



**Network Technologies (NTECH);
Autonomic network engineering
for the self-managing Future Internet (AFI);
Autonomicity and Self-Management in the Backhaul
and Core network parts of the 3GPP Architecture**

Reference

DTR/NTECH-AFI-0017-GANA-3GPP

Keywords

3GPP, architecture, autonomic networking

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - NAF 742 C
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° 7803/88

Important notice

The present document can be downloaded from:
<http://www.etsi.org/standards-search>

The present document may be made available in electronic versions and/or in print. The content of any electronic and/or print versions of the present document shall not be modified without the prior written authorization of ETSI. In case of any existing or perceived difference in contents between such versions and/or in print, the only prevailing document is the print of the Portable Document Format (PDF) version kept on a specific network drive within ETSI Secretariat.

Users of the present document should be aware that the document may be subject to revision or change of status. Information on the current status of this and other ETSI documents is available at
<https://portal.etsi.org/TB/ETSIDeliverableStatus.aspx>

If you find errors in the present document, please send your comment to one of the following services:
<https://portal.etsi.org/People/CommiteeSupportStaff.aspx>

Copyright Notification

No part may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm except as authorized by written permission of ETSI.

The content of the PDF version shall not be modified without the written authorization of ETSI.
The copyright and the foregoing restriction extend to reproduction in all media.

© European Telecommunications Standards Institute 2016.
All rights reserved.

DECT™, **PLUGTESTS™**, **UMTS™** and the ETSI logo are Trade Marks of ETSI registered for the benefit of its Members.
3GPP™ and **LTE™** are Trade Marks of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners.
GSM® and the GSM logo are Trade Marks registered and owned by the GSM Association.

Contents

Intellectual Property Rights	5
Foreword.....	5
Modal verbs terminology.....	5
1 Scope	6
2 References	6
2.1 Normative references	6
2.2 Informative references.....	6
3 Definitions, symbols and abbreviations	9
3.1 Definitions.....	9
3.2 Symbols.....	10
3.3 Abbreviations	11
4 Background	13
4.1 Introduction	13
4.2 The GANA Reference Model in Brief	14
5 The 3GPP Reference Architecture considered for fusion with GANA Functional Blocks (FBs).....	15
5.1 Working reference architecture	15
5.2 3GPP entities and reference points.....	18
5.3 Assumptions for the present document	21
5.3.1 Assumptions for the EPC.....	21
5.3.2 Assumptions for the backhaul.....	22
6 Instantiation of GANA reference model to 3GPP EPC architecture and basic functions	22
6.1 Introduction	22
6.2 Self-Configuration/Commissioning	22
6.2.1 General description and use cases.....	22
6.3.2 Protocol-level entities (GANA Level 1) to be managed (orchestrated dynamically and adaptively (re)-configured to adapt to changes)	23
6.3.3 Identification of GANA DEs for Self-configuration	24
6.3.4 Mapping of the key Self-configuration functions	25
6.3.4.1 MME.....	25
6.3.4.2 HSS	27
6.3.4.3 IP/MPLS router.....	29
6.4 Self-Optimization	31
6.4.1 General Description and use cases.....	31
6.4.2 Protocol-level entities (GANA Level 1) to be managed (orchestrated dynamically and adaptively (re)-configured to adapt to changes)	31
6.4.3 Identification of GANA DEs for Self-Optimization.....	32
6.4.4 Mapping of Self-Optimization functions	33
6.4.4.1 MME.....	33
6.4.4.2 P-GW	34
6.4.4.3 S-GW	35
6.4.5 Indicative example of instantiation - MME pooling management.....	36
6.4.5.1 General Description	36
6.4.5.2 MME pooling revisiting	36
6.4.5.3 Hooks for GANA instantiation	38
6.4.5.3.1 Instantiation of DEs and indicative parameters that need to be dynamically managed and controlled by the DEs	38
6.4.5.3.2 Policies distribution	40
6.5 Self-Healing	40
6.5.1 General Description and use cases.....	40
6.5.2 Protocol-level entities (GANA Level 1) to be managed (orchestrated dynamically and adaptively (re)-configured to adapt to changes)	41
6.5.3 Identification of GANA DEs for Self-Healing	41
6.5.4 Mapping of Self-Healing functions	42

6.5.4.1	Mapping of instantiated GANA DEs to protocols and mechanisms for MME and PGW	42
6.5.4.2	MME	43
6.5.4.3	P-GW	44
6.5.5	Indicative examples of GANA instantiation - MME Self-Healing	45
6.6	GAN Network Level view with the Knowledge Plane (KP)	47
6.6.1	Relationship between C-SON and GANA KP, and KP southbound interfaces instantiations for 3GPP Core	47
6.6.2	Instantiation of the GANA Network Governance Interface of the KP for the 3GPP Core Network	47
7	Autonomic behaviours across multiple segments to enable End-to-End (E2E) Self-Optimization	49
7.1	Introduction	49
7.2	Reference points and associated protocols	50
7.3	Load Control	52
7.4	Congestion resolution	55
7.5	S-GW failure (without restart)	59
8	Recommendations	63
8.1	Overview	63
8.2	Reference Point between C-SON and GANA KP for the Core Network	63
8.3	Enhancing the GANA KP for EPC with analytics capabilities	64
8.4	Policies-related considerations	65
8.5	Other considerations	65
9	Conclusion and Further Work	67
Annex A:	Change History	69
History		70

Intellectual Property Rights

IPRs essential or potentially essential to the present document may have been declared to ETSI. The information pertaining to these essential IPRs, if any, is publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (<https://ipr.etsi.org>).

Pursuant to the ETSI IPR Policy, no investigation, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Foreword

This Technical Report (TR) has been produced by ETSI Technical Committee Network Technologies (NTECH).

Modal verbs terminology

In the present document "**should**", "**should not**", "**may**", "**need not**", "**will**", "**will not**", "**can**" and "**cannot**" are to be interpreted as described in clause 3.2 of the [ETSI Drafting Rules](#) (Verbal forms for the expression of provisions).

"**must**" and "**must not**" are **NOT** allowed in ETSI deliverables except when used in direct citation.

1 Scope

The present document aims at providing recommendations for the introduction of autonomics (management and control intelligence) in the 3GPP Core and Backhaul network architectures. To this effect, it covers the instantiation of the reference model for Autonomic Networking, Cognition and Self-Management, called GANA (Generic Autonomic Networking Architecture), onto the architecture defined in ETSI TS 123 401 [i.2] and ETSI TS 123 402 [i.35]. It superimposes GANA Decision Elements (DEs) into node/device architectures and the overall 3GPP network architecture, so that the DEs and their associated Control-Loops can be further designed to perform autonomic management and control of the specific resources (Managed Entities) in the target architecture. It develops recommendations on the basic behaviours of the GANA Functional Blocks (FBs) in the above context. 3GPP specifications on policy control (ETSI TS 123 203 [i.3]) and network management (TS 123 32x series) are taken into account into the working reference architecture. It also involves the backhaul network and associated interactions between the different entities for an optimization with an end-to-end perspective.

2 References

2.1 Normative references

Normative references are not applicable in the present document.

2.2 Informative references

References are either specific (identified by date of publication and/or edition number or version number) or non-specific. For specific references, only the cited version applies. For non-specific references, the latest version of the referenced document (including any amendments) applies.

NOTE: While any hyperlinks included in this clause were valid at the time of publication, ETSI cannot guarantee their long term validity.

The following referenced documents are not necessary for the application of the present document but they assist the user with regard to a particular subject area.

- [i.1] ETSI TS 123 002 (V12.5.0): "Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); LTE; Network architecture (3GPP TS 23.002 version 12.5.0 Release 12)".
- [i.2] ETSI TS 123 401 (V13.6.1): "LTE; General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) access (3GPP TS 23.401 version 13.6.1 Release 13)".
- [i.3] ETSI TS 123 203 (V13.8.0): "Digital cellular telecommunications system (Phase 2+) (GSM); Universal Mobile Telecommunications System (UMTS); LTE; Policy and charging control architecture (3GPP TS 23.203 version 13.8.0 Release 13)".
- [i.4] ETSI TS 132 501 (V13.0.0): "Universal Mobile Telecommunications System (UMTS); LTE; Telecommunication management; Self-configuration of network elements; Concepts and requirements (3GPP TS 32.501 version 13.0.0 Release 13)".
- [i.5] ETSI TS 132 541 (V13.0.0): "Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); LTE; Telecommunication management; Self-Organizing Networks (SON); Self-healing concepts and requirements (3GPP TS 32.541 version 13.0.0 Release 13)".
- [i.6] ETSI NTECH AFI White Paper (draft): "The GANA (Generic Autonomic Networking Architecture) Reference Model for Autonomic Networking, Cognitive Networking and Self-Management of Networks and Services".

- [i.7] ETSI Draft DTR/NTECH-AFI-0018-GANA-Mesh: "Network Technologies (NTECH);Autonomic network engineering for the self-managing Future Internet (AFI); Autonomicity and Self-Management in Wireless Ad-hoc/Mesh Networks".
- [i.8] ETSI GS AFI 002: "Autonomic network engineering for the self-managing Future Internet (AFI); Generic Autonomic Network Architecture (An Architectural Reference Model for Autonomic Networking, Cognitive Networking and Self-Management)".
- [i.9] NGMN Alliance (2008): "NGMN Recommendation on SON and O&M Requirements".
- NOTE: Available at https://www.ngmn.org/uploads/media/NGMN_Recommendation_on_SON_and_O_M_Requirements.pdf.
- [i.10] Andrea Fabrizi: "NFV self-adaptive networks; Integrate your control plane with analytics", HP Viewpoint paper, July 2015.
- NOTE: Available at <https://www.hpe.com/h20195/V2/getpdf.aspx/4AA4-6147ENW.pdf>.
- [i.11] Ranganai Chaparadza, Tayeb Ben Meriem, John Strassner, Steven Wright and Joel Halpern: "Industry Harmonization for Unified Standards on Autonomic Management & Control of Networks and Services, SDN, NFV, E2E Orchestration, and Software-oriented enablers for 5G", Report from the Joint SDOs/Fora Workshop hosted by TMForum during TMForum Live 2015 Event Nice, France (June 4th, 2015).
- [i.12] NGMN Alliance: "RAN Evolution Project, Backhaul and Fronthaul Evolution", 31 March 2015.
- NOTE: Available at <http://docplayer.net/10579127-Ran-evolution-project-backhaul-and-fronthaul-evolution.html>.
- [i.13] ETSI TS 129 274 (V13.6.0): "Universal Mobile Telecommunications System (UMTS); LTE; 3GPP Evolved Packet System (EPS); Evolved General Packet Radio Service (GPRS) Tunnelling Protocol for Control plane (GTPv2-C); Stage 3 (3GPP TS 29.274 version 13.6.0 Release 13)".
- [i.14] ETSI TS 129 281 (V13.2.0): "Universal Mobile Telecommunications System (UMTS); LTE; General Packet Radio System (GPRS) Tunnelling Protocol User Plane (GTPv1-U) (3GPP TS 29.281 version 13.2.0 Release 13)".
- [i.15] ETSI TS 129 272 (V13.6.0): "Universal Mobile Telecommunications System (UMTS); LTE; Evolved Packet System (EPS); Mobility Management Entity (MME) and Serving GPRS Support Node (SGSN) related interfaces based on Diameter protocol (3GPP TS 29.272 version 13.6.0 Release 13)".
- [i.16] ETSI TS 129 061 (V13.4.0): "Digital cellular telecommunications system (Phase 2+) (GSM); Universal Mobile Telecommunications System (UMTS); LTE; Interworking between the Public Land Mobile Network (PLMN) supporting packet based services and Packet Data Networks (PDN) (3GPP TS 29.061 version 13.4.0 Release 13)".
- [i.17] ETSI TS 129 212 (V13.6.0): "Universal Mobile Telecommunications System (UMTS); LTE; Policy and Charging Control (PCC); Reference points (3GPP TS 29.212 version 13.6.0 Release 13)".
- [i.18] ETSI TS 132 251 (V13.5.0): "Digital cellular telecommunications system (Phase 2+) (GSM); Universal Mobile Telecommunications System (UMTS); LTE; Telecommunication management; Charging management; Packet Switched (PS) domain charging (3GPP TS 32.251 version 13.5.0 Release 13)".
- [i.19] IETF RFC 4006: "Diameter Credit-Control Application".
- [i.20] ETSI TS 132 295 (V13.0.0): "Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); LTE; Telecommunication management; Charging management; Charging Data Record (CDR) transfer (3GPP TS 32.295 version 13.0.0 Release 13)".
- [i.21] ETSI TS 129 219 (V13.3.0): "Digital cellular telecommunications system (Phase 2+) (GSM); Universal Mobile Telecommunications System (UMTS); LTE; Policy and charging control: Spending limit reporting over Sy reference point (3GPP TS 29.219 version 13.3.0 Release 13)".

- [i.22] ETSI TS 132 102 (V13.0.0): "Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); LTE; Telecommunication management; Architecture (3GPP TS 32.102 version 13.0.0 Release 13)".
- [i.23] 3GPP TR 23.705 (V13.0.0) (12-2014): "3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Study on system enhancements for user plane congestion management (Release 13)".
- [i.24] 4G Americas white paper: "Self-Optimizing Networks: The Benefits of SON in LTE", October 2013.
- NOTE: Available at http://www.4gamericas.org/files/7714/0759/1264/Self-Optimizing_Networks-Benefits_of_SON_in_LTE_10.7.13.pdf.
- [i.25] ETSI TS 132 426 (V13.0.0): "LTE; Telecommunication management; Performance Management (PM); Performance measurements Evolved Packet Core (EPC) network (3GPP TS 32.426 version 13.0.0 Release 13)".
- [i.26] 3GPP TR 23.857 (V11.0.0) (12-2012): "Technical Specification Group Core Network and Terminals; Study of Evolved Packet Core (EPC) nodes restoration (Release 11)".
- [i.27] ETSI TS 129 211 (V6.4.0): "Universal Mobile Telecommunications System (UMTS); Rx Interface and Rx/Gx signalling flows (3GPP TS 29.211 version 6.4.0 Release 6)".
- [i.28] ETSI TS 132 299 (V13.5.0): "Digital cellular telecommunications system (Phase 2+) (GSM); Universal Mobile Telecommunications System (UMTS); LTE; Telecommunication management; Charging management; Diameter charging applications (3GPP TS 32.299 version 13.5.0 Release 13)".
- [i.29] ETSI TS 129 214 (V13.6.0): "Universal Mobile Telecommunications System (UMTS); LTE; Policy and charging control over Rx reference point (3GPP TS 29.214 version 13.6.0 Release 13)".
- [i.30] ETSI TS 132 101 (V13.0.0): "Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); LTE; Telecommunication management; Principles and high level requirements (3GPP TS 32.101 version 13.0.0 Release 13)".
- [i.31] IETF Draft: "IP based Generic Control Protocol (IGCP): draft-chaparadza-intarea-igcp-00.txt".
- NOTE: Available at <https://tools.ietf.org/html/draft-chaparadza-intarea-igcp-00>.
- [i.32] NGMN Alliance: "Recommended Practices for Multi-vendor SON Deployment", Deliverable D2, 28 January 2014.
- NOTE: Available at https://www.ngmn.org/uploads/media/P-Small_Cells_WS2_Multivendor_Recommended_Practices_v1_0.pdf
- [i.33] ETSI TS 103 371: " Network Technologies (NTECH); Autonomic network engineering for the self-managing Future Internet (AFI); Proofs of Concept Framework".
- [i.34] ETSI TS 103 194: "Network Technologies (NTECH); Autonomic network engineering for the self-managing Future Internet (AFI); Scenarios, Use Cases and Requirements for Autonomic/Self-Managing Future Internet".
- [i.35] ETSI TS 123 402 (V13.6.0): "Universal Mobile Telecommunications System (UMTS); LTE; Architecture enhancements for non-3GPP accesses (3GPP TS 23.402 version 13.6.0 Release 13)".
- [i.36] ETSI TS 129 215: "Digital cellular telecommunications system (Phase 2+) (GSM); Universal Mobile Telecommunications System (UMTS); LTE; Policy and Charging Control (PCC) over S9 reference point; Stage 3 (3GPP TS 29.215)".
- [i.37] ETSI TS 123 228: "Digital cellular telecommunications system (Phase 2+) (GSM); Universal Mobile Telecommunications System (UMTS); LTE; IP Multimedia Subsystem (IMS); Stage 2 (3GPP TS 23.228)".

[i.38] ETSI ES 283 035: "Telecommunications and Internet converged Services and Protocols for Advanced Networking (TISPAN); Network Attachment Sub-System (NASS); e2 interface based on the DIAMETER protocol".

3 Definitions, symbols and abbreviations

3.1 Definitions

For the purposes of the present document, the following terms and definitions apply:

Autonomic Behaviour (AB): process which understands how desired Managed Entity (ME) behaviours are learned, influenced or changed, and how, in turn, these affect other elements, groups and network [i.34]

NOTE: In the GANA model, an autonomic behaviour is any behaviour of a DE that is observable on its interfaces. Autonomic function (AF) desired element's behaviour in order to learn, influence, or changed on how it turns its effect on other elements or groups of element or network.

autonomic networking: networking paradigm that enables network devices or elements (physical or virtual) and the overall network architecture and its management and control architecture to exhibit the so-called self-managing properties, namely: auto-discovery of information and entities, Self-configuration (auto-configuration), Self-diagnosing, Self-repair (Self-healing), Self-optimization, and other self-* properties

NOTE 1: Autonomic Networking can also be interpreted as a discipline involving the design of systems (e.g. network nodes) that are self-managing at the individual system levels and together as a larger system that forms a communication network of systems.

NOTE 2: The term "autonomic" comes from the autonomic nervous system (a closed control loop structure), which controls many organs and muscles in the human body. Usually, human are unaware of its workings because it functions in an involuntary, reflexive manner - for example, human do not notice when their heart beats faster or their blood vessels change size in response to temperature, posture, food intake, stressful experiences and other changes to which human are exposed. And their autonomic nervous system is always working [i.34].

Complex Event Processing (CEP): data processing discipline which correlates data from multiple sources to identify patterns of events

context aware: capability of a component or system to be aware of its execution environment, the objectives it is supposed to meet and possibly the consequences of not delivering on the objectives, and be able to react to changes in the environment

Decision Making Element (DME): functional entity designed and assigned to autonomically manage and control its assigned Managed Entities (MEs) by dynamically (re)-configuring the MEs and their configurable and controllable parameters in a closed-control loop fashion.

NOTE 1: Decision Making Elements (DMEs) [i.15] referred in short as Decision Elements (DEs) fulfil the role of Autonomic Manager Elements.

NOTE 2: In GANA a DE is assigned (by design) to very specific MEs that it is designed to autonomically manage and control (ETSI GS AFI 002 [i.8] provides more details on the notion of ownership of MEs by specific DEs required in a network element architecture and the overall network architecture).

Managed Entities (Mes): physical or logical resource that can be managed by an Autonomic Manager Element (i.e. a Decision Element) in terms of its orchestration, configuration and re-configuration through parameter settings [i.34]

NOTE: MEs and their associated configurable parameters are assigned to be managed and controlled by a concrete DE such that an ME parameter is mapped to one DE. Mes can be protocols, whole protocol stacks, and mechanisms, meaning that they can be fundamental functional and manageable entities at the bottom of the management hierarchy (at the fundamental resources layer in a network element or node) such as individual protocols or stacks, OSI layer 7 or TCP/IP application layer applications and other types of resources or managed mechanisms hosted in a network element (NE) or in the network in general, whereby an ME exposes a management interface through which it can be managed. Mes can also be composite Mes such as whole Nes themselves (i.e. Mes that embed sub-Mes).

non-aggregated scenario: scenario of 3GPP architecture without the aggregation of other types of networks, e.g. previous generations of mobile networks

overlay: logical network that runs on top of another network

NOTE: For example, peer-to-peer networks are overlay networks on the Internet. They use their own addressing system for determining how files are distributed and accessed, which provides a layer on top of the Internet's IP addressing.

self-advertising: capability of a component or system to advertise its self-model, capability description model, or some information signalling message (such as an Ipv6 router advertisement message) to the network in order to enable other entities to discover it and be able to communicate with it, or to enable other entities to know whatever is being advertised

self-awareness: capability of a component or system to "know itself" and be aware of its state and its behaviours. Knowledge about "self" is described by a "self-model"

self-configuration: capability of a component or system to configure and reconfigure itself under varying and unpredictable conditions

self-descriptive: capability of a component or system to provide a description of its self-model, capabilities and internal state

self-healing: capability of a component or system to detect and recover from problems (manifestations of faults, errors, failures, and other forms of degradation) and continue to function smoothly

self-monitoring: capability of a component or system to observe its internal state, for example by monitoring quality-of-service metrics such as reliability, precision, rapidity, or throughput

self-optimization: capability of a component or system to detect suboptimal behaviours and optimize itself to improve its execution

self-organizing function: function that includes processes which require minimum manual intervention

self-protecting: capability of a component or system to be capable of detecting and protecting its resources from both internal and external attack and maintaining overall system security and integrity

self-regulation: capability of a component or system to regulate its internal parameters so as to assure a quality-of-service metric such as reliability, precision, rapidity, or throughput

3.2 Symbols

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

Gx	Reference point between PCEF and PCRF
Gy	Reference point between OCS and PCEF
Gyn	Reference point between OCS and TDF
Gz	Reference point between OFCS and PCEF
Gzn	Reference point between OFCS and TDF
Itf-N	Interface between the Network (Element Manager or NEs with an embedded EM) and the Network Manager
Itf-P2P	Interface between peer Domain Managers

Np	Reference point between RCAF and PCRF
Rx	Reference point between PCRF and AF
S3	Reference point between MME and SGSN
S4	Reference point between S-GW and SGSN
S5	Reference point between S-GW and P GW
S6a	Reference point between MME and HSS
S6d	Reference point between SGSN and HSS
S11	Reference point between MME and S-GW
Sd	Reference point between TDF and PCRF
Sgi	Reference point between PDN GW and packet data networks
Sp	Reference point between SPR and PCRF
Sy	Reference point between OCS and PCRF

3.3 Abbreviations

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

3GPP	3 rd Generation Partnership Project
AC_DE	Auto-Configuration Decision Element
AF	Autonomic Function
AFI	Autonomic network engineering for the self-managing Future Internet
AMC	Autonomic Management and Control
AN	Access Network
ANDSF	Access Network Discovery and Selection Function
ANIMA	Autonomic Networking Integrated Model and Approach
AP	Access Point
APN	Access Point Name
BG	Border Gateway
BGP	Border Gateway Protocol
BH	BackHaul
CAPEX	Capital Expenditures
CEP	Complex Event Processing
CMIP	Common Management Information Protocol
CN	Core Network
CN-EMS	Core Network-Element Management System
CN-NMS	Core Network-Network Management System
COPS	Common Open Policy Service
CORBA	Common Object Request Broker Architecture
CORBA IIOP	Common Object Request Broker Architecture Internet Inter-ORB Protocol
CPU	Central Processing Unit
CSO	Centralized SON
C-SON	Centralized Self-Organizing Network
DE	Decision-making-Element
DHCP	Dynamic Host Configuration Protocol
DL	DownLink
DPI	Deep Packet Inspection
e2e	end-to-end
EIR	Equipment Identity Register
EM	Element Manager
EMS	Element Management System
EPC	Evolved Packet Core
EPC-KP	Evolved Packet Core-Knowledge Plane
EPS	Evolved Packet System
E-UTRAN	Evolved Universal Terrestrial Radio Access Network
FB	Functional Block
FM_DE	Fault Management Decision Element
GAN	Generic Autonomic Network Architecture
GERAN	GSM EDGE RAN
GMLC	Gateway Mobile Location Centre
GPRS	General Packet Radio service
GRASP	GeneRic Autonomic Signalling Protocol

GS	Group Specifications
GTP	GPRS Tunnelling Protocol
GTP-C	GPRS Tunnelling Protocol control plane
GTP-U	GTP user plane
GW	GateWay
HLR	Home Location Register
HSS	Home Subscriber Server
HW	HardWare
IDE	Incident Dissemination Engine
IETF	Internet Engineering Task Force
IGCP	ICMPv6 based Generic Control Protocol
IIOP	Internet Inter-Object request broker Protocol
IMS	IP Multimedia Subsystem
IP	Internet Protocol
IP-CAN	IP Connectivity Access Network
KP	Knowledge Plane
KPI	Key Performance Indicator
LBI	Linked Bearer ID
LS	Liaison Statement
LTE	Long Term Evolution
LTE-A	LTE-Advanced
MBTS	Model-Based-Translation Service
ME	Managed Entity
MIB	Management Information Base
MM	Mobility Management
MME	Mobility Management Entity
MNO	Mobile Network Operator
MPLS	Multi-Protocol Label Switching
ND	Neighbour Discovery
NE	Network Element
NFV	Network Function Virtualisation
NGMN	Next Generation Mobile Networks
NM	Network Manager
NMS	Network Management System
OAM	Operation Administration and Maintenance
OCS	Online Charging System
OFCS	Offline Charging System
ONIX	Overlay Network system of information servers for Information eXchange
OPEX	Operational Expenditures
OSI	Open Systems Interconnection (model)
OSPF	Open Shortest Path First
OSS	Operation Support Systems
P Router	Provider Router
PCC	Policy and Charging Control
PCEF	Policy and Charging Enforcement Function
PCRF	Policy and Charging Rules Function
PDN	Packet Data Network
PDN-GW	Packet Data Network-GateWay
PDP	Policy Decision Point
PE	Provider Edge
PEP	Policy Enforcement Point
PGW	Packet GateWay
P-GW	Packet Data Network Gateway
PS	Packet Switched
QoS	Quality of Service
RAN	Radio Access Network
RAT	Radio Access Technology
RCAF	RAN Congestion Awareness Function
REST	Representational State Transfer
RFC	Request For Comments
Rfp	Reference point
RNC	Radio Network Controller

RPL	Routing Protocol for Low-Power and Lossy Networks
RRC	Radio Resource Control
S1-AP	S1 Application Protocol
SC	Self-Configuration
SCTP	Stream Control Transmission Protocol
SDN	Software Defined Networking
SGSN	Serving GPRS Support Node
SGW	Service GateWay
S-GW	Serving Gateway
SH	Self-Healing
SNMP	Simple Network Management Protocol
SO	Self-Optimization
SOA	Service Oriented Architecture
SOAP	Simple Object Access Protocol
SON	Self Organizing Network
SPR	Subscription Profile Repository
STF	Specialized Task Force
SW	SoftWare
TC	Technical Committee
TcoSH	Trigger Condition of Self-Healing
TCP	Transmission Control Protocol
TDF	Traffic Detection Function
TSG-SA	Technical Specification Group-System Aspects
UDC	User Data Convergence
UDP	User Datagram Protocol
UDR	User Data Repository
UE	User Equipment
UL	UpLink
UMTS	Universal Mobile Telecommunications System
UTRAN	Universal Terrestrial Radio Access Network
WG	Working Group
WLAN	Wireless Local Area Network
WS	Web Services

4 Background

4.1 Introduction

Autonomic management and control (AMC) of Network & Services is intended to help operators and enterprises in reducing OPEX and handling the increasing complexity of network management. The ETSI AFI WG of TC NTECH produces specifications for the Autonomic Networking & Services Management, namely Use Cases and Requirements for AMC in various architectures, the Generic Autonomic Network Architecture (GANA) reference model, and instantiations of the GANA model onto various implementation-oriented reference architectures that enable developers to innovate and implement algorithms for the autonomies enabling GANA Functional Blocks (FBs). The TC is now progressing in producing technical reports on instantiation of the GANA Reference Model onto existing network architectures and emerging ones to embed Self-management capabilities.

