S38.3115 Signaling Protocols – Lecture Notes

Lecture 12 – Motivation for Voice and Video over IP

Contents

Introduction

In this lesson we discuss the business impact of the Internet and IP networks on voice services. This is useful because broadband changes the nature of network business dramatically compared to the narrow band era and because these changes are reflected on the technical architectures that are created to provide services over the broadband networks.

There are actually two markets that define the future of Voice (and video) services over IP networks. One is the public services market intended for residential and mobile users as well as corporate customers. The systems are put in place and maintained by the operators that sell services on the market. The second is that of corporate ICT services market. On this market the systems are bought, put in place and usually also maintained by the IT departments in corporations. We will first look at the public services market and then come back to the corporate market.

We already discussed the ongoing changes in the transmission and transport infrastructure. Broadband networks are spreading and the business driver is the demand for data traffic. Data traffic comes from two sources. More valuable to the operators is the capacity sold to companies for interconnecting their sites. The second source is the residential and mobile broadband subscriber data traffic. Once the broadband infrastructure has been built, the operators start to see their old circuit-switching infrastructure as an inferior cost burden. At the same time, one should not forget that even today most of the revenues the operators earn come from voice services.

Fundamental constraints

To have a sustainable service business the following must hold:

Cost of service < price < Willingness to pay for the service

It turns out that the customer's willingness to pay greatly varies between different services. Docent Kalevi Kilkki has estimated that the willingness to pay per bit varies as:

Data : voice : sms = 1 **:** 1000 **:** 1000 000.

Ideally, a service provider wishes to set the price as close to the willingness to pay as possible. The willingness varies between customers. By customer segmentation the sellers of goods and services try to group their customers by willingness to pay so that they could maximize their revenues. The nature of the market and the competition on the market may hinder the sellers pricing their goods close to the willingness to pay.

PSTN service paradigm vs Internet service paradigm

The service paradigm expresses the nature of the financial obligations between the service operator and the customers.

The *paradigm in the PSTN, ISDN and GSM* networks has been such that the operator allocates a limited capacity on-demand for a call end to end and charges a tariff that may depend on the distance between the communicating parties. The tariff is defined as *money per time unit*. The operator may and usually does in addition charge a montly connection fee. The users can send within the capacity as they please or not send anything, the tariff does not change.

Circuit networks have limited capacity. Dimensioning of the network ensures that during the busy hour, some expected maximum traffic can be carried. The circuit switching technology goes into great lenghts in order to measure and control the use of network resources. Charging is closely related to resource use.

In economic terms the kind of service that is produced under the circuit network paradigm is both *depletable* and *excludable*. The first property means that the service has limited availability. Once the first call above the dimensioned capacity attempts to gain access, the result is a busy tone. Excludable means that once the operator has allocated the capacity for a pair of users, the operator can not allocate the same capacity to anyone else.

As a result of these two properties, economic theory tells us that markets will establish a price for the service that is higher than the cost and less than the willingness to pay for the service. Because of the low cost of technology, a long term sustainable business is possible. This is confirmed by the history of telephony services provisioning. Due to

these two properties, markets of circuit switched services are similar to the markets of physical goods. The latter are also depletable and excludable.

The *Internet paradigm* is different: the operator sells access to a network but does not guarantee any quality of the service. IP network service is characterised as *Best Effort*. The nature of the technology forces the operators to overprovision their core networks. Otherwise, the service would collapse under high load[1].

Let us first look at the Internet paradigm in light of the *classical economic theory*[2]. One can argue that it is always possible to add another flow into the network for a previously inactive customer. The service is *non-depletable*. This is true under the constraint of overprovisioning. The rate of customer flows is not guaranteed. If more flows arrive, the rates are dynamically adjusted by TCP such that even more flows can be accommadated. This means that the service *non-excludable*. Another way to look at the situation is to note that in the Internet there is a constant state of oversupply of services.

Due to the non-depletability and non-excludability, the marginal cost of providing the best effort service in the Internet approaches zero. Naturally, this ignores the long term fact that the network will become congested and investing into additional capacity from time to time is a must. It also ignores the fact that a new customer brings additional operational costs in terms of customer assistance, billing etc. Nevertheless, additional usage by existing customers will usually not increase operator costs. That being said, to avoid network collapse and thus to avoid going out of business the Internet network operator is forced to invest into additional capacity from time to time. To minimize operational costs, it makes sense to increase network capacity in significant steps and the previously congested network is back in the state of overprovisioning.

