University of Glasgow | School of Computing Science

RTCP Feedback for Congestion Control in Interactive Multimedia Conferences

draft-ietf-rmcat-rtp-cc-feedback-03

Colin Perkins

RTCP feedback timing – can we report often enough?

- Congestion control requires reasonably rapid feedback
  - On packet loss or ECN-CE marking
  - On packet timing, for delay-based algorithms
- Using the feedback packets in draft-dt-rmcat-feedback-message-01 can RTCP send timely feedback with acceptable overhead?

- Assume a modern RTCP implementation:
  - RTP/AVPF or RTP/SAVPF profile
  - Non-compound RTCP packets
  - RTCP XR
  - RFC 7022 format for SDES CNAME items
  - Report aggregation fixes from draft-ietf-avtcore-rtp-multi-stream
  - Reporting groups (draft-ietf-avtcore-rtp-multi-stream-optimisation)

Scenario 1: VoIP

- Two-party point-to-point VoIP call

- Speech frames sent every T_f seconds; both participants sending
- Want to send congestion feedback every N_r frames

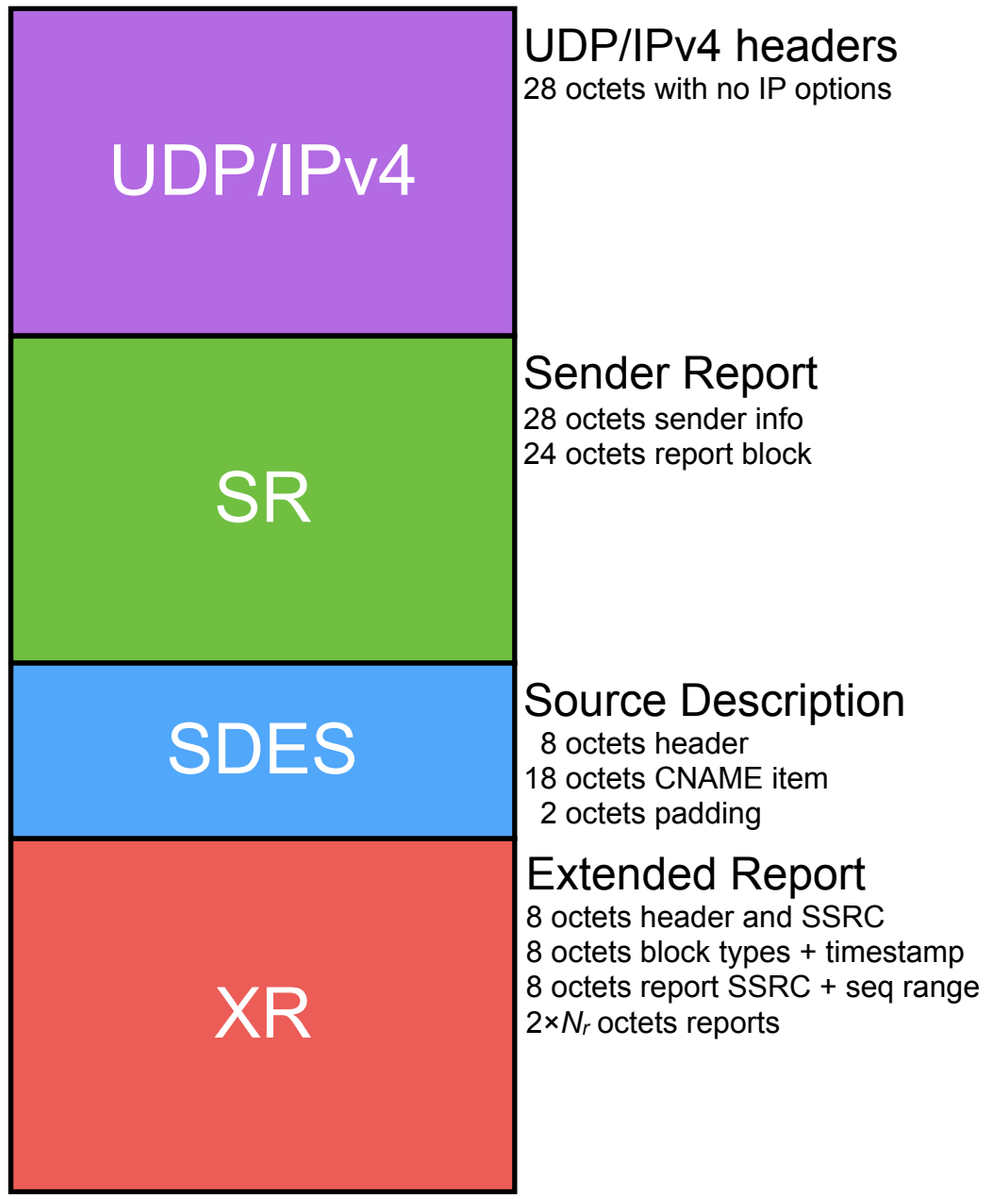
- Desire RTCP reporting interval = $T_f \times N_r$ seconds

