



MongoDB for DBA

Petr Medonos, Lukáš Heřbolt

Petr Medonos etnetera

- 6 let v ETN
- dohled nad přiváděním projektů k životu
- DBA, performance, security
- EWA (ewa.etnetera.cz)
- RHCE, M102
- Soyka (soyka.etnetera.cz)

Lukáš Heřbolt



- ATSE @ GSS @ redhat
 - cluster
 - storage
 - filesystem
- ~ 2 roky v ETN
 - projekty
 - HTTP balancing
 - MongoDB
 - pod vedením PMe
- ~ 2 roky na DCGI, FEE, CTU

- co je MongoDB?
- JSON a BSON
- instalace a konfigurace
- provoz
- bezpečnost a zálohování
- HA (Replikace, Sharding)

Co je MongoDB



- 10gen -> mongo
- nestrukturovaná (schema-less) noSQL databáze
- databáze, kolekce, dokument
- JSON (BSON)
- žurnály
- atomické operace nad jedním záznamem
- nativní replikace, sharding

- JavaScript Object Notation

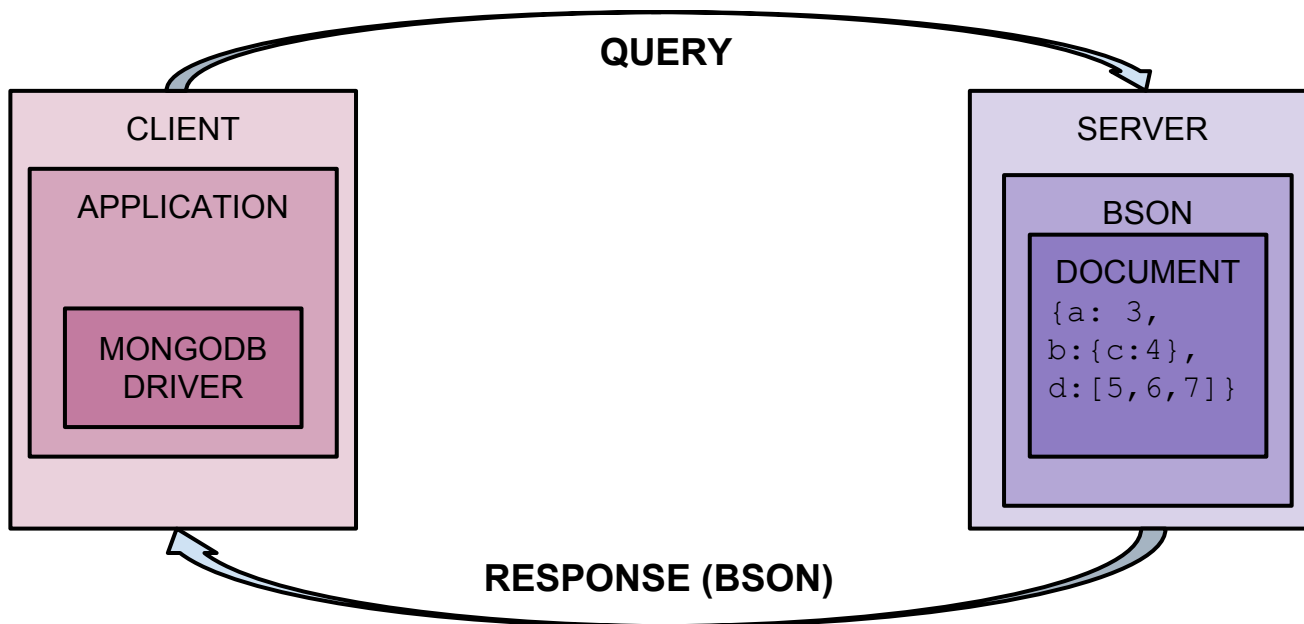
```
{ "data types": {  
  "document": {},  
  "array": ["element1", "element2", "element3"],  
  "bool": <true|false>,  
  "number": 13798465,  
  "string": "Hello Word",  
  "null": null  
}  
}
```

JSON 2 BSON



{ "a": 4, "b": "XYZ" }

length of document 32 bits	type: int32 8 bits	A 16 bits	4 32 bits	type: str 8 bits	B 16 bits	length of string 4	XYZ\0 32 bits	0 8 bits
-------------------------------	-----------------------	--------------	--------------	---------------------	--------------	-----------------------	------------------	-------------



Instalace



- Balíčky:
<http://www.mongodb.org/downloads>
- MacPorts, Homebrew, FreeBSD, ArchLinux, Debian, Fedora, CentOS, Gentoo

Konfigurace



/etc/mongo.conf

port: 27017

http-api: port + 1000

default 28017

rest=false

objcheck since 2.4 true

oplogsize ~ 5% FS size

replSet=<name>

slowms=100ms

Mongo shell



- DB admin
 - system.users
 - administrátorský účet
 - `--setParameter enableLocalhostAuthBypass=0`
- DB local
 - data pro replikaci
 - specifická data pro danou instanci
 - kolekce:
 - startup_log
 - oplog.rs
 - system.replset

CRUD



```
> db.<collection>.find()  
> db.<collection>.findOne()  
> db.<collection>.findAndModify()  
> db.<collection>.insert()  
> db.<collection>.update()  
> db.<collection>.remove()  
> db.getLastError()
```

