|  |  |
| --- | --- |
| 3GPP TR 38.879 V18.0.0 (2022-09) | |
| Technical Report | |
| 3rd Generation Partnership Project;  Technical Specification Group Radio Access Network;  NG-RAN;  Study on enhancement for Resiliency of gNB-CU-CP  (Release 18) | |
|  | |
|  |  |
|  | |
| The present document has been developed within the 3rd Generation Partnership Project (3GPP TM) and may be further elaborated for the purposes of 3GPP. The present document has not been subject to any approval process by the 3GPPOrganizational Partners and shall not be implemented. This Specification is provided for future development work within 3GPPonly. The Organizational Partners accept no liability for any use of this Specification. Specifications and Reports for implementation of the 3GPP TM system should be obtained via the 3GPP Organizational Partners' Publications Offices. | |

|  |
| --- |
|  |
| ***3GPP***  Postal address  3GPP support office address  650 Route des Lucioles - Sophia Antipolis  Valbonne - FRANCE  Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16  Internet  http://www.3gpp.org |
| ***Copyright Notification***  No part may be reproduced except as authorized by written permission. The copyright and the foregoing restriction extend to reproduction in all media.  © 2022, 3GPP Organizational Partners (ARIB, ATIS, CCSA, ETSI, TSDSI, TTA, TTC).  All rights reserved.  UMTS™ is a Trade Mark of ETSI registered for the benefit of its members  3GPP™ is a Trade Mark of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners LTE™ is a Trade Mark of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners  GSM® and the GSM logo are registered and owned by the GSM Association |

Contents

Foreword 4

1 Scope 5

2 References 5

3 Definitions of terms, symbols and abbreviations 5

3.1 Terms 5

3.2 Symbols 5

3.3 Abbreviations 5

4 General 5

5 Study failure scenarios associated with the gNB-CU-CP 7

6 Conclusion 7

Annex <A> (informative): Failure Scenario from each contribution’s view 8

A.1 Scenario and issue description (R3-224787) 8

A.1.1 CU SW failure 8

A.1.2 gNB-CU-CP HW failure 8

A.1.3 Power outages 8

A.1.4 Transport network failure (out of 3GPP scope) 8

A.1.5 Summary of failure scenarios 9

A.2 Scenario (R3-224281) 9

A.2.1 Scenarios A, B, C 9

A.2.2 Scenario D 10

A.2.3 Scenario E 10

A.3 Scenario and issue description (R3-224303) 11

A.4 Scenario and issue description (R3-224324) 11

A.5 Failure Case (R3-224576) 12

A.6 Scenario and issue description (R3-224627) 12

A.6.1 Scenario #1: Geographical Redundancy 12

A.6.1.1 General description 12

A.7 Scenario and issue description (R3-224754) 13

A.7.1 Local Redundancy 13

A.7.2 Failure Scenarios 13

A.7.2.1 Server Failure 13

A.7.2.2 VNFC (Virtual Network Function Component) Failure 13

A.7.2.3 Transport Link Failure (out of 3GPP scope) 13

A.7.3 Failure Impact 13

A.8 Scenario and issue description (R3-224898) 14

A.8.1 Possible failure scenarios 14

Annex <B> (informative): Change history 14

# Foreword

This Technical Report has been produced by the 3rd Generation Partnership Project (3GPP).

The contents of the present document are subject to continuing work within the TSG and may change following formal TSG approval. Should the TSG modify the contents of the present document, it will be re-released by the TSG with an identifying change of release date and an increase in version number as follows:

Version x.y.z

where:

x the first digit:

1 presented to TSG for information;

2 presented to TSG for approval;

3 or greater indicates TSG approved document under change control.

y the second digit is incremented for all changes of substance, i.e. technical enhancements, corrections, updates, etc.

z the third digit is incremented when editorial only changes have been incorporated in the document.

# 1 Scope

The present document provides descriptions of failure scenarios associated with the gNB-CU-CP based on current architecture for the NG-RAN.