- RTCP packets can be regular compound packets or non-compound packets sent using RTP/AVPF early feedback
  - Send N_{nc} non-compound packets between every compound packet

Scenario 1: VoIP – compound RTCP packets

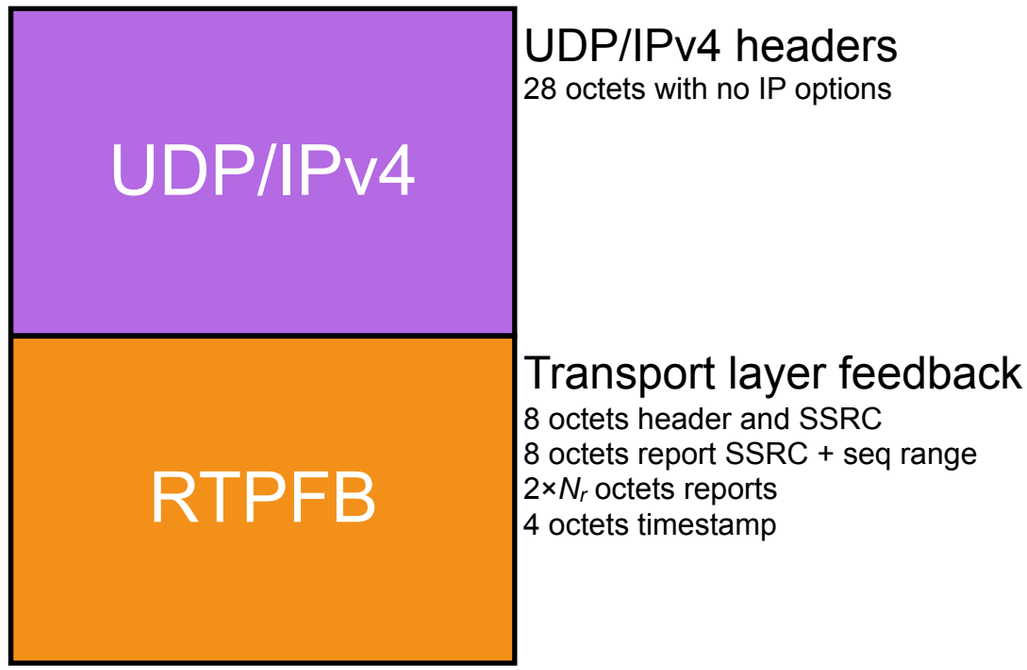
- Compound RTCP packets contain:
  - Sender Report (SR)
  - Source Description (SDES) with CNAME item
  - Extended Report (XR) with congestion control feedback (draft-dt-rmcat-feedback-message-01)

