

Ethernet over SONET/SDH

GFP, VCAT and LCAS

Technology White Paper

Serge-Paul Carrasco

Abstract

Enterprise networks are demanding Ethernet services to their Service Providers. In a time when capital is constrained, Service Providers want to use their present SONET/SDH infrastructure to transport Ethernet in their access and metro networks.

Three new technologies are emerging, which combine, can optimize SONET/SDH for data transport: the Generic Framing Protocol (GFP), Virtual Concatenation (VCAT) and LCAS (Link Capacity Adjustment).

GFP provides a new and more efficient Layer 2 encapsulation scheme for data traffic over SONET/SDH. VCAT allows the logical concatenation of multiple SONET/SDH frames. Whereas, LCAS allows to right size those virtual concatenated paths. Together VCAT and LCAS allow SONET/SDH to better fit busy data traffic.

Table of Content

Introduction

Ethernet Services

SONET/SDH Transport

Ethernet over SONET

Generic Framing Procedure Overview

Virtual Concatenation Overview

Link Capacity Adjustment Overview

Generic Framing Procedure

GFP Frame Structure

GFP Client Frames

GFP Client Independent Processes

GFP Client-Specific Processes

Ethernet MAC Encapsulation

Virtual Concatenation

High Order Concatenation

Low Order Concatenation

Link Capacity Adjustment Scheme

Conclusion

References

Abbreviations

Introduction

Worldwide Service Providers have significantly reduced their capital expenditures for the last three years. That was the first step for them to return to profitability. As a second step, they are presently trying to operate their network and services in a more efficient way to increase their gross margins.

In parallel to decrease their costs, Service Providers are under tremendous pressure to seek new streams of revenues. With declining wire line voices services, Service Providers are investing in the creation of new services to capture the growing demand for Internet and data connectivity.

Ethernet Services

Ethernet is the ubiquitous transport in LAN for computer networking. All enterprise data traffic starts off and ends up as Ethernet. It provides a cost-effective interface, generally inexpensive compared to other network interfaces. As a broadcast-oriented medium, Ethernet is a good match for IP. It can scale from 10 Mb/s to 10 Gb/s of bandwidth. Furthermore, it is geographically independent.

Ethernet can be the foundation for the Service Providers to develop a portfolio of new data services such as:

- LAN Interconnect or Transparent LAN services (TLS)
- Internet Access
- Ethernet Private Line
- Virtual Private LAN services (VPLS).

These primary data services can lead to other more sophisticated IP managed services for storage or security. RHK forecasted that for North America, Ethernet could be a \$4 B market for Service Providers in 2006.¹

SONET/SDH Transport Networks

Today's metro networks are built with SONET/SDH rings. SONET/SDH was designed in the early 90s for circuit-switching. It provides a digital TDM transmission hierarchy for optical networks. SONET was not designed for data services. New transport technologies such as RPR (Resilient Packet Rings) and 10 Gigabit Ethernet, IEEE 802.3ae also called Optical Ethernet, are certainly by design more data centric than SONET.

But since Service Providers have invested so much in SONET/SDH transport networks, they know how to operate and maintain them that transport networks based on RPR or Optical Ethernet will likely be deployed by new Service Providers and for green field applications by incumbent Service Providers.

¹ RHK: Optical Ethernet to bring \$4B, press release March 26, 2002.

Ethernet over SONET (EoS)

A number of significant enhancements have been made recently to better use the existing SONET/SDH transport infrastructure for data services. These include the Generic Framing Protocol (GFP), Virtual Concatenation (VCAT) and the Link Capacity Adjustment (LCAS). Ethernet over SONET/SDH by combining GFP, VCAT and LCAS provides improved bandwidth efficiently for data transport while allowing the Service Provider to operate its SONET/SDH transport network.

Before EoS, equipment vendors have used a number of proprietary encapsulation techniques to transport IP/Ethernet over SONET/SDH.

The first method has been to use ATM AAL 5 over SONET/SDH. ATM is a very efficient switching and multiplexing technology which speeds and feeds scale with SONET/SDH but requires a high overhead with the ATM “cell tax” of 5 byte header and heavy software burden because mainly of its connection-oriented capability.

Other methods have been focused mainly on using PPP. The IP traffic coming to an Ethernet port is encapsulated over a PPP link and multiple ports can be encapsulated over ML-PPP links. By using an HDLC framing, the PPP traffic is transporting over the SONET/SDH payload.

