

Uniwersytet Śląski
Wydział Matematyki, Fizyki i Chemii

Szymon Bluma

Praca licencjacka

Porównanie wydajności popularnych skryptów forów internetowych

Promotor:
dr Tomasz Powierża

Katowice, 2008

Autor: Szymon Bluma, www.BlueMan.pl

Zabrania się kopiowania, rozpowszechniania, udostępniania bez podania autora pracy

1. Wstęp teoretyczny	3
1.1. Oprogramowanie	3
1.1.1. Debian	4
1.1.2. Apache	5
1.1.3. PHP	7
1.1.4. MySQL	8
1.2. Infrastruktura aplikacji internetowych	9
1.2.1. wielordzeniowce	10
1.2.2. pamięci ECC	11
1.2.3. dyski HDD SAS i SATA	12
1.3. Sposoby podziału mocy	13
1.3.1. klasteryzacja	13
1.3.2. load balancing	13
1.3.3. decentralizacja usług (apache, mysql, pliki)	14
2. Opis/historia skryptów	15
2.1. phpBB	15
2.2. phpBB2 by Przemo	17
2.3. SMF	17
2.4. MyBB	18
3. Opis teoretyczny testów	19
3.1. aplikacja WAPT	19
3.2. aplikacja munin	21
3.3. sprzęt użyty w testach	21
3.4. opis ścieżki użytkownika	22
3.5. porównanie skryptów	25
4. Część praktyczna – testy	28
4.1. Munin	28
4.1.1. porównanie RAM	28
4.1.2. porównanie SQL	31
4.1.3. porównanie CPU	33
4.1.4. porównanie LOAD	35
5. Wnioski	37
6. Bibliografia	39

1. Wstęp teoretyczny

Niniejsza praca poświęcona jest mierzeniu wydajności najpopularniejszych darmowych skryptów forów internetowych. Temat odpowiedniej optymalizacji tych skryptów jest bardzo często poruszany w sieci Internet, ale niestety nie ma opracowania dotyczącego porównania skryptów względem siebie. Zagadnienia i problemy z tym związane są mi bardzo dobrze znane ze względu na prowadzenie kilku forów dyskusyjnych^{[13][14]} zainstalowanych na różnych skryptach.

Omawiając termin wydajności skryptów (nie tylko forów dyskusyjnych) warto wspomnieć o tym, jakie podejście do limitów wykorzystanych zasobów mają firmy hostingowe, w których wykupujemy usługi na hostowanie strony www. Nigdzie w umowach zawieranych z jakąkolwiek firmą nie ma jasno podanych limitów procesora, pamięci, itd. jakie możemy zużyć. Limity te są w pewien sposób „wirtualnie” przydzielane i wynoszą zwykle 5% mocy całego serwera. W skład tych 5% wchodzi głównie wykorzystanie procesora CPU. Po przekroczeniu tego progu i dłuższym utrzymywaniu się takiego obciążenia najprawdopodobniej klient dostanie list z informacją o tym, że jego strona przeciąża serwer i wymagana jest optymalizacja strony, lub rozwiązanie umowy. Firmy nie mogą sobie pozwolić, aby jeden użytkownik wykorzystywał większość mocy serwera. Warto więc zastanowić się nad prawidłowym wyborem skryptu dla forum – aby niewłaściwy wybór nie generował problemów w przyszłości.

Omawiając temat wydajności aplikacji internetowych należy najpierw zapoznać się z rzeczami, jakie są niezbędnymi składnikami do prawidłowego działania strony oraz które mają znaczący wpływ na szybkość działania aplikacji.

1.1. Oprogramowanie

Strona www, aby została wyświetlona na ekranie odbiorcy musi być obsługiwana przez odpowiedni system operacyjny, a także szereg innych podstawowych aplikacji, które zostaną omówione w tym podrozdziale. Prawidłowa instalacja, jak również odpowiednia konfiguracja tych usług ma znaczący wpływ na otrzymanie końcowego rezultatu w postaci wydajnego serwera obsługującego dynamiczne strony internetowe. Do poszczególnych składników istnieje wiele innych (oficjalnie niewspieranych przez wydawcę) dodatków, które także potrafią przyspieszyć działanie stron www, np.: eAccelerator, XCache, itp.

Autor: Szymon Bluma, www.BlueMan.pl

Zabrania się kopiowania, rozpowszechniania, udostępniania bez podania autora pracy

1.1.1. Debian

System operacyjny jest głównym składnikiem każdego serwera. Dzięki jego możliwościom i funkcjonalności, jakie dostarcza jest możliwa instalacja i konfiguracja dodatkowego oprogramowania potrzebnego do kompletnego wydajnego systemu serwerowego. W testach korzystano z systemu Debian. Poniżej została zamieszczona krótka charakterystyka tego systemu.

Projekt systemu operacyjnego Debian^[1] jest otwartym ogólnosięciowym projektem, w którym uczestniczą wolontariusze - wspólnie dążąc do utworzenia i udoskonalenia systemu operacyjnego, w którym zamieszczone jest tylko i wyłącznie legalne, wolne, otwarte oprogramowanie.



Rys. 1 – Logo systemu Debian

Projekt^[2] Debian GNU/Linux został oficjalnie założony 13 sierpnia 1993 roku (pierwsza wersja 0.01) przez Iana Murdoca. Nazwa Debian pochodzi od imienia jego dziewczyny (w tamtych czasach), a obecnie byłej żony: Debra, oraz jego własnego imienia: Ian. Najważniejszą cechą wyróżniającą system Debian spośród innych systemów z rodziny linux jest łatwość instalacji dodatkowego oprogramowania, dzięki potężnemu zbiorowi gotowych pakietów. Pakiety są to pliki z rozszerzeniem *.deb, które posiadają program instalacyjny, oraz spis wszystkich zależności (bibliotek), które są wymagane do prawidłowej, pełnej instalacji oprogramowania oraz jego uruchamiania. W razie potrzeby instalator automatycznie doinstaluje brakujące biblioteki w systemie. Wersja 1.1 o kodowej nazwie Buzz opublikowana 17 czerwca 1996 roku była pierwszą wersją w historii dystrybucji Debiana, która rozpoznawana była przez jej kodową nazwę. Nazewnictwo wszystkich wersji Debiana pochodzi z animowanej bajki Toy Story. W czasie pierwszego publicznego wydania systemu, pieczę nad projektem przejął od Ian'a, Bruce Perens, który równorzędnie pracował dla

Autor: Szymon Bluma, www.BlueMan.pl

wytwórni filmowej Pixar – wydawcy w/w bajki. Debian w tej wersji posiadał już pełne wsparcie dla ELF (format plików wykonywalnych, linkowanych, bibliotek zawierający skompilowany kod). Buzz posiadał dołączone 474 pakietów oprogramowania.

Kolejne wersje dystrybucji (1.2-rex, 1.3-bo, 2.0-hamm, 2.1-slink, 2.2-potato, 3.0-wood, 3.1-serge, 4.0-etch, 5.0-lenny), ukazywały się przeciętnie raz na rok, łatając niebezpieczne luki w systemie, dodając nowe architektury na których Debian potrafił działać oraz wprowadzając nowe możliwości były rozwijane przez coraz większą grupę entuzjastów tego systemu.

Główne cechy systemu Debian/GNU Linux:

- podstawowy priorytet to bezpieczeństwo systemu
- wykorzystanie tylko legalnego oprogramowania
- 10 wspieranych architektur: alpha, amd64, arm, hppa, i386, ia64, mips, powerpc, sparc, s390
- łatwość instalacji dodatkowego oprogramowania¹. Wstęp teoretyczny [serwer]

1.1.2. Apache HTTP Server

Podstawową aplikacją, która będzie przyjmować i obsługiwać ruch internetowy po stronie serwera jest Apache^[3].

Nazwa Apache pochodzi od indiańskiego plemienia, jego logo to kolorowe piórko (widoczne na Rys. 2).



Rys. 2 – Logo serwera Apache

Apache HTTP Server - często nazywany po prostu *Apache* jest obecnie najpopularniejszym serwerem spotykanym w internecie. Serwer Apache był pierwszą alternatywą dla serwera Netscape Communication Corporations (obecnie znanego jako Sun Java System Web Server). Aplikacja Apache została stworzona przez Roberta McCool, który pracował na zlecenie National Center for Supercomputing Applications. Kiedy w połowie 1994 roku odszedł z tej

Autor: Szymon Bluma, www.BlueMan.pl

Zabrania się kopiowania, rozpowszechniania, udostępniania bez podania autora pracy

firmy, projekt serwera został porzucony z powodów licznych wad i luk bezpieczeństwa jakie w nim występowały. W styczniu 1995 niewielka grupa osób, które zgłaszały nieprawidłowości w pracy serwera postanowiła kontynuować zaczęłą pracę nad serwerem i wspólnymi siłami uzupełnić luki bezpieczeństwa w aplikacji i propozycje, jakie były zgłaszane i udostępnianie przez inne osoby w globalnej sieci internet. W ten sposób Brian Behlendorf, Roy T. Fielding, Rob Hartill, David Robinson, Cliff Skolnick, Randy Terbush, Robert S. Thau, Andrew Wilson wraz z dodatkową pomocą Eric Hagberg, Frank Peters oraz Nicolas Pioch stworzyli Fundację Apache Group. Korzystając z wersji serwera 1.3 NCSA dodali wszystkie łaty bezpieczeństwa, oraz proponowane sugestie funkcjonalności serwera zgłaszane przez użytkowników. Pierwsza publiczna oficjalna publikacja nowej wersji oprogramowania miała miejsce w kwietniu 1995 roku i była oznaczona numerem 0.6.2. Podczas dalszych prac nad projektem, które odbywały się od maja do czerwca i wysłuchaniu setek sugestii wciąż rosnącej grupy użytkowników nowego oprogramowania światło dzienne ujrzała kolejna wersja serwera z numerem 0.7.x. Twórcy zdawali sobie sprawę, że kod źródłowy programowania jest napisany bardzo chaotycznie i wymaga ponownego uporządkowania i opracowania nowych struktur i technologii, które będą odpowiednie dla celu istnienia wydajnego oprogramowania serwerowego. Dlatego też Robert Thau opracował nową architekturę, która została wdrożona w październiku w wersji 0.8.8.

Po kolejnych miesiącach testów, dodaniu nowych funkcjonalności, napisaniu dokumentacji, światło dzienne ujrzała wersja 1.0 – było to 1 grudnia 1995 roku. Wersja 2.0, która również polegała na kompletnym przepisaniu od nowa kodu aplikacji 1.0, oraz wносиła całkiem nową logikę struktury aplikacji (m.in. modułowość, obsługę wielordzeniowych maszyn, lepszą integrację z systemami z grupy poza UNIX'owymi, wsparcie dla IPv6 oraz nowe API) została finalnie wydana w marcu 2000 roku, jako zamknięta dla wybranych osób wersja alfa. 6 kwietnia 2002 została udostępniona dla całego internetu.

