amadeus

MongoDB At Amadeus IT Group

Startup technology and the enterprise

Attila Tozser, Amadeus Data Processing GmbH, NoSQL Nap
Budapest, 20/10/2015

© 2015 Amadeus IT Group SA

Agenda

1. Amadeus IT Group
2. MongoDB at Amadeus
3. Use-cases
4. Future plans

© 2015 Amadeus IT Group SA

amadeus

1 Amadeus IT Group

© 2015 Amadeus IT Group SA

amadeus

Amadeus in a few words

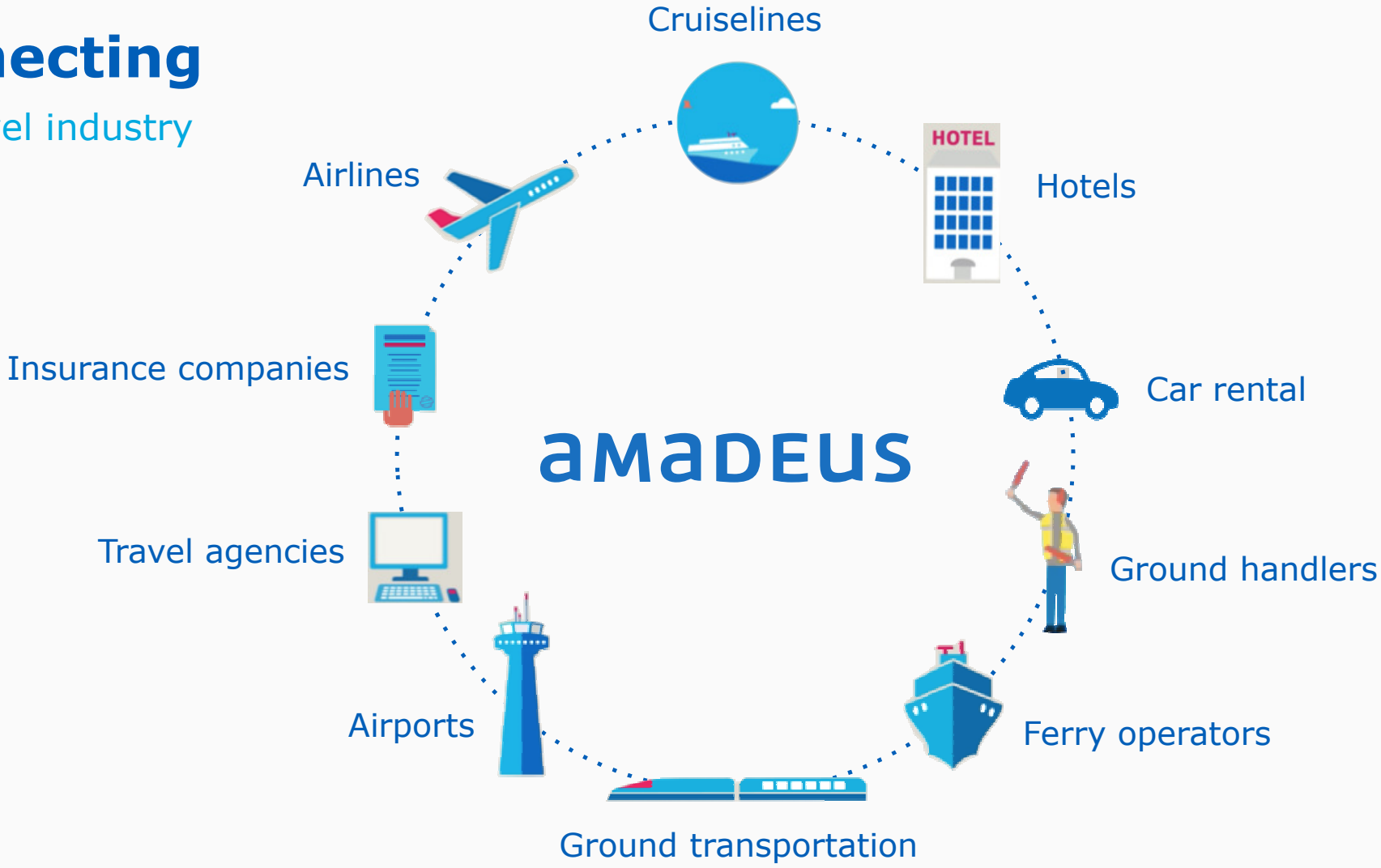
Amadeus is a technology company dedicated to the
global travel industry.

