Miika Heini

Charging Model Development For Data Center Network Environment

School of Electrical Engineering

Thesis submitted for examination for the degree of Master of Science in Technology.

Espoo 11.10.2011

Thesis supervisor:

Prof. Heikki Hämmäinen

Thesis instructor:

M.Sc. Tero Latvakangas

Aalto University
School of Electrical
Engineering

AALTO UNIVERSITY
SCHOOL OF ELECTRICAL ENGINEERING

ABSTRACT OF THE
MASTER'S THESIS

Author: Miika Heini

Title: Charging Model Development For Data Center Network Environment

Date: 11.10.2011 Language: English Number of pages: 11+68

Department of Communications and Networking

Professorship: Networking technology Code: S-38

Supervisor: Prof. Heikki Hämmäinen

Instructor: M.Sc. Tero Latvakangas

The purpose of this thesis is to investigate how an organization can improve and organize its charging activities in data center network environment. Theory part of this study introduces relevant concepts for this case and discusses their properties. Case study in this thesis is conducted to form an overall picture about current challenges in the current environment and compare it to general practices used in the industry. Based on the case study results, future state and improving actions are proposed.

Theory part of this study is divided in three topics. First, service management practices based on ITIL processes are explored. Second part concentrates on network services that are relevant in data center environment. Third and last theory part explores the economical models related to network services. Theories related to cost collection and allocation, pricing and charging are discussed. The empirical part of this study reflects the explored theories to recognize the largest challenges and analyzes possible solutions to solve these challenges. The current state in the organization is mapped by interviewing key personnel and observing the daily working routines of network unit.

Case study reveals that the main challenges lie in the service definitions and cost collection and allocation practices. Based on the results, service definitions are proposed to be written out in more detail and service structure is to be divided into vertical technology areas. This structure will allow more accurate cost collection using horizontal activities. Based on the new costing structure, pricing and charging practices are simplified using cost-based pricing.

Keywords: Network service, data center, charging, pricing, ITIL, Activity-based costing

AALTO-YLIOPISTO
SÄHKÖTEKNIIKAN KORKEAKOULU

DIPLOMITYÖN TIIVISTELMÄ

Tekijä: Miika Heini

Työn nimi: Verkkopalveluiden veloitusmallin kehitys datakeskusympäristöön

Päivämäärä: 11.10.2011 Kieli: Englanti Sivumäärä: 11+68

Tietoliikenne- ja tietoverkkotekniikan laitos

Professuuri: Tietoverkot Koodi: S-38

Valvoja: Prof. Heikki Hämmäinen

Ohjaaja: DI Tero Latvakangas

Tämän diplomityön tarkoituksena on tutkia, kuinka organisaatio voi kehittää ja järjestää verkkopalveluiden veloituskäytäntöjä datakeskusympäristössä. Työn teoriaosuus esittelee työn kannalta keskeiset konseptit sekä niiden ominaispiirteet. Empiirisessä osuudessa kartoitettaan nykytilanteen haasteet ja verrattaan niitä yleisiin käytäntöihin. Kartoituksen pohjalta luodaan suunnitelma ja kehitysehdotukset veloituskäytäntöjen kehittämiseksi.

Työn teoriaosuudessa esitellään tutkimuskysymyksen kannalta relevantit konseptit. Ensiksi käydään läpi palvelunhallintakäytäntöjä pohjaten ITIL-prosesseihin. Tämän jälkeen perehdytään datakeskusympäristön kannalta oleellisiin verkkopalveluiden piirteisiin. Lopuksi esitellään verkkopalveluihin liittyviä ekonomisia malleja. Esiteltävät mallit liittyvät kustannusten allokointiin, hinnoitteluun ja veloituskäytäntöihin. Työn empiirisessä osiossa pyritään tunnistamaan organisaation suurimmat haasteet peilaamalla tämänhetkisiä toimintatapoja teoriaosuudessa esiteltyihin konsepteihin. Tältä pohjalta analysoidaan ratkaisuvaihtoehtoja, joilla tunnistetut haasteet voidaan ratkaista. Nykytilan selvittämiseksi haastatellaan organisaation avainhenkilöitä ja seurataan verkkoyksikön päivittäisiä työrutiineja.

Tutkimus paljastaa suurimpien haasteiden liittyvän palvelukuvauksiin, kustannusten keräämiseen sekä allokointiin. Tulosten pohjalta esitetään, että palvelukuvausten sisältöä tarkennetaan ja palvelurakenne jaettaan selkeämmin vertikaalisiin teknologia-alueisiin. Tämä rakenne mahdollistaa aktiviteettipohjaisen kustannusmallin käyttöönoton nojaten horisontaalisiin aktiviteetteihin. Näin kustannukset saadaan kohdistettua palveluille tarkemmin. Uuden kustannusmallin pohjalta hinnoittelu- ja veloituskäytännöt yksinkertaistuvat käyttämällä kustannuspohjaista hinnoittelua.

Avainsanat: Tietoliikennepalvelu, datakeskus, veloitus, hinnoittelu, ITIL, Aktiviteettipohjainen kustannusmalli

Preface

This Master's Thesis has been written as a partial fulfillment for the Master of Science degree in Aalto University School of Electrical Engineering. This thesis was conducted for Tieto Corporation as a part of internal development project.

I want to express my gratitude to those who have contributed to this work. First, I want to thank my supervising professor Heikki Hämmäinen for his insights, guidance and valuable feedback throughout the process. Secondly, I want to thank my instructors, Tero Latvakangas and Kim Imponen, for the possibility to conduct this study at Tieto and the great support I received. I also want to thank all my colleagues for the discussions and feedback about the subject.

This thesis now concludes one period in my life. I want to thank my family and friends for all the support during this long road.

Last but definitely not least, I want to thank my beloved wife, Anna: You have been my guiding light through the tough last years.

Otaniemi, 11.10.2011

Miika Heini

Contents

List of Figures

List of Tables

Abbreviations

ABC	Activity-Based Costing
CAPEX	Capital Expenditures
CMDB	Configuration Management Database
DNS	Domain Name System
DWDM	Dense Wavelength Division Multiplexing
FCC	The Federal Communications Commission
FDC	Fully Distributed Cost approach
FTP	File Transfer Protocol
HTTP	HyperText Transfer Protocol
ICT	Information and Communications Technology
IDS	Intrusion Detection System
IETF	Internet Engineering Task Force
IP	Internet Protocol
IPS	Intrusion Prevention System
ISO	International Organization for Standardization
ISP	Internet Service Provider
ITIL	Information Technology Infrastructure Library
LAN	Local Area Network
LRIC	Long-Run Incremental Cost
OPEX	Operational Expenditures
OSI	Open Systems Interconnection reference model
PPP	Point-to-Point Protocol
PBA	Patterns of Business Activity
QoS	Quality of Service
RAS	Remote Access Service
RTT	Round Trip Time
SLA	Service Level Agreement
SMTP	Simple Mail Transfer Protocol
SNMP	Simple Network Management Protocol
SPM	Service Portfolio Management
SWOT	Strengths, Weaknesses, Opportunities and Threats
SSL	Secure Sockets Layer
TCP	Transmission Control Protocol
UDP	User Datagram Protocol
VPN	Virtual Private Network

1 Introduction

Data centers have become more and more important for enterprises by hosting business critical services. Continuously increasing demand of availability and connectivity has set new challenges for service providers to organize their services in the best possible way. Though the cloud services are gaining popularity all the time and many companies are transferring their services in to the cloud, the traditional managed services are not fading away any time soon.

Network services can be seen as a skeleton of the data centers. They provide fast, secure, reliable and scalable environment for servers that reside in data centers. External connectivity services guarantee connections for customers and their partners to reach their business critical services. Internal connectivity provides means of constructing complex customer environments.

Economical and technical demands have set challenges for providers to develop cost efficient ways of working. Shared network infrastructure is one example of utilizing network capacity more efficiently. Network services of multiple customers are provided using same equipments. To make such solution possible, different virtualization techniques are employed to separate customers from each other.

Managing shared network environment differs from the traditional dedicated networks. Maintenance windows and SLAs (Service Level Agreements) for network services are common and have to be compromised not to disturb the activities of customers. Failure situations often affect more than one customer, and thus it is crucial to see the consequences of failures and map their effects. IT service management practices are in a key role of ensuring continuous service: standard and common ways of working are emphasized together with standardized equipment and design principles.

In an organization that is providing IT services, the network is seen as a backbone to build actual end-user services. The value of network services for customers is rather challenging to define. Nevertheless, these services are in a key role, and generate significant share of the total costs. Thus it is crucial to know the production costs as precisely as possible and to be able to share these costs fairly to all network components and services. Fairly allocated costs between services allow many ways for setting the price and deciding how the services could be charged.

Service definitions and packaging choices are to help in creating the service offerings. By taking care of the cost allocations and charging models for each service separately, the overall offering development becomes easier and requires less customization. All these actions can be seen as an aim to improve customer value and quality.

1.1 Objectives and scope

This thesis has been conducted for Tieto Corporation, Data center services business unit and Networks team. Tieto is one of the leading IT service providers in Northern

Europe.

The main objective of this Master's thesis is to gather information about how an organization can organize and improve its charging of network services in data center environment. The main question to answer is what actions are required to simplify and develop the charging activities of network services? The research question can be broken down into two parts. First part focuses on the current situation in the case organization:

1. How are the data center networks organized?
2. What network services are running in the data centers?
3. What are the current challenges to overcome in the organization?

Second part concentrates on the target state that organization aims to reach in the future.

5. How should the service offering be developed?
6. What costs are related to service provisioning and how the costs should be allocated and followed?
7. How should the costs be retrieved by pricing the services and charging?

To reach the objectives set to this thesis, relevant frameworks and models are explored from the literature to gain understanding in charging processes of network services. Issues will be addressed from the strategic point of view.

The scope for this work is limited to discuss the shared network infrastructure and continuous services, but the results are to be kept as general as possible to be usable in other environments too. Due to time constrains the deployment of the model will not be conducted as a part of this thesis.

1.2 Research methods

This thesis will be conducted using literature review and case study method. The literature review will concentrate on two main topics: 1) Introducing data center network solutions and the services in the data center networks. 2) Exploring general cost allocation, pricing and charging models that are used in telecommunications.

The case study will be performed about an organization that is offering managed IT services. First, the current situation in the case organization will be identified. Based on the analysis, challenges are to be identified and as a result, a target state will be created. Actions to reach this target state will be proposed. These results are gathered using interviews and observation in the networks team.

1.3 Structure

This thesis is divided into 8 chapters. Chapters 2, 3 and 4 explore the theoretical background relevant for this study. Chapter 2 introduces IT service management practices that relate to the research questions. Chapter 3 discusses characteristics of network infrastructure in data center environment and gives an outlook of network services. In chapter 4, costs related to network services are discussed and classified. Cost allocation methods are presented and connected with different pricing and charging models.

Chapter 5 introduces the case organization and describes the case study methods, conduction of the study and the results. Chapter 6 analyzes the results and compares the presented theoretical models and their suitability to the needs of the organization.

Chapter 7 presents the proposed future vision and proposes an action plan to reach the target. Last chapter concludes the work done and gives recommendations for future research.

2 IT Service Management

This section discusses the relevant practices proposed by ITIL and other relevant literature. Section begins with listing the characteristics of services and introducing the nature of service provisioning. Then the topics of service portfolio management, demand management and financial management are discussed. As a last part of this section, typical service operation processes are introduced.

Information Technology Infrastructure Library (ITIL) is a collection of best practices in IT industry. ITIL is widely used in IT industry as a base for own processes. Current version of ITIL that is in use is version 3, which is usually referenced as ITILv3. ITILv3 consists of five books: Service strategy, service design, service transition, service operation and continual service improvement. These five books cover the whole service life cycle from the process perspective.

2.1 Service characteristics

The definition of service is diverse, and thus challenging to present. From the literature we can find several different definitions. ITIL (OGC, 2007a, p. 16) defines service to be *a means of delivering value to customers by facilitating outcomes customers want to achieve without the ownership of specific costs and risks*. To put it in other words, customers seek outcomes that have positive effect on activities, objects and tasks to create conditions for better performance, but do not want to be accountable or have the ownership of all the associated costs and risks.

Services also have characteristics which separate them from traditional goods. Wolak et al. (1998) made an investigation about the four most common characteristics that are used to describe services. Following characteristics are often referenced as *IHIP*:

- *Intangibility* – Services are imperceivable by senses in physical manner and challenging to describe in mental manner.
- *Heterogeneity* – It is relatively challenging to standardize service outcomes and thus there might be variability in individual units of service production. Each service user subjectively evaluates service outcomes against their expectations.
- *Inseparability* of production and consumption – Services are produced and consumed at the same time. Often the service delivery or production cannot start before the user or some other party has defined some inputs.
- *Perishability* – One unit of service can be consumed only once. Service provisioning is crucial for matching capacity and demand.

The idea of ITIL is based on a Service Lifecycle (illustrated in Figure 1). Service Design, Service Transition and Service Operation are progressive phases of the Lifecycle that represent change and transformation. Service Strategy represents policies and objectives. Continual Service Improvement represents learning and improvement. As it is clear that the business needs and technology change over time, the Lifecycle structure is important for service management. Without a structure it is difficult to learn from gained experience and improve the services and offerings constantly.

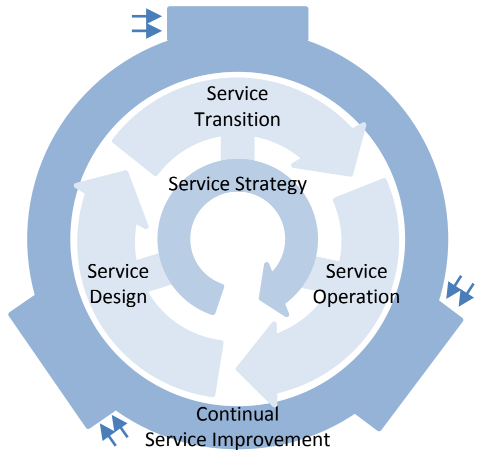


Figure 1: The Service Lifecycle according to ITIL. (OGC, 2007a, p. 24)

2.1.1 Interaction in service provisioning

Figure 2 illustrates the interactions between customers, business and technology in service provisioning. Kilkki and Pohjola (2006) distinguish two actors in this framework. These two actors have independent aims: Customers want to maximize their benefit and the business owners want to maximize their income. Also two main tasks are identified in this framework: Service provider deals with the customer relationship and designs service portfolio, pricing and marketing. Network operator builds and manages the network and divides the resources between customers and applications. These two tasks may be offered by two separate companies or by separate parts of one company.

In business-to-business environment, business needs steer user behavior and create needs to the users. User needs are satisfied with technological solutions. Customers are to be separated from the users; customer in this case is the one that pays the service and is the one who generate the business need. Thus the business aspect is to satisfy the needs of customer behavior by service offerings.

2.1.2 Service contracts

Services are provisioned according to service contracts between the network provider and the customer. Service contract defines the framework for the service. Baseline for contract is that *the network commits to deliver a service with given quality and performance characteristics, and the customer commits to interact with the network in a given way.* (Courcoubetis and Weber, 2003, p. 33) Table 1 lists three different contract examples to define the service quality. In this example, contract A is

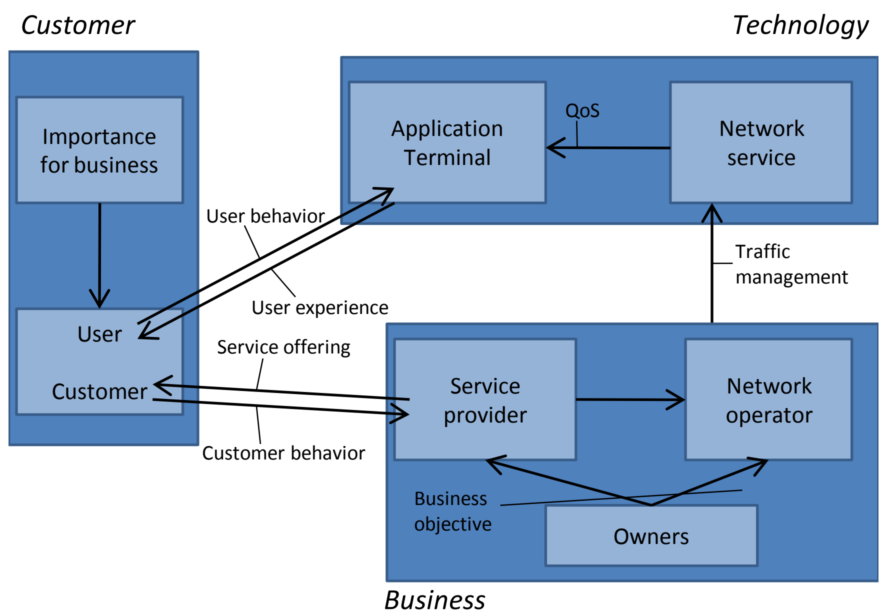


Figure 2: Interaction in service provisioning, adapted from Kilkki and Pohjola (2006).

Table 1: Three examples of network service contract in terms of quality.

Contract A	Contract B	Contract C
There will be no data loss provided the rate of the source stays below 100 Mbps.	There will be no data loss provided the rate of the source stays below h Mbps, where the network can vary h dynamically between 100 and 110 Mbps.	The data loss rate will be less than 0.00001% provided the rate of the source stays below h Mbps, where the user can vary h dynamically between 100 and 110 Mbps.

defined in terms of *static parameters*. Static parameters are set at the time when a contract is established and remain constant throughout the life of the contract. In this scenario, network guarantees the resources to be available for the service. Contracts B and C are defined in terms *dynamic parameters*. Dynamic parameters are updated during the life of the contract. Contract B has a static part that guarantees 100 Mbps, but includes also a dynamic extra part for rate between 0 and 10 Mbps. Contract C has a static parameters of 100 Mbps and 0.00001%. The probability in contract C makes it lossy even with rate below h Mbps. By pricing the dynamic parameters, the network gives customers incentive to purchase smaller value of them, thus the network ends up needing to fulfill only the minimal static part of the contract. (Courcoubetis and Weber, 2003, pp. 37-38)

Especially between large corporations, service contracts are usually called as *Service*

Level Agreement (SLA). SLAs often include also other service related functions, such as service desk and customer support. Table 2 presents different availability requirements for service and allowed downtimes.

