Data Retention in Norway
An industry perspective and proposal for efficient implementation

NRDB White Paper
Based on the European Directive on Data Retention and experiences from implementation of national initiatives such as Number Portability and Emergency Caller positioning

February 2008,
The authors
Companies behind this White Paper:

Nasjonal Referansedatabase AS, NRDB, is owned by the seven largest telecom service providers in Norway. NRDB was established in 2000 to offer services to telecom service providers on a non-profit basis. NRDB has grown from Number Portability to offer a multitude of services. The services offered are either driven by the service providers needs or by regulatory initiatives. NRDB is currently offering Norwegian Number Portability, Number lookup, DSL authorization database and emergency caller address and positioning lookup services (police / ambulance / fire emergency calls).

Systor Trondheim AS has a long experience from delivery and operations of mission critical real time database systems and has specialized in cost effective solutions for customers with high expectations and requirements with respect to availability and capacity. Systor’s customers range from private companies to large governmental and semi-governmental organizations. Systor Trondheim AS is audited with respect to the Personal Data Act in Norway.

Infostrada AS is a Norwegian consulting firm focusing on telecom, IT and content provision with in-depth knowledge in technology, products & services, and business development. Infostrada is offering senior consultants with a strong combination of operational experience and theoretical knowledge. NRDB product management and service operations are provided by Infostrada.

Authors:

Even Hovde
Infostrada AS

Even has a Degree in Telecommunication. He is managing partner at Infostrada and head of NRDB’s daily management team. Even has extensive experience from several leading telecom services providers in Norway and has worked with Number Portability from the start.

Mattias Lindstedt
Infostrada AS

Mattias has a Master Degree in Data Communication and Industrial Economics. Mattias is partner at Infostrada and a part of NRDB management team. He has a broad experience from world renowned industrial companies as well as local telecom services providers.

Lars Svenkerud
Infostrada AS

Lars has a Master Degree in Telecommunication and Computer Science, and a Degree in economics. He has 20 years experience as Technology Officer in several major international players with responsibility for development of cutting edge telecommunication solutions. Lars is senior consultant at Infostrada and serves as project manager at NRDB.

Karl Oddvar Balvold,
Systor Trondheim AS

Karl Oddvar Balvold is Senior Systems Engineer at Systor Trondheim since 2000. His main responsibility is towards the telecom sector. Karl Oddvar has experience from development, system design, project management and security management. He has been involved with NRDB since the start.
Table of contents

0. EXECUTIVE SUMMARY .................................................................................................. 4

1. BACKGROUND AND OBJECTIVE ............................................................................... 5

2. THE EUROPEAN DATA RETENTION DIRECTIVE ....................................................... 5
   2.1 THE RELEVANT REQUIREMENTS FOR THE INDUSTRY ........................................... 5
   2.2 NATIONAL ASSUMPTIONS ..................................................................................... 6

3. CHALLENGES FOR IMPLEMENTATION ....................................................................... 6
   3.1 NATIONAL CONSIDERATIONS – NORWAY ................................................................. 6
   3.2 CONTRADICTORY INTERESTS: THE INDUSTRY AND THE AUTHORITIES ................ 6
   3.3 THE INDUSTRY ........................................................................................................ 6
   3.4 THE USERS – POLICE DEPARTMENT AND OTHERS ................................................ 6
   3.5 THE REGULATING AUTHORITIES ........................................................................... 6
   3.6 OTHER IMPLICATIONS .......................................................................................... 6

4. IMPLEMENTATION PROPOSALS AND ANNUAL COST ............................................ 6

5. KEY SUCCESS FACTORS FOR IMPLEMENTATION ................................................... 6

6. CONCLUSION ................................................................................................................ 6