- Packet size, $S_c = 132 + 2 \times N_r$ octets

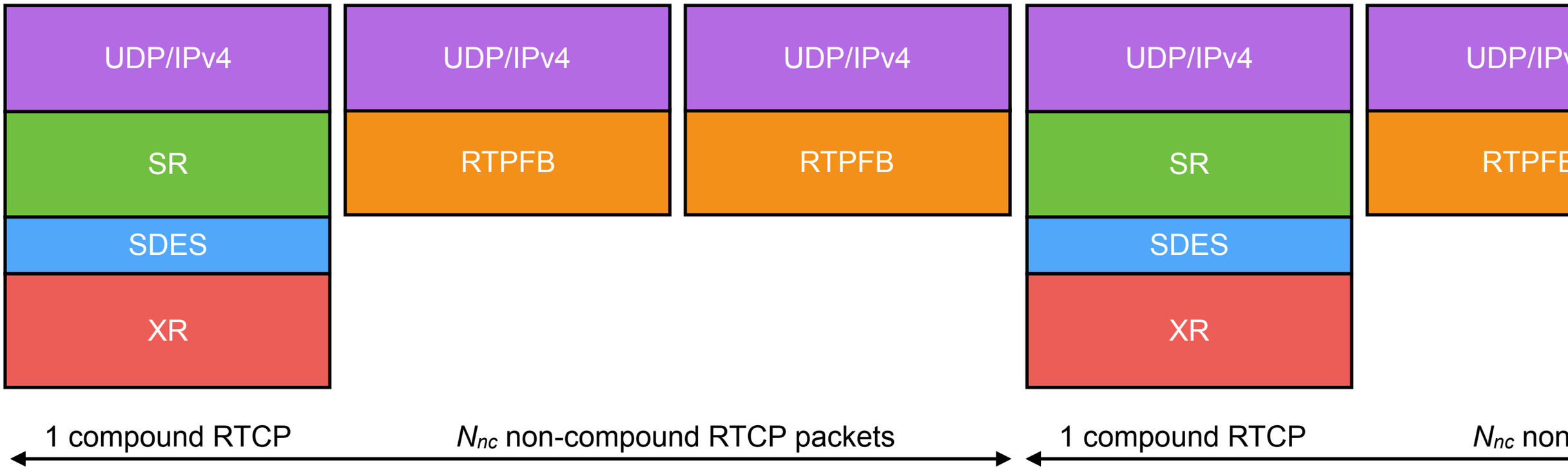


Scenario 1: VoIP – non-compound RTCP packets

- Non-compound RTCP packets contain:
  - RTP/AVPF transport layer feedback packet (draft-dt-rmcat-feedback-message-01)
- Packet size, $S_{nc} = 48 + 2 \times N_r$ octets



Scenario 1: VoIP – average RTCP size



- Average RTCP packet size, $S_{rtcp} = (S_c + N_{nc} \times S_{nc}) / (1 + N_{nc})$
  where $N_{nc} = 0$ if non-compound packets are not sent

Scenario 1: VoIP – RTCP bandwidth

- From RFC 3550: RTCP reporting interval, $T_{rtcp} = n \times S_{rtcp}/B_{rtcp}$ where:
  - n is the number of participants (n = 2 in this scenario)
  - $S_{rtcp} = (S_c + N_{nc} \times S_{nc}) / (1 + N_{nc})$ is the average RTCP packet size in octets
  - B_{rtcp} is the bandwidth allocated to RTCP in octets per second

- To report every N_r frames, we want $T_{rtcp} = N_r \times T_f$

  $\Rightarrow N_r \times T_f = n \times S_{rtcp}/B_{rtcp}$

  $\Rightarrow B_{rtcp} = (n \times (S_c + N_{nc} \times S_{nc})) / (N_r \times T_f \times (1 + N_{nc}))$

Scenario 1: VoIP – RTCP bandwidth requirements (1)

T_f (seconds)	N_r (frames)	B_{rtcp} (kbps)
20ms	2	53.1
20ms	4	27.3
20ms	8	14.5
20ms	16	8.0
60ms	2	17.7
60ms	4	9.1
60ms	8	4.8
60ms	16	2.7

Sending only compound RTCP packets

- Chart gives the required RTCP bandwidth, B_{rtcp}, to send a report after every N_r frames with frames being sent every T_f seconds
  - Total RTCP bandwidth for the session: each participant gets half of this
  - Compound packets only: $N_{nc} = 0$
- Sending an RTCP report every 2nd frame with 20ms frames → 53kbps RTCP bandwidth
- Sending an RTCP report every 16th frame with 60ms frames → 2.7kbps RTCP bandwidth
  - This is 1 RTCP packet per second from each SSRC in the VoIP call

Scenario 1: VoIP – RTCP bandwidth requirements (2)

T_f (seconds)	N_r (frames)	B_{rtcp} (kbps)
20ms	2	36.7
20ms	4	19.1
20ms	8	10.4
20ms	16	6.0
60ms	2	12.2
60ms	4	6.4
60ms	8	3.5
60ms	16	2.0