Indexy



- B-tree, hash, text, geospatial, TTL, unique, ...
- single, compound, multikey
- covering indexes
- background/foreground

```
db.collection.ensureIndex({ orderDate: 1 })  
db.collection.ensureIndex( { a: 1 },  
{ background: true } )
```

Provoz MongoDB



```
> db.currentOp()  
> db.killOp(<opid>)  
> db.<collection>.findOne().explain()  
> db.<collection>.find().hint({date:1})  
> db.setProfilingLevel(1,100)  
  > db.system.profile.find({op:"update"}).pretty()  
> mongostat  
> mongotop
```

- uživatelé jsou uloženi v konkrétní databázi
- kolekce `system.users`
- role:
 - read, readWrite, userAdmin, dbAdmin, clusterAdmin
 - any

```
db.addUser( { user: "<username>", pwd:
"<password>", roles: [
"userAdminAnyDatabase" ] } )
```

Bezpečnost v RS



```
openssl rand -base64 60 > keyfile
```

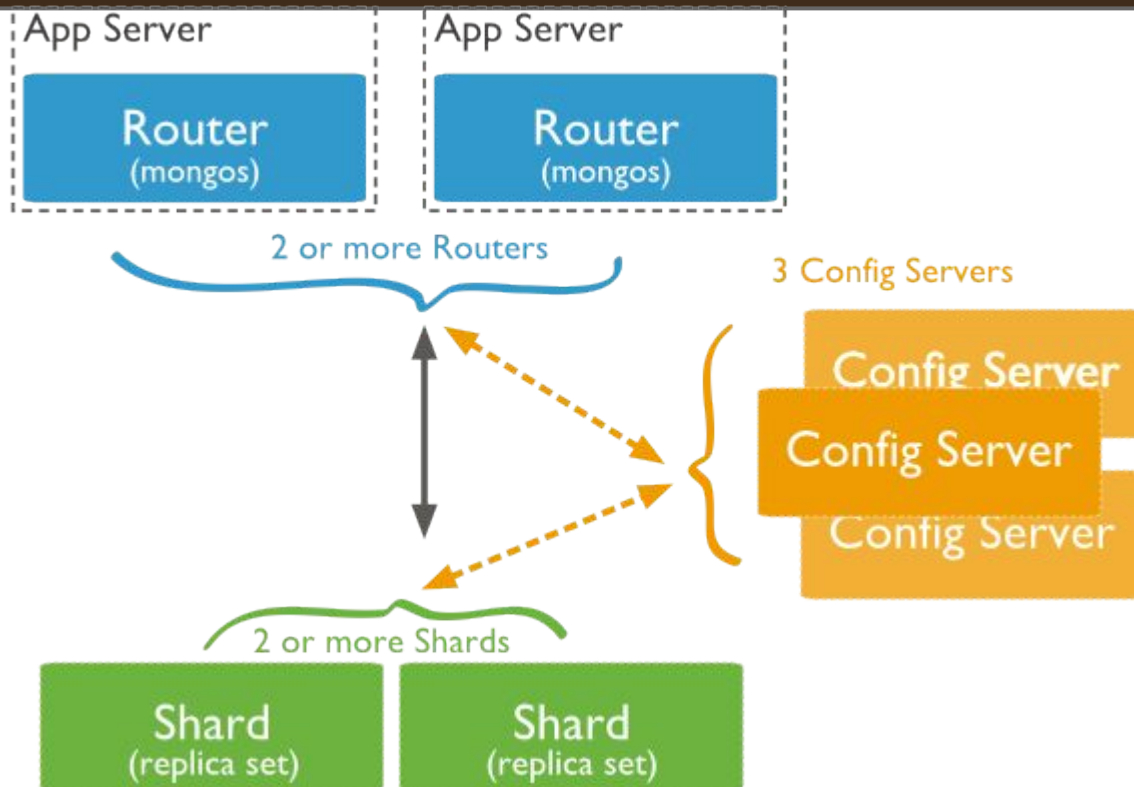
```
mongod --dbpath data --auth --replSet z  
--keyfile keyfile
```

- **Master**
 - v replikasetu vždy jeden
 - read/write operace
- **Slave**
 - volí nového mastera, pokud k tomu nastane situace
 - read only
 - hidden, delayed
- **Arbiter**
 - pouze pro volbu nového masteru
 - bez dat
 - od verze ~2.4.5 nereplikuje system.users collection

Různé uzly můžou mít různou velikost dat na disku

- **až 12 uzlů v RS**
 - maximálně 7 volí
 - asynchronní (cluster-wide commit)
 - “row-based”
- **maximální počet operací v RS**
 - 4,294,967,295/s
 - optime
 - 64bit number
 - higher 32 bits time (unix since epoch)
 - lower 32 bits is operation index within 1 second

Sharding



Sharding



- chunk (64MB, --chunkSize)
- shared key
- range-based vs. hash-based partitioning
- operace:
 - split
 - balancing (migrate)

- **mongodump/mongorestore**
 - replicaset oplog / oplogreplay
 - celé databáze nebo konkrétní kolekce
- **mongoexport/mongoimport**
 - export/import CSV,TSV,JSON
- **snapshots**
 - db.fsyncLock
 - <snapshot/copy db files>
 - db.fsyncUnlock

Díky za pozornost!



Q&A