ABSTRACT
IP Flow Mobility (IFOM) is a feature defined in the 3GPP standards. In this paper we propose a Converged Gateway (CGW) which enables IFOM to perform seamless offload from licensed spectrum to unlicensed spectrum. The IFOM functionality within the CGW takes into account the current conditions, rules provisioned from the Evolved Packet Core (EPC) network and local administrator rules that are merged with the rules provisioned from the core network. This functionality will manage IP Flows from their inception until their end. It will actively decide the optimal link to use for the delivery of data to an end-user device. Routing rules are used to decide the initial transport assignment of an IP Flow. Then once an IP Flow has been identified by Deep Packet Inspection (DPI), it will be placed on the transport according to its type and the rule for that type. Periodically, the CGW will perform load balancing to most effectively utilize the links. We then show a few examples based on a set of rules and flow types.

KEY WORDS
LTE; WiFi AP; seamless offload; IP Flow Mobility.

1. Introduction

WiFi and cellular technologies are nearly ubiquitous, available almost anytime and almost anywhere. 3GPP standards define IP Flow Mobility (IFOM) that is to allow cellular operators to offload certain traffic over WiFi in a non-seamless manner. IFOM benefits the operators and users as it allows for offloading some traffic from licensed spectrum, cellular, to unlicensed spectrum, WiFi. There are, however, some limitations to the currently defined IFOM that prevents it from being an ideal solution. These limitations, and the benefits provided by a CGW are explored in this paper.

Specifically, this paper describes the architecture of the CGW and the functionality within the CGW. It then describes the interaction between an eNode B and the EPC and how they are affected by the CGW. After this

Figure 1. Current LTE EPC Architecture
discussion, IFOM is examined in detail and then several examples are provided that show the benefits of offloading certain traffic to WiFi.

2. Existing LTE Evolved Packet Core
The current LTE EPC architecture is shown in Figure 1. This architecture supports eNode B’s that are deployed as part of the core network itself as well as supporting local eNode B’s, whether they be Home eNode B’s deployed in a residence or a small enterprise or a “small cell” eNode B deployed in a metro environment. The core network eNode B’s provide the bulk of coverage while the Home eNode B’s are deployed by customers within their premise to provide in-building local coverage. Small cell eNode B’s deployed outside of the mobile core network, usually by operators or other large entities provide capacity boosts to users at the area covered by one of these small cells. For instance, a mass transit station is an ideal location for a small cell to be deployed as it is a location where many users may be concentrated. A traditional cell tower, serviced by an eNode B may be insufficient to cover this metro environment, so an operator may deploy an eNode B to provide its customers with the needed services.

3. Converged Gateway
3.1 Introduction
A method to further enhance the capacity in any of these scenarios is via the introduction of a CGW. The CGW can be introduced within the EPC or deployed locally, outside the EPC, depending on the use case, as shown in Figure 2. The CGW provides this improved capacity by allowing for IFOM dynamic flow management between the WiFi and cellular transports. In addition, the CGW has other benefits:

- Allows for local control and management of the cellular and WiFi spectrum
- Allows for superior interference mitigation
- Supports simultaneous use of Selected IP Traffic Offload (SIPTO) and IFOM

These benefits are attained by requiring no change to any existing EPC elements and requiring no change to any existing protocols used by any of these EPC elements. Furthermore, the CGW supports all procedures defined in the current standards.

There are two salient features of the CGW that allow it to manage the WiFi and cellular accesses and the flows that traverse these accesses. These features include:

- Transparent to both the EPC and femto-cell
- WiFi/3G association

Both of these features are explored in depth in the following sections on this document.

3.2 Placement within LTE Architecture
The introduction of a CGW between an eNode B and the Serving Gateway (SGW) and Mobility Management Entity (MME) is transparent to both the eNode B and EPC. No protocols or procedures require modification to allow this placement of the CGW. However, it is necessary to properly provision the CGW so that it can act as an intermediary. The absence of a requirement to

Figure 2. LTE Architecture in the Presence of the CGW
change any of the existing procedures and protocols is a significant advantage of the CGW solution.

