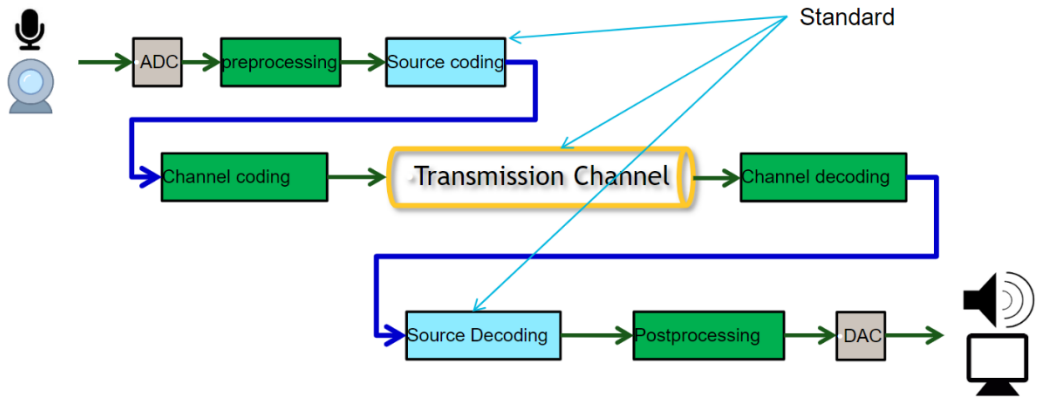
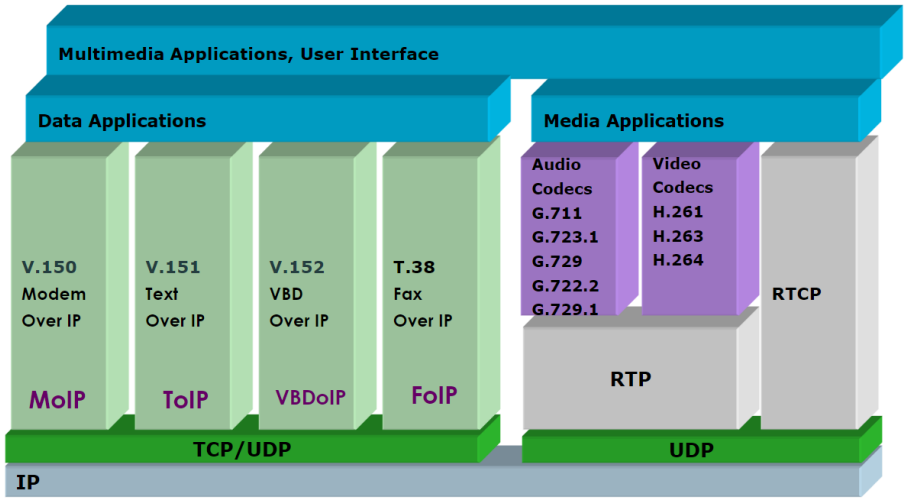
End To End media transmission overview



Audio, Video & Data VoIP protocol Stack



Audio Media Coding Standards Summary

Formal name	ITU-T G.711	ITU-T G.722	ITU-T G.722.1	ITU-T G.722.2, GPP AMR-WB	ITU-T G.723.1	ITU-T G.726	ITU-T G.728	ITU-T G.729	GSM FR (Full-Rate)	GSM HR (Half-Rate)	GSM / PCS 1900 EFR (Enhanced Full-Rate)	3GPP AMR (Adaptive Multi-Rate)
Speech Model? (Y/N)	N	N	N	Y	Y	N	Y	Y	Y	Y	Y	Y
Audio Bandwidth (KHz)	0.3-3.4	0.05-7	0.05-7	0.05-7	0.3-3.4	0.3-3.4	0.3-3.4	0.3-3.4	0.3-3.4	0.3-3.4	0.3-3.4	0.3-3.4
Primary Application	T	T, RN, TC	VC	M, S, T, V, VC, SVD	VT	D	T, D	T, D, SVD, V	M, V	M, V	M, T, V	M, S, T, V, SVD
Bitrate(s) (kbits/sec)	56 64	48 56 64	24 32	8.85 12.65 23.85	5.3 6.4	16 24 32 40	16	6.4 8 11.8	13	5,6	12,2	5.15, 7.4 10.2 12.2
VAD	?	N	N	Y	Y	N	N	Y	Y	Y	Y	Y
CNG	?	N	N	Y	Y	N	N	Y	Y	Y	Y	Y
DTX	?	N	N	Y	Y	N	N	Y	Y	Y	Y	Y
Frame loss concealment	?	?	N	Y	Y	?	Y	Y	Y	Y	Y	Y
Embedded Scalability	N	Y	N	N	N	N	N	N	N	N	N	N
Sample Rate (kHz)	8	16	16	16	8	8	8	8	8	8	8	8
Frame Length (msec)	0,125	0,125	20	20	30	0,125	0,625	10	20	20	20	20
Algorithmic Delay (msec)	<< 1	1,625	40	25	37,5	0,125	< 2	15	20	24,4	20	25

