Department of Computer Science and Engineering

Techno-Economic Analysis of Novel Opportunities for Mobile Networks

Open Innovation and Cloud Computing

Yrjö Raivio

Aalto University

DOCTORAL DISSERTATIONS

Aalto University publication series
DOCTORAL DISSERTATIONS 135/2012

Techno-Economic Analysis of Novel Opportunities for Mobile Networks

Open Innovation and Cloud Computing

Yrjö Raivio

A doctoral dissertation completed for the degree of Doctor of Science (Technology) to be defended, with the permission of the Aalto University School of Science, at a public examination held at the lecture hall T2 of the school on January 18th, 2013, at 12 noon.

Aalto University
School of Science
Department of Computer Science and Engineering
Data Communications Software

Supervising professor
Professor Antti Ylä-Jääski

Thesis advisor
Docent Sakari Luukkainen

Preliminary examiners
Professor Pieter Ballon, Vrije Universiteit Brussel, Belgium
PhD Frédéric Desprez, LIP-ENS Lyon, France

Opponent
Professor Jan Bosch, Chalmers University of Technology,
Gothenburg, Sweden

Aalto University publication series
DOCTORAL DISSERTATIONS 135/2012

ISBN 978-952-60-4840-6 (printed)
ISBN 978-952-60-4839-0 (pdf)
ISSN-L 1799-4934
ISSN 1799-4934 (printed)
ISSN 1799-4942 (pdf)
http://urn.fi/URN:ISBN:978-952-60-4839-0

Unigrafia Oy
Helsinki 2012

Finland

Publication orders (printed book):
yrjo.raivio@aalto.fi

NORDIC ECOLABEL
441 697
Printed matter

Aalto University

Aalto University, P.O. Box 11000, FI-00076 Aalto www.aalto.fi

Abstract

Author
Yrjö Raivio

Name of the doctoral dissertation
Techno-Economic Analysis of Novel Opportunities for Mobile Networks – Open Innovation and Cloud Computing

Publisher School of Science

Unit Department of Computer Science and Engineering

Series Aalto University publication series DOCTORAL DISSERTATIONS 135/2012

Field of research Data Communications Software

Manuscript submitted 10 August 2012 **Date of the defence** 18 January 2013

Permission to publish granted (date) 28 November 2012 **Language** English

☐ **Monograph** ☒ **Article dissertation (summary + original articles)**

Abstract

Mobile operators are financially in good shape. However, there are weak signals, which can be regarded as an early warning about the future prospects of the operators. Examples of the mobile terminal industry have proven that the incumbent mobile actors must be prepared for a new market, which is no longer built on the old achievements. This dissertation proposes two complementing propositions, which can be utilized for the operator renewal process. Open innovation has been one of the main drivers in the success of the internet companies, while the introduction of cloud computing has lowered the risk of entrepreneurship to a new level that has attracted in a vast number of novel startups.

The telecom industry can learn from these initiatives. Firstly, operators can apply open innovation to mobile networks, in accordance with open application programming interfaces (APIs), which offer developers novel opportunities to create services based on the operators' currently hidden network assets. We call this initiative Open Telco, which enables Network as a Service (NaaS) service model. Opening the network assets paves the way for a two-sided platform, also enabling new business models.

Secondly, operators should carefully review the opportunities and challenges of cloud computing. Clouds can be harnessed, not only as a service for customers, but also in the active utilization of the operators' software, platform and infrastructure systems, e.g. in SaaS, PaaS and IaaS service models. According to this dissertation, cloud computing is well suited to interactive SaaS services, and also the sharing of cloud computing PaaS platforms between operators offers interesting opportunities. The Open Telco platform can be built in PaaS fashion, too. Besides, even public IaaS clouds, including database services, can provide carrier grade performance.

In summary, mobile operators should, without prejudice, evaluate the novel opportunities and experiment with new concepts, technologies and business models. The telecommunication sector has been used to working on well-defined standards, but in the networked and digital economy, the old approach is too slow. Instead, the mobile industry should adapt the best practices of the Internet and transform their business to meet the future challenges.

Keywords mobile network, cloud computing, open innovation, performance, service

ISBN (printed) 978-952-60-4840-6 **ISBN (pdf)** 978-952-60-4839-0

ISSN-L 1799-4934 **ISSN (printed)** 1799-4934 **ISSN (pdf)** 1799-4942

Location of publisher Espoo **Location of printing** Helsinki **Year** 2012

Pages 182 **urn** http://urn.fi/URN:ISBN:978-952-60-4839-0

A! **Aalto-yliopisto**

Aalto-yliopisto, PL 11000, 00076 Aalto www.aalto.fi

Tiivistelmä

Tekijä
Yrjö Raivio

Väitöskirjan nimi
Teknis-taloudellinen analyysi matkapuhelinverkkojen uusista mahdollisuuksista – avoin innovaatio ja pilvilaskenta

Julkaisija Perustieteiden korkeakoulu

Yksikkö Tietotekniikan laitos

Sarja Aalto University publication series DOCTORAL DISSERTATIONS 135/2012

Tutkimusala Tietoliikenneohjelmistot

Käsikirjoituksen pvm 10.08.2012 **Väitöspäivä** 18.01.2013

Julkaisuluvan myöntämispäivä 28.11.2012 **Kieli** Englanti

☐ **Monografia** ☒ **Yhdistelmäväitöskirja (yhteenveto-osa + erillisartikkelit)**

Tiivistelmä

Matkapuhelinoperaattorit ovat taloudellisesti vielä hyvässä kunnossa, mutta niiden tulevaisuus ei ole välttämättä yhtä positiivinen. Viimeaikaiset esimerkit matkapuhelinteollisuudesta todistavat, että markkinat saattavat muuttua nopeasti, jolloin myös yritysten toimintatapojen on muututtava. Tämä väitöskirja esittelee kaksi toisiaan täydentävää ehdotusta, jotka edesauttavat matkapuhelinoperaattoreiden uusiutumista. Avoin innovaatio on ollut useimpien internet-yritysten kantava strategia ja toisaalta, pilvilaskenta on pienentänyt kasvuyritysten aloitusriskiä.

Myös matkapuhelinoperaattorit voivat hyödyntää samoja aloitteita. Ensiksi, operaattorit saattavat avata rajapintoja toistaiseksi suljettuihin tietovarantoihin, joita palveluiden kehittäjät kykenevät hyödyntämään uusissa innovaatioissa. Kutsumme tätä aloitetta nimellä avoin tietoliikennejärjestelmä eli Open Telco, mikä mahdollistaa NaaS-verkkopalvelumallin tarjoamisen. Avoimet rajapinnat luovat kaksipuolisen ohjelmistoalustan, joka avaa uusia liiketoimintamahdollisuuksia.

Toiseksi, operaattoreiden tulisi huolellisesti analysoida pilvilaskennan tarjoamat mahdollisuudet ja haasteet. Nykyisin operaattorit tarjoavat asiakkailleen pilvilaskentaa itsenäisenä palveluna, mutta pilvilaskenta soveltuu myös operaattoreiden omiin ohjelmisto-, alusta- ja infrastruktuuri-järjestelmiin eli SaaS-, PaaS- ja IaaS-palvelumalleihin. Tämän tutkimuksen mukaan pilvilaskenta sopii hyvin vuorovaikutteisiin SaaS-ohjelmisto-palveluihin ja operaattoreiden keskenään jakamiin PaaS-alustoihin. Myös Open Telco-ohjelmistoalusta voidaan rakentaa PaaS-palvelumallin mukaisesti. Lisäksi, jopa julkiset IaaS-laskenta- ja tietokantapalvelut täyttävät matkapuhelinverkkojen tiukat suorituskykyvaatimukset.

Yhteenvetona voidaan todeta, että matkapuhelinoperaattoreiden pitäisi ennakkoluulottomasti kokeilla uusia palveluita, teknologioita ja liiketoimintamalleja. Tähän saakka tietoliikennesektori on toiminut standardien ohjaamana, mutta tulevaisuuden kehitystä on vaikeampi ennakoida. Internet tarjoaa erinomaisia esimerkkejä onnistuneista innovaatioista. Niitä voidaan soveltaa myös operaattoriympäristöön, jotta ala pystyy vastaamaan tulevaisuuden haasteisiin.

Avainsanat matkapuhelinverkko, pilvilaskenta, avoin innovaatio, suorituskyky, palvelu

ISBN (painettu) 978-952-60-4840-6 **ISBN (pdf)** 978-952-60-4839-0

ISSN-L 1799-4934 **ISSN (painettu)** 1799-4934 **ISSN (pdf)** 1799-4942

Julkaisupaikka Espoo **Painopaikka** Helsinki **Vuosi** 2012

Sivumäärä 182 **urn** http://urn.fi/URN:ISBN:978-952-60-4839-0

Preface

First of all, I want to thank my supervisor, Professor Antti Ylä-Jääski, who kindly offered me a doctoral student position in his outstanding research team. His support and advice were most important in executing this work.

Secondly, the passion of my instructor, Docent Sakari Luukkainen, was enormously helpful from the start to the very end of the work. His inspiration and guidance were indispensable in overcoming the difficulties. In addition, our Networking Business team including, in alphabetic order, Koushik Annapureddy, Rushil Dave, Antero Juntunen, Ramasivakarthik Mallavarapu, Rasmus Päivärinta, Saku Seppälä and Alberto Vila Tena offered major support in co-authoring the publications. The team mates also spurred me and got involved creating the key research results of the dissertation.

Thirdly, the pre-examiners Professor Dr. Pieter Ballon and PhD in C.S. Frédéric Desprez did efficient work, giving the extremely valuable feedback on the dissertation. I am also honoured that Professor Jan Bosch promised to act as my opponent. I thank all anonymous and known reviewers for their useful comments on the publications. Special thanks I give to Soili Adolfsson and Kristiina Hallaselkä for practical arrangements, Edward Bonney for proofreading and Miika Komu for sharing his most valuable experiences of how to write a dissertation. I cannot either forget my former superior Lauri Oksanen and Services Research team, who laid the foundation of the whole research work.

Financially, the work was supported by Tekes (the Finnish Funding Agency for Technology and Innovation, `www.tekes.fi`) as a part of the Cloud Software Program (`www.cloudsoftwareprogram.org`) of Tivit (Strategic Centre for Science, Technology and Innovation in the Field of ICT, `www.tivit.fi`). I want to thank everyone in the Cloud Software Program for their invaluable support. Warmest thanks I owe to co-authors Oleksiy Mazhelis and Professor Pasi Tyrväinen, and the business case owners Antti Leinonen, Pekka Markkula and Mika Raitola, who continuously tested my ideas and gave constructive feedback.

Finally, I express my greatest gratitude to my parents, Saara and Erkki, who encouraged to life-time studies. Last but definitely not least, I thank my spouse Merja for patiently supporting the writing process and friends for inspiring my intellectual life.

> ”A new idea is first condemned as ridiculous and then dismissed as trivial, until finally, it becomes what everybody knows.”

William James (1842–1910)

Helsinki, December 9, 2012,

Yrjö Raivio

Contents

Preface i

Contents iii

List of Publications v

Author's Contribution vii

List of Abbreviations viii

List of Figures xiii

List of Tables xv

1. Introduction 1
- 1.1 Research Questions and Scope 7
- 1.2 Research Methodology 8
  - 1.2.1 Case Study 9
  - 1.2.2 Survey 10
  - 1.2.3 Experiment 11
- 1.3 Research Framework 11
- 1.4 Contributions 12
- 1.5 Structure of the Dissertation 14

2. Background 15
- 2.1 Service 15
  - 2.1.1 Innovation Strategy 16
  - 2.1.2 Open Innovation 18
  - 2.1.3 Mashups 19
  - 2.1.4 The Long Tail 20
  - 2.1.5 Challenges in Mobile Services 21
- 2.2 Technology 22
  - 2.2.1 Mobile Networks 22
  - 2.2.2 Distributed Computing 26
  - 2.2.3 Cloud Computing 28
  - 2.2.4 Dynamic Resource Provisioning 34
  - 2.2.5 Performance Testing 35
- 2.3 Organization 36

2.3.1 Value Chain and Value Network 37
2.3.2 Ecosystem 38
2.3.3 Value Network vs. Ecosystem 39
2.3.4 Network Effect and Two-Sidedness 40
2.3.5 Broker Models 41
2.4 Finance 42
2.4.1 Consumer Theory 42
2.4.2 Internet Revenue Models 43
2.4.3 Cloud Cost Analysis 43
2.4.4 Cloud Markets 44
2.4.5 ARPU Break-Even Point 44

3. Results 45
3.1 Service 45
3.1.1 Open APIs 45
3.1.2 Event Experience Mashup 46
3.2 Technology 47
3.2.1 Open Telco Architecture 47
3.2.2 Performance Evaluation 48
3.2.3 MVNO and OSS/BSS 50
3.2.4 Proof of Concepts 52
3.3 Organization 53
3.3.1 Virtual Broker 53
3.3.2 Two-Sided Platform 55
3.4 Finance 56
3.4.1 Cost Analysis of Hybrid Cloud 57
3.4.2 Cost Comparison of Hosted Solution and Public Cloud 58

4. Discussion 59
4.1 Strategic Recommendations 61
4.2 Summary of Contributions 62
4.3 Limitations 63
4.4 Future Work 63

5. Conclusions 65

Bibliography 67

Publications 81

List of Publications

This thesis consists of an overview and of the following publications which are referred to in the text by their Roman numerals.

I Raivio, Y., Luukkainen, S., and Juntunen, A. Open Telco: A New Business Potential. In *Proceedings of the 6th International Conference on Mobile Technology, Application & Systems (Mobility '09)*, Nice, France, pp. 2:1–2:6, September 2–4, 2009.

II Raivio, Y., Luukkainen, S., and Seppälä, S. Towards Open Telco – Business Models of API Management Providers. In *Proceedings of the 44th Hawaii International Conference on System Sciences (HICSS-44)*, Koloa, Kauai, Hawaii, pp. 1–11, January 4–7, 2011.

III Raivio, Y. and Luukkainen, S. Mobile Networks as a Two-Sided Platform – Case Open Telco. *Journal of Theoretical and Applied Electronic Commerce Research*, vol. 6, no. 2, pp. 77–89, August 2011.

IV Raivio, Y. and Dave, R. Cloud Computing in Mobile Networks – Case MVNO. In *Proceedings of the 15th International Conference on Intelligence in Next Generation Networks (ICIN 2011)*, Berlin, Germany, pp. 253–258, October 4–7, 2011.

V Raivio, Y., Mazhelis, O., Annapureddy, K., Mallavarapu, R., and Tyrväinen, P. Hybrid Cloud Architecture for Short Message Services. In *Proceedings of the 2nd International Conference on Cloud Computing and Services Science (CLOSER 2012)*, Porto, Portugal, pp. 489–500, April 18–21, 2012.

VI Paivarinta, R. and Raivio, Y. Applicability of NoSQL Databases to Mobile Networks: Case Home Location Register. In *Cloud Computing and Services Science*, I. Ivanov, M. van Sinderen, and B. Shishkov, Eds., Service Science: Research and Innovations in the Service Economy, pp. 225–242, May 2012.

Author's Contribution

Publication I: "Open Telco: A New Business Potential"

The author of this thesis proposed the original research idea, and wrote all other Sections of the paper except Section 2.

Publication II: "Towards Open Telco – Business Models of API Management Providers"

The author of this thesis wrote Sections 1, 3, 4.4, 5 and 6. In addition he provided half of the analysis in Sections 2.4 and 4.5.

Publication III: "Mobile Networks as a Two-Sided Platform – Case Open Telco"

The author of this thesis proposed the original manuscript for the article. Based on the review feedback, the author wrote all other Sections except Section 2, which was written together with the second author.

Publication IV: "Cloud Computing in Mobile Networks – Case MVNO"

The author of this thesis proposed the original research idea, and managed the work. In addition, he wrote Sections 1, 2, 5 and 6, and participated in the drawing up of the survey questions. The proof of concept was based on the ideas of the author of this thesis.

Publication V: "Hybrid Cloud Architecture for Short Message Services"

The author of this thesis provided the original research idea to find the cost optimal hybrid architecture, and managed the project where the results were created. The author wrote Sections 1, 2.1, 2.2, 5, 6 and 7, and Section 2.3 was written together with the 3rd and 4th authors. In addition, the author of this thesis compiled the whole text together and defined the parameter values used in equations of Section 4.

Publication VI: "Applicability of NoSQL Databases to Mobile Networks: Case Home Location Register"

The author of this thesis proposed the original research idea to explore the performance of cloud databases in mobile networks, and wrote Sections 1, 2.1 and 3.1. Sections 5 and 6 were written together with the first author. In addition, the author of this thesis acted as a corresponding author, compiling the whole text together based on an older conference paper, and making the revisions and updates to the original text.

List of Abbreviations

3GPP	3rd Generation Partnership Project
ACID	Atomicity, Consistency, Isolation, Durability
API	Application Programming Interface
APU	Average Profit per User
ARFIMA	Autoregressive Fractionally Integrated Moving Average
ARIMA	Autoregressive Integrated Moving Average
ARMA	Autoregressive Moving Average
ARPU	Average Revenue Per User
AS	Application Server
ASF	Apache Software Foundation
ATCA	Advanced Telecommunications Computing Architecture
AWS	Amazon Web Services
BASE	Basically Available, Soft state, Eventually consistent
BFT	Byzantine Fault Tolerance
BS	Base Station
BSS	Business Support System
CAP	Consistency, Availability and Partition tolerance
CDMA	Code Division Multi Access
CDN	Content Delivery Network
CLI	Command Line Interface
CN	Core Network
CRM	Customer Relationship Management
DaaS	Database as a Service
DVB	Digital Video Broadcasting
EC2	Elastic Compute Cloud
GFS	Google File System
GGSN	Gateway GPRS Support Node
GPRS	General Packet Radio Service

GSM	Global System for Mobile Communications
GSMA	GSM Association
GSR	Goods, Services and Revenue
HDFS	Hadoop Distributed File System
HLR	Home Location Register
HSPA	High Speed Packet Access
IaaS	Infrastructure as a Service
IMS	IP Multimedia Subsystem
IN	Intelligent Network
IPR	Intellectual Property Rights
KPI	Key Performance Indicator
LAN	Local Area Network
LBS	Location based services
LTE	Long Term Evolution
MNO	Mobile Network Operator
MSC	Mobile Switching Centre
MVNO	Mobile Virtual Network Operator
NaaS	Network as a Service
NE	Network Element
NGN	Next Generation Networks
NIST	National Institute of Standards and Technology
NoSQL	No SQL or Not only SQL
OLAP	Online Analytical Processing
OLTP	Online Transaction Processing
OSS	Operations Support System
OSS/BSS	Operations and Business Support Systems
OTT	Over The Top
P2P	Peer-to-Peer
PaaS	Platform as a Service
QoE	Quality of Experience
QoS	Quality of Service
RAN	Radio Access Network
RCS	Rich Communication Suite
RDBMS	Relational Database Management System
REST	Representational State Transfer
RNC	Radio Network Controller
RPV	Resources, Processes and Values
RSS	Really Simple Syndication

SaaS	Software as a Service
SDP	Service Delivery Platform
SEPA	Single Euro Payments Area
SGSN	Serving GPRS Support Node
SIP	Session Initiation Protocol
SLA	Service Level Agreement
SME	Small and Medium Enterprise
SMS	Short Message Service
SMSC	Short Message Service Center
SOAP	Simple Object Access Protocol
SQL	Structured Query Language
SS7	Signalling System number 7
STOF	Service, Technology, Organization and Finance
SWOT	Strength, Weakness, Opportunity and Threat
TATP	Telecommunication Application Transaction Processing
TDM	Time-Division Multiplexing
UMTS	Universal Mobile Telecommunications System
UTRAN	UMTS terrestrial RAN
VCE	Value Chain Evolution
VLR	Visitor Location Register
VM	Virtual Machine
VoIP	Voice over IP
VPC	Virtual Private Cloud
WLAN	Wireless Local Area Network
YCSB	Yahoo! Cloud Serving Benchmark

List of Figures

1.1 Operator ARPU development [Publication III, adapted from Figure 1]. 2
1.2 Research focus. 4
1.3 Growth of open APIs [50]. 5
1.4 Contributions mapped to mobile network architecture. . . . 13

2.1 The long tail of mobile services [152, adapted from Figure 1]. 20
2.2 Mobile network architecture [180, adapted from Figure 1]. . 22
2.3 Virtualization. 30
2.4 Service models [80, adapted from Figure 17]. 31
2.5 Hybrid model. 37
2.6 Demand curve [167, adapted from Figure 4.1]. 42

3.1 Open Telco architecture [Publication I, adapted from Figure 1]. 47
3.2 Impact of database size [Publication VI, Figure 5]. 50
3.3 MVNO systems in cloud [Publication V, adapted from Figure 8]. 51
3.4 Virtual broker architecture for the Open Telco [Publication II, Figure 7]. 54
3.5 Open Telco as a two-sided platform. 56
3.6 Hybrid costs [Publication V, adapted from Figure 7]. 57

List of Tables

1.1 Summary of publications. 12

2.1 TATP transactions [Publication VI, Table 1]. 25

2.2 Comparison of value networks and ecosystems. 39

3.1 Examples of open APIs and services [Publication I, Table 2]. 45

3.2 Analysis of the Open Telco as a two-sided platform [Publication III, adapted from Table 1 and Table 3]. 55

4.1 SWOT analysis. 60

1. Introduction

The mobile industry has created one of the most profitable businesses in the modern world. Both mobile operators and device manufacturers have made considerable profits with their networks and handsets. Mobile operators have utilized the incremental network evolution from 2G, e.g. Global System for Mobile Communications (GSM) and Code Division Multi Access (CDMA), to 3G for ensuring a solid foundation for their business. Recently the incumbent device vendors such as Motorola and Nokia have lost their dominant position to new challengers. The internet giants, including Apple and Google, have changed the rules of the game. The importance of the ecosystem has been underestimated by the old device vendors, enabling fast changes in the market share of companies.

However, mobile operators have not yet suffered from a major fall in income. The mobile networks involve a heavy and long investment burden, making the life of network challengers very difficult due to the high barriers to entry. Nor have new network technologies such as Wireless Local Area Network (WLAN), cable networks or Digital Video Broadcasting (DVB) provided the necessary traction or short cuts for successful challenges to be mounted. Does this all mean that mobile operators do not have any worries ahead? Unfortunately for the incumbents, this is not the case. In this dissertation the aim is to explore whether mobile operators can avoid the fate of the previous device incumbent vendors.

Although the mobile operator business case has been positive almost in all respects, there exists weak signals [116] of new threats. In the past, Average Revenue Per User (ARPU) has been one of the Key Performance Indicator (KPI) of the operator business. In most markets, however, the ARPU is decreasing and simultaneously customer churn has been high. As an example, Figure 1.1 illustrates how these key parameters have developed in TeliaSonera's, the medium size Nordic operator, Finnish mobile market during the last decade. The key problem is that the usage of data is increasing exponentially [40], but the revenues gathered from the increased data usage do not yet compensate for the decreases in voice revenues [196].

