

White Paper

Evolved Packet Core & Policy Management for LTE

Prepared by

Gabriel Brown
Senior Analyst, *Heavy Reading*



www.heavyreading.com

On behalf of



www.cisco.com

July 2010

TABLE OF CONTENTS

I.	EVOLVED PACKET CORE & SERVICE INNOVATION	3
1.1	Functions of the Mobile Packet Core	3
1.2	Evolved Packet Core Deployment Options	4
1.3	EPS Bearers & QoS Model	5
1.4	Policy & Charging Control	6
II.	BACKGROUND TO THIS PAPER	8
2.1	Original Research.....	8
2.2	About the Author	8
2.3	About <i>Heavy Reading</i>	8

LIST OF FIGURES

Figure 1	Deployment Options for Evolved Packet Core	4
Figure 2	EPS Bearer Model.....	5
Figure 3	Policy & Charging Control Architecture	6
Figure 4	Importance of Policy Management Features to LTE	7

I. Evolved Packet Core & Service Innovation

Proto-4G Long Term Evolution (LTE) technology offers tremendous opportunities for mobile service innovation and business model evolution. The Evolved Packet Core (EPC) is at the heart of this new network and is essential to an operator's ability to support a diverse application mix on an end-to-end packet infrastructure. This paper examines EPC from the perspective of 2G/3G operators planning the migration to LTE. It covers deployment options, the quality-of-service (QoS) model, and associated policy and charging control functions.

There are clearly promising mechanisms in the LTE core network that will allow for differentiated and personalized service offerings, and which place operators at the center of the wider mobile services ecosystem. The challenge is one of timing: The technical and commercial complexities of fine-grained, dynamic service control are such that the ability to phase-in functionality as technology matures and market requirements evolve should inform an operator's EPC strategy.

1.1 Functions of the Mobile Packet Core

The mobile packet core is what makes a wireless access network a truly mobile network. It performs functions fundamental to the delivery of mobile data services and provides connectivity to the Internet and the carrier's own services environment. These attributes make the packet core influential to services innovation and a cornerstone to operator monetization strategies.

The EPC was defined in 3GPP Release 8, ostensibly to support LTE services; but as the name suggests, it builds on previous generations of mobile packet core. EPC is designed, ultimately, to also incorporate 3G, 2G, and non-3GPP accesses. Key functions of the mobile packet core, common across the generations, include:

- **Mobility Management.** The ability to track users as they move between cell sites (a.k.a. across routing areas or tracking areas) and route traffic accordingly. In EPC this functionality, potentially, also has application to non-cellular access, such as WiFi.
- **Session Management.** Establishing bearer setups and managing the "information flow" of a particular service or application is a primary function of the packet core. This "session layer" is critical to delivering differentiated service quality.
- **Security & Privacy.** Authentication, encryption, and user privacy are primary functions of the core network. Insofar as the operator positions itself as a trusted provider, these are commercially valuable service attributes above and beyond their primary functions.
- **Policy & Charging.** The mobile packet core has always had a role in charging for usage and content. This is extended significantly in LTE, with sophisticated policy management inherent to the architecture. Policy is tightly linked to session management.

All the above functions are instantiated in various network elements. However, the "classic" packet core architecture, which was defined for 2G data services and prevails in 3G to this day, changes significantly in EPC.

LTE is now formally called Evolved Packet System (EPS). The most obvious difference from 2G/3G is the elimination of the circuit-switch domain to create an end-to-end packet network. This puts new and demanding requirements on the packet core, which must now support real-time, delay-sensitive services in addition to best-effort data. The formal separation of control and bearer planes in the EPS system architecture, and the removal of the RNC node, to create a "flatter" network are two other major differences from the legacy 2G/3G packet core.

The network elements that make up the EPC are the Mobility Management Entity (MME) and Policy and Charging Rules Function (PCRF) in the control plane; and the Serving Gateway (S-GW) and Packet Data Network Gateway (P-GW) in the bearer plane. S-GW and P-GW nodes require

