Introductuion to High Performance Reactive Architectures & Spring WebFlux

Disclaimer

All information presented in this document and all supplementary materials and programming code represent only my personal opinion and current understanding and has not received any endorsement or approval by IPT - Intellectual Products and Technologies or any third party. It should not be taken as any kind of advice, and should not be used for making any kind of decisions with potential commercial impact.

The information and code presented may be incorrect or incomplete. It is provided "as is", without warranty of any kind, express or implied, including but not limited to the warranties of merchantability, fitness for a particular purpose and non-infringement. In no event shall the author or copyright holders be liable for any claim, damages or other liability, whether in an action of contract, tort or otherwise, arising from, out of or in connection with the information, materials or code presented or the use or other dealings with this information or programming code.

About Me

Trayan Iliev

– CEO of IPT – Intellectual Products & Technologies

– Oracle® certified programmer 15+ Y

– End-to-end reactive fullstack apps with Go, Python, Java, ES6/7, TypeScript, Angular, React and Vue.js

– 12+ years IT trainer

– Voxxed Days, jPrime, Java2Days, jProfessionals, BGOUG, BGJUG, DEV.BG speaker

– Lecturer @ Sofia University – courses: Fullstack React, Multimedia with Angular & TypeScript, Spring 5 Reactive, Internet of Things (with SAP), Multiagent Systems and Social Robotics, Practical Robotics & Smart Things

– Robotics / smart-things/ IoT enthusiast, RoboLearn hackathons organizer

Where to Find The Code and Materials?

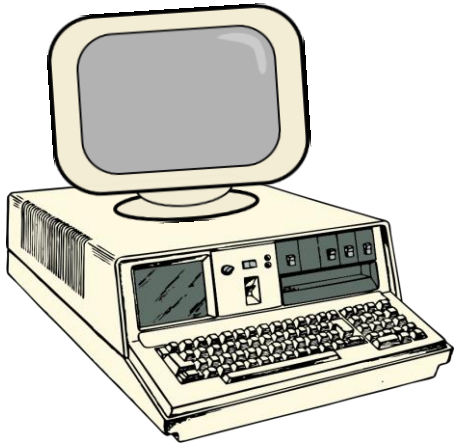
https://github.com/iproduct/course-ml

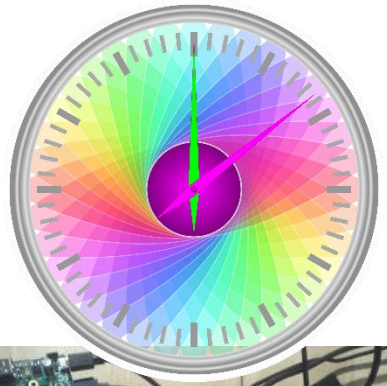
Machine Learning + Big Data in Real Time + Cloud Technologies

=> The Future of Intelligent Systems

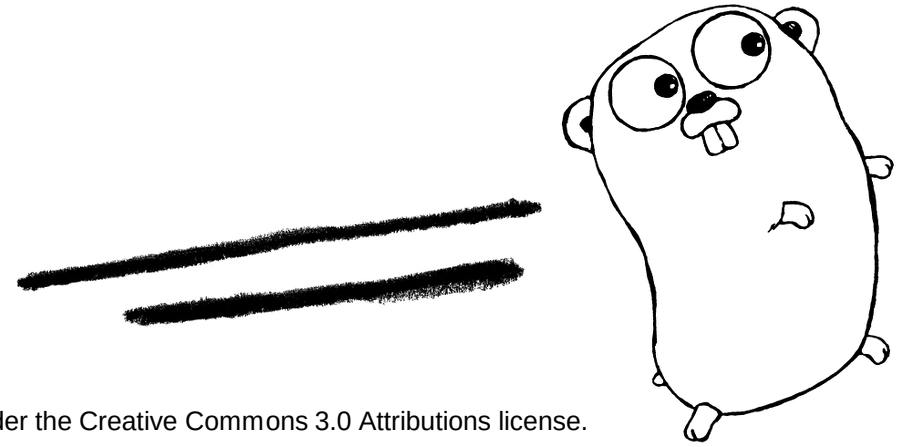
Asynchronous Event Streams in Real-time

Need for Speed :)

