

Handling Excessive Service Based Interface Signalling in a Core Network

TECHNICAL FIELD

[0001] The present disclosure relates to wireless communications and more specifically to the handling of excessive Service Based Interface, SBI, signalling in a core network.

BACKGROUND

[0002] A wireless communications system may include one or multiple network communication devices or nodes, such as base stations, which may support wireless communications for one or multiple user communication devices, which may be otherwise known as user equipment (UE), or other suitable terminology. The wireless communications system may support wireless communications with one or multiple user communication devices by utilizing resources of the wireless communication system (e.g., time resources (e.g., symbols, slots, subframes, frames, or the like) or frequency resources (e.g., subcarriers, carriers, or the like). Additionally, the wireless communications system may support wireless communications across various radio access technologies including third generation (3G) radio access technology, fourth generation (4G) radio access technology, fifth generation (5G) radio access technology, among other suitable radio access technologies beyond 5G (e.g., sixth generation (6G)).

[0003] The Service-Based Architecture (SBA) for core 5G networks is defined in 3GPP Technical Specification (TS) 23.501 — “System Architecture for the 5G System”. This architecture uses service-based interfaces between control-plane functions, while user-plane functions connect over point-to-point links. The SBA architectural approach enables 5G network functionality to become more granular and decoupled. Individual services can be updated independently with minimal impact on other services and can be deployed on demand to provide a number of advantages and efficiencies including:

- Any service can be supported by a different Network Function type. Different Network Function vendors can interoperate via a common Service Based Interface (SBI).
- Simpler operation by using application programming interface (APIs).

- Leverages the use of standard protocols such as HTTP instead of using the legacy Diameter protocol that was deemed complicated to develop and which often depends on vendor specific proprietary procedures.

[0004] Figure 1 illustrates four models A-D for providing communication between a pair of Network Functions (NFs) within an SBA architecture, where an NF that requests and consumes services and / or information is referred to as a “consumer” and an NF that responds to a request for services and / or information is referred to as a “producer”. [Reference to the “NRF” is reference to the Network Repository Function which acts as an index that can be consulted by other NFs to allow them to discover information regarding other entities present in the core network, as well as service capabilities that may be required. Reference to the “SCP” is reference to Service Communication Proxy.] The illustrated models are:

[0005] Model A - Direct communication without NRF interaction: Neither NRF nor SCP are used. Consumers are configured with producers' "NF profiles" and directly communicate with a producer of their choice.

[0006] Model B - Direct communication with NRF interaction: Consumers do discovery by querying the NRF. Based on the discovery result, the consumer does the selection. The consumer sends the request to the selected producer.

[0007] Model C - Indirect communication without delegated discovery: Consumers do discovery by querying the NRF. Based on the discovery result, the consumer performs the selection of an NF Set or a specific NF instance of NF set. The consumer sends the request to the SCP containing the address of the selected service producer pointing to a NF service instance or a set of NF service instances. In the latter case, the SCP selects an NF Service instance. If possible, the SCP interacts with the NRF to obtain selection parameters such as location, capacity, etc. The SCP routes the request to the selected NF service producer instance.

[0008] Model D - Indirect communication with delegated discovery: Consumers do not do any discovery or selection. The consumer adds any necessary discovery and selection parameters required to find a suitable producer to the service request. The SCP uses the request address and the discovery and selection parameters in the request message to route the request to a suitable producer instance. The SCP can perform discovery.

[0009] Use of the SCP has not completely resolved the potential for a signaling “storm” of control plane messages within the core network and it remains an issue for network operators. This is reflected in numerous discussion papers for Rel19 prioritization workshop that were submitted to the SA#100 Plenary (SWS-230013, SWS-230016, SWS-230034, SWS-230035, SWS-230050, SWS-230068).

SUMMARY

[0010] An article “a” before an element is unrestricted and understood to refer to “at least one” of those elements or “one or more” of those elements. The terms “a,” “at least one,” “one or more,” and “at least one of one or more” may be interchangeable. As used herein, including in the claims, “or” as used in a list of items (e.g., a list of items prefaced by a phrase such as “at least one of” or “one or more of” or “one or both of”) indicates an inclusive list such that, for example, a list of at least one of A, B, or C means A or B or C or AB or AC or BC or ABC (i.e., A and B and C). Also, as used herein, the phrase “based on” shall not be construed as a reference to a closed set of conditions. For example, an example step that is described as “based on condition A” may be based on both a condition A and a condition B without departing from the scope of the present disclosure. In other words, as used herein, the phrase “based on” shall be construed in the same manner as the phrase “based at least in part on. Further, as used herein, including in the claims, a “set” may include one or more elements.

[0011] Some implementations of the method and apparatuses described herein may include a network node comprising at least one memory and at least one processor coupled with the at least one memory and configured to cause the network node to receive an analytics request relating to Service Based Interface, SBI, signalling exchanges generated or received for each of a first set of network functions, subscribe via an Operation, Administration and Management, OAM, service to receive first information on a volume of SBI exchanges for each of the first set of network functions, derive for each of the first set of network functions a threshold volume of SBI signalling exchanges based on the first information, subscribe to a Management Data Analytics Service, MDAS, to receive fault information for a second set of network functions, upon receiving fault information from the MDAS for one of the second set of network functions, monitor via said OAM service a volume of SBI signalling

exchanges for one of the first set of network functions to determine whether or not the associated threshold volume is reached, and provide an analytic output for the one of the first network functions in response to the threshold volume being reached.

[0012] The threshold volume may be a ratio of an expected number of SBI signalling exchanges. The expected number of SBI signalling exchanges may be included in the analytics request.

[0013] The analytics request may include network event information indicating an event taking place that will cause a network outage.

[0014] The network event information may include one or more of an area of interest, a validity time, a first set of NF information, and a second set NF information.

[0015] The analytic output may include one or more of information relating to a network outage event, statistics or predictions of a percentage increase in SBI signalling when a network outage event occurs, impacted NF-types, impacted NF instances, and priority of the SBI message.

[0016] The second set of network functions may include one or more of the network functions of the first set of network functions.

[0017] The network node may provide or operate as a Network Data Analytics Function, NWDAF.