Główne cechy serwera Apache:

- kod źródłowy dostępny w formie OpenSource (napisany w całości w C)
- modułowość, wsparcie dla dodatkowych modułów, które poszerzają funkcjonalność serwera
- kompresja odebranych/wysłanych danych „w locie”
- udostępnianie kilku aplikacji internetowych na jednej uruchomionej instancji Apache

1.1.3. PHP

Kolejnym składnikiem serwera internetowego jest interpreter PHP^[4], który odpowiada za przetwarzanie plików PHP aplikacji serwisu. Dzięki niemu mamy możliwość tworzenia, dynamicznych stron internetowych.



Rys. 3 – Logo projektu PHP

Nazwa ta jest rekursywnym akronimem słowa *PHP Hypertext Preprocessor*.

PHP jest skryptowym językiem programowania stworzony do generowania dynamicznych stron www. Jest językiem, który działa po stronie serwera, ale również może zostać wywoływany z linii poleceń systemowych.

Pierwotnie PHP został stworzony w 1995 roku przez Rosmusa Lerdorfa, a obecnie jest rozwijany przez PHP Group. Lerdorf zaczął tworzyć swój projekt „Personal Home Page Tools”, ponieważ szukał alternatywy do języka Perl, którego używał na swojej domowej stronie internetowej. Potrzebował nowych funkcjonalności do wyświetlania swojego curriculum vitae, oraz analizowania ile ruchu (ściągnięć, itp.) generuje jego strona oraz samo CV. Tworząc PHP zaimplementował dużą ilość bibliotek napisanych w języku C, dzięki którym była możliwość komunikacji z bazami danych dostępnymi na serwerze. Rosmus zdecydował się 8 czerwca 1995 udostępnić swoje dzieło, aby przyspieszyć znajdowanie luk bezpieczeństwa i poprawę kodu, wykorzystując do tego pomoc innych internautów. Wydanie to zostało nazwane wersją PHP2 i posiada swoją podstawową funkcjonalność po dzień dzisiejszy. Składnia nowego języka podobna jest do Perla – posiada zmienne, możliwość umieszczania kodu HTML. Mimo podobieństwa do języka Perl, PHP jest bardziej ograniczony i prostszy.

Andi Gutmans oraz Zeev Suraski w 1997 dokonali sporo zmian w PHP na potrzeby projektu, nad którym pracowali u siebie na uniwersytecie. Przepisali od nowa PHP2 tworząc w ten sposób wersję PHP3. Wprowadzili między innymi obsługę wielu baz danych, protokołów oraz własne API. PHP3 umożliwiło także wszystkim programistom dopisywanie kolejnych modułów. Nowa wersja PHP została także nazwana inną nazwą: PHP Hypertext Preprocessor, która ujrzała światło dzienne w czerwcu 1998 roku.

Autor: Szymon Bluma, www.BlueMan.pl

Tuż po tym Gutmans i Suraski rozpoczęli pracę (w 1999 r.) nad przepisaniem swojego dzieła do projektu nazwanego Zend Engine. Projekt miał być głównym silnikiem dla nowej wersji PHP. 22 maja 2000 roku powstało PHP4 oparte na Zend Engine 1.0, które zostało rozszerzone względem PHP3 o następujące właściwości: lepsza wydajność/optymalizacja, wsparcie dla wielu aplikacji serwerowych, sesje HTTP, bezpieczne przechowywanie danych przesyłanych przez użytkowników (przez formularze na stronie www) oraz wiele innych właściwości.

Prace nad kolejnymi wersjami PHP4 trwają nadal. W międzyczasie (13 lipiec 2004) został opublikowany następca – PHP5, który został oparty na silniku Zend Engine II. Najważniejszą zmianą jest przejście na obiektowe programowanie. Podczas ciągłych prac nad aktualną stabilną wersją PHP5 trwają również równorzędnie prace nad PHP6, który obecnie udostępniony jest w wersji testowej i nie jest zalecany do instalacji na produkcyjnych maszynach.

1.1.4. MySQL

Ostatnią niezbędną częścią serwera jest baza danych, w której składowane są wszystkie dane tekstowe napisane przez użytkowników na forum. Dzięki idei/strukturze baz danych, znacznie prostsze i wydajniejsze są operacje pobierania/zmiany/zapisywania danych, niż ma to w miejsce w podobnych operacjach na plikach.

MySQL^[5] jest wielowątkowym, wolnodostępowym (Open Source), wieloużytkownikowym systemem bazy danych (ang. DBMS – DataBase Management System), który działa na wszystkich dostępnych platformach systemów operacyjnych.



Rys. 4 – Logo serwera bazy danych

SQL to strukturalny język zapytań (ang. Structured Query Language). Jest to najbardziej rozpowszechniony standaryzowany język dostępu do systemów zarządzania relacyjnymi bazami danych (RDBMS). SQL służy do zapisywania i pozyskiwania danych do, i z bazy. Język ten jest rozpoznawany przez niemal wszystkie systemy baz danych, m.in. MySQL, Oracle, PostgreSQL, Sybase, Microsoft SQL Server. W 1986 roku SQL stał się oficjalnym

Autor: Szymon Bluma, www.BlueMan.pl

Zabrania się kopiowania, rozpowszechniania, udostępniania bez podania autora pracy

standardem, wprowadzonym przez Międzynarodową Organizację Normalizacyjną (ang. ISO – International Standard Organization) oraz członka tej organizacji - Amerykański Narodowy Instytut Normalizacji (ang. ANSI - American National Standards Institute).

Pierwszą firmą, która wprowadziła SQL do swojego produktu, była korporacja Oracle. Wprowadzanie SQL-a w produktach konkurencyjnych firm wiązało się z modyfikacją pierwotnego kodu i zwlekano z tymi zmianami ze względu na skomplikowanie tej operacji. Wkrótce, utrzymanie dalszej jednolitości kodu oraz wywierana presja rynku zmusiał producentów oprogramowania bazodanowego do wprowadzenia tego standardu do swoich projektów.

1.2. Infrastruktura aplikacji internetowych

Kluczowym składnikiem sprzętowym wykorzystywanym w aplikacjach internetowych są serwery. Zostanie opisane kilka kluczowych cech, które wpływają na ostateczną wydajność aplikacji. Większość współczesnych serwerów jest oparta na wielordzeniowej architekturze procesorów firmy Intel.

Najczęściej kupowanym typem serwerów przez klientów z całego świata są jednostki Intel Core2Duo, oraz bliźniaczy Intel Dual-Core z zegarami ok 2GHz (na każdy rdzeń). Są to najpopularniejsze procesory dostępne dla klientów indywidualnych (osób prywatnych). W ofercie każdej firmy hostingowej, u której jest możliwość wynajęcia serwera dedykowanego są również dostępne modele z procesorami typowo serwerowymi (Intel Xeon), ale ze względu na ich cenę, oraz zbyt duży potencjał mocy, który nie zostałby wykorzystany w całości, wybór najczęściej pada na tańsze maszyny. Core2Duo posiada obsługę wielowątkowych procesów, oraz ma 64 bitową infrastrukturę.

Drugim ważnym elementem serwera jest ilość i jakość pamięci operacyjnej RAM.

W typowo serwerowych maszynach występują pamięci RAM ECC-Registered w ilości nie mniejszej, niż 1GB. System operacyjny dzięki funkcji cachowania i buforowania danych, jakie musi obsłużyć jest w stanie zapełnić całą dostępną pamięć – dzięki czemu ponowny dostęp do danych będzie znacznie szybszy. Minimalną ilością pamięci RAM jaka powinna znaleźć się w serwerze to 512MB. Mniejsza ilość pamięci może stwarzać problem z uzyskaniem odpowiedniej wydajności systemu z włączonymi/uruchomionymi wszystkimi niezbędnymi usługami (apache, php, mysql, counter, smtp, pop3, spamassistant, bind, itp.).

Kolejnym elementem sprzętowym serwera jest dysk twardy, na którym zapisujemy wszystkie dane - pliki, bazy danych, system, itp.

Najlepszymi dyskami dostępnymi na rynku dla serwerów są dyski z rodziny SAS, które charakteryzują się przede wszystkim bardzo dużą szybkością obrotową talerzy, a co za tym

Autor: Szymon Bluma, www.BlueMan.pl

idzie szybszymi transferami wyszukiwania i przesyłu danych. Mimo wszystko przy dzierżawieniu serwera, ze względu na wysokie koszty dysków SAS, stosuje się tańsze dyski SATA, które mają odpowiednio zadowalającą wydajność dla niskobudżetowych projektów serwisów utrzymywanych na prywatnym serwerze dedykowanym.

1.2.1. Procesory wielordzeniowe

Współczesne serwery oparte są o procesory wielordzeniowe. Dzięki temu możliwe jest jednoczesne uruchamianie wielu wątków, procesów i aplikacji.

Wątek (z ang. thread) jest to jednostka wykonawcza w ramach jednego procesu.

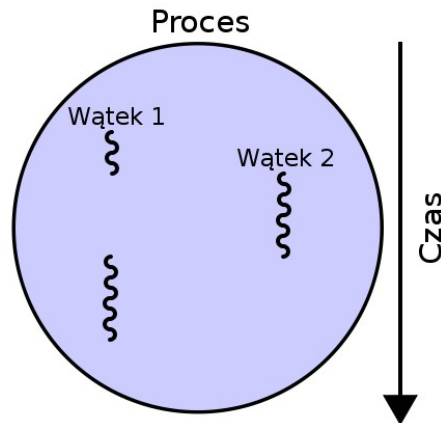
Wielowątkowość daje możliwość podzielenia, lub łączenia kilku operacji w dwa (lub więcej) symulowanych (lub pseudo-symulowanych) wykonywanych zadań.



Rys. 5 – Logo najpopularniejszego serwerowego procesora wielordzeniowego

Wielowątkowość jest zależna od systemu operacyjnego i polega na tym, że w ramach jednego wykonywanego procesu może obsłużyć kilka wątków. Nowe przychodzące wątki, to kolejne ciągi instrukcji tego samego programu wykonywane oddzielnie. W systemie nieobsługującym wielowątkowości pojęcie wątek i proces znaczą to samo, ponieważ są tym samym.

Wielowątkowość w procesorach możliwa jest dzięki temu, że nie wszystkie części jednostki wykonawczej są w jednakowym stopniu wykorzystywane przez pojedynczy wątek (ciąg instrukcji). Nieaktywne części jednostki wykonawczej mogą w tym czasie wykonywać inny wątek zwiększając efektywność wykorzystania całego procesora.



Rys. 6 – Ilustracja podziału procesu na wątki

Cechy wielowątkowości:

- wszystkie wątki wykonują się w ramach tylko jednego programu - to znaczy, że wykonują ten sam kod wykonawczy (przykładowo: w systemach Linux każdemu wątkowi przydziela się konkretną funkcję do wykonania)
- wątki zostały wprowadzone aby zwiększyć wydajność, gdzie zachodzi potrzeba wykonania wielu zadań jednocześnie
- komunikacja między wątkami w odróżnieniu od procesów jest bardzo łatwa do wykonania - w przypadku wątków wystarczy odwoływać się do tych samych zmiennych i obiektów - komunikacja między procesami wymaga zastosowania mechanizmów IPC
- współdzielenie wirtualnej przestrzeni adresowej niesie duże zagrożenie - jeden "wadliwy" wątek może zagrozić wykonaniu całego programu;

1.2.2. Pamięć DDR ECC

Wydajność serwera zależy w dużej mierze od użytej pamięci. W zależności, jaka pamięć zostanie wykorzystana można spodziewać się różnych wyników końcowych w szybkości działania serwera.