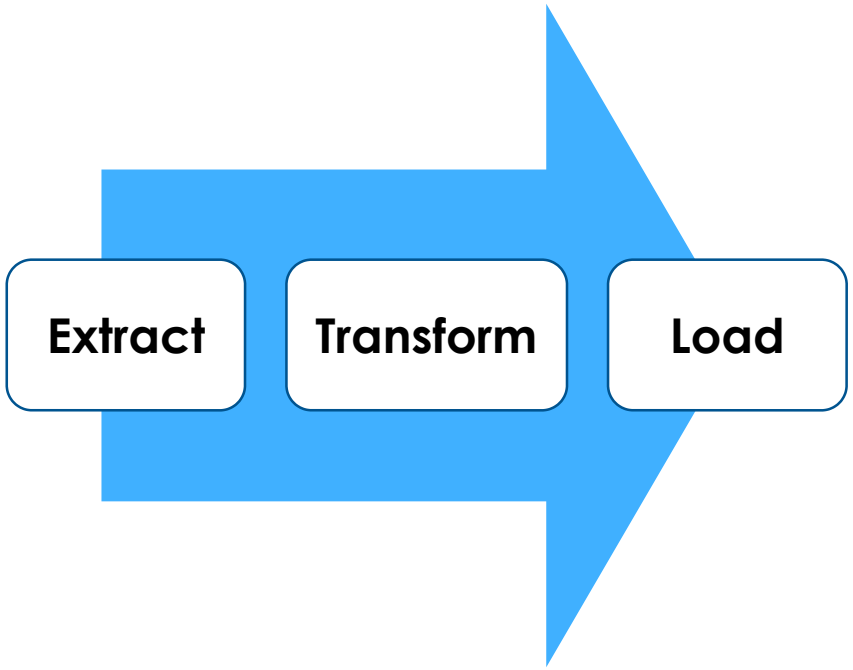


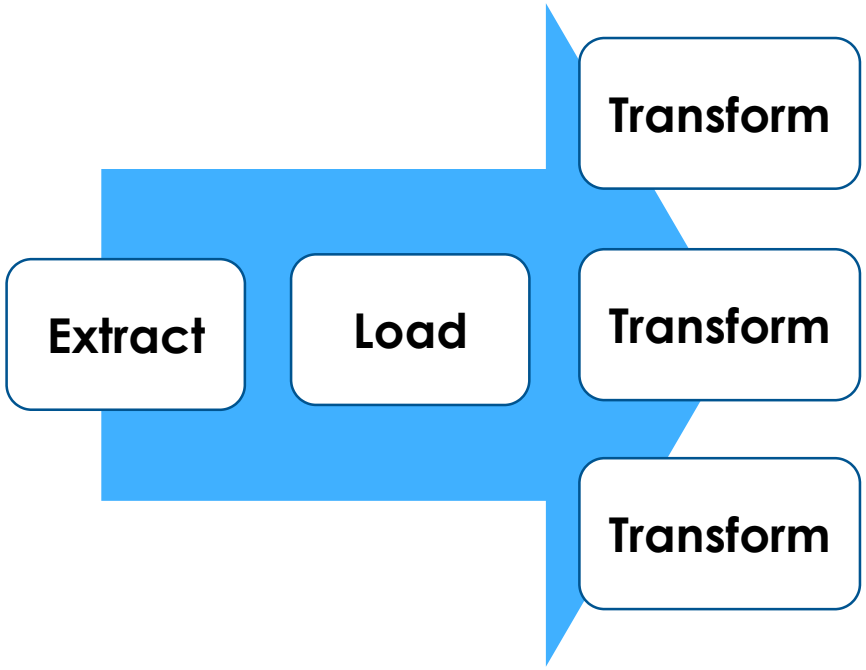


?

