Diameter Routing Agent (DRA)

Technology primer

White paper

Diameter Routing Agents (DRAs) are poised to become increasingly important as Diameter signaling traffic grows — in some cases, to the point of congestion. Heightening their importance are other factors, such as the ongoing growth of LTE, expanded roaming and IP Exchange (IPX) interconnection, network functions virtualization, as well as the Internet of Things. This white paper serves as a primer on the role and function of DRAs.

Contents

Introduction

Early adopters of Long Term Evolution (LTE) reported that the amount of Diameter signaling traffic generated across multiple network elements (NEs) caused routing and congestion problems. Sources of this traffic include:

- Home subscriber server (HSS)
- Policy and charging rules function (PCRF)
- Public data network gateway (P-GW)
- Mobility management entity (MME)
- Online charging system (OCS), and
- IP multimedia subsystem (IMS).

Strong growth in the use of smartphones and connected devices, including watches, sensors, and vehicles, drives LTE capacity growth, resulting in a dramatic rise in Diameter signaling traffic. The challenge, though, is that the deployment of Diameter-based NEs without a DRA creates a complicated control topology, as well as an inefficient and unpredictable signaling plane. This, in turn, degrades performance and the subscriber's quality of experience (QoE), in addition to increasing integration and maintenance costs for communications service providers (CSPs). In the event of a network fault or a disaster, congestion could flood the network, causing a signaling storm and a network outage.

The industry's answer to these problems is the Diameter Routing Agent (DRA). The importance of DRAs will further increase due to continued growth of LTE, expansion of roaming and IP Exchange (IPX) interconnection, network functions virtualization (NFV), and the Internet of Things (IoT).

What is a Diameter routing agent?

Diameter is a control plane message protocol that enables communication among IP NEs, replacing Remote Authentication Dial-In User Service (RADIUS), which provided authentication, authorization and accounting (AAA) for dial-up remote access.

In legacy networks, Signalling System Number 7 (SS7) performed the role that Diameter currently plays in both fixed and LTE networks. Diameter signaling messages are used for subscription, policy, charging, and control, and help provide carrier-grade services in the LTE network.

Definitions:

- The Diameter Routing Agent (DRA), as defined by the 3GPP, provides traffic management, load balancing, and session binding. A DRA is also referred to as a Diameter Signalling Controller (DSC).
- The Diameter Edge Agent (DEA), as defined by the GSMA, provides an interworking capability for Diameter and SS7, as well as vendor variants of both.
- The Diameter Interworking Function (IWF), defined by 3GPP and GSMA, enables 4G LTE networks to interoperate with 2G/3G networks.

Together, the DRA routes traffic while the DEA manages activity at the edge of the network. For its part, the IWF translates signaling between 2G/3G networks and the 4G LTE network. For simplicity, the term ‘DRA’ often encompasses all three definitions, as it does in this paper.

Diameter signaling messages help enable:

- Authentication for access to LTE mobile broadband and VoLTE communication services
- Mobility management, so the subscriber can move while registered or connected to the network
- Policy decisions and enforcement
- Subscribers’ credit control, including limits, where they can connect, and service quality.

Overview of Diameter signaling

In the early days of LTE, it quickly became apparent that a mechanism was needed to control communication between NEs. The industry selected Diameter to provide this mechanism. Diameter told NEs what privileges to allow — or not allow — for individual subscribers. From the beginning of LTE, focusing on the end-user’s context was an important characteristic of Diameter.

The next three sub-sections provide more information on Diameter’s messages, elements, and routing.

Diameter message structure

At the transport level, the Diameter protocol uses either Transmission Control Protocol/Internet Protocol (TCP/IP) or Stream Control Transmission Protocol (SCTP). Two secure variations also exist: Transport Layer Security (TLS) for TCP/IP and Datagram Transport Layer Security (DTLS) for SCTP.

Like other messages that flow on top of TCP/IP or SCTP, control plane messages have a defined structure. As shown in Figure 1, the first 8 bits of the message structure indicate the protocol version. Other information in the message structure includes: the length of the message, the command code (whether the message is a Diameter Request “R”, Proxyable “P”, Errored “E”, or Re-Transmit “T”), the application ID, the hop-by-hop ID, the end-to-end ID, as well as the Attribute-Value Pairs (AVPs).

Figure 1. Sample Diameter message structure

Bit offset	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	26	26	27	28	29	30	31
0	version								message length																							
32	R	P	E	T					command code																							
64	application ID																															
96	hop-by-hop ID																															
128	end-to-end ID																															
160	AVPs																															
...	...																															