It follows from the economic theory that if the markets are ”efficient” the price of the service will approach to the marginal cost. This means that the price approaches zero. To make the business sustainable and to be able to invest into new capacity, the operators must somehow ”tax” the users. Tax, as opposed to price, does not depend on usage. This corresponds to the idea that the economically efficient way of charging for the Intenet use is a *monthly flat rate*.

This is confirmed by the history of the Internet service provisioning. History also tells us that ISPs have not been particularly profitable. This can be expected because the network provides a simple service that offers very limited ways for differentiating the service.

On the other hand, when the IP network has been seen as a replacement to traditional data services provided by incumbent network operators and the skills needed to build large IP-networks have been scarce, the tax that a competent ISP has been able to collect, has been

[1] Contrary to this, circuit switched exchanges are required not to collapse under load much heavier than what they can process. This is achieved by elaborate overload control implemented in the exchanges. Since traffic arrivals are call attempts, it suffices to block extraneous call attempts at the edges of the exchange. The result is that the system works at dimensioned maximum capacity although the offered load is several times higher.
[2] For example, “Internet Economics” by MIT researchers.

high compared to cost. Historically, it has been more typical that vendors have been able to attract the skills better than ISPs and thus able to earn high returns.

Trying to predict the future, one can claim that during the broadband era people have grown to depend on the availability of the ISP service and the willingness to pay is growing in developed countries. It is a different matter whether the operators are able to collect a price that would come close to the willingness to pay.

Contrary to what researchers working on the basis of classical economic theory are saying, we can argue that a non-depletable service requires *infinite capacity*. Naturally, no network has infinite capacity. The argument leads to the conclusion that Internet service is depletable. Therefore, the classical theory fails to give a satisfactory interpretation to the historical fact that flat rates dominate Internet service.

The critics look for answers in the theory of *behavioral economics*. Observations from markets of information goods show that users try to *avoid loss* and therefore *like predictable pricing*. Data services are difficult if not impossible to price based on volume or time so that the user would feel the price predictable. Moreover, adoption of a service requires experimentation by the users. If prices are unpredictable, users are afraid to experiment. This slows down adoption of the service. The result is that users prefer flat rate prices for data services. The argument goes on to claim that it is not so much a question of the level of the price as the pricing structure that determines the adoption of the service.

Observations of network services adoption in many markets show that when the market moves from volume prices on data services to flat rates, the adoption of the service grows about ten-fold in a short period of time.

Whatever theory we apply, the fact we are trying to explain is that flat rate pricing for data services leads to fast adoption of the service and that Internet has historically been dominated by flat rate pricing.

The economic fundamentals are such that under ideal markets it is challenging to sustain a long term business with the Internet best effort paradigm. Clearly, the business logic differs from what incumbent operators are used to in circuit switched networks. Moreover, with broadband the operators are facing the threat of cannibalization of their cash cow – the voice service. In order to understand what the operators can do to survive and to sustain their business, we must look at the nature of competition in the information economy and in particular in the network services business.

Let us first analyse the nature of competition in information economy.

Competition in Information Economy

The classical model of competition in an industry was formulated by professor Michael Porter in his *five forces model*. Let us apply this model into an information industry. Our model is presented in Figure 12.1.

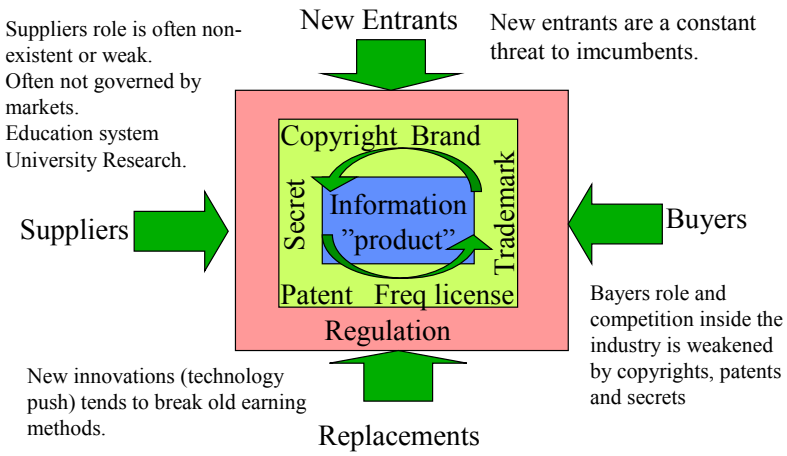

Figure 12.1: Competition in an Information industry

According to Porter, the five forces shaping the competition in an industry are: *customers' buying power, the suppliers' role, the impact of replacement products emerging for example due to innovation, the role of new entrants into the business and intra-industry competition.*