[0018] The analytics request may explicitly identify said first set of network functions, or identifies parameters that implicitly identify said first set of network functions.

[0019] The network node may be a node of a 5G core network.

[0020] A threshold volume may be indicative of a volume of SBI signalling exchanges greater than an expected volume during normal operation of the associated NF or an associated group of NFs.

[0021] Some implementations of the method and apparatuses described herein may include a processor for a network node, comprising at least one controller coupled with at least one memory and configured to cause the processor to receive an analytics request relating to Service Based Interface, SBI, signalling exchanges generated or received for each of a first set of network functions, subscribe via an Operation, Administration and Management, OAM, service to receive first information on a volume of SBI exchanges for each of the first set of network functions, derive for each of the first set of network functions

a threshold volume of SBI signalling exchanges based on the first information, subscribe to a Management Data Analytics Service, MDAS, to receive fault information for a second set of network functions, upon receiving fault information from the MDAS for one of the second set of network functions, monitor via an OAM service a volume of SBI signalling exchanges for one of the first set of network functions to determine whether or not the associated threshold volume is reached, and provide an analytic output for the one of the first network functions in response to the threshold volume being reached.

[0022] Some implementations of the method and apparatuses described herein may include a network node comprising at least one memory and at least one processor coupled with the at least one memory and configured to cause the network node to send an analytics request relating to Service Based Interface, SBI, signalling exchanges generated or received for each of a first set of network functions to a Network Data Analytics Functions, NWDAF, and receive in response from the NWDAF an analytic output indicating that a threshold volume of SBI signalling exchanges for one or more of said first set of network functions has been reached.

[0023] The network node may implement or operate as a Service Communication Proxy, SCP, the processor being configured to use said analytic output to throttle or prioritise control plane traffic of the network.

[0024] The network node may implement or operate as a Network Function, NF, the processor being configured use said analytics output to modify signalling to be sent to other NFs of the network.

[0025] Some implementations of the method and apparatuses described herein may include a method performed by a network node and comprising receiving an analytics request relating to Service Based Interface, SBI, signalling exchanges generated or received for each of a first set of network functions, subscribing via an Operation, Administration and Management, OAM, service to receive first information on a volume of SBI exchanges for each of the first set of network functions, deriving for each of the first set of network functions a threshold volume of SBI signalling exchanges based on the first information, subscribing to a Management Data Analytics Service, MDAS, to receive fault information for a second set of network functions, upon receiving fault information from the MDAS for one of the second set of network functions, monitoring via said OAM service a volume of SBI signalling

exchanges for one of the second set of network functions to determine whether or not the associated threshold volume is reached, and providing an analytic output for the one of the first network functions in response to the threshold volume being reached.

[0026] In some implementations of the method and apparatuses described herein, rather than or in addition to subscribing to the OAM service, the processor may be configured to cause a network node to collect the first information from one or more NFs and / or a Service Communication Proxy, SCP.

BRIEF DESCRIPTION OF THE DRAWINGS

[0027] Figure 1 illustrates various models A-D for providing communication between a pair of Network Functions (NFs) within an SBA architecture;

[0028] Figure 2 illustrates the provision of Management Data Analytics (MDA) in a management plane and which acts as an enabler of automation and intelligence for services management and orchestration;

[0029] Figure 3 illustrates schematically an example of a wireless communications system;

[0030] Figure 4 illustrates potential NFs impacted to allow a NWDAF to determine analytics for SBI signalling;

[0031] Figure 5 illustrates a procedure for providing output analytics;

[0032] Figure 6 illustrates an example of a network equipment (NE) 300 in accordance with aspects of the present disclosure;

[0033] Figure 7 illustrates an example processor 400 in accordance with aspects of the present disclosure; and

[0034] Figure 8 illustrates a flowchart of a method performed by a NE in accordance with aspects of the present disclosure.

DETAILED DESCRIPTION

[0035] The 5G Service-Based Architecture, SBA, architecture relies on HTTP/2 protocol and RESTful APIs to exchange information between Network Functions, NFs. Any NF can

access a service of another NF by utilizing the SBI interface. This creates a scaling issue as it has been observed by network operators that control plane signalling traffic using the HTTP/2 protocol has increased exponentially.

[0036] While the SBA architecture provides many benefits, operators are concerned about the number of request-responses required in order to exchange information between NFs via the SBI interfaces. This has the potential to generate SBI signalling “storms” where sudden excessive volumes of SBI signalling give rise to service failures and/or a deterioration of service levels. Techniques such as the dropping of service requests and the throttling (buffering) of requests may not fully overcome this problem.

[0037] 3GPP has done some work to address the issue of signalling storm traffic scenarios by introducing the Service Communication Proxy (SCP). The SCP acts as intermediary function able to receive a number of control plane service requests and route them to an appropriate NF according to the NF consumer request as shown in Figure 1.

[0038] The Operation, Administration and Management, OAM, service is already able to provide statistics with performance measurements for each 5G core network function. The performance measurements provide statistics on the number of requests received or sent by a NF by/to a different NF. This information is currently used for statistics and the OAM cannot currently provide predictions on the number of requests expected to be received or sent by/from a NF.

[0039] Analytics are also generated at the management plane. Management Data Analytics (MDA) is an enabler of automation and intelligence for services management and orchestration. An MDA MnS (also referred to as MDAS) defined in 3GPP TS 28.102 v2.0.0 enables any authorized consumer to request and receive analytics. The architecture is shown in Figure 2.

[0040] A management function (MDAF) may perform the roles of MDA MnS producer, MDA MnS consumer, other MnS consumer, Network Data Analytics Functions (NWDAF) consumer and LMF service consumer, and may also interact with other non-3GPP management systems.

[0041] The internal business logic related to MDA leverages the current and historical data related to:

Performance Measurements (PM) as per TS 28.552 and Key Performance Indicators (KPIs) as per TS 28.554.

Trace data, including MDT/RLF/RCEF, as per TS 32.422 and TS 32.423.

QoE and service experience data as per TS 28.405 and TS 28.406.

Analytics data offered by NWDAF as per TS 23.288 including 5GC data and external web/app-based information (e.g., web crawler that provides online news) from AF.

Alarm information and notifications as per TS 28.532.

CM information and notifications.

UE location information provided by LMF as per TS 23.273.

