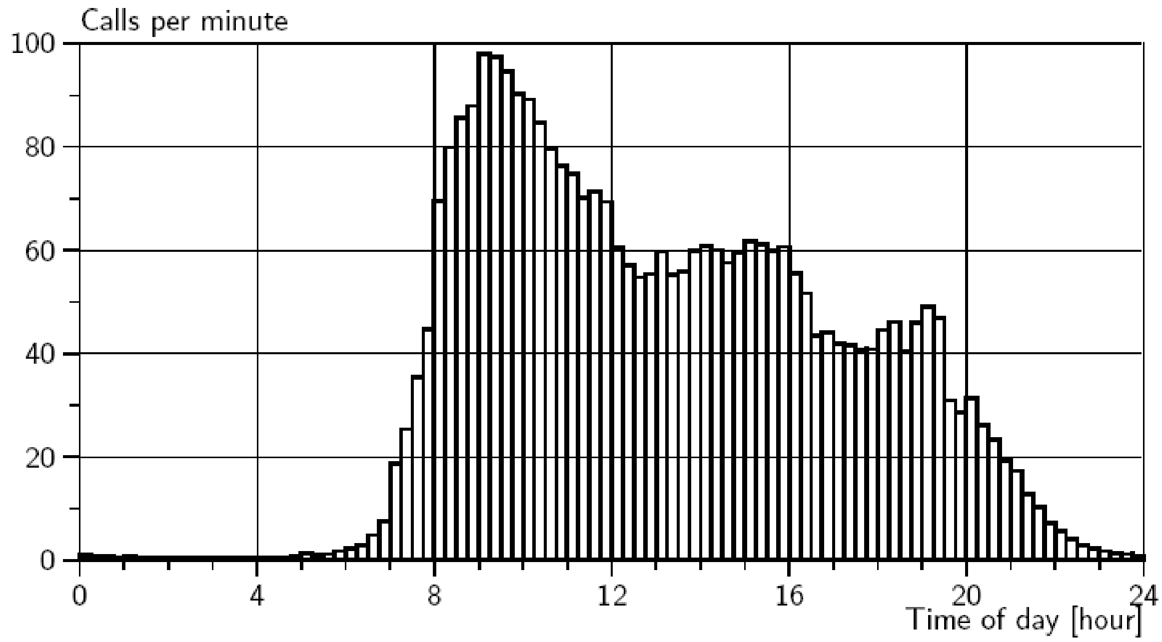


Figure 4.8.1: The mean number of calls per minute to a switching center taken as an average for periods of 15 minutes during 10 working days (Monday - Friday).

In Fig. 4.8.2 we show Internet traffic measurements. Cellular mobile telephony has a different profile with maximum late in the afternoon, and the mean holding time is shorter than for wire-line calls. By integrating various forms of traffic in the same network we may therefore obtain higher utilization of the resources.

The highest traffic does not occur at same time every day. We define the concept time consistent busy hour, TCBH as the 60 minutes (determined with an accuracy of 15 minutes) which on the average has the highest traffic during a long time period.

A network is dimensioned for the time consistent busy hour. The load of processors may be proportional to the number of occupations, whereas the load of a link is proportional to the traffic.

Different parts of a network may have different busy hours. Only from measurements are we able to get knowledge of the actual traffic variations and the busy hour load, which is the basis for dimensioning.

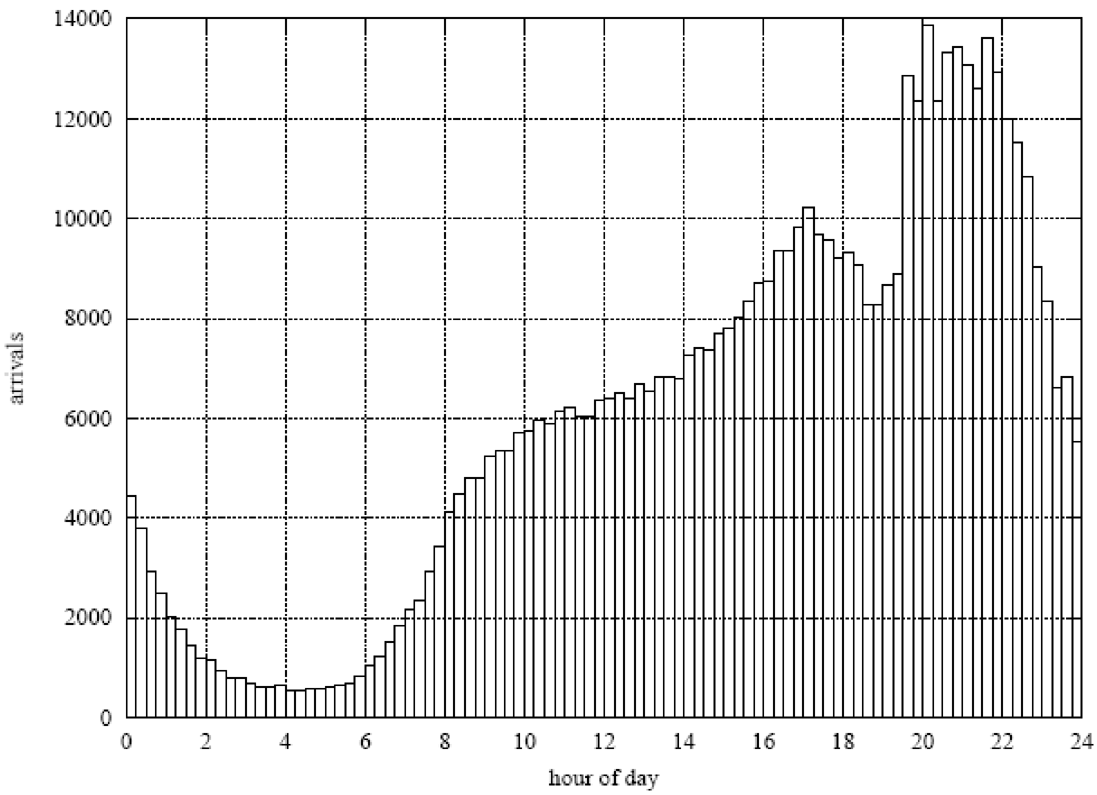


Figure 4.8.2: Number of calls per 15 minutes to a modem pool.

We consider both loss systems and delay systems. Loss systems are common in circuit-switched telecommunication networks whereas delay systems are common in data communication networks.

4.8.4 Loss systems

4.8.4.1 Grade of Service parameters

We distinguish between several performance parameters depending on the system and strategy considered.
For loss systems the main performance parameter is the blocking or congestion probability. This can be defined in several ways:

- The *time congestion E* denotes the proportion of time the system is blocked
- The *call congestion B* denotes the proportion of call attempts which are blocked.

- The *traffic congestion C* denotes the proportion of the offered traffic which is not carried.

The call congestion *B* is typically observed by the user who initiates call attempts. For traffic engineering the relevant measure is the traffic congestion *C*.

4.8.4.2 Erlang's loss systems

The most successful and simple model is Erlang's loss system where the blocking probability is given by Erlang's B-formula. The traffic is described by the offered traffic *A*, the system (only one link) by the number of channels, and the strategy is full accessibility with lost calls cleared.

The above-mentioned three elements of the model are each described by only one parameter. Only single-channel calls are considered. The network performance is described by the blocking probability $E_{1,n}(A)$, i.e. the probability that a call attempts is blocked because all n channels are busy.

$$E_{1,n}(A) = \frac{\frac{A^n}{n!}}{1 + A + \frac{A^2}{2!} + \cdots + \frac{A^n}{n!}} .$$

For Erlang's loss system time, call, and traffic congestions are equal. The state probabilities are given by the truncated Poisson distribution, and when the number of channels is very large this becomes a Poisson distribution.

This model has been very successful for traffic engineering. The background for this success is that the traffic is very well modeled by one parameter only. The underlying mathematical assumption is a Poisson arrival process. This is fulfilled when the traffic is generated by many independent users, which is the case for telephony. If the arrival process is a Poisson process, then the model is insensitive to the holding time distribution, which means that only the mean holding time is of importance. So the model is very robust to the traffic and models the real world extremely well.

Improvement function:

This denotes the increase in carried traffic when the number of channels is increased by one from n to n + 1:

$$F_n(A) = Y_{n+1} - Y_n \,, = A\{1 - E_{n+1}\} - A\{1 - E_n\}\,,$$
$$F_n(A) = A\,\{E_n(A) - E_{n+1}(A)\}$$

4.8.4.3 Engset's loss system

The Poisson arrival process is the most random process, and the calls are generated by a very large number of independent sources, each having an infinitesimal calling rate. In many real systems the number of users is limited, and the arrival process is more regular or smooth than random traffic.

This is modeled by Engset's loss system where we have a finite number S of users (traffic sources) which alternates between the states *off* (= idle) and *on* (= busy). When a source is idle it generates γ calls per time unit (mean inter-arrival time = $1/\gamma$. It is on during a mean holding s. When it is on it generates no new calls. If we let $\beta = \gamma \,.\, s$ and consider the strategy lost calls cleared, then the blocking probability is given by Engset's formula:

$$E_{n,S}(\beta) = p(n) = \frac{\binom{S}{n} \cdot \beta^n}{\sum_{j=0}^{n} \binom{S}{j} \cdot \beta^j}\,, \qquad S \geq n\,.$$

The state probabilities are given by the truncated Binomial distribution, and when $S \geq n$ this becomes the Binomial distribution. For the same number of channels and the same offered traffic, the Engset system will have lower blocking probability than the Erlang system because the offered traffic is more smooth. For Engset's loss system we have $E \geq B \geq C$.

It can be shown that the call congestion is the time congestion when the number of users S is reduced by one:

$$B_{n,S}(\beta) = E_{n,S-1}(\beta)\,, \quad S \geq n\,.$$

The traffic congestion can be obtained by:

$$C_{n,S}(\beta) = \frac{S-n}{S} \cdot E_{n,S}(\beta)\,.$$

For practical applications we should always use the traffic congestion.

4.8.4.4 Peakedness

We characterize variations of traffic by peakedness. Even if the traffic is stationary, i.e. there are no variations of the parameters describing the traffic process, then the traffic intensity is fluctuating around the mean value *m* (measured in channels) because we can only describe the traffic by a statistical distribution. The fluctuations around the mean value are described by the variance *v*.

The *peakedness Z* is defined as $Z = v/m$ and has the unit [channel]. For the offered traffic in Erlang's loss system the peakedness is $Z = 1$, because the Poisson distribution has $m = v$. The offered traffic is the carried traffic when number of channels is unlimited and the carried traffic is the mean value of number of busy channels ($m = A$). For the Engset model (and Binomial distribution) $Z < 1$. In fact, we have $Z = 1-A/S$ and number of sources *S* must always be greater than the offered traffic *A*.

For $Z = 1$ the traffic is random, whereas for $Z < 1$ the traffic is smooth. Below we consider overflow traffic with $Z > 1$ which is peaked or bursty traffic. The traffic congestion will be almost proportional with *Z*. We will characterize a traffic stream by mean and variance or peakedness.

It is noticed that peakedness has the dimension [channels]. Therefore, it is proper for circuit-switched networks, whereas for packet-switched network the coefficient of variation v/m^2 is more appropriate.

Above we have used the parameters (S, β) to characterize the traffic streams. Alternatively we may also use (A,Z) related to (S, β) by the formulae:

$$\begin{aligned} A &= S \cdot \frac{\beta}{1+\beta}, \\ Z &= \frac{1}{1+\beta}, \\ \beta &= \frac{1-Z}{Z}, \\ S &= \frac{A}{1-Z}. \end{aligned}$$

In addition to Erlang and Engset model we also have the Pascal model which has peakednes $Z > 1$.

If we let *S* and β be negative in the above formulae, then we get the Pascal model.

Another model with $Z > 1$ is the Interrupted Poisson process [4.1].

4.8.4.5 Overflow traffic

For planning circuit switched networks with e.g. alternate routing we have to be able to characterize the traffic which is blocked from one link and routed via another link.

The basic methods for this problem is the Equivalent Random Traffic (ERT) method by Wilkinson and the equivalence method of Fredericks-Hayward.

Given an Erlang loss system with *n* channels and offered traffic *A* we are able to derive the mean value and peakedness of the blocked traffic:

$$\begin{aligned} m &= A \cdot E_n(A)\,, \\ \frac{v}{m} &= Z = 1 - m + \frac{A}{n+1-A+m} \geq 1\,. \end{aligned}$$

We may also for given mean value *m* and peakedness *Z* solve the two equations and _find (*A*, *n*) which is called the equivalent group. The idea of the ERT-method is to find the total mean value and variance of all traffic streams offered to a group, and then replace this system by an equivalent Erlang loss system.

The method of Fredericks-Hayward is easier to apply. This method proposes that a system with *n* channels, which are offered *A* erlang with peakedness *Z*, has the same blocking probability as an Erlang loss system with *n*=*Z* channels, offered traffic *A*=*Z* (and thus peakedness *Z*=1):

$$E(n, A, Z) \sim E\left(\frac{n}{Z}, \frac{A}{Z}, 1\right) \sim E_{\frac{n}{Z}}\left(\frac{A}{Z}\right).$$

There are several other methods to deal with overflow traffic. Using the above Erlang-Engset-Pascal models (BPP-traffic models) and traffic congestion we get results similar to the above methods. Also Interrupted Poisson processes are used to model bursty traffic processes.

4.8.4.6 Principles of dimensioning

When dimensioning service systems we have to balance grade-of-service requirements against economic restrictions.

In telecommunication systems there are several measures to characterise the service provided. The most extensive measure is Quality-of-Service (QoS), comprising all aspects of a connection as voice quality, delay, loss, reliability etc. We consider a subset of these, Grade-of-Service (GoS) or network performance, which only includes aspects related to the capacity of the network.

For proper operation, a loss system should be dimensioned for a low blocking probability. In practice the number of channels *n* should be chosen so that $En(A)$ is 1-5% to avoid overload due to many non-completed and repeated call attempts which both load the system and are a nuisance to subscribers.

If Erlang's B-formula is applied with a fixed blocking probability for dimensioning trunk groups, then we will observe that

a. The utilisation per channel is, for a given blocking probability, highest in large trunk groups, but very low in small groups At a blocking probability $E = 1$ % a single channel can at most be used 36 seconds per hour! See Fig. 4.8.3.

b. Large trunk groups are more sensitive to a given overload than small trunk groups. This is explained by the low utilisation of small groups, which therefore have a higher spare capacity.

Thus two conflicting factors are of importance when dimensioning trunk groups: we may choose among a high sensitivity to overload or a low utilisation of the channels.

4.8.4.6.1 Improvement principle (Moe's principle)

If we replace the requirement of a fixed blocking probability with an economic requirement, then the improvement function $Fn(A)$ should take a fixed value so that the extension of a trunk group with one additional channel increases the carried traffic by the same amount for all groups.

We will then notice that the utilisation of small groups becomes better corresponding to a high increase of the blocking probability. On the other hand the congestion in large groups decreases to a smaller value.

$$F_B = \frac{c}{g} = \frac{\text{cost per extra channel}}{\text{income per extra channel}} \ .$$

F_B is called the improvement value.

4.8.5 Delay systems

In loss systems users are either served immediately or lost. In delay systems a user finding all servers busy may wait in a queue (buffer) until a server becomes idle.

4.8.5.1 Grade of Service parameters

Also for delay systems we distinguish between time, call, and traffic averages.

The main performance measure is:

- Mean *waiting time W* for all customers
- Mean *waiting time w* for delayed customers
- *Delay variation = delay jitter*

Later we also consider a finite buffer size so that customers may (1) be served immediately, (2) be served after delay, or (3) be blocked with being served.

4.8.5.2 Erlang's delay systems

As for Erlang's loss system the number of channels is *n* and the offered traffic is *A*. Calls which finds all channels busy wait in a buffer (queue) until they are served. The probability that a call attempt is delayed is given by *Erlang's C-formula*:

$$E_{2,n}(A) = \frac{\frac{A^n}{n!}\frac{n}{n-A}}{1+\frac{A}{1}+\frac{A^2}{2!}+\cdots+\frac{A^{n-1}}{(n-1)!}+\frac{A^n}{n!}\frac{n}{n-A}}, \quad A < n\,.$$

This delay probability depends only upon *A*, the product of *λ* and *s*, not upon the parameters *λ* and *s* individually. The formula is also called *Erlang's second formula* The waiting time depends on the mean value *s* of the service time distribution which must be exponentially distributed. Where loss systems in general are insensitive to the service time distribution and only depends on the mean service time (offered traffic), then delay systems are very sensitive to the distribution of the service time.

The mean waiting time for all customers is:

$$W_n = E_{2,n}(A)\cdot\frac{s}{n-A}\,.$$

The mean waiting time for delayed customers is:

$$w_n \;=\; \frac{s}{n-A}\,.$$

4.8.5.3 Palm's delay systems

In models of computer and data networks it is common to have a fixed number of users (jobs, packets). As we for loss systems consider a finite number of users, we may do the same for delay systems. Then we get a model where we have n servers and S users which are *on/off* users. This model is widely used for closed systems with fixed number of users (e.g. packets).

Let the mean idle (*off*) time be $m_t = \gamma^{-1}$ and the mean service time (*on*) be $s = \mu^{-1}$. We introduce $\varrho = m_t/s$.
If we consider a single server system (n =1) and assume exponentially distributed service times, then the mean waiting for all customers becomes:

$$W_{1,S} = \frac{S}{1 - E_{1,n}(\varrho)} - 1 - \varrho \quad [\text{mean service times}],$$

where $E_{1,n}(\varrho)$ is the *Erlang-B formula*. With a finite number of sources the mean waiting time becomes less than for the *Erlang-C* case with Poisson arrival process. The mean waiting time is independent of (insensitive to) the idle time distribution. Only the mean value of the idle time is of importance.

The service time has to be exponential, but later we introduce processor sharing and then it also becomes insensitive to the service time distribution.

It is easy to generalize the model to n servers [4.1].

4.8.5.4 Processor sharing strategies

By Processor Sharing (*PS*) all users equally share the available capacity. No users are waiting, but all get some service at reduced rate, depending on the number of users. Delay systems are in general very sensitive to the service time distribution, but if we introduce processor sharing, then the systems become insensitive to the service time distribution. In comparison with the service time realized if the user obtained the required capacity, the service time (sojourn time) is increased, and the increase corresponds to a virtual waiting time. Applying processor sharing strategy to the single server queue with general service time distribution (*M/G/1*) we get the same mean delay as for exponential service times (*M/M/1*) which are easy to deal with.

If we apply processor sharing to a queueing system with n servers, the queueing system (*M/G/n*) experience the same mean waiting time as Erlang's waiting time system (*M/M/n*) as the system becomes insensitive to the service time distribution. A user never requires more capacity than one channel, even if more channels are idle. We may consider one channel as the access capacity of a user and n as the total capacity of the system. A modified model includes multi-rate traffic so that if possible a multi-rate user obtains more channels, but during overload multi-rate calls are first restricted, and everybody gets the same capacity. This is called Generalized Processor Sharing (*GPS*).

In a similar way processor sharing applied to Palm's waiting time system with *n* servers will make this insensitive to the service time distribution.
Thus the system becomes insensitive to both the *off* and the *on* time distribution. In this way we get a very robust model appropriate for modeling real-life systems.

So far we only considered single-slot (single-rate) traffic. In the following we generalize this to multi-rate traffic by generalizing processor sharing to *reversible scheduling* and obtain new models applicable for evaluating future generation networks.

4.8.6 Multi-rate (multi-service) loss systems

In classical traffic models we only consider one service (voice) and all connections use one channel on each link. In service-integrated systems with N services, service i has individual bandwidth requirement d_i ($i = 1; 2; : : : ;N$).

There are two classes of exact algorithms to deal with these systems:

- convolution algorithms based on aggregation of services, and
- state-based algorithms based on aggregation of states.

4.8.6.1 Convolution algorithm

The convolution is described in details in [4.1]. It is based on the product-form property. Let the state probability of the system be described by number of channels x_i occupied by service i. Then the product form implies:

$$\begin{aligned} p(\overline{X}) &= p(x_1, x_2, \ldots, x_N) \\ &= p(x_1) \cdot p(x_2) \cdot \ldots \cdot p(x_N)\,, \end{aligned}$$

By convolution we aggregate the services and end up with two services.
One is the aggregation of all services except service i which we want to calculate the performance of. The number of states are thus reduced to a two-dimensional state transition diagram independent of the total number of services.
For example we aggregate services 1 and 2 $p(x_{12}) = p(x_1) * p(x_2)$ as follows:

$$p(x_{12} = j) = \sum_{i=0}^{j} p(x_1 = i)\, p(x_2 = j - i)\,.$$

For each service we may both guarantee a minimum number of reserved channels and restrict the number of connections by an upper limit.

The algorithm is applicable for calculating end-to-end blocking in circuit-switched multi-rate networks with Binomial-Poisson-Pascal traffic with minimum guaranteed and maximum allowed number of connections end-to-end.

More details are given in chapter 10 in [4.1].

4.8.6.2 State space based algorithms

Another approach is to aggregate the state space into global state probabilities.

4.8.6.2.1 Fortet & Grandjean (Kaufman & Robert) algorithm

We still consider multi-rate traffic streams. In case of Poisson arrival processes the algorithm becomes very simple.
Let *pi*(*x*) denote the contribution of stream *i* to the global state probability *p*(*x*):

$$p(x) = \sum_{i=1}^{N} p_i(x) \, .$$

Thus the average number of channels occupied by stream *i* when the system is in global state *x* is $x \, . \, p_i(x)$.
Let traffic stream *i* have the slot-size d_i. Due to reversibility we will have local balance for every traffic type.
The local balance equation for state *x* becomes:

$$\frac{x\, p_i(x)}{d_i}\, \mu_i = \lambda_i \cdot p(x - d_i) \, , \qquad x = d_i, d_i + 1, \ldots n \, .$$

The left-hand side is flow from state [*x*] to state [*x* - *di*] due to departure of type *i* calls. The right-hand side is the flow from global state [*x* - *di*] to state [*x*] due to arrivals of type *i*.
It does not matter whether *x* is a integer multiple of *di*, as we only consider average values.

From the equation above we get:

$$p_i(x) = \frac{1}{x}\, d_i\, A_i \cdot p(x - d_i) \, .$$

The total state probability *p*(*x*) is obtained by summing over all traffic streams :

$$p(x) = \frac{1}{x} \sum_{i=1}^{N} d_i\, A_i\, p(x - d_i) \, , \quad p(x) = 0 \quad \text{for} \quad x < 0 \, .$$

This is *Fortet & Grandjean's algorithm*. The algorithm is usually called *Kaufman & Roberts' algorithm*, as it was re-discovered by these authors in 1981.

More details are given in [4.1].

4.8.7 Multi-rate traffic and reversible scheduling

A completely new revision of the classical teletraffic theory is being published in [4.3]. It simplifies and generalizes all classical theory for circuit switching and packet switching networks. It includes multi-rate traffic. For buffer size zero it is equivalent to multi-rate loss systems, which are insensitive to service time distribution. For systems with infinite buffer and single-rate traffic it corresponds to generalized processor sharing, and for multi-rate delay systems the users share the capacity so that broadband calls are allocated more resources than narrow-band traffic. Broadband calls are reduced more than narrow-band calls and when the overload increases in the limit every user get the same capacity

We consider a system with *N BPP* traffic streams, *n* servers, *k* buffers.

The offered traffic is *BPP* multi-rate traffic. The basic bandwidth unit is equivalent to one channel.
A generalized recursion formula for state-probabilities is derived by Iversen :

$$p(x) = \begin{cases} 0 & x < 0 \\ 1 & x = 0 \\ \sum_{i=j}^{N} p_j(x) & x = 1, 2, \ldots, k+n \end{cases}$$

where

$$p_j(x) = \max\left\{\frac{x}{n},\, 1\right\} \cdot \left\{\frac{d_j}{x} \cdot S_j\,\beta_j \cdot p(x-d_j) - \frac{x-d_j}{x} \cdot \beta_j \cdot p_j(x-d_j)\right\}$$

or by replacing the parameters (*Sj ; βj*) by (*Aj ;Zj*) :

$$p_j(x) = \max\left\{\frac{x}{n},\, 1\right\} \cdot \left\{\frac{d_j}{x} \cdot \frac{A_j}{Z_j} \cdot p(x-d_j) - \frac{x-d_j}{x} \cdot \frac{1-Z_j}{Z_j} \cdot p_j(x-d_j)\right\}$$

The initialization values of $p_j(x)$ are $\{p_j(x) = 0;\ x < d_j\}$. This is a simple general recursion formula covering all classical models.
As a special case we get Erlang's loss system, and the recursion formula for evaluating this.
The approach is mathematically very simple and general, and it allows for simple numerical evaluation. The properties of the algorithm is analysed in section 4.8.7.2.

Remark: For infinite number of buffers we require *A* < *n* to attain statistical equilibrium. For Pascal arrival processes there are more strict requirements. For a system with buffers the Pascal case may result in a carried traffic which is bigger than the offered traffic, because the arrival rate increases linearly with the number of customer being served or waiting.

4.8.7.1 Performance measures

We consider a system with *n* channels and *k* buffers, both given in basic bandwidth units. The model includes loss systems (blocked calls cleared and blocked calls held), classical delay systems, processor sharing systems etc.
The performance measures become rather diversified, and we only derive the basic performance measures.

4.8.7.1.1 Time average performance measures

The time congestion *Ebj* of stream *j* is defined as the proportion of time new connections are blocked:

$$Eb_j = \sum_{x=n+k-d_j+1}^{n+k} p(x)\,, \quad j = 1, 2, \ldots, N\,.$$

In a similar way the proportion of time new calls are delayed becomes:

$$Ed_j = \sum_{x=n-d_j+1}^{n+k-d_j} p(x)\,, \quad j = 1, 2, \ldots, N\,.$$

The proportion of time new calls get full service at the time of arrival becomes:

$$Es_j = \sum_{x=0}^{n-d_j} p(x)\,, \quad j = 1, 2, \ldots, N\,.$$

Of course, we have *Ebj* + *Edj* + *Esj* = 1. When the system operates as a classical single-slot delay or loss system these measures are simple to understand. However, we should remember that in case of processor sharing systems, calls arriving later may influence the service of existing calls. The above performance measures are time averages. The more useful call averages are derived below.

4.8.7.1.2 Traffic average performance measures

It should be noticed that these mean values are much more important than time and call average values.

The carried traffic for stream *j* measured in channels is given by:

$$Y_j = \sum_{x=0}^{n} x \cdot p_j(x) + n \cdot \sum_{x=n+1}^{n+k} p_j(x)$$

As the offered traffic of type *i* measured in channels is *dj .Aj* , the lost traffic is (*dj .Aj -Yj*).

The traffic congestion *Cj* for stream *j*, which is the proportion of offered traffic blocked, becomes:

$$C_j = \frac{d_j \cdot A_j - Y_j}{d_j \cdot A_j}, \quad j = 1, 2, \ldots, N\,.$$

The total traffic congestion is:

$$C = \frac{A - Y}{A}$$

where

$$A = \sum_{j=1}^{N} d_j \cdot A_j \quad \text{and} \quad Y = \sum_{j=1}^{N} Y_j = \sum_{x=0}^{n} x \cdot p(x) + \sum_{x=n+1}^{n+k} n \cdot p(x)\,.$$

4.8.7.1.3 Call average mean values

For systems with processor sharing the probabilities that a random call attempt is served without delay, delayed and served, or blocked are not well defined as calls may get the required capacity at the start of service, but be delayed by later arrivals.

For classical queueing systems, where a call keeps the full capacity from start of service, we find the following probabilities (call averages).

For Engset and Pascal traffic the average number of idle sources type *j* is *Sj -(Yj +Lj)/dj* . So the average number of call attempts of stream *i* per time unit is (*Sj* - (*Yj* + *Lj*)/*dj*) γ_j , where γ_j is the call intensity of an idle source type *j*.

The probability that a random call attempt of type *j* get full service at the time of arrival becomes:

$$X_j = \frac{\displaystyle\sum_{x=0}^{n-d_j} \left\{ S_j \cdot p(x) - \frac{x}{d_j} \cdot p_j(x) \right\}}{S_j - \dfrac{Y_j + L_j}{d_j}}, \quad j = 1, 2, \ldots, N \,.$$

The probability that a random call attempt of type *j* is delayed at the time of arrival becomes:

$$D_j = \frac{\displaystyle\sum_{x=n-d_j+1}^{n+k-d_j} \left\{ S_j \cdot p(x) - \frac{x}{d_j} \cdot p_j(x) \right\}}{S_j - \dfrac{Y_j + L_j}{d_j}}, \quad j = 1, 2, \ldots, N \,.$$

The probability that a random call attempt of type *j* is blocked becomes using $\beta j = \gamma_j / \mu j$:

$$B_j = \frac{\displaystyle\sum_{x=n+k-d_j+1}^{n+k} \left\{ S_j \cdot p(x) - \frac{x}{d_j} \cdot p_j(x) \right\}}{S_j - \dfrac{Y_j + L_j}{d_j}}, \quad j = 1, 2, \ldots, N \,.$$

For a random call attempt we of course have $Bj + Dj + Xj = 1$.

For Poisson arrival processes the above probabilities are obtained directly by summation of the proper global state probabilities because of the *Pasta*-property, and we get the same results as in Sec. 4.8.7.1.1.

4.8.7.1.4 Mean waiting times and queue lengths

The mean queue length of stream *j* (traffic of stream *j* carried by the queueing positions) measured in unit of channels becomes:

$$L_j = \sum_{x=n+1}^{n+k} (x - n)\, p_j(x)\,, \quad j = 1, 2, \ldots, N \,.$$

As the same calls are waiting (including no waiting time) and served, the mean waiting time for all accepted customers of type *j* becomes:

$$W_j = s_j \cdot \frac{L_j}{Y_j}\,, \quad j = 1, 2, \ldots, N \,,$$

where s_j is the man service time of type *i* calls, and both Lj and Yj are measured in channels.

The total mean queue length measured in channels is:

$$L = \sum_{x=n+1}^{n+k} (x-n) \cdot p(x) = \sum_{j=1}^{N} L_j \, .$$

The overall mean waiting time for all accepted customers is given by:

$$W = s \cdot \frac{L}{Y} \, .$$

The mean service time *s* for all accepted customers is only obtainable by proper weighting of accepted calls.

The mean waiting time of delayed calls type *j* excluding blocked calls is obtained from the formulae for *Wj* above:

$$w_j = \frac{W_j}{D_j} \cdot (X_j + D_j) \, .$$

We notice that mean waiting times are measured in mean service times. Let us for transfer of a fixed amount of data (bytes) using bandwidth $d = 1$ denote the mean service time by *s*. Then the mean service time when using a bandwidth *dj* will be *s/dj* , and also the mean waiting time will be reduced. By choosing a bigger bandwidth we may give priority to a traffic stream (reduce transfer time) and/or increase the amount of data transferred (goodput).
Only when the system is heavily overloaded all connections will be allocated the same capacity (generalized processor sharing).

4.8.7.1 Properties of the algorithm

The above theory results in a simple algorithm with the following basic features:

1. Initialization of variables

2. Let $x := x + 1$

3. Calculate $pi(x)$ and $p(x)$ using formulae described in the beginning of subchapter 4.8.7

4. Normalize all states by dividing by $(1 + p(x))$

5. Go to step 2 if $x < n + k$

6. Calculate performance measures

If we know the normalized state probabilities *pj*(*x*-1), then we can calculate *pj*(*x*) for *x*. As we know *pj*(*x*) = 0 for *x* < *dj* we thus are able to calculate the state probabilities by recursion. The implementation is described elsewhere.
To find the performance measures we only need to know the *dj* previous states of traffic stream *j*, as we may accumulate the necessary information on carried traffic and other statistics in a few variables.

Thus memory requirements of the algorithm is of the order of size:

$$m_m = O\left\{\sum_{j=1}^{N} d_j\right\} .$$

The number of operations is of the order of size:

$$m_c = O\{(n+k) \cdot (m_m + N)\}$$

as we for a given number of channels need to calculate *N* new terms from max*{dj}* previous global states and normalize m_m terms.

Thus the algorithm requires very little memory and is linear in the number of channels and number of services. The accuracy is optimal as we always operate with normalized values and always normalize (divide) with constants greater than one.

It may be mentioned that the famous recursion formula for Erlang-B formula is a special case with one single-slot Poisson traffic stream.

Also the recursion formula for Engset is a special case as well as Delbrouck's formula.

Chapter 5 – Economical modelling and business plans

In this chapter an overview on the business plan objectives is given with the main activities and results to be used for the technical planning. Also an overview is summarised for the economic modelling needed to evaluate different alternatives and to model telecom equipment for the purpose of optimization

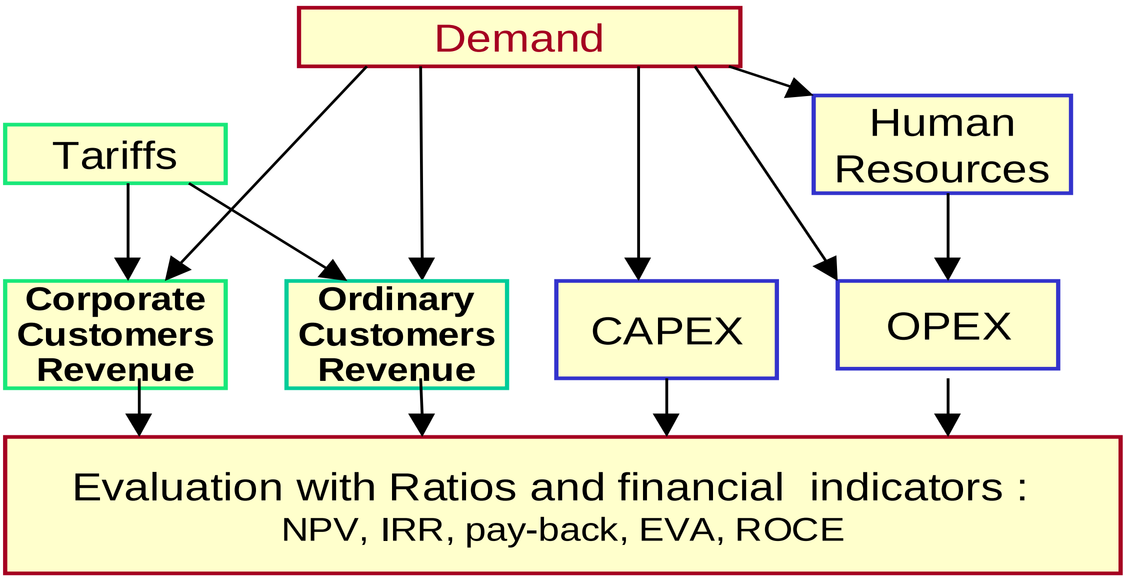
5.1. Business planning

A Business Plan presents the calculation of the financial indicators that enable the managers to evaluate the financial performances of an enterprise in order to take best decisions for the overall operation. Due to the high number of alternatives today and the need to find economical feasibility in competition, the business evaluations are being used not only for the business plan itself but as an iterative evaluation of those techno-economical alternatives to select the ones that perform better in the competitive market.

A Business Plan summarises the results of the planning process:
- the objectives to reach (subscribers demand, sales)
- the future revenues expected from the plan and per service class;
- the planned expenses (investment and operations) as overall and per service class;
- the accounting statements and the financial indicators characterising the profitability of the project.

The framework structure for the evaluation follows the model of the figure. Each box is expanded with more degree of detail as a function of the plan time frame with the corresponding des-aggregation.

Fig 5.1: Business model structure for planning



5.2. Economic modelling for planning

- Economical modelling to evaluate network solutions (modelling tariffs, equipment costs per type, economy of scale, lifecycle, equipment deployment, elasticity, trends with time, etc.)

5.3. Economic concepts and terms

- The following terms and associated concepts are the most frequently used to analyze the Telecom business and decide on best project alternatives with its specific properties due to the multiplicity of diverse equipments, life cycles and operational practices. For a more detailed economic terms definition refer to classical books of Economy.

Amortisation

Amortisation refers to the paying off of a debt with regular payments and it also has the meaning of the accounting procedure that gradually reduces the cost value of an intangible asset, that is, depreciation.

Amortization is the method of liquidating a debt on an installment basis; for example an amortized loan would be one where the principal amount of the loan would be paid back in installment over the life of the loan. Sometimes used as an alternative term for depreciation, in particular with regard to the process of writing off the cost of an intangible asset, such as a lease or patent, over its useful life.

Assets

Resources owned by an enterprise. In the balance sheet assets are listed in rising order of liquidity. They include fixed assets (land and buildings, plant and machinery, etc), current assets (inventories, account payable, etc.) and liquid assets (cash in hand, cash in banks, cheques, etc).

Breakeven period

Time required for project revenues (after deduction of operating expenses) to offset investment expenditure. This method of comparing project avoids the need for discounting calculations. It takes account, however, neither of the effects of the time factor of the different alternatives, nor of what happens after breakeven.

Capacity

Capability of an equipment, network or sub network to handle the traffic flows with an associated grade of service.

Maximum capacity is the limit that may be reached over a short period of time without overload protection

Nominal capacity is the value of the proper engineering to account for the natural statistical variations over sustained periods and the overload protection to guarantee the traffic handling in a sustained manner

Operational occupancy is the actual value of the resource occupancy referred to the nominal capacity at any point of time taking into account the reserved capacity for extensions and the partial occupancy due to the demand grow, the natural equipment modularity and the needed time lag between consecutive installations.

Capex

Capital Expenditures due to the purchase of a fixed asset to be installed at the different network segments and layers:

- Typically: land, building, exchange, cabinet, duct, fiber, cable, transmission system, tower, BTS, computer, IT platform, car, etc.
- Costing more than a threshold defined internally in any company and following current financial best practices in order to allow consideration as an asset and not as a consumable or operation expense.
- Having an expected life of more than one year (value subject to specific parameters of industry sectors and companies)

Cash flow

Cash receipts and cash disbursements over a given period.
Also funds generated internally by the activity of an enterprise or a project equivalent to the balance between the inflow of funds arising from revenues and the outflow of funds arising from expenditures.

The following diagram illustrates the main generators for the inflows to the company classified in three categories: The first one at the left is the specific business operating income due to the selling of services to customers and the most interesting to analyze when comparing projects or evaluating strategies for the operator evolution. The other two consider the generic financing capital increases either due to the shareholders by a capital increase or to the external sources of capital by credits or loans.

Typical originators for the outflows in a company are also summarized in the diagram with the first three concepts due to the proper activities of the Telecom activity itself like laborforce, network equipment investment and all technical, operation and administrative expenses. The other three concepts reflect the generic outflows due to taxes, debt payment and dividends for the shareholders that are needed to have an overall company running.

A detailed analysis of the specific Telecom associated inflows and outflows is the nucleus of the operational business analysis when a decision has to be taken in a modernization of the network, migration to NGN, introduction of new services, etc. Yearly cash flows are taken as the main base for the evaluation of a company value, capability to generate business and calculation of the Net Present Value (NPV) when transforming into present values and decide which evolution alternative is recommended.

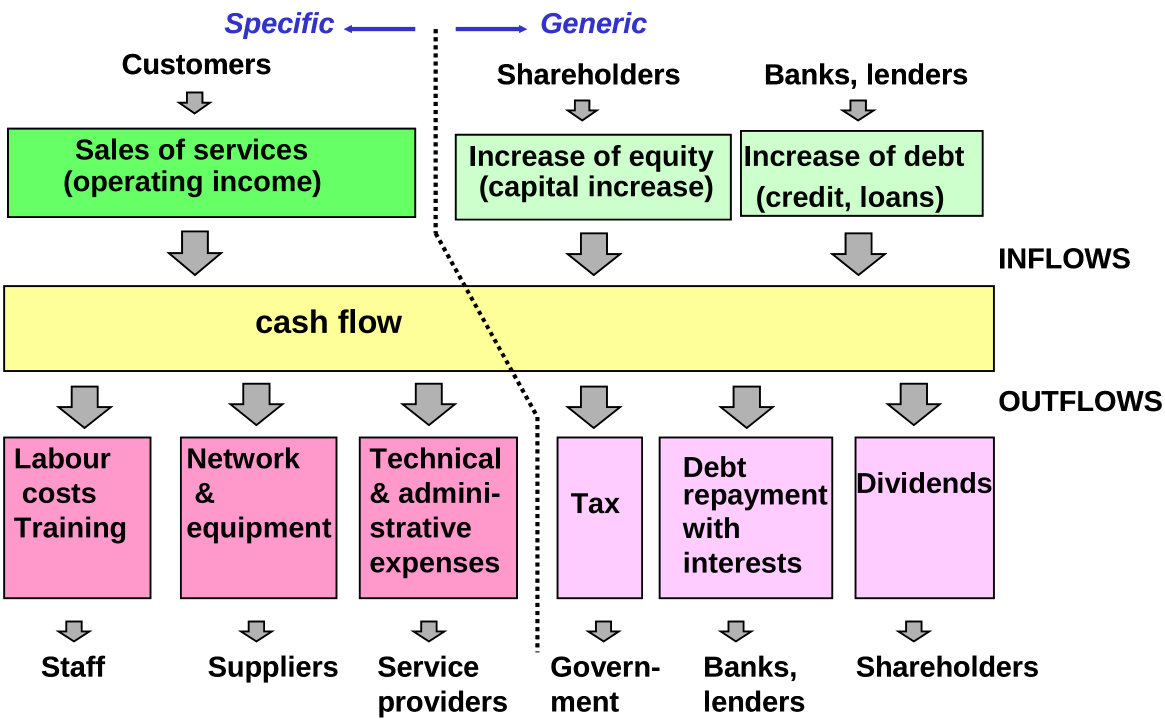


Fig:5.3.1 Main Inflows and Outflows contributing to the cash flow generation

A typical set of components for an evaluation of a given project over a period of time is illustrated in the diagram below which considers inflows and outflows at the year of generation without time distribution effects due to financing, amortization, depreciation, etc.

Main sources of inflows are due to the revenues of the different operation services and at the end of evaluation period also the terminal value of those network elements that did not reached the end of life cycle have to be taken into account as they have a positive value.

As main components of the outflows we have the major capex equipment investment at the project start with the corresponding equipment extensions or upgrades for capacity increase in subsequent years as well as equipment substitution when some of the elements reached their end of life cycle. Opex increases as a function of the cumulative invested Capex through time and is the main outflow component at the medium long term.

Net cash flow is derived from the difference of the inflows and outflows and provides the main input for a more detailed dynamic evaluation of the project added value to a company. Higher cash flows at the end of the evaluation period and a prompt turn into positive values are good indicators for a better project.

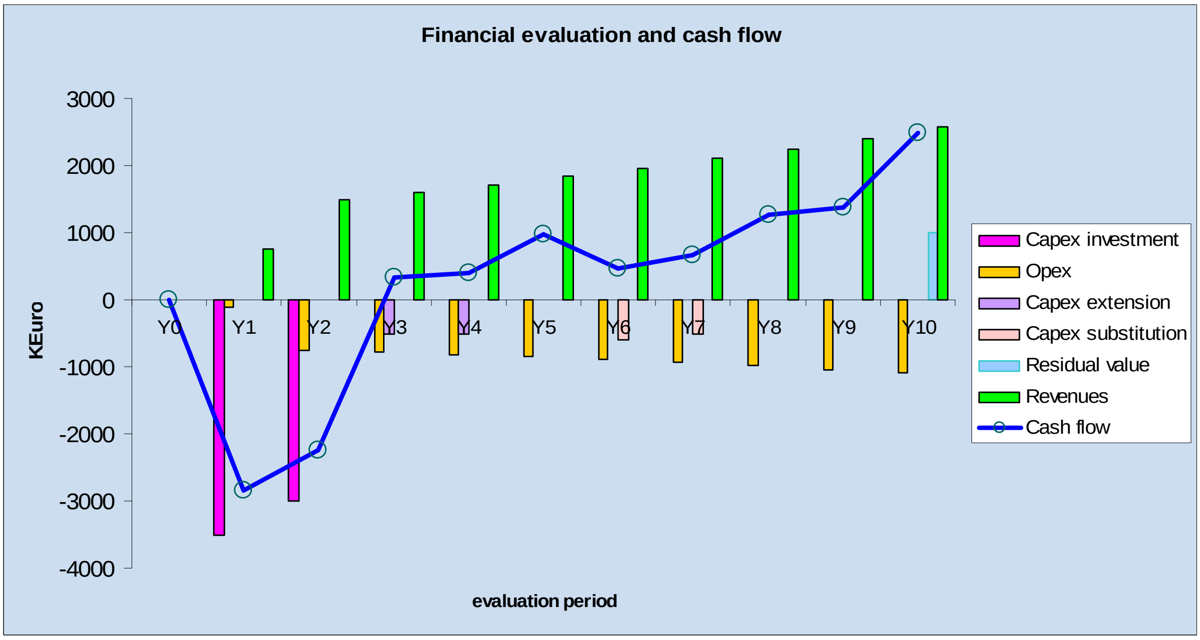


Fig:5.3.2 Typical components as a basis for evaluation of a company or project through the years.

Churn

Annual rate at which the own customers or subscribers leave the service either to move to a competitor, to migrate to other service or leave the market.

Depreciation

Loss of value of an asset over time, as a result of wear, aging or obsolescence. With the method of linear (or straight-line) depreciation, the loss of value of an asset is spread uniformly over the number of years of its useful life. Depreciation charges do not give rise to an actual outflow of funds and the sums remain available to the enterprise.

Diagram below illustrates the residual value of an asset through the years as a function of the depreciation law, either linear as the most common for network elements, accelerated (for the elements with very rapid evolution) or delayed (for the stable and robust network elements)

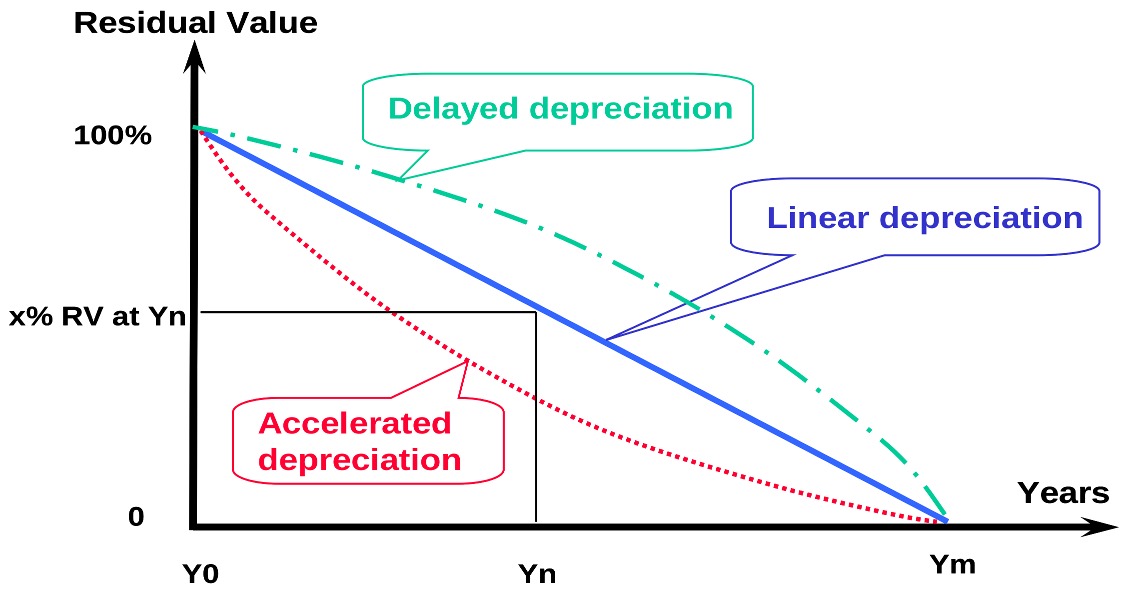


Fig:5.3.3 Residual value as a function of depreciacion law

Discounted Cash Flows

An investment appraisal technique which takes into account both the time value of money (i.e. the conversion of cash flows that occur over time to an equivalent amount at a particular point in time) and the total profitability over a project' life.

Discount factor

The discount rate used to calculate the net present value of a company or project. This rate has to consider the cost factors for the capital of the company such as interest rate, and expected inflation in a strict sense. In a wider sense has to consider also the risk rate for long term evaluation in large projects and new scenarios with uncertainty.

EBITDA

Annual Earnings Before Interest, Tax, Depreciation and Amortization. This means all the revenues minus operating costs that is the basic information for the evaluation of a business from its own specific factors and the first indicator to be calculated and analysed. The following diagram illustrates a simplified interrelation among main generators for the EBITDA and the sequence to proceed in the obtention of the Net Income

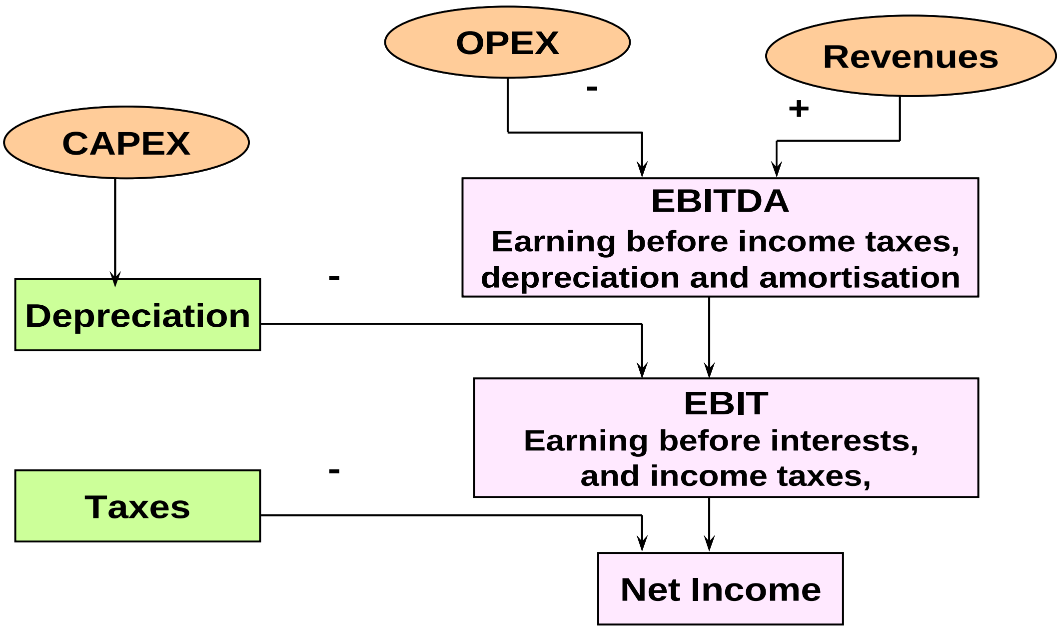


Fig:5.3.4 Relation between EBITDA and main business concepts

EVA

Economical Value Added or net operating profit (after tax) minus the cost of the capital used to generate that profit either in debt or in equity. It is a good indicator for the point of view of the investors

Future Value (FV)

The value of a present amount at a future date. It is found by applying compound interest over a specified period of time.

$$FV = PV * (1 + k)^n$$

<u>Internal rate of return (IRR)</u>

This is the discount rate that equates the present value of investment outflows with the present value of inflows produced by the investment. The internal rate of return may be considered as the highest rate of interest admissible for a project wholly financed by borrowing.
IRR is the discount rate that equates the present value of cash inflows with the initial investment associated with a project, thereby causing NPV = 0.

$$IRR = 0 = \sum [CF/(1 + IRR)^t] – \text{Initial Investment}$$

<u>Life cycle costing</u>

The full cost of an asset over its life. This includes all costs associated with acquiring, controlling, operating and disposing of the asset.

<u>Net Present Value (NPV)</u>

A global capital budgeting technique; found by subtracting a project’s initial investment from the present value of its cash inflows discounted at a rate equal to the firm’s cost of capital.

$$NPV = \text{present value of cash inflows - initial investment}$$

$$NPV = \sum [CF/(1 + k)^t] – \text{Initial Investment}$$

According to the consideration of the final value of the network at the end of the evaluation period, basically two procedures are the most frequent in the analysis:

- <u>NPV with zero terminal value</u> that ignores the terminal value of the network investments as a function of equipment life cycles, amortization status and future values of cash flows.
- <u>NPV at perpetuity rate</u>: that considers terminal value estimated as a projected cumulative discounted cash flow for the remaining years based on future perpetuity and discount rates.

The selection of one of these values or other intermediate ones is a function of the size of the evaluation period, strategy of the operator, importance of the investments performed during the period and expectations of revenues in the market.

Calculating the present value of all cash flows during different years allows valid comparisons and measurement to be made between them.

The decision criterion when using the net present value approach to make accept-reject decisions is as follows: if NPV > 0, accept the project; otherwise reject the project. When comparing different projects, those with higher value at the evaluation period will be selected as the ones with more potential to generate business and to ensure being profitable in a competitive market.

Following diagram illustrates the typical NPV curves over the evaluation period as a function of the discount rate that is applicable in a given country and financial context.

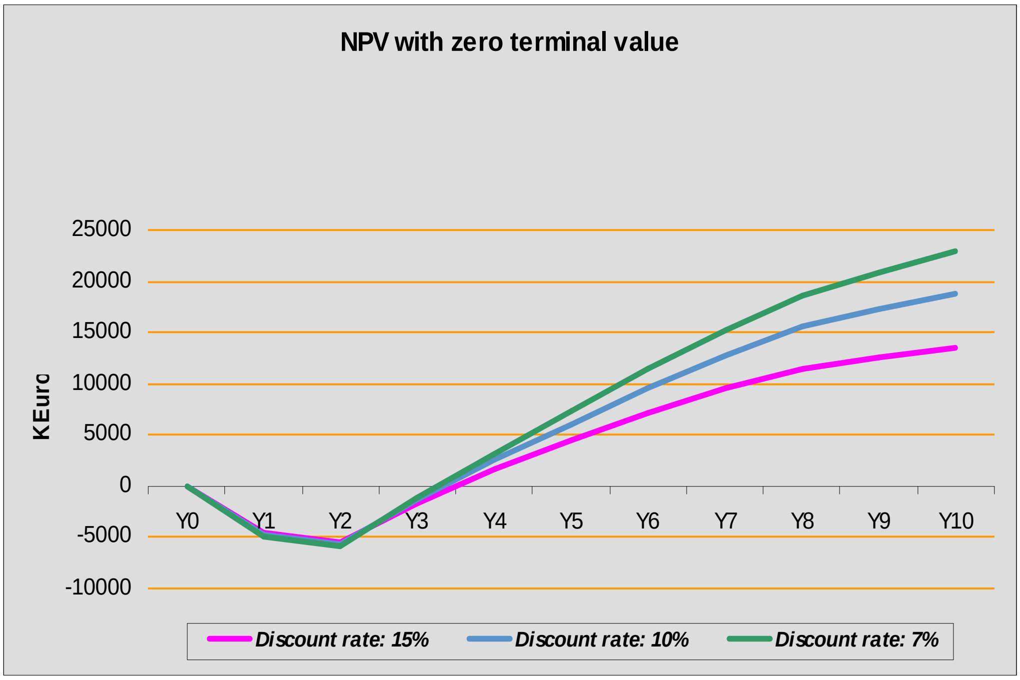


Fig:5.3.5 Typical NPV evolution for a new project as a function of the discount rate

Diagram below illustrates the typical NPV curves for two frequent deployment strategies in a given network. Conservative one invests at a low speed over the geographical area and requires less capital but decreases the future business potential. This is typical for a very short term view. The ambitious strategy invests at a higher speed and requires higher initial capital with the consequence of sooner increase of revenues and higher business potential. This is more proper for a medium-long term view of the operator

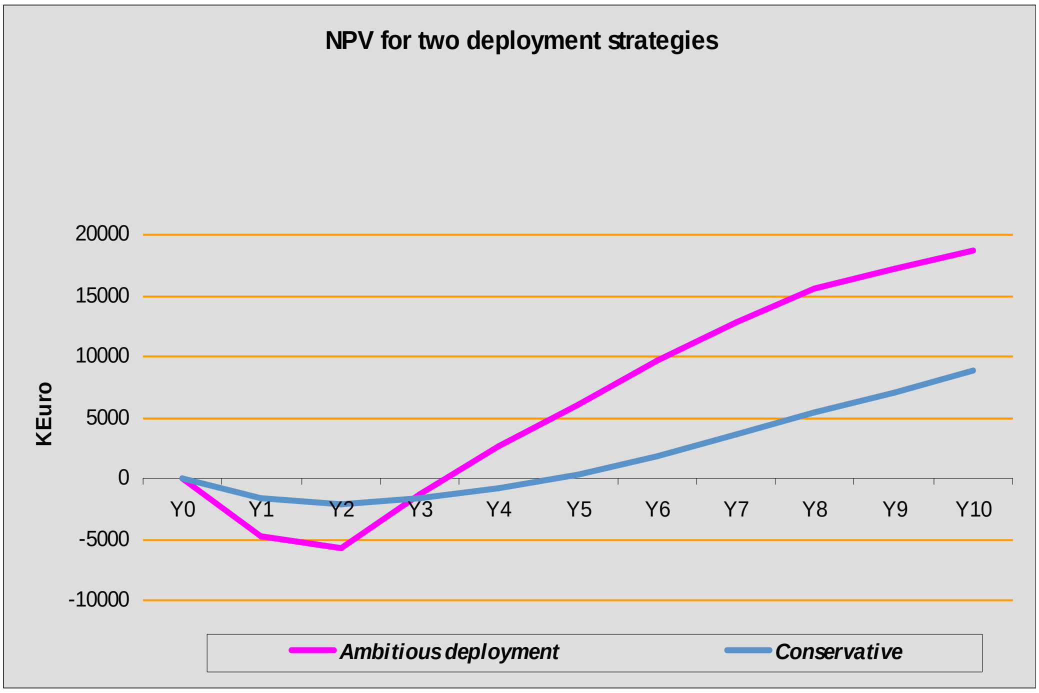


Fig:5.3.6 Typical NPV evolution for two network deployment strategies

Opex

Operational Expenditure or operations costs: All non-capitalised costs of operating the network either associated to each network element, to running the services or generic company activities.

Typical operation cost associated to the network elements are:
- Maintenance
- Connection
- Rental
- Technical operation
- Decommissioning

Services associated operations include:
- Service activation
- Commercial operation
- Service marketing campaigns
- Balance of international traffic (if negative)
- Compensation to content providers, etc.

Generic operation costs consider:
- Labor costs
- Social charges

- Training
- Company marketing
- Administrative expenses
- Bad debt

Payback period

The number of years required for a firm to recover the initial investment required by a project from the cash inflows it generates. Short payback periods are preferred.

Like internal rate of return, the payback period metric takes essentially an "Investment" view of the action, plan, or scenario, and its estimated cash flow stream. Payback period is the length of time required to recover the cost of an investment (e.g. purchase of computer software or hardware), usually measured in years. Other things being equal, the better investment is the one with the shorter payback period.

Also, payback periods are sometimes used as a way of comparing alternative investments with respect to risk: other things being equal, the investment with the shorter payback period is considered less risky.

Present Value (PV)

The current monetary value of a future amount. The amount of money that would have to be invested today at a given interest rate over a specified period to equal the future amount.

$$PV = FV / (1 + k)^n$$

PV is the currency value today of some future inflow, outflow, or balance of funds. In essence, it is the discounting of future funds to their present value by taking into account the time value of money. It is useful in providing a common basis for comparing investment alternatives. See also discounted cash flow, future value, and net present value.

Profit

Another term for Net Income or Earnings.
Surplus of sales revenues over costs or expenditure during an accounting period or operating cycle. Leads to an increase in owners' equity, though not necessarily to an increase in cash. It may be reflected in increased assets or decreased liabilities. Net profit may refer to profits after tax (on profits) or to profits less financial costs, depending on the purpose of the analysis.

Residual value

Value of an investment at the end of its economic or estimated life. At the end of the period, residual value may be treated as a positive cash flow, and discounted as such. The present value of the business attributable to the period beyond the forecast period.

Revenues

An income statement term, referring to the sum of money owed the company for sales of goods and services. Revenues (or “Sales”) are ordinarily the top line in the income statement, against which most other costs and expenses are subtracted to calculate income. In Britain, the term turnover is often used in place of revenues.
The term revenues generally mean "gross revenues,” that is, revenues before adjustments for customer discounts and allowances.
(see Profit)

ROCE

Return on Capital Employed or net income divided by the sum of fixed assets and working capital. Shows the company profitability from the point of view of the owners.

5.4. Economic modelling for services

Due to the high number of new services and large variety of marketing characteristics, one of the most interesting planning activities when operating NGN is the economic modelling of the services costs, revenues and profitability, either per single service or for groups of services in the case of a bundle offer.

The service related modelling is a subset of the overall NGN economic modelling that represents with less detail the analysis of technological variants and for a given network solution models with more detail the demand, dimensioning, tariffs, revenues and profitability of the services themselves.

In order to consider correctly the per service impacts on the network and business, it should be differentiated in all economic evaluations those values due to:

- the overall operation company
- the global network solution
- the broadband platform needed for all new services
- the specific platform for each service type

Modelling and differentiation of resources, cost drivers, revenue drivers with the corresponding cost allocation per service type is a must for the knowledge of impact from each service and to decide the introduction strategy per service or service bundle.

- Major cost drivers for multimedia service require the evaluation of all dimensioning units like:

- Number of Users

- Number of BB ports
- CAPS (Call Attempts Per Second) and Sessions rate
- Equivalent Sustained Bit Rate at the different network segments
- Required storage memory

- Those units, dimensioning and costing has to be performed for the following network resources:

- Specific service platforms to be dimensioned per service with the corresponding cost evaluation are needed for the applications to be implemented such as: VoIP, VoD, IP Centrex, Unified messaging, Content Delivery, Multimedia messaging, etc.
- Common platforms and network resources to all services requiring the broadband platform should be dimensioned and cost with the corresponding aggregated flows of all services grouped by affinity of QoS or Service Level Agreement
- Operational expenses also need to be modeled with differentiation per service type and service group, both for the technical operation, maintenance, software upgrades as well as for the marketing, promotion, training, etc. that will be important at the first years of the new services.
- For all the common costs either due to BB platforms as well as for common network resources and overall company operation, a sharing factor has to be evaluated also as a function of partial consumed resources that will aggregate to the specific costs

- Major revenue drivers for multimedia services need to consider the contract fee, monthly fee and usage/traffic dependent fee, either in erlangs, Mbs, time, delivery unit (ie: video film) etc. Due to confluence of many new multimedia services of heterogeneous types, it is fundamental to keep track of consumed resources per service type in order to be able to perform backward cost allocation as a function of utilization.

This will allow a latter calculation of proper tariffs to obtain the service profitability either for single services or services bundles and to ensure fulfillment regulation principles when activity based costing is used. For those services provided in a shared mode by multiple players, like video content, gaming, etc. sharing factors among players have to be applied for the evaluation of global revenues and partial revenues to be incorporated to the operator income.

Decision making of which service is introduced first, which grouping or services and for what customer types will be function of the profitability of those alternatives that have a high sensitivity to the services mix by the economy of scale factors. That high impact on the economy of scale is the main driver for convergence of services and the interest of “triple play” and ”multiple play” strategies.

5.5. Cycle life amortization versus modernization

The telecom equipment and infrastructure have to be installed at periodical intervals as a function of the demand evolution. That intervals or provisioning periods for a given technology are a function of the demand growing rate, the systems capacity and modularity, the life cycle of the installed equipment and the corresponding associated costs. The diagram below illustrates three provisioning scenarios for a given demand over time:

- In scenario A, the provisioning is performed at short regular periods (say yearly or quarterly) minimizing the spare capacity but increasing the installation costs by the high number of intervention; in addition an unexpected demand grow at higher rate will imply under provisioning and lost of quality of service as well as of customers.

- In scenario B, the provisioning is performed for the expected demand over a long period or "once for all" with a high start-up cost in CAPEX and high maintenance for an installed equipment with low utilization rate

- In scenario C, the number and volume of provisioning is optimized to minimize Cost of Ownership that considers both CAPEX and OPEX while maintaining adequate utilization rates and Quality of Service

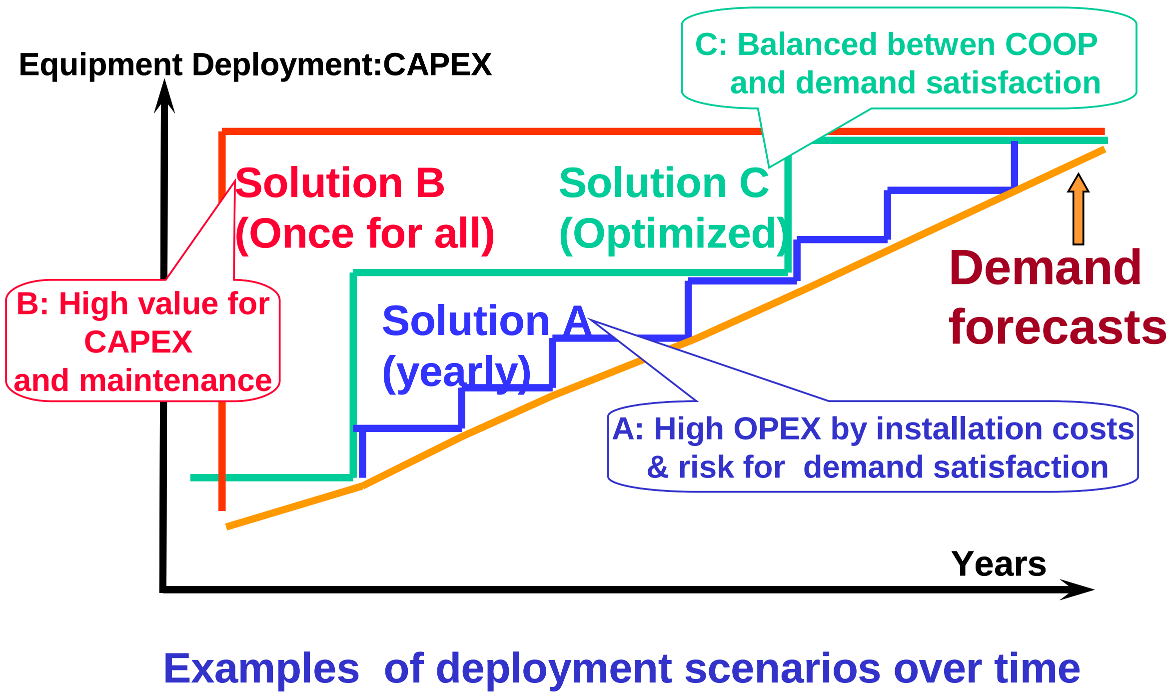


Fig: 5.5.1 Strategies of network resources deployment according to life cycle and economics

The dynamic optimization problem in network planning takes care for the optimum provisioning periods and volumes in order to minimize costs and is a function of the following parameters:

- Demand growing rate and forecast reliability
- Equipment capacity and modularity
- Technical life cycle and capability to provide new services
- Economical lifecycle
- Equipment fixed costs and incremental costs
- Operational and maintenance costs
- Labour costs
- Interest rate and inflation rate

In the case of availability of new technologies of the same functionality with larger capacities or when new functionalities appear like in the network migration towards NGN, the decisions to be taken by the planner introduce additional scenarios:

- Increase of capacity with same technology modularity (ie: more STM-1 systems) versus jumping to the next technology modularity (i:e: substituting STM-1 by STM-4 or STM-16)
- Substitution of existing technology functionality by the next generation functionality with larger capacities and new functions like in the NGN case.

In these types of scenarios, in addition to the previous parameters, the possibility to provide same services with additional sub networks and the possibility to provide new services not feasible with installed technologies, introduce additional complexity in the techno-economical evaluation by the need to incorporate in the evaluation the differential revenues by those new services.

The evaluation process has to consider and compare the *Net Present Value (NPV)* of potential alternatives with all costs and revenues per alternative with the discounted cash flows. The proper selection of those alternatives that: first fulfil the profitability requirements and second provides better NPV and *Internal Rate of Return (IRR)* will give the decision to the planner on the equilibrium between the amortization versus modernization or substitution by the new generation technology.

Key parameters that most influence on the decision are the degree of equipment obsolescence or remaining time of the life cycle period, the new customers grow rate and the expected cash flows by the new services

The large variety of scenarios in actual networks do not allow a generic recommendation, although it is common that in new Greenfield areas and with obsolete equipment that has to be renovated anyhow it may be recommended the installation of new generation systems once all technical capabilities are available and proved. Also it is frequent that for modern equipment with required functionalities and many years remaining to fulfil the life cycle, substitution is delayed until economically feasible.

This combination of decisions is called *"cap and grow"* and is the more frequent today, but for every country, region and services demand scenario, the NPV and IRR have to be evaluated in order to ensure not only positive business results but also results in line with benchmarking values in order to survive in a competitive environment.

Chapter 6 – Network architectures and technologies

Importance of fluent and economical migration path, as well as interoperability of currently existing and new technologies should be properly addressed.

The needs of

1) established "multimedia driven" markets and service regions, and

2) the emerging basic services regions with just a limited need for high-end services

are somewhat different.

As the needs and requirements for design of the high-end (#1 above) networks are very widely discussed and well documented, but the learnings from early times of those networks have -- due to the rapid pace of telecoomunications evolution -- much vanished in the dust, not having been very well documented; this information would however now be very valuable for the bodies who face the service coverage expansion questions in context with #2 solutions above.

In this chapter different network architectures are described - existing telephony network architectures, data network architectures, data invasion of the telecommunication network, the future telecommunication network architectures. Special attention is drown on the next generation network (NGN) and the migration scenarios from the current TDM networks to this goal.

6.1. Network architectures

6.1.1 Metropolitan Area Network – MAN technology

The Evolution of Metro Networks

Regarding the physical layer, fibre optics dominates the core and metro networks. 99% of core networks are already optical. The remaining 1% is satellite and point-to-point microwave used in well defined specific situations, usually in geographically remote areas, which are sparsely populated and have very rough terrain [6.1].
In the next 15 years the number of optical channels is expected to increase from the presently common 40-80 channels to 200 channels and the bitrate per optical channels is expected to increase from the presently common 2.5-10 Gbits/s to 40-160 Gbit/s.
In parallel with the above outlined development of pure "volume" increase, the optical layer will become smarter, and the functionality implemented in the optical layer will also increase. For example, in many instances protection is already realised in the optical layer.

The protocol stack will continue to converge (eg. from IP-over-ATM-over-SDH-over-WDM to IP-over-WDM). This will bring increased efficiency through reduced functionality duplication/redundancy.

Optical Transport Networking (OTN) represents a natural next step in the evolution of transport networking. For evolutionary reasons, OTNs will follow many of the same high-level architectures as followed by SONET/SDH, ie. optical networks will remain connection-oriented, multiplexed networks. The major differences will derive from the form of multiplexing technology used: TDM for SONET/SDH vs. wavelength division for OTN. To satisfy the short-term need for capacity gain, the large-scale deployment of WDM point-to-point line systems will continue. As the number of wavelengths grows, and as the distance between terminals grows, there will be an increasing need to add or drop wavelengths at intermediate sites. Hence, flexible, reconfigurable Optical Add-Drop Multiplexers (OADMs), will become an integral part of WDM networks. As more wavelengths become deployed in carrier networks, there will be an increasing demand to manage capacity. In much the same way that digital cross-connects emerged to manage capacity into the electrical layer, Optical cross-connects (OXCs) will emerge to manage capacity at the optical layer.

Figure 6.1.1 depicts an OTN architecture covering the core, metro, and high-capacity access domains. Initially the need for optical-layer bandwidth management was most acute in the core environment, but increasingly the access network at the client or server is becoming the bottleneck for data transfer. The logical mesh-based connectivity found in the core will be supported by way of physical topologies, including OADM-based shared protection-rings, and OXC-based mesh restoration architectures. As bandwidth requirements grow for the metro and access environments, OADMs will be used there too.

It is expected that the core and metro network will evolve to consist only of IP- and WDM-technologies. The architecture of the next generation network will take advantage of the provision of an integrated IP network layer directly on top of a WDM transport layer. The encapsulation of IP over WDM can be accomplished in different ways with simplified network stacks deploying protocols such as Packet over SONET/SDH, Gigabit Ethernet or Simple Data Link.

The basic guideline for the integrated IP/WDM architecture is that WDM is considered as a backbone technology and IP is interconnected to the WDM equipment at the edges of the Core network. Such a network is mainly considered by ISPs and in particular, Competitive Operators, deploying optical infrastructure, leased or owned, willing to provide IP services on top of it using IP Points of Presence (PoPs).

The optical infrastructure will gradually evolve from ATM/SDH. Different topologies of WDM equipment may be deployed in the metropolitan and backbone areas. Incumbent operators could also deploy such a network, where in that case they integrate their existing ATM and SDH infrastructure with the DWDM equipment by using the WDM backbone or core to carry the ATM and SDH traffic.

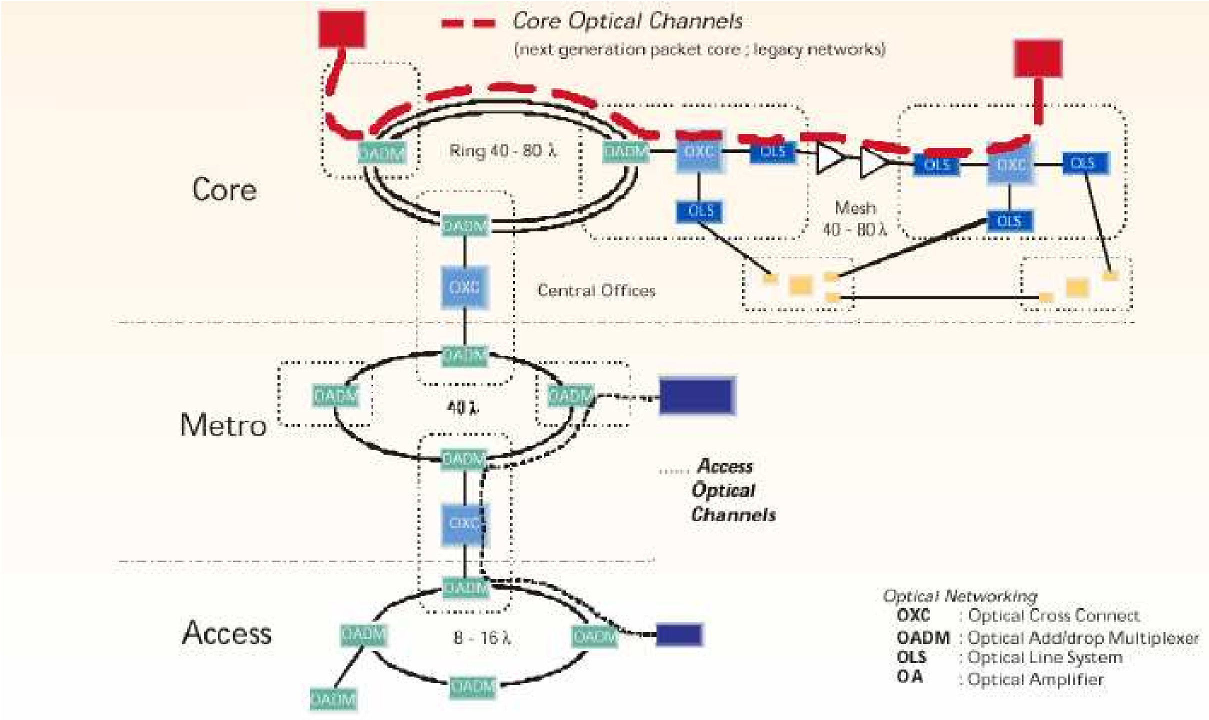


Figure 6.1.1 Optical Transport Network Architecture

Three main areas are considered in this integrated IP over WDM network architecture:

- **Backbone area**, consisting of core level IP PoPs, which are interconnected via the WDM backbone network. WDM backbone network topologies heavily depend on the distances of the IP PoPs. For long distances with significant power losses (partial) mesh networks or concatenated rings of point-to-point WDM systems are most common, while for smaller distances similar topologies to the Metro area (eg. rings) are applicable.
- **Metro area**, consisting of an optical WDM metro core with ring topologies dominating, and metro access area, where the IP PoPs are located. IP PoPs can be of 2 categories:
  - edge level ones are the gateways to the Customer Premises IP equipment
  - core or transit ones are used to groom traffic and forward it to the IP backbone
- **Access area**, where main Business/Enterprise customers or smaller Residential/Small Office/ Home Office IP customers are interconnected to the ISP acquiring Internet access.

Figure 6.1.2 depicts a future ISP's metropolitan network consisting of a WDM optical Metro core and IP Metro access. The IP section is composed of a number of IP PoPs, where customers can access the IP network services and traffic is groomed and forwarded to other PoPs or networks through the backbone. Access is facilitated to customers through the interconnection of the ISP's Provider Edge (PE) IP routers with the Customer Edge (CE) IP routers. Existing ATM and SDH equipment is shown for completeness. Provider equipment can be collocated or not with the customer equipment, depending upon the distance between customer and provider premises and on the amount of traffic generated by the customer, and the tele-housing policies.

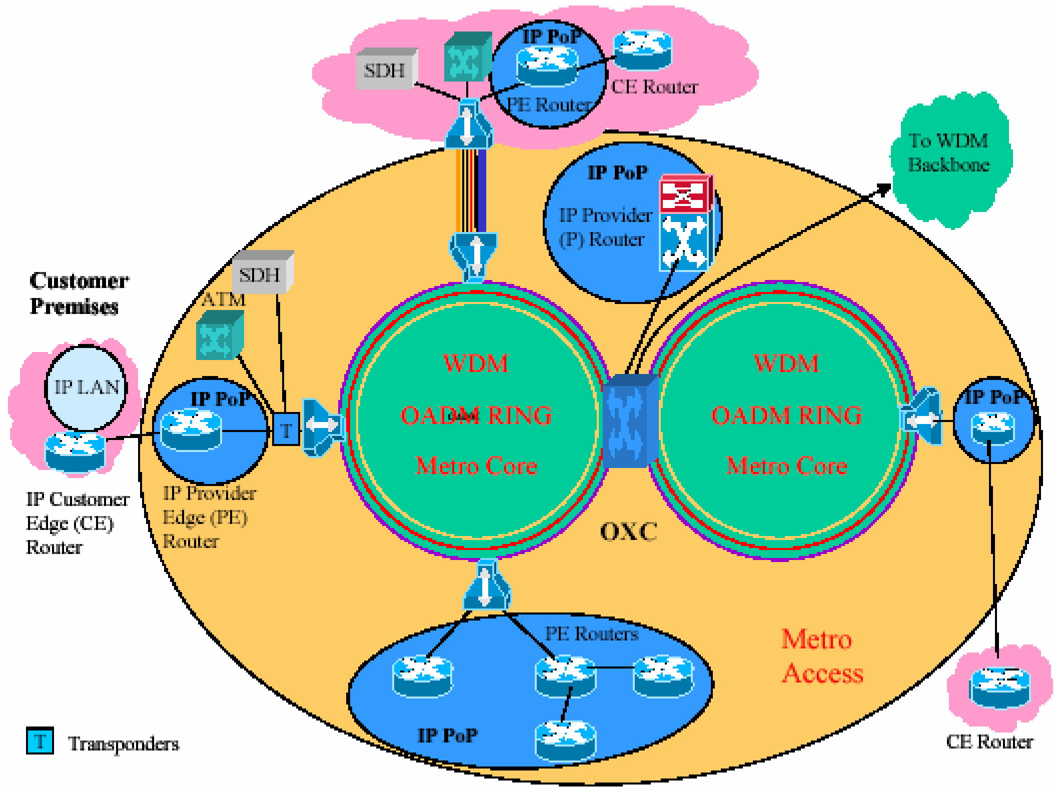


Figure 6.1.2. Metropolitan Area IP over WDM Example

The optical WDM metro core is usually composed of a ring of re-configurable OADMs, while additional point-to-point WDM links with Terminal Multiplexers can be considered for large customers. OADMs offer management interfaces so that they can be remotely re-configured to add and drop wavelengths (optical channels) to the ring through the tributary cards and multiplex them in the form of optical line signal in the corresponding line cards of the ring in each direction.

In the case where there are two WDM metro core rings, then an optical cross-connect is needed, to route wavelengths from one ring to the other supporting all-optical networking. Such cross-connects are the most expensive pieces of optical networking equipment, capable of performing additional tasks, such as wavelength switching and conversion for hundreds of ports in an all-optical form without O-E conversion.

The metropolitan network should extend the transparency and the scalability of the LAN through to the optical core network. The IP Metro access is composed of a set of PE routers interconnected via optical interfaces with OADMs. At the access side of the metropolitan network, Fast Ethernet is becoming commonplace.

However, a more-compatible methodology would be the use of optical Ethernet (40-Gigabit speeds (SONET OC–768) have already been demonstrated). Network operators may limit their customers to a few Mbit/s, but the links are gigabit-capable; and someday the fees for gigabit-scale Ethernet services will be affordable. In the meantime, the protocols and techniques for bandwidth segregation over shared links exist, work well, and are used in thousands of sites. It is a simple step to run parallel optical Ethernet trunks, each on a separate wavelength, all multiplexed over a single fibre pair using DWDM technology. In this way, a point-to-point Ethernet link could have scores of 10 Gbit/s channels, with an aggregate Ethernet bandwidth of perhaps 400 Gbit/s. Of course, this kind of network requires very large Ethernet switches at the ends of the fibres.

The limits on optical Ethernet bandwidth may be only the limit of fibre optic bandwidth (perhaps 25 Tbit/s per second for the available spectrum on today’s fibre) which is still well beyond the capabilities of today’s lasers and electronics. However, extrapolating from recent trends brings us to that level in only 5 or 10 years..

In the case that the router provides interfaces working in 15xx nm for transmission and reception, there is no need for a transponder in the OADM. The usual case, however, is that the routers’ optical interfaces work in 1310 nm and there is a need to adapt this wavelength to the 15xx, which is done by the corresponding two-way transponder. The transponder converts the optical signal of 1310 nm to electrical and back to optical.

The Wide Area Network is usually composed of a partial mesh-type optical WDM network. Transmission rates of more than 10 Gbit/s per wavelength are providing access to terabits of bandwidth between metropolitan areas. The power budget is generally sufficient for distances up to 1000km without regeneration, reshaping and retiming. Optical Amplification is deployed either to boost the aggregate multiplexed optical line signal (eg. with an Erbium Doped Fibre Amplifier) or to separately regenerate each optical channel at the corresponding tributary.

6.1.2 Access Network Technology

The Evolution of Access Networks

Broadband access needs are changing very rapidly [6.1]. Content-intensive applications are driving up the need for speed. New peer-to-peer applications such as instant messaging with text, voice and - in the future - video will push the envelope even further since they require bidirectional data streaming.
IP with Quality of Service differentiation (Differentiated Services, Integrated Services and Multi-Protocol Label Switching) is expected to become necessary to handle a range of different services.

6.1.2.1 Fixed Access Network Technologies

- ADSL - Asymmetric Digital Subscriber Line - enables a broadband always-on connection to be provided over the copper pair originally installed for POTS (typically 1-2 Mbit/s downstream and 128-512 kbit/s upstream, depending upon the distance from the exchange, and the quality of the copper pairs).

- VDSL - provides very high speed symmetric communication over short copper pairs (or co-ax CATV) for the last few hundred metres to the user and may be used in conjunction with fibre.

- Cable modems - provide a shared broadband interactive link over (upgraded) Cable TV networks, and are capable of similar data flow rates upstream and downstream to ADSL. Being a shared medium, however, the instantaneous throughput experienced is dependent upon the number of simultaneous users and their usage pattern.

- Fibre - is penetrating into access areas, but the dream of fibre to the home (FTTH) or desktop has yet to materialise, mainly because of the cost-sensitive nature of this part of the network. Passive Optical Networks provide fibre communications without expensive electronics. They are well suited to enhancing existing networks by replacing the copper between the Local Exchange and a flexibility point. A similar approach can be used with CATV networks, for instance in a Hybrid Fibre Co-ax system.

- Ethernet and fibre optics – the combination of these two technologies would provide an almost unlimited bandwidth to individual users, but the economics are still not clear.

- Powerline. Some operators have provided services using the electricity distribution network for communications. This has great potential (especially for in-home networking) but there are a number of problems to overcome.

Concerning wired access networks to the home, up to now mainly use is made of existing infrastructures of telephone-companies (phone-line), broadcast-companies (cable), and utility-companies (power-lines), using dedicated protocols like ATM, ADSL, DOCSIS, and allowing for speeds up to Mbit/s. Some companies have started investing in new infrastructures to cover the 'last mile' to the homes, notably using fibre-optic cabling, allowing for true broadband access, but requiring huge investments in the infrastructure.
The following challenges can be seen for fixed networks:

- Deal with heterogeneity, which requires bridging solutions, or a common network abstraction layer

- How to support isochronous data-transfer and plug-and-play on top of Ethernet (and IP)

- Increasing the bandwidths to support future application needs

6.1.2.2 Mobile Access Network Technologies

The increase in mobile communications and user expectations for diversified wireless services has led to the development of a variety of wireless access systems.

In particular, IEEE802.11b wireless LANs supporting up to 11 Mbit/s have become popular in the home/business area, and this technology is now being used to serve public “hot spots”. HSCSD and GPRS - are enhancements to GSM to provide a mobile service more suited to data.

UMTS, the 3rd generation of mobile systems, promises to allow data communications at up to 2 Mbit/s.

Considerable effort is underway to reconcile the different standards, typically by using multimode terminals and interworking devices. However, this approach does not seem to have all the ingredients to make the multiple existing and emerging mobile access technologies appear to the user as a single, seamless, and homogeneous network.

A possible way forward is the development of an open radio-access concept; ie. an access network which on one hand is based on a versatile air interface, and on the other hand is capable of satisfying different applications in different radio environments, when combined with IP-based backbone networks.

Besides flexibility in the air interface, such an open network paradigm requires a corresponding redefinition of layers above the physical one. In order to integrate heterogeneous mobile access networks, it is necessary to break the tie between mobile users and networks, and to move towards ways of operating that are:
- compatible with IP-based networks
- scalable; and
- distributed.

The resource management should provide an independent performance calibration ("tuning knobs") allowing network operators to set target levels, tailored to user needs, on a unified IP-based access interface.

There will be a lot of different technologies and systems that will be used for the cellular communications. Therefore in the future, software radio solutions will be developed to enable dynamic reconfiguration (for all layers) and to offer a multifrequency and multimode system.

The IP protocol will be used by all types of terminals and by all networks. The 4G terminals will be a mobile and a wireless terminal with integrated Mobile IP and Cellular IP protocols.

6.1.2.3 Dynamic handover between wireless networks

In wireless networks, terminal devices make connections to base-stations. By enabling handover, consumers are offered improved mobility. We discriminate between horizontal and vertical hand-ver.

Horizontal handover is a transition of this connection from one base station to another, within the same type of network. This may be necessary to deal with situations such as the motion of the mobile terminal, interference or a request for a different service.

A difference can be made between hard horizontal handover and soft horizontal handover between base stations:

- A hard horizontal handover is described as a horizontal handover involving a frequency change. The transition occurs in one instance, ie. the mobile will give up its connection to one base station completely before it established a connection to the second. In a well-managed network it should know where to find the second connection but unfortunately calls can be lost. Hard handover is also called ‘break-before-make’ handover.
- A soft horizontal handover is described as a handover where the second connection is established before the first is dropped, using the same frequency. In this way a mobile may during this period be communicating simultaneously with two base stations, and calls should not be dropped. Soft handover is also called ‘make-before-break’ handover.

Vertical handover is the process of handover between different types of networks. This may be a handover between standards of the WAN/cellular type, like GSM and UMTS, but also of the LAN/hot-spot type, like WiFi, and the PAN/personal type, like Bluetooth. This type of handover typically is hard (involving a frequency change).

Multi-band (via hard horizontal handover) and multi-standard (via vertical handover) mobile terminals are already available on the market, although they only cover the second-generation (2G) GSM mobile cellular standards that operate at 900 MHz and 1.8 GHz with the addition of 1.9 GHz for use in the USA. Currently, terminals are being developed that will also support the third-generation (3G) standard UMTS, that operates at 2 GHz, together with some of the aforementioned 2G standards.

Roaming (via soft horizontal handover) is possible with all contemporary terminals.
Regarding the multi-band, the benefit offered to the user by the hard-horizontal handover is an extended service provision in terms of geographical area, since the services offered by each network (on a different band) are very similar (voice, SMS and in some cases data if GPRS is involved).

Regarding the multi-standard, in all these terminals the system operates using only one standard at any one time, since components are shared within a single radio architecture. This limits their handover to hard vertical handover between the various standards.

Terminals that can operate simultaneously on more that one wireless standard are not yet available. Simultaneous operation may however only be a perception by the user and achieved within the terminal using soft vertical handover between the various standards.
An additional need arises to support wireless LAN standards within the terminal in addition to the existing mobile cellular 2G and emerging 3G standards. Work has already commenced in the ETSI BRAN project and in the 3GPP on the inter-working between the 3G and Hiperlan/2 standards. The reason behind this is that basically LANs provide support for higher data rates that are offered by the cellular networks. They tend, however, to be restricted to short-range and mobility, implying no contiguous coverage and not suitable for high speeds of mobility. So, cellular WANs and “hot-spot” LANs can be seen as complementary, in terms of data rates but also service provisioning.

The following challenges can be seen for dynamic handovers:
- Develop terminals that can operate simultaneously on more than one wireless standard
- Enable seamless handover between WAN/cellular, LAN/”hot-spot”, and PAN/personal networks

6.1.2.4 Wireless LAN Market Trends

Most WLANs are employed to augment rather than replace wired LANs. They provide connectivity to a LAN in places where wiring is difficult, costly or inconvenient to employ. Common applications for WLANs may include the following:

- Museums and archaeological places

- Hospitals, recording patient information at bedside

- Car rental companies, to input car-return information
- Warehouse and retail shops, to keep inventories
- Restaurants placing orders
- Offices that extend networks into boardrooms and libraries
- Schools
- Wireless meeting rooms
- Wireless business centres
- Wireless small offices

The total WLAN revenue was $839 million in 2000 and $1,056 million in 2001 (IDC).

6.1.2.5 Fixed-Wireless Access Technologies

Fixed wireless access (eg. LMDS) - systems use radio links to provide connections to customers in fixed locations. It is suitable for broadcast applications as well as broadband telecommunications.

The wireless Internet markets are opening quickly after 2001. Although most of the attention has been focused on cellular telephony, the fixed wireless Internet will be also extremely important market area. In fact, the fixed wireless will be the first technology that is implementing a real wireless IP connectivity. The transitional technologies such as WAP or UMTS are not inherently TCP/IP-compatible in the sense that they do not allow for the transparent flow of Internet traffic. The large bandwidth, relatively good channel conditions with fixed wireless etc. will make it possible to build IPv6 compatible wireless links well before actual mobile Wireless Internet is possible.

Overall, LMDS compares favourably with competing options on the basis of both performance and cost, but it lacks the wide support and financial backing which other platforms possess. Industry support has flooded behind cable modem and ADSL technologies, and this could prove to be significant.

The computing industry also seems to be supporting ADSL, and to a lesser extent cable modems, as a means of delivering multimedia content to homes and businesses. This industry has a lot to gain from the success of broadband delivery, and logic suggests, that its members will go to great lengths to bring the most likely broadband technologies to the mass market. ADSL seems to be that technology.

Only a few companies have publicly committed to supporting LMDS platform and those, which have, lack the distribution, name-brand awareness, and financing which supporters of ASDL and cable modems possess. This difference is likely to result in an LMDS CPE that costs more than - and lacks the distribution of - cable and ADSL modems. In addition, there will be lower visibility for LMDS. Time could also be an issue. If cable modems and ADSL services become widely accessible within the next two years, then the deployment of LMDS could prove unattractive in areas that already possess other alternatives.

There are a number of variables that could drastically alter the market and the fortunes of LMDS providers. ADSL and cable modem deployment could lag considerably behind expectations, and satellite and many LMDS operators may not even build out their networks. However, developments in the market today suggest that leading LMDS auction winners will deploy networks, and that these Service Providers will concentrate their efforts on business and well-to-do residential customers. The high-cost of CPE will preclude deployment to other residential areas, at least initially.

- Geostationary satellites and terrestrial broadcasting - can now provide broadband (asymmetric) interactive capability using the fixed network (eg ISDN) for the upstream path.

- Low Earth Orbit satellites and High Altitude Platform Stations – considerably reduce the problems caused by the transmission time to and from geostationary satellites but have not yet been proven commercially viable.

In the near future, residential access is expected to remain copper-based, using technologies such as xDSL to boost the capacity of traditional copper lines. However, for business offices, optical technology is already being used to bring high bandwidth to the end-user, with ATM and SDH access equipment at the customer premises. The next step is to use WDM technology for these environments. WDM will first be used in industrial and campus LAN environments. The DWDM network at the Microsoft headquarters in Redmond is a good example of a trial of these latest technologies, which use DWDM in the enterprise environment. This will become technically and economically feasible due to the very large number of wavelengths that a single fibre can carry, thus spreading the cost to more subscribers. Introducing more wavelengths per fibre can also lead to new topologies for home access by using ring or bus like structures with an add/drop port per home so that each home has its own wavelength.

Nevertheless, the point-to-point optical fibre star structure is preferred for business customers with critical security requirements. In major cities, fibre already connects most big business offices (FTTO) and some residential buildings with ring or star structures. Fibre is getting closer to small business customers and residential customers with double star or tree-branch structures. Fibre to the cabinet (FTTCab) and fibre to the curb (FTTC) are becoming more common, also fibre to the town (FTTT) and fibre to the village (FTTV) are increasingly popular.

6.1.3 The evolution of home networks

Homes contain many kinds of network technologies, for example:

- analogue/ISDN/ADSL/CATV/Ethernet/WLAN for communicative, interactive services

- CATV, satellite links, etc. for entertainment services

- various low speed smart devices, interconnected and controlled by radio, fixed, infrared, … types of network.

6.1.3.1 Fixed home networks

A lot of in-home networking standards require cabling between the devices. One option is to install new cabling in the form of galvanic twisted-pair or coaxial wires, or optical fibres. The alternative is to use existing cabling, like power-lines and phone-lines. Especially for the existing cabling a wide range of proprietary standards exist, but we will limit ourselves to the more interesting open standards.

Using existing cabling in the home is very convenient for end-users: «no new wires!». For in-home networking via the phone-line, HomePNA22 has become the de-facto standard, providing up to 10 Mbit/s, where 100 Mbit/s is expected. For power-line networking, low-bandwidth control using X1023, and (high) bandwidth data transfer using CEBus and HomePlug are the most prominent ones, offering from 10 kbit/s up to 14 Mbit/s.

Premium performance is obtained when using new cabling. New cabling requires an additional effort of installation, but has the advantage that premium-quality cabling can be chosen, dedicated to digital data-transport at high rates. The IEEE-1394a standard (also called Firewire and i.Link) defines a serial bus that allows for data transfers up to 400 Mbit/s over a twisted-pair cable, and extension up to 3.2 Gbit/s using fibre is underway. Similarly, USB defines a serial bus that allows for data transfers up to 480 Mbit/s over a twisted-pair cable, but using a master-slave protocol instead of the peer-to-peer protocol in IEEE-1394a. Both standards support hot plug-and-play and isochronous streaming, via centralised media access control, which are of significant importance for consumer-electronics applications. A disadvantage is that this sets a limit to the cable lengths between devices.

Another major player is the Ethernet (also known as IEEE 802.3) which has evolved via 10 Mbit/s Ethernet and 100 Mbit/s Fast Ethernet, into Gigabit Ethernet, providing 1 Gbit/s using fibre. Ethernet notably does not support isochronous streaming since it lacks centralised medium-access control. Also it does not support device discovery (plug-and-play). It is however widely used, also because of the low cost.

Currently there is no dominant wired networking standard for in the home, and networks are likely to be heterogeneous, incorporating multiple standards, both wired and wireless.

6.1.3.2 Wireless home networks

As opposed to wired networks, wireless systems are far easier to deploy. Already widely deployed in Europe is the well-known DECT technology, notably for voice communication. For services other than voice, they can be divided into two categories: Wireless PANs and LANs.

Wireless Personal Area Networks (PANs) typically have a short range-of-use (10-100 meters), and are intended to set up connections between personal devices. The most widely deployed standard in this class is Bluetooth. Its capability is providing 1 Mbit/s for few connected devices in a small network, called piconet. Its range is between 10 and 100 meters depending on the transmission power. The used transmission-band for Bluetooth lies in the 2.4 GHz ISM band (license-free).

The IEEE 802.15 standard is intended to go a step further. It integrates the Bluetooth standard and harmonizes it with the IEEE 802 family, such that it is IP and Ethernet compatible. The objectives are a high-bit rate solution (IEEE 802.15.3) providing up to 20 Mbit/s, and a low bit-rate one (IEEE 802.15.4, also known as ZigBee).
The HomeRF standard, like Bluetooth, also works in the 2.4 GHz ISM band. From an initial maximum data rate of 1.6 Mbit/s, it has been extended to 10 Mbit/s. HomeRF has a range of 50 meters at this speed. It is not interoperable with its strongest competitor, IEEE 802.11b (see below), however.

Wireless local area networks (LANs) have a broader application area: their purpose is to provide a wireless connection for networked devices like laptops or even handheld devices, not restricted to one person. The IEEE 802.11 series of standards are leading in this area: The IEEE 802.11b (WiFi) standard uses the 2.4 GHz band, and the IEEE 802.11a standard the 5 GHz band. Notably the 802.11b standard is gaining market share. Capabilities of 802.11 are to provide up to 54 Mbit/s over 300 meters distance. ETSI former Hiperlan2 standard has now merged with 802.11a, giving some features that were already considered like power control and QoS.

Concerning wireless networks to the home, the driving and most deployed systems are DVB-based access networks. Currently, they are mainly deployed through satellite transmission, for Digital TV broadcasting services. Interactive services are provided through the use of eg. telephone-lines for the (narrow-band) return channels. The technology has the main characteristic to be a broadcast and reliable (with very low error rate) link supporting around 1Gbit/s in total, and thereby able to transport hundreds of compressed TV programs. In parallel, some data-based services can be carried, adding extra features around the TV programs, such as electronic program guides (EPGs) and encryption keys. For terrestrial transmission of digital TV, the DVB-T standard has been standardised and will be deployed in the near future progressively. Its purpose is the same, but the number of carried TV programs will be limited to about 40.

Some wireless fixed broadband-access solutions have also been standardised, with relatively poor success. The local multipoint distribution service (LMDS) is being used for point-to-multipoint applications, like Internet access and telephony. It only has a 3-mile coverage radius, however. The multichannel multipoint distribution service (MMDS) was initially used to distribute cable television service. Currently it is being developed for residential Internet service. However, installations have not been profitable and service delays have been widespread. Currently, new standards have being defined: e.g. the IEEE 802.16 (WirelessMAN) standard addresses metropolitan-area networks; amendment 802.16a expands the scope to licensed and license-exempt bands from 2 to 11 GHz. ETSI is following a similar track for Europe.

The following challenges can be seen for wireless home networks:

- To deal with governmental regulations that vary widely throughout the world, and prevent interference, especially in the license-free spectrum bands, to ensure optimal network performance.

- Power consumption for mobile devices: since wireless networks enable mobile applications, their success relies on the duration and limited weight of the devices batteries. One of the requirements driving the development of Bluetooth was to have low-cost, low power consumption devices.

- To enable the use of wireless networks in consumer applications of every day life, seamless integration of new devices is critical. This involves interoperability for both low level protocols (plug-and-play devices) as well as higher-level functionality.

- Wireless networks have specific features such as loss of packets and bit rate modifications that have a significant impact on some applications requiring a constant QoS such as video. Adaptation of data transport to the constraints of wireless networks with techniques such as error resilience, scalability or joint source-channel coding is therefore critical.

Interworking and interoperability, as well as the seamless provision of services, independent of the underlying networks is the most challenging topic to be addressed in the access and home network environments. The standards arena of home networks is another area, which is currently too diversified and hence there is a number of proprietary technologies and interfaces. This is not a cost-effective solution that can exist in the long term.

6.1.3.2.1 Heterogeneous in-home networks

From the previous sections it is clear that several technologies for in-home networking exist. These standards and technologies differ in:

- Application domain (home control, communication, infotainment, entertainment)

- Middleware technology (HAVi, UPnP, Jini, etc.)

- Connection technology (based on new wiring (coax, twisted-pair, fibre), on existing wiring like power-line and phone-line, or wireless).

At least for the coming years, but even in the long run, there will not be a clear winner, and it is expected that several technologies will co-exist. Moreover, since there is no main player dominating the home infrastructure, all kind of technology combinations will co-exist within a single home network making it fully heterogeneous. This implies that networked devices, services and applications will only be successful, if they are prepared to run within a heterogeneous environment. Therefore, a strong research need arises to develop bridges and gateways that can couple the different clusters in a heterogeneous home network.

The heterogeneity can appear both at the lower (data transport focused) and higher (middleware) layers of the ISO OSI protocol stack.

An overview of the various wired and wireless data transport focused standards we provided before. Here we focus on heterogeneity at the middleware layer, which means that different middleware standards are present dealing diversely with fundamental issues like:
- Device models and definitions
- Resource management
- Event management
- Stream management
- Plug-and-Play mechanism
- User-interface concepts

The available (combinations of) middleware standards heavily influence both the architecture of the devices, as well as the architecture of services and applications in the network.
To make a proper architectures for devices and applications it is essential to have adequate knowledge of the various middleware technologies they may end up working on. The state-of-the-art middleware technologies for entertainment/infotainment are: HAVi, UPnP, Jini, Bluetooth, WAP. They vary in the protocol stack, some of them like UPnP or Jini, are bound to a specific layer (network, eg. IP) in the OSI stack. Others, like HAVi, Bluetooth and WAP, are tightly coupled to a specific communication medium but stretch out far into the application layers.

There is a pervasive use of sensors in all areas of our lives, and these will be increasingly miniaturised, equipped with embedded intelligence and capabilities for being networked so that they can communicate with other devices.
Besides the in-home networks, there are the various access networks to the outside world, via telephone, cable, satellite, etc. These will have different characteristics and have an impact on the provided external services. To make good architectures for devices and applications it is essential to have good knowledge of the properties from the different access networks and the services offered through them.

The following challenges can be seen for heterogeneous in-home networks:

- To deal with heterogeneity at lower layers requires development of bridging solutions, or a common network abstraction layer like IP

- To deal with heterogeneity at higher layers requires development of gateways coupling different middleware standards

6.1.3.3 Ad-hoc Networks

Ad-hoc networking enables users to co-operatively form dynamic and temporary networks without any pre-existing infrastructure, using capabilities of the underlying, mostly wireless, network. This contrasts with infrastructure-based networks.

In its most primitive form, ad-hoc networking enables direct communication between any two mobile nodes that are in the wireless transmission range. This may be the only way of communicating. This is, for instance, the case for Bluetooth, which facilitates ad-hoc connections for stationary and mobile communication environments. On the other hand, the long-range IEEE 802.11 standard defines two modes of operation: base-station mode (BSS, basic service set), and ad-hoc mode (IBSS, independent basic service set).

Every mobile node operating in the BSS mode must be in the transmission range of one or more base stations, which are responsible for buffering and forwarding traffic between nodes. Nodes can send outgoing traffic to the base station anytime and periodically poll the base station to receive incoming traffic, while being in the sleep mode the remaining time. The ad hoc operation mode does not use any base station infrastructure; nodes communicate directly with all other nodes that are in the wireless transmission range. Because there is no base station to moderate communication, nodes must always be ready to receive traffic from their neighbours.

Ad-hoc routing protocols extend the basic one-hop ad-hoc networking by managing the routing of messages among mobile nodes. Proposed protocols are typically implemented over IEEE 802.11 and maintain a routing table from which the path of mobile nodes, to be followed for sending a message from one mobile node to another, may be retrieved.
The main issue to be addressed in the design of an ad-hoc routing protocol is to compute an optimal communication path between any two mobile nodes. This computation must minimise the number of control messages that are exchanged among mobile nodes in order to avoid network congestion but also to minimise energy consumption.

There exist two basic types of ad-hoc routing protocols: proactive and reactive. Proactive protocols (eg. OLSR [JMQL+01]) update their routing table periodically. Reactive protocols (eg. AODV , DSR , TORA) a-priori reduce the network load due to the traffic of control messages, by checking the validity of, and possibly computing, the communication path between any two mobile nodes only when a communication is requested between the two. ZRP is a hybrid protocol that combines the reactive and proactive modes. The design rationale of ZRP is that it is considered advantageous to accurately know the neighbours of any mobile node (ie. mobile nodes that are accessible in a fixed number of hops, whose optimal value is of 3-4). Hence, ZRP implements both a proactive protocol for communicating with mobile nodes in the neighbourhood (referred to as the zone), and a reactive protocol for communicating with the other nodes.

Ad-hoc networking is a very cost-effective solution. In general, it is very convenient for accessing services and data that are present in the local (physically close) area, and for possibly reaching a WLAN base station, all at little or no cost for the user.
Ultimately, the user may decide to pay for communication using (global) mobile-service networks, if the connectivity using the ad-hoc network connection to a wireless LAN is of poor quality.

The following challenges can be seen for ad-hoc networks:

- Middleware over ad hoc networks. Ad-hoc networking allows for setting up collaborative networks among mobile nodes based on their geographical proximity and/or sharing of interest. However, middleware that sets up such a collaborative environment still needs to be devised, offering in particular base services for the management of ad hoc grouping and of security. In addition, the middleware must be designed so as to minimise energy consumption by and optimise the performance of the devices involved in the collaboration.

- Setting up an ad hoc group of mobile nodes using the underlying ad hoc network relies on the design of adequate services for service discovery and lookup and for dynamic group configuration. These issues are further discussed in the sections on Lookup and discovery of services and devices and Service composition.

- Ad-hoc networking raises the issue of ensuring end-to-end privacy and integrity of the users' data. However, strong security enforcement must be balanced with consumption of resources and in particular of energy.

- The increasingly powerful hardware embedded in mobile nodes and the relatively slow increase of battery capacity require devising adequate solutions to energy saving on he mobile nodes for all the constituents of the mobile environment, ie. application software, network operating system and hardware. In particular, communication is one of the major sources of energy consumption. This has thus led to the design of wireless communication protocols that reduce energy consumption. However, most of the communication protocols that have been proposed for ad hoc networks are assessed in terms of bandwidth usage and not energy consumption. There are actually few protocols that are specifically aimed at reducing energy consumption, which cannot be managed by simply optimising bandwidth usage between the sender and the receiver. In addition, it is mandatory for these protocols to be coupled with distributed application software (ie. application and middleware) that are designed so as to minimise energy consumption, including the one associated with communication.

- It is crucial to design the middleware with performance improvement in mind regarding both resource usage and response time, which are often contradictory. Such a concern relates in particular to how the aforementioned open issues are addressed. In addition, it requires integrating well-known techniques for performance improvement such as caching.

- Integrating ad-hoc and base-station modes.

6.2. New network technologies

6.2.1. On information carrying procedure and associated routing: IP, MPLS, MGW, etc.

6.2.1.1 Generation IP (IPv6) and Transition Strategies from IPv4 to IPv6

IPv6 is the next generation protocol designed by the IETF to replace the current version of the Internet Protocol, IPv4. During the last decade, IP has conquered the world's networks. Most of today's Internet uses IPv4, which is now more than twenty years old. IPv4 has been remarkably resilient in spite of it's which has been remarkably resilient in spite of its age, but it is beginning to have problems.

Introduction to IPv4

The Internet Engineering Task Force published the IPv4 specification (RFC 791) in the fall of 198l. When the IPv4 specification was released, the Internet was a community of approximately one thousand systems. The IPv4 specification called for every IP address to be represented by a 32-bit number made up of four groups of eight-bit numbers. This provides a total of just over four billion addresses, although only a few hundred million are actually available due to hierarchic allocation schemes.
Since the release of IPv4, the Internet population has grown to over 100 million computers, increasing far faster than anticipated. As the pool of available addresses decreases, it will become increasingly difficult to obtain IPv4 addresses.
Furthermore, the pace of this growth is expected to continue for years to come. The bottom line is this: *The Internet is running out of addresses*. And by some hard estimates, this could happen as soon as 2002.Early IP assignments reserved addresses for some corporations and institutions in very large blocks. These “Class A” and “Class B” network assignments were issued in the early days when the current growth was not anticipated. While some early adopters may still have addresses available for internal usage, the pool of unissued addresses is becoming smaller every day. The addresses that were handed out to some of the early large corporate networks cannot now be reissued to other users.

		7 bits	**24 bits**
Class A	**0**	**Network**	**Host**

		7 bits	**24 bits**
Class B	**0**	**Network**	**Host**

		21 bits	**8bits**
Class C	**0**	**Network**	**Host**

Figure-6.2.1 Mapping of network and host identifier in the 32-bit IPv4 protocol

Why IPv6?

Introduction to IPv6

IPv6 vs. IPv4

Transition Strategies
There are three general strategies to deal with transitioning to IPv6. They can be used independently of each other, or in combination.

Dual-Stack approach
Dual-Stack Strategies, which allow IPv4 and IPv6 to co-exist in the same devices and networks. When adding IPv6 to a system, do not delete IPv4. This multi-protocol is well understood (IPX, AppleTalk) and applications or libraries choose the IP version to use based on a DNS response (AAAA or A6 record) or the version of the initiating packet.

Ways for the dual stack approach operates

IPv4 and IPv6 Interoperability

Tunneling   approach

Tunnels allow you to get packets through IPv6 ignorant Routers and Switches by encapsulating IPv6 packets inside of IPv4 packets or MPLS frames.

Methods for establishing a tunnel
A. Manual configuration
B. Tunnel brokers
C. 6-over-4
D. 6-to-4

Types of tunnels

A. Static point to point tunneling

B. Automatic tunneling

C. 6to4 tunneling

D. 6to4 to 6to4 tunneling

E. 6to4 upstream tunneling

F. downstream tunneling

Translation approach

Translation is for use for those who may prefer to use IPv6-IPv4 protocol translation for new kinds of internet devices (cell phones, cars, appliances), or for those who need to entirely shed the IPv4 stack on their LANS.

Translation Requirements

System Architecture

NAT-PT(Network Address Translation- Protocol Translation)

Phases for transition

Conclusion

6.2.1.2 MPLS (Multiprotocol Label Switching)

MPLS is an end-to-end forwarding paradigm; it essentially establishes a tunnel across the network. When a Datagram arrives at the ingress edge device; it is tagged with a “label”. The label represents the path or Route the datagram will take to reach its destination. At each node in the path, only the label is used by Devices to make forwarding decisions about the datagram. In contrast, unlabeled datagrams require each Node to extract and interpret several fields in the datagram’s encapsulation to make forwarding decisions, some requiring several look-ups and computations to be performed.
MPLS networks use labels to forward packets . The ingress MPLS node assigns a packet to a particular Forwarding Equivalence Class(FEC). The FEC to which the packet is assigned is encoded as a short fixed-length value known as a label. The packets are labeled they are forwarded. At subsequent hops, there is no further analysis of the packets network layer header . The label is used as an index in to a table, which specifies the next hop, and a new lable. The old label is replaced with the new label , and the packet is forwarded to it's next hop.
In MPLS networks, labels drive all forwarding. This has a number of advantages over conventional network layer forwarding:

- MPLS forwarding can be performed by switches, which can do label lookup and replacement but can't analyze the network layer headers. ATM switches perform a similar function by switching cells based on VPI/VCI values found in the ATM header. If the VPI/VCI values are replaced with label values, ATM switches can forward cells based on label values. The ATM switches would need to be controlled by an IP based MPLS control element such as Label Switch Controller(LSC). This forms the basis of integrating IP with ATM using MPLS.

- A packet is assigned to a FEC when it enters the network. The ingress router may use any information it has about the packet, such as ingress port or interface,even if that information cannot be obtained from the network layer header. A packet that enters the network at a particular router can be labeled differently than the same packet entering the network at a different router. As a result, forwarding decisions that depend on the ingress router can be made easily. This cannot be done with conventional forwarding, because the identity of a packet's ingress router does not travel with the packet. For example, packets arriving on different interface connected to CPE routers might be assigned to different FECs. The attached labels would represent the corresponding FECs. This functionality forms the basis for the building of MPLS Virtual Private Networks.
- Traffic-engineered networks force packets to follow a particular path, such as an underutilized path. This path is explicitly selected when or before the packet enters the network, rather than being selected by normal dynamic routing algorithm as route, so the identity of the explicit route need not be carried with the packet. This functionality forms the basis of MPLS traffic engineering.
- A packet's "class of service" may be determined by the ingress MPLS node. An ingress MPLS node may then apply different discard thresholds or scheduling disciplines to police different packets. Subsequent hops may enforce the service policy using a set of per-hop behaviors(PHBs).MPLS allow(but does not require) the precedence or class of service to be fully or partially inferred from the label.in this case, the label represents the combination of a precedence or class of service. This functionality forms the basis of MPLS Quality of service (QoS).

MPLS Node Architecture

MPLS nodes have two architectural planes:
1-the MPLS forwarding plane
2-the MPLS control plane.
MPLS nodes can perform Layer 3 routing or Layer 2 switching in addition to switching labeled packets. Figure 6.2.2 shows the basic architecture of an MPLS node.

