RECAP 5GC: Resilience and CAP aware 5G Core for Consistent and High Availability Service

Siddhesh Pratim Sovitkar, Shwetha Vittal, Antony Franklin A
Department of Computer Science and Engineering, Indian Institute of Technology Hyderabad
Email: {cs21mtech12005, cs19resch01001, antony.franklin}@iith.ac.in

NOMS 2024-2024 IEEE Network Operations and Management Symposium | 979-8-3503-2793-9/24/$31.00 ©2024 IEEE | DOI: 10.1109/NOMS59830.2024.10575767

***Abstract*—Being scalable, distributed by nature, and stateless by design, the Service Based Architecture (SBA) of 5G Core (5GC) has become the popular approach today, to provide reliable services to 5G users. However, the well-known Consistency (C), Availability (A), and Partition (P) tolerance (CAP) theorem places a limit on 5GC, as it typically does for any distributed system to choose between consistency and availability, whenever a partition occurs in the system. Hence, choosing between ensuring consistency for users' state information and the availability of the service while being resilient against unforeseen sudden failures becomes a critical need for Beyond 5G (B5G) and 6G networks. In this paper, we present RECAP 5GC: a REsilience and CAP (RECAP)-aware 5GC that keeps up the High Availability (HA) service by being resilient and consistent during sudden failures of stateless functioning. To prove the feasibility and effectiveness of the proposed solution, we implement it entirely on the 3GPP-compliant control plane of the 5GC testbed between the Access and mobility Management Function (AMF) and Unstructured Data Storage Function (UDSF). Experiments on the minimal two nodes active/active cluster configuration prove that the proposed RECAP 5GC is not only resilient by dropping 47% fewer user requests than the CP mode and 86% fewer user requests than the AP mode but also 1.6% additionally available than the CP mode while remaining consistent on the user's state and information.**

I. Introduction

The current Service Based Architecture (SBA) of 5G Core (5GC) is distributed by nature with various Network Functions (NFs) split between the control plane and user plane separation framework. However, the overwhelming arrival of a variety of users (User Equipments (UEs)) connecting to the 5GC has urged the Mobile Network Operator (MNO) to dive deep into studying the key factors that challenge the robustness, scalability, consistency, resilience, and high availability of the service. In this regard, the distributed nature of 5GC with the stateless functioning of different NFs in it has aided to a large extent. While the distributed nature of 5GC with multiple NFs for their roles has assisted in independent scaling on the control plane and data plane, the stateless functioning of these NFs has intuitively helped in the scaling and High Availability (HA) [1, 2] at the individual NF level. For example, the scalable and stateless working of the Access and Mobility Management Function (AMF) consists of a load balancer, multiple serving instances, and an Unstructured Data Storage Function (UDSF) [3]. Here, the serving instances work in the stateless mode, by storing the users' information at the UDSF [1, 4]. However, as shown in the work [5], the very distributed nature of 5GC and the inter-NF dependencies with tight coupling between them can bottleneck the entire 5GC's service. Therefore, in this regard, we dive deep into evaluating the impacts of studying the Consistency, Availability, and Partition tolerance (CAP) theorem [6] on the stateless 5GC. While the CAP theorem (see Fig. 1) states that "A distributed system can deliver only two of the three desired characteristics: Consistency (C), Availability (A), and Partition tolerance (P) at any given time", the ultimate service delivery from such a system depends on the HA with consistency and resilience to the best possible level.

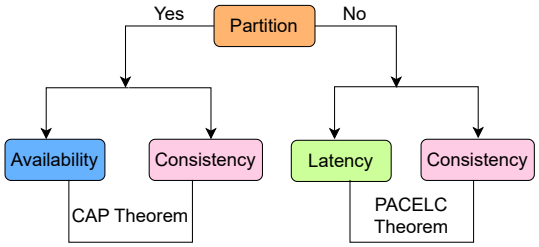


Fig. 1: View of CAP and PACELC Theorems.

In this work, we utilize our previous study of CAP theorem on the stateless 5GC [7] where we identified the possible impacts on the resilience and HA service of the 5GC, specifically, when the UDSF goes unreachable (a situation of partitioning). Specifically, we propose a REsilience and CAP (RECAP)-aware 5GC to prevent the identified impacts on the resilience and HA service in [7]. The key idea of RECAP 5GC is that it transitions from the stateless handling approach to the stateful handling until the UDSF becomes reachable again. Essentially, whenever the UDSF goes unreachable, all the available serving instances of the NF store and update the UE state information locally, and the load balancer too, transitions from the existing stateless load balancing scheme to the stateful load balancing scheme. So, unlike the single stateful instance approach, the RECAP 5GC does not sacrifice the availability of multiple serving instances. Hence, the RECAP 5GC is not only resilient enough to cater to the HA service but also can serve more UE registration and PDU session establishment requests for immediately accessing the data services thereafter. Therefore, to summarize, the key contributions of this work are:

- Design of RECAP 5GC as a CAP-aware solution to facilitate resilience, consistency, and HA while preserving the architectural and scalability needs.