Pamięć DDR ECC jest to pamięć stosowana w specjalnie przystosowanych do tego serwerach. Moduł pamięci jest wyposażony w ECC (ang. Error Correction Code), czyli kod korekcji błędów, dzięki czemu oprócz przesyłanych danych po szynie transferowane są również kody kontrolne, pozwalające wykryć przekłamanie pojedynczych bitów w danych. Takie rozwiązanie zapewnia większą stabilność systemu, niż w przypadku wykorzystania normalnych kości pamięci (nie obsługujących ECC). Kosztem zastosowania dodatkowej korekcji jest spadek wydajności pamięci, który oscyluje w granicach 1-2%, a także większa

Autor: Szymon Bluma, www.BlueMan.pl

Zabrania się kopiowania, rozpowszechniania, udostępniania bez podania autora pracy

ilość tranzystorów dla tej samej pojemności, gdyż każda komórka pamięci zawiera oprócz danych kod korekcyjny. Aby skorzystać z takich pamięci, chipset oraz BIOS, jaki znajduje się na płycie głównej musi obsługiwać tę dodatkową funkcjonalność.

1.2.3. Dyski twarde HDD SAS, SATA

Kolejnym składnikiem, który ma duży wpływ na szybkość pracy serwera jest dysk twardy.

Typowymi szybkimi dyskami serwerowymi są dyski z interfejsem SAS. Charakteryzują się one głównie szybszą prędkością obrotową talerzy – 15'000 obrotów, a także znacznie mniejszą awaryjnością. Dyski SAS są kontynuacją rozwoju technologicznego magistrali SCSI, która pierwszy raz ujrzała światło dzienne w 1980 roku. Szacuje się, że dyski SAS (Serial Attached SCSI) będą w stanie pracować bezawaryjnie przez 15 lat. Natomiast dla dysków SATA (Serial ATA) wynik ten wynosi ok.6 lat.



Rys. 7 – wygląd złącza SAS przy dysku twardym

1.3. Podział mocy

W tym rozdziale zostanie opisane kilka technik, które mogą zostać wykorzystane w zaawansowanych systemach internetowych. Techniki te w znaczący sposób poprawiają wydajność aplikacji.

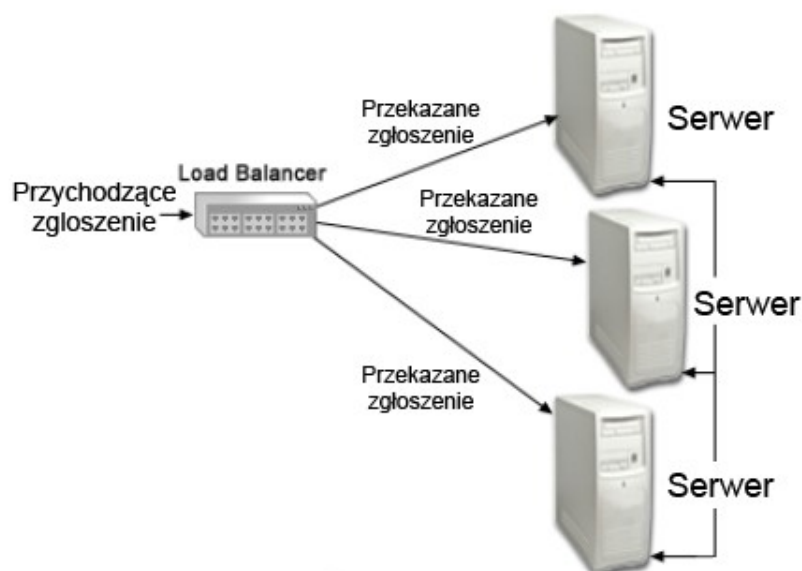
Aby serwer obsłużył jeszcze większą ilość odwiedzających witrynę, nie wystarczy wyposażyć go w odpowiedni nowoczesny i dobrze dobrany sprzęt – szybki procesor, pojemną pamięć, czy wydajny dysk twardy. Z czasem najnowsza technologia nie będzie już na tyle wystarczająca, aby zapewnić dynamikę serwisu i należy szukać innych, alternatywnych rozwiązań, którymi są: klasteryzacja, load balancing, decentralizacja usług.

1.3.1. klasteryzacja

Klaster to grupa komputerów połączonych ze sobą w celu powiększenia dostępnej mocy obliczeniowej. Połączenie komputerów odbywa się w sposób równoległego przyłączania kolejnych maszyn. Istnieje także połączenie szeregowe – opisany poniżej load balancing. Większość superkomputerów w dzisiejszych czasach działa właśnie na zasadzie połączeń równoległych. Klastry buduje się po to, aby z kilku słabszych komputerów uzyskać jeden „wirtualny”, o skumulowanej mocy wszystkich jego „węzłów” (czyli komputerów podpiętych pod główną stację roboczą). Najczęściej klastry buduje się z wykorzystaniem systemu z grupy Linux i zainstalowanym dodatkowym oprogramowaniem, oraz prekonfigurowanym jądrem. Jedną z najbardziej popularnych implementacji klastrów obliczeniowych jest klaster o nazwie Beowulf, gdzie rolę węzłów pełnią popularne komputery typu PC.

1.3.2. load balancing

To technika polegająca na rozpraszaniu generowanego obciążenia pomiędzy kilka maszyn (2 i więcej). Na każdej z tych maszyn stoi taka sama, lustrzana kopia aplikacji. Najczęściej rolę load balancera przejmuje specjalnie do tego przeznaczony switch, który na każde wezwanie użytkownika sprawdza obciążenie każdej z maszyn do niego podpiętych i przekierowuje na nią ruch. Optymalny wybór zazwyczaj prowadzi do równomiernego obciążenia wszystkich serwerów. Load balancing zapewnia także, że awaria jednej z maszyn nie wypłynie na funkcjonowanie całego serwisu i będzie niezauważalna dla użytkownika.



Rys. 8 – schemat działania load balancer'a^[6]

1.3.3. decentralizacja usług

Decentralizacja usług systemowych jest kolejnym sposobem na przyśpieszenie działania witryny internetowej. Polega ona na rozszczepieniu jednego całego systemu na kilka mniejszych, które będą wyspecjalizowane w konkretnych czynnościach/usługach na nich działających.

Podział usług polega na emigracji niektórych części usług na inne serwery, które znajdują się w sieci lokalnej głównego serwera (aby komunikacja między serwerami była możliwie jak najszybsza). Nie ma przyjętego schematu, w jaki sposób powinniśmy podzielić system. Jednak najbardziej newralgicznym elementem systemu zazwyczaj jest system bazy danych, który potrzebuje stosunkowo dużo mocy obliczeniowej procesora na operacjach na danych. Często także serwis, który oferuje wysyłanie i ściąganie plików, lub ogólnie bazuje na dużej ilości statycznych plików korzysta z dodatkowego serwera na te pliki – aby odciążyć odczyt/zapis na dysku twardym. W tym wypadku nie ma wymogu, aby ten dodatkowy serwer stał w tej samej serwerowni, a wręcz nie powinien, ponieważ kolejnym „wąskim gardłem” będzie przepustowość łącza. W przykładzie mysql, gdzie dane wysłane/pobrane z bazy danych muszą być ponownie przetworzone przez kod PHP, nie ma obawy o wąskie gardło łącza dostępowego serwerowni – dane te są nieporównywalnie małe w stosunku do przesyłanego zdjęcia/obrazka/pliku. Żądania pobrania/wysłania dużych plików nie muszą współpracować z innymi modułami Apache/PHP, więc ich obsługa (wyświetlenie) jest jedynie zadaniem realizowanym po stronie użytkownika przeglądającego serwis.

Autor: Szymon Bluma, www.BlueMan.pl

Możliwy jest więc podział usług na kilka serwerów w ten sposób:

- główny serwer z zainstalowanym Apache i PHP
- osobny serwer bazy danych MySQL (lub innej)
- serwer obsługujący statyczne pliki użytkowników, np. zdjęcia

2. Opis/historia skryptów

Forum internetowe – czyli następca grupy mailingowej, w dobie przyjętej mody web 2.0 jest najpopularniejszą metodą tworzenia wirtualnej społeczności ludzi, którzy zostawiając swoje wypowiedzi, wiedzę, opinie w internecie budują swoje drugie życie.

2.1. phpBB

Skrypt ten jest jednym z najpopularniejszych skryptów używanym w internecie. Popularność ta spowodowana jest długim czasem istnienia tego oprogramowania, jak i licznym modyfikacjom, jakie są do niego udostępniane.



Rzeczony skrypt `phpBB`^[7] został zapoczątkowany przez Jamesa Atkinsona w czerwcu 2000 roku. Zaczął on tworzyć to oprogramowanie na potrzeby biuletynu informacyjnego na stronie swojej żony. Kiedy skrypt został udostępniony jako otwarte oprogramowanie w popularnym serwisie SourceForge zyskał bardzo szybko duże zainteresowanie. Do grupy programistów tworzących ten skrypt dołączyli Nathan Coddington oraz John Abela i w grudniu tego samego roku ukazała się pierwsza, pełna wersja `phpBB 1.0.0`. Kolejne wersje, tj. 1.2 i 1.4 zostały odpowiednio wydane w lutym i kwietniu 2001 roku. Natomiast finalna wersja gałęzi 1.x, tj. 1.4.4 została opublikowana w listopadzie 2001 roku.

Pracę nad wersją `phpBB 2.0.x` rozpoczęły się w lutym 2001 i trwały równocześnie z udoskonalaniem stabilnej wersji 1.x. Nowy skrypt został napisany od zera. Po roku produkcji i intensywnych testach nowej wersji, wydano ją oficjalnie 4 kwietnia 2002 roku.

Wersja `phpBB 2.x` zyskała największą popularność nie tylko w krajach anglojęzycznych, ale również na całym świecie dzięki wielu wersjom językowym. Końcową wersją serii 2.2.x została ogłoszona 14 stycznia 2005. W ogłoszeniu tym zakomunikowano, że wersja `phpBB 2` dobiega końca i rozpoczęto pracę nad nową wersją `phpBB 3.0.0` o kodowej nazwie „Olympus”.

Pierwsza wersja beta nowego skryptu Olympus powstała w marcu 2007, a 13 grudnia 2007 została upubliczniona stabilna wersja `phpBB 3`.

2.2 phpBB2 by Przemo

Skrypt phpBB2 by Przemo^[8] bazuje na skrypcie z rodziny phpBB w wersji 2.x.x. Zyskał on popularność wśród rzeszy użytkowników internetu dzięki prostocie instalacji i wielu dostępnych preinstalowanych modyfikacji, którymi w prosty sposób można zarządzać wprost z poziomu panelu administracyjnego forum.