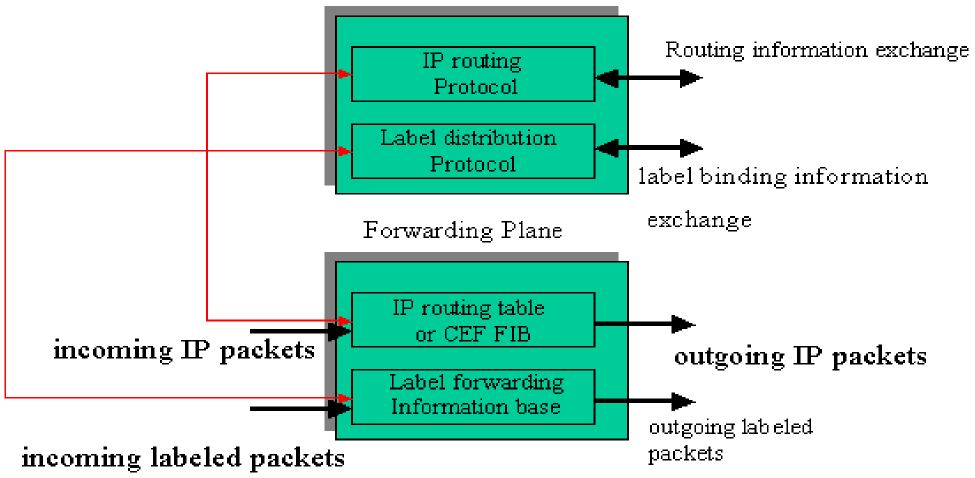


Figure 6.2.2

Forwarding plane

The MPLS forwarding plane is responsible for forwarding packets based on values contained labels. This forwarding plane uses a label forwarding information base (LFIB) maintained by the MPLS node to forward labeled packets. The algorithm used by the lable switching forwarding component uses information contained in the LFIB as well as the information contained in the label value. Each MPLS node maintains two tables relevant to MPLS forwarding
1-the label information base (LIB)
2-the label forwarding information base (LFIB)
The LIB contains all the labels assigned by the local MPLS node and the mapping of these labels to labels receiving from its MPLS neighbors. The LFIB uses a subset of the labels contained in the LIB for actual packet forwarding.

MPLS Label

The label is a condensed view of the header of an IP packet, although contained within it is all of the information needed to forward the packet from source to destination. Unlike the IP

header, it does not contain an IP address, but rather a numerical value agreed upon by two MPLS nodes to signify a connection along an LSP. The label is a short, fixed-length, physically contiguous identifier that is used to identify a FEC, usually of local significance. A packet assigned to a given FEC
Is usually based on its destination address, either partially or completely. The label, which is put on a particular packet, represents the FEC to which that packet is assigned. Within some transport mediums, there are existing labels that can be used by MPLS nodes when making forwarding decisions, such as ATM's virtual path identifier/virtual circuit identifier (VPI/VCI) field and frame relay's data link connection identifier (DLCI). Other technologies, such as Ethernet and point-to-point links, must use what is called a shim label, shown in Figure 6.2.3 The shim label is a 32-bit, locally significant identifier used to identify a FEC.

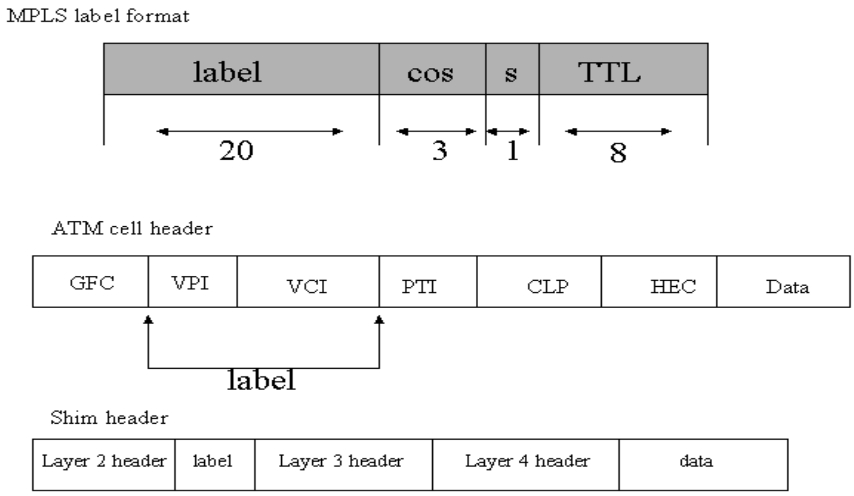


Figure 6.2.3

Label: MPLS label
Cos: Class of service
S: Bottom of stack
TTL: Time to Live

The MPLS lable contains the following fields:

- *lable field(20 bits):* Carries the actual value of the MPLS lable.

-*CoS field(3 bits):*Affects the queing and discard algorithms applied to the packet as it is transmitted through the network.

-*stack field(1 bit):*Supports a hieraachical label stack.

-*TTL(Time-to-Live)field(8bit):*provides conventional IP TTL functionality.

Label Stack

The label stack mechanism allows for hierarchical operation in the MPLS domain. It basically allows MPLS to be used simultaneously for routing at the fine-grain Each level in a label stack pertains to some hierarchical level. This facilitates a tunneling mode of operation in MPLS.unicast IP routing does not use stacked lables,but MPLS VPNs and traffic engineering utilize stacked lables for their operation.

TTL

The TTL field is similar to the time-to live carried in the IP header.the MPLS node only processes the TTL field in the top entry of the lable stack.the IP TTL field contains the value of the Ipv4 TTL field or the value of the Ipv6 Hop Limit field whichever is applicable.

Label Forwarding Information Base(LFIB)

The label forwarding base(LFIB) maintained by an MPLS node consist of a sequence of entries.As illustrated in figure 6.2.4 each entry consists of an incoming label and one or more subentries.the LFIB is indexed by the value contained in the incoming label.each subentry consisits of an outgoing interface and next hop address.subentries within an individual entry may have the same or different outgoing labels. Multicast forwarding requires subentries with multiple outgoing labels,where an incoming packet arriving at one interface needs to sent out on multiple outgoing interfaces.in addition to the outgoing label,outgoing interface and next hop information an entry in the forwarding table may include information related to resources the packet may use such as an outgoing queue that the packet should be placed on.an MPLS node can maintain a single forwarding label per each its interfaces,or a combination of both.in the case of multiple forwarding table instances packet forwarding is handled by the value of incoming label as well as the ingress interface on which the packet arrives.

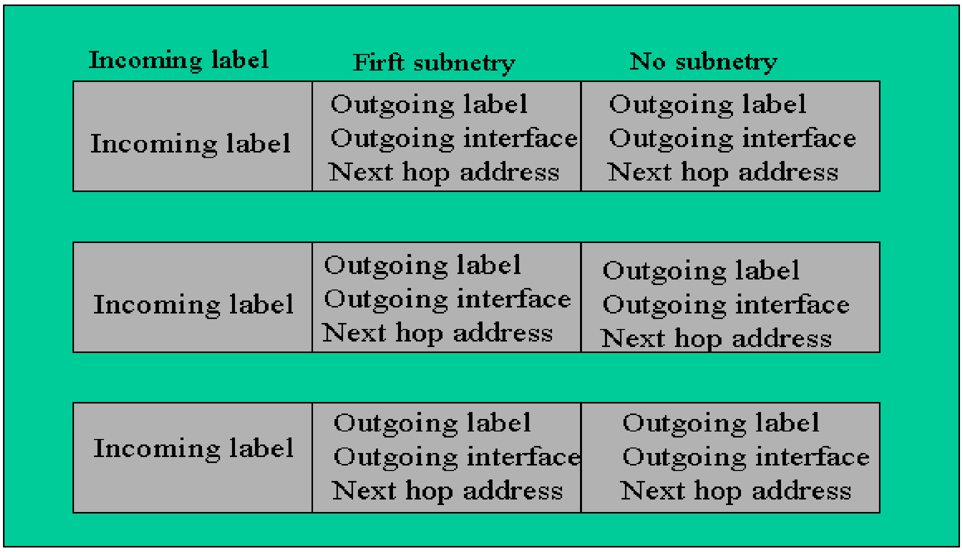


Figure 6.2.4

Labels and Label Bindings

A label, in its simplest form, identifies the path a packet should traverse. A label is carried or encapsulated in a Layer-2 header along with the packet. The receiving router examines the packet for its label content to determine the next hop. Once a packet has been labeled, the rest of the journey of the packet through the backbone is based on label switching. The label values are of local significance only, meaning that they pertain only to hops between LSRs. Once a packet has been classified as a new or existing FEC, a label is assigned to the packet. The label values are derived from the underlying data link layer. For data link layers (such as frame relay or ATM), Layer-2 identifiers, such as data link connection identifiers (DLCIs) in the case of frame-relay networks or virtual path identifiers (VPIs)/virtual channel identifiers (VCIs) in case of ATM networks, can be used directly as labels. The packets are then forwarded based on their label value.
Labels are bound to an FEC as a result of some event or policy that indicates a need for such binding. These events can be either data-driven bindings or control-driven bindings. The latter is preferable because of its advanced scaling properties that can be used in MPLS

Control Plane

The MPLS control plane is responsible for populating the LFIB. All MPLS nodes must run an IP routing protocol to exchange IP routing information with all other MPLS nodes in the network. MPLS enabled ATM nodes would use an external Label Switch Controler(LSC) such as 7200 or 7500 rouer or use a built-in Router Processsor Module (RPM) in order to participate in the routing process. The lable exchanged with adjacent MPLS nodes are used to build the LFIB. MPLS use aforwarding paradigm based on label swapping that can be combined with a range of different control modules. Each control madule is responsible for assigning and distributing a set of labels, as well as for maintainning other relevant control information. IGPs are used to define reachability, binding and mapping between FEC and next-hop addresses.
MPLS control modules include:

Unicast routing module
the unicast routing module builds the FEC table using conventional Interior Gateway Protocol (IGPs) such as OSPF, IS-IS, and so on. The IP routing table is used to exchange lable bindings with adjacent MPLS nodes for subnets contained in the IP routing table.the lable binding exchange is performed using LDP.

Multicast routing module

The multicast routing moudule builds the FEC table using a multicast routing protocol such as protocol Independent Multicast(PIM). The multicast routing table is used to exchange label binding with adjacent MPLS nodes subnets contained in the multicast routing table. The label binding exchange is perfomed using the PIMv2 potocol with MPLS extensions.

Traffic engineering module

The traffic-engineering module lets explicitly specified label-switched paths be set up through a network for traffic engineering purposes. It uses MPLS tunnel definitions and extensions to IS-IS or the OSPF routing protocol to build the FEC table. The label-binding exchange is performed using the Resource Reservation Protocol (RSVP)or Constraint-based Routing LDP(CR-LDP), which is a set of extensions to LDP that enables constraint-based routing in an MPLS network.

Virtual Private Network (VPN)module

The VPN module uses per VPN routing table for the FEC tables, which are built using routing protocols run between the CPE routers and service provider edge MPLS nodes. The label binding exchange for the VPN-specific routing table is performed using extended multiprotocol BGP inside the service provider network.

Quality of service(QoS)module

The QoS module builds the FEC table using conventional Interior Gateway Protocols(IGPs) such as OSPF,IS-IS,etc. The IP routing table is used to exchange label bindings with adjacent MPLS nodes for subnets contained within the IP routing table. The lable binding exchange is performed using extension to LDP.

<u>MPLS Components</u>

In MPLS, data transmission occurs on label-switched paths (LSPs). LSPs are a sequence of labels at each and every node along the path from the source to the destination. LSPs are established either prior to data transmission (control-driven) or upon detection of a certain flow of data (data-driven). The labels, which are underlying protocol-specific identifiers, are distributed using label distribution protocol (LDP) or RSVP or piggybacked on routing protocols like border gateway protocol (BGP) and OSPF. Each data packet encapsulates and carries the labels during their journey from source to destination. High-speed switching of data is possible because the fixed-length labels are inserted at the very beginning of the packet or cell and can be used by hardware to switch packets quickly between links.

LSRs and LERs

The devices that participate in the MPLS protocol mechanisms can be classified into label edge routers (LERs) and label switching routers (LSRs).
An LSR is a high-speed router device in the core of an MPLS network that participates in the establishment of LSPs using the appropriate label signaling protocol and high-speed switching of the data traffic based on the established paths.
An LER is a device that operates at the edge of the access network and MPLS network. LERs support multiple ports connected to dissimilar networks (such as frame relay, ATM, and Ethernet) and forwards this traffic on to the MPLS network after establishing LSPs, using the label signaling protocol at the ingress and distributing the traffic back to the access networks at the egress. The LER plays a very important role in the assignment and removal of labels, as traffic enters or exits an MPLS network.

FEC(Forward Equivalence Class)

The forward equivalence class (FEC) is a representation of a group of packets that share the same requirements for their transport. All packets in such a group are provided the same treatment en route to the destination. As opposed to conventional IP forwarding, in MPLS, the assignment of a particular packet to a particular FEC is done just once, as the packet enters the network. FECs are based on service requirements for a given set of packets or simply for an address prefix. Each LSR builds a table to specify how a packet must be forwarded. This table, called a label information base (LIB), is comprised of FEC–to-label bindings.

Label Creation

There are several methods used in label creation:

- *topology-based method*—uses normal processing of routing protocols (such as OSPF and BGP)
- *request-based method*—uses processing of request-based control traffic (such as RSVP)
- *traffic-based method*—uses the reception of a packet to trigger the assignment and distribution of a label The topology- and request-based methods are examples of control-driven label bindings, while the traffic-based method is an example of data-driven bindings.

Label Distribution

MPLS architecture does not mandate a single method of signaling for label distribution. Existing routing protocols, such as the border gateway protocol (BGP), have been enhanced to piggyback the label information within the contents of the protocol. The RSVP has also been extended to support piggybacked exchange of labels. The Internet Engineering Task Force (IETF) has also defined a new protocol known as the label distribution protocol (LDP) for explicit signaling and management of the label space. Extensions to the base LDP protocol have also been defined to support explicit routing based on QoS and CoS requirements. These extensions are captured in the constraint-based routing (CR)–LDP protocol definition.

A summary of the various schemes for label exchange is as follows:

- *LDP*—maps unicast IP destinations into labels
- *RSVP, CR–LDP*—used for traffic engineering and resource reservation
- *protocol-independent multicast (PIM)*—used for multicast states label mapping
- *BGP—external labels (VPN)*

Label-Switched Paths (LSPs)

-*A collection of MPLS*—enabled devices represents an MPLS domain. Within an MPLS domain, a path is set up for a given packet to travel based on an FEC. The LSP is set up prior to data transmission. MPLS provides the following two options to set up an LSP.

- *hop-by-hop routing*—Each LSR independently selects the next hop for a given FEC. This methodology is similar to that currently used in IP networks. The LSR uses any available routing protocols, such as OSPF, ATM private network-to-network interface (PNNI), etc.
- *explicit routing*—Explicit routing is similar to source routing. The ingress LSR (i.e., the LSR where the data flow to the network first starts) specifies the list of nodes through which the ER–LSP traverses. The path specified could be nonoptimal, as well. Along the path, the resources may be reserved to ensure QoS to the data traffic. This eases traffic engineering throughout the network, and differentiated services can be provided using flows based on policies or network management methods.

The LSP setup for an FEC is unidirectional in nature. The return traffic must take another LSP.

Label Spaces

The labels used by an LSR for FEC–label bindings are categorized as follows:

-*per platform*—The label values are unique across the whole LSR. The labels are allocated from a common pool. No two labels distributed on different interfaces have the same value.

*-per interface***—**The label ranges are associated with interfaces. Multiple label pools are defined for interfaces, and the labels provided on those interfaces are allocated from the separate pools. The label values provided on different interfaces could be the same.

Label Merging
The incoming streams of traffic from different interfaces can be merged together and switched using a common label if they are traversing the network toward the same final destination. This is known as stream merging or aggregation of flows.
If the underlying transport network is an ATM network, LSRs could employ virtual path (VP) or virtual channel (VC) merging. In this scenario, cell interleaving problems, which arise when multiple streams of traffic are merged in the ATM network, need to be avoided.

Label Retention
MPLS defines the treatment for label bindings received from LSRs that are not the next hop for a given FEC. Two modes are defined.
- *conservative***—**In this mode, the bindings between a label and an FEC received from LSRs that are not the next hop for a given FEC are discarded. This mode requires an LSR to maintain fewer labels. This is the recommended mode for ATM–LSRs.
- *liberal***—**In this mode, the bindings between a label and an FEC received from LSRs that are not the next hop for a given FEC are retained. This mode allows for quicker adaptation to topology changes and allows for the switching of traffic to other LSPs in case of changes.

Label Control
MPLS defines modes for distribution of labels to neighboring LSRs.
- *independent*—In this mode, an LSR recognizes a particular FEC and makes the decision to bind a label to the FEC independently to distribute the binding to its peers. The new FECs are recognized whenever new routes become visible to the router.
- *ordered***—**In this mode, an LSR binds a label to a particular FEC if and only if it is the egress router or it has received a label binding for the FEC from its next hop LSR. This mode is recommended for ATM–LSRs.

Signaling Mechanisms
- *label request***—**Using this mechanism, an LSR requests a label from its downstream neighbor so that it can bind to a specific FEC. This mechanism can be employed down the chain of LSRs up until the egress LER (i.e., the point at which the packet exits the MPLS domain).
- *label mapping***—**In response to a label request, a downstream LSR will send a label to the upstream initiator using the label mapping mechanism

Label Distribution Protocol (LDP)
The LDP is a new protocol for the distribution of label binding information to LSRs in an MPLS network. It is used to map FECs to labels, which, in turn, create LSPs. LDP sessions are established between LDP peers in the MPLS network (not necessarily adjacent). The peers exchange the following types of LDP messages:
- *discovery messages***—**announce and maintain the presence of an LSR in a network
- *session messages***—**establish, maintain, and terminate sessions between LDP peers
- *advertisement messages***—**create, change, and delete label mappings for FECs
- *notification messages***—**provide advisory information and signal error information
- *Traffic Engineering*

Traffic engineering is a process that enhances overall network utilization by attempting to create a uniform or differentiated distribution of traffic throughout the network. An important result of this process is the avoidance of congestion on any one path. It is important to note that traffic engineering does not necessarily select the shortest path between two devices. It is possible that, for two packet data flows, the packets may traverse completely different paths even though their originating node and the final destination node are the same. This way, the less-exposed or less-used network segments can be used and differentiated services can be provided.
In MPLS, traffic engineering is inherently provided using explicitly routed paths. The LSPs are created independently, specifying different paths that are based on user-defined policies. However, this may require extensive operator intervention. RSVP and CR–LDP are two possible approaches to supply dynamic traffic engineering and QoS in MPLS.

CR(Constraint-based routing)
Constraint-based routing (CR) takes into account parameters, such as link characteristics (bandwidth, delay, etc.), hop count, and QoS. The LSPs that are established could be CR–LSPs, where the constraints could be explicit hops or QoS requirements. Explicit hops dictate which path is to be taken. QoS requirements dictate which links and queuing or scheduling mechanisms are to be employed for the flow.
When using CR, it is entirely possible that a longer (in terms of cost) but less loaded path is selected. However, while CR increases network utilization, it adds more complexity to routing calculations, as the path selected must satisfy the QoS requirements of the LSP. CR can be used in conjunction with MPLS to set up LSPs. The IETF has defined a CR–LDP component to facilitate constraint-based routes.

6.2.2. On the mobile technology: Edge, 3G, etc.

6.2.3. On the access segment: xDSL, FTTC, FTTP, FTTH, etc.

6.2.4. On the transmission technology: FO, WDM, SDH, Ethernet

6.2.4.1. Ethernet Technologies

Abstract

The term *Ethernet* refers to the family of local-area network (LAN) products covered by the IEEE 802.3 standard that defines what is commonly known as the CSMA/CD protocol. Four data rates are currently defined for operation over optical fiber and twisted-pair cables:

- 10 Mbps—10Base-T Ethernet
- 100 Mbps—Fast Ethernet
- 1000 Mbps—Gigabit Ethernet
- 10-Gigabit Ethernet

The Ethernet Physical Layers

Because Ethernet devices implement only the bottom two layers of the OSI protocol stack, they are typically implemented as network interface cards (NICs) that plug into the host device's motherboard. The different NICs are identified by a three-part product name that is based on the physical layer attributes.

The naming convention is a concatenation of four terms indicating the transmission rate, the transmission method, and the media type/signal encoding. For example, consider this:

- 10Base-T = 10 Mbps, base band, over two twisted-pair cables
- 100Base-T2 = 100 Mbps, base band, over two twisted-pair cables
- 100Base-T4 = 100 Mbps, base band, over four-twisted pair cables
- 1000Base-LX = 100 Mbps, base band, long wavelength over optical fiber cable

10-Giga Ethernet Technology

10 Gigabit Ethernet might well become the technology of choice for enterprise, Metropolitan and wide area networks. In terms of physical media, 10 Gigabit Ethernet will support distances to 300 meters on multimode fiber and 40 km or more on single mode fiber. With 10 Gigabit Ethernet, enterprise network managers and service providers will be able to build LANs, MANs, and
WANs using Ethernet as the end-to-end Layer 2 transport. Long-distance reach on single mode fiber enables enterprise network managers and service providers to build simple, low-cost, metropolitan sized Networks with Layer 3-4 switches and 10 Gigabit Ethernet backbones. In addition, 10 Gigabit
Ethernet will support an optional SONET/SDH-friendly PHY to enable transmission of Ethernet over the SONET/SDH transport infrastructure.
For enterprise LAN applications, 10 Gigabit Ethernet will enable network managers to scale their Ethernet networks from 10 Mbps to 10,000 Mbps, while leveraging their investments in Ethernet as they increase their network performance. For service provider metropolitan and wide-area applications, 10 Gigabit Ethernet will provide high-performance, cost-effective links that are easily managed with Ethernet tools. 10 Gigabit Ethernet matches the speed of the fastest technology on the WAN backbone, OC-192, which runs at approximately 9.5 Gbps. 10-Gigabit Ethernet (IEEE 802.3ae) will define a standard that guarantees interoperation between different vendors' implementations. Essentially, the standard will specify physical layers (PHY);only a very slight change will be made to the medium access control (MAC). A major theme of earlier versions of Ethernet has been the pragmatic adoption of cost effective but robust technologies. In large part, this enabled Ethernet to dominate the LAN market. One of the major challenges addressed by the standards effort has been the development of specifications that are friendly to directly modulated lasers—it is believed this will facilitate very cost effective implementations. It is important to note that 10-Gigabit Ethernet represents the coming together of both data communications and telecommunications. Some of the important features adopted by 10-Gigabit Ethernet are:

• Wide range of cost/reach options

• much longer maximum reach than previous Ethernets

• a four bit wide electrical bus extender (XAUI)

• a very low overhead, scrambler-based,64B66B code

• an option for transport in SONET/SDH like frames

• two serial physical layer types

• a coarse or wide wavelength division multiplexed (WWDM) physical layer

• line rates of 10.3125 (LAN), 9.95328 (WAN, OC-192 rate) and 3.125 (LAN, 4 avelengths) GBd

10-Gigabit Ethernet Standard

The 10-GE standard specifies seven port types as listed in Table 1. Six of the port types use bit serial optical transmission whilst the remaining port type multiplexes MAC data across four wavelengths. The WWDM physical layer can support both multimode and singlemode fiber.

10-Gigabit Ethernet Port Types

As can be seen from Table 6.2.1, two categories of port types are defined:
• LAN PHY for native Ethernet applications
• WAN PHY for connection to the installed base of SDH/SONET 10 Gb/s networks

10-Gigabit Ethernet MAC

Obviously, the normal MAC data rate (the rate at which the MAC transfers its information to the PHY) for 10-Gigabit Ethernet is 10 Gb/s

Description	Name	Comment
85o nm serial LAN PHY	10 GBASE-SR	Directly Modulated VCSEL ,MMF,2-300m
1310 nm serial LAN PHY	10 GBASE-LR	Directly Modulated DFB Laser,SMF,2-10KM
1550 nm serial LAN PHY	GBASE-ER	Modulator , DFB Laser , SMF ,2-40km
1310 nm WWDM LAN PHY	10 GBASE-LX4	Directly Modulated VCSEL ,MMF,2-300m
850 nm serial LAN PHY	10 GBASE-SW	Directly Modulated VCSEL ,MMF,2-300m
1310 nm serial LAN PHY	10 GBASE-LW	Directly Modulated DFB Laser , SMF , 2-10 km
1550 nm serial LAN PHY	10GBASE-EW	Modalator, DFB Laser, SMF, 2-40KM

Table 6.2.1. 10-Gigabit Ethernet Port Types.

Layered model for 10-Gigabit Ethernet

The layered model for 10-Gigabit Ethernet is shown in Fig. 6.2.5. Sublayers for the two families of PHY (LAN and WAN) are included in the diagram.
Also shown are the specified interfaces as follows:

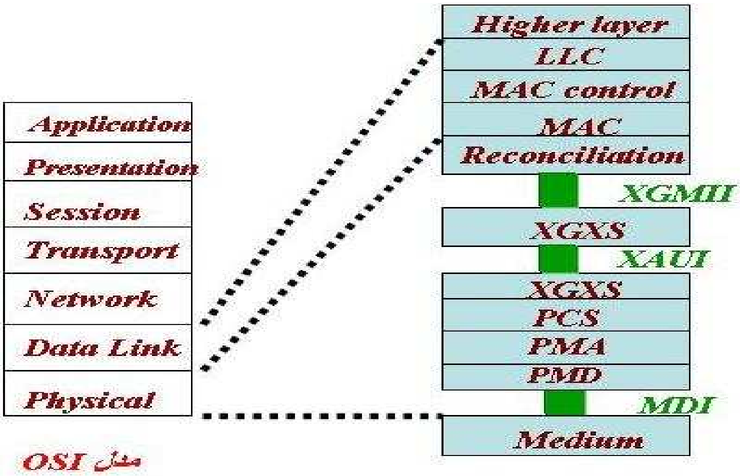


Figure 6.2.5: Layered model for 10 GE

MDI: Medium Dependent Interface
XGMII: 10 Giga bit Media Independent Interface

***XAUI*: 10 Giga bit Attachment Unit Interface
PCS : Physical Coding Sub Layer
XGXS: XGMII Extender Sub layer
PMA: Physical Medium Attachment
PHY: Physical Layer Service
XSBI: 10-Gigabit 16-bit Interface
PMD: Physical Medium Dependent

The reconciliation sublayer (RS) adapts the protocol of the Ethernet MAC into the parallel encoding of the 10 Gb/s PCS.Although the physical implementation of the XGMII is op-tional,for the purposes of specifying 10-Gigabit Ethernet,the XGMII is assumed to be the inter-face between the RS and PCS sublayers.TheXGMII uses 32 bit data paths that are partitioned into four transmit and four receive lanes,8 bits per lane.Also,each lane has a control bit associ-ated with it.The RS maps MAC data octets to (from)the lanes of the XGMII in round-robin order.At the request of the MAC or PHY the RS also maps MAC control signals to (from)the XGMII.Optionally,the transmission distance of the XGMII can be extended using the XGXS and XAUI.Both XGXS and XAUI use the 10GBASE-X (see Fig.6.2.5),PCS and PMA.XAUI associates one of its serial 8B10B lanes,operating at a data rate of 3.125 Gb/s,to each XGMII lane.Essentially,the XGXS and the XAUI interface provide a narrow 4 bit wide,self timed,full duplex,data bus.Repetitive XAUI control signals (for example,idle)are scrambled to prevent exces-sive electromagnetic interference.The optional sixteen bit wide interface between the serial PCS or WIS and the serial PMA is called the 10-Gigabit sixteen-bit interface (XSBI).The XSBI is a fully differential,LVDS,clocked interface that is very similar to the SFI-4 inter-face of the Optical Internetworking Forum (OIF).For specification convenience the standard is written in t rms of the XSBI.

The application's of 10-Gigabit Ethernet

Initially, 10-Gigabit Ethernet will be a switch-to switch interconnection for statistically multiplexing packet traffic from lower data rate (10/100/1000 Mb/s) Ethernets. Therefore, 10-Gigabit Ethernet is primarily a backbone technology that is targeted at the enterprise LAN or the telecom WAN.10 Gigabit Ethernet targets three Application spaces: the LANs (including storage area networks), MANs, and WANs.

LAN Applications

10 Gigabit Ethernet has many potential applications for both service provider and enterprise networks. Figure 6.2.6 shows the standard LAN applications for 10 Gigabit Ethernet, which includes the following:
· Storage area networking (SAN) applications - Server interconnect for clusters of servers.
· Aggregation of multiple 1000BASE-X or 1000BASE-T segments into 10 Gigabit Ethernet downlinks.
· Switch-to-switch links for very high-speed connections between switches in the equipment room, in the same data center, or in different buildings.

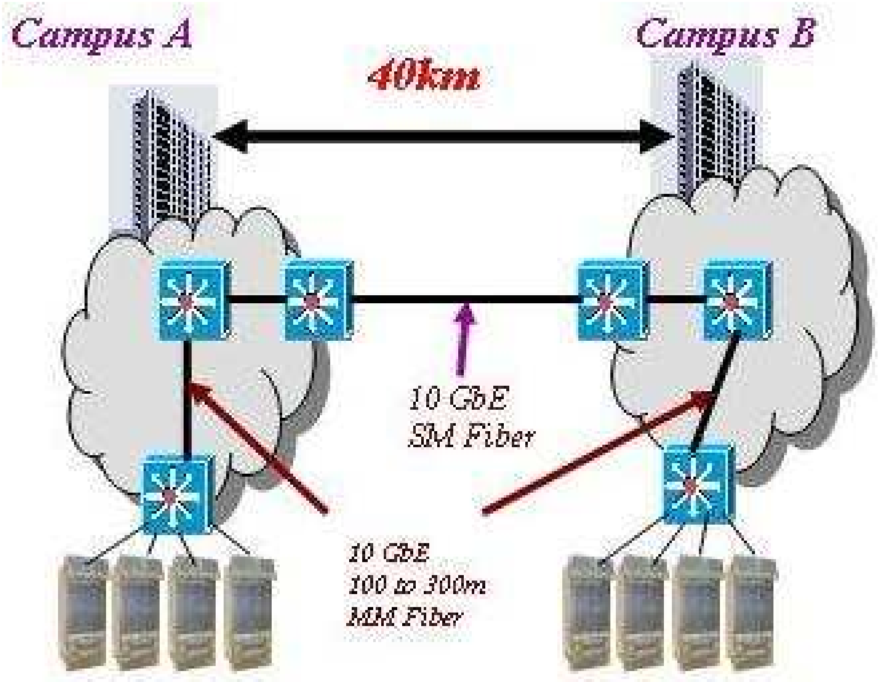


Figure 6.2.6: LAN application

Dark Fiber Metro Applications

One of the most exciting innovations in the Gigabit space has been the growth of the deployment of long distance Gigabit Ethernet using long wavelength optics on dark fiber to build network links that reach metropolitan distances.

10 Gigabit Ethernet, as a fundamental transport for facility services, will be deployed in MAN applications over dark fiber, and over dark wavelengths. The term "dark fiber" refers to unused singlemode fiber capacity from fiber that has been installed for long distance applications that usually reach up to 100 kilometers without amplifiers or optical repeaters. This fiber is not currently "lit," meaning that it is not carrying traffic and is not terminated to equipment. 10 Gigabit Ethernet metropolitan networks will enable service providers to reduce the cost and complexity of their networks while increasing backbone capacity to 10 Gbps.

This will be accomplished by eliminating the need to build out an infrastructure that contains not only several network elements required to run TCP/IP and data traffic, but also the network elements and protocols originally designed to transport voice. Reduction in the number of network elements and network layers lowers equipment costs, lowers operational costs, and simplifies the network architecture. With 10 Gigabit Ethernet backbone networks, service providers will be able to offer native 10/100/1000/10,000 Mbps Ethernet as a public service to customers, namely offering the customer twice the bandwidth of the fastest public MAN services OC-3 (155 Mbps) or OC-12 (622 Mbps) with no need for the added complexity of SONET or ATM, nor protocol conversions.

Dark Wavelength Metro Applications with DWDM

10 Gigabit Ethernet will be a natural fit to the dense wave division multiplexing

(DWDM) equipment, which is deployed for metropolitan area applications. For enterprise networks,access to 10 Gigabit Ethernet services over DWDM will enable serverless buildings, remote backup,and disaster recovery. For service providers, 10 Gigabit Ethernet in the MAN will enable the provisioning of dark wavelength gigabit services at very competitive costs. The terms “dark wavelength” or “dark lambda” refer to unused capacity available on a DWDM system.WDM is a long established technology in the WAN backbone that enables multiple data streams to be transformed into multiple, independent wavelengths. DWDM refers to systems that apply the tight wavelength spacing specified by the International Telecommunications Union (ITU), which is normally less than a nanometer (nm). Coarse or wide wavelength division multiplexing (CWDM or WWDM) refers to less costly optics that use wider spacing between wavelengths. The WDM device then multiplexes the multiple (16, 32, and 64) streams into one stream of “white light” across one fiber pair, increasing the bandwidth capacity of the link by a factor of 16, 32, or 64. At the opposite end, the multiple wavelengths are demultiplexed into the original data streams. Many MANs and much of the WAN backbone today contain installed DWDM equipment that has unused capacity or dark wavelengths.

10 Gigabit Ethernet WAN Applications

WAN applications for 10 Gigabit Ethernet look very similar to MAN applications: dark fiber, dark wavelength, and support for SONET infrastructure. 10 Gigabit Ethernet in WAN application is included multilayer switches and terabit routers attached via 10 Gigabit Ethernet to the SONET optical network, which includes add drop multiplexers (ADMs) and DWDM devices. When dark wavelengths are available, 10 Gigabit Ethernet can be transmitted directly across the optical infrastructure, reaching distances from 70 to 100 km.SONET/SDH is the dominant transport protocol in the WAN backbone today, and most MAN public services are offered as SONET OC-3 (155 Mbps) or OC-12 (622 Mbps). Most of today’s installed optical infrastructure is built out with a specific architecture and specific timing requirements to support OC-192 SONET. To make use of the SONET infrastructure, the IEEE 802.3ae Task Force specified a 10 Gigabit Ethernet interface (WAN PHY) that attaches to the SONET-based TDM access equipment at a data rate compatible with the payload rate of OC-192c/SDH VC-4-64c. This is accomplished by means of a physical layer link based on the WAN PHY between Gigabit or Terabit switches and Ethernet line-terminating equipment (LTE), which is attached to the SONET network. The WAN PHY interface does not attach directly to a SONET OC-192 interface.The WAN PHY interface will allow the construction of MANs and WANs that connect geographically dispersed LANs between campuses or POPs through the SONET transport network. In other words, 10 Gigabit Ethernet interfaces that are compatible with SONET OC-192 payload rate facilitate the transport of native Ethernet packets across the WAN transport network, with no need for protocol conversion. Reducing the need for protocol conversion increases the performance of the network,makes it simpler and easier to manage, and less costly, because protocol conversion is CPU intensive,adding complexity and additional elements to the network.

10 Gbps Ethernet Physical Layer Specifications

The 10 Gigabit Ethernet physical layer specifications, referred to as the “PHY”, provides the network manager and cabling distribution designer with the basic information required to select the appropriate optical transceiver types based on their network distance requirements, cabling performance, and types of network connections. The 10 Gigabit Ethernet Standard defines two unique physical layer specifications associated with the types of network connections: the LAN physical layer (LAN PHY) and the WAN physical layer (WAN PHY).

The physical layer (PHY) contains the types of transmitters and receivers and the functions that translate the data into signals (encoding), which are compatible with the cabling type used. The encoding function is performed in the physical coding layer (PCS) of the PHY. The LAN PHYs use 64B/66B encoded data; the WAN PHYs implement an encapsulation of the 64B/66B encoded data for compatibility with OC-192c/SDH VC-4-64c. The encapsulation is performed in the wide area network interface sublayer (WIS).

In Ethernet-speak, the transceiver types, which are cabling media dependent, are referred to as the physical media dependent (PMD) types. Examples of Ethernet optical fiber PMD types are 10BASE-F (10 Mbps), 100BASE-FX (100 Mbps), and 1000BASE-SX (1000 Mbps).

The 10 Gigabit Ethernet PMDs include both serial and wavelength division multiplexing (WDM) fiber optic transceiver types.

10 Gigabit Ethernet operating distances are specified for both multimode and single mode fiber. The minimum operating distance for each option is associated with the targeted operating environment,i.e., LAN/MAN/WAN.

6.2.4.2. Next Generation SDH

Introduction

There is no doubt that most of the present transmission networks are based on SDH/SONET technology. Although the demand growth for higher bit rates and the increase in traffic growth in communication networks have caused the introduction of WDM/DWDM technology so it can cover most the demands, but still traffic aggregation is continuously done by SDH/SONET systems. Even in networks with more than 500Gb/s capacity, almost 90% of the traffic is aggregated on STM-16 interfaces (Fig.6.2.7).

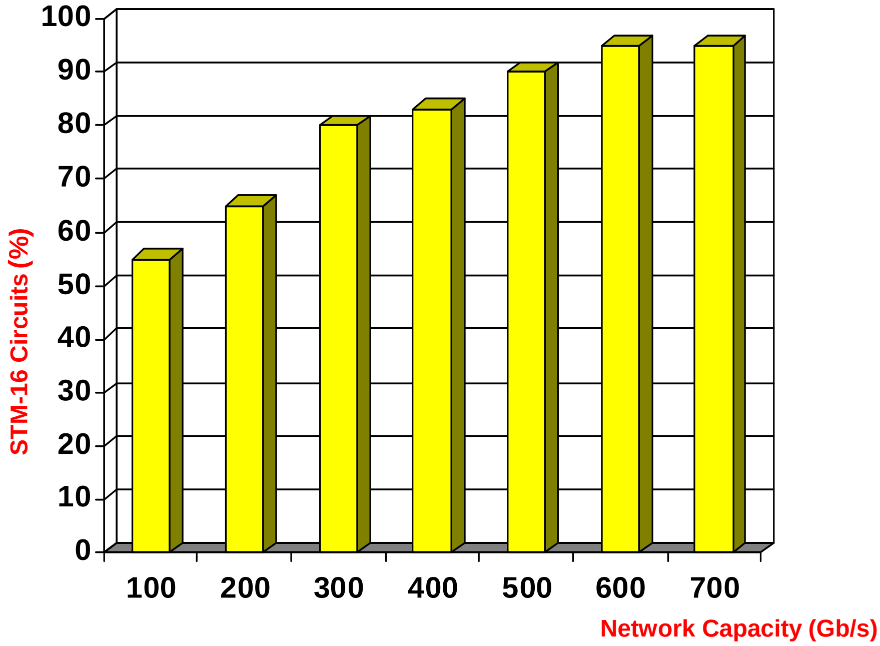


Fig .6.2.7-Fraction of circuits STM-16 of aggregated network capacity

In addition, most of the present and future service's request from the transmission networks has been forecasted over STM-1 and STM-4 interfaces. This shows the importance of SDH/SONET systems in the future.

On the other hand, one of the important challenges in developing transmission networks is the growth of data traffic compared with voice traffic. This challenge has become evident when most of data traffic produced by subscribers is done via Ethernet protocol. Ethernet optimizes traffic transport, in other words the role of SDH systems in future transmission networks has been adapted with Ethernet. At the present time, the transport of Ethernet over SDH systems has the following problems:

- Lack of flexibility

The defined bit rates for Ethernet are 10Mb/s, 100Mb/s, 1Gb/s and 10Gb/s and for SDH the bit rates are 155Mb/s, 622Mb/s, 2.5Gb/s and 10Gb/s.

If Ethernet traffic has carried on SDH system, part of the capacity will become useless. For example (Fig.6.2.8), a fast Ethernet service (100Mb/s) should be carried via STM-1 interface, which means that 30% of the transmission network capacity is lost.

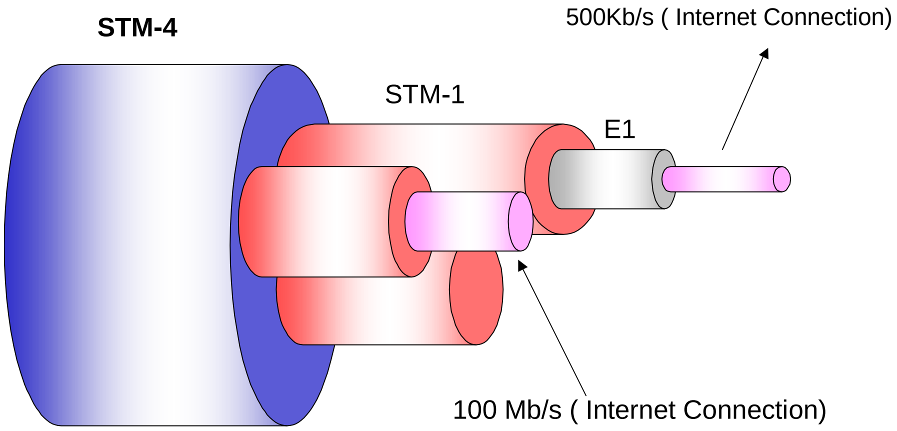


Fig.6.2.8-Lack of flexibility

- Low band with utilization over a ring

In an SDH ring, to transport certain traffic on part of the ring between two nodes, the needed capacity will occupy the whole ring. This causes low utilization of bandwidth (Fig.6.2.9).

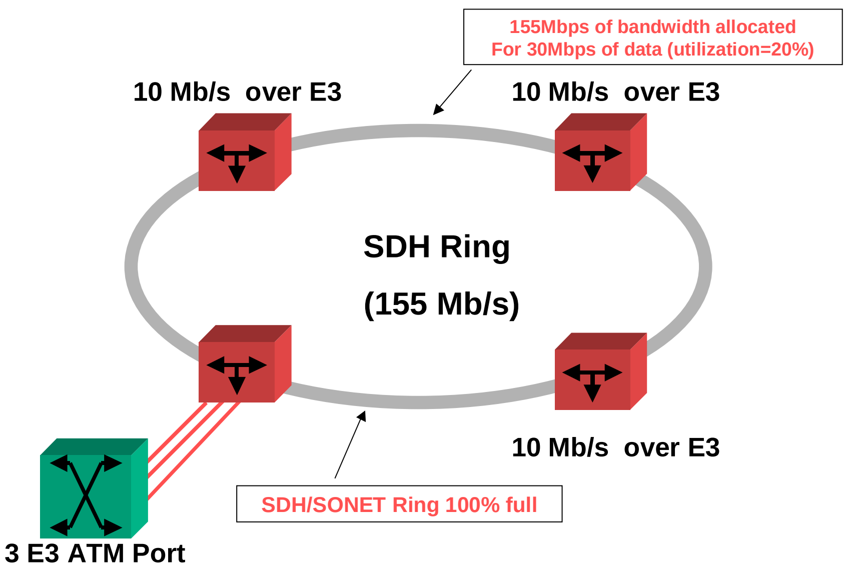


Fig. 6.2.9 - Low bandwidth utilization

- Variety in platforms

A variety in services comes from various technologies and hardware. In other words, new services in central offices add a new set of hardware on a rack that results in increase of auxiliary costs.

<u>Next Generation SDH</u>

The above-mentioned problems introduce a new aspect to solve them, which is called “Next Generation SDH”. The development and deployment of SDH system offers new services using new protocols such as RPR, GFP, VC and LCAS.
Changing services landscape shows that Ethernet services have its importance in the future. The evolved Ethernet offers higher bit rates (10Mb/s, 100Mb/s, 1Gb/s, 10Gb/s), but speed is only part of the story. Ethernet still has its own disadvantages:

Poor facility utilization
No edge to edge Qos
Slow services restoration
Standard capacity (spanning tree)
In addition, in a transmission network some transparent disadvantages appear in transport circuits and the number of nodes over SONET/SDH ring is limited up to 16 nodes.

Ethernet architecture

- Meshed Ethernet

In a meshed Ethernet architecture, STP (Spanning Tree Protocol) limits protection response to minutes, and Qos is applied at each hop.

- Ethernet Rings

In Ethernet ring, there is a limitation in the number of nodes and protection schemes. Also, packets are processed at each hop and no high-priority transit allowed. The attempt to solve the above problems or offer Ethernet services on present ring architectures has resulted in emergence of RPR (Resilient Packet Rings).

- RPR

RPR is a new media access control protocol based on a ring topology that has developed by 1EEE802.17 working group. The goal behind it is to provide an efficient use of network bandwidth and a resilient network with < 50 ms recovery time. Also, it supports up to 128 nodes in a ring.
In RPR, packets take the shortest path to the destination. The entire ring belongs to one subnet; this reduces many of the inter-subnet issues. RPR supports multicasting such that targeted multicast group nodes will copy the multicast source and pass them through the ring to the next node. Also, using RPR, the multicast source nodes will remove the multicast packets.
The resilience or proactive span protection automatically avoids failed spans within 50 ms. RPR supports both latency/jitter sensitive traffic such as voice & video services and committed information rate (CIR) services. Another RPR feature is the high efficiency that

comes from spatial reuse. Unlike SDH/SONET, bandwidth is consumed only between the source and destination nodes. Packets are removed at their destination, leaving this bandwidth available to down stream nodes on the ring. RPR supports topologies of more than 100 nodes per ring and automatic topology discovery mechanism.

Packet over SDH

At the present time, transports of Ethernet traffic over SDH via HDLC protocols are not quiet efficient. The main protocols that encapsulate and frame data traffic for transport over next generation SDH infrastructure are as follows:

GFP (Generic Framing Procedure)

GFP is a traffic adaptation protocols, which is an ITU-T standard (ITU-T G704I).
It enables mapping of any data type to a SONET/SDH byte-synchronous channel. The GFP recommendation defines procedures for transporting various variable length client frames over legacy SDH transport and enables multiplexing of different client signals on to a single transport.

VC (Virtual Concatenation)

This protocol allows, instead of making a continuous allocation for a client signal, the transport path is created using a concatenation of smaller transport channels with a defined capacity like STM-1. VC compensates differential network delays up to 32ms. Only termination nodes need to support this feature.

LCAS (Link Capacity Adjustment Scheme)

It is a method to modify VCG size at the end points of transport path by using a specific signaling procedure. The signaling messages are transported in H4 byte. LCAS controls hitless addition/ removal of STM-N's (VC-n's) to/from VCG under management control. It also addresses the dynamic management of bandwidth for data transport services over SONET/SDH. LCAS works best on point-to-point links and supports virtual channel protection through "load sharing" on STM-n's.

TrafficType	SONET		SDH	
	Contiguous	Virtual	Contiguous	Virtual
10Mbps Ethernet	STS-1 (20%)	VT-1.5-7v (89%)	VC-3 (20%)	VC-12-5v(92%)
1Gbps Fibre Channel	STS-21c (85%)	STS-1-18v (95%)	VC-4-16c (35%)	VC-4-6v(95%)
100Mbit/s FastEthernet	STS-3c (67%)	STS-1-2v (100%)	VC-4 (67%)	VC-3-2v (100%) VC-12-46v 100%)
200Mbit/s (ESCON)	STS-6c (66%)	STS-1-4v (100%)	VC-4-4c (33%)	VC-3-4v (100%) VC-4-2v (66%)
1Gbit/s Ethernet	STS-24c (83%)	STS-1-21v (92%)	VC-4-16c (42%)	VC-4-7v (95%)

Tabele.6.2.2 GFP, Virtual Concatenation & LCAS Transport Efficiency

New architecture of SDH/SONET platform

As mentioned so far, rapid change in services result in deployment of various systems and hardware at central offices. This made the suppliers to offer new platforms for (ATM, IP, Ethernet services) that are capable of producing various signals in the transmission network (STM-n or λ).

These architectures are as follows:

Single switch architecture

In this case a single fabric switch is defined on a system and interface adaptation is used over input/output card.

Hybrid multi-switch Architecture

In this case I/O traffic directed to a fabric switch based on service type.

Multi layer SONET/SDH architecture

In this case ATM, IP switching is identified on I/O cards. This architecture identifies the cross connect and add/drop function over one platform and introduces new topology for new SDH metropolitan networks.

Conclusion

In the evolution of networks, the significant principle is to save and utilize the existing infrastructure. Fiber ring network is the most deployed network and the infrastructure should adapt with evolved trends. Traffic patterns are changing from voice towards data, and networks based on SONET/SDH inherent planned and implemented for transport of voice traffic (or circuit service).Saving and utilization of SONET/SDH network depends on the change of transport capabilities in data traffic.

Next generation SDH includes RPR implementation, utilization of VC, GFP, LCAS protocols, and same platforms can play effective role in development and operation of SONET/SDH networks.

6.2.5. On the Radio technologies: TDMA, CDMA, WI-FI, etc.

6.2.6. On the service and applications platforms

6.3. NGN solutions and migration steps

6.3.1. NGN concepts definition and NEs (reference ITU-T Recommendation Y 2001 (12/2004)

The concept of an NGN (Next Generation Network) has been introduced to take into consideration the new realities in the telecommunications industry, characterized by factors such as: competition among operators due to ongoing deregulation of markets, explosion of digital traffic, e.g., increasing use of "the Internet", increasing demand for new multimedia services, increasing demand for a general mobility, convergence of networks and services, etc. The NGN (Next Generation Network) is conceived as a concrete implementation of the GII (Global Information Infrastructure).

The target of NGN is to ensure that all elements required for interoperability and network capabilities support applications globally across the NGN while maintaining the concept of separation between transport, services and applications.

A major goal of the NGN is to facilitate convergence of networks and convergence of services. The common understanding is that the NGN has to be seen as the concrete realization of concepts defined for the GII .

<u>Definitions</u>

- **Next Generation Network (NGN): A packet-based network able to provide telecommunication services and able to make use of multiple broadband, QoS-enabled transport technologies and in which service-related functions are independent from underlying transport-related technologies. It enables unfettered access for users to networks and to competing service providers and/or services of their choice. It supports generalized mobility which will allow consistent and ubiquitous provision of services to users.**

- **Generalized mobility: The ability for the user or other mobile entities to communicate and access services irrespective of changes of the location or technical environment. The degree of service availability may depend on several factors including the Access Network capabilities, service level agreements between the user's home network and the visited network (if applicable), etc. Mobility includes the ability of telecommunication with or without service continuity.**

<u>Objectives of the NGN</u>

NGN should fulfil the requirement of the environment described in ITU-T Recs Y.100 [1], Y.110 [2], Y.130 [3] and Y.140 [4] or Y.140.1 [5], for example to:

- **promote fair competition;**
- **encourage private investment;**
- **define a framework for architecture and capabilities to be able to meet various regulatory requirements;**
- **provide open access to networks;**

while:

- **ensuring universal provision of and access to services;**
- **promoting equality of opportunity to the citizen;**
- **promoting diversity of content, including cultural and linguistic diversity;**
- **recognizing the necessity of worldwide cooperation with particular attention to less developed countries.**

Fundamental characteristics of NGN

The term NGN as defined in clause 3 is commonly used to give a name to the changes to the service provision infrastructures that have already started in the telecommunication industry. The NGN can be further defined by the following fundamental characteristics:

- **packet-based transfer;**
- **separation of control functions among bearer capabilities, call/session, and application/ service;**
- **decoupling of service provision from transport, and provision of open interfaces;**
- **support for a wide range of services, applications and mechanisms based on service building blocks (including real time/ streaming/ non-real time and multimedia services);**
- **broadband capabilities with end-to-end QoS (Quality of Service);**
- **interworking with legacy networks via open interfaces;**
- **generalized mobility (see 3.2 and 8.7);**
- **unrestricted access by users to different service providers;**
- **a variety of identification schemes;**
- **unified service characteristics for the same service as perceived by the user;**
- **converged services between fixed/mobile;**

- **independence of service-related functions from underlying transport technologies;**

- **support of multiple last mile technologies;**

- **compliant with all regulatory requirements, for example concerning emergency communications, security, privacy, lawful interception, etc.**

<u>NGN Capabilities</u>

NGN shall provide the capabilities (infrastructure, protocols, etc.) to make the creation, deployment and management of all kinds of services (known or not yet known) possible.

One of the main characteristics of NGN is the decoupling of services and transport, allowing them to be offered separately and to evolve independently.Therefore in the NGN architectures, there shall be a clear separation between the functions for the services and the functions for the transport. NGN allows the provisioning of both existing and new services independently of the network and the access type used.

Interworking between NGNs of different operators and between NGN and existing networks such as PSTN (Public Switched Telephone Network), ISDN (Integrated Services Digital Network) and GSM (Global System for Mobile communications) is provided by means of gateways.

NGN will support both existing and "NGN aware" end terminal devices. Hence terminals connected to NGN will include analogue telephone sets, fax machines, ISDN sets, cellular mobile phones, GPRS (General Packet Radio Service) terminal devices, SIP [13] (Session Initiation Protocol) terminals, Ethernet phones through PCs (Personal Computers), digital set top boxes, cable modems, etc.

Specific issues include the migration of voice services to the NGN infrastructure, Quality of Service related to real-time voice services (with guaranteed bandwidth, guaranteed delay, guaranteed packet loss, etc.) as well as Security. NGN should provide the security mechanisms to protect the exchange of sensitive information over its infrastructure, to protect against the fraudulent use of the services provided by the Service Providers and to protect its own infrastructure from outside attacks.

A major feature of NGN will be generalized mobility, which will allow a consistent provision of services to a user, i.e., the user will be regarded as a unique entity when utilizing different access technologies, regardless of their types. Generalized mobility means providing the ability of using different access technologies, at different locations while the user and/or the terminal equipment itself may be in movement allowing users to use and manage consistently their applications/customer services across existing network boundaries.

At present mobility is used in a limited sense such as movement of user and terminal and with or without service continuity to similar public accessed networks (such as WLAN, GSM, UMTS, etc.) and service discontinuity to some wired line accessed networks with

strong limitations (such as UPT). In the future, mobility will be offered in a broader sense where users may have the ability to use more access technologies, allowing movement between public wired access points and public wireless access points of various technologies. This means that this movement will not necessarily force an interruption of an application in use or a customer service.

Terms and Definitions

- **Media Server (MS):** A network element providing the media resource processing function for telecommunication services in NGN.
- **AG:** Access Gateway, that allows end users with various accesses (e.g., PSTN, ISDN, V5.x) connection to the packet node of
- **Remote User Access Module (RUAM**): A unit that physically terminates subscriber lines and converts the analogue signals into a digital format. The RUAM is physically remote from the Local Exchange.
- **User Access Module (UAM):** A unit that physically terminates subscriber lines and converts the analogue signals into a digital format. The UAM is collocated with a Local Exchange, and is connected to the Local Exchange.
- **Trunking Media Gateway (TMG):** A unit that provides interfaces between the packet node of NGN and the circuit-switched nodes (e.g. transit exchange, local exchange, international exchange) of PSTN/ISDN for bearer traffic
- **Signalling Gateway (SG):** A unit that provides out-of-band call control signalling conversion between the NGN and other networks (e.g., between a call server in NGN and an STP or SSP in SS7)

Abbreviations and acronyms

- This draft uses the following abbreviations.
- ACS Access Call Server
- AG Access Gateway
- AS Application Server
- AN Access Network
- ATM Asynchronous Transfer Mode
- BCS Breakout Call Server
- BICC Bearer Independent Call Control
- BTV Broadband TV
- CAS Channel Associated Signalling
- CBR Constant Bit Rate
- CC Content of Communication
- CCS Common Channel Signalling
- CDR Call Detail Record
- CPE Customer Premise Equipment
- CS Call Server
- DSL Digital Subscriber Line
- DSLAM Digital Subscriber Line Access Multiplexer
- ETS Emergency Telecommunications Services
- FTTC Fibre-To-The-Curb
- FTTH Fibre-To-The-Home

- GCS Gateway Call Server
- GoS Grade of Service
- ICS Interworking Call Server
- IMS IP Multimedia Subsystem
- IN Intelligent Network
- INAP Intelligent Network Application Part

- Access Gateways

- Allows the connection of subscriber lines to the packet network

- Converts the traffic flows of analogue access (Pots) or 2 Mb/s access devices into packets

- Provides subscriber access to NGN network and services

- Trunking Media Gateways

- Allows inter-working between classical TDM telephony network and Packet-based NGN networks,

- Converts TDM circuits/ trunks (64kbps) flows into data packets, and vice versa

- Packet networks for NGN

- Trend is to use IP networks over various transport possibilities (ATM, SDH, WDM…)

- IP networks for NGN must offer guarantees of Quality of Service (QoS) regarding the real time characteristics of voice

- Softswitch/MGC

- Referred to as the Call Agent or Media Gateway Controller (MGC).

- Provides the “service delivery control” within the network

- In charge of Call Control and handling of Media Gateways control (Access and/or Trunking) via H.248 protocol

- Performs signalling gateway functionality or uses a signalling gateway for interworking with PSTN N7 signalling network

- Provides connection to Intelligent Network /applications servers to offer the same services as those available to TDM subscribers

- H.248 Protocol

- Known also as MEGACO: standard protocol, defined by ITU-T, for signalling and session management needed during a communication between a media gateway, and the media gateway controller managing it

- H.248/MEGACO allows setting up, keeping, and terminating calls between multiple endpoints as between telephone subscribers using the TDM

6.3.2. Examples of NGN solutions and migration steps

Example of Network migration from today existing TDM towards full NGN (each solution depending on each situation) :

Fig 6.1 : Step1. Network consolidation and optimization at topological level

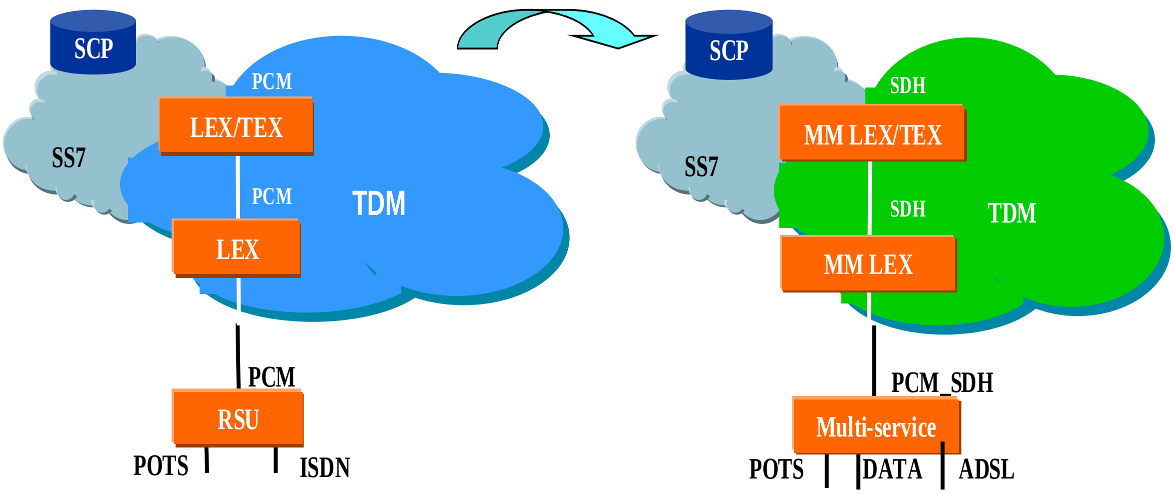


Fig 6.2 : Step2. Network migration at C4 first and C5 with service compatibility

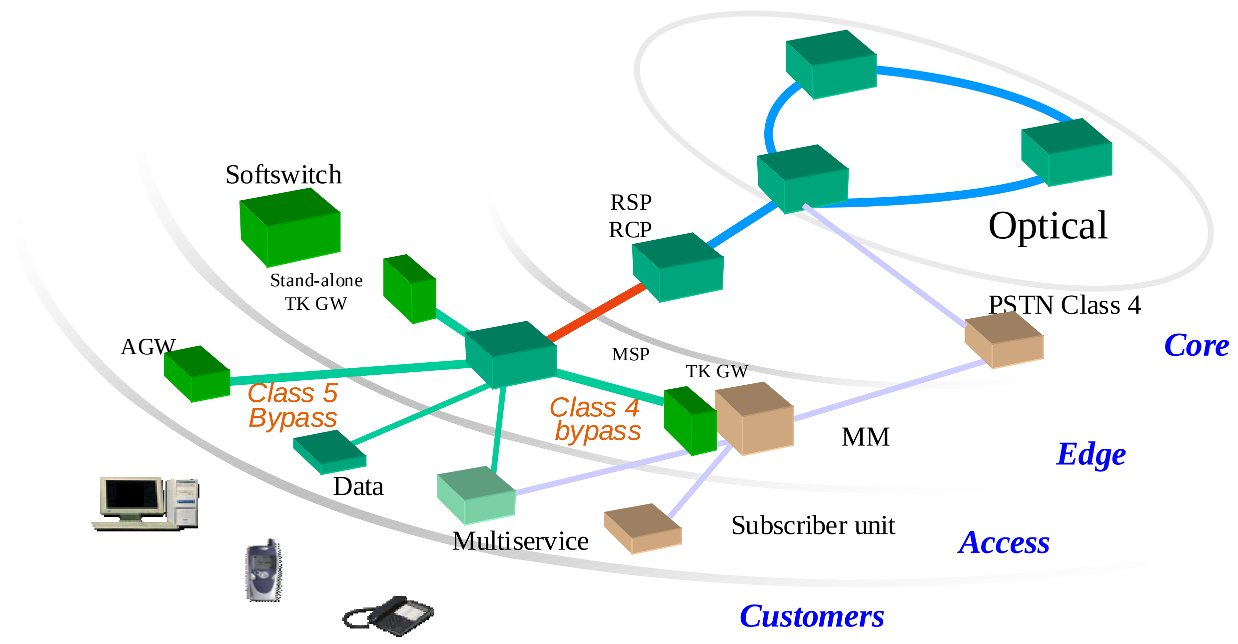
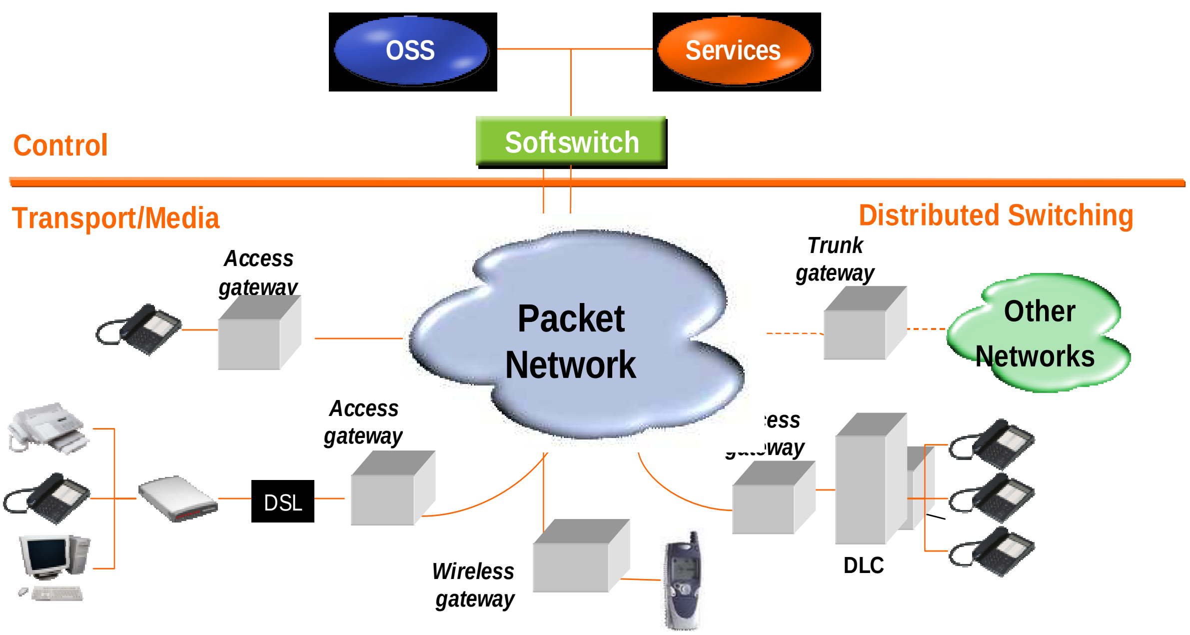


Fig 6.3 : Step3. Converged Network at all layers



This is a purpose for step by step migration from a TDM-based public switched telephone network (PSTN) to a packet-based next generation network, from an economic point of view, it identifies the drivers and benefits for an established carrier to consolidate its current network and to migrate to an NGN from a technology point of view.

In a more detaild manner the following migration phases could be defined:

- Phase 1 (current PSTN network):

The starting point for the migration to NGN is today's public switched telephone network. In TDM and SS7 network all voice traffic is transported over TDM, and controlled by a hierarchy of local (Class 5) and transit (Class 4) circuit switches. The voice-related signaling network (ISUP and INAP) is handled by the SS7signaling network. Value-added services are provided inside the switches, or through the intelligent network (IN). Widely spread IN services include calling card services, number translation and routing services (such as free phone, premium rate and universal access number), and enterprise network services such as virtual private networks (VPNs). With the growing number of Internet users, carriers are providing connectivity to Internet service providers (ISP) either through narrowband (PSTN or ISDN) dialup services, or through introduction of broadband ADSL (with voice split off as a separate service).

- Phase 2:

Network infrastructure optimization will reduce carriers' operational expenses and allow them to generate additional revenues. Deployment of a small number of large exchanges (local and transit) with increased switching capacity, and high speed interfaces (SDH, ATM) reduces the operator's and enables faster deployment of new services. "Redundant" switches may be converted to additional remote access concentrators.
Introduction of new technology with smaller footprint, or packet fabrics inside the exchanges, allows the carrier to reduce expenses and reuse the switching infrastructure for new data services. Adding new access nodes and upgrading the existing ones lets the carrier capitalize on his PSTN, while extending the coverage area and the bandwidth offered to individual subscribers (fiber closer to the end user). New access technology provides seamless multiservice access to voice (POTS, ISDN) and data (ADSL, ATM, IP, FR, etc.) services and paves the way to NGNs. Optimization of the ADSL access infrastructure is realized through introduction of voice over DSL (VoDSL) loop-emulation services (inverse gateway, with a V5.2/GR303 connection to the LEX).

- Phase 3:

As one of the basic goals of NGN introduction is to move to a unique, packet-based infrastructure voice transport will smoothly migrate to IP (or ATM) technology. Initially, carriers will focus on trunking scenarios to offload long-distance voice from their TDM network. The first step toward VoIP migration is extending the existing local exchanges with integrated trunking gateways (TGW) for converting TDM voice into packets (IP). This approach guarantees full protection of TDM investments, while providing the operator with a full fledged trunking-over-packet solution, as well as continued access to switch based and IN-based value-added services. In order to address existing switches without integration of a gateway, external trunking gateways, controlled by a Class 4 softswitch (through the H.248 or Megaco protocol), may be added. From a functional point of view, the softswitch performs like a Class 4 (Toll/Transit) exchange, with similar features (e.g., screening and routing), signaling interfaces (ISUP, INAP) and access to value-added services (IN).

- Phase 4:

In fast growing and deployment of broadband access (ADSL, LMDS, cable) operators may introduce voice-over-packet technology to capture growth in the access network, or as a means to offload the local exchanges from DSL. The Class 5 softswitch with local features (e.g., CLASS, custom calling) will be a shared control element, but several alternatives for voice gateways (depending on end user topology, density, service requirements, etc.) may be deployed. Just as in the Class 4 case, the softswitch will address the gateways using the H.248 or Megaco protocol. ADSL subscribers may install a residential gateway (RGW) or integrated access device (IAD) with VoIP coding capability. Contrary to the ADSL with split-off voice or VoDSL loop emulation solutions, the RGW provides the broadband user with end-to-end voice-over-packet. As an alternative to upgrading the CPE of its subscribers, an ADSL operator may choose to extend the DSLAMs with VoIP gateway functionality. Another solution for connecting voice subscribers directly to the data network is to introduce new access gateways [AGW] or to upgrade the existing access nodes with AGW functionality. In order to address new generation voice terminals (IP phones), the Class 5 softswitch can also terminate emerging user-to-network signaling protocols such as H.323 and SIP.
[It is difficult to differentiate with Phase 4 clearly except use Class 5 softswitch. Is this major message of this pahse?]

- Phase 5:

There is no doubt that, in the near (and even midterm) future, voice will be the predominant service, even in NGNs. The introduction of broadband access in the network, however, enables the deployment of a new range of data and multimedia services. These new services will allow carriers to differentiate and compete with new entrants. A prerequisite for the deployment of multimedia services is the general availability of appropriate terminals. Today's personal computers are a good starting point, but it is expected that the convergence of computer, consumer and communications technology will result in a number of new multimedia devices. These new terminals will communicate with the softswitch through emerging multimedia signaling protocols such as H.323 and SIP. In order to fully support the new network and terminal capabilities, the softswitch is extended with mixed-media session and QoS control. With the introduction of new business models and new players (e.g., virtual network operators, third party application providers, content providers), there is a need for application access (for authentication, authorization, accounting, roaming, subscriber profiles, etc.) and service brokering platforms (terminal capabilities negotiation, bandwidth brokering, content aggregation, etc.).
Such portals not only provide the network operator with new business opportunities as a service retailer, but also clearly separate network control from services functionality. In a full-fledged NGN architecture, applications and network will interface through standardized protocols (e.g., SIP) and APIs. It is even assumed that voice services offered on VoP

networks will have fewer features than the ones on circuit networks (especially in an H.323 environment).

- Phase 6:

As a final migration step toward the full NGN, the remaining legacy PSTN equipment is transformed to or replaced by NGN ‘compliant’ network components. The aim of this ultimate (though optional) transformation, is to capitalize on access concentrators connected to local exchanges while further reducing the packet-only network for transport and signaling. At the end of their life, remaining TDM exchanges and access nodes are gracefully transformed to or replaced by trunking gateways, access gateways and softswitches as outlined in the previous sections. While keeping the upper layers (SCCP, ISUP, TCAP, INAP), the lower layers of the SS7 signaling network are replaced by a packet-based equivalent, as defined by the IETF.

The "migration steps" towards NGN could be presented in a macroscopic way, in which the migration steps implies major changes in topology/architecture (giving 3 major migration steps) or in a more detaild way as the second one, where are considered also applications (as phase 5) which may be parallel to the others.

According to the level of detail and considering all underlying technologies and layers we may also define many more steps/phases.

Main role for the planner is to define those "migration steps" in a coordinated manner for each scenario and country ensuring the objectives of the business plan for the corresponding Services, Capacity, Quality and Economy.

- **Evolution of PSTN/ISDN to NGN : Reference ITU-T Recommendation Y.2261 - PSTN/ISDN evolution to NGN**

NGN (Next Generation Network) is believed to provide new opportunities for and capabilities to the network and service providers. Considering that existing networks have different life span and vast amount of capital has been spent on them, complete replacement of their components is not considered to be either advisable or possible. So, a phased approach should be considered for evolution of existing networks to NGN.

PSTN/ISDN Public Switched Telephone Network/Integrated Services Digital Network) being one of the first networks, is considered to be prime candidate for evolution. For PSTN/ISDN evolution to NGN a phased approach is considered

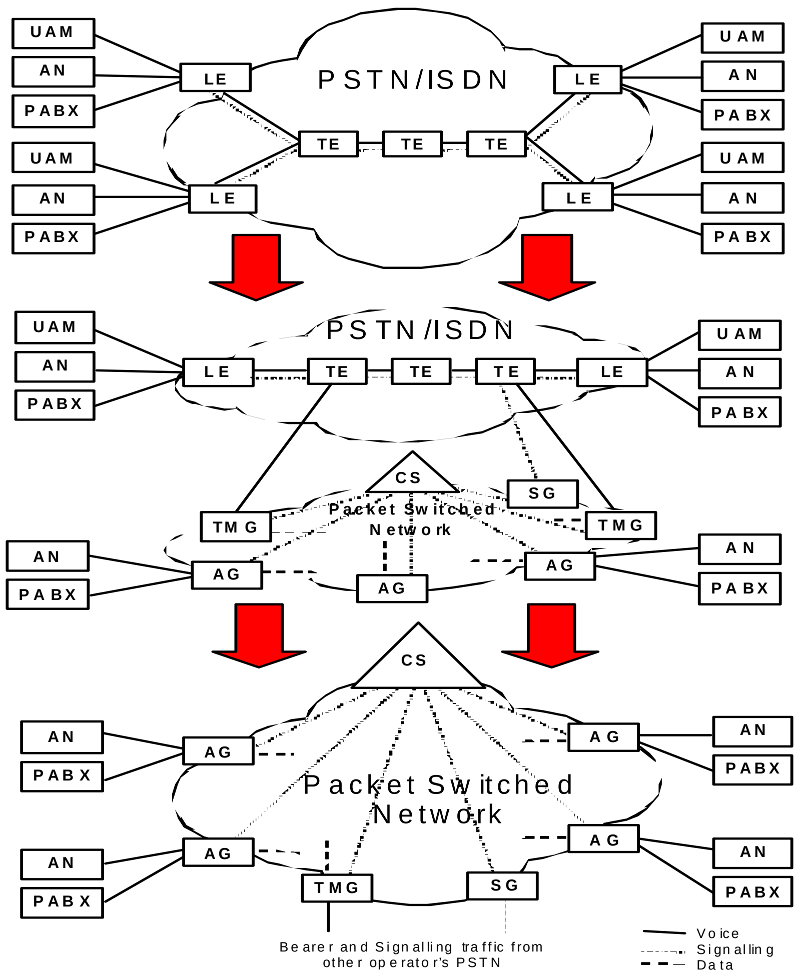
ITU-T Recommendations Y.2261, describes possible ways of evolving PSTN/ISDN to NGN. Both IP multi-media sub-system (IMS-based) and call server (CS-based) are described. It describes aspects, which need to be considered including evolution of transport, management, signalling and control parts of PSTN/ISDN to NGN. Examples of evolution scenarios are also provided in this Recommendation.

Examples: Core Network evolution to NGN

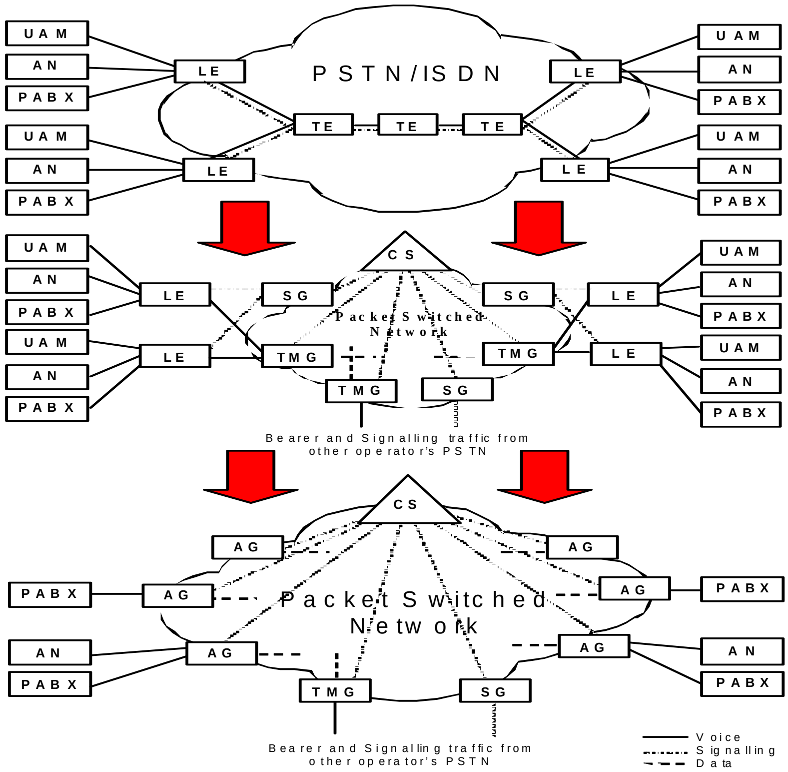
a) CS-based evolution to NGN

The call server (CS) is the core element for PSTN/ISDN emulation. It is responsible for call control, gateway control, media resource control, routing, user profile and subscriber authentication, authorization and accounting. The Call Server may provide PSTN/ISDN basic service and supplementary services, and may provide value added services through service interaction with an external service control point (SCP) and/or application server (AS) in the service/application layer.

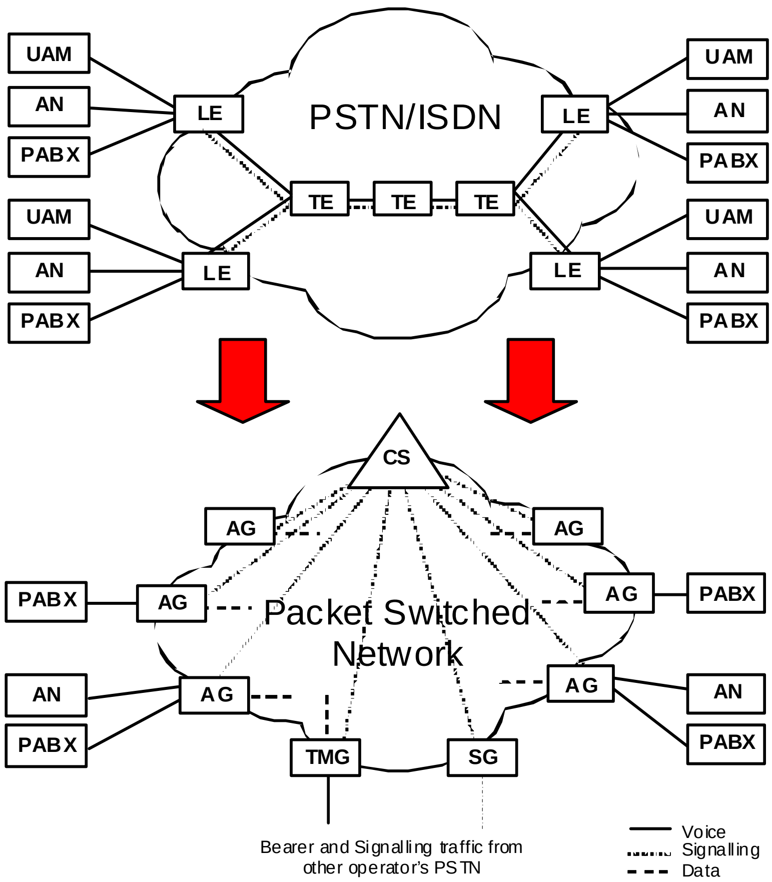
Scenario 1 – Co-existence of PSTN/ISDN and PSN



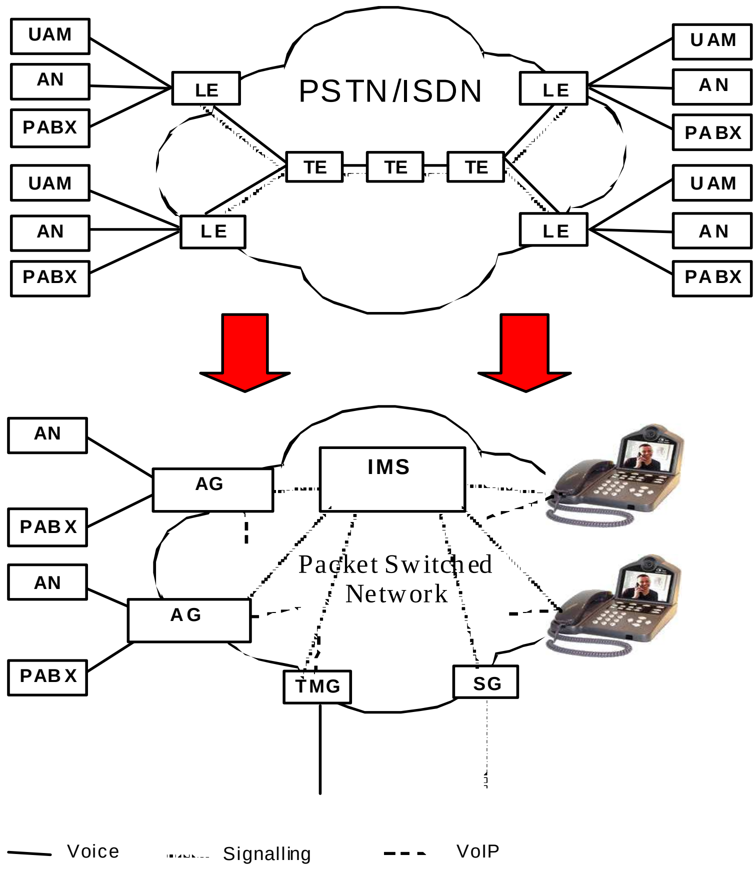
Scenario 2 – Immediate use of PSN, initially via SGs and TMGs



Scenario 3 – one-step approach

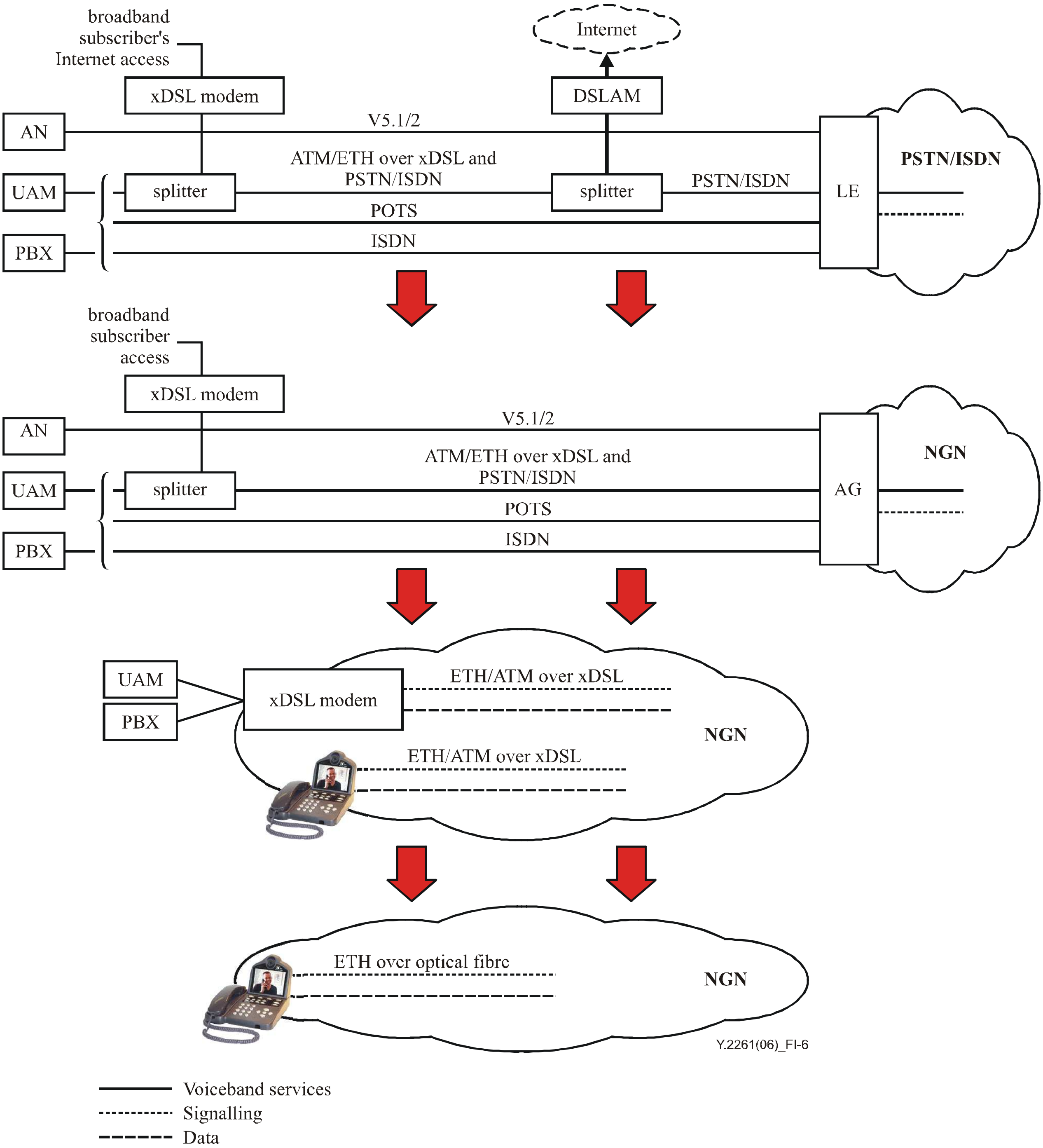


b) IMS-based evolution to NGN



Examples: Access Network evolution

Evolution of xDSL access to NGN



NOTE – Data indicate packetized voiceband services.

6.4. Converged Networks

Convergence at telecom networks is driven by a number of motivations for the customers, the operators and the equipment suppliers. From the overall economical perspective for all of them, the major advantage of convergence is the economical savings due to the important economies of scale by:

- Larger systems capacities and sizes are cheaper per unit and new convergent technologies facilitate those larger capacities

- Higher traffic efficiency of bigger groups due to better system utilization for a given GoS

- Higher density of customers per geographical area will produce infrastructure savings with a quadratic rule

- Higher volumes of purchasing imply significative volume related discounts

From the service provider's point of view:

- Higher utilization of the installed infrastructure capacity
- Better customer's retention by providing multiple service types
- Better utilization of capital when modernizing the network
- Coordinated management and operation for the multiple domains

From the equipment provider's industry point of view:

- Lower complexity due to the alignment of requirements for fixed, mobile, nomadic, etc.
- Harmonized solutions avoiding multiple parallel developments

From the customer's point of view:

- Utilization of multiple services with same terminals at different domains
- Personalization of user profiles across domains
- Service usage simplification for subscription, billing, etc.
- Accessibility to new multimedia services

From the industry and operators initiatives, a number of developments and implementations are being developed that form important pillars for the network convergence. The IP multimedia system, the Fix Mobile Convergence and Mobile Broadcasting convergence are fundamental for the convergence.

6.4.1 IMS architecture for convergence

The IP Multimedia Subsystem (IMS) standard defines a generic architecture for offering multimedia services with generic applications for many technologies and access systems like: GSM, WCDMA, CDMA2000, Wire line broadband access, WiMAX, etc.. It is an international recognized standard, first specified by the Third Generation Partnership Project (3GPP/3GPP2) and involving key actors like operators, equipment suppliers and standard organizations like ETSI/TISPAN, ITU-T, ATIS, and IETF. Protocols initially developed for mobile applications are extended and generalized for fixed networks implying effort and time saving for all multimedia services.

Specific functionalities to be provided include:

- Delivery of person-to-person real-time IP-based multimedia communications, Person-to-person, person-to-machine

- Integration of real-time with non-real-time multimedia communications like live streaming and chat

- Interaction of functionalities for different services and applications like combined use of presence and instant messaging

- Enabling easy user setup of multiple services in a single session, or multiple synchronized sessions

- Facilitation of better control by operators of service value chain and end-to-end QoS.

Extensions of original IMS in the FMC for a wider NGN include:

- The control of IP Connectivity Access Networks (QoS, admission control, authentication, etc.);

- The co-ordination of multiple control components to a single core transport for resource control;

- The inter-working and interoperability with legacy and other networks;

- Mutual de-coupling of the applications from the session/call control and the transport;

- Access technology independence of session/call control and applications.

From the supplier's design and operator's point of view, IMS expands the layered architecture by defining a horizontal architecture, where service enablers and common functions for many services can be reused for multiple applications. The horizontal architecture in IMS provides bearer control and also specifies interoperability, roaming, charging and security. These functionalities positions IMS as a fundamental enabler for fixed-mobile convergence and for the other convergence dimensions, including the ones at the IT domain.

ITU-T have defined, in the ITU-T SG13, a global structure and functionality for an NGN transport and services network that has that functional blocks, interfaces and interrelation flows as indicated in the figure below. Standardization of those functional blocks and interfaces will be a facilitator for the reusability, third party developments and exploiting the associated economies of scale.

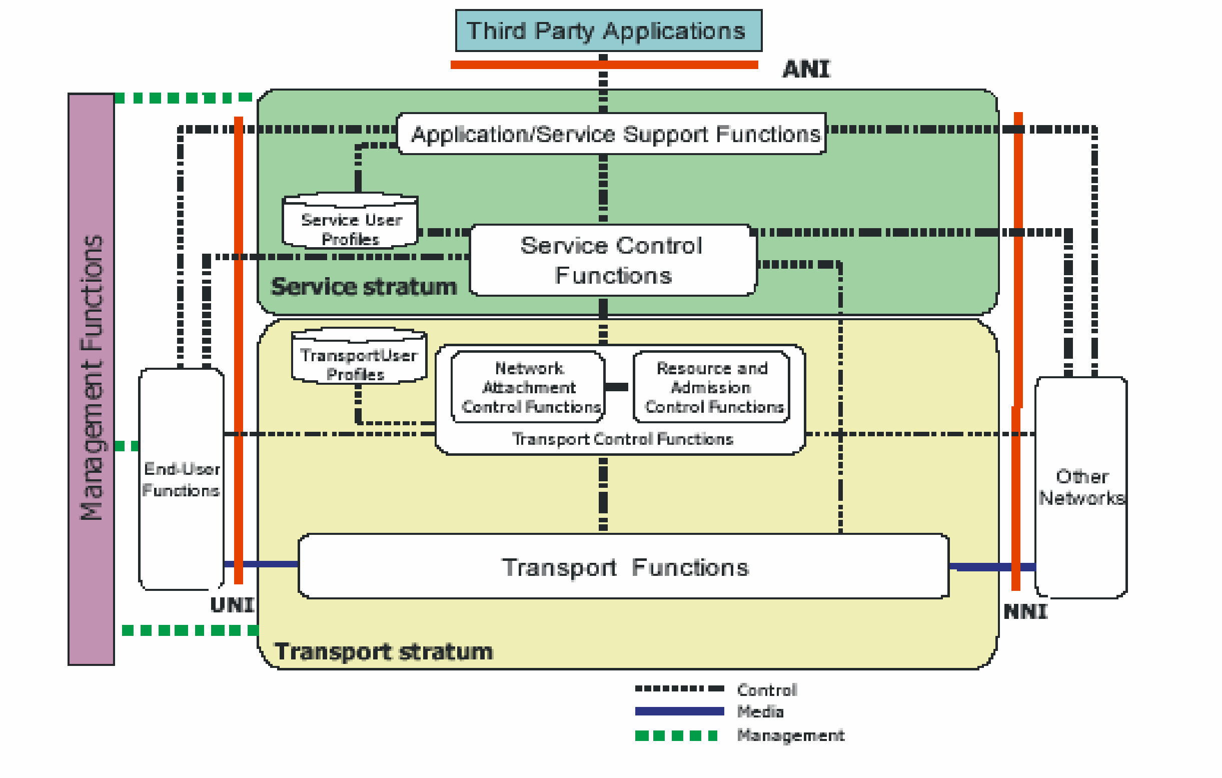


Fig: 6.4.1 Global structure and functional blocks for NGN transport and services by the ITU-T Recommendation Y.2011

From the IT point of view and with more level of detail, the layered architecture of the IMS is organized in three levels as follows:

- The "application layer" that comprises application and content servers to execute the value-added services for the user. IMS defines 3 types of Application Servers:
    - SIP Application Server
    - OSA/Parlay Service Capability Server
    - Application Server for IN-like services

- The "control layer" comprises network control servers for managing call or session set-up, modification and release. The most important of these is the CSCF (Call Session Control Function), also known as a SIP server. This layer also contains a full suite of support functions, such as provisioning, charging and operation & management (O&M). Interworking with other operators' networks and or other types of networks is handled by border gateways.

- The "media-connectivity layer" comprises media conversion and protocol adaptation for the different gateways, routers and switches, at the different network segments either backbone, local or access.

Diagram below illustrates the most typical resources and functionalities related to the IMS:

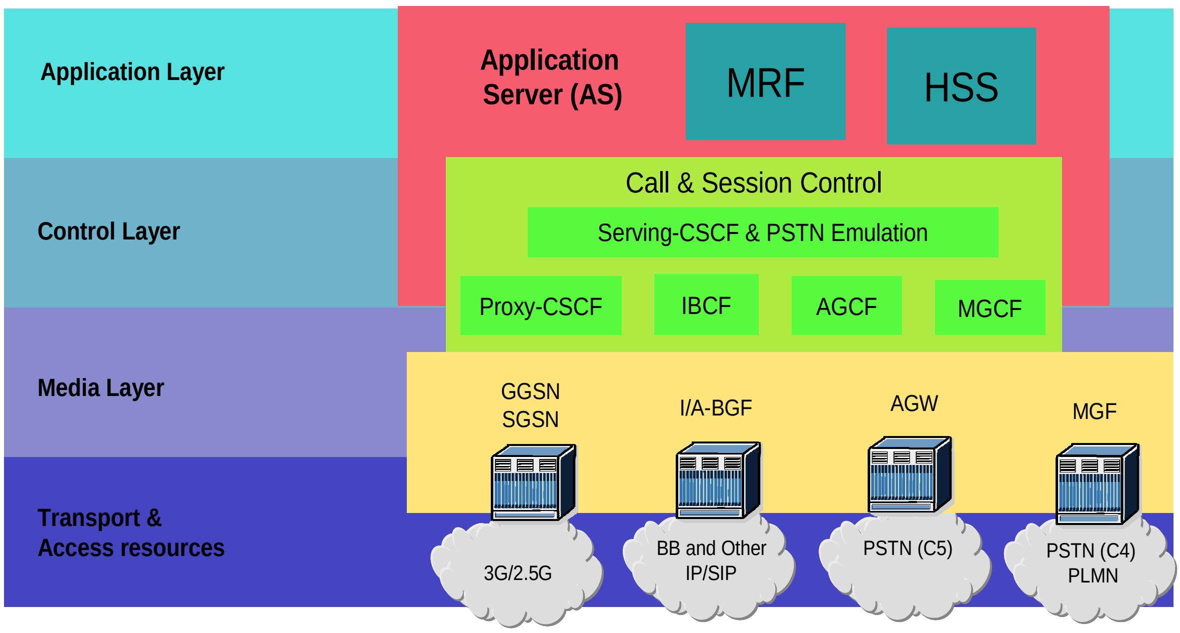


Fig: 6.4.2 Example of typical structure of IMS for the most common functions

- Application Server (AS), implements the value-added services

- Home Subscriber Server (HSS) with a unique service profile for each user and the AAA functionality.

- Multimedia resource function (MRF), which controls media stream resources

- The "S-CSCF & PSTN emulation" (Serving CSCF) is the serving call state session control function for IMS and PSTN/ISDN simulation system subscribers. Its main function is to control the session states

- The Proxy-CSCF supports roaming and nomadism by supporting SIP-based registration of terminals from different accesses

- Interconnecting border control function (IBCF) converts signalling protocols when inter-working with other application service provider networks is necessary. It also provides the function of information hiding and call filtering to prevent illicit use of the network.

- Border gate function (BGF) is a packet-to-packet gateway for access (A) and for interconnection (I)

- Access gateway control function (AGCF) converts SIP signalling from the S-CSCF to MGCP/H.248 towards an RGW or AGW

- Media gateway control function (MGCF) provides inter-working between ISUP and other PSTN NNI protocols and SIP used inside the SP network

6.4.2 Fixed Mobile convergence

FMC is essential for seamless delivery of services and allows two key telecom industries to share experiences. However, FMC must not create “disruptive dynamics” but cater for an evolutionary path taking into account competitive environment FMC is essential for high level “vision” to help resolve societal, business and standardization issues of NGN in a multi-operator environment.
Main topics here are the integration at the level of Applications and Interfaces to the NGN at IP mode.

Fixed Mobile Convergence (FMC) is concerned with the provision of network capabilities which are independent of the access technique. This does not imply necessarily the physical convergence of networks. It is concerned with the development of converged network architecture and supporting standards. This set of standards may be used to offer fixed, mobile or hybrid services.

An important feature of fixed mobile convergence is the separation of the subscriptions and services from individual access points and terminals and to allow users to access a consistent set of services from any fixed or mobile terminal.

An extension of this principle is related to inter-network roaming; users should be able to roam between different networks and to be able to use the same consistent set of services through those visited networks. Both, levels of mobility and variety of networks to be integrated in the FMC are illustrated in the figures below:

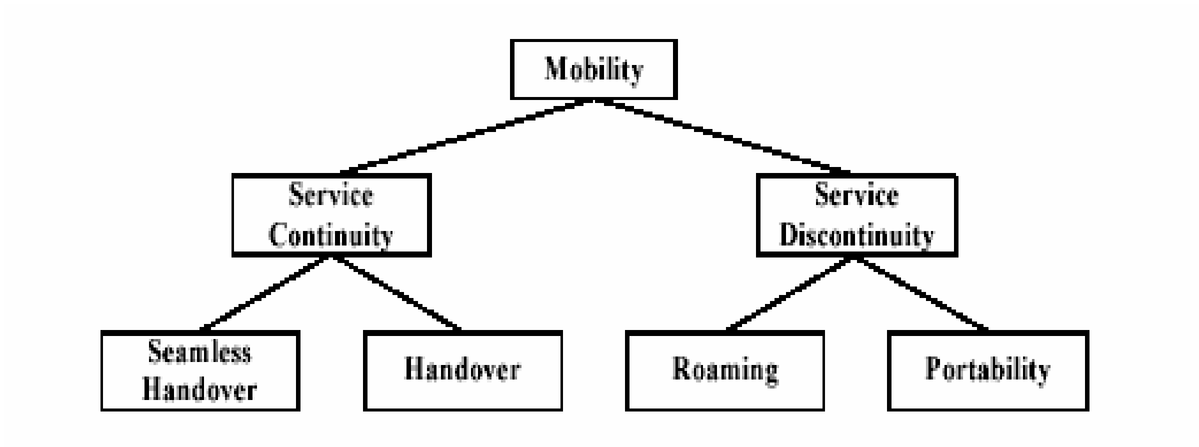


Fig: 6.4.3. Different levels of mobility to be managed in convergence

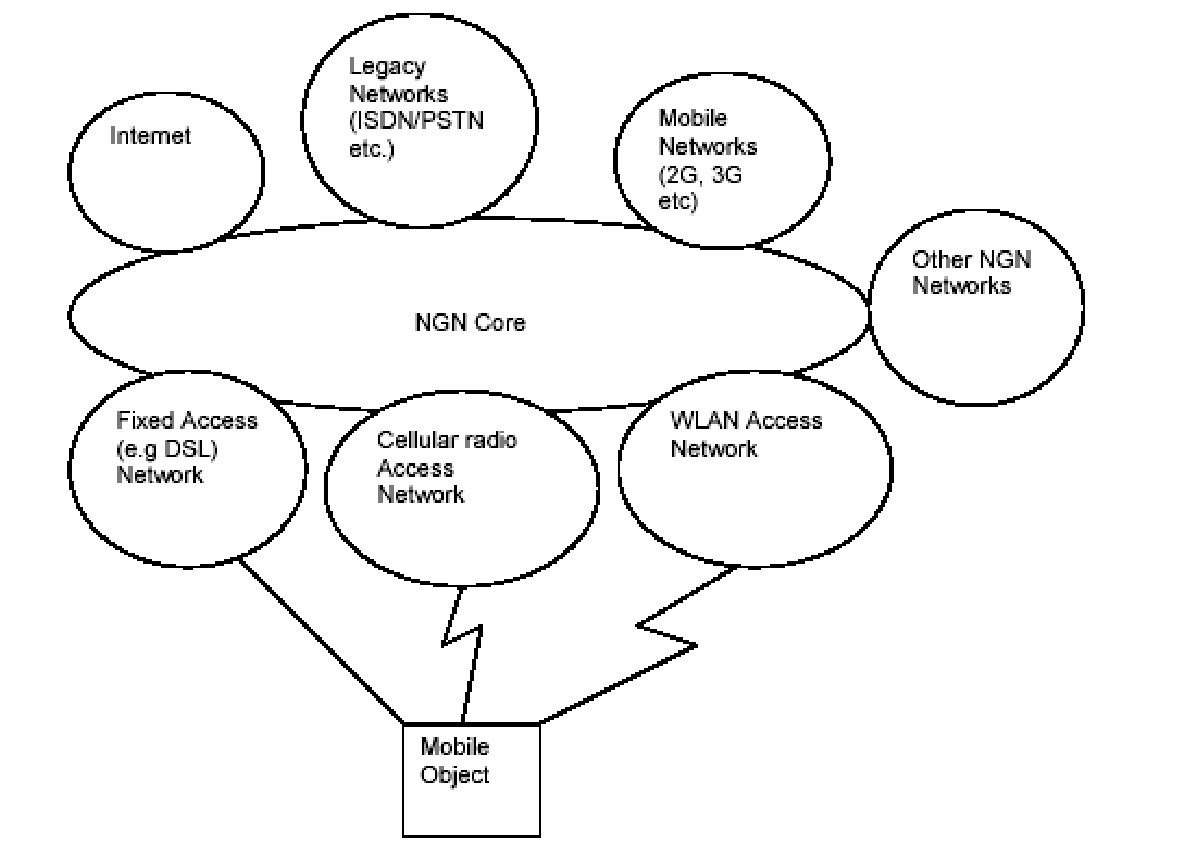


Fig: 6.4.4. Variety of network types coexisting in the FMC

In addition to the generic driving factors for all convergence solutions related to the economies of scale, specifically for the FMC additional facilitators are the user interest in service ubiquity, availability of mobile facilities also in the fixed terminals, single billing, same handset, etc. The fact of high penetration of mobile and wireless services is another factor that motivates a common trend towards the convergence.

Convergence level from the current status should evolve according to the standardization process at the level of applications, services and interfaces. An initial stage is expected at the first releases of IMS definition including SIP for control session, IMS authentication, security, charging and QoS. At further stages more functionalities will be standardized like the inter-working with no-IMS networks, IMS Group Management, IMS conferencing, Lawful Interception, emergency calls, etc.

The implementation of convergent applications at the network resources will be a function of the initial status of the network modernization: Some applications will be implemented at the Softswitches (SSW) in those networks were the SSW were implemented before the standardization of the IMS functionalities while at a later stage and when convergence takes place at the same time than evolution towards NGN, most applications will be implemented at the IMS itself. Due to the lifecycle of the different types of network resources, it is expected a quicker convergence at the application and services layers than at the overall network infrastructure and physical level.

6.4.3 Broadcasting convergence

The evolution of broadcasting technologies towards digital and the generalization of multimedia services in all domains, paves the way for another convergence of services between the mobile solutions and the broadcast solutions. Several benefits will be obtained from the synergies and complementarities of both alternatives.

In addition to the generic convergence and economy of scale advantages given at the convergence chapter, in a more specific manner for broadband services that require the availability of spectrum, sharing of spectrum and related network resources will allow optimization of media and higher capacities.

- Mobile networks are characterized by the bidirectional one-to-one communication with full mobility, on-demand call establishment and billed as a function of utilization. Content in 3G and further versions evolves to Broadband applications reach in video applications were many ones related to TV channels, films, events, etc. coincide with the distributive contents.

- Broadcast networks are characterized by the one-to-many unidirectional communication, restricted mobility, high capacity content and time independent low cost. When evolving to digital (DVB), content is compatible with other media and expand the number of distribution alternatives, but with the requirement of an additional return channel to enable interactive services.

- Mentioned evolution from both sides: incorporating video related applications in mobile and incorporating interactivity in broadcasting prepares the path for a convergence in which “synergies” and “complementarities” of the two alternatives may collaborate for a better service provision, capacity increase and savings by the derived economies of scale.

- The diagram below illustrates the structure of a cooperating platform between a broadcast network operator and a mobile network operator as proposed by the Digital Video Broadcast project that is an industry-led consortium of over 270 broadcasters, manufacturers, network operators, software developers and regulatory bodies. TM-CBMS subgroup on Technical Module for Convergence Broadcast Mobile Services specific located for convergence. Digital Video Broadcast Handheld (DVB-H) technology as an enhancement of the DVB-T for terrestrial digital TV is used to transmit video content to hybrid 3G-DVB terminals that use one or other media as a function of availability and efficiency according to the content type. DVB-T is IP based and allows handheld portable and mobile reception with low consumption and mobility at high data rates.

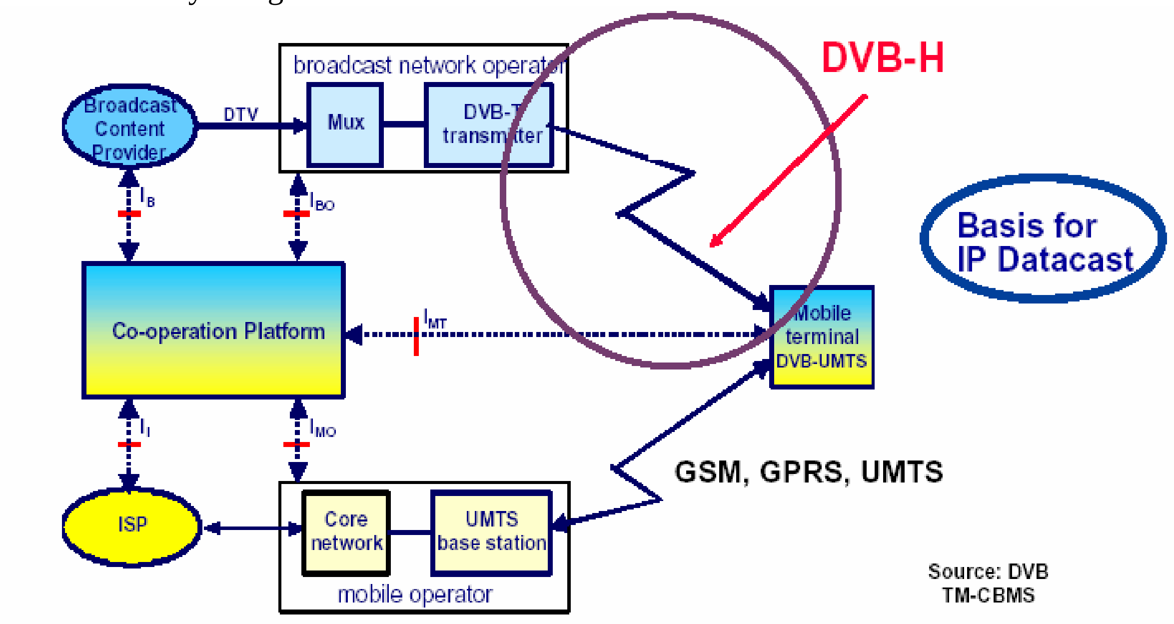


Fig: 6.4.5. Example of functional architecture for hybrid mobile and DVB-H networks

- Convergence in this case between mobile and broadcast will imply a set of benefits as indicated:

- Customers will increase the capacity by the use of downstream heavy traffic flows and decrease cost by the utilization of best media according to application type.

- Mobile operators will increase the set of services and content, extend the business field, unload the 3G networks of heavy downstream flows, optimize the investments on infrastructures and enhance positioning to attract customers in a competitive environment.

- Broadcast operators also benefit from the extension of new services, level of interactivity, new sources of revenue and accessibility to a wide population of customers derived from the ones of the mobile operators.

6.4.4. IMS development in NGN and benefits

Currently, several trials and initial deployments are being implemented for the transition from PSTN towards IMS and specific extensions of IMS are being developed for NGN either to serve the provision of full SIP-based multimedia services to NGN terminals or the provision of PSTN/ISDN simulation services for existing legacy technologies that still will stay during all transition period. Among those extensions we have the following ones:

- The control of IP Connectivity Access Networks (QoS, admission control, authentication, etc.)
- The co-ordination of multiple control components to a single core transport for resource control
- The interworking and interoperability with legacy and other networks
- Mutual de-coupling of the applications from the session/call control and the transport
- Access technology independence of session/call control and applications.

6.4.4.1 Functionalities

Functional entities of an IMS may be used by an operator in support of network scenarios in the transition phases. For instance, the routing may be performed based on signalling information, configuration data, and/or data base lookup as a function of the traffic type and the entity being used.

IMS, as defined within ITU-T, is comprised of a number of functional entities that together can provide support for the capabilities of the service stratum of NGN as described in section 6.4.1 above. The current IMS functional entities and their environment are illustrated in the following figure with a short description of main ones afterwards:

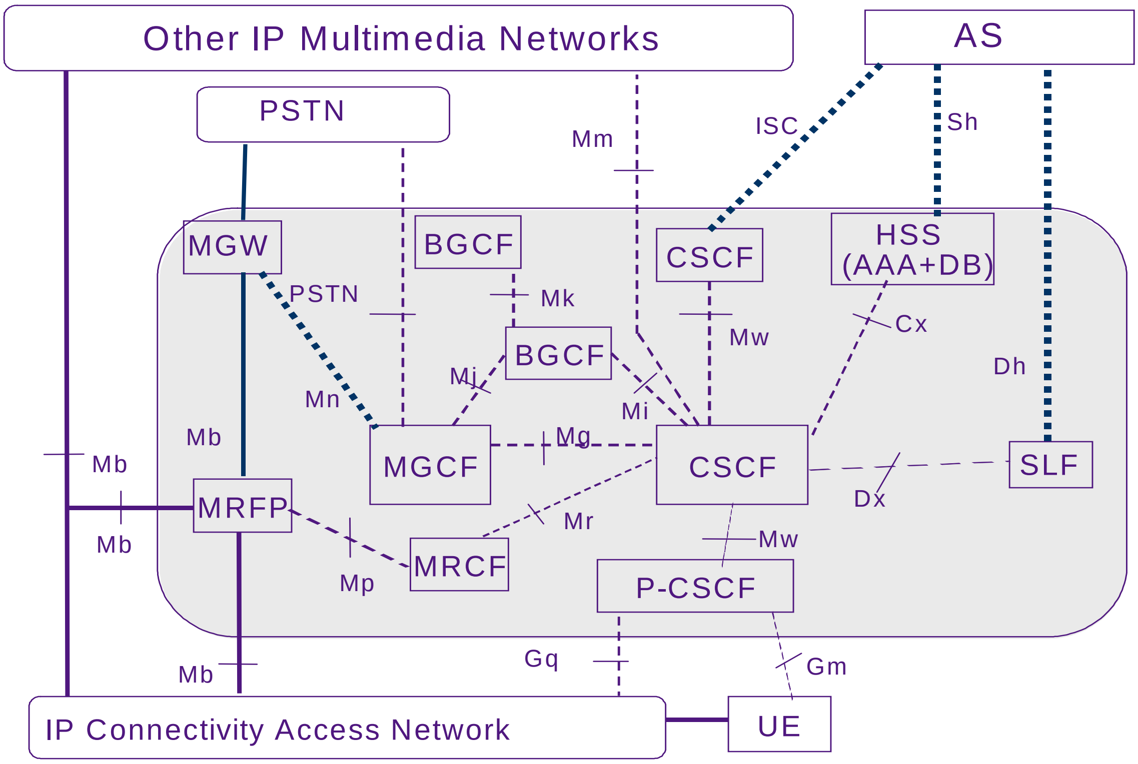


Figure 6.4.4.1: Current logical functionalities and interfaces of the IMS

- *Call Session Control Function (CSCF)*

The Call Session Control Function (CSCF) establishes, monitors, supports and releases multimedia sessions and manages the user's service interactions. The P-CSCF or Proxy CSCF is the first contact point within the IMS that forwards the SIP messages received from the User Equipment (UE). The CSCF interfaces through SIP with the Application Servers (AS) that host and execute the services.

- *Media Gateway Control Function (MGCF)*

The Media Gateway Controller Function (MGCF) provides the ability to control a trunking media gateway functional entity through a standardized interface. Such control includes allocation and deal location of resources of the media gateway, as well as modification of the usage of these resources. The MGCF communicates with the CSCF, the BGCF, circuit-switched networks and performs protocol conversion between ISUP and SIP. It also supports interworking between SIP and non-call related SS7 signalling (i.e. TCAP-based signalling for supplementary services such as Call Completion Busy Subscriber).
In case of incoming calls from legacy networks, the MGCF determines the next hop in IP routing depending on received signalling information. In case of transit the MGCF may use necessary functionality for routing transit traffic. A node implementing this functional entity in an NGN network and a node implementing it in a 3GPP network may differ in terms of supported resources (e.g. codecs) and configuration.

- *Multimedia Resource Function Controller (MRFC)*

The Multimedia Resource Function Controller (MRFC), in conjunction with a Media Resource Processing Functional Entity (MRP-FE) located in the transport layer, provides a set of resources within the core network for supporting services. The MRFC, in conjunction with the MRP-FE, provides e.g., multi-way conference bridges, announcement playback, media transcoding.

- *Multimedia Resource Function Processor (MRFP)*

The Multimedia Resource Function Processor (MRFP) processes the mixes of media streams and transcoding when interworking is required under the control of the MRFC.

- *Breakout Gateway Control Function (BGCF)*

The Breakout Gateway control function (BGCF) selects the network in which PSTN breakout is to occur and - within the network where the breakout is to occur - selects the MGCF. In case of transit the BGCF may have extra functionality for routing transit traffic.

- *Subscriber locator Function (SLF)*

The Subscriber Locator Function (SLF) identifies a user's Home Subscriber Server when multiple HSSs are being used and each one maintains a unique collection of users.

6.4.4.2 Convergence to IMS and phasing

A common IMS for all services of mobile and fixed networks is an ambitious target with disruptive implementations that imply a phasing approach from the current network status. The following issues have to be solved by the planner:

- How IMS functionalities will impact the network architecture?
- What services will be the first to be implemented with IMS?
- Which will be the impact of the IMS deployment on the network flows and load?
- How the functions of IMS will be distributed through the network?
- At what phases and speed will the services be implemented at IMS?
- What will be the benefits for the service provider and for the customer of an IMS based solution?

