HELSINKI UNIVERSITY OF TECHNOLOGY
Department of Electrical and Communications Engineering
Networking Laboratory

Nicklas Beijar

Distribution of Numbering Information in Interconnected Circuit and Packet Switched Networks

Thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Engineering

Espoo, Finland, January 31, 2002

Supervisor Professor Raimo Kantola

Instructor Jose Miguel Costa Requena, M.Sc.

HELSINKI UNIVERSITY OF TECHNOLOGY ABSTRACT OF THE MASTER'S THESIS

Author:	Nicklas Beijar	
Name of thesis:	Distribution of Numbering Information in Interconnected Circuit and Packet Switched Networks	
Date:	31.1.2002	**Number of pages:** 105
Faculty:	Department of Electrical and Communications Engineering	
Professorship:	S-38 Networking Technology	
Supervisor:	Prof. Raimo Kantola	
Instructor:	Jose Miguel Costa Requena, M.Sc. (Tech)	

Because of the introduction of Internet Protocol (IP) telephony and the increasing complexity of the telephony networks, the burden of managing routing information is growing. Number portability, which allows users to change operators, locations and services without changing their numbers, is now being required in many countries. An automatic approach of generating and distributing routing information is motivated.

This master's thesis examines the issue of how a routing protocol can be used to distribute numbering information in an interconnected circuit and packet switched network. An architecture based on the Telephony Routing over IP (TRIP) protocol for distributing routing information is developed.

The first part introduces some of the theory and literature related to the subject. The principles of routing telephone calls on the circuit switched network and IP networks are presented. The TRIP protocol and ENUM (tElephone NUmbering Mapping) directory are major issues, since they form the base for the architecture. Number portability and its implementations are described.

In the second part, the solution is developed. The architecture is defined and some scenarios are presented. A new protocol named CTRIP (Circuit Telephony Routing Information Protocol) is developed as a counterpart to TRIP for the circuit switched network. The attributes and messages of CTRIP are defined. The TRIP protocol is extended with some new optional attributes. Further, the process of converting routing information between TRIP and CTRIP is defined. A new network element named numbering gateway is introduced. To some extent, the ENUM directory is integrated into the solution. Finally, some applications and scenarios based on the solution are discussed.

Keywords: routing, routing protocol, number portability, TRIP, CTRIP, ENUM, NP, VoIP

TEKNILLINEN KORKEAKOULU **DIPLOMITYÖN TIIVISTELMÄ**

Tekijä:	Nicklas Beijar
Työn nimi:	Numerointitietojen levitys yhdistetyissä piiri- ja pakettikytkentäisissä verkoissa
Päivämäärä:	31.1.2002 **Sivumäärä:** 105
Osasto:	Sähkö- ja tietoliikennetekniikan osasto
Professuuri:	S-38 Tietoverkkotekniikka
Työn valvoja:	Prof. Raimo Kantola
Työn ohjaaja:	DI Jose Miguel Costa Requena

IP (Internet Protocol) -puheluiden yleistymisen ja puhelinverkkojen kasvavan monimutkaisuuden johdosta reititystietojen hallinta vaikeutuu. Numeron siirrettävyys, jonka avulla käyttäjät voivat vaihtaa operaattoria, paikkaa ja palveluja vaihtamatta puhelinnumeroaan, vaaditaan nyt pakollisena useissa maissa. Tarvitaan menetelmä, jolla luodaan ja levitetään reititystietoja automaattisesti.

Tämä diplomityö tutkii, miten reititysprotokollaa voidaan soveltaa reititystietojen levitykseen yhdistetyssä piiri- ja pakettikytkentäisessä verkossa. Työssä kehitetään TRIP (Telephony Routing over IP) -protokollaan perustuva arkkitehtuuri numerointitietojen levitykseen.

Työn ensimmäinen osa käsittelee aiheeseen liittyvää teoriaa ja kirjallisuutta. Tässä osassa esitetään pääperiaatteet, miten puhelut reititetään piirikytkentäisissä verkoissa ja IP-verkoissa. Erityisesti kuvataan TRIP-protokollaa ja ENUM (tElephone NUmbering Mapping) -hakemistoa, koska ne muodostavat arkkitehtuurin perustan. Tämän lisäksi käsitellään numeron siirrettävyyttä ja sen toteutuksia.

Toisessa osassa ratkaisu kehitetään. Arkkitehtuuri määritetään ja muutamia skenaarioita esitetään. Uusi CTRIP (Circuit Telephone Routing Information Protocol) -niminen protokolla kehitetään TRIP:in vastineeksi piirikytkentäiseen verkkoon. CTRIP:in attribuutit ja sanomat määritellään. TRIP-protokollaa laajennetaan lisäämällä uusia valinnaisia attribuutteja. Lisäksi määritellään reititystietojen muunnosprosessi TRIP:in ja CTRIP:in välillä. Uusi verkkoelementti, nimeltään numerointiyhdyskäytävä esitellään. ENUM-hakemisto liitetään ratkaisuun tietyissä määrin. Lopuksi käsitellään muutamia ratkaisuun perustuvia sovellutuksia ja skenaarioita.

Avainsanat: reititys, reititysprotokolla, numeron siirrettävyys, TRIP, CTRIP, ENUM, NP, VoIP

Acknowledgements

This Master's Thesis has been written in the Networking Laboratory of Helsinki University of Technology within the IMELIO project. IMELIO is funded by Tekes, Nokia Networks, Nokia Research Center and Elisa Communications.

I would like to thank Professor Raimo Kantola for his guidance and support during writing of the thesis.

I want to express my thanks to my instructor Jose Costa Requena for valuable comments, ideas and support.

I would also like to thank the people within the team who have contributed to my thesis with knowledge, ideas and implementation work: Ignacio González Olías, Julio Ramirez, Jari Huttunen, Juho Haapala and Antti Paju. I want to thank my colleagues at the lab for an innovative atmosphere to work in.

Last, but most importantly, I would like to thank my friends and close relatives for their support during the studies. Special thanks go to Minna Niilonen for her encouraging and loving support.

January 31st, 2002 in Espoo, Finland

Nicklas Beijar

Abbreviations

ACQ	All Call Query
ANSI	American National Standards Institute
AS	Autonomous System
BGP-4	Border Gateway Protocol version 4
CCF	Call Control Function
CLIP	Calling Line Identification Presentation
codec	Coder Decoder
CTAD	Circuit Telephony Administrative Domain
CTRIB	Circuit Telephony Routing Information Base
CTRIP	Circuit Telephony Routing Information Protocol
DNS	Domain Name Service
DPC	Destination Point Code
DSP	Digital Signal Processor
ENUM	tElephone NUmber Mapping (IETF working group)
ETSI	European Telecommunications Standards Institute
FICORA	Finnish Communications Regulatory Authority
HLR	Home Location Register
IANA	Internet Assigned Numbers Authority
IETF	Internet Engineering Task Force
IN	Intelligent Network
INAP	Intelligent Network Application Part
IP	Internet Protocol
IPTEL	Internet Protocol Telephony (IETF working group)
ISDN	Integrated Services Digital Network
ISP	Internet Service Provider
ISUP	ISDN User Part
ITAD	Internet Telephony Administrative Domain
ITSP	Internet Telephony Service Provider
ITU-T	International Telecommunications Union - Telestandardization Sector
LAN	Local Area Network
LDAP	Lightweight Directory Access Protocol
LS	Location Server
MAP	Mobile Application Protocol
NAPTR	Naming Authority Pointer
NP	Number Portability
NPDB	Number Portability Database
OR	Onward Routing
OSPF	Open Shortest Path First
PBX	Private Branch Exchange
POTS	Plain Old Telephone System
PLMN	Public Land Mobile Network
PSTN	Public Switched Telephone Network
QoR	Query on Release
QoS	Quality of Service

RAS	Registration, Admission, Status
RIP	Routing Information Protocol
RFC	Request For Comments
RTCP	Real Time Control Protocol
RTP	Real Time Protocol
SCF	Service Control Function
SCF	Service Control Point
SCN	Switched Circuit Network
SCSP	Server Cache Synchronization Protocol
SDF	Service Data Function
SIP	Session Initiation Protocol
SMS	Service Management System
SS	Signaling Server
SS7	Common Channel Signaling System No. 7
SSF	Service Switching Function
SSN	Subsystem Number
SSP	Service Switching Point
TAD	Telephony Administrative Domain
TCP	Transmission Control Protocol
TIPHON	Telecommunications and Internet Protocol Harmonization Over Networks
TRIB	Telephony Routing Information Base
TRIP	Telephony Routing over IP
TSAP	Transport Service Access Point
TUP	Telephony User Part
UAC	User Agent Client
UAS	User Agent Server
URI	Universal Resource Identifier
URL	Universal Resource Locator

Table of contents

Index of figures

Index of tables

1. Introduction

1.1 Background

The popularity of Internet Protocol (IP) telephony technology has been increasing. As the technology has matured, it has become a considerable alternative for use in the trunk and access networks as well as in local area networks. With the many advantages of IP telephony, such as a single network for data and voice, application integration, efficient network utilization and new services, it seems that the development is going towards a replacement of traditional circuit switched technologies with IP telephony. Before this completely IP-based network can be realized, the IP telephony network has to coexist with the switched circuit network (SCN) for a long time. During this transition time, the IP telephony network has to be a full peer to the SCN.

The coexistence of two network types puts additional demands on routing. When a call crosses the technology border, it has to pass through a gateway, which converts between packet-switched and circuit-switched transfer modes and performs necessary coding conversions. Conversions cause delay and jitter, which degrade the voice quality. It is desirable to reduce the number of conversions on the media path. The problem of selecting a suitable gateway is a non-trivial process depending on several factors. Although automatic gateway selection is available for calls from the IP network to the SCN, the gateways to use for calls in the opposite direction must be manually configured.

Number portability allows a number to move geographically within a network, between different operator's networks and, as IP telephony is introduced, between network technologies. The governments in many countries are making number portability mandatory in order to enable and encourage competition. The requirement applies to IP-based telephony networks as well. Furthermore, number portability is an important enabler for IP telephony by allowing smooth transfer of subscribers to the IP network.

Number portability requires that the exchanges have access to additional routing information, which is shared between operators. Many of the current solutions cause inefficient routing due to routing through the previous network. The situation is further complicated when numbers are allowed to move between network technologies. To maintain efficient routing and good voice quality, a different gateway may need to be selected. An automated approach for the gateway selection is motivated.

IP telephony allows current Internet service providers (ISP) to open telephony service in their networks. IP telephony technology lowers the threshold for opening new networks and services. Due to multiple network technologies, an increasing number of operators and number portability, the load of managing routing information increases. With more revenues coming from services rather than voice transport, service management becomes more important. Currently management is performed more or less manually. Therefore, an automatic approach is considered necessary.

1.2 The goals and objectives of the thesis

In this thesis, we analyze the problems and inefficiencies seen in routing in an interconnected SCN and IP-network. The aim is to develop a solution for application layer routing in an interconnected SCN and IP network. Telephone numbers are used as the common addressing method.

The main goal is to examine:

- How can a routing protocol be utilized for distributing numbering information in an interconnected SCN and IP network?

The secondary goals are to define:

- The protocols used for distributing numbering information in the SCN and IP networks.
- The conversion process between numbering information used in the SCN and the IP networks.

We want the solution to include the definition of the architecture, the protocol specification and the required network elements. The solution is developed and analyzed using some scenarios where number portability plays an important role.

The study is based on the following background conditions:

1. The existence of two interconnected network technologies: SCN and IP.
2. Number portability is mandatory in the SCN, in IP networks and between them.
3. The TRIP protocol is used for gateway location in IP networks.
4. Information about terminals on the IP network is stored in DNS, or alternatively in TRIP.

The solution automates the creation of routes to destinations on both the IP network and the SCN. The motivation is to reduce the configuration load of the operator, and to provide close to optimal routing in a hybrid SCN-IP network. The requirement is that the network operators should be able to influence on routing by defining policies. The selection of gateways should be automatic for calls in both directions between the two network technologies. The number of media conversions in gateways should be minimized, since they degrade voice quality. By using a routing protocol on the SCN, it is also possible to automatically generate the information distributed by TRIP.

In the development of the solution, the following objectives are considered:

1. Existing protocols and solutions are preferred instead of new ones.
2. The solution should be compatible with existing protocols and solutions

3. Modification of existing protocols should be avoided.
4. The solution should be extendable for inclusion of future technologies and protocols.

1.3 Scope of the thesis

Conforming to our objectives, we limit the scope to solutions based on existing protocols. There are several existing routing protocols used in different types of networks. Most of them are network layer routing protocols. We have chosen the TRIP protocol [RFC 2871] as the base, since it is an application layer protocol and it performs similar functions on the IP network.

Although the solution could be extended to include other types of networks, such as second and third generation mobile networks, we limit the scope to only SCN and IP networks. The solution must, however, be made general enough to allow inclusion of these later. Other network types are only discussed in brief.

As a central application of the solution, we will study number portability. Especially interesting is number portability between the SCN and the IP network. The scope is limited to slow changes. Fast movement, such as roaming, should be supported by other protocols instead.

The solution is developed with the plans of the ENUM working group in mind. However, interworking with ENUM will only be specified on a conceptual level. ENUM is still in a state of change, and the exact definition of its integration into the solution is left for further research.

1.4 The structure of the thesis

This thesis is divided into two parts. The first four chapters form the first part and they cover some of the theory that is related to the topic. They are based on literature research. The chapters form a ground for understanding the rest of the thesis.

Chapter 2 describes the central concepts related to routing and numbering on the switched circuit network.

Chapter 3 presents number portability and its relation to routing. The emphasis is on the number portability solution used in Finland. However, most countries have adopted similar solutions.

Chapter 4 describes routing of Internet Protocol Telephony calls. It focuses on the problems of locating terminals and gateways, and the solutions developed to solve these problems.

In the second part, consisting of chapters five to ten, the solution is developed. A stepwise approach is used to clarify the ideas behind the choices.

Chapter 5 introduces the idea of using a distributed routing protocol in the switched circuit network and discusses various issues and scenarios related to the approach. It lays the foundation for the specification work in the following chapters.

Chapter 6 describes the proposed architecture.

Chapter 7 defines the operation of the protocols and the format of the attributes of the protocols. It mainly describes the proposed protocol named CTRIP, but it also suggests some additions to the existing TRIP protocol.

Chapter 8 specifies the operation of the numbering gateway.

Chapter 9 discusses scenarios and applications of the designed architecture. The main topics are how to implement carrier selection and how larger networks are managed.

Chapter 10 concludes the thesis, describes the advantages and limitations of the solution, and suggests ideas for further work.

2. Switched circuit networks

To get an overview how switched circuit networks work, we will present the central concepts related to routing and numbering. We present the E.164 numbering scheme and describe how numbering is related to routing. We explain the difference between directory numbers and routing numbers. We also give a very brief introduction to Intelligent Network (IN) technology.

2.1 E.164 numbering

On the switched circuit network (SCN), telephone numbers are used to identify subscribers. The public numbering system for the PSTN and ISDN is defined in ITU-T recommendation E.164 [ITU-T E.164]. In practice it also includes mobile networks. The number has a hierarchical format, beginning with the country code and ending in the subscriber number. The structure of an E.164 number is exemplified in Figure 1. [ETSI TR 101 326, Understanding 1997].

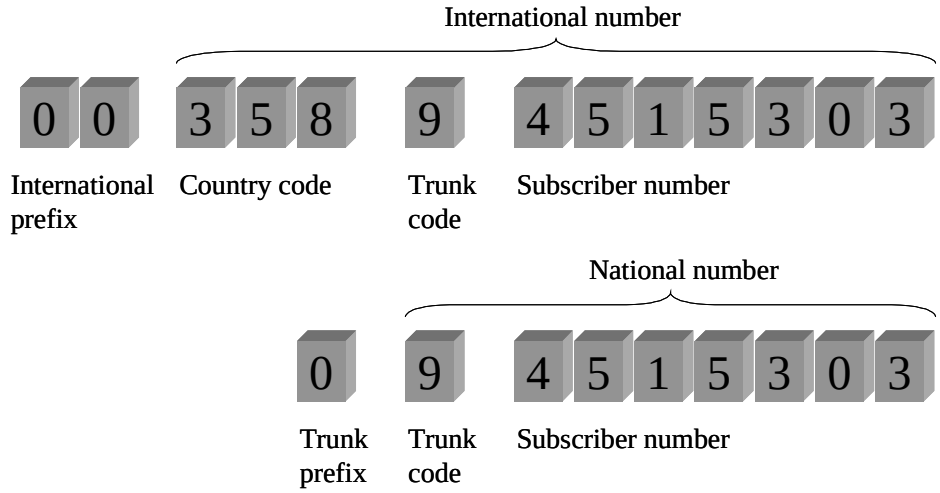

Figure 1. International and national numbers

A telephone number consists of a country code, a trunk code and a subscriber number. In an international number, it is preceded by an international prefix. In a national number, the country code is omitted and the number may be preceded by a trunk prefix. The combination of a trunk prefix and a trunk code is called an area number. [Understanding 1997]

Different bearer networks have different numbering plans. Bearer networks include the PSTN, PLMN and PSPDN. ISDN is exceptionally integrated with the numbering plan of PSTN. The opening of the telecommunications market has allowed competition in the telecommunications

networks. Especially in the PLMN, the existence of several operators is visible also in the numbering plans: the prefix not only indicates the bearer network, but also the operator [Understanding 1997]. As a result of deregulation of trunk and international traffic, the user can select the carrier using a carrier selection code [THK 2000].

Numbering within countries is managed by the national administration or numbering authority. In Finland, the Finnish Communications Regulatory Authority (FICORA) defines the numbering plan. The numbering space in the regional areas is divided to the operators in blocks of three most significant numbers. These are distributed to the operators when needed. The International Telecommunications Union (ITU) allocates codes to countries and geographical regions, global services and special networks sharing a global code. The lengths of the trunk codes and subscriber numbers vary from country to country. The subscriber number may have fixed or variable length. Fixed length numbers simplify the call setup procedure but require often longer number than necessary. Starting from year 1997, the maximum length of an international telephone number is 15 digits. [Understanding 1997, E.164, THK 1996]

In addition to switches and terminals, also different services, such as 800 services, are assigned E.164 numbers. When a user dials a service number, the switch recognizes the number as a service number and accesses a database using Intelligent Network (IN) technology. Further processing of the call follows the instructions obtained from the IN database. [Sugino 1999]

2.2 Routing

The format of the telephone number is closely related to the routing process. A geographic number (GN) indicates the geographical location of the subscriber: the first digits tell the country area, the second group of digits indicates the area, the following group of digits indicates the switch, and the last numbers identify the subscriber line within the switch. When the subscriber is connected to a PBX, the last digits indicate the extension and the preceding digits identify the PBX. [ETSI TR 101 119, Understanding 1997]

On the switched circuit networks, the telephone number is used to route the call to the terminating switch. It is convenient to attach numbering blocks (e.g. 358 9 345 xxxx) or ranges (e.g. 358 9 345 – 358 9 347) to switches. The number analysis is performed step by step on the route to the destination. The switches usually only examine the first digits that are necessary to know the route to the next switch. The number analysis tree is followed according to the digits until a leaf is reached. The routing instructions for the match are extracted and executed. The next switch then analyzes the same or the following digits, and the process repeats until the terminating switch is reached. The terminating switch locates the hardware address of the line card. [ETSI TR 101 326]

The number analysis process is thus distributed among switches, and the total required switch processor power is minimized. Additionally, maintenance of up-to-date routing tables is easier. A disadvantage of using numbering blocks is the uneconomic use of the available numbering space. [ETSI TR 101 326]

2.3 Routing addresses

The number dialed by the user to reach a destination is called a *directory number*, since it is the number that can be found in telephone directories. Traditionally the telephone number performs several functions. It indicates the party that the caller wants to reach, and it determines how the telephone network sets up the route of the call. The caller's number indicates the calling number presented to the callee, but it also indicates the party that will be charged for the service.

Services such as number portability require the functions to be separated. Due to number portability, the subscriber may move away from the location, operator or network that the telephone number indicates. Thus, the subscriber moves away from the area that the hierarchical directory number indicates, and into an area where it may not belong hierarchically. The address space becomes flat, since all numbers are at the same hierarchical level. To utilize the hierarchical structure that is used for routing, a different number than the directory number may be required for routing. That number is called the *routing number*. A translation function is required to map the directory number into a routing number. [ETSI TR 101 326]

The European Telecommunications Standards Institute (ETSI) separates between the addresses used for identifying destinations and the addresses used for routing by using the terms *name* and *address*. ETSI formally defines an address as a string or combination of digits and symbols which identifies the specific termination points of a connection/session and is used for routing. The address is a specification of the location in terms of network structure. A name is defined as a combination of letters, digits or symbols that is used to identify end-users. [ETSI TR 101 326]

The main difference between these is that a name is an identifier for the end user, while an address is a locator. An address should typically have some structure that allows aggregation for routing purposes. According to this definition, we can classify the directory number as a type of name and the routing number as a type of address.

For various historical reasons, E.164 numbers are a mixture of names and addresses. The trend is however to make them more names than addresses by reducing the degree of address information. Especially the requirements of operator and location portability stand behind this development. [ETSI TR 101 326]

With both a directory number and a routing number, there are various methods for giving routing information in signaling messages. Routing information can be given with a concatenated address, with separate addresses, with partly separated addresses, as a routing number only, or as a directory number only.

A concatenated address consists of a routing number followed by the directory number. The resulting address is transported in the Called Party Number field. The separate addresses approach uses two fields in the signaling message: one for the routing number and another for the directory number. The partly separated address approach also uses two fields. One field is used for the part of the routing number that is used for routing to the right network. The part of the routing number

that consists of detailed routing information is concatenated with the directory number in a second field. Finally, in many cases, it is enough to use only a routing number. However, this is only possible across network boundaries. [ETSI TR 101 122]

Existing ISUP versions can only transport one number, which is the reason for combining the routing number with the directory number. Newer ISUP versions can transport both numbers. However, some networks put the routing number in the directory number field, and the directory number in the new field, and some put the directory number in the old field and the routing number in the new field for compatibility reasons. [ETSI TR 101 122]

The routing number is generated through a query to an IN database or through on-switch processing. In addition to number portability, IN-like queries are used to generate the Mobile Station Roaming Number (MSRN), which also is a type of routing number. [ETSI TR 101 326]

2.4 Intelligent Networks (IN)

Supplementary services benefit both users and operators. These are for instance call waiting, call forwarding, automatic callback, freephone and third-party billing calls. Also number portability, which is a major issue in this thesis, is a supplementary service. The services are either distributed, i.e. deployed in the local exchanges, or centralized. Central implementation has the advantage that not every exchange has to support the service. When a new service is introduced, only one element must be modified. Introduction of services is thus quick and simple. Centralized supplementary services are implemented using Intelligent Network (IN) technology. [Understanding 1997]

The IN architecture consists of service switching points (SSP), service control points (SCP) and the service management system (SMS). The SSP is an exchange equipped with service switching functions (SSF). The SCP is a centrally located node containing logic and data for IN services. IN services are invoked when a triggering point in the SSP is detected. The SCP executes service requests received from the SSP and returns information how the call should be handled in the SSP. The functions of the SCP are named service control functions (SCF). The SCP and SSP can be combined to form a service switching and control point (SSCP). [Understanding 1997]

3. Number Portability

Since number portability is a major application of the architecture that we design, and also an application that already utilizes number translation databases, we give a presentation of number portability in this chapter. The chapter begins with a description of number portability and a classification into different types of number portability. We describe the relationship between routing and number portability, and present some different implementation schemes. Especially we concentrate on the IN-based number portability model that is proposed for switched circuit networks in Finland. Understanding number portability and routing with separate routing numbers is essential for designing the architecture based on routing protocols.

3.1 Introduction to number portability

Traditionally, when a subscriber moved from one location to another, his telephone number changed. Now, when operators are competing for subscribers in the local loop and in the mobile networks, an additional reason for changing the telephone number has appeared. The subscriber must change his telephone number when he chooses to be served by another operator. This poses considerable inconveniences and costs that reduce the interest in moving. Number portability solves this problem. With number portability, the subscriber can keep his old telephone number when changing to another geographical region or to another operator.

Number portability can be classified into three categories. The ability to change the service provider while keeping the number is called *service provider portability* or *operator portability*. The ability to change the location while keeping the number is called *location portability*. *Service portability* allows the subscriber to change the subscribed services while keeping the same number. One example of this is a subscriber changing from POTS to ISDN service. [Foster 2000]

Service provider portability has become mandatory in many countries. The aim is to liberalize the competition between operators, especially for local service. Competition between operators was opened in most European countries in the beginning of 1998. With the introduction of competition on the telecommunications market, many regulators have made number portability mandatory to lower the threshold for changing operators. Number portability allows subscribers to choose the operator based on tariffs and available services. It greatly simplifies the change to the operator offering the best price and service level, thus encouraging competition. According to the European Union consortium, number portability for subscribers that change operators without moving physically must be operational in 2003. [THK 1996, ETSI TR 101 119]

The increasing interest in IP telephony and the emerging IP based telephone networks, also create needs for number portability between different network technologies. Market entrants, whose telephone network is based on IP technology, may be concerned to give their customers numbers that looks as similar as possible to numbers used on existing SCN networks. Number portability is in general very essential to market entrants, since it will allow them to compete for customers from the existing operators. In return, they have to offer number portability to other operators from their own networks [Rosbotham 1999]. We consider the ability to move between SCN and IP networks as a type of service portability.

3.2 Terms

We now introduce some terms that will be used in the following discussion. The *directory number* is the number that is dialed by the user. It corresponds to the normal telephone number found in telephone directories. It is in the E.164 format, although it may also be given in a local format. A number in E.164 format uniquely and globally identifies the subscriber. The directory number does not include the area code itself or any possible operator prefix. A *routing number* is a number that is used within the network for routing purposes. For non-ported numbers the directory number and the routing number are generally the same. [Foster 2000]

We define a *moved number* or *ported number* as a telephone number that has been moved from one operator, location or network technology to another. These numbers have at least once moved. The *donor network* is the network that first assigned the number to a subscriber from a number range that was assigned to it administratively. The ported numbers thus belong to a number range that belongs to the donor network. The *old serving network* is the network that served the number before it was ported. The *new serving network* is the network that currently serves the subscriber. The old serving network is not necessarily the same as the donor network. [Foster 2000]

3.3 Number portability and routing

In the SCN, the exchanges only have a local view of the network. The exchange analyses the destination number of an incoming call setup and forwards the setup to the next exchange. Usually only the first few digits of the number are analyzed, since they indicate the neighbor exchange where the setup is forwarded. Only for terminating calls in local exchanges, the last digits are analyzed.

Introduction of number portability requires significant changes to numbering administration, signaling, call routing, billing and service management. Number portability changes the role of the dialed number from a hierarchical physical routing address to a virtual address. Since the number cannot be split into parts, which step by step tell where to route the call setup, the whole number must be analyzed at once. The number has no hierarchical structure that can be used for routing and the number space is more or less flat.

If only a few numbers are moved, the exchange can check from a table containing all moved numbers to find where to send the call setup. If it is not found in the table, the call is routed normally. If many numbers are moved, the size of the table grows. The lookup must be performed in every exchange on the way, which is an enormous burden.

The solution it to separate between directory numbers and routing numbers. The directory number, dialed by the subscriber, is used to retrieve a routing number. The routing number is used for routing internally in the network. It has a hierarchical structure that identifies the topological location.

The separation between directory numbers and routing numbers solves the problem using hierarchical numbering for routing numbers and flat numbering for directory numbers. A mapping from directory number to routing number must therefore be performed at least once per call. The directory number is translated to a hierarchical routing number at an early stage of call setup. The hierarchical structure of the routing number can be utilized so that the call can be routed by analyzing the number step-by-step for locating the following exchange. The subscriber has the ability to move without changing the number, since only the mapping has to be modified. The mapping is stored in a number portability database.

Depending on the implementation, the subscriber may have to change the number to a portable number the first time he moves. The subscriber obtains a special portable number, which belong to a numbering space reserved for number portability. This is simpler from the implementation point of view, since the number can directly be recognized as a portable number. The drawback of this solution is that the subscriber must change his number once. This drawback can be avoided with a more sophisticated implementation, which allows the original number to be moved. [THK 1996]

3.4 Number portability implementations

The implementations of numbering portability can be grouped into four main types, as presented in [Foster 2000].

In the *All Call Query (ACQ)* scheme the originating network queries a number portability database upon receiving a call. The database is centrally administered, but the network usually contains a copy of it. The database returns a routing number, which is used to route the call to the new serving network.

In the *Query on Release (QoR)* scheme, the call is first routed to the donor network. The donor network releases the call and indicates that the number has been ported out from the switch. Then the call is processed similarly to the ACQ scheme: a number portability database is queried to obtain a routing number, which is used to route the call to the new serving network.

In the *Call Dropback* scheme, the originating network routes the call to the donor network. The donor network detects that the number has been ported out. It queries an internal network-specific database and obtains a routing number. The donor network releases the call and returns the routing number, which the originating network uses to route the call to the new serving network.

The *Onward Routing (OR)* scheme is least efficient in using network resources. The originating network routes the call to the donor network. The donor network detects that the number has been ported out. It queries an internal network-specific database, which returns a routing number. The donor network uses the routing number to route the call to the new serving network.

In all schemes, except the ACQ scheme, the call is first routed to the donor network. The OR scheme is the least efficient, since it requires the setup of two physical call segments: from the originating network to the donor network, and from the donor network to the new serving network. In both the QoR and Call Dropback schemes the call segments to the donor network is released after the query. [Foster 2000]

3.5 IN-based number portability in Finland

The Finnish Communications Regulatory Authority (FICORA) compared a number of solutions for number portability. The work was divided into one working group examining solutions based on IN technology and one for solutions without IN. The work of the IN-based number portability group resulted in a report [THK 1996], presenting an IN-based solution for Finland. The solution is based on a model shown in Figure 2.

