F5 and the Growing Role of GTP for Traffic Shaping, Network Slicing, IoT, and Security

For years, F5 BIG-IP solutions have managed GPRS Tunneling Protocol (GTP) traffic. Now, in part due to EU regulations, GTP traffic growth is much higher than ever before. International trends show that mobile data is increasing, as is the demand for smart and secure GTP traffic handling. In this document, we will explain GTP, where it is typically used, and where BIG-IP solutions provide significant value.

What is GTP?

GTP is a group of IP-based communication protocols used to carry General Packet Radio Service (GPRS) within GSM, UMTS, and LTE networks. In 3GPP architectures, GTP and Proxy Mobile IPv6-based interfaces are specified on various interface points.

GTP in GPRS (2.5G and 3G) Networks

GTP is a set of three separate protocols: GTP Control (GTP-C), GTP User (GTP-U), and GTP Prime (GTP').

GTP-C is used within the GPRS core network for signaling between gateway GPRS support nodes (GGSN) and serving GPRS support nodes (SGSN), as demonstrated by the Gn interface on the diagram below. GTP-C allows the SGSN to activate and deactivate a user's session, adjust quality-of-service parameters, or update sessions when subscribers arrive from another SGSN.

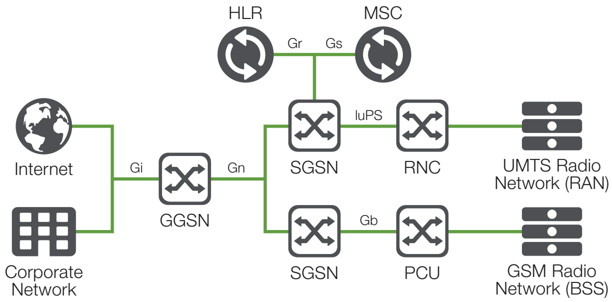


Figure 1: GTP in GPRS (2.5G and 3G networks).

GTP-U carries user data within the GPRS core network, and between the radio access network and core network. The user data is transported in IPv4, IPv6, or PPP formats.

GTP' uses the same message structure as GTP-C and GTP-U, but carries charging data from the charging data function (CDF) of the GSM or UMTS network to the charging gateway function (CGF). In most cases, it carries charging data from many individual network elements such as the GGSN to a centralized computer that delivers the charging data more conveniently to the network operator's billing center.

Note: this report covers GTP' for completeness; F5 does not offer specific solutions for managing GTP'.

Different GTP variants are implemented by RNCs, SGSNs, GGSNs, and CGFs within 3GPP networks. GPRS mobile stations (MSs) are connected to an SGSN without being aware of GTP.

GTP version 1 is used only on UDP transport. GTP version 2 can be used with UDP or TCP, with UDP either recommended or mandatory. GTP version 0 still technically exists, but it's not clear who is using it; very few operators request 100% backward compatibility for GTP v0.

GTP beyond GPRS

GTP was originally used in GPRS (2.5G networks), later developing a similar role in 3G and 4G networks. For 4G, the key nodes have different names and, to a certain extent, are comparable to nodes used in 3G networks. So, Serving Gateway (SGW) compares to the SGSN and the Packet Gateway (PGW) compares to the GGSN. The diagram below describes a basic 4G packet core network where interface S5 (or S8 for roaming, when SGW and PGW are in different networks) is the main GTP-based interface between SGW and PGW. S11 is the GTP-based interface between SGW and the Mobility Management Entity (MME). The use of GTP is not limited to S5, S8, and S11; however, further details (like GTP interfaces for VoWifi) are purposely left out.

