Proposal for Group Scheduling Support in IEEE 802.16m SDD

Document Number: C80216m-08/712
Date Submitted: 2008-07-07
Source: Rath Vannithamby, Shweta Shrivastava, Roshni Srinivasan, Shailender Timiri, Hujun Yin
rath.vannithamby@intel.com
Intel Corporation
Venue: Denver, USA
Re: IEEE 802.16m-08/024, Call for Comments and Contributions on Project 802.16m System Description Document (SDD), DL Control Structure
Purpose: For discussion in TGm and adoption of proposed SDD text the IEEE 802.16m System Description Document

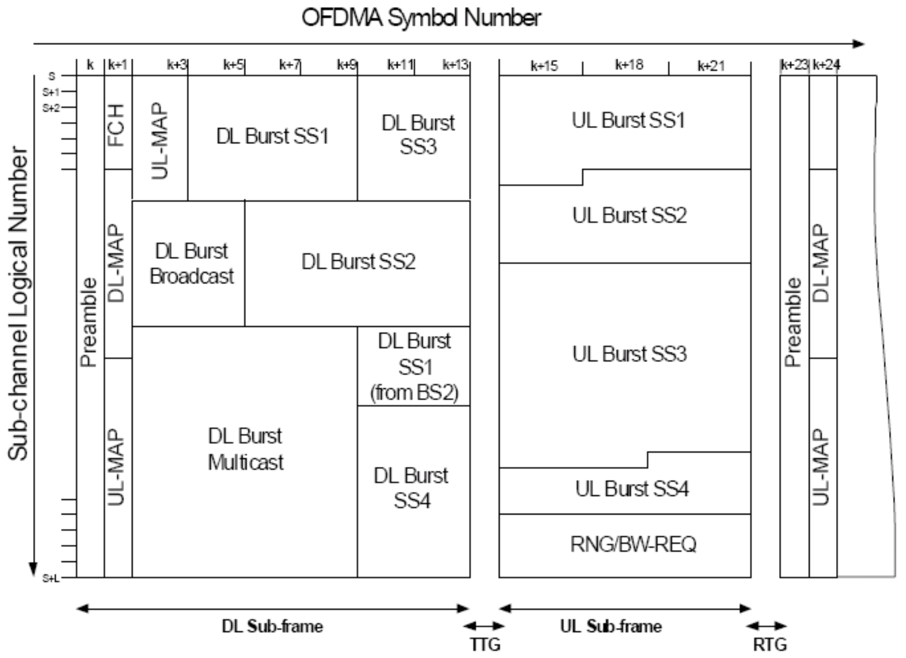
Outline

- Introduction
- Group-based Scheduling
- HARQ Optimization with Group Scheduling
- Simulation results

VoIP over WiMAX

- VoIP is envisioned as an important service provided by WiMAX network
- IEEE 802.16m requirement – 60 VoIP users/sector/FDD MHz → 300 bidirectional VoIP users for 10 MHz bandwidth for TDD.

Challenges with Supporting VoIP over WiMAX



- Motivation: Minimizing the overhead instead of maximizing throughput for VoIP

- DL-MAP and UL-MAP describe all allocations made in a frame.
- MAP Size is proportional to number of allocations.
- VoIP → Smaller Packets → More allocations per frame → Larger MAP Size.
- The MAP is typically sent with most robust (least efficient) MCS → large MAP overhead.
- Simulations show MAP overhead can be as much as 40-50% of DL subframe.

MAP Overhead Reduction with Persistent Scheduling

- VoIP traffic is periodic in nature and vocoders generate constant size packets.
- Resources can be allocated *persistently* in multiple frames rather than dynamically in every frame.
- Resource allocation information can be specified in MAP of first frame and skipped in subsequent frames.
- Persistent resource allocation will work for both DL and UL.

