Life of a Packet within a GENI Experiment

GDD-07-47

GENI: Global Environment for Network Innovations

April 2, 2007

Status: Draft (Version 0.4)
Note to the reader: this document is a work in progress and continues to evolve rapidly. Certain aspects of the GENI architecture are not yet addressed at all, and, for those aspects that are addressed here, a number of unresolved issues are identified in the text. Further, due to the active development and editing process, some portions of the document may be logically inconsistent with others.

This document is prepared by the Backbone and Wireless Working Groups.

Editors:

Jennifer Rexford, *Princeton University*

T.V. Lakshman, *Lucent Technologies Bell Laboratories*

Contributing workgroup members:

Sarit Mukherjee, *Lucent Technologies Bell Laboratories*

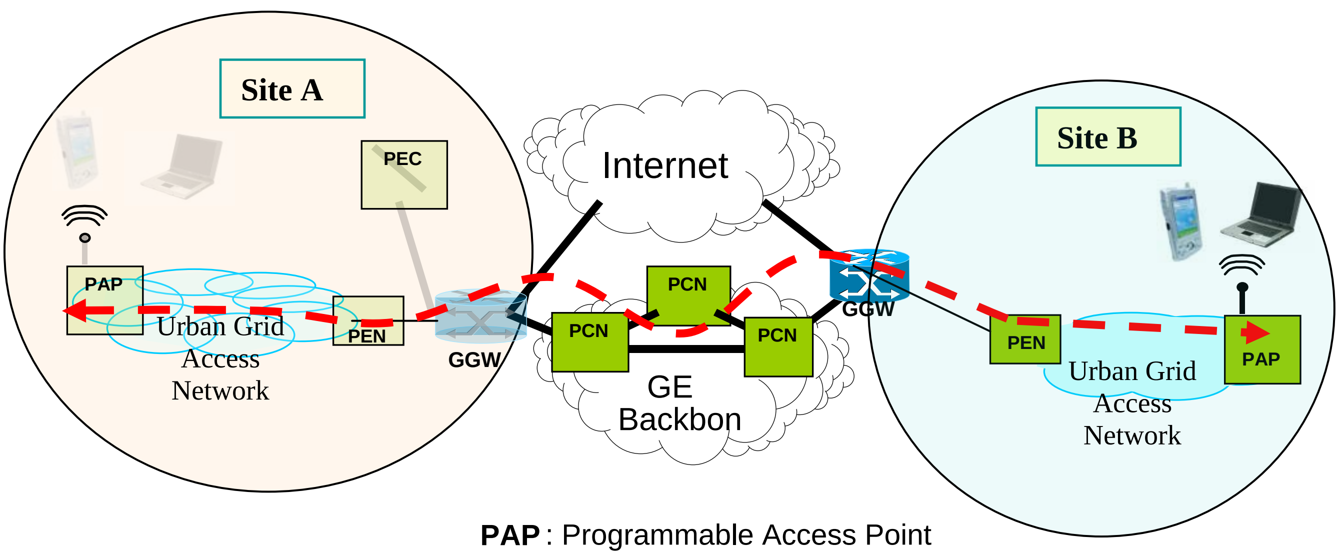
Sampath Rangarajan, *Lucent Technologies Bell Laboratories*

1 Introduction

In this document, we detail some experimental scenarios and consider end-to-end packet flows within the GENI infrastructure for each of these scenarios. To start with (the intention is to add more scenarios to this document later), we describe the assumed end-to-end architecture first and consider one specific scenario where packets flow between two wireless networks located within two different edge networks over the GENI backbone. For this example, the backbone serves as a cut-through where the PCNs at the GENI PoPs are used to forward traffic through a "fast-path" without any extra processing.

2 Example Scenario and Architecture

The architecture diagram shows two edge-networks connected over a GENI backbone. The edge networks consist of Programmable Access Points (PAP), which are the end-to-end producers and consumers of packets in our example, Programmable Edge Nodes (PEN), which serve as gateways to the PAPs, Programmable Edge Clusters (PEC), which run distributed services (these are not required for our example, but included for completeness) and the GENI Gateway (GGW) which is a gateway between the edge network and GENI PoP and terminates the physical tail circuit from the GENI PoP. The GENI backbone consists of a set of Programmable Core Nodes (PCN) on which slivers can be run.



