

# STATISTIKY MOODLE

## MOODLE STATISTICS

Petr Grolmus

### Abstrakt

Článek shrnuje možnosti sledování on-line statistik a přehledů jak z provozu systému, tak cílené vyhledávání konkrétních údajů – např. využití systému Moodle v konkrétním akademickém roce, využití kurzu studenty v průběhu semestru, atp. Přehledy provozu a využití LMS mohou vést ke zlepšení efektivity on-line výuky. Lze snadno zjistit, které kurzy jsou málo frekventované, v kterých otázkách testů studenti nejčastěji chybují a do jisté míry lze i odhadovat, kolik času studenti věnují materiálům v LMS.

**Klíčová slova:** *LMS, Moodle, (on-line) statistiky, Google Analytics, Matomo, Piwik, Icinga, Excel.*

### Abstract

This article summarizes the possibilities of monitoring on-line statistics and overviews both from system operation and targeted searching of specific data – eg. usage of Moodle system in a particular academic year, usage of the course by students during the semester, etc. Traffic summaries and usage reports of LMS can improve the effectiveness of online learning. It is easy to find out which courses are not very frequent, in which test questions the students most often make mistakes, and to some extent it is also possible to estimate how much time students spend on materials in LMS.

**Key words:** *LMS, Moodle, (on-line) statistics, Google Analytics, Matomo, Piwik, Icinga, Excel.*

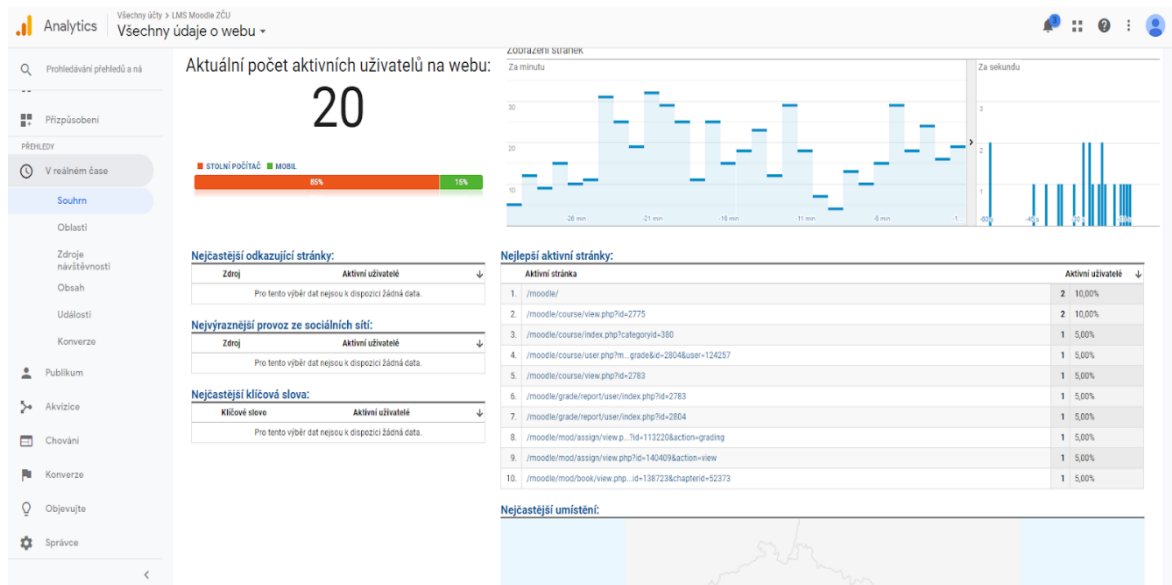
## 1 Google Analytics

Google Analytics (GA) zcela jistě nebyly prvním nástrojem pro sledování a generování statistik návštěvnosti webových stránek. Díky kapitálu společnosti se však velmi rychle dostaly do popředí a staly se standardem na poli produktů on-line statistik.

GA byly v srpnu 2006 zpřístupněny zdarma pro využití veřejností. V roce 2009 zavádí GA [1] tzv. *Tracking Code Snippets* - sledovatelné kódy pro přesnější určování návštěvnosti webových služeb, stránek, nebo dokonce jen jejich částí - např. konkrétních *tweetů*.

Princip použití sledovatelných kódů ve stránkách je v detailu popsán v kapitole věnující se softwaru Matomo (dříve Piwik) [2], který využívá stejného principu, a který jsme na ZČU nasadili do rutinního provozu. Použití sledovacích snippetů v GA je tedy totožné, jako u produktu Matomo.