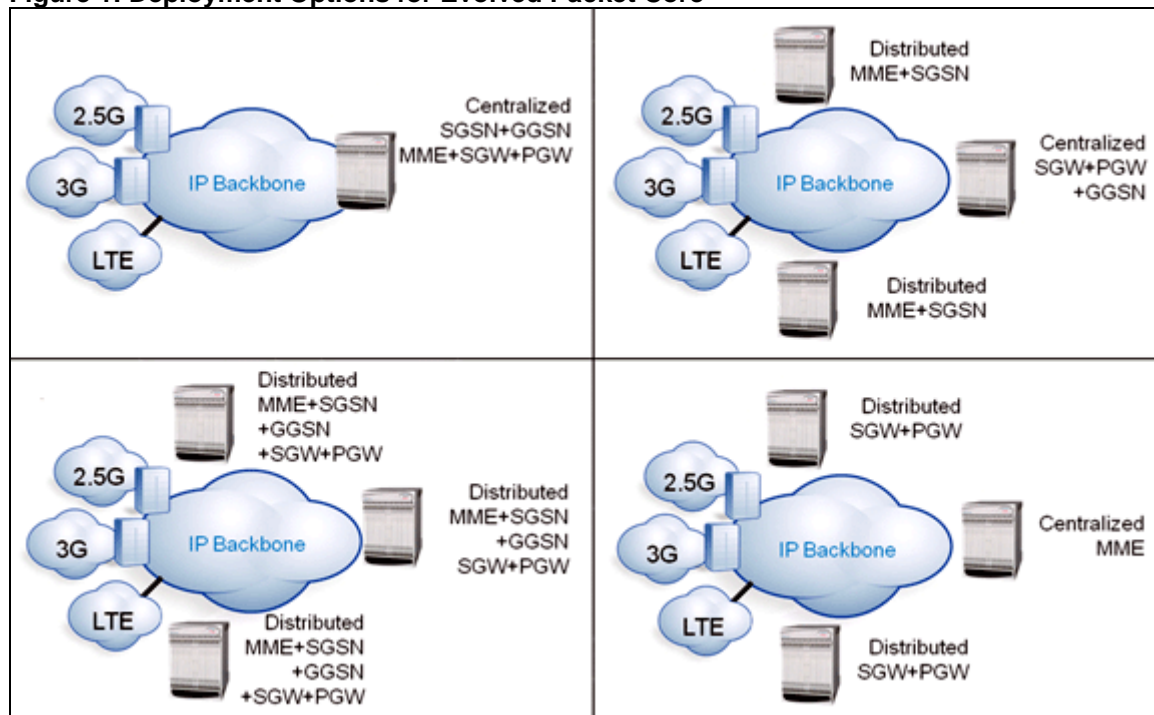
significant transaction processing capacity to support signaling, QoS, bearer setups, and so on, and not simply throughput capacity. This has implications for the choice of hardware, and in some cases vendors have elected to integrate applications on a common platform.

EPC elements interact with, and are part of, the wider policy architecture, which makes decisions about how to allocate network resources. Policy handles information provided by the lower-layer packet core nodes and acts on it according to operator preferences. It is fundamental to network management and is an important tool to help operators evolve their service portfolios.

1.2 Evolved Packet Core Deployment Options

EPC will be introduced to support the launch of LTE. There are a number of ways in which the various logical network elements can be instantiated in equipment and deployed in the network. **Figure 1** below shows some potential options.

Figure 1: Deployment Options for Evolved Packet Core



Source: Cisco

Emerging orthodoxy, according to *Heavy Reading* operator surveys, is that operators will initially deploy EPC elements in centralized locations, as shown in the top left of **Figure 1** – for example, PDN gateways may be collocated at a GGSN site. Then over time, as subscribers and traffic grow and applications become more demanding, many operators anticipate migrating to a distributed bearer plane/centralized control plane model, as shown in the bottom right.

Naturally there will be variations on this model. In geographically large networks, a more distributed model may be needed from the outset to address latency challenges, while in other networks, specific requirements and preferences will drive different outcomes.

In early deployments, EPC is being introduced as an overlay to the existing packet core to lower the risk of introducing LTE and minimize 3G service disruption. It's clear, however, that a majority of operators intend to migrate to a common core that supports LTE, 3G, and 2G access in one or two years following the initial EPC deployment. Some operators will even target a common core

from the start of LTE operations. This will reduce operational costs and make it simpler to deliver a common set of services across 3G and LTE access networks.