Primary Applications:

DVD: DVD-video, **M**: Mobile, **RN**:Radio news, **S**:Streaming, **SVD**:Simultaneous voice/data, **T**:Telephony, **TC**:Teleconferencing, **TV**:Television, **V**:Voice on IP, **VC**:Video conferencing, **VT**:Video telephony, **W**:Wireless LAN

Wideband codecs

Wideband codec standard	Bandwidth (Hz)	codec Bit rate (Kbps)											
ITU-T G.722	50-7000	48	56	64									
ITU-T G.722.1	50-7000	24	32										
ITU-T G.722.1 Annex C	50-14000	24	32	48									
ITU-T G.722.2, 3GPP AMR-WB	50-7000	6.6	8.85	12.7	14.3	15.9	18.3	19.9	23.1	23.9			
ITU-T G.729.1	50-7000			14	16	18	20	22	24	26	28	30	32
OPUS (non ITU)	Up to 20kHz	6	..									..	510

16 sur 114

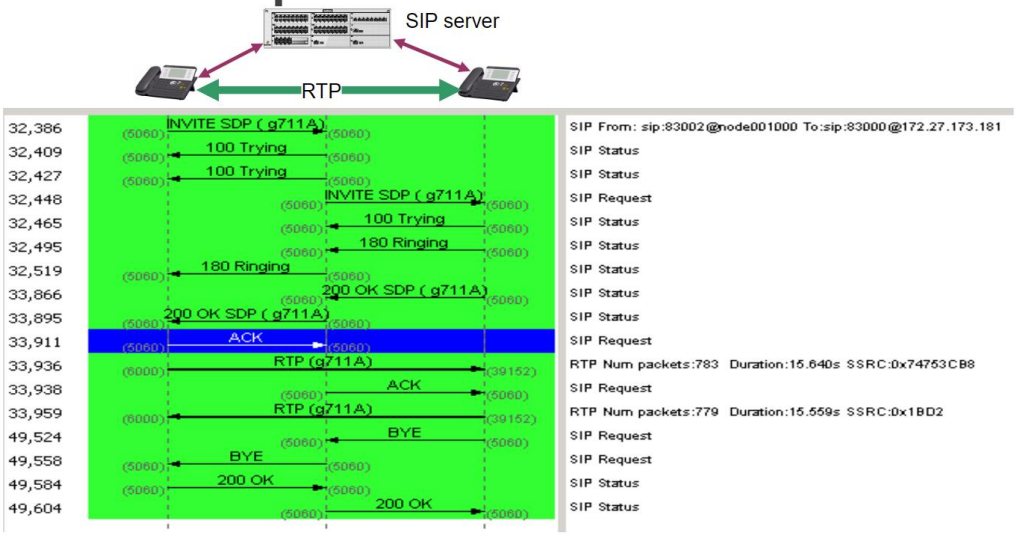
Video Media Coding Standards Summary

Formal name	ITU-T H.261	ISO/IEC 11172-2	ITU-T H.262, ISO/IEC 13818-2	ITU-T H.263	ISO/IEC 14496-2	ITU-T H.264, ISO/IEC 14496-10
Interlace coding?	N	N	Y	Y	Y	Y
Progressive coding?	Y	Y	Y	Y	Y	Y
Optimized Bitrate Range (bits/sec)	64k+	1M-2M	4M-20M	>= 10k	>= 10k	>= 10k
Primary Applications	VC	VCD	TV, DVD	M	S, W, M	M
Min. Picutre Size (width, height)	172x144	16x16	16x16	16x16	16x16	16x16
Max. Picture Size (width, height)	352x288	4kx4k	64kx64k	2048x1152	64kx64k	4096x2048
Variable aspect ratio?	N	Y	Y	Y	Y	Y
Variable frame rate?	Y	N	difficult	Y	Y	Y
4:2:0 Chrominance?	Y	Y	Y	Y	Y	Y
4:2:2 Chrominance?	N	N	Y	N	Y	N
4:4:4 Chrominance?	N	N	N	N	Y	N
Motion Comp.	Y	Y	Y	Y	Y	Y
Transform coding	Y	Y	Y	Y	Y	Y
Bitrates supported (kbits/s)	any	any	any	any	any	any
PR Status	1	1	2	1	2	1, 2

