

Wykorzystanie
zaawansowanych technologii
BigData (NoSQL, Cloud)
i skalowanie rozwiązań
na przykładzie systemu **Marketing Automation**

Konrad Pawlus
Benhauer, SALESmanago

Intro

- * SALESmanago – pierwszy w polsce i jeden z pierwszych w Europie systemów MA
- * 400 klientów
- * <http://www.salesmanago.pl>



BIG: Marketing Automation

- * monitoring zachowania na stronie WWW
- * monitoring transakcji
- * wysyłka maili
- * profile klientów

BIG DATA

Not so BIG: Marketing Automation

- * CRM: kontakty, grupy
- * reguły automatyzacji
- * szablony maili e.t.c.

Not so BIG: SQL

- * Sort, find, change...
- * Skala → wiele baz danych
- * Narzędzia, wiedza, ludzie...

BIG: NoSQL

- * DynamoDB
- * MongoDB
- * HBase

MongoDB

- * C++
- * zorientowana na dokumenty
- * indeksy
- * dane często zmieniające się
(zapis)
- * dane nie liniowe

MongoDB – kto?

- * Disney: repozytorium informacji o stanie gier,
- * Git Hub: wewnętrzne aplikacje raportujące,
- * Craigslist, Foursquare,

HBase

- * Java
- * Kolumny
- * Hadoop + Map/Reduce
- * Analiza długich zbiorów danych (dane liniowe)
- * losowy zapis/odczyt

Hbase – kto?

- * Facebook: wiadomości real-time
- * Adobe: wewnętrzne raportowanie i analizy,

DynamoDB

- * MODEL: HASH, RANGE
- * 3 kliki = tabela
- * T1 zapytanie API = tabela
- * Serwer? OS? Problem z dyskiem? A co to?

DynamoDB

- * Java
- * Integracja z Amazon S3
- * Hive (import, eksport)
- * Elastic MapReduce
- * Dodatkowe indeksy
- * Max 64kB/element

DynamoDB – kto?

- * AdRoll – reklama,
- * IMDb – ranking filmów,
- * SmugMug – archiwizacja zdjęć + „social”,
- * Elsevier – informacje dla naukowców,

To Cloud or not to Cloud

- * APP - tak
- * DB – nie bardzo
- * NoSQL – bardzo nie bardzo

- * SAN
- * CPU

Pytania?

Dziękuję za uwagę

konrad.pawlus@benhauer.pl

<http://www.linkedin.com/in/konradpawlus>