Table 2: Availability measures in terms of percentage.

Availability requirement	Allowed annual downtime	Allowed weekly downtime
90%	36.5 d	16.85 h
95%	18.25 d	8.42 h
98%	7.3 d	3.37 h
99%	3.65 d	1.68 h
99.5%	1.825 d	50.54 min
99.8%	17.52 h	20.22 min
99.9%	8.76 h	10.11 min
99.95%	4.38 h	5.05 min
99.99%	52.56 min	1.01 min
99.999%	5.256 min	6.06 s

Service level parameters have a large impact on the production costs of a service. The higher the requirements for bandwidth and delays are, the more costly it is to provide the service. The same principle applies also to the availability. 99.99% service availability allows the service to be unreachable about one minute per week. That time window is not enough for maintenance work without special arrangements. High availability usually requires redundant infrastructure, which leads to increase in production costs.

2.2 Service portfolio management

According to ITIL service strategy (OGC, 2007a, p. 119), a service portfolio should describe a provider's services in terms of business value. It should articulate business needs and the provider's response to those needs. Service portfolio is to clarify or help to clarify the following strategic questions:

- Why should a customer buy these services?
- Why should they buy these services from us?
- What are the pricing or chargeback models?
- What are our strengths and weaknesses, priorities and risk?
- How should our resources and capabilities be allocated?

Service Portfolio Management (SPM) should be a continuous process, as conditions and markets change constantly. Services may no longer be optimal due to events such as mergers and acquisitions, redeployed strategy and changing technologies. Thus there is a constant need to monitor, measure, reassess and rebalance the portfolio. ITIL (OGC, 2007a, p. 123) proposes following working methods (also illustrated in Figure 3) for dynamic portfolio management:

Define – Process starts with collecting information from all existing services as well as proposed services. Documenting the existing service information helps to understand the opportunity costs of the existing portfolio. When the current state is understood, it is easier to assess whether to reallocate capabilities and resources or to stick to existing portfolio. Due to a cyclic nature of SPM, definition phase is not only to create an initial inventory of services but to also validate the data on recurring basis. Each service in the portfolio should include a business case. Business case is a model of what a service is expected to achieve and its justification for pursuing a course of action to meet stated organizational goals.

Analyse – This part of the process is where the strategic top-down questions are to be answered:

- What are the long-term goals of the service organization?
- What services are required to meet those goals?
- What capabilities and resources are required for the organization to achieve those services?
- How will we get there?

These questions are to address the four Ps: *Perspective*, *Position*, *Plan* and *Patterns*. The answers guide not only the analysis but also the desired outcomes of SPM.

Approve – Previous phases should have led to a well-understood future state. This phase is where decisions to approve or disapprove the future state take place. If the state is approved, it includes the corresponding authorization for new services and resources. The outcomes of approvals can be divided into six categories:

- *Retain* services that are aligned with and are relevant to the organization's strategy.
- *Replace* services that have unclear and overlapping business functionality.
- *Rationalize* services that are composed of multiple releases of the same operating system, multiple versions of the same software and/or multiple versions of system platforms providing similar functions.
- *Refactor* services that meet the technical and functional criteria of the organization but display fuzzy process or system boundaries.
- *Renew* such services that meet the functional fitness criteria, but fail technical fitness.
- *Retire* services that do not meet minimum levels of technical and functional fitness.

Charter – Start with a list of decisions and action items. These items are to be communicated to the organization clearly and unambiguously. Communicated decisions should be correlated to budgetary decisions and financial plans. Budget allocations should enforce the allocation of resources.

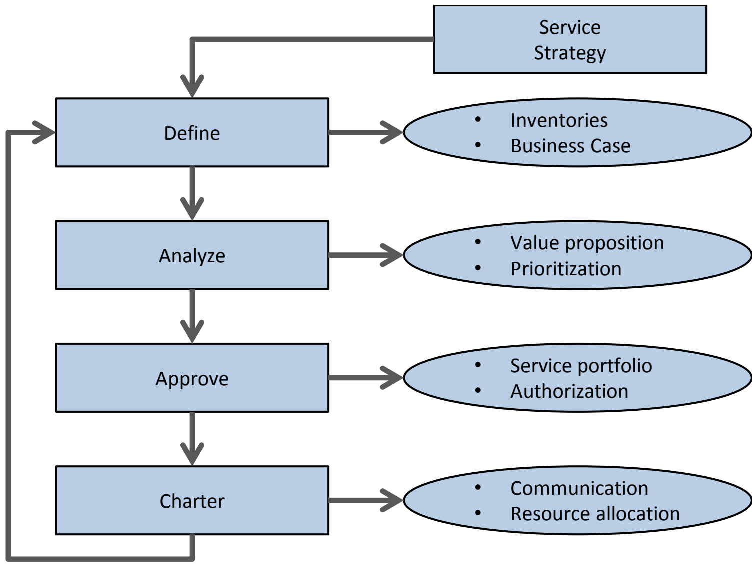


Figure 3: Service Portfolio Management process (OGC, 2007a, p. 124).

2.3 Demand management

Managing the demand of services is crucial. Poorly managed demand forms a risk for service providers because of the uncertain nature of demand. Excess capacity creates costs that the customers are reluctant to pay for, but at the same time there must exist unused capacity to guarantee service levels. Insufficiently managed capacity may have impact on the quality and growth possibilities of delivered services. Services are by nature different than traditional goods: their production cannot occur without demand that consumes the output. This leads to the fact that demand and capacity are tightly coupled in service systems. ITIL gives recommendations for Demand Management process in service strategy part. (OGC, 2007a, p. 129)

2.3.1 Sources of demand

Business processes are the primary sources of demand for services. Patterns of business activity (PBA) influence the demand patterns seen by the service providers. Thus it is crucial to study and understand the business of the customers to be able to respond demand patterns with capacity. The idea is illustrated in Figure 4. Demand patterns have a tendency to occur at multiple levels. ITIL presents Activity-based Demand Management to address the business activity patterns. Analyzing the patterns of business activity brings benefits in form of inputs to service management functions and processes: (OGC, 2007a, pp. 130-131)

- Service Design can optimize designs to suit demand patterns.
- Service Catalogue can map demand patters to appropriate services.

- Service Portfolio Management can approve investments in additional capacity, new services, or changes to services.
- Service Operation can adjust allocation of resources and scheduling.
- Service Operation can identify opportunities to consolidate demand by grouping closely matching demand patterns.
- Financial Management can approve suitable incentives to influence demand.

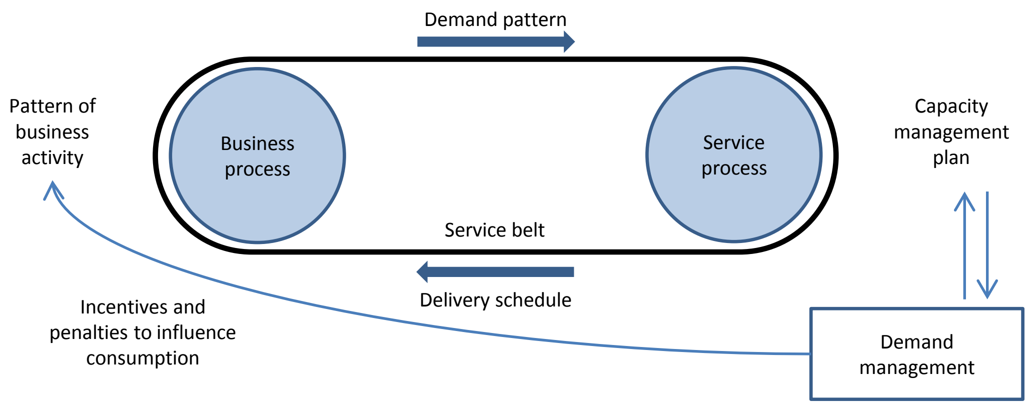


Figure 4: Business activity influences patterns of demand for services. (OGC, 2007a, p. 130)

As the business activities drive the demand for services, such customer assets as people, processes and applications generate patterns of business activity. PBA can be identified, codified and shared across processes for clarity and completeness of detail. Attributes such as frequency, volume, location and duration can be used to describe business activity. These attributes are associated with other requirements such as security, privacy and latency or tolerance for delays. As each PBA differs substantially from each other, a catalog of patters can be developed.

User Profiles can be based on roles and responsibilities within organizations for people, functions and operations for processes and applications. In many contexts, business processes and applications are treated as users, as they are not actively executed or controlled by personnel. When PBAs are well defined, user profiles can be associated with one or more PBA. This way aggregations and relations can be made between patterns of business activities and their respective user profiles. Matching the patterns using PBA and user profiles, allows systematic approach to understand and manage demand from customers. This leads to an ability to sort and serve the demand with appropriately matched services, service levels and service assets, and thus improves value for both, customers and provider by eliminating waste and poor performance. (OGC, 2007a, p. 132)

2.3.2 Offerings to match demand

Offerings to match demand are developed based on the service portfolio. Offering can be seen as a package of services that satisfy the needs of customer's business pattern activities. Services can be divided into core services and supporting services:

Core services are to deliver the basic outcomes desired by the customer. They represent the value that the customer wants and for which customers are willing to pay. Core services anchor the value proposition for the customer and provide the basis for their continued utilization and satisfaction.

Supporting services either enable or enhance the value proposition. Enabling services are basic factors and enhancing services are excitement factors. Supporting services will either provide the basis for differentiation or represent the minimum requirements for operation. As excitement factors, enhancing services provide differentiation. As basic factors, enabling services only qualify the provider for an opportunity to serve customers. Enabling services are necessary for customers to utilize the core services effectively. As basic factors, customers take their availability for granted and do not expect to be charged for the value of such services.

Packaging of core and supporting services is an essential aspect of market strategy. Bundling of core services with supporting services has implications for the design and operation of services. Decisions have to be made whether to standardize on the core or the supporting services. The same level of differentiation can be achieved either way (see Figure 5), but the costs and risks involved may be different.

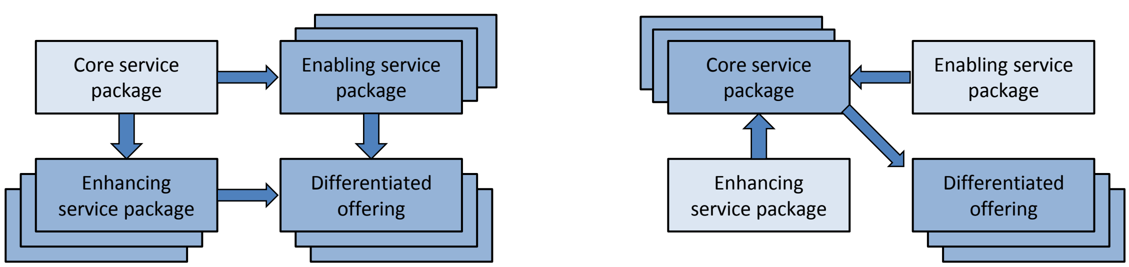


Figure 5: Same level of differentiation can be achieved either standardizing core services or supporting services. (OGC, 2007a, p. 133)

Service providers may adopt different strategies for core and supporting services. Standardization and consolidation can be driven in supporting services to leverage economies of scale and to reduce operating costs, while core service packages are specifically designed for particular customers. Another possibility is to standardize the core service packages and use supporting services to differentiate the offerings across customers and market segments. This aspect is important for service providers who need to balance the differing needs of several enterprises and business units while trying to keep costs down across the portfolio to remain competitive.

2.4 Road map of implementing financial management

ITIL Service Strategy part (OGC, 2007a, pp. 111-112) propose a checklist of implementing financial management process. It is intended to be a representation of a phased approach to the implementation, not a direct project plan. Tracks are illustrated in Figure 6.

Track 1 – Plan addresses the very fundamental questions of the service provisioning, stakeholders and processes related:

- Identify all parties that receive and/or provide financial information
- Identify the service deliverables and chargeback model to be used
- Determine systems that contribute and provide financial data
- Determine the operating model to be used to set the goal for accounting and valuation
- Prepare policy and operating procedures list

Track 2 – Analyse includes gathering in-depth details of the previously identified items:

- Are all processes and information required to produce deliverables accounted?
- What are the current costs and how and where are they accounted? Centralization of expenditures is a prerequisite.
- Service valuation should be performed

Track 3 – Design creates the outputs that are expected from financial management implementation:

- Identify processes in place within IT and design clear hooks into financial management
- Prepare and test valuation models for their appropriateness to the business environment
- Design the accounting processes according to the outcome of previous learnings
- Create the chosen chargeback methodology
- Complete the design of policies and procedures
- Define required roles and responsibilities

Track 4 – Implement the planned processes by activating them to use.

Track 5 – Measure the success of the implementation:

- Audit credibility gaps between value being received and price being paid

- Use tools such as balanced scorecard, regular communication, meaningful data and mappings to business strategy

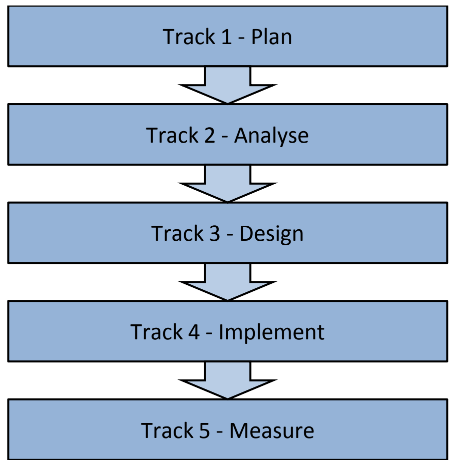


Figure 6: Financial Management implementation checklist according to ITIL service strategy. (OGC, 2007a)

According to ITIL, chargeback model for IT should provide accountability and transparency. Visibility is to be brought through identification of Service Portfolios and catalogs, and valuing those IT services. Accountability refers to the ability to provide expected services as agreed with business. Charging without taking into account the operating model typically does not deliver desired levels of accountability and visibility.

Charging should be done to encourage behavioral changes related to steering demand for IT Services. Charging must add value to the business and be in business terms, and it should have a degree of simplicity appropriate to the business culture. (OGC, 2007a, p. 110) ITIL Service strategy proposes five different scenarios for chargeback models:

Notional charging – these alternatives address whether an entry to financial systems should be made. For example, two book method can be used to record costs in one fashion, while second book is kept but not recorded. Second book provides the same information but reflects what would have happened if alternative method had been used.

Tiered subscription – involves varying levels of warranty and/or utility for a service or service bundle, all of which have been priced, with the appropriate charge-back models applied. Most commonly referred to as gold, silver and bronze levels of service.

Metered usage – involves a more mature financial environment and operational capability, where demand modeling is incorporated with the utility computing capabilities to provide confidence in the capture of real-time usage. This consumption information is then translated into customer charging based on various service increments that have been agreed, such as hours, days or weeks.

Direct Plus – simplistic model where costs that can be allocated directly to a service are charged accordingly with some percentage of indirect costs shared amongst all.

Fixed or user cost – most simplistic model. Takes the costs and divides by an agreed denominator such as number of users. This model contributes little to affecting customer behavior, or identifying true service demand or consumption, but does allocate the costs to the bottom line in the easiest fashion.

2.5 Service operation

ITIL also defines service operation processes, which are covered in the Service Operation book. Service operation is responsible for coordinating and carrying out activities and processes required to deliver and manage services at agreed levels to business users and customers. Service operation is also responsible for ongoing management of technology that is used to provide the services. The scope of service operation includes the following four aspects (OGC, 2007b, pp. 13-14):

Services – Any activity that forms part of a service is included in service operation, regardless of who is performing such activity.

Service management processes – Management and execution of service management processes are performed in service operation, though some of the processes originate from other levels of service lifecycle.

Technology – Managing the technology that is required by the services is an integral part of the service management.

People – Regardless of the other aspects, it is always people that drive the demand for services and decide how things will be done. It is also people who manage the other aspects, thus it is all about people.

Service operation consists of processes. Below is described the key processes that must link together to enable effective IT service support structure.

Incident management
Incident is a disruptive event that has or could have effect on service. The purpose of incident management is to restore the normal service operation as quickly as possible and minimize the effect of incident on business operations. Normal service levels are defined in SLAs. (OGC, 2007b, pp. 46-55)

Problem management
Problem is an unknown cause of one or more incidents. Problem management

is responsible for managing problems during their lifecycle. Primary objectives are to prevent problems and resulting incidents from occurring, eliminate recurring incidents and minimize the effects of incidents that cannot be prevented. (OGC, 2007b, pp. 58-68)

Change management

Change management is mainly part of service transition processes, but some of the aspects are tightly tied in service operation. Often service operation issues are drivers for changes, and the changes are implemented by service operation staff. (OGC, 2007b, pp. 72-73)

Capacity and Availability management

Capacity management as an umbrella term is more of longer term strategic planning. Nonetheless, it is also tightly related to daily work as the service capacity and performance need to measured. Capacity planning should also be a continuous process as the requirements set in design phase tend to change due to the volatility of services. Service operation is responsible of making the IT service available to users according to the SLAs. Availability management is responsible for conforming the defined SLAs. (OGC, 2007b, pp. 73-77)

Continuity management

Continuity management relies on testing and executing the recovery plans as determined. Activities that service operation staff is involved are such as risk assessment, execution of risk management measures and assistance in plan creation. (OGC, 2007b, p. 77)

3 Data center networks

Data centers are the home of computational power, storage and business critical applications. Network infrastructure is the critical enabler of data centers. First part of this chapter discusses the network architecture and topology and the services in the network. Second part explores the current research done in the field of traffic characteristics. Last part of this section introduces traffic measurement methods and metrics and describes how they can support charging activities.

