**Source: Huawei (Rapporteur)**

**Title: KI#4 and KI#5, Summary of company views**

This document is to summary company views on key questions of KI#4 and #5 to facilitate the following conclusion discussion.

### Q1: How does UPF identify DL PDU Set info?

* Option 1: use existing IETF RTP/SRTP RFC and draft
* Option 2: Define/extend N6 protocols to carry related info
  + Option 2.1: extend GTP-U protocol
  + Option 2.2: extend HTTP header (S2-2205830)
  + Option 2.3: extend RTP header
* Option 3: UPF implementation based on e.g. traffic characteristics.
* Option 4: UPF interacts with NWDAF(S2-2205838)

**[Company inputs]**

|  |  |  |
| --- | --- | --- |
| Company | Position | Justifications |
| Apple | Support Option 1. Support Option 2.2/2.3 for an enhanced PDU Set identification | Option 1 would be the least required for PDU Set identification. However there may be limitations to the extent of information that could be derived from interpreting the headers. Option 2.2 using MASQUE for identification of DL PDU Set on N6 or Option 2.3 by extending RTP header would offer more flexibility for deriving PDU Set information at the UPF for wider range of application communication. We also support these options as a longer term solution if achieving this in Rel-18 is difficult. We don’t support Option 3 and Option 4. |
| CATT | Support Option 1 and option 3 in this TR. | Option 1 is well defined by the IETF. And Option 3 can let UPF to implement the DPI for the XRM services.  Option 2.2 and 2.3 will need wait for IETF to define new IETF. And HTTP is not used for the real-time interactive XRM services. Option 2.3 can be considered in the further release. It is still unclear what general information to be defined for the GTP-U.  Option 4 cannot work at all and cannot provide RT information. |
| China Mobile | prefer option1, option 2.1, option2.3, option3 | For option1, it is better to reuse the existing RTP protocol.  For option 2.1, the close cooperation between application and operator’s edge platform is expected. So option 2.1 is also preferred.  For option2.3, the RTP extension can bee accepted.  For option3, this is also an acceptable way to let the UPF identify the PDU set. |
| China Telecom | support option 1, with option 3 as a supplement. | Option 1: The existing RTP and SRTP header/header extensions can be effective in identifying DL PDU Set.  Option 2.1: Current GTP-U protocol is used between UPF and gNB, while the identification of PDU Set is at the UPF. The extension of GTP-U protocol may introduce too much impact on current 5GS.  Option 2.2: Since HTTP protocol is mainly used for encrypted data transmission, it is still not mature enough to adopt extended HTTP header in this Release.  Option 2.3: Extended RTP header is also not mature enough to be realized in this Release.  Option 3: AF knows the traffic characteristics and can send this information to assist the identification of PDU Set. This can be used as a supplement for option 1.  Option 4: Since this option is intended for encrypted XR traffic, its feasibility needs further discussion. We are neutral. |
| Ericsson | Support options 2.2 (MASQUE) and 2.3 (RTP ext header) | Option 1:  There are multiple alternatives to identify the PDU Sets and that the parameters need to be adapted depending on the specific media payload to be detected (e.g. video frame vs. H.264, H.265 and H.266 video slices, see solution 12). Since there are many variants and combinations and the assumptions at UPF and AS may be different, the UPF can make wrong interpretations. This mechanism does not work for media transferred on SRTP.  Currently packet filters in dynamic PCC rules are limited to IP and Ethernet header parameters. Filters on a predefined PCC rule may use parameters beyond IP+port and Ethernet headers, using DPI and other advanced techniques. However, the definition of such DPI filters is not standardized by 3GPP.  If the rules are to be provided by the AF, they need to be encoded so that they are properly interpreted by the UPF and there are impacts on the AF, PCF, SMF and UPF.  Supporting new or evolved media protocols and NALUs may require new development in the UPF and AF and may also impact the PCF and the SMF. This can slow down innovation, especially if there are new matching protocol and payload parameters and combining expressions to be standardized.  Detection of PDU Sets requires continuous inspection of packets for the associated media. The more complex and varied the matching rules, the more CPU and memory requirements are placed on the UPF.  Option 2.1  The GTP-U tunnel based solution is based on a pre-defined configuration of GTP-U tunnels towards AS for the different applications (between UPF and AS), since otherwise the tunnels should be managed by the CP, which entails further impacts on the 5GS. However, it does not cover the assignment of tunnel endpoint identifiers needed to handle the traffic for different UEs and requires additional OPEX for the operator to add, delete and maintain GTP-U tunnels that are manually configured.  Option 2.2  All the knowledge on RTP, RTP extension and media NALUs for the different payload types used by a specific application is known by the AS, so it can provide the assistance information as metadata on the media packets.  Option 2 approach is more future proof than option 1, since any new payload types can be mapped by the application with media knowledge into PDU set identification and boundaries. This may also foster and speed up innovation, in addition to minimizing impacts on 5GS.  Also, CPU requirements on UPF are kept to a minimum, since all metadata can be read from a single place: the extension header, without any complicated parsing conditions.  This approach is more deterministic than the other ones, since the PDU Set identification, size and boundaries are directly signaled by the AS. XRM applications are willing to provide the necessary information, and this solution is more straightforward than providing matching rules that need to be validated through interoperability tests. Application enhancements and activation and deactivation of features come with no or less impacts in the interworking.  The MASQUE implementation has the following additional advantages:  •It does not require CP involvement for tunnel management  •It works regardless of the protocol used for real-time media (i.e. can be different from RTP) and the level of encryption  •It provides integrity protection and encryption of both the payload and the metadata  •It provides the means to insert metadata once per PDU Set, by means of registering the metadata for a MASQUE context and referencing the context for subsequent PDUs in the set, therefore optimizing signaling  •Only one tunnel needs to be established between the UPF and the Application Server  Option 2.3  The RTP extension option has the advantages detailed above for option 2 plus the fact that it is more straightforward for current applications based on RTP or SRTP.  Option 3  Providing assistance information in the form of traffic characteristics, such as burst arrival time, periodicity/separation, last PDU packet size, etc. has applicability limitations, meaning that the needed information may not be inferred for every XRM payload.  It may require frequent updates as applications evolve, be more difficult to develop and maintain than solutions based on parameters provided by collaborative XRM applications and may require ad-hoc detection per media type.  Option 4  Similar comments as for option 3 |
| Futurewei | Support option 1.  Methods in option 2 need to be considered only beyond Rel18 as HTTP/3 matures. (For completeness, option 2.4 (extend UDP options, QUIC media header) should also be evaluated.)  Do not support options 3, 4. | Option 1 with RTP and SRTP where header/header extensions are visible are well defined and should be specified in the normative phase.  Option 2 is for use cases with encrypted media payloads/ no visible meta-data. Since XRM focuses on low latency, our view is that option 2 applies in practice to HTTP/3 (QUIC transport). TCP based transports for WebRTC, DASH suffer from HOL blocking and do not offer unreliable delivery.  QUIC media header - MoQ (https://datatracker.ietf.org/group/moq/about/) has just started standardization work in IETF. While this method will likely not be ready for Release 18, other methods will compete with a general solution in MoQ. The best approach for now would be to engage IETF, provide input such that 3GPP requirements and concerns are considered during the MoQ standardization.  If MoQ cannot meet XRM needs, 3GPP specific options such as MASQUE (Option 2.2), GTP-U (Option 2.1), or UDP options (Option 2.4) should be evaluated. A short evaluation of each below:  -MASQUE will require 3GPP specific adaptation by the application domain (AF – 5GC), incurs per packet overhead of L7 encapsulation and TLS encryption overhead and negotiation overhead per application flow.  In addition, L7 load balancer/reverse proxy at application domain would need to encode MASQUE meta-data for media streams on path from AS to UE (both meta-data and reverse proxy behavior need new stds)  -GTP-U requires CP support for tunnel termination, and AS to support GTP-U.  -UDP options while standardized in IETF would require new standards for XRM meta-data.  (Figures below show comparison of the different options)  Option 3 (traffic characteristics) is not practical to implement at UPF as transport segments between AS and UPF introduce distortions to the expected traffic patterns due to variability in forwarding and out-of-order packet arrivals. In addition, the AS may adjust sending rates (codec rates) dynamically to respond to network congestion.  Option 4 (S2-2205838 has been noted), but similar comments as for option 3. |
| Google | Option 1 as the baseline for the unencrypted media header of RTP/SRTP based protocol for normative in R18. Security aspect needs further investigation at SA3.  Option 2.2/2.3 can be considered for the encrypted media header of HTTP/RTP based protocol using TLS/DTLS/QUIC. The Coordination and collaboration with IETF/SA4 are needed. | Option 2.1: the impacts on the application server cannot be ignored for broader deployment and supports. The tunnel-based solutions generally have scalability issues, e.g. GTP-U over N6 based solution requires CP plane (re)configuration if tunnel setting is not static.  Option 3/Option 4 would need assistant information of accurate/coarse traffic characteristic from application server, which seems not practical for variants of XRM services. |
| Huawei | Support Option 1 and 3.  Support Option 2.3 if SA4 agrees to define necessary extension within Rel-18 timeline. Support for this option might be done in WI phase or next release if SA4 progress is not enough when SA2 close study phase in Nov. | Option 1 reuses existing IETF specs, which means wider support from application aspect.  Option 2 defines 3GPP specific protocol headers, and relies on support of application on such protocol headers:   * Option 2.1 request application server to support GTP-U protocol, which is questionable considering GTP-U was mainly defined and used in 3GPP domain. * Option 2.2 requests UPF to support MASQUE proxy, which we have concerns on the performance downgrade caused by the HTTP proxy in UPF, and also the maturity of MASQUE which is still under development in IETF. It’s also not clear which SDO should define extension to HTTP to transfer the necessary PDU Set information between AS and UPF. * Option 2.3 is fine if SA4 agrees to define the corresponding extensions. However, option 2.3 relies on SA4 progress hence might be only done during WI phase or next release depends on SA4 progress.   Option 3 should be supported to allow vendor-specific UPF implementation to identify necessary information based on standardized or property protocols used by the application.  Option 4 may cause some performance issue if a separated NWDAF is used. If UPF/NWDAF must be collocated, it should be considered as an UPF with big data analysis capability, as a subset of option 3. |
| Intel | As a minimum, Option 2.1 to be progressed to normative work. Other options are FFS. | Option 2.1 has minimum impact on the 5G system as the burden of PDU Set identification is delegated to the Application Server.  In contrast, Option 1 requires “PDU Set identification rules”, which implies PDR extensions to include packet fields beyond the traditional IP 5-tuple which are specific to the upper layer protocol (RTP, SRTP, NAL). Option 3 does not require any specification. Option 4 is unclear. |
