使用MongoDB构建物联网应用

目录

关于讲师

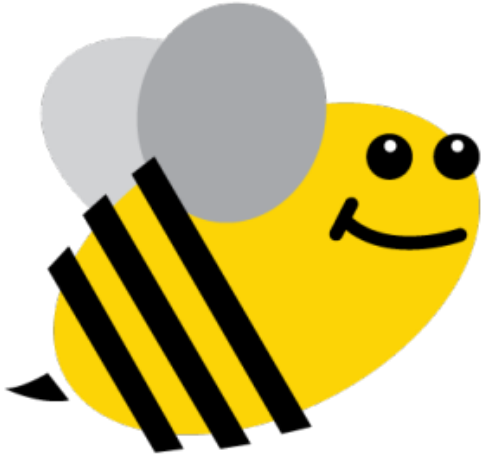
TJ TANG

高级方案架构师

jianfa.tang@mongodb.com

http://www.mongoing.com

什么是物联网？

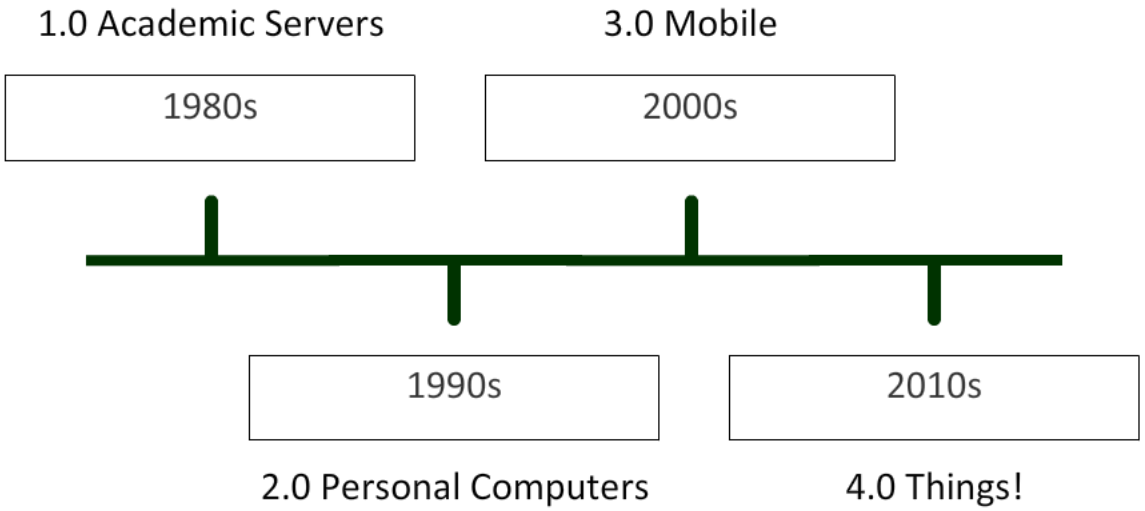


是一个Buzzword！

最”Buzz”的词汇！

Internet 4.0

互联网的演化

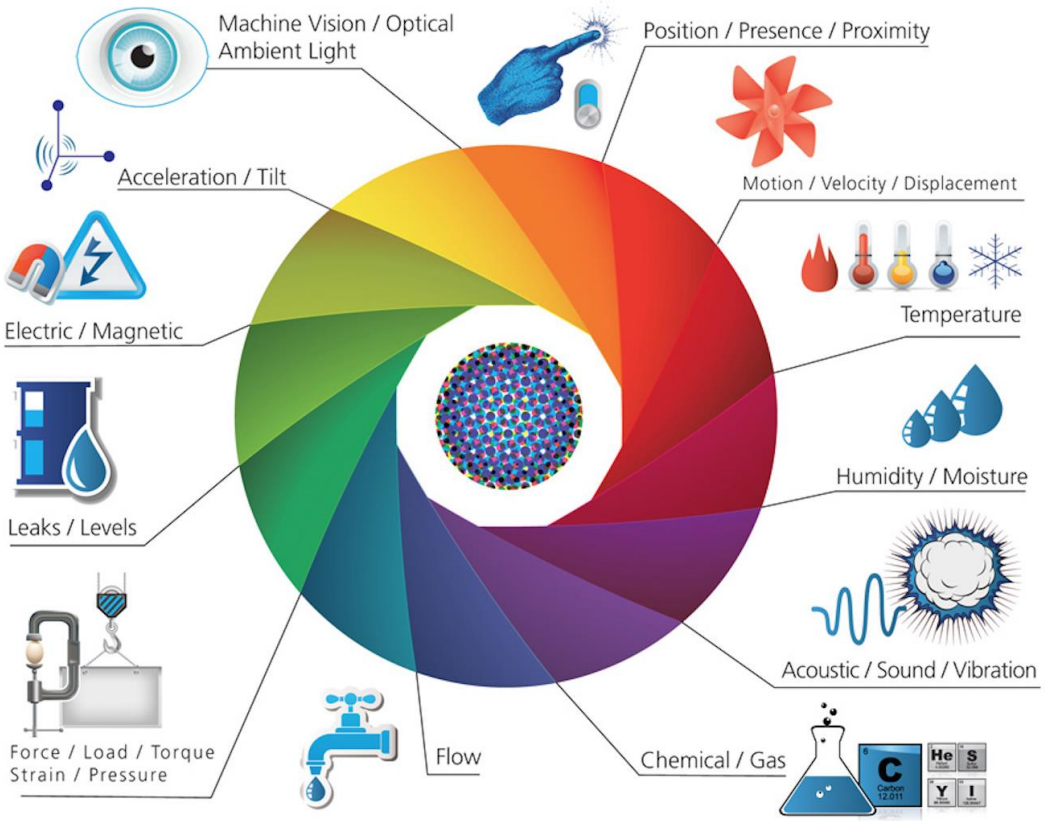


物联网中有什么？

来源: http://postscapes.com/what-exactly-is-the-internet-of-things-infographic