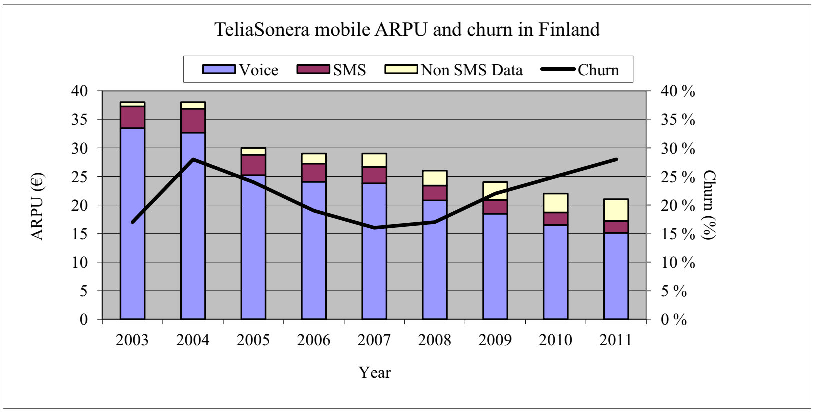

Figure 1.1. Operator ARPU development [Publication III, adapted from Figure 1].

Even a more worrying trend has been the decline in Average Profit per User (APU) [146]. These numbers are not officially reported, but dividing operating profit with the number of subscribers reveals that the trend is similar to the ARPU one. The high churn also has a direct negative impact on the APU. In the Finnish market one of the main reasons for the increase of churn in 2004 was the introduction of a number portability that enabled subscribers to easily switch operators while keeping the old number. A similar negative impact on churn was detected in 2011, when a new law forced operators to release subscribers from long term 3G agreements if the subscribers wanted to change their operator in the middle of the agreement period. New network technologies such as High Speed Packet Access (HSPA) and Long Term Evolution (LTE) also bring new profit challenges to operators [81].

The mainstream mobile services, voice and messaging, are continuously being challenged by the internet alternatives. Skype gave the first serious warning, and other similar services have been introduced. In a questionnaire by STL Partners, mobile players forecast that by 2015 voice revenues will drop by 20%, and short messaging income will shrink twice as much, by 40% due to Over The Top (OTT) competitors such Facebook, Apple, Google and Yahoo!. Operators' own alternative, Rich Communication Suite (RCS), did not receive much support either, and the responders said that RCS is "too little and too late" [176]. Data users fully utilize flat rate data packages, giving operators only limited opportunities for additional earnings. In these circumstances operators are forced to search for alternative operational models or otherwise they will – sooner or later – lose their leading role. On the other hand, mobile operators have a large number of unused assets that can be used to gain new revenues.

This dissertation introduces two complementing solutions to respond to the challenges. Both ideas are widely used by the internet companies, but for some reason they have been neglected in the telecommunication space. The first proposal, open innovation, accelerates innovativeness [120], and it has become the cornerstone of most internet success stories through open APIs. The second one, cloud computing, lowers the investment costs, optimizes the operability and improves the scalability of computing platforms [79]. Both open innovation and cloud computing can be regarded as major prerequisites of all new internet enterprises. Services such as Gmail, Facebook, MySpace, Twitter or YouTube would not have ever been born without the introduction of these ideas.

Amazon was one of the first successful adapters of open APIs and cloud computing. The company's closed business model was changed to a new one by opening the core assets to partners and even competitors. This was a disruptive strategic change, but the company's recent success has proven the decision correct. By offering open and simple access to its own data resources, Amazon created a new ecosystem that was also supporting its own core business. Moreover, Amazon managed to open a marketplace that provided both the head and the long tail items for purchasing. Amazon built its own cloud computing infrastructure to support the marketplace, but lately they have successfully offered the infrastructure for other users as well. At the same time it is good to remember that the infrastructure business, Amazon Web Services (AWS), brings less than 4% of company's total turnover[1], highlighting the fact that selling books and other merchandise is still Amazon's core business.

Amazon succeeded in transformation, although the decision to open its platform to rivals was a difficult one [61]. The question follows whether mobile operators can follow the Amazon example. Operators have launched a few interesting initiatives, but so far the success has been modest. Compared to the internet service providers, the transformation task is not easy. Operators have a huge installed base, a successful business to run

[1] `http://phx.corporate-ir.net/phoenix.zhtml?c=176060&p=irol-newsArticle&ID=1688177&highlight=`

and thus, no urgent need to change anything. However, similar actions to Amazon could be worth trying. Figure 1.2 depicts the position of operator networking resources and cloud computing in relation to developers and end users. By offering network assets operators can integrate themselves to service ecosystems. In addition, operators can utilize cloud computing in their own service platforms, too [29].

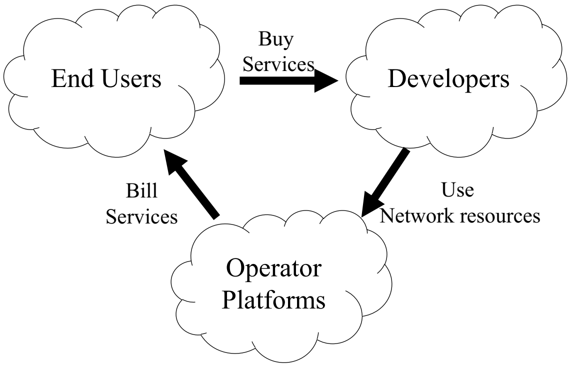

Figure 1.2. Research focus.

Open innovation was a term introduced by Chesbrough [37] to describe a development where enterprises can utilize novel ideas also from outside a company. Enterprises can use several methods to screen open innovation, but in this context we focus on the usage of open APIs. Open APIs provide a safe and well-controlled entry point to service providers for harnessing novel innovations from developer communities, bringing considerable income and support for the core business. On the Internet, open APIs have already been widely available for a long time. By the middle of 2012, the number of APIs exceeded the 6000 mark, with a growth speed close to the Moore's law, see Figure 1.3. The majority of the APIs, around 70%, of the APIs are based on the Representational State Transfer (REST) approach, while 21% support the Simple Object Access Protocol (SOAP) standard[2].

Mobile networks have been based on the closed innovation, also known as the walled garden, approach. The main idea has been that the core services are well standardized and only offered by the operators. The auxiliary services for enterprises are provided by a group of qualified partners, which use telecom APIs such as Parlay-X and web services framework. These technologies are complex and not designed for individual develop-

[2] http://www.programmableweb.com

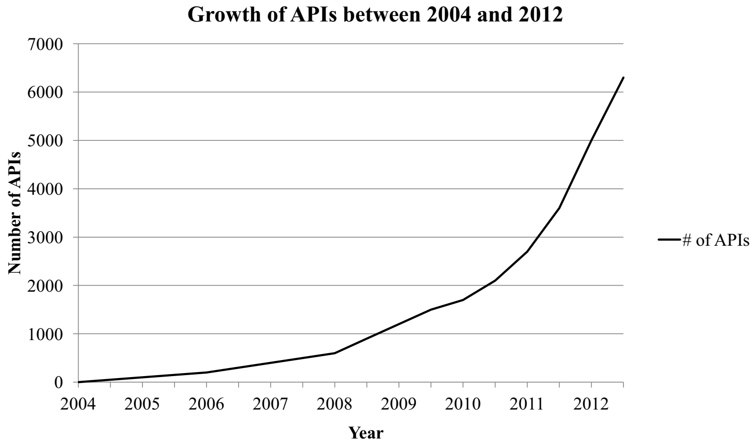

Figure 1.3. Growth of open APIs [50].

ers; thus, they do not enable open innovation. This means that mobile networks lack the long tail and are missing novel ideas. Instead, disruptive innovations are invented and developed on the Internet, bypassing the operator control. However, this progress can pose a threat to the operators. They will lose control and, simultaneously, possess only a few options to compete and earn more revenues. As a result, operators are restricted to minimizing the Capex and Opex, and competing on price, but in the end this progress will result in declining ARPU and an even worse APU.

In the Open Telco concept mobile operators open their network resources to developers, partners and even themselves. Similar ideas have been presented within initiatives such as Mobile Web 2.0 [93], Telco 2.0[3], Network as a Service (NaaS)[4] and OneAPI[5]. The key idea in all these proposals is to provide mobile network resources through well defined, easy-to-use and secure REST APIs, available to interested parties. Developers and partner companies can use APIs for novel service creation, but operators themselves can also utilize the same APIs for analyzing customer data. The basic idea of the Open Telco is simple, but unfortunately the implementation includes several challenges in respect to technology, business, security and regulation issues.

The operator market is dispersed, but a successful Open Telco initiative will require smooth cooperation across the operator boundaries. The open

[3] `http://www.stlpartners.com/telco2_index.php`
[4] `http://www.aepona.com/`
[5] `http://oneapi.gsma.com/`

API technology as such is straightforward, but bringing the actual data from systems that were originally designed to be closed is not an easy task. Besides, the business case is questionable, and it is hard to prove that the investment in the Open Telco system will ever result in a return of the money invested. Finally, telecommunication operators are forced to act according to the strictest security requirements, a comment which holds true for open APIs, too [112]. In fact, the telecommunication sector is struggling under a heavy burden of regulation [103].

Cloud computing owes its heritage to open innovation as well. Most of the key technologies behind are invented and developed by the open community. Cloud computing as a term is quite new, although already in 1969 Leonard Kleinrock [100] predicted that "computer utilities" will spread across the world. Computing has been regarded as the 5th utility in addition to water, electricity, gas and telephony [28]. Compared to its predecessors, Peer-to-Peer (P2P) computing, grid computing and clusters, cloud computing provides a novel business model that enables on-demand and fast provisioning of computing resources. Hiring one machine for a thousand hours costs the same as hiring a thousand machines for one hour. For new businesses the pay-as-you-go principle lowers the entry barrier, offering great flexibility both in the up-scaling and down-scaling phases.

Telecommunication networks are used to running their systems on dedicated and isolated computing units. Backend systems are often operating in silos that do not interwork very well. Cloud computing is based on a completely opposite approach where computers are shared by several, usually independent customers, which may also operate in different time zones, enabling a statistical multiplexing advantage [80].

Cloud computing can be utilized in mobile networks [63], too, although there are also concerns that cloud computing is not yet mature for critical services [105]. The operator business includes several areas where cloud computing can offer an integrated view of various computing needs. In all service models, including Software as a Service (SaaS), Platform as a Service (PaaS) and Infrastructure as a Service (IaaS), cloud technologies and business models can provide considerable opportunities for improvements and enhancements. As an example, the success of network sharing initiatives paves the way for the cloud computing approach. Besides, mobile networks have very variable ICT needs, changing from strict real-time services to customer care and billing, which are typical non real-time applications, and thus favourable to clouds. However, successful cloudification is not a straightforward task, but it requires careful planning [77, 97].

It is clear that cloud computing has several challenges ahead in the mobile domain. Lowenstein [114] underlines the fact that mobiles are not ready for clouds and vice versa. The continuous connectivity, pricing and usability are the issues to be solved first. On the other hand, telecom vendors have warned that cloud computing does not meet the Service Level Agreement (SLA) requirements for which the mobile networks are designed [131]. It can be expected that operators will favour so-called telco clouds [185] that are tailored to meet telecom requirements and are provided by trusted vendors. The current market leaders, such as Amazon Elastic Compute Cloud (EC2) and Rackspace Cloud Servers, spur the telecom options and bring valuable experiences for the purposes of comparison. In addition to public clouds, operators can simultaneously utilize private clouds, leading to a hybrid cloud setup.

Looking forward, cloud computing may just be an intermediate step to the next phase called utility computing [28], a term also known as sky computing [96]. Utility computing, which was already foreseen as early as in 1969, can add P2P networking to cloud computing, enabling a fully virtualized and globally accessible computer network. According to this vision, computing tasks can be run on any computer. In current clouds interoperability is already a major challenge, and customers are suffering from a vendor lock-in problem. De facto cloud access standards are available, but most cloud resources are still aligned to technology silos. These problems will most probably disappear in the long run, transforming the distributed computing units into a harmonized marketplace.

1.1 Research Questions and Scope

This dissertation reviews how open innovation and cloud computing can be applied to mobile networks. These initiatives are complementary but also partly independent, meaning that operators can introduce them in phases. Open innovation and cloud computing are still quite novel terms for the mobile industry, and very little related public research has been conducted before. Mobile vendors have already analyzed new opportunities for years, but until today new initiatives have been rare, partly due to a stable and profitable financial outlook. However, keeping the weak signals in mind, it can be expected that in the near future the innovation cycle will accelerate, and the mobile networks will become mature for new ideas.

Open APIs can be utilized both in networks and terminals. Terminal APIs are already widely available, but unfortunately they are still vendor or operating system dependent. Besides, in several cases network APIs can provide a more economic choice as well as coherent access to operator terminals, regardless of device capabilities and manufacturers. A drawback is that markets can become fragmented, not per terminal type but across the operator boundaries. However, network and terminal APIs can complement and support each other in creating a globally harmonized developer market and ecosystem for developers and customers. In this context we will focus on the mobile networks only, excluding the mobile terminals and associated innovation ideas, including marketplaces such as Google Play, Apple App Store and Nokia Store.

Originally the following research questions were drawn up to address how open innovation and cloud computing can support the operator business case.

Q1: Can mobile networks benefit from open innovation and especially open APIs, referred to here by the term Open Telco? Open APIs have already been offered for several years by individual operators. However, the success of the open API initiatives and trials has been modest. There exists several reasons for the slow start-up, and the main challenges and solutions are reviewed in this dissertation.

Q2: How can mobile networks utilize the opportunities of cloud computing in infrastructure, platform and software delivery computing units? Mobile networks have usually been built on dedicated and tailored computing systems that are expensive. Lately more standard solutions conforming to the Advanced Telecommunications Computing Architecture (ATCA) standard have become common. As a next step, mobile networks can also utilize cloud computing utilities in the computation tasks. The research explores how SaaS, PaaS and IaaS service models adapt to different technology and business needs in mobile networks.

1.2 Research Methodology

Research of social sciences can be divided into exploratory, descriptive and explanatory strategies. According to the standard view, case studies are especially suitable for exploring new phenomena, surveys and history data can be used to describe the research object and finally, experiments are applied to explain the new topics. However, this structure is not rigid, but different research strategies can be applied to different research phases [199]. This work uses all the above research methodologies, which are discussed in detail in the following subsections.

1.2.1 Case Study

A case study is a popular research methodology for evaluating new phenomena that do not exist yet or which cannot be otherwise emulated or simulated [52]. Case studies can be based on single- or multiple-case approaches. The single-case study, which is also known as intrinsic case study [174], analyzes one case in depth using a holistic design and explanatory method with a single-unit of analysis. The single-case research includes five different types, namely critical, unique, representative, revelatory or longitudinal cases.

A multiple-case study improves the validity of analysis, because a replication logic can be used to test the generalization of the results. Yin [199] compares experiments with case studies, where a single-case design refers to single and a multiple-case design to multiple experiments respectively. If similar results are obtained in several cases, a replication has taken place, and the findings have an expanded external generalizability compared to inferences from a single-case design. The case study method has received criticism and several researchers do not value the results [199]. However, the case study has its advantages, and when the limitations of the research are acknowledged, results can be as valuable as the results of experiments or surveys.

Case studies are well suited for the research of the Open Telco phenomenon, because the concept is novel and it consists of unknown factors. This dissertation uses case studies in two publications. First of all, the business models of API management providers in the Open Telco context were analyzed in Publication II, applying multiple-case studies, where history data was partly used as a source of analysis. Secondly, Publication III continued the Open Telco research by analyzing whether the Open Telco creates a two-sided platform. The research was conducted using a representative single-case approach, to where the research framework was deduced from the literature review.

1.2.2 Survey

The survey is a useful research method to analyze a new phenomenon that involves people. More specifically, Pinsonneault and Kraemer [147] identified that survey research can be applied to evaluate relationships of the people quantitatively. Furthermore, survey research uses a selected sub-group of the population from which the findings can later be generalized back to the whole population. More formally, survey research is used [91] to evaluate what exists, future comparisons and trends. Glasow [67] continues that in survey research, variables, which are beyond the control of the researcher, are used to define the scope of study. Furthermore, the method is cost efficient when gathering a large variety of samples from a large population.

At the same time the limitations of surveys have to be kept in mind. The results are only estimates of the whole population, not exact measurements [161]. Pinsonneault and Kraemer [147] note that surveys are not suitable for cases where historical background of the context of the phenomena is required. Lastly, Bell [19] warns that surveys are vulnerable to errors due to lack of respondents, misreporting of results or novel phenomena that respondents cannot fully understand.

The applicability of cloud computing to business support systems of mobile networks was analyzed in Publication IV with a survey methodology that utilizes the Delphi method [159], which is a flexible research technique that can be applied both to quantitative and qualitative research. The Delphi method has four prerequisites: anonymity, iteration, controlled feedback and statistical aggregation [171]. This research fulfills all those conditions. The target group consisted of nine experts, who were first personally interviewed, without revealing their identity to other responders. The answers were analyzed, and the aggregated results, followed by more focused second round questions were distributed. The second round was carried through as a web questionnaire, which dropped the answer rate to 66%. However, because the answers already saturated and the research questions were answered, no more survey rounds were required and process was stopped. Finally, the answers were quantitatively and qualitatively analyzed.

1.2.3 Experiment

There are several research methods that are suitable to evaluate a software code in computer science. Depending on the available resources and time, code can be reviewed using formal methods, simulation, emulation, experimental or software development. The first two methods in the list are cheaper and leave more control to a researcher, but they are more theoretical and far from real world conditions. On the other hand, real software development in the context of mobile networks is really challenging due to complicate architecture, large number of specifications and implementation of the Signalling System number 7 (SS7) [129].

However, emulation and experimental development provide a compromise between a theoretical and practical approach, offering a verifiable, reliable, repeatable and reproducible research method. Experiments are most applicable to the explanatory research phase in order to answer research questions like ”how” and ”why”. Experiments can be used in which a researcher can adjust research parameters, for example, in the laboratory environment where an experiment can be focused on certain variables. The novel test results can be compared to measurements made with old systems, and in this way new solutions can be evaluated.

For previous reasons emulations and experiments were carried out in Publication V and Publication VI to explain and clarify cloud computing opportunities in the service and infrastructure service models. The chosen services, text messaging and database services were implemented using emulators, and the performance of test systems were evaluated in the cloud environment. This research was complemented with formal methods to verify the cost analysis of the systems.

1.3 Research Framework

The Service, Technology, Organization and Finance (STOF) business model framework [75] offers a complete toolbox for analyzing novel service innovations. Although the model was initially designed to analyze mobile services, it can be applied to any digital services. The framework consists of four interrelated domains: service, technology, organization, and finance. According to Bouwman et al. [24, Chapter 2], the service domain and the customer value of the service are the key issues of the model, while technology only acts as an enabler for novel services. The organization domain explains the roles of the value chain, and finally, the finance domain evaluates prices and revenue models. All domains have important relations to each other, meaning that all domains are required to create a complete view of a business model. The STOF model supports business model dynamics as well, because the model can be iterated in different product phases. Besides, the STOF method provides a a-four-step mechanism for designing new business models [24, Chapter 5].

Unlike some other business model frameworks such as Osterwalder's business model canvas [137], the STOF method takes into account techno-economic interdependencies, which are the key aspects of the Open Telco research. Due to the comprehensive approach, set of tools and specific applicability to mobile services, the STOF model was selected for the basic research framework. The Open Telco concept was first evaluated in Publication I using the STOF model, and later it was utilized also in Publication II to review the individual cases. Finally, the STOF categorization was applied to Chapter 2 and Chapter 3 for grouping the background and related results.

1.4 Contributions

This dissertation includes six publications, including four conference papers, one journal article and one article in collection. The detailed contributions can be found in Chapter 3. Figure 1.4 presents a high level vision of how the cloud computing stacks and the Open Telco concept map to the mobile core network. The Figure also shows the areas where to the attached publications are contributing.

The attached publications address different research questions and utilize various research methodologies and frameworks. Table 1.1 summarizes the main contributions of each publication relating to research questions, research approach and main content.

Table 1.1. Summary of publications.

Publication	Question	Method	Main content
PI	Q1	STOF	Open Telco concept
PII	Q1	Case study, STOF	Business case
PIII	Q1	Literature, Case study	Two-sided platform
PIV	Q2	Survey, Experiment	PaaS, Cloud MVNO
PV	Q2	Experiment	SaaS, Hybrid cloud
PVI	Q2	Experiment	IaaS, NoSQL

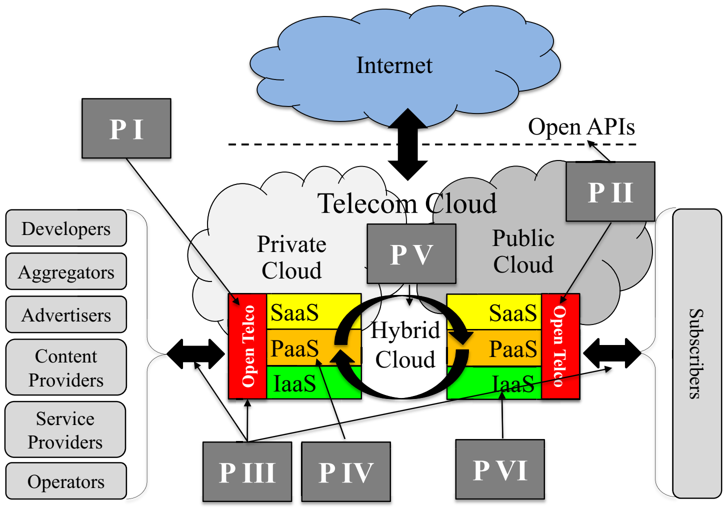

Figure 1.4. Contributions mapped to mobile network architecture.

In summary, the following contributions were made:

Publication I reviews the critical factors of the Open Telco concept using the STOF model. The work was based on the author's own ideas from his earlier career.

Publication II evaluates the business case of the broker concept required in the successful Open Telco concept. The analysis is based on multiple-case analysis, supported by the STOF model, using three internet samples and one mobile open API trial.

Publication III analyzes whether the Open Telco concept is able to set up a two-sided platform. Single-case analysis is applied to a representative usage scenario that was selected from a list of candidates by an expert group. The chosen scenario, called Event Experience, was implemented and reviewed, and critical aspects for the two-sided platform creation were listed and analyzed.

Publication IV uses survey research methodology to review whether mobile networks, and especially Mobile Virtual Network Operators (MVNOs), can utilize cloud computing principles in their backend, namely Operations and Business Support Systems (OSS/BSS). The publication also proposes a billing-as-a-service trial system that is based on a hybrid cloud PaaS system.

