CHAPTER 2

IEEE 802.16 Standard

In this chapter, the IEEE 802.16 standards will be introduced. Both PMP and mesh mode will be described, while the majority context shall focus on mesh mode. The IEEE 802.16 is a time-slotted network. The time is divided into frames, and a frame is divided into timeslots, which serve as the very unit of transmissions. However, the detailed definition of frames defers depending on the transmission mode chosen.

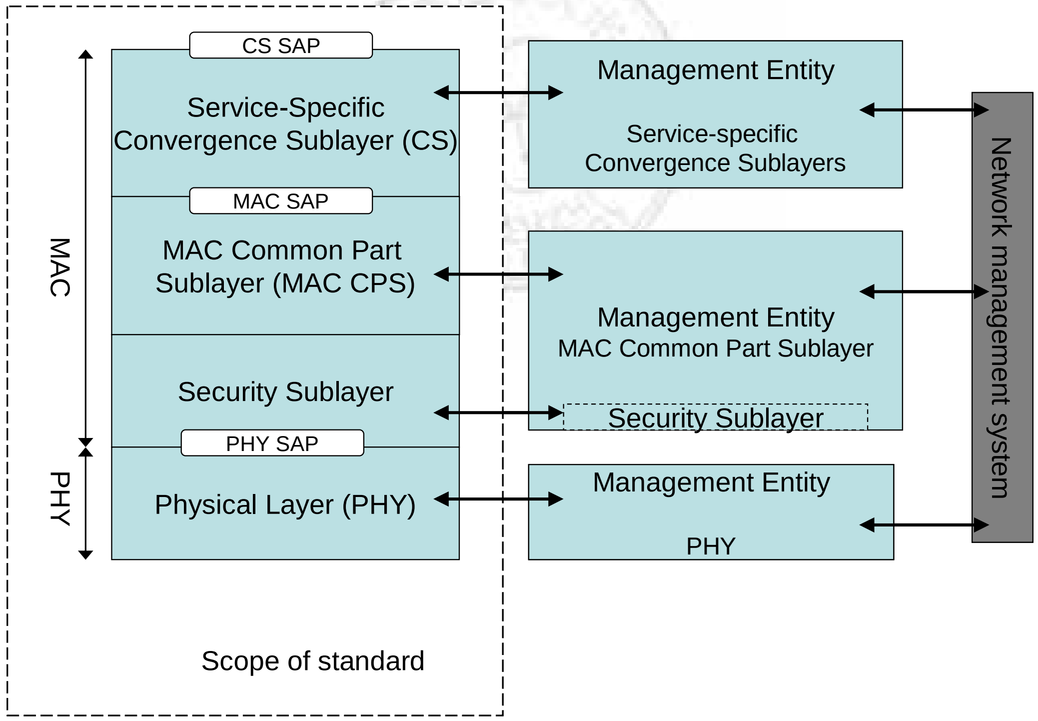

Figure 2.1: Scope of IEEE 802.16 Standards.

The IEEE 802.16 standard includes both physical and MAC layer specifications. The MAC layer comprises three sublayers: CS (Convergence Sublayer), MAC CPS (Common Part Sublayer) and Privacy Sublayer (also called Security Sublayer). SAP stands for Service Access Point, where the service of a lower layer is available to its next higher layer. The MAC and PHY layer model is as Figure 2.1.

The Service-Specific Convergence Sublayer (CS) provides transformation or mapping of external network data, received through service access point (SAP), into MAC SDUs (Service Data Units) used by MAC Common Part Sublayer. This includes classifying external network packets into SFID (Service Flow Identifier) and CID (Connection Identifier). Therefore, different CSs are needed when the system is dealing with various upper-layer protocols. The MAC CPS provides the core MAC functionality of system access, bandwidth allocation, connection establishment and maintenance. And, Security Sublayer provides authentication, secure key exchange, and data encryptions.

