**3GPP TSG SA2 Meeting #153e S2-220xxxx**

**Electronic, October 10th-14th, 2022**

**Source: MediaTek Inc.**

**Title: KI#9: Evaluation and Conclusion**

**Document for: Approval**

**Agenda Item: 9.19**

**Work Item / Release: FS\_XRM /Rel-18**

***Abstract of the contribution:*** *This contribution provides an evaluation and conclusion for the following solutions:*

* *#34 "5GC and UE Assistance to RAN for CDRX Optimization"*
* *#35 "Provision of conditional QoS profile(s) "*
* *#61"Policy Determination Considering UE Battery Level"*
* *#74 "Power mode based QoS adjustment "*
* *#75 "QoS Enhancement to Support Trade-off of QoE and Power Consumption"*

# Introduction

Solutions #34, #35, #61, #74 and #75 were introduced to address KI#9: Trade-off of QoE and Power Saving Requirements. In addition, solution #34 addresses KI#8: Enhancements to power savings for XR services.

**Solution #34** "5GC and UE Assistance to RAN for CDRX Optimization" suggests the use of multiple CDRX configuration or adaptable CDRX configuration. The CN (SMF) provides RAN with information of the “(DL) communication pattern” allowing in turn the RAN to configure the UE with multiple CDRX configurations or adaptable CDRX configuration (e.g. as a function of frame rate, periodicity, GOP structure). In case of multiple CDRX configurations, the UE thereafter selects a CDRX configuration as a function of the DL traffic it has *started* receiving (upon notification from the application layer). The UE then signals to the RAN via AS signalling which CDRX configuration it uses. Thereafter, the GOP structure is used in the UE and the RAN to switch between CDRX configurations. In case of adaptable CDRX configuration, the UE can decide extending/reducing the CDRX cycle length as a function of the DL traffic it receives and UL traffic it is about to send.

**Solution #35** "Provision of conditional QoS profile(s)" describes a mechanism where the SMF provides conditional QoS profile(s) (and associated event i.e. condition) to the RAN. When an event is detected by the RAN (incl. UE interaction), it activates the conditional QoS profile associated with that event, it deactivates the previously used QoS profile, and it notifies the SMF accordingly. This solution relies on the AF or the RAN being able to associate UE power consumption or memory status to QoS profiles.

**Solution #61** "Policy Determination Considering UE Battery Level" argues the UE battery level may impact the user's experience since the high throughput requires the high-power consumption in the UE (and conversely low throughput leads to lower power consumption – which is in fact not always true since low throughput may be associated to high power consumption e.g. in poor coverage). The battery level value or the percentage of the battery life reported by the UE can be considered when determining the policy to adjust the QoS level of some specific services depending on the reported battery level.

**Solution #74** "Power mode based QoS adjustment" describes a mechanism to assist PCF to adjust the policy based on the power mode selected by the user (/UE). It relies on the UE informing with NAS signalling (PDU Session establishment/modification) what power mode it uses, *assuming* in turn the PCF adjust its policies/rules, and thus claims a trade-off throughput/latency/reliability and power consumption can be achieved.

**Solution #75** proposes CN provides multiple QoS profiles to assist RAN select QoS parameters based on UE overheating status.

This contribution discusses the above.

# Discussion

## Observations on the solutions

**Solution #34** relies on multiple signalling handshakes between the UE and the network in order to constantly change the applied CDRX configuration. These signalling handshakes cause additional UE power consumption. This solution claims it provides gains, but fails to demonstrate what power saving gains, if any, are actually achieved over existing mechanisms. It should be noted that power saving in the UE when entering low power mode is inherent to the low power mode itself (not in 3GPP remit), NOT to the proposed solution. The SA4 LS (s2-2203658) has also cast doubt on the use of GOP based traffic patterns: "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)". Last, this solution relies on GOP pattern tracking in the UE – which may or may not be feasible in the UE depending on UE implementation and architecture.

**Solution #35** fails to explain how the association event/QoS profile is actually done, what relationship exists between power consumption or memory status and QoS profile, and what granularity is used. It also fails to explain what memory in the UE is at play and how reporting of memory status is of any use. This solution also relies on time-related events without defining how these are defined, based on what criteria, how they are used and what benefits they provide. This solution also incurs additional signalling from the UE to the RAN and related increase of power consumption which is function of the event a) granularity b) frequency and c) hysteresis. Importantly, this solution fails to demonstrate, let alone illustrate, what power saving gains if any are achieved on the UE side – worse, this solution does not even claim any gain is achieved or what trade-off is actually obtained.

**Solution #61** fails to explain how policy adjustments are actually made depending on the reported battery level, with what granularity and based on what criteria. The solution also incurs additional signalling from the UE (and related increase in power consumption) which is function of the reporting granularity and hysteresis. Solution #61 also fails to demonstrate how any improvement is reached over the UE *itself* e.g. throttling the application, reducing its capabilities etc.

