

Na tomto místě bude oficiální zadání vaší práce

- Toto zadání je podepsané děkanem a vedoucím katedry,
- musíte si ho vyzvednout na studijním oddělení Katedry počítačů na Karlově náměstí,
- v jedné odevzdané práci bude originál tohoto zadání (originál zůstává po obhajobě na katedře),
- ve druhé bude na stejném místě neověřená kopie tohoto dokumentu (tato se vám vrátí po obhajobě).

České vysoké učení technické v Praze
Fakulta elektrotechnická
Katedra kybernetiky



Bakalářská práce

Ověření škálovatelnosti Google App Engine aplikací

Jakub Škvára

Vedoucí práce: Ing Marek Šmíd

Studijní program: Softwarové technologie a management, Bakalářský

Obor: Softwarové inženýrství

19. května 2011

Poděkování

Zde můžete napsat své poděkování, pokud chcete a máte komu děkovat.

Prohlášení

Prohlašuji, že jsem práci vypracoval samostatně a použil jsem pouze podklady uvedené v přiloženém seznamu.

Nemám závažný důvod proti užití tohoto školního díla ve smyslu §60 Zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon).

V Praze dne 27.5.2011

.....

Abstract

Translation of Czech abstract into English.

Abstrakt

Abstrakt práce by měl velmi stručně vystihovat její podstatu. Tedy čím se práce zabývá a co je jejím výsledkem/přínosem.

Očekávají se cca 1 – 2 odstavce, maximálně půl stránky.

Obsah

Seznam obrázků

Seznam tabulek

Kapitola 1

Úvod

Úvod charakterizující kontext zadání, případně motivace.

Výsledná struktura vaší práce a názvy a rozsahy jednotlivých kapitol se samozřejmě budou lišit podle typu práce a podle konkrétní povahy zpracovávaného tématu. Níže uvedená struktura práce odpovídá *práci implementační*.

Kapitola 2

Teorie: Obecný popis cloudu

2.1 Novinka jménem cloud computing

V poslední době se čím dál tím více začíná mluvit o cloud computingu. Jedná se o nový typ hostingu a ukládání webového obsahu vůbec. Oproti klasickému způsobu, kde máme jeden konkrétní server, na určeném místě, se svojí danou pamětí, procesorem a pevným diskem, nám tento nový přístup umožňuje nezabývat se hardwarem, ale mluvíme takto o platformě.

Definice se dosti různí, takže použiji verzi od Národního institutu standardů a technologií (National Institute of Standards and Technology) [?], která volně přeložena zní: cloud computing je způsob poskytování sdílených škálovatelných zdrojů (výpočetní kapacita, uložení, služby, aplikace, ...), k nimž je přístupováno skrz síť a které jsou uživateli k dispozici ihned na vyžádání.

Mezi hlavní výhody je považováno snižování nákladů a zvyšování efektivity. Nemusíme vlastnit hardware, za jehož pořízení a správu je potřeba vynaložit nemalé finanční náklady, přičemž většina zdrojů není plně zatížena. Zvyšování efektivity se projevuje hlavně placením jen za využití zdroje. Pokud bychom měli vlastní infrastrukturu, tak v době nižší aktivity nevyužíváme možnosti serverů naplno a platíme vlastně za nevyužití zdroje. Naopak v cloudu jsou naše prostředky sdíleny s ostatními a v době neaktivity můžou být nabídnuty někomu jinému.

2.2 Infrastructure as a Service, Platform as a Service, Software as a Service

Existují různé nabídky cloudových řešení pro efektivní využívání zdrojů hardware pro více aplikací. Nejzákladnější je Iaas - Infrastructure as a Service (Infrastruktura jako Služba) - jedná se například o Amazon EC2 [?] cloud, kde platíme jen za spotřebované zdroje, které reálně využijeme a na hardware si můžeme sami instalovat co potřebujeme.

U PaaS - Platform as a Service (Platforma jako Služba) již nemáme přístup k hardwaru, to znamená že nelze instalovat žádný software, ale máme zde připravená API pro různé služby které můžeme využívat a většinou i další nástroje pro vývoj na lokálním stroji a pro deployment.

Nejvíce jsme od fungování služby odstíněni u SaaS - Software as a Service (Software jako Služba) - jedná se například o online e-mailové služby jako `gmail.com` anebo `seznam.cz`, obrázkové galerie jako `flickr.com` anebo `rajce.cz`, tedy služby které používáme prostřednictvím internetu a nezajímá nás jak a kde jsou data uložena a nemáme ponětí jak jsou naprogramované.

2.3 Horizontální a vertikální škálování

Nové služby a především sociální sítě s rychlým nárůstem uživatelů a potřebě dynamicky měnit počet serverů donutily programátory a správce přemýšlet o novém způsobu ukládání a organizace dat. Pokud náš server nestíhá, tak máme dvě možnosti, jak tento problém řešit. Prvním z řešení je vertikální škálování, to znamená že koupíme silnější hardware, přidáme procesor, paměť a další komponenty podle potřeby. Nevýhodou tohoto řešení ale je, že takto nejde infrastruktura rozšiřovat do nekonečna, protože po čase narazíme na hardwarové limity.

Druhou z možností je nakoupení více serverů, nemusí být ani velmi výkonné, ale řešení spočívá v propojení těchto počítačů dohromady, čímž můžeme zvětšovat naši infrastrukturu bez omezení. Můžeme toto řešení přirovnat k problému z běžného života, kdy potřebujeme převést určitý počet osob z jednoho místa na druhé. Můžeme objednat autobus, který jednoduše řeší tento problém. Pokud počet osob naroste, tak můžeme objednat autobus s větší kapacitou, jenže pokud se bude počet osob zvyšovat, tak časem již nenajdeme tak velký autobus pro přepravu všech osob. Takže jako řešení budeme muset objednat více vozidel, ale ty poté budeme moci efektivně zaplňovat podle počtu osob. Nevýhoda tohoto řešení spočívá v složitějším správě infrastruktury, potřebujeme software, který se stará o propojení, synchronizaci a spolupráci všech částí systému, nebo pokud nějaký stroj přestane fungovat, musíme zajistit aby ho ihned zastoupil jiný se stejnou funkcí jako předchozí.

Pro cloud computing se obecně vžila značka mraku (v angličtině znamená cloud mrak) a vznikla proto, že na obrázcích a schématech se ve většině případů mrakem značí internet a vzdálená zařízení, které nejsou uloženy u nás. A to právě z toho důvodu, že jsou tyto služby většinou přístupné skrze internet a přistupujeme k nim vzdáleně.

2.4 Různé druhy pohledu na cloud computing

Cloud computing se jako každá jiná novinka potýkal s různými názory od těch pozitivních až po ty z negativní. Někteří tvrdili že se jedná jen o buzzword¹, který má nalákat nové zákazníky, jiní predikovali že se takovýto princip nemůže nikdy uchytit. Zde je pár výroků z doby, kdy nebyl tolik rozšířený:

The interesting thing about Cloud Computing is that we've redefined Cloud Computing to include everything that we already do... I don't understand what we would do differently in the light of Cloud Computing other than change the wording of some of our ads.

Larry Ellison, Oracle Corporation CEO, Wall Street Journal, 26. září 2008

¹módní slovo

Larry Ellison říká, že cloud computing je jen pojmenování toho, co již dávno používáme a že jediná změna, která je potřeba je změna textů u reklam. Je pravda, že některé velké společnosti, jako třeba Oracle anebo Google používali tento přístup již mnohem dříve než vzniknul samotný název, ale pravdou je, že v posledních dvou letech se začal cloud computing používat masivněji a to hlavně díky možnosti pronájmu cloudů. Nyní si mohou programátoři vyzkoušet pracovat s cloudy a použít je i pro své menší aplikace, bez nutnosti spravovat a starat se o velké množství strojů.

A lot of people are jumping on the (cloud) bandwagon, but I have not heard two people say the same thing about it. There are multiple definitions out there of “the cloud.”

