

# NoSQL – Not Only SQL, CouchDB

Apache CouchDB has started. Time to relax.

Igor Wojnicki



Katedra Informatyki Stosowanej, Akademia Górniczo-Hutnicza w Krakowie

29 października 2014



# CouchDB

- [www.couchdb.org](http://www.couchdb.org)
- Apache Foundation
- Erlang

I. Wojnicki, NoSQL

# WWW?

*Django may be built for the Web, but CouchDB is built of the Web. I've never seen software that so completely embraces the philosophies behind HTTP. CouchDB makes Django look old-school in the same way that Django makes ASP look outdated.*

— Kaplan-Moss, Django developer

# Charakterystyka

- HTTP
- Dokument
- Brak schematów
- JSON
- Replikacja
- Multi-Version Concurrency Control (MVCC)
- Bezstanowe połączenia

# Komunikacja z bazą

Dokumentacja API znajduje się pod adresem:

[http://wiki.apache.org/couchdb/HTTP\\_Document\\_API](http://wiki.apache.org/couchdb/HTTP_Document_API)

Przykładowa sesja z wykorzystaniem polecenia curl:

```
$ curl -X GET http://awing.kis.agh.edu.pl:5984/
```

```
{"couchdb":"Welcome","version":"1.0.1"}
```

```
$ curl -X GET http://awing.kis.agh.edu.pl:5984/_all_dbs
```

```
["_users"]
```

# Tworzenie bazy danych

```
$ curl -X PUT http://awing.kis.agh.edu.pl:5984/ogloszenia  
  
{"ok":true}
```

```
$ curl -X GET http://awing.kis.agh.edu.pl:5984/_all_dbs  
  
["ogloszenia", "_users"]
```

```
$ curl -X PUT http://awing.kis.agh.edu.pl:5984/ogloszenia  
  
{"error": "file_exists",  
 "reason": "The database could not be created,  
 the file already exists."}
```

# Usuwanie bazy danych

```
$ curl -X PUT http://awing.kis.agh.edu.pl:5984/  
do_skasowania  
  
{"ok":true}  
  
$ curl -X DELETE http://awing.kis.agh.edu.pl:5984/  
do_skasowania  
  
{"ok":true}  
  
$ curl -X GET http://awing.kis.agh.edu.pl:5984/  
_all_dbs  
  
["ogloszenia", "_users"]
```

# Pobieranie danych

```
$ curl -X GET http://awing.kis.agh.edu.pl:5984/  
ogloszenia/_all_docs  
  
{"total_rows":0,"offset":0,"rows":[]}
```

I. Wojnicki, NoSQL



# Wstawianie danych, nowy dokument

```
$ curl -X POST http://awing.kis.agh.edu.pl:5984/  
          ogloszenia  
  -H "Content-Type: application/json"  
  -d  
  '{"kierunek":"PAW" , "tresc":"zajecia odwolane"}'  
  
{ "ok":true,  
  "id":"5474fe40a47e9e554f03bb2def74f425",  
  "rev":"1-952bd6fd9cf92592c0370f2c30713dd2" }
```

**Revision:** kolejny\_numer-MD5 (dokument)

# Wstawianie danych, zawartość bazy

```
$ curl -X GET http://awing.kis.agh.edu.pl:5984/ogloszenia/_all_docs

{"total_rows":1,"offset":0,"rows":[
{"id":"5474fe40a47e9e554f03bb2def74f425",
 "key":"5474fe40a47e9e554f03bb2def74f425",
 "value":{"rev":"1-952bd6fd9cf92592c0370f2c30713dd2"}}
]}

$ curl -X GET http://awing.kis.agh.edu.pl:5984/ogloszenia/5474fe40a47e9e554f03bb2def74f425

{"_id":"5474fe40a47e9e554f03bb2def74f425",
 "_rev":"1-952bd6fd9cf92592c0370f2c30713dd2",
 "kierunek":"PAW","tresc":"zajecia odwolane"}
```

# Modyfikacja dokumentu

Dodanie pola data.

```
$ curl -X PUT http://awing.kis.agh.edu.pl:5984/  
      ogloszenia/  
      5474fe40a47e9e554f03bb2def74f425  
-H "Content-Type: application/json"  
-d '{"_id":"5474fe40a47e9e554f03bb2def74f425",  
    "_rev":"1-952bd6fd9cf92592c0370f2c30713dd2",  
    "data":"2012-03-01"}'  
  
{"ok":true,"id":"5474fe40a47e9e554f03bb2def74f425",  
 "rev":"2-acc9738539da4e04b0af5366131ba0b2"}
```

# Modyfikacja dokumentu, wynik

Dodanie pola data.

```
$ curl -X GET http://awing.kis.agh.edu.pl:5984/  
ogloszenia/  
5474fe40a47e9e554f03bb2def74f425
```

```
{"_id": "5474fe40a47e9e554f03bb2def74f425",  
  "_rev": "2-acc9738539da4e04b0af5366131ba0b2",  
  "data": "2012-03-01"}
```

Ooops...

