Graylog (GIM) Schema

Release 3.4

Jun 29, 2023

Contents

This guide is a reference for the schema used in Graylog Illuminate. We will keep this updates as changes are made, but if you feel a change should be here, please open a GitHub issue HERE.

CHAPTER 1

How To Use this Guide

Welcome to the Graylog Information Model (GIM). The Schema is broken down into sections, like “File” or “Destination” with all meta data fields below that link.

Graylog Illuminate will utilize this schema for all the content it will be creating. This guide will the be official reference, and should be used if you are creating your own content, and want it to match with our content we create.

If you take any log source, and put the user name into a field called `user_name`, then any dashboard or alert created with Illuminate will work as well.

CHAPTER 2

Graylog Template

During processing of the logs, data from the logs are inserted into Elasticsearch as "keywords", meaning they are not modified in any way, and stored as-is. This means the follow data points are unique:

`Administrator` or `administrator`

If you are doing a search in the Graylog UI, you would have to search for both of the terms, or know exactly which one to search for. Fields like `user_name` make sense to have the ability to search without worrying about the case of the word.

In order to ensure these options are accounted for, a custom analyzer has been included in the Graylog Schema template, called "loweronly". Fields normalized with "loweronly" will be converted to lowercase before the data is indexed, and search query strings for these fields will be converted to lowercase as well when ran. Pages in the schema, will list these fields as *keyword/loweronly* for reference.

2.1 Information Model: Entities

2.1.1 Alert Fields

- For messages that are an alert, such as an IDS alert
- For Vendor alert severity levels the vendor_event_severity* fields will be used

Table 1: Alert Fields

Field Name	Example Values	Field Type	Notes
alert_definitions_version	2020.1, 4092348	keyword	Version or identification value that indicates the version a collection of signatures (A/V, etc.) is in use
alert_category	malware, trojan, ransomware	keyword	Future: How do we define this field considering vendors will have their own categories? Or is that not a concern? Possibly movie this to derived fields & set only allowed values
alert_indicator	malware.exe, http://badsite	keyword	A filename, URL, packet snippet or other artifact that is related to the event that caused the alert to be generated.
alert_response_level	0, 1, 2	byte	Numeric value representing the type of action taken in response to an alert/threat. 0 = Nothing (allowed, ignored), 1 = prevent (blocked, quarantined), 2 = eradicate (deleted). This allows the use of numeric functions to detect unblocked threats where products may log multiple events for a single threat.
alert_signature		keyword	Vendor-provided Alert text description
alert_signature_id		keyword	Vendor specific unique identifier for alert signature (e.g., 1:1905345:5 for Snort signatures.)

Table 2: Derived and Enriched Fields (values will be derived or added from external sources)

Field Name	Example Values	Field Type	Notes
alert_severity	critical, high, medium, low, informational	keyword	Severity of Alert
alert_severity_level	1-5	byte	Numeric representation of the severity rating of the source message: 1 = informational, 2 = low, 3 = medium, 4 = high, 5 = critical

2.1.2 Application Fields

Table 3: Application Fields

Field Name	Example Values	Field Type	Notes
application_name	Facebook, SQL, windows_rdp	keyword (normalized:lowercase only)	Name of the application, this can be a layer 7 application name for network traffic, the name of an authenticating service/program for authentication, etc.
application_response_time		keyword	Amount of time Applications Take to give response to a request
application_sso_signonmode		keyword	For Single Sign-On (SSO) events this is the method used to access the application
application_sso_target_name		keyword	For SSO events this is the name of the application being accessed

2.1.3 Autonomous System (AS) Sub-Fields

- Autonomous System (AS) fields for the Internet (Nested as needed)
- AS fields have data referencing organization information related to an IP address

- AS fields apply to source, destination, and host entities

Table 4: Autonomous System (AS) Sub-Fields

Field Name	Example Values	Field Type	Notes
. . . _as_number	15169	keyword	Unique number. ASN identify each network on internet
. . . _as_organization	Graylog	keyword	Organization Name
. . . _as_isp		keyword	ISP associated with IP address
. . . _as_domain		keyword	Domain associated with IP address

2.1.4 Associated Fields

Table 5: Associated Fields

Field Name	Example Values	Field Type	Notes
associated_category		keyword	TBD: Not sure if this is useful
associated_hash	449afb466e043b9f3635827ce44fe3c	keyword	All associated md5,sha1,sha256,sha512,imp hashes from a log message
associated_host	10.1.2.3,corpdc01,corpdc01.corp.com	keyword	EVERYthing of any identifying host information - IP, Hostname, etc. from a log message, not implmented yet.
associated_ip	10.1.2.3,fe80:5cc3:11:4::2c	ip	Associated IP addresses for a log message
associated_mac	0cb4:44:01:a9:d1	keyword	Associated MAC addresses for a log message, colon-delimited and lower case
associated_session_id	0xa72c	keyword	Associated session IDs for a log message
associated_user	999,S-1-5-18	keyword	This will be a field that maps to all user ID values (uids, SIDs, etc.) that are associated with a user context. This can/may eventually be populated from the user framework.
associated_user_email	administrator,administrator@corp.local	keyword (normalized:lowercase only)	Any associated/alternate user ID or email, can be a set of multiple values.

2.1.5 Container Fields

Table 6: Container Fields

Field Name	Example Values	Field Type	Notes
container_id		keyword	Unique container ID
container_name		keyword	Container Name
container_namespace		keyword	Container's Namespace it is running in

