MongoDB Backup Overview

Sam Weaver

Solution Architect, MongoDB

mongoDB

Disasters do happen

U. S. Disaster Avoidance Map

U.S. DISASTER AVOIDANCE MAP

- UNITED STATES SAFETY ZONE
- ICE STORMS
- TORNADOS
- FLOODS
- EARTHQUAKES
- HURRICANES
- POWER OUTAGES
- TSUNAMI
- NUCLEAR POWER PLANTS

U.S. Safety Zone

Las Vegas
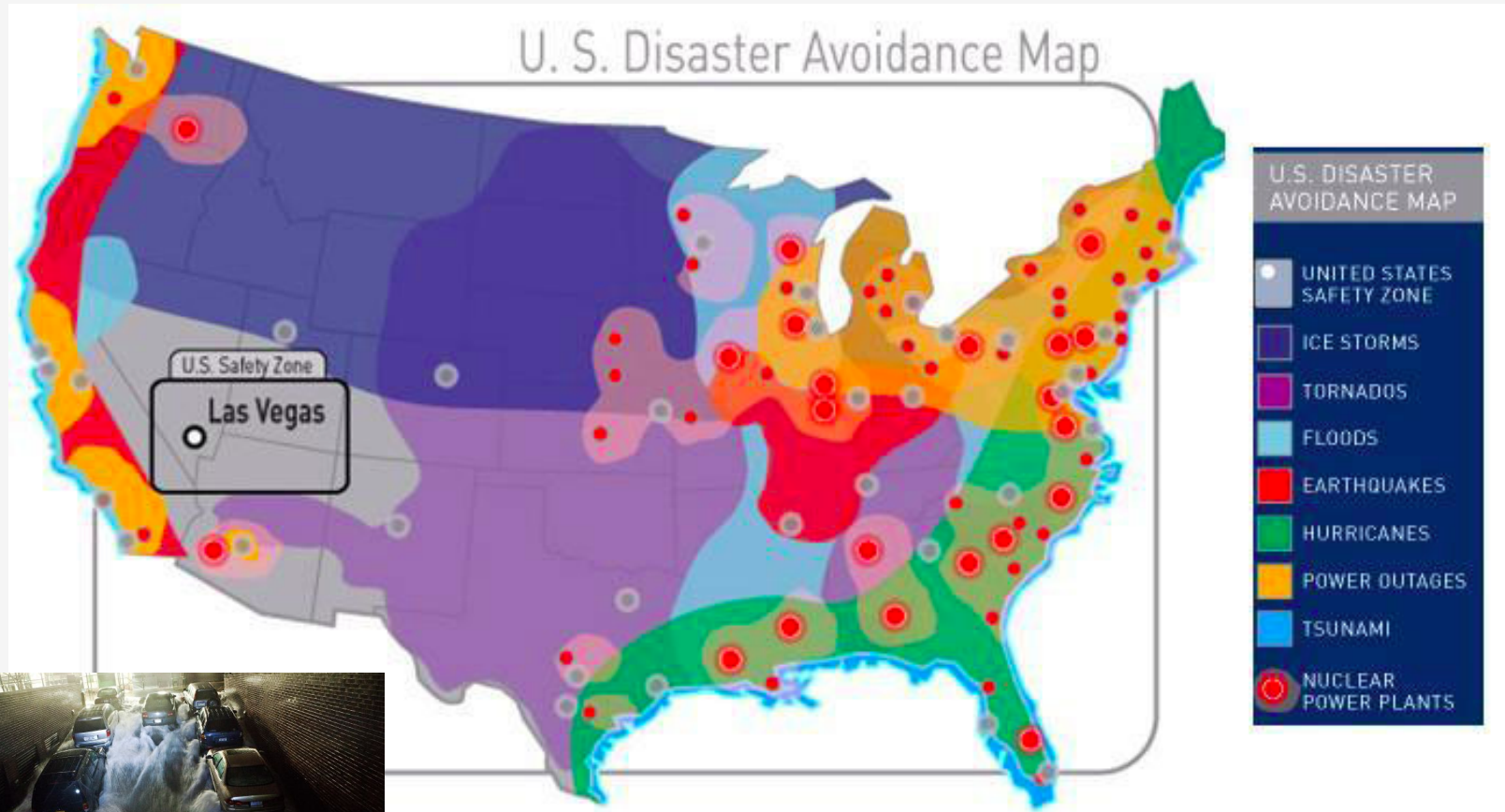
Sometimes they are our fault

Availability and Durability

- Risks
  - Storage failure
  - Power outage
  - Programmer error
  - Hardware failure
  - Data centre failure
  - Cyber attack

- Storage level options

- MongoDB Backup Service

mongoDB

Addressing Risks

- Relative to any particular risk
  - How much data can you afford to lose? (RPO)
  - How long can you afford to be offline? (RTO)
  - What price are you willing to pay to remove the risk?