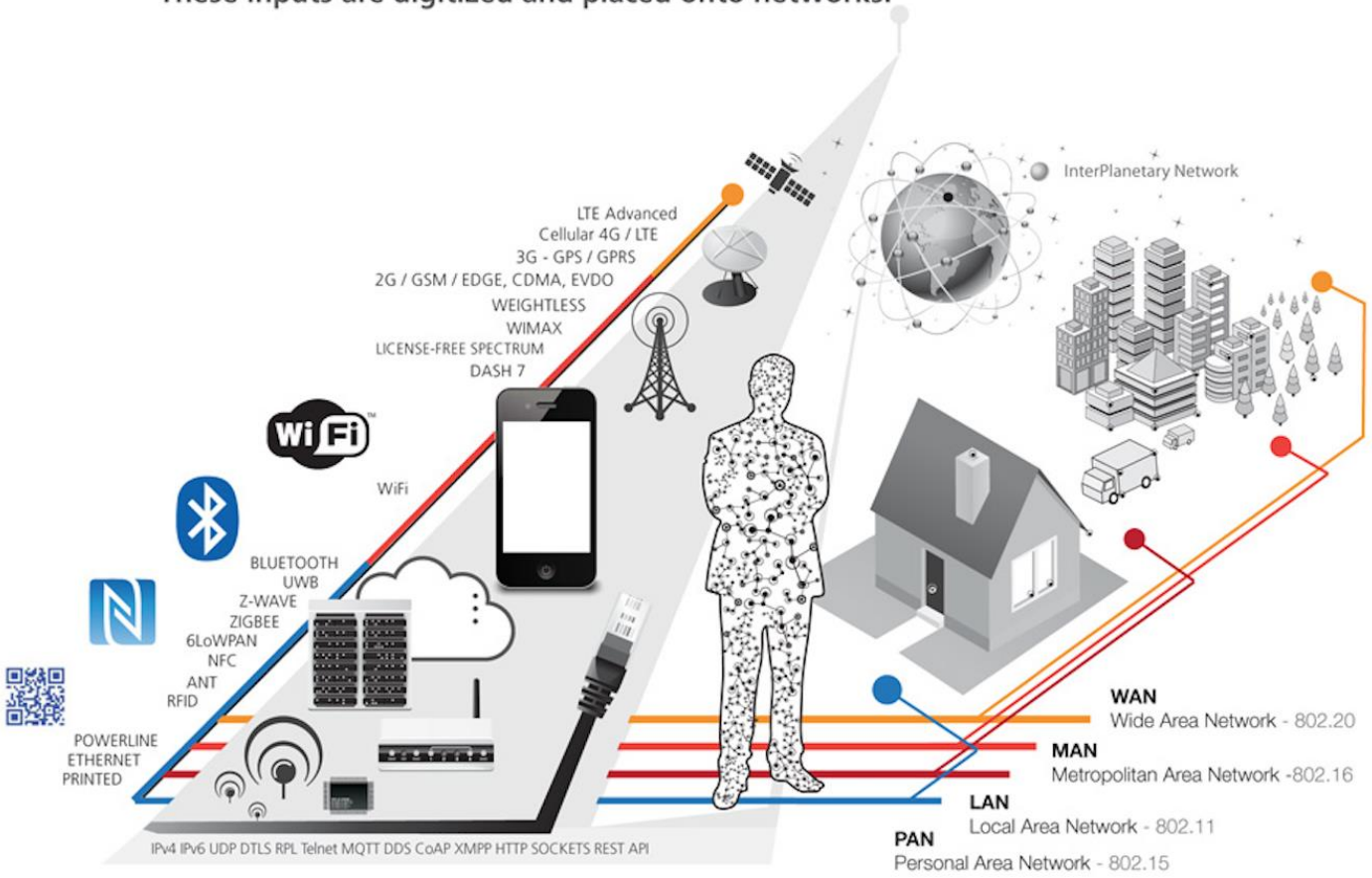
1 SENSORS & *ACTUATORS*

We are giving our world a digital nervous system. Location data using GPS sensors. Eyes and ears using cameras and microphones, along with sensory organs that can measure everything from temperature to pressure changes.



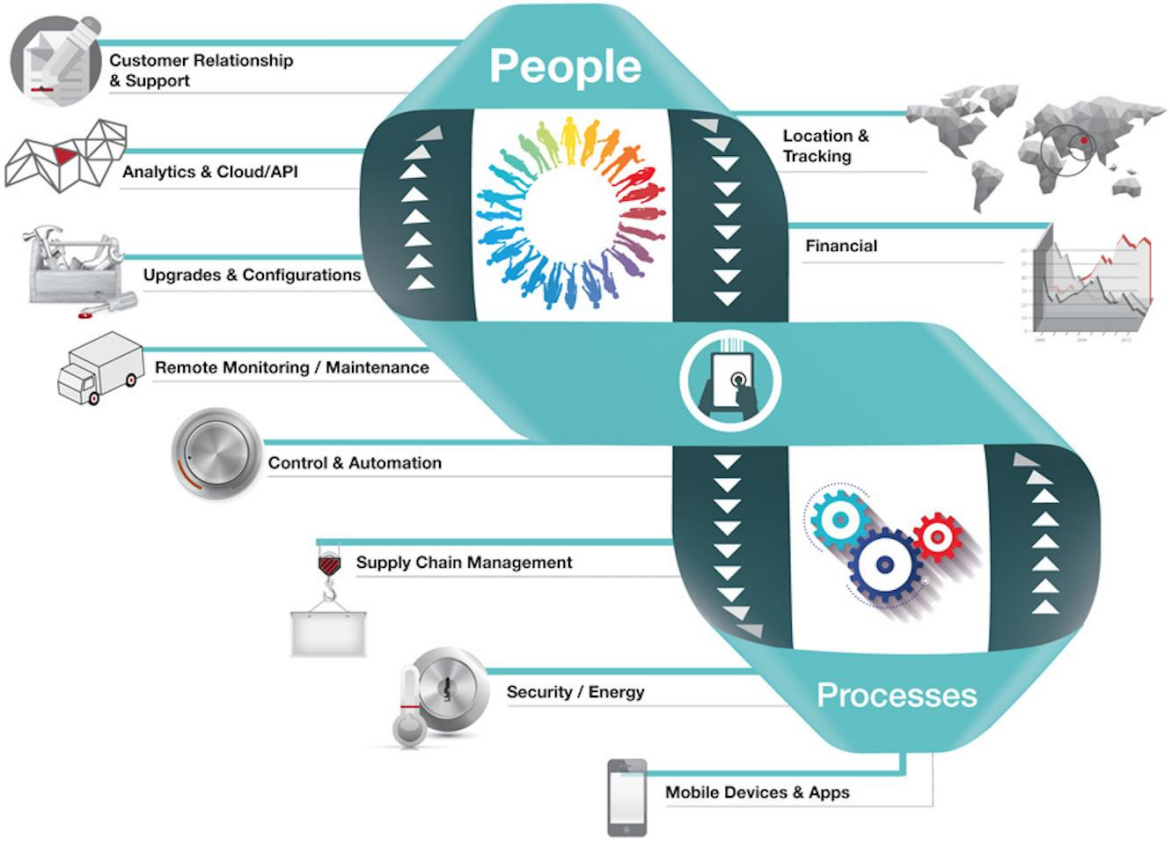
2 CONNECTIVITY

These inputs are digitized and placed onto networks.