For an EPC-based placement, the CGW does not require security provisioning since it is internal to the EPC. But in a metro or premise environment, the CGW will be provisioned with the security keys needed to establish a security association to both eNode B and the Secure Gateway (SeGW) at the edge of the EPC. The IPSec Security Associations (IPSec SA) that normally exists between the eNode B and SeGW will be split into two IPSec SAs. The first is between the eNode B and the CGW. The second is between the CGW and the SeGW. It is expected that each interface would have unique keys. Regardless of the deployment configuration, the CGW acts as a proxy for the signalling between the eNode B and EPC elements.

For a non-EPC based configuration, when the eNode B has a signal to send to the EPC, it sends it through the IPSec SA to the CGW. The CGW with terminate the IPSec encapsulating the signal, then encapsulates the signal with IPSec, and sends the signal to the SeGW. When the SeGW receives the signal, it will terminate the IPSec encapsulating the signal and will forward it into the EPC. When an element within the EPC has a signal to send to the eNode B, it occurs in a similar fashion. For an EPC based configuration, the signalling passes through the CGW but does not require an IPSec tunnel.

A similar mechanism occurs when data is sent between an end-user and the EPC. As the data travels through the cellular network a GPRS Tunnelling Protocol (GTP) tunnel will be used between the eNode B and the SGW. This tunnel is used to ferry data between the eNode B and SGW. With the CGW in place, there are two GTP tunnels, one between the eNode B and CGW and the other between the CGW and SGW. Downlink data is sent from the SGW via a GTP tunnel to the CGW. The CGW terminates the GTP protocol and, if it decides to route that packet via cellular, places the data packet into the GTP tunnel between the eNode B and CGW. The eNode B receives the packet and delivers it to the end-user. If the CGW decides to route the packet via WiFi, the CGW will send the packet to the WiFi AP. In the uplink, an end-user device many send the packet over the WiFi or cellular transports. If it is sent via cellular, the packets are sent from the eNode B to the CGW via a GTP tunnel. The CGW terminates the GTP protocol and places the packet into the GTP tunnel that exists between the CGW and the SGW. If it is sent via WiFi, the CGW will receive the packet and place in the GTP tunnel that exists between the CGW and the SGW.

This tunnelling concept is shown in Figure 3.

### 3.3 Data Plane

The current, sans CGW, data plane protocol/transport between the SGW and eNodeB is shown in Figure 4.

With a CGW in place, the data plane between the SGW and eNode B changes as shown in Figure 5. Notice that the eNodeB and SGW are not impacted with the insertion of the CGW. In an EPC based CGW deployment, the figures are the same except there is no IPSec layer.

As a result of IFOM, some data will be routed through WiFi between the SGW and the end-user device. The protocol/transport is shown in Figure 6.
4. WiFi AP

The architecture in this paper employs a WiFi Access Point (AP) that is managed and controlled by the CGW. To facilitate this control and management, the AP is configured to use the Dynamic Host Configuration Protocol (DHCP) Server in the CGW. This setting allows the CGW to manage and be cognizant of the local IP Addresses assigned to the WiFi interface of devices connected through the managed WiFi AP. The use of the DHCP Server within the CGW is a key to determining that a device is reachable over both WiFi and cellular transports as is described below.

5. End-User Device Management

5.1 End-User Device Attachment

When an end-user device connects to an eNode B, the end-user device and Packet Domain Network (PDN) Gateway (PGW) exchange signals to allow the end-user device to attach to the network. As part of the Initial Attach procedure, a PDP context will be activated to allow for the exchange of data with, perhaps, an application server on the public Internet via the EPC. In the absence of the CGW, this exchange of data between the end-user device and SGW is carried within a GTP tunnel that exists between the eNode B and SGW. With the CGW in place, the exchange of data between an end-user device and an entity on the public Internet is performed by two GTP tunnels. The first is between the CGW and eNode B while the second is between the CGW and the SGW.

5.2 Client Addressing

When the end-user device associates with the WiFi AP, it will request an IP address from the WiFi AP using the DHCP protocol. Since the DHCP Server in the WiFi AP is disabled, the request for an IP address is forwarded to the CGW by the WiFi AP. The DHCP Server within the CGW assigns a local IP address to the WiFi modem of an end-user device.

For the cellular modem, the end-user device will be assigned an IP address by the PGW as part of the Initial Attach procedure (or a subsequent PDP context activation procedure).

