



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ

FACULTY OF INFORMATION TECHNOLOGY

ÚSTAV INFORMAČNÍCH SYSTÉMŮ

DEPARTMENT OF INFORMATION SYSTEMS

**ANALÝZA VYUŽÍVÁNÍ TECHNOLOGIE
NETWORK ERROR LOGGING**

ANALYSIS OF NETWORK ERROR LOGGING DEPLOYMENT

BAKALÁŘSKÁ PRÁCE

BACHELOR'S THESIS

AUTOR PRÁCE

AUTHOR

MATEJ JURÍK

VEDOUCÍ PRÁCE

SUPERVISOR

Ing. LIBOR POLČÁK, Ph.D.

BRNO 2024

Zadání bakalářské práce



155130

Ústav: Ústav informačních systémů (UIFS)
Student: **Jurík Matej**
Program: Informační technologie
Název: **Analýza využívání technologie Network Error Logging**
Kategorie: Web
Akademický rok: 2023/24

Zadání:

1. Seznamte se s technologií Network Error Logging (NEL). Vyhledejte publikované informace o nasazení této technologie a zkušenostech s jejím provozem.
2. Seznamte se s daty poskytovanými HTTP Archive a možnostmi analýzy těchto dat. Nastudujte mechanismy procházení nejčastěji navštěvovaných stránek a technologie umožňující automatizaci prohlížeče jako je Selenium. Analyzujte výhody a nevýhody využití vlastních a externích dat o stavu webu.
3. Navrhněte postup a nástroje vedoucí k analýze nasazení technologie NEL, prezentujte je vedoucímu práce a zapracujte jeho připomínky.
4. Navrhnuté nástroje implementujte.
5. Analyzujte nasazení technologie NEL, dílčí výsledky průběžně konzultujte s vedoucím a na základě zpětné vazby vylepšete nástroje pro analýzu.
6. Vyhodnoťte dosažené výsledky a navrhněte možná pokračování práce.

Literatura:

- POLČÁK Libor a JEŘÁBEK Kamil. Data Protection and Security Issues with Network Error Logging. In: *Proceedings of the 20th International Conference on Security and Cryptography*. Řím: SciTePress - Science and Technology Publications, 2023, s. 683-690. ISBN 978-989-758-666-8.
- JEŘÁBEK Kamil a POLČÁK Libor. Network Error Logging: HTTP Archive Analysis. Preprint dostupný na <https://arxiv.org/abs/2305.01249>.
- BURNETT, S., CHEN, L., CREAGER, D. A., EFIMOV, M., GRIGORIK, I., JONES, B., MADHYASTHA, H. V., PAPAGEORGE, P., ROGAN, B., STAHL, C., a TUTTLE, J. Network Error Logging: Client-Side Measurement of End-To-End Web Service Reliability. In 17th USENIX Symposium on Networked Systems Design and Implementation, NSDI 2020, str. 985–998. USENIX Association.
- The World Wide Web Consortium. Network Error Logging. <https://w3c.github.io/network-error-logging/>, interní návrh editora ze 3. července 2023.

Při obhajobě semestrální části projektu je požadováno:
První tři body zadání včetně vypracování technické zprávy.

Podrobné závazné pokyny pro vypracování práce viz <https://www.fit.vut.cz/study/theses/>

Vedoucí práce: **Polčák Libor, Ing., Ph.D.**
Vedoucí ústavu: Kolář Dušan, doc. Dr. Ing.
Datum zadání: 1.11.2023
Termín pro odevzdání: 9.5.2024
Datum schválení: 30.10.2023

Abstrakt

V tejto práci sa zameriavame na nasadenie protokolu zvaného Network Error Logging do prevádzky stránok verejne dostupných na internete a rôzne NEL konfigurácie, s ktorými sa vyskytuje

Abstract

In this thesis, we dive deep into the state of usage of a protocol named Network Error Logging on websites that are publicly available on the internet. We also study the various configurations that the NEL protocol can be encountered with.

Klíčové slova

Network Error Logging, HTTP, Selenium, Web Crawling, Analýza Nasadenia Protokolu

Keywords

Network Error Logging, HTTP, Selenium, Web Crawling, Protocol Deployment Analysis

Citácia

JURÍK, Matej. *Analýza využívání technologie*

Network Error Logging. Brno, 2024. Bakalářská práce. Vysoké učení technické v Brně, Fakulta informačních technologií. Vedoucí práce Ing. Libor Polčák, Ph.D.