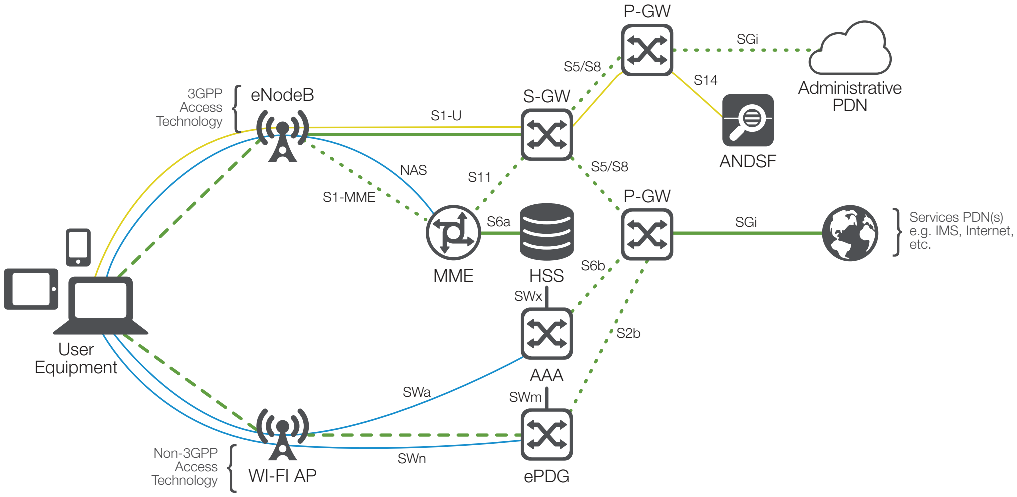


Figure 2: Evolved packet core

GTP protocol stack

The diagram below shows how the protocol stacks look for GTP-C and GTP-U. Both are relying on UDP, but note the end-user communication between, say, an application on a handset and a server on the network side travel well over TCP. As such, a web-browsing session from a mobile device is typically TCP based.

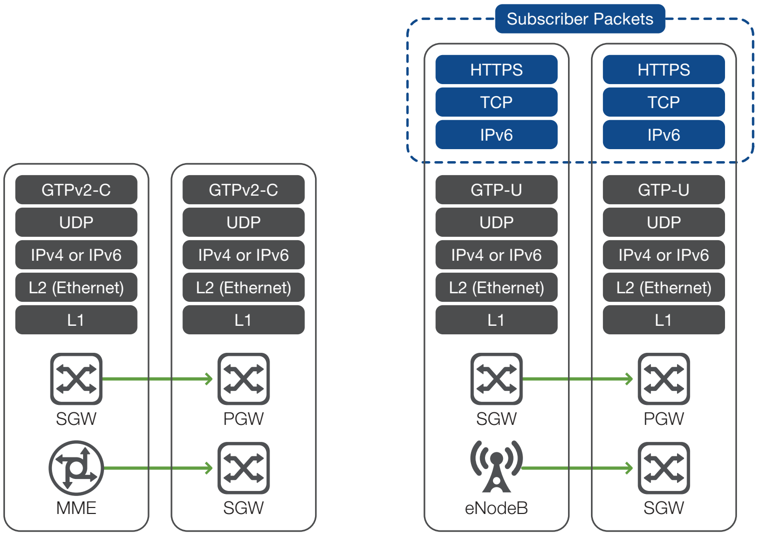
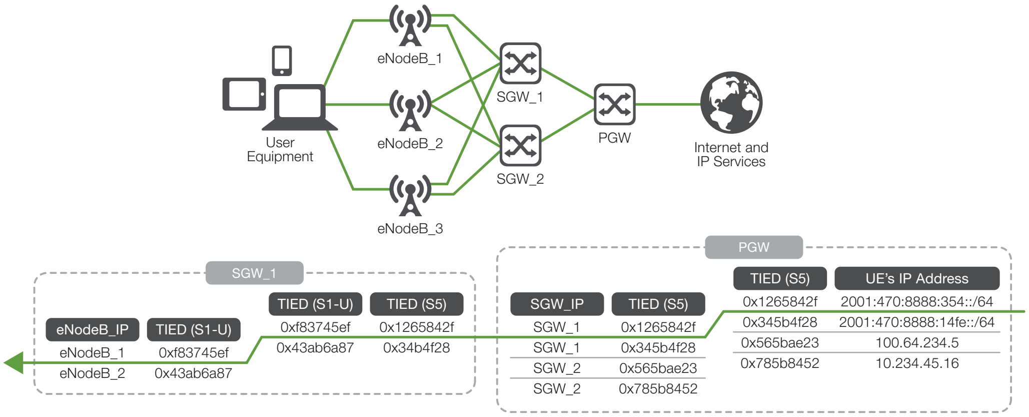


Figure 3: GTP protocol stack

GTP Tunnel Endpoint Identifier

In GTP, the Tunnel Endpoint Identifier (TEID) used to transmit GTP-U data is first signaled via GTP-C. TEID is a 32-bit field used to multiplex different connections in the same GTP tunnel, and serves as a fundamental aspect in allowing a fixed point of attachment (IP Address) to move around a large geographic area.