- Solutions
  - Replication
  - Application engineering?
  - Backup

mongoDB

Replication

- Built into MongoDB, only ops and infrastructure cost

- Very little to zero data loss in failure

- Down for a very short interval

- BUT… programmer errors will replicate almost instantly

mongoDB

Application Engineering

- Flexible

- Expensive, time and money

- Good use of engineer?

mongoDB

Backup

- Slower to restore

- Can suffer from being out of date

- Fairly cheap

- Isolated

- Covers most risks

mongoDB

Backup Approaches

- Mongodump/mongorestore

- Storage level options

- MongoDB Backup Service

Mongodump/MongoRestore

- Can be run online or offline

- Oplog aware for point in time restores

- Filter in, filter out

- Considerations
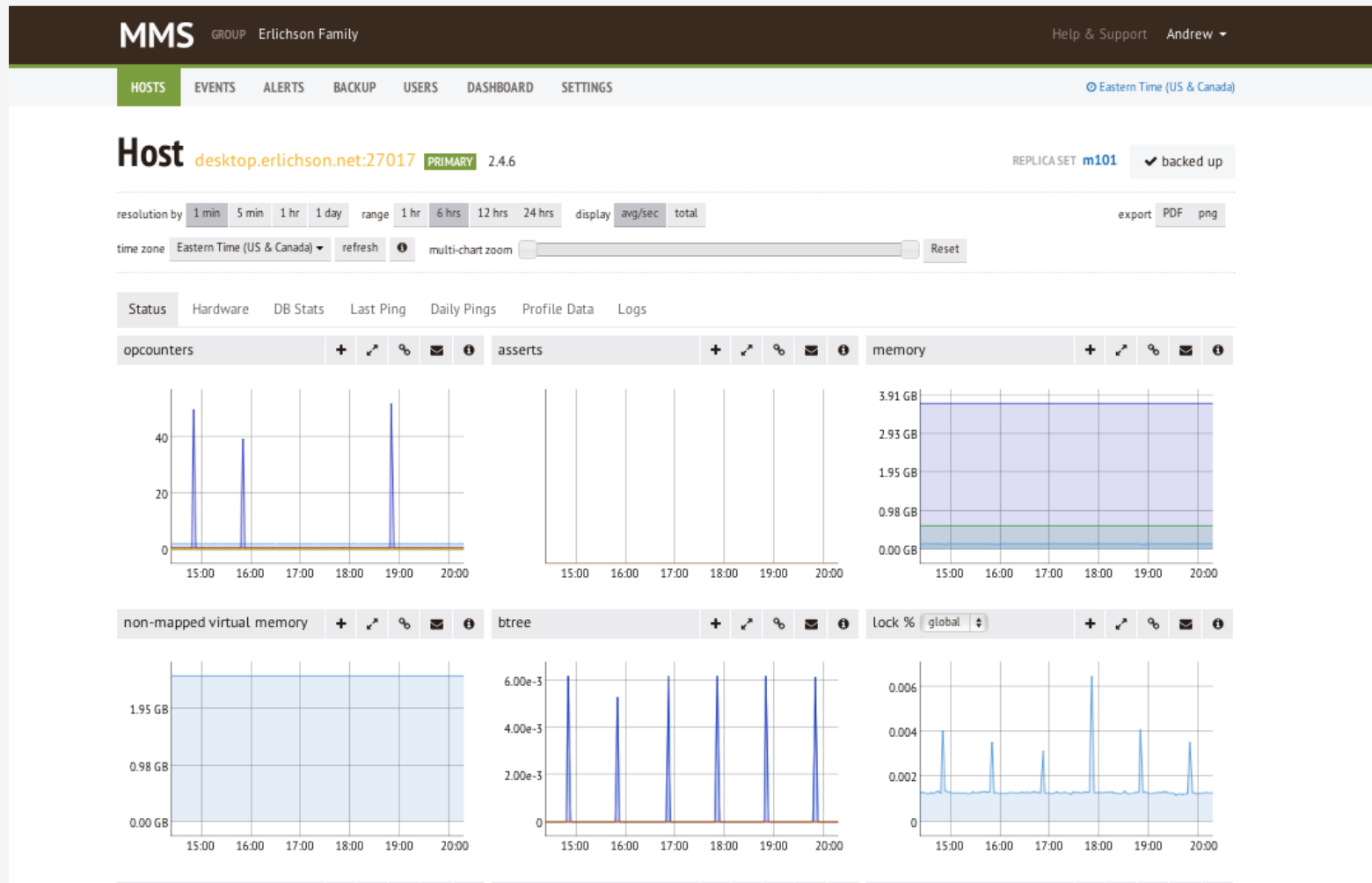 - Data size
  - Sharding