Authorized licensed use limited to: Indian Institute of Technology Hyderabad. Downloaded on October 10,2026 at 15:38:54 UTC from IEEE Xplore. Restrictions apply.

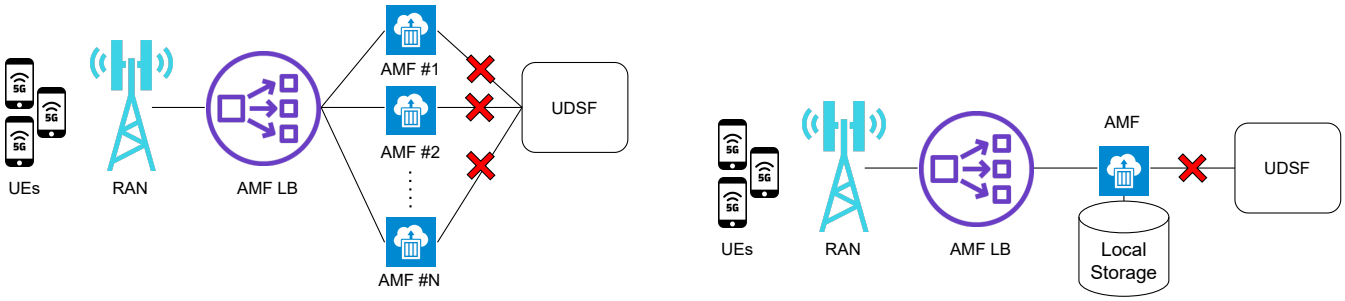


Fig. 2: View of AMF functioning with stateless AP mode in a) and stateful CP mode in b), both taken from [7].

- Evaluation of the proposed solution on the 3GPP-compliant 5G testbed to illustrate the HA, consistency, and resilience in serving the fresh user service requests.

The rest of the paper is organized as follows. First, Section II describes the required motivation and related work. Section III then proposes the RECAP 5GC to address the trade-off identified in [7]. Further, in Section IV we detail the implementation of RECAP 5GC which will be used to evaluate its performance for resilience and availability in Section V. Finally, in Section VI we conclude our paper with the future work. Note that, for further discussions in this paper, we consider the stateless 5GC architecture only and refer to it as 5GC for brevity.

II. Motivation

A. Overview of CAP and PACELC Theorem

CAP theorem [6] introduced in Section I has been known for the last two decades to help network operators make careful decisions on the working of any distributed systems. In addition to the CAP theorem, for completeness, we also consider the PACELC (Partition-tolerance, Availability, Consistency, Else Latency or Consistency) theorem, as it (PACELC theorem [8]) extends the CAP theorem by considering latency in addition to the consistency when there is no partition. Fig. 1 depicts the flow of CAP and PACELC theorems.

PACELC theorem states that "When there is Partition (P), one of Availability (A) and Consistency (C) is sacrificed (similar to CAP theorem), else (E), even when the system is running normally in the absence of partitions, one has to choose between Latency (L) and Consistency (C)." Evaluating this on the 5GC when there is no partition, we notice that the working of 5GC by default already falls under the consistency path of the PACELC theorem. This is because, the UDSF being the only external data store for all the UEs' states, mandates all the serving instances of NF (for example AMF) to fetch and update the UEs' states at the UDSF as per the standard procedures specified by 3GPP [9]. Therefore, in the rest of the paper, we focus on proposing RECAP 5GC, a CAP-aware solution to provide a consistent and highly available 5GC that is in turn resilient.

B. Results of Studying the CAP Theorem on the 5GC

Recently, we have studied the CAP theorem on the 5GC in our previous work [7]. As the current work is a continuation of this work [7], we briefly describe the highlights and results of studying the CAP theorem on the 5GC from [7]. First, we identified that the UDSF is an important NF, to cater to the consistent service as it stores all the users' state information who register and establish the PDU session. Hence, it should be reachable all the time for serving different types of user requests. Taking this observation, we injected the partition into this critical path, making UDSF unreachable, and performed the CAP theorem analysis on it. To check if consistency and availability can be maintained at the same time or not, we defined Availability with Partition tolerance (AP) mode as shown in Fig. 2a and Consistency with Partition tolerance (CP) mode as shown in Fig. 2b. For testing these setups described in work [7] and shown in Fig. 2, we continued to feed the UE registration and PDU session establishment requests, when the partition on UDSF was injected. The key observation was that 5GC in CP mode served more PDU session requests while remaining less available as compared to the 5GC in AP mode.