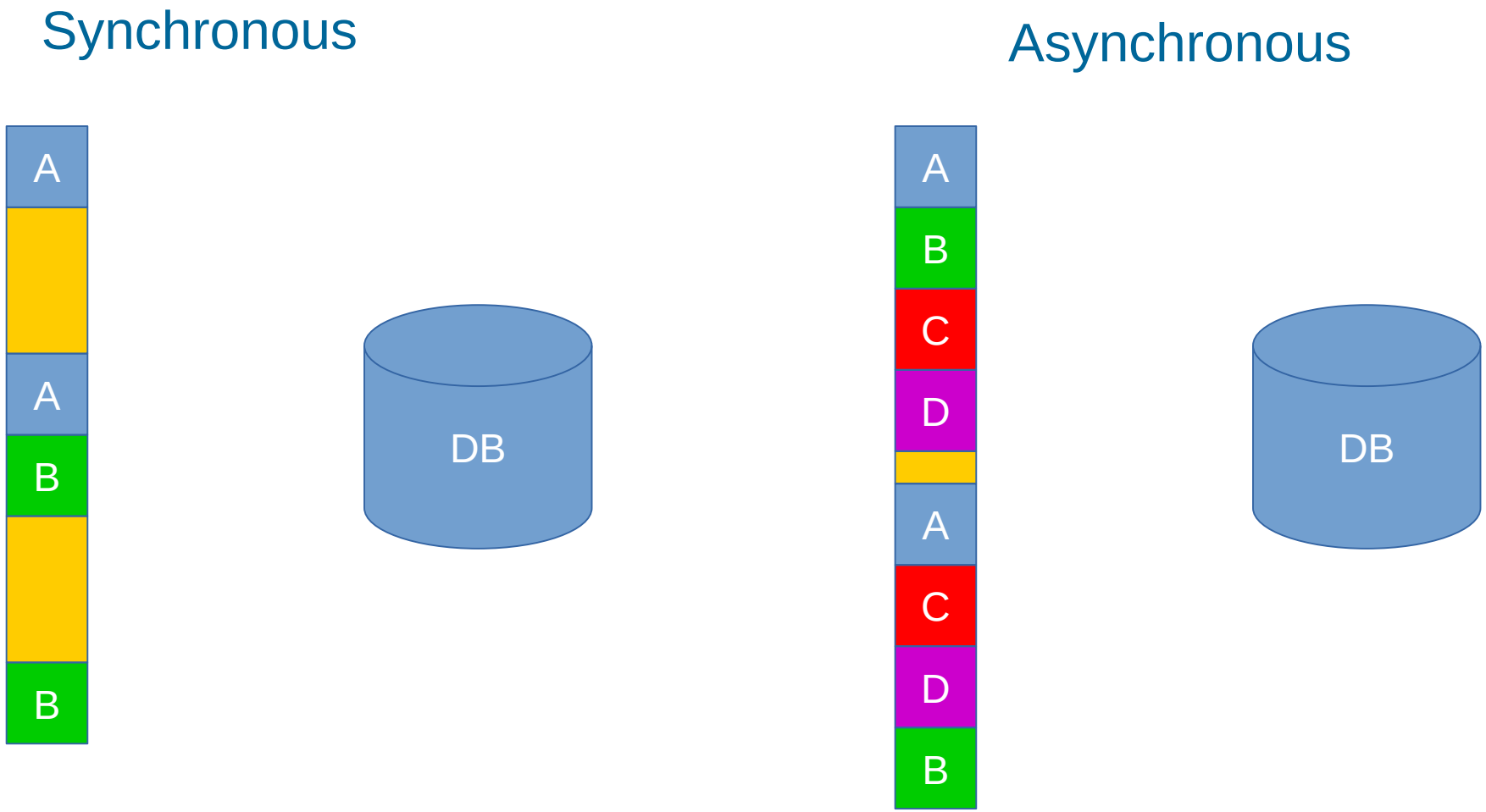


Batch Processing



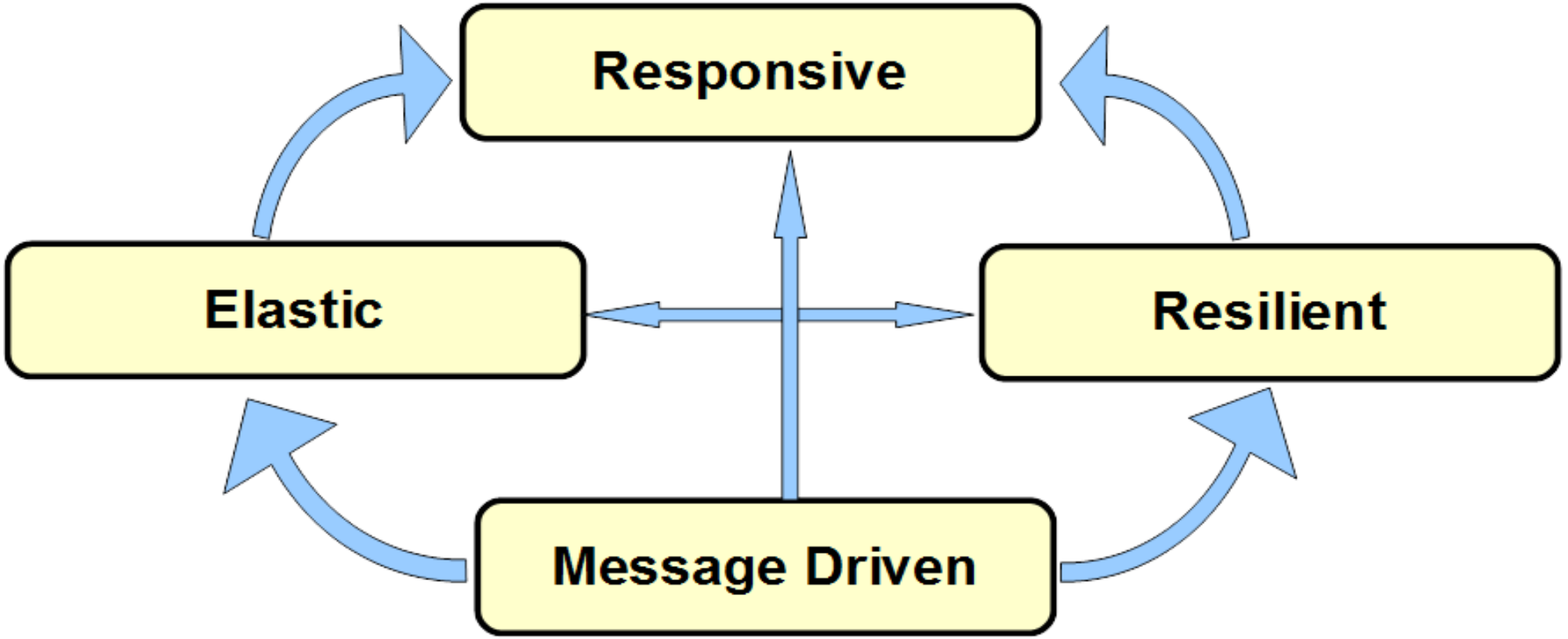


Synchronous vs. Asynchronous IO



Reactive Manifesto

http://www.reactivemanifesto.org



Scalable, Massively Concurrent

- **Message Driven** – asynchronous message-passing allows to establish a boundary between components that ensures loose coupling, isolation, location transparency, and provides the means to delegate errors as messages [Reactive Manifesto].
- The main idea is to separate concurrent producer and consumer workers by using **message queues.**
- **Message queues** can be **unbounded** or **bounded** (limited max number of messages)
- **Unbounded** message queues can present memory allocation problem in case the producers outrun the consumers for a long period → **OutOfMemoryError**

What's High Performance?

- **Performance** is about 2 things (Martin Thompson – http://www.infoq.com/articles/low-latency-vp):
  - **Throughput** – units per second, and
  - **Latency** – response time
- **Real-time** – time constraint from input to response regardless of system load.
- **Hard real-time system** if this constraint is not honored then a total system failure can occur.
- **Soft real-time system** – low latency response with little deviation in response time
- **100 nano-seconds** to **100 milli-seconds**. [Peter Lawrey]

Data / Event / Message Streams

“Conceptually, a stream is a (potentially never-ending) flow of data records, and a transformation is an operation that takes one or more streams as input, and produces one or more output streams as a result.”