IEEE 802.16 supports different transmission rate by different burst profiles. According to these burst profile, modulation can be chosen from QPSK, 16 QAM or 64 QAM.

2.1. IEEE 802.16 PMP mode

In IEEE 802.16, the node with wired connection is called “Base Station”, abbreviated as BS. The client-side node that associates with BS is called Subscriber Station, or SS. In PMP mode, the network architecture is similar to traditional cellular network architecture, where the BS centrally controls the transmissions of its associating SSs. The network architecture is in Figure 2.2. The MAC mechanism can be TDD (Time Division Duplex) or FDD (Frequency Division Duplex).

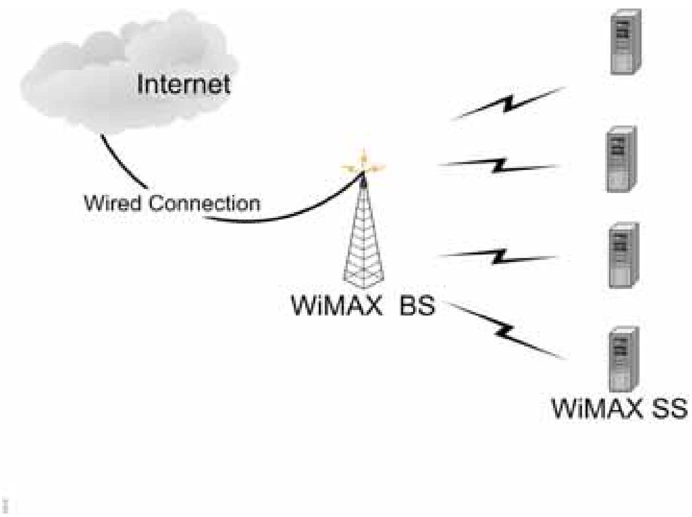

Figure 2.2: PMP Network Architecture.

There are four classes of traffic supported in 802.16 PMP mode networks:

- Unsolicited Grant Service (UGS)
- real-time Polling Service (rtPS)
- non-real-time Polling Service (nrtPS)
- Best Effort (BE).

The UGS is designed to support real-time data stream that comprises fixed-sized packets issued at periodic intervals, such as VoIP packets. This type of packets is sent at a constant interval without any requests, thus, the delay shall be easily calculated.

The rtPS is designed to support real-time data streams consisted of variable-sized packets that are issued at periodic intervals, such as real-time video. The nrtPS is designed to support delay-tolerant data streams, such as FTP transmissions. And, BE is designed to support the data streams that has no minimum data rate requirements.

The nrtPS and BE are designed for non-real-time connections. The nrtPS is to support delay-tolerant data streams with minimum data rate is required. And BE is to support data streams with no minimum service level is required. Table 2.1 shows these four traffic types.

Table 2.1: QoS Levels in IEEE 802.16 PMP Modes.

Class name	Traffic type	Application
UGS	real-time CBR	Voice over IP
rtPS	real-time VBR	real-time video
nrtPS	non-real-time	FTP
BE	non-real-time	HTTP

Each frame has a downlink subframe and an uplink subframe for downlink and uplink transmissions, respectively. Downlink transmission means that the direction of the transmission is flowing from BS to SS. Uplink transmission mean the transmission is happening in an opposite direction. In the beginning of downlink subframe, DL-map and UL-map fields are placed. With DL-map and UL-map, each SS can have the information of the exact time to access the medium for downlink or uplink transmissions. In the uplink transmissions, the BS polls each SS or connections, depending on the granularity. When the network is heavily-loaded, polling may not be efficient, especially for low-priority flows. Therefore, the bandwidth-request contention period is in the beginning of uplink subframe. This period is used for low-priority flows to get the opportunities of sending BW-req. to BS by contention. The frame structure of PMP mode is in Figure 2.3.