Alternating compound and non-compound RTCP

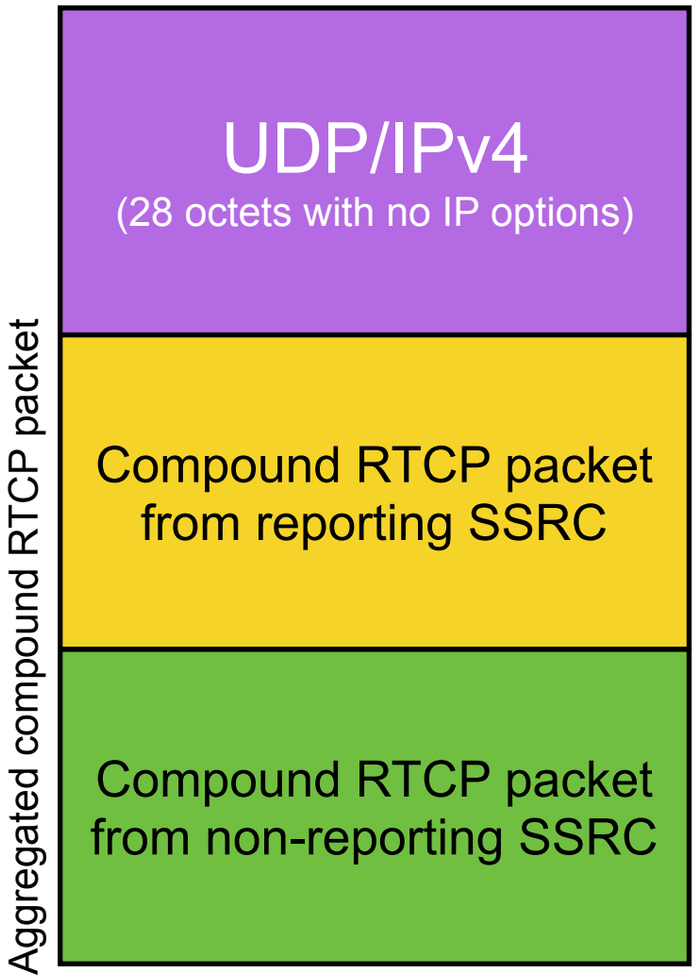
- Required RTCP bandwidth is reduced if a non-compound packet is sent between compound packets
- Reduced header overheads – due to not sending SR/RR and SDES packets in some reports

Scenario 2: Video conference

- Point-to-point video conference
- Two parties, each sending audio and video
- Media bundled onto single 5-tuple → 4 SSRCs
- 1 audio SSRC, 1 video SSRC, for each party