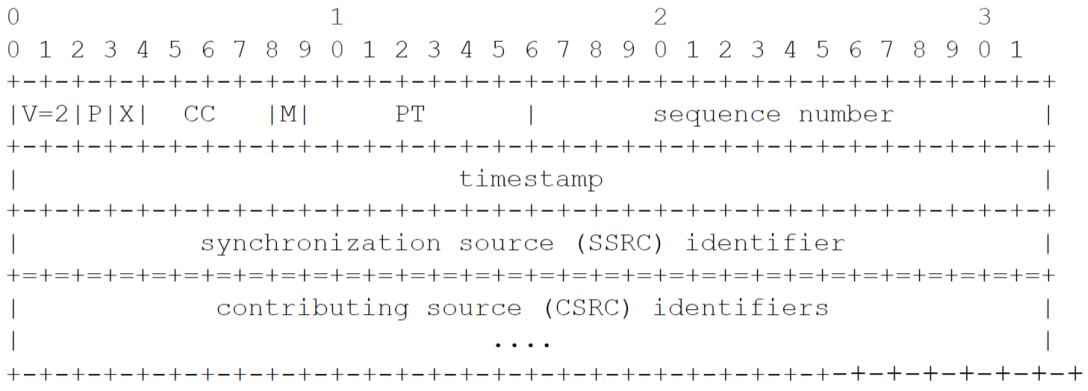
Primary Applications:

DVD: DVD-video, **M**: Mobile, **RN**:Radio news, **S**:Streaming, **SVD**:Simultaneous voice/data, **T**:Telephony, **TC**:Teleconferencing, **TV**:Television, **V**:Voice on IP, **VC**:Video conferencing, **VT**:Video telephony, **W**:Wireless LAN