We have added into the classical 5 forces model the entry barriers to competing on the markets. We observe that in all major information industries one or another type of entry barrier plays an important role. For example, mobile cellular operators have *frequency licenses*, the movie and recording industries have *copyrights* to content, providers of commercial software can keep their source code *secret* from customers and competitors and it is possible to *patent inventions*. All these legal protections give a monopoly right to the seller. The scope of the monopoly right is always limited and there may exist alternatives on the market but nevertheless, we can see that these rights are crucial to many industries and that some of the industries could not survive at least in the present form without these rights.

The significance of these monopoly rights in different areas of information industry varies. The ways they are used to earn money also vary. One method is to embed a set of patents into gadgets and sell the gadgets. One phenomenon is common: all information industries that depend of any form of monopoly right spend a lot of efforts for the sake of protecting the rights.

We also show one more de-jure entry barrier that in itself is information, namely *trademarks*. They are usable both for physical and for information goods.

We also observe that there are *de-facto* entry barriers in some information industries. For example, in most cases it is not economically viable to build many fixed wire line access networks to residential customer homes. Once someone has lain down the cables to homes, another fixed network operator is very unlikely to come and build a parallel access network reaching the same homes. This is called a "natural monopoly". Fair, enough there are exceptions. In countries were cable-TV has a major role, there may be two parallel copper wireline networks reaching a significant portion of homes. For example this is the case in the UK and the USA.

Even softer entry barrier that is commonplace also for physical goods is *brand*. This barrier in itself is information regardless whether it relates to information or physical goods.

The box in the model has one more layer. If the products are protected by monopoly rights, one can expect that an ideal market will not emerge. On the other hand, if there are no monopoly rights at all, then it may be that no market is possible at all because copying of information is so cheap. On the outer border of the pink box, the target of pink area activity is help creating "normal market conditions".

The governments are protecting the monopoly rights but we live in a market economy – at least this is a popular belief. Also, we know that monopoly rights tend to slow down innovation and make the economy inefficient. To counter these undesirable tendencies, governments have competition policies, anti-trust laws and particularly for telecommunications they have industry specific *regulation*.

Characterizing the forces, one observation is that electronics, computer and network technology are creating innovations and bringing replacement products and delivery methods into information industries at a very fast pace. Due to this, the industry has been in a constant state of turmoil from the 1980's. The impact on business is drastic when the base technology that is used to process, store, present and transfer information has been doubling its capability in one or two years and this has continued for several decades.

Concerning the de-jure entry barriers, one should note that a business cannot be sustainable on ideal markets if the price for the goods tends to zero. The copyrights, patents and licenses are necessary in order to create an economy. However, politicians and governments have a difficult task in balancing the impacts of these monopoly rights with consumer rights in order to keep the economy innovative and avoid inefficiencies emerging under the protection of the monopoly rights.

Over the past decade we have seen legal battles around the phenomena of file sharing using peer-to-peer networks and the copyrights protected industries. These battles should be looked at in light of this model of competition.

The model we have presented is generic and not all barriers apply to all types of information industries. The external forces also vary. Next, let us apply this generic model into the network services industry.

Network Economics

Figure 12.2 shows the 5 forces model applied to the network services industry. The set of entry barriers are limited to the most relevant ones for this industry.