MDA reports from other MDA MnS producers.

Management data from non-3GPP systems.

[0042] Analytics output from the MDA internal business logic are made available by the MDAFs playing the role of MDA MnS producers to the authorized consumers, (including but not limited to other management functions, network functions/entities, NWDAF, SON functions, optimization tools and human operators).

[0043] The analytics provided by MDAS include analytics for fault management predictions/statistics or generally the MDA can assist in fault management. The MDA can supervise the status of various network functions and resources, and predict the running trend of network and potential failures to intervene in advance. These predictions can be used by the management system to autonomously maintain the health of the network, e.g., speedy recovery actions on a network function related to the predicted potential failure.

[0044] As part of the analytic output the information elements shown in Table 1 below can be provided by the MDAF.

[0045] It is advantageous to be able to generate and provide an analytic output for a given NF or set of NFs that is indicative of some predefined threshold volume of SBI signalling being reached. Such output may for example provide an alert to some function or service that can be used to initiate some preemptive action to avoid or at least mitigate the effects of an SBI signalling storm.

[0046] Aspects of the present disclosure are described in the context of a wireless communications system.

[0047] Figure 3 illustrates an example of a wireless communications system 100 in accordance with aspects of the present disclosure. The wireless communications system 100 may include one or more NE 102, one or more UE 104, and a core network (CN) 106. The wireless communications system 100 may support various radio access technologies. In some implementations, the wireless communications system 100 may be a 4G network, such as an LTE network or an LTE-Advanced (LTE-A) network. In some other implementations, the wireless communications system 100 may be a NR network, such as a 5G network, a 5G-Advanced (5G-A) network, or a 5G ultrawideband (5G-UWB) network. In other implementations, the wireless communications system 100 may be a combination of a 4G network and a 5G network, or other suitable radio access technology including Institute of Electrical and Electronics Engineers (IEEE) 802.11 (Wi-Fi), IEEE 802.16 (WiMAX), IEEE 802.20. The wireless communications system 100 may support radio access technologies beyond 5G, for example, 6G. Additionally, the wireless communications system 100 may support technologies, such as time division multiple access (TDMA), frequency division multiple access (FDMA), or code division multiple access (CDMA), etc.

[0048] The one or more NEs 102 may be dispersed throughout a geographic region to form the wireless communications system 100. One or more of the NEs 102 described herein may be or include or may be referred to as a network node, a base station, a network element, a network function, a network entity, a radio access network (RAN), a NodeB, an eNodeB (eNB), a next-generation NodeB (gNB), or other suitable terminology. An NE 102 and a UE 104 may communicate via a communication link, which may be a wireless or wired connection. For example, an NE 102 and a UE 104 may perform wireless communication (e.g., receive signaling, transmit signaling) over a Uu interface.

[0049] An NE 102 may provide a geographic coverage area for which the NE 102 may support services for one or more UEs 104 within the geographic coverage area. For example, an NE 102 and a UE 104 may support wireless communication of signals related to services (e.g., voice, video, packet data, messaging, broadcast, etc.) according to one or multiple radio access technologies. In some implementations, an NE 102 may be moveable, for example a satellite associated with a non-terrestrial network (NTN). In some implementations, different geographic coverage areas 112 associated with the same or different radio access

technologies may overlap, but the different geographic coverage areas may be associated with different NEs 102.

[0050] The one or more UEs 104 may be dispersed throughout a geographic region of the wireless communications system 100. A UE 104 may include or may be referred to as a remote unit, a mobile device, a wireless device, a remote device, a subscriber device, a transmitter device, a receiver device, or some other suitable terminology. In some implementations, the UE 104 may be referred to as a unit, a station, a terminal, or a client, among other examples. Additionally, or alternatively, the UE 104 may be referred to as an Internet-of-Things (IoT) device, an Internet-of-Everything (IoE) device, or machine-type communication (MTC) device, among other examples.

[0051] A UE 104 may be able to support wireless communication directly with other UEs 104 over a communication link. For example, a UE 104 may support wireless communication directly with another UE 104 over a device-to-device (D2D) communication link. In some implementations, such as vehicle-to-vehicle (V2V) deployments, vehicle-to-everything (V2X) deployments, or cellular-V2X deployments, the communication link 114 may be referred to as a sidelink. For example, a UE 104 may support wireless communication directly with another UE 104 over a PC5 interface.

[0052] An NE 102 may support communications with the CN 106, or with another NE 102, or both. For example, an NE 102 may interface with other NE 102 or the CN 106 through one or more backhaul links (e.g., S1, N2, N2, or network interface). In some implementations, the NE 102 may communicate with each other directly. In some other implementations, the NE 102 may communicate with each other or indirectly (e.g., via the CN 106). In some implementations, one or more NEs 102 may include subcomponents, such as an access network entity, which may be an example of an access node controller (ANC). An ANC may communicate with the one or more UEs 104 through one or more other access network transmission entities, which may be referred to as a radio heads, smart radio heads, or transmission-reception points (TRPs).

[0053] The CN 106 may support user authentication, access authorization, tracking, connectivity, and other access, routing, or mobility functions. The CN 106 may be an evolved packet core (EPC), or a 5G core (5GC), which may include a control plane entity

that manages access and mobility (e.g., a mobility management entity (MME), an access and mobility management functions (AMF)) and a user plane entity that routes packets or interconnects to external networks (e.g., a serving gateway (S-GW), a Packet Data Network (PDN) gateway (P-GW), or a user plane function (UPF)). In some implementations, the control plane entity may manage non-access stratum (NAS) functions, such as mobility, authentication, and bearer management (e.g., data bearers, signal bearers, etc.) for the one or more UEs 104 served by the one or more NE 102 associated with the CN 106.

[0054] The CN 106 may communicate with a packet data network over one or more backhaul links (e.g., via an S1, N2, N2, or another network interface). The packet data network may include an application server. In some implementations, one or more UEs 104 may communicate with the application server. A UE 104 may establish a session (e.g., a protocol data unit (PDU) session, or the like) with the CN 106 via an NE 102. The CN 106 may route traffic (e.g., control information, data, and the like) between the UE 104 and the application server using the established session (e.g., the established PDU session). The PDU session may be an example of a logical connection between the UE 104 and the CN 106 (e.g., one or more network functions of the CN 106).