| InterDigital | Option 2.2 | We prefer option 2 over option 1 because it is more extensible / future proof. However, options 1 and 2 can be considered complimentary. Of the three options under option 2, option 2.2. is the most flexible.  We understand Option 3 to mean that PDRs in the UPF are configured to point to application detection rules and the application detection rules would not be standardized. We do not think that much can be achieved with this approach. It cannot be guaranteed that packets will be received in order and behaviour would not be consistent because the application detection rules would not be standardized.  Our concerns with Option 4 are similar to our concerns with option 3. Additionally, the added delay due to the interaction with the AnLF is concerning. |
| KDDI | Support Option 1 and Option 3. | For Option 1, UPF can make the best use of RTP/SRTP header /NAL unit header information as defined in IETF specs. Taking the SA4 LS as an example, SA4 has analyzed to map RFC7798 (H.265) and RFC6184 (H.264) to PDU Set concept. The LS also shows the operation for H.265/H.264 Fragmentation Unit Packets as a potential scenario (The first packet of the PDU Set is lost, all other packets of the fragmentation units are useless, whereas of the last packet is lost, the decoder can use all packets except the last one. Obviously, in this example scenario, the first packet is “more important” than the last one). This operation is also applicable to the RTP header identification as defined in draft-ietf-avtext-framemarking.  Option 2.2 and Option 2.3 may be useful in encrypted traffic. However, these options depend on IETF works and need to be considered to future release.  If the RTP/SRTP header is encrypted, Option 3 may be applicable.  Option 4 is similar to Option 3 except for NWDAF analytics. The NWDAF can be useful for no-real time analytics. So Option 4 is not suitable for PDU Set identification. |
| Lenovo | Option1 and Option 3 are preferred at this stage. | Option 1 utilizing existing IETF RFC and draft, which has only 3GPP impact. Option 2 is either out of 3GPP scope or interworking with IETF is necessary, which is not feasible at this stage. Option 3 depends on UPF implementation based on traffic characteristics, which is also feasible. Option 4 is not feasible which does not provide real-time information. |
| MediaTek | Option 1 along with Option 2.1 | Option 1: The IETF has issued a last call for draft-ietf-avtext-framemarking version 13 for it to become a recognised RFC. The framework described in the draft is codec agnostic.  Option 2.1: The parameters contained in the draft framework (of option 1) can be included in the extended GTP-U header.  Option 2.2: HTTP is a protocol to fetch or push (HTTP3) resources through web links provided in a manifest file and is not suited to XR.  Option 2.3: A single extension per RTP packet is allowed according to RFC 5285 without needing IETF approval. |
| Meta | Support Option 1, Option 2.2, and Option 2.3. | Option 1 should be supported when related media headers are not encrypted.  Option 2.2 should be supported when media headers are encrypted; otherwise, 3GPP is not able to deal with encrypted transport protocol e.g, with QUIC.  Option 2.3 should be supported for extendibility and efficiency when PDU set attributes identification are used.  Option 3 is purely based on vendor’s implementation; hence, outside 3GPP scope. Implementation (not standard based) will introduce a wide variety of requirements/supports in the ecosystem. This is not preferred.  Option 4: is not preferred as the outcome is highly dependent on the AI/ML engine. A more precise solution is preferred (option 1 or 2). |
| Nokia | Support Option 1 and Option 3. Option 2 should be pursued in the longer term | Option 1 supports existing mechanisms for encoding PDU Set information on N6, and hence is likely to have the widest applicability and the greatest ease of adoption for applications. Option 1 has several sub-options, any of which may be supported according to implementation:   * Option 1.1: RTP is unsecured, frame marking extension header is not in use * Option 1.2: RTP is unsecured, frame marking extension header is in use * Option 1.3: RTP is secured (SRTP), frame marking extension header is not in use * Option 1.4: RTP is secured (SRTP), frame marking extension header is in use   In Options 1.1 and 1.2 the 5GS can identify rich set of information about the application traffic, e.g., for video individual slices and their types (H.26x NAL units) can be identified as well as information about the frame and its dependencies. In Option 1.4 slice level information is not visible, but frame start, end, type and dependencies can be identified. In Option 1.3 the information is scarce, possibly only frame start and end can be identified. This means the GTP-U header should have optional fields and RAN will get a different level of information in the GTP-U header depending on the case.  Option 3 does not require standardization of PDU set identification and is thus supported by default.  Option 2 requires standardization on N6, which only adds to the many alternatives already defined in Option 1, and in the case of Options 2.2 and 2.3, creates a dependence on outside standards bodies. Option 2.3 is important as it can provide more information about the traffic, especially about the video slices which are not identifiable when Secure RTP is used (see Options 1.3 and 1.4 above). It should be done in collaboration between SA4 and the IETF . The same applies to Option 2.2.  Option 4 is beyond the current and planned (Rel. 18) scope and capability of the NWDAF. While eNA\_Ph3 (Rel 18, TR23.700-81 KI#2) is considering NWDAF assisted application detection, there is no consideration in the KI of real-time detection of media layer attributes that would be classified as a PDU Set. |
| OPPO | Support option 1, do not support option 2, 3, 4. | The existing RTP layer header including RTP fixed header, RTP extended header, RTP payload header can be used to identify the PDU Set info with no impact to the XR applications, and the RTP layer header is common used for any access technologies, e.g. 3GPP access or WIFI access. Although in some cases the RTP payload header and RTP extended header may be encrypted, the RTP Fixed Header can still be read, e.g. Marker, sequence number, etc. Therefore, we propose to use existing RTP/SRTP header, especially for the use of RTP Fixed Header Fields.  GTP-U protocol extension affects both 3GPP specs and application server-side app development, i.e., the application needs to support the packet encapsulation specially used for 3GPP access. We don’t think it’s reasonable for application to design different encapsulation headers for different access technologies.  HTTP/RTP header extension should be determined in IETF which is outside the scope of 3GPP. Hence, cannot be supported in this release.  Option3 and option4, the expected traffic patterns may not be matched considering the various conditions for the route between AS and UPF, e.g. jitter, packet lost, out of order packet, and UPF may spend extra time on packet buffering or interaction with other NFs for traffic characteristics detection. |
| Qualcomm | Option 1, Option 2.3 and Option 3 | Options 1 and 3 should be supported without requesting to from SA4. Option 2.3 should be confirmed by SA4.  Option 2.1 is not recommended because it would make the PDU Set enhancements strictly dependent on the adoption of a 3GPP defined interface.  Option 2.2 is not recommended because HTTP is not used for XR traffic.  Option 4 is not needed, because Options 1, 2.3 and 3 can address the issue without interacting with the NWDAF. |
| Samsung | Support Option 1, and Option 2.3 | Option 1 seems not to have standardization effort and if required e.g in SA4, RTP header can be extended to deliver some PDU set info, which can be updated and implemented in IETF by CT WGs.  XR server can deliver header information to UPF via existing or some new extended header. And option 3 implies non-standardization method which can not be excluded specfically. |
| Sony | Option 1 and 2.3 |  |
| Tencent | Option 1 should be supported. Option 2 can be further considered. | Option 1: The protocol in Option 1 has been ready and used in the industry. Supporting option 1 should be firstly considered for the XR traffic at this stage. The network needs to be cooperated with AF to get some information on the payload to be interpreted.  Option 2: We are basically neutral and open to option 2. All the options require enhancements and acceptance in the application side. Option 2.2 requires the further support in IETF; Option 2.3 requires further work in SA4 and IETF. Coordination with IETF is needed to support option 2.2 and 2.3.  Option 3: It’s not perfect enough to get the precise information, but it could be considered as a backup solution. |
| T-Mobile US | Option 1 and Option 2 (Prefer 2.2 and 2.3) | Operators will support many different XRM customers/applications, some will utilize the IETF existing markings, and thus implementations will need to be able to support these, even though they may be less optimal than applications that apply the HTTP or RTP extensions. The HTTP and RTP extensions are required to support operator provided XRM services (e.g. IMS XRM communications under development in FS\_RTC), and enable optimal performance for providers choosing to use these extensions. Extending GTP-U beyond the UPF is problematic as it changes the nature of the N6 interface. |
| Vivo | Option 1 when the RTP/SRTP is not transmitted in the encrypted protocol. | Whether RTP/SRTP header is always cleartext or not is related to SA4 reply for Q6. If the reply is yes, then we need to see to other solutions (e.g. option2, 3, 4) for traffic with encrypted RTP/SRTP header  “Q6: SA2 would like to ask SA4 whether there are XRM use cases where RTP/SRTP could be transferred over TLS/DTLS/QUIC and whether it implies encryption of the entire XRM media packet headers?”  Option2 requires close coordination between App server and 5GS and server need to support 3GPP user plane specification.  Both Option3 and Option4 are applied to traffic with specific characteristic, so can only applied to specific XR flow. Option4 applies ML learning based on the traffic characteristic. So its detection should be more accurate than option3. But the delay of Option4 is longer than option3 when the NWDAF is not collocated with UPF. |
| Xiaomi | Prefer Option 1 and Option 2.3 | Option 2.2: Typical scenarios for XR requires the content to be encoded in real-time in response to user interaction (e.g. head movement, interaction with objects, etc.). For these use cases, HTTP is thus not suited for delivering the real-time encoded content since it introduces incompressible delays due to the request/response pattern of the protocol. HTTP is better suited for on-demand content delivery and live content without any feedback loop between the receiver and the source. In addition, HTTP header are typically encrypted when HTTP is used over TLS.  Option 1 and 2.3: Since RTP headers may be visible to intermediate nodes, this solution may be considered for carrying such information from a media delivery point of view. However, other aspects should also be considered such as security, architecture, scalability, etc.  Option 3: Traffic characteristics may be very diverse from one XR service to another. For instance, one service may be using I frames which typically represent the largest type of frame in terms of bytes while another service may configure an video encoder to smooth out peaks by not using I frames such that the bitrate is more even over time. Additionally, bitrate may also vary based on the user interaction when the content is real-time encoded/packaged for each user based on the user pose. Therefore, this could be difficult to derive relevant information from traffic characteristics as the bitrate shapes can vary from one service to another and even per service. |
| ZTE | Support Option 1, Option 2.1 and option 3 in this TR. | Option 1 and option 3 do not have standardization effort.  Option 2.1 needs to specify the N6 interface which is normally out scope of 3GPP. However the XR server can deliver header information to UPF via non-standard header to assist XR traffic detection. We think this method falls into option 3.  Option 2.2 and 2.3 have dependency on IETF which seems not possible at this stage.  Option 4 is not feasible because NWDAF only provides not real time analysis. The prediction cannot be used by the UPF to do traffic binding. |