Až s příchodem sledovatelných kódů, které jsou vlastně do webových stránek vložené krátké (nejčastěji javascriptové) programy, které aktivně a v reálném čase kontaktují server sbírající data, bylo možné generovat on-line statistiky a mít tak přehled, co se děje na sledovaných stránkách v daném, současném okamžiku.



Obrázek 1- Základní přehledová stránka Google Analytics. Zdroj: vlastní.

Uvedená přehledová stránka (viz obrázek 1) je základní, tj. bez dalších úprav, které jsou nicméně v GA samozřejmostí. Zobrazená statistika odpovídá reálnému provozu a zobrazuje aktuální počet aktivních uživatelů sledované služby, včetně právě zobrazených stránek. Pod číselným údajem o počtu uživatelů je také dostupná informace, jaké zařízení bylo použito pro zobrazení stránek LMS Moodle. Pod pojmem "stolní počítač" jsou započítány také zařízení typu notebook. V drtivé většině časových okamžiků stolní počítače tvoří 100 % aktuálně přistupujících zařízení, z čehož lze snadno vyvodit, že pro e-learningové systémy, a tedy i pro samotný proces učení, je častěji využívána větší zobrazovací plocha. Tato zařízení též mají uvedenou delší dobu sezení (angl. session) - doba od přihlášení uživatele po odhlášení nebo opuštění stránek.

Menší zobrazovací plochy - tablety a smartphony - mívají výrazně kratší dobu sezení a zřejmě tedy bývají používány spíše pro nalezení konkrétní informace ve studijních materiálech a memorování znalostí, než pro cílený proces učení.

## 2 Matomo / Piwik

Aplikace Piwik byla roku 2007 vytvořena jako *open source* alternativa ke službě Google Analytics. Aplikace je provozována v prostředí s podporou skriptovacího jazyka PHP a databázového systému, nejčastěji MySQL. V této konfiguraci je též provozován i na Západočeské univerzitě.

Matomo lze velmi snadno integrovat téměř do libovolné stránky. Oficiální stránky projektu prezentují snadné použití ve 100+ technologiích, včetně Content Management Systems (CMS), Online Shops, Forums, Frameworks, atd.

Princip použití Matomo je jednoduchý. Do sledované služby, například LMS Moodle, je zapotřebí na vhodné místo, ideálně do záhlaví každé stránky zobrazované uživateli v okně prohlížeče, vložit krátký kousek javascriptového kódu, který umožňuje předávat aplikaci Matomo informace pro zpracování, viz Obrázek 2. Tento javascriptový kód vygeneruje pro nově registrovanou službu provozovanou v prostředí

World Wide Web (WWW) sama aplikace Matomo, stačí jej jen vložit na vhodné místo sledované webové stránky.

```

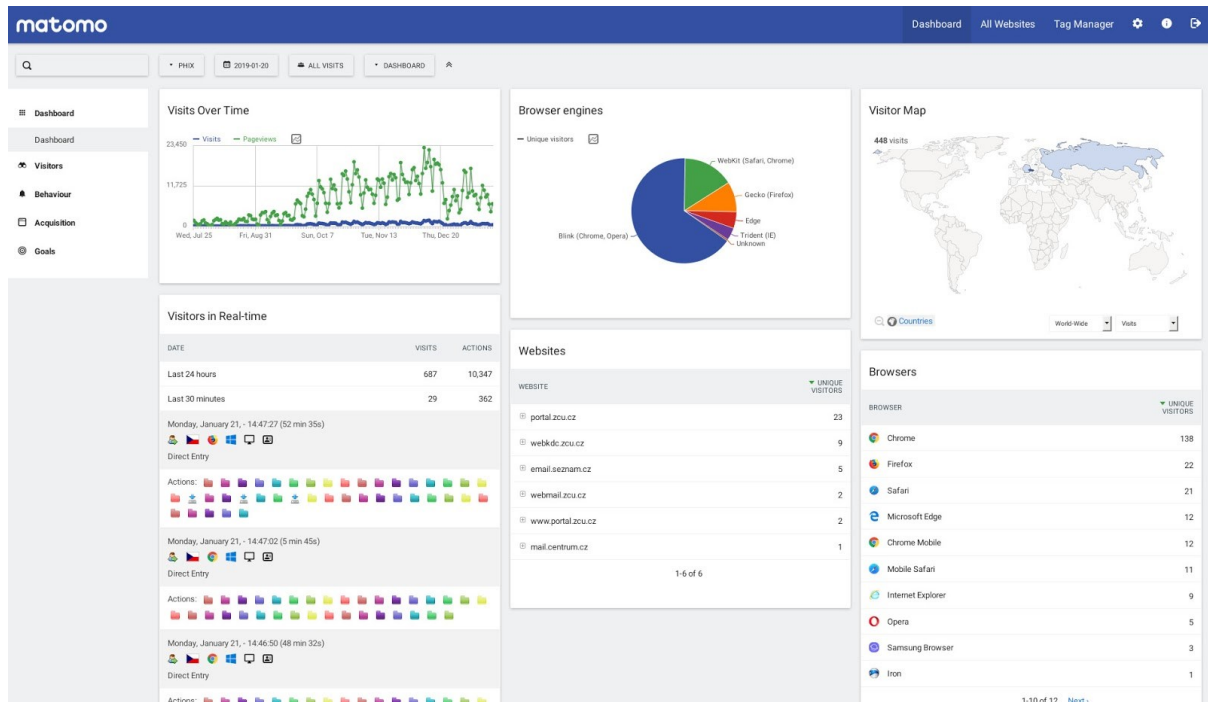
<!-- Piwik -->
<script type="text/javascript">
var _paq = _paq || [];
_paq.push(["setDomains", ["moodle.zcu.cz", "*.zcu.cz"]]);
_paq.push(['trackPageView']);
_paq.push(['enableLinkTracking']);
(function() {
var u="//piwik.zcu.cz/";
_paq.push(['setTrackerUrl', u+'piwik.php']);
_paq.push(['setSiteId', '5']);
var d=document, g=d.createElement('script'),
s=d.getElementsByTagName('script')[0];
g.type='text/javascript'; g.async=true; g.defer=true;
g.src=u+'piwik.js'; s.parentNode.insertBefore(g,s);
})();
</script>
<noscript><p></p></noscript>
<!-- End Piwik Code -->