Application identifier

For a Diameter message, the application ID is a 32-bit number that specifies the application to which the message applies. If no application is associated with the message, the application ID is 0.

The application ID also influences the command code, which is only relevant in the context of an application ID.

Hop-by-hop identifier

A Diameter message is a routable message that can travel through several intermediate NEs before arriving at its destination. When a message is transmitted, a hop-by-hop ID is generated. When the message arrives at the first hop, the message is read. If that NE is assigned to process the message, the originating NE will receive a confirmation. If the subsequent NE is not identified as the destination, it will change the hop-by-hop ID and pass the message to another NE. This collection of hop-by-hop IDs serve as a trail, in case the message needs to return to its source NE.

Diameter is designed from the ground up as a routable message protocol. As a result, the Diameter protocol does not require that NEs process all received messages. This is an improvement over RADIUS — a user datagram protocol (UDP)-based transport protocol, which requires that all messages are processed. With RADIUS, messages are sometimes proxied, but the expectation is that, if a message is received, it is processed. This protocol soon became complicated and the need for a routable protocol became apparent.

End-to-end identifier

When a message is transmitted, an end-to-end ID is generated. The end-to-end ID is an unsigned 32-bit number that, in conjunction with the attribute-value pair (AVP), detects duplicate messages. The end-to-end ID is set to a locally unique value and is never modified. As a result, the same value is returned in the corresponding reply.

Attribute-Value Pairs (AVPs)

In Diameter, data is carried within a message as a collection of AVPs. An AVP consists of multiple fields: a code, a length, some flags, and data. Some AVPs are used by the Diameter-based protocol, others are intended for an associated Diameter application, and still others are used by the higher-level end-system application that employs Diameter. The AVP is Vendor Specific (“V”), Mandatory (“M”), or Protected (“P”).

Diameter elements

Diameter is based on a peer-to-peer architecture. The protocol defines three distinct types of nodes: clients, servers, and peers (or agents). The Diameter NE that receives the user connection request, (such as a network access server), is referred to as the client. The Diameter NE that processes the request is referred to as the server. Intermediary nodes are referred to as peers or agents.

Clients

Diameter clients initiate requests by either connecting directly with servers or sending messages using intermediaries. It is the job of the intermediate agent — a DRA, for example — to determine how to send the message to the destination. Typically, clients do not handle or process messages. Examples of a Diameter client include a network access server (NAS), packet gateway (P-GW), or MME.

Servers

Servers process messages generated by clients. There are many kinds of servers, including edge access systems (EASs), online charging systems (OCSs), policy and charging rules functions (PCRF), emergency call routing services (ECRSs), and AAA systems.

Peers

Peers (or agents) are directly connected NEs between which messages are delivered. Not all messages are processed by a peer; a downstream intermediate agent —a DRA — may process the message.

Diameter routing

Routing occurs when a message is transmitted to an NE that is not a peer. When a message goes beyond the next hop, it is called a ‘routed message’. In this case, the message must specify the correct path, using one of two routing tables.

A ‘peer table’ maintains information about the next peer, including its state, (such as open, closed, and idle), as well as how its entry is configured (such as statically or dynamically).

A ‘realm-based routing table’ contains the routing and processing information of all peers. Most of the routing that takes place in Diameter is realm-based routing, meaning that the destination is specified in the message. The priority may be indicated, but this is not standards-based functionality; it is specific to the Diameter NE being used.

Diameter signaling challenges

Diameter signaling traffic is growing rapidly because of subscribers’ enthusiasm for user experiences brought on by LTE’s mobile broadband combined with new devices, things, and applications.

As a result, the key Diameter signaling challenges faced by CSPs are:

- **Simplifying the control plane design:** Increasingly complex topologies, driven by the addition of services and interconnections, can result in inefficiency, unpredictability, difficulty with multi-vendor networks, and service outages.
- **Scaling the network:** Growth in advanced broadband services and the Internet of Things (IoT) requires load balancing of Diameter traffic across NEs, or else server partitioning.
- **Securing the network for roaming between networks:** Roaming services for LTE mobile broadband and/or voice over LTE (VoLTE) require a robust, scalable, and secure control plane across the roaming infrastructure.

Solutions for Diameter signaling challenges

In response to the key Diameter signaling challenges, the following four use cases illustrate the advantages of using a DRA in the network. Inserting a DRA into the control plane helps to simplify, scale, and secure the control topology.