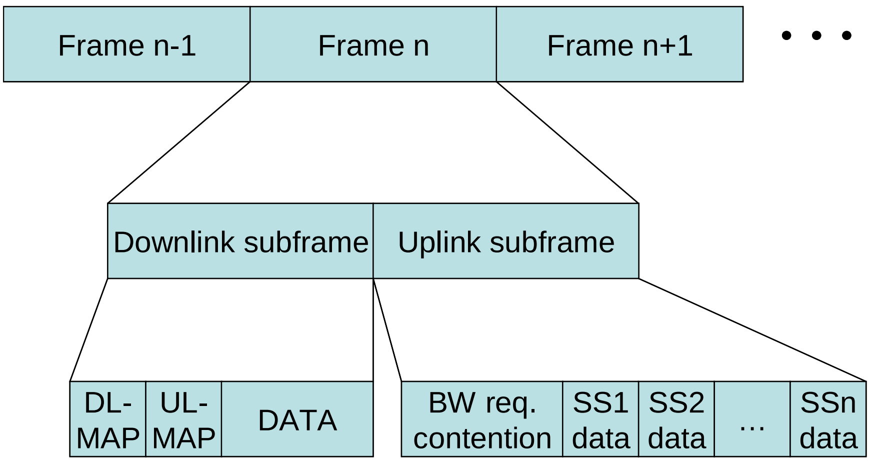

Figure 2.3: PMP Frame Structure.

2.2. IEEE 802.16 mesh mode

Mesh mode is an optional mode in IEEE 802.16 standard. The main difference between PMP mode and mesh mode is that in mesh mode, transmission can occur between any two stations, while in the PMP mode, transmission can occur only between the BS and SSs. The network architecture is in Figure 2.4.

The terms used in mesh mode are also slightly different from those in PMP mode. The node that has a wired connection to internet is called “Mesh Base Station”, abbreviated as MBS, or simply BS. All the other nodes are termed “Mesh Subscriber Station”, short for MSS or simply “SS”. Also, the “extended neighborhood” means the neighborhood formed by all two-hop neighbors. The uplink and downlink are defined as traffic in the direction of the Mesh BS and traffic away from the Mesh BS, respectively.

There are two modes to access the medium: centralized and distributed. Each mode uses different scheduling or controlling messages to exchange the available resources of the nodes. However, there is no difference in the mechanism used in determine the schedule for uplink and downlink. The QoS is provisioned over links on a message-by-message basis. No strict QoS classification in mesh mode is given as in PMP mode. The QoS level in mesh mode is controlled by three fields of CID (Connection ID):

- Reliability: 1 bit. To indicate if the message needs re-transmission or not.
- Priority: 3 bits. Simply represents the class level of the message.
- Drop Precedence: 2 bits. Messages with larger Drop Precedence have higher drop likelihood during congestion.