C. Related Work

Resilience and HA service are gaining quite an attention recently in the 5G research. In [5], authors show the impact on the procedure completion time, due to tight inter-NF dependency in the 5GC, for which the same authors have proposed a Per Procedure 5G system (PP5GS) [10]. Here, they have developed Per-Procedure NFs in the 5GC, such as an NF for serving UE registration, another NF for serving PDU session establishment, and so on. With this, the authors remove the dependency on multiple NFs for every UE procedure and illustrate resilience by remaining highly available to the best possible level. However, even in this work, UDSF is stateful and remains an external NF without getting absorbed into any of the PP5GS NFs. Recently in [11] authors highlight the significant need for handling short-time scale traffic variation in the 5GC and its impact on resilience. In [12], the authors perform an empirical study of the CAP theorem in a 4G mobility situation and confirm that the

Authorized licensed use limited to: Indian Institute of Technology Hyderabad. Downloaded on October 10,2026 at 15:38:54 UTC from IEEE Xplore. Restrictions apply.

incurred data unavailability (i.e., data service suspension) time is comparable to that induced by wireless connectivity setup. They also highlight that the desirable mobility control for the 5G networks should take a paradigm shift. In this regard, no work so far has specifically evaluated the CAP on 5GC. Hence, in our previous work [7], we studied the CAP theorem on the 5GC and identified the trade-off between 5GC in AP and CP modes of functioning when the partition occurs with UDSF going unreachable. Further, to address this trade-off, we propose RECAP 5GC in this work.

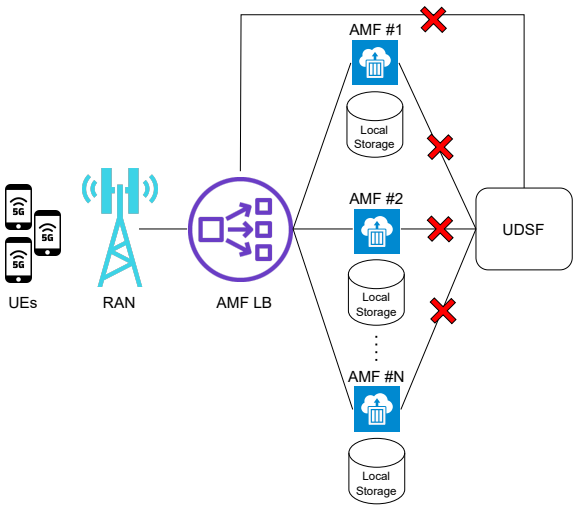


Fig. 3: Proposed framework of RECAP 5GC at AMF.

III. Proposed Resilience and CAP Aware 5G Core (RECAP 5GC)

The proposed RECAP 5GC aims for consistency and HA of service to the best of its ability, by being resilient and CAP-aware. Awareness of resilience [11] defines its potential to handle unexpected failures and recover from them to continue its operation. However, CAP awareness defines its ability to deal with availability and consistency trade-offs due to sudden partition happening on the critical path. Hence, in this context, resilience and CAP awareness connect and converge.

We propose RECAP 5GC making it proactive by design in the AMF, both in its serving instances as well as in its load balancer. Specifically, when the partition occurs on the identified critical path (with UDSF going unreachable), both, the load balancer and serving instances of AMF switch from stateless to stateful functioning to successfully handle the fresh user requests. Fresh users are those who are not registered yet and hence perform fresh registration. Further, they could request to establish the PDU session, to access the data services.