3 PEOPLE & *PROCESSES*

These networked inputs can then be combined into bi-directional systems that integrate data, people, processes and systems for better decision making.



CONNECTED COW by VITAL HERD

电子药丸注射到胃中

传输心率，温度及化学成分

检测到异常情况时通知农场主

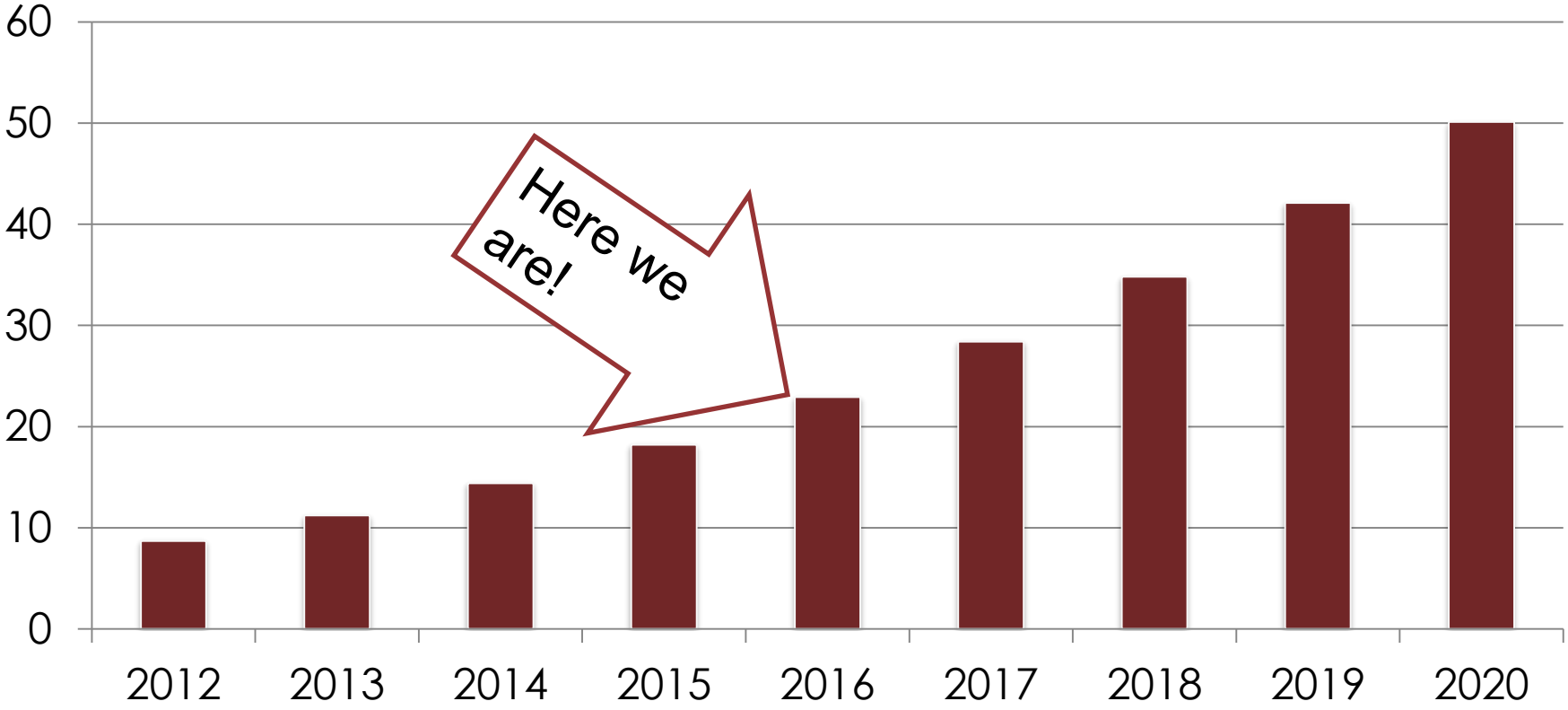
健康管理

美国9400万头奶牛，降低了数以亿计的成本

MyJohnDeere

物联网总连接数

（单位：十亿）

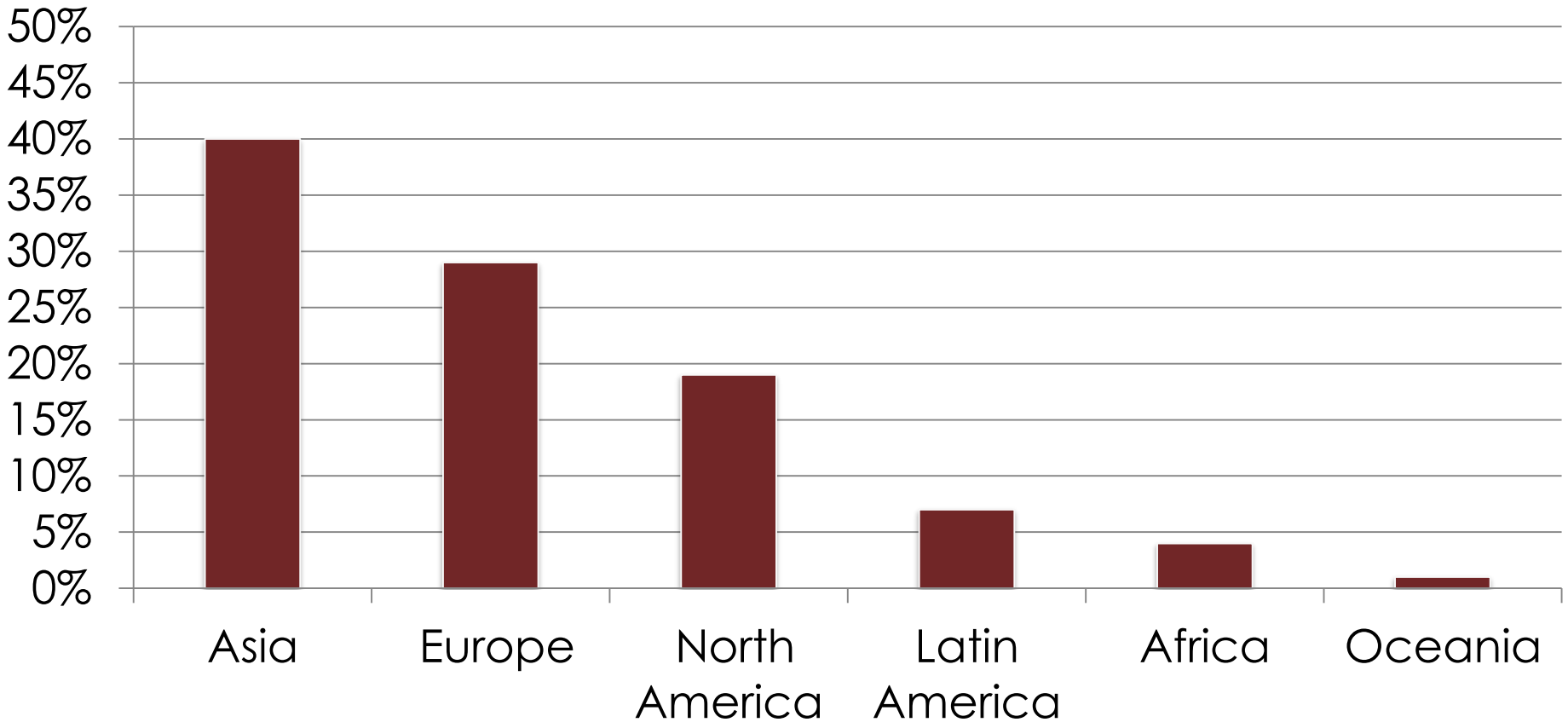


Source: Cisco.

物联网市场规模
（单位：万亿美元）

Source: IDC.

机器对机器连接百分比



Source: GSMA.