SIP call flow example

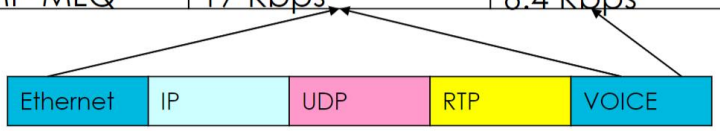


RTP packet structure

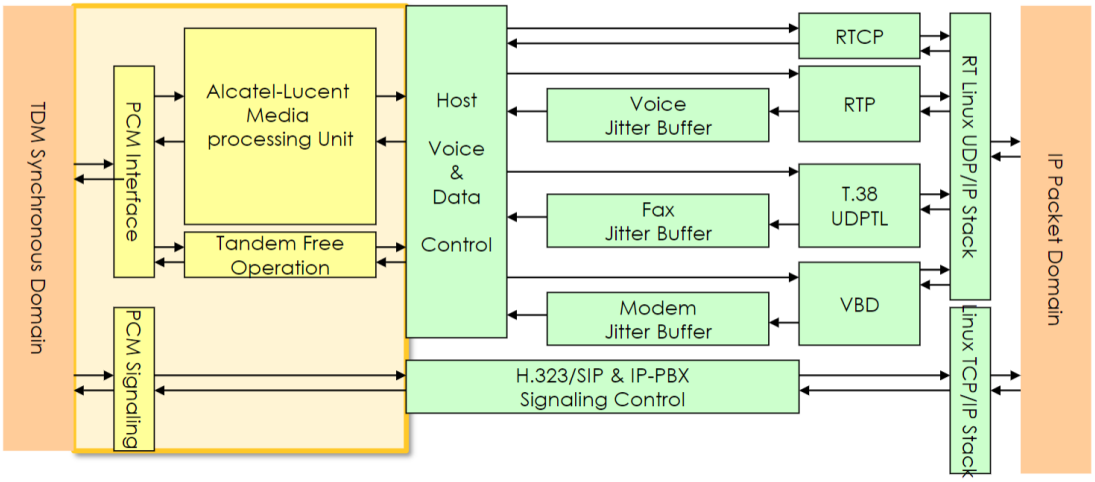


VoIP Bandwidth Usage over Ethernet

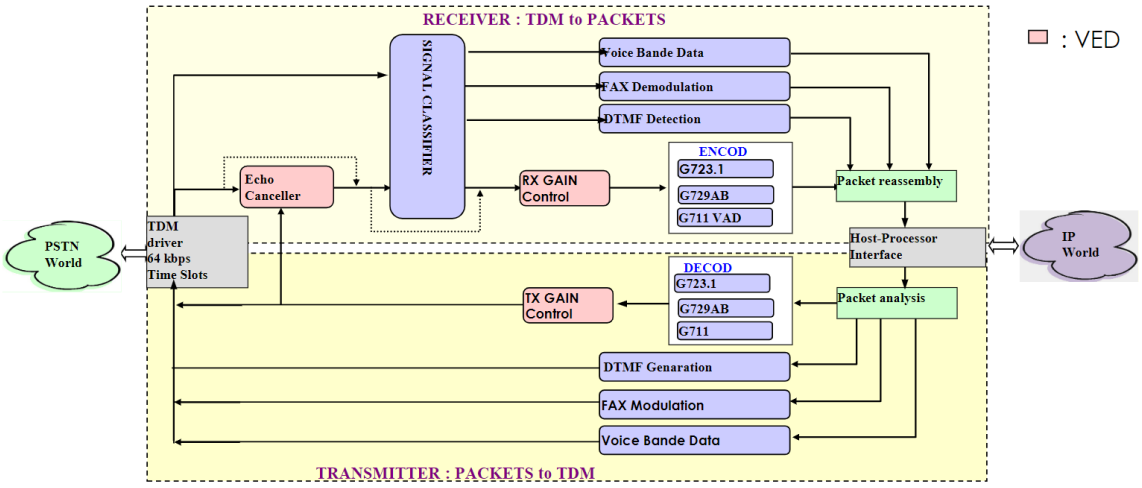
codec	IP network	Voice bandwidth
G.711 PCM	80 Kbps	64 Kbps
G.729AB CS-ACELP	24 Kbps	8 Kbps
G.723.1 ACELP	16 Kbps	5.3 Kbps
G.723.1 MP-MLQ	17 Kbps	6.4 Kbps



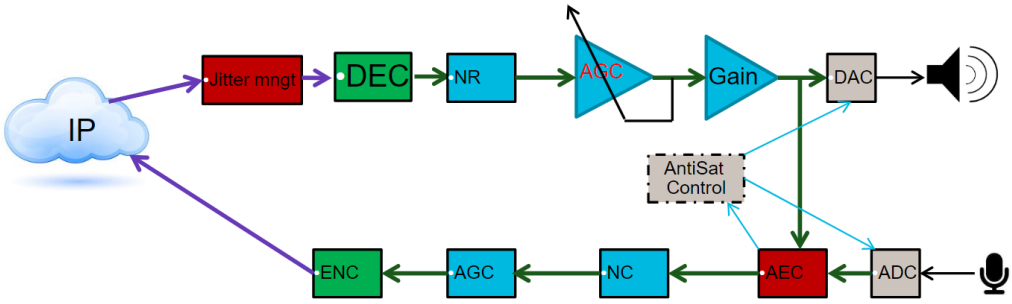
OmniPCX VoIP Gateway Architecture



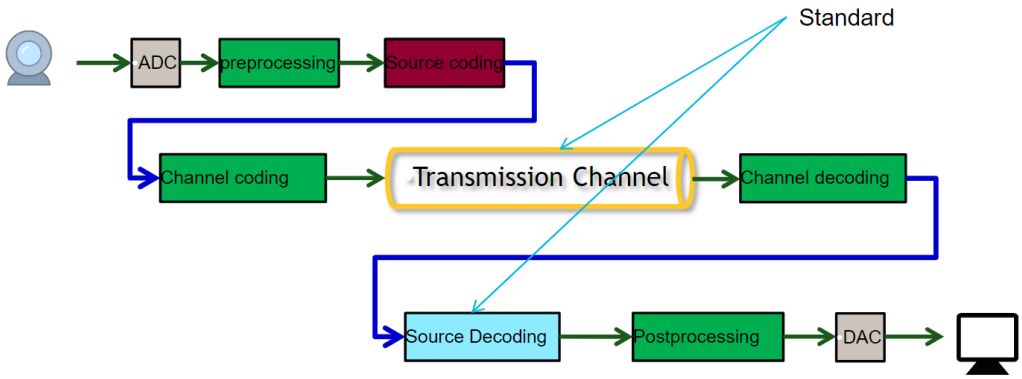
Alcatel-Lucent Vocoder System Architecture