7. ARCHITECTURAL MODELS ........................................................................................ 6
   7.1 DISTRIBUTED SYSTEM .......................................................................................... 6
      7.1.1 Discussion ........................................................................................................ 6
   7.2 DISTRIBUTED SYSTEM WITH CENTRAL PROXY ................................................... 6
      7.2.1 Discussion ........................................................................................................ 6
   7.3 CENTRAL SYSTEM ................................................................................................... 6
      7.3.1 Discussion ........................................................................................................ 6
      7.3.2 Example configuration of a Data Retention Central ........................................... 6
   7.4 COMBINED SYSTEM .............................................................................................. 6
      7.4.1 Discussion ........................................................................................................ 6
      7.4.2 Centralized services ....................................................................................... 6
   7.5 NRDB DATA RET. CENTRAL UTILIZING NRDB RESOURCES FOR INFO LOOKUP .... 6
      7.5.1 Numbering plan and porting table ................................................................. 6
      7.5.2 Number lookup and emergency location ...................................................... 6
0. Executive summary

The European Union formally adopted the European Directive on Data Retention on March 15, 2006. This directive is addressed to the Member States to harmonise their national legal frameworks for data retention, i.e. the storage of call detail records of telephony, location information, Internet traffic, and transaction data. The users of data are the crime fighting authorities. Norwegian legislation is not yet developed, but work is in progress and the telecommunications industry will face new requirements.

The Internet and telephony service providers will have to comply with this new legislation and all providers will benefit from an efficient implementation. With this background, the Board of Nasjonal Referansedatabase AS, NRDB, has initiated this White Paper in order to identify efficient approaches to solving this challenge for the industry. NRDB is owned by seven of the largest telephony service providers and offer services to telecom service providers on a non-profit basis. Among services offered are Number Portability and Emergency Caller Address and Positioning Lookup services.

There are a number of challenges for the industry and for the individual service providers with respect to Data Retention as described in the European Directive. All service providers will have major challenges in fulfilling the directive individually, since they do not have all information that is requested (B-subscriber name, address and IMEI, for example). In order to fulfill the requirements in the directive, and also to avoid communication and interfaces to more than 200 service providers, there is a clear need for a centralised solution. A centralised solution will also allow the regulator to have superior control of data retrieval and management of the balance between crime fighting efficiency and citizens’ rights and respect of personal data.

NRDB today operates several variants of centralised solutions with interfaces to many providers. The experience is good, with respect to quality, cost efficiency and other important aspects. NRDB has evaluated several Data Retention solutions and propose a centralised approach to implementation, with local options for service providers who prefer this. The proposed structure has many similarities with the well-functioning system for Number Portability.

It is estimated that NRDB can develop and run the proposed solution for around 11-17 MNOK per year, all inclusive. With 8.9 million subscribers (telephony and Internet combined) the yearly cost of data retention will be between NOK 1,20 and NOK 1,90 per subscription.
1. **Background and objective**

This White Paper is a result of a request from the board of Nasjonal Referansedatabase AS, NRDB, the Norwegian organisation handling number portability and related issues on behalf of service providers. The background of this request is the European directive on Data Retention, and the following need for information and analysis of the consequences and proposal of efficient solutions for Norwegian service providers.

The objective of this White Paper is to identify and propose solutions for Data Retention, based on previous experiences in NRDB, solutions that can be effective and efficient both for the industry and for the authorities. The objective is also, based on experience from other regulatory initiatives in the industry, to address issues that will be crucial for a successful implementation of the Data Retention directive in Norway.

The document is written in English, considering that there are both Norwegian and International service providers and that there may be a European interest.

2. **The European Data Retention Directive**

The Directive on Data Retention was formally adopted by the EU on March 15, 2006. The Directive is addressed to the Member States to harmonise their national legal frameworks for data retention. Data retention refers to the storage of call detail records of telephony, location information, Internet traffic and transaction data. The users of data are crime fighting authorities. A national supervisory authority shall be appointed for each member state and statistics on the data use shall be sent to the European Commission each year.

2.1 **The relevant requirements for the industry**

The Directive requires national laws in the European Economic Area (EEA) i.e. EU, Iceland, Liechtenstein and Norway, to be valid for public operators of:

1. fixed telephony
2. mobile telephony
3. SMS, EMS, MMS
4. Internet access
5. Internet email
6. Internet telephony
7. other Internet communication service

The Directive requires Member States to ensure that telecommunication service providers retain necessary data from minimum 6 months to maximum 2 years. Retained data shall only be provided to competent national authorities in specific cases. Typically retained data on one or more persons are provided to a police investigation. No specific response time have been defined. The Directive states that data shall be retained in such a way that it can be transmitted upon request to the competent authorities without “undue delay”.

