Scalability Analysis of Metro Ethernet

Heikki Mahkonen
Helsinki University of Technology
Email: heikki.mahkonen@piuha.net

***Abstract*— This document gives a scalability analysis of Metro Ethernet architecture. The purpose is to identify architectural issues in Metro Ethernet Network (MEN) that are scalability constraints when extending the carrier grade Ethernet service provided by the MEN to a routed Ethernet service. The intention of the scalability analysis presented in this document has not been to present detailed theoretical or mathematical analysis of the issues identified. The aim has been to identify the possible constraints for routed Ethernet service and give some examples how these issues could be resolved. The detailed design for resolving these issues is left as a future work items. In addition, the first part of this work topic has been to learn internal workings of Metro Ethernet and its architectural elements.**

I. Introduction

Metro Ethernet Forum (MEF) is responsible of standardization a Metro Ethernet architecture that provides carrier grade Ethernet service that can connect two or more distant LANs. Basic Metro Ethernet architecture is built from service providers Metro Ethernet Network (MEN) and two or more subscriber LANs. The MEN consists of Ethernet switches or bridges that forward Ethernet frames between subscriber LANs. The subscriber LANs are connected to the MEN with Customer Edge (CE) equipment. The Ethernet traffic flow from one of the subscriber's LANs is encapsulated by the service providers MEN into Metro Ethernet frames and delivered to the other LANs via their CE equipment. This service is invisible to the LAN users of the subscriber and the interconnected LANs seem to be all part of the same LAN.

To extend carrier grade Ethernet service to a routed Ethernet service the functional elements of the Ethernet architecture needs to be extended in order to provide a network architecture that can support scalable services like the Internet. The Internet routing architecture has two big advantages when compared to Ethernet. One of these advantages is the hierarchical routing architecture that provides true global routability. The other advantage is the nature of the IP address. The IP address itself provides enough global location information for a IP router to make independent routing decision for individual IP packet. The Ethernet is a flat architecture and the MAC addresses are not globally routable entities nor can they be used to uniquely identify a node.

The next section takes detailed look into metro Ethernet architecture. It defines all the functional elements needed to provide carries grade Ethernet service. The following section gives scalability analysis of these architectural elements. The functional elements of the Metro Ethernet architecture are analyzed. In the next section suggestions are made on how these elements might be extended to form routed Ethernet service. Finally the last section concludes the work.

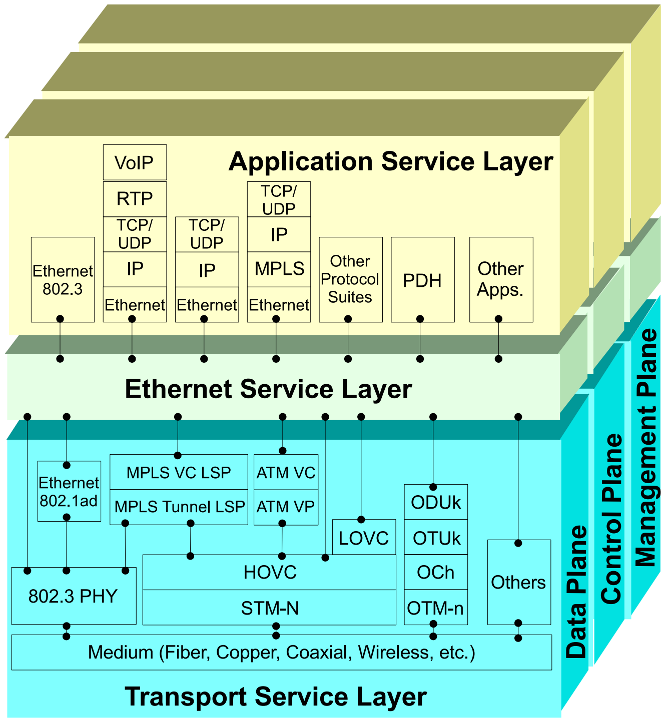

Fig. 1. Layers of metro Ethernet architecture.