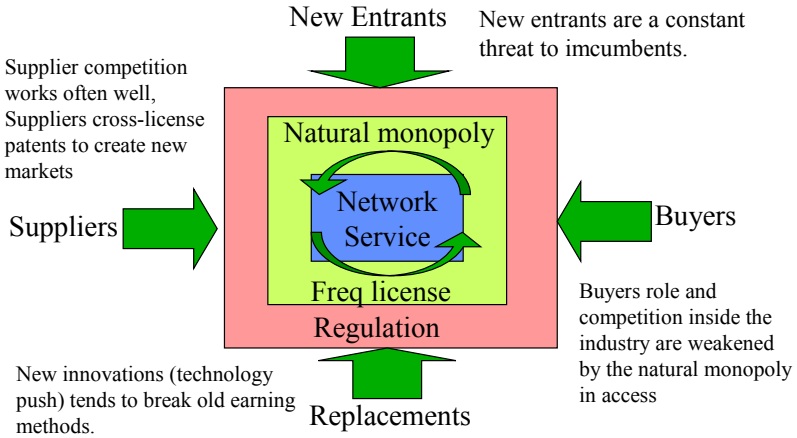

Figure 12.2: Five forces model applied to network services industry

In fixed network services the natural monopoly created by cabling the homes limits the competition among the operators (intra-industry competition). It also limits the buyers' buying power. This, however, is not that relevant when it comes to providing network access to companies particularly in cities. The value of data services connectivity to companies is so high that the price operators are able to collect covers the costs of even parallel cabling to companies in densely populated areas. The role of the natural monopoly based on wired access to homes has been decreasing lately due to wireless technologies that make is economically feasible to build several parallel access networks in most geographical areas. With increased wireless capability and unwillingness of operators to invest into Fiber in many countries, wireless has become more and more a broad purpose replacement for wired access.

Radio frequencies are a scarce natural resource. Frequency licenses so far have been necessary to control the use of radio frequencies for different purposes. However, like we mentioned, the economies of radio access differ from the economies of fixed access. It is economically feasible to build several parallel radio access networks. As a result, real competition between mobile operators for customers is possible.

Users place a great value on mobile services. Due to high willingness to pay, the mobile industry has grown phenomenally. Due to high demand that lasted for several years (1990's), it was not necessary for operators to compete on price. In countries where the penetration has saturated because almost everybody already has a mobile phone, price competition has appeared as the model predicts.

The motivation of mobile broadband is problematic for the mobile operators. Technology push leads to a situation where they can not stop investing but the increase of supply

created by broadband mobile access threatens the cash cow of voice services. That is why the operators are trying to strike a fine balance between supply and demand among their customers when making investment decisions. This is normal market behavior and only shows that we have come a long way from the era of the 3G hype around the year 2000, when major European operators sunk billions of Euros into 3G frequency licenses.

Broadband IP networks are a replacement to circuit switched fixed networks. They create a major threat to the way operators earn money from voice services. We will see how this threat works and how the operators are trying to tackle the threat. IP networks also create opportunities for new entrants to offer services to consumers. This is due to the fact that the impact of distance is diminished by broadband. This happens because the cost of transferring a bit of information over a distance has been falling faster than the cost of computing power that has been following the famous Moore's law. The winners are the buyers of network services including consumers and corporations. The high value of communication services tends to fall into their hands leading to a high demand for the services and rapid growth of traffic.

Drivers for End-to-end vs. Operator managed services

Services in the Internet can be provided by host computers owned by users themselves or by third parties. We call this end-to-end services provisioning. The traditional model of services provisioning has been that the network operator provides the service and sells it to its customers. We call such services operator managed services.

The driver that is changing signaling systems is broadband and the Internet. So, let us try to analyze how services can be provided in IP networks and how this is different from circuit networks.

An ideal IP network gives each user an IP address and the user becomes reachable by any other user or server that also has an IP address. The operator's role is doing its best to carry packets end-to-end without much means to control who sends what to whom or how many packets are sent. The basic role of the operator is to provide a bit pipe for the users. Operators try to differentiate by bundling additional services such as e-mail accounts etc to their service offering. Markets, however, very quickly force every operator to offer the same bundles. Entry barriers to the additional services business are very low. Once the technology for new services is available from the vendors, it is quickly offered to all operators. There are typically no de-jure means for the operators to protect their way of making money on the additional services.

Recently, several global players have emerged in the area of providing additional services over the Internet. Most prominent of them earn their revenues from advertisements rather than from the users of its service. This development confirms the diminishing importance of distance. (Prime examples are Google and Facebook).

It is worth to note that for companies that provide "free services" end users are not customers. Their customers are the advertisers who pay the money. What are the end-

users then? A nice definition is “potential customers of the advertisers”. Another way to look at them is that they (=most of us) are victims.

On the other hand, terminals in IP networks are computers. It is thus possible for the users to implement their own services on their own end systems. We say that IP facilitates user innovation. Since the users know best what their needs are, they can benefit from this service innovation opportunity. Examples of services that can be attributed to user innovation are the World Wide Web and Peer-to-Peer overlay networks.

Users and service providers that provide services over the Internet independent of the incumbent operators that own and build the network infrastructure are in the same position. This has created new opportunities for new entrants to the network services business. Examples are Google, Amazon and Skype. Often the core competence of such service providers comes from outside the network services industry. These companies just make use of the network to deliver a service to the masses. Some of these service providers (e.g. SKYPE) offer a service that is a direct replacement to what the network operators owning the infra have been offering all along.