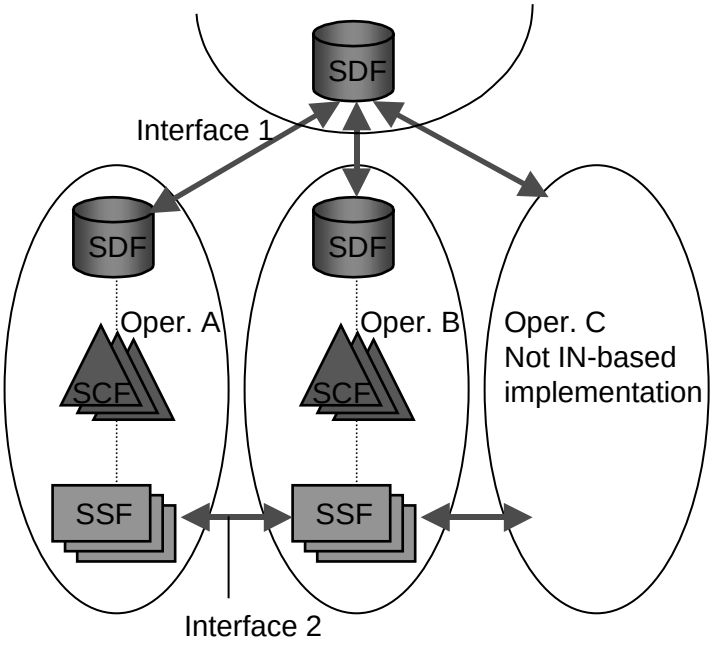

Figure 2. Model for IN-based number portability

3.5.1 Architecture and interfaces

A central database, called master database, contains the mappings between directory numbers and routing numbers for the subscribers within a number portability routing domain. In the first phase, only moved numbers are stored in the database, but at a later stage when about 10 – 20% of the subscribers are moved, the database will contain all numbers. The database is administered by an outside organization. The operators maintain copies of the master database in their SDF, and all database queries are made to the copies. At regular intervals, the contents of the master database are transferred to the operator's databases. The interface between the master database and the operator's database is named interface 1, but it has not been defined. No queries are made directly to the master database. Only information about changes such as subscribers moving between operators is transferred in interface 1. [THK 1996]

The master database contains the following fields:

- The directory number of the subscriber (key field)
- The operator serving the subscriber
- Information about the number portability routing domain where the subscriber is located

The last field is only necessary if the same database serves several number portability routing domains.

Two interfaces are defined in the model: interface 1 is the database interface between the master database and the SDF of the operator, and interface 2 is the TUP/ISUP interface between the operators. The internal implementation can be made in other ways than the proposed, for instance without using IN technology, but both the interfaces have to exist.

In interface 1, information is transferred in three different situations:

- The old operator informs about the subscriber moving to another operator. The message contains the directory number, the new operator and possibly the old operator.
- The master database informs all operators within the domain about the change. The message is similar to the above, but also contains the activation time of the change.
- All operators acknowledge the activation to the master database. The message contains the directory number, the new operator and activation time.

Call setup signaling between operators passes through interface 2, which is the normal signaling interface between operators. TUP or ISUP signaling is used in interface 2. The interface is used to establish calls between the operators, when the receiving subscriber uses a portable number. In addition to the normal information needed for call establishment, number portability needs at least two additional fields: the operator of the B-subscriber and a flag whether the database query has been performed. An alternative would be to add the information as new fields to TUP and ISUP signaling. However, this would require changes to the signaling standards, which would delay the deployment of number portability. The solution proposed in [THK 1996] sends the information as

an additional prefix before the number of the B-subscriber. It is used during the first stage of number portability.

The number of the B-subscriber has the following format in interface 2:

1D + operator-id + service-id + b-directory-number

The first two digits, “1D”, indicate that the number contains signaling information. The operator-id is a one or two-digit number uniquely identifying the operator. The service-id is a two-digit number, which is defined as 01 for the number portability service. The service-id allows for future extensions to the model, for implementation of services between operators.

Since the number signaled across interface 2 always is a directory number, the operators can choose the routing numbers used within their own network freely. The routing number can consist of normal digits between 0 and 9, but to extend the available number space the operator may also want to use the pentadecimal digits “*”, “#”, “A” and “B”. The routing number identifies the destination exchange of the subscriber and the equipment position of the subscriber’s line. The subscriber should not be able to dial the routing number directly.

The drawback of this choice is that two queries are required for calls between operators. It would be more efficient if the first query would return the routing number instead of the directory number. In that case, only one query in the originating exchange would be required. However, this would require that the databases of all operators would contain complete information about all subscribers within the routing domain.

The proposal defines a solution for number portability that can rapidly be implemented. It utilizes as much of the existing solutions and technology as possible. At its final stage, the solution requires SSP capabilities in all exchanges, but the proposal also defines a first phase solution, which only requires SSP capabilities in some of the exchanges. It can be implemented in smaller steps towards the final solution. We first describe the final solution and then the first-phase solution.

3.5.2 Call setup procedure in the final solution

The following rules are used for number translation for an incoming call setup:

1. For an incoming call setup, the database query is performed in the first exchange in the trunk area. That means that the query is performed in the local exchange of the A-subscriber for calls within the trunk area, and in the first exchange in the destination trunk area for calls between trunk areas.

2. The query returns the final routing number if the B-subscriber’s operator is the same as the one that performed the query. If the B-subscriber is on another operators network, the query returns the destination operator and the call setup is sent trough interface 2. The destination operator then uses an additional query to map the directory number to a routing number.

3. If the B-subscriber is not found in the database, the subscriber has not moved and the query returns the directory number as routing number. A flag is used to indicate that the query has been performed, to inform the next exchanges that a new query is unnecessary.

When a call is set up from the PSTN, also the A-subscriber must be identified for the CLIP service. The CLIP (Calling Line Identification Presentation) service displays the number of the caller to the called party. The calling party is identified with the number on which he can be reached, which is the directory number, not the routing number. When the caller uses a portable number, the directory number is different from the routing number. The routing number of the A-subscriber has to be translated into a directory number before it reaches the end terminal or destination network. The IN based number portability solution solves the problem with an IN lookup in the first exchange. After that, the A-subscriber's number is always forwarded as a directory number.

3.5.3 Call setup procedure in the first phase solution

In the first phase solution, SSP capabilities are only required in some exchanges. For an incoming call setup, the query is performed in the destination exchange, compared to the first exchange in the final solution. If the exchange has SSP capabilities, the same procedure as in the final solution is followed. Otherwise, if the call is to one of the operator's subscribers, the setup is directed to an exchange with SSP. If the call is to another operator's subscriber, the exchange will direct the setup to the correct operator using the basic capabilities. If the originating exchange is not an SSP, the number translation of the A-subscriber number is directed to an exchange with SSP capabilities.

3.5.4 Number portability with dedicated portable numbers

The report [THK 1996] describes how a dedicated numbering space could be used for portable numbers. Many of the technical problems in the above solution would be solved, but it would require the subscriber to change his number when he moves to another operator or location the first time. A dedicated number space limits the need for database queries, since only the portable numbers require queries. Otherwise, all call setups require database queries. Another advantage is that a countrywide portable number does not cause routing through the old operator's exchange for calls between trunk areas, since the query can be performed correctly in the originating exchange.

A dedicated number space also allows for countrywide number portability. The number does no longer tell the geographical location of the subscriber. The situation is similar to the mobile networks. Because the number does not tell the location of the subscriber and consequently does not tell the tariff, the solution would require the tariffs to be independent of the location.

3.5.5 Restrictions of the IN-based number portability proposal

If we analyze the solution and compare it to the schemes presented in [Foster 2000], we conclude that the solution is a variant of All Call Query scheme for calls within the trunk area. This is an efficient solution, although the additional query required in the destination network slightly

degrades performance. However, for calls between trunk areas the solution works like the inefficient Onward Routing scheme. Calls to the moved subscribers are thus routed through the old destination exchange, which is a problem in the IN-based number portability proposal. Since the exchange is located in the old operator's network, this not only creates unnecessarily long routes, but also may raise a question about compensation for the used network resources if the load is unequally distributed. The problem can be solved by distributing the database across the whole country, which also would make countrywide number portability possible. The query could then be performed in the originating exchange and it would return the correct destination.

Another problem is the lack of standardized interfaces. Interface 1 between the master database and the operator's own database is not defined, and it is expected that the master database will probably have a different interface to each operator. With the increasing number of operators due to IP telephony, the need for standardization of the interfaces increases. Interface 2 is based on existing TUP/ISUP signaling, but uses a special number format and does not conform to international standards. [THK 1996, Kantola 1997]

4. IP telephony routing

In this chapter, we describe routing of IP telephone calls. The chapter begins with an introduction to IP telephony. We describe the addressing methods on the IP network generally and the specific issue how signaling protocols locate destinations. We explain the need for a global telephone numbering mapping and the solution of the ENUM working group. To provide calls from the IP network to the SCN, a gateway must be used. The gateway location problem is discussed. The Telephone Routing over IP (TRIP) protocol is described in detail, since it is the base for the protocol developed in later chapters. Finally, we discuss number portability in IP telephony networks and in hybrid IP-SCN networks.

4.1 Introduction to IP telephony

Contrary to the circuit-switched telephony network, IP telephony is packet based. Instead of sending a constant stream of samples, the samples are grouped into packets, which are sent separately from each other. In the receiving end, the packets are collected, buffered and played back as a voice stream. In the Internet, media streams are usually sent using the Real Time Protocol (RTP) [RFC 1889] and controlled by the Real Time Control Protocol (RTCP) [RFC 1889]. The main advantages of packet transmission are statistical multiplexing, efficient compression, variable bandwidth and silence suppression. As a result of variable network delay, multiple routes and lost packets in congested networks, the quality can be poor. The quality is improved by buffering, which in turn increases the delay. Different protocols are developed to improve the Quality of Service (QoS), among others the Resource Reservation Protocol (RSVP). The QoS issue is currently a popular research topic, and a lot of material can be found on it.

Signaling protocols are used to establish, control and terminate calls. They also perform address translation, location lookup and feature negotiation. In IP telephony, the signaling protocol have laid the foundation for the architecture and determined the distribution of functions between network elements. The main signaling protocols for IP telephony are Session Initiation Protocol (SIP) [RFC 2543] and H.323 [ITU-T H.323].

The H.323 protocol is developed by the International Telecommunications Union (ITU). It is a part of a larger series of communication standards called the H.32x series, which is used for multimedia conferencing over different types of networks. The H.323 specification has since its first version, which was approved in 1996, developed into its current version 3. The recommendation represents a traditional circuit-switched approach, based on the ISDN signaling protocol Q.931.

The Session Initiation Protocol (SIP) was developed by the Internet Engineering Task Force (IETF). SIP represents a more Internet-centered approach based on a more lightweight protocol. It reached the RFC status in 1999. The current version of SIP is version 2. Most signaling functions are integrated into a single protocol.

Although the protocols use different names for the network elements, the architectures are basically similar. H.323 defines four types of entities: terminals, gatekeepers, gateways and multipoint control units (MCU). SIP defines terminals, proxy servers, redirection servers and registrars. The proxy server and the redirection servers are signaling servers. SIP further divides the functionality of a terminal into a user agent server and a user agent client.

Terminals are either stand-alone IP telephones or programs running in a computer. The terminals are able to receive and place calls. Calls can be established between terminals directly, but usually the call setup signaling passes through a gatekeeper (in H.323) or signaling server (in SIP). The functions of the gatekeeper and signaling server are similar, so in this thesis we will use the name signaling server (SS) for both. The most important functions of the signaling server are address translation and location of the destination. They may also perform admission control and bandwidth management.

For calls between the IP telephony network and the SCN, a gateway must be used. The gateway converts the media stream and the signaling between the circuit switched network and packet network. The functionality of the gateway has later been divided into separate signaling gateways and media gateways. Further, the gateway can be controlled by a gateway controller.

The registrar of SIP is a server that accepts registration requests. It can be co-located with a signaling server, and it may provide a location service.

4.2 Addressing

In the Internet, the network layer provides connectivity between every host in a connected network. The intra-domain routing protocols, such as RIP [RFC 1058] and OSPF [RFC 1583], create and maintain routing tables for hosts in a single autonomous system (AS). The inter-domain routing protocols, of which BGP-4 [RFC 1771] is used today, distribute routes between autonomous systems. Every host is thus able to send signaling messages and establish connections to all other hosts in the network, provided that the destination is identified by an IP address. IP addresses are network layer addresses that are either 32-bit or 128-bit binary strings, depending on whether IPv4 or IPv6 is being used.

However, in practice it is much easier to use and remember textual names. The Domain Name System (DNS) [RFC 1034, RFC 1035] maps textual host names into IP-addresses. Host names are hierarchical and consist of domain names separated by dots. The mapping between the host name and an IP address is stored in DNS databases in a distributed fashion, corresponding to the hierarchical levels. To resolve a host name, the root DNS server is queried, which returns the address of a server with more specific information. The process is repeated until the IP address can

be obtained. To reduce network traffic and to speed up the resolving process, the results are cached. Using the definition of ETSI [ETSI TR 101 326], the IP address is an address and the host name is a name.

Since several users may be contactable on a single host, several applications add a user name to the host name or IP address. The address is then given in the format “user@host” or “user@ip-address”. E-mail is a typical example. The host is located using DNS and the user is contacted by the daemon or server running in the host. For e-mail, DNS allows for specifying a mail server responsible for mails for the whole domain, so that the host name can be a domain name instead of a complete host name. E-mail can also be forwarded to other hosts by the mail server.

The IP telephony signaling protocols support both IP addresses and host names. SIP uses a Universal Resource Locator (URL) [RFC 2396] in the format “user@host” or “user@ip-address”. The user name is optional, but often required. For calls to the SCN, the user part of the address specifies the telephone number and the host part the gateway, which gives an address in the format “number@gateway”.

Since the H.32x series covers several network technologies, it has a more generic approach to addressing. Each network element has a network address, which uniquely identifies the element on the network. In the IP-network it corresponds to the IP address. The entities also have a TSAP identifier, which is used for multiplexing several channels sharing the same network address. In the IP-network the TSAP identifier consists of the combination of an IP address and a port number. H.323 also supports limited use of alias addresses, which are textual names valid within one gatekeeper.

The “user@host” format is very suitable for Internet telephony, since the same address can be used for e-mail, telephony and other services. Functions like forwarding to another address can directly be applied to telephone calls in the same way as mail. DNS also provides some degree of portability, because the mapping can be changed to assign another IP address to the host name. Further, textual names are easy to memorize and familiar to users from e-mail.

Although textual names in the “user@host” format are suitable for IP telephony calls, they are not appropriate for calls from the SCN to the IP network. Textual names cannot be entered using a numeric keypad without using some complicated scheme for entering letters. Practically all SCN phones have only a numeric keypad, and the billions of old telephones will not disappear. Telephone number is the most widely used addressing and naming scheme for communications in the world [Shockey 2001]. SCN users are limited to numeric addresses, which must be in E.164 format to be compatible with existing numbering plans.

The problem was recognized by TIPHON, which chose to equip IP-terminals with E.164 numbers. Textual names can still be used for calls within the IP network. Because the user does not necessarily know the type of the destination network, if even knowing the type of his own network, it is recommended to use only E.164 numbers. In the same way as SCN-operators, Internet Telephony Service Providers (ITSP) can obtain ranges of E.164 numbers to assign to their

subscribers. Furthermore, a special country code has been reserved for global Internet telephony.

The use of E.164 numbers in IP networks requires a new mapping. An E.164 number must be mapped to a host name, or directly to an IP-address. This mapping is best performed by a directory. A client makes a query to the directory to map an E.164 number to an address in the call setup phase. In the case of a call from the SCN, the gateway performs the query. In addition to the mapping, a directory can contain other information about the subscriber, such as location, street address, alternative telephone numbers and other information that is usually found in traditional telephone directories. Directories can contain global information or only information about a specific network.

Directories are not a new invention. The X.500 directory model [CCITT X.500] was proposed in 1988 to provide a global directory for generic types of information. The X.500 did never succeed in large scale, mostly because of its heaviness. The directory access protocol (DAP) was too heavy for the clients of that time. Later, a lighter and simpler version of the directory access protocol was developed, called the Lightweight Directory Access Protocol (LDAP) [RFC 2251]. LDAP is more suitable for Internet applications.

In IP telephony, different types of directories have been used for different purposes and with varying success. In this thesis, we will describe two types of directories that are relevant to the subject. Firstly, the local directories used by the signaling servers are essential to the understanding of the IP telephony architecture. These are described in the following section. Secondly, a DNS-based directory has been proposed for global mapping of E.164 number into addresses. This directory, called ENUM after the IETF working group developing it, is presented in section 4.4.

4.3 Address resolution in the signaling protocols

In the IP telephony architecture, signaling servers (in SIP) or gatekeepers (in H.323) have access to directories over a limited group of terminals. The group may for example be subscribers of a Internet telephony service provider or employers of a company. Specifically, the SIP architecture defines the directory as an entity called location server (LS), which can be a separate element or integrated into the signaling server. The terminals register to a registrar, which obtains the current IP address of the terminal. The registrar provides the location server with information. In H.323 the directory is maintained by the gatekeeper. For generality, we will use the term location server for the directory of H.323 as well.

The location server contains a directory of the terminals that have registered with it. As calls are set up through the signaling server, the address is obtained from the location server. The protocol for querying a location server is not specified, but a popular candidate is LDAP.

Usually all terminals in an administrative domain or a subnet use the same location server. H.323 defines this as a zone, as illustrated in Figure 3. A zone consists of the network elements controlled by one or a group of gatekeepers.

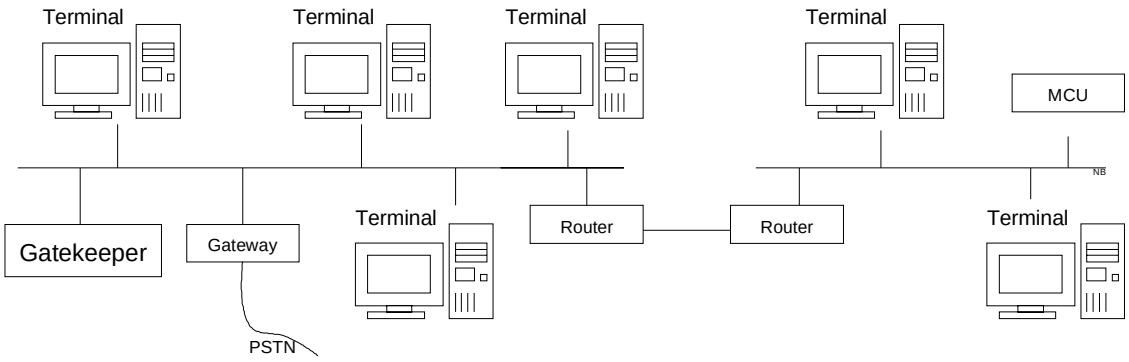

Figure 3. An example H.323 zone

In H.323, the concept of zones is important when alias addresses are used. The alias address is an alternative method for addressing an endpoint. Using ETSI terminology, it is a name. Alias addresses can be E.164 (partyNumber) addresses or H.323 IDs. H.323 IDs are strings, which represent e-mail addresses, conference names or user names. According to the H.225.0 specification, also other types of aliases are possible. Alias addresses are unique within a zone and an endpoint can have several alias addresses. Gatekeeper location is either automatic through multicast or manual. When a gatekeeper has been located, the endpoint registers its transport address and alias address with the gatekeeper. Hence, the gatekeeper contains a mapping between alias addresses and transport addresses for all endpoints in its zone.

SIP does not define any similar zone concept. A terminal can register to any registrar, provided that it has the permission to use it. A user may move between several locations over time. The new location is registered with the new SIP server using the register request of SIP [1]. The signaling server can redirect or proxy calls to another signaling server or to the end-terminal. Depending on capabilities and configuration, the server operates as either a proxy server or a redirect server. A proxy server may also fork the request to several servers simultaneously. To speed up the setup, proxy servers may perform parallel searches by issuing multiple requests without waiting for the result of the previous request. [RFC 2543]

A user will be located after a number of translations, as shown in Figure 4. A single address may lead to different host locations depending on the time of day, media to be used and other factors. SIP resolves addresses using the SRV, MX or CNAME records of DNS. If these fail, an SMTP server may be contacted to obtain an alternative address. The session description can even be sent as email, if all of the above fails.

[1] The server may additionally use other protocols to determine the location, such as finger [RFC 1288], rwhois [RFC 2167], LDAP [RFC 1777], multicast-based protocols or operating system dependent methods.

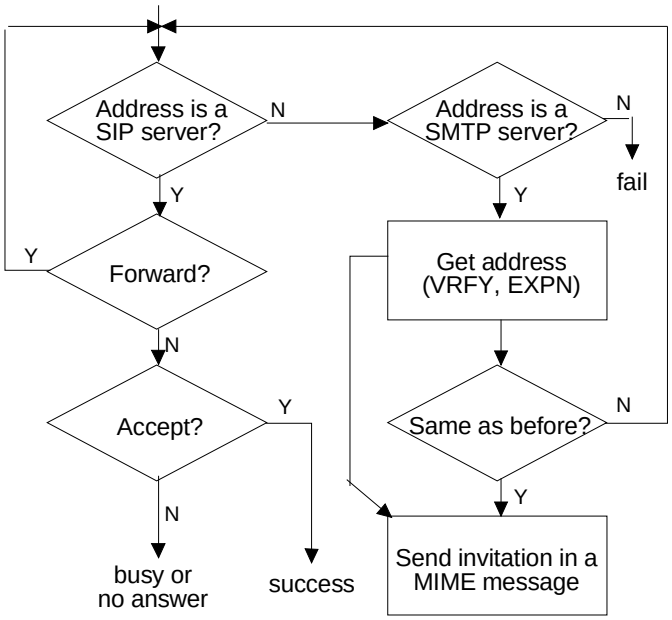

Figure 4. SIP address resolution

E.164 numbers can be used with both SIP and H.323 natively. The number is given as the user part of a SIP address, and as an alias address in H.323. The number is translated to an IP address using the mapping in the location server or gatekeeper. However, the mapping is local to the terminals using the same location server (in SIP) or belonging to the same zone (in H.323). Therefore, the E.164 number cannot be used globally.

To be globally accessible, the mapping must be distributed between all location servers. Although it is not its purpose, the Telephony Routing over IP (TRIP) [Rosenberg 2000a] protocol could be used. TRIP is described in section 4.5. Another approach is to use a global directory, as described in the following section.

4.4 Telephone Number Mapping

Many solutions for a global E.164 to address mapping have been suggested. Some of them are based on existing protocols like BGP [RFC 1771], X.500 [CCITT X.500], finger [RFC 1288], whois++ [RFC 2167], SAP [RFC 2974], search engines and indexing tools. Some are based on propriety solutions like ULS [Microsoft], ILS [Microsoft], Active Directories [Microsoft] and ICQ [Mirabilis].

The most complete and promising solution for the mapping is based on the Domain Name System (DNS) [RFC 1034, RFC 1035]. The solution is specified by the tElephone NUmber Mapping (ENUM) working group of the IETF. ENUM uses the existing DNS to map E.164 numbers into URIs. The URIs describe different ways to contact a host with a given telephone number. In addition to IP telephony, applications like email can utilize the service. A query returns an ordered list of URIs, which can represent SIP URIs, email addresses or LDAP servers, among others. The main document describing the procedure is [RFC 2916]. The draft [Brown 2000] describes a

telephone number directory service based on ENUM.

ENUM utilizes the hierarchical structure of DNS to distribute the directory between the authorities administering the numbering space. In a domain name, the domain name parts are given in increasing hierarchical order. For example, in the name “tele.tct.hut.fi”, the “fi” part is at a hierarchically higher level than the “tele” part. In E.164 numbers, the digits are in decreasing hierarchical order. By reversing the order of the digits, the hierarchy of DNS can be used to store E.164 numbers.

In the first stage of the mapping process, the telephone number is transformed into a domain name. ENUM uses the domain “e164.arpa” for storing E.164 numbers. The number must be in its full form, including the country code. All characters and symbols, except the digits, are removed. Dots are put between the digits. The order of the digits is reversed and the string “.e164.arpa” is added to the end. Using the procedure, the number +358-9-4515303 is mapped to the domain name 3.0.3.5.1.5.4.9.8.5.3.e164.arpa. [RFC 2916]

DNS uses different types of records to store information. Number mappings can be made with any type of record, but ENUM especially makes use of the Naming Authority Pointer (NAPTR) record [RFC 2915]. The NAPTR record is used to identify the available ways to contact a node with a given name. It tells what services exist for a given domain name. The fields of the record are shown in Table 1. ENUM defines a new service named “E.164 to URI”, which maps an E.164 number to a list of URIs. [RFC 2916]

Table 1. Fields of the NAPTR record

Name	Description
Order	The order for processing the records if the response contains several records.
Preference	The order for processing the records if several records have the same order value.
Service	The resolution protocol and resolution service that is available if the rewrite of the regexp or replacement field is applied. ENUM defines the service “E2U”, which stands for “E.164 to URI”.
Flags	Modifier affecting the way the next lookup is performed.
Regexp	Regular expression for the rewrite rule.
Replacement	Replacement for the rewrite rule.

Some example NAPTR records are shown in Figure 5. The records describe a telephone number that can be contacted by SIP, SMTF or using the “tel” URI scheme [RFC 2806] in the given order. The rewrite returns a URL, as indicated by the “u” flag.

```
$ORIGIN 3.0.3.5.1.5.4.9.8.5.3.e164.arpa.
 IN NAPTR 10  10 “u” “sip+E2U”  “!^.*$!sip:nbeijar@sipserver.tct.hut.fi!”  .
 IN NAPTR 100 10 “u” “smtp+E2U” “!^.*$!mailto:nbeijar@tct.hut.fi!”         .
 IN NAPTR 100 10 “u” “tel+E2U”  “!^.*$!tel:+35894515303!”                  .
```

Figure 5. Example NAPTR records

At higher levels, the hierarchical distribution is implemented using the NS, DNAME and CNAME records. The e164.arpa zone contains the addresses of regulators in different countries. The regulator performs the next DNS mapping into service providers. The service provider can further direct the query to other DNS servers, for example at a company. The entries in the regulator's name server could look like in Figure 6. [Brown 2000]

```
1.e164.arpa      IN NS ns.NANP.phone.net ; North America
3.3.e164.arpa    IN NS ns.FR.phone.net   ; France
8.5.3.e164.arpa  IN NS ns.FI.phone.net   ; Finland
```

Figure 6. Sample top-level delegations

4.5 Gateway location

For calls from the IP network to the SCN, a gateway must be located. The gateway functionality is usually split into the signaling gateway, the media gateway and the media gateway controller [RFC 2805, RFC 2885]. The signaling gateway converts signaling messages between the protocols on the SCN and IP networks. The call setup is sent to the signaling gateway with SIP or H.323. The signaling gateway allocates space for the call in a media gateway, and the media stream of the call is directed to the media gateway using the signaling protocol.

To specify a destination on the SCN, SIP uses an URL in the format "number@gateway". The number is the E.164 number of the destination on the SCN. The gateway is the host name or IP address of the gateway. The gateway acts like an end terminal, which terminates the call on the IP side and establishes a call on the SCN. [RFC 2543]

The gateway address must be known by the caller. Depending on the implementation, it is also possible to let the local signaling server choose a gateway. If the gateway address is omitted and the URL is only an E.164 number, the signaling server can recognize the destination as an SCN terminal, and send the call through a gateway. The signaling server then either uses a predefined gateway or selects a gateway from a configured list of available gateways. This solution is in practice limited to small networks since the signaling server must be manually configured with the available gateways.

In H.323 gateways are used and addressed in a similar way. The caller can establish a call to a given E.164 number on a given gateway. The gateway must be chosen by the client or the gatekeeper.

4.5.1 The gateway location problem

The terminal or signaling server must be manually configured with the available gateways. The list of available gateways must be updated when new gateways become available. As the number of gateways increases and their capacities and properties become more varied, the management

becomes more difficult. In a situation where the size of the IP telephony network approaches the size of the SCN, a large part of the calls passes through a gateway. There may be several gateways willing to complete a call, and selecting the most suitable is a non-trivial process. Several factors influence on the selection, including physical location, business relationship, capacity, protocols, media codecs and user requirements. Additionally, gateways may become blocked when all lines are in use. This problem is called the gateway location problem. [RFC 2871]

Above all, route selection is dependent on business relationships. Gateways may only be available for users with some established relationship with the gateway provider. The end user will generally not pay for the gateway service and the call termination directly. Instead, the user may have a relationship with an IP telephony service provider (ITSP) that may have own gateways, or that may act as an intermediary to gateway providers. The selection is thus largely driven by various policies. Therefore, a global directory of available gateways is not suitable. The information must be exchanged by providers and distributed according to policies. A protocol is needed for a policy driven distribution between service providers.

4.5.2 TRIP

To address the gateway location problem, the IETF working group IPTEL began developing a protocol for distributing information between gateway providers and IP telephony providers. The protocol was first called the Gateway Location Protocol (GLP), but as the scope was extended to more general path selection, the protocol was renamed to Telephony Routing over IP (TRIP). The protocol advertises routes to telephony destinations and the properties of these. The attention is put on calls from IP networks to the SCN. The main documents describing TRIP are the TRIP framework [RFC 2871] and the draft protocol specification [Rosenberg 2000a].

TRIP is an inter-domain protocol in the sense that it is used for exchange of gateway routing information between peers of telephony service providers. Policies control the distribution between these administrative domains. TRIP is based on the Border Gateway Protocol 4 (BGP-4) [RFC 1771] and has a similar transport mechanism, state machine and message formats. It contains additional functions for synchronizing information within a provider's network. The intra-domain flooding mechanism is similar to those of OSPF [RFC 1583] and SCSP [RFC 2334].

The service providers have a number of gateways and location servers. The set of resources under the control of an administrative authority constitutes an Internet Telephony Administrative Domain (ITAD). The location servers act as TRIP nodes and they announce routes to gateways in their domain. In addition to advertising own gateways, a location server also advertises routes to gateways in other domains by forwarding modified or unmodified advertisements. A route is defined as the combination of a set of reachable destination addresses, an address family and an application protocol (SIP or H.323). Additionally, routes are associated with a number of attributes. Every ITAD is identified by a unique numeric ITAD identifier.