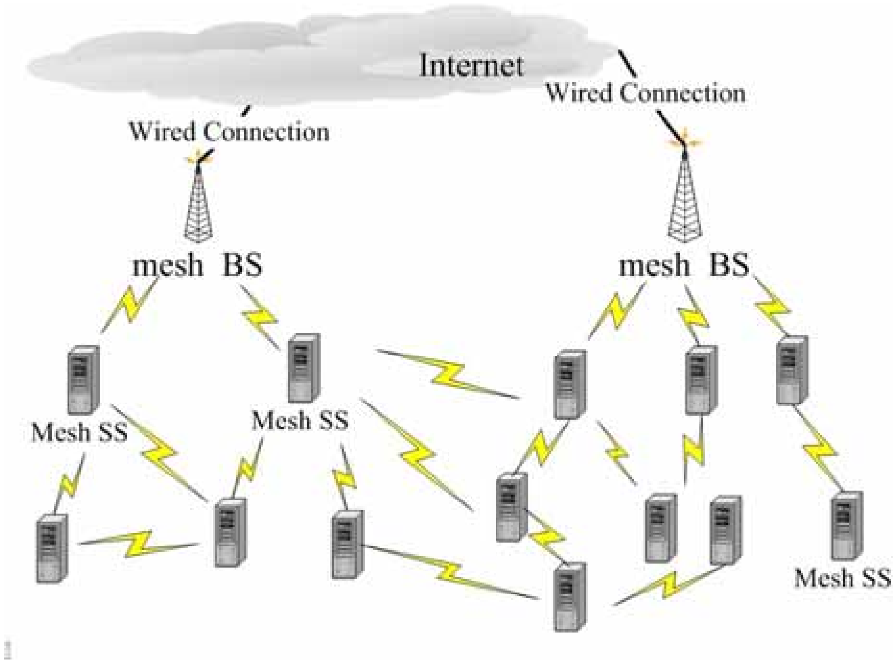

Figure 2.4: Mesh Network Architecture..

Similar to PMP mode, the network is time-slotted as well. Each frame is composed with control subframe and data subframe. The mesh frame structure is in Figure 2.5. The actual data transmission occurs in the timeslots of the data subframes by a scheduling mechanism running in the control subframes.

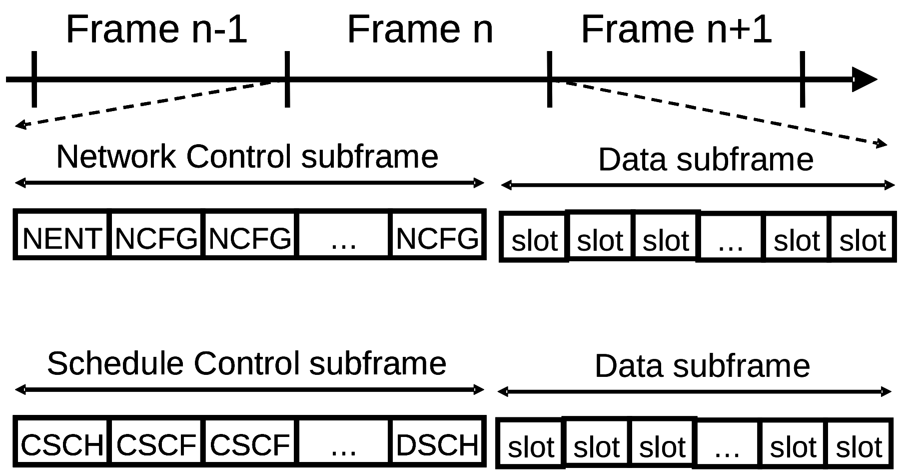

Figure 2.5: Mesh Frame Structure.

The Control subframe includes Network Control subframe and Schedule Control subframe. In Network Control subframe, MSH-NENT (Mesh-Network Entry) message is used to provide the entries of new-coming stations. MSH-NCFG (Mesh-Network Configuration) messages are sent by each station periodically. By MSH-NCFG messages, the basic information about the mesh networks can be exchanged between SSs; the information includes the identifier of BS, hops to BS, neighbor number of reporting station …, etc.

Two scheduling modes are defined in IEEE 802.16 mesh networks: Centralized and Distributed modes. In centralized scheduling mode, the BS is in charge of allocating network resource over links, and all the transmissions must be granted by BS. The scheduling information is carried by the MSH-CSCH (Mesh- Centralized Scheduling) and the MSH-CSCF (Mesh Centralized Scheduling Configuration) messages. And, In distributed scheduling mode, each node sends MSH-DSCH (Mesh-Distributed Scheduling) message periodically to convey the requests and grants of the data timeslots.

The centralized and distributed scheduling mode will be introduced in the following context. As this work is done over distributed mesh mode, the more emphasis will be put on distributed scheduling mode.