```

**Obrázek 2 - Ukázka vkládaného sledovacího kódu Matomo do provozované služby. Zdroj: vlastní.**

Jakmile Matomo započne sběr dat nad konkrétní službou, je ihned možné začít sledovat grafické výstupy, které aplikace Matomo na vlastní provozované webové stránce poskytuje. Pro každého přihlášeného uživatele Matomo je nespornou výhodou možnost upravit si přehledovou stránku sledované služby - tzv. dashboard - dle svých představ. Přehledová stránka nabízí velkou škálu sledovaných kritérií, které je možné dle uvážení ve formě *portletů* přidávat a ubírat. Často si lze zvolit, zda údaje budou zobrazovány v přehledné grafické formě nebo v textové formě v podobě výpisu datových údajů. Příklad vizuálního vzhledu přehledové stránky lze nalézt na obrázku č. 3.

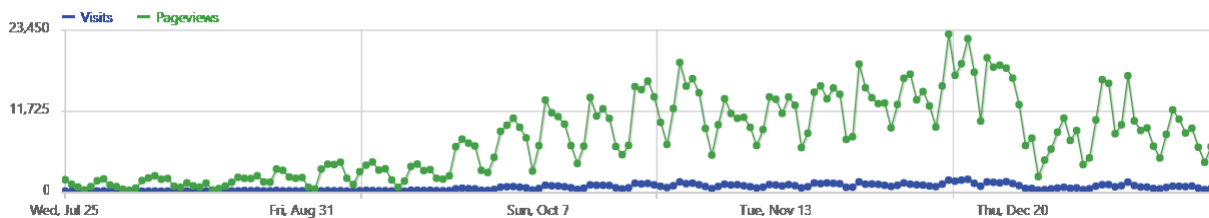
Dle výše uvedeného popisu je patrné, že použití Matomo je zcela nezávislé na konkrétní verzi nasazeného LMS Moodle a lze je s úspěchem využít na všech běžně provozovaných verzích.



Obrázek 3- Příklad přehledové stránky Matomo. Zdroj: vlastní.

Možnosti sledovaných charakteristik na provozované službě jsou opravdu velmi široké a je na každém, jaké přehledy jsou pro něj zajímavé, a jakým způsobem si je na přehledové stránce poskládá. Většinu "portletů" (dlaždic) v dashboardu je možné si dále rozkliknout, takže pak lze následně vidět více detailů nebo změnit nastavení, např. změnit rozsah zobrazovaného období v *portletu*.

Příkladem může být *portlet* "Visits Over Time", který se po rozkliknutí stane přehlednějším a je na něm vidět více vypovídajících podrobností - viz obrázek č. 4.



Obrázek 4- Detailní návštěvnost Moodle. Zdroj: vlastní.

V tomto detailu je snadno vidět, jak se mění celková návštěvnost v Moodle v průběhu času. V době letních prázdnin, období července v levé části je provoz minimální. Postupně se provoz od začátku zimního semestru zvyšuje, jak učitelé seznamují své studenty s kurzy a nabádají je k jejich používání. Vše kulminuje v závěru semestru v období před vánočními svátky, kdy se studenti z elektronických kurzů připravují na testy, případně je přímo v LMS realizují.