The objective of the present document is to produce specifications for the instantiation of the GANA model onto the mobile backhaul and the Evolved Packet Core (EPC) deployed as per 3GPP Architecture specifications. Depending on the specific implementation of the EPC components by the various operators, the present document can be used in order to enable the mapping of GANA reference model functions onto the EPC components, thus enabling autonomic functions (AF) to be introduced in the EPC modules.

The work has been divided into several tasks which are reflected in the following clauses. The first task consisted in extracting from 3GPP specifications the 3GPP reference architecture that has been a basis for GANA instantiation in the report. This is documented in clause 5. In a second step, a mapping of the GANA model to the 3GPP architecture was defined for a basic scenario with the 3GPP EPC and the mobile backhaul networks as standalone networks. This is reported in clause 6. The final task consisted in considering autonomic behaviours across multiple segments. This is presented in clause 7.

4.2 The GANA Reference Model in Brief

To enable the reader to understand the GANA instantiation on the 3GPP core and backhaul architectures, the present clause provides a short overview of the GANA reference model specified in ETSI GS AFI 002 [i.8] and described in [i.6].

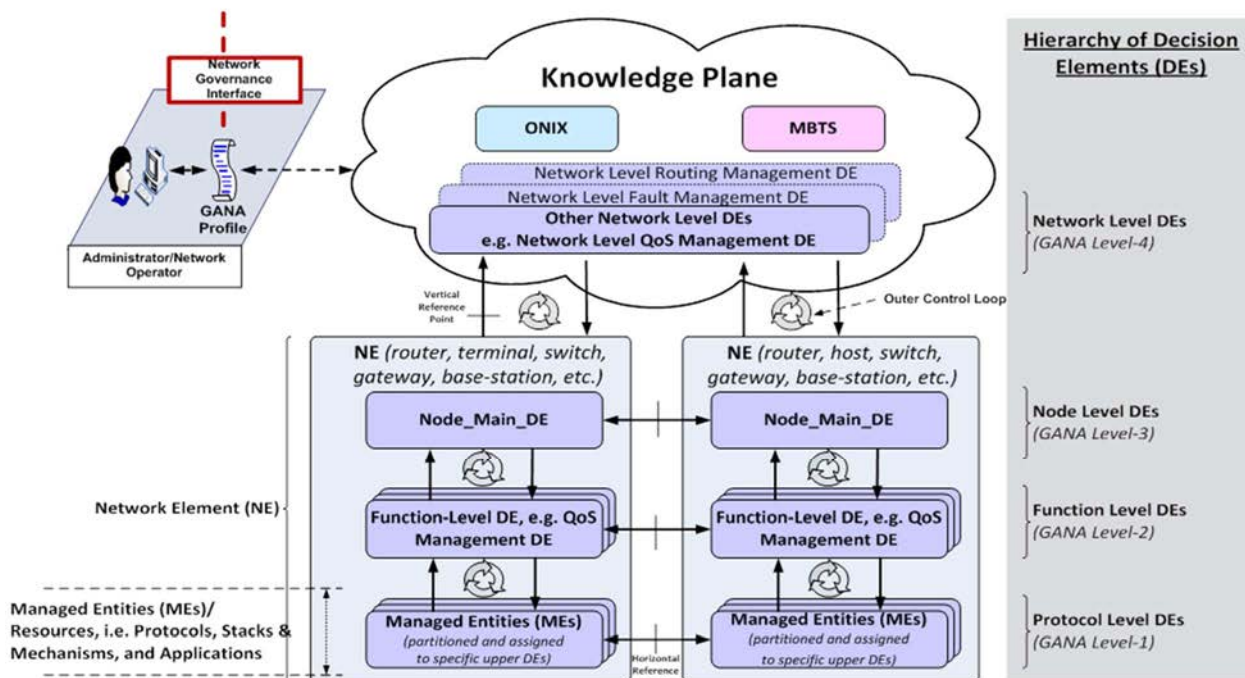


Figure 1: Snapshot of the AFI GANA Reference Model

Self-manageability in GANA is achieved through the dynamic and context-aware orchestration and management and control of Mes by collaborative Decision-making-Elements (DEs) (see definition of an ME and a DE in the Definitions clause). GANA defines a hierarchy of such DEs in four basic levels: the protocol, function, node, and network levels. At each level, a DE manages one or more lower-level DEs through a control loop. These lower-level DEs are therefore considered as Managed Entities (Mes) by the DE that controls them. Over the control loop, a DE sends commands, objectives, and policies to its lower-level DEs and receives feedback in the form of monitoring information or other type of knowledge. The Protocol Level DEs represent protocols, services, and other fundamental mechanisms running in the target network as Mes that may exhibit intrinsic control-loops (DE logic) and associated DE-as is the case for some of today's protocols such as OSPF, which can be considered an example of the instantiation of a protocol-level DE (though such autonomic-like feature in OSPF is not cognitive (learning and reasoning) in its operation and by design). As discussed in [i.6], the GANA Specification puts forward a recommendation to primarily focus on the three higher GANA levels of hierarchical control-loops (Level-2 to Level-4) when introducing autonomies in architectures and considers the protocol level DEs as Mes. The GANA hierarchy emphasizes only the three other levels, which should collaboratively work together. The argument, put forward in the GANA white paper [i.6] and ETSI GS AFI 002 [i.8], is that three levels of hierarchical control-loops (GAN level-2 to level-4) demonstrate how AMC can be gracefully (non-disruptively) introduced in today's existing networks and architectures and even in new network architectures that follow the approach of designing and employing protocols to build protocol stacks in which individual protocols are rather simple and do not embed any intrinsic control-loops.

At the lowest level in the management hierarchy in GANA is the resource layer in the Network Elements (Nes), which can be physical or virtual, that consists of Managed Entities (Mes) such as protocols or stacks, OSI layer 7 or TCP/IP application layer applications and other types of resources or managed mechanisms hosted in a network element (NE). They are managed by Function Level DEs (present in every Network Element, or NE) e.g. Routing Management DE. The orchestration of the Function Level DEs is performed by the Node Main DE. A Node main DE is present in every NE, for example a router. At the highest DE level, the Network Level DEs address similar aspects as the Function Level DEs but on a wider scope. Therefore there is a Network Level Routing Management DE, Network Level Monitoring DE, Network Level QoS Management DE, etc. The GANA Knowledge Plane (KP) is constituted by the Network Level DEs, together with a distributed, scalable Overlay Network system of information servers for Information eXchange (ONIX) and a Model-Based-Translation Service (MBTS) for translating information and commands/responses towards Nes. The ONIX is useful for enabling auto-discovery of information/resources of an autonomic network via "publish/subscribe/query&find" protocols. DEs can make use of ONIX to discover information and entities (e.g. other DEs) in the network to enhance their decisions. More details on ONIX are given in ETSI GS AFI 002 [i.8]. The ONIX itself does not have network management & control decision logic (as DEs are the ones that exhibit decision logic for AMC). MBTS (Model-Based Translation Service) is an intermediation layer between the GANA Knowledge Plane (KP) and the Nes (physical or virtual) for translating vendors' specific and technology specific raw data onto a common data model for use by network level DEs, based on an accepted and shared information model.

Moreover, governance is implemented through the Network Governance Interface. The network administrator uses this interface to manage the operation of the whole autonomic network by authoring, validating and submitting conflict-free policies, high-level network objectives and some configuration data to the KP, all encapsulated together in the form of a "GANA Network Profile" that is generated using automation tools. This GANA Profile is then used by the Network Level DEs to configure themselves and issue commands and lower level policies that are issued to lower level DEs for enforcement (including relaying sub-profiles that are used by lower level DEs to configure themselves and their Mes). More details on the GANA Network Profile creations and use can be found in the GANA white paper [i.13] and ETSI GS AFI 002 [i.8]. The GANA Profile can be augmented/extended with run-time related information by the Network Level DEs and MBTS and stored/maintained in the ONIX.

5 The 3GPP Reference Architecture considered for fusion with GANA Functional Blocks (FBs)

5.1 Working reference architecture

The identification of a working reference architecture for the 3GPP LTE-Advanced (LTE-A) Core Network (CN) is important in order to have a consolidated basis for the instantiation of the GANA Framework. Moreover, the consideration of the mapping of the GANA model and its Knowledge Plane (KP) concept onto this reference architecture should encompass the relevant layers and respective functionalities. The working reference architecture, based on ETSI TS 123 002 [i.1], ETSI TS 123 401 [i.2] and ETSI TS 123 203 [i.3] is illustrated in Figure 2, where the main functional and network entities of the core network, the EPC and PCC are depicted. The EPC contains the following logical network elements. The Mobility Management Entity (MME) is the control plane entity that mainly supports mobility and security. The Serving Gateway (S-GW) is the user plane gateway point that interconnects the radio side (Radio Access Network, or RAN, e.g. the Evolved Universal Terrestrial RAN, or E-UTRAN in case of LTE) and the CN. The Packet Data Network Gateway (P-GW) is the user plane gateway to the packet data network (PDN) i.e. the Internet or the operator's IP Multimedia Subsystem (IMS). In addition, the Home Subscriber Server (HSS) database contains user-related and subscriber-related information for supporting mobility management, call and session setup, user authentication and access authorization.

In Figure 2, the control plane and data plane interactions are highlighted. For completeness, the RAN entities, i.e. the GSM EDGE RAN (GERAN), UTRAN and E-UTRAN and non-3GPP RANs, e.g. Wireless Local Area Network (WLAN), which are not in scope of this work, are marked in grey.

Last but not least, the Policy and Charging Control (PCC) architecture [i.3] is an indispensable part of the 3GPP CN architecture. The central entry in PCC is the Policy and Charging Rule Function (PCRF), which is responsible for making policy and charging decisions based on inputs and policies received from the mobile operator, information on user subscription and services and so on. These decisions are then communicated to the Policy and Charging Enforcement Function (PCEF) (typically embodied within P-GW) in the form of PCC rules. In brief, a rule contains information for identifying a specific IP traffic flow, parameters for charging this traffic and QoS parameters to be applied to that traffic flow. PCRF and PCEF are complemented by other entities illustrated in Figure 2 and described in ETSI TS 123 203 [i.3]. Therefore it is important to follow through the reference architecture the evolution of the knowledge and policy management after the mapping of the GANA related knowledge plane and control loops.

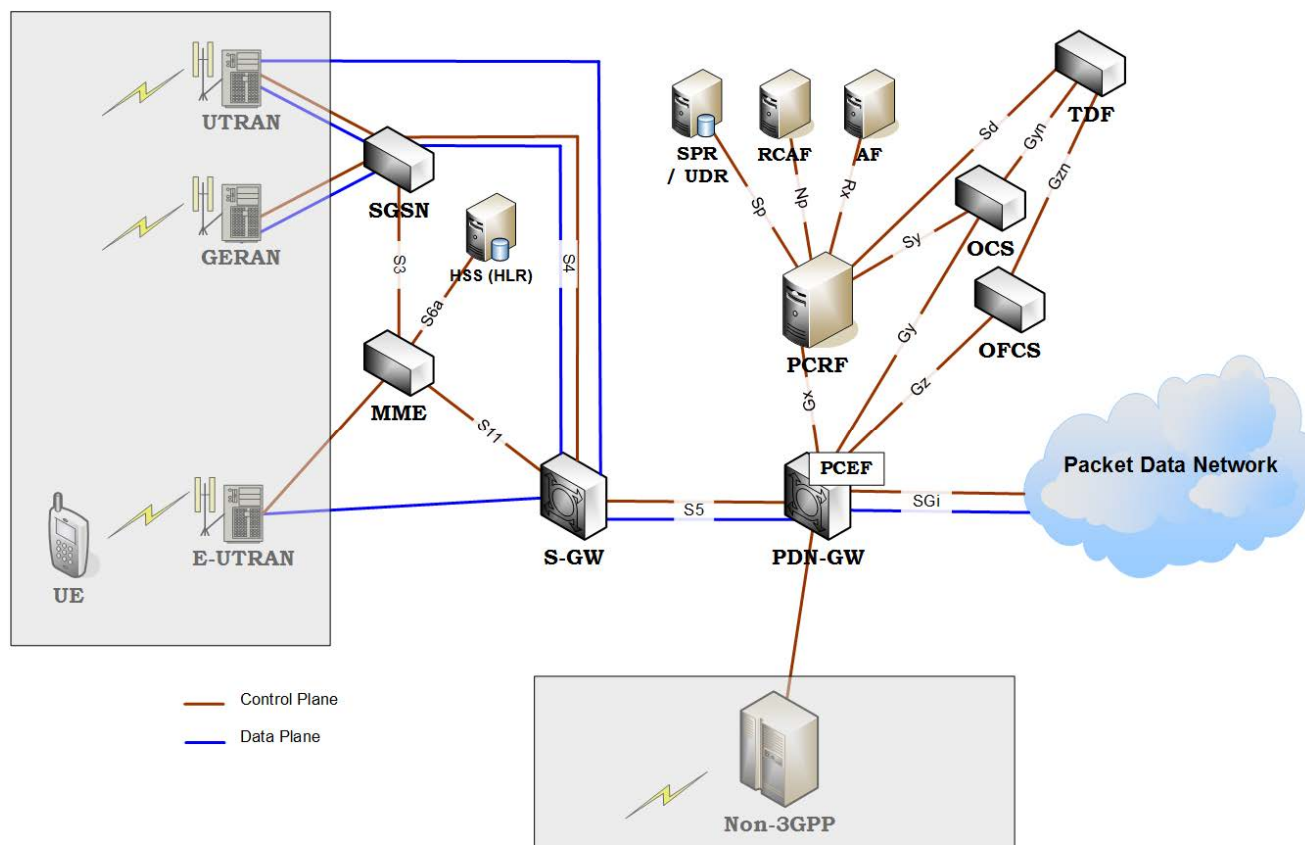


Figure 2: Reference 3GPP architecture

In Figure 3, the 3GPP Core Network reference architecture is presented, highlighting the management plane along with the control and data plane providing a more consolidated view of the reference entities and interfaces that have been considered during the GANA instantiation. The typical management plane entities namely, Network/Element Management System (N/EMS) and Operation Support Systems (OSS) together with the Northbound (Itf-N) and southbound (Itf-S) interfaces specified by 3GPP in ETSI TS 132 101 [i.30] are also depicted in Figure 3. The relation of GANA to these systems is crucial for showing the migration of today's management plane towards the GANA decision plane. Figure 3 highlights the interactions between the CN and the legacy management plane, providing a consolidated view of the reference entities and interfaces to be considered during the GANA instantiation with legacy network.

The mapping of GANA model will consider in subsequent clauses of the present document the placement of decision-making entities, the coordination of the entities and collaboration of functional blocks, feedback loops and policies. This may affect the three different views of the 3GPP CN: control, data and management plane.

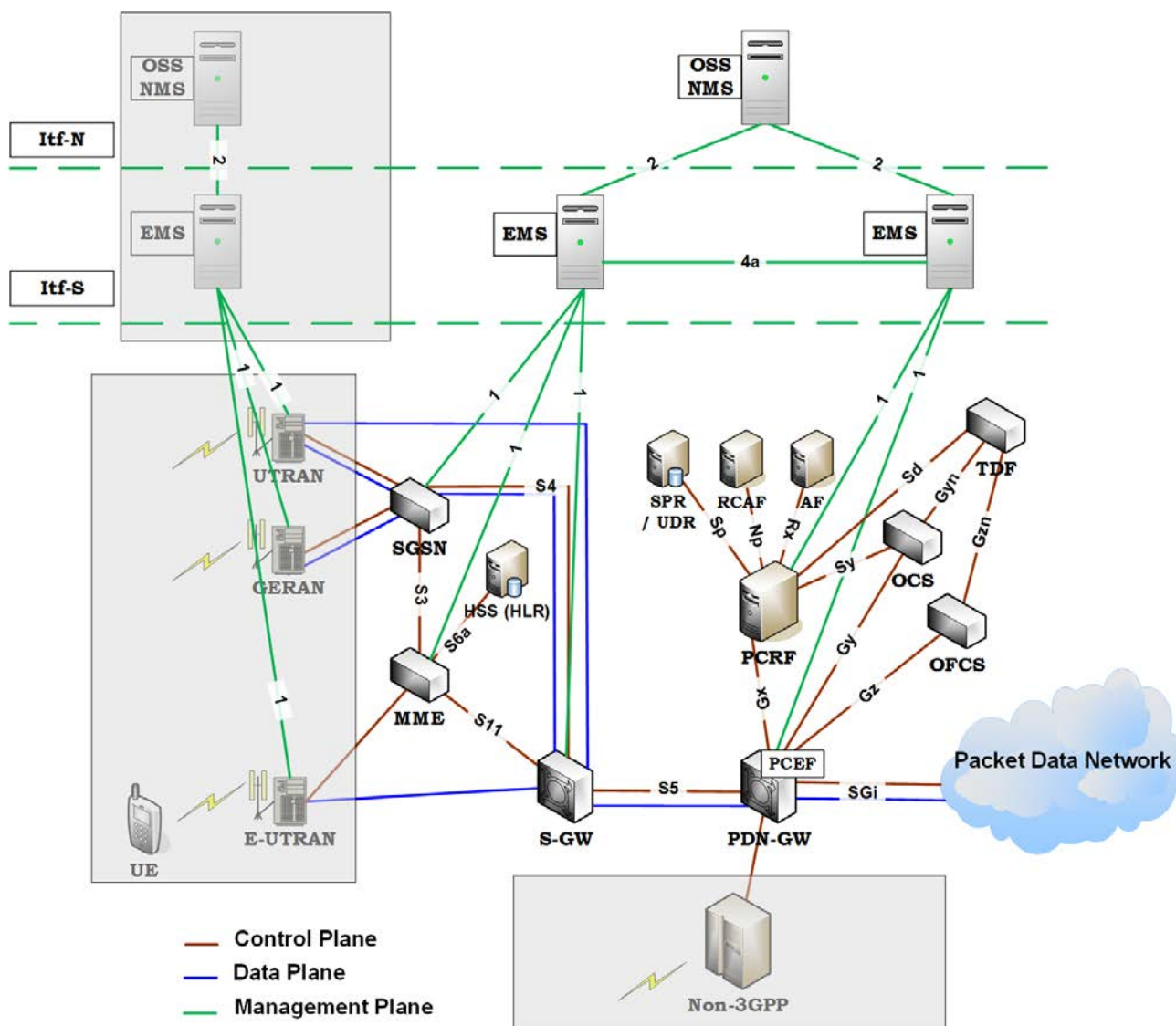


Figure 3: Management plane for the non-aggregated scenario

Figure 4 depicts the assumed reference architecture for the mobile backhaul part of the network. In principle this covers IP/MPLS enabled routers that are typically used for backhauling the traffic from the 3GPP access networks (including controllers in case of earlier releases) to the EPC core. Figure 4 also depicts the systems responsible for the management of the backhaul network in whole and per element.

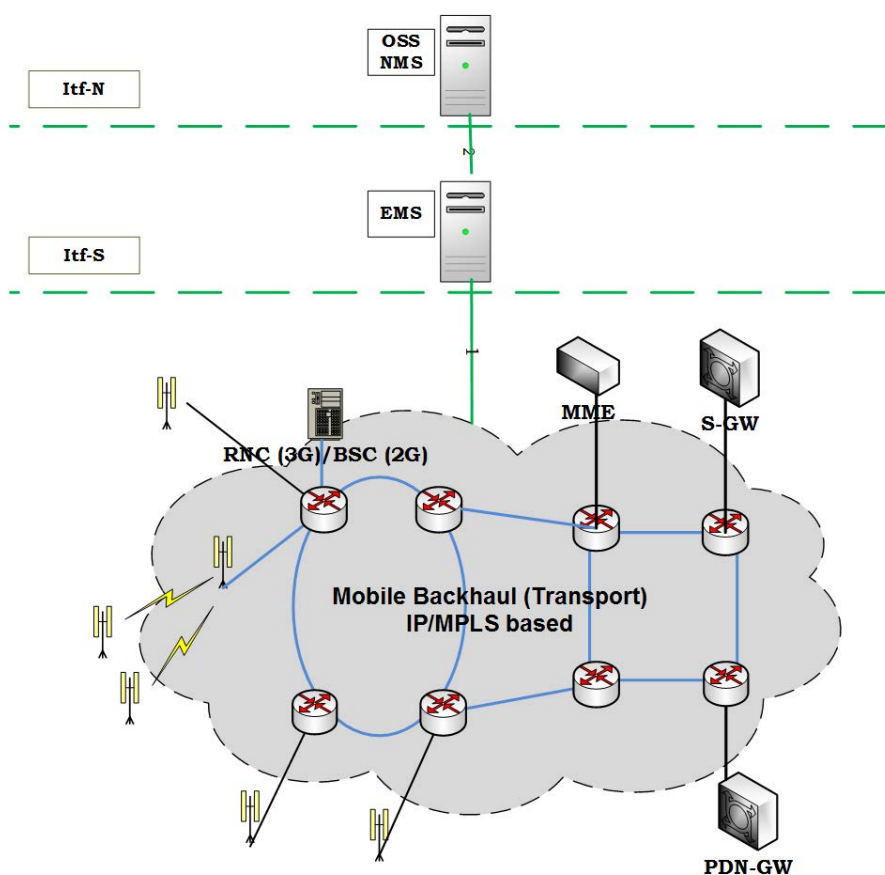


Figure 4: Reference architecture for the backhaul

Current requirements and evolution for the backhaul can be assessed by the efforts depicted in the NGMN evolution report [i.12]. Moreover, other approaches can be addressed, for example Self-backhauling for SMARTER, e.g. in recent 3GPP TSG-SA WG1 work items.

