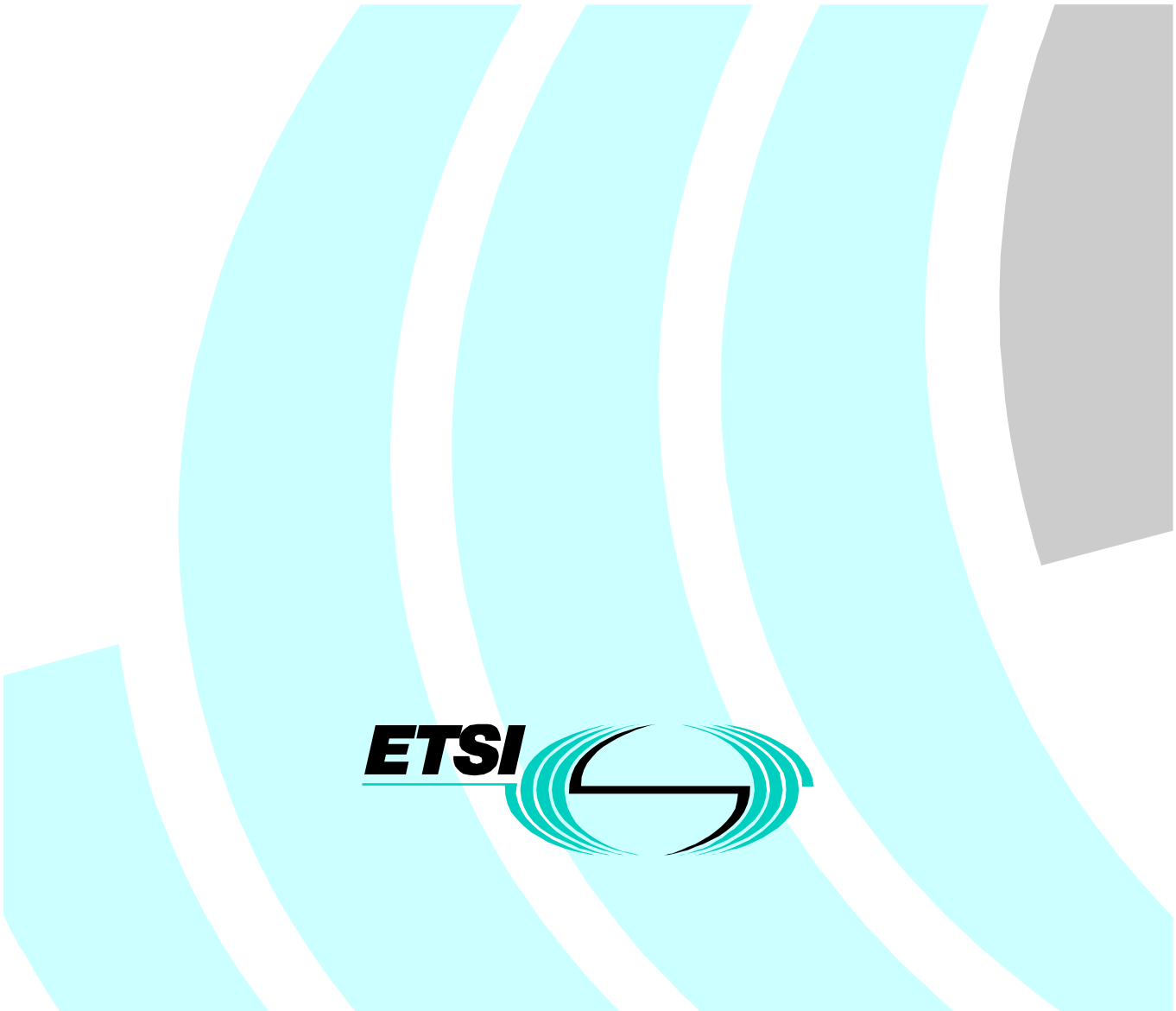
ETSI TR 101 876 V1.1.1 (2001-01)

Technical Report

Telecommunications security;
Lawful Interception (LI);
Description of GPRS HI3



Reference
DTR/SEC-003012

Keywords
GPRS, security

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - NAF 742 C
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° 7803/88

Important notice

Individual copies of the present document can be downloaded from:
http://www.etsi.org

The present document may be made available in more than one electronic version or in print. In any case of existing or perceived difference in contents between such versions, the reference version is the Portable Document Format (PDF). In case of dispute, the reference shall be the printing on ETSI printers of the PDF version kept on a specific network drive within ETSI Secretariat.