物联网的技术栈

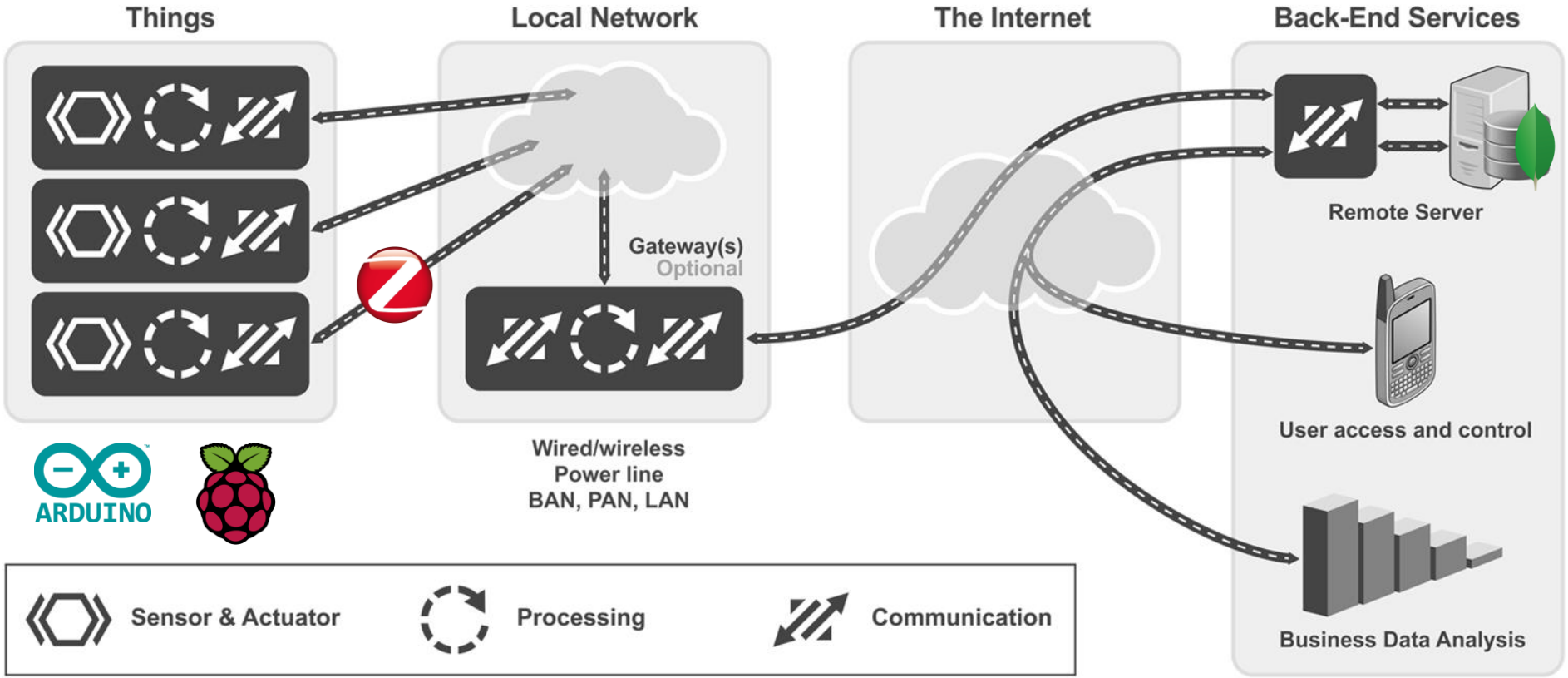
价值传递：业务分析，用户存取&控制

中间件和存储：应用服务器，数据库服务器

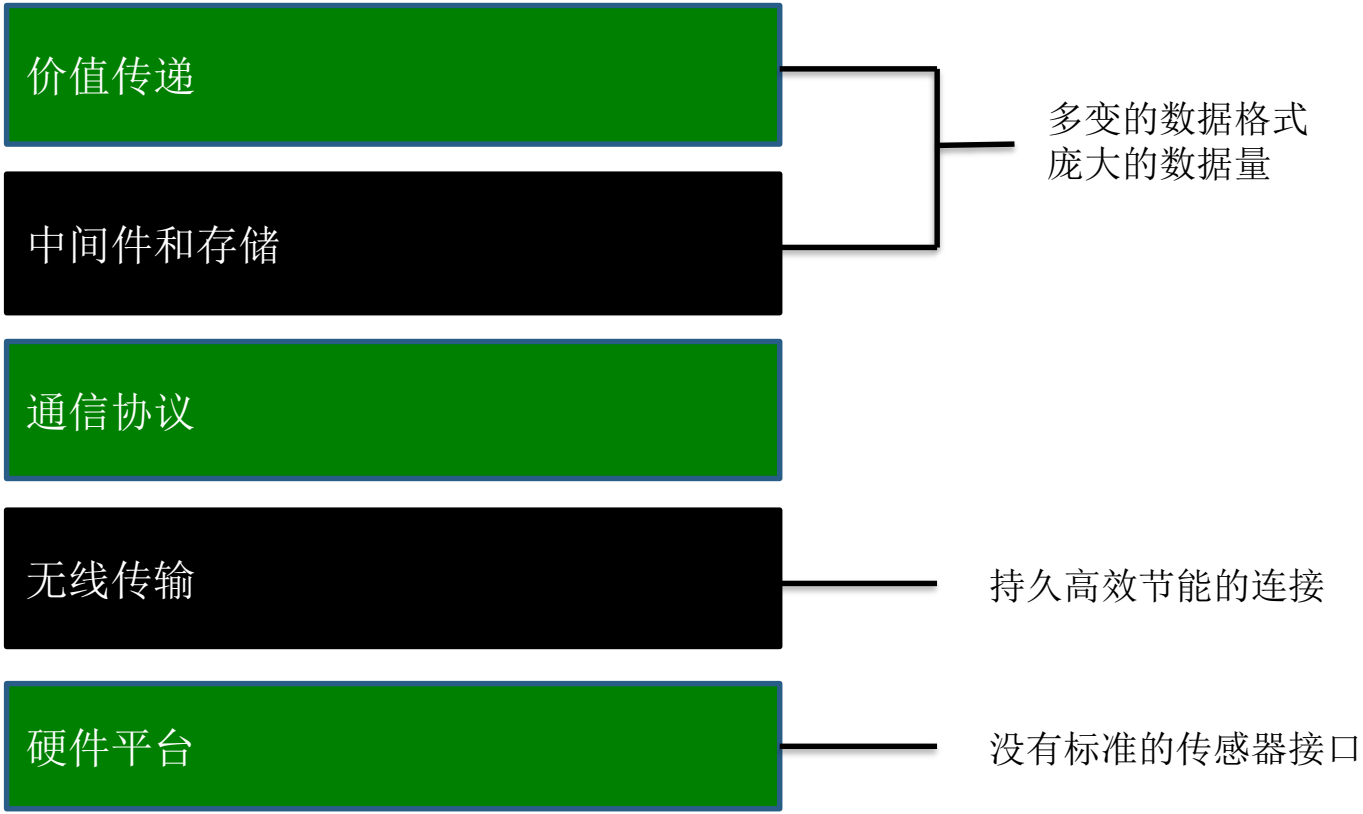
通信协议：MQTT, CoAP, XMPP, AMQP, RESTful

无线传输：Zigbee，Z-Wave, WIFI, GPRS, Bluetooth-LE

硬件平台：Arduino, Raspberry Pi, Intel Edison


Things
Local Network
The Internet
Back-End Services
Gateway(s)
Optional
Remote Server
User access and control
Business Data Analysis
ARDUINO
Wired/wireless
Power line
BAN, PAN, LAN
Sensor & Actuator
Processing
Communication

挑战


价值传递
中间件和存储
通信协议
无线传输
硬件平台
多变的数据格式
庞大的数据量
持久高效节能的连接
没有标准的传感器接口


malaysia AIRLINES
malaysia
OUR PRAYERS ALWAYS WITH YOU!
GOD Bless
and be safe

多变的数据结构

多种来源

ADS-C, HFDL, ASDI, EUROCONTROL, ACARS

各种形式

```
location: [ 38.2031,  -120.4904 ] ,
speed: 750,
altitued: 29384,
engine:
        fuel_level: 78% ,
        temperature: 89,
        EPR: xx
        N-value: { N1: xxx, N2: xxx, N3: xx}
...
```

设计方案1

将每个指标作为关系表里的一个列来建模

EVENT_ID	PLANE_ID	TIMESTAMP	LAT	LONG	ENGINE TEMP	FUEL LEVEL	…	SPEED
100001	3902	1437297148810	38.2031	-124.4904				
100002	3902	1437297149213						750

巨大的表格，由于空值所造成的大量空间浪费

增加新的指标时会带来频繁的模式修改及数据迁移

设计方案 2

在实体-属性-值表格中存储多变的指标

EVENT_ID	PLANE_ID	TIMESTAMP
100001	3902	1437297148810

EVENT_ID	METRIC_NAME	METRIC_VALUE
100001	LAT	38.2031
100001	LONG	-124.4904
100002	SPEED	750

METRIC_VALUE 必须定义为 TEXT 字段

对METRIC_VALUE 字段进行索引可能带来隐含问题

必要的话，需要有多个自联接

巨大的数据量

一个航班，每秒：

3 * 60 * 100 = 18K 数据点/航班

每天100，000 个航班:

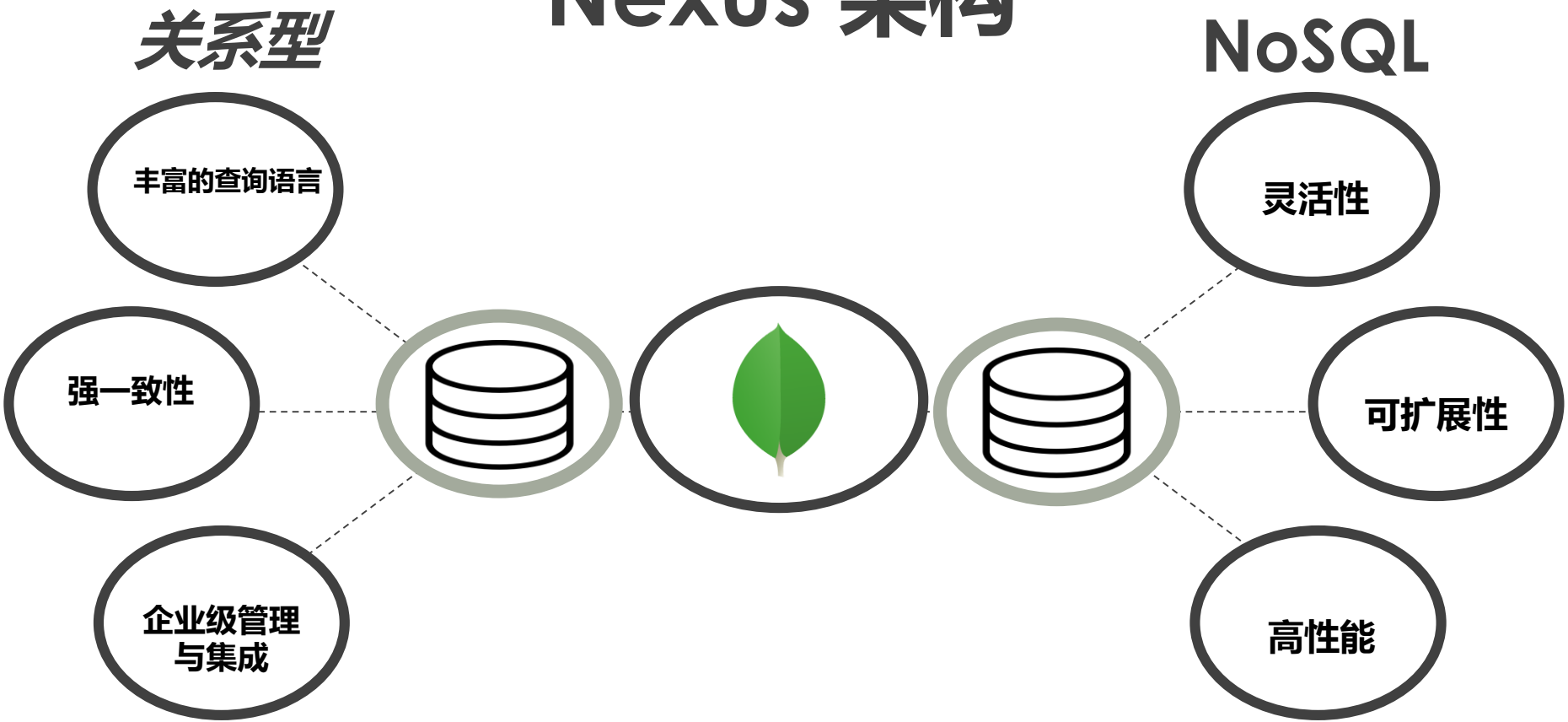
18 亿, 1.8TB /天
21,000 条查询

管理海量，形式多变的数据是非常困难的一件事情。

MEET

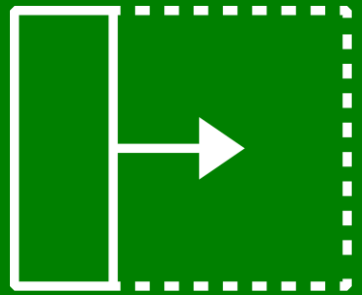
mongoDB

Nexus 架构



灵活、敏捷性

可扩展性



敏捷性

不需要漫长的实体关系设计，可以立刻开始编码

动态地根据需求简单灵活修改模式

毫不费力地使用富文档对异构数据进行建模

location: (-84.2391, 34.1039)

speed: 750

engine:
fuel_level: 100 ,
temperature: 88.48

1. 可变的数据结构
2. 稀疏索引
3. 动态模式

数据模型

```
{
    "_id" : ObjectId("55a99ceee4b05e3843afd9e9"),
    "plane_id" : "3902",
    "ts" : ISODate("2015-07-17T22:25:16.409Z"),
    "metrics" : {
            "fuel_level": 100,
            "engine_temp": 88.48
    }
},
{
    "_id" : ObjectId("55a99cece4b05e3843afd97a"),
    "plane_id": "3902",
    "ts" : ISODate("2015-07-17T22:25:16.409Z"),
    "metrics" : {
       "location": [ -84.2391, 34.1039 ]
    }
}
db.events.ensureIndex({"plane_id":1, "ts":-1});
db.events.ensureIndex({"ts":-1, "metrics.location": "2d"}, {sparse: true});

### sometime in future, a new metric "elevation" is introduced:

{
    "_id" : ObjectId("55a99ceee4b05e3843afd9e9"),
    "plane_id" : "3902",
    "ts" : ISODate("2016-07-17T22:25:16.409Z"),
    "metrics" : {
            "fuel_level": 90,
            "elevation": 20000
    }
},
```

1

2

3

```
### Find all engine fuel events from 15:00 to draw a histogram

db.events.find({
    plane_id: 3902,
    ts: {$gt: ISODate("2015-07-29 15:00:00")},
    "merics.fuel_level": {$exists: true}
});

### Find all planes currently within 20km of New York city

db.events.aggregate([
   {
     $geoNear: {
        near:   [ -84.0000, 34.0000 ] ,
        distanceField: "distance",
        maxDistance: 20000,
        query: { ts: {$gt: ISODate("2015-07-29 15:00:00")}  }
     }
   },
   {$group: {_id: "$plane_id", location: {$last: "$metrics.location"}} }
])
```

查询示例

优化

使用文档模型

时间序列是
— **系列数据点**，
一般由 一个 **时间段** 内生成的
连续度量结果 组成。

时间序列的示例包括海洋潮汐，太阳黑子的数量以及每日道琼斯指数的平均收市指数。

--wikipedia

```
{
  plane_id: ”3209",
  ts: ISODate("2014-07-03T16:00:00.000Z”)
  metrics: { engine_fuel: 99 }
},
```

```
{
  plane_id: ”3209",
  ts: ISODate("2014-07-03T16:01:00.000Z”)
  metrics: { engine_fuel: 98.5 }
},
```

```
{
  plane_id: ”3209",
  ts: ISODate("2014-07-03T16:02:00.000Z”)
  metrics: { engine_fuel: 98 }
}
```

. . .

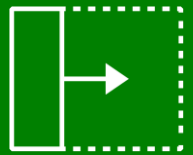
```
{
  plane_id: ”3209",
  ts: ISODate("2014-07-03T16:59:00.000Z”)
  metrics: { engine_fuel: 69 }
}
```


```
{
  plane_id: ”3209",
  hour: ISODate("2014-07-03T16:00:00.000Z"),
  metrics: {
    engine_fuel:  {
       “0”:  99,
       “1”:  98.5,
       “2”:  98,
       ...
       “59”: 69
    },
    avg: 81.4
  }
}
```