II. Metro Ethernet Architecture

Metro Ethernet architecture is built on top of the layered network model shown in figure 1. This layered network model consists of three layers, the Application Service Layer (APP), Ethernet Service Layer (ETH), and Transport Service Layer (TRAN). The same layer structure is used for, Data Plane, Control Plane, and Management Plane. [1]

The data plane (figure 1) provides functional elements needed to steer subscriber flows and transport traffic units between MEN network entities. The control plane supports distributed flow management for network entities in MEN. In addition, the control plane defines the signaling mechanism, supervision and connections release operations for distributed set up. The management plane includes fault, configuration, account, performance, and security functions and support for OAM tools. [1]

Application Service Layer (APP) (figure 1) gives support for legacy applications carried in the ETH layer through MENs

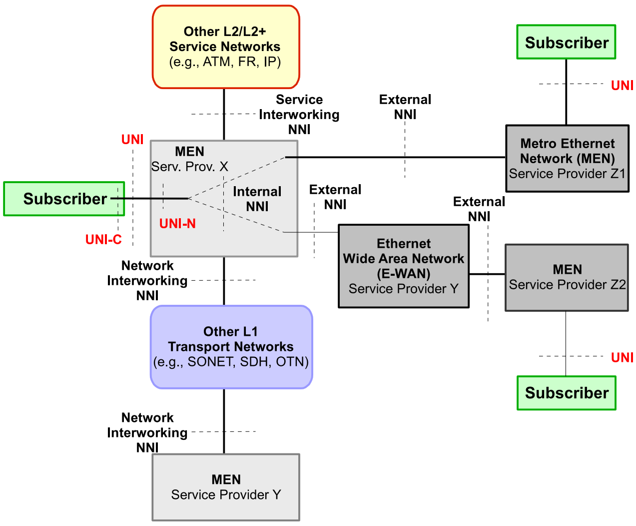

Fig. 2. Metro Ethernet Network reference points.

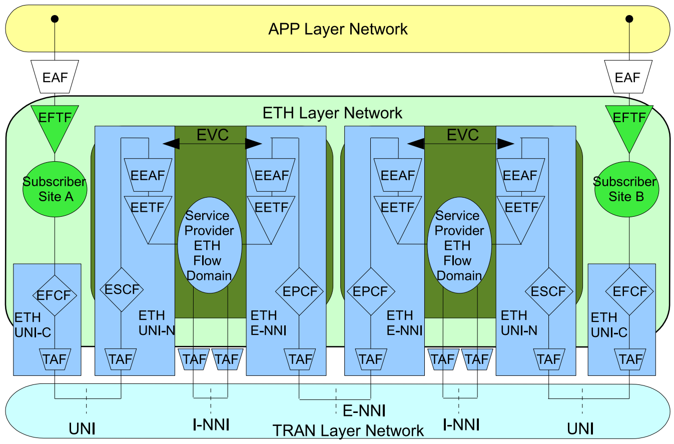

Fig. 3. Functional elements of metro Ethernet architecture.

and add-on functions to support Ethernet Service Layer (ETH). Metro Ethernet architecture supports basically any kind of application layer service (TCP/UDP, IP, MPLS, etc.) to be carried in ETH layer. [1]

Ethernet Service Layer (ETH) (figure 1) defines Ethernet (MAC address) connectivity service and handles the delivery of ETH frames. ETH layer also handles service aware operations, administration, maintenance, and provisioning capabilities to support Ethernet type connectivity service. ETH layer supports broadcast, multicast and unicast Ethernet frame formats defined in IEEE 802.3-2002. [1]

Transport Service Layer (TRAN) (figure 1 gives support for ETH layer connectivity in a transport mechanism independent manner. Many different transport layer networks and their respective server layers (802.3PHY, 802.1 bridged, SONET/SDH, ATM VC, OTN ODUk, PDH, etc.) can be used to transport ETH layer flows. TRAN layer functions in between the ETH layer and the physical transmission mediums. [1]

A. MEN Reference Points

Metro Ethernet architecture defines reference points that must be used when sending or receiving Ethernet frames across different network domains. Figure 2 shows a example usage of these reference points and how they are used to connect different parts of a Metro Ethernet architecture.