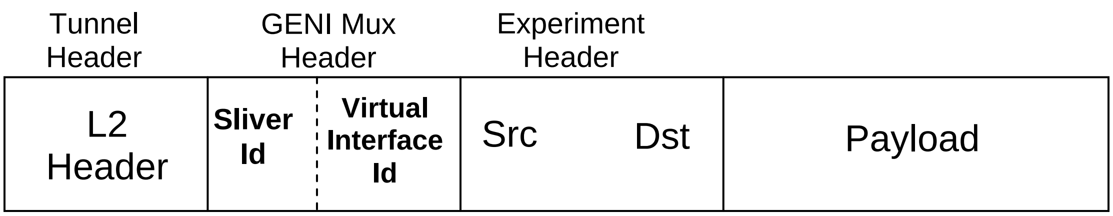
PAP : Programmable Access Point
PEN : Programmable Edge Node
PEC : Programmable Edge Cluster
PCN : Programmable Core Node
GGW : GENI Gateway

For an illustrative life-of-a-packet discussion, we consider an end-to-end experiment that requires the experimenter to receive a GENI slice with slivers running on the PAP and the PEN on site A and the PAP and PEN on site B but only requires a QoS guaranteed tunnel between the PAPs on site A and site B for bi-directional traffic flow between the PAPs (shown by the red dashed arrow). For this experiment, the experimenter does not require sliver configurations on the PCN to process packets but would like the PCN to cut-through data through a "fast-path".

In this experiment, a mobile within site A sends a packet to a mobile within site B. The packets are received at the PAP at site A and are forwarded to the PAP at site B through a QoS guaranteed tunnel. The mobiles use geo-location addresses to identify one another (for eg. latitude-longitude information). For this example, we assume that the mobiles know the geo-location of the other mobiles.

3 Packet Format

Before we describe the packet flow for the example scenario, we consider a generic structure for a packet shown below.



- The experiment header provides the source and destination geo-locations (among other information) of the mobiles. It is added to the packet by the experimental sliver running at the source of the packet. The experiment header, as it passes through the different GENI components, would be processed by "slivers" that run on these components corresponding to this experiment.

- A GENI Mux header is required to de-multiplex the packet at a GENI component and forward it to the appropriate sliver. The GENI Mux header may consist of a sliver id and a virtual interface id for the interface within the sliver that should receive this packet.

- Encapsulating the GENI Mux header would be a tunnel header which indicates the QoS guaranteed tunnel that has been setup to forward the packet. Normally, the tunnel header would be a Layer 2 header such as an Ethernet header with the appropriate VLAN tag to differentiate this traffic from other traffic or an MPLS header where the Exp bits specify the DiffServ class for the packets carried within the MPLS tunnel.

The following points need consideration:

- The type and format of the GENI Mux header and the tunnel header could be link-local, i.e., could be different on the links between any two GENI components. This issue is discussed further in Section **Error! Reference source not found.**.
- The L2 header could also change between different GENI components depending on the layer 2 technology used.

- It is also conceivable that the L2 header and the GENI Mux header are combined together into one header (on one or more of the links) where the L2 header serves the dual-purpose of de-multiplexing between the slivers as well enabling traffic differentiation.

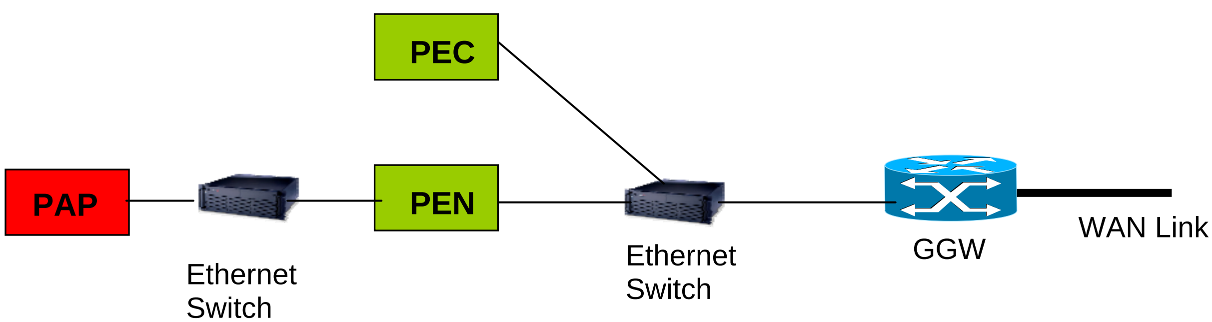
Encapsulating the L2 header would be a L1 header (such as a SONET header) that is specific to the physical link technology and is not shown in the figure.