- 更少的文档-节省空间
- 写入性能-更少的索引条目
- 可查询性&更优秀的分析支持

时间序列数据的分桶优化

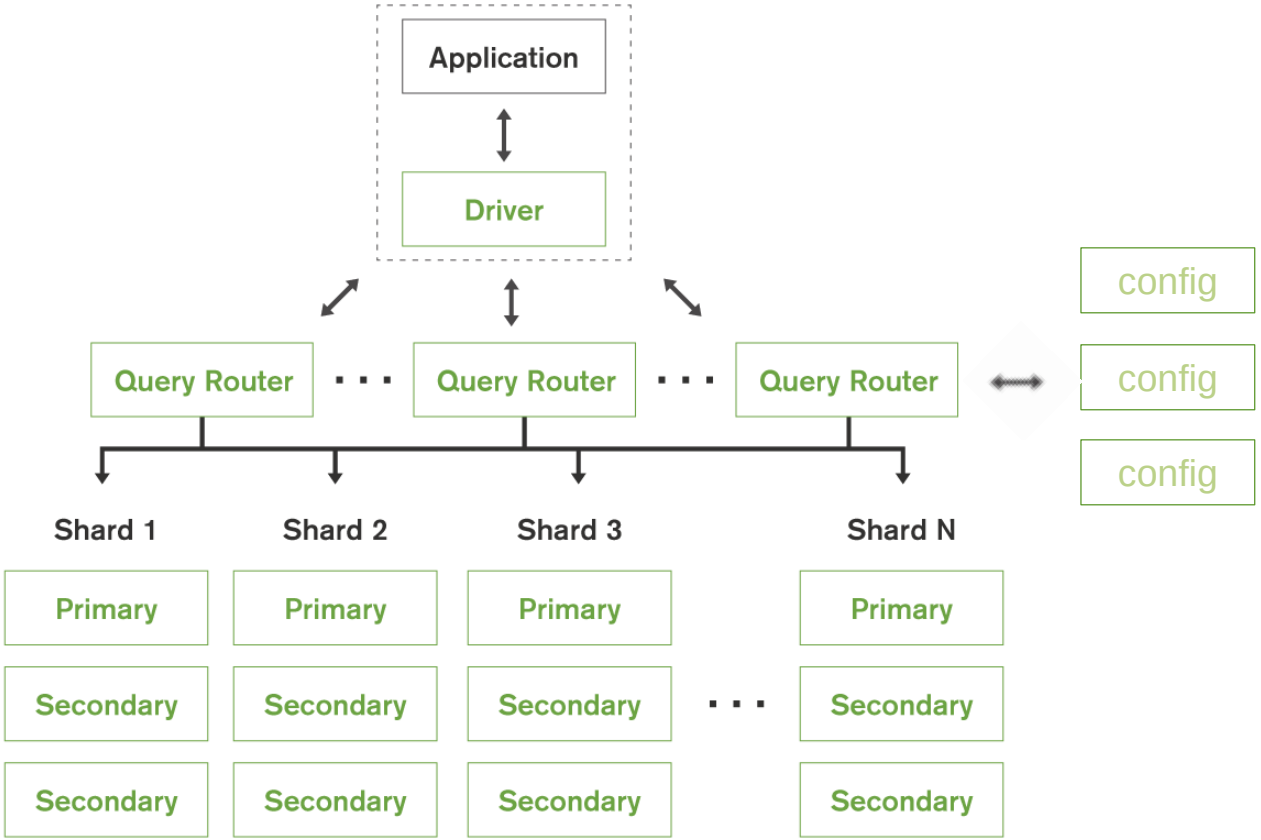
水平扩展



非共享，水平、线性拓展

自动均衡以保证一个均匀分布的集群

分片集群


Application
Driver
Query Router
Query Router
Query Router
config
config
config
Shard 1
Shard 2
Shard 3
Shard N
Primary
Primary
Primary
Primary
Secondary
Secondary
Secondary
Secondary
Secondary
Secondary
Secondary
Secondary

