

POZNAŃSKIE CENTRUM SUPERKOMPUTEROWO-SIECIOWE



AFILIOWANE PRZY INSTYTUCIE CHEMII BIOORGANICZNEJ PAN

h t t p : / / w w w . m a n . p o z n a n . p l



Skalowanie i monitorowanie działania systemu dLibra 5.0

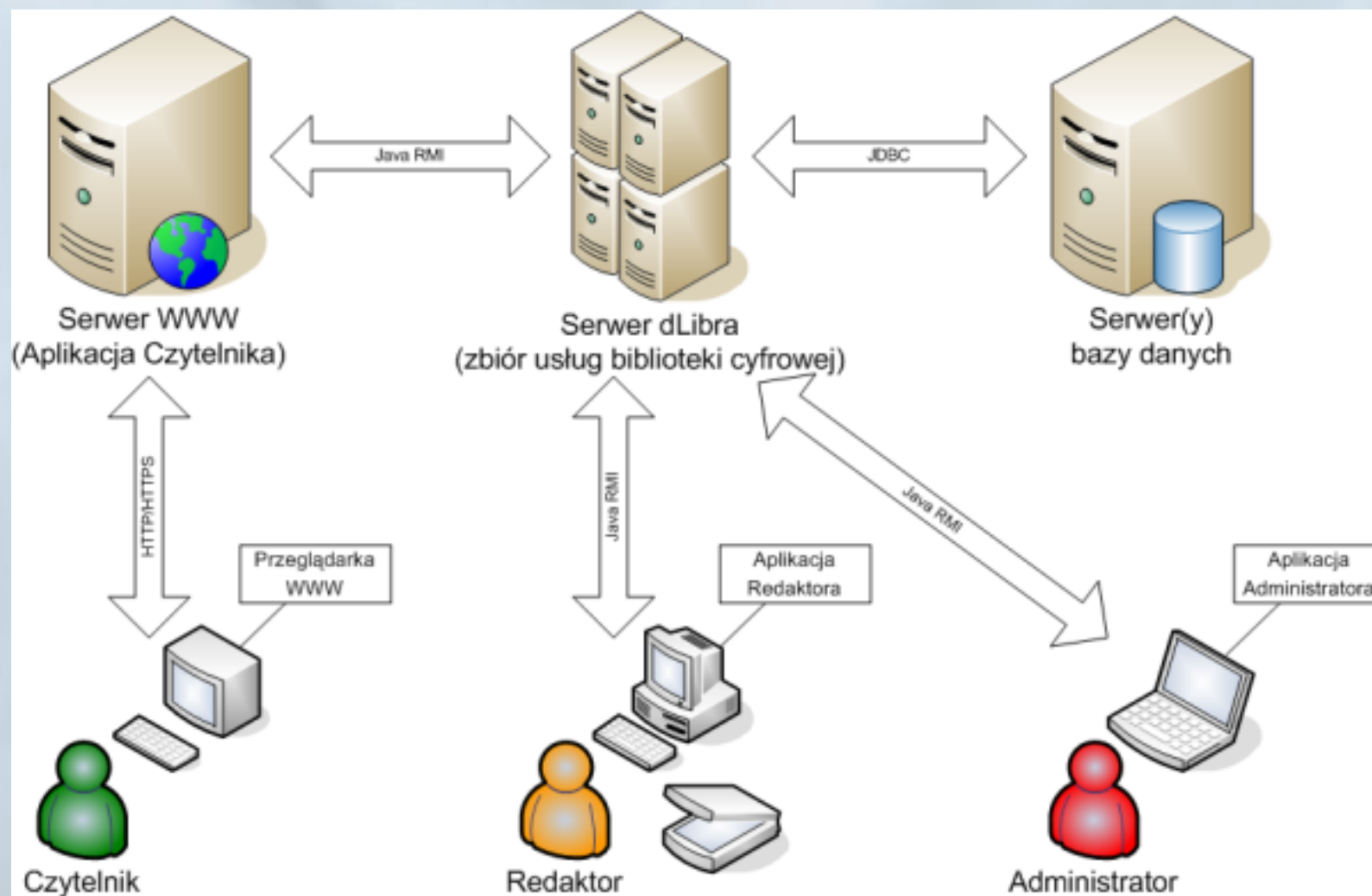
Marcin Mielnicki

marcinm@man.poznan.pl

Poznańskie Centrum Superkomputerowo-Sieciowe

22 października 2010, VII Warsztaty „Biblioteki Cyfrowe”, Poznań

Architektura systemu dLibra



Usługi serwera dLibra

Funkcjonalne

- Content Server (CS)
- Index Server! (IS)
- Message Server! (MS)
- Metadata Server (ME)
- Profile Provider (PP)
- Search Server (SE)
- User Server (US)

Wspierające

- Event Manager (EM)
- System Services (SS)
- dLibra JMX Management Service (MX)



Podstawowa konfiguracja

- Na jednym serwerze
 - Aplikacja Czytelnika – Tomcat
 - Serwer dLibry z wszystkimi usługami
 - Baza danych
- Stosowana w
 - większości mniejszych bibliotek cyfrowych
 - oraz w Kujawsko-Pomorskiej Bibliotece Cyfrowej, trzeciej w Polsce pod względem liczby publikacji

Metody monitorowania

- Narzędzia systemu operacyjnego
 - Linux: top, ps, ...
 - Windows: Menadżer zadań
- Logi
- Panel administracyjny serwera (JMX)

Logi

Apache log4j – możliwości

- Identyfikacja źródła (usługa, zadanie)
- Poziom ważności wiadomości (np. INFO, WARN, ERROR)
- Miejsce zapisu logów (plik, konsola, e-mail)