Dokładna historia tego skryptu autorstwa Przemysława Matogi nie jest dokładnie znana i nigdzie nie została opublikowana. Wiadomo jednak, że Matoga zaczął modyfikować dla własnych potrzeb istniejący wtedy skrypt phpBB2 dodając do niego kolejne swobodnie dostępne modyfikacje. Z biegiem czasu zainteresowanie skryptem, który był już lekko zmodyfikowany wzrosła, więc Matoga postanowił opublikować swoje dzieło – wtedy właśnie rozpoczęła się ekspansja i jeszcze większe zainteresowanie tą modyfikacją. Ludzie, po prostu potrzebowali skryptu, który spełnia ich wymagania dzięki licznym dodatkowym funkcjonalnościom, a niestety standardowy skrypt phpBB2 był dość ubogi w tego typu możliwości.

Przemysław Matoga zapowiadał, że w miesiąc po wydaniu oficjalnej wersji phpBB3 przepisze także i swoją modyfikację na nowy silnik. Niestety do dnia dzisiejszego (mimo kolejnego wydania phpBB 3.0.1) nie doczekaliśmy się jego nowego zmodyfikowanego skryptu.

2.3 SMF (Simple Machines Forum)

Skrypt ten powstał w odpowiedzi na działający w obecnych czasach (przed wrześniem 2003r.) skrypt YaBB SE, który to miał bardzo złą reputację ze względu na problemy z kodem napisanym w Perlu,



bardzo podobnej nazwy innego oprogramowania YaBB, a także wieloma lukami bezpieczeństwa jakie w nim występowały. Prace nad SMF^[9] zaczęły się jako prace nad małym projektem zapoczątkowanym przez deweloperów YaBB. Ich głównym założeniem było rozszerzyć swoje nowe dzieło (SMF) o bardziej zaawansowaną obsługę szablonów graficznych, większe bezpieczeństwo skryptu oraz uczynić go bardziej popularnym. Pierwsza wersja beta skryptu SMF 1.0 Beta 1a została udostępniona 30 września 2003 tylko dla wybranej grupy testerów („Charter member”).

Dzisiaj aktualną publiczną, stabilną wersją jest 1.1.5, która została wydana 30 kwietnia 2008, a także trwają równorzędne prace nad nowym wydaniem SMF 2.0, które obecnie jest w wersji beta 4.

Autor: Szymon Bluma, www.BlueMan.pl

2.4 MyBB (MyBulletinBoard)

Historia skryptu MyBB^[10] rozpoczęła się przed 2002 rokiem, kiedy to Chris Boulton, web34rk oraz b0ndman rozpoczęli programowanie skryptu, który nazwali XMB. Po jakimś czasie web34rk i b0ndman zrezygnowali z kontynuacji programowania dla tego projektu. Przez kolejne miesiące następowały następne rotacje w zespole programistów, aż w końcu napisano wersję o kodowej nazwie „DevBB”, którą można uznać za ojca dzisiejszego MyBB. W okolicach września 2002 roku została wydana pierwsza nieoficjalna wersja MyBB. Rok później świat ujrzał MyBB 1.00 RC1. Podczas kolejnych, przemijających lat w zespole programistów tego oprogramowania trwały liczne transfery osób, aż w końcu 2 września 2006 roku światło dzienne ujrzała wersja 1.2.0 skryptu MyBB.



Obecnie najnowszą, stabilną wersją skryptu jest 1.2.12 wydana 20 stycznia 2008, a także trwają nieprzerwane prace nad nową wersją 1.4, której przypuszczalna data wydania nadal nie jest znana, a deweloperzy jeszcze nie potrafią jej określić.

3. Opis teoretyczny testów

Część praktyczną pracy dyplomowej stanowiły testy wydajnościowe skryptów prezentowanych w poprzednim rozdziale.

Testy kolejnych skryptów forów dyskusyjnych przebiegały w 4 godzinnych sesjach. W skład każdej sesji wchodziły:

- 1 godzina pracy serwera bez obciążenia
- 2 godziny testów obciążeniowych aplikacją WAPT
- 1 godzina ustabilizowania serwera po testach

Zgodnie z zaprezentowanym wyżej wzorcem, na wszystkich wykresach (oprócz użycia serwera SQL) po godzinie pracy serwera widać ewidentny skok, a pod koniec wygenerowanego wykresu można zauważyć obniżenie wszystkich wartości. Każdy z testów został wykonany 3-krotnie, a średnie wyniki zostały zaprezentowane w dalszej części. Wydajność mierzona jest za pomocą dwóch aplikacji. Jedna działa po stronie serwera, druga na komputerze klienckim.

3.1. Aplikacja WAPT

Aby wygenerować reprezentatywny ruch użytkowników potrzebny do przeprowadzenia testów wydajności stron internetowych, należy skorzystać z odpowiedniego, gotowego oprogramowania służącego do tego celu.

Aplikacja WAPT^[11] została wybrana, ponieważ daje największą możliwość konfiguracji względem innych darmowych aplikacji do generowania obciążenia stron www.

Najważniejszymi walorami tego oprogramowania są:

- ustawienie ilości iteracji w ciągu jednego wywołania
- generowanie własnej trasy przejść użytkownika pomiędzy stronami
- łatwość edycji istniejącej trasy użytkownika, dzięki otwartemu źródłowi pliku
- ustawienie czasu trwania testu
- ustawienie daty rozpoczęcia testu (harmonogram zadań)
- generowanie raportów po wykonaniu testów

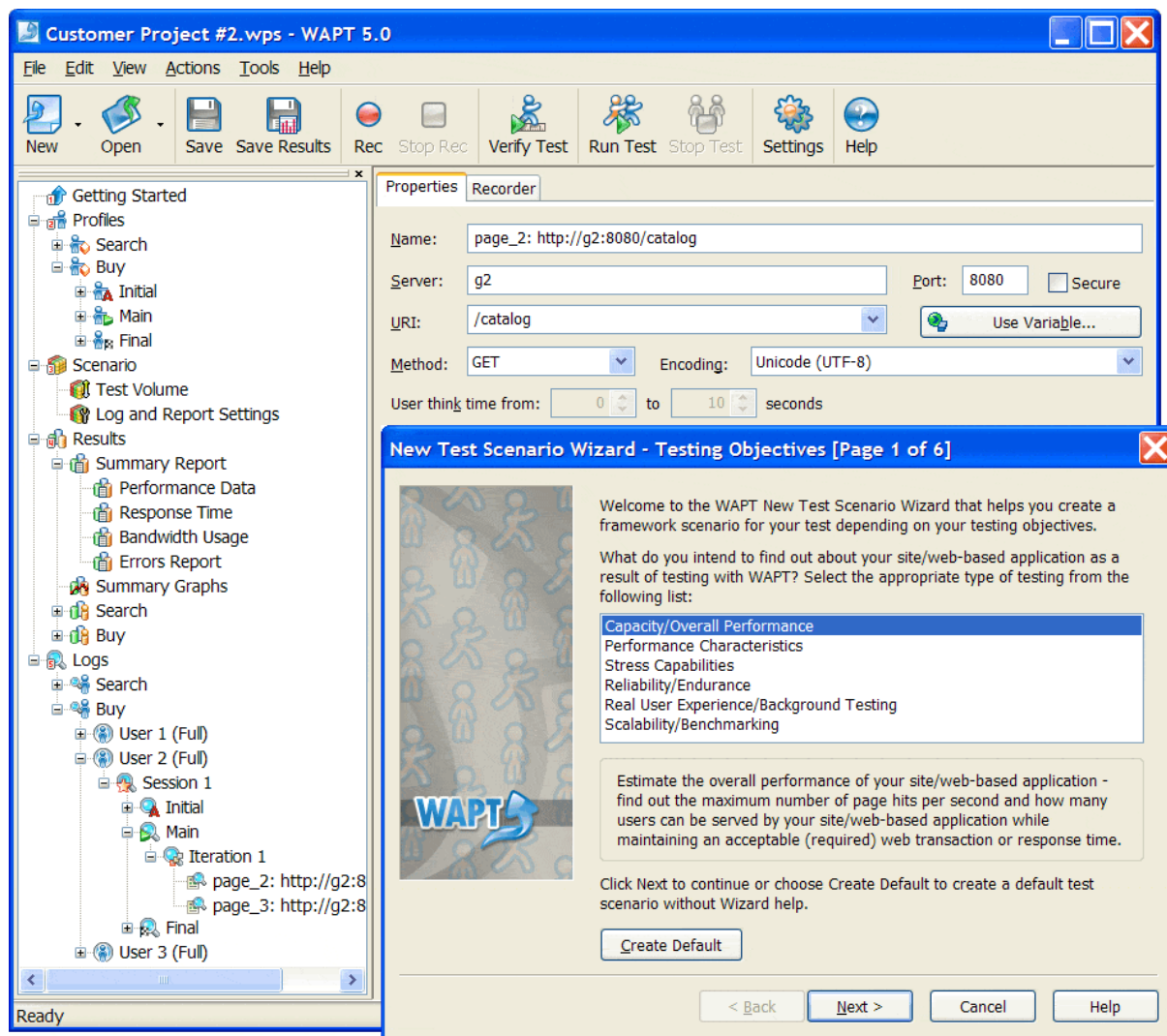
Autor: Szymon Bluma, www.BlueMan.pl

Zabrania się kopiowania, rozpowszechniania, udostępniania bez podania autora pracy

- utworzenie kilku różnych typów profili użytkowników przeglądających stronę

WAPT jest darmową w pełni funkcjonalną aplikacją z licencją 30-dniową. Jedyne ograniczenia w tej wersji dotyczą limitu wirtualnych użytkowników – 20 na jeden test.

Program działa pod wszystkimi systemami operacyjnymi z rodziny Windows z wyłączeniem Windows 95 i Windows Vista.



Rys. 9 – Główne okno aplikacji WAPT

Autor: Szymon Bluma, www.BlueMan.pl

Zabrania się kopiowania, rozpowszechniania, udostępniania bez podania autora pracy

3.2. aplikacja munin

Munin^[12] to system statystyk obciążenia/wykorzystania serwera napisany w Perlu. Działa on na serwerze zbierając na bieżąco dane, oraz generując co 5 min wykresy, które można zobaczyć korzystając z każdej graficznej przeglądarki internetowej.

Dzięki jego funkcjonalności mamy możliwość uzyskania przejrzystych informacji o:

- ilości uruchomionych procesów Apache
- procentowej zajętości poszczególnych partycji na dysku twardym
- ilości i rodzajach zapytań do serwera bazy danych
- ilości wolnych zapytań do bazy danych (wykonywanych dłużej, niż >2 sec)
- wykorzystanej przepustowości karty sieciowej
- ilości uruchomionych procesów w systemie
- temperaturze procesora
- użyciu procesorów
- średnim obciążeniu systemu (tzw. load)
- użyciu pamięci RAM
- użyciu dodatkowej pamięci SWAP
- i innych informacjach

3.3. sprzęt użyty w testach

Środowiskiem do testowania będą dwa komputery typu PC połączone w sieci lokalnej LAN, aby uzyskać możliwie jak najszybszy przesył dużych ilości pakietów.