SGW_1

eNodeB_IP	TIED (S1-U)
eNodeB_1	0xf83745ef
eNodeB_2	0x43ab6a87

TIED (S1-U)	TIED (S5)
0xf83745ef	0x1265842f
0x43ab6a87	0x34b4f28

PGW

SGW_IP	TIED (S5)
SGW_1	0x1265842f
SGW_1	0x345b4f28
SGW_2	0x565bae23
SGW_2	0x785b8452

TIED (S5)	UE's IP Address
0x1265842f	2001:470:8888:354::/64
0x345b4f28	2001:470:8888:14fe::/64
0x565bae23	100.64.234.5
0x785b8452	10.234.45.16

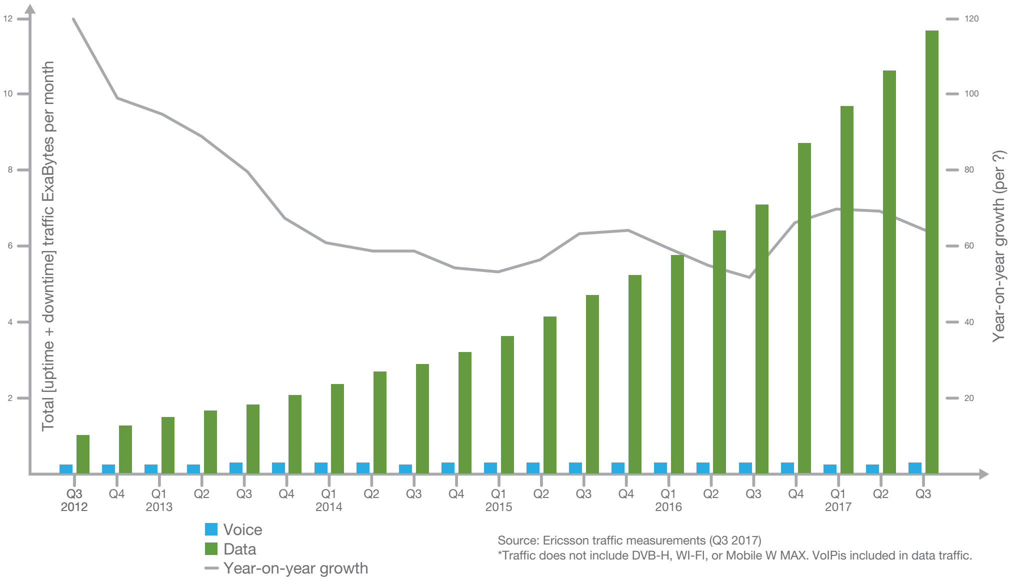
Figure 4: GTP tunnel endpoint identifier (TEID)

Why is GTP Management Required?

There are multiple cases that require GTP traffic management; BIG-IP solutions can play an important role by offering GTP management and GTP security functions.

GTP load balancing

As data over mobile networks continues to grow (see the graph below), more network nodes are needed to handle it. As explained earlier, the majority of the mobile data traffic is transported over GTP, so the number of nodes handling GTP are expanding to deal with this increased traffic. Load balancing comes in to play to help operators scale their network and its traffic handling.



Source: Ericsson traffic measurements (Q3 2017)
*Traffic does not include DVB-H, WI-FI, or Mobile W MAX. VoIPis included in data traffic.

Figure 5: Data over mobile networks continues to grow.

GTP security

Securing GTP traffic is a fundamental requirement for mobile operators. This is especially true for interfaces exposed to other networks, like for roaming, Gp interface for 3G networks, and S8 interface for 4G networks. For the last two or three years, the GSM Association (GSMA) has paid very close attention to GTP Security. The GSMA Fraud and Security Group released an FS.20 document, GTP Security, to serve as a guide for its members.

Common GTP security issues include confidential data disclosures, denial of service, network overloads, and a range of fraud activities. While security strategies vary from case to case, all organizations should implement solutions that provide full traffic visibility and comprehensive distributed denial-of-service (DDoS) protection.