Only data handled or stored should be retained. No data revealing the content of the communication may be retained. Data considered are data that identify

1. communication source
2. communication destination
3. communication type
4. communication date, time and duration
5. the device used for communication
6. mobile communication equipment position

The table below indicates what kind of data are to be retained.
<table>
<thead>
<tr>
<th>Data to be retained</th>
<th>Fixed and mobile telephony</th>
<th>Internet Access, Internet e-mail and Internet telephony</th>
</tr>
</thead>
<tbody>
<tr>
<td>Communication source.</td>
<td>1. The calling telephone number</td>
<td>1. The user ID(s)</td>
</tr>
<tr>
<td></td>
<td>2. The name and address of the subscriber or registered user.</td>
<td>2. The user ID and telephone number to any communication entering the public telephone network.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. The name and address of the subscriber to whom an Internet Protocol (IP) address, user ID or telephone number was allocated at the time of the communication</td>
</tr>
<tr>
<td>Communication destination.</td>
<td>1. The telephone number(s) called, and, in cases involving supplementary services such as call forwarding or call transfer, the number or numbers to which the call is routed</td>
<td>1. The user ID or telephone number of the intended recipient(s) of an Internet telephony call</td>
</tr>
<tr>
<td></td>
<td>2. The name(s) and address(es) of the subscriber(s).</td>
<td>2. The name(s) and address(es) of the subscriber(s) or registered user(s) and user ID of the intended recipient of the communication</td>
</tr>
<tr>
<td>Date, time and duration of a communication.</td>
<td>1. The date and time of the start and end of the communication</td>
<td>1. The date and time of the log-in and log-off of the Internet access service, based on a certain time zone, together with the IP address, whether dynamic or static, allocated by the Internet access service provider to a communication, and the user ID of the subscriber or registered user.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. The date and time of the log-in and log-off of the Internet e-mail service or Internet telephony service, based on a certain time zone.</td>
</tr>
<tr>
<td>Communication type</td>
<td>The telephone service used</td>
<td>The Internet service used</td>
</tr>
<tr>
<td>The device used for communication</td>
<td>1. Fixed network telephony: The calling and called telephone numbers.</td>
<td>1. The calling telephone number for dial-up access.</td>
</tr>
<tr>
<td></td>
<td>2. Mobile telephony:</td>
<td>2. The digital subscriber line (DSL) or other end point of the originator of the communication.</td>
</tr>
<tr>
<td></td>
<td>a. The calling and called telephone numbers.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>b. The International Mobile Subscriber Identity (IMSI) of the calling and called party.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>c. The International Mobile Equipment Identity (IMEI) of the calling and called party.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3. In the case of pre-paid anonymous services:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>The date and time of the initial activation of the service and the location label (Cell ID) from which the service was activated.</td>
<td></td>
</tr>
<tr>
<td>Mobile communication equipment position</td>
<td>1. The location label (Cell ID) at the start of the Communication</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2. Data identifying the geographic location of cells by reference to their location labels (Cell ID) during the period for which communications data are retained.</td>
<td></td>
</tr>
</tbody>
</table>
2.2 National Assumptions

While the European Directive on Data Retention defines the European framework, it is the task of national legislation to implement on a national level. At this point in time, it is not known what will be the specifics of the Norwegian legislation and regulation when it comes to timeframe, scope and level of detail. The Police push for 12 months data retention. The most likely timeframe for a Norwegian legislation is early 2009.

This White Paper assumes that the basics of the Directive will be implemented within a foreseeable future, and that the industry must adapt to this. This White Paper also describes some of the vital factors for successful implementation with an industry point of view.

3. Challenges for implementation

There are a number of challenges for implementation of the Data Retention Directive. In this White Paper we focus on the most important issues for the industry from a practical point of view. We do not focus issues related to information security (i.e. personal information in the wrong hands), legal and regulatory processes.