Andy Isherwood, HP's vice president, ZDnet News, 11. prosinec 2008

Andy Isherwood naznačuje, že není přesně daná definice toho, co ještě cloud je a co už není. Je to způsobeno tím, že po vzniku tohoto názvu chtěl každý s více než jedním serverem označovat svoje služby jako cloudové. To vedlo spíše ke zmatení, ale v poslední době se toto slovní spojení ustálilo pro farmu serverů se snadnou škálovatelností a jednoduchou možností přidat novou instanci.

It's stupidity. It's worse than stupidity: it's a marketing hype campaign. Somebody is saying this is inevitable — and whenever you hear somebody saying that, it's very likely to be a set of businesses campaigning to make it true.

Richard Stallman, Founder of GNU Project and Free Software Foundation, The Guardian, 29. září 2008

Richard Stallman má na cloud poněkud negativní názor a pro The Guardian vyslovil názor že se jedná o hloupost a jde jen o nafouknutou bublinu podpořenou businessovými kampaněmi. Kritizoval hlavně uložení dat mimo naši vlastní kontrolu a nutnost spolehnout se na společnost, které dáváme naše data a aplikace k dispozici. Nikdo nám nemůže na sto procent zaručit, že bude tato společnost fungovat i po několika letech. Navíc jsme většinou vázáni na API rozhraní, služby a možnosti platformy určené danou společností.

Pokud budeme chtít přenést službu k jinému poskytovateli cloudu, bude nám to s největší pravděpodobností činit nemalé potíže, v angličtině se používá termín *lock-in*, což znamená doslova zamknutí. Sice v poslední době vznikají návrhy na jednotná rozhraní a sjednocení rozhraní těchto služeb, aby byl přechod co nejjednodušší, ale ty se bohužel zatím nesetkaly s větším rozšířením. Do budoucna by to mohla být jedna z klíčových vlastností při rozhodování kterou službu zvolit.

2.5 Porovnání Amazon Web Services, Windows Azure a Google App Engine

Největší konkurenti Google App Engine jsou Amazon Web Services - EC2 (Elastic Compute Cloud) a Microsoft Azure. Pokud budeme mít zakázku, pro kterou je nejvýhodnější použít cloudovou infrastrukturu, budeme se pravděpodobně rozhodovat mezi těmito třemi, jedná

se o velké a známé společnosti s rozsáhlým zázemím. Je tedy málo pravděpodobné, že by ze dne na den přestaly provozovat svoje služby, což může být velký problém u menších nebo méně známých společností.

2.5.1 Amazon Web Services - Elastic Compute Cloud

Amazon EC2 je spíše blíže modelu IaaS, takže dostaneme hardware s kterým si můžeme dělat co chceme, instalovat libovolný software a musíme si ho sami spravovat. Největší výhodou je rychlé přidávání nových serverů kdykoliv potřebujeme a platba jen za spotřebované zdroje. Navíc není tento cloud vázán žádnými API a omezeními, takže pokud budeme chtít přesunout naši aplikaci na náš server, tak nebudeme muset měnit aplikaci a to platí i naopak, tedy pro přesun aplikace na cloud. Jedná se čistě o pronájem hardwaru. Neýhoda Amazon Web Services je v tom, že platíme hned jakmile nahrajeme naši aplikaci, neexistují žádné volné kvóty jako u App Engine. Amazon v rámci svých služeb nabízí i další možnosti, například speciální relační i nerelační uložště a další produkty, celý seznam je možné najít na stránce <http://aws.amazon.com/products/>.

2.5.2 Microsoft Azure

Microsoft Azure je již více podobný App Engine, jedná se o PaaS, máme zde již připravené prostředí pro několik jazyků: .NET (C# a VisualBasic), C++, PHP, Ruby, Python a Java. Výhodou je, že můžeme použít klasickou relační databázi nazvanou SQL Azure Database (SAD), což se vyplatí pokud migrujeme nějaký projekt postavený na relační databázi, ale tento typ je hůře škálovatelný. Vedle SAD můžeme použít i Azure Storage, která obsahuje nerelační tabulky, tabulky pro velké objemy dat (blobs) a fronty (queues). Azure má speciální staging prostředí, kde můžeme přímo na cloudu vyvíjet aplikaci a nedostane se k ní nikdo, dokud není připravena na spuštění. V tomto prostředí také můžeme spouštět aplikaci v debug režimu, což nám umožňuje například nastavovat breakpointy, pokud potřebujeme aplikaci ladit přímo na cloudu. Azure také umožňuje propojení s Microsoft Live službami a s vývojovým prostředím Microsoft Visual Studio. Stejně jako u Amazonu platíme ihned jakmile nasadíme aplikaci do plného provozu, aby ji mohli vidět i ostatní. Azure je určitě výhodné použít, pokud vyvíjíme aplikace v technologiích od Microsoftu, naše infrastruktura je na těchto technologiích založena anebo pokud používáme jako vývojový nástroj Visual Studio.

2.5.3 Google App Engine

Oproti dvěma předchozím má App Engine hlavní výhodu v tom, že nemusíme platit ihned jak aplikaci nahrajeme na cloud. Jsou zde nastavené kvóty, které jsou velmi vysoké a je potřeba velká návštěvnost pro jejich přesáhnutí, což je pro začínající aplikaci výhodné obzvláště v České republice. Takže pokud začínáme se startupem, nemusíme se v počátcích obávat velkých investic a pokud bude náš projekt úspěšný a bude obsluhovat velký počet požadavků, tak poté budeme muset platit jen za přesáhnuté limity, které se každých 24 hodin vynulují. Některé limity jsou nastaveny napevno a nejdou zvýšit ani za poplatek, je to kvůli tomu, aby se nemohlo stát, že jedna aplikace zaneprázdní celý cloud, což by mělo

za následek zpomalení i ostatních aplikací. Tyto limity jsou naštěstí velmi vysoké, například datastore API má maximálně 141 241 791 volání za den a 784 676 volání za minutu. [tabulka kvót, cena] Standardně se aplikace nahraje na adresu jako doména třetího řádu ve tvaru `jmeno-aplikace.appspot.com` a pokud potřebujeme můžeme nastavit i doménu vlastní. Další výhodou App Engine je možnost provozovat více verzí stejné aplikace najednou. Nahrají se pak jako poddomény, takže například `verze.jmeno-aplikace.appspot.cz`. Tyto verze mohou běžet na cloudu zároveň a v administraci se dá nastavit, která bude výchozí. Toto do jisté míry nahrazuje staging area z Azure, výhodou zde je, že můžeme mít libovolný počet verzí. App Engine bohužel podporuje jen jazyky Java, Python a Go. Díky různým projektům, které jsou schopny přeložit další jazyky do javovského bytekódu, zde tedy můžeme spouštět velké množství dalších jazyků i když je to vykoupeno nižší rychlostí. Můžeme zde tedy používat jazyky jako: Groovy, Scala, Ruby s pomocí JRuby, PHP díky projektu Quercus (viz dále), JavaScript za pomoci Rhino a další. Jedním z důvodů přidání Javy do App Engine byla právě možnost běhu dalších jazyků nad Java Virtual Machine. Co se Javy týká, tak nejsou povoleny všechny třídy, nemůžeme například vytvářet nová vlákna, nemůžeme vytvářet nové soubory a některá volání třídy `System` nedělají nic, například `System.exit()` a `System.gc()`. Seznam všech povolených tříd je možné najít na adrese <http://code.google.com/appengine/docs/java/jrewhitelist.html>. Kvůli těmto omezením bohužel nemůžeme použít všechny knihovny, anebo musíme použít upravenou verzi pro App Engine. Seznam nejpoužívanějších frameworků a knihoven s popisem zda jsou kompatibilní a případné nastavení pro App Engine je dostupný na adrese <http://groups.google.com/group/>. Na App Engine nemáme jistotu, že bude naše aplikace přímo připravena, protože kvůli tomu, aby se šetřily prostředky, jsou nahrány jen aktivní a využívané aplikace, se po určité době neaktivity aplikace odstaví a nahradí ji jiná. Můžeme si za poplatek zařídit, že naše aplikace bude vždy k dispozici, protože každé nahrání stojí čas. Můžeme tedy využívat většinu frameworků, ale kvůli tomuto nahrávání se každý kód navíc negativně projeví na době prvního přístupu k aplikaci, což je velmi znatelné především u rozsáhlých frameworků.