Pierwszym komputerem – klientem będzie komputer PC o następujących parametrach:

- Procesor: Athlon XP Barton 2500+
- Pamięć RAM: 1GB DDR RAM
- Dysk HDD: WD 80GB HDD 7200obr.
- K. graficzna: ATI Radeon 9200 64MB
- System operacyjny: Windows XP
- aplikacja do testowania: WAPT 5.0

Autor: Szymon Bluma, www.BlueMan.pl

Zabrania się kopiowania, rozpowszechniania, udostępniania bez podania autora pracy

Komputerem pełniącym rolę serwera, na którym zostały umieszczone strony jest laptop firmy Hewlett-Packard:

- Procesor: Intel Core2Duo 1,66MHz
- Pamięć RAM: 2GB DDR2 RAM
- Dysk HDD: Toshiba 160GB HDD 5200 obr.
- K. graficzna: Intel X3100
- System operacyjny: Ubuntu 7.10, 32 bit (wraz z Apache 2.2.4, PHP 5.2.3, MySQL 5.0.45)
- monitoring pracy serwera Munin 1.2.5

3.4. opis ścieżki użytkownika

Z uwagi na to, że większość ruchu na forum dyskusyjnym generują osoby niezarejestrowane/niezałogowane, a właściwie roboty wyszukiwarek internetowych, symulowane będzie przypuszczalne zachowanie tylko takich użytkowników.

Przy generowaniu trasy użytkownika na forum zostanie wzięte pod uwagę:

- a) wędrówka użytkownika po subforach
- b) wędrówka użytkownika po postach
- c) przejście oglądającego przez podstrony w danym temacie
- d) dynamiczne przeskakiwanie pomiędzy subforami (z menu rozwijanego)
- e) przeglądanie listy zarejestrowanych użytkowników forum
- f) wyszukiwanie wszystkich wypowiedzi na forum danego użytkownika
- g) korzystanie z wyszukiwarki forum
- h) oglądanie profili użytkowników
- i) wejście na stronę główną

W testach wykorzystana zostanie następująca trasa:

typ zachowania	ID w bazie	opis	czas [s]
i)		Strona główna forum	18
a)	8	Forum „Przywitalnia”	4
b)	10987	Temat „Powrót po latach”	11
h)	3683	Profil użytkownika „smoczyca”	4
f)	smoczyca	Szukanie wszystkich wypowiedzi napisanych przez użytkownika „smoczyca”	12
b)	163835	Skok do postu w temacie „Zdrada psychiczna”	9
b)	10741 2 strona	Druga podstrona tematu „Zdrada psychiczna”	5
b)	10741 3 strona	Trzecia podstrona tematu „Zdrada psychiczna”	9
d)	59	Skorzystanie z rozwijanego menu dostępnych działów, aby przeskoczyć do działu „Gry i zabawy”	7
b)	7214	Temat „1000 powodów, żeby się napić”	5
b)	7214 2 strona	Druga podstrona tematu „1000 powodów, żeby się napić”	3
b)	7214 3 strona	Trzecia podstrona tematu „1000 powodów, żeby się napić”	4
b)	7214 4 strona	Czwarta podstrona tematu „1000 powodów, żeby się napić”	8
b)	7214 5 strona	Piąta podstrona tematu „1000 powodów, żeby się napić”	4
b)	7214 6 strona	Szósta podstrona tematu „1000 powodów, żeby się napić”	5
a)	59	Forum „Gry i zabawy”	6
b)	7300	Temat „Powody, dla których się nie napiję :)”	4
h)	1870	Profil użytkownika „privace”	6
f)	privace	Szukanie wszystkich wypowiedzi napisanych przez	5

Autor: Szymon Bluma, www.BlueMan.pl

Zabrania się kopiowania, rozpowszechniania, udostępniania bez podania autora pracy

		użytkownika „privece”	
f)	privece 2 strona	Szukanie wszystkich wypowiedzi napisanych przez użytkownika „privece”, druga strona	5
e)		Lista wszystkich użytkowników forum. Posortowana rosnąco wg daty rejestracji.	4
e)	2 strona	Lista wszystkich użytkowników forum. Posortowana rosnąco wg daty rejestracji.	3
e)	3 strona	Lista wszystkich użytkowników forum. Posortowana rosnąco wg daty rejestracji.	3
e)	4 strona	Lista wszystkich użytkowników forum. Posortowana rosnąco wg daty rejestracji.	6
i)		Strona główna forum	3
g)	komputer	Wyszukiwanie wyrazu „komputer” na forum	13
g)	dypłom	Wyszukiwanie wyrazu „dypłom” na forum	13
g)	przyjaciół	Wyszukiwanie wyrazu „przyjaciół” na forum	16
a)	52	Forum „Kobieta i Mężczyzna”	7
b)	11006	Temat „Internetowa miłość”	2
i)		Strona główna forum	0

Tab. 1 – Ścieżka ruchu użytkownika podczas testów

Każdy ze skryptów został uzupełniony bazą postów, tematów, użytkowników forum^[13] i będzie posiadać takie statystyki:

- Liczba postów: **163'012**
- Liczba wątków: **10'262**
- Liczba użytkowników: **4'429**

Wartości te mogą nieznacznie różnić się w zależności od tego, w jaki sposób konwerter dostosował domyślną bazę do swojego skryptu.

Trzeba tu wspomnieć o sposobie konwersji jaki jest wykorzystywany w MyBB. Oryginalna baza z podanymi wyżej ID w tabelce została w ten sposób przekonwertowana, że każda wypowiedź, temat, forum dostało swój własny, nowy wpis ID w bazie. Konwerter zamiast przepisać istniejące już ID, wykonywał samodzielnie od początku auto inkrementację numerów. Dlatego trasy ruchu użytkownika wewnątrz forum MyBB została opracowana

Autor: Szymon Bluma, www.BlueMan.pl

oddzielnie, na podstawie nazw odwiedzanych profili, tematów, działów.

Konwersja zaproponowana przez MyBB nie jest pozytywnym zjawiskiem, ponieważ nie wpływa na lepszą wydajność bazy danych, jak i na inne cechy funkcjonalności forum. Jest ona natomiast bardzo niekorzystna pod względem spójności identyfikatorów, jakie istniały przed konwersją oraz są po konwersji. Tracimy w ten sposób kontrolę nad kluczami głównymi wpisów, wątków, użytkowników, itp.

3.5. porównanie skryptów

Mimo że testowane skrypty posiadają bardzo zróżnicowane funkcjonalności (od najbardziej rozwiniętego *phpBB by Przemo*, po najprostszy *phpBB2*), to jednak osoby wybierające skrypt do swojego nowego forum powinny wziąć pod uwagę i rozważyć, jakich funkcjonalności potrzebują, i ile odwiedzających przewidują podczas największych szczytów oglądalności swojej społeczności.

Nie można z góry zaproponować gotowego rozwiązania, ponieważ nie będzie ono dla wszystkich administratorów uniwersalne. Do takiej decyzji trzeba podejść indywidualnie i po zapoznaniu się z różnicami funkcji, jakie oferuje oprogramowanie, wybrać najbardziej odpowiedni skrypt dostosowany do swoich wymagań lub dzięki możliwości wgrania dodatkowych modyfikacji na forum rozbudować forum na własną rękę. Niestety takie nieumiejętne prywatne działania mogą doprowadzić do znacznego pogorszenia wydajności skryptu. Najczęściej nieautoryzowane przez dystrybutora modyfikacje nie są prawidłowo zoptymalizowane i potrafią generować gigantyczne obciążenie systemu.

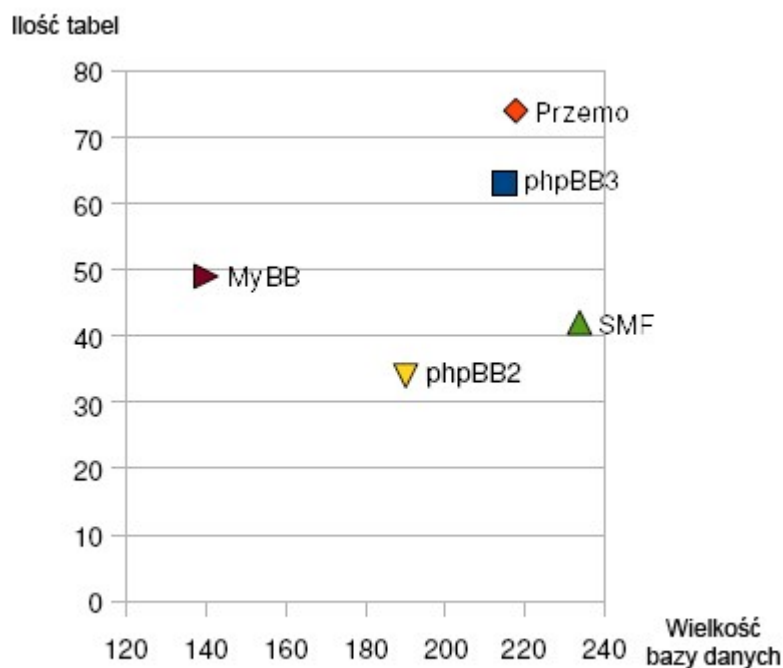
Do rozważanego przykładu uzupełnionej bazy danych wielkości danych w poszczególnych systemach kształtują się następująco:

Skrypt	Ilość tabel w bazie	Wielkość bazy danych
phpBB2 by Przemo	74	217.8 MB
phpBB2	34	190.1 MB
phpBB3	63	214.9 MB
SMF	42	233.8 MB
MyBB	49	139.9 MB

Tab. 2 – Wielkość bazy danych zajmowanej przez skrypty

Autor: Szymon Bluma, www.BlueMan.pl

Co można zilustrować poniższym wykresem:



Rys. 10 – wykres zależności ilości tabel w bazie (oś Y), do wielkości bazy danych (oś X)

Można zauważyć, że nie występuje zbieżność między wielkością bazy, a ilością tabel, co świadczy o różnym sposobie modelowania danych. Wielkość bazy danych nie ma znaczącego wpływu na wydajność skryptu. Jedyne fakty, na który może mieć wpływ, to korzystanie z wyszukiwarki.

Każdy ze skryptów mimo prostej idei – wymiany poglądów i dyskusji zarejestrowanych użytkowników, cechuje dodatkowa funkcjonalność, która z jednej strony ułatwia prace administracyjne oraz daje dodatkowe możliwości przy pisaniu postów, ale z drugiej strony powoduje większe obciążenie serwera.