Matomo umožní i v tomto náhledu získat doplňující informace, pokud kurzorem najedeme na jednotlivé body tvořící spojnicí grafu. Kromě konkrétního datumu přidá i další sledované charakteristiky (jejich výběr lze opět snadno změnit). V tomto případě graf kulminuje v pondělí 10. prosince 2018, kdy Moodle navštívilo 1667 jedinečných uživatelů, kteří celkem zobrazili 22 763 stránek.

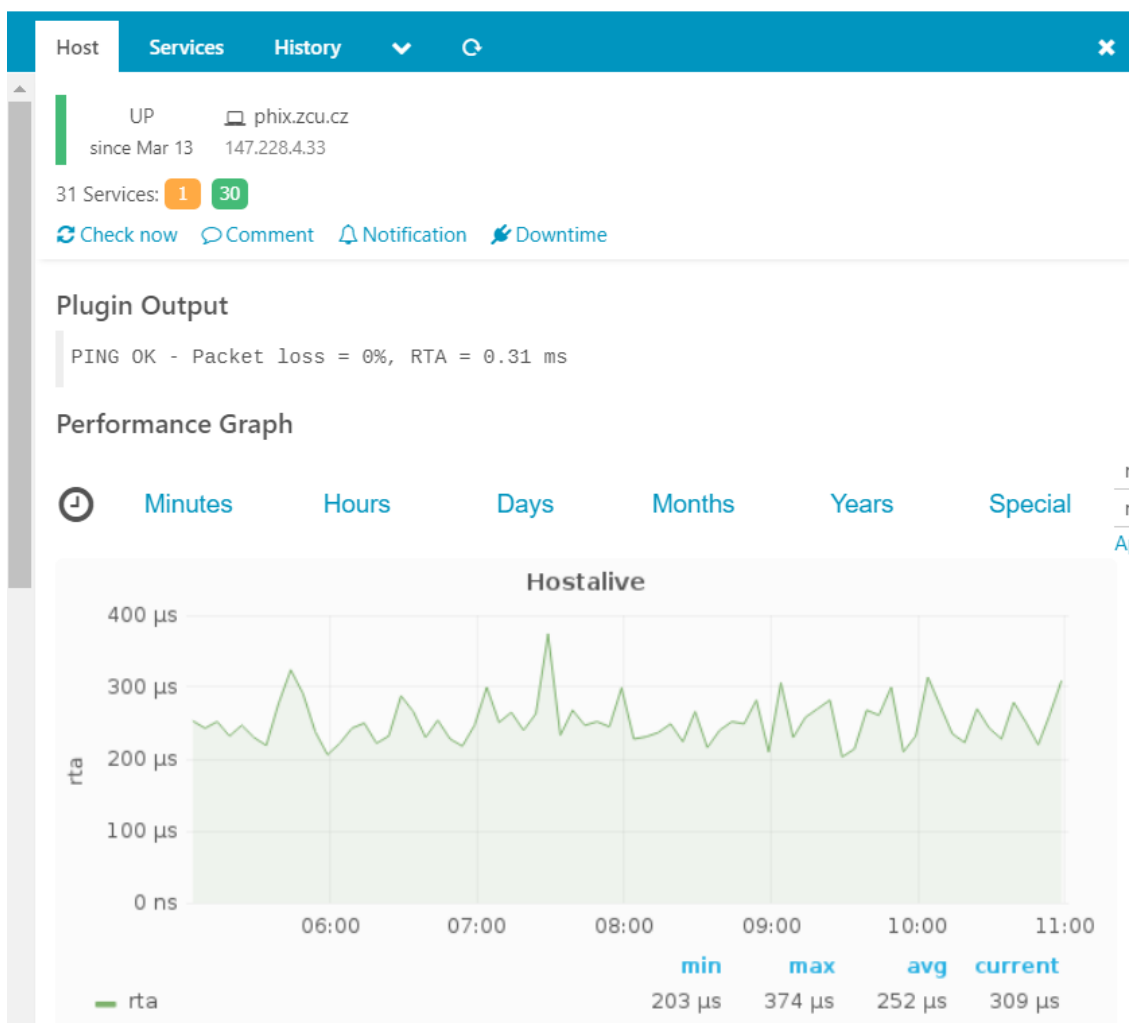
Na první pohled v uvedeném grafu zřejmě upoutá zřetelná týdenní perioda v používání Moodle. Chování uživatelů v různých týdnech je téměř shodné, liší se jen celkovými hodnotami, které se mění v průběhu akademického roku. Lze snadno nahlédnout, že lokální minimum využití je téměř vždy v sobotu a maximum v pondělí. Výjimku v roce 2018 tvořily jen lokální minima v pondělí 24. prosince a 31. prosince.