# 2 References

The following documents contain provisions which, through reference in this text, constitute provisions of the present document.

- References are either specific (identified by date of publication, edition number, version number, etc.) or non‑specific.

- For a specific reference, subsequent revisions do not apply.

- For a non-specific reference, the latest version applies. In the case of a reference to a 3GPP document (including a GSM document), a non-specific reference implicitly refers to the latest version of that document *in the same Release as the present document*.

[1] 3GPP TR 21.905: "Vocabulary for 3GPP Specifications".

[2] 3GPP TS 38.401: "NG-RAN;Architecture description".

[3] 3GPP TS 23.501: "System architecture for the 5G System (5GS)".

[4] RP-221438: “Revised SID: Study on enhancement for resiliency on gNB-CU-CP”

# 3 Definitions of terms, symbols and abbreviations

## 3.1 Terms

For the purposes of the present document, the terms given in TR 21.905 [1] and the following apply. A term defined in the present document takes precedence over the definition of the same term, if any, in TR 21.905 [1].

## 3.2 Symbols

For the purposes of the present document, the following symbols apply:

<symbol> <Explanation>

## 3.3 Abbreviations

For the purposes of the present document, the abbreviations given in TR 21.905 [1] and the following apply. An abbreviation defined in the present document takes precedence over the definition of the same abbreviation, if any, in TR 21.905 [1].

<ABBREVIATION> <Expansion>

# 4 General

The 3GPP split architecture for NG-RAN is described in TS 38.401 [2]. The NR gNB disaggregated architecture is characterized by the presence of a single gNB-CU-CP, which is in control of one or more gNB-DU(s) and one or more gNB-CU-UP(s) (Figure 4-1). Similarly, a gNB-DU may interact with multiple gNB-CU-UPs simultaneously for the same user context as long as they are controlled by the same gNB-CU-CP.



Figure 4-1: Overall architecture for separation of gNB-CU-CP and gNB-CU-UP TS 38.401 [2]

The gNB-CU-CP may have specific configured data for each supported cell in the gNB, and a common gNB ID per PLMN. This identity is known in the gNB-DU and in neighbor gNBs, and it is also broadcasted in the served NR cells as part of the cell identity.

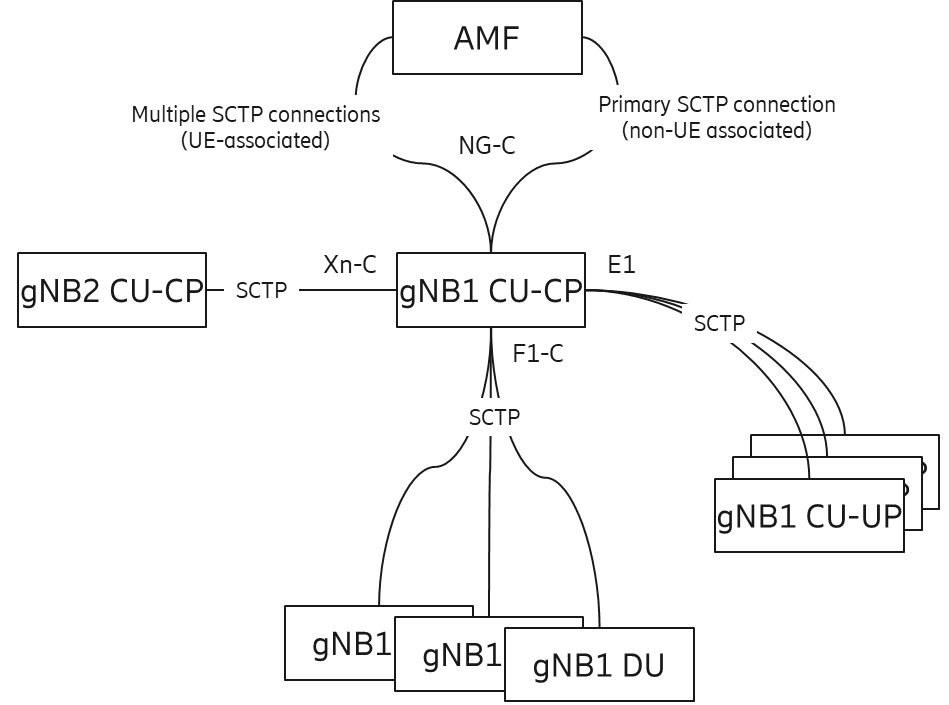


Figure 4-2: CP connections involving the AMF, gNB-CU-CP, gNB-CU-UP, and gNB-DUs.