A location server may modify the attributes of an advertisement forwarded from another domain. In this way, it can apply policies to routes that pass through the domain. To reduce the size of databases and messages, aggregation on the advertised routes is possible. Several routes with similar properties are then combined to a single route, obeying defined aggregation rules.

4.5.3 Protocol operation of TRIP

As TRIP is based on BGP-4, it borrows most of its functionality from BGP. The nodes form peer relationships using TCP connections. When a connection is established, the complete routing tables are exchanged. As the tables change, incremental updates are sent. The tables are not periodically refreshed, so all received routing entries must be stored.

TRIP separates between internal and external peers. Internally TRIP uses link-state mechanisms to flood updates. The information is synchronized between the internal nodes. Policies are applied on the information exchanged between external peers.

Location servers process three types of routes:

1. External routes, which are received from a location server in another ITAD.
2. Internal routes, which are received from another location server in the same ITAD.
3. Local routes, which are originated within the ITAD. These are locally configured or received from another routing protocol.

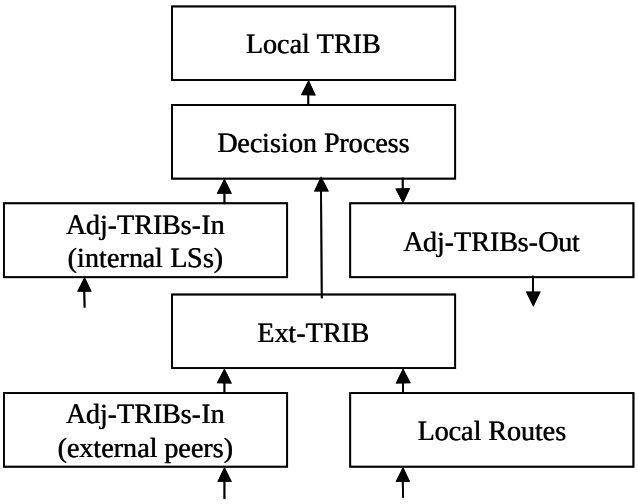

Figure 7. The structure of a TRIP node

The structure of a TRIP node is shown in Figure 7 [Rosenberg 2000a]. The arrows show the information flow within the node. The node consists of four types of databases, called Telephony Routing Information Bases (TRIB):

1. The *Adj-TRIBs-In* databases contain unprocessed routing information that has been received from other peers. These routes are available as input to the decision process. Routes from each external and internal peer location server are maintained independently in the database. There

is an Adj-TRIBs-In for every location server of the domain, also for those for which there is no direct peer relationship.

2. The *Ext-TRIB* contains the preferred route for each destination, as selected by the route selection algorithm.

3. The *Loc-TRIB* stores the local routing information that is selected by applying the local policies to routes from Adj-TRIBs-In and Ext-TRIB.

4. The *Adj-TRIBs-Out* store routing information that is selected for advertisements to external peers. There is one database for each peer.

TRIP uses the same messages as BGP-4. The *Open* message is sent after the TCP connection has been established. It contains the TRIP version, ITAD identifier, TRIP identifier, Hold time and a number of optional parameters. The optional parameters can be used to indicate capabilities, such as the supported route types. To confirm the Open message, a *Keep-alive* message is sent. Keep-alive messages are also sent regularly to confirm that the connection is working. The maximum interval of Keep-alive messages is given in the hold time parameter. *Notification* messages are used to inform about errors. After a Notification message the connection is closed, and can only be opened with a new Open message. [Rosenberg 2000a]

The *Update* message carries the routing information as a list of attributes. The attributes are identified with an integer code. New optional attributes can be added later to allow for expansion. Like most information in TRIP messages, the attributes are coded in type–length–value format. [Rosenberg 2000a]

Attributes are equipped with flags that indicate how they are handled if they are unrecognized by some node. The specification [Rosenberg 2000a] defines the flags *Well-known, Transitive, Dependent, Partial* and *Link-state Encapsulated*.

A well-known attribute must be supported by all implementations. An attribute that is not well-known is non-transitive, dependent transitive or independent transitive. An unrecognized non-transitive attribute must not be propagated by intermediate location servers, while an unrecognized independent transitive attribute can be propagated by intermediate location servers. An unrecognized dependent transitive attribute may only be propagated if the next-hop server attribute is not changed. The partial flag is set when the attribute is propagated by a location server that does not recognize the attribute.

The link-state encapsulated flag only applies to the Reachable Routes, Withdrawn routes and ITAD topology attributes. It indicates that link-state encapsulation is used instead of standard encapsulation. In link-state encapsulation, a sequence number and the TRIP identifier of the originator are added to the attribute to detect duplicates and old routes in the flooding process.

4.5.4 TRIP attributes

Attributes contain information used for routing and for protocol operation. The attributes form the basis for route selection. A route selection algorithm can for example prioritize a route that contains or does not contain a specific ITAD.

The following mandatory well-known attributes are defined in the draft specification [Rosenberg 2000a]:

- The *Reachable Routes* attribute represents routes that become reachable. The attribute contains prefixes reachable through the advertising domain, and it functions as a key for the rest of the update message. For every prefix an application protocol and address family is given as well. The application protocol is SIP or some variant of H.323. The address families are defined in [RFC 1700].

- The *Next Hop Server* attribute contains the IP address where signaling messages should be sent for call setup. The address may identify either a signaling server or a gateway. The attribute also contains the ITAD identifier of the next hop.

- The *Advertisement Path* attribute indicates the path that the route advertisement has traversed. It is used to prevent advertisements from looping. Like in BGP-4, it is given as a list of path segments. The segments are either sets or sequences of ITAD identifiers.

- The *Routed Path* attribute indicates the path that the signaling messages will follow. Routed Path is a subset of Advertisement Path, since it only includes the hops that have modified the Next Hop Server attribute. The format is similar to the Advertisement Path attribute.

- The *Withdrawn Routes* attribute is exceptional, since it is independent of the other attributes. It defines routes that are no longer reachable and must be removed. The format is similar to the Reachable Routes.

In addition to the mandatory attributes described above, the draft defines a number of optional attributes:

- The *Atomic Aggregate* attribute indicates that the Routed Path does not include all ITADs that the signaling may traverse. This situation may arise due to aggregation.

- The *Local Preference* attribute is only used between internal peers to indicate the calculated preference for the routing object to other location servers in the domain.

- The *Multi Exit Disc* attribute is only used between external peers. It is used to specify the preference for routes received over one link compared to routes received over another link, when several links connect two ITADs.

- The *Communities* attribute is used to group destinations sharing some common property so that the routing decision can be based on the identity of the group.

- The *ITAD Topology* attribute is used by the location server to advertise its internal topology to other location servers within the domain. It is used for failure detection.
- The *Converted Route* attribute indicates that some location server on the path has changed the application protocol field in the Reachable Routes attribute.

TRIP can be extended with new attributes. Especially a pricing metric has been suggested. The pricing information should be an optional attribute, since it should be possible to use different pricing metrics [RFC 2871]. Also security is implemented with optional attributes. The *Authentication* attribute defined in [Rosenberg 2000b] contains a list of signatures, which are used to verify the originator and the contents of selected attributes. If an implementation does not recognize an optional attribute, the attribute is handled according to the attribute flags.

Since TRIP routes describe numbers reachable through gateways, it would be useful to attach properties of gateways to the advertisement. Such information would for example describe the currently available capacity of the gateway. The draft [Rosenberg 2000c] defines a *DSP Capacity* attribute and a *Circuit Capacity* attribute as part of the TRIP-for-Gateways proposal. Because of their dynamic nature, these are only sent from the gateway to the first location server.

4.6 Number portability with ENUM and TRIP

Number portability is also required in the IP network. Basic number portability within the IP network can be implemented with ENUM by modifying the mapping in DNS. The directory service solution defined in [Brown 2000] describes number portability on three conceptual levels: between authorities, between service registrars and between service providers.

Number portability in an interconnected SCN and IP network scenario is discussed in [Lind 2000]. It analyses call setup from SCN to IP and from IP to SCN in the scenarios when a number moves within the SCN, within the IP network, and between the SCN and the IP network.

Number portability within the SCN is handled with IN based number portability. However, to provide efficient routing of calls from the IP network to the SCN, the information in TRIP may need to be updated since another gateway may be more suitable due to the move. In case TRIP contains routing addresses, these addresses must be updated. Number portability within the IP network can be handled by updating DNS information. For numbers moving between the SCN and IP network, the information in ENUM and the IN network must be updated. In many cases also the TRIP information is modified.

Due to number portability, situations often arise where the information in several protocols must be updated simultaneously. Updating is currently performed manually. The information in TRIP, ENUM and IN databases must be consistent to prevent misrouted and lost calls. Further, the update must be coordinated between operators to avoid inconsistencies.

TRIP provides automatic selection of gateways for calls from the IP network to the SCN. Nevertheless, in the other direction, there is no similar method. The gateways must be manually configured into the routing tables, or some proprietary method must be used.

5. A routing protocol approach for the SCN

Hitherto, we have described IP telephony solutions including signaling protocols, number directories and a gateway location protocol. We have also described the basics of E.164 numbering, number portability and routing on the SCN. In this chapter, we discuss the motivation and requirements for a distributed routing protocol for the SCN. We describe the principles and operation of a TRIP based solution for distribution of routing information within the SCN and between the SCN and IP networks. A counterpart to TRIP for the SCN is developed. We then integrate ENUM into the solution. In this chapter, the solution is described at a conceptual level, which allows analyzing usage scenarios and establishing the requirements for the more detailed protocol specification.

5.1 Motivation

As we have seen, IP telephony is becoming increasingly more mature. The many benefits of IP telephony make it a commercially feasible solution. A popular view is that the telephony network will gradually become completely IP based. Before this completely IP based scenario is possible, there will be a long transition time when both the networks co-exist. Probably the SCN will never be completely replaced. It is therefore necessary to make the IP telephony network a full peer of the SCN. This requires that the networks are connected with multiple connection points, so that services on one network can be accessed from the other and a suitable connection point can be automatically located. [Kantola 2000]

The development of IP telephony is reaching a complete set of protocols, allowing calls to be established between terminals on the IP network and between the SCN and IP networks. The endpoint location and route selection processes, which previously were performed completely by the signaling protocols, have been extended with two new protocols: TRIP and ENUM. TRIP is used to automatically locate suitable gateways for the calls and to form routes towards destinations on the SCN. The information distributed with TRIP is used in call setup, where the client queries a location server to obtain the IP address of the next hop server on the path towards a given E.164 number. The route selection is driven by policies. TRIP works like an application layer routing protocol. It could also be used to distribute information about IP terminals, but according to current plans, ENUM is used for that. The DNS-based directory ENUM is used to locate terminals on the IP network by performing a mapping between E.164 numbers and IP addresses. Basic number portability can be implemented using the directory. ENUM relies solely on the network layer for routing.

The current TRIP draft does not specify the origin of the reachablity data. The data describes the numbers reachable through the gateways, and the properties of the gateways and the routes towards them. Currently this data must be manually configured for each gateway. With a routing protocol, the data could be collected from the SCN automatically. [Kantola 2000]

In the SCN, much of the work related to route configuration and service deployment is done manually or by using proprietary solutions. Generally, significant amounts of manual configuration are required to introduce a new numbering range, a new operator relationship or a new route. With more revenues coming from services rather than voice transmission, the management of service related routing is increasing. The introduction of IP telephony increases the configuration and maintenance load even more. Gateways require configuration and new routes must be installed. Due to number portability and mobility, a number may move between locations and operators. With IP telephony, a number may also move between the IP and SCN networks. Number portability leads to more complexity in management and routing. In addition, the number of operators is expected to increase when IP telephony and other network technologies such as 3G and WLAN are introduced.

With an automatic routing protocol of the type that is used on the Internet, much of the route configuration and maintenance could be automated. An application layer routing protocol similar to TRIP could automatically form routes according to the policies of the operator. This layer provides high-level routing based on application layer addresses, i.e. directory numbers in E.164 format. It relies on the network layer for routing in the network topology using routing numbers.

The routing process could consider network topology and minimize crossing of technology borders. An important issue is to minimize or avoid conversion between packet and circuit switched transmission and between codecs. Each media conversion decreases the quality by adding jitter, delay and other forms of degradation. In IP telephony, these occur in the media gateways between the SCN and IP networks, but similar media conversions also arise in third generation mobile networks. Further, new services put increasing demands on routing.

The benefit of automatic routing grows with a larger number of operators. As numbers can move between networks and operators, it is important to know where the destination currently resides at an early stage in call setup to be able to choose an optimal route. A goal would be to eliminate the need of routing calls through the donor network when number portability is used.

5.2 Requirements

Numbering information, such as the mapping between names and addresses, is currently stored in the IN databases on the SCN, and in location servers and DNS on the IP network. The planned solution should be based on these elements.

Enough information must be shared between operators and network technologies to allow for efficient routing. However, the information amount must be limited to improve scalability and to hide the network topology from other operators. To meet these requirements, the solution must have some form of aggregation of information. With aggregation, the information amount is

reduced and topological details are hidden. Further, the policies of the operators and service providers must be considered. All information should be in standardized formats.

According to our objectives, existing protocols should be used wherever it is possible. Since TRIP and ENUM already are specified for the IP network, our solution should include them without any significant modifications. For the SCN we must design a new protocol, since there are no routing protocols. This protocol should be based on existing protocols as much as possible, and it must be compatible with the protocols on the IP side. We choose to use TRIP as the base, since it is an application level routing protocol performing similar functions. A routing protocol has several advantages, of which the most important might be the ability to have different information at different locations. This is necessary for implementing policies and aggregation.

In order to reach the goal, we need to

1. Specify routing and addressing
2. Examine usage scenarios and the required information
3. Design the architecture
4. Define the information
5. Develop a protocol for distributing numbering information on the SCN

The first three steps are mainly done in this chapter. The following chapter concentrates on the last two.

5.3 Routing and addressing

The solution adds a new layer of routing above the network layer. Both the network technologies have their own network layer routing protocols suitable to their respective technology. However, the application level routing is common to both technologies. The division of functions between application and network layer is distinct: the network layer routes the media streams while the application layer provides routing for the telephony service. The application layer uses directory numbers in E.164 format as addresses (“names” according to ETSI). E.164 numbers can be entered with IP terminals as well as with traditional SCN phones, and they serve as the common denominator in identifying destinations on both network types. In the future, textual addresses would be a considerable alternative as application layer addresses.

The primary goal of the planned solution is to provide automatic distribution of information between the network types and protocols in order to make routing more efficient. This means populating the databases of the Intelligent Network with information from TRIP and ENUM. In the opposite direction, it means dynamic creation of the information in the location servers, that is the information distributed by TRIP. TRIP describes the paths towards destinations on the SCN network with a list of ITADs that signaling will pass through. It gives the address of the next signaling server on the path, and for the last hop it gives the address of the gateway. In our solution we should be able to create similar paths on the SCN.

5.3.1 Telephony administrative domains

On the SCN, networks owned by different operators are interconnected with each other according to contracts. Similar contracts are made between Internet telephony service providers (ITSP). The operators and ITSPs are the administrative owners of the corresponding networks. TRIP uses the term ITAD for the network resource owned by an ITSP. To be able to use the same term for the networks owned by both operators and ITSPs, we will use a more general term: Telephony Administrative Domain (TAD). The TAD can be identified with a similar numeric identifier as the ITAD. In the description of our solution, the TAD identifier is used to identify the collection of network resources owned by a single administrative unit. These must be of a single network technology. Thus, if an operator owns both an IP telephony network and an SCN network, they are identified by different TAD identifiers. Thereby, the TAD identifies one operator's network based on a single technology.

Each TAD has at least one numbering database. If the TAD contains several numbering databases, i.e. for redundancy, these must contain the same routing information. This is consistent with the TRIP approach. The information in the databases represents routes towards the destination for each numbering range. The central piece of information in a route is the path to the destination. The path is described as a list of TAD identifiers towards the destination.

5.3.2 Prefixes and aggregation

Routes are created for groups of telephone numbers, which are described using prefixes. The prefix acts as the key for the route. In the context of our solution, we extend the concept of prefixes: we don't separate between prefixes and individual numbers. Where a prefix is required, also a single number can also be given. Consequently, a prefix can represent either an individual number (for example 35894515303) or an actual prefix as used on the SCN (for example 3589451). Using this notation, the prefix 3589451 includes all numbers beginning with the digits 3589451, with the length of seven or more digits. For example, the numbers 35894518932 and 35894515303 matches this prefix. On the other hand, the prefix 35894515303 might only match the number 35894515303.

Using this extended definition of prefixes, we can mark prefixes and individual numbers in the same way. This generic notation is also used by TRIP. We do not need any flag indicating whether the route is for a prefix or a number. The second major advantage is that this makes aggregation possible. With aggregation, it is possible to describe the numbers 35894515300 to 35894515309 as a single prefix 3589451530. This prefix describes ten numbers. Further aggregation may combine the prefixes 3589451530 to 3589451539 into 358945153, which describes 100 individual numbers with a single prefix. There is thus no need to separate between subscriber numbers and prefixes.

Aggregation reduces database size, the required processing power and the amount of information to be distributed and synchronized. Aggregation requires that the properties of the information to aggregate are to a defined degree similar. The conditions and rules for aggregation are defined separately for each attribute. Aggregation of decimal numbers needs ten numbers of length *n* to be

combined into one of length *n-1*. Aggregation of pentadecimal requires fifteen numbers. Pentadecimal numbers are mostly used for routing addresses. For aggregation, we need to know whether the number is decimal or pentadecimal.

5.3.3 Alternatives to prefixes

This type of aggregation has the disadvantage that it requires 10 or 15 numbers to form one. Corresponding aggression in IP routing protocols needs only two, because of the binary representation. Binary representation of telephone numbers is, however, not suitable. Alternative methods would be to describe the numbers as ranges or as patterns (regular expressions).

A range matches the numbers between the given bounds, the bounds included. For example, the range [35894515300, 35894515303] matches the numbers 35894515300, 35894515301, 35894515302 and 358945153003. The range in the previous example could also match the longer number 358945153014. The question how to handle numbers of different length must be defined separately. In this sense, ranges work like two prefixes, and matching works like prefix matching. The range [1230, 1239] corresponds to the prefix 123, which also can be described as the range [123, 123].

Aggregation of ranges can be more efficient than aggregation of prefixes. It takes 10 sequential numbers to form one aggregated prefix, but a range can also aggregate shorter sequences. For instance, the numbers 35894515300 to 35894515308 can be described as the range [35894515300, 35894515308] but cannot be aggregated into 3589451530 since the number 35894515309 is excluded. However, assuming longest match, the range can be described as two prefixes, where 3589451530 specifies the given range and 3589451309 specifies some other range. Nevertheless, handling of ranges is not as explicit as handling of prefixes.

Regular expressions [IEEE 1993] are powerful, and can be used to match a wide variety of number patterns with a short expression. For example, the regular expression "^35894153\d*" matches numbers beginning with 35894153 and continuing with zero or more other digits. That is, it matches the same numbers as the prefix 35894153. Another example is the expression "^358941530[1-4]", which matches numbers beginning with 35894153 followed by one of the digits 1, 2, 3 or 4. Further, the expression "^3589451\d{4}" matches 11 digits long numbers beginning with 3589451. The power of regular expressions is superior to both prefixes and ranges, but the required processing power is higher.

Although ranges and regular expressions are more efficient, we choose to use prefixes in our solution. The reason for choosing prefixes is compatibility. Prefixes are the common method for representing a group of numbers on the SCN. Exchanges analyze the numbers digit-by-digit beginning from the most significant digit. The number is compared to the alternatives digit by digit until a single alternative can be selected unambiguously. If a number matches several prefixes, the longest prefix is selected. The longest prefix is said to be the most specific.

Prefixes are also used by TRIP. It is a significant advantage to use the same format for the key field on both the SCN and IP networks. Although regular expressions provide powerful tools for matching telephone numbers and prefixes, the conversion to and from prefixes used by TRIP and ENUM would complicate the solution and create overhead in the conversion process. Matching of regular expressions requires much processing power. Because of the wide expression possibilities, it is not possible to use the regular expression as an unambiguous key, and a lookup could require matching with all keys in the databases in the worst case. Compared to ranges, prefixes need fewer comparisons for matching and less storage space if the degree of aggregation is low. Matching of prefixes is well defined.

The selection to use prefixes has some consequences. A prefix of length *n* matches numbers with all values in the digits following the *n*:th digit. If it does not cover all those numbers, it is said to contain holes. To correctly match only existing numbers, the prefix must not contain holes unless the holes are known by all other nodes. A way to make the holes known is to represent them as entries with longer prefixes.

5.3.4 Routing addresses

TRIP defines the Next Hop Server as the IP address, where signaling messages are sent for establishing the call. The address represents a signaling server, a gateway or some other element processing signaling messages. The Next Hop Server is given on a hop-by-hop basis, so signaling messages can traverse several servers on the path to the destination.

On the SCN we need a different method for finding the path to the destination, mainly because of services like number portability. Number portability maps a directory number into a routing number, which is used for routing. The routing number can be created from the directory number by adding a prefix in front of the directory number or a part of the directory number.

The routing number can also be obtained from a directory through a query. Different services may map a directory number into a routing number depending on external parameters, for example time of the day or subscriber information. The address of the next hop should therefore be possible to generate dynamically or obtained through a query. Furthermore, different protocols can be used for the query.

Consequently, we need several methods to locate the next hop. We have the following alternatives:

- To generate a routing number by modifying or replacing the directory number
- To locate the next hop address through a query with a given protocol to a given address on the SCN
- To locate the next hop address through a query with a given protocol to a given address on an IP network

For each alternative, we must examine the required information and define the mapping process.

5.3.4.1 Generation of routing numbers

Calls to moved numbers are established using a routing number. In our solution, a different routing number can be used for each hop. The numbers are generated from the directory number using a pattern. Parts of the directory number can be inserted and new digits added. A usual case is to keep the last digits of the directory number and add new digits in front of the number for routing to another network. In this way, we can generate the routing numbers of IN-based number portability.

It would be possible to define new methods for describing the pattern, but we prefer to use standardized ones. The natural choice for this is to use regular expressions. In this case we are not limited because of the above compatibility reasons, since the pattern is only used on the SCN. Neither the higher processing power is a limit, since the next hop address is not a key field. It is only processed on call establishment.

Now we need to perform substitution with regular expressions instead of matching. Substitution rules can be written as strings in various manners, but we prefer to use the notation used in the NAPTR resource record of DNS [RFC 2915], which performs a similar function. The NAPTR resource record is also used by ENUM. The substitution expression defined in [RFC 2915] has the format shown in Figure 8.

```
subst_expr   = delim-char  ere  delim-char  repl  delim-char  *flags
delim-char   = "/" / "!" / ... <Any non-digit or non-flag character
               other than backslash '\'. All occurances of a delim_char
               in a subst_expr must be the same character.>
ere          = POSIX Extended Regular Expression
repl         = 1 * ( OCTET /  backref )
backref      = "\" 1POS_DIGIT
flags        = "i"
POS_DIGIT    = %x31-39
```

Figure 8. Substitution expression format

The expression consists of two parts: the regular expression and the replacement. Optional flags can be added. The parts are separated by an arbitrary character, but we use the “!” character in our examples, since it is used by ENUM. The POSIX extended regular expression is defined in [IEEE 1993]. The back-reference refers to a numbered matched sub-pattern in the regular expression. The pattern can be divided into sub-patterns using parenthesis. The flag “i” (ignore case) is not applicable to number substitution.

We demonstrate the application by an example. The directory number 35894515303 can be mapped into routing numbers with a variety of expressions. The expression “!(\d{4})(\d*)!5555\2!” gives the routing number 55554515303 by removing the first four digits and replacing them with 5555. The same routing number is obtained with “!(\d*)(451)(\d+)!5555\2\3!”, which removes any digits before the sequence 451 and creates a number consisting of 5555 and the following digits. The pattern “!(\d{3})(\d)(\d*)(\d{4})!5555\2\4!“ extracts the fourth digit and the last four digits and adds the prefix “5555” in front. The frequently used digit matching “\d” can be replaced by “.”,

which matches any character, since the number is only expected to contain digits. In the case of pentadecimal numbers, it is however necessary to use character matching.

5.3.4.2 SCN queries

Queries on the SCN can be made with Signaling System 7 (SS7) to a HLR (home location register) or SCP (service control point). We need to have the following information:

- Destination point code (DPC) and subsystem number (SSN)
- Query number
- Protocol (MAP, INAP, LDAP)
- Protocol version
- Additional protocol specific information

The query number is formed like the routing number described above. With a regular expression, the directory number can be rewritten and a prefix can be added, resulting in a number, for which the query is made. We get more expression power without using the combination of prefix and cut-first rule. The required information can be coded into the next hop attribute as separate fields. To allow for future protocols, the protocol field must be expandable.

5.3.4.3 IP queries

With IP queries, queries can be made directly to services implemented on the IP network or to location servers. Queries to hosts on an IP network require the following information:

- IP address (IPv4, IPv6) or host name
- Query number
- Protocol (MAP, INAP, LDAP)
- Protocol version
- Additional protocol specific information

On the IP network, this type of queries are usually written as an URL. URLs are defined for LDAP [RFC 2255] but not for INAP and MAP. Instead of defining new URLs, we choose to use the same approach as for SCN queries by defining fields for each parameter.

5.4 Routing scenario

To describe how the information is used in our model, we present a scenario containing two IP telephony networks and three switched circuit networks. The IP and SCN networks are interconnected through two gateways. Connections between the IP networks are provided through routers in a normal way. Pairs of SCN networks are interconnected. The network is illustrated in Figure 9.

Each IP network contains a signaling server (called SS1 and SS2) responsible for routing incoming calls. All networks have a numbering database, which is used in the call setup to locate the next hop to the destination. At this stage, we will not discuss how the information is obtained or distributed. We concentrate on how the information is used in the call setup. The networks are identified by their TAD identifier (numbered from 1 to 5 in the figure). In this scenario, routing prefixes (marked RP1-RP6) are used to locate the neighboring networks and gateways. These routing prefixes are added in front of the destination number. If the call arriving from a neighbor network already has a routing prefix, it is removed before the new prefix is added. This function is easy to implement using regular expression substitution.

In the notation, we describe the routing information databases as a collection of routes. The route has a mapping between directory number prefix and next hop address. It also has a path list, with the TAD identifier and network type of each hop on the path. The path is given as a list in the format {(identifier1, type1), (identifier2, type2), …, (identifierN, typeN)}.

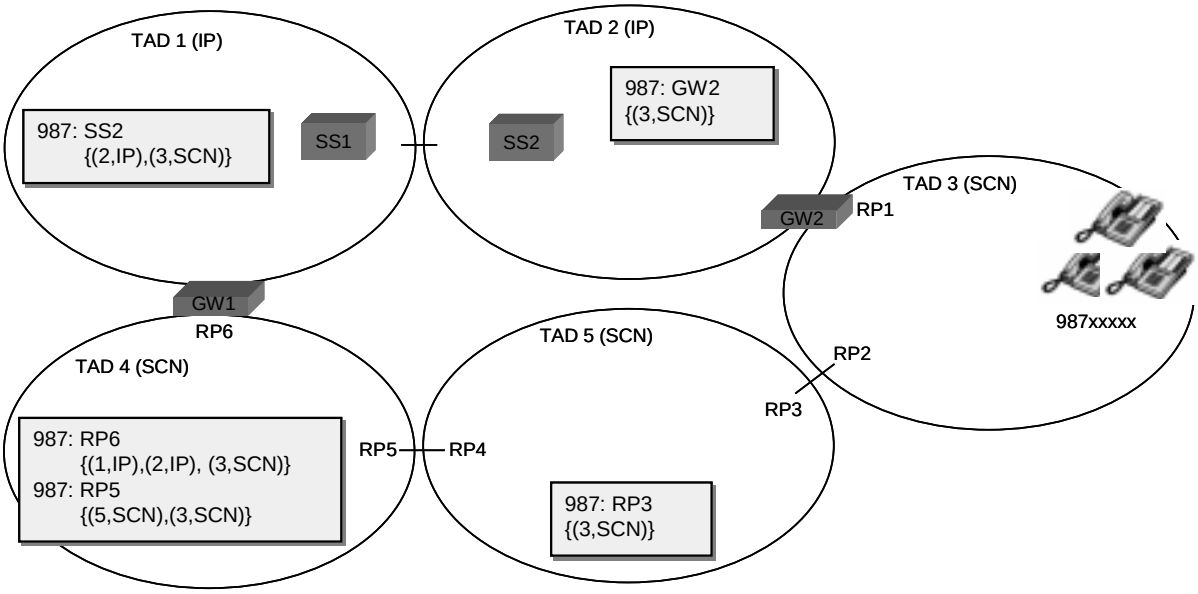

Figure 9. Routing information scenario

In this scenario, a range of phones, whose numbers begin with the digits 987 are located in the TAD identified by number 3. Thus, they share the common prefix “987”. The prefix has moved to this network with number portability. The network of this TAD is a switched circuit network. Routes towards destinations beginning with this prefix are stored in the databases of all other TADs. The TADs numbered 2 and 5 can reach TAD 3 through one single hop. TAD 2 is an IP telephony network. It reaches the neighboring TAD 3 through a gateway. Therefore, it stores the IP address of the gateway as the next hop address. The path list contains a single element, which is the SCN-based TAD 3. TAD 5 is an SCN, which is connected to TAD 3. To reach the destination network, a routing prefix is added in front of the destination number. The prefix is marked as RP3 in the figure. The next hop address is the regular expression for generating this routing prefix. The path list contains a single element.

The IP based TAD 1 is connected to the neighboring IP network TAD 2. The numbers beginning with 987 can be reached through the domains TAD 2 and TAD 3. Signaling messages are sent to the signaling server SS2, whose IP address is the next hop address. The signaling server handles the incoming calls to TAD 2.