These methods have been standardized within the IETF through RFC 1662, RFC 1990 and RFC 2615. The ITU-T expanded this work by specifying the use of LAPS (very similar protocol to PPP/HDLC) and specifying IP over LAPS in X.85/Y.1321 and Ethernet over LAPS in X.86/Y.1323.

All these encapsulation mechanisms suffer from the inherent deficiencies of HDLC framing which introduces variable packet sizes because of its trailer and suffers from limited protection from corruption of flag, address etc...

Now to better optimize the transport of Ethernet and other data services over SONET, GFP has been standardized taking into account both the pros and cons of ATM and PPP/HDLC and leveraging two new emerging SONET/SDH capabilities: VCAT and LCAS.

Generic Framing Procedure (GFP) Overview

GFP defines a mapping of client data signals into SONET/SDH payloads in order to allow SONET/SDH to transport non-TDM traffic more efficiently. GFP defines two types of client signals:

- **Frame-mapped GFP** for PDU-oriented signals such as IP/ PPP or Ethernet MAC;
- **Transparent-mapped GFP** for block-oriented signals such as Fiber Channel and ESCON.

GFP provides a flexible and robust encapsulation technology that supports both fixed and variable length frame. Unlike HDLC, GFP does not use any special character for frame delineation. GFP provides a more deterministic encapsulation scheme than HDLC whose overhead is data dependant. GFP generalizes the ATM frame delineation mechanism to encapsulate variable length frames. Its frame delineation is based on the length of the current payload and an error control check.

Benefits of GFP

GFP provides two major benefits. First, it gives one uniform mechanism to transport any data type over SONET/SDH. Second, its encapsulation mechanism is superior to HDLC without the layer processing of ATM.

Virtual Concatenation (VCAT) Overview

Two approaches exist for concatenation: contiguous and virtual. Both solutions provide concatenated bandwidth of X-times Container-N at the path termination.

However, contiguous concatenation keeps the concatenated SONET payload through the whole SONET/SDH transport. Therefore, network elements must support contiguous concatenation from the source to the destination and, at every intermediate node.

Virtual Concatenation relaxes the “rigidity” of SONET/SDH payloads originally designed for TDM traffic. VCAT allows the concatenation of multiple payload frames from VT1.5 to STS-3c SPE to better scale the requirements for incremental data bandwidth.

VCAT breaks the initial SONET payload at the source into individual Virtual Containers (VC). Each VC is part logically of a Virtual Concatenated Group (VCG). Each VCG member is routed and transported individually across the SONET/SDH transport network and is recombined with the other VCs at the destination node to form the whole VCG.

Therefore, network elements must support virtual concatenation at the source and at the destination. But intermediate nodes do not need to be aware of the virtual concatenation.

VCAT can be provided in two different ways:

- **High-Order:** for **STS-M-N_v** where n indicates the number of STS-m virtually concatenated.
 - M can be equal to 1 (STS-1) or 3 (STS-3c);
 - N can vary from 0 to 255;
- **Low-Order:** for **VT-M-N_v** where n indicates the number of VT-m virtually concatenated.
 - M can be equal to 1.5, 2, 3 and 6;
 - N can vary from 0 to 64.

Benefits of VCAT

VCAT provides a much more efficient use of the transport bandwidth for data user interfaces. With VCAT, an OC-48 link can carry two full Gb Ethernet with 95% of the link used through 7 virtual STS-3c instead of one Gb Ethernet with 42% of the link used through an STS-48c.

Service	SONET Payload Without VCAT	Bandwidth Efficiency Without VCAT	SONET Payload With VCAT	Bandwidth Efficiency With VCAT
Ethernet 10 Mb/s	STS-1	20%	VT1.5-7v	89%
Fast Ethernet 100 M/s	STS-3c	67%	STS-1-2v	100%
Gigabit Ethernet 1000 Mb/s	STS-48c	42%	STS-3c-7v	95%
Fiber Channel 1000 Mb/s	STS-48c	42%	STS-3c-7v	95%

Figure 1: VCAT Efficiency Comparisons

Furthermore since VCG members are not constrained to the same path, VCAT allows a more efficient use of the capacity of the different routes in the network.

Link Capacity Adjustment Scheme (LCAS) Overview

Standardized in ITU-T G.7042/Y.1305, LCAS is a signaling protocol for sizing virtually concatenated paths. With LCAS, VCG can be resized at any time without disturbing network traffic.

LCAS signaling messages are exchanged to change the number of VC between the source and the destination of the path. The number of VC can be increased or decreased without any frames lost therefore increasing or decreasing the capacity of the VCG link.