Summary:

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Apple | CATT | China Mobile | China Telecom | Ericsson | Futurewei | Google | Huawei | Intel | InterDigital | KDDI | Lenovo | MediaTek | Meta | Nokia | OPPO | Qualcomm | Samsung | Sony | Tencent | T-Mobile US | Vivo | Xiaomi | ZTE | Totol(24) |
| Option 1 | Y | Y | Y | Y |  | Y | Y | Y |  |  | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | 21 |
| Option 2.1 |  |  | Y |  |  |  |  |  | Y |  |  |  | Y |  |  |  |  |  |  |  |  |  |  | Y | 4 |
| Option 2.2 | Y |  |  |  | Y |  | Y |  |  | Y |  |  |  | Y |  |  |  |  |  |  | Y |  |  |  | 6 |
| Option 2.3 | Y |  | Y |  | Y |  | Y | Y |  |  |  |  |  | Y |  |  | Y | Y | Y |  | Y |  | Y |  | 11 |
| Option 3 |  | Y | Y | Y |  |  |  | Y |  |  | Y | Y |  |  | Y |  | Y |  |  |  |  |  |  | Y | 9 |
| Option 4 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | 0 |

Observation:

There is a majority support on option 1(based on existing RTP related RFC/draft) as a baseline. Some companies clarified option 1 is only for the applications that not fully encrypted.

11 companies are supportive on option 2.3(defining new RTP header) while some other companies show concerns on the progress of SA4 on new RTP header definition comparing to the Rel-18 timeline.

9 companies are supportive on option 3(based on UPF implementation) as an alternative method.

6 companies are supportive on option 2.2 (HTTP/MASQUE based N6) while others shown concerns on the matureness and feasibility.

4 companies are supportive on option 2.1 (using extended GTP-U tunnel in N6).

Proposed way forward:

1. Support option 1 as baseline for conclusion.

2. Option 2.3 can be supported in this release only in case SA4 confirms that the necessary extension can be defined within Rel-18.