3.1 National considerations – Norway

The telecommunications industry in Norway consists of approximately the following (source: Post- og Telestilsynet statistics and 120 second average call length):

- 7.5 million telephony subscriptions and 1.4 million Internet access subscriptions.
- 215 registered suppliers, offering one or several services.
- 10 billion calls and 6 billion SMS yearly. For Internet, statistics are not published, but there is an estimate of 20 billion e-mails yearly.

Regarding retrieval of data today, the Police department gets a court order and then contacts the individual supplier, asking for information. This is not an efficient process.

3.2 Contradictory interests: The industry and the authorities

Basically, the industry has no benefits from the EU Data Retention Directive. The Directive has been defined to have benefits for society, but for the industry as such it is purely a requirement where the industry has to comply. The speed and quality of implementation will depend on the legislation, the regulation, and the willingness to expose the industry with consequences for non-compliance.

For the authorities, it is of utmost importance that the implemented solutions are structured in such a way that the objective with the Directive is met in real life. Storage of data in itself will not be meaningful. The core is for the relevant authorities to gain access to relevant qualitative data in a meaningful way, so that they will assist in fighting crime in a timely, effective and efficient way. For this to happen, both accessibility and quality of data must be sufficient, and with 215 service providers in the market, this is a challenge.

There is also an important contradiction between crime fighting efficiency and citizens’ rights and respect of personal data. Efficient data retrieval increases the vulnerability of the data, as the solution can be misused by the police and by unfaithful servants. In Norway the Data Inspectorate will have a strong interest to monitor how and what data is actually retrieved by whom.

3.3 The industry

In this White Paper, the industry is defined as those providers that are defined by law as providers of telecommunication services and registered with the regulator. Today, this does not include hotels, cafes, companies and others that provide local Internet access (WLAN café, for instance). These players will, if affected by the Directive, have no chance to comply with regulations in a short term.
For the industry, the challenges are primarily related to the following;

- How to find and collect all data
- How to store and protect data
- How to access data – technical and procedures

**How to find and collect all data**

If the regulative is put into effect as it is written, all service providers will have major challenges to comply. Primarily, this relates to two areas;

- Providing user information for B-number subscribers. The Directive requires both name and address, and for mobile calls also IMSI and IMEI numbers to be stored. Service providers only have data for own subscribers, not for competitors’ customers, so providing data for B-number subscriber will be impossible.
- Providing cell ID for mobile calls. This is not always transferred from the radio network provider to the service provider, so the service provider does not have this data today in all cases
- Many providers do not match usage data with customer name in real time, so this can be a challenge if only short delays are allowed.

Most other data can probably be found and collected without major obstacles.

**How to store and protect data**

For a large provider, there will be large amounts of data. However, if there are no requirements for near real-time storage and access to the data, storage is not thought to be a major issue.

If the data is thought to have the same degree of sensitivity as call logs and billing details, all providers in the market will already have systems, localities and routines for adequate protection of data. If, however, there will be more stringent requirements, this will be a challenge for the smaller providers that does not have state of the art physical and logical data security.

The weakest point for the individual provider will typically be the ability to delete obsolete data on a continuous basis. In case of a distributed system, this will also be difficult for the authorities to monitor.

**How to access data**

This is a key to systems requirements and design. The Directive is weak on definitions in this area. It is not clear what will be asked for and how results shall be presented.

If only basic data is requested with simple questions, the provider can make a more or less manual or automated search in one or several systems and send the plain data back. If, however, there are requirements to formatting, level of detail and aggregation, electronic interfaces and response time, the provider will have to put significant effort into designing systems and procedures for this.

**3.4 The users – Police department and others**

How to ask for and receive data is vital in developing regulations for Data Retention. It will be a balance between effectiveness/efficiency and information security. The crime fighting authorities will benefit from fast, accurate and complete data, and also a wide selection of search, aggregation and wildcard criteria across service providers.

The exact same factors will complicate for the industry to comply.

