3. Custom sub-rules where VMID is NOT equal to a specific value.
4. System sub-rules where VMID is NOT equal to a specific value.

Parse Fields and Tags

The following tables provide lists of all the metadata fields LogRhythm can parse, as well as their associated parsing tags, and default regex. The fields are grouped by how they appear in the Web Console. If you do not see a field in the Web Console in the same tab as this document, you may have tagged the field as a favorite, in which case the field will appear in the Favorites tab instead of the main group tab as shown in this document. If necessary, the default regex can be overridden, as described in Override the Default Regex.

> All Mapping and Parsing tags are lower case.
>
> Fields denoted with † are available for parsing and investigations, and are viewable in the Web Console.

Field	Description	Tags	Default Regex
Application Tab			
Application	Application derived by IANA protocol and port number or directly assigned in MPE processing settings.	N/A	N/A
Object	The resource (i.e., file) referenced or impacted by activity reported in the log.	<object>	\w+
Object Name	The descriptive name of the object. Do not use unless Object is also used.	<objectname>	\w+
Object Type †	A category type for the object (e.g., file, image, pdf, etc.).	<objecttype>	\w+
Hash †	The hash value reported in the log. Choose MD5 > Sha1 > Sha256.	<hash>	\w+

Field	Description	Tags	Default Regex
Policy †	The specific policy referenced (i.e., Firewall, Proxy) in a log message.	<policy>	\w+
Result †	The outcome of a command operation or action. For example, the result of *quarantine* might be *success*.	<result>	\w+
URL	The URL referenced or impacted by activity reported in the log. You may need to override the default regex for URLs that are not HTTP/HTTPS.	<url>	https?://.+
User Agent †	The User Agent string from web server logs.	<useragent>	\w+
Response Code †	The explicit and well-defined response code for an action or command captured in a log. Response Code differs from Result in that response code should be well- structured and easily identifiable as a code.	<responsecode>	\w+
Subject	The subject of an email or the general category of the log.	<subject>	\w+
Version	The software or hardware device version described in either the process or object.	<version>	\w+

Field	Description	Tags	Default Regex
Command	The specific command executed that has been recorded in the log message.	<command>	\w+
Reason †	The justification for an action or result when not an explicit policy.	<reason>	\w+
Action †	Field for "what was done" as described in the log. Action is usually a secondary function of a command or process.	<action>	\w+
Status †	The vendor's perspective on the state of a system, process, or entity. Status should NOT be used as the result of an action.	<status>	\w+
Session Type †	The type of session described in the log (e.g., console, CLI, web). Unique from IANA Protocol.	<sessiontype>	\w+
Process Name	System or application process described by the log message.	<process>	\w+
Process ID	Numeric ID value for a process.	<processid>	\d+
Parent Process ID †	The parent process ID of a system or application process that is of interest.	<parentprocessid>	\w+

Field	Description	Tags	Default Regex
Parent Process Name †	The parent process name of a system or application process.	<parentprocessname>	\w+
Parent Process Path †	The full path of a parent process of a system or application process.	<parentprocesspath>	\w+
Quantity	A numeric count of something. For example, there are 4 lights (quantity is 4).	<quantity>	[0123456789\.]+
Amount	The qualitative description of quantity (percentage or relative numbers) For example, half the lights are on (amount is .5 or 50). Amount is also used for currency.	<amount>	[0123456789\.]+
Size	Numeric description of capacity (e.g., disk size) without a specific unit of measurement. Size is generally used as a limit rather than a current measurement. Use Amount for non- specific measurements.	<size>	[0123456789\.]+
Rate	Defines a number of something per unit of time without a specific unit of measurement. Always expressed as a fraction.	<rate>	[0123456789\.]+