Analýza využívání technologie Network Error Logging

Prehlásenie

Prehlasujem, že som túto bakalársku prácu vypracoval samostatne pod vedením pána Ing. Libora Polčáka, Ph.D. Ďalšie informácie mi boli poskytnuté jeho kolegom, OPONENT.NAME, ktorý pomáhal s OPONENT.POMAHAL... Uviedol som všetky literárne pramene, publikácie a ďalšie zdroje, s ktorých som čerpal.

.....
Matej Jurík
01.05.2024

Podakovanie

Chcem sa týmto poďakovať vedúcemu mojej práce, Ing. Liborovi Polčákovi, Ph.D., ktorý mi poskytol svoju pomoc a odborné rady vo všetkom, kde to bolo potrebné.

Obsah

1	Úvod	2
2	Network Error Logging a relevantné technológie	3
2.1	Všeobecná problematika zlyhaní v komunikáciach typu klient-server	3
2.2	End to End Reliability Monitoring - Nevhodný nadpis, preložiť	3
2.3	Network Error Logging	4
	Základné funkčné prvky	4
	Konfigurácia	5
	Politiky a ich uskladňovanie na strane klienta	5
	Rozšírenia plánované v budúcnosti	5
2.4	Reporting API	5
3	Spôsoby získavania dát pre vyhodnocovanie aktuálneho stavu	6
3.1	HTTP Archive (+ iné zdroje v prípade potreby)	6
3.2	Teória k čomukoľvek ďalšiemu, čo sa bude používať	6
3.3	Nástroje pre automatickú analýzu stavu nasadenia NEL - teória (selenium, bs4, matplotlib...)	6
4	Súvisiace práce	7
4.1	Vedecké články a výsledky predošlých analýz	7
4.2	Hodnotenie aktuálneho stavu	7
5	Možnosti analýzy nasadenia technológie NEL	8
5.1	Výčíslenie dostupných metód pre analýzu	8
5.2	Voľba a ohraňenie konkrétnych metód skúmaných v tejto práci	8
6	Výsledky vykonanej analýzy	9
7	Testovanie vytvorených nástrojov pre automatickú analýzu	10
8	Záver	11
	Literatúra	12
A	Obsah priloženého pamäťového média	13
B	Manuál	14
C	Plagát ?	15

Kapitola 1

Úvod

Sample Úvod

[[Nasledujúca sekcia (2) svojim začiatkom splýva so záležitosťami, ktoré majú byť uvedené tu]]

Kapitola 2

Network Error Logging a relevantné technológie

[[Sekcie môžu byť zatiaľ nekoherentné a bude nutné ich prekontrolovať a napraviť]]

V tejto kapitole popisujem problematiku, v ktorej sa táto práca venuje. Jedná sa tu o možné nástrahy pri komunikácii typu klient-server, ktoré môžu nastať a spôsobiť problémy ako napríklad nedosiahnuteľnosť servera, s ktorým klient komunikáciu pôvodne nadviazal. O takýchto a podobných problémoch sa server ziaľ nemá ako dozvedieť, a preto vzniká potreba nájsť nejaký spôsob ako takéto problémy pri ich vzniku identifikovať a nahlásiť formou štruktúrovanej správy vývojárom zodpovedným za prevádzku daného servera. V prípade, že na problém takto poukázané, v nasledujúcich krokoch je možné postaviť sa k nemu s vhodnými protiopatreniami a úspešne ho odstrániť, a teda tak zároveň aj znovuspojzdiť predtým nefunkčnú komunikáciu s klientami, u ktorých sa tento problém prejavoval.

V tejto kapitole začnem s úvodom do širších spektra zavedenej problematiky, kde spomeniem technológie s podobným účelom ako Network Error Logging (ďalej označovaný iba ako NEL), aké nedostatky aktuálneho stavu problematiky riešia, no hlavne, aké majú nedostatky. Tým sa prepracujem k podstate a motivácií k nasadeniu samotného NEL-u a ďalších technológií priamo súvisiacich s ním. Jedná sa tu čisto o teoretický základ nutný pre pochopenie nasledujúcich sekcií tejto práce. Aktuálny stav využívania technológií spomínaných nižšie bude uvedený a detailne rozpísaný v kapitole 4.

2.1 Všeobecná problematika zlyhaní v komunikáciach typu klient-server

Aké problémy môžu nastať

Ako sa to prejavuje

Dôkaz, že server takéto problémy nedokáže detekovať

2.2 End to End Reliability Monitoring - Nevhodný nadpis, preložiť

Sekcia bude vyplývať z článku priloženého k odporúčenej literatúre zadania tejto BP [2]
Existujúce populárne spôsoby riešenia uvedených problémov - napr: DVP, JavaScript...