2.1.6 Destination Fields

Table 7: Destination Fields

Field Name	Example Values	Field Type	Notes
destination_application_name	facebook, twitter	keyword	Describes the target application
destination_bytes_sent	203948	long	Network bytes sent by destination to the source. Some sources may present this as source bytes received, bytes received, or similar.
destination_device_model	iPad	keyword	Device Model Name
destination_device_vendor	Apple, ASUS	keyword	Device Vendor Name
destination_domain	corporate	keyword (normalized:lowercase only)	Destination domain context
destination_hostname	hostname01	keyword (normalized:lowercase only)	
destination_ip	10.1.2.3, fe80:5cc3:11:4::2c	ip	IPv4 and IPv6 addresses
destination_nat_ip	10.1.2.3, fe80:5cc3:11:4::2c	ip	translated IP address assigned by a network device performing the NAT function
destination_nat_port	2356	integer	translated network port assigned by a network device performing the NAT function
destination_os_name	iOS, Android	keyword	Operating System Name
destination_os_version	iOS 10.0	keyword	Version number of Operating System
destination_packets_sent	73458824	long	Number of packets delivered to the destination endpoint
destination_port	80, 443	integer	Service port associated with a network connection port, 0-65535
destination_service_name	smtp, ftp	keyword	The IANA-registered service name associated with the network application. Illuminate Core will use this value to define destination_port in events that have destination_ip defined, if destination_port is not already defined
destination_region	us-east-1	keyword	Name of region source device is located in
destination_serial_number	0HVX93DD	keyword	Identifying value for the destination such as a serial number
destination_type		keyword	Destination device information such as model number
destination_vm_name		keyword	Virtual system name (not to be confused with the hostname)
destination_vsys_uuid	f5398cc7-4d84-4499-84ee-d5e9246c52f8	keyword	Destination virtual system UUID
destination_zone	internal	keyword	Network zone for the destination

Table 8: Derived and Enriched Fields (values will be derived or added from external sources)

Field Name	Example Values	Field Type	Notes
destination_as_*			See: *as_* fields*
destination_category		keyword	Future: from entity mapping
destination_geo_*			See: *geo_* fields*
destination_location_name	Chicago, IL, Datacenter 01, Bismark - Finance	keyword	Field is derived either from an internal enterprise network definition or the Geo location fields if availble
destination_mac	0:b4:44:01:a9:d1	keyword	MAC address of host, colon-delimited and lower case
destination_priority	critical, high, medium, low	keyword	Future: from entity mapping
destination_priority_level	1-4	byte	Numeric value representing the priority of the destination device, 1 = low, 2 = medium, 3 = high, 4 = critical
destination_reference	IPv4, IPv6, hostname,fqdn	keyword (normalized:lower only)	Automatically mapped from the following fields: destination_ip, destination_hostname, destination_target, destination_vm_name, desination_mac

2.1.7 Email Fields

Table 9: Email Fields

Field Name	Example Values	Field Type	Notes
email_message_id		keyword	
email_subject	RE: FWD: Testing	keyword	

2.1.8 Event Fields

Table 10: Event Fields

Field Name	Example Values	Field Type	Notes
event_action	blocked, allowed, scan_start, scan_end, scan_pause, scan_cancel, scan_resume	keyword	Action that was described in a log such as a firewall log or an antivirus agent log
event_code	4624, 1	long	Numeric event defined by the vendor representing the source message type, e.g. EventCode/Event ID for Microsoft. This field is treated as a numeric value in order to support ranged queries. Any leading 0 values will be removed
event_created	2020-02-20 08:23:15.102, 1602080607	date	Date/time that the event actually occured or when the original event message was created
event_duration	10293874	long	Length of time, in seconds, for the event being described
event_end	2021-03-26T11:25:13.113	date	Date/time that event described in the log message had concluded, usually associated with an event that has a duration.
event_error_code	0xC0000008	keyword	Vendor-provided error code associated with the current message
event_error_description	ERROR_ACCESS_DENIED, Not Found	keyword	Description of error associated with the current message
event_id	0023425, 90EF8	keyword	Vendor-provided identifier representing a message type. This is similar to event_code but is instead mapped as a lateral string value. Ranged searches are not supported but the ID values will not be modified in any way.
event_log_name	security, auth.log	keyword	Reference to log, such as 'Security', 'auth.log', etc. - this differs from vendor_subtype as it refers more to the original source the log was collected from.
event_log_path	/var/log/syslog	keyword	Full path of log file source
event_observer_hostname	SERVER01.server.local, ids.local	keyword.lowercase	Hostname or FQDN of a system such as an IDS or IPS that generates an message (such as an alert) based on inspection of a thing, such as network traffic.
event_observer_id	234-id78sc	keyword	Unique ID of the Observer Device, Serial Number, etc
event_observer_ip	10.1.2.3, fe80:5cc3:11:4::2c	ip	IP address of the event observer
event_observer_uid		keyword	Unique identifier (such as a serial number or asset ID) associated with the event observer
event_received_time	2020-02-20 08:00:00, 1602080607	date	Date/time that the event was received by the reporting host. Normally applicable to logs relayed by a centralized log server.
event_repeat_count	5, 3, 9185	long	Count of times a message has been repeated
event_reporter	SERVER01.server.local	keyword.lowercase	Hostname or IP for system that delivered the message to Graylog - a WEC server, syslog collector, etc.
event_source	LAPTOP01,laptop01.corp.int	keyword.lowercase	Hostname or IP of source system that generated the event
event_source_api_version		keyword	API version of source where logs are collected via API
event_source_product	windows, linux, okta	keyword	System responsible for generating the event, e.g. "windows", "okta", etc.
event_start	2020-02-20 08:00:00, 1602080607	date	Beginning time of an event described in a log message, usually associated with an event that has a duration.
event_uid	112352564, 0122e2b3-9923-11ea-ab51-061b68b4ea16	keyword	Unique identification associated with a single event/message (e.g, "record number" from Windows event logs, a Graylog message ID)

Table 11: Derived and Enriched Fields (values will be derived or added from external sources)

Field Name	Example Values	Field Type	Notes
event_outcome	success, failure	keyword	The outcome (success/failure) of the action described by event_action.
event_severity	critical, high, medium, low, informational	keyword	This will be added by Illuminate Core if only the event_severity_level is defined. This can be mapped from vendor severity levels that do not use the same severity definitions.
event_severity_level	1-5 level	byte	Numeric representation of the severity rating of the source message: 1 = informational, 2 = low, 3 = medium, 4 = high, 5 = critical. This will be added by Illuminate core when only event_severity is defined.