2.2.1. Centralized Scheduling Mode

The schedule using centralized scheduling is determined in more of a centralized manner than in distributed scheduling mode. In IEEE 802.16 centralized mesh networks, all scheduled transmissions for the SSs shall be defined by the BS. The BS determines the flow assignments from the resource requests from the SSs. Subsequently, the SSs determines schedule from these flow assignments by using a common algorithm that divides the frame proportionally to assignments. Therefore, the BS in centralized mesh mode acts just as in PMP mode, except that the BS does not have a direct link to the SSs. Centralized scheduling ensures that transmissions are collision-free over the links in the routing tree to and from BS.

The MSH-CSCH and MSH-CSCF are the messages used in centralized scheduling mode. In MSH-CSCH messages, two parameters are carried: *Uplink/Downlink Flow* and *Flowscale Exponent*. Using these parameters, the bandwidth of each link is assigned by the Mesh BS by the following equation.

$$BW_{toBS} = UplinkFlow \cdot 2^{FlowscaleExponent+14}$$

$$BW_{fromBS} = DownlinkFlow \cdot 2^{FlowscaleExponent+14}$$

The MSH-CSCH and MSH-CSCF are both broadcasted by the BS to its neighbors, and the neighbors forward the message if the hop count is lower than a pre-defined threshold. The MSH-CSCH is used to carry the request and grant information. And the MSH-CSCF message is used to update the information of the nodes.

2.2.2. Distributed Scheduling Mode

In distributed scheduling mode, each node regularly transits MSH-DSCH message to synchronize the clock and exchange the schedule information. There are four Information Elements (IEs) in a MSH-DSCH message:

- Scheduling IE
- Request IE
- Availability IE
- Grant IE

The scheduling IE is used to exactly determine the next transmission time of the MSH-DSCH message. The other three information elements are used to reserve the free time slots among the neighbors by a three-way mechanism.

The distributed scheduling mode can be coordinated or un-coordinated. If the mesh network is running under a coordinated scheduling mode, the MSH-DSCH messages can only be sent in control subframe. In un-coordinated distributed scheduling mode, the MSH-DSCH can be sent in both control and data subframes. The differences between coordinated and uncoordinated distributed scheduling are as follows: In the coordinated case, the MSH-DSCH messages are scheduled in the control subframe in a collision-free manner; whereas, in the uncoordinated case, MSH-DSCH may collide. Nodes responding to a MSH-DSCH:request should, in the uncoordinated case, wait a sufficient number of minislots of the indicated availabilities before responding with a grant, such that nodes listed earlier in the request have an opportunity to respond. The MSH-DSCH:grant for confirmation is sent in the minislots immediately following the first successful reception of an associated Grant packet.

The coordinated distributed scheduling information is carried by MSH-DSCH with two parameters of its own and all its neighbors. The scheduling IE includes the following parameters:

- Next Xmt Mx
- Xmt Holdoff Exponent

With these two parameters, the Next Xmt Time, which is the next MSH-DSCH eligibility interval for this node, is computed as the range:

$$2^{\exp} \cdot M \le NextXmtTime \le 2^{\exp}(M+1)$$

, where *exp* is Xmt Holdoff Exponent, and *M* is Next Xmt Mx. If a node discovers that its Next Xmt Time is overlapped with some other stations' Next Xmt Time. The scenario is considered "collision". To prevent that, a *meshelection()* algorithm is given in the standard:

```
Boolean MeshElection( uint32 XmtTime, unit16 MyNodeID, uint16 NodeIDList[]){
        uint32 nbr_smear_val, smear_val1, smear_val2;
        smear_val1=inline_smear(MyNodeID ^ XmtTime);
        smear_val2=inline_smear(MyNodeID + XmtTime);
        For each NodeID nbrsNodeID in NodeIDList Do{
                nbr_smear_val=inline_smear(nbrsNodeID ^ XmtTime);
                if(nbr_smear_val > smear_val1)
                        return false;
                else if(nbr_smear_val == smear_val1){
                                nbr_smear_val=inline_smear(nbrsNodeID + XmtTime);
                                if(nbr_smear_val1>smear_val1)
                                        return false;
                                else if(nbr_smear_val1 == smear_val2){
                                        if((XmtTime is even &&(nbrsNodeID>MyNodeID))||
                                        (XmtTime is odd &&(nbrsNodeID<myNodeID)))
                                                return false;
                                }
                        }
        return true;
        }
```