Smart routing

Another key challenge is to effectively route and distinguish between GTP traffic. Subscriber traffic on the home network is different from MVNO traffic and from IoT (sliced) traffic. Charting a course for GTP traffic is usually based on the content of GTP messages and other aspects, like source and destination. A smart GTP routing function selects the PGW or network slice best suited to a specific service.

Existing technology is capable of harnessing advanced routing, proxy, and security functionalities while being able to access GTP Information Elements. For example, BIG-IP devices can tap into over 100 types, including APN, IP address, MS-ISDN, RAT type, PDN Type (v4/v6), user location info, aggregate max bit rate, and quality of service. If a terminal wants web access, it can select the Internet APN, which is conveyed via GTP. The network can select the right PGW to route the traffic to the Internet. It is also possible to support various GTP proxy cases such as routing MVNO and a service provider's own traffic to different destinations using the same APN. While doing this, GTP information can be modified to better suit its specific purpose (e.g., overriding the default APN a specific type of handset selects if it will cause problems for an MVNO who wants to manage its traffic based on the APN that a customer selected).

GTP and 5G

GTP is staying with us, and it is vital that service providers grasp how to optimize its capabilities, both now and as an important protocol in the new 5G Core Network.

While the 5G architecture was defined in Release 15 of the 3GPP specifications, Release 14 defined Control and User Plane Separation (CUPS) of EPC nodes. CUPS provides the architecture enhancements for the separation of functionality in the Evolved Packet Core's SGW, PGW, and TDF. This enables flexible network deployment and operation by distributed or centralized deployment and the independent scaling between control plane and user plane functions, without affecting the functionality of the existing nodes subject to this split.

See the diagram below for 3GPP 23.214 architecture enhancements for control and user plane separation of EPC nodes and stage 3 specification 29.244 interface between the control plane and the user plane of EPC nodes.

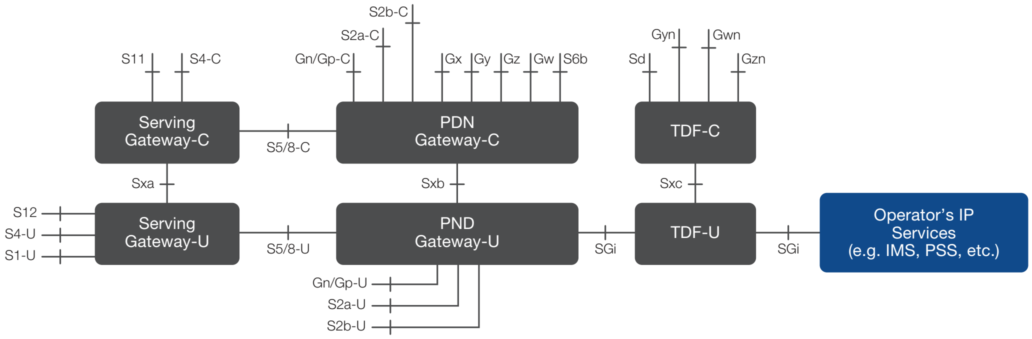


Figure 6: 3GPP 23.214 architecture enhancements.

GTP signaling between control plane nodes (indicated in the diagram as '-C') and user plane nodes (indicated in the diagram as '-U') stays untouched, but between the control and user planes, a series of new Diameter interfaces (Sxa, Sxb and Sxc) has been defined to allow communication between these different nodes, or at least different logical entities.

GTP is still being used in 5G Core Networks with the N9 interface as the reference point between two core user plane functions (UPFs). However, there is another GTP-like protocol proposed for other interfaces, such as N4, the reference point between the session management function (SMF) and the UPF. The 3GPP 29.891 v110 uses Packet Forwarding Control Protocol (PFCP), a GTP-U-based protocol. Using PFCP and enhancing 3GPP TS 29.244 to support the N4 interface is optimal.

Conclusion

GTP is a vital protocol for signaling and transporting mobile data, whether in the initial GRPS networks via 3G and 4G, or the developing 5G network. Strong mobile user data growth requires better scaling of the networks while operators respond to the GSMA-driven initiative to be more aware of the GTP's security risks and better able to respond to vulnerabilities.

To learn more about Service Provider signaling solutions, visit f5.com.