**3.5 The regulating authorities**

For the industry it is not important who will be the regulating authority. The crucial factor is that the regulating authorities have a realistic picture of how the industry works and what the industry can comply with.
However, an important perspective is the protection of citizens’ rights and respect of personal data. The Data Inspectorate will be monitoring this, and with a solution where data can be retrieved from more than 200 sources, monitoring will be a difficult task. If the data and/or the data retrieval process are protected and controlled by a trusted third party, all access and retrieval can be controlled, monitored and documented efficiently. Thereby, not only data retention and data retrieval will be efficient, but the regulation of what data is allowed to be used will also be practiced efficiently.

3.6 Other implications
There should probably be a tracing functionality, so that it is possible to find out what has been searched for and by whom. For this function to be trusted, it cannot be implemented at the retrieving side (the Police). It must be implemented at a trusted third party and/or in respective systems where data is stored and retrieved. This will potentially complicate for the industry. With a centralised solution, this can be integrated with no effort for the service providers.

Information security, especially regarding unauthorized access, loss of data, and removal of outdated data is an important perspective. With 215 service providers in the market, this will be a challenge. With a centralised solution, this can be controlled in a much better way.

4. Implementation proposals and annual cost
Today there are about 200 public (very different sized) telecom service providers in Norway. Telephone numbers are allocated in number series to service providers and numbers are ported between service providers as the customer changes provider. It is therefore not possible to find a service provider based on a telephone number without having access to NRDB porting table. The NRDB porting table also includes porting history that may be needed to find the right service provider after a number has been ported. For national wide data retention solution it is not efficient, for any party, if all information requests are sent to all service providers. Thus a central proxy service is needed.

Different architectural models can be assumed for a national data retention solution. Data can be stored and managed locally at each service provider or stored and managed by a central data retention service. From a theoretical view point we can divide in central and distributed solutions. However, experience from prior NRDB projects has shown that the most efficient solution is a combined solution consisting of a central data retention central with central database and NRDB porting table look up, and local nodes with local databases. Please see figure above and technical chapter for more information and elaboration on different alternatives that can be realised with the NRDB structure.

A combined solution will provide the right flexibility for all parties involved. Small service providers will get a simple and cost efficient solution via the central data retention central with central database. A local node will fulfill the internal business security requirement of large service providers and avoid large amount of data transported to a central database. The crime fighting authority will have a single information request interface. The following Technical Section gives an in depth description of the possible architectural models with pro’s and con’s for each model.

A combined solution consisting of 20 local nodes with local databases and 100 service providers uploading data to a central data base has an estimated yearly cost of 11 – 17
MNOK. This solution will cover 7.5 million telephony subscriptions and 1.4 million Internet subscriptions. The remaining providers in the market will most probably be small virtual service providers, where the wholesale provider will handle uploading or local storage for them. Thus the yearly cost of Data Retention will be between NOK 1,20 and NOK 1,90 per subscription.

Implementation of a centralised solution does not rule out use of other platforms, for instance in cases where a provider prefer proprietary solutions.

5. Key Success Factors for implementation

In this White Paper, successful implementation is defined as a solution that
- can be delivered by the industry without large effort
- contributes to crime fighting
- has a acceptable balance between efficiency and information security

In order to make this happen, many interests must be balanced. The most important issues for the industry are the following;

- The right balance regarding what data to collect from service providers.
  - Providers typically cannot collect B-number name, address, IMEI and IMSI, since these are not in the providers databases. If this is a requirement, it must be provided from a centralized register
  - Service providers typically cannot collect cell information for mobile calls, since only the network providers has this information
  - Limiting Internet communication logs to services that are provided by the provider, and not for all 3rd party e-mail providers etc.
- A clear expectation on data retrieval. What data shall be possible to retrieve, how shall it be retrieved, who shall ask for data and how, and what is the response time.
  - What shall be the search criteria? Only ID (number/IP), or also names and misspelled names?
  - Should the crime fighting authorities be able to search in all providers’ databases in all cases?
  - Should the crime fighting authorities only search for individual items or it is expected to handle list search, for instance based on a server log with many IP-addresses?
- A clear expectation on the balance between data retrieval efficiency and the protection of citizens’ rights and respect of personal data
  - To secure efficient use of data, and at the same time controlling possible misuse of information
  - To secure deletion and non-retrieval of old data
- The correct systems design, based on what data is to be retrieved and how
  - How to assure a consistent fulfilment of the requirements set forth in the directive, and how to track down and sanction eventual breaches or shortcomings with respect to the requirements?
  - If there is a distributed system with local storage, specifications must be clear or automated retrieval will be impossible. Even with clear specifications, it must be planned for manual work in all steps of the retrieval process
  - If the decision is a centralized system in order to solve multi-provider searches, B-number identity, tracing of search activity, etc, retrieval processes can be standardised