Fig. 3 shows the complete framework of the proposed RECAP 5GC for AMF, with a load balancer and serving instances. Additional storage is reserved here for each of the AMF instances to store the UE state information locally when the UDSF goes unreachable. Fig. 4 and Fig. 5 show

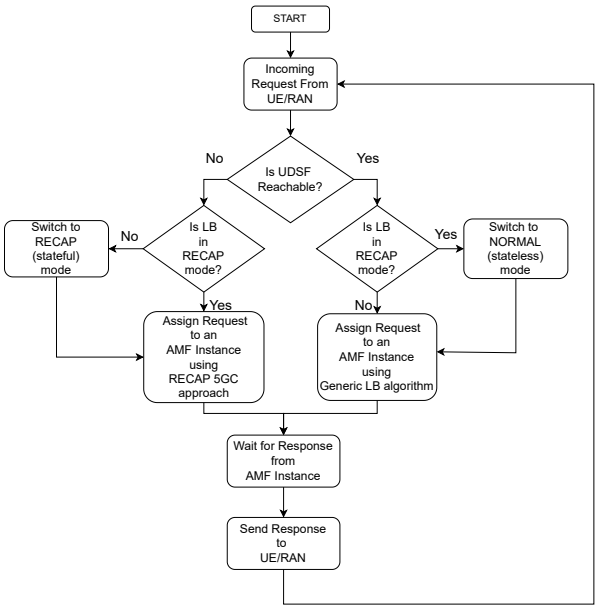


Fig. 4: Working of AMF's load balancer in the RECAP 5GC.

the working of RECAP 5GC at the AMF load balancer and each of the AMF instances respectively. When UDSF goes unreachable, the load balancer performs UE-specific load balancing. i.e., it first selects one of the AMF instances for scheduling the fresh UE registration request. Further, when the same UE requests to establish a PDU session, it forwards the PDU session request(s) of that UE to the same AMF instance that had handled the same UE's registration. More details on the implementation are given in Section IV. To align with this, each of the serving instances of AMF stores the UE state locally once they complete handling the UE's registration so that they can identify the UE when the same UE sends the PDU session request. The process continues, till the UDSF is reachable again, where each instance first updates the stored UEs' contexts to UDSF and then continues with its serving operation. This UE state update is synchronization-issue safe as each of the AMF instances would have handled the same UE's requests unlike in stateless mode.

In this way, when the partition occurs in the RECAP 5GC, both consistency and availability are not sacrificed to the best of their ability. i.e., availability is not sacrificed because we do not deactivate any of the available AMF instances. Consistency is also not sacrificed because each of these available instances can access the UE state from its local storage. Thus, the RECAP 5GC can continue to serve UE registration requests and allow them to establish the PDU session as well for access to data services immediately.

IV. Implementation of RECAP 5GC

A. Overview of In-house 5GC Prototype

RECAP 5GC is implemented on top of native 5GC. In this regard, we pick the in-house 5GC prototype that we used in our previous work too [7]. Similar to openly available 5G

Authorized licensed use limited to: Indian Institute of Technology Hyderabad. Downloaded on October 10,2026 at 15:38:54 UTC from IEEE Xplore. Restrictions apply.

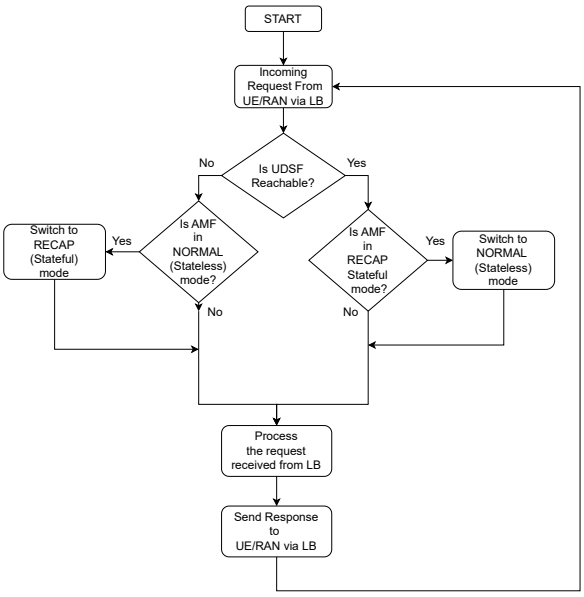


Fig. 5: Working of each of the serving instances of AMF in the RECAP 5GC.

cores like free5GC [13], OAI-5G Core [14], the in-house 5GC prototype is 3GPP-compliant including AMF along with other control plane NFs, and a User Plane Function (UPF) for the data plane services. For service-based interactions between the control plane NFs, nghttp2 [15] has been leveraged to support HTTP/2. In addition to the 5GC, a RAN+UE emulator is also developed as part of this prototype, for end-to-end testing. The emulator has Non Access Stratum (NAS) and Next Generation Application Protocol (NGAP) functions without a radio stack. It emulates the UE behavior for UE registration and PDU session establishment procedures as per NAS protocol and is provisioned by RAN using NGAP messages. While all the NFs of 5GC including the RAN+UE emulator can be run either on docker containers or Virtual Machines (VMs), in this work, we have deployed them as docker containers as they are the popular approach today supporting cloud-native deployments. So, on this native 5GC prototype, we add the functionality required for RECAP 5GC at the load balancer and serving instance of AMF.

