Jihočeská univerzita v Českých Budějovicích
Ekonomická fakulta
Katedra aplikované matematiky a informatiky

Teze bakalářské práce

Zpracování logů pomocí ELK stacku

Vypracoval: David Slavík

Vedoucí práce: doc. Ing. Ladislav Beránek CSC., MBA

České Budějovice

2023

Klíčová slova: elasticsearch, kibana, logstash, beat, log management
Anotace: Tato práce popisuje problémy s počítačovými logy a jak s nimi pracovat. Hlavním cílem této práce je ukázat, jak shromažďovat, transformovat, ukládat a zkoumat logy vytvořené aplikací nebo systémem. Představuje řešení pro správu logů s názvem ELK stack, které se skládá ze tří hlavních komponent: Elasticsearch, což je NoSQL databáze a vyhledávací engine, Kibana používaná jako grafické uživatelské rozhraní pro prohlížení, analýzu dat a také pro vytváření dashboardů. Logstash je součástí stacku, který transformuje a připravuje strukturu zpráv pro Elasticsearch. Další komponentou jsou beaty, které slouží ke sběru dat ze serverů. Součástí této práce je také návod na instalaci, konfiguraci a vytvoření pracovního prostoru s daty. Výsledek této práce by měl být užitečný pro společnosti, které se potýkají s logy, při implementaci vlastního řešení pro správu logů.

Úvod a přehled literatury

Cíl práce

Tato bakalářská práce má za úkol představit komplexní řešení Log Managementu od společnosti Elasticsearch za použití ELK stacku. Toto řešení je vhodné pro snadné auditování a ke zpětnému trasování chyb, které se mohou v IT systému vyskytnout. Kvůli centralizaci logů a možnosti pracovat s daty pomocí jednoduchého rozhraní, je možné tato data zpřístupnit i kolegům s menšími technickými znalostmi.

V praktické části je cílem práce popsat komplexně celé řešení pomocí daného log managementu a co vše je možné tímto způsobem řešit. Zvolené řešení je detailně popsáno a obsahuje implementaci jednotlivých komponent včetně jejich instalace a konfigurace. Dále je zde uvedeno, jak efektivně pracovat v grafickém rozhraní a jak toto rozhraní využít pro analýzu dat. Na závěr je uvedeno zhodnocení včetně možných alternativních řešení.