TAD 4 is an SCN network. This TAD can establish a call to the given number through TAD 5 and TAD 3. A routing prefix RP5 is added in front of the number to route the call to the neighboring SCN network. In TAD 5 it is replaced by another routing prefix RP3, which routes the call to the destination network. TAD 4 only needs to know the regular expression for generating the routing prefix of the neighboring domain. In addition to this path, the call can also be made through TAD 1, 2 and 3. In this case, the first hop is a gateway. To establish a call through the gateway, a specific routing prefix is added in front of the directory number. The switches in the SCN are programmed to route numbers beginning with this routing prefix to the gateway. The numbering database of TAD 4 thus contains the routing prefix as the next hop address. Call signaling from this TAD passes through the IP based TADs 1 and 2, as indicated by the routing path. Of these two alternatives only one is used. The selection of the used path is performed according to the policies of the operator. The shorter path is completely switched circuit based, while the longer passes through two IP telephony networks and two gateways. The operator might want to transfer the call to the IP network to save costs by using IP telephony. This is especially attractive if the operator also owns the IP network of TAD 1, or in another way has a relationship with the Internet telephony provider. On the other hand, the operator might want to use only SCN networks, for example for quality reasons. As described in chapter 9, also the dialer can choose the route through carrier selection.

For routing, only the next hop address is needed. The location server or IN-database is queried with the directory number, whereas the next hop address of the longest matching prefix is returned. As we can see from the scenario, the next hop addresses are IP addresses on the IP telephony networks. The switched circuit networks use regular expressions to generate routing prefixes as next hop addresses. In each network, the call is routed to the address given as the next hop address. The call can proceed from a network type to another. The path list is used for route selection, but it is not critical for the success of call routing.

Note that in the above scenario there are more ways to set up a call than the described ones. For example, a call from TAD 1 to TAD 2 could also be established through the domains TAD 4, TAD 5 and TAD 3. Moreover, a call setup from TAD 2 to TAD 3 could traverse the longer path through TADs 1, 4, 5 and 3. This path could be used if gateway GW2 were down. The availability of these alternative paths depends on the policies used by the different service providers as explained later.

5.5 Information distribution

To create routing information of the type described above, protocols are needed to distribute the information between the numbering databases. In our solution, we would like to utilize the existing protocols as much as possible.

On the IP network, TRIP and ENUM cover all terminals reachable from the IP-network. These protocols are included into the solution. Some adaptation may be required to make them suitable for use in an interconnected SCN and IP network.

We would like to utilize the framework for IN-based number portability in the SCN. In the SCN, numbering information is used by the nodes of the Intelligent Network. The information is stored in databases in the SDF. In the IN-based number portability model, the information is distributed in one direction from a master database to the SDFs of each operators network. The updates are made to the master database. Our goal is to develop a distributed approach without any master database, and where all databases have equal value. The information is located in the nodes where it is needed and the updates are performed by the entities involved in the transaction. A protocol is required to distribute and synchronize information between the databases. On the SCN there is no standardized protocol that can be used to distribute numbering information. As we find it required, we will develop such a protocol. The protocol is used to distribute information needed for the number mapping.

Since the purpose of the protocol is the same as the purpose of TRIP, we expect that much of the functionality of TRIP can be borrowed. TRIP provides a quite general mechanism for both intra- and inter-domain distribution of routing information. Also the routing information infrastructure is similar. We plan to use TRIP with a different set of attributes in the solution. Since the resulting protocol will be used to distribute routing information over the switched circuit network, we call it *Circuit Telephony Routing Information Protocol* (CTRIP). CTRIP carries numbering information about both SCN- and IP-terminals. It acts like the SCN counter-part to TRIP. The protocol operation, node structure and advertisements are similar to TRIP. Similarity to TRIP guarantees smooth interworking between the protocols. Only the carried information differs.

5.5.1 Information distribution scenario

To illustrate how the combination of TRIP and a TRIP-like protocol on the SCN can be used to create and distribute routing information, we present a scenario. This scenario uses the same set of networks as in Figure 9. TRIP is used between domains on the IP network. The new protocol CTRIP is used between SCN operators. On the border between IP and SCN based networks, some type of protocol converter is needed. This element will be described in the next section. Similarly to TRIP, CTRIP is a peer-to-peer protocol used between networks that have an agreement on exchanging routing information. The protocol connections are shown in Figure 10.

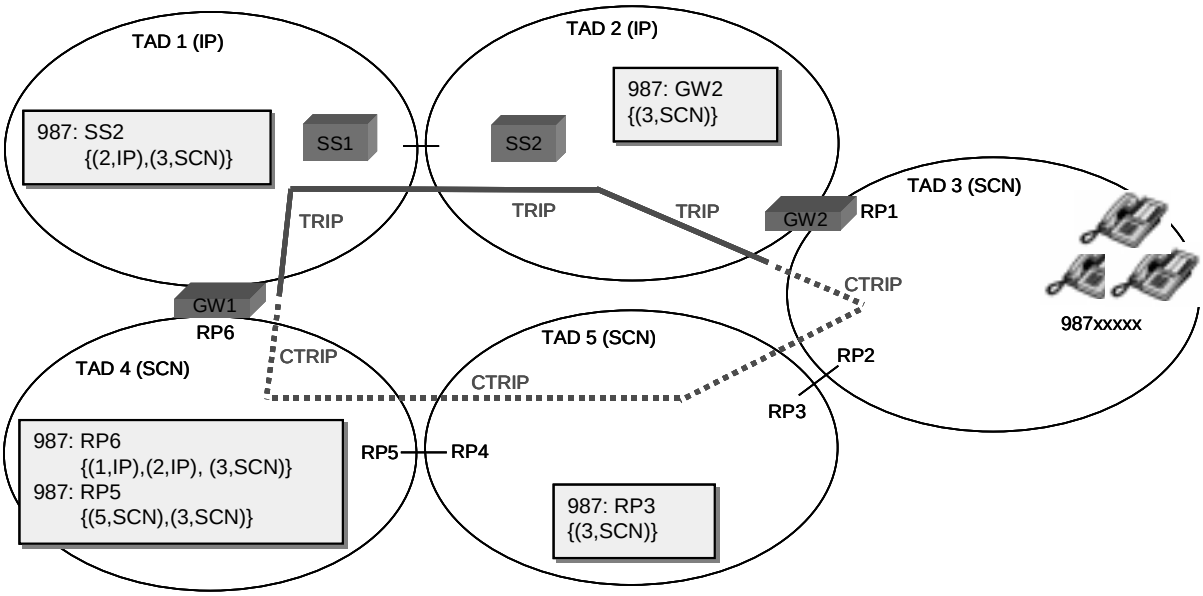

Figure 10. Protocol connections

The operator with telephony administrative domain (TAD) number 3 has subscribers whose numbers begin with the digits “987”. The operator advertises to all neighboring networks that these numbers exist on his network. Therefore, a CTRIP advertisement is made for the prefix 987. One of the neighboring networks is TAD 5, which receives the advertisement and updates it into its numbering database. Since the advertisement arrives from a network, with which the operator has a peer relationship for exchanging routing information, the routing prefix used as next hop address is known. As defined in the TRIP specification, the originating TAD puts its own identifier as the path before sending the advertisement [Rosenberg 2000a]. In our solution, also the network type of each hop in the path is added. Thereafter, the advertisement is propagated to the neighbor of TAD 5, which has the TAD identifier 4. Also this time the TAD identifier and network type is added to the path before sending it to the neighbor. The routing prefix for domain TAD 5 is set as the next hop address. Likewise, the advertisement is sent to the neighbors of TAD 4 as well, but we will not consider it in this scenario.

Another path is formed as the advertisement propagates from originating domain TAD 3 to the neighboring TAD 2. Since TAD 2 is an IP telephony network, it uses TRIP. Therefore, the CTRIP protocol used by the SCN-based neighbor must be converted to TRIP. In the conversion process, the routing prefix advertised by CTRIP is translated into an IP address of the gateway. This is performed at the border between TRIP and CTRIP before the advertisement arrives to a TRIP node of TAD 2. The conversion is presented in the next section. The gateway address is the next hop address in the advertisement arriving to TAD 2. The path consists of the TAD identifier and network type of the originating network. As TAD 2 advertises the route, the IP address of signaling server of the domain is given as the next hop address. In this case, it is the address of SS2. The identifiers of TAD 2 are added to the path, and the path of the advertisement arriving to TAD 1 contains two elements.

TAD 1 advertises reachability to prefix 987 to its neighbors, among others to TAD 4. The advertisement contains the address of the signaling server of TAD 1, called SS1. Again, the advertisement crosses the technology border and the TRIP protocol is converted to CTRIP. In the conversion the signaling server address is replaced by the routing prefix used for the gateway. The advertisement arriving to TAD 4 contains the path with TADs 1, 2 and 3. The TAD 4 already has an advertisement for this prefix, so it has to choose one of them to use. This is done according to its policies. Although only one route is chosen, all incoming advertisements are stored. If the preferred route becomes unavailable, another route can be chosen.

5.5.2 Distribution between TRIP and CTRIP

The described distribution process works much like the situation where only TRIP is used. The main difference is the conversion at the technology border between two protocols. This conversion is not only required because two protocols (CTRIP and TRIP) are used. It is also necessary to include the gateway on the path. Since routing prefixes are used on the SCN and IP addresses are used on the IP network, the conversion process performs this adaptation by making the next hop address point to the gateway. Another major deviation from the TRIP specification is that the path also contains the network type for each hop.

To transfer information between TRIP and CTRIP, an automatic system is a necessity. Updating the numbering databases separately for every change would not only be an unnecessary burden, but would also result in inconsistency in the end. The converter generates the TRIP information automatically from the CTRIP information. Thus, the TRIP information describing the numbering space of the SCN and the gateways connecting the network is generated at its source. Similarly, the CTRIP information about IP terminals is generated automatically from the TRIP information. The converter translates information between the two quite similar protocols TRIP and CTRIP while adding gateway information. We will call this entity *Numbering Gateway*.

5.5.3 Route selection and policies

A domain can receive route advertisements for the same prefix from several neighboring domains. Generally, a domain will receive one advertisement from each neighbor for a given prefix, since every domain stores the reachability information for all prefixes, and in most cases the domain forwards one selected route to all its neighbors. The domain must select one route to use within the domain and one route to advertise to the neighboring domains for each prefix. A different route can be advertised to different neighbors.

The selection of these routes is driven by policies. The policy of an ITAD is a set of rules describing how the priority of different routes is calculated. Based on the calculated priorities, the decision process of the node chooses the route for internal use and puts it into the Local TRIB. The priority of a route is calculated separately for each neighbor, and the route with the highest priority is chosen for advertisement to this neighbor. The chosen route is placed in the corresponding Adj-TRIB-Out. All nodes belonging to a given ITAD use the same policy functions. [Rosenberg 2000a]

Since CTRIP is based on TRIP, it has an identical route selection process. Received and local routes are prioritized both for local use and for advertising to different neighbors. The Local TRIB is populated with the routes with highest priority for local use. The Adj-TRIB-Out for each neighbor is populated with the selected routes for advertisement. The priority calculation of a route can be based on any combination of attributes, but it must not be dependent on the existence or the properties of other routes. [Rosenberg 2000a]

The policy is only implemented in the nodes. The conversion between TRIP and CTRIP does not perform any route selection. It only translates the protocol and adds the gateway properties. Thus, the converter will not be equipped with policy rules. No databases are required, and the converter has only one peer node on each side. The conversion is performed between two nodes, but nothing prevents an implementation from integrating this process into a node for the peers requiring it. The gateway is described in chapter 8.

The main factor in the decision process is the relationship between providers. Providers can have agreements with each other for routing telephony traffic. A provider may prefer a specific network on the path, and may as well avoid a specific network. Also the signaling gateways and media gateways themselves are owned by one of the providers that they interconnect. The RFC 2871 [RFC 2871] distinguishes three types of relationship: clearinghouses, confederations and wholesales. These are depicted in Figure 11 with M1 to M6 marking members. Routing information is exchanged between the members through peer relationships. In addition, also other structures are possible. Peer relationships in CTRIP can be assumed to be similar.

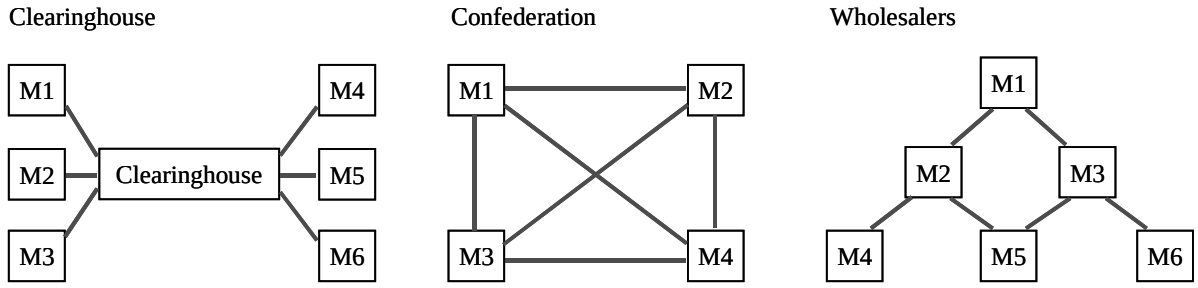

Figure 11. TRIP peer relationships

In a scenario with both SCN and IP networks, the technology used by the different networks on the path has a significant influence on the decision. Firstly, each media conversion between packet and circuit switched voice introduces delay, jitter and other voice degradation. The number of conversions must be kept low. Secondly, the cost of voice transmission over IP networks is generally lower, while quality of circuit switched voice is currently better in most cases. Thirdly, the capacity of the gateways is limited. Further, some services might require a particular network technology, for example video conferencing. Therefore, it is important to provide information about the type of each network on the path. However, it is possible to set up a voice call even without the complete path description, although it may suffer from worse quality than necessary.

5.6 Distribution between three protocols

So far, our solution has only included two protocols: TRIP and CTRIP. Since the development of ENUM seems promising, we also want to be able to interoperate with this protocol. Each protocol carries information about E.164 numbers to be used in a specific call setup situation. The information used by the two network technologies must be consistent. A change in the addressing information on one network must progress to the other network. Information is exchanged between the two network types in different situations. In the current IP telephony architecture there is no interaction between TRIP and ENUM, since they contain information about different types of numbers. CTRIP on the SCN must be able to interact with both TRIP and ENUM on the IP network.

With three protocols, there are six update cases. According to the specifications, the information carried by TRIP and ENUM is mutually exclusive. If updates between TRIP and ENUM are excluded, four update cases remain: for each protocol pair (CTRIP-TRIP, CTRIP-ENUM) in two directions. TRIP carries information about SCN destinations and ENUM stores information about IP terminals. With this separation, only two of the update cases make sense:

1. Update from CTRIP to TRIP
2. Update from ENUM to CTRIP

This relationship is depicted in Figure 12. For new destinations, information will only pass in one direction. TRIP contains information about SCN numbers reachable through gateways, so updates from CTRIP to TRIP are necessary. Updates from CTRIP to ENUM are not motivated, since ENUM is not related to SCN-terminals. In the opposite direction, CTRIP can obtain information about IP-terminals from ENUM. An update from TRIP to CTRIP would seem unnecessary, since CTRIP already contains information about SCN destinations. There would not seem to be any need to transfer information back to CTRIP that originally originated from CTRIP.

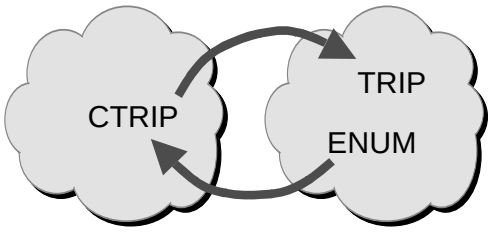

Figure 12. Interaction between CTRIP, TRIP and ENUM

Nevertheless, there are good reasons to transfer information in both directions between CTRIP and TRIP. Calls are generally possible between all SCN networks. Likewise, calls should be possible between all IP telephony networks. Thus, signaling connects all networks and there are no islands isolated from the other networks. However, not all networks support TRIP and CTRIP, especially not in the introduction stage. There may be islands in the routing information exchange even though all networks are interconnected from the signaling perspective.

In Figure 13, there is a group of SCN networks sharing routing information using CTRIP and a group of IP networks sharing TRIP information. Let us call the first one cloud A and the second one cloud B. These clouds exchange routing information through a numbering gateway. Additionally there are some SCN networks, cloud C, that share routing information with cloud B through a numbering gateway. CTRIP information is not exchanged between cloud A and cloud C. Such a situation can for example arise, when two SCN networks in different countries are using IP telephony for cheaper international calls. In this case, the information about terminals in cloud A reaches cloud B. But cloud C will not get the information if the direction from TRIP to CTRIP is not enabled.

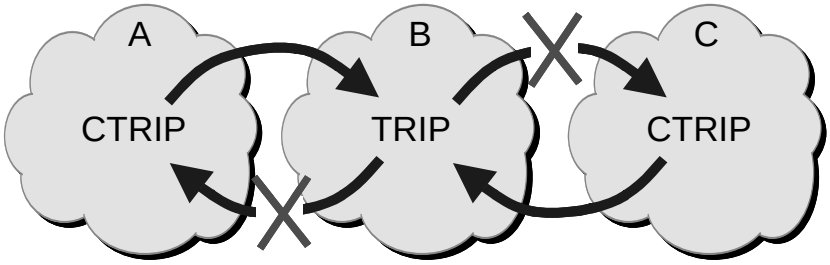

Figure 13. Islands of routing information

The solution is to convert information in both directions between CTRIP and TRIP. In the above example, cloud C then receives information about terminals in cloud A. The route to these terminals passes the IP network. To skip the IP network, routing information exchange is required between clouds A and C.

The two-way conversion also makes more routes available, since routes passing through several network types are possible. This includes long distance transport of SCN calls.

To avoid routing loops, information obtained from one domain must not be passed back to the same domain. Within the TRIP and CTRIP protocols, this is guaranteed since every domain checks that the advertisement path attribute does not contain its own domain identifier. The method can easily be extended to work between TRIP and CTRIP by passing the complete advertisement path between the protocols.

Another reason for two-way conversion is that it is possible to store information about IP terminals in TRIP. Although it is against the current plans, it makes sense in some cases. This is discussed in section 5.6.1.

Based on this premise, we design the architecture to include TRIP to CTRIP conversion as well as CTRIP to TRIP conversion. The updated interaction scheme is shown in Figure 14.

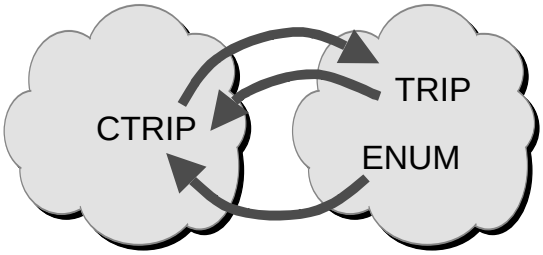

Figure 14. Protocol interaction with two-way TRIP-CTRIP conversion

5.6.1 IP terminals and ENUM

Numbering information for terminals on the IP network is stored in DNS as defined in the ENUM specification [RFC 2916]. Since ENUM is a global directory, every IP connected network has access to it. Before a call is established from an IP network, the destination number is looked up in DNS. The query can be performed by the caller's terminal or by the signaling server used by the caller. If the destination is an IP terminal, the caller can establish the call to the IP address returned by the query. If the destination is on the SCN, the query does not return any information, or it returns an URL in the TEL: format [RFC 2806]. In this case, the call is routed using TRIP information.

Since many SCN operators have IP connectivity, the same approach could be used on the SCN. The originating network could check the ENUM information to find that the destination is on the IP network. If the destination is an IP terminal, the call is routed to a gateway in an appropriate way; otherwise the call is routed using CTRIP information. This operation, which prioritizes IP destinations, can also be performed in reverse order: by first checking CTRIP and then ENUM routes the SCN destinations are prioritized.

Another approach is to transfer the ENUM information into CTRIP. This can be done at several places. A numbering gateway could convert ENUM data to CTRIP routes at the border of the domain. Alternatively, an operator with IP connectivity could create the routes using a converter inside his domain. The functionality is the same in both cases. Since the ENUM-to-CTRIP conversion is independent of the TRIP-to-CTRIP conversion, a special converting entity could perform the function. Let us call it an *ENUM-CTRIP gateway*. This is a one-way conversion, because neither CTRIP nor TRIP information can be used in ENUM, which only stores mappings for IP terminals.

One could also choose not to use ENUM at all for mapping E.164 numbers to IP addresses for IP terminals. It is possible to store routes to IP terminals in the location databases and only use TRIP. The location server to which the terminals are registered (the "home" location server) has the exact mapping from E.164 number to IP address for its registered terminals. When the numbers are announced to neighbors, the information can usually be heavily reduced by aggregation, since most numbers are consecutive and have a common prefix. Routes to the terminals are formed by TRIP in the same way as routes to SCN destinations are formed. Although this solution was abandoned by the IETF, it is still a feasible alternative.

The alternatives are:

1. ENUM is used in call setup.
2. ENUM data is transported over CTRIP.
3. ENUM is not used at all.

We need to examine the position of these alternatives in our solution.

5.6.2 Comparison of methods for storing information about IP terminals

The first of the described alternatives can be used by any operator with access to DNS information. An operator can use this solution independently of the choices of other operators. If other operators store information about IP terminals in CTRIP or TRIP, an operator can choose to use ENUM directly. It can even set the policies so that CTRIP entries for IP terminals do not enter the domain. An advantage is that no additional information needs to be transported over CTRIP and the databases can be kept smaller. The operator still has the mapping for all IP terminals available through ENUM. However, a DNS query must be performed for each call setup, the switches must have access to DNS and two separate protocols (ENUM and CTRIP) must be used.

Another major disadvantage of the first alternative is that the powerful policy-based routing process of TRIP/CTRIP is not available. In both the second and the third alternatives, routes to IP destinations can be handled like other TRIP/CTRIP routes passing through several networks, and gateways are selected in accordance with the policies.

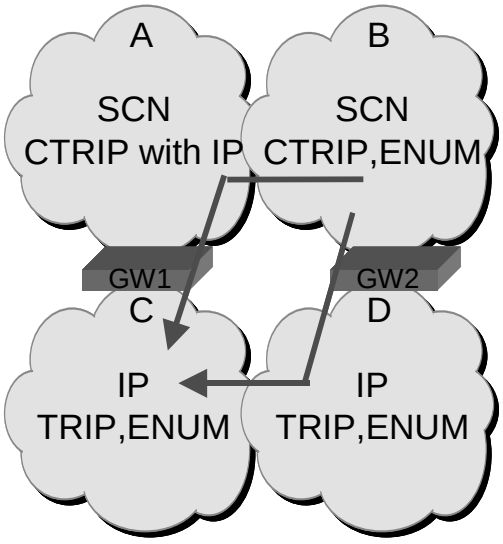

Figure 15. Routes to IP terminals

In the first alternative, an SCN to IP call must be routed to the gateway of the originating operator. In the other alternatives, a gateway of another operator could be used, depending on the policies. The Figure 15 illustrates this situation. The destination is an IP terminal residing on network C. The originating network B is connected to an IP network through a gateway. If network B does not have routes to IP terminals available on CTRIP, and uses ENUM information directly instead, it finds

out that the destination is on the IP network. It does not know any other routes, so it uses its own gateway to route the call to the IP network (through network D). On the other hand, if network B had information about routes to IP terminals on CTRIP, it would have two available routes. In addition to the previous route, it would see a route passing through network A. The second route would be useful for redundancy or as an alternative for higher quality voice transmission. In larger scenarios the advantages are even more visible.

In the second alternative, information about IP terminals is carried on CTRIP. The disadvantage of this is that more information must be carried over TRIP and CTRIP, the database sizes are larger and the overhead of processing routing information is higher. The alternative also leads to higher complexity of CTRIP, and the process of transferring information from ENUM to CTRIP must be defined.

The difference between alternative two and three is smaller. In both, routes to IP terminals are transported by CTRIP. They differ in the method how the routes are obtained. In transporting the routes over CTRIP, the solutions are similar and they share the same pros and cons. The cost of being able to build routes across both the network technologies are routing overhead, larger numbering databases and complex algorithms. If the routes are obtained from ENUM through ENUM-CTRIP gateways, an additional network element is required. The process of detecting changes in DNS is complicated since DNS is distributed between several providers, which manage a small part of the total DNS database. Storing IP terminal information in TRIP is against the planned IP telephony architecture, so at most, only a few providers might use this approach. If only some provider uses the solution, the information in ENUM must be extracted anyway. Even some type of ENUM-TRIP gateway would be possible in this solution.

The advantages and disadvantages of the alternatives are shown in Table 2.

Table 2. Comparison between alternatives for IP terminal information

	Advantages	Disadvantages
1. ENUM used in call setup	Low CTRIP overhead and small database size	Limited possible routes DNS lookup for each call setup
2. ENUM data transported over TRIP	Multiple possible routes Policy-based	Larger CTRIP overhead and database size ENUM-CTRIP gateway required
3. ENUM not used	Multiple possible routes Policy-based Only two protocols needed	Larger CTRIP overhead and database size Not in IETF architecture

Our goal is to leave the choice open and design the CTRIP protocol to allow for all alternatives. The operators can choose the alternative suitable for them. However, our focus is on the second solution, with ENUM-CTRIP gateways. The design of the actual gateway is out of the scope of the thesis. We will only specify the TRIP-CTRIP gateway, assuming that CTRIP routes can be generated from ENUM data.

5.6.3 CTRIP with routes to IP terminals

As we allow all of the above solutions, we must examine how the call setup scenario is affected with the different alternatives in Table 2. The first alternative does not affect the CTRIP routes, since ENUM is used directly without modifying CTRIP information. We need to examine the second and third alternatives only. We will use a scenario similar to the scenario in Figure 9, but with TAD 3 replaced by an IP network. The destination range exists on TAD 3. These IP terminals are registered to a server SS3 in the domain. A gateway connects TAD 3 and TAD 5.

5.6.3.1 Scenario with IP terminal information in TRIP

First, we look at a scenario where TRIP carries routes to IP destinations, corresponding to the third alternative in Table 2. We get the routes shown in Figure 16. Routes are formed according to policies in a similar manner as before. The only difference is that the last entry in the path list is an IP network. TAD 1 receives routes from two of its neighbors and chooses one to use. One is an IP-only path, and the other passes through two SCN networks. The planned solution needs no modifications to take account of this scenario.

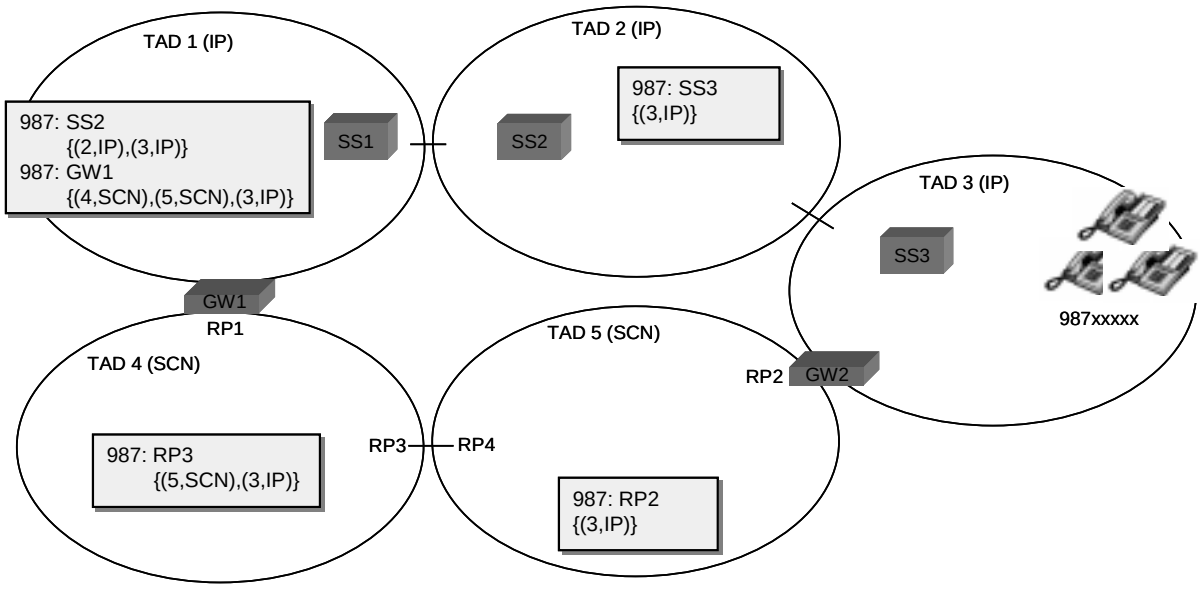

Figure 16. Routing scenario when TRIP carries information about IP destinations

5.6.3.2 Scenario with ENUM information transferred to CTRIP

Then, we look at a scenario where the destination is registered in ENUM. This corresponds to the second alternative in Table 2. The resulting routes are shown in Figure 17. TAD 4 has an ENUM-CTRIP gateway, which is configured to direct calls to IP terminals through gateway GW1 to the neighboring IP-network TAD 1.

At first sight, the routes in TAD2 and TAD3 seem erroneous, or at least inefficient. For example, the routing table of TAD 2 has a route that passes through all domains in the order 2-3-5-4-1, and then uses normal IP routing through the domains 1-2-3 to the terminal. The next-hop server is an address in TAD3 (GW2 or SS3), since TAD2 received the entry from TAD3. These routes are the result when gateway RP2 feeds routes created by the ENUM-CTRIP gateway back to the IP network. Namely in this configuration, the numbering gateways translates all routes between CTRIP and TRIP regardless of their source. Another inefficiency can be seen, since TAD 5 uses the gateway of the neighbor network instead of its own gateway, which is directly connected to the destination domain. The inefficiencies can be corrected by proper configuration and by using policy rules.