Publication V continues the hybrid cloud research, but now the main focus is on the service layer. The use case is taken from the mobile industry where the Short Message Service Center (SMSC) is a key element of the successful text messaging service. The publication uses emulator software for investigating how a hybrid cloud can optimize the computation resources both in performance and cost.

Publication VI evaluates how cloud computing can meet the strictest SLA requirements of the mobile networks. The use case is a Home Location Register (HLR) that is the most critical single point of failure in the network. The test environment is built using a HLR benchmark tool that is tailored to the No SQL or Not only SQL (NoSQL) database and Amazon EC2.

1.5 Structure of the Dissertation

Chapter 2 introduces the background and latest development within the research focus. Chapter 3 summarizes the main results of the attached publications, and Chapter 4 discusses the strategic recommendations, summary of contributions, research limitations and future research ideas. Finally, Chapter 5 concludes the dissertation.

2. Background

This chapter presents background for the dissertation. It is divided to four sections, following the STOF structure: service, technology, organization and finance. The first section describes the main principles of service innovation, starting from service definition and innovation strategy, continuing with open innovation phenomenon and ending to challenges in mobile services. The technology section addresses the key technology enablers of mobile networks and cloud computing, followed by the organization section, which describes theories on value networks, ecosystems, two-sidedness and broker models. The chapter is concluded with a finance section, which outlines consumer theory and internet revenue models, in addition to cloud cost, market and ARPU break-even analysis.

2.1 Service

Products cover both goods and services. Traditionally, services differ from physical goods due to their intangible nature. A service has various definitions depending on the author and background. A few of the most referred to are the one by Quinn et al. [151, p. 50]:

> ”Most authorities consider the services sector to include all economic activities whose output is not a physical product or construction, is generally consumed at the time it is produced, and provides added value in forms (such as convenience, amusement, timeliness, comfort, or health) that are essentially intangible concerns of its first purchaser.”

and by Grönroos [73, p. 46]:

> ”A process consisting of a series of more or less tangible activities that normally, but not necessarily, take place in interactions between the customer and service employees and/or physical resources or goods systems of the service provider, which are provided as solutions to customer problems.”

These definitions are adequate in this context. Furthermore, customers more or less participate in the service creation process. Because services require a considerable amount of human activity, they rarely adhere to a predefined process. Services are perceived as the output of a process as well as a process itself. As a summary, products contain parts from both goods and services [24]. According to Zeithaml et al. [201], services can be identified according to five characteristics: intangibility, inseparability, perishability, heterogeneity and ownership. Services are intangible, which means they cannot be touched. Secondly, services are produced and consumed at the same time. Thirdly, services are perishable, underlining that they cannot be stored. Fourthly, services have a heterogeneous nature, meaning that they change all the time. Finally, services cannot be owned, but usually a consumer has merely a usufruct, or a right to use the service for a certain time.

2.1.1 Innovation Strategy

Innovation is surely one of the most popular words in the current leadership vocabulary. Schumpeter [166] mentioned as early as 1934 that innovation is an important accelerator of new technology cycles. He identified five different innovation types: new products, new methods for production, new sources of supply, exploration of new markets and new ways to organize business. In modern literature innovation can refer, not only to technological, but also to organizational and institutional innovation. Hekkert et al. [83] note that these innovation types and their complex and dynamic interactions form the basis of the innovation process. According to Chesbrough [37], however, innovation is changing towards open innovation from a closed, centrally-lead and mainly technology-driven approach. [24].

According to Chesbrough [37, pp. 13–14], former IBM research director James McGroddy compares market changes to chess and poker. McGroddy says that companies play chess in a stable market, but they switch to poker when markets change. Chess rules are clear and strategic decisions can be made based on available knowledge, while in poker correct decisions are harder to make, because you must take into account an opponent's possible bluff. The same thinking applies to innovation strategy, too. Incremental innovations need a different strategy than disruptive innovations do.

Five Forces Analysis

A good vision without a good strategy and implementation is worth nothing, and this is true in the internet business as well. Porter created the Five Forces analysis in 1980, but he later applied the theory to the Internet as well [148]. The paper highlights the fact that the Internet is a complementary delivery channel that can effectively support existing trading methods. In this respect, the Internet should not be regarded as part of the new economy, because the old business laws are also valid in this case. The Internet does not inevitably guarantee economies of scale and furthermore financial success, but finally the well-being depends on the correct strategy. The Five Forces analysis is as valid for the internet business as for any kind of traditional trading. However, Hamilton [78] reminds that the Internet offers tools and competitiveness models for disruptive innovations to explore novel strategic actions and to deliver personalized services to individual customers.

Porter continues that the bargaining power of various channels is weakening, but in general, from an enterprise's point of view, most of the forces are negative. The Internet offers consumers new methods to compare prices over wider geographic areas and to select the cheapest ones. In this respect, nation-wide regulatory rules do not support internet trading very well. The threat of substitutes is apparent, at least in cases where the price is much cheaper. E-commerce empowers the competition, prices fall and tend to converge. The dominant market position gives a competitive edge to incumbents, making entry for newcomers extremely difficult. Competition with existing competitors will get fiercer, and will make it difficult to differentiate the competitors from each other in the eyes of the consumer. Any good idea can be copied immediately. [148]

Disruptive Innovation

Schilling [165] defines that dominant designs and incremental changes converge due to technological discontinuities towards an era of ferment. She classifies innovations to competence enhancing and destroying. Disruptive innovations are typically simple, cheap and revolutionary. According to Christensen [39], new market entrants can utilize either low-end disruption or new market disruption strategies. The first alternative fulfills the needs of low-end users, which strategy helps to gain a market position. The second alternative attracts customers who are looking for product features that are not offered by the current incumbents.

Disruptive innovations can replace and cannibalize older services. Popular internet services such as instant messaging, microblogging and Voice over IP (VoIP) can impact on Short Message Service (SMS) and circuit switched voice revenues. McGrath [123] separates cannibalization into two categories: positive and negative. Companies make mistakes because they do not know when they should avoid or accept cannibalization. Cannibalization can be harmful, especially for the market leader, if the new product brings less profit or incorporates unfavourable economics. But cannibalization can also be positive. For a challenger operator a radical innovation with an advanced technology offers an opportunity to attack the incumbent. Also the market leader can utilize cannibalization. They can use the pricing tool and frequently update their platforms or specific product segments. The main risk for cannibalization is the timing. For the correct decision, companies must utilize an analytical framework with a sensitivity analysis to determine the critical break-even points.

2.1.2 Open Innovation

Open innovation [37] and crowdsourcing [87] are closely related terms, which have been widely used in the modern business literature. Both terms inherit their reasoning from a sentence written by Chesbrough [37, p. xxvi]:

> ”Not all the smart people work for us. We need to work with smart people inside and outside our company.”

Open innovation means that developers, residing outside of a company, openly utilize the resources provided by a company. The resources can be offered through open APIs, which companies can fully control. However, the companies cannot control the outcome of services that utilize the APIs [37]. Crowdsourcing can also benefit from open APIs, but in this context companies have tighter control of the final result [87].

Grodal [72] warns that closed networks will reduce innovativeness. Open innovation promotes more free exchange of Intellectual Property Rights (IPR), and if the optimum of internal and external ideas can be found, a win-win situation can be established. Besides, openness makes it easier to test new ideas. According to Thomke [182] and Gaynor [64], the success of new innovations is difficult to predict. For that reason the best way to develop new services happens through experimentation, by successive approximations. They also pave the way for accidental innovations, including examples such as Flickr and mobile text messaging. Often a better strategy is to try and then fail quickly than to use long term planning.

However, enterprises should first carefully evaluate which assets they open. Boudreau and Lakhani [23] propose a three-step review process to help in making the correct decision. Firstly, decide which innovations will be shifted to external developers and how the developers are approached, either through collaborative communities or competitive markets. Secondly, clarify the motivations of those developers, whether they are extrinsic or intrinsic, and thirdly, review the business model of the developer platform. Boudreau and Lakhani present three different platform business models, that is, integrator, product and two-sided platform, defined by the question ”who sells to whom”. The first alternative offers the highest control for the platform provider, while the two-sided platform provides the other extreme, e.g. the highest autonomy for the external developers.

Open innovation is, however, challenged on the basis that it endangers the first mover advantages [165]. On the other hand, early adopters and even late entrants can learn from the first mover mistakes, and thus create better products. Chesbrough [37] continues that a better business model is more important than being the first on the market. Microsoft CEO Steve Balmer recently said[1]: ”Being fast is not always the right answer, being right is always the right answer”. In addition, service providers must take into account privacy concerns [11]. The definition of openness is not always clear either. Jaokar and Fish [93] underline that, for example, AOL and Microsoft utilize open standards such as Really Simple Syndication (RSS) and Session Initiation Protocol (SIP), although these companies in several cases act on behalf of the closed strategy.

2.1.3 Mashups

The Web 2.0 offers services with mashups that utilize internet resources through public APIs [200]. As a result, web developers have begun to develop software applications that merge separate APIs and data sources into one integrated interface, called a mashup. By definition, anyone can create a mashup. Mashups are not anymore designed to the fixed internet only, but they can be applied to mobile internet as well. Mobile operators can also compete in the mashup space. O'Reilly [135] argues that mobile

[1] http://news.in.msn.com/exclusives/it/article.aspx?cp-documentid=250143320

operators should offer end users' history data through open APIs.

Stecca and Maresca [175] divide mashups to three categories: event, map and data. Mashups can be executed on the client or server side. Server side execution may be installed to cloud on PaaS to save Capex, Opex and energy. Salminen et al. [162] propose a mobile multimedia mash-up ecosystem that consists of home, mobile and cloud domains. The cloud domain includes an aggregator that combines various APIs to be used by clients. The cloud infrastructure has, however, a few challenges relating to networking, security and servers. The aggregator must support continuous changes in APIs and also scale to variable loads. REST-type APIs are proposed to tackle the complexity challenges. In addition to server side technologies, the system must meet the legal requirements in licensing and terms of service.

2.1.4 The Long Tail

Several natural phenomena follow from the so-called Zipf function or the Power law. In its simplest form it means that the quantity of an item is inversely relational to the rank of it. Also in the Internet, the Power law seems to be the rule rather than an exception [4]. Anderson [8] came up with the idea of applying the long tail phenomenon for online services. He found that a wider selection lengthens and fattens the tail. An example of the long tail in the context of mobile services is presented in Figure 2.1.

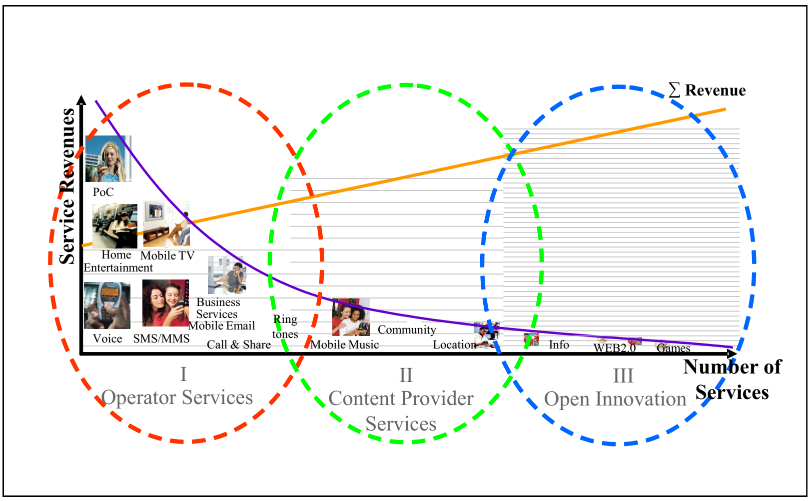

Figure 2.1. The long tail of mobile services [152, adapted from Figure 1].

The long tail drives the consumption of certain products and services from hits into niches by bringing previously unknown products and services to public knowledge. For example, a typical large music store can contain up to 25 000 titles for sale, while an online store can offer millions of songs. The tail may also be absent due to a non-existing sales channel. Mutanen [132] calls this type of market the invisible tail.

Verkasalo [188] has identified tails with mobile browsing data, while Jaokar and Fish [93] and Raivio [152] apply the long tail also to mobile applications. In that context operators provide the mass market services, such as voice and SMS, while the long tail consists of the niche services created by the open community. Kilkki [99] has found that the long tail curve can be drafted even with limited source information.

2.1.5 Challenges in Mobile Services

Maitland et al. [115] note that new telecom services put pressure on value chains, corporate resources and skills, economics and patterns of regulation. Kuo and Yu [102] continue with the following challenges: correct business and investment strategy, limitations in mobile devices, constraints in networks and infrastructure, security concerns, decreasing ARPU and user distrust. Mulligan [130] concludes the list with business instability, changing economies of scale, investment requirements and the migration of mass media into telecommunication networks. She also addresses the fact that there is clear link between open APIs, developers and speed of innovation.

The media convergence challenges the current regulation pattern. On the other hand, Reuver et al. [46] claim that regulation does not have a major role in driving changes on business models unlike technology and market forces. Gaynor [64] claims that the network operators cannot support all customer segments. Sánchez and Rodríguez [163] underline the importance of user-centred innovation and social trends that are coming into the mobile networks, bringing the end users into a controlling position as content providers.

Location based services (LBS) [113] are one of the most promising, but at the same time also critical, new services available through open APIs. According to studies [190, 168], the lack of a proper and simple privacy enforcement system is one of the biggest concerns among mobile users in expanding the usage of LBS. Location services are usually used in conjunction with other services such as social media and advertising, which makes the implementation of a well-controlled service even more challenging.

2.2 Technology

This section elaborates the key technologies behind the research. The review is started with an overview of mobile networks and their dedicated requirements. Next, the major theories of distributed computing are highlighted, and finally, cloud computing principles are described.

2.2.1 Mobile Networks

The current mobile networks have evolved from 2G to 3G within the period of twenty years. The starting point was the GSM network defined early 1990s. The 3G [1] and Next Generation Networks (NGN) [56] architecture specifications have inherited much from the GSM model on the core network side. Both the circuit and packet switched network elements are called the same in 2G as in 3G, including the Mobile Switching Centre (MSC) and the General Packet Radio Service (GPRS) elements Serving GPRS Support Node (SGSN) and Gateway GPRS Support Node (GGSN), but the data rates in 3G mode are capable of supporting data rates of up to 2 Mbit/s. The core network elements are connected to the 3G radio network via the Iu interface, which is very similar to the A-interface used in GSM [180]. Figure 2.2 presents the combined 2G and 3G mobile architecture.

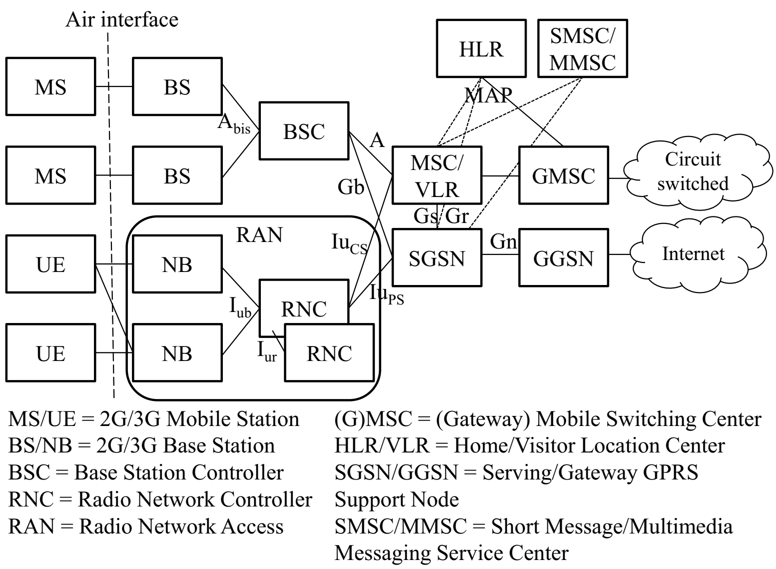

Figure 2.2. Mobile network architecture [180, adapted from Figure 1].

The major changes in the new architecture are in the Radio Access Network (RAN), which is also called UMTS terrestrial RAN (UTRAN). There is a new interface called Iur, which connects two neighbouring Radio Network Controllers (RNCs) together. This interface is used for macro-diversity combining, which is a new WCDMA-based function implemented in the RNC. Base Stations (BSs) are connected to the RNC via the Iub interface. Throughout the standardization process, extra effort has been made to ensure that most of the 2G core elements can smoothly support both generations, and potential changes are as minimal as possible. A good backward compatibility is one of the key targets in mobile networks [180].

The development of the network has not stopped here. The old architecture supports even higher bandwidths. Different versions of HSPA reach speeds of up to 50 Mbit/s. On the Core Network (CN) side an innovation called MSC server separates the control and signalling planes from each other, enabling new transport technologies and cost reductions. The next step in development, 4G or LTE will change the architecture completely by simplifying the whole design and basing the whole transport technology on top of a packet network.

Voice

Voice is the most important mobile service, bringing in the majority of the operator income, as shown in Figure 1.1. The mobile speech services are based on Time-Division Multiplexing (TDM) technologies, where a 20 ms time slot is applied to carry voice samples over the mobile network. The call control is provided by MSCs that manage voice connections in both 2G and 3G mobile networks. The TDM technology provides a highly optimal solution for fixed and small voice packets. The solution does not have major performance problems because the delay variations are small and the delay budget is carefully controlled. On the other hand, the maximum capacity is prefixed, and for that reason sudden traffic peaks will inevitably result in challenges. Besides, the existing voice systems are rigid and will not support auxiliary, internet style services such as sharing of presence status, group calls or in call messaging.

For those reasons 3rd Generation Partnership Project (3GPP) started already early in 2000 the work to design a new voice machinery that would respond to the challenges created by internet competitors. The new system was built on the SIP and the call control was provided by the IP

Multimedia Subsystem (IMS). The new concept offers several additional services, and it is also extendible for future innovations. IMS was already designed for the 3G networks, but the success of VoIP technology has so far been minor. However, IMS will be the de facto speech solution in the 4G networks where circuit switched TDM networks are not anymore available. In addition to novel services, SIP based voice systems enable autoscaling functions that can also support sudden traffic peaks [187]. In this case voice quality can become a problem, because server virtualization increases overhead and challenges to real-time applications, but voice quality can be significantly improved by the usage of virtual RAM [109].

Messaging

Text messaging is the second most popular mobile service after voice. The key element of the text messaging system in mobile networks is the SMSC. It is connected to the MSC, the Visitor Location Register (VLR), the HLR, and further to the Application Servers (ASs) and the other SMSCs. Text messages are first sent to the closest MSC, which locates the home SMSC where the message is to be forwarded. Depending on the content and address of the message, it is sent either to the next SMSC, to the AS or – if the receiver belongs to the same SMSC domain – directly to the receiver. Basically SMSC acts as a store-and-forward router. If the SMSC is overloaded, the MSC will store the messages and will resend them to the SMSC later. In the case when the message receiver suddenly is outside the network, the SMSC will store the message and send it to the receiver the moment it joins the network again [129, pp. 557-558].

Text messaging is still a very popular service, although other competitive communication methods such mobile email and social media, have become widespread. According to ITU-D [92], 6 trillion SMS messages were sent in 2010. Traffic peaks are not anymore as common as they used to be during festivals or emergencies, but still today the load can increase up to tenfold compared to usual base load [202].

Home Location Register

The HLR is one of the most critical core network elements in the mobile architecture. It is responsible for subscriber authentication and roaming functionalities. HLR also incorporates a risk of single point of failure, resulting in very high SLA requirements. The 3GPP has defined the HLR performance requirements in their general specification [2].

Telecommunication Application Transaction Processing

The Telecommunication Application Transaction Processing (TATP) benchmark tool is designed to measure the performance of a database under load, which is typical in telecommunication applications. In particular, it is modelled after the type of queries that are processed in HLR on a GSM network [178, 90]. The TATP tool encompasses seven different transactions, of which three are reads and four are writes. The description gives probabilities at which each of the transactions is executed in the client. Broadly, 80% are reads and 20% are writes. Table 2.1 summarizes the transactions, their types, distribution and effected tables based on information given by Strandell [178, Table 2].

Table 2.1. TATP transactions [Publication VI, Table 1].

Transaction name	Type	%	Tables
Get-Subscriber-Data	Read	35	Subscriber
Get-New-Destination	Read	10	Special Facility, Call Forwarding
Get-Access-Data	Read	35	AccessInfo
Update-Subscriber-Data	Write	2	Subscriber, Special Facility
Update-Location	Write	14	Subscriber
Insert-Call-Forwarding	Write	2	Call Forwarding
Delete-Call-Forwarding	Write	2	Call Forwarding

Operations and Business Support Systems

Operation Sub-System or Operations Support System (OSS) is not part of the GSM specifications [129, p. 105]. The 3GPP has not specified the support systems for the 3G networks either, but each operator and vendor can themselves define their own OSS/BSS systems. However, recently most operators have realized that tailored OSS/BSS systems are not anymore required, but they prefer to concentrate on their core businesses, including network and customer care. For that reason they have mainly outsourced OSS/BSS software development and maintenance to the service providers, software vendors and integrators [184]. In addition, the fact that the OSS/BSS market has matured and products have commoditized has paved the way for usage of open source software and cloud service models [62].

Mobile Virtual Network Operator

A MVNO buys network capacity from a Mobile Network Operator (MNO), and pays for resources based on a pay-per-use principle. A MVNO has its own billing and marketing functions and focus on selling services to customers. MVNOs usually compete on price rather than service differentiation. They can work across multiple geographical locations, too, and thus have location based differentiation [30].

A true or full MVNO owns HLR as well as switching and Intelligent Network (IN) platforms in addition to services, billing and marketing systems [18]. A true MVNO has control of service and tariff design, service implementation, marketing and differentiation, as well as branding. In contrast, a weak or light MVNO owns billing, customer care and marketing systems, while HLR, switching and service components are shared with an MNO. In this case the MVNO controls service and tariff design, service marketing, as well as branding. The reseller provides network services by implementing marketing and branding only. It sells prepackaged services that are outsourced from MNOs. Usually resellers own just marketing, branding and, in some cases, customer management integrated with other operators [173].

2.2.2 Distributed Computing

Distributed computing has a long history in computer science. Already the predecessor of Internet, Arpanet, and the introduction of Local Area Network (LAN) such as Ethernet paved the way for distributed computing. There are several different forms of distributed computing such as mainframe, client-server, peer-to-peer, grid, cloud and utility computing. Mainframes were popular until the mid 1990s, but work stations replaced the large mainframes in late 1990s. P2P technologies became familiar with the introduction of Skype and also several content sharing services. Grid computing attracted numerous researchers in the early 2000s, but commercial success did not ever happen. Cloud computing filled this gap with a very similar technology, but the business idea was better drafted. Finally, utility computing can be seen as a future that will improve the interoperability of clouds and makes computation the 5th utility, similarly to water, electricity, gas and telephony [28].