We are present in **195 countries**
with a worldwide team of more than 13,200 people.

Our solutions help improve the
business performance
of travel agencies, corporations, airlines,
airports, hotels, railways and more.

Connecting

The travel industry



amadeus

Supporting

The traveler life cycle

amadeus

We have a clear purpose

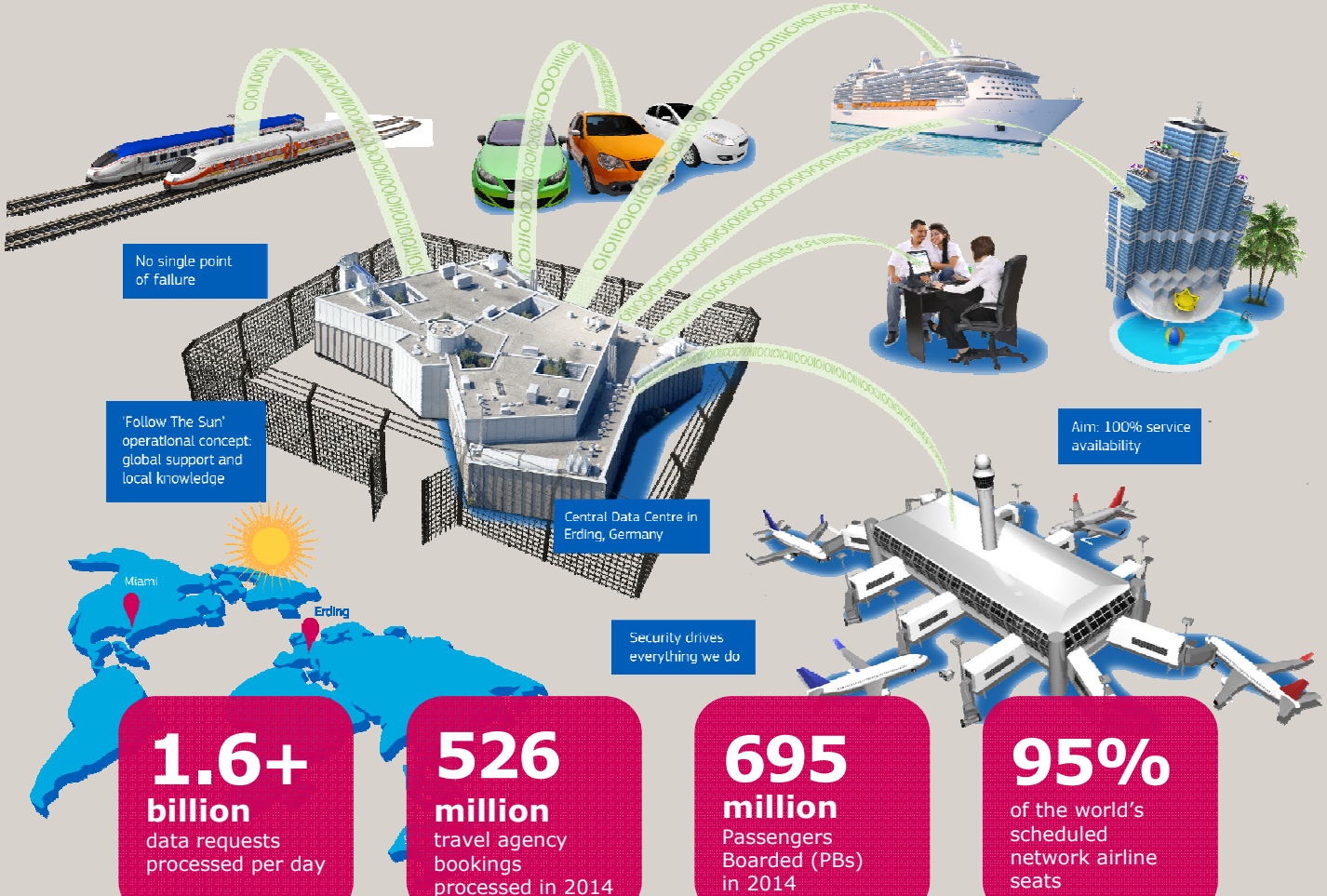
Our solutions facilitate journeys and enrich the travel **experience** for hundreds of millions of people every year.