Aké problémy tieto riešenia stále nepokrývajú - napr. JavaScript sa vôbec nemusí spustiť (klient ho neobdrží od servera).

Motivácia používania NEL - aké aktuálne riešenia ponúka pre možné chyby a perspektíva do budúcnosti.

[[Table 1: Properties satisfied by different approaches for detecting service reachability problems at scale.]]

Odkaz na Table 1 - Jednou z motivácií pre vývoj štandardu NEL bolo, že ako jediný by bol schopný presne vyčíslieť koľko klientov je v danom čase (aktuálne) afektovaných výpadkom v komunikácii. Práve táto schopnosť ručí za pridanú hodnotu korektného určenia závažnosti, tým pádom aj priority konkrétného zlyhania. Správci serveru, ktorý interne používa NEL, majú dostatočný prehľad o stave siete a na základe toho môžu rozhodovať o tom, kam zamerajú svoje úsilie o riešenie závad.

2.3 Network Error Logging

[[Náhľad na NEL z hľadiska jeho špecifikácie]]

NEL je štandard pre zachytávanie a získavanie chýb a zlyhaní na úrovni webového prehliadača navrhnutý organizáciou World Wide Web Consortium (ďalej v skratke iba W3C). Tento štandard bol prvýkrát popísaný v článku publikovanom 11/02/2014 a je dodnes aktualizovaný s tým, že posledná verzia jeho špecifikácie bola vydaná práve v dobe vzniku tejto bakalárskej práce, a to presne 5/10/2023.

Hlavným cieľom NEL je poskytovať prevádzkovateľom webových serverov priamy pohľad na chybové stavy, ktoré môžu vzniknúť pri snahe klientských aplikácií komunikovať s nimi. Pri takejto klient-server komunikácii môže nastať hneď niekoľko kategórií problémov, kde v každej z nich si zaslúžia jednotlivé chyby svoje vysvetlenie. Dôležité je, že server v obyčajnom scenári takejto komunikácie nemá žiadnu možnosť dostať sa k informácií o tom, že sa niečo pokazilo, ani čo konkrétne bol samotný problém. Práve toto je cieľom napraviť pre tento štandard a tým pádom sa tvorcom tejto technológie jedná o postupné zvýšenie dostupnosti služieb poskytovaných na internete sprevádzkovaním NEL na svojich webových serveroch.

V nasledujúcich sekciách chcem vysvetliť jednotlivé detaily týkajúce sa tejto technológie, ktoré sú naprosto potrebné k porozumeniu implementačnej časti tejto práce - kapitole č.6. Pokiaľ nebude uvedené inak, v tejto kapitole čerpám hlavne z poslednej dostupnej verzie dokumentu špecifikácie NEL[3].

Základné funkčné prvky

[[Špecifikácia NEL, ktorá je jednoznačne potrebná pre účel pochopenia implementačnej časti (analýzy)]]

NEL je možné využiť pri komunikácii klienta so serverom pomocou protokolu HTTP. Jeho funkcionality je zapracovaná do užívateľských prehliadačov založených na open-source?? distribúcií projektu **Chromium**, a to menovite napríklad: Google Chrome, Opera alebo Microsoft Edge. **[[CITE]]**

Jeho vnútorné mechanizmy sú spustené práve vtedy, keď server pri vyhotovovaní novej požiadavky zašle v svojej odpovedi spolu s ostatnými aj hlavičku NEL. Tento proces sa nazýva **Policy delivery** a sprostredkuje dohodu o zbieraní, udržiavaní a nahlasovaní chýb vzniknutých pri komunikácii. Samotné politiky sú popísané vo väčšom detaile v sekcii 2.3.

Hlavička NEL vo svojej najjednoduchšej podobe musí obsahovať nasledovné položky:

- `report_to` - menom označená skupina zberačov reportov (collectors, viď 2.4)
- `max_age` - doba platnosti zaslanej politiky NEL

Tieto hlavičky musia byť zadané vo formáte `application\json`, takže ukážkový obsah HTTP hlavičky NEL môže vyzeráť napríklad takto:

```
{"report_to": "network-errors", "max_age": 604800}
```

Výpis 2.1: Ukážka obsahu najjednoduchšej/minimálnej HTTP hlavičky NEL. Akékoľvek chyby budú hlásené do skupiny `network-errors` po dobu platnosti tejto politiky, ktorá bola nastavená na 7 dní (604 800 / 60s / 60min / 24h)

[[center only the listing without the caption]]

[[V návrhu NEL je v detaile opísané, ako požiadavky na bezpečnosť ovplyvnili jeho návrh a funkcionality. Toto je nutné spomenúť, no myslím, že sa v tejto práci tomu nie je potreba venovať]]

Konfigurácia

Všetky položky, ktoré je možné v NEL hlavičke uviesť.

