MongoSoup SensorLab

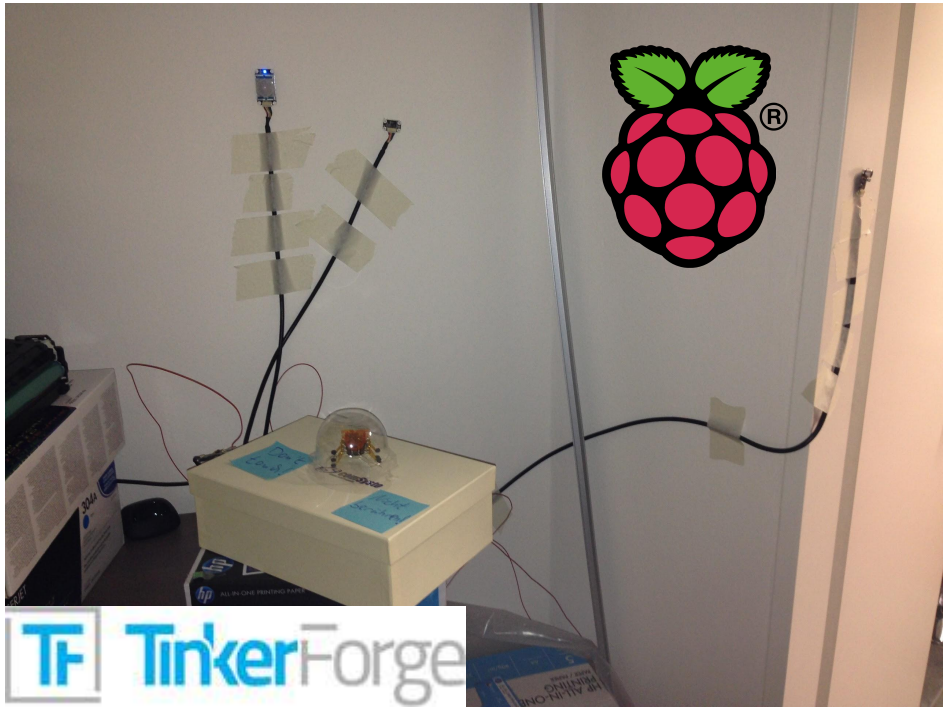
MongoDB IoT Hackathon

Erster deutscher MongoDB-as-a-Service

https://www.mongosoup.de

Physical Setup / Sensors

- Raspberry Pi
- Tinkerforge Sensors
  - Motion
  - Sound Intensity
  - Temperature
  - Multi Touch



Software Stack

- Python:
  - Tinkerforge API on Raspberry
  - Flask as Webframework
  - PyMongo
- MongoDB as a Service: MongoSoup
- Heroku (PaaS)
- R, D3

Stream Data

MongoDB data model

- Follows MongoDB proposal for timeseries data
- Nested document structure
- Fits for frequent, periodic data
- Preallocation

Stream Data

Write

```
ipcon = IPConnection() # Create IP connection
ipcon.connect(config.BRICKD_HOST, config.BRICKD_PORT) # Connect to brickd

sound_brick = SoundIntensity(config.UID_sound_intensity, ipcon) # Create device object
sound_brick.set_intensity_callback_period(100)
sound_brick.register_callback(sound_brick.CALLBACK_INTENSITY, stream_handler.cb_intensity_SI)
```

```
self.collection.update({"start": start_hour.isoformat(),
                        "type": type},
                       {"$set": {"values." + minute + "." + second + "." + tenth: str(value)}})
```

mongoDB

MongoSoup
YOU CODE, WE COOK.

Event Data

MongoDB data model

- “Bucket model”
- Fits for sporadic, sparse data
- Preallocation

```
{
    "_id" : ObjectId("538476f074fece0978e8bc48"),
    "start" : ISODate("2014-05-27T13:22:35.585Z"),
    "end" : ISODate("2014-05-27T13:42:19.608Z"),
    "type" : "MD",
    "values" : [
        {
            "start" : ISODate("2014-05-27T13:28:43.448Z"),
            "length" : 5,
            "end" : ISODate("2014-05-27T13:28:48.861Z")
        },
        ...
        {
            "start" : ISODate("2014-05-27T13:42:09.737Z"),
            "length" : 9,
            "end" : ISODate("2014-05-27T13:42:19.608Z")
        }
    ]
}
```

Event Data

Write

- Callback from Tinkerforge API
- Update at next “undefined” value in array

```
db.SensorEventDataFiltered.update(
    {"start" : ISODate("2014-05-23T06:56:42.909Z"),
     "type": "MD",
     "values":{
            "start" : 0,
            "length" : -1,
            "end" : 0
        }},
    {'$set': {
        "values.$": {
            "start" : ISODate("2014-05-23T10:05:58.700Z"),
            "length" : 832,
            "end" : ISODate("2014-05-23T10:05:59.531Z")
            }
        }
    }
)
```

Stream/Event Data

Read

- Get all documents for a certain sensor and time interval in one query
- Iterate through the result set in memory (different for stream and event data)
- Filter the relevant values, aggregate, etc.