Fallacies of Distributed Computing

The research on distributed computing includes various topics. The fallacies of distributed computing, originally listed by Peter Deutsch in 1994 and complemented by James Gosling in 1997, offers a good starting point for research work [68, 158].

1. The network is reliable.
2. Latency is zero.
3. Bandwidth is infinite.
4. The network is secure.
5. Topology doesn't change.
6. There is one administrator.
7. Transport cost is zero.
8. The network is homogeneous.

The fallacies of distributed computing apply to grid and cloud computing as well as to mobile networks [140]. In the real networks, especially mobile ones, most – if not all – of the statements are false. Bit errors happen every now and then, the transmission delay can vary significantly, and especially bandwidth in mobile networks is scarce. Distributed computing is very vulnerable to security risks [95]. The network management is a continuous concern due to routing updates and different management procedures. Transmission seems to be cheap, but large file transfers are costly and actually slow to transmit. Finally, the introduction of mobile terminals has made the network fragmented.

In this research, the main emphasis is on the first three fallacies. We focus on the network SLA measurements, but in more detail on throughput and response times that are usually part of the SLA [142]. Other network SLA parameters such as availability, jitter and sustainability [15, 33] are beyond the scope of this research due to their extended requirements on research time and resources. Relating to other fallacies of distributed computing, security is a major research topic and thus worth its own thesis. The same comment applies to the network topology, management, cost or structure. Instead, here the focus is on server side computing performance [84] and costs [118].

There are also other fallacies relating to cloud computing. Sankaralingam and Arpaci-Dusseau [164] list the following ones: software is driven by hardware, multicores will be everywhere and multicore hardware implies parallelism is exposed to all developers. They conclude that using parallel hardware to execute concurrent and unrelated tasks will solve the cloud computing challenges in performance, energy-efficiency and cost.

CAP Theorem

One of the most relevant theories in distributed data services has been the Consistency, Availability and Partition tolerance (CAP) theorem. It was originally introduced by Eric Brewer in 2000 [26], and proven in 2004 by Gilbert and Lynch [66]. The main idea of the theorem says that it is impossible to optimize the distributed system in all three aspects at the same time.

Firstly, data operations are consistent if they can be linearized, which means that clients will not detect any difference in the state of the data between single and distributed systems. Secondly, availability defines the probability that each data node responds to a client request in a certain time interval. Finally, the partition tolerance defines how well the system can recover from a message loss or data corruption.

In cloud computing systems, Structured Query Language (SQL) databases are often replaced by NoSQL solutions that support more relaxed design approaches than the SQL ones. Stonebraker [177] has claimed that the NoSQL systems have dropped the consistence requirement in favour of availability and partition. Robinson [156] counters that actually the partition is fixed, and designers can select either C or A. Although CAP is originally a database theory, similar ideas can be adapted to other factors as well. Vajda [185] applies the theorem to cost of computation, networking and storage. He says that all three parameters cannot be simultaneously optimized.

Byzantine Fault Tolerance

Although the distributed systems can be made highly reliable, a few threats cannot be fully avoided. An internal malicious attack, operator mistakes or software errors are common problems that further create arbitrary behaviour, also known as Byzantine faults [31]. There errors have become more common because society depends on ICT systems that have grown increasingly complex, leading to higher risk of software and operability errors. On the other hand, open connectivity through the Internet has exposed the systems to malicious attacks [31]. Byzantine Fault Tolerance (BFT) has been questioned in the cloud computing context [20]. However, utility computing, consisting of multiple cloud providers, can increase the need for BFT measures [192].

2.2.3 Cloud Computing

Cloud computing is one form of distributed computing in addition to client-server, peer-to-peer, grid and utility computing. The underlying concept of cloud computing originates from 1950s when scientist Herb Grosch said that the world computing needs can be supported by dumb terminals and

15 data centres. Although the inventor of word ”cloud computing” is unknown, ”clouds” were often used in the early 2000s in telecommunication network architectures to present a high level abstraction of complex networks. Especially ideas about IP or ATM based mobile networks [180] were fully based on clouds. Frameworks such as All IP or Full IP used IP clouds to connect BSs to network servers and other networks.

The ICT industry was introduced to the term cloud computing again in 2006 when Amazon EC2 started to offer computing capacity as a cloud. After that the success of cloud computing has been phenomenal. In addition to service providers, the cloud providers including Amazon, Google, Microsoft, Rackspace and Salesforce are key players in the cloud ecosystem. These companies have their special products and strategies in the cloud space, but all of them offer tools and resources for distributed computing.

Definition and Architecture

Cloud computing does not have any official single definition, but the one by U.S. National Institute of Standards and Technology (NIST) is often referred to [124, p. 2]:

> ”Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.”

The key words in the definition are on-demand, shared, rapidly provisioned and minimal management. Relating to on-demand, the novel IT business model called pay-per-use has also increased the popularity of cloud computing, especially among startups. Enterprises do not need to invest anymore in computing resources that soon become outdated.

Virtualization

One of the key technology concepts behind cloud computing is virtualization [45]. It separates computation from the operating system and hardware used. This functionality improves utilization, simplifies management, lowers costs and saves energy. Virtualization requires that hardware is virtualization enabled and, in addition, a piece of software called a hypervisor is needed between the virtualized infrastructure and OS. Hypervisor enables multiple OSs to share the same infrastructure, e.g. Virtual Machines (VMs), including processors, memories and storage services. It provides partitioning of resources and ensures that VMs and their applications are isolated from each other. In the next phase also networks can be virtualized [133]. Figure 2.3 illustrates the layers relating to virtualization [45].

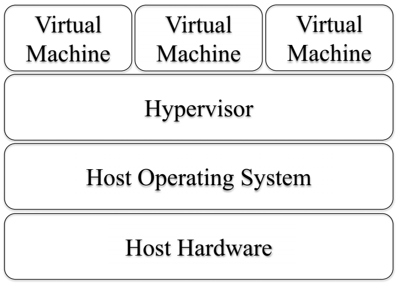

Figure 2.3. Virtualization.

Service Models

Cloud computing offers three different types of services [13]. The most common, Software as a Service (SaaS) service model provides software applications such as email or Customer Relationship Management (CRM). Secondly, the cloud can offer a toolkit that enables customers to create their own services within Platform as a Service (PaaS) model. Finally, cloud providers can sell traditional computing services such as computation and data storage through an open network access. This service model is called Infrastructure as a Service (IaaS).

Each model has different strengths and weaknesses that meet different business needs. The SaaS model offers fast adoption, but it is the most rigid model in providing minimal interoperability functions which leads to lock-in problems. The IaaS model provides the other extreme with the greatest flexibility, but simultaneously the IaaS solution requires more from the implementation side. The PaaS model presents a compromise between the two extremes. Figure 2.4 explains the differences of cloud service models compared to traditional IT. Depending on the model, all, part or none of the layers are self-serviced or offered as a service.

In the SaaS model, the user utilizes a service provider's application that is running in the cloud. The service provider offers its customers the complete service package, including the hardware, software and interface. Customers typically access the service using a web browser or a dedicated software client. Instead of using their own infrastructure, most companies run their applications on the SaaS model [12]. The model hides all infras-

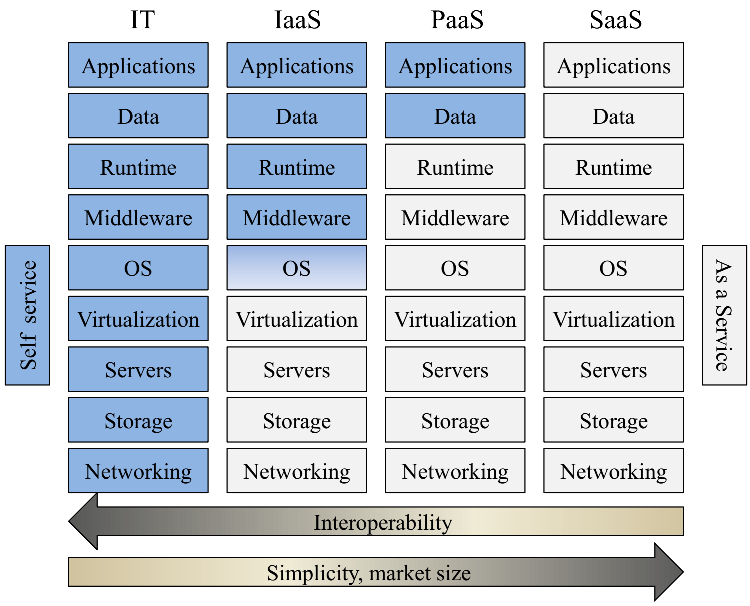

Figure 2.4. Service models [80, adapted from Figure 17].

tructure and software details from the customers. The SaaS model is simple to use, offers access to various end-user devices and provides elasticity in the face of sudden load changes. On the other hand, customers have limited opportunities to tailor the services to meet their individual requirements.

The PaaS model provides cloud users more flexibility than the SaaS one. The end users can deploy their own applications on the cloud using programming languages, libraries, services and tools supported by the PaaS provider [124]. The users do not need to care about infrastructure management or scalability, but they can fully control their applications. At the same time the SPs get rid of costs and work related to hardware and tool support.

The third service model, IaaS, provides the most flexibility for its users. It offers virtual machines with unique IP addresses, and different storage alternatives in providing for a user's need. The hardware layer is fully abstracted. Compared to PaaS, the IaaS users can select their own development tools and operating systems [124]. On the other hand, compared to their own IT, the IaaS model liberates the developers from the chore of hardware management. IaaS offers fast scalability to the users, but the software developers, on the other hand, must themselves manage the load balancing functions.

Deployment Models

Clouds can be deployed through several models [13]. The most common model is called a public cloud. In this model the cloud infrastructure is offered for open use by an entity that owns and runs the service on its premises. Public clouds on the IaaS service model are offered, for example, by Amazon EC2 and Rackspace Cloud Servers, while Google App Engine and Microsoft Azure are the major PaaS providers. The business model is described as a pay-per-use method, which means that a cloud user can buy resources on a need basis. There are also other business models, where long term agreements are used in parallel to hourly charging.

The private cloud infrastructure is available exclusively for a single entity, including possibly multiple users. The cloud owners can run the infrastructure from their own premises, use a hosted solution or a combination of the previous alternatives. There are several open source software stacks on private clouds, and the most well known projects have been Eucalyptus, OpenNebula and OpenStack. All alternatives offer interoperability with Amazon EC2, which enables a smooth transition from a private cloud to a public one and vice versa. Common APIs enable hybrid solutions as well.

The third deployment model is called a community cloud, which is a less common choice. In this model a group of enterprises share the common infrastructure that they manage and operate themselves on or off premises. A common denominator of the group can be, for example, location, mission or security requirements. Public, private and community clouds can also be combined, in which case the architecture is called a hybrid cloud. We will review the hybrid cloud option in detail next.

Hybrid Cloud

Public and private clouds have their own advantages, but also a few disadvantages. One of them is performance unpredictability [12]. A hybrid cloud, however, can be an answer in searching for an optimal solution. Public clouds meet scalability needs when dealing with a large variation in load, but in the long run the usage of public clouds can become expensive [193] and furthermore, the performance is not on the carrier grade level [84]. On the other hand, private clouds can be designed for the strictest performance requirements, but unfortunately, without an illusion of infinite resources.

These thoughts ultimately lead to the idea of utilizing a hybrid cloud.

The hybrid cloud is the most suitable for unpredictable traffic such as video [32, 108]. Video is also the major single application of Content Delivery Networks (CDNs) that can benefit from cloud computing [111, 194]. It is mathematically proven that the hybrid cloud is the most economic solution upon variable traffic load [195].

The basic idea of the hybrid cloud for improving the performance is simple: use the private cloud to manage the basic load and the public cloud for excessive traffic. Applying public clouds for short term usage is not straightforward due to the relatively long start-up time of VMs. A smooth transition of load within the hybrid cloud requires that public cloud VMs can be controlled from a private cloud [128] and launched well in advance before they are needed. Furthermore, this means that traffic characteristics must be predicted and modelled to build an effective load balancing functionality [60]. In general, mapping of VMs into a hybrid cloud is a complicated task [125].

Storage

Distributed databases are part of the infrastructure services and also one of the key concepts of cloud computing. Database as a Service (DaaS) was already invented early 2000s [76], but the introduction of cloud databases have popularized the DaaS abbreviation. Cloud databases have specific needs that differ from ordinary IT data service requirements. The term NoSQL describes the family of distributed databases that do not support the SQL interface [139]. These databases scale horizontally and are capable of storing massive data amounts. Originally NoSQL databases were designed to support social media services, but nowadays several open source software implementations are available to developers.

Compared to the SQL database that is typically used in telecom networks, a NoSQL database provides different interfaces, transaction guarantees and scalability. SQL solutions are known as Relational Database Management System (RDBMS), and they are optimal for Online Transaction Processing (OLTP) in contrast with NoSQL versions that are better suited to Online Analytical Processing (OLAP) [3]. A SQL database confirms Atomicity, Consistency, Isolation, Durability (ACID) requirements, but NoSQL databases typically support Basically Available, Soft state, Eventually consistent (BASE) principles [149].

The first NoSQL implementation was presented by Google, who published in 2003 the details of its Google File System (GFS) [65]. Later in

2006, the company published an article describing Bigtable [34], a distributed storage system built on top of GFS. Amazon published its own Dynamo [47], and similar open source versions were soon introduced by the Apache Software Foundation (ASF), including, for example, Cassandra [104], HBase[2] and the Hadoop Distributed File System (HDFS) [170].

Comparing different data serving systems is very difficult. There are a few benchmark tools, but often they are tailored to certain applications, and for that reason the results cannot be generalized. The Yahoo! Cloud Serving Benchmark (YCSB) is one example of a new type of cloud storage benchmark services [42]. It provides a generic and open source software-based framework for performance testing of new generation cloud serving systems. Another paper [44] evaluates database performance over HBase. Replication is the standard way of achieving durability of data in NoSQL databases. Replication makes confidentiality policies harder to enforce because it creates more copies of the data that are potentially vulnerable to attack [205].

2.2.4 Dynamic Resource Provisioning

Frequently changing traffic loads set strict requirements for optimal cluster schedulers. While a PaaS solution releases designers from the burden of a scheduler design, an IaaS alternative forces system architects to build a proper scheduler algorithm. Ali-Eldin et al. [6] define that the algorithm must be fast, scalable, adaptive and robust. All these requirements are essential in the context of variable workloads, but the fourth requirement especially addresses the oscillation problem. The research on real workloads has suffered from a lack of real traces, but recently Google released a cluster trace from a period of 29 days[3]. Initial analysis [153] points out that most tasks are short, lasting only a few minutes, requiring just a small part of a VM capacity and consisting of one component only.

A load balancing algorithm can be designed on a single, e.g. black box view, or interconnected applications. On the latter, performance monitoring [35] is the typical measurement approach, while deadline of application execution is another method [117, 186]. Cost-wise the optimization can be done both from a user's or SP's point of view. In addition to cost, also energy consumption can be added as an additional factor in the equation [122]. Dynamic resource provisioning algorithms can be divided into

[2] http://hbase.apache.org/
[3] http://goo.gl/5uJri

reactive and proactive models. Reactive models can suffer from deteriorating Quality of Experience (QoE) due to delays in VM start-up time, but proactive models can launch the VMs in advance depending on the predicted results. The theory of predictive algorithms is old and common in scientific fields such as econometrics, machine learning and neural networks.

Cloud infrastructure can be modelled as a simple G/G/N queue, where both the interval of incoming requests and the service time are random, while the number of servers varies upon need [6]. Applying proactive models for autoscaling in cloud computing has become a popular research area. For example, linear regression is applied to solve optimal cost parameters for deploying business software in the hybrid cloud [179]. A probabilistic model checker called PRISM evaluates the performance of VM live migration [98]. The PRESS algorithm [70] uses signal processing and statistical methods to identify repeating patterns and a discrete-time Markov chain to predict load in the near future. Markov models are well suited to analyze http traffic that is often of the heavy-tail type [138]. Finally, statistical methods such as the Box-Jenkins models [25, 49], supported by a Fourier Transfer analysis [203], are highly efficient in modelling stationary time series data. The monitored load can be centralized or distributed [198, 10].

2.2.5 Performance Testing

Performance testing in a cloud computing context is rather challenging. Most systems are closed, and there are, moreover, no standardized measurement systems available. Throughput and delay are straightforward measurement factors, but the performance evaluation of availability requires more extensive and time-consuming laboratory setups. It is notable that high throughput and low latency are both desired but often contradictory features [139]. Based on the CAP theorem it is difficult, if not impossible, to optimize the system in all three above mentioned factors [185].

SLAs are typically used between service providers and users to agree the Quality of Service (QoS). In cloud computing SLAs have an even more central role due to continuous changes in cloud infrastructures. Typically cloud providers just promise the availability figure, but for critical services, stricter conditions are applied. In any case the customer has the burden of proving any SLA violations, and for that reason continuous performance monitoring is an essential tool for cloud users [142]. Throughput is measured as the number of operations within a time interval. VM performance can be measured as messages per second, and database performance as transactions per second. Load is a similar concept, but it describes how much the clients stress the system [139].

Latency or response time is usually calculated with the 95th percentile method rather than an average to filter out occasional deviations. In addition, all measurements must be repeated several times to minimize occasional errors. The delay measurements must verify which delay components are based on the network and the processing time to prevent wrong conclusions. It is a well-known fact that increasing load improves the throughput up to a certain limit but after that the results get worse. This point is called a knee and there the performance with respect to delay and throughput is optimal [191].

2.3 Organization

This section summarizes the main theories on organizational questions. The key parts describe how organizations are transforming from value chains to value networks, and furthermore to business ecosystems. On the other hand, one-sided platforms are replaced with two- or even N-sided platforms, where business transactions can be made on several directions. Organization consists of resources and capabilities, which relate to technologies, marketing and finance. These factors are required to deliver a service. Usually the organization must be capable of cooperating with other organizations, because services produced solely by a single unit are rare [24]. In addition to exploratory strategy, enterprises must also execute exploitative actions to run the current business. O'Reilly and Tushman [136] call this kind of duopoly organization ambidextrous. The key idea of this proposal is to separate the existing and new organizations to independent units, for guaranteeing the wealth of the traditional business but also, at the same time, for achieving breakthrough innovations.

The Resources, Processes and Values (RPV) theory [39, pp. xvii-xviii] explains why incumbent companies have difficulties in reacting to disruptive innovations. Organizational resources and processes are usually tailored for the current products. Company values prioritize factors such as customer requirements, cost structure and size of opportunity, giving less emphasis to new innovations. On the other hand, Value Chain Evolution (VCE) theory [39, p. xix] assesses whether a company organizational structure is optimal for the operations. The core idea of the VCE theory says that companies should drive the performance along the dimensions that are most valuable for their customers. Radical innovations create a problem for companies because the company organization may be optimal to execute old products, but completely unsuitable for new ideas [38].

Eisenmann et al. [54] emphasize that new and mature markets require different approach in openness. Next, they discuss which roles of platform-mediated networks should be opened. The roles can be divided into demand-side and supply-side users, platform providers and platform sponsors. The examples from existing industries present that any combination can lead to a successful business, meaning that no general statement can be made about the attractiveness of open versus closed strategies. Eisenmann et al. conclude that, in the long run, forces tend to push both closed and open platforms toward a hybrid model where the platform technology is centrally controlled, but where end users are served in a shared manner. Figure 2.5 illustrates the hybrid model.

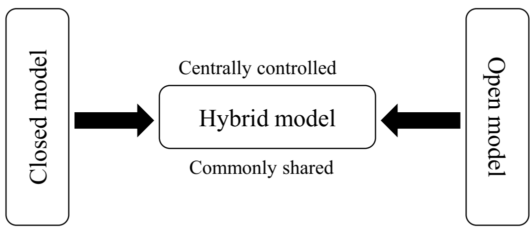

Figure 2.5. Hybrid model.

2.3.1 Value Chain and Value Network

Allee [7] defines that value is created through the exchange of Goods, Services and Revenue (GSR), knowledge and other intangible benefits. The value created by products and services is usually evaluated by the value chain analysis, which was originally created for processes produced within a single company, but later it has been applied to multiple companies or the whole industry as well. However, Peltoniemi [143] notes that the value chain analysis partly ignores the dependencies between the resources of different companies. The value network offers a better option for new economy organizations, particularly for those where both the product and the supply and demand chain is digitized [146].

The value network does not create value in transforming the objects, but in their mediation [143]. The strength of the value network originates from cooperation and interaction among participating companies. Companies believe that revenues increase and, on the other hand, costs decrease. Companies become members of the value network based on their unique competencies. A value network is not tied to any specific region, but it can be even global. Additionally, the same value network can include companies from different industries. The value network analysis is also well suited to the telecommunication industry, where system complexity continuously expands due to frequent strategy updates with multiple entry and exit points [107].

2.3.2 Ecosystem

Natural and business ecosystems are complex adaptive systems that share some fundamental properties and follow the same laws. However, the ecosystems also have differences. Business ecosystems can make conscious decisions and compete over possible members [106]. A natural ecosystem, instead, aims to survive, while a business ecosystem targets to create innovations [89]. An ecosystem consists of loosely interconnected participants that simultaneously compete and cooperate, leading to coevolution of the system [144]. Enterprises depend on each other, which means that they share their fate with each other [89]. Being a member of an ecosystem brings advantages to the companies because they can make alliances and protect against possible invaders. The setup also includes a risk that a change in the ecosystem propagates through the network leading to death [106], instead of self-renewal [127].

Business ecosystems share a few complexity factors such as self-organization, emergence, coevolution and adaptation [143, 144]. These factors will adapt to external changes. Self-organization in the business ecosystem context means that an ecosystem organizes itself without an external controller, and events happen spontaneously due to local interventions [126]. Emergence refers to the links between individual organizations and to the fact that the long-term outcome of a system is unpredictable [172]. Coevolution means that the evolution of one organization impacts the evolution of other companies [145]. As an example, the capacity of higher transmission links will be consumed by new and heavier services, e.g. Jevons' paradox "every bit will be used" holds true [5]. Finally, ecosystems are adaptive to external changes, such as regulation [145].

Ecosystem success factors include productivity, robustness to manage changes and the ability to create niches and opportunities for new firms [89]. These factors can be analyzed both within the ecosystem and at company level. According to Hartigh [82], company level productivity is the key measure to identify the health level. Unproductive companies have difficulties to survive, and as a result, these companies will lower the health of the whole ecosystem. At a network level a healthy ecosystem is represented by the number of connections between the partners that also form clusters of different type of companies and interact intensively. Finally, Hartigh says that a healthy ecosystem should have a few central players that possess a high visibility in the market.

Mobile networks are typical examples of highly complex business ecosystems. Various incumbent and emerging players such as network operators, integrators, content providers, advertisers and developers are connected to subscribers [17]. The ecosystem inter-connections are critical, and constantly changing [197]. Currently mobile social networking and the end users' role have become more important in the mobile ecosystems [59].