3. Continue to discuss whether to support Option 3, Option 2.2 and Option 2.1. (Prepare for SoH during SA2 #153e to make decisions)

### Q2. How to deliver PDU Set importance information to RAN:

* Option 1: use different QoS Flows with different priority level. PDU Set importance is mapped to existing QoS flow priority.
* Option 2: use one QoS flow for different PDU Set with different priority level
  + Option 2.1: use different sub-QoS Flow within one QoS Flow, and using sub-QoS flow Identifier in GTP-U header
  + Option 2.2: use PDU Set importance information in GTP-U header

**[Company inputs]**

|  |  |  |
| --- | --- | --- |
| Company | Position | Justifications |
| Apple | Support option 1 | Option 1 relies on the existing QoS model and is least complex when realized without increasing the max number of DRBs. If PDU sets of different importance are mapped to different QoS flows, some enhancement may be needed for in-order delivery at the RAN side. But this could be manageable provided the right set of header information is available. As a second priority we also support Option 2.2. We do not support Option 2.1. |
| CATT | Support Option 2 and Option 2.1 is preferred. | For the Option 1, if different QoS Flows are used for the same PDU Set, it will introduce the complexity to re-order the received PDU set to get the same order. And it will introduce more delay and more dependencies between these QoS Flows.  In fact, the Option 2.2 can be implemented by the Option 2.1 since the importance information can be treated as the "sub-QoS Flow Identifier". |
| China Mobile | prefer option1, can also accept 2.1 | For option1, there is no specific requirement for NG-RAN, and is easy to be introduced.  For option 2.1, some enhancement in RAN maybe needed. For option2.2, seems no need to support. |
| China Telecom | prefer option 1. | Using different sub-QoS flows within one QoS Flow can basically achieve the same results as using different QoS flows, while sub-QoS profile configuration and sub-QoS Identifier are too complex. What’s more, if the concept of sub-QoS flow is introduced, this statement “The QoS Flow is the finest granularity of QoS differentiation in the PDU Session” in current QoS model would no longer work. |
| Ericsson | Option 1, but in general not supportive to introduce ‘importance’ information into the 5GS (see justification). | No input showing benefits from using ‘importance’ information was provided. Simulations we conducted shown that using importance as QoS flow priority has actually negative impact on the capacity and thus in our view it should not be introduced.  Brief analysis of the options:  --  Option 1: keeps the premise of the current QoS model that the QoS flow is the finest level of QoS differentiation. This approach has no impacts on the RAN. The main impact is on the SMF, which basically uses the same procedure for QoS flow selection with the addition of the importance as input parameter. SMF then just provides UPF with the mapping of importance values to QFIs in the N4 rules.    Option 2.1: a structure of QoS flows and sub-QoS flows is used to select a main QoS profile and a sub-QoS profile, with a mixture of QFI as such and XQFI as GTP-U extended header.    The same can be achieved by using in option 1 with separate QoS flows for PDUs with different importance, where the QoS profile for each QoS flow will have the settings of equivalent main QoS profile plus the ones of the corresponding sub-QoS profile.    Option 1 is much more straightforward and allows to keep the same handling in the RAN, with the only addition of the importance as input parameter for QoS flow selection at the SMF.    Option 2.2: in this option, a QoS flow is no longer the finest level of QoS differentiation, a premise of the current QoS model. Handling of priorities in the RAN needs to be modified to take into account this new parameter, while option 1 proves that this is unnecessary. |
| Futurewei | Support Option 2.  Option 1 is complex. | Importance information (PDU set priority) along with PDU set sequence number (PSSN), PDU set GBR (PS-GBR) and PDU set MBR (PS-MBR) are necessary for gNB in our view.  Prefer option 2.2 - importance plus PSSN, PS-GBR, PS-MBR, but option 2.1 where sub-QoS flow indicates importance/priority for the PDU set plus PSSN, PS-GBR, PS-MBR is quite close (the concept of a sub-QoS “flow” is a bit confusing though).  Option 1 is complex.  The limiting factor regardless of the number of QoS flows is eventually DRB queues (so whether there are 10 queues or 1 queue, it will be mapped to the same limited set of DRBs.) And handling more (virtual) queues is relatively more complicated. |
| Google | Support Option 2 including both of Option 2.1 and Option 2.2 | Option 1 has issue of handling out of order PDU Sets in different QoS flows for the same SDF of a XRM traffic at RAN.  Option 2 has no issue of out of order PDU Sets in a QoS flow with different Sub-flows. This option provides more granularities of QoS flow to RAN.  Both Option1 and Option2 require to provide QoS profile to RAN for the QoS requirement/characteristics of PDU Set. |
| Huawei | Support Option 1 | PDU Sets with different priority level can be transferred via different QoS Flows with corresponding QoS parameters (e.g. PER/PSER). Option 1 can reuse existing RAN capability for such PDU Set importance-based scheduling. Option 2.1 and 2.2 break current QoS Flow framework without clear benefits, and will also cause significant impacts to RAN DRB based scheduling.  One main concern on option 1 was the potential packet dis-ordering caused by different QoS flows. However, we don’t think it’s a really issue due to the following reasons:  1) Different QoS Flows can have different priority but same PDB/PSDB values, since the delay requirements are same no matter whether the PDU Set is important or not. Once PDB/PSDB is satisfied, disorder should be avoidable.  2) The intervals between frames are usually > 10ms (e.g.16.6 ms for 60fps video), where the PDB requirement of XRM services are on the same level or even lower. The possibility of disordering between sequent frames should be very low considering the big interval comparing to the PDB.  3) Media layer protocols like RTP, are able to deal with disorder if it happens using the timestamp info in the RTP header. |
| Intel | Option 2.2. | Option 2.2 of Q2 (together with Option 2.1 of Q1) relies on the traditional IP 5-tuple for QoS Flow binding, the additional information for PDU Set importance being conveyed in-band. This option preserves the order of packet arrival. When combined with Option 2.1 of Q1 it implies simple copy/paste of the PDU Set importance information received on N6. In our view the PDU Set importance information should primarily be understood as “discard eligibility” i.e. it is used by NG-RAN primarily for the purpose of discarding entire PDU Sets of lower importance in presence of congestion. It is possible that NG-RAN could use the PDU Set information for additional purpose (e.g. setting of more relaxed parameters for transmission on the radio interface), but this is up to RAN2 WG to discuss.  In contrast, Option 1 increases the number of QoS Flows and may lead to out-of-order delivery if the QoS Flows are mapped to different DRBs by NG-RAN. Option 2.1 introduces unnecessary complexity implying that a “QoS Subflow” could have a set of 5QI characteristics. |
| InterDigital | Option 1 | Option 1 is preferred because it maintains the principle that packets that require the same treatment get mapped to the same QoS Flow. We are not opposed to conveying additional information in the GTP-U header. We do not see a good reason to introduce QoS sub-flows. |
| KDDI | No strong opinions for each option. |  |
| Lenovo | Option1 and Option2.2 are preferred. | * Option 1 simplify RAN node handling and may complicate UPF handling. As we know, RAN node schedules based on logical channel. By binding logical channel with specific frame type/importance in Option1, RAN node is able to perform differentiated QoS handling without introducing new scheduling mechanism. That is, QoS flow should first be associated with specific frame type/importance. Then 1:1 mapping of QoS flow and radio bearer (i.e., logical channel) makes logical channel to be associated with specific frame type/importance, which facilitates RAN to perform differentiated QoS handling. For slice based traffic model defined in TR 38.838, there’re I-stream and P-stream respectively. It is easily to map one stream into one QoS flow. However, if the I-frame and P-frame are in the same service data flow, then it is a bit complicated for UPF to generate I-frame specific PDU set SN and P-frame specific PDU set SN respectively. * Option 2.2 simplifies UPF handling but complicates RAN node handling. That is, there’s different PDU set importance within a single logical channel. The logical channel based scheduling is not applicable anymore. A more complicated scheduling mechanism shall be designed by taking the PDU set importance information in GTP-U header into consideration.   In our point of view, it is hard to decide which one is better. So it is preferred to support both options. |
| MediaTek | Option 2.2 is preferred, but Option 1 is also viable | Option 1: Is feasible without further changes to the specification.  Option 2.2: When packet losses occur, the importance information allows the RAN to opportunistically discard less important PDU Sets on the fly. In addition, complexity is lower. |
| Meta | Support Option 1 | Option 1 is preferred to minimize the impact to current 5GS by reusing existing QoS framework as much as possible.  Option 2.1: It is not clear on how this resolves the problem as sub-QoS flow identifiers will have to be mapped to a “relative factor” somehow for RAN to consume this info.  Option 2.2: Not preferred as it is not clear why this is better than option 1. |
| Nokia | Support Options 2.1 (preferred) or Option 2.2. | Options 2.1 and 2.2 are similar. Both have an indication that distinguishes PDU Sets in the GTP-U header. How PDU Sets are handled in the NG-RAN should be determined by PDU Set QoS parameters, such as PSDB, PSER and PDU Set Priority sent via NG-AP, where a separate set of QoS parameters may be provided for each “importance” or “QoS Sub-Flow” designation in the GTP-U header. Option 2.2 provides the additional feature that the UPF has knowledge of relative importance of PDU Sets it identifies.  Option 1 introduces out-of-order PDUs at the UE / application client, with possible unanticipated consequences. |
| OPPO | both Option1 and Option 2.2 could be possible. We do not support option 2.1 | Whether to use different QoS flows or one QoS flow should depend on the design of application server. If an SDF generated by AF only contains the PDU sets with the same importance, the PDU sets with different importance carried in different SDF can be mapped into different QoS flows, e.g., I-frames use the same IP 5 tuple filter, while P-frames use the other IP 5 tuple filter. If an SDF generated by AF contains the PDU sets with different importance, the PDU sets are mapped into a single QoS flow, e.g., I-frames and P-frames share the same IP 5 tuple filter.  In our view, the Q2 is strongly correlated with the Q3. if the answer to Q3 is to identify accurate dependency relationship between PDU Sets for scheduling, it would be better to use one QoS flow, as it is difficult to associate different PDU Sets in different QoS Flows. |
| Qualcomm | Option 1 | The addition of sub-QoS Flows or PDU Set importance information in the GTP-U header would introduce unnecessary complexity in the system. |
| Samsung | Support Option 2.2 | We don’t see the need of sub-QoS Flow concept, but such PDU Set info can be additionally used for better scheduling. |
| Sony | Option 2.2 and possible 2.1 |  |
| Tencent | Support option 2. | Option 1: It’s not preferred to separate the PDUs within the same service data flow into the multiple QoS flows. It may also need to define some cooperation between these QoS flows. This option would limit the further PDU set identification and usage within the XR service.  Option 2: Using one QoS flow for different PDU sets is more preferred and aligned with the service expectation. Option 2.1 and 2.2 are similar, and would be fine for us. |
| T-Mobile US | Option 2 (Prefer 2.2) | Option one while simple will quickly exhaust QOS Flows in the network, complicate RAN mapping, and make troubleshooting difficult. We prefer Option 2.2 as this will provide better backwards compatibility to existing services (which don’t use sub-QOS flows) and allowing RAN to continue to provide prioritization and other QOS actions for each packet. |
| Vivo | Option1 is preferred. | The importance information is provided because of the importance differentiation based scheduling at RAN. The scheduling granularity is DRB based.  In Option1, although it doesn’t align the existing service data flow/QoS flow mapping, mapping one service data flow into multiple QoS flow can avoid extra RAN’s complexity. RAN can reuse the existing QoS flow/DRB mapping at most.  Option2 follows the existing mapping rule between service data flow to QoS flow. But RAN further needs identify the sub QoS flow ID or importance information per PDU and design new mapping mechanism, which is extra complexity to RAN.  Also both Option2.1 and Option2.2 need extra bits overhead in GTP-U for each PDU. The burden for N3/N9 backhaul cannot be ignored if the XR data amount is great. Option2.2 is more straightforward than Option2.1. |
| Xiaomi | Prefer Option 1 and Option 2.2. | Option 1 and 2.2 have less impact with the procedures handling from our point of view. Option 1 can reuse the current QoS model. Option 2, the RAN can reuse the priority handling with the importance information considering. |
| ZTE | prefer Option 2.2. | Option 1 may cause out of order problem since the RAN may bind the QoS Flows into different DRBs.  We don’t see the need of define sub-QoS Flow concept. It is not meaningful to support multiple priorities within one QoS flow. Actually we don’t see any difference with multiple QoS flows.  We also think Option 2.2 should be considered together with Q3. |

