What is Logagent

Logagent is a modern, open-source, light-weight log shipper. It is like Filebeat and Logstash in one, without the JVM memory footprint. It comes with out of the box and extensible log parsing, on-disk buffering, secure transport, and bulk indexing to Elasticsearch, Logsene, and other destinations. Its low memory footprint and low CPU overhead makes it suitable for deploying on edge nodes and devices, while its ability to parse and structure logs makes it a great Logstash alternative.

Features

This project contains a library and patterns for log parsing and cli tools and installers to use logagent-js as log shipper with the following features:

Pluggable inputs

- Standard input (stdin) that can read the output stream from any Linux cli tool
- TCP input that can read from a TCP socket
- File input, watching a list of files defined by glob patterns (e.g. /var/log/**/*.log). The tail file functions can deal with logrotate (renaming/moving files) and stores the last position read for each file descriptor to recover after a service restart.
- Syslog Server (UDP) listener - Logagent can also act as a syslog server and receive Syslog messages via UDP. The parser is applied to the message field.
- Heroku Log Drain makes it easy to receive Heroku logs for processing
- Cloud Foundry Log Drain
- Command - Run commands frequently and process the command output
- Elasticsearch - run Elasticsearch queries frequently and process teh search results
- Database - Run MySQL, PostgresSQL, MS-SQL commands frequently and process query results
- Custom plugins via Logagent plugin API

Pluggable processors

- Grep input filter - filters raw messages by regular expression
- SQL output filter - transforms or aggregates parsed log messages
- Access.watch - Enrich web server logs with robot detection and traffic intelligence. Access Watch is the reference intelligence for the thousands

of good and bad robots active on the web. With Reveal, easily integrate this knowledge in your logs and dashboards.
- Custom input and output processors via Logagent plugin API, defined in separate npm modules or inline JavaScript in the configuration file

Pluggable outputs

- Standard output in JSON, line delimited JSON or YAML format.
- Elasticsearch output - Efficient bulk indexing and reliable transmission.
- Logagent doesn't lose data. It stores parsed logs to a disk buffer if the network connection to the Elasticsearch API fails. Logagent retries shipping logs later, when the network or Elasticsearch is available again.
- Log routing functions are available to map log inputs to different Elasticsearch servers or different indices. Full SSL/TLS support, authentication via SSL client certificates or basic authentication.
- rtail output: UDP forwarding to rtail server for realtime log view

Build-in data parser

- Log format detection and intelligent pattern matching
- Pattern library included, covering a set of standard application like databases, web servers and message queue services
- Easy to extend with custom patterns and JS transform functions
- Hot reload of changed pattern definitions without service restart
- Recognition of Date and Number fields
- Replace sensitive data with configurable hashing algorithms (SHA-1, SHA-256, SHA-512, …)
- GeoIP lookup with automatic GeoIP db updates (maxmind geopip-lite files)

Command-line tool

- Works with Unix pipes (stdin/stdout)

- Log parser and format converter (e.g. text to JSON, line delimited JSON or YAML) `cat access.log | logagent --yaml`
- Import files into Elasticsearch `cat access.log | logagent -n nginx -e http://localhost:9200 -i logs`
- Watch multiple log files in the terminal `logagent -yaml -g '/var/log/**/*.log'`
- Store logs in Elasticsearch and watch them in real-time in the Web Browser `logagent -e http://localhost:9200 -i logs --rtailWebPort 8080 --rtailPort 9999 /var/log/*.log`

Plugins

The architecture of logagent is modular and each input or output module is implemented as plugin for the logagent framework. Plugins are loaded on demand depending on the configurations.

Plugin	Type	Description
stdin	input	Reads from standard input
files	input	Watching and tailing files
syslog	input	Receive syslog messages (UDP)
input-tcp	input	Receive data via TCP
heroku	input	Receive logs from Heroku log drains (HTTP)
cloudfoundry	input	Receive logs from Cloud Foundry log drains (HTTP)
command	input	Receive logs from the output of a command, which could run once or periodically
mysql-query	input	Receive results from SQL queries, which could run once or periodically
mssql-query	input	Receive results from SQL queries, which could run once or periodically
postgresql-query	input	Receive results from SQL queries, which could run once or periodically
elasticsearch-query	input	Receive results from Elasticsearch queries, which could run once or periodically
input-kafka	input	Receiver messages from Apache Kafka topics
grep	Processor / input filter	Filters text with regular expressions before parsing
sql	Processor / output filter	Transforms and aggregates parsed messages with SQL statements
access-watch	Processor / output filter	Enriches web server logs with robot detection and traffic intelligence
stdout	output	Prints parsed messages to standard output. Supported formats: YAML, JSON, Line delimited JSON (default).
elasticsearch	output	Stores parsed messages in Elasticsearch

Plugin	Type	Description
rtail	output	Sends parsed messages to rtail servers for real-time view of logs in a web browser
output-kafka	output	Sends parsed messages to Apache Kafka topics

Reliable log shipping with disk buffer

Logagent doesn't lose data. It stores parsed logs to a disk buffer if the network connection to the Elasticsearch API fails. Logagent retries shipping logs later, when the network or Elasticsearch is available again.

Deployment options

- Deployable as a system service: systemd, upstart (Linux), or launchd (Mac OS X)
- Docker Container
- Deployement to Heroku as Heroku Log drain
- Deployement to Cloud Foundry as Cloud Foundry Log drain (thus usable with Pivotal, Bluemix, etc.)

API

- Node.js module to integrate parsers into Node.js programs
- logagent-js is a part of Sematext Docker Agent to parse container logs

Related packages

- Sematext Agent for Docker - collects metrics, events and logs from Docker API and CoreOS. Logagent-js is a component of sematext-agent-docker. More Information: Innovative Docker Log Management
- Logsene-CLI - Enables searching Logsene log entries from the command-line.
- SPM Agent for Node.js - collects performance metrics for Node and io.js applications
- Custom Metrics - Custom Metrics for SPM
- Winston-Logsene - Logging for Node.js - Winston transport layer for Logsene

Support

- Twitter: [@sematext](http://twitter.com/sematext)
- Blog: sematext.com/blog
- Homepage: sematext.com