There is one (active) CP anchor node per UE in the 5GC (AMF) and one (active) CP anchor per gNB in the NG-RAN (gNB-CU-CP). Figure 4-2 highlights the various CP connections going through the same gNB-CU-CP towards the AMF, to neighbor gNB-CU-CPs, to gNB-CU-UPs, and to gNB-DUs. All CP connections terminate in the gNB-CU-CP; if we only look at the logical architecture, the gNB-CU-CP may indeed be considered as a single point of failure.

For Operator, resiliency of public network is under national regulations or other operational constraints.

With regard to resiliency TS 38.401[2] includes only the following note:

NOTE: For resiliency, a gNB-DU may be connected to multiple gNB-CUs by appropriate implementation.

Furthermore, multiple TNL associations toward the gNB-CU-CP are currently supported by the standard.

In the event of a failure at gNB-CU-CP, the likely outcome is that all user contexts could be affected and experience user service interruption.

This study handles failure scenarios for gNB-CU-CP, based on the current NG-RAN architecture.

Solutions for failure scenarios will not be addressed in this study.

The nodes shown in Figure 4-1 are logical nodes, i.e., they can be implemented as physical network functions (PNFs) using dedicated hardware (HW) infrastructures or as virtual network functions (VNFs) running as software (SW) functions (e.g., virtual machines (VMs) or cloud-native containers (CNFs)) on general purpose processors (GPPs; with or without HW acceleration support), e.g., in a cloud environment of a data center.

# 5 Study failure scenarios associated with the gNB-CU-CP

Various perspectives were considered by each company (see the informative Annex A for details).

A failure scenario for a deployment case with gNB-CU-CP at the antenna site (Distributed RAN/D-RAN) should not be of relevance for this study as an outage would impact the network performance only in a limited local area (similar to an outage of the co-located gNB-DU). Local resiliency measures may therefore be realized in the same way for both logical nodes (dependent on PNF/VNF implementation).

A more important scenario is the case when a gNB-CU-CP is deployed in a central office location of an operator’s network. Those central locations may have transport network (TN) connections (e.g., via tree- or ring-based fibers) to several gNB-DUs located at different antenna sites. In such deployment case the gNB-CU-CP has the responsibility for proper operation of a large number of cells covering a wider regional area, i.e., an outage has a strong impact on overall user experience in a larger part of the mobile network. Each gNB-CU-CP may be connected via multiple TNL associations. As implicitly stated by the note in TS 38.401[2] gNB-DUs at antenna sites may be connected via the same or different TN links to one or several central locations (geo-redundancy) hosting gNB-CU-CP instances.

The following failure scenarios have been discussed:

- Hardware failure

- Software failure

- Transport network failure

- Power failure

- Entire gNB-CU-CP failure, e.g., due to natural disaster

It was confirmed that some failures in the gNB-CU-CP may be addressed by using virtualization technology.

On the other hand, there is a scenario with possibly an entire gNB-CU-CP failure (e.g. due to natural disasters). Resiliency mechanisms to recover from such failure may be achieved via gNB-DU’s connection to multiple gNB-CUs by appropriate implementation, as captured in TS 38.401[2].

# 6 Conclusion

In this TR, failure scenario in the existing gNB-CU-CP architecture from various perspectives were studied.

A failure affecting the entire (logical) gNB-CU-CP would impact all UEs served by the cells of the gNB-DUs connected to the gNB-CU-CP, resulting in those UEs being out of service for potentially long periods of time (dependent on type of failure and available resiliency measures). Current specification allows for some resiliency mechanisms via appropriate implementation (as described in TS 38.401[2]) which can address a number of the failure scenarios.