Summary:

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Apple | CATT | China Mobile | China Telecom | Ericsson | Futurewei | Google | Huawei | Intel | InterDigital | KDDI | Lenovo | MediaTek | Meta | Nokia | OPPO | Qualcomm | Samsung | Sony | Tencent | T-Mobile US | Vivo | Xiaomi | ZTE | Totol(24) |
| Option 1 | Y |  | Y | Y | Y |  |  | Y |  | Y |  | Y | Y | Y |  | Y | Y |  |  |  |  | Y | Y |  | 13 |
| Option 2.1 |  | Y | Y |  |  | Y | Y |  |  |  |  |  |  |  | Y |  |  |  | Y | Y | Y |  |  |  | 8 |
| Option 2.2 |  |  |  |  |  | Y | Y |  | Y |  |  | Y | Y |  | Y | Y |  | Y | Y | Y | Y |  | Y | Y | 13 |

Observation:

No clear majority view is identified. It seems option 1 and 2.2 have more supports than 2.1.

Proposed way forward:

Continue to discuss on the key concerns of each option, e.g. whether mis-ordering could happen for option 1, impacts on RAN scheduling for option 2. (Prepare for SoH during SA2 #153e to make decision)

### Q3: Support to PDU Set dependency-based scheduling

* Option 1: Identify accurate dependency relationship between PDU Sets for scheduling.
* Option 2: In some scenario (e.g. closed GOP), the decoding of the non-I frames between two successive I frames always directly or indirectly relies on the 1st I frame of the two successive I frames. If the 1st I frame is in error, the non-I frames can be dropped until the next I frame. (proposed in S2-2205839)
* Option 3: If a PDU Set is depended by others, it can be considered as more important during scheduling. But the scheduling will not further consider the accurate dependency relationship.