Users of the present document should be aware that the document may be subject to revision or change of status. Information on the current status of this and other ETSI documents is available at http://www.etsi.org/tb/status/

If you find errors in the present document, send your comment to:
editor@etsi.fr

Copyright Notification

No part may be reproduced except as authorized by written permission.
The copyright and the foregoing restriction extend to reproduction in all media.

© European Telecommunications Standards Institute 2001.
All rights reserved.

Contents

Intellectual Property Rights 4
Foreword 4
1 Scope 5
2 References 5
3 Definitions and abbreviations 6
3.1 Definitions 6
3.2 Abbreviations 6
4 Functional architecture 7
5 Correlation 7
6 HI3 (Delivery Content of Communication (CC)) 8
6.1 GTP* 8
6.1.1 Introduction 8
6.1.2 Definition of GTP* Header 8
6.1.3 Exceptional Procedure 9
6.1.4 Other Considerations 10
6.2 FTP 10
6.2.1 Introduction 10
6.2.2 Usage of the FTP protocol 10
6.2.3 Exceptional procedures 12
6.2.4 CC Contents for FTP 12
6.2.4.1 Fields 12
6.2.4.2 Information Element Syntax 13
6.2.5 Other Considerations 14
History 15

Intellectual Property Rights

IPRs essential or potentially essential to the present document may have been declared to ETSI. The information pertaining to these essential IPRs, if any, is publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (http://www.etsi.org/ipr).

Pursuant to the ETSI IPR Policy, no investigation, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Foreword

This Technical Report (TR) has been produced by ETSI Technical Committee Security (SEC).

1 Scope

The present document specifies the HI3 (interface for the delivery of CC to the LEMF) for GPRS. This will be an informative annex for GPRS HI3, in ES 201 671 [12].

2 References

For the purposes of this Technical Report (TR) the following references apply:

[1] GSM 01.04: "Digital cellular telecommunications system (Phase 2+); Abbreviations and acronyms".

[2] GSM 01.33: "Digital cellular telecommunications system (Phase 2+); Lawful Interception requirements for GSM".

[3] GSM 02.33: "Digital cellular telecommunications system (Phase 2+); Lawful Interception stage 1".

[4] GSM 03.33: "Digital cellular telecommunications system (Phase 2+); Lawful Interception stage 2".

[5] GSM 03.60: "Digital cellular telecommunications system (Phase 2+); GPRS Service description stage 2".

[6] GSM 09.02: "Digital cellular telecommunications system (Phase 2+); Mobile Application Part (MAP) specification".

[7] GSM 09.60: "Digital cellular telecommunications system (Phase 2+); General Packet Radio Service (GPRS); GPRS Tunneling Protocol (GTP) across the Gn and Gp Interface".

[8] GSM 12.15: "Digital cellular telecommunications system (Phase 2+); General Packet Radio Service (GPRS); GPRS Charging".

[9] IETF STD 09: "File Transfer Protocol" (RFC 0959).

[10] IETF STD 05: "Internet Protocol". : "Empty".

[11] IETF STD 07: "Transmission Control Protocol".

[12] ETSI ES 201 671: "Telecommunications security; Lawful Interception (LI); Handover interface for the lawful interception of telecommunications traffic".

[13] ETSI TS 132 015: "Digital cellular telecommunications system (Phase 2+) (GSM); Universal Mobile Telecommunications System (UMTS); Telecommunication Management; Charging and billing; 3G call and event data for the Packet Switched (PS) domain; (3GPP TS 32.015 version 3.3.0 Release 1999)".

[14] ETSI TS 129 060: "Digital cellular telecommunications system (Phase 2+) (GSM); Universal Mobile Telecommunications System (UMTS); General Packet Radio Service (GPRS); GPRS Tunnelling Protocol (GTP) across the Gn and Gp Interface; (3GPP TS 29.060 version 3.6.0 Release 1999)".

3 Definitions and abbreviations

3.1 Definitions

See definitions in GSM 02.33 [3] and ES 201 671 [12].

3.2 Abbreviations

See abbreviations in GSM 03.33 [4] and ES 201 671 [12].