# Annex <A> (informative): Failure Scenario from each contribution’s view

## A.1 Scenario and issue description (R3-224787)

### A.1.1 CU SW failure

**Failure reason**

gNB-CU-CP SW failure caused by software in gNB-CU-CP's equipment. This is mainly caused by information inconsistencies in the software, which in most cases can be resolved by restarting the system (or, in case of a virtualized implementation, by respawning the relevant software parts).

**Failure impact**

The impact of the failure and the extent of the restart depends on the configuration of the equipment and where the problem occurs. In some cases, the effect may be limited to only some functions in the gNB-CU-CP, while in other cases, all functions in the gNB-CU-CP may be affected.

### A.1.2 gNB-CU-CP HW failure

**Failure reason**

gNB-CU-CP HW Failure is a problem caused by the hardware of the gNB-CU-CP, mainly due to aging and deterioration of the HW of the gNB-CU-CP, and in most cases can be resolved by replacing some of the HW in the equipment.

**Failure impact**

The impact of the failure and the extent of the replacement depends on the configuration of the equipment and the part of the problem. In some cases, only some functions of the gNB-CU-CP will be affected, but it is expected that the whole gNB-CU-CP will be affected, especially in the case of vRAN scenarios.

### A.1.3 Power outages

**Failure reason**

Power outages are caused by power grid failures due to various disasters (earthquakes, lightning, tsunamis, fires, windstorms, snowstorms, etc.), as well as by malfunctions of substation equipment or cable breaks in the station building.

**Failure impact**

The impact of a failure depends on the part of the system where the failure occurs, but often affects multiple gNB-CU-CPs. The higher degree of centralization, the more gNB-CU-CPs will be affected.

### A.1.4 Transport network failure (out of 3GPP scope)

**Failure reason**

Transport network failures result from cable breaks in transmission lines, network equipment failures, and　misconfigurations. These can also occur as software or hardware malfunctions, disasters, or human error.

**Failure impact**

The impact of a failure depends on the part of the system where the failure occurs. When the failure occurs outside line, it often affects multiple gNB-CU-CPs. The higher degree of centralization, the more gNB-CU-CPs will be affected.

### A.1.5 Summary of failure scenarios

Failures affecting only one gNB-CU-CP can be addressed by existing countermeasures.

For failures affecting multiple gNB-CU-CPs simultaneously, in particular when the cause of the failure is closely related to the region, such as a disaster, a solution that takes advantage of the different locations for deploying redundant infrastructure should be considered.

## A.2 Scenario (R3-224281)

The most relevant scenarios in regard to gNB-CU-CP are those that will incur a failure of the entire gNB-CU-CP. Such cases would affect all the existing UE contexts under the gNB-CU-CP where the failure occurred. Likewise, the disaggregated gNB architecture allows for very large configurations. A given gNB-CU-CP can host 512 gNB-DUs, and a total of 16,384 cells with existing specifications. Furthermore, each cell may be serving hundreds of UEs Therefore, a gNB-CU-CP failure can incur a very high impact in service availability and user experience for a very large number of UEs. Further, these failure scenarios can incur further problems by generating very high signalling loads from e.g., re-establishment of connections and signalling interfaces.

In contrast issues affecting only limited portions of the gNB-CU-CP, for example, trouble incurred from individual hardware blades in a virtualized environment, are not the focus of this study.

### A.2.1 Scenarios A, B, C

**(A) Node breakdown leading to a gNB-CU-CP failure**

This failure scenario includes cases where the gNB-CU-CP becomes completely unresponsive. This could be e.g., due to hardware failure, or power source becoming lost and unrecoverable.

**(B) Natural disaster leading to loss of the gNB-CU-CP**

This failure scenario includes cases where the gNB-CU-CP becomes unrecoverable due to the node itself becoming destroyed or its required connectivity, hence becoming unrecoverable. This could be result of e.g., an earthquake, tsunami, or major fire.