2.1.9 File Fields

Table 12: File Fields

Field Name	Example Values	Field Type	Notes
file_company	Microsoft	keyword	Company name associated with a file taken from the file metadata
file_compile_time		date	Compiled date/time that a binary file was compiled
file_contents		keyword	Contents of a file
file_description	VMI	keyword	Description of file
file_is_executable	true, false	boolean	Flag indicating if file is executable
file_is_signed	true	boolean	Flag indicating if file has been digitally signed
file_name	file.zip, file.exe, file	keyword	File name, not including path
file_path	C:\temp\file.exe	keyword	Full path and file name
file_product		keyword	Product name the file was shipped with
file_product_version		keyword	Product version the file was shipped with
file_signature_status	valid	keyword	Status of file signature
file_signed_by	Microsoft Windows	keyword	Title of file signer
file_size	23894713	long	File size in bytes
file_type	gzip compressed data, application/pdf	keyword	Description of file contents
file_version	10.0.14393.4169 (rs1_release.210107-1130)	keyword	Version of file

2.1.10 Geolocation Sub-Fields

- Geo fields have data referencing location of event/host/ip
- Geo fields apply to source, destination, and host entities

Table 13: Geolocation Sub-Fields

Field Name	Example Values	Field Type	Notes
. . . _geo_city	Hamburg, Houston	keyword	City Name
. . . _geo_continent	America	keyword	Continent Name
. . . _geo_country_iso	US, DE, CA	keyword	Country ISO Alpha-2 code
. . . _geo_country	USA, Canada	keyword	Country Name
. . . _geo_coordinates	34.1186,-118.3004	keyword	Latitude, Longitude Coordinate
. . . _geo_name	Hamburg, DE	keyword	Location Name, can be derived by combining other values
. . . _geo_state	Hamburg	keyword	State name

2.1.11 GIM Fields

The gim fields are meta fields used by Graylog to assign a standard category, subcategory, and type to messages.

Table 14: GIM Fields

Field Name	Example Values	Field Type	Notes
gim_event_type_code	100000	long	This field is assigned during the normalization process. Based on this field messages will have category, subcategory, and type fields applied.

Table 15: GIM Derived fields (These fields are added to messages during the enrichment process)

Field Name	Example Values	Field Type	Notes
gim_event_category	process, audit, authentication	keyword	The category the associated log message falls under. Message categories are groupings of related messages that often have common fields.
gim_event_class	endpoint, protocol	keyword	This is an optional field that is used for related categories. For example, the process and service categories are part of the Endpoint gim_event_class, among others.
gim_event_type	network connection	keyword	A description of the event described in the associated log message.
gim_event_subcategory	subcategory, validation, process	keyword	A secondary grouping of events under a category where individual events share many common characteristics.

2.1.12 Hash Fields

Table 16: Hash Fields

Field Name	Example Values	Field Type	Notes
hash_md5	4c583e00d4710[illegible]	keyword	MD5 hash value
hash_sha1	5d4d04eff6aba84[illegible]	keyword	SHA1 hash value
hash_sha256		keyword	SHA256 hash value
hash_sha512		keyword	SHA512 hash value
hash_imphash	0b2803c4e9a2102[illegible]	keyword	IMP hash value

2.1.13 Host Fields

Table 17: Host Fields

Field Name	Example Values	Field Type	Notes
host_device	\Device\Harddisk0\Volume2	keyword	Identifier for a device (drive, network adapter) connected to a system
host_hostname	corpdc01, corpdc01.local, lab01.corpdomain.com	keyword (normalized:lowercase only)	NetBIOS or dns hostname
host_id		keyword	Host unique identifier (e.g. SID for Microsoft)
host_ip	10.1.2.3, fe80:5cc3:11:4::2c	ip	IPv4 and IPv6 addresses
host_ipv6	fe80:5cc3:11:4::2c	ip	IPv6 addresses
host_mac	02:a1:f9:c2:d5:04	keyword	MAC address of host, colon-delimited and lower case
host_reference	127.0.0.1, corpdc01, corpdc01.local, lab01.corpdomain.com	keyword (normalized:lowercase only)	Mapped from host_ip or host_hostname in that order - allows a common field to reference for messages that do not provide both (note: CIDR search will not work against this field)
host_region	us-east-1	keyword	Name of region source device is located in
host_type_version		keyword	Operating sytem version of host
host_virtfw_hostname		keyword/lowercase only	For firewalls that operate as partitioned services this is the name of the logical device
host_virtfw_id		keyword	For firewalls that operate as partitioned services this is the ID value of the logical device
host_virtfw_uid		keyword	Unique identifier such as a UUID value representing a virtual host
host_vm_name		keyword	Virtual system name (not to be confused with the hostname)

Table 18: Derived and Enriched Fields (values will be derived or added from external sources)

Field Name	Example Values	Field Type	Notes
host_as_*			See: *as_* fields*
host_category		keyword	Future: from entity mapping
host_geo_*			See: *geo_* fields*
host_location_*	Chicago, US, Datacenter 01, Bismark - Finance	keyword	Field is derived either from an internal enterprise network definition or the Geo location fields if availble
host_priority	critical, high, medium, low	keyword	Future: from entity mapping
host_priority_level	2	byte	Numeric value representing the priority of the host device, 1 = low, 2 = medium, 3 = high, 4 = critical
host_reference	IPv4,IPv6, hostname,fqdn	keyword (normalized:lower only)	Automatically mapped from the following fields: host_ip, host_hostname, host_vm_name, host_mac
host_type		keyword	Machine “type”

2.1.14 HTTP Fields

Table 19: HTTP Fields

Field Name	Example Values	Field Type	Notes
http_application	facebook	keyword	Layer 7 application name
http_bytes	29347485	Long	Sum of request + response bytes
http_content_type	application/octet-stream	keyword	Mime type of http content https://developer.mozilla.org/en-US/docs/Web/HTTP/Basics_of_HTTP/MIME_types
http_headers		keyword	Full list of http headers https://developer.mozilla.org/en-US/docs/Web/HTTP/Headers
http_host	Host: wwww.mycorp.local	keyword	host: … header from request, if present
http_referrer	http://mycorp.local/	keyword	"referer" header value if present
http_request_bytes	239478	long	SIze of request
http_request_method	GET, POST	keyword	HTTP request method from https://tools.ietf.org/html/rfc7231
http_request_path	/path/to/resource?option=test	keyword	Need to review field length/truncation at 8192 characters (consider utf-8). Some may consider the path not to include the "query" (text after the last "/") but this value may include it.
http_response_bytes	498274	long	Size of response
http_response	OK, Moved Permanently	keyword	Text response mapped from the response code https://www.w3.org/Protocols/rfc2616/rfc2616-sec6.html
http_response_code	200, 404, 500	integer	Numeric server response code
http_uri	https://www.graylog.org, https://www.graylog.org/blog, https://www.mycorp.local/workspaces/team#posts	keyword	Full request string; Need to review field length/truncation at 8192 characters (consider utf-8)
http_uri_category	Suspicious, Games	keyword	Categorization of associated web site/URL
http_uri_stem	Default.htm	keyword	The target of the request. For Example: http://www.test.com/test.jsp?hello=y the URI stem is /test.jsp
http_uri_query	hello=y	keyword	The query the client was trying to perform. Example http://www.test.com/test.jsp?hello=y the query is hello=y
http_user_agent	Mozilla/5.0 (Macintosh; Intel Mac OS X 10.15; rv:74.0) Gecko/20100101 Firefox/74.0)	keyword	User Agent string
http_user_agent_name	Firefox	keyword	Attempted identification of the browser client usually based on user agent analysis
http_user_agent_os	Windows 10	keyword	Operating System of User Agent
http_version	1.0, 1.1, 2.0	keyword	HTTP version
http_xff	X-Forwarded-For: 10.1.2.3	keyword	HTTP x-forwarded-for header value. Future: May map as IP, need to account for different ways this is presented.