3.1 Network architecture

The most used design principle of data centers is three level tree topology. (Al-Fares et al., 2008; Benson et al., 2010) In this model, the functionality is divided into three layers: Core, aggregation and edge. Model is illustrated in Figure 7.

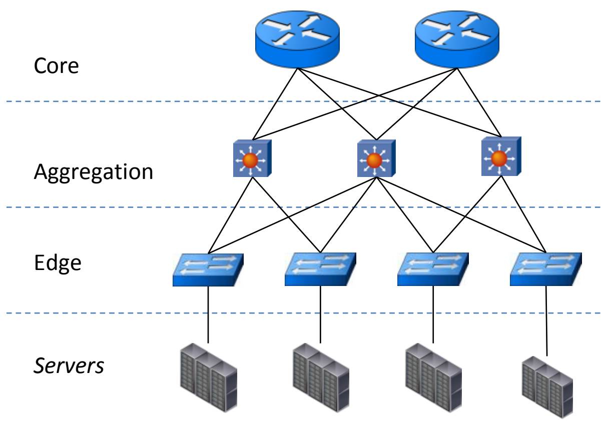


Figure 7: Basic three layered data center topology.

Core layer provides the high-speed packet switching backplane for all the flows going in and out of the data center. Core layer provides connectivity to multiple aggregation level modules and a resilient Layer 3 routed fabric with no single point of failure. Interior routing protocol is run in core layer. Core layer routers are normally redundant to guarantee efficient and reliable routing. (Cisco, 2007)

Aggregation layer has the primary responsibility of aggregating the thousands of sessions leaving and entering data center because the access layer uplinks are connected to it. Aggregation layer switches carry the workload of spanning tree processing in the network. Aggregation layer is typically the most critical layer in data center network because of port density, over-subscription values and CPU processing. Aggregation layer is also used to provide value-added services such as load balancing and firewalling. Aggregation layer is connected to the core layer with redundant links to guarantee reliability. (Cisco, 2007)

Edge or **access layer** is where the servers are physically attached to the data center's network fabric. Switches at this layer are usually either Top-of-Rack or End-of-Row type of switches located in server racks. Servers are usually connected at least to two physically different edge layer switches to guarantee the operations in case of link or hardware failure. Also the edge level switches are connected to at least two aggregation layer switches for the same reasons. In both cases, one of the links is acting as a master and the other one as a secondary. (Cisco, 2007)

3.2 Traffic characteristics

There is rather small amount of research done in the field of data center network traffic characteristics. Benson et al. (2009) conducted a study in 19 data centers that were different in sizes but all followed two or three tiered network architecture. Benson et al. studied the macroscopic and microscopic views of the data center traffic. They found out that roughly 60% of the core and edge links are actively being used, but the utilization is significantly higher in the core than in the aggregation and edge layers. When studying the link loss rates, they found out that despite the higher utilization rate in the core, the core links suffered the least rate of loss. Links near the edges suffered the greatest degree of loss. The microscopic view to the data center traffic revealed that the traffic is ON/OFF by nature.

Benson et al. (2010) studied more deeper the traffic characteristics of data centers. They came up with some key findings:

- A wide variety of applications were detected.
- Most flows in the data centers were small in size (less than 10 KB) and a significant fraction of flows lasted under a few hundred milliseconds.
- Despite the differences in size and usage of data centers, traffic was ON/OFF by nature with properties that fit heavy-tailed distributions.
- Link utilizations are rather low in all layers, but the core. A subset of core links often experience high utilization, the number of links varies over time, but never exceeds 25% of the core links in any data center.
- Losses occur within the data centers, however losses are not localized to links with persistently high utilization. Losses tend to occur at links with low average utilization, which implicates momentary spikes as the primary cause of losses. The magnitude of losses is greater at the aggregation layer than at the edge or the core layers.
- Link utilizations are subject to time-of-day and day-of-week effects across all data centers.

Benson et al. studied also the flow of the traffic generated by the servers. Study includes the traffic that stays within a rack versus the traffic that leaves the rack for either other rack or to external destination. They found that significant amount of traffic in commercial data centers is generated and stays inside a rack. Benson et al. explain the founding with the nature of services that are typically provided from commercial data centers: applications are more customer-facing, like web applications and file sharing services.

3.3 Services in the network

Network services can be divided in higher and lower level services. Higher level services can be considered to be such services that interact directly with the user. Lower level services are the enablers of the higher level services, which users use indirectly. Web browsing is an example of higher level service, that is enabled by lower level services which exchange data between user and servers. Lower level services tend to define the quality that higher level services experience. (Courcoubetis and Weber, 2003, p. 24)

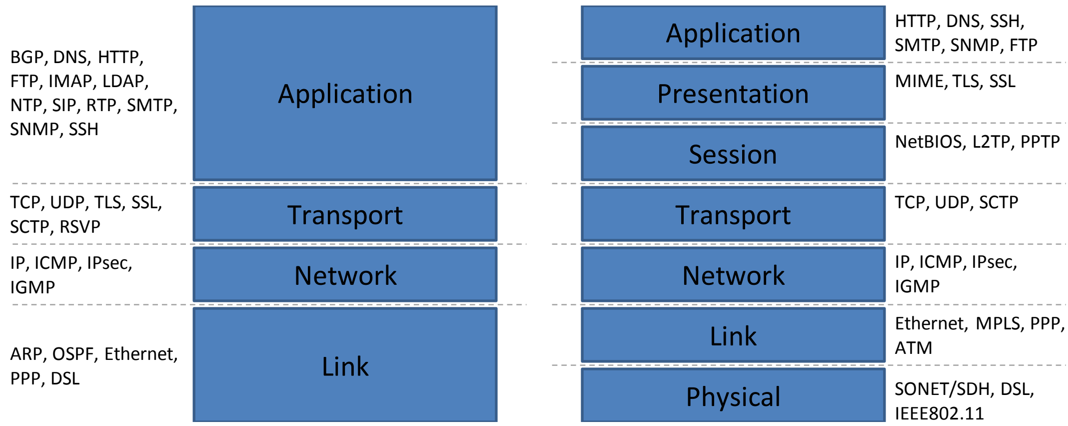
The Federal Communications Commission (FCC) gives recommendations of terms to be used in the context of network services in The Telecommunications Act of 1996. The term *information services* is tied to value-added services and *telecommunications services* for lower level transport services. The act defines telecommunications service as *"the offering of telecommunications for a fee directly to the public, or to such classes of users as to be effectively available to the public regardless of facilities used"*. Information service is defined as *"the offering of a capability for generating, acquiring, storing, transforming, processing, retrieving, utilizing or making available information via telecommunications"*. (FCC, 1996)

As the division between higher and lower level services has been made, different layering models have emerged. Layering models are used to define functionality and guide service development. Standardized layering models allow the services in different layers to be sold separately. When buying a certain level service, models assume that the lower level services provide the means automatically to implement the higher level service. Two commonly used layering models are presented in section 3.3.1.

3.3.1 TCP/IP and OSI model

Basically two models for computer networks exist: TCP/IP stack and Open Systems Interconnection (OSI) reference model. TCP/IP stack was developed in 1970s by United States Department of Defense and consists of four layers, which are defined in RFC1122 (Braden, 1989): Application, Transport, Network/Internet and Link layer. The TCP/IP stack is illustrated in Figure 8(a).

Application layer is where network applications and their application-layer protocols are located. Application layer protocols include such protocols as HTTP, SMTP and FTP. In addition, some network functions are also specified as a application-layer protocol, like DNS. Application-layer protocols are distributed over multiple end systems and the end systems exchange messages



(a) TCP/IP stack (b) OSI reference model

Figure 8: TCP/IP stack and OSI reference model. Adapted from Kurose and Ross (2008, p. 74)

with each other.

Transport layer is responsible of transporting the application-layer messages between application endpoints. TCP and UDP protocols are used in message transportation in different cases. TCP provides connection-oriented service, guarantees delivery of messages and flow control. TCP also provides congestion control mechanisms. UDP provides connectionless service, with no guaranteed reliability, flow control or congestion control.

Network/Internet layer takes care of moving network-layer packets from host to host. Transport layer segments are encapsulated inside network layer IP packets. The network layer also contains the routing protocols that determine the routes used between hosts.

Link layer is responsible of delivering datagrams between network nodes. Services provided by the link layer depend on link-layer protocol employed over the link. Link-layer protocols are for example Ethernet, WiFi and PPP (Point-to-Point Protocol). (Kurose and Ross, 2008, pp. 75-76)

Another model used in computer networking is OSI reference model, developed by ISO (International Organization for Standardization). OSI model is divided in seven layers and is illustrated in Figure 8(b). Also this reference model was developed in the 1970s, during the time when Internet protocols were still in their infancy. Largest differences in OSI model compared to TCP/IP stack are application, presentation and session layers. TCP/IP stack allows more freedom to application developers while OSI model can be seen to restrict and steer the development with these higher layers. Thus the protocols built according to OSI model have not gain a lot of popularity, OSI model is used more widely in teaching purposes. (Kurose and Ross, 2008, p. 77)

3.3.2 Connectivity services

At the most basic level, network provides services for transporting data between two points in the network. Data can be transported between two points as a *unicast* service, or from one point to many points as a *multicast* service. Chapter 3.3.1 discusses the two commonly used layering methods in network services. Connectivity services can be identified to reside on link and network layers in TCP/IP model (see Figure 8(a)). In OSI model, connectivity services are often referenced to reside on layers 1 to 3 (see Figure 8).

An important distinction need to be made between services that do and do not come with guarantees. Providing *guaranteed services* usually require some reservation of resources to ensure performance of the service according to certain parameters. For example, to guarantee a minimum transmission rate may require reservation of capacity on certain set of links. On the other hand, service might not make any guarantees for its performance, by not reserving any resources. In this case services are called *best-effort* services, as the performance of the service depends on the dimensioning of the network and other services that are competing of the same resources. (Courcoubetis and Weber, 2003, p. 32) Parameters that are used to guarantee the service are defined in service contracts. The role of service contracts is described in section 2.1.2.

3.3.3 Security services

Security services are provided to ensure reliable and secure communications between network hosts. Security services include for example firewalls and intrusion detection/prevention systems (IDS/IPS). These services typically reside on network layer and are based on inspecting network traffic and reporting all suspicious traffic to network administrators.

3.4 Measuring network traffic

Metering network traffic consists of selection of appropriate metrics, metering points and the measurement granularity and organizing measurement activities. Collection and analysis of network traffic is fundamental ability for the network operator to design and operate their networks. (CAIDA) Long-term aggregated statistics and short-term per flow statistics provide necessary insights to:

- Network provisioning
- Peering arrangements
- Per-customer accounting and SLA verification
- Per-peer accounting (traffic balance)
- Performance management
- Tracking topology and routing changes
- Other troubleshooting

All these items are valuable for the operator as the network is always a long-term investment and the provider should be interested in optimizing technical and economical performance of the network. For charging purposes, needs for measuring relate to per-customer accounting with long-term aggregated statistics.

3.4.1 Metrics

Metrics can be divided in to four main categories: utilization, performance, availability and stability metrics. These categories are described in more detail in Table 3. (Lambert, 1995)

Table 3: Network metric categorization.

Category	**Characteristics**	**Examples**
Utilization	Describes different aspects of the total traffic being forwarded through the network.	Packet and byte counts, peak metrics, protocol and application distribution
Performance	Relate to the quality of service issues such as delays and congestion situations.	RTT (at different layers), throughput, packet drop rate
Availability	Long term accessibility on different layers.	Line availability, route availability, application availability
Stability	Describe short-term fluctuations in the network. Changes in traffic patterns can be identified.	Line status transitions, route changes, next hop stability, short term ICMP behavior

For basic usage-based charging purposes, utilization metrics are the most relevant. Charging is often based on packet or byte counts or to peak rates. These metrics are also a useful tool in capacity planning. For customers, the most interesting metrics relate to availability and performance. They are basically interested in that their services are reachable and respond as defined in SLAs.

3.4.2 Measurements

Network measurements are crucial for network operator, but also other stakeholders are interested in the state of the network and services offered to them. Table 4 presents these stakeholders, their motivation and possibilities to measure their interests.

Traffic measurements in networks can be divided into active and passive measurements. In active measurements, test traffic is generated in to the network. Traffic can be created by real applications or it can be test-only traffic. Active measurements are suitable for testing the network, for example prior taking it in to production or to do troubleshooting.

Table 4: Motivation for network measurements for different stakeholders (Claffy and Monk, 1997).

	Goal	Measure
ISPs	- capacity planning - operations - value-added services - usage-based billing	- bandwidth utilization - packets per second - round trip time (RTT) - RTT variance - packet loss - reachability - circuit performance - routing diagnosis
Users	- monitor performance - plan upgrades - negotiate service contracts - optimize content delivery - usage policing	- bandwidth availability - response time - packet loss - reachability - connection rates - service qualities - host performance
Vendors	- improve design/configuration of equipment - implement real-time debugging/diagnosis of deployed hardware	- trace samples - log analysis

In passive measurements, existing traffic is used as a base for measurements. It provides insights to what is really happening in the network. Passive measuring provides information for the goals that both, provider and customer want to actively monitor. Figure 9 illustrates a common capturing-analysis work-flow. Work-flow can be divided into three parts: monitoring (marked in the figure with red), data collection (blue) and analysis (green).

The downside in passive measurements is the amount of generated measurement data. Depending on the environment, the amount can be huge. For example, when monitoring two 1 Gbit/s links, storing only IP and transport headers (40 bytes per packet) and assuming that the average packet size is 300 bytes. The amount of stored measurement data would be 30 MB/s. At this data rate, 1 TB database can hold less than nine and half hours of data.

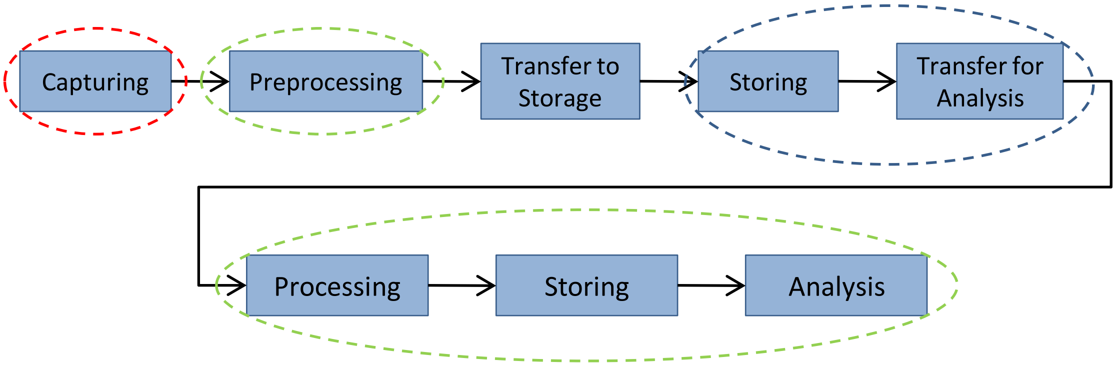


Figure 9: Flow of passive network measurement. (Järvinen, 2009)

Kandula et al. (2009) studied the nature of data center traffic by capturing the network related events from servers. They choose 1500 servers that represented a logical cluster in an operational data center. The measurement lasted two months, and roughly a petabyte of measurement data was gathered. Since this amount of data was gathered only from a fraction of data center's servers, the total amount of data would have been huge.

3.5 Energy consumption in data centers

Research company Gartner estimated that the share of carbon dioxide emissions generated by ICT-industry was approximately 2% in 2007. Although the share of emissions generated by ICT is about to increase in the future, Gartner predicts that efficient use of information services will remarkably decrease the total energy consumption of society. (VihreäICT, 2007)

Greg Schulz (2009, p. 20) claims that until recently, the energy efficiency of servers, networks, software and storage has been a little concern to IT organizations. Now this is changing as the price of electricity is increasing. Figure 10 illustrates electricity price development in Finland between 1995 and 2010 (Eurostat, 2011).

James Hamilton (2010b), vice president of Amazon web services, shows by calculations that in a modern high-density data center, networks constitute less than 10% of the total energy consumption. Koomey (2007) also claims that the total energy costs in data centers are roughly 10% of the total costs. Koomey also presents calculations that represent the total cost of ownership of a data center. Shares of energy consumption by different devices in data centers are illustrated in Figure 11(a). Figure 11(b) presents the actual electricity consumption during one year (note logarithmic y-axis).