User-Network Interface (UNI) is an interface that connects MEN service subscribers to MEN service providers networks. UNI is functional element that consist of client (UNI-C) and network (UNI-N) side elements (figure 2). UNI-C supports functionality to exchange data, control and management plane information with MEN service provider. The UNI-C is entirely in subscribers domain. UNI-N is entirely in the MEN service providers domain. The UNI-N supports functionality to exchange data, control, and management plane information with the MEN. [1]

External Network-to-Network Interface (figure 2) (E-NNI) is used to interconnect two MENs. The E-NNI provides reference point between network equipment inside two MENs that provide the physical connection between the networks. The E-NNI also defines the protocol exchange that is needed in the interconnected MENs. [1]

Internal Network-to-Network Interface (I-NNI) (figure 2) interconnects Network Elements (NE) inside an MEN service providers network. The I-NNI provides reference point for two directly connected NEs. The I-NNI provides the protocol exchange that exists between NEs inside the MEN. [1]

Network Interworking Network-to-Network Interface (NI-NNI) (figure 2) is an interface that can be used to support Ethernet service and virtual connections over transport networks not involved in end-to-end Ethernet service. The NI-NNI is used to preserve the ETH layer frames (as payload) while transported over transport networks not directly supporting Ethernet transport flows. The NI-NNI defines the protocol exchange needed to connect MENs to other transport networks. [1]

Service Interworking Network-to-Network Interface (SI-NNI) supports interconnection of MENs with other service enabling technologies (e.g., Frame Relay, ATM, IP, etc.). The SI-NNI defines the reference point and protocol exchange between MEN and another service network. [1]

B. MEN Functional Elements

Ethernet Virtual Connection (EVC) (figure 3) associates two or more UNIs to each other. The Metro Ethernet architecture supports three types of EVCs. These are point-to-point EVC, multipoint-to-multipoint EVC, and rooted-multipoint EVC. The first of these EVC types can be used to associate not more than two UNIs between each other, the second type supports association between two or more UNIs, and the rooted-multipoint EVC is an asymmetric association between one or more root UNI(s) and at least one leaf UNI. In the rooted-multipoint EVC root UNIs can send ETH frames to any UNI but the leaf UNIs can only send ETH frames to root UNIs not to other leaf UNIs. The point-to-point and multipoint-to-multipoint EVCs are symmetric connections and bi-directional meaning that UNIs associated with the EVC can send ETH frames to each other. UNI must have at least one EVC to be able to communicate with another UNI. One UNI may have more than one EVC in it but this demands service multiplexing to be applied in the UNI. [3], [2]

Each UNI maps ETH frames into the EVCs with the use of Customer Edge VLAN ID (CE-VLAN ID) per EVC map. In this map every CE-VLAN ID is mapped to one EVC (UNI can support bundling when all or some CE-VLAN IDs are mapped to one EVC).In addition, UNIs must support association of untagged frames to an EVC. The CE-VLAN ID value is carried in the 802.1QTag field that holds 12 bit space for VLAN-ID. Because of the 12 bit restriction only 4095 CE-VLAN IDs can be supported at each UNI. [3], [2]

The CE-VLAN ID per EVC maps are predetermined and normally MEN service providers dictate these maps to subscribers UNIs. The Metro Ethernet Forum has not specified a dynamic way of setting up an EVC with a purpose built signaling protocol. However, this is seen as one of the future work items. [2]

APP to ETH Adaptation Function (EAF) (figure 3) is a class of processing entities that provide adaptation of the APP layer PDUs to ETH layer. EAFs are application (e.g., IP, voice, video, TDM, etc.). specific entities. The EAF is the logical interface between the APP and ETH layer. The EAF consists of source and sink processes. The former is responsible for LLC PDU formation, EtherType allocation, padding, and multiplexing adapted client PDUs to EFTF function. The later is responsible for de-multiplexing client PDUs from EFTF, EtherType processing and decapsulation, and LLC PDU extraction and forwarding to client processes. [3]

ETH Flow Termination Function (EFTF) (figure 3) is a functional entity that creates and terminates ETH network flows. It also functions as a protocol interface between APP and ETH layers. The EFTF consists of source and sink functionality. The source functionality is used for ETH frame preparation (destination and source MAC addresses, 802.1QTag, and user data preparation), Formatting of ETH PDUs, and forwarding the PDUs to Ethernet flow domain. The sink functionality is responsible for receiving the ETH PDUs from Ethernet flow domain, extracting the user data, and forwarding the user data to EAF. [3]