Nespornou výhodou produktu Matomo je skutečnost, že nasbírané údaje o návštěvnosti vašich stránek jsou tyto ukládány na vlastních serverech a zdrojích.

### 3 Icinga

Pro generování nebo lépe sledování on-line statistik je možné použít též i software Icinga [3], který obecně slouží k monitorování strojů a služeb ve výpočetním prostředí. Není tedy specializován přímo na LMS Moodle, nicméně o chování tohoto systému může mnoho prozradit.

Softwarový produkt Icinga na Západočeské univerzitě používáme cca 2 roky, kdy nahradil předchozí monitorovací systém Nagios, který byl provozovaný déle než 15 let.



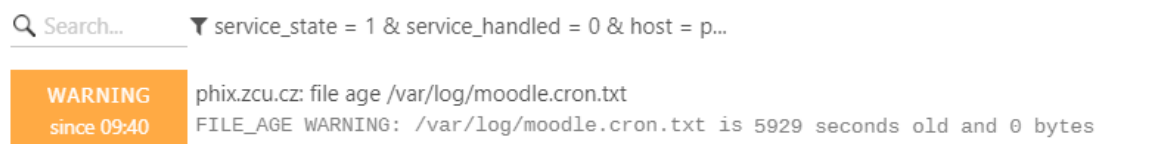
Obrázek 5- Základní přehled stroje s LMS Moodle v Icinge. Zdroj: vlastní.

Monitorovací systémy obecně obsahují velkou množinu sledovacích prvků, které je možné snadno rozšířit o svou vlastní kontrolu konkrétní běžící služby nebo

projevu správné funkčnosti této služby. Tyto monitorovací prvky periodicky kontrolují odezvu služeb běžících na serverech nebo dostupnost těchto serverů jako celek. V případě, že dojde k odchýlení od normálního stavu, jsou správci daného serveru nebo konkrétní služby o této události informováni e-mailovou zprávou.

Díky tomu, že předchozí stavy každé monitorované služby jsou dlouhodobě uchovávány, umí Icinga snadno a rychle pro konkrétní službu vygenerovat přehledný graf hodnot. Základní přehled stroje `phix.zcu.cz`, na kterém je provozován LMS Moodle v Icinze je k nahlédnutí na obrázku č. 5.

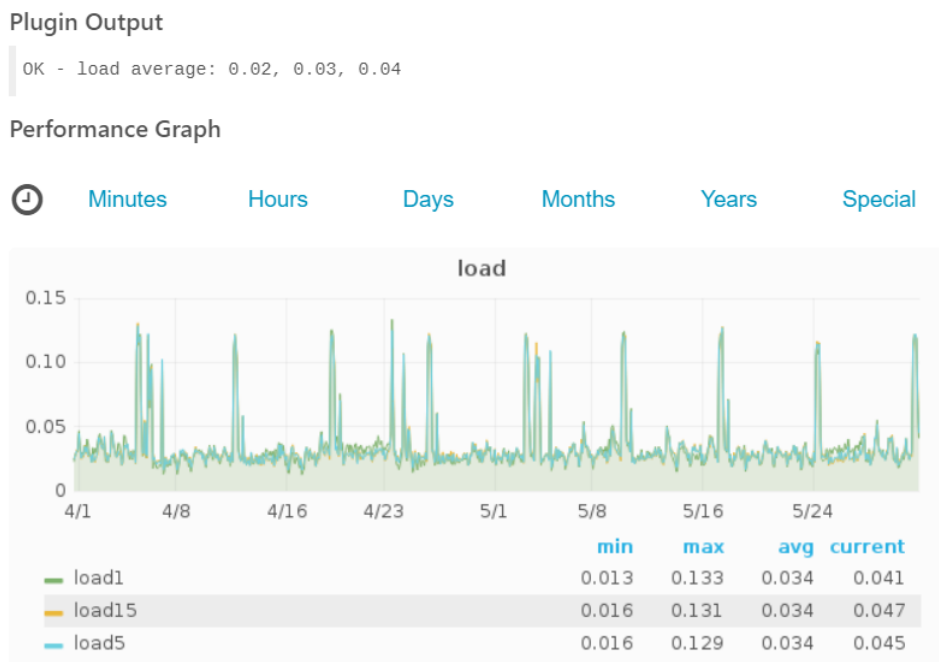
Z náhledu je patrné, že na serveru `phix.zcu.cz` je monitorováno průběžně 31 parametrů - např. odezva serveru, obsazenost jednotlivých disků, platnost https certifikátu, teplota na procesorech, vytížení paměti, atp. Z uvedených 31 parametrů je 30 z nich v povolených (zelených) hodnotách a jeden parametr je zvýrazněn žluto-oranžovou barvou - tzv. warning - varování. Po rozkliknutí lze zjistit, která sledovaná služba má potíže – viz obrázek č. 6.



