Using Fluentd as an alternative to Splunk

As infrastructure within organizations grows in size and the number of hosts, the cost of Splunk may become prohibitive. I created this document to demonstrate, step-by-step, how to implement Fluentd, Elasticsearch and Kibana to be used as a replacement for Splunk.

This is a simple configuration and does not include clustering. I may add clustering in later versions. The OS used for this configuration is CentOS 6.7.

The instructions for newer versions of CentOS/RHEL, and for Ubuntu may vary.

All commands should be run as root or be preceded by “sudo “.

Table of Contents

Pre-requisites

Create a Logical Volume and Mount Point

Because this system will be used for log aggregation, the usage of disk space will grow over time. To accommodate this future growth and prevent planned downtime for expanding disks, it makes a lot of sense to create a logical volume that can be extended on the fly. For this purpose I added a second virtual disk to the VM (/dev/sdb).

1. Run the command “fdisk /dev/sdb”
    a. “n” to create a new partition.
    b. “p” to make it a primary partition.
    c. Partition number “1”.
    d. Accept the defaults and allow the partition to be created.
    e. When the processes if finished, use “t” to change the partition type.
    f. Choose “8e” to change to LVM partition type.
    g. “w” to write the changes to disk.
2. Run the command “pvcreate /dev/sdb1” to create an LVM physical volume on the new partition.
3. Create a new volume group on the physical volume using “vgcreate vglogs /dev/sdb1”.
4. Create a new logical volume using 100% of the space on the new volume group using “lvcreate -l 100%FREE -n lv_logs vglogs”.
5. Create a filesystem on the new logical volume using “mkfs -t ext4 /dev/vglogs/lv_logs”.
6. Create an empty directory for the new filesystems “mkdir /logs”.
7. Add the following to /etc/fstab so that the new filesystem is mounted on boot:
   /dev/vglogs/lv_logs /logs ext4 defaults 0 0
8. Mount the new filesystem using “mount /logs”.
9. OPTIONAL: Disable regular fsck checks using “tune2fs -c 0 -i 0 /dev/mapper/vglogs-lv_logs”
   *** This is not recommended, but can increase performance and decrease reboot times on occasion.

Create Directory Structure for Log Storage

Create the following directories:

1. /logs/data
2. /logs/work
3. /logs/logs
4. /logs/syslog

Install or Update to the Latest Version of Java

Java is a prerequisite for installing Elasticsearch.

In this case, you don’t need to install Sun Java; you can install the Open Source version of the JRE.

Simply run “yum –y install java”. This will install the latest version of Open Source Java, or update your existing environment to the latest.

NTP

If you are using NTP in your environment, you should configure /etc/ntp.conf to sync with your NTP servers. This is strongly recommended and will prevent differences in the timestamps within collected logs.

Tweak Network Parameters

1. Add the following to /etc/sysctl.conf if your environment will have heavy loads:
   net.ipv4.tcp_tw_recycle = 1
   net.ipv4.tcp_tw_reuse = 1
   net.ipv4.ip_local_port_range = 10240 65535
2. Run "sysctl –w" to reload the configuration.

Increase Limit on File Descriptors

1. Run "ulimit –n" to view the currently configured limit. If the result is 1024, you will need to change it in the next steps.
2. Edit the file "/etc/security/limits.conf" and add the following lines (before "# End of file"):
   root soft nofile 65536
   root hard nofile 65536
   * soft nofile 65536
   * hard nofile 65536
3. Reboot the server for these to take effect

Install and Configure Elasticsearch as a Service

The easiest way to install Elasticsearch and run it as a deamon is to download and extract it directly in the /etc directory.

Download and Install

1. Run “cd /etc”.
2. Run “curl -O https://download.elastic.co/elasticsearch/elasticsearch/elasticsearch-1.7.2.tar.gz” to download. (check the Elasticsearch website for newer versions).
3. Extract the application using “tar zxvf elasticsearch-1.7.2.tar.gz”.
4. To keep the directory clean, remove the gzip file using “rm -f elasticsearch-1.7.2.tar.gz”.
5. Rename the extracted directory using “mv elasticsearch-1.7.2 elasticsearch”.