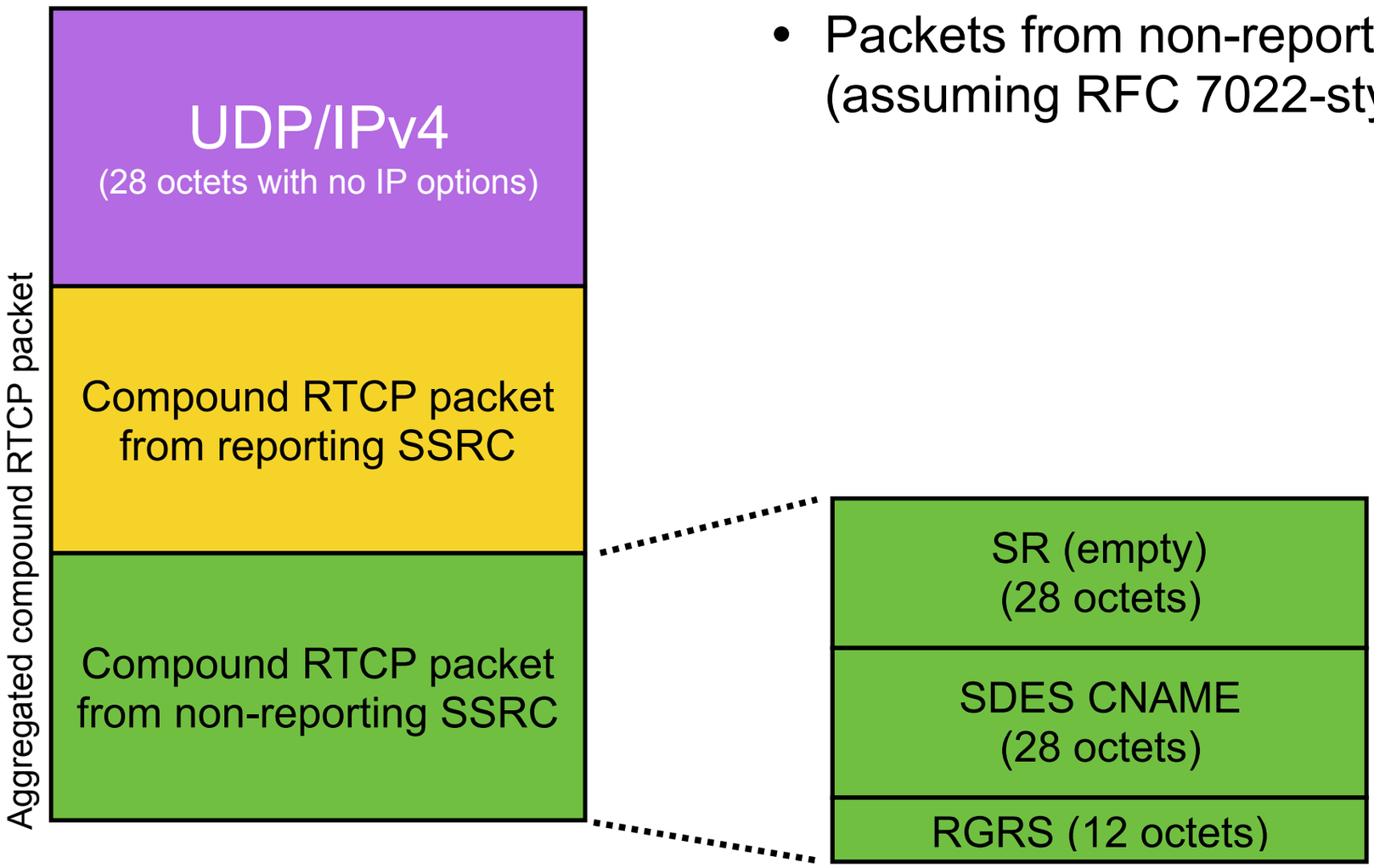
- Video frame interval = T_f (i.e., frame rate = $1/T_f$ frames per second)
- Desire RTCP reporting interval = $N_r \times T_f$
  - If $N_r = 1$, report every frame
  - If $N_r = 2$, report every other frame
  - …
- Packets can be sent as compound or reduced size (non-compound) RTCP packets