As the calculations show, networks are in a minor role in energy consumption in data centers. Nevertheless, network equipment vendors and standardization organizations are driving development of greener network solutions. For example, in 2006

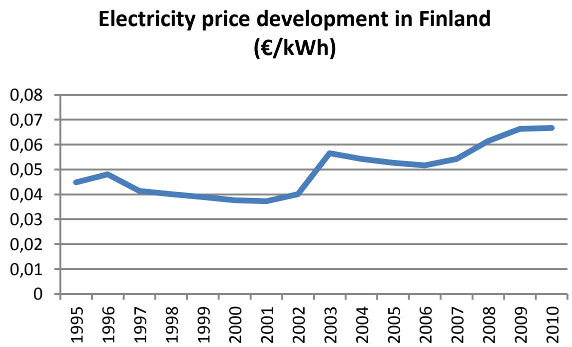
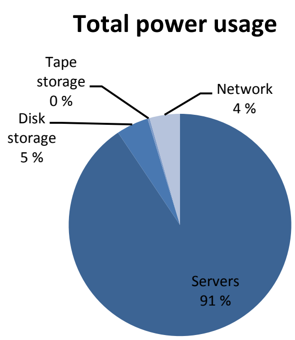
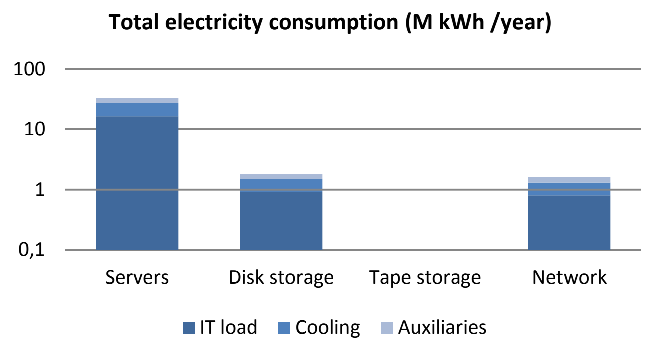


Figure 10: Electricity price development in Finland for medium size industries between 1995 and 2010. (Eurostat, 2011)



(a) Total power usage. (Koomey, 2007)



(b) Total electricity consumption. (Koomey, 2007)

Figure 11: Energy and electricity consumption in data centers.

IEEE started working with energy efficiency of Ethernet, which lead to standard IEEE P802.3az Energy Efficient Ethernet (EEE). (Christensen et al., 2010).

Jan Hof (2007) points out three different ways to reduce power consumption in networks, one of them being the design of network topologies. He highlights that introduction of more powerful network switches and technological breakthroughs such as 10 Gigabit Ethernet speeds and system resiliency allow to flatten the network topology. Instead of current three tier networks, aggregation layer can be taken out leaving only core and edge tiers. This will lead not only to capital savings in network devices, but also save the energy consumed by aggregation layer.

4 Cost allocation, pricing and charging

This chapter begins with discussion of costs that are caused by provisioning of network services and models that can be used to allocate these costs to services. After that, the chapter addresses possibilities of pricing the services. Last part of the chapter discusses charging models that are applicable in the context of network services.

4.1 Cost allocation

Identifying different costs that the provisioning of services causes, is the first step when developing charging activities. Even the identification of costs is not enough: costs have to be allocated in a fair manner to all services to be able to see the economical performance of services, and to be able to create a solid foundation for service pricing.

4.1.1 Cost classification

In the broadest end, division between Capital Expenditures (CAPEX) and Operational Expenditures (OPEX) is made. CAPEX contribute to the fixed infrastructure of the company which is depreciated over time. OPEX do not contribute to the infrastructure itself, and thus is not subject to depreciation. OPEX represent the costs to keep the infrastructure and company operational.

Although the classification between CAPEX and OPEX is necessary to allocate total costs effectively, they are very deeply interconnected. Trade-offs between CAPEX and OPEX can be made by technology and equipment choices. With higher acquisition costs (higher CAPEX), it is possible to acquire technologies and equipment that provide large amount of automated maintenance and provisioning tasks: higher level of automation enables cost savings in manual labor (OPEX). Network type has also effect on the cost structure. Backbone networks are rather easy to set up and maintain, but the acquisition cost per device is relatively high. Access networks on the other hand require more maintenance but the acquisition cost per device is lower. (Casier et al., 2006)

Casier et al. (2006) divide capital expenditures of a network operator in three categories. Classification is illustrated in Figure 12. Normally the largest share of capital expenses consist of network infrastructure acquisitions, which includes also infrastructure residing outside the data center, such as fibers and large IP routers. Only criteria is that the infrastructure is bought, not leased. Leasing costs from infrastructure (for example buildings and network equipment) are counted to operational expenses in this model. Second category of CAPEX is the network software that is needed in order to build and run the network. Network software includes network management systems and network planning tools. Third category is all the other infrastructure costs that cannot be linked in to the core business. This

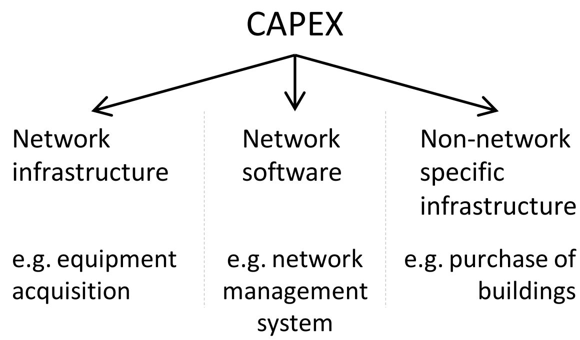


Figure 12: Capital expenditure classification in case of network operator. (Casier et al., 2006)

includes purchasing of land and buildings to house personnel, for example.

Casier et al. (2006) define a classification for operational expenses. This classification is presented in Figure 13. First the operational expenses are divided in three higher level categories. Each category is then divided in to subcategories that explain the model in more detail.

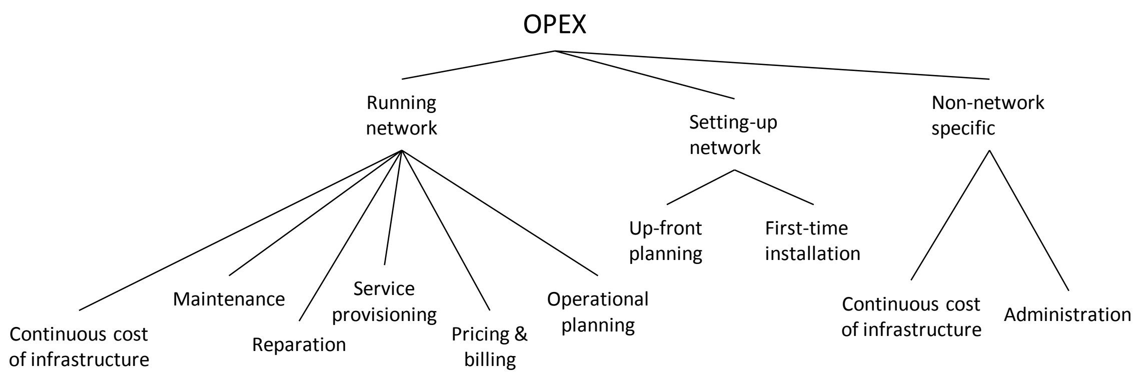


Figure 13: Operational expenditure classification in case of network operator. (Casier et al., 2006)

OPEX for network which is up and running consists of seven different costing bases, which are described below. *Continuous costs of infrastructure* are the costs to keep the network operational in failure free situations. These costs include costs for paying needed space, power and cooling energy and leasing network equipment (such as fiber rental). *Maintenance* includes costs to maintain or operate the network in case of a failure. The main actions here aim at monitoring the network and its services, such as direct and indirect polling of equipment and logging status information. Also stock management, software management, security management, change management and preventive replacement are included in maintenance costs. *Reparation* costs consist of actually repairing failures in the network, if the reparation cannot be performed in routine operation. Actions include diagnosis and analysis, technicians traveling to the place of failure, actual fixing the failure, and performing needed tests to verify the results. *Service provisioning* costs are given by the process of provisioning and service management. This means providing a previously defined and negotiated service to a customer. *Pricing and billing* includes such actions as collecting information on service usage per customer, calculating tariff per customer, sending bills, and checking payments. *Operational planning* costs are realized by day-to-day planning in existing network, such as re-optimization and planning upgrades. *Marketing* costs are caused by acquiring of new customers to a specific service of the operator. Actions include for example promoting new services and providing information about pricing.

OPEX associated with setting-up the network can be divided in to two subcategories. *Up-front planning* includes all the costs before the investment decision is made. For example, studies to plan the network, changing the network topology, and introducing a new technology or a new service platform are tasks considered under this category. Also the choice of an appropriate equipment vendor and costs related to vendor selection are included here. *First-time installation* is constituted from equipment installation and the operational part of the costs. This category includes connecting the devices in the network as well as necessary testing.

Non-network specific OPEX relates to costs that are not specially related to network company. *Non-network specific continuous cost of infrastructure* represents continuous costs that are not related to the network itself. Including buildings to house the personnel, energy for desktop PCs, heating and so on. *Non-network specific administration* costs include the administrative costs that every company has. These include such costs as payment administration for employees and human resources.

4.1.2 Mapping of CAPEX and OPEX

As the previous section shows, it is not enough to allocate just the network equipment costs to the services. Provider should aim to allocate all occurring costs to the offered services, and do that in a fair manner. Figure 14 illustrates the mapping of CAPEX and OPEX to direct, shared and common network costs. (Casier et al., 2006)

Some of the costs can directly be linked to a certain network service. The costs from provisioning the service and pricing and billing the service can be seen as direct costs. Direct costs can be explained so, that if the service would not be provided, then the costs would not occur at all. Common costs are costs that are not specific for the network service provider. These cost are often considered as overhead costs.

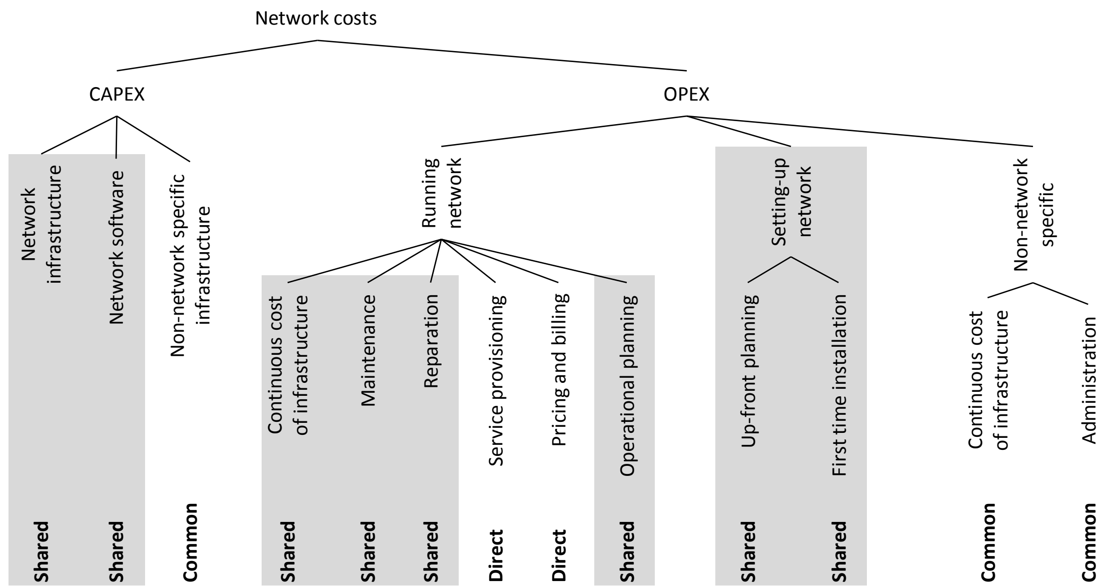
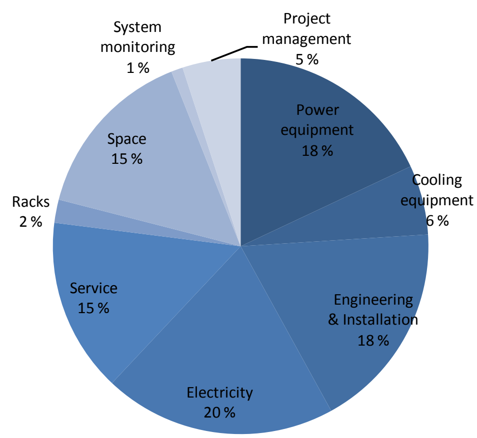


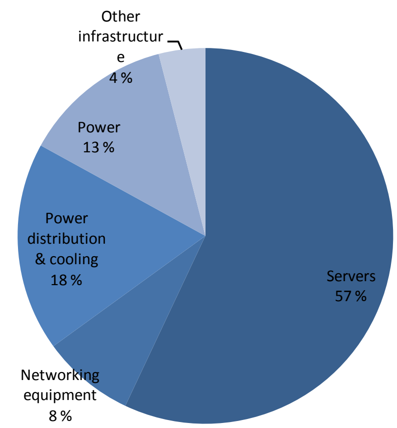
Figure 14: Mapping CAPEX and OPEX parts on shared costs. (Casier et al., 2006)

4.1.3 Data center cost structure

Hamilton (2010a) has studied the overall cost structure of a data center. He presents that the cost structure of a modern high-scale data center is dominated by the server costs. Second largest share of costs is constituted from power distribution and cooling. Hamilton's calculations are illustrated in Figure 15(b).



(a) Server rack. (APC, 2005)



(b) Whole data center. (Hamilton, 2010a)

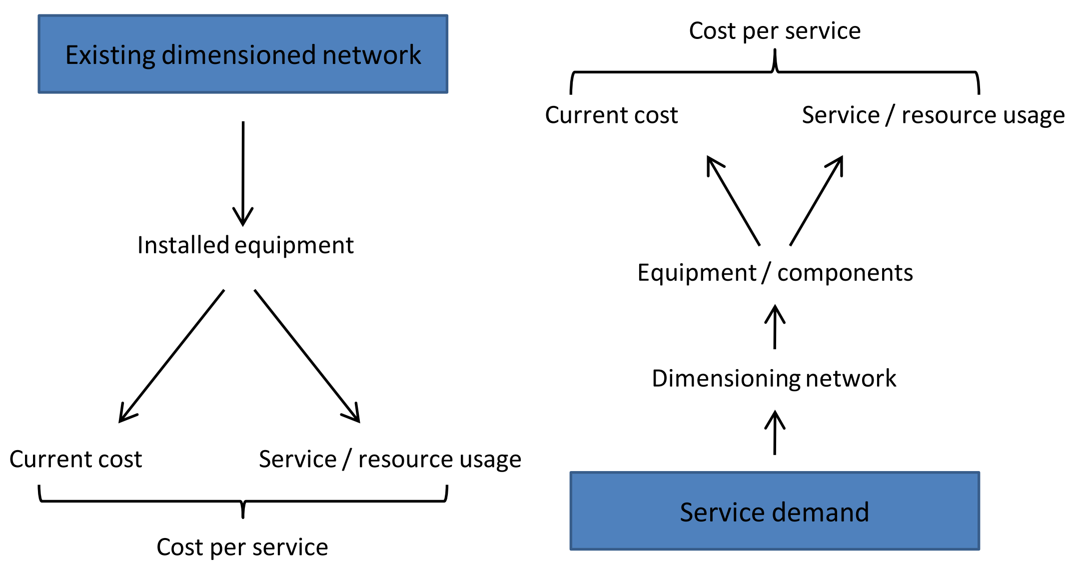
Figure 15: Cost structures for the data center environments.

Figure 15(a) illustrates cost structure of one server rack. (APC, 2005) This illustration also confirms that electricity, power and cooling equipment form almost 40% of the total costs. What this illustration does not show, is the share of servers versus network equipment costs. But it can be assumed, that the structure is relatively same on the rack level.

4.2 Top-down and bottom-up approaches

Shared network infrastructure helps to decreases operating costs of the network as a result of declining maintenance and repair costs and the standardization of equipment. But as the infrastructure evolves, based on the need of multiple customers and multiple services, infrastructure cost allocation becomes challenging. Two common principles in cost allocation can be identified according to many literature sources (see for example, Casier et al. (2006); Courcoubetis and Weber (2003); Kilkki and Pohjola (2006)): fairness and fullness. Fairness in this case can be considered in two ways: on one hand, costs need to be allocated fairly towards customers so that no customer is subsidizing any other. On the other hand, cost allocation needs to be done fairly between services, no service should be used to subsidize each other. Fullness in this context means that all costs and cost drivers need to be identified and allocated to gain realistic and holistic picture of the costs.

Casier et al. (2006) propose two approaches to cost allocation. These approaches are top-down and bottom-up and illustrated in Figure 16. According to Casier et al., the approach is dependent on the considered starting point of the network modeling process.



(a) Top-down (b) Bottom-up

Figure 16: Cost modeling approaches by Casier et al. (2006)

Top-down method (Figure 16(a)) starts from the existing network infrastructure. Network dimensioning is often a result from fluctuations in historic and current demand. New customers have been added to network, and traffic volumes for services have increased. At the same time, demand for some services has declined freeing resources to other services. This kind of network is usually less efficient than a network that is dimensioned specifically for the current needs. The costs of existing network equipment are allocated to the elements needed to deliver the service. Cost allocation is done through appropriate cost drivers, which requires accurate identification of real cost drivers. Identifying these cost drivers may be difficult in practice and it may lead to less efficient and less fair allocations.

Second approach to cost modeling is bottom-up method (Figure 16(b)). This method requires demand of services as a starting point. The model assumes that the network is dimensioned in such a way that it is optimal for the current demand.