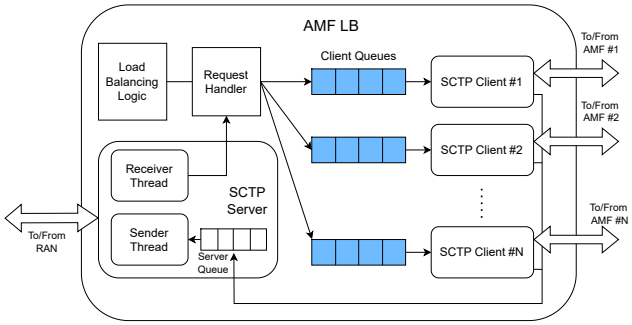


Fig. 6: Internal view of AMF load balancer in the user space.

B. AMF Load balancer

To show the effectiveness of the proposed RECAP 5GC, we have developed an application-based AMF load balancer that runs in the user space. As per 3GPP [16], AMF has NGAP and NAS for all its interactions towards RAN and UE respectively. NGAP works on SCTP. Since the AMF load balancer is internal to the design of AMF, and RAN is not aware of it, the AMF load balancer has to interact with the RAN as if it is an AMF by itself. Therefore, in the proposed implementation, the AMF load balancer has $N+1$ interfaces, where N represents the number of AMF instances.

Fig. 6 shows the internal design of the AMF load balancer that is implemented here. One interface acts as an SCTP server towards the RAN, and the remaining N interfaces act as SCTP clients towards the respective serving instances of AMF. On either side, the load balancer has a standard SCTP socket interface from the *libsctp* library. The available AMF instances' information like their IP and port are configured at the load balancer. So, upon start, the load balancer, sets up SCTP connections followed by NG Setup [16] to each of the available AMF instances, so that these instances are ready to accept the UE-associated NGAP messages further on their respective SCTP server sockets.

When a fresh UE registration request arrives, the load balancer chooses one of the AMF instances to serve it. To perform the stateful (i.e., UE aware) load balancing, the load balancer maintains a mapping of the AMF UE NGAP ID field fetched from the NGAP portion of the response for the UE registration request received from the respective chosen AMF instance. Further, when the same UE sends other messages like PDU session establishment, the load balancer forwards these UE-associated NGAP messages to the previously chosen AMF serving instance identified with the help of the IP of the AMF Instance and AMF UE NGAP ID mapping, by pushing the message to the sender queue of the SCTP client serving the respective AMF instance. Similarly, on the downlink, the respective SCTP client in the load balancer pushes the response received from the AMF instance to the SCTP server's sender queue (see Fig. 6) to forward it to the RAN. Note that, the load balancer also keeps a TCP-based heartbeat with UDSF for it to be aware if the UDSF is reachable or not. This heartbeat works continuously on a periodicity basis, to help decide, switching between stateless to stateful functioning and vice-versa when UDSF goes unreachable or back reachable again.

C. AMF Serving Instance

AMF serving instance is nothing but the standard AMF having all the functionalities to serve various UE procedures at the entry of 5GC. For RECAP 5GC, we add the required changes to it, as per the description in Section III. Additionally, similar to the load balancer, the AMF's serving instance in the RECAP 5GC also has a TCP-based heartbeat with UDSF. Accordingly, when UDSF is detected unreachable by heartbeat, the instance switches its handling mode of user requests from stateless to stateful, until the UDSF is reachable

Authorized licensed use limited to: Indian Institute of Technology Hyderabad. Downloaded on October 10,2026 at 15:38:54 UTC from IEEE Xplore. Restrictions apply.

again. i.e., upon switching this way, the instance starts storing the UE state and related information locally, so that it can continue to serve the same UEs whose registration requests it served earlier. Note that, similar to the load balancer the heartbeat to UDSF functions continuously to track the reachability of UDSF (see Fig. 5). Note that, we set the heartbeat frequency to 2 seconds both in the load balancer and serving instance. This value can be set as desired by the MNO by assessing the situations of UDSF going unreachable and deployment conditions.

V. Performance Evaluation

A. Experimental Setup

To evaluate the performance of the proposed RECAP 5GC, we set up the end-to-end test framework as shown in Fig. 3. Here, we evaluate resilience, consistency, and HA during partitioning. For comparison, we pick the 5GC in the AP setup (shown in Fig. 2a), and 5GC in the CP setup (shown in Fig 2b). Additionally, we also consider PP5GS [10] for comparison, as this work resolves the inter-NF dependency issues (see Section II-C) in the 5GC but keeping UDSF outside. So, the partitioning situation of the CAP theorem with UDSF going unreachable applies to PP5GS too.