For our example, the geo-location addresses would constitute the Experiment header. The GENI Mux header specifies the sliver and virtual interface ids for the sliver that should process packets at the PAP and the PEN (as indicated earlier, for this example, the PCNs do not run experiment sliver for this example; they may run “control slivers” to perform MPLS label switching which we address later).

4 Packet Flow

The tunnel from the PAP at site A to site B may consist of a concatenation of multiple “tunnels”. For our example, we assume the following:

- a single Ethernet hop from PAP to PEN where the traffic is differentiated using VLAN tags,
- another Ethernet hop from the PEN to the GGW, VLAN tags again differentiating the traffic,
- an MPLS tunnel from the GGW at site A to site B which traverses multiple PCNs which serve as “cut-through” and only perform MPLS label switching to forward the packet over the end-to-end MPLS tunnel.

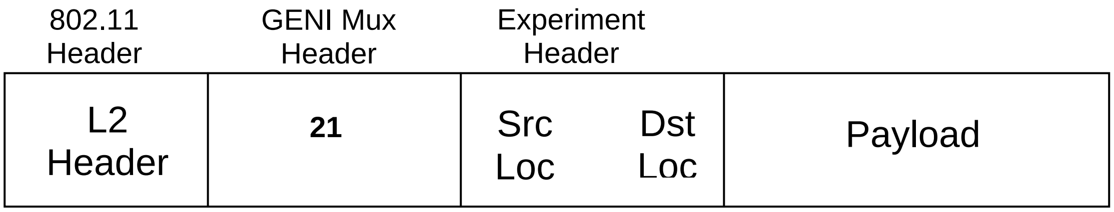


Traffic from the GGW at site B to the PAP would also be carried over two Ethernet hops with traffic differentiated using VLANs.

The figure above shows the network connectivity among different GENI components in site A. The PAP is connected through an Ethernet link to the PEN. The PEN is connected to the GGW through an Ethernet link as well. The PAP and PEN run slivers specific to the experiment.

4.1 Packet Originating at Mobile Node

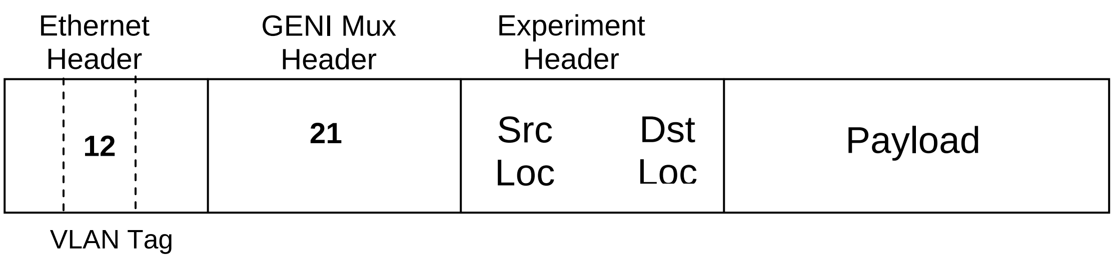
The packet originating at the mobile is received over an 802.11 link at the PAP as shown below.



The GENI Mux header that denotes the sliver and virtual interface ids is included by an experiment specific client on the mobile and is assumed to be 21 for this example.