Obrázek 6- Detail problematické služby v monitoringu. Zdroj: vlastní.

Problematickým je soubor `/var/log/moodle.cron.txt`, kam zapisuje pravidelně spouštěný skript tzv. “*Moodlecron*”, standardní součást LMS Moodle, která periodicky vykonává různé akce - např. rozesílá všem zúčastněným uživatelům zprávy z diskusního fóra kurzu nebo otevírá přednastavené testy podle předem definovaného času. *Moodlecron* bývá pouštěn periodicky, např. každých 5 minut, a do logovacího souboru `moodle.cron.txt` vždy zapíše minimálně čas spuštění. Delší časový úsek, kdy tento soubor nebyl změněn, je tedy podezřelý a Icinga správně rozpoznala potenciální problém. V tomto případě ale *Moodlecron* spustil zálohování kurzů, které může trvat i několik hodin, což je plně závislé na celkovém počtu kurzů provozovaných v LMS a aktuálním vytížení serveru. Po dokončení této časově náročné operace se spuštění *Moodlecronu* opět stane pravidelným a i tento warning bude změněn na zelený “OK” stav.

Sledování údajů nasbíraných Icingou může včas odhalit problémy, které mohly vzniknout jak v provozu, tak i třeba již při rozhodování při pořizování vhodného hardwaru na provoz LMS. V tomto případě je vhodné občas zkontrolovat vytěžování serveru - tzv. load. Pohledem na graf vytíženosti procesorů, např. za poslední dva měsíce provozu, lze konstatovat, že hardware je více než dostačující, když vytížení serveru i při špičkách při zálohování systému za provozu dosahuje maximální vytížení cca hodnoty 0.15 (= opravdu nepodstatné vytížení) - viz obrázek č. 7.



Obrázek 7- Vytížení hardwaru LMS Moodle. Zdroj: vlastní.

## 4 Excel

Pomocí Excelu lze vizualizovat různými způsoby obrovské množství dat. Data pro zpracování je však nutné získat pracnějším způsobem. Základní údaje k jednotlivým kurzům lze do určité míry získat přímo z Moodle ve vhodném formátu - např. CSV4.

Pokud chceme zobrazit nějaké složitější náhledy a statistiky, neobejdeme se bez znalosti databázového jazyka SQL<sup>5</sup> a také znalosti přístupového jména a hesla do databázového systému, který je pro provoz Moodle používán. Za tento pracnější způsob získávání dat však budeme odměněni tím, že dokážeme vytvořit statistiky na míru dle požadavku.

Na Západočeské univerzitě v Plzni používáme jako databázový backend pro LMS Moodle databázový systém MySQL, resp. nově její open source odnož pod názvem MariaDB. Díky shodnému API<sup>6</sup> rozhraní MySQL a MariaDB lze externí nástroje totožně použít pro obě varianty.

Pro dolování dat Moodle z databázového systému lze použít různé grafické a webové nástroje - MySQL Workbench, PHPMyAdmin, Adminer a jistě i spoustu dalších. V nouzi však postačí nejjednodušší nástroj, který je součástí instalačního balíku databázového systému MySQL (i MariaDB) a tím je CLI<sup>7</sup>. I v tomto základním prostředí, lze při běžné znalosti jazyka SQL a přehledu tabulek Moodle vystavět dotaz, kterým se dostaneme k požadovaným datům.

Příklad - zajímá nás, kdy je nejhodnější provádět každodenní zálohování Moodle dat uložených v MySQL databázi.

<sup>4</sup> CSV - Comma Separated Value - hodnoty oddělené čárkou, běžný formát pro export dat pro statistické zpracování.

<sup>5</sup> SQL - Structured Query Language - standardizovaný databázový dotazovací jazyk.

<sup>6</sup> API - Application Programming Interface - definované rozhraní pro programování aplikací, pojem ze softwarového inženýrství.

<sup>7</sup> CLI - Command Line Interface - prostředí příkazové řádky.