**(C) Human-made disaster leading to loss of the gNB-CU-CP**

This failure scenario is similar to that caused by natural means but with the source of the failure being due to human involvement. This failure could be result e.g., of war, civil war, terrorism or social unrest.



Figure A.2.1-1: Example of gNB-CU-CP Failure Scenarios (A)(B)(C)

### A.2.2 Scenario D

**(D) Signalling Interface Link Failure**

This failure includes cases where the control plane signalling link becomes unavailable. The issue may be temporary (e.g., intermittent issues at a switch/router at a given communication path), or (semi-)permanent.



Figure A.2.2-1: Example of gNB-CU-CP Failure Scenario (D)

### A.2.3 Scenario E

**(E) Planned Maintenance Causing Downtime**

This use case corresponds to maintenance events (e.g., Software upgrade, new feature activations, public protests or temporary unrest). Software upgrade may result in a full reboot of the gNB-CU-CP as well.



Figure A.2.3-1: Example of gNB-CU-CP Failure Scenario (E)

## A.3 Scenario and issue description (R3-224303)

In the legacy network architecture, a gNB-CU-DU is only connected to one gNB-CU-CP. With the increasing number of connected gNB-CU-UP and gNB-DU, gNB-CU-CP will be in the risk of failure caused by the large amount of signalling processing. So this SI will introduce a mechanism to allow the gNB-CU-UP and gNB-DU to recovery service. Higher-layer split between gNB-CU and gNB-DU would enable highly-centralized gNB-CU deployment with large coverage area per gNB-CU, especially from C-Plane perspective. Likewise, Higher-layer split between gNB-CU-CP and gNB-CU-UP would enable highly-centralized gNB-CU-CP deployment with large coverage area per gNB-CU-CP.

As described in the draft SID, such a centralized gNB-CU(-CP) would be a single point of failure. Hence the resiliency of the gNB-CU(-CP) is highly important. Given the limited time allocated for gNB-CU resiliency, only gNB-CU-CP resiliency should be considered in this release, other network nod, e.g. AMF, gNB CU(non-split CU architecture) should not be considered in this release if time not allowed.

Failure condition:

Failure condition is the trigger of gNB-CU-CP resiliency. If a network node is considered a failure when it is crashed completely, then it would be too late to recovery. The network should have scheme to switch part or all of the services/connections to another resilient network node to offload before it completely crashes down. The decision of offloading the service/connection can be either of the below:

1) Up to the implantation of gNB-CU-CP

2) The configured threshold from OAM

Both options work and has no impact to RAN3. So the failure condition can be left to network implementation.

In addition, if any of the adjacent nodes of gNB-CU-CP, e.g. AMF, gNB-CU-CP and gNB-DU lost the connection to gNB-CU-CP, due to an unexpected gNB-CU-CP failure, adjacent node should be able to activate the resilient gNB-CU-CP automatically.

## A.4 Scenario and issue description (R3-224324)

The most critical cases are failures that may happen if a gNB-CU-CP is deployed in a central office location with responsibility for a high number of gNB-DUs and therefore of cells. There are different failure scenarios that may happen in such a deployment scenario:

(1) SW or HW failure in a PNF-based implementation;

(2) VNF (SW) or GPP (HW) failure in a virtualized environment (e.g. cloud-based implementation);

(3) Failures of NW connections, e.g., TN or server NW card failures;

(4) Total failure of central location, e.g., caused by a power shutdown or a disaster case.

For failure scenarios (1) and (2) the gNB-CU(-CP) functionality can be recovered by taking local redundancy in the same central location into account (e.g., based on usage of spare HW (PNF/GPP) for redundancy purposes). A slow recovery can be initiated via OAM orchestrating a new gNB-CU(-CP) instance on the spare HW (e.g., in combination with cloud-based mechanisms in a virtualized environment). A fast recovery approach to bring the outage time down would require active/stand-by operation of a mirror gNB-CU-CP instance in parallel which may be realized in a local environment by appropriate implementation (incl. also cloud-based mechanism in case of virtualization).