Due to the initial stages of IMS implementations, a phased approach is required that will take, as in all network transitions, several years. It has to be taken into account that availability of a core NGN IP based network is a prerequisite for an IMS solution and an end to end all IP needed for a fully fledged IMS solution. The following diagram illustrates a feasible scheme for migrating from current networks towards a fully based IMS case.

- Open Service architecture and a basic Home Subscriber Server (HSS) are mandatory from the starting process in the so called Pre-IMS or Early-IMS that will implement the easiest services for the coexisting networks in fixed and mobile technologies without the need of the user entity supporting IPv6. PSTN emulation and simulation are used in order to extend services to all customers either connected to the full IP mode NGN or conventional TDM.

Services expected to have priority in this phase includeVoIP, Location Based Services (LBS), Presence Based Services (PBS), Instant Messaging (IM) and Push to Speak PtS).

- An ambitious fully fledged IMS solution requires, in addition, a more complete end to end NGN infrastructure, complete coverage of SIP with a full functionality of the Application Servers and HSS. User entities need to support IPv6 and the corresponding facilities of the IPsec are exploited. Services expected at this stage include Peer to Peer video (P2P), Service Broker (SBr) function to manage interactions among applications, Service Blending (SBle) for services grouping and personalization, Resource Acceptance Control Function (RACF) to ensure QoS with a common network policy for resource management across network subsystems and Intelligent Content Delivery (ICD).

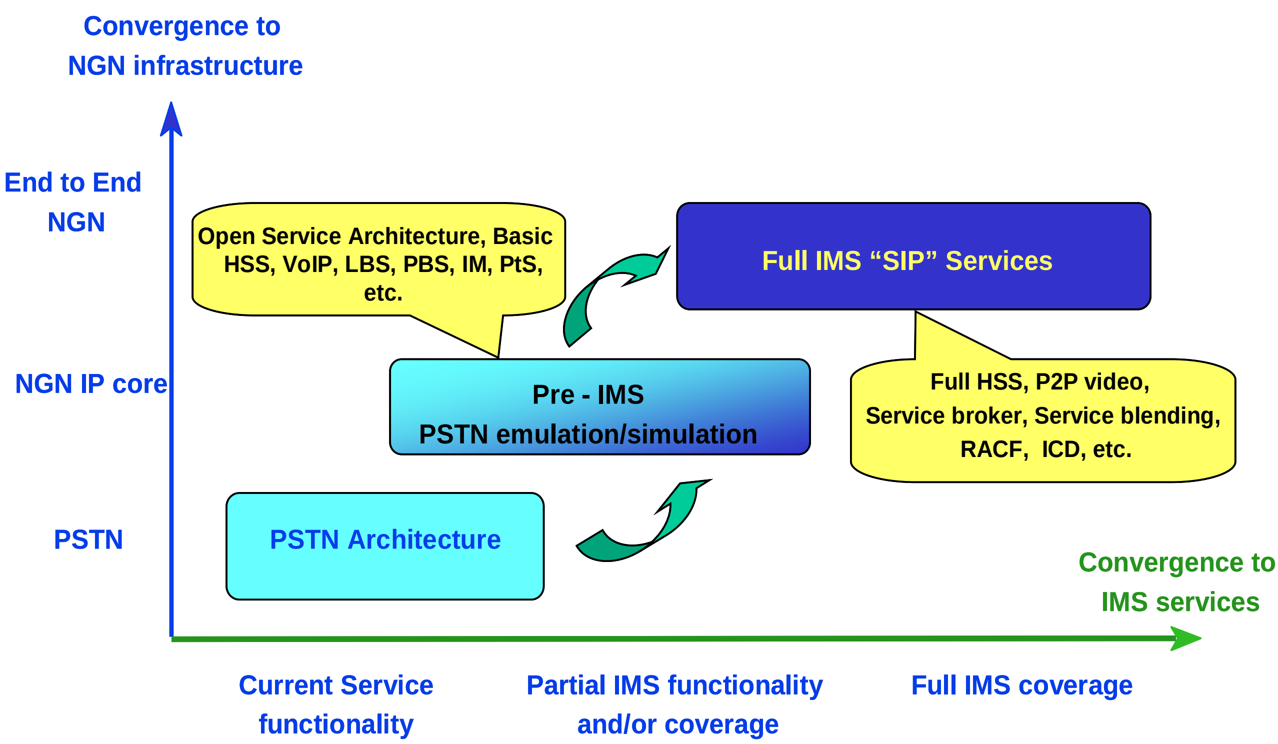


Figure 6.4.4.2: Phased approach for an evolution towards IMS services implementation

6.4.4.3 IMS Benefits

- Service delivery with IMS integrated architecture provides important advantages over the classical "pile" separated applications and functionalities once that the structure is implemented. First advantage is the higher flexibility of the IMS functionality to adapt to the customer services, irrespective of the technology they use and the access method to reach the network.

- Saving in effort and time for the development and deployment of a new service is considerably reduced once the architecture is ready at the network, implying economic savings and better Time to Market for a given service provider in a competitive market. This advantage in the Time to Market will allow the service provider a higher capture of new customers, better market share and reduction of churn.

- Efficient introduction on new services at a lower cost will increase the service provider revenues, ARPU and profitability which is the major business driver for the healthy operation, market grow and financial results.

- From the end customers' point of view, a common use and feel for all services and applications will facilitate the higher utilization of services and better personalization of functions to specific requirements.

Chapter 7 – Network design, dimensioning and optimization

Heterogeneous networks require a variety of planning methods in order to assure the conformance of a composite network with its specification while keeping network costs low. In this chapter an overview on the diverse models and methods used in the telecommunication network planning is given. Network design and planning is based on mathematical theories related to optimization and operations research. In fact, the most important mathematical framework for formulating and resolving network design problems is called the theory of multi-commodity flow networks. The presentation of Sections 7.1 and 7.2 (partly) is based on this framework. In these sections we introduce a number of mathematical models related to most important telecommunication network design problems. The basic optimization methods applicable for network design are presented in Section 7.3.

7.1. Core Network

By core networks we roughly mean wide-area networks, usually spread over a large geographical area, and connecting a set of access/local networks. In this section we shall make a concise survey of problems related to core network design. The presentation will use the multi-commodity flow network design and is based on book [7.1] (which means that all problems considered below are discussed in detail in [7.1]). Basic optimization methods applicable for the discussed problems are described in Section 7.3.

7.1.1. Single layer design

In this subsection we shall discuss selected optimization models related to single-layer networks, applicable for the classical dimensioning and allocation problems of core communication networks involving the nominal (normal) state of network operation. We start with classical problems in Paragraph 7.1.1.1. Then, in Paragraph 7.1.1.2, we will demonstrate how these models can be extended to the shortest-path routing of the OSPF type.

7.1.1.1. Classical problems

Dimensioning Problems

Dimensioning problems (also called capacitated problems) require simultaneous optimization of flows and link capacities. We start the presentation with a simple dimensioning problem, which we will further extend in various directions throughout Subsection 7.1.1. The problem is referred to as DP1-LP (Dimensioning Problem 1 - Linear Programming formulation) and assumes that the lists of candidate paths for the demands are given in advance.

DP1-LP (Dimensioning Problem 1 - Linear Programme)

indices

$d=1,2,...,D$ demands

$j=1,2,...,m(d)$ candidate paths for flows realizing demand d

$e=1,2,...,E$ links

constants

a_{edj} = 1 if link e belongs to path j realizing demand d
= 0 otherwise

h_d volume of demand d, $\boldsymbol{h} = (h_1,h_2,...,h_D)$

c_e marginal (unit) cost of link e, $\boldsymbol{c} = (c_1,c_2,...,c_E)$

variables

x_{dj} non-negative continuous flow allocated to path j of demand d,
$\boldsymbol{x} = (x_{dj}: d=1,2,...,D, j=1,2,...,m(d))$

y_e non-negative continuous capacity of link e, $\boldsymbol{y} = (y_1,y_2,...,y_E)$

objective

minimize $\boldsymbol{C}(\boldsymbol{y}) = \Sigma_e c_e y_e$ (7.1.1a)

constraints

$\Sigma_j x_{dj} = h_d \qquad d=1,2,...,D$ (7.1.1b)

$\Sigma_d \Sigma_j a_{edj} x_{dj} \le y_e \qquad e=1,2,...,E.$ (7.1.1c)

Objective function (7.1.1a) is interpreted as the cost of the link capacity. Constraint (7.1.1b) is called the demand constraint; it assures that all demand volumes are realized by means of flows assigned to their paths (of course, not all flows have to be non-zero). Constraint (7.1.1c) is called capacity constraint and it requires that the link load (left hand side) does not exceed the link capacity. As indicated by its name, problem DP1-LP is a linear programming problem, since all the functions defining cost (7.1.1a), demand constraints (7.1.1b), capacity constraints (7.1.1c), and the non-negativity constraints (7.1.1d) are linear, and all variables are continuous. It is quite easy to see that problem (7.1.1) can be solved in a straightforward way, by allocating the demand volumes to their shortest paths with respect to the link unit costs. To demonstrate this, we first note that for any optimal solution $(\boldsymbol{x}^0,\boldsymbol{y}^0)$ of the problem all constraints (7.1.1c) become equalities (otherwise we would unnecessarily pay for unused capacity of links). Then, we can eliminate the capacity variables, inserting the left hand sides $\Sigma_d \Sigma_j \boldsymbol{a}_{edj} x_{dj}$ of (7.1.c) (link loads) instead of y_e into (7.1.1a). The reduced problem (only in variables $\boldsymbol{x}$) reveals that the non-zero optimal flows x_{dj}^0 can be assigned only to the paths with the minimum length $\Sigma_e \boldsymbol{a}_{edj} c_e$ (we leave the details as an easy exercise for the reader). Thus, an optimal solution of problem (7.1.1) can be easily found by using the *shortest path allocation*. Observe, that if demand d has several shortest paths, then its volume h_d can be split into the flows assigned to the shortest paths in an arbitrary way.

In the sequel, unless stated otherwise, we will assume non-negativity of all optimization variables.

In most cases the link capacities are not continuous, since the link capacity in majority of network technologies is composed of certain capacity modules. For instance, in an SDH/SONET network, links are typically dimensioned in STM-1 (OC-48) modules corresponding to the transmission rate of 155.52 Mbps. Then we have to change accordingly the formulation of DP1-LP by imposing the requirement of integrality (modularity) of the link capacity variables and changing constraint (7.1.1c) into:

$$\Sigma_d \Sigma_j a_{edj} x_{dj} \le M y_e \quad e=1,2,...,E \tag{7.1.1c′}$$

where M is the given module size. The resulting, modified problem will be referred to as DP1-MIP (Dimensioning Problem 1 - Mixed Integer Programming formulation). The name, MIP, reflects the fact that some variables (flows) are continuous, and some (link capacities) are integral. It is important to note that the above, seemingly simple, modification makes the resulting problem very difficult to solve in the exact way, especially for large networks. It is well known that problem DP1-MIP is NP-complete (NP-complete problems as DP1-MIP are very difficult to solve exactly in a computationally effective way; see [7.2] and Appendix B in [7.1] for the discussion of the notion of NP-completeness), which means that all exact algorithms for resolving the problem have exponential complexity, i.e., the time required to find the exact solution grows exponentially with the size of the problem. In practice, the applicable exact methods are based on the branch-and-bound approach [7.3], especially on its modification called branch-and-cut [7.4]. We should emphasize that the branch-and-cut algorithms are pretty complicated and require generating the so called valid inequalities in the nodes (subproblems) of the branch-and-bound tree (see [7.5]), in order to effectively account for the integrality of link capacities. Such equalities can be obtained by Benders' decomposition, or by some special, problem-specific methods (cf. [7.6]). For approximate solutions one can always try to use the shortest path allocation approach (applied for the unit module costs $\boldsymbol{c}$ used as the link metrics for the shortest paths computation), and then dimension the links for the resulting link loads. This approach, however, may sometimes result in quite poor solutions, with the cost (7.1.1a) much higher than optimal.

In many cases, also the demand volume is modular. For instance, if the demand for digital telephone circuits (64 kbps each) is to be realized in the STM-1 links, then, for the European version of the PCM system, we assume that one demand volume unit (DVU) corresponds to 30 circuits, and then $M = 63$, since one STM-1 transport module can carry 63 VC-12 containers, each carrying one PCM basic module. Then, DP1-MIP is simply modified by the requirement that the flows must be non-negative integers and assuming $M = 63$. In effect, we arrive at an all-integer problem DP1-IP, referred to as Dimensioning Problem 1 - Integer Programming formulation. DP1-IP can be approached in a similar way as DP1-MIP, i.e., with the branch-and-cut algorithms or by linear approximation.

In many cases, the modularity of the link capacity may be more complicated than being the multiple of just one module M. First of all, the link capacity can be built from more than one type of the module, as for example in an SDH network with transmission systems STM-1, STM-4 and STM-16. In such case we have three modules: M, $4M$ and $16M$, for $M = 155.52$ Mbps. In the general case we assume that there K module types with module sizes M_k, $k=1,2,..,K$. Then, we have to introduce more link capacity variables and this leads to the following problem.

DP2-MIP (Dimensioning Problem 2 - Mixed Integer Programme)

indices

$d=1,2,...,D$ demands

$j=1,2,...,m(d)$ candidate paths for demand d
$e=1,2,...,E$ links
$k=1,2,...,K$ number of different link capacity modules

constants

a_{edj} = 1 if link e belongs to path j realizing demand d
= 0 otherwise
h_d volume of demand d, $\boldsymbol{h} = (h_1,h_2,...,h_D)$
M_k capacity of module no. k expressed in DVU's
c_{ek} cost of one module of type k on link e

variables

x_{dj} continuous flow allocated to path j of demand d,
$\boldsymbol{x} = (x_{dj}: d=1,2,...,D, j=1,2,...,m(d))$
y_{ek} integer number of modules of type k realized on link e,
$\boldsymbol{y} = (y_{ek}: e=1,2,...,E, k=1,2,...,K)$

objective

minimize $\boldsymbol{C}(\boldsymbol{y}) = \Sigma_e \Sigma_k c_{ek}y_{ek}$ (7.1.2a)

constraints

$\Sigma_j x_{dj} = h_d$ $d=1,2,...,D$ (7.1.2b)
$\Sigma_d \Sigma_j a_{edj}x_{dj} \le \Sigma_k M_k y_{ek}$ $e=1,2,...,E.$ (7.1.2c)

Problem (7.1.2) is more complicated than DP1-MIP, still the optimization approaches for DP1-MIP can be extended to DP2-MIP. Obviously, integer flows can be assumed as well.

The next, more general type of modularity, can be introduced using a general step-wise dimensioning function in the following way.

DP3-MIP (Dimensioning Problem 3 - MIP)

---

indices

$d=1,2,...,D$ demands
$j=1,2,...,m(d)$ candidate paths for demand d
$e=1,2,...,E$ links
$k=1,2,...,K$ number of incremental link capacity modules

constants

a_{edj} = 1 if link e belongs to path j realizing demand d
= 0 otherwise
h_d volume of demand d, $\boldsymbol{h} = (h_1,h_2,...,h_D)$
m_k incremental capacity of module no. k expressed in DVU's
c_{ek} incremental cost of one module of type k on link e

variables

x_{dj} continuous flow allocated to path j of demand d,

$\boldsymbol{x} = (x_{dj}: d=1,2,...,D, j=1,2,...,m(d))$

ε_{ek} binary variable indicating whether incremental module of type k is realized on link e, $\boldsymbol{\varepsilon} = (\varepsilon_{ek}: e=1,2,...,E, k=1,2,...,K)$

objective

minimize $\boldsymbol{C(\varepsilon)} = \Sigma_e \Sigma_k c_{ek}\varepsilon_{ek}$ (7.1.3a)

constraints

$\Sigma_j x_{dj} = h_d$ $d=1,2,...,D$ (7.1.3b)

$\Sigma_d \Sigma_j a_{edj}x_{dj} \leq \Sigma_k m_k\varepsilon_{ek}$ $e=1,2,...,E$ (7.1.3c)

$\varepsilon_{e1} \geq \varepsilon_{e2} \geq ... \geq \varepsilon_{eK}$ $e=1,2,...,E.$ (7.1.3d)

Note that the new constraint (7.1.3d) makes sure that if, for some k, incremental module of type k is installed, then also all modules of type i, $i < k$, are installed. Hence, the actual capacity of link e is the sum of all m_k with $k=1,2,...,k_{max}$, where k_{max} is the greatest index k such that $\varepsilon_{ek} = 1$ (cf. constraint (7.1.3c)). Of course, it is assumed that $\Sigma_k m_k$ is the maximal capacity of each link. It may seem that problem DP3-MIP is harder to solve than problem DP2-MIP, due to additional “monotonicity” constraints (7.1.3d). In practice the opposite holds. Additional constraints allow for more effective use of the branch-and-bound tree and lead to shorter execution times.

Finally, we note that in some cases we may need to use concave dimensioning functions $y_e = F_e(\underline{y}_e)$ in the problem formulations ($\underline{y}_e$ denotes load of link e). For instance, if the link load $\underline{y}_e$ is expressed in Erlangs (1 DVU = 1 Erl.), then its capacity can be computed as the number of circuits y_e such that

$$B_e = E_{ye}(\underline{y}_e) \qquad (7.1.4)$$

where is $E_{ye}(\underline{y}_e)$ the Erlang loss formula giving the link call blocking probability when traffic of $\underline{y}_e$ Erlangs is offered to y_e circuits, and B_e is the assumed fixed link call blocking. The resulting dimensioning function $y_e = F_e(\underline{y}_e)$ is the inverse of (7.1.4) for fixed B_e, and is known to be a concave function. With concave dimensioning function the dimensioning problem DP1-LP takes the form:

DP4-CV (Dimensioning Problem 4 - Concave Programming formulation)

---

constants

a_{edj} = 1 if link e belongs to path j realizing demand d
= 0 otherwise

h_d volume of demand d, $h = (h_1,h_2,...,h_D)$

$F_e(\cdot)$ concave dimensioning function for link e

c_e marginal (unit) cost of link e, $c = (c_1,c_2,...,c_E)$

variables

x_{dj} continuous flow allocated to path j of demand d,
$\boldsymbol{x} = (x_{dj}: d=1,2,...,D, j=1,2,...,m(d))$

$\underline{y}_e$ continuous load of link e, $\underline{\boldsymbol{y}} = (\underline{y}_1,\underline{y}_2,...,\underline{y}_E)$

objective

minimize $\boldsymbol{C}(\underline{\boldsymbol{y}}) = \Sigma_e\, c_e F_e(\underline{y}_e)$ (7.1.5a)

constraints

$\Sigma_j\, x_{dj} = h_d$ $d=1,2,...,D$ (7.1.5b)

$\Sigma_d\, \Sigma_j\, a_{edj} x_{dj} = \underline{y}_e$ $e=1,2,...,E.$ (7.1.5c)

The concave programming problems are very difficult to solve for global minimum, as they are usually characterized by enormous number of local minima. In fact, as discussed in detail in Section 4.3 of [7.1], problem DP4-CV can be transformed to a MIP problem and solved accordingly (by branch-and-cut). Also, we may try use stochastic meta-heuristics discussed in Section 7.3.3.

Another type of extensions of the basic problem DP1-LP is obtained when flow routing is constrained in some way (so far we have not imposed any constraint on the flows). The first requirement which constraints the flow distribution is the path diversity requirement: the demand volume must be split among at least a certain number of disjoint paths.

DP5-PD-LP (Dimensioning Problem 5 - Path Diversity - LP)

---

indices

$d=1,2,...,D$ demands

$j=1,2,...,m(d)$ candidate disjoint paths for demand d

$e=1,2,...,E$ links

constants

a_{edj} = 1 if link e belongs to path j realizing demand d
= 0 otherwise

h_d volume of demand d, $\boldsymbol{h} = (h_1,h_2,...,h_D)$

n_d minimal number of paths for splitting volume h_d (e.g. $n_d = 3$)

c_e marginal (unit) cost of link e, $c = (c_1,c_2,...,c_E)$

variables

x_{dj} continuous flow allocated to path j of demand d,
$\boldsymbol{x} = (x_{dj}$: $d=1,2,...,D$, $j=1,2,...,m(d))$

y_e continuous capacity of link e, $\boldsymbol{y} = (y_1,y_2,...,y_E)$

objective

minimize $\boldsymbol{C}(\boldsymbol{y}) = \Sigma_e\, c_e y_e$ (7.1.6a)

constraints

$\Sigma_j\, x_{dj} = h_d$ $d=1,2,...,D$ (7.1.6b)

$x_{dj} \le h_d / n_d$ $d=1,2,...,D$ (7.1.6c)

$\Sigma_d\, \Sigma_j\, a_{edj} x_{dj} \le y_e$ $e=1,2,...,E.$ (7.1.6d)

Path diversity is assured jointly by the requirement that the candidate paths are disjoint (link or node disjoint) and by constraint (7.1.6c) which does not allow to put more than $1/n_d$ of the demand volume on one path. This requirement is a commonly used means for protecting the demands against link (node) failures.

The next problem requires that the entire demand volume of each demand is assigned to only one path.

DP6-SPR-MIP (Dimensioning Problem 6 - Single-Path Routing - MIP)

variables

x_{dj} flow allocated to path j of demand d
$\boldsymbol{x} = (x_{dj}: d=1,2,...,D, j=1,2,...,m(d))$

ε_{dj} binary variable associated with flow variable x_{dj}
$\boldsymbol{\varepsilon} = (\varepsilon_{dj}: d=1,2,...,D, j=1,2,...,m(d))$

y_e integer capacity of link e, $\boldsymbol{y} = (y_1,y_2,...,y_E)$

objective

minimize $\boldsymbol{C}(\boldsymbol{y}) = \Sigma_e c_e y_e$ (7.1.7a)

constraints

$x_{dj} = h_d u_{dj}$ $d=1,2,...,D$ $j=1,2,...,m(d)$ (7.1.7b)

$\Sigma_j \varepsilon_{dj} = 1$ $d=1,2,...,D$ (7.1.7c)

$\Sigma_d \Sigma_j a_{edj} x_{dj} \le M y_e$ $e=1,2,...,E.$ (7.1.7d)

Note that since link capacities are modular, the single-path allocation to the shortest paths with respect to link weights $\boldsymbol{c}$ does not in general solve the problem, so the single-path routing must be forced explicitly. This is done by the binary variables $\boldsymbol{\varepsilon}$ and constraint (7.1.7c). Observe also that the flow variables are auxiliary in formulation (7.1.7) since they can be eliminated by substituting x_{dj} in (7.1.7d) with the right hand side of (7.1.7b). DP6-SPR-MIP is an example of another NP-complete problem, and hence is difficult to solve. Again, for exact solutions the branch-and-cut approaches are applicable here [7.7]. For approximate solutions meta-heuristic methods are applicable [7.8].

Other routing restrictions can also be taken into account by appropriate MIP formulations. For instance, we may require that non-zero flows must be greater than a certain fraction of the demand volume (not to use too small flows), or that the demand volume must be must be split among at most n_d paths. Such formulations can be found in Chapter 4 of [7.1].

Allocation Problems

In allocation problems, called also capacitated problems, link capacities are given (installed) and fixed; the task is to allocate flows for given demand volumes in such a way that the resulting link loads do not exceed link capacities. Although certain additional objective function can be added to allocation problems, the main issue is to find a feasible solution, i.e., to be able to allocate demands in the existing link capacity, as in the following problem.

AP1-LP (Allocation Problem 1 - Linear Programme)

indices

$d=1,2,...,D$ demands
$j=1,2,...,m(d)$ candidate paths for flows realizing demand d
$e=1,2,...,E$ links

constants

a_{edj} = 1 if link e belongs to path j realizing demand d
= 0 otherwise
h_d volume of demand d, $\boldsymbol{h} = (h_1,h_2,...,h_D)$
y_e capacity of link e, $\boldsymbol{y} = (y_1,y_2,...,y_E)$

variables

x_{dj} continuous flow allocated to path j of demand d,
$\boldsymbol{x} = (x_{dj}: d=1,2,...,D, j=1,2,...,m(d))$

constraints

$$\Sigma_j x_{dj} = h_d \qquad d=1,2,...,D \tag{7.1.8a}$$
$$\Sigma_d \Sigma_j a_{edj}x_{dj} \le y_e \qquad e=1,2,...,E. \tag{7.1.8b}$$

Note that the above problem has no objective function. There is no such a simple way to solve AP1-LP as the shortest path allocation for DP1-LP, as there are no link unit costs involved in the problem and also optimal solutions may have to be bifurcated (the reader is asked to find an example). In fact, AP1-LP must be solved by the general LP methods (simplex algorithm). A typical objective that can be added to problem (7.1.8) accounts for the cost of flows:

$$\text{minimize} \qquad \boldsymbol{C}(\boldsymbol{x}) = \Sigma_d \Sigma_j c_{dj}x_{dj} \tag{7.1.9}$$

where c_{dj} is the cost of realizing of one DVU of demand d on its path j. Another example of an additional objective is to maximize the total unused capacity of links left after feasible allocation of flows - formulation of this objective is left for the reader as an exercise.

Certainly, additional routing restrictions such as path diversity or single-path routing can be added to the problem in the same way as for the dimensioning problems. Also integral (modular) demand volumes can be assumed; then, however, the problem may become NP-complete.

Finally let us notice that in the allocation case it can be important to have access to all the paths in the network graph in order to be able to use the available capacity of the links with no limitations. One way to do it is to use the so called *column generation method* of LP (called path generation in our context) to adjust the candidate path lists (cf. [7.6]).

Another (and somewhat simpler for a reader not used to more sophisticated use of linear programming) is to differently formulate the optimization problem, using the so called *node-link formulation* given below. We point out here that so far we have used the *link-path* *formulations*, with candidate path lists given explicitly in the problem. The node-link formulation assumes the directed graph, so that the links (often called arcs in this case) are directed and this direction must be followed when the link is used in a path. The node-link formulation can be easily adapted for undirected links (see Chapter 4 in [7.1]). Note that in the link-path formulation it is not important if links are directed or undirected; what is important is only whether the paths are correctly constructed.

AP1-NL-LP (Allocation Problem 1 - Node-Link formulation - LP)

indices

$d=1,2,...,D$ demands
$v=1,2,...,V$ nodes
$e=1,2,...,E$ links

constants

A_{ev} = 1 if link e originates at node v, 0 otherwise
B_{ev} = 1 if link e terminates in node v, 0 otherwise
s_d source node of demand d
t_d sink node of demand d
h_d volume of demand d
y_e capacity of link e, $\boldsymbol{y} = (y_1,y_2,...,y_E)$

variables

x_{ed} continuous flow realizing demand d allocated to link e
$\boldsymbol{x} = (x_{de}: d=1,2,...,D, e=1,2,...,E)$

constraints

$$\Sigma_e A_{ev}x_{ed} - \Sigma_e B_{ev}x_{ed} \begin{cases} = h_d & \text{if } v = s_d \\ = 0 & \text{if } v \neq s_d, t_d \\ = -h_d & \text{if } v = t_d \end{cases} \qquad v=1,2,...,V \qquad d=1,2,...,D \tag{7.1.10a}$$

$$\Sigma_d\, x_{ed} \leq \boldsymbol{y_e} \qquad e=1,2,...,E. \tag{7.1.10b}$$

This time the flows realizing a particular demand d are associated with links, not with the paths pre-allocated to the demand. Hence the capacity constraints and link loads take a different form, see (7.1.10a). Also, demand constraints are different, and take the form of the flow conservation law (cf. (7.1.10b)). We note that the node-link formulation (7.1.10) usually has more constraints than the link-path formulation (when candidate path lists are limited). Another important remark here is that there exist more sophisticated node-link formulations involving less flow variables (e.g. with link flows associated only with destination nodes, not with the demands), see Chapter 4 in [7.1].

Topological Design

We end this paragraph by defining two versions of the topological design problem. The problem assumes that there are two components of the link cost: fixed cost of installing the link, and the capacity dependent factor considered so far.

TDP1-LP (Topological Dimensioning Problem 1 - Mixed Integer Programme)

---

indices

d=1,2,...,D	demands
j=1,2,...,$m(d)$	candidate paths for flows realizing demand d
e=1,2,...,E	links

constants

a_{edj}	= 1 if link e belongs to path j realizing demand d
	= 0 otherwise
h_d	volume of demand d, $\boldsymbol{h} = (h_1,h_2,...,h_D)$
c_e	unit of the capacity-dependent cost factor of link e, $\boldsymbol{c} = (c_1,c_2,...,c_E)$
κ_e	installation cost of link e, $\boldsymbol{\kappa} = (\kappa_1,\kappa_2,...,\kappa_E)$
Δ	an upper limit for link capacity (usually a large number)

variables

x_{dj}	continuous flow allocated to path j of demand d,
	$\boldsymbol{x} = (x_{dj}$: d=1,2,...,D, j=1,2,...,$m(d))$
y_e	continuous capacity of link e, $\boldsymbol{y} = (y_1,y_2,...,y_E)$
ε_e	binary variable if link e is installed ($\varepsilon_e = 1$) or not ($\varepsilon_e = 0$), $\varepsilon = (\varepsilon_1,\varepsilon_2,...,\varepsilon_E)$

objective

minimize $\boldsymbol{C}(\boldsymbol{y},\boldsymbol{\varepsilon}) = \Sigma_e\, c_e y_e + \Sigma_e\, \kappa_e \varepsilon_e$ (7.1.11a)

constraints

$\Sigma_j\, x_{dj} = h_d$	d=1,2,...,D	(7.1.11b)
$y_e \le \Delta\varepsilon_e$	e=1,2,...,E	(7.1.11c)
$\Sigma_d\, \Sigma_j\, a_{edj}x_{dj} \le y_e$	e=1,2,...,E.	(7.1.11d)

We note that there appears a new type of constraint, (7.1.11c), which forces that the capacity y_e of a non-installed link e (with $\varepsilon_e = 0$) is equal to 0. A variant of the above problem introduces the budget constraint for the installation cost.

TDP2-LP (Topological Dimensioning Problem 2 - MIP)

---

additional constant

B	budget limit for the installation cost

variables

x_{dj}	continuous flow allocated to path j of demand d,
	$\boldsymbol{x} = (x_{dj}$: d=1,2,...,D, j=1,2,...,$m(d))$

y_e continuous capacity of link e, $\mathbf{y} = (y_1,y_2,...,y_E)$
ε_e binary variable if link e is installed ($\varepsilon_e = 1$) or not ($\varepsilon_e = 0$), $\varepsilon = (\varepsilon_1,\varepsilon_2,...,\varepsilon_E)$

objective

minimize $\boldsymbol{C}(\boldsymbol{y}) = \Sigma_e c_e y_e$ (7.1.11a)

constraints

$\Sigma_j x_{dj} = h_d$ $d=1,2,...,D$ (7.1.11b)
$y_e \leq \Delta\varepsilon_e$ $e=1,2,...,E$ (7.1.11c)
$\Sigma_d \Sigma_j a_{edj}x_{dj} \leq y_e$ $e=1,2,...,E$ (7.1.11d)
$\Sigma_e \kappa_e\varepsilon_e \leq B.$ (7.1.11e)

Both variants of the topological design problem are NP-complete. In fact, problem TDP1-MIP is similar to the modular dimensioning problem DP1-MIP (problem (7.1.1) with integral link capacity variables and constraint (7.1.1c)), and can be solved by similar optimization techniques (see [7.9]).

7.1.1.2. Shortest-Path Routing Allocation Problems

The OSPF (Open Shortest Path First) packet routing protocol is one of the most commonly used Interior Gateway Protocols in today's IP networks. OSPF uses shortest paths for routing the packets and applies the Equal-Cost Multi-Path (ECMP) principle to deal with multiple shortest paths. The packet routing mechanism is relatively simple, and can essentially be summarised as follows: all the packets arriving at an intermediate node v and destined for node t are directed to the next hop along the shortest path from v to t, regardless of the packets' originating nodes. If there is more than one link outgoing from node v and belonging to the shortest paths from v to t, then the traffic is distributed evenly among these links. The shortest paths to destinations are identified at the network nodes on the basis of the current links' weight (metric) system $\boldsymbol{w}$: each link e is assigned a positive number w_e (weight) and, as a result of the OSPF link-state flooding mechanism, all the nodes are aware of the weights $\boldsymbol{w} = (w_1,w_2,...,w_E)$ of all network's links.

In this paragraph we consider the issue of the existence of a feasible OSPF link weight system for given demand matrix and link capacities. In other words, we ask whether there exists a weight system $\boldsymbol{w}$ that generates flows realising the demands such that the resulting link loads do not exceed the given link capacities. We consider the following Allocation Problem for the Shortest-Path Routing.

AP2-SPR (Allocation Problem 2 - Shortest-Path Routing)

indices

$d=1,2,...,D$ demands
$j=1,2,...,m(d)$ candidate paths for flows realizing demand d
$e=1,2,...,E$ links

constants

a_{edj}	= 1 if link e belongs to path j realizing demand d
	= 0 otherwise
h_d	volume of demand d, $\boldsymbol{h} = (h_1,h_2,...,h_D)$
y_e	capacity of link e, $\boldsymbol{y} = (y_1,y_2,...,y_E)$
W	set of admissible weight systems

variables

x_{dj}	continuous flow allocated to path j of demand d,
	$\boldsymbol{x} = (x_{dj}: d=1,2,...,D, j=1,2,...,m(d))$
w_e	weight of link e, $\boldsymbol{w} = (w_1,w_2,...,w_E)$

constraints

$x_{dj} = x_{dj}(\boldsymbol{w})$	$d=1,2,...,D$	$j=1,2,...,m(\mathrm{d})$	(7.1.12a)
$\Sigma_j\, x_{dj} = \boldsymbol{h_d}$	$d=1,2,...,D$	(7.1.12b)	
$\Sigma_d\, \Sigma_j\, \boldsymbol{a}_{edj}x_{dj} \le y_e$	$e=1,2,...,E$		(7.1.12c)
$\boldsymbol{w} \in W.$			(7.1.12d)

Above, $x_{dj}(\boldsymbol{w})$ denotes the flow realising demand d on path j, implied by the link weight system $\boldsymbol{w}$. For a given weight system $\boldsymbol{w}$, the flows $x_{dj}(\boldsymbol{w})$ are computed according to the ECMP rule. The rule is illustrated in Figure 7.1.3: the shortest paths s-a-c-t and s-a-d-t realise 0.25 of the total demand volume between nodes s and t, while the shortest path s-b-e-t realises the remaining 0.5 of the volume. Note that the flow functions $x_{dj}(\boldsymbol{w})$ are not given explicitly and hence problem AP2-SPR is not an optimization problem in the standard form (optimization problems in the standard form are called mathematical programmes). In order to make the ECMP flow splitting procedure consistent, we assume that the link weights are positive, which eliminates loops in the shortest paths. Consequently, constraints (7.1.12b) guarantee that all demands are realised, and constraints (7.1.12c) - that links' loads do not exceed their capacities. The weight system space W in constraint (7.1.12d) limits the set of link weights systems; for instance, a state space assuring the consistency of the weight systems is:

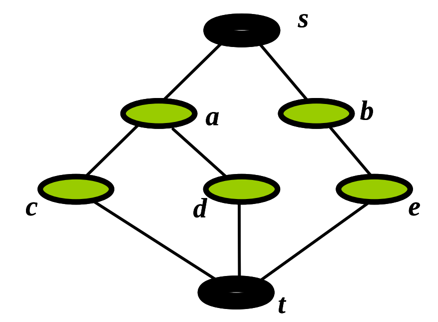


Fig.7.1. Equal-split rule

$$1 \le w_e \le K \ \text{ and } \ w_e \text{ - integer}, \quad e=1,2,...,E \qquad (7.1.13)$$

for some integer K.

As shown in Chapter 7 of [7.1], problem (7.1.12) is NP-complete. Recently, several approximate (and exact, based on the branch-and-cut approach) methods for solving this problem appeared, as well as for its various modifications (for a survey see Chapter 7 in [7.1]).

7.1.2. Multi-state restoration/protection design

An important extension in the nominal design of single-layer networks is to take into account failure situations at the design stage and plan the installation of link (and node) capacity sufficient not only for the nominal state of network operation (i.e., the state when all resources are available) but also for the assumed major failure states (e.g., cable cuts) when some part of link capacity is failed and not available. Needless to say such a design will need more capacity than for the nominal design considered in Section 7.1.1, as spare capacity (on top of the nominal capacity) to be used in failure situations to restore failed demands is required.

7.1.2.1. Failure situations

Following [7.1], we shall label the considered failure situations with $s = 0,1,...,S$ where $s = 0$ denotes the nominal (normal) state. Each failure situation s is characterized by a vector of link availability coefficients $\boldsymbol{\alpha}_s = (\alpha_{1s},\alpha_{2s},...,\alpha_{Es})$ with $0 \le \alpha_{es} \le 1$. Coefficient α_{es} specifies the fraction of the nominal capacity y_e of link e, $\alpha_{es}y_e$, that is available on link e in situation s. As we will see, in certain problems it will be important to assume that the availability coefficients are binary, i.e. $\alpha_{es} \in \{0,1\}$. In any case, we in general assume that more than one link can fail at a time (in a particular situation s). Note that $\alpha_{e0} = 1$ for $e = 1,2,...,E$, i.e., in the nominal state $s = 0$ all links are fully available.

It is important than the demand in failure situation s can be different (typically reduced) from nominal. Hence, we denote the demand volume of demand d in situation s by h_{ds} (with $h_{ds} \le h_d$). Note that link availability coefficients can be used to model node failures. If a node v fails, we simply put $\alpha_{es} = 0$ for all links incident with the failed node v, and, what can be important, $h_{ds} = 0$ for all demands d incident with the failed node (node v is one of the end nodes of d).

7.1.2.2. Restoration (protection) mechanisms

Restoration (protection) mechanisms are responsible for restoring (protecting) demand volumes into the assumed degree (for demand d and situation s this degree is determined by the difference of the nominal demand volume h_d and the situation-dependent demand volume h_{ds}). Roughly speaking, the protection mechanisms are “passive” and provide protection of flows by splitting (path-diversity) or duplicating (hot-standby) all, or a part of nominal flows.

In turn, the restoration mechanisms are “active” and can reroute failed flows around the failed links. Here two basic types of mechanisms are used: link protection and path protection. Link protection is used to protect networks against single (but total) link failure; when a link fails its capacity is re-routed on a “detour” path (in this way all demand flows that use the failed link are restored automatically). Path protection is more complicated: when one or more link fails, the affected flows are re-routed (individually, on the end-to-end basis) using the spare capacity installed for this purpose. Certainly, the latter mechanism in general requires less protection capacity than the former; on the other hand path protection is more complicated to implement.

Spare capacity may be dedicated to resources (when for example a path-flow has a dedicated capacity to protect it in all situations) or shared (when the whole pool of spare capacity can be used for restoring all flows/links in all situations). Below, we shall present most representative design problems related to networks robust to failures.

7.1.2.3. Path diversity

Probably the simplest way to achieve some degree of robustness is path-diversity. Recall that path diversity is a requirement to split demand volumes into several (link or node disjoint) paths and has been discussed in Problem DP5-PD-LP (7.1.6). An assumed degree of protection can be achieved through path-diversity at the expense of realizing more nominal demand than really required, as illustrated below:

RDP-GPD-LP (Robust Dimensioning Problem - Generalised Path Diversity - LP)

---

indices

d=1,2,...,D	demands
j=1,2,...,$m(d)$	link- or node-disjoint candidate paths for demand d
e=1,2,...,E	links
s=0,1,...,S	situations (s = 0 denotes the nominal state)

constants

a_{edj}	= 1 if link e belongs to path j realizing demand d; 0, otherwise
h_{d0}	nominal volume of demand d, $h_{d0} = h_d$
h_{ds}	demand volume of demand d in situation s
c_e	marginal (unit) cost of link e
α_{es}	binary availability coefficient of link e in situation s ($\alpha_{es} \in \{0,1\}$)
δ_{djs}	binary availability coefficient of path (d, j) in situation s, $\delta_{djs} = \Pi_{e:\, aedj=1}\, \alpha_{es}$

variables

x_{dj0}	continuous nominal flow allocated to path j of demand d
y_e	continuous capacity of link e

objective

minimize $\quad \boldsymbol{C}(\boldsymbol{y}) = \Sigma_e\, c_e y_e$ (7.1.13a)

constraints

$\Sigma_j\, \theta_{djs} x_{dj0} \geq h_{ds} \qquad d=1,2,...,D \qquad s=0,1,...,S$ (7.1.13b)

$\Sigma_d\, \Sigma_j\, a_{edj} x_{dj0} \leq y_e \qquad e=1,2,...,E.$ (7.1.13c)

Crucial to understanding the above problem are the path availability coefficients δ_{djs}. Such a coefficient for situation s is equal to 1 if, and only if, all links composing the considered path no. j of demand d (denoted in the sequel by P_{dj}) are fully available in situation s. For that to work we need binary link availability coefficients α_{es}.

Certainly, modular link capacities and/or modular flows can be assumed instead of continuous quantities. This, as usual, makes the problem much more difficult to solve.

7.1.2.4. Hot-standby

With hot-standby, basic, nominal flows are protected by their "copies" realized on dedicated paths capacities. This mechanism is simple to realize, yet expensive in terms of additional spare (dedicated) link capacity required.

RDP-HS-MIP (Robust Dimensioning Problem - Hot-standby - MIP)

---

indices

$d=1,2,...,D$	demands
$k=1,2,...,m(d)$	candidate nominal paths for demand d
$l=1,2,...,n(d,k)$	candidate backup paths for nominal path P_{dk} (each path Q_{dkl} is disjoint with path P_{dk})
$e=1,2,...,E$	links
$s=0,1,...,S$	situations

constants

a_{edk}	= 1 if link e belongs to nominal path P_{dk}; 0, otherwise
b_{edkl}	= 1 if link e belongs to backup path Q_{dkl} protecting nominal path P_{dk}; 0, otherwise
h_d	volume of demand d
c_e	marginal cost of link e
α_{es}	binary availability coefficient of link e in situation s ($\alpha_{es} \in \{0,1\}$)
$\delta_{\boldsymbol{dks}}$	binary path availability coefficient indicating whether nominal path P_{dk} is available in situation s, $\delta_{dks} = \Pi_{e:\, aedk=1}\, \alpha_{es}$

variables

x_{dkl0}	continuous nominal flow of demand d allocated to pair (P_{dk},Q_{dkl})
u_{dkl}	binary variable corresponding to flow x_{dkl0}
y_e	continuous capacity of link e

objective

minimize $\boldsymbol{C}(\boldsymbol{y}) = \Sigma_e\, c_e y_e$ (7.1.14a)

constraints

$\Sigma_k\, x_{dkl0} = h_d$	$d=1,2,...,D$		(7.1.14b)
$\Sigma_l\, u_{dkl} \le 1$	$d=1,2,...,D$	$k=1,2,...,m(d)$	(7.1.14c)
$x_{dkl0} \le h_d u_{dkl}$	$d=1,2,...,D$	$k=1,2,...,m(d)$ $l=1,2,...,n(d,k)$	(7.1.14d)
$\Sigma_d\, \Sigma_k\, \Sigma_l\, (\delta_{dks}\, a_{edk} + (1-\delta_{dks}) b_{edkl}) x_{dkl0} \le \alpha_{es} y_e$	$e=1,2,...,E$ $s=0,1,...,S.$		(7.1.14e)

Note that the above formulation is pretty complicated although the mechanism itself is simple. The formulation assures that for each nominal flow is routed simultaneously on exactly one of its backup paths.

7.1.2.5. Link protection

For the link protection mechanism it is guaranteed that the network demands are fully protected in the case of any total failure of any of the links. In such a case all the failed capacity is restored using spare capacity, which is shared for restoring of all links (in different situations).

RDP-LP-LP (Robust Dimensioning Problem - Link Protection - LP)

indices

$d=1,2,...,D$	demands
$j=1,2,...,m(d)$	allowable paths for flows realizing demand d
$e,f=1,2,...,E$	links
$k=1,2,..., n(e)$	restoration paths for link e

constants

a_{edj}	= 1 if link e belongs to path j realizing demand d; 0, otherwise
h_d	volume of demand d
c_e	marginal (unit) cost of link e
b_{fek}	= 1 if link f belongs to path k restoring link e; 0, otherwise

variables

x_{dj0}	continuous nominal flow allocated to path j of demand d
y_e	continuous nominal capacity of link e
z_{ek}	continuous flow restoring nominal capacity of link e on restoration path k
y_e'	continuous spare, protection capacity of link e (not used in the nominal state)

objective

minimize $\boldsymbol{C}(\boldsymbol{y}) = \Sigma_e c_e(y_e + \boldsymbol{y_e}')$ (7.1.15a)

constraints

$\Sigma_j x_{dj0} = h_d$	$d=1,2,...,D$	(7.1.15b)
$\Sigma_d \Sigma_j a_{edj}x_{dj0} \le y_e$	$e=1,2,...,E$	(7.1.15c)
$\Sigma_k z_{ek} = y_e$	$e=1,2,...,E$	(7.1.15d)
$\Sigma_k b_{fek}z_{ek} \le y_f'$	$f=1,2,...,E \quad e=1,2,...,E \quad f \ne e.$	(7.1.15e)

7.1.2.6. Path protection

The path protection mechanism is more complicated than link protection. It does not assume any particular type of failures (as single link failures) and consists in restoring the demand volumes on the path-flow basis. There are several variations of path protection. Below we will discuss three of them.

The first variation assumes that in case of failure all flows can be disconnected and restored from scratch (into the assumed degree) in the surviving link capacity.

RDP-PP1-LP (Robust Dimensioning Problem - Path Protection 1 - LP)

indices

$d=1,2,...,D$	demands
$j=1,2,...,m(d)$	allowable paths for flows realizing demand d
$e=1,2,...,E$	links
$s=0,1,...,S$	situations

constants

a_{edj}	= 1 if link e belongs to path j realizing demand d; 0, otherwise
h_{ds}	volume of demand d in situation s
c_e	unit cost of link e
α_{es}	fractional availability coefficient of link e in situation s ($0 \le \alpha_{es} \le 1$)

variables

x_{djs}	continuous flow allocated to path j of demand d in situation s
y_e	continuous capacity of link e

objective

minimize $\boldsymbol{F} = \Sigma_e c_e y_e$ (7.1.16a)

constraints

$\Sigma_j x_{djs} = h_{ds}$ $d=1,2,...,D$ $s=0,1,...,S$ (7.1.16b)

$\Sigma_d \Sigma_j \delta_{edj} x_{djs} \le \alpha_{es} y_e$ $e=1,2,...,E$ $s=0,1,...,S.$ (7.1.16c)

The second variation assumes that the unaffected flows are not moved and only the broken flows are restored (individually). Note that the surviving capacity "released" by broken flows is used for the restoration. Also, only total link failures are assumed.

RDP-PP2-LP (Robust Dimensioning Problem - Path Protection 2 - LP)

indices

$d=1,2,...,D$	demands
$j=1,2,...,m(d)$	candidate paths for demand d
$e=1,2,...,E$	links
$s=0,1,...,S$	situations

constants

a_{edj}	= 1 if link e belongs to path j realizing demand d; 0, otherwise

h_{ds} volume of demand d in situation s
c_e marginal (unit) cost of link e
α_{es} binary availability coefficient of link e in situation s ($\alpha_{es} \in \{0,1\}$)
δ_{djs} binary availability coefficient of path (d, j) in situation s, $\delta_{djs} = \Pi_{e:\, aedj=1}\, \alpha_{es}$

variables (all continuous non-negative)
x_{dj0} nominal flow allocated to path j of demand d
x_{djs} flow allocated to path j of demand d in situation s (these flows are provided on top on the surviving nominal flows)
y_e capacity of link e
z_{ds} volume of demand d surviving in failure situation s
y_{es}' capacity of link e not occupied by surviving nominal flows in situation s (provided e is not failed in situation s)

objective
minimize $\boldsymbol{C}(\boldsymbol{y}) = \Sigma_e\, c_e y_e$ (7.1.17a)

constraints
$\Sigma_j\, x_{dj0} = h_{d0}$ $d=1,2,...,D$ (7.1.17b)
$\Sigma_d\, \Sigma_j\, a_{edj} x_{dj0} \le y_e$ $e=1,2,...,E$ (7.1.17c)
$z_{ds} = \Sigma_j\, \delta_{djs}\, x_{dj0}$ $d=1,2,...,D$ $s=1,2,...,S$ (7.1.17d)
$\Sigma_j\, x_{djs} \ge h_{ds} - z_{ds}$ $d=1,2,...,D$ $s=1,2,...,S$ (7.1.17e)
$y_{es}' = y_e - \Sigma_d\, \Sigma_j\, a_{edj}\delta_{djs}x_{dj0}$ $e=1,2,...,E$ $s=1,2,...,S$ (7.1.17f)
$\Sigma_d\, \Sigma_j\, a_{edj}x_{djs} \le \alpha_{es}y_{es}'$ $e=1,2,...,E$ $s=1,2,...,S.$ (7.1.17g)

Finally, the simplest (by not in terms of the formulation) path protection mechanism assumes that each broken nominal flow is restored only one link, common to all situations. Thus, it is assumed that the nominal (basic) paths and their backup paths never fail simultaneously.

RDP-PP3-LP (Robust Dimensioning Problem - Path Protection 3 - LP)

---

indices
$d=1,2,...,D$ demands
$j=1,2,...,m(d)$ pair (P_{dj},Q_{dj}) of situation disjoint paths for flows realizing demand d, nominal path P_{dj}, and backup path Q_{dj} (each such pair is disjoint)
$e=1,2,...,E$ links
$s=0,1,...,S$ situations

constants
a_{edj} = 1 if link e belongs to nominal path P_{dj}; 0, otherwise
b_{edj} = 1 if link e belongs to backup path Q_{dj}; 0, otherwise
h_d volume of demand d
c_e marginal (unit) cost of link e
α_{es} binary availability coefficient of link e in situation s ($\alpha_{es} \in \{0,1\}$)
δ_{djs} binary availability coefficient of path P_{dj} in situation s, $\delta_{djs} = \Pi_{e:\, aedj=1}\, \alpha_{es}$

variables

x_{dj0} continuous flow allocated to basic path j of demand d in the nominal state
y_e continuous capacity of link e

objective

minimize $\boldsymbol{C}(\boldsymbol{y}) = \Sigma_e c_e y_e$ (7.1.18a)

constraints

$\Sigma_j x_{dj0} = h_d$ $d=1,2,...,D$ (7.1.18b)
$\Sigma_d \Sigma_j (a_{edj}\delta_{djs} + b_{edj}(1 - \delta_{djs}))x_{dj0} \leq \alpha_{es}y_e$ $e=1,2,...,E$ $s=0,1,...,S.$ (7.1.18c)

This completes our presentation of the restoration/protection design problems. Of course, many other valid variants of the presented problems can be considered. For example, such elements as modular link capacity, modular flows or single-path routing can be taken into account.

7.1.3. Design of Multi-Layer Networks

In this subsection we shall present selected problems on multi-layer design, involving more than one layer of resources (one layer of resources has been assumed in Subsections 7.1.1 and 7.1.2). In fact, we will consider the case of two resource layers plus the demand layer.

7.1.3.1. Nominal design of multi-layer networks

In this section we will present two nominal design problems for three-layer networks: a dimensioning problem and an allocation problem.

The following formulation is a counterpart of the simple single-layer design problem DP1-LP (7.1.1). Although the considered model actually involves only two layers of resources (Layers 1 and 2) we refer to it as a three-layer network, including the auxiliary Layer 3 used for modelling the demands. This allows for the unified interpretation of the constraints.

TLDP-LP (Three-Layer Design Problem - Linear Programme)

---

indices

d=1,2,...,D demands (links of Layer 3)
$j=1,2,...,m(d)$ allowable paths in Layer 2 for flows realizing demand d
$e=1,2,...,E$ links of Layer 2
$k=1,2,...,n(e)$ allowable paths in Layer 1 for flows realizing link e
$g=1,2,...,G$ links of Layer 1

constants

h_d volume of demand d
a_{edj} = 1 if link e of Layer 2 belongs to path j realizing demand d, 0 otherwise

b_{gek} = 1 if link g of Layer 1belongs to path k realizing link e of Layer 2, 0 otherwise
c_g unit cost of link g of Layer 1

variables

x_{dj} continuous flow allocated to path j realizing volume of demand d
y_e continuous capacity of link e
z_{ek} continuous flow allocated to path k realizing capacity of link e
u_g continuous capacity of link g, $\mathbf{u} = (u_1,u_2,...,u_G)$

objective

minimize $\boldsymbol{C}(\mathbf{u}) = \Sigma_g\, c_g u_g$ (7.1.19a)

constraints

$\Sigma_j\, x_{dj} = h_d$ $d=1,2,...,D$ (7.1.19b)
$\Sigma_d\, \Sigma_j\, a_{edj}x_{dj} \le y_e$ $e=1,2,...,E$ (7.1.19c)
$\Sigma_k\, z_{ek} = y_e$ $e=1,2,...,E$ (7.1.19d)
$\Sigma_e\, \Sigma_k\, b_{gek}z_{ek} \le u_g$ $g=1,2,...,G.$ (7.1.19e)

Because of the presence of two upper layers there are two sets of demand-flow constraints (7.1.19b and 7.1.19d), and because of two lower layers there are two sets of load-capacity constraints (7.1.19c and 7.1.19e). The rule is that the capacity of links of the upper layer are realized by means of the path flows in the neighbouring lower layer; this is expressed by the demand constraints. Also, both resource layers (Layers 1 and 2) are networks on their own, so the link capacity constraints must be obeyed in each of them.

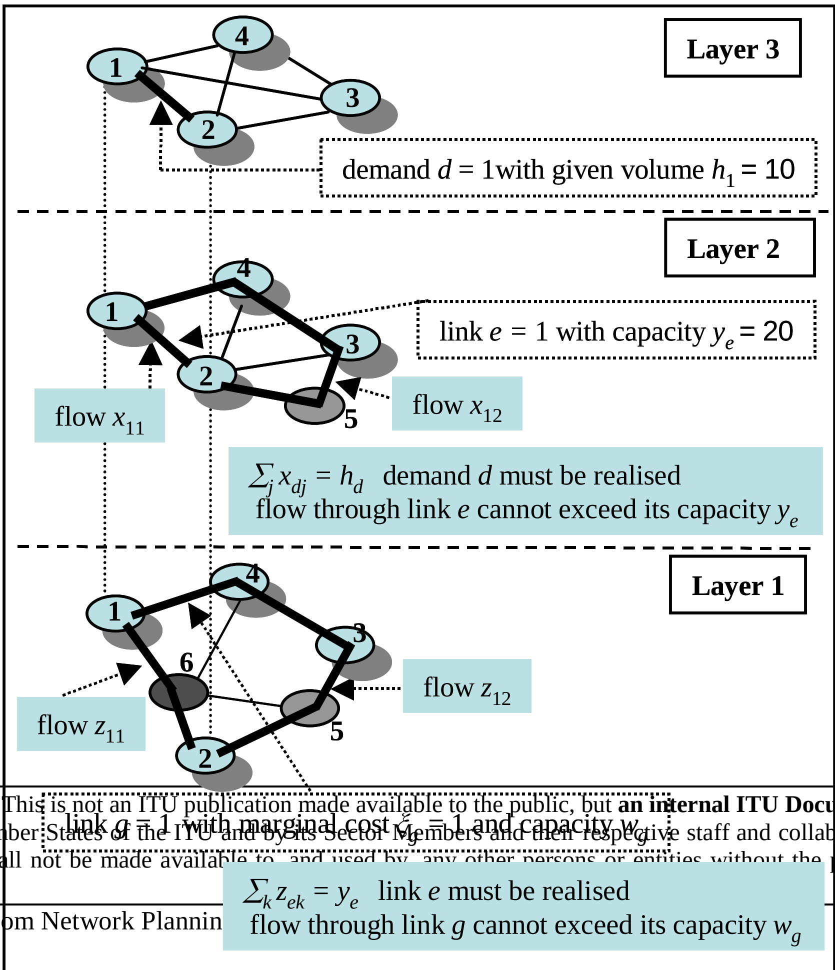


Fig.7.2. Three-layer network

Problem TLDP-LP is illustrated in Figure 7.2 which shows that the volume h_1 of demand d = 1 between nodes 1 and 2 can be realized by means of two Layer 2 flows (direct flow x_{11} and flow x_{12} on path 1-4-3-5-2). Then the capacity of link e = 1 resulting from its load (the sum of all flows through the link) can be realized by means of two Layer 1 flows (flow z_{11} on path 1-6-2 and flow z_{12} on path 1-4-3-5-2). The resulting loads of the Layer 1 links determine their capacities and hence the network cost.

As DP1-LP, Problem TLDP-LP can be easily and effectively solved using the generalized version of the shortest path allocation rule described in Section 7.1.1 for the former problem. (In fact in can be generalized to arbitrary number of layers.)

In practice, certain requirements on links' modularity in one of the layers, or in both layers can be imposed, as well as on the integral flows. The IP problem resulting from such full integrality requirements is obtained when the constraints

$\Sigma_d\, \Sigma_j\, a_{edj}x_{dj} \le My_e \quad e=1,2,...,E$ (7.1.19f)
$\Sigma_e\, \Sigma_k\, b_{gek}z_{ek} \le Nu_g \quad g=1,2,...,G$ (7.1.19g)
all variables are non-negative integers (7.1.19h)

are used. Above, M (Layer 1) and N (Layer 2) are link capacity modules. The integrality of variables makes the considered problems NP-complete.

The next generalization is the counterpart of the single-layer allocation problem AP1-LP (7.1.8) for the case of two layers of resources.

TLAP-LP (Three-Layer Allocation Problem - LP)

constants

h_d volume of demand d
a_{edj} = 1 if link e of Layer 2 belongs to path j realizing demand d, 0 otherwise
b_{gek} = 1 if link g of Layer 1belongs to path k realizing link e of Layer 2, 0 otherwise
u_g capacity of link g

variables

x_{dj} continuous flow allocated to path j realizing volume of demand d
y_e continuous capacity of link e
z_{ek} continuous flow allocated to path k realizing capacity of link e

constraints

$\Sigma_j\, x_{dj} = h_d \quad d=1,2,...,D$ (7.1.20a)
$\Sigma_d\, \Sigma_j\, a_{edj}x_{dj} \le y_e \quad e=1,2,...,E$ (7.1.20b)

$\Sigma_k z_{ek} = y_e$ e=1,2,...,E (7.1.20c)
$\Sigma_e \Sigma_k b_{gek}z_{ek} \le u_g$ g=1,2,...,G. (7.1.20d)

Note that in TLAP-LP capacities of the Layer 1 links are fixed, whilst the Layer 2 links' capacities are variables. TLAP-LP as an LP problem and can be solved accordingly. In general, there are instances of TLAP-LP with only bifurcated feasible solutions.

Observe that in the considered case, modularity of the link capacity variables in Layer 2 can be required (and possibly integrality of the flow variables in both layers). The IP problem resulting from such full integrality requirements is obtained when the constraint

$\Sigma_d \Sigma_j \delta_{edj}x_{dj} \le My_e$ e=1,2,...,E. (7.1.20e)

is used.

7.1.3.2. Restoration design for three-layer networks

In this paragraph we present an example of a restoration design three-layer problem. The concerns designing a three-layer network robust to failures, where flows of Layers 1 and 2 are assumed to be reconfigurable, and the reconfiguration is unrestricted. As will become clear in a while, this assumption implies that the capacities of the links of the upper layer (Layer 2) are flexible, and in general situation-dependent. Links of the lowermost Layer 1 are not flexible: what can only happen is that a part or entire of their nominal capacity can be lost in a failure situation.

TLRDP-LP (Three-Layer Restoration Design Problem - LP)

---

indices as in TLDP-LP (7.1.19), and
s=1,2,...,S failure-demand situations (including the nominal state)

constants
h_{ds} volume of demand d in situation s
a_{edj} = 1 if link e of Layer 2 belongs to path j realizing demand d, 0 otherwise
b_{gek} = 1 if link g of Layer 1 belongs to path k realizing link e of Layer 2, 0 otherwise
α_{gs} fractional availability coefficient of link g of Layer 1 in situation s ($0 \le \alpha_{gs} \le 1$)
c_g unit cost of link g of Layer 1

variables (all variables are continuous and non-negative)
x_{djs} flow allocated to path j of demand d in situation s
y_{es} capacity of link e in situation s
$\boldsymbol{z_{eks}}$ flow allocated to path k realizing capacity of link e in situation s
u_g capacity of link g

objective
minimize $\boldsymbol{C}(\boldsymbol{u}) = \Sigma_g c_g u_g$ (7.1.21a)

constraints

$\Sigma_j\, x_{djs} = h_{ds}$	$d=1,2,...,D$	$s=1,2,...,S$	(7.1.21b)
$\Sigma_d\, \Sigma_j\, a_{edj}x_{djs} \le y_{es}$	$e=1,2,...,E$	$s=1,2,...,S$	(7.1.21c)
$\Sigma_k\, z_{eks} = y_{es}$	$e=1,2,...,E$	$s=1,2,...,S$	(7.1.21d)
$\Sigma_e\, \Sigma_k\, b_{gek}z_{eks} \le \alpha_{gs}u_g$	$g=1,2,...,G$	$s=1,2,...,S.$	(7.1.21e)

The demand-flow constraints assure that the situation-dependent demand volumes h_{ds} (which are equal to the capacities of the corresponding Layer 3 links) are realized by the situation-dependent flows in Layer 2 (constraints (7.1.21b)), and that the demand imposed on Layer 1 and specified by the Layer 2 links' capacities y_{es} is realized by the situation-dependent Layer 1 flows (constraints (7.1.21d)). The load-capacity constraints (7.1.21c) and (7.1.21e) take care about the feasibility of flows in Layers 2 and 1, respectively, i.e. they assure that the situation-dependent loads of links do not exceed their capacities. Observe that although for any optimal solution to TLRDP-LP (which is an LP problem) in constraints (7.1.21c) equalities will always hold, this is not the case for constraints (7.1.21e). In the latter case come links may not be saturated in some situations even for an optimal solution.

Problem TLRDP-LP is illustrated in Figure 7.3. You can notice that the volume h_{ds} of demand d in situation s can be realized by means of two flows x_{d1s} and x_{d2s} in Layer 2. In the considered situation two of the Layer 1 links fail totally, so the capacity y_{es} can be realized only by flow z_{e3s} in Layer 1 since the two remaining flows, z_{e1s} and z_{e2s}, use failed links.

As before, many other valid variants of the problems presented in this subsection can be considered. Besides such elements as modular link capacity, modular flows or single-path routing, the extensions to more layers, the use of different restoration mechanisms, etc., can be taken into account.

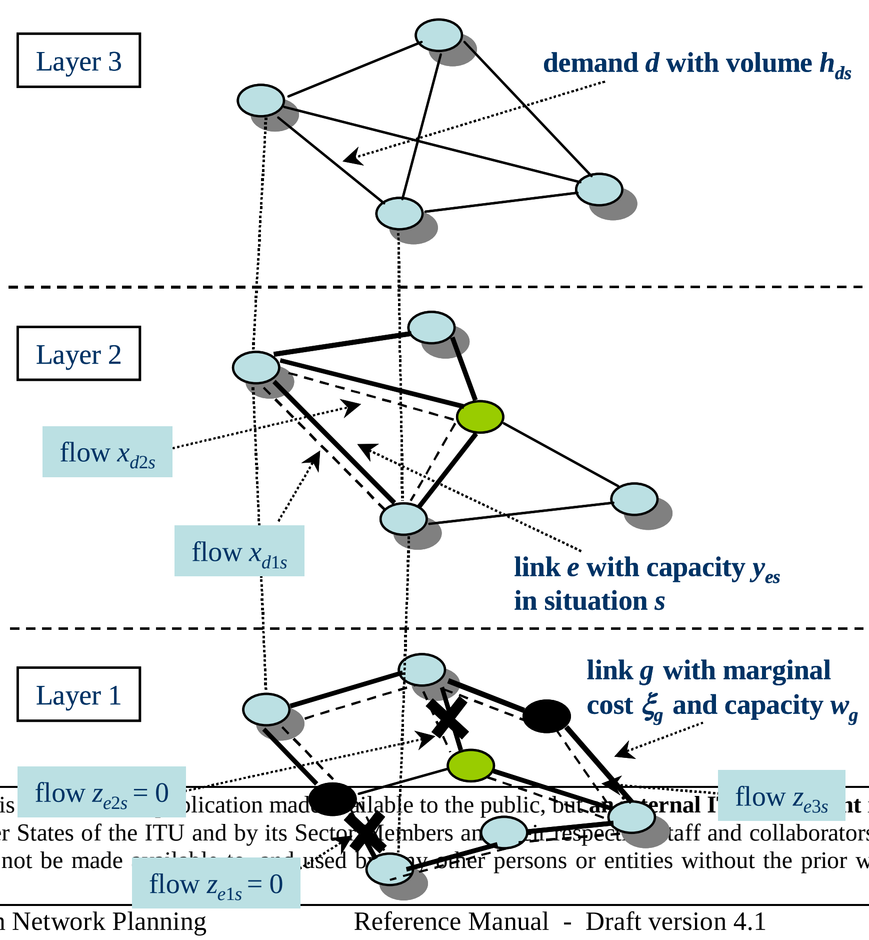


Fig.7.3. Three-layer network with failures

7.2. Access Network

The role of access network is the transfer of services originating at the core network to user terminals and vice versa. Access networks (AN) are the most costly part of the network (typically 70-80% of the overall network cost). This factor imposes a strong pressure on the optimization in AN planning and design. Unfortunately because of many factors of technological and economical nature, and despite many efforts [7.2.1], [7.2.2], [7.2.3], [7.2.6], [7.2.26], there is no unified and efficient approach to access network design, planning and dimensioning. Generally the planning methodology described in chapters 7.4 and 7.3 can be tailored to access networks planning. In practice, due to multiple constraints, even simple AN planning tools handled by experienced experts are of practical value.

7.2.1 Key factors and constraints in access networks deployment

Typically the access network design contains a large number of different problems. During the access network planning process the network optimization procedure has to be supplied by a number of input parameters. Some of them come from operator’s service oriented targets, some are imposed be the selection of access network technology, some are area dependent, and some are of economic meaning.

Typically, the factors that have to be taken into account in access network planning include:

- the service portfolio to be handled by the designed access network (narrowband, broadband) – in fact specific access network solutions can impose limitations on the set of offered services;
- access network technology (copper, fiber, coax, wireless);
- the assumed strategy of evolution of the already installed access network;
- traffic demands forecast;
- geographical subscribers dislocation and the area type on which AN is to be installed (rural, urban);
- segmentation of subscribers in the business context (large business, small business, residential customers);
- topological constraints of access network solutions (star, bus, tree);
- greenfield approach or upgrade of the existing (legacy) AN infrastructure;
- regulatory issues;
- time framework required to access network deployment, incremental deployment possibility;
- the relative cost of different access network technologies (CAPEX);
- the access network operation and maintenance cost (OPEX);
- the compatibility with already installed solutions;
- assumed access network availability/reliability;
- the overall cost of AN deployment.

The relative importance of a specific item from the above list is operator or project dependent. In some cases decision about the preferred choices should be taken at the start of the access network planning process, and in some cases project specific constraints has to be taken into account.

The operator service offer has strong impact on the technology choice for AN. The classical, narrowband AN solutions are designed for telephone voice services.
Broadband solutions require much more bandwidth and should be able to handle high quality video communication of point-to-point, point-to-multipoint or distributive nature for example for broadcasting of video services. The service convergence is seen as key factor in modernization of access of networks in developed countries. This convergence is often referred as Triple Play i.e. the convergent network operator is able to offer Internet services, VoIP and VoD/TV services using the same network infrastructure. In practice the operator should has the possibility to offer a rich service mix, which has to be considered during the network planning and dimensioning process.

The complex service model makes the broadband access network design process much harder than in narrowband case. The traffic optimization for broadband services may require the placement of active/service nodes (servers and so on). That approach has a strong impact on network planning.

An overview of the factors and constraints that have to be considered during access network planning is presented in the following sections. As it was stated before it is the operator’s role to assign the weights to the mentioned factors according to its preferences and the project specific constraints.

7.2.2 Access networks - technology specific issues

There are many technological solutions of access networks, and to every solution appropriate design and network planning methods has to be chosen and applied. There is a fundamental difference in planning methodology between wired and wireless AN solutions. In wireless systems not only the capacity is a subject of optimization but also the radio coverage. A specific kind of planning has to be applied to mobile networks, in which the users mobility has to be taken into account.

7.2.2.1 Impact of the physical layer on the access network design

7.2.2.1.1 Wired access solutions

The wired access networks are classics of the fixed access. Typically two types of medium are used for access networks – copper wires and optical fibres.
In the last years access networks have evolved significantly and the technology progress made possible the use of cable TV networks (so called hybrid fibre-coax solution) and power grid networks as access solutions for telecommunications and data services (see chapter 61.2.1). In these cases the mentioned access network functionality is provided as a kind of overlay to the existing infrastructure.
At present, in most fixed access networks there is a combination of the cooper part and the optical fibre based part. Such hybrid solution makes the planning more complicated than in the homogeneous case. The fibre part typically uses point-to-point, ring or tree topology,

while the copper part is based on the star topology with dedicated wires to every subscriber. Such an approach is economically viable and provides possibility of smooth and scalable introduction of fibres into AN area. Because the conversion between optical and electrical signals introduces a significant cost the number of conversions has to be minimized.
The common problem with wired AN deployments is a long time framework and the cost of the cable installations (civil works). The problem is especially important in urban areas where the digging is hard to perform (for example in city centres). The possibility of the use of existing ducts (the upgrade case) is of greatest importance, because it reduces the cost of the cabling infrastructure significantly. Thus the reuse of existing ducts has the highest priority in network planning. The planning tools should use all the information related to the existing infrastructure (ducts, masts etc.) in order to limit civil works and to speed up the access network deployment, which is another important factor for AN deployment.

7.2.2.1.1.1 Copper networks

Twisted pairs based copper networks are typically used as a basic solution for access networks. In access networks based on the pure copper infrastructure a star network topology is common. The drawback of the cooper cabling is the limited bandwidth and the susceptibility to electromagnetic interferences. The subscriber line quality is decreasing with the subscriber loop length (signal attenuation, interferences) and the maximum access network range must be taken into account during network planning using this medium. The most important constraint that has to be taken into account during copper network planning is thus the length of the local loop, which in a typical twisted pair based access solution is a dedicated medium (non shared).
Because of broadband services there is also a requirement to implement or rebuild the legacy access infrastructure with gradually increasing the part based on the fibre (hybrid copper-fibre solutions).
The use of unshielded twisted pairs for broadband data transmission is possible due to the use of xDSL modems. In this approach the bandwidth bottleneck of the network part based on the cooper medium is removed. The shortening of the local loop is in this case of great importance – shorter loops provide higher data rates. This approach is mostly applied to existing infrastructure as a short or medium term access network upgrade. The specific constraint of this approach is the placement of the broadband concentrator at the centre of the copper cabling star topology.

7.2.2.1.1.2 Optical networks

Last years the optical fibre technology has been successfully deployed for the long-distance communication and now is the technology of choice in this area. The most important fibre properties, which make them so popular, are: the capability to transmit information at very high bit rates, insensitivity to electromagnetic interferences, which provides very low bit error rates, high reliability, low the optical signal attenuation and a small diameter and weight of the optical cables. The cost of optical fibres has reached an acceptable value and is no more a prohibition factor for the deployment of fibre based access networks [7.2.5]. In practice, due to high bandwidth, fibres are used as a shared communication medium.
Dedicated point-to-point fibre access networks are too costly, and except Ethernet based solutions (for example the MAN case, which is described in section 6.1.1), ring or tree topologies are widely used for all-optical and hybrid copper-fibre access solutions.

In optical ANs critical is the number of splitting points, due their significant cost and the introduced signal attenuation. In hybrid optical-copper based solution the cost of electro-optical transceivers is also important, so during planning their number should be minimized.

Planning of optical networks is less critical in the context of dimensioning - the cabling introduces no bandwidth limit and a quite simple bandwidth increase can be done by the use of so called Coarse Wave Division Multiplexing technology (CWDM) designed especially as a cheap WDM technology for broadband access and metropolitan networks.

In typical, copper based ANs with a dedicated medium there was generally no protection. In fibre based ANs of tree topologies, due to high possible traffic aggregation levels, the protection of links should be considered, however such protection is not the operators present practice.

7.2.2.1.2 Fixed wireless access

The main advantage of the radio access is fast network deployment and potentially low initial investment (CAPEX). In opposite to CAPEX, OPEX of wireless networks is generally more significant than in wired case due to the cost of the radio license. The main disadvantage of radio networks is relatively small bandwidth.

The wireless AN approach (Fixed Wireless Access – FWA, characterized in chapter 6.1.2.5) has to cope with radio propagation problems in the access system specific band, which have influence on the system capacity and the quality of radio links. In order to increase the system capacity it is possible to use mechanisms like: multipath mitigation, directional antennas, advanced link quality improving mechanisms and so on.

The planning target in FWA case is to find the localization of base stations, which will provide the requested coverage and traffic capacity. It is realised by appropriate allocation of radio channels radio, selection of antennas etc. The transmission between base stations is typically provided by the core network (via radio point-to-point links or fibres). It is a common vendor practice to deliver radio network planning tools for their FWA systems, thus no generic planning rules can be provided. The description of mobile network planning and cell dimensioning in WCDMA case is presented in chapter 7.4.

7.2.2.2 Impact of networking technology on the access network design

7.2.2.2.1 TDM networks

In TDM networks, which were designed mainly for voice services, an important requirement is to determine the number of voice circuits needed for a given call traffic demand while meeting a certain grade-of-service (GoS). Fortunately, there is an elegant result due to Danish mathematician A. K. Erlang that he developed almost a hundred years ago. In this context, the demand is often referred to as offered traffic or offered load, and is given in the dimensionless unit, Erlang. The traffic offered can be best characterized in Erlangs by the following product:

*Offered traffic (a) = Average call arrival rate * average call duration time*

Then, for **c** circuits, the call blocking probability is given by the following Erlang-B loss formula:

$$E\,(a,c) = a^c / c! \,/\, (\Sigma_k\,(a^k / k!)) \qquad (7.2.1)$$

where the summation is from k=0 to c. It may be noted that this formula is developed under the assumption that the arrival traffic follows a Poisson Process, while the result being insensitive to the actual statistical distribution of the call duration time. It is easy to see that this result is applicable to a network link.
Often, we're interested in determining the number of circuits if the offered traffic and the acceptable grade-of-service (in blocking probability threshold) is given. It is not hard to see that a simple iterative test method can be employed using (7.2.1) so the proper number of circuits required can be determined. A commonly used value for the grade-of-service is 1% call blocking probability. Thus, for 100 Erl of offered traffic, and for 1% call blocking probability, we can iteratively use Formula (7.2.1) to determine that 117 circuits are needed.

Nowadays, the teletraffic software includes the Erlang-B calculation. The interested reader may also use the freely usable web-based Erlang calculator available at http://www.erlang.com

In many cases, we're primarily interested in busy-hour offered traffic and its impact on performance. In many networks, the average Erlang offered traffic per customer can be determined from operational measurements. For example, if average offered traffic per customer is 0.03 Erl, then for an offered traffic of 100 Erl, the average number of customers that can be supported with 117 circuits is a little over 3,300 customers (An astute reader may note that Eng-set model may be more appropriate since a finite population is eventually considered due to 0.03 Erl per subscriber; however, since the population size is large, the Erlang-B loss formula still turns out provide a very good approximation.).

In many practical networks, the offered traffic is hard to estimate. Only, measured traffic, or carried load can be determined. Interestingly, it can be shown that the carried traffic is nothing but the average number of busy circuits. Thus, for a measured carried load a' and given number of circuits, c, we have the following relation

$$a' = a\,(\,1 - E\,(a,c)\,) \qquad (7.2.2)$$

The above equation can be iteratively solved to determine the offered traffic, **a**.

To summarize, the advantage of the Erlang-B loss formula is that it gives us the relation between offered traffic, call blocking, and the number of circuits. This model is applicable to any single link design problem; in particular, the formula can be applied for designing access network links, both aggregated wired and wireless.

7.2.2.2.2 ATM Networks

The Asynchronous Transfer Mode (ATM) network is a session oriented cell switching network, capable of carrying many different services. It is based on virtual connections (VC), and each of the virtual connection of the ATM network is characterized by a set of parameters such as maximum required bitrate, burstiness and different quality requirements (QoS) with respect to allowable delay and cell loss rate. In opposite to the IP network, in the ATM network call (session) admission mechanisms are implemented in order to prevent the network congestion. A set of services attributes for ATM-based networks has been defined by ITU-T. The following list consist the most important ones in the context of the ATM network design and dimensioning.

- The type traffic, which has to be handled by the network: Constant Bit Rate (CBR), Variable Bit Rate (VBR), etc. When the access network performs the traffic aggregation of VBR connections the aggregation gain is obtained. If no concentration of the traffic is used or all of the virtual connections are considered as CBR then the TDM dimensioning rules can applied.
- The traffic attributes of ATM networks specify the character of the traffic and can be represented as:
    - Peak Cell Rate (PCR)
    - Cell Delay Variation Tolerance (CDVT)
    - Sustainable Cell Rate (SCR)
    - Burst Tolerance (BT)
    - Minimum Cell Rate (MCR)

  The CBR traffic is described by Peak Cell Rate attribute only.

- Establishment of communication (dynamic, static). This attribute should be taken into account in the access network dimensioning process.
- The traffic symmetry level (traffic: unidirectional, bi-directional symmetric, bi-directional asymmetric).
- Communication configuration (point-to-point, multipoint, broadcast). In most applications only point-to-point links are used.
- Quality of Service (QoS) defines the transport quality of ATM based access networks as well as the availability level. The QoS can be specified by the following parameters:
    - Maximum Cell Transfer Delay (MCTD)
    - Mean Cell Transfer Delay (Mean CTD)
    - Cell Delay Variation (CDV)
    - Cell Loss Rate (CLR)

Initially the ATM technology was used as a core network technology for IP and Frame Relay access/edge networks. At present there are also popular solutions as xDSL, which are native ATM access solutions, thus there is possibility to construct end-to-end native ATM networks. The methodology, which is used for designing and dimensioning of ATM core networks, can be successfully adopted for broadband, full-service ATM access networks.

7.2.2.2.3 Frame Relay Networks

Frame Relay (FR) is a medium-speed (up to 2 Mbps) connection-oriented packet data transfer service. It uses permanent virtual connections (PVCs) to establish logical connections between terminals to provide end-to-end FR services. The typical parameters that describe the FR connection are: the interface bit rate, Committed Information Rate (CIR), which describes the guaranteed mean bitrate and the total information transfer delay.
It is a common operators practice to use ATM as a core network for the Frame Relay traffic aggregation and to use ISDN layer 2 technologies in order to reach end-users (HDSL). So, the

design of Frame Relay networks combines the design of the core/edge ATM networks and the classical narrowband design for the ISDN network.

7.2.2.2.4 Native IP networks

The design and dimensioning of IP networks is a troublesome process due to lack of native IP long-range access solutions and the lack of user models. The best effort approach is at present applied to all Internet services. No resource reservation is used and the congestion is a typical behaviour in IP networks. These factors make hard or impossible to use advanced mathematical tools for IP networks dimensioning. There is however a common opinion that in Internet the access part is the network bottleneck. As Internet access widely used are all existing narrowband networks (PSTN, ISDN) and broadband networks like Frame Relay or ATM (including xDSL) . In opposite to the generic telecommunications architecture in the IP network there is no strict distinction between the access and the core.
Sometimes the placement of edge routers makes such separation.
As a native IP access technology the metropolitan networks (MANs) can be considered. The design and the dimensioning of mesh-like MANs is similar to the design and dimensioning of IP core networks. There are ongoing works on introducing of different class of services using advanced IP network mechanisms (the DiffServ model). At the time of writing of this document there are no indications related to dimensioning of such a network.

7.2.2.3 The impact on density of population on network design

The density of population has important influence on the access network design. Typically the area which has to be covered by AN can be defined according to the following parameters:

- the number of potential users (business and residential),
- the area size to be covered (km^2),
- the type of urban infrastructure on the AN area (residential buildings, multi-flat buildings and so on).

A common practice in modern designs of access network is the use of GIS information, which should include all required by planning tools information about the density of population and its distribution within the interested area. Some assumption about future network development has to be taken into account a priori.

Using density of population as the main criterion we may identify the following types of areas (EURESCOM):

1. Downtown area, which is characterized by a short average copper loop length, typically 500 m - 1.000 m, corresponding to density of 9,200 -2,300 subscribers/km^2.
2. Urban area, where the average copper loop length is in range 1,000 - 2,000 m. It corresponds to 2,300 -570 subscribers/km^2.
3. Suburban area, 2,000 - 3,000 m average copper loop length. It corresponds to 570 -250 subscribers/km^2 respectively. The subscribers are located typically in single house dwellings.

4. Rural area. This area is characterized by average copper loop longer than 3,000 m and corresponds to less than 250 subscribers./km^2. The subscribers are located in single house dwellings.

Radio access deployment planning requires some additional information with respect to the average number of buildings, the average number of flats per building, the average number of floors per building and the average number of flats per floor.

7.2.2.4 Possible access networks evolution strategy

All installed access networks should take into account future growth in the sense of number of served customers and services. Another aspect of the evolution is a smooth migration towards broadband, typically fibre based networks. The evolution toward the target access network should be studied techno-economically [7.2.4], [7.2.27].

7.2.2.5 The time to deploy target access network

The time to deploy of access networks is in many cases of premium importance. In classical wired systems the most important factors responsible for long installation times are civil works and the installation of cables. Using of trenchless approaches or using of existing ducts gives a possibility to reduce the AN installation time significantly. So in order to speed up the network installation all the existing infrastructure, which can be used for cable installations, should be evaluated carefully.
Radio access networks solutions are solutions of choice when the time to network deploy is the most important factor. They however possess some limitations (low BER of channels, limited bitrate) and generally they are not the solution of choice in a long-term perspective.

7.2.2.6 Access Networks availability

The core and the access networks are involved in providing services to end-users. Typically the less reliable is the access part of the network, and typically there is no redundancy and the link protection within the access area. A failure of an individual link has impact on one or several subscriber lines only, which justifies the lack of AN protection mechanisms.
The high reliability of the access network is however important in the context of OPEX reduction and customer satisfaction. The reliability level of the access network can be typically increased by the deployment of advanced diagnostics mechanisms built in AN solutions, which are able to provide detailed and early reporting of cabling and devices failures. Of a special importance are failures of AN links, which carry the aggregated traffic. In such a case (wireless or wired) protection mechanisms should be installed in order to eliminate a single point of failure.
The important factor in the context of active access devices is the high availability of the power supply and the installation of batteries as back-up sources of electricity is recommended. The placement of access devices should take into account environmental conditions.

7.2.2.7 Greenfield access network installation versus network upgrade

There are typically two different scenarios of the access network deployment. The first one is a greenfield approach i.e. building of the access infrastructure from scratch, whilst the other

one is an upgrade of the existing infrastructure. The network upgrade can be motivated by the increase in number of subscribers, the introduction of new services, the conversion a narrowband access infrastructure to a broadband one or modernization of existing access solutions using different access technology than previously installed. Both approaches differ significantly in the context of planning, dimensioning, deployment time frame and the cost. The cost of the upgrade of access networks is sought minimized, constrained by dependability and traffic requirements, i.e. QoS requirements, from each end-user.
In the context of planning the upgrade approach introduces many additional constraints related to maximum reuse of existing infrastructure for new subscribers, and potential rebuilding of the existing infrastructure while adding new lines. Such the network modernization can be performed via gradual installation of fibres in the access area and shortening of the cooper part of the network (FTTx solutions). The upgrade policy is operator strategy dependent and should be made as rare as possible via significant and over-dimensioned steps [7.2.23], [7.2.24], [7.2.25].

In case of increased demands in wireless access networks to provide the higher network capacity it is recommended to upgrade already installed base stations, rather than installing of new ones. This approach typically provides the best economy.

7.2.2.8 Access network deployment cost

The access network deployment cost is the most significant part of the network overall costs, thus minimizing this cost is of premium importance and the problem was a subject of many studies [7.2.30], [7.2.31], [7.2.33], [7.2.32]. The deployment cost of fixed access networks can be divided into the part related to civil works, the cost of the cabling, the cost of the active access equipment and other costs. The mutual relation between mentioned costs are country, area type and system dependent and should be checked in the preliminary phase of the network design.
The average civil works cost related to cable installations depends on the surface type and the method used for cable deployment. The labour cost is typically country and even region dependent. The cable deployment techniques include digging trenches in asphalt roads, digging trenches in tarmac-free areas, no-dig cable installation techniques and the suspension of aerial cables.

Generally there are four main types of cable deployment:

1. Digging trenches in asphalt roads. This deployment technique is the most expensive one.
2. Digging trenches in tarmac-free area. In this case the cost of the civil work can be reduced about 2 times, comparing to the previous case.
3. Digging of shallow trenches with direct burial of cables in the tarmac-free area. This approach provides about 2 times reduction of the cost of civil works than in the previous case.
4. Suspension of aerial cables. The is the fastest and the cheapest technique of cable installation.

The use of no dig techniques in telecommunications had been for many years limited to the use of railways/metro and roads for the deployment of the cabling infrastructure. From several years the use of ducts created for other purposes, for example sewer ducts is also possible for installation of fibre-based cables. In some cases due to the use of no dig techniques the ducts can be installed underground without digging deep trenches. In such tunnels fibres are subsequently pulled.

In wireless access systems the typical approach is to maximize the number of subscribers served by each base station during the first stage of the radio network deployment.

The minimizing of number of different access networks devices and configurations, plays essential role in the overall cost management and speeds up the network deployment

7.2.2.9 Access networks OA&M costs

That implementation of access networks involves operations, maintenance and administrative running costs (OPEX) for the network operator.
The operations costs are generated by operational staff, which operate and manage the access networks. Using modern access technologies with intrinsic enhanced diagnostics the maintenance cost can be significantly reduced in comparison to older ones. Thus network modernization can be motivated by the reduction of the OA&M costs.

7.2.2.10 Market oriented issues

For operators the speed of the network and services rollout is a critical factor in the competing market. The return of investment (ROI) in case of access networks takes typically many years, so the deployment strategy of the network and services is extremely important. The network operator has to make a proper decision that should shorten the ROI time, but which will also address all present and future customers needs and enable a simple and cost-effective network growth in number of access network terminations (customers). To make the correct choice a market analysis, which will assume the potential interest in offered services and ARPU is required. The widely used approach, which makes that type of analysis more accurate, is to divide the subscribers into several categories. Each of the categories should describe customer’s potential interest in specific services and the potential intensity of the use of services. The following classification of customers is common:

- Business customers divided into two groups: Small and Medium-size Enterprises (SME) and large businesses.
- Residential customers (next divided into urban, suburban and rural)

The more detailed statistical information about subscribes can be also processed by the marketing tools. Such statistical information about the average subscribes income value, subscriber’s occupation and education, demographic characteristics, (e.g. age of residents and family composition), etc. can help in the estimation on the money amount, which the customers might spend on new services.

In order to keep the initial costs at the low level operators may introduce basic services with limited quality in the first place. The more advanced and high quality services can be introduced later in a gradual manner.

The increasing competition has not been until recently taken into account as a factor important in the access network design. Latest technology progress has made possible of the use of cable TV networks, power-lines or data transmission networks (in a Voice over IP manner) for telecommunication services. This situation has led to the reduction of prices of the telecommunication services and positively stimulated the market by making a further increase of demands [7.2.23], [7.2.34]. This phenomenon, called “price elasticity” has to be taken into account in demand analysis. Mobile communication systems may in a sense cannibalize classical telecommunication, reducing the usage and ARPU of the wireline access.

7.2.3 Access network planning methodology

The deployment of access networks with a reasonable cost requires planning and implementation of the design in a well-organized manner. Generally there are two basic approaches to the network planning.

The classical approach, which is currently used for the planning of narrowband networks, is based on the forecast of demands and the dimensioning of the network in order to minimize the required network resources (cabling, equipment). This approach is oriented towards the minimization of the investments required to provide a given amount of resources and is very suitable for situations where the traffic and number of subscribers increases smoothly.

The new transport technologies, like DWDM have made the optimization of resources less critical - the operators have obtained a technology, which provides the ability to rapid and substantial increase of the network capacity at the moderate cost level. On the other hand the growing popularity of the Internet has led to the fast and substantial traffic increase in short time frame and problems with reliable forecasting of demands. In this situation, the uncertainty in the demand forecast and the relatively low cost of the bandwidth have limited the applications of the traditional planning - it produces a network that is optimized for a given demand matrix, but it does not guarantee that the network is ready for upgrade in case the demands were underestimated or growing [7.2.28], [7.2.29].

The second design approach is to build a network that can provide a considerable excess of resources (transport, switching or routing) thus providing nearly seamless network upgrade at the price of higher initial cost.

The presented network design approaches can be combined together in order to find the solution that better fits the operator's needs and strategy.

Network planning should address aspects related to the network evolution. Traditionally the planning activities can be divided into Short Term Planning (STP), Medium Term Planning (MTP) and Long Term Planning (LTP) as described in chapter 2.4.1.

Short term planning is realized in a response to present needs and generally should be applied for solutions characterized by a short deployment time, for example for the wireless access.

Middle term planning takes into account the network upgrade in the context of capacity upgrading of network links and nodes.

LTP’s objective is to design and dimension the network in a long time-frame [7.2.31].

The relation between LTP, MTP and STP is extremely important in network evolution - all planning methods should act coherently. The coherency is in practice hard to obtain and operators of the access networks are often applying a more pragmatic and simpler approach for the network design.

A planning process of access networks is strongly dependent on the network technology, network architecture, offered functionality and resource allocation strategies. As more technological solutions are available on the market, the operator has more options for the design of a suitable and cost-effective access network. Unfortunately it makes the access network design process more complicated.

7.2.4 Mathematical foundations of access network planning

It is possible to formulate mathematically most of the network optimization problems. The mathematical formulation of the problem has to use appropriate network models, necessary simplifications and general assumptions. In practice there is hard to find the optimal solution for the network, but the obtained solutions can be close to the optimum. The applicability of a particular solving technique is in close relation with the size of the problem, which depends on the type and number of input variables and constraints, and on the type of the used cost/energy function, which is minimized during the optimization process [7.2.7]. Some of input variables are to be subject of change in a long or medium term.

In order to keep the network optimization process reasonable complex it is advisable to:

- simplify the planning problems via reduction of number of input parameters and creation of appropriate network models;
- the acceptance of the suboptimal planning - the uncertain growth of the traffic and the number of customers justifies it.

Choosing appropriate design techniques is non trivial and generally cannot be done automatically. Every access network problem is unique and requires operator’s knowledge about the overall network planning methodology and optimization tools. Some of access network solutions allow the use of a specific category of planning algorithms only.

There are several models that have been traditionally used in network optimization: flow formulation for multi-commodity flow model, path-flow formulation and path formulation. To solve the network optimization problem the following classes of algorithms can be applied:

- mathematical programming - usually based on multi-commodity flow formulation of the capacitated routing problem;
- simulated annealing approach and it's variant of simulated allocation, which is the simulated annealing idea applied for capacitated network planning problems;
- simple heuristic algorithms - these algorithms are applied in order to reduce the complexity of the multi-commodity flow problem. Heuristic algorithms are based on the extraction of practical rules from the way an expert solves a specific problem. These rules may not have a mathematical proof as they are based on good results obtained in the practice.

Table 7.2.1 consists of mentioned above models with applicable optimization algorithms. All they have been described in details in chapter devoted to the core network planning. They are fully reusable in the context of the access network planning. Detailed network optimization models and algorithms description can be found in [7.2.1], [7.2.3], [7.2.4] and [7.2.6].

Table 7.2.1 Summary of the available planning techniques

Models	Optimization Algorithms
Multi commodity flow model	Enumeration

Path-flow model	Branch and bound
Path formulation model	Relaxation techniques
	Marginal improvement techniques
	Simulated annealing [7.2.8], [7.2.9], [7.2.10], [7.2.11]
	Simulated allocation [7.2.12] [7.2.13]
	Tabu search [7.2.14], [7.2.15]
	Stochastic geometry
	Successive smooth cost approximation
	Evolutionary algorithms [7.216], [7.2.17], [7.2.18], [7.2.19]
	Heuristics

7.2.5 A pragmatic approach to access network design

The access network design process is a process that has to cope with many initial assumptions and also with many constraints. Some of them are based on the strategic assumptions while others are related to the existing infrastructure and technological solutions, which have been chosen a priori. The access network solutions, even if belong to the same class differ significantly in technical details and parameters which should be set up during the planning process. So in practice vendors of access networks solutions offer planning tools, which are tailored to their solutions. In many cases relatively simple spreadsheets can be effectively used for network planning. Of a great importance is the use of GIS databases, but such information is typically available in developed countries only.
Heuristic optimization is very popular in the operators practice, due to their practical basis.

7.2.6 Access network planning tools

Network planning tools are offered by the access network solution vendors, planning tools vendors and academic institutions [7.2.19]. Some of them are described in Annex 1 (STEM, NetWORKS, VPIlifecycleManager). This issue was also a subject of several European level projects [7.2.20]. The most important one are: RACE 2087/TITAN project, ACTS 226 OPTIMUM, and ACTS 364 TERA. All mentioned projects were focused on techno-economic analyses of networks and were developed in evolutionary manner (i.e. there is a evolution from TITAN through OPTIMUM to TERA).

RACE 2087 TITAN project developed a methodology and a tool for the techno-economic evaluation for the introduction of new narrowband and broadband services for both residential and SME customers.
The TITAN project started in 1992 and ended in 1996. The developed within this project TITAN network optimization tool is dedicated for techno-economic analysis, and a demand forecast for access network. The tool is based on generalized access network models and utilizes the geometrical model for the calculation of cable lengths and the cost of civil works.
It covers (non exhaustive list) the following aspects:

- the evaluation of access network evolution scenarios based on existing networks and fibre/radio/copper access technologies;
- the comparison of scenarios and strategies for introducing the fibre in the loop (FITL) for residential customers;
- the calculation of the life-cycle cost and the overall system budget.

The TITAN tool is based on the Excel® spreadsheet. It has two operating modes - the main mode and the database mode. The database mode is used for the estimation of the cost of components and services, while the main mode is used for the definition of the network architecture, services and performing all calculations.
The TITAN database contains several sections:

- cost components, containing all component specific information;
- learning curve classes, which define a specific learning curve behaviour;
- volume classes, which define a specific market volume evolution;
- OA&M classes, which define a certain operations, administration and maintenance;
- write-off period class which defines a component lifetimes for calculation of depreciation;
- uncertainty class, which defines the relative uncertainty for the risk assessment.

The objective of another European project, AC226/OPTIMUM was the calculation of the overall financial budget of all kinds of access solutions. In result a network-planning tool has been developed, which takes into account: the system cost, operation and maintenance costs, life-cycle costs and the cash balance of the project. In the OPTIMUM project the network evolution costs are also considered. The tool combines low level, detailed network parameters with high level parameters, which is seen as the key feature of the OPTIMUM methodology and the tool. The structure of the OPTIMUM tool is shown in the figure 7.2.1.

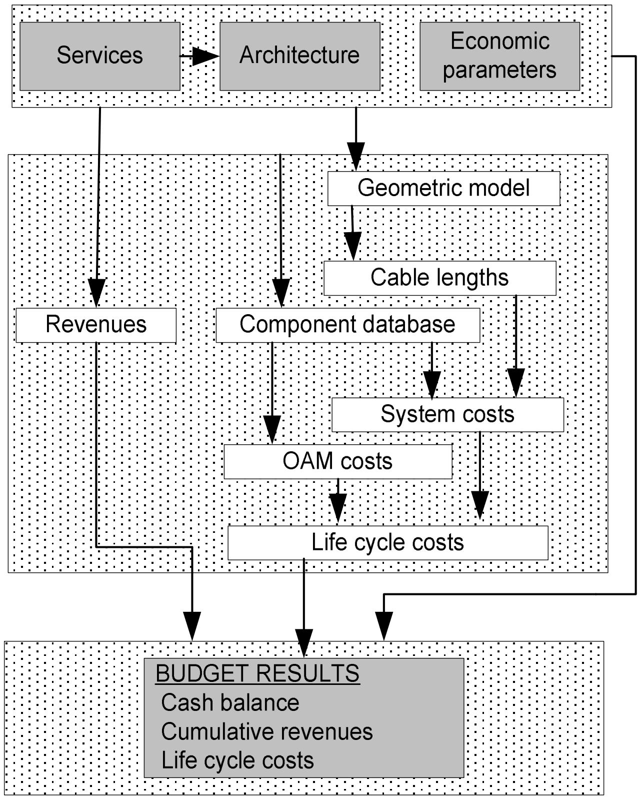


Figure 7.2.2 The OPTIMUM tool structure

TERA was a project realized within the ACTS Programme (1994-2000). Compared to its predecessor, the TERA tool aims at the study of architectures spanning the whole telecom network, not only the access network part (http://www-nrc.nokia.com/tonic/description/bg.htm).

The framework of the TERA techno-economic evaluations is shown in the Figure 7.2.2

The geometric planning model is an optional part of the TERA tool. The model is used to estimate the amount of cable/fibre and ducting required in the network. On the conceptual level the geometric model is a function that takes several inputs such as subscriber density, network topology (star, ring, bus), average cable over length, duct availability, etc. and gives two outputs, which are the total amount of cable/fibre required in the network and the total amount of new duct required. The TITAN network model is a part of the TERA tool.

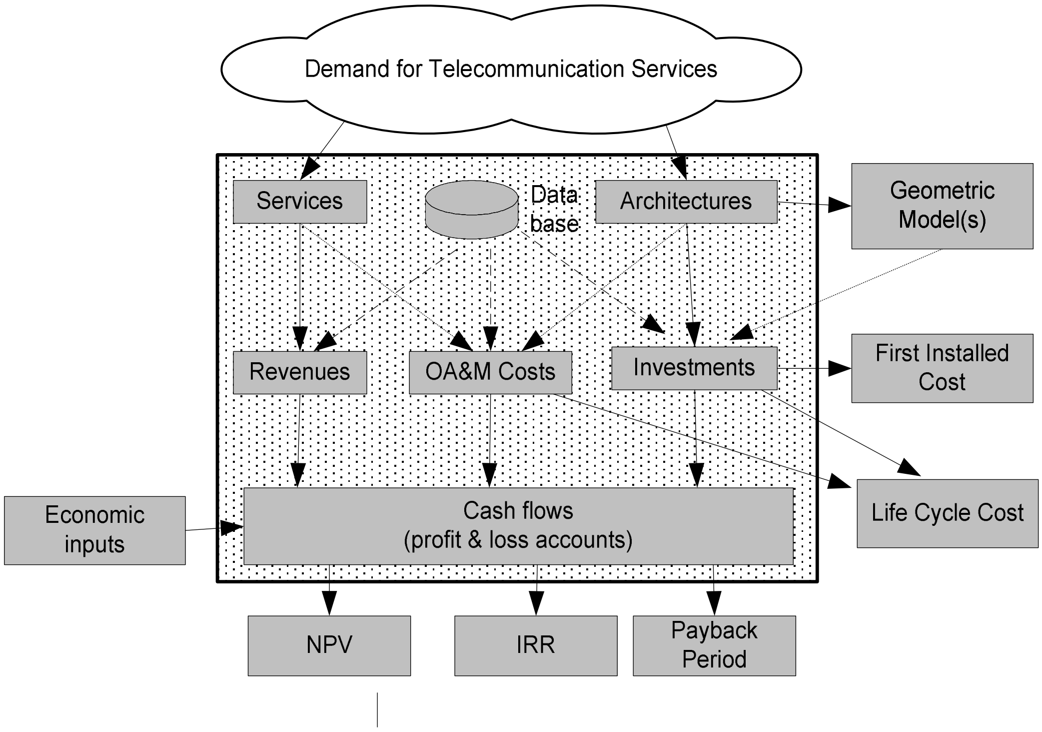


Figure 7.2.2. The TERA information flow.

7.2.7 Example of the access network design algorithm

In this section, we present an optimization model for cost-effective design of a generic access network that can be applicable for both wire-line and wireless access networks.
Simply put, the design problem is: we are to connect N sites through a list of M possible concentrator locations so that the total access network cost is to be minimized. We are given the cost information for connecting each site to each possible concentrator location. Secondly, we are given that any concentrator location, if opened, can handle only up to a certain number of site terminations and that each site needs to be connected to only one concentrator location, means there is only a single access link for a site to the network. Mathematically, we are given the following cost information

c_{ij} := *cost of connecting site i to possible location j (i=1,2,…,N; j=1,2,…,M)*
b_j := *cost of location j, if opened (j=1,2,…,M)*

We have two sets of unknowns, one is the decision (binary) variable that is used for identifying sites to locations

x_{ij} := *1, if site i is connected to location j; 0, otherwise*
y_j := *1, if location j is chosen; 0, otherwise*

We also need another piece of information, the capacity of each location (if opened), given by K_j.
Then, the cost-effective design formulation for the access network design (to determine decision variables, x & y) is written as follows:

$$\begin{array}{lll} \textit{minimize} & \Sigma_i \Sigma_j c_{ij} x_{ij} + \Sigma_j b_j y_j & \\ \textit{subject to} & & \\ & \Sigma_j x_{ij} = 1, & i=1,2,\ldots,N \\ & \Sigma_i x_{ij} \leq K_j y_j, & j=1,2,\ldots,M \\ & x_{ij} = 0/1 & \\ & y_j = 0/1. & \end{array}$$

We now explain the constraints described above. The first constraint $\Sigma_j x_{ij} = 1$ says that one site can be connected to only one of the locations since the decision variable take only the 0/1 values. The constraint, $\Sigma_i x_{ij} \leq K_j y_j$, says that multiple sites can be connected to a location 1) if the location is open ($y_j = 1$), and 2) the capacity of the location is not violated; if the location is not open, so sites are connected. Finally, the cost in the objective function is the cost of both the connectivity and location opening cost.

The above problem is classified as an integer programming problem. While commercial available software such as CPLEX or XPRESS-LP can be used for solving moderate size problems, we have given below an Add Heuristic which is computational efficient.

Algorithm: Add Heuristic

<u>Step 0</u> Select an initial location j' and assume that all sites are connected to this location. Compute total cost $\boldsymbol{F}^0$ *with this configuration. Set* $S_0 = \{ j' \}$ *and iteration count to* $k = 0$. *Set* $c_i' = c_{ij}$ $i=1,2,\ldots,N$. *Let* $\boldsymbol{M}$ *denote the set of locations.*

<u>Step 1</u> For j in $\boldsymbol{M} \setminus S_k$, *do*

$$F^{k+1}_j = F^k_j + \Sigma_{\{i \text{ in } Ij\}} (c_{ij} - c_i') + b_i, \quad \text{where } I_j = \{ i \mid c_{ij} - c_i' < 0 \}$$

<u>Step 2</u> Determine a new <u>j</u> such that

$$F^{k+1}_{j'} = \min_{\{ j \text{ in } \boldsymbol{M} \setminus Sk\}} F^{k+1}_j < F^k$$

If there is no such j', go to Step 4.

<u>Step 3</u> Update

$$S_{k+1} = S_k \text{ union}\{ j' \}$$

and

$$c_i' = c_{ij}' \qquad \text{for } i \text{ in } I_j'$$

Set

$$\boldsymbol{F}^{k+1} = F^{k+1}_{j'} \text{ and } k := k + 1, \text{ and go to step-1.}$$

<u>Stop 4</u> No more improvement possible; stop.

7.2.6.1 An example of planning of a wireline access network

In this section, we discuss the applicability of the models described in Section 7.2.1.

First, for a single wireline access link, link dimensioning is required – this is where the Erlang-B loss formula is applicable when offered traffic estimate is determined, and a grade-of-service is selected. Typically, for wireline networks, an acceptable grade-of-service value is 1% or 0.1% call blocking probability.

Another important problem in wireline access networks is the local loop plant design that connects to the central office switch. In this network design, the goal is to do cost effective design so that an access network layout with concentrators can be used. Interestingly, the access network design model presented in 7.2.1.2 is applicable here.
In order to use this model, we first set the very first index of the location j=1 to be the site of the central office; furthermore, we assume that, for the access design purpose, its site cost, b_1, is set to zero. Note that in actuality, the central office does have a cost. Since, in the access design part, we want this to be a selected site anyway, we need to adjust the location cost accordingly so that this is enforced in the design model. With this special treatment, the access design can be performed which will select a subset of concentrator location in the optimality along with the central office. Obviously, all the concentrator locations are connected eventually to the central office. Pictorially, the eventual design would look as shown in the following figure:

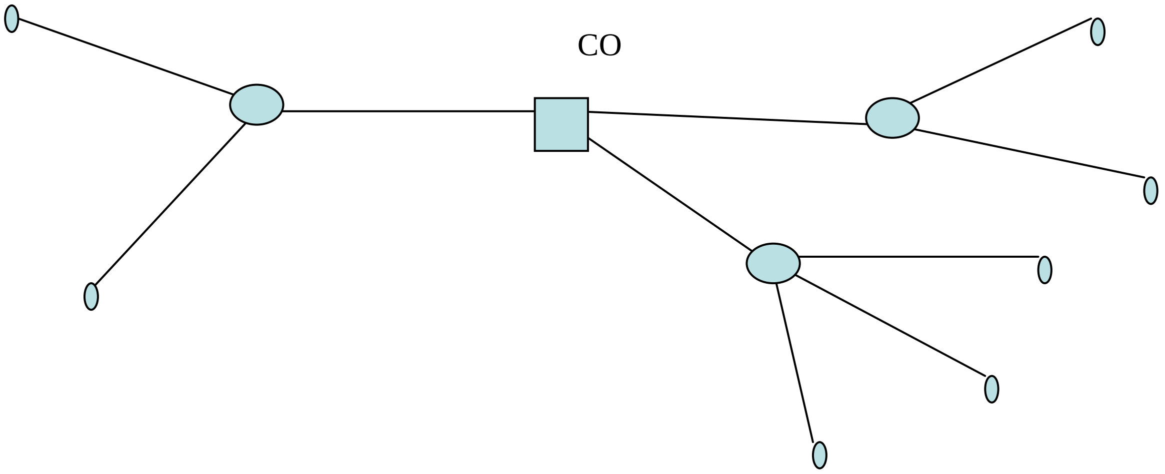


Fig. 7.2.3 Sample wireline network topology

Certainly, this is only a possible access network design layout. More complex design can be developed where there is a tertiary network starting from one of the concentrator nodes (acting as core connect node like the role of the central office, without its functionality) which itself can be designed using the above access network design model for the tertiary access network design.

7.2.6.2 Access Planning Example for Wireless Access Networks

The access network design model is also applicable in wireless access networks. In this case, we have a Mobile Switching Centre (MSC), sites for Base Stations (BS), and possible locations for Base Station Controllers (BSC). Furthermore, the MSC can be indexed as the first location j=1 with location cost $b_1 = 0$ since the MSC needs to be in the network, regardless. In this sense, this set up is similar to what has been discussed for the wireline case. Topologically, the layout looks the same (with renaming of different entities):

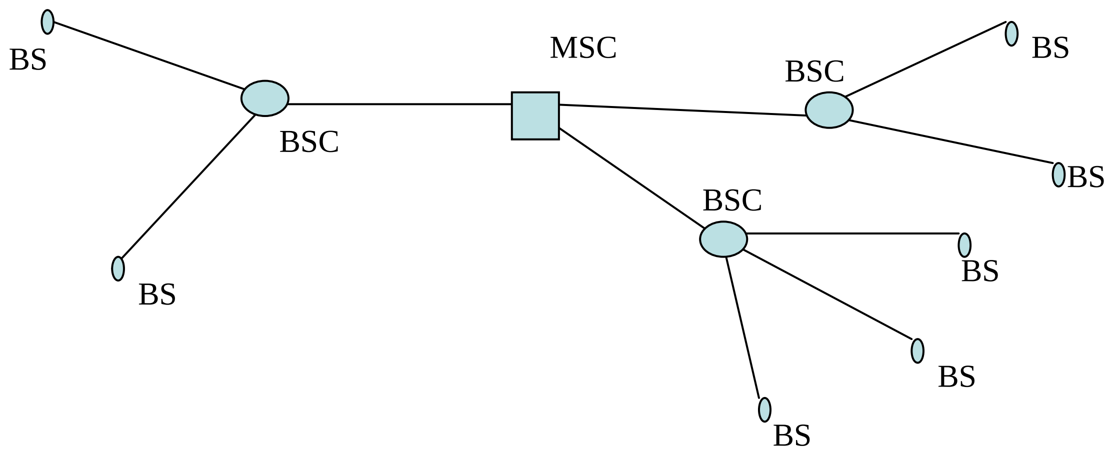


Fig. 7.2.4 Sample wireless network topology

An important issue in the case of wireless networks is the fact that there is limited frequency space. Since the actual frequency allocation can conceivably be different from one country to another, which will impact how many voice channels, this will translate to (and depending on TDMA or CDMA technology), we'll show a simple general case to illustrate the point here. Suppose, for a given frequency space, 100 circuits may be the limit. Then, with the help of the Erlang-B loss formula, we can find that for 1% call blocking GoS, the maximum offered traffic that can be handled is 84.06 Erl. If we assume 0.03 Erlang per customer busy-hour usage, then the maximum number of subscribers that can be handled is 2800.

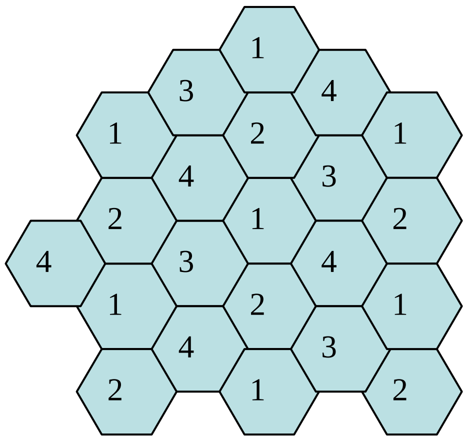


Fig. 7.2.3 Frequency allocation example for wireless systems

On the other hand, the frequency space can be better utilized through spectrum efficiency through the re-use factor process. Typically, re-use factor of 4 (or 3) is suitable for current generation wireless networks. Assume that the re-use factor to be 4; then, the frequency space will lend us in four groups, each of 25 circuits. Now the area which could serve earlier only 100 circuits of capacity (which translated to 2800 subscribers), can be now 22 times for the same geographical area (see figure above, for the frequency re-use layout with N=4 to avoid frequency interference). In each cell area, at 1% blocking GoS and with 0.03 Erl per subscriber, we can accommodate 16.12/.03 = 537 subscribers; this translates to being able to accommodate about 537 x 22 = 11,814 subscribers in the same geographical region due to spectrum efficiency.

It is important to point out that with the re-use factor based layout, each cell (hexagon) will have a base-station – they will be connected through base station controllers to the Mobile Switching Centre (MSC) – this backhaul part (from the base station onward) can be accomplished through the access design approach discussed earlier in this section.
Finally, as the customer base grows, the size of each cell can be made smaller by tuning the power range to create more smaller cell sites, and thereby increase the capacity of the overall system (by again taking advantage of re-use factor).

7.3. Basic optimisation methods

In this section we shall discuss selected optimization methods applicable to the design problems discussed in Sections 7.1 and 7.2 of this chapter. The presented methods are discussed in many operation research books and papers on optimization. In particular, an extended survey of these methods can be found in Chapter 5 of [7.1].

7.3.1. Linear Programming

Many problems defined in Section 7.1 are linear programming problems (linear programmes, LP, in short). This is denoted in the acronyms by ”-LP”. Such problems are commonly known and by far the most frequently used. A general form of a LP problem is as follows:

$$\textbf{minimize} \quad z = \Sigma_i \, c_i x_i \qquad (7.3.1a)$$
$$\textbf{subject to} \quad \Sigma_i \, a_{ji} x_i \le b_j \quad j=1,2,...,J \qquad (7.3.1b)$$
$$\Sigma_i \, x_i = e_k \quad k=1,2,...,K. \qquad (7.3.1c)$$

In the above formulation z and x_i (i=1,2,...,N) are continuous variables (unknowns). There are J non-equality constraints (with the left-hand side coefficients a_{ji}, j=1,2,...,J, i=1,2,...,N and right-hand sides b_j, j=1,2,...,J) and K equality constraints (with the left-hand side coefficients d_{ki}, k=1,2,...,K, i=1,2,...,N and right-hand sides e_k, k=1,2,...,K).

For LP problems a very efficient (in practice) method, simplex algorithm, is well known. This algorithm is implemented in all commercially available LP solvers (there are also freeware LP solvers available on the Web) and can be easily used for solving the LP design problems formulated in Section 7.3. It should be noted that most of these implementations (as CPLEX or XPRESS) are able to deal with large linear programmes with several thousands of variables and constraints, yielding exact optimal solutions in acceptable time.

Still, for large telecommunication core networks with several tens of nodes, the resulting LP problems (especially for the restoration/protection problems) can have too many variables and constraints even for the most advanced LP solvers. In such a case, appropriate LP decomposition methods of column generation and Benders’ decomposition (see [7.6] and any modern handbook on LP) can be used.