We are working together with our customers and partners **to shape the future of travel**

Close to our customers

amadeus

Robust global operations



1.6+ billion data requests processed per day

526 million travel agency bookings processed in 2014

695 million Passengers Boarded (PBs) in 2014

95% of the world's scheduled network airline seats

© 2015 Amadeus IT Group SA

amadeus

2 MongoDB at Amadeus

© 2015 Amadeus IT Group SA

amadeus

Business needs

Gathered from customers

user-friendly way to **query any data**
in our main operational databases

- **Unpredictable** ad-hoc search
- Many **advanced reporting** requirements

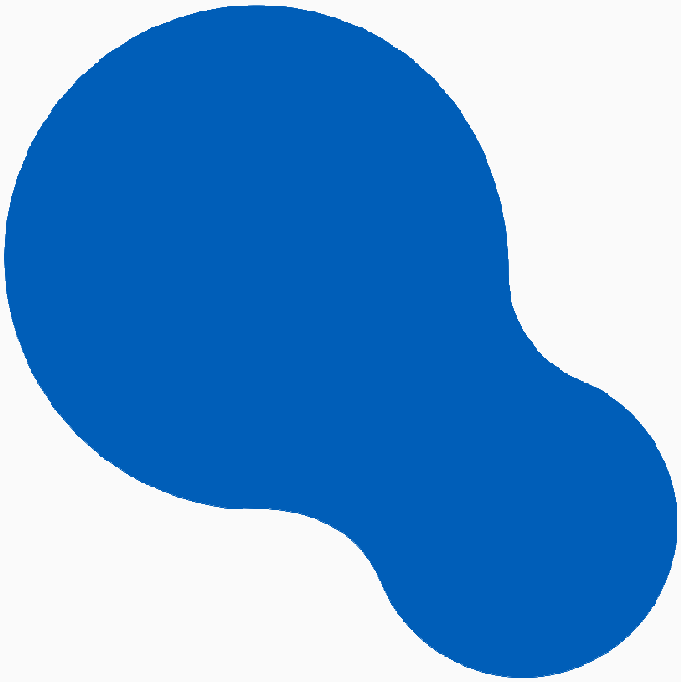
A way to store orders of magnitude more different data from various sources, but still maintain **low latency and high throughput**

Sustain Organic growth of data and throughput (time-to-time 2X or more a year) without degradation

amadeus

Integration

In the Amadeus standards



Runs on **standard x86 architecture**

C++, **Python** & Java **drivers**

Enterprise-grade **security**

- **SSL** encryption
- **Kerberos** authentication
- **Data-at-rest encryption**
- **Advanced tooling for backup and monitoring**
- **Enterprise support**

amadeus

Considered alternatives

To MongoDB

Oracle

Mounting all data in memory is **irrelevant** for **cost** & **hardware** reasons: 90TB for our biggest prospect.

MySQL cluster

Technical & functional **limitations**, **complex** to implement & maintain.

Impala

Still **young**, with a **steep learning curve.**

Distributed data analysis not exactly matching our use-case.

Couchbase

Slightly behind MongoDB for document search (**index mandatory**).

N1QL not finalized.

Key-value store not exactly matching our use-case.

Crescando

Amadeus in-house R&D database engine (**index-less**, **main-memory only**, partitioning data at CPU core level).

Project terminated.

2 First Use-case

© 2015 Amadeus IT Group SA

amadeus

Passenger Revenue Accounting

What for?

