Q&A from Webinar

Q1. Metric beats를 사용할 때 CPU나 메모리 등 얼마나 리소스를 소모하는지 궁금합니다.
A1. Metric beats를 포함한 beats들이 자원을 얼마나 사용하는지 공식적으로 공유한 자료는 없어 보입니다. Metric beat 설치하는 장비의 종류와 사양, Metircbeat 설정에 따라서 천차만별이이기 때문입니다. (예를 들어 수집하려는 데이터의 개수와 전송 interval 등이 큰 영향을 미칩니다.) 내부 기술지원 엔지니어들이 자체 테스트할때는 그리 신경쓸만한 수준은 아니라는 입장이 지배적입니다만, 정확한 것은 직접 대상 서버와 비슷한 상황에서 원하는 설정과 함께 테스트를 해보시는 것이 좋겠습니다. 설정 관련 정보는 본 링크에서 참고해주시면 됩니다.
https://www.elastic.co/guide/en/beats/metricbeat/current/metricbeat-configuration.html

Q2. 쿠버네티스의 Pod 단위의 모니터링에 더해서 Pod 내 컨테이너들에 대한 모니터링도 가능한가요? 가능한 경우 쿠버네티스 컨테이너 업데이트 시에 recreate 되는 이슈가 있는데 이부분은 어떻게 처리될까요?
A2. Pod 하나에 Container 하나가 관리측면에서 이상적인 deployment 모습이기는 하지만 경우에 따라 싱글pod에 멀티컨테이너를 러닝하는 deployment 모습도 있습니다.
Elasticsearch기준으로 보면 현재 docker 엔진의 기본디렉토리인 /var/lib/docker/container/*/*.log를 수집대상으로 하기때문에 컨테이너 모니터링이 가능하다고 말씀드릴수 있고 recreate 된다고 하더라도 리소스 모니터링에는 이슈가 없어보입니다.

Q3. MetricBeat가 데이터를 수집하는 최소 주기가 몇 인가요? 1초 미만도 가능한가요?
A3. 수집 주기는 Defualt가 있지만 원하시는 대로 설정하실 수 있습니다. 1초까지는 가능합니다. 1초 이하도 세팅은 가능하지만 오버헤드가 많다는 경고가 있습니다.

Pre-planned Questions:

- Where should I install Beats when I'm using Docker/K8S?
-

IRC Questions:

Additional Notes:

Questions from English webinar chatroom:

- i have filebeat streaming streaming docker enhanced logs as just discussed. however, now i have an issue of huge log files and no rotation. is that a solved problem? otherwise, i need to instead stream logs to systemd, which handles this for me, and stream from there
    - unlike his example, im single host docker-composing* ATM, with migration to swarm soon
- how many beat processes will run in the pod, when I start 100 nginx pods?
- how difficult is it to implement own modules for autodiscovery?
- For time series data does elastic solve this problem completely? for example, with influx the concept of continuous queries and different retention policies allows for the capture of fine-grained to more coarse grained metrics
- Is there a JVM / JMX beat e.g. for Cassandra monitoring? How does metric beat find JMX port of cassandra?
- can you please organize a presentation on containers that log to files with dynamic names (ie appfilelog-a3c87f.log or appname-20180215.log)
- I'm running JBOSS/wildfly, can I use Beats* to monitor ?
- what is pack licensing thing you mentioned in the beginning? which part is free to use if I like it
- does filebeat support multiline from docker log
- how to monitor the application specific logs in kubernetes which are not written in the files but are getting written in standard console output?
    - by default stdout/stderr are what are monitored by default. the docker prospector collects logs from only stdout/stderr

- do you plan to support consul or/and nomad?
- I am using the 6.1 version of the Filebeat daemonset which used the "file" prospector and picked up messages from /var/lib/docker/containers/*/*.log. // what would be the advantage to use the docker prospector ?
- Is it possible to manage docker container logs on jelastic too?
- q: Do you have any plans to add the "metadata processors" into the Logstash pipeline if e.g. a customer want to continue to use docker's fluentd log export -> logstash -> elasticsearch?
- what would be the advantage to use the docker prospector ?
- How did he build a histogram of response codes?
- how would you parse the "log line" with log stash in this workflow?
- jboss/wildfly monitoring plz (is this a question or feature request?)
- Support for Docker swarm is available?

- using gelf to logstash is the right way or should we use gelf to filebeat is correct?
-
-
- --