For the purposes of the present document, the following abbreviations apply:

FTP	File Transfer Protocol
GGSN	Gateway GPRS Support Node
GSN	GPRS Support Node
GTP	GPRS Tunneling Protocol
GTP*	GTP star
xGSN	SGSN or GGSN
IP	Internet Protocol
LIID	Lawful Interception Identifier
MF	Mediation Function
SGSN	Serving GPRS Support Node
TCP	Transmission Control Protocol
TID	Tunnel Identifier
T-PDU	tunneled PDU
PDU	Protocol Data Unit

4 Functional architecture

The following picture contains the reference configuration for lawful interception (see GSM 03.33 [4]).

There is one Administration Function (ADMF) in the network. Together with the delivery functions it is used to hide from the xGSN that there might be multiple activation's by different Law Enforcement Agencies (LEAs) on the same target.

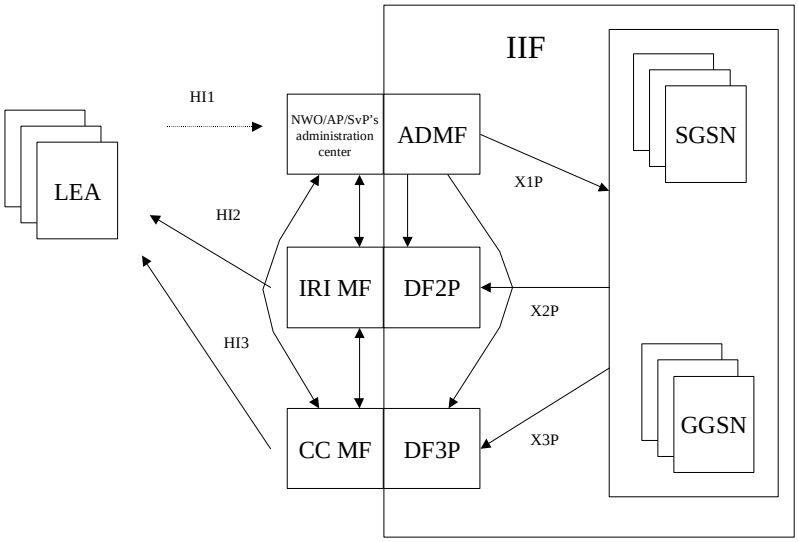


Figure 1: Reference configuration

NOTE: GGSN interception is a national option

The reference configuration is only a logical representation of the entities involved in lawful interception and does not mandate separate physical entities. This allows for higher levels of integration.

A call could be intercepted based on several identities (MSISDN, IMSI, IMEI) of the same target.

For the delivery of the CC and IRI the xGSN provides a correlation number and target identity to the DF2P and DF3P which is used there to select the different LEAs where CC/IRI shall be delivered to.

5 Correlation

Correlation of ES 201 671 [12] ID's to GSM ID's of GSM 03.33 [4]:

- Lawful interception identifier (LIID) → Warrant reference number;
- Network identifier (NID) → xGSN address.

6 HI3 (Delivery Content of Communication (CC))

There are two possible methods for delivery l content of communication to the LEMF:

- GTP* (see clause 6.1);
- FTP (see clause 6.2).

According to national requirements al least one of these methods have to be provided.

6.1 GTP*

6.1.1 Introduction

The header and the payload of the communication between the intercepted subscriber and the other party (later called: Information Element) are duplicated. A new header (later called: GTP*-Header, see table 1) is added (see table 3) before it is sent to LEMF.

For data transfer over the HI3 interface in GPRS the existing GTP protocol is used in a modified way. The only adaptation to GSM 09.60 [7] is to replace the TID by a correlation number identifying the intercepted product (context) unambiguously even in case of SGSN change. The modified protocol will be called GTP*.

GTP* could be used via UDP or TCP/IP.

6.1.2 Definition of GTP* Header

GTP* header contains the following attributes:

- Correlation Number;
- Message Type (analogue to GTP a value of 255 is used for HI3-PDU's);
- Direction;
- Sequence Number;
- Length;
- T-PDU contains the intercepted information.

Table 1: Outline of GTP* header

Octets	Bits 8	7	6	5	4	3	2	1
1	Version ('0 0 0')			'1'	Spare '1 1'		DIR	'0'
2	Message Type (value 255)							
3-4	Length							
5-6	Sequence Number							
7-8	not used (value 0)							
9	not used (value 255)							
10	not used (value 255)							
11	not used (value 255)							
12	not used (value 255)							
13-20	correlation number							

For interception tunneling the GTP* header shall be used as follows:

Version shall be set to 0 to indicate the first version of GTP*.