5.2 3GPP entities and reference points

Table 1 shows the distribution of the basic entities of the backhaul and Core networks that will be taken into consideration for the mapping of GANA Functional Blocks (FBs) onto the 3GPP core and backhaul network architectures. For completeness, it additionally shows the entities of the RAN and the external PDN, but they are greyed to indicate that only their impact on the CN is considered.

Table 1: Identification of entities of the CN and backhaul network to be managed

Basic entities			Service Specific Entities	Backhaul	Management Plane	External PDN entities
RAN	GPRS-PS	EPC PS				
e-UTRAN UTRAN GERAN non-3GPP Access UE	SGSN	MME S-GW P-GW HSS (HLR) EIR	PCEF PCRF OCS OFCS TDF SPR/UDR RCAF AF	IP/MPLS Router (P, PE)	CN-EMS CN-NMS	BG

Table 2 lists the reference points (Rfp) standardized in the 3GPP network architecture defined in ETSI TS 123 002 [i.1] that operate at the functional level. These reference points allow visualizing the interactions that take place between the CN entities and their associated functionalities. In the table, they are shown with their 3GPP names, the entities that they connect as well as a short description of the functionalities and mechanisms they provide, with a reference to the 3GPP standard where they are specified.

Table 2: Identification of functional reference points

Name	Entities	Function
S3	MME and SGSN	Enables user and bearer information exchange for inter 3GPP access network mobility in idle and/or active state (ETSI TS 129 274 [i.13]).
S4	S-GW and SGSN	Provides related control and mobility support between GPRS Core and the 3GPP Anchor function of Serving GW. In addition, if Direct Tunnel is not established, it provides the user plane tunnelling (ETSI TS 129 274 [i.13] and ETSI TS 129 281 [i.14]).
S5	S-GW and P GW	Provides support for functions for packet data services towards end users (ETSI TS 129 274 [i.13] and ETSI TS 129 281 [i.14]).
S6a	MME and HSS	Used to exchange the data related to the location of the mobile station and to the management of the subscriber (ETSI TS 129 272 [i.15]).
S6d	SGSN and HSS	Used to exchange the data related to the location of the mobile station and to the management of the subscriber (ETSI TS 129 272 [i.15]).
S11	MME and S-GW	Used to support mobility and bearer management between the MME and S-GW (ETSI TS 129 274 [i.13]).
Sgi	PDN GW - packet data networks	may be an operator external public or private packet data network or an intra operator packet data network, e.g. for provision of IMS services (ETSI TS 129 061 [i.16]).

Table 3 lists the reference points standardized in the 3GPP Policy and charging control (PCC) architecture [i.3]. They are shown with their 3GPP names, the entities that they connect as well as a short description of the functionalities and mechanisms they provide, with a reference to the 3GPP standard where they are specified.

Table 3: Identification of PCC reference points

Name	Entities	Function
Rx	PCRF - AF	Allows for dynamic QoS and charging-related service information to be exchanged between the Policy and Charging Rules Function (PCRF) and the Application Function (AF). This information is used by the PCRF for the control of service data flows and IP bearer resources (ETSI TS 123 203 [i.3]).
Gx	PCEF - PCRF	Provides transfer of policy and charging rules from PCRF to PCEF in the P-GW (ETSI TS 129 212 [i.17]).
Gy	OCS - PCEF	Allows for the online charging based on the Diameter credit control application between the Online Charging System (OCS) and the PCEF (ETSI TS 132 251 [i.18] and IETF RFC 4006 [i.19]).
Gyn	OCS - TDF	Allows for the online charging based on the Diameter credit control application between the Online Charging System (OCS) and the TDF (ETSI TS 132 251 [i.18]).
Gz	OFCS - PCEF	Allows for the offline charging based on the charging record data transfer between the Offline Charging System (OFCS) and the PCEF (ETSI TS 132 295 [i.20]).
Gzn	OFCS - TDF	Allows for the offline charging based on the charging record data transfer between the Offline Charging System (OFCS) and the TDF (ETSI TS 132 295 [i.20]).
Np	RCAF - PCRF	Enables transport of RAN User Plane Congestion Information sent from the RCAF to the PCRF (ETSI TS 123 203 [i.3]).
Sp	SPR - PCRF	Allows for the PCRF to request subscription information from the SPR and the SPR to notify the PCRF when the subscription information has changed if the PCRF has requested such notifications.
Sd	TDF - PCRF	Enables a PCRF to have dynamic control over the application detection and control behaviour at a TDF (ETSI TS 129 212 [i.17]).
Sy	OCS - PCRF	Enables transfer of information relating to subscriber spending from OCS to PCRF (ETSI TS 129 219 [i.21]).

Table 4 lists the management plane interfaces identified in the 3GPP Telecommunication Management Architecture [i.22] that operate on top of the Core Network. They are shown with their 3GPP names, the entities that they connect as well as a short description of the functionalities and mechanisms they provide.

Table 4: Identification of Management plane interfaces

Name	Entities	Function
Type 1	NE - EM	Allows the use of certain management application layer protocols (CORBA IIOP, NETCONF, SNMP, SOAP).
Type 2 (Itf-N)	EM - NM	Used by the network and service management systems to transfer messages, notifications and service management requests via the EM to the NE.
Type 4a (Itf-P2P)	EM - EM	The approach for interfaces of type 4a is the same as for interfaces of type 2.

The interfaces that are most relevant for the mapping with the GANA model entities are the Itf-N, i.e. the Type 2 management interface and Itf-P2P, i.e. the Type 4a management interface, since they can be associated with the interfaces and reference points between node level DEs and network level DEs (defined in the GANA model), depending also on the implementation choices. Management application layer protocols used in 3GPP are based on SNMP, CORBA IIOP, SOAP (WS), REST.

Moreover, these reference points can also be taken into account when considering policies and the policy exchange and respective protocols. Actually, the policy based management and the respective protocols that are used for the policy exchange between Policy Decision Points (PDPs) and Policy Enforcement Points (PEPs) has been an integral aspect of the network management model in 3GPP. In order to address similar aspects in the GANA model, it is essential to address the potential reuse of the 3GPP protocols. Based on the respective use of the aforementioned protocols in several instantiations of the 3GPP reference management model, the mapping of these protocols to relevant GANA instantiations will be addressed. In order to summarize the main policy exchange protocols used in 3GPP models, an example instantiation of a Policy-based QoS Management System for 3GPP networks is illustrated in Figure 4.

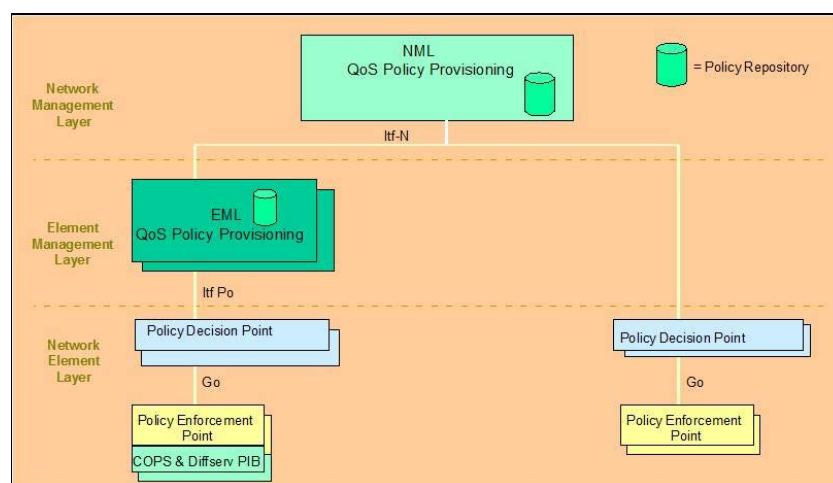


Figure 5: Conceptual architecture for a policy-based QoS Management System (reproduced from ETSI TS 132 101 [i.30])

In Figure 5 it is apparent how the policy continuum can be addressed, between PDPs and PEPs. In the different levels of policy exchange, there are currently several protocols used, some of them are of more general purpose than others, but all contribute to the enablement of different implementations for the policy exchange.

On the other hand, in a 3GPP network architecture the functions that are primarily related to policies and policy exchange are PCEF and PCRF [i.25]. The Gx reference point is located between the PCRF and the PCEF. The Gx reference point is used for provisioning and removal of PCC rules from the PCRF to the PCEF and the transmission of traffic plane events from the PCEF to the PCRF. The Gx reference point can be used for charging control, policy control or both by applying AVPs relevant to the application. The distribution of PDPs/PEPs and the GANA policy derivation and delivery in general should be seen in accordance with these architectural entities and interfaces and the respective policy exchange protocols used therein i.e. COPS (Common Open Policy Service) protocol, BGP (Border Gateway Protocol) and Diameter.

Particularly for the latter, a list of interfaces is defined in 3GPP LTE architecture. They are based on Diameter and are applicable in the present document. These interfaces are for example: S6a-interface - Authentication (specified in ETSI TS 129 272 [i.15]); Gy Reference Point - Prepaid charging (specified in ETSI TS 123 203 [i.3]); Gz Reference Point - Postpaid charging; Gx Reference Point - QoS/Policy (specified in ETSI TS 129 211 [i.27]); Rf Reference Point and Ro Reference Point - Charging (both specified in ETSI TS 132 299 [i.28]); Rx Reference Point - QoS/Policy (specified in ETSI TS 129 214 [i.29]); S6d-interface - Authentication (specified in ETSI TS 129 272 [i.15]); S9-interface - QoS/Policy (specified in ETSI TS 129 215 [i.36]); Sh Reference Point and Cx Reference Point - Subscriber Profile (specified in ETSI TS 129 228 [i.37]) and e2 Reference Point - Physical location (specified in ETSI ES 283 035 [i.38]). These interfaces are also defined in ETSI TS 123 002 [i.1].

5.3 Assumptions for the present document

5.3.1 Assumptions for the EPC

As shown in the previous clause, the EPC includes an extensive number of nodes and functionalities, which cover a large extent of use cases. In order to focus on its main features, some assumptions have been made to avoid complicating the mapping work and its presentation. The remainder of the present document focuses on specific use cases, which comply with the following assumptions.

- 3GPP EPC architecture to consider
 - Non-Roaming: the simplest EPC architecture case
 - Include PCC architecture (PCRF, PCEF)

Rationale: Considering both the EPC and the PCC architectures provides a good coverage of the main functions to be managed in the Core Network. More elaborated architecture models, including the roaming cases, would be an extension of the GANA functionalities provided by the entities defined in the present mapping.

- EPC interactions with 2G/3G PS Core nodes
 - Consider 2G, 3G (GANA Knowledge Plane (KP) only), 4G (GANA KP + DEs)
 - 3GPP RAT mobility + QoS and their impact

Rationale: The present document addresses the latest version of Mobile Network Operator (MNO) networks, i.e. the LTE-Advanced release of 3GPP. However, since 2G and 3G accesses are part of these networks, they have been included in the study, focusing mainly on their possible impact on the KP, e.g. decisions made in the KP about Radio Access Technology (RAT) mobility or QoS control for example. Taking into account this impact implies a close collaboration of the EPC GANA entities with the centralized Self-Organizing Network (C-SON) and the RAN nodes.

- Non-3GPP access (e.g. WLAN)
 - Focus only on the CN (not the Access Network, or AN), impact of non 3GPP Access on the CN (decision placement, interaction/collaboration/coordination between DEs, policies, etc.)

Rationale: Same as for 2G and 3G RAN, the non-3GPP access is included in the study mainly for its impact on the decisions made at the CN level.

- Consider all 3 different views
 - User/Data plane view
 - Control plane view
 - Management plane view

Rationale: With the objective to be comprehensive, the present document considers the three planes that are involved in the operation of a CN.

5.3.2 Assumptions for the backhaul

The backhaul network is responsible for the forwarding of the messages between the different entities of the MNO. It may be constituted of a very varied set of peripheral nodes, operating either at layer 2 or layer 3 (network layer) of the OSI model. The present document focuses on a restricted scenario, the results provided in the next clauses being extendable in a similar fashion to more complex scenarios. This scenario implies the following assumptions:

- Mobile backhaul network based on IP/MPLS enabled (P, PE) routers
- No Convergence Architecture Scenario ("Current Scenario" in NGMN document [i.9])
- With 3 independent components : Access Network, Backhaul, Core Network
- RAN OSS/NMS/EMS is involved only for its impact on the CN Management plane

6 Instantiation of GANA reference model to 3GPP EPC architecture and basic functions

6.1 Introduction

The present clause proposes an instantiation of the GANA reference model to the EPC. The analysis is split between the phases of the life-cycle of EPC, which leave room for simplification through automation, namely Self-Configuration/Commissioning (SC), Self-Optimization (SO) and Self-Healing (SH). All the relevant GANA DEs are identified and a mapping to the EPC core Network Elements (NEs) is proposed.

Based on the GANA reference model, an instantiation of specific GANA DEs and GANA Knowledge Plane Functional Blocks (FBs) towards an autonomic EPC is attempted in the sequel.

6.2 Self-Configuration/Commissioning

6.2.1 General description and use cases

The tasks involved in the Self-configuration [i.4] cover the initial and final life cycle of the Network Elements (NEs), as well as the specific cases of updates and upgrades of these NEs. Firstly, these operations manage the dynamic set-up of newly installed Nes from the point of view of their internal pieces of executable programs (software and firmware), and their functional insertion in the network. At the end of the Self-configuration, the NE is able to execute all its planned duties in the network. Secondly, they manage the dynamic update, upgrade or the removal of the NE from the network. Finally, the Self-configuration also covers the storage of the configuration data locally in the NE (as described in ETSI TS 123 401 [i.2], clause 5.7) and/or in the management plane.

The list of protocols and mechanisms involved in these tasks depend on the type of the NE. They are presented in the clauses 6.3.2.

The subsequent clauses focus on the addition of a NE in the mobile operator network and its initial setup. The other operations, such as deletion, update and upgrade are covered indirectly as the mapping to the GANA DEs would apply in a similar manner in these cases.

6.3.2 Protocol-level entities (GANA Level 1) to be managed (orchestrated dynamically and adaptively (re)-configured to adapt to changes)

Table 5 shows the operations related to the Self-configuration use case and identifies the entities of the CN and the backhaul network to which they apply. These operations are mainly related to the installation of a new node or a new function in a server and to the recording of configuration data in the node and in the management entities storage.

Table 5: Configuration-related use cases for the mapping

Use cases	Operation	Network Entity								
		IP/MPLS Router	MME	SGSN	S-GW	P-GW	HSS/EIR	PCRF	TDF	OCS
Node/function installation in the system	Addition/deletion/update/upgrade	*	*		*	*	*	*	*	*
Knowledge information storage	Store the configuration data of the node		*	*	*	*	*			

Some examples of the various protocols and mechanisms (MEs) involved in the operation of Self-configuration are described below, classified into a smaller set of generic high-level mechanisms. Examples of these mechanisms are given in the mapping tables of clause 6.3.4. The management aspects that need to be automated and the methods that can be applied autonomically (dynamically and adaptively to configure the MEs according to policy changes, context and operational environment changes and any faulty conditions) are as follows:

- Network Protocols Configuration
 - For the configuration of backhaul routers: subnets definition, configuration of routing protocols, routing table convergence, NETCONF invocations.
 - For the allocation of IP address and/or port number: DHCP server discovery, IP address and default gateway acquisition, Ipv6 Neighbour Discovery, TCP/UDP port number allocation and discovery.
- Security
 - Access control mechanisms, certificate passwords mechanisms, encryption key dissemination mechanisms.
- Connectivity
 - To management plane entities: OAM sub-system discovery (IP address and port number, network management protocol).
 - To neighbouring entities: TCP/IP protocols, layer 2 protocols, physical layer protocols.
- Firmware and Software Management
 - Software update and dissemination mechanisms, firmware update and dissemination mechanisms.
- CN Functional Configuration
 - Parameters required for the operation of the CN mechanisms and protocols, e.g. S1-AP, GTP-C, GTP-U, and Diameter applications. Examples of such parameters are user profiles, weight factors, APN, etc.
- Configuration Information Storage
 - Network management configuration data collection mechanisms, inventory systems mechanisms, network element capabilities descriptions

6.3.3 Identification of GANA DEs for Self-configuration

The present clause lists the key GANA DEs that should orchestrate the autonomy for the Self-configuration use case. This list has been built using Figure 17 and Table 1 of ETSI GS AFI 002 [i.8] as a reference. Each DE is explained with the type of Managed Entities (MEs) under its control as of the GANA architecture. The logic of a DE should keep monitoring the behaviour and state of the respective MEs it orchestrates and manages, such that in a closed control-loop structure (the heart of autonomy) it dynamically adjusts and re-configures the MEs or their parameters to meet objectives (which may change when policies get changed by the network operator). More complementary details that help developers of DEs, concerning the behaviour of a DE during initiation and its lifetime, are found in the GANA white paper [i.6] and ETSI GS AFI 002 [i.8].

Function level DEs

- Mobility-Management-DE [FUNC_LEVEL_MOM_DE] - Typical MEs managed by this DE are mobility management mechanisms and protocols.
- Service-Management-DE [FUNC_LEVEL_SM_DE] - Typical MEs managed by this DE are services and applications.
- Routing-Management-DE [FUNC_LEVEL_RM_DE] - Typical MEs managed by this DE are routing protocols and mechanisms.
- Generalized-Control-Plane-DE [FUNC_LEVEL_GCP_M_DE] - Typical MEs managed by this DE are the other control plane protocols that are not classified (by abstraction) as routing protocols.

Node level DEs

- AutoDiscoveryAndAutoConfiguration-DE [NODE_LEVEL_AC_DE] - Typical MEs managed by this DE are discovery protocols and mechanisms for the neighbours and the network.
- Security-Management-DE [NODE_LEVEL_SEC_M_DE] - Typical MEs managed by this DE are security protocols, algorithms and mechanisms operating at node level.
- Resilience-and-Survivability-DE [NODE_LEVEL_RS_DE] - Typical MEs managed by this DE are proactive and resilience mechanisms or restoration and protection mechanisms

Network level DEs (key ones)

- AutoDiscoveryAndAutoConfiguration-DE [NET_LEVEL_AC_DE] - Typical MEs managed by this are discovery protocols and mechanisms operating at network level.
- Mobility-Management-DE [NET_LEVEL_MOM_DE] - Typical MEs are managed by this mobility management mechanisms and protocols running at network level.
- Security-Management-DE [NET_LEVEL_SEC_M_DE] - Typical MEs are managed by this security protocols, algorithms and mechanisms operating at network level.
- Resilience-and-Survivability-DE [NET_LEVEL_RS_DE] - Typical MEs are managed by this proactive and resilience mechanisms or restoration and protection mechanisms running at network level.

Other Network Level DEs that also need to be initialized and activated in self-configuration phase, such that they configure their MEs and participate in self-configuration phase.

- Network-Level-Data Plane and Forwarding Management-DE [NET_LEVEL_DPM_DE]
- Network-Level-Routing Management-DE [NET_LEVEL_RM_DE]
- Network-Level-Quality of Services Management-DE [NET_LEVEL_QoS_M_DE]
- Network-Level-Monitoring Management-DE [NET_LEVEL_MON_DE]

6.3.4 Mapping of the key Self-configuration functions

6.3.4.1 MME

The present clause and the following ones introduce a mapping of some specific protocols and mechanisms in the case of the addition of a node or a function, which is a typical case for Self-configuration. They define this mapping for these three different types of Nes: a node (e.g. the MME), a function installed in a pre-configured server (e.g. the HSS) and a backhaul router.

Table 6 provides a one to one mapping between the instantiated DEs and the key managed protocols and mechanisms for Self-configuration in the case of an MME NE. The MME involves node mechanisms as well as several embedded functions, so this is the most complex example of mapping for the Self-configuration operation. The same would apply to an S-GW or a P-GW for example.

Table 6 is built on the model of Table 1 in the ETSI GS AFI 002 [i.8]. It shows on the right the protocols and mechanisms that are involved in the self-configuration of an MME, as well as some examples of these mechanisms (further right). It also shows on the three left columns which are the DEs which could serve as orchestrating entities for these Mes. Tables built on the same model can be found in subsequent clauses of clause 6.

Table 6: DE-to-ME Mapping table for the MME

Network-Level DEs	Node-Level DEs	Function Level DEs	Protocols & Mechanisms that need to be orchestrated and dynamically managed by their respective DEs	Examples
NET_LEVEL_SEC_M_DE	NODE_LEVEL_SEC_M_DE		Authentication of entity	Access control mechanisms, certificate passwords mechanisms, encryption key dissemination mechanisms
NET_LEVEL_AC_DE	NODE_LEVEL_AC_DE		Allocation of its own IP address and port number	DHCP server discovery, IP address and default gateway acquisition, Ipv6 Neighbour Discovery, TCP/UDP port number discovery and allocation
NET_LEVEL_MOM_DE		FUNC_LEVEL_MOM_DE	Connectivity to management plane entities	OAM sub-system discovery (IP address and port number, network management protocol)
NET_LEVEL_RS_DE	NODE_LEVEL_RS_DE		Firmware and software levels management	software update and dissemination mechanisms, firmware update and dissemination mechanisms
NET_LEVEL_AC_DE	NODE_LEVEL_AC_DE		Functional configuration in the CN	MME parameters configuration, e.g. weight factor for MME pooling (see clause 6.4.5.3)
		FUNC_LEVEL_MOM_DE	Connectivity to neighbouring entities: E-UTRAN, SGSN, HSS, S-GW	Activation of the links underlying the Core Network protocols
		FUNC_LEVEL_GCP_DE	Activation of the CN protocols with neighbour entities	S1-AP, GTP-C, and Diameter applications
		FUNC_LEVEL_SM_DE	Upload of configuration information in the management plane entities (e.g. MIB)	Network management configuration data collection mechanisms, inventory systems mechanisms
NET_LEVEL_MON_DE		FUNC_LEVEL_MON_DE	Configuration of monitoring tools and mechanisms that need to disseminate monitoring data to DEs	Configuration of the different monitoring data that need to be retrieved by DEs within the node
NET_LEVEL_FM_DE		FUNC_LEVEL_FM_DE	Configuration of the alarms scheme	Threshold trigger to send alarms alarm aggregation to send meaningful alarms

Figure 6 illustrates the mapping in the case of an MME NE. Figure 6 shows the three lower levels of the GANA model that may be instantiated in a NE. In the case of the MME, the DEs are complemented with a local information/knowledge repository that stores the information necessary to the operation of the node [i.2]. The grey-scale components represent the inactive components.