2.3.3 Value Network vs. Ecosystem

Value networks differ from ecosystems in several aspects. Table 2.2 summarizes the main differences between value networks and ecosystems.

Table 2.2. Comparison of value networks and ecosystems.

Factor	Value network	Ecosystem
Control	one actor	self-organized
Stability	static	dynamic
Mode	cooperation	coopetition
End user role	passive	active
Knowledge sharing	operational data	open technologies
Entry barrier	high	low

In value networks one actor is usually more dominant than the others, while the ecosystems are self-organized [145]. Furthermore, in value networks the relationships between companies are more stable and predefined, whereas in ecosystems the interactions of companies are more dynamic and companies have a low barrier to joining or leaving the ecosystem [144]. Value networks are typically cooperative where only operational information is shared. The major working mode in ecosystems is parallel competition and cooperation, e.g. coopetition, and knowledge is shared through open technologies [89]. Finally, the role of end users is different. In value networks they are passive, but in ecosystems end users actively participate in the operations [59].

2.3.4 Network Effect and Two-Sidedness

According to Shapiro and Varian [167, Chapter 7], network effects or externalities affect the demand-side economies of scale, meaning that the demand of a service or goods defines its value. Tuunainen and Tuunanen [183] summarize that the platform's value to users largely depends on the number of users on the other side of the network, and the value of the platform relatively grows when both sides are matched. They stress that both same and cross-side network effects must be understood, as well as different types of costs related to the platform. On the same-side, a positive network effect can be obtained by economies of scale, while a cross-side phenomenon brings indirect network effects. Parker and Alstyne [141] give examples where temporal or skill-based discounts are given to young or novice users, on behalf of experienced or professional users.

Two-sided markets differ from one-sided in several ways. The key distinction for two-sidedness is that network effects must cross market populations. Rochet and Tirole [157] have defined that a two-sided market with network externalities exists if a platform can cross-subsidize end users that belong to different categories and are parties to a transaction. Similarly, Ballon [14, p. 114] defines the two-sided network theory as describing cases where two different types of customers interact on a platform where the interaction is affected by indirect network externalities located on the opposite sides of the platform.

According to Evans and Schmalensee [58], many diverse industries, including telecom networks [14, p. 115], are occupied by businesses that operate two-sided platforms. As an example, content delivery networks comply with two-sidedness [204]. Furthermore, two-sided platforms share a few common characteristics defined by Evans and Schmalensee [58]: indirect network effects, economies of scale, congestion, platform differentiation and multihoming.

Two-sided platforms also have a few threats. Correct pricing in two-sided markets is important due to the subsidization of unpaid users [141]. Compared to proprietary and company controlled software platforms, open source software platforms challenge the profitability of software platforms [51]. Eisenmann et al. [53] warn about envelopment risk, which means that the platform provider may want to exclude other players from the market. Multihoming that refers to parallel service providers prohibits the exclusion risk from materializing. The credit card market is a good example of a multihoming setup.

2.3.5 Broker Models

Barros and Dumas [16] define a service ecosystem in terms of five actors: provider, user, broker, mediator and specialist intermediaries. A broker connects service providers and users together, offering additional services such as a single-sign-on and payment in one service offering. A mediator provides service adaptations for different data formats and service functions, while a specialist intermediary offers specialized service components for providers. These can be, for example, monitoring and payment services. In addition to the above, Riedl et al. [154] mention that there are platform providers who offer an overall platform for other providers to work with.

Evans and Schmalensee [57] divide brokers to three categories: digital marketplaces, audience builders and cost minimizers. Another way to make the distinction is to analyze what is brokered: network, content or both. In mobile networks network brokering incorporates topics such as messaging, payment, location, profile, authentication and security, while content brokering requires actions on content formats, copyright protection, hosting and distribution mechanisms, to name a few examples [152].

The European bank industry has used a so-called 4-corner model within the Single Euro Payments Area (SEPA)[4] agreement. The main idea of the approach is that each customer makes a deal only with her or his own bank, and banks have a clearing system between each other. This setup simplifies clients' agreement actions because each customer needs just one agreement to transfer money between any customer and a bank. Furthermore, banks conclude an agreement between themselves according to the inter-connection procedures and agree the fees within the SEPA agreement that is based on ISO 20022 and XML standards. The 4-corner model has been recently applied in a Finnish mobile certificate service[5].

[4] http://ec.europa.eu/internal_market/payments/sepa/
[5] http://www.mobiilivarmenne.fi

2.4 Finance

The last background section presents the introduction of the financial part of the work. First, the relevant consumer theory is introduced, followed by a summary of internet revenue models. Next, cloud cost and market analysis are presented. Finally, a discussion on ARPU break-even point is offered.

2.4.1 Consumer Theory

Consumer theory provides important background information for the business model analysis. Consumer theory belongs to microeconomics that defines how consumers' preferences meet their demands. The first phenomenon under study is the demand curve [167]. The baseline case A in Figure 2.6 shows the status quo, where the revenue is optimized with a certain price. More liberal terms and conditions, shown in case B, increase the value of the service to the customers, which enables higher prices. Although the demand will reduce, the total revenue remains the same, because the higher prices compensate for the reduced sales. A sustained business model seeks the optimal solution between these forces.

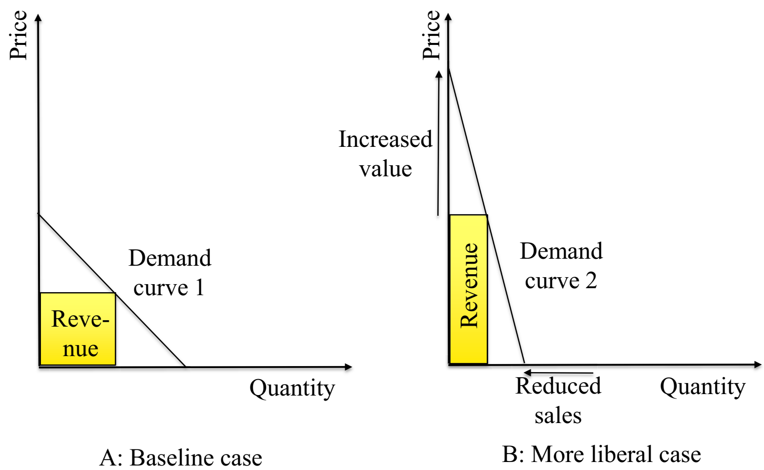

Figure 2.6. Demand curve [167, adapted from Figure 4.1].

The substitution, network and exposure effects are other important consumer theories. The substitution effect defines how well two products substitute each other. For example, text messaging and email can substitute each other, but between TV and communication services the similar correlation cannot be necessarily detected. The network effect is the impact that one user of a good or service has on the value exposed to other people using the same good or service [148]. Finally, according to the exposure effect end users tend to create a preference for the products that they are familiar with.

2.4.2 Internet Revenue Models

One strength of the Internet has been the innovativeness on revenue models. According to Shuen [169], the Web 2.0 world has six different revenue models: subscription, advertising, transaction, volume, licensing and sponsorship. Anderson [9] suggests one additional revenue model: free. Free is very effectively used in conjunction with subscription and advertising business models in various Web 2.0 services. This model can be called freemium[6], combining the words free and premium. Free access is the key to accelerate the positive momentum of the network effect. Parallel discontinuous technological change, ignored by the incumbents, may enable cost effective products that initially attract cost sensitive low-end customers [38]. For example, Amazon, Skype, eBay, Yahoo and Apple have been successful in utilizing novel revenue models that are well suited to a new economy.

2.4.3 Cloud Cost Analysis

Cloud computing involves a complicated cost analysis question. Public and private clouds provide a heterogeneous pricing system, which makes a neutral cost comparison almost impossible. The situation is even more complicated when a cost analysis of a hybrid cloud is performed. The costs of public cloud can be transparently available, but price variations between on-demand and reserved instances make the cost calculation challenging. Spot pricing [119] forms its own challenge, creating fully unpredictable costs.

Several approaches are available for cost-optimally distributing the load within a hybrid cloud, for example, by optimally allocating computing tasks to cloud resources at each time slot [179] or by determining the fixed cost-optimal threshold for the private cloud resources and allocating the excess load to the public cloud [121]. On the other hand, Weinman [195] proved that the optimal time to use the public cloud is inversely re-

[6] `http://www.avc.com/a_vc/2006/03/my_favorite_bus.html`

lational to the utility premium of the public cloud. For example, if the public cloud is four times more expensive than the private cloud, the public cloud should be used 25% of the time period. Finally, Walker [193] has made extensive studies on the cost comparison of leased or own systems, relating to storage and CPU platforms.

2.4.4 Cloud Markets

Cloud computing offers a perfect match to Small and Medium Enterprises (SMEs) [55]. They are less capable to handle high up-front costs, but also more flexible to adapt their business models to changing market conditions. Especially startups clearly benefit from cloud computing: Suster [181] claims that the initial investment costs of startups have dropped between 2000 and 2011 from $5M to $5k. On the other hand, small countries with a well-developed public infrastructure are well suited to adopt cloud computing to attract substantial capital expenditures [55]. The ICT market has not yet fully utilized the potential of cloud computing, because according to the statistics [55], only 9% of the total ICT market, worth $494B in 2012, has been conquered by cloud computing. The highest investment potential is seen in business applications and, especially, in data storage systems.

2.4.5 ARPU Break-Even Point

New radio technologies, such as LTE, enable new mobile business models. The price per bit will significantly drop. According to the study by Hoikkanen [86], LTE includes significant improvements into the radio interface. The advanced radio reduces the cost of data, and enables several, potentially important, new services. If the average data usage is 20 MB per month, and doubling every second year, a break-even point can be achieved with an ARPU level of 10 €, whereas a level of 15 € is already very profitable. Bohlin [22] confirms this estimate by noting that according to his simulations, a minimum ARPU level of 15 to 19 € will be required for the LTE business case to become viable. However, Bohlin warns that a more precise estimate depends on various factors. For example, the average data consumption has a big impact on the results. According to Blennerud [21], a price of one gigabyte will drop by 2012 to 1 € if the average data consumption is two gigabytes per user.

3. Results

This chapter summarizes the main results gathered from the attached publications in accordance with the STOF classification.

3.1 Service

The Service section includes two major results, a list of proposed open APIs and a description of a mashup service, called Event Experience.

3.1.1 Open APIs

Open telecommunication APIs can provide various novel services and a few high level examples are shown in Table 3.1 that is extracted from Publication I, Table 2. Messaging, location and payment APIs are themes common to several operator trials, see Publication I, Table 1, and the first GSM Association (GSMA) OneAPI specification, version 1, covers those APIs. More advanced services built on profile, presence and call control APIs are available in GSMA OneAPI versions 2 and 3.

Table 3.1. Examples of open APIs and services [Publication I, Table 2].

API	Example services
Call control	Call rerouting & triggering
Messaging	Message rerouting & triggering
Location	Advertising, timetable, find, friend & family
Payment	Micropayment
Network Presence	Roaming state, free & busy
Profile	Advertising, recommendations
SLA	Quality of Service, security, content caching and delivery

At this moment, operator offering on open APIs varies a lot, see Publication I, Table 1. Messaging API is commonly provided, but otherwise most

APIs are operator dependent. However, the introduction of the GSMA OneAPI standards has improved the situation. Location in combination with social media services has attracted most of the developer interest [163]. However, end users need flexible and easy-to-use tools to set their location visibility and accuracy [168, 190]. Payment and billing APIs are also highly valued by the developer community, but on these services the regulation sets various limitations on innovativeness [130]. Also, operators must first carefully analyze which assets they offer to developers, and what are the best business models to support the open innovation [23].

3.1.2 Event Experience Mashup

Publication III evaluates eight different usage scenarios, which use open APIs most extensively. All cases were analyzed by Chang and Kaasinen [36], who are looking at the cases from a user innovation viewpoint. For the purposes of this study, only the most representative use case, the mobile ticket was reviewed in detail. The main idea of the Facebook-based application, later renamed Event Experience, is to provide a compact event mashup, which also offers, in addition to basic event data, payment, ticketing, navigation and notification services. The service utilizes operator APIs, such as messaging, location and payment, although the payment API was not available in reality due to regulative restrictions.

The application utilizes two different two-sided platforms. Firstly, Facebook itself is a two-sided internet platform that connects end users and applications together. Secondly, the Open Telco platform creates a two-sided market, where on one side subscribers use services, while on the other side service developers and content providers create novel services, based on the data provided by the APIs. The platform provides both same- and cross-side network effects [141, 183] that positively accelerate supply- and demand-side offering [167]. The Event Experience application can be modelled with a value network [146], but because the factors of the value network can be replaced or bypassed [46], the system may transform into an ecosystem [143].

The conclusions of the representative case study claim that, firstly, the Open Telco platform must support operator interoperability to maximize the network effects [58] and to cope with internet competitors [93]. Secondly, the pricing model must be carefully analyzed to meet both end user and developer expectations [189] and also the challenges of two-sided platforms [141]. Thirdly, operators have an access to subscribers' profile data, but the usage of sensitive information sets pressure on the privacy rules [11]. Finally, operators suffer from a heavy regulation burden [103] when compared to internet challengers. The current regulative laws should be adapted to better match the current markets and novel mashup business innovations, for example in the payment and privacy sections.

3.2 Technology

This section summarizes the technology innovations from the attached publications. The novel ideas include the Open Telco architecture, performance evaluation of NoSQL databases in the mobile environment, and the usage of proactive resource provisioning algorithms in the load balancing of the hybrid cloud architecture. Furthermore, results from the survey research on how to map OSS/BSS functions to the cloud are presented. Finally, a set of proof of concepts are briefly introduced.

3.2.1 Open Telco Architecture

Based on the mashup ecosystem recommendation by Stecca [175], the Telco Platform should be built using cloud computing principles in PaaS fashion. The PaaS solution, which can be implemented with a Service Delivery Platform (SDP), consists of three main layers: secure APIs, support functions and adaptation layer, as shown in Figure 3.1.

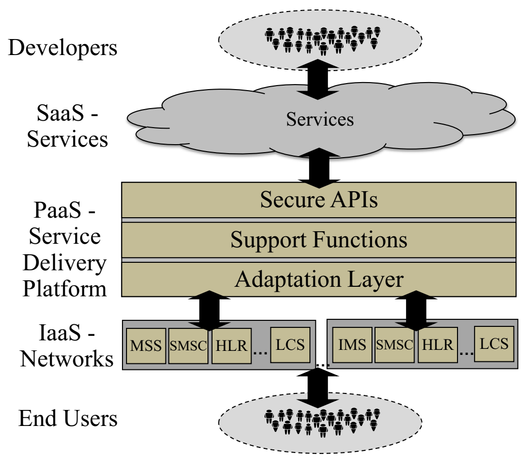

Figure 3.1. Open Telco architecture [Publication I, adapted from Figure 1].

Secure and simple, REST-type [162] APIs provide an external view of the network resources, which also fulfill the coopetition [167, Chapter 8] and backward compatibility requirements [180]. The Mashery case, presented in Publication II, in particular, highlights the importance of introducing open APIs without affecting the existing services and running new business from a public cloud to eliminate investment risks.

Support functions are responsible for background processes, while the adaptation layer provides intelligent data filters to manage the massive information flow gathered and accessed from the network layer that also can be called the IaaS layer. The actual services are created by developers on the SaaS layer.

3.2.2 Performance Evaluation

The contributions Publication V and Publication VI included performance measurements on cloud computing and storage systems. In this context, the measurements are reflected in a SLA, including factors such as throughput, delay and availability. The main interest is on the throughput and delay that can be easily measured with the emulators. However, research into availability is very challenging in a laboratory environment and requires long term performance measurements on the sustainability of the cloud computing systems.

Dynamic Resource Provisioning
Publication V, Section 2.3 evaluates whether the Autoregressive Moving Average (ARMA) algorithm can be applied to model SMS traffic for the purposes of dynamic resource provisioning. In this use case SMS traffic prediction is trivial, but nevertheless, the data can be used to test and calibrate different prediction algorithms. The SMS data is not stationary as required by the ARMA algorithm, but it can be stationarized. Moreover, the model assumes that data is assumed to be centralized and not distributed [198, 10]. The first order ARMA filter is represented by Publication V, Equations 1–2.

The parameters α and β were selected with an off-line method, using a tool called Forecast package for R[1]. The tool used maximum likelihood estimation to optimize the prediction with the real data. The results were good, presenting on average 4.5% prediction error. See Publication V, Figure 3 for more details. The chosen method works well with periodic data,

[1] `http://robjhyndman.com/software/forecast/`

but it is clear that with more random traffic pattern, the results with the ARMA model would not be as good as here, and other models such as Autoregressive Integrated Moving Average (ARIMA) or more complicated Autoregressive Fractionally Integrated Moving Average (ARFIMA) should be verified [49]. The second improvement step can be done by selecting the ARMA parameters through an on-line procedure.

The measurement in Publication V, Figure 4 calibrated the SMS traffic performance in the public cloud, in accordance with throughput and delay. The results indicated the threshold value for the VM capacity, a fact which was required in the cost analysis and design of the load balancing functionality. Finally, the last measurement compared the performance of reactive and proactive models in the hybrid solution. The results in Publication V, Figure 8 imply that the proactive model, on average, provides a smaller message delay, and also more stable performance with higher loads than the reactive model.

The performance analysis has a few limitations. The research is dedicated to a black box approach and does not look at the dependencies between the sub modules [117, 186]. On the other hand, the presented Algorithms 1–2 in Publication V are novel, and unlike the previous research [35, 198, 160], the special requirements of the hybrid cloud in the telecommunication context are highlighted in the algorithms.

NoSQL Databases in Mobile Networks

The main performance measurements were carried out within the public cloud, where six VMs were allocated for the research, using the setup shown in Publication VI, Figure 2. The measurements verified the throughput and delay in the function of replication factor, load and database size. The first measurement in Publication VI, Figure 4 evaluates the impact of the replication factor to performance. The results indicate that the replication factor does not have a meaningful impact on the results. Also, the knee [191] is very visible, proving that after the threshold load, which in this measurement is roughly on the level of 4000 requests per second, the performance dramatically deteriorates. It is notable that commercial, carrier grade HLR systems have reported similar values, although the numbers are already a few years old [74].

The results of the second measurement, where the research object is the impact of database size, are shown in Figure 3.2, taken from Publication VI, Figure 5. The diagram logically shows that the increased database size results in worse performance. The threshold of good performance is at the level of one million subscribers. One client does not have enough capacity to load a large database, and for that reason a higher delay is detected with a large database and low load. Finally, Publication VI, Figure 6 verifies how HDFS can resist node failures. The results show that the system can recover, due to replication, from failures in a few seconds. The value is reasonable although not enough for real-time communication services. However, the results can be improved by fine tuning the timeout parameters.

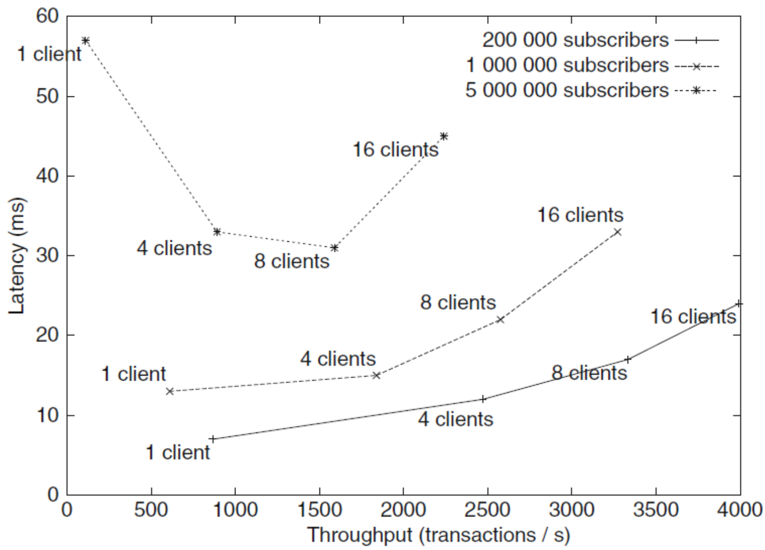

Figure 3.2. Impact of database size [Publication VI, Figure 5].

3.2.3 MVNO and OSS/BSS

The research in Publication V aimed to verify whether cloud computing is well suited to MVNO OSS/BSS functions in the mobile support systems. Due to the complicated nature of the support systems, survey was chosen for the research methodology. A group of OSS/BSS experts were invited for two rounds of interviews that were conducted in accordance with the Delphi method [171].

The first questions asked what the critical parameters in the MVNO and cloud context are. According to the answers, cost benefits and flexible architecture are the key requirements for the MVNOs as seen in Publication V, Figure 6. In the cloud environment, concerns were raised about security and performance parameters but on the positive side, cloud computing can improve energy efficiency as shown in Publication V, Figure 7. The results of the OSS/BSS mapping to the cloud are presented in Figure 3.3, which is adapted from Publication V, Figure 8. As a conclusion, the responders agreed that most Business Support System (BSS) functions can be mapped to the clouds. However, OSS or prepaid and network systems were not recommended for the clouds due to performance doubts.

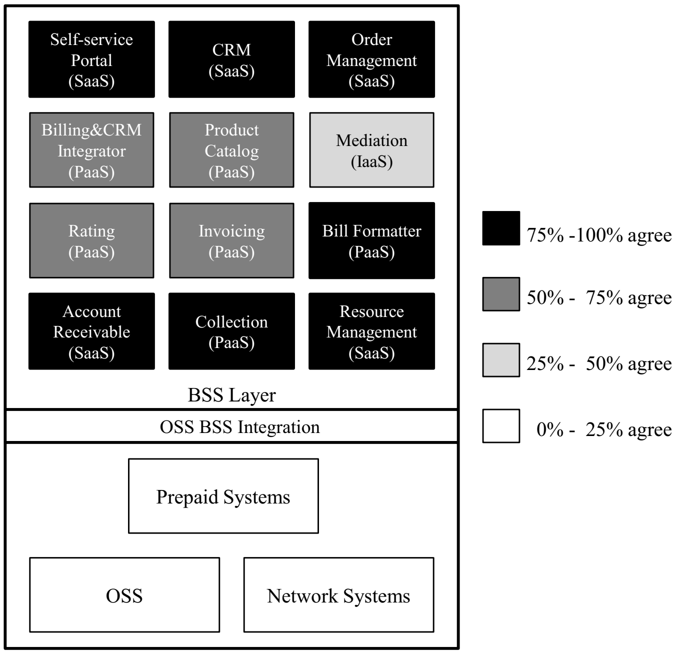

Figure 3.3. MVNO systems in cloud [Publication V, adapted from Figure 8].