mongodb
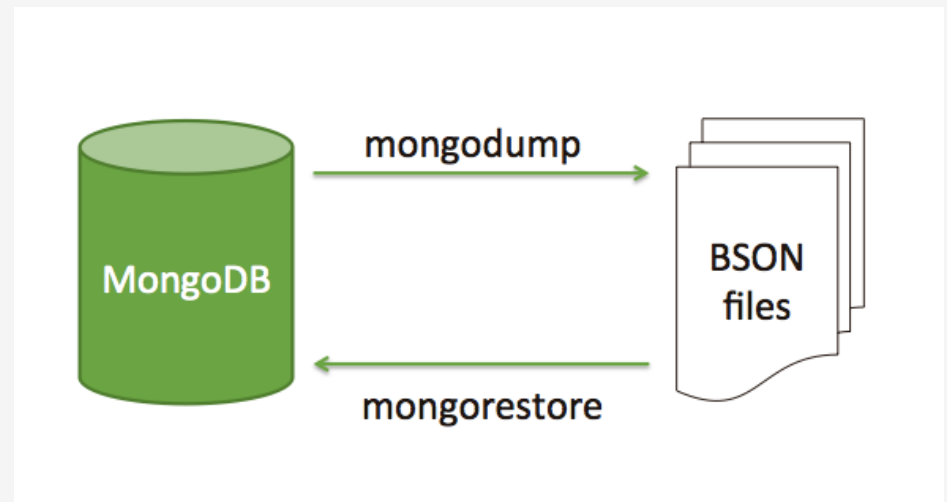
Storage-level Backups

- Copy files in your data directory (e.g. /data/db)