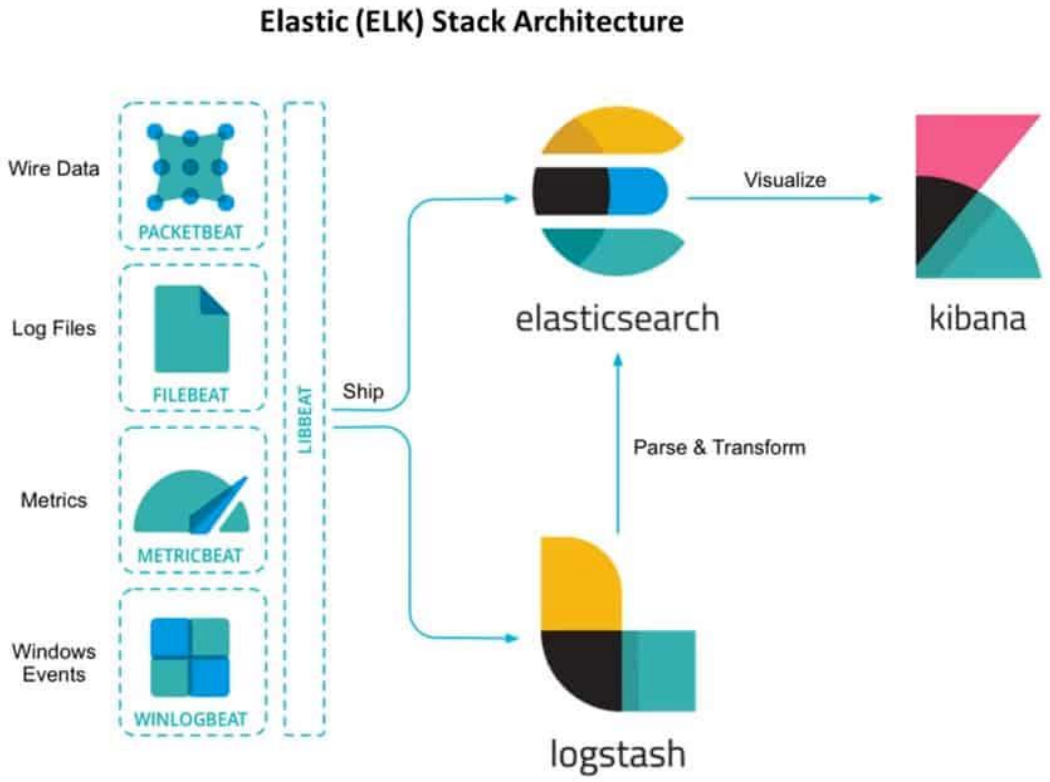
Metodika

1. Studium odborné literatury
2. Úvod do problematiky správy logů a obecný popis zvoleného řešení
3. Teoretický popis jednotlivých komponent
4. Popis implementace řešení
5. Zhodnocení a možnosti alternativ

Výsledky

Úvodem této práce je problematika logů, jejich rozřazení a jaké typy logů rozeznáváme. Jde tak o představení jejich podstaty, proč je dobré tuto oblast nezanedbat a smysluplně s logy pracovat. Efektivní práce je stejně důležité jako samotný sběr těchto dat. Pokud nebudeme mít přehledné log záznamy, ze kterých lze čerpat důležité informace, jsou záznamy spíše na obtíž. Výsledkem práce je detailně popsané řešení ELK stacku se všemi komponentami beats, logstash, elasticsearch, kibana a představení práce s logy. To obsahuje sběr dat, transformaci na požadovaný formát, parsování textu do jednotlivých polí pro lepší analýzu dat. Také fungování NoSQL databáze elasticsearch a náhled do fungování grafického rozhraní kibana.

Součástí této práce je vyřešení automatizované správy těchto dat bez potřeby věnovat značné množství času na provoz.



Obrázek 1: ELK stack architektura [1]

Pro porovnání možných řešení jsou na konci teoretické části uvedeny alternativy pro správu logů. Jedná se tak o možné doporučení, pokud by někomu nevyhovovalo toto řešení. Dále je zde uvedeno porovnání grafických rozhraní pro analýzu dat, a to Grafana a Kibana. Kdy Grafanu je možné použít jako univerzální rozhraní pro práci s mnoha typy zdrojů, kdežto kibanu je možné implementovat pouze v rámci ELK stacku.

Tabulka 1: Porovnání Grafana a Kibana [2]

Parametr hodnocení	Grafana	Kibana
Instalace	Podporuje Linux, MacOS, Windows. Jednoduché na instalaci.	Podporuje Linux, MacOS, Windows. Jednoduché na instalaci.
Procesování dat	Je používána převážně na monitoring a metriky při porovnávání v čase. Analýza dat v reálném čase. Nabízí tedy funkcionality v rámci tohoto zaměření.	Nabízí velmi efektivní možnosti vyhledávání nad daty a jejich filtrování a vizualizaci. Hodí se spíše pro vizualizaci a prohlížení uložených dat například logů. Pracuje s širokou škálou zobrazení dat pomocí mnoho nástrojů.
Procházení dat	Pro procházení dat a query dotazy používá jazyk zvaný Grafana DSL. Tento jazyk je	Používá jazyk Elasticsearch Query DSL. Jeho nevýhodou