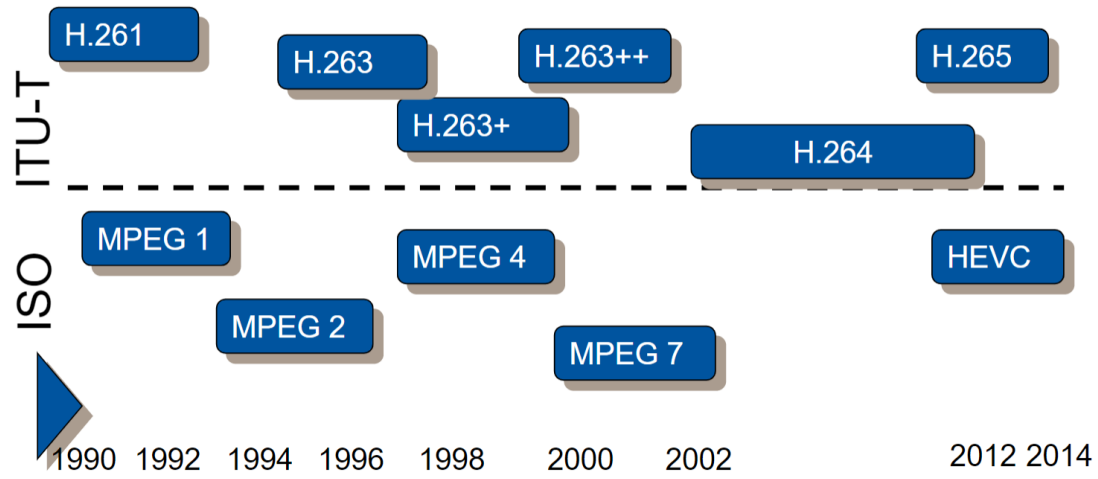
Exemple of Media Architecture of SWPhone