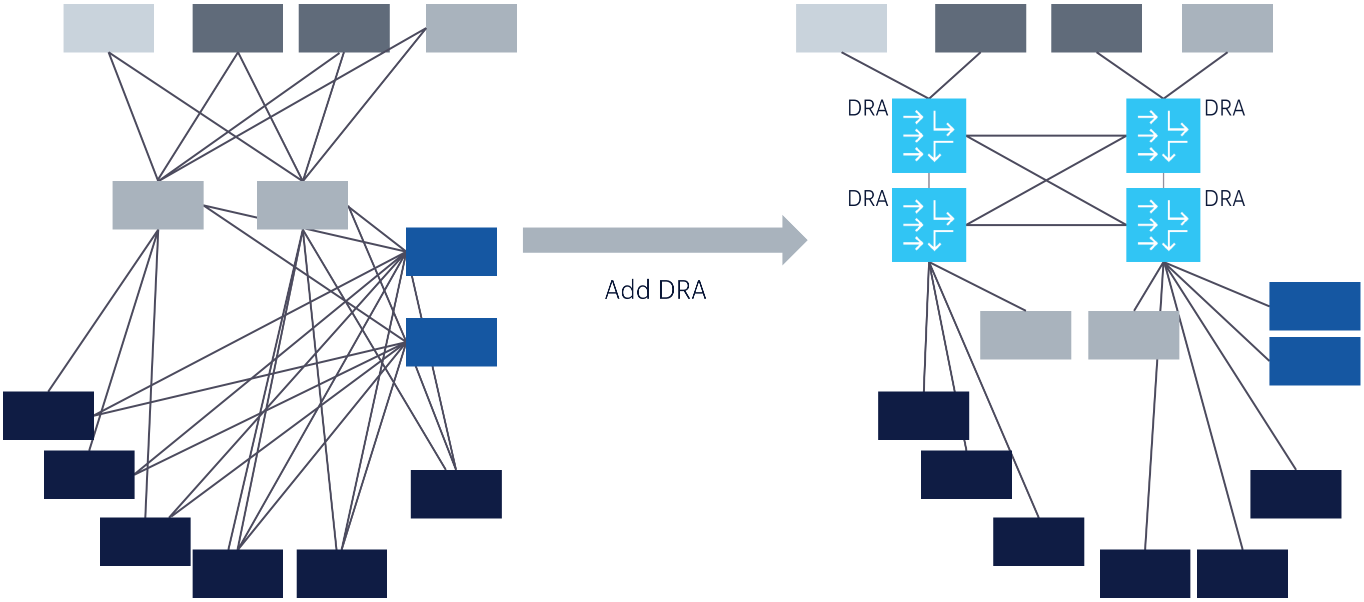
Simplify the signaling mesh

There are two aspects to simplifying Diameter signaling. The first is to remove complexity and the second is to ease multi-vendor interoperability.

Without a DRA, the signaling links between a network's many Diameter-based NEs create a complicated mesh. A DRA collapses that mesh, creating a simpler hub-and-spoke approach that:

- Reduces signaling complexity in the control plane
- Removes or collapses the complicated peer-to-peer Diameter mesh, and
- Optimizes Diameter flows across the network.