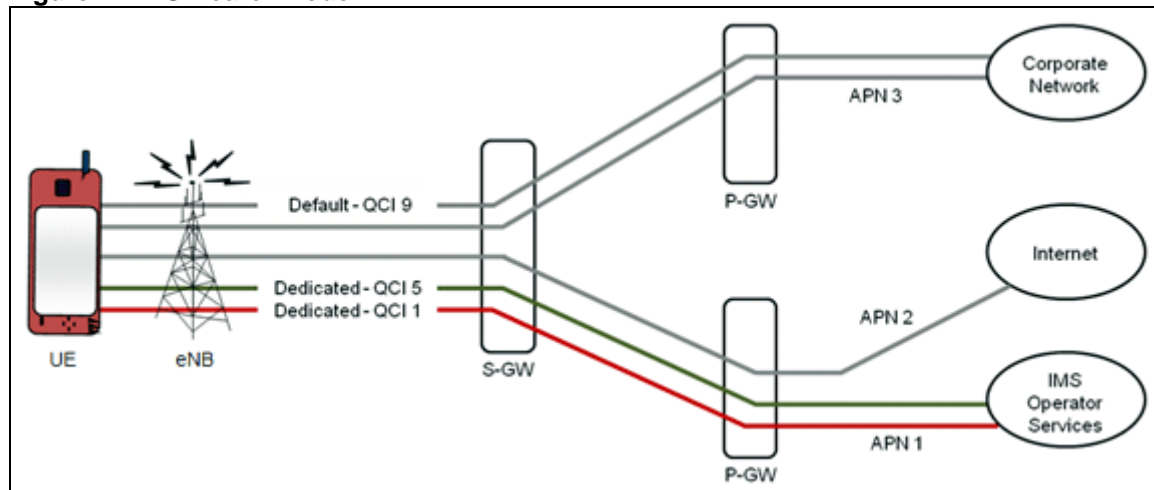
This objective of a common mobile packet core will inform EPC deployment decisions. One likely scenario is that operators will cap investment in the "legacy" core and start migrating 3G traffic to the EPC-based common core as they require capacity expansion. How this process occurs is heavily dependent on the age of the operator's legacy equipment and the vendor's capabilities. In some cases the same hardware platforms are used for 3G and LTE, and it may be that 3G packet core equipment can be upgraded in software to support LTE in certain cases.

1.3 EPS Bearers & QoS Model

Because of the requirement to offer services with different performance attributes, 3GPP has defined an extensive "bearer model" for EPS, such that services can be allocated a specific EPS bearer type. Each EPS bearer is assigned one of nine QoS Class Identifiers (QCIs) that define bit rate, packet loss, and delay.

End-user services (a.k.a. Service Data Flows) are bound to specific types of bearer according to the operator's network management policies and the subscriber's entitlement. It is worth noting that S-GW and P-GW are logical elements that can be combined for some traffic (APNs) and separated for others. The basic concept is illustrated in the diagram below:

Figure 2: EPS Bearer Model



Source: Heavy Reading, after Punz 2010

EPS bearers assigned to voice are shown in red. Described as "conversational voice," the bearers are assigned a QCI of 1, which means a dedicated bit rate (to be determined), 100ms delay, 10^{-2} packet loss, and Priority 2 in the overall bearer model. Signaling bearers, shown in green, are assigned QCI 5, with 100ms packet delay, 10^{-6} packet loss, and Priority 1. Best-effort Internet, meanwhile, is assigned to default bearers, shown in grey, using QCI 9, with 300ms latency and non-guaranteed bit rates.

There are nine different QCI classes specified in EPS. Most operators are likely to begin with these first three service classes (Signaling, Voice, Data), since this is relatively straightforward to implement. Over time there is potential to add other bearers (e.g., for premium video service or machine-to-machine applications) once operational processes are better understood.

Mechanisms exist in EPS to drop or downgrade lower-priority bearers in the event that the "pipe" is full and a specific user's service, such as voice or premium video, is granted priority access. The Allocation and Retention Priority (ARP) function provides input into the decision about

whether a bearer establishment/modification request should be accepted or rejected in case of resource limitation.

The more service classes are introduced, the more complex this decision becomes, and the more important the design of the scheduler in the eNodeB (LTE base station) is to the overall performance of the system. The eNodeB scheduler is responsible for allocating resources over the most constrained link in the network – the radio interface. This is challenging because what might be optimal from a radio efficiency point of view may not align with the subscriber's entitlement or operator's policy. For example, a user may be entitled to request a premium video stream, but if that user is poorly located in the cell, providing the service may unduly degrade cell throughput.

For this and other reasons, the industry is still quite some way from being able to implement sophisticated QoS schemes commercially. The basic voice and data model is the focus initially, and even this is only just starting to be stress-tested in trial networks. It will be some time before more dynamic QoS models are introduced, and it is likely that only a subset of what is theoretically defined by 3GPP will ever be implemented. Nevertheless, this inherent capability of the EPC to manage sessions (information flows, essentially) is the basis for ensuring network performance and evolving the mobile data business model.