[0055] In the wireless communications system 100, the NEs 102 and the UEs 104 may use resources of the wireless communications system 100 (e.g., time resources (e.g., symbols, slots, subframes, frames, or the like) or frequency resources (e.g., subcarriers, carriers)) to perform various operations (e.g., wireless communications). In some implementations, the NEs 102 and the UEs 104 may support different resource structures. For example, the NEs 102 and the UEs 104 may support different frame structures. In some implementations, such as in 4G, the NEs 102 and the UEs 104 may support a single frame structure. In some other implementations, such as in 5G and among other suitable radio access technologies, the NEs 102 and the UEs 104 may support various frame structures (i.e., multiple frame structures). The NEs 102 and the UEs 104 may support various frame structures based on one or more numerologies.

[0056] One or more numerologies may be supported in the wireless communications system 100, and a numerology may include a subcarrier spacing and a cyclic prefix. A first numerology (e.g., $\mu=0$) may be associated with a first subcarrier spacing (e.g., 15 kHz) and

a normal cyclic prefix. In some implementations, the first numerology (e.g., $\mu=0$) associated with the first subcarrier spacing (e.g., 15 kHz) may utilize one slot per subframe. A second numerology (e.g., $\mu=1$) may be associated with a second subcarrier spacing (e.g., 30 kHz) and a normal cyclic prefix. A third numerology (e.g., $\mu=2$) may be associated with a third subcarrier spacing (e.g., 60 kHz) and a normal cyclic prefix or an extended cyclic prefix. A fourth numerology (e.g., $\mu=3$) may be associated with a fourth subcarrier spacing (e.g., 120 kHz) and a normal cyclic prefix. A fifth numerology (e.g., $\mu=4$) may be associated with a fifth subcarrier spacing (e.g., 240 kHz) and a normal cyclic prefix.

[0057] A time interval of a resource (e.g., a communication resource) may be organized according to frames (also referred to as radio frames). Each frame may have a duration, for example, a 10 millisecond (ms) duration. In some implementations, each frame may include multiple subframes. For example, each frame may include 10 subframes, and each subframe may have a duration, for example, a 1 ms duration. In some implementations, each frame may have the same duration. In some implementations, each subframe of a frame may have the same duration.

[0058] Additionally or alternatively, a time interval of a resource (e.g., a communication resource) may be organized according to slots. For example, a subframe may include a number (e.g., quantity) of slots. The number of slots in each subframe may also depend on the one or more numerologies supported in the wireless communications system 100. For instance, the first, second, third, fourth, and fifth numerologies (i.e., $\mu=0$, $\mu=1$, $\mu=2$, $\mu=3$, $\mu=4$) associated with respective subcarrier spacings of 15 kHz, 30 kHz, 60 kHz, 120 kHz, and 240 kHz may utilize a single slot per subframe, two slots per subframe, four slots per subframe, eight slots per subframe, and 16 slots per subframe, respectively. Each slot may include a number (e.g., quantity) of symbols (e.g., OFDM symbols). In some implementations, the number (e.g., quantity) of slots for a subframe may depend on a numerology. For a normal cyclic prefix, a slot may include 14 symbols. For an extended cyclic prefix (e.g., applicable for 60 kHz subcarrier spacing), a slot may include 12 symbols. The relationship between the number of symbols per slot, the number of slots per subframe, and the number of slots per frame for a normal cyclic prefix and an extended cyclic prefix may depend on a numerology. It should be understood that reference to a first numerology

(e.g., $\mu=0$) associated with a first subcarrier spacing (e.g., 15 kHz) may be used interchangeably between subframes and slots.

[0059] In the wireless communications system 100, an electromagnetic (EM) spectrum may be split, based on frequency or wavelength, into various classes, frequency bands, frequency channels, etc. By way of example, the wireless communications system 100 may support one or multiple operating frequency bands, such as frequency range designations FR1 (410 MHz – 7.125 GHz), FR2 (24.25 GHz – 52.6 GHz), FR3 (7.125 GHz – 24.25 GHz), FR4 (52.6 GHz – 114.25 GHz), FR4a or FR4-1 (52.6 GHz – 71 GHz), and FR5 (114.25 GHz – 300 GHz). In some implementations, the NEs 102 and the UEs 104 may perform wireless communications over one or more of the operating frequency bands. In some implementations, FR1 may be used by the NEs 102 and the UEs 104, among other equipment or devices for cellular communications traffic (e.g., control information, data). In some implementations, FR2 may be used by the NEs 102 and the UEs 104, among other equipment or devices for short-range, high data rate capabilities.

[0060] FR1 may be associated with one or multiple numerologies (e.g., at least three numerologies). For example, FR1 may be associated with a first numerology (e.g., $\mu=0$), which includes 15 kHz subcarrier spacing; a second numerology (e.g., $\mu=1$), which includes 30 kHz subcarrier spacing; and a third numerology (e.g., $\mu=2$), which includes 60 kHz subcarrier spacing. FR2 may be associated with one or multiple numerologies (e.g., at least 2 numerologies). For example, FR2 may be associated with a third numerology (e.g., $\mu=2$), which includes 60 kHz subcarrier spacing; and a fourth numerology (e.g., $\mu=3$), which includes 120 kHz subcarrier spacing.

[0061] Considering now a mechanism to address the issue of SBI signalling storms, the following steps are proposed.