Fully Distributed Cost approach

Fully Distributed Cost (FDC) approach is a top-down approach in which costs are attributed to services using existing cost accounting records. This makes it simple and auditable method, as the information can easily be checked for its accuracy. The idea of FDC is simply to divide the total occurring costs amongst the sold services. The computation is done by taking the actual costs of the various operating factors and dividing them to each service correspondingly. Basic principle of FDC is illustrated in Figure 17. (Courcoubetis and Weber, 2003, p. 181)

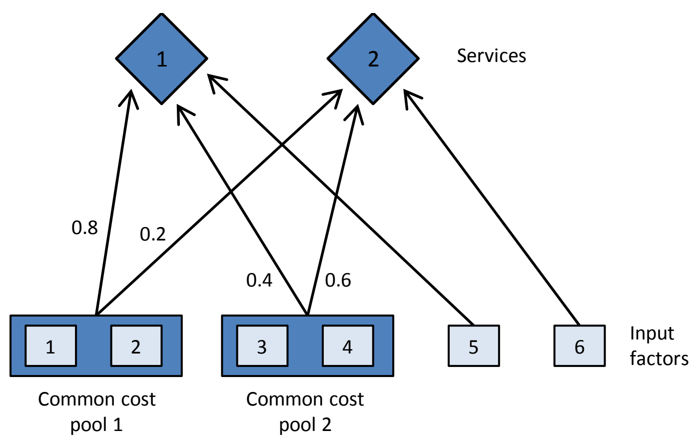


Figure 17: FDC approach. The costs of input factors are assigned to services according predefined coefficients. (Courcoubetis and Weber, 2003, p. 182)

All the costs of factors that are not uniquely identified with a single service are put into a number of common cost pools. Since only a small part of costs is directly attributable to a single service, most of the costs will be common. Next, coefficients to apportion the common costs among services need to be defined. These coefficients are always dependent on the particular cost pool. (Courcoubetis and Weber, 2003, p. 182) To express this in a formal way, suppose service i, that is produced in quantity y_i and has a variable cost $VC_i(y_i)$ that is directly attributable to that service. There is also a shared cost $SC(y)$, $y = (y_1, ..., y_n)$, that is attributable to all services (in this case, shared costs are assigned to a single cost pool for simplicity). Using this notation, the price for the quantity y_i of service i is defined to be its costs:

$$p_i(y_i) = VC_i(y_i) + \gamma_i SC(y), where \sum_i \gamma_i = 1 \quad (1)$$

After the selection of γ_i, the construction of prices is rather trivial, and can be done automatically based on accounting data. This way there is no need to build all from the scratch, as required by bottom-up models.

Due to simplicity and auditability of price construction process, FDC has been popular with network operators and regulators. However, there is also challenges that are related to FDC based pricing. There are no guarantees, that the prices are optimal or stable. This may be a result from the construction of the coefficients, if they are constructed without taking into account important information. Secondly, FDC based prices may hide potential inefficiencies of the network, such as excess capacity, out-of-date equipment, inefficient operation, bad routing and resource allocation. This is basically due to the fact that all the historical costs are taken into account. (Courcoubetis and Weber, 2003, pp. 190-191)

Activity-based Costing

Activity-based Costing (ABC) is also a top-down approach assigning the actual costs to services. ABC is based on a four-level hierarchy and is a refinement of the traditional FDC approach. The hierarchy of ABC is illustrated in Figure 18. Bottom level consists of input factors that are consumed by the operator, such as salaries, depreciation and cost of capital. Second level consists of activities that are required to operate the network and produce services. Third level consists of network elements, such as routers, switches and transmission links. The cost for each element is calculated by dividing input factors that are related to particular element, and the activities that are concerned with the operation and management of the network element. Last level consists of services that are sold to customers and make use of the network elements and service related activities. (Courcoubetis and Weber, 2003, p. 185)

In ABC, the definition of coefficients is also needed to apportion the costs of different input factors to activities and network elements, and the costs of activities and network elements to services. (Courcoubetis and Weber, 2003, p. 185)

ABC makes it easier to understand various occurring costs. It enables the analysis of costs and identification of activities that add value and that do not. Based on analysis, improvements can be implemented and benefits can

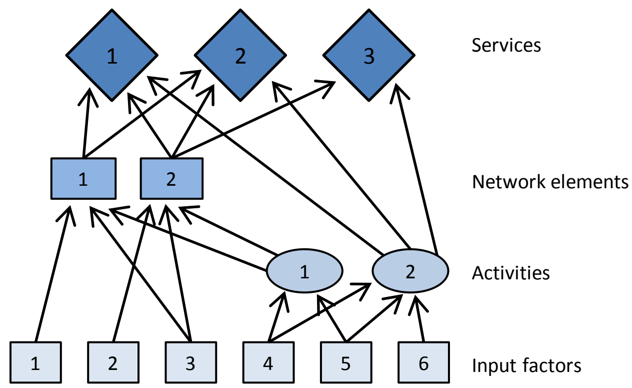


Figure 18: Activity-based Costing method. (Courcoubetis and Weber, 2003, p. 184)

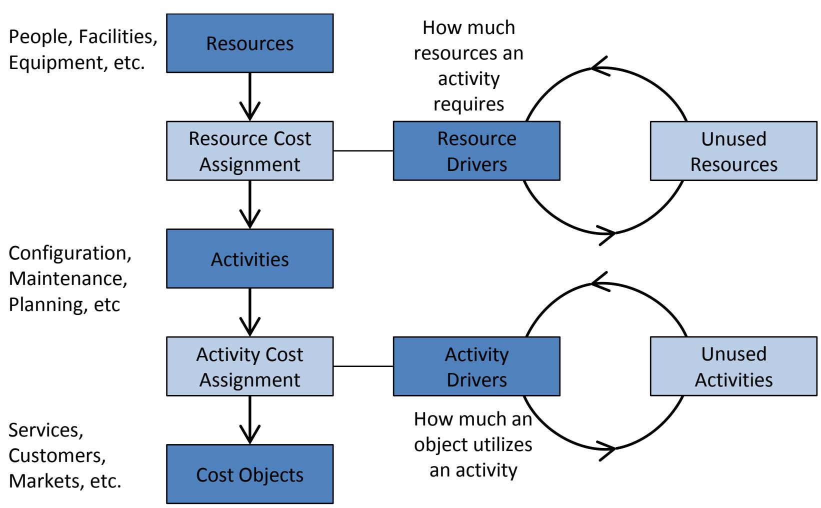


Figure 19: Continuous improvement process using ABC (12Manage).

be realized. This forms a continuous improvement process that helps to achieve higher overall efficiency by decreasing costs and reducing or eliminating non-value adding activities (see Figure 19). (12Manage)

Same problem that concerns FDC, exists also for ABC; it hides potential inefficiencies of the network. Even the network devices are underutilized, they might seem to be profitable, as the costs are fully shared to services. Thus there might not be incentive to improve the efficiency. This problem is common for top-down approaches, as the costs of existing facilities are allocated amongst sold services. Another problem that originate from top-down approach, is that it provides a one-way function to calculate prices from the input factors and coefficients. It does not provide the means to assess the situation, if a service was not produced. If this question want to be answered, one has to go backwards and check all the cost factors that contribute to the given service. (Courcoubetis and Weber, 2003, p. 186)

Long-Run Incremental Cost

LRIC (Long-Run Incremental Cost) is often used as bottom-up approach for cost modeling. Incremental costs are all costs that are associated only with the production of a specific service. In other words, LRICs are the costs that a producer would avoid, if a specific service (or some other increment) would not be produced. (Dewenter and Haucap, 2007)

The long run in this context assumes that all costs are considered variable. Thus the production can be optimized in the long run so that production of a given output takes place at minimum cost levels. Bottom-up models basically require the network to be designed from the scratch using the most cost-effective technologies available. After the design or modeling, the costs can be computed similarly as in activity-based approach by apportioning the cost of the network elements and by adding the cost of labor and the rest of the overheads as a simple markup on the cost of the infrastructure. (Courcoubetis and Weber, 2003)

Prices based on LRIC are subsidy-free and likely to prevail in an actual contestable markets. LRIC-based prices provide also signals about the economic efficiency. Although LRIC seems effective approach to construct the prices, it also has its downsides. It requires much from the accounting system that it is based on, and quite often the traditional accounting systems are not able to provide all the needed information. In this case, models need to be built from the scratch. This also leads to the challenge of auditing the LRIC-based prices.

4.3 Pricing

Pricing has been traditionally viewed as a method to cover the production costs and generate revenue and profits. Technology and market development has lead to new possibilities and needs in terms of service pricing. By service pricing, the provider can affect the way services are used and consumed. Pricing can also be used to provide input for the service provider about the current design of the network and its efficiency in terms of economics and technology.

According to Courcoubetis and Weber (2003), there are several practical requirements that must be met when charging for network services. Courcoubetis and Weber divide these requirements in three groups: 1) the end-user who pays the charges, 2) the service provider who defines the charges, and 3) the underlying technology that is used to produce the charge. Recipients of charges appreciate charges that are *predictable, transparent* and *auditable*. With predictable charges, the user knows in advance the total cost of using the service. This has lead to the popularity of flat-rate charges. Transparency in network services is challenging; the value of one kilobyte is basically unknown as the users consume it indirectly using higher-level services. Auditability walks hand in hand with transparency. Charges that are transparent are detailed in an itemized bill, and the origin can also be audited from the accounting systems.

Terminology used in this context consists basically of three terms. *Charge* is the total amount that is billed for using a certain service. *Price* refers to the amount of money that is associated with a unit of service and is used to compute the charge. *Tariff* indicates the general structure of prices and charges. An example of a tariff is $a + pT$, where a is a price for setting up the service, p is a price per minute for using the service, and T is the duration of the service usage in minutes. (Courcoubetis and Weber, 2003, p. 16)

4.3.1 Cost-based pricing

Cost-based pricing has historically been the most common pricing scheme. It is easy to understand and justify from the technical perspective. The goal of cost-based pricing is to create such prices that cover all the costs related to provisioning a certain service, but at the same time are fair and stable under potential competition. Figure 20 shows the basic principle of cost-based pricing method. (Nagle et al., 2010, p. 2)

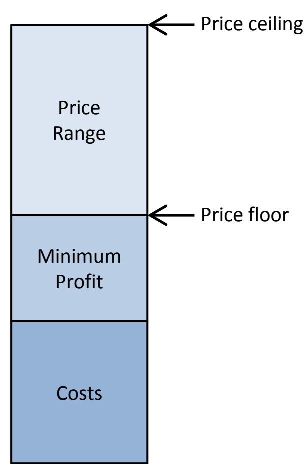


Figure 20: Basic idea of cost-based pricing.

Cost-based pricing tries to solve the need to ensure that prices guarantee profitability of company, or at least cover all the costs. The fundamental difficulty in defining cost-based prices is that services are usually produced jointly. To create stable cost-based prices, Courcoubetis and Weber (2003) state that prices need to be subsidy-free and sustainable. To satisfy these requirements, cost allocation and mapping to services are in a key role.

Cost-based prices can also provide efficiency information about service production. If the organization is able to decrease its production costs, price floor will decrease. On the other hand, if the production costs begin to increase, profit margin begins to decrease.

4.3.2 Value-based pricing

Traditionally customers have not considered connectivity and information security services as core services, but rather as enablers included in the end-user service. Lately, as the value-based thinking has emerged in other fields of connectivity and IT services, data center network providers have also got interested in the possibilities of value-based pricing mechanisms.

Value-based pricing tries to reflect the additional value obtained by customer by using certain service. Kilkki and Pohjola (2006) propose that for customers the essential motive to purchase a service, is their perceived value from the purchased service, not the cost that occurs to provider to produce the service. In value-based pricing, the goal of the service provider is to maximize the value created to the customer and then to use appropriate pricing mechanisms to capture as much profit from the created value.

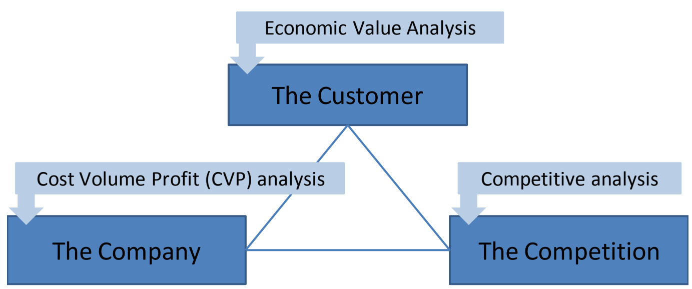


Figure 21: A framework for value-based pricing, adapted from Hinterhuber (2004).

As Figure 21 shows, a holistic view from the whole economic environment is needed to successfully establish value-based pricing. Economic value analysis should help to understand sources of economic value tied to a service to different customer clusters. Service provider needs to capture the company-internal perspective. Understanding the implications of price and volume changes on profitability is needed. Competitive analysis inside the industry is used to provide the understanding of trends in competitive pricing, product offerings, and strategies.

The concept of economic or customer value is not easy to define. Simpson et al. (2001) and Zeithaml (1988) define the customer value to be the difference between perceived benefits and sacrifices that are caused by used product or service. In other words, customer value is seen as the difference between the customer's willingness to pay and the actual price paid. This difference can be defined to be *consumer surplus*. Consumer surplus can also be defined to be the net benefit of using a certain service. Another way to define the value is a broader definition. Nagle et al. (2010) define customer value to be the maximum amount that the customer would pay to obtain a given product or service. This value is the price that would leave the customer indifferent between the purchase and foregoing the purchase. Figure 22 illustrates one possible scenario of user-perceived value from using certain service.

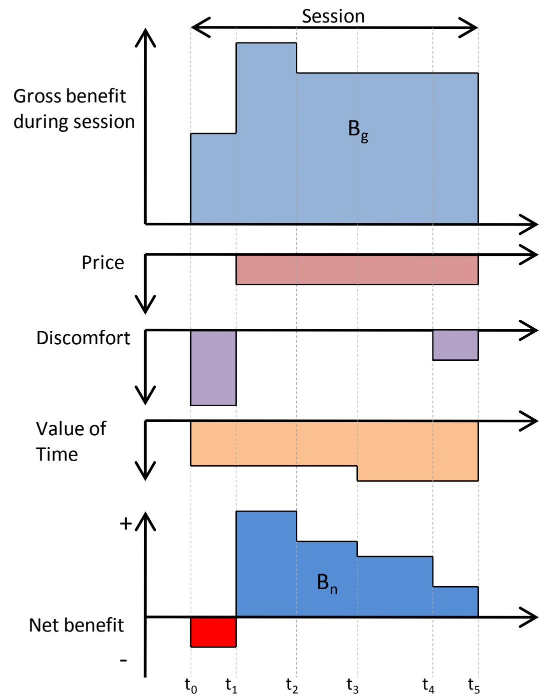


Figure 22: Customer perceived value when using service.

4.4 Charging

The purpose of charging is to generate revenue for the service provider. Charging can also be used to steer the behavior of customers and thus affect the demand of services. This chapter introduces two most common charging schemes that are used in communication services: flat rate and usage-based charging.

4.4.1 Flat rate charging

In flat rate charging, total charge for customer for a service contract is fixed at the time the contract is made. That is, the charge is determined *a priori*, even though the actual costs of the contract to the service provider could be known at a later time, for example after the contract has expired. (Courcoubetis and Weber, 2003, p. 191) Usually flat rate charges tend to reflect the average costs of the past.

According to Courcoubetis and Weber (2003) there are two advantages that have led to the wide use of flat fees. First, a flat fee is easy to implement and second, customers often prefer the predictability of flat fee. But there is also serious drawbacks in flat rate charging. According to Courcoubetis and Weber (2003) and Kilkki and Pohjola (2006) flat rate tends to produce high social cost because of the waste of resources. Figure 23 illustrates the social waste caused by flat rate charging. Assuming that each customer has its own demand function that is linear and starts from the same point on the price axis and ends on the monthly volume axis. On the left, the customer is charged a price $p = MC$ (Marginal Cost) and consumes x^*. In this case, the social welfare (SW) is the area A. The situation changes if the customer is charged a flat rate ($p = 0$). Now the customer has no incentive to reduce the consumption, and consumes x_{flat}. This makes the social welfare $SW = A$ – W, where W is the social waste. If the demand is greater and the user consumes x'_{flat}, the social waste W' grows even larger.

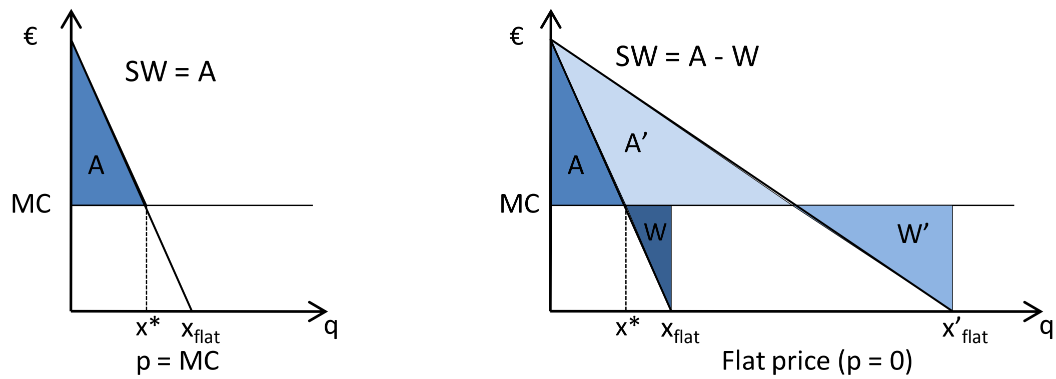


Figure 23: Social welfare and waste under flat rate charging (Courcoubetis and Weber, 2003, p. 192).

Flat rate also tend to be unstable under competition, because the light users are easily used to subsidize the heavy users. Figure 24 illustrates the dilemma of cross subsidization with a flat rate pricing. A flat fee is computed to cover the costs from providing the services. Flat fee is computed using the average consumption of services $x_{flat}(average)$ but regardless of the consumption, the same fee is charged from all customers. In this case, customers that consume less than the average amount are having negative net benefit. Negative net benefit means that the customer is experiencing the service unprofitable, because their utility for using the service is less than what they must pay.

Courcoubetis and Weber (2003, p. 193) propose a solution to reduce the bad effects of flat rate charging, by restricting the range of allowed resource usage, v. This will lead to fairer charges of individual customers. Contract constraints need to be enforced, so the service provider must police the service usage of users. If the maximum allowed resource usage is exceeded, some measures must be taken to keep the pricing fair. Measures can be for example, extra charging, traffic dropping or blocking and bandwidth restriction. This motivates the customer to predict their service usage as accurately as possible and to assign the cheapest contract that accommodates the needs. This scheme suits well in cases, when the customer has a low variance in service usage, and hence predictable resource usage. This leads to reduction in social waste and cross-subsidization between heavy and light users.