Apache Flink: Dataflow Programming Model

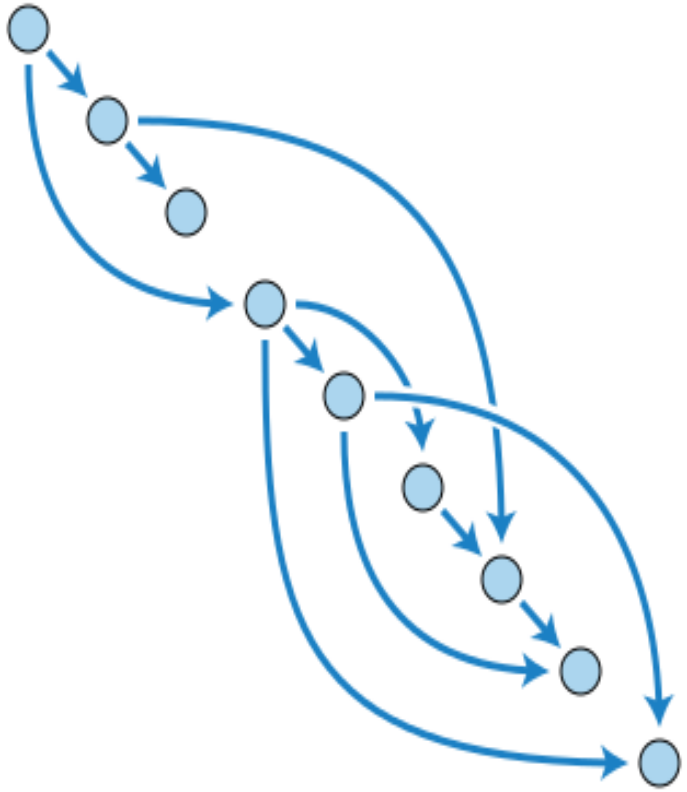
Data Stream Programming

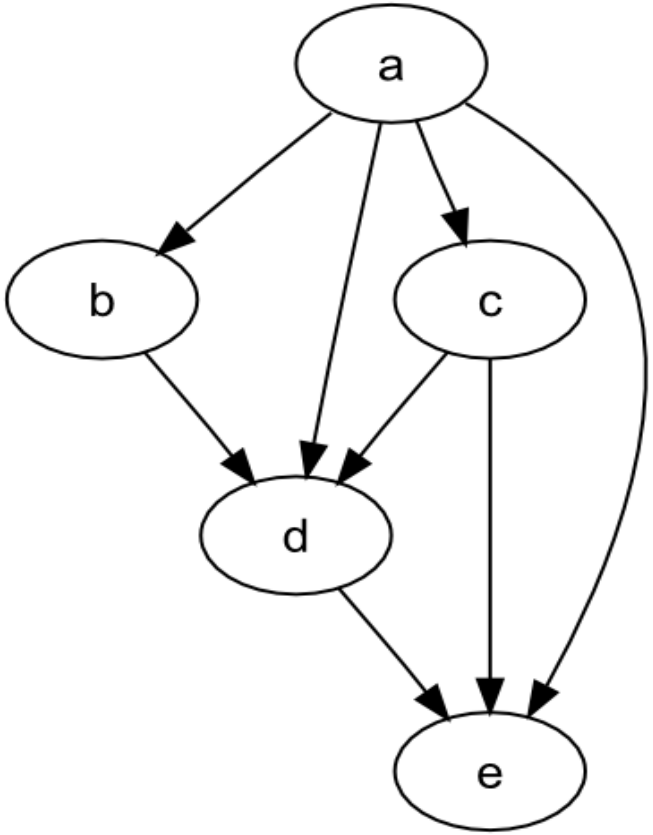
The idea of abstracting logic from execution is hardly new -- it was the dream of SOA. And the recent emergence of microservices and containers shows that the dream still lives on.

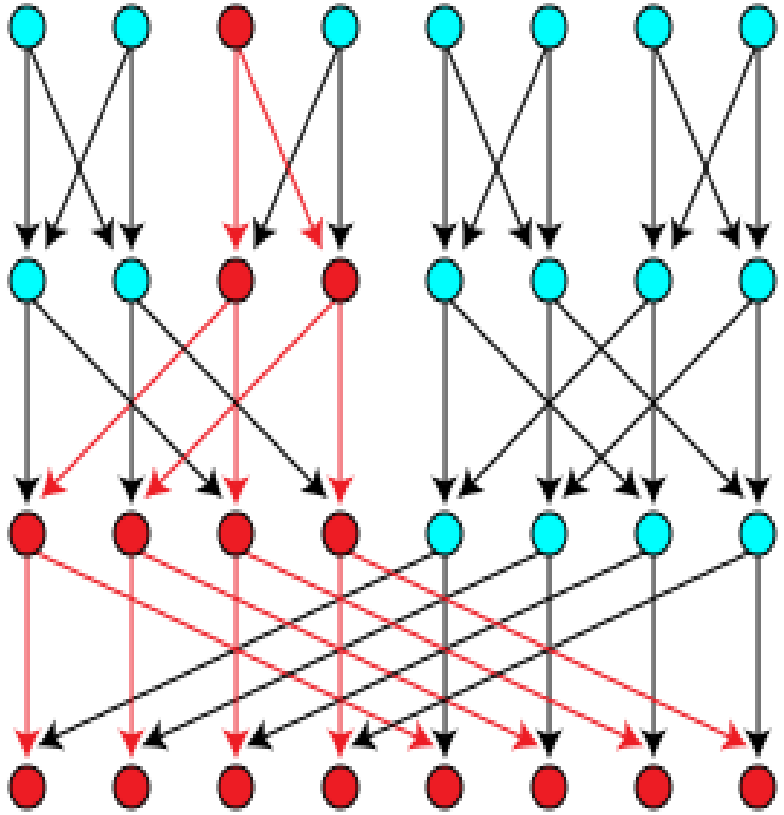
For developers, the question is whether they want to learn yet one more layer of abstraction to their coding. On one hand, there's the elusive promise of a common API to streaming engines that in theory should let you mix and match, or swap in and swap out.