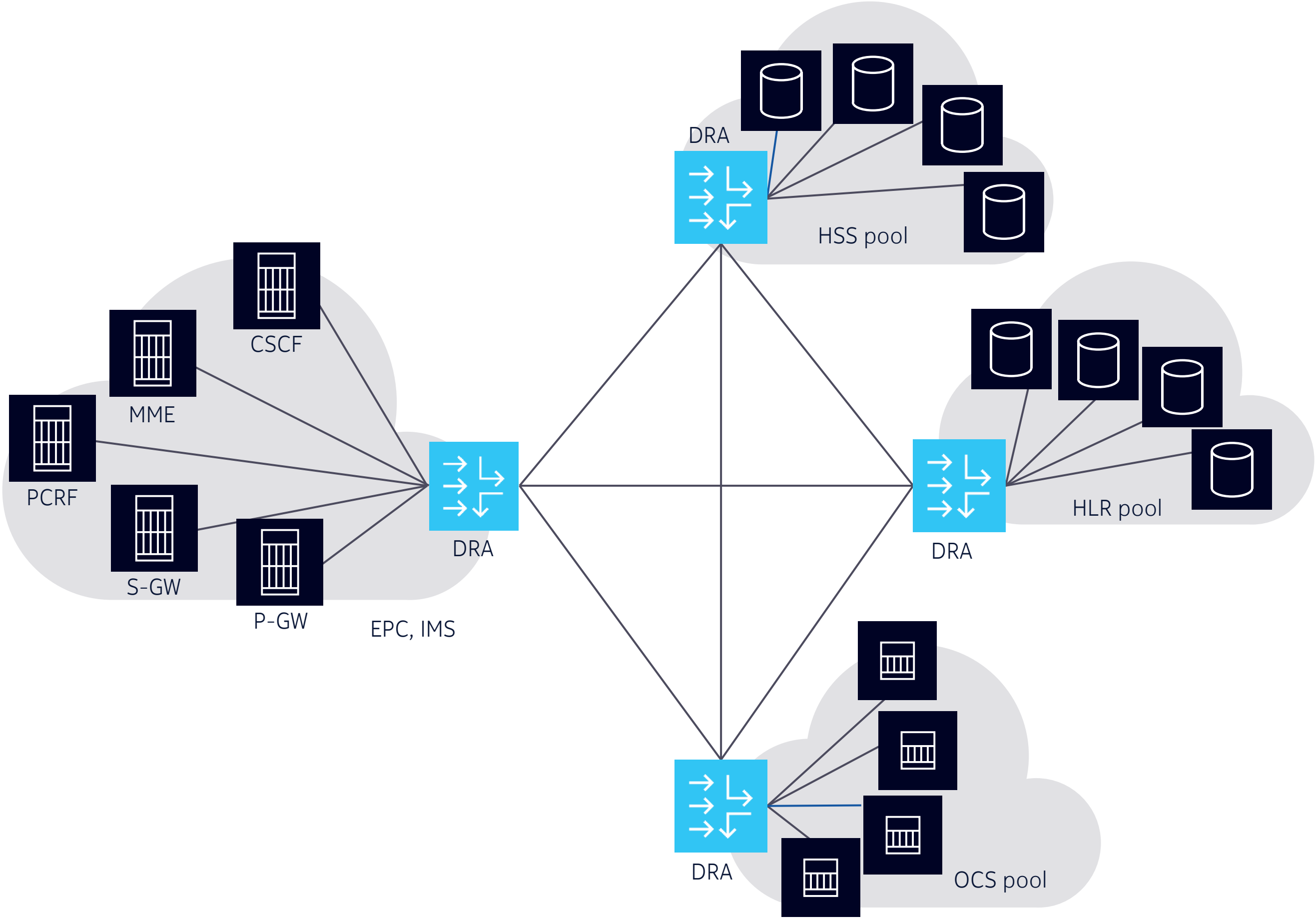
Figure 2. Simplifying the signaling mesh



A DRA helps resolve compatibility differences between network elements coming from multiple vendors, by mediating variances in how different NEs create and format Diameter messages. The DRA:

- Transforms messages to overcome issues due to different versions or interpretations of Diameter standards, and
- Reduces integration costs when deploying new NEs or performing a software update of existing NEs.