Scenario 2: Video conference – compound packets



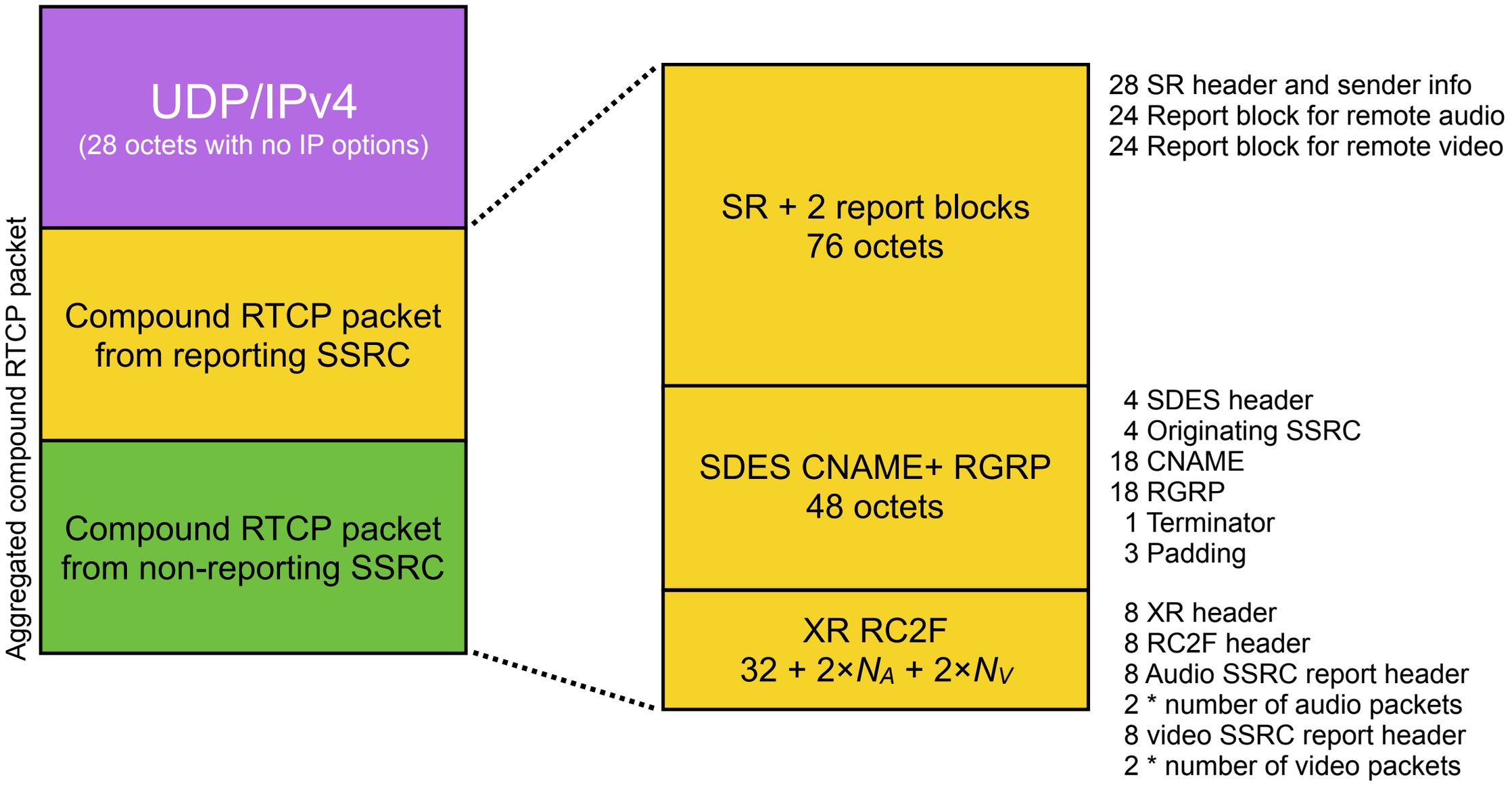
- Two SSRC → need to aggregate feedback into a single RTCP packet
  - Each packet is an aggregation of a compound RTCP packet from the audio SSRC and a compound RTCP packet from the video SSRC
- RTCP reporting groups are used:
  - One SSRC is designated as the reporting SSRC
  - The other SSRC delegates its reports to that SSRC
  - The reports are aggregated, so it doesn't matter which is chosen as reporting SSRC

Scenario 2: Video conference – compound packets



- Packets from non-reporting SSRC are 68 octets (assuming RFC 7022-style CNAME)

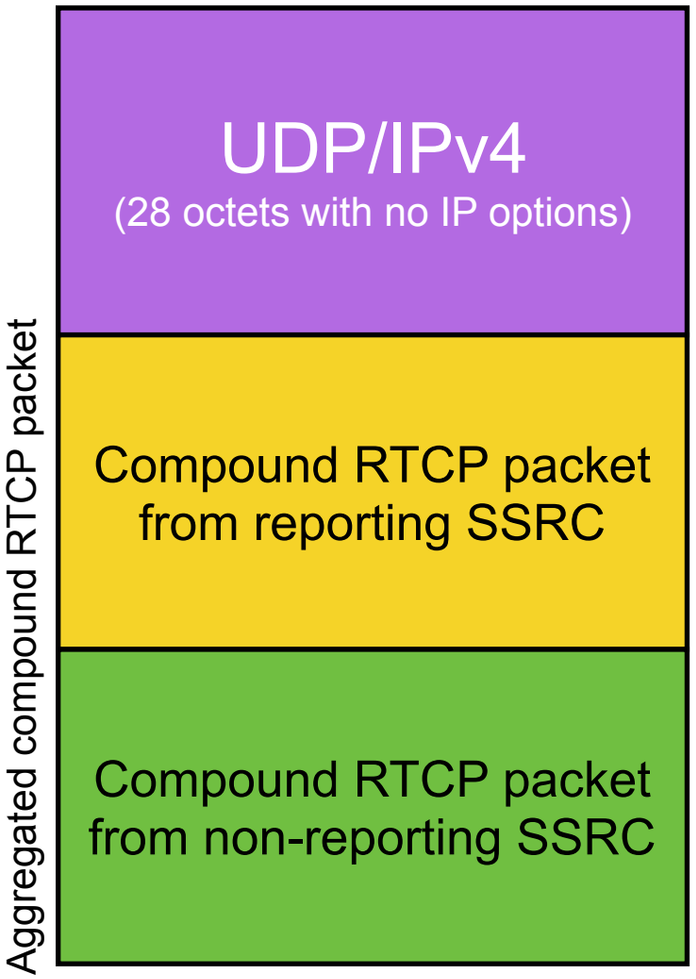
Scenario 2: Video conference – compound packets