Tony Baer (Ovum) @ ZDNet - Apache Beam and Spark: New coopetition for squashing the Lambda Architecture?

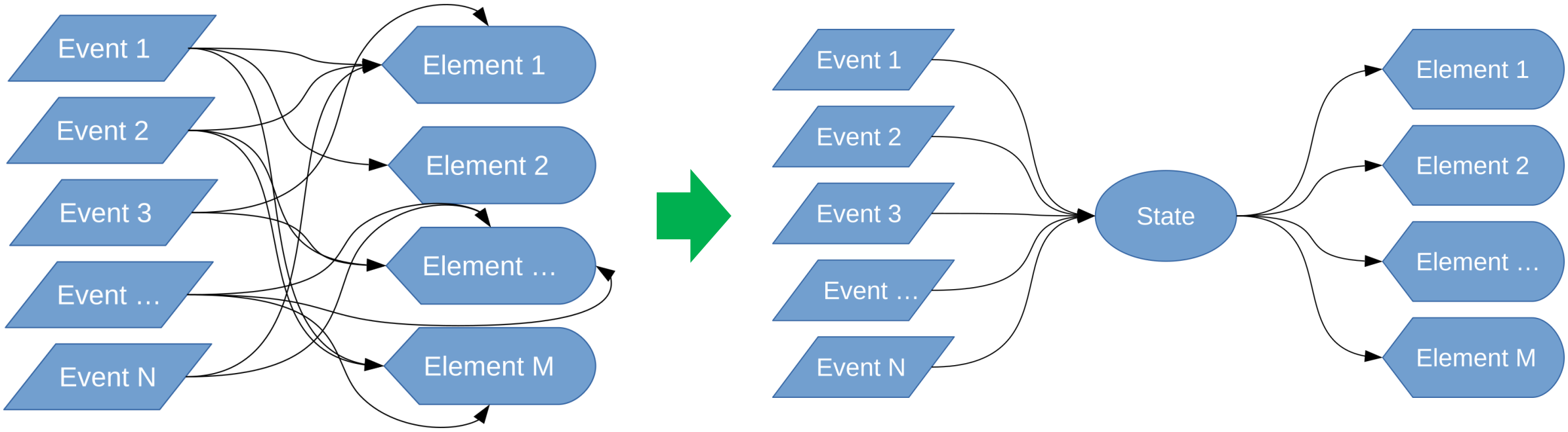
Direct Acyclic Graphs - DAG





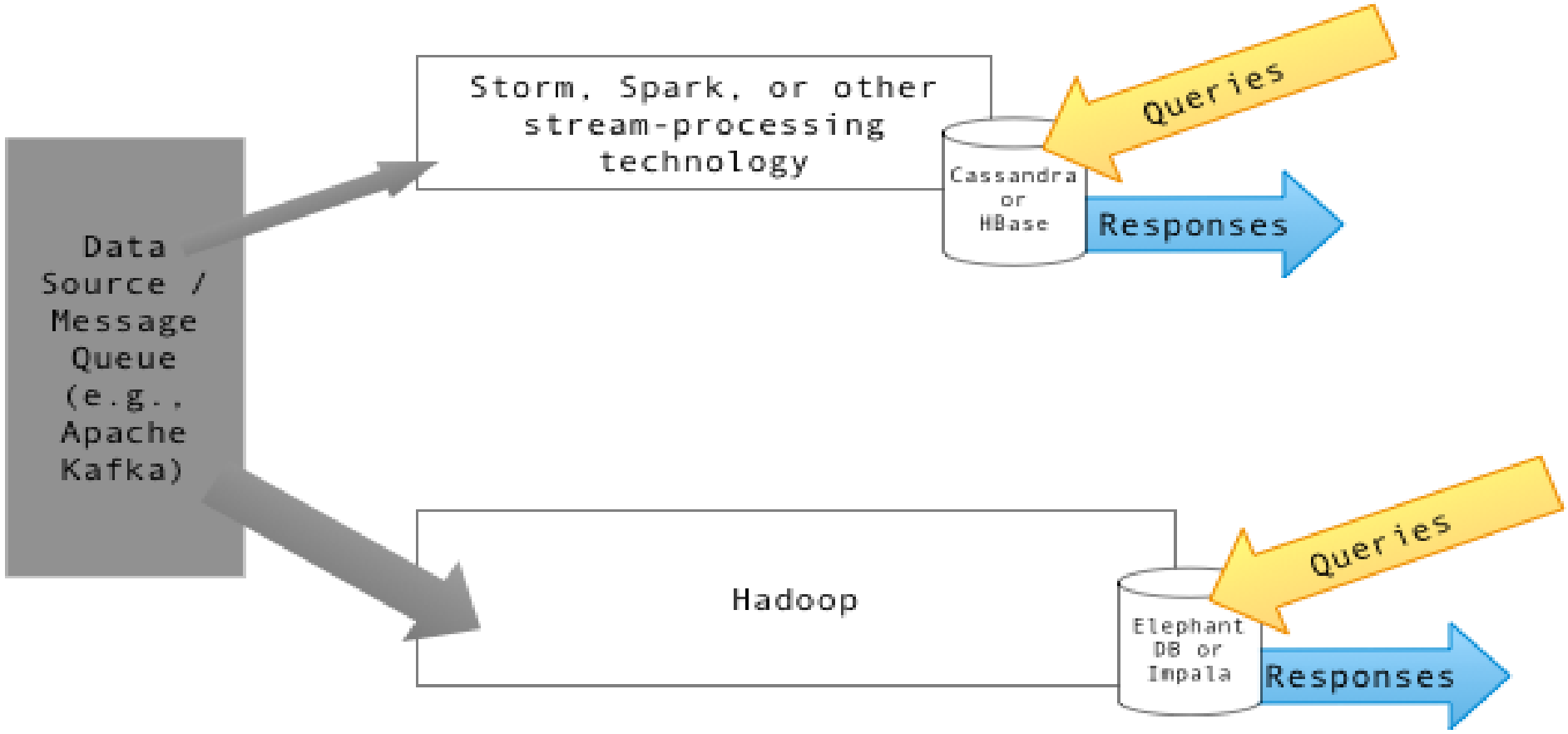


Event Sourcing – Events vs. Sate (Snapshots)



