Elasticsearch

Elasticsearch to rozproszony silnik wyszukiwania oparty o Apache Lucene. Elasticsearch to tak naprawdę niezależna baza danych, stworzona w Javie. Rozproszony dlatego ponieważ umożliwia klastrowanie czyli uruchomienie wielu instancji tej samej aplikacji tworzącą jedną, dzięki czemu możemy stworzyć naprawdę dużą i wydajną bazę danych, która przyjmie dane wsadowe w wielu zaawansowanych formatach, a co najważniejsze umożliwi szybkie wyszukiwanie pełnotekstowe, a na dodatek będzie łatwa w obsłudze dla deweloperów, ponieważ komunikacja z Elasticsearch następuje za pomocą JSON.

Technicznie rzecz ujmując, Elasticsearch jest niezwykłą bazą danych. Na jego wyjątkowość wpływa kilka czynników:

- dane przechowywane są w tzw. indeksach, czyli zbiorach dokumentów (całość oparta jest na bibliotece Lucene),
- dane można wydobywać za pomocą zapytań Elasticsearch DSL wysyłanych za pomocą REST API,
- jest systemem rozproszonym i bardzo dobrze skaluje się horyzontalnie – możemy dokładać kolejne node'y, zaś Elasticsearch zajmie się resztą (np. odpowiednim rozszerzeniem klastra i podziałem danych pomiędzy serwery) i przyspieszy działanie.

Icinga

Icinga to narzędzie aktywnie monitorujące podzespoły danego systemu. W klasycznym podejściu, gdzie aplikacje, bazy i inne komponenty aplikacji webowej są umieszczone na serwerach, mogą być one monitorowane przez

agenta NRPE. Serwer Icinga przesyła do nich zapytania o konkretne czujki i metryki sprawdzające np. obciążenie procesora, ilość dostępnej pamięci dyskowej, długość kolejek, czas ładowania strony czy ilość zapytań do bazy danych.

W zależności od ustawień Icingi, możemy zdefiniować zadania jakie mają zostać wykonane w przypadku wystąpienia sytuacji krytycznej: od powiadomienia odpowiednich osób po wykonanie automatycznej akcji.

Grafana

Grafana to narzędzie bardziej rozbudowane graficznie, jednak gorsze pod kątem powiadomień niż Icinga.

Narzędzie to pozwala na wizualizacje danych w postaci wykresów, wskaźników, diagramów obrazujących kluczowe metryki działania aplikacji.

Źródłem danych do Grafany może być baza danych ElasticSearch, InfluxDB, SQL czy AWS CloudWatch. Dobrze skonstruowany dashboard pozwala określić obecny oraz historyczny stan całego środowiska.

Wykresy mogą być uzupełniane odpowiednimi adnotacjami, takimi jak uruchomienie nowszej wersji aplikacji, co pozwala na łatwe odnajdywanie zależności pomiędzy wydajnością środowiska, a wersją aplikacji.

Grafana nadaje się również do tworzenia wizualizacji metryk sprawdzanych rutynowo, pozwalając sprawnemu administratorowi na przewidzenie awarii – np. kończącej się pamięci czy przestrzeni dyskowej – przed ich wystąpieniem. Grafana, umożliwia także definiowanie prostych reguły powiadomień, choć znacznie bardziej ubogich niż Icinga.

Kibana

Kibana jest narzędziem służącym do eksploracji zbiorów danych, które umożliwia agregację logów z różnych źródeł.

Poza funkcjami takimi jak przeglądanie i sprawne przeszukiwanie logów, pozwala tworzyć widoki z wykresami, a także posiada moduł timelion, pozwalający na analizę trendów czy analizę danych w czasie.

Poza tworzeniem wykresów i składaniem z nich dashboardów, można skonstruować zapytanie timelion, które będzie wyszukiwało anomalie między danymi, na przykład tydzień do tygodnia.

Bazą danych dla Kibany jest zazwyczaj indeks ElasticSearch, który może być użyty również przez wspomnianą wcześniej Grafanę. Oba te narzędzia dają w połączeniu świetny zestaw do przeglądania i wizualizacji zbieranych metryk.

Prometheus

Prometeus jest bardzo dobrym narzędziem dedykowanym nowoczesnym środowiskom: automatycznie skalowanych, dynamicznych, konteneryzowanych, o niedeterministycznym czasie życia komponentów.

W relatywnie prosty sposób, można dokonać integracji z katalogiem usług wystawionych przez service discovery (np. consul, eureka). Za pomocą Prometheusa umożliwione jest także zbieranie danych, używając protokołu http(s) z wykorzystaniem rozmaitych exporterów.

Prometheus integruje się z Grafaną, posługując się nią jako silnikiem do wizualizacji danych oraz z alert managerem, jako modułem do agregacji alarmów i wysyłania powiadomień.

Graylog

Graylog to opensourcowy projekt, który jest rozwijany od 2009 roku. Jego twórcy od samego początku chcieli stworzyć system do analizy logów pochodzących z przeróżnych źródeł: systemów operacyjnych, serwerów aplikacji, firewalli sprzętowych i programowych. Dzięki temu Graylog znajduje zastosowanie zarówno podczas monitoringu stron internetowych, aplikacji webowych i wielu obszarów infrastruktury informatycznej. Jego użytkownikami mogą być nie tylko pracownicy zespołów IT, ale również działy sprzedaży i marketingu, zainteresowane m.in. wizualizacją trendów.

Graylog przetwarza logi pochodzące z wielu różnych źródeł. Podstawowym warunkiem jest jednak zgodność plików dzienników zdarzeń z powszechnie stosowanym standardem opisanym w dokumentach RFC 5424 i 3164, które definiują m.in. takie reguły jak sposób zapisu daty i nazwy hosta źródłowego. Obsługiwane są logi pochodzące z serwerów Linux/Unix, wysyłane za pomocą protokołów TCP i UDP przez serwisy syslog oraz syslog-ng. Konfiguracja tych klientów nie odbiega zbytnio od standardowej, definiującej przesyłanie logów do centralnego serwera stworzonego za pomocą wspomnianych serwisów.

Aplikacja	**Wykresy**	**Analiza logów**	**Powiadomienia**
Icinga	Nie	Nie	Tak
Grafana	Tak	Nie	Tak (w ograniczonym zakresie)
Kibana	Tak	Tak	Nie
Prometheus	Tak (można podpiąć grafanę)	Nie	Tak (przez manager)
Graylog	Tak	Tak	Nie