

Mendelova univerzita v Brně
Provozně ekonomická fakulta

Webová architektura v prostředí vysoké zátěže

Diplomová práce

Vedoucí práce:
Ing. Michael Štencl, Ph.D.

Bc. Jakub Škrha

Brno 2012

Text poděkování

Text prohlášení

místo a datum prohlášení

.....

Abstract

Citace práce v anglickém jazyce

abstrakt práce v anglickém jazyce **Abstrakt**

Citace práce v českém jazyce

abstrakt práce v českém jazyce

Contents

1	Úvod	6
2	Nežádoucí vlivy a důsledky vysoké zátěže	7
3	Zneužití vysoké zátěže	8
4	Tří a vícevrstvá architektura webových systémů	9
5	Aplikační vrstva	10
5.1	Webová aplikační architektura MVC	10
5.2	Druhy aplikačních vrstev	10
5.3	Webový server Apache	10
5.4	Programovací jazyk PHP	10
5.5	Ajax a webové služby	10
6	Dabázová vrstva	11
6.1	Optimalizace SQL dotazů	11
6.2	Indexace	11
6.3	Partitioning	11
6.4	Replikace	11
6.5	Druhy databází	11
6.6	PostgreSQL	11
7	Webová cache	12
7.1	Druhy cache	12
7.2	Typy obsahu	12
7.3	Proxy cache a cache prohlížeče	12
7.4	Reverzní proxy cache	12
7.4.1	Nginx	12
7.5	Aplikační a distribuovaná cache	12
7.5.1	Memcached	12
8	Další vrstvy aplikace	13
8.1	CDN	13
8.2	NoSQL Databáze	13
8.3	Vyhledávání	13
9	Virtualizace	14
10	Load balancing	15
11	Cloud Computing	16

12 Praktická část a experimenty	17
12.1 Aplikace a její vrstvy	17
12.2 Testovací nástroje	17
12.2.1 XHPProf	17
12.2.2 Siege	17
12.2.3 PostgreSQL Explain	17
12.3 Aplikační vrstva PHP	17
12.3.1 Optimalizace PHP pomocí APC	17
12.3.2 Dosažené výsledky	17
12.4 Databázová vrstva	17
12.4.1 Optimalizace databáze	17
12.4.2 Dosažené výsledky	18
12.5 Aplikační cache	18
12.5.1 Optimalizace aplikace pomocí Memcached	18
12.5.2 Diagram tříd pro aplikaci s podporou Memcached	18
12.5.3 Dosažené výsledky	18
12.6 Reverzní proxy cache	18
12.6.1 Nasazení a konfigurace NGINX s Memcached	18
12.6.2 Význam Ajax a webových služeb pro NGINX	18
12.6.3 Dosažené výsledky	18
13 Diskuze	19
14 Závěr	20

1 Úvod

Něco na téma jak vzniká vysoká zátěž na internetu, jaký je vliv internetu, rostoucí zájem o internet a stoupající zátěž. Uvidíme, napíšu to jako poslední.

2 Nežádoucí vlivy a důsledky vysoké zátěže

Co může způsobit vysoká zátěž, pád serverů, projektů, škoda pro firmy

3 Zneužití vysoké zátěže

Různé typy útoků DoS, Anonymous

4 Tří a vícevrstvá architektura webových systémů

Třívrstvá architektura je začátek, ovšem v práci nestačí. Obrázky a příklady architektur projektů s vysokou zátěží. Lehké představení dalších možných vrstev.

5 Aplikační vrstva

Aplikační vrstva je základ. Realizuje a sestavuje odpovědi na požadavky uživatelů.

5.1 Webová aplikační architektura MVC

Webové projekty jsou dnes tvořeny na této architektuře. Proč a co to je.

5.2 Druhy aplikačních vrstev

Něco o JAVA Spring, C sharp, Pearl, Python a PHP.

5.3 Webový server Apache

Popis webového serveru apache, jeho úlohy, proč se používá, proč o něm píšu a proč ho já používám pro svoji práci.

5.4 Programovací jazyk PHP

Je to rozšířený jazyk pro svoji jednoduchost, čitelnost, ohebnost a dynamičnost.

5.5 Ajax a webové služby

Ajax na straně klienta, webová služba na straně aplikace. Stránky jsou dynamické a skládané z více požadavků.

6 Dabázová vrstva

Úloha perzistence dat. Stará se o data, odlehčení aplikace, apod.