2.6 API služby

Abychom mohli propojit naši aplikaci s ostatními máme na App Engine velké množství API sloužících k různým účelům. Pojďme si nyní projít jaké možnosti máme. Některé z nich asi vůbec nevyužijeme, ale některé jsou velmi užitečné a důležité.

Asi jednou z nejpoužívanějších služeb, pokud pomineme Datastore, je Memcache. Jedná se o možnost, jak zrychlit častý přístup do databáze. Jedná se o key-value cache, která je přibližně desetkrát až stokrát rychlejší, než přístup k Datastore. Nehodí se ale pro vše, protože data z ní po vypršení zmizí. Místo klasického cachování na disk, tedy máme možnost ukládat data do paměti. Implementace je podle standardu JSR-107, takže bude kompatibilní s dalšími knihovnami.

Mezi další užitečné služby patří Mail, posílání mailů funguje klasicky pomocí tříd `javax.mail`. Můžeme přidávat i přílohy. Některé soubory jsou z bezpečnostních důvodů zakázány, ale všechny běžně používané jsou povoleny. Přijímání e-mailů se ošetřuje pomocí servletu. Ve `web.xml` se nastaví servlet pro url `/_ah/mail/jmeno-emailu` a e-mail má tvar: `jmeno-emailu@id-aplikace` a to bohužel i v případě, že máme nastavenou naši vlastní doménu. Příchozí e-mail se chová

jako HTTP požadavek, takže v servletu se musíme zpracování postarat sami, podle toho co potřebujeme.

Podobně jako Mail funguje i služba pro práci s XMPP protokolem². Jedná se o otevřený standardizovaný protokol Jabberu postavený na XML. Princip na App Engine je podobný jako s e-mailem, identifikátor příjemce je JID, který se dá získat z e-mailu. Podporovány jsou i další funkce, jako posílání pozvánek, nastavování statusů a další. Přijímáme pomocí servletu nastaveného na adresu: `/_ah/xmpp/message/chat/`.

Kvůli absenci JMS³ máme na App Engine možnost použít Task Queues. Jedná se o frontu úloh, které by mohly zpomalit náš systém, takže je výhodnější je zpracovat později. Fronta funguje následovně: pomocí Task Queues API přidáme do fronty URL naší aplikace, můžeme jí předávat parametry stejně jako u klasické URL⁴. Provedení úlohy z fronty je záležitostí servletu⁵, na který je URL nastavena pomocí `web.xml`. Vykonání úlohy je omezeno deseti minutami, toto by měl být dostatečný limit pro běžné úlohy. Pokud servlet vrátí HTTP status mimo rozmezí 200 - 299, což znamená chybu, tak se úloha zavolá znovu aby proběhla v pořádku. Pokud potřebujeme informovat aplikaci o dokončení úlohy, musíme to řešit pomocí Datastore.

Pokud potřebujeme naše stránky integrovat s nějakou webovou službou, anebo používat veřejná REST API, použijeme URL Fetch. Jedná se o klasické java.net API, můžeme použít HTTP i HTTPS, většinu běžných portů a samozřejmě i všechny HTTP metody: GET, POST, PUT, HEAD i DELETE pro správné fungování REST rozhraní. K požadavku můžeme nastavovat i vlastní HTTP hlavičky.

Na některá data, jako například obrázky, nebo velké soubory se Datastore nehodí, maximální velikost entity je 1MB. Právě kvůli tomuto účelu můžeme na App Engine použít Blobstore, jedná se o úložiště pro velké soubory do velikosti až 2GB. Blobstore je plně oddělen od Datastore. Nahrávání je velice jednoduché, stačí použít formulář s prvkem `<input type="file" />` a atribut action formuláře nastavíme pomocí `<%= blobstoreService.createUploadUrl("/upload")%>`. Blobstore už se sám postará aby se soubor nahrál na správné místo. Zobrazovat soubory můžeme pomocí `blobstoreService.serve(blob-key)`, potřebujeme k tomu klíč souboru. Tato služba umí vybrat všechny uložené soubory. Práce je velmi podobná jako s Datastore, akorát s tím rozdílem, že zde pracujeme s velkými soubory. Jedná se o užitečnou službu, protože dříve než existovala tato služba, se ukládání řešilo rozdělením do mnoha samostatných entit o velikosti 1MB a při zobrazení jsme je museli zase nazpět složit. Toto všechno se dělo na aplikační úrovni, takže jsme museli ošetřit všechny chybové případy a bylo vše velmi zdlouhavé a nepohodlné. Takto se o ukládání souborů stará AppEngine sám a nám stačí jednoduché API.

S předchozím Blobstore souvisí i další služba: Images. Jedná se o možnost úpravy obrázků přímo na serveru. Obrázky se načítají z Blobstore anebo můžeme službě předat přímo pole `byte[]`. Můžeme takto aplikovat jednoduché transformace jako je změna velikosti, otočení, oříznutí, skládní obrázků a také magická funkce "I'm feeling lucky", která změní nastavení tmavých a světlých barev a k tomu také zvýší kontrast obrázku, výsledkem jsou pak více barevnější obrázky. Upravené obrázky můžeme přímo posílat uživatelům anebo uložit do

²Extensible Messaging and Presence Protocol (<http://xmpp.org/about-xmpp/technology-overview/>)

³Java Message Service (http://en.wikipedia.org/wiki/Java_Message_Service)

⁴Uniform Resource Locator (http://en.wikipedia.org/wiki/Uniform_Resource_Locator)

⁵Servlet je komponenta napsaná v jazyce Java, určená pro spouštění na webovém serveru

Blobstore, pokud se budou zobrazovat čtětji. Služba obsahuje základní transformace, ale na vytvoření náhledů nebo menší úpravy jako zvětšení a zmenšení obrázku, které jsou pravěpodobně na webových stránkách nejpoužívanější, se Image API hodí výborně.

Pokud potřebujeme u naší aplikace vytvořit sekci jen pro přihlášené uživatele, nabízí nám k tomu App Engine možnost použít Users API a interní přihlašovací mechanismus Googlu využívaný u všech aplikací této společnosti, například tedy `gmail.com`, `youtube.com` a dalších. Použití je jednoduché, pokud uživatel není přihlášený, tak ho přesměrujeme na vygenerovanou přihlašovací stránku. Ta je stejná pro všechny služby Googlu, zadáme e-mail a heslo. Poté můžeme nastavit, které všechny údaje o sobě chceme aplikaci, do které se právě přihlašujeme, poskytnout. Nakonec nás stránka přesměruje zpět na naši aplikaci. Nyní můžeme o uživateli zjistit základní informace: přihlašovací e-mail a jednoznačný identifikátor ID. Odhlašování funguje stejně jako přihlašování, přesměrujeme uživatele na odhlašovací stránku Google, která nás následně po odhlášení znovu přesměruje, tentokrát na naši aplikaci.

Pokud chceme dát možnost přihlašování i pro uživatele, kteří nevlastní účet u Google, můžeme použít OAuth protokol. Ten není vázaný na konkrétní společnost, takže si můžeme vybrat kterého poskytovatele si vybereme. Jedná se o možnost přihlášení uživatelským jménem a heslem jiné aplikace a v naší aplikaci jen kontrolujeme token. Výhoda tohoto způsobu je, že uživateli stačí jeden účet pro více aplikací, nemusí si tak pamatovat hesla pro každou stránku na které má účet. S touto službou se pracuje velmi podobně jako s předchozím API.

App Engine obsahuje Capabilities API pomocí nějž můžeme zjistit, zda daná služba běží anebo ne. Můžeme tak ošetřit případ, kdy zrovna probíhá údržba anebo výpadek a služby jsou nedostupné. Jsou zde dostupné informace o těchto službách: Blobstore, čtení z Datastore, zápis do Datastore, Images, Mail, MemCache, TaskQueue, Url Fetch a XMPP.

Pro lepší spolupráci s klientskou stranou máme k dispozici Channel API. To se stará o trvalé spojení JavaScriptu na stránce se serverem Googlu, aniž by se musel klient stále dotazovat serveru. Toto se hodí, pokud chceme uživatele informovat o nastalé akci, toto se hodí například na hry pro více hráčů a internetové chaty.

