IBM **Guardium**
Discover and Classify

IGDC NETWORK DISCOVERY OVERVIEW

IBM GUARDIUM DISCOVER AND CLASSIFY

VERSION 4.3.3

TABLE OF CONTENTS

Table of Contents **2**
About Discovery **3**
NAE Requirements 3
Discovery Module Functions 4
Discovery Results 4
Discovery Modes **6**
PCAP (Virtual-NIC) Mode 6
DPDK Mode 6
About DPDK 6
DPDK Impact on Network Traffic Performance 6
Appendix A: NAE Specifications **8**
Appendix B: Supported Protocols **9**

ABOUT DISCOVERY

The discovery module is an IGDC service for the analysis of your network traffic that is deployed separately as a dedicated physical server - network analytical engine (NAE). At least one NAE per site is required.

NAE analyzes the network and data in motion (data traffic) and automatically discovers data repositories and network topology.

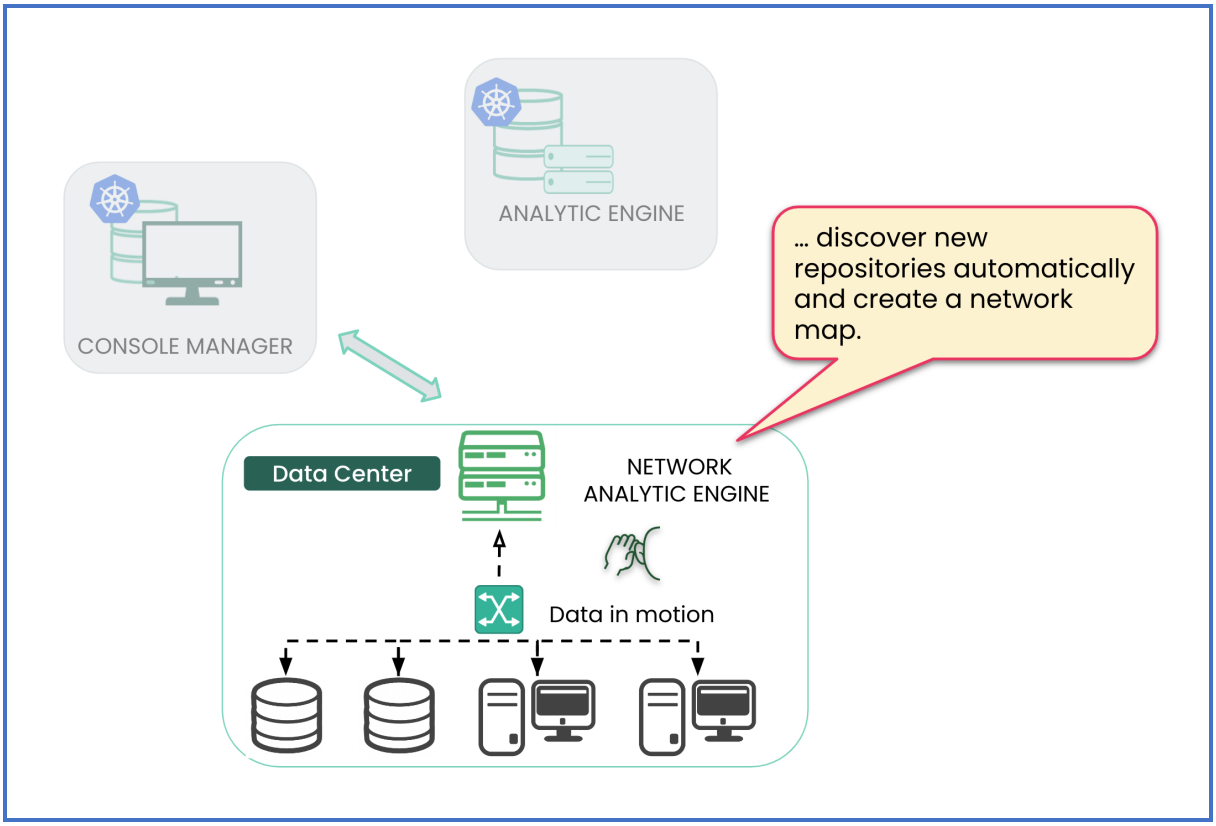


Figure 1: Discovery module capabilities

NAE REQUIREMENTS

The network analytical engine can run only on a physical server. See detailed requirements in Appendix A.

IGDC NAE stands behind Network TAP or SPAN/Mirror Port. Supported network cards:

- Ethernet Connection X722 for 10GbE SFP+
- Ethernet Controller X710/X557-AT 10GBASE-T
- VMXNET3 virtual NIC



10GBps interface vs multiple 1GBps interfaces

We recommend using the 10GBps interfaces over multiple 1GBps interfaces. If a 1GBps network card receives 10GBps traffic, NIC (network interface card) will drop 90% of the traffic, which leads to decreased accuracy and increases time to collect the real picture of the network.