ETH Flow Conditioning Function (EFCF) (figure 3) is a processing entity that is used for in general conditioning (meaning classification, filtering, metering, marking, shaping, and conditioning) subscriber flows into and out of Ethernet flow domain. The EFCF processing entity consists of ingress and egress functions where the former operates on flows from a MEN and later on flows to a MEN. The egress processing entity includes reception of service frames from subscribers Ethernet flow domain, multiplexing of these service frames into one or more ingress flows, ingress flow conditioning based on metering, marking, and policing, and forwarding the ingress flows to the TAF of the UNI-C. The ingress processing entity is responsible for reception of egress service frames from TAF of the UNI-C, egress service frame classification and conditioning and forwarding the egress service frames to subscribers Ethernet flow domain. [3]

ETH Subscriber Conditioning Function (ESCF) (figure 3) is a processing entity responsible for conditioning of subscriber flows into and out of service providers Ethernet flow domain. The egress process of the ESCF is responsible for reception of ETH frames from the Ethernet flow domain of the service provider and classification of these frames. In addition, the egress process is responsible for forwarding these frames towards the TAF of the UNI-N. The ingress process of the ESCF is responsible of receiving ETH frames from the TAF of the UNI-N, classification of these frames, Class-of Service instance determination, ETH frame conditioning, ETH service frame shaping per resource management requirement, and forwarding egress ETH frames towards the Ethernet flow domain of service provider (EEAF). [3]

ETH Provider Conditioning Function (EPCF) (figure 3) does conditioning of ETH flows between two MENs. The ingress process of EPCF is responsible of receiving ETH frames from TAF of the E-NNI and forwarding these frames to the Ethernet flow domain (EEAF) of the service provider. The ingress process also does classification, CoS instance determination, ingress frame conditioning. The egress process of EPCF is responsible for receiving ETH frames from EEAF and forwarding them to the TAF of the E-NNI. The egress processing also includes service frame classification and conditioning. [3]

ETH EVC Adaptation Function (EEAF) (figure 3) is a processing entity that adapts ETH frames into and out of EVCs. The EEAF is responsible of instantiation of an EVC. The EEAF consists of separate source and sink processes. The source process maps conditioned ETH frames into their EVC PDUs and forwards the PDUs towards the EVC(s). The source process of the EEAF also adapts the subscriber CoS ID into service provider CoS, multiplexes ETH frames from various service instances into their corresponding EVC, and buffer management and scheduling based on the CoS. The sink process of the EEAF handles the de-multiplexing ETH frames from the EVCs into their service flow instances and forwards these frames to TAF of the UNI or E-NNI (ESCF or EPCF). The sink process also changes the CoS of the service provider back to the CoS information of the subscriber. [3]

ETH EVC Termination Function (EETF) (figure 3) creates and terminates EVC trails. To handle EVCs the EETF also handles the instantiation of EVCs. The source process of EETF forwards the adapted PDUs towards the Ethernet flow domain of the service provider. The sink process receives EVC PDUs from Ethernet flow domain of the service provider and forwards the frames to EEAF. [3]

ETH Connection Function (ECF) handles the steering of EVC PDUs inside the MEN. The ECF switches traffic through ETH links to create point-to-point and multipoint connections. [3]

ETH to TRAN Adaptation Function (TAF) (figure 3) represents processing entities responsible for adaptation of ETH frames to serving TRAN layer PDUs. The TAF is technology specific as there can be multitude of server layer networking technologies used in MENs that can be used to instantiate ETH links. The source process of TAF handles buffering and scheduling of TRAN PDUs, if needed allocation of VLAN ID, payload padding, generation of service frames, encapsulation/encoding, multiplexing EVC PDUs to ETH link, rate adaptation, and insertion of the EVC PDUs to data stream payload as TRAN layer signals. The sink process of the TAF does Ethernet MAC frame FCS verification, Ethernet MAC frame filtering of subscriber flows not intended to be passed to the UNI, extraction of EVC PDUs from the TRAN layer signals, and de-multiplexing of encapsulated EVC PDUs. [3]

III. Scalability of Metro Ethernet Architecture

The Metro Ethernet architecture provides carrier-grade Ethernet service that supports global, reliable, and scalable connectivity for customer LANs. The scalability issues that have been concentrated on when designing this architecture were service bandwidth scalability, scalability of geographical area where the service can be offered, scalability of connecting big variety of physical infrastructures used by wide range of service providers, and scalability of amounts of subscribers that can be handled in the MENs. This section tries to analyze the scalability issues that arise when Internet like service is to be provided for individual end nodes on top of this architecture. The scalability analysis provided in this section looks into the different architectural elements described in the preceding section (Section II) to see what kind of scalability constraints these elements issue for routed Ethernet service.