**Solution #74** fails to explain how such trade-off is achieved, how policies are adjusted, with what granularity and importantly whether any gain is obtained at all, whilst it however requires additional signalling handshakes between the UE and the CN (and related increase in power consumption). This solution also speaks about "ultimate performance mode" assuming it is a well defined notion – which it is not, given performance can be associated to e.g. data rate, latency, reliability, power consumption and any combination thereof i.e. ultimate performance is indefinite and certainly does NOT necessarily imply highest power consumption.

**Solution #75** fails to explain how reporting the overheating status actually helps, what trade-off is achieved, how policies are adjusted, with what granularity and importantly whether any gain is obtained at all, whilst it requires additional signaling between the UE and the RAN (and related increase in power consumption). Last it should be noted that this solution can also be counterproductive when combined with other proposed solutions (some UEs do in fact overheat when being charged i.e. plugged in).

**Observation 1:** Common denominators to the above Solutions #34, #35, #61, #74, #75 are:

- No single solution has demonstrated *any* gains will be achieved over existing solutions, let alone explained how gains or trade-off would materialize.

- No single solution has explained how the reporting/usage of battery level / power mode / memory status / overheating is used by the network in adjusting its policies under which criteria and with which granularity.

- All solutions require additional signaling (some with signaling handshakes) from the UE to the RAN or CN which incurs additional UE power consumption and additional cost in terms or radio capacity. Some of these solutions esp. reliant on battery level, memory status could incur significant signaling overload.

**Observation 2:** CDRX is a RAN concept – SA2 is not equipped to meddle with such notion. Any discussion, evaluation and conclusion pertaining to CDRX is out of SA2 scope.

## Battery Level

Neither "battery level" nor "relative battery life of the device" provide any deterministic quantity that can reliably be used on the network side. Indeed, an x% battery life with a high-capacity battery is clearly not equivalent to an x% battery life with a small-capacity battery. Battery consumption may also vary greatly depending on the activity of the device. A battery at x% could also be charging without there being therefore any "battery consumption" issue. Thus that x% battery remains provides no indication as to how long the battery will continue to run in the device for a given activity, and therefore provides no usable information that can be exploited by the network. Nor is there any reliable information allowing the network to reliably determine that dropping the QoS level would markedly improve the remaining battery life.

In addition, proposals relying on battery level events being triggered will lead to unexpected signalling load in the network. This could of course be mitigated with some hysteresis, however given the observations above, we question this solution is a solution to an *existing* problem. It costs additional signalling, UE power consumption and precious radio capacity without *fundamentally* resolving any problem.

**Observation 3:** Battery level and relative battery life of a device are not deterministic quantities that can be used reliably in a network.

## Power Mode

Today’s devices can, upon user input, enter a "low power mode" when their battery level drops below a given threshold – thus optimizing their behavior to conserve the remaining battery power. Entering/exiting low power mode is typically prompted to the end user who retains the ability to decide what to do. However, the key point of low power mode is that the device *itself* optimizes its own behavior incl. application behavior: for example such device can reduce its radio capabilities, its display activity, some animations, interrupt background application refresh, throttle down applications (e.g. reduced codec rates) etc. The device alone knows what should be done to preserve its battery life. No *direct* network involvement is necessary – i.e. the power mode is transparent (and rightly so) to the network.

**Observation 4:** The device alone knows (and may be able to learn) what should be done to preserve its battery life when entering a low power mode and the device activity is very dependent on the device itself. No direct network involvement is necessary.

Should the device decide to run an application at a lower resolution and/or lower frame rate to save power, it can decide to do so at the application, with no *direct* network involvement. If a device decides to deactivate some radio capabilities, it could also do so, using existing means.

**Observation 5:** The application in the UE can be throttled down when low power mode kicks in, transparently to the network. Existing means also allow a UE to deactivate some radio capabilities if needed.

Power mode is not a universally accepted definition but is instead very specific to a particular device (at least make and model). Low power mode in one device may have a very different meaning from low power mode in another device. The meaning of low power mode can also vary in time domain depending on the UE activity, application status etc. Similar to battery level, low power mode is simply not a deterministic variable that can be used reliably by the network.

As argued earlier, ultimate performance mode is not deterministic either – it is indefinite.

**Observation 6:** Power mode / performance mode are not deterministic attributes that can be used reliably in a network.

## Memory Status

One solution proposes reporting of a UE memory status – however it is important to remember that

a) there is not a single memory in the device; and

b) a memory status is useless when not associated to a known memory capacity; and

c) memory status is useless if not knowing precisely how memory usage is managed by the device; and

d) memory management is a highly optimized process in a device and very device-dependent.