DIR indicates the direction of the T-PDU:

"1" indicating uplink (from observed mobile user); and

"0" indicating downlink (to observed mobile user).

Message Type shall be set to 255 (the unique value that is used for T-PDU within GTP [6]).

Length shall be the length, in octets, of the signaling message excluding the GTP* header. Bit 8 of octet 3 is the most significant bit and bit 1 of octet 4 is the least significant bit of the length field.

Sequence Number is an increasing sequence number for tunneled T-PDUs. Bit 8 of octet 5 is the most significant bit and bit 1 of octet 6 is the least significant bit of the sequence number field.

Correlation Number consists of two parts: GGSN-ID identifies the GGSN which creates the Charging-ID.

Charging-ID is defined in [7] and assigned unique to each PDP context activation on that GGSN (4 octets).

The correlation number consists of 8 octets and guarantees a unique identification of the tunnel to the LEA over a long time. The requirements for this identification are similar to that defined for charging in [7], clause 5.4. Therefore it is proposed to use the Charging-ID, defined in [7], clause 5.4 as part of correlation number. The Charging-ID is signaled to the new SGSN in case of SGSN-change so the tunnel identifier could be used "seamless" for the HI3 interface.

Table 2: Outline of correlation number

0 1 2 3 4 5 6 7	8 9 0 1 2 3 4 5	6 7 8 9 0 1 2 3	4 5 6 7 8 9 0 1	
Charging –ID Octet 1	Charging –ID Octet 2	Charging –ID Octet 3	Charging –ID Octet 4	Octet 13-16
GGSN-ID				Octet 17-20

The GTP* header is followed by a subsequent payload information element. Only one information element is allowed in a single signaling message.

Table 3: GTP* header followed by the subsequent payload Information Element

	Bits							
Octets	8	7	6	5	4	3	2	1
1 – 20	GTP*-Header							
21 – n	Information Element							

The Information Element contains the header and the payload of the communication between the intercepted subscriber and the other party.

6.1.3 Exceptional Procedure

With UDP and GTP*: the delivering node doesn't take care about any problems at LEMF.

With TCP and GTP*: TCP tries to establish a connection to LEMF and resending (buffering in the sending node) of packets is also supported by TCP.

In both cases it might happen that call content gets lost (in case the LEMF or the transit network between MF and LEMF is down for a long time).

6.1.4 Other Considerations

The use of Ipsec for this interface is recommended.

The required functions in LEMF are:

- collecting and storing of the incoming packets inline with the sequence numbers;
- correlating of CC to IRI with the use of the correlation number in the GTP* header.

6.2 FTP

6.2.1 Introduction

At HI3 interface FTP protocol is used over TCP/IP stack for the delivery of the result of interception. The FTP protocol is defined in the IETF standard STD 09 "File Transfer Protocol" (RFC 0959). The TCP is defined in the STD 07 "Transmission Control Protocol". The IP is defined in the STD 05 "Internet Protocol".

FTP supports reliable delivery of data. The data may be temporarily buffered in the sending node (MF) in case of link failure. FTP protocol is independent of the payload data it carries.

6.2.2 Usage of the FTP protocol

In the packet data LI the MF acts as the FTP client and the receiving node (LEMF) acts as the FTP server. The client pushes the data to the server.

The receiving node LEMF stores the received data as files. The sending entity (MF) may buffer files.

Several smaller intercepted data units may be gathered to bigger packages prior to sending, to increase bandwidth efficiency.

The following configurable intercept data collection (= transfer package closing / file change) threshold parameters should be supported:

- frequency of transfer, based on send timeout, e.g. X ms;
- frequency of transfer, based on volume trigger, e.g. X octets.

There are two possible ways how the interception data may be sent from the MF to the LEMF. One way is to produce files that contain interception data only for one observed target (ref: "File naming method A)"). The other way is to multiplex all the intercepted data that MF receives to the same sequence of general purpose interception files sent by the MF (ref: "File naming method B)").

The HI2 and HI3 are logically different interfaces, even though in some installations the HI2 and HI3 packet streams might also be delivered via a common transmission path from a MF to a LEMF. It is possible to correlate HI2 and HI3 packet streams by having common (referencing) data fields embedded in the IRI and the CC packet streams.

File naming:

The names for the files transferred to a LEA are formed according to one of the 2 available formats, depending on the delivery file strategy chosen (e.g. due to national convention or operator preference).

Either each file contains data of only one observed target (as in method A) or several targets' data is put to files common to all observed target traffic through a particular MF node (as in method B).