Field	Description	Tags	Default Regex
Duration	The elapsed time reported in a log message, derived from multiple fields. Timestart and Timeend need custom parsing patterns.	If log has start/end use: (?<timestart>*pattern)* (?<timeend>*pattern)* If log has elapsed time use: <days> <hours> <minutes> <seconds> <milliseconds> <microseconds> <nanoseconds>	[0123456789\.]+ *Note: Time Start and Time End tags must be overloaded to function properly.*
Session	Unique user or system session identifier.	<session>	\w+
Known Application	Application derived from IANA protocol and port number. If a known application cannot be derived, it is displayed as unknown.	N/A	N/A
Kbytes/Packets Tab			

Field	Description	Tags	Default Regex
• **Host (Impacted) KBytes Rcvd** • **Host (Impacted) KBytes Sent** • **Host (Impacted) Kbytes Total**	The number of bytes sent or received in the context of the Impacted Host. • Rcvd – Bytes received by impacted host • Sent – Bytes sent by impacted host • Total – Total bytes in session as seen by impacted host	Use the appropriate tags based upon the units and direction represented by the log data: <bitsin>, <bitsout> <bytesin>, <bytesout> <kilobitsin>, <kilobitsout><kilobytesin>, <kilobytesout><megabitsin>, <megabitsout> <megabytein>, <megabyteout><gigabitsin>, <gigabitsout><gigabytein>, <gigabyteout> <terabitsin>, <terabitsout><terabytesin>, <terabytesout><petabitsin>, <petabitsout> <petabytesin>, <petabytesout>,<bits>, <bytes>, <kilobits>,<kilobytes>, <megabits>, <megabytes>, <gigabits>,<gigabytes>, <terabits>,<terabytes>, <petabits>,<petabytes>	[0123456789\.]+
• **Host (Impacted) Packets Rcvd** • **Host (Impacted) Packets Sent** • **Host (Impacted) Packets Total**	The number of packets sent or received in the context of the Impacted Host. • Rcvd – Packets received by impacted host • Sent – Packets sent by impacted host • Total – Total packets in session as seen by impacted host	<packetsin>, <packetsout>, <packets>	[0123456789\.]+
Classification Tab			

Field	Description	Tags	Default Regex
Classification	Value is determined based on the MPE Rule's assigned Common Event.	N/A	N/A
Common Event	Value is determined based on the MPE Rule's Assigned Common Event.	N/A	N/A
Priority	Value is determined based on the Risk-Based-Priority (RBP) calculation.	N/A	N/A
Direction	Indicates the directional flow of data between the Origin Host and the Impacted Host — Inbound, Outbound, Internal, External, or Unknown.	N/A	N/A
Severity	The vendor's view of the severity of the log.	<severity>	\w+
Vendor Message ID	Specific vendor for the log used to describe a type of event.	<vmid>	\w+
Vendor Info †	Description of a specific vendor log or event identifier for the log. Human readable elaboration that directly correlates to the VMID.	<vendorinfo>	\w+
MPE Rule Name	Name of rule that matched, assigned on rule creation.	N/A	N/A
Threat Name †	The name of a threat described in the log message (e.g., malware, exploit name, signature name). Do not overload with Policy.	<threatname>	\w+

Field	Description	Tags	Default Regex
Threat ID †	ID number or unique identifier of a threat. Note that CVE is stored separately.	<threatid>	\w+
CVE †	CVE ID (i.e., CVE-1999-0003) from vulnerability scan data.	<cve>	\w+
Host Tab			
Host (Origin)	Origin host derived from Origin IP Address and/or Origin Hostname.	N/A	N/A
Host (Impacted)	Impacted host derived from Impacted IP Address and/or Impacted Hostname.	N/A	N/A
MAC Address (Origin)	The MAC address from which activity originated (i.e., attacker, client).	<smac>	(\w{2}(:|-)?){6}
MAC Address (Impacted)	The MAC address that was affected by the activity (i.e., target, server).	<dmac>	(\w{2}(:|-)?){6}
Interface (Origin)	The network port/interface from which the activity originated (i.e., attacker, client).	<sinterface>	\w+
Interface (Impacted)	The network port/interface that was affected by the activity (i.e., target, server).	<dinterface>	\w+