DISCOVERY MODULE FUNCTIONS

- Traffic capture via interface (Virtual-NIC or psychical DPDK-NIC);
- Captured packet analysis (see supported protocols in Appendix B);
- Retrieval of sensitive data from the captured packets.

DISCOVERY RESULTS

IGDC visualizes the following information based on input from the discovery module:

1. For the network topology, the discovery module provides all connection and network entities by their protocols.

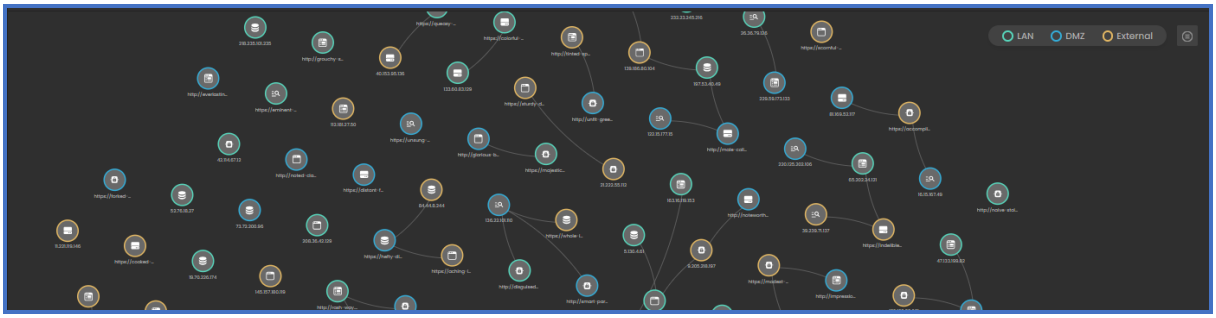


Figure 2: Network topology

2. For the data source catalog, the discovery module provides a list of potential data sources and their metadata. For example, Oracle databases, CIFS file shares, etc.

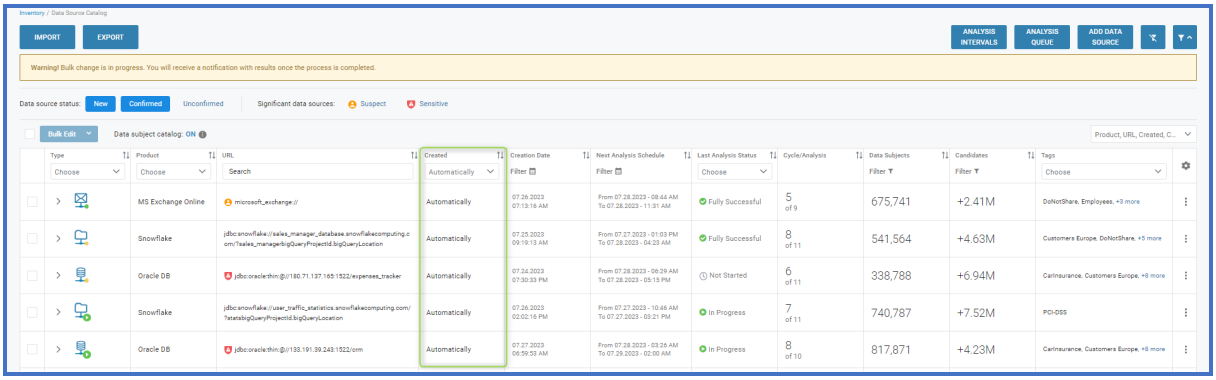


Figure 3: Automatically created data sources in the Data source catalog

3. The discovery module extracts files from network trafic for futher classification.

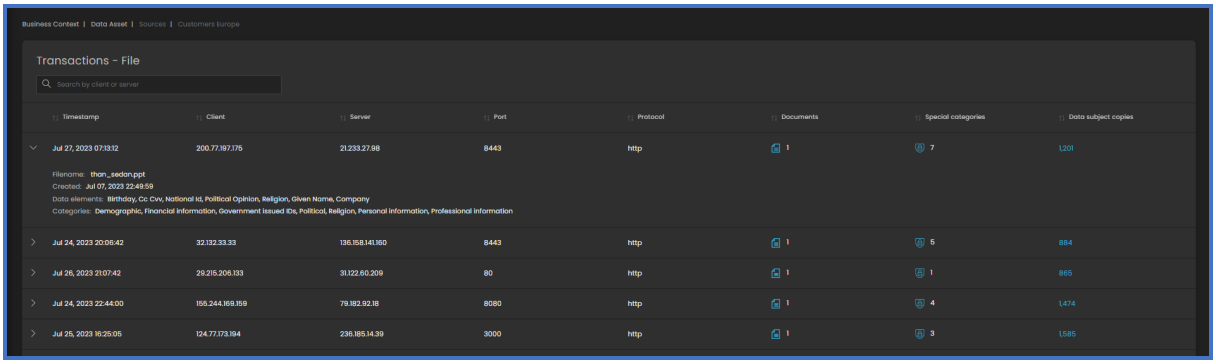


Figure 4: List of files extracted from transactions

DISCOVERY MODES

The discovery module has two modes: PCAP (Virtual-NIC) mode and DPDK mode.