Poniżej przedstawiam niektóre różnice w funkcjonalności poszczególnych skryptów:

Funkcja	phpBB by Przemo	phpBB2	phpBB3	SMF	MyBB
shoutbox	Tak	Nie	Nie	Nie	Nie
System ostrzeżeń	Tak	Nie	Tak	Nie	Tak**
Własne przyciski BBcode	Nie	Nie	Tak	Nie	Nie
Załączniki w postach	Tak	Nie	Tak	Tak	Tak
Sondy z możliwością wielokrotnego wyboru	Tak	Nie	Tak	Tak	Tak
Własne pola w profilu	Tak	Nie	Tak	Nie	Tak
Subforum jako odnośnik	Tak	Nie	Tak	Nie	Nie
Zagnieżdżanie subfor	Tak	Nie	Tak	Tak	Tak
Wielu liderów grup	Nie	Nie	Tak	Nie	Tak
Kolor grup	Tak	Nie	Tak	Tak	Tak
Zgłaszanie niestosownych postów	Tak	Nie	Tak	Nie	Tak
Ogłoszenia globalne	Tak	Nie	Tak	Nie	Nie
Szybka odpowiedź	Tak	Nie	Nie	Tak	Tak
Cytowanie selektywne	Tak	Nie	Tak*	Tak	Nie
Wykorzystanie obrazków captcha	Tak	Nie	Tak	Tak	Tak
RAZEM	13 – Tak 2 – Nie	0 – Tak 15 – Nie	13 – Tak 2 – Nie	7 – Tak 8 - Nie	10 – Tak 5 - Nie

Tab. 3 – Spis możliwości skryptów forów internetowych

* - cytowanie selektywne działa w momencie pisania odpowiedzi na temat na osobnej podstronie (bez użycia modułu „szybkiej odpowiedzi”) w głównym wątku tematu.

** - system ostrzeżeń w MyBB działa na zasadzie wystawiania publicznej opinii (punktacji). Funkcjonalność ta jest dostępna dla każdego zalogowanego użytkownika, nie tylko dla moderatorów.

Autor: Szymon Bluma, www.BlueMan.pl

Zabrania się kopiowania, rozpowszechniania, udostępniania bez podania autora pracy

4. Część praktyczna – testy

W tej części zostaną przedstawione wyniki przeprowadzonych testów.

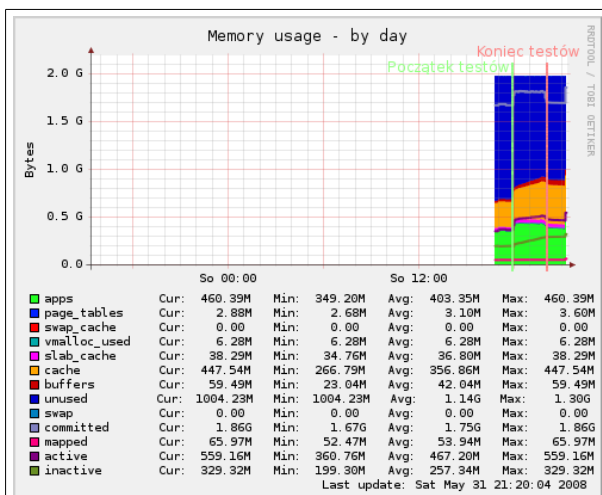
Najmłodszym skryptem jest phpBB3, który jest następcą rozwijanego przez kilka lat phpBB2. Z racji wieloletniego doświadczenia zespołu oraz świeżości wydania skryptu, jest on podejrzany o najbardziej wydajne działanie. Z drugiej strony najbardziej naszpikowany dodatkami phpBB by Przemo powinien być najmniej wydajnym skryptem.

4.1. munin

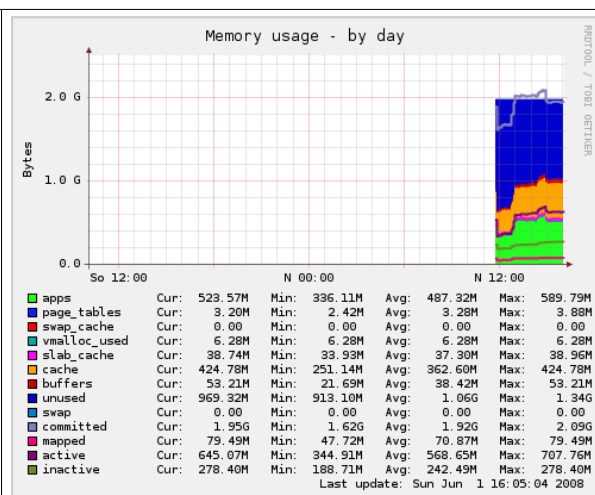
4.1.1. porównanie RAM

Mimo, że ilość zużywanej pamięci RAM nie jest najważniejszym czynnikiem, który może zadecydować o rozwiązaniu umowy z firmą hostingową, to jednak warto mieć na uwadze, to w jakim stopniu forum internetowe korzysta z tych zasobów.

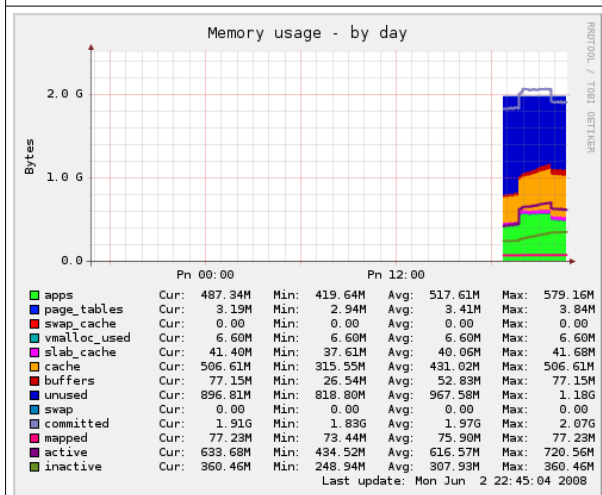
Charakterystyczny skok wykorzystania zasobów został zaprezentowany na pierwszym rysunku.



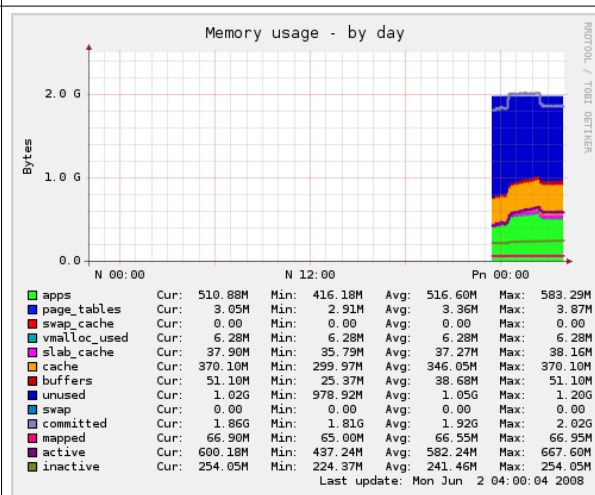
phpBB2



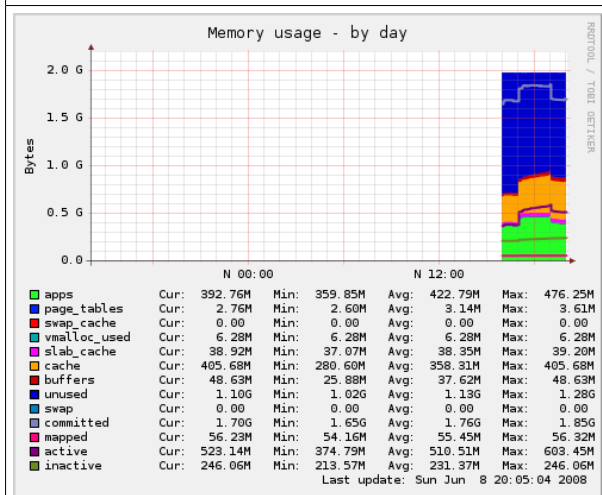
phpBB3



phpBB2 by Przemo



SMF1



MyBB

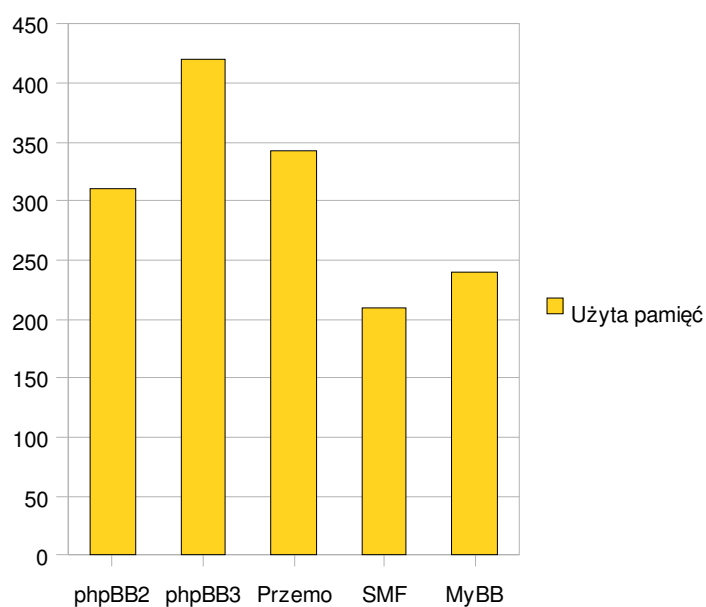
Autor: Szymon Bluma, www.BlueMan.pl

Zabrania się kopiowania, rozpowszechniania, udostępniania bez podania autora pracy

	RAM		
	cache [MB]	apps [MB]	razem [MB]
phpBB2	200	110	310
phpBB3	170	250	420
phpBB2 by Przemo	185	160	345
SMF	70	140	210
MyBB	125	115	240

