Flexilink

Support for internetworking

Clean slate architecture

In this section the “vision”, i.e. how the system would be if built from scratch with all the component networks using Flexilink for the network layer, is described. Migration from, and interworking with, current technology is covered later.

The system is assumed to consist of inter-connected wired networks (ISP, mobile core, CDN, internet exchange, tier 1, …) with UEs connected via wireless PHY standards (802.11, 802.15, 3GPP radio, Bluetooth, DECT-2020, …). Scenarios such as a LAN that includes several Wi-Fi base stations, or UEs that are hard-wired, are obvious extensions but aren’t covered directly to keep the description simpler.

The wireless networks use their respective PHY layers as far as possible. Their MAC layers are adapted to provide the service specified in ETSI GS NIN 004; in many cases that will be possible with current silicon because it can be done as a software upgrade.

Wireless links to UEs

Communication with a UE is assumed to be over a “link” to a base station which is connected to a wired network. The link is set up when the UE comes within range of the base station, and the set-up process includes verifying that the UE is authorised to access the network to which the base station is connected, e.g. Wi-Fi password or mobile SIM. Extensions omitted for simplicity include peer-to-peer links between UEs, multicasting on a wireless LAN, and UEs forwarding data for other UEs that are out of range of the base station.

Each packet transmitted over the link is part of a “flow”. There is one flow that carries signalling (control plane) messages between the UE and the base station. Other flows are set up using signalling messages, in a process similar to the Berkeley Sockets API. The messages include provision for authentication and specifying or negotiating (either between the endpoints or with the network) a wide range of aspects including QoS parameters and the protocol and data format to be used at the higher layers. The destination can be specified as a conventional network address or by other means of identification such as a URL, service name, or content identifier. The set-up message is propagated through the system to the destination entity as described in ETSI GS NIN 005.

Handover

A UE can have links to more than one base station.[1] At any time it can move a flow from one link to another by requesting a flow on the second link with a flow identifier that differs only in the Route Identifier field.[2,3] It then switches to using the second link, and clears the first one down. This is a similar process to re-routing around a failed link on a wired network.

[1] I’m assuming that’s supported by the silicon in this age of MIMO.

[2] The flow identifier would be created by the system code; all the application sees is the socket handle.

[3] Some of the protocol elements that implement the handover will require extensions to ETSI GS NIN 005.

There are three scenarios that can occur: the route of the new flow meets the old one at some point in the network (the “switching point”), or it arrives at the same destination (in which case the destination is the switching point), or it arrives at a different destination. This last case can occur if the destination is identified as a piece of content which is cached at the edge, and the two base stations find different copies, also if it is a service that is available in multiple places. Protocols for accessing such content or services need to be such that any required state is set up by the signalling messages.

The process at a switching point depends on the service and direction, as follows.

For the “basic” service (best-effort, label routed) in the uplink direction, packets arriving on the new route are forwarded in the same way (i.e. to the same output and with the same label) as on the old route. In the downlink direction, forwarding is immediately switched to the new route. When a ClearDown request arrives for the old route, it is not propagated to the rest of the route.

For the “guaranteed” service (low-latency, TDM-like) in the downlink direction, the normal multicasting mechanism is used to send incoming packets on both routes. The process in the uplink direction is more complex; once the new part of the route is established its slot timing must be adjusted to match the existing route and switching to the new route co-ordinated with the UE.[4,5]

Peering connections between networks

Signalling is the same on connections between networks as within networks. However, external ports should be configured as such before the cable is connected.

Routing information is exchanged using signalling messages. The process is not currently defined,[6] but will use one of the transaction types that are shown as reserved in Table 5.4.1 of GS NIN 005.

The information is only required when processing signalling messages, principally FindRoute transactions, not for packet forwarding. It can therefore be kept in dRAM, avoiding the space constraints of the TCAM memory used for BGP tables. Information for some address types can be kept in servers rather than copied to every network node, for instance something similar to a DNS look-up will be needed for URLs.

Security aspects

A server receiving a flow set-up request can verify the client’s identity before making the forwarding plane connection, and ought to be able to assume that the flow remains attached to that client until it is cleared down. The handover process, particularly between different networks (e.g. mobile to Wi-Fi), will need to ensure that that is the case. This might require an option to limit which network nodes can act as switching points, e.g. restricting it such that if the handover cannot occur within the network to which the client is currently connected (using similar procedures to those used in current mobile networks) the route will be replaced all the way to the server.

[4] There are implementation-dependent constraints on slot positioning, for instance in the gigabit implementation an allocation period is 1 ms but the forwarding buffer only holds 16 µs of data so each outgoing slot must be within 16 µs of the slot in which the packet arrives.

[5] It gets worse if the stream is multicast, so I think there need to be restrictions here; see also under “Security aspects”. One option is always to re-route all the way to the destination UE (or at least to its base station) in which case the process gets a lot easier, for instance the base station can send packets from both routes to the UE with the same label provided it doesn’t forward empty slots (because the inactive route will be sending all empty slots, and we don’t want those taking up space on the wireless link; this ought to be a requirement of GS 004 though I don’t see it in the draft).

[6] There being only 30 switches in existence, and with the routing protocols being quick to execute and robust as regards loops etc, the prototype implementation simply flooded FindRoute requests to all nodes.

Flow set-up messages support information about the networks across which a route passes. Peering connections can be configured to add information including how trustworthy the remote network is considered to be and in which jurisdiction the equipment is located.

Clocks and timing

Flexilink networks include several different timing references. As far as possible, they are kept independent of each other and avoid the requirement to synchronise physical clocks.

Slot allocations for the “guaranteed” service are controlled by an arbitrarily-chosen node and propagate through the network by each node starting outgoing allocations at a defined point in the incoming allocations on one of its links. This provides system-wide hard alignment to support the TDM-like routing while allowing each link to use a local clock; differences in frequency only affect the number of idle symbols between frames.

Absolute time similar to PTP is distributed as part of the framing.

There are several methods for distributing “house clock” information to which audio and video clocks can be synchronised. The clocks can also be recovered from the incoming packet flow, though the requirement for good stability and jitter rejection means they take several minutes to stabilise when a new flow is connected.

Migration path

Flexilink can be tunnelled over IP and vice versa. At Birmingham City University we are providing the Flexilink service over fibre links that were installed to carry IP over Ethernet. A Flexilink switch is inserted between the Ethernet switch and the fibre at each end of the link; a “basic service” tunnel is configured so that the Ethernet packets are carried between the two Ethernet switches exactly as before, except to the extent that Flexilink traffic takes some of the capacity. Packets can be routed between IP and Flexilink flows, or the two networks can be kept completely separate to keep the Flexilink network secure. This configuration could provide a way of gradually introducing the new services into fixed networks such as CDNs.

ETSI GR NIN 002 discusses how Flexilink can be used with 3GPP radio. The most likely scenario would be a private network for an application with a requirement for efficient use of spectrum or controlled latency (or both). The network can be linked to an IP network by a gateway which stores the IP addresses etc for each flow, in a table similar to that used for the NAT function in Figure 2 of ETSI GS NGP 001.

In general, a Flexilink internetwork can be built up in parallel with existing technology with cross-connection between the two where required, in the same way that internet exchanges formed a parallel system to the original structure. The signalling messages setting a flow up can include requirements for QoS parameters such as latency, and report the values actually achieved; this allows an application to adapt its performance according to whether a media flow has had to be routed over a legacy network.