6. Conclusion

The industry has to comply with national legislation based on the European Directive on Data Retention. For the industry however, there are no benefits other than compliance with regulations to implement Data Retention. The benefits rest solely on the crime fighting authorities.
For successful implementation, national legislation must be clear on what data to retain. Also, expectation on data retrieval efficiency must be defined. This regards what and how data is retrieved, as well as requirements for response time. Citizens’ rights and respect of personal data must also be secured, including the monitoring of the data retrieval process.

After evaluating several solution models, NRDB have concluded that centralised handling is necessary for data retention to work efficiently on a national level. Without centralised handling, there will be too many interfaces and too much integration, and too little monitoring of the data retrieval process. The centralised model can be supplemented with local nodes for the service providers that prefer this model. If regulators agree, other NRDB resources such as routing/porting tables and name/address functions can be utilized.

For Number Portability, the telecommunications industry has proven that it is possible to develop and manage a centralised platform for a low cost and with a high degree of availability. Data Retention could be handled in a similar way.

A solution as proposed in this document does not rule out other possibilities. From a technical perspective, provider proprietary platforms can be used in parallel, as well as other initiatives. The actual implementation will be a consequence of the regulation and service providers’ preferences.
Technical considerations
Models for implementation, utilising the NRDB framework
7. Architectural models

In the following, different architectural models are described and discussed. Whether data are stored and managed locally at each service provider or stored and managed by a central data retention service is one of the fundamental questions we try to give an answer to.

The architectural models may be analysed from two points of view:

- The service providers’ point of view; - the need to fulfil the requirements set forth by the directive 2006/24/EC in relation to the existing routines for generation and processing of traffic data.
- The authorities’ point of view; - the need for quick access to the retained data including sufficient search and lookup capabilities.

7.1 Distributed system

A distributed system is defined as follows: Each service provider implements its own, independent data retention service and provides access to the authorities for information disclosure via a user interface specific to this service provider. The authorities find out (in some way) where the actual traffic data is stored and send an information request to the actual service provider (possibly more than one service provider).

![Figure 1 Infrastructure for distributed system](image)

7.1.1 Discussion

A distributed system is flexible in the way that each service provider can choose its own preferred solution for data retention. Each service provider has full control with the use of its traffic data, but the overall fulfilment of the requirements becomes unclear, as the responsibility and control is fully distributed.

The overall cost (counting the costs for each service provider) for this solution is very high as all service providers must establish a sufficient solution on their own, - also taking into account information, security requirements and access management. This might also lead to inconsistent fulfilment of the directive.

The authorities are facing a number of independent data retention query interfaces, - complicating the information lookup. A standardized information request API would improve the overall solution considerably, - but the authorities would still need to find out which service provider to request.
7.2 Distributed system with central proxy

This type of system is an improvement of the plain distributed system, in particular from the authorities' point of view. A distributed system with central proxy is illustrated below.

Figure 2 Infrastructure for distributed system with proxy

7.2.1 Discussion

The authorities may use the proxy to work around the problem created by a large number of independent data retention interfaces. The proxy will be the only interface used by the authorities and will pass on all requests on their behalf, as well as pass on all responses (traffic data) on behalf of the actual service provider. The underlying network of different service providers is hence abstracted away.