# Wersje

```
$ curl -X GET http://awing.kis.agh.edu.pl:5984/  
ogloszenia/  
5474fe40a47e9e554f03bb2def74f425?revs_info=true
```

```
{ "_id": "5474fe40a47e9e554f03bb2def74f425",  
  "_rev": "2-acc9738539da4e04b0af5366131ba0b2",  
  "data": "2012-03-01",  
  "_revs_info": [  
    { "rev": "2-acc9738539da4e04b0af5366131ba0b2",  
      "status": "available" },  
    { "rev": "1-952bd6fd9cf92592c0370f2c30713dd2",  
      "status": "available" }  
  ]  
}
```

# Dostęp do poprzednich wersji

```
curl -X GET http://awing.kis.agh.edu.pl:5984/  
ogloszenia/5474fe40a47e9e554f03bb2def74f425  
?rev=1-952bd6fd9cf92592c0370f2c30713dd2
```

```
{"_id": "5474fe40a47e9e554f03bb2def74f425",  
 "_rev": "1-952bd6fd9cf92592c0370f2c30713dd2",  
 "kierunek": "PAW", "tresc": "zajecia odwolane"}
```

## Edycja, cd.

```
$ curl -X PUT http://awing.kis.agh.edu.pl:5984/
ogloszenia/5474fe40a47e9e554f03bb2def74f425
-H "Content-Type: application/json"
-d '{ "_id": "5474fe40a47e9e554f03bb2def74f425",
      "_rev": "2-acc9738539da4e04b0af5366131ba0b2",
      "kierunek": "PAW",
      "tresc": "zajecia odwolane",
      "data": "2012-03-01" }'

{"ok": true, "id": "5474fe40a47e9e554f03bb2def74f425",
 "rev": "3-a5d979e8767925bcf7de6c27b767ffd7" }
```

## Edycja, cd.

```
$ curl -X GET http://awing.kis.agh.edu.pl:5984/  
ogloszenia/5474fe40a47e9e554f03bb2def74f425?  
revs_info=true  
{ "_id": "5474fe40a47e9e554f03bb2def74f425",  
  "_rev": "3-a5d979e8767925bcf7de6c27b767ffd7",  
  "kierunek": "PAW",  
  "tresc": "zajecia odwolane",  
  "data": "2012-03-01",  
  "_revs_info": [  
    { "rev": "3-a5d979e8767925bcf7de6c27b767ffd7",  
      "status": "available" },  
    { "rev": "2-acc9738539da4e04b0af5366131ba0b2",  
      "status": "available" },  
    { "rev": "1-952bd6fd9cf92592c0370f2c30713dd2",  
      "status": "available" } ] }
```

Wszystko OK.



# Usuwanie dokumentu

```
$ curl -X DELETE http://awing.kis.agh.edu.pl:5984/  
ogloszenia/5474fe40a47e9e554f03bb2def74f425?  
rev=3-a5d979e8767925bcf7de6c27b767ffd7
```

```
{"ok":true,  
"id":"5474fe40a47e9e554f03bb2def74f425",  
"rev":"4-a903837c8891e7d400a91e44d4291e2e"}
```

# Alternatywne tworzenie dokumentów

## Użycie własnych identyfikatorów

```
$ curl -X PUT http://awing.kis.agh.edu.pl:5984/  
          ogloszenia/moje1  
-H "Content-Type: application/json"  
-d '{"kierunek":"PAW",  
    "tresc":"koniec semestru",  
    "data":"2012-04-01"}'  
  
{"ok":true,  
 "id":"moje1",  
 "rev":"1-b1805cae2fbae27e6bc929a5b16672b1"}
```