Za dlouhou dobu provozu LMS Moodle jeho data uložená v databázi poměrně nabobtnala na téměř 9 GB. Zálohování takto velkého množství dat je více než vhodné dělat v průběhu minimálního provozu, aby nedocházelo k prodlužování reakční doby z důvodu používání zámků v databázi při zálohování jednotlivých tabulek. Zámek - LOCK TABLE <tablename> - je běžná praxe v databázových systémech kdy proces modifikující záznam v tabulce si vynutí exkluzivní přístup k záznamu nebo celé tabulce, aby nemohlo dojít k vícenásobné změně stejného záznamu. V tomto případě tabulku uzamyká zálohovací proces, aby ukládaná data byla konzistentní a nemohlo tak dojít ke změně jiným procesem.

Pro tento účel můžeme vytvořit přehledný graf počtu prováděných akcí pro jednotlivé dny v týdnu a pro konkrétní hodinu.

Pod pojmem “akce” Moodle identifikuje např. přihlášení uživatele, jeho odhlášení, návštěvu/zobrazení konkrétní stránky nebo části kurzu, editaci kurzů, modifikaci uživatelských rolí, oznámkování studenta, odevzdání úkolu, zodpovězení testové otázky, atp.

Všechny tyto akce jsou přehledně uloženy v jediné tabulce s pojmenováním “mdl\_logstore\_standard\_log”. Přestože na Západočeské univerzitě provozujeme systém Moodle v centrální správě již déle než 10 let, a při pravidelných upgradech jsou všechna data zachovávána, tak v této tabulce jsou data jen cca za posledních 5 let. V dřívějších verzích Moodle používal jiné způsoby a jiné tabulky pro logování akcí uživatelů.

Prvním nutným krokem je získání relevantních dat z databázového systému za dobu posledních 12 měsíců. SQL příkaz, který tato data získá, je zde:

```
select dayname(from_unixtime(timecreated)) as den,
hour(from_unixtime(timecreated)) as hodina, count(*) as pocet
from mdl_logstore_standard_log where timecreated >
(unix_timestamp(now()) - (365 * 86400)) group by den, hodina
INTO OUTFILE '/tmp/actions-per-day-and-hour.csv' FIELDS
TERMINATED BY ',' LINES TERMINATED BY '\n';
```

V atributu (položce) timecreated tabulky mdl\_logstore\_standard\_log je uložen čas vzniku akce v tzv. UNIXovém formátu<sup>8</sup>. Jedná se o velké celé číslo, které udává počet vteřin od půlnoci ze středy 31. prosince 1969 na čtvrtek 1. ledna 1970 - tzv. UNIX Epoch Time [4]. Z této položky lze získat vhodnými SQL funkcemi konkrétní den, hodinu a počet akcí v kombinaci těchto dvojic.

V závěru příkazu je podmínka omezující vyhledávání akcí jen na posledních 365 dní zpětně od aktuálního data. Na samém konci příkazu je definování výstupu nalezených výsledků do souboru ve formátu CSV. Tento formát je vhodný přímo pro načtení do Excelu a následné zpracování. Výsledkem je právě 168 záznamů (týden má 168 hodin).

Prvních několik řádek z výsledného souboru actions-per-day-and-hour.csv:

```
Friday,0,7911
Friday,1,4434
```

<sup>8</sup> UNIX - jméno operačního systému vyvinutého AT&T v roce 1969 - <https://cs.wikipedia.org/wiki/Unix>.