The proxy can also provide the following services on behalf of the service providers:

- **Access management** (authenticate and authorize lookup requests from the authorities)
- **Event logging** (all information requests should be logged, including client ID and source IP address)
- **Statistics and reports for lookup request**

The proxy becomes more complex the more service providers that are connected, - especially if there is no standardized information request API. To fully benefit from a proxy solution, it is required that requests can be handled via a machine-to-machine interface. The proxy should be equipped with a kind of catalogue service providing information on the relation between for example phone number (range) and responsible service provider, and IP range and ISP. Alternatively the proxy could initiate a request against all information sources simultaneously, and “merge” the result before returning to the caller.

The proxy is located on top of the distributed system and needs to be financed and operated separately. Telecommunications service providers need to implement a near full solution on their own, - as the solution from a service provider’s point of view is practically the same as for the distributed system.

7.3 Central system
A central system entails that all traffic data to be retained are stored and maintained by a Data Retention Central. Telecommunications service providers upload traffic data to the central which also handles all information requests. A central system is illustrated below.

![Diagram](image)

**Figure 3 Infrastructure for central system**

### 7.3.1 Discussion

A central system will assure a consistent management of data retention as all data are stored and managed by the same entity. All requirements set forth in the directive (except from the generation of data) will be fulfilled on behalf of the service providers.

The data volume to be handled by the Data Retention Central will be summed up to the total volume of data generated by all connected service providers. Large data transfers (I/O) will require broad bandwidth.

A central store with all traffic data will ease the processing of information requests and assure high quality on the returned data since the same query can operate on data from all service providers. This fact will on the other hand also make the database more sensible and possibly require a higher degree of information security (encrypted data should be considered).

### 7.3.2 Example configuration of a Data Retention Central

Suggested modules and APIs to be provided by a Data Retention Central are described in this section.

The lower part of the figure shows different data sources managed by the actual service provider. The Data Retention Central provides a “Data Collection API” for uploading traffic (and customer) data. Once uploaded, the service provider may delete the original data according to internal (and existing) routines. The lifecycle for retained data (according to the directive) will be assured by the Data Retention Central.

The upper part of the figure shows the interface towards the authorities.
An example system design for the Data Retention Central is shown below.

Uploaded data are being processed (and saved in the database) by the “Data Collection and Processing Server”, which typically provides a secure FTP Server or a Web services interface to the service providers.

The database server should be placed in a dedicated security (trust) zone and duplicated with a standby replica of the primary database. The standby server could also be placed on a geographically separated “Disaster Recovery Site”.

In addition to “hot” and online backup to a replica server, data should also be backed up to tape. Due to the large amount of data, a tape robot will be needed for this task.

The Web server (application server) will handle the incoming information requests from the authorities, including authentication.
7.4 Combined system

In today’s complex telecommunication services community, both with respect to the number and diversity of actors and the different types of telecommunications media it might be difficult to limit the implementation to a fully distributed or a fully centralized system. A combined solution is illustrated in this section, where some traffic data are stored centrally and others are stored at each service provider.

There is, however, possible to find a compromise between provider-internal data retention and centralized data retention via the use of Local Data Retention Nodes. This is typically local versions of the Data Retention Central, consisting of:

- A Data Collection API that the service provider uses to upload traffic data
- A database for retained data
- A Web Services based lookup interface that is used by the Data Retention Central for delegated information lookup

A Local Data Retention Node may fulfill the data retention requirements on behalf of the service provider, as long as traffic data are uploaded and a sufficient infrastructure (including Internet access) is provided for the solution. The Local Data Retention Node may act as a “certified” data retention solution that is integrated with the Data Retention Central.

A Local Data Retention Node installation should as a minimum consist of two servers: one primary database server/application server and one standby server. Depending on the data volume to be handled the system could be extended with a separate storage system and a separate application server.

Figure 5 Data Retention Central system design
7.4.1 Discussion
For a service provider selecting the central alternative this combined service is exactly the same as the one described for the central system. The service is also practically the same for a service provider opting for a Local Data Retention Node solution (with the obvious difference that data is stored and processed locally).