	založen na PromQL (Prometheus Query Language).	je, že je kompatibilní pouze v ELK stacku.
Integrace datových zdrojů	Popdoruje databáze jako Prometheus, InfluxDB, Elasticsearch, SQL databáze, CSV soubory, AWS CloudWatch.	Podporuje pouze Elasticsearch.
UI a UX	Uživatelsky přívětivé prostředí, kde je možné vytvářet dashboardy, grafy a metriky. Spravovat konfiguraci.	Uživatelsky přívětivé prostředí, kde je možné vytvořit dashboardy, grafy. Lze zde také spravovat nastavení. Více uživatelsky přívětivé, jelikož umožňuje také drag and drop.
Vizualizace	Obsahuje mnoho různých typů vizualizace jako grafy, metriky a mapy.	Obsahuje mnoho různých typů vizualizace jako grafy, metriky, koláčové grafy, heat mapy, mapy, liniové grafy.
Alerting	Podporuje alerting a incident management.	Podporuje alerting a incident management.
Týmová spolupráce	Grafana umožňuje vytvářet uživatele, skupiny a jednoduše sdílet data v týmu. Je možné sledovat změny.	Kibana může pracovat s prostředím pro určité skupiny lidí, kde sdílí informace. Týmy si mohou nastavit alerty a notifikace pro sdílení informací o změnách. Kibana také nabízí nástroje jako je Slack pro komunikování v reálném čase.
Dokumentace a podpora	Velice přehledně a podrobně napsaná dokumentace. Větší podpora komunity.	Velice přehledně a podrobně napsaná dokumentace.
Licencování	Pro self-managed řešení se jedná o open source řešení zcela zdarma. Je možné provozovat základní cloudové řešení zcela zdarma. Poté jsou zde další 2 řešení Pro a Advanced. Ty nabízí funkcionality jako SSO, LDAP. Cloudové SLA a podporu. Reporty a exporty.	V rámci Basic licence je řešení zdarma. Toto řešení je nabízeno v self-managed režimu. Při potřebě více funkcionalit je možné povýšit licenci na placenou. Zde se cena odvíjí od potřeby funkcionalit. V rámci cloudového řešení spravované od Elasticu se též platí dle preferencí.

Závěr

Tato práce uvedla čtenáře do problematiky počítačových logů a jejich zpracování. Vyzdvihla důležitost logů a jak s nimi efektivně pracovat a spravovat. Po přečtení této práce by měl čtenář rozumět základům log managementu a také daného řešení od firmy Elastic. Pomocí postupu v praktické části by pak měl být dále schopen zprovoznit si své vlastní clusterové řešení, které bude plně funkční.

Seznam literatury použité v tezích

Plný seznam použité literatury je uveden v textu bakalářské práce.

1 - Elk-arch [Online]. Retrieved March 16, 2023, from https://i0.wp.com/sysadminxpert.com/wp-content/uploads/2018/05/elk-arch.jpg?ssl=1

2 - Grafana vs Kibana: How to Choose in 2023 [Online]. Retrieved March 16, 2023, from https://betterstack.com/community/comparisons/grafana-vs-kibana/

Další bibliografické údaje

Type of thesis: bachelor

Author: David Slavík

Academic Year: 2022/2023

Department: Department of Economic Informatics, Faculty of Economics, University of South Bohemia in Ceske Budejovice, Czech Republic

Thesis supervisor: doc. Ing. Ladislav Beránek CSc., MBA

Number of pages: 44

Language od thesis: CZ

Title of Thesis: Log management ELK stack

Annotation: This work describes problematic of computer logs and how to treat them. The main goal of this work is to show how to collect, transform, save, and explore logs created by application or system. It presents the log management solution named ELK stack that consists of three main components Elasticsearch which is NoSQL database and search engine, Kibana graphical user interface for viewing, analyzing data also as creating dashboards, Logstash part of the stack that transform and prepare the message structure for the Elasticsearch. Additional component is beat used for collecting the data from servers. From the theoretical part is slowly comes to the practical part. There is presented installation, configuration and creating workspace with data. The result of this work should be helpful to implement its own log management.

Key words: elasticsearch, kibana, logstash, beat, log management