Lambda Architecture - I

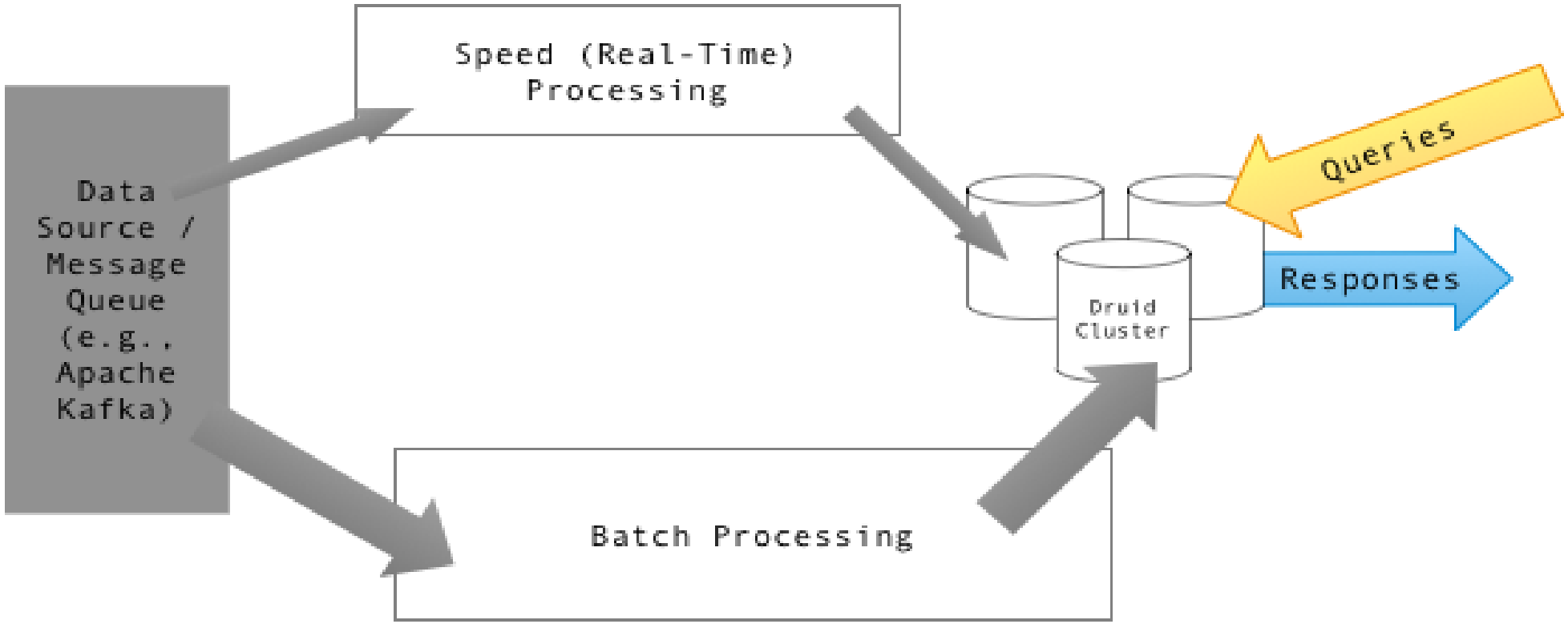
Query = λ (Complete data) = λ (live streaming data) * λ (Stored data)



https://commons.wikimedia.org/w/index.php?curid=34963987, By Textractor - Own work, CC BY-SA 4