Install Plugins

1. Change into the bin directory, “cd elasticsearch/bin”.
2. Run the following commands:
    a. ./plugin -install karmi/elasticsearch-paramedic
    b. ./plugin -install mobz/elasticsearch-head

Configure Elasticsearch

1. Change into the config directory “cd ../config”.
2. Open the configuration file elasticsearch.yml
3. Uncomment and change the following lines to be:
    a. path.data: /logs/data
    b. path.work: /logs/work
    c. path.logs: /logs/logs
    d. path.plugins: /etc/elasticsearch/plugins

Configure Elasticsearch to Run as a Service

1. cd back to the root directory.
2. Run “curl -L http://github.com/elasticsearch/elasticsearch-servicewrapper/tarball/master | tar –xz”.
3. Run “mv elastic-elasticsearch-servicewrapper-8513436/service /etc/elasticsearch/bin/”. ** Note that “elastic-elasticsearch-servicewrapper-8513436” may have a different name as newer versions are released.
4. Run “rm –rf elastic-elasticsearch-servicewrapper-8513436” to keep the filesystem clean.
5. Run “/etc/elasticsearch/bin/service/elasticsearch install”.
6. Run “chkconfig elasticsearch on” to set the service to start at boot.
7. Run “service elasticsearch start” to start the service.
8. Run “ps -ef | grep elastic” to verify that the service is running. You should see lots of output.
9. OPTIONAL: Reboot the server “init 6” then run the command “ps -ef | grep elastic” after the server reboots to verify that the service started on boot.

Install and Configure Kibana to Run as a Service

Download and Install Kibana

1. Run “cd /etc”
2. Run “curl -O https://download.elastic.co/kibana/kibana/kibana-4.1.2-linux-x64.tar.gz”
3. Run “tar zxvf kibana-4.1.2-linux-x64.tar.gz”
4. Run “mv kibana-4.1.2-linux-x64 kibana”

Create and Secure the User for the Kibana Service

1. Run “useradd kibana”.
2. Edit the file /etc/passwd and change the home directory for the user “kibana” to “/sbin/nologin”.

Configure Kibana to Run as a Service

1. Run “cd /etc/init.d”.
2. Run “curl –LO https://github.com/cjcotton/init-kibana/raw/master/kibana”.
3. Run “chmod 755 kibana”
4. Edit the file “kibana” and change these lines to the following values:
    a. KIBANA_BIN=/etc/kibana/bin
    b. DAEMON_USER=kibana
5. Run “chkconfig kibana on” to ser the service to start on boot.
6. Run “service kibana start” to start the service.
7. Run “ps -ef | grep kibana” to verify that the service is running.
8. OPTIONAL: Reboot the server “init 6” then run the command “ps -ef | grep kibana” after the server reboots to verify that the service started on boot.

Install and Configure Fluentd to Run as a Service

This install will be for the application td-agent, which is the stable release of Fluentd.

Install td-agent

Installation is very easy: download and run a script that does all the work!

1. Run “curl -L https://td-toolbelt.herokuapp.com/sh/install-redhat-td-agent2.sh | sh”.
2. Run “chkconfig td-agent on” to configure td-agent to start at boot.

Install and Configure the Elasticsearch plugin for td-agent

1. Run the command “/usr/sbin/td-agent-gem install fluent-plugin-elasticsearch”
2. Modify the file /etc/td-agent/td-agent.conf to have the contents below:

```
<source>
  type syslog
  port 42185
  tag syslog
</source>

<source>
  type forward
</source>

<match syslog.**>
  type elasticsearch
  logstash_format true
  flush_interval 10s # for testing
</match>
```

Start the td-agent Service

1. Run “service td-agent start” to start the service.
2. Run “ps -ef | grep td-agent” to verify that the service is running.