LCAS provides as well a means of removing links that have experienced failures. The VC in fault is detected and removed automatically from the VCG.

LCAS has been designed to operate both with management systems for the set-up and release of VC but can also operate with emerging GMPLS-based control plane responsible for network path set-up and teardown.

Benefits of LCAS

The use of LCAS provides an effective way for the Service Provider to change the bandwidth allocated. Provisioning quickly the right bandwidth at any time is a major operations management goal of Service Providers.

Generic Framing Procedure (GFP)

GFP Frame Structure

A standard GFP frame format is defined as shown in Figure 1 below:

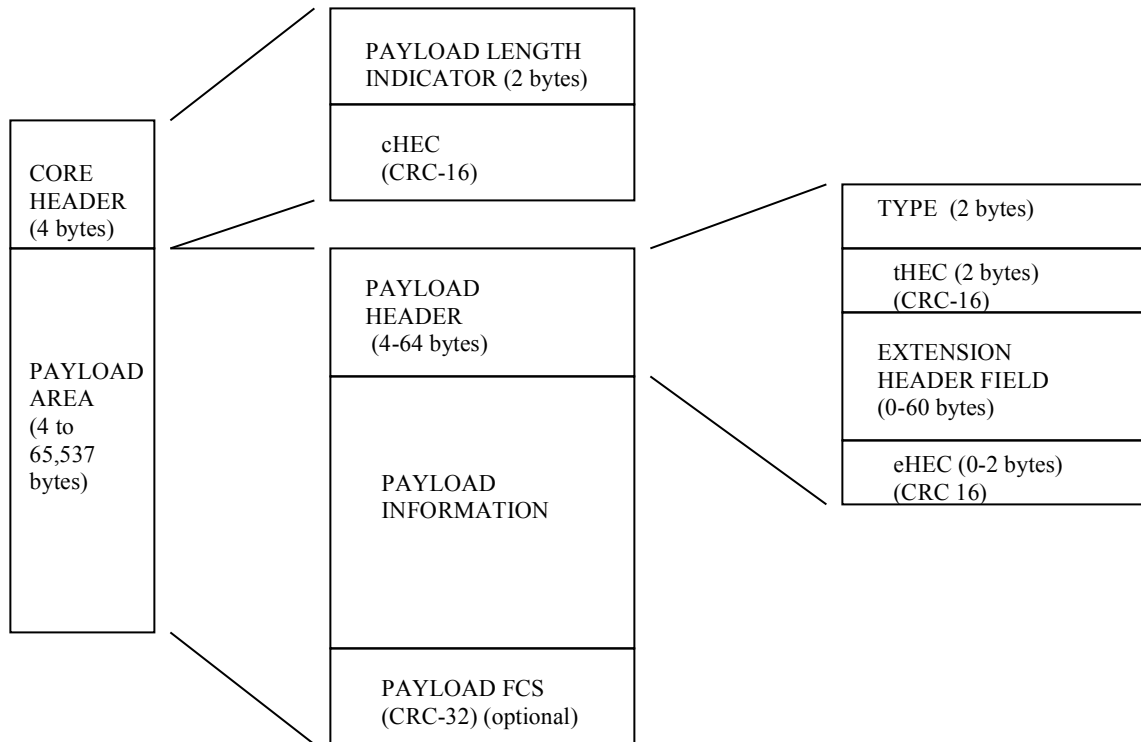


Figure 2: Frame Format for GFP Client Frames

The GFP core header supports GFP specific data link information. The core header allows GFP frame delineation independently of the content of the higher layer PDUs. It has two fields:

- **Payload Length Indicator (PLI)** (2 bytes): indicate the PDU length, that is, the number of octets in the GFP payload area;
- **Core Header Error Control (cHEC)** (2 bytes): provide a CRC-16 single bit error correction and multi-bit error detection to protect the integrity of the Core Header.

The GFP payload contains all octets after the core header. It includes three fields:

- **Payload Header (PH)** (4 to 64 bytes): support data link management procedures specific to higher layer client signals;
- **Payload Information:** contain the client signal in an octet-aligned packet stream. This field may include from 0 to 65,535-X, octets where X is the size of the payload header and the payload FCS if present.

- **Payload Frame Check Sequence (pFCS):** provide an optional CRC-32 single bit error correction and multi-bit error detection to protect the integrity of the Core Header.