Table 20: Derived and Enriched Fields (values will be derived or added from external sources)

Field Name	Example Values	Field Type	Notes
http_request_path_analyzed	** TBD		Need to review best analyzer configuration for HTTP paths / consider truncation
http_uri_analyzed	ftp://ftp01.server.internal/file.tar.gz, https://www.graylog.org, https://www.graylog.org/blog	text/standard	Optionally copied when a URL must be tokenized. Future: will have to research best analyzer config / consider truncation
http_uri_length	9283	long	String length of HTTP user agent
http_user_agent_analyzed		text/standard	This is a copy of the http_user_agent field but processed with text analysis
http_user_agent_length	54	long	String length of original user agent

2.1.15 Network Fields

Table 21: Network Fields

Field Name	Example Values	Field Type	Notes
network_application	facebook, instagram	keyword/lowercase	Application name - Facebook, etc.
network_bytes	71238	long	Bytes transferred during a connection, may be calculated by summing bytes sent/received (source_bytes_sent/destination_bytes_sent) - some vendors may report this as packet_length
network_bytes_rx			DEPRECATED - use destination_bytes_sent
network_bytes_tx			DEPRECATED - use source_bytes_sent
network_community_id		keyword	See: https://github.com/corelight/community-id-spec
network_connection_duration	00:23:45	keyword	Duration of time a network connection was established
network_connection_id	CMdzit1AMNsmfAIiQc	keyword	Unique identifier value for a network connection
network_data_bytes	71238	long	Total bytes of the data payload
network_direction	inbound, outbound, lateral	keyword	Indicates the direction of the observed network flow. Must be either inbound or outbound, this should be mapped to these values if vendors provide network direction differently .
network_forwarded_ip	10.1.2.3, fe80:5cc3:11:4::2c	ip	
network_header_bytes	71238	long	Total bytes of packet header information
network_iana_number	6, 17, 41	integer	https://www.iana.org/assignments/protocol-numbers/protocol-numbers.xhtml
network_icmp_type	echo, time exceeded	keyword	https://www.iana.org/assignments/icmp-parameters/icmp-parameters.xhtml
network_inner			TBD
network_interface_in	eth0	keyword/lowercase	Name of interface traffic receiving traffic
network_interface_out	eth0	keyword/lowercase	Name of interface traffic sending traffic
network_ip_ver	4, 6	keyword	IPv4 or IPv6
network_name			TBD
network_packets	71238	long	Count of packets transferred during a connection, may be calculated by summing packets sent/received (source_packets_sent/destination_packets_sent)
network_packets_rx			DEPRECATED - use destination_packets_sent
network_packets_tx			DEPRECATED - use source_packets_sent
network_protocol	ipv4, ipv6, icmp	keyword/lowercase	Protocol names, preferrably from the Keyword column in https://www.iana.org/assignments/protocol-numbers/protocol-numbers.xhtml
network_transport	udp, tcp	keyword/lowercase	transport layer protocol of packet/connection
network_tunnel_type	gre, ipsec	keyword/lowercase	tunnel type
network_tunnel_duration	209384	long	time in seconds for tunnel duration
network_type			TBD - maybe not needed since network_protocol

2.1.16 Policy Fields

- Related to system/device policies for operating systems, firewalls, etc.

Table 22: Policy Fields

Field Name	Example Values	Field Type	Notes
policy_id	6da61e4c-84a8-4136-900d-f86c09bb3774	keyword	Unique identifier of a policy
policy_uid		keyword	
policy_name	admin-user-template	keyword	Name of a policy

2.1.17 Process Fields

- Process is related to the execution of binaries
- The *process_* names can also be prefixed with *target_...* and *parent_...* e.g, parent_process_id, target_process_name, etc.

Table 23: Process Fields

Field Name	Example Values	Field Type	Notes
process_description	Windows Command Line Utility	keyword	Description of executed process
process_command_line	c:\tmp\runme.exe, /tmp/runme	keyword/lowercase	Full command line of executed process
process_command_line_length	29347	long	Length of of process_command_line
process_id	2045,0x3e7	keyword/lowercase	Process identifier associated with executed process
process_integrity_level	high, trusted	keyword	Integrity level of executed process
process_parent_command_line	c:\tmp\runme.exe, /tmp/runme	keyword/lowercase	Full command line of parent process
process_parent_id	2045,0x3e7	keyword/lowercase	Process identifier associated with parent process
process_parent_name	whoami, whoami.exe	keyword/lowercase	File name of parent process, excluding path
process_parent_path	C:\Windows\system32\whoami.exe, /usr/bin/whoami	keyword/lowercase	Full path of parent process
process_parent_uid	{73123815-5caa-4e39-90dc-d25d4013bf15}	keyword	GUID or unique identifier for parent process that is not the process_id
process_name	whoami, whoami.exe	keyword/lowercase	File name of executed process, excluding path
process_path	C:\Windows\system32\whoami.exe, /usr/bin/whoami	keyword/lowercase	Full path of executed process
process_target_id	2045,0x3e7	keyword	The process ID of the targeted process of some action that was taken against that process
process_target_name	whoami, whoami.exe	keyword	The name of the targeted process of some action that was taken against that process
process_target_path	C:\Windows\system32\whoami.exe, /usr/bin/whoami	keyword/lowercase	The full path and name of the targeted process of some action that was taken against that process
process_target_uid	{73123815-5caa-4e39-90dc-d25d4013bf15}	keyword	The process unuqie identifier of the targeted process of some action that was taken against that running process
process_uid	{73123815-5caa-4e39-90dc-d25d4013bf15}	keyword	GUID or unique identifier for executed process that is not the process_id
process_working_directory	C:\Windows\Temp	keyword	The current working directory that the process was called from