The inline_smear() function is below:

```
uint    inline_smear(uint16 val){
        val +=(val<<12);
        val ^=(val>>22);
        val +=(val<<4);
        val ^=(val>>9);
        val +=(val<<10);
        val ^=(val>>2);
        val +=(val<<7);
        val ^=(val>>12);
        return(val);
}
```

The *meshelection()* algorithm is similar to a hashing function, only one station shall win the mesh election, while the other lose stations will set their TempXmtTime to the next MSH-DSCH opportunity. And check if any collision occurs, if any, the *meshelection()* is executed individually in these two stations. After the above mentioned procedures, the MSH-DSCH messages can be sent in a collision-free manner.

The request IE, availability IE and grant IE in the MSH-DSCH message are used to convey the channel resources. Both the coordinated and un-coordinated distributed scheduling employ a three-way handshake:

- MSH-DSCH:request is made along with MSH-DSCH:availabilities, which indicate potential slots for replies and actual schedule.
- MSH-DSCH:grant is sent in response indicating a subset of the suggested availabilities that fits, if possible, the request. The neighbors of this node not involved in this schedule shall assume the transmission takes place as granted.
- MSH-DSCH:grant is sent by the original requester containing a copy of the grant from the other party, to confirm the schedule to the other party. The neighbor of this node not involved in this schedule shall assume the transmission takes place as granted.

In the MSH-DSCH:request IE, two fields are defined: Demand Level and Demand Persistence. The Demand Level represents the demand in timeslots. And the Demand Persistence is the persistency for the demands, that is, the number of frames wherein the demand exists. The possible values are 1 to 128 frames. In the MSH-DSCH:availabilities and MSH-DSCH:grant IEs, the parameters are similar: start frame number, minislot start, minislot range, direction, and persistence. The field of start frame number is the eight least significant bit of frame number. The minislot start indicates the first usable timeslot and the minislot range indicates the number of slots for grants. Direction means the timeslot is available for transmission or reception. The persistency field still represents the persistency for availability or grant. The three way handshake is in Figure 2.6.

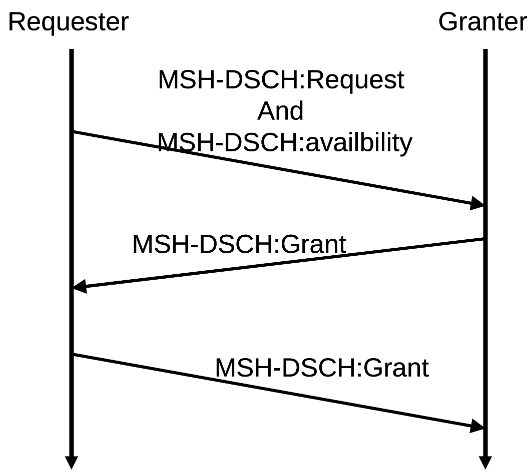

Figure 2.6: Three-way Handshake.

By the three way handshake, each station determines if there is sufficient timeslots when MSH-DSCH:Request is received. If there is enough timeslots, the MSH-DSCH:Grant message will be transmitted to grant the schedule. On the other hand, if there is no enough timeslots, the MSH-DSCH:Grant message is transmitted to reject the request as well. In the IEEE 802.16 mesh networks, the service parameters are defined in CID fields. When the priority of flow A is higher than flow B, the timeslots that reserved for B can be preempted by A. In all, the station first checks the empty slots when receiving a request. If there is no enough timeslots, the station will try to preempt those timeslots that is from lower-priority flows. When there is no available timeslots, the request shall be rejected.