Revenue of a flight ticket is **shared**

- Travel agent
- Governments
- Airlines: many can be involved (marketing & operating)

Amadeus Revenue Accounting handles cash flows **on behalf of airlines**

- Tracking
- Error handling & optimisation
- Reporting: analysis & audit

amadeus

Business needs

Gathered from a launch partner

One of our launch partners is a **large European airline**

- transporting **35m+ passengers** a year
- **key player** in the revenue accounting industry

Migrating

- from their **in-house** data warehouse
- to our **cloud-based** solution

They requested a **user-friendly way** to **query any data** in our main operational database

- **Unpredictable** ad-hoc search
- Many **advanced reporting** requirements

amadeus

Metadata Search

The main promises

- **Graphical** user interface
  edit, import, save & share queries
- **Data warehouse**
  fed in **real time**
  4 years history (**500m+ documents**, versioned)
- **Interactive** response times
- Search further using
  chained queries (patent pending)

amadeus

Expecting fast answer

to unpredictable queries

No index, no hint *(almost)*

- Fields to be scanned **unknown**
- **Main-memory full scans** to decrease response time

Need to **scale out**
for **sustainable performances**

Support **mainstream SQL DML** statements

- Aggregation
- Cross-column comparison, Boolean logic
- Sort

amadeus

Project milestones

And possible impacts

November **2013** **User acceptance** testing

December **2014** Migration & parallel running **validation** on **production**

Summer **2015** **Production cut-over**

Post **cut-over** **SLA** & **optimisation** based on **usage statistics**

Any **delay** or **functional gap** may
impact the whole project
as application is used to **validate**
migration and parallel running phases.

amadeus

Production response times

And their lessons learnt

Full scan aggregation is **CPU-bound**, with a **fixed entry cost** for unwinds.

- **no** unwind **3s**
- **unwinds** on 1, 2 or 3 levels **70s**

Interactive response times promise is complied with on **basic use-cases**

In the **absence of concurrency**, response times are **consistent** across all tests.

Indexes have a **linear impact** on response times.

Complex query with 4 match criteria

- **full scan** **100s**
- **index**, 40% selectivity **40s**

Complex query with 4 match criteria, including field-on-field comparison

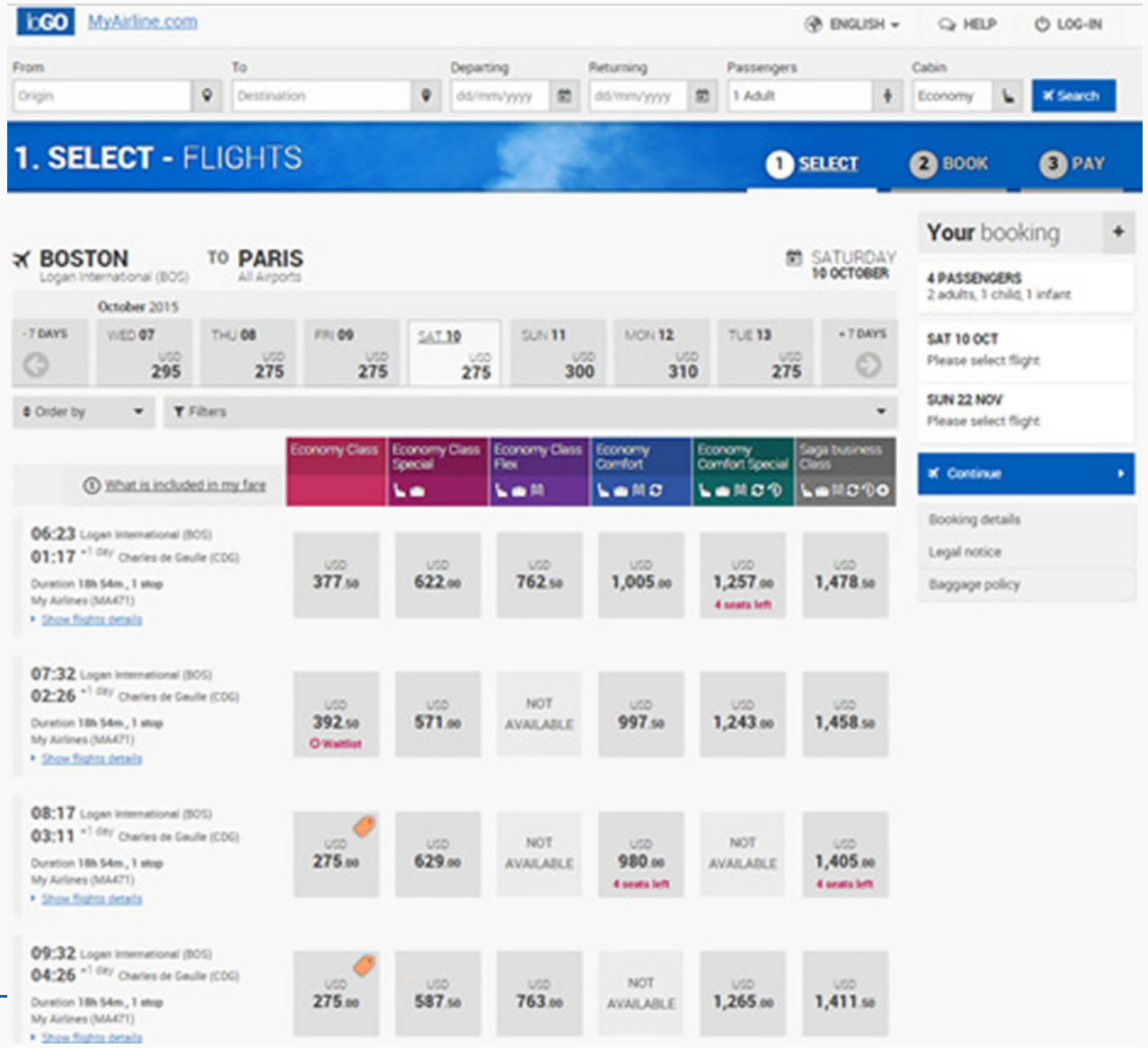
- **full scan** **190s**
- **index**, 40% selectivity **70s**
- **index**, 75% selectivity **145s**