Outline

- Introduction
- Group-based Scheduling
- HARQ Optimization with Group Scheduling
- Simulation results

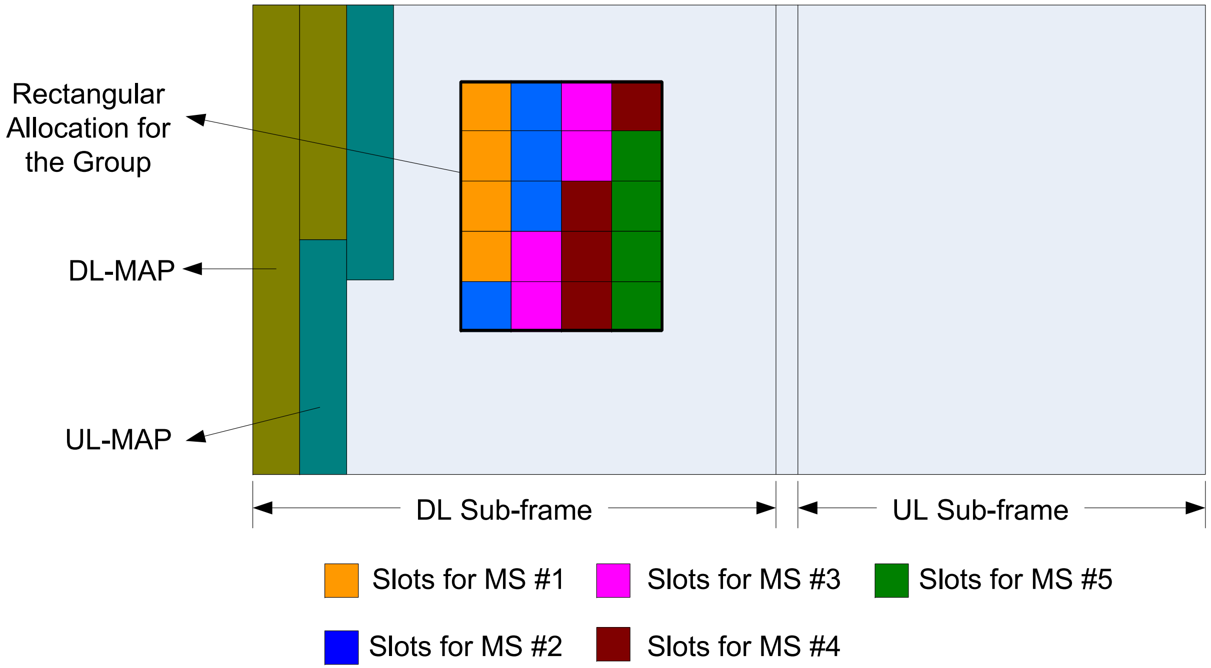
Limitations of Individual Persistent Scheduling

- Since resource locations of users are fixed within the frame, it might result in resource "holes" in the frame resulting in inefficient resource utilization
  - When a user enters silence, a hole is created in the location of his allocation. The hole can only be filled by another user needing same # RBs entering talk at the same time. Otherwise, all subsequent users need to be moved causing large overhead.
  - When a user's MCS changes, his resource requirements change, resulting in reallocation of resources to that user. This can again result in holes problem as above.

Grouping – Efficient Resource Utilization

- Users can be clustered into multiple groups based on certain criteria e.g. similar channel conditions
- Instead of individual allocations, allocations are made to a group of users.
- The location of resources allocated to a user within the resources allocated to a group is fixed.
- Once a user's group and resource location within a group are specified, this information need not be repeated in subsequent frames.
- The amount and location of each group's resources within the frame can be flexibly allocated.
- Grouping enables efficient resource utilization by avoiding allocation gaps in a frame.