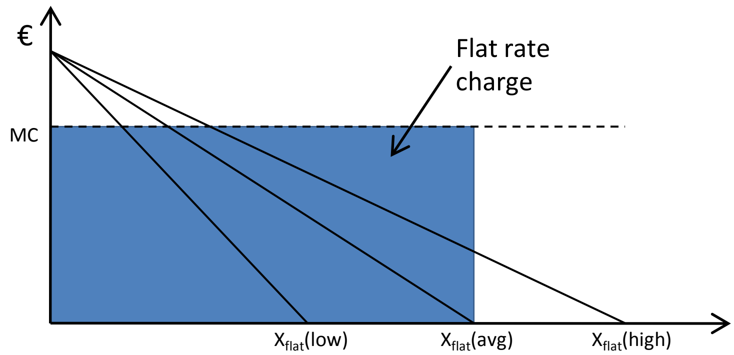


Figure 24: Cross subsidization dilemma with a flat rate charging (Courcoubetis and Weber, 2003, p. 193).

4.4.2 Usage-based charging

Usage-based charging relies on metrics that can be explicitly measured and accounted. Usage-based charges provide customers more control over their choices of how to use services and better value-for-money.

Usage-based charges typically constitute of two-part tariffs, in which the fixed part is proportional to the capacity available for use. Variable part represents the usage fee to cover the marginal costs (see Figure 25). Relationship to capacity makes usage-based charging appropriate for industries that have high capacity costs due to high capacity peaks. (Dewenter and Haucap, 2007)

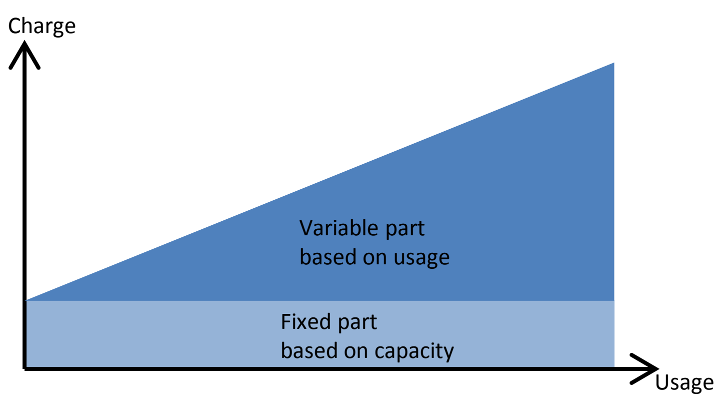


Figure 25: Basic idea of usage-based charging.

Usage-based charging in network services has its downsides. Network users may have challenges predicting their accurate traffic volumes, which results from interaction with services. As customers appreciate predictable and stable bills, the nature of network usage may lead to surprises. Same applies for service providers: they want to have predictable and stable incomes. (Courcoubetis and Weber, 2003) On the other hand, usage-based charging ensures fairness for customers, if they are able to monitor and control their usage.

Usage-based charging in a complex network environment sets high requirements for metering and accounting. Systems performing such actions are often very heavy and complex, and thus very expensive. In case of a single link or a small set of links, the collection of data is relatively easy. In case of data center environment, where network traffic traverses between many servers and services, collecting, aggregating and accounting the data is very challenging.

5 Case study: Network services and charging in data center environment

This chapter covers the case study conducted in the case organization. First, background information about the organization and provided services are presented. Second part of this chapter describes study methods and how the study was performed. Chapter is concluded with the results from the case study.

5.1 Case organization

Tieto Corporation started its operations in 1968 under the name of Tietotehdas Oy. In the early days, the company produced data center services for its owners: mainly for Union Bank of Finland and a few forest companies. During 1970s the company started growing as midi-computers were introduced alongside with the mainframes. In the 1980s personal computers became common and besides the mainframe computer services and software, the development of IT systems became a central operations for Tietotehdas. During 1990s the company experienced a rapid growth through a number of acquisitions, mergers and strategic alliances. The name of the company was changed to TT Tieto in 1995 and to Tieto in 1998. In 1999 Finnish Tieto and Swedish Enator merged to form TietoEnator. During the 2000s, the IT industry became global and the growth was huge. In 2009 TietoEnator changed its name back to Tieto. (Tieto Oyj, 2011b)

Today Tieto is an IT service company employing nearly 18,000 employees. Tieto's main markets are in Nothern Europe, Russia and Poland. Tieto also serves its customers globally at certain areas of expertise and has industry-specific activities in selected countries. Tieto had net sales of 1,713.7 million euros in 2010 with operating profit of 72.4 million euros. (Tieto Oyj, 2011a)

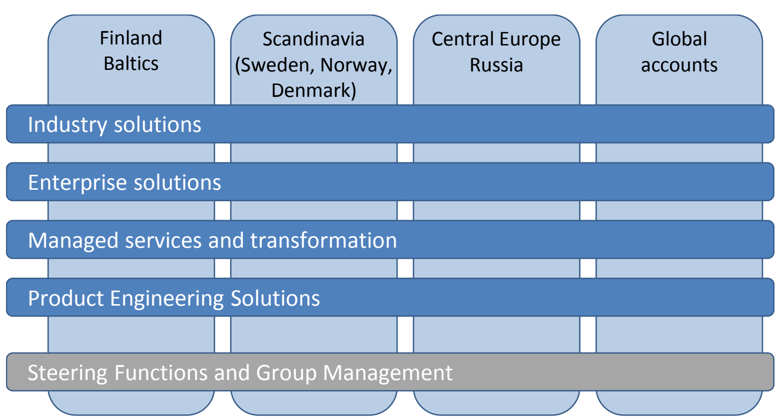


Figure 26: Tieto's operating model.

Figure 26 illustrates the current operating model of Tieto. The model is divided in to vertical market units and horizontal business lines. Market units drive the sales and develop customer relationships in selected market areas. Business lines are responsible for offering development, sales support and project and service deliveries to customers.

Connectivity and information security services are a part of the infrastructure services portfolio. Purpose of these services is to ensure reliable and secure communication to enterprise business systems and data, while ensuring that business operations are supported by agile and adaptable enterprise network architecture. The positioning of connectivity and information security services in the service portfolio is illustrated in Figure 27, and the connectivity services portfolio is described in Table 5.

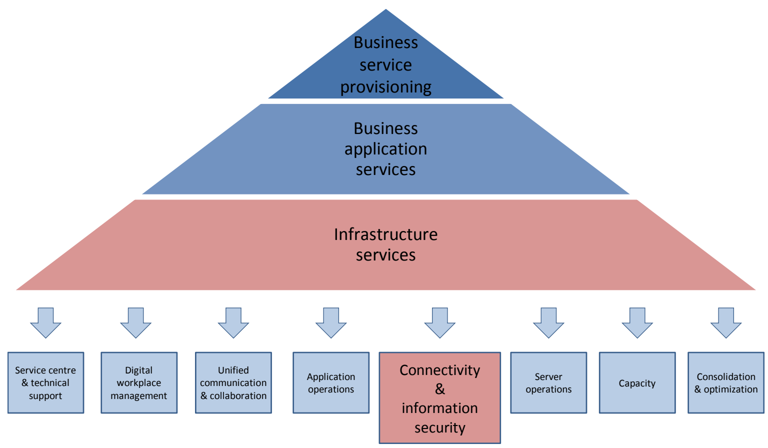


Figure 27: Connectivity and information security services position in service portfolio.

5.2 Methods

Study is conducted using two main research methods: evidence reviews and interviews. Evidence reviews are used to gather information about network infrastructure and design, and current cost allocation and pricing practicalities. Internal documents and records are explored to form a picture about these areas.

Interviews are chosen as a second main research method to discover opinions and insights of personnel regarding the context of the study. Interviews are performed as a *in-depth interviews* (Yin, 2009). No strict interview sessions will be held, but the interviewees are more like informants and their opinions are gathered during multiple discussions. Table 6 lists the people that are interviewed during the case study. Key questions for interviewees are listed in Appendix A.

In addition to evidence reviews and interviews, daily working methods, approaches, opinions and processes will observed and the observations will be used to complement

Table 5: Service portfolio of connectivity services.

Service	Description
Network Backbone Management	Consists of the management, maintenance and problem-solving related to the corporate network backbone infrastructure, which includes network backbone and internal connections. The service is based on planning, provisioning, monitoring and management of various network active components (typically routers, backbone switches and LAN-to-LAN VPN routers).
Partner and Customer Connections	Covers the authorized customer, business partner, and other external party connections to the corporate data communication network. This service is based on management and monitoring of gateways and firewall systems. Typically the connections are one of the following: VPN or other secure extranet type of connection from the Internet, leased line, or the operator's TCP/IP service.
Internet Connections	Includes access to the Internet from the corporate network and vice versa. The service is based on firewalls and proxy/cache systems with an access rules database.
Remote Client Terminal Connections	Offer a remote client connections from the Internet to the corporate communications network. Service is based on remote access points and strong user authentication. Typical remote access points are VPN access servers supporting VPN clients, SSL VPN or RAS servers.
DNS Management	Offers both internal and external DNS services for customers.
Data Center LAN Management	Service includes single- or dual-site data center LANs with server connections. Service delivery is based on the use of LAN switches and other equipment (such as LAN routers), and the related management and monitoring. With the Data Center LAN Management service, the supplier creates and operates an optimal data center LAN architecture, one that is able to handle large traffic volumes and provide high-speed and high-availability server-to-LAN connections.

the results of two main research methods.

5.3 Current state and largest challenges

This section presents the results that were gathered in the study. First, the network architecture and design principles are described based on the explored evidence. Then, currently provided services and their charging principles are listed. Last part of the results brings forth the points that were found in the interviews and

Table 6: List of interviewees in the case organization.

Position	Role in organization
Director	Director of network services in the global data center services business unit.
Manager	Manager of connectivity and information security services in Finland.
Manager	Manager of shared network infrastructure team in Finland.
Infrastructure architect	Lead technical architect of shared network infrastructure.
Technical specialist	Responsible of tool development for network management.

observations.

5.3.1 Data center network

Two possible data center network choices are provided for customers: depending on needs, either shared or dedicated network infrastructure is be used. Customer dedicated network is build for specific customer purposes either in organization's own data centers or to customer's premises. All the devices and connections are dedicated solely for one customer allowing custom tailored SLAs and maintenance windows. Network management in case of dedicated environment is performed as requested by the customer. For example, all the management work can be performed from Finland and only by authorized personnel. Customers can also have access or visibility to devices and configurations. As dedicated networks are always built according to customer needs, they are not discussed here.

Following chapter discusses the shared network infrastructure in more detail. There is no clear rule of thumb, when customer selects dedicated or shared network solution. In practice, most customer implementations are combinations of both.

Shared network infrastructure

In shared environment, network resources are shared among multiple customers. Network is dimensioned using the principle, that there is always enough capacity to fill the needs of customers, and thus resembles the *best-effort*-principle. Only tested and proofed-to-work network equipment and software are selected to guarantee reliable service provisioning.

Shared network infrastructure is operated by network specialists of the organization locally and in offshore locations. Unlike in dedicated infrastructures, customers are not allowed to have any access to shared infrastructure itself. These principles aim to high reliability, availability and cost efficiency.

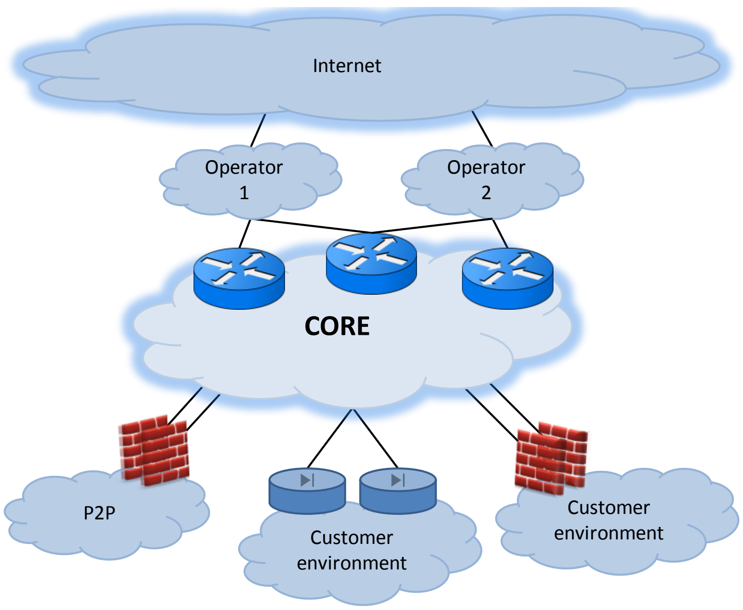


Figure 28: General routing model used in the case network.

The routing model of the network is illustrated in Figure 28. Network is divided into logically separated blocks, that are illustrated as clouds in the figure. Each block is connected to core, that acts as a default gateway. For the core, the Internet acts as default gateway. Each logical block is secured with redundant firewalls that allow all traffic towards the core. P2P -cloud in the figure represents a network block for dedicated customer or partner connections. Otherwise connections are routed in the network from the Internet and through firewalls to customer environments.

Case environment consists of multiple data centers, that are located apart from each other. Data centers are interconnected with redundant DWDM (Dense Wavelength Division Multiplexing) ring. DWDM-ring provides high availability: there is no single point of failure as signal is always sent to physically separate devices that reside in different locations. All components in the system are redundant.

Connectivity in data centers is guaranteed with layer 2 switch network that covers all sites. Layer 2 network allows a great deal of flexibility as servers can be freely deployed. On the other hand, this generates more traffic inside and between the data centers.

5.3.2 Cost structure of the case organization

The overall cost structure of the network unit is illustrated in Figure 29. All costs that occur are considered to be fixed, as there are no real variable components. Largest amount of costs are generated by investments in new equipment, personnel expenses and operating expenses.

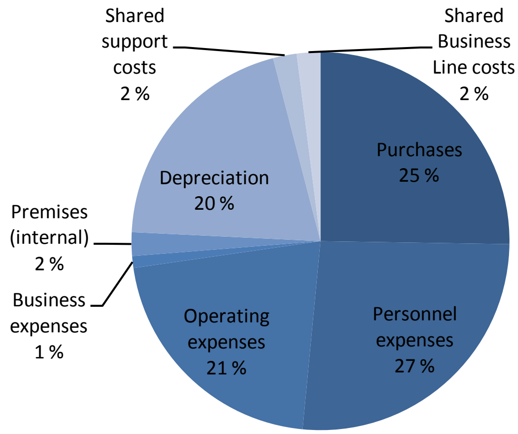


Figure 29: Cost structure of network unit in the case organization.

As the figure shows, premise costs generate only 2% share of the total costs. This cost pool includes all the costs that are related to housing of network equipment in the data centers. These costs include floorspace, cooling and power distribution, and electricity.

5.3.3 Typical combination of services

Figure 30 illustrates a typical combination of network services that are provided to customer when its IT services are provisioned from the data centers of the organization. Services can roughly be divided into two parts: *external connectivity* services enable the connectivity to the data center core, and *internal services* provide connectivity inside the data center.

External connectivity

External connectivity services can be described to enable the connectivity to the edge of the data center. Customer connections are routed to the core either using the Internet connection or customer dedicated lines. Internet capacity is either divided in to 2 MB increments or to shared capacity, which are provided to customers according to their needs. Internet capacity is charged according to dedicated increments or with a flat-rate, if a shared capacity is used. If customer dedicated lines are used, all associated costs are charged from the customer.

When connections are routed in to the network, they go through Intrusion Detection/Prevention (IDP) sensors, which investigate the traffic in case of anomalies. IDP service is charged with a basic monthly fee according to the amount of sensors that are utilized by the customer.

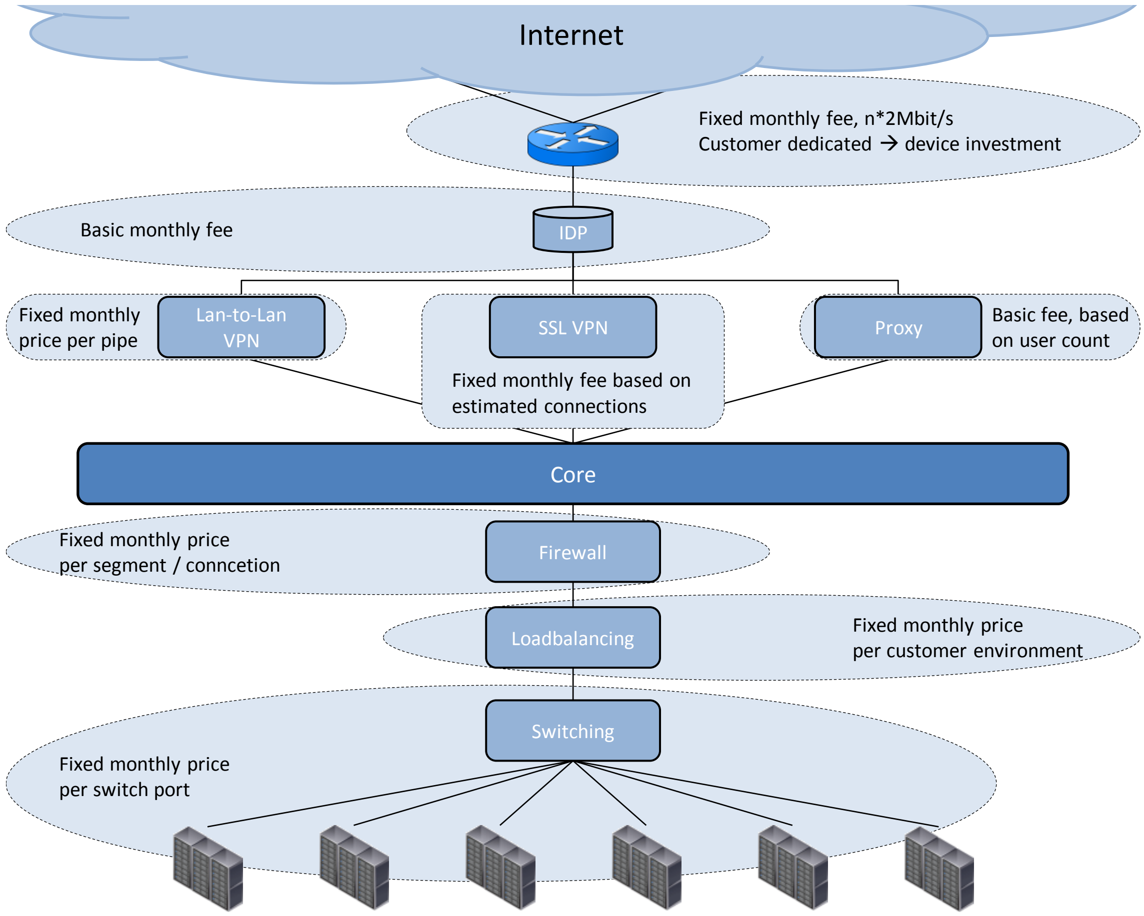


Figure 30: Simplified presentation of currently offered services and their charging principles.