To summarize: if the problem at hand can be formulated (or its reasonably approximated) as a linear programme then there is very good chance to solve in an exact way using commercial (or freeware) solvers. Otherwise, in general we may face serious problems.

7.3.2. Branch-and-Bound method for Mixed-Integer problems

Recall that mixed-integer programming problems (MIP), already considered in Section 7.1, are obtained from LP formulation when variables in a certain subset of all variables $\boldsymbol{x}$ = (x_1,x_2,...,x_N) are allowed to assume only integral (binary) values. When we cannot formulate a problem in an LP form and we are forced to use an integral MIP formulation then in most cases of network design this means that we are dealing with a difficult (likely NP-complete) problem. Still, if we have a MIP formulation then we can again use commercial MIP solvers (as CPLEX or XPRESS), although this time their efficiency can be very limited only to small networks. Roughly speaking, MIP solvers are based on the branch-and-bound approach, in many cases extended to the so called branch-and-cut approach (we will not discuss the later extension since it is a rather advanced one, see [7.1] and [7.10]).

The branch-and-bound (B&B) method (see [7.1] and [7.11]) is based on scanning the nodes of the so called B&B tree. These nodes correspond to all possible (partial) combinations of the values assumed by integral variables (note that the number of such nodes can be enormous as it grows exponentially with the number of integral constraint). In each such node, certain integral variables are fixed (if all these variables are fixed then the corresponding node is a leaf of the B&B tree) and the rest are assumed continuous and they define a relaxed LP sub-problem, which is solved by an LP solver. The basic feature of a good B&B algorithm is that it does not visit all the nodes but rather does not enter many branches of the B&B tree, substantially reducing the total number of visited nodes (and the number of relaxed LPs that have to be solved). This feature of B&B is based on the observation that if the optimal solution of a relaxed LP problem corresponding to a B&B tree node has a value larger than or equal to the best solution of the MIP problem achieved so far, then we can skip the branch emanating from the considered node. This is turn is implied by the fact that the optimal solution of each relaxed problem is a lower bound of the original MIP problem.

In fact, the branch-and-cut enhancements of B&B are based on certain ways of improving the lower bounds in question by generating additional constraints for the relaxed problems imposed by the nature of the MIP problem.

We end this subsection with listing a Pascal-like pseudo-code of version of the recursive B&B algorithm presented in Chapter 5 of [7.1]. We assume that all decision variables $\boldsymbol{x}$ are binary: $x_i \in \{0,1\}$, $i = 1,2,...,N$. In the following description we use the following notation:

$N_U \subseteq \{1,2,...,N\}$	set of indices corresponding to unspecified values of binary variables, the binary requirement for these variables is relaxed so that for $i \in N_U$, x_i is a continuous variable from interval [0,1]
$N_0 \subseteq \{1,2,...,N\}$	set of indices corresponding to binary variables equal to 0
$N_1 \subseteq \{1,2,...,N\}$	set of indices corresponding to binary variables equal to 1.

Function *objective*(N_0,N_1) used in the algorithm returns the optimal solution $z^0 = \boldsymbol{C}^0$ of the LP sub-problem being a relaxed, LP version of the original MIP problem defined by (7.3.1) and the following variable constraints:

$$\begin{array}{lll} - \ 0 \le x_i \le 1,\ x_i \text{ continuous} & \text{for } i \in N_U & \\ - \ x_i = 0 & \text{for } i \in N_0 & \qquad (7.3.2) \\ - \ x_i = 1 & \text{for } i \in N_1. & \end{array}$$

If for given N_0 and N_1 such a sub-problem is infeasible (which can frequently happen) then, by definition, *objective*(N_0,N_1) = $+\infty$. To initialize the procedure we put $N_0 = N_1 = \emptyset$, $N_U = \{1,2,...,N\}$, and assign a large number (greater than the expected optimal solution of the problem (7.3.1) and (7.3.2)) to $\boldsymbol{C}^0$. Note that during execution of B&B, it always holds that $N_0 \cap N_1 =$ _ (N_0 and N_1 are disjoint) and that $N_U = \{1,2,...,N\} \setminus (N_0 \cup N_1)$. Any such triple

(N_0,N_1,N_U) is a node of the BB tree; each such node is associated with the LP problem defined by (7.3.1) and (7.3.2).

---

procedure *Branch_and_Bound*(N_U,N_0,N_1)
begin
    if N_U = _ **and** *objective*$(N_0,N_1) < \boldsymbol{C}^0$ **then**
        begin
            $\boldsymbol{C}^0 := objective(N_0,N_1)$; $N_0{}^0 := N_0$; $N_1{}^0 := N_1$
        end
    else { N_U is not empty }
        if *objective*$(N_0,N_1) \geq \boldsymbol{C}^0$ **then**
            return
        else
            begin
                choose $i \in N_U$;
                $BBB(N_U \backslash \{i\}, N_0 \cup \{i\}, N_1)$;
                $BBB(N_U \backslash \{i\}, N_0, N_1 \cup \{i\})$
            end
end { procedure }

In the introduced procedure the lower bounds used for "pruning" the BB tree are computed by solving the node LP sub-problems (function *objective*(N_0,N_1)). Still, for particular problems the lower bounds can be found with other, specific (and more effective) means. Generally, the quality of the lower bounds (the greater the better) and the time required for their computation is decisive for the efficiency of the approach.

7.3.3. Stochastic Meta-heuristics

Stochastic meta-heuristic methods can be used for approximate solution of the network design problems such as MIPs and concave problems. Such meta-heuristics include such methods as simulated annealing (SA), evolutionary (genetic) algorithms (EA), tabu search (TS) and others. For a survey of these methods refer to [7.2] and [7.3]. Stochastic heuristics can be pretty effective in such cases as modular dimensioning problems DP2-MIP (7.1.2) and DP3-MIP (7.1.3), concave dimensioning DP4-CV (7.1.4) or single-path dimensioning DP6–SPR-MIP (7.1.7).

Below, as an example, we will present a Pascal-like pseudo-code of a version of SA specified in Chapter 5 of [1]. The procedure assumes the following discrete combinatorial optimization problem:

$$\textbf{minimize } \boldsymbol{C}(\boldsymbol{x}) \qquad \textbf{subject to } \boldsymbol{x} \in X, \qquad (7.3.3)$$

where X is a finite set of feasible solutions $\boldsymbol{x}$. The procedure uses the notion of "neighbourhood" of point $\boldsymbol{x}$, $N(\boldsymbol{x})$ with the properties $N(\boldsymbol{x}) \subseteq X$ and $\boldsymbol{x} \notin N(\boldsymbol{x})$. The choice of the neighbourhood is problem-dependent and is a very important element in problem solving with SA.

```
procedure Simulated_Annealing
begin
    choose_initial_point(x);
    C^0:= C(x); x^0:= x;
    set_initial_temperature(t);
    while stopping_criterion not true
        begin
            l:= 0;
            while l < L do
                begin
                    y:=  random_ neighbor(N(x));
                    ΔC:= C(y) - C(x);
                    if ΔC ≤ 0 then
                        begin
                            x:= y;
                            if C(x) < C^0 then begin C^0:= C(x); x^0:= x end;
                        end
                    else if random(0,1) < e^(-ΔC/t)  then  x:= y;
                    l:= l + 1
                end
            t:= τ×t
        end
end { procedure }
```

The procedure starts with selecting an initial point $\boldsymbol{x} \in X$ and setting the initial temperature t (initial temperature is a parameter, usually a large number). Then the algorithm proceeds to the main outer **while-end** loop for which the temperature is fixed. Then the inner **while-end** loop is executed L times (L is another parameter of the algorithm, also a large number). Each execution of the inner loop consists in selecting a neighbour $\boldsymbol{y}$ of the current point $\boldsymbol{x}$ ($\boldsymbol{y} \in N(\boldsymbol{x})$) at random and performing a test in order to accept a move from $\boldsymbol{x}$ to $\boldsymbol{y}$ or not. The move is always accepted if it does not increase the objective function $\boldsymbol{C}(\boldsymbol{x})$. Moreover, the (uphill) move is accepted with probability $e^{-\Delta \boldsymbol{C}/t}$ even though it results in an increase of $\boldsymbol{C}(\boldsymbol{x})$. For a fixed t, the acceptance probability is an exponentially decreasing function of $\Delta\boldsymbol{C}$ so the acceptance probability quickly becomes very small with the increase of $\Delta\boldsymbol{C}$. The use of the test makes it possible to leave local minima encounter during the process of wondering around the solution space within the inner loop. Once L steps of the inner loop are performed the temperature is decreased ($t := \tau \times t$, for some fixed parameter τ from interval (0,1), e.g., $\tau = 0.95$) and the inner loop started again. An example of the temperature reduction function is, for some parameter τ from). For fixed $\Delta\boldsymbol{C}$ the acceptance probability decreases, so in the consecutive executions of the inner loop the uphill moves are more and more rare. The stopping criterion can be the lack of significant improvement of the objective function in two consecutive executions of the outer loop.

Now we will show how to apply SA to the concave problem DP4-CV (7.1.5). Since it can be shown that in the optimal solution set of DP4-CV there are always single-path (non-bifurcated) solutions, the solution space X is finite and its points can be coded as vectors $\boldsymbol{x} = (x_1,x_2,...,x_D)$ where for d=1,2,...,D, $1 \le x_d \le m(d)$, and x_d denotes the number of the path on the candidate path list actually used to allocate demand volume h_d. A neighbourhood of $\boldsymbol{x} \in X$ is defined as:

$$N(\boldsymbol{x}) = \{ \boldsymbol{y} \in X: \boldsymbol{x} \text{ and } \boldsymbol{y} \text{ differ exactly at one position } \}. \quad (7.3.3)$$

Having defined the neighbourhood structure, the SA procedure can be run from some starting point $\boldsymbol{x}^0 \in X$ with some fixed parameters t, L and τ.

7.3.4. Other Optimization Methods

Besides general approaches of LP, B&B, and stochastic meta-heuristics, there exist optimization methods specialized for convex problems (such as reduced gradient method and gradient projection method, see [7.6]) and for concave problems (such as iterative algorithms of [7.12] and [7.9]). The methods for convex optimization problems are discussed in any general book on optimization as the convex problems play a central role in the optimization theory. Algorithms for concave problems are less known (and much less efficient). For a survey of both types of methods see Chapter 5 of [7.1].

7.3.5. Shortest Path Algorithms

For the link-path problem formulations used throughout Section 7.1 we need to generate the lists of candidate paths. This can pose some problems but in general a reasonable way is to generate, for each demand d, $m(d)$ shortest paths in terms of weights being the link marginal costs (c_e, e=1,2,...,E) for dimensioning problems of the DP type, or in terms of unit weights (c_e = 1, e=1,2,...,E) for allocation problems of the AP type. This can be achieved by using a K-shortest path algorithm which generates a list of consecutive K shortest paths $P_{d1},P_{d2},...,P_{dK}$ with the property that P_{d1} is the shortest path demand d, P_{d2} is the second shortest path, and so on. We point out that such an algorithm is not simple at all (refer to [7.13] and Appendix C in [7.1.]).

In the K-shortest path algorithm we do not assume that the generated paths are disjoint. Disjoint paths are of interest for the problems involving path-diversity (see DP5-PD-LP (7.1.6) and RDP-GPD-LP (7.1.13)). An algorithm for finding the so called shortest sets of K link- or node-disjoint paths can be found in [7.14]. They are based on iterative use of any of classical shortest-path algorithms, as the famous Dijkstra algorithm (with which the reader should be familiar).

Still another type of problem is to generate a shortest path (or a set of K-shortest paths) with a limited number of hops (i.e., intermediate nodes). This can be done by a relatively simple modifications of the appropriate algorithms described above.

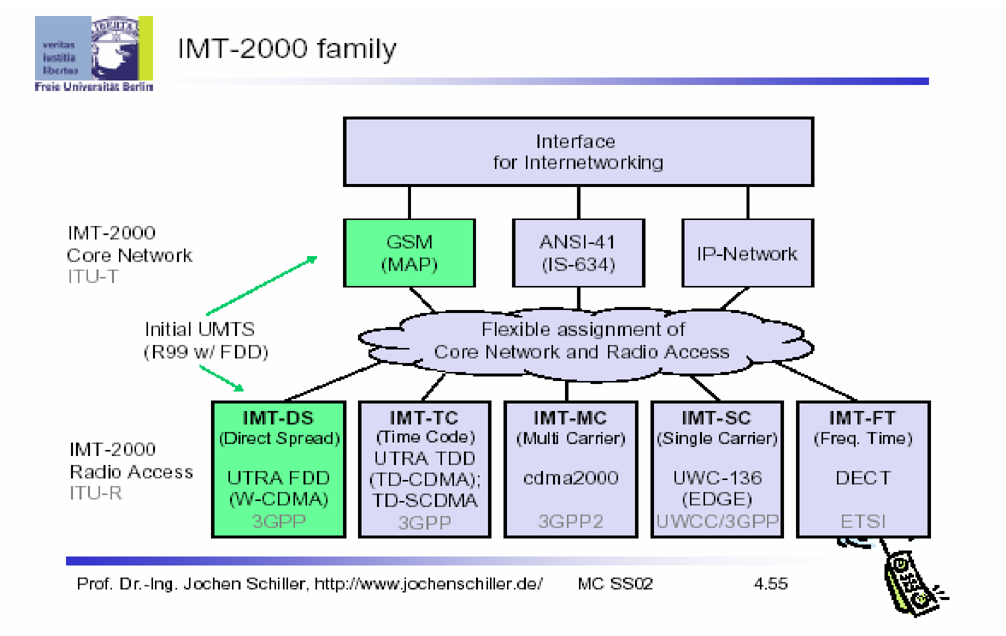
For a survey of the shortest paths algorithms relevant for network design refer to Appendix C in [7.1].

7.4 Specific Issues of Radio Network Planning

7.4.1. Introduction to radio network planning

7.4.1.1 Introduction to IMT2000

IMT-2000 [1], formerly called future public land mobile telecommunication system (FPLMTS), aimed to establish a common worldwide standard communication system allowing for terminal and user mobility, supporting the idea of universal personal telecommunication (UPT). Within this context, ITU has created several recommendations for FPLMTS systems, e.g., network architectures for FPLMTS (M.817), Requirements for the Radio Interfaces for FPLMTS (M.1034). The number 2000 in IMT2000 should indicate the start of the system (year 2000+x) and the spectrum used (around 2000 MHz). IMT-2000 includes different environments such as indoor use, vehicles, satellites and pedestrians. The World Radio Conference (WRC) 1992 identified 1885-2025 and 2110-2200 MHz as the frequency bands available worldwide for the new IMT-2000 systems. Within these bands, two times 30MHz have been reserved for the mobile satellites services (MSS). ITU originally planned to have a common global system, but after many political discussion and fights about the patents the idea of so-called family of 3G standards was adopted. What are the IMT-2000 family members? The ITU standardized five groups of 3G radio access technologies. The figure below gives an overview.

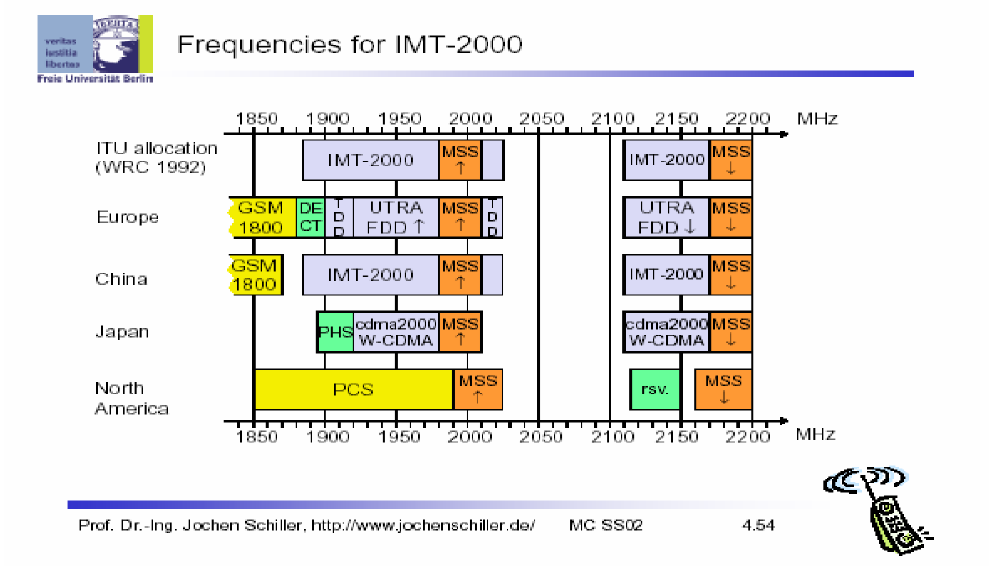


- IMT-DS: The direct spread technology comprises wideband CDMA (W-CDMA) systems. This is the technology specified for UTRA-FDD and used by all European providers and Japanese NTT DoCoMo for the 3G wide area services. To avoid complete confusion, ITU's name for this technology is IMT-DS, ETSI calls it UTRA-FDD in the UMTS context, and the technology used is W-CDMA (in Japan this is promoted as FOMA, freedom of mobile multimedia access). Today, standardization of this technology takes place in 3GPP (Third generation partnership project, 3GPP, 3GPP2).

- IMT-TC: Initially, this family member, called time code, is contained only in the UTRA-TDD system which use the time-division CDMA (TD-CDMA). Later the Chinese proposal TD-synchronous CDMA was added. Now both standards have been combined and 3GPP fosters the development of these technologies. Up to now, it is still unknown what future perspectives this technology will introduce. The initial UMTS installations are based on W-CDMA.
- IMT-MC: The American 3G standard cdma2000 is a multi-carrier technology standardized by 3GPP2 (Third generation partnership project 2, 3GPP2, 2002), which was formed shortly after 3GPP to represent the second main stream in 3G technology. Version cdma2000 EV-DO has been accepted as the 3G standard.
- IMT-SC: The enhancement of US TDMA systems, UWC-136 is a single carrier technology originally promoted by the Universal Wireless Communications Consortium (UWCC). It is now integrated into the 3GPP efforts. This technology enhance the 2G IS 136 standard.
- IMT-FT: As the frequency/time technology, an enhanced version of the cordless telephone standard DECT has also been selected for the applications that do not require high mobility. ETSI is responsible for the standardization Of DECT.

The main driving forces in the standardization process are 3GPP and 3GPP2. ETSI has moved its standardization process to 3GPP and plays a major there. 3GPP tends to be dominated by European and Japanese manufacturers and standardization bodies, while 3GPP2 is dominated by the company Qualcomm and CDMA network operators. The figure above shows more than just the radio access technologies. One idea of the IMT-2000 is the flexible assignment of a core network to a radio access system. The classical core network uses SS7 for signaling which is enhanced by ANSI-41 (cdmaOne, cdma2000, TDMA) or MAP (GSM) to enable roaming between different operators.

The figure below shows the ITU frequency allocation (from the world administrative radio conference. 1992) together with examples from several regions that already indicate the problems of worldwide common frequency bands.



7.4.1.2 A brief look at cellular history

The history of mobile communications started with the experiments of the first pioneers in the area. In late 1800 century, the studies of Hertz inspired Marconi to search market for the new commodity. The needs for communication in the first and second world wars were also aiding the start of cellular radio, especially in terms of utilisation of ever higher frequency. Bell Laboratory first introduced the cellular concept as known today and demonstrated how the cellular system could be designed in 1971 [7.4.2].

The first operational cellular system in the world was in Tokyo, Japan, in 1979 and the network was operated by NTT. The system utilised 600 duplex channels in the 800 MHz band, with channel separation of 25 kHz.
Two years later, the cellular era reached Europe. The Nordic Mobile Telephone at 450 MHz band (NMT-450 system) started operation in Scandinavia. Total Access Communication System (TACS) was launched in United Kingdom in 1982 with an extended version deployed in1982. Subsequently the C-450 cellular system was introduced in Germany in 1985. Therefore, at the end of 1980s there were several different cellular systems in Europe, which are known as first generation **(1G)** cellular systems. One of the disadvantages of 1G is lack of interoperation in terms of different countries having different cellular standard. Thus, in the early 1990s, with the development of integrated circuit technology, the second generation cellular systems (**2G**) began to be deployed throughout the world. GSM intended to provide a single unified standard in Europe which enables seamless speech service throughout Europe in terms of international roaming. In United States the analogue first-generation system called Advanced Mobile Phone System (AMPS) was launched in 1983. In 1991 and 1996 respectively, the IS-54 and IS-136 was introduced as the first digital cellular system in US. The other standard IS-95, also known as CDMA-One was introduced in 1993. Both of these standards operate at the same band as AMPS.

Over the decade the world of telecommunications has been changed drastically for various technical and political reasons. The wide spread use of digital technology in telecommunications bas brought about radical changes in services and networks. Furthermore, as time has been passing by, the world has become smaller: roaming in Japan, roaming in Europe or Roaming in the United States is not anymore enough. Globalisation is having its impact also in the cellular world. In addition to this, current strong drive towards wireless internet access through mobile terminals generates needs for universal standard, Universal Mobile Telecommunication Standard, **3G**. The third generation networks being developed by integrating the features of telecommunications and internet protocol (IP) based networks. Networks based on IP, initially designed to support data communications, have begun to carry streaming signals such as voice/sound traffic, although with limited voice quality and delay tolerance. Commentaries and predictions regarding wireless broadband communications and wireless internet services are cultivating visions of unlimited services and applications that will be available to customers ‘ anywhere and anytime’. Consumers expect to surf the web, check the email, download files, have real time videoconferencing calls and perform various other tasks through wireless communication link. The consumers expects a uniform user interface that will provide access to wireless link whether shopping at the mall, waiting at the airport, walking around the town, working in the office or driving on the highway!

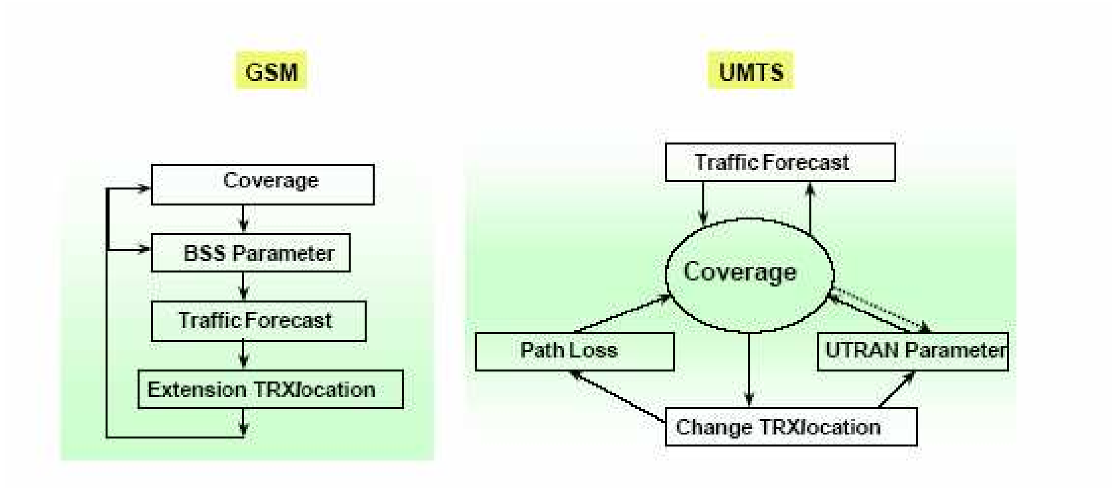
7.4.1.3 Evolution of radio network planning—From 1G to 3G

The radio network planning and its development have always been mapped to the development of the access technologies and requirements set by those [7.4.3]. Main characteristics such as: analogue transmission, high power transmitter, voice only, national wide usage, enables the first generation analogue mobile networks to be planned based on low capacity requirement. The radio network planning was based purely on coverage. Sites were high enough to keep the site density low and omni-directional antennas were used. The Okumura-Hata propagation model was and still is widely used for coverage calculation in the macro-cellular network planning. The model was developed by Y.Okumura in Tokyo based on the measurement at frequencies up to 1920 MHz combing with measurement result fit into the mathematical model by M.Hata. In the original model the path loss was computed by calculating the empirical attenuation correction factor for urban areas as a function of the distance between the base station and mobile station and frequency. Later on the factors were added to the free space loss. Further correction factors were provided for street orientation, suburban, open areas, and irregular terrain. The range of usability with different land use and terrain types and for different network parameters has made the Okumura-Hata propagation very useful in many different propagation studies.

With the evolution of the second generation mobile system, site density was getting higher due to the increasing capacity requirements. The key characteristics of 2G mobile systems are digital transmission, dense site placement, low power transmitter, semi-compatible, and voice focus with data service support. All this forced the cellular network to replace the omni directional site structure and lead to the introduction of cell splitting, for example one site consisting of three sectors instead of just one. Owing to the increased spectral efficiency requirements, the interference control mechanism became more important. Such a mechanism like antenna tilting was introduced to reduce co-channel interference. Furthermore, the macro-cellular propagation model was no more accurate enough. New models were needed to support microcellular planning. The Walfisch-Ikegami is such a new model based on assumption that the transmitted wave propagates over the rooftop by a process of multiple diffraction. The buildings in the line between the transmitter and receiver are characterised as diffracting half screens with equal height and range separation.

The general planning process for GSM network can be divided into three main subsequent steps: (1) coverage, (2) parameter, and (3) capacity planning. Coverage planning consists of the selection of the location and configuration of the antennas. The coverage area achieved by a single antenna depends mainly on the propagation conditions as it is independent from other antennas in the network. During the following parameter planning process all radio parameters (frequency, hand-over configuration and power control parameters) are defined. Once a cell is in operation traffic measurements are made to yield to the prediction of required number of channels. The increased traffic does not affect the coverage area or the parameter settings – at least to a reasonable good approximation. In this case, an additional TRX has to be installed and new parameter settings for this TRX have to be provided. Only when an additional site may be required for capacity reasons the increase of traffic has an influence on the coverage area. For GSM well-developed algorithms both for the synthesis and analysis of the networks exist and a lot of appropriate planning tools are commercially available now.

In contrast, the situation for planning 3G mobile network such as UMTS is much more complicated. The cell range in a WCDMA system does not only depend on propagation conditions but also on the traffic load of the cell. Furthermore, the amount of interference received from other cells depends on their traffic load as well. Additionally, the traffic load of a cell is influenced by the soft hand-over areas, which are mainly defined during the parameter planning step. Coverage, parameter, and capacity are thus a highly coupled process requiring integrated planning of these three steps. The fundamental difference between the planning process in GSM and UMTS is displayed below.



1.4 New Challenges in 3G/4G network planning

3G mobile network will be used to provide broadband communications and pervasive computing infrastructure and to form wireless mobile internet (WMIT). 3G UMTS technology supports a large range of services with different bit rates and quality requirements, asymmetric links, mixed traffic scenarios, coverage and capacity dependency, making the design of the network a difficult and challenging task [7.4.2].

Before looking into more detail what actually will be new (and different) in WCDMA radio network planning and optimization, it is useful to summarise some of the defining characteristics of 3G multi-service radio network. One can characteristic 3G radio access with the following attributes [7.4.2]:

- Highly advanced radio interface, aiming at great flexibility in carrying and multiplexing a large set of voice and in particular data services. Furthermore the throughput ranging from low to very high data rates, ultimately up to 2Mbit/s.

- Cell coverage and service design for multiple services with largely different QoS requirements. Due to the large difference in the resulting radio link budgets, uniform coverage and capacity designs as practised in today’s voice-only radio network, can no longer be obtained. Traffic requirement and QoS targets will have to be distinguished among the different services.

- A large set of sophisticated features and well-designed radio link layer. Example of this are: various radio link coding/ throughput adaptation schemes; support for advanced performance enhancing antenna concepts, such as BS transmission diversity, or the enabling of interference cancellation schemes.

- Efficient mechanism for interference averaging and robustness to operate in a strongly interference limited environment. High spectral efficiency operation will require good dominance of cells by proper choices for site locations, antenna beamwidths, tilts, orientation, ect.

- Extensive use of 'best effort' provision of packet data capacity, i.e. temporarily unused radio resource capacity shall be made available to the packet data connections in a flexible and fair manner.

- In order to be able to provide ultimately high radio capacity, 3G networks must offer efficient means for multi-layered network operation. Furthermore, seamless interoperation of 2G and 3G is required.

- Another very important aspect is the possibility of co-existence of 3G cells and 2G cells, reducing cost and overhead during site acquisition and maintenance. Consequently, 2G-3G co-site gives new challenges for radio network planners.

With those new characteristics in 3G mobile network, general challenges to face in 3G network planning are based on the fact that a lot of issues are interconnected and should be considered simultaneously.

- Planning methods should not only meet the current network demands, but also be adaptive to new and future requirement for the next generation mobile network. Furthermore, the network management mechanism must support the operators with the real time network performance and indicates not only the coverage or capacity limited area, but also identify the areas where new services could be introduced within the existing infrastructure GSM, EGPRS, IMT2000.

- Analysing the new traffic demands and distribution will be more and more important because of the uncertain of the traffic growth. This is not only a questions about the total amount of traffic growth, but also how to determine the distribution and characteristic of the new traffic mixed with previous traffic

- All CDMA systems (CDMA2000, WCDMA, TD-SCDMA) have an interconnection between capacity and coverage, and thus quality. Therefore the network planning will be based on both propagation estimations and interference situation in the network.

- There are also some practical constraints in 3G network planning. Due to economic issues or technical reasons, operators who already have a network tend to use the site-co-location. For the greenfield operator, there are more practical limitations set by site acquisition process.

7.4.2 General Process of 3G radio network planning

7.4.2.1. Introduction

A cellular mobile communications network consists of radio access network (RAN) and core network (CN). Network planning is the most important issue when building a new network or expanding an existing network. The assistance of good and capable planning tools plays a very important role in cellular network planning. The planning of a cellular network can be divided into three stages [7.4.6]:

- Cell planning:t The number, locations and parameters of the base stations are determined. It is the most important, the most challenging and the most tedious part of cellular network planning. The cell planning process can be divided in the order: cell dimensioning, detailed capacity and coverage planning and cell optimization. Nowadays, cell planning issue are most popular theme to which most research on radio network planning has been devoted [7.4.6].

- Radio network controller (RNC) planning: We need to decide the number and location of RNCs. We need to decide the link topology and capacity in each RNC area. One example is how to design primary and secondary routes that connect each BS to its RNC. RNC planning is a promising and potential research area in radio network planning [7.4.6].

- Core network (CN) planning: We need to decide CN transmission topology and capacity; we also need to dimension a number of CN interfaces. Most problems in CN planning can be found in previous chapter on traditional core network planning and thus we will not look into this [7.4.6].

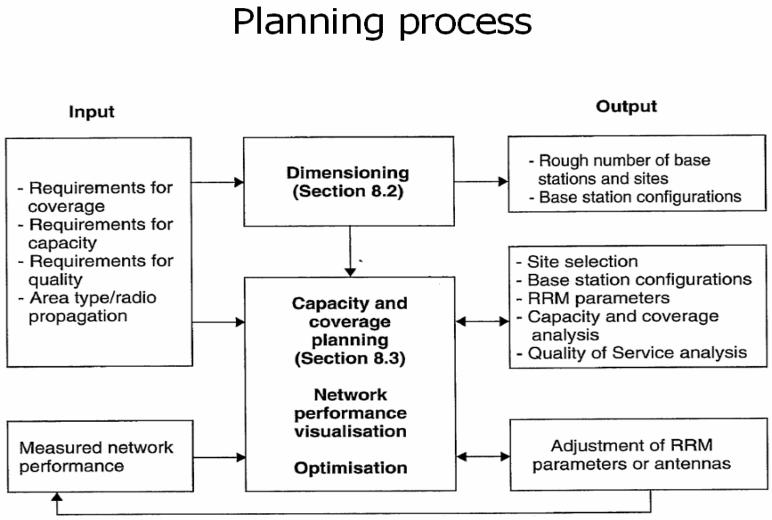


Figure 8.1. WCDMA radio network planning process

7.4.2.2 Cell dimensioning

3G radio network dimensioning is a process through which possible configurations and amount of network equipment are estimated, based on the operators' requirements related to the following [7.4.7]:

Coverage:

- coverage regions
- area type information
- propagation conditions

Capacity:

- spectrum available
- subscriber growth forecast
- traffic density information

Quality of Service:

- area location probability (coverage probability)
- blocking probability
- end user throughput

Dimensioning process includes the following steps:

1. Radio link budget (RLB) and coverage analysis
2. Capacity estimation
3. Estimations on the amount of sites and base station hardware, radio network controllers and core network elements.

Systematic dimensioning provides the first and rapid evaluation of the possible network configuration. This includes both the radio access network and the core network. Here we focus on the access network part dimensioning since core network part has been discussed in previous chapters. The radio link budget calculation is done for each service, and is the tightest requirement that determines the maximum allowed path loss. There are a big difference between uplink radio link budget and down link radio link budget in 3G UMTS mobile network because of the asymmetric traffic demands and WCDMA air interface property [7.4.7].

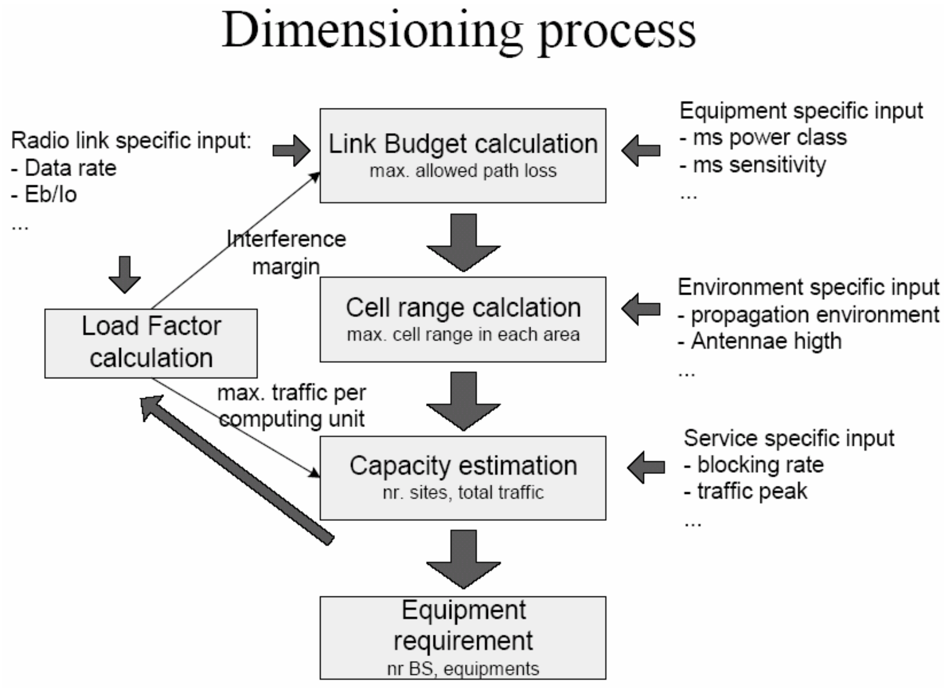
- Uplink Radio Link Budget
- Downlink Radio Link Budget
- Load factors
- Soft capacity

7.4.2. WCDMA capacity

In this section, we mainly focus on studying the parameters used in the calculation of WCDMA capacity. Most of work is within the dimensioning process of WCDMA radio network planning [7.4.7].

7.4.2.1 Radio Link Budget

The purpose of calculating radio link budget is to find the allowed propagation loss coverage area. In WCDMA radio link budget, some new parameters have been defined. Interference margin is the counter for the maximum allowed load of the system. The higher the load is in the system, the more the interference margin will be, and the smaller the coverage will be in uplink. For coverage limited case, we should have a small interference margin, but for capacity limited case, a big interference margin is recommended. Fast fading margin is needed for low speed terminals to overcome fast fading by adding extra power, maintaining adequate closed loop fast power control. Below we show a particular table of link budget reference table. There are several steps for calculating it. First, we get the Max path loss between mobile and NodeBs according to the system parameters like interference margin, antenna gain, receiver sensitivity, transmission power, required E\N. We should notice that the max path loss includes indoor loss, allowed propagation loss, and normal fading loss. Or we can see the max path loss is predetermined at the beginning by the physical properties of the communication system, and (the indoor loss) + (allowed propagation loss)+(normal fading loss), which is caused by environment, cannot be more than the max path loss. Thus we get the allowed propagation loss, which is path loss due purely o to propagation, by a simple calculation: allowed propagation loss = (max path loss) – (the indoor loss) –- (normal fading loss). After that, based on a particular propagation model, we can get the distance which is used for calculating the cell range.



Reference link budget for 12.2 kbps in WCDMA system

12.2 kbps voice service (120 km/h, in car)		
Transmitter (mobile)		
Max. mobile transmission power [W]	0.125	
As above in [dBm]	21	a
Mobile antenna gain [dBi]	0	b
Cable/Body loss [dB]	3	c
Equivalent Isotropic Radiated Power	18	d=a+b-c
Receiver BS		
Thermal noise density [dBm/Hz]	-174	e
Base station recever noise figure [dB	5	f
Receiver noise density [dBm/Hz]	-169	g=e+f
Receiver noise power [dBm]	-103.2	h=g+10*log10(3840000)
Interference margin [dB]	3	i
Receiver interference power [dBm]	-103.2	j=10*log10(10^((h+1)/10)-10^(h/1(
Total effective noise + interference [	-100.2	k=10*log10(10^(h/10)+10^(j/10))
Processing gain [dB]	25	l=10*log10(3840/12.2)
Required Eb/No [dB]	5	m
Receiver sensitivity [dBm]	-120.2	n=m-l+k
Base station antenna gain [dBi]	18	o
Cable loss in the base station [dB]	2	p
Fast fading margin [dB]	0	q
Max. path loss [dB]	154.2	r=d-n+o-p-q
Coverage probability [%]	95	
Log normal fading constant [dB]	7	
Propagation model exponent	3.52	
Log normal fading margin [dB]	7.3	s
Soft handover gain [dB], multi-cell	3	t
In-car loss [dB]	8	u
Allowed propagation loss for cell ran	141.9	v=r-s+t-u

7.4.2.2 Uplink Load factor and Uplink Capacity

CDMA defines two kinds of channels—"forward" and "reverse" channels. When signal transmits from Mobile Station to Base Station, this channel is called Reverse channel, which is uplink.When the Base station transmit signal to Mobile Station, it is called forward channel, which is downlink. The coverage is often limited by the uplink, while the capacity is downlink limited (Radio network planning). Furthermore, the load factor or load equation (both uplink and downlink) is commonly used to make a semi-analytical prediction of the average capacity of a WCDMA cell, without going into system level capacity simulation.

In the following, the capacity of a WCDMA cell with multiple services classes is calculated. Since the downlink capacity can be counted in the same way with different parameters, we can omit it. In the following considering a single UMTS cell for capacity analysis, the influence from neighbour cells is modelled as by the noise, which is expressed by neighbour cell interference ratio introduced below. The admission control is performed on the basis of measured noise rise. Noise rise is the ratio of the interference to the interference of an unloaded system, which corresponds to the thermal noise. The total interference density consists of own-cell interference, the other-cell interference, and also the thermal noise. The admission control estimates the increase of the noise rise that would be caused by accepting a new connection and blocks it if the result exceeds the pre-defined threshold value. The noise rise is a value that is actually measured by a Base Station.

The relevant WCDMA cell capacity defining parameters are the following:

– WCDMA chip rate, W=3.84Mcps. The chip sequence
has a much faster data than the information signal and thus spread the signal bandwidth beyond its original bandwidth. The term chip is used so as to differentiate between coded bits and the longer encoded bits of the information signal. The digital information signal is directly multiplied by a code sequence with a very high chip rate, which is called the spreading process.

– Noise-rise is defined as the ratios of the total received wide-band power to the noise power; It is functionally related to the cell uplink utilization factor $\eta_{uL.}$ The Value of 3dB noise-rise corresponds to $\eta_{uL.=}$ 50% utilization.
---Neighbor cell interference ratio *i* (explained below) :

$$i = \frac{other\_cell\_interference\_}{own\_cell\_interference\_} \quad (2.0)$$

Interference ratio depends on cell environment, types of antenna used, and other factors. Common assumed values are 0.55 or 0.65.
We first define the E_b/N_0, the energy per user bit divided by the noise spectral density, and the processing gain of user j is $W/v_J R_J$.

$$(E_b/N_0)_J = \text{Processing gain of user j} \bullet \frac{Signal\_of\_user\_j}{Total\_received\_power(excl..own\_signal)}$$

This can be written as: $(E_b/N_0)_J = \frac{W}{v_j R_j} \bullet \frac{P_j}{I_{total} - P_j}$

Where W is the chip rate (explained below), P_J is the received signal power from user *j*. v_J is the activity factor of user *j*. I_{total} is the total received wideband power including thermal noise power in the base station. By solving P_J, we can derive

$$P_J = \frac{1}{1 + \frac{W}{(E_b/N_0)_j \bullet R_j \bullet v_j}} \bullet I_{total}$$

We define $P_J = L_J \bullet I_{total}$ and the load factor L_J of one connection

$$L_J = \frac{1}{1 + \frac{W}{(E_b/N_0)_j \bullet R_j \bullet v_j}}$$

The total received interference consisting of own-cell interference, other cell interference and the thermal noise can be written as the sum of the received powers from all *N* users in the same call.

$$I_{total} - p_N = \sum_{j=1}^{N} P_j = \sum_{j=1}^{N} L_j \bullet I_{total}$$

The noise rise is defined as the ratio of the total received wideband power to the noise power

Noise rise= I_{total} / P_N (2.6)
Then by using the Equation (2.5), we can obtain:

Nose rise = $I_{total} / P_N = \frac{1}{1 - \sum_{j=1}^{N} L_j} = \frac{1}{1 - \eta_{UL}}$

The Noise rise approaches to infinity and the system reached its pole when η_{UL} becomes close to 1.

In addition, if we consider the interference from other cells, it can be calculated as theRatio of other cell to own cell interference i defined in equation (2.0). What this equation actually does is comparing the interference intensity of other cells to own cell by giving a factor *i* which indicates other cells interference is only 1/*i* times as strong as the original cell interference. And then it converts the other cells interference into the following equation:

$$\eta_{UL} = (1+i) \bullet \sum_{j=1}^{N} L_j = (i+1) \bullet \sum_{j=1}^{N} \frac{1}{1 + \frac{W}{(E_b/N_0)_j \bullet R_j \bullet v_j}} \quad (2.8)$$

N is the total number of active users in the cell for all service. We can also define the service-related entities product by k_J:
$k_J = (E_b/N_0)_J R_J v_J$ (2.9)

Table 2.1 Parameters used in uplink capacity calculation

	Definitions	Recommended values
N	Number of users per cell	
v_j	Activity factor of user j at physical layer	0.67 for speech, assumed 50% voice activity and DPCCH overhead during DTX 1.0 for data
E_b/N_0	Signal energy per bit divided by noise spectral density that is required to meet a predefined QoS (e.g. bit error rate). Noise includes both thermal noise and interference	Depend on service, bit rate, multipath fading channel, receive antenna diversity, mobile speed, ect
W	WCDMA chip rate	3.84 Mcps
R_j	Bit rate per user j	Depend on service
i	Other cell to own cell interference ratio seen by the base station receiver	Macro cell with omni directionalantenna:55%. Macro cell with 3 sectors: 65%

The maximum number of channels of service j in one cell is:

$N_{\max,j} = \eta_{UL}\frac{1}{1+i}(1+\frac{W}{k_j})$ (2.9), which can be derived from (2.8)

In fact, the maximum number of channels of service j is calculated assuming that the cell only provide service j. The maximum number is set to guarantee the Quality of Service since η_{UL} is closely related to Noise rise. When all N users in the cell has a low bit rate of R (e.g. a classical all-voice-service network), we can get that

$$\frac{W}{E_b/N_0 \bullet R \bullet v} \gg 1, \Rightarrow \eta_{UL} = \frac{E_b/N_0}{W/R} \bullet N \bullet v \bullet (1+i)$$

In this case, we can simplify equation (2.9) to be

$$N_{\max,j} = \eta_{UL}\frac{1}{1+i} \bullet (\frac{W}{k_j})$$

Finally, we define a parameter $R_{s,j}$ the service j channel spread bit rate, the proportion of W utilized by one service R_j channel:

$$R_{s,j} = \frac{W}{N_{\max,j}} = \frac{1+i}{\eta_{UL}}(k_j - \frac{k_j^2}{k_j + W})$$

The $R_{s,j}$ is the bandwidth of each channel of the total $N_{\max,j}$ channels. If we specify the total capacity and bandwidth of one cell by a common bandwidth unit, say 20.5 kcps, the $R_{s,j}$ can be counted or quantized as numbers of unit channels. If we choose a smaller unit bandwidth, more accuracy will be achieved. The equations above give an approach to obtain unified measure unit for calculating traffic for in multi-service and multi-QoS WCDMA system. An example using the formulas above is provided below.

Table 2.2 Capacity of a WCDMA cell. Assuming i=0.55, $\eta_{UL} = 0.5$, W=3840 kcps

j	Service	R_J [kbps]	V_J	$(E_b/S_0)_{j\ (dB)}$	$(E_b/S_0)_j$	k_j	$N_{max,j}$	$R_{s,j}$ [kcps]
1	Voice	7.95	0.67	4	2.51	13.38	92.9	41
2	Voice/moving	7.95	0.67	7	5.01	26.70	46.7	82
3	Data	32	1	3	2.00	63.85	19.7	195
4	Data	64	1	2	1.58	101.43	12.5	306

The downlink load factor can be calculated in the same way. What we should notice is that in downlink MS suffers multiple access interference from both other users within one cell but also from other cells, which is the case in the uplink. The reason is that poor synchronization makes the orthogonality.

Finally, we come to compare the coverage and capacity relation in both downlink and uplink. Some interesting results come out. In uplink, the coverage remains too small although either we increase or decrease the load. Therefore, we should study how to enhance the uplink coverage in the future. On the contrary, the downlink is capacity limited case. Whatever we reduce the coverage, the capacity always has a limit value
But in general, the capacity and coverage always has a trade-off relation in both downlink and uplink. Increasing capacity will inevitable result in loss of coverage. We may solve the downlink capacity limited problem by power splitting between frequencies or power splitting between sectors.

7.4.2.3 Soft capacity for both WCDMA and GSM

The soft capacity can be explained as follows. We have a cell in the middle of neighbouring cells. The less interference is coming from the neighbouring cells, the more channels are available in the middle cell. When the service rates in the neighbouring cells are quite high, we may have a low spreading factor and thus small number of scrambling codes for each user. Therefore, number of the users in the cells is small. With a low number of channels per cell, the average load must be low to guarantee low blocking probability. Since the average loading is low, there is typically extra capacity available in the neighbouring cells, therefore interference gives soft capacity.

Also in GSM system, we have the outage-limited cell if the cell is micro-cell and with a low reuse factor. In this case, cells of the same frequency share the interferences, instead of the neighbouring cells, and one cell may borrow the capacity from other cells within the same frequency if load is low. Below we show the steps for calculating soft capacity

- Soft capacity: the increase of Erlang capacity with soft blocking compared to that with hard blocking with the same maximum number of channels per cell.

$$SoftCapacity = \frac{Erlang\ capacity\ with\ soft\ blocking}{Erlang\ capacity\ with\ hard\ blocking} - 1$$

Algorithm for estimation:

- Calculate the number of channels per cell, N, in the equally loaded case, based on the uplink load factor.
- Multiply total number of channels by $1+i$ to obtain the total pool in the soft blocking case.
- Calculate the maximum offered traffic from the Erlang B formula.
- Divide the Erlang capacity by $1+i$.

7.4.2.4 Detailed cell planning and Optimization

7.4.2.4.1 Iterative Capacity and Coverage Prediction

In the detail planning phase real propagation data from the planned area is needed, together with the estimated user density and user traffic. Also information about the existing base station sites is needed in order to utilize the existing site investments. The output of the detailed capacity and coverage planning are the base station locations, configurations and parameters. Network optimization includes: performance measurement, analysis of the measurement results, updates in the network configurations and parameters. A comprehensive treatment can be found in Chapter 3 in [7.4.6]. In this section, we only give general overview of cell planning; recent research on cell planning (e.g. automatic cell planning) will be discussed in detail in next section

Since in W-CDMA all users are sharing the same interference resource in the air interface, they cannot be analysed independently. Each user is influencing the others and causing their transmission to change. These changes themselves again cause changes, and so on. Therefore, the whole prediction process has to be done iteratively until the transmission power is stabilised. This iterative process is illustrated in Figure 8.16. Also, the mobile speeds, multi-path channel profiles, and bit rates and types of services used play a more important role in TDMA/FDMA systems. Furthermore, in W-CDMA fast power control in both uplink and downlink, soft/softer handover and orthogonal downlink channels are included, which also has impact on the system performance. In current GSM coverage planning processes the base station sensitivity is typically assumed to be constant and the coverage threshold is the same for each base station. On contrary, in the case of W-CDMA the base station sensitivity is cell and service specific since it depends on the number of users and used bit rates in all cells.

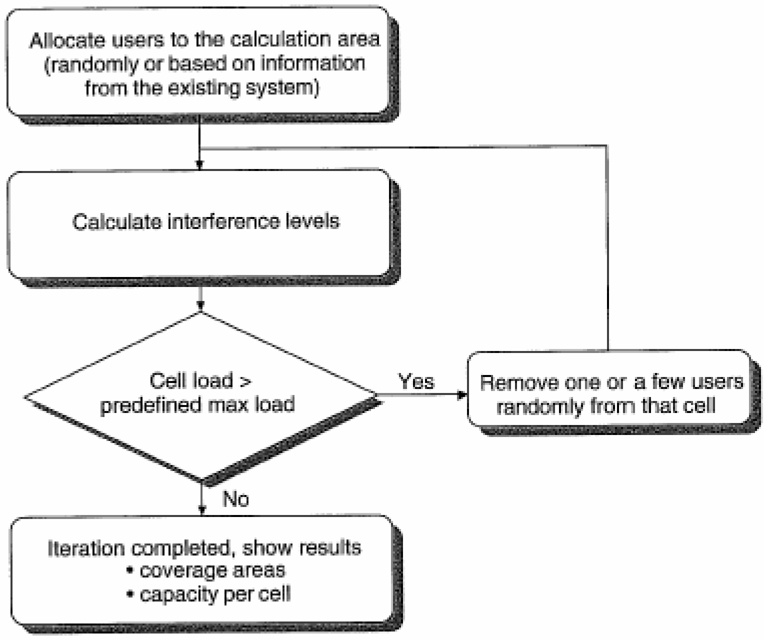


Figure 8.16. Iteration capacity and coverage calculations

7.4.2.4.2 Planning Tool for radio network optimization

In third generation systems, a more detailed interference planning and capacity analysis than simply coverage optimization is needed. The tool should aid the planner to optimize the base station configurations, the antenna directions and even the site locations, in order to meet the quality of service and the capacity and service requirement at the minimum cost. To achieve the optimum result the tool must have knowledge of the radio resource algorithms in order to perform operations and make decisions, like the real network. Uplink and downlink coverage probabilities are determined for a specific service by testing the service availability in each location of planning area. A detailed description of one planning tool can be found in [7.4.7]. New planning methods, so-called automatic radio planning, have an important impact on radio network planning. In automatic radio network planning, the optimization process has to rely completely on the result of the propagation prediction. The accuracy of these propagation prediction methods has a crucial impact on the overall quality of the planning and optimization results. As an example, one of the aims in the IST project ‘MOMENTUM’ is to develop an adaptive propagation model for automatic cell planning [7.4.7].

The planning tool described here differs from the dynamic simulator introduced in radio resource management. The planning tool is a static simulator based on average conditions, and snapshots of the network can be taken. The dynamic simulator includes the traffic and mobility models which make it possible to develop and test the real-time radio resource management (RRM) algorithms. The dynamic simulations can be used to study the performance of the RRM algorithms in realistic environments and the results of those simulations can be used as an input to this network planning tool. For example, the practical performance of the handover algorithms with measurement errors and delays can be tested in the dynamic tool and the results fed into the network planning tool. Testing of RRM algorithms requires accurate modelling of WCDMA link performance, and therefore a time resolution corresponding to power control frequency of 1.5 kHz is used in the dynamic simulator. Such a high accuracy makes the dynamic simulation tool complex and simulations are still too slow – using current top-line high speed workstations – for practical network planning purpose. The accurate dynamic simulation tool can be used to verify and develop the more simple performance modelling in the network planning tool. When enough results from large-scale WCDMA networks are available, those results will be used in calibratiing the network planning tool [7.4.7].

7.4.2.4.3 Radio Network Optimization

Network optimization is a process to improve the overall network quality as experienced by the mobile subscribers and to ensure that network resources are used efficiently. Optimization includes: (1) Performance measurements. (2) Analysis of the measurement results. (3) Updates in the network configuration and parameters [7.4.7].

The optimization process is shown below in Figure below.

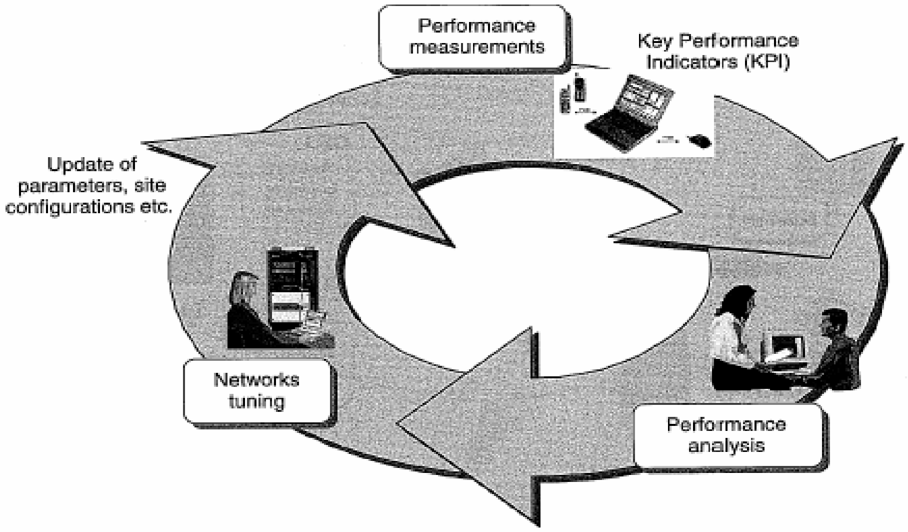


Figure 8.19. Network optimisation process

A clear picture of the current network performance is needed for the performance optimization. Typical measurement tools are shown in Figure 8.20. The measurements can be obtained from the test mobile and from the radio network elements. The WCDMA mobile can provide relevant measurement data, e.g. uplink transmission power, soft hand-over rate and probabilities, CPICH E_c/N_0 and downlink BLER. Also scanners can be used to provide some of the downlink measurements, like CPICH measurements for the neighbourly optimizations. The radio network can typically provide connection level and cell level measurements. Examples of the connection measurements include uplink BLER and downlink transmission power. Connection level measurements, both from the mobile and from the network, are important to get the network running and provide the required quality of service for the end users. The cell level measurements may include the total received power and total transmitted power, the same parameters that are used by the radio resource management algorithms [7.4.7].

The measurement tools can provide lots of results. In order to speed up the measurement analysis it is beneficial to define the most important measurements called Key Performance

Indicator (KPI). Examples of KPI include the total base station transmission power, soft handover overhead, drop call rate and packet data delay. The comparison of KPIs and desired target values indicates the problem areas in the network where the network tuning can be focused. The network tuning can include the updates of RRM parameters, e.g. handover parameters, common channel powers or packet data parameters. The tuning can also include changes of antenna directions. It may be possible to adjust the antenna tilts remotely without any site visits. With advanced Operations Support System (OSS) the network performance monitoring and optimization can be automated. OSS can point out the performance problems, propose corrective actions and even make some tuning actions automatically [7.4.7].

The network performance can be best observed when the traffic load is high. With low load some of the problems may not be visible. Therefore, we need to consider artificial load generation to emulate high loading in the network. A high uplink load can be generated by increasing the Eb/No. If we increase manually the Eb/No target, e.g. 10 dB higher than normal operation point, that uplink connection will cause 10 times more interference and converts 32 kbps connection into 320 kbps high bit rate connection from the interference point of view. Another load generation approach in downlink is to transmit dummy data in downlink with a few code channels even if there are no mobiles receiving that data. The approach is called Orthogonal Channel Noise Source, OCNS [7.4.7].

7.4.2.5 Radio Network Subsystem (RNS) planning

Radio Network Subsystem is also called base station access network or Cellular Transmission access network. This section highlights the main issues that should be taken into account when planning future microwave radio networks for base station access or radio network subsystem (RNS) in 3G. The base station access part of 3G cellular network basically consist of two types of network elements called RNS (Radio Network Controller) and BTS or Node-B (Base Station). The task of the RNC is to manage the radio channels of BTS connected to it, and it also concentrates the traffic flows of connections and trunks them to the upper level core network. The base station handles the radio channels and forwards the traffic of lower level BTS towards dedicated RNC. BTSs are connected to the RNC, either directly or via some other BTSs in a cascaded way [7.4.8].
With the introduction of 3G, the increased capacity requirement will affect both the individual radio network links and the network topology. GSM network BTS capacities are on average 0.5-1Mbit/s while in 3G BTS capacities are in range of 2-10 Mbit/s. This huge capacity increase will force mobile operators to introduce fibre based solutions in regional networks and possibly change the existing network structure towards a more star type of topology. The introduction of 3G will also force operators to use the existing frequency bands as efficient as possible and also to find new frequency band. One of the key issues in coming years is how to optimize the usage of radio spectrum and how to guarantee the required quality of service. Microwave access, based on point-to-point microwave radios, is the dominating technology in the base station access network. It offers the fastest means for the network roll-out and capacity expansion. When using microwave radio transmission, an operator saves on the operational expenses compared to laying his own cable or leasing connections. At least two-thirds of the base station connections area based on the microwave radios. The guidelines for radio network planning are similar to other telecommunication networks in terms of traffic demand/capacity analysis and network topology selection. Radio network are subject to two more restrictions: 1. In order to build a microwave link between two stations, there must exist a Line-of-Sight (LOS) between them. 2. There must be available frequency pairs between the

two stations to be connected. Therefore LOS evaluation and link availability calculations are also essential steps in building a radio link [7.4.8].
The following facts effect to the access network and use of point-to-point radio links:

- Use of existing infrastructure; it is very costly and time consuming to modify network topology. Equipment should also be re-used as much as possible.
- Capacity often requires the use of certain media and equipment. Capacity depends on the topology and can not be chosen freely. The increase of capacity will force to change some topology models.
- Connectivity technology (TDM, ATM or IP) as such does not change the topology and the number of interfaces, etc.
- Frequency band for last mile media radio links are becoming difficult to get, driving to higher frequencies, i.e. also towards shorter hops.
- Fibres will be more and more available within few kilometres from any given base station in city area making star configuration more feasible in last mile connection.

The base station access network has currently a lot of tree and chain topologies, especially in rural areas. The requirement of increased capacities and use of radio link makes the last mile layer shorter – -suggesting a wider use of fibre loops and big radio link stars, see in Figure one. This topology most probably will be the favourite one in the future, starting from city areas [7.4.8].

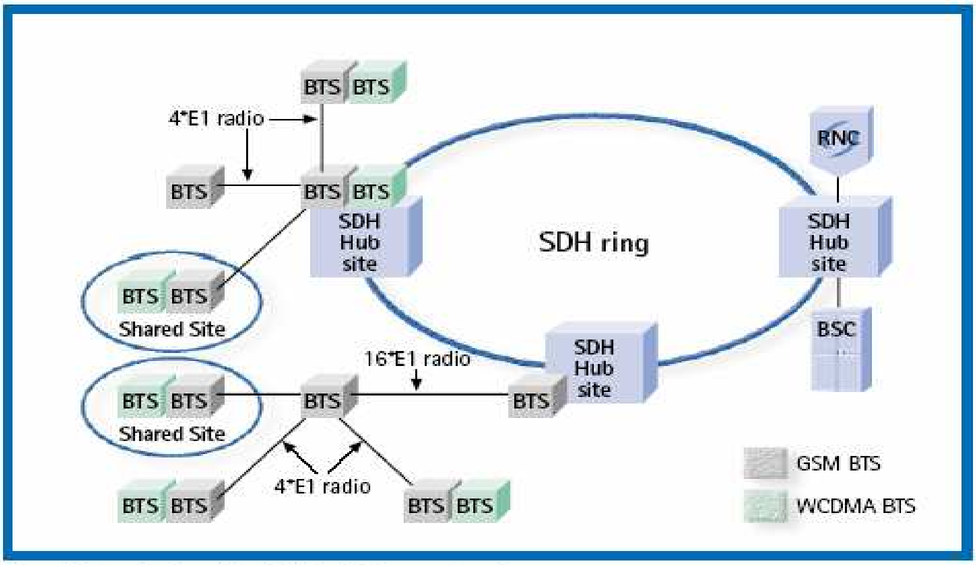


Figure 1. Example of combined GSM and 3G access network.

- Planning of Point to Point Microwave

The designers often have to make the network meet some performance objectives under some limiting facts like available sites stations and spectrum band [7.4.8].

Operating band
The network planning is mainly based on the availability at frequencies above about 17GHz. Below 17GHz design is normally dominated by error performance. In the tropics or other areas of heavy rainfalls this limit frequency may be lower (near 10GHz). The lower

frequencies allow longer hops, while at higher frequency high antenna gains are easier to achieve which makes handling of the interference easier [7.4.8].

Choice of frequency and polarization

Often it is advisable to choose higher frequency for shorter hops and use lower frequency for longer loops, if possible. One example is 30G Hz for hops below some 5-10 km. In particular, for very short hops below 1 km, we might use 58GHz radios as the interference is well under control due to high atmospheric attenuation. Besides, the unlicensed use of frequency band gives some flexibility to designers. The attenuation caused by the rain is lower for vertical polarization than for a horizontal one, so vertical polarization should be used for the long hops in the network while horizontal polarization may provide good network spectrum efficiency when used for shorter loops [7.4.8].

Path Design

Clearance for the hop is designed as usual, i.e. the first Fresnel zone should be free at normal k-value 1.33. It should be noted that relatively small obstructions, like a single tree in the radio path, might prevent signal reception at proper levels. Similarly, due to the high frequencies and corresponding small Fresnel zones, relatively small areas may act as reflecting surfaces. This is contrary to the design at lower frequencies (below 10 GHz) [7.4.8].

Modulation method

By choosing a modulation method with few states (for instance: 4QAM, MSK,ect.) or a system with good error correction capability, one might have relatively high tolerance against noise and interference, i.e. a low receiver threshold power Prxth. That will allow longer hops to be built and lead to best areal spectrum efficiency independently of the hop lengths. In some cases, point to point can have more weight, which may justify using modulation methods with higher number of states. Combing coding and high state modulation, e.g. in trellis-coding modulation (TCM) – may sometimes give a good compromise between point to point and areal spectral efficiency [7.4.8].

Transmitter power Ptx

When selecting transmission power Px, one should avoid using unnecessarily high power, although it is true that higher transmission power can improve availability and error performance of the system. Sometimes extra attenuators are needed to adjust the power. A more convenient way is to use transmitters with selectable or programmable power levels. Another effective way to avoid generating unnecessary interference into the network is to use adaptive transmitter power (ATP) where high power is used only during fading periods and otherwise a low power is used. A working ATP scheme requires that there is a return path in order to send information about receiving conditions to the transmitter. Use of error monitoring as a control parameters is crucial in achieving good network performance. At star points, where several paths converge to the same station, good rule of thumb for design is to have equal received powers for each path. Usually this means that at least some of the far and transmitter power should be adjusted. For very short hops one might even use somewhat smaller received levels than for the longer ones [7.4.8].

Receiver threshold power Prxth

This is mainly dictated by the selected capacity, noise figure and used modulation method. In heavy interference environment the effective receiver threshold may degrade considerably—

in tightly built networks about 3dB or even more. It should be reminded that low threshold powers enable longer hops if the transmission power remains the same. It also directly raises, in addition to filtering and other things, interference tolerance [7.4.8].

- **Current research on RNS planning and Optimization**

Ericsson Research Hungary
In the paper 'Planning of Tree-Topology UMTS Terrestrial Access Network', researchers at Ericsson Research, Hungary modelled each radio network subsystem as a tree, with the RNS node at the root. Basically in each tree there is an RNC as root node and given number of RBS. Due to some technical reasons there is a constraint on how many Base stations can be connected to an RNC, and on how many RBS can be connected to another one. In the paper these types of constraints are called degree constraints. There is also constraints on the depth of RBS sub-tree, which means how many other base stations, can be placed between any base station and the RNC, in other words how many base stations can cascade in the network. Further on this constraint is called cascading constraints. The level value of a base station means how many link hops are placed between itself and the RNC in case of the current network topology [7.4.9].

They suggest a method of planning a multi-constrained and capacitated sub-network tree with a previous dedicated root node. A heuristic planning algorithm that combines Simulated Annealing (SA) and a local improvement strategy was used to find a sub-optimal solution of a single RNS with a previous dedicated root node, i.e. they assumed the site of the RNC was already decided and its BSc already assigned. By detailed performance analysis of the proposed method using different network configurations and traffic demands, they get the conclusion that the proposal method is capable of planning the RNS or Base Station Access network in a cost-optimal way very close to global optimal. Furthermore, they demonstrate that, with optimization model and process, the algorithm can only work in UMTS technical background, equipment constraints and specific cost functions, but also be able to work in case of any other multi-constrained capacitated tree optimization problems with non-linear function. The drawbacks of this method is how to decide RNC locations and how to cluster BSs into RNC areas are not discussed ,which are very important for RNS optimization and more challenging than optimizing a single RNC tree.

Siemens Research
In Siemens' paper 'Network Planning and Topology Optimization of UMTS Access Networks', they also model each RNS as a tree. They divided RNS planning into two separate stages and develop algorithms to cluster BSs into RNC areas and algorithms for cluster internal optimization, Proximity graph (with the idea that BSs that are closed to each other should be clustered into the same RNC) is used for clustering, and greedy heuristic is used to find a cheap tree topology for each RNS. The disadvantage of their paper is there are no means to estimate the optimal cluster sizes, which are important guides when clustering BSs into RNCs. The separation of the clustering process and cluster internal optimization can speed up the optimization process, but it is less likely to find the optimal configuration [7.4.10].

Automatic RNS planning proposed by Luton University
In the paper ' Automation and Optimization of 3G Radio Network Planning' by Luton University proposed a automatic RNS planning method which make it possible to get the

optimal configurations for network of realistic size in reasonable time frames. Basically there are two aspects to consider: (1) Estimation of optimal RNS sizes and (2) RNS internal optimizations. When deciding the optimal RNS size (i.e. RNC coverage), a trade-off between RNC cost and link cost should be found. With large RNCs, the RNC cost is small since there are fewer RNCs, but the link cost is big since averagely each BS will send more hops to reach its RNC [7.4.6]. See figure 2

In the RNS internal optimization, optimization algorithm such as simulated annealing (SA) and genetic algorithms (GA) can be used to optimize RNS internal connections, the constraints that need to be considered in the optimization process are as follows:

- The constraints on the number of hops between a BS and the RNC
- The constraints on how many BSs can be aggregated to an aggregation node.
- The constraints on how many BSs can be connected to a RNC.
- There must be a pair of frequency between two stations in order to establish a link.
- Link capacities are only available in discrete values. E.g. Mbps E1 link and 4Mbps E2 link.

They also have the following future prediction in 4G RNS planning

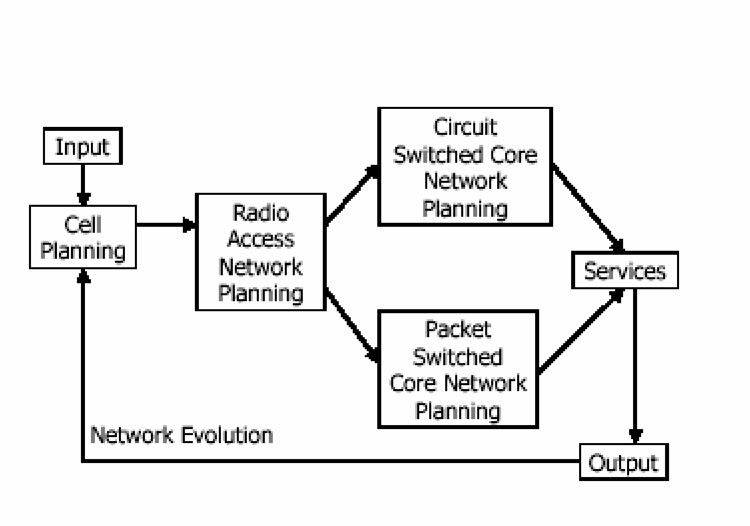
- Frequency planning will be a major focus in the future microwave networks planning.
- The topology of RNS can be not only tree, but also a hybrid of tree, star, and ring.
- Point-to-multipoint (PMP) links can be used in densely populated areas to reduce cost.

7.4.3. 2/2.5G Radio network planning for GSM /GPRS

7.4.3.1 Introduction to general planning process

GSM is the most widespread, most commonly deployed and fastest growing system standard for mobile telephony in the world. Even though UMTS, the third generation mobile system, makes its entry into market, GSM undergoes continuous evolution and development. High Speed Circuit Switched Data (HSCSD), Enhanced Data for GSM Evolution (EDGE) and General Packet Radio Service (GPRS) can be added to GSM network resulting in new features, new functionality and increased data rates. A GSM system with implemented GPRS functionality provides the core network required also by UMTS. Hence, the platform us already prepared for the migration into third generation mobile system, UMTS.

Figure below shows the network planning process for GSM/GPR network [7.4.1]. We will look at the input that will be required and then we will present each of the planning processes and the required information to build the network plan.



Input

We will start with the information that is used as input in the network planning process.

The requirement can be broken into three categories.

The first is the landscape which includes[7.4.13]:
- Geography
- Market information
- Size of the area to be covered by the wireless network
- Type of area (i.e. is this going to cover down town core area or rural or both)

The second category is the requirements that a subscriber puts on the network. They include:
- The subscriber (sub) profile
- Subscriber usage model
- Mobility model
- Data demand model
- Grade of Service (GOS)

Lately, the technology and operational requirements includes:
- Technology capabilities
- Component capabilities
- Cost to purchase the equipment
- Cost to obtain the facilities
- Building rental space may also be required

Cell Planning
Cell planning is based on a number of models. This includes the frequency-planning model (depending on the type of technology used). This also includes the call traffic capacity model and spectrum efficiency that is used depending on the technology and the available spectrum. The cell coverage model is used to estimate the coverage of a cell based on the type of area. For example, a cell covers more area if there are no buildings in the way. [7.4.13]

Radio Access Network Planning
Radio access network planning consists of estimating the capacity of the base station controller (BSC), the physical location of the BSCs, and the planning of the backhaul of radio access network. This includes an optimization of the physical location of the BSC based on backhaul cost [7.4.13].

Circuit Switched Core network Planning
Circuit Switch Core Network Planning includes the MSC and HLR capacity planning, the planning of the SS7 network, the planning of the transport facilities between the components, and the planning of the locations of the MSC and HLR, based on the cost of the network. [7.4.13].

Packet Switched Core Network Planning
The packet switched core network planning includes the Packet Data Node (PDN), the data subscriber profile database (i.e., the HLR in GPRS and AAA server in CDMA 2000) capacity planning, the planning of the IP network, the planning of the ISP transport facilities between components, and the planning location of the packet data nodes based on cost [7.4.13].

Output
The output of the network planning process is used by the network deployment and engineer teams to deploy the network. This output includes the network requirement and topology. With the provided topology, an estimate is provided of the network capacity and an analysis of the potential bottlenecks. The network and nodal performance, including the sensitivity analysis, is provided so that deployment engineering will know what area of network will need to be visited as the number of subscribers in the network increase [7.4.13].

Service Network Planning
Service network planning is used to provision the number and type of service nodes that are required. This includes the location for the service nodes and the required transport faciliteis

Network Evolution
As briefly mentioned with the network output, the number of subscribers will hopefully grow over time. This will force the Wireless Service Provider (WSP) to change the network to meet those needs. Some of the changes include future services (high speed packet data). The call model used to do all of the original planning was based on a number of assumptions (i.e. subscriber traffic patterns, mobility models, ect.) After the network is deployed, we are able to look at the specifics of real call model and may have to change the network to meet the real needs versus the projected needs. The technology of network will also evolves as the vendors increase the capacity based on the new processors and software optimizations. These will cause the capacity of the network increase without having add new components to the network and may potentially change bottlenecks of the network.[7.4.13].

7.4.3.2 Cell planning for GSM/GPRS

In GSM/GPRS cellular network, the cell planning can determine how effectively the allocated spectrum is used dependent on the cell plan that includes: the number, the location, and the configuration of the base station. Poor cell planning will waste frequency spectrum a lot. Within a selected geographic area where the radio network must be installed or extended, operators define the number of frequency to assign to the area. These parameters are then used for radio site positioning and configuration and frequency assignment. The purpose of the Site Positioning and configuration is to optimize the radio coverage of an area. On the other hand, the main objective of the Frequency Assignment Problems is to minimize the electromagnetic interference due to multiple uses of frequencies in different parts of the network [7.4.14].

Cell Planning Model

A complete cell model can be found in [7.4.15]. A radio access network is composed of three objects:

- A discrete geographical working area, where signals and traffic are measured:
- Receiver equipment. e.g. mobile telephone, which define the service requirements within the working area;
- Antenna, which can be located on some pre-defined sites within the geographical working area.

Working area

A number of points are defined on the working area [7.4.15]:
- A set, R, of reception test points (RTP) at which propagation information is recorded.
- A set, S, of service test point (STP), where the radio signal must be higher than a specified threshold e.g. –90dB (the exact threshold is based on the type of receiver equipment at the STP).
- A set, T, of traffic points (TTP), which gives the traffic demand, measured in Erlang's, at the point. This measurement the peak capacity requirements at the point.
- A set, Z, of candidate sites at which antennae could be placed. It is assumed in this work that at most three antenna could be placed t particular candidate site (different values could be used as appropriate).

The relation is $T \subseteq S \subseteq R$

Receiver Equipment

A network provides a service for all receiver equipment represented in the working area. A service threshold, S_q defines the required level of service at each STP (which can vary throughout the working area if different services are provided). The exact value of the threshold at a STP is dependent on the equipment at the point. Examples are given below[7.4.2]:

Types of Service	Threshold S_q (dBm)
8 Watt outdoor	-90
2 Watt outdoor	-83
2 Watt in car	-82
Indoor	-75
Deep indoor	-65

Antenna

Many types of antenna exist with different characteristics such as radiation pattern and transmission gains and losses. In this chapter we will assume that three types of antenna are available that could be deployed at a candidate site. These are

- An omni-directional antenna (omni)
- A small directional antenna (small)
- A large directional antenna (large)

The associated gains and losses of these antennas are shown in table below

Antenna	Gain (dBm)	Loss (dBm)
Omni	11.15	7.00
Small	17.15	7.00
Large	15.65	7.00

To configure an antenna, values for power, azimuth (horizontal direction of antenna relative to North) and tilt (vertical angle of the antenna, relative to the horizontal) need to be allocated by the cell planner.. The range for each of these values, used in this work, is as follows [7.4.14]:

- Power: 26 to 55 dBm, in steps of 1dBm,
- Azimuth: 0 degree to 359 degree (for directional antenna only), in steps of 1 degree
- Tilt: 0 degree to –15 degree (for directional antenna only), in steps of 1 degree.

Network Requirements

When modeling a cell plan two distinct issue need to be addressed. These are the RF requirements and the teletraffic capacity requirements. Most of the papers on cell planning have paid considerable attention to coverage problems. That is, designing cellular networks that ensure that all STP are served by at least one antenna with adequate signal levels, i.e. area coverage problem are satisfied. Economic factors are then portrayed by minimizing the number of sites and antenna used. In practice, cellular networks are capacitated, that is, the level of traffic that can be assigned to a particular antenna is limited. This generally forces a limit on the cell size and increase the number of required antenna sites. In FTDMA system such as GSM each channel can support up to 8 calls, hence in a busy street there is likely to be a requirement for two or even more channel to cope with the anticipated traffic. The number of channels that an antenna can emit is limited by the technology involved, hence the size of high traffic cells may have to be reduced despite that the RF environment could provide coverage a much wider area. A detailed analysis can be found in [7.4.14].

The number of channels available to an antenna is limited, hence also is the traffic that can be successfully serviced by the antenna. By examining the Erlang-B table the number of channels required to support the varied traffic demands can be found. An antenna can support traffic demands of up to 43.0 Erlang. These traffic capacities refer to the number of channels required to support that demand in the absence of interference. These are shown in table below [7.4.2]

Channels Required	1	2	3	4	5	6	7
Erlang Carried	2.9	8.2	15	22	28	35.5	43

If the sum of the traffic values of TTP in a cell (its traffic load) exceeds the maximum given value (43 is used as default parameter), traffic cannot be guaranteed to be fully serviced or some traffic may be dropped i.e. users will experience a "busy" tone. The amount by which an antenna exceeds its maximum capacity limit is termed the overload.

Design Objectives

The objectives of a network design (cell plan) are as follows.
Coverage: All STP are covered i.e. receive at least one signal above its service threshold.
Traffic Capacity: The traffic load within all cells should be less than some maximum value corresponding to the capacity of the maximum number of TRX devices available. Currently the value used corresponding to 43 Erlang, the capacity of 7 TRX devices (see table above)
Interference: This objective attempts to minimize a measure of potential interference in the design. Our measure is to minimise the number of overlap STP i.e. at each reception point, minimize the number of interfering base stations, where an interfering base station is defined as a base station providing a signal strength at a reception point that is greater than the receiver sensitivity (-99dBm) but is not the best server or one of the base stations providing handover.[7.4.14]

Optimization

Details of the optimization framework used for generating cell plans can be found in [7.4.3]. The only difference is the components used in the cost function (since only coverage, capacity, and interference are considered here). The components used in the cost function (to be minimize) include the following [7.4.14]:

- The number of uncovered STP
- The amount of uncovered capacity required i.e. the difference between the total traffic capacity required minus the traffic capacity of the current network design.
- The number of interfering base stations over all STP.

Each component is normalized and weighted with a value between 0 and 1.

7.4.3.3 GPRS planning over GSM

GPRS is a new packet data service and operated on top of existing GSM networks. As the first trial GPRS is already available in some countries, GPRS network planning becomes an imminent issues for GSM network operator. In this section, the principles of GPRS network planning will be discussed in details. First we will look at the general requirement of GPRS service, then the remaining capacity of existing network will be calculated after which the issue of GPRS capacity and downlink performance will be investigated. Finally we will discuss the coverage and capacity planning will be addressed. [7.4.15][7.4.16]

Service Characteristics and Requirements

GPRS is a set of new GSM bearer service that provides packet mode transmission within the PLMN and internetworks with external networks. GPRS should support wireless application such as retrieving information from wireless information centers and mobile offices, including WWW surfing, file transfer, remote network login, stock market information transfer, lottery transaction, ect. The traffic of those applications indicates the characteristic from frequent transmission of small volumes to intermittent and non-periodic (transmission of median volumes of data, and infrequent transmission of large volumes of data. There are several GPRS traffic models. One is called FUNET model which maps packet size X (in Kbytes) to a truncated Cauchy distribution. In the model of Railway, the packet size is treated as a truncated exponential distribution. For GPRS network planning we can apply those models, but how we model WWW surfing? A model has been suggested by ETSI which introduce the ‘packet service session’ concept. For WWW browsing, the traffic load in downlink is higher. The asymmetric of the traffic should be taken into account when planning GPRS network since downlink always have higher load than the uplink. [7.4.15]

There are two types of GPRS services: connectionless and connection-oriented point-to-point transmission (PTP) & point to multi-point (PTM) services including multicast, group call and IP multicast transmission. The resource allocated to GPRS is flexible depending on the local traffic condition and can be online negotiated. This negotiation set the quality of service parameters such as user data throughput, QoS class (transfer delay, priority) and reliability of transmission to a certain values or default values. In a word, the GPRS planning should provide the dynamic range of the resource enabled to be used by the network based on the system parameters [7.4.15].

Since the GPRS is operated on top of GSM network, there are two principles we should consider when planning GPRS. Firstly, the introduction of GPRS may effects the performance of existing voice service, such as increasing blocking probability and call dropping rate. Therefore we should try to minimize it in GPRS network planning process. The second one is that the reconfiguration of the existing network due to the introduction of GPRS must be kept as small as possible to minimize the additional GPRS deployment cost[7.4.15].

The procedures for are as follows:

- Firstly we evaluate the signal to interference ratio (SIR) level in existing network by measurement or simulation or both. A simple way to obtain the SIR level of the existing network is to utilize the signal quality measurement report in BSC, such as the RXQUAL report. There are also other ways in [7.4.16].
- After analyzing the SIR level of the network, we may conclude if the network can accommodate the GPRS traffic or not. But still we need to know the average capacity obtained by overlaying GPRS onto an existing GSM network; the maximum resource can be used for GPRS without causing the quality of voice service under the acceptable limit.

Capacity Planning

As GPRS application vary from high-bit-rate services provided for business application to low-bit-rate services like the normal file transfer, the resource occupied by a GPRS user may vary from 1 to 8 timeslots. The traffic density generated will depend on the environment characteristic (downtown, suburban, business area, rural area), the mix of terminal types, the demographic situation, the penetration factor of GPRS service. According to these factors, the distribution of the traffic demand can be estimated. The traffic can be sorted into groups according to the basic bandwidth or timeslots required for each group [7.4.15].

Since the busty characteristic of the traffic, the variable data rate and the packet switched transmission mode, the GPRS traffic capacity required can not longer be specified by a single unit (Erlang) based on voice service. The transmission rate can be represented as kilobits per second per square kilometer (kb/s/km^2). The main task of GPRS capacity planning are not only to try to meet the capacity required by traffic with an acceptable blocking probability and delay, but also to provide decision criteria to network for online QoS negotiation. To minimize the effort and cost for the network operator, the original network configuration including cell planning, frequency planning, setting of power and other cell parameters, must not require extensive modification. In the initial phase of GPRS launch, the network reconfiguration might not be needed, As the GPRS getting more and more popular the reconfiguration might not be avoided, e.g. more base station are needed to be installed[7.4.15].

Due to the bursty characteristic of the GPRS traffic and multi-rate parallel service provided, the Erlang B or Erlang C formula can not be applied in capacity planning. GPRS is supposed to utilize those resources which are not used by the voice services. It does not require permanently or temporarily some physical channels for GPRS traffic. In addition, rather than having a dedicated PCCCH, the GPRS may utilize the existing GSM paging and control channels. For such a dynamically variable resource and the bursty traffic, the capacity offered by the network main only be able to obtain from simulations[7.4.15].

Coverage Planning

The main purpose of coverage planning is to achieve the required radio coverage with specified time and location probability. In traditional voice service it is achieved by the link budget within the range of the transmitted power level. Since GPRS is deployed on top of an existing GSM network, the same link budget as that of voice services may be used for GPRS. However, we need to consider that the outage probability near to the cell border area will increase as more channels used for GPRS. If the outage exceeds the network target it will reduce the real served area of a cell and may cause a higher dropping rate of inter-cell handover after introduction of GPRS [7.4.15].

The required SIR values for GPRS coding

Coding Scheme	Code Rate	Data Rate (kbps)	Req.SIR without FH	Req.SIR with FH
CS-1	1/2	9.05	13dB	9dB
CS-2	2/3	13.4	15dB	13dB
CS-3	3/4	15.6	16dB	15dB
CS-4	1	21.4	19dB	23dB

For the four coding schemes used for GPRS, the specification of GSM 05.05 has defined the required SIR values (the table above) corresponding to a BLER (block error rate) value not more than 10% for systems with and without frequency hopping (FH). Those SIR values have been included an implementation margin of 2dB. The coding scheme CS-1 is mainly used for the GPRS signaling and is the same scheme which is used in the GSM signaling channel SDCCH. Almost every required SIR value for those coding scheme is higher than the 9 dB target value required for GSM voice service. In addition, when more channels are allocated to GPRS, it has an impact on the quality of both voice services and GPRS. And the SIR may also degrade to level of 9-13 dB. Therefore, the GPRS service might not have coverage in some areas in a cell for some networks. The network planning should at least try to cover the areas with a high service demand. If the degradation is a temporary issue, it will not be a problem for GPRS because the transmission can be delayed and retransmitted. However if the effect is a long term problem, we need to recon-figurate the networks, e.g. installing new base stations, or having a larger reuse factor and more frequency carriers. The varying SIR target values for different coding schemes may let the users to achieve a higher data rate near the base station and a lower data rate near the cell border in generally. We do not need to define which coding scheme should be used beforehand [7.4.15].

Conclusion

In GPRS network planning, the most important issue is to evaluate correctly the remaining capacity of an existing network considering interference constraints. Otherwise, it may damage the GSM voice services. The maximum radio resources that can be allocated to GPRS are defined by the network remaining capacity and are dependent on the outage or interference level of the existing network. It needs to consider that the outage probability near to the cell border area will increase and the real served area of a cell may be reduced due to the introduction of GPRS. That implies that GPRS may cause a higher handover dropping rate to GSM voice services. GPRS capacity performance in downlink is quite different from that in uplink because of the difference in transmission protocols. The GPRS transmission efficiency is highly affected by the packet size of the data traffic. It will cause both low efficiency of transmissions and high signalling load if the data packet size is too small [7.4.15].

7.4.4. 2/2.5G radio network planning

7.4.4.1 Introduction to automatic cell planning

One of the most important cellular planning activities is to select a number of sites from a list of candidate sites that have been identified as potential sites by marketing. The selected sites form the basis of a network that must satisfy certain network requirements such as high area coverage and high traffic capacity but that minimize the infrastructure cost. The configuration of the selected base stations is also a complex problem and involves choosing among different antenna types, e.g. various directional or omni-directional antennas, power control, tilt, and azimuth. Operators can get competitive advantage by minimising commitment to infrastructure while maintaining appropriate levels of spatial and temporal access to wireless services. Cell planning is challenging due to inherent complexity, which ranges from requirements concerning radio modelling and optimization. This is particularly acute for UMTS services where WCDMA protocol leads to interdependency between cell coverage and, quality of service and capacity. Power control and multi-service traffic lead to dynamic cell footprints which significantly complicate the cell planning and dimensioning process [7.4.18].

Traditional cell planning methods

Methodologies for planning a new network vary from operator to operator, although there are certain consistencies in approach. Generally, a planner will be given a set of financially and technical requirements and will set about developing a plan that meets them, by choosing sites locations from candidate alternatives, antenna types, tilts, powers, ect. Conventionally, experienced planners can use their knowledge to constrain the number of configurations considered (such as reducing the number of potential sites), however in doing so, many potentially high performance cell plans may be removed from consideration. On a manual basis, a natural way to proceed is via construction of a plan using simple “rules-of-thumb”, making visual inspections of plots and then resolving problem areas iteratively. The end result is generally one single plan which represents a potential feasible solution, although the operator may ideally wish to see a number of alternative plans, to access the trade-off between cost and benefit. The accepted plan is also susceptible to failure due to site acquisition problems, requiring further iteration of the plan which introduces the danger of degenerative plan performance and increases in cost. Therefore there are needs in overcome those problems in traditional cell planning [7.4.19]

1. rationalise the planning process so that it is not entirely dependent on a radio engineers wide-area cell planning skills, a task which may infrequently be taken;
2. better use a radio engineers training in the cell process, by focusing the radio engineer on analysing and interpreting optimal alternative plans for potential deployment, rather than engineer performing the site selection task, which is a combinatorial rather than radio engineering problem

Automatic cell planning

Automatic cell planning is defined as a cell planning process where the computer has complete autonomy in the selection of sites and the configuration of the transmission infrastructure, subject to constraints and objectives imposed as input to the process. Basically automatic cell planning including the following aspects: adaptive propagation model, a reasonable mathematical model of 3G radio network, a couple of sophisticated heuristic methods to short-cut the enormous amount calculations and mathematical optimisation methods [7.4.20].

An adaptive propagation model can enable the fully automatic generation of accurate predictions for all different operational environments. With a general framework for an adaptive propagation model, one can process digital terrain data of different resolution and granularity and contains methods and criteria for unsupervised selection of proper methods for the different operational environments [7.4.20].

In order to be able to handle a complex system like UMTS radio network, a mathematical model for feasible network configuration is extremely needed. This model can be used for analytical studies of structure of the optimization problem as well as a starting point to develop the optimization methods. In the IST project MOMENTUM, two principal approaches for the optimization problems have been developed. The mixed-integer programming approach contains heuristics for both mobile assignment and installation selection. For the heuristics different methods have been tested revealing Tabu Search, Greedy and Set Covering as the most promising approaches [7.4.20].

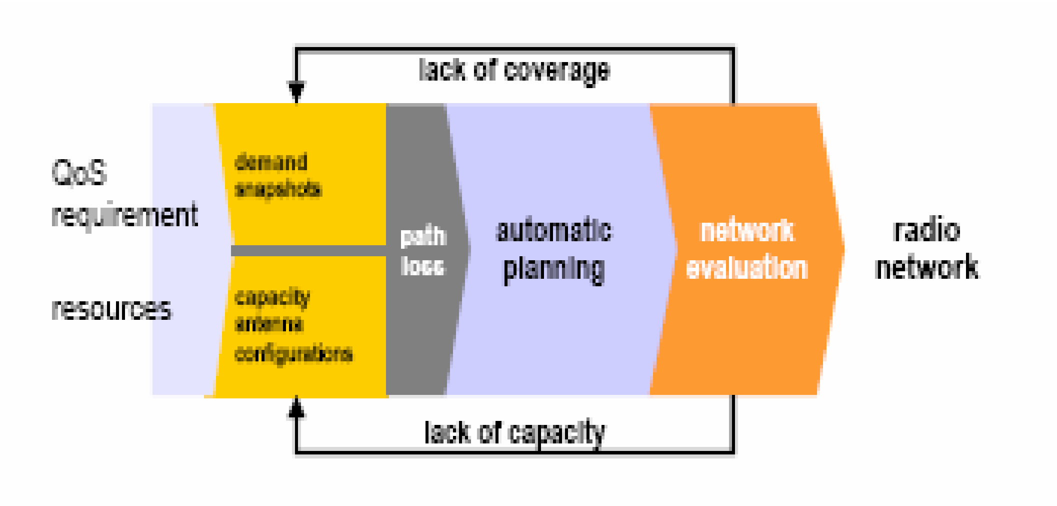


Figure Components of automatic cell planning

7.4.4.2. Activities on adaptive propagation model selection

There are many planning tools these days for cellular planning purposes, within which propagation plays an important role, for either coverage problem or interference estimation. In these tools, the radio planner still has a key role, for example when determining in which area a specific propagation model should be used, since no automatic choice is given. Although the science of propagation models is an area of intensive research, a universal propagation model applicable to all possible propagation situations is not available. Reasonable results in propagation modelling are achieved by finding more or less accurate models for the most dominate propagation phenomena observed for specific applications. The specific application area of a propagation is described for example by the carrier frequency, the typical antenna heights for both base station(BS) and mobile station (MS), the distance between them and the structure of the environment (indoor/outdoor, build-up/open/forested, etc.) in the reception area of the signal. One consequence of the requirements for this specific application is the availability of propagation models applicable only within a restricted validity range. Furthermore, these propagation models require digital terrain models (DTM) which may be different either in content (e.g. land use vs. detailed building data), granularity (e.g. different number of land use classes and/or attributes) and/or resolutions 7.4.20].

In the IST project “Models and Simulations for Network Planning and Control of UMTS” short for MOMENTUM, an automatic radio network planning approach is investigated. This approach covers all major aspects of automatic planning including an adaptive propagation model applicable for all relevant deployment scenarios, supplicated heuritics to reduce calculation times, and mathematical optimization methods. Here we focus on the adaptive models developed in this project, which is a key aspect in automatic radio network planning. In MOMENTUM, a general framework for a fully automatic and adaptive selection of propagation models has been introduced addressing the integration of different models in terms of [7.4.20]

- different deployment scenarios
- use of different digital terrain databases
- the identification of parameters for the selection of different models and /or the transition between models.

Basic Components of an adaptive propagation model

In a UMTS network the full range of cell types will be used. This covers macro-cells, which are deployed in rural and suburban areas, small macro-cells, and micro-cells occurring in urban areas as well as pico cells in hot spot areas like airport and exhibition halls. In the latter case indoor solutions are applied. These indoor base stations are at least a potential interfering source in the outdoor area. Since also signals from outdoor base stations can be received within a certain penetration loss at indoor environments a complete description of the interaction between indoor and outdoor areas is important [7.4.20].

Typically low resolution data is available for all environments, whereas the more expensive high resolution data is typically available for the dense urban areas only. The corresponding areas can be defined as follows:
A1: Area where only low-resolution is available
A2: Area where also high-resolution data is available. Subdivision into:

- A2a: outdoor areas
- A2b: indoor areas

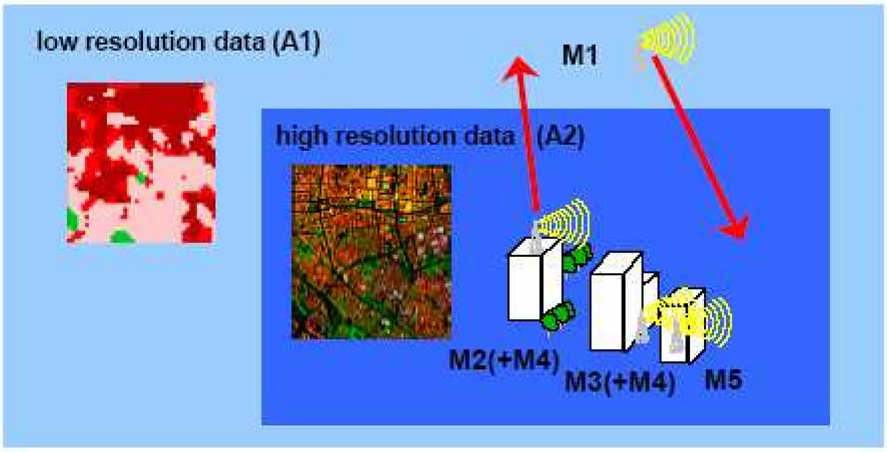


Figure 2-1: Propagation environments and configurations of practical interest

Theoretically all possible combinations of cell types, deployment mode and DTM availability have to be considered. However, not all possible configurations are of practical interests. Therefore only the following propagation models will be used in adaptive UMTS propagation model, see figure above [7.4.20].

Macro cell models using low-resolution data (M1)
Small macro cell models using high-resolution data (M2)
Micro cell models using high-resolution data (M3)
Outdoor-indoor models using high-resolution data (M4)
Indoor-to-indoor models using high-resolution data (M4)
Indoor-models using high-resolution data (M5) and their extension to the indoor-to-outdoor scenario

General Approach

For complete interference calculation of the network, a matrix containing the mutual coupling of cells is required. This coupling matrix is computed by a superposition of predictions with different models and DTM. Therefore a couple of model extensions, transition models and switching criteria between models are required. The necessary developments can be grouped into three main tasks [7.4.20]:

1. For those cases where the prediction area of a cell covers different DTM appropriate model extensions for the transition between different data source are required. This includes a model extension of M1-type models in order to exploit high-resolution data in some parts of the prediction area (BS in A1, MS in A2), whereas M2-type models need an extension to low-resolution data (BS in A2, MS in A1).
2. In dense urban areas three different cell types-small macro, micro and pico cells-may be deployed. Therefore switching criteria between M2-, M3-, and M5-type are required. Since the decision between M2 and M3 is not binary smooth transition functions are required.
3. The interaction between indoor and outdoor configurations (outdoor coverage by indoor base stations an indoor coverage by outdoor base stations) requires model extensions (M4) to the corresponding models M1/M2/M3 (BS in A1 or A2a, MS in A2b) and M5 (BS in A2b, MS in A2a) respectively.

For detailed description about adaptive propagation models, please refer to [7.4.20]

7.4.4.3 Review different algorithms used in cell planning

In cell planning fields, we need to consider the traffic demand to cover a specific region, availability of base station sites, available channel capacity at each station, and the service quality at various potential traffic demand areas (TDAs). Selections of good base station sites and channels will result in acceptable coverage area at base station both in coverage area and in signal quality. There have been a lot of researches on how to optimize the base stations sites and configurations.

In [7.4.22], the radio coverage optimization problem is converted to a maximum in dependent set problem. The objective is to achieve a large coverage of TDAs with a small number of base stations. A simulation method is employed to examine the relationship between the number of base stations and relative coverage of TDAs
Optimal location of transmitters for microcellular system is studied by Sherali [7.4.23]. The path loss at each area is represented as a function of the base station location. A nonlinear programming problem is presented which minimizes a measure if weighted path-losses. Several nonlinear optimization algorithms are investigated to solve the problem. Tutschku [24] proposed an automatic cellular network design algorithm without considering the capacity of transmitter. The network design problem is converted to a maximal covering location problem by using demand node concept. The location of the transmitter is optimized by minimizing co-channel interference. In [7.4.25], he also proposed a greedy heuristic to solve the maximal coverage location problem of transmitters. The heuristic takes into account all the RF design objectives as well as the capacity and network deployment constraints. In [7.4.26], [7.4.27] a genetic algorithm approach is presented by Calegari. The selection of base station is represented in a bit string. Selection based on fitness value, one-point crossover and mutation operators are employed. The fitness value combines two goals of maximizing the cover rate and minimizing the number of transmitters. To speed up the procedure a parallel

genetic algorithm is implemented by using island model. Their computational results show that the solution quality is significantly influenced by the number of islands [7.4.21].

7.5. Additional design and dimensional problems

The changing standard of quality

Quality is a subjective term, with benchmarks and the perceived level changing over time and place.

In an area where very limited or no basic telecommunication services are currently available, it is perceived as a significant improvement when basic services are first introduced, even if the service was a bit clumsy to access (e.g. need to find a location where network coverage is available), somewhat congested from time to time, or available only part of the time (e.g. due to daily outages in electricity in that region).

However, as people become accustomed to the service, or have had the chance to enjoy better service (e.g. as result of competition of the service providers), the perception and standard of quality increases, and finally reaches the level currently enjoyed in many of the highly populated cities.

Therefore, to keep the initial costs down and meet the constraints of time and competition, operators must recognize and appreciate the need for a balanced level of compromise in order to make the service available in the first place. Again, such compromises must not sacrifice the ability to further grow and develop the networks to the current market benchmark.
Radio access – smart choices are required to keep network costs down
Often more than $^{2}/_{3}$ of both the network investment and operational expenses come from the base station network, with the remaining part from the core network, including switching, network management and so on. The natural reason for this is that the radio base station sites far outnumber those needed for the more centralized parts of the network, and are located all around the network area.

To implement network services at an affordable cost per end-user, it is of the utmost importance to have the radio access components built and dimensioned correctly.

Minimizing the number of sites, maximizing subscribers per site

A logical first step, particularly in a low traffic density region, is to minimize the number of base station sites, and to maximize the number of subscribers served by each site (thus contributing to minimized cost per subscriber).
As an example, building a low-capacity base station site (1+1+1 TRX in GSM) provides 4-5 times higher cost per subscriber than having the higher number of subscribers served by a high-capacity base station site (4+4+4 TRX) where there are more subscribers to share the site cost.

Getting more out of the network

It is beneficial for the operators in highly populated regions to apply new spectrum efficiency increasing techniques, to stretch the network capacity even further. Typically, the approach of providing more capacity from the existing sites instead of having to build more sites clearly provides the best economy.

Planning and implementing the network

Operators continue to report speed of rollout as one of their key differentiators in a competing market. The operator with the fastest rollout or ability to expand the capacity in the optimal manner is the more likely to attract new (and loyal) subscribers.

The quality of site implementation is a key element in limiting the number of site visits required by maintenance teams to fix or finalize their work. A vendor's experience in *getting it right the first time* reduces the need for additional site visits, hence vastly improving the time-line and cost plans of even the most experienced project manager.

The way to combine rapid deployment with good cost control is to do the planning and implementation in a systematic manner. Factors such as; choosing the target area and planning criteria, preparing for future coverage and capacity expansions, planning the needed network solution using standard site configurations, and proceeding only when the work is fully completed, are critical to a rapid implementation.

The role of standardized site solutions, i.e. keeping to a minimum number of different site configurations, is essential in overall cost management, as it gives savings throughout the process: standardized requirements for sites enable easier and faster site acquisition, a minimum number of base station configuration variants speeds up the network planning and simplifies staff training, controls the logistics cost and enables fast and correct site implementation in volume. Also network maintenance will be easier due to the usage of standard site solutions.

Selection of the right base station site concept

Choosing the correct site concept provides significant site-level savings. A new base station site of marginal profitability could turn into a good investment, if a more economical site solution was found.

If ready premises for the equipment do not exist, the options are as follows: to apply an indoor site solution using equipment rated for use in an indoor environment, thus requiring a weatherproof housing to be built or rented; or whether to implement an outdoor site solution with all the equipment readily installed in a weatherproof cabinet suitable for environments such as a rooftop, on an existing tower structure or at ground level. The latter is usually more economical, considering the overall cost for putting up a new site, and may contain battery backup and transmission solutions built in.

Lack or quality of electricity as a challenge

A restricting factor in many cases that increases the radio access site cost is the unavailability or poor quality of electricity at base station sites. The traditional mobile network approach has been to overcome this by installing generators and battery backup systems. This is a costly approach, considering the initial investment (generator, housing for protection from weather and theft), ongoing operating expenses (regular test use, maintenance, fuel management) and social compatibility (permission to store enough fuel, working noise, management of exhaust gases and vandalism).

However, in regions with no previous telephony services, the operator's income stream may not allow for such costly implementations. The high power drain of sites, combined with the difficulty of providing a proper supply of electricity, may contribute to the decision not to extend the service to the region at all, or to put the expansion on hold for an undetermined period of time.

While the base station equipment is the main cause for energy consumption on a base station site, the second important drain on supply is the systems to support the base station, primarily the air conditioning system. As air conditioning systems are also a major maintenance problem, and target for theft, it makes sense to try and avoid their use whenever possible.

Working around the electricity problem

Passive or fan-only cooling is typically not an option for traditional high-power indoor base stations in hot climate countries. A potential alternative is to instead consider the use of temperature-tolerant outdoor base station types; either a high-power traditional base station or a low-power-consumption siteless base station, even allowing the use of solar energy as a source for electricity, and not requiring the use of air conditioning equipment.

The concurrent impact in site cost reduction and reduced power consumption overcomes the majority of the electricity and site maintenance concerns.

When applying a battery backup system, it is worth considering whether to dimension the backup capacity for all-time full-power use of the base station, or for the more likely typical partial load, e.g. 50-60% of the maximum site power consumption, as this makes a difference particularly from the investment point of view.

The batteries required for a solar-powered system can, in favourable cases, be buried in the ground to achieve a more optimal operating temperature, thus further decreasing the maintenance cost.

Transmission in less densely populated areas

It is common practice to extend the transmission system to the base station in conjunction with the rollout of electricity or roads to the site. Traditionally, optical fibers are laid at the same time as other services, to be rented or sold as an asset for providing a backbone transmission network. From centralized hub sites, the telecommunication operators would then typically extend this to their sites using microwave radio link technology.

Reliability considerations

It makes natural sense to pay strict attention to the reliability of the base station equipment at the site, as this is often one of the largest hidden costs in the network. As the conditions at sites may be harsh over the years, the distances to sites long or access difficult during some seasons of the year, and any maintenance laboriously expensive, the failure of equipment has a greater impact on the availability of service, as well as overall operational expenses.

Remote preventive maintenance, diagnostics and manipulation practices should also be standard features for the base station sites, so that time-consuming traveling to and from sites can be taken to the absolute minimum.

The "art of starting with minimum investment"

The "art of starting with minimum investment" -- each service will surely then evolve to fulfill more demanding requirements later on -- but the question if the entry step can be made economically may decide if there will be (the capability to invest into) any new service at all.

This potentially touches also the work of telecommunication regulators -- whether to allow initially e.g. somewhat higher blocking rates, and only require improvement when the service has properly established.

CAPABILITY FOR (MASS TRANSACTION) Prepaid IS OF KEY IMPORTANCE

The majority of new subscribers in next years are expected to have prepaid as the payment method. As the anticipated volume of subscribers grows significantly, it is vital that processes and cost management practices for prepaid subscriber provisioning and prepaid account recharge have been tuned for maximum efficiency.

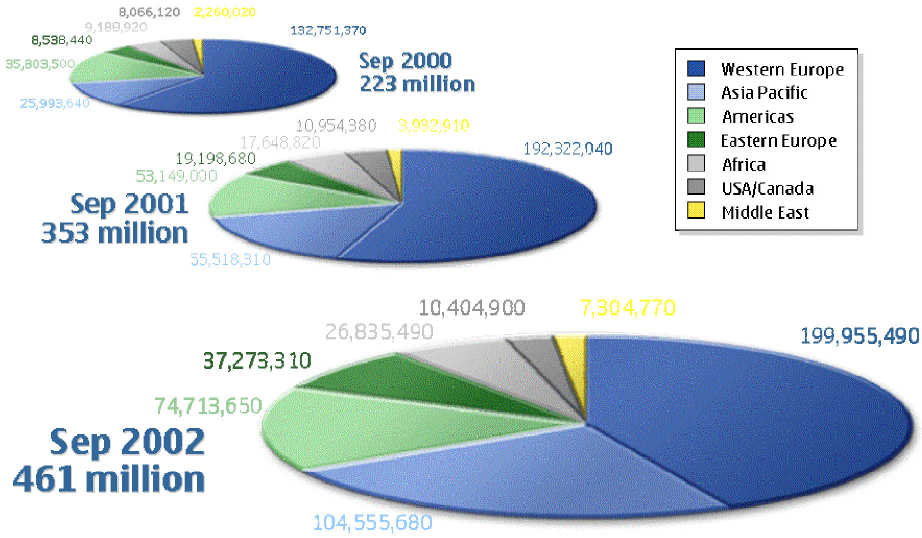


Figure 3. Prepaid has grown rapidly as a charging mechanism (source: EMC Cellular Subscribers)

A significant factor affecting the operator's market strategy is the value of the lowest prepaid denomination. Using Asia as a yardstick here, operators who are more successful in attracting volumes on entry subscriptions are now offering prepaid subscription values starting from as little as USD 0.50, in contrast to the previous dollar plus or USD 5 minimum top-up values.

By encouraging continuous recharging in this manner, an operator keeps the subscribers active in the network - even during periods of lower communications activity. Thus the customer is kept in reach of in-coming calls, enabling the operator to collect revenue from

other subscribers and possibly even from other operators' call-termination network fees - even when the end-user him or herself did not spend the money on communications.
The technical and financial challenges resulting from a higher number of low-value top-up transactions naturally include a greater number of top-up transactions to be managed by the operator. If the operator's operating expenses per transaction are high, this can soon turn into a poor business case requiring further proactive actions.

The costly process of handling paper format prepaid recharge-coupons (scratch cards), and the potential economical risks linked to their handling and distribution have led to the rapid development of electronic recharging. This has not been with the more advanced solutions, such as ATMs in countries with extensive electronic banking networks, but with the electronic reselling of airtime from small distribution points, such as village kiosk keepers or other individual entrepreneurs. In these instances, the reseller tops-up the prepaid account of a consumer by exchange of short messages with the operator's airtime accounting system. The reseller may have purchased from the operator bulk airtime at a discount price (on a prepaid or postpaid basis), reselling it further in smaller portions and making his profit out of the price difference.

This approach also complements the needs of cash and exchange economies, typical in most of the rapid growth markets.

7.6. Special issues for rural networks

This subchapter discusses planning models suitable for very sparsely populated, mountainous areas with very low population density, and application of the models for appropriate access equipment.

A considerable proportion of the world's population lives in rural areas, especially in Africa and Asia, as it can been seen from Fig. 7.6.1.

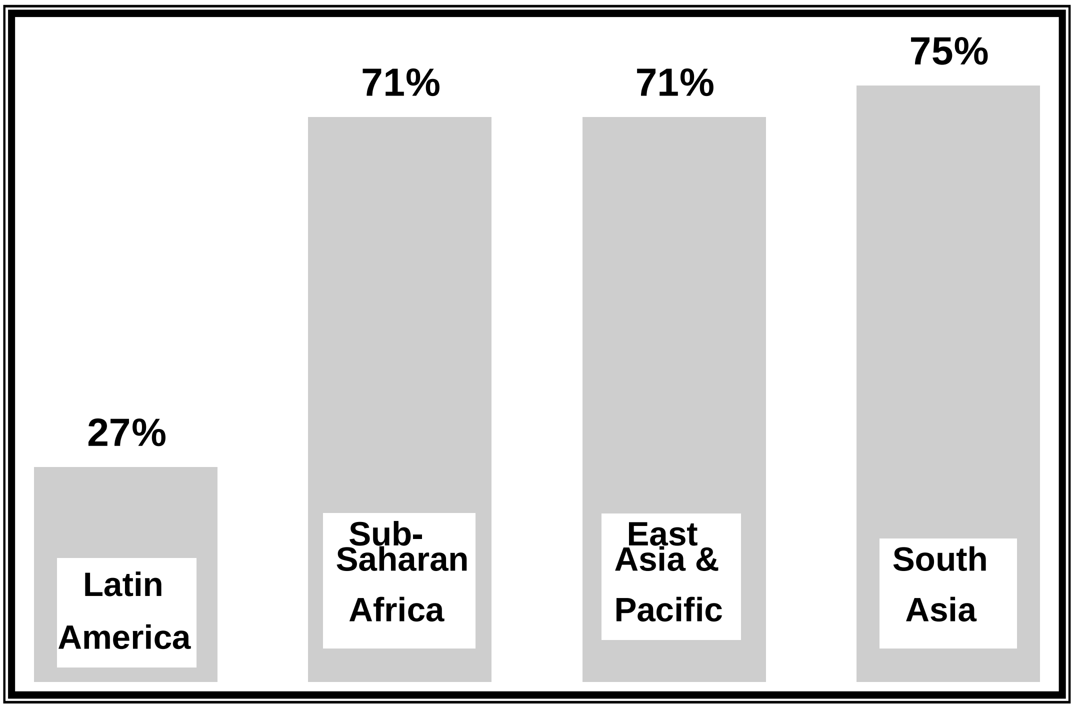


Figure 7.6.1: Percentage of the population living in rural areas. *Source:* The World Bank(1992)

There is a considerable difference between rural and urban telecom development, which depends mainly upon the country development, e.g. for Low Income and Lower Middle Income of the countries (see Fig. 7.6.2):

- Low Income: 9,3 % teledensity versus 2,1 %

- Lower Middle Income: 24,8% teledensity versus 7,3 %

One explanation could be seen through the population growth of the countries.

From the findings of the United Nations :

- all growth in population will concentrate in urban areas, no growth in rural areas
- most of the growth will concentrate in urban areas of less developed regions

i.e., attention will be primarily on the urban areas.

	Population of large cities as %	Large city teledensity [%]	Rural areas teledensity [%]	Overall teledensity [%]
Low Income	**6,0**	**9,26**	**2,15**	**2,54**
Lower Middle Income	**5,8**	**24,84**	**7,30**	**8,77**
Upper Middle Income	**16,1**	**30,77**	**21,10**	**22,94**
High Income	**10,8**	**57,49**	**54,83**	**55,21**
Africa	**12**	**6,42**	**1,39**	**1,99**
Americas	**13,6**	**34,8**	**21,72**	**11,39**
Asia	**4,8**	**25,97**	**6,94**	**7,84**
Europe	**10,9**	**48,24**	**30,19**	**31,98**
Oceania	**17,8**	**45,97**	**36,77**	**38,38**
WORLD	**7,7**	**17,4**	**25,25**	**9,20**

Figure 7.6.2: Teledensity diversity - largest cities vs. rural areas . *Source:* ITU WTID 2002

This section is devoted to modeling and optimization methods for designing rural networks. We wish to note here that this issue is much less developed (as can be seen in the literature) than analogous problems for core/backbone nation-wide and metropolitan networks, despite the fact that a great proportion of the world’s population lives in rural areas, especially in Africa and Asia. Having this in mind, we will concentarte on these aspects of rural network design/optimization which are different from wide-area large core networks covered in the previous sections of this chapter.

7.6.1. Rural networks – specific features

The most imporatnt part in rural network planning is designing of the layout of transmission facilities, both for wired (fixed) and wireless networks. In the first case, the transmission facility planning has to include the underlying cable layer, and in the second case (radio newtorks) – the underlying cellular/base station structure. In the balance of this section we will concentrate on the fixed, cable-based transmission networks appropriate for the rural areas.

With the contemporary cable (optical) transmission technologies, the link and node capacity dimensioning issue is less important in the rural network planning/design. The capacity of a single-fibre WDM system (or even a SDH/SONET system) would be in most cases more than sufficient to fulfil the requirement for bandwidth in a typical rural area. This implies that in the rural network design the most important aspect is the topological design (see Subsection 7.1.1), i.e., the question where to put ducts/cables/transmssion links, rather than how much capacity to install on transmission links built on existing ducts/cables. In fact, this leads to such problems as the minimum Hamiltonian cycle problem, the minimum spanning-tree problem, or the minimum two-connected graph problem (when network resiliance is to be provided).