The Payload Header includes the following fields:

- **TYPE:** indicate the content and format of the GFP Payload Information field. It distinguishes between GFP frame types and between different services;
- **TYPE HEC(tHEC):** provide a CRC-16 single bit error correction and multi-bit error detection to protect the integrity of the TYPE field;
- **Extension Header (EHF):** support technology specific data link headers such as virtual link identifiers, source/destination addresses, port numbers, class of service...
- **Extension HEC (eHEC):** provide a CRC-16 error control code to protect the integrity of the EHF field.

The TYPE field itself consists of several sub- fields:

- **Payload Types Identifiers (PTI)** (Table 1) (3 bits): identify the type of user frames either client data frame or client management frames;
- **Extension Header Identifier (EXI)** (4 bits) (Table 2): three kinds of extensions headers are currently defined: a null extension header, a linear extension header and a ring extension header;
- **Payload FCS Indicator (PFI)** (1 bit): indicate the presence or absence of the Payload FCS field;
- **User Payload Identifier (UPI):** identify the signal type defined in the GFP Payload Information Field (Table 3).

Payload Type Identifiers (PTI)	Usage
000	Client Data
100	Client Management
Others	Reserved

Table 1: GFP Payload Type Identifiers.

Extension Header Identifier (EXI)	Usage
0000	Null Extension Header
0001	Linear Frame
0010	Ring Frame
Others	Reserved

Table 2: GFP Extension Header Identifiers

GFP Client Frames

Two types of GFP client frames are currently defined: Client Data and Client Management.

Client Data frames are defined by PTI=000. To distinguish between the many transported signals the User Payload Identifier (UPI) is used as shown in Table 3.

PTI = 000	
User Payload Identifier	GFP Frame Payload Area
0000 0000 and 1111 1111	Reserved and not available
0000 0001	Frame Mapped Ethernet
0000 0010	Frame Mapped PPP
0000 0011	Transparent Fibre Channel
0000 0100	Transparent FICON
0000 0101	Transparent ESCON
0000 0110	Transparent Gb Ethernet
0000 0111	Reserved
0000 1000	Frame Mapped Multiple Access Protocol over SDH (MAPOS)
0000 1001 through 1110 1111	Reserved for future standardization
1111 0000 through 1111 1110	Reserved for proprietary use.

Table 3: User Payload Identifiers for GFP Client Frames

Client Management frames are defined by PTI=100. To distinguish between the many transported signals the User Payload Identifier (UPI) is used as shown in Table 4.

PTI = 100	
User Payload Identifier	GFP Frame Payload Area
0000 0000 and 1111 1111	Reserved
0000 0001	Loss of Client Signal
0000 0010	Loss of Character Synchronization
0000 0011 through 1111 1110	Reserved

Table 4: Client Management Frame User Payload Identifier.

GFP Client-Independent Processes

GFP client-independent processes involve a number of mechanisms common to all payloads: frames delineation, header/payload scrambling, frames multiplexing and client management.

Frames Delineation

GFP uses a modified version of the HEC check algorithm specified in ITU-T I.432, clause 4.5.1.1, to provide GFP frame delineation. The frame delineation algorithm used in GFP differs from that in ITU-T I.432 in two basic ways:

- The algorithm uses the PDU Length Indicator field of the GFP Core Header to find the end of the GFP frame; and
- HEC field calculation uses a 16-bit polynomial and, consequently, generates a two-octet cHEC field.

GFP frame delineation is performed based on the correlation between the first two octets of the GFP frame and the embedded cHEC field.

Header and Payload Scrambling

The core header is scrambled on transmission by an exclusive-OR operation with the hexadecimal number B6AB31E0. The core header scrambling ensures high bit transmission during idle transmission periods.

Scrambling of the GFP payload is required to provide security against payload information replicating scrambling word (or its inverse) from a frame synchronous scrambler such as those used in the SONET line layer.

Frames Multiplexing

GFP frames from multiple ports and multiple clients types are multiplexed on a frame-by-frame basis. Client data frames are always sent first over client management frames. When there are no other frames available for transmission, idle frames shall be inserted, thus providing a continuous stream of frames.

Client Management

Client management frames (CMF), presently defined are related to the propagation of client interface failure conditions, referred to as client signal fail (CSF).

CSF provides an indication to the far-end GFP client-specific sink adaptation process upon failure/degradation detection in the ingress client signal. Two types of CSF can be reported:

- Loss of client signal
- Loss of client character synchronization.