2.1.18 Query Fields

Table 24: Query Fields

Field Name	Example Values	Field Type	Notes
query_class	IN	keyword	Class of name query, usually IN for DNS
query_record_type	A, AAAA, MX, SRV	keyword	Record type being requested
query_record_type_code	1, 3, 5	keyword	IANA assigned code for record type being requested
query_request	www.graylog.org	keyword	Name being resolved in DNS request
query_request_length	25	Long	Length of name resolution request
query_response		keyword	Name resolution answer
query_response_length	25	Long	Length of name resolution response
query_result	NXDOMAIN, NOERROR	keyword	Status of name resolution request
query_result_code	0, 3	keyword	IANA assigned DNS RCODE

2.1.19 Rule Fields

- Related to system/device rules for operating systems, firewalls, etc.

Table 25: Rule Fields

Field Name	Example Values	Field Type	Notes
rule_id	6da61e4c-84a8-4136-900d-f86c09bb3774	keyword	Unique identifier of a rule
rule_name	admin-user-template	keyword	Name of a Rule (ex. Outbound Web Traffic)

2.1.20 Service Fields

- Service describes the service/application for which the data was collected from.

Table 26: Service Fields

Field Name	Example Values	Field Type	Notes
service_name	graylog-server.service, sshd ,graylog-sidecar	keyword	Name of service
service_version	1.0, 1054	keyword	Version Number of service or underlying application
service_state	running, started, stopped	keyword	State of service

2.1.21 Session Fields

- A network session, logon session, any kind of thing with a beginning and and end.

Table 27: Session Fields

Field Name	Example Values	Field Type	Notes
session_id		Keyword	Vendor-provided unique identifier. This can be a random alphanumeric string, a hex value, a GUID value, etc.

2.1.22 Source Fields

Table 28: Source Fields Schema

Field Name	Example Values	Field Type	Notes
source_bytes_sent	2983471	long	Network bytes sent by source, some sources may present this as source bytes tx, bytes tx or something similar.
source_device_model	iPad	keyword	Device Model Name
source_device_vendor	Apple, ASUS	keyword	Device Vendor Name
source_hostname	corp-pdc01, corp-pdc01.local, lab01.corpdomain.com	keyword (normalized:lowercase only)	NetBIOS or dns hostname, converted to lowercase
source_id	09VX93DD	keyword	Identifying value for the source such as a serial number
source_ip	10.1.2.3, fe80:5cc3:11:4::2c	ip	IPv4 and IPv6 addresses
source_ipv6	fe80:5cc3:11:4::2c	ip	Only IPv6 addresses
source_nat_ip	10.1.2.3, fe80:5cc3:11:4::2c	ip	translated IP address assigned by a network device performing the NAT function
source_nat_port	2384	integer	translated network port assigned by a network device performing the NAT function
source_os_name	IOS, Android	keyword	Operating System Name
source_os_version	iOS 10.0	keyword	Version number of Operating System
source_packets_sent	2309483	long	Count of packets sent by source
source_port	45392	integer	numeric port, 0-65535
source_port_name	ssh, ftp	keyword	The IANA-registered service name associated with the network application. Illuminate Core will use this value to define source_port in events that have source_ip defined, if source_port is not already defined.
source_region	us-east-1	keyword	Name of region source device is located in
source_type		keyword	Source device information such as model number
source_vm_name		keyword	Virtual system name (not to be confused with the hostname)
source_vsys_uuid		keyword	
source_zone		keyword	

Table 29: Derived and Enriched Fields (values will be derived or added from external sources)

Field Name	Example Values	Field Type	Notes
source_as_*			See: *as_* fields*
source_category		keyword	Future: from entity mapping
source_geo_*			See: *geo_* fields*
source_location	Chicago, US, Datacenter 01, Bismark - Finance	keyword	Field is derived either from an internal enterprise network definition or the Geo location fields if availble
source_mac	a0:b4:44:01:a9:d1	keyword	MAC address of host, colon-delimited and lower case
source_priority	critical, high, medium, low	keyword	Future: from entity mapping
source_priority_level	4	byte	Numeric value representing the priority of the source device, 1 = low, 2 = medium, 3 = high, 4 = critical
source_reference	IPv4, IPv6, hostname, fqdn	keyword (normalized:lower only)	Automatically mapped from the following fields: source_ip, source_hostname, source_vm_name, source_mac

2.1.23 Threat Fields

- Information Around Threats

Table 30: Threat Fields

Field Name	Example Values	Field Type	Notes
threat_category	malware, trojan	Keyword	
threat_detected	true, false	Keyword	Is a threat detected

2.1.24 Trace Fields

- Tracing makes it possible to track events across multiple logs on a unique ID (Micro-service, Web App)

Table 31: Trace Fields

Field Name	Example Values	Field Type	Notes
trace_id		Keyword	Unique ID of multiple events belonging together.
trace_call		Keyword	Stack trace related to process call

2.1.25 User Fields

- Possible Field Prefixes: source_* (e.g., "source_user_name") or destination_* (e.g., "destination_user_name")
- Where messages describe an action taken by one account impacting another account, the actor (account taking the action) will be described by the "source_user_*" fields and the subject (account for which the action was taken) will be described by the "user_*" fields; Examples include:

- Authentication, where the authenticating service account context is provided
- IAM events, where a user or service has performed an action that impacts a user or group

Table 32: User Fields

Field Name	Example Values	Field Type	Notes
user_command		keyword	
user_command_path		keyword	
user_domain	mycorp.internal	keyword	AD or LDAP domain
user_email	user@mycorp.internal	keyword	
user_id		keyword	Mapped to SID or UID, etc.
user_name		keyword (normalized:lowercase only)	
user_session_id	0x534, 1055	keyword	User logon session identifier

Table 33: Derived and Enriched Fields (values will be derived or added from external sources)