7.6.2. Customers distribution in the rural areas

Population in rural areas is typically distributed over large geographical territories in sparsely populated areas, very often with difficult access and poor infrastructure, such as mountainous areas.

Figure 7.6.3: Rural area – geographical data presented with a raster map

Userts are primarily residential, in some cases there could be some kind of community shared access for using the telecom services, e.g.,telecenters.

The users' distribution in such areas is characterized by a very low density.

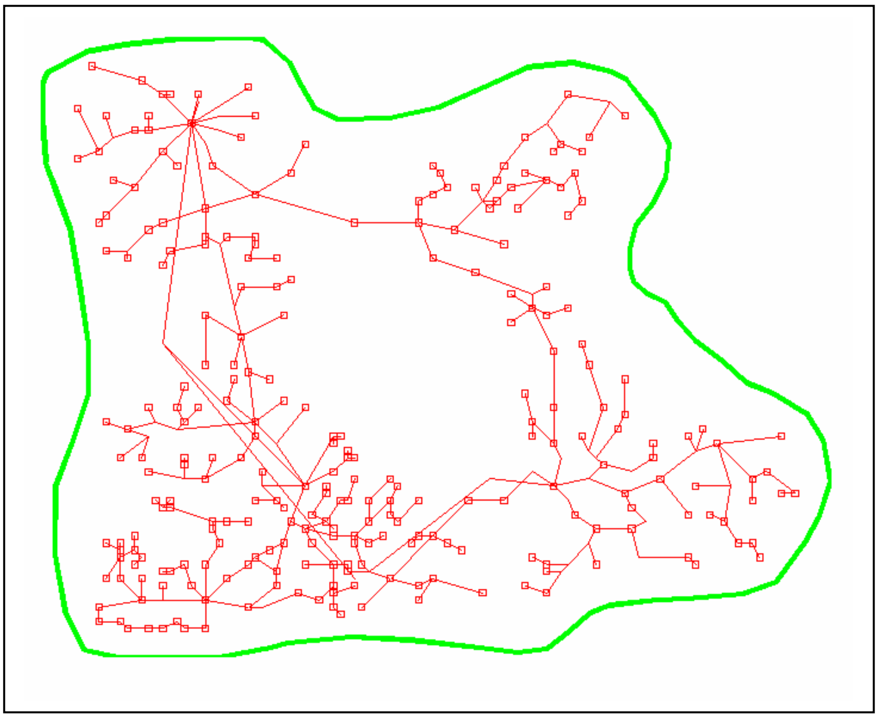


Figure 7.6.4: Rural area – modeling of the customers

Most appropriate seems modeling of the customers as concentrated in nodes (sites) – Fig. 7.6.4.

7.6.3. Services and traffic intensity in rural areas

In many rural areas the basic service is still voice, but it is expected in the future that the broadband access (giving access to Internet, E-services, etc.) will start penetrating such areas.

Special case is the community shared access, as the telecenters, where broadband connection will be very important.

7.6.4. Telecommunication technologies for rural networks

7.6.5. Structure of rural networks

In the case of rural areas with sparse demand point distriburtion and low overall traffic/demand requirement, a typical transmission network can be represented with a one-level graph with no hierarchical structure typical for nation-wide or large metropolitan-area networks. This certainly makes the modeling issue relatively simple and leads to such typical network structures as rings, trees (stars), or sparse mesh graphs (see Figure 7.6.5).

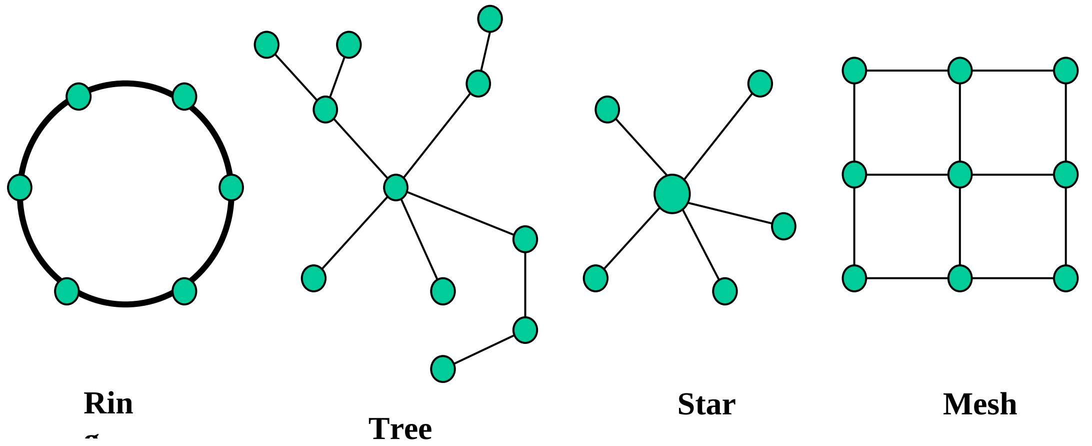


Fig. 7.6.5. Rural network

When a rural area is too large to be covered by a single ring, a multi-ring structure can be used. Such a structure is depicted in Figure 7.6.6. It should be noted that rings are mostly installed to assure resilience (see Paragraph 7.6.6.4), in particular robustness against any single link (link on a ring is called a segment) or any single node failure. If such a robustness is to be maintained in a multi-ring structure, the rings must be connected by at least two nodes. A suitable multi-ring network structure for a rural area is depicted in Figure 7.6.6, with one master ring and a number of suboridnate rings.

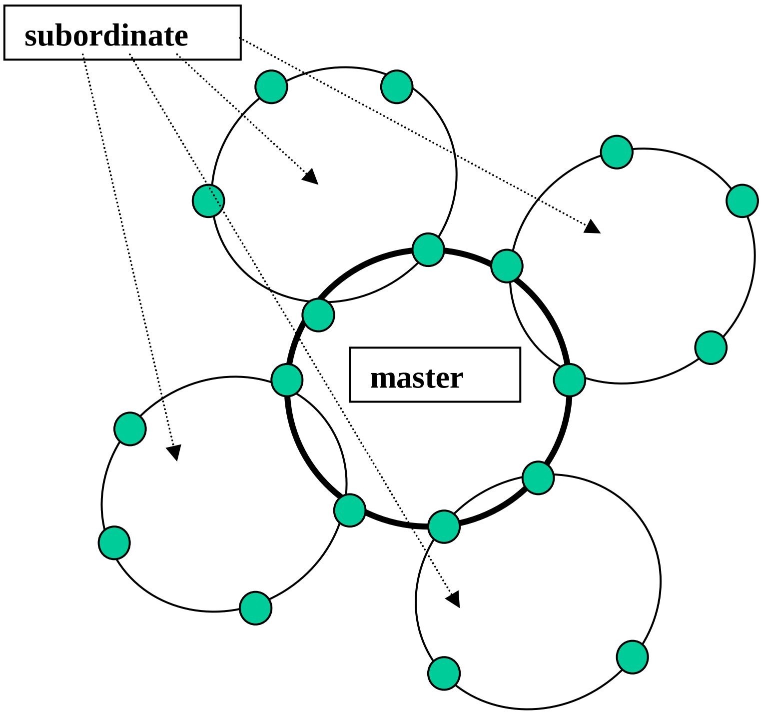


Fig. 7.6.6. A hierarchical multi-ring

7.6.6. Optimization models for fixed rural networks

7.6.6.1. Ring networks

Rings (see Figure 7.6.5) are typical network configurations constructed using the SDH/SONET transmsission systems and add-drop mutiplexers (we note that currently WDM rings are considered). A ring is very well suited for serving rural areas with few nodes and limited bandwidth requirement. A typical ring can support up to 16 nodes and has capacity up to 2.5 Gb/s (these figures determine the range of the use of such single ring network configurations).

Rings can be configured in several ways, e.g., as unidirectional (one-fiber) or bidirectional (two-fiber) rings, with (then the number oof fibers is doubled) or without the self-healing capability. In the unidrectional case all the demands betwen the node pairs on the ring use the VC containers throught the whole ring so the capacity of the ring (e.g., 2.5 Gb/s) must be greater or equal to the sum of the required bandwidth between all node pairs on the ring. In the bidiractional case this is not that simple and the demands between the node pairs are to be configure, leading to the following (non-trivial problem).

BRDP (Bidirectional Ring Dimensioning Problem)

indices

$v,w,s,t = 1,2,...,V$ nodes (V - number of nodes on the ring)

constants

h_{vw} bandwidth (capacity expressed in modules, e.g., in VC-12 = 2,048 kb/s) to be realized on the ring between node v and w ($v < w$) in both directions (form v to w and from w to v); this bandwidth must be realized between v and w on the clockwise fiber, and between w and v on the counter-clockwise fiber

C capacity (in the same modules as above for node pairs) of the ring

variables

x_{vw} number of modules allocated between node v and node w on the clockwise fiber of the ring

y_{vw} number of modules allocated between node w and node v on the counter-clockwise fiber of the ring

constraints

$$x_{vw} + y_{vw} = h_{vw} \qquad v,w=1,2,...,V,\ \ v < w \tag{7.6.1a}$$

$$\Sigma_{(v,w)}\, x_{vw} + \Sigma_{(s,t)}\, y_{st} \le C, \tag{7.6.1b}$$

where one constraint of the form (7.6.1b) is posed for each segment u (u=1,2,...,V) of the ring (segment u is the linkon the ring that joins node number u and node number u+1 (*mod V*), see Figure 7.6.7), and for each such segment u (u=1,2,...,V) the summations are carried out for:

- all pairs (v,w) such that $v < w$, and the clockwise fiber of the ring form v to w contains the segment u

all pairs (s,t) such that $s < t$, and the counter-clockwise fiber of the ring form t to s contains the segment u.

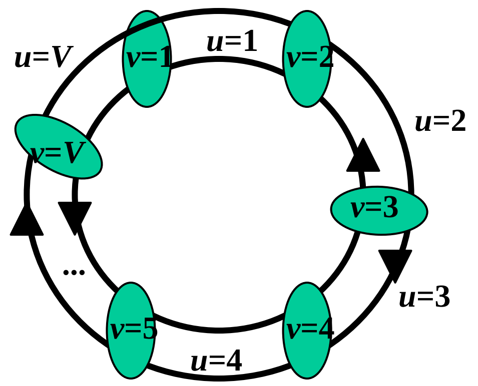


Fig. 7.6.7. Bidirectional ring

In general it is not a trivial task to resolve BRDP (which is a integer programming problem), especially when the whole capacity C has to be used in the feasible solution.

Anyhow, the main design problem for a rural ring is to find its location so as to minimize the cost of links/ducts. The problem is in fact equivalent to finding the minimal Hamiltonian cycle (this problem is frequently called the travelling salesman problem – TSP). Recall that a Hamiltonian cycle in a given graph is cycle that traverses all the nodes such that each node is traversed exactly once. TSP is as follows. Let v=1,2,...,V be a set of nodes with given locations and let g_{vw} be the cost (weight) of installing a direct optical cable (link) between nodes v and w. We are to find a Hamilitonian cycle with the minimum sum of the weights of the links it uses. In general, TSP is a NP-hard problem. Still it is quite easily solvable when the number of nodes is reasonable (see Paragraph 7.6.7.1).

7.6.6.2. Tree/star networks

When resilience is not an big issue then a tree network can be a good solution for a rural area. Recall that a tree spanning a set of nodes v=1,2,...,V is a connected graph with exactly V-1 links (see Figure 7.6.5). Certainly, if any link of a tree is deleted, then the resulting grpah becomes not connected. When we decide to install a rural network of the tree structure it is highly desirable to use the minimum (synonims: lightest, shortest) spanning tree, i.e. the tree that has the minimum weight where the weight of the tree is the sum of the weights of all its links g_{vw}. Note that a ring with one link removed (any link, in fact) is a tree, but obviously not necessarily the lightest tree.

Finding a shortest tree is an easy task when all nodes have to be in the tree. Otherwise, when only a predefined subset of V' nodes has to be connected by a tree (but possibly using nodes outside this subset) then we encounter a NP-hard problem called the Steiner tree problem (see Paragraph 7.6.7.2).

A star network (see Figure 7.6.5) is a special case of a tree with a central node (called a root) with all other nodes connected to the central node by means of a direct link. The simple structure of a star network has a lot of advantages from the management viewpoint, and hence it is preferable, provided its cost can be made close to that of the minimum spanning tree.

7.6.6.3. Mesh networks

In rural areas mesh networks are provided instead of rings and trees for two main reasons: (i) when the area is large (many nodes, high traffic demand), and (ii) resilience. In a large rural area a single ring network structure may not be sufficient (because of the excessive number of nodes in the area, or for insufficient capacity of the ring compared to the traffic demand generated in the area). In such a case a multi-ring structure or a pure mesh transmission networks has to be installed instead. Also, when resilience is an issue, then a tree network is not acceptable and the tree structure must be converted into a mesh structure.

Optimization models for mesh networks have been decribed in detail in Section 7.1.

7.6.6.4. Resilience issues

Both unidirectional and biderectional rings can be made robust to single segment/node failures by doubling the number of fibers they use. This leads to USHR (unidirectional self-healing ring) and BSHR (bidirectional self-healing ring), built upon two and four fibers, respectively. Although the restoration pricinples are different for USHR and BSHR, both types of rings restore 100% demands in the case of a single segment failure and in the case of a single node failure, using extra fibers.

When for some reasons the single ring structure cannot be used in a rural area, then either a multi-ring structure or a mesh network should be installed. A multi-ring structure can consist of a master ring with subordinate rings attached to it (see Figure 7.6.6). Each subordinate ring is connected to the master ring in two different nodes so that no single node failure can disconnect the subordinate ring from the rest of the network.

Let us now consider the case of meshed networks. As we have mentioned above, for rural mesh networks the aspect of link capacity dimensioning is less important than that of finding link locations. Hence, the most important aspect is to find the subset of links (and corresponding routing) to be installed, such that the resulting network is cheapest and at least two-connected. An example of a representative rsilient design problem is formulated below. It consists of finding a pair of paths (basic path, backup path) for each traffic demand such that at least one of these two paths is working in each failure stuation.

RDPRN (Robust Dimensioning Problem for Rural Networks)

indices

$d=1,2,...,D$	demands
$j=1,2,...,m(d)$	pair (P_{dj},Q_{dj}) of situation disjoint paths for flows realizing demand d, nominal path P_{dj}, and backup path Q_{dj}
$e=1,2,...,E$	links
$s=0,1,...,S$	situations

constants

a_{edj}	= 1 if link e belongs to nominal path P_{dj}; 0, otherwise
b_{edj}	= 1 if link e belongs to backup path Q_{dj}; 0, otherwise
κ_e	installation cost of link e
α_{es}	binary availability coefficient of link e in situation s ($\alpha_{es} \in \{0,1\}$)
δ_{djs}	binary availability coefficient of path P_{dj} in situation s, $\delta_{djs} = \Pi_{e:\, aedj=1}\, \alpha_{es}$

variables

u_{dj}	binary routing variable indicating whether pair j of demand d is used (u_{dj}=1) or not (u_{dj}=0)
ε_e	binary variable indicating whether link e is installed (ε_e =1) or not (ε_e =o)

objective

minimize $\boldsymbol{C}(\boldsymbol{\varepsilon}) = \Sigma_e\ \kappa_e\varepsilon_e$ (7.6.2a)

constraints

$\Sigma_j\ u_{dj} = 1$ $d=1,2,...,D$ (7.6.2b)

$\Sigma_d\ \Sigma_j\ (a_{edj}\delta_{djs} + b_{edj}(1 - \delta_{djs}))u_{dj0} \leq D\alpha_{es}\varepsilon_e$ $e=1,2,...,E$ $s=0,1,...,S.$ (7.6.2c)

Note that when the above problem is solved then we can use the pair (basic path, backup path) either within a hot-standby mechanism or a path-restoration mechanism (see Paragraphs 7.1.2.4 and 7.1.2.6, respectively).

7.6.7. Optimization methods for fixed rural networks

In this subsection we will describe the basic optimization methods required for designing ring, tree, and mesh rural networks.

7.6.7.1. Ring network optimization

The basic problem for designing a single-ring rural network is finding the route for the ring connecting the set of nodes with given locations. This, as metioned before, is equivalent to finding the minimum Hamiltonian cycle (circuit) in a given (fully connected) graph. Hence, we assume that for each pair of nodes v and w ($v,w,=1,2,...,V$, $v < w$) the cost (wieght) of lying down the cable is equal to gvw, and we look for the least expensive Hamiltonian cycle, that a permutation $\varphi: \{1,2,...,V\} \rightarrow \{1,2,...,V\}$ of nodes' indices such that the cost

$$g(\varphi) = \Sigma_{v=1,2,...,V}\ g_{\varphi(v)\varphi(w)} \qquad (7.6.3)$$

is minimal among all such permutations.

It is well known that the above problem (recall that this problem is called TSP – travelling salesman problem) is NP-hard. Consequently, we do not know of any algorithm that solves TSP effectively in a general case (i.e., in time polynomial with the number of nodes). Nevertheless, TSP can be solved effectively even for very large instances, of the order of ten thousand of nodes [7.6.1]. The basic technique used in such practical algorithms is called branch-and-cut [7.6.2], and is an extension of the branch-and-bound method described in Subsection 7.3.2, enhanced with stochastic heuristics such as simulated annealing (SA, see Subsection 7.3.3).

The branch-and-cut technique uses an IP fromulation of TSP and generates additional inequalities (so called valid inequalities) at the consecutive nodes of the branch-and-bound process. These inequalities improve the lower bounds of the process and considerably speed up the time of reaching an optimal solution. Generation of valid inequalities is based on particular properties of TSP.

The use of simulated annealing for TSP can be based on the following definition of the neighbourhood of a cycle: the neighbourhood of a Hamiltonian cycle φ^1 is the set $N(\varphi^1)$ of all cycles φ^2 that can be obtained form by interchanging links between any four nodes, as illustrated in Figure 7.6.8. Using the SA algorithm for a graph corresponding to a rural area would usually lead to good sub-optimal solutions.

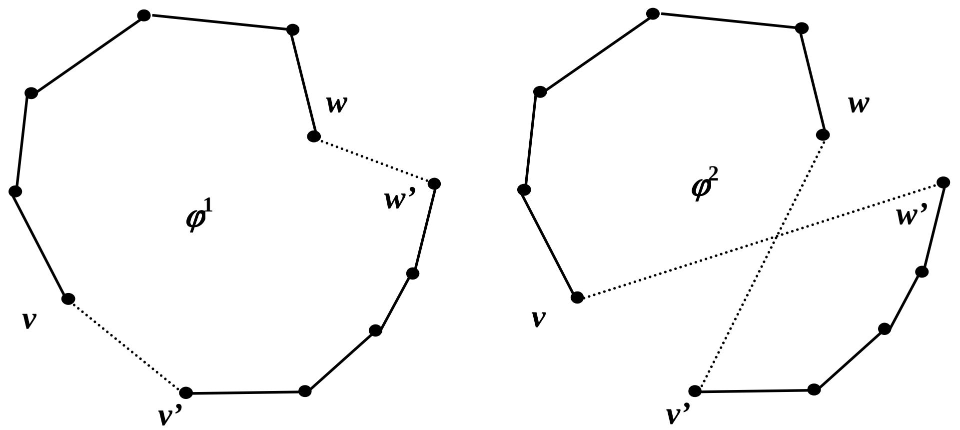


Fig.7.6.8. Neighbouring Hamiltonian cycles

7.6.7.2. Tree network optimization

The basic problem in design of tree structures for rural networks is the problem of finding the minimum spanning tree (MSTP – minimum spanning tree problem). MSTP can be resolved effectively when all nodes of the graph have to be spnned by the tree. Then, for example, the Kruskal algorithm can be used (see [7.6.3]). To apply the algorithm, we first order the links of the graph in the non-decreasing order by their length g_e.

Step 1: Select a shortest link and color it blue. Set $k = 1$.

Step 2: Among the links not yet considered, select a shortest a shortest link $\{v,w\}$. If nodes v and w belong to the same subtree, color link $\{v,w\}$ red; otherwise, color it blue and set $k = k+1$.

Step 3: If $k = V$-1, stop. Otherwise, return to Step 2.

Hence, intially, each node forms a subtree of its own, and then the consecutive subtrees are merged until the final minimum spanning tree is formed. Note that detecting whether a link is in the same subtree or joins two different subtrees is easy and can be done by mainting the subtrees as ordered lists.

7.6.7.3. Mesh network optimization

Optimization methods for mesh networks have been decribed in detail in Section 7.3.

Chapter 8 – Data gathering

Network planning, especially performed with NP tools, requires collection of numerous data.

Main input data are:

8.1. Geographical information for the studied area

Geographical Information Systems (GIS) are nowadays one of the most import tools for network planning, maintenance and monitoring. GIS applications try to build a model of the world inside a computer, allowing displaying and analyzing complex geographical correlations and details. The computer model of the world, which is the basis for all analyses, consists of geographical information.
Geographical information is the combination of geometrical, locatable details and collected information. Their coordinates as geometrical detail and their length or name as collected information, e.g. describe roads.

8.1.1. Vector and Raster data

There exist various types of geographical information mainly stored in 2 formats: raster or vector. Both formats have their justification inside the GIS world. Each format has its advantages and is intended for specific tasks. This chapter will explain both formats for the storage of geographical information and give an overview of the use and advantages of each format.

8.1.1.1. Vector data

Vector data has its name from the method by which the geometry is collected, starting with a point a vector with a certain length and direction points to the second location where another vector directs to a next location and so on.
Geometrical objects of a vector can be:

- Points
- Lines: two points connected by a line
- Poly lines: several connected points in a row, e.g. roads
- Polygons: several connected points in a row, the last and first points are connected. Polygons are closed poly lines and are used to surround areas, e.g. country borders


Figure 8.1: Real World

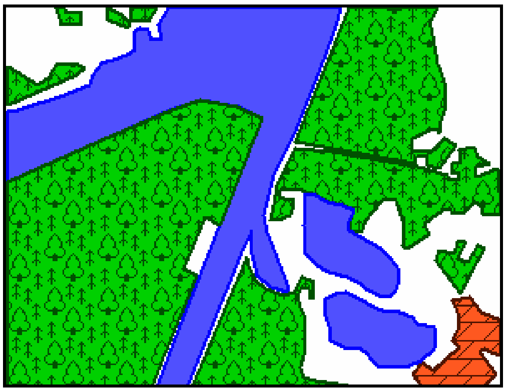


Figure 8.2: Vector model or the Real World

The geometry indicates points or follows lines that go along or surround objects of one and the same kind. Lines of the same type can be roads, rivers or microwave links. Closed polygons can surround areas like political districts, city contours, coverage areas, etc. Besides the geometry of vectors, collected information about these objects can be attached. Each vector object can have a lot of different information stored. The collected information for a country border could be for example: the country name, country size, number of inhabitants, inhabitants divided by male / female, income, etc.

8.1.1.2. Raster data

The second format in which geographical information is stored is the raster or grid format. The area of interest is divided into a regular grid of equally spaced small rectangular regions. One region of the raster computer model is called a pixel. Each pixel stores one (1) information about the surrounded area. This can be the mean height above sea level, the number of inhabitants living there, the field strength, etc. If more than one (1) information of the same type is available in a pixel area, a decision has to be made, which information will represent this area. A sub-classification of pixels is not possible because each pixel can indicate only one (1) information about its area content. In case the stored information should be more detailed, the pixel resolution, the area size of one (1) pixel, has to be reduced.


Figure 8.3: Real World


Figure 8.4: Raster model of the Real World

A collection of pixels containing the same information type is called a layer. Therefore one (1) layer can only store one type of information. If more than one information about a pixel object is required, several layers with different contents have to be created.

8.1.1.3. Comparison

Vector and raster format have different advantages and disadvantages. The table below gives an overview of the most significant differences.

	Vector format	**Raster format**
Available information	Several information types available per vector object	1 information value per raster pixel 1 information type per raster layer
Memory size	Low Memory size dependent on complexity of terrain	High Fixed memory size for terrain data area 1 full layer for each information type
Zooming Range	Flexible zooming of geometrical information	Zooming in multiples of pixel resolution
Zooming in > 100%	Possible, contents limited by accuracy of geometry	Zooming in leads to enlarging of pixel size on screen
Data access	Slow, valid information for specific point must be detected out of all vector objects	Fast, valid information at specific point detected from nearest pixel

Table 8.1: Comparison of geographical information formats

The table above points out that vector data has the great advantage of low memory consumption and the multiple availability of information regarding geometrical objects compared to raster data. Therefore it is predetermined for displaying purposes and the analysis of complex correlations of geographically large objects. If fast access to information at a certain point is needed or the terrain is very complex in terms of changes, the raster format is preferable. This is the reason why GIS applications which perform field strength calculations use raster format for the storage of most of the terrain data layer, like elevation-, land-use-, population-data or calculation results. Terrain data in vector format are usually used by these application for displaying political maps, roadmaps or networks, or the specification of areas for computer analyses.

8.1.2. Background Maps for Display and Visualization

Radio network planning with GIS applications allows the easy geo-correct overlay of sites, links or planning results over maps. Therefore overview and detailed maps are required.

8.1.2.1. Overview maps

Overview maps are needed to display countrywide or large area coverage or even multinational networks for international coordination. The maps can be in vector as well in raster format. The maps in raster format are most of the time scanned paper maps, which are

post processed in order to display the scanned information at the geographically correct places. The scanned paper maps have the advantage that they have a higher recognition factor than vector maps. Vector maps do usually not show so many details and their representation strongly depends on the visualization functionality of the CIS tool.
Overview maps should show only the necessary details, like political borders, main cities, rivers, lakes, and roads that are needed to identify known locations at the first view. Map scales from 1:500,000 to 1:4,000,000 are suggested depending on the area of interest.

8.1.2.2. Medium scale maps

Medium scale maps at a scale between 1:50,000 and 1:250,000 are needed for more detailed network investigations in rural areas or the presentation of locally restricted networks. Like the overview maps they can be in vectors or in raster format. Again scanned raster maps have the advantage of a higher recognition factor, which makes it easier to present planning results to spectators that are not familiar with vector based maps. On the other hand vector based maps are often more up-to-date than raster maps which have an update circle of 5 or more years. Vector based maps are often generated out of most recent data collected for route planning software.

8.1.2.3. Detailed maps

Detailed maps of a scale of 1:30,000 and lower are used for detailed network-, micro-cell- and PMP planning in rural and urban areas. They can be in raster and vector format. Scanned and geo-corrected city maps are a very good and cheap source for detailed maps of the most important cities around the world. For a lot of countries paper maps at a scale of 1:10,000 to 1:25,000 are available. Unfortunately, they are often more than 10 years old or restricted by military. Therefore the usage of these maps is very often limited.
A second very useful source for detailed maps is a combination of most recent orthorectified images and precise vector data. The images are created from satellite imagery or aerial photography. The orthorectification changes the point of view from one (1) point camera lens perspective towards parallel top view perspective. This mechanism allows looking at every point of the image right down onto the top of the building instead of viewing it from the side. The vectors overlaying the images mark well known points of interest, highlight rivers and important streets or display names of cities, streets and house numbers. The available details in the images depend on the source of the maps. Aerial photography usually has a pixel resolution of 0.1 m to 2 m. A pixel resolution of 0.64 m to 5m for optical satellites is available on the market for nearly every area in the world.

Figure 8.5: IRS 1C/D satellite image

8.1.3. Elevation data

Elevation data can be divided into 3 categories:

- DTM (Digital Terrain Model)
- DEM (Digital Elevation Model)
- DBM (Digital Building Model)

Elevation data is most of the time stored in raster format. Only DBM data is often stored in vector format. The advantage of raster format for elevation data is that a preprocessed height value is easily detectable for every point of planning area. The exact height can be detected by the next neighboring pixel or bilinear interpolated by the four neighboring pixel. Elevation data in vector format would require a detection of all neighboring elevation contours needed to interpolate the exact height at that position.
The most important and most used elevation type is the DTM. There is no standardization for these names. A digital Terrain Model is often called DEM or simply Elevation Model. It represents the height of the surface of the earth above sea level.

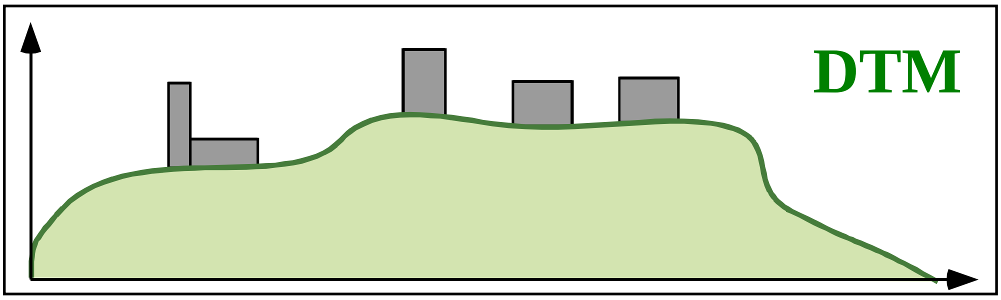


Figure 8.6: Digital Terrain Model – Surface height above sea level

The DTM in combination with land-use data is the most used source for field-strength calculations. For GSM and microwave link planning a pixel resolution of 50 m for rural areas and of 20m – 25m for urban areas is suggested.

Digital Elevation Model (DEM) is often called Digital Surface Model (DSM) or Digital Building Model (DBM). A DEM is most of the time stored in raster format. Each pixel of the DEM represents the height of the surface of the earth above sea level plus the rooftop height, where buildings are.

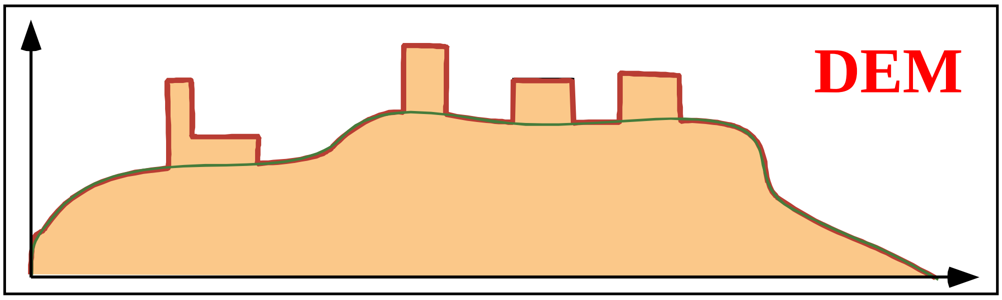


Figure 8.7: Digital Elevation Model – Surface height above sea level plus rooftop height

The recommended pixel resolution for DEM data is between 1 m and 5 m. A pixel resolution of 5m implies that a building represented by 1 pixel has a minimum size of 5 m x 5 m. Smaller buildings will not be detected and therefore not included into the DEM.

Digital Building Model (DBM) is available in raster as well as in vector format. The height given represents the rooftop height above sea level. All areas, which are not covered by buildings, are marked with a certain default value, which is most of the time –9999 or –999.

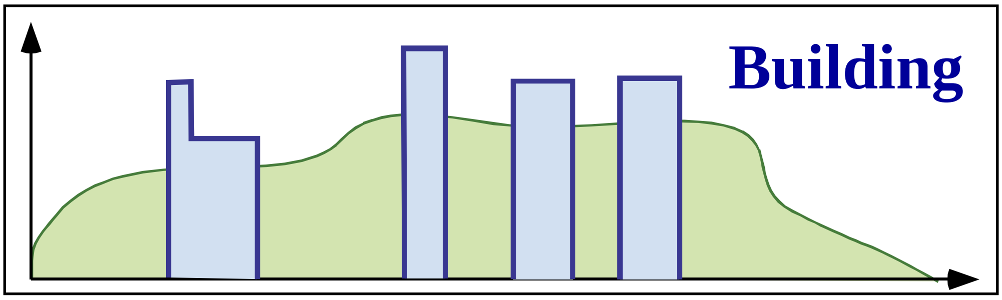


Figure 8.8: Digital Building Model – Rooftop height above sea

The recommended pixel resolution for DBM data is between 1 m and 5 m. A pixel resolution of 5m implies that a building represented by 1 pixel has a minimum size of 5 m x 5 m. Smaller buildings will not be detected and therefore not included into the DBM. Some new 3D propagation model for micro- and pico-cell planning are based on a DBM in vector format. Besides the rooftop height above sea level other information about that building is included. This additional information can be: rooftop height above ground, wall material parameter, roof type, etc. The geometry and file format for this DBM in vector format strongly depends on the software used and can be very complex, since there exists no industrial standard for 3D-building data yet.

	DTM	DEM	DBM
Given height	Ground height above sea level	Ground height or rooftop height above sea level	Roof-top height above sea level, else default value
Resolution for rural areas	50 m		
Resolution for urban areas	20m – 25 m	1m – 5 m	1m – 5 m
Sources	Topographical maps	Aerial photography Stereo high resolution satellite imagery	Aerial photography Stereo high resolution satellite imagery Aerial photography
Usage	Macro- cell and microwave link planning	Micro- and pico-cell planning, PMP	Micro- and pico-cell planning, PMP
Price indication	0.05 – 2 Euro / km^2	50 – 300 Euro / km^2	50 – 300 Euro / km^2

Table 8.2: Comparison of elevation data types

8.1.4. Clutter / Land-use data

Besides the DTM some propagation models take the land coverage into consideration for the calculation of the field-strength. Depending on the type how the earth is covered at a place, these models calculate complex correction coefficients to adopt the free-space field strength on measurements. There exist different synonyms for this kind of information:

- Land-use data
- Land-coverage data
- Clutter data
- Morpho data

Clutter data stores the information on how the earth is covered. The supported clutter classes are not standardized and therefore strongly depend on the software used and the propagation model supported. The most important clutter classes are listed below.

Clutter class	Classification
Sea	Tidal water
Inland water	Lakes, reservoirs, rivers, streams
Open wet areas	Areas liable to flooding usually situated beside rivers
Open	Agricultural land with sparse dwelling; uncultivated land
Forest	Woodlands; forest
Open in Urban	Open areas located within urban areas
Industrial areas	Factories, warehouses, surface quarries, docklands
Villages	Urban areas located approximately > 2 km from the urban fringe unless clearly separated by a feature such as a major river. Villages can be scattered throughout the image
Suburban Residential	Areas of housing mixed with wooded areas and streets. Often situated on the outskirts of urban areas
Urban	Medium density buildings or housing often mixed with areas of open.
Dense Urban	Dense buildings with narrow side streets, mostly found in city center areas.
Parks Recreational areas	Urban parks, sports stadiums, golf courses
Block Buildings	Groups of generally narrow buildings, parallel and separated by open space

Table 8.3: Most important clutter classes

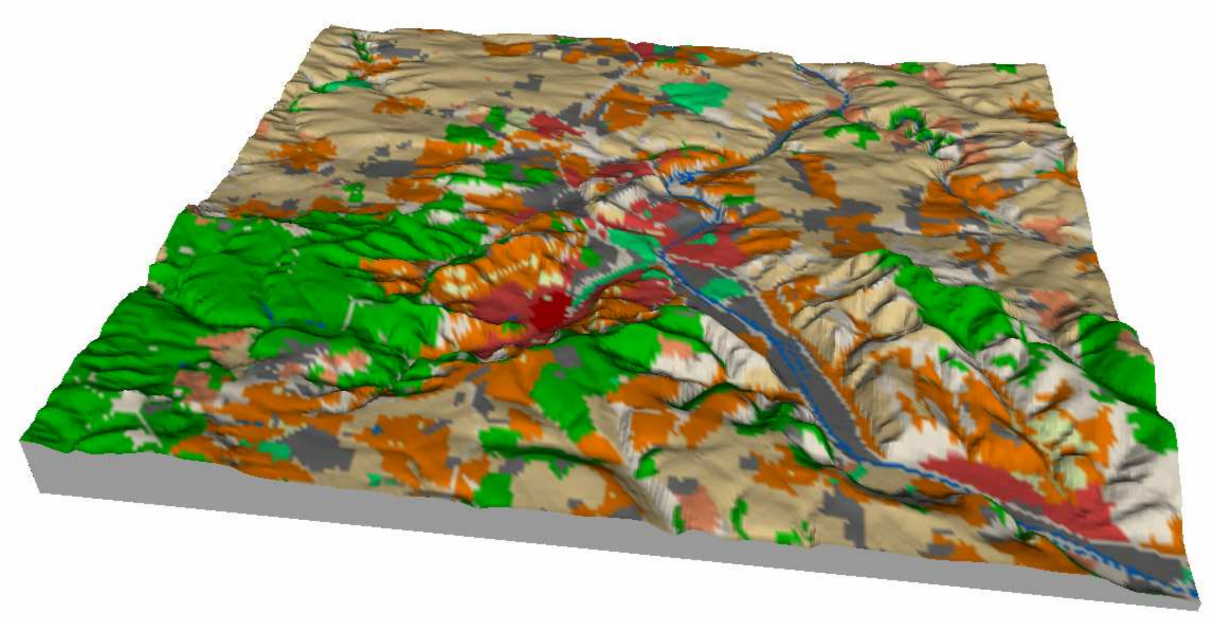


Figure 8.9: 3D-model of terrain overlaid with clutter data

Clutter data for network planning is most of the time stored in raster format. Till the year 2000 most of the clutter has been extracted from paper maps of various scale. The still most used paper maps are Russian topographic maps, which are available for approximately 80 - 90% of the world. The big disadvantage of the Russian maps is the date of creation, which is

between 1980 and 1995 for European and areas of the former Soviet Union. For all other countries the maps are even older. The use of paper maps from local sources like ordnance surveys is often restricted by military or only allowed after paying certain license fees. More recent clutter data can be generated out of satellite imagery. Clutter data with 25m resolution per pixel or less is often generated out of Spot or IRS 1C/D satellite imagery. Specially trained people according to visually dividable classification rules carry out the classification manually. The nowadays most used source is Landsat7 multi-spectral satellite imagery, which allows a semiautomatic classification. Typical resolutions for clutter data created out of Landsat7 multi-spectral satellite imagery are 25m, 50, or lower. Comparing to Spot or IRS 1C/D clutter data from Landsat7 is much cheaper because of 3 reasons:
the semiautomatic process for creation of the clutter data
the satellite scenes are bigger
there are less license fees to pay.

8.1.5. Demographic data

Statistical information about the population distribution or income situation inside a specific area is called demographic data. This statistical information can often be obtained from the statistical ministries. Most of the time they are in simple table format. For the use in GIS applications beside the statistical information the geographical geometries of the related areas are needed. Demographic data can be stored in vector and in raster format. In case of raster format, several layers have to be created, one for each type of statistical information, e.g. one layer for total population, one for households, one for income per family, etc.

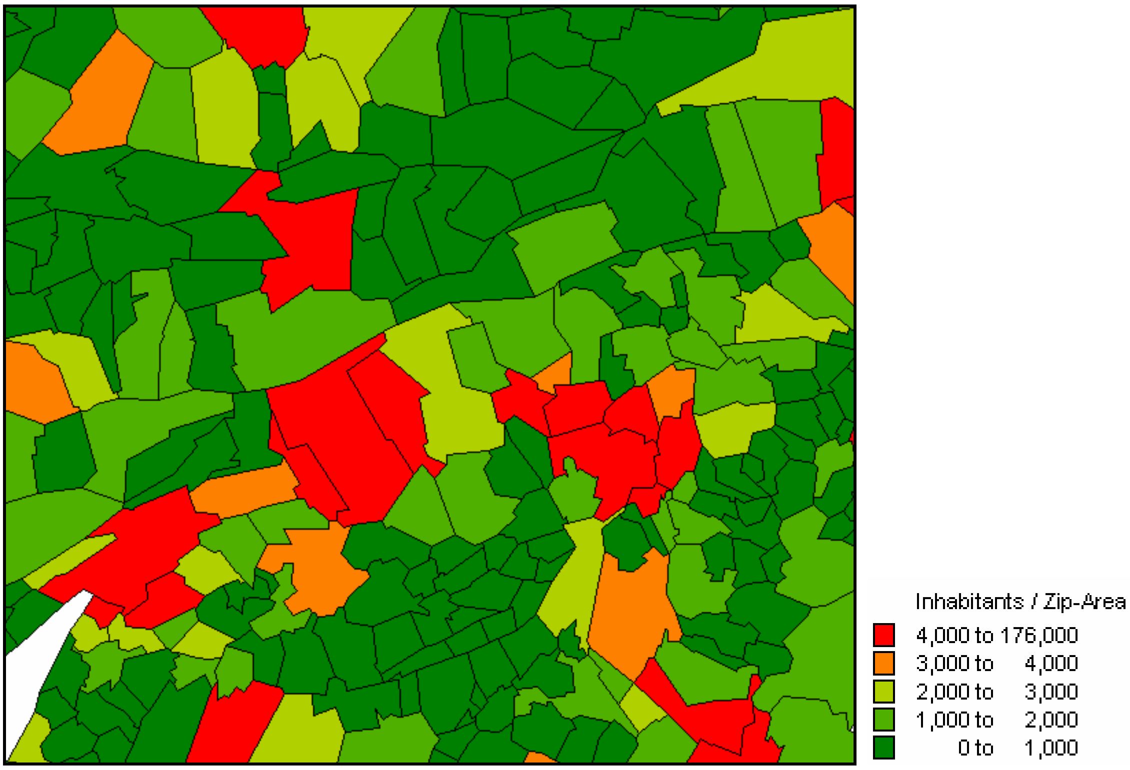


Figure 8.10: Total population based on zip-code vectors

Attention: This is not an ITU publication made available to the public, but **an internal ITU Document** intended only for use by the Member States of the ITU and by its Sector Members and their respective staff and collaborators in their ITU related work. It shall not be made available to, and used by, any other persons or entities without the prior written consent of the ITU.

There is demographic data available for each country of the world based on the information about the entire country, or about their main provinces. The degree of accuracy of demographic maps depends on the size of the geometrical areas, where statistical information is available. The smaller the areas, the more accurate the statistical information. For a lot of countries demographic data based on county-, zip code area, or electoral district is available.
If the geometrical areas are too large, the demographic information can be merged with existing clutter data in order to distribute the information only in those areas, which are indicated as populated by the clutter data. Some software packages allow to use additional weighting coefficients for a more realistic distribution of the statistical information.

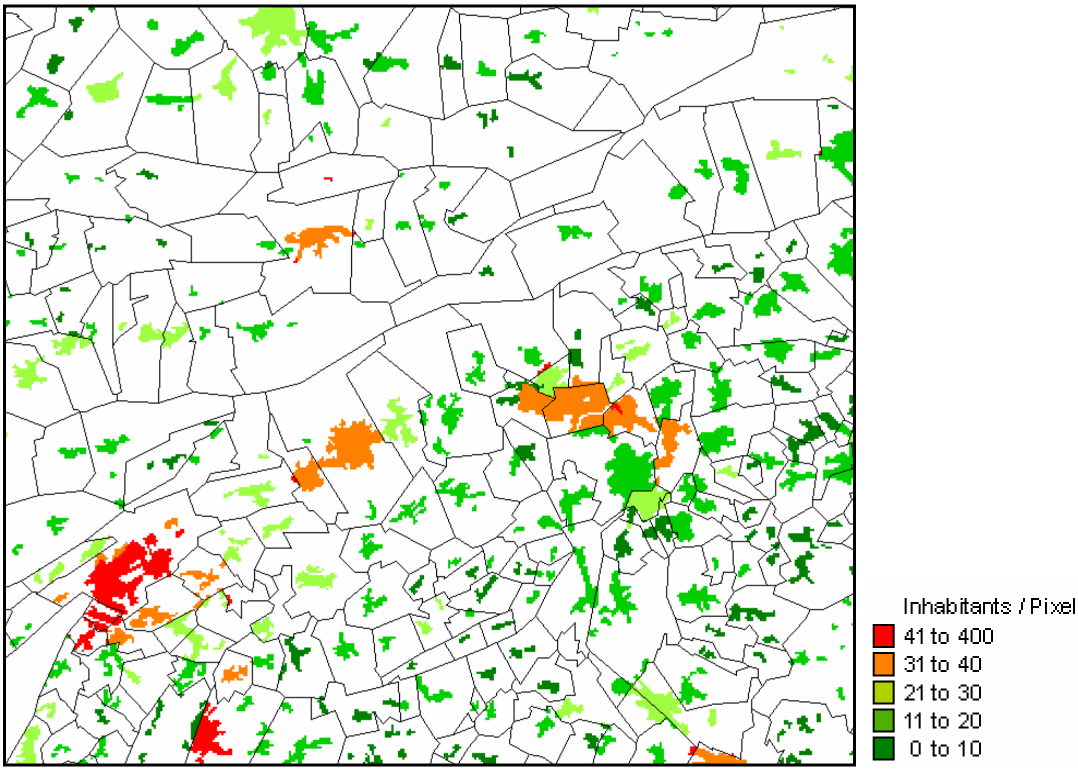


Figure 8.11: Total population based on zip-code vectors merged with populated areas from clutter layer

8.1.6. Locations of populated places / buildings, Service areas, Clutter data

8.1.7. Digitizing of maps

8.2. Demand of services in relative penetration per customer category

8.2.1. Offered services

8.2.2. Customer categories

8.2.2. Demand per site and per area

8.2.4. Demand per time point

8.3. Demand of traffic, usually expressed as traffic matrices

8.3.1. Traffic per service per customer class

8.3.2. Traffic matrices per service

8.3.3. Traffic per time point

8.4. Information for the existing network and infrastructure

8.4.1. Exchanges, routers, concentrators, etc.

8.4.2. Transmission equipment, cables, etc.

8.4.3. Buildings, duct system, etc.

8.5. Telecommunication equipment characteristics and capabilities

8.5.1. Max capacities for utilization

8.5.2. Technical characteristics, e.g. cable attenuation/km

8.6. QOS requirements

8.6.1. For the system - e.g. congestion criteria

8.6.2. For the technology – e.g. permitted attenuation

8.7. Telecommunication equipment fixed and variable costs

8.7.1. Equipment structuring and modeling

8.7.2. Cost models - linear, step functions, etc.

8.8. Economical and Operational data

8.8.1. Purchasing, installation costs

8.8.2. Operational and maintenance expenses

Annex 1 – Network planning tools

Network planning tools are essential for network planners to be able to design multi-service networks that can meet today's growth in user applications and traffic. These networks also have to provide guaranteed service quality, availability, reliability, ensure minimal delay, and must be optimized according to various cost constraints.

Since the complexity is also growing in a very fast way, the planning tools must be powerful and flexible, to handle all the different network designing issues. The possibility of extension in a very easy way is also required, since the user's work must be only related to developing new algorithms and applications, instead of programming the integration of his methods into the main framework.

A portfolio selection of planning tools to support those planning activities is provided. The selection criteria are: capability to model modern technologies, commercial availability and being well proven in the field.

Used structure to describe the tools:

Tool xxx :

- Objective
- Domain of applicability: planning activity, network layer(s), technologies, etc.
- Capabilities: network type, size, routing type, etc.
- Required inputs
- Provided results
- Example cases

LIST OF NETWORK PLANNING TOOLS:

A1.1. Application of EXCEL

Microsoft (MS) Excel, as it is well known, is a powerful spreadsheet that is easy to use and allows you to store, manipulate, analyze, and visualize data.

It is a standard part of all versions of Microsoft Office, and is currently in its tenth version, known as Excel 2003.

At its most basic level, Excel simply provides a structured way to store data.
An Excel workbook is made up of worksheets, each of which contains an array of 256 columns and 65536 rows. That's a total of over 16.7 million cells for data.
In a workbook there are maximum 255 worksheets.

Excel has efficient ways for managing the stored data.
Excel allows to sort data by as many columns as one may need, provides shortcuts to allow to move instantly to any cell, or to find any piece of data in any cell and methods of filtering stored data.

The real power of Excel is in analyzing rather than simply storing data.

As with any spreadsheet programme, Excel can handle mathematical calculations of any complexity.

It also has many Functions built in for specific tasks.
The most common of these, as SUM, MAX, MIN, AVERAGE can be calculated with a click of a button.

There are also over a hundred functions, grouped by category, invoked from dialog window.

Each function can be manually typed, or if preferred, Excel can guide the user through the process, prompting for the required information and giving extra help where needed (see figure A1.1).

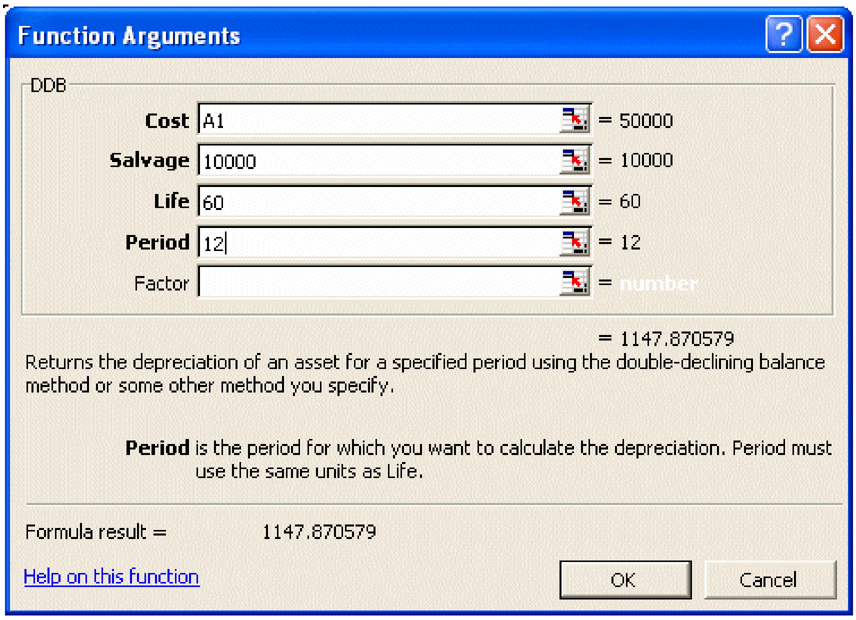


Figure A1.1 – Dialog window of Excel function calculating the depreciation of an asset

There are over 80 functions available within the Statistical category – mean, modal and median averages, standard deviations, ranking functions, etc.

In addition to analyzing data, Excel provides a number of very powerful tools for summarizing purposes.
Excel puts subtotal information below each category specified, and can further break down these categories, or add averages, maximums and so on as necessary.

Another major area in Excel application is the creation of charts.
Bar charts, pie charts, line charts, column charts, radar charts, surface charts, area charts, etc. (see figure A1.2).

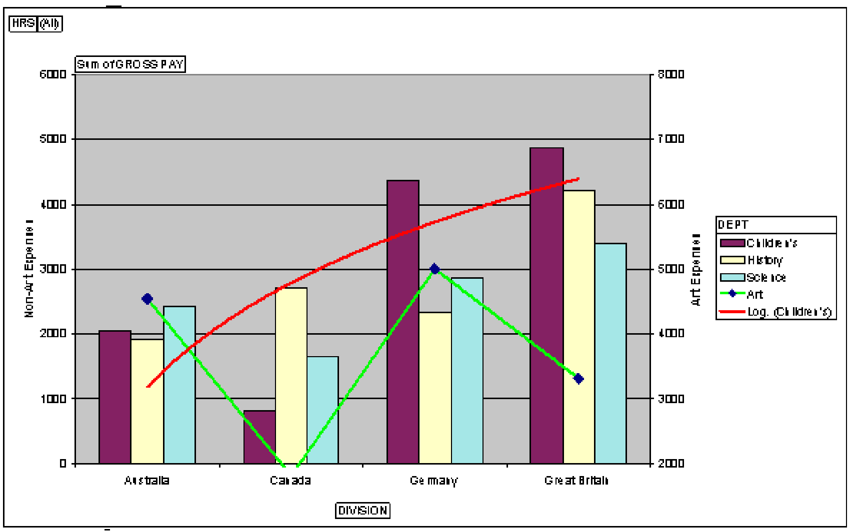


Figure A1.2 – Diagrams produced by the Excel specialized charting capabilities

Many of the charts can be presented in three dimensions, allowing for a variety of special effects.

The charts can be easily moved around the workbook or copied into Word documents or PowerPoint presentations.

As kind of Network Planning tool EXCEL could be used for:

- Entering and storing of network data, e.g. node coordinates, traffic volumes, equipment costs

- Application of simple network planning and forecasting methods, e.g. demand forecasting with trend methods, simplified methods for exchange locations optimization, etc.

- Presenting of tables and charts for quantities of customers, network elements, cost results, etc.

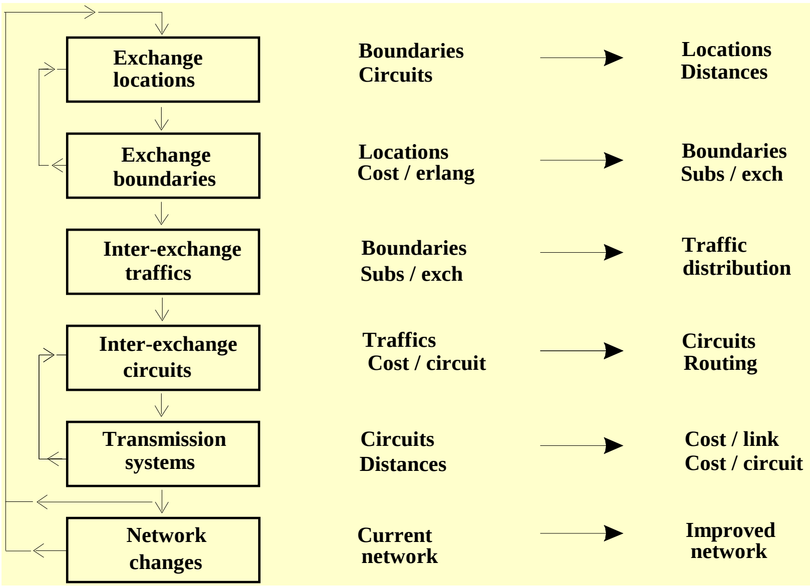
A1.2. PLANITU - ITU

Objective:

PLANITU is a tool for optimisation and dimensioning of circuit-switched telecom networks, based on an integrated interactive approach for finding minimum cost solutions for:

- location and boundaries of exchanges
- selection of switching and transmission equipment
- circuit quantities, traffic routing, switching hierarchy
- choice of transmission paths.

Fig A1.2.1. : PLANITU iterative optimization

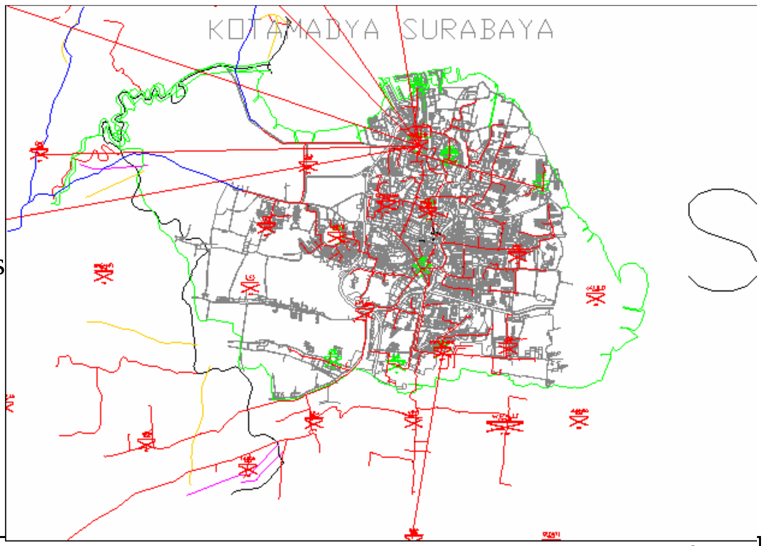


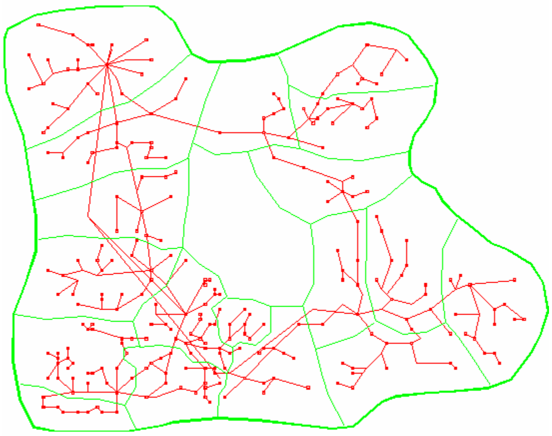
Coverage

Local Networks

Exchange locations

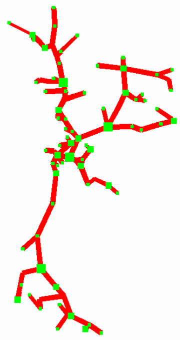
- Exchange boundaries
- RSU locations & boundaries
- Inter-exchange network
- Exchange hierarchy
- Transmission systems





Rural Networks

- Exchange locations & boundaries
- Exchange hierarchy
- Inter-exchange network
- Transmission systems



National & International Networks

- Traffic routing
- Exchange hierarchy
- Inter-exchange network
- Transmission systems

Access network optimization

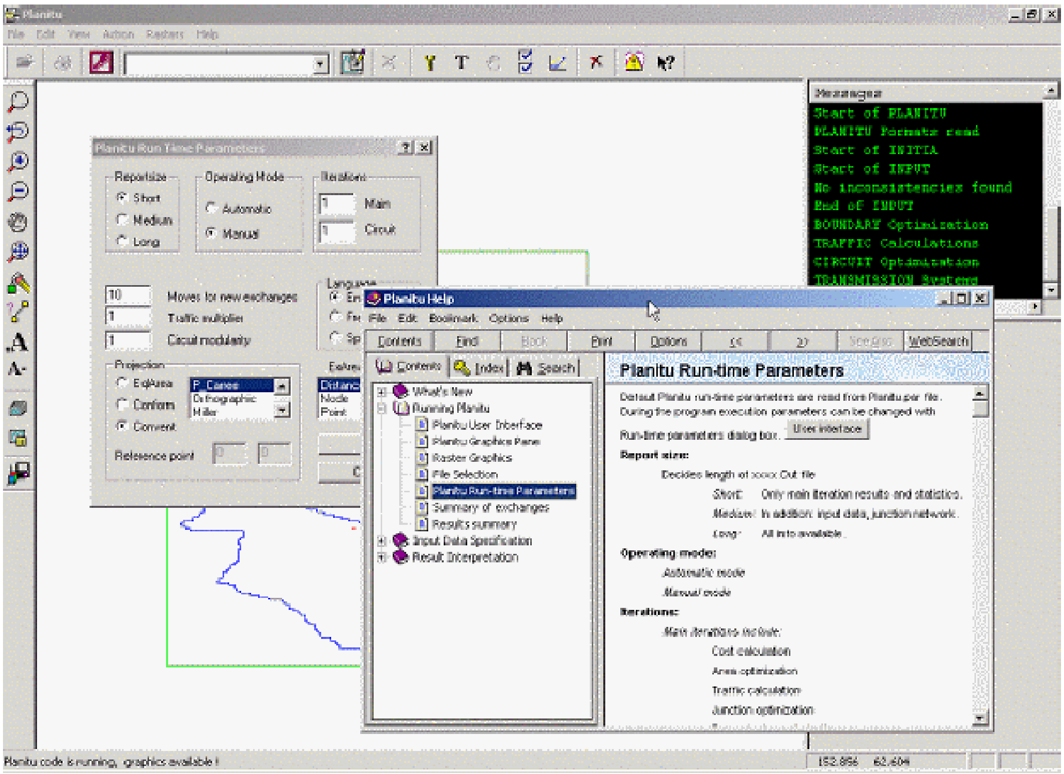
- Dial-up Internet subscriber planning
- Broadband access planning
- Planning of cabinet areas

Backbone network optimization

- Dual homing (load sharing)
- Design of nonhierarchical circuit-switched networks
- Optimization of the fixed part of mobile (GSM) networks
- Optimization of Ring/ Mesh SDH/ SONET transport networks

- Design of ATM, IP MPLS, WDM networks using equivalent bandwidth paradigm

Fig A1.2.2. : Tool appearance and windows



Application :

- PSTN circuit-switched (TDM) networks
- Data (packet) networks – very limited
- Evolution to NGN – limited
- Training tool for network planning

A1.3. STEM

Objective:

STEM is a business decision making support tool that enables the analysis of business models and cost assignment for Telecommunication Networks and services over a period of time**.** Business planning is based on a service demand model that is categorised by market segment, service type and geo-type; That demand drives the resource dimensioning, replacements, costing and revenue generation as well as the calculation of all financial parameters and ratios.

Platform:

In order to run STEM, you must have a PC running Microsoft Windows 95, Windows 98, Windows 2000 or Windows NT, and at least 13Mb of free disk space. Platform is Object-oriented with an editing interface that associates data directly with icons and links between elements. Interfacing is provided to other MS Windows applications like Excel and Word.

(The Analysys STEM network investment modeling tool is a product of Analysys Consulting Ltd, Cambridge, UK see: www.analysys.com)

Issues addressed by the tool

- What are the main cost drivers of the business?
- Which technology offers the most cost-effective solutions?
- What are the average and long-run incremental costs of this network?
- What is the return on investment?
- What impact will changes in busy-hour traffic have on service profitability?
- How will economies of scale and volume influence financial results?
- How sensitive is the business to changes in demand and technology?
- What are the investment implications of varying speeds of network roll-out?

Applications concerned in business planning

- Network strategy planning
- Cost allocation
- Incremental service costing
- Market and competition analysis
- Systems design and sale Research and development
- Business case planning

- Network technology acquisition
- Tariff planning
- Technology and cost comparisons Service roll-out

Modeling Features

The following are the main features and characteristics for the telecom network solutions:

- Service Demand Projection per customer class
- All network layers and technologies considered
- Network modeling with several degrees of detail according to planning objectives
- Any network type and size
- Evaluation of network resources and associated investment (CAPEX)
- Evaluation of revenues for given tariffs and installation rate
- Modeling multiple resource lifetimes
- Modeling multiple time periods
- Modeling of demand elasticity to tariffs
- Interrelation between network growth and operational cost (OPEX)
- Cost assignment as a function of resource utilization rates
- Audit function to have explicit track of precedence, hierarchies, interrelations and causes for primitive or derived results
- Produces automatically the standard financial results like Cash Flow,
- Profit & Loss, Balance Sheet and up to 100 typical business related parameters
- Allows to add new results in tables or graphical formats as customer needs.

Fig A1.3.1 : Activity flow for the STEM techno-economical economical evaluations

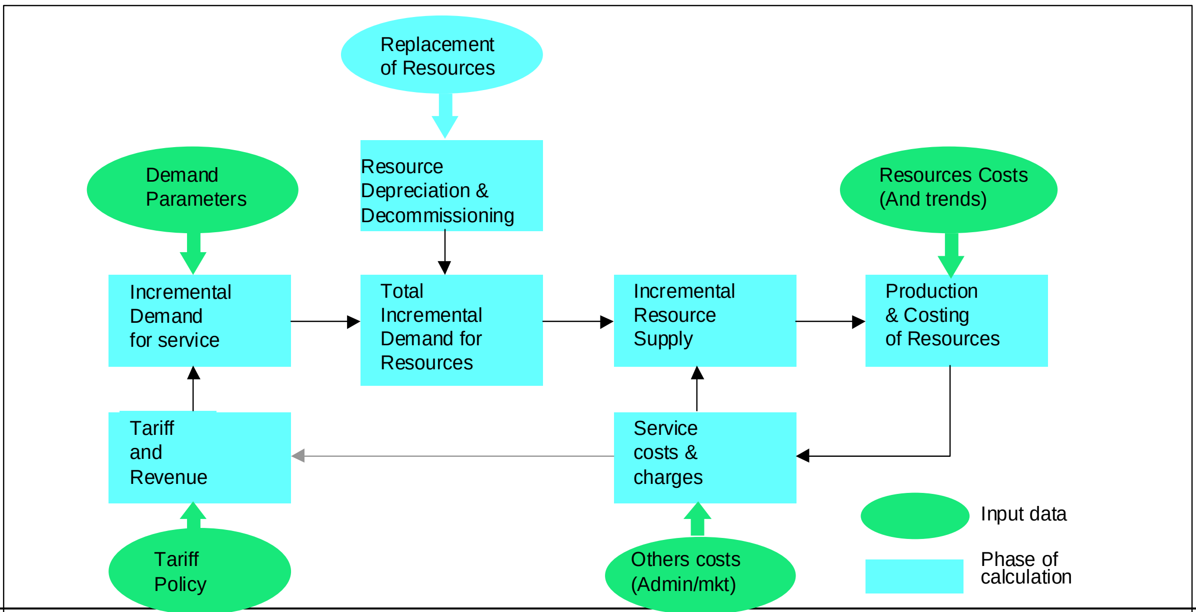
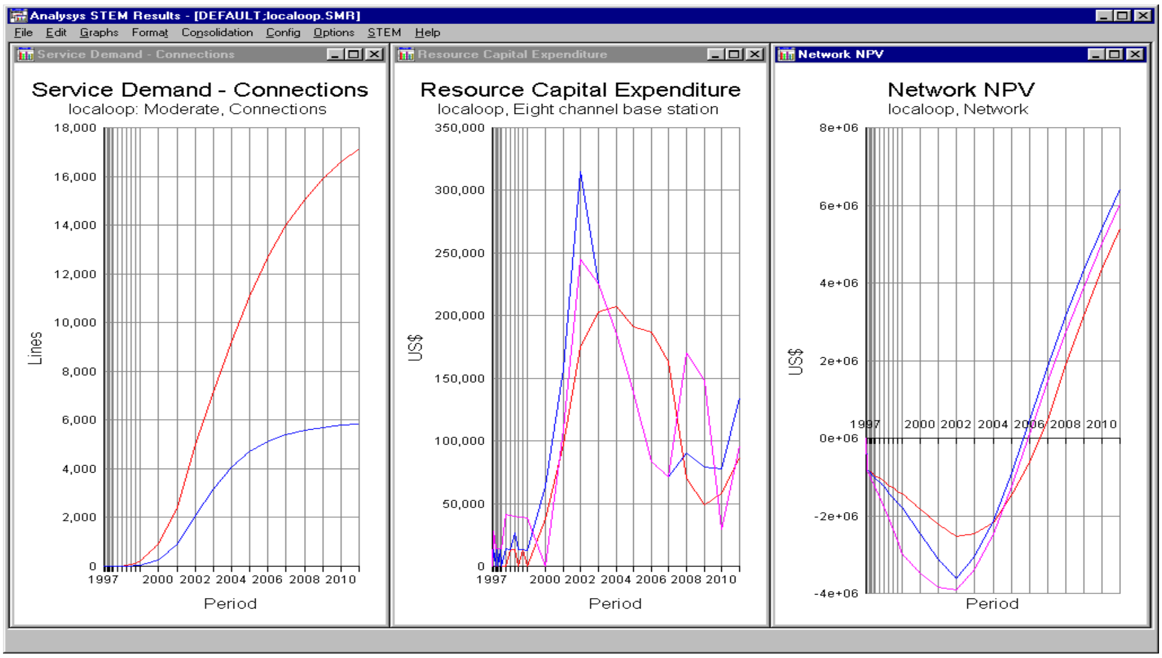


Fig A1.3.2 : Example of tool results

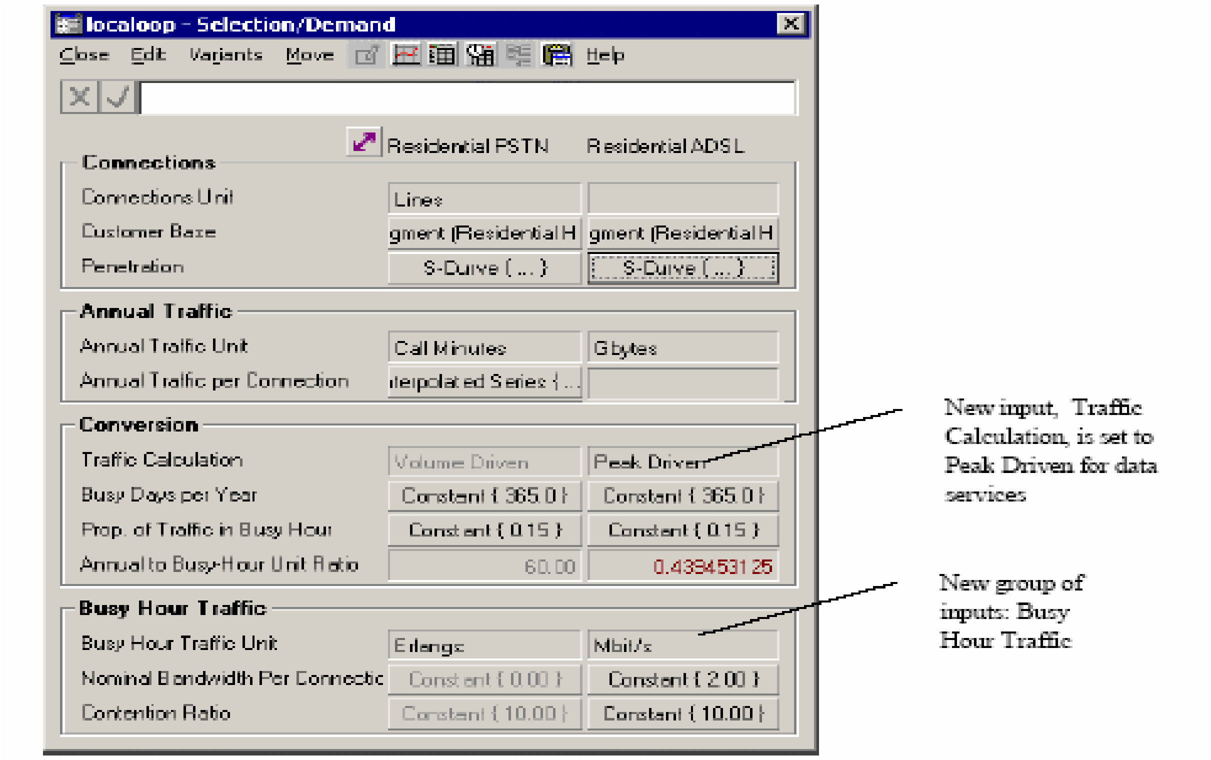


Capabilities of the New version 7.0

STEM version 7.0 was released in September 2004 at the STEM User Group Meeting. This milestone development embraces a new depth of analysis, including intrinsic support for data services, a range of new financial calculations, and the addition of Service results broken down by individual Resources. Clarity is the focus of this release, enhanced by new formatting options in both the Editor and Result programs. Main new capabilities are summarized:

- ***Demand parameters for data services***

STEM 7.0 makes it straightforward to model data services directly. The Service Demand dialog has been extended to allow more intuitive entry of demand parameters for data services. The new 'Peak Driven' calculation option works from new bandwidth and contention inputs and makes annual traffic volume a calculated quantity.



- Multiple contention ratios

In general, there will be a contention ratio at the first point of aggregation, on the immediate access pipe, where traffic may aggregate across, for example, other dwellers of a multi-tenant unit. However, deeper in the core network, there may be a further averaging effect due to multiple nodes coming together, resulting in a higher overall contention ratio. This can be readily modelled through the multiplier of an intermediate Transformation.

- New financial modelling options

-Time-series unit costs and inline trend for Capital Cost

STEM models are typically forward-looking and made up of a combination of known quantities and estimates. So unit costs were originally defined as a spot value for a calibration period, linked to per-Resource and global trends, and multiple cost indices for capital costs. All trends and cost indices were normalised separately for each Resource according to its calibration period. This simple approach allowed for the rapid development of cost models, with efficient re-use of global cost trends. However, this approach did not accommodate other ways of working, such as linking to time-series costs calculated from other sources, or capturing known values and historical costs.
STEM 7.0 provides the flexibility users want. Unit costs are now time-series in their own right, allowing you to enter known or linked time series directly and leave cost trends unset. But you can still enter constant calibration values and use inline or global trends when this approach is easier or quicker.

A new Use Global Trends input also makes it possible to completely decouple cost assumptions for certain Resources, while retaining global trends as an efficient common hypothesis for the remainder.

-Modelling loans and bonds with individual debt facilities

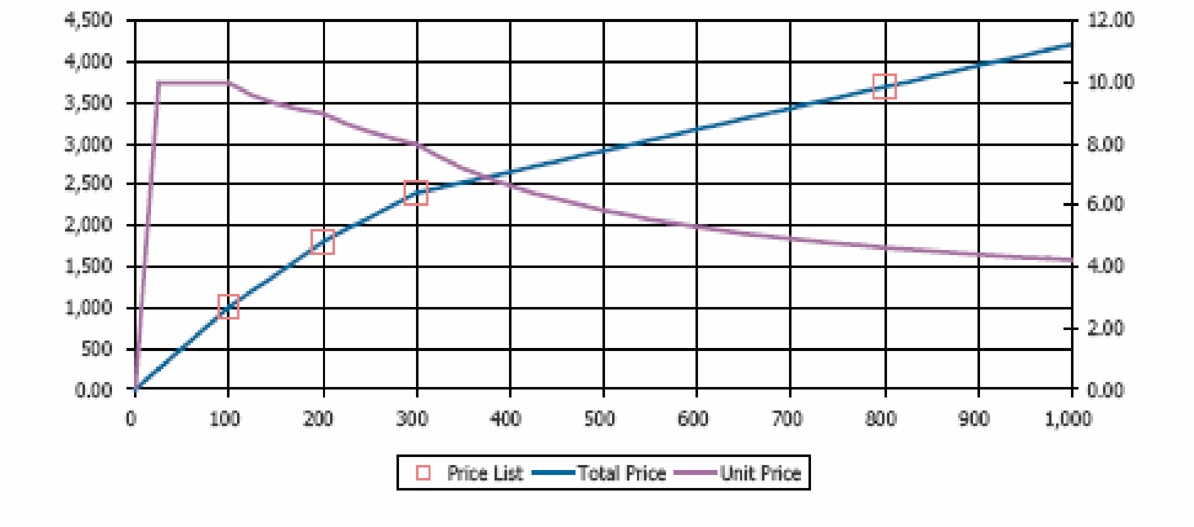
In recognition of the need to model mixes of loans variety, STEM 7.0 allows for the definition of individual debt facilities, with separate terms and schedules, to support the funding requirements for a network.

A typical debt facility has a fixed *term* (defined as a *start date* plus n years) within which funds may be borrowed, up to an optional *credit limit*. Separate *grace periods* are allowed for the principal repayment or amortisation (gp) and the interest payments (gi). New borrowing is only allowed up to gp where gp ε gi. After the principal grace period, amortisation may be paid in one or more equal annual instalments (*straight-line amortisation*), or there may be an *optional repayment profile*.
There is typically a *fixed interest rate* for the term, with payments calculated from the year-beginning balance. Small adjustments are made for the interim amortisation: a.(n – 1) / 2n, where a = annual amortisation, and n = number of payments each year:

-Price lists for Resource capital costs

The smooth logarithmic model for economies of scale from earlier versions of STEM has been replaced with a straightforward and intuitive price-list model which can capture specific supplier price points. (The logarithmic model was fine in theory but inflexible and almost impossible to match to real-life data.) STEM 7.0 has a new per-Resource Price List input which makes it possible to specify a series of cost/cumulative volume breakpoints.

The total cost for intermediate quantities is calculated by linear interpolation, that is, linear from zero to the first breakpoint, and also beyond last breakpoint at same gradient as the last segment:



Linear interpolation of total costs is equivalent to a reciprocal interpolation betwee unit cost breakpoints. Therefore costs defined as unit costs or discount factors are first scaled by breakpoint quantities before interpolating as total costs;

- ***Intelligent formatting options***

-Number format and orders of magnitude

Feedback from users showed that they would prefer to suppress excessive precision in the formatting of numbers, for the sake of clarity and consistency. In line with this, STEM 7.0 has a customisable number formatting system which allows the user to specify:

– an order of magnitude for a graph, such as millions or billions, which will be used automatically to qualify the y-axis scale

– the number of digits and decimals to be shown in a table

– whether values should appears as percentages.

These options can be applied to individual graphs and tables and are also stored with all pre-defined graphs in the Results configuration.

However, the selection of a number format, and especially the choice of orders of magnitude, depends critically on the actual numbers at hand. In other tools it may be necessary to manually format each graph or table to suit the specific data. In contrast, STEM 7.0 includes a system that will *automatically* select an appropriate order-of-magnitude label and the optimal number of decimals, depending on the actual values displayed. This system is controlled by a set of global parameters which you can control with a new Number Format dialog, which is accessed from the Options menu in the Results program.

-Routine preview of inputs with Auto Graph

In the Editor, a chosen input can be graphed by clicking on the Graph button in the appropriate dialog. STEM 7.0 extends this feature with the addition of a window which will remain on screen and *automatically* display a graph of whatever input parameter is selected.

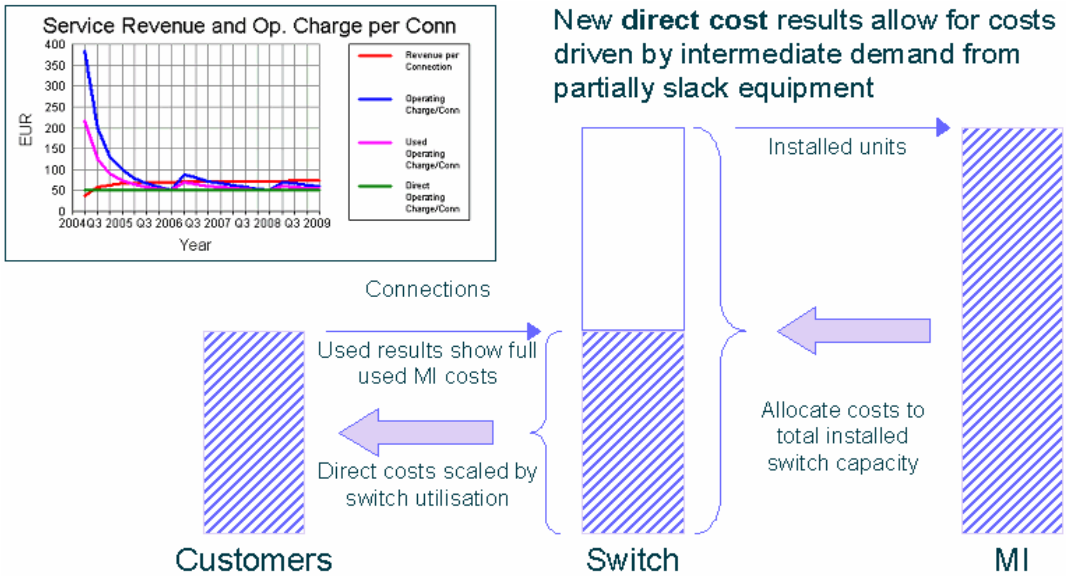
The Graph button on a data dialog showing parameters for an individual time series immediately graphs that time series.

-Visual grouping with colour blocks
Another improvement to the user interface in STEM 7.0 is the option to use blocks of colour to group together a set of model elements. This makes it easier to see at a glance which elements belong together, for example the Services grouped within a particular Collection. Colour blocks may be drawn around the members of Collections, Market Segments, Functions, Dimensions and Templates.

- Advanced service costing
A defining feature of STEM version 7.0 is the generation of allocated Service cost results, broken down by individual originating Resources, based on a scaleable implementation which is designed to avoid a combinatorial impact on performance.

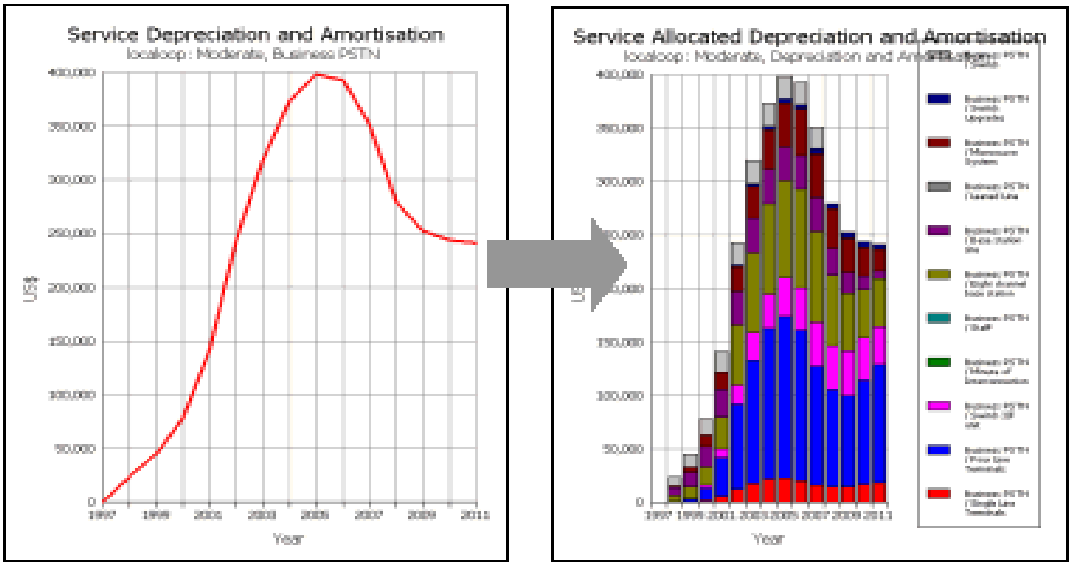
STEM performs detailed cost-allocation through the calculation framework for handling incremental demand which is generated when a model is run. Appropriate shares of Resource costs are routed to the respective Services responsible for used and slack capacity. However, the 6.2 model engine only stored results for the total allocations, mainly because costs were aggregated by intermediate Transformations before being passed back to Services (though originally also to limit the size of results files on disk).

New direct cost results allow for costs driven by intermediate demand from partially slack equipment, and vary more closely in relation to the underlying service demand than used cost results. Direct costs represent the cost of a fully efficient network and are the keenest indicator for tactical pricing.



STEM 7.0 stores separate results for the costs of each separate Resource which are allocated to a given Service. No changes to the inputs are required; a STEM 6.2 model must simply be re-run with the new model engine. In the new Results program, we have added two new types of result, 'Service / Resource' and 'Transformation / Resource', for which all the usual cost results are available, as well as a Used Capacity

result which can be used to understand respective Service shares of the installed capacity of a Resource.



A1.4. VPIsystems

Planning and Engineering Solutions from VPIsystems

VPIsystems, Inc. is a recognized expert in engineering decision support software used for modeling and optimizing communications network capacity and improving the efficiency of the network planning and engineering process.

VPIsystems provides network operators with solutions for improving the processes associated with network resource and capacity planning, deployment, and network optimization that result in major savings in CapEx and OpEx. VPI solves problems such as improving capacity utilization,optimizing new builds, streamlining network planning, optimizing tariffed network bids and central-office de-risking.

VPIsystems provides communications equipment vendors with solutions that automatically design and specify optimum equipment configuration in response to customer proposals and quotation requests. VPI software provides a competitive sales advantage, allowing you to preserve margins, eliminate design errors and be more customer responsive.

VPIsystems provides optical subsystem and communications equipment designers powerful simulation tools to optimize product designs. VPI software can enable you to rapidly compare multiple designs and predict performance without time consuming lab prototyping, leading to better and more cost effective designs.

VPIsystems has offices in Berlin, Melbourne, Minsk and U.S.A. Its Berlin office is located at: 6, Carnot Strasse, Berlin 10587, Germany. For more information about VPIsystems and its range of products contact: info@vpisystems.com

Overview of Integrated Network Planning using VPI products

The entire product range offered by VPIsystems is shown in Figure 1 below . The product range is offered as individual modules that are integrated on a platform called the

VPIlifecycleManager^{TM} . Each module addresses a unique planning function in the multi-technology, multilayer, telecommunications network hierarchy.

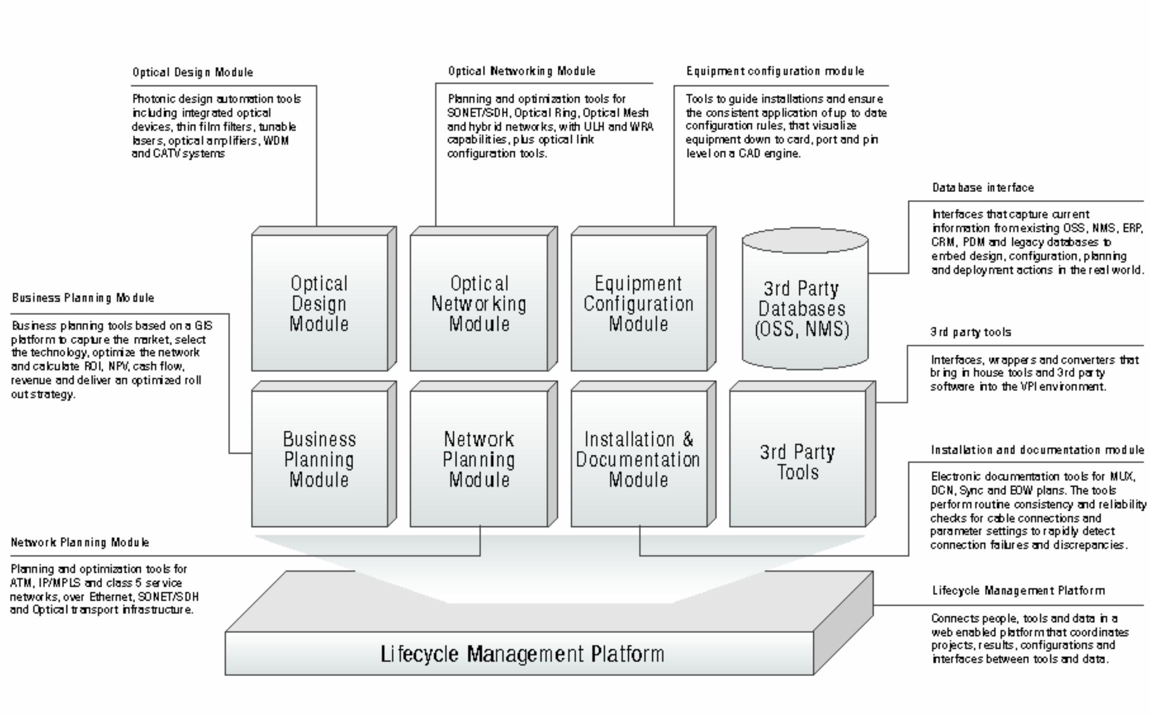


Figure 1 VPIlifecycleManager and supported modules

New Planning Environment

Several factors have combined to create a unique opportunity and need for integrated network planning today. Salient among these are:

1. market competition due to deregulation
2. the transition from voice to data-centric networks
3. the introduction of new technologies like IP/MPLS, ROADM, and MSPP.

Effective network planning and design is essential if the best use is to be made of new technologies while at the same time taking into account traffic distribution patterns, changing economic conditions, and new network concepts. The primary objective is to reduce investment and operational costs while improving service quality and flexibility.

Increased competition, the emergence of low cost service alternatives such as packet-voice, and the advent of high-volume/low-revenue services, such as Internet access, have significantly eroded the profit margins traditionally enjoyed by network and service providers. Consequently, a high degree of efficiency in all aspects of service and transport network operation—from service provisioning to infrastructure planning—is a must for network and service providers to stay competitive. It is becoming increasingly clear that an *integrated* network planning environment—one that provides a common platform to support all aspects of service and transport network planning—is the answer to the problem of improving the efficiency of the planning process. Providing access to a suite of expert design tools that optimize individual service and transport networks, while at the same time affording simplified access to network data via seamless integration with Network Management

Systems (NMSs), is the most effective way of enhancing a network provider's ability to compete.

Current Network Planning Process

Most network planning activity today is a widely varying mix of individual decisions based on guesswork or intuition, ad hoc back-of-the-envelope calculations, and occasional use of sophisticated stand-alone tools focusing on specific parts of the overall network planning problem. There is a clear lack of an integrated, systematic, quantitative approach that allows the planning process to be modeled in its entirety. Groups responsible for the design of specific service networks—such as circuit-switched voice, IP or ATM—design their networks as best as they can and rely on transport network planning groups to provide them with the bandwidth needed as cost-effectively and reliably as they can.

The transport network planning groups, in turn, attempt to optimize the transport network based on traffic forecasts received from the various service network planning groups. Typically, these groups make decisions on the network technologies and architectures to be deployed based on a combination of needs, such as getting the lowest cost per unit bandwidth, meeting the SLA requirements imposed by specific services, the type of restoration needed, the ease of operation, and so on.

This two-step process, involving individual service network planning and transport network planning, is further complicated by the lack of coordination among different organizations and areas of expertise, and by the existence of a plethora of NMSs each containing a subset of the network data. There are often multiple, inconsistent copies of a certain set of data and to achieve data consistency is in itself a problem of considerable magnitude for most network providers. As a consequence, the combined planning process often tends to be fragmented, inefficient, and inordinately long.

The Solution to Network Planning

By providing a complete suite of planning tools together with integrated NMS access and database support on a common platform, the VPIlifecycleManager™ radically transforms the network planning process, rendering it simple, transparent, fast, and cost-efficient. It does this by enabling an integrated network planning approach that eliminates guesswork, cross-organizational miscommunications, and manual processes that are costly, inefficient, and time consuming. The collaborative planning environment offered by VPIlifecycleManager™ supports teamwork across different organizations and functional areas, and promotes the seamless gathering of network data from different NMSs. This allows multi-organizational teams to effectively share their expertise and needs, and to collaboratively evaluate alternative service and transport network architectures to arrive at a solution that achieves the cost-performance compromise in a manner best suited to the overall company goals.

VPIlifecycleManager supports the following set of tools:

VPIserviceMaker™ATM

VPIserviceMaker™IP

VPIserviceMaker™Switch
VPIserviceMaker™SS7
VPIaccessmaker™
VPIserviceMaker™Distribution
VPItransportMaker™
VPItransportMaker™Sync
VPIlinkConfigurator™

These tools are brought together for collaborative employment by platform software known as VPIlifecycleManager™. Figure 2 below shows the interrelations between these tools.

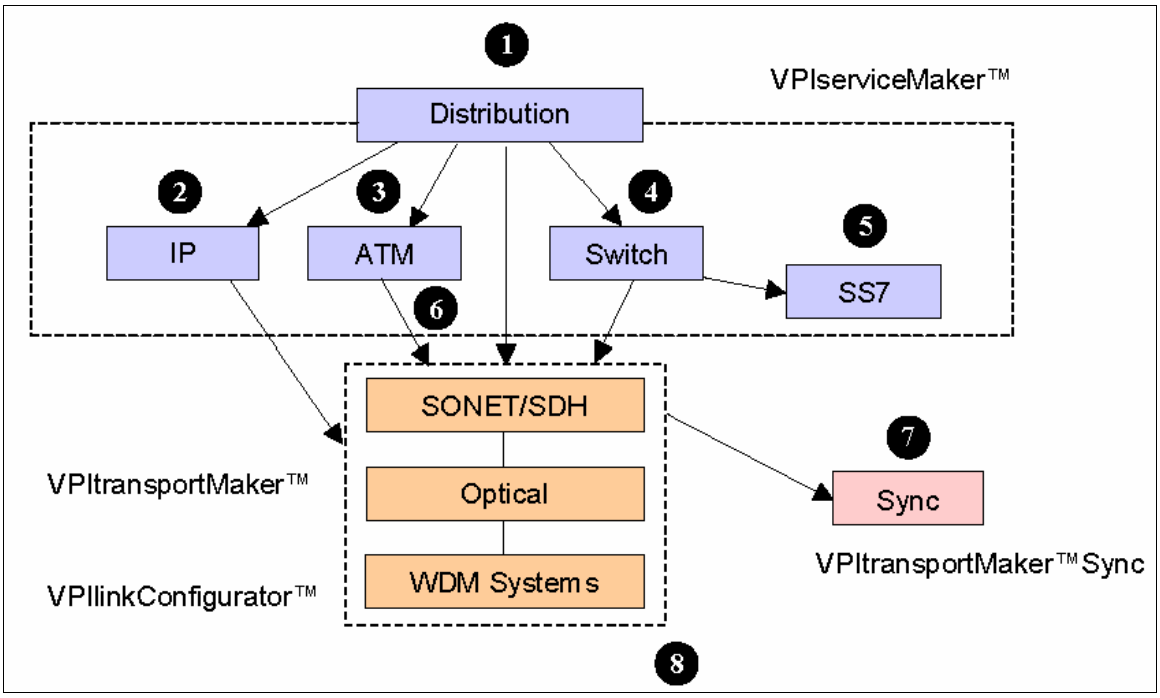


Figure 2 Interrelations between VPI Tools

VPIaccessMakerTM is a tool used for modeling business plans and feasibility studies for the deployment of access technologies. The tool captures subscriber information, models different service combinations and technologies and selects the best technology for the task, to calculate ROI, NPV, cash flow, revenue and an optimized roll out strategy. VPIaccessMaker uses a Geographical Information System (GIS) system, based on Map Objects ESRI® and MapX - MapInfo® to provide direct visualization of maps, network equipment and data**.**

VPIserviceMaker™Distribution is used to generate point-to-point traffic matrices based on various demographic and geographic assumptions. In addition, this tool can be used to estimate unknown traffic for the next planning period based on certain traffic growth assumptions. This tool offers an excellent means of generating an initial traffic matrix from uncertain data and offers various controls to perform what-if analyses.

VPIserviceMaker™IP is used for planning IP networks. It can perform IP network capacity design for various service classes supported on the Internet, such as best effort services (web browsing, HTTP, FTP, e-mail, etc.) and VoIP (Voice over IP). VPIserviceMaker™IP further supports OSPF (Open Shortest Path First) topology, effective bandwidth calculations for bursty traffic, network bottleneck identification, failure simulation, and modeling of link interface costs.

VPIserviceMaker™ATM is used for ATM network planning. The key functions provided include equivalent bandwidth calculation for bursty data, topology design based on traffic, link and equipment costs, Virtual Path (VP) classification based on traffic, VP dimensioning to determine the capacity to be assigned to a given VP, Virtual Path Connection routing and bundling, shared capacity restoration computation, network bottleneck identification and overload correction, equipment list and network cost generation, and call level simulation to verify that performance objectives are met.

VPIserviceMaker™Switch is a tool for planning and dimensioning circuit switched networks. It supports optimized trunk group dimensioning based on topology, traffic volume, and alternative routing strategies. It can also be used to determine a cost-optimized routing scheme based on a specified cost factor for each trunk group. The tool can also evaluate the traffic carried, and trunk group blocking, in existing networks.

VPIserviceMaker™SS7 is used to design common channel Signaling System No. 7 networks. It can optimize the signaling network topology, dimension signaling link sets, and route signaling traffic. VPIserviceMaker™SS7 can also analyze an existing SS7 network for routing and reliability problems, undertake failure analysis, and perform bottleneck analysis.

VPItransportMaker™ supports transport network design based on PDH, SONET/SONET, and optical networking technologies. It covers a wide range of network architectures, including ring, mesh, and ring–mesh hybrids. Using VPItransportMaker™, a planning engineer can optimize the network topology and determine the appropriate routing, protection, restoration, and equipment. VPItransportMaker™ allows you to define a variety of technology, architecture, and network constraints that must be explicitly honored during the design. The tool also supports an analysis module for checking design results and for performing what-if analyses. VPItransportMaker™ can be used for planning metro and long-haul networks.

VPItransportMaker™Sync is used for planning synchronization networks, including the optimization of clock distribution in DS1/E1 synchronous communication networks. The purpose of a synchronization network is to ensure that all DS1/E1 elements in a network have a timing source that is accurate to within acceptable tolerances. VPItransportMaker™Sync enables synchronization planning by generating a master clock distribution plan, creating alternative back-up timing supply routes, and providing an analysis and evaluation capability that checks for timing loops and for breaks in the timing supply paths under various failure scenarios.

VPIlinkConfigurator™ is used to design links in optical networks. It supports the synthesis of multiple design alternatives given a set of network planning requirements. VPIlinkConfigurator™ enables you to compare and rank link design alternatives. You can easily modify the synthesized link design, create new designs, and perform what-if analyses. You can also undertake sophisticated analyses of a link's optical performance and cost, taking into account the WDM system technology used, the fiber type, and the type of optical amplifiers.

Integrated Design Flows Supported in VPIlifecycleManager™

Depending on the network provider's environment and network planning needs, various design flows can be instigated using the VPIlifecycleManager™. The most common design scenario involves the following steps:

1. Generation of traffic matrices using VPIserviceMaker™Distribution. One or more traffic matrices may be generated for each service layer network to be designed. For example, you can generate a VoIP matrix for the Voice over IP application, an ATM CBR (Continuous Bit Rate) matrix for ATM network design, a point-to-point Erlang traffic load matrix for circuit-switched voice network design, and a point-to-point bandwidth matrix for designing a transport network at any desired layer (PDH, Ethernet, SONET/SONET, or Optical). Alternatively, you could begin your design process with VPIaccessMaker to model customers and service characteristics and derive the service demand for each access serving area. This can then drive VPIserviceMaker™Distribution to generate the traffic matrices for each service type.

2. The traffic matrix generated is used in conjunction with the appropriate VPIserviceMaker™ tool to produce an optimized service layer design, approximate equipment interface, and link/trunk costs. The link/trunk dimensioning provides a point-to-point bandwidth demand requirement that is then fed into VPItransportMaker™ to be incorporated into the design of the transport network.

3. Each point-to-point bandwidth demand matrix resulting from step 2, along with those generated in step 1, together with associated requirements for routing, protection, topology, restoration, available equipment types, and so on is fed to VPItransportMaker™. Here, every possible combination of transport network technology and architecture can be analyzed to determine the most cost-effective solution that meets the various SLA requirements.

While steps 1–3 above constitute the most straightforward workflow for undertaking network design, a number of iterative design loops may be executed to improve a specific design or to try out different combinations of service and transport network technologies and architectures. Thus, we may go back to step 2 after step 3 and revise the assumed link/trunk costs based on the total network cost determined during step 3. This leads to a revised set of service network designs, which in turn leads to a modified transport network design whose total network cost is now different. You can repeat this process by revising your service network designs using the newly derived average link/trunk costs. Perfect convergence, wherein the average service layer link/trunk cost in step 2 is in perfect agreement with the newly derived link/trunk cost after step 3, is rarely achieved. The iterative process is usually halted when the designer is satisfied that agreement within acceptable limits has been achieved.

The above process of collaboratively determining the needs of different service networks and the transport network needed to carry the services enforces a planning discipline that will lead to a vastly improved planning process. The network provider realizes immediate benefits in terms of fewer technical staff, more cost-effective designs, and shorter planning times.
Many other design refinements are possible even with the relatively simple three-step process described above. After step 3, you could go back to step 1, modify the assumptions behind traffic matrix generation, and then investigate the resulting impact on total network cost after step 3. Alternatively, you could revisit step 1 after step 2, change the traffic generation assumptions, and see how this affects service layer designs.

More sophisticated analyses could take into account the effect of new equipment that achieves layer integration. Thus, after step 3 you could examine equipment architectures that integrate service and transport layer functionality. This would allow you to derive, in addition to average service layer link/trunk costs, new service layer interface costs, and then to repeat steps 2 and 3. For example, an IP/WDM designer might first assume stand-alone IP router port costs at step 2 and, after doing a WDM mesh design at step 3, consider equipment that combines the functionality of an IP router, DWDM (Dense Wave Division Multiplexing), and OXCs (optical cross connects). The average router port costs derived from this new equipment type might be sufficiently different to warrant revisiting step 2. This may lead to a significantly improved IP layer design, a correspondingly improved WDM network design, and much lower overall network costs. It is evident, even in this simple illustration, that the collaborative planning methodology coupled with the use of powerful planning tools can lead to very significant cost savings.
Several other design flows are feasible, and would make sense in many practical situations. For example, you could first design a circuit-switched voice network using VPIserviceMaker™Switch and then design the associated signaling network using VPIserviceMaker™SS7. Another common scenario involves designing a PDH or SONET/SONET transport network and then undertaking synchronization planning using VPItransportMaker™Sync. Using the VPIlinkConfigurator™ tool, a network provider could then obtain a complete characterization of the network’s embedded fiber database and use it, in conjunction with various vendor implementations of WDM systems, to determine the best optical network that may be supported by its installed fiber base.
As one example of such a planning process, consider an optical sub-network consisting of old fiber carrying only OC-48 signals, a second sub-network carrying a mix of OC-48 and OC-192 signals, and a third consisting of a transparent optical island comprising the latest wavelength banded Ultra-long Haul WDM systems with OADMS (Optical Add/Drop Multiplexers) that provide band add/drops and support multiple WDM system branches.

Demonstration of Integrated Network Planning

This section illustrates how you can conduct integrated network planning across multiple layers and technologies using VPIlifecyclemanager™. The example selected represents a typical network planning problem from a nationwide network provider’s perspective, namely the USA backbone network shown in Figure 3 below . This is used as the underlying physical network where nodes represent traffic generation/switching points and links represent fibers connecting the nodes. To show how VPIlifecycleManager™ can model the multi-layer multi-technology complexity of today's telecommunication networks, we will design multiple service networks and one transport network. The service networks are an IP network, an ATM network, a switched voice (PSTN) network and an SS7 signaling network. All are supported over a common transport infrastructure, being a SONET ring network carried over a DWDM transport infrastructure.

Before starting any network design, you need to determine the traffic that the designed network will have to serve. Traffic information is the fundamental driving force behind any network design. Normally traffic, if not available, can be forecast from historical data, or estimated by market projection and subscriber population. Here we assume both cases; that is, some traffic is known, and some traffic is unknown and has to be estimated.

The unknown traffic is estimated using VPIserviceMaker™Distribution, which provides many models for estimating traffic. The data flows across a number of tools in this example, as illustrated in Figure 2 . The service layer networks (IP, ATM, Switch) are designed first, one by one. Then all the point-to-point bandwidth pipes required by service layer networks are mapped to SONET/SONET layers to design a common transport network. Finally SONET/SONET pipes are bundled into point-to-point DWDM transmission links.

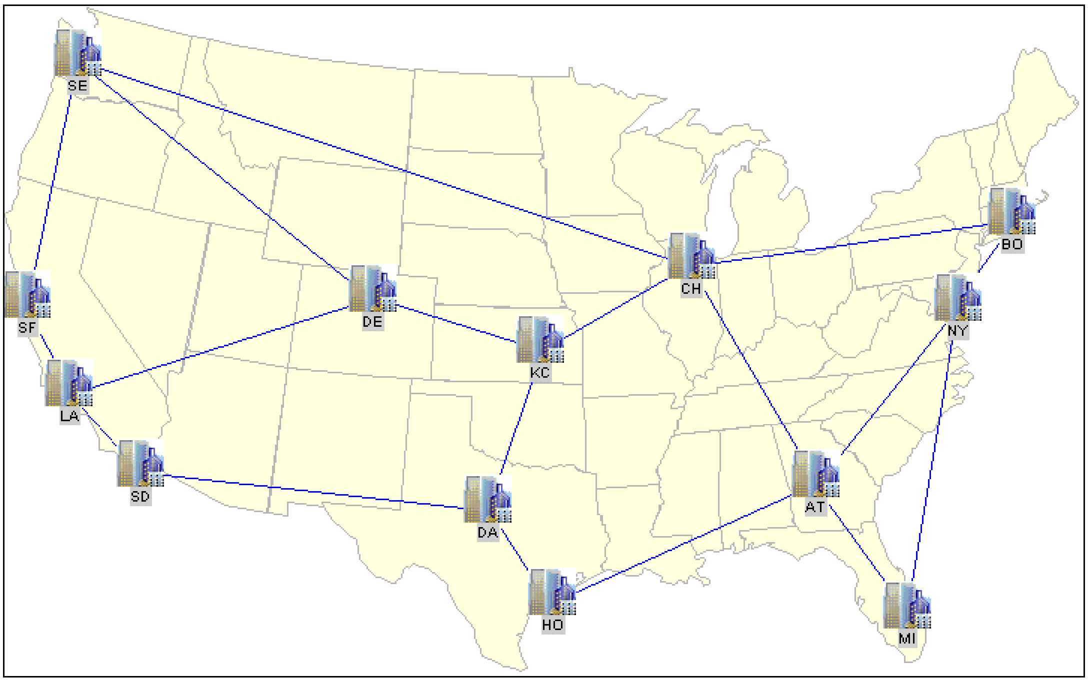


Figure 3 USA national backbone network

VPIserviceMaker™Distribution

In this section we show how to use VPIserviceMaker™Distribution to generate a matrix of point-to-point voice traffic (Erlangs) for the switched network design. We will assume that 10% of this voice traffic is VoIP traffic This will be the input data fed into the IP network design. In addition, we generate a traffic matrix of point-to-point constant bit rate services which will be fed into our ATM network design. Furthermore, we also want to create a traffic matrix of point-to-point circuits representing leased line services to feed into the transport network design.

We start by importing the network node data from the US national backbone network (see Figure 4). Each node represents a major city.

Note that all VPIlifecycleManager™ tools have the same underlying network model. Hence, network topology can be shared across different tools, either by explicitly using the import/export functionality in each tool, or by simply opening another tool's project file directly.

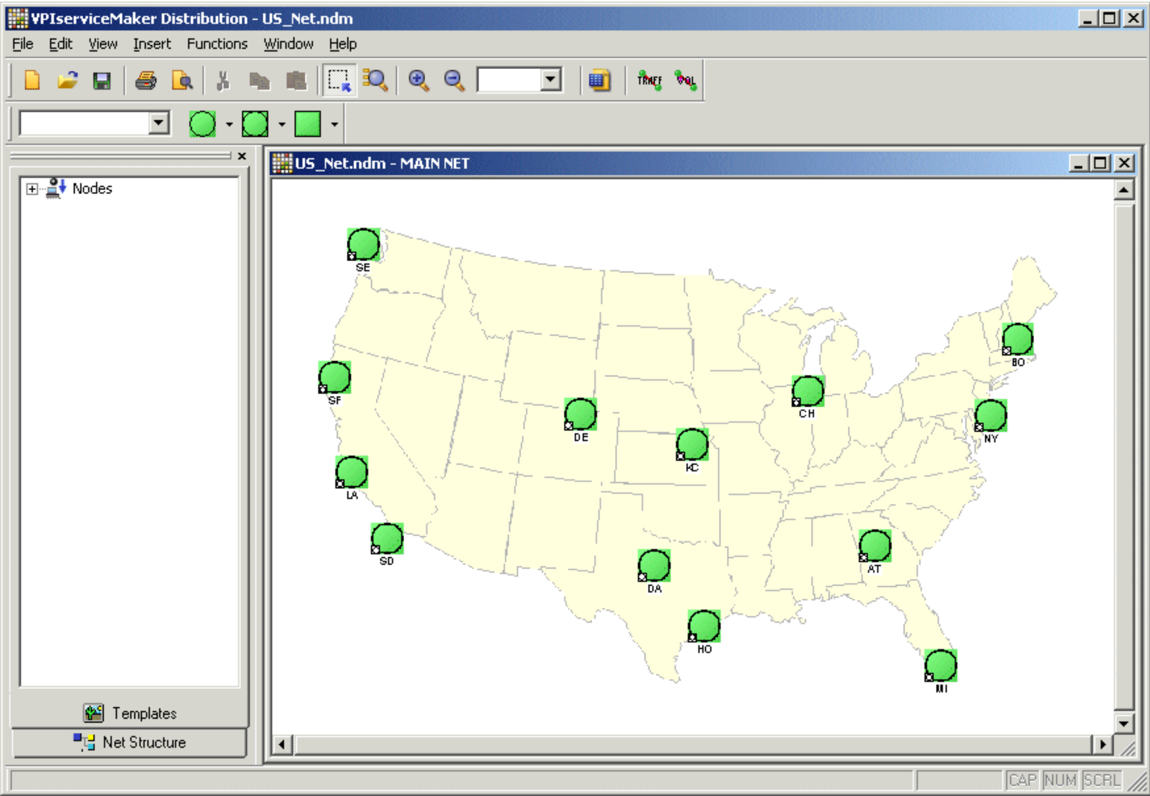
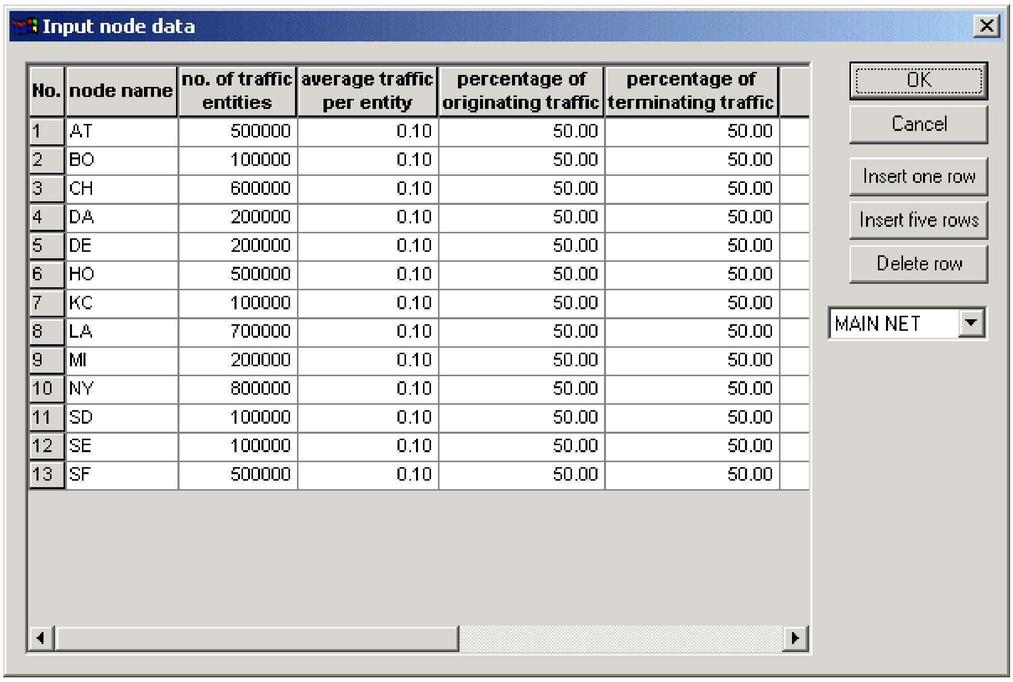


Figure 4 The set of nodes for traffic distribution

Next we specify the subscriber population of each city (in proportion to the city's population) and the average traffic per subscriber, as shown in Figure 5.



Input node data

No.	node name	no. of traffic entities	average traffic per entity	percentage of originating traffic	percentage of terminating traffic
1	AT	500000	0.10	50.00	50.00
2	BO	100000	0.10	50.00	50.00
3	CH	600000	0.10	50.00	50.00
4	DA	200000	0.10	50.00	50.00
5	DE	200000	0.10	50.00	50.00
6	HO	500000	0.10	50.00	50.00
7	KC	100000	0.10	50.00	50.00
8	LA	700000	0.10	50.00	50.00
9	MI	200000	0.10	50.00	50.00
10	NY	800000	0.10	50.00	50.00
11	SD	100000	0.10	50.00	50.00
12	SE	100000	0.10	50.00	50.00
13	SF	500000	0.10	50.00	50.00

OK
Cancel
Insert one row
Insert five rows
Delete row
MAIN NET

Figure 5 User population and profile

VPIserviceMaker™Distribution will then create a matrix of point-to-point Erlang loads, as shown in Figure 6. This voice matrix is then exported to a text file that can be used later in the switched voice network design.

We will come back to VPIserviceMaker™Distribution again when we need to create a traffic estimate for ATM and SONET network designs. IP design assumes 10% of the voice matrix as its VoIP traffic by specifying one tenth of subscriber population.

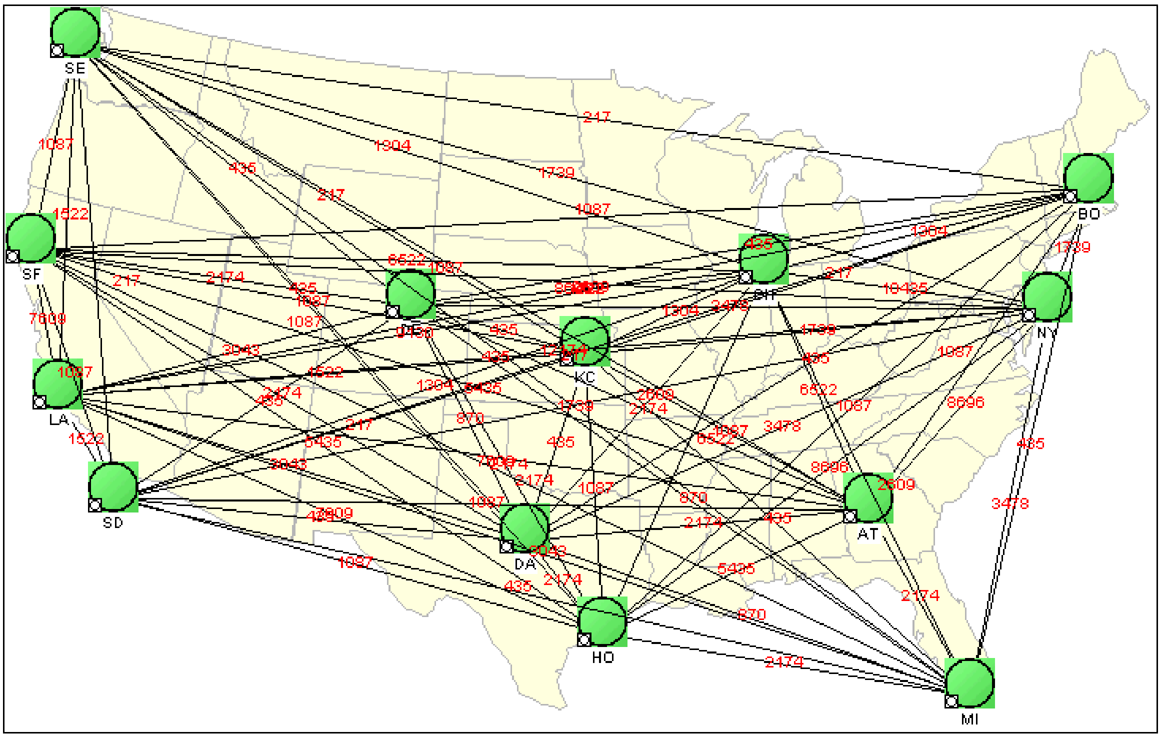


Figure 6 Pomit-Point Voice traffic

VPIserviceMaker™Switch

Now we design a switched voice network (PSTN) using the voice traffic generated by VPIserviceMaker™Distribution. Again we begin by importing network node and topology data into VPIserviceMaker™Switch. The result is shown in Figure 7 . We use the physical network topology as the initial topology. This can be changed after running the topology improvement function of VPIserviceMaker™Switch.