Posledním rozšířením je Multitenancy API, to nám dává možnost používat jmenné prostory pro naše data, můžeme je aplikovat na: Datastore, Memcache, Task Queue a Blobstore. Můžeme tak provozovat více oddělených stránek z jedné aplikace a pro každou stránku budeme mít speciální jmenný prostor. Data se tak nebudou překrývat a budou správně oddělena.

2.7 Omezení cloudu

Pokud se rozhodneme naši službu provozovat na cloudu, tak musíme již od návrhu počítat s jistými omezeními a odlišnou strukturou aplikace, než na jakou jsme zvyklí z klasických aplikací. Většina z těchto omezení plyne z požadavku na škálovatelnost aplikací.

Pro většinu programátorů je největším problémem databáze, používá se totiž poměrně nový typ - NoSQL databáze (pro češtinu nejlépe asi sedí překlad: nerelační databáze). Databáze používaná na App Engine se nazývá Big Table⁶, jedná se o vícerozměrnou distribuovanou mapu optimalizovanou pro rychlé čtení a pomalejší zápis, protože u běžných

⁶<http://en.wikipedia.org/wiki/BigTable>

aplikací je čtení dat mnohem častější operace. Google navíc toto uložisko využívá i pro své ostatní služby. Naprostá většina dnešních aplikací využívá relační databázi, pravděpodobně jednu z trojice nejpoužívanějších: Oracle, PostgreSQL, MySQL anebo MS-SQL, všechny tyto databáze mají tabulky a pomocí konstrukce JOIN je můžeme navzájem libovolně spojovat. Nevýhoda tohoto řešení ale spočívá v tom, že pro tuto operaci potřebujeme všechny tabulky, které v dotazu spojujeme. V praxi se tedy používá samostatný stroj jen pro databázi. U škálovatelných aplikací, se ale nemůžeme spolehnout na to, že jsou všechny tabulky na jednom místě, mohou totiž být v různých datacentrech na různých kontinentech. Řešením je tedy ukládání všech potřebných dat do jedné tabulky anebo přiřazování do skupin, které se budou spojovat a databáze se sama postará o to, aby byla data uložena ve stejném datacentru. App Engine proto používá speciální dotazovací jazyk šitý na míru tomuto uložišti: GQL - Google Query Language⁷, což je podmnožina SQL jazyka pro App Engine Datastore. Nenajdeme v něm samozřejmě operátor JOIN a s ním spojené konstrukce.

Mezi další omezení patří žádná anebo omezená možnost vyhledávání nad sloupci databáze. Není totiž zaručeno jakou strukturu sloupců bude tabulka mít. Toto jde obejít vytvořením speciální tabulky se slovy a v kterých sloupcích se vyskytují, ale je to dosti složité a musíme se o vše starat sami. Pokud potřebujeme vyhledávat na internetové stránce, je jednodušší použít internetový vyhledávač například Google, Bing anebo český Seznam. Všechny jmenované mají nástroj pro vyhledávání podle domény, takže stačí jen přidat formulář na naše stránky. Pokud potřebujeme vyhledávat v našich interních datech budeme muset použít nějaké rozsáhlejší řešení.

2.8 Vývoj pro App Engine

Pokud se rozhodneme vytvářet naše aplikace pro App Engine, tak máme k dispozici poměrně vyspělou infrastrukturu. Google oficiálně podporuje Eclipse plugin pro App Engine⁸, ale dostupné jsou i plně funkční pluginy pro NetBeans IDE⁹ a také pro vývojové prostředí IntelliJ IDEA¹⁰. Pomocí těchto pluginů můžeme jednoduše vyvíjet aplikaci na našem domácím stroji bez nutnosti připojení k internetu. Součástí je totiž jednoduchý webový server simulující App Engine, jedná se o upravený Jetty server¹¹. Máme zde úplně stejné API jako na produkčním serveru a pomocí URL `http://localhost/_ah/admin` máme k dispozici jednoduchou administraci obsahující správce Datastore, správce Task Queues a další. Data se lokálně ukládají do souboru `.bin` přímo ve složce `/build` projektu. Deploy na lokální prostředí je stejný jako u jakéhokoli jiného serveru, tedy pomocí tlačítka v IDE, anebo můžeme použít některý z buildovacích nástrojů, jako například ANT anebo Maven. Pro upload přímo na produkční prostředí je možné také použít přímo plugin, stejně tak jako je integrováno tlačítko pro lokální upload, tak je zde možnost uploadu přímo na App Engine. Vše probíhá nahráním výsledného `war` souboru aplikace na speciální URL. Pokud bychom chtěli integrovat nahrávání do jiného nástroje, máme možnost provést upload pomocí skriptu pro příkazovou řádku. Ten provádí upload pomocí `jar` souboru, takže není problém celý deployment integrovat

⁷<http://code.google.com/appengine/docs/python/datastore/gqlreference.html>

⁸Google Plugin for Eclipse - <http://code.google.com/appengine/docs/java/tools/eclipse.html>

⁹NetBeans support for Google App Engine - <http://kenai.com/projects/nbappengine/pages/Home>

¹⁰Google App Engine Integration for IntelliJ - <http://plugins.intellij.net/plugin/?id=4254>

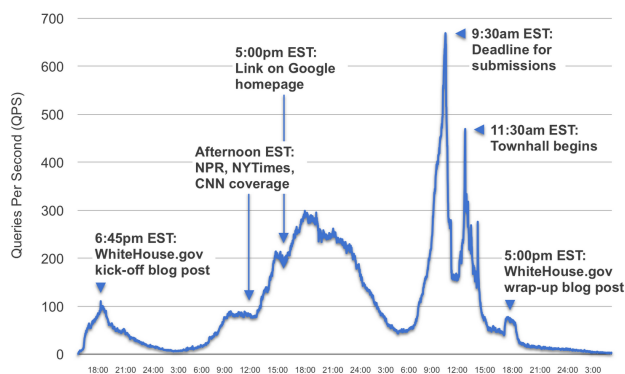
¹¹Jetty je odlehčený open source HTTP server a servlet kontejner napsaný v Javě

do naší infrastruktury. Dále můžeme mít libovolný počet verzí aplikace, všechny jsou na URL `verze.jmeno-aplikace.appspot.com`, kde `verze` je jakýkoliv řetězec definovaný ve `web.xml` a výchozí možnost se nastavuje v administraci přímo na App Enginu. Máme zde navíc oproti `localhost`¹² mnoho různých nastavení a statistik. Pro každou aplikaci, kterých můžeme k jednomu Google účtu mít až deset je zde podrobný přehled návštěv a spotřebovaných prostředků, počet aktivních instancí, logy, přehled a správa Datastore, nastavení aplikace a další.

2.9 Zajímavé aplikace

Nyní představím zajímavé aplikace a stránky, které můžeme na App Engine cloudu najít. Nacházejí se zde zajímavé experimenty, jako například běh různých jazyků nad JVM (Java Virtual Machine) až po stránky s velkým zatížením. Nejzajímavější z nich je pravděpodobně stránka www.officialroyalwedding2011.org, založená k příležitosti svatby anglického prince Williama a Catherine Middleton. V pátek 29. dubna, tedy v den oddání, bylo na hosting generováno 2 000 požadavků za vteřinu a dohromady bylo zobrazeno 15 milionů stránek od 5,6 milionu uživatelů. I přes tento nápor stránka běžela bez problémů a bez ovlivnění více jak 200 000 dalších aplikací běžících na stejném cloudu, které všechny dohromady za den vygenerovaly více jak 1,5 miliardy stránek. [?]

Další podobnou zkouškou pro App Engine byla aplikace Google Moderator. Tato aplikace běžela dva dny v březnu roku 2009 na stránce www.whitehouse.gov. Jednlo se o hlasovací systém určený pro obyvatele USA, kde může kdokoli vložit svůj dotaz a další uživatelé pak hlasují o tom, které dotazy jsou nejlepší. Vítězné otázky byly dne 29. března zodpovězeny prezidentem Barackem Obamou. Během 48 hodin zadalo 92 934 uživatelů 104 073 otázek a ohodnotilo je 3 605 984 hasy. Při nejvyšší zátěži obsloužil App Engine 700 dotazů za vteřinu. [?]