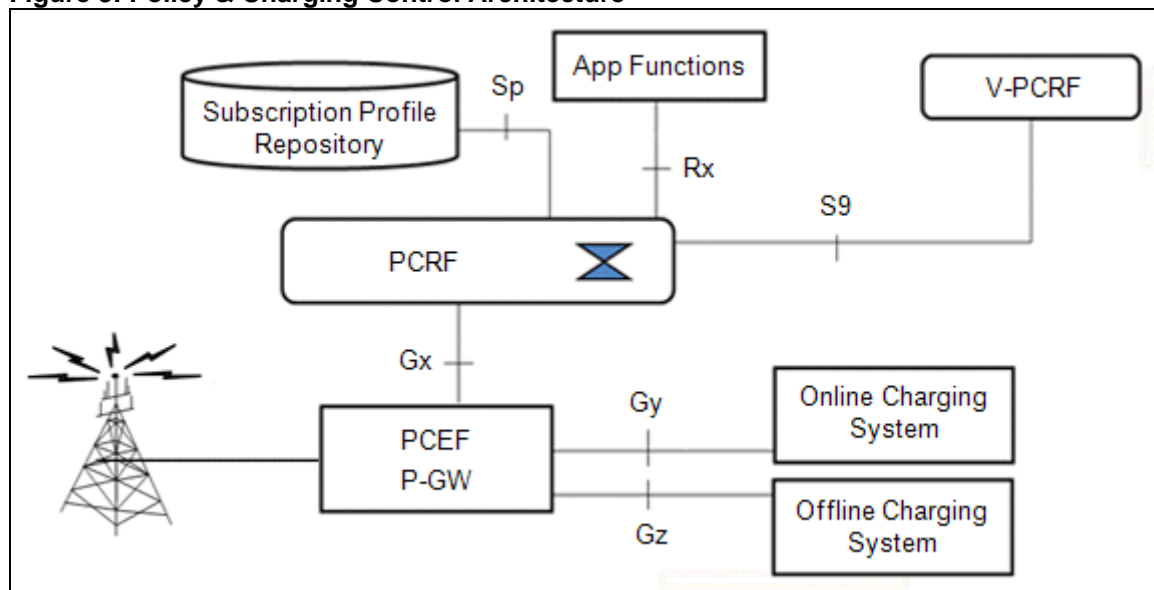
1.4 Policy & Charging Control

On top of the "session layer" (EPS bearers) LTE can make use of an extensive policy management architecture that provides operators with fine-grained control over users and services. This is integrated, via standardized interfaces, to online and offline charging systems, and therefore offers opportunities for monetization.

This capability can be used by operators to develop different service types. This starts with the tiered service plans common today, and evolves into more sophisticated service models, such as non-user-paid services or other "upstream" business models, where connectivity is bundled with a third-party service or content. For example, a video download could be zero-rated against the subscriber's data cap because the cost of transit is encapsulated in the content fee.

The policy architecture is shown below. At a basic level, the Policy Enforcement Function (PCEF) interacts with the Policy Server (PCRF) to provide a service class to the subscriber.

Figure 3: Policy & Charging Control Architecture

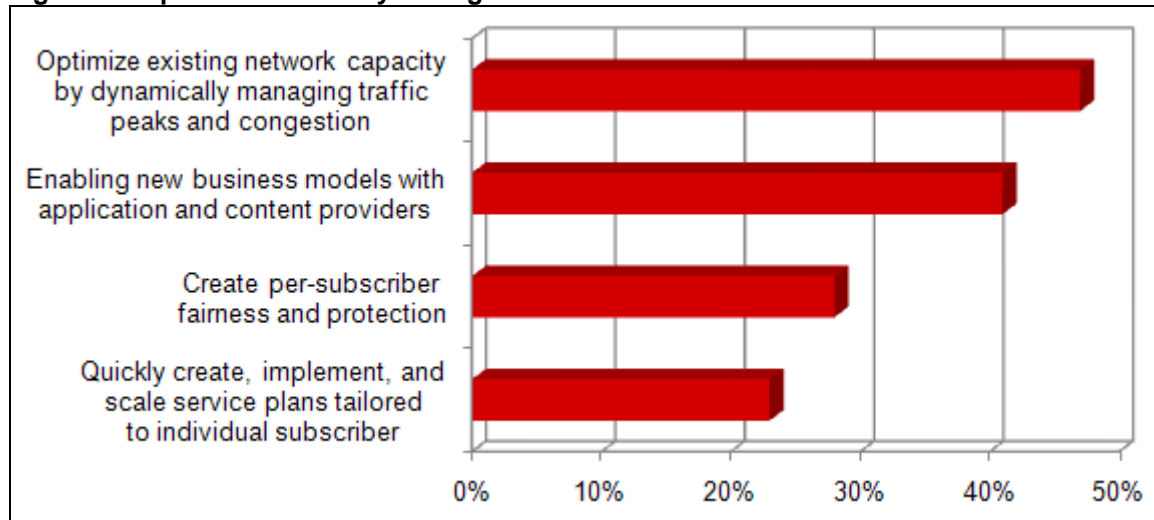


The PCEF is located in the P-GW, which is the device that connects the mobile network to external packet networks. This makes the P-GW the logical element in which to perform traffic management functions such as deep packet inspection, and makes it a strategic network element. Even where services are running on default EPS bearers (e.g., best-effort Internet) the P-GW is able to process and manage that traffic. So, for example, it could count or "shape" traffic and share information with the higher-layer policy and charging systems.