The BSS functions can be distributed to SaaS, PaaS and IaaS service models depending on the functionality as shown in Figure 3.3. User centric functions such as the self-service portal and CRM can be implemented in SaaS functions, while asynchronous, background computation tasks, including billing, are well suited to the PaaS model. Computation intensive batch jobs such as mediation should be implemented in the IaaS model. In general, cloud computing is well suited to background processes, and this idea is not restricted only to MVNOs, but also MNOs can benefit from the cost benefits of applying clouds to BSS. Similar ideas were already presented by Jim Crookes from BT in 1997, well before the invention of clouds [43].

The Billing-as-a-Service concept that is illustrated in Publication IV, Figure 10, is designed to be used in PaaS fashion, shared by MVNOs or even MNOs. The proposed architecture can utilize the hybrid cloud architecture in the mediation and rating engine functions. Due to security and access reasons databases should locate in the private cloud, while the public cloud can justify load peaks.

3.2.4 Proof of Concepts

Along with the research process, a few proof of concepts were developed to verify the research claims. Publication IV includes a hybrid cloud test bed for verifying a CRM solution, shown in Publication IV, Figure 9. Publication V continued the hybrid cloud research and evaluates, using fully open source software components, the SMSC operability, shown in Publication V, Figure 5. The third implementation, simulating the HLR operations and presented in Publication VI, was based on an existing benchmark tool [178], which was ported from a SQL to NoSQL solution [139], and tested mainly in a public cloud environment. Proof of concepts did not include the SS7 implementation due to complexity reasons. The lack of signalling support is one challenge for the cloud providers. Proof of concepts are described shortly next and in detail in Publication IV, Publication V and Publication VI respectively.

Customer Relationship Management
A generic hybrid cloud architecture was tested with a CRM application. See Publication IV, Figure 9 for details. The private domain includes analytics and a database as well as controller images and a server that also includes the load balancing functionality. The public cloud counterpart supports additional controller instances that are connected through secure links to a database. The IP addresses of the public cloud instances are added to the load balancer. The load balancer logic has been modified to switch the load from private to public cloud instances whenever the private cloud instances reach their maximum capacity, in terms of memory, processing or other attributes.

Short Message Service Center
The SMSC functionality was implemented using the open source software components that are presented in Publication V, Section 3.1. The hybrid architecture follows the same principles as described above, and the soft-

ware blocks are shown in Publication V, Figure 5. The Resource Controller includes two blocks: the Monitor software and an optional Forecast functionality. The Monitor periodically reports the load detected by the Load Balancer to the Resource Allocator. The Forecast block can improve the monitoring data by analyzing the history data in order to predict the future data points. Based on the received information, the Resource Allocator, with help of the Virtual Infrastructure Manager, scales up or down the resources available to the system. A virtual bridge inside the private cloud enables communication between the SMSCs, while on the other hand, the Load Balancer provides a gateway to the public network.

Home Location Register
The original TATP benchmark database, representing HLR, includes four different tables called subscriber, access info, special facility and call forwarding as shown in Publication VI, Table 1 [178]. MSC clients use seven different transactions that either read or write data. The transaction distribution is known from real networks. The NoSQL version of TATP was implemented using Java and ported for HBase [139]. The benchmark tool includes a user Command Line Interface (CLI), separate entities for creating a table, populating it and running the benchmark. A user can define the number of client threads, which further execute queries according to the TATP probability distribution.

3.3 Organization

This section presents the novel organizational contributions from the publications. The novelties include two main ideas, a virtual broker and a two-sided platform, both proposals of which can be applied to the Open Telco concept.

3.3.1 Virtual Broker

Already the original Open Telco architecture, shown in Publication I, Figure 1, visualized the need to connect multiple and competing operators to the system in order to enable the birth of a successful ecosystem. Publication II continues the research with a multiple-case study of internet API brokers and GSMA OneAPI trial. Operators may provide APIs directly to their customers or use a broker. The first alternative is cumbersome for the developers because they have to manage the agreement process separately with each operator. The broker option simplifies the agreement bureaucracy involved in the process, but, on the other hand, a new entity is added to the ecosystem, which may decrease the transparency of the services and also increase the transaction costs.

A virtual broker enables the innovation space described in Figure 3.4, extracted from Publication II, Figure 7. It has a few critical advantages when compared to standalone or broker solutions. It supports both two-sided business models and also a single point of contract, between operators, users, developers and service providers. Moreover, the virtual broker provides at the same time a cooperative and competitive business environment, where cooperation maximizes network externalities and competition assures effectiveness of the environment [167, Chapter 8].

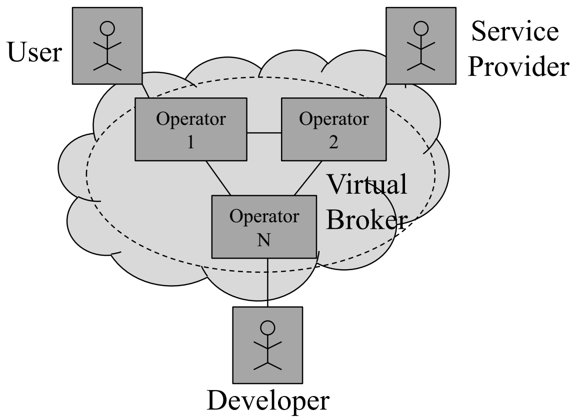

Figure 3.4. Virtual broker architecture for the Open Telco [Publication II, Figure 7].

The presented model follows the idea of a hybrid model presented by Eisenmann et al. [54]. The virtual broker is a compromise between closed and open platforms, where each operator owns its own platform, but the creativeness is shared by many participants. The chosen architecture also provides flexibility because at later phases, the virtual broker can be replaced with one or several physical brokers. The virtual broker can been regarded as an intermediate step from a value network to a full business ecosystem [143]. The broker partly supports the business ecosystem characteristics such as coevolution, adaptation and emergence, although the self-organization aspect can be supported with multiple, parallel brokers. Moreover, the virtual broker enables the usage of ambidextrous organization [136] where new and existing business can be run by separate teams.

3.3.2 Two-Sided Platform

Publication III, Tables 1 and 3 summarize the results of the single-case study of the Open Telco as a two-sided platform. The results are combined in Table 3.2. The analysis includes a framework of eight dimensions. In the Open Telco context network effect, scalability, pricing and regulation are the most critical factors. The platform must gather positive network effects and ensure good scalability by offering, at least, interoperability of local operators, but preferably roaming should work over country borders as well. The pricing mechanisms must be flexible to meet various needs, and finally, regulation should facilitate, not prohibit, novel innovations.

Table 3.2. Analysis of the Open Telco as a two-sided platform [Publication III, adapted from Table 1 and Table 3].

No	Dimension	References	Result
1	Network effects	[58, 157, 183]	Positive network effects available both the same and the cross-side, but at least national or preferably a global coverage required
2	Economies of scale	[58, 157]	Initial capital investment can be high, but operability costs are low
3	Scalability	[58, 157]	Open APIs can support scalability well, but the developer support and global extensions must be planned carefully, using, for example, cloud computing resources
4	Differentiation	[58, 157]	Tailored to open telecom APIs but can be ported to other markets and can support other APIs
5	Multihoming	[58, 157]	Not expected in the first phase but requires mature markets with global roaming support
6	Pricing	[53, 58, 141]	Must be flexible, supporting pay-per-use, flat and variable commission charging alternatives
7	Openness	[37, 41, 64, 167, 182, 189]	Fully open, no need for formal agreement or application approval, simple REST technology
8	Regulation	[94, 134]	Payment and privacy legislation must be reviewed

The results of the analysis can be criticized based on the chosen research methodology, because it was based on a single-case study with a representative usage scenario. More reliable results can be achieved with multiple-case study methodology [199], or using design science principles [101]. However, the applied research framework utilized several steps that are familiar from design science methodology. Moreover, the extensive literature analysis and usage of experts to review and filter the usage scenarios, improved the quality of the created framework.

Figure 3.5 presents a high level view of the Open Telco two-sided platform. The Figure combines the ideas from Publication I, Figure 1 and Publication III, Figure 5.

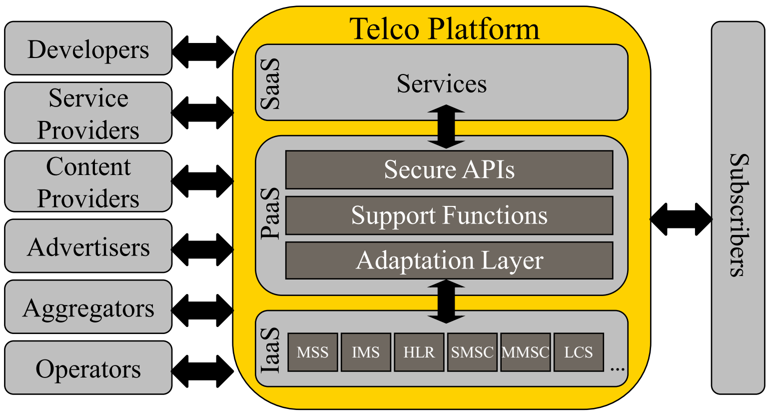

Figure 3.5. Open Telco as a two-sided platform.

3.4 Finance

Public clouds offer flexibility due to the pay-per-use charging principle. However, in the long run and due to the flexibility costs, a public cloud is usually more expensive than a dedicated, private cloud. Nevertheless, comparing prices in a neutral manner is extremely difficult. Already public clouds alone have a diverse list of cost options, leading to a complicated optimization equation. Moreover, a hybrid cloud summarizes all possible cost combinations. The second comparison can be made between hosted and public cloud solutions. The hosted solution is based on a web hotel that offers various hosting options.

3.4.1 Cost Analysis of Hybrid Cloud

As said above, hybrid clouds possess a complicated cost optimization challenge. In addition to computation, data storage, transport and support service costs must be taken into account as well. Publication V includes an analysis of the SMSC in a hybrid cloud use case, utilizing the work done earlier by Mazhelis and Tyrväinen [121], which is also presented in Publication V, Equation 6. The equation defines the optimal time for using the public cloud in the hybrid cloud context. As already discussed in Section 2.4.3, Weinman [195] proved that the optimal time to use the public cloud is inversely relational to the utility premium of it. To support this proof, also Publication V, Equation 6 reduces to the same format if transmission cost, e.g. k is set to zero. The results of the analysis are shown in Figure 3.6, which is adapted from Publication V, Figure 7.

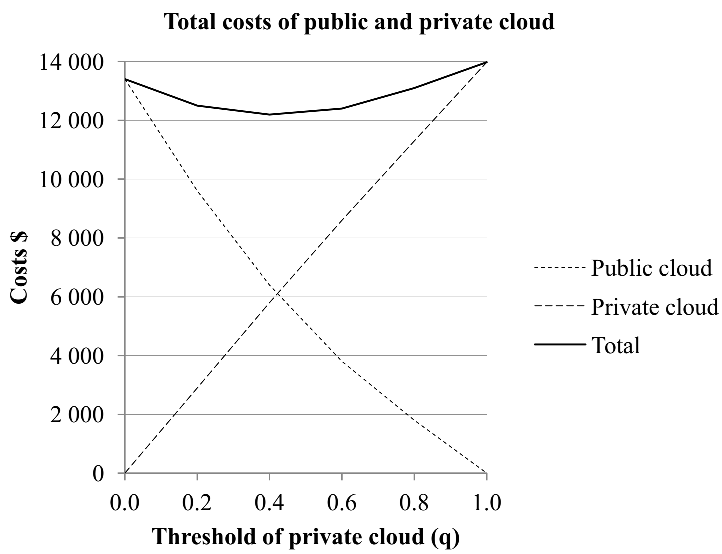

Figure 3.6. Hybrid costs [Publication V, adapted from Figure 7].

With the given parameter values and the SMSC use case, the optimal hybrid solution is achieved when a private cloud capacity is 40% and a public is 60% of all needs. It is notable that the cost advantage of the hybrid solution is limited compared to pure public or private alternatives. This is true due to the fact that the messaging traffic is periodic and the absolute differences in volume between minimum, base and peak loads are not large. On the other hand, extreme peak workloads are very rare.

In addition, the actual message size is moderate, accumulating a very small transmission load. The computation power required to manage an average SMS market segment is not large either. The advantages of the hybrid cloud will not be clearly visible in case of low traffic volume or variation. The message size as such is not a very critical parameter [84].

3.4.2 Cost Comparison of Hosted Solution and Public Cloud

Publication VI, Section 4.4 evaluates in financial terms whether a database should be run in a public cloud or in a hosted rack space. The approach follows ideas presented by Walker [193]. The database can be placed in private premises, too, but in that case the cost structure becomes very complex due to additional dynamic factors such as the property, workforce and energy [71]. In our case the focus is, however, in small clusters, and therefore, the model only includes the cost of hosted rack space, server hardware and disk replacements. The formulas and related parameters are presented in Publication VI, Equations 1–3 and Publication VI, Table 2 respectively.

The public cloud example, Amazon EC2, provides two main alternative payment options: pay-per-use or a one-time fee with a reduced hourly cost, which is also called a reserved instance purchasing option. Publication VI, Figure 8 presents the results of the comparison, showing the monthly costs in a function of the amortization period. The results prove that the Amazon pay-per-use approach is an economic alternative for a short usage, but longer term usage will become cheaper with a long term agreement. On the other hand, a hosted solution is expensive for short usage, but in the longer term it is a competitive alternative.

4. Discussion

This dissertation explorers what mobile operators can learn from internet companies to survive, renew and transform. The financial status of mobile operators is still strong and they have a large customer base, but there are clear signals such as declining ARPU and increasing churn, which indicate that without decisive actions future can bring various challenges. The networking business also tends to carry high moment of inertia, but nevertheless, operators should be prepared for the downward trend. Instead of only concentrating on developing the current core business, companies should actively, independently and simultaneously [136] run both the existing businesses and also search for new opportunities provided by disruptive innovations [39].

The first research question asked whether open innovation and especially open APIs can help to improve the operator business case. Open APIs can be regarded as the first step towards a Network as a Service (NaaS) paradigm, where operators can offer their network resources according to the cloud computing service model [29]. The research shows that open APIs, referred to here by the term Open Telco, represent a promising initiative, which enables a two-sided platform and helps operators to involve in a service ecosystem. Open innovation helps operators to accelerate the speed of innovations and fill the gaps in their internet skills. Usage of open APIs possess a privacy risk [11], but on the other hand, operators enjoy the trust and esteem of the public [63]. The actual business impact of the Open Telco remains to be seen, and no concrete conclusions can be made before large scale deployments are available.

The second research question asked how cloud computing can be adapted to mobile networks. The observations are that SaaS and PaaS service models offer promising opportunities despite the security concerns, while the IaaS service model still raises questions. According to a study [84], the IaaS option remains beyond the performance requirements of the real-time, carrier grade systems, but this dissertation indicates that at least cloud database systems can provide high performance. Furthermore, mobile networks support a large number of applications, including messaging, video streaming and encoding, browsing and customer care that do not require strict SLA [2].

Cloud computing also enables sharing of network functions [124]. Virtual operators already share various support services with each other, but in the future sharing initiatives can be extended to core functions such as voice machinery as well [187]. Besides, the hybrid cloud deployment model provides in financial terms [195] a noteworthy alternative to pure private or public IaaS clouds although the challenges in cloud interoperability [48] still restrict the applicability. This problem, however, is almost invisible in the mobile networks. Voice and text messaging work over national borders thanks to roaming agreements. In addition, the number portability feature has enabled easy switching of operator. These advanced functions have not necessarily always favoured the operator business targets.

Finally, operators suffer from a heavy regulation burden [134], a fact which currently restricts their operations. In addition, the investment period of mobile networks is often up to 30 years, which means that it can be difficult to update the installed base to the cloud grade. For that reason, novel open API and cloud based solutions should not interfere with older systems, and also backward compatibility must be ensured [180].

Table 4.1 summarizes the operator status using traditional Strength, Weakness, Opportunity and Threat (SWOT) analysis.

Table 4.1. SWOT analysis.

Strengths - Customer base - Financial status - Standardization - Trust	Weaknesses - High churn - Innovation speed - Installed base - Internet skills
Opportunities - Cloud computing - Network sharing - Open Innovation and APIs - Two-sided markets	Threats - Cloud SLA & security - Decline of ARPU - Internet competition - Regulation

4.1 Strategic Recommendations

Mobile operators can utilize the observations of this dissertation in several ways. Firstly, the correct time to change the strategic game from chess to poker is now. Like in chess, the rules of the telecom business have been stable for decades, but the weak signals indicate that markets are rapidly changing towards a poker game where opponents may use bluff. Operators must respond to challenges and turn them to novel opportunities. Already Joseph Schumpeter [166] promoted the importance of innovation, and the Open Telco concept and cloud computing can help operators to innovate new products, production methods, supply sources, markets and business organizations. Operators have learned that they cannot only innovate in-house, but so far their main corrective action has focused on decreasing the R&D staff. This research encourages operators to compensate for the lack of internal resources with open innovation [37].

Secondly, the Open Telco concept provides the first step towards a Network as a Service (NaaS) paradigm. It creates a new service model that attracts developers into the operator ecosystem. Operators must maximize the benefits of open innovation, while simultaneously minimizing the risks [150]. One of those risks is cannibalization of existing services such as voice and text messaging with substitutes enabled by the Open Telco system. However, in this setup operators have full control of services, which is not the case with the services offered by the Over The Top (OTT) competitors.

Moreover, Open Telco paves the way for a two-sided platform, which can be implemented in a Platform as a Service (PaaS) fashion. In the next phase, operators must focus on creating a successful and attractive ecosystem, where they are the dominant players. Operators should provide added value to the ecosystem by offering a so-called smart pipe, which developers can access through well established and secure methods. In addition to messaging, location and payment APIs, mobile operators can aim higher and provide sub networks and bandwidth using the NaaS approach.

Finally, operators are used to the Software as a Service (SaaS) service model through call and messaging services. However, more opportunities can be found in the PaaS layer. The SaaS service model is typical to the one-sided business model, but the PaaS service model enables a two-sided business model, which offers the highest autonomy and attraction to developers [69]. The Open Telco platform is a good example of a service that can be well run in the PaaS service model. Depending on the market status, the platform can be shared by several operators or they can lower the initial risk with the virtual broker concept presented in this dissertation. In any case the market fragmentation should be avoided and global markets should be targeted instead.

The research results show that cloud computing can meet the strictest performance requirements, and thus even the Infrastructure as a Service (IaaS) service model can be considered for the execution of carrier grade telecom services. However, real systems require the SS7 signalling support and its implementation in a cloud environment can be difficult. The telecom industry has doubts in trusting cloud providers, too, and to overcome these concerns cloud providers should offer telecom grade cloud computing systems. The hybrid cloud deployment model can be one answer to the challenges.

On the other hand, this research indicates that even the current mobile networks include various ICT elements that can be based on cloud computing principles. One interesting approach is to utilize clouds in network sharing initiatives. Except for Mobile Virtual Network Operators (MVNOs), most operators currently build their networks independently, although the network sharing approach could save costs. However, unlike in the Amazon case, mobile operators have to take into account the installed base and regulative obligations.

4.2 Summary of Contributions

The Open Telco concept summarizes the PaaS type of platform that connects the software and infrastructure layers together, offering network assets through secure and open APIs to developers, content creators and service providers. The platform can form a two-sided platform that offers additional revenues to operators. More importantly, the Open Telco platform accelerates service creativeness, adding the third segment, the long tail, to the accessible markets. In addition, the virtual broker approach simplifies the agreement processes and eases operator cooperation, lowering the barrier to starting a novel service.

Cloud computing as such can bring considerable cost reductions for the network infrastructure. The research indicates that the IaaS service model, including especially the database services, can reach carrier grade performance level, but in practice the best results can be achieved in the SaaS and PaaS service models. The deployment model of the hybrid cloud provides the cost optimal solution, but due to current interoperability and security challenges, the hybrid cloud is not yet a mature option.

Although the research focus here is on the mobile network sector, the results are applicable to most ICT industries. The ultimate goal of cloud computing is to move towards utility computing. It means that the customers of the clouds can let the cloud providers compete with each other in prices, technology and features, enabling an easy comparison and switch of providers.

4.3 Limitations

The biggest concern is that, so far, the mobile sector has carried out only a few large-scale open API or cloud computing trials, to say nothing of product launches. The result is that a very limited amount of research data is available, which means that it is necessary to utilize concept studies, case studies from other industries, surveys, mathematical analysis and proof of concepts. Research community has difficulties to implement real mobile systems in cloud environment due to complex network architectures and partly non-IP signalling systems.

However, the success of internet services such as Amazon offers strong evidence that open innovation and continuous experiment are the key drivers for the success. Also, numerous examples of internet startups have shown that cloud computing has paved the way to faster growth together with reasonable performance. Whether this performance is enough for the telecommunication networks, is the topic of another debate. The same comment applies to the security concerns. On the other hand, the carrier grade clouds may support the extreme performance and security requirements of the mobile networks.

4.4 Future Work

This dissertation acts as an introduction to a novel area of research, which is very much in its early phases. Mobile networks are moving towards a new generation, the 4G or LTE networks, which also provides novel opportunities for disruptive innovations. Although the next generation will mainly result in upgrades in the radio interface and transport networks, solutions for the service domain can be enhanced as well. The new network architecture that combines voice and data services under one technology paves the way for new challengers, e.g. greenfield operators who do not have the burden of an installed base.

The new networks will increase the bit rates but, unfortunately, also add to the concerns about energy consumption. It is estimated that in terms of energy consumption, LTE will be 23 times worse than WiFi, and even poorer when compared to 3G [88]. Moreover, the GreenTouch project[1] forecasts that if the usage of mobile data increases until 2020 in accordance with predictions, and if the energy efficiency of the networks remains on the current level, operators will have major difficulties in profitability due to high energy costs [155]. In general, it is also easy to see that the energy efficiency of data centres will be a major research topic in the coming years [27, 71, 110].

Finally, the main challenge for the cloud providers is the lack of interoperability [48]. Cloud customers have limited opportunities to avoid the lock-in problem. However, research in this area has started, and substantially more results can be expected in the future. The role of standardization remains open in a market that is used to working according to de facto standards. However, various standardization bodies are working to create a set of widely accepted cloud computing standards, which may resolve the interoperability challenge in the long run [85]. In this matter, the role of the EU and the US markets and specifically the US Government, as one of the largest customers of cloud services, is very decisive.

[1] http://www.greentouch.org/

5. Conclusions

Mobile operators run one of the most profitable ICT business sectors. Until today, they have avoided the pitfalls of the incumbent enterprises, but the future can change rapidly. This dissertation presents two complementing propositions to improve the operator business case. Firstly, the Open Telco concept, utilizing open innovation and open APIs, helps developers to utilize operators' hidden assets for creating novel services and innovations for their customers. The platform, furthermore, provides a cross domain architecture that enables simple end user, developer and operator interactions, which leads from one-sided to two-sided markets, and possibly to a new Open Telco ecosystem.