Politiky a ich uskladňovanie na strane klienta

Rozšírenia plánované v budúcnosti

Nová verzia NEL, ktorá bude schopná spolupracovať s Reporting API v1 (momentálne funguje iba s v0)

2.4 Reporting API

Doplnkové info, ktoré je skrátka nutné spomenúť [4]

Kapitola 3

Spôsoby získavania dát pre vyhodnocovanie aktuálneho stavu

V tejto kapitole sa budem venovať...

3.1 HTTP Archive (+ iné zdroje v prípade potreby)

KJerabek Repository

3.2 Teória k čomukoľvek ďalšiemu, čo sa bude používať

Našiel som napríklad **crawler.ninja** - web crawled data od 2021 do 1.10.2023 dostupné na stiahnutie

3.3 Nástroje pre automatickú analýzu stavu nasadenia NEL - teória (selenium, bs4, matplotlib...)

Selenium je super, tu je k nemu nejaka teoria

Kapitola 4

Súvisiace práce

V tejto kapitole sa budem venovať...

4.1 Vedecké články a výsledky predošlých analýz

LPolčák a KJerabek [5][6]

Ďalšie články a stránky, ktoré som našiel

4.2 Hodnotenie aktuálneho stavu

Čo sa zistilo

Čo chýba

Čo by bolo fajn zistiť

Čo je potrebné zdôrazniť

Kde sa dá s nasadením pomôcť

...

Kapitola 5

Možnosti analýzy nasadenia technológie NEL

V tejto kapitole sa budem venovať...

5.1 Vyčíslenie dostupných metód pre analýzu

Domény, dodávatelia, zdroje dát apod [1].

Stratégie získavania dát

Časová náročnosť a výdavky

Potrebné technológie

5.2 Voľba a ohraničenie konkrétnych metód skúmaných v tejto práci

Na čo sa v práci sústredím

Čo sa javilo ako vhodná voľba

Čo bolo dosiahnuteľné

Návrh postupovania v analýze a vývoja potrebných nástrojov

Kapitola 6

Výsledky vykonanej analýzy

Dáta, percentá, konfigurácie, top N XY, grafy

Kapitola 7

Testovanie vytvorených nástrojov pre automatickú analýzu

Loading...

Kapitola 8

Záver

Sample Záver

Literatúra

- [1] BEDNÁŘ, M. *Automatické testování projektu JavaScript Restrictor*. Brno, 2020. Diplomová Práce. Vysoké učení technické v Brně, Fakulta informačních technologií. Vedúci práce ING. LIBOR POLČÁK, P. Dostupné z: <https://www.fit.vut.cz/study/thesis-file/22376/22376.pdf>.
- [2] BURNETT, S., CHEN, L., CREAGER, D. A., EFIMOV, M., GRIGORIK, I. et al. Network Error Logging: Client-Side Measurement Of End-To-End Web Service Reliability. *Proceedings of the 17th USENIX Symposium on Networked Systems Design and Implementation, NSDI 2020*. USENIX Association. Február 2020.
- [3] CREAGER, D. a CLELLAND, I. *Network Error Logging* [online]. 5. októbra 2023 [cit. 2023-11-01]. Dostupné z: <https://www.w3.org/TR/2023/WD-network-error-logging-20231005/>.
- [4] CREAGER, D., CLELLAND, I. a WEST, M. *Reporting API* [online]. 12. mája 2023 [cit. 2023-11-04]. Dostupné z: <https://www.w3.org/TR/2023/WD-reporting-1-20230512/>.
- [5] JEŘÁBEK, K. a POLČÁK, L. Network Error Logging: HTTP Archive Analysis. SciTePress - Science and Technology Publications ??? Máj 2023. Dostupné z: <https://arxiv.org/abs/2305.01249>.
- [6] POLČÁK, L. a JEŘÁBEK, K. Data Protection and Security Issues With Network Error Logging. *Proceedings of the 20th International Conference on Security and Cryptography*. Rím: SciTePress - Science and Technology Publications. Máj 2023. ISBN 978-989-758-666-8. Dostupné z: <https://arxiv.org/abs/2305.05343>.

Príloha A

Obsah priloženého pamäťového média

Sekcia bude vyplnená podľa návodu

Príloha B

Manuál

Vytvorený nástroj X: manuál na jeho sprevádzkovanie

Príloha C

Plagát ?

Bude nutné spraviť plagát ?