# Alternatywne tworzenie dokumentów, sprawdzenie

```
$ curl -X GET http://awing.kis.agh.edu.pl:5984/  
ogloszenia/_all_docs
```

```
{ "total_rows":3, "offset":0, "rows": [  
  { "id":"5474fe40a47e9e554f03bb2def74f42d",  
    "key":"5474fe40a47e9e554f03bb2def74f42d",  
    "value":{"rev":"2-8fafd9f0ee6bbcdf1753d5e310f02c75"}},  
  { "id":"5474fe40a47e9e554f03bb2def750eaa",  
    "key":"5474fe40a47e9e554f03bb2def750eaa",  
    "value":{"rev":"1-02adafaebe85d80139c3dde97e47b367"}},  
  { "id":"mojel",  
    "key":"mojel",  
    "value":{"rev":"1-b1805cae2fbae27e6bc929a5b16672b1"}}  
]}
```

# Generacja UUID

## Universally Unique Identifier

```
$ curl -X GET http://awing.kis.agh.edu.pl:5984/_uuids
```

```
{"uuids":["5474fe40a47e9e554f03bb2def75401d"]}
```

I. Wojnicki, NoSQL

# Załączniki

```
$ curl -X PUT http://awing.kis.agh.edu.pl:5984/  
  ogloszenia/moje1/  
zdjecie.jpg?rev=1-b1805cae2fbae27e6bc929a5b16672b1  
  --data-binary @zdjecie.jpg  
  -H "Content-Type: image/jpeg"
```

```
{"ok":true,  
  "id":"moje1",  
  "rev":"2-29b17212e3635e445f1a009c4d04e384"}
```

# Załączniki, sprawdzenie

```
$ curl -X GET http://awing.kis.agh.edu.pl:5984/  
ogloszenia/moje1
```

```
{"_id":"moje1",  
  "_rev":"2-29b17212e3635e445f1a009c4d04e384",  
  "kierunek":"PAW",  
  "tresc":"koniec semestru",  
  "data":"2012-04-01",  
  "_attachments":{  
    "zdjecie.jpg":{"content_type":"image/jpg",  
                  "revpos":2,  
                  "length":5240,  
                  "stub":true}}}
```

**Dostępne:** <http://awing.kis.agh.edu.pl:5984/ogloszenia/moje1/zdjecie.jpg>

# Usuwanie załącznika

```
$ curl -X DELETE
http://awing.kis.agh.edu.pl:5984/
  ogloszenia/moje1/
  cos.odt?rev=2-29b17212e3635e445f1a009c4d04e384

{"ok":true,"id":"moja1",
 "rev":"3-a158c2c2a0ae7f31effb0402844515d1"}
```

# Widoki (Views)

- Tymczasowe (temporary): zapytania ad hoc.
  - Tylko dla testów.
- Stałe (permanent): przechowywane w bazie.
- Język: JavaScript.
- Rezultat działania buforowany w bazie danych.
- Sortowanie rosnąco w/g klucza ()
- Malejąco z atrybutem `descending="true"`
- Ograniczenie ilości rezultatów: `limit=5`
- Pominięcie rezultatów: `skip=30`



# Widok, przykład

Argument: dokument, wyjście: emit ()

```
function(doc) {  
  if (doc.kierunek=='PAW') {  
    emit(doc.data, { tresc: doc.tresc,  
                   kierunek: doc.kierunek });  
  }  
}
```

# Widok, przykład, JSON

W pliku: paw.json

```
{ "map" : "  
function(doc) {  
  if (doc.kierunek=='PAW') {  
    emit(doc.data, { tresc: doc.tresc,  
                    kierunek: doc.kierunek });  
  }  
}  
"}
```

# Odpowiedź, widok tymczasowy I

```
$ curl -X POST http://awing.kis.agh.edu.pl:5984/ogloszenia/_temp_view  
-H "Content-Type: application/json"  
-d @paw.json
```

```
{  
  "total_rows":6,"offset":0,"rows":[  
    {  
      "id":"5474fe40a47e9e554f03bb2def755ef5",  
      "key":"2011-10-01",  
      "value":{"tresc":"rozpocz\u0119cie semestru",  
              "kierunek":"PAW"}},  
    {  
      "id":"5474fe40a47e9e554f03bb2def75508f",  
      "key":"2011-11-01",  
      "value":{"tresc":"rozpocz\u0119cie kursu",  
              "kierunek":"PAW"}},  
    {  
      "id":"5474fe40a47e9e554f03bb2def74f42d",  
      "key":"2012-03-20",  
      "value":{"tresc":"zajecia odwolane",
```