Client-Specific Processes

Ethernet	IP/PPP	Other Client Signals
GFP – Client Specific Aspects (Payload Dependent)		
GFP – Common Aspects (Payload Independent)		
SDH VC-<i>n</i> Path	Other octet-synchronous paths	OTN ODUk Path

Figure 3: Relationships Client Signals and Transport Paths

They are two types of client signals payloads:

- Frame-mapped GFP (GFP-F) for PDU oriented clients such as Ethernet MAC, IP and PPP;
- Transparent-mapped GFP (GFP-T) for block-code oriented clients such as Fiber Channel, ESCON frame is mapped entirely into one GFP frame.

Frame-mapped GFP clients consist of variable length packets. In that mode, each client frame is mapped entirely into one GFP frame.

The SONET node encapsulates the entire Layer 2 frame into the GFP frame. Frame-mapped GFP uses the basic frame structure of a GFP client, including the required payload header. The payload FCS is optional.

For transparent-mapped GFP clients, the individual characters of the signal are de mapped from the 8B/10B client block and then mapped into periodic 64B/66B fixed-length GFP frames rather than buffering an entire frame of client data into its own GFP.

Transparent-mapped GFP uses the same structure as the Frame-mapped GFP, including the required payload header. The payload FCS is optional.

Ethernet MAC Encapsulation

The source deletes the gap between the Ethernet packets, known as Inter Packet Gaps (IPGs).

And, the fields between the Destination Address (DA) to the Frame Check Sequence (FCS) are encapsulated into the GFP payload.

The destination node restores the IPGs and the Ethernet MAC is then forwarded.

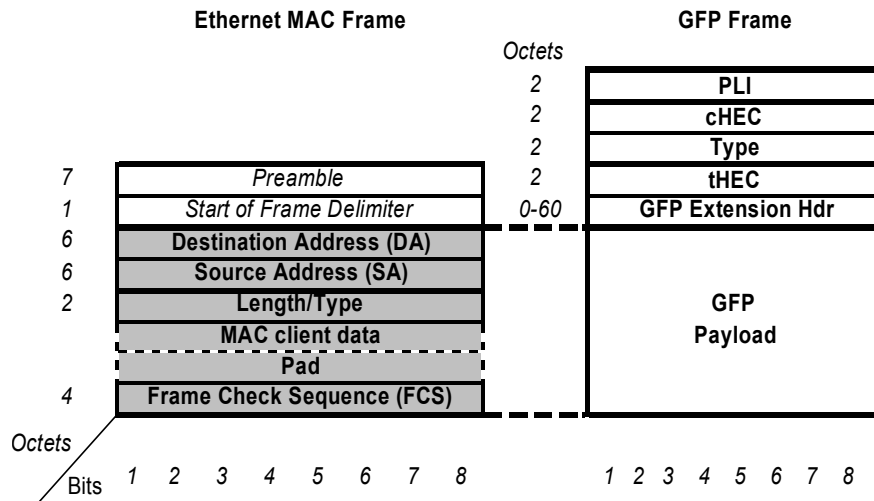


Figure 4: Ethernet MAC encapsulation into GFP

Virtual Concatenation (VCAT)

Because different VCAT members can take different routes in the network, VCAT needs to deal with two major constraints:

- propagation delays associated with individual VCs
- keeping the logical sequence of the VCs

The differential delay has to be compensated and the individual VCs have to be realigned in order to reassemble the original byte stream at the termination path.

To do so, VCAT defines two indicators the Multi-Frame Indicator (MFI) and the Sequence Indicator Field (SQ).

High Order Concatenation (HO)

For HO, MFI and SQ are communicated using the H4 byte into the SONET path overhead.

MFI is a two stages multi-frame indicator that is used to realign the payload for all VCs that may not arrive at the receiving node in the same order in which they were transmitted.

The first stage MF11 uses 4 bits, bits 5-8 of the H4 byte. MF11 is incremented every basic frame and counts from 0 to 15.

The second stage MF12 uses 8 bits, bits 1-4 of frames 0 and 1 of MF11. MF12 is incremented once every multiframe of the first stage and counts from 0 to 255.

So the total bit combination of MF11 and MF12 are: $16 \times 256 = 4,096$ which can provide up to: $4,096 \times 125 \text{ microseconds} = 512 \text{ milliseconds}$ of differential delays.

SQ contains the sequence number assigned to a specific VC member in a VCG. Each VC is assigned a unique sequence number so that it can be correctly reassembled at the terminating path.

SQ is carried in the 14th and 15th frames of the MF11 sequence. For HO, the SQ is an 8-bit field and its range is 0-255. It is carried in the H4 byte overhead.