The maximum set of allowed characters in interception file names are "a"…"z", "A"…"Z", "-", "_", ".", and decimals "0"…"9".

File naming method A):

```
<LIID>_<seq>.<ext>
```

LIID = as defined in the ES 201 671 [12]. This field has a character string value, e.g. "ABCD123456". This is a unique interception request identifier allocated by the ADMF. It will be given by the ADMF to the LEA via the HI1 interface after the ADMF has been authorized to command the start of the interception of a specific target. The possible network operator identifier part used should be agreed with (and allocated by) the regulatory organization administrating the local telecommunication practises.

Seq = integer ranging between [0..2^64-1], in ASCII form (not exceeding 20 ASCII digits), identifying the sequence number for file transfer from this node per a specific target.

Ext = ASCII integer ranging between ["1".."7".] (in hex: 31H…37H), identifying the file type. The possible file type codings for intercepted data are shown in table 4. But for the HI3 interface, only the types "2", "4", and "6" are possible.

Table 4: Possible file types

File types that the LEA may get	Intercepted data types
"2" (in binary: 0011 0010)	CC(MO)
"4" (in binary: 0011 0100)	CC(MT)
"6" (in binary: 0011 0110)	CC(MO&MT)

(The least significant bit that is '1' in file type 1, is reserved for indicating IRI data.) The bit 2 of the **ext** tells whether the Mobile Originated (MO) Content of Communication (CC) is included to the intercepted data.

The bit 2 of the **ext** tells whether the Mobile Originated (MO) Content of Communication (CC) is included to the intercepted data.

The bit 3 of the **ext** tells whether the Mobile Terminated (MT) Content of Communication (CC) is included to the intercepted data.

Thus, for Mobile Originated Content of Communication data, the file type is "2", for MT CC data "4" and for MO & MT CC data "6".

This alternative A is used when each target's intercepted data is gathered per observed target to dedicated delivery files. This method provides the result of interception in a very refined form to the LEAs, but requires somewhat more resources in the sending node than alternative B. With this method, the data sorting and interpretation tasks of the LEMF are considerably easier to facilitate in near real time than in alternative B.

File naming method B):

The other choice is to use monolithic fixed format file names (with no trailing file type part in the file name):

```
<filenamestring>   (e.g. ABXY00041014084400006)
```

where:

ABXY = Source node identifier part, used for all files by the mobile network operator "AB" from this MF node named "XY";

00 = year 2000;

04 = month April;

10 = day 10;

14 = hour;

08 = minutes;

44 = seconds;

0000 = extension;

6 = file type. Coding: "2" = CC(MO), "4" = CC(MT), "6" = CC(MO&MT). (The type "1" is reserved for IRI data files).

This alternative B is used when several targets' intercepted data is gathered to common delivery files. This method does not provide the result of interception in as refined form to the LEAs as the alternative A, but it is faster in performance for the MF point of view. With this method, the MF does not need to keep many files open like in alternative A.

6.2.3 Exceptional procedures

Overflow at the receiving end (LEMF) is avoided due to the nature of the protocol.

In case the transit network or receiving end system (LEMF) is down for a reasonably short time period, the local buffering at the MF will be sufficient as a delivery reliability backup procedure.

In case the transit network or receiving end system (LEMF) is down for a very long period, the local buffering at the MF may have to be terminated. Then the following intercepted data coming from the intercepting nodes towards the MF would be discarded, until the transit network or LEMF is up and running again.

6.2.4 CC Contents for FTP

6.2.4.1 Fields

The logical contents of the CC-header is described here.

CC-header = (Version, HeaderLength, PayloadLength, PayloadType, PayloadTimeStamp, PayloadDirection, CCSeqNumber, CorrelationNumber, LIID, PrivateExtension).

The Information Element CorrelationNumberforms the means to correlate the IRI and CC of the communication session intercepted.

The first column indicates whether the Information Element referred is Mandatory, Conditional or Optional.

The second column is the Type in decimal.

The third column is the length of the Value in octets.