The Metro Ethernet architecture was designed to connect two or more distant LANs together to form one VLAN. The connection between the connected LANs is invisible to the nodes inside these LANs. This means that based on the Service Layer Agreement (SLA) all the Ethernet frames originating from a node in one of the LANs are delivered to all the other LANs that are included in the same EVC. For example, if a node A in LAN1 is sending a packet to a node B that is in LAN2, the MEN would deliver this frame to LAN2 and to LAN3 that could also be part of the same EVC.

This is why the basic EVC mapping of UNIs in a MEN must be broken down to form routed Ethernet service into the MENs. The routed Ethernet service will mean that at some point in the traffic path of a ETH frame a routing decision(s) has to be made in order to deliver the ETH frame only to the destination node or only into the LAN of the destination node. The routing decision should result to a EVC (or similar virtual connection entity) between the UNI of the sending node and the UNI of the receiving node.

As stated before it is not possible to globally route an Ethernet frame based on the MAC addresses information. This is because the MAC addresses are not globally unique and because of the fact that the Ethernet architecture is flat (single hierarchy level). This means that there is simply not enough information for a switch/bridge/router to make a routing decision based on the Ethernet frame content. Additionally, it is hard to see how globally unique routing tables could be put together with just MAC address information. In the IP architecture, IP address itself holds enough location information to route the IP packets through series of IP routers. The IP routers are connected to each other to form a routing hierarchy that has multiple levels of hierarchy. In this sort of a architecture routers that do not have enough routing information to make a specific routing decision can always pass the IP packet to higher hierarchy level. The top level of the hierarchy has been built up from routers that have huge routing tables basically spanning the globe in order to find global routes for IP packets. One way to go with the routed Ethernet is to build some form of hierarchy into the network to support global routing. Additionally, node location and identity information must be separated and some form of global location information needs to be put together, for instance, from the node MAC address and the location of the LAN the node is in.

If Metro Ethernet architecture would be able to globally route ETH frames between individual nodes another scalability issue would arise from the fact that in the MEN all the traffic generated by the nodes in a one LAN are in MEN domain seen as one ETH flow. This is why the CE-VLAN ID per EVC mapping in the MEN can be seen as quite straight forward task. However, in routed Ethernet destination of all the traffic flows from one node might be going to different LANs and different nodes. This means that EVC need to be dynamically mapped to some IDs and there need to be scalable amount of these IDs. As there can in the carrier grade Ethernet be only 4095 mappings between CE-VLAN IDs and EVCs.

A. Scalability Issues in MEN Reference Points

The most important reference point in the Metro Ethernet architecture is the UNI. This reference point connects the subscriber domain into the service provider domain. In addition, the Metro Ethernet traffic model is designed for traffic switching in the MENs between two or more UNIs. The traffic switching path between two UNIs is meant to be predetermined and a Spanning Tree Protocol (STP) is used to prepare loop free switching path between the two (or more) connected UNIs. One scalability constraint in UNIs is also the amount of CE-VLAN IDs that can be used to differentiate traffic flows originating from one UNI. [4], [5]

Because the EVCs that connects the UNIs together are made manually by the service provider per subscription, the EVC model does not scale to facilitate routed Ethernet model as it is. In a routed Ethernet the connections between UNIs, in more detail between two nodes, need to be made when either a new communication between two nodes is started or per Ethernet frame sent between these two nodes.

In addition, it is not clear whether the STP used in the Metro Ethernet today can facilitate Spanning Trees (ST) that span over any number of MENs. If the ST are only valid in one MEN this will result into a traffic model where EVC between two UNIs will be made from many different STs in different MENs. These EVCs need to be set up manually by number of service providers before any traffic between UNIs can be sent.