Field Name	Example Values	Field Type	Notes
user_category	vip, default account, finance, help desk	keyword	Future: From entity mapping
user_name_mapped	Built-in Administrators	keyword (normalized:lowercase only)	When a user identity or identities is mapped from a source outside of the message itself it is written to this field. This is where Windows well-known SIDs are resolved.
user_priority	critical, high, medium, low	keyword	Future: From entity mapping
user_priority_level	1-4	byte	Numeric value representing the priority of the user account, 1 = low, 2 = medium, 3 = high, 4 = critical
user_type	user, computer, well-known sid, group, {any vendor-provided value}	keyword	Experimental field ** This is still being researched - need to look at what winlogbeats/nxlog may provide in terms of SID resolution in different configurations, and consider different technologies use of "types"

2.1.26 Vendor Fields

- The vendor fields are to capture data provided by source, as-is
- The vendor fields are intended to capture information that is either used in the content we develop, or can be used to provide background on how a field such as event_outcome was defined

Table 34: Vendor Fields

Field Name	Example Values	Field Type	Notes
vendor_alert_severity	critical, high, medium, low	keyword	When the message is an alert this is the vendor-provided text description of the alert severity
vendor_alert_severity_level	4, 3, 2, 1	integer	When the message is an alert this is the vendor-provided numeric value for the alert severity
vendor_authentication_provider	Active Directory	keyword	Vendor defined action - Quick description of the service providing credential validation
vendor_credential_type	password, token	keyword	Vendor-defined credential type
vendor_event_action	allow, deny, pass, fail	keyword	Vendor defined action - this should be a short, typically one-word, description of what action the event is describing. The value is to be used verbatim, including case, from the source log.
vendor_event_category	Removable Media, Registry, File System	keyword	Vendor defined category of an event
vendor_event_description		keyword	Vendor defined description of the action with more detail than is included in vendor_event_action
vendor_event_outcome	block, drop, report, allow, reject	keyword	Vendor-defined result of the action defined in the message
vendor_event_outcome_reason		keyword	Vendor-provided text detailing the reason for the vendor-provided action and/or outcome the message is describing
vendor_event_severity	critical, high, medium, low, informational	keyword	Vendor-defined text description of the severity rating
vendor_event_severity_level	0, 1, 5, 10	integer	Vendor-defined numeric severity rating for this event
vendor_private_ip		ip	
vendor_private_ipv6		ip	
vendor_public_ip		ip	
vendor_public_ipv6		ip	
vendor_signin_protocol		keyword	
vendor_subtype	ipsec, dnsmasq, kernel, threat	keyword	Vendor-defined subtype of log - this differs from event_log_name as it refers more to the subject or category of log message.
vendor_threat_suspected		keyword	
vendor_transaction_id		keyword	
vendor_transaction_type		keyword	
vendor_user_type		keyword	

2.1.27 Vendor Entities

Included here are fields specific to a vendors technology, which does not fall under the common schema.

Palo Alto Fields

Table 35: Palo Alto Fields

Field Name	Example Values	Field Type	Notes
pan_alert_direction		keyword	Indicates the direction of the attack, client-to-server or server-to-client: 0—direction of the threat is client to server. 1—direction of the threat is server to client
pan_after_change_detail		keyword	This field is in custom logs only; it is not in the default format. - It contains the full xpath after the configuration change.
pan_assoc_id		keyword	Number to idetify all connections for an association between to SCTP endpoints
pan_auth_method		keyword	A string showing the authentication type, such as LDAP, RADIUS or SAML
pan_before_change_detail		keyword	This field is in custom logs only; it is not in the default format. - It contains the full xpath after the configuration change.
pan_cloud_hostname		keyword	FQDN of WildFire appliance or Cloud where file was uploaded
pan_dev_group_level_[1-4]		keyword	ID Numbers that indicate the device groups location within DG Hierarchy
pan_dynusergroup_name		keyword	Name of the dynamic user group that contains the user who initiated the session.
pan_event_name		keyword	String showing the name of the event.
pan_event_object		keyword	Name of the object associated with the system event.
pan_evidence		keyword	A summary statement that indicates how many times the host has matched against the conditions defined in the correlation object. For example, Host visited known malware URl (19 times).
pan_flags		keyword	32-bit field that provides details on session
pan_gp_client_version		keyword	The client's GlobalProtect app version.
pan_gp_connect_method		keyword	A string showing the how the GlobalProtect app connects to Gateway, (for example, on-demand or user-login)
pan_gp_error		keyword	A string showing that error that has occurred in any event.
pan_gp_error_code		keyword	An integer associated with any errors that occurred
pan_gp_error_extended		keyword	Additional information for any event that has occurred.
pan_gp_hostname		keyword	The name of the GlobalProtect portal or gateway.
pan_gp_hostid		keyword	Unique ID GlobalProtect assigns to identify the host.
pan_gp_location_name		keyword	A string showing the administrator-defined location of the GlobalProtect portal or gateway.
pan_gp_reason		keyword	A string that shows the reason for the quarantine
pan_hip		keyword	Name of the HIP object or profile.
pan_hip_type		keyword	Whether the hip field represents a HIP object or a HIP profile.
pan_http2		keyword	Identifies if traffic used an HTTP/2 Connection by displaying one of the following values: Parent session ID—HTTP/2 connection. OR. 0—SSL session
pan_link_changes		keyword	Number of link flaps during session
pan_link_switches		keyword	Contains up to four link flap entries, with each entry containing the link name, link tag, link type, physical interface, timestamp, bytes read, bytes written, link health, and link flap cause.
pan_log_action		keyword	Log Forwarding Profile Applied to Session
pan_log_panorama		keyword	A bit field indicating if the log was forwarded to Panorama
pan_log_subtype		keyword	Subtype of Given Log
pan_module		keyword	It provides additional information about the sub-system generating the log
pan_monitor_tag		keyword	IMEI 15/16 Digit number
pan_object_id		keyword	Name of the object associated with the system event.
pan_objectname		keyword	Name of the correlation object that was matched on.