M	130	2	**Version** = the version number of the format version to be used. This field has a decimal value, this enables version changes to the format version. The values are allocated according to national conventions.
M	131	2	**HeaderLength** = Length of the CC-header up to the start of the payload. Length: 2 octets.
M	132	2	**PayloadLength** = Length of the payload following the CC-header.
M	133	1	**PayloadType** = Type of the payload, indicating the type of the CC. Type of the payload. This field has a decimal value. The possible PDP Type values can be found in the standards, e.g. 3G TS 29.060 [14]. The value 255 is reserved for future PDP Types and means: "Other".
O	134	4	**PayloadTimeStamp** = Payload timestamp according to intercepting node. (Precision: 1 second, timezone: UTC). Format: Seconds since 1970-01-01 as in e.g. Unix (length: 4 octets).
C	137	1	**PayloadDirection** = Direction of the payload data. This field has a decimal value 0 if the payload data is going towards the target (i.e. downstream), or 1 if the payload data is being sent from the target (i.e. upstream). If this information is transferred otherwise, e.g. in the protocol header, this field is not required as mandatory. If the direction information is not available otherwise, it is mandatory to include it here in the CC header.
O	141	4	**CCSeqNumber** = Identifies the sequence number of each CC packet during interception of the target. This field has a 32-bit value.
M	144	8 or 20	**CorrelationNumber.** Identifies an intercepted session of the observed target. This can be implemented by using e.g. the Charging Id (4 octets, see TS 3GPP 32.015 [13]) with the (4-octet/16-octet) Ipv4/Ipv6 address of the PDP context maintaining GGSN node attached after the first 4 octets.
			<Possible future parameters are to be allocated between 145 and 253.>
O	254	1-20	**LIID** = Field indicating the LIID as defined in the ES 201 671 [12]. This field has a character string value, e.g. "ABCD123456".
O	255	1-N	**PrivateExtension** = An optional field. The optional Private Extension contains vendor or LEA or operator specific information. It is described in the document 3G TS 29.060 [14].

6.2.4.2 Information Element Syntax

The dynamic TypeLengthValue (TLV) format is used for its ease of implementation and good encoding and decoding performance. Subfield sizes: Type = 2 octets, Length = 2 octets and Value = 0…N octets. From Length the T and L subfields are excluded. The Type is different for every different field standardized.

The octets in the Type and Length subfields are ordered in the little-endian order, (i.e. least significant octet first). Any multioctet Value subfield is also to be interpreted as being little-endian ordered (word/double word/long word) when it has a (hexadecimal 2/4/8-octet) numeric value, instead of being specified to have an ASCII character string value. This means that the least significant octet/word/double word is then sent before the more significant octet/word/double word.

TLV encoding:

TLV encoding can always be applied in a nested fashion for structured values.

(The small "v" refers to the start of a Value field that has inside it a nested structure.)

The first octet of the first TLV element will start right after the last octet of the header of the protocol that is being used to carry the CC information.

The first TLV element (i.e. the main TLV IE) comprises the whole dynamic length CC information, i.e. the dynamic length CC header and the dynamic length CC payload.

Inside the main TLV IE there are at least 2 TLV elements Header of the payload and the Payload itself. The Header contains all the ancillary Ies related to the intercepted CC packet. The Payload contains the actual intercepted packet.

There may be more than one intercepted packet in one GPRS HI3 delivery protocol message. If the Value of the main TLV IE is longer than the 2 (first) TLV Information Elements inside it, then it is an indication that there are more than one intercepted packets inside the main TLV IE (i.e. 4 or more TLV Ies in total). The number of TLV Ies in the main TLV IE is always even, since for every intercepted packet there is one TLV IE fistits header and one TLV Iistor its.

6.2.5 Other Considerations

The FTP protocol mode parameters used:

Transmission Mode: stream;

Format: non-print;

Structure: file-structure;

Type: binary.

The FTP service command to define the file system function at the server side: STORE mode for data transmission.

The FTP clien– (=user -FTP process at the MF) uses e.g. the default standard FTP ports 20 (for data connection) and 21 (for control connection), 'passive' mode is supported. The data transfer process listens the data port for a connection from a server-FTP process.

For the file transfer from the MF to the LEMF(s) e.g. the following data transfer parameters are provided for the FTP client (at the MF):

- transfer destination (IP) address, e.g. "194.89.205.4";
- transfer destination username, e.g. "LEA1";
- transfer destination directory path, e.g. "/usr/local/LEA1/1234-8291";
- transfer destination password;
- interception file type, e.g. "2" (this is needed only if the file naming method A is used).

LEMF may use various kind directory structures for the reception of interception files. It is strongly recommended that at the LEMF machine the structure and access and modification rights of the storage directories are adjusted to prevent unwanted directory operations by a FTP client.

The use of IPSec services for this interface is recommended.

History

Document history		
V1.1.1	January 2001	Publication