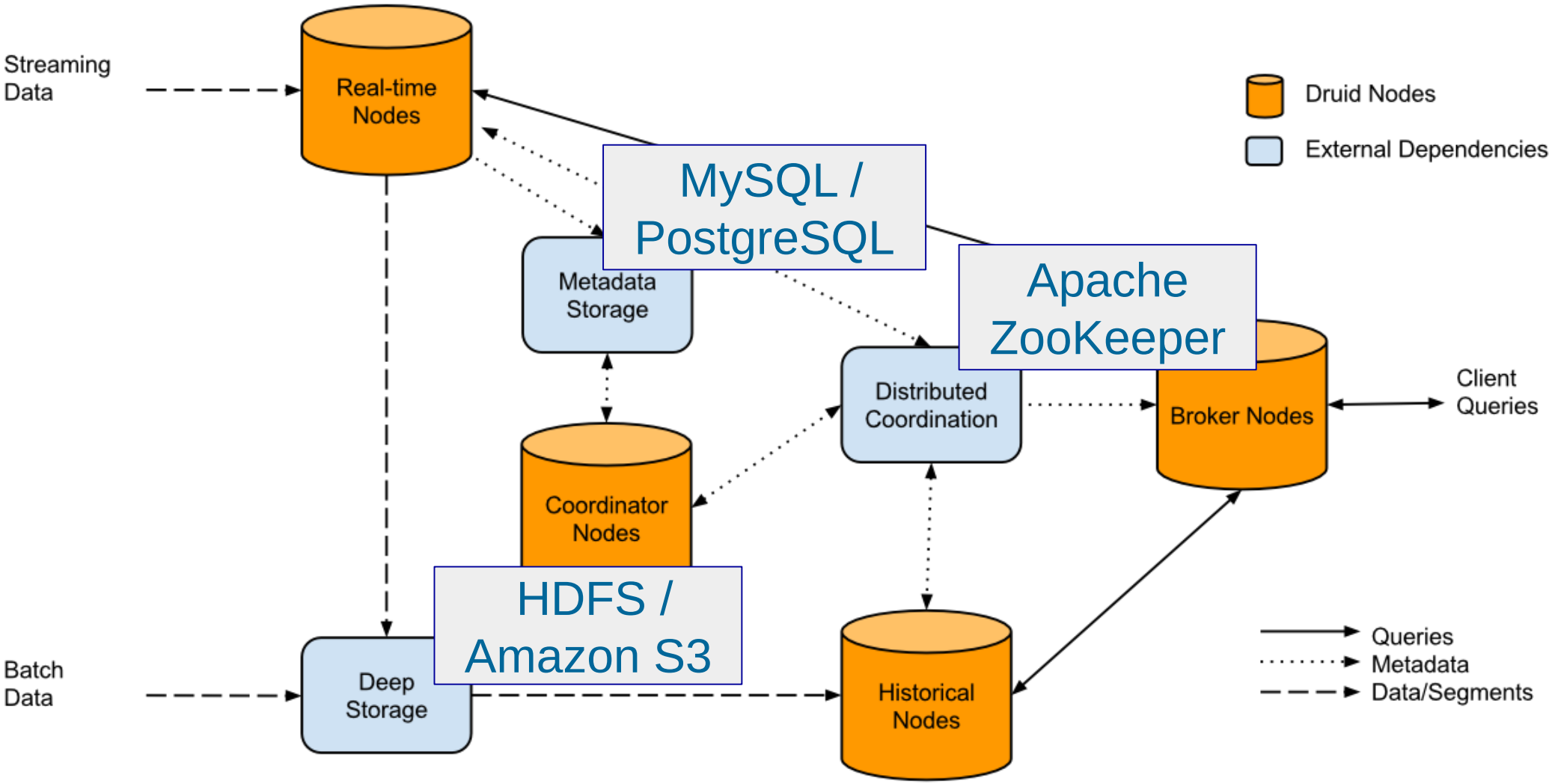
Lambda Architecture - II

Query = λ (Complete data) = λ (live streaming data) * λ (Stored data)



https://commons.wikimedia.org/w/index.php?curid=34963986, By Textractor - Own work, CC BY-SA 4

Lambda Architecture - Druid Distributed Data Store



https://commons.wikimedia.org/w/index.php?curid=33899448 By Fangjin Yang, GFDL

Kappa Architecture

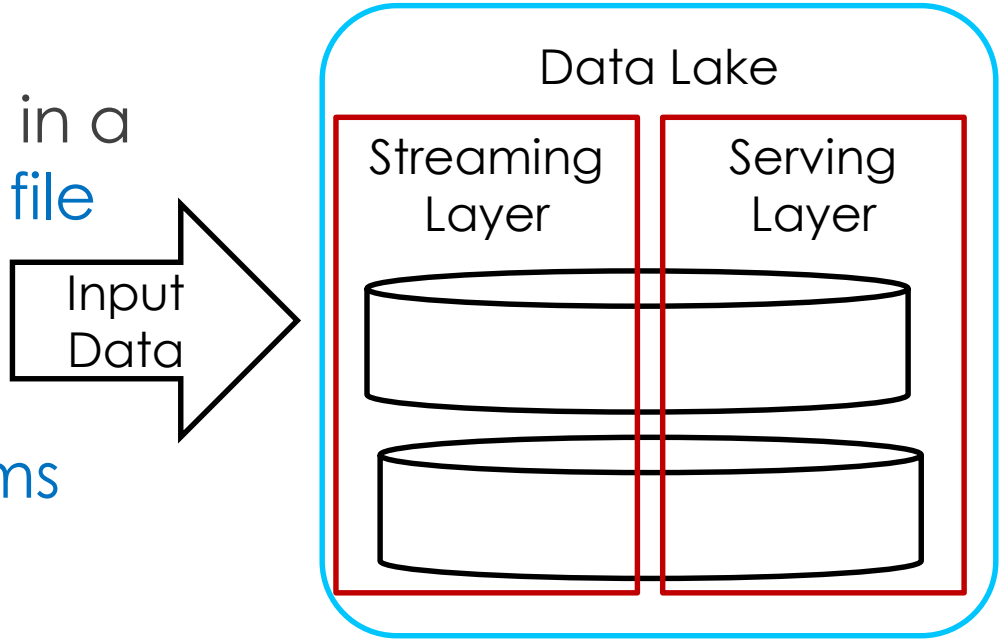
Query = K (New Data) = K (Live streaming data)