Position of the **match operator** in the aggregation pipeline can **impact index usage**.

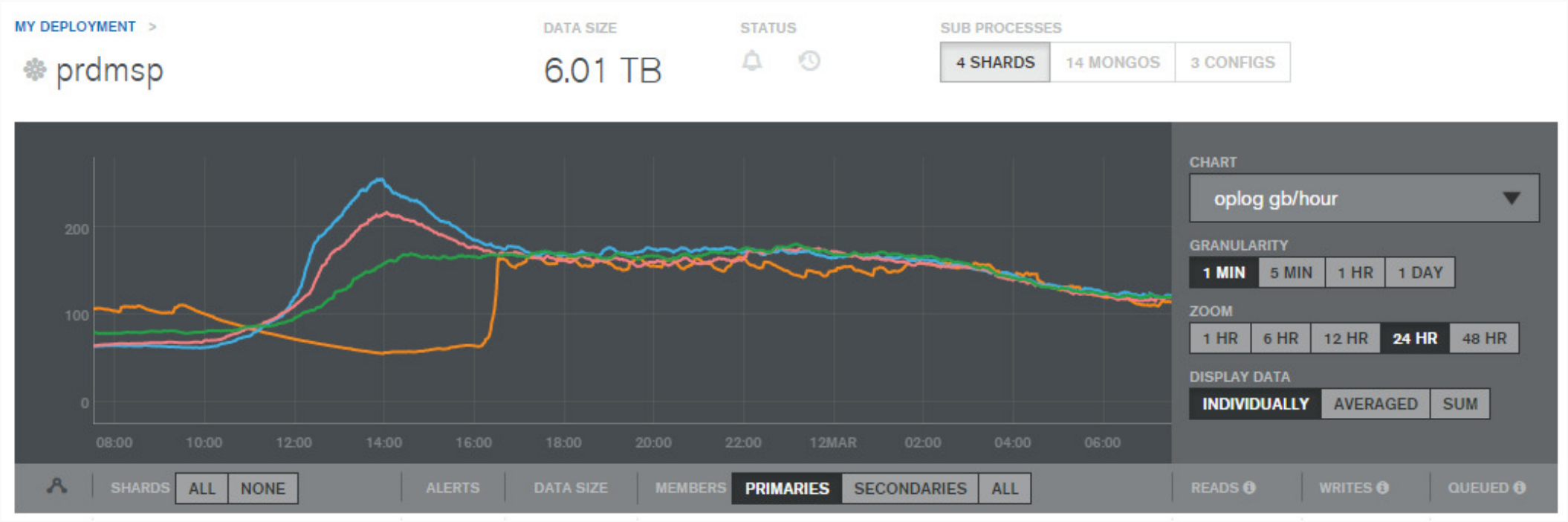
3 One other Use-case

© 2015 Amadeus IT Group SA

amadeus


MyAirline.com
ENGLISH
HELP
LOG-IN
From
Origin
To
Destination
Departing
dd/mm/yyyy
Returning
dd/mm/yyyy
Passengers
1 Adult
Cabin
Economy
Search
1. SELECT - FLIGHTS
1 SELECT
2 BOOK
3 PAY
BOSTON
Logan International (BOS)
TO PARIS
All Airports
SATURDAY
10 OCTOBER
October 2015
-7 DAYS
WED 07 USD 295
THU 08 USD 275
FRI 09 USD 275
SAT 10 USD 275
SUN 11 USD 300
MON 12 USD 310
TUE 13 USD 275
+7 DAYS
Order by
Filters
What is included in my fare
Economy Class
Economy Class Special
Economy Class Flex
Economy Comfort
Economy Comfort Special
Saga business Class
06:23 Logan International (BOS)
01:17 +1 day Charles de Gaulle (CDG)
Duration 18h 54m, 1 stop
My Airlines (MA471)
Show flights details
USD 377.50
USD 622.00
USD 762.50
USD 1,005.00
USD 1,257.00
4 seats left
USD 1,478.50
07:32 Logan International (BOS)
02:26 +1 day Charles de Gaulle (CDG)
Duration 18h 54m, 1 stop
My Airlines (MA471)
Show flights details
USD 392.50
Waitlist
USD 571.00
NOT AVAILABLE
USD 997.50
USD 1,243.00
USD 1,458.50
08:17 Logan International (BOS)
03:11 +1 day Charles de Gaulle (CDG)
Duration 18h 54m, 1 stop