The network operators, however, have some competitive advantages. They are usually perceived by the consumers and corporate customers as trusted companies and they do have a strong presence on the market. Consumers tend to go to them because they are used to do so. A customer of a network operator expects to know exactly the nature of the service he or she is getting and that for example the privacy of the service is never under doubt. The operators face the challenge of meeting this expectation.

Particularly, to corporations trusting their service provider is important. Corporations expect a reliable service that does not stop them running their own business. Protecting their privacy is also very important to corporations. Corporations use IT systems that they see critical to their business. IT systems have different and sometimes quite strict requirements on the network connectivity that they expect between the sites of the company.

The way to formulate the relationship of a (corporate) customer and the data network provider is called a Service Level Agreement (SLA). An SLA will detail requirements for the operator to *restore the service in case of failure* within a limited time. An SLA may also lay down *throughput*, *delay* and *packet loss criteria* for the service. To verify that the operator is meeting the requirements of an SLA, both the operators and the customer use elaborate active and passive measurement methods.

Once, the operator starts to guarantee or statistically assure the quality of its service, the nature of the business changes from non-depletable and non-excludable service to something closer to traditional physical goods kind of business with depletable and excludable goods leading to higher returns. The move, however, is not complete because the corporation can dynamically manage risk. A corporation can bet that its operator has an over-provisioned network and it will have enough capacity to carry its traffic although there are no written guarantees.

However, if the value of the business over the data network the company is conducting is high, the cost of the data service can be insignificant to the company. This is often the case. Under such circumstances any betting on the quality of service is out of the question. The company will rather invest in an IT infra that allows it to conduct its high value business at all times. (For example, take the stock exchange where the volume of transactions in a few minutes is measured in millions of Euro). This calls for highly robust services and fast reaction to network failures. This also calls for technologies that allow virtualizing the network operator's infrastructure. As a result, the operator can sell or rent a slice of its infrastructure to its corporate customers and earn a reasonable rate of return for the service.

Impact of end to end on operator business

Due to the openness of services provisioning in the Internet, new entrants such as SKYPE that provides a replacement service and Google that provides a new kind of service that was not available in circuit networks have come to play an important role. Value of connectivity is not falling into the hands of the operators like before.

At the same time the basic service expected from the operator, in a sense, is non-depletable and non-excludable. Providing voice over IP in a best effort network does not seem to make sense economically to the operators. VOIP would just cannibalize the operators' current cash cow, namely the PSTN and GSM service. This logic has led to slow adoption of VOIP by incumbent operators and has created an opportunity for new entrants to step in. In developed countries where mobile penetration has saturated, this has led to declining revenues to operators or at least cut the growth. PSTN is loosing customers. Instead, for example in Finland a high percent of users (>30% according to some sources) are using VOIP. Most of this is end-to-end rather than managed operator services.

At the moment this end-to-end trend is strong in the industry.

It remains to be seen whether the incumbent operators can win these users back. This depends for example on how much value the users will place on the quality of the service. When a home has its TV connected to the Broadband access, there are several TV set in the house, the teens are downloading entertainment on their own machines and still the parents need to do their own communications, it may be that QoS for the access will become more important.

Deployment scenarios for public services VOIP

First VOIP came to wireline networks that had sufficient capacity to carry voice among other services. Historically, all kinds of services that first became technically feasible in wireline networks have emerged also in mobile networks. The distance in time between the introduction of new services in wireline and mobile networks seem to remain more or less constant over time.

On the Internet that has lots of capacity, VOIP is a service among others. SKYPE has shown that it is feasible to provide it as an end-to-end service. Most of the time, the quality is very good. Fair enough, sometimes the quality collapses or is less than satisfactory.

VOIP can also be provided as a managed service by the operators. For example, BT is building the so called Next Generation Network. In the NGN customers are connected to Multiple Services Access Nodes (MSAN) instead of Digital Subscriber Line Access Multiplexers (DSLAM). An MSAN can house analogue subscriber line cards, ISDN subscriber line cards, ADSL line cards etc and is connected to an IP core network with high speed links. Telephony services are provided by a new IMS (IP Multimedia Subsystem) based infrastructure irrespective of whether the customer is using an analogue phone, ISDN phone or a VOIP client on some device at home.

VOIP can also appear as a feature of some software that was originally built for other purposes like games.