- Packets from reporting SSRC are 156 + 2×N_A + 2×N_V octets

13

Scenario 2: Video conference – compound packets



- 28
- $156 + 2\times N_A + 2\times N_V$ octets
- 68 octets
- Total = $252 + 2\times N_A + 2\times N_V$ octets
- Since this reports on two SSRCs, it is halved before use: $S_c = (252 + 2\times N_A + 2\times N_V)/2$

Scenario 2: Video conference – B_{rtcp} calculation

- Assume:
  - Constant rate media
  - Video frames equal size
  - Audio at 50 packets per second (20ms frames)
  - MTU around 1500 octets
- RTCP bandwidth calculation as for scenario 1:

$$B_{rtcp} = (n \times (S_c + N_{nc} \times S_{nc})) / (N_r \times T_f \times (1 + N_{nc}))$$

with

$S_c = (252 + 2{\times}N_A + 2{\times}N_V)/2$

$N_{nc} = 0$

T_f based on chosen video frame rate

$N_r = 1$ (report on every frame)

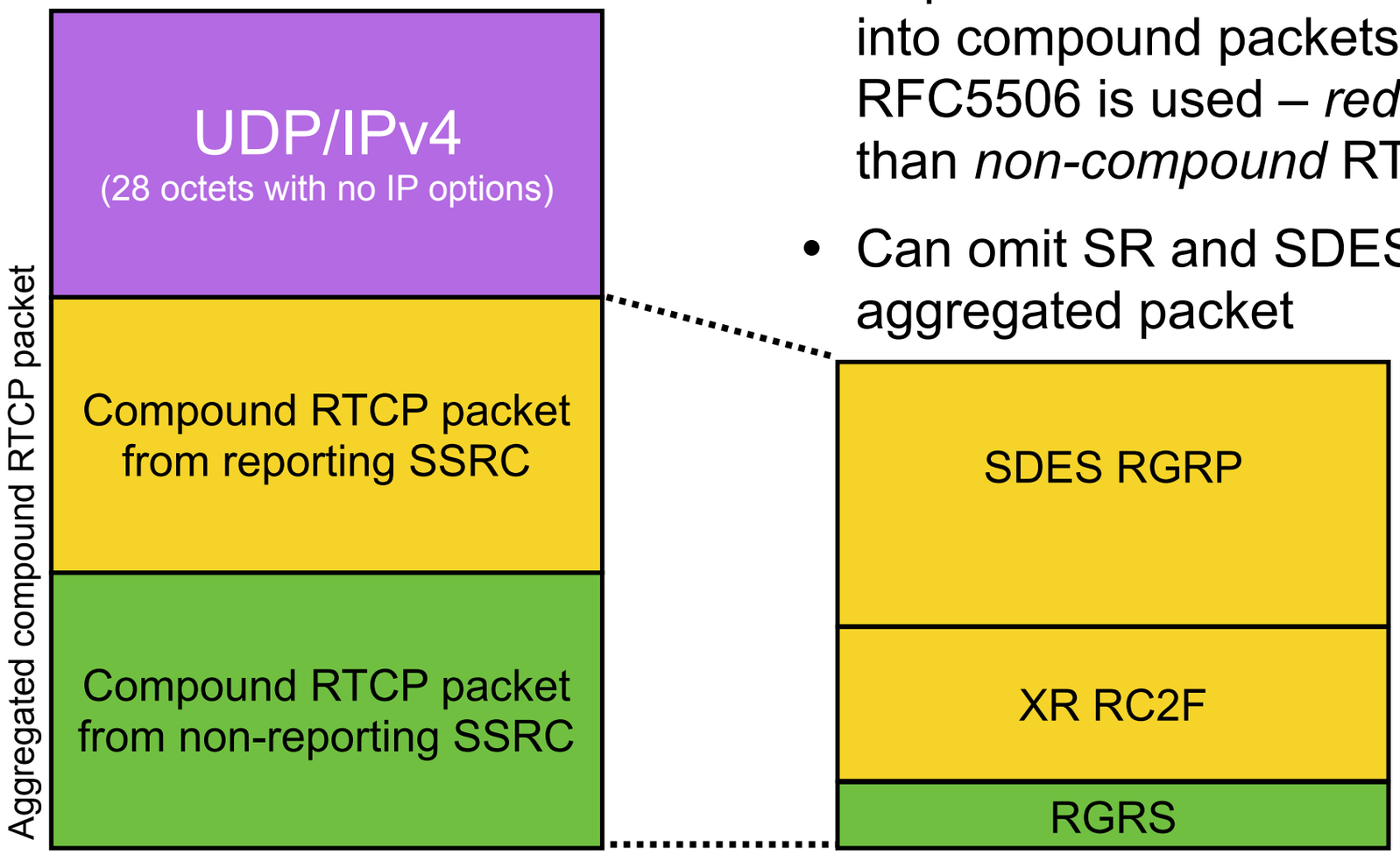
Scenario 2: Video conference – required RTCP bandwidth