Obrázek 2.1: Zatížení App Engine serverů v době spuštění aplikace Google Moderator pro Bílý Dům [?]

¹²označení serveru, který běží na našem lokálním počítači

Nyní bych rád zmínil některé zajímavé a užitečné aplikace, které jsou ideální pro umístění v cloudu. Jednou z nich je i Socialwok (www.socialwok.com), jedná se o obdobu facebooku pro práci, můžeme zde sdílet naši práci v Google Apps (Docs, Calendar, Spreadsheet) se spolupracovníky a ti mohou do našich dokumentů zasahovat. Jedná se o zajímavou myšlenku a praktické využití sociálních sítí.

Podobnou službou je i Giftag (www.giftag.com), ta umí uložit část webové stránky a tu pak sdílet s dalšími. Existuje doplněk pro internetové prohlížeče, který zjednoduší uložení stránky. Všechny uložené části navíc můžeme organizovat a přidávat do seznamů. Tato služba nám může pomoci, pokud pracujeme na výzkumu anebo potřebujeme udělat prezentaci na které pracuje více lidí.

Nyní z jiného soudku, největší uplatnění získaly cloudy díky sociálním sítím, další služba je určena právě pro ně. Jedná se o BuddyPoke! (www.buddypoke.com), umožňuje nám dát si na náš profil trojrozměrný obrázek postavičky s popisem jak se cítíme. Tuto aplikaci můžeme najít na všech používaných sociálních sítích, které vkládání obrázků dovolují, tedy: Facebook, Orkut, MySpace, hi5, Netlog, Ning a dalších. Tato aplikace nemá žádný přínos, ale díky velkému množství podporovaných sociálních sítí a jejich velké oblibě v poslední době je tato služba velice úspěšná.

Poslední zmíněná aplikace je naopak velmi přínosná. Mnoho vývojářů by rádo vidělo na App Engine podporu pro skriptovací jazyk PHP, ten se bohužel vývojáři v nejbližším čase přidat neplánují. Projekt Quercus (quercus.caucho.com) umožňuje právě běh PHP na JVM a vývojáři připravili speciální verzi pro App Engine, kde můžeme spouštět PHP s podporou některých Java frameworků.

Další zajímavé a úspěšné projekty jsou na stránce: <http://code.google.com/appengine/casestudies.html>

Kapitola 3

Implementace: Vývoj v App Engine

3.1 Výběr cloudu

Pro praktickou ukázkou a prověření implementace jsem vybral cloud App Engine od společnosti Google. Nejdůležitějším důvodem pro mne byla možnost nahrát plnohodnotnou aplikaci bez nutnosti vynaložení jakýchkoliv nákladů na hosting. Nastavené kvóty od kterých je nutné platit jsou vysoké (odpovídají několika stovkám tisíců požadavků za den) a stále se postupně zvyšují. Je tedy možné s tímto cloudem libovolně experimentovat a nahrávat různé testovací aplikace. Mnoho Java vývojářů hledá pro svoje malé aplikace hosting, kde by nemuseli platit, anebo mohli nahrát více aplikací. To je ale problém webového Java světa. Server počítá pouze s jednou aplikací a je často nutné si pamatovat prostředky a další zdroje v rámci aplikace mezi požadavky. Javovský server tak počítá s tím, že si alokuje většinu dostupné paměti systému. Výhodou je, že nemusíme při každém požadavku vytvářet nové spojení do databáze a k ostatním prostředkům. Naproti tomu u skriptovací jazyků (jako například PHP nebo Python), se při každém požadavku provede celý kód znovu. Na serveru tak může běžet několik aplikací a zároveň se neovlivňují. V praxi se toho běžně využívá a proto je možné provozovat velké množství takovýchto aplikací na jediném hosting. App Engine nám umožňuje stejný princip pro Javu, běh více aplikací na stejném hardware. Cena za tuto možnost je nutnost uvolnění zdrojů z nepoužívaných aplikací, aby nezabírali místo ostatním. To App Engine sám hlídá a po určité době neaktivity aplikaci odstaví a nahraje aktivní.

3.2 Aplikace

Cílem této bakalářské práce je ověřit možnosti a omezení škálovatelných aplikací. Nejdůležitějším hlediskem je celková rychlost a odezva aplikace v závislosti na počtu souběžných požadavků. Tedy jak bude aplikace a celý cloud reagovat na zvýšený nápor požadavků, jak se s tím cloud vyrovná a zda bude služba stále použitelná. Také jsem se zaměřil na rychlost čtení a ukládání do Datastore, pokud je klasická aplikace vystavena vysoké zátěži, je právě databáze úzkým hrdlem (takzvaný *bottleneck*) a přestává stíhat jako první.

Kromě ověření rychlosti a škálovatelnosti bylo také cílem vytvořit rozsáhlejší aplikaci a vyzkoušet všechny úskalí, které nám platforma App Engine staví do cesty. To znamená im-

plementace běžných požadavků na aplikace, například M:N¹ relace mezi entitami, zamykání dat proti přepsání a další požadavky, které jsou na aplikace běžně kladeny. Po vytvoření tohoto základu a překonání všech zádrhelů již bude jednoduché takovou aplikaci upravit pro jiné požadavky, anebo rozšířit o nové funkce.

Jako nejvhodnější řešení pro vyzkoušení implementace jsem vybral jednoduchý systém pro správu obsahu (CMS - Content Management System). Můžeme si tak vytvořit webovou stránku s plnou administrací, tedy s možností přidávání, úpravy a mazání stránek. Ke každé stránce můžeme přidat šablonu. Mimo samotných stránek lze přidávat, upravovat a mazat i novinky, jedná se o kratší zprávy, pro které je zbytečné vytvářet samostatnou stránku. Ke každé stránce lze také přiřadit štítek (tag), tato možnost nahrazuje kategorie. Výhodou oproti kategoriím, kde může být jeden článek pouze v jedné kategorii je, že článek může být označen více štítky a zároveň jeden štítek může být přiřazen k více článkům. Z těchto údajů pak můžeme generovat takzvané *tag cloudy* (oblaky štítků)², kde velikost štítku určuje jeho význam, čím více článků štítek označuje, tím je důležitější a výraznější. Z hlediska implementace se jedná o vazbu M:N, která se v NoSQL databázích musí implementovat složitěji než v relačních a bylo potřeba vymyslet nejlepší řešení.

3.3 Výběr frameworku: Slim3

Než jsem začal aplikaci psát, zjišťoval jsem, které knihovny a frameworky³ bych mohl použít na ulehčení práce. Kvůli omezením nejsou podporovány všechny knihovny a některé potřebují speciální úpravy anebo nastavení. Další nevýhodou je, že čím je náš kód rozsáhlejší, tím déle bude trvat načítání, pokaždé když bude App Engine nahrávat aplikaci do aktivního stavu. Tímto jsem eliminoval známé, ale velké frameworky jako jsou Spring⁴ a Seam⁵. Poté jsem hledal mezi menšími, ale žádný nevyhovoval úplně potřebám App Engine. Bohužel je tento cloud relativně mladý a tak pro něj neexistují frameworky, anebo jsou málo známé. Myslel jsem tedy že použiji přímočaré řešení pomocí servletů a naprogramuji aplikaci od základů. Naštěstí jsem ale náhodou narazil na framework Slim3⁶ určený přesně pro potřeby App Engine.

Slim3 je MVC⁷ framework optimalizovaný pro App Engine, přináší rozšíření i jednodušší práci s Datastore API a další vylepšení. Jeho hlavní koncepty jsou: *simple* (jednoduchý) a *less is more* (méně je více), tedy že jednoduchost a srozumitelnost vedou ke správnému designu. Framework se drží Paretova pravidla, tedy že 80% aplikace, pramení z 20% práce, takže se snaží co nejvíce zjednodušit oněch 20% a ve zbytku nechává programátorovi volnost.