Friday, 2, 2562  
 Friday, 3, 2378  
 Friday, 4, 2554

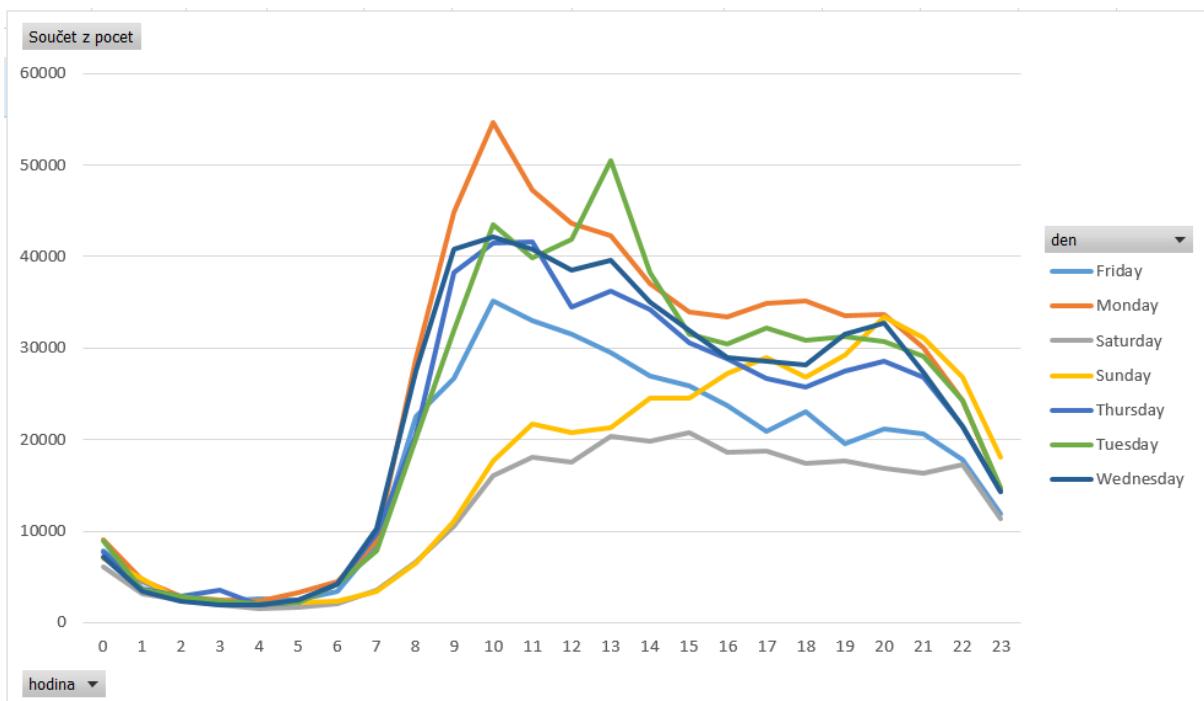
Z prvního řádku vyplývá, že za poslední rok každý pátek (kumulativně) od času 00:00:00 do času 00:59:59 Moodle zpracoval 7911 akcí, z druhého řádku každý pátek od času 01:00:00 do času 01:59:59 Moodle zpracoval 4434 akcí, atd.

Je vhodné před tato data přidat řádek se jmény jednotlivých sloupců. Pojmenované sloupce se budou prospěšné při dalším zpracování v Excelu. Přidáme tedy na úplný začátek souboru řádek:

den, hodina, počet

Tato data následně načteme do Excelu a vytvoříme kontingenční tabulku, kde do sloupců kontingenční tabulky přetáhneme položku `den`, do řádků pak položku `hodina` a do hodnot umístíme položku `počet`. Nyní již stačí jen data zobrazit ve vhodném kontingenčním grafu, viz obrázek 8.

Z uvedeného grafu jasně vyplývá odpověď na původní otázku “kdy je vhodné provádět zálohování databáze Moodle?”. Minimální provoz v systému Moodle je mezi 1. až 6. hodinou ranní. Zálohování dat Moodle z databázového systému MySQL na Západočeské univerzitě v Plzni provádíme každý den v čase 1:00. Zálohování samotné trvá necelých 5 minut, kdy je Moodle funkční, ale může mít pro uživatele o něco delší odezvy než je obvyklé. Tyto potenciální potíže ale ovlivní v daném čase jen minimum uživatelů.



Obrázek 8 - Rozložení počtu akcí v LMS Moodle za poslední rok. Zdroj: vlastní.

## 5 Závěr

Pro účely zpracování statistik LMS Moodle a jejich následné převedení do graficky vhodné podoby je možné využít širokou paletu nástrojů, jak spojených přímo s LMS

Moodle, coby součást hlavního balíku Moodle nebo externích modulů ve formě rozšíření základní funkcionality Moodle, tak i externích nástrojů nebo celých softwarových balíčků pro zpracování surových dat exportovaných nebo získaných přímo z relačního databázového systému použitého pro provoz LMS.

Věřím, že se podařilo zmapovat základní použitelné nástroje pro vytvoření graficky vhodných reprezentací provozu a dat LMS Moodle.

### **Použité zdroje**

1. The History of Google Analytics - Infographic. Digital Agency - Higher Education Marketing | Converge Consulting. [Online] <https://convergeconsulting.org/2016/09/21/infographic-the-history-of-google-analytics/>.
2. Secure Open Web Analytics Platform [online]. Copyright © 2019 matomo.org [cit. 07.08.2019]. Dostupné z: <https://matomo.org/>. [Online] <https://matomo.org/>
3. Icinga Monitoring Software. [Online] <https://icinga.com/>.
4. Unix time - Wikipedia. [Online] [https://en.wikipedia.org/wiki/Unix\\_time](https://en.wikipedia.org/wiki/Unix_time).

### **Kontaktní údaje**

Ing. Petr Grolmus

Západočeská univerzita v Plzni, Centrum informatizace a výpočetní techniky

Univerzitní 20, 306 14 Plzeň

e-mail: [indy@civ.zcu.cz](mailto:indy@civ.zcu.cz)