Media Rate (kbps)	Video Frame Rate ($1/T_f$)	Video packets per report: N_v	Audio packets per report: N_a	Required RTCP bandwidth, B_{rtcp} in kbps (and as % of media rate)
100	8	1	6	33.3 (33%)
200	16	1	3	65.0 (33%)
350	30	1	2	120.1 (35%)
700	30	2	2	121.9 (17%)
700	60	1	1	240.0 (34%)
1024	30	3	2	122.8 (12%)
1400	60	2	1	241.8 (17%)
2048	30	6	2	125.6 (6%)
2048	60	3	1	243.8 (12%)
4096	30	12	2	131.3 (3%)
4096	60	6	1	249.4 (6%)

Sending only compound RTCP packets

B_{rtcp} scales linearly with N_r (i.e., reporting every 2nd frame halves the required RTCP bandwidth)

Scenario 2: Video conference – reduced size packets



- Reports from two SSRCs are being aggregated into compound packets irrespective of whether RFC5506 is used – *reduced size* RTCP rather than *non-compound* RTCP
- Can omit SR and SDES CNAME from the aggregated packet
- Gives $S_{nc} = (96 + 2{\times}N_v + 2{\times}N_a)/2$
- Repeat calculation with $N_{nc} = 1$ indicating that we alternate regular and reduced size RTCP

Scenario 2: Video conference – required RTCP bandwidth

Media Rate (kbps)	Video Frame Rate ($1/T_f$)	Video packets per report: N_v	Audio packets per report: N_a	Required RTCP bandwidth, B_{rtcp} in kbps (and as % of media rate)
100	8	1	6	23.5 (23%)
200	16	1	3	45.5 (23%)
350	30	1	2	84.4 (24%)
700	30	2	2	85.3 (12%)
700	60	1	1	166.9 (24%)
1024	30	3	2	86.2 (8%)
1400	60	2	1	168.8 (12%)
2048	30	6	2	89.1 (4%)
2048	60	3	1	170.6 (8%)
4096	30	12	2	94.7 (2%)
4096	60	6	1	176.3 (4%)

Alternating regular and reduced-size RTCP packets

B_{rtcp} scales linearly with N_r (i.e., reporting every 2nd frame halves the required RTCP bandwidth)

Conclusions

- RTCP can be used for congestion control feedback with reasonable overhead, provided:
  - Care is taken with session configuration
  - Feedback rates scale with media rates – low rate sessions may need to report on a smaller fraction of media frames

- Questions:
  - Can congestion control candidates operate with the amount of feedback available with reasonable overheads?
  - Are the overheads/configurations acceptable?
  - What guidance do we need to provide to implementers?