[0062] A. A consumer requests analytics identified by a specific Event ID from the NWDAF to provide statistics or predictions of the volume (e.g. number) of extra control plane signalling expected when a "network outage event" occurs. A network outage event may happen due to an NF or a set of NFs malfunctioning in a slice, or when a specific action takes place that places some NFs out of service (e.g. when software is upgraded or migration

of data from one NF instance to another takes place). The request may include the following additional parameters:

- When a specific NF-type(s) malfunctions or is out of service (due to a network update). This is an indication for the NWDAF to provide analytics of extra SBI signalling only when the specific NF-type (e.g. SMF, AMF) malfunctions or is out of service. It may also include NF instance ID and/or an NF-set ID corresponding to a list of NFs that could malfunction or be out of service.
- A network outage event identifier. This identifies a specific event that occurs in the network where extra SBI signalling is expected to be generated. This identifier is known to the network operator and can be an identifier or a name. For example a network outage event identifier corresponding to a firmware update to a set of NFs at a specific location.
- A time of day range where a failure event occurs or is expected to occur.
- An area of interest where a failure event occurs or is expected to occur.
- Impacted NF-types from the extra SBI signalling. This field is used by the NWDAF to monitor the SBI signalling load of the indicative NF-type(s) and determine whether extra signalling load is created/received by such NFs when a failure event occurs. It may also include NF instance ID and/or NF set ID corresponding to the NF for which SBI signalling needs to be monitored. If this field is not included the NWDAF monitors all NF-types.
- A network slice where the failure event may occur. This is an indication to the NWDAF to identify failure events only for NFs serving a specific slice.
- A Service ID (e.g. Application ID, or IMS indicator). This is an indication to the NWDAF to identify a failure event if NFs are serving a specific service (i.e. IMS service).
- A percentage increase threshold of additional SBI signalling. This field is used by the NWDAF to determine when to provide analytic output to the analytics consumer.

-An expected SBI signalling load. This may include information on the number of signalling requests exchanged between NFs or NF-types in a specific period of time. This information is used by the NWDAF to determine when an SBI signalling spike occurs. This information may also be provided by the AF/OAM as input data.

B. The NWDAF as input data collects the following data to derive the requested analytics:

-OAM: Collects performance data as described in 3GPP TS 28.552 by collecting information on the number of SBI request received/requested, for example by the "impacted NF-types" and the time of day and / or area of interest. If the analytics request included no impacted NF-types the NWDAF collects data from all NFs and identifies the NFs where SBI signalling has considerably increased.

-MDAS: Collects Fault Prediction Analytics and the time it took for the NF to recover based on the specific NF-type(s) that malfunctioned according to the analytics request.

-In alternative embodiments the NWDAF may interface directly with the impacted NF types or the SCP to retrieve data on the number of SBI requests sent or received.

-NWDAF may also collect as input data the number of HTTP/2 requests sent/received to establish a session with a NF.

-NWDAF may also collect as input data the priority information of each SBI message between NFs.

-If an area of interest is included but the analytics request does not include information of a specific NF-type malfunctioning, the NWDAF may find all NFs serving the area of interest and determine whether an NF serving the area of interest malfunctions and find the extra SBI signalling introduced due to a network outage event.

C. The NWDAF derives the average SBI signalling traffic load when the network operates in nominal mode or takes as input data the expected SBI signalling load and compares it the

signalling load generated when a network outage event occurs. Output analytics may include the following:

- Information on the network outage event (e.g. NF-type failure/outage or use case context id).
- Statistics or predictions of percentage increase of SBI signalling when a network outage event occurs.
- Impacted NF-types (the NF-types and potentially the NF instance IDs that created or received additional SBI signalling).
- Priority of the SBI message: Indicates whether signalling overload was due to high or lower priority SBI messages.
- A consumer of such analytics may be an SCP. The SCP uses this information to determine how to handle the extra signalling storm generated (e.g. by throttling traffic or prioritising control plane traffic essentially to the operation of the network, e.g. prioritising discovery request signalling).
- A consumer of such analytics may also be any NF that needs to select an NF-type that usually may be overloading with signalling when a network outage event occurs (i.e. an NRF/UDM).

Figure 4 shows the potential NFs impacted to allow the NWDAF 200 to determine analytics for SBI signalling and in response to an analytics request received from a consumer 205. The NWDAF 200 may obtain data from the MDAS 201 (failure events), the SCP 202 (data input for SBI requests sent / received to NF types), the NF 203 (data input for SBI requests sent received to this NF), the OAM 204 (to obtain performance measurements for SBI requests sent received to NF types).

Figure 5 further illustrates a procedure for providing output analytics, specifically showing the following steps:

1. An analytics consumer requires analytics to determine whether an SBI signalling traffic overload will occur when an NF malfunctions or when a network outage event occurs.

2. Analytics consumer sends an analytics request setting the analytic ID to "SBI signalling traffic load" and including one or more of the parameters described as Analytics request information in step A.
3. The NWDAF subscribes to collect data on the number of SBI requests sent or received to NF-types, OAM or SCP. The NWDAF monitors the impacted NF-types or NF-set ID if included in step 2.
4. Based on the data collection, the NWDAF prepares analytics on expected SBI signalling traffic load for the NF-types that the NWDAF collects data. The NWDAF may for example derive the average SBI signalling load (for the or each NF of interest) when the network operates in nominal mode (or the network operator provides expected signalling load).
5. The NWDAF subscribes for fault information from MDAS as described in 3GPP TS 28.104 taking into account the information included in step 2. The NWDAF may include in the request to the MDAS, the NF-type(s) to monitor if there is a fault if provided in step 2.
6. The NWDAF identifies whether a network outage takes place either based on fault information received from MDAS or based on network outage time of day/location area information provided in the request.
7. The NWDAF identifies NF types that experience heavy SBI signalling traffic load compared to the expected signalling traffic load as determined in step 4. The NWDAF may use as a trigger any fault information received from MDAS (information that MDAS can provide is included in Table 1), or network event outage information included in step 2 or by monitoring any unexpected increase of SBI signalling traffic load compared to the expected SBI signalling traffic load determined in step 4.
8. NWDAF derives analytics for SBI signalling traffic load.
9. NWDAF provides analytics to the consumer.
10. The Analytics consumer may take some appropriate action in response to receiving the analytics output and depending upon the contents of the output.

[0063] Considering further point 10, a consumer of such analytics may be an SCP. The SCP uses this information to determine how to handle the extra signalling storm generated (e.g. by throttling traffic or prioritising control plane traffic essentially to the operation of the

network, e.g. prioritising SBI signalling related to discovering and selecting a network function). A consumer of such analytics may also be any NF that needs to send a request to an NF-type that usually experiences heavy SBI signalling. For example, if the consumer receives an analytic output that an NRF will experience heavy signalling at a specific time of day then the consumer NF can postpone sending any signalling to that NF for any non-priority related communication.