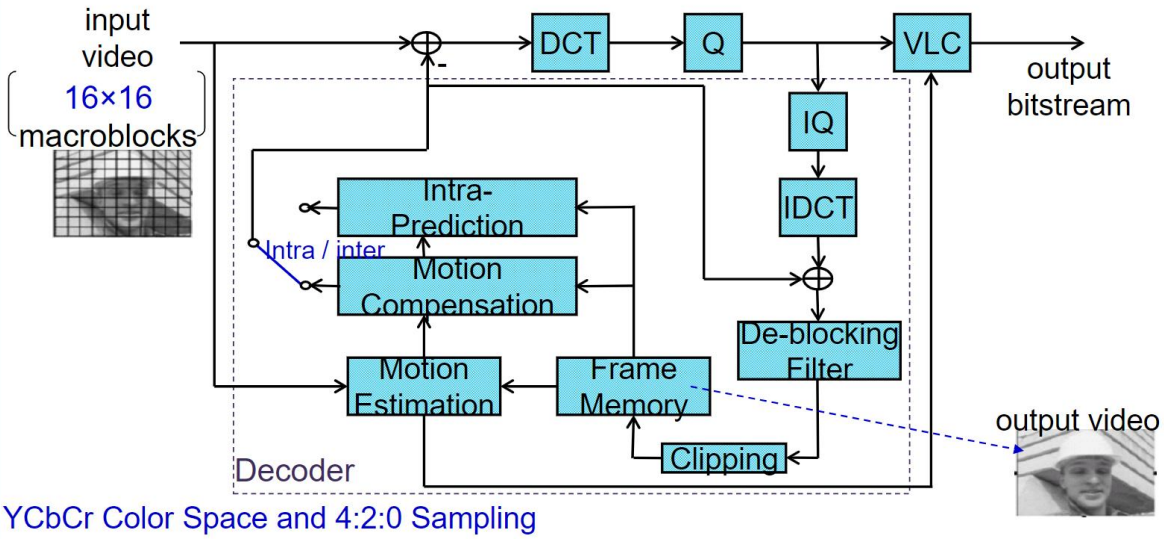
End To End Video transmission overview



Chronology of Video Standards

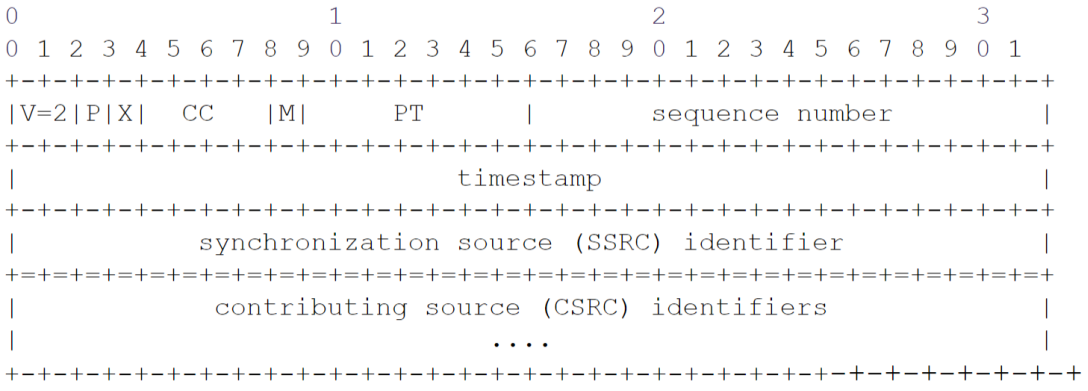


VCL (Video Coding Layer)

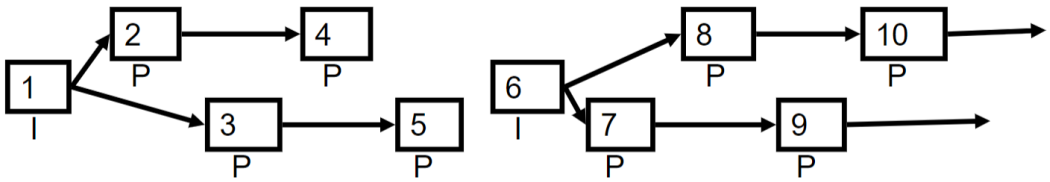


YCbCr Color Space and 4:2:0 Sampling

RTP packet structure



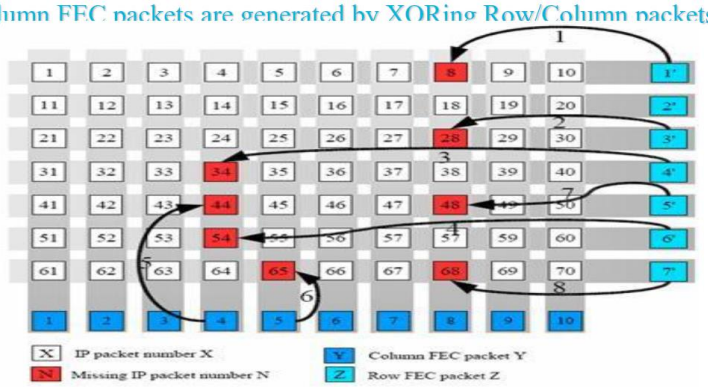
Multi-threaded Video



- Reference pictures are interleaved to create two or more independently decodable threads.
- If a frame is lost, the frame rate drops to 1/2 rate until a sync frame is reached.
- Same syntax as Reference Picture Selection, but without ACK/NACK.
- Adds some overhead since prediction is not based on most recent frame.

FEC Illustration: simple Matrix based FEC

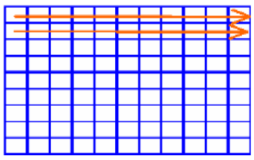
- Packet ranged in matrix in sender side
- Row/Column FEC packets are generated by XORing Row/Column packets



- One missed packet per Row or Column can reconstructed by XORing other packets and FEC packet
- By iterative it's possible to reconstruct more than one missing packet by Row and Column

H264 Slice Groups & FMO Technique

- A picture is divided into a number of mackroblocks:
- Slices are a sequence of macroblocks that are scaned in order when not using FMO
- Slices are self-contained



Example of block mapping without FMO

- FMO modifies the partition of mackrobloks into slices using slice groups in the encoding side
- Many mapping can be used
- If a slice is lost, the decoded picture will have holes in the positions of missed mackroblocks
- A concealment algorithm will use correctly decoded data (spacially very close) to reconstruct missed block

block mapping using FMO and 2 silces group