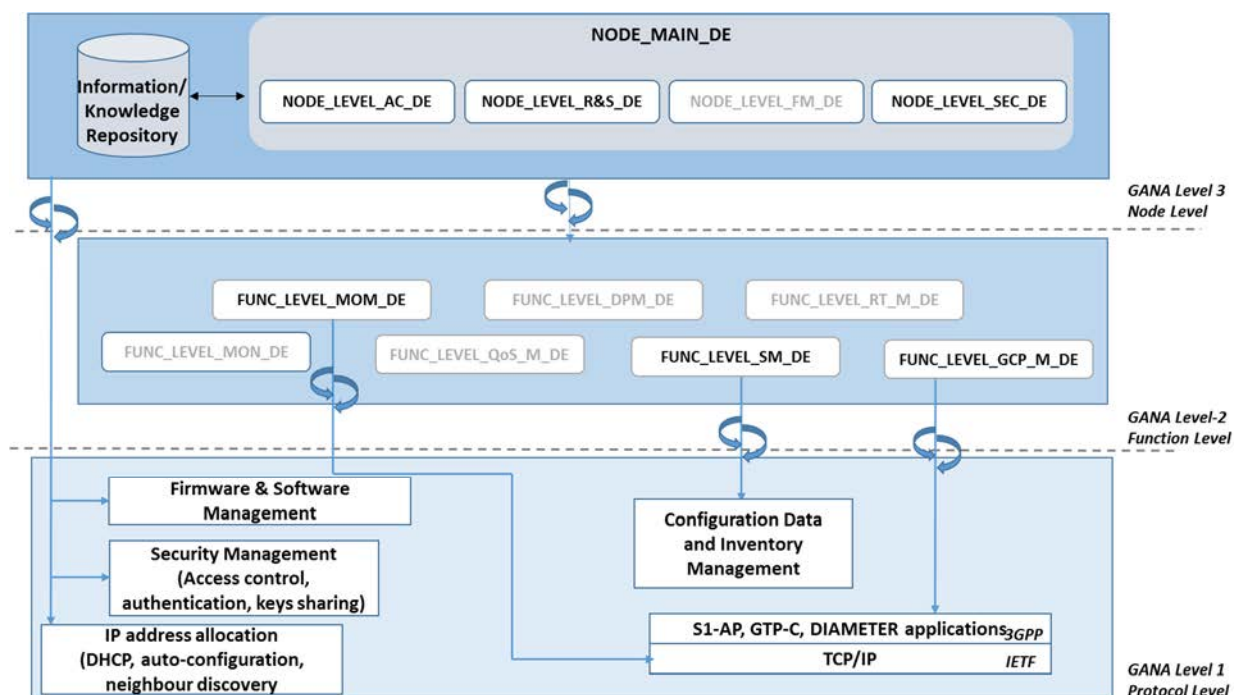


Figure 6: Node level view of Self-configuration case for the MME

Regarding implementation options for GANA DEs internal to an NE (network node), the GANA White Paper [i.6] discusses the options that implementers may take in choosing between implementing DEs in an NE as standalone processes or combined together as a single process or as executable behavioural models. While DEs and their algorithms that are meant to be implemented and run in a NE (node) can be designed and simulated individually for each DE and its interactions with its Mes and with other DEs (particularly its upper DE), the actual final implementation of the run-time DE instances can take the different approaches [i.16]. This option applies to the mapping shown in Figure 6 and to all subsequent mappings shown in Figure 7, Figure 8, Figure 11, Figure 12 and Figure 13.

6.3.4.2 HSS

Table 7 provides a one to one mapping between the key instantiated DEs and the managed protocols and mechanisms for Self-configuration in the case of an HSS NE installed in a pre-configured server. The HSS involves only embedded software functionalities; so in this case, the mapping does not need to address the node configuration mechanisms and protocols.

Table 7: DE-to-ME Mapping table for the HSS

Network-Level DEs	Node-Level DEs	Function Level DEs	Protocols & Mechanisms that need to be orchestrated and dynamically managed by their respective DEs	Examples
NET_LEVEL_SEC_M_DE	NODE_LEVEL_SEC_M_DE		Authentication of HSS entity	Certificate passwords mechanisms,
NET_LEVEL_AC_DE	NODE_LEVEL_AC_DE		Allocation of port number	TCP/UDP port number discovery and allocation
NET_LEVEL_MOM_DE		FUNC_LEVEL_MOM_DE	Connectivity to management plane entities	OAM sub-system discovery (IP address and port number, network management protocol)
NET_LEVEL_RS_DE	NODE_LEVEL_RS_DE		Software levels management	software update and dissemination mechanisms, firmware update and dissemination mechanisms
		FUNC_LEVEL_SM_DE	Functional configuration of the CN parameters	Upload of user profiles, Diameter server setup
NET_LEVEL_MOM_DE		FUNC_LEVEL_MOM_DE	Connectivity to neighbouring entities: MME	SCTP
		FUNC_LEVEL_GCP_DE	CN protocols activation	S6a/S6d application initiation
		FUNC_LEVEL_SM_DE	Upload of configuration information in the management plane entities (e.g. MIB)	Network management configuration data collection mechanisms, inventory systems mechanisms

Figure 7 illustrates the mapping in the case of the HSS NE.

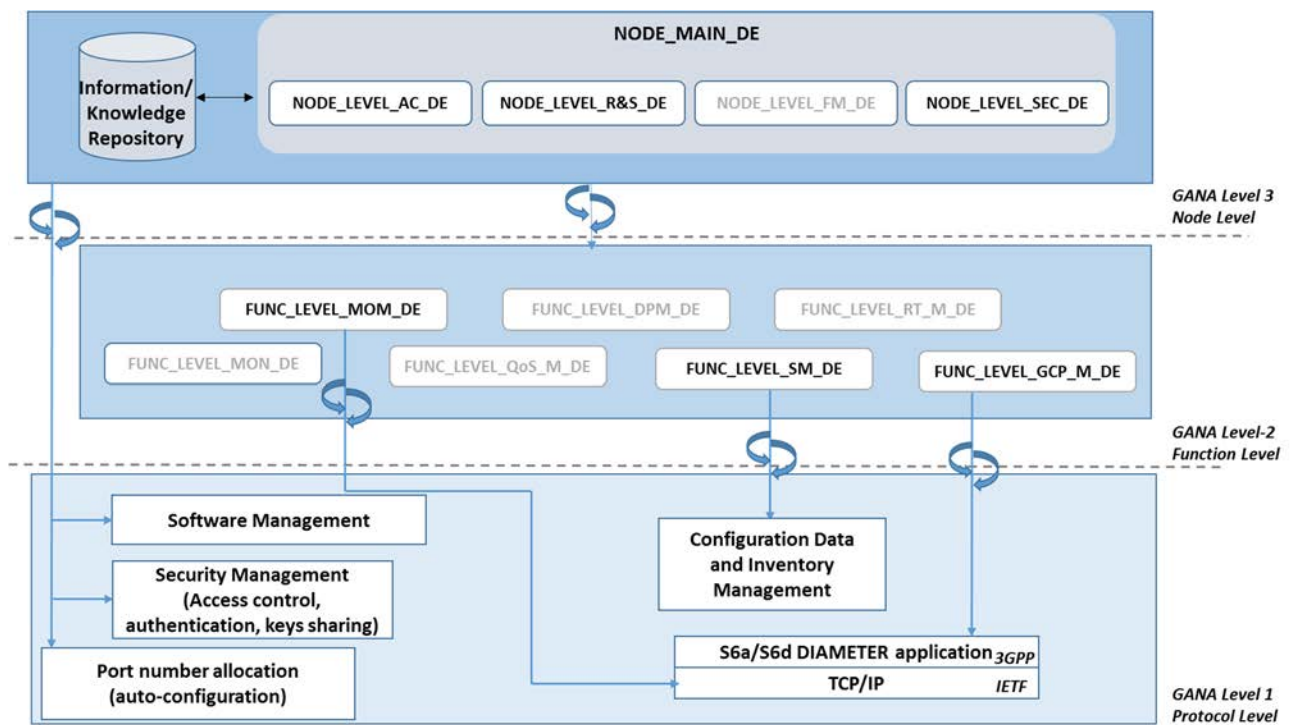


Figure 7: Node level view of Self-configuration case for the HSS

6.3.4.3 IP/MPLS router

Table 8 provides a one to one mapping between the instantiated DEs and the managed protocols and mechanisms for Self-configuration in the case of a backhaul router NE. The specificity in this case lies in the fact that the backhaul router acts only as a transport level node. It is not directly involved in the functional operation of the 3GPP Core Network, which restricts the configuration to be executed to the establishment of the node and of the routing protocols.

Table 8: DE-to-ME Mapping table for a router of the backhaul network

Network-Level DEs	Node-Level DEs	Function Level DEs	Protocols & Mechanisms that need to be orchestrated and dynamically managed by their respective DEs	Examples
NET_LEVEL_RT_M_DE		FUNC_LEVEL_RT_M_DE	Configuration of backhaul routers	Subnets definition, configuration of routing protocols, routing table convergence
NET_LEVEL_SEC_M_DE	NODE_LEVEL_SEC_M_DE		Authentication of entity (CN node/function)	Access control mechanisms, certificate passwords mechanisms, encryption key dissemination mechanisms
NET_LEVEL_AC_DE	NODE_LEVEL_AC_DE		Allocation of IP address and/or port number	IP address acquisition
NET_LEVEL_RS_DE	NODE_LEVEL_RS_DE		Firmware and software levels management	software update and dissemination mechanisms, firmware update and dissemination mechanisms

Figure 8 illustrates the mapping in the case of an IP/MPLS router NE.

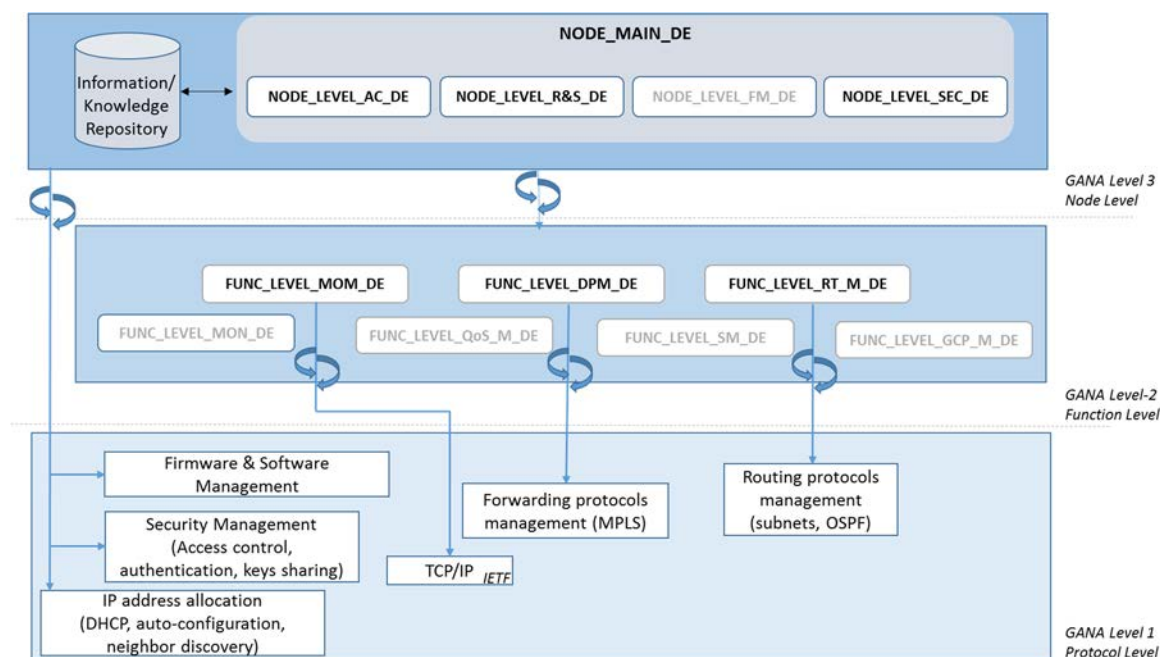


Figure 8: Node level view of Self-configuration case for a router of the backhaul network

6.4 Self-Optimization

6.4.1 General Description and use cases

The initial network configuration is typically made based on assumptions that deviate from the live operation. Changes in traffic volume and behaviour in time and space, newly deployed elements and other environmental variations render the initially configured parameters suboptimal, thus deteriorating network performance. Self-optimization caters for mechanisms that rely on measurements from the operating network in order to provide a much more accurate picture of the real conditions and reveal probable problems, which could not have been predicted under the initial assumptions, whereas they decide for actions that adapt the parameters in order to track those changes and eventually improve the performance of the network.

6.4.2 Protocol-level entities (GANA Level 1) to be managed (orchestrated dynamically and adaptively (re)-configured to adapt to changes)

The present clause lists examples of mechanisms and protocols (Mes) that can be seen as protocol-level entities (GANA Level-1 entities) to be managed by key GANA DEs for (self-) optimization purposes. Examples of these mechanisms are given in the mapping tables of clause 6.4.4. The management aspects that need to be automated and the methods that can be applied autonomically (dynamically and adaptively to configure the Mes according to policy changes, context and operational environment changes and any faulty conditions) are as follows:

- Mobility Management related
 - management of Tracking areas and of the reachability of Ues in the mobile service area
 - mechanisms for supporting the mobility (handovers) among eNodeBs and among different 3GPP networks e.g. among 3G and 4G
- Network Management (O&M) related
 - Mechanisms are there to support O&M functions related to the EPS such as MME pooling management, Node Load (Re)Balancing, Offloading, Overload control

- Selection related
 - Mechanisms for selection and reselection of the network elements (e.g. MME, P-GW, S-GW, SGSN) that will constitute the network path which can optimally serve the Ues
- QoS related
 - Mechanisms that deal with the management (activation, medication, release) of bearers i.e. the main model for enabling QoS provisioning and differentiated treatment of traffic in the EPC
 - QoS decision policies and QoS enforcement actions taken into elements in terms of filtering, marking, shaping, gating control, etc.
 - Charging decision policies and enforcement
- Energy savings
 - Mechanisms for lowering the network element power consumption when for example traffic is lower e.g. by shutting down certain CPUs, interfaces
- Monitoring protocols, mechanisms and tools
 - Mechanisms and tools associated with monitoring of load/traffic behaviour in the network elements
 - Mechanisms and tools for network resource usage Accounting e.g. for charging (creation and propagation of charging data records and respective tools for analytics)
 - Deep Packet Inspection (DPI), analytics

6.4.3 Identification of GANA DEs for Self-Optimization

The present clause lists the GANA DEs that orchestrate the autonomicity for the Self-optimization use case. This list has been built using Figure 17 and Table 1 of ETSI GS AFI 002 [i.8] as a reference. Each DE needed for the Self-optimization is explained with the type of Managed Entities (MEs) under its control as of the GANA architecture. The logic of a DE should keep monitoring the behaviour and state of its respective MEs it orchestrates and manages, such that in a closed control-loop structure (the heart of autonomicity) it dynamically adjusts and re-configures the MEs or their parameters to meet objectives (which may change when policies get changed by the network operator). More complementary details, that help developers of DEs, concerning the behaviour of a DE during initiation and its lifetime, are found in [i.17] and ETSI GS AFI 002 [i.8].

NOTE: Therefore, a DE's dynamic behaviour and lifecycle may be associated with various self-* related features and behaviours, ranging from self-configuration, self-optimization, through to some self-healing, etc. (refer to ETSI TS 132 251 [i.18] for more insights on this subject).

Function level DEs

- Mobility-Management-DE [FUNC_LEVEL_MOM_DE] - Typical MEs associated to the DE are mobility management mechanisms and protocols.
- QoS-Management-DE [FUNC_LEVEL_QoS_M_DE] - Typical MEs associated to the DE are mechanisms and protocols associated with the QoS control and provisioning.
- Monitoring-DE [FUNC_LEVEL_MON_DE] - Typical MEs associated to the DE are mechanisms, protocols and tools for monitoring the status of the elements e.g. in terms of signalling/user load, CPU usage, etc. as well as disseminating the monitoring data to entities (local or remote) that may need the data.

Node level DEs

- AutoDiscoveryAndAutoConfiguration-DE [NODE_LEVEL_AC_DE] - Typical MEs are mechanisms for the selection of other nodes e.g. GWs, for OAM related mechanisms such as load balancing, or for saving energy at node level.

Network level DEs

- Mobility-Management-DE [NET_LEVEL_MOM_DE] - Typical MEs are mobility management mechanisms and protocols managed at network level perspective (scope). The algorithm at the network level should complement an algorithm introduced/implemented at the similar DE (MOM_DE) at GANA Function-Level (inside a GANA node).
- QoS-Management-DE [NET_LEVEL_QoS_M_DE] - Typical MEs are mechanisms and protocols associated with the QoS control and provisioning at the network level.
- Monitoring-DE [NET_LEVEL_MON_DE] - Typical MEs are mechanisms, protocols and tools for monitoring the status of the network e.g. in terms of signalling/user load in all the interfaces. The Network-Level Monitoring DE works in tandem with the Function-Level-DE in GANA nodes (in network elements (NEs)) to orchestrate/trigger and dynamically manage monitoring tools/sub-systems and mechanisms across the network, influencing the MEs to disseminate monitoring data to where it is needed. In a closed control-loop fashion they continue to monitor the monitoring MEs and adjusting them accordingly.

6.4.4 Mapping of Self-Optimization functions

6.4.4.1 MME

Table 9 provides a one to one mapping between the instantiated DEs and the managed protocols and mechanisms for Self-optimization in the case of MME NE. As expected, there exist plenty of mechanisms and protocols within MME, which are associated with mobility management i.e. the main function that an MME typically offers. Inherently, these could be assigned to FUNC_LEVEL_MOM_DEs, being responsible for Mobility Management. However, decisions such as the ones made e.g. for selection of nodes for achieving load balancing, etc. could be also assigned to Node Level DEs (NODE_LEVEL_AC_DE), since the managed mechanism affects the behaviour of the node (MME) as a whole.

Table 9: DE-to-ME Mapping table for MME (control plane)

Network-Level DEs	Node-Level DEs	Function Level DEs	Protocols & Mechanisms that need to be orchestrated and dynamically managed by their respective DEs	Examples
NET_LEVEL_MOM_DE		FUNC_LEVEL_MOM_DE	Mobility Management protocols & mechanisms	Reachability, Tracking Area list management, Inter CN node signalling for mobility between 3GPP access networks
NET_LEVEL_MOM_DE	NODE_LEVEL_AC_DE	FUNC_LEVEL_MOM_DE	Network Management (O&M) oriented mechanisms	MME pooling management, Load (Re)Balancing, Offloading, Overload control
NET_LEVEL_MOM_DE	NODE_LEVEL_AC_DE	FUNC_LEVEL_MOM_DE	Selection mechanisms	P-GW selection, S-GW selection, MME selection for handovers with MME change, SGSN selection for handovers to 2G or 3G 3GPP access networks
NET_LEVEL_QoS_M_DE		FUNC_LEVEL_QoS_M_DE	QoS protocols & mechanisms	Bearer management
	NODE_LEVEL_AC_DE		Energy saving mechanisms	Traffic-aware processing power consumption e.g. by shutting down certain CPUs, interfaces
NET_LEVEL_MON_DE		FUNC_LEVEL_MON_DE	Monitoring protocols, mechanisms and tools	Monitoring of load/traffic, Accounting, DPI, analytics

6.4.4.2 P-GW

Table 10: DE-to-ME Mapping sub-table for P-GW (data plane)

Network-Level DEs	Node-Level DEs	Function Level DEs	Protocols & Mechanisms that need to be orchestrated and dynamically managed by their respective DEs	Examples
NET_LEVEL_MOM_DE		FUNC_LEVEL_MOM_DE	Mobility Management protocols & Mechanisms	Mobility anchoring for inter-working with non-3GPP technologies
	NODE_LEVEL_AC_DE	FUNC_LEVEL_MOM_DE	Selection mechanisms	PCRF selection
NET_LEVEL_QoS_M_DE		FUNC_LEVEL_QoS_M_DE	QoS & Charging Protocols & Mechanisms	Traffic Detection Function, Traffic Shaping, Transport level packet marking in the UL and DL, gating control, Traffic Redirection, Flow based charging support OCS/TDF
NET_LEVEL_MON_DE		FUNC_LEVEL_MON_DE	Monitoring protocols, mechanisms and tools	Monitoring of load/traffic, Accounting, DPI, analytics
	NODE_LEVEL_AC_DE		Energy saving mechanisms	Traffic-aware processing power consumption e.g. by shutting down certain CPUs, interfaces

6.4.4.3 S-GW

Table 11: DE-to-ME Mapping sub-table for S-GW (user plane)

Network-Level DEs	Node-Level DEs	Function Level DEs	Protocols & Mechanisms that need to be orchestrated and dynamically managed by their respective DEs	Examples
NET_LEVEL_MOM_DE		FUNC_LEVEL_MOM_DE	Mobility Management protocols & Mechanisms	Mobility anchoring for inter-eNB Handovers (Local), Mobility anchoring for interworking with other 3GPP technologies, packet buffering
NET_LEVEL_QoS_M_DE		FUNC_LEVEL_QoS_M_DE	QoS & Charging Protocols & Mechanisms	Transport level packet marking in the uplink and the downlink, Charging support
NET_LEVEL_MON_DE		FUNC_LEVEL_MON_DE	Monitoring protocols, mechanisms and tools	Monitoring of load/traffic, DPI, analytics
	NODE_LEVEL_AC_DE		Energy saving mechanisms	Traffic-aware processing power consumption e.g. by shutting down certain CPUs, interfaces

6.4.5 Indicative example of instantiation - MME pooling management

6.4.5.1 General Description

This clause provides an example in order to get into some more details on the instantiation and mapping process described in the previous sub-clauses. The example relates to the automated/autonomic management of mobility management entity (MME) pooling in 3GPP EPC networks.

6.4.5.2 MME pooling revisiting

The connectivity between Core and Access network part in legacy 2G/3G cellular networks was defined as "one to many" hierarchical relationship i.e. each core network node connected to its own set of controllers in a fixed manner. In Release 5 of 3G/UMTS standards, a feature that allowed for more flexibility in the inter-connection between access and core nodes, thus breaking the usual network hierarchy, was added.

This feature is well known as MME Pooling or S1-Flex (S1 Flexibility). MME pooling defines an MME pool architecture that enables eNBs to access and be served by multiple MMEs, which are grouped in a structure known as a "pool."

According to ETSI TS 123 401 [i.2]:

- An MME Pool Area is defined as an area within which a UE may be served without need to change the serving MME.
- An MME Pool Area is served by one or more MMEs ("pool of MMEs") in parallel.
- MME Pool Areas are a collection of complete Tracking Areas (Tas).
- MME Pool Areas may overlap each other.

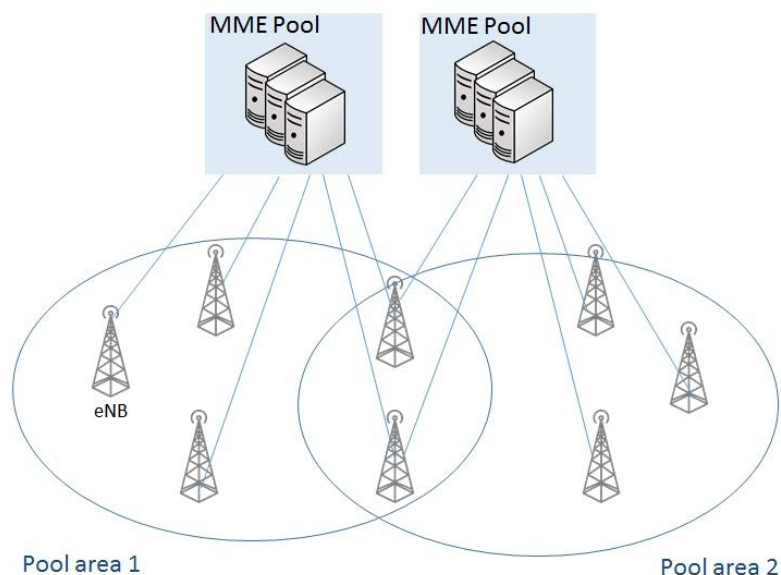


Figure 9: MME pooling concept

MME pooling is actually opening the opportunity to select the most appropriate MMEs (within an MME pool) to serve the UEs in the network based on different context and objectives, with the choice having a great impact on user service latency and general quality of experience. It can take place both while in initial configuration (S1 setup) but also during operation (optimization) and is mainly based on a controllable parameter defined as "weight factor" (e.g. ranging from 1 to 100). Each MME is assigned with such a weight factor that is conveyed to eNB during initial S1 setup. Based on the MME weight factors, an eNB can decide which MME in the pool can serve the calls and to what level. The probability of the eNB selecting an MME is proportional to its weight factor. The weight factor is typically set according to the capacity of an MME node relative to other MME nodes. The weight factor is sent from the MME to the eNodeB via S1-AP messages.

A numerical example follows. For instance, assuming that:

- an eNB is communicating with 2 MMEs in a pool
- MME1 has a weight factor of 100
- MME2 has a weight factor of 50
- 3 UE attachments are pending

In this case, the total sum of weights is 150 and the relative priorities per MME become $100/150 (=2/3)$ and $50/150 (=1/3)$ for MME1 and MME2, respectively. This practically means that in our example, the eNB will forward 2 out of 3 UEs to MME1 and 1 of 3 UEs to MME2.