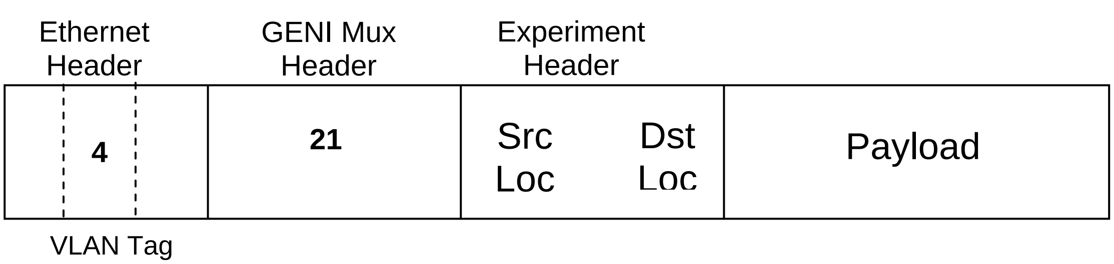
4.2 Packet Processing at PAP

Based on this header, the packet is forwarded to the appropriate sliver at the PAP. The sliver at the PAP looks up a "routing table" based on the experiment header and determines that the destination mobile is at the PAP at site B. (Note: The control plane mechanisms for the sliver at the PAPs to exchange geo-location based routing information to populate this routing table is beyond the scope of this document). Based on this information, the sliver determines the next hop (which is the PEN) and sends the packet to the PEN after marking the VLAN tags appropriately to differentiate this traffic on the link between the PAP and the PEN. For our discussion, we assume that the GENI Mux header is global; as indicated earlier, this could be link-local. The packet from the PAP destined to the PEN would appear as shown below. It is assumed that the VLAN tag used is 12.



4.3 Packet Processing at PEN

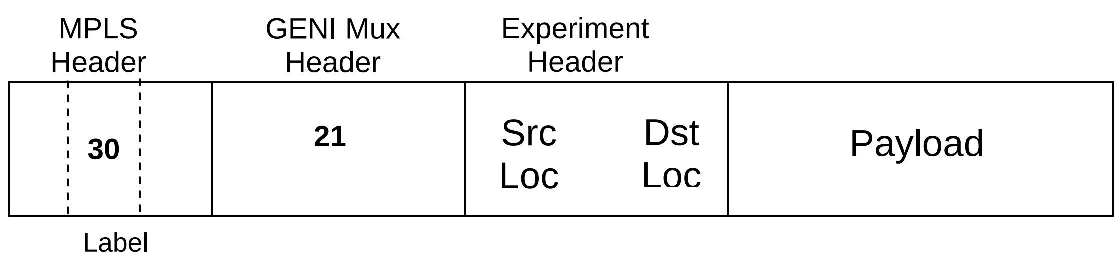
When the packet is received at the PEN, it is forwarded to sliver/interface 21. The sliver processes the packet and determines that it needs to be sent to the GGW as the next hop. To differentiate the packet on the Ethernet link between the PEN and the GGW, it tags the packet with VLAN Tag 4. The packet from the PEN to the GGW would then appear as follows. Note that VLAN tags chosen are configured into the slivers running on the PAP and the PEN and the mechanism for this will be discussed in the corresponding GMC documents.



4.4 Packet Processing at GGW

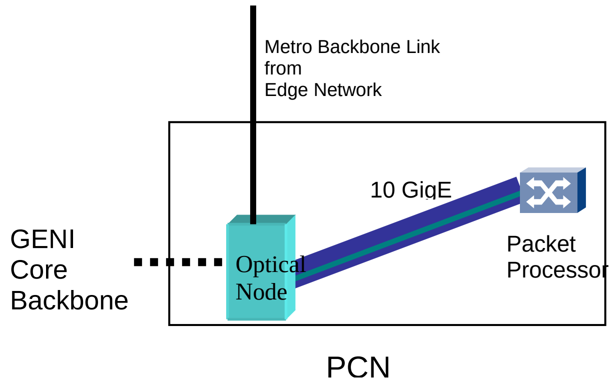
We assume that the GGW is not programmable but is capable of establishing MPLS tunnels. Through control mechanism not discussed in this document, we assume that an MPLS tunnel has been set up between the GGWs at sites A and B. We further assume that the GGW is configured to look at the VLAN tags and encapsulate the packet with an appropriate MPLS tunnel header so that packets can be differentiated over the WAN link. (Again note that these capabilities of GGW are assumed for this example and may vary in other scenarios.) This header may be used by the PCN to perform cut-through. Note that the GGW does not have to look at the GENI Mux header as it does not run any slivers on behalf of the user. The MPLS tunnel between the GGWs traverses the three PCNs shown in the figure. We assume that at each PCN, when a packet is received, it is label switched towards the destination GGW.

When the packet is received at the GGW on Site A, the Ethernet header will be stripped and an MPLS header with the appropriate label will be added. For the example, we assume that the MPLS label is 30 as shown in the packet format below. The packet is then sent out over the WAN link. Lower layer headers such as a POS header (Packet-over-SONET) and a SONET header/trailer will be added to the packet before it is sent out, but these are not shown in the figure.