A specific data request API between a service provider using the distributed alternative and the Data Retention Central will be defined. This machine-to-machine interface will secure that some of the directive’s requirements, such as data retrieval, will be fulfilled by the Data Retention Central, - also on behalf of the service provider selecting the distributed alternative.

7.4.2 Centralized services
The following table lists different requirements set forth by or inferred from the directive. The “Centrally fulfilled” columns indicate whether the actual requirement is fulfilled by the Data Retention Central (or a Local Data Retention Node) on behalf of the service provider, depending on which connection that is selected (distributed or centralized/node storage of traffic data).

<table>
<thead>
<tr>
<th>Requirement</th>
<th>Centrally fulfilled?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collect traffic data by processing data sources (files)</td>
<td>Distributed</td>
</tr>
<tr>
<td>Store traffic data for retention</td>
<td>No</td>
</tr>
<tr>
<td>Assure confidentiality (encryption of stored data if required)</td>
<td>No</td>
</tr>
<tr>
<td>Assure availability (redundancy, backup)</td>
<td>No</td>
</tr>
<tr>
<td>Assure integrity (access control and consistent processing of input data, audit logs)</td>
<td>No</td>
</tr>
<tr>
<td>Assure deletion according to current specifications for retention periods</td>
<td>No</td>
</tr>
<tr>
<td>Provide lookup and searching capabilities</td>
<td>No</td>
</tr>
<tr>
<td>Maintain and update data retention procedures according to current requirements (laws and regulations)</td>
<td>No</td>
</tr>
<tr>
<td>Manage lookup access control (from authorities) including authentication and authorization</td>
<td>Yes</td>
</tr>
<tr>
<td>Event logging (logging all information requests including client ID and source IP address)</td>
<td>Yes</td>
</tr>
</tbody>
</table>
Manage data disclosure | Yes | Yes |
--- | --- | --- | --- |
Generate and provide statistics on data usage | Yes | Yes |

7.5 NRDB Data Ret. Central utilizing NRDB resources for info lookup

A Data Retention service implemented and maintained as part of the NRDB service “suite” may also be designed in a way that utilizes existing NRDB resources, namely:

- Numbering plan, providing information on the allocation of number ranges to telecommunications service providers, - for unported phone numbers
- Porting table, providing information on the responsible telecommunications service providers, - for ported phone numbers
- Number lookup services, providing information on the terminating service provider (access provider) for a specific (active) phone number
- Emergency location service, providing information on the customer and user (including name and address) of a specific phone number. Currently, the addresses for about 2 million phone numbers are registered.

7.5.1 Numbering plan and porting table

The numbering plan and porting table can be used to find the source to the information related to a phone number. The correct service provider (that is responsible for the actual subscription) can always be found, - regardless of whether the phone number is ported or not.

The numbering table and porting table also contain historical data, - meaning that the range holder or responsible service provider at any time (since the start of NRDB) can be found from the registers, - regardless of how many times a number range has been transferred or a phone number has been ported.

This resource can be used for a distributed solution or a Local Data Retention Node solution to direct an information lookup request to the correct service provider, based on phone number.

7.5.2 Number lookup and emergency location

The number lookup services can be used to look up the terminating provider (access provider) that is related to the actual service provider in case the service provider is reselling a telephone service (fixed or mobile).

The emergency location service does also provide the terminating provider (access provider) in addition to the customer name and address related to a phone number. Access to the emergency location service should not be involved with the access to a data retention service, - but the imported data directed to the emergency location database could also be directed to a data retention database making a copy that can be used exclusively for data retention purposes.

A service provider that is reselling telephone services only can then fulfil the data retention requirements with no extra effort, given that:

1. The actual service provider is a registered user of the NRDB Data Retention Central.
2. The actual service provider is a user of the NRDB Emergency Location service, - and uploads address data according to the solution’s requirements (centralized solution).
3. The terminating provider (access provider) is providing a machine-to-machine interface for data retention lookup (for example via a Local Data Retention Node), or uploads all traffic data to a Data Retention Central (central system).

The NRDB Data Retention Central will then find address data by using the “emergency location” data, and find traffic data by issuing a request to the terminating provider.