[0064] Figure 6 illustrates an example of a NE 300 in accordance with aspects of the present disclosure and which may, for example, operate as the NWDAF. The NE 300 may include a processor 302, a memory 304, a controller 306, and a transceiver 308. The processor 302, the memory 304, the controller 306, or the transceiver 308, or various combinations thereof or various components thereof may be examples of means for performing various aspects of the present disclosure as described herein. These components may be coupled (e.g., operatively, communicatively, functionally, electronically, electrically) via one or more interfaces.

[0065] The processor 302, the memory 304, the controller 306, or the transceiver 308, or various combinations or components thereof may be implemented in hardware (e.g., circuitry). The hardware may include a processor, a digital signal processor (DSP), an application-specific integrated circuit (ASIC), or other programmable logic device, or any combination thereof configured as or otherwise supporting a means for performing the functions described in the present disclosure.

[0066] The processor 302 may include an intelligent hardware device (e.g., a general-purpose processor, a DSP, a CPU, an ASIC, an FPGA, or any combination thereof). In some implementations, the processor 302 may be configured to operate the memory 304. In some other implementations, the memory 304 may be integrated into the processor 302. The processor 302 may be configured to execute computer-readable instructions stored in the memory 304 to cause the NE 300 to perform various functions of the present disclosure.

[0067] The memory 304 may include volatile or non-volatile memory. The memory 304 may store computer-readable, computer-executable code including instructions when executed by the processor 302 cause the NE 300 to perform various functions described herein. The code may be stored in a non-transitory computer-readable medium such the

memory 304 or another type of memory. Computer-readable media includes both non-transitory computer storage media and communication media including any medium that facilitates transfer of a computer program from one place to another. A non-transitory storage medium may be any available medium that may be accessed by a general-purpose or special-purpose computer.

[0068] In some implementations, the processor 302 and the memory 304 coupled with the processor 302 may be configured to cause the NE 300 to perform one or more of the functions described herein (e.g., executing, by the processor 302, instructions stored in the memory 304). For example, the processor 302 may support wireless communication at the NE 300 in accordance with examples as disclosed herein.

[0069] The NE 300 may be configured to support a means for operating as a network node configured to receive an analytics request relating to Service Based Interface, SBI, signalling exchanges generated or received for each of a first set of network functions; subscribe via an Operation, Administration and Management, OAM, service to receive first information on a volume of SBI exchanges for each of the first set of network functions; derive for each of the first set of network functions a threshold volume of SBI signalling exchanges based on the first information; subscribe to a Management Data Analytics Service, MDAS, to receive fault information for a second set of network functions; upon receiving fault information from the MDAS for one of the second set of network functions, monitor via the OAM service a volume of SBI signalling exchanges for one of the first set of network functions to determine whether or not the associated threshold volume is reached; and provide an analytic output for the one of the first network functions in response to the threshold volume being reached.

[0070] The controller 306 may manage input and output signals for the NE 300. The controller 306 may also manage peripherals not integrated into the NE 300. In some implementations, the controller 306 may utilize an operating system such as iOS®, ANDROID®, WINDOWS®, or other operating systems. In some implementations, the controller 306 may be implemented as part of the processor 302.

[0071] In some implementations, the NE 300 may include at least one transceiver 308. In some other implementations, the NE 300 may have more than one transceiver 308. The

transceiver 308 may represent a wireless transceiver. The transceiver 308 may include one or more receiver chains 310, one or more transmitter chains 312, or a combination thereof.

[0072] A receiver chain 310 may be configured to receive signals (e.g., control information, data, packets) over a wireless medium. For example, the receiver chain 310 may include one or more antennas for receive the signal over the air or wireless medium. The receiver chain 310 may include at least one amplifier (e.g., a low-noise amplifier (LNA)) configured to amplify the received signal. The receiver chain 310 may include at least one demodulator configured to demodulate the receive signal and obtain the transmitted data by reversing the modulation technique applied during transmission of the signal. The receiver chain 310 may include at least one decoder for decoding the processing the demodulated signal to receive the transmitted data.

[0073] A transmitter chain 312 may be configured to generate and transmit signals (e.g., control information, data, packets). The transmitter chain 312 may include at least one modulator for modulating data onto a carrier signal, preparing the signal for transmission over a wireless medium. The at least one modulator may be configured to support one or more techniques such as amplitude modulation (AM), frequency modulation (FM), or digital modulation schemes like phase-shift keying (PSK) or quadrature amplitude modulation (QAM). The transmitter chain 312 may also include at least one power amplifier configured to amplify the modulated signal to an appropriate power level suitable for transmission over the wireless medium. The transmitter chain 312 may also include one or more antennas for transmitting the amplified signal into the air or wireless medium.

[0074] Figure 7 illustrates an example of a processor 400 in accordance with aspects of the present disclosure. The processor 400 may be an example of a processor configured to perform various operations in accordance with examples as described herein. The processor 400 may include a controller 402 configured to perform various operations in accordance with examples as described herein. The processor 400 may optionally include at least one memory 404, which may be, for example, an L1/L2/L3 cache. Additionally, or alternatively, the processor 400 may optionally include one or more arithmetic-logic units (ALUs) 406. One or more of these components may be in electronic communication or otherwise coupled

(e.g., operatively, communicatively, functionally, electronically, electrically) via one or more interfaces (e.g., buses).

[0075] The processor 400 may be a processor chipset and include a protocol stack (e.g., a software stack) executed by the processor chipset to perform various operations (e.g., receiving, obtaining, retrieving, transmitting, outputting, forwarding, storing, determining, identifying, accessing, writing, reading) in accordance with examples as described herein. The processor chipset may include one or more cores, one or more caches (e.g., memory local to or included in the processor chipset (e.g., the processor 400) or other memory (e.g., random access memory (RAM), read-only memory (ROM), dynamic RAM (DRAM), synchronous dynamic RAM (SDRAM), static RAM (SRAM), ferroelectric RAM (FeRAM), magnetic RAM (MRAM), resistive RAM (RRAM), flash memory, phase change memory (PCM), and others).

[0076] The controller 402 may be configured to manage and coordinate various operations (e.g., signaling, receiving, obtaining, retrieving, transmitting, outputting, forwarding, storing, determining, identifying, accessing, writing, reading) of the processor 400 to cause the processor 400 to support various operations in accordance with examples as described herein. For example, the controller 402 may operate as a control unit of the processor 400, generating control signals that manage the operation of various components of the processor 400. These control signals include enabling or disabling functional units, selecting data paths, initiating memory access, and coordinating timing of operations.