H4 Byte								1st multi-frame number	2nd multi-frame number
Bit 1	Bit 2	Bit 3	Bit 4	Bit 5	Bit 6	Bit 7	Bit 8		
				1st multiframe indicator MFI1 (bits 1-4)					
Sequence indicator MSB (bits 1-4)				1	1	1	0	14	n-1
Sequence indicator LSB (bits 5-8)				1	1	1	1	15	
2nd multiframe indicator MFI2 MSB (bits 1-4)				0	0	0	0	0	n
2nd multiframe indicator MFI2 LSB (bits 5-8)				0	0	0	1	1	
Reserved ("0000")				0	0	1	0	2	
Reserved ("0000")				0	0	1	1	3	
Reserved ("0000")				0	1	0	0	4	
Reserved ("0000")				0	1	0	1	5	
Reserved ("0000")				0	1	1	0	6	
Reserved ("0000")				0	1	1	1	7	
Reserved ("0000")				1	0	0	0	8	
Reserved ("0000")				1	0	0	1	9	
Reserved ("0000")				1	0	1	0	10	
Reserved ("0000")				1	0	1	1	11	
Reserved ("0000")				1	1	0	0	12	
Reserved ("0000")				1	1	0	1	13	
Sequence indicator SQ MSB (bits 1-4)				1	1	1	0	14	
Sequence indicator SQ LSB (bits 5-8)				1	1	1	1	15	
2nd multiframe indicator MFI2 MSB (bits 1-4)				0	0	0	0	0	n+1
2nd multiframe indicator MFI2 MSB (bits 5-8)				0	0	0	1	1	
Reserved ("0000")				0	0	1	0	2	