Lan-to-Lan VPN service is provided according to customer needs. LAN-to-LAN VPN connections are static by their nature, and always configured between two sites. This service is charged with a monthly fee per site.

SSL VPN service is used to authorize user access in to the network. Service is charged using a two part charge: basic monthly fee for authentication profile, plus a fee according to the estimated amount of simultaneous connections due to the need of licenses.

Proxy services are used as cache between services and customers. This service is charged based on the user amount: basic monthly fee includes 500 users, and the number of users can be increased in increments of 500.

Internal services

Internal connectivity services can be considered to enable the internal connectivity inside the data centers. Internal connectivity consists of switching, load-balancing and firewalls.

Switching is used to provide the connectivity for servers in data centers.

Switching is done in layer 2 and it follows the basic principle of 3-tiered infrastructure. The costs of switching infrastructure is retrieved by charging the used ports at the edge tier. Fixed monthly price per port is based on the investments and monthly operating costs.

Load-balancing is used to provide redundancy and scalability to the network. This service is used when a multi-site solution is built. Load-balancing service is charged per customer environment with a fixed monthly price.

Firewall service is used to provide security for the customer environments. Every network block is secured with a firewall and the firewall service is charged based on the environment size behind the firewall. Basically this is done by calculating firewall segments and number of firewall rules.

5.3.4 Largest challenges

This section describes the largest challenges that were identified according to the interviews and observations. Summary of the largest challenges is presented in Table 7.

Table 7: Largest identified challenges to overcome in order to develop charging activities.

Area	**Challenges**
Network design	- Lack of standardization - Too general service and offering descriptions - Packaging of service offerings
Network management	- Fragmented management tools - Large share of manual work
Accounting	- The share of manual work - Time reporting accuracy - Auditability
Cost follow-up	- Mapping performed work to services - Costs are not followed on a level of single service
Pricing and charging	- Relationship between prices and costs is not clear - Charging practices are vague

Network design and management

Strategy of the corporation is to move focus on standardized and packaged service offerings, but yet there is no standardized network design used as a baseline for new implementations. There is too much variation in the solutions, which leads to challenges in network management. Since many device and software versions are in use, network management becomes complex. Complexity causes possible pitfalls what becomes to service quality. Another downside is the increase in operational costs. These challenges are partially implications of too general service and offering descriptions. Descriptions do not make enough comments or recommendations on how the services should be produced. This causes a communication gap between sales representatives and network architects, and leads to tailoring in the selling phase.

Situation in the area of network management tools is very fragmented. There is no single integrating solution, that allows easy and simple view of the network and its status. Network specialist have to do large amounts of manual work to find out all dependencies of one device, which makes fault investigation very challenging. Partially this is caused by the lack of centralized configuration management.

Observations and comments from the network specialists show frustration about the current situation. Workload on certain areas has become too large, which eats resources from proactive process development. This leads to shortcomings in processes while networks are constantly becoming more complex. Since operational processes do not cover the whole service management and challenges, the result can be seen as increased operating costs.

As mentioned in chapter 3.5, energy consumption of network devices and its follow-up has gained more attention in recent years. Greenness and energy efficiency have been one of the key design principles also for Tieto. Latest great example about the success in this area was the Green Enterprise IT award that Tieto received from the Uptime Institute in 2011. (UptimeInstitute, 2011) Although the overall energy efficiency is at decent level, the energy consumption of network equipment is not followed separately at the moment. This is a clear area of improvement. Development for the future has actually been started, but there is not yet reliable information available on the actual energy consumption, only iterated estimates.

Accounting and cost follow-up

Accounting of services and configurations relies on manual work using spreadsheets. This requires a lot of manual work, since after each configuration changes need to be updated to the spreadsheets. The amount of manual work and large spreadsheets make this accounting style very vulnerable to human errors. At the moment, performed work is very difficult to trace back to certain service based on the accounting information.

Time reporting is based on filing used time to projects and performed tasks. Projects are divided by customer or continuous service and tasks are defined to represent performed work. Time is not reported based on services, which skews the profitability and efficiency of services. Although the task structure is defined to represent performed work, it is not uniform across the projects. Thus it does not allow simple auditing of possible defects in efficiency of service provisioning. One important requirement for accounting system is auditability. Due to the auditability issues, the cost follow-up has become challenging: occurring costs and work are hard to map directly to services, which makes the profitability analysis of services challenging.

Pricing and charging

Current service prices have evolved during the time, and are not clear nor unambiguous. Prices are constructed to cover the overall costs. Nevertheless the relationship between costs and prices has faded, and the prices do not directly reflect costs related to certain service. Prices constructed this way are not sustainable, nor likely to last in the long rung competition.

Pricing of services should also be used to steer customers towards standardized services. Customization always increases service provisioning costs, which can be addressed by using service-related costs as a base for prices. Currently there is no noticeable difference in prices of standardized and customized solutions.

Charging principles of services differ from each other remarkably. It is not clearly stated, where the services are and should be charged from. Service units also differ remarkably from each other, which requires a lot of effort when determining monthly charges for different services and for the complete service offering.

5.4 Market benchmarking

To form a better view of the best practices of charging the network services, market benchmarking is performed. Information is gathered via public material such as websites and offering catalogs. For the benchmarking, services of providers listed in Table 8 are examined.

Table 8: Benchmarked service providers.

Provider	**Type**
SAVVIS	Global managed hosting and colocation service provider.
ServerChoice	UK-based web-hosting and managed service provider.
Hosting.com	North American hosting and managed service provider.
Rackbase	German-based colocation and data center service provider.
Rackspace	Global cloud and managed services provider.
Amazon EC2	Global cloud service provider.
Windows Azure	Global cloud service provider.

There are similarities how providers of same type provide the services. Cloud service providers were chosen to benchmarking to be able to compare the future models for current practicalities. Table 9 presents the results of benchmarking by comparing the services of the case organization to other providers.

Table 9: Market benchmarking results.

Service	**Traditional providers**	**Cloud providers**
Data Center LAN	Included in server price.	Included in computing instances.
Firewall management	Dedicated firewall or virtual firewall per environment.	N/A
Load balancing	Per connected server.	Load balancing time and data volumes.
Backbone network	Reserved bandwidth, monthly flat rate.	Data volumes between data centers.
Internet connections	Reserved bandwidth, monthly flat rate.	Data volume OUT of the data centers.
Remote connections	Per remote user.	N/A
Cross site connections	Per managed site.	N/A
Proxy	Based on the number of users.	N/A
Remote LANs	Per managed LAN device.	N/A
WAN connections	Per managed site.	N/A

This information shows how the services are charged from the customers. Some of the services are provided internally to other units of the company (see chapter 2.1.1). For example, data center network is basically always included in prices of servers. Internal pricing components and charging information are usually based on the production costs, and thus organizations are not willing to share them publicly.

6 Analysis

This chapter analyses the challenges that were identified. Interviews and observations in the organization revealed important issues that need to be addressed in order to be able to improve the charging process. Following sections discuss these challenges, their implications and compare possible solutions.

6.1 Network service standardization and management

Messages in the service descriptions and what came across in the interviews and observation are somewhat conflicting. Service descriptions of the shared network infrastructure propose that only tested and proofed-to-work equipment and software versions and configurations are used. In reality, the commonly agreed standard ways of working seem to be vague. If the future strategy is to provide standardized network solutions and keep the shared infrastructure manageable, scalable and reliable, there is a definite need to improve common ways of working.

Service offerings to match demand were discussed in section 2.3.2. Currently there is no clear division between core and supporting services. Services are packaged without clear structure, and each case requires thorough analysis. To create more structured service packages to match demand of different network environments, business activity and user profile catalogs (section 7.6) and clearer division of core and supporting services are proposed (section 7.5).

6.2 Accounting and cost follow-up

Section 4.2 discusses different methods of allocating costs to services. Table 10 presents a SWOT (Strengths, Weaknesses, Opportunities and Threats) analysis of these different modeling approaches and their suitability for the needs of the case organization. SWOT analysis (Johnson et al., 2005, pp. 148-150) is a simple but effective tool to summarize and compare the key issues related to analyzed methods and their relations to business environment.

Bottom-up approaches have their benefits as costing methods, but in this particular case they are not very suitable. To successfully implement a bottom-up costing method, such as LRIC, would require large modifications to current accounting systems and network design. LRIC-based prices may also be challenging to audit, as they are based on long-run increments considered as variable costs. In ABC, there is no need to consider variable costs, all occurring costs can be handled as fixed.

Table 10: SWOT-analysis of suitability of FDC, ABC and LRIC to this case.

	FDC	ABC	LRIC
Strengths	- Simple and easy to implement - Represents actual costs - Auditable	- Maps activities and network elements to services - Represents actual costs - Auditable	- Sustainable, subsidy-free prices
Weaknesses	- Prices may not be optimal or stable - May hide potential inefficiencies of network - Doesn't reflect the amount of labor used to each network component	- May hide potential inefficiencies of network	- Auditing - Requires large modifications to accounting system - Historical fluctuations in network design
Opportunities	- Requires very little changes to accounting system	- Requires very little changes to accounting system - Better follow-up of cost/labor -ratio - Allows continuous improvement based on the continuous analysis	- Provide continuous feedback of service efficiency as long as increments are defined correctly
Threats	- Construction of coefficients may skew prices	- Granularity of activities is too low/high	- Defining the increments to be used - Assumes minimum cost-levels

6.3 Pricing and charging

General pricing and charging models are discussed in chapters 4.3 and 4.4. For pricing, cost-based and value-based pricing methods are introduced. These methods are compared in Table 11. Focus of the analysis is on the suitability and implementability of these two pricing methods for the environment of the case organization.

Cost structure of the case organization is presented in chapter 5.3.2. Due to the challenges in cost collection, it is very challenging to divide the costs in capital and operating expenses. Nevertheless, it can be assumed that the cost structure does not remarkably differ from the one that is presented in chapter 4.1.3. Cost structure has remained rather static and large changes are not expected to happen in the near future.

Table 11: SWOT-analysis of pricing models.

	Cost-based pricing	**Value-based pricing**
Strengths	- Simple model - Can be based on occurring costs - Transfer to cost-plus pricing is easy (only add markup)	- Represent the real value that the service is worth, not the costs that occurs
Weaknesses	- Defining the unit price may be challenging	- Requires thorough analysis about the value that network services generate for customers
Opportunities	- Changes in production costs can be transferred to prices almost instantly	- Customer and provider value can be maximized
Threats	- If cost allocation is done poorly, will lead to skewed prices	- Services are provided to multiple industries, which may lead to complex value analysis - Sudden changes in production costs may lead to serious errors in prices

The overall goal of service provisioning is to provide value to the business of customer. Different parts that constitute the whole service experience and deliver the value, would require very thorough analysis. This kind of analysis is not plausible to conduct during this thesis. It requires cross-unit collaboration from all units that participate in the overall service provisioning. For now, cost-based pricing is simple to implement and it provides information about the profitability of services.

Flat-rate and usage-based charging models are selected for comparison, since they are the two most commonly used ones. These models are compared in Table 12 from the perspective of the case organization.

During the interviews, transition towards usage-based charging was discussed. It was argued that the target state should be to charge everything based on the used capacity. There are two main arguments that are against the usage-based charging. First, network is dimensioned so, that there is a certain reserved capacity for a customer. For example, when a server is connected to the network, 1Gb link is reserved, which makes no difference, whether the whole capacity of the link is used or not. Secondly, usage-based approach would require massive changes to the accounting system.

Flat-rate charging can be justified with the nature of services. All services that are provided, have a static nature and their production generate fixed costs. Without pure variable costs, it is challenging to justify the usage-based component in service charging. It is true that flat-rate charging includes several possible efficiency defects,

Table 12: SWOT-analysis of flat-rate and usage-based charging schemes.

	Flat-rate charging	**Usage-based charging**
Strengths	- Simple to implement - Predictable income for provider - Predictable bills for customers - Current environment doesn't hold any pure variable costs	- Fair charging according to usage - Encourage customers to follow and plan their usage
Weaknesses	- Doesn't give customers incentive to plan or follow network usage - May hide network inefficiencies and possible bottlenecks	- Requires investments to measurement devices - Requires capacity and scalability from the accounting system - No pure variable production costs
Opportunities	- Tariffs can be used to reduce the possible generation of social waste	- Using tariffs in addition to usage-based charging, provider may be able to steer customer's behavior - May provide information about possible bottlenecks in the network
Threats	- May generate social waste - Light users easily subsidize heavy users	- May lead to unpredictable income for provider - May generate unpredictable bills for customers

by generating possible social waste and not giving incentive to follow and restrict service usage. These defects on the other hand are not that relevant in the given environment, as the services are almost always business critical and their usage is not tightly tied to network charges.

7 Future vision and action plan

This chapter proposes a future vision how to address the challenges that were identified to need attention. In the future, the IT will be consumed more and more as a packaged services.

At the moment, there is a larger business service management development project ongoing, that concentrates on improving the whole service management at the corporation level. The goals for the development project are to:

- Make sure that continuous services are up and running
- Standardize and centralize
- Increase automation level
- Support offering development
- Enable more business possibilities

These goals are kept in mind when proposing the solution for the future. Main focus is on the areas that have the largest impact on the charging of network services. Summary of proposed solutions is presented in Table 13.

7.1 Network design and management

As analyzed in section 6.1, definitions of standard ways of working and communicating them to all relevant stakeholders is crucial first step. To standardize and package the services in more manageable form, first step is to have service descriptions written out in more technical manner. Descriptions should take account the preferred infrastructure choices and comment the way environments are configured. This material should provide unambiguous definitions about the services to sales personnel and customers. Such definitions will also improve processes and reduce the working time spent in design and implementation.

Improvements in network design and management processes should aim to balance the workload of specialists and shift the focus to more proactive direction. When the continuous service processes are under a proactive assessment and improvement, quality and cost efficiency will constantly improve.

The shortly mentioned ongoing service management development project, will bring new possibilities to network and service management. This project should give possibilities to centralize and integrate management systems and processes. Work flow can be automated and simplified, but it requires thorough roll-out and active participation.

7.2 Service definitions

As the analysis shows, the current service descriptions are pretty abstract and do not give unambiguous description about their contents. The idea for the future is to

Table 13: Comparison of current and future state.

Area	Current	Target
Network design and management	- Very general and abstract service descriptions - Fragmented management and accounting tools	- Define clear design and implementation guidelines for environments - Define standard equipment, hardware and software versions
Cost allocation and time reporting	- Costs are gathered in one entity - Very challenging to map costs directly to services - Auditability of accounting records is challenging	- Define cost pools for ABC - Define activities for ABC - Define projects and tasks to time reporting portal based on activities - Allocate costs to services using ABC
Pricing and charging	- Relationship between costs and prices unclear - Vague charging practices	- Construct service prices based on ABC allocations - Base the charging on a monthly flat-rate - Direct the charges straight to customers
Service packaging	- Each case is started from a scratch - No predefined service packages or catalogs for mapping needs	- Identify and define core and supporting services - Divide supporting services to enabling and enhancing services - Identify and codify Patterns of Business Activity - Define User Profiles for servers and applications in collaboration with server and application management - Associate UPs with matching PBA

define services in such way that they represent vertical technology areas. This way the contents of the service will become clearer. This will also help to map costs to services using horizontal activities that are same across the services.

7.3 ABC and time reporting

For the future, accounting methods are to be developed to allow the cost records to be linked accurately to certain service components. Section 4.2 discusses the different approaches of allocating costs to services. To keep the accounting arrangements rather simple and enable simplistic cost follow-up, Activity-based Costing method is proposed to be used as basis. ABC is preferred over the Fully-Distributed Cost

Table 14: Services and their definitions.

Service	**Definition**
Data Center LAN management	Service that provides the connectivity to production network. Includes the management of LAN network, consultancy and technical expertise.
Firewall management	Firewall service used to provide secure connections to customer environments.
Load balancing management	Load balancing service is used to provide dual-site solution for customers and their environments.
DWDM management	Service includes the management of the optical DWDM network that is used to connect the different data center sites.
Internet connections management	Includes the management and provisioning of Internet capacity for the customer environments based on the needs.
Remote connection management	Includes different solutions to enable connectivity to customer's network environment from remote locations.
LAN-to-LAN VPN management	Service is used to connect customer's LANs using VPN techniques.
Proxy management	Proxy management service is used to enable proxy services for the customer environments.
Remote LAN management	Service to manage the customer LANs, such as office LAN and WLAN networks.
WAN management	WAN management service covers the management of customer's WAN connections between multiple sites.
DNS management	Domain Name System service for customer environments.

because it takes activities into account in more detailed level than FDC.