Nejvíce nám framework pomáhá při práci s Datastore. Pokud s uložištěm chceme pracovat, máme na App Engine možnost použít klasické JPA⁸ s těmito omezeními: nemůžeme

¹Many-to-Many - druh relace, kde více entit může být spojeno s více entitami (M:N), v relačních databázích je pro propojení použita spojovací tabulka

²http://en.wikipedia.org/wiki/Tag_cloud

³Framework je ucelený soubor knihoven a kódu pomáhající vytvořit aplikaci

⁴Spring Framework - <http://www.springsource.org/>

⁵Seam Framework - <http://seamframework.org/>

⁶Slim3 - <http://sites.google.com/site/slim3appengine/>

⁷Model View Controller - <http://en.wikipedia.org/wiki/Model-view-controller>

⁸JPA - Java Persistence API - http://en.wikipedia.org/wiki/Java_Persistence_API

použít vztah *many-to-many*, v dotazu nemůžeme použít `JOIN` a agregační funkce (`GROUP BY`, `HAVING`, `SUM`, `AVG`, `MAX` a `MIN`) a polymorfické dotazy (slouží k získání podtřídy). Druhou možností je použít JDO⁹, jedná se o obecnější obdobu JPA, kde můžeme náš doménový model ukládat do různých uložišť - relačních, nerelačních, objektových databází, XML i do obyčejných souborů. Programátorská práce je stejná jako s JPA, definujeme model pomocí anotací¹⁰ a dotazujeme se pomocí objektu `Query`. Poslední možností je použít přímo Datasore API, které je optimalizované pro práci s BigTable, Dovoluje nám měnit strukturu entit za běhu. Znamená to, že se nemůžeme spolehnout na to, že je daný sloupec v entitě obsažen, ani na jeho typ. Sloupce u entity je vlastně kolekce objektů, v Javě by BigTable vypadala následovně: `Map<Key, Set<Object>>`. Toto je důležité uvědomit si a může nám to přinést potíže v začátcích, dobrou radou je ukládat si přímo do objektu i verzi schématu, při změně v aplikaci se totiž schémata uložených entit nezmění.

Další změnou jsou transakce, pro správné fungování potřebujeme, aby byl všechny typy entity používané v transakci na jednom stroji. Toto může být problém, jelikož Google má několik datacenter a není tak jisté, že budou všechny entity pohromadě. Kvůli tomuto zavádí App Engine takzvané *entity groups* (skupiny entit). Entity přiřadíme do stejné skupiny tak, že nové entity nastavíme tzv. *parent entity* (rodičovskou entitu), její klíč pak bude složen z klíče samotné entity a klíče rodičovské entity. Pokud se pokusíme provést operace v transakci na entitách z rozdílných skupin, vyhodí App Engine výjimku.

⁹JDO - Java Data Objects - http://en.wikipedia.org/wiki/Java_Data_Objects

¹⁰Anotace v Jazyce java představují speciální konstrukce začínající `@` a přidávající ke kódu dále zpracovatelné metainformace

Kapitola 4

Realizace

Popis implementace/realizace se zaměřením na nestandardní části řešení.

Kapitola 5

Testování

- Způsob, průběh a výsledky testování.
- Srovnání s existujícími řešeními, pokud jsou známy.

Kapitola 6

Závěr

- Zhodnocení splnění cílů DP/BP a vlastního přínosu práce (při formulaci je třeba vzít v potaz zadání práce).
- Diskuse dalšího možného pokračování práce.

Příloha A

Testování zaplnění stránky a odsazení odstavců

Tato příloha nebude součástí vaší práce. Slouží pouze jako příklad formátování textu.

Určitě existuje nějaká pěkná latinská věta, která se k tomuhle testování používá, ale co mají dělat ti, kteří se nikdy latinsky neučili? Určitě existuje nějaká pěkná latinská věta, která se k tomuhle testování používá, ale co mají dělat ti, kteří se nikdy latinsky neučili? Určitě existuje nějaká pěkná latinská věta, která se k tomuhle testování používá, ale co mají dělat ti, kteří se nikdy latinsky neučili?

Určitě existuje nějaká pěkná latinská věta, která se k tomuhle testování používá, ale co mají dělat ti, kteří se nikdy latinsky neučili? Určitě existuje nějaká pěkná latinská věta, která se k tomuhle testování používá, ale co mají dělat ti, kteří se nikdy latinsky neučili? Určitě existuje nějaká pěkná latinská věta, která se k tomuhle testování používá, ale co mají dělat ti, kteří se nikdy latinsky neučili?

Určitě existuje nějaká pěkná latinská věta, která se k tomuhle testování používá, ale co mají dělat ti, kteří se nikdy latinsky neučili? Určitě existuje nějaká pěkná latinská věta, která se k tomuhle testování používá, ale co mají dělat ti, kteří se nikdy latinsky neučili? Určitě existuje nějaká pěkná latinská věta, která se k tomuhle testování používá, ale co mají dělat ti, kteří se nikdy latinsky neučili?

Určitě existuje nějaká pěkná latinská věta, která se k tomuhle testování používá, ale co mají dělat ti, kteří se nikdy latinsky neučili? Určitě existuje nějaká pěkná latinská věta, která se k tomuhle testování používá, ale co mají dělat ti, kteří se nikdy latinsky neučili? Určitě existuje nějaká pěkná latinská věta, která se k tomuhle testování používá, ale co mají dělat ti, kteří se nikdy latinsky neučili? Určitě existuje nějaká pěkná latinská věta, která se k tomuhle testování používá, ale co mají dělat ti, kteří se nikdy latinsky neučili? Určitě existuje nějaká pěkná latinská věta, která se k tomuhle testování používá, ale co mají dělat ti, kteří se nikdy latinsky neučili? Určitě existuje nějaká pěkná latinská věta, která se k tomuhle testování používá, ale co mají dělat ti, kteří se nikdy latinsky neučili?

Určitě existuje nějaká pěkná latinská věta, která se k tomuhle testování používá, ale co mají dělat ti, kteří se nikdy latinsky neučili? Určitě existuje nějaká pěkná latinská věta, která se k tomuhle testování používá, ale co mají dělat ti, kteří se nikdy latinsky neučili? Určitě existuje nějaká pěkná latinská věta, která se k tomuhle testování používá, ale co mají dělat ti, kteří se nikdy latinsky neučili? Určitě existuje nějaká pěkná latinská věta, která se k tomuhle testování používá, ale co mají dělat ti, kteří se nikdy latinsky neučili? Určitě existuje nějaká pěkná latinská věta, která se k tomuhle testování používá, ale co mají dělat ti, kteří se nikdy latinsky neučili?

Určitě existuje nějaká pěkná latinská věta, která se k tomuhle testování používá, ale co mají dělat ti, kteří se nikdy latinsky neučili? Určitě existuje nějaká pěkná latinská věta, která se k tomuhle testování používá, ale co mají dělat ti, kteří se nikdy latinsky neučili? Určitě existuje nějaká pěkná latinská věta, která se k tomuhle testování používá, ale co mají dělat ti, kteří se nikdy latinsky neučili? Určitě existuje nějaká pěkná latinská věta, která se k tomuhle testování používá, ale co mají dělat ti, kteří se nikdy latinsky neučili? Určitě existuje nějaká pěkná latinská věta, která se k tomuhle testování používá, ale co mají dělat ti, kteří se nikdy latinsky neučili?

Určitě existuje nějaká pěkná latinská věta, která se k tomuhle testování používá, ale co mají dělat ti, kteří se nikdy latinsky neučili? Určitě existuje nějaká pěkná latinská věta, která se k tomuhle testování používá, ale co mají dělat ti, kteří se nikdy latinsky neučili? Určitě existuje nějaká pěkná latinská věta, která se k tomuhle testování používá, ale co mají dělat ti, kteří se nikdy latinsky neučili? Určitě existuje nějaká pěkná latinská věta, která se k tomuhle testování používá, ale co mají dělat ti, kteří se nikdy latinsky neučili? Určitě existuje nějaká pěkná latinská věta, která se k tomuhle testování používá, ale co mají dělat ti, kteří se nikdy latinsky neučili?

Příloha B

Pokyny a návody k formátování textu práce

Tato příloha samozřejmě nebude součástí vaší práce. Slouží pouze jako příklad formátování textu.