To extend the Metro Ethernet architecture to routed Ethernet service new UNI type needs to be specified. The new type would specify that this UNI is capable of routing Ethernet frames between other routing UNIs. The routing UNI would not perform the routing service with the use of CE-VLAN ID/EVC mapping. This would need a new type of mapping between subscriber flow and service provider virtual connections. This mapping could be performed for instance with the source and destination MAC address to a virtual connection.

Similar scalability issues as described above can be seen for also the E-NNI that needs to connect two MENs together. If the EVCs spanning over the E-NNI reference point are manually mapped to E-NNI equipment by the service provider the model does not scale to allow routed Ethernet service. This means that E-NNI interface needs some form of automated participation mechanism in the routing done in MEN domain.

Similar mechanism for participation in the routing decision need to be also designed for NI-NNI and SI-NNI reference points. In these reference points, however, the routing signaling needs to be tunneled/encapsulated over the network as otherwise this signaling could mix up the internal workings of the service provided by these networks.

B. Scalability Issues in MEN Functional Elements

The EVC has several scalability issues as it is designed right now in the Metro Ethernet architecture. The EVC is scalable for carrier grade Ethernet service but it is not really usable at all for routed Ethernet service. This is because of its static nature. Service providers setup EVCs per subscription and it is not needed to be changed until new UNIs are connected to it or it is removed completely. The EVC points to a specific ST(s) that can be used to switch the ETH frames to their destination.

The EAF and the APP layer in the carrier grade Ethernet is functioning in the Internet scale, as there is IP layer services that this layer can use like ARP, DNS, DHCP, Neighbor Discovery (in IPv6). These functions work as in normal LANs when MEN just connects two or more LANs to form a bigger LAN. In addition, for the APP layer the whole Metro Ethernet architecture is invisible. For instance IP layer can function in a normal fashion by learning the default gateway/router and pass all the global IP traffic to this node.

In a routed Ethernet however none of these service are needed nor present. This means that even before a node is able to transmit a frame to a destination it needs some way to determine at least the MAC address of the destination. For this a totally new way of resolving destination addresses and location information with the identity of the destination need to be designed. In addition, the next functional element in the chain the EFTF needs to know the results from this resolving process to be able to form the ETH frames.

The EFCF that is a functional element in the UNI-C and the ESCF in the UNI-N in addition to the EPCF in E-NNI are designed to multiplex and condition all the traffic coming/going from/to subscribers flow domain into a single flow that is passed to the flow domain of the service provider. In a routed Ethernet the traffic flows from/to subscribers domain cannot be multiplexed into a single flow. All the flows that are coming from and going to subscribers domain need to be kept separate flows also inside the MENs. This is because these flows will have separate paths through the MENs also. This means that most of the functionality in these elements need to be disabled or totally re-designed to support service that can be used to perform the needed functions (like traffic shaping) for multiple separate traffic flows going through these elements.

The EEAF and EETF that reside in UNI-N and in E-NNI are responsible for mapping the subscriber flows to EVCs. Because the EVC is not usable in a routed Ethernet service these functional elements need to be re-designed completely or the routing service need to be added on top of the carrier grade service. The EEAF needs to be in charge of initiating the routing between the UNIs (nodes) that are communicating. The routing protocol to be used in the routed Ethernet is out of scope of this document. However, the routing protocol used will dictate how this functional element needs to be changed/designed. If the routing protocol construct EVC like virtual connections or if the routing is done by next hop mechanism where all switches/bridges/routers make their own routing decision, it will effect the functionality of the EEAF. For instance if next hop method is used the EEAF and EETF can be quite simple as they just pass the ETH frames to the next hop inside the MEN. However, if virtual connections are used this node needs to keep state information about flow paths and mapping between subscriber flows and these flow paths. In addition, these paths might need some lifetimes and other management and control functionality.

The final functional element considered here is the ECF that handles the switching/bridging/routing inside the MEN. Some IEEE switching and bridging protocols have been defined from this functional element for carrier grade service. However, these protocols will not be enough for routed Ethernet service. This functional element will need extensions and design work to provide routed service inside the MENs.