TABLE I: Test setup parameters for performance evaluation.

Parameter	Value
Number of PDU session establishment requests	Uniform distribution with numbers from 50 to 200 requests/second. Note that the reference value for this number of requests is taken from works in [10] and [17].
PDU session establishment processing capacity per AMF instance	50 requests/second (see Section V-A for description).
Memory required per UE registration	58 KB
Memory required per UE PDU session establishment	2 KB
Periodicity of heartbeat to the UDSF	2 seconds

For hosting the end-to-end testing, we pick an Intel® Xeon® CPU E5-1650 v4 @ 3.60GHz system. Here, we bring up the 5GC first with required NFs, including AMF with 2 instances (for RECAP 5GC and 5GC in AP mode), and the rest of the needed NFs of the 5GC. Then we start the AMF load balancer, followed by the RAN+UE emulator in all the modes.

Upon start, the RAN+UE emulator feeds a different number of UEs which register and try to establish the PDU session in sequence to the 5GC. Note that with a single thread running in every AMF instance, the average time taken to complete a single PDU session establishment procedure in this 5GC testbed is '20 ms'. Using this '20 ms' value, we considered the ground truth capacity of processing 50 such requests per second for every AMF instance. Specific parameters considered for testing are tabulated in Table I. During this testing, we inject a partition into the test setup, by shutting down the AMF-UDSF momentarily. So now for 5GC in CP mode, there is only a single AMF instance present as per Fig. 2b. Whereas, in the RECAP-5GC, the AMF load balancer and the serving AMF instances immediately switch their functioning from stateless mode to stateful mode as the heartbeat to UDSF fails on TCP. At this point, we measure the different metrics from individual setups.

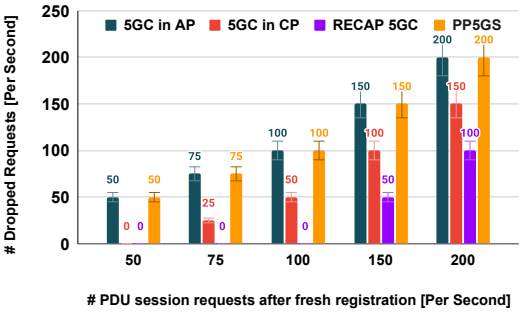


Fig. 7: Comparison of number of dropped PDU session requests.

B. Evaluating Resilience with Consistency

As this work is the continuation of our work in [7], once again we capture the 5GC's resilience by measuring the number of impacted UEs (users). Fig. 7 compares the number of PDU session establishment requests dropped by various modes. We can observe that 5GC in AP mode and PP5GS drop all these requests, as none of the AMF instances nor procedure-specific NFs in the PP5GS can access the UE state, due to UDSF's unreachability. On the other hand, as observed in [7], the 5GC in CP mode drops only the excess requests beyond its capacity, or in some situations, it could serve more, but delaying their procedure completion time. However, the proposed RECAP 5GC can serve all of these PDU session requests when they arrive within its capacity of processing them, as the AMF load balancer forwards this request of the UE to the same AMF instance that had served its registration (which can serve them successfully by accessing the respective UE state from the local data store). Note that, when requests arrive more in number, RECAP 5GC also drops excess requests. This essentially is a scaling criterion and is beyond the scope of this paper.

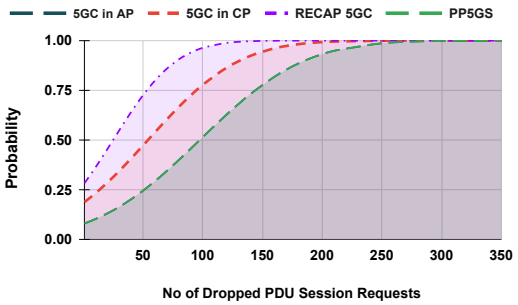


Fig. 8: CDF of the number of dropped PDU session requests.

Further, based on results shown in Fig. 7, Fig. 8 shows the Cumulative Distribution Function (CDF) of the number of

Authorized licensed use limited to: Indian Institute of Technology Hyderabad. Downloaded on October 10,2026 at 15:38:54 UTC from IEEE Xplore. Restrictions apply.

dropped PDU session requests among all the modes. Here, we confirm that RECAP 5GC shows a higher probability of dropping less number of such requests followed by 5GC in CP, 5GC in AP mode, and PP5GS. Note that 5GC in AP and PP5GS overlap here in CDF, as both these modes drop the same number of requests with the same number of serving instances (2) being available, though they differ in their implementation (while PP5GS has separate NFs for handling UE registration and PDU session procedures, the 5GC in AP is not procedure specific).