Používat se dají všechny příkazy systému L^AT_EX. Existuje velké množství volně přístupné dokumentace, tutoriálů, příruček a dalších materiálů v elektronické podobě. Výchozím bodem, kromě Googlu, může být stránka CSTUG (Czech Tech Users Group). Tam najdete odkazy na další materiály. Většinou dostačující a přehledně organizovanou elektronikou dokumentaci najdete například na.

Existují i různé nadstavby nad systémy T_EX a L^AT_EX, které výrazně usnadní psaní textu zejména začátečníkům. Velmi rozšířený v Linuxovém prostředí je systém Kile.

B.1 Vkládání obrázků

Obrázky se umísťují do plovoucího prostředí **figure**. Každý obrázek by měl obsahovat **název** (`\caption`) a **návěští** (`\label`). Použití příkazu pro vložení obrázku `\includegraphics` je podmíněno aktivací (načtením) balíku `graphicx` příkazem `\usepackage{graphicx}`.

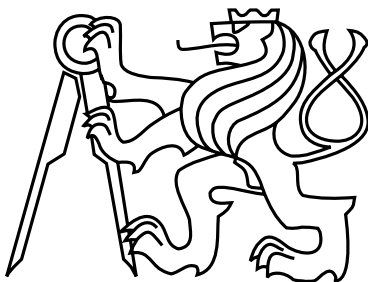
Budete-li zdrojový text zpracovávat pomocí programu `pdflatex`, očekávají se obrázky s příponou `*.pdf`¹, použijete-li k formátování `latex`, očekávají se obrázky s příponou `*.eps`.²

Příklad vložení obrázku:

```
\begin{figure}[h]
\begin{center}
\includegraphics[width=5cm]{figures/LogoCVUT}
\caption{Popiska obrazku}
\label{fig:logo}
```

¹`pdflatex` umí také formáty PNG a JPG.

²Vzájemnou konverzi mezi snad všemi typy obrázku včetně změn velikostí a dalších vymožeností vám může zajistit balík ImageMagick (<http://www.imagemagick.org/script/index.php>). Je dostupný pod Linuxem, Mac OS i MS Windows. Důležité jsou zejména příkazy `convert` a `identify`.



Obrázek B.1: Popiska obrázku

DTD	construction	elimination
	in1 A B a:sum A B in1 A B b:sum A B	case([_:A] a) ([_:B] a) ab:A case([_:A] b) ([_:B] b) ba:B
+	do_reg:A -> reg A	undo_reg:reg A -> A
*, ?	the same like and + with empty_el:empty	the same like and + with empty_el:empty
R(a,b)	make_R:A->B->R	a: R -> A b: R -> B

Tabulka B.1: Ukázka tabulky

```
\end{center}
\end{figure}
```

B.2 Kreslení obrázků

Zřejmě každý z vás má nějaký oblíbený nástroj pro tvorbu obrázků. Jde jen o to, abyste dokázali obrázek uložit v požadovaném formátu nebo jej do něj konvertovat (viz předchozí kapitola). Je zřejmě vhodné kreslit obrázky vektorově. Celkem oblíbený, na ovládání celkem jednoduchý a přitom dostatečně mocný je například program Inkscape.

Zde stojí za to upozornit na kreslicí programe Ipe, který dokáže do obrázku vkládat komentáře přímo v latexovském formátu (vzroce, stejné fonty atd.). Podobné věci umí na Linuxové platformě nástroj Xfig.

Za pozornost ještě stojí schopnost editoru Ipe importovat obrázek (jpg nebo bitmap) a krelit do něj latexovské popisky a komentáře. Výsledek pak umí exportovat přímo do pdf.

B.3 Tabulky

Existuje více způsobů, jak sázet tabulky. Například je možno použít prostředí `table`, které je velmi podobné prostředí `figure`.

Zdrojový text tabulky ?? vypadá takto:

```

\begin{table}
\begin{center}
\begin{tabular}{|c|l|l|}
\hline
\textbf{DTD} & \textbf{construction} & \textbf{elimination} \\
\hline
 $\mid$  & \verb+in1|A|B a:sum A B+ & \verb+case([_:A]a)([_:B]a)ab:A+\\
& \verb+in1|A|B b:sum A B+ & \verb+case([_:A]b)([_:B]b)ba:B+\\
\hline
 $\$$  & \verb+do_reg:A -> reg A+ & \verb+undo_reg:reg A -> A+\\
\hline
 $\$,?\$$  & the same like  $\mid$  &  $\$$  & the same like  $\mid$  &  $\$$  \\
& with \verb+empty_el:empty+ & with \verb+empty_el:empty+\\
\hline
R(a,b) & \verb+make_R:A->B->R+ & \verb+a: R -> A+\\
& & \verb+b: R -> B+\\
\hline
\end{tabular}
\end{center}
\caption{Ukázka tabulky}
\label{tab:tab1}
\end{table}
\begin{table}

```

B.4 Odkazy v textu

B.4.1 Odkazy na literaturu

Jsou realizovány příkazem `\cite{odkaz}`.

Seznam literatury je dobré zapsat do samostatného souboru a ten pak zpracovat programem bibtex (viz soubor `reference.bib`). Zdrojový soubor pro bibtex vypadá například takto:

```

@Article{Chen01,
  author   = "Yong-Sheng Chen and Yi-Ping Hung and Chiou-Shann Fuh",
  title    = "Fast Block Matching Algorithm Based on
              the Winner-Update Strategy",
  journal  = "IEEE Transactions On Image Processing",
  pages    = "1212--1222",
  volume   = 10,
  number   = 8,
  year     = 2001,
}

@Misc{latexdocweb,

```

```

author = "",
title = "{\LaTeX} --- online manuál",
note = "\verb|http://www.cstug.cz/latex/lm/frames.html|",
year = "",
}
...

```

Pozor: Sazba názvů odkazů je dána BibTeX stylem (`\bibliographystyle{abbrv}`). BibTeX tedy obvykle vysází velké pouze počáteční písmeno z názvu zdroje, ostatní písmena zůstanou malá bez ohledu na to, jak je napíšete. Přesněji řečeno, styl může zvolit pro každý typ publikace jiné konverze. Pro časopisecké články třeba výše uvedené, jiné pro monografie (u nich často bývá naopak velikost písmen zachována).

Pokud chcete BibTeXu napovědět, která písmena nechat bez konverzí (viz `title = "{\LaTeX} --- online manuál"` v předchozím příkladu), je nutné příslušné písmeno (zde celé makro) uzavřít do složených závorek. Pro přehlednost je proto vhodné celé parametry uzavírat do uvozovek (`author = "..."`), nikoliv do složených závorek.

Odkazy na literaturu ve zdrojovém textu se pak zapisují:

```

Podívejte se na \cite{Chen01},
další detaily najdete na \cite{latexdocweb}

```

Vazbu mezi soubory `*.tex` a `*.bib` zajistíte příkazem `\bibliography{}` v souboru `*.tex`. V našem případě tedy zdrojový dokument `thesis.tex` obsahuje příkaz `\bibliography{reference}`.

Zpracování zdrojového textu s odkazy se provede postupným voláním programů `pdflatex <soubor>` (případně `latex <soubor>`), `bibtex <soubor>` a opět `pdflatex <soubor>`.³

Níže uvedený příklad je převzat z dříve existujících pokynů studentům, kteří dělají svou diplomovou nebo bakalářskou práci v Grafické skupině.⁴ Zde se praví:

```

...
j) Seznam literatury a dalších použitých pramenů, odkazy na WWW stránky, ...
Pozor na to, že na veškeré uvedené prameny se musíte v textu práce
odkazovat -- [1].

```

Pramen, na který neodkazujete, vypadá, že jste ho vlastně nepotřebovali a je uveden jen do počtu. Příklad citace knihy [1], článku v časopise [2], stati ve sborníku [3] a html odkazu [4]:

```
[1] J. Žára, B. Beneš;, and P. Felkel.
```

```

    Moderní počítačová grafika. Computer Press s.r.o, Brno, 1 edition, 1998.
    (in Czech).