Once both these events occur, the end-user device has two IP addresses, one for each modem. The EPC assigned IP address is assigned to the cellular modem while the IP address assigned by the DHCP Server within the CGW is assigned to the WiFi modem. The CGW is cognizant of both connections since both were established with its knowledge. However, the CGW does not know that these two IP addresses terminate in the same end-user device. Therefore, a method is required to link these two IP addresses to the same device. We describe two such methods below.
5.3 WiFi/LTE IP Address Association at CGW

As part of the DHCP procedure that resulted in the WiFi modem of the end-user device being assigned an IP address, the DHCP Server became aware of the Media Access Control (MAC) of the WiFi modem.

To link the WiFi parameters to the cellular IP address, the CGW will issue an Address Resolution Protocol (ARP) Request using the cellular IP address, and in most cases, the WiFi card within the wireless device responds with its MAC address. Once this occurs, the CGW knows the WiFi MAC address, local WiFi IP address, and cellular IP address are all associated with the same end-user device.

In order to ensure that it is the same device, the CGW sends an Internet Control Message Protocol (ICMP) Echo Request message via the WiFi AP with the destination IP address set to the cellular IP address extracted during the setup of the PDP context.

If the end-user device is connected through the WiFi AP managed by the CGW, the end-user device responds with an ICMP Echo Response message with the source and destination IP addresses reversed. From this information, the CGW infers that the end-user device has both the cellular and local WiFi connection "active".

It is necessary for the CGW to periodically issue the ARP Request and ICMP Echo Request messages to ensure that the WiFi connection is still active. Another reason for periodically performing this process is when the cellular PDP context is activated after the WiFi has associated. It is expected that the Operating System (OS) in the CGW will take care of managing the linkage of all this information.

A benefit of this logic is that the CGW still supports devices that support only WiFi connections as well as end-user devices that only support LTE connections. For a WiFi only device, it will not respond when the CGW sends a query related to an EPC assigned IP address. Similarly, a cellular only device, when the CGW issues a query related to that EPC assigned IP address, there is no WiFi modem to respond to that particular IP address.

If a device does not respond to the ARP request even when it does have both modems yet still supports IFOM, there are other methods for the CGW to learn that the device has both WiFi and cellular modems active. When the end-user device requests a policy from the CGW, it can inform the CGW of that it does indeed have two modems active.

6. IP Flow Mobility

The CGW will have policies that define how traffic is to be routed for particular users and particular traffic. The CGW will have its own default policies and it may also get policies from interactions with the Access Network Discovery and Selection Function (ANDSF) Server located within the EPC. The CGW also has logic which blends the policies defined locally with those retrieved from an ANDSF Server (or other entity). The policy conforms to the Long Term Evolution (LTE) standards that define the criteria that should be contained within a policy. The policy allows for defining which types of IP Flows are permitted (or not permitted) on which accesses. For any IP Flow or the characteristics of an IP Flow, the CGW may have a routing rule that dictates one of the following:

- Cellular Only
- WiFi Only
- Cellular Preferred
- WiFi Preferred
- No Preference

Even if no specific routing rule applies to a specific IP Flow, the CGW will have a default rule which it can apply. Based on some characteristic of the IP Flow, such as application (i.e. Voice over IP (VoIP), File Transfer Protocol (FTP) or gaming protocol such as Xbox LIVE), or the application server IP address range (such as YouTube), the operator can enforce one of the above access routing rules. This granularity allows the operator great flexibility in managing flows and effectively controlling load balancing between licensed and unlicensed spectrum. Additionally, the policy has a catch-all setting which is used for flows that are either unknown or new, prior to DPI being performed.

An example policy is shown below:

- IMSI = "12345678901234"
- Rule 1
  - Type = 'Skype'
  - Routing Rule = 'Cellular Only'
- Rule 2
  - Type = 'Xbox LIVE'
  - Routing Rule = 'WiFi Only'
- Rule 3
  - Type = 'Streaming Video'
  - Routing Rule = 'No Preference'
- Rule 4
  - Type = 'Default'
  - Routing Rule = 'Cellular Only'

In this user-specific policy, Skype traffic for this user is routed over the cellular access. Meanwhile, all Xbox LIVE traffic will be routed over WiFi the access only. Finally, streaming video will be routed over either the cellular or WiFi access as a function of the loading on each access. Any traffic that is not any of these types is only routed over the cellular access. This is a typical scenario that a person may desire. They would like to get the QoS benefit of cellular for their Skype sessions but push the kids Xbox traffic to WiFi where it will not interfere.