C. Evaluating High Availability

Once again, for calculating the total availability of the AMF, we do it as done in our previous work [7] which compared 5GC in AP and 5GC in CP modes, while studying the CAP theorem. Availability of individual serving instances of AMF is given by

$$A_{AMFi} = \frac{Uptime\ of\ AMFi}{Uptime\ of\ AMFi + Downtime\ of\ AMFi} \quad (1)$$

The total availability of AMF without AMF-UDSF is given by

$$A_{AMF} = A_{AMF-LB} * (1 - (1 - A_{AMFi})^n)) \quad (2)$$

where 'n' is the total number of serving AMF instances.

TABLE II: High availability of AMF

Availability of 5GC in AP and 5GC based on PP5GS	Availability of 5GC in CP	Availability of RECAP 5GC
99.972%	98.33%	99.972%

As it takes 5 seconds for an AMF instance to come up, during the total time of 5 minutes of testing, we restarted the AMF instance once, resulting in 98.33% availability for it, as per (1). Further, using (2)), Table II lists the total availability calculated for each mode. Here, we can realize that with the presence of 2 instances of AMF, the proposed RECAP 5GC, 5GC in AP mode, and PP5GS prove higher availability as compared to the 5GC in CP mode which sacrifices the availability with only a single alive instance of AMF.

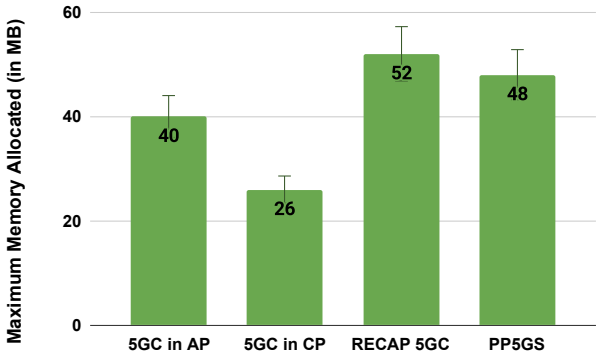


Fig. 9: Comparison of memory requirement.

D. Evaluating Memory Requirement

Fig. 9 compares the total memory requirement among different modes, for handling 100 UEs in the AMF instances. A ground truth value of 60 KB is the memory requirement for serving a single UE registration and PDU session establishment successfully (see Table I) on this testbed. 16 MB is the default memory requirement for running a single AMF instance. About 4 MB memory is required to serve a total of 50 UEs per instance working in stateless mode. While the 5GC in CP has the least memory of 26 MB (as it has a single AMF instance working in the stateful mode for consistency), the 5GC in AP has 40 MB in total for 2 AMF instances in the setup, but working in stateless mode. On the other hand, the PP5GS demands about 48 MB[1] in total as it builds a procedure-specific NF for UE registration and PDU session establishment. But, with a total of 52 MB (for 2 AMF instances in the setup), we confirm that RECAP 5GC demands additional local memory at every available AMF instance to store the UEs' information whenever the UDSF is unreachable.

On a final note, for the study of the CAP theorem, we have not considered the other NFs' unreachability but only the UDSF, as it falls on the critical path of handling UE procedures in the currently happening stateless 5GC. Though, PP5GS [10] solves the inter-NF dependencies for other NFs by absorbing the NFs in a single procedure-specific NF, UDSF, UDM, and NRF remain stateful and hence external NFs, even there. So, overall, we have addressed the CAP theorem for UDSF unreachability, but NRF and UDM are left out. For addressing the unexpected unreachability of NRF, the required NFs can be discovered via NRF in the beginning itself, before the depending NFs start serving the UE requests. However, when UDM is unreachable, UE registration will fail in the first place even in the RECAP 5GC. We plan to revisit this as part of our future work.

VI. Conclusion and Future Work

Studying the CAP theorem on the stateless 5GC is a critical need to evaluate the HA and resilience impacts in Beyond 5G (B5G) and 6G. This paper presented RECAP 5GC to address the trade-off observed from the study of the CAP theorem in [7]. Our experiments on the 3GPP-compliant 5GC testbed show that the proposed RECAP 5GC not only addresses the trade-off issues between AP and CP modes, but also serves the maximum number of fresh users arriving during such a situation, thereby catering to the consistency, resilience, and HA service overall. In the future, we plan to study the CAP theorem challenges in the inter-slice handovers and possibly extend the proposed RECAP 5GC there.