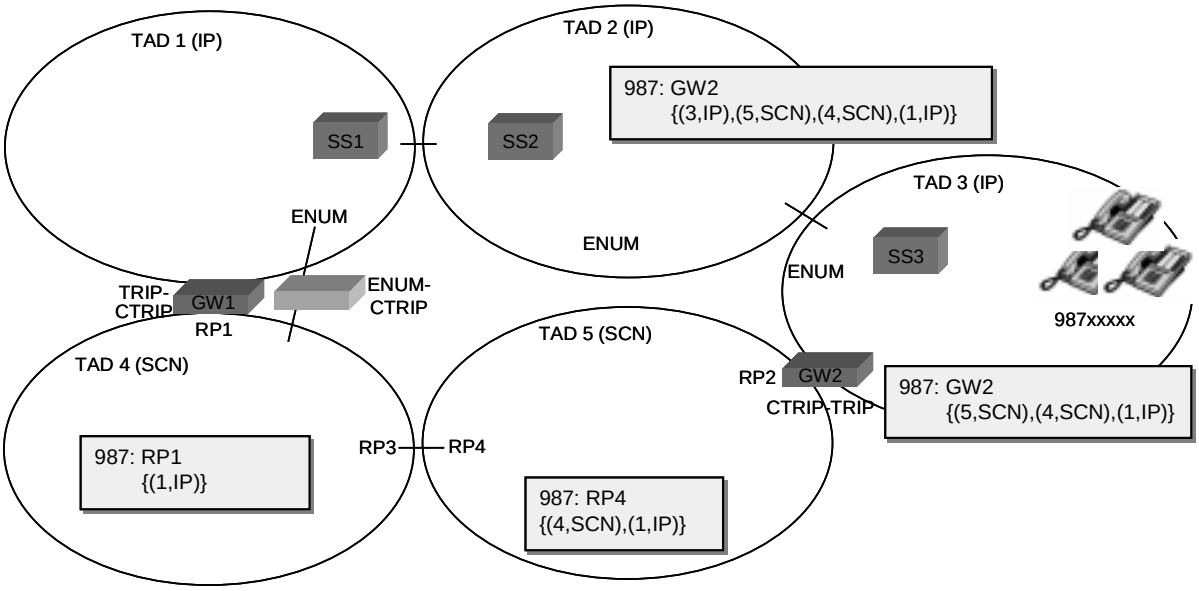

Figure 17. Routing scenario with ENUM-CTRIP gateway

Firstly, we can draw the conclusion that if an SCN domain has a gateway to the IP network, it also requires an ENUM-CTRIP gateway of its own. If TAD 5 had an ENUM-CTRIP gateway, it would route IP destined calls through gateway GW2 to the correct destination network. For terminals in the IP-based TAD 1, it has two choices: either it uses its own gateway or then it uses the gateway of TAD 4 through the route distributed by CTRIP. The selection of route is difficult, since it does not know the exact location of the end terminal. It only knows that the terminal is on the IP network. It is usually desired to use the domain's own gateway for calls to IP telephones so that most of the path consists of IP networks. However, the same gateway would be used for all IP terminals regardless of their location on the IP network, and the policies would not be very useful.

The problem is that the location of the terminal within the IP network is unknown. ENUM only tells that the terminal is on the IP network and what its IP address is. The service provider (TAD) and the geographical location are unknown. Even if it were possible to obtain the TAD identifier of the terminals from ENUM, the optimal routing would not be available. The actual path through the IP network would still be unknown. Policies could be statically configured to use a specific path for specific IP telephony domains, but the configuration must be done manually. Smart ENUM-CTRIP

gateways could learn a part of the topology of the IP network from routes distributed with for example TRIP or BGP-4. The gateway would then add the path information to the generated CTRIP entries. In this scenario, the entry generated by the ENUM-CTRIP gateway would then have the path 987: {(1, IP), (2, IP), (3, IP)} when it arrives to TAD 4. Due to its complexity the solution is however not feasible in real scenarios. Since ENUM neither distributes routes nor uses policies, routes and policies will not be available for ENUM destinations.

Secondly, the CTRIP routes distributed to the IP network as TRIP entries are far from optimal in the scenario. These are a result of that routes created from ENUM information are sent back to the IP network as TRIP routes. Inefficiency also results from the distance to the ENUM-CTRIP gateway in TAD 4 that created them; all IP destinations seen by TAD4 are marked as being on TAD1. The solution suggested in the previous paragraph would create more efficient routes, but it is not practically feasible. It is more important to notice that these TRIP routes are not used on the IP network. The DNS lookup is performed before the TRIP query, so the ENUM information is used instead. Therefore, the TRIP information for IP destination can in this case be filtered out. Alternatively, it can be used for backup routes.

Note that network TAD 1 will not accept the entry received from the TAD2 because of the loop prevention mechanism: the receiving TRIP node detects that the identifier of the domain is included on the path, and the entry is ignored. This is because the ENUM-CTRIP gateway creates entries that are similar to entries coming from TAD 1. As an alternative, the ENUM-CTRIP gateway could create entries with an empty path. These deviate from normal TRIP entries in that they contain a next hop address but the path is empty. Both approaches work, but we prefer the more standardized method to have a virtual path to the neighboring IP network.

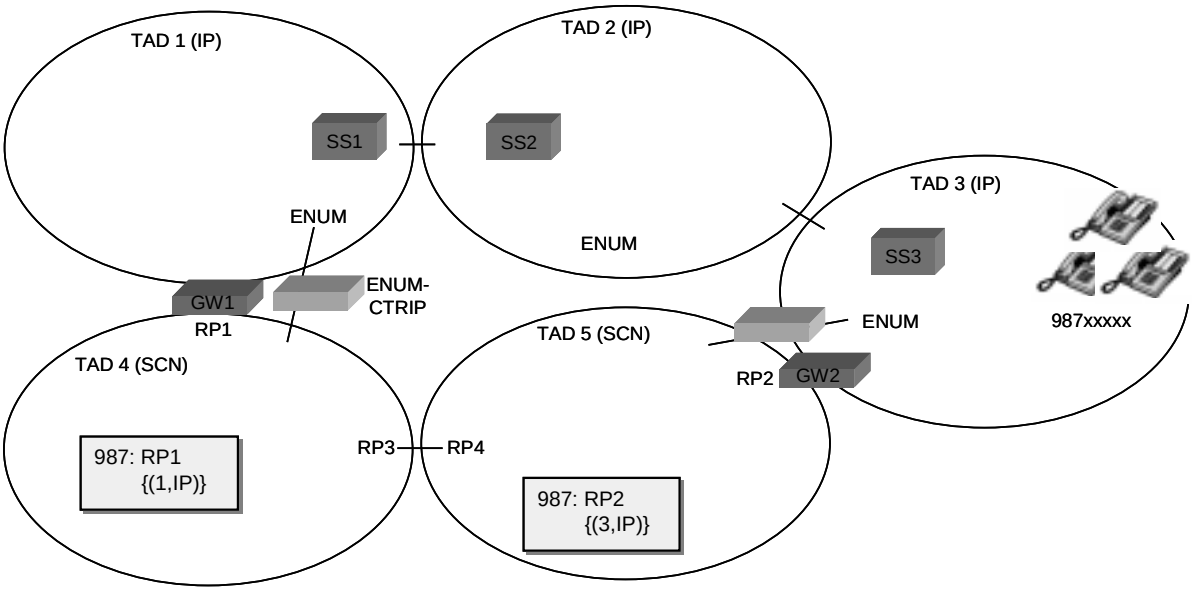

Figure 18. Improved scenario with ENUM-CTRIP gateways

An implementable scenario with the mentioned improvements is shown in Figure 18. Firstly, in this scenario every SCN domain with a gateway to an IP domain has an ENUM-CTRIP gateway. This gateway is used for IP destined calls. Naturally, also routes through other domain's gateways are received, so depending on policies also these can be used. For example if the policy function has access to topology information about the IP network, it could choose another route instead of using its own gateway. Domains without gateways can choose a route through some other domain using regular CTRIP policies. Secondly, routes originating from ENUM information are filtered out and not sent back to the IP network as TRIP entries.

5.6.3.3 Conclusions about the scenarios

As we have seen, the simplest method for storing information about IP terminals is either to use ENUM only (alternative 1) or not to use ENUM at all (alternative 3).

With proper configuration, conversion of ENUM information to CTRIP (alternative 2) is also a feasible solution. To obtain efficient routes, each domain with a gateway should also have an ENUM-CTRIP gateway. Routes to IP terminals that are reflected back to the IP network as TRIP entries are not normally useful and can be filtered out. However, if the topology contains IP- or SCN-islands that are separated from each other they are required. They also serve as backup routes. They can be left unfiltered if the database size and distribution overhead is not too large. Since the exact location of IP terminals and the IP network topology are unknown, most SCN domains will route calls to IP terminals to the nearest gateway. If additional information is available, other routes can be configured using policy rules.

We can also conclude that more efficient routes can be created if information about IP terminals is available on CTRIP (alternatives 2 and 3). Routes are then formed according to policies. Further, storing IP terminal information on TRIP allows for policies in the IP network too.

The policy knows whether a route leads to an IP terminal by examining the last hop in the path. If it is an IP domain, the entry has been obtained either from ENUM or from TRIP carrying IP terminal information. For these entries, it is in many cases important to separate between these two possible sources. A new attribute, IP destination, could be used to tell which is the source of the entry. The attribute would take the values 1 for ENUM and 2 for TRIP. The value 0 is reserved for non-IP sources. Because the attribute indicates whether the destination is an IP terminal, the path does not need to be searched to examine the network type of the last hop. Consequently, the processing work can be reduced.

5.7 Number portability scenarios

Number portability is a major application of the routing protocol based architecture. To examine the need for attributes that are specific to number portability, we look at some scenarios where a number is moved from one domain to another.

5.7.1 Number portability with TRIP and CTRIP

TRIP allows that several sources send advertisements for the same prefix. This is needed, since typically several gateways can reach the same numbers. For number portability, it is necessary to remove the previous route when a new route is advertised. In TRIP, old routes must be removed with a Withdrawn Routes attribute. No routes are automatically deleted, since routes are not regularly refreshed and timeouts are not used.

Number portability based on TRIP is implemented by removing the previous advertisement and generating a new advertisement from the new domain. The new domain sends route advertisements for the new number with the Reachable Routes attribute. These are added to the neighbors' Adj-TRIBs-In databases, and fed to the decision process. At the same time, the previous domain must send a Withdrawn route. After a transition time, when the advertisements are forwarded between the domains and the decision processes are run, the new information has been distributed to all other domains. The nodes have removed routes based on the removed advertisement and generated new routes leading to the new domain.

If the old route is removed after the new advertisement is sent, both the old and new routes will coexist for a time. The route used is during this time unpredictable and depends on the policies along the path. After the old route is removed, new routes are generated based solely on the new information. On the other hand, if the old route is removed some time before the new advertisement is sent, there will be no route to the specific destination during the transition time. Frequently there is another shorter prefix that matches the number. If this prefix is used, the call is probably routed to wrong destination. For this reason, the time between removal and generation of the advertisement must be minimized, and the old prefix must only be removed when the new prefix is operational. In the optimal case, the events would be synchronized.

Since CTRIP is similar to TRIP, the same procedure also applies to CTRIP. Moving prefixes are removed from the previous domain and advertised from the new domain. In a scenario based on only TRIP and CTRIP, number portability between the network types is easy to implement. For a number moving from the SCN to the IP network, the CTRIP entry is removed and a new TRIP entry is created. The procedure is similar for numbers moving in the opposite direction. The time between removal and creation of a new entry should be minimized in this case as well. Figure 19 illustrates the update process when a number moves from the SCN to an IP network when TRIP carries information about IP terminals.

In practice, a complete numbering range seldom moves. More often only one number out of a prefix moves. For example, the moving number 987654 belongs to a previously advertised prefix 987. Although one number moves, the rest of the numbers in the prefix remain. In this case, the advertisement for the prefix is not removed. The nodes contain one entry for the prefix 987 and one for the number 987654. The correct route is selected by normal longest prefix matching. Accordingly, the previous prefix is only removed if every number in that prefix has moved, that is, if the moved prefix and the previous prefix are identical.

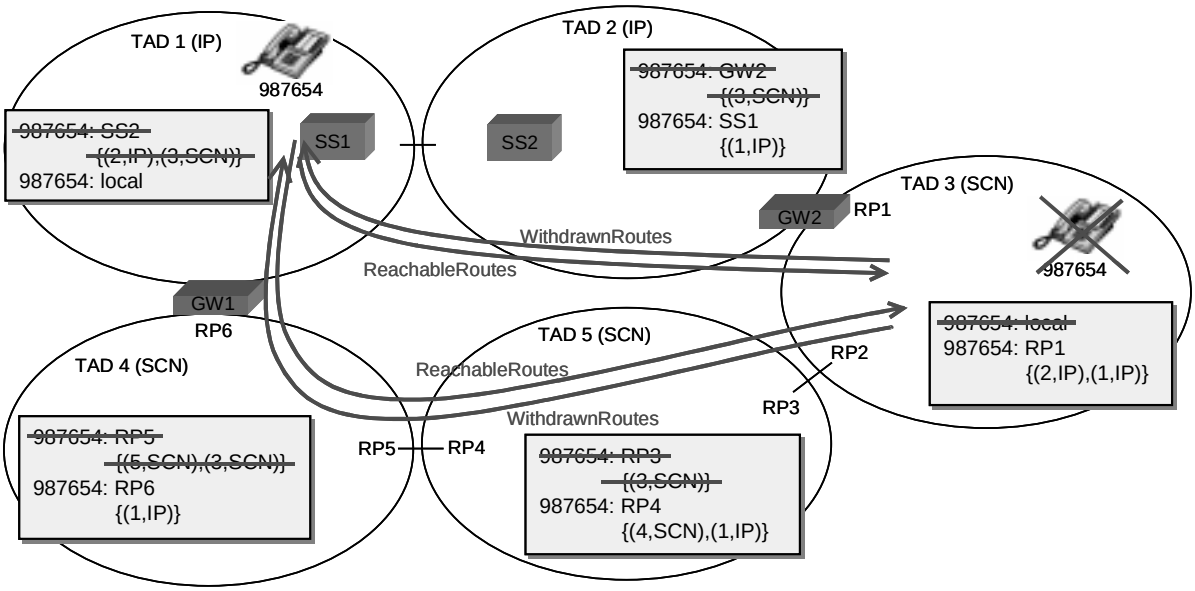

Figure 19. Number moving from SCN to the IP network

5.7.2 Signaling for number portability

No additional information is required to implement number portability if the donor provider and new provider are informed about the movement using some out-of-band method. Nevertheless, to synchronize removal and creation of entries and to automate number movement, some signaling is required. The new domain could signal to the donor domain that a new entry is created. Upon receiving this signal, the donor domain removes the previous entry. The entry should only be removed if the complete prefix is moved, thus, the donor domain keeps the previous entry if only a number out of the range is moved.

The signaling can be implemented in several ways. One approach is to send a signal from the new domain to the donor domain when the new number has been installed. In this way, the signal is sent when the new domain is ready to realize the move. The donor domain is always ready to realize the move, since it only needs to remove the previous route. If the signal were sent in the opposite direction, it would not be guaranteed that the new domain has its new entry ready when the donor domain removes its entry and sends the signal.

A straightforward way would be to send the signal with the new entry as an attribute in addition to the Reachable Routes attribute. The suggested attribute indicates that this advertisement replaces previously advertised routes. When an advertisement with the attribute is received by the donor domain, the previous route for the advertised prefixes is removed. Only an exact match is removed. The donor domain then distributes the removal by using the Withdrawn Routes attribute. This method has the advantage that the signal is synchronized with the sending of the replacing route.

The first problem with the approach is to ensure that the donor really receives the attribute. TRIP and CTRIP advertisements are neither broadcasted globally nor sent from host to host. Instead, their distribution is determined by the policies of the intermediate domains. The distribution from the donor domain is different from the distribution from the new domain, especially due to

aggregation. If the prefix is announced by only one domain and if every domain has a route for every prefix, the distribution is a spanning tree. However, during the transition the prefix is advertised by two domains. The only guaranteed way to reach the donor domain is to travel the signaling path by following the next hop addresses in each intermediate domain.

The second problem is related to security. If it is possible to remove the entry from the route originator with a single advertisement, it must be guaranteed that the new domain is authorized. It should not be possible to announce a prefix that has not been released by the previous domain.

Because of these problems, we propose a two-phase approach. In the first phase, the donor network announces the number as ready to move. In the second phase, the new network announces that the move is completed. The advertisement can be made with an attribute called *Number Portability State*. If the attribute is included in the advertisement, it can take one of the values 1 for *Ready To Move* and 2 for *Move Completed*. When the donor operator releases the number for movement, the prefix is advertised with the attribute Ready To Move. The new domain waits for receiving a route with this attribute. When the attribute is received, the new number can be installed. When the number is operational, the new domain sends the new route for the prefix with the attribute Move Completed. Advertisements with the Number Portability State attribute are not aggregated with other advertisements, which keeps them separated from normal advertisements. Later the new domain can advertise the route as a normal advertisement, which can be aggregated.

The intermediate domains will at most have two routes for the same prefix. These are in separate Adj-TRIBs-In databases. To perform the replacement, the route with the Move Completed must be given higher priority than the route with the Ready To Move attribute. The priority relative to routes without the attribute is irrelevant. A route with Ready To Move attribute will replace the previously advertised route in the intermediate nodes, since the source of the advertisement is the same. The route advertised by the new domain replaces the routes in all intermediate domains, because of the higher priority of the Move Completed attribute. Later, when the new route is advertised as a normal route, it will replace the routes with the Move Completed attributes, since the advertiser is the same. When the donor domain receives an advertisement marked as Move Completed for a route advertised as Ready To Move, it must remove the route and send a Withdrawn Routes advertisement to its peers.

Both the case when a complete advertised prefix is moved and the case when only a number from a prefix is moved can be handled in a similar way. The actual moving number is advertised with the Ready To Move and Move Completed attributes. The advertisement with the Ready To Move attribute is removed by the donor domain when the advertisement marked Move Completed is received. If the whole prefix was moved, the original advertisement was automatically removed by the Ready To Move advertisement. Signaling and database contents for the two cases are presented in Figure 20 and Figure 21.

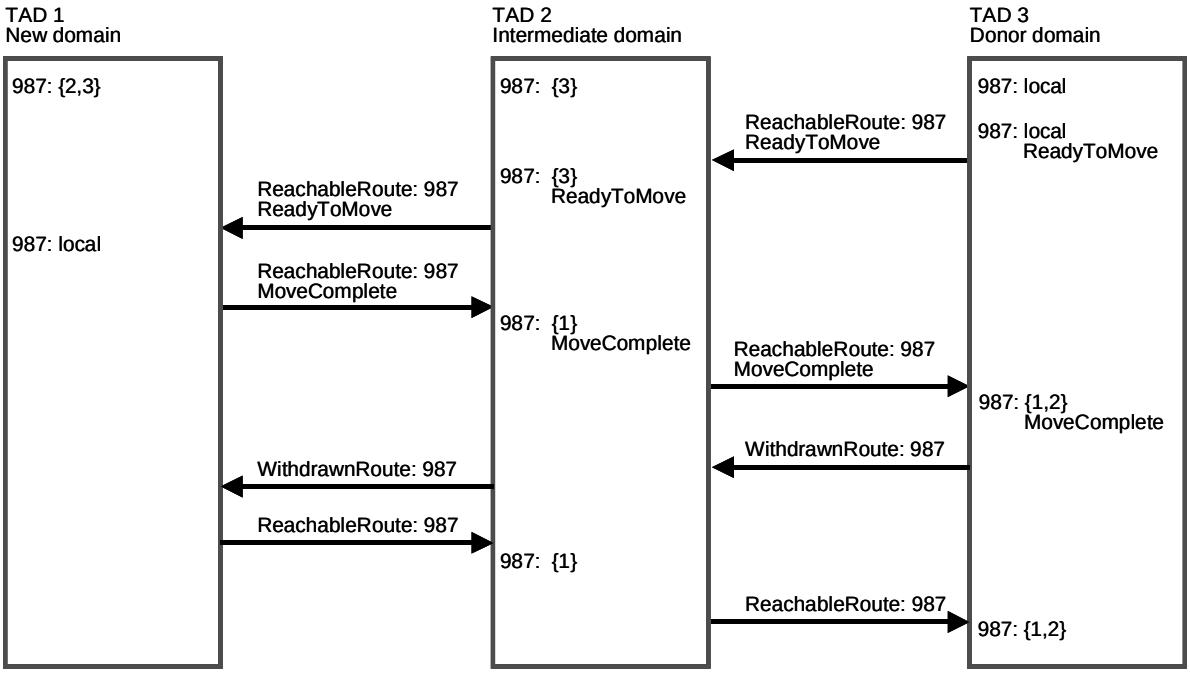

Figure 20. Moving prefix

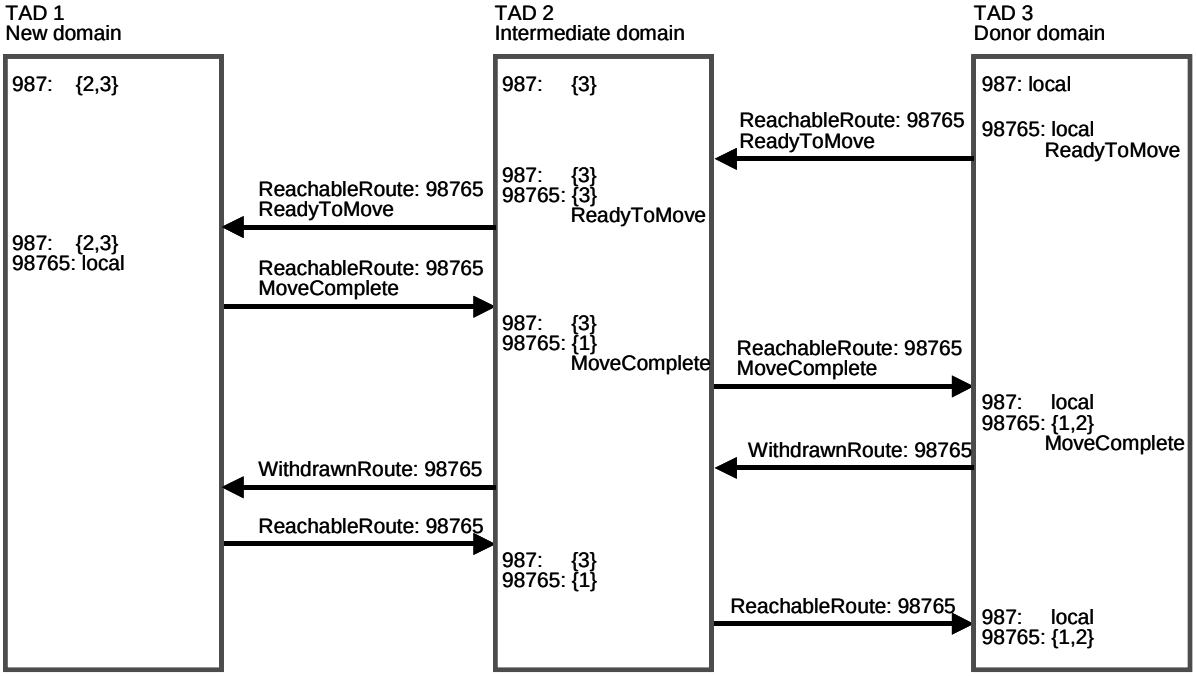

Figure 21. Moving number

Firstly, this approach prevents numbers from being advertised before they are released by the donor. The new network installs the number only when it has received the advertisement with the Ready To Move attribute. Likewise, the previous advertisement is removed only when the Move Complete is received. Although no real security is provided, it prevents mistakes and provides better synchronization. Additionally, the Ready To Move attribute can be authenticated to ensure that the number really is released.

More importantly, the advertisement from the new domain reaches all other domains and replaces routes from the donor domain. The advertisement also reaches the donor domain, which can remove the previous route with a Withdrawn Routes attribute. This is a result of using a separate attribute, which is not aggregated with normal routes and which gives the new route higher priority than the old had. It is unambiguous which route is used during the process, since the new route has higher priority.

5.7.3 Number portability with CTRIP and ENUM

Number portability between CTRIP and TRIP is a good example because of the symmetry between the protocols. As it seems, however, information about IP terminals will be stored in DNS. Therefore, we must study the cases when numbering information is moved between ENUM and CTRIP in both directions. We assume that the movement between ENUM and TRIP is a rare case, which can be ignored.

As IP telephony networks are becoming popular, it is expected that more numbers move from the SCN to the IP network than in the opposite direction. To encourage using new technology, it is beneficial to simplify movement by retaining the previous number. To implement number portability from SCN to IP in our solution, the previous CTRIP entry is removed and a new ENUM entry is created.

In the initial state, the route is advertised by CTRIP and corresponding TRIP advertisements are made by the numbering gateways. As the number is prepared for transfer, a Ready To Move attribute is issued with the route advertisement. The numbering gateways forward a corresponding TRIP advertisement. To synchronize the move, the new domain should wait for this advertisement before generating an ENUM entry. The ENUM-CTRIP gateway generates a route advertisement with the Move Complete attribute upon detecting the new ENUM entry. It must not generate route advertisements if the donor domain has not issued any Ready To Move attribute. The donor domain removes its advertisement with a Withdrawn Routes attribute.

If a number moves from an IP network to the SCN, its ENUM entry must be removed or replaced by an URL in the TEL: format [RFC 2806]. A new CTRIP entry is then created. Since ENUM stores less information about numbers, the mapping to CTRIP is difficult. ENUM entries have two states, they exist or they do not exist.

The ENUM-CTRIP gateway generates CTRIP advertisement for existing ENUM entries. A solution based on existence of ENUM information could generate an advertisement with the Ready To Move attribute when an ENUM entry is removed. The gateway generates a Withdrawn Routes advertisement for the removed entry after a specific timeout, to take into account the case when a number is completely removed. When the new domain has received the Ready To Move attribute, it can begin advertising the moved prefix. The first advertisement has a Move Complete attribute. As the ENUM-CTRIP gateway that issued the Ready To Move attribute receives the new advertisement, it must withdraw its own advertisement. An alternative is to let the ENUM-CTRIP

gateway withdraw its initial advertisement directly when the removal of the ENUM entry is detected. This alternative is simpler since no timer is needed, but it does not give any explicit indication to the new domain.

6. Architecture

In this chapter, we present the reference network architecture for the solution we are developing. We define the terminology that is used in the rest of the thesis. Architectural concepts that already have been described, such as the TAD, are briefly summarized. We describe the different elements, and how they are used to store and distribute numbering information. The architecture is fundamentally similar to that defined by TRIP.

6.1 Domains

The architecture consists of SCN networks and IP-networks owned by different operators. We use the term *network technology* to differentiate between SCN and IP technology. The term *network* refers to the network of one operator based on one network technology. Each network runs at least CTRIP or TRIP, depending on the network technology. Additionally the IP network can use ENUM.

On the IP side, the resources belonging to the same administrative authority (operator) constitute an *IP Telephony Administrative Domain* (ITAD) as defined by TRIP. The ITADs are identified by a numeric value assigned by IANA [Rosenberg 2000a]. To identify operators on the SCN, we need a similar numeric identifier in CTRIP. Let us consistently call it *Circuit Telephony Administrative Domain* (CTAD). A more general term, *Telephony Administrative Domain* (TAD), refers to both ITADs and CTADs. The ITAD identifiers and CTAD identifiers are non-overlapping. Consequently, a numeric value can either be used for an ITAD or a CTAD, but not both. If C is the set of all CTADs, I is the set of all ITADs and T is the set of all TADs, then

$T = I \cup C$

$I \cap C = \varnothing$

If an operator owns both an SCN network and an IP network, they are treated as separate networks and have different TAD identifiers.

In TRIP, each node has a 4-octet long *TRIP identifier*, which is unique within the ITAD [Rosenberg 2000a]. Similarly, we equip each CTRIP node with a 4-octet long *CTRIP identifier*, which is unique within the CTAD. The value of the identifier can be set to the IPv4 address of the node.

6.2 Reference architecture

The reference architecture is depicted in Figure 22. It consists of a number of SCN- and IP-domains sharing numbering information with TRIP and CTRIP. The main elements are the nodes, the protocol connections and the numbering gateways.

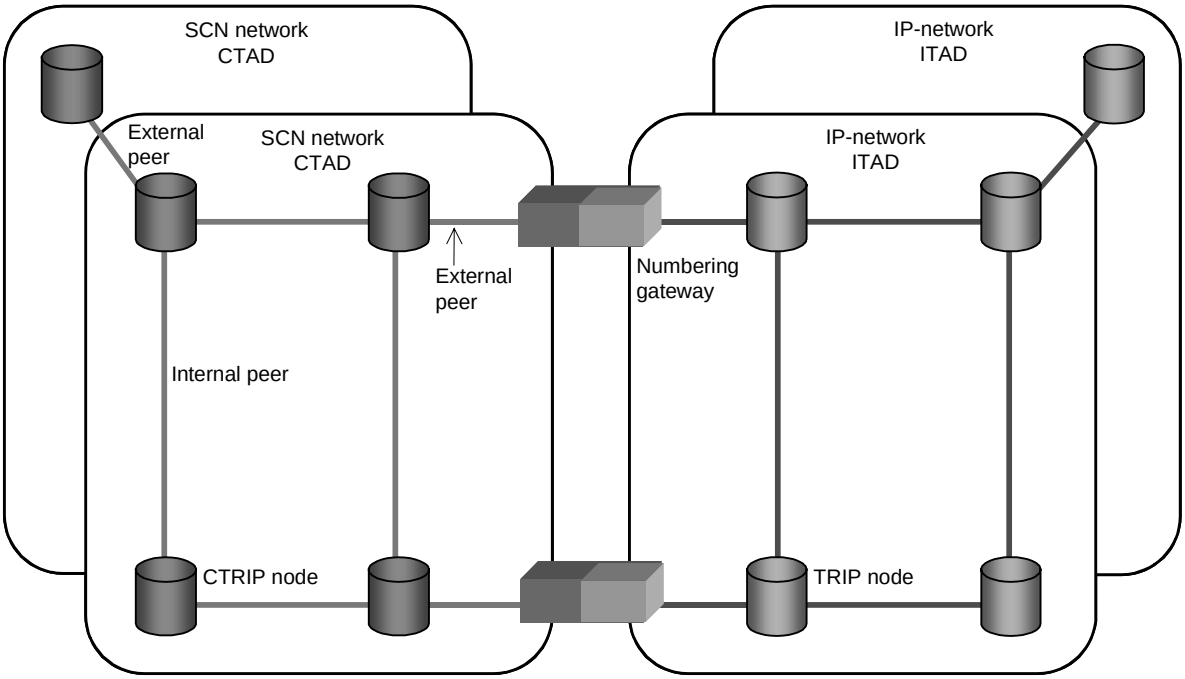

Figure 22. Reference architecture

The fundamental element in the numbering architecture is the node. A domain can contain one or several nodes. The nodes are CTRIP nodes on the SCN and TRIP nodes on the IP based network. The nodes store numbering information, which they send to and receive from other nodes. They reply to queries by mapping telephone numbers into addresses, which are used in routing.