Slow recovery may also work in case of failure scenarios (3) and (4) using a second central location for orchestrating a new gNB-CU-CP instance via OAM, but the critical and more important issue from an operator’s perspective is fast recovery in the geo-redundant case as 3GPP specifications do not provide suitable support for it.

Therefore, the failure case where a benefit of a resiliency enhancement in 3GPP specifications is seen would be:

- Outage of gNB-CU(-CP) when deployed in a central location without sufficient local redundancy, but availability of geo-redundancy (SW/HW) is given and fast recovery of gNB-CU(-CP) functionality is required by the operator.

Note: Fast recovery would also incorporate the gNB-CU-UP functionality in a redundant central location if it was running in the primary central location experiencing the failure.

## A.5 Failure Case (R3-224576)

*If a gNB-CU-CP fails, the network support for the NR cells is lost and UEs cannot communicate*. More specifically, the following can be observed:

● CP signaling redundancy is possible in RAN, but still only one RRC-anchor (gNB-CU-CP) is defined:

○ SRB duplication is possible in DC (PDCP duplicates RRC messages, which are sent via different carriers to/from UE), but losing the main connection to the UE still means CP failure

○ RLF handling:

▪ Rel-15: if SCG connection fails, inform the network; if MCG fails, the connection is lost

▪ Rel-16: if SCG/MCG connection fails, inform the network via MCG/SCG. There will be MCG/SCG UP downtime due to DRBs suspension.

● CP failure leading to UP failure (UP failure, leading to service interruption time, is probably the most important KPI associated with a CU-CP failure):

○ Generally, an absence of RRC messages is not noted by the UE, so the UE does not know that the CP has failed

▪ RRC timer expiry procedures in RRC work as long as RRC connectivity is up. In some cases DRBs suspension is caused by the expiration of timers

○ If the gNB-CU-CP goes down, CP connectivity is not re-established automatically until the failure is detected

○ UP connection could be lost if RRC reconfiguration (critical to maintain UP) is not possible

▪ e.g. Handover required; bearer reconfiguration required; meet QoS or other RAN reconfiguration required

There is always a single RRC anchor even for the case of DC; furthermore, there is currently no way for the UE to detect a gNB(-CU(-CP)) failure.

A CU-CP failure is essentially a node internal failure (i.e. it may not be the whole logical node that fails).

## A.6 Scenario and issue description (R3-224627)

### A.6.1 Scenario #1: Geographical Redundancy

#### A.6.1.1 General description

Geographical redundancy is the distribution of mission-critical components or infrastructures, such as the servers across multiple data centers that reside in different geographic locations, which is able to ensures high availability and disaster recovery. Geographical redundancy will replicate the data and store it in other databases located in the separate physical locations. Even one of the locations fails or simply needs to be taken offline, the other location with the replicated data will not be affected.

Considering the geographical redundancy for gNB-CU-CP, the backup gNB-CU-CP could be allocated in the separate physical locations. In this case, even the original gNB-CU-CP detects the failures and cannot work, the appearance of the backup gNB-CU-CP is able to avoid the interruption of multiple UP traffic or the disconnection of multiple UEs.

## A.7 Scenario and issue description (R3-224754)

### A.7.1 Local Redundancy

The split NG-RAN architecture consists of single logical gNB-CU-CP connected to multiple logical gNB-CU-UPs and multiple logical gNB-DUs, and the gNB-CU-CP is connected to multiple AMFs in 5GC. The logical gNB-CU-CP function can be implemented and deployed with the Network Function Virtualization as shown in Figure A.7.1-1. And the logical gNB-CU-CP function can be split further with sub-functions, e.g. network interface sub-function, and each sub-function can be implemented as a virtual network function component (VNFC).