Acknowledgment

This work was supported by the Department Of Science & Technology, India as part of the "Next Generation Wireless Research and Standardization on 5G and Beyond" project.

[1]As PP5GS code was not openly available when this work was done, we had to extrapolate this number from our testbed used in this experiment.

Authorized licensed use limited to: Indian Institute of Technology Hyderabad. Downloaded on October 10,2026 at 15:38:54 UTC from IEEE Xplore. Restrictions apply.

REFERENCES

[1] Shwetha Vittal and A. Antony Franklin, "HARNESS: High Availability Supportive Self Reliant Network Slicing in 5G Networks", *IEEE Transactions on Network and Service Management*, vol. 19, no. 3, pp. 1951–1964, 2022.

[2] ETSI, "Network Functions Virtualisation (NFV) Release 3;Reliability;Report on NFV Resiliency for the Support of Network Slicing", Tech. Rep. ETSI GR NFV-REL 010 V3.1.1, 2019.

[3] 3GPP, "Unstructured Data Storage Services", Tech. Rep. TS 29.598, 2023.

[4] Sourav Sarkar, Shwetha Vittal, and Antony Franklin A, "Locomotive 5g core for 6g ready resilient and highly available network slices and sfcs", in *2022 18th International Conference on Network and Service Management (CNSM)*, 2022, pp. 367–373.

[5] Endri Goshi, Michael Jarschel, Rastin Pries, Mu He, and Wolfgang Kellerer, "Investigating inter-nf dependencies in cloud-native 5g core networks", in *2021 17th International Conference on Network and Service Management (CNSM)*, 2021, pp. 370–374.

[6] Seth Gilbert and Nancy Lynch, "Brewer's conjecture and the feasibility of consistent, available, partition-tolerant web services", *SIGACT News*, vol. 33, no. 2, pp. 51–59, jun 2002.

[7] Siddhesh Pratim Sovitkar, Shwetha Vittal, and Antony Franklin A, "Evaluating CAP Theorem on a Stateless 5G Core", in *(press) 2024 IEEE 21st Consumer Communications & Networking Conference (CCNC)*, 2024, pp. 876–879.

[8] Daniel Abadi, "Consistency tradeoffs in modern distributed database system design: Cap is only part of the story", *Computer*, vol. 45, no. 2, pp. 37–42, 2012.

[9] 3GPP, "Procedures for the 5G System", Tech. Rep. TS 23.502.

[10] Endri Goshi, Raffael Stahl, Hasanin Harkous, Mu He, Rastin Pries, and Wolfgang Kellerer, "PP5GS—An Efficient Procedure-Based and Stateless Architecture for Next-Generation Core Networks", *IEEE Transactions on Network and Service Management*, vol. 20, no. 3, pp. 3318–3333, 2023.

[11] Rui Li, Bertrand Decocq, Anne Barros, Yi-Ping Fang, and Zhiguo Zeng, "Estimating 5g network service resilience against short timescale traffic variation", *IEEE Transactions on Network and Service Management*, pp. 1–1, 2023.

[12] Yuanjie Li, Zengwen Yuan, Chunyi Peng, and Songwu Lu, "CAP on Mobility Control for 4G LTE Networks", in *Proceedings of the 3rd Workshop on Hot Topics in Wireless*, New York, NY, USA, 2016, HotWireless '16, p. 33–37, Association for Computing Machinery.

[13] "free5GC", https://github.com/LABORA-INF-UFG/NetSoft2020-Tutorial4-Demo2-Exp2/tree/master/free5gc.

[14] "OpenAirInterface 5G core network development", https://gitlab.eurecom.fr/oai/cn5g.

[15] Nghttp2, "Nghttp2: HTTP/2 C Library", https://nghttp2.org/, 2021.

[16] 3GPP, "NG-RAN; NG Application Protocol (NGAP)", Tech. Rep. TS 38.413, 2019.

[17] I. Alawe, A. Ksentini, Y. Hadjadj-Aoul, and P. Bertin, "Improving Traffic Forecasting for 5G Core Network Scalability: A Machine Learning Approach", *IEEE Network*, vol. 32, no. 6, pp. 42–49, 2018.

Authorized licensed use limited to: Indian Institute of Technology Hyderabad. Downloaded on October 10,2026 at 15:38:54 UTC from IEEE Xplore. Restrictions apply.