**[Company inputs]**

|  |  |  |
| --- | --- | --- |
| Company | Position | Justifications |
| Apple | Support Option 1 as a first preference. | Option 1 is the preferred solution as it provides RAN with accurate information for dropping packets, driving capacity management. As a second option, we can go for option 3 (or a combination of Option 1 and 3 depending on the information available in the 5GS) |
| CATT | Support Option 3 | In the new video codec H.265/266, it is very hard to accurate identify the dependency relationship between PDU Sets and these dependency is changed very quickly if multi-view, multi-layer video codec are used, i.e. the Option 1 cannot be well implemented.  For the Option 2, it does not consider the multi-layer, multi-temporal video codec used in H.265/H.266, and the solution is workable for very limited part of the video stream and cannot work for most part of the video stream.  Option 3 is more general and can work for all kinds of video codecs. |
| China Mobile | Prefer option2, can also accept option3 | We consider option2 as a valid scenario for media service, even this is not a common feature for all media service. We can also accept option3 and considered as a more simple way to help consider the PDU set dependency feature. |
| China Telecom | support Option 2 and Option 3. | Option 1: It is too complex to explore all kinds of dependency relationship between PDU Sets.  Option 2: Since the accurate dependency relationship is too complex and this is a possible scenario, we can start with such a simple relationship.  Option 3: In our understanding, if a PDU Set is depended by others, AF/UPF will consider this PDU Set to be important. Based on such “importance” information, the RAN then perform scheduling for this PDU Set. This is reasonable. |
| Ericsson | Not supportive to providing dependency information in general. | There is no evidence that this information is providing any benefits compared to the basic functionality relaying on PSDB.  Hence in our view there is no justification to add complexity to XRM related functionality in the specifications. |
| Futurewei | Support Option 3. Options 1, 2 are more complex and may be too rigid to easily adapt to different/future encoding schemes. | The extensions to QoS for XR should reflect the handling and optimizations needed for general patterns of media such as treating a set of packets and their associated characteristics as a group. Additions to QoS consisting of importance and related PDU set parameters that RAN can use to optimize XR handling are needed.  Since XR and media applications are continuously evolving, specific mechanisms for a particular encoding may be better to avoid as the application space, codecs and packetization techniques are likely to change at a faster rate than lower layer network enhancements (such as for gNB). As applications change, richer meta-data should be provided and the UPF should “map” the meta-data in N6 to N3 accordingly. But the QoS enhancements over N3 should only have to evolve as the gNB/RAN capabilities evolve. |
| Google | Support Option 3. | Option 3 is simple and allows more general use cases for applications. Other options need great efforts at UPF but may provide limited optimization for RAN. SA2 can wait for RAN decision on whether such scheduling optimization is needed/beneficial and then decide whether to support identifying accurate dependency at 5GC. |
| Huawei | Prefers to Option 3, but open to discuss the feasibility of Option 1/2. | We have concerns on the feasibility of option 1. Identifying accurate dependency relationship between frame/slice. Option 2 tries to simplify the solution by identify the I-frames and non I-frames only, while it is limited to certain scenarios.  PDU Set importance could denote whether a PDU Set is depended by others or not (see RTP NRI, ‘I’/‘D’ fields as explained in e.g. solution 24/17/18). But we are open to discuss the feasibility of supporting option 1/2. |
| Intel | Leaning towards Option 3, but open for further discussion. | In our reading of the SA4 reply LS (S2-2203658), SA4 are advising against any handling of PDU Set dependencies at transport level because corresponding use cases that would benefit from such handling seem difficult to identify (refer to the text below from S2-2203658). According to Option 3, the “reference PDU Set” could be indicated explicitly (e.g. in the GTP-U header) or can be indicated implicitly by assigning a high PDU Set priority to it. |
| InterDigital | Option 1 | Option 1 is preferred because it can provide the same benefits as option 2 and 3. However, option 1 is more flexible in the sense that it is not tied specifically to the I frame / non-I frame scenario and it provides more information to the RAN than option 3. Although it should be acknowledged that option 1 requires more interaction with both SA4 and RAN WGs. SA2 should work with SA4 to identify what dependency information can be conveyed to RAN. RAN can then decide if and how the dependency information can be used. |
| KDDI | Support Option 3 | For Option 1, it’s difficult for UPF to identify accurate dependency relationship, especially spatial-temporal coding use cases. Option 2 is only used for limited use cases such as closed GOP. Option 3 approach is simple and available to various video coding use cases. |
| Lenovo | more related with RAN WG | Lenovo thinks Q3 is more related with RAN WG. And from our point of view, how dependency information can be provided to RAN shall be first discussed. E.g., to introduce “dependency information” in GTP-U header or to configure dependency over control plane. However, Option2 depends on SA4’s response for Q5 in LS out S2-2207887. If it is true, it is more CP like solution. E.g., AF indicates or specification states that the decoding of non-I frames between two successive I frames always directly or indirectly relies on the 1st I frame of the two successive I frames.  Regarding Option 1 and Option 3, whether accurate dependency relationship will be further considered for scheduling depends on RAN WG’s decision. |
| MediaTek | No support for the options listed. | Option 1: Undesirable since it requires deep packet inspection of the NAL units in each PDU Set. Also from SA4 reply LS (S2-2203658 Q4): “For example, the handling of dependent PDU Sets once a leading PDU Set is lost is not universally defined and depends on the operation of the application”  Option 2: The LS reply from SA4 (s2-2203658, Q2) cast doubt on this scenario: “In particular, low-latency XR and cloud gaming video services such as Split-Rendering or Cloud Gaming typically would not use the traditional coding structure with a fixed Group of-Picture (GOP).”  Option 3: The method by which the dependency is derived, as well as the metric used to determine its significance, are not explained, and requires coordination with SA4. |
| Meta | Support Option 1 | Option 1 is preferred as it makes the system more predictable (i.e., provides accurate information to RAN for RRM).  Option 2 is not preferred as this only applies to some scenarios. We don’t think this type of fixed relationship is good for future evolution when compared to option 1.  Option 3 can also be considered as a basic option as this seems to be less complex. |
| Nokia | Support Option 1. | The GTP-U header should support dependency information as optional parameter(s). Whether it is possible for an implementation to included dependency information depends on which Options in Q1 are supported by the implementation and what protocols the actual traffic is using, see the sub-options 1.1 to1.4 in the answer to Q1. Dependency should be explicitly indicated without assumptions in the NG-RAN of the media-type. The NG-RAN is not aware of the specific media (e.g. that a PDU set is an I-Frame or a P-Frame). Option 2 does not work, for example if the PDU Set is a slice (where slices within a Frame are independent) rather than a Frame. Option 3 does not allow the coupling of dependent PDU Set transmission on the successful transmission of the PDU Set on which it is dependent. |
| OPPO | Support option 3. | Although identifying accurate dependency relationship facilitates scheduling, it’s difficult to figure out the accurate dependency relationship between PDU Sets. In addition, Option-1and Option-2 require PDU Sets with different PDU Set importance to be mapped into the same QoS flow. Otherwise, it is complicated to support coordination between multiple QoS flows.  Option-3 is much simpler and more general than others. |
| Qualcomm | None of the above | The scheduling in the RAN should be the regular scheduling based on the QoS characteristics values of the QoS flow, n particular the QoS Flow’s priority level. |
| Samsung | Support Option 1 and possibly Option 2 | RAN need to know dependency and importance of each PDUs which belong to PDU set(i.e. PDU set’s dependency and importance) then other PDU sets. Option 2 can be a partial solution. So need more general way(i.e Option1) |
| Sony | Option 3, if time permit, we could also do option 1 and/or 2 |  |
| Tencent | Support option 1, 2, 3. | Considering dependency would be helpful for the scheduling.  Option 1: Identifying dependency relationship can be possible for certain options in Q1. This option has higher complexity than option 3.  Option 2 is similar with option 1, and also relies on the LS out from SA4.  Option 3 has lower complexity than option 1 and 2. It is not clear how the dependency information would be used for the scheduling. Some clarification may be needed. |
| T-Mobile US | Option 3 | Simpler to use a priority; if priority is low and the available BW is limited then the unimportant packets/frames will be dropped first, if BW is sufficient then packets/frames don’t need to be dropped. |
| Vivo | Option2(pending to SA4 reply) and also OK without PDU Set dependency | The controversial point is PDU Set dependency-based data dropping, not about the scheduling.  Firstly, PDU Set dependency-based data dropping is not naturally applied but need consent from App. Based on that, for some XR traffic with AF providing consent for PDU Set dependency-based data dropping, the three options can be evaluated:  -Option1 has ideal motivation but dependency relationship cannot be exhausted.  -Option2 relies on SA4 reply on Q5. If the reply is yes, this comparably simple dependency can be considered.  - “Q5: SA2 would like to ask SA4 whether the following scenario exists for some XR service flow: The non-I frames (e.g., P frame or B frame) transmitted/decoded between two successive I frames directly or indirectly refers to the 1st I frame of the two successive I frames?”Option3 in more related to importance section, it is not clear whether the less importance packet can be dropped or not  If there is no possibility of APP consent for PDU Set dependency-based data dropping, we are also OK without PDU Set dependency. |
| Xiaomi | Open for Option 3. | This kind of assumptions are hard to make since it depends on what is in the PDU set and then what exact dependency exists between the PDU sets, i.e. coding dependency, parsing dependency, etc. |
| ZTE | Option 1 | The RAN needs to know which PDU set that the current PDU set is depending on, otherwise it is meaningless to use this parameter for scheduling.  For option 2 we believe it is possible scenario.  For Option 3, it only addresses how to schedule the more important PDU set, but it does not address how to schedule the less important PDU set if the more important PDU set is not transmit successfully. So it is not full solution option. |

Summary:

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Apple | CATT | China Mobile | China Telecom | Ericsson | Futurewei | Google | Huawei | Intel | InterDigital | KDDI | Lenovo | MediaTek | Meta | Nokia | OPPO | Qualcomm | Samsung | Sony | Tencent | T-Mobile US | Vivo | Xiaomi | ZTE | Totol(24) |
| Option 1 | Y |  |  |  |  |  |  |  |  | Y |  |  |  | Y | Y |  |  | Y | Y | Y |  |  |  | Y | 8 |
| Option 2 |  |  | Y | Y |  |  |  |  |  |  |  |  |  |  |  |  |  | Y | Y | Y |  | Y |  |  | 6 |
| Option 3 |  | Y | Y | Y |  | Y | Y | Y | Y |  | Y |  |  |  |  | Y |  |  | Y | Y | Y |  | Y |  | 13 |

Observation:

A slight majority (13 companies) supports option 3, which basically means we don’t support dependency between PDU Sets, but only consider the importance of PDU Set instead (using methods discussed in Q2).