4.5 Packet Processing at PCN

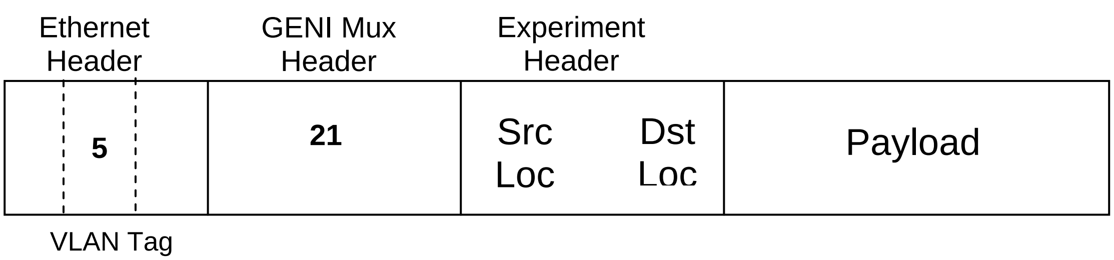
The PCN consists of two main components, the Packet Processor and the Optical Node. This component level split is not shown in the original figure. WAN links are terminated at the Optical Node. The Optical Node connects to the Packet Processor using Ethernet or Packet-over-SONET (POS). Packets that need processing at the PCN are sent over this.



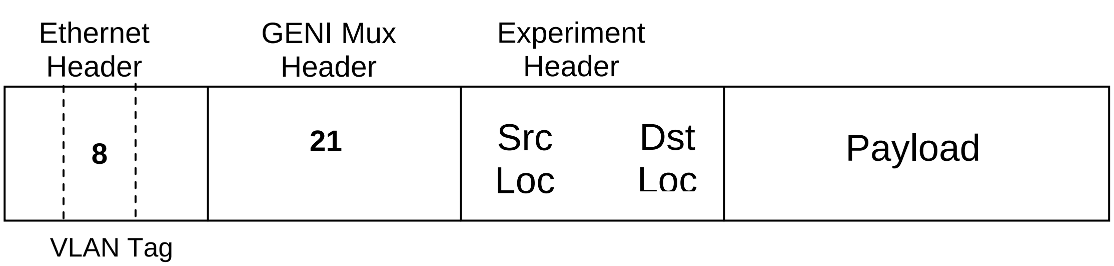
In this example, we assume that the PCN support cut-through routing at the optical domain for the traffic flow and, therefore, packets do not get out of the optical domain within the PCN. The packet traverses multiple PCNs and ultimately is sent to a destination GGW on Site B over a tail circuit.

4.6 Packet Processing at Site B

At the GGW at Site B, the MPLS header will be stripped off and an appropriate VLAN tag will be included and sent to the PEN as shown.



The PEN will strip the VLAN tag and look at the GENI Mux header to decide on the sliver to forward this packet. After processing, the packet will be forwarded to the PAP at site B with the appropriate VLAN encapsulation as shown.



When the packet is received at the PAP, it is delivered to sliver/virtual interface 21. The sliver processes the packet and based on the experiment header, delivers the packet to the mobile located at the destination location.

5 Discussion on GENI Mux Header

- We mentioned in Section 3 that a GENI Mux header is required to de-multiplex the packet at a GENI component and forward it to the appropriate sliver. The GENI Mux header may consist of a sliver id and a virtual interface id for the interface within the sliver that should receive this packet. It is possible to have the GENI Mux header global across the network (i.e., it does not change within a slice) or link local (i.e., it changes at each hop and is associated with the link the packet traverses). There are pros and cons for each configuration.
    a. GENI Mux header as global across the network: The global format and global id space makes allocation of ids simple if there is a central authority (like GMC). Monitoring and tracking of experiments, e.g., to trace back bad behavior to the offending slice, is easier with global ids. However, if there is no central authority to allocate ids, then global coordination is required to limit fragmentation of the limited id space.
    b. GENI Mux header as link local: Link local id space avoids global coordination of limited id space. Local formats allows different link technologies to use different sized id spaces (e.g., just a few bits in wireless/sensor, and maybe many bits in backbone, consistent with the differences in bandwidth and the differences in the expected number of slivers per node or link). Local formats would also allow different existing technologies to be used (e.g., MPLS, VLAN, IP, IP/UDP, time/frequency slots) as appropriate for the media/device. On the other hand, link local ids would require the appropriate mapping to be downloaded into each of the GENI elements that run a sliver on behalf of an experiment.