PCAP (VIRTUAL-NIC) MODE

The PCAP (Virtual-NIC) is designed for testing and demo (POC) purposes. A user feed captured traffic files with **.pcap** extension to IGDC, and the discovery module will analyze the known traffic and provide output (network topology, data sources and files for further classification).

DPDK MODE

The DPDK mode is the default mode that analyzes real traffic on the fly.

The discovery module captures and analyzes traffic from all network interfaces bound to DPDK.

The discovery module uses the DPDK library to capture mirrored network traffic (packets) for further analysis. All interfaces are handled sequentially (one by one). After one cycle for all target interfaces, the discovery module accumulates statistics and sends it to the discovery controller.

ABOUT DPDK

DPDK (Data Plane Development Kit) is an open-source software library and a set of drivers that provides high-performance packet processing for network applications. It allows applications to run in userspace, bypassing the kernel for network packet processing, resulting in improved throughput and reduced latency.



For reference, see the links below:

DPDK documentation.

DPDK Programmer's Guide.

DPDK GitHub repository.

DPDK IMPACT ON NETWORK TRAFFIC PERFORMANCE

IGDC NAE stands behind Network TAP or SPAN/Mirror Port in passive mode. Duplicated network traffic is consumed from such devices for analyzing.

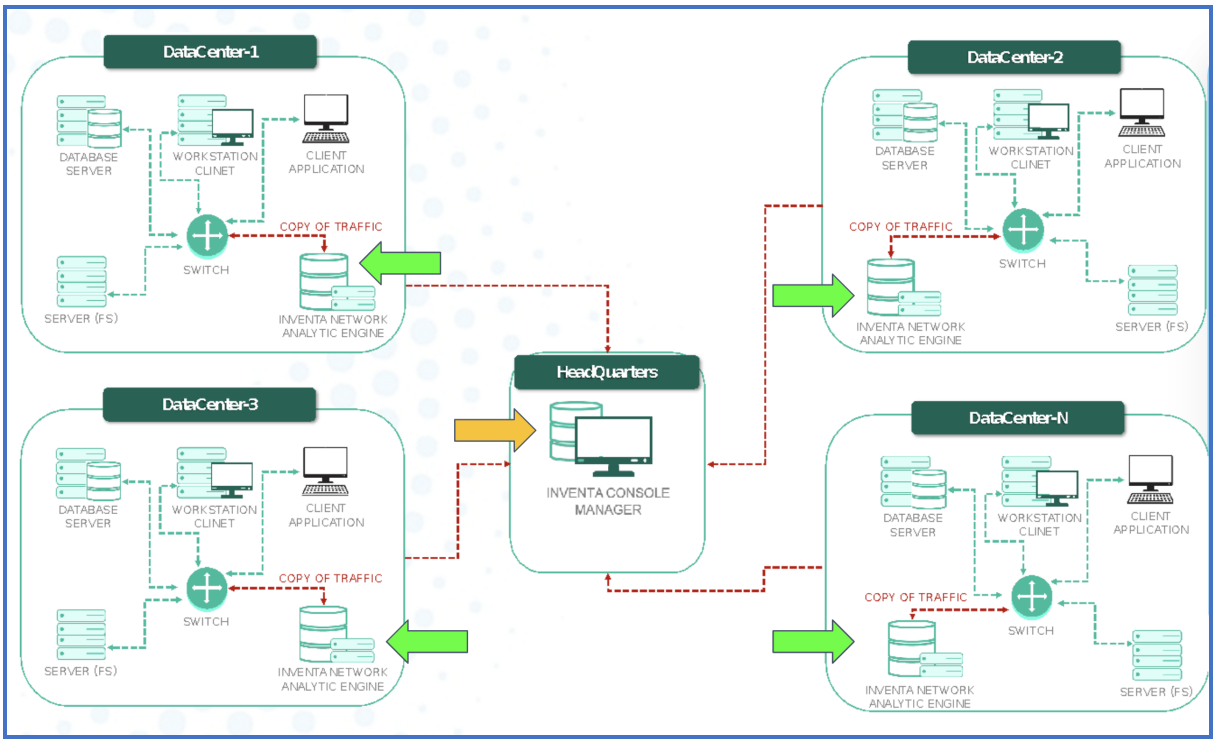


Figure 1: Architecture example with discovery

Network TAP and SPAN/Mirror Port in passive mode do not have a noticeable impact on network traffic. These methods operate by making a copy of network packets without affecting the original traffic flow.