Continued on next page

Table 35 – continued from previous page

Field Name	Example Values	Field Type	Notes
pan_parent_session_id		keyword	ID of the session in which this session is tunneled
pan_parent_start_time		keyword	Time the Tunnel Session began
pan_pcap_id		keyword	Packet Capture ID
pan_ppid		keyword	ID of the protocol for the payload of the data chunk
pan_sctp_chunks_sum		keyword	Sum of SCTP chunks sent and received for an association.
pan_sctp_chunks_tx		keyword	Number of SCTP chunks sent for an association.
pan_sctp_chunks_rx		keyword	Number of SCTP chunks received for an association.
pan_sdwan_cluster		keyword	Name of the SD-WAN cluster.
pan_sdwan_cluster_type		keyword	Type of cluster (mesh or hub-spoke)
pan_sdwan_device_type		keyword	Type of device (hub or branch)
pan_sdwan_policy_id		keyword	Name of the SD-WAN policy.
pan_sdwan_site_name		keyword	Name of the SD-WAN site
pan_session_end_reason			The reason the session was terminated
pan_source_region		keyword	The region for the user who initiated the session.
pan_tunnel_id		keyword	International Mobile Subscriber Identity Number
pan_tunnel_stage		keyword	A string showing the stage of the connection (for example, before-login, login, or tunnel)
pan_url_index		keyword	Counter allowing you to correlate order of log entries in URL Filtering/WildFire
pan_wildfire_hash		keyword	Binary Hash of file sent to WildFire
pan_wildfire_report_id		keyword	Identifies the analysis request on Wildfire Cloud/Appliance

Microsoft Windows Fields

Table 36: Windows Fields

Field Name	Example Values	Field Type	Notes
source_user_sid_authority1	S-1-0-0	keyword	Initial "authority" with SID preamble. For well-known non-domain SIDs this will be the only field contianing SID information.
source_user_sid_authority2		keyword	The domain authority portion of the SID
source_user_sid_rid	500	keyword	This is the user RID
target_user_sid_authority1	S-1-0-0	keyword	Initial "authority" with SID preamble. For well-known non-domain SIDs this will be the only field containing SID information.
target_user_sid_authority2		keyword	The domain authority portion of the SID
target_user_sid_rid		keyword	This is the user RID
user_sid_authority1		keyword	Initial "authority" with SID preamble. For well-known non-domain SIDs this will be the only field containing SID information.
user_sid_authority2		keyword	The domain authority portion of the SID
user_sid_rid		keyword	This is the user RID
windows_authentication_lm_package_name		keyword	This field is defined only when the windows_authentication_package_name = "NTLM"
windows_authentication_package_name		keyword	Authentication information from Event ID 4624/4625
windows_authentication_process_name		keyword	Authentication information from Event ID 4624/4625
windows_logon_type	2, 3, 10	byte	https://docs.microsoft.com/en-us/windows/security/threat-protection/auditing/event-4624
windows_logon_type_description		keyword	Description mapped to the logon type field
windows_kerberos_encryption	0x12	keyword	The Windows kerberos encryption hex value
windows_kerberos_encryption_type		keyword	Kerberos ticket encryption types https://docs.microsoft.com/en-us/windows/security/threat-protection/auditing/event-4768
windows_kerberos_service_name		keyword	Name of service targeted for Kerberos ticket requests

2.2 Graylog Message Categories

The below table shows how Graylog is mapping `gim_event_type_code` created in a pipeline, to a normalized category in our Illuminate Content. Normalized categories allow for dashboards, searches, alert rules to use a common name across all device types utilizing this format. An example of how this line looks like in the lookup tables is:

```
"100000","|authentication|","|logon|","logon"
```

The Code 100000, is attached to the log in the processing pipeline, and allows for the lookup function to attach a category, sub cateory and event type further down the processing chain.

The Category in the above case is `|authentication|`, where many types of events can fall. Logon, Logoff and Session Disconnect all fall under authentication for easy grouping on dashboards. A Subcategory of `|logon|` is applied as well to this log to signify this is happening during the logon process. There can be many under logon, like a logon sucess, logon failure. Finally the event type is added `logon` for further granularity of what this event was processed as.

2.2.1 Notes:

- This document is a work in progress and fields will be added as content is developed. If you have a suggestion, please open a GitHub ticket HERE.

Table 37: Graylog Message Categories

gim_event_type_code	gim_event_category	gim_event_subcategory
0	|message|	|message| message
100000	|authentication|	|authentication.logon| logon
100003	|authentication|	|authentication.logon| logon with alternate credentials
100004	|authentication|	|authentication.logon| session reconnect
100500	|authentication|	|authentication.credential validation| credential validation
100501	|authentication|	|authentication.credential validation| error
100502	|authentication|	|authentication.credential validation| mfa
100503	|authentication|	|authentication.credential validation| sms_send_message
100504	|authentication|	|authentication.credential validation| voice_call
101000	|authentication|	|authentication.notice| special logon
101001	|authentication|	|authentication.access notice| error
101500	|authentication|	|authentication.access policy| access policy violation
101501	|authentication|	|authentication.access policy| device policy violation
101502	|authentication|	|authentication.access policy| account policy violation
102000	|authentication|	|authentication.kerberos request| service ticket renewed
102001	|authentication|	|authentication.kerberos request| service ticket requested
102002	|authentication|	|authentication.kerberos request| tgt request
102003	|authentication|	|authentication.kerberos request| error
102500	|authentication|	|authentication.logoff| logoff
102501	|authentication|	|authentication.logoff| session disconnect
109500	|authentication|	|authentication.logon|authentication.credential validation| logon
109501	|authentication|	|authentication.kerberos request|authentication.credential validation| tgt request
109999	|authentication|	|authentication.default| authentication message
110000	|iam|	|iam.object create| account created