6.1 Addition of New IP Flows

When an IP Flow starts, the flow type is unknown and its packets are sent by the IFOM functionality within the CGW to the default access. Once DPI is performed within the CGW on an IP Flow, the IP Flow type may be
known. For example, the IP Flow type may be Skype, FTP, or Xbox LIVE data. If the DPI is not able to discern what the IP Flow is, the flow will be labelled as unknown. Once the DPI has concluded, the policy can be consulted for this specific user. As was described earlier, the policy contains the routing rule associated with the identified flow type. If the access indicated in the rule is the same as the default access, then no change is made. The IFOM will continue to route the data to the end-user device via the same access. Should the routing rule indicate a different access, the IFOM functionality will begin routing the data associated with this IP Flow to the end-user via the selected access. When a new flow is identified by the DPI, the IFOM routes the data associated with that IP Flow to the end-user with the selected accesses based on the following rules:

- Cellular Only and Cellular Preferred IP Flows are assigned to the cellular access
- WiFi Only and WiFi Preferred IP Flows are assigned to the WiFi access
- No Preference IP Flows are assigned to the least loaded accesses as a function of the number of existing IP Flows, the measured throughput and of the capacity of the access

Using the example policy shown in the previous section, if a new IP flow is detected, the CGW will place it on cellular access while DPI is performed on the IP Flow. To illustrate the functionality of the CGW, let us assume that the new IP Flow is Xbox LIVE traffic. When the Xbox LIVE traffic starts, it will be assigned to the cellular transport since it is the default transport. Once DPI has determined that the traffic is indeed Xbox LIVE traffic, the CGW will consult the routing rules, learn that this type of traffic is to be assigned to WiFi and move this traffic to the WiFi access.

### 6.2 Dynamic Flow Management of Existing IP Flows

Periodically, the CGW performs load balancing to ensure no access is overburdened. If an access is found to have an unfair proportion of the current IP Flows, the CGW attempts to perform load balancing without violating any of the routing rules of the existing IP Flows. When performing load balancing, the IFOM within the CGW takes into account the following factors:

- Number of IP Flows on each access
- Bandwidth of IP Flows on each access
- Bandwidth capacity of each access
- The policy for each IP Flow

An example helps illustrate this logic. Let us assume that a CGW in used to manage a WiFi AP and Home eNode B in a premise. Further assume there are three users, a parent and two children, each having a dual mode device, and each desiring to perform some task which requires bandwidth. Finally, assume that the policy in place within the CGW for all three of these users is as shown in the introductory paragraphs of Section 6.

The first child starts a video streaming session. Since the default access is cellular, this IP Flow is started on the cellular access. After DPI has identified this as streaming video, it remains on the cellular access as there as the policy for streaming video is “No Preference.” After load balancing has occurred, this IP Flow remains on the cellular access since there is no other traffic.

The parent then starts a Skype session. Since the default access is cellular, this IP Flow is started on the cellular access. Once it is identified by DPI to be Skype, the policy is consulted and since the rule is “Cellular Only”, the Skype session remains on the cellular access. When load balancing is performed by the CGW, it will recognize that there is congestion on the cellular access as compared to the WiFi access. It will review all IP Flows and decide to move the streaming video session to WiFi since its policy allows it to use either transport. Therefore, once load balancing occurs, the Skype session will remain on cellular and the streaming video session will be on WiFi.

The second child then starts using his or her Xbox and generates traffic that has the Xbox LIVE protocol. Prior to DPI being performed, this IP flow is assigned to the cellular access. After DPI, which identifies the IP Flow as Xbox LIVE, it is moved to the WiFi access as per the policy. After this assignment, there is one IP Flow on the cellular access – Skype – and two IP Flows on the WiFi access – streaming video and Xbox LIVE. Let us assume that the Skype session concludes which leaves the cellular access having no traffic while the streaming video and Xbox LIVE remain on the WiFi access. When load balancing occurs within the CGW, it will decide to move the streaming video session to cellular since its policy allows for it to be on either transport. The Xbox LIVE traffic remains on the WiFi access. In this instance, the CGW has attempted to balance the loading of the existing sessions to maximize usage of the current capacity of the access.