Field	Description	Tags	Default Regex
IP Address (Origin)	The IP address from which activity originated (i.e., attacker, client).	<sip> (parses IPv4 and IPv6)	((?<sipv4>(?<sipv4>1??(1??\d{1,2}|2[0-4]\d|25[0-5])\.(1??\d{1,2}|2[0-4]\d|25[0-5])\.(1??\d{1,2}|2[0-4]\d|25[0-5])\.(1??\d{1,2}|2[0-4]\d|25[0-5])))|(?<sipv6>(?<sipv6>1??((?:(?:[0-9A-Fa-f]{1,4}:){7}[0-9A-Fa-f]{1,4}|(?=(?:[0-9A-Fa-f]{1,4}:){0,7}[0-9A-Fa-f]{1,4}\z)|(([0-9A-Fa-f]{1,4}:){1,7}|:)((:[0-9A-Fa-f]{1,4}){1,7}|:))))))
IP Address (Impacted)	The IP address that was affected by the activity (i.e., target, server).	<dip> (parses IPv4 and IPv6)	((?<dipv4>(?<dipv4>1??(1??\d{1,2}|2[0-4]\d|25[0-5])\.(1??\d{1,2}|2[0-4]\d|25[0-5])\.(1??\d{1,2}|2[0-4]\d|25[0-5])\.(1??\d{1,2}|2[0-4]\d|25[0- 5])))|(?<dipv6>(?<dipv6>1??((?:(?:[0-9A-Fa-f]{1,4}:){7}[0-9A-Fa-f]{1,4}|(?=(?:[0-9A-Fa-f]{1,4}:){0,7}[0-9A-Fa- f]{1,4}\z)|(([0-9A-Fa- f]{1,4}:){1,7}|:)((:[0-9A-Fa-f]{1,4}){1,7}|:))))))
NAT IP Address (Origin)	The Network Address Translated (NAT) IP address from which activity originated (i.e., attacker, client).	<snatip>	Same as IP Origin (<sip>)
NAT IP Address (Impacted)	The Network Address Translated (NAT) IP address that was affected by the activity (i.e., target, server).	<dnatip>	Same as IP Impacted (<dip>)

Field	Description	Tags	Default Regex
Hostname (Origin)	The hostname from which activity originated (i.e., attacker, client).	<sname> (or DNS resolved from IP)	([^\s\.]+\.?)+
Hostname (Impacted)	The hostname that was affected by the activity (i.e., target, server).	<dname> (or DNS resolved from IP)	([^\s\.]+\.?)+
Known Host (Origin)	A value determined by mapping parsed origin host identifiers, such as IP address or hostname, to a LogRhythm host record.	N/A	N/A
Known Host (Impacted)	A value determined by mapping parsed impacted host identifiers, such as IP address or hostname, to a LogRhythm host record.	N/A	N/A
Serial Number †	The hardware or software serial number in a log message. This value should be a permanent unique identifier.	<serialnumber>	\w+
Identity Tab			
User (Origin)	The originating user or system account of the activity reported in the log.	<login>	\w+
User (Impacted)	The user or system account impacted by activity reported in the log.	<account>	\w+
Sender	The sender of an email or the "caller number" for a VOIP log. This value must relate to a specific user or unique address in the case of a phone call or email.	<sender>	[^\s]+@[^\s]+
Recipient	The recipient of an email or the dialed number for a VOIP log.	<recipient>	[^\s]+@[^\s]+

Field	Description	Tags	Default Regex
Group	The user group or role impacted by activity reported in the log. Do not use for entity group (zone or domain).	<group>	\w+
Location Tab			
Entity (Origin)	A value determined based on the origin host's assigned entity.	N/A	N/A
Entity (Impacted)	A value determined based on the impacted host's assigned entity.	N/A	N/A
Zone (Origin)	A value determined based on the zone of the origin host — Internal, External, DMZ, or Unknown.	N/A	N/A
Zone (Impacted)	A value determined based on the zone of the impacted host — Internal, External, DMZ, or Unknown.	N/A	N/A
Location (Origin)	A value determined by resolving the parsed origin IP address against a Geo-IP database.	N/A	N/A
Location (Impacted)	A value determined by resolving the parsed impacted IP address against a Geo-IP database.	N/A	N/A
Country (Origin)	The country in which the determined origin location exists.	N/A	N/A
Country (Impacted)	The country in which the determined impacted location exists.	N/A	N/A
Log Tab			