The nodes are connected in peer-to-peer relationships, and exchange numbering information using CTRIP and TRIP. A connection between two TRIP peers uses TRIP. Correspondingly, CTRIP is used between two CTRIP peers. A numbering gateway connects a CTRIP node to a TRIP node. The numbering gateway acts like a protocol converter with one CTRIP peer and one TRIP peer.

Nodes belonging to the same domain are internal peers. Internal peers store the same routing information and use the same policies. Thus, the internal peers use TRIP or CTRIP to synchronize information, without applying policies. Peers of different domains are external peers. External peers apply policies and optionally aggregation to the received and sent information. For routing, only external distribution is of interest since the information of all internal peers is identical. The routing scenarios studied so far were based on external distribution only. Aggregation and policy rules are thus only applied at the borders of the domains.

6.3 Nodes

On the switched circuit network, we use a solution much like the IN-based number portability proposal. The numbering databases are contained in the Service Data Functions (SDF) elements. The databases are accessed by the Service Control Functions (SCF). The main exchanges are equipped with Service Switching Functions (SSF). Together these elements constitute the Intelligent Network (IN). Signaling for IN-based functions uses INAP.

In our solution, it is also possible to integrate numbering databases into the exchanges. Making the information available directly to the exchange, makes the IN query unnecessary and thereby notably reduces the load per call. With this approach, the routing on the SCN is similar to routing on IP networks. The CTRIP nodes can thus be SDFs, exchanges or any other element generating or using the numbering information.

On the IP network, numbering information is stored in the Location Servers (LS). As calls are being set up, the signaling passes through one or several Signaling Servers (SS) and the numbering information is fetched from the location servers. In this work, we will use the term signaling server also for the gatekeeper defined in the H.323 recommendation. Regardless of the different names, their functionality is practically similar. In implementations the location server functionality can be integrated with the signaling server. If they are not integrated, the signaling server uses some directory access protocol, such as LDAP, for accessing information in the location servers.

The nodes control the information about the local destinations. These are the numbers within the domain. The numbering information can be manually entered into the databases or obtained from an operation and maintenance system. On an IP telephony network, the terminals register to a registrar which provides the location server with the information.

6.4 Protocol connections

Within the domain, information is synchronized so that all nodes contain the same information. Intra-domain synchronization is only performed between nodes of the same type, i.e. TRIP or CTRIP nodes, and no gateways are used. Intra-domain synchronization uses link state mechanisms to flood updates over an arbitrary topology. The internal topology must be connected and the peer connections should be configured to provide sufficient redundancy. When a route is received from an internal peer, the routes are checked to determine if they are newer than the versions in the database. Newer routes are forwarded further to the other peers in the domain. The mechanism is based on techniques used in OSPF [RFC 1583] and SCSP [RFC 2334].

Two domains are connected to each other through a peer-to-peer relationship. One of the nodes establishes a connection for exchanging routes. CTRIP is used to transport routing information between two SCN domains and TRIP is used between two IP domains. For redundancy, peer domains can have connections between several nodes, but because of the intra-domain synchronization, the transmitted information is identical. The mechanism for inter-domain distribution is based on BGP-4 [RFC 1771].

The TRIP protocol has been summarized in section 4.5. The detailed definition can be found in [Rosenberg 2000a]. The CTRIP protocol will be defined in chapter 7.

6.5 Connections between CTRIP and TRIP

Between domains based on different network technology, a numbering gateway is required to translate between TRIP and CTRIP. Optionally a special version of TRIP, called TRIP-for-Gateways [Rosenberg 2000c], can be used for the connection between the numbering gateway and the TRIP node.

We choose to design the gateway without policies, since we want to make the operation of the gateway independent of its owner. The gateway is located on the border between two domains, and the owner of the gateway would have too large influence on the transmitted information if the policies were implemented in the gateway. The result would then be that both domains would get their own gateway. If the gateway is a well-defined element without policies, it is not relevant who is the owner of the gateway. The implementation requires that the peer nodes see the gateway as a peer of the same protocol that they use. The gateway is thus invisible to the peers.

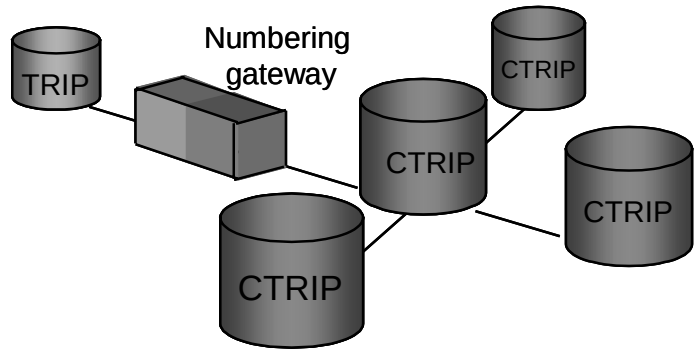

Figure 23. Separate gateway

The numbering gateway does not have to be a separate element as in Figure 23. Its functionality can be integrated into a node, whose peers can be of different type as illustrated in Figure 24. Either way, we define the numbering gateway as a separate logical entity.

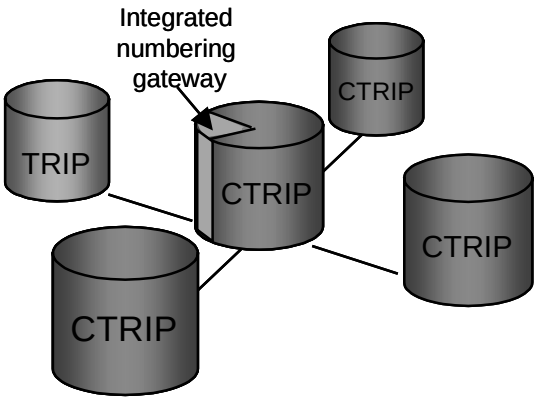

Figure 24. Node with integrated numbering gateway

6.6 Distribution between CTRIP and ENUM

To make information about IP terminals available on the SCN, an ENUM-CTRIP gateway is used. The gateway operates in only one direction; it transfers information only from ENUM to CTRIP. The gateway generates routes to terminals on the IP network. Routes should be generated for the existing entries in the ENUM directory and for all new entries that are created. Routes corresponding to removed ENUM entries must be removed.

ENUM uses the Domain Name Service (DNS) to store numbering information. DNS differs from TRIP and CTRIP in that it does not replicate the information to the nodes in the domain. There is no distribution and synchronization protocol. The problem is to know when the DNS information has changed. Normally DNS uses time-outs to remove old information, and new information is fetched the next time it is needed. A DNS database thus only caches information contained in higher level DNS directories. Because of the large volume of data, it is impossible to regularly poll the DNS servers for all IP destinations to get a complete view over all IP terminals. Instead, the ENUM-TRIP gateway could cache a part of the ENUM information by transferring it to CTRIP. Route information would be generated when it is needed, by querying DNS servers for requests to unknown numbers. The ENUM-CTRIP gateway would rely on the timeout of the DNS entries to check the validity of entries or alternatively remove all routes based on expired DNS entries. Caching of DNS entries is the normal way of operation by clients and lower hierarchy DNS servers. We will not define the exact operation of the gateway in this thesis.

7. The protocols

The architecture is mainly based on two protocols: TRIP and CTRIP. TRIP is an existing protocol defined by the IETF. The CTRIP protocol will be designed based on the TRIP protocol. Since the distribution and synchronization process are identical in the two protocols, the protocol operation needs no modification for CTRIP. Because of the tight interoperation between the protocols, the similarity is an advantage. The difference is in the transported information. Most of the attributes of TRIP can be used in CTRIP as such.

In this chapter, we examine the transported information. We define the CTRIP attributes, their usage, their format and their flags. We also define new TRIP attributes to improve the performance. Finally, we define the messages of the CTRIP protocol.

7.1 Information

In this section, we discuss the information required by the directory based routing architecture on both SCN- and IP-networks. The main goal is to determine the information required in the nodes on the SCN, and thus the information that is distributed with CTRIP. As CTRIP is based on TRIP, we must examine the TRIP attributes and examine whether the information of TRIP is adequate for our purpose. Because of the ease to add new attributes to TRIP, the CTRIP protocol can be evolved from TRIP by adding and modifying attribute definitions without changing the protocol operation.

7.1.1 Information for routing on a hybrid SCN/IP network

By studying different scenarios, we have gathered information needed for an architecture based on routing protocols between IP telephony and switched circuit networks. This information must be transmitted between operators with TRIP and CTRIP. The CTRIP protocol can be specified on the basis of the required information. In the case of TRIP we try to avoid changes to the existing protocol, although some changes would make the solution more efficient.

The most important piece of information needed for call setup is the next hop towards the destination. The next hop is the domain to which the call is set up. TRIP uses a Next Hop Server attribute, which contains the ITAD of the next hop domain and, more importantly, the address of the next signaling server or gateway. On the SCN the next hop address is the expression for generating the routing number for the given directory number. Alternatively it is an address to query with a given protocol for obtaining the routing number.

We saw that information about the technology of the networks along the path is necessary to provide efficient routes, with a minimum of media conversions. Therefore, we attach the information on network technology to every hop along the path. In this case, it is the path used for routing, called Routed Path in TRIP terminology. This path affects the preference of the route calculated by the policies. It is not crucial for successful routing, but it provides invaluable information for preventing media conversions and thereby improving voice quality. The information should be transported by both TRIP and CTRIP.

Based on previous discussion, we suggest a new attribute named IP Destination to indicate that the route leads to a prefix on the IP network, and above all, from where the route originates. The attribute takes the value 1 for an ENUM originated route and 2 for a TRIP originated route. The attribute is used by the nodes at the border between SCN and IP. It should be transported on both the SCN and the IP network, so it should be defined for both TRIP and CTRIP.

To make number movement more efficient, we propose a Number Portability State attribute. The attribute would take one of two values: Ready To Move and Move Completed. The attribute should be available on both TRIP and CTRIP.

7.1.2 Analysis of TRIP attributes

Let us review the attributes currently defined for TRIP and analyze their function in routing. The attributes are defined in [Rosenberg 2000a] and described in section 4.5.4.

The attributes in TRIP are the following:

- Withdrawn Routes
- Reachable Routes
- Next Hop Server
- Advertisement Path
- Routed Path
- Atomic Aggregate
- Local Preference
- Multi Exit Disc
- Communities
- ITAD Topology
- Converted Route

The Withdrawn Routes and Reachable Routes work like key values for the entry, determining which numbering ranges are affected. The Withdrawn Routes is exceptional in that it is not a part of the route and never stored in the routing databases. It is only used to indicate that the previously advertised routes have become unavailable. In addition to the prefix, also an address family and an application protocol are given in the Reachable and Withdrawn Routes attributes. The address

family allows decimal routing numbers and pentadecimal routing numbers to be used in addition to E.164 numbers. The application protocol is H.323 in one of its three setup models or SIP. Together the prefix, the address family and the application protocol form the key.

We can group the rest of the attributes into two groups: attributes required for the protocol operation and attributes carrying routing information. The protocol operation depends on the Advertisement Path for loop detection, Local Preference for transmitting the preference value within the domain and ITAD topology for being aware of other nodes in the same ITAD. The attributes carrying routing information are: Next Hop Server, Routed Path, Atomic Aggregate, Multi Exit Disc, Communities and Converted Route. Most of the optional attributes are likely to contain routing information and properties used for route selection. An exception is the Authentication attribute, which is used to authenticate other attributes. The Authentication attribute is defined as an extension in a separate draft [Rosenberg 2000b].

The Next Hop Server indicates the next signaling server or gateway on the path to the destination. It is responsible for routing signaling along the correct path. The attribute contains the address of the next hop server as an IPv4 address, IPv6 address or host name. The attribute also contains the ITAD identifier of the next hop.

The Next Hop server together with the key fields would be enough for routing calls to the correct destination. The other routing information attributes are used in the policy decision. The Routed Path and Atomic Aggregate inform the location server about the domains on the path towards the destination. The Converted Route indicates that the route contains path segments using another application protocol. The Multi Exit Disc attribute contains the relative preference of a route if several routes are available between two domains. The communities attribute allows grouping routes into communities to simplify management.

7.1.3 Comparison with IN-based number portability

Since the planned architecture builds on a similar call establishment procedure as that of IN-based number portability [THK 1996], we can expect that the same call setup information is required in CTRIP. The fields used in IN-based number portability are the following:

- The directory number of the subscriber (key field)
- The operator serving the subscriber
- The number portability routing domain, where the subscriber is located.

All the queries are based on the directory number, which therefore is required as the key field. To allow for whole ranges of numbers to be portable, a directory number prefix can be used instead of single numbers. As expected, the key field in CTRIP should be the directory number prefix, which is the key field of both TRIP and IN-based NP

In the IN-based routing model, there are fields indicating the operator and the number portability routing domain of the destination. Using our terminology, we obtain the TAD of the subscriber. If

we compare to TRIP, we see that TRIP gives the next hop network on the path, whereas IN-based NP specifies the destination network. In a situation with one or several intermediate networks, which is a likely situation in a hybrid IP-SCN scenario, it is more useful to know the next hop network. If only the destination network is known, all the networks are required to know the routes to all other networks. If the next hop network is indicated, the originating network does not need to know the route to the destination network, but only to the neighboring networks. Therefore, we use the TRIP approach instead of the IN-based NP approach, and indicate the next hop instead of the final network in our solution. In TRIP, this information is located in the Next Hop Server attribute.

In TRIP, the Next Hop Server attribute contains the IP-address of the next hop signaling server in addition to the ITAD. The corresponding address in the SCN would be the routing prefix to next operator. In the IN-query, this prefix is added to the number to reach the next operator on the path. The operator and number portability routing domain must be mapped to a routing prefix or routing number.

7.1.4 Identifiers in CTRIP

As both TRIP and CTRIP serve the same purpose, it is obvious that we can use most of the features of TRIP. We can use the TRIP specification as a framework and make modifications to make it suitable for our purpose. The first required enhancement is to form counterparts for the identifiers of TRIP.

We already defined CTAD as the administrative domain identifier corresponding to the ITAD of TRIP. The ITAD is a 16-bit identifier. It would be possible to use existing operator codes in CTRIP. However, since the CTRIP routes will spread to the IP network and vice versa, we must be able to specify administrative domains on the opposite network. Thus, we must be able to use CTAD identifiers in TRIP and ITAD identifiers in CTRIP without confusion. To implement this without modifications in the TRIP protocol and without conversion tables in the gateways, we must use non-overlapping identifiers that are 16-bit long.

A special number range can then be reserved for CTADs and another for ITADs. However, the 16-bit number space is very small considering the number of telephony operators and Internet telephony service providers, both the current and the future. To save space, we propose to combine the identifiers to a single telephony administrative domain identifier. We call the identifier the telephony administrative domain (TAD). The identifiers are allocated by a limited number of bodies. For our solution, we use the TAD also in TRIP where the protocol specification requires an ITAD.

The nodes of TRIP are location servers. In CTRIP, the nodes can be of various types, for example SDFs and exchanges. Therefore, we will use the more general term *node* where TRIP uses location servers (LS). TRIP uses a 32-bit identifier, called TRIP identifier, to separate between several location servers within an ITAD. In CTRIP the same type of identifier is required, which consistently is called CTRIP identifier.

7.1.5 Network technology and application protocol

In TRIP, every destination is located on the SCN, and the intermediate signaling servers reside on the IP network. Call direction is always from IP to SCN. Differently from TRIP, CTRIP distributes information about both IP and SCN numbers. A call may be set up to either an SCN- or an IP-terminal. To optimize routing efficiency and quality of service, we need to know the network technology of the networks along the path towards the destination. The selection of path is handled by the policy functions, which can examine the type of the intermediate networks when the route priority is calculated.

To further complicate things, at least two different signaling protocols are currently used on IP-networks. The protocols considered in this work are H.323 and SIP. Currently, the protocols divide the world of Internet telephony into two separate realms, since there is no open interconnection between them. Looking at media transport, both protocols are packet switched and use similar codecs. From signaling point of view, a network operated using a different protocol can be considered a separate network type since calls cannot be established between SIP and H.323 terminals without using a gateway. Gateways connecting the SIP and H.323 protocols must be on the call path in the same way as the gateways connecting SCN and IP technologies. Although the SIP-H.323 gateways do not necessarily degrade voice quality, these are a scarce resource and are undesired on the path. Thus, with two signaling protocols for IP telephony, we have three types of destinations: SCN, IP using H.323 signaling and IP using SIP signaling. Moreover, the TRIP specification has taken one step further. It separates between the three variants of H.323.

The application protocol field of the reachable and withdrawn routes attribute of TRIP can take the values shown in Table 3.

Table 3. TRIP application protocols [Rosenberg 2000a]

Code	Protocol
1	SIP
2	H.323 - H.225.0 - Q.931
3	H.323 - H.225.0 – RAS
4	H.323 - H.225.0 - Annex-G

The application protocol field is used to specify the IP-side signaling protocol towards the destination. It is thus the application protocol of the originating network. This is adequate for TRIP, since all routes go from IP to the SCN. The call may have different application protocols on the IP side but the signaling protocol on the SCN does not need to be considered. Neither are the application protocols on the path considered, since normally only one application protocol is used on the IP side and the call crosses the technology border only once. Routes using several application protocols are marked with the Converted Route attribute which can be detected by the policies.

In our solution, the call can take an arbitrary path. Calls can be originated and terminated on networks based on any signaling protocol, and they can pass different types of networks along the path.

Therefore, we must answer the questions:

1. How is the type of route indicated?
2. How is the type of intermediate networks indicated?

Beginning with the first question, we should decide how the application protocol field is used in CTRIP. In TRIP, this field allows to have a different route to the destination for different application protocols. A call using H.323 would be set up using a different signaling path than a SIP call. This is obvious, since a H.323 gatekeeper is not necessarily co-located with a SIP signaling server and vice versa. The protocol of the originating network, which is IP based, determines which route is used. The different routes can have different properties and take different paths.

In a similar way, the corresponding CTRIP field would have the value of a signaling protocol used on the SCN. On the SCN, different protocols are used in the access network and in the trunk network. The calling terminal may be analog, an ISDN or a mobile terminal, amongst others. In the trunk network ISUP, TUP and MAP are used. These have different versions and they can be transported over different protocols (e.g. ISUP over IP).

The SCN signaling protocol interoperate better than the IP signaling protocols, and calls can be established from any SCN terminal to any SCN destination. With existing signaling, there would be no need to separate between the signaling protocols on the SCN. Nevertheless, we want to separate between the SCN signaling protocols in order to provide detailed information for efficient routing. With this approach, there can be separate routes for the different protocols of the originating network.

At any stage it is possible to convert signaling by changing the application protocol field and adding a Converted Route attribute. Since CTRIP is hop-by-hop based, each network is the originating network for the rest of the path, and may have a different set of routes available.

Next, we need to define codes for the protocols used on the SCN. To leave space for future application protocols added to TRIP, we chose higher codes for the SCN protocols. The field is 16 bit long and the values 32768-65535 are reserved for vendor-specific applications. In our implementation, we chose to leave this area to vendor-specific applications and use values from the area controlled by IANA. By choosing codes beginning from 128, the first bit in the second octet can be used to indicate SCN.

The values of the modified application protocol field are shown in Table 4. TUP, ISUP, ISUP over IP and MAP are considered as separate signaling protocols on the SCN. A separate version field for the application protocol would be useful, but it would require changes in TRIP. Therefore, different versions are registered as separate application protocols. This is rarely required: only versions that

are to some extent incompatible need to be specified as separate protocols. TRIP and CTRIP are application layer routing protocols and the details of signaling protocol incompatibilities are left to lower layers.

There is also a general SCN entry for the cases where the signaling protocol is still unknown or undefined. These entries may be generated by gateways connected to the SCN through an access network. The signaling protocols of the access network are not considered as important for the route selection.

The query protocols (INAP, LDAP) are not used for call signaling, so they are not in the list of application protocols. Instead, queries are specified with the Next Hop Address attribute. Many unused values are left in the application protocol field, which can be used for future application protocols, used for example in mobile networks. The MAP entry is included for short messages, not as a query protocol.

Table 4. Application protocols

Code	Protocol
1	SIP
2	H.323 - H.225.0 - Q.931
3	H.323 - H.225.0 - RAS
4	H.323 - H.225.0 - Annex-G
128	SCN general
129	TUP
130	ISUP
131	MAP
132	ISUP over IP

Proceeding to the second question, we need a method to indicate the network type along the path. It is necessary to know the type for the networks that the signaling will traverse. The path in question is the Routed Path of TRIP. It is not necessary to know the type of the domains that the advertisement has passed, which is the Advertisement Path. Consequently, we only need to attach the type to each network in the Routed Path field. For indicating the network type, the values of the application protocol field (Table 4) are adequate. The information about network technology and signaling protocol are thereby combined.

7.2 CTRIP attributes

The aim of this section is to define the attributes of CTRIP. The definitions are based on the discussion in chapters 5 and 7. Because of the similarity with TRIP, we will concentrate on the differences between the protocols. The current protocol specification of TRIP [Rosenberg 2000a] is used as a base for the definition of CTRIP.

The following considerations for defining new TRIP attributes, listed in [Rosenberg 2000a], also apply to CTRIP:

- The use of the attribute
- The attribute's flags
- The attribute's syntax
- How the attribute works with route origination
- How the attribute works with route aggregation
- How the attribute works with route dissemination and the attributes scope

Additionally, each attribute has to be assigned an attribute code. The properties of the attributes are summarized in the end of this section. We study the attributes grouped according to their function.

7.2.1 The key fields

The Reachable Routes and Withdrawn Routes act like key fields for the advertisement by indicating the prefixes that the advertisement affects. These attributes can be similar to their TRIP counterparts. The similarity is a major advantage, since it simplifies the conversion between TRIP and CTRIP. The attributes have the same required and potential flags as in TRIP and the same type code. As in TRIP, neither is conditional mandatory. The format of the attribute and the address family values are the same. However, the application protocol field is extended, and it uses the values in Table 4.

7.2.2 Attributes for protocol operation

The protocol operation of CTRIP is identical to TRIP. The attributes controlling the protocol operation are Advertisement Path, Local Preference and ITAD Topology. The first of these is used for loop prevention, while the other two are used for intra-domain operation.

In TRIP, the Advertisement Path consists of a list of ITAD identifiers grouped in path segments. A path segment can be of one of two types (set, sequence), depending on the performed aggregation. The attribute is similar to the TRIP counterpart, with the exception that the ITADs are replaced by more general TAD identifiers. The similarity is necessary to prevent loops when routes are distributed between the two protocols. The flags, mandatory requirement and type code are identical as well.

The Local Preference must be included for intra-domain flooding and must not be used over inter-domain links. It is identical to the corresponding TRIP attribute. In CTRIP the internal topology is communicated with the CTAD topology attribute. It is a copy of the ITAD Topology attribute with only a new name.

7.2.3 Routing attributes

We consider the Routed Path, Atomic Aggregate and Next Hop Server as routing attributes, since they are the main attributes in route selection. Routing in the interconnected SCN-IP network is different from the TRIP case, so consequently the routing attributes are different.

7.2.3.1 Routed Path

The Routed Path attribute consists of a list of path segments. Each path segment is composed of a type value, a length value and a list of administrative domain identifiers. Depending on the type value, the domain identifier list represents either an unordered set or an ordered sequence. As earlier described, we attach the network type information to each hop on the Routed Path. We use the 16-bit long application protocol identifier with the values defined in Table 4. The identifier is added after the 16-bit long administrative domain identifier. No changes are required to handle the increased size of the attribute, since the length value tells the number of domains and not the octet length. With this addition, each hop is associated with both a TAD and a type.

Since the structure of the Routed Path attribute of CTRIP is different from the one of TRIP, we need a new attribute code. The type code is 8-bit long. Again, we leave some space for new TRIP attributes and use 128 as the value for the Routed Path attribute of CTRIP. We also rename the ITAD identifier as a more general TAD identifier. Like in TRIP, the attribute is conditional mandatory if the Reachable Routes attribute is present. The Well-known flag is required. The route origination, route selection and route dissemination rules of TRIP's Routed Path attribute are applied as such.

7.2.3.2 Atomic Aggregate

The Atomic Aggregate attribute is identical to its TRIP equivalent. The attribute must be included if a node selects the less specific route without selecting the more specific route from a set of overlapping routes. It indicates that a route may traverse domains not listed in the Routed Path.

7.2.3.3 Next Hop Server

In TRIP, the Next Hop Server attribute contains the next hop ITAD and a variable length server address. Various types of addresses can be used, including domain names, IPv4 and IPv6 addresses. A port number is optional. The format is shown in Figure 25.

```
Server  = host [':' port ]
        host    = <   A legal Internet host domain name
                   or an IPv4 address using the textual representation
                      defined in Section 2.1 of RFC 1123
                   or an IPv6 address using the textual representation
                      defined in Section 2.2 of RFC 2373. The IPv6
                      address MUST be enclosed in '[' and ']'
                      characters.>
        port    = *DIGIT

If the port is empty of not given, the default port is assumed.
```

Figure 25. Format of the server address in TRIP [Rosenberg 2000a]

Based on the discussion in section 5.3.4, we define three different types of next hop addresses for CTRIP. The type is identified by a one-octet value. The type field can take the values presented in Table 5. In the same table, also the parameters are listed. More types can be defined later as they are required.

Table 5. Next Hop Address types in CTRIP

Type value	Address type	Parameters
1	Routing number	Regular expression
2	SCN query	Regular expression, Protocol, Protocol version, DPC, SSN, Parameters
3	IP query	Regular expression, Protocol, Protocol version, Address, Parameters

In CTRIP, the routing number and the number used in the queries can be modified by a regular expression. The regular expression is similar to the substitution expression of the NAPTR record, defined in [RFC 2915]. The regular expression is a string given as a length-value pair, where length is a one-octet value.

In both the query types, a query protocol and a protocol version is given. The query protocol and the version are given as one-octet identifiers. The interpretation of the version identifier depends on the value of the protocol identifier. In this thesis we use the values given in Table 6 for the protocol, and we leave the definition of version identifiers open. Since the same protocols may be used on both the SCN and the IP networks, we use a common table of identifiers for them.

Table 6. Query protocols in the Next Hop Address attribute of CTRIP

Identifier	Query protocol
1	LDAP
2	INAP
3	MAP

The destination point code (DPC) and the subsystem number (SSN) form the address where the SCN query is sent. Since ITU defines the length of point codes as 14 bits, ANSI uses 24 bits and some countries use other lengths [Nortel 1998], the DPC field must be length-value encoded. In this

case, the length specifies the number of bits, and additional padding bits are added to the value field to fill the last octet. For the SSN a 32-bit integer is enough. The IP queries are sent to an IP address, which can be an IPv4 address, an IPv6 address or a host name. For coding this into a string, the same rules are used as in the Next Hop Server attribute of TRIP, presented in Figure 25. In the attribute, this string is given as a length-value pair, where length is two octets long as defined for TRIP.

Both queries can have additional parameters given as a string, whose interpretation is dependent on the query protocol value. The string is given as a length-value pair, where length is one octet long.

The TRIP definition of address family for the Withdrawn and Reachable Routes attributes allows for three address families: decimal routing numbers, pentadecimal routing numbers and E.164 numbers. The separation is required for aggregation and for avoiding duplicate number portability mappings. Although the same address families can be used in the Next Hop Address attribute, any indication about which address family is used is not required. The number portability mapping is already performed and the field is not aggregated.

The Next Hop Server attribute is only used within one protocol. That is, the attribute is never transferred as such from CTRIP to TRIP or vice versa. It is regenerated by the numbering gateway. There is no risk for confusion between the CTRIP and TRIP versions of the attribute, so they may use the same type code. For clarity, however, we define a new code to the CTRIP version: the value 129. For generality, we also rename the attribute to Next Hop Address instead of Next Hop Server. The flags are identical to the TRIP version. The attribute is conditional mandatory if a Reachable Routes and/or Withdrawn Routes attribute is present.

7.2.3.4 Usage

In TRIP, signaling can take a different path than the path that the advertisement has traversed. A domain can choose to add itself on the path by modifying the Routed Path and Next Hop Server attributes. A signaling server on the domain is given as the Next Hop Server and the ITAD is added to the Routed Path. The ITAD is added to the path if and only if the Next Hop Server attribute is modified. In CTRIP the corresponding operation is to give a new routing prefix in the Next Hop Address attribute, and to add the CTAD to the Routed Path.

The requirement for the regular expression in the Next Hop Address is that the receiving domain should be able to route calls to the address generated by the subscriber number and the expression. The range of possible subscriber numbers is limited by the prefix of the route.

If the next hop address is modified on each hop, it is enough to ensure that the neighboring domain can route to the generated routing number. The regular expression could in this case be negotiated in opening the connection, so that each domain tells its neighbor which expression it uses to route calls to it. However, if a hop does not add itself on the signaling path, the routing prefixes must be known globally. This requires the routing prefixes to be standardized globally to indicate a specific network in a given country. The requirement should not pose any serious problems, since global

country codes already exist, and in all countries the network operators can be identified by a routing prefix.