VOIP has started to appear on advanced mobile phones. It remains to be seen how successful that will be.

The most recent emerging driver for VOIP as a public service is the newest mobile broadband technology, LTE – long term evolution, that provides capacities of around 100Mbit/s to users (see max throughputs measured in NetRadar provided by Comnet). Contrary to all earlier mobile BB technologies LTE does not support circuit switching at all. So, the only way to bring telephony to LTE is to use packets. The plan is that IMS is the technology that will be used. In the interim, the operators can use the alternative: mobile devices support many radio technologies such as GSM, 3G WCDMA, HSPA and LTE and the voice telephony service is provided using the older radio technologies and LTE is reserved for data only. Because LTE is more cost efficient than the older technologies, this is only a temporary position for conservative operators.

Threats to VOIP that is just one service among others

If a user gets a static IP address for an ADSL service, it is a matter of minutes when someone makes the first attempt to break into the user's computer. This is proven by honeypot experiments that have been conducted e.g. by Comnet.

E-mail was the first successful service on the Internet. Due to unsolicited e-mail, or Spam, and poor culture of sending e-mail inside a company, e-mail may be coming to the end of its useful life.

Spamming is a business that is based on the fact that the cost of sending is insignificant compared to the cost of receiving. For example 1M sent advertisements of Viagra may result in several sold packets of the drug making the business case profitable to the seller and to the spammer.

Spamming is more efficient if the sender can make use of other users' hacked computers. So, spamming is actually one starting point of a value chain of grey economy on the Internet. Other players in the game are hackers that break into other people's PCs and form botnets of PCs that they sell to spammers etc.

Question is, if VOIP becomes widely adopted, what new unwanted network phenomena will emerge. For example, telemarketing becomes more profitable than it is now. The next question is whether it is possible to protect users from unwanted calls by technical means or whether senders need to be made pay for their actions?

If a host with a VOIP client is hacked, many kinds of fraud become possible. Particularly vulnerable are computers that house many services and are managed by the user himself.

A few years ago, an M.Sc thesis worker of Comnet placed a honey-pot into a C-block of addresses in a corporate network in Finland. The honey-pot successfully downloaded more than 1.4M Trojans over 4 months, i.e. one in about 8 seconds.
Attacks are launched from bot-net machines taken over by Trojans. All this activity is highly automated and professional. This should be taken into account when one thinks about the viability of new mission critical services using the Internet.

Over 2013, it became clear to the general public that governmental organizations such as National Security Agency (NSA) in the US, GCHQ in the UK and FRA in Sweden have access to lots of the Internet content. For example, if an Internet service is provided to foreigners from a datacenter in the US, by law, NSA has the right to read anything that the datacenter sees. Equally, FRA of Sweden has the right to read all Internet traffic originated abroad or passing through Sweden. It remains to be seen what the citizens and consumers will make of this violation of their privacy.

VOIP and Video on Corporate Markets

Traditionally voice services for corporate customers were provided using PABXs with their own signaling systems such as DPNSS, DASS and Q.SIG. These systems supported a very large set of supplementary services in a single PABX and also in a PABX network for multi-site companies.

With the emergence of VOIP, PABXs started to become hybrid systems that along the traditional telephony like voice services provided VOIP using H.323 and also SIP signaling.

Nowadays, on these markets, voice services fall under the more generic term of *collaboration services* that include application sharing, desktop sharing, shared editing, video and voice services. High quality delivery of these services set certain requirements on the underlying corporate and inter-corporate networks. Lately, this product segment shows healthy growth in the tens of percent per year.

Conclusion

VOIP puts the operators in front of a dilemma: either cannibalize their own cash cow by themselves or let new entrants take over the voice services business first in wireline and with the emergence of mobile broadband also in mobile services networks.

Users are attracted to “free” services such as peer-to-peer based voice over IP. At the moment the users are putting their trust on global services such as Skype making one company a global winner and in charge of a lot of data about lots of users. Question is whether this is a sustainable model in the long run.

The opportunity that the mobile operators have is that the cost of providing mobile broadband is low and will become even lower due to technology development in the coming years. At the same time the value of the service from the users’ perspective and the users’ willingness to pay for the service is high. The operators must learn to collect a sufficient portion of that value while they adopt the business logic emerging from the users’ preferences of flat rates and advertisement driven earning models.

Interactive Voice and Video over IP services have found a growth market in the corporate segment. For more intensive collaboration of teams in globalized companies reducing traveling and speeding up work processes, these services are essential.