

# Data In, Data Out



Nuri Halperin

@nurih | [www.plusnconsulting.com](http://www.plusnconsulting.com)

# Stuff Happens!





Oh No!  
Disaster!!!

# Stuff Happens!



“

**Back It Up!**

— Anyone

”

**mongodump** [destination]

- Extracts data from running mongod
- Needs directory with sufficient space

## mongodump

### Default options

- Connect to local mongod running on default port
- Files will be stored in "dump" directory under current working directory

```
mongodump --host [...] --port [...] --out [...]
```

## Explicit options

- **--out** names the directory, not the file.
- Directory-per-dump "2014-10.11-backup"



Long Running Backup

Dump Document "A"

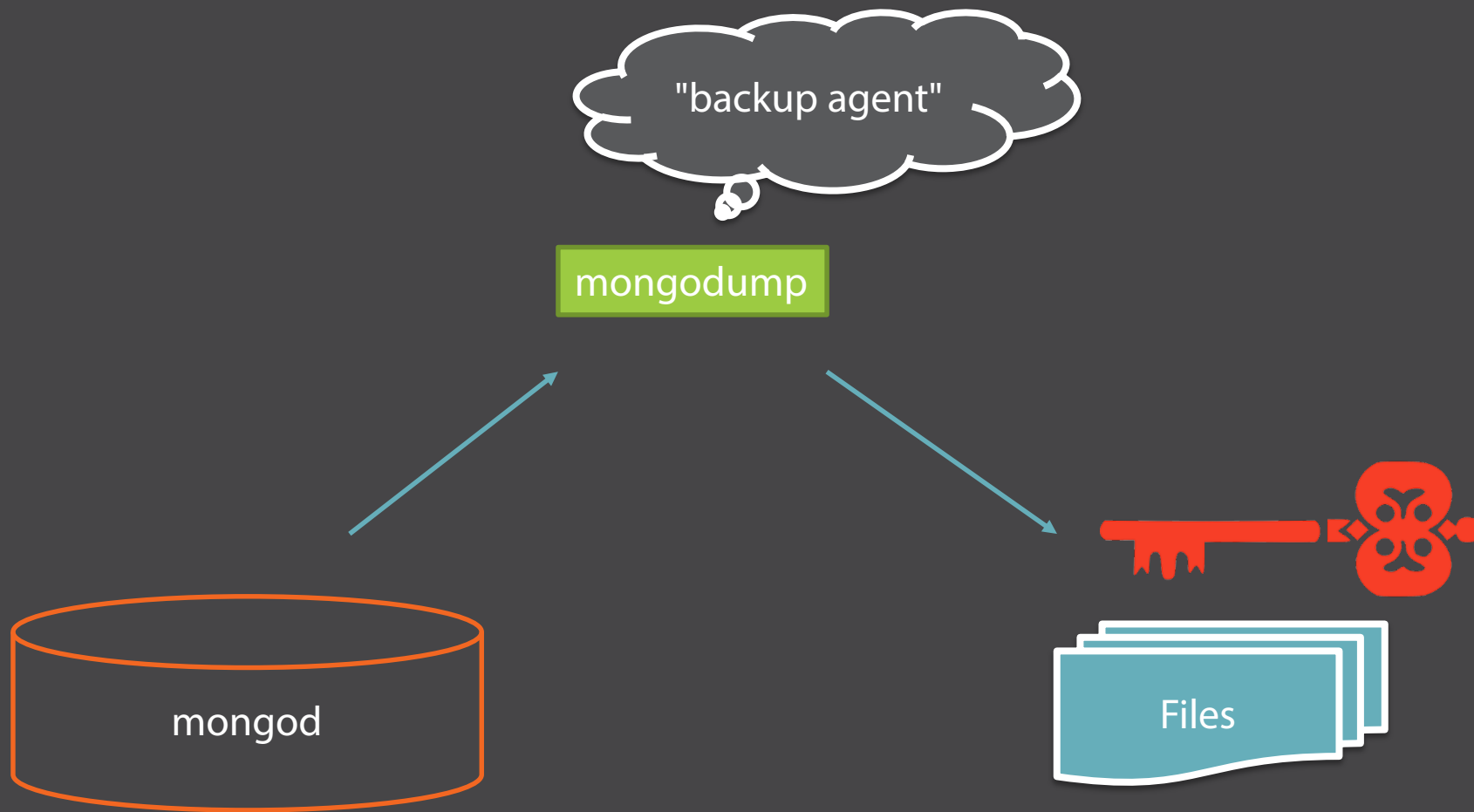
Dump Document "B"

Dump Document "C"

Modify Document A

Other...

Live Data Operations



`mongodump --dbpath`

## Deprecated syntax

- Direct mongo data file dump without mongod is deprecated.
- Must connect to a ***running*** instance of mongo[s].

**mongorestore** --oplogReplay

## Replay operations occurring during backup

- Point in time restore
- Available for replica set dumps created with **--oplog** option

# Practice Restoring Routinely

Master the syntax

Ensure backup file  
integrity

Know the time to  
restore

---

# Bulk data in / out

## mongoimport & mongoexport

---

```
mongoimport --type [json|csv|tsv]
```

Import data from a file into a collection

- Default format JSON

# No Source Id

## Source Document

```
{  
  SKU: 'xyz',  
  Price: 19.99  
}
```

## Mongo Document

```
{  
  _id: 123,  
  SKU: 'xyz',  
  Price: 19.99  
}
```



`--filedFile [file name]`

```
_id  
The Sku  
Unit Price
```



A diagram illustrating the mapping of field names to data rows. A teal box on the left contains the field names: `_id`, `The Sku`, and `Unit Price`. Three green arrows originate from these names: one from `_id`, one from `The Sku`, and one from `Unit Price`. These arrows point to an orange box on the right that contains three rows of data: `1, "A10", 70`, `2, "B20", 60`, and `3, "X30", 30`. The arrows indicate that the first column of data corresponds to `_id`, the second to `The Sku`, and the third to `Unit Price`.

```
1, "A10", 70  
2, "B20", 60  
3, "X30", 30  
...
```

## Field names in a file

- One name per line
- Names can contain spaces (not recommended)

# Summary