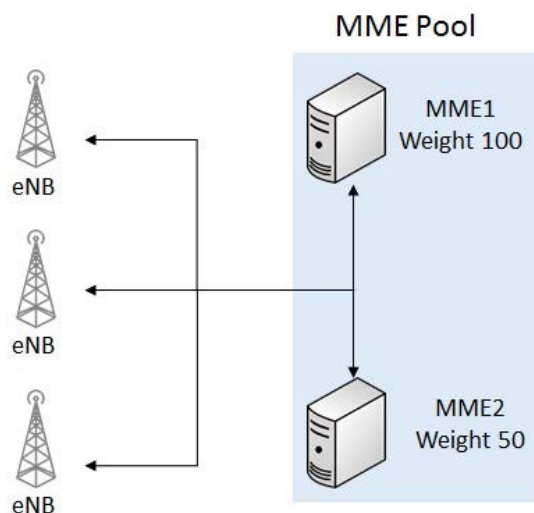


Figure 10: MME weight usage example

The weight factors (a.k.a. relative capacity) can be modified and enhanced by specific algorithms e.g. based on load measurements. These are reported to eNBs and the latter select the MME from the pool based on these updated values per each MME. In this way, the following mechanisms can be supported:

- **Load balancing** i.e. balancing the load among MMEs in the pool during initial configuration or after the introduction of a new MME and/or after removal of a MME from the network.
- **Load re-balancing** i.e. balancing the load among MMEs in the pool, for instance when an MME enters a maintenance/upgrade state.
- **UE offloading** i.e. moving certain (number of) UEs to less loaded MMEs.
- **Overload control** i.e. restricting any new signalling connections by MMEs due to overload conditions.

In general, optimal load distribution of the signalling traffic towards MME nodes can be achieved through autonomic management of the mechanisms above.

In the latter case, proper reconfiguration may be required:

- a) after the introduction of a new MME and/or after removal of a MME from the network,
- b) for balancing the load among MMEs in the pool,
- c) for re-balancing the load among MMEs e.g. when there is a need to offload traffic from an MME that may happen for instance when an MME enters a maintenance/upgrade state.

6.4.5.3 Hooks for GANA instantiation

6.4.5.3.1 Instantiation of DEs and indicative parameters that need to be dynamically managed and controlled by the DEs

In addition, load balancing can take place both in a centralized and in a distributed mode.

In the first case, a centralized management entity (NM/EM) has the knowledge of the whole picture of the network i.e. the MME statuses, and based on this knowledge, it can take decisions for (re)configuring the mechanisms. Network-Level DEs can play the role of enhanced NM/EM by involving cognitive features e.g. advanced monitoring and predictive models for predicting MME loads in certain periods and location and exploiting this new knowledge for proactively configuring the MME weights and in general for automating the whole process. Network-Level-DEs can be designed as modules that are loaded into the NMS/EMS to enhance management and control intelligence in such systems (more details in IETF RFC 4006 [i.19], [i.20] and [i.21]), or the DEs may be designed to run as standalone run-time entities that interact with NMS/EMS (via e.g. APIs) to drive them in writing transactions/configurations into the Network Elements (NEs) as is done on C-SON driving NMS/EMS via some APIs.

More distributed schemes can also apply. In this case, the responsibility of monitoring the MME loads and of accordingly setting weight factors, etc. is undertaken solely by each MME node and its DEs, of course by respecting the policies propagated by the operator.

In this case, decisions taken at the Node Level DEs e.g. by a `NODE_LEVEL_AC_DE` could be envisaged since the managed mechanism affects the behaviour of the node (MME) as a whole.

Alternatively, the mechanisms above (MEs) could be assigned to `FUNC_LEVEL_MOM_DE` being responsible for Mobility Management which is actually the main function the MME is logically offering, involving also a `FUNC_LEVEL_MON_DE` for managing monitoring and measurements associated with the mechanisms.

Last but not least, the distributed case requires coordination and interaction among MMEs (their DEs) belonging to the same pool.

By monitoring the MME relative capacities, load status, overload situations, etc., the involved DEs should be able to manage the mechanisms and the parameters that each of the mechanisms is actually managing an indicative set of which is listed in Table 12.

Table 12: DE-to-ME Mapping table for DEs instantiation in MME

Network Level DE	Node Level DE	Function Level DE	Managed Entity (Mechanism) that needs to be orchestrated and dynamically managed the respective DEs	Example parameter
NET_LEVEL_MOM_DE NET_LEVEL_MON_DE	NODE_LEVEL_AC_DE	FUNC_LEVEL_MOM_DE	MME Load (Re)Balancing	MME weight factor
		FUNC_LEVEL_MOM_DE	MME Offloading	Percentage of Ues to be offloaded
		FUNC_LEVEL_MOM_DE	MME Offloading	Duration of offloading
		FUNC_LEVEL_MOM_DE	MME Overload control	eNB from which to reject signalling connections
		FUNC_LEVEL_MOM_DE	MME Overload control	Types of RRC connections to reject (Overload Action)
		FUNC_LEVEL_MOM_DE	MME Overload control	Duration of signalling rejections
		FUNC_LEVEL_MON_DE	MME Load (Re)Balancing	Measurement parameters e.g. period, Predictions
		FUNC_LEVEL_MON_DE	Offloading	Measurement parameters e.g. period, Predictions
		FUNC_LEVEL_MON_DE	MME Overload control	Measurement parameters e.g. period, Predictions

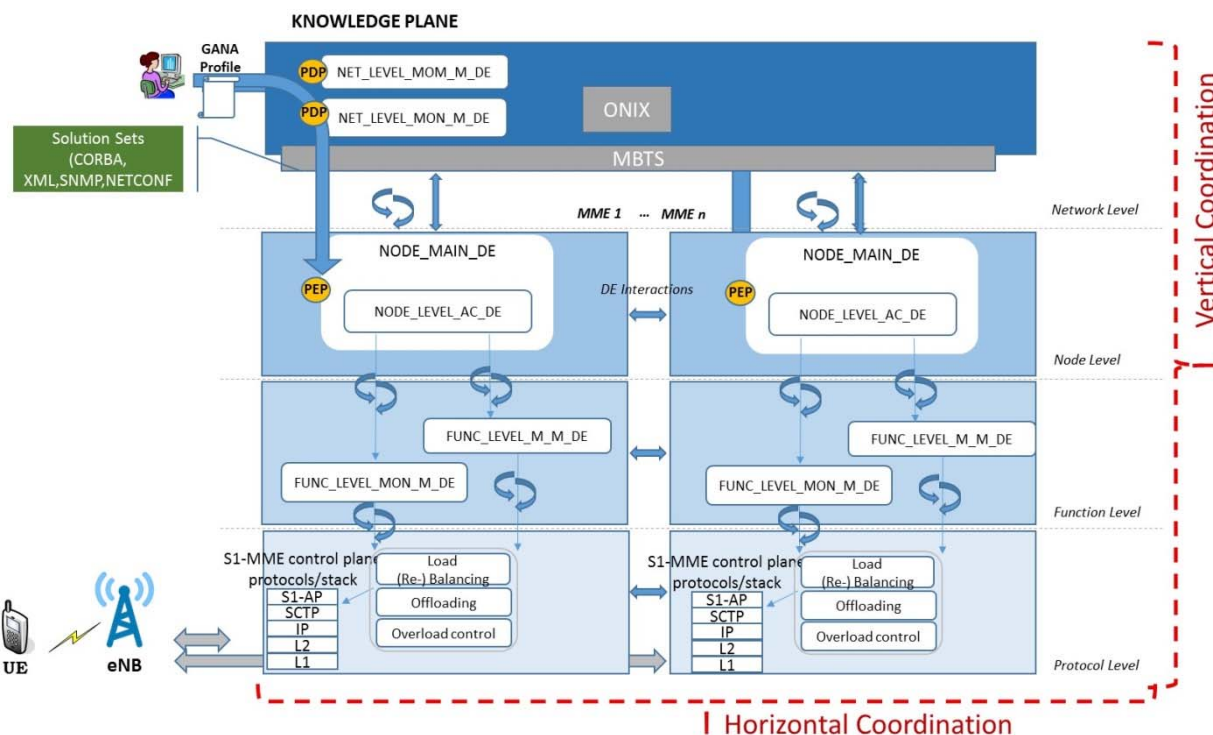


Figure 11: Instantiation of DEs and indicative parameters in the MME pooling management example

6.4.5.3.2 Policies distribution

Different policies may be decided and be enforced based on the operator objectives. The policies can be propagated from the network level DEs (policy decision points) down to the node and/or function-level-DEs (policy enforcement points) and be respected by the decisions being made therein. Examples could be as follows:

- Favouring the use of a specific MME among others belonging in the same pool.
- Controlling the frequency of weight factor changes.
- Newly installed MMEs can be assigned with bigger weight factor, to give it higher priority and thus making them faster to increase their load.
- Offloaded MMEs can be assigned with zero weight factor so as to prohibit new UEs to be allocated to this MME.
- The duration of offloading is also configurable and can be managed in an autonomic manner.
- The overload control could only apply to non-emergency and non-high priority mobile originated services.

6.5 Self-Healing

6.5.1 General Description and use cases

Self-healing is the process which detects problems (e.g. operational, HW/SW faults, etc.) and performs diagnosis of the problems (fault-diagnosis/location/isolation), performs repairing procedures for self-repair, or mitigates these to avoid impact on user services and overall health of network's systems and to significantly reduce maintenance costs.

Self-healing is triggered by alarms generated by the faulty network elements. If it finds alarms that it might be able to resolve and correct the problem or minimize the effects of, it gathers more necessary correlated information (e.g. measurements, testing results, and so forth), does deep analysis (fault-diagnosis/localization/isolation), and then triggers the appropriate actions. Self-Healing Decision making elements may incorporate similar mechanisms as employed in SON in order to reach its full functional capacity.

The two major areas where the self-healing concept could be applied are as follows:

- a) Self-diagnosis: create a model to diagnose faults or perform basic fault-diagnosis accordingly, learning from past experiences.
- b) Self-healing: automatically start the corrective actions to solve the problem.

In SON, self-healing is also defined but for the core network and backhaul self-healing takes a slightly different nature as explained above.

The process of self-healing can be divided into four main phases:

- 1) Detection - Where in the network a problem has appeared.
- 2) Diagnosis - Identification of the root cause of failure and if some healing is required. False alarms are eliminated.
- 3) Recovery planning - a set of the necessary recovery actions is selected.
- 4) Recovery execution - phase when the actions are taken to handle the particular type of the detected problem and its outcome is assessed.

The list below consists of the self-healing functions which will be executing when a fault symptom (alarm) is detected. In clause 6.5.5, there is an analytical flow-chart in which all the appropriate steps that will be executed in each case are depicted.

- Self-Healing Input Monitoring Function (SH_MON_F): This self-Healing function monitors the TCoSHs (Trigger Condition of Self-Healing) continuously and when a TCoSH is reached, then an appropriate Self-Healing process will be triggered.

- Self-Healing Diagnosis Function (SH_DG_F): Based on the TCoSH and gathered information, this Self-healing Function does deep analysis and diagnosis related to the root cause identification, and gives the result that will identify the failure and the potential action/s for recovery. If the result includes recovery action/s, then go to next step of the healing process part, if not, go to End (termination of process).
- Triggering Recovery Action/s Function (SH_TG_F): If the result from previous functions includes recovery action/s, then this Self-healing Function triggers the execution of the recovery action/s.
- Self-Healing Evaluating Function (SH_EV_F): This Self-healing Function evaluates the result of the self-healing recovery action/s. If the fault has not been solved and the stop condition/s is not reached, then the self-healing runs again. This process is followed when more faults have been diagnosed and need to be resolved. If the fault has been solved, then this function emits a notification to report the result.
- Self-Healing Fallback Function (SH_FB_F): This function is executed when the stop condition of the Self-Healing function is reached. In addition, if necessary, a fall-back situation is executed.
- Self-Healing Monitoring and Management Function (SH_MMF): This function monitors the self-healing process and provides the operator with the necessary information of the self-healing process. This function will be able to get information about all other functional blocks. In addition to this, it allows the operator to control the execution of the self-healing process.
- Self-Recovery of NE software Function (SR_NSW_F): This function handles the self-healing function of recovery of NE software.
- Self-Healing of Board Fault Function (SH_BF_F): This function handles the self-healing function for board fault.

There are three concrete Use Cases which are defined by 3GPP for the scope of the self-healing:

- a) Self-Recovery of Network Elements
 - [All the EPC components]
- b) Self-Healing of board faults
 - [All the EPC components]
- c) Self-Healing of Cell Outage (this use case refers to the RAN, therefore is out of scope to the present document due to the fact that the focus area is the EPC, not the E-UTRAN of LTE (which is already covered by SON).

6.5.2 Protocol-level entities (GANA Level 1) to be managed (orchestrated dynamically and adaptively (re)-configured to adapt to changes)

The Self-Healing process affects the HW and SW of a network entity in terms of fault and resilience management when the monitoring process detects TCoSH. This is a condition which is used to judge whether a Self-healing Process needs to be started. In addition, this condition could be an alarm or the detection of a fault in general. To this end, the respective 3GPP and non-3GPP protocols and mechanisms may be affected and healing actions may be performed on these managed entities (protocols and mechanisms).

6.5.3 Identification of GANA DEs for Self-Healing

This clause presents the key GANA DEs which are part of the self-healing use case. The abovementioned list has been built using Figure 19 of ETSI GS AFI 002 [i.8] as the generic reference structural model. In addition, each DE is linked with the type of Managed Entities under its control as of the GANA architecture overview. The logic of a DE should keep monitoring the behaviour and state of its respective Mes it orchestrates and manages, such that in a closed control-loop structure (the heart of autonomicity) it dynamically adjusts and re-configures the Mes or their parameters to meet objectives (which may change when policies get changed by the network operator). More complementary details that help developers of DEs, concerning the behaviour of a DE during initiation and its lifetime, are found in ETSI TS 132 102 [i.22] and ETSI GS AFI 002 [i.8].

NOTE: Therefore, a DE's dynamic behaviour and lifecycle may be associated with various self-* related features and behaviours, ranging from self-configuration, self-optimization, through to some self-healing, etc. (refer to 3GPP TR 23.705 [i.23] for more insights on this subject).

Function Level DEs

- Monitoring DE: Typical MEs are monitoring protocols/mechanisms which are responsible for monitoring the condition of each Network Element (NE).

Network Level DEs

- Fault-Management DE: Typical MEs are mechanisms or protocols which are responsible for firing an action when a fault alarm occurred, as well as various fault-detection methods that need to be orchestrated, fault-localization/isolation/diagnosis methods and fault-removal techniques/methods that need to be employed when a fault condition has been detected (e.g. by reception of an alarm).
- Resilience and Survivability DE: Typical MEs are proactive and resilience mechanisms or restoration and protection mechanisms.

6.5.4 Mapping of Self-Healing functions

6.5.4.1 Mapping of instantiated GANA DEs to protocols and mechanisms for MME and PGW

Table 12 and Table 13 provide a one to one mapping between the instantiated DEs and their managed protocols and mechanisms for MME and PGW. These entities have been initially chosen for the instantiation due to the key role they have in the 3GPPP CN.

Table 12: DE-to-ME Mapping sub-table for MME

Network-Level DEs	Node-Level DEs	Function Level DEs	Protocols & Mechanisms that need to be orchestrated and dynamically managed by their respective DEs	Examples
NET_LEVEL_RS_DE			3GPP protocols/mechanisms	Software Failure in the MME entity
NODE_LEVEL_FM_DE			No 3GPP protocols/mechanisms	Hardware Failure in the MME entity
	NODE_LEVEL_FM_DE		3GPP or no 3GPP protocol/mechanism	Software/Hardware Failure in the Node Level of MME
		FUNC_LEVEL_MON_DE	3GPP or no 3GPP protocol/mechanism	Monitoring Component

Table 13: DE-to-ME Mapping sub-table for PGW

Network-Level DEs	Node-Level DEs	Function Level DEs	Protocols & Mechanisms that need to be orchestrated and dynamically managed by their respective DEs	Examples
NET_LEVEL_RS_DE			3GPP protocols/mechanisms	Software Failure in the PGW entity
NODE_LEVEL_FM_DE			No 3GPP protocols/mechanisms	Hardware Failure in the PGW entity
	NODE_LEVEL_FM_DE		3GPP or no 3GPP protocol/mechanism	Software/Hardware Failure in the Node Level of PGW
		FUNC_LEVEL_MON_DE	3GPP or no 3GPP protocol/mechanism	Monitoring Component

6.5.4.2 MME

Figure 12 illustrates the mapping of GANA functions in the case of an MME network element. As mentioned below in the figure, the bolded "boxes", which represent the Function Level of the GANA model, are the active components in each GANA Level. The grey-scale components represent the inactive components.

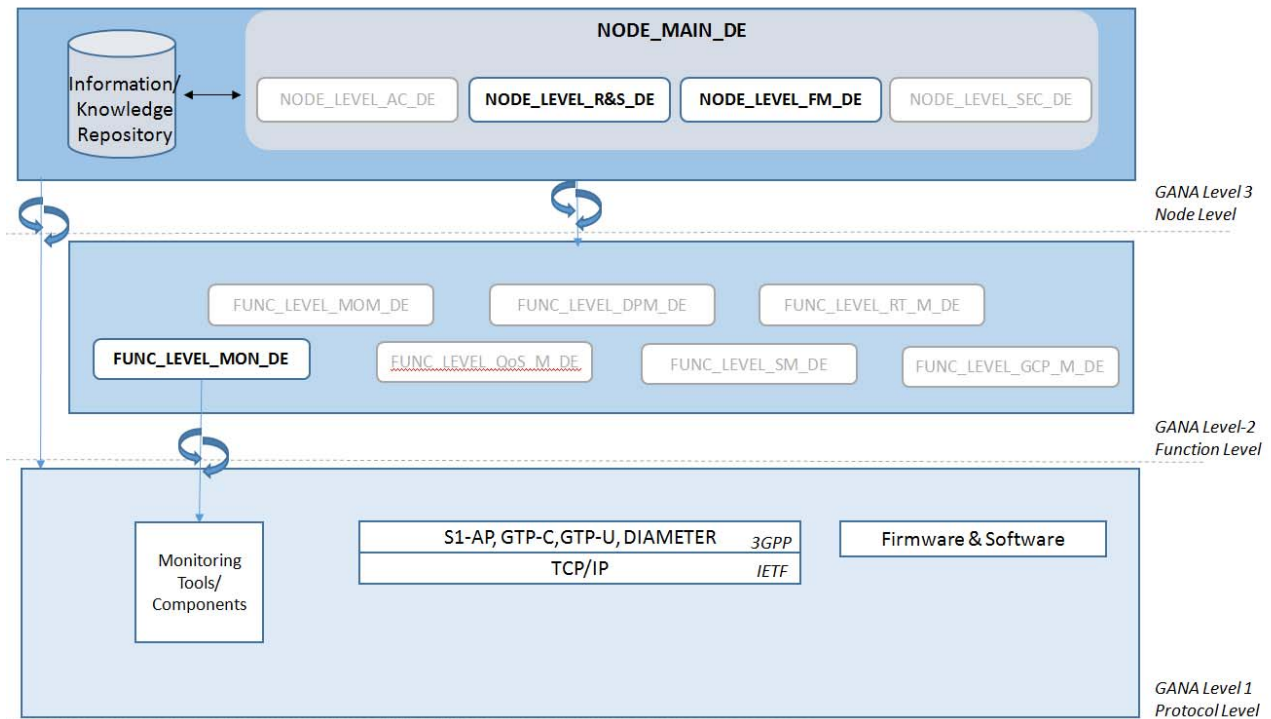


Figure 12: Node level view of self-healing case for the MME

6.5.4.3 P-GW

Figure 13 illustrates the mapping of GANA functions in the case of a PDN Gateway network element. As mentioned below in the figure, the bolded "boxes", which represent the Function Level of the GANA model, are the active components in each GANA Level. The grey-scale components represent the inactive components.

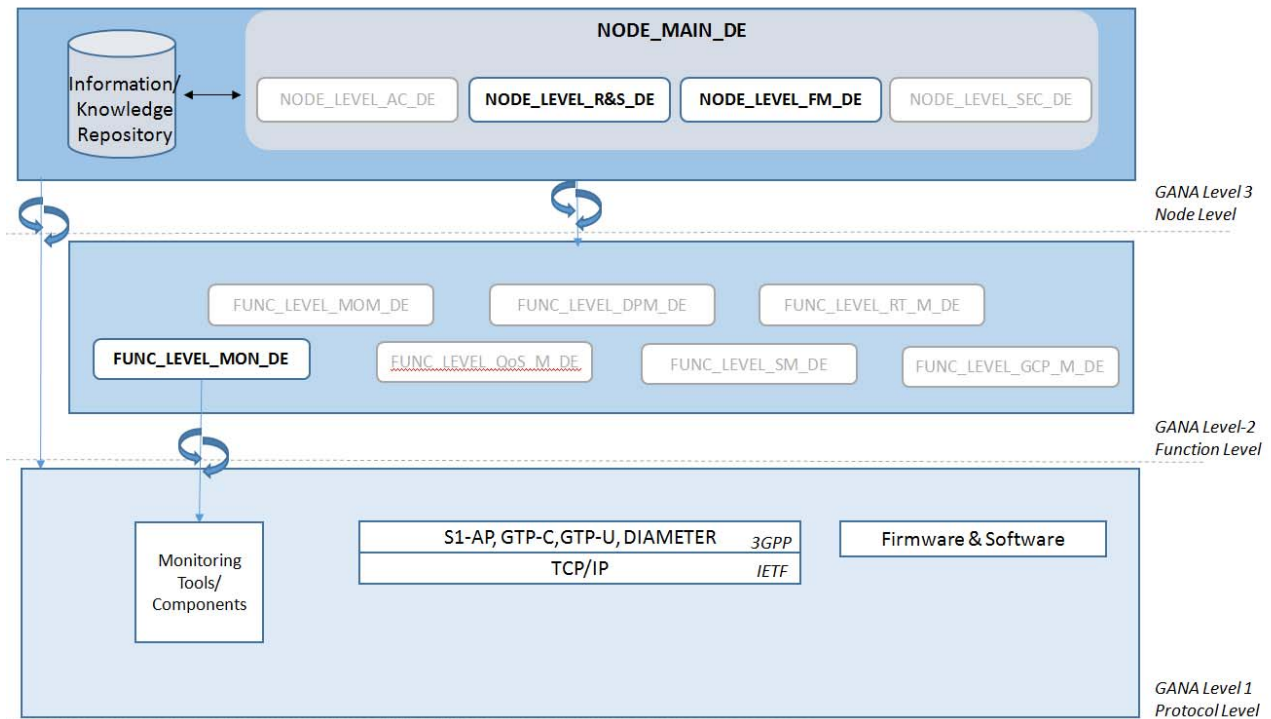


Figure 13: Node level view of self-healing case for the P-GW

6.5.5 Indicative examples of GANA instantiation - MME Self-Healing

As referred in the ETSI TS 132 541 [i.5], the self-healing process has two parts: the monitoring part and the healing process part. The first one monitors, through a control loop, all the time the network element in order to alarm the appropriate healing process if a TCoSH (Trigger Condition of Self-Healing) is reached. Afterthought, the healing process part, collects the necessary information, measure them and inform the appropriate self-healing processes. Below an initial example of how the GANA and the self-healing functions respond in a fault alert situation, e.g. in the MME entity of the LTE EPC, is listed.