Continued on next page

Table 37 – continued from previous page

gim_event_type_id	gim_event_category	gim_event_subcategory	gim_event_type
110001	|iam|	|iam.object.create|	error
110002	|iam|	|iam.object.create|	group created
110500	|iam|	|iam.object.delete|	account deleted
110501	|iam|	|iam.object.delete|	group deleted
111000	|iam|	|iam.object.modify|	account modified
111001	|iam|	|iam.object.modify|	privileges assigned
111002	|iam|	|iam.object.modify|	privileges removed
111003	|iam|	|iam.object.modify|	account renamed
111004	|iam|	|iam.object.modify|	password change
111005	|iam|	|iam.object.modify|	administrative password reset
111006	|iam|	|iam.object.modify|	error
111007	|iam|	|iam.object.modify|	group member added
111008	|iam|	|iam.object.modify|	group member removed
111009	|iam|	|iam.object.modify|	group properties modified
111500	|iam|	|iam.object.disable|	account locked
111501	|iam|	|iam.object.disable|	account disabled
112000	|iam|	|iam.object.enable|	account unlocked
112001	|iam|	|iam.object.enable|	account enabled
112002	|iam|	|iam.object.enable|	error
119500	|iam|	|iam.information|	group membership enumerated
119999	|iam|	|iam.default|	iam message
120000	|network|	|network.network connection|	network connection
120100	|network|	|network.routing|	network routing
120500	|network|	|network.flow|	flow record
129999	|network|	|network.default|	network message
130000	|messaging|	|messaging.email|	email sent
130500	|messaging|	|messaging.email|	email blocked
131000	|messaging|	|messaging.email|	email rejected
131500	|messaging|	|messaging.email|	email quarantined

Continued on next page

Table 37 – continued from previous page

gim_event_type_code	gim_event_category	gim_event_category_type
132000	|messaging|	|messaging.email|	email deleted
139999	|messaging|	|messaging.default|	message
140000	|name resolution|	|name resolution.dns request|	dns query
140100	|name resolution|	|name resolution.dns transaction|	dns query and response
140200	|name resolution|	|name resolution.dns answer|	dns response
140300	|name resolution|	|name resolution.error|	dns error
140500	|name resolution|	|name resolution.ddns update|	ddns update
149999	|name resolution|	|name resolution.default|	dns message
150000	|database|	|database.query|	database query
150500	|database|	|database.update|	update rows
151000	|database|	|database.add|	insert rows
151001	|database|	|database.add|	add table
151002	|database|	|database.add|	create database
151500	|database|	|database.delete|	delete rows
151501	|database|	|database.delete|	drop table
151502	|database|	|database.delete|	drop database
159999	|database|	|database.default|	database message
160000	|endpoint|	|endpoint.process|	process started
160001	|endpoint|	|endpoint.process|	process stopped
160002	|endpoint|	|endpoint.process|	image loaded
160003	|endpoint|	|endpoint.process|	process accessed
160004	|endpoint|	|endpoint.process|	process altered
160005	|endpoint|	|endpoint.process|	remote thread created
160500	|endpoint|	|endpoint.ports|	port open
160501	|endpoint|	|endpoint.ports|	port closed
160502	|endpoint|	|endpoint.ports|	open ports
161000	|endpoint|	|endpoint.file.system|	file created
161001	|endpoint|	|endpoint.file.system|	file deleted
161002	|endpoint|	|endpoint.file.system|	file modified
161003	|endpoint|	|endpoint.file.system|	file timestamp modified
161004	|endpoint|	|endpoint.file.system|	file stream created
161010	|endpoint|	|endpoint.file.system|	raw access
161011	|endpoint|	|endpoint.file.system|	file access

Continued on next page

Table 37 – continued from previous page

gim_event_type_id	gim_event_type_code	gim_event_category	gim_event_type
161500	|endpoint|	|endpoint.service|	service started
161501	|endpoint|	|endpoint.service|	service stopped
161502	|endpoint|	|endpoint.service|	configuration change
161503	|endpoint|	|endpoint.service|	service installed
161504	|endpoint|	|endpoint.service|	service removed
161505	|endpoint|	|endpoint.service|	service error
162000	|endpoint|	|endpoint.configuration|	registry key added
162001	|endpoint|	|endpoint.configuration|	registry key removed
162002	|endpoint|	|endpoint.configuration|	configuration file modified
162003	|endpoint|	|endpoint.configuration|	system configuration modified
162004	|endpoint|	|endpoint.configuration|	system time changed
162005	|endpoint|	|endpoint.configuration|	registry value set
162006	|endpoint|	|endpoint.configuration|	registry value added
162007	|endpoint|	|endpoint.configuration|	registry value removed
162008	|endpoint|	|endpoint.configuration|	registry object renamed
162020	|endpoint|	|endpoint.configuration|	driver loaded
162500	|endpoint|	|endpoint.audit|	audit log cleared
162501	|endpoint|	|endpoint.audit|	auditing stopped
162502	|endpoint|	|endpoint.audit|	audit error
162503	|endpoint|	|endpoint.audit|	audit policy changed
162600	|endpoint|	|endpoint.pipe|	pipe created
162601	|endpoint|	|endpoint.pipe|	pipe connected
162700	|endpoint|	|endpoint.wmi|	wmi filter
162701	|endpoint|	|endpoint.wmi|	wmi consumer
162702	|endpoint|	|endpoint.wmi|	wmi binding
162800	|endpoint|	|endpoint.agent.activity|	agent activity
162900	|endpoint|	|endpoint.agent.update|	agent update
163000	|endpoint|	|endpoint.agent.status|	agent status
169800	|endpoint|	|endpoint.system.health|	system health
169900	|endpoint|	|endpoint.dlp|	clipboard changed
169999	|endpoint|	|endpoint.default|	endpoint message
170000	|alert|	|alert.network.alert|	ids alert
170001	|alert|	|alert.network.alert|	network alert
170002	|alert|	|alert.network.alert|	network dlp alert
171000	|alert|	|alert.host.alert|	malware alert
171001	|alert|	|alert.host.alert|	host dlp alert
171002	|alert|	|alert.host.alert|	hips alert
171003	|alert|	|alert.host.alert|	fim alert
179999	|alert|	|alert.default|	alert message
180000	|http|	|http.default|	http message

Continued on next page

Table 37 – continued from previous page

gim_event_type_id	gim_event_category	gim_event_type
180100	|http|	|http.request|	http request
180200	|http|	|http.communication|	http communication
180300	|http|	|http.proxied|	http proxied communication
900000	|http|network|	|http.default|network.default|	network default http default
900001	|http|network|	|http.default|network.http.default|connection|	network http default
900002	|http|network|	|http.request|network.default|	network default http request
900003	|http|network|	|http.request|network.http.request|connection|	network http request
900004	|http|network|	|http.communication|network.default|	network default http communication
900005	|http|network|	|http.communication|network.http.communication|connection|	network http communication
900006	|http|network|	|http.proxied|network.default|	network default http proxied
900007	|http|network|	|http.proxied|network.http.proxied|connection|	network http proxied