Secondly, cloud computing can enhance several ways operator networks. The Software as a Service (SaaS) and Platform as a Service (PaaS) service models offer a good match to most operator service requirements, including the Open Telco platform itself, but also the Infrastructure as a Service (IaaS) service model can be applied to mobile platforms. The SaaS service model suits well to user-interactive, PaaS to shared and IaaS to computation-heavy applications. Initiatives towards network sharing also increase interest in the cloud computing approach. However, current cloud interoperability challenges must first be solved, a process which later may lead to utility computing.

The dissertation uses various research methodologies, but due to the broadness of the research area, the findings provide only preliminary results. The presented ideas are still quite novel and they have not been thoroughly tested in real networks with real, paying customers. The introduction of Long Term Evolution (LTE) unifies the circuit and packet switched networks, bringing a natural disruption, a fact which also offers opportunities to novel innovations. The next step requires determined and prompt execution from operators and their vendors.

Bibliography

[1] 3GPP. Network architecture TS 23.002 v4.8.0 (2003-06).

[2] 3GPP. Technical performance objectives TS 43.005 v9.0.0 (2010-02).

[3] ABADI, D. J. Data management in the cloud: Limitations and opportunities. *IEEE Data Engineering Bulletin 32*, 1 (March 2009), 3–12.

[4] ADAMIC, L. A., AND HUBERMAN, B. A. Zipf's law and the Internet. *Glottometrics*, 3 (2002), 143–150.

[5] ALCOTT, B. Jevons' paradox. *Ecological Economics 54*, 1 (2005), 9–21.

[6] ALI-ELDIN, A., KIHL, M., TORDSSON, J., AND ELMROTH, E. Efficient provisioning of bursty scientific workloads on the cloud using adaptive elasticity control. In *Proceedings of the 3rd workshop on Scientific Cloud Computing Date* (New York, NY, USA, 2012), ScienceCloud '12, ACM, pp. 31–40.

[7] ALLEE, V. Reconfiguring the value network. *Journal of Business Strategy 21*, 4 (2000), 1–6.

[8] ANDERSON, C. *The Long Tail: How Endless Choice is Creating Unlimited Demand*. Random House Business Books, 2006.

[9] ANDERSON, C. *Free: The Future of a Radical Price*. Random House Business Books, 2009.

[10] ARDAGNA, D., CASOLARI, S., AND PANICUCCI, B. Flexible distributed capacity allocation and load redirect algorithms for cloud systems. In *Proceedings of the 4th International Conference on Cloud Computing* (Washington, DC, USA, July 2011), Cloud 2011, IEEE Computer Society, pp. 163–170.

[11] ARJAMAA, H. Privacy implications of public APIs. Master's thesis, Helsinki University of Technology, Faculty of Information and Natural Sciences, Finland, May 2010.

[12] ARMBRUST, M., FOX, A., GRIFFITH, R., JOSEPH, A. D., KATZ, R., KONWINSKI, A., LEE, G., PATTERSON, D. A., RABKIN, A., STOICA, I., AND ZAHARIA, M. Above the clouds: A Berkeley view of cloud computing. Technical report UCB/EECS-2009-28, Feb 2009.

[13] BADGER, L., GRANCE, T., PATT-CORNER, R., AND VOAS, J. Draft Cloud Computing Synopsis and Recommendations, Recommendations of the National Institute of Standards and Technology. Available from `http://csrc.nist.gov/publications/drafts/800-146/Draft-NIST-SP800-146.pdf`, May 2011. NIST Special Publication 800-146.

[14] BALLON, P. *Control and value in mobile communications: A political economy of the reconfiguration of business models in the European mobile industry*. PhD thesis, Vrije Universiteit Brussel, Belgium, 2009.

[15] BARKER, S. K., AND SHENOY, P. Empirical evaluation of latency-sensitive application performance in the cloud. In *Proceedings of the first annual ACM SIGMM conference on Multimedia systems* (New York, NY, USA, 2010), MMSys '10, ACM, pp. 35–46.

[16] BARROS, A. P., AND DUMAS, M. The rise of web service ecosystems. *IT Professional 8* (2006), 31–37.

[17] BASOLE, R. C., AND KARLA, J. On the evolution of mobile platform ecosystem structure and strategy. *Business Information Systems Engineering 3*, 5 (2011), 313–322.

[18] BEDHIAF, I., CHERKAOUI, O., AND PUJOLLE, G. Third-generation virtualized architecture for the MVNO context. *Annals of Telecommunications 64* (2009), 339–347.

[19] BELL, S. *Learning With Information Systems: Learning Cycles in Information Systems Development*. Routledge Studies in Information and Library Management Systems. Routledge, 1996.

[20] BIRMAN, K., CHOCKLER, G., AND VAN RENESSE, R. Toward a cloud computing research agenda. *SIGACT News 40*, 2 (June 2009), 68–80.

[21] BLENNERUD, G. The recipe for mobile broadband profitability. *Ericsson Business Review 3* (2009), 46–50.

[22] BOHLIN, E. Business models and financial impacts of future mobile broadband networks. *Telematics and Informatics 24*, 3 (Aug 2007), 217–237.

[23] BOUDREAU, K. J., AND LAKHANI, K. R. How to manage outside innovation. *MIT Sloan Management Review 50*, 4 (2009), 69–76.

[24] BOUWMAN, H., DE VOS, H., AND HAAKER, T. E. *Mobile Service Innovation and Business Models*. Springer, 2008.

[25] BOX, G. E. P., JENKINS, G. M., AND REINSEL, G. C. *Time Series Analysis, Forecasting and Control, 3rd ed.* Prentice Hall, 1994.

[26] BREWER, E. A. Towards robust distributed systems. In *Proceedings of the Nineteenth annual symposium on Principles of distributed computing* (New York, NY, USA, 2000), PODC '00, ACM, p. 7. Keynote.

[27] BRYNJOLFSSON, E., HOFMANN, P., AND JORDAN, J. Cloud computing and electricity: beyond the utility model. *Commun. ACM 53*, 5 (May 2010), 32–34.

[28] BUYYA, R., YEO, C. S., VENUGOPAL, S., BROBERG, J., AND BRANDIC, I. Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility. *Future Gener. Comput. Syst. 25*, 6 (June 2009), 599–616.

[29] CÁCERES-EXPÓSITO, J.-A., HIERRO-SUREDA, J.-J., VAQUERO-GONSÁLEZ, L. M., AND IGLESIA-MEDINA DE LA, F. Key factors for the adoption of cloud technologies by telco operators. *Upgrade 11*, 3 (2010), 30–32.

[30] CAMARÁN, C., AND DE MIGUEL, D. Mobile virtual network operator (MVNO) basics: What is behind this mobile business trend. Technical report, Valoris, Oct 2008.

[31] CASTRO, M., AND LISKOV, B. Practical byzantine fault tolerance and proactive recovery. *ACM Trans. Comput. Syst. 20*, 4 (Nov 2002), 398–461.

[32] CERVIÑO, J., ESCRIBANO, F., RODRÍGUEZ, P., TRAJKOVSKA, I., AND SALVACHÚA, J. Videoconference capacity leasing on hybrid clouds. In *Proceedings of the 4th International Conference on Cloud Computing* (Washington, DC, USA, July 2011), Cloud 2011, IEEE Computer Society, pp. 340–347.

[33] CERVIÑO, J., RODRÍGUEZ, P., TRAJKOVSKA, I., MOZO, A., AND SALVACHÚA, J. Testing a cloud provider network for hybrid P2P and cloud streaming architectures. In *Proceedings of the 4th International Conference on Cloud Computing* (Washington, DC, USA, July 2011), Cloud 2011, IEEE Computer Society, pp. 356–363.

[34] CHANG, F., DEAN, J., GHEMAWAT, S., HSIEH, W. C., WALLACH, D. A., BURROWS, M., CHANDRA, T., FIKES, A., AND GRUBER, R. E. Bigtable: a distributed storage system for structured data. In *Proceedings of the 7th symposium on Operating systems design and implementation* (Berkeley, CA, USA, 2006), OSDI '06, USENIX Association, pp. 205–218.

[35] CHANG, F., REN, J., AND VISWANATHAN, R. Optimal resource allocation in clouds. In *Proceedings of the 3rd International Conference on Cloud Computing* (Washington, DC, USA, 2010), Cloud 2010, IEEE Computer Society, pp. 418–425.

[36] CHANG, T.-R., AND KAASINEN, E. Three user-driven innovation methods for co-creating cloud services. In *Proceedings of the 13th IFIP TC 13 international conference on Human-computer interaction - Volume Part IV* (Berlin, Heidelberg, 2011), INTERACT'11, Springer-Verlag, pp. 66–83.

[37] CHESBROUGH, H. *Open Innovation: The New Imperative for Creating and Profiting from Technology*. Harvard Business School Press, 2003.

[38] CHRISTENSEN, C. M. *The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail*. The Management Of Innovation And Change Series. Harvard Business School Press, 1997.

[39] CHRISTENSEN, C. M. *Seeing What's Next - Using the Theories of Innovation to Predict Industry Change*. Harvard Business School Press, 2004.

[40] CISCO. Cisco visual networking index: Global mobile data traffic forecast update, 2010-2015. White Paper published 01.02.2011, available from http://newsroom.cisco.com/ekits/Cisco_VNI_Global_Mobile_Data_Traffic_Forecast_2010_2015.pdf, March 2011.

[41] CLARK, K. B. The interaction of design hierarchies and market concepts in technological evolution. *Research Policy 14*, 5 (October 1985), 235–251.

[42] COOPER, B. F., SILBERSTEIN, A., TAM, E., RAMAKRISHNAN, R., AND SEARS, R. Benchmarking cloud serving systems with YCSB. In *Proceedings of the 1st symposium on Cloud computing* (New York, NY, USA, 2010), SoCC '10, ACM, pp. 143–154.

[43] CROOKES, J. Multiservice billing system - a platform for the future. *BT Technology Journal 15*, 1 (Jan 1997), 98–113.

[44] CRYANS, J.-D., APRIL, A., AND ABRAN, A. Criteria to compare cloud computing with current database technology. In *Proceedings of the International Conferences on Software Process and Product Measurement* (Berlin, Heidelberg, 2008), IWSM/Metrikon/Mensura '08, Springer-Verlag, pp. 114–126.

[45] DANIELS, J. Server virtualization architecture and implementation. *Crossroads 16*, 1 (Sep 2009), 8–12.

[46] DE REUVER, M., BOUWMAN, H., AND MACINNES, I. Business model dynamics: a case survey. *Journal of Theoretical and Applied Electronic Commerce Research 4*, 1 (Apr 2009), 1–11.

[47] DECANDIA, G., HASTORUN, D., JAMPANI, M., KAKULAPATI, G., LAKSHMAN, A., PILCHIN, A., SIVASUBRAMANIAN, S., VOSSHALL, P., AND VOGELS, W. Dynamo: Amazon's highly available key-value store. In *Proceedings of the twenty-first SIGOPS symposium on Operating systems principles* (New York, NY, USA, 2007), SOSP '07, ACM, pp. 205–220.

[48] DILLON, T., WU, C., AND CHANG, E. Cloud computing: Issues and challenges. In *Proceedings of the 24th International Conference on Advanced Information Networking and Applications* (Washington, DC, USA, April 2010), AINA, IEEE Computer Society, pp. 27–33.

[49] DINDA, P. A., AND O'HALLARON, D. R. Host load prediction using linear models. *Cluster Computing 3*, 4 (Oct 2000), 265–280.

[50] DUVANDER, A. 6,000 APIs: It's business, it's social and it's happening quickly. Blog, May 22, 2012, available from `http://blog.programmableweb.com/2012/05/22/6000-apis-its-business-its-social-and-its-happening-quickly/`, 2012.

[51] ECONOMIDES, N., AND KATSAMAKAS, E. Two-Sided Competition of Proprietary vs. Open Source Technology Platforms and the Implications for the Software Industry. *SSRN eLibrary* (2005).

[52] EISENHARDT, K. M. Building theories from case study research. *Academy of Management Review 14*, 4 (1989), 532–550.

[53] EISENMANN, T. R., PARKER, G., AND VAN ALSTYNE, M. W. Strategies for two-sided markets. *Harvard Business Review 84*, 10 (2006), 92–101.

[54] EISENMANN, T. R., PARKER, G., AND VAN ALSTYNE, M. W. Opening Platforms: How, When and Why? *HBS Working Paper Number: 09-030* (August 2008).

[55] ETRO, F. The economic impact of cloud computing on business creation, employment and output in Europe. *Review of Business and Economics 54*, 2 (2009), 1–40.

[56] ETSI TISPAN. Next generation network ES 283.026 v1.1.1 (2009).

[57] EVANS, D. S., AND SCHMALENSEE, R. *Catalyst Code: The Strategies Behind the World's Most Dynamic Companies*. Harvard Business School Press, 2007.

[58] EVANS, D. S., AND SCHMALENSEE, R. Markets with two-sided platforms. *Issues in Competition Law and Policy (ABA Section of Antitrust Law) 1* (October 2008), 667–693. Chapter 28.

[59] FEIJOO, C., PASCU, C., MISURACA, G., AND LUSOLI, W. The next paradigm shift in the mobile ecosystem : Mobile social computing and the increasing relevance of users. *Computing 58*, 75 (2009), 57–77.

[60] FEITELSON, D. G. *Workload Modeling for Computer Systems Performance Evaluation*. School of Computer Science and Engineering, The Hebrew University of Jerusalem, 91904 Jerusalem, Israel, 2011.

[61] FINNIE, G. Amazon's lessons for telcos. Blog published 15.05.2009, available from `http://www.lightreading.com/document.asp?doc_id=176804`, May 2009.

[62] FRANK, L., LUOMA, E., MAZHELIS, O., PULKKINEN, M., AND TYRVÄINEN, P. Software business in the telecommunications sector. In *Telecommunication Economics*, A. Hadjiantonis and B. Stiller, Eds., vol. 7216 of *Lecture Notes in Computer Science*. Springer Berlin / Heidelberg, 2012, pp. 148–155.

[63] GABRIELSSON, J., HUBERTSSON, O., MÁS, I., AND SKOG, R. Cloud computing in telecommunications. *Ericsson Review 1* (2010), 29–33.

[64] GAYNOR, M. *Network Services Investment Guide. Maximizing ROI in Uncertain Times*. John Wiley & Son, 2003.

[65] GHEMAWAT, S., GOBIOFF, H., AND LEUNG, S.-T. The Google file system. In *Proceedings of the nineteenth ACM symposium on Operating systems principles* (New York, NY, USA, 2003), SOSP '03, ACM, pp. 29–43.

[66] GILBERT, S., AND LYNCH, N. Brewer's conjecture and the feasibility of consistent, available, partition-tolerant web services. *SIGACT News 33*, 2 (2002), 51–59.

[67] GLASOW, P. A. Fundamentals of survey research methodology. *Distribution* (2005), 1–5.

[68] GOFF, M. K. *Network distributed computing: fitscapes and fallacies*. Professional technical reference. Prentice Hall PTR, 2004.

[69] GONÇALVES, V., AND BALLON, P. Adding value to the network: Mobile operators' experiments with software-as-a-service and platform-as-a-service models. *Telematics and Informatics 28*, 1 (2011), 12–21.

[70] GONG, Z., GU, X., AND WILKES, J. PRESS: PRedictive Elastic ReSource Scaling for cloud systems. In *Proceedings of the 6th International Conference on Network and Service Management* (445 Hoes Lane, Piscataway, NJ 08854, USA, 2010), CNSM 2010, IEEE, pp. 9–16.

[71] GREENBERG, A., HAMILTON, J., MALTZ, D. A., AND PATEL, P. The cost of a cloud: research problems in data center networks. *SIGCOMM Comput. Commun. Rev. 39*, 1 (2009), 68–73.

[72] GRODAL, S. Towards a dynamic model of networks and innovation. In *Proceedings of the DRUID Summer Conference 2004 on Industrial Dynamics, Innovation and Development* (2004), pp. 1–29.

[73] GRÖNROOS, C. *Service management and marketing: a customer relationship management approach*. John Wiley & Sons, Ltd, 2000.

[74] GUPTA, N. *Enabling High Performance HLR Solutions*. Sun Microsystems, Inc., 2006.

[75] HAAKER, T., FABER, E., AND BOUWMAN, H. Balancing customer and network value in business models for mobile services. *International Journal of Mobile Communications 4*, 6 (July 2006), 645–661.

[76] HACIGUMUS, H., IYER, B., AND MEHROTRA, S. Providing Database as a Service. In *Proceedings of the 18th International Conference on Data Engineering* (Washigton, DC, USA, 2002), ICDE '02, IEEE Computer Society, pp. 29–38.

[77] HAJJAT, M., SUN, X., SUNG, Y.-W. E., MALTZ, D., RAO, S., SRIPANIDKULCHAI, K., AND TAWARMALANI, M. Cloudward bound: planning for beneficial migration of enterprise applications to the cloud. In *Proceedings of the ACM SIGCOMM 2010* (New York, NY, USA, 2010), SIGCOMM '10, ACM, pp. 243–254.

[78] HAMILTON, J. Porter's: 'Strategy And The Internet' Revisited. In *Proceedings of the International Conference on Service Systems and Service Management* (2007), IEEE, pp. 1–6.

[79] HAMILTON, J. Cloud computing economies of scale, 2010. Keynote at AWS Genomics & Cloud Computing Workshop, Seattle, WA, 08.06.2010, available from http://www.mvdirona.com/jrh/TalksAndPapers/JamesHamilton_GenomicsCloud20100608.pdf.

[80] HARMS, R., AND YAMARTINO, M. The economics of the cloud. White paper, Microsoft Corporation, Nov 2010.

[81] HARNO, J. Techno-economic analysis of beyond 3G mobile technology alternatives. In *Proceedings of the 6th Conference on Telecommunication Techno-Economics* (Piscataway, NJ, USA, 2007), CTTE, IEEE, pp. 1–9.

[82] HARTIGH DEN, E., TOL, M., AND VISSCHER, W. The health measurement of a business ecosystem. In *Proceedings of the ECCON 2006 Annual meeting* (2006), pp. 1–39.

[83] HEKKERT, M., SUURS, R., NEGRO, S., KUHLMANN, S., AND SMITS, R. Functions of innovation systems: A new approach for analysing technological change. *Technological Forecasting and Social Change 74*, 4 (2007), 413–432.

[84] HILL, Z., AND HUMPHREY, M. A quantitative analysis of high performance computing with Amazon's EC2 infrastructure: The death of the local cluster? In *Proceedings of the 10th IEEE/ACM International Conference on Grid Computing* (Washington, DC, USA, 2009), Grid 2009, IEEE Computer Society, pp. 26–33.

[85] HOGAN, M., LIU, F., SOKOL, A., AND TONG, J. NIST Cloud Computing Standards Roadmap. Available from http://www.nist.gov/customcf/get_pdf.cfm?pub_id=909024, July 2011. NIST Cloud Computing Standards Roadmap Working Group, NIST Special Publication 500-291.

[86] HOIKKANEN, A. Economics of 3G Long-Term Evolution: the Business Case for the Mobile Operator. In *Proceedings of the IFIP International Conference on Wireless and Optical Communications Networks* (2007), WOCN '07, IEEE, pp. 1–5.

[87] HOWE, J. The rise of crowdsourcing. *Wired* (June 2006).

[88] HUANG, J., QIAN, F., GERBER, A., MAO, Z. M., SEN, S., AND SPATSCHECK, O. A close examination of performance and power characteristics of 4G LTE networks. In *Proceedings of the 10th international conference on Mobile systems, applications, and services* (New York, NY, USA, 2012), MobiSys '12, ACM, pp. 225–238.

[89] IANSITI, M., AND LEVIEN, R. Strategy as ecology. *Harvard Business Review 82*, 3 (2004), 68–78.

[90] IBM. *Telecom Application Transaction Processing (TATP) Benchmark Description*, 2009. version 1.0.

[91] ISAAC, S., AND MICHAEL, W. B. *Handbook in Research and Evaluation: A Collection of Principles, Methods, and Strategies Useful in the Planning, Design, and Evaluation of Studies in Education and the Behavioral Sciences*. EdITS, 1997.

[92] ITU-D. The world in 2010 - the rise of 3G. White paper, Oct 2010.

[93] JAOKAR, A., AND FISH, T. *Mobile Web 2.0: The Innovator's Guide to Developing and Marketing Next Generation Wireless/Mobile Applications*. Futuretext, 2006.

[94] JORDAN, S. The application of net neutrality to wireless networks based on network architecture. *Policy & Internet 2* (2010), 125–158.

[95] KAUFMAN, L. M. Data security in the world of cloud computing. *IEEE Security and Privacy 7* (July 2009), 61–64.

[96] KEAHEY, K., TSUGAWA, M., MATSUNAGA, A., AND FORTES, J. Sky computing. *IEEE Internet Computing 13* (2009), 43–51.

[97] KHAJEH-HOSSEINI, A., GREENWOOD, D., SMITH, J. W., AND SOMMERVILLE, I. The cloud adoption toolkit: Supporting cloud adoption decisions in the enterprise. *CoRR abs/1008.1900* (2010).

[98] KIKUCHI, S., AND MATSUMOTO, Y. Performance modeling of concurrent live migration operations in cloud computing systems using PRISM probabilistic model checker. In *Proceedings of the 4th International Conference on Cloud Computing* (Washington, DC, USA, 2011), Cloud 2011, IEEE Computer Society, pp. 49–56.

[99] KILKKI, K. A practical model for analyzing long tails. *First Monday 12*, 5 (2007).

[100] KLEINROCK, L. A vision for the Internet. *ST Journal for Research 2*, 1 (November 2005), 4–5.

[101] KUECHLER, B., AND VAISHNAVI, V. Extending prior research with design science research: two patterns for DSRIS project generation. In *Proceedings of the 6th international conference on Service-oriented perspectives in design science research* (Berlin, Heidelberg, 2011), DESRIST'11, Springer-Verlag, pp. 166–175.

[102] KUO, Y.-F., AND YU, C.-W. 3G telecommunication operators challenges and roles: A perspective of mobile commerce value chain. *Technovation The international journal of technological innovation entrepreneurship and technology management 26*, 12 (2006), 1347–1356.

[103] LAAKSO, K., RUBIN, A., AND LINTURI, H. Delphi method analysis: The role of regulation in the mobile operator business in Finland. In *Proceedings of the Technology Management for Global Economic Growth* (2010), PICMET '10, Emerald Group Publishing Limited, pp. 1–7.

[104] LAKSHMAN, A., AND MALIK, P. Cassandra: a decentralized structured storage system. *SIGOPS Oper. Syst. Rev. 44*, 2 (2010), 35–40.

[105] LEAVITT, N. Is cloud computing really ready for prime time? *Computer 42*, 1 (Jan 2009), 15–20.

[106] LEWIN, R. *Complexity: Life at the Edge of Chaos*. The University of Chicago Press, 1999.

[107] LI, F., AND WHALLEY, J. Deconstruction of the telecommunications industry: From value chain to value networks. *Telecommunications Policy 26*, 9-10 (2002), 451–472.