The next step is to import the voice matrix. To do this, we must first create an empty traffic matrix. We then open the matrix window and import the voice matrix data from VPIserviceMaker™Distribution. After running design optimization, we get the trunking design of the switched network shown in Figure 8, where the number by each link indicates the size of the link in DS0 units. A file of this data will be used later by VPItransportMaker™ as one of its input traffic matrices.

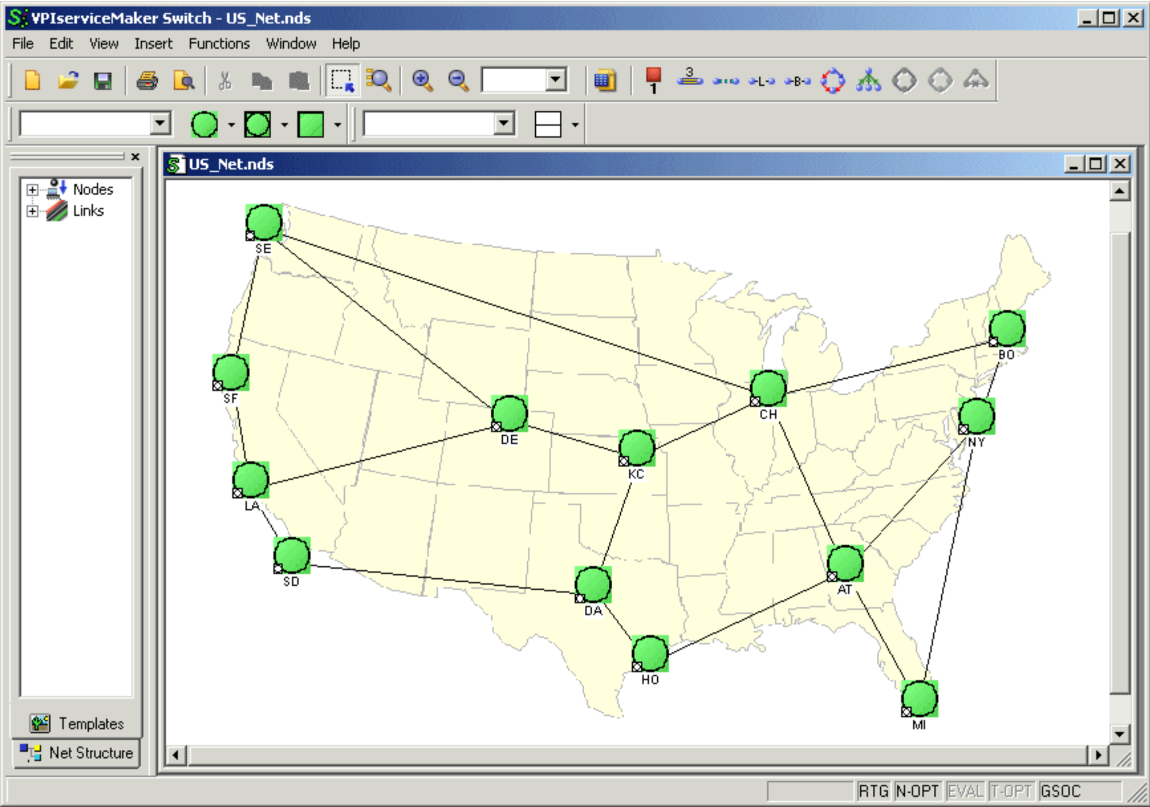


Figure 7 Switched voice network initial topology

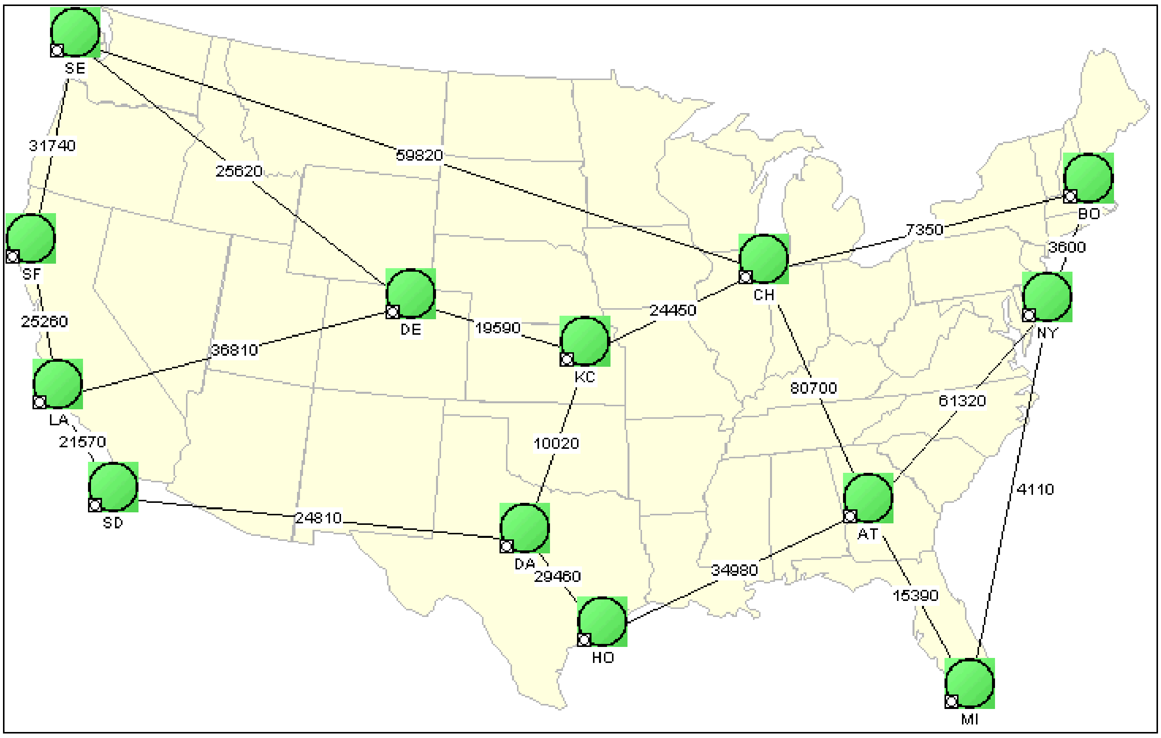


Figure 8 Trunk size of switched network

A SS7 signaling network can be designed in association with the switched voice network design. But since the bandwidth requirements produced by the SS7 signaling network are negligible for the SONET transport network design, the SS7 design is not included here.

VPIserviceMaker™IP

Next we design the IP network. Again we start by importing the basic network data into VPIserviceMaker™IP. But since the IP tool models traffic at the individual service level, we need to create user populations (each one represented on screen by a pair of computers) and servers (represented on screen by a single computer). These are new nodes added to the network. The original nodes are considered as core switching centers (and represented on screen by core routers). We also purposely add more connectivity. (Unused links—that is, links with no load—will be automatically deleted at the end of the design process.) Figure 9 shows the IP network setup.

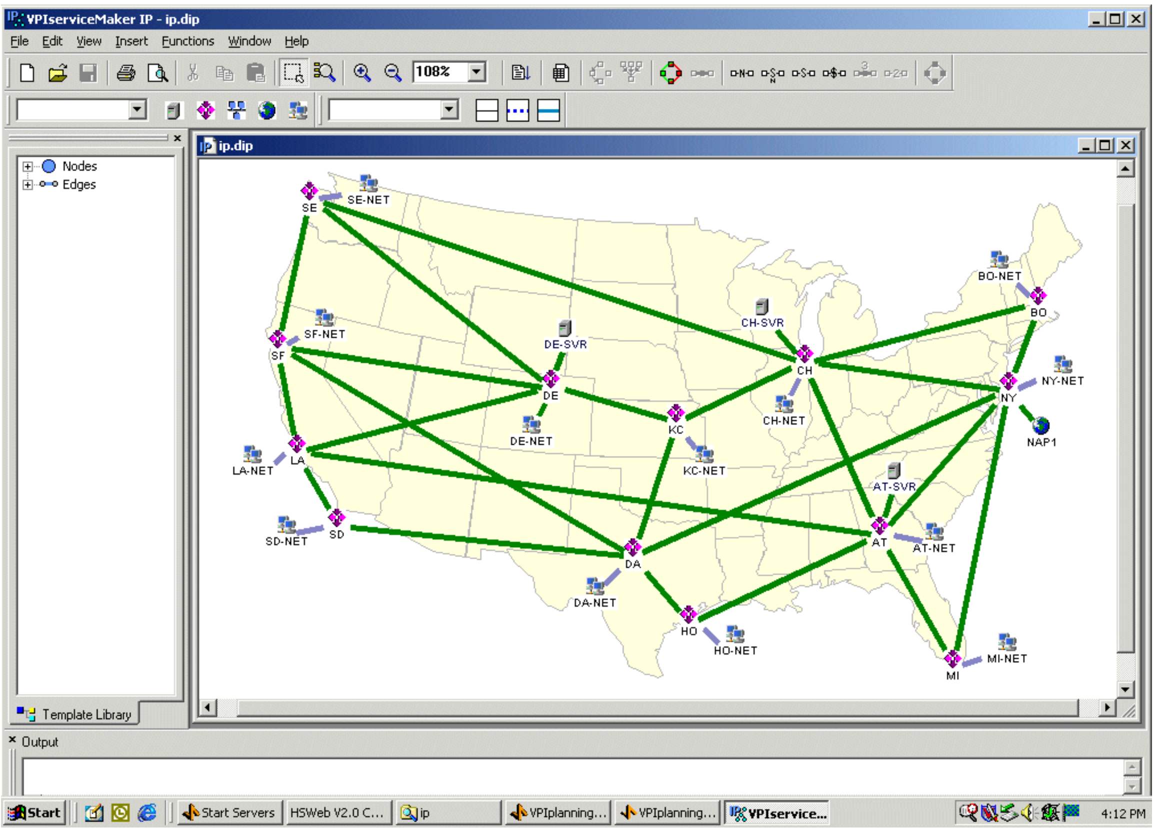


Figure 9 IP network topology

We now specify a traffic profile for each user population, each server, and each type of service. The only exception is VoIP traffic, whose traffic profile will be imported from VPIserviceMaker™Distribution.

The import procedure is similar to that of VPIserviceMaker™Switch; that is, we create an empty traffic matrix and use the import function. We then run analysis, routing, and effective bandwidth calculation. The resulting IP network dimensioning is shown in Figure 10.

Finally we export the IP network to VPItransportMaker™. A new VPItransportMaker™ project is created with IP network links being treated as a matrix of point-to-point bandwidth demands for the transport network design.

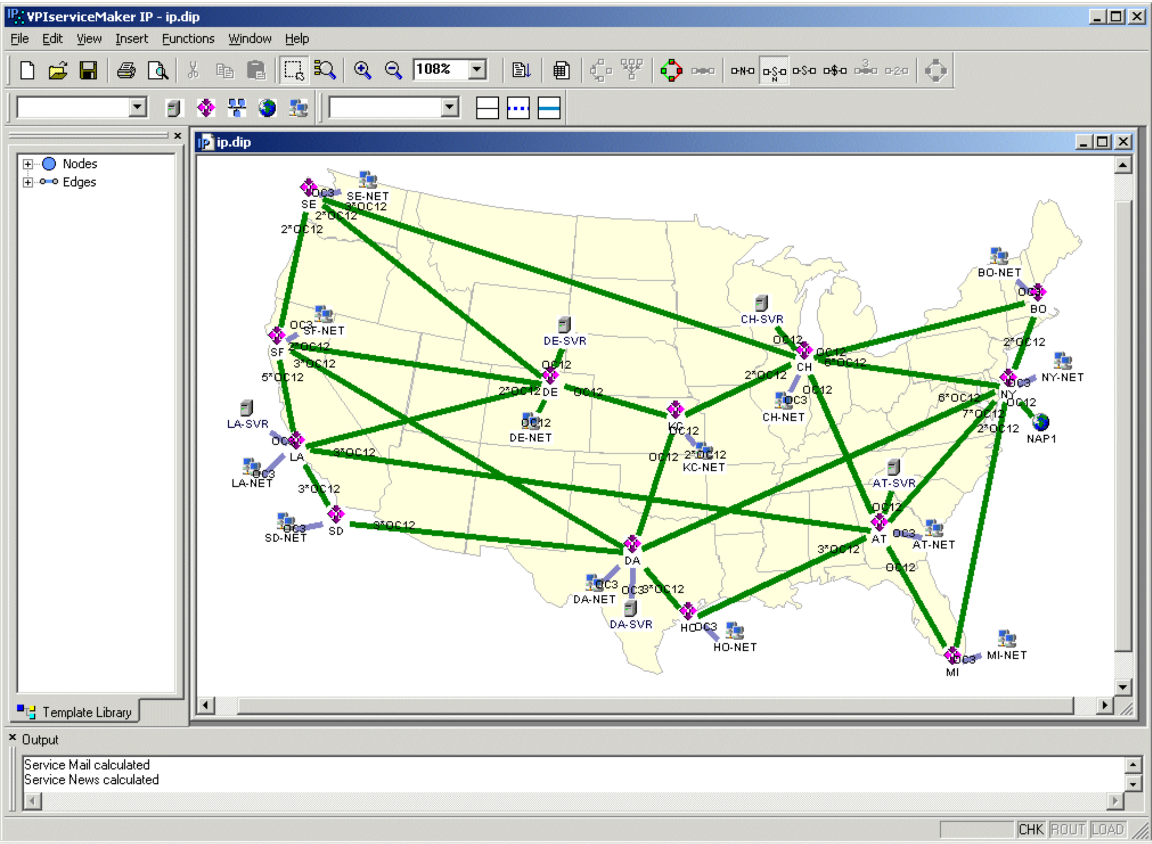


Figure 10 IP network dimensioning

VPIserviceMaker™ATM

Now we design the ATM network. We import the same network topology as before (see Figure 11). Our task is to model three known classes of services: Permanent Virtual Circuit (PVC) service, Metropolitan Area Network (MAN) data service, and Switched Voice service. The traffic matrix for each class is known. The fourth class of traffic is a constant bit rate service. Its traffic matrix is unknown, but it will be generated by VPIserviceMaker™Distribution.

Unlike voice traffic generation (where we need to specify user populations and average usages), for constant bit rate data traffic we only need to specify the total amount of terminating traffic at each node. We then choose, using VPIserviceMaker™Distribution, the single matrix method and the gravity model to produce a traffic distribution matrix. The traffic generated is then exported for use as the fourth traffic matrix in our ATM design.

After combining all traffic inputs, the aggregated traffic pattern for the ATM network design is as shown in Figure 12 .

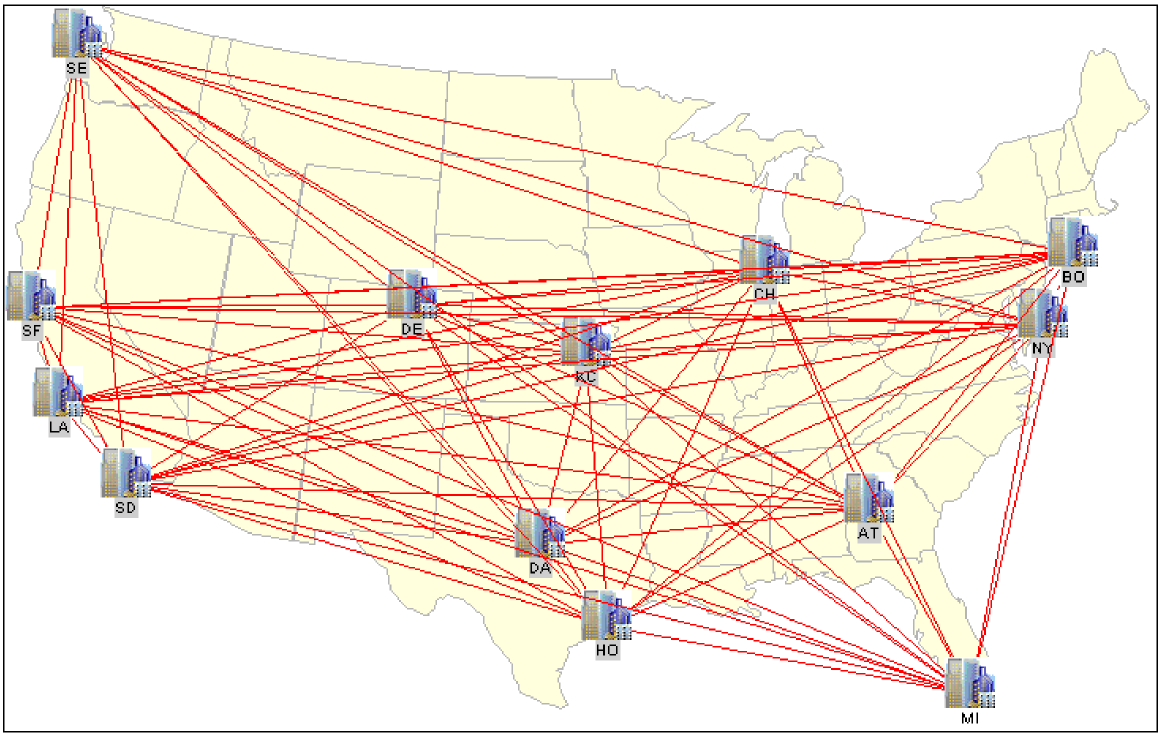


Figure 11 ATM traffic pattern

After a load optimized design, we arrive at a network with link capacities as shown in Figure 12.

Again we export the data to VPItransportMaker™ to create a VPItransportMaker™ project where the ATM network link sizes are grouped as one of the traffic matrices for transport network design.

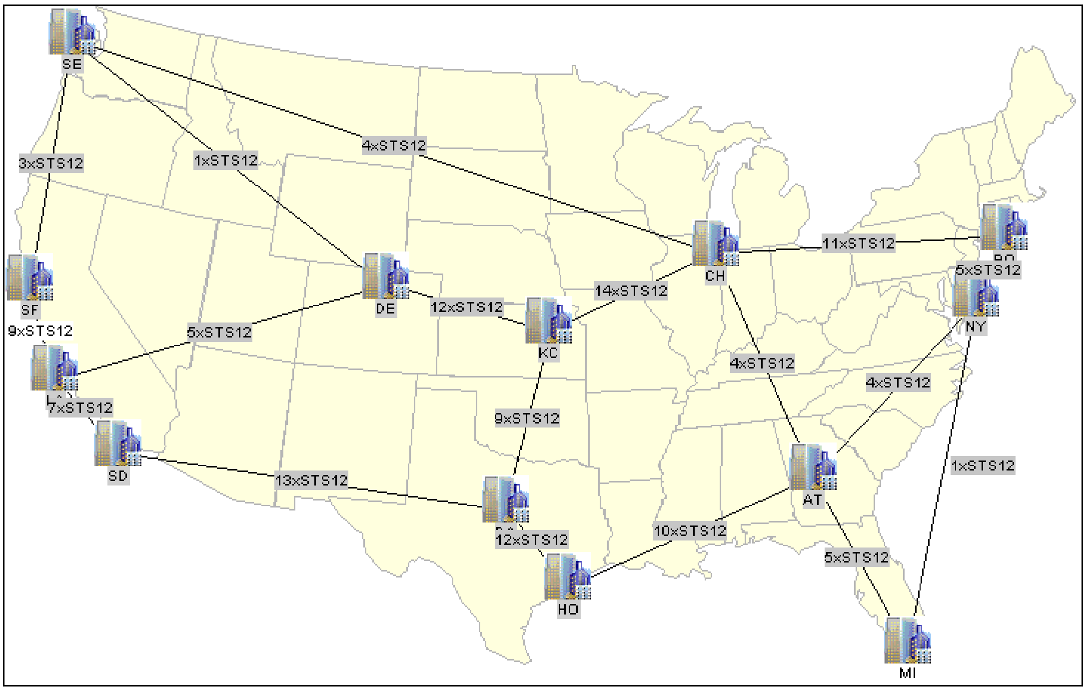


Figure 12 ATM network dimensioning

VPItransportMaker™

Now we are ready to design the transport network based on SONET over DWDM. We would like to provide protection at the SONET layer, and choose to design a SONET ring network.

We start with the VPItransportMaker™ project created by the VPIserviceMaker™ATM export. The fiber network topology is shown in Figure 13 . The ATM network bandwidth requirement is included as the first traffic matrix (with the name ATM_STM4).

The reason we do not start our transport design with the VPItransportMaker™ project created by IP export is that we added some new nodes—representing user groups and servers—to the IP network which we do not want to consider in our transport network design. But we can import the IP network links as traffic for VPItransportMaker™.

The nodes added during IP design are not recognized as VPItransportMaker™ network nodes and the traffic between them is automatically ignored. This is precisely what we want. Nonetheless, the IP network bandwidth requirement is represented as the second traffic matrix (called IP_STM4) imported into VPItransportMaker™ .

We also need to import the results of the switched voice network into VPItransportMaker™. To do so we import the VPIserviceMaker™ matrix. A new traffic matrix (called SWITCH) is now created with a DS0 bit rate.

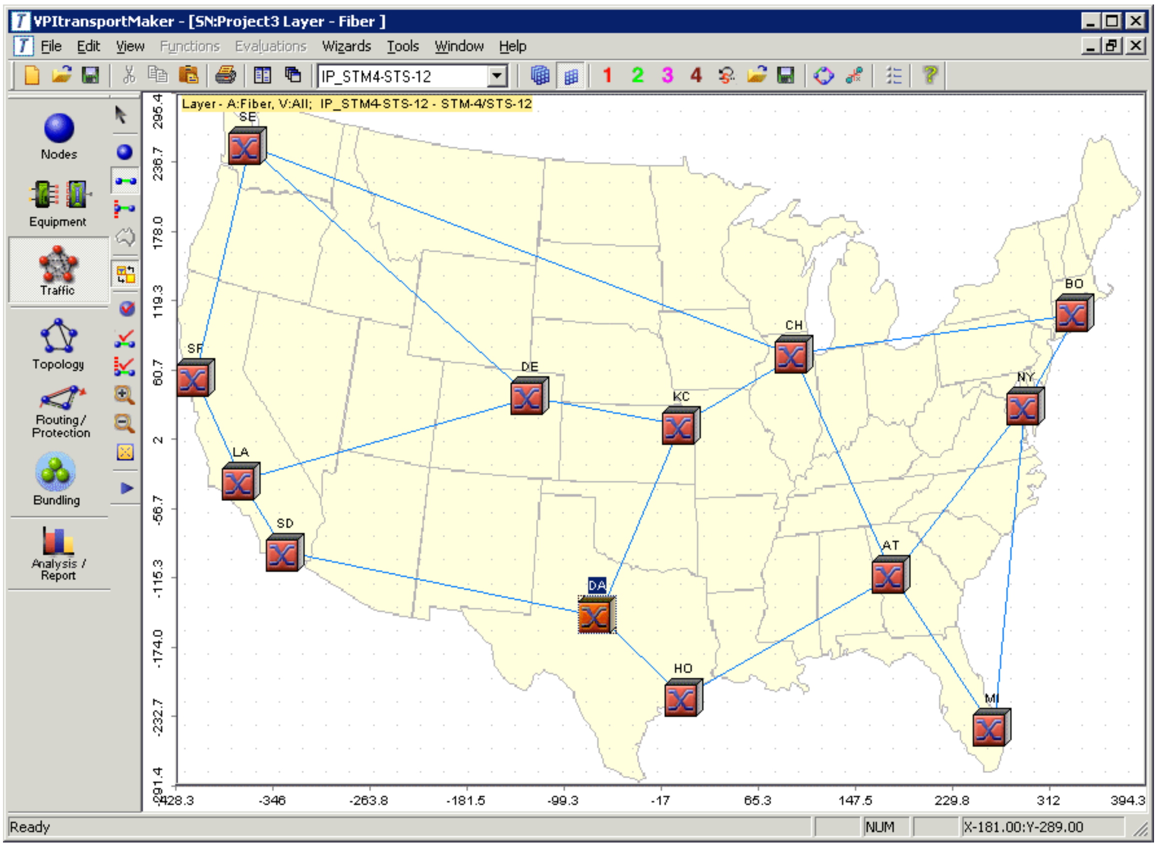


Figure 13 Physical network topology

We also need a fourth traffic matrix as an estimate of unknown leased line services. So we go back to VPIserviceMaker™Distribution and create a matrix of point-to-point OC-3 (STM-1) demands for the SONET network. For this purpose we only need to specify the total amount of terminating traffic at each node (in units of OC-3). A new traffic matrix named DISTRIBUTION with bit rate OC3 (STM1) is created.

We have now imported four traffic matrices: ATM, IP, SWITCH, and DISTRIBUTION. They are illustrated in Figures 14-17.

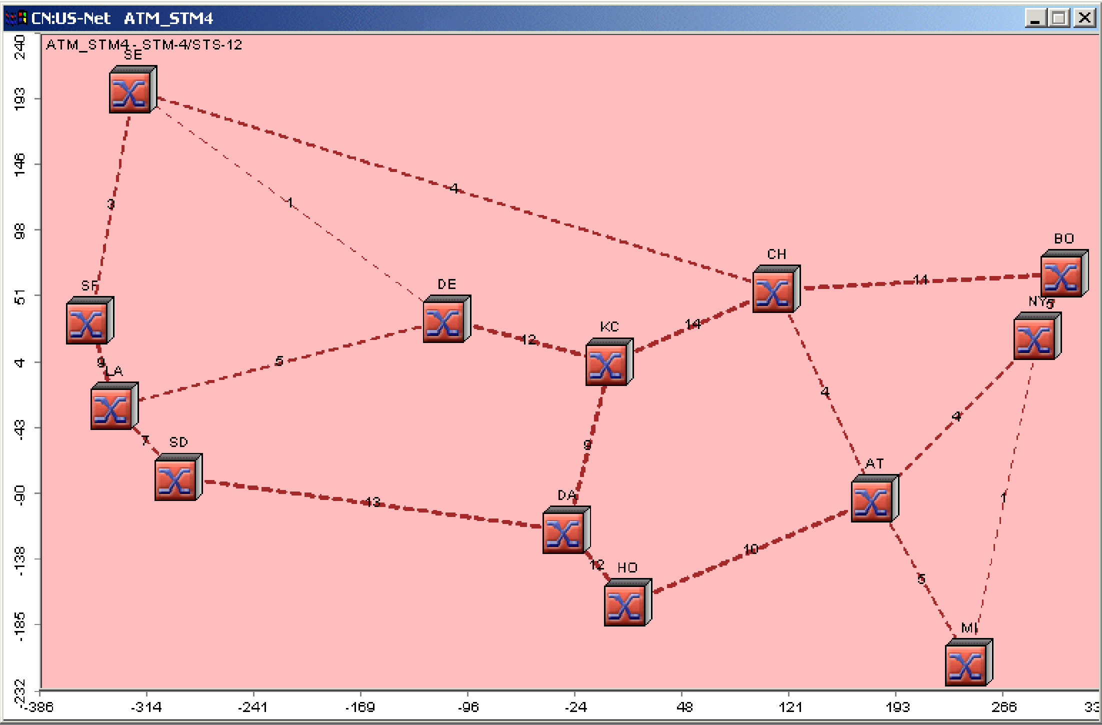


Figure 14 ATM bandwidth (STM4)

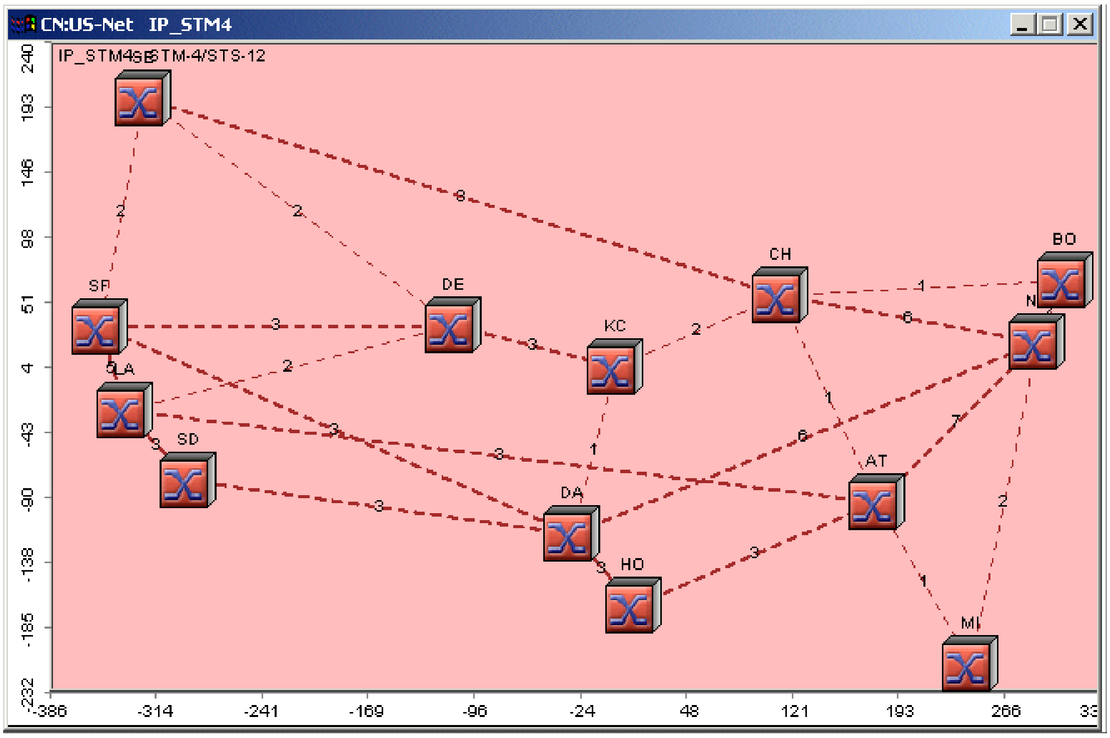


Figure 15 IP bandwidth (STM4)

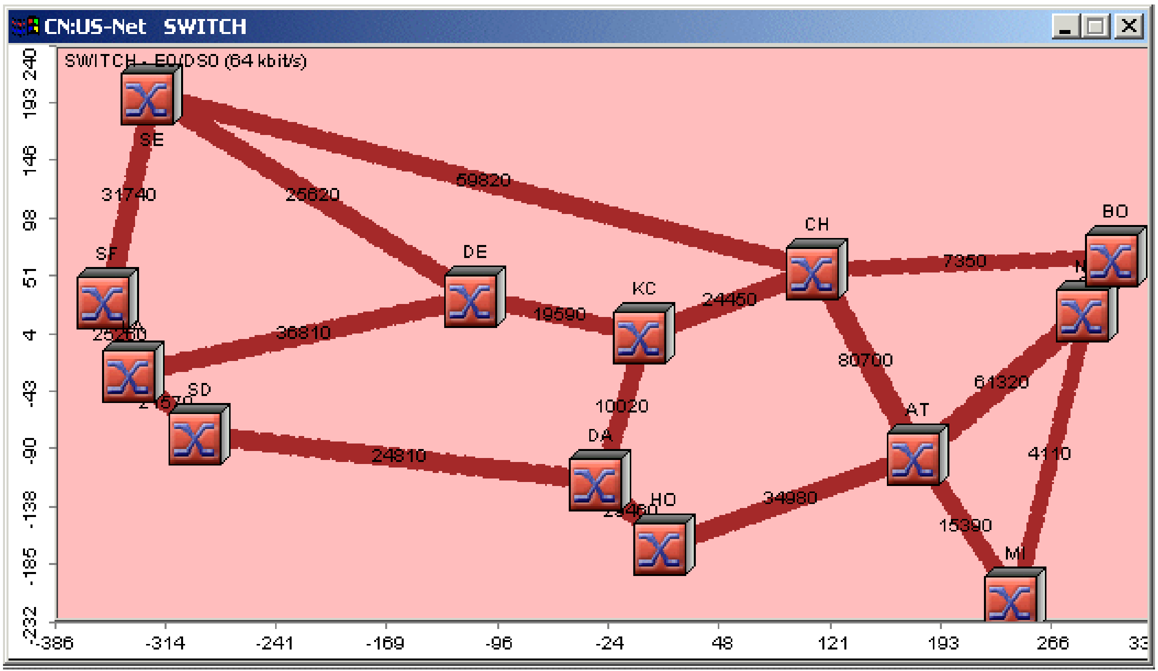


Figure 16 Switched voice bandwidth (DS0)

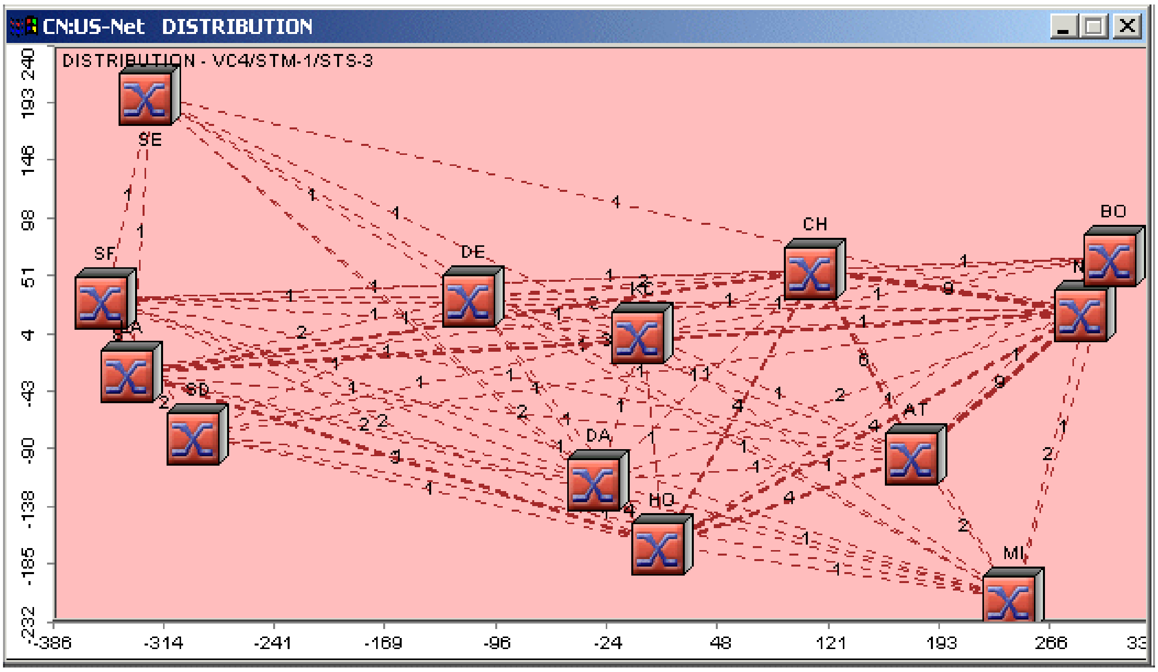


Figure 17 Distribution bandwidth (leased STM1 lines)

The SWITCH matrix from the switched voice network design is in units of DS0. DS0 cannot be handled by SONET add/drop multiplexers (ADMs) since the speed is too low. So we need to multiplex DS0s to higher-speed signals. We do this by first routing the traffic and then using the end-to-end bundling functionality in VPItransportMaker™ .

After bundling, the new SWITCH matrix is in STM-1 units (OC-3) and is renamed Switch_STM1 (as in Figure 18).

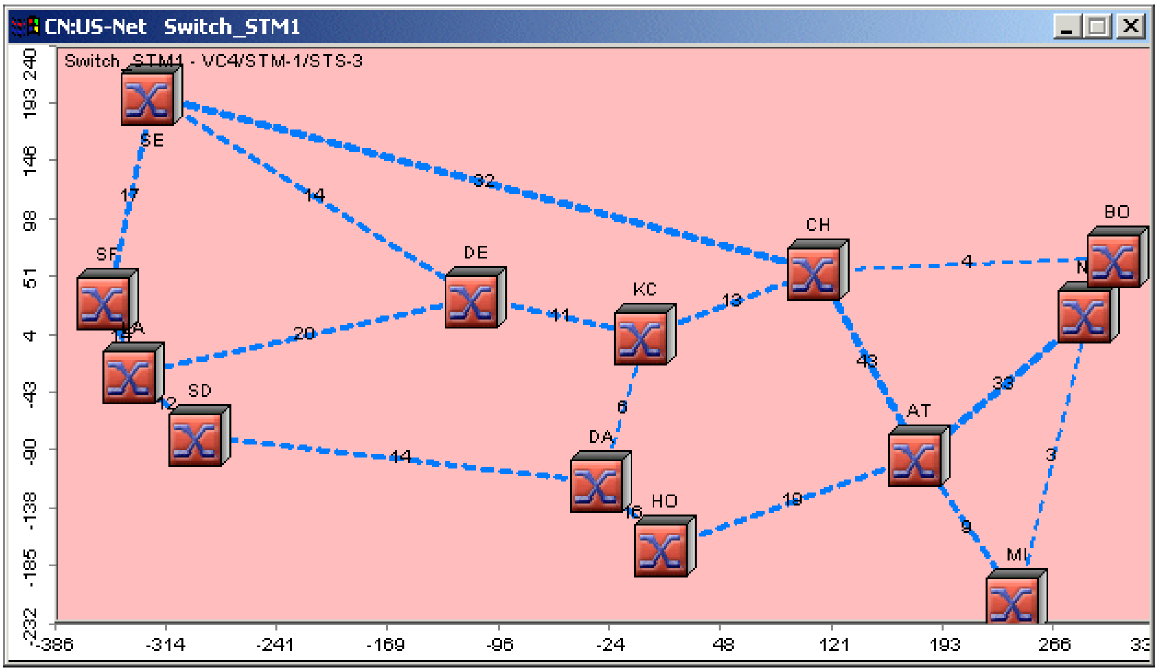


Figure 18 Switched voice bandwidth in STM1s

We are now ready to design the SONET and DWDM network. The basic steps are:

- Define a particular ring technology (such as a 4-fiber, STM-64, BLSR ring).
- Specify ring constraints (such a ADM add/drop capacity and granularity) and cost parameters.
- Specify general design parameters (such as enabling demand splitting, joint OMS optimization and cross connects for inter-ring traffic).
- Specify the DWDM system technology, which provides point-to-point wavelengths for the SONET ring network. You can specify such parameters as maximal number of wavelengths, optical amplifier spacing, regenerator spacing and cost.
- Run the ring designer to produce all the topological rings, modular rings, their interconnections, underlying DWDM systems, and traffic routing. (Here a modular ring refers to a physical SONET ring—as constituted by SONET ADMs—and a topological ring represents the topology of a modular ring. So multiple modular rings may traverse the same topological ring.)

Figure 19 shows the DWDM layer topology where each link is one or more DWDM systems. The number beside the link indicates the number of such systems. For this particular design, only one DWDM system is needed over each link. This is because each system provides 40 lambdas (OC-192) of capacity, which is more than sufficient for the size of our traffic.

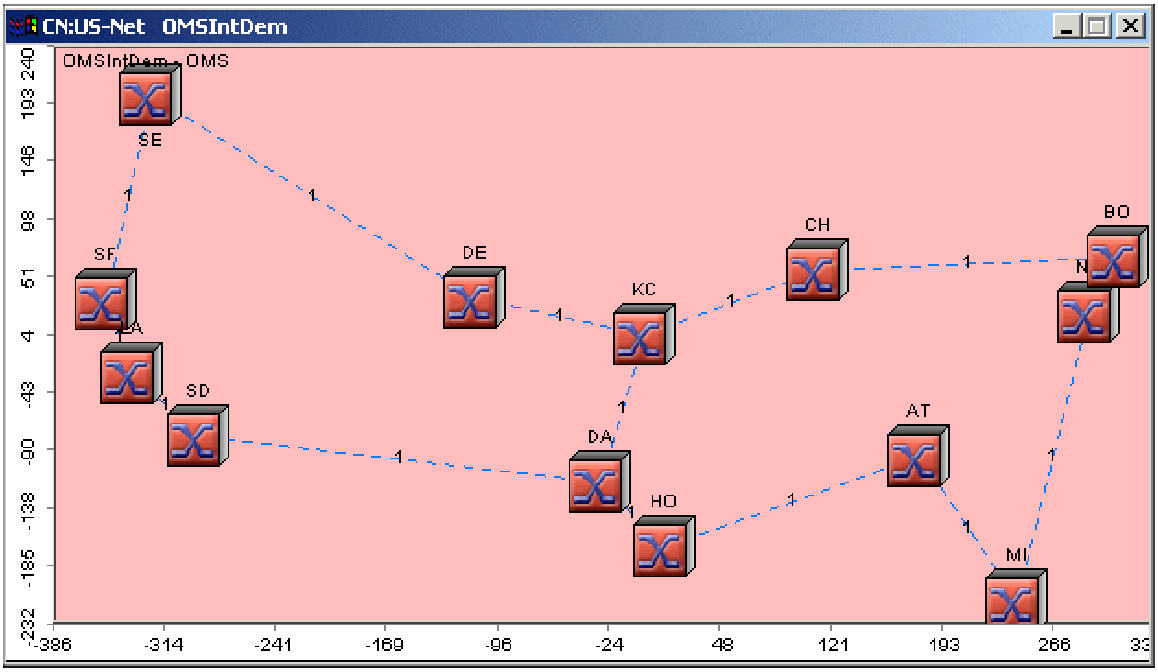


Figure 19 DWDM layer deployment

In the SONET layer, two topological rings and five modular rings are created. The two topological rings are routed exactly along the two loops of the DWDM topology shown in Figure 19, one on the left and one on the right. There are three modular rings over the right loop, and two modular rings over the left loop.

Summary of Integrated Planning Procedure

In this design exercise we have shown how the VPIlifecycleManager™ tools can be used to undertake integrated network planning across multiple layers. We have demonstrated a smooth end-to-end planning procedure, showing how the various tools can inter-work, and how VPIserviceMaker™Distribution can help estimate traffic where traffic data is not available.

The data flows in this exercise have been one-way: from traffic distribution to service layer design to transport layer design. It is possible to have a feedback process from the transport layer to service layers. For example, by looking at the spare capacity of the transport layer, you can estimate how to expand a service network. You could then adjust transport bandwidths allocated to the various service layer networks. You could also conduct what-if analyses of the potential impact of service layer changes on the transport layer.

Next Generation Products from VPIsystems

So far we have described the integrated planning solution supported by VPIsystems' products supported on the VPIlifecycleManager™ platform. In this section we present the next generation of VPIsystems products and solutions. The first of these is VPI's Network Lifecycle management solution which provides an integrated and end-to-end solution for a telecommunications operator's network planning and engineering process. The second is VPInetworkConfigurator™ a product aimed at automating the equipment vendor's pre-sales process. We provide a brief overview of these products and the motivation behind them, below.

VPIsystems Network Lifecycle Management Solution

Introduction

Revenue is lost when customer demands cannot be fulfilled by the available network infrastructure. Carriers usually overbuild to resolve this problem, which results in needless waste of capital resources, while not necessarily completely eliminating hot spots in the network. For example, switch trunk planners typically estimate growth forecasts with built-in spare. These estimates are then passed on to the transport planners who put various trunk estimates together to develop a transport plan that also includes some spare, resulting in spare-on-spare in the overall network. Yet operations still experience hot spots in the network with unexpected customer demand, resulting in lost revenue.

Unfortunately for most carriers, the processes required to identify idle capacity and bottlenecks, to analyze today's traffic for preparing forecasts, and to optimize and re-configure the embedded infrastructure for minimizing new investments, are fragmented, slow and inefficient, resulting in under-utilized infrastructure, and needless capital expense. Network Lifecycle Management (NLM) is an Engineering Support System (ESS) concept pioneered by VPIsystems in order to solve these problems and facilitate the automation of a Telecommunications Operator's Network Planning and Engineering process. NLM allows network operators to carry more traffic with less equipment and less people involved. By bringing all of the various planning and engineering process elements into one single generic process, NLM will ensure interaction between the functions with a single overall view of the process from end to end.

While telecommunications Engineering & Operations have traditionally been supported by a variety of software systems, the introduction of software systems to aid the integrated network Planning and Engineering processes is relatively new. Figure 20 helps understand how NLM relates to other well established Operations Support Systems (OSSs) on the telecommunications software landscape. Probably the best-known OSS systems that interface to NLM are the "Service Fulfillment", "Service Assurance", and "Activation" systems, which automate processes concerned with managing, monitoring, and utilizing the existing network resources. These systems mainly facilitate management of already-deployed network resources and usually act on logical and physical operational inventory databases. These types of OSS are designed to support many thousands of short transactions everyday and hence can be categorized as "Transaction Support Systems[5]".

In sharp contrast to the above, the network planning and engineering processes require a new type of software system, called Engineering Support Systems. Network Planning & Engineering provides the intelligence on how to build and extend service provider networks. It involves creation of new logical and physical resources to accommodate the forecasted

[5] On the other hand, a different type of OSS, known as Business Support Systems (BSS), manages customer-facing processes such as "Service Order Management", "Customer Relationship Management", and "Billing"[5]. BSS systems act on operational product and service databases and facilitate administration, sales, marketing and billing of existing network resources.

traffic within the constraints imposed by the existing logical and physical network infrastructure. The aim is to design, develop and deploy a CapEx-optimized network infrastructure that meets service requirements. For a given set of forecasted demands, the key network planning and engineering activities consist of checking available network capacity to see if it can accommodate the projected growth, generating network new-build plans where necessary, and producing detailed engineering designs for network upgrade.

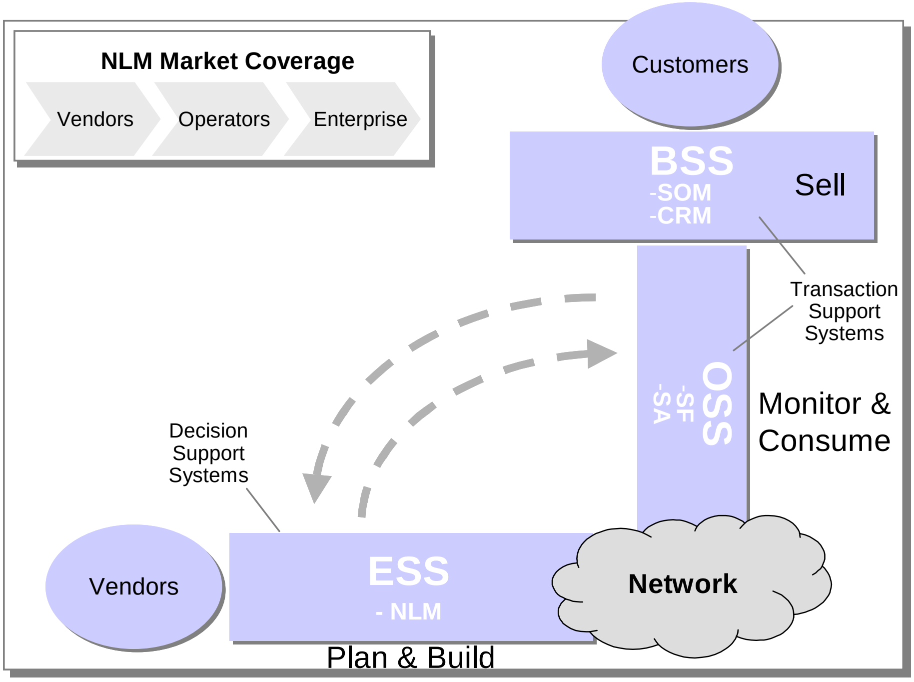


Figure 20: Telecommunications Software Landscape

In contrast to the OSS, Engineering Support Systems involve relatively few transactions, but each transaction may launch computation intensive algorithms operating on a large amount of data. Consequently, Engineering Support Systems have to be designed like “Decision Support Systems” so as to support few computation-intensive transactions with large volumes of data.

Key Requirements of NLM

A key requirement to NLM is to provide the necessary “Network Design” and “Configuration Modules” that can create feasible designs which ensure that when the equipment are deployed as per the design guidelines, the network will function as designed with near zero fall-outs. For this to happen, the designs have to be driven by the physical and logical constraints of the equipment. Capturing all details of not only the legacy equipment but also of latest technologies for producing optimum and feasible designs requires creation of a comprehensive equipment library.

Another key requirement is that NLM should be capable of automating detailed planning and engineering processes in carriers. VPI’s NLM solution achieves this by providing the necessary “Workflow Automation Modules” that streamline the design, configuration, and deployment of network infrastructure. Driven by forecast network demand, the process of checking available capacity, generating network new-build and producing detailed engineering designs for network upgrades constitute key components of NLM.

Aside from the necessary design, configuration, and workflow automation modules, another key component of NLM is the underlying database, on which all modules operate. Network planning & engineering business processes require detailed data on planning, availability, usage and state of various products, services, and resources in the network. This data is usually kept in various inventory systems. NLM should take the as-built network data from the operational inventory systems (logical and physical inventory) and combine them with detailed equipment data to build the Planning & Engineering Database (PED). The PED is the underlying database of NLM that keeps an off-line historical model of the operational databases, and hence needs to be periodically synchronized. As output, NLM issues and delivers "new build plans" and "Work Orders" for deployment of new-build network and re-optimization of the as-built network and initiates updates to the corresponding inventory systems about the changes made to the network.
In summary, NLM enables operators to deliver better services, more quickly and consistently, and make better use of already-deployed network resources. With large networks and many complex processes, the CapEx and OpEx impact on operators is potentially huge.

VPIsystems Pre-sales Process Automation Solution

A serious problem faced by equipment vendors has to do with the currently prevalent pre-sales proposal generation process. Sales opportunities are lost when product experts are overloaded and when the sales and support team lacks consistent, up to date product knowledge. This results in proposal responses that are non-competitive or sub optimal, lengthened sales cycles that jeopardize business opportunities, more costly solutions, and design errors creating additional post-sale expenses.

The VPI solution to this problem is VPInetworkConfigurator™ an expert software program that increases the productivity, responsiveness and competitiveness of the sales organization. It enables sales support and proposal personnel to rapidly provide responses to sales opportunities and proposal requests, which are optimized to lowest cost and are 100% accurate. Productivity improvements of 2X for sales support personnel are typical. Benchmarking against manual designs created by product experts typically shows 80% time savings and 20% lower cost of completed designs when compared to existing techniques while totally eliminating product configuration errors.

This solution is achieved by embedding VPI's world class planning and engineering expertise in an easy to use web enabled software that allows the vendor's global sales force access to the vendor's scarce product and design expertise. In addition to the core design engines developed by VPI, the software consists of a customizable product library module that can be populated by product experts, and a customizable design wizard that encapsulates design rules developed by vendor experts with a thorough knowledge of the product capabilities. Once the customer network data is captured by the sales person, the pre-built product library and design wizard take over and generate optimized designs that meet the customer needs. A key innovation of this software is the integration of the planning, link engineering and equipment configuration steps which enables the sales person to produce designs that are 100% compatible with the products chosen, and to generate orderable Bill Of materials (BOM). A schematic depiction of this process is shown in Figure 21 below.

Figure 21:: Integrated planning, link engineering and equipment configuration

Value Proposition: Extensive benchmarking has shown that for complex network configurations, the overall design improvements result in network designs that are 25% - 30% less expensive than those generated by previous methods. This enables the sales force to be much more competitive, putting tremendous pressure on competitors without any lowering of product prices. New product training for sales staff has been minimized, and field sales personnel can immediately take advantage of newly introduced products and design rules to address existing customer opportunities, via software updates over the Intranet. Additionally, it has enabled the vendor to avoid expanding the sales support team to handle increased sales opportunities, thereby providing significant annual OpEx savings.

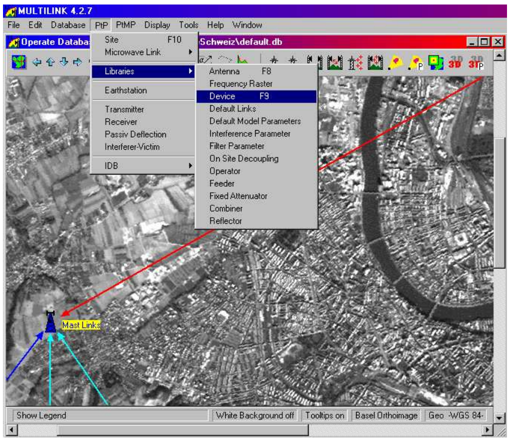
Conclusion: How to insure that the best solution is always proposed to every customer has been a classic problem of sales organizations. VPInetworkConfigurator is the first expert system that has solved this problem for vendors of network equipment. In the hands of the sales organization, it turns everyone in sales support into product experts with little or no training. It selects the proper network topology, equipment family and product configuration to meet unique customer requirements based on product design rules, prices and network constraints. It is 100% accurate down to the individual line card level. It can be programmed to learn new design techniques from ongoing competitive experiences in the field. It enables new product additions and design rules to be implemented instantaneously across the worldwide sales organization. It pays for itself both through OpEx savings and as quickly as with one win of a major competitive customer RFQ

A1.5. LStelcom

MULTIlink

Introduction

MULTIlink is a proven microwave link-planning tool from LS telcom AG, which provides the network planner with state-of-the-art and easy-to-use functions during the entire design process. MULTIlink is a complete solution for fast microwave link engineering and designing of PMP/WLL/LMDS networks. It can be used for planning and optimizing single links (e.g. path loss, coverage and availability calculations) as well as for doing network-wide analysis (e.g. interference calculation, channel assignment).



<u>Features and Highlights</u>

- Powerful Database containing Sites, Links, Receivers, Transmitters, Antennas, Devices, ITU/ETSI Frequency Plans.
- Advanced Map Handling / Network Viewing
- Interactive Link Engineering Desktop Environment
- Technologies: FDMA and TDMA
- Passive / Reflector Back-to-Back Link Profile
- Clearance of Fresnel Zone
- Reflection Points Determination
- Excellent Profile Handling
- Link Analysis
- Free Space Loss, Atmospheric Absorption, Clear Air fading, Obstruction Loss, Rain Attenuation Calculation according to ITU-R Recommendations
- Dispersive Fade Margin, Outage due to Dispersive Fade Margin
- Area-wide Field Strength Coverage Prediction
- Availability Calculations considering Rain and Multipath based on worst month and average annual Statistics
- Interference Analysis Calculation for both Long and Short Term
- Microwave Link Reports Creation

Special Modules

- Point-to-Multipoint Links / LMDS
- Satellite-Earth-Station Coordination
- Short-wave Point-to-Point

<u>Implemented ITU Recommendations and Regulations</u>

- ITU-R P.530-9
- ITU-R P.452-10
- ITU-R P.676 Atmospheric absorption

- ITU-R P.837 Attenuation due to rain
- ITU-R P.526 Diffraction mechanism (knife edge)

Satellite Earth-Station Berlin: (optionally)
- Ap S7 (Sat-ES)
- Ap 28 (Sat-ES)

Short Wave: (optionally)
- ITU Rec. P.533
- ITU Rec. P.372
- ITU Rec. BS.705

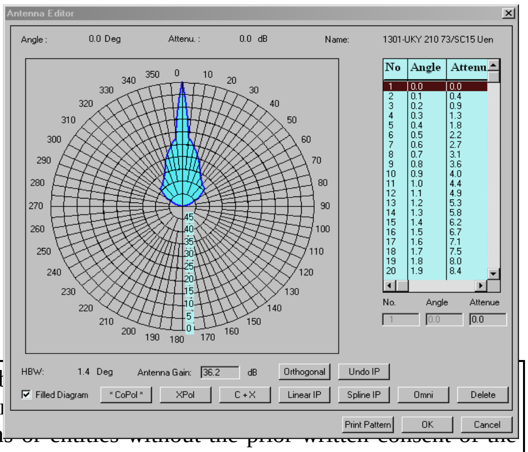
Frequency Range
The MULTIlink software package includes propagation models used in the typical frequency range from 700 MHz to 40 GHz. Optional modules covering 200 -700 MHz are also available. In addition, a shortwave module is provided to cover the frequency range from 150 kHz to 30 MHz.

Database

MULTIlink is able to store all used databases (e.g. terrain maps, technical data) either locally or on a network server. Based on Client/Server structure several users can be linked to a central database. Each user has a local working database for planning purposes. Easy copy and update processes are built in. Data integrity is established using record locking procedures.

Terrain Map Database
MULTIlink supports both raster and vector maps. Calculation results can be visualized as vectors (such as field strength or interference contours) or as raster data (for example area calculations). Raster or vector calculation results can be overlaid on background maps with an adjustable transparency factor.

Technical and Subscriber Data
All data is kept in the LS-MULTIBASE database or optionally in an Oracle database. Data can be entered either manually in spreadsheets, editors or using the standard ASCII interfaces. Customer tailored import solutions can be provided if request.
Antennas / Devices
The database comprises important technical information, like: antennas, transmitter, receiver, feeders, combiner or reflectors
Antenna pattern can be either imported or created manually by entering the angular discrimination values in a table. An antenna editor is provided for modifying the antenna pattern.
The horizontal and cross-polar radiation patterns of the antenna are essential parameters for interference analyses.
NSMA-Format is supported. Besides, many other kind of antenna formats are also supported because



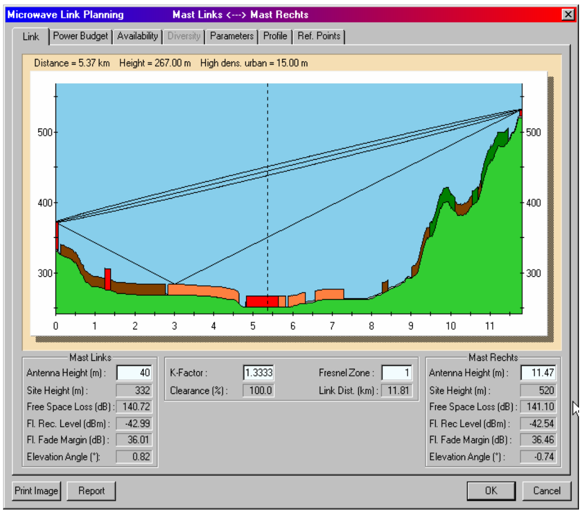
a generic ASCII interface is used. Customer specific input formats can also be implemented if request.

Frequency Raster
The definition of frequencies or channels is based on frequency plans described in ETSI Standards or ITU Recommendations.
"Manual frequency plans" are supported as well. This type of frequency raster allows the definition of arbitrary frequencies.

Planning

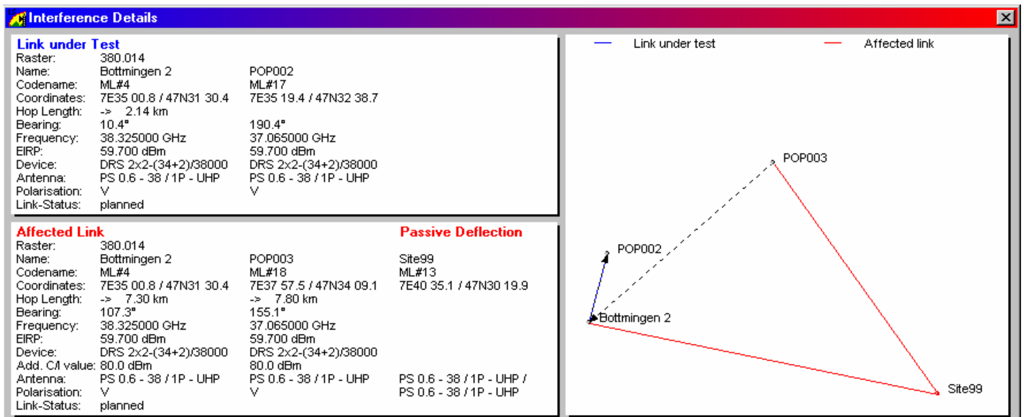
Microwave Link Planning

Using MULTIlink a network planner is able to design his microwave link directly in the interactive Link Planning Window. All necessary calculations, like Fresnel zone, path loss, power budget, link availability, diversity, reflection points, etc, could be done in this window.
The calculations incorporate the effects of atmospheric absorption at higher frequencies and the effects of different rain zones.
By a simple mouse click a microwave link report can be generated, containing all important data of the new designed microwave link.



Interference Analysis

MULTIlink allows detailed interference analysis of microwave networks. Short and long-term interference analysis is performed in accordance with the ITU-R 452 propagation model, using terrain and clutter data. The interference analysis can be performed to Point-to-Point links and also to the PMP scenario.



	Transmitter	Receiver	hop/km	bear.	ILL/dBm	ILS/dBm	DF/dB	AD/dB	PiRXL/dBm	PiRXS/dBm	C/IL/dB	C/IS/dB	TD/dB	Add. IF
1	POP003 2	*Bottmingen 2 1	6.1	216.9°	-32.1	-30.9	0.0	120.8	-152.9	-151.7	74.9	120.1	0.0	X
2	Bottmingen 2 1	*POP002 2	2.1	10.4°	-25.0	-24.6	0.0	47.0	-72.0	-71.6	-6.0	39.7	30.6	

MULTIlink generates a detailed interference report, which can be printed or saved.

Channel Assignment

The channel assignment function is used to find usable frequencies for a selected link in a specific frequency raster. The Channel Assignment Algorithm performs an automatic interference analysis for the whole frequency band and creates a list of recommended channels (channels with minimum interference). For each channel, a “quality measure” is calculated.

Modules

Basic Calculations

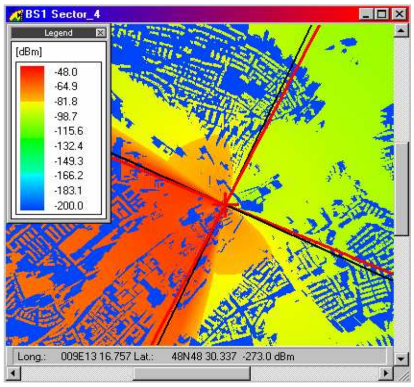
Line-of-Sight (LOS) Check

For each site or point on a map, a LOS Check can be executed.

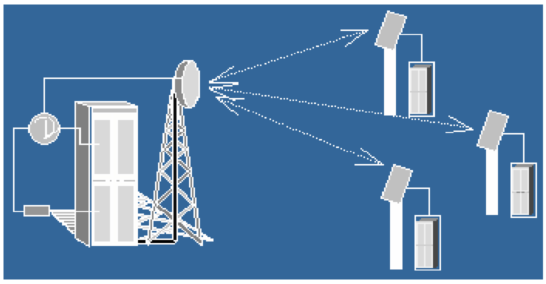
The height of antenna above ground level is specified as parameters. Results are overlaid on a map.

Coverage Prediction

For coverage prediction several propagation modules are available. The propagation parameters used by the different modules (e.g. rain zone, k-factor, etc.) are configurable. Coverage prediction results are the basis of all further network investigations.

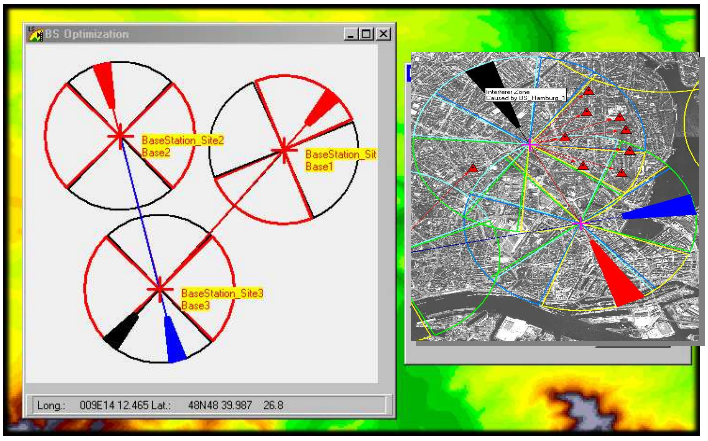


Point-to-Multipoint (PMP)



MULTIlink_PMP models a PMP / WLL network in a hierarchical structure using several network elements. The “Base Station” consists of several sectors. To each sector several subscribers (terminals) can be linked. Each sector covers a defined area (depending on the used antenna).

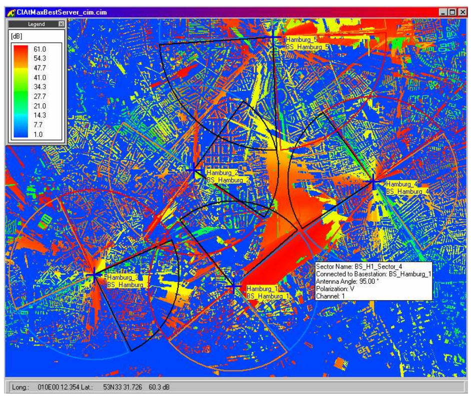
Fast Base Station / Sector Optimization



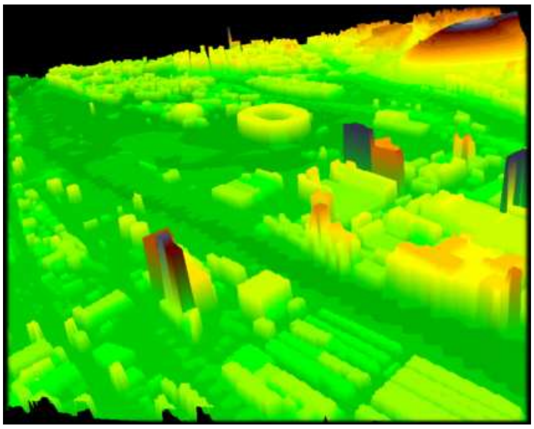
The PMP module in MULTIlink includes a fast sector optimization algorithm, which is based on free space propagation and geometrical assumptions. This function detects possible interferer zones in a fast manner and the user can eliminate the interferer zones by rotating base stations.

PMP Network Processor

The PMP network processor allows performing detailed network analysis to design and optimise PMP/WLL networks. Typical network analysis results are: Maximum Field Strength, Maximum Server, Best Server, C/I, etc.



3D View



The digital terrain model even overlaid with calculation results can be displayed three-dimensionally.

xG-Planner

Introduction

xG-Planner is a mobile network planning tool from LS telcom AG, which is applicable to 2G (GSM, TETRA), 2.5G (GPRS, EDGE), and 3G (UMTS). It is intended to cover the whole range of mobile network planning aspects including:
early project studies
license application or offer phases
first rollout
detailed network phase planning including site management
network operation
network optimization
It is a state of the art product with advantages relative to other existing tools on the market through its main functions to:
Provide fast and high quality coverage planning (several propagation and refraction algorithms, optional support for field strength calculations due to the Vienna agreement)
Hierarchical database (Network, Project, Project status and Cell level) to structure large projects
Storing of all results in a result database, filtering & context sensitive regaining of information
Spreadsheet visualization of database contents
Graphical and/or tabular management of all network elements
User role management to administrate user privileges
Configurable menu actions, access restrictions to certain user groups
produce optimum frequency plans
supply quality of service (QoS) predictions
xG-Planner provides the possibility to plan pure macro- and microcellular networks, as well as multilayer networks consisting of a combination of micro-, macrocells, extended cells or concentric cells.
Furthermore xG-Planner supports the planning of different mobile services like GSM900, GSM1800, GSM1900 (PCS), DECT, TETRA and UMTS. This is due to the fact, that network technologies as well as calculation algorithms can be allocated on a per cell basis. The tool is also open for keeping up with future system technologies.

Basic xG-Planner Module

<u>Graphical User Interface (GUI)</u>

The development of xG-Planner was focused on a modern, ergonomic, user-friendly software handling, which allows for convenient and comfortable user operation.
Some of xG-Planner GUI features are:
main windows with the complete tree of features (Menu bars, toolbar buttons)
context-sensitive menus & tool tips
overview cartographic window with fast navigation possibility
Simultaneous multiple map windows for different views

user configurable tool bar with icons providing fast access to functions which are often needed by the user; key shortcuts for fast access to the menu items; ergonomic arrangement of features similar to the general Windows standard.
Context sensitive mouse menus on objects displayed on the map window, direct editing of items in the plot or getting quick information on certain pixels; status line with actual information and cursor readouts.
definition of user individual layouts (table layouts)
Restoring of all user settings, reopening of last open windows
All cartographic windows use coupled cursor readout

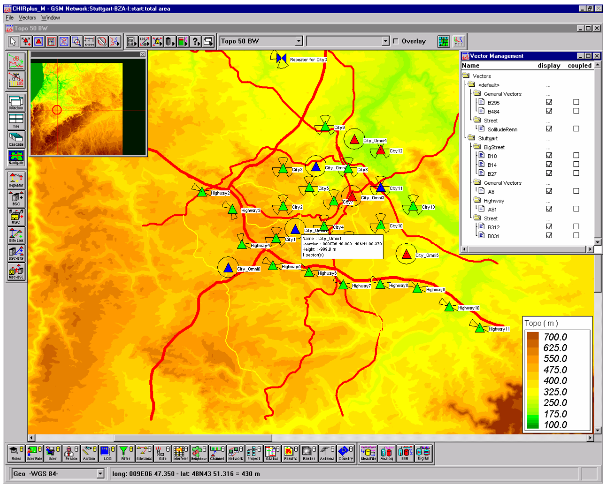


Fig. 1.6.1: Topographical map with sites and streets displayed. A small legend and overview window are also displayed

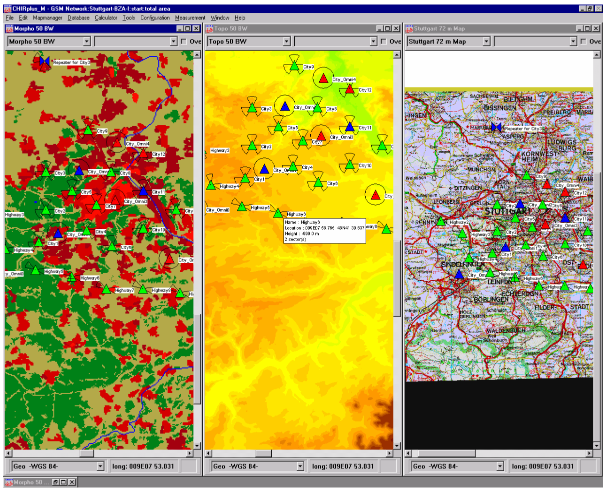


Fig. 1.6.2: Multiple cartographic views showing different types of information

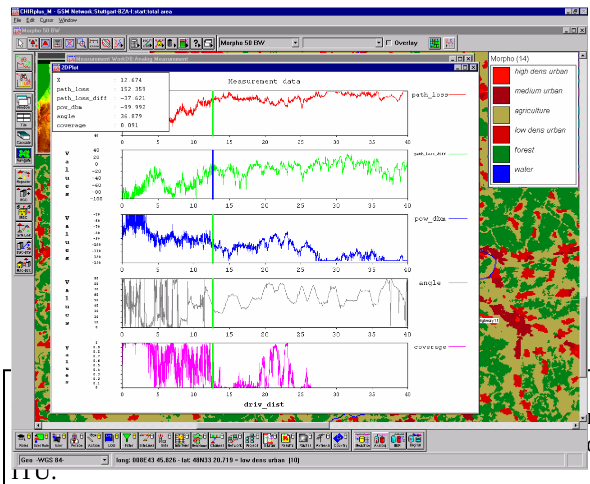


Fig. 1.6.3: Measurements in 2D plot representation

Database Model

A suitable radio network planning tool has to handle a huge amount of input and output data. Hence, the database structure is very important.
xG-Planner stores all the planning data in a fully relational database system.

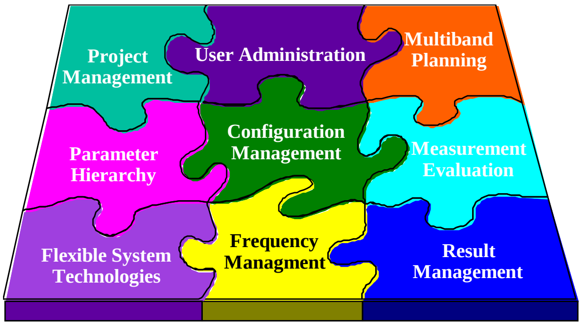


Fig. 1.6.4: Database Model

Going more into detail, this means that the powerful optimised relational database structure provides:
highest data integrity and consistency
hierarchical data organization
flexibility through:
user profile assignment
user role (user mode) assignment
customizable and transparent views on all data which is related directly to the planning
Further comforts are:
editing the database items using the following methods:
editing via forms directly from the displayed objects
editing via report windows (e.g. in tables) using powerful spreadsheet functions
automatic update of information in databases during interactive changes in plot and vice versa
log book functionality

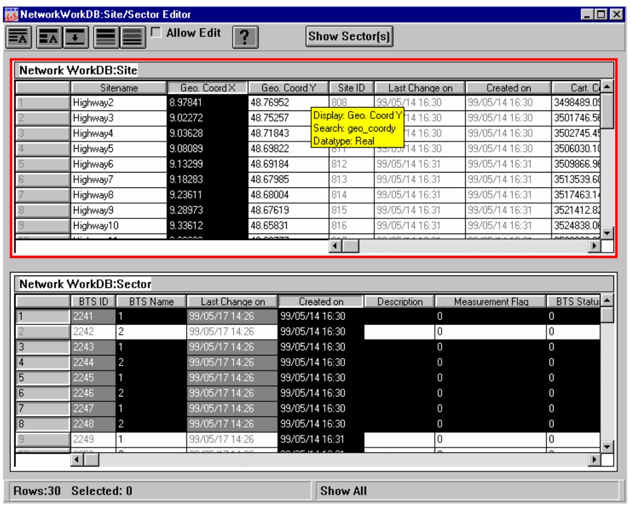


Fig. 1.6.5: Site/sector spreadsheet

Coverage Prediction Module

The computations within this module are based on the cell design and source databases (topography, morpho structure, traffic) and yield the field strength prediction and its derivations on a cell and network basis. Furthermore different cell shape calculations are performed.
In particular the following tasks can be performed:
site calculations like LOS calculation point to point and on an area
cell power calculation (downlink)
median power calculation for cells and repeaters (choice among several propagation models like free space propagation, flat earth model, LS model, enhanced Okumura Hata, Walfish Ikegami, etc)
signal delay difference and signal level difference (for repeaters)
coverage probability calculated from the power result
network-wide power calculations
cell shape calculations

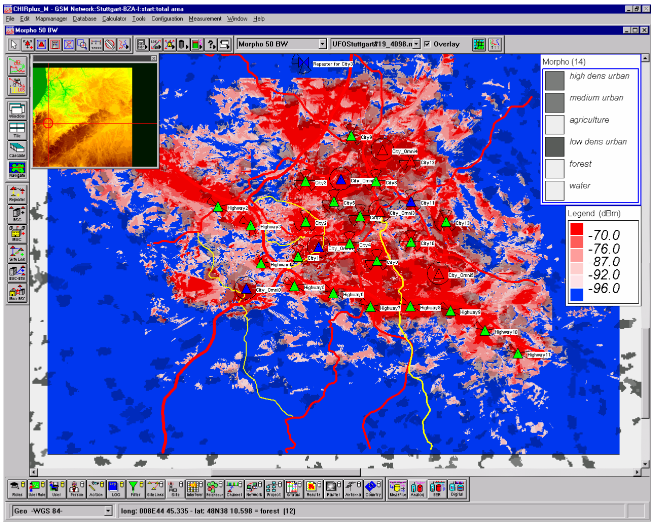


Fig. 1.6.6: Networkwide power overlaid on background morpho map

Microcell Module

xG-Planner provides the full amount of features necessary to plan microcellular networks. The ray tracing model searches for different propagation paths and sums up the contribution of different rays for each receiver pixel. Reflections on floors, walls, roofs as well as diffraction on horizontal / vertical edges are computed. The number of reflections and diffractions taken into account can be selected individually by the network planner beside a wide variety of options which are available within the 3D ray tracing model.

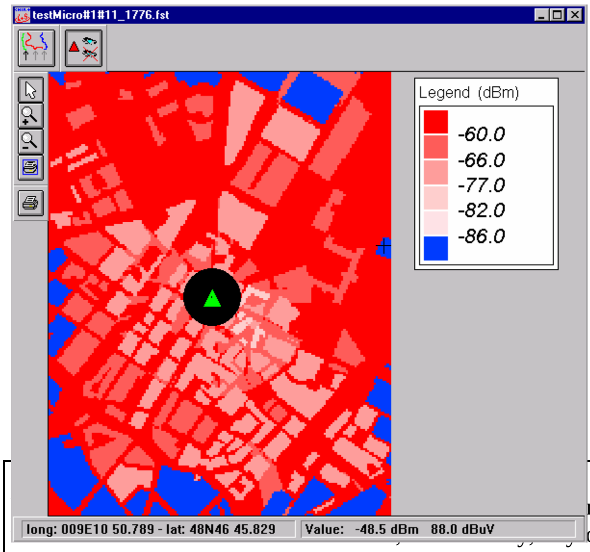


Fig. 1.6.7: Field Strength Prediction for Microcells

Ray Launching

In contrary to the ray tracing algorithm which searches for different propagation paths (ray classes) receiver-pixel-wise, the ray launching approach considers the rays sent by the transmitter with a fixed angle increment and follows them until a defined termination criteria is fulfilled (e.g. maximum number of reflections). Furthermore, during ray launching only reflected rays are considered, whose contribution to the field strength are summed up receiver-pixel-wise.

Field Strength Prediction with COST-231 Walfish Ikegami Model

Besides the above mentioned deterministic models a fast empirical model is provided: the COST231 Walfish Ikegami Model. Basically this model can be utilised in combination with the 3D Ray Tracing Mode for speeding up calculations.
This model can be calibrated by measurements.

Indoor Prediction

Indoor prediction can be applied in combination with the fully deterministic ray tracing model as well as with the COST 231 Walfish Ikegami model. It can be performed during standard prediction or separately during post processing
The prediction model can be calibrated by adding an offset to the predicted field strength, by modification of the diffraction model parameters and by setting the breakpoint.

Frequency Planning Module

The automatic frequency assignment is based on: the channel requirement for each BTS, the matrix of mutual interference probabilities for co-channel or adjacent channel, the total number of frequencies in the allocated bandwidth and the given constraints for frequency planning.
Types of calculations performed during frequency planning:
traffic calculation: calculation of required frequency channels out of the traffic density database, the assignment probability of the required number of traffic and signaling channels
setup of channel separation matrix as a preparatory step for the automatic frequency assignment
automatic frequency assignment: choice between Box algorithm (I and II) and Simulated Annealing as frequency allocation algorithms; assignment of actual channel numbers considering the structure of the available frequency band
optionally available: usage of frequency groups and frequency hopping algorithm

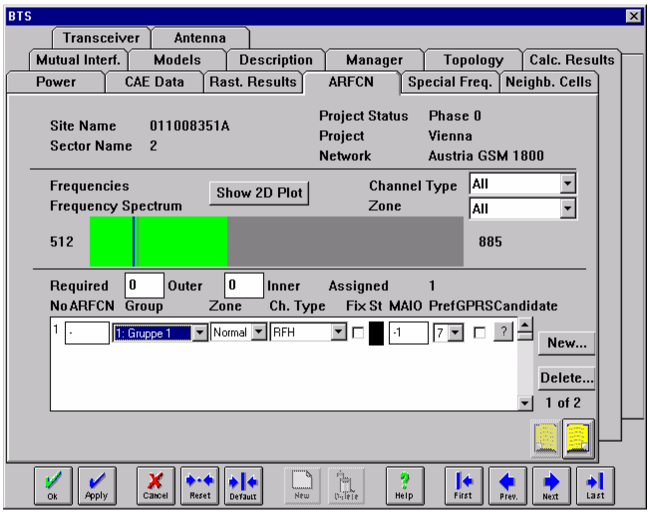


Fig. 1.6.8: Frequency Planning

Furthermore various additional frequency constraints can be considered for frequency planning:

comfortable individual cell definition of constraints for the automatic frequency assignment via graphical Interface: fixed frequencies; co-site separation, co-cell separation, free definable frequency bands; allowed, not allowed frequencies (e.g. at country borders, at borders of neighboring networks etc.)

The GUI offers different options for:

manual editing of the channel separation matrix

manual assignment of frequencies with the help of an interference ranking list (best possible choice of available channels)

fixing previously assigned frequencies against accidental tempering

statistical evaluation of frequency distribution and reuse

QoS Prediction Module

Aim is to provide a quality of service prediction of the network design based on interference coverage holes and BER according to the frequency plan. Its evaluation checks system performance and defines optimization possibilities.

The analysis includes the following types of calculations:

the traffic capacity analysis

powerful cell performance calculations:

final cumulative interference calculation based on the actual frequency assignment; output of results are C/I values in dB and probabilities %

quality of service prediction achieved by an outage calculation considering the combined effects of interference and coverage holes, optionally weighted by traffic density;

output of results as probabilities (BER and FER calculations in preparation)

network-wide performance calculations: final interference and outage calculations

averaging of interference and outage probability over serving cell area or over network

BSS Parameter Module

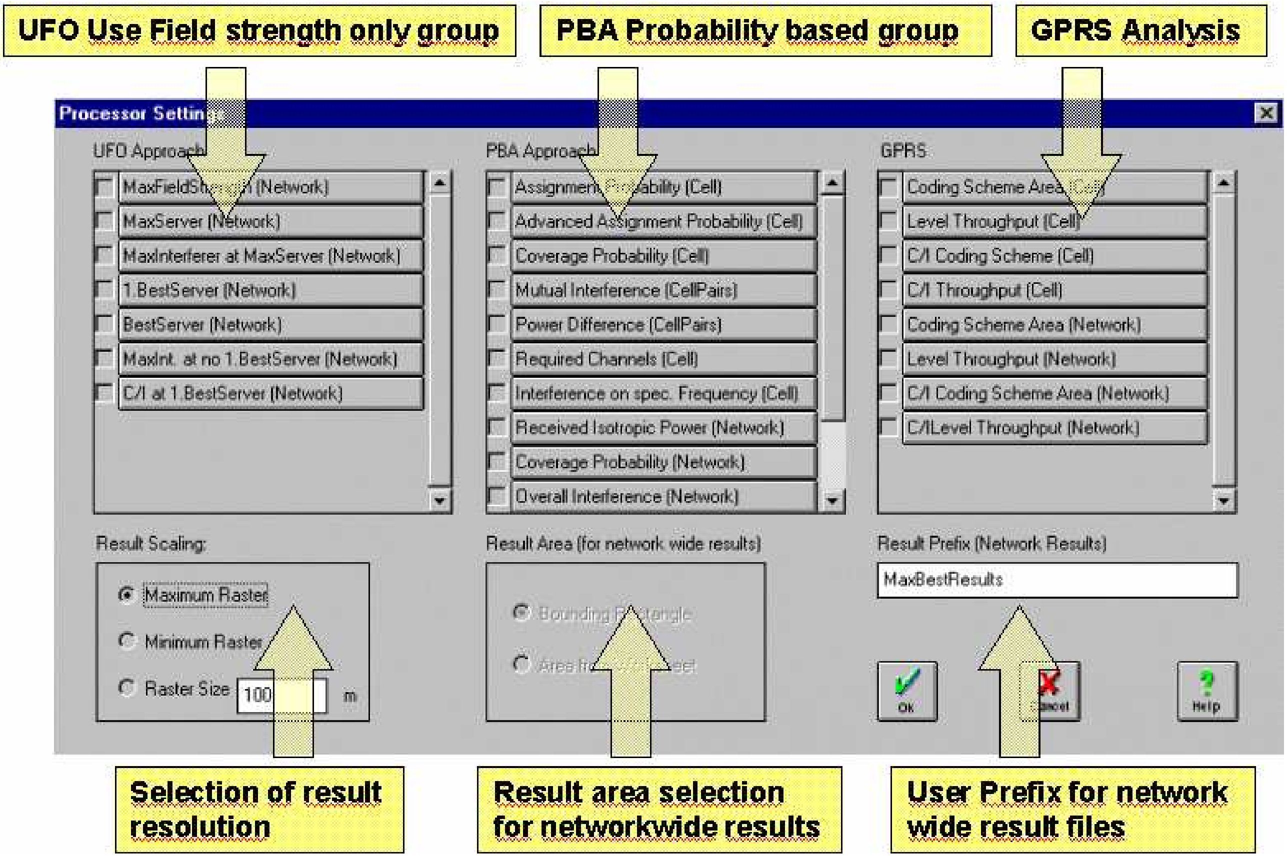
The module is responsible for creation of system dependent BSS parameters, which is the first step towards implementing the real network. The BSS parameter files are loaded in the OMC-R via a corresponding interface. This interface and the parameters that are passed to it are strongly dependent on the operator's hardware.
Types of calculations in the BSS parameter generation module:
location areas (LAC) allocation and Cl allocation
calculation of neighborhood relationships/neighbor cell lists
geometric
best server based
interference based
link planning parameters (traffic, number of channels, link distance,...)

Measurement Evaluation Module

Major purpose of the measurement evaluation is the visualization of the real network behavior and the comparison (calibration) with network planning requirements / prediction. The visualization of measurements is done by mapping them together with cartographic background information and statistical analysis.
Analogue and digital measurements can be imported into the measurement database and evaluated.
The following tasks to be performed at different stages of radio network:
propagation model calibration: adaptation of parameters used in the field strength prediction model to the real conditions in the planning environment; comparison of measured and predicted field strength under variation of the parameter(s) of the OH model to be calibrated
site verification: verification of possible transceiver sites regarding the fulfilment of propagation requirements
installation error detection: detailed analysis for problem detection after installation of the network (e.g. missing HO-relationships, ping-pong-HO etc.)
acceptance test: analysis of the huge amount of test drives in the service area of the network

GPRS Module

xG-Planner optionally also supports 2.5G functionality like GPRS. For simplicity it is included in the tool's central network processor as a third function collection:



The GPRS part of the processor thus collects the GPRS based. Currently, the following algorithms are already included:
Calculation of GPRS coding scheme areas CS1..CS4 (best case, based on field strength level)
Throughput of data in kbps/s per cell
Coding scheme areas network wide
Throughput network wide
Real CS and throughput calculations based on interference situation

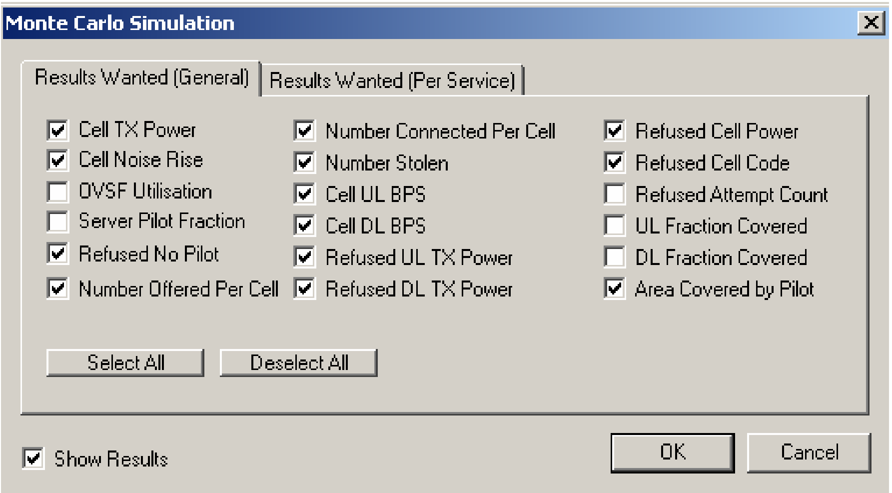
UMTS Module

xG-Planner addresses the UMTS planning part with different modules, which are intended to serve the planner for his different needs during the project phases. These represent different levels of detail:
A propagation based planning module for initial bid-planning estimations
A Monte Carlo based planning module for regular network roll-out
Optionally, there is also a third module available, intended for academic hotspot evaluation purposes:
A UMTS Real Time simulator for chip-level analysis of deterministic single UMTS users.

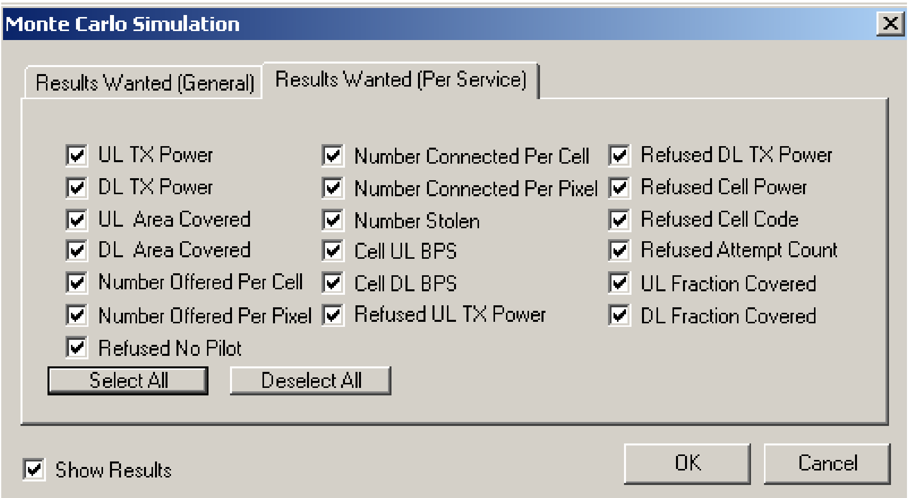
Monte-Carlo Simulation

In the following we focus on Monte Carlo (MC) simulation and its corresponding approach to plan UMTS networks as Monte Carlo simulation is the standard method for UMTS network

planning. For UMTS, downlink and uplink must be analyzed, but as the mobiles are (by their very definition) not static in position, it requires a simulation (instead of a calculation as was in GSM).
The following screenshots give an overview of the result types that are calculated in the framework of xG-Planner' Monte Carlo Simulation. The first figure shows the available cell related results the second figure reflects the "per service" results (that is, one plot for each defined service per MC simulation) that are calculated for each service.



Cell related (general) results of the Monte Carlo Simulation



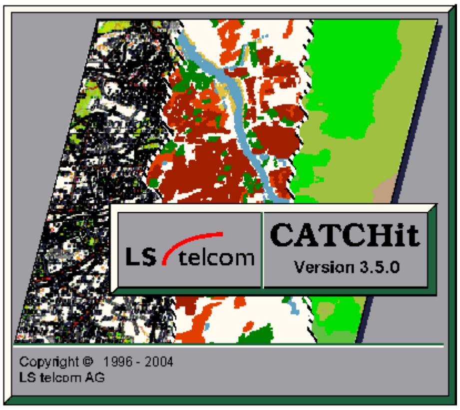
Service related results of the Monte Carlo Simulation

Scrambling Code Planning

xG-Planner's Scrambling Code Planning Engine is responsible for the allocation of cell-specific downlink scrambling codes within a UMTS (3GPP WCDMA variety) compliant terrestrial network. The Scrambling Code Planning Engine is capable of handling an arbitrary number of ranges of allowed codes. These codes are assigned in a manner such that there is minimal interference between cells with the same scrambling code.

CATCHit - Terrain Database Management Tool

Network planning, especially performed with RNP tools, requires the collection and management of numerous data. RNP planning tools generally don't offer more than the standard terrain database functionality like:
Open data layers
Create simple polygons
Change color, size and, line type of vector layers
Overlay vector and raster maps
These basic functionalities are not sufficient to fulfill the requirements for efficient management of terrain databases for RNP purposes.
CATCHit, the digital terrain database management tool of LS telcom, is the perfect assistant for all tasks coming up around digital terrain databases. Generally the functions of CATCHit can be divided into 3 main categories.
Generation of digital terrain data information
Conversion of existing digital terrain data information
Maintenance of existing digital terrain data information



Digital Terrain Data Generation

CATCHit is offering the following functionalities for the creation of new terrain information.

<u>Geo-referencing of scanned paper maps</u>

Paper maps often show a coordinate system grid for measuring distances and positions. After scanning these paper maps, this information is only visible for the user but not accessible for the GIS application. In order to combine the scanned map with the correct coordinate system usable in GIS applications, the map has to be geo-referenced. At places where the real world coordinates are known, reference points have to be placed. These reference points combine the visual scanned map with real world coordinates. Eight to twelve reference points have to be placed all over the map.

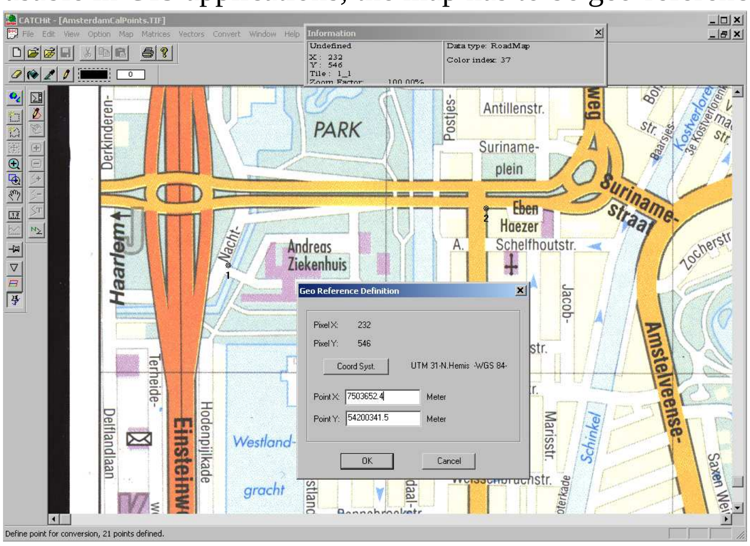


Figure 1.6.9: Placing a reference point and specifying coordinate system and coordinates

CATCHit allows using different coordinate systems for each reference point. This is very useful for the geo-referencing of city maps, which do not have a regular coordinate system grid included. The reference points can be taken from GPS receivers, where the coordinates are given in Longitude – Latitude (WGS84). Additional points can be written out of 1:20,000 or 1:50,000 scale paper maps, which usually use a metric coordinate system like UTM and a local Datum. CATCHit will automatically transform all the reference points into the coordinate system that will be used as coordinate system for the scanned map.

Vector tools

Vectors plus attributes are an important way to store geographic information. Most of the time the vector information is taken out of scanned paper maps. The extraction process is called digitization.
CATCHit serves the following digitization tools for vector creation and handling:
Creation of points, lines, poly-lines and closed polygons, with or without attributes
Move point or line
Add new points / remove point
Continue the digitization at the end of an existing poly-line
Break a poly-line or polygon into two parts
Combine two poly-lines to one seamless poly-line
Open / Closed poly-/ polygon conversion
Change of attribute
Change of line color, width and style
Import of and export to standard vector file formats

Automatic vector to raster conversion

Terrain Data information can be stored in vector or in raster (grid) format. Most RNP applications are grid based tools, because of the faster access to the geographic information at a certain point. Therefore it is often important to convert vector information into raster data. CATCHit serves an automatic process for the vector to raster conversion.

Interpolation of DTM's using a ray tracing algorithm

A Digital Terrain Model (DTM) in raster format stores the height above sea level as information. It is the most important terrain data layer for RNP analysis. Still 90% of the DTM's are created out of paper maps, using digitized contour lines and elevation points as source. These contour lines and elevation points are stored in vector format. The height between the contours has to be filled using interpolation methods. CATCHit is using a ray-tracing algorithm for the interpolation of DTM's. Its special source data access ensures the creation of a seamless DTM without needing much RAM, even if the target area covers several map sheets.

Demographic Data Processor

Demographic data, like total population, population divided by male/female, or number of income is very useful in RNP's daily work. Such kind of data is typically available on the market in vector format. The demographic information is always related to some areas like zip- code or district area. These areas cover not only built up areas but also contain forest, open land or water. If the demographic information were equally distributed over the entire zip code or district area, the digital model would simulate that people are living in the forest, lake or on agricultural land. This simulation doesn't represent the real situation.

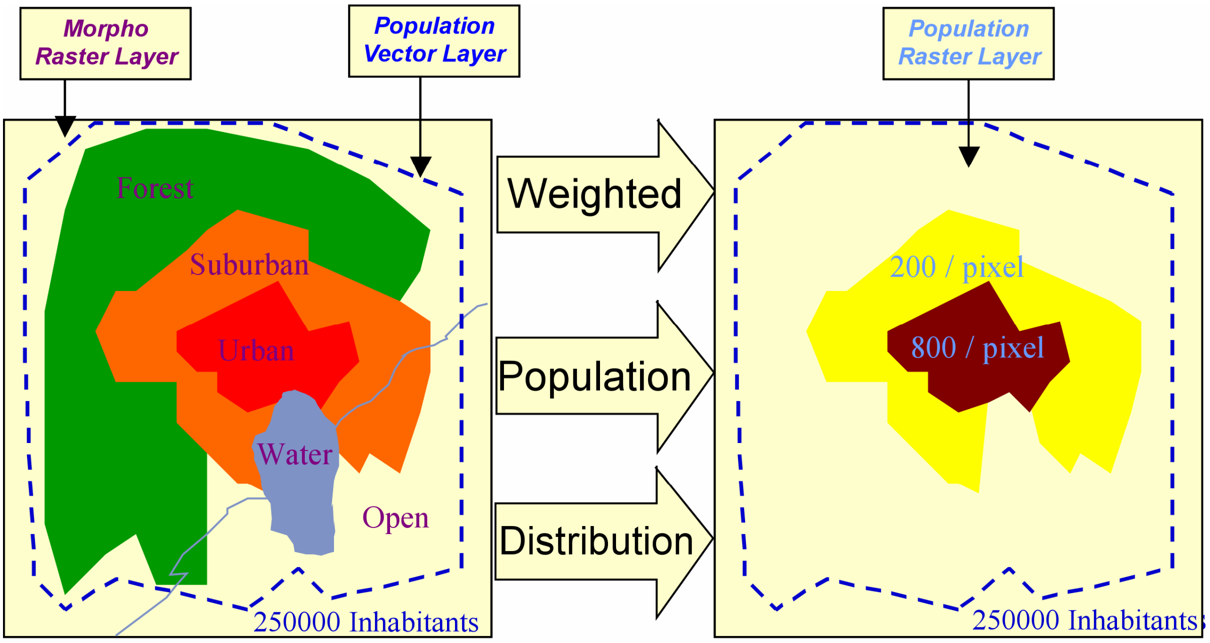


Figure 1.6.10: Weighted distribution of demographic information in populated clutter areas

CATCHit offers a special processor for the distribution of the demographic information that comes closer to the reality. The processor uses a clutter map for the detection the populated areas. The demographic information will be distributed only in these areas. Additional weighting coefficients for each clutter class allow the creation of different scenarios, like daytime and nighttime.

Terrain Data Conversion

The "standard file format" did not exist, there are several important file formats on the market. Therefore it is important to be able to read this "standard" file formats and convert them into the file format, which can be used by your RNP applications.

Import / Export

CATCHit supports several "standard" file formats for vector and for raster data.
Some supported vector formats are:

- LS telcom ASCII
- LS telcom Binary
- MapInfo TAB
- MapInfo MIF
- Planet

Some supported raster formats are:

- ESRI ASCII Grid

- ESRI ASCII XYZ
- ESRI BIL
- LS telcom Binary Grid
- Planet
- TIF plus TFW or TAB file for coordinate system information

Coordinate System Transformation

Besides the file format conversion, terrain data often has to be transformed to a specific coordinate system. Most countries have their own coordinate system, which is limited by the area of this specific country, International operating data suppliers usually use global coordinate systems for the terrain data like UTM- or geographical projection based on the WGS84 Datum (Mathematical model of the earth).

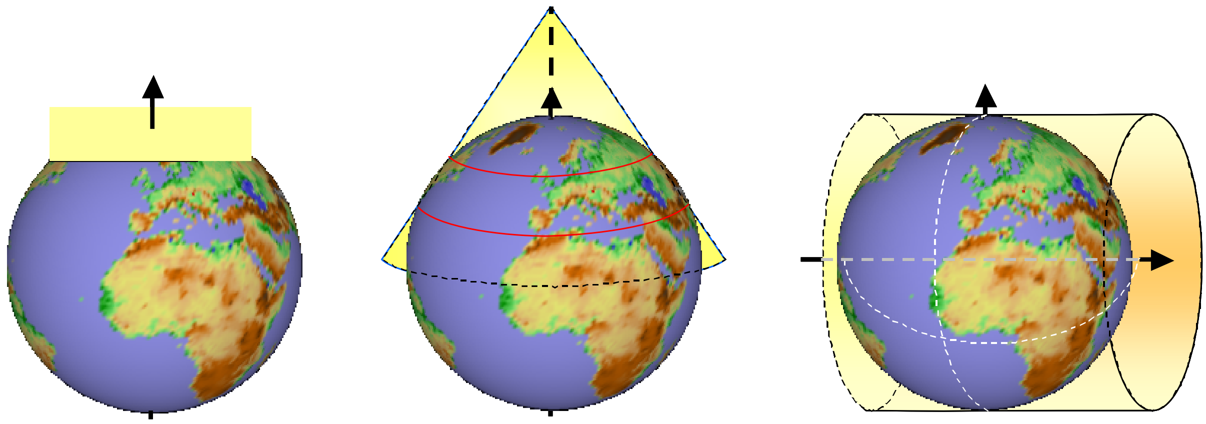


Figure 1.6.11: The main projection categories

A coordinate system transformation is needed in order to store all terrain data layer in one and the same coordinate system. The recent version of CATCHit supports over 600 different local and global coordinate systems and mare than 300 different Datums.

Terrain Puzzler

The file format conversion and coordinate system transformation of huge terrain databases is very time consuming, if one process is done after the other. Therefore CATCHit offers the powerful module “TerrainPuzzler”, which is able to perform both processes in one step. Four items are characterizing the main tasks of the TerrainPuzzler

1. File format conversion
2. Coordinate system transformation
3. Merging of different terrain data files and source
4. Change of resolution

These four tasks can be carried out in one step. Sequential read and write allows to handle files that are even bigger than 1 GB. Recently 3 interpolation methods are supported which are used to determine the new value for each pixel of the resulting map layer. The supported interpolation methods are:

- Next neighbour

- Bi-linear
- Cubic convolution

<u>Change Value</u>

Each RNP tool and every data supplier has its own clutter data classification. The numbers of classes available, the categorization of the classes, and the corresponding class numbers, which are used to store the clutter information, are user specific. Therefore new clutter layers always need an adoption of the clutter. This reclassification can be carried out with CATCHit with the “Change value” tool.

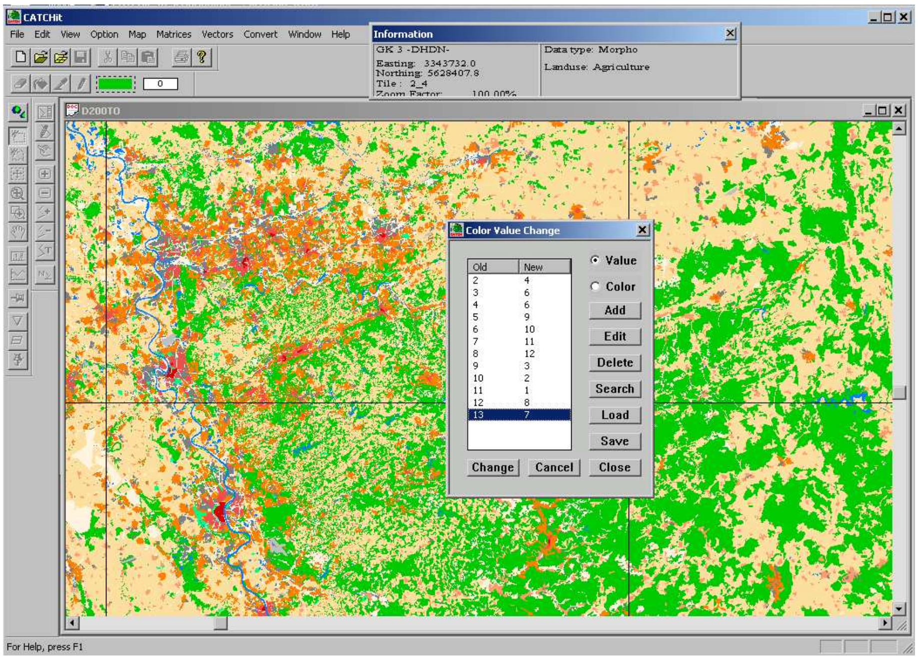


Figure 1.6.12: Change Value – Reclassification of clutter data according to RNP tool specification

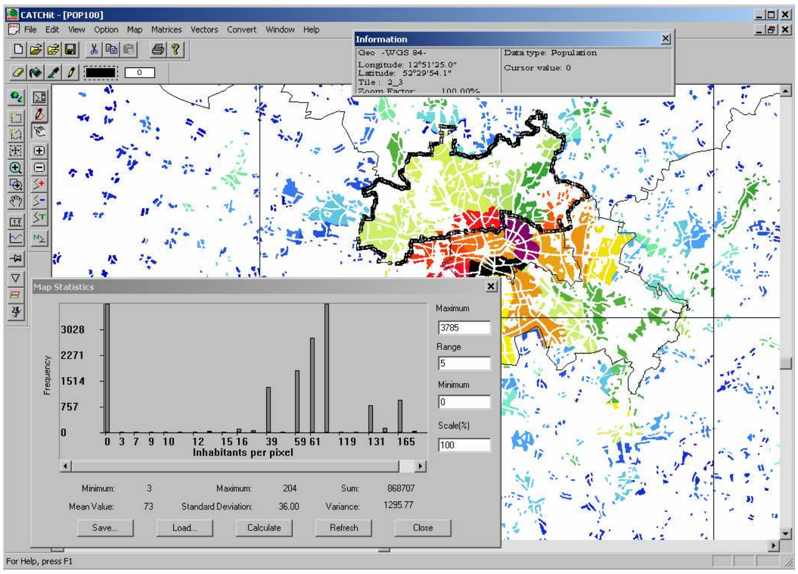
This tool allows searching for every appearing value within a selected area. The found values are listed in a table. Each value can now be changed to another one. This way a correspondence list can be created and stored for later use. The correspondence list can be created in another tool like Excel as well and loaded from a simple text file. Pressing the calculate button will start the change of the values for all pixel within the selected area according to the correspondence list.

Terrain Data Analysis and Maintenance

Every terrain database needs to be maintained. Population is growing. Industrial areas and cities are changing their faces continuously. Since the accuracy of analyses and simulations depends very hard on the topicality of the applied source, it is of utmost importance keeping the sources up-to-date. On the other hand some errors might be detected in the recent terrain data layer, which need to be corrected. Therefore CATCHit has implemented tools for terrain data analysis and maintenance.

Data analysis

For the analysis of a terrain data layer the “Map Statistics” function of CATCHit can be used. It calculates for a selected area the following parameters:
Minimum value, Maximum value, Sum, Mean value, Standard deviation, Variance
The selected area to be analyzed can be limited by a polygonal selection; a 2D-plot shows the value distribution in a graphical way and can be stored in a text file for further analysis in other tools like Excel, etc.

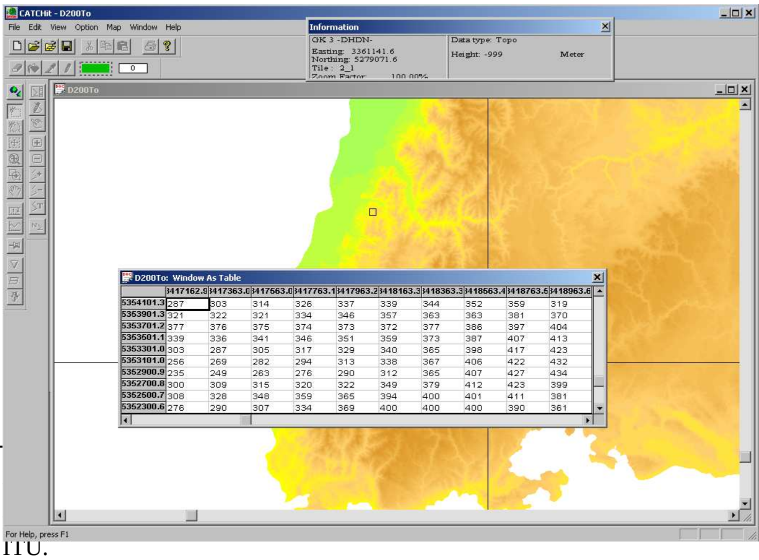


Updates of terrain data layer always need a quality check. Therefore the new version of the layer has to be compared with the old one. CATCHit offers the geo-correct copy and paste tool with mathematical overlay function. The sum or the differences of both data layer can be calculated and colored displayed. This allows a graphical detection of the changes and developments of the new data layer in a graphical manner

Figure 1.6.13: Map Statistics for polygonal selection area in population layer

Data Maintenance

Errors in the terrain database are most of the time difficult to detect but could have great impact on the RNP results. Therefore a tool is needed that allows checking the data of a certain suspected area pixel by pixel. CATCHit offers a “Window as Table” tool. The pixel values are written in a table view of a 10 x 10 pixel big area. The coordinates area displayed besides the table elements and can be changed to any coordinate system that is valid for this area.



	3417162.9	3417363.0	3417563.0	3417763.1	3417963.2	3418163.3	3418363.3	3418563.4	3418763.5	3418963.6
5354101.3	287	303	314	326	337	339	344	352	359	319
5353901.3	321	322	321	334	346	357	363	363	381	370
5353701.2	377	376	375	374	373	372	377	386	397	404
5353501.1	339	336	341	346	351	359	373	387	407	413
5353301.0	303	287	305	317	329	340	365	398	417	423
5353101.0	256	269	282	294	313	338	367	406	422	432
5352900.9	235	249	263	276	290	312	365	407	427	434
5352700.8	300	309	315	320	322	349	379	412	423	399
5352500.7	308	328	348	359	365	394	400	401	411	381
5352300.6	276	290	307	334	369	400	400	400	390	361

Figure 1.6.14: “Window as Table” view of the terrain data with coupled selection area

The table area is directly coupled with the map layer. The values can be changed within the table. Changes will automatically be updated within the map window.

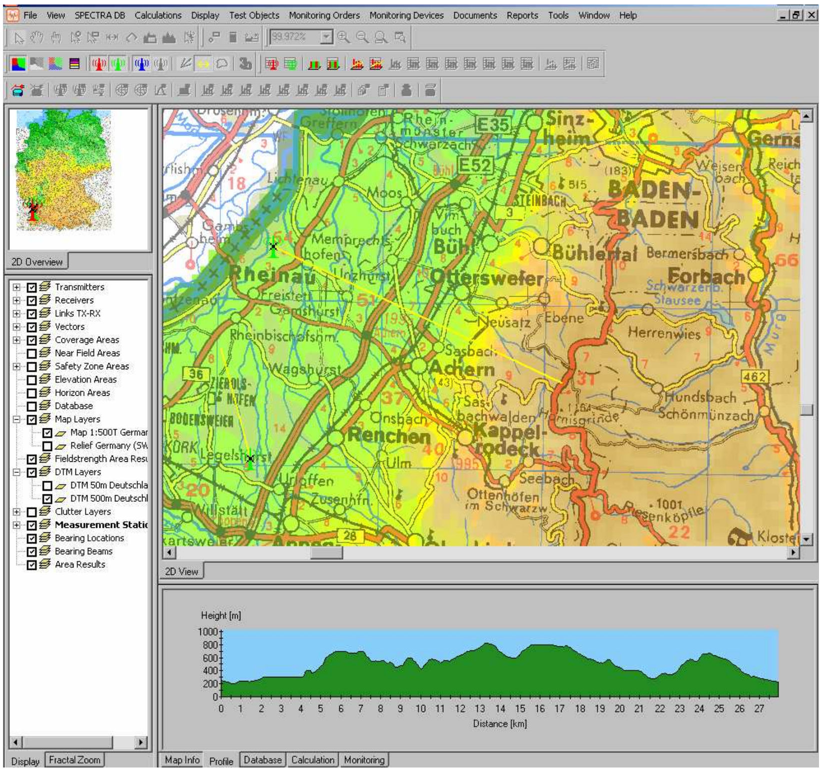
SPECTRAemc

SPECTRAemc is a comprehensive and easy to use visualization and calculation tool providing additional interface to different contemporary database systems. It enables the user to perform EMC calculations and general network planning tasks in conjunction with a powerful wave propagation calculation library for intra and inter radio service applications.