On option 1 and 2, some companies claim the benefits and flexibilities, while other companies express concerns on feasibility and benefits.

Proposed way forward:

Continue the discussion on the feasibility and benefits of option 1/2. Option 3 will be a default conclusion (i.e. no special support on dependency between PDU Sets) if no consensus can be reached. (Prepare for SoH during SA2 #153e to make decision)

### Q4. Support to hierarchical PDU Set:

* Option 1: introduces PDU Set group. (S2-2205938)
* Option 2: not support.

**[Company inputs]**

|  |  |  |
| --- | --- | --- |
| Company | Position | Justifications |
| Apple | Option 2 | We are not convinced whether this added complexity would bring in any benefit. |
| CATT | Select Option 2: Not Support | The proposed solution only works well for very limited scenarios. |
| China Mobile | Suggest clarifying the concept of PDU set group and the necessity of PDU set group related parameters | If we consider PDU set as frame, the PDU set group can be considered as GOP. For GOP, the possible use case is if the I frame of the GOP is dropped, the other frames of GOP should be considered with low priority or also be dropped. |
| China Telecom | neutral | On the one hand, we accept that PDU Set Group can indicate frame that may be dependent on each other; on the other hand, the introduction of PDU Set group would certainly make things more complicated. Therefore, we take a neutral stance. |
| Ericsson | Not supportive | In our understanding there is no dependency relation that can’t be expressed by means of the PDU Set concept.  Furthermore, there is no need to understand what information is carried by some PDU Set and have understanding of frame or slice concepts (as recommended by SA4). |
| Futurewei | Support Option 2. Options 1 is more complex and may be too rigid to easily adapt to future encoding schemes. | The extensions to QoS for XR should reflect the handling and optimizations needed for general patterns of media such as treating a set of packets and their associated characteristics as a group. Additions to QoS consisting of importance and related PDU set parameters that RAN can use to optimize XR handling are needed.  Since XR and media applications are continuously evolving, specific mechanisms for a particular encoding may be better to avoid as the application space, codecs and packetization techniques are likely to change at a faster rate than lower layer network enhancements (such as for gNB). As applications change, richer meta-data should be provided and the UPF should “map” the meta-data in N6 to N3 accordingly. But the QoS enhancements over N3 should only have to evolve as the gNB/RAN capabilities evolve. |
| Google | Option 2 | this is related to Q3 and we stay the same views as Q3. No need to add complexity at UPF at this stage, which may provide limited optimization at RAN. |
| Huawei | Support option 2. | Hierarchical PDU Set brings complexities considering one PDU might belongs to multiple PDU Sets in different levels, which leads to a complicated handling determination for the PDU. The use cases are not fully justified yet. |
| Intel | Option 2 | Unnecessary complexity without a clear benefit. |
| InterDigital | Option 1 | Similar to our answer for Q3, closer coordination with SA4 and RAN is required. |
| KDDI | Support Option 2. | NAL unit header contains video slice type (e.g. I slice, P slice) and the dependency information of the video flame to which the video slice belongs. When the PDU Set is a video slice, the PDU Set can indicate the flame dependency by using NAL unit header or RTP header (defined in draft-ietf-avtext-framemarking). |
| Lenovo | Option2 | S2-2205938 introduces PDU set group concept serves as an alternative to dependency information in the GTP-U header. For slice based traffic model, the PDU set is a “slice” and the PDU set group is a “frame”. It introduces “PDU set group SN” into dependency information. It is too complex by introducing the concept of PDU set group. There may be “PDU set group SN”, “PDU set SN” and “SN within the PDU set” in one GTP-U header. In our understanding, P-slice of one frame dependents on the I-slice of the previous frame. Therefore, we can still use “PDU set SN” to infer dependency information for the slice based traffic model.  Besides, dependency information in GTP-U header is not the only option. In LS out S2-2207887, Q4 is also related with dependency, i.e., the non-I frames (e.g., P frame or B frame) transmitted/decoded between two successive I frames directly or indirectly refers to the 1st I frame of the two successive I frames?  If the answer is yes, we can refer the dependency information from the frame type and the time relation without introducing the “dependency information”. |
| MediaTek | Option 2 | The value of grouping has not been adequately demonstrated, and its application to XR media transportation requires the support of SA4. SA4 have frequently mentioned the need for coordination in their LS (S2-2203658). |
| Meta | Still TBD on option 1. | It is still unclear on how hierarchy is enhancing the scheduling capability in RAN when compared to just having QoS attributes related to PDU sets (and their relative dependency information). |
| Nokia | Support Option 1 | PDU Set group is necessary to indicate dependency of P-Frames on I-Frames when the PDU Set is a Slice. In addition, the PDU Set Group concept is extensible to support independent and dependent Temporal and Spatial layers for scalable media. |
| OPPO | Prefer option 2. | PDU Set group may be useful in some cases, but it is relatively complicated to introduce PDU Set group. It’s also difficult for UPF to identify the PDU Sets which belong to the same group. We don’t expect UPF to read the info from the layers beyond RTP layer, e.g. H.264, it’s more complex for the NF to do that. |
| Qualcomm | Option 2 (do not support) | The addition PDU Set groups would introduce unnecessary complexity in the system |
| Samsung | Support Option 2 | We don’t see the need of PDU Set group without considering interworking/handling multiple service flows. |
| Sony | Option 2 |  |
| Tencent | support option 1 with some conditions. | The PDU Set group is not clearly defined in the TR. It’s not clear to us which PDU sets can be formed as the PDU Set group. It would be better to have clear definition on it before our final conclusion. If PDU set group can help the handling of the PDU set, we are fine to support it. |
| T-Mobile US | Option 2 | PDU set group further complicates the network and UE. |
| Vivo | Option2 in this release | RTP and SRTP transmits the NAL Unit, which is slice based. When a frame includes multiple slices, it is unclear any co-operation is needed among those slices of one frame. Note: the non-I slice is not necessary refers to the I slice within the same frame. |
| Xiaomi | Prefer Option 2 for this release. | Handling of PDU set is already hard to grasp from a media delivery point of view. It is even more unclear how a PDU Set group can be beneficial. In addition, if introducing PDU Set group, whether the importance or priority between the PDU Set groups or not. It can be introduced in next release after further study about the details of PDU Set group. |
| ZTE | Option 1 | In order to indicate the accurate dependency relationship between PDU Set, we believe PDU Set Group is one simple solution, i.e. the first PDU set within the group is more important than the other PDU set within same PDU set group. |

Summary:

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Apple | CATT | China Mobile | China Telecom | Ericsson | Futurewei | Google | Huawei | Intel | InterDigital | KDDI | Lenovo | MediaTek | Meta | Nokia | OPPO | Qualcomm | Samsung | Sony | Tencent | T-Mobile US | Vivo | Xiaomi | ZTE | Totol(24) |
| Option 1 |  |  |  |  |  |  |  |  |  | Y |  |  |  |  | Y |  |  |  |  | Y |  |  |  | Y | 4 |
| Option 2 | Y | Y |  |  | Y | Y | Y | Y | Y |  | Y | Y | Y |  |  | Y | Y | Y | Y |  | Y | Y | Y |  | 17 |

Observation:

The majority view (17/24) is not to support hierarchical PDU Set or PDU Set group concept in this release.

Proposed way forward:

Not support to hierarchical PDU Set or PDU Set group.

### Q5. On “*Whether to drop a PDU Set in case PSDB is exceeded*”, do we need further define “*PDU Set Discard Time*” (A PDU Set shall be dropped in case this time is exceeded (sol 25 etc):

* Option 1: Support
* Option 2: not support.