Configure rsyslog to Receive Logs and Forward to td-agent

1. To load the TCP and UDP listener modules, edit the file /etc/rsyslog.conf to contain the following:

```
#### MODULES ####
$ModLoad imuxsock # provides support for local system logging (e.g. via logger command)
$ModLoad imklog   # provides kernel logging support (previously done by rklogd)
#$ModLoad immark  # provides --MARK-- message capability

# Provides UDP syslog reception
$ModLoad imudp
$UDPServerRun 514
# Provides TCP syslog reception
$ModLoad imtcp
$InputTCPServerRun 514

#### GLOBAL DIRECTIVES ####
# Use default timestamp format
$ActionFileDefaultTemplate RSYSLOG_TraditionalFileFormat
# Include all config files in /etc/rsyslog.d/
$IncludeConfig /etc/rsyslog.d/*.conf

#### RULES ####
# Log anything (except mail) of level info or higher.
# Don't log private authentication messages!
*.info;mail.none;authpriv.none;cron.none        /logs/syslog/messages
# The authpriv file has restricted access.
authpriv.*                                      /logs/syslog/secure
# Log all the mail messages in one place.
mail.*                                          -/var/log/maillog
# Log cron stuff
cron.*                                          /logs/syslog/cron
# Everybody gets emergency messages
*.emerg                                         *
# Save news errors of level crit and higher in a special file.
uucp,news.crit                                  /logs/syslog/spooler
# Save boot messages also to boot.log
local7.*                                        /logs/syslog/boot.log

# ### begin forwarding rule ###
*.* @127.0.0.1:42185
```

2. To forward to td-agent, add the following to end of the file /etc/rsyslog.conf:

```
*.* @127.0.0.1:42185
```

3. Restart the rsyslog service using “service rsyslog restart”.

Test the system

1. Send a test log message using the command “logger –t test this thing”.
2. Log into the Kibana interface on the server using the URL http://hostname.domain:5601 where hostname and domain are the FQDN for your server.
3. At this time, you will see a green “Create” button in the Kibana interface. Push the button.
4. Click on “Discover” at the top of the page. You should see the test log message you sent to the system and you may also see some other syslog messages from the local host.

Configure Remote Clients to Send Logs

This section includes how to configure Linux and Windows clients to forward logs and events to the new logging server.

For other use-cases such as MongoDB, please see the quickstart guide on the Fluentd website: http://docs.fluentd.org/articles/quickstart#step2-use-cases

Configure RHEL / CentOS Linux Servers as Forwarding Clients

1. Edit the file “/ect/rsyslog.config” to uncomment and modify the line (near the bottom of the file) to use the FQDN of the logging server you have set up:
   . @@hostname.domainname:514
2. Restart the rsyslog service using “service rsyslog restart”.
3. On the client, run “logger -t test this from <hostname>.
4. Verify in the Kibana interface that the message arrived.

Configure Ubuntu Servers as Forwarding Clients

1. Edit the file “/ect/rsyslog.config” to uncomment and modify the line (near the bottom of the file) to use the FQDN of the logging server you have set up:
   . @hostname.domainname:514
2. Restart the rsyslog service using “service rsyslog restart”.
3. On the client, resatart a service to generate a log event in /var/log/syslog.
4. Verify in the Kibana interface that the message arrived.

Configure Windows Servers as Forwarding Clients

To integrate Windows logging into the new system, the free application Event Log Forwarder for Windows will be used.

*** NOTE: It might be wise to clear the event logs (if you can) prior to installing the application. When the service starts, it reads through all of the event logs on the server, and will spike the CPU to 100% for longer if the logs are large.

1. Download the free Event Log Forwarder for Windows from Solarwinds: http://www.solarwinds.com/products/freetools/log-forwarder.aspx
   ** You will have to Provide your name, email and phone number
2. Run the file "SolarWinds_Event_LogForwarder_Setup.exe" as Administrator.
3. Go through the wizard until the installation is complete.
4. Open the program "SolarWinds Event Log Forwarder for Windows" as Administrator – this is the dashboard.
5. Click on the "Syslog Servers" tab and Click the "Add" button.
6. Enter the server name and IP address for the new logging server and click the "Create" button.
7. Click on the "Subscriptions" tab and click the "Add" button.
8. In this part of the application, you can configure what events from what logs are forwarded. The interface is fairly intuitive.
9. Once you have configured one or more subscriptions, click on the "Test" tab.
10. In the "Test" tab, create a test event that matches the subscription criteria.
11. In the Kibana interface, verify that the test event has been received.