Grouping Example

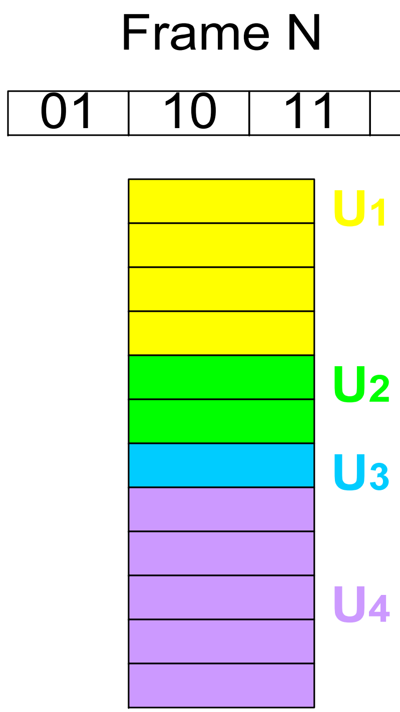


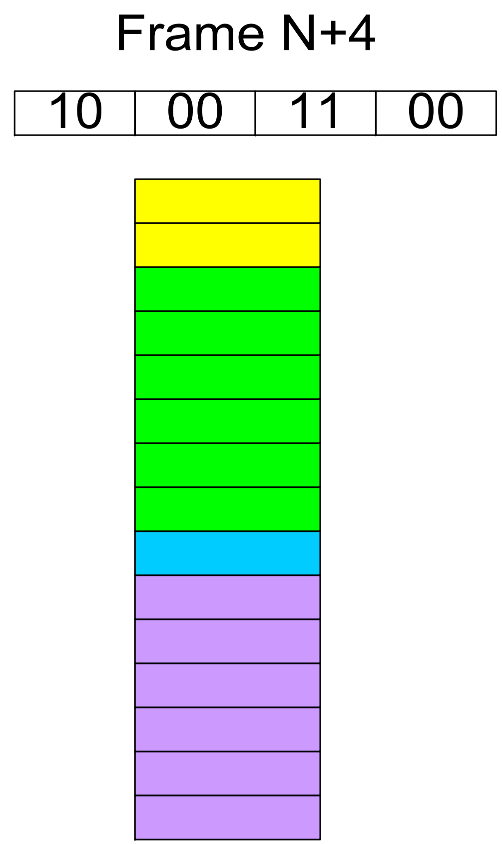
MAP Contents – Dynamic Scheduling (every frame):	MAP Contents – Grouping (first frame):	MAP Contents – Grouping (subsequent frames):
For every user CID MCS & Rep Coding Resource Information	Group ID MCS & Rep Coding Group's resource info *For every user in group* CID Resource Location	Group ID MCS & Rep Coding Group's resource info

Grouping Mechanism

- Group users with multiple adjacent MCS into a single group.
  - Avoids having a large number of groups in a frame.
  - E.g. we can group users with 4 MCS in one group.
  - Since there are a maximum of 19 MCS, there will not be more than 5 groups (per codec).
  - Saves overhead of assigning user to new group when MCS changes.
- Convey MCS of each user in the group using a bitmap
  - Bitmap has 2 bits for every user; this 2-bit code conveys the MCS
  - Mapping of MCS to 2-bit code is done at connection setup.
  - No need to send bitmap in every frame; send only when it changes: saves overhead.
- If users goes into silence:
  - Since activity change is not very frequent, remove user from group; reassign him to a group when he becomes active again.
  - For codecs with frequent activity change, one value of 2-bit code in the bitmap can be used to indicate silence. E.g. 00 can indicate silence and the other 3 values can represent 3 different MCS levels.

Bitmap example





00 MCS_0 require 6 RBs
01 MCS_1 requires 4 RBs
10 MCS_2 requires 2 RBs
11 MCS_3 requires 1 RB

Allocation, De-allocation, Group Change

- Allocation of a new user is done by including the following information for the group the user is joining
  - Position in the bitmap
  - Bits are added in the bitmap for the user to indicate its MCS level
- De-allocation of a user from a group is done by mentioning the position of the user in the bitmap for the group that the user is leaving
  - Other users in the group can calculate their new position in the bitmap from this information
  - The bitmap size reduces accordingly
- When group change is needed
  - De-allocation from the original group and allocation to the new group are done as indicated above

High Level MAP Structure

```
HARQ_DL_MAP_IEs ()
        Resource Information for HARQ region
        Number of groups
        For each group (Group_Allocation_IE)
                Group ID
                Starting Location of Group
                Size of packet
                For each new user
                        CID
                        Position in bitmap
                For each removed user
                        Position in bitmap
                MCS Bitmap
                }
```

Advantages of Grouping and Bitmap

- Grouping allows flexibility in assignment of frame resources
  - Group size is flexible: Users can be easily removed from or added to group if needed, without affecting other users in the group.
  - Group's resource location is flexible: Group's resource as a whole can be moved around in the frame to accommodate changes due to other groups' sizes.
  - Avoids resource wastage due to small gaps in frame caused by users going into silence mode or changing MCS.
- Bitmap allows overhead reduction
  - No need to de-allocate /allocate resources with every MCS change: saves considerable overhead.