ABC as an accounting method helps employees to see and understand the different costs that occur in service provisioning. Different costs can be divided further into value-adding and non-value-adding activities. By continuously analyzing these activities, employees and management can form an overall picture of current situation, implement changes and improve overall efficiency (see figure 19).

ABC answers also to the need of more accurate time reporting. Time will be reported by activity. It is also important that each and every employee reports the working hours on a daily basis to guarantee 30 minutes accuracy. Following structure is proposed for allocating costs of continuous services. For each service (listed in Table 14), own project is to be established in the accounting system. For each service, tasks are defined based on ITIL service operation processes. These tasks are presented in Table 15.

Table 15: Tasks to be defined for continuous service projects in the accounting systems.

Task	**Description**
Continuous service management	Includes time that is used for identifying and solving incidents and problems, time that is used in change management activities, time that is used in planning, evaluating and implementing capacity changes and time used for ensuring service continuity.
Customer & Sales support	Time that is used in miscellaneous activities such as sales support and proposal preparations (if customer-specific project is not available)
Vendor cooperation	Time that is used to collaborate with subcontractors, contractors and vendors
Reporting	Time that is used to generate and analyze reports
External operating expenses	All external expenses are reported under this task (such as operator invoices, maintenance fees and training)
Assets	Asset-related invoiced expenses (basically equipment investments)

For cross-service type activities concerning process management, that are performed by incident and problem managers, own project will be established. Under this project, tasks for incident, problem and change management are to be defined.

Proposed task structure allows rather simple auditability of costs. These activities are chosen as they represent the basic ITIL processes: they are familiar to everyone, and thus it is easy to identify under which activity the performed work belongs. On the other hand, they are general enough which allows the number of activities to remain reasonable.

Next steps on this area are to implement these changes to the accounting system and communicate them across the organization. During a transition period, accounting records need to be followed and analyzed, whether any changes are needed. After the period, construction of new service prices can be done by auditing the cost records.

7.4 Unit prices and charging principles

Network environment is based on fixed costs as there are no real variable components. Given this fact and the environment design principles, it is sufficient to rely on cost-based prices. As proposed in earlier section, Activity-Based Costing method should be used to construct the service cost pools. From the service cost pools it is rather trivial to derive the unit prices.

Table 16 lists the services and their charging principles. The goal in the following structure is to unify and simplify the practicalities across services. Unifying charging practicalities enable better visibility for customers.

Table 16: Charging principles for services.

Service	**Charging principle**
Data Center LAN management	Internally per switch port.
Firewall management	Externally per environment size.
Load balancing management	Externally per environment size.
DWDM management	Internally/externally per reserved bandwidth.
Internet connections management	Externally per reserver bandwidth.
Remote connections management	Externally per user amount.
LAN-to-LAN VPN management	Externally per managed site.
Proxy management	Externally per user amount.
Remote LAN management	Externally per site size.
WAN management	Externally per managed site.
DNS management	Externally per environment.

As mentioned in section 7.3, construction of cost pools will take a transition period. Unit price for each service will then be calculated based on the defined unit of each service.

7.5 Service packaging

Successful transition to packaged service offering requires dealing with challenges identified in the field of standardization and network management. Division in core and supporting services will be done according to following proposition.

Core service package

The goal of data center network services is to provide production network connection for a server that reside in the data center. Core service package should address this goal by combining network services in one entity. Following services are proposed to be combined as a core service package which is always used as a base for a new environment:

- *Data Center LAN management* to connect the servers in the network infrastructure, provide management connections and basic network monitoring
- *Firewall management* to ensure the security of customer blocks

Enabling services

External connectivity tend to be more dependent on preferences and requirements of customers, and thus there should be more options to complement the core service. These services are to be defined as enabling services:

- *Internet connections management*
- *Remote connections management*
- *LAN-to-LAN VPN management*

Enhancing services

In data center environment, value adding services relate either to security, connectivity or capacity. These services are to complement the core service package and enabling services. For example, if customer requests that the environment need to be a dual-site solution, then the core service is complemented with DWDM and load balancing services. In the current service portfolio, following services are proposed to be defined as enhancing services:

- *DWDM service*
- *Load balancing service*
- *Proxy service*
- *WAN management*
- *Remote LAN management*
- *DNS management*
- *On-shore service management*

Although this kind of packaging model is proposed, it is not intended to restrict service offering. Combining the service packaging with patterns of business activity and user profile catalog, which are discussed in the following section, offering development is intended to become simpler as requirements can easily be mapped to corresponding service packages.

7.6 Patterns of business activity and user profile catalog

Environments and the number of different servers vary between customers, but server roles are rather similar in different environments. These different needs and roles are addressed in section 2.3.1 by introducing the terms of Patterns of Business Activity (PBA) and User Profiles (UPs).

Table 17 presents an example of a table that can be used to identify PBA. This list is constructed from the network point of view and helps to assess the needed services for a certain server.

For UPs, a catalog that matches the applicable PBA to a certain profile should look like an example in Table 18.

Table 17: Example of tool to codify the patterns of business activity.

PBA No. xx Activities	Activity levels					
	Hi	3	2	1	Lo	N/A
Network bandwidth requirements						
Data volumes						
Availability requirements						
Connectivity requirements						
Information sensitivity						
Redundancy requirements						

Table 18: Example of User Profile catalog that matches applicable PBA to profile.

User Profile (UP)	Applicable Pattern of Business Activity (PBA)	PBA code
Web server	Large number of connections	xxY
Database server	High data volumes, high availability, sensitive data	zzY
Application server	High number of transactions, large number of connections	ppY
File server	High availability, sensitive data	iiY

8 Conclusions

This chapter summarizes the conducted work. Results are assessed and their exploitativity is evaluated. As a final part of this section, recommendations for the future work are given.

8.1 Results

The objective of this thesis was to investigate the possibilities to develop charging model of network services in data center environment. First, the starting state was explored and the main issues were recognized. A target state was created according to the literature and future expectations inside the organization. The target state aims to answer the following questions, that were set in the beginning of this thesis:

- How should the service offering be developed?
- What costs are related to service provisioning and how those should be allocated and followed?
- How should the costs be retrieved by pricing the services and charging?

Case study revealed many challenges that have effect on service charging. Not all of them are even tightly related to accounting or charging, but are so fundamental that they have effect on the whole service provisioning. Due to the scope of this thesis, only such challenges that have direct relation to charging were discussed.

Service definitions and cost collection and allocation were seen as the largest challenges. To overcome these challenges, improvements in service definitions and service structure were proposed. More detailed service definitions and clearer division of services by their technological area were proposed as a solution. This solution is the base for all other proposed actions. It will simplify the service offering, but also help to map occurring costs to services.

Activity-based costing as a cost collection and allocation method was proposed to solve the identified challenges. New service structure was taken as a base for the new cost pools, and activities were defined according to common operational processes. The new way of collecting the service related costs is to help cost follow-up and to give better visibility on efficiency and profitability of different services.

Defined higher level services and cost pools were taken as a base for pricing. Cost-based pricing was proposed to be used as pricing method because of its simplicity. Defined services were divided into service units and the total production costs of a service were divided to the units accordingly. Prices generated this way represent the actual costs and also give feedback about the performance of the organization.

8.2 Assessment of results

The results of this case study are related to the organization in which the case study was conducted. This study concentrated on identifying the largest challenges in accounting, pricing and charging activities and to address them with a simple solution, that can be implemented with relatively small effort and costs. The aim was also to fit the proposal to the overall strategy of the organization.

The subject of this thesis turned out to be rather challenging. There is rather little research material available in this field. Industry benchmarking was also challenging, since cost structures and pricing principles are not publicly available information, and thus hard if even possible to obtain. The lack of automation in network management and reporting in the organization also partially lead to challenges when gathering material. These reasons caused the material remain relatively constrict.

Although the material was relatively constrict, the interviews and observations showed very consistent results. If the same study would be performed again in the same organization, the results would be somewhat identical. Interviews and observations are always subject to personal habits of the interviewer and observer, which is a fact that cannot be neglected. The questions were tried to keep rather general, to get the voice of the interviewee to stand out. This succeeded pretty well, as every interviewee signaled same kind of challenges from different aspects.

Case study part of this thesis is mainly applicable in the organization that it is conducted about. On the other hand, the issues that were found in this study, are likely to be rather common and other organizations may also recognize them. Proposed solution and actions are not to be taken as definitive, but organization specific proposals.

8.3 Future work

There has been many attempts to create a comprehensive Configuration Management Database (CMDB) for the needs of the organization. This has failed due to the differing needs of many units. Now there is a new project beginning, that aims to solve this problem. Decent CMDB solution would help to solve many challenges listed also in this study. Thus it is important to participate actively in defining the requirements from the network's point of view.

Network services in data center environments are in a key role when value is provided to customers. There is still rather little or no research at all available on field of value-based pricing in such environment. As the longer-term strategy of the organization is to increase the customer value, parts that really affect the total value should be investigated in more detail.

Traditional data center environments tend to be rather static in sense of capacity reservation. Thus the charging principles are also based on static measures. To be able to move towards dynamic gathering of charging data, further studies are required in the field of measurement and accounting arrangements.

One clear area of improvement in the case organization is the energy consumption follow-up of network equipment. A solution to clearly follow the energy consumption would give input for network design and selection of devices and technologies. The ability to follow the energy consumption would also enable new possibilities in the field of charging, as the energy is getting more and more expensive all the time. Price of energy will emerge to a real variable component to network service costs soon and probably become a dominant component.

References

12Manage. Activity based costing (abc). URL http://www.12manage.com/methods_abc.html.

Mohammad Al-Fares, Alexander Loukissas, and Amin Vahdat. A scalable, commodity data center network architecture. *SIGCOMM Comput. Commun. Rev.*, 38:63–74, August 2008. ISSN 0146-4833. URL http://doi.acm.org/10.1145/1402946.1402967.

APC. Determining total cost of ownership for data center and network room infrastructure, 2005. URL http://www.linuxlabs.com/PDF/Data%20Center%20Cost%20of%20Ownership.pdf.

Theophilus Benson, Ashok Anand, Aditya Akella, and Ming Zhang. Understanding data center traffic characteristics. In *Proceedings of the 1st ACM workshop on Research on enterprise networking*, WREN '09, pages 65–72, New York, NY, USA, 2009. ACM. ISBN 978-1-60558-443-0. URL http://doi.acm.org/10.1145/1592681.1592692.

Theophilus Benson, Aditya Akella, and David A. Maltz. Network traffic characteristics of data centers in the wild. In *Proceedings of the 10th annual conference on Internet measurement*, IMC '10, pages 267–280, New York, NY, USA, 2010. ACM. ISBN 978-1-4503-0483-2. URL http://doi.acm.org/10.1145/1879141.1879175.

Robert Braden. Requirements for internet hosts – communication layers. RFC 1122, Network Working Group, October 1989. URL http://tools.ietf.org/html/rfc1122.

CAIDA. Preliminary measurement spec for internet routers. Technical report. URL http://www.caida.org/tools/measurement/measurementspec/.

Koen Casier, Sofie Verbrugge, Raf Meersman, Jan Van Ooteghem, Didier Colle, Mario Pickavet, and Piet Demeester. A fair cost allocation scheme for capex and opex for a network service provider. June 2006. URL http://biblio.ugent.be/input/download?func=downloadFile&fileOId=570460.

K. Christensen, P. Reviriego, B. Nordman, M. Bennett, M. Mostowfi, and J.A. Maestro. Ieee 802.3az: the road to energy efficient ethernet. *Communications Magazine, IEEE*, 48(11):50 –56, november 2010. ISSN 0163-6804. doi: 10.1109/MCOM.2010.5621967.

Cisco. *Data Center Infrastructure 2.5 Design Guide*. 2007.

KC Claffy and Tracie Monk. What's next for internet data analysis? status and challenges facing the community. volume 85, pages 1563 – 1571, October 1997.

Costas Courcoubetis and Richard Weber. *Pricing Communication Networks: Economics, Technology and Modelling*. Wiley, Chichester, UK, 2003. ISBN 0-470-85130-9.

Ralf Dewenter and Justus Haucap. *Access Pricing: Theory and Practice*. Elsevier, Amsterdam, The Netherlands, 2007. ISBN 978-0-444-52803.

Eurostat. Electricity prices by type of user, 2011. URL `http://epp.eurostat.ec.europa.eu/tgm/refreshTableAction.do?tab=table&plugin=1&pcode=tsier040&language=en`.

FCC. Telecommunications act of 1996, 1996. URL `http://www.fcc.gov/Reports/tcom1996.pdf`.

James Hamilton. Overall data center costs, September 2010a. URL `http://perspectives.mvdirona.com/2010/09/18/OverallDataCenterCosts.aspx`.

James Hamilton. Energy proportional datacenter networks, August 2010b. URL `http://perspectives.mvdirona.com/2010/08/01/EnergyProportionalDatacenterNetworks.aspx`.

Andreas Hinterhuber. Towards value-based pricing – an integrative framework for decision making. *Industrial Marketing Management*, 33(8):765 – 778, 2004. ISSN 0019-8501. doi: DOI:10.1016/j.indmarman.2003.10.006. URL `http://www.sciencedirect.com/science/article/B6V69-4B94KC0-1/2/f0a494ebd0f12cf9766e9abaf91396e4`.

Jan Hof. Three smart ways for the network to save power. 2007. URL `http://tnc2008.terena.org/core/getfile.php?file_id=191`.

Juha Järvinen. Testing and troubleshooting with passive network measurements. Master's thesis, Helsinki Universtity of Technology, 2009.

Gerry Johnson, Kevan Scholes, and Richard Whittington. *Exploring Corporate Strategy*. Pearson, Essex, England, 2005. ISBN 0-273-68734-4.

Srikanth Kandula, Sudipta Sengupta, Albert Greenberg, Parveen Patel, and Ronnie Chaiken. The nature of data center traffic: measurements & analysis. In *Proceedings of the 9th ACM SIGCOMM conference on Internet measurement conference*, IMC '09, pages 202–208, New York, NY, USA, 2009. ACM. ISBN 978-1-60558-771-4. URL `http://doi.acm.org/10.1145/1644893.1644918`.

Kalevi Kilkki and Olli-Pekka Pohjola. Value-based methodology to analyze communication services. June 2006. URL `http://kilkki.net/files/50ajatelmaa.ajatukset.fi/tiedostot/pohjola_kilkki_ctte06.pdf`.

Jonathan Koomey. A simple model for determining true total cost of ownership for data centers. 2007.

James F. Kurose and Keith W. Ross. *Computer Networking – A Top-Down Approach*. Pearson, New York, NY, USA, 2008. ISBN 978-0-321-51325-0.

Michael H. Lambert. A model for common operational statistics. RFC 1857, Network Working Group, December 1995. URL http://www.ietf.org/rfc/rfc1857.txt.

Thomas T. Nagle, Joseph Zale, and John E. Hogan. *The Strategy and Tactics of Pricing*. Pearson, Upper Saddle River, NJ, USA, 2010. ISBN 978-0-13-610681-4.

OGC. *ITIL, Service Strategy*. The Stationary Office, Norwich, UK, 2007a. ISBN 978-0-11-331045-6.

OGC. *ITIL, Service Operation*. The Stationary Office, Norwich, UK, 2007b. ISBN 978-0-11-331046-3.

Greg Schulz. *The Green and Virtual Data Center*. Auerbach Publications, Boca Raton, FL, USA, 2009. ISBN 978-1-4200-8666-9.

Penny M. Simpson, Judy A. Siguaw, and Thomas L. Baker. A model of value creation: Supplier behaviors and their impact on reseller-perceived value. *Industrial Marketing Management*, 30(2):119 – 134, 2001. ISSN 0019-8501. doi: DOI:10.1016/S0019-8501(00)00138-3. URL http://www.sciencedirect.com/science/article/pii/S0019850100001383.

Tieto Oyj. Tieto business in 2010, 2011a. URL http://www.tieto.com/annualreport/our-business-in-2010.

Tieto Oyj. Tieto history, 2011b. URL http://www.tieto.com/about-us/history.

UptimeInstitute. Green enterprise award 2011: Distributed heating using data center waste heat, 2011. URL http://symposium.uptimeinstitute.com/images/stories/symposium_2011_files/GEITcasebriefs/Tieto_p.pdf.

VihreäICT. Numerotietoa, 2007. URL http://www.vihreaict.fi/fi/fi_6.html.

Russel Wolak, Stavros Kalafatis, and Patricia Harris. An investigation into four characteristics of services. *Empirical Generalisations in Marketing Science*, 3:22 – 43, 1998.

Robert K. Yin. *Case study research: design and methods*. Sage Publications, Thousand Oaks, CA, USA, 2009. ISBN 978-1-4129-6099-1.

Valarie A. Zeithaml. Consumer perceptions of price, quality and value: A means-end model and synthesis of evidence. *Marketing*, 52:2 – 22, 1988. URL http://areas.kenan-flagler.unc.edu/Marketing/FacultyStaff/zeithaml/Selected%20Publications/Consumer%20Perceptions%20of%20Price,%20Quality,%20and%20Value-%20A%20Means-End%20Model%20and%20Snthesis%20of%20Evidence.pdf.

Appendix A – Interview form

Interviewee, title and position in organization

Charging Model Development For Data Center Network Environment

The purpose of this interview is to gather information about the current state of network services and their charging practices in the organization. These questions are guidelines for the discussion, they are not intended to restrict the discussion but to introduce the topic. Discussion results will be analyzed to find the largest challenges to overcome in order to improve the charging practicalities.

1. How would you describe the area that you are responsible of at the moment?
2. What do you think are the largest challenges at the moment what comes to charging?
3. What aspects are organized well at the moment?
4. How do you see the overall service structure?
5. What is your opinion about service packaging?
6. What things you want to emphasize?

Thank you for your participation!

Miika Heini