**[Company inputs]**

|  |  |  |
| --- | --- | --- |
| Company | Position | Justifications |
| Apple | Option 2 | An option to drop PDU sets in quite important from RAN capacity point of view. However, looking at the existing QoS model we only have PDB and currently it does not require another parameter for discard timer. How the discard timer is set (e.g., in the RAN node or in the UE’s PDCP) is up to the RAN. This can be the same for XR. In other words, PSDB is essential, but PSDT is not absolutely needed. More important would be the information if a PDU Set can tolerate loss of packets. |
| CATT | Select Option 2: not support | In case PSDB is exceeded, it is not clear how to handle the first part packets of the PDU set that have been delivered to UE. The remaining buffered PDUs of the PDU set can be dropped if the PDB of these PDUs is exceeded or may be continually sent to the UE if the importance of the PDU set (e.g. the I-Frame) is very high. |
| China Mobile | Option2 | This aspect can be realized by RAN implementation. |
| China Telecom | option 2 | We should first reach an agreement on the definition of PSDB. And in my understanding, RAN can drop packets once the transmission time exceed PDB/PSDB. To avoid duplicated effort, it is unnecessary to define PDU Set Discard Time for now. |
| Ericsson | Not supportive. | An option to drop PDU sets in quite important from RAN capacity point of view. However, looking at the existing QoS model we only have PDB and currently it does not require another parameter for discard timer. How the discard timer is set (e.g., in the RAN node or in the UE’s PDCP) is up to the RAN. This can be the same for XR. In other words, PSDB is essential, but PSDT is not absolutely needed. More important would be the information if a PDU Set can tolerate loss of packets. |
| Futurewei | Support Option 2 (but qualified, please see description). | In our view, both PDB and PDU set delay budget are not necessary as it resolves to the same (since each packet of the PDU set is from media that has the same timestamp.) PDB is therefore sufficient.  Regarding the discarding of packets after the expiry of a discard timer, we agree that RAN should continue to forward packets in some cases even beyond the delay budget. This can be related to the importance of the packet and perhaps a delta value by which to extend the discard time.  Note: PDU set size (or burst size) may be useful in addition to MDBV (Maximum Data Burst Volume) for RAN to reserve resources accurately. However, PDU set size/burst size information is not available in RTP/SRTP headers and statically configured values are not useful as the PDU set size can vary based on network congestion, codec level changes and content. RAN should have accurate values for it to be useful. |
| Google | Option 2 | unnecessary complexity for introducing two related parameters. Whether to drop a PDU Set in case PSDB is exceeded is RAN implementation as what it does when PDB is exceeded and depending on the resource type. |
| Huawei | Support option 1 | The media server could provide more accurate valid time of PDU Set, i.e. how long time a PDU Set will considered as useful by the application. Such a time may be different based on the application codec configuration, e.g. ‘--ref’ parameter in x265 codec (https://x265.readthedocs.io/en/master/cli.html#cmdoption-ref). |
| Intel | Option 2 (or FFS). | SA2 should focus on the PSDB definition. Any need for “discard time” will be assessed based on the agreed PSDB definition. |
| InterDigital | Option 1 | In some cases, the PDU(s) become worthless if the PSDB is exceeded, thus it is preferred to discard them. |
| KDDI | Open for further discussion. |  |
| Lenovo | Option2 | As stated in 23.501, PDB can be utilized to handle the “expired” packets, i.e., discard or delivered. We don’t see the necessity of introducing a new parameter for packet discarding.  -The PDB for Non-GBR and GBR resource types denotes a "soft upper bound" in the sense that an "expired" packet, e.g. a link layer SDU that has exceeded the PDB, does not need to be discarded and is not added to the PER. However, for a Delay-critical GBR resource type, packets delayed more than the PDB are added to the PER and can be discarded or delivered depending on local decision. |
| MediaTek | Option 1 | When PSDB is exceeded, which is primarily due to congestion, we support enabling discarding PDU Sets if discarding PDU Sets is allowed. When such an event occurs, we support lowering the encoding bitrate as indicated in the SA4 reply LS (S2-2203658 Q4): “typically, video applications prefer reducing the encoding bitrate in order to minimize congestion-related packet losses”. |
| Meta | Option 1 | PSDT allows more flexibility in the RRM to determine if a PS needs to be dropped or not (i.e., not just based on PSDB but also based on the expected arrival period of the incoming media packets). |
| Nokia | Option 1: Support | If the PDSB of an I-Frame is exceeded, it may not be useful for rendering, but it can be used for decoding of subsequent P-Frames. A separate PDU Set Discard Time or validity time should be provided as a QoS Parameter, where a separate “PDU Set Discard Time” or “PDU Set Validity Time” may be provided for each “importance” or “QoS Sub-Flow” designation in the GTP-U header (see response to Q2). This allows, for example, PDU Sets comprised of I-Frame PDUs (marked e.g. with importance=1) to have a different PDU Set Discard time from PDU Sets comprised of P-Frame PDUs (marked e.g. with importance=2). |
| OPPO | Have no strong opinions. |  |
| Qualcomm | Option 1 (support), but with clarification. | A PDU Set shall NOT be dropped if its PSDB is exceeded; on the contrary the 5G system should keep trying to deliver a PDU-set even after its PSDB is exceeded, since the PDU-set can be useful for decoding subsequent frames. The PSDT can be introduced and used to support the configuration of scheduling and link layer functions (e.g., the setting of scheduling priority weights and HARQ target operating points). |
| Samsung | Support Option 2 | The question can be answered after we conclude the PSDB |
| Sony | Option 1 |  |
| Tencent | Open now. The PSDB is not clearly defined. The definition needs to be firstly fixed. | We are open to the first question. The definition of PSDB is not quite clear now.  PDU Set Discard Time is not needed for now. |
| T-Mobile US | Option 1 | This is probably needed if only priority is used to ensure any un-needed packets/frames in the PDU set are dropped. Different services and PDU set types may need this timer to be different depending on setup configuration of the PDU set type. This will be something determined by the network operator or communicated via the NEF when the external AS setups the PDU set. |
| Vivo | Option2 | PDU Set discard time should be determined by RAN not in SA2.  Whether to drop a PDU Set in case PSDB is exceeded depends on whether the PDU Set is delay critical or not. |
| Xiaomi | Support (Option 1) conditionally. The discard time may be difficult to correctly set. | In theory, having a maximum time of reception is easy to understand and could be in general useful. However, it is more complex in case of video decoding. A frame (assuming that this is matching a PDU set) could be too late for its expected presentation time but could be needed for decoding the future frames. As a result, the discard time may be difficult to correctly set since it would require to look into the dependency between frames.  Another case is for instance a user who wants to pause the content for a certain duration. Assuming that the application will resume playback where it stopped the discard time may be completely different now for the next frames to come. As a result, the discard time may even be dynamic and dependent on user interaction with the application and not a property of the video stream itself which seems to make the use of it even more complex in practice.  In general, the PSDB is agreed as one of the baseline parameters of the PDU Set in the TR, then we support that it can be dropped conditionally in case PSDB is exceeded. With regard to the discard time, it is difficult to set correctly as justified above. |
| ZTE | prefer option 2 | The question can be answered after we conclude the PSDB.  So far we don't think the need to define PDU Set Discard Time. |

Summary:

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Apple | CATT | China Mobile | China Telecom | Ericsson | Futurewei | Google | Huawei | Intel | InterDigital | KDDI | Lenovo | MediaTek | Meta | Nokia | OPPO | Qualcomm | Samsung | Sony | Tencent | T-Mobile US | Vivo | Xiaomi | ZTE | Totol(24) |
| Option 1 |  |  |  |  |  |  |  | Y |  | Y |  |  | Y | Y | Y |  | Y |  | Y |  | Y |  | Y |  | 9 |
| Option 2 | Y | Y | Y | Y | Y | Y | Y |  | Y |  |  | Y |  |  |  |  |  | Y |  |  |  | Y |  | Y | 12 |

Observation:

No clear majority view is identified (9 vs. 12).

Proposed way forward:

Continue to discuss on whether “PDU Set discard time/valid time” is useful together with the definition of PSDB. (Prepare for SoH during SA2 #153e to make decision)