IV. Extending to Routed Ethernet

This section gives some examples of how the identified constraints can be resolved in the Metro Ethernet architecture in order to provide routed Ethernet service. The view point is that if IP protocol is not used in the Internet for nodes using routed Ethernet then everything from node attachment to access networks to frame delivery between nodes need to be changed.

A. Node Attachment and Registration

In a routed Ethernet nodes need some form of default gateway/router for manging global communications that the node has. The UNI-C could perform this function with introduction of new features. To reach the UNI-C the nodes need to know the MAC address of the UNI-C serving LAN the node is connected to. Either the UNI-C could have a specific MAC address that can be pre-configured to nodes or the UNI-C could send periodic advertisements to let all the nodes in the LAN know the MAC address of it.

Because the MAC addresses are not globally routable or unique the nodes need unique identifier and some form of location information that can pin point their location in global scale. The nodes could use, for instance, Host Identity Tag (HIT) specified in Host Identity Protocol (HIP) [6] as an identifier. The use of HIP could be justified further by the fact that with HIP also mobility of the nodes could be supported directly. In addition, HIP defines rendezvous mechanism that can support also global registration of nodes for address and location resolution based on the HIT.

The problem of global location is much more bigger. As the MAC addresses are not globally routable these addresses cannot directly be used as Care-of Addresses (CoA). Still the MAC address needs to be presented in the rendezvous mechanism for a node that uses this mechanism to be able to format the destination MAC address into the ETH frame. This would mean that some location specific information needs to be added to the rendezvous mechanism in addition to the MAC address. This information could be the identifier of service providers MEN or UNI-N or both. In addition, UNI-C could have globally unique identifier as well. How these identifiers would be constructed is out of scope of this document, but some form of authorized body needs to specify these identifiers to service providers and the UNIs inside the service providers domain could be uniquely identified further by the service provider.

When the node is attaching to the LAN it would first either associate itself with the UNI-C and then send a registration message containing it's identity, MAC address, and the location information learned while authenticating with UNI-C. The registration would be sent via UNI-C. Because the UNI-C is the gateway for the node it would be possible that the first message the node sends would be a registration message to the UNI-C. The UNI-C could validate the node and then further send the ID and MAC of the node to a registration server with its own location information. If the registration would be acceptable the acknowledgement would be passed back to the node by the UNI-C. After a successful registration the node would know that it is globally reachable and that it is able to send and receive frames to other nodes.

B. Mapping Node Identities to Location Information

When a node has something to send in a routed Ethernet it would first need to resolve the destination MAC address. This resolution is initiated by the EAF function. The destination MAC address is enough information to format the destination field in the ETH frame in the EFTF of the node and pass it to the UNI-C. The UNI needs additionally the location information of the destination MAC address in order to route the ETH frame to the destination. The node sending the ETH frame needs to use the destination domainname or HIT to get the MAC and location information of the destination. This resolution can be done with the rendezvous mechanism defined in HIP. The node would send a specific frame once again via UNI-C that is meant for name/HIT resolution. The UNI-C would pass this resolution further and in successful resolution case an answer would come via the UNI-C. The answer will contain the MAC address of the destination and the location information. The location information could be concealed from the node as only the MAC address is needed by the node. When a answer to a resolution comes to the UNI-C it could immediately start the Ethernet routing procedure as the answer contains the location information of the destination UNI. The routing procedure is initiated EEAF to open virtual connection to the destination UNI of the destination node. In addition, to the opening of the virtual connection the UNI-C will map the source and the learned destination MAC addresses to the new virtual connection.

When the first ETH frame is received by the UNI-N with the specific MAC addresses stored in the MAC to virtual connection map the frame is passed to the correct switching path by the EETF. If there would be no entry in the map for the MAC addresses the EEAF would initiate the routing and the frame would be buffered.

C. Ethernet Level Routing

To enable routing in Ethernet level the routers (switches, bridges, routers, etc.) participating in the routing of a ETH frame need some way of learning about addresses and paths of each other. This will demand some form of routing signaling to be passed between these routers. In addition, the routers need to be divided into border routers (UNI, E-NNI, NI-NNI, SI-NNI) and core routers (the switches and bridges) inside the MENs.