Features and Highlights

SPECTRAemc includes the following main functions:

- Selection of licensed transmitters/receivers from a central database
- Visualization of the selected transmitters/receivers on a map
- Access to the properties of the selected transmitters/receivers
- Various calculations based on the on board wave propagation library
- WAVE PROPAGATION LIBRARY: including specially tuned wave propagation models for the complete radio frequency range from VLF to EHF (3 kHz - 300 GHz)



- Field strength calculation at a single point or in an area
- Calculation of coverage areas, contours, safety zone and near field
- Interference analysis (interference, desensitization, frequency pre-selection and area based analysis)
- Inter-modulation analysis
- Reports on calculation results and selections of technical data based on powerful Crystal Reports templates
- 3D view of DTM maps including transparent overlay with other types of map layers

Implemented ITU-Recommendations

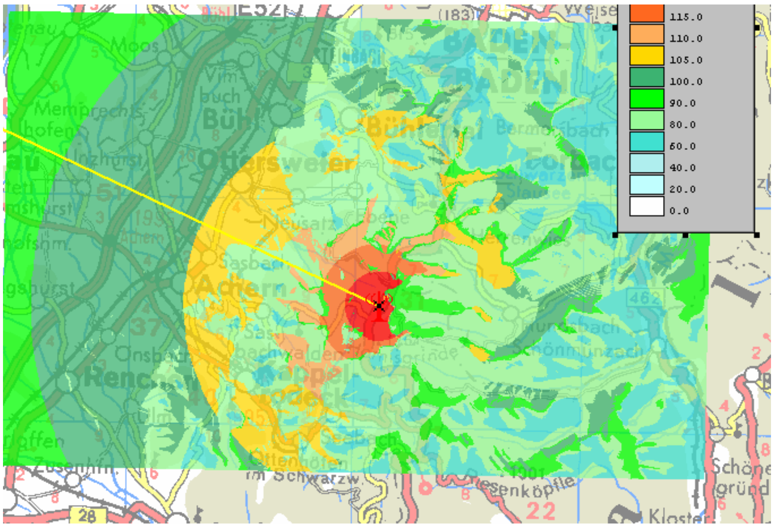
ITU-R SM.1370 SMS
ITU-R P.1546 Terrestrial services
ITU-R P.530 Planning of microwave links
ITU-R P.533 HF propagation
ITU-R P.452 Microwave interference
ITU-R P.676 Atmospheric absorption
ITU-R P.837 Attenuation due to rain
ITU-R P.526 Diffraction mechanism (knife edge)

Terrain Map Database

SPECTRAemc supports both raster and vector maps. Calculation results can be visualized as vectors (such as field strength or interference contours) or as raster data (for example area calculations). Raster or vector calculation results can be overlaid on background maps with an adjustable transparency factor. Several third party map formats are supported.

Coverage Prediction

For coverage prediction a variety of propagation modules for the frequency range VLF to EHF are available. The propagation parameters used by the different modules (e.g. rain zone, k- factor, etc.) are user configurable. This includes standard ITU recommendations as well as modules used in GSM, DCS and UMTS services (e.g. Okumura Hata)



Interference analysis

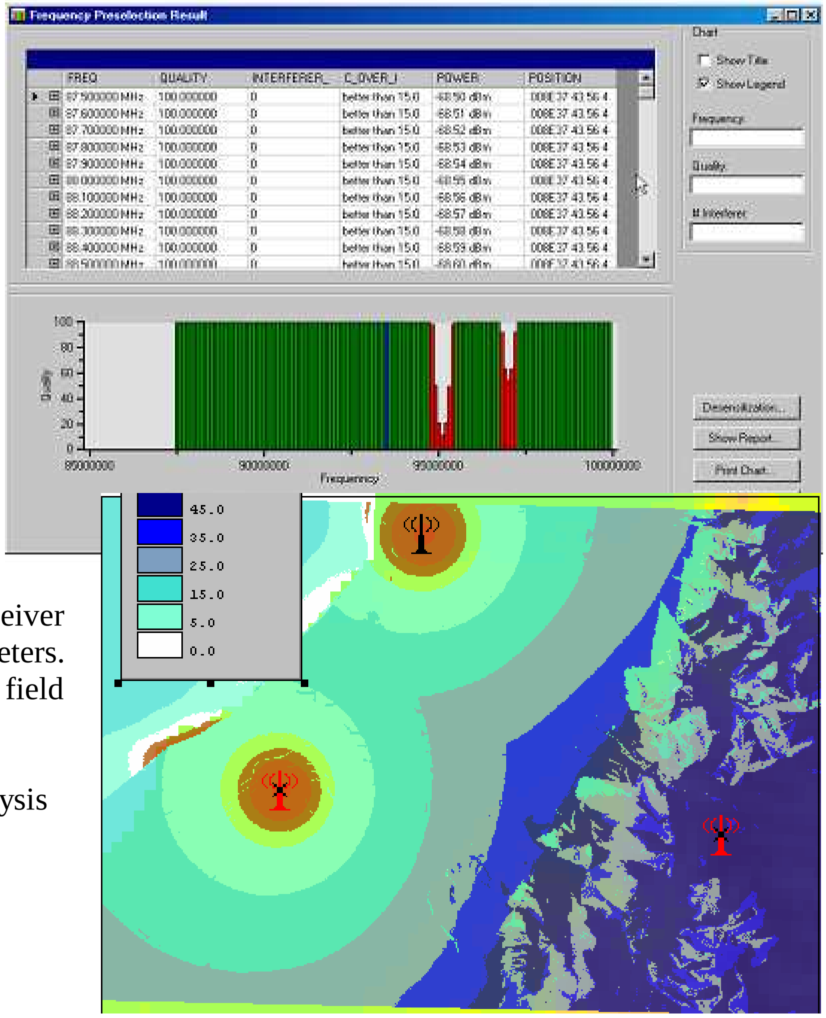
Based on ITU-R SM.1370 interference analysis and frequency assignment within existing networks is done in threesteps.

- Receiver desensitization analysis
- Transmitter noise analysis
- Frequency interference (co-channel, interstitial and adjacent channel or distant channel) analyses

An extensive report on the results is issued for these cases.

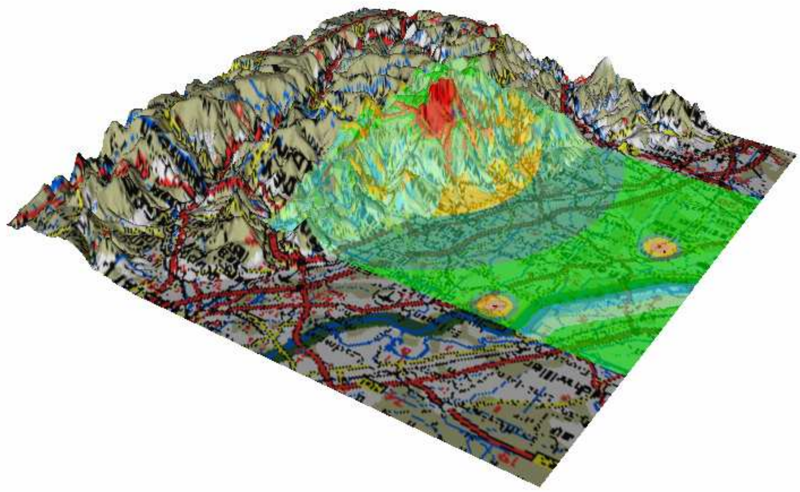
In addition to these results, area based interference evaluation can be calculated, using a typical receiver with user configurable parameters. It enables the user to perform field strength based analysis as:

- C/I evaluation
- Maximum server analysis
- Composite coverage analysis



3D-View

SPECTRAemc offers for a visual evaluation the possibility to display area results on top of terrain data as a 3D-view. Transmitter, receivers and links are displayed also. Using this feature the user can immediately detect areas of interest for a further detailed analysis.



Annex 2 – Case Studies

A selection of most frequent case studies (ie: Network extension, transmission, signalling, migration to NGN, mobile, etc.) is provided in order to illustrate the application process.

List of case studies:

A2.1. Forecasting of services

On the national transit level of a network the volume of services used by the network customers, which transit over the national network, is presented with the national traffic matrix.

For the consolidation of the semi-meshed national transit (LD) circuit switched network, shown in Case Study A2.2, such traffic matrix is constructed and forecasted, using as tool EXCEL.

To construct the traffic matrix data from different sources are used:

- Subscriber traffic per category of populated place, per customer class and per network level
- Measurements of the traffic load on the interexchange trunks and of the traffic interest between different network areas
- Evaluation of the monthly traffic variations within one year

Initial data are the national traffic matrix of the incumbent network operator, extended with the traffic from/to the mobile operators.
Traffic data are formed as 34 x 39 EXCEL table with total traffic of 9120.59 Erlang (see extract in Table A2.1.1).

	TRD_B													
TRD_A	Sliven	Smol	Sofia	SPRINT	St. Zag	STTS	Targ	Varna	VAS	V. Tar	Vidin	Vraca	Yambol	Grand Total
Blagoevg	0,46	0,49	1,99	0,63	1,09	45,73	0,07	2,55	1,60	1,50	0,44	0,48	0,35	200.83
Burgas	11,43	0,71	0,98		5,84	73,04	0,29	19,94	4,40	2,04	0,12	0,30	10,06	391.78
Dobrich	0,24	0,14	1,11	0,17	0,45	23,85	0,19	30,57	2,09	0,85	0,07	0,76	0,26	149.99
Dupnica	0,02	0,08	0,32	0,35	0,03	13,51	0,04	0,11	0,79	0,01	0,67	0,10	0,05	63.51
Gabrovo	0,44	0,13	0,41		2,99	23,67	0,35	3,40	2,24	11,63	1,37	1,11	0,15	133.88
GLOBUL	2,47	1,84	3,95		5,50	32,34	1,00	7,70		2,74	1,87	1,88	2,43	133.44
Haskovo	0,96	2,52	1,39	0,18	7,38	24,88	0,02	2,76	2,60	1,63	0,49	0,59	0,90	180.84
International	10,98	2,88	8,20		16,48	109,58	3,77	25,82		8,98	7,62	7,54	6,53	675.16
Lovech	0,07	0,09	2,03	0,86	1,46	30,76	0,20	1,76	2,41	5,63	0,04	1,70	0,15	137.28
MOBICOM	1,99	1,91	2,27		5,25	17,68	0,75	4,69	0,01	1,94	0,65	1,32	1,90	90.45
MOBILTEL	6,98	6,69	12,51		12,30	165,30	4,25	28,27	0,07	14,71	5,91	10,02	5,14	602.32
Montana		0,12	1,24	0,05	0,52	33,61	0,06	1,99	1,56	1,01	3,46	9,80	0,54	113.47
Pazardjik	0,89	1,75	3,39	0,24	2,26	32,70	0,10	1,69	3,20	0,40	0,39	0,18	0,06	188.49
Pernik	0,16	0,14	3,28	0,10	0,58	40,67	0,28	1,49	1,09	0,74	0,35	0,46	0,10	**110.99**
Pleven	0,33	0,07	1,75	0,55	1,38	45,97	0,70	4,87	3,82	14,20	1,90	6,26	0,14	**220.72**
Sliven		0,29	0,32	0,82	9,91	16,10	0,43	3,29	2,14	2,99	0,33	0,29	7,68	**134.14**
Smolyan	0,13		0,17		1,02	12,21	0,30	1,55	0,75	0,24	0,25	0,34	0,10	**69.11**
Sofia	0,99	0,29		0,27	0,78	117,95	0,01	2,20	2,73	1,53	0,11	2,33	0,15	**247.64**
Stara Zagora	11,52	0,71	0,51			39,87	0,24	4,50	2,54	3,25	0,12	0,60	6,77	**243.42**
STTS	19,18	15,53	121,92	5,66	26,34		10,84	68,38	24,19	35,65	18,06	37,65	10,83	**2423.18**
Targovishte	0,39		0,26		0,42	9,01		8,09	1,15	3,20	0,05	0,08	0,56	**81.40**
Vraca	0,63	0,22	3,97	0,58	1,19	44,00	0,11	3,28	3,53	2,74	5,59		0,03	**178.19**
Yambol	3,35	0,06			2,29	12,12		0,74		0,65	0,06			**59.15**
Grand Total	**86,52**	**55,56**	**183,85**	**11,81**	**155,11**	**1349,29**	**41,74**	**296,09**	**85,22**	**150,23**	**55,47**	**94,51**	**64,47**	**9120.59**

Table A2.1.1 Extract from the measured national traffic

Corresponding matrix of traffic interest between network areas with 0 to 1 coefficients is calculated.

The data are rearranged as 27 x 27 EXCEL table to reflect the national network structure.

Also, in this table the total traffic per direction is calculated (see extract in Table A2.1.2).

TRD_A	T_SfA	T_SfB	Blag	Burgas	Dobrich	Haskovo	Kardjali	Pazar	Shumen	Yambol	GSM_SA	GSM_SB	Internat.	Total_Nat	GSM	Grand Total
T_SfA	0.00	0.23	0.07	0.06	0.01	0.02	0.01	0.04	0.01	0.01	0.50	0.50	0.11	0.47	0.42	1124.52
T_SfB	0.23	0.00	0.07	0.06	0.01	0.02	0.01	0.04	0.01	0.01	0.50	0.50	0.11	0.47	0.42	1124.52
Blagoevg	0.35	0.35	0.00	0.02	0.02	0.02	0.00	0.02	0.01	0.00	0.50	0.50	0.11	0.44	0.45	266.48
Burgas	0.22	0.22	0.01	0.00	0.02	0.02	0.00	0.02	0.03	0.06	0.00	0.00	0.05	0.54	0.41	308.04
Smolyan	0.18	0.18	0.03	0.02	0.00	0.02	0.02	0.04	0.01	0.00	0.00	0.50	0.04	0.57	0.39	59.56
St.Zagora	0.17	0.17	0.02	0.06	0.01	0.10	0.03	0.01	0.01	0.05	0.00	0.00	0.05	0.55	0.41	271.61
Varna	0.18	0.18	0.01	0.07	0.13	0.01	0.01	0.01	0.09	0.01	0.00	0.00	0.06	0.55	0.39	425.24
V.Tarnovo	0.20	0.20	0.01	0.03	0.01	0.02	0.00	0.01	0.04	0.01	0.00	0.00	0.06	0.53	0.41	366.18
Vraca	0.33	0.33	0.02	0.02	0.01	0.01	0.00	0.02	0.01	0.00	0.50	0.50	0.06	0.53	0.41	298.57
Yambol	0.20	0.20	0.00	0.17	0.00	0.02	0.00	0.01	0.00	0.00	0.00	0.00	0.03	0.65	0.33	46.68
GSM_SA	0.34	0.34	0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.30	0.70	0.00	244.47
GSM_SB	0.26	0.26	0.07	0.00	0.00	0.04	0.02	0.04	0.00	0.00	0.00	0.00	0.25	0.75	0.00	302.51
GSM_VT	0.00	0.00	0.00	0.00	0.14	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.00	1.00	0.00	49.21
GSM_PD	0.00	0.00	0.00	0.00	0.00	0.16	0.07	0.16	0.00	0.00	0.00	0.00	0.00	1.00	0.00	58.04
GSM_VN	0.00	0.00	0.00	0.14	0.06	0.00	0.00	0.00	0.06	0.04	0.00	0.00	0.00	1.00	0.00	110.56
GSM_SZ	0.00	0.00	0.00	0.26	0.00	0.00	0.00	0.00	0.00	0.08	0.00	0.00	0.00	1.00	0.00	61.34
International	0.15	0.15	0.05	0.04	0.02	0.02	0.02	0.03	0.02	0.02	0.50	0.50	0.00	0.57	0.43	675.67
Grand Total	766.58	766.58	161.21	181.96	90.45	106.40	52.03	114.61	85.01	64.47	814.98	1024.89	671.23			7828.07

Table A2.1.2 Extract from the traffic interest matrix

The total traffic in this EXCEL table is 7 828.07 Erlang, as far as all internal for the 27 areas traffic is excluded .

Forecasted traffic matrix is based on the measured and rearranged traffic matrix, multiplied with a coefficient K,

K= K1x K2 x K3 x K4 = 1,34

Where,

K1 compensates the traffic measuring method precision : **K1 = 1,05**

K2 is yearly traffic variations (e.g. January compared to July) : **K2 = 1,10**

K3 reflects the subscriber forecast : **K3 = 1,05**

K4 is forecast for the transit traffic growth, based on improved economic situation and lowered tariffs : **K4 = 1,10**

Microsoft Excel - Tab_3.2_3.3.xls

File Edit View Insert Format Tools Data Window Help

C9 = =1.34*MAT!C9

	M	N	O	P	Q	R	S	T
1	9.35	13.52	10.60	25.44	47.29	48.30	61.83	7.36
2	9.35	13.52	10.60	25.44	47.29	48.30	61.83	7.36
3	0.95	0.65	0.86	1.93	4.84	4.53	4.08	0.52
4	6.01	15.31	0.95	9.53	26.71	5.44	1.92	13.48
5	3.16	0.32	0.19	0.78	40.96	2.32	1.41	0.35
6	0.26	1.29	3.38	11.88	3.70	3.03	1.68	1.20
7	0.18	0.44	2.47	2.33	0.61	1.21	0.53	0.99
8	0.40	1.19	2.35	3.92	2.27	1.43	1.51	0.08
9	0.27	0.22	0.19	1.01	2.00	1.61	1.45	0.14
10	1.71	0.54	0.22	5.30	8.89	39.71	16.24	0.38
11	4.62	3.39	18.73	25.84	18.46	14.49	8.65	5.32
12	7.83	3.96	0.69	4.32	33.84	32.15	2.89	1.64
13	0.00	0.90	0.14	2.12	24.49	9.13	1.54	0.80
14	1.50	0.00	0.39	14.97	4.41	5.35	1.20	10.29
15	0.43	0.17	0.00	1.52	2.08	0.98	0.91	0.14
16	1.06	17.97	1.31	0.00	6.76	11.13	1.94	9.53

Table A2.1.3 EXCEL view of forecasted traffic matrix calculation

The total forecasted traffic is 10 489.6 Erlang (see extract in Table A2.1.4).

Forecasted traffic matrix

TRD	T_SfA	T_SfB	Blagoev	Burgas	Dobrich	Haskovo	Kardj	Pazard	Pernik	Pleven	Plovdiv	Ruse	Shumen	Sliven	Smolyan
T_SfA	0,00	160,72	51,05	40,11	10,26	16,28	4,95	25,43	25,25	44,62	71,67	34,92	9,35	13,52	10,60
T_SfB	160,72	0,00	51,05	40,11	10,26	16,28	4,95	25,43	25,25	44,62	71,67	34,92	9,35	13,52	10,60
Blagoevg	54,43	54,43	0,00	2,39	2,68	2,39	0,22	2,64	5,45	3,95	8,10	0,86	0,95	0,65	0,86
Burgas	49,60	49,60	1,85	0,00	3,66	4,80	0,72	5,10	0,24	2,90	15,77	8,52	6,01	15,31	0,95
Dobrich	16,72	16,72	0,56	2,26	0,00	0,47	0,11	0,37	0,14	1,88	1,54	8,63	3,16	0,32	0,19
Haskovo	17,60	17,60	1,74	3,26	3,39	0,00	13,29	2,83	0,56	1,42	21,69	1,97	0,26	1,29	3,38
Kardjali	6,99	6,99	0,44	1,13	0,30	9,22	0,00	0,94	0,02	0,85	9,77	1,23	0,18	0,44	2,47
Pazardjik	24,18	24,18	3,12	1,61	0,21	0,44	0,78	0,00	0,02	1,42	43,53	3,74	0,40	1,19	2,35
Pernik	29,45	29,45	8,70	0,75	0,04	0,33	0,00	0,51	0,00	1,33	2,09	1,27	0,27	0,22	0,19
Pleven	53,94	53,94	2,39	2,79	1,03	0,69	0,96	1,29	0,71	0,00	12,04	9,04	1,71	0,54	0,22
Plovdiv	83,61	83,61	7,88	16,07	2,65	22,55	11,08	35,89	1,44	9,25	0,00	10,07	4,62	3,39	18,73
Ruse	42,01	42,01	1,31	4,52	7,58	3,41	0,63	0,88	0,21	10,90	9,27	0,00	7,83	3,96	0,69
Shumen	12,38	12,38	0,99	3,33	2,66	0,37	0,22	0,55	0,69	1,46	4,24	9,95	0,00	0,90	0,14
Sliven	11,00	11,00	1,20	11,33	0,53	1,03	0,16	0,88	0,14	1,32	4,76	1,01	1,50	0,00	0,39
Smolyan	8,30	8,30	1,33	0,77	0,05	0,71	1,11	1,70	0,02	0,73	15,55	0,58	0,43	0,17	0,00
St. Zagora	33,99	33,99	3,38	11,21	1,10	19,97	5,93	2,47	0,56	3,20	29,60	3,43	1,06	17,97	1,31
Varna	56,49	56,49	2,98	22,95	40,89	3,15	1,99	2,25	1,32	10,32	14,97	24,78	27,89	4,73	2,09
V. Tarnovo	51,32	51,32	3,29	7,35	3,24	4,83	0,10	2,07	0,93	30,50	14,16	29,97	10,33	2,42	0,71
Vraca	70,65	70,65	3,23	4,67	1,82	1,17	0,39	3,20	1,32	17,43	6,44	6,36	2,77	0,87	0,64
Yambol	8,12	8,12	0,04	6,97	0,06	0,94	0,07	0,55	0,03	1,20	2,69	0,89	0,13	4,48	0,09
GSM_SA	78,40	78,40	22,59	0,00	0,00	0,00	0,00	0,00	7,73	20,38	0,00	0,00	0,00	0,00	0,00
GSM_SB	78,40	78,40	22,59	0,00	0,00	12,62	5,39	12,74	7,73	20,38	40,03	0,00	0,00	0,00	6,99
GSM_VT	0,00	0,00	0,00	0,00	9,41	0,00	0,00	0,00	0,00	0,00	0,00	23,90	8,44	0,00	0,00
GSM_PD	0,00	0,00	0,00	0,00	0,00	12,62	5,39	12,74	0,00	0,00	40,03	0,00	0,00	0,00	6,99
GSM_VN	0,00	0,00	0,00	21,06	9,41	0,00	0,00	0,00	0,00	0,00	0,00	23,90	8,44	7,66	0,00
GSM_SZ	0,00	0,00	0,00	21,06	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	7,66	0,00
International	78,91	78,91	24,29	18,12	9,96	8,32	11,26	13,11	5,63	16,00	55,41	34,92	8,84	14,71	3,86

Grand Total	**1027,2**	**1027,2**	**216,0**	**243,8**	**121,2**	**142,6**	**69,7**	**153,6**	**85,4**	**246,1**	**495,0**	**274,9**	**113,9**	**115,9**	**74,4**

Table A2.1.4 Extract from the forecasted traffic matrix

All dimensioning results shown in Case study A2.2. are based on the forecasted traffic matrix, calculated as shown above with the standard features of the EXCEL.

A2.2. Consolidation of national transit network

Consolidation of a semi-meshed national transit (LD) circuit switched network consists of applying of hierarchical routing with dual homing for the transiting of the traffic.

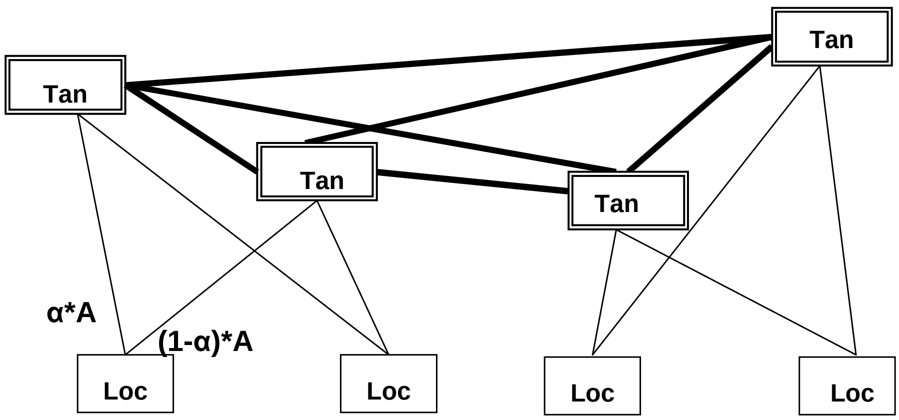
<u>Dual homing (load sharing) :</u>

In the hierarchical routing one option is to overflow/transit traffic through two different tandems (Tan), i.e. to implement dual homing for the source of the traffic (Loc).

General rule is to divide traffic in equal portions, i.e. 50% to 50%.

More universal approach will be to use coefficient **α, 0< α <1.**

Fig A2.2.1. : Dual homing (load sharing)



<u>Dimensioning of the LD network :</u>

Dimensioning of the national LD network consists of routing of the traffic over the hierarchy and dimensioning of each trunk for the required QoS.

A planning tool will be needed for a real practical case, because of the large number of cases and the complicated routing.
PLANITU is appropriate tool for such task.

Main input data are source-destination traffic matrix of the LD network.

Results:

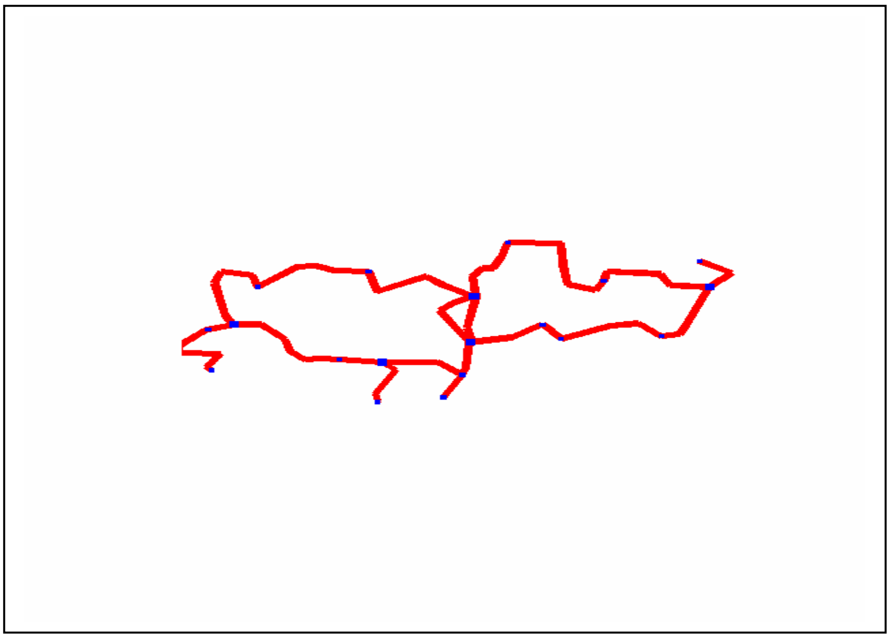
Fig A2.2.2. : Example of routing of traffic and dimensioning of dual homing

```
Exchange   1      Blagoev

 J  Name         A       M     V/M      Circs  Avl     Cong     Total    Routing

 1  Blagoev    0.00    0.00   1.00        0     0     0.000    0.000     D
 2  Pernik     5.45    5.45   1.00        0     0     1.000    0.004     T 26 27
 3  Vratza     4.08    4.08   1.00        0     0     1.000    0.009     T 26 27
 4  Pleven     3.95    3.95   1.00        0     0     1.000    0.004     T 26 27
 5  Pazarjik   2.64    2.64   1.00        0     0     1.000    0.007     T 26 27
 6  Smoljan    0.86    0.86   1.00        0     0     1.000    0.008     T 26 27
 7  Haskovo    2.39    2.39   1.00        0     0     1.000    0.012     T 26 27
 8  Kurdjali   0.22    0.22   1.00        0     0     1.000    0.007     T 26 27
 9  Ruse       0.86    0.86   1.00        0     0     1.000    0.004     T 26 27
10  Dobrich    2.68    2.68   1.00        0     0     1.000    0.005     T 26 27
11  Shumen     0.95    0.95   1.00        0     0     1.000    0.012     T 26 27
12  Sliven     0.65    0.65   1.00        0     0     1.000    0.005     T 26 27
13  Jambol     0.52    0.52   1.00        0     0     1.000    0.005     T 26 27
14  Burgas     2.39    2.39   1.00        0     0     1.000    0.013     T 26 27
15  GSM_SA    80.83   80.83   1.00        0     0     1.000    0.006     T 26 27
16  GSM_SB    80.82   80.82   1.00        0     0     1.000    0.014     T 26 27
17  GSM_VT     0.00    0.00   1.00        0     0     1.000    0.004     T 26 27
18  GSM_PD     0.00    0.00   1.00        0     0     1.000    0.008     T 26 27
19  GSM_VN     0.00    0.00   1.00        0     0     1.000    0.009     T 26 27
20  GSM_SZ     0.00    0.00   1.00        0     0     1.000    0.003     T 26 27
21  Intern    39.54   39.54   1.00        0     0     1.000    0.003     T 26 27
22  Tr_C_VT    4.53    4.53   1.00        0     0     1.000    0.004     T 26 27
23  Tr_C_PD    8.10    8.10   1.00        0     0     1.000    0.007     T 26 27
24  Tr_C_VN    4.84    4.84   1.00        0     0     1.000    0.005     T 26 27
25  Tr_C_SZ    1.93    1.93   1.00        0     0     1.000    0.002     T 26 27
26  TrC_SfA   54.43  178.55   1.00      210     0     0.002    0.002     D
27  TrC_SfB   54.43  178.55   1.00      210     0     0.002    0.002     D

Total:                                  420     0
```

Fig A2.2.3. : Dimensioning of the LD network

	SF-A	SF-B	BL	PK	VR	PL	VT	RS	VN	DB	SH	BS	YA	SL	SZ	HS	KD	PD	PZ	SM	Σ
SF-A		870	150	60	150	150	150		180						120			150			1980
SF-B	870		150	60	150	150	150		210						120	90	60	210	120	60	2400
BL	210	210																			420
PK	90	90																			180
VR	240	240																			480
PL	240	240																			480
VT	150	150						180	240	60	60				60			60			960
RS							240		240												480
VN	180	180					150	180		120	90	150	60	90	240			60			1500
DB							90		150												240
SH							90		120												210
BS									240						240						480
YA									60						60						120
SL									90						120						210
SZ	120	120					60		210			150	60	90				90			900
HS		120																150			270
KD		90																90			180
PD	150	360					60		90						90	90	60		120	60	1080
PZ		120																150			270
SM		60																60			120
Σ	2250	2850	300	120	300	300	990	360	1830	180	150	300	120	180	1050	180	120	1020	240	120	12960
DIE	420	420					60		60						60			90			1110

Fig A2.2.4. : Transmission layer of the LD network



The consolidation of the national LD network results in:

- transition from semi-meshed towards dual-homing network structure
- more robust and reliable traffic handling and routing
- simplifying the network management
- readiness for smooth transition towards Class 4 NGN solutions, deploying MGW in the location of the existing upper level transit exchanges.

A2.3. Business planning

Application of STEM

A2.3.1. Problem of network evolution and study case scope

- One of the most common planning issues today in mobile networks is the design of network migration from the 2G technology and capabilities towards the 3G technology and associated services. In particular, decisions have to be taken on the target network and technology to evolve, the adequate timing for the migration, the policy to introduce new services and the estimate of the expected investments and profitability.
- The main objective of this case study for a reference medium developed country was to characterize the involved parameters in the evaluation for an scenario that has already a GSM solution with an intermediate level of penetration in the population, the impacts of the migration and the assesment on an evolution either towards and intermediate solution like the EDGE or towards a full 3G UMTS.
- In order to assess convenient migration alternatives, estimate needed investments for infrastructure modernization, project revenues, project NPV and business feasibility; a techno-economic modeling was performed and later implemented with the STEM tool as described in following sections.

A2.3.2. Modeling scenarios for business analysis

This evaluation contemplates the overall market size of the country with the corresponding weighted average values for each business parameter. In order to have enough observation period for the migration from 2G to 3G, business evaluation is performed in a 10 years time frame from Y1 up to Y10. Consolidated reference inputs from the network correspond to the year Y0 and evaluation starts in an initial population around 3 Million inhabitants and growing at a cumulative rate of the order of 2, 5% per year. Current mobile services penetration is low as compared to similar countries in development and a large space exist for new customers grow that will more than double in the 10 years period.

From the processing of the geographical information in the country, 3 type of geo-scenarios are defined on the base of population, area, population density, subscribers penetration, number of current radio cells and traffic per subscriber: Urban Type comprising the nucleus or major cities with high/moderate values of previous variables, Suburban type including cities or areas with medium/low values and finally Rural type with very low values of those parameters. This characterization will imply a differentiation in the modelling of the mobile infrastructure, number of sites and related costs.

According to the case study objectives, two migration scenarios are evaluated in order to know the impact of investments and revenues in the evolution:

- Case A as a conservative plan evolving to EDGE within the 900 MHz and 1800 MHz frequencies and the services and Quality that may be carried according to the possible Peak Rates and Sustained Bit Rates in this technology. This will save investments in the infrastructure mainly in low density regions but still will be able to serve many of the new data services with low and medium speed.
Additional services to be handled in EDGE as compared to GSM/GPRS include data up to 144 Kbps peak rate, audio streaming, video download, video messaging, news, location based services, etc.

- Case B as a more ambitious plan to exploit all potential in new multimedia services with higher capacity in voice traffic and higher Peak Rates, Sustained Bit Rates and Quality in data but implying higher investments in infrastructure and new frequency assignment in the 2GHz band. This option will generate additional revenues mainly from the high data rate services and real time multimedia services that require high Grade of Service
Additional services to be handled in UMTS as compared to GSM/EDGE include data up to 384 Kbps or 2 Mbps peak rate, video calling, video mailbox, intranet/extranet, enhanced quality news, enhanced location services, videoconferencing, enhanced video streaming, mobile TV, etc.

From the point of view of traffic flows consideration and taking into account the confluence of a multimedia environment, the following modelling was developed:

- Packet mode flows are characterized both by the Peak rates and Sustained Bit Rates (SBR) that are used to define which services may be handled. Sustained Bit Rate is defined as the bit rate of a service that may be carried in a sustained manner by the network resources with the fulfilment of the Grade of Service conditions required by that service in terms of packet delay, jitter and loss probability. Period for measurement of SBR is typically 5 minutes. SBR is also used as the link between the information flow for dimensioning purposes and the information volumes that generate service revenues. It has to be noted that SBR for each service is lower than the peak rate with a proportion that is a function of the service or service class characterization.

Dimensioning of the upper layer network resources follows the standard circuit mode or packet mode criteria as a function of the origin/destination traffic flows and is equivalent for all scenarios under analysis.

When analyzing migration scenarios to different radio type solutions, access dimensioning requires special care to ensure validity of the comparisons and the following criteria have to be taken into account for the number of base station sites in all technologies: GSM/GPRS, GSM/EDGE, UMTS and CDMA:

- a) Coverage dominated by Frequency of the radio system used, that decreases the radius with a quadratic relation to the frequency.

- b) Coverage dominated by Bandwidth of data services to be provided that are very sensitive to the signal to noise ratio and specially for the upper radio signal. This condition being more restrictive that the previous one and also with a quadratic behaviour is common to all technologies under consideration.

- c) Coverage dominated by Topography as a function of the terrain conditions of each area to be covered and very dependent of each particular region. This coverage is to be considered jointly to the previous ones and may be modelled with some extrapolation on the basis of the existing network.

- d) Coverage dominated by Voice Traffic capacity measured by the capacity in erlangs of each cell for a given grade of service. This criterion is a function of the service mix in the area and number of subscribers.

- e) Coverage dominated by Data Traffic capacity measured by the capacity in Sustained Bit Rate of a given service type or service class that ensures the agreed Grade of Service. This criterion is the most restrictive one in the IMT-2000 systems as both capacity in Mbps and signal to noise ratio by the distance to the source are applicable. Special care has to be taken to ensure the service level agreement (SLA) in all the cell area and not only in the close vicinity to the BS. This criteria has a quadratic behaviour and is frequently the most critical to fulfil in a mature network, being dominant when traffic of high speed real time services have to be provided in all CDMA based technologies.

The overall BS dimensioning is a combination of the previous 5 factors according to each geo-scenario type and has the following combined effects:

In a high density area of urban type the design at the start year is already dominated by the capacity in erlangs of the GSM solution with an initial cell radius in the order of magnitude of 1 to 4 km and no increase of BS is needed for the radio frequency. An initial saving is obtained due to the higher erlang capacity per site in 3G that nevertheless has the counterpart of an increase that is required by the high speed data services. The final effect is a slight increase of sites.

In a medium density area of suburban type**,** existing design at start has cell radius between 4 and 10 km and the same combined effects appear, although the capacity saving is lower and the criteria e) for the SBR may imply a significant increase in sites.

In a low density area of rural type**,** existing cell radius larger than 10 km and up to 20 km is common and all the radio coverage criteria are dominant as compared with the erlang capacity one. In theses type of areas, an important increase of sites and investment in infrastructure is needed both for the higher frequencies and the higher service bandwidth/speed rates.

A2.3.3. Business evaluation with STEM tool

The business evaluation of a number of migration scenarios like the ones described above, require powerful tools to be able to implement all dynamic evolution for demand, installed or substituted equipments, associated operation costs, equipment life cycles and all related financial calculations. The tool selected for this case study was STEM, as described in section A 1.3 of this document due to the fact that fulfils all needed requirements. Application is based on a reference transition reference case developed by Analysys.

Network is modelled in 7 STEM resource layers for all elements with significant influence in the business parameters as follows:

- Market segments, customers and services for all demand related flows
- Radio access network for all Base Stations equipment per type of geo-scenario
- Base Station sites for the physical structure and location
- Backhaul transmission from the Base Stations to the core network
- Transport network among high level network nodes
- Core network with the MSC, HLR, RNC, SGSN, GGSN, etc.
- Interconnection to other networks

Business modelling is considered with the associated techno-economical parameters specific to the above mentioned network layers and the ones generic for all projects and country dependent. In those aspects related to new services, traffics, tariffs and trends for the 10 years period, a benchmarking has been performed with leading companies worldwide that already implemented 3G type of solutions. As an indication, a summary of both business and technical assumptions is given below:

- Voice tariffs are today over the average benchmark and are expected to decrease per year due to competition at a rate of 7%
- Data tariffs decrease due to maturity at a rate of 7% per year.
- PSTN and Internet interconnection charge with a decrease of 5% per year
- Hardware equipment costs evolving with a yearly decrease of 7%
- Business to residential customers proportion at the start of the study period is of 1:2
- Business busy hours/days per year of 250

- Residential busy hours/days per year of 360

- Voice traffic and circuit mode data flows are characterized by the traffic in erlangs and loss probability models are applied for the dimensioning of the network resources

- Proportion of traffic in the busy hour in relation to the full day of 20%

- Voice call blocking in the busy hour less or equal than 2%

- Activity user utilization rate in transmission state for the packet mode applications of 10%

- Migration to new solutions stars in the Y2 with few pilot sites in urban and suburban areas opening to the entire network since Y3.

- Churn rate from the existing GSM/GPRS technology and terminals to the 2.5G/3G scenarios is of 20% per year

A2.3.4. Evaluation Results

A high amount of results was obtained within the case study as provided by STEM for all technical, business and financial parameters and special analysis was done with the following ones:

- Evolution of mobile customers demand per main category that will grow at a Cumulative Average Grow Rate of 8%.

- Projection for multiservice traffic flow demands per service class:
    - Voice
    - Data circuit type at very low (V-low spd CS) and low rates (Low spd CS)
    - Data packet type at low (Low spd PS), medium (Medium spd PS) and high speed rates (High spd PS). This last case of high speed only feasible with UMTS technology.

- Projection for capacity requirements per flow class in Packet/s, BHCA, Mbs/s, etc.

- BS, TRX and Carriers per geo area: Urban, Suburban, and Rural

- Operation and maintenance costs

- Main projections for average tariffs

- Overall revenue projection per main user/service types

- CAPEX and OPEX projections

- Financial values and ratios: (Cash flows, NPV, IRR. Etc.)

For illustration purposes a selection of those results are included herewith in order to derive main conclusions and assessment for both scenarios: evolution to EDGE and evolution to UMTS

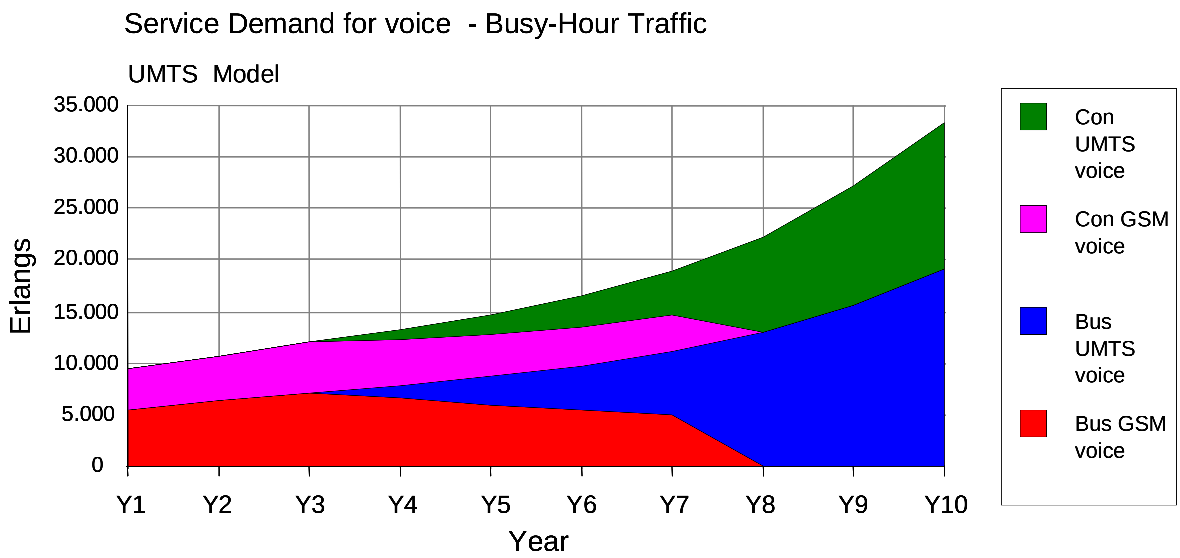


Fig:2.3.1 Service demand for voice at the Busy hour for business and consumer customers

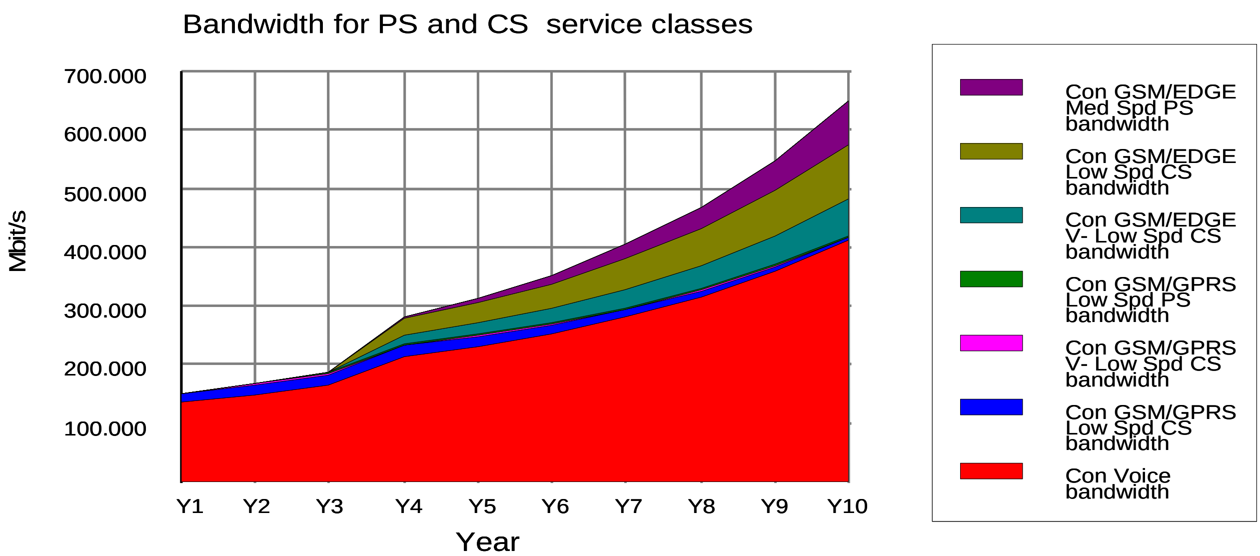


Fig:2.3.2 Bandwith evolution per service class of consumer customers in the EDGE scenario

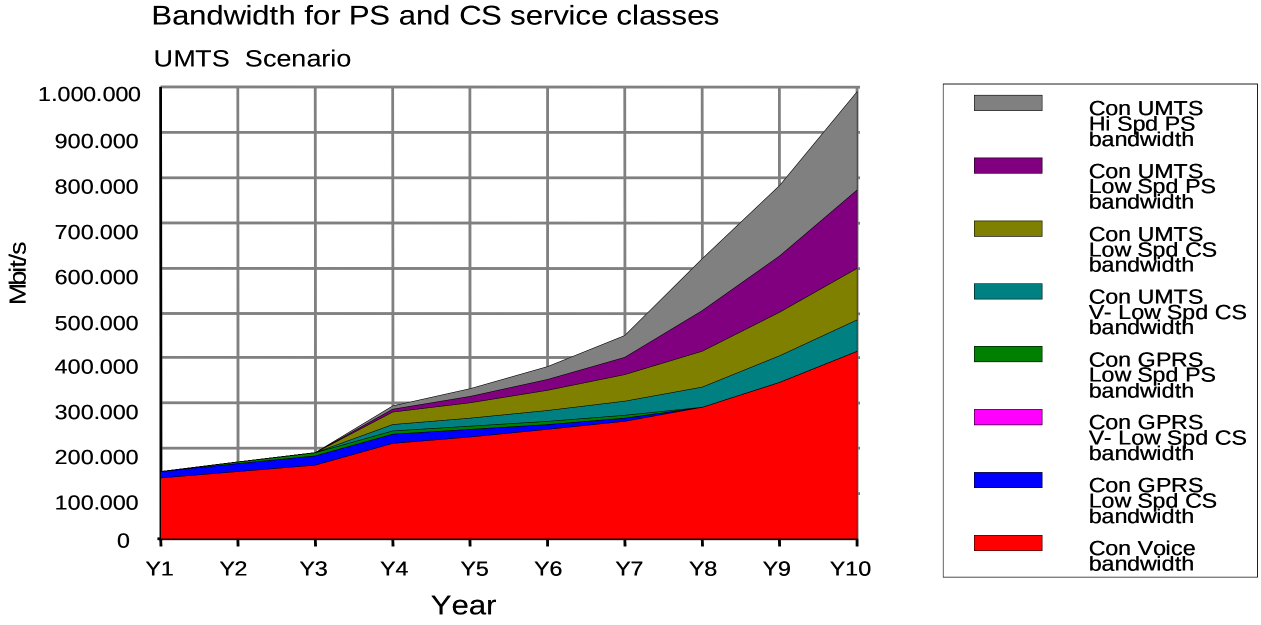


Fig:2.3.3 Bandwith evolution per service class of consumer customers in the UMTS scenario

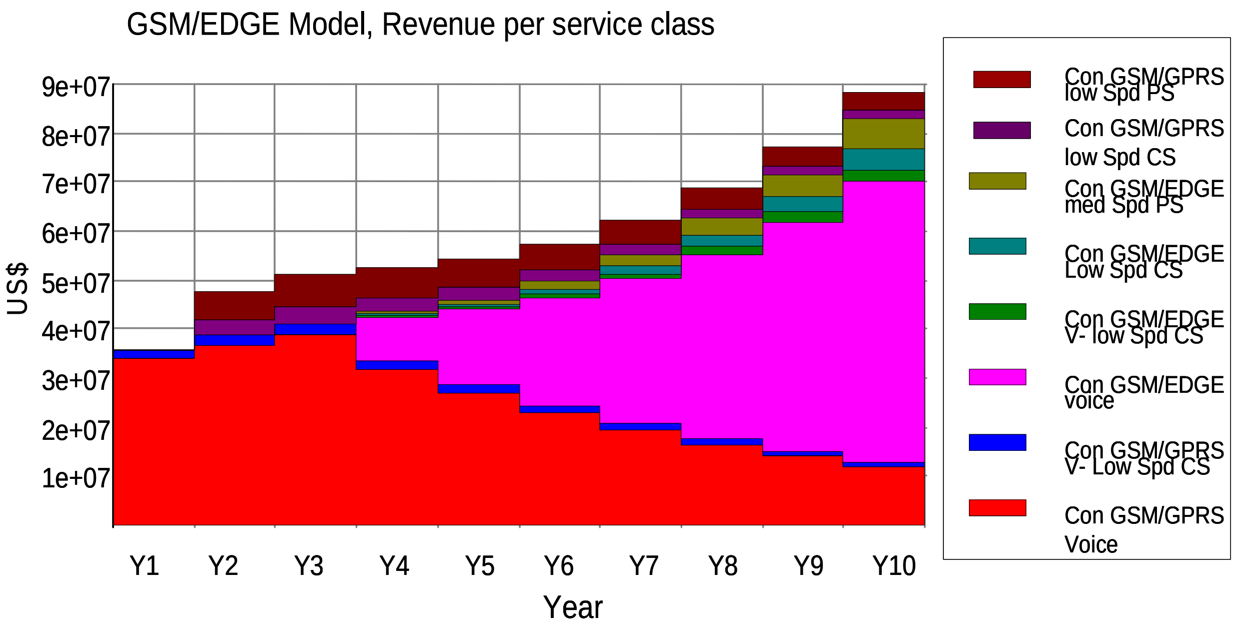


Fig:2.3.4 Service revenue evolution for consumer customers in the EDGE scenario

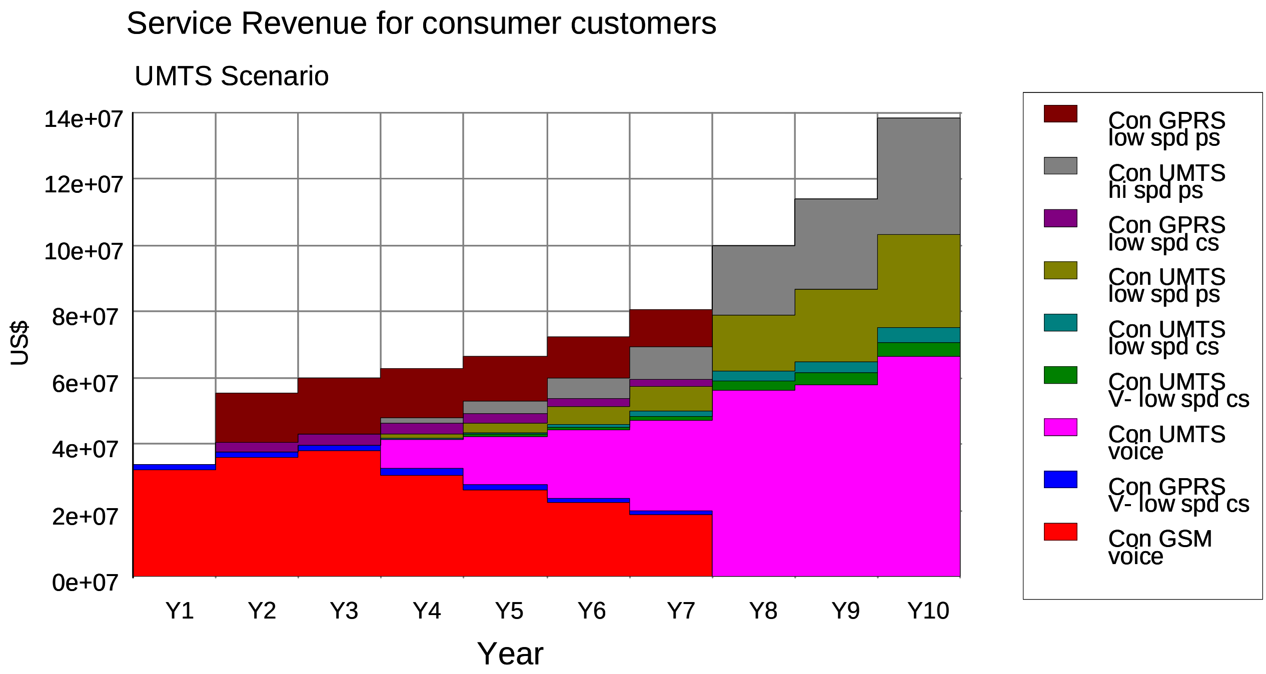
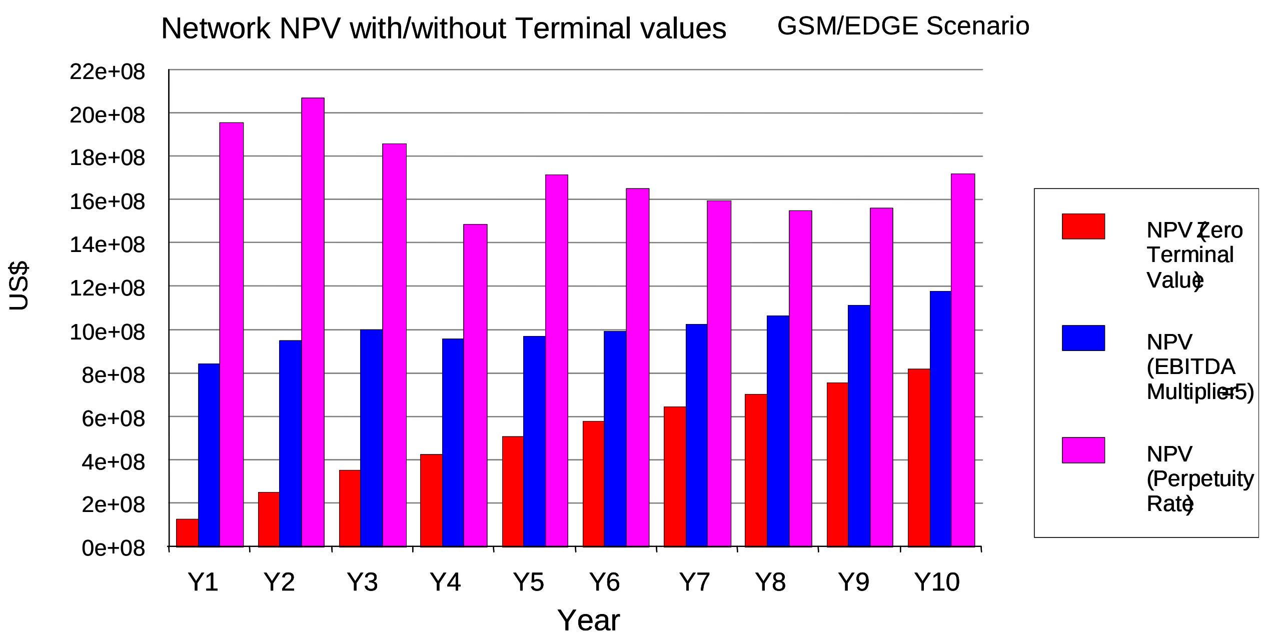


Fig:2.3.5 Service revenue evolution for consumer customers in the UMTS scenario



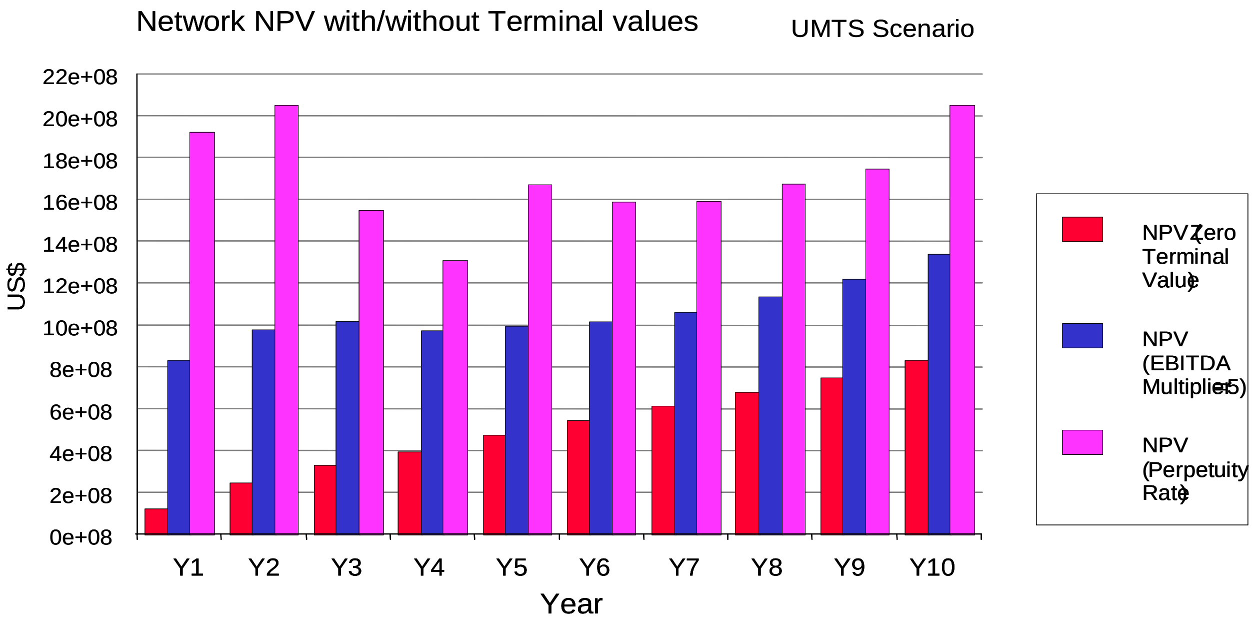


Fig:2.3.7 Net Present Value for the migration case in the UMTS scenario

A2.3.5. Results assessment

- From the analysis of the two scenarios of the study case it is derived that business is feasible in both scenarios with the corresponding economies of scale at the global country market. A simultaneous growth due to customers for better market penetration and to new services with higher bandwidth implies investments and business grow in the two dimensions: Users and Services that implies important investment in infrastructure but facilitates the business feasibility.

- For voice traffic, UMTS solution gives more spectrum efficiency and saving in number of sites for the high density areas of urban type, while EDGE requires less number of sites in the low density areas. In a more detailed observation, important differences appear between the two options from the technical and economical perspectives that are summarized as follows:

  - For data traffic, EDGE solution needs less number of sites at the suburban and mainly at rural areas at the cost of lack of provision of high speed data services and corresponding multimedia applications.

- For multimedia services (voice+ data +video), UMTS solution facilitates a more efficient integration both at user interface and end to end network as well as an easier operational behaviour. This will favour not only to carry new high speed services but also a higher penetration of the medium speed ones.

- Network revenues at the UMTS solution start to surpass the EDGE ones after Y3 of the migration start up to an advantage of 30% at the end of the period. Profitability in relative terms has the same order of magnitude in both cases as the larger the larger revenues in UMTS are compensated by the larger investment required.

- The NPV behaviour, that provides a good complete vision along the 10 years period, shows network values for medium and long term of for UMTS are higher than for EDGE with the additional advantage of a higher generation potential is in the UMTS case and better capability to converge with other radio technologies like the DVB-H

A2.3.6. Case study summary

- A detailed modelling of the mobile network migration towards 3G was performed for a medium size country with business evaluation of the main factors in a 10 years period.

- Case study shows a series of results to support the decision making in the migration of the mobile network. The main key business factor during the period is the important increase in data services, bandwidth and revenues that will compensate the decrease in voice tariffs and support the business profitability and network modernization to be competitive and satisfy customer demands.

  - EDGE case is more conservative and appropriate for the short term with less financing needs but with less multimedia services and business potential.
  - UMTS case represents a more ambitious decision that requires more investments and operating expenses but also will incorporate more multimedia services and higher capacity to take benefits of convergence at network and access speed quality and revenues, especially at medium and long term.

  Decision to follow requires strategic considerations at each country level but will be straightforward with all provided information in this evaluation plus the consideration of financing capabilities and economies of scale at country level.

- Due to the high number of involved parameters, interrelations and the dynamic character of the migration, a powerful tool is required to be able to analyze scenarios and impact of decisions. In this case STEM tool was used and showed high flexibility in all the process.

It is strongly recommended to perform periodical business and technical planning (i.e.: every two years) to adjust market assumptions, assurance of new services adaptation to country culture, quick market promotion and terminals compatibility with new applications to avoid delays in the grow of incomes and exploit business potential.

A2.4. Broadband access planning for major cities

According to the described above scenarios for evolution to NGN in the access network planning studies for introduction of xDSL equipment in the main Bulgarian cities of Plovdiv and Sofia have been carried out.

All studies are performed with specialized network planning tool, VPI AccessMaker, which has unique parameters and capabilities.

VPIaccessMaker™ quickly generates business plans and feasibility studies.

The tool captures subscriber information, models different service combinations and technologies and selects the best technology for the task, to calculate ROI, NPV, cash flow, revenue and an optimized roll out strategy.

For the city of Plovdiv it is foreseen to serve business (SOHO and SME) and limited number of residential customers with xDSL equipment.

Three categories of services are assumed - ADSL Basic (256/64 kbit/s) for residential and SOHO customers, ADSL Gold (512/128 kbit/s) and SHDSL (512 kbit/s) for SME customers.

The forecasting results for the possible customers in the period 2003 – 2007 are shown in Table A2.4.1.

Table A2.4.1.

Year	ADSL Basic	ADSL_Gold	SHDSL	Total
2003	476	402	132	1010
2004	946	763	159	1868
2005	1416	1122	184	2722
2006	1887	1481	211	3579
2007	2357	1842	238	4437

For the period 2003 – 2007 it is assumed to have:100% growth of SOHO customers, 90% growth of SME customers, 10% growth of SHDSL customers.

With the planning tool an optimization of the DSLAM locations is performed. Important condition is to use existing exchange buildings if possible.

Planning results with locations of the DSLAM equipment could be seen on Fig A2.4.1.

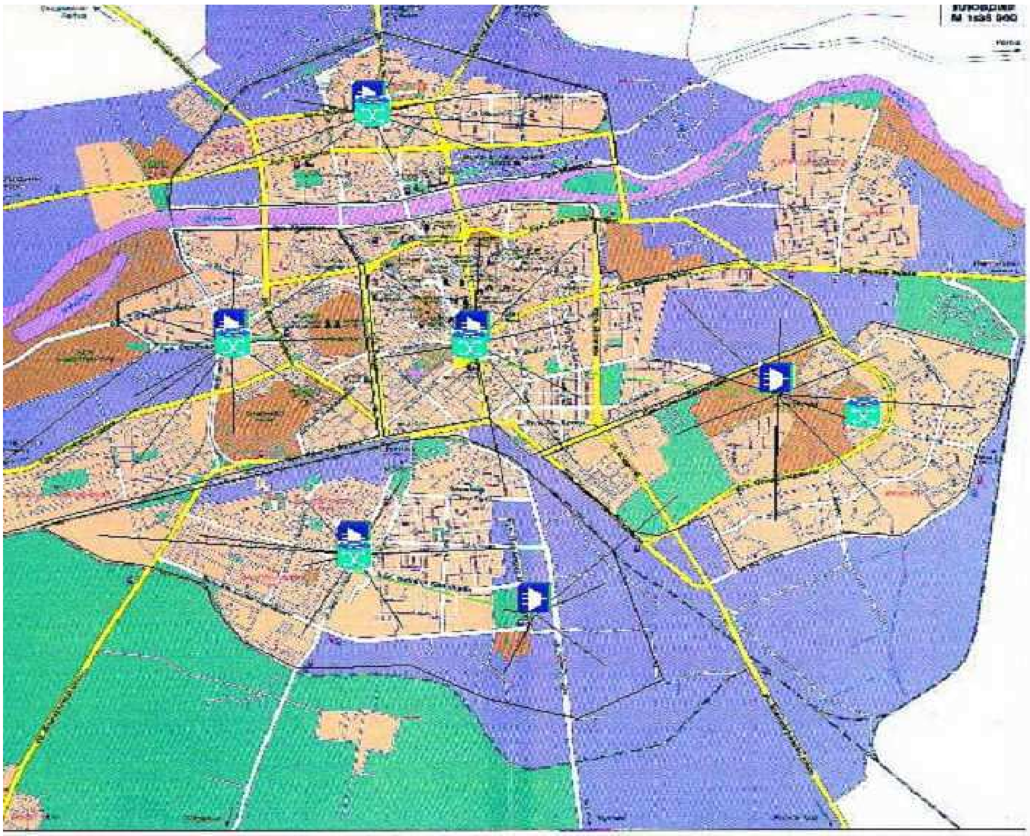
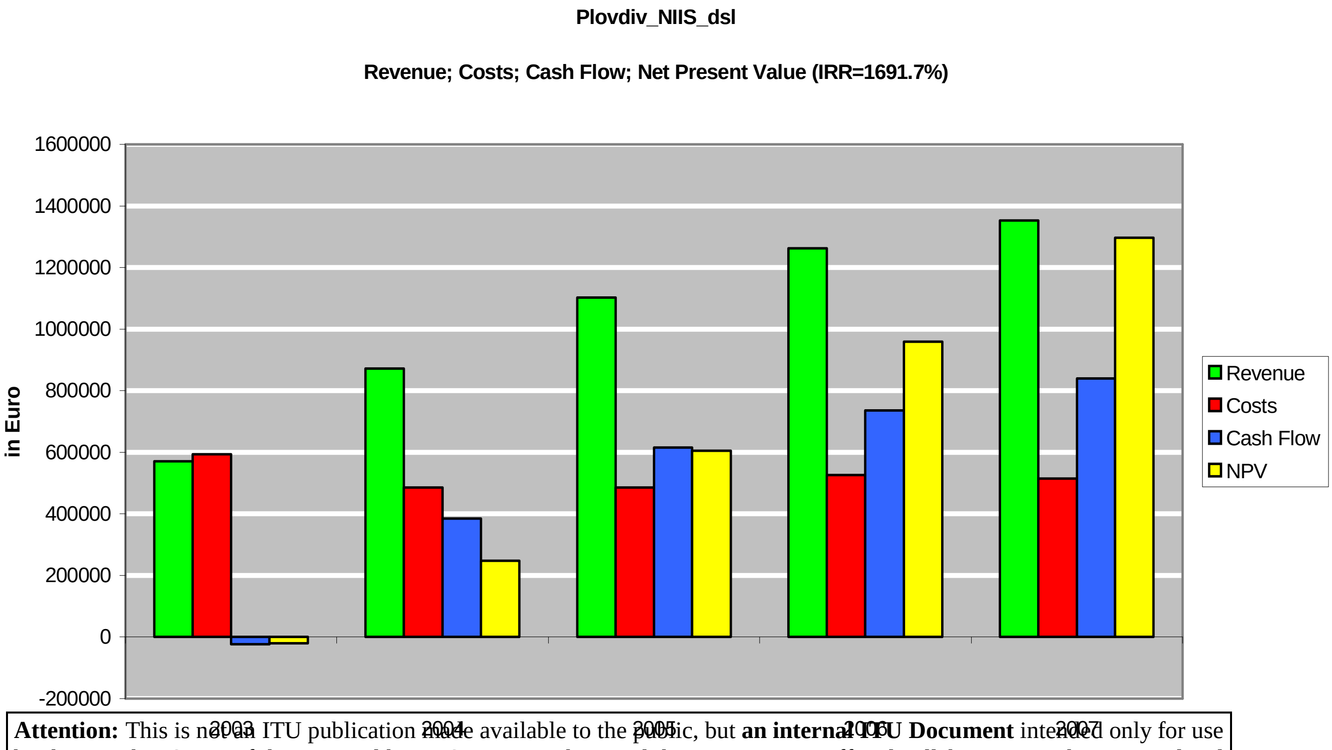


Fig. A2,4,1 Optimized DSLAM locations in Plovdiv

Optimization result shows one DSLAM for each existing exchange area, situated in the exchange building. There is one area, where a second DSLAM will be needed and its location is optimized.

Fig. A2.4.2 Economic analysis of xDSL in the access network of Plovdiv



Planning tool has produced also economic analysis with calculation of costs, revenues, cash flow NPV and IRR, which are shown on Fig. A2.4.2.

For the economic analysis it is assumed 20% decrease of the annual installation fees,
10% decrease of the annual subscription fees and 16% discount factor.

One preliminary study of the city of Sofia was also carried out.

Services to be offered were ADSL-Basic (for residential and SOHO customers), ADSL-Gold for SME customers, SHDSL for business users and FE for Large Enterprises. Forecasted services for the period 2003 - 2007 are shown on Fig. A2.4.3.

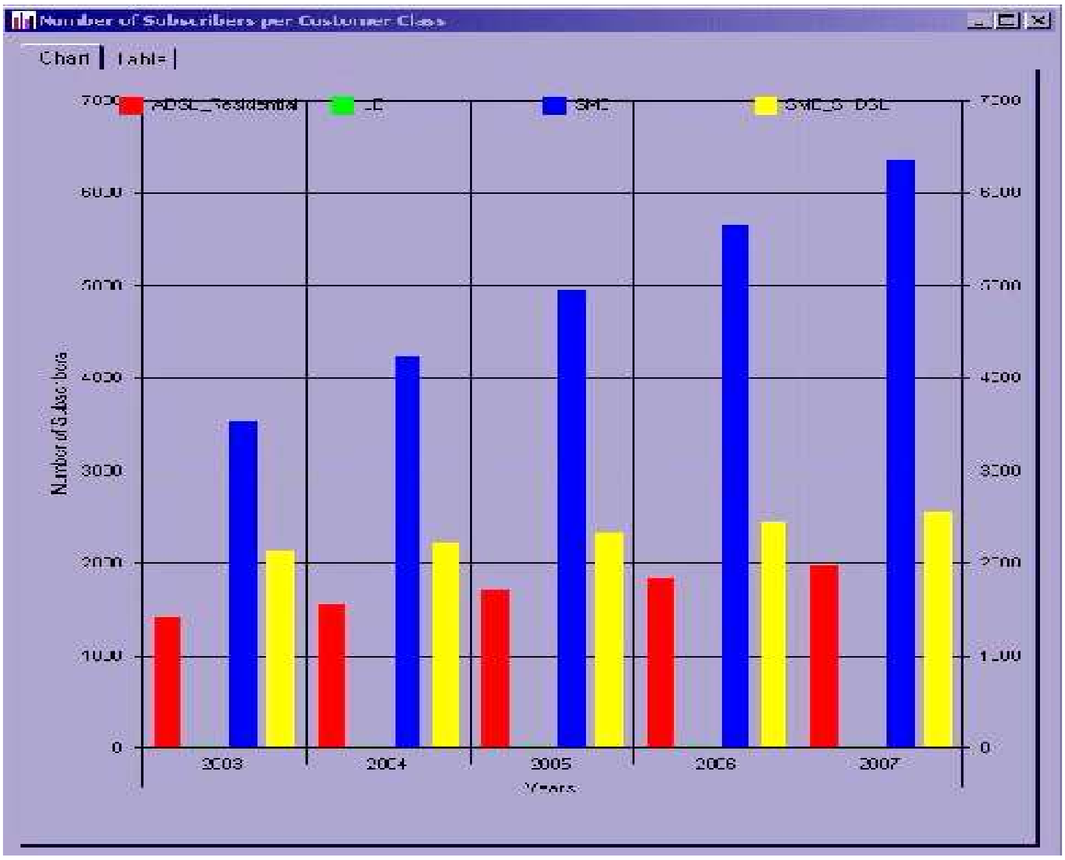


Fig. A2.4.3 Forecasted services for the period 2003 - 2007 in Sofia

Planning results for Sofia with locations of the DSLAM equipment and routers could be seen on Fig A2.4.4.

Network Design is based on xDSL technology with DSLAMs, Routers, FE, GE.

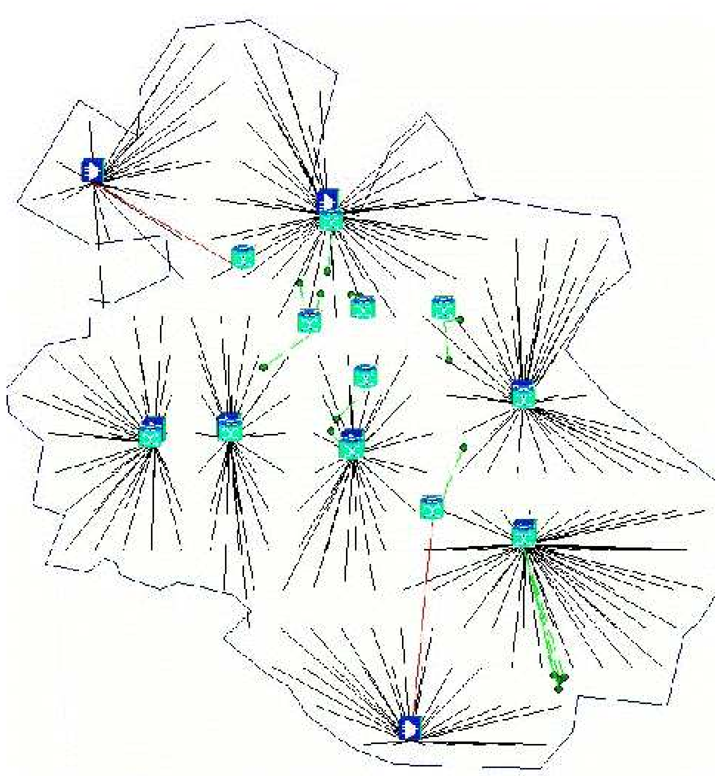


Fig. A2.4.4 Optimized DSLAM and routers locations for Sofia

The optimization result shows that eight DSLAMs could serve the city of Sofia. Locations of 12 routers are also result of the optimization.

A2.5. Voice over IP over WDM

Introduction

This example is meant to demonstrate how one can conduct an integrated network planning across multiple layers and technologies. The example is selected with the consideration that it should represent a typical network planning problem from a nationwide network provider's perspective as realistically as possible. Therefore a representative USA backbone network as shown in figure 1 is chosen as the baseline topology. This topology is used as the underlying physical network where nodes represent traffic generation/switching points and links represent fibers connectivity between nodes. Since there is a real need to model the multi-layer multi-technology complexity of today's telecommunication networks, in this exercise we would like to do a multi-layer network planning: Voice over IP over DWDM. The traffic for the voice over IP service is generated first.

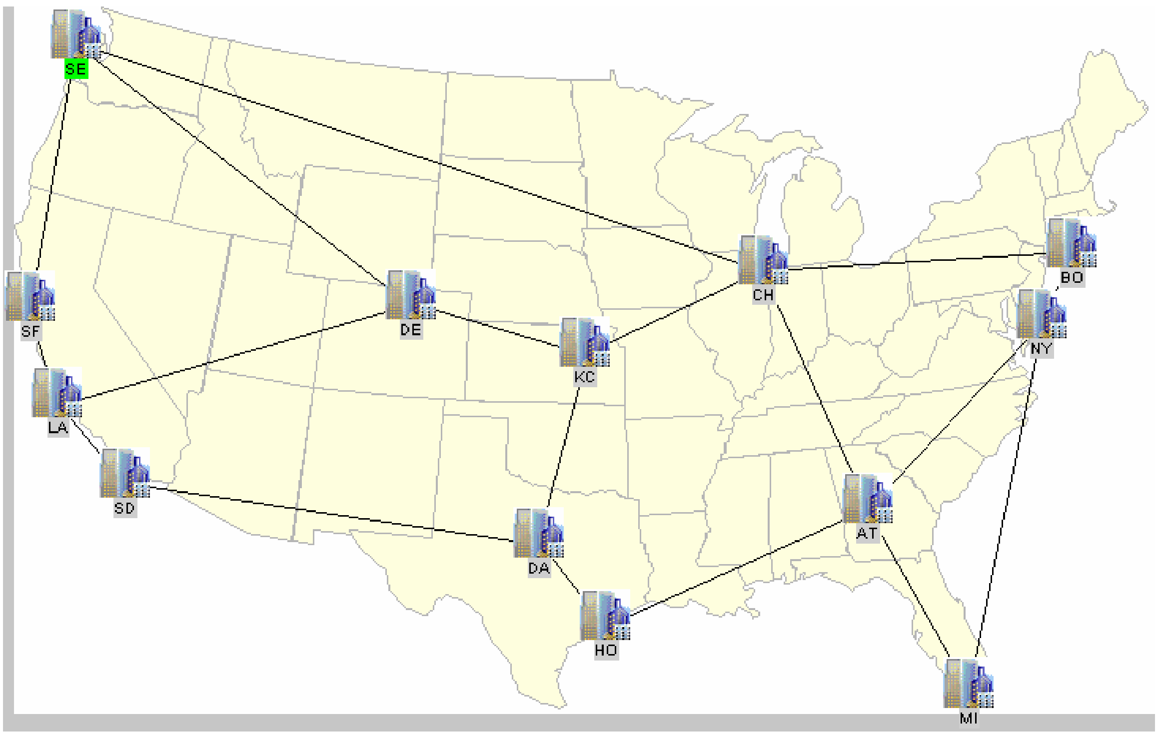


Figure A2.5.1. USA national backbone network

A high speed IP network (with OC192 links) is then planned by incorporating the voice over IP traffic and other native IP traffic such as HTTP, MAIL, FTP etc. Finally the bandwidth pipes of the IP links are viewed as point to point wavelength demands and an optical transport network is designed with switching provided by optical cross connects and multiplexing sections provided by DWDM transmission links. Now the planning of each layer and the inter-layer data exchange are described one by one using the following tools: VPIserviceMaker™IP, VPIserviceMaker™Distribution, and VPItransportMaker™. The IP network design is conducted first by assuming voice over IP traffic matrix available. Then the voice traffic design is performed by VPIserviceMaker^{TM}Distribution. Finally, the optical network design is carried out.

IP Network Planning

Since our aim is to plan an multi-layer carrier scale national backbone network, the IP layer network should represent a typical nationwide internet service provider's (ISP) network. Typical ISPs (such as AOL and AtHome) not only provide Internet connectivity services, but also content services. The Internet connectivity services can create both intra-network and inter-network traffic, while content services only create intra-network traffic between users and content servers. Estimating traffic is the first step and a key to network planning. To model the traffic characteristics of such ISPs, we need to consider the following elements:

1. Subscriber population by city.
2. Subscriber usage profile
3. Content servers
4. Intra and inter network traffic percentages.

In terms of types of traffic, we will consider the native IP services such as HTTP, EMAIL, FTP etc., as well as VoIP.

We choose a typical US ISP backbone network as show in figure A2.5.2. with reference to the physical fiber topology as shown in figure A2.5.1. The nodes of the backbone network are points of presence (POP) at all major cities. In this example there are 13 POPs, representing 13 major metropolitan areas as shown in the US map. Each POP may be composed of a number of routers of different types. Here we use only one router icon to represent each POP because what we want to model is the backbone network, not the intra POP structure.

To model subscriber population by city, we create a computer group icon for each city representing all the users served by the ISP in the city. Each computer group is named by adding the suffix NET to the city's name as the computer group also represents the regional access network. We also select a few cities as the locations of content server centers of the ISP and represented by the server icon. Similarly server names are city names followed by a suffix -SVR. To be more specific, the ISP has three web server centers: one in Los Angeles (LA-SVR), one in Denver (DE-SVR), and one in Atlanta (AT-SVR). It has two other server centers to provide other services (FTP, Email, News etc.): one in Chicago (CH-SVR) and one in Dallas (DA-SVR).

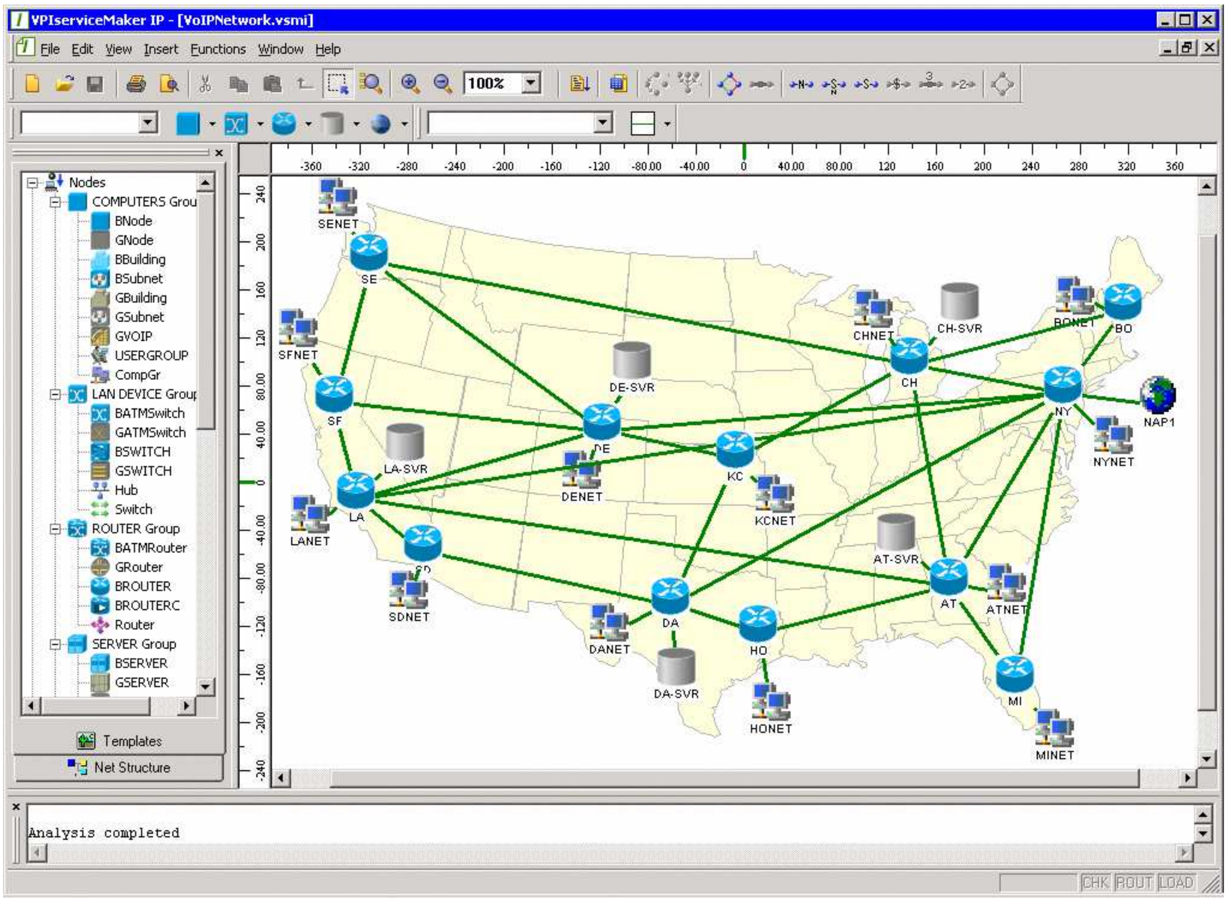


Figure A2.5.2. A Sample US Nationwide ISP Backbone Network

We assume that the network has only one gateway to the global internet and therefore the network has no transit traffic. The gateway is located at New York and is connected to other peering ISPs at a peering site named NAP1 (Network Access Point). The peering site is represented by the globe icon. Please note that the tool is only focused on the internal network of an ISP. Therefore it does model the actual locations of peering sites. In fact the globe icon represents the global internet at large. The planning scope of the tool ends at the outgoing link of each gateway. Anything beyond that is hidden into global internet. We could specify two or more gateways, such as one on the east coast and one on the west coast, by setting the routers to be AS boundary routers and connecting them to the globe icon. But in that way we need to consider transit traffic which neither originate nor terminate within the ISP's network. Since in this demo we are only interested in seeing the network planning impact by IP traffic created by the ISP's own subscribers, we consider only one gateway.

In each city, the computer group, the server (if exits) and the global gateway (if exits) are directly connected to the POP of the city.

The topology shown in figure 2 is the initial set of links inter-connecting connecting POPs which we assumed based on the underlying fiber topology (shown in figure 1) and estimated traffic pattern. The IP topology normally has richer connectivity than the underlying physical network because of desirable express (logical) links connecting distant routers. Topology optimization is an iterative process. If a link is not used by the design, it should be deleted.

Similarly if a link's load is too heavy, other links should be added to off load the congested link. Since we are planning an IP over DWDM network, we assume all links are high speed links with the OC192 granularity.

The computer group representing the subscribers of each city not only enables us to model the subscriber population in the city, but also allows us to specify the usage profiles of these subscribers from which network traffic can be derived. We first need to specify subscriber population for each city by the following assumptions. Among the 13 cities, 7 of them are large cities and each of them has a subscriber population of half million: San Francisco (SF), Los Angeles (LA), Dallas (DA), Houston (HO), Chicago (CH), Atlanta (AT), and New York (NY). The other 6 cities are mid sized and each has a subscriber population of a quarter million: Seattle (SE), San Diego (SD), Denver (DE), Kansas City (KC), Boston (BO), and Miami (MI).

To set the parameters for each computer group, just double click the icon. As an example, figure A2.5.3 shows the parameter settings of computer group NY-NET. We should pay special attention to the following parameters:

- The user population, which is 500,000 for New York.
- Use of services: For all computer groups, we choose all available services including VoIP.
- Peak rate: For all cities, we assume subscribers only use telephone dial up to access the network (like AOL) and therefore the peak rate is set at 64 Kbits/s uniformly. The only exception is VoIP. Each channel of VoIP is considered as a constant bit rate connection and the bit rate is stored in the peak rate field. The VoIP bit rate is assumed to be 8 Kbits/s after taking into consideration bandwidth saving factors such as silence suppression and statistical multiplexing.
- Percent: It is a breakdown of amount of usage across all the services for an average user. For this demo example these percentages are also uniformly set for all computer groups because it is not our intention to model user behavior differences across cities. Here the only difference between cities is the one in subscriber population.
- Individual service settings (not shown): Specific settings for each individual service can be made by clicking the corresponding type. Here we set that all http traffic are equally distributed across the three servers, and a 50-50 split between internal traffic and external traffic. Similar assumptions are set for other services except for VoIP.
- For VoIP, we check the Use Voice Matrix box because VoIP traffic will be provided by a separate voice matrix which will be generated by VPIserviceMaker^{TM}Distribution. For now we just assume the voice matrix is available. The next section will describe how the matrix is created.

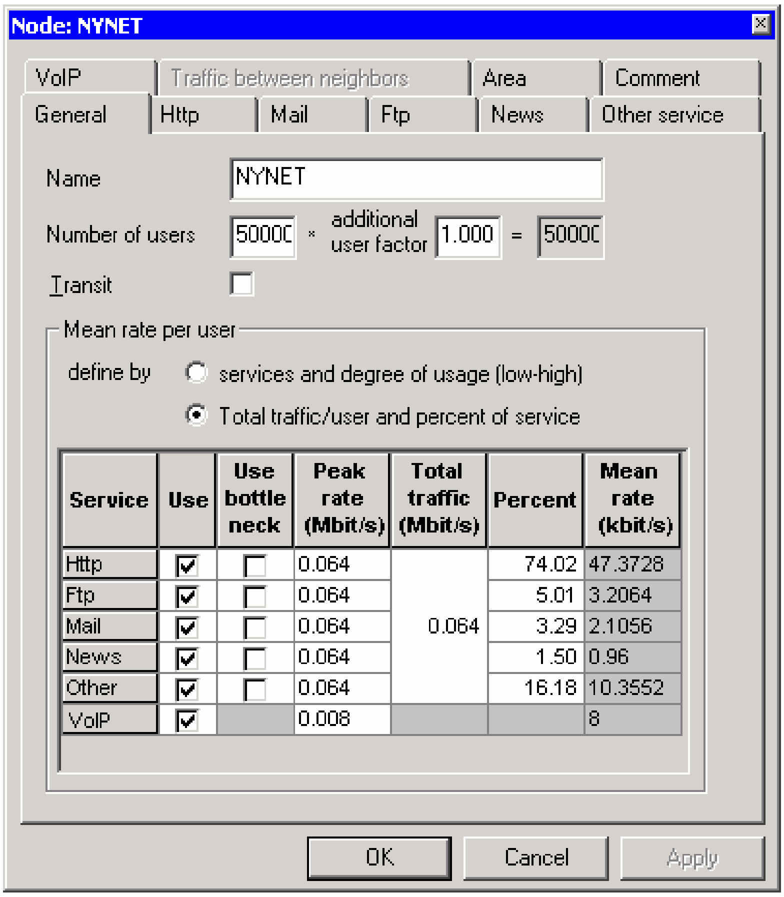


Figure A2.5.3. Sample Computer Group Setting

To set parameters for POPs, double click the router icons. As an example, figure A2.5.4 shows the parameter settings of NY POP. The maximal switching capacity is set to 500 Gbit/s. Currently the largest core router product available in the market place is about 100 Gbit/s. But here the switching capacity does not represent the capacity of a single router, but that of a POP which normally consists of multiple routers. For this demo we assume all the seven large POPs uniformly have a switching capacity of 500 Gbit/s and all the six medium POPs uniformly have a switching capacity of 100 Gbit/s. It should also be pointed out that for the purpose of this design the OSPF area assignment is not important. All the backbone links should be in area 0. The access networks as represented by computer groups and intra-POP networks should have other area numbers but our interest is not to model the internal structures of the access networks or POP offices. Therefore one backbone area suffices.

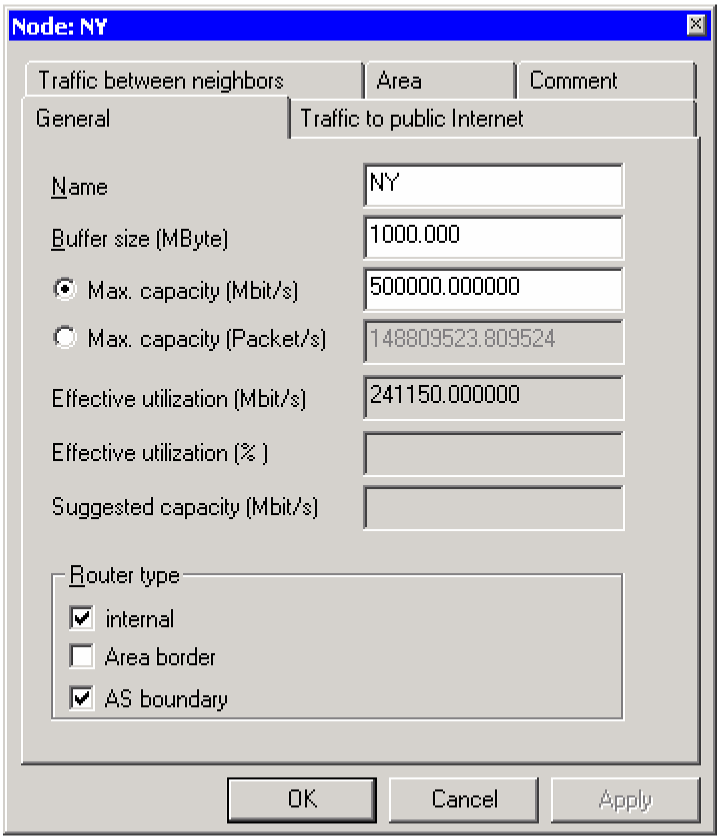


Figure A2.5.4. Sample POP Setting

Server settings are not described here because they are the simplest and can be populated in a straightforward way.

The last step of data input is to import the VoIP matrix. To do so open Edit menu and select VoIP Matrix option. A table pops up which shows the current VoIP matrix. Here only the computer group nodes which participate in VoIP service (the parameters are set) are included in the matrix. For this example all the computer groups participate VoIP.

We assume there already exists a voice matrix, which is generated by VPIserviceMaker™Distribution (to be described in the next section). To import, open the File menu of the table, select Import and choose Replace Values option, the matrix will be populated by the imported values as shown in figure A2.5.5.

Matrix: Traffic Matrix - Simultaneous VoIP user

File Edit

from/to	ATNET	BONET	CHNET	DANET	DENET	HONET	KCNET	LANET	MINET	NAP1	NYNET	SDNET	SENET
ATNET		1250.00	2500.00	2500.00	1250.00	2500.00	1250.00	2500.00	1250.00	0.00	2500.00	1250.00	1250.00
BONET	1250.00		1250.00	1250.00	625.00	1250.00	625.00	1250.00	625.00	0.00	1250.00	625.00	625.00
CHNET	2500.00	1250.00		2500.00	1250.00	2500.00	1250.00	2500.00	1250.00	0.00	2500.00	1250.00	1250.00
DANET	2500.00	1250.00	2500.00		1250.00	2500.00	1250.00	2500.00	1250.00	0.00	2500.00	1250.00	1250.00
DENET	1250.00	625.00	1250.00	1250.00		1250.00	625.00	1250.00	625.00	0.00	1250.00	625.00	625.00
HONET	2500.00	1250.00	2500.00	2500.00	1250.00		1250.00	2500.00	1250.00	0.00	2500.00	1250.00	1250.00
KCNET	1250.00	625.00	1250.00	1250.00	625.00	1250.00		1250.00	625.00	0.00	1250.00	625.00	625.00
LANET	2500.00	1250.00	2500.00	2500.00	1250.00	2500.00	1250.00		1250.00	0.00	2500.00	1250.00	1250.00
MINET	1250.00	625.00	1250.00	1250.00	625.00	1250.00	625.00	1250.00		0.00	1250.00	625.00	625.00
NAP1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00		0.00	0.00	0.00
NYNET	2500.00	1250.00	2500.00	2500.00	1250.00	2500.00	1250.00	2500.00	1250.00	0.00		1250.00	1250.00
SDNET	1250.00	625.00	1250.00	1250.00	625.00	1250.00	625.00	1250.00	625.00	0.00	1250.00		625.00
SENET	1250.00	625.00	1250.00	1250.00	625.00	1250.00	625.00	1250.00	625.00	0.00	1250.00	625.00	
SFNET	2500.00	1250.00	2500.00	2500.00	1250.00	2500.00	1250.00	2500.00	1250.00	0.00	2500.00	1250.00	1250.00

Figure A2.5.5. Imported VoIP matrix

Now we are ready to run the design functions. We can either do design step by step by running functions (analysis, routing and effective bandwidth calculation) one at a time, or run all of them together one clicking on the “run all the planning functions” button. In order to perform the dimensioning for all defined services select the services as indicated in figure A2.5.6.

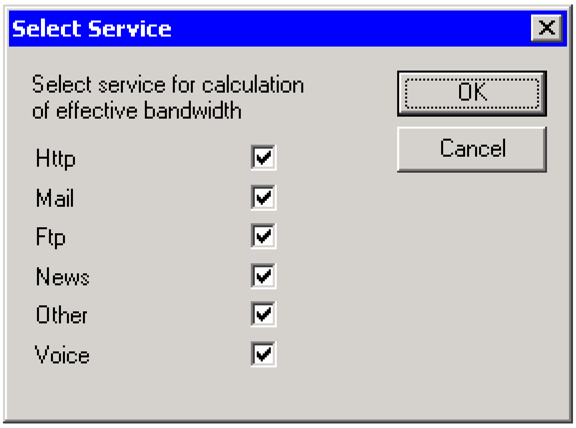


Figure A2.5.6. Select Service Dialog

The design results are compiled into a set of detailed reports. The results can also visually displayed in many different ways by selecting corresponding options under the View menu. For example, figure A2.5.7 shows the resulting IP network dimensioning. The IP network is sized in two dimensions: link bandwidth requirement and router (POP) switching capacity requirement. The number on each link indicates how many OC12s the link needs in order to carry the traffic with adequate performance. The color codes of the routers indicate the router utilization. Green and purple mean that no one is over loaded. Both routers and links can be viewed by utilization, by traffic load, and by suggested capacity.

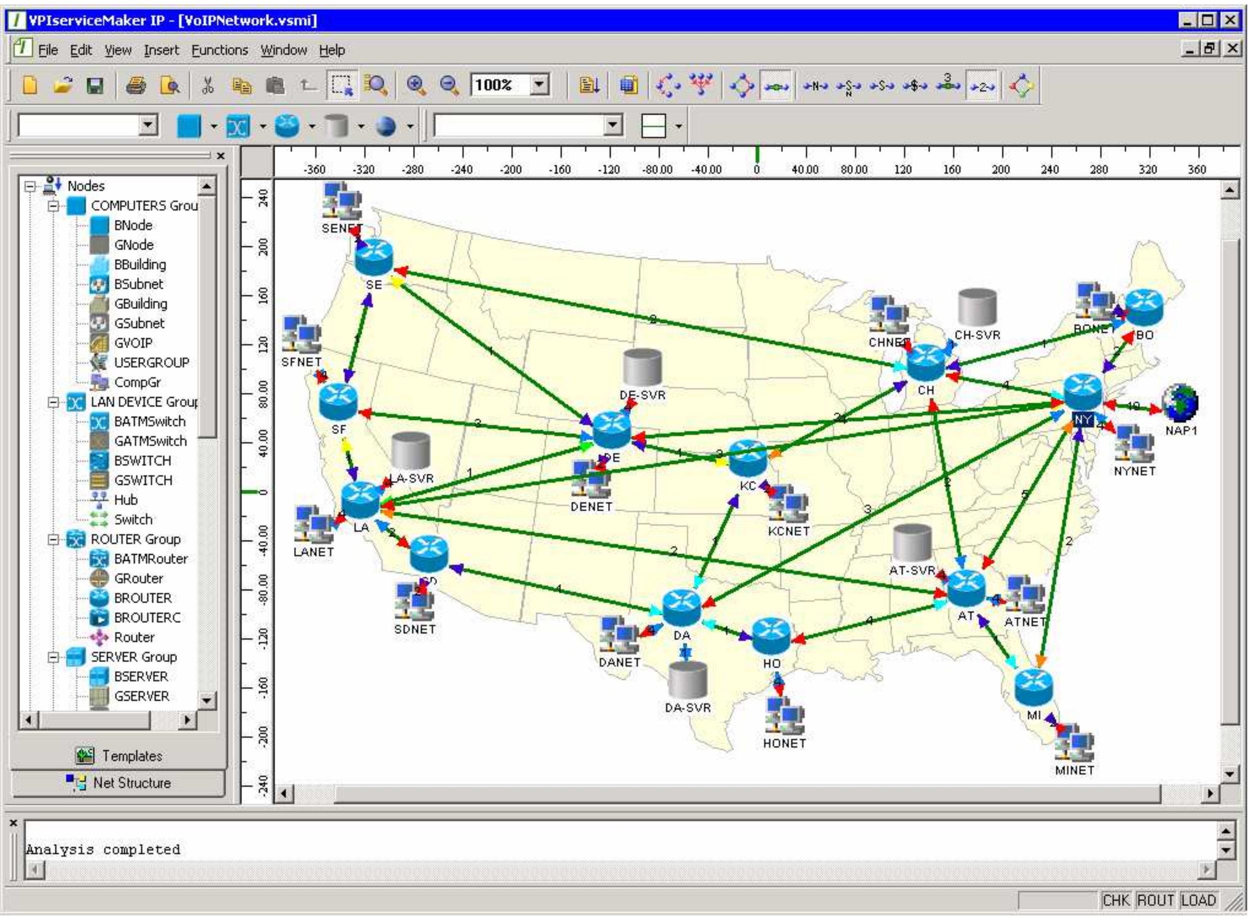


Figure A2.5.7. IP Network Dimensioning

After the design, network bottlenecks and under-utilized resources can be identified. Topology improvement can be made by simply looking at link load. For example, the link between SE and SF is very much under-utilized and can be removed. Some scenario analysis can also be conducted. To simulate the impact of a failed link, simply disable the link and rerun the tool. To project the impact of the subscriber growth at certain cities, go to the corresponding computer groups and re-set the user population, or simply use the "additional user factor" feature to scale up subscriber population.

Finally once the IP network design is satisfactory, we would like to design an underlying transport network to provide the required bandwidth for IP links. To do so the first step is to export the IP network to VPItransportMaker™ using the IP tool's export feature (Export to TransportMaker under file menu). A new TransportMaker™ project will be created with IP network links being treated as a traffic matrix of point to point bandwidth demands for the transport network design.

VoIP Matrix Generation

Here we will show how to use VPIserviceMaker™Distribution to generate a matrix of point to point voice traffic as VoIP traffic for the IP network design. Normally the units of traffic quantities are Erlangs but for VoIP we interpret them as simultaneous voice channels between a pair of nodes.

Since VoIP traffic is only between computer (subscriber) groups, we do not need to consider the other nodes. The corresponding network with all VoIP relevant nodes are shown in figure A2.5.8.
The network data can be loaded from project "VoIP_Network_Dist".

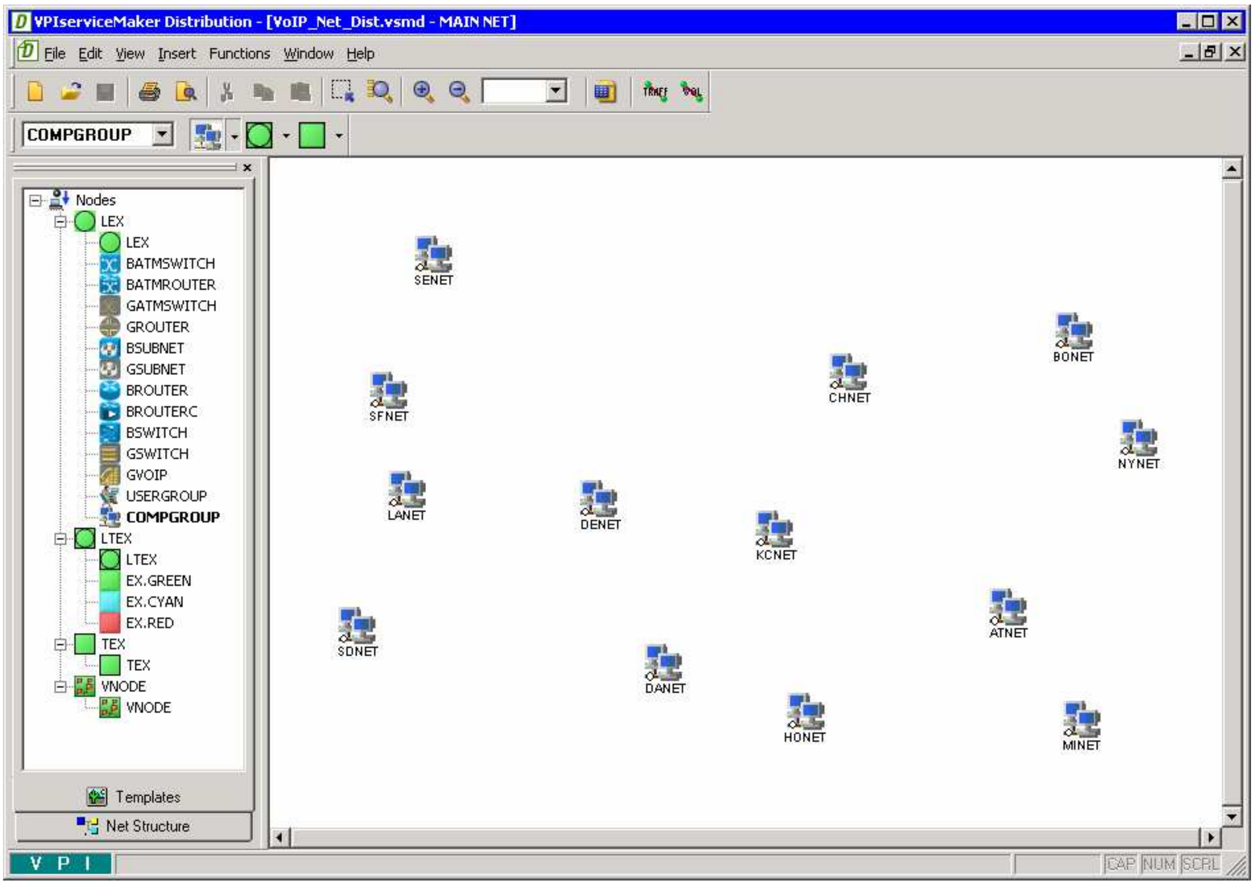
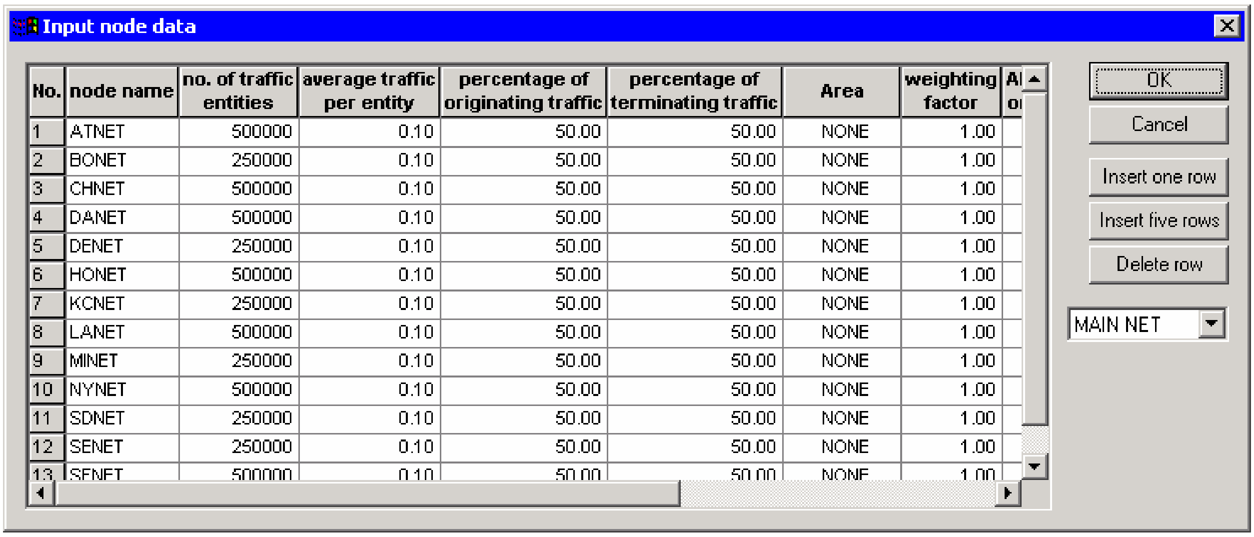


Figure A2.5.8. The set of VoIP nodes

Now we are ready to design the voice traffic. First open the Functions menu and select Input Data, an input table pops up asking for design parameters, as shown in figure A2.5.9. Since it is assumed that the amount of voice traffic between two cities is proportional to the populations of the two cities, the subscriber population for each city should be specified here. We simply use the same population numbers as those we used in the IP project. We also need to specify other parameters such as average traffic per subscriber and call originating/termination percentages which are all shown in figure A2.5.9.



No.	node name	no. of traffic entities	average traffic per entity	percentage of originating traffic	percentage of terminating traffic	Area	weighting factor	A| o
1	ATNET	500000	0.10	50.00	50.00	NONE	1.00	
2	BONET	250000	0.10	50.00	50.00	NONE	1.00	
3	CHNET	500000	0.10	50.00	50.00	NONE	1.00	
4	DANET	500000	0.10	50.00	50.00	NONE	1.00	
5	DENET	250000	0.10	50.00	50.00	NONE	1.00	
6	HONET	500000	0.10	50.00	50.00	NONE	1.00	
7	KCNET	250000	0.10	50.00	50.00	NONE	1.00	
8	LANET	500000	0.10	50.00	50.00	NONE	1.00	
9	MINET	250000	0.10	50.00	50.00	NONE	1.00	
10	NYNET	500000	0.10	50.00	50.00	NONE	1.00	
11	SDNET	250000	0.10	50.00	50.00	NONE	1.00	
12	SENET	250000	0.10	50.00	50.00	NONE	1.00	
13	SFNET	500000	0.10	50.00	50.00	NONE	1.00	

Figure A2.5.9. Input Data

Next choose the single matrix method by clicking the Single option and select the homogeneous matrix model, as indicated in figure A2.5.10.

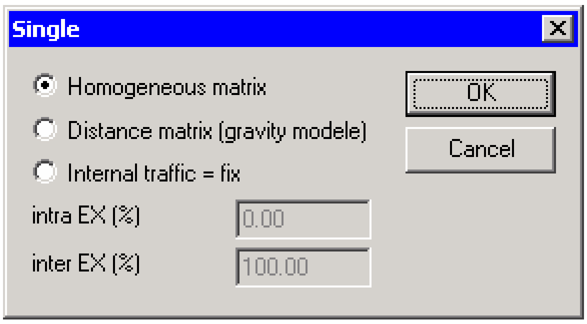


Figure A2.5.10. Input Data

The tool will create a matrix of point to point voice traffic. The traffic matrix is shown graphically in figure A2.5.11 where each line represents the voice traffic between the two end nodes and the number on the line indicate the number of simultaneous voice channels. This voice matrix is then exported to a text file in the service matrix format to be used by the IP network design, by selecting File → Export → Service Matrix.

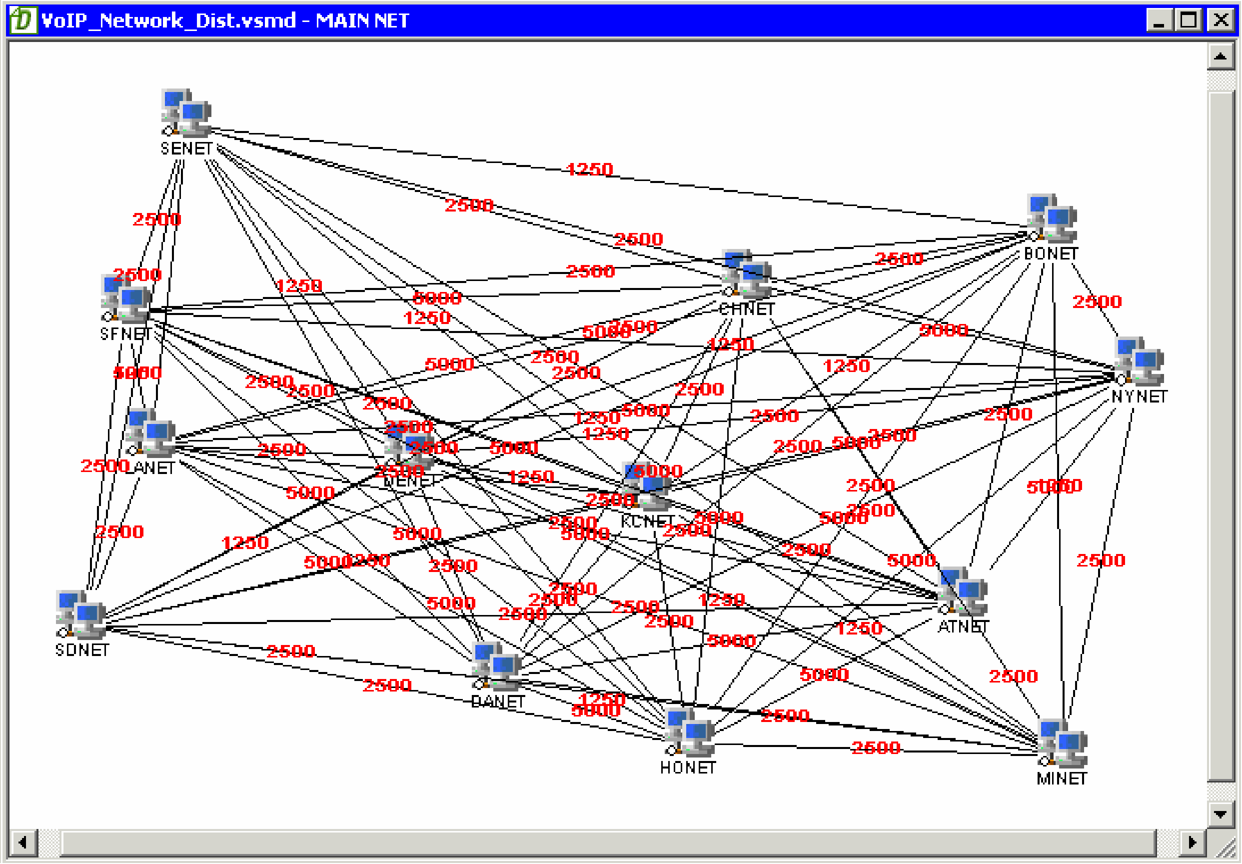


Figure A2.5.11. Voice traffic matrix

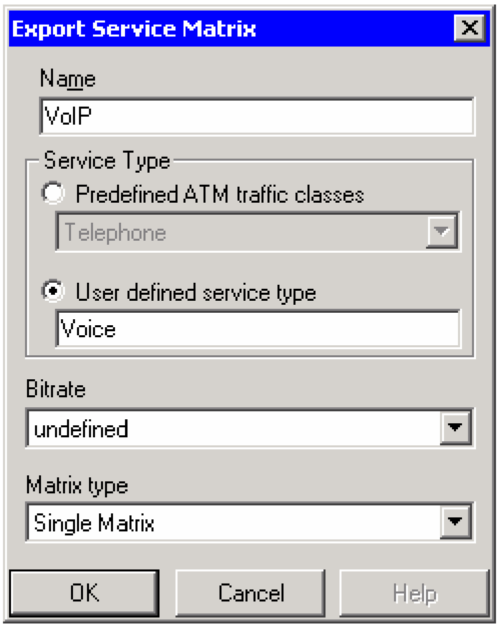


Figure A2.5.12. Export service matrix

Transport Network Planning

Figure A2.5.13 shows the newly created VPItransportMaker™ project by exporting the IP network. The IP links are now grouped into a traffic matrix called STM64 (to indicate their signal rate of OC192). The server window (right) initially contains only nodes, without OMS links. The OMS links can be either manually created or imported from a topology file. The client window (left) represents graphically the traffic matrix, where each link represents a demand and the number by the link represents quantity in OC192 (STM64) units.