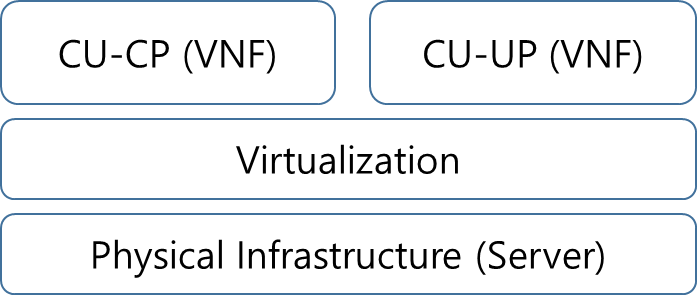


Figure A.7.1-1: gNB-CU-CP failure scenarios

### A.7.2 Failure Scenarios

#### A.7.2.1 Server Failure

The server may not work properly because of hardware failure or virtualization platform failure. The hardware failure is caused owing to some deterioration or the instant failure of the hardware. Depending on the failure cause, the hardware needs to be replaced or be recovered by rebooting the system. The virtualization platform failure is caused by software error, so it could be also recovered by rebooting the system or by the software upgrade.

#### A.7.2.2 VNFC (Virtual Network Function Component) Failure

The logical gNB-CU-CP function can be implemented with sub-functions, e.g. network interface sub-function, and they can be separately implemented as VNFC (Virtual Network Function Component) and interact each other. One or more VNFCs consisting a gNB-CU-CP may not work properly owing to software error or other reason. Mostly the VNFC failure could be recovered by rebooting the system or by the software upgrade.

#### A.7.2.3 Transport Link Failure (out of 3GPP scope)

Transport link failure is caused by the malfunction of transport link hardware or software. The failure also comes from the transmission line problem or the transport network equipment failure. If the transport link failure caused in the gNB-CU-CP system, the failure could be recovered by rebooting the system or by the software upgrade. If the failure caused outside the gNB-CU-CP system, the transmission line needs to be substituted or the network equipment needs to be rebooted, upgraded or substituted.

### A.7.3 Failure Impact

In any failure scenarios, there exist solutions to recover the logical gNB-CU-CP operation, e.g. system reboot, software upgrade, hardware replacement. Depending on the recovery solution, it may take a few minutes or some days and it may break the service continuity of the gNB-CU-CP.

So to minimize the interruption of UP traffic and disconnection of UEs, the virtualized gNB-CU-CP function could be duplicated, and the data used is shared and synchronized between duplicated gNB-CU-CP function. And if a failure is detected, the role of duplicated gNB-CU-CP functions are exchanged to support fast switch-over by implementation.

## A.8 Scenario and issue description (R3-224898)

### A.8.1 Possible failure scenarios

The purpose of this SI is to Study and identify failure scenarios associated with the gNB-CU-CP, based on the current architecture for the NG-RAN, which could be referred in 38.401 [2], see below:



Figure A.8.1-1 Overall architecture for separation of gNB-CU-CP and gNB-CU-UP

gNB-CU-CP as a logical node, could be deployed in different way which is up to operator and vendor’s strategy, e.g. could be software plus dedicated hardware within a physical box or, could be software instance run over virtualized environment (generic hardware). Thus, the failure could happen at any single point distributed at any place in Figure A.8.1-1, i.e. either to software or to hardware which would finally lead to the unavailable of gNB-CU-CP; in addition, power is also a main factor causing failure.

# Annex <B> (informative): Change history

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Change history | | | | | | | |
| Date | Meeting | TDoc | CR | Rev | Cat | Subject/Comment | New version |
| 2022-08 | RAN3#117-e | R3-225156 |  |  |  | TR skeleton | 0.0.1 |
| 2022-08 | RAN3#117-e | R3-225212 |  |  |  | Updated draft TR | 0.1.0 |
| 2022-08 | RAN3#117-e | R3-225251 |  |  |  | Editorial changes in Change history | 0.1.1 |
| 2022-09 | RAN#97-e | RP-222271 |  |  |  | For approval in plenary | 1.0.0 |
| 2022-09 | RAN#97-e |  |  |  |  | TR approved by TSG RAN plenary | 18.0.0 |