1) Monitoring Part:

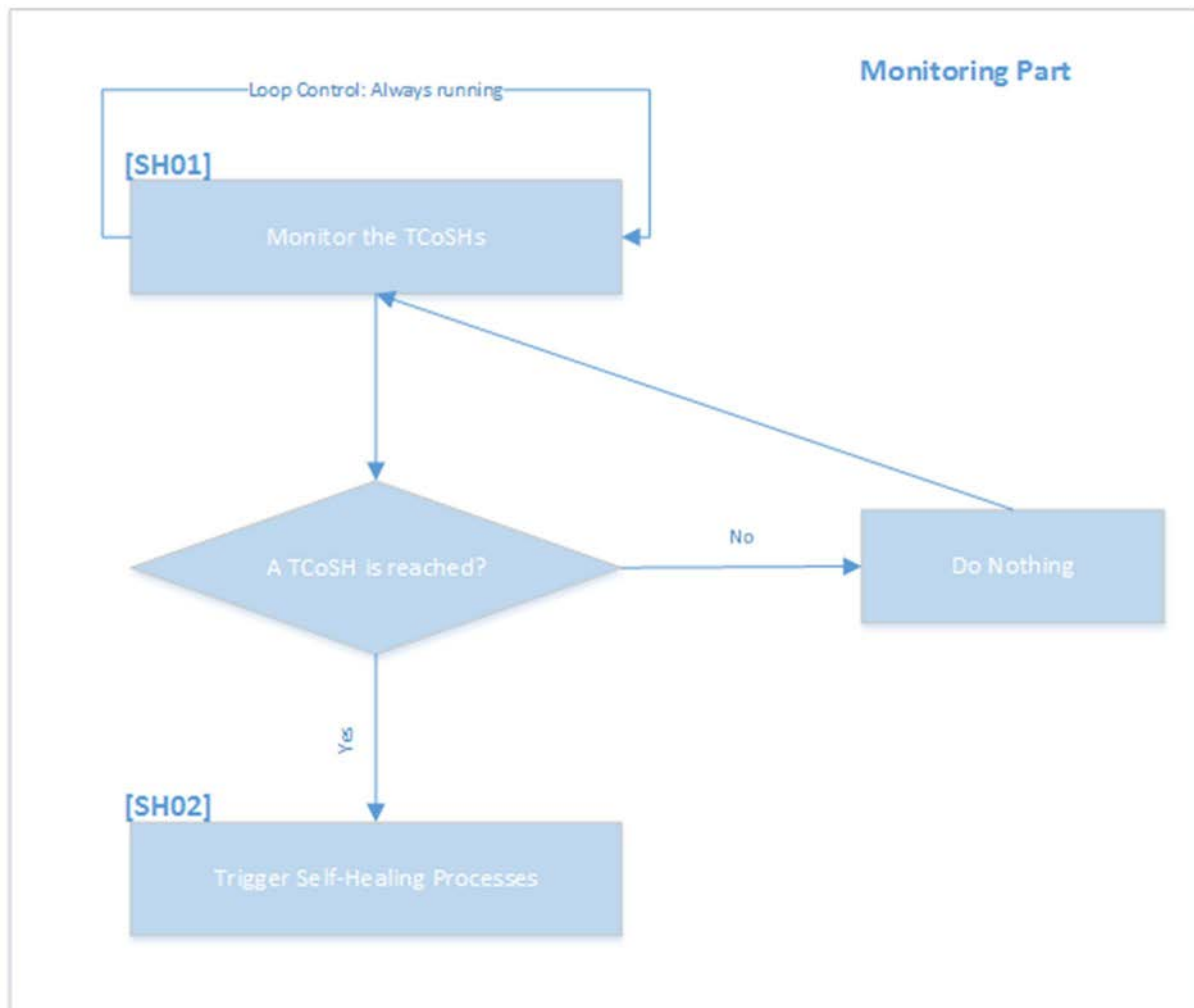


Figure 14: Self-healing monitoring part

[SH01] and [SH02] are corresponding to the Self-Healing functions 01 and 02 respectively. The [SH01] may collaborate with the GANA functions of MME, especially the **Monitoring Tools/Components**, in order to monitor the state of each NEs component.

2) Healing process Part:

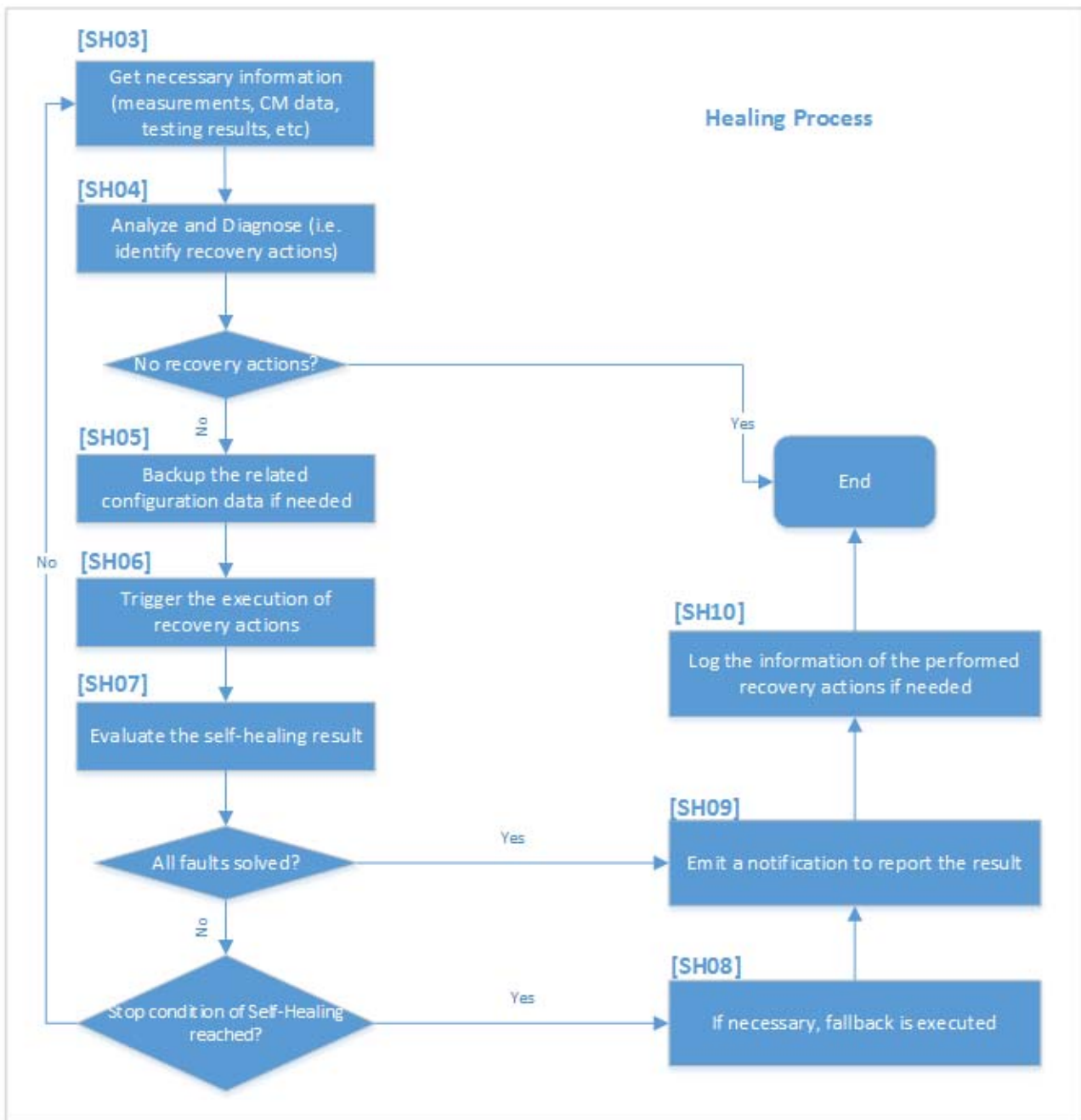


Figure 15: Self-healing processing part

The [SH03] is responsible for collecting all the necessary information, such as measurements, testing results, etc., in order to alert the appropriate self-healing function. The [SH03] may collaborate with the GANA functions of MME, the **FUNC_LEVEL_MOM_DE** and **FUNC_LEVEL_GCP_M_DE**, by collecting all the needed information through these two functions.

6.6 GANA Network Level view with the Knowledge Plane (KP)

6.6.1 Relationship between C-SON and GANA KP, and KP southbound interfaces instantiations for 3GPP Core

In Figure 16, the Network level view for the mapping of relevant functionalities is depicted (here, showing the DEs involved in the self-healing functions), with an emphasis to the GANA knowledge plane (KP).

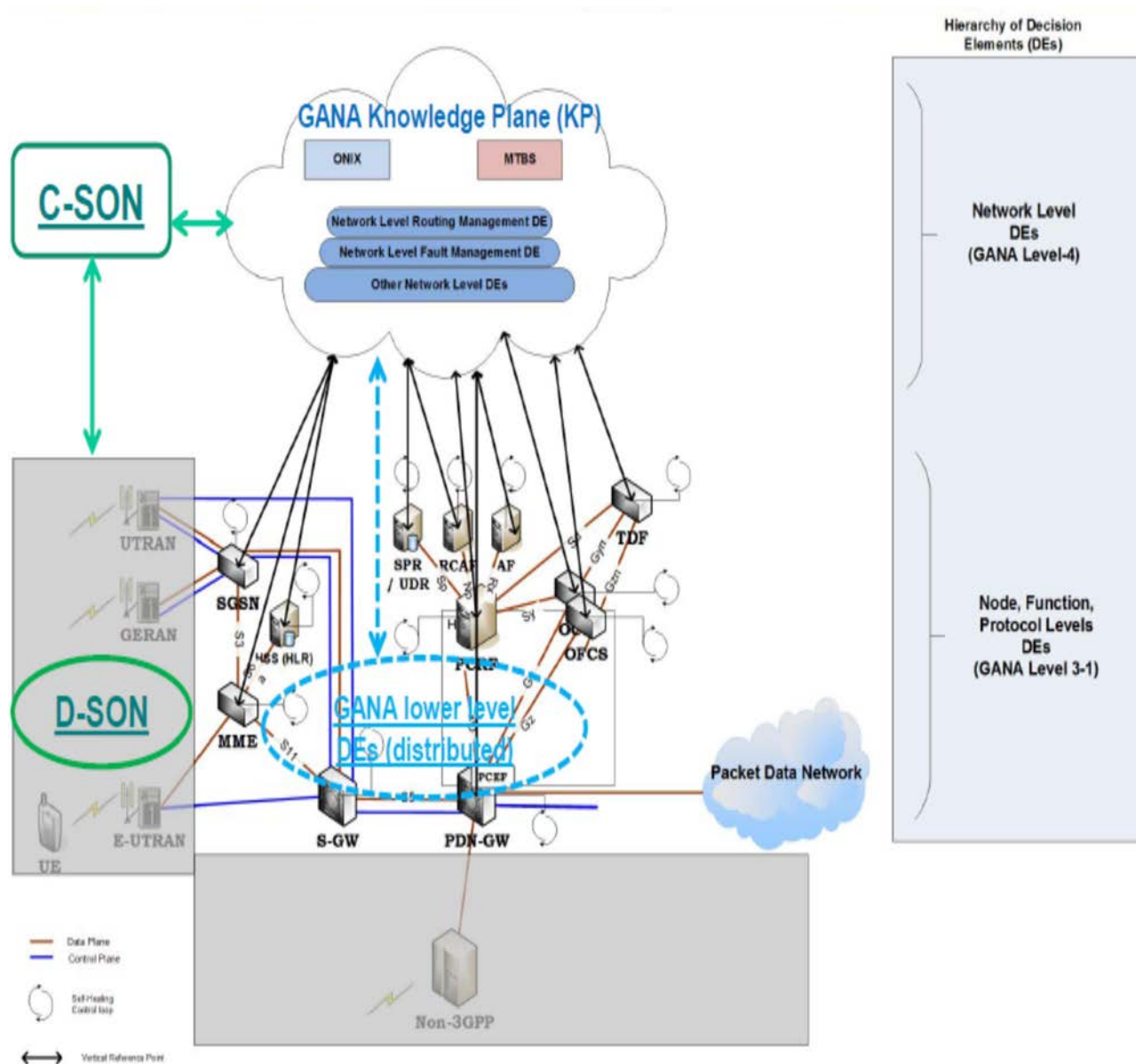


Figure 16: Instantiation of the GANA Knowledge Plane for the 3GPP Core Network

6.6.2 Instantiation of the GANA Network Governance Interface of the KP for the 3GPP Core Network

The GANA Network Governance Framework provides coherence of policies and synchronization between and across operator and administration domains. This is schematically depicted in Figure 17.

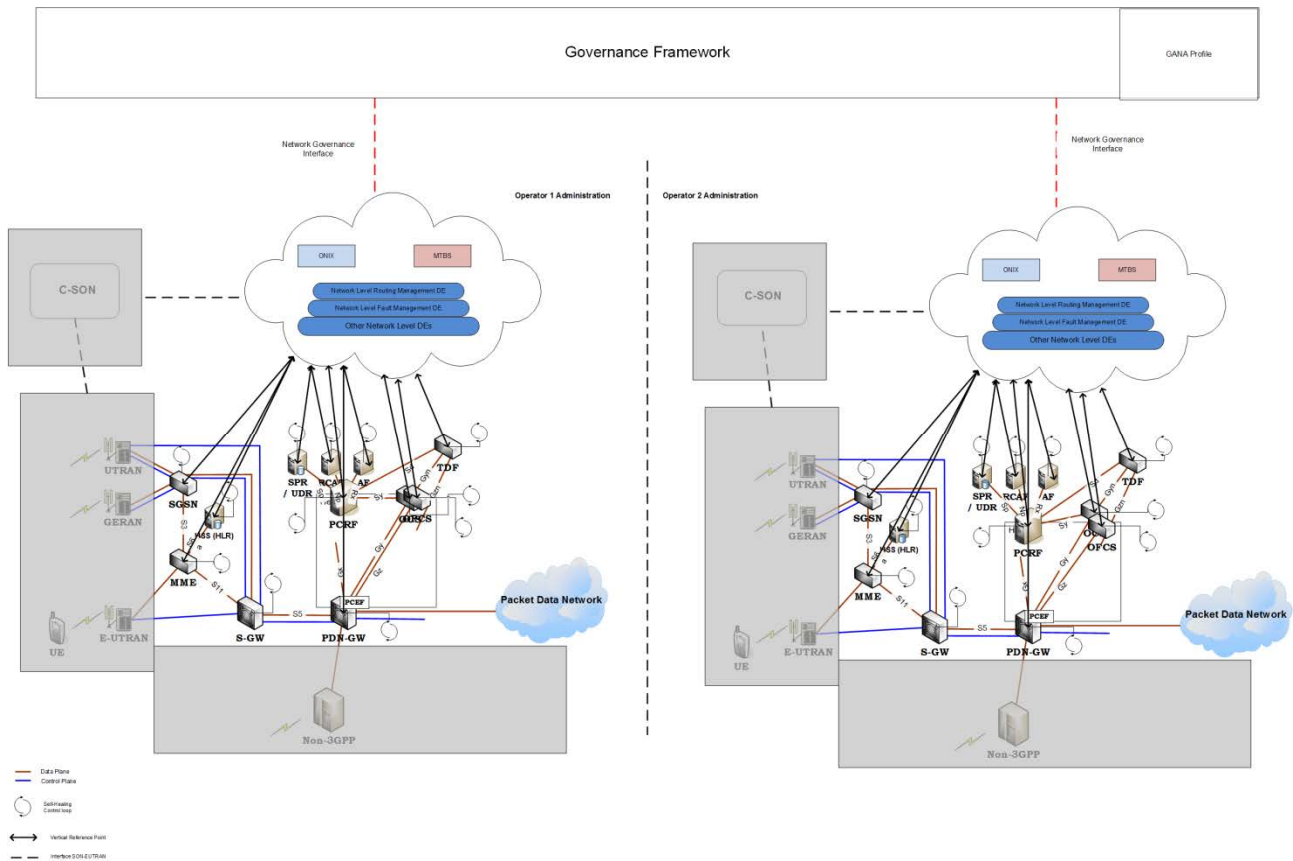


Figure 17: Instantiation of the GANA Network Governance Interface of the KP for the 3GPP Core Networks

An important issue for the governance framework and for the GANA Knowledge Plane is the policy management for enabling the autonomic network behaviour. In general, policies are very important for the operation and configuration of autonomic networks, since the policy is the rule defining a choice of the network behaviour. Policies have been widely used for the introduction of some level of automation in the control and configuration of network equipment operational behaviour. Usually they are based on predefined event-condition-action rules. They provide a deliberate plan of action for the decision enforcement and for the achievement of desired behaviours. Policy based behaviour of a network can be either static, or dynamic, depending on whether the policies can be updated automatically by the network systems, based on the evolution of network knowledge base. The application of either static or dynamic policies in autonomic systems may affect different decision levels. Conflict resolution mechanisms may be required as well.

Policies can be addressed as higher level policies (more abstract network view or business level) and low level policies (can be derived from system transformations). Moreover, policies can be also adaptive to new situations that demand changes in the network operation or configuration in order to accomplish its goals or to have a better performance. Policies also play a big role in interacting self-governing networks, for example for trust management between networks and domains and for achieving better operation between these networks from a wider network level perspective.

The GANA KP generates the GANA Profile(s) using automation and translation tools (which could have some translation capabilities that may be similar to those employed for the MBTS-meaning that the MBTS could potentially be used for extending the GANA Profile with additional information). The GANA Profile is stored in the ONIX and is then used by the Network Level DEs to configure themselves and issue commands and lower level policies that are provided to lower level DEs for enforcement (including relaying sub-profiles that are used by lower level DEs to configure themselves and their MEs). More details on the GANA Network Profile creations and use can be found in GANA White Paper [i.6] and ETSI GS AFI 002 [i.8]. The GANA Profile can be augmented/extended with run-time related information by the Network Level DEs and MBTS and stored/maintained in the ONIX.

7 Autonomic behaviours across multiple segments to enable End-to-End (E2E) Self-Optimization

7.1 Introduction

The second phase of the GANA instantiation on the 3GPP EPC consists in devising the information exchanges and interactions between the entities identified in the mapping, i.e. the entities inside the KP, the DEs at the three higher levels of GANA, and the MEs. Interactions also take place across the different domains involved in the end-to-end operation of the mobile network, spanning RAN, backhaul and EPC NEs.

At network level, this study defines how objectives and policies are transferred from PDPs to PEPs, how knowledge is shared and stored, which decisions are taken and where they are enforced. The study also covers interactions among GANA nodes instantiated in NEs located in the backhaul as well as in the EPC segments. When needed it expands to the RAN, covering for example autonomic decisions that are taken to optimize the EPC segment, but consider information retrieved from management functions responsible of the RAN, mainly referring to the C-SON.

These interactions are summarized in Figure 18.

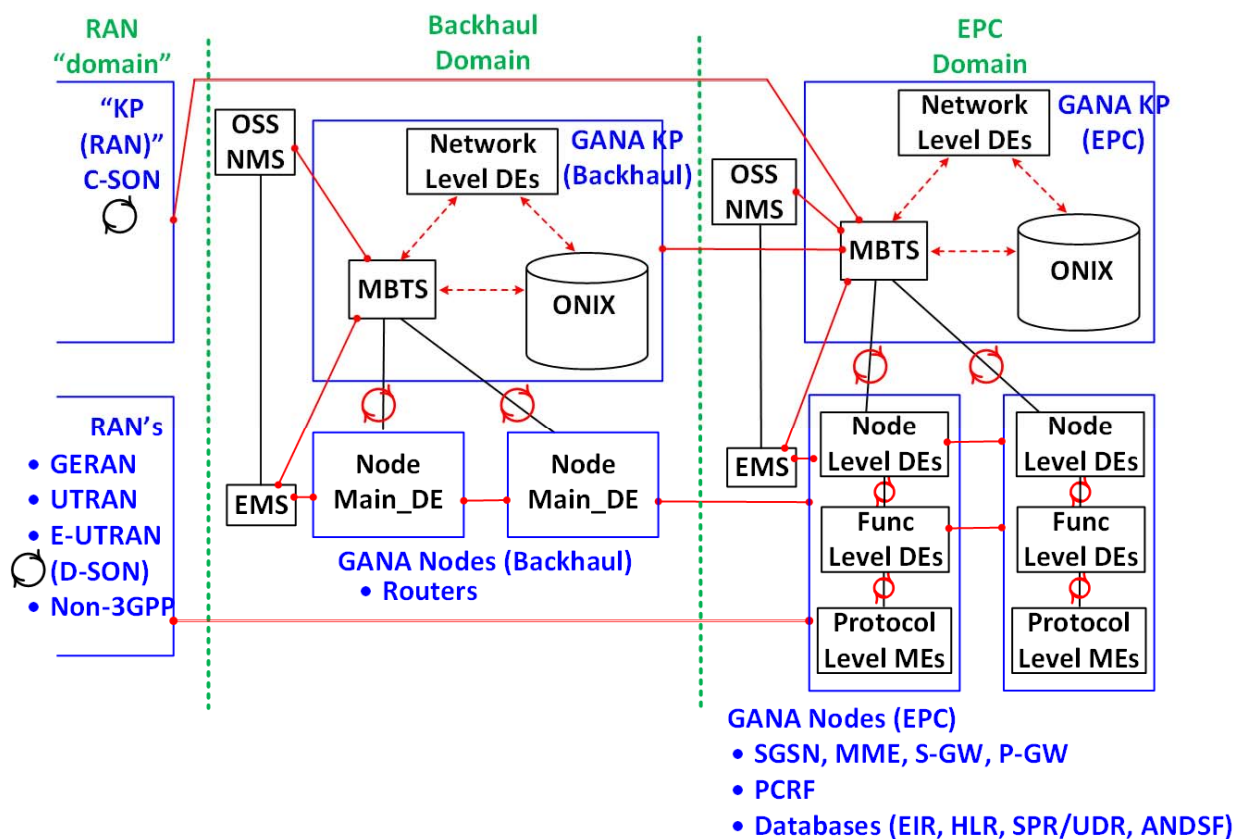


Figure 18: Interactions involving GANA entities for 3GPP Core Network and backhaul mapping

Thus, this clause describes the transfer of policies, characteristic information and knowledge that happens as part of the interactions among entities both horizontally (among different nodes and entities in the nodes) and vertically (among centralized management systems/knowledge plane and nodes) so as to support the decision taking and eventually the enforcement of autonomic decisions.

This is shown using a set of exemplary use cases. These are not all the possible use cases. Others may be studied, which would involve other DEs as defined in the GANA architecture [i.8] and/or nodes from the reference working architecture, but yet they provide a guide on how to implement the 3GPP instantiation. For each use case, the following elements are identified and analysed:

- a) the objective of the use case,
- b) the list and tasks of involved DEs,
- c) the information, in terms of parameters/groups of parameters, that is to be exchanged.

The examples are preceded in the next sub-clause with elements common to all use cases:

- a) the reference points at which information needs to be exchanged,
- b) the candidate associated protocols that can be used to exchange the information.

7.2 Reference points and associated protocols

In order to highlight the potential exchange of parameters in the respective Reference points of the architecture of Figure 18, the main reference points are summarized hereafter. Then the protocols that are used to exchange messages and information over these reference points are presented.

Reference points:

- Reference points are illustrated in Figure 18. Unless otherwise specified, they enable bi-directional interactions. The data exchanged are described in the next clause.
- Network level reference points:
 - From CN OSS i.e. the NMS at network level, or the EMS entities inside the nodes to MBTS;
 - Between RAN C-SON and MBTS;
 - For building the EPC KP: from MBTS to ONIX, between the Net_level DEs and ONIX, between the Net_level DEs and the MBTS, between the different Net_level DEs;
 - From the EPC KP to the Backhaul KP.
- Node level reference points:
 - Between the different entities of the KP EPC and Node_Main_DE, targeting the orchestration of the Func_level DEs;
 - Between the KP EPC and the Node level DEs;
 - Between the Node-Main DE and the involved Func_level DEs;
 - Between Node level DE and respective Function level DE;
 - Between each of the Node_level DE and its corresponding Mes;
 - Between each of the Func_level DE and its corresponding Mes;
 - Between the peer Node_level_DEs, across the GANA nodes;
 - Between the peer Func_level_DEs, across the GANA nodes;
 - Between the different Node_level DEs, inside the GANA node;
 - Between the different Func_level DEs, inside the GANA node.