Tab. 4 – Tabelka wykorzystanej pamięci RAM

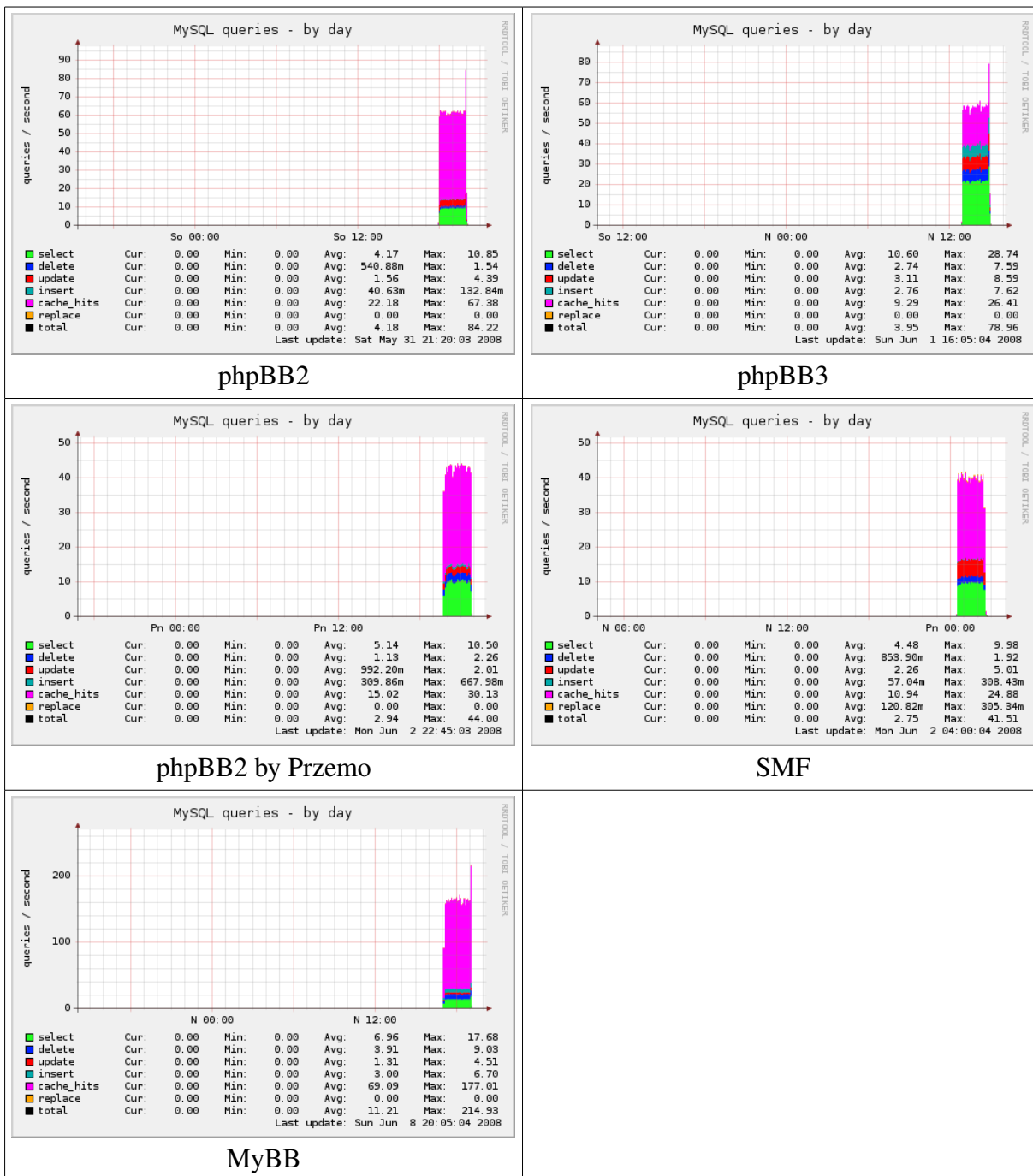
Powyższą tabelkę można przedstawić w postaci wykresu słupkowego:



Rys. 11 – wykres słupkowy całkowitej wykorzystanej pamięci RAM

4.1.2. porównanie SQL

Poniżej przedstawione zostały wykresy ilości zapytań, jakie są wykonywane przez serwer bazy danych. Wysokość słupka odpowiada łącznej liczbie zapytań SELECT, DELETE, INSERT, UPDATE.



Autor: Szymon Bluma, www.BlueMan.pl

Zabrania się kopiowania, rozpowszechniania, udostępniania bez podania autora pracy

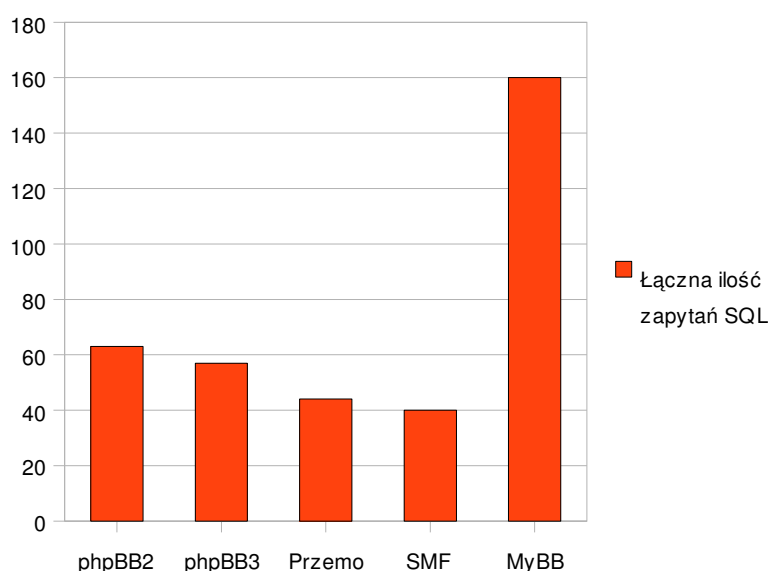
Przy wykresach serwera SQL nie ma nagłego skoku charakterystycznego dla specyfikacji testu, tj. 1 godz. przed testem, 2 godz. testu oraz 1 godz. po teście. Wynika to z tego, że zarówno przed, jak i po serwer SQL nie jest wykorzystywany - nikt nie korzysta/odwiedza testowanych forów internetowych, więc nie są wywoływane żadne zapytania do bazy danych.

W tym wypadku niektóre różnice pomiędzy poszczególnymi typami zapytań do bazy danych w różnych skryptach są tak niewielkie, że można je pominąć. Zostanie więc wzięta pod uwagę tylko całkowita suma wszystkich typów zapytań.

Zapytania SQL	
phpBB2	63
phpBB3	57
phpBB2 by Przemo	44
SMF	40
MyBB	160

Tab. 5 – Tabelka średnich wartości wykonywanych zapytań SQL

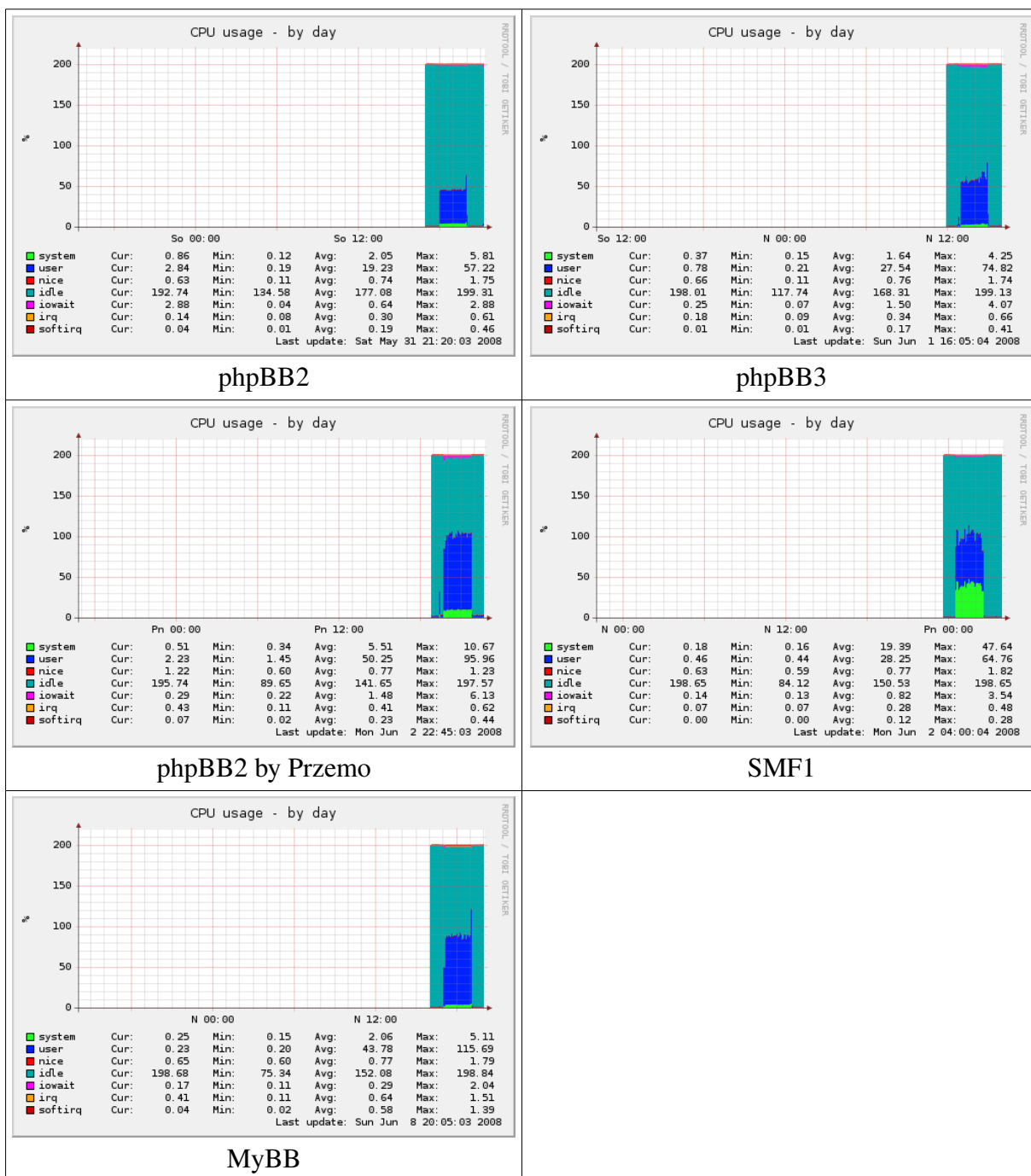
Wykres porównujący otrzymane wartości:



Rys. 12 - wykres ilości używanych zapytań do bazy danych

4.1.3. porównanie CPU

Aby obsłużyć serwer bazy danych, serwer apache, interpreter PHP, serwer potrzebuje odpowiedniej mocy obliczeniowej procesora CPU. Poniżej zostały przedstawione uzyskane wartości obciążenia procesora.



Autor: Szymon Bluma, www.BlueMan.pl

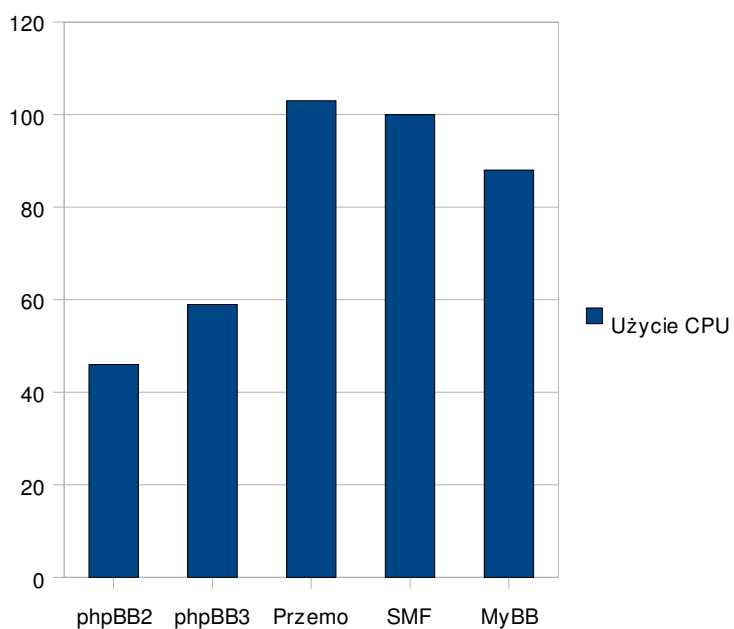
Zabrania się kopiowania, rozpowszechniania, udostępniania bez podania autora pracy

Średnie zużycie procesora [%]	
phpBB2	46
phpBB3	59
phpBB2 by Przemo	103
SMF	100
MyBB	88

Tab. 6 – Tabelka średniego zużycia procesora

Ze względu na specyfikację sprzętową serwera, tj. wykorzystanie procesora dwurdzeniowego, system, jak również inne oprogramowanie widzi fizycznie dostępne 2 procesory. Każdy z procesorów może zostać obciążony w 100%, a więc łącznie dopuszczalne obciążenie wynosi 200%. Uzyskany wynik 103% w skrypcie phpBB2 by Przemo mówi o pełnym wykorzystaniu jednego z procesorów, podczas gdy drugi miał jeszcze wolne 97% mocy.