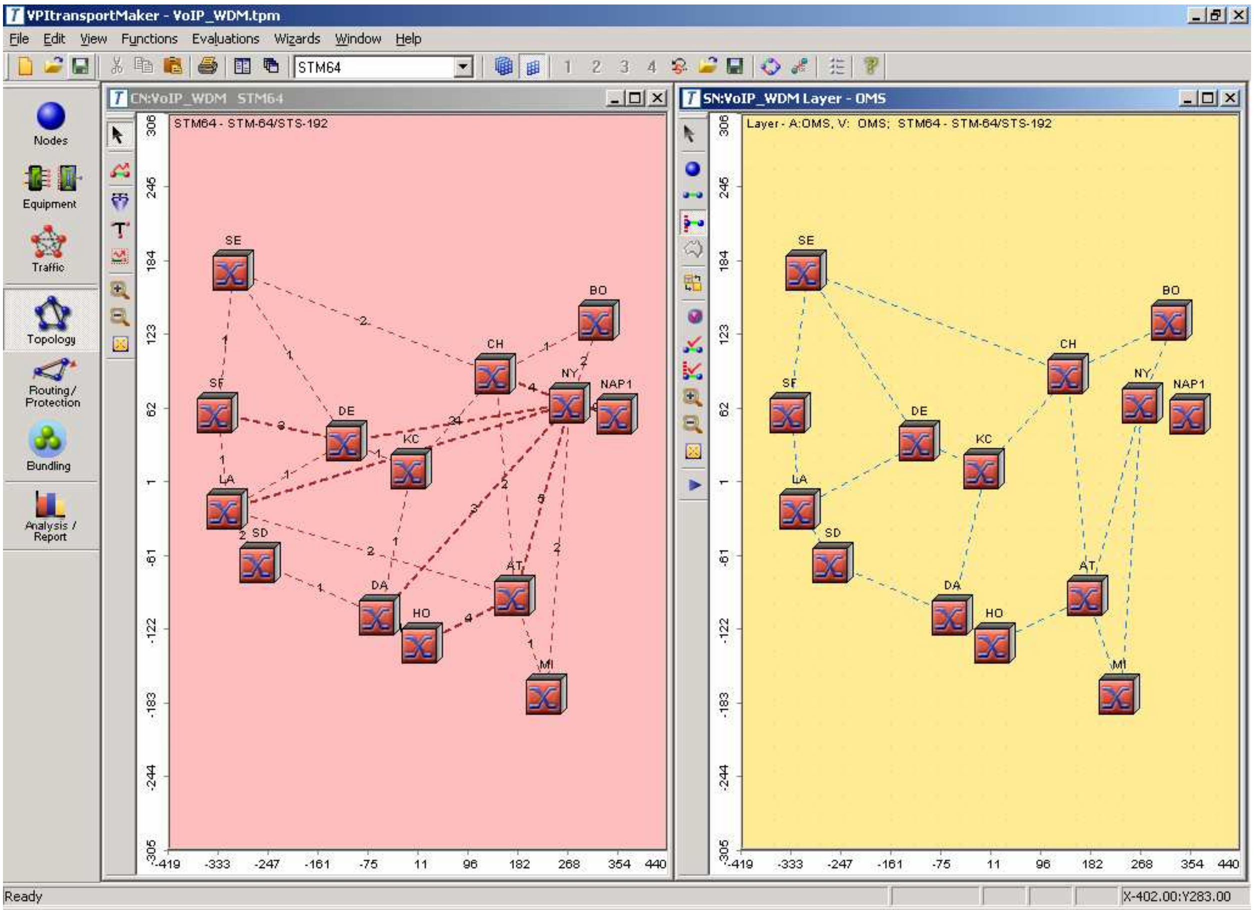


Figure A2.5.13. Transport project exported from IP

We want to plan a two layer IP over optical network. For the optical layer, we want to design a mesh network with wavelength routing and protection capabilities. Further down, DWDM systems provide point to point optical multiplexing section links. For this purpose, when we create the server layer topology (the right window), we need to specify that all links are multiplexing links at the OMS layer.

Before create the OMS links, we want to set the number of channels available to each DWDM system. We go to Bundling mode, open the Edit menu, select Set Bundling Factors option. A parameter input window shown in figure A2.5.14 appears, we set No. of Channels equals to 40 for DWDM system.

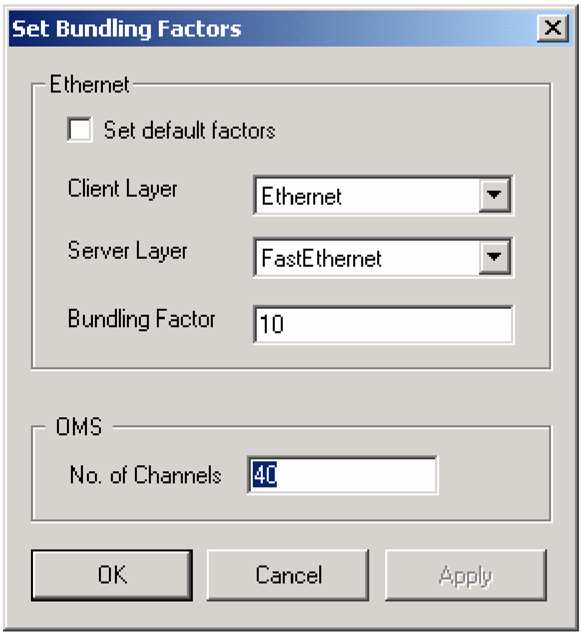


Figure A2.5.14. Set Bundling Factors

To design a wavelength routed optical network, we go to Routing/Protection mode, open the Functions menu, select the Wavelength Routing option. A parameter input window pops up as shown in figure A2.5.15. For the Algorithm Options, we choose Planning for better optimization (at the price of more computation). We select no wavelength conversion because we assume that optical cross connects are not capable of converting wavelengths based on the reality of today's technology. We also select the 1+1 path protection option because no protection is considered when we design the IP network. After the run, the design result is compiled into a report.

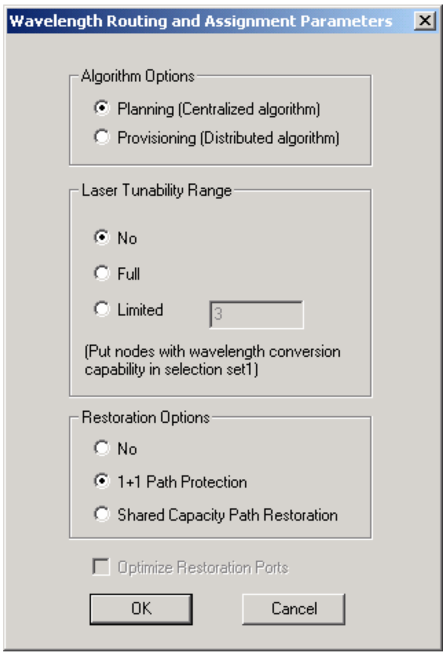


Figure A2.5.15. Wavelength routing parameters

The design can also be visually displayed on the GUI. Figure A2.5.16 shows the network of DWDM facilities where each dashed line represents one DWDM system and the number by the line indicate the number of wavelengths carried by the system. It can be seen that there is only one line over each fiber link, which means there is only one DWDM system deployed over each fiber route. This is because each DWDM system offers a capacity of 40 wavelengths (OC192s) which is more than sufficient for our traffic.

Figure A2.5.16. Wavelength load on each link

By opening the details table for each demand and select routing option, the routing of a demand can also be displayed. Figure A2.5.17 and figure A2.5.18 show the working route and the protection route of the wavelengths from Los Angeles to New York respectively.

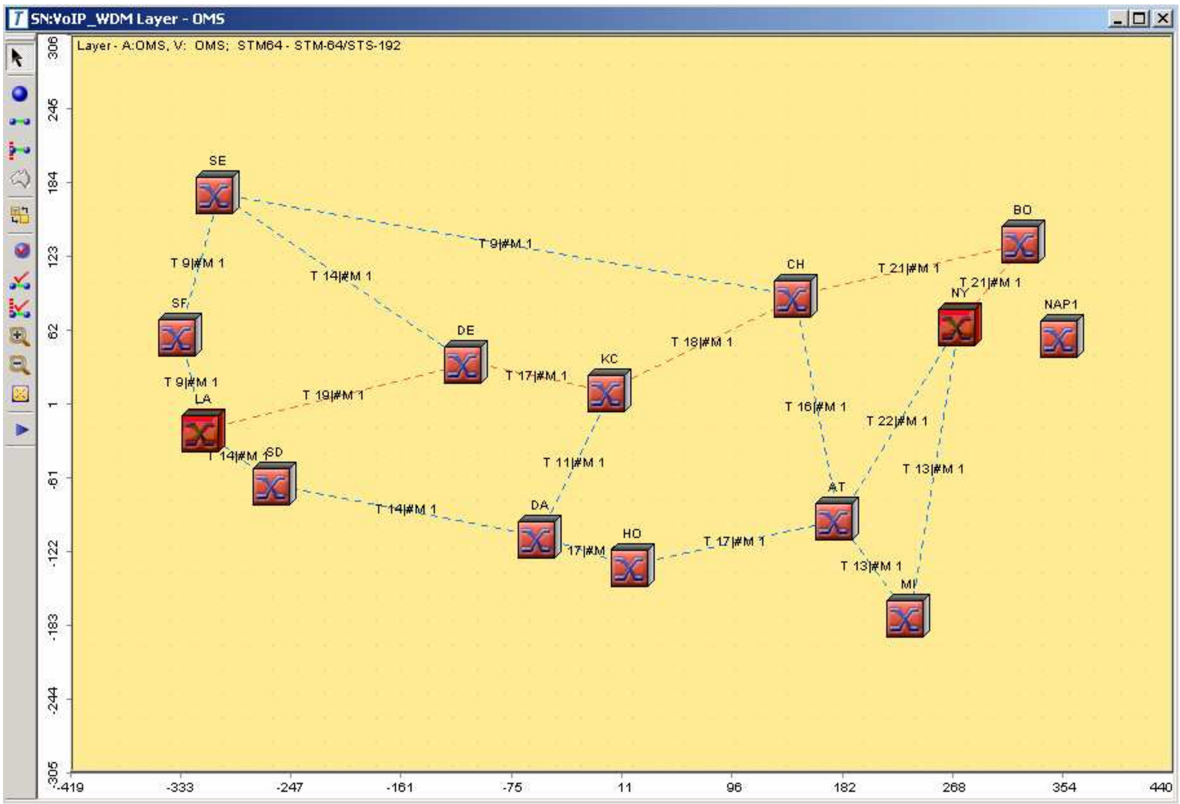


Figure A2.5.17. Working path for demand between LA and NY

Figure A2.5.18. Protection path for demand between LA and NY

Summary

In this design exercise we have shown how to leverage the VPI software and the complete suite of design tools it provides to do an integrated network planning across multiple layers. We have demonstrated a smooth end to end network planning procedure by showing how different tools can inter-work by exporting one tool's output to another tool as input.

So far in this design exercise the data flows are one way: from traffic distribution to service layer design to transport layer design. It is also possible to have a feedback process from transport layer to the service layers. For example, by looking at the spare capacity of the transport layer, one can estimate how to expand a service network, or to adjust transport bandwidths allocated to different service layer networks. Another way is to conduct what-if scenario analysis on potential impact of any service layer changes over the transport layer.

A2.6. Mobile network coverage

Case Study - GSM Network Planning Papua New Guinea

<u>Introduction</u>

The following document presents the results of a case study performed for a GSM 900 Network in Papua New Guinea. The scope of the study is to determine the number of sites needed to cover the urban area and the costal region around Port Moresby, the urban area and the catchment area of Popondetta and the connection road between these two cities (See Figure 4). Additionally the fixed network has been planned for the found radio sites. Result of the fixed network planning is the number of needed links divided into long hop- and short hop links.
The planning task has been split up in three different phases.
Please note that the case study has been based on rough mapping data and hypothetical site locations; no site survey has been done. Operational aspects like availability of power and site accessibility has not been included in the site selection process. An examination of the general economical feasibility of implementing a mobile network in the used regions (especially the connection roads and the rural areas) has not been subject of the study.

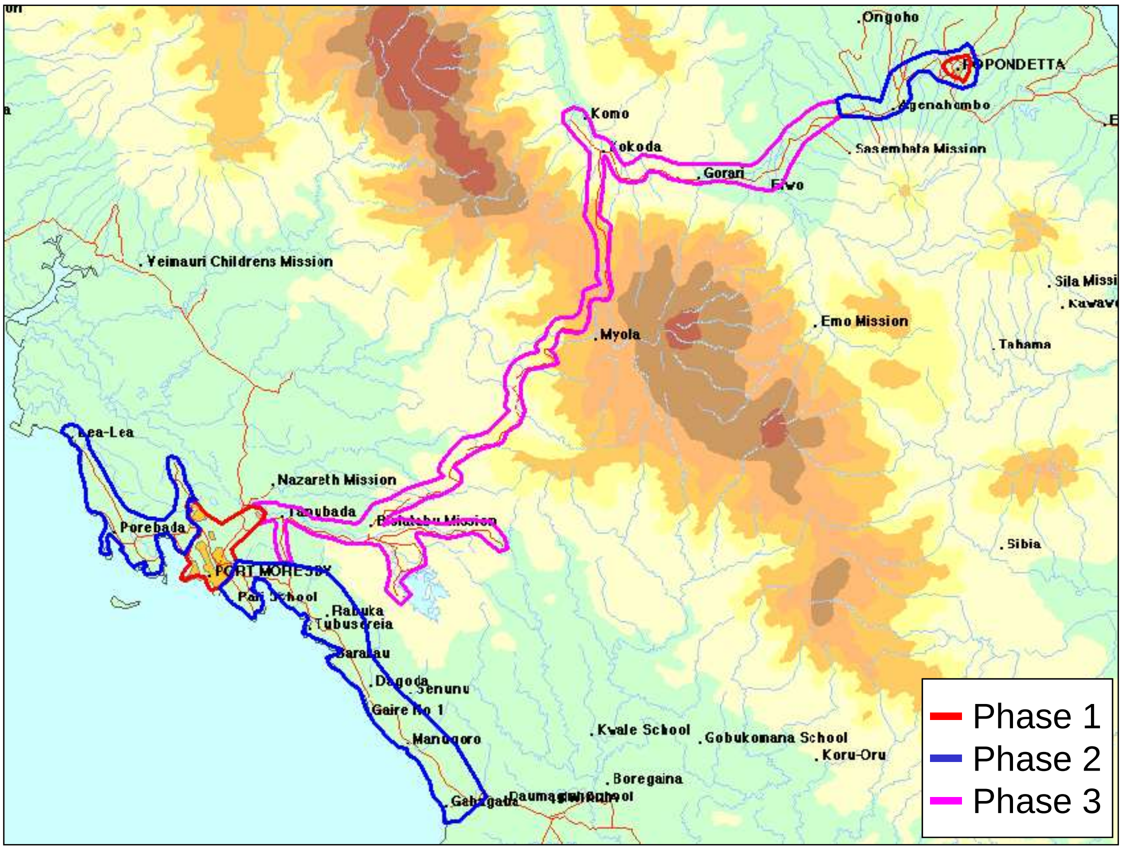


Figure 4: Rollout Phases

<u>Planning Guideline</u>

Scope of the Document

The scope of the following document is to identify and fix all technical and organizational parameters used by the planning staff during the elaboration of the GSM and Fixed network for Papua New Guinea. The document has been worked out on proposals of the experts of LS telcom based on typical system parameters achieved from vendors and ETSI / 3GPP documents.

Network Parameter Settings

Coverage Area

The coverage planning will be divided into three phases:

- During a first phase the capital of Papua New Guinea, the city Port Moresby, and the area of Popondetta in the western part of the country have to be covered. The required coverage in these regions is an acceptable indoor coverage in urban areas of Port Moresby and Popondetta.
- Goal of the second phase is to cover the catchment areas of Port Moresby and Popondetta. Aim is to achieve good outdoor coverage; focus is not set on indoor.
- In the third phase the connection road between Port Moresby and Popondetta will be covered with good outdoor coverage. At the same time indoor coverage in Port Moresby and Popondetta will be optimized if necessary.

The regions are shown in gives the different types of land use for the planning regions.

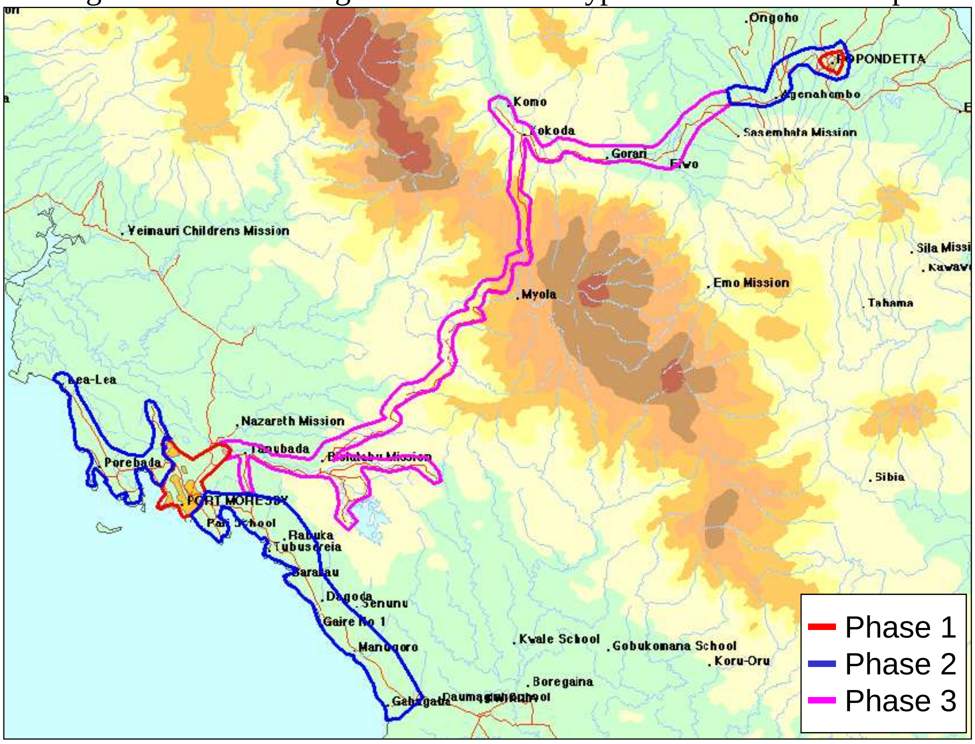


Figure 5: Phases for Rollout

Clutter Classes	Port Moresby	Popondetta	Port Moresby Area	Popondetta Area	Inter-connection

Urban	59.2 %	42.0 %	0.3 %	4.4 %	0.1 %
Open	14.0 %	1.8 %	47.1 %	9.0 %	41.3 %
Forest	26.5 %	56.2 %	47.6 %	86.6 %	53.8 %
Agricultural	0 %	0 %	0 %	0 %	3.0 %
Inland Water	0 %	0 %	0 %	0 %	1.8 %
Sea Water	0.3 %	0 %	5.0 %	0 %	0 %

Table 2: Land use in planning regions

Network design rules

Assumption for the design is that the network will be limited by coverage not by traffic. Therefore, base stations will be placed in a way that coverage will be maximized with a minimum number of base stations. As well as indoor coverage is expected in the urban areas of Port Moresby the focus will not be set on an overall indoor coverage. In urban areas sectorized sites with 65° antennas will be used, rural sites will use 90° antennas as default. Antenna types and antenna configuration may be adjusted to optimize site count and coverage. Coverage along roads should be achieved from sites inside build up areas. Nevertheless, green field sites will be defined where necessary.
For base station antenna height, default heights will be used for the first approach. During coverage optimization phase for selected sites, the height will be minimized in order to save cost for infrastructure.
The sites will be connected by microwave links. If necessary additional sites for fixed network repeaters will be introduced. Sites may be connected in star, loop or daisy chain configuration. No redundancy concept has been foreseen at this stage of the project.
Traffic calculations and frequency planning will not be performed.

System Parameter for Radio Access Network (GSM)

The following paragraphs gives on overview about the used parameter for the GSM radio access network.

Frequency Band

The used frequency band is 870 to 980 MHz, 200 kHz channel spacing and 45 MHz channel separation between uplink and downlink. Field strength calculations will be performed for downlink on a center frequency of 925 MHz.

Parameter for Handsets

Power Classes for Handsets
The calculations will be performed for GSM 900 handsets of class 4 (33 dBm, 0 dBi)

Handset Antenna
An isotropic antenna for the handsets with 0dBi has been assumed.

Receiver Height for Handsets
The used receiver height for handsets is 1.5 m

Parameter for Mobile Stations

Power Classes for Base Stations
The calculations will be performed for GSM900 base stations with an output power of 46 dBm.

Base Station Antennas
Three different types of sector antennas with horizontal beam widths of 33°, 65° and 90° and an omni-directional antenna are available. The technical parameter of the used antennas and the preferred environments are given in Table 2 Mechanical downtilt will be applied where necessary.

Antenna Type	**Antenna Gain**	**Electrical Tilt**	**Preferred usage**
Omni Antenna	11 dBi	0°	Usage in rural areas
Sector Antenna 33°	18 dBi	0°	Usage along roads
Sector Antenna 65°	17 dBi	2°	Usage in urban areas
Sector Antenna 90°	15,5 dBi	0°	Usage in rural areas

Table 3 Base Station Antenna Types

Base Station Antenna Height
An antenna height of 20 meters in sub-urban areas and of 30 meters in rural areas will be used.

Sector Configuration
The following numbers of TRX will be used as default for planning according to the land use of the coverage area of the site.
Sites in Rural area – 1 TRX for each sector
Sites in Urban area – 2 TRX for each sector
No further traffic calculations have been performed at this stage of the project.

Power Budget for planned System
The following tables give the power budget for uplink and downlink. As different TRX configurations are used in rural and urban areas in the latter case additional attenuation for combining has been introduced. The budgets are given for sector antennas with a gain of 17 dBi; as antennas with other gains will not change the balance between up and downlink, separate tables for each antenna type are not necessary. The link budged is done for a coverage probability of 50% without any correction factors for different types of land use. For rural areas, the system is balanced whereas in urban areas the system is limited by the downlink.

Downlink	Rural Areas (1TRX)	Urban Areas (> 1 TRX)	
BTS-Power-Output	46	46	dBm
Losses for Cabling and Combining	4	7	dB
Antenna Gain	17	17	dBi
EIRP	59	56	dBm
Mobile Rx Level	-102	-102	dBm
Antenna Gain	0	0	dBi
Isotropic Received Power at MS (50%)	-102	-102	dBm
Max Path Loss Downlink	161	158	dB

Table 4: Link Budget for Downlink

Uplink	Rural Areas (1TRX)	Urban Areas (> 1 TRX)	
Mobile-Power-Output	33	33	dBm
Antenna Gain	0	0	dBi
EIRP	33	33	dBm
BTS Rx Level	-110	-110	dBm
Antenna Gain	17	17	dBi
Losses for Cabling etc.	3	3	dB
Diversity Gain	4	4	dB
Isotropic Received Power at BS (50 %)	-128	-128	dBm
Max Path Loss Uplink	161	161	dB

Table 5: Power Budget for Uplink

<u>Design field strength values for coverage calculation</u>

The following table gives the design values for the field strength that will be used inside the planning tool to check the coverage. The field strength values are given for different classes of land usage with a coverage probability of 95 percent. For rural areas, the coverage will be optimized for outdoor usage of mobiles. For urban areas acceptable indoor coverage for class 4 mobile phones will be achieved.

	Rural	Incar	Suburban	Urban	
Minimum required Isotropic Power	-95	-85	-83	-75	dBm

Table 6: Signal Strength Design Values

System Parameters for Fixed Network

The paragraph gives an overview over the used parameter for the fixed network planning.

Frequency Band

The following frequency bands will be used according to the length of the link:
Short hops: 23 GHz
Long hops: 8 GHz

Availability

The planned availability for the links will be 99,99 %.

Network Structure

Sites can be connected in star, loop or daisy chain configuration.

Equipment

Technical parameter from Standard Equipment will be used for the planning.

Rain Rate

A rainfall intensity of 110 mm/h and a time percentage of 0.01 will be used.

Radio Network Planning Tools

Used Software

During the Planning work the following software will be used:

Radio Network Planning:	LS telcom xG-Planner	Version 4.6.0
Microwave Link Planning:	LS telcom Multilink	Version 4.3.0

Data sources

The following data sources will be used for additional information

Antenna pattern	Typical GSM-Antennas (Kathrein)
Mapping Data	LS telcom mapping department
Technical Parameter	ETSI / 3GPP and vendor documentation

Used Data Maps

The following mapping data has been available for the case study
Overview Map of Papua New Guinea

Topo Map	200m x 200m per Pixel
Morpho Map	200m x 200m per Pixel

The used rectangular coordinate system is UTM 54 Southern Hemisphere - WGS 84.

Results

In this part of the document, the results of the planning work will be discussed. The results are based on the parameters that have been fixed in the planning guideline.

Results for GSM Network

Results for Phase 1

Site Count and Site Configurations

The found network for the specified area for phase 1 in 12 sites to achieve a good coverage. All sites have three sectors. 29 of these 36 sectors are equipped with antennas with half power beam width of 65° and 17 dBi and the other 7 sectors are equipped with antennas with half power beam width of 90° and 15.5 dBi, see Figure 6.

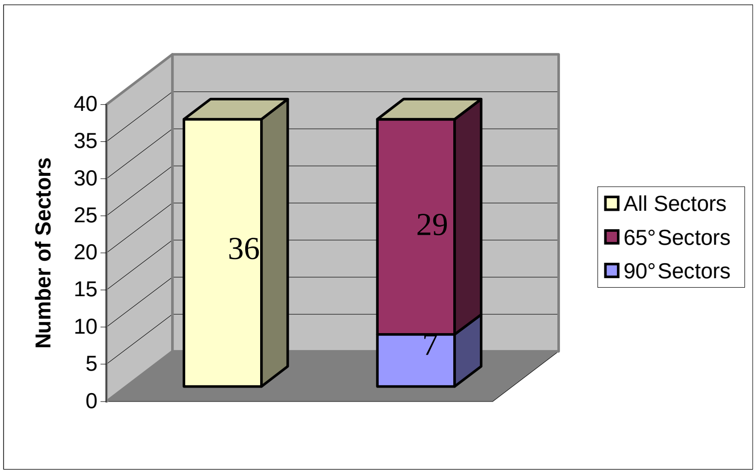


Figure 6: Number of Sectors in Phase 1

Antenna heights of 20m and 30m are used in phase 1 of the network. 24 sectors of the 36 sectors have a height of 20 m and the other sectors have a height of 30 m, see Figure 7.

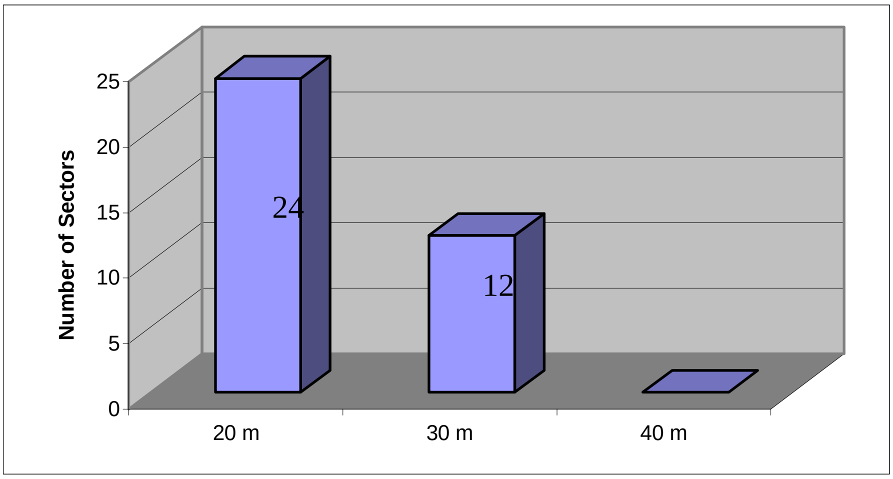


Figure 7: Sector Height Phase 1

Coverage Statistic

Table 7 shows the covered area at the end of phase 1. Values are given in percent of the overall area defined by the polygons for Port Moresby and Popondetta.

Coverage Classes	**Port Moresby**	**Popondetta**	**Port Moresby Area**	**Popondetta Area**	**Inter-connection**
Rural	98.7%	98.5%	40.6%	39.4%	7.9%
Incar	89.2%	90.5%	24.2%	18.5%	4.5%
Suburban Indoor	82.1%	78.8%	19.9%	12.3%	2.9%
Urban Indoor	36.5%	34.8%	8.7%	3.5%	1.4%

Table 7: Covered Area in Percent / Phase 1

<u>Coverage Plot</u>

The coverage at the end of phase 1 in the urban area of Port Moresby is shown in Figure 8. The goal in Phase 1 was to reach acceptable indoor coverage in the urban areas of Port Moresby and Popondetta.

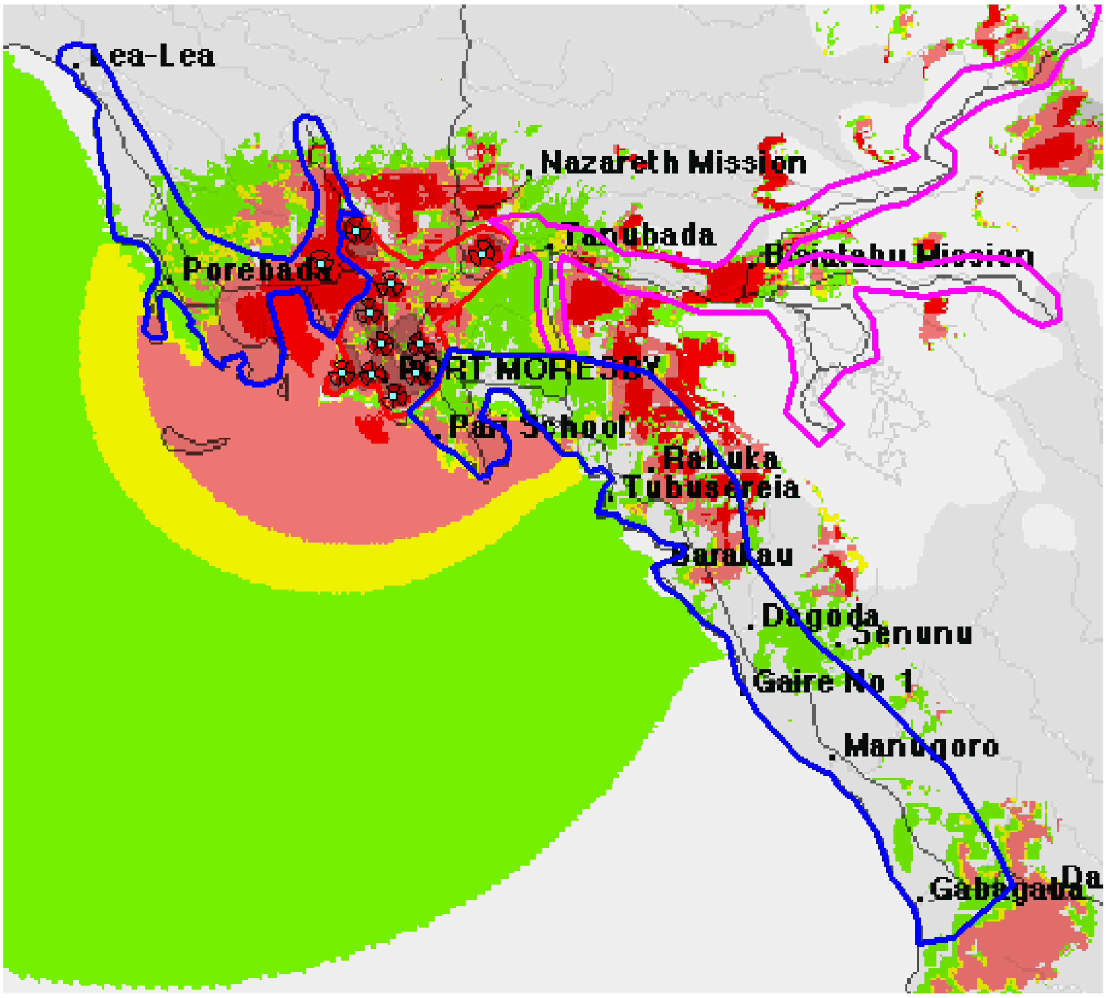


Figure 8: Coverage Phase 1 Port Moresby

The coverage for phase 1 in the urban area of Popondetta is shown in Figure 9.

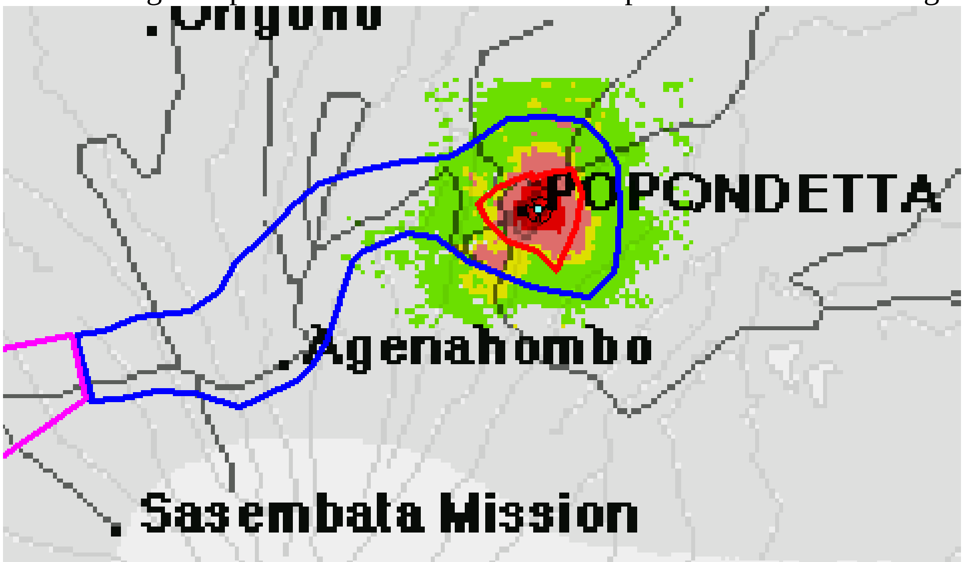


Figure 9: Coverage Phase 1 Popondetta

Site List

Table 8 lists the sites used to achieve coverage in the urban areas of Port Moresby and Popondetta. The used rectangular coordinate system is UTM 54 Southern Hemisphere - WGS 84.

Site Name	X Coordinates	Y Coordinates
Port Moresby 1	55517424.80	8952319.12
Port Moresby 2	55518426.03	8962100.94
Port Moresby 3	55520087.56	8954323.51
Port Moresby 4	55519309.62	8956522.88
Port Moresby 5	55522571.71	8952313.47
Port Moresby 6	55520973.88	8950719.43
Port Moresby 7	55522818.91	8954307.52
Port Moresby 8	55527061.18	8960542.25
Port Moresby 9	55515921.30	8959840.90
Port Moresby 10	55519467.41	8952263.32
Port Moresby 11	55520736.00	8958422.70
Popondetta1	55636991.48	9030657.05

Table 8: Site List of Phase 1

Results for Phase 2

Site Count and Site Configurations

The extension of the network in phase 2 consists of 7 additional sites with 20 sectors. All sectors are using antennas with half power beam width of 90° and gain of 15.5 dBi, see Figure 10.

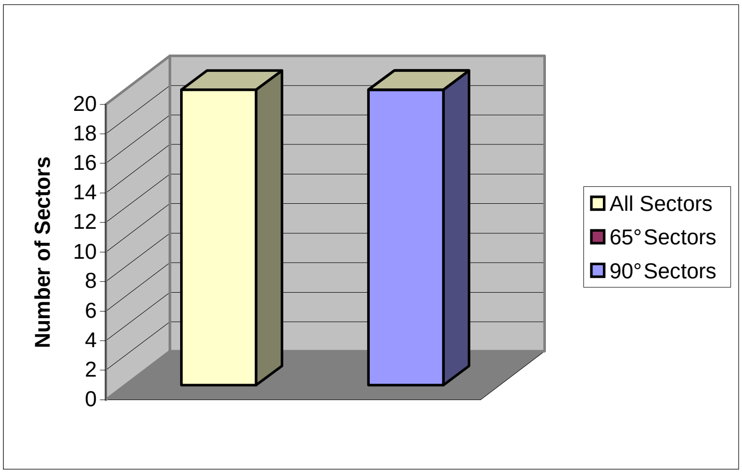


Figure 10: Number of added Sectors in Phase 2

The complete network in Phase 2 includes all in all 19 sites with 56 sectors. The antenna with half power beam width of 90° is used in 27 cases. The other 29 sectors are equipped with a 65° antenna; see Figure 11.

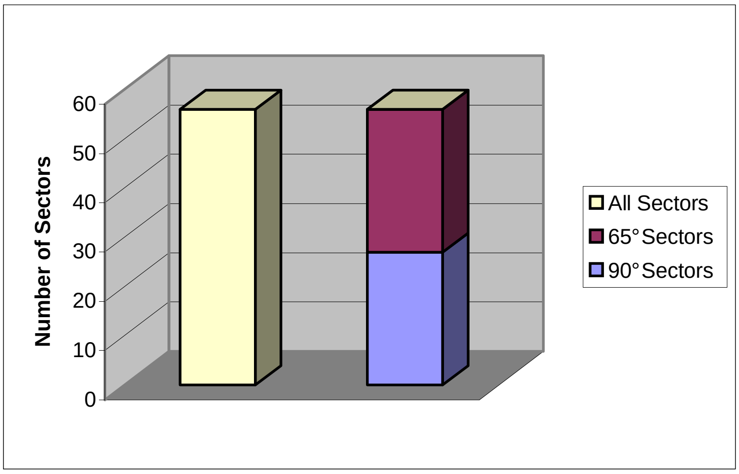


Figure 11: Number of Sectors in complete Phase 2

Mainly antenna heights of 30 m are used. In 6 cases a height of 40 m was selected. The antenna heights of the additional sites in phase 2 are shown in Figure 12.

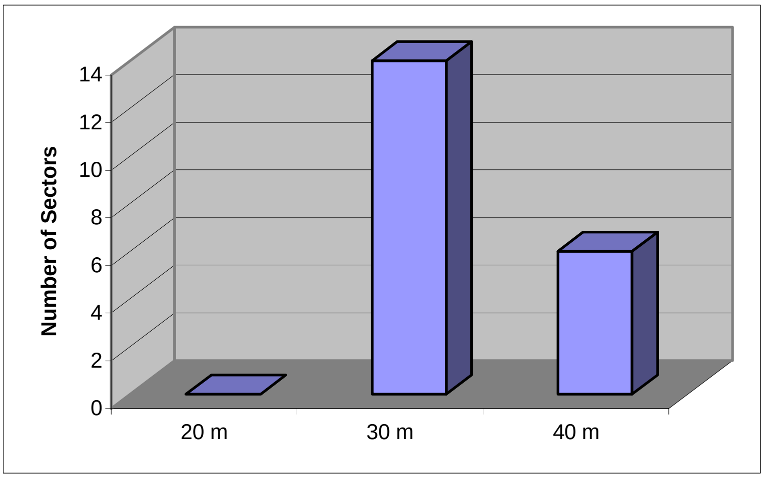


Figure 12: Sector Heights of added Sites / Phase 2

The number of sectors for each antenna height for the complete phase 2 is shown in Figure 13.

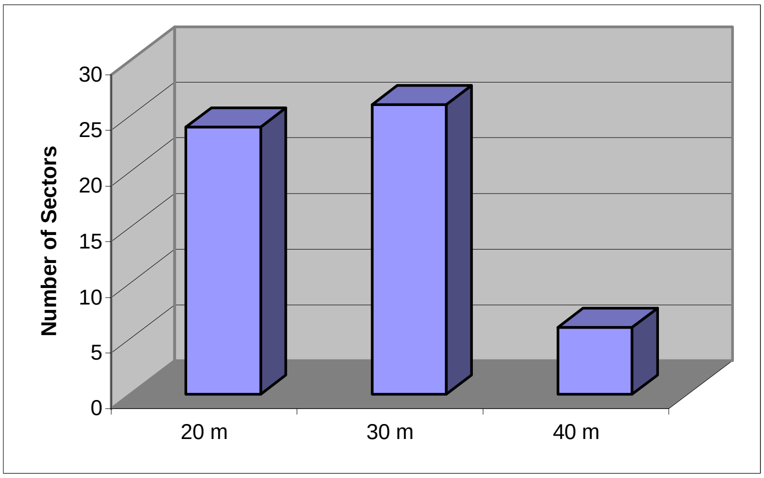


Figure 13: Sector Heights of complete Sites in Phase2

<u>Coverage Statistic</u>

Coverage Classes	**Port Moresby**	**Popondetta**	**Port Moresby Area**	**Popondetta Area**	**Inter-connection**
Rural	98.7%	100%	95.9%	96.6%	9.8%
Incar	89.2%	90.5%	65.7%	48.9%	4.6%
Suburban Indoor	82.1%	78.8%	56.7%	37.1%	3.0%
Urban Indoor	36.5%	34.8%	25.4%	15.3%	1.4%

Table 9: Covered Area in Percent / Phase 2

Coverage Plot

The coverage at the end of phase 2 in the rural area of Port Moresby is shown in Figure 14. The goal in this phase was to achieve good outdoor coverage; focus was not set on indoor.

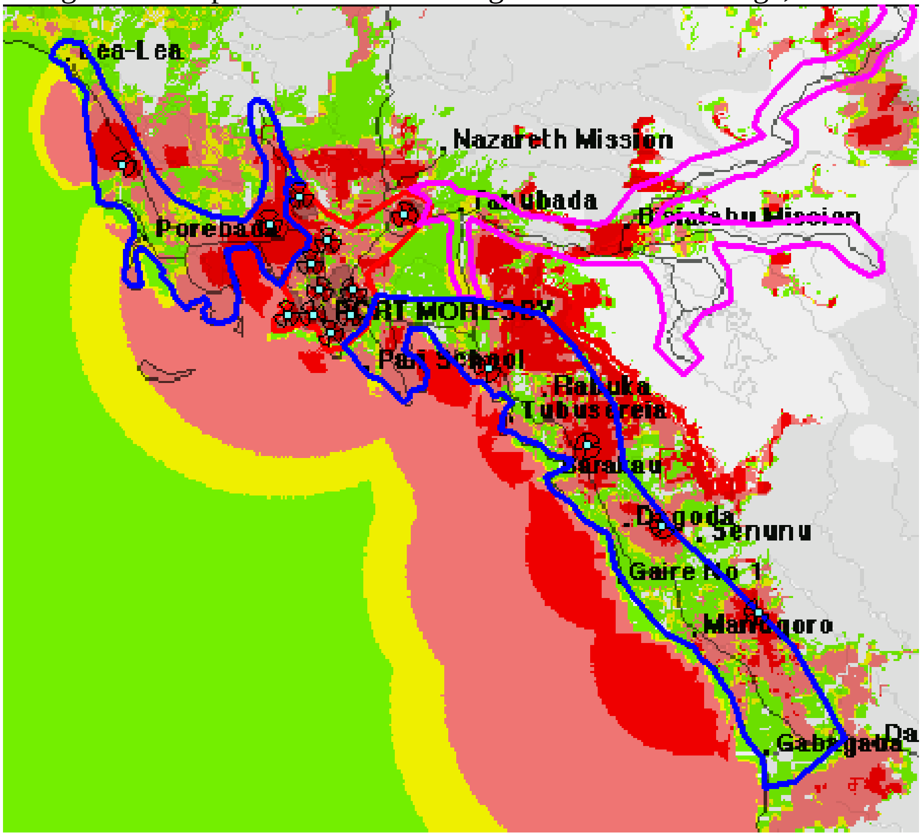


Figure 14: Coverage Phase 2 Port Moresby Area

The coverage of phase 2 in the rural area of Popondetta is shown in Figure 15.

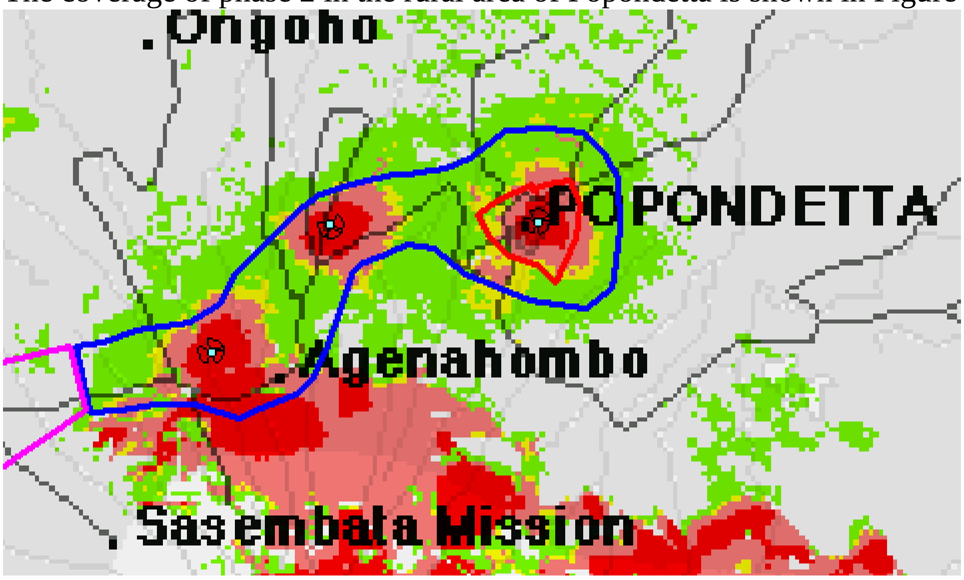


Figure 15: Coverage Phase 2 Popondetta Area

Site List

The following table shows the sites added during phase 2 to cover the rural area of Port Moresby and Popondetta. The used rectangular coordinate system is UTM 54 Southern Hemisphere - WGS 84, see Table 10.

Site Name	X Coordinates	Y Coordinates
Port Moresby Area 1	55503779.78	8964730.75
Port Moresby Area 2	55556409.37	8927602.40
Port Moresby Area 3	55534019.80	8947773.39
Port Moresby Area 4	55542259.83	8941408.50
Port Moresby Area 5	55548507.64	8934821.32
Popondetta Area 1	55628561.14	9030545.94
Popondetta Area 2	55623712.66	9025382.31

Table 10: Site List of Phase 2

Results for Phase 3

Site Count and Site Configurations

In the third and last network phase, 13 sites with 28 sectors were added. In 23 cases, the antenna with a half beam width of 33° was used to cover the interconnection area between the Port Moresby and Popondetta. In 5 cases the sectors were equipped with the 90° antenna, see Figure 16.

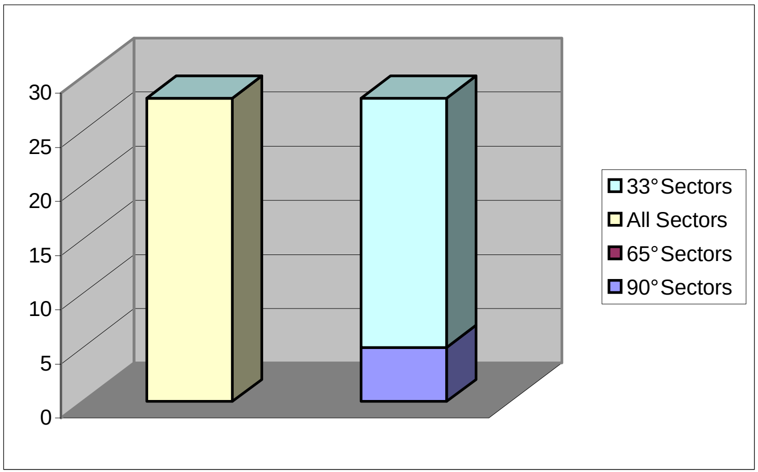


Figure 16: Number of added Sectors in Phase 2

At the end of phase three, the network consists of 32 sites with 84 sectors. Three different antenna types are used. Figure 17 shows the distribution of the antenna types for the sectors. The 33° antenna was used in 23 cases, the 65° antennas in 29 cases and the 90° antenna in 23 cases.

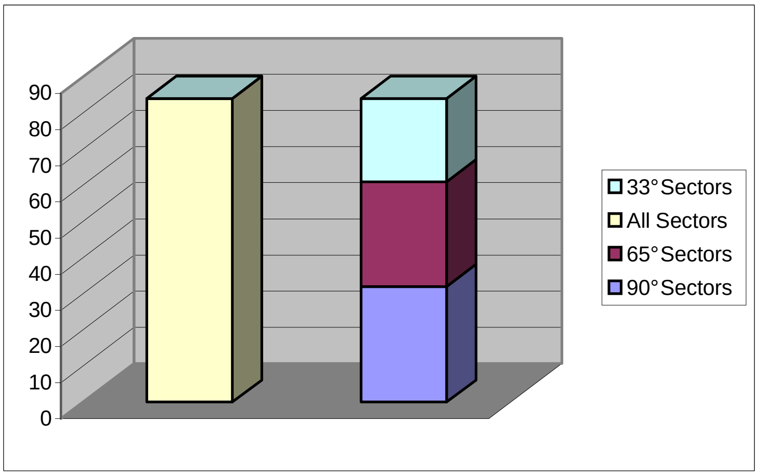


Figure 17: Number of Sectors in complete Phase 3

During the third planning phase only two times an antenna height of 40 m was used. The remaining 26 sectors are having a height of 30 m, see Figure 18. Antenna heights below 30 m has shown to be ineffective in the hilly terrain of the interconnection route between the areas of Port Moresby and Popondetta.

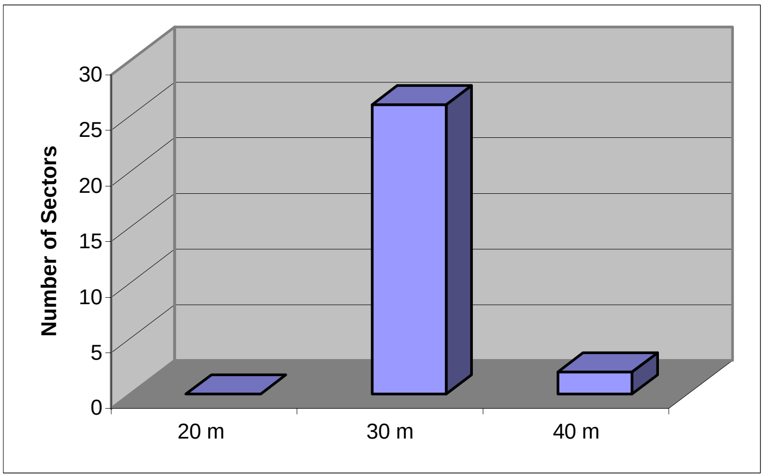


Figure 18: Sector Heights of added Sites / Phase 3

Three different antenna heights were used during the whole planning phases. Figure 19 shows the distribution of the antenna height. Roughly, two thirds of the sectors were using the height of 30 m. 24 sectors are 20 m high and 8 sectors 40 m.

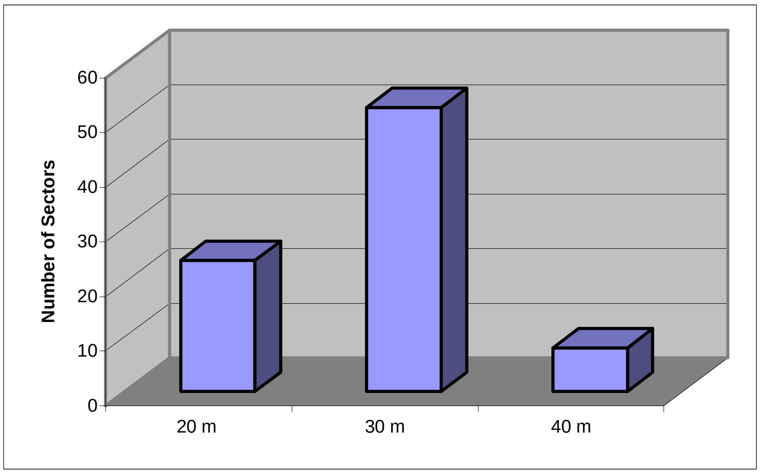


Figure 19: Sector Heights of complete Sites in Phase 3

Coverage Statistic

Table 11 shows the covered area in percent for the result network in phase 3.

Coverage Classes	**Port Moresby**	**Popondetta**	**Port Moresby Area**	**Popondetta Area**	**Inter-connection**
Rural	99.1%	100%	95.9%	97.8%	87.9%
Incar	89.8%	90.5%	65.7%	48.9%	57.8%
Suburban Indoor	82.6%	78.8%	56.7%	37.1%	51.1%
Urban Indoor	36.5%	34.8%	25.0%	15.3%	29.6%

Table 11 Covered Area in Percent / Phase 3

Coverage Plot

The coverage plot shows the achieved coverage for the network after phase 3, see Figure 20:

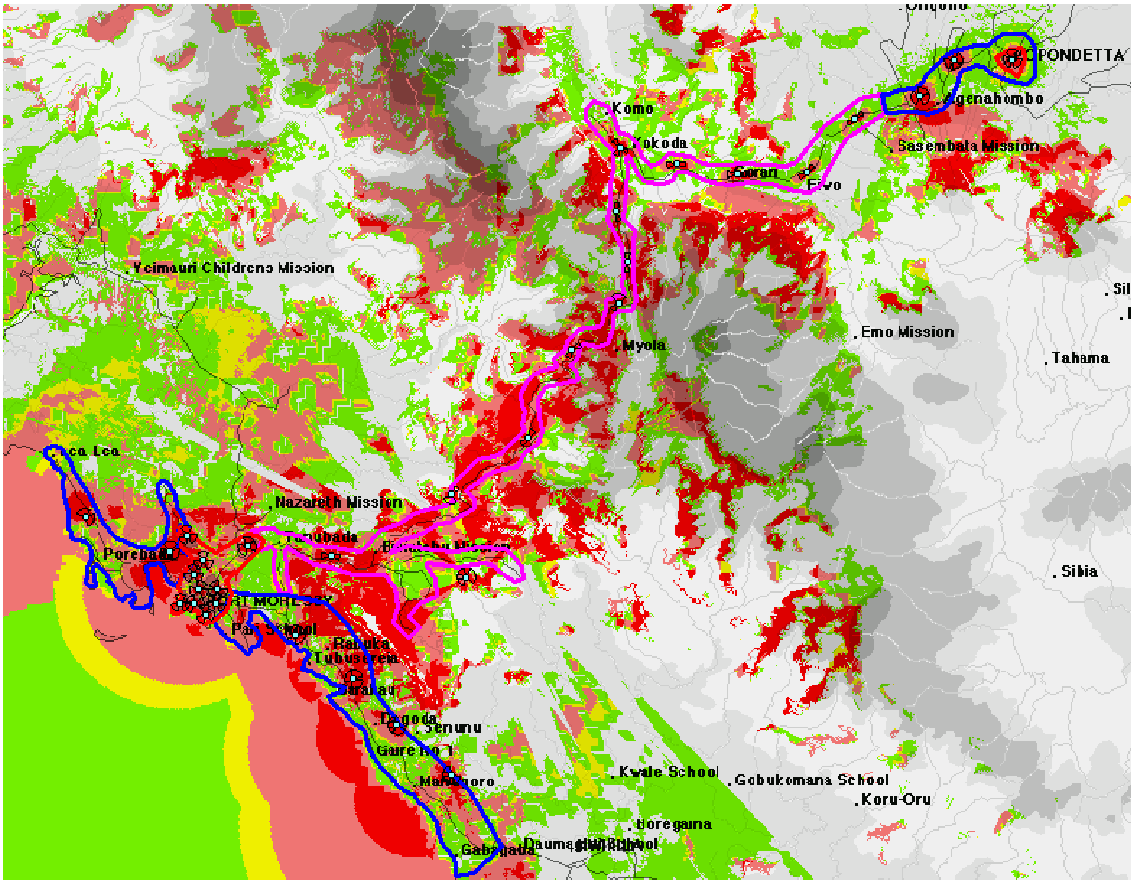


Figure 20: Coverage Phase 3

Site List

The following table shows the used sites needed to achieve the coverage in the interconnection area between Port Moresby and Popondetta. The used rectangular coordinate system is UTM 54 Southern Hemisphere - WGS 84.

Site Name	X Coordinates	Y Coordinates
Interconnection 1	55614358.80	9022109.60
Interconnection 2	55607492.18	9014330.90
Interconnection 3	55597409.24	9014145.95
Interconnection 4	55588740.65	9015527.77
Interconnection 5	55580654.80	9018000.50
Interconnection 6	55580104.05	9008776.11
Interconnection 7	55580392.44	8995565
Interconnection 8	55581685.96	9001492.51
Interconnection 9	55567382.71	8976130.13
Interconnection 10	55558457.45	8956070.41
Interconnection 11	55556385.70	8967977.59
Interconnection 12	55573709.59	8988901.21
Interconnection 13	55539063.00	8959142.60

Table 12: Site List of Phase 3

Results for Fixed Network
- - to be finalized --

Results for Phase 1

Link Count and Link Configuration

Plot with links

Link List

Results for Phase 2

Link Count and Link Configuration

Plot with links

Link List

Results for Phase 3

Link Count and Link Configuration

Plot with links

Link List

Summary

A2.7. Case study from developing country

BROADBAND ACCESS & BACKBONE SOLUTIONS.

CASE STUDY FROM PAPUA NEW GUINEA (PNG)

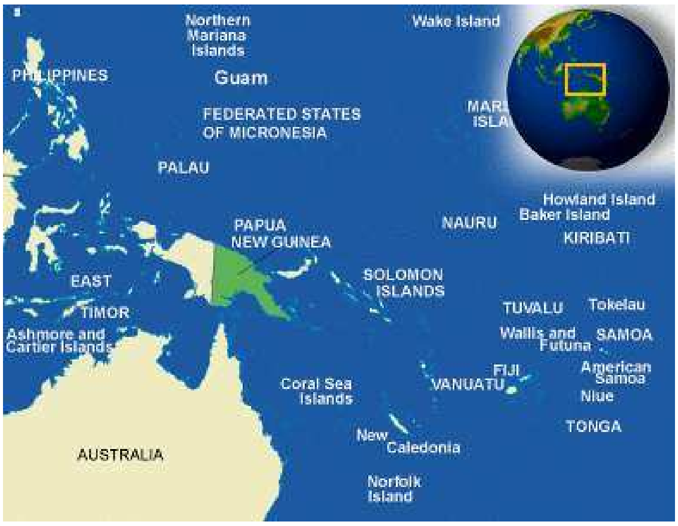
TELECOMMUNICATIONS INDUSTRY IN PNG

The Telecommunication Industry in Papua New Guinea is a regulated industry. The sole provider of telecommunication services is Telikom PNG Limited (TPNG) that is a fully state owned Company. Plans are currently underway to partially privatize it. Independent Consumer and Competition Commission (ICCC) is the super regulator (economic and social regulator) while PANGTEL is the sole Technical Regulator.

The telecommunications industry in PNG has a potential for 1.038 million additional subscribers for basic telecommunications services, current connection capacity of 96,000 lines and 150,000 mobile cellular (GSM) capacities with 67,000 telephone subscriber lines connected and 54,000 mobile subscribers connected. A teledensity of 13 lines per 1000 people and mobiledensity of 11 mobiles per 1000 people.

In the past two years, the Telecommunication industry has been under going a number of new reforms. Some of these reforms are to regulate the industry by providing a legislative framework where TELIKOM PNG LTD will be a regulated entity by way of a regulatory contract. PANGTEL will retain responsibility for the technical aspects of the telecommunications industry.

SOCIO-ECONOMIC ASPECTS



Papua New Guinea is an Island country comparable to but slightly larger than the city of California, USA and borders the main island of New Guinea with Indonesia on the west. It has a land area of 452,860 sq.km of 0.084 sq.km. per person. It has a population of 5.4 million people. The country's population density is 11.92 persons per sq.km. Its terrain is mostly mountains with coastal lowlands and rolling foothills.

The language is Melanesian pidgin as lingua franca, English spoken by 1-2% and is the language of commerce and business with motu in the coastal region. Life expectancy is 64.56 years. Port Moresby is the largest and the capital city and has the largest population of 325,000 people.

Its natural resources are gold, copper, silver, natural gas, timber, oil and fisheries. PNG is richly endowed with natural resources but exploitation has been hampered by rugged terrain and high cost of developing infrastructure. Agriculture provides a subsistence livelihood for 85% of the population. Minerals deposits including oil, copper, gold account for 72% of export earnings.

PNG receives 7.2% of GDP as aid, has budget expenditures (Total Expenditure US$1.35 billion) of US$996.8 million including capital expenditures of US$344 million and total revenue of US$1.33 billion. The economy has improved over the past two years after a prolonged period of instability. Australia annually supplies US$240 million in aid which accounts for 20% of the national budget. Challenges facing the country include further gaining investor confidence, privatizing government assets and balancing relations with Australia the former colony. The current exchange rate is one Kina = 3.2225 USA dollar. PNG income category is low income. The population level below poverty line is 37%.

TELECOMMUNICATIONS SERVICES & COMMUNICATIONS

Fixed services

TPNG's Public Switched Telecommunications Network services include:

- Switched telephone international services
- Switched telephone domestic (local and national distance) services
- Public pay phones
- Leased circuit services
- Packet switched data
- Outstation HF radio services
- Public radio-communications services (coastal radio services/maritime and aeronautical services).
- Other services provided include mobile satellite services and paging.

Mobile services

Mobile network and services are provided by Pacific Mobile Communications Ltd (PMC), a wholly owned subsidiary of TPNG. Like the fixed services, mobile services are currently provided on an exclusive basis.

PMC has replaced the old AMPS analogue system with the "second generation" GSM technology. The GSM system now services 6 major urban towns and is expanding due to its preference over fixed phones by customers. The service is provided with pre-paid option.

The demand and preference for mobile phones over fixed phones saw the customer base (150,000) exceed the fixed phone customer base (96,000) recently and is growing. With more and more demand for the service, the network is currently experiencing significant congestion. The congestion was caused by a number of factors including:

- Insufficient call handling capacity of the mobile exchange;
- Insufficient capacity of the prepaid billing platform;
- Insufficient transmission capacity between the mobile exchange and TPNG prepaid billing platform;
- Insufficient transmission between the mobile exchange and TPNG fix line exchange;
- Insufficient radio capacity on the mobile base station network; and
- Network not being optimised.

System optimisation to be completed by end of August 2005 with the mobile network upgrade.

Internet gateway service

TPNG also provides Internet Gateway Service to five (5) Internet Service Providers (ISP's). ISP's are not licensed as Value Added Service providers however they operate under an agreed arrangement with TPNG. One ISP is currently providing email service, which makes use of HF radio, particularly in the rural areas. The satellite link provides an aggregate bandwidth of 8 Mbps while the cable network has a capacity of 2.0 Mbps.

Private Communication Networks

The demand for ICT and broadband applications from various sectors in PNG can not be emphasised enough. So much so that certain entities, have installed and are operating their own Very Small Terminal Aperture (VSAT) based communication systems..

Radio broadcasting services

AM and SW Broadcasting services are provided by the government operator, National Broadcasting Corporation (NBC). A non-commercial community SW radio service commenced operation in June2005.
There are also two nationwide commercial FM radio networks while in the main city and towns. Four FM radio stations (commercial and non-commercial) also operate locally as well, two of which broadcast in the Pidgin and Motu languages. A number of non-commercial community FM radio stations operate in remote rural locations, which use the local language of the serviced area to broadcast its programs.

Television and Multi-channel Multi-Distribution System (MMDS)

There is only one nationwide TV broadcaster and a 27-channel Pay TV service, which operates only in Port Moresby, the national capital, using Microwave Multipoint Distribution System. At least two major Cable TV networks serve Port Moresby and other major towns.

DEVELOPMENT OF TELECOMMUNICATIONS SERVICES

Government Policy & Privatisation

The government has also commenced privatization of state-owned entities including Telikom PNG. The objectives of this exercise include;

- Improving the provision of essential services in accordance with priority policy objectives
- Promoting economic efficiency and encouraging competition, and
- Improving the performance of state-owned enterprises.
- Future Plans for Information Communications

With the privatization of Telikom PNG, the government is including the Community Service Obligation (CSO) in the sale process as a major condition. This would constitute a deed of agreement between the State and the privatized Telikom PNG binding it to provide accessibility of basic telecommunication services to the rural sector in which 80% of the total population are based. The main objective is raising the current access level to the ITU USO acceptable levels as much as possible.

Major Set-Backs and Hurdles to Development of Telecommunications

The diversity of culture/languages, scattered villages, the rural based population and the rugged terrain of the country is a major hurdle to the development of the Telecommunications in PNG. Natural disasters and theft also have a part to play in the already very difficult area.

Rural Based Population

One major barrier in the development of the industry is that the vast majority (80%) of the total population of PNG is rural based and depend heavily on subsistence farming. The rural populations do not really see the need to own a phone and even if they are in the vicinity of an access point will not afford a phone. Due to this, Telikom PNG does not see such an exercise as profitable.

Scattered Villages

In PNG, the definition of a village would differ from the village in other countries. The villages in PNG are small and are scattered making it very difficult to extend the current PSTN network out to the rural areas. Even if an access point were located in one of the PNG villages, the inhabitants of the other villagers would have to travel kilometers just to have access to a telephone. The targets as set by the USO are very difficult to meet in such a scenario.

Rugged Terrain

The natural terrain of PNG makes the provision of service difficult and the extensive use of radio a necessity with radio repeaters often being sited on inaccessible mountain peaks serviceable only by helicopters. High compensation demands by landowners for land on which repeaters are located has forced Telikom to introduce satellite system to provide a more efficient and reliable network that is very costly.

PNG NETWORK PLANNING CASE STUDIES

The organisation of the PNG case study was in two parts:

1. Broadband solution for the capital city of Port Moresby,
2. Backbone solution for the whole country.

The case studies were performed and run with highly professional Network Planning Tools provided by VPISystems. Other companies' tools partners of the ITU in NP programs in NetWORKS, LSTelcom were supposed to be used but unfortunately were not run or study performed using them.

1 Broadband Solution Study of Boroko suburban area in Port Moresby.

Figure 1: Boroko commercial area. The Study Area is on the top right hand of the picture.

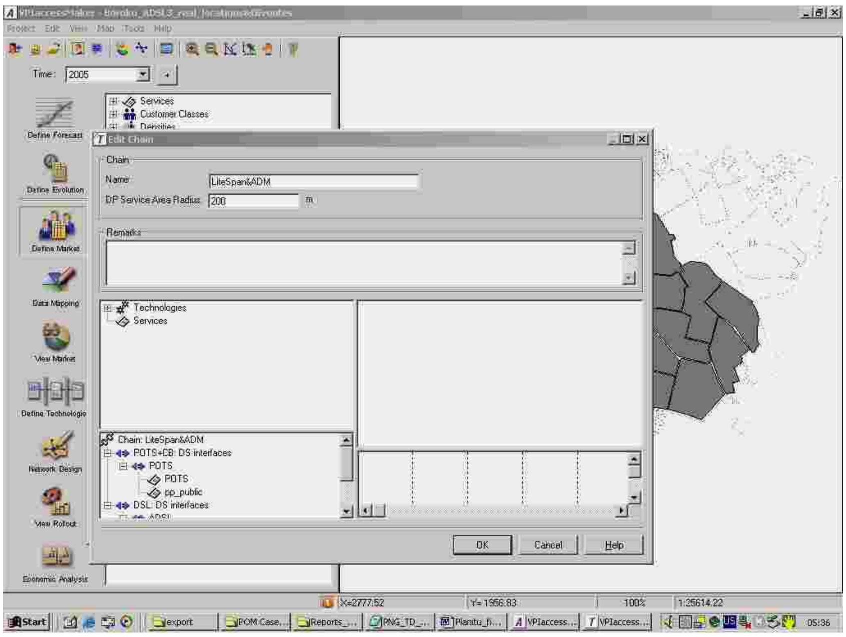
Figure 2: Service areas of Boroko, Business, Residential and Government according to Distribution Areas (DA).

The customer classes were segmented into the following:

- Business customers_ Bus
- Residential POTS_ Res_POTS
- Residential Low_Res_Low
- Coin Boxes_CB

The services respectively were:

- ADSL Business/ 512Kbps
- POTS/64kbps
- ADSL Residential / 128kbps
- Payphones / 64kbps



Attention: This is not an ITU publication made available to the public, but **an internal ITU Document** intended only for use by the Member States of the ITU and by its Sector Members and their respective staff and collaborators in their ITU related work. It shall not be made available to, and used by, any other persons or entities without the prior written consent of the ITU.

Figure 3: The VPIAccessMaker™ Tools is used to study the potential and distribution of the new services.

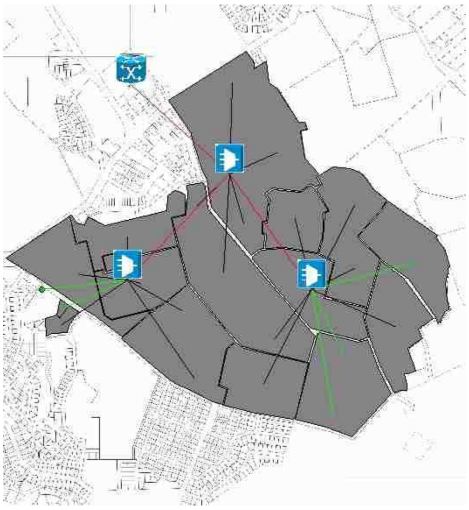


Figure 4: The Optimised DSLAM & Router locations in Boroko. Only three DSLAM with one Router at the exchange is needed.

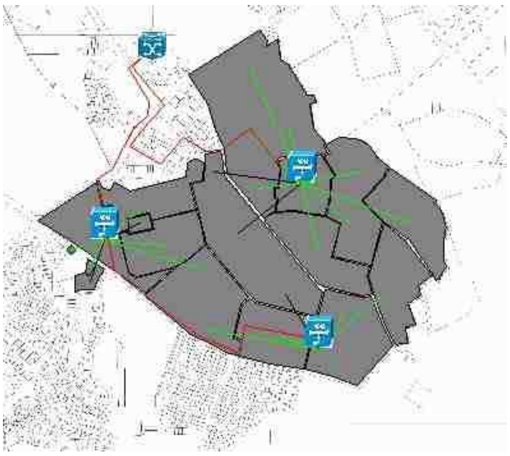


Figure 5: The Optimised Location within existing buildings with optical fibre links.

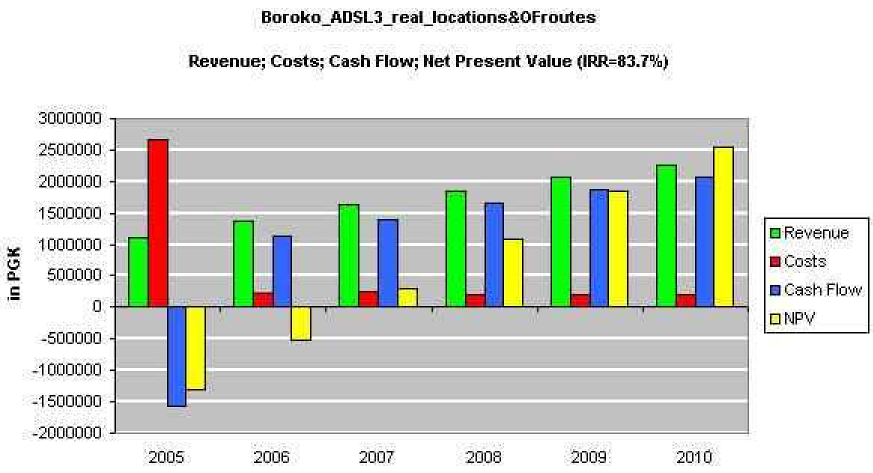


Figure 6: Economic Analysis of ADSL in the access network of Boroko

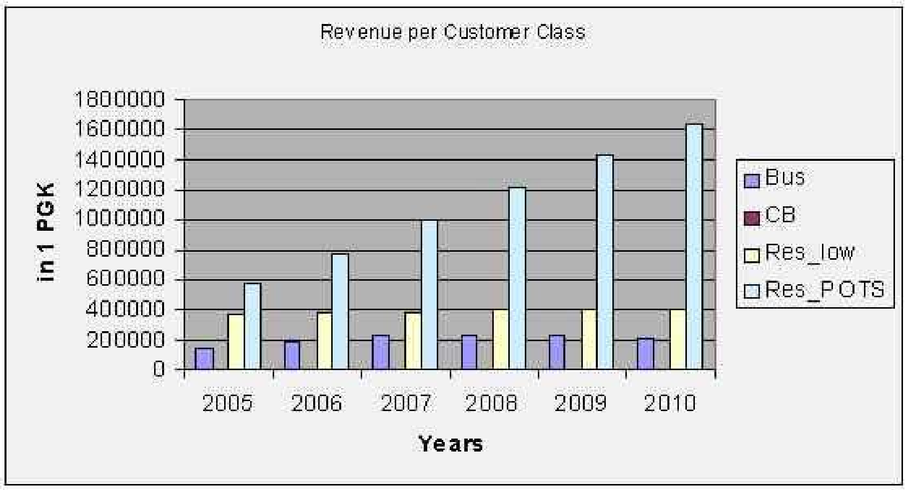


Figure 7: The forecasted revenue of the services. The POTS segment of the services will provide most of the revenue

2 Backbone Solution for the Whole Country.

The backbone solution study was to optimize the routing and dimensioning of the network. The numbers of exchanges of 27 were used in the study.

The transmission medium used in the networks are Radio, Satellite, Fibre optic and cable. About 90% of the links are radio and satellite transmission systems covering the whole country. Apart from diversity on satellite it also provide means to counter demands for compensation using vandalism and forceful disruptions to service by the local land owners on whom the repeaters are installed . The Data network is over lay on the PSTN. The satellite links are for diversity and for oil, copper, gold mines were radio links are not covered or inaccessible due to rugged terrain.

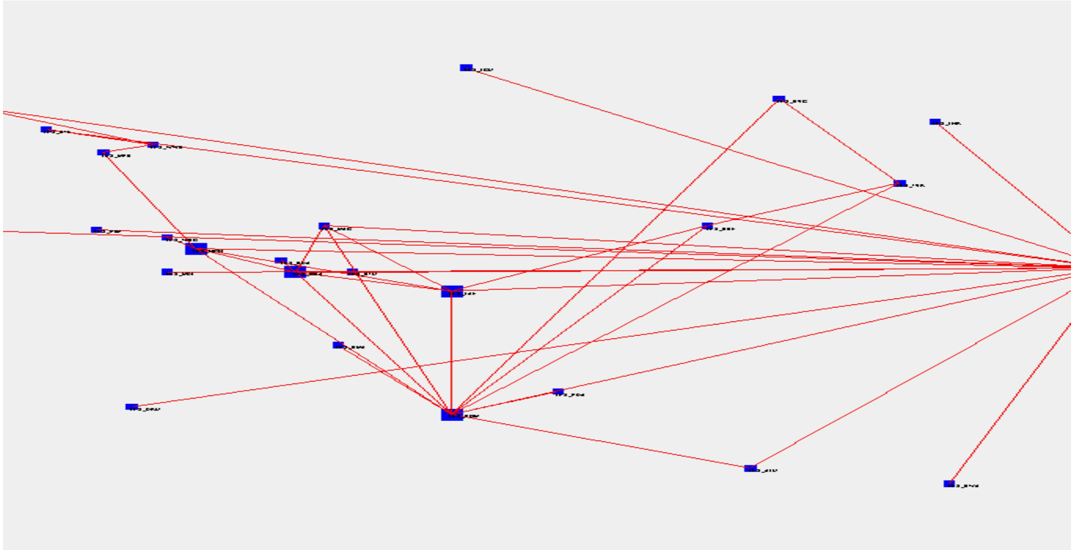


Figure 8: Physical Links of the Papua New Guinea Telecommunications Networks showing Satellite and Radio links.

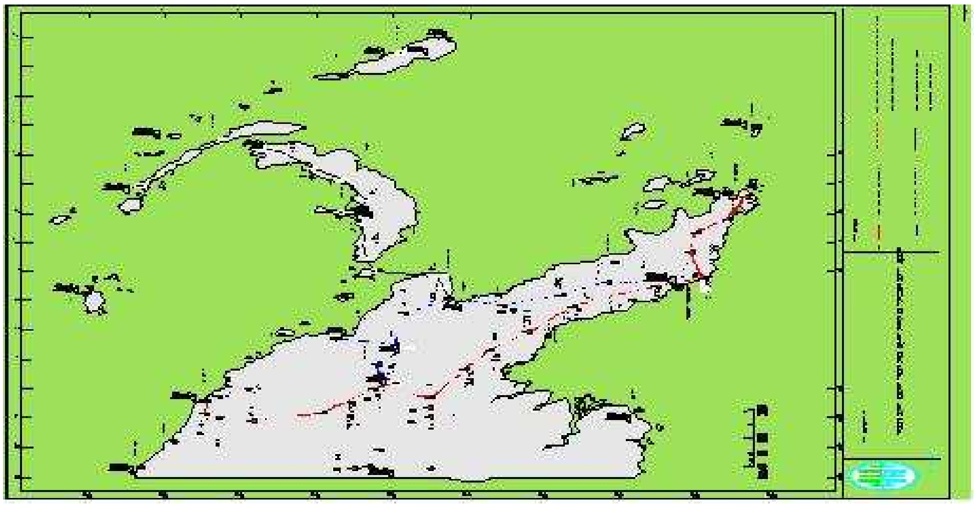


Figure 9: PNG National Microwave Trunk Network.

The complete digitization of the transmission network is undergoing completion by 2005/2006 planning period.

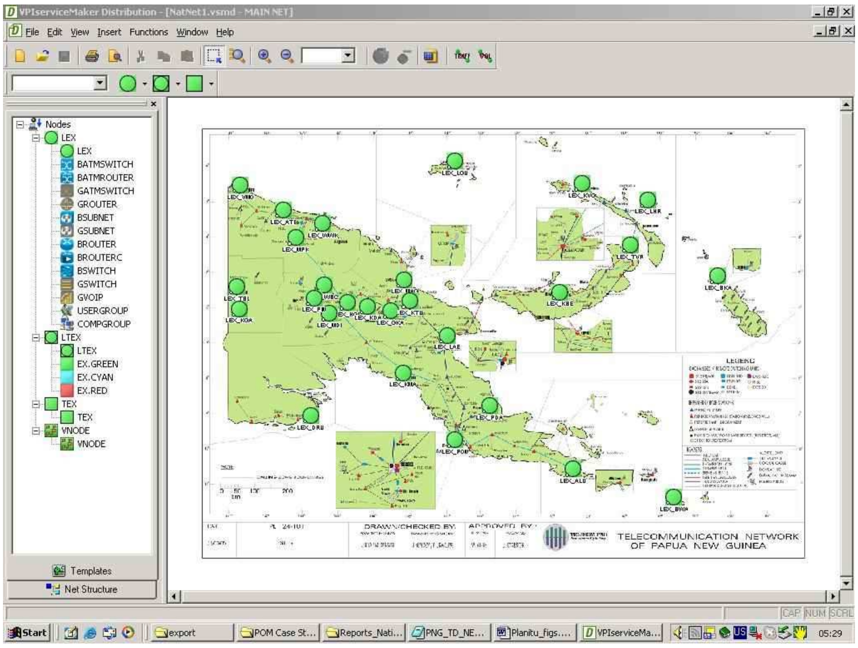


Figure 10: The Switching Network of Papua New Guinea showing existing exchange nodes locations.

The Telikom Switching Network is presently comprised of a five level hierarchical structure as follows;

- International Switching Centers – Port Moresby & Lae
- Sector Primary Exchanges
- District Local/Transit Exchanges
- Remote Switching/Concentrator Units

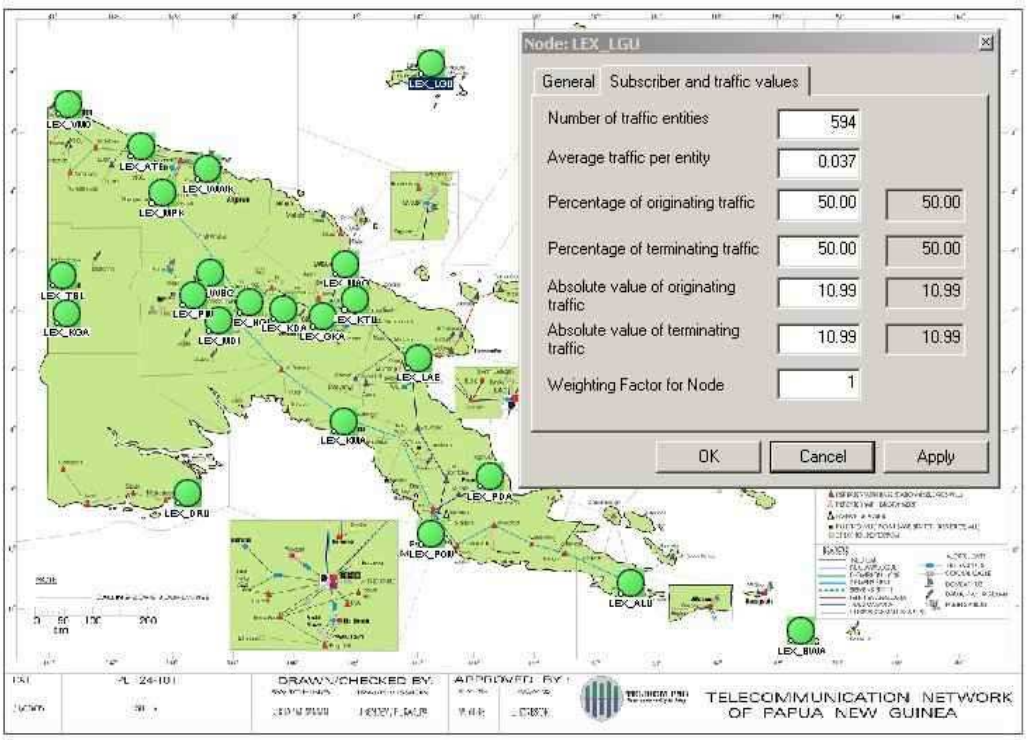


Figure 11: The specialised Network Planning tool of VPIserviceMaker™ Distribution was used to perform traffic forecast and distribution. The generated traffic distribution matrix for the network is below. The parameters were required:

- Subscriber numbers for each of the 27 nodes,
- Average traffic outgoing for national network from the 27 nodes
- Proportion of traffic terminating and originating from the 27 nodes

MM	LEX_ALU	LEX_ATE	LEX_BKA	LEX_BWA	LEX_DRU	LEX_GKA	LEX_HGN	LEX_KBE	LEX_KDA	LEX_KGA	LEX_KMA	LEX_KTU	LEX_KVG	LEX_LAE	LEX_I
LEX_ALU	0.00	0.10	0.19	0.07	0.19	0.93	1.36	0.67	0.13	0.14	0.11	0.12	0.33	3.77	0
LEX_ATE	0.10	0.00	0.03	0.01	0.03	0.17	0.25	0.12	0.02	0.03	0.02	0.02	0.06	0.69	0
LEX_BKA	0.19	0.03	0.00	0.03	0.06	0.32	0.46	0.23	0.04	0.05	0.04	0.04	0.11	1.29	0
LEX_BWA	0.07	0.01	0.03	0.00	0.03	0.13	0.18	0.09	0.02	0.02	0.02	0.02	0.05	0.51	0
LEX_DRU	0.19	0.03	0.06	0.03	0.00	0.32	0.47	0.23	0.05	0.05	0.04	0.04	0.12	1.31	0
LEX_GKA	0.93	0.17	0.32	0.13	0.32	0.00	2.33	1.14	0.22	0.24	0.19	0.21	0.57	6.46	0
LEX_HGN	1.36	0.25	0.46	0.18	0.47	2.33	0.00	1.67	0.33	0.35	0.28	0.31	0.84	9.43	0
LEX_KBE	0.67	0.12	0.23	0.09	0.23	1.14	1.67	0.00	0.16	0.17	0.14	0.15	0.41	4.62	0
LEX_KDA	0.13	0.02	0.04	0.02	0.05	0.22	0.33	0.16	0.00	0.03	0.03	0.03	0.08	0.90	0
LEX_KGA	0.14	0.03	0.05	0.02	0.05	0.24	0.35	0.17	0.03	0.00	0.03	0.03	0.09	0.97	0
LEX_KMA	0.11	0.02	0.04	0.02	0.04	0.19	0.28	0.14	0.03	0.03	0.00	0.03	0.07	0.78	0
LEX_KTU	0.12	0.02	0.04	0.02	0.04	0.21	0.31	0.15	0.03	0.03	0.03	0.00	0.08	0.85	0
LEX_KVG	0.33	0.06	0.11	0.05	0.12	0.57	0.84	0.41	0.08	0.09	0.07	0.08	0.00	2.32	0
LEX_LAE	3.77	0.69	1.29	0.51	1.31	6.46	9.43	4.62	0.90	0.97	0.78	0.85	2.32	0.00	1
LEX_LGU	0.23	0.04	0.08	0.03	0.08	0.40	0.58	0.28	0.06	0.06	0.05	0.05	0.14	1.60	0
LEX_LHR	0.24	0.04	0.08	0.03	0.08	0.40	0.59	0.29	0.06	0.06	0.05	0.05	0.14	1.63	0
LEX_MAG	1.13	0.21	0.38	0.15	0.39	1.93	2.82	1.38	0.27	0.29	0.23	0.25	0.69	7.81	0
LEX_MDI	0.16	0.03	0.06	0.02	0.06	0.28	0.40	0.20	0.04	0.04	0.03	0.04	0.10	1.12	0
LEX_MPK	0.07	0.01	0.02	0.01	0.02	0.12	0.17	0.08	0.02	0.02	0.01	0.02	0.04	0.48	0
LEX_PDA	0.31	0.06	0.11	0.04	0.11	0.54	0.79	0.38	0.08	0.08	0.07	0.07	0.19	2.18	0
LEX_PIM	0.10	0.02	0.03	0.01	0.03	0.17	0.25	0.12	0.02	0.03	0.02	0.02	0.06	0.68	0
LEX_POM	12.21	2.22	4.17	1.65	4.25	20.93	30.53	14.97	2.92	3.14	2.53	2.75	7.50	84.61	5
LEX_TBL	0.54	0.10	0.19	0.07	0.19	0.93	1.36	0.67	0.13	0.14	0.11	0.12	0.33	3.77	0
LEX_TVR	1.24	0.23	0.42	0.17	0.43	2.13	3.11	1.52	0.30	0.32	0.26	0.28	0.76	8.62	0
LEX_VMO	0.22	0.04	0.07	0.03	0.08	0.37	0.54	0.26	0.05	0.06	0.04	0.05	0.13	1.50	0
LEX_WBG	0.12	0.02	0.04	0.02	0.04	0.21	0.31	0.15	0.03	0.03	0.03	0.03	0.08	0.86	0
LEX_WWK	0.66	0.12	0.23	0.09	0.23	1.13	1.65	0.81	0.16	0.17	0.14	0.15	0.41	4.58	0
Debit	25.90	4.72	8.85	3.49	9.01	44.40	64.75	31.74	6.20	6.66	5.37	5.84	15.91	179.45	10
sum TERM	25.3562	4.6970	8.7827	3.4831	8.9433	42.8014	61.3499	30.9249	6.1639	6.6241	5.3467	5.8144	15.7049	153.3302	10.8
Corr	0.9790	0.9962	0.9928	0.9972	0.9927	0.9640	0.9475	0.9742	0.9950	0.9946	0.9957	0.9953	0.9871	0.8544	0.9
Delta	0.0210	0.0038	0.0072	0.0028	0.0073	0.0360	0.0525	0.0258	0.0050	0.0054	0.0043	0.0047	0.0129	0.1456	0.0

OK
Cancel
Update Debit
Calc Single
IT-Factor
Target Factor
Fix matrix
gen TAM
Int Matrix
Undo Int Matrix
Kruithof
MinQS

Figure 12: The Network Traffic Matrix generated by the VPIserviceMaker™ Distribution Planning Tool for originating, terminating and total traffic in the Network.

The Dimensioned & Routing of the Network

The Dimensioning of traffic routes and routing for the whole network is performed by the PLANITU Planning Tool..

PNG National Telecoms Network 2005 Tr&D new 2005- 7-22 21

Statistics for Circuits & Traffics

Exchange		Circuits		Peak Traffics		
#	Name	Inc	Outg	Inc	Outg	Transit
1	LEX_ALU	60	60	38.	38.	0.
2	LEX_ATE	30	30	7.	7.	0.
3	LEX_BKA	30	30	13.	13.	0.
4	LEX_BWA	30	30	5.	5.	0.
5	LEX_DRU	30	30	13.	13.	0.
6	LEX_GKA	90	90	65.	65.	0.
7	LEX_HGN	600	600	95.	95.	291.
8	LEX_KBE	60	60	46.	46.	0.
9	LEX_KDA	30	30	9.	9.	0.
10	LEX_KGA	30	30	10.	10.	0.
11	LEX_KMA	30	30	8.	8.	0.
12	LEX_KTU	30	30	8.	8.	0.
13	LEX_KVG	60	60	23.	23.	0.
14	LEX_LAE	540	540	258.	258.	139.
15	LEX_LGU	30	30	16.	16.	0.
16	LEX_LHR	30	30	16.	16.	0.
17	LEX_MAG	120	120	79.	79.	0.
18	LEX_MDI	30	30	11.	11.	0.
19	LEX_MPK	30	30	5.	5.	0.
20	LEX_PDA	60	60	22.	22.	0.
21	LEX_PIM	30	30	7.	7.	0.
22	LEX_POM	1140	1140	613.	613.	140.
23	LEX_TBL	60	60	38.	38.	0.
24	LEX_TVR	360	360	87.	87.	133.
25	LEX_VMO	30	30	15.	15.	0.
26	LEX_WBG	30	30	9.	9.	0.
27	LEX_WWK	60	60	46.	46.	0.

Total traffic offered : 1562.

Total traffic lost : 3.

Table 1: Optimised Dimensioning results from PLANITU Planning Tool. The results show that the Total Network Traffic generated is 1562 erlang, and 3 Erlang of traffic lost implying no congestion in the Network.

27 exchanges

LEX_ALU	1	0.0	0.0	1	2	2	22	0.	0.
LEX_ATE	2	0.0	0.0	1	2	2	7	0.	0.
LEX_BKA	3	0.0	0.0	1	2	2	22	0.	0.
LEX_BWA	4	0.0	0.0	1	2	2	22	0.	0.
LEX_DRU	5	0.0	0.0	1	2	2	22	0.	0.
LEX_GKA	6	0.0	0.0	2	2	2	0	0.	0.
LEX_HGN	7	0.0	0.0	2	2	2	0	0.	0.
LEX_KBE	8	0.0	0.0	1	2	2	22	0.	0.
LEX_KDA	9	0.0	0.0	1	2	2	6	0.	0.
LEX_KGA	10	0.0	0.0	1	2	2	22	0.	0.
LEX_KMA	11	0.0	0.0	1	2	2	22	0.	0.
LEX_KTU	12	0.0	0.0	1	2	2	22	0.	0.
LEX_KVG	13	0.0	0.0	1	2	2	22	0.	0.

LEX_LAE	14	0.0	0.0	2	2	2	0	0.	0.
LEX_LGU	15	0.0	0.0	1	2	2	22	0.	0.
LEX_LHR	16	0.0	0.0	1	2	2	22	0.	0.
LEX_MAG	17	0.0	0.0	1	2	2	14	0.	0.
LEX_MDI	18	0.0	0.0	1	2	2	22	0.	0.
LEX_MPK	19	0.0	0.0	1	2	2	7	0.	0.
LEX_PDA	20	0.0	0.0	1	2	2	22	0.	0.
LEX_PIM	21	0.0	0.0	1	2	2	22	0.	0.
LEX_POM	22	0.0	0.0	2	2	2	0	0.	0.
LEX_TBL	23	0.0	0.0	1	2	2	22	0.	0.
LEX_TVR	24	0.0	0.0	1	2	2	22	0.	0.
LEX_VMO	25	0.0	0.0	1	2	2	22	0.	0.
LEX_WBG	26	0.0	0.0	1	2	2	7	0.	0.
LEX_WWK	27	0.0	0.0	1	2	2	7	0.	0.

Table 2: The Routing policy from PLANITU with all other exchanges routing traffic to the regional four transits.

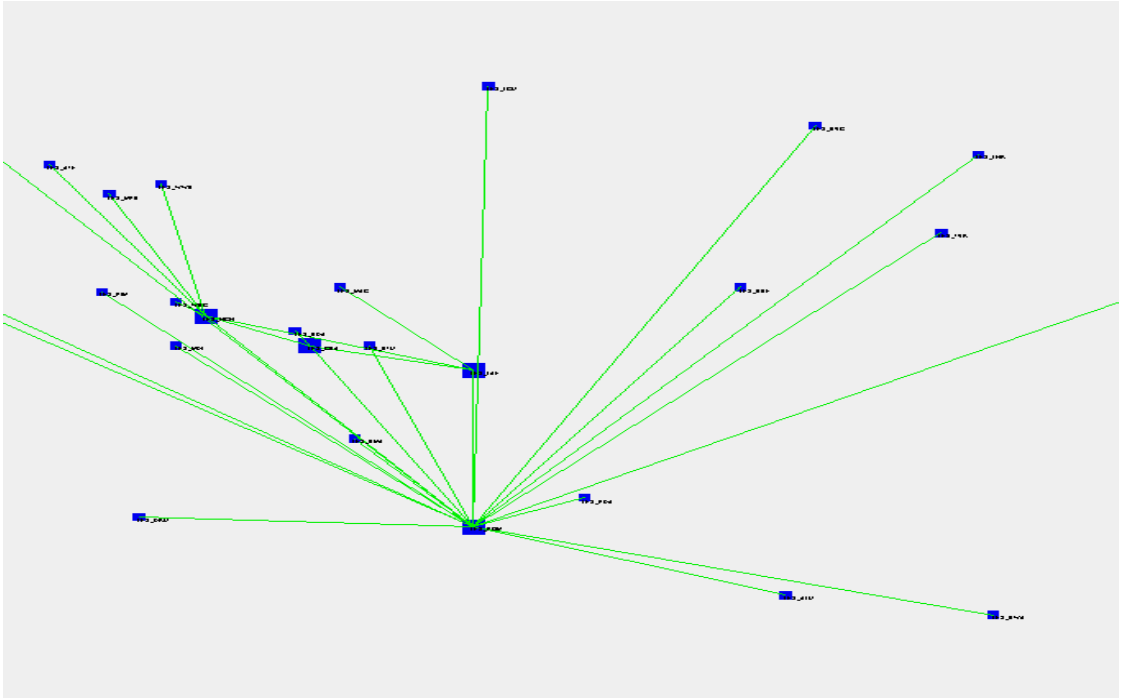


Figure 13: Optimised Traffic Routing from using PLANITU Planning Tool.

The Optimised routing transforms the routing into a three level hierarchical routing from a four level national network of the following;

- National Transit – Located in Port Moresby
- Regional Transits – Port Moresby, Lae, Mt.Hagen and Tomavatur
- Local Transits/Local Exchanges.

CONCLUSIONS

PNG as a developing country is faced with the following challenges of migration of present network, anticipated competition and regulation, and increased planning complexity. Critical planning issues and aspects of network development are needed to be addressed in transmission, signalling, routing, traffic management and quality of service.

The following case studies have addressed the need for attention to be given to the use of planning tools as compulsory if it is required to implement new, better and recognized methods for improving planning methods.

The development of telecommunications strategy for evolving of present networks and the transition to Next Generation Networks (NGN) is facilitated by the role of planning tools. The PNG Case studies show this by using specialised tools as the VPIsystems AccessMaker^{TM} and VPIsystems ServiceMaker^{TM} and for specific planning cases for broadband access solutions and backbone routing optimisation and dimensioning, using PLANITU.

Annex 3 – References

A3.1. Direct references within the text

[1]	Sami Uskela, Key Concepts for Evolution Toward Beyond 3G Networks, IEEE Wireless Communications, Feb, 2003
[2.1]	Eurescom, Project P709, Planning of Optical Network, Deliverable 3, Optical Network Planning, Volume 1 of 9: Main Report, March 2000
[4.1]	Teletraffic Engineering Handbook, www.itu.int/ITU-D/study_groups/free-publications-2.html, Document TTEhandbook1.pdf ITU, 2003
[4.2]	Brockmeyer, E. & Halstr¿m, H.L. & Jensen, Arne: The Life and works of A.K. Erlang. Transactions of the Danish Academy of Technical Sciences, 1948, No. 2. 278 pp. Available at < http://www.com.dtu.dk/teletraffic/Erlang.html >
[4.3]	Iversen, Villy B.: Reversible fair scheduling: the teletra±c theory revisited. Paper submitted to 20th International Teletraffic congress, ITC20, Ottawa, Canada, June 17- 21, 2007. 12 pp.
[6.1]	NGNI, 3rd Roadmap on technology developments or evolutionary trends, Deliverable, D3.7, Dec. 2002
[7.1]	Pioro, M. and Medhi, D.: *Routing, Flow and Capacity Design in Communication and Computer Networks*, Morgan Kaufmann Publishers (Elsevier), ISBN 0125571895, April 2004
[7.2]	Garey, M.R. and D. R. Johnson, D.R.: *Computers and Intractability: A Guide to the Theory of NP-Completeness*, W H Freeman & Co, 1979
[7.3]	Bertsekas, D.P.: *Network Optimization. Continuous and Discrete Models*, Athena Scientific, 1998
[7.4]	Wolsey, L.A..: *Integer Programming*, John Wiley & Sons, 1998
[7.5]	Dahl, G. and Stoer, M.: A Cutting Plane Algorithm for Multicommodity Survivable Network Design Problems, *INFORMS Journal on Computing*, vol. 10, pp. 1-11, 1998
[7.6]	Minoux, M.: *Mathematical Programming: Theory and Algorithms*, John Wiley & Sons, 1986
[7.7]	Alvelos, F. and Valério de Carvalho, J.M.: Comparing Branch-and-price Algorithms for the Unsplittable Multicommodity Flow Problem. *International Network Optimization Conference INOC 2003*, pp. 7-12, Evry, France, 2003
[7.8]	Anderson, C.A., Fraughnaugh, F., Parker, M. and Ryan, J.: Path Assignment for Call Routing: An Application of Tabu Search, *Annals of Operations Research*, ed. F. Glover et al., vol. 41, 1993
[7.9]	Minoux, M.: Network Synthesis and Optimum Network Design Problems: Models, Solution Methods and Applications, *Network*, vol. 19, pp. 313-360, 1989
[7.10]	Wolsey, L.A.: Strong Formulations for Mixed Integer Programming: A Survey, *Mathematical Programming* **45**, pp. 173-191, 1989
[7.11]	Nemhauser, G.L. and Wolsey, L.A.: *Integer and Combinatorial Optimization*, John Wiley & Sons, 1988

[7.12]	Yaged, B.: Minimum Cost Routing for Static Network Models, *Networks*, vol. 1, pp. 139-172, 1971
[7.13]	Yen, J.Y.: Finding the k shortest loopless paths in a network, Management Science, vol. 17, pp. 712-716, 1971
[7.14]	Bhandari, R.: *Survivable Networks - Algorithms for Diverse Routing*, Kluwer, 1999
7.2.1	EURESCOM, "Optical Networking in the local loop", Project P-413, March 1996
7.2.2	EURESCOM, "Elaboration of common FTTH guidelines", Deliverable 8, Project P-614, July 1998
7.2.3	EURESCOM, "Evaluation of broadband home networks for residential and small business users", Deliverable 11, Project P-614, 1998
7.2.4	Ims, L. A. (editor), “Techno-economic analysis of major factors of B-ISDN/ATM upgrades”. EURESCOM P614 Deliverable 3. January 1998
7.2.5	EURESCOM “Basic factors influencing optical networks”, Deliverable 2, Project P709, 1998
7.2.6	J. Stern, J.A. Quayle, S.A. Cooper, "Full Services Access Networks Requirements Specification", Eighth International Workshop on Optical/Hybrid Access Networks, Atlanta, March 2-5, 1997
7.2.7	M. Minoux, "Network synthesis and optimum network design problems: Models, solution methods and applications", Networks Vol. 19 pp. 313-36
7.2.8	N. Rosenblouth, A. Rosenbluth, M., Teller, Metropolis, E. 1953, Journal of Chemical Physics, vol.21, pp 1087-1092.
7.2.9	Otten R.H.J.M., Van Ginneken L.P.P.P., The Annealing Algorithm, 1989, (Boston:Kluwer)
7.2.10	D. Johnson, C. Aragon, L. McGeoch and C. Schevon, Optimization by simulated annealing: An experimental evaluation, Operation Research Vol. 37, No 6, Nov-Dec 1989.
7.2.11	S. Kirkpatrick, C.D. Gelatt Jr. and M. P. Vecchi, Optimization by simulated annealing, Science Vol. 220 (1983) pp. 671-680.
7.2.12	M. Pióro, Simulation Approach to the Optimization of Integral Flow Networks, International Conference on Optimization and Simulation, Singapore, September 1997.
7.2.13	M. Pióro, P. Gajowniczek, Solving Multicommodity Integral Flow Problems by Simulated Allocation, Selected Proceedings of the INFORMS Telecommunications Conference, Telecommunication Systems, Malltzer Science Publishers, 1997.
7.2.14	Glover F., Tabu search Part I. ORSA Journal of Computing 1, 1989, pp 190-206.
7.2.15	Glover F., Tabu search Part II. ORSA Journal of Computing 2, 1990, pp 4-32.

7.2.16	Goldberg D.E, Genetic algorithms in search, optimization and machine learning, Addison Wesley Pub Co Inc, 1989.
7.2.17	J. Holland, Genetic Algorithms Scientific American, July 1992.
7.2.18	L. Davis (ed.), Handbook of genetic algorithms, Van Nostrand Reinhold (1991).
7.2.19	Poul E. Heegaard , Bjarne E. Helvik, rvald Stol, „Genetic algorithms for dimensioning of full service access networks”,
7.2.20	Olsen, B.T. “OPTIMUM Methodology”. In proc. OPTIMUM workshop on Multimedia Techno-economics, Aveiro, Portugal, 20-22 October, 1997.
7.2.21	L. A. Ims, D. Myhre, B. T. Olsen, “Investment Costs of Broadband Capacity Upgrade Strategies in Residential areas”, Networks ’98, Sorrento, Italy, Oct. 17-23, 1998, pp. 641-646.
7.2.22	Ims, L A, Stordahl K, Olsen B T. “Risk analysis of residential broadband upgrade in a competitive environment”. IEEE Communications Magazine, June 1997.
7.2.23	L. A. Ims, D. Myhre, B.T. Olsen, **”**Costs of upgrading the residential telecommunications infrastructure to broadband”, In Proc. Globecom ’98, Sydney, Australia, Nov 8-12, 1998.
7.2.24	L. A. Ims, D. Myhre, B.T. Olsen, K. Stordahl, A. Ruiz-Cantera, B. Jacobs, B. Rose, M. C. Combes, J. Mononen, M. Lähteenoja, L. Budry, M. Salerno, U. Ferrero, V, Benedetto, D. Collins, **”**Economics of broadband access network upgrade strategies: the European perspective”, In Proc. OHAN ’98, Sydney, Australia, Nov 8-12, 1998.
7.2.25	L. A. Ims, J. Mononen, M. Lähteenoja, L. Budry, M. Salerno, U. Ferrero, D. Myhre, B. T. Olsen, K. C. Aas, “Key factors influencing investement strategies of broadband access network upgrades”, In Proc. ISSLS ‘98, Venice, Italy, 22-27 March 1998.
7.2.26	L. A. Ims (editor), “Broadband Access Networks - Introduction strategies and techno-economic analysis”, Chapman-Hall, January 1998.
7.2.27	Ruiz, A, Ims, L. A, Olsen, B. T, Myhre, D, Jacobs, B, “A techno-economic analysis of broadband access network evolutionary paths”, International Conference on Telecommunications 98 (ICT ‘98), Chalkidiki, Greece, June 22-24, 1998.
7.2.28	L. A. Ims, D. Myhre, B.T. Olsen, ”Investment costs of broadband capacity upgrade strategies in residential areas”, in Proc. Networks 98, Sorrento, Italy, Oct 18-23, 1998.
7.2.29	Ims, L A et al. Key factors influencing investment strategies of Broadband Access Network Upgrades, In Proc ISSLS ’98, Venice, Italy, March 22-27, 1998.
7.2.30	Ims, L A, Myhre, D, Olsen B T. ”Investment costs of broadband capacity upgrade strategies in residential areas”, in Proc. Networks 98, Sorrento, Italy, Oct 18-23,

	1998.
7.2.31	K. Stordahl et al ”Evaluating broadband strategies in a competitive market using risk analysis”, in Proc. Networks 98, Sorrento, Italy, Oct 18
7.2.32	K. Stordahl, et al Overview of risks in Multimedia Broadband Upgrades. Globecom ’99, Rio de Janeiro, Brazil, Dec 5
7.2.33	K. Stordahl, L. A. Ims, N. K. Elnegaard, F. Azevedo, B.T. Olsen, **”**Broadband access network competition - analysis of technology and market risks”, In Proc. Globecom ’98, Sydney, Australia, Nov 8-12, 1998.
7.4.1	Jochen Schiller, Mobile Communications, Second Edition, Addison-Wesley, 2003, pp 136-156
7.4.2	Jaana Laiho, Achim Wacker, Tomas Novosad, Radio Network Planning and Optimization for UMTS, John Wiliey& SONS, LTD, 2002, pp 1-6
7.4.3	Final Report on Automatic Planning and Optimization, MOMENTUM project 2003
7.4.4	www.itu.int
7.4.5	www.3GPP.org
7.4.6	J.Zhang, L.Guo, J.Yang & J.Y.Wu, Automation of 3G/4G Cellular Network Planning, 2004
7.4.7	Harri Holma, Antti Toskala, WCDMA for UMTS, 2002 John Wiley & Sons
7.4.8	Nokia White Paper, Microwave Radio Network Planning in 3G
7.4.9	Janos Harmatos (2000): Planning of Tree-Topology UMTS Terrestrial Access Networks. PIMRC 2000, London, UK, Sept.2000
7.4.10	Heller,T.Winter & M.Ziegelmann (2002): Network Planning and Topology Optimization of UMTS Access Network
7.4.11	Jaana Laiho, ‘Radio Network planning and optimization for UMTS’, John Wiley & Sons 2002
7.4.12	Wacker, A., Laiho-Steffens, ‘ Static Simulation for Studying WCDMA Radio Network Planning issues’, Proceedings of VTC’99, Houston, Texas, May 1999, pp 2436-2440
7.4.13	Chris Reece, Award Solution .Inc, White Paper “Introduction to wireless internet planning”
7.4.14	S.M.Allen, S.Hurley and R.M.Whitaker, “Spectrally Efficient Cell Planning in Mobile Wireless Networks” IEEE Transaction on communication, 2001
7.4.15	Shaoji Ni, “GPRS Network Planning on the Existing GSM System”, IEEE Transaction on Communication, 2000
7.4.16	S.Ni, Y.Liang, “Outage Probability in GSM-GPRS Cellular System with and without Frequency Hoping,” Wireless Personal Communications, Vol.14 Sept.2000, pp.215-234
7.4.17	GSM 03.30, V4.2.0, “Digital Cellular tele-communication system (Phase 2) ; Radio Network Planning Aspects,” 7/1993
7.4.18	Stephen Hurley, “Planning Effective Cellular Mobile Radio Network, IEEE Transactions on Vehicular Technology, 51(2): 243-253, 2002
7.4.19	Stuart Allen, Bachir Belloul, Hurley, Saunders, M.Whitaker, “smart cell planning and optimization for UMTS”, 2004, 3G UK

7.4.20	Eisenblatter, Greerdes, Junglas, Koch, Kurner, Martin, “Final Report on Automatic Planning and Optimization” MOMENTUM project 2003
7.4.21	Chae Y.Lee, Hyon G.Kang, “Cell Planning with Capacity Expansion in Mobile Communications: A Tabu Search Approach”, IEEE Transaction on Vehicular Technology, 2000
7.4.22	B.Chamaret, S.josslin, P.Kuonen,, “ Radio network optimization with maximum independent set search”, IEEE VTC, pp, 770-774, 1997
7.4.23	H.D.Sherali, “Optimal location of transmitters for microcellular radio communication systems design”
7.4.24	K.Tutschku, “Interference minimization using automatic design of cellular communication networks,” IEEE VTC, pp 634-638, 1998
7.4.25	K.Tutschku, “Demand-based radio network planning of cellular communication systems”, in Pro.IEEE Infocom’ 98, 1998, pp.1054-1061
7.4.26	P.Calegari, “Genetic approach to radio network optimization for mobile systems”, IEEE VCT, pp.755-759, 1997
7.4.27	P.Calegari, “Parallel island-based genetic algorithm for radio network design”, J.Parallel Distrib.Comput, vol.47, pp.86-90, 1997
7.6.1	M. Padberg: „Classical cuts for mixed-integer programming and branch-and-cut”, *Mathematical Methods in Operations Research*, 2001
7.6.2	L. Wolsey: „*Integer Programming*”, Wiley,-Interscience, 1998
7.6.3	W-K. Chen: „Flows in Networks”, Wiley-Interscience, 1990

A3.2. Additional references for extension

Additional reference documents will be available in ITU, collected as not published (or officially published) contributions from countries and companies.

[1]	Hadi Maleki Parast, Computing of bandwidth necessary for ring creation, Data Communication of Iran , Head of Design & Consultative, Iran, Tehran
[2]	Nokia Multi-Access to IP Multimedia, White paper, 2003
[3]	Nokia HSDPA Solution, White paper, 2003
[4]	Intelligent Edge, White paper, 2003
[5]	4G mobile networks, Telecomunication Company of Iran / TCI
[6]	Programme du Réseau des Télécommunications et le Développement de son Architecture, GABON TELECOM S.A., Libreville, Gabon, Novembre 2003
[7]	Contributions to the ITU-T Ref. Manual , *Networks Design, Dimensioning and Optimization (Chapter 7,)* DINH Van Dzung, Vietnam
[8]	RURAL AREAS DIGITALIZATION, Ashot Mamyan, ARMENTEL, Armenia

A3.3. ABBREVIATIONS/GLOSSARY

The list of abbreviations is intended to include all those used in these Reference Manual.

The list will be further completed and updated

A	
ADM	Add/Drop Multiplexer
ADSL	Asymmetric Digital Subscriber Line
ATM	Asynchronous Transfer Mode
B	
BW	Bandwidth
BWA	Broadband Wireless Access
C	
CAPEX	Capital Expenditure
CATV	Cable Television
CDMA	Code Division Multiple Access
CN	Core Network
CS	Circuit Switching
CSCF	Call Session Control Function
D	
DECT	Digitally Enhanced Cordless Telephone
DLC	Digital Loop Carrier
DSLAM	Digital Subscriber Line Access Multiplexer
DWDM	Dense Wavelength Division Multiplexing
E	
EBIT	Earnings Before Interest and Taxes
EBITDA	Earnings Before Interest and Taxes, Depreciation and Amortisation
F	
FDD	Frequency Division Multiplexing
FDMA	Frequency Division Multiple Access
FO	Fiber Optics
FTTH	Fiber To The Home
FTTU	Fiber To The Unit
G	
GPRS	General Packet Radio Service
GSM	Global System for Mobile Communications
GW	Gateway
H	
HDSL	High bit Rate Digital Subscriber Line
HFC	Hybrid Fiber Coax
I	
IN	Intelligent Network
IP	Internet Protocol
IRR	Internal rate of return
ISDN	Integrated Services Data Network

IT	Information Technology
ITU	International Telecommunication Union
ITU-D	International Telecommunication Union – Development Sector
ITU-R	International Telecommunication Union – Radiocommunication Sector
ITU-T	International Telecommunication Union – Telecommunication Sector
J	
K	
L	
LEX	Local Exchange
LMDS	Local Multipoint Distribution System
M	
MGCF	Media Gateway Control Function
MM	Multimedia
MPLS	Multi Protocol Label Switching
N	
NAS	Network Access Server
NGN	New Generation Network
NPV	Net Present Value
NWA	Network Wireless Access
O	
OPEX	Operational Expenditure
P	
PDSN	Packet Data Serving Node
PDSN	Public Data Switched Network
PLMN	Public Land Mobile Network
POTS	Plain Old Telephone Service
PS	Packet Switching
PSTN	Public Switched Telephone Network
Q	
QoS	Quality of Service
R	
RAN	Radio Access Network
RFP	Request for Proposal
RSU	Remote Service Unit
S	
SDH	Synchronous Digital Hierarchy
SDMA	Space Division Multiple Access
SDSL	Symmetrical Digital Subscriber Line
SLA	Service Level Agreement
SME	Small Medium Enterprises
SOHO	Small Office Home Office
SS	Signaling System
STM	Synchronous Transfer Module
T	
TEX	Transit Exchange
TMN	Telecommunications Management Network
U	
UMTS	Universal Mobile Telecommunication System
V	
VC-12	Virtual Channel/Virtual Container
VNO	Virtual Network Operator
VoIP	Voice Over IP
W	

WDM	Wavelength Division Multiplexing
WIP	Wireless Internet Protocol
WL	Wireline
WLL	Wireless Local Loop
WLAN	Wireless Local Area Network
X	
xDSL	Generic Digital Subscriber Line
Y	
Z	