大型分片集群

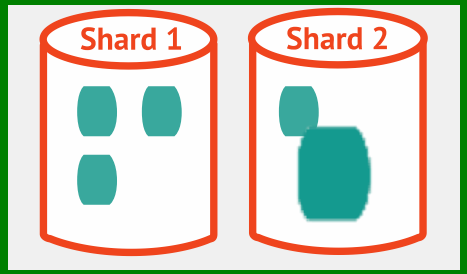
在线设计云服务提供商

超过 1,000,000 的每秒查询率

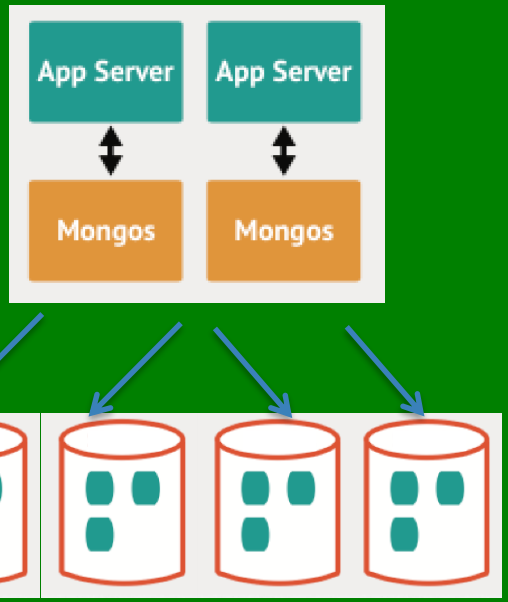
个人存储云的元数据

20个分片上存储 ~1000 亿条记录

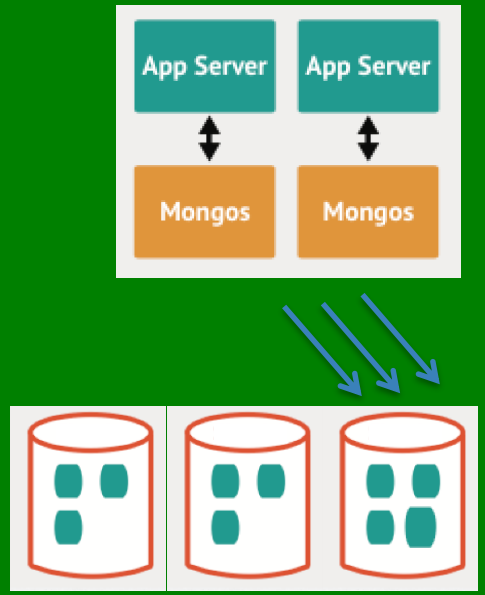
片键的选择 – 避免以下情况



基数太小导致
jumbo chunk



Scatter/Gather 影响范围查询性能



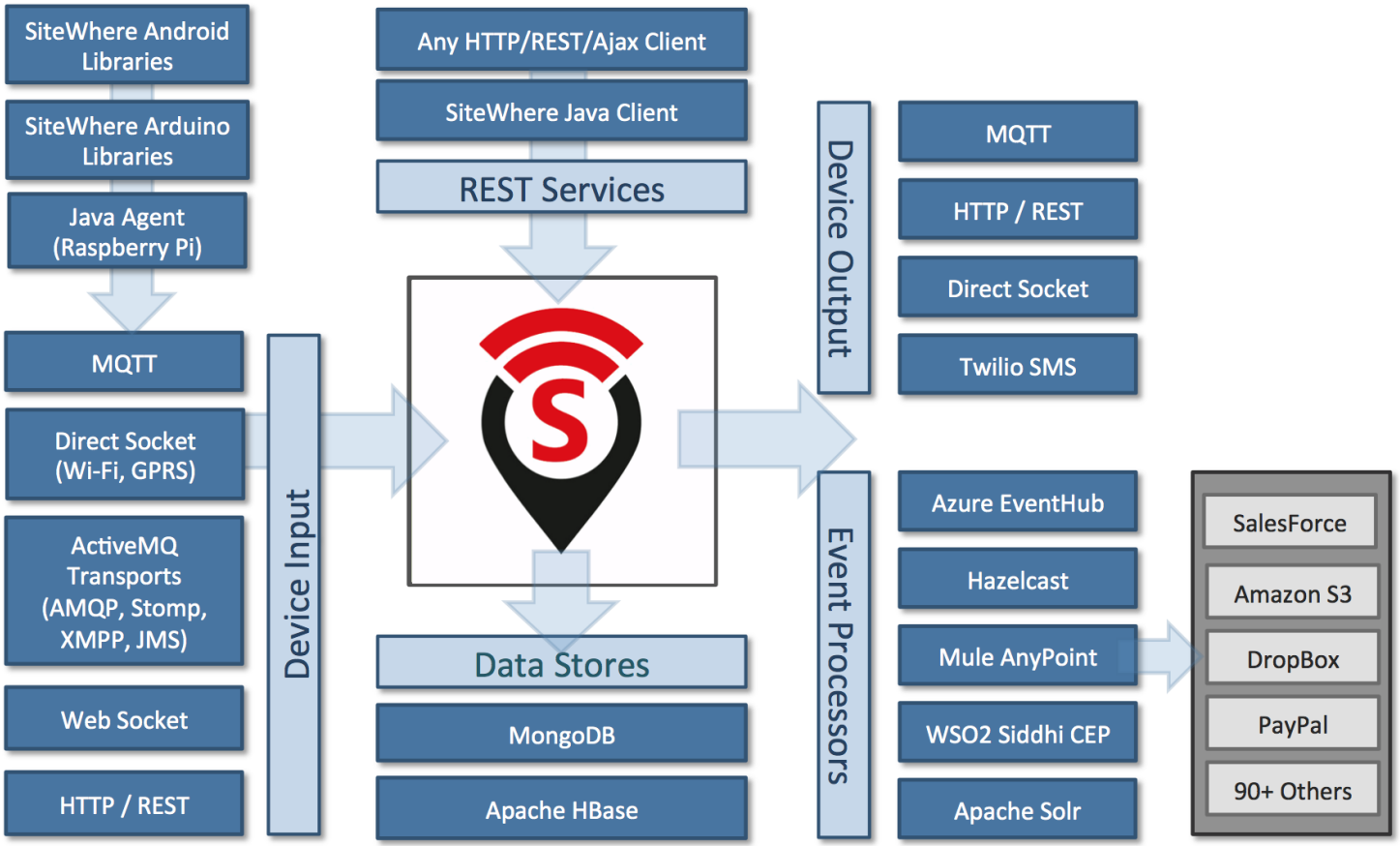
写入分布不均，
产生热分片

片键的选择 – 物联网数据

	基数	写分发	范围查询
_id	文档级别，基数大	热分片	Scatter/gather
hash(_id)	文档级别，基数大	所有分片	Scatter/gather
asset_id	一个Asset有很多文档	所有分片	定向
asset_id, ts	文档级别	所有分片	定向


KAA
APPLICATION
MIDDLEWARE
HARDWARE
KAA
Say "HELLO" To World
IoTgo
An Open Source IoT
SiteWhere
The Open Platform for the Internet of Things™
powered by
ongoDB

SiteWhere Connectivity Diagram



Air Traffic Example

localhost:8080/airtraffic/

Search

Map | Satellite

SiteWhere

AIR TRAFFIC MONITORING EXAMPLE

Flight SW-624

Boeing 737-700

Flight Details | Air Speed History | Fuel Level History

Google

Imagery ©2015 TerraMetrics | Terms of Use | Report a map error

2 使用Sitewhere的SDK/API将数据发送到Sitewhere

传感器
数据
网关

1 在XML中定义需要存储的数据格式

SiteWhere

追踪应用

3 通过Sitewhere的API检索飞机数据

SiteWhere - Manage Devic...

localhost:8080/sitewhere/admin/devices/list.html

Search

SiteWhere

Sites | Specifications | Devices | Device Groups | Batch | Users | Admin User

Manage Devices

Filter Results | Add New Device

Device	**FMZ 2000**	Asset	**Boeing 737-700**	**Created:**	2015-07-26 17:52:00
	Id: 4f6a93df-45e3-41f7-aeb9-90c...		**Assigned:** 2015-07-26 17:52:00	**Updated:**	N/A
	Spec: FMZ 2000 Specification		**Status:** Active		
Device	**FMZ 2000**	Asset	**Airbus A330-300**	**Created:**	2015-07-26 17:52:00
	Id: 96e67cb1-2828-4176-8a0d-80...		**Assigned:** 2015-07-26 17:52:00	**Updated:**	N/A
	Spec: FMZ 2000 Specification		**Status:** Active		