My Airlines (MA471)
Show flights details
USD 275.00
USD 629.00
NOT AVAILABLE
USD 980.00
4 seats left
NOT AVAILABLE
USD 1,405.00
4 seats left
09:32 Logan International (BOS)
04:26 +1 day Charles de Gaulle (CDG)
Duration 18h 54m, 1 stop
My Airlines (MA471)
Show flights details
USD 275.00
USD 587.50
USD 763.00
NOT AVAILABLE
USD 1,265.00
USD 1,411.50
Your booking
4 PASSENGERS
2 adults, 1 child, 1 infant
SAT 10 OCT
Please select flight
SUN 22 NOV
Please select flight
Continue
Booking details
Legal notice
Baggage policy

amadeus


MY DEPLOYMENT >
prdmsp
DATA SIZE
6.01 TB
STATUS
SUB PROCESSES
4 SHARDS
14 MONGOS
3 CONFIGS
200
100
0
08:00
10:00
12:00
14:00
16:00
18:00
20:00
22:00
12MAR
02:00
04:00
06:00
CHART
oplog gb/hour
GRANULARITY
1 MIN
5 MIN
1 HR
1 DAY
ZOOM
1 HR
6 HR
12 HR
24 HR
48 HR
DISPLAY DATA
INDIVIDUALLY
AVERAGED
SUM
SHARDS
ALL
NONE
ALERTS
DATA SIZE
MEMBERS
PRIMARIES
SECONDARIES
ALL
READS
WRITES
QUEUED

amadeus

4 Future plans

© 2015 Amadeus IT Group SA

~10 use-cases running currently
So far no showstopper issues identified

We still miss some features:

- Cross cluster replication
  - Filtered / delayed / batched
- Multimaster (AP operation)
- Multithreaded queries
- Transactions
  - On multi update

Thank you!
Questions?

© 2015 Amadeus IT Group SA

amadeus