Field	Description	Tags	Default Regex
Log Date	Timestamp when the log was generated or received, corrected to UTC.	N/A	N/A
Log Count	The number of identical log messages received.	N/A	N/A
Log Source Entity	The entity to which the log source belongs.	N/A	N/A
Log Source Type	The device or application type from which a log was received.	N/A	N/A
Log Source Host	The origin host from which the log was received.	N/A	N/A
Log Source	The assigned name of a log source.	N/A	N/A
Log Sequence Number	The sequence in which a log was collected, generated by the Agent.	N/A	N/A
Log Message	The raw log message.	N/A	N/A
First Log Date	Timestamp when the first identical log message was received.	N/A	N/A
Last Log Date	Timestamp when the last identical log message was received.	N/A	N/A
Network Tab			
Network (Origin)	A value determined by mapping the origin IP address to a LogRhythm network record.	N/A	N/A
Network (Impacted)	A value determined by mapping the impacted IP address to a LogRhythm network record.	N/A	N/A

Field	Description	Tags	Default Regex
Domain (Impacted) †	The Windows or DNS domain name referenced or impacted by activity reported in the log.	<domain> or <domainimpacted>	\w+
Domain (Origin) †	The Windows or DNS domain where the logged activity originated.	<domainorigin>	\w+
Protocol	The IANA protocol name or number.	<protnum>,<protname>	1??\d{1,2}|2[0-4]\d|25[0-5] \w+
TCP/UDP Port (Origin)	The port from which activity originated (i.e., client, attacker port).	<sport>	\d+
TCP/UDP Port (Impacted)	The port to which activity was targeted (i.e., server, target port).	<dport>	\d+
NAT TCP/UDP Port (Origin)	The Network Address Translated (NAT) port from which activity originated (i.e., client, attacker port).	<snatport>	\d+
NAT TCP/UDP Port (Impacted)	The Network Address Translated (NAT) port to which activity was targeted (i.e., server, target port).	<dnatport>	\d+

Map Tags

Five additional tags are available for identifying data in the log specifically for sub-rules. These tags do not parse text into metadata fields, so they do not appear in Investigations, Reports, and so on. These tags are intended only to identify portions of the log message that should be used in the development of sub-rules.

Tag	Field Type	Default Regex
<tag1>	Text	.*
<tag2>	Text	.*
<tag3>	Text	.*

Tag	Field Type	Default Regex
<tag4>	Text	.*
<tag5>	Text	.*

Override the Default Regex

The default regex is applied by using only the named group tag. For example, <account> will apply the regex pattern \w+ in the rule, as shown in the table above.

If the default regex for a parsing tag will not properly parse the correct data out of the log message or is not the optimal regex from a performance perspective, the default should be overridden. To override the default regex, the following syntax should be used:

```
(?<[tagname]>[regex])
```

For example, suppose your regex needs to match file names with a specific extension such as the sample log message below:

```
User john.doe opened AnnualReport.pdf
```

If the base rule was written as:

```
User <login> opened <object>
```

The value parsed for login would be john and the value for object would be AnnualReport. This is due to the fact that a period is not a word character and the default regex of “\w+” would only match up to the period. Instead, the default regular expressions should be overridden, and the base rule should be:

```
User (?<login>\w+\.?\w*) opened (?<object>\w+\.pdf)
```

Now, the base rule will parse anything for login starting with a word character that optionally contains a period followed by additional word characters.

> Do not override the default regex for fields which parse an IP address, such as <sip>, <dip>, <sipv6>, and so on.