- Proposed by Jay Kreps in 2014
- Real-time processing of distinct events
- Drawbacks of Lambda architecture:
  - It can result in coding overhead due to comprehensive processing
  - Re-processes every batch cycle which may not be always beneficial
  - Lambda architecture modeled data can be difficult to migrate
- Canonical data store in a Kappa Architecture system is an append-only immutable log (like Kafka, Pulsar)

Kappa Architecture II

Query = K (New Data) = K (Live streaming data)

- Multiple data events or queries are logged in a queue to be catered against a distributed file system storage or history.
- The order of the events and queries is not predetermined. Stream processing platforms can interact with database at any time.
- It is resilient and highly available as handling terabytes of storage is required for each node of the system to support replication.
- Machine learning is done on the real time basis



Zeta Architecture

- Main characteristics of Zeta architecture:
  - file system (HDFS, S3, GoogleFS),
  - realtime data storage (HBase, Spanner, BigTable),
  - modular processing model and platform (MapReduce, Spark, Drill, BigQuery),
  - containerization and deployment (cgroups, Docker, Kubernetes),
  - Software solution architecture (serverless computing – e.g. Amazon Lambda)
- Recommender systems and machine learning
- Business applications and dynamic global resource management (Mesos + Myriad, YARN, Diego, Borg).

Distributed Stream Processing – Apache Projects:

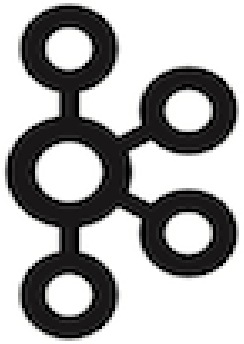
- Apache Spark is an open-source cluster-computing framework. Spark Streaming, Spark Mllib

- Apache Storm is a distributed stream processing – streams DAG

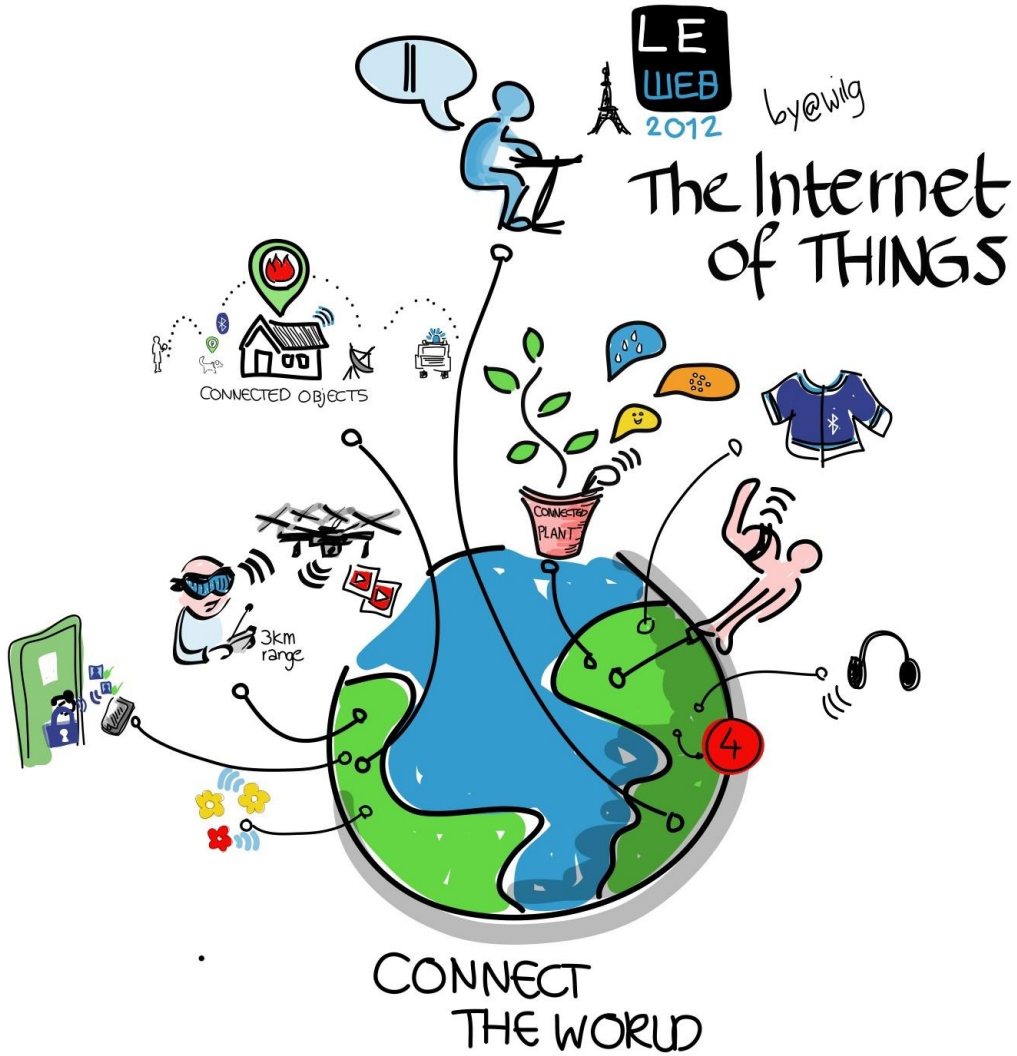
- Apache Samza is a distributed real-time stream processing framework.

Distributed Stream Processing – Apache Projects II

- Apache Flink - open source stream processing framework – Java, Scala
- Apache Kafka - open-source stream processing (Kafka Streams), real-time, low-latency, high-throughput, massively scalable pub/sub
- Apache Beam – unified batch and streaming, portable, extensible