Wykres uzyskanych wartości:



Rys. 13 – wykorzystanie procesora

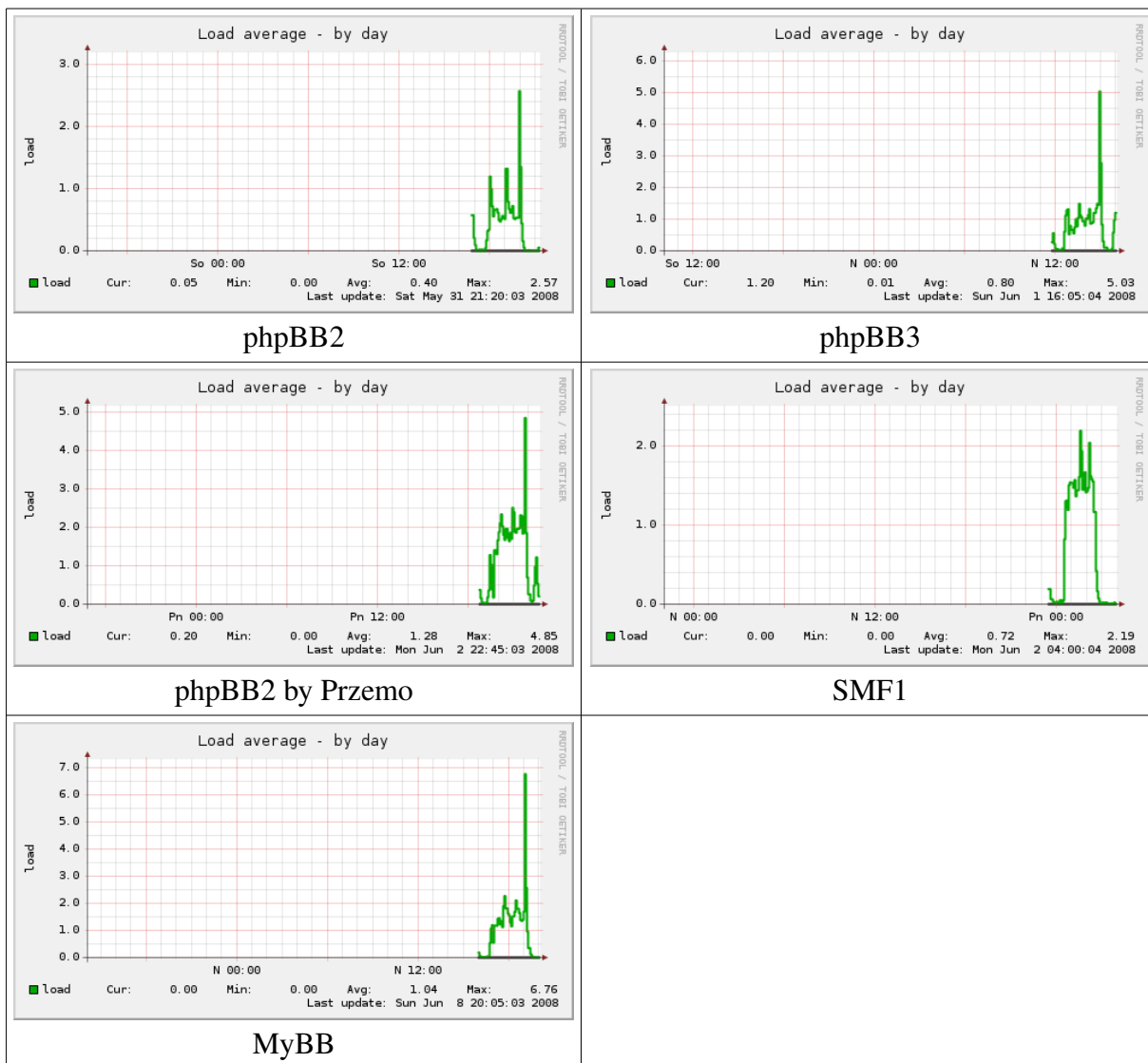
Autor: Szymon Bluma, www.BlueMan.pl

Zabrania się kopiowania, rozpowszechniania, udostępniania bez podania autora pracy

4.1.4. porównanie LOAD

Load jest to linuxowy wyznacznik średniego obciążenia serwera. Bierze on pod uwagę procesy aktualnie obsługiwane przez procesor, procesy czekające na obsłużenie, oraz procesy czekające na dostęp do danych na dysku twardym komputera.

Nagły skok load, jaki występuje w chwili kończenia testów jest cechą charakterystyczną oprogramowania WAPT. Skok ten nie zależy od testowanego skryptu forum.



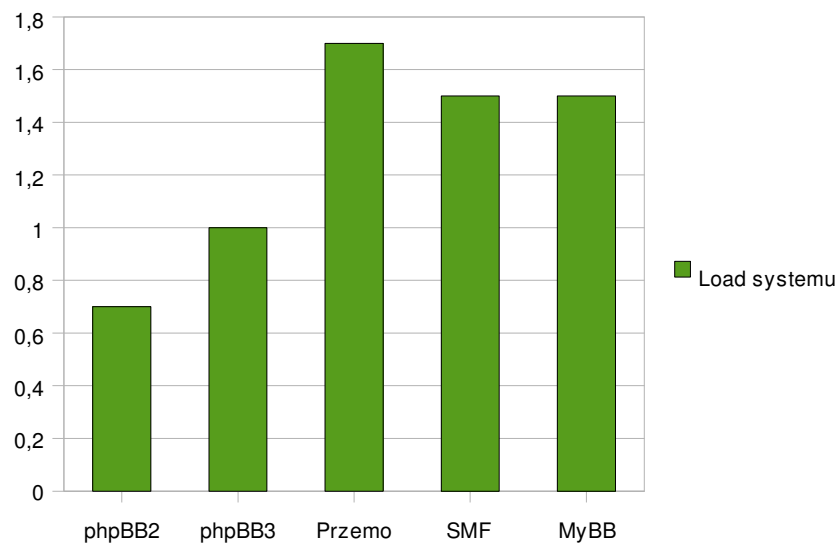
Autor: Szymon Bluma, www.BlueMan.pl

Zabrania się kopiowania, rozpowszechniania, udostępniania bez podania autora pracy

Średnie obciążenie serwera	
phpBB2	0.7
phpBB3	1.0
phpBB2 by Przemo	1.7
SMF	1.5
MyBB	1.5

Tab. 7 – Tabelka średniego obciążenia serwera

Średnie obciążenie skryptów przedstawione zostało na wykresie:



Rys. 14 – wykres obciążenia serwera

5. Wnioski

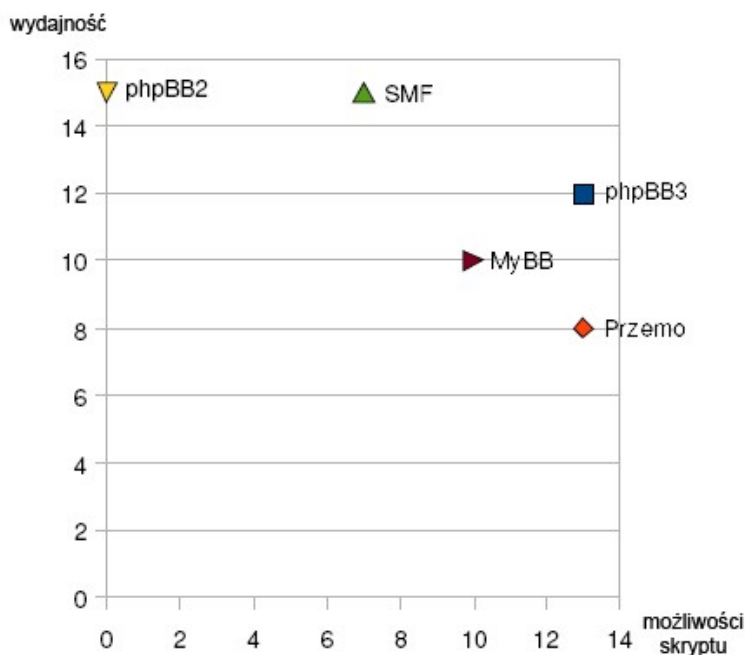
	RAM	SQL	CPU	LOAD	możliwości skryptu
phpbb2	3	2	5	5	0
phpbb3	1	3	4	4	13
przemo	2	4	1	1	13
smf	5	5	2	3	7
mybb	4	1	3	2	10

Tab. 8 – podsumownie, punktacja za poszczególne testy

Punktacja została przydzielona na podstawie każdego z testów osobno. Za najbardziej optymalne wartości była przyznawana wartość 5, za najmniej 1. Maksymalnie można, więc było osiągnąć 20 punktów.

Za każdą funkcjonalność^[tab.3] jaką spełnia dany skrypt był przyznawany 1 (słownie: jeden) punkt, za brak takiej możliwości nie przyznawano żadnego punktu.

Wykres zależności możliwości do wydajności skryptu wygląda następująco:



Rys. 15 – wykres porównawczy możliwości i funkcjonalności skryptów

Autor: Szymon Bluma, www.BlueMan.pl

Skrypt phpBB2 uzyskał wynik 0 (słownie: zero) za dostępne w nim funkcjonalności z tego względu, że nie posiada on żadnej cechy, którą posiadają inne skrypty. Nie ma także żadnej innej dodatkowej możliwości, która byłaby unikatowa na tle konkurencyjnego oprogramowania.

Wykres^[rys. 15] w klarowny sposób przedstawia zależność wydajności od dostępnych możliwości. Dla osób, które cenią sobie przede wszystkim wydajność najlepszym wyborem będzie skrypt phpBB2. Osoby chcące uzyskać maksymalną dostępną funkcjonalność, którym nie zależy na odpowiedniej optymalizacji kodu, powinny wybrać phpBB by Przemo.

Warto podkreślić, że wyniki powyższych testów są potwierdzeniem obiegowych opinii administratorów serwerów internetowych. Opinie takie można przeczytać na licznych listach dyskusyjnych^[15].

6. Bibliografia

- [1] <http://www.debian.org/>
- [2] <http://www.debian.org/doc/manuals/project-history/>
- [3] <http://httpd.apache.org/>
- [4] <http://php.net/>
- [5] <http://www.mysql.com/>
- [6] <http://publib.boulder.ibm.com/infocenter/eserver/v1r2/topic/ewlminfo/eicaawkldbaling.htm>
- [7] <http://www.phpbb.com/>
- [8] <http://www.przemo.org/phpBB2/>
- [9] <http://www.simplemachines.org/>
- [10] <http://www.mybboard.net/>
- [11] <http://www.loadtestingtool.com/>
- [12] <http://munin.projects.linpro.no/>
- [13] <http://www.blueforum.pl/>
- [14] <http://www.forumorange.net/>
- [15] <http://www.webhostingtalk.pl/>