Figure 5: H4 Multi-Frame Byte

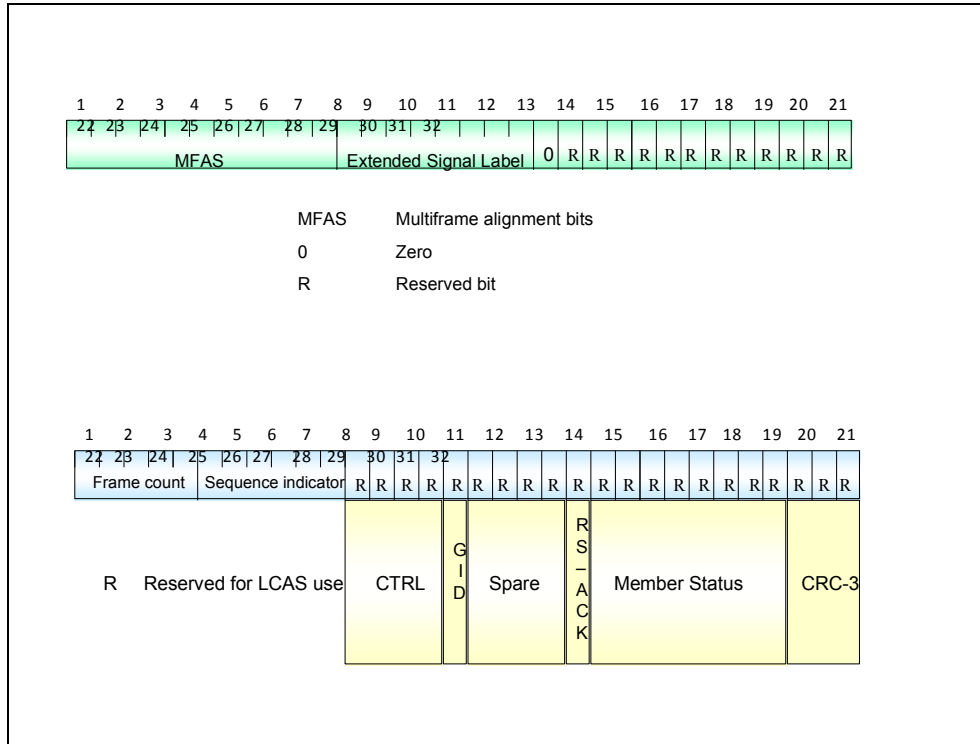


Figure 6: K4(Z7) Bits 1 and 2 Multi Frame

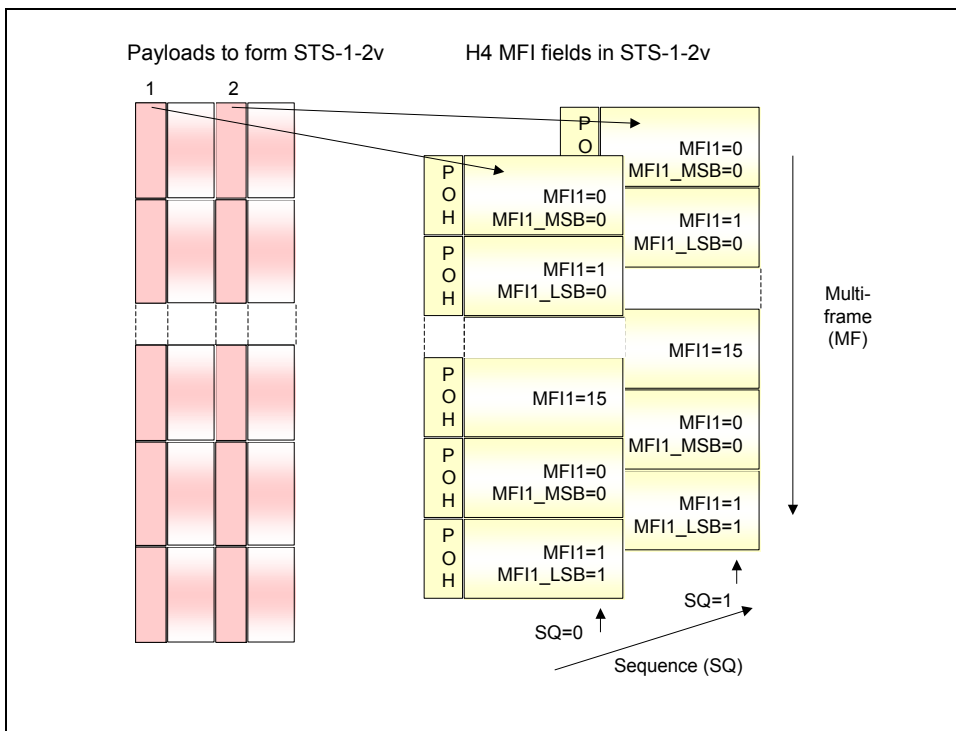


Figure 7: STS-1-Xv Multi Frame and SQ when X = 2

Low Order Concatenation (LO)

For LO, MFI is communicated using the K4 (Z7) byte path overhead, bits 1 and 2. SQ is a 6-bit field and its range is 0-64. It is carried in the bit 2 of the K4 (Z7) byte.

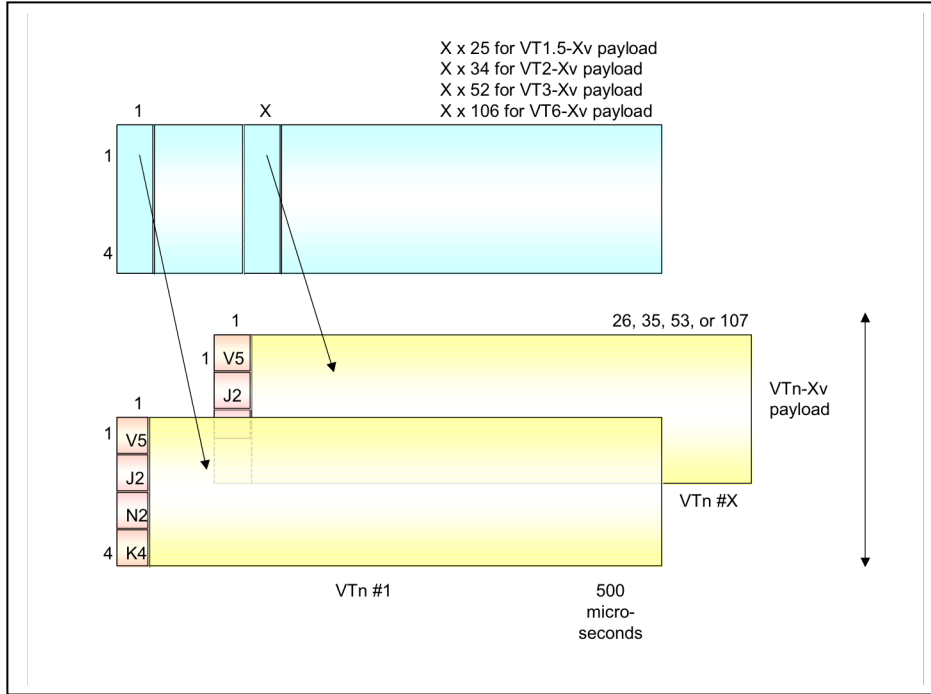


Figure 8: Low-Order Virtual Concatenation VT-m-nv

Link Capacity Adjustment Scheme (LCAS)

Virtual concatenation can be used without LCAS, but LCAS itself requires virtual concatenation. LCAS signaling is provided through control packets that enable the synchronization of changes. Control packets contain information sent from So to Sk and from Sk to So in the capacity of the transmitter (source) (So) and the receiver (sink) (Sk) that are path terminating. Control packets consist of fields dedicated to a specific function.