## Odpowiedź, widok tymczasowy II

```
      "kierunek": "PAW" }},  
{ "id": "5474fe40a47e9e554f03bb2def750ea",  
  "key": "2012-03-22",  
  "value": { "tresc": "kolowkum zaliczeniowe",  
            "kierunek": "PAW" }},  
{ "id": "5474fe40a47e9e554f03bb2def7545cc",  
  "key": "2012-03-22",  
  "value": { "tresc": "wystawianie ocen",  
            "kierunek": "PAW" }},  
{ "id": "moje1",  
  "key": "2012-04-01",  
  "value": { "tresc": "koniec semestru",  
            "kierunek": "PAW" } }  
]}
```

## Reduce, przykład

```
map:  
function(doc) {  
  if (doc.kierunek='PAW') {  
    emit(doc.data, 1);  
  }  
}  
  
reduce:  
function(key,value) {  
  return value;  
}
```

## Reduce, przykład, wyjście

```
"2012-04-01" [1]
"2012-03-22" [1, 1]
"2012-03-20" [1]
"2011-11-01" [1]
"2011-10-01" [1]
```

I. Wojnicki, NoSQL

# Poprawiona funkcja

```
function(key,value) {  
    return sum(value);  
}
```

## Wyjście:

```
"2012-04-01" 1  
"2012-03-22" 2  
"2012-03-20" 1  
"2011-11-01" 1  
"2011-10-01" 1
```

## Alternatywnie:

```
function(key,value) {  
    n=0;  
    for (var i=0; i<value.length; i++) {  
        n+=value[i];  
    }  
    return n;  
}
```

# Map-Reduce

- map: uruchamiana na każdym dokumencie → klucz, wartość
- reduce: klucz, lista wartości → wartość



- [http://wiki.apache.org/couchdb/HTTP\\_view\\_API](http://wiki.apache.org/couchdb/HTTP_view_API)
- *Design Documents*
- .../baza\_danych/\_design/nazwa

I. Wojnicki, NoSQL

# Dokument definiujący widoki I

```
{
  "_id": "_design/paw",
  "language": "javascript",
  "views":
  {
    "wszystkie": {
      "map": "function(doc) {
        if (doc.kierunek == 'PAW')
          emit(null, doc) }"
    },
    "data": {
      "map": "function(doc) {
        if (doc.kierunek == 'PAW')
          emit(doc.data, doc.tresc) }"
    },
    "ile": {
      "map": "function(doc) {
```

## Dokument definiujący widoki II

```
    if (doc.kierunek == 'PAW')  
      emit(doc.data, 1) }",  
  "reduce": "function(keys, values) {  
    return sum(values) }"  
}  
}  
}
```

# Umieszczenie w bazie

```
curl -X PUT http://awing.kis.agh.edu.pl:5984/  
ogloszenia/_design/paw  
-H "Content-Type: application/json"  
-d @views.json
```

I. Wojnicki, NoSQL

# Używanie widoków I

```
$ curl -X GET http://awing.kis.agh.edu.pl:5984/  
ogloszenia/_design/paw/_view/data
```

```
{ "total_rows": 5, "offset": 0, "rows": [  
  { "id": "5474fe40a47e9e554f03bb2def75508f",  
    "key": "2011-11-01",  
    "value": "rozpocz\u0119cie kursu" },  
  { "id": "5474fe40a47e9e554f03bb2def74f42d",  
    "key": "2012-03-20",  
    "value": "zajecia odwolane" },  
  { "id": "5474fe40a47e9e554f03bb2def750eaa",  
    "key": "2012-03-22",  
    "value": "kolowkum zaliczeniowe" },  
  { "id": "5474fe40a47e9e554f03bb2def7545cc",  
    "key": "2012-03-22",  
    "value": "wystawianie ocen" },  
  { "id":
```

## Używanie widoków II

```
"moje1", "key": "2012-04-01",  
"value": "koniec semestru"}  
]}
```

I. Wojnicki, NoSQL

## Używanie widoków, cd.

```
$ curl -X GET http://awing.kis.agh.edu.pl:5984/  
ogloszenia/_design/paw/_view/ile
```

```
{"rows": [  
{"key": null, "value": 5}  
]}
```