- File system or block storage snapshots

- Fastest way to backup/restore

- Considerations
  - Journal
  - Consistency
  - Backup granularity (whole file system back up?)
  - Ops expertise
  - Storage of snapshots or data file backups
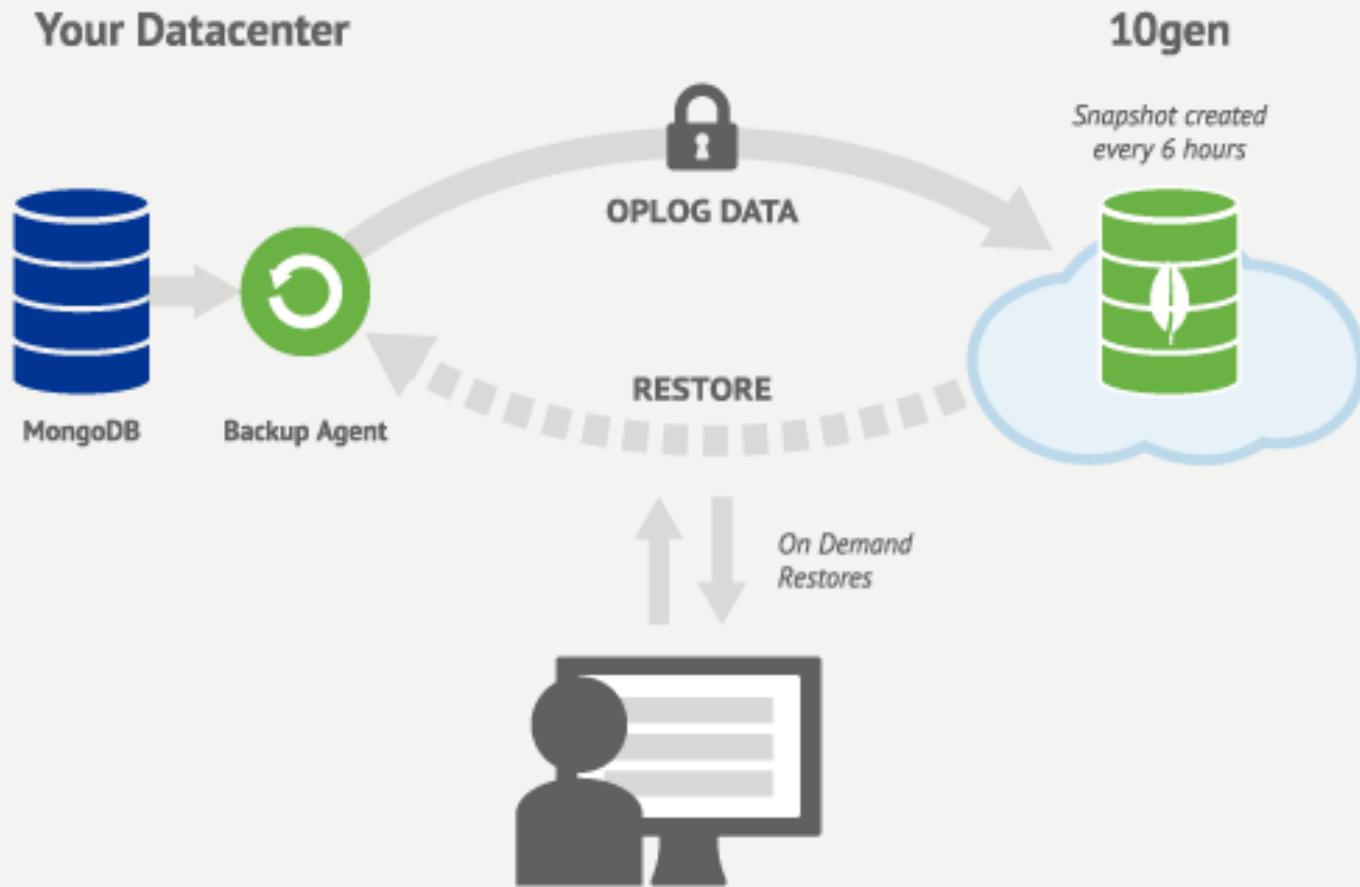
mongoDB

What is MongoDB Management Service (MMS)?

Free monitoring

How it Works

mongoDB

Overview

Your Datacenter

MongoDB → Backup Agent

OPLOG DATA

RESTORE

On Demand Restores

10gen

Snapshot created every 6 hours

mongoDB

Under the Hood

- From the initial sync, we rebuild your data in our datacenters and take a snapshot

- We take snapshots every 6 hours

- Oplog is stored for 48 hours

mongoDB

Sharded Clusters

- Balancer paused every 6 hours

- A no-op token is inserted across all shards, mongoses and config servers

- Oplog applied to replica sets until point in which token was inserted

- Provides a consistent state of database across shards

Recovery Approaches

Recovery Approaches

	Mongodump	File system	MMS Backup
Initial complexity	Medium	High	Low
Confidence in Backups	Medium	Medium	High
Point in time recovery of replica set	Sort of ☺	No	Yes
System Overhead	High	Can be low	Low
Scalable	No	With work	Yes
Consistent Snapshot of Sharded System	Difficult	Difficult	Yes

Key Benefits

Point in Time

- Oplog stored for 48 hours

- Restore your replica set to any point-in-time in the last 48 hours by creating a custom snapshot

Snapshot Policy

- Every 6 hours for 48 hours

- Every day for 1 week

- Every week for for a month

- Every month for a year

Easy to Restore

- Pull from custom URL

- Push via SCP

mongoDB

Unlimited Restores

- Confidence in your restore process

- Build development, QA, analytics environments without impacting production

Fully Managed

- Created by the engineers that build MongoDB

- No need to write or maintain custom backup scripts

mongoDB

Getting Started

Getting Started

1. Create an account at mms.mongodb.com

2. Install MMS Monitoring on your deployment

3. Register at mms.mongodb.com/backup

4. Install MMS Backup on your deployment

5. Start initial sync

6. Rest easy!

mongoDB

Free Month –Register Using MONGODB_1MONTHFREE

mongoDB

Questions?

Thank You

Sam Weaver

Solution Architect, MongoDB

mongoDB