From So to Sk, the control packets are:

- **MFI:** as used in VCAT;
- **SQ:** as used in VCAT
- **Control Field (CTRL):** used to synchronize the Sk with the So and to provide the status of the individual member of the group.

Value msb...lsb	Command	Remarks
0000	FIXED	This is an indication that this end uses fixed bandwidth (non-LCAS mode)
0001	ADD	This member is about to be added to the group
0010	NORM	Normal transmission
0011	EOS	End of Sequence indication and Normal transmission
0101	IDLE	This member is not part of the group or about to be removed
1111	DNU	Do Not Use (the payload) the Sk side reported FAIL status

Figure 9: CTRL Words

- **Group Identification Group (GIG):** used for the identification of the VCG in order to distinguish multiple VCG on a single STS-N. The GIG provides the receiver with a means of verifying that all the arriving members originated from one transmitter;

From Sk to So, the control packets are:

- **Member Status Field (MSF):** used to provide information about the status of all members of the same VCG: OK or fail;
- **Re-Sequence Acknowledge Bit (RS-Ack):** used to send from So to Sk to indicate that the changes initiated by So were accepted and that So can begin accepting the new member status information.

Used in both directions:

- **CRC:** used to protect each control packet.

For HO, LCAS messages are carried in bits 1-4 of the H4 byte. For LO, they are communicated in bit 2 of the Z7 byte.

Conclusion

By combining GFP, VCAT and LCAS, Service Providers have a more efficient way to optimize their SONET/SDH transport network for Ethernet services. GFP, VCAT and LCAS do not require end-to-end upgrades to the existing SONET/SDH network. EoS network elements using GFP, VCAT and LCAS can be deployed at the ingress and the egress of the Service Provider's transport network.

References

ITU-T G.707/Y.1322: Network Node Interface for the Synchronous Digital Hierarchy (SDH).

ITU-T G 7041/Y1303: Generic Framing Procedure (GFP).

ITU-T G.7042/Y.1305: Link Capacity Adjustment Scheme (LCAS) for virtual concatenated signals.

Abréviations

ATM	Asynchronous Transfer Mode
cHEC	Core HEC
CID	Channel ID
CRC	Cyclic Redundancy Check
CSF	Client Signal Fail
CTRL	Control word sent from source to sink
DE	Discard Eligibility
DNU	Do Not Use
DP	Destination Port
DST	Destination
eHEC	Extension HEC
EOF	End of Frame
EOS	End of Sequence
ESCON	Enterprise Systems Connection
EXI	Extension Header Identifier
FC	Fiber Channel
FCS	Frame Check Sequence
FICON	Fiber Connection
GFP	Generic Framing Procedure
GFP-F	Frame mapped GFP
GFP-T	Transparent GFP
GID	Group Identification
HDLC	High-level Data Link Control
HEC	Header Error Check
IFG	Inter-Frame Gap
IPG	Inter-Packet Gap
LCAS	Link Capacity Adjustment Scheme
LCC	Last Control Character

LSB	Least Significant Bit
LOL	Loss of Light
LOS	Loss of Signal
LOM	Loss of Multiframe
MAC	Media Access Control
MFI	Multiframe Indicator
MSB	Most Significant Bit
MST	Member Status
NORM	Normal Operating Mode
PDU	Protocol Data Unit
PFI	Payload FCS Identifier
PLI	Payload Length Indicator
PTI	Payload Type Identifier
PPP	Point-to-Point Protocol
RD	Running Disparity
RS-Ack	Re-sequence acknowledge
Sk	Sink
So	Source
SQ	Sequence Indicator
SBCON	Single-Byte Command Code Sets Connection
SDH	Synchronous Digital Hierarchy
SSF	Server Signal Failure
SOF	Start of Frame
SONET	Synchronous Optical Network
SP	Source Port
SPE	Synchronous Payload Envelop
SRC	Source
STS	Synchronous Transport Signal
tHEC	Type HEC
TSD	Trail Signal Degraded
TSF	Trail Signal Fail
TTL	Time-to-Live
UPI	User Payload Identifier
VC	Virtual Container
VCG	Virtual Concatenation Group