Candidate Protocols to be used at the reference points:

- The main protocols used are listed in clause 7.2. The type of protocol to be used for each of the reference points is listed below. Furthermore, the protocols can be categorized in several sets, depending on the targeted outcome:
 - Protocols that focus on the policy or rule exchange: for example, COPS, BGP, DIAMETER;
 - Protocols that focus on the configuration: for example, SNMP, CMIP, Netconf, GTP-C, RestConf, Netflow™, OpenFlow®;
 - Protocols that collect, aggregate, and move large amounts of streaming event data: for example, Flume™;
 - Protocols that support application components to provide services to other components via a communications protocol, typically over a network: for example, SOA;
 - Protocols that support neighbour discovery and configuration: for example, Ipv6 Neighbour Discovery (ND); ANIMA (GRASP, RPL).
- Network level reference points:
 - From CN OSS (NMS, EMS) to MBTS: the protocols used are mainly management protocols, for example: COPS, SOA, Flume™, SNMP;
 - Between RAN C-SON and MBTS: the protocols used are mainly sharing a large amount of data for example Flume™, SNMP;
 - From MBTS to ONIX: the schemes used are mainly publish/subscribe, secure query/search, data mining and find mechanisms;
 - Between the Net_level DEs and ONIX: the schemes used are mainly publish/subscribe, secure query/search, data mining and find mechanisms;
 - Between the Net_level DEs and the MBTS: the protocols that can be used are mainly management and command protocols for example, COPS, SOA, SNMP, CMIP, Flume™, etc.
 - From the EPC KP to the Backhaul KP: the protocols that can be used are mainly sharing a large amount of command/operation and data/primitives e.g. COPS, SOA, BGP.
- Node level reference points:
 - Between the EPC KP and the Node_Main_DE/Node level DEs: the protocols that can be used are mainly management/command protocol e.g. COPS, CMIP, SOA, Diameter, Netconf, SNMP;
 - Between the Node-Main DE and the Func_level DEs: Node internal secure communication mechanisms;
 - Between each of the Func_level DE and its corresponding Mes in the EPC node the protocols that can be used are mainly management/command protocol e.g. Diameter, SNMP, Netconf, RestConf, Netflow™, OpenFlow®; and also the IGCP [i.31] could be used for such purposes if it gets further developed and standardized (and possibly an evolved IGCP protocol for supporting both Ipv4 and Ipv6 could be developed in the future);
 - Between the peer Func_level_DEs/Node_level_DEs, across the GANA nodes the protocols that can be used are mainly management/command protocol within a secure relationship e.g. GTP-C, ANIMA control plane protocol (GRASP, RPL), legacy protocols such as BGP, COPS, Diameter, SNMP, CMIP, SOA;
 - Between the different Func_level DEs, inside the GANA node the protocols that can be used are mainly management/command protocol within a secure relationship e.g. Node internal secure communication mechanisms.

7.3 Load Control

This clause defines the objectives, entities and interactions at network level for the Load control example described in clause 6.4.5.

Load Balancing and in general load control refers to the process whereby the load of the traffic that is going to be shared by more than one similar network elements is equally shared or generally controlled so as to serve specific network/operator needs. The similar network elements can be anything from P-GW, S-GWs, MMEs but also RAN elements such as base stations. As a matter of fact, the procedure is much intertwined with the selection functions described in ETSI TS 123 401 [i.2], in the sense that each time a decision is to be made that involves the selection of an MME/P-GW or S-GW, etc., respectively, it is important to take into account load information among the multiple candidate elements to be chosen. Accordingly and in order to increase system capacity, MME pools may share user traffic load across different MMEs, P-GW may be selected to serve PDN connections according to their load and capacity, while at the same time base stations may share/offload traffic to neighbouring cells, etc. As a result, different DEs (algorithms) per segment or even per node, can simultaneously take load balancing decisions as per need.

Objective:

Thus the objective of the sub-clause is to propose an assignment of the load control/balancing and gateway selection decisions taken by EPC Nes into GANA specified DEs, while at the same time to drill into the details with respect to the involved reference points, as well as the policies to be injected, the load specific parameters to be delivered and the protocols to be used for delivering these parameters in order to support such load control decisions.

Involved DEs:

- In the KP, Network level DEs that include: the Net_Level_MON_DE, for tuning the monitoring behaviours and tools of lower level monitoring DEs and the Net_Level_MOM_DE for taking decisions on the balancing and control of the load in the NEs and eventually in the network as a whole.
- In the GANA nodes, the equivalent Function level DEs are involved i.e. Func_Level_MON_DE and Func_Level_MOM_DE.
- Net_Level_MON_DE coordinate and inject policies to lower level monitoring DEs in nodes/devices to guide them on load related information that needs to be disseminated to external data collectors and requesting DEs like the Net_Level_MOM_DE.
- Net_Level_MOM_DEs essentially take decisions on the load assignment per NE, based on high level operator policies, but also based on load information collected from the involved nodes, typically representing the mid and longer term traffic measurements of each node.
- Analogous decisions can be taken at the node level, based on short term measurements of load and distributed exchange of load information at node level or among the respective Function level DEs.

Parameters or groups of parameters exchanged at the reference points

- Network level reference points:
 - From OSS (NM, EM) to MBTS. The MBTS translates the policies and key performance indicators (KPIs) received from the OSS and add into the GANA profile (by extending the initial GANA profile generated by human operator's tools for GANA Profile generation [i.25]), and also including processed KPIs. This acquired knowledge is transferred and stored into the ONIX and is also sent to DEs which use this knowledge to process real time decisions. In general, the GANA Profile should be generated using automation and translation tools (which could have some translation capabilities that may be similar to those employed for the MBTS-meaning that the MBTS could potentially be used for extending the GANA Profile (then stored in ONIX) with additional information).

- Between RAN C-SON and MBTS. This reference-point/interface allows to retrieve RAN related information, parameters, streams of KPIs and/or decisions made at RAN level by C-SON algorithms. The real-time data are directly consumed by the network level DEs to run the decision and data processing algorithms related to the desired network objectives, while aggregated data are stored in the ONIX. In the other direction, the network level DEs may either provide their decisions to the C-SON or collaborate with the C-SON decision logic functions, e.g. the load balancing or handover optimization, to coordinate the decisions made in order to achieve an e2e load balancing. This interface also transfers the control for the KPIs flow to be received, i.e. which parameters and with what periodicity required by different DEs.
 - Between the Net_level DEs and ONIX. This interface is used to transfer history of decisions, predicted network behaviour and traces. The ONIX provides GANA profiles to the Net_level DEs. Updated GANA profiles are exchanged in the other direction to be stored in the ONIX.
 - Between the Net_level DEs and MBTS. The Net_level DEs use the node level views, retrieved from the GANA nodes through the MBTS, to complement their own knowledge and knowledge stored in the ONIX. In the case where a decision impacts one or several entities in the network, the actuator part of the control loop triggers the corresponding commands, which are translated by the MBTS and forwarded to the related entities in the GANA nodes.
 - From the EPC KP to the Backhaul KP. The decisions made by the Net_level DEs may also require an adaptation of the routes or resource provisioning in the backhaul transport network. Policies and decision data are then provided to the backhaul KP, which triggers its own network level DEs.
- Node level reference points:
 - Between the different entities of the KP EPC and Node_Main_DE as the highest level entity in the GANA hierarchy at node level. In the upwards direction, the Node_Main_DE provides the node level view of the network, e.g. monitoring information and measurement results regarding the load in the node. In the downwards direction, it receives and forwards policies and commands regarding node weight factors, load thresholds, variant load estimation algorithms. This interface is also used to push the node configuration and parameters into the KP, and more specifically the ONIX, through the MBTS.
 - Between the Node-Main DE and the involved Func_Level DEs: The Node-Main DE orchestrates the Func_level DEs, forwarding in both directions the information exchanged with the KP, and more specifically the MBTS, to retrieve the required parameters or apply the decisions made at network level.
 - Mon_DE -ME: measurement commands (what (load), when, periodicity, duration), measurement reports.
 - Between the Func_Level_MON_DE and its corresponding Mes in the EPC node. The Func_Level_MON_DE uses the information it receives to run its own control loops. It communicates with the MEs it manages (monitoring mechanisms) to provide commands on how to measure the load, at which time, with what periodicity, with what granularity, for how long, etc.
 - Between the Func_level_MOM_DE and its corresponding Mes in the EPC node: The Func_Level_MOM_DE uses the load information it receives to run its own control loops. It communicates with its related MEs, i.e. 3GPP nodes, mainly for updating the weight factors based on which the traffic load is appropriately shared, directed, or offloaded among the involved nodes.
 - Between the peer Func_Level_DEs, across the GANA nodes. These DEs exchange local policies from network level, load information/weight factors at node level, or streams of Func_Level load related decisions or types of load estimation algorithms used.
 - Between the different Func_level DEs, inside the GANA node. These DEs collaborate internally to exchange decisions and parameters request/parameters report (values). The MOM_DE may request the MON_DE to configure data sources that could feed such data back to the MOM_DE.

Table 13: Examples of parameters exchanged for load control

From (vertical)/To (horizontal)	MME/S-GW/P-GW + MON-Func + Node-DE	MME/S-GW/P-GW + MOM-Func + Node-DE	MBTS + MON-Net	MBTS + MOM-Net	MBTS + C-SON	Backhaul KP
MME/S-GW/P-GW + MON-Func + Node-DE	Monitoring configuration (type of load metrics measured, granularity, periodicity)	Monitoring configuration (type of load metrics measured, granularity, periodicity)	Monitoring configuration (type of load metrics measured, granularity, periodicity)			
MME/S-GW/P-GW + MOM-Func + Node-DE	Requests for configuring the load specific measurements (e.g. collect specific metrics, select data sources, configure granularity, etc.)	Weight factors, local node load control/balancing decisions)		Used weight factors, used load control/balancing algorithms, local node load balancing decisions		
MBTS + MON-Net	Requests for Load metric/value and assigned weight factors			Information on monitoring configuration	List of KPIs of interest for retrieval	
MBTS + MOM-Net		Node weight factors, load thresholds, load estimation algorithms for GW selection			Update of Load Balancing related C-SON on load control/balancing decisions, load and weights factors in the elements of the core	Update of Load Balancing related C-SON on load control/balancing decisions, load and weights factors in the elements of the backhaul
C-SON + MBTS			Information on monitored RAN load metrics and associated metadata (averages, aggregates, periodicity)	Load related parameters (streams, aggregates) and load balancing decisions applied in RAN		
OSS + MBTS + ONIX			Policies for load monitoring behaviour (what metric, current vs predicted, periodicity, thresholds, triggers)	Policies for load control/balancing behaviour (dynamic assignment of weight factors, activation of load balancing algorithms) Customer/Subscriber related info		

7.4 Congestion resolution

Load balancing allows transferring load from capacity restricted resource to others which have available capacity. This capability is also very important in the case where the congestion of one node occurs in the network [i.23]. Using autonomicity to measure network quality and detect and resolve congestion situations would allow MNOs to react in a very short time frame to this type of events, improve the efficiency of their network operation and planning, maximize network utilization, avoid over-provisioning, and thus reduce their CAPEX in a way not possible manually. Considering the LTE network deployments, congestion affects mostly the user plane, in the RAN, or in extreme cases, in the S-GW. However, for completeness, the present document also considers it may occur in the P-GW and in the MME for the control plane. Existing mitigation techniques consist in disrupting the traffic in excess, which causes inconvenience to the MNO customer. Using autonomicity and GANA architecture, it is possible to resolve these situations by automatically offload part or whole of the UE traffic, or moving it between NEs, while maintaining the enforcement of the operator policies and authorizations.

Objective:

- The congestion resolution example covers the detection of congestion situations and their resolution by traffic offloading or UE mobility.
- The effect of the decision is to apply policies to execute the triggering of the resolution mechanisms, to control the monitoring behaviour, to define off-loading policies and mobility policies in the target nodes.
- In this example, the information collection also takes into account the reporting from the RCAF in the PCC and the UE subscription data from one of the databases installed in the EPC: HSS, HLR, UDC, User equipment information (EIR), or the User location information in the GMLC. Non-3GPP policies from the Access Network Discovery and Selection Function (ANDSF) may be considered as well.
- The impacted MEs are functions in the MME, the S-GW, and the P-GW, with some potential effect on the PCC entities as well.

Involved DEs:

- In the KP, this use case requires the Net_Level_MON_DE for monitoring and the Net_Level_MOM_DE for managing the offloading and the UE mobility. The Net_Level_QoS_M_DE may be involved as well to provide functionalities like Self-optimization and Self-adaptation.
- In the GANA nodes, the equivalent Function level DEs are involved: Func_Level_MON_DE, Func_Level_MOM_DE and potentially Func_Level_QoS_M_DE.
- At network level, the DEs take decisions such as: retrieve monitoring information, trigger congestion resolution, selectively offload part of the traffic from one or more UEs, relocate some PDN connections, start UE mobility to a new NE with available resources.

Parameters or groups of parameters exchanged at the reference points:

- Network level reference points:
 - From OSS (NMS, EMS) to MBTS. The MBTS translates the network policies and key performance indicators (KPIs) received from the OSS and add into the GANA profile (by extending the initial GANA profile generated by human operator's tools for GANA Profile generation [i.26]), and also including processed KPIs. This acquired knowledge is transferred and stored into the ONIX and is also sent to DEs which use this knowledge to process real time decisions. In general, the GANA Profile should be generated using automation and translation tools (which could have some translation capabilities that may be similar to those employed for the MBTS-meaning that the MBTS could potentially be used for extending the GANA Profile (then stored in ONIX) with additional information).

- Between RAN C-SON and MBTS. This reference point allows retrieving RAN parameters and streams of KPIs. The real-time data are directly consumed by the network level DEs to run the decision and data processing algorithms related to the desired network objectives, while aggregated data are stored in the ONIX (see more details in clause 7.7). In the other direction, the network level DEs may either provide their decisions to the C-SON or collaborate with the C-SON functions, e.g. the load balancing or handover optimization, to coordinate the decisions to be performed in order to mitigate the issue. This interface also transfers the control for the KPIs flow to be received, i.e. which parameters and with what periodicity.
 - Between the Net_level DEs and ONIX. This interface is used to transfer history of decisions, predicted network behaviour and traces. The ONIX provides GANA profiles to the Net_level DEs. Updated GANA profiles are exchanged in the other direction to be stored in the ONIX.
 - Between the Net_level DEs and ONIX. The Net_level DEs use the node level views, retrieved from the GANA nodes through the MBTS, to complement the knowledge already stored in ONIX. In the case where a decision impacts one or several entities in the network, the actuator part of the control loop triggers the corresponding commands, which are translated by the MBTS and forwarded to the related entities in the GANA nodes.
 - From the EPC KP to the Backhaul KP. The decisions made by the Net_level DEs may also require an adaptation of the routes or resource provisioning in the backhaul transport network. Policies and decision data are then provided to the backhaul KP, which triggers its own network level DEs.
- Node level reference points:
 - Between the different entities of the KP EPC and Node_Main_DE as the highest level entity in the GANA hierarchy at node level. In the upwards direction, the Node_Main_DE provides the node level view of the network, e.g. monitoring information and measurement results (see "List of parameters" below), while in the downwards direction, it receives traffic steering or UE mobility commands, duration of offloading, updated parameters, updated algorithms and policies. This interface is also used to push the node configuration and parameters into the KP, and more specifically the ONIX, through the MBTS.
 - between the Node-Main DE and the involved Func_level DEs: The Node-Main DE orchestrates the Func_level DEs, forwarding in both directions the information exchanged with the KP, and more specifically the MBTS, to retrieve the required parameters or apply the decisions made at network level.
 - Between the Func_Level_MON_DE and its corresponding Mes in the EPC node. The Func_Level_MON_DE uses the information it receives to run its own control loops. It communicates with the Mes it manages to provide commands and their parameters (e.g. measurement commands: what (load, UE location information), when, periodicity, duration), or receive corresponding responses and measurement reports.
 - Between the Func_level_MOM_DE and its corresponding MEs in the EPC node: The Func_level_MOM_DE uses the information it receives to run its own control loops. It communicates with its controlled MEs, i.e. the EPC nodes and/or functionalities, to provide commands and parameters related to offloading or UE mobility, e.g. parameters for traffic steering, overload control parameters, candidate routes for offloading, candidate anchors for UE mobility, or receive corresponding answers or reports.
 - Between the peer Func_level_DEs, across the GANA nodes (e.g. Func_Level_MON_DE in the MME and Func_Level_MON_DE in the S-GW). These DEs exchange local policies from network level, node-level monitoring information, or streams of Func_level decisions such as offloading decisions.
 - Between the different Func_level DEs, inside the GANA node e.g. between the Func_Level_MON_DE and the Func_Level_MOM_DE. These DEs collaborate internally to exchange decisions and parameters request/parameters report (values).

List of exchangeable parameters between GANA KP and C-SON

The detailed list of C-SON KPIs parameters is not standardized. Therefore, it depends on the vendor implementations and is often considered as confidential information. On the other hand, the functions performed by the different products available on the market are well-known: automatic neighbour relations, load balancing, interference control, coverage and capacity optimization, or handover optimization [i.24]. Service assurance platforms may be programmed to export the whole stream of KPIs or process them with their internal algorithms, while defining specific events where an aggregated view of the KPIs is exported to the GANA KP for the core network (EPC-KP). This corresponds to a flow of parameters or specific knowledge information with a time interval lower than with raw data. C-SON vendors may provide some means by which KPIs can be retrieved from C-SON.

C-SON - EPC KP Rfp and operation primitives exchange should be specified for interoperability's issues and for reducing the flow of information exchange. Otherwise EPC-KP should retrieve RAN views from probes, robots, network element sensor information, and the EPC-KP and the C-SON may not collaborate for decision coordination. Furthermore, if this Rfp is not standardized, interoperability between multivendor solutions will be difficult to reach. As at the RAN level different vendor s solutions are deployed. C-SON KPI may be more relevant to exchange with EPC KP in order to not flood the Rfp between C-SON and EPC-KP and share only relevant/useful information for decisions.

On the EPC side, ETSI TS 132 426 [i.25] specifies the performance measurements to be performed in the different nodes: MME, P-GW, S-GW, but also MBMS-GW and PCRF. For example, the following measurements are relevant in the congestion resolution example:

- MME
 - Failed EPS attach procedures
 - MME control of overload related measurements for EPC (Start procedure, Stop procedure)
 - EMM-Registered subscribers (Mean number, Maximum number)
 - Handover related measurements (Inter RAT handover)
 - Number of dedicated EPS bearers in active mode (Mean, Maximum)
 - Failed dedicated bearer activation procedures by MME
 - MME Processor usage (Mean, Peak)
- P-GW
 - Failed number of PDN-GW initiated Dedicated Bearer Creation
 - Failed PDN-GW initiated Dedicated Bearer Modification with QoS update
 - Active EPS Bearers related measurements for EPC (Mean Number, Max Number)
 - Number of EPS bearers (Mean number, Maximum number)
 - Sgi related measurements
- S-GW
 - GTP related measurements (GTP data/signalling packets, data volume on S4 and S12 interfaces)
 - S1-U data volume related measurements (GTP data packets on the S1-U interface)
 - Bearer resource Usage related measurements
- PCRF
 - IP-CAN session establishment related measurements

Table 14: Examples of parameters exchanged for congestion resolution

From (vertical)/To (horizontal)	RCAF/PCRF/PCC entities + MON-Func + Node-DE	MME/S-GW/P-GW + MON-Func + Node-DE	MME/S-GW/P-GW + MOM-Func + Node-DE	MBTS + MON-Net	MBTS + MOM-Net	MBTS + C-SON	Backhaul KP
RCAF/PCRF/PCC entities + MON-Func + Node-DE	User plane congestion alarm (UE subscription data, congestion level)	User plane congestion alarm (UE subscription data, congestion level)	UE subscription data retrieval	User plane congestion alarm (UE subscription data, congestion level)			
MME/S-GW/P-GW + MON-Func + Node-DE	UE mobility context retrieval (e.g. cell location)	Setup and reports for QoS measurement from CN Mes; retrieval of node level monitoring information	Provision of congestion information; CN QoS measurements and parameters reports	Provision of CN QoS measurements and parameters reports; Node level view of the network			
MME/S-GW/P-GW + MOM-Func + Node-DE		Measurements and Parameters request for impacted Mes	Updated local policies, node level mobility decisions; Local commands to Mes for UE offloading or mobility (candidate routes, candidate anchors, GW relocation)		Information about node level decision (mobility, offloading or GW relocation)		
MBTS + MON-Net		Parameters requests from node level DEs			Congestion and global network level information	List of parameters of interest for KPIs retrieval	
MBTS + MOM-Net			Network level mobility decision (Traffic steering, UE mobility, WiFi offload)			Synchronous update of C-SON functions with mobility decisions	Update of mobility decisions (routes changes, new resources provisioning)
C-SON + MBTS				RAN parameters retrieval (stream or aggregated)			
OSS + MBTS + ONIX				Policies for congestion trigger (e.g. congestion thresholds)	Policies for congestion recovery (e.g. authorization for WiFi offload)		

7.5 S-GW failure (without restart)

Following the capabilities that GANA model enables in Autonomic networks, the autonomic behaviours of the nodes can provide several solutions in cases of nodes restoration. An interesting example is when a PGW/MME/SGSN detects that a peer SGW has undergone failure; it deletes all PDN connection table data/MM bearer contexts associated with the peer node that fails as well as freeing any internal PGW/MME/SGSN resources associated with those PDN connections. Since there are no bearer/MM contexts in the PGW, the mobile terminated service cannot be delivered to the UE for long time. If the MME/SGSN performs implementation specific action to clean up external resources, including at eNodeB/RNC and UE, the UE will be reattached to the network and the EPS services will be restored. Incorporating the GANA model and extending the autonomic behaviour of the nodes, in the abovementioned case, an autonomic approach could be adopted for the restoration of the resources and the connections. The general approach would encompass the Autonomic Fault-Management activities in a GANA network. These are executed by a set of node level DEs and a network level DE on top of them. These DEs, together with the Resilience and Survivability DE on node/network level, and additional supporting components (e.g. incident dissemination mechanisms), form the Unified Architecture for Autonomic Fault-Management, Resilience, and Survivability (UAFAReS). The Network-Level-Fault-Management DE (NET_LEVEL_FM_DE) is responsible for Self-configuration of diverse aspects of the NODE_LEVEL_FM_DEs in a certain network scope (e.g. OSPF area). Furthermore, the NET_LEVEL_FM_DE is the instance that keeps the global knowledge with respect to incidents and alarms for the network. This enables the NET_LEVEL_FM_DE to react to network wide erroneous states in case the NODE_LEVEL_FM_DEs are not capable of resolving the occurred faults and to escalate faulty conditions to the network operation personnel in case they are not resolvable by means of Autonomic Fault-Management.

Objective:

- Address the autonomic behaviour of nodes in cases of failures and recovery attempts. The self-healing actions should be driven by policies, decision taking and knowledge sharing. The failures can be partial or total, can involve the SW or the HW or both and can have varying recovery delays and remedies. Depending on the policies that are fed in the node or network level, in combination with the Knowledge building performed in KP, the policies that apply in each case create also a feedback loop in order to produce new policies based on history and dependency of events and on system knowledge.
- Consider examples like Node Failures, e.g. S-GW failure:
 - Policies relevant to S-GW failure: S-GW selection, UE services consumed (Subscription Profile Repository (SPR)), re-allocation of the paths from/to eNB + P-GW.
- Processes and tasks that can benefit from autonomicity (automation using closed control-loops):
 - Perform self-healing and recovery either at node level or network level, using policies and knowledge from previous situations in order to tackle the healing and recovery process more efficiently (pro-actively).
 - In case of node failures, to automate the transfer of the traffic over to new connections/tunnels and/or allocate flows, users, tasks, etc. to other nodes seamlessly.

Involved DEs:

- **The GANA model enables Self-healing/Self-repair functionality.** The Autonomic Fault-Management control loop realized in the network devices is complemented by a control loop on the network level. The network level control loop facilitates the self-* features of self-healing and self-repair in case that the control loops inside the devices fail to introduce a successful self-healing/repair behaviour. The main functionalities involved in these processes are the **Fault-Detection and Incident Dissemination and Fault-Isolation**. The generic approach for these operations is the following:
 - The main processes of Fault-Detection and Incident Sharing, are carried out by the NET_LEVEL_FM_DE which is supported by a set of repositories and an Incident Dissemination Engine (IDE) that facilitate the sharing of alarm/incident events in a network scope. The NET_LEVEL_FM_DE keeps all incident events (across the network) in its repositories, i.e. it has the global view regarding faults. On the other hand, the NODE_LEVEL_FM_DEs get informed only about information that they are concerned with.

- The NET_LEVEL_FM_DE runs a Fault-Isolation process on the global view of incident events stored in its repositories. This enables the NET_LEVEL_FM_DE to identify network wide faults and initiate the required process of Fault-Removal.
- Therefore, the main DEs involved in the S-GW failure are:
 - In the Network level : The Fault-Management DE is used for the identification of failures, while Resilience and Survivability DE is used for the application of network wide repair schemes that will allow seamless fault removal and network operation.
 - The respective Node level DEs involved are: Fault Mngt, for the identification, repair and removal of faults, based on the parameters that are monitored by the Mon_DE.
 - The main Function level DE is the : Mon_DE, providing parameter monitoring.