[0077] The controller 402 may be configured to fetch (e.g., obtain, retrieve, receive) instructions from the memory 404 and determine subsequent instruction(s) to be executed to cause the processor 400 to support various operations in accordance with examples as described herein. The controller 402 may be configured to track memory address of instructions associated with the memory 404. The controller 402 may be configured to decode instructions to determine the operation to be performed and the operands involved. For example, the controller 402 may be configured to interpret the instruction and determine control signals to be output to other components of the processor 400 to cause the processor 400 to support various operations in accordance with examples as described herein. Additionally, or alternatively, the controller 402 may be configured to manage flow of data

within the processor 400. The controller 402 may be configured to control transfer of data between registers, arithmetic logic units (ALUs), and other functional units of the processor 400.

[0078] The memory 404 may include one or more caches (e.g., memory local to or included in the processor 400 or other memory, such RAM, ROM, DRAM, SDRAM, SRAM, MRAM, flash memory, etc. In some implementations, the memory 404 may reside within or on a processor chipset (e.g., local to the processor 400). In some other implementations, the memory 404 may reside external to the processor chipset (e.g., remote to the processor 400).

[0079] The memory 404 may store computer-readable, computer-executable code including instructions that, when executed by the processor 400, cause the processor 400 to perform various functions described herein. The code may be stored in a non-transitory computer-readable medium such as system memory or another type of memory. The controller 402 and/or the processor 400 may be configured to execute computer-readable instructions stored in the memory 404 to cause the processor 400 to perform various functions. For example, the processor 400 and/or the controller 402 may be coupled with or to the memory 404, the processor 400, the controller 402, and the memory 404 may be configured to perform various functions described herein. In some examples, the processor 400 may include multiple processors and the memory 404 may include multiple memories. One or more of the multiple processors may be coupled with one or more of the multiple memories, which may, individually or collectively, be configured to perform various functions herein.

[0080] The one or more ALUs 406 may be configured to support various operations in accordance with examples as described herein. In some implementations, the one or more ALUs 406 may reside within or on a processor chipset (e.g., the processor 400). In some other implementations, the one or more ALUs 406 may reside external to the processor chipset (e.g., the processor 400). One or more ALUs 406 may perform one or more computations such as addition, subtraction, multiplication, and division on data. For example, one or more ALUs 406 may receive input operands and an operation code, which determines an operation to be executed. One or more ALUs 406 be configured with a variety of logical and arithmetic circuits, including adders, subtractors, shifters, and logic gates, to process and

manipulate the data according to the operation. Additionally, or alternatively, the one or more ALUs 406 may support logical operations such as AND, OR, exclusive-OR (XOR), not-OR (NOR), and not-AND (NAND), enabling the one or more ALUs 406 to handle conditional operations, comparisons, and bitwise operations.

[0081] The processor 400 may support wireless communication in accordance with examples as disclosed herein. The processor 400 may be configured to or operable to support a means configured to receive an analytics request relating to Service Based Interface, SBI, signalling exchanges generated or received for each of a first set of network functions, subscribe via an Operation, Administration and Management, OAM, service to receive first information on a volume of SBI exchanges for each of the first set of network functions; derive for each of the first set of network functions a threshold volume of SBI signalling exchanges based on the first information; subscribe to a Management Data Analytics Service, MDAS, to receive fault information for a second set of network functions; upon receiving fault information from the MDAS for one of the second set of network functions, monitor via an OAM service a volume of SBI signalling exchanges for one of the first set of network functions to determine whether or not the associated threshold volume is reached; and provide an analytic output for the one of the first network functions in response to the threshold volume being reached.

[0082] Figure 8 illustrates a flowchart of a method in accordance with aspects of the present disclosure. The operations of the method may be implemented by a NE as described herein. In some implementations, the NE may execute a set of instructions to control the function elements of the NE to perform the described functions.

[0083] A. At 500, the method may include receiving an analytics request relating to Service Based Interface, SBI, signalling exchanges generated or received for each of a first set of network functions. The operations of 500 may be performed in accordance with examples as described herein. In some implementations, aspects of the operations of 500 may be performed by a NE as described with reference to Figure 6.

[0084] B. At 504, the method may include subscribing via an Operation, Administration and Management, OAM, service to receive first information on a volume of SBI exchanges for each of the first set of network functions. The operations of 504 may be performed in

accordance with examples as described herein. In some implementations, aspects of the operations of 504 may be performed by a NE as described with reference to Figure 6.

[0085] C. At 506, the method may include deriving for each of the first set of network functions a threshold volume of SBI signalling exchanges based on the first information. The operations of 506 may be performed in accordance with examples as described herein. In some implementations, aspects of the operations of 506 may be performed a NE as described with reference to Figure 6.

[0086] D. At 508, the method may include subscribing to a Management Data Analytics Service, MDAS, to receive fault information for a second set of network functions. The operations of 508 may be performed in accordance with examples as described herein. In some implementations, aspects of the operations of 508 may be performed a NE as described with reference to Figure 6.

[0087] E. At 510, the method may include upon receiving fault information from the MDAS for one of the second set of network functions, monitoring via an OAM service a volume of SBI signalling exchanges for one of the first set of network functions to determine whether or not the associated threshold volume is reached. The operations of 510 may be performed in accordance with examples as described herein. In some implementations, aspects of the operations of 510 may be performed a NE as described with reference to Figure 6.

[0088] F. At 512, the method may include providing an analytic output for the one of the first network functions in response to the threshold volume being reached. The operations of 512 may be performed in accordance with examples as described herein. In some implementations, aspects of the operations of 512 may be performed a NE as described with reference to Figure 6.

[0089] It should be noted that the method described herein describes A possible implementation, and that the operations and the steps may be rearranged or otherwise modified and that other implementations are possible.

[0090] The description herein is provided to enable a person having ordinary skill in the art to make or use the disclosure. Various modifications to the disclosure will be apparent to

a person having ordinary skill in the art, and the generic principles defined herein may be applied to other variations without departing from the scope of the disclosure. Thus, the disclosure is not limited to the examples and designs described herein but is to be accorded the broadest scope consistent with the principles and novel features disclosed herein.