[108] LI, H., ZHONG, L., LIU, J., LI, B., AND XU, K. Cost-effective partial migration of VoD services to content clouds. In *Proceedings of the 4th International Conference on Cloud Computing* (Washington, DC, USA, 2011), Cloud 2011, IEEE Computer Society, pp. 203–210.

[109] LIAO, Y. *Examining Effects of Virtual Machine Settings on Voice over Internet Protocol in a Private Cloud Environment*. PhD thesis, Indiana State University, Terre Haute, Indiana, USA, 2011.

[110] LIM, S.-H., SHARMA, B., TAK, B. C., AND DAS, C. A dynamic energy management scheme for multi-tier data centers. In *Proceedings of the IEEE International Symposium on Performance Analysis of Systems and Software* (Washington, DC, USA, 2011), ISPASS, IEEE Computer Society, pp. 257–266.

[111] LIN, C.-F., LEU, M.-C., CHANG, C.-W., AND YUAN, S.-M. The study and methods for cloud based CDN. *Cyber-Enabled Distributed Computing and Knowledge Discovery, International Conference on 0* (2011), 469–475.

[112] LIU, K., AND XU, K. OAuth Based Authentication and Authorization in Open Telco API. In *Proceedings of International Conference on Computer Science and Electronics Engineering* (Washington, DC, USA, March 2012), vol. 1 of *ICCSEE*, IEEE Computer Society, pp. 176–179.

[113] LORETO, S., MECKLIN, T., OPSENICA, M., AND RISSANEN, H.-M. Service broker architecture: Location business case and mashups. *Communications Magazine, IEEE 47*, 4 (April 2009), 97–103.

[114] LOWENSTEIN, M. Is mobile ready for the cloud? Blog, September 26, 2011, available from `http://www.fiercewireless.com/story/lowensteins-view-mobile-ready-cloud/2011-09-26`, 2011.

[115] MAITLAND, C. F., BAUER, J. M., AND WESTERVELD, R. The European market for mobile data: evolving value chains and industry structures. *Telecommunications Policy 26*, 9–10 (2002), 485–504.

[116] MÄKINEN, S., JÄRVENPÄÄ, H. M., USKALI, T., AND OJALA, J. Spotting weak signals considering new technological innovations: An empirical search for appropriate sources. *Innovation Journalism 4*, 3 (July 2007), 1–15.

[117] MAO, M., AND HUMPHREY, M. Auto-scaling to minimize cost and meet application deadlines in cloud workflows. In *Proceedings of the 2011 International Conference for High Performance Computing, Networking, Storage and Analysis* (New York, NY, USA, 2011), SC '11, ACM, pp. 49:1–49:12.

[118] MARTENS, B., AND TEUTEBERG, F. Decision-making in cloud computing environments: A cost and risk based approach. *Information Systems Frontiers* (Jul 2011), 1–23.

[119] MATTESS, M., VECCHIOLA, C., AND BUYYA, R. Managing peak loads by leasing cloud infrastructure services from a spot market. In *Proceedings of the 12th International Conference on High Performance Computing and Communications* (Washington, DC, USA, 2010), HPCC '10, IEEE Computer Society, pp. 180–188.

[120] MAXWELL, E. E. Open standards, open source, and open innovation: Harnessing the benefits of openness. *Innovations* (June 2006), 119–176.

[121] MAZHELIS, O., AND TYRVÄINEN, P. Role of data communications in hybrid cloud costs. In *Proceedings of the 37th EUROMICRO Conference on Software Engineering and Advanced Applications* (Washington, DC, USA, 2011), IEEE Computer Society, pp. 138–145.

[122] MAZZUCCO, M., DYACHUK, D., AND DETERS, R. Maximizing cloud providers' revenues via energy aware allocation policies. In *Proceedings of the 3rd International Conference on Cloud Computing* (Washington, DC, USA, 2010), Cloud 2010, IEEE Computer Society, pp. 131–138.

[123] MCGRATH, M. E. *Product Strategy for High Technology Companies*. McGraw-Hill, 2001.

[124] MELL, P., AND GRANCE, T. The NIST Definition of Cloud Computing, Recommendations of the National Institute of Standards and Technology. Available from http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf, Sep 2011. NIST Special Publication 800-145.

[125] MISHRA, M., AND SAHOO, A. On theory of VM placement: Anomalies in existing methodologies and their mitigation using a novel vector based approach. In *Proceedings of the 4th International Conference on Cloud Computing* (Washington, DC, USA, 2011), Cloud 2011, IEEE Computer Society, pp. 275–282.

[126] MITLETON-KELLY, E. Ten principles of complexity & enabling infrastructures. In *Complex Systems and Evolutionary Perspectives on Organisations: The Application of Complexity Theory to Organisations*, E. Mitleton-Kelly, Ed. Pergamon, 2003, pp. 23–50.

[127] MOORE, J. F. Predators and prey: A new ecology of competition. *Harvard Business Review 71*, 3 (May 1993), 75–86.

[128] MORENO-VOZMEDIANO, R., MONTERO, R. S., AND LLORENTE, I. M. Elastic management of cluster-based services in the cloud. In *Proceedings of the 1st workshop on Automated control for datacenters and clouds* (New York, NY, USA, 2009), ACDC '09, ACM, pp. 19–24.

[129] MOULY, M., AND PAUTET, M.-B. *The GSM System for Mobile Communications*. Telecom Publishing, 1992.

[130] MULLIGAN, C. E. A. Open API standardization for the NGN platform. *Communications Magazine 47*, 5 (May 2009), 108–113.

[131] MURPHY, M. Platform-as-a-Service for Telcos. Keynote at Cloud Asia, Singapore, May 5, 2010, available from http://cloudasia.ngp.org.sg/2010/programs_track2.php, 2010.

[132] MUTANEN, U.-M. The invisible tail: How free product codes can democratize the global market. Blog, October 27, 2005, available from http://www.hobbyprincess.com/2005/10/the_invisible_t.html, 2005.

[133] NIEBERT, N., KHAYAT, I. E., BAUCKE, S., KELLER, R., REMBARZ, R., AND SACHS, J. Network virtualization: A viable path towards the future internet. *Wirel. Pers. Commun. 45*, 4 (June 2008), 511–520.

[134] NOAM, E. M. Regulation 3.0 for Telecom 3.0. *Telecommunications Policy 34*, 1-2 (Feb 2010), 4–10.

[135] O'REILLY, T. Static on the dream phone. *The New York Times* (Dec 2007). published 15.12.2007.

[136] O'REILLY III, C., AND TUSHMAN, M. L. The ambidextrous organization. *Harvard Business Review 82*, 4 (2004), 74–81.

[137] OSTERWALDER, A., AND PIGNEUR, Y. *Business Model Generation: A Handbook for Visionaries, Game Changers, and Challengers*. Wiley Desktop Editions. Wiley, 2010.

[138] PACHECO-SANCHEZ, S., CASALE, G., SCOTNEY, B. W., MCCLEAN, S. I., PARR, G. P., AND DAWSON, S. Markovian workload characterization for QoS prediction in the cloud. In *Proceedings of the 4th International Conference on Cloud Computing* (Washington, DC, USA, 2011), Cloud 2011, IEEE Computer Society, pp. 147–154.

[139] PÄIVÄRINTA, R. Applicability of NoSQL databases to mobile networks. Master's thesis, Aalto University, School of Science and Technology, Finland, Dec 2010.

[140] PALMER, N., KEMP, R., KIELMANN, T., AND BAL, H. Ibis for mobility: solving challenges of mobile computing using grid techniques. In *Proceedings of the 10th workshop on Mobile Computing Systems and Applications* (New York, NY, USA, 2009), HotMobile '09, ACM, pp. 17:1–17:6.

[141] PARKER, G., AND VAN ALSTYNE, M. W. Two-sided network effects: A theory of information product design. *Management Science 51*, 10 (2005), 1494–1504.

[142] PATEL, P., RANABAHU, A., AND SHETH, A. Service level agreement in cloud computing. *Cloud Workshops at OOPSLA09* (2009), 1–10.

[143] PELTONIEMI, M. Cluster, value network and business ecosystem: Knowledge and innovation approach. *Conference Paper*, September (2004), 9–10.

[144] PELTONIEMI, M. Preliminary theoretical framework for the study of business ecosystems. *E:CO 8*, 1 (2006), 10–19.

[145] PELTONIEMI, M., AND VUORI, E. Business ecosystem as the new approach to complex adaptive business environments. *Power* (2004), 1–15.

[146] PEPPARD, J., AND RYLANDER, A. From value chain to value network: Insights for mobile operators. *European Management Journal 24*, 2–3 (2006), 128–141.

[147] PINSONNEAULT, A., AND KRAEMER, K. L. Survey research methodology in management information systems: An assessment. Working paper, available from `http://escholarship.org/uc/item/6cs4s5f0`, 1993.

[148] PORTER, M. E. Strategy and the Internet. *Harvard Business Review 79*, 3 (2001).

[149] PRITCHETT, D. BASE: An ACID Alternative. *Queue 6*, 3 (2008), 48–55.

[150] PURDY, M., ROBINSON, M. C., AND WEI, K. Three new business models for "the open firm". *Strategy & Leadership 40* (2012), 36–41.

[151] QUINN, J. B., BARUCH, J. J., AND PAQUETTE, P. C. Technology in services. *Scientific American 257*, 6 (Dec 1987), 50–58.

[152] RAIVIO, Y. The broker - a solution for global mobile services. In *Proceedings of the 11th International Conference on Services, Enablers and Architectures Supporting Business Models for a New Open World* (2008), ICIN 2008, NeuStar.

[153] REISS, C., TUMANOV, A., GANGER, G. R., KATZ, R. H., AND KOZUCH, M. A. Towards understanding heterogeneous clouds at scale: Google trace analysis. Technical report, Intel Science & Technology Center for Cloud Computing, April 2012.

[154] RIEDL, C., BÖHMANN, T., LEIMEISTER, J. M., AND KRCMAR, H. A framework for analysing service ecosystem capabilities to innovate. In *Proceedings of the 17th European Conference on Information Systems* (2009), ECIS '09, pp. 1–12.

[155] RITTENHOUSE, G. Green wireless networks. Keynote at WCNC 2012, Paris, France, April 2, 2012, available from `http://www.ieee-wcnc.org/2012/Keynote.html`, 2012.

[156] ROBINSON, H. CAP confusion: Problems with 'partition tolerance'. Blog, August 26, 2010, available from `http://www.cloudera.com/blog/2010/04/cap-confusion-problems-with-partition-tolerance/`, 2010.

[157] ROCHET, J.-C., AND TIROLE, J. Platform competition in two-sided markets. *Journal of the European Economic Association 1*, 4 (2003), 990–1029.

[158] ROTEM-GAL-OZ, A. Fallacies of distributed computing explained. White paper, May 2006.

[159] ROWE, G., AND WRIGHT, G. The delphi technique as a forecasting tool: issues and analysis. *International Journal of Forecasting 15*, 4 (1999), 353–375.

[160] ROY, N., DUBEY, A., AND GOKHALE, A. Efficient Autoscaling in the Cloud using Predictive Models for Workload Forecasting. In *Proceedings of the 4th IEEE International Conference on Cloud Computing* (Washington, DC, USA, 2011), Cloud 2011, IEEE Computer Science, pp. 500–507.

[161] SALANT, P., AND DILLMAN, D. A. *How to conduct your own survey*. Business Reference. Wiley, 1994.

[162] SALMINEN, A., KALLIO, J., AND MIKKONEN, T. Towards mobile multimedia mashup ecosystem. In *Proceedings of the International Conference on Communications Workshops* (2011), ICC, pp. 1–5.

[163] SÁNCHEZ, L. A. G., AND RODRÍGUEZ, J. S. Overview of the new user centred mobile applications. In *Proceedings of the Second International Conference on Next Generation Mobile Applications, Services, and Technologies* (Washington, DC, USA, 2008), NGMAST '08, IEEE Computer Society, pp. 167–173.

[164] SANKARALINGAM, K., AND ARPACI-DUSSEAU, R. H. Get the parallelism out of my cloud. In *Proceedings of the 2nd USENIX conference on Hot topics in Parallelism* (Berkeley, CA, USA, 2010), HotPar'10, USENIX Association, pp. 1–6.

[165] SCHILLING, M. A. *Strategic Management of Technological Innovation*. McGraw-Hill/Irwin, 2009.

[166] SCHUMPETER, J. A. *The Theory of Economic Development: An Inquiry Into Profits, Capital, Credit, Interest, and the Business Cycle*. Harvard Economic Studies. Transaction Books, 1934.

[167] SHAPIRO, C., AND VARIAN, H. R. *Information Rules: A Strategic Guide to the Network Economy*. Harvard Business School Press, 1999.

[168] SHIN, K. G., JU, X., CHEN, Z., AND HU, X. Privacy protection for users of location-based services. *IEEE Wireless Communications 19*, 1 (February 2012), 30–39.

[169] SHUEN, A. *Web 2.0: A Strategy Guide - Business thinking and strategies behind successful Web 2.0 implementations*. O'Reilly Media, Inc., 2008.

[170] SHVACHKO, K., KUANG, H., RADIA, S., AND CHANSLER, R. The Hadoop distributed file system. In *Proceedings of the 26th Symposium on Mass Storage Systems and Technologies* (Washington, DC, USA, 2010), IEEE Computer Society, pp. 1–10.

[171] SKULMOSKI, G. J., HARTMAN, F. T., AND KRAHN, J. The Delphi method for graduate research. *Journal of Information Technology Education 6*, 1 (2007), 1–21.

[172] SMITH, M. Y., AND STACEY, R. Governance and cooperative networks: An adaptive systems perspective. *Technological Forecasting and Social Change 54*, 1 (1997), 79–94.

[173] SMURA, T., KIISKI, A., AND HÄMMÄINEN, H. Virtual operators in the mobile industry: a techno-economic analysis. *Netnomics 8*, 1-2 (Oct. 2007), 25–48.

[174] STAKE, R. E. *The Art of Case Study Research*. Sage Publications, 1995.

[175] STECCA, M., AND MARESCA, M. A cloud-based platform for enterprise mashup ecosystems. In *Proceedings of the Conference on Commerce and Enterprise Computing* (washington, DC, USA, 2011), IEEE Computer Society, pp. 244–251.

[176] STL PARTNERS. Your text is on Fire: OTT's to burn 40% SMS revenue by 2015. Newsletter, November 24, 2011, available from http://www.telco2research.com/articles/AN_text-sms-on-fire-OTT-burn-revenue_summary, November 2011.

[177] STONEBRAKER, M. In search of database consistency. *Communication ACM 53*, 10 (2010), 8–9.

[178] STRANDELL, T. Open source database systems: Systems study, performance and scalability. Master's thesis, University of Helsinki, Faculty of Science, Department of Computer Science and Nokia Research Center, Finland, August 2003.

[179] STREBEL, J., AND STAGE, A. An economic decision model for business software application deployment on hybrid cloud environments. In *Proceedings of the Multikonferenz Wirtschaftsinformatik* (2010), M. Schumann, L. M. Kolbe, M. H. Breitner, and A. Frerichs, Eds., Universitätsverlag Göttingen, pp. 195–206.

[180] SUBBIAH, B., AND RAIVIO, Y. Transport architecture for UMTS/IMT-2000 cellular networks. In *Proceedings of the International Performance, Computing, and Communications Conference* (Washington, DC, USA, 2000), IPCCC '00, IEEE Computer Society, pp. 208–214.

[181] SUSTER, M. It's Morning in Venture Capital. Blog, May 23, 2012, available from http://www.bothsidesofthetable.com/2012/05/23/its-morning-in-venture-capital/, 2012.

[182] THOMKE, S. H. *Experimentation Matters: unlocking the potential of new technologies for innovation*. Harvard Business School Press, 2003.

[183] TUUNAINEN, V. K., AND TUUNANEN, T. IISIn - A Model for Analyzing ICT Intensive Service Innovations in n-Sided Markets. In *Proceedings of the 44th Hawaii International Conference on System Sciences* (Washington, DC, USA, 2011), HICSS '11, IEEE Computer Society, pp. 1–10.

[184] TYRVÄINEN, P., AND MAZHELIS, O. *Vertical Software Industry Evolution: Analysis of Telecom Operator Software*. Contributions to Management Science. Physica-Verlag HD, 2009.

[185] VAJDA, A. Cloud and Telecoms. Presentation at Aalto University, Finland, 13.03.2012, available from https://noppa.aalto.fi/noppa/kurssi/t-110.6000/luennot, 2012.

[186] VAN DEN BOSSCHE, R., VANMECHELEN, K., AND BROECKHOVE, J. Cost-optimal scheduling in hybrid IaaS clouds for deadline constrained workloads. In *Proceedings of the 3rd International Conference on Cloud Computing* (Washington, DC, USA, 2010), Cloud 2010, IEEE, pp. 228–235.

[187] VENUGOPAL, S., LI, H., AND RAY, P. Auto-scaling emergency call centres using cloud resources to handle disasters. In *Proceedings of the Nineteenth International Workshop on Quality of Service* (Washington, DC, USA, 2011), IWQoS '11, IEEE Computer Society, pp. 34:1–34:9.

[188] VERKASALO, H. Open mobile platforms: Modeling the long-tail of application usage. In *Proceedings of the Fourth International Conference on Internet and Web Applications and Services* (Washington, DC, USA, 2009), ICIW '09, IEEE Computer Science, pp. 112–118.

[189] VERYZER, R. W. Discontinuous innovation and the new product development process. *Journal of Product Innovation Management 15*, 4 (1998), 304–321.

[190] VILA TENA, A., AND RAIVIO, Y. Privacy challenges of open APIs: Case location based services. In *Proceedings of the Ninth Annual International Conference on Privacy, Security and Trust* (Washington, DC, USA, 2011), PST 2011, IEEE Computer Society, pp. 213–220.

[191] VUKADINOVIĆ, V., AND KARLSSON, G. Video streaming in 3.5G: On throughput-delay performance of proportional fair scheduling. In *Proceedings of the 14th International Symposium on Modeling, Analysis, and Simulation* (Washington, DC, USA, 2006), MASCOTS '06, IEEE Computer Society, pp. 393–400.

[192] VUKOLIĆ, M. The byzantine empire in the intercloud. *SIGACT News 41*, 3 (Sep 2010), 105–111.

[193] WALKER, E. The real cost of a CPU hour. *Computer 42*, 4 (2009), 35–41.

[194] WANG, Y., WEN, X., SUN, Y., ZHAO, Z., AND YANG, T. The content delivery network system based on cloud storage. In *Proceedings of the 2011 International Conference on Network Computing and Information Security - Volume 01* (Washington, DC, USA, 2011), NCIS '11, IEEE Computer Society, pp. 98–102.

[195] WEINMAN, J. Mathematical proof of the inevitability of cloud computing, 2011. Online report, 08.01.2011, available from `http://www.joeweinman.com/Resources/JoeWeinmanInevitabilityOfCloud.pdf`.

[196] WIRELESS INTELLIGENCE. European mobile ARPU falls 20% - Mobile data revenue not yet compensating for declining voice prices in EU mobile markets. Available from `http://www.wirelessintelligence.com/analysis/2011/03/european-mobile-arpu-falls-20/`, March 2011.

[197] XIA, R., ROST, M., AND HOLMQUIST, L. E. Business models in the mobile ecosystem. In *Proceedings of the 2010 Ninth International Conference on Mobile Business / 2010 Ninth Global Mobility Roundtable* (Washington, DC, USA, 2010), ICMB-GMR '10, IEEE Computer Society, pp. 1–8.

[198] YAZIR, Y. O., MATTHEWS, C., FARAHBOD, R., NEVILLE, S., GUITOUNI, A., GANTI, S., AND COADY, Y. Dynamic resource allocation in computing clouds using distributed multiple criteria decision analysis. In *Proceedings of the 3rd International Conference on Cloud Computing* (Washington, DC, USA, 2010), Cloud 2010, IEEE Computer Society, pp. 91–98.

[199] YIN, R. K. *Case study research: design and methods*. Applied social research methods series. Sage Publications, 2003.

[200] ZANG, N., ROSSON, M. B., AND NASSER, V. Mashups: who? what? why? In *CHI '08 extended abstracts on Human factors in computing systems* (New York, NY, USA, 2008), CHI EA '08, ACM, pp. 3171–3176.

[201] ZEITHAML, V. A., BITNER, M. J., AND GREMLER, D. D. *Services Marketing*. McGraw-Hill Irwin, 2008.

[202] ZERFOS, P., MENG, X., WONG, S. H., SAMANTA, V., AND LU, S. A study of the short message service of a nationwide cellular network. In *Proceedings of the 6th ACM SIGCOMM conference on Internet measurement* (New York, NY, USA, 2006), IMC '06, ACM, pp. 263–268.

[203] ZHANG, H., JIANG, G., YOSHIHIRA, K., CHEN, H., AND SAXENA, A. Intelligent workload factoring for a hybrid cloud computing model. In *Proceedings of the Congress on Services - I* (Washington, DC, USA, 2009), SERVICES '09, IEEE Computer Society, pp. 701–708.

[204] ZHANG, N., LEVÄ, T., AND HÄMMÄINEN, H. Two-Sidedness of Internet Content Delivery. In *Proceedings of the 10th Conference on Telecommunications Internet and Media Techno-Economics* (Berlin, Offenbach, 2011), CTTE2011, VDE Verlag GMBH, pp. 1–6.

[205] ZHENG, L., CHONG, S., MYERS, A. C., AND ZDANCEWIC, S. Using replication and partitioning to build secure distributed systems. In *Proceedings of the Symposium on Security and Privacy* (Washington, DC, USA, 2003), SP '03, IEEE Computer Society, pp. 236–250.

Mobile operators are financially in good shape. However, there are weak signals, which can be regarded as an early warning about the future prospects of the operators. This dissertation proposes two complementing propositions, which can be utilized for the operator renewal process. Open innovation has been one of the main drivers in the success of the internet companies, while the introduction of cloud computing has lowered the risk of entrepreneurship to a new level that has attracted in a vast number of novel startups. In summary, mobile operators should, without prejudice, evaluate the novel opportunities and experiment with new concepts, technologies and business models. The telecommunication sector has been used to working on well-defined standards, but in the networked and digital economy, the old approach is too slow. Instead, the mobile industry should adapt the best practices of the Internet and transform their business to meet the future challenges.

Aalto-DD 135/2012

9 789526 048406

ISBN 978-952-60-4840-6
ISBN 978-952-60-4839-0 (pdf)
ISSN-L 1799-4934
ISSN 1799-4934
ISSN 1799-4942 (pdf)

Aalto University
School of Science
Department of Computer Science and Engineering
www.aalto.fi

BUSINESS +
ECONOMY

ART +
DESIGN +
ARCHITECTURE

SCIENCE +
TECHNOLOGY

CROSSOVER

DOCTORAL
DISSERTATIONS