**Observation 7:** Memory status is not a deterministic attribute that can be used reliably in a network.

# pCR 23.700-60 v1.1.0 Evaluation and Conclusion

\*\*\*\* FIRST CHANGE (all new text) \*\*\*\*

## 7.9 KI#9: Evaluation

### 7.9.1 Solution #34 Evaluation

CDRX is under full responsibility of RAN2. This solution cannot be evaluated or concluded by SA2.

It can simply be observed this solution incurs multiple signalling handshakes between the UE and the network in order to constantly change the applied CDRX configuration, therefore causing additional UE power consumption and costing precious radio capacity. No evidence has been shown the solution provides gains over existing mechanism.

### 7.9.1 Solution #35 Evaluation

Solution #35 proposes SMF provides conditional QoS profile(s) to RAN, where each conditional QoS profile associates with an event. When RAN detects the event (possibly based on UE signalling) the associated QoS profile is applied. This solution relies on the AF or the RAN being able to associate UE power consumption or memory status to QoS profiles.

This solution claims: "it assumes AF can foresee the relationship of power consumption and time related events e.g. "time interval [start, end]", "duration of the service=x", or NG-RAN can get power consumption (e.g. device battery life) from UE, thus the conditional QoS profile(s) with event can be seen as the trade-off of throughput/ latency/ reliability and power consumption."

This solution does not explain how or why such time-related events are useful esp. for XR application – it does not explain what these events are based on what criteria they are defined.

This solution relies on reporting of UE memory status, without any explanation as to how this information can be used reliably by the network. Memory status is not a deterministic attribute that can be used reliably in the network: there is not a single memory in the device; a memory status is useless when not associated to a known memory capacity; memory status is useless if not knowing precisely how memory usage is managed by the device; and memory management is a highly optimized process in a device and very device-dependent.

This solution also incurs additional signalling from the UE to the RAN and related increase of signalling load and UE power consumption which is function of the event a) granularity b) frequency and c) hysteresis.

Solution #35, does not provide any evidence that the above events are necessary nor any explanation as to what trade-off is actually achieved.

### 7.9.1 Solution #61 Evaluation

Solution #61 proposes to report events pertaining to the Battery level (relative battery life) of a device as a trigger for the network to adjust its policy accordingly e.g. trigger QoS level adjustment.

The solution does not explain how policy adjustments are made depending on the reported battery level, with what granularity and based on what criteria and fails to explain how this in turn provides any benefit.

Battery level and relative battery life of a device are not deterministic quantities that can be used reliably in the network for policy adjustment. The device and the device alone knows what it should do to preserve battery life. An application in the UE can e.g. be throttled down when the device low power mode is activated, and/or the UE could decide to deactivate some radio capabilities.

Solution #61 also introduces additional signalling triggered by battery level changes, that besides increasing the signalling load in the network also costs additional power consumption in the UE without any evidence shown of any gains being achieved.

### 7.9.1 Solution #74 Evaluation

Solution #74 introduces a mechanism to assist PCF to adjust the policy based on UE or user preference, in the form of power mode that can be either power saving mode or ultimate performance mode.

The solution does not explain how a trade-off is achieved between QoE and power consumption, how policies are adjusted, with what granularity and importantly whether any gain is obtained at all, whilst it however requires additional signalling handshakes between the UE and the CN with the corresponding increase in power consumption and radio capacity usage.

"Power mode" is not a universally accepted definition but is instead very specific to a particular device (make and model). Low power mode in one device may have a very different meaning from low power mode in another device. Similar to battery level, low power mode is not a deterministic variable that can be used reliably by the network.

"Ultimate performance mode" is not a universal accepted definition notion, given performance can be associated to e.g. data rate, latency, reliability, power consumption and any combination thereof. In other words ultimate performance is indefinite and certainly does not necessarily imply highest power consumption.

### 7.9.1 Solution #75 Evaluation

Solution #75 proposes CN provides multiple QoS profiles to assist RAN select QoS parameters based on UE overheating status. RRC signalling is used to report UE overheating status to the RAN.

The solution does not explain how reporting the overheating status helps, what trade-off is achieved, how policies are adjusted, with what granularity and importantly whether any gain is obtained at all, whilst it requires additional signaling between the UE and the RAN with the corresponding increase in power consumption and radio capacity usage. This solution can be counterproductive when combined with other proposed solutions (some UEs do in fact overheat when being charged i.e. plugged in).

\*\*\*\* NEXT CHANGE (All new text) \*\*\*\*

# 8 Conclusions

Editor's note: This clause will list conclusions that have been agreed during the course of the study item activities.

## 8.9 KI#9: Conclusion

Solution #34, #35, #61, 74 and #75 do not proceed to normative work.

\*\*\*\* END OF CHANGES \*\*\*\*