Figure 3. Simplifying multi-vendor interoperability

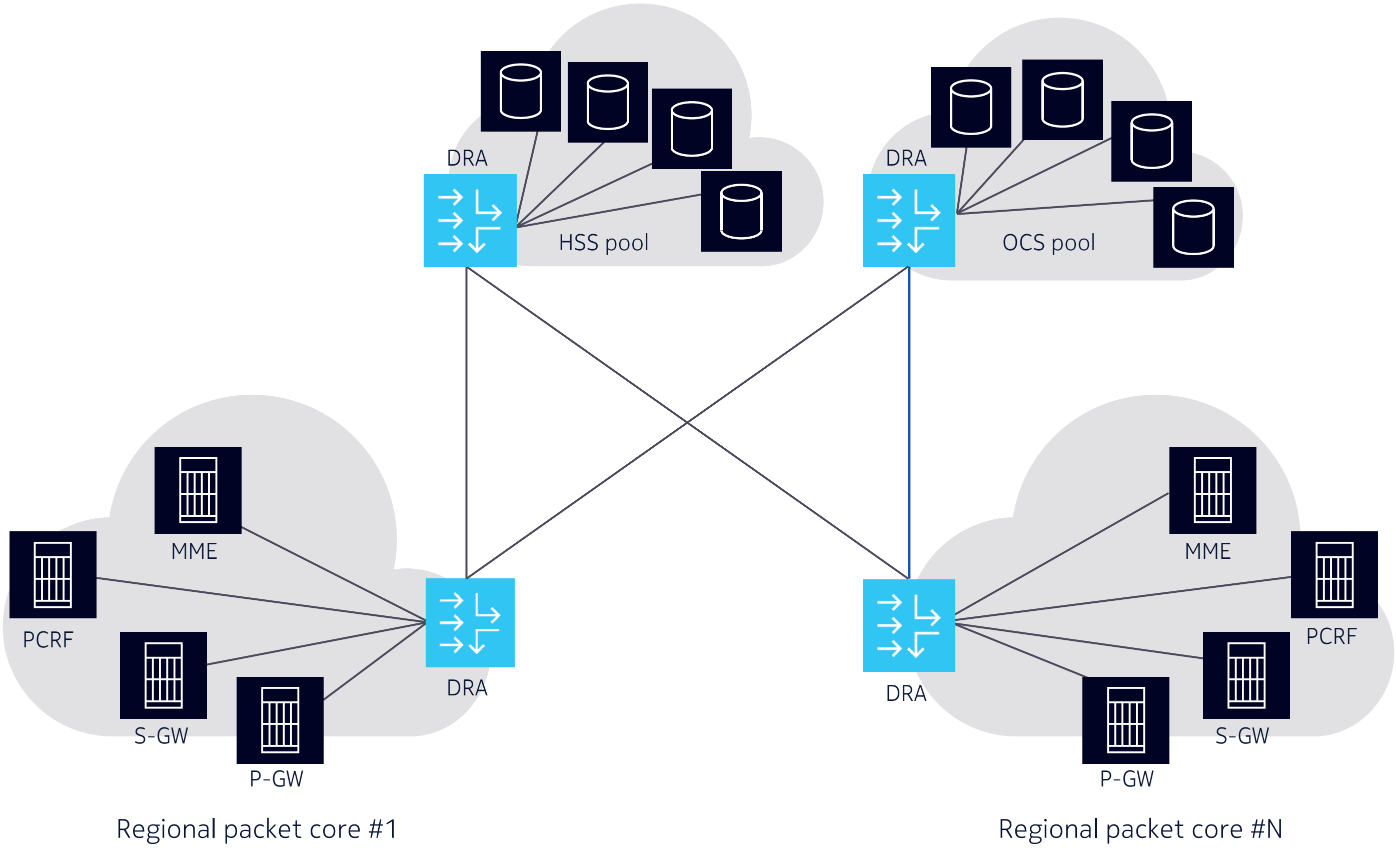


Scale the network

A DRA defines a more efficient, predictable, and scalable control plane by implementing Diameter routing capabilities, including:

- Server selection, using either partitioning (divide the load in a predictable manner) or load balancing (divide the load as needed), and
- Session management.

