Monitoring and metering of Enterprise Content Management systems using Telegraf, Influxdb and Grafana

Motivation and Objective

To monitor the performance of the ECM system and uncover the problems whenever there is a drop in performance.

To design a **service provider dashboard** to view and analyze the below.

- Overall system load across all and by each tenant: CPU, I/O, Network, Memory

- Access patterns: Log-on events, Workload: across and by each tenant

- Number of CRUD-S operations by time, Service Level Agreement (SLA) and related infringement

Motivation and Objective

To design a **tenant dashboard** with the below objectives,

- Current logged on users and failed log-on cases in the tenant

- Access patterns of the tenant and of its users

- Workload, Number of CRUD-S (operations) by time, by type, by user

Setup

- Open source stack Telegraf, Influxdb, Chronograf and Kapacitor (TICK) along with Grafana is deployed in openstack infrastructure as docker containers

- Telegraf plugins are installed to populate the system metrics into the database

- Data from ECM systems is populated using python scripts

- Influxdb is configured as a data source in Grafana and Chronograf and thereby acts as a source for all the dashboards

Service provider dashboard

- Monitors the health of the ECM system in terms of CPU usage percentage, disk space used and gives an overview of the tenants' activities in terms of CRUDS operations

- Main objective is to monitor if the system resources are well under the threshold and the services comply to the agreements between the client and the provider

- Following are the sections under which panels are grouped:

    - ECM - System overview

    - ECM - System performance

    - ECM - Access pattern by tenants

ECM - System overview

- All the instant values related to - are grouped in this section

    - system usage

    - SLA

    - current users across tenants and of the tenant filtered

- In order to show the numeric values, single stat panel is used

- A quick glance of this section equips the service provider with information about the current performance of the system and SLA infringement in the past 30 days

ECM - System performance

- System performance is tracked by comparing the system metrics with the ECM metrics over time with four different panels

  - Average CPU usage % and CRUDS operations - All Tenants

  - IO wait %, Disk Read/Write rate - All tenants

  - Number of operations - grouped by type - All tenants

  - Average response time - Create and Retrieve operations

ECM - Access pattern by tenants

- To monitor the tenant activities in terms of the log on events and workload across tenants and by each tenant

- Provides a history of tenant activities over a selected time range

- Helps to correlate system resource usage with tenant activities

- A time-based Graph panel is used to visualise the data

Tenant Dashboard

- Monitors the activities of a selected tenant in the ECM system in terms of.

  - Logged on users

  - Workload based on CRUDS operations

- Similar to the Service Provider dashboard in terms of System Access Patterns but also drilled down to an user level

- Provides an additional panel to view the current logged on users in the tenant

Features explored in InfluxDb

- Continuous queries

  - Since Influxdb does not support joins over multiple measurements, continuous queries are used to create new measurements and populate it at the same time with the information from different measurements

- Using Holt winters for predictive analytics

  - Holt winters comes in built as a part of Influxdb

  - Query for prediction extracted data for the last 30 days and fitted the actual curve approximately

  - Configured the seasonal pattern based on recurring pattern of data which in our case was 24 hours

Features explored in Grafana

- Panel alerts

  - Sent alert messages via Telegram notification channel

- Dashboard links

  - To link any panel or external URL to the dashboards

- Ad-hoc filters

  - Provision to use any number of value/key filters that will be automatically applied to all queries

- Template variables

  - Filter options provided as drop downs which allows to query data based on the value selected

Limitations of Influxdb

- Cannot use aggregators on the time column

  - For example, cannot subtract two values in the time column to find the time difference

- DISTINCT function operates only on fields, not on tags

- ORDER BY only works against timestamps

- SQL JOINs aren't available for InfluxDB measurements

- Time column is the constant primary key

Limitations of Grafana

- Alerts in panels cannot be configured for queries having template variables

- Template variables are fixed at the top of a dashboard and cannot be moved near the corresponding panels

- Positioning of panels inside a dashboard is not very efficient and adaptive

- Alert rules can be configured only for limited panels

Grafana vs Chronograf

- External plug-ins are not available in Chronograf but can be added in Grafana

- More visualization options available for panels in Grafana than in Chronograf

- Dynamic panel creation supported in Grafana for multiple values selected in template variables but not in Chronograf

- Alerts in Grafana are configured in the alert tab of panels whereas in Chronograf, it is stored as TICKScripts

- Alerts can be made to trigger when it changes state in Chronograf but Grafana sends alerts whenever an alert condition is met

References

- https://docs.influxdata.com/influxdb/v1.6/

- http://docs.grafana.org/features/datasources/influxdb/

- https://github.com/influxdata/docs.influxdata.com

- https://docs.influxdata.com/chronograf/v1.6/introduction/getting-started/