It is not necessary to select one of the above approaches to standardize. A route dissemination rule including both can be defined as follows. If the domain knows that it is reachable with a global routing prefix, it can choose to skip adding itself on the path. In other cases, the domain must add itself on the path with the regular expression negotiated in the open message or defined in the policy. Looking at routing efficiency, it is desirable to skip adding intermediate domains on the path to avoid signaling through more domains than required. The issue should be agreed by the parties exchanging information with CTRIP.

7.2.4 Extended routing attributes

The Converted Route and IP Destination attributes are used when several application protocols or network technologies are combined. We call them extended routing attributes.

The Converted Route attribute was recently added to the TRIP specification. The attribute indicates that an intermediate location server has altered the route by changing the route's application protocol. It is mandatory to include the attribute if the application protocol is changed. It has no value fields.

In a hybrid SCN-IP network, the application protocol attribute is changed at the technology border. It is changed from an IP telephony application protocol to an SCN protocol and vice versa. Consequently, the attribute must be added to all routes passing through a numbering gateway. The CTRIP version is identical to the corresponding TRIP attribute.

As described in section 5.6.3, an attribute is needed to indicate from where a route to an IP terminal originates. The attribute is named IP Destination. As defined in Table 7, it takes the values 1 for ENUM and 2 for TRIP. The value 0 is reserved for non-IP sources. The attribute indicates whether the destination is an IP terminal, so the path does not need to be examined, thus reducing processing work. If the attribute has the value 0 the route is not to an IP terminal. If there is no IP terminal attribute, the originating protocol is unknown.

Table 7. IP Destination values

Code	Route
0	The terminal is not on an IP network
1	The route leads to IP terminals. The route for this prefix originates from ENUM
2	The route leads to IP terminals. The route for this prefix originates from TRIP

The IP Destination attribute consists of a single code, which currently has the values 0, 1 and 2 defined. The length of the value is 8 bits. It is not conditional mandatory. It has the CTRIP flags Well-known. It has no potential flags. The type code is 130.

Routes to SCN terminals *must* be originated with an IP Destination attribute of value 0 or without IP Destination attribute. Only numbering gateways and ENUM-CTRIP gateways can originate routes with different values. A TRIP-CTRIP gateway *must* set the value of the IP Destination attribute to 2 for routes leading to a terminal on the IP network. If the route received from the TRIP protocol carries a corresponding IP Destination attribute, the same value *must* be used. All routes converted by an ENUM-CTRIP gateway *must* have an IP Destination attribute with the value 1.

The IP Destination attribute may be used in route selection. It is mainly used for filtering out routes to IP terminals before they are advertised back to the IP network. A CTRIP node may filter out routes before sending them to a TRIP node. Correspondingly, a TRIP node can filter out these routes if they are received from a CTRIP node.

Routes with different values in the IP Destination attribute should not be aggregated. Routes with missing IP Destination attribute may be aggregated with route with IP Destination value 0. On route dissemination the IP Destination attribute remains unchanged.

7.2.5 Peer relationship attributes

The Communities attribute of TRIP is also useful in CTRIP. The attribute is used to group destinations sharing some common property, so that the routing decision can be based on the identity of the group. The grouping can be used to define alliances between operators and to define a scope for aggregation as described in the TRIP specification.

To convert the TRIP Communities attribute to a CTRIP attribute, the ITAD numbers are replaced by TAD numbers. In a hybrid SCN-IP scenario, communities can span both SCN and IP networks. No other modifications are needed. The flags, type code, route origination, route selection, aggregation and route dissemination rules apply.

The Multi Exit Disc attribute can be used to specify relative preference for the routes between two domains, if several links connect the domains. In CTRIP, the attribute is an exact copy of the corresponding TRIP attribute.

7.2.6 Number portability state attribute

For number portability we suggested a new attribute called Number Portability State. The attribute can take two values: Ready To Move and Move Completed. The attribute is used for signaling the state of a moving number. The donor domain advertises the route as Ready To Move. When the new domain has installed the new route, it advertises Move Completed. The main intention is to be able to override the old route with the new route using different priorities for the attributes. By modifying the priority and preventing aggregation, the replacing route reaches the donor domain, which can remove the old route. The transition is speeded up, and only one route is defined at the same time in each node.

To ensure a large priority difference between the new and the old routes, we define the priority calculation as follows. A route with the Ready To Move attribute is assigned one fourth (¼) of the priority of the similar route without the attribute. A route with the Move Completed attribute has four times higher priority than a route without the attribute. Consequently, the new route has 16 times higher priority than the old one. The attribute codes and priority scaling factor are shown in Table 8.

Table 8. Number Portability State attribute values and relative priorities

Attribute value	Name	Priority scaling factor
No attribute	Normal state	1
1	Ready To Move	¼
2	Move Completed	4

The attribute is not conditional mandatory. It is not necessary to make the attribute well-known since the forwarding behavior can be controlled through flags. However, we make it well-known to guarantee that all nodes calculate the priority correctly. CTRIP implementations must support well-known attributes. No currently defined potential flags can be applied to a well-known attribute. The syntax of the attribute is simple: it only carries a one-octet unsigned numeric value. The code is 131.

Routes are only originated with the Number Portability attribute in two situations. Firstly, when the previous operator releases the number for moving, the released prefix is advertised with the Number Portability attribute with value Ready To Move. The advertisement has exactly the same attributes as the prefix had in previous advertisements, but with the Number Portability attribute added.

The second situation is when the new operator starts advertising the number. The advertisement then contains the Number Portability attribute with value Move Complete. This is exactly the same advertisement that will be used for further advertisement of the route, with only the Number Portability attribute added. Routes must not be advertised with the Number Portability attribute in any other situation.

When the domain, which has advertised the route as Ready To Move receives a Move Complete for the advertised prefix, the route must be removed. The route is removed by sending an advertisement with a Withdrawn Routes attribute.

Advertisements with a Number Portability attribute are always time limited. Advertisements with the Ready To Move attribute should only be active for one day, and must never be active longer than 3 days. If no Move Complete attribute for the same prefix has been received within this time, the route is removed. Also the Move Complete attribute should only be active for at most 5 days, to allow the propagation to be complete. After this time, the same route is advertised without Move Complete attribute.

In route selection the calculated preference of the routes must be scaled with the factors in Table 8. Other factors may be used, provided that they separate the preference between Ready To Move and Move Complete routes with at least a factor of 16.

Routes with different states in the Number Portability State attribute must not be aggregated. Routes with a Number Portability attribute and routes without this attribute must not be aggregated. Routes with the same code in the Number Portability State attribute may be aggregated together and the common Number Portability State attribute code is attached to the resulting routes. Because of the short lifetime, this aggregation is not recommended, however.

Routes with a Number Portability State attribute must be propagated unmodified to peers in the same and in other domains. A Number Portability State attribute must not be removed or added during the propagation of a route.

7.2.7 Summary of CTRIP attributes

A summary of the attributes defined for CTRIP is given in Table 9. The numbers in parenthesis indicate the length in octets. Variable length field are indicated with a “v”. Variable length fields that are preceded with a two-octet length value are indicated as “L+v”. Variable length fields that are preceded with a one-octet length value are indicated as “l+v”. An exception is the DPC field marked with “b+v”, where the length specifies the number of bits.

The codes and flags of the attributes are summarized in Table 10. An attribute must have the required flags when it is originated. Potential flags can be added where they are appropriate. A conditional mandatory attribute must be included in an update massage if another attribute is included in that message [Rosenberg 2000a].

Table 9. Summary of CTRIP attributes

Attribute	Fields	Short description
Withdrawn Routes	List of routes: - Address family (2) - Application protocol (2) - Address (L+v)	Routes becoming unreachable
Reachable Routes	(see Withdrawn Routes)	Routes becoming reachable
Next Hop Address	Next hop TAD (4) Type (1) If type == 1: - Routing number pattern (l+v) If type == 2: - Routing number pattern (l+v) - Query protocol (1) - Query protocol version (1) - DPC (b+v) - SSN (4) - Parameters (l+v) If type == 3: - Routing number pattern (l+v) - Query protocol (1) - Query protocol version (1) - Address (L+v) - Parameters (l+v)	The TAD and address to reach next hop. Depending on the type, the address is generated from a routing number pattern or with a query.
Advertisement Path	List of path segments: - Type (1) - List of hops: - TAD identifier (2)	Path traversed by the advertisement

Routed Path	List of path segments:
	- Type (1)
	- List of hops:
	  - TAD identifier (2)
	  - Application Protocol (2)
		Path of signaling messages

Atomic Aggregate	(no fields)	Indicates that the path can contain domains not included in Routed Path
Local Preference	Preference value (4)	Intra-domain preference
Multi Exit Disc	Preference value (4)	Inter-domain preference
Communities	List of community values: - Community TAD (2) - Community ID (2)	Communities the destination belongs to
Domain Topology	List of CTRIP identifiers (4)	Lists the internal peers of a node
Converted Route	(no fields)	Indicates that the Application Protocol has been changed
IP Destination	Code (1)	Indicates the origin of the route to a prefix on the IP network
Number Portability State	Code (1)	Indicates the state of a moving number

Table 10. Properties of CTRIP attributes

Attribute	Code	Conditional Mandatory	Required Flags	Potential Flags
Withdrawn Routes	1	False	Well-known	Link-state encapsulation (when flooding)
Reachable Routes	2	False	Well-known	Link-state encapsulation (when flooding)
Next Hop Address	129	True (if Reachable Routes and/or Withdrawn routes)	Well-known	None
Advertisement Path	4	True (if Reachable Routes and/or Withdrawn routes)	Well-known	None
Routed Path	128	True (if Reachable Routes)	Well-known	None
Atomic Aggregate	6	False	Well-known	None
Local Preference	7	False	Well-known	None
Multi Exit Disc	8	False	Well-known	None
Communities	9	False	Independent Transitive	None
Domain Topology	10	False	Well-known, Link-state encapsulation	None
Converted Route	12	False	Well-known	None
IP Destination	130	False	Well-known	None
Number Portability State	131	False	Well-known	None

7.3 New attributes in TRIP

To make interoperation between CTRIP and TRIP smoother, we must add some attributes to TRIP. The added attributes are counterparts of the attributes specific to CTRIP. The aim is to make the extended TRIP version fully compatible with the existing version. Therefore, well-known or mandatory attributes cannot be added. Existing attributes cannot be changed or replaced. In every place where ITAD is required in TRIP, a TAD identifier may be given as well.

7.3.1 Network technology and routed path

To be able to describe routes across the two network types, we need to extend the Routed Path attribute to be similar to the corresponding attribute of CTRIP. Each hop on the path must be described with an application protocol. Let us call the attribute Extended Routed Path. For generality, the ITAD identifier is replaced with a TAD identifier, which allows identification of both IP and SCN domains.

We could replace the existing Routed Path attribute with a modified version. However, this would

no longer follow the attribute definition in the specification, and the extended TRIP version would be incompatible with the original one. Instead, we must add the Routed Path as a new attribute.

If a new Routed Path attribute is added, the previous attribute is either discarded or used in parallel. To be compatible with original TRIP specification, the first is unacceptable since the attribute is well-known. The new attribute must be used in parallel with the original to track the path of a route propagating through domains using original TRIP. These domains only update the original Routed Path attribute. When the route reaches a domain that understands the Extended Routed Path attribute, the Extended Routed Path attribute must be complemented with the missing hops from the original Routed Path attribute. This delayed update causes no harm, since the attribute is not used until it reaches an extended TRIP node, which can update the attribute before it is used. No information is lost if the application protocol is kept the same within the zone of original TRIP nodes.

The Extended Routed Path is identical to CTRIP's Routed Path attribute. We choose to use the same attribute code, since the attribute can be passed between CTRIP and TRIP without conversion. In TRIP, we cannot make the attribute well-known. Instead, we make it independent transitive. An unknown independent transitive attribute may be propagated by any intermediate node. If the unknown attribute is forwarded, the partial flag is set to indicate that not all nodes on the path have processed the attribute.

The other rules for using the attribute are similar to those of CTRIP's Routed Path with some modifications. A node supporting the Extended Routed Path attribute must upon receiving an advertisement compare the Extended Routed Path and Routed Path attributes for hops only found in the latter. The Extended Routed Path must be updated to match Routed Path. For each missing hop, a corresponding hop must be added to the Extended Routed Path. Both the attributes must have the same division into path segments. The application protocol for the missing hops is obtained from the Application Protocol field of the Reachable Routes attribute.

A node supporting the Extended Routed Path attribute must update both path attributes if the Next Hop Address is modified. On aggregation, both attributes must be updated simultaneously. The Atomic Aggregate attribute is common to both the Routed Path and Extended Routed Path attribute.

7.3.2 IP destination attribute

When routes to IP destinations reach the IP network again after crossing the SCN, they can be filtered out since in most cases IP networks have direct connectivity between each others. For some reasons, they can be distributed back to the IP network over TRIP. The reason might be better quality, business relationship or redundancy. The IP destination attribute must then be transported over TRIP. Therefore, we add the CTRIP attribute as a new TRIP attribute. The code is the same to simplify conversion. The only difference is the flags, which cannot be well-known. Instead, the attribute is independent transitive.

7.3.3 Number portability state attribute

The described signaling for number portability was based on an attribute named Number Portability State. The attribute is not necessary for number portability to work, but it helps synchronizing the movement process and minimizes the transition time. To use the solution in an interconnected SCN and IP network, the attribute must also be defined in TRIP.

The definition, usage, and rules are identical to the corresponding ones for CTRIP. Since the attribute cannot be well-known in TRIP, the attribute has the independent and transitive flags. Because the attribute is not part of the TRIP specification, it is expected that only a few implementations support it. A node that does not understand the attribute will not calculate the priority correctly, and will not necessarily prefer the new route to the old. In such a case, the chain may break, and the signaling does not reach to the end. The domains before the break send signaling to the new domain and the ones after the break send to the old. There is, however, no risk of routing loops. Because of the timers, the old route will be removed anyway. However, to be able to use the advantages, all nodes between the old and the new domain of a moving number must support the attribute.

7.3.4 Summary of TRIP attributes

Table 11 presents the attributes defined in the TRIP specification [Rosenberg 2000a] complemented with the additional attributes defined in this chapter. For generality, names of identifiers such as ITAD are changed for compatibility with CTRIP. The same length indications are used as in Table 9. The codes and flags of the attributes are summarized in Table 12.

Table 11. Summary of TRIP attributes

Attribute	Fields	Short description
Withdrawn Routes	List of routes: - Address family (2) - Application protocol (2) - Address (L+v)	Routes becoming unreachable
Reachable Routes	(see Withdrawn Routes)	Routes becoming reachable
Next Hop Server	Next hop TAD (4) Server (L+v)	The TAD and server address of the next hop
Advertisement Path	List of path segments: - Type (1) - List of hops: - TAD identifier (2)	Path traversed by the advertisement
Routed Path	List of path segments: - Type (1) - List of hops: - TAD identifier (2)	Path of signaling messages
Atomic Aggregate	(no fields)	Indicates that the path can contain domains not included in Routed Path
Local Preference	Preference value (4)	Intra-domain preference
Multi Exit Disc	Preference value (4)	Inter-domain preference
Communities	List of community values: - Community TAD (2) - Community ID (2)	Communities the destination belongs to
Domain Topology	List of TRIP identifiers (4)	Lists the internal peers of a node
Converted Route	(no fields)	Indicates that the Application Protocol has been changed
Extended Routed Path	List of path segments: - Type (1) - List of hops: - TAD identifier (2) - Application Protocol (2)	Path of signaling messages containing application protocol
IP Destination	Code (1)	Indicates the origin of the route to a prefix on the IP network
Number Portability State	Code (1)	Indicates the state of a moving number

Table 12. Properties of TRIP attributes

Attribute	Code	Conditional Mandatory	Required Flags	Potential Flags
Withdrawn Routes	1	False	Well-known	Link-state encapsulation (when flooding)
Reachable Routes	2	False	Well-known	Link-state encapsulation (when flooding)
Next Hop Address	3	True (if Reachable Routes and/or Withdrawn routes)	Well-known	None
Advertisement Path	4	True (if Reachable Routes and/or Withdrawn routes)	Well-known	None
Routed Path	5	True (if Reachable Routes)	Well-known	None
Atomic Aggregate	6	False	Well-known	None
Local Preference	7	False	Well-known	None
Multi Exit Disc	8	False	Well-known	None
Communities	9	False	Not well-known Independent Transitive	None
Domain Topology	10	False	Well-known, Link-state encapsulation	None
Converted Route	12	False	Well-known	None
Extended Routed Path	128	False	Not well-known Independent Transitive	
IP Destination	130	False	Not well-known Independent Transitive	None
Number Portability State	131	False	Not well-known Independent Transitive	None

7.4 The CTRIP protocol

The operation of the CTRIP protocol is similar to TRIP. It uses the same messages: Open, Update, Notification and Keepalive. The Open and Notification messages need minor modifications to be suitable for CTRIP. Both other messages are identical to their corresponding TRIP messages.

7.4.1 Open message

The Open message is sent by both sides after the transport connection has been opened. The format of CTRIP's Open message is similar to TRIP's Open message. An implementation that adhere to the definitions in this document, has the value one in the version field. The My ITAD field and the

TRIP identifier are replaced by My TAD and CTRIP identifier. Otherwise these fields as well as the hold time are identical to TRIP.

It is expected, that in most cases globally available routing addresses or query addresses are used in the next hop parameter. Every other domain that will receive the advertisement must be able to set up a call with a routing number generated by this regular expression.

However, in some cases a domain might want to use a method that only is available in some domains. In section 7.2.3 we suggested that the domains should be able to negotiate the prefixes used to reach the peer domains using the open message. Although the importance of this feature is minimal, we want to define it. Generally in these cases, when domain A sets up calls to domain B it uses the routing address pattern in the advertisement sent from domain B. To generate an advertisement, domain B must know the routing prefix used by domain A to set up calls to B. If it is not globally known, domain A can tell the prefix it uses for calls to domain B in the open message. This information can be passed as an optional parameter. Let us call it Routing Address Pattern.

Optional parameters are identified by a 2-octet identifier and a variable length string for the contents. We use the identifier 128 for the Routing Address Pattern. The high value is chosen to avoid overlapping between optional parameters in TRIP's and CTRIP's Open messages. The content of the parameter is the routing address pattern in the same format as described in section 7.2.3. The use of the parameter is optional.

The optional Capability Information parameter defined in [Rosenberg 2000a] also applies to CTRIP. Currently only two capabilities are defined: Route Types Supported and Send Receive. The Route Types Supported capability lists the route types supported by the transmitting node. Route types are given as a combination of address family and application protocol. The Send Receive capability specifies the mode in which the node will operate with a particular peer.

7.4.2 Notification message

The notification messages of TRIP can be used in CTRIP without modifications. We only replace the term ITAD with TAD and the term TRIP identifier with CTRIP identifier.

Although the Routing Address Pattern parameter is optional, a node may require it if no routing address pattern has been given administratively. If a node requiring this parameter receives an Open message without a routing address pattern, it replies with a notification message. The notification message has the error code 2 for Open message errors and the subcode 128. The subcode is again chosen to allow for future TRIP subcodes without overlapping. The data field of the notification may contain a suggested routing address pattern, for example a global routing prefix. If the suggestion is verified and accepted the node can send a new open message with this routing address pattern. The error description of the error is "Routing Address Pattern required".

8. The numbering gateway

We earlier presented two types of gateways that transfer routing information between protocols. One forms a connection between TRIP and CTRIP, the other between ENUM and CTRIP. Since the second type a gateway is out of the scope of this thesis, we will concentrate on the TRIP-CTRIP gateway. This gateway, which provides bi-directional conversion of information between the two similar protocols CTRIP and TRIP, was called numbering gateway.

In this chapter, we will define the structure and operation of the gateway in detail. We define how messages and attributes are translated between CTRIP and TRIP. Finally, we describe the numbering gateway's relation to the signaling and media gateway and how the TRIP-for-Gateways protocol can be applied.

8.1 Gateway structure

A gateway can be implemented either as a back-to-back node (Figure 26) or as a protocol translator (Figure 27). A back-to-back node implementation consists of a TRIP node and a CTRIP node with a translation process in between. The nodes are equipped with full functionality and contain databases for storing routing information. A protocol translator is simpler. It converts TRIP messages into CTRIP messages and vice versa by performing simple operations on the message contents.

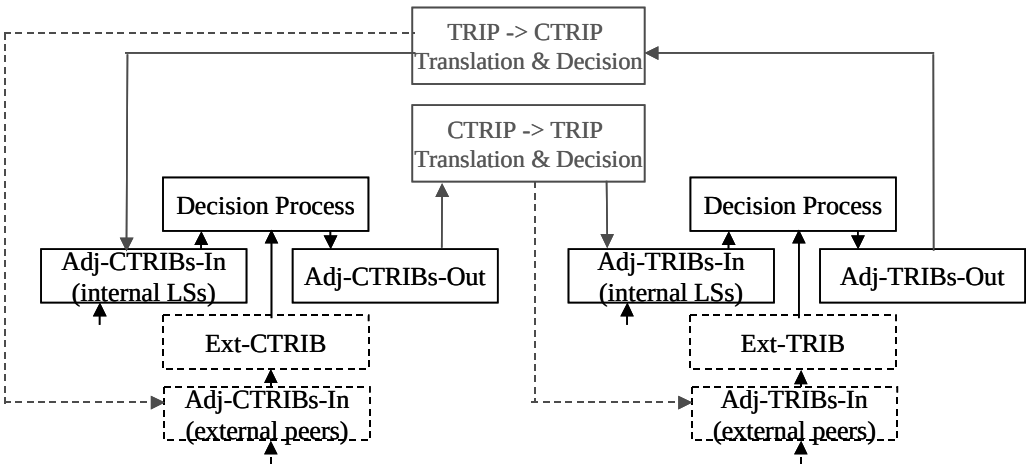

Figure 26. Back-to-back node gateway

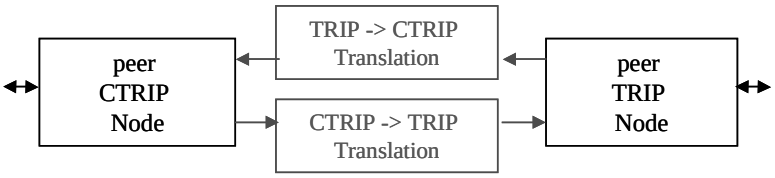

Figure 27. Protocol translator gateway

A back-to-back node is the only alternative if translation of one route requires information of some other route. The gateway must then have a database containing all routes. On the other hand, a protocol translator only operates on one entry at a time. It can be made much simpler, since it does not need any database and other functionality typical to a node. Moreover, a protocol translator is stateless, while a back-to-back node usually needs to keep state of the routes.

Because of the similarity between CTRIP and TRIP, it is possible to translate a route without information about other routes. Therefore, it is possible to use a protocol translator. It is an advantage, since the simplicity allows us to integrate the functionality into a node. With integrated numbering gateways, there is no need for separate devices, which allows us to use a large number of gateways to provide more seamless integration between TRIP and CTRIP. Moreover, the gateway can operate in two directions without doubling its resource needs.

The choice to use a protocol translator also allows us to define the operation in detail. The rules for translating attributes can be exactly defined, which meets our goal to make the numbering gateway independent of the owner. The operation is well defined and it uses no policies. Therefore, it will operate in the same way independently on who is the owner.

The numbering gateway is closely related to the signaling gateway, since it generates advertisements with the next hop address of the signaling gateway. One numbering gateway may translate the advertisements for several signaling gateways. Several advertisements are then generated, and the receiving node filters them according to its policies. Although it is outside of the standard TRIP specification, attributes describing the gateway properties can be added. Such an addition is TRIP-for-Gateways [Rosenberg 2000c]. The information for these attributes may be obtained through configuration, but more likely it will be fetched directly from the gateway. The protocol for this is currently not defined.

8.2 Gateway operation

Because of the similarity between the protocols, the translation process is simple. The main functions of the conversion are the following:

- Convert the messages.
- Convert the application protocol of the Reachable Routes and Withdrawn Routes attributes.
- Insert the gateway on the path by generating the Next Hop Address and the Next Hop Server attributes.

- Generate the Routed Path attribute on TRIP and remove it on CTRIP.
- Complete the Extended Routed Path by adding the missing hops.
- Convert attribute codes and flags.
- Add gateway properties as attributes.

In the following two sections we will define these functions.

8.3 Message translation

The numbering gateway is transparent between a TRIP and CTRIP node, that is, the nodes do not necessarily know the existence of a numbering gateway between them. The TRIP node communicates with the peer using TRIP messages, and the CTRIP node uses CTRIP messages. The Update message of CTRIP and TRIP are identical, but the attributes must be translated as described in section 8.4. The simple Keep Alive message is passed unchanged. However, both the Open and the Notification messages need conversion.

8.3.1 Open message

When the numbering gateway receives an incoming connection from one node, it opens the connection to the other node. When receiving an open message from one node, an open message is sent to the other node. Since the numbering gateway is transparent, the numbering gateway must convert the open message to a standard CTRIP and TRIP message.

The numbering gateway is visible in that it must indicate the supported TRIP and CTRIP version number to its peers. The gateway must know the well-known attributes passing through it, so it must not accept connections to peers with higher TRIP/CTRIP version number than it supports itself. That means, the CTRIP and CTRIP version number sent by the gateway indicates the supported version by the gateway and not the other node. Likewise, if the received version number is higher than the supported version, the numbering gateway must send an “Unsupported Version Number” notification message containing the highest supported version.

Because of the compatibility between TAD and ITAD identifiers and CTRIP and TRIP identifiers, these are passed unmodified. Also the hold time is passed between the nodes without modification. If needed, the gateway may request longer hold times using the “Unacceptable Hold Time” notification. However, the hold time must be the same on both sides of the gateway, since the Keep Alive messages are passed through the gateway directly.

Optional parameters and capabilities can be passed between the TRIP and CTRIP in the open message. Unknown optional parameters are rejected with the “Unsupported Optional Parameter” notification and unknown capabilities with the “Unsupported Capability” notification. The nodes can reject unknown optional parameters and capabilities without help from the numbering gateway. However, to avoid coordination between parameters and capabilities specific to only one protocol, only parameters known by the gateway to be used on both protocols should be passed through.

Others are rejected. Currently, the capabilities supported by both protocols are Route Types and Send Receive Capability.

CTRIP nodes can negotiate the routing prefix pattern in the Open message. The routing prefix pattern is unknown by TRIP, so this parameter must not be forwarded to TRIP. The gateway can participate in the negotiation if a Routing Address Pattern parameter is received from its CTRIP peer. The gateway only needs to receive this parameter. The parameter is never sent, since routing prefix are only used in the direction from the SCN to the IP network. The gateway sends the "Routing Address Pattern Required" notification if no routing address pattern has been configured and the parameter was not received.

8.3.2 Notification message

The notification messages of CTRIP and TRIP are compatible. The only difference is the "Routing Address Pattern required" notification added in CTRIP, which is handled correctly since its error code identify it as a open error. A numbering gateway participating in the routing address pattern negotiation, may use the error to choose a routing address pattern. In this case, a notification is not forwarded to the TRIP node until the connection attempt fails completely.

8.4 Attribute translation

The main function of the numbering gateway is to translate between the attributes of TRIP and CTRIP. These are transported in the Update message.

8.4.1 Reachable Routes, Withdrawn Routes and Converted Route

The application protocol of the Reachable Routes and Withdrawn Routes attributes must be changed when the advertisement passes between TRIP and CTRIP. Several application protocols, including SIP and H.323, are used on the IP network. Likewise, several application protocols are used on the SCN, including TUP, ISUP and IMAP. When an advertisement is converted, the application protocol of both sides of the signaling gateway must be known. A signaling gateway is only capable of connecting certain application protocol pairs. This list of pairs must be configured into the numbering gateway or obtained from the signaling gateway itself. If several application protocol combinations are available, a CTRIP advertisement is generated for each combination. Thus, several CTRIP advertisements can be generated for one TRIP advertisement and vice versa. This indicates that the prefix can then be reached with different application protocols.

With this approach, a single advertisement can cause several advertisements to be sent on the other side. If also the prefix is reachable through several application protocols on the receiving side, the amount of generated advertisements is even larger. This is a natural consequence of using several protocols on both sides, and the maximum number of generated advertisements equals to the number of application protocol pairs. It is up to the policy function of the receiving node to select the advertisements to use.

Every time the application protocol is changed, a Converted Route attribute must be added. Thus, every entry passing through a numbering gateway has this attribute.

8.4.2 Next Hop Server and Next Hop Address

The Next Hop Server attribute of TRIP and Next Hop Address of CTRIP have the same function. The gateway must convert between an IP address and a routing number pattern. Both are strings.

When a route is distributed from CTRIP to TRIP, the IP address of a signaling gateway must be given as the Next Hop Server. The address can be configured into the numbering gateway or obtained from the signaling gateway itself. In the case when a numbering gateway is common to several signaling gateways, several TRIP advertisements with different Next Hop Server attributes may be generated. Alternatively, one signaling gateway is selected for the given route, for example based on the routing number pattern in the corresponding CTRIP advertisement.

For a route distributed from TRIP to CTRIP, the routing address pattern of the Next Hop Address must be given. The routing number pattern is either administratively configured or obtained by negotiation in the connection opening. For multiple signaling gateways, several advertisements can be generated with different Next Hop Address or one can be chosen, potentially based on the signaling server IP address.

8.4.3 Routed Path and Extended Routed Path

The Routed Path on CTRIP corresponds to the Extended Routed Path on TRIP. They have the same format and attribute code. The flags differ between the variants: the Routed Path of CTRIP is well-known while the Extended Routed Path is independent transitive. Since the Extended Routed Path attribute is known by only a few TRIP nodes, if any, the hops of TRIP's Routed Path attribute must be inserted into the Extended Routed Path before converting the attribute to CTRIP. The Routed Path attribute is then removed. In the opposite direction, the TRIP Routed Path attribute is generated from CTRIP Routed Path attribute by removing application protocol information.