The border routers need to advertise them selves to let every other border router in the MEN to know the presence of other border routers. This signaling could include the type of the router and some other needed information like UNI or network aggregation information that can be reached behind this router. In addition, the border routers that connect different networks (E-NNI, NI-NNI, and SI-NNI) need to have signaling with their peer routers in the other MEN. Via this signaling specific virtual connections could be requested or learned.

When a UNI-C (EEAF) initiates a construction of virtual path it would first check whether it already knows the path to the destination UNI. If yes it could reuse this information. If not it would first check whether it has learned prior that the UNI is in the the same MEN. If the UNI is in the same MEN a ST could be used to reach it. When the UNI is not in the same MEN the UNI could check if some E-NNI is advertising an aggregate route to the UNI if this functionality is supported or try to request path with a specific signal via different border routers that connect the MEN to other MENs. While the path request message is being forwarded through the border routers, these routers could add state for this virtual connection to their memory after they see the answer coming back that the UNI has been found. After the originating UNI receives the answer the virtual connection can be regarded as open and any ETH frame that is mapped to it can be sent through it. The virtual connections need some form of lifetime or other control and management mechanism to refresh and remove them so the state information doesn't over flow the memory of the routers.

Another way of routing the ETH frames would be the next hop method. In this case the ETH frame would need an additional field to carry the location information of the destination. This however, would demand big changes to all the existing switches and bridges that are designed to work with a specific format of IEEE 802.1 frames. One option would be to add this information to the first bytes in the payload and routers that need this information could read it from there. This approach would then lead to similar kind of routing architecture that is used in IP routing today. The MENs would need to be constructed in a hierarchical levels and any router that cannot route a ETH frame based on the location information would pass the frame to the higher level in the architecture.

V. Conclusion

The scalability issues identified in the carrier grade Ethernet service described in this document are seen as constraints when designing a routed Ethernet service. The biggest scalability issue in Metro Ethernet architecture comes from the fact that EVCs are meant for mapping distant LAN segments to each other. This means that the distant LAN segment look like the same LAN to the nodes connected to one of the LANs. This traffic model itself is opposed to the Internet or routed Ethernet service architecture. Also, the functional entities in the Metro Ethernet architecture are designed for carrier grade service meaning that these entities either have functionality not needed for routed Ethernet, or the functionality must be updated in order for these elements to be usable for routed Ethernet, or new functionality is needed to provide routed Ethernet service.

This document also gave some example solutions for the scalability issues in the carrier grade Ethernet service. These solutions must be worked upon in the future to be able provide the routed Ethernet service with Metro Ethernet architecture. The most important issues that must be solved are the Ethernet routing protocol design, how this routing is initiated in the UNIs and how it is handled in the other routing entities in the MENs. This Ethernet routing protocol can be used to break the traffic model of carrier grade Ethernet to a traffic model that is concentrating on separate flows between Ethernet nodes. In addition, some form of registry must be defined for the nodes subscribing to the routed Ethernet service. This demands node identity and location information to be registered together with the MAC address of the node. For node identity this document proposes the HIT of HIP because on top of providing the identity for Ethernet nodes, HIP also provides registry service and mobility protocol features that can be reused. The proposals for location information in this document is seen very preliminary ideas and the schematics of providing global location information to an Ethernet node need more work.

It is seen that with sufficient Ethernet routing protocol used in MENs, changes to the functional elements of the Metro Ethernet architecture, and adding global communication enabling functions in place, the routed Ethernet service can be achieved with the Metro Ethernet architecture.

References

[1] MEF4: *Metro Ethernet Network Architecture Framework Part 1: Generic Framework*, Metro Ethernet Forum, May 2004

[2] MEF10.1: *Ethernet Service Attributes Phase 2*, Metro Ethernet Forum, November 2006

[3] MEF12: *Metro Ethernet Network Architecture Framework Part 2: Ethernet Services Layer*, Metro Ethernet Forum, April 2005

[4] Chiruvolu, G, *Issues and approaches on extending Ethernet beyond LANs*, IEEE Communications Magazine Vol. 42, Mar 2004

[5] Ali, M., *Traffic engineering in metro Ethernet*, IEEE Network volume: 19, March 2005

[6] P. Jokela ed., *Host Identity Protocol*, Internet-Draft, work in progress, IETF, October 30, 2007