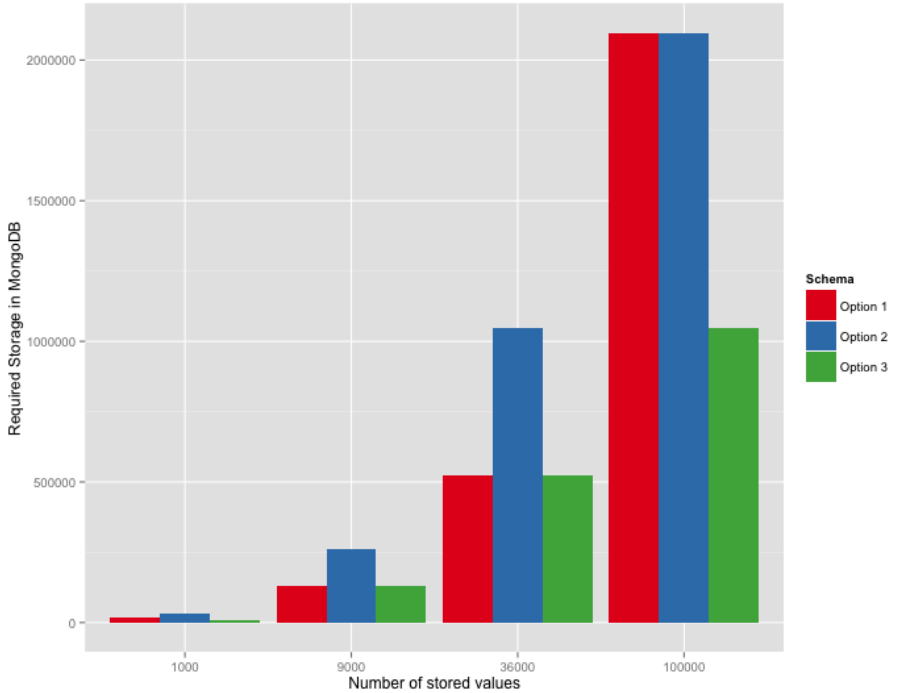
```
query = {
    'type': database_id_for_sensor(sensor_name),
    '$or': [
        {'end': {'$exists': False}},
        {'end': {'$gt': from_time}}],
    'start': {
        '$lt': to_time
    }
}
```

Comparison of data models

1. Plain array of values

2. Array of values with subdocuments
= event data model

3. Nested subdocuments
= streaming data model



Comparison of data models

- Plain arrays are inefficient, due to the JS object they are converted to
- Optimize storage by optimizing the average key chars per data point
- Streaming data model

  -> average of 1.2 chars / dp

  -> document size 433KB for 1hour with 10 data points/s (=36.000 values)

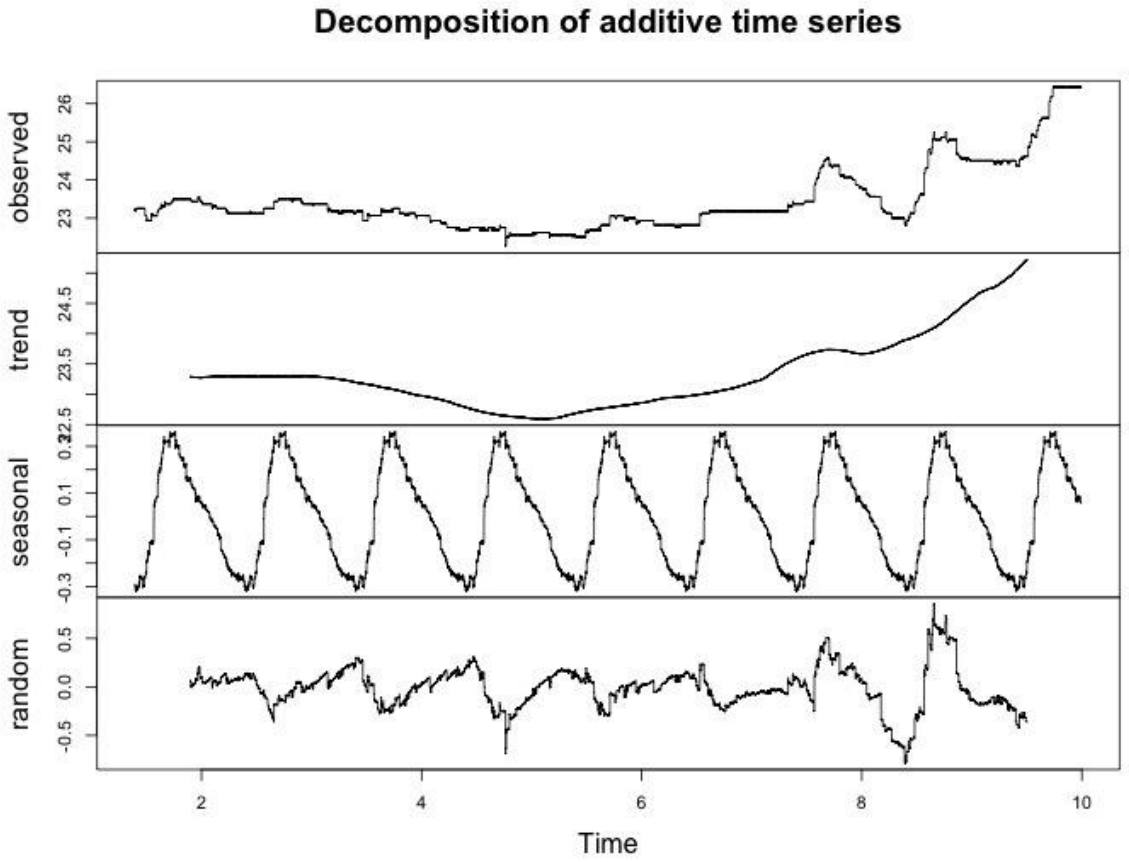
  -> Aggregation within MongoDB very limited (no aggregation on keys)
- Event / bucket data Model:

  -> Limit amount of elements per array -> avoid overly long keys.

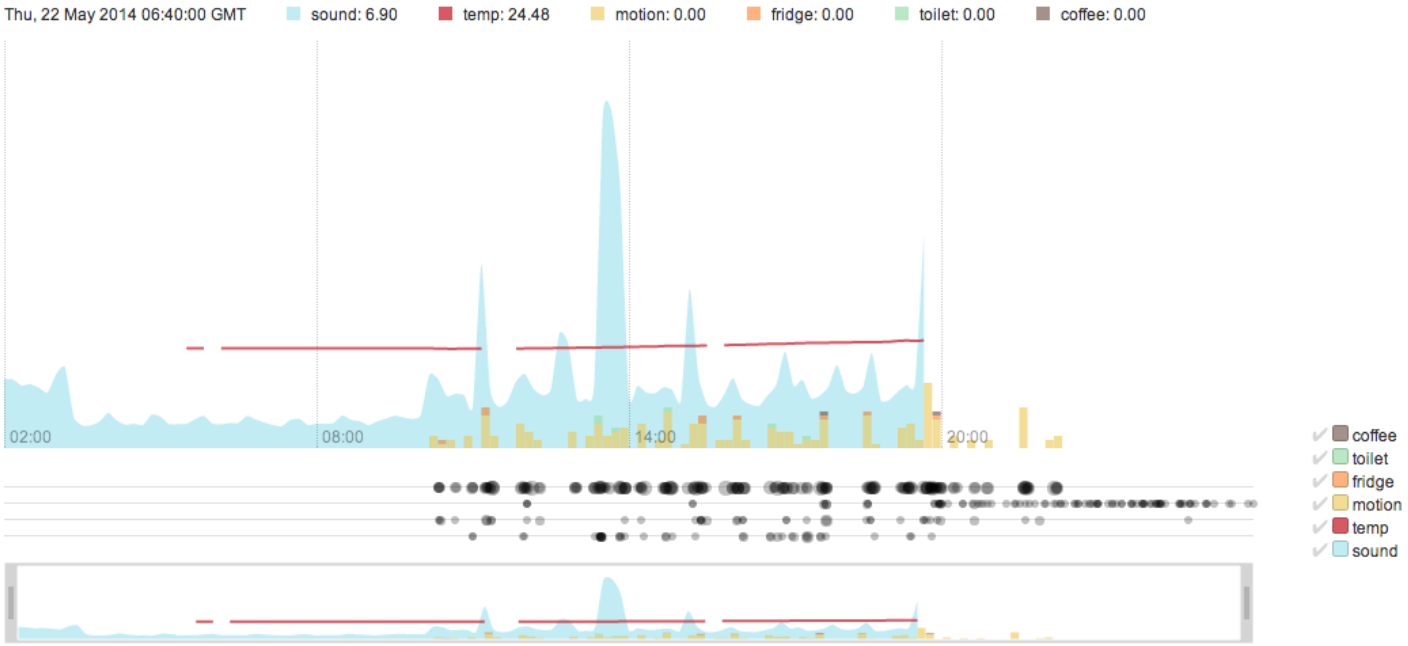
  -> E.g. 100 elements in plain array -> 2 chars / dp

Analytics

Temperature



Visualization



- Realized with Rickshaw (built on d3.js), e.g. http://code.shutterstock.com/rickshaw/examples/multi.html

References

SensorLab members:

Julio Herce, Cindy Lamm, Mario Koppen, Christian Kroemer, Ronny Steinbrenner, Lars Haferkamp

comSysto Blog about SensorLab

- http://blog.comsysto.com/2014/05/16/processing-and-analysing-sensor-data-a-diy-approach/
- http://blog.comsysto.com/2014/07/14/processing-and-analysing-sensor-data-a-diy-approach-part-ii/

MongoSoup Blog about data model comparison

- https://www.mongosoup.de/blog-entry/Storing-Large-Lists-In-MongoDB.html

MongoDB Blog about timeseries data model

- http://blog.mongodb.org/post/65517193370/schema-design-for-time-series-data-in-mongodb