8.4.4 Other attributes

All the other attributes can be passed between TRIP and CTRIP as such. However, the flags of the IP Destination and Number Portability State attributes must be changed, since these attributes are well-known on CTRIP but independent transitive on TRIP.

8.4.5 Unrecognized attributes

Both TRIP and CTRIP are expandable protocols. New attributes can be added by defining a new attribute code. The attribute flags define how the attribute is handled by an implementation that does not recognize the attribute. Transitive attributes can be forwarded to other nodes by a node that does not recognize it. Other unknown attributes are not forwarded. An unknown attribute

marked as well-known generates an error notification.

The problem is to define what unrecognized attributes are forwarded between TRIP and CTRIP. There is no flag telling whether a specific unrecognized attribute should be forwarded to the other protocol. The idea of the numbering gateway is that the nodes themselves should not need to know, that they are communicating with a peer of another protocol type. Furthermore, the numbering gateway may receive a transitive attribute that is unrecognized by the node that sent it, consequently marked with the Partial flag. In some cases, an unrecognized attribute could be forwarded between the protocols (CTRIP and TRIP), and in some cases it must not be forwarded.

In CTRIP, a flag could be defined that tells whether the attribute is forwarded to a TRIP node if it is unknown. In this case, the decision must be made by the forwarding node, not by the numbering gateway. In TRIP, the introduction of a new flag is more difficult and requires a new version. In CTRIP it could be defined natively. We will leave the question for further studies. If the attributes have non-overlapping codes, unrecognized attributes from another protocol can be handled in the same way as unrecognized attributes from another node of the same protocol. Therefore, the Transitive flag is enough for defining handling of unrecognized attributes. However, in some cases a new flag might be motivated.

It is simpler to define the procedure how to handle well-known attributes defined in a newer version of CTRIP or TRIP than the numbering gateway supports. A well-known attribute must be known by the entities they pass through. Therefore, they must also be known by the numbering gateway. If the attribute is not known by the gateway, a notification message is returned, with the error code “Unrecognized Well-known Attribute”. That means that even though both nodes connected to the gateway recognize the attribute, it cannot be forwarded if the gateway does not support it. This is the correct behavior, since the gateway might be required to perform some translation on a well-known attribute.

8.5 Obtaining signaling and media gateway properties

The TRIP framework [RFC 2871] assumes that location servers generate the advertisements for the routes. It leaves the question open, how the location servers learn about the gateways. The draft protocol TRIP-for-Gateways [Rosenberg 2000c] proposes to use a subset of TRIP for exporting this information from the gateways to location servers. It defines two attributes describing gateway properties: circuit capacity and DSP capacity. The circuit capacity describes the number of free SCN circuits and DSP capacity the amount of available DSP resources. Because of their dynamic nature, these are only transmitted to the nearest location server, and not propagated with the route. The usage of the protocol is illustrated in Figure 28.

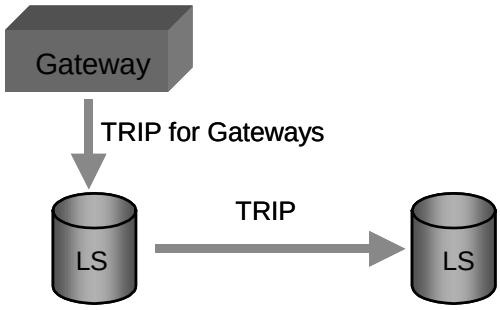

Figure 28. Usage of TRIP-for-Gateways

In our architecture, the numbering gateway originates the TRIP routes. If the route contains information about the gateway, the numbering gateway is the entity that is responsible for adding it. The gateway information may describe the availability and capacity of the gateway. The selection of signaling and media gateways to use must not be performed by the numbering gateway, since it is not allowed to use policies. Instead, the numbering gateway generates routes equipped with gateway information and the receiving node is the entity that selects the specific routes to use. The situation is illustrated in Figure 29.

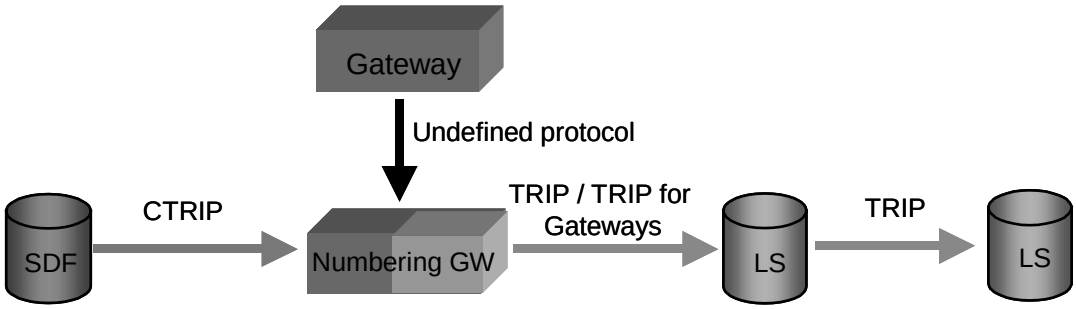

Figure 29. Adding gateway properties to routes obtained from CTRIP

The information of the originated TRIP routes is thus obtained from two directions: the route arrives from CTRIP where the destination resides and the properties of the gateway are obtained from the gateway. The protocol to use for obtaining gateway properties is in this model undefined. It can no longer be TRIP-for-Gateways, since it would be difficult to synchronize the route from CTRIP with the TRIP-for-Gateways route from the gateway. Besides, in this model we only want to obtain the status of the gateway, not actual routes as the TRIP-for-Gateways protocol describes. Therefore, a much simpler protocol is adequate for obtaining the information from the gateway. We will not specify the protocol in this thesis.

The TRIP-for-Gateways protocol is better suited for distributing the information from the numbering gateway to the peer location server. The route contains the capacity and availability properties, which are used only in the first location server. The policy of the location server selects the gateway to use using the capacity and availability information. This is exactly the purpose of TRIP-for-Gateways. However, since TRIP-for-Gateways was designed to operate in send-only mode and we need to normally send TRIP toward the SCN, we will not use the protocol as such. We only use the attributes of TRIP-for-Gateways, but the protocol is normal TRIP.

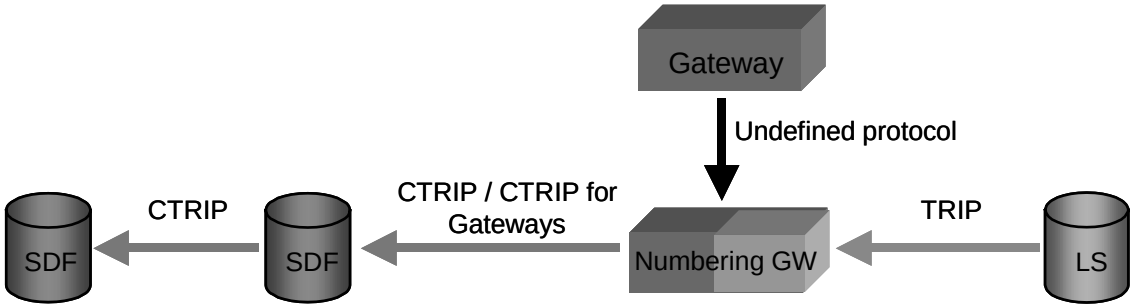

Figure 30. Adding gateway properties to routes obtained from TRIP

The situation is similar in the opposite direction, as shown in Figure 30. The same information about the gateway can be used in both directions. The corresponding protocol CTRIP-for-Gateways does not need to be specified separately. Only the attributes of the TRIP-for-Gateways protocol are added to the generated CTRIP routes. They are removed in the following hop, as stated in [Rosenberg 2000c].

9. Scenarios and applications

In this chapter, we go through some scenarios and applications of the designed architecture. First, we describe how carrier selection for long distance calls can be implemented. Then we discuss how large networks can be split into smaller areas. We explain some aspects related to the geographical scope of the information exchange. Finally, we briefly discuss some considerations for service implementation.

9.1 Carrier selection

Due to the policies, only one route is selected for each prefix. The operator has the freedom to choose the policies and can decide the transit network and the network technology of the call. The user is not able to influence on the path of the call.

In the SCN, the user has been able to choose the transit network for long distance calls using a prefix for carrier selection. In Finland, for example the prefix 101 selects Sonera's network, while the prefix 109 selects Kaukoverkko Ysi's network [Ficora].

Carrier selection can be implemented with TRIP and CTRIP if the carrier selection prefix is added in front of the directory number. The long distance network adds its carrier selection prefix to the routes before they are advertised. The receiving network then gets a route for each long distance network.

This is illustrated in Figure 31. The long distance networks TAD 2 and TAD 3 have the carrier selection prefixes 102 and 103 respectively. Before the route is advertised to the neighboring networks, the prefix of the long distance network is added. The routing table of TAD 4 contains a route for both the long distance networks. Here, the keys of the routes do not correspond to directory numbers, but rather to dialed digits.

There is also one entry for the prefix without carrier selection prefix. The long distance networks also advertise their reachability to the prefix as routes without carrier selection prefix. These routes are subject to the policy function of the receiving network. Only one route for each prefix is selected, which allows the operator to choose the route for calls made without carrier selection prefix.

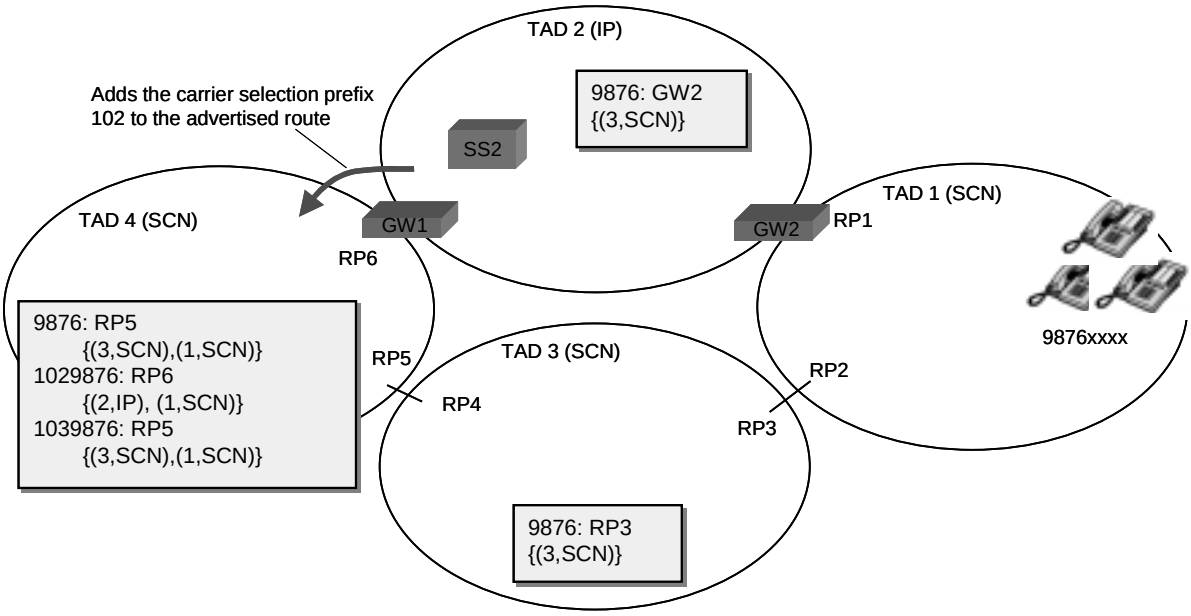

Figure 31. Scenario with carrier selection prefixes

In Finland, amongst other countries, calls dialed without carrier selection prefix are routed to a randomly chosen long distance network in proportion to the market share of the respective operator. This would seem to require several routes for the same prefix to be stored. However, it is against the objective of TRIP and CTRIP to have several routes for the same prefix: the prefix is an unambiguous key. The problem must be solved in another way. To distribute calls between carriers, a randomly chosen carrier selection prefix must be added before the route is selected by TRIP/CTRIP. If a customer has an agreement to use a specific carrier, the agreed carrier prefix is added instead of a random prefix. The procedure is illustrated in Figure 32. The carrier selection prefix is added to the dialed digits before the query to the SDF or location server.

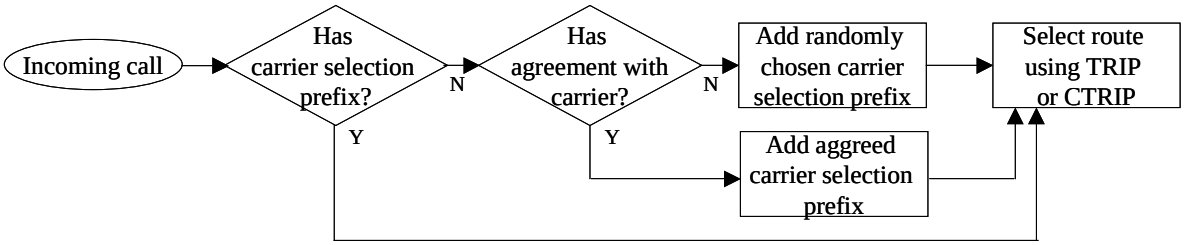

Figure 32. Procedure for adding carrier selection prefixes

9.2 Large networks

Within one domain, the information in all the databases is identical. The intra-domain synchronization ensures that every node sees the same routes. The policy functions in every node in the domain must be identical.

In large networks it may be preferred to divide the network into several areas. Each area then constitutes a separate TAD. The resulting domains contain fewer nodes, so the synchronization

becomes faster. More importantly, different policies can be defined in different domains and the administration becomes easier. Further, with domains corresponding to geographical regions, more efficient routes can be formed. Policies are applied within the network to better control the traffic. Higher preference can be given to routes with shorter path, for example by dividing the preference value by the hop count.

Another advantage of dividing the network is that aggregation can be performed on the routes distributed between the domains. Aggregation is only possible on the border of the domain. In the area containing the terminals (the home area of the terminals), the routes contain the complete mapping from subscriber number to routing address in the Next Hop Address. When the routes are distributed to the neighboring areas, the complete mapping is not required. Instead, only a routing prefix for reaching the home area is required as the Next Hop Address. The routes can be aggregated together since their attributes are similar. If the amount of moved numbers is low, very efficient aggregation can be performed.

Inter-domain distribution is a heavier operation than intra-domain synchronization. Further, the number of possible TAD identifiers is limited. Therefore, the network should not be split into too many areas. The suitable size depends on the network size, topology, the proportion of moved numbers, routing number scheme and policy requirements, amongst others.

The division into areas can be implemented without changes to the protocols. Each area has a unique TAD identifier. The TAD identifiers are 32 bits long. The last eight bits could be used to identify the area, leaving 24 bits for the operator, which should be enough. The actual allocation of TADs is however out of scope of the thesis.

The policies should be defined to give the operator's own TADs a higher preference in order to prevent calls from passing the networks of other operators where it is possible to use own resources.

9.3 Geographical scope

The purpose of TRIP is not to establish global connectivity across all ITADs. Instead, there can be many small islands of TRIP connectivity. The networks in the island have administrative relationships with each other. Still, each island can still have complete connectivity to all existing numbers. [RFC 2871]

Likewise, there may exist islands in our scenario. It is not necessary that all networks running TRIP and CTRIP would be interconnected. The solution can be used for example within one country or between a limited group of operators.

Default routes to destinations not residing within the island, can be created with short prefixes. Routes to prefixes with the length of one (i.e. 0, 1, 2, 3, …) are chosen only if there exist no more specific prefix. With these, each operator can define where calls to numbers outside the island should be routed. The next hop address can for example be a LDAP query to another database.

Also routes for prefixes corresponding to long distance prefixes and international prefixes should be defined.

In international use, the information in TRIP and CTRIP can be heavily aggregated. Number portability is not generally possible between countries; so all numbers within the country can be aggregated to one entry. The databases then contain one route to each country. On the other hand, this entry can be implemented as a static route. The selection of international network is performed by the policy, or with carrier selection as described above.

9.4 Service implementation

Some services have several points of presence. Depending on the location of the dialer, the call is directed to the nearest point that is available. With conventional IN technology, the call is directed to the routing number obtained from an IN query. For the same purpose, the Next Hop Address field of a TRIP and CTRIP route can indicate a database to query. The result of the query is the routing address of the selected point of presence.

Another way to implement this with TRIP and CTRIP is to advertise reachability to the same number from different locations. If one point is blocked, its advertisement is removed and consequently another point is selected by every node. To locate the nearest point, the policies are defined in such a way, that they give a higher preference value for a shorter length of the routed path or advertisement path. However, this solution requires the network to be divided into areas, since only one route to the number can be advertised in each area. Because of the transition time, the fact that a point is blocked is not immediately known by all domains.

Generally, services are only accessible from within one country, or within an even more limited region. To prevent routes describing services from spreading across borders, the policies must be defined to filter them out.

10. Conclusions and further work

Our main goal was to examine how a routing protocol can be utilized for distributing numbering information in an interconnected SCN and IP network. The secondary goals were to define the protocols used for distributing numbering information in the SCN and IP networks, and to define the conversion process between numbering information used in the SCN and the IP networks.

In this thesis, we have developed a solution based on TRIP, which is a distributed application layer routing protocol. The solution is built on an architecture, which includes protocols for distributing routing information, nodes for storing and managing information and gateways for converting information between protocols. The architecture covers both the SCN and the IP network.

The central part of the solution is the CTRIP protocol, which has been specified in terms of protocol operation and transported information. CTRIP is similar to TRIP with a different set of attributes and minor changes in protocol operation. For the IP network, the already existing TRIP protocol has been extended with optional attributes. The numbering gateway, which performs the conversion process between TRIP and CTRIP, has been defined in detail.

10.1 Advantages and limitations

The solution has numerous advantages. It simplifies the management of routing information and services by automating distribution, gateway selection and generation of routes. This can be considered necessary because of the increasing load due to the increasing number of services, the increasing number of service providers and number portability. With the CTRIP protocol, information about gateways and the numbers on the SCN is collected, and TRIP routes are automatically generated.

The generated routes are determined by the policy functions of the TRIP and CTRIP nodes. The same policy functions are used in all nodes within a domain. The policy can thus be defined centrally, and by modifying the policy, the operator has control over routing.

The usage of a routing protocol has the advantages that it reacts to changing conditions in the network. The protocol searches for optimal paths and the most suitable gateways. It routes around failed gateways and signaling servers. Since information about both SCN and IP numbers is available, it is possible to reduce the number of media and signaling protocol conversions. A situation where such a conversion can be avoided is when a number has moved to an IP telephony network and the originating IP network is not aware of it.

Further, the solution provides number portability across the technology border. It simplifies the transition toward IP based technology. This is necessary for moving subscribers to the IP network without causing them the inconvenience of changing telephone numbers. The solution maintains efficient routes to moving subscribers by selecting a new gateway.

The solution adds application layer routing for the telephony service. Telephone numbers can be considered as application layer addresses, compared with routing numbers and IP addresses on the network layer. Using a similar protocol on both network types has the advantage that both networks can be considered as a single entity from routing perspective. The protocols are expandable, and new attributes can be added as new services and network types are introduced.

The cost of realizing the solution is reasonable. Implementation of CTRIP in an SCN network requires new software in the SCPs, but the exchanges do not need updates. Numbering gateways are software components that can be integrated into the CTRIP nodes. CTRIP and TRIP can be deployed gradually. At an early stage, the protocols can be used between a small group of operators or between the IP- and SCN networks of a single operator. Later, if needed, the CTRIP functionality can be integrated into the exchanges, whereas the function of exchanges becomes analogous to IP routers.

A limitation is that it is not possible to start routing before most of the digits of a number have been received. However, the importance of overlap sending is expected to decrease with the increasing usage of mobile networks, IP telephony and number portability. Overlap sending in a scenario where a large fraction of the numbers is moved is not efficient. Due to number portability, routes cannot be made before getting most of the digits. New terminals with built-in directories also reduce the share of numbers that are typed in manually.

10.2 Considerations

The following additional considerations were made:

- The network should be divided into areas to utilize policies and routing capabilities within the network of an operator.
- The gateway should be simple and have well-defined behavior to make it independent on ownership and to allow integration into the nodes.
- The attributes of the TRIP-for-Gateways protocol can be used from the numbering gateway to the first node.
- TRIP and CTRIP can support several carriers by adding the carrier selection prefix to the prefixes.
- To be able to replace a route being advertised by another node, all nodes must prioritize the new route. This approach is taken in the definition of the Number Portability State attribute.

- If TRIP scales well enough, it would be preferred to use TRIP to distribute information about IP terminals instead of ENUM. This would allow using policies for routes to IP destinations, and it would simplify the suggested architecture.

10.3 Future research

The subject of using a distributed routing protocol in the SCN is not very actively researched. To our knowledge, no other group has suggested a TRIP-like protocol for the SCN. In any case, we see substantial potential advantages in using a routing protocol in the SCN, and our research done on the subject shows that the solution is technically feasible.

There is, however, still much research to do before real implementations are possible. The scenarios must be examined with topology and subscriber data from existing networks. The scalability of the solution must be tested. The actual cost and advantages must be determined numerically.

Above all, the solution must be tested and verified. We are developing an implementation of the TRIP and CTRIP protocols, as well as the numbering gateway. The implementation has reached the stage, where it is possible to test the protocol operation and simple scenarios. The results made so far seem promising and have verified that the ideas are working. However, there is still much testing and verification to be done. The implementation and the test results will be published separately.

Another topic for further work is the integration of ENUM into the solution. ENUM has been taken into considered at a conceptual level in this thesis, but the exact specification needs to be done. The definition of the ENUM-CTRIP gateway is part of that topic as well. The usage of ENUM is still evolving and the integration work should be performed when the result reaches a more stable state. We also see TRIP as a potential candidate for replacing ENUM in distributing information about IP terminals.

In this thesis, we added some attributes to TRIP as optional, since we did not want to change existing protocol. The performance could be improved by making at least the Extended Routed Path attribute mandatory in a future version of TRIP.

The definition of policy functions is left to operators. However, some testing with various policies would be necessary to examine the routes generated for different networks.

References

[Brown 2000] A.Brown, G. Vaudreuil, “ENUM Service Specific Provisioning: Principles of Operation”, IETF Internet Draft, October 2000, Work in progress, draft-ietf-enum-operation-00.txt

[CCITT X.500] CCITT, “The Directory: Overview of Concepts, Models and Service. CCITT Recommendation X.500”, 1988

[ETSI TR 101 119] European Telecommunications Standards Institute, “ETSI TR 101 119 V1.1.1 – Network Aspects (NA); High level description of number portability”, November 1997, Technical report

[ETSI TR 101 122] European Telecommunications Standards Institute, “ETSI TR 101 122 V1.1.1 – Network Aspects (NA); Numbering and addressing for number portability”, November 1997, Technical report

[ETSI TR 101 326] European Telecommunications Standards Institute, “ETSI TR 101 326 V1.1.1 – Telecommunications and Internet Protocol Harmonization Over Networks (TIPHON); The procedure for determining IP addresses for routeing packets on interconnected IP network that support public telephony”, September 2000, Technical report

[Ficora] Viestintävirasto, ”Kaukoverkot”, www-page, http://cgi.ficora.fi/numerointi/kaukoverkot_dat.htm

[Foster 2000] Mark Foster, Tom McGarry, James Yu, “Number Portability in the GSTN: An Overview”, IETF Internet Draft, March 2000, Work in progress, draft-foster-e164-gstn-np-0.txt

[Gallant 2000] A. Gallant, “The Number Portability Supplement to ITU-T Recommendation E.164”, IETF Internet Draft, July 2000, Work in progress, draft-ietf-enum-e164s2-np-00.txt

[IEEE 1993] IEEE, “IEEE Standard for Information Technology – Portable Operating System Interface (POSIX) – Part 2: Shell and Utilities (Vol. 1)”, IEEE Std 1003.2-1992, January 1993

[ITU-T E.164] International Telecommunications Union Telecommunication Standardization Sector, “The international public telecommunication numbering plan”, ITU-T Recommendation E.164, Geneva, May 1997

[ITU-T H.323] International Telecommunications Union Telecommunication Standardization Sector, Study group 16, “Packet-based multimedia

communications systems", ITU-T Recommendation H.323, February 1998

[Kantola 1997] Raimo Kantola, "A Routing Protocol Approach to Address Mapping for IP voice/ISDN Interworking", Memo, Helsinki University of Technology, May 1997

[Kantola 2000] Raimo Kantola, Jose Costa Requena, Nicklas Beijar, "Interoperable routing for IN and IP Telephony", Computer Networks, Volume 35, Issue 5, pp. 597-609, April 2001

[Kyrönaho 1999] Jukka Kyrönaho, "Interworking in an IP telephony gateway", Master's Thesis, Helsinki University of Technology, 1999

[Lind 2000] S. Lind, "ENUM Call Flows for VoIP Interworking", November 2000, Work in progress, draft-lind-enum-callflows-01.txt

[Loughney 2000] J. Loughney, "The Use of ENUM Services by Signaling Transport Protocols", IETF Internet Draft, July 2000, Work in progress, draft-loughney-enum-sigtran-00.txt

[Microsoft] http://www.microsoft.com/

[Mirabilis] http://www.icq.com/

[Nortel 1988] Nortel Networks, "Signaling Transfer Point – In SS7 Networks", November 1998, http://www.nortelnetworks.com/products/01/signaling/collateral/stp101.pdf

[RFC 1034] P. Mockapetris, "Domain Names – Concepts and Facilities", RFC 1034, November 1987

[RFC 1035] P. Mockapetris, "Domain Names – Implementation and Specification", RFC 1035, November 1987

[RFC 1058] C. Hedrick, "Routing Information Protocol", RFC 1058, June 1988

[RFC 1583] J. Moy, "OSPF Version 2", RFC 1583, 1994

[RFC 1771] Y. Rekhter, T. Li, "A Border Gateway Protocol 4 (BGP-4)", RFC 1771, March 1995

[RFC 1777] W. Yeong, T. Howes, S. Kille, "Lightweight directory access protocol", RFC 1777, March 1995

[RFC 1288] D. Zimmerman, "The finger user information protocol", RFC 1288, December 1991

[RFC 1889] H. Schulzrinne, S. Casner, R. Frederick, V. Jacobson, “RTP: A Transport Protocol for Real-Time Applications”, RFC 1889, January 1996

[RFC 2167] S. Williamson, M. Kosters, D. Blacka, J. Singh, K. Zeilstra, “Referral whois (rwhois) protocol V1.5”, RFC 2167, June 1997

[RFC 2251] M. Wahl, T. Howes, S. Kille, “Lightweight Directory Access Protocol (v3)”, RFC 2251, December 1997

[RFC 2255] T. Howes, M. Smith, “The LDAP URL Format”, RFC 2255, December 1997

[RFC 2396] T. Berners-Lee, R.T. Fielding, L. Masinter, “Uniform Resource Identifiers (URI): Generic Syntax”, RFC 2396, August 1998

[RFC 2543] M. Handley, H. Schulzrinne, E. Schooler, J. Rosenberg, “SIP: Session Initiation Protocol”, RFC 2543, 1999

[RFC 2334] J. Luciani, G. Armitage, J. Halpern, N. Doraswamy, “Server Cache Synchronization Protocol (SCSP)”, RFC 2334, April 1998

[RFC 2805] N. Greene, M. Ramalho, “Media Gateway Control Protocol Architecture and Requirements”, RFC 2805, April 2000

[RFC 2806] A. Vaha-Sipila, “URLs for Telephone Calls”, RFC 2806, April 2000

[RFC 2871] J. Rosenberg, H. Schulzrinne, “A Framework for a Gateway Location Protocol”, RFC 2871, June 2000

[RFC 2885] F. Cuervo, N. Greene, C. Huitema, A. Rayhan, B. Rosen, J. Segers, “Megaco Protocol version 0.8”, RFC 2885, August 2000

[RFC 2915] M. Mealling, R. Daniel, “The Naming Authority Pointer (NAPTR) DNS Resource Record”, RFC 2915, September 2000

[RFC 2916] P. Faltstrom, “E.164 number and DNS”, RFC 2916, September 2000

[RFC 2974] M. Handley, C. Perkins, E. Whelan, “Session Announcement Protocol”, RFC 2974, October 2000

[Rosbotham 1999] Paul Rosbotham, “WG4 FAQ”, Temporary Document, European Telecommunications Standards Institute, TIPHON

[Rosenberg 2000a] J. Rosenberg, H. Salama, M. Squire, “Telephony Routing over IP (TRIP)“, IETF Internet Draft, November 2000, Work in progress, draft-ietf-iptel-trip-04.txt

[Rosenberg 2000b] J. Rosenberg, H. Salama, “Authentication Attribute for TRIP”, IETF

Internet Draft, December 2000, Work in progress, draft-ietf-iptel-trip-authen-00.txt

[Rosenberg 2000c] J. Rosenberg, H. Salama, “Usage of TRIP in Gateways for Exporting Phone Routes”, IETF Internet Draft, July 2000, Work in progress, draft-rs-trip-gw-01.txt

[Shockey 2001] Richard Shockey, “ENUM: Phone Numbers Meet the Net”, Communications Convergence, July 2001, http://www.cconvergence.com/article/CTM20010618S0010/2

[Sugino 1999] Isao Sugino, “Addressing System for Internet Telephony”, Master’s thesis, Massachusetts Institute of Technology, June 1999

[THK 1996] Telehallintokeskus, “IN-tekniikkaan perustuva puhelinnumeron siirrettävyys”, May 1996

[THK 1999] Telehallintokeskus, ”Määräys puhelinnumeron siirrettävyydestä”, Regulation, Helsinki, February 1999

[THK 2000] Telehallintokeskus, ”Määräys yleisen valintaisen televerkon numeroinnista”, Regulation, Helsinki, December 2000

[Understanding 1997] Ericsson Telecom AB, Telia AB, Studentlitteratur AB, “Understanding Telecommunications 1”, 1997

[Understanding 1998] Ericsson Telecom AB, Telia AB, Studentlitteratur AB, “Understanding Telecommunications 2”, 1998