Ooops, wszystkie dane na wejściu funkcji reduce

## Używanie widoków, cd.

```
$ curl -X GET http://awing.kis.agh.edu.pl:5984/  
ogloszenia/_design/paw/_view/ile?group="true"
```

```
{"rows": [  
{"key": "2011-11-01", "value": 1},  
{"key": "2012-03-20", "value": 1},  
{"key": "2012-03-22", "value": 2},  
{"key": "2012-04-01", "value": 1}  
]}
```



## Używanie widoków, cd.

```
$ curl -X GET http://awing.kis.agh.edu.pl:5984/  
ogloszenia/_design/paw/_view/wszystkie
```

```
{"total_rows":5,"offset":0,"rows": [  
{"id":"5474fe40a47e9e554f03bb2def74f42d",  
 "key":null,  
 "value":{"_id":"5474fe40a47e9e554f03bb2def74f42d",  
          "_rev":"2-8fafd9f0ee6bbcdf1753d5e310f02c75",  
          "kierunek":"PAW",  
          "tresc":"zajecia odwolane",  
          "data":"2012-03-20"}},  
{"id":"5474fe40a47e9e554f03bb2def750eea",  
 "key":null,  
 "value":{"_id":"5474fe40a47e9e554f03bb2def750eea",  
          ...
```

# Dostęp do dokumentów... raz jeszcze I

## Dodatkowe parametry określające zakres wartości klucza:

- key, startkey, endkey, keys, skip, limit.
- [http://wiki.apache.org/couchdb/HTTP\\_view\\_API](http://wiki.apache.org/couchdb/HTTP_view_API)

```
curl -X GET 'http://awing.kis.agh.edu.pl:5984/  
ogloszenia/_design/ogolny/_view/kierunek?key=null'
```

```
curl -X GET 'http://awing.kis.agh.edu.pl:5984/  
ogloszenia/_all_docs?  
key="5474fe40a47e9e554f03bb2def75508f"'
```

```
curl -X GET 'http://awing.kis.agh.edu.pl:5984/  
ogloszenia/_all_docs?  
startkey="5474fe40a47e9e554f03bb2def75508f"'
```

## Dostęp do dokumentów... raz jeszcze II

```
curl -X GET 'http://awing.kis.agh.edu.pl:5984/ogloszenia/_all_docs?startkey="5474fe40a47e9e554f03bb2def75508f"&endkey="5474fe40a47e9e554f03bb2def755ef5"'
```

```
curl -X GET 'http://awing.kis.agh.edu.pl:5984/ogloszenia/_all_docs?limit=2'
```

```
curl -X GET 'http://awing.kis.agh.edu.pl:5984/ogloszenia/_all_docs?key="_design/paw"'
```

# Wiele dokumentow w jednym zapytaniu

`http://wiki.apache.org/couchdb/HTTP_Bulk_Document_API`

I. Wojnicki, NoSQL

# Reduce raz jeszcze

Przykładowa funkcja typu *reduce*.

```
function (key, values, rereduce) {  
    return sum(values);  
}
```

rereduce może przyjmować wartości:

- **false**, wartości z funkcji *map*

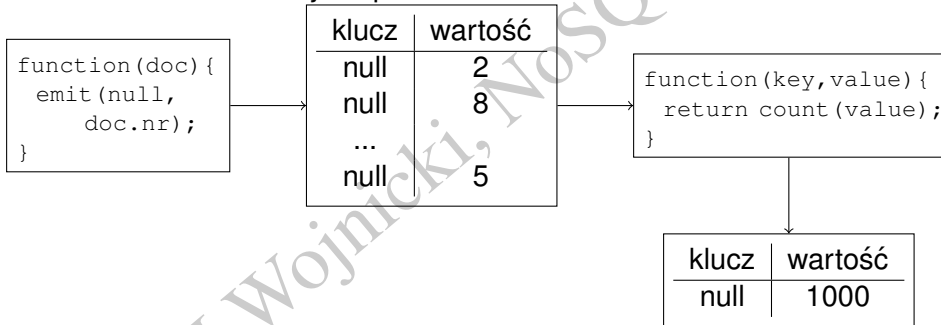
```
reduce([ ... ],  
       [value1,value2,value3],  
       false)
```

- **true**, wartości z funkcji *reduce*

```
reduce(null,  
       [value1,value2,value3],  
       true)
```