```

³První volání `pdflatex` vytvoří soubor s koncovkou `*.aux`, který je vstupem pro program `bibtex`, pak je potřeba znovu zavolat program `pdflatex (latex)`, který tentokrát zpracuje soubory s příponami `.aux` a `.tex`. Informaci o případných nevyřešených odkazech (cross-reference) vidíte přímo při zpracovávání zdrojového souboru příkazem `pdflatex`. Program `pdflatex (latex)` lze volat vícekrát, pokud stále vidíte nevyřešené závislosti.

⁴Několikrát jsem byl upozorněn, že web s těmito pokyny byl zrušen, proto jej zde přímo necituji. Nicméně příklad sám o sobě dokumentuje obecně přijímaný konsensus ohledně citací v bakalářských a diplomových pracích na KP.

- [2] P. Slavík. Grammars and Rewriting Systems as Models for Graphical User Interfaces. *Cognitive Systems*, 4(4--3):381--399, 1997.
- [3] M. Haindl, Š. Kment, and P. Slavík. Virtual Information Systems. In *WSCG'2000 -- Short communication papers*, pages 22--27, Pilsen, 2000. University of West Bohemia.
- [4] Knihovna grafické skupiny katedry počítačů:
<http://www.cgg.cvut.cz/Bib/library/>

... abychom výše citované odkazy skutečně našli v (automaticky generovaném) seznamu literatury tohoto textu, musíme je nyní alespoň jednou citovat: Kniha , článek v časopisu , příspěvek na konferenci , [www odkaz](#).

Ještě přidáme další ukázkou citací online zdrojů podle české normy. Odkaz na wiki o frameworkích a ORM. Použití viz soubor `reference.bib`. V seznamu literatury by nyní měly být živé odkazy na zdroje. V `reference.bib` je zcela nový typ publikace. Detaily dohledal a dodal Petr Dlouhý v dubnu 2010. Podrobnosti najdete ve zdrojovém souboru tohoto textu v komentáři u příkazu `\thebibliography`.

B.4.2 Odkazy na obrázky, tabulky a kapitoly

- Označení místa v textu, na které chcete později čtenáře práce odkázat, se provede příkazem `\label{navesti}`. Lze použít v prostředích `figure` a `table`, ale též za názvem kapitoly nebo podkapitoly.
- Na návěští se odkážeme příkazem `\ref{navesti}` nebo `\pageref{navesti}`.

B.5 Rovnice, centrováná, číselovaná matematika

Jednoduchý matematický výraz zapsaný přímo do textu se vysází pomocí prostředí `math`, resp. zkrácený zápis pomocí uzavření textu rovnice mezi znaky `$`.

Kód `$ S = \pi * r^2 $` bude vysázen takto: $S = \pi * r^2$.

Pokud chcete nečíselované rovnice, ale umístěné centrovane na samostatné řádky, pak lze použít prostředí `displaymath`, resp. zkrácený zápis pomocí uzavření textu rovnice mezi znaky `$$`. Zdrojový kód: `$$$ S = \pi * r^2 $$$` bude pak vysázen takto:

$$S = \pi * r^2$$

Chcete-li mít rovnice číselované, je třeba použít prostředí `equation`. Kód:

```
\begin{equation}
S = \pi * r^2
\end{equation}
```

```
\begin{equation}
V = \pi * r^3
\end{equation}
```

je potom vysázen takto:

$$S = \pi * r^2 \quad (\text{B.1})$$

$$V = \pi * r^3 \quad (\text{B.2})$$

B.6 Kódy programu

Chceme-li vysázet například část zdrojového kódu programu (bez formátování), hodí se prostředí *verbatim*:

```

      (* nickname2 *)
Lego> Refine in1
      (do_reg (nickname1 h));
Refine by in1 (do_reg (nickname1 h))
    ?4 : pcddata
    ?5 : pcddata
      (* surname2 *)
Lego> Refine surname1 h;
Refine by surname1 h
    ?5 : pcddata
      (* email2 *)
Lego> Refine undo_reg (email1 h);
Refine by undo_reg (email1 h)
*** QED ***
```

B.7 Další poznámky

B.7.1 České uvozovky

V souboru `k336_thesis_macros.tex` je příkaz `\uv{}` pro sázení českých uvozovek. „Text uzavřený do českých uvozovek.“

Příloha C

Seznam použitých zkratk

API Application Programming Interface - rozhraní pro programování aplikací

CEO Chief executive officer - ředitel

URL Uniform Resource Locator - jednoznačný identifikátor u webových stránek, například:
`http://www.example.com`

Příloha D

UML diagramy

Tato příloha není povinná a zřejmě se neobjeví v každé práci. Máte-li ale větší množství podobných diagramů popisujících systém, není nutné všechny umísťovat do hlavního textu, zvláště pokud by to snižovalo jeho čitelnost.

Příloha E

Instalační a uživatelská příručka

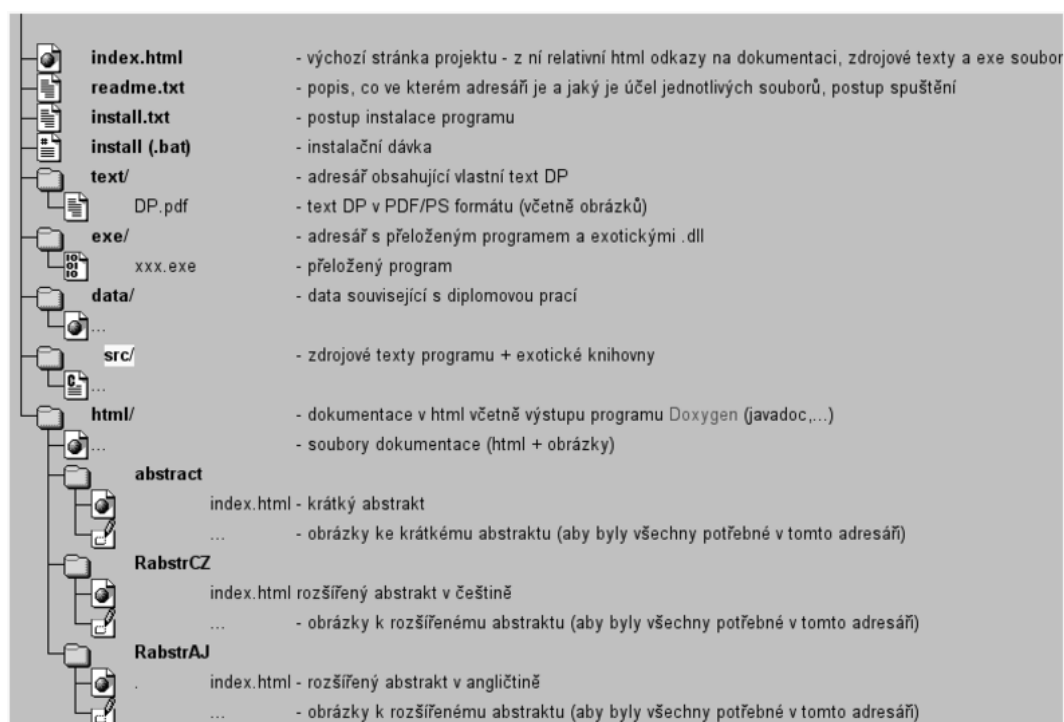
Tato příloha velmi žádoucí zejména u softwarových implementačních prací.

Příloha F

Obsah příloženého CD

Tato příloha je povinná pro každou práci. Každá práce musí totiž obsahovat příložené CD. Viz dále.

Může vypadat například takto. Váš seznam samozřejmě bude odpovídat typu vaší práce.:



Obrázek F.1: Seznam příloženého CD — příklad

Na GNU/Linuxu si strukturu příloženého CD můžete snadno vyrobit příkazem:

```
$ tree . >tree.txt
```

Ve vzniklém souboru pak stačí pouze doplnit komentáře.

Z **README.TXT** (případně **index.html** apod.) musí být rovněž zřejmé, jak programy instalovat, spouštět a jaké požadavky mají tyto programy na hardware.

Adresář **text** musí obsahovat soubor s vlastním textem práce v PDF nebo PS formátu, který bude později použit pro prezentaci diplomové práce na WWW.