Network TAPs are physical devices that split the network traffic into two separate streams, one for monitoring and the other for regular network traffic. As a result, there is minimal impact on network traffic since the monitoring stream is isolated from the primary network traffic. This allows for full-duplex monitoring without introducing any latency or dropping packets.

On the other hand, SPAN (Switched Port Analyzer) or mirror ports replicate network packets from one or more source ports to a designated monitoring port. While SPAN ports are more commonly available on network switches, they can introduce some degree of impact on network performance. When traffic exceeds the bandwidth limit of the monitoring port, packet drops may occur, potentially leading to decreased network performance.

The specific impact of using SPAN/Mirror ports on network traffic depends on various factors such as the switch's capabilities, available bandwidth, the amount of traffic being mirrored, and the monitoring device's capacity to handle the incoming mirrored traffic. It is recommended to check the specifications of the network switch and monitoring device for their capacity limitations.

Since both methods are passive, they do not introduce additional latency or cause network interruptions. However, the monitoring infrastructure receiving the copied traffic should have sufficient resources to handle the increased data load.

In summary, the impact on network traffic from using a network TAP or SPAN/Mirror Port in passive mode is minimal and should not affect the normal operation of the network.



Discovery has impact on network performance by sending DNS queries, which can be throttled using max parallel queries and query timeout under DNS settings.

APPENDIX A: NAE SPECIFICATIONS

Table 1: Hardware specifications

ITEM	PARAMETER		
HARDWARE	Function		Network Analytical Engine
	CPU		2x16-core (32 threads)
	MEMORY		256GB
	STORAGE		960GB
	NETWORK		Optional: 4 x 10GbE BASE-T/SFP+ 2x40Gbe BASE-T/SFP+
	HW ENCRYPTION		YES
	REMOTE MANAGEMENT		Integrated Dell Remote Access Controller 9 (iDRAC9)
	FAULT TOLERANCE		Dual, Hot-plug, Redundant Power Supply (1+1), 1100W, 4 X hot swap Disks (RAID S140)
PHYSICAL	FORM FACTOR		1U
	DIMENSIONS		H: 42.8mm W: 482.0mm D: 808.5mm
ELECTRO-MAGNETIC COMPATIBILITY	ANZ		ANZ AS/NZS 3548 Emissions
	Canada		Canada ICES-003 Emission Class A Limit
	Europe		Low Voltage Directive (2006/95/EC) and EMC Directive 2004/108/EC). The product has been marked with the CE Mark to illustrate its compliance.
	Taiwan		BSMI CNS13438 Emissions
	USA		FCC Emission Class A Limit
	International		CISPR 32 Emission, CISPR 24 Immunity, Class A Limit
ENVIRON-MENTAL TOLERANCE	TEMPERA-TURE	ASHARE Class A2 - continuous operation. +10° C to +35° C (+50° F to +95° F)	ASHARE Class A2 - continuous operation. +10° C to +35° C (+50° F to +95° F)
		-40° C to +70° C (-40° F to +158° F)-	-40° C to +70° C (-40° F to +158° F)-
	HUMIDITY (SHIPPING)		50% to 90%, non-condensing with a maximum wet bulb of 28° C (at temperatures from 25° C to 35° C)
	ALTITUDE (OPERATING)		Up to 3050m with ASHRAE class deratings

APPENDIX B: SUPPORTED PROTOCOLS

Table 1: Supported protocols and products

PROTOCOL	PRODUCT	DATA SOURCES	NETWORK TOPOLOGY	SENSITIVE DATA
TNS	Oracle Database	✓	✓	
SMB; SMB v.2	General-CIFS/SMB	✓	✓	✓
PgSQL	PostgreSQL	✓	✓	✓
TDS	Microsoft SQL Server	✓	✓	✓
MySQL	MySQL Database Server	✓	✓	✓
HTTP	General-HTTP, Apache Solr, Microsoft Dynamics, Percona XtraDB		✓	✓
Couchbase	Couchbase Server	✓	✓	✓
HTTP	Couchbase Server	✓	✓	✓
TLS	General-TLS		✓	
SMTP	General-SMTP		✓	✓
Kafka	Apache Kafka		✓	✓
DRDA	IBM DB2	✓	✓	
NFS (v.2, 3, 4)	NFS General	✓	✓	✓
MariaDB	MariaDB	✓	✓	✓
MONGO	MongoDB	✓	✓	✓
FTP	FTP		✓	✓
OneDrive	SharePoint Online and OneDrive for Business	✓	✓	
ICAP	HTTP		✓	✓

IBM, the IBM logo, and IBM Guardium Discover and Classify are trademarks or registered trademarks of International Business Machines Corporation, in the United States and/or other countries. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on ibm.com/trademark.