6.1 Optimalizace SQL dotazů

Proč a jak optimalizovat dotazy

6.2 Indexace

Proč jsou dobré indexy

6.3 Partitioning

Kdy a jak používat partitioning

6.4 Replikace

Něco o replikacích a proč se používají.

6.5 Druhy databází

Něco málo o MySQL, Oracle, PostgreSQL, MSSQL.

6.6 PostgreSQL

Co je to za databázi a proč jsem si ji vybral pro svoji práci.

7 Webová cache

Proč webovou cache, proč se používá a k čemu je dobrá. Urychlení odezvy, odlehčení jiným požadavkům na aplikaci, apod.

7.1 Druhy cache

Popsat jednotlivé druhy browser, proxy, reverse proxy a aplikační cache

7.2 Typy obsahu

Popsat typy obsahu pro cache. Dynamický vs. statický

7.3 Proxy cache a cache prohlížeče

Klientská část, kde a kdo ji instaluje a kdy a jak se používá.

7.4 Reverzní proxy cache

Serverová část, kde a kdo ji instaluje a kdy a jak se používá.

7.4.1 Nginx

Popis NGINX a proč jsem ho vybral.

7.5 Aplikační a distribuovaná cache

Popis aplikačních a distribuovaných cache, jaký je jejich smysl a kde je jejich použití

7.5.1 Memcached

Popis Memcached a proč jsem ji vybral

8 Další vrstvy aplikace

K čemu jsou další vrstvy

8.1 CDN

Content delivery network obrázky a stream.

8.2 NoSQL Databáze

K čemu slouží a kde najdou své uplatnění.

8.3 Vyhledávání

Z vyhledávání se také dělá další vrstva.

9 Virtualizace

Projekty dnes neběží vždy na jednom serveru, ale na více virtualizovaných serverech. Proč tomu tak je.

10 Load balancing

Nevím jestli k této kapitole se vůbec dostanu, uvidíme. Každopádně serverů bývá vždy několik a jak zajišťovat toto rozložení zátěže.

11 Cloud Computing

Budoucnost projektů, startupů, vše řešeno cloudem. AWS

12 Praktická část a experimenty

Úvod do toho, že se budu praktickou částí snažit dosáhnout nasymolování vytížené webové architektury a optimalizovat jednotlivé vrstvy v rámci možností.

12.1 Aplikace a její vrstvy

Představení aplikace, její síťové schéma, jednotlivé vrstvy s popisem, domény, apod.

12.2 Testovací nástroje

Popis toho co sleduji testovacími nástroji

12.2.1 XHProf

K profilování

12.2.2 Siege

Pro generování zátěže

12.2.3 PostgreSQL Explain

Vysvětlení sql dotazů

12.3 Aplikační vrstva PHP

Tady se budeme snažit optimalizovat PHP kód.

12.3.1 Optimalizace PHP pomocí APC

Jak se chovala aplikace bez APC a co dosahnu APC

12.3.2 Dosažené výsledky

Toto bude vždy na konci každého experimentu, grafy, časy, screeny, apod.

12.4 Databázová vrstvy

Tady se budeme snažit optimalizovat databázi.

12.4.1 Optimalizace databáze

Co jsem použil pro optimalizaci

12.4.2 Dosažené výsledky

Toto bude vždy na konci každého experimentu, grafy, časy, screeny, apod.

12.5 Aplikační cache

Tady se budeme snažit optimalizovat databázy.

12.5.1 Optimalizace aplikace pomocí Memcached

Co jsem udělal s Memcached, jednotlivé vrstvy modelu, popis toho čeho chci dosáhnout.

12.5.2 Diagram tříd pro aplikaci s podporou Memcached

Diagram tříd mé aplikace

12.5.3 Dosažené výsledky

Toto bude vždy na konci každého experimentu, grafy, časy, screeny, apod.

12.6 Reverzní proxy cache

Tady se budeme snažit použít reverzní proxy cache.

12.6.1 Nasazení a konfigurace NGINX s Memcached

Jak jsem co dělal s NGINX, problémy, řešení, návrh architektury a jak to ovlivňuje aplikační vrstvu

12.6.2 Význam Ajax a webových služeb pro NGINX

Ajax komunikuje s NGINX, proč a jak.

12.6.3 Dosažené výsledky

Toto bude vždy na konci každého experimentu, grafy, časy, screeny, apod.

13 Diskuze

Diskutovaná řešení, jak je možné je kombinovat, apod.

14 Závěr

Závěr ve smyslu nákladů a přínosů, kdy je lepší co. Uvidíme, napíšu jako poslední.