Figure 4. Enabling network scalability

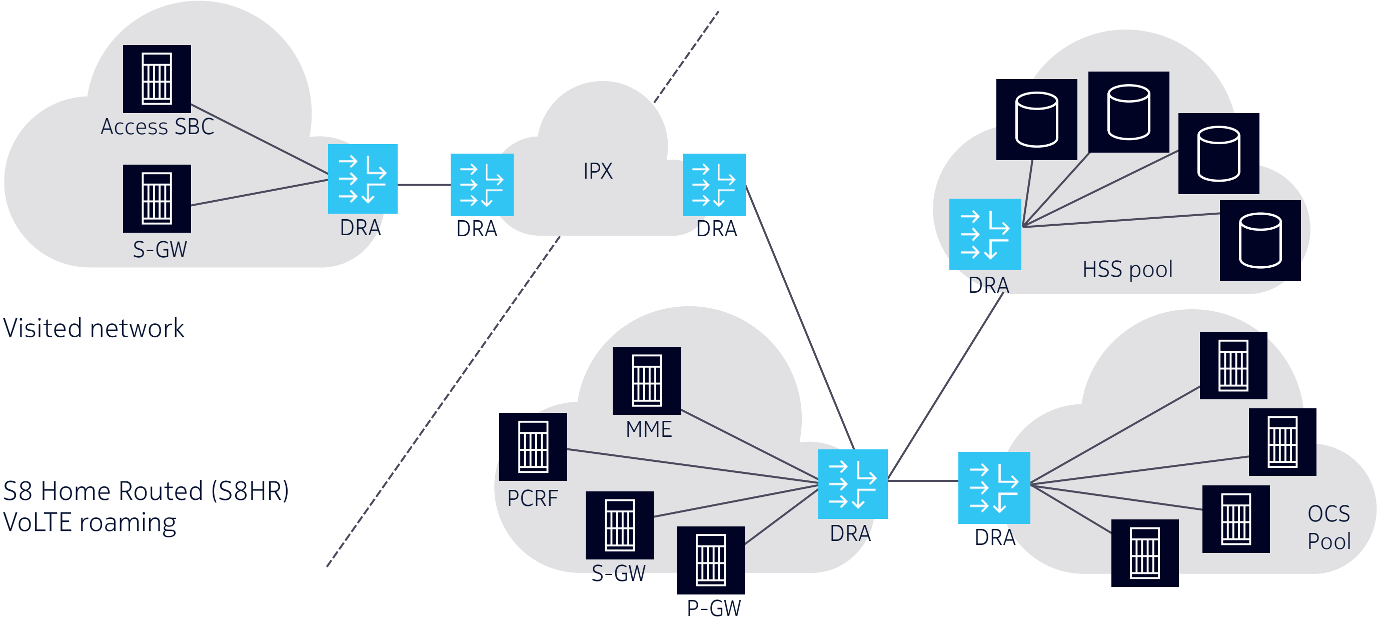


Secure roaming infrastructures

A DRA protects and secures the exchange of signaling information between your network and roaming partners, which:

- Enables complex roaming agreements and stringent performance requirements
- Ensures security by hiding your network's topology, and
- Fixes key message content, filters out unknown and unwanted data, and discards poorly formatted messages.

Figure 5. Secure roaming



The future of Diameter routing agents

With the growing popularity of LTE roaming, security and topology hiding is required between your network and roaming and IP Exchange (IPX) partners — much as the session border controller (SBC) does in IMS networks. The DRA acts as the entry point to the LTE network and supports all Diameter interfaces from the roaming subscriber's home LTE network.

Diameter signaling is already a vital component of LTE networks; however, the importance of DRAs will only increase as the amount of signaling traffic continues to grow, driven by a number of factors, including:

- NFV and distributed cloud adoption: Allows CSPs to serve different customers with different network options, such as private cloud solutions for enterprises. This pushes the control plane to the edge, which requires network traffic management and validation, interconnectivity, as well as enhanced security.
- The growth of LTE roaming: Subscribers that rely on LTE mobile broadband and/or VoLTE, when roaming from or onto the network, use capabilities enabled by inter-network Diameter messaging.
- The Internet of Things (IoT): IoT will result in more signaling and increased demand for additional security as CSPs look to control, authenticate, and verify billions of new connected devices.

Conclusion

With ubiquitous access to LTE-based personal multimedia devices and the emerging growth of IoT, subscribers' hunger for services and applications are driving the next generation of network capabilities, resulting in a huge rise in Diameter control traffic. The deployment of Diameter-based NEs creates a complicated control topology and increases risk due to an inefficient and unpredictable signaling control plane. This, in turn, can degrade performance, reduce the subscriber's quality of experience, cause outages, as well as increase integration and maintenance costs for CSPs. Inserting DRAs into the network fixes these problems by simplifying, scaling, and securing the network.

Software-based and cloud native, Nokia Cloud Signaling Director[1] (previously known as the Nokia Dynamic Diameter Engine) is a flexible, cost-effective solution that handles LTE Diameter traffic in the core, at the edge, and in interconnect scenarios. It includes service for mobile data access, VoLTE/VoWiFi roaming, IoT, and evolves to 5G services implemented with HTTP/2 and JSON signaling protocols.

Acronyms

AAA	Authentication, authorization and accounting
DRA	Diameter routing agent
IMS	IP Multimedia Subsystem
IoT	Internet of Things
IPX	IP Exchange
NE	Network element
VoLTE	Voice over LTE
VoWiFi	Voice over WiFi

[1] https://networks.nokia.com/products/cloud-signaling-director

About Nokia

We create the technology to connect the world. Powered by the research and innovation of Nokia Bell Labs, we serve communications service providers, governments, large enterprises and consumers, with the industry's most complete, end-to-end portfolio of products, services and licensing.

From the enabling infrastructure for 5G and the Internet of Things, to emerging applications in digital health, we are shaping the future of technology to transform the human experience. networks.nokia.com

Nokia Oyj
Karaportti 3
FI-02610 Espoo, Finland
Tel. +358 (0) 10 44 88 000

Document code: SR1706012276EN (February)