Parameters or groups of parameters, policies exchanged at the reference points:

- Network Level Reference Points:
 - From OSS (NMS, EMS) to MBTS, mainly business level and network wide operational level policies recovery objectives and KPIs are exchanged.
 - Between C-SON and MBTS: C-SON to MBTS: parameters and RAN KPIs from C-SON for recovery and self-healing, load balancing, etc., if necessary - MBTS to CSON: alarms to C-SON Self-healing function for fault incidents, like e.g. S-GW failure:
 - S-GW performs per-bearer user plane tunnelling and inter-eNodeB path switching during handover. S1-U is the interface between E-UTRAN and S-GW. The transport protocol over this interface is GPRS Tunnelling Protocol-User plane (GTPv1-U). In case of S-GW failure, the C-SON should receive an alarm. Then, depending on the recovery policies adopted, the e-NBs may be prompted to establish a connection to another anchor S-GW. Relevant parameters can be: dedicated bearer identity, Linked Bearer Identity to identify the associated default bearer, the traffic flow template, and the associated QoS parameters.
 - Between MBTS - ONIX, mainly GANA profiles, translated KPIs and knowledge are exchanged. The MBTS translates the network policies and key performance indicators (KPIs) received from the OSS and adds into the GANA profile (by extending the initial GANA profile generated by human operator's tools for GANA Profile generation [1.6]), and also including processed KPIs. This acquired knowledge is transferred and stored into the ONIX and is also sent to DEs which use this knowledge to process real time decisions. DEs also subscribe for receiving information from ONIX using its publish/subscribe services.
 - Between the Net_level DEs and ONIX: history of decisions, healing and recovery schemes and optimizations at network level, traces, and network resilience schemes. This interface is used to transfer history of decisions, predicted network behaviour, healing and recovery schemes, network resilience schemes and traces. The ONIX provides GANA profiles to the Net_level DEs. Updated GANA profiles are exchanged in the other direction to be stored in the ONIX. Between Net level DEs-ONIX: The Net_level DEs use the node level views, retrieved from the GANA nodes through the MBTS, to complement the knowledge stored in ONIX . In the case where a decision impacts one or several entities in the network, the actuator part of the control loop triggers the corresponding commands, which are translated by the MBTS and forwarded to the related entities in the GANA nodes by updated GANA profiles.
 - Between Net-level DEs - MBTS: The Net_level DEs use the node level views, retrieved from the GANA nodes through the MBTS, to complement the knowledge stored in ONIX. In the case where a decision impacts one or several entities in the network, the actuator part of the control loop triggers the corresponding commands, which are translated by the MBTS and forwarded to the related entities in the GANA nodes.

- EPC KP - Backhaul KP exchange policies, decisions, views (e.g. KPIs). The decisions made by the Net_level DEs may also require an adaptation of e.g. the routes or resource provisioning in the backhaul transport network. Policies and decision data are then provided to the backhaul KP, which triggers its own network level DEs. The MON_DE (node or network level) and the Fault-Management_DE (node or network level) exchange mainly TcoSH (Trigger Condition of Self-Healing).
 - Fault Management DE and Resilience & Survivability DE exchange mainly data analytics and policies for decisions related to network robustness and fault tolerance and resolution, as well as schemes for self-Healing and recovery.
- Node Level Reference Points:
- Fault Management DE (network level) and Fault Management DE (node level) exchange mainly policies for self-Healing and recovery and TCoSH (Trigger Condition of Self-Healing).
 - Peer Node level DEs exchange mainly monitoring data, as well as policies for recovery.

An example table illustrating the use of the identified parameters between respective reference points in the case of S-GW failure is shown hereafter. The rows of Table 15 describe the initial (triggering) nodes to be considered for the exchange of the parameters, while in conjunction with the columns the targeted nodes can be identified.

Table 15: Examples of parameters exchanged for S-GW failure (without restart)

From (vertical)/To (horizontal)	eNodeBs+FM-Node	MME/SGSN+FM-Node	SGW1+FM-Node	SGW2+FM-Node	PGW+FM-Node	MBTS	CSON	OSS/ONIX/FM_Node/Node_DE/BH KP
eNodeBs+FM-Node						Path reallocation: LBI, QoS		history of decisions
MME/SGSN+FM-Node						Path reallocation: LBI, QoS		history of decisions
SGW1+FM-Node	Failure trigger(ID)	Failure trigger(ID)	Failure ID, TcoSH		Failure trigger(ID)	Failure trigger(ID), TcoSH		history of decisions
SGW2+FM-Node						Path reallocation: LBI, QoS		history of decisions
MBTS+FM-Net	Reallocation of the paths to SGW2:measureme bearer identity, Linked Bearer Identity to identify the associated default bearer, the traffic flow template, and the associated QoS parameters	Reallocation of the paths to SGW2:measureme bearer identity, Linked Bearer Identity to identify the associated default bearer, the traffic flow template, and the associated QoS parameters		Recovery scheme: path reallocation, TcoSH	Reallocation of the paths to SGW2: LBI, QoS params, TcoSH		alarm for fault incident S-GW failure measureme KPIs raw KPIs	history of decisions, Failure(ID), recovery scheme: path reallocation TcoSH, Policies and decision data Policies and decision data , healing and recovery schemes and optimizations at network level, traces, and network resilience schemes
CSON						Recovery scheme Failure ID (e.g. S-GW failure)		
OSS/FM_DE/NET_DE/ONIX						Policies for recovery, business level,network wide operational level policies, KPIs	GANA Profiles, KPIs, knowledge	GANA Profile, Failure(ID), recovery scheme, TcoSH measurement of new setup, Policies, decision data, knowledge

8 Recommendations

8.1 Overview

The main outcome of the study in the present document is that it is possible to introduce autonomies (management and control intelligence) in the Core and the Backhaul without impacting the existing 3GPP reference points and architecture. Existing protocols and management mechanisms can be leveraged to execute the necessary interactions between the Core Network standardized nodes and functionalities and the overlay of atomic autonomic components (software) entities introduced by the GANA framework mapping. The autonomic components may be launched in the network as intelligence-enhancing management and control software. There is no impact on the user plane or the control plane. However the work identified the need for a reference point in the management plane to be placed between the C-SON and the GANA Knowledge Plane for the Core network. This would enable information (e.g. some KPIs) or decision sharing between the two entities so as to implement collaborative end-to-end Self-optimization by C-SON and GANA Knowledge Plane for the Core Network.

This interaction and more specific recommendations regarding implementation alternatives of the GANA model are derived in the following clauses.

8.2 Reference Point between C-SON and GANA KP for the Core Network

The cooperation/collaboration between the EPC KP and the respective C-SON entity(-ies) being responsible for the RAN segment is apparently very important first, for ensuring the maximum impact of introducing GANA/Autonomic in the EPC and second, for ensuring that the latter can be obtained with the minimum impact to the currently standardized 3GPP interfaces. However, it also calls for the definition of a Reference Point between them. This interaction is illustrated in Figure 16 for the case of Self-healing and highlighted in the three exemplary use cases developed in clause 7. An attempt to describe this reference point, at least in terms of required characteristic information is presented below.

C-SON may communicate/share the following information:

- Situation Description and associated Decisions executed by C-SON (situation is assumed to be pending at the time C-SON exchanges this information).
- C-SON may query the Core Network GANA KP of "current situation" of the core network, using pull model.
- For certain decisions of which C-SON may need to synchronize its actions with the Core Network GANA KP, a "synchronization message" could be sent by C-SON to the Core Network GANA KP so that the two can coordinate the approval/disapproval and execution of the tentative actions. The "synchronization message" should convey the following parameters: "situation", "decision in form of tentative actions".

NOTE: The subject of synchronizations between C-SON and Core Network GANA KP versus the "autonomous" operations (by which influence is achieved only through information exchanged via a push/pull model without explicit enforcement of synchronization), is subject for further study.

- C-SON may aggregate some KPIs that relate to various SON functions or RAN state and communicate the aggregate KPIs to the Core Network GANA KP using a "push model", or C-SON may push various SON or other RAN KPIs (which may be raw i.e. being in the initial form in which they were collected without being further processed) or other measurements to Core Network GANA KP. Aggregate KPIs reduce the need to stream raw data and/or raw KPIs. The frequency by which a push operation occurs may be configurable. Examples of aggregate KPIs are average load level forecasts in a cluster of cells, and examples of raw KPIs or measurements are attachment success rate or Counter values.
- For both aggregate KPIs and some raw KPIs, or measurements, C-SON may provide a means for the Core Network GANA KP to pull the information at its own will.

Core Network GANA KP may communicate/share the following information:

- Situation Description and associated Decisions executed by the Core Network GANA KP (situation is assumed to be pending at the time the Core Network's KP exchanges this information).
- Core Network GANA KP may query the C-SON of "current situation" of the RAN, using pull model.
- For certain decisions of which the Core Network GANA KP may need to synchronize its actions with C-SON, a "synchronization message" should be sent by the Core Network GANA KP to the C-SON so that the two can coordinate the approval/disapproval and execution of the tentative actions. The "synchronization message" should convey the following parameters: "situation", "decision in form of tentative actions".

NOTE: The subject of synchronizations between C-SON and Core Network GANA KP versus the "autonomous" operations (by which influence is achieved only through information exchanged via a push/pull model without explicit enforcement of synchronization), is subject for further study.

- Core Network GANA KP may aggregate some KPIs or data that relate to autonomic management and control aspects or core network state in general and communicate the information retrieved from aggregate data (e.g. aggregate KPIs) to the C-SON using a "push model", or Core Network GANA KP may push a stream of raw data (e.g. some raw KPIs) to C-SON (which then needs to process the raw data/KPIs on its own to retrieve or semantically aggregate the information and create knowledge that requires to make local decisions). Aggregate KPIs reduce the need to stream raw data and/or raw KPIs. The frequency by which a push operation occurs may be configurable. Example of aggregate KPIs is an alarm for faulty node (alarm ID) or a Key performance indicator processed from a various types of measurements, and examples of raw KPIs are node level measurements performed by the EPC nodes.
- For both aggregate KPIs information and some raw KPIs or data, Core Network GANA KP may provide a means for the C-SON to pull the information at its own will.

Example GANA KP and C-SON interactions in the use case of Self-healing - S-GW failure

The C-SON Self-healing function and the GANA Self-healing related functions have similarities and complementarities. A harmonisation of the detailed alarm/s which a self-healing function monitors should be performed during the implementation phase. For example C-SON should be able to monitor alarms from the GANA Network level DEs. The recovery actions can be different, although a harmonisation of these actions between core and access is also preferable. So, it is recommended that the solutions of Self-healing would be implemented in a harmonised way in terms of monitoring and recovery schemes between C-SON and GANA Self-healing functions. Following this, it is recommended to define a general framework for the Self-healing function, including the monitored faults that a Self-healing function may resolve and the general method of the Self-healing function. The alarms monitored by the self-healing functions would benefit from a harmonised framework and the recovery actions of the self-healing function may be specified for all self-healing use cases which need cooperation between any DE and ME(s). The general framework may be used by developers of DEs and C-SON to develop a trusted solution which ensures interoperability.

8.3 Enhancing the GANA KP for EPC with analytics capabilities

GANA comprises a hybrid model that defines the abstraction levels at which interworking control-loops can be designed, while enabling to design nested and distributed control-loops to address complexity in design of autonomic functions. It is quite challenging to continuously analyse and merge raw data streams so that the results make actionable and operational sense.

The GANA KP and abstraction model can be viewed as a Modularized Analytics Engine, where each of the DEs may perform analytics and based on these analytics, the decisions and respective policies may be updated in the DEs. This is the part that incorporates the feedback loops and this may incorporate Complex Event Processing (CEP), which may support the harmonisation of the respective updates in policies and decision making, following the identification of a series of events. Examples of real-time and predictive analytics by specific DEs, may be addressed especially in Network Level DEs for the GANA Knowledge Plane like: QoS and QoE Management DE, Mobility Management DE, Fault Management DE or Resilience & Survivability DE. For example fault analytics may be gathered by Fault Management DE. After performing and evaluating the data analysis, CEP for Decision-Making in individual DEs can be realized in two ways, either at each DE, or an instance of a CEP Module may undertake the assessment overhead and feed all the relevant DEs. This implies the update of the policies in order to enable updated decision making based on the feedback loop.

The data analytics can be based on Monitoring tools gathering various measurements and data from the network operations and services. The data can be of big volume or it can also be history and real time data. In any case, the respective application of data analytics can have influence on the update of policies and decision making capabilities of the DEs, resulting in changing network or node behaviour.

The report from Joint SDOs/Fora Workshop [i.11] discusses the relationship between analytics engines/processes designed specifically for management and control of network resources, parameters and services, and the broader paradigm of autonomics whereas examples of input to real-time and predictive analytics by specific DEs (including example use cases for decision-making) can be found in the paper [i.10].

8.4 Policies-related considerations

Need for dynamic/advanced decision and policy derivation

To fully deploy the potential of an autonomic EPC, the distribution of GANA DEs in the EPC nodes, the KP and/or management systems is the first and important step but it is not sufficient. Decisions and policy rules are today limited to network segments or proprietary network equipment without considering the whole network. Moreover, they do not learn from historical situation, whereas most of policy systems are based on static condition-action rules. These DEs and policy systems can be associated to create a more powerful and advanced closed-loop decision system that:

- a) takes into account information exchanged between segments in an e2e manner considering the whole network, and
- b) combines knowledge deriving from learning historical data and real-time analytics in order to dynamically adapt to the varying and unforeseen network situations.

Example: the knowledge plane and the feedback loops in order to update the policies in Self-Healing - S-GW failure

The knowledge plane and the policies have a strong interdependence. In self-healing functions, the respective business level policies will be incorporated in network level and node level DEs, in order to enable the identification and prioritization of Faults/Failures as well as the appropriate recovery schemes. Following the actions taken in each case, then this input will be feedback to the knowledge plane along with the results of the actions, e.g. whether the failure was repaired, whether the issue has been escalated to the administrators and required human intervention, etc. In this case the feedback loop is able to generate new policies that depend on the situation awareness knowledge enriched by the results of the actions triggered by the default policies. The feedback loop will be able to help update the diagnosis process (based on e.g. structured/unstructured data and/or events). So, it is recommended that the policies and respective solutions of self-healing should be implemented in case by case manner.

8.5 Other considerations

MBTS and ONIX

The role of the MBTS is not to store knowledge, but rather to process ontologies from different vendors, technology domains in order to translate the policy/data received at its interfaces with OSS and GANA node/Network Element and add into the GANA profile, and create information which can then be processed with a common meaning for human (developers/administrator) and DEs (algorithms). In general, the GANA Profile should be generated using automation and translation tools (which could have some translation capabilities that may be similar to those employed for the MBTS-meaning that the MBTS could potentially be used for extending the GANA Profile (then stored in ONIX) with additional information). On the other hand, ONIX is a distributed system providing secure services for storing, retrieving and disseminating information. It does not contain any decision logic.

In the 3GPP CN mapping, the ONIX stores information with a larger time scale, e.g. in the granularity of days, months or even years:

- Profiles from OSS translated into cleaned GANA profiles (as an xml structure)
- updated GANA profiles after DE algorithm execution
- history of DE commands and their context

- Information pushed from GANA nodes at the EPC nodes (MME, S-GW, P-GW, PCRF): capabilities description, resources, configuration
- configuration information pushed by 3GPP databases (HSS, HLR, EIR, UDC, ANDSF, etc.) when relevant
- Information shared with BH, EPC and RAN e.g. aggregated KPIs information from RAN and measurement from EPC nodes
- predicted network behaviour

Exchanged parameters for the development of GANA algorithms and autonomous functions

Sub-clauses 7.3 to 7.5 have provided exemplary use cases and indicated the flows of parameters and information exchanged at the different reference points, including the link between the DEs and their respective Mes. Developers willing to implement specific algorithms and autonomous functions related to the present mapping should design a set of GANA entities corresponding to the target scenario and fill a parameters table constructed on the models given in Table 13, Table 14 and Table 15. Additional recommendations can be found in clause 3.6 of the AFI GANA White Paper [i.6]. These recommendations help to design an autonomic network that could be trusted by network providers.

Minimum DEs in each EPC node: Monitoring

Monitoring is an important aspect of the management of 3GPP CN. Thus all EPC nodes should at least host a GANA node including the Node-Main DE and the Func_Level_MON_DE to be able to report the node level view of the network to the network level DEs through the MBTS, which also then translates information from the network nodes, including some monitoring data into knowledge that is stored in the ONIX.

On the allocation of PDPs and PEPs

Because of the dynamic nature of the GANA model (meaning that there is no restriction in the model to address only static PDPs and PEPs), the attribution of PDP vs. PEP functions may vary over time, mainly for the entities related to the KP and for the DEs in the GANA nodes. From a general point of view, a DE acts as a PDP for its lower level DE and as PEP for its upper level DE. Each ME (e.g. 3GPP EPC node) is considered as the PEP of its respective Function level DE which acts as a PDP in this interaction. In the 3GPP CN mapping, the network level DEs may act as PEP or PDP for the OSS and the RAN C-SON, depending on the target function and the status of the network. In this case, the notion of cascading PDP is introduced, when OSS acts as PDP for the network level DEs which, in their turn, are PDPs for the DEs in the GANA nodes. However, a reverse setup can also be perceived as well, in which the PDP is the GANA Knowledge Plane DEs and the OSS being a PEP, whereby the Knowledge Plane DEs command the OSS via APIs of the OSS to execute actions (e.g. writing management and control transactions) towards the network. The references [i.15] and [i.27] discuss this perspective.

Insights on GANA Implementation Guide that implementers need to consider

Quoting [i.28]:

It does not mean that in order to implement Autonomic Functions (Afs), meaning DEs, every hierarchical level in GANA has to be implemented in the target architecture. Because in incremental implementation of autonomies in a network architecture, one particular GANA level or multiple GANA levels and associated DEs may be collectively considered at a time. A full implementation of interworking autonomies at multiple levels (especially Level-2 to Level-4) may simply emerge over time.

More discussions on this subject are provided in the clause on the Implementation Guide for GANA in GANA White Paper [i.6].

On enabling both, Multi-Party Providers (Suppliers) of GANA DEs (i.e. Autonomic Manager Software Modules) versus Single Party Providers (Mono-vendor) for a given scope of a network and its management architecture

For a given network architecture and its management and control architecture, carrier (network operator) may go for multiple providers (suppliers) of DEs and their algorithms and software modules, or a single (mono) supplier covering a certain scope or entire architecture. A mono (single) DE vendor may cover a certain scope in terms of DEs that they can instantiate to operate and be coordinated by a Coordination Function (s) to achieve global optimization objectives for the scope covered. Mono-vendor option has the potential of reducing the burden on addressing interoperability of autonomic functions (DEs and other GANA Functional Blocks).

A concrete vendor scope of coverage in terms of their DEs they provide (as a supplier), and global scope of optimization and coordination of the Autonomic Functions (DEs) in a network and its management and control architecture it is expected to be well understood by targeted customers (network operators, enterprises, etc.). Example of network architecture is 3GPP Core Network architecture, mobile backhaul network architecture, or a wireless Mesh Network architecture [i.7]. The following points provide more insights on the notion of scope of coverage of autonomics space by a vendor of autonomics:

- [i.30] discusses various potential suppliers of DEs for the levels defined in GANA, such as ISVs (Independent Software Vendors) e.g. OSS vendors, and Traditional networking equipment vendors.
- For a given network architecture and its management and control architecture, it is possible that there may be multiple providers of DEs and their algorithms and software modules or a single provider covering a certain scope or entire architecture.
- For a given network architecture and its management and control architecture, a single DE vendor may cover a certain scope in terms of DEs that they can instantiate to operate and be coordinated by a Coordination Function (s) to achieve global optimization objectives for the scope of network architecture covered. The coverage scope should be communicated to targeted customers of the whole autonomics solution. Such a scenario on single provider of autonomics software for a given network architecture and its associated management and control architecture has the potential of reducing the burden on addressing interoperability of autonomic functions (DEs and other GANA Functional Blocks). This is because global synchronization of autonomic functions on actions and policies towards achieving stability and global optimization objectives may be easily addressed by the single vendor over the scope of coverage than in the case with multiple vendors needing to agree on how to interoperate the autonomic functions (NGMN tried to address the similar challenge in relation to multivendor SON deployment recommendations (in NGMN deliverable D2 [i.32])).
- Mapping Autonomic Managers Components to GANA Levels of Abstractions for Self-Management Operations helps to understand the scope ("operating-region of the Control-Loop(s)-a concept from control theory" implemented by a particular Autonomic Manager (DE)):
 - The ETSI NTECH AFI GANA PoC Template [i.33] provides a table that can be used by developers of Autonomic Manager Components (DEs) to map their "autonomic managers" to specific GANA Levels.

9 Conclusion and Further Work

The present document describes the introduction of autonomics in the 3GPP Evolved EPC and backhaul network segments. Both network segments are beyond the scope currently covered by SON. SON itself is part of the broader picture of the autonomics concept. Introducing autonomics in these network segments complements the existing 3GPP management architecture with add-on automated and autonomic management & control (AMC) features that target dynamic management and adaptation of resources in these two segments.

The present document is based on the GANA framework, which also provides for coordination and synchronization between Decision making Elements (DEs) of control loops to ensure the stability of the network and prevent conflicts between the decisions that might happen. It considers an e2e perspective to optimize the operation of the network. Implementing autonomics in the EPC and backhaul networks, together with interactions with the C-SON allows to ensure this e2e perspective and to improve the global operation of the network. The backhaul network itself is considered as a black box and no direct impact is foreseen on its nodes, e.g. the IP routers. Whenever a relation is needed between the backhaul KP and the C-SON e.g. when the C-SON needs some information from the backhaul KP, it can be processed through the EPC KP.

The present document provides use case examples and recommendations for implementers to help them develop their algorithms and plan the testing of their products. It aims at providing a guide on how to implement the instantiation of GANA Functional Blocks (FBs) and their reference points, which enable autonomicity (automation with Decision-making-Elements (DEs) and their associated closed-control loops) in the target architecture.

The present document may be completed with further studies covering, for example:

- the whole set of possible use cases involving the EPC nodes. In particular, other use cases showing the mapping with 3GPP architecture may be studied, which would involve the other DEs defined in the GANA architecture. For example, further use cases may involve billing and charging functions, which would include directly the PCRF and the PCC architecture as MEs, while involving DEs such as Monitoring, Data Plane or QoS management DEs. This is illustrated in Figure 18, where the PCRF appears as one of the potential GANA nodes;
- a thorough study of the synchronizations between C-SON and EPC KP versus the "autonomous" operations (by which influence is achieved only through information exchanged via a push/pull model without explicit enforcement of synchronization);
- virtualised 3GPP EPC and backhaul network architectures in terms of autonomic functions: a further study should seek to cover other use cases for autonomies in the backhaul and core networks, as well as GANA integration with NFV (Network Functions Virtualisation) and SDN (Software-Defined Networking) frameworks for 3GPP core network and backhaul network architectures, based on applying the ongoing work in ETSI on GANA integration with NFV and SDN.

The result of such studies will be documented in the form of a revision of the present document or in a new report.

NOTE: The stable draft version of the present document was shared with 3GPP SA2 (architecture) and SA5 (telecommunication management) working groups, and presentations were made to both groups in the timeframe of June-July 2016 in order to solicit comments. According to the initial feedback received, the efforts around this work (reported by the present document) have been acknowledged, and no specific suggestion for modifying any part of the content or rework of the present document has been put forward. However, any further comments that pertain to improving the mapping of GANA to 3GPP may be addressed in a revised version of the present document. Moreover, the ETSI NTECH technical committee has been invited by the 3GPP SA2 and SA5 working groups to keep them informed of the results of any further study on introducing autonomies in the 3GPP backhaul and core network architectures. The prospective topics for further study (listed above) were also presented to both 3GPP SA2 and SA5.

Annex A: Change History

Date	Version	Information about changes
December 2015	0.0.1	First publication of the TR for TC NTECH #13
February 2016	0.0.2	STF Internal version including new Chapter 7
February 2016	0.0.3	Improvements to be submitted to steering committee on 03/03/2016
March 2016	0.0.4	Update for TC NTECH #14
March 2016	0.0.5	Update for LS to 3GPP
April 2016	0.0.6	Publication of a stable draft in the ETSI TC NTECH portal
May 2016	0.0.7	Update of the stable draft with comments received from NTECH
June 2016	0.0.8	Distribution to 3GPP SA2 and SA5, followed by a presentation of the draft to SA2 and SA5 (in the June-July 2016 timeframe)
July 2016	0.0.9	Final draft for distribution to NTECH AFI, updated mainly based on the initial feedback received from 3GPP SA2 and SA5

History

Document history		
V1.1.1	October 2016	Publication