Logi

Apache log4j – serwer dLibra

- wrapper.log – w formacie tekstowym
- server.log – w formacie XML
- error-log.xml – tylko błędy
- informacje na temat konkretnego zadania np.
consistency-check-log.xml, index-backup-log.xml
- query.log – zapytania wyszukiwawcze czytelników

Panel administracyjny – JMX

- Standardowe komponenty zarządzania udostępniane przez środowisko wirtualnej maszyny Java
 - Poziom użycia pamięci, procesora, wątków i klas
- Dedykowane komponenty zarządzania serwerem
 - Event Manager – EventMonitoringBean
 - NoOfEveEvents
 - EventProcessingDelay
 - OldestEventDate

Aplikacja czytelnika

- Tomcat Monitor – Server Status
 - Czas obsługi żądań HTTP
- Panel administracyjny – statystyki mechanizmu WWW
 - Liczba wywołań i średni czas generowania stron i komponentów oraz wykonywania akcji
- Statystyki
 - Liczba wszystkich publikacji
 - Liczba nowych publikacji
 - Liczba wyszukiwań
 - Liczba wyświetlonych publikacji

Baza danych

APPARENT DEADLOCK w PostgreSQL

```
INFO | jvm 7 | 2010/09/07 02:44:49 | WARN:  
com.mchange.v2.async.ThreadPoolAsynchronousRunner$DeadlockDetector@1322e40 -- APPARENT DEADLOCK!!! Creating  
emergency threads for unassigned pending tasks!  
  
INFO | jvm 7 | 2010/09/07 02:44:49 | 2010.09.07 02:44:49 [Timer-0] ThreadPoolAsynchronousRunner  
  
INFO | jvm 7 | 2010/09/07 02:44:49 | WARN:  
com.mchange.v2.async.ThreadPoolAsynchronousRunner$DeadlockDetector@1322e40 -- APPARENT DEADLOCK!!!  
Complete Status: [num_managed_threads: 6, num_active: 6; activeTasks:  
com.mchange.v2.resourcepool.BasicResourcePool$5@66bd2d  
(com.mchange.v2.async.ThreadPoolAsynchronousRunner$PoolThread-#1),  
com.mchange.v2.resourcepool.BasicResourcePool$5@10de0d8  
(com.mchange.v2.async.ThreadPoolAsynchronousRunner$PoolThread-#2),  
com.mchange.v2.resourcepool.BasicResourcePool$5@249f17  
(com.mchange.v2.async.ThreadPoolAsynchronousRunner$PoolThread-#0),  
com.mchange.v2.resourcepool.BasicResourcePool$5@48ca29  
(com.mchange.v2.async.ThreadPoolAsynchronousRunner$PoolThread-#4),  
com.mchange.v2.resourcepool.BasicResourcePool$5@d06282  
(com.mchange.v2.async.ThreadPoolAsynchronousRunner$PoolThread-#5),  
com.mchange.v2.resourcepool.BasicResourcePool$5@f383cc  
(com.mchange.v2.async.ThreadPoolAsynchronousRunner$PoolThread-#3); pendingTasks:  
com.mchange.v2.c3p0.stmt.GooGooStatementCache$2@151c9a4,
```

Nowe możliwości skalowania w systemie dLibra 5.0

- Dlaczego wprowadziliśmy zmiany?
 - Częste problemy z wyszukiwaniem
 - Wydzielenie usługi SE w wersji 4.0 nie zawsze rozwiązywało problem
 - Funkcja indeksowania w ramach usługi SE
- Zakres zmian w wersji 5.0
 - Wydzielenie funkcji indeksowania do nowej usługi – Index Server (IS)

Skalowanie systemu dLibra

dwa serwery

- Serwer 1:
 - Aplikacja Czytelnika (Serwer WWW)
 - Serwer dLibra:
 - CS, MS, ME, PP, SE, US, EM, SS, MX
 - Baza danych
- Serwer 2
 - Serwer dLibra:
 - Index Server (IS)
 - dLibra JMX Management Service (MX)

Konfiguracja usług serwera

- [dlibra-server]/conf/server.xml
 - Adres serwera
 - Adres usługi System Services
 - Lista usług, które będą uruchamiane w ramach danego serwera
- Baza danych – tabela SYS_SERVICES
 - Lista wszystkich usług dostępnych na wszystkich serwerach

Przykład – przeniesienie usługi

- 1) Wyłącz system dLibra (Tomcat, serwer)
- 2) Skopiuj na nowy serwer pliki aplikacji serwera systemu dLibra
- 3) Skopiuj pliki związane z daną usługą (np. pliki obiektów cyfrowych w przypadku usługi CS)
- 4) Zaktualizuj pliki konfiguracyjne (conf/server.xml) na obu serwerach
- 5) Zaktualizuj wpisy w tabeli SYS_SERVICES
- 6) Wgraj nowy plik licencji na nowy serwer

Kolejne usprawnienia

- Przeniesienie serwera WWW z Aplikacją Czytelnika na odrębny serwer
 - [dlibra-webapp]/WEB-INF/conf/services.properties
 - Uaktualnienie adresu usługi PS w tabeli SYS-SERVICES
- Przeniesienie bazy danych na odrębny serwer
 - [dlibra-server]/conf/database.properties

Konfiguracja usług w WBC

- Serwer 1: Aplikacja Czytelnika (Tomcat)
- Serwer 2: SS, UI, ME, MS, PP, US, EM, MX
- Serwer 3: SE, MX
- Serwer 4: IS, MX
- Serwer 5: CS, MX
- Serwer 6+7: Baza danych Oracle



Dziękuję za uwagę!
Pytania?