### 6.3 Link-down handling of existing IP Flows

Part of the support required to effectively manage the IP Flows will be the end-user device reporting channel metrics to the CGW. For example, the CGW will configure the end-user device to report the Received Signal Strength Indicator (RSSI) of the active accesses when they pass certain thresholds. This allows the CGW to induce the end-user device to send an alert to the CGW when the signal quality has passed through a CGW-defined limit. As the end-user device performs this monitoring, should the quality of the signal pass through this limit, the end-user device will inform the CGW. The CGW will then attempt to move as many IP Flows as possible for this user from the degraded access to the other access. There are several criteria that are used when evaluating each IP Flow for the particular user:

- Routing Rule for this IP Flow
- Quality of the non-degraded access
If an end-user reports that an access has degraded, the CGW will examine the routing rule for each IP Flow. The CGW will only move those flows away from the degraded access that have a routing rule that allows the IP Flow to move, such as WiFi-Preferred, Cellular-Preferred, or No Preference. IP Flows that have a rigid routing rule, either WiFi-Only or Cellular-Only will not be moved even in this case. Additionally, the CGW will examine the quality of the non-degraded access, if the end-user device has that access available. If the non-degraded access is also of poor-quality, the CGW will not move any IP Flows. This prevents the scenario where both accesses become degraded and the CGW attempts to mitigate this situation by moving IP Flows back-and-forth between two degraded accesses.

An example will help clarify the use of this logic in a practical case. As we concluded the load balancing discussion, there were two IP Flows, an Xbox LIVE session that was assigned to the WiFi access and a streaming video session that was assigned to the cellular access. Let’s assume that the quality of the cellular access becomes poor for that user, not poor enough to drop the call, but whose signal strength has been degraded. When this occurs, the CGW will determine if the “other” access, in this case, the WiFi access is of good quality. For the purpose of this description, we assume it is of good quality; therefore, the CGW will review the policy for each IP Flow on the cellular access. Since the policy for streaming video is that it can use either access, the CGW will move the streaming video session away from the cellular access and move it to the WiFi access.

For the same example, if the WiFi access had become poor while the cellular access had remained strong, the CGW would evaluate each of the IP Flows on WiFi and if permitted by the rules, would move IP Flows from the WiFi access to the cellular access. However, in this case, since the Xbox LIVE IP Flow has a rule that only allows it use WiFi, the CGW would not move this IP Flow from WiFi to cellular.

7. Other Areas of Innovation

There are several other areas of innovation within the CGW that we are currently exploring. These are briefly summarized in this section.

7.1 Local Control and Management of the Cellular and WiFi Spectrum

Since the CGW accesses the policies from the ANDSF Server within the EPC, it can apply, or overlay, local policies onto the policies stored within the core network. This allows for a local administrator to control and manage both the licensed and unlicensed spectrum.

7.2 Interference Mitigation

The X2 interface between eNode’s is routed through the CGW. This allows the CGW to monitor the interference experienced by specific users and by the eNode B’s. When the CGW detects interference in the cellular spectrum, it will assess if it can move some flows from cellular to WiFi to reduce the interference being experienced.

7.3 Simultaneous use of SIPTO and IFOM

The current LTE standards do not adequately support simultaneous SIPTO and IFOM. If SIPTO is used by the core network, then IFOM is not supported and vice versa. Since IFOM is performed in the CGW, between the SGW and eNode B, SIPTO and IFOM can both be done simultaneously.

8. Conclusion

In this paper we have detailed how the CGW is situated within the cellular network and explained how it is configured. We also have shown how it is manages the communications between an eNode B and the EPC elements. We then detailed IFOM and gave examples of the CGW processing for different use case scenarios. We then listed and briefly listed other areas of innovation that are currently being explored.

References

[3] 3GPP TS 29.274, “Evolved GTPv2-C; Stage 3 (Release 10)”, v10.3.1
[5] 3GPP TS 32.592, “HeNB OAM&P; Information model for Type 1 interface HeNB to HeMS (release 10)”, v10.4.0
[6] 3GPP TS 32.593, “HeNB OAM&P; XML definitions for Type 1 interface HeNB to HeMS (Release 10)”, v10.0.0
[8] 3GPP TS 36.411, “S1 Layer 1 (Release 10)”, v10.1.0
[9] 3GPP TS 36.412, “S1 Signaling Transport (Release 10)”, v10.1.0
[10] 3GPP TS 36.413, “S1 Application Protocol (S1AP) (Release 10)”, v10.2.0