Realizing the potential of policy is far from straightforward. The complexity of system integration, itself a considerable challenge, is multiplied when one factors in the commercial and marketing aspects of creating end-user services. For this reason, operators will take a phased approach to introducing policy into their networks, starting with congestion management (e.g., of peer-to-peer traffic) and then moving to more sophisticated per-subscriber policies later.

This view is largely supported by *Heavy Reading's* "2010 LTE/SAE Mobile Operator Survey." The chart below (**Figure 4**) shows the results of a question that asked respondents to rate the importance of policy management features to their LTE deployments.

Figure 4: Importance of Policy Management Features to LTE



* Chart shows responses ranked "5: Very Important," on a scale of 1 to 5; n=102

A primary finding is that the two options that were closer to network-level or wholesale-type uses of policy scored significantly higher than the use of policy to develop subscriber services at an individual level. This indicates that operators:

- Place more importance on the ability to manage network capacity at an aggregate level – to "optimize existing network capacity by dynamically managing traffic peaks and congestion" – than on per-subscriber management.
- Take a pragmatic view that reflects the complexity of per-subscriber uses of policy and recognizes that to "quickly create service plans tailored to the individual" is challenging at present and therefore not an immediate priority.

In summary, EPC and policy management will underpin and contribute to service innovation in next-generation 4G/LTE networks. For operators, deployment choices will be informed by the desire to move to a common core for 2G/3G and LTE access, and by the flexibility they need to adopt dynamic, sophisticated policy and QoS mechanisms as technology and market requirements evolve.

II. Background to This Paper

2.1 Original Research

This *Heavy Reading* White Paper was commissioned by Cisco, but is based on independent research. The research and opinions expressed in the report are those of *Heavy Reading*.

2.2 About the Author

GABRIEL BROWN
SENIOR ANALYST, *HEAVY READING*

Brown's coverage at *Heavy Reading* focuses on wireless data networking technologies, including WLAN, 3G/HSPA, WiMax, and LTE, with reference to how these technologies impact the wider mobile data services market. Brown has covered the wireless data industry since 1998. Before moving to *Heavy Reading*, Brown was Chief Analyst of the monthly *Insider Research Services*, published by *Heavy Reading's* parent company *Light Reading*.

Prior to joining *Light Reading Communications Group*, Brown was the editor of *IP Wireline and Wireless Week* at London's Euromoney Institutional Investor. He often presents research findings at industry events and is regularly consulted by wireless networking technology leaders. Brown is based in the U.K. and can be reached at Brown@HeavyReading.com.

2.3 About Heavy Reading

Heavy Reading (www.heavyreading.com) is an independent market research organization offering quantitative and qualitative analysis of telecom technology to service providers, technology suppliers, and investors. Its mandate is to provide comprehensive competitive analysis needed for the deployment of profitable networks based on next-generation hardware and software. *Heavy Reading* produces nearly 100 research reports per year, including industry-leading technology assessments, market tracker reports for emerging technologies, and concise research reports focusing on the telecom industry's most dynamic vertical market sectors.

Heavy Reading offers a wide range of custom/consulting services aimed at identifying market and revenue opportunities for telecom industry clients. These services include in-depth product and marketing strategy assessments, independent surveys to assess and validate demand and spending trends for new products and services, and consultations on specific product and go-to-market strategies.

Heavy Reading's network of research resources also includes Pyramid Research (www.pyr.com), a leading provider of research, data, and custom/consulting services covering emerging market and service opportunities. Pyramid offers in-depth forecast and market performance analysis for more than 100 countries and is uniquely positioned across emerging markets, emerging technologies, and emerging business models. Together, the *Heavy Reading* and Pyramid Research team includes more than 50 industry-leading analysts tracking global telecom market, technology, and service trends.

Heavy Reading
11 W. 19th Street, 3rd Floor
New York, NY 10011
United States of America
Phone: +1 212-600-3000
www.heavyreading.com