Outline

- Introduction
- Group-based Scheduling
- HARQ Optimization with Group Scheduling
- Simulation results

HARQ Optimization with Group Scheduling

- Group-based Scheduling of initial VoIP transmissions saves considerable overhead and improves capacity.
- However, about 30% of the frame allocations could be retransmissions.
- MAP information for HARQ retransmissions can still cause considerable overhead.
- Need a mechanism to reduce overhead for HARQ retransmissions.

HARQ Optimization with Group Scheduling (contd.)

- Frame-synchronized, non-adaptive HARQ is assumed
  - Packet size is the same for new transmission and retransmissions.
  - MCS is the same for new transmission and retransmissions.
  - Fixed delay in retransmitting, say, M frames (M is a function of ACK delay defined in DCD/UCD).
  - If these need to be changed, dynamic HARQ ReTx may be scheduled.
- Solution
  - All parameters related to first transmission are specified in Group_Allocation_IE
  - Mobile stations in the group that have retransmissions can simply refer to the Group_Allocation_IE for the new transmission.
  - A bitmap will convey which MS have retransmissions.
    - Size of bitmap for first transmission (in frame n + M) is equal to the number of MSs in group during first transmission (in frame n). The bits in bitmap represent MSs in the order in which they are allocated for first transmission.
    - For second retransmission, the number of bits in bitmap (in frame n + 2M) is equal to the number of MSs that had first retransmissions. Same pattern is followed for subsequent retransmissions.

Example

Frame n
First Tx

MS # 1, 2, 3, 4 and 5, belonging to a group have first transmissions

Frame n+M

Bitmap

1	2	3	4	5
0	1	0	1	1

MS # 2, 4 and 5 have first retransmissions

Frame n+2M

Bitmap

2	4	5
0	1	0

MS # 4 has second retransmission

High Level MAP Structure

```
HARQ_DL_MAP_IEs ()
        Resource Information for HARQ region
        Number of groups
        For each group (Group_Allocation_IE)
                Group ID
                Starting Location of Group
                Retx_num (2 bits)
                if (Retx_num == 0) {
                        Size of packet
                        For each new user
                                CID
                                Position in bitmap
                        For each removed user
                                Position in bitmap
                        MCS Bitmap
                } else {
                        Retx Bitmap
                }
```

Outline

- Introduction
- Group-based Scheduling
- HARQ Optimization with Group Scheduling
- Simulation results

Simulation Assumptions

- Speech source model compliant with 802.16m EMD
  - Packet size of 44 bytes includes compressed UDP/IP headers (3 bytes), MAC header (6 B) and CRC (2 B) in addition to 33 bytes of VoIP payload.
- Compressed MAP/ Submap structure in compliance with 802.16e
- Uplink is not simulated
  - Uplink MAP overhead is modeled by assuming same number of uplink users as the number of downlink users in each frame
- Synchronous HARQ retransmission
  - Time synchronous, but location flexible within the frame
  - HARQ retransmissions are dynamically allocated (proposed optimizations not incorporated for the results shown)
- SID frames dynamically allocated
- DL:UL ratio 26:21 OFDMA symbols
- Outage 2%, delay budget 50 ms
- Algorithm optimizations for link adaptation, dynamic scheduling of asynchronous HARQ retransmissions and MAP overhead reduction using submaps have not been modeled extensively in estimating VoIP capacity for comparative simulations

Simulation results

	Dynamic Scheduling	Individual w/o repacking	Individual with repacking	Grouping (Grp size = 1)	Grouping (Grp size = 4)
VoIP Capacity (3km/h)	135	146	153	165	182
Capacity Gain	-	8.2%	13.3%	22.2%	34.8%
Per frame Avg. MAP OH (slots) at full capacity	183.8	154.7	157.4	141.8	119.2
Per frame Avg. MAP OH slots Per User	1.36	1.05	1.02	0.86	0.65
Reduction of MAP OH/User	-	22.7%	25%	36.7%	52.2%

SDD Text

Add the following text in Section 11.6.2.3.1.2 of the SDD on Page 43, Line 20:

- User specific control information is used to allocate radio resources by taking account the channel conditions at the MS identified through measurements made at the BS and/or reported by the MS.
- Resources can be allocated persistently to MSs. An indicator may be used to identify persistent allocations and the periodicity of the allocation may be configured.
- A group message is used to allocate resources and/or configure resources to one or multiple mobile stations within a user group. Each group is associated with a set of resources. VoIP is an example of the subclass of services that use group messages.