Information element	Definition	Support qualifier	Properties
failurePredictionObject	Indication of NR cells or NFs where the failure related issues occurred or potentially occur.	M	type: DN multiplicity: 1..* isOrdered: N/A isUnique: N/A defaultValue: None isNullable: False
potentialFailureType	Indication of type of issues that can cause the failures. NOTE: The values can be defined as a list of example values: "Operational Violation", "Physical Violation" and "Time Domain Violation". See alarmType described in TS 28.532.	M	type: string multiplicity: 1 isOrdered: N/A isUnique: N/A defaultValue: None isNullable: False
eventTime	This field holds the time of potential failure predicted. Examples: "20:15:00", "20:15:00-08:00" (for 8 hours behind UTC).	M	type: DateTime multiplicity: 1 isOrdered: N/A isUnique: N/A defaultValue: None isNullable: False
issueID	This field holds the ID of this failure prediction which is reported. When reports, this identifier can be used to provide the information to management system to maintain.	M	type: string multiplicity: 1 isOrdered: N/A isUnique: N/A defaultValue: None isNullable: False
perceivedSeverity	This field holds the value to indicate relative level of urgency for operator attention. NOTE: the value can be Critical, Major, Minor, Warning, Indeterminate, Cleared, see ITU-T Recommendation X.733.	M	type: ENUM multiplicity: 1 isOrdered: N/A isUnique: N/A defaultValue: None isNullable: False

Table 1: Analytics output for fault prediction analysis (source 3GPP TS 28.104)

CLAIMS

1. A network node comprising:
at least one memory; and
at least one processor coupled with the at least one memory and configured to cause the network node to:
 - receive an analytics request relating to Service Based Interface, SBI, signalling exchanges generated or received for each of a first set of network functions;
 - subscribe via an Operation, Administration and Management, OAM, service to receive first information on a volume of SBI exchanges for each of the first set of network functions;
 - derive for each of the first set of network functions a threshold volume of SBI signalling exchanges based on the first information;
 - subscribe to a Management Data Analytics Service, MDAS, to receive fault information for a second set of network functions;
 - upon receiving fault information from the MDAS for one of the second set of network functions, monitor via said OAM service a volume of SBI signalling exchanges for one of the first set of network functions to determine whether or not the associated threshold volume is reached; and
 - provide an analytic output for the one of the first network functions in response to the threshold volume being reached.
2. A network node according to claim 1, wherein said threshold volume is a ratio of an expected number of SBI signalling exchanges.
3. A network node according to claim 2, wherein the expected number of SBI signalling exchanges is included in the analytics request.
4. A network node according to any one of the preceding claims, wherein the analytics request includes network event information indicating an event taking place that will cause a network outage.

5. A network node according to claim 4, wherein the network event information includes one or more of an area of interest, a validity time, a first set of NF information, and a second set NF information.
6. A network node according to any one of the preceding claims, wherein said analytic output includes one or more of:
 - information relating to a network outage event;
 - statistics or predictions of a percentage increase in SBI signalling when a network outage event occurs;
 - impacted NF-types;
 - impacted NF instances; and
 - priority of the SBI message.
7. A network node according to any one of the preceding claims, wherein said second set of network functions includes one or more of the network functions of the first set of network functions.
8. A network node according to any one of the preceding claims, the network node providing a Network Data Analytics Function, NWDAF.
9. A network node according to any one of the preceding claims, wherein said analytics request explicitly identifies said first set of network functions, or identifies parameters that implicitly identify said first set of network functions.
10. A network node according to any one of the preceding claims, the network node being a node of a 5G core network.
11. A network node according to any one of the preceding claims, wherein a threshold volume is a indicative of a volume of SBI signalling exchanges greater than an expected volume during normal operation of the associated NF or an associated group of NFs.

12. A processor for a network node, comprising:
at least one controller coupled with at least one memory and configured to cause the processor to:

receive an analytics request relating to Service Based Interface, SBI, signalling exchanges generated or received for each of a first set of network functions;
subscribe via an Operation, Administration and Management, OAM, service to receive first information on a volume of SBI exchanges for each of the first set of network functions;
derive for each of the first set of network functions a threshold volume of SBI signalling exchanges based on the first information;
subscribe to a Management Data Analytics Service, MDAS, to receive fault information for a second set of network functions;
upon receiving fault information from the MDAS for one of the second set of network functions, monitor via an OAM service a volume of SBI signalling exchanges for one of the first set of network functions to determine whether or not the associated threshold volume is reached; and
provide an analytic output for the one of the first network functions in response to the threshold volume being reached.

13. A network node comprising:
at least one memory; and
at least one processor coupled with the at least one memory and configured to cause the network node to:

send an analytics request relating to Service Based Interface, SBI, signalling exchanges generated or received for each of a first set of network functions to a Network Data Analytics Functions, NWDAF; and
receive in response from the NWDAF an analytic output indicating that a threshold volume of SBI signalling exchanges for one or more of said first set of network functions has been reached.

14. A network node according to claim 13, wherein the network node implements a Service Communication Proxy, SCP, the processor being configured to use said analytic output to throttle or prioritise control plane traffic of the network.
15. A network node according to claim 13, wherein the network node implements a Network Function, NF, the processor being configured use said analytics output to modify signalling to be sent to other NFs of the network.
16. A method performed by a network node and comprising:
 - receiving an analytics request relating to Service Based Interface, SBI, signalling exchanges generated or received for each of a first set of network functions;
 - subscribing via an Operation, Administration and Management, OAM, service to receive first information on a volume of SBI exchanges for each of the first set of network functions;
 - deriving for each of the first set of network functions a threshold volume of SBI signalling exchanges based on the first information;
 - subscribing to a Management Data Analytics Service, MDAS, to receive fault information for a second set of network functions;
 - upon receiving fault information from the MDAS for one of the second set of network functions, monitoring via said OAM service a volume of SBI signalling exchanges for one of the second set of network functions to determine whether or not the associated threshold volume is reached; and
 - providing an analytic output for the one of the first network functions in response to the threshold volume being reached.