Device	**FMZ 2000**	Asset	**Airbus A330-300**	**Created:**	2015-07-26 17:52:00
	Id: 0bb87103-42f2-48bf-b4d8-a2...		**Assigned:** 2015-07-26 17:52:00	**Updated:**	N/A
	Spec: FMZ 2000 Specification		**Status:** Active		
Device	**FMZ 2000**	Asset	**Airbus A330-300**	**Created:**	2015-07-26 17:52:00
	Id: b5266b45-be97-4aef-b2de-0c...		**Assigned:** 2015-07-26 17:52:00	**Updated:**	N/A
	Spec: FMZ 2000 Specification		**Status:** Active		
Device	**FMZ 2000**	Asset	**Airbus A330-200**	**Created:**	2015-07-26 17:52:00
	Id: 3cee8679-6374-4118-9b1e-a5...		**Assigned:** 2015-07-26 17:52:00	**Updated:**	N/A

SiteWhere - View Assignm...

localhost:8080/sitewhere/admin/assignments/detail.html?token=ad385b93-6570-40d6-8936-d0addf6a465

Search

SiteWhere

Sites | Specifications | Devices | Device Groups | Batch | Users | Admin User

View Assignment

Emulate Assignment | Edit Assignment

Asset

Token:	ad385b93-6570-40d6-8936-d0addf6a4657
Assigned Hardware:	Boeing 737-700
Assigned device:	FMZ 2000
Created Date:	2015-07-26 17:52:00
Updated Date:	N/A
Active Date:	2015-07-26 17:52:00
Released Date:	N/A
Assignment Status:	Active

Locations | Measurements | Alerts | Command Invocations

Device Locations

Filter Results | Refresh

Latitude	Longitude	Elevation	Event Date
37.188108	-120.352560	18371.050000	2015-07-26 18:06:00
37.188108	-120.352560	18371.050000	2015-07-26 18:06:00
37.157330	-120.208100	19289.080000	2015-07-26 18:05:59
37.157330	-120.208100	19289.080000	2015-07-26 18:05:59

总结

物联网的数据特点	MongoDB的应对
海量	分片集群
时序数据	分桶优化
多态异构	动态模式
实时分析	聚合框架

mongoDB

MongoDB 官方为您提供…

MongoDB 高级企业版

在您的数据中心运行MongoDB的最佳最安全的方式

MongoDB Ops Manager

监控、备份及自动化部署

生产支持

生产及管理

开发支持

让我们帮助您运行MongoDB

咨询

我们解决问题

培训

提高团队的开发效率

TJ TANG
jianfa.tang@mongodb.com