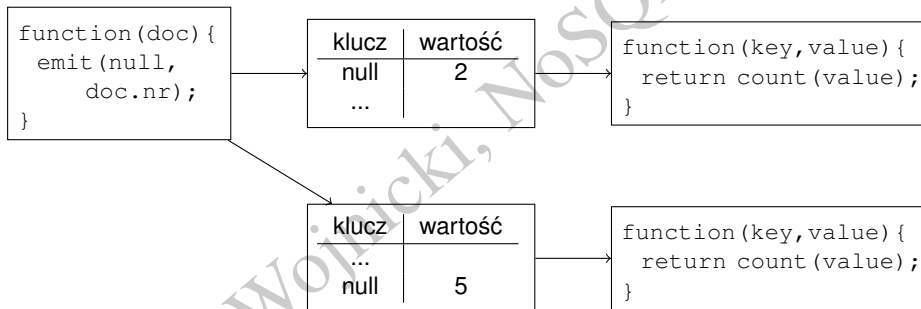
# Reduce, PROBLEM!

1000 rezultatów z funkcji map:

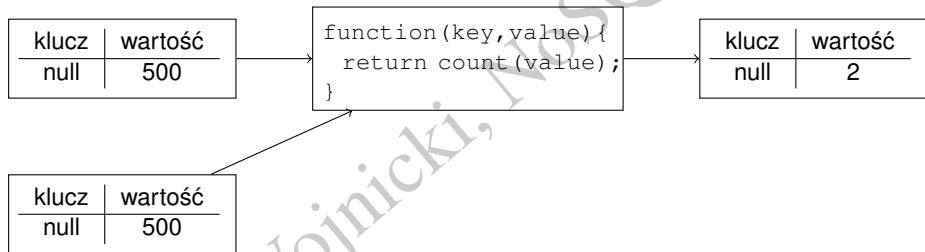


Uwaga: może nie działać!

# Jak na prawdę działa reduce? I



## Jak na prawdę działa reduce? II





# Rerreduce, przykład

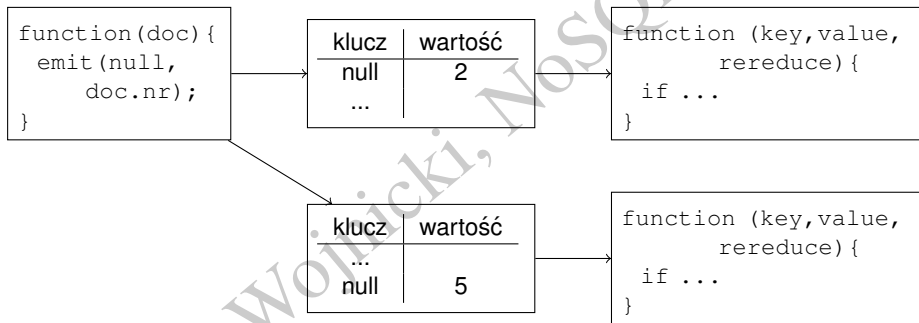
Źle:

```
function (keys, values) {  
  return count(values);  
}
```

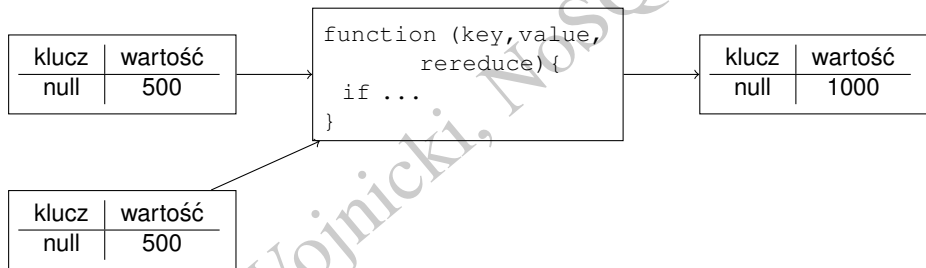
Dobrze:

```
function (keys, values, rerreduce) {  
  if (rerreduce==false) {  
    return count(values);  
  } else {  
    return sum(values);  
  }  
}
```

# Jak na prawdę działa rereduce? I



## Jak na prawdę działa rereduce? II



# Wbudowane funkcje reduce

- `_sum`
- `_count`
- `_stats`

I. Wojnicki, NoSQL

- Interfejs WWW (?!)
- `http://serwer.gdzies:5984/_utils`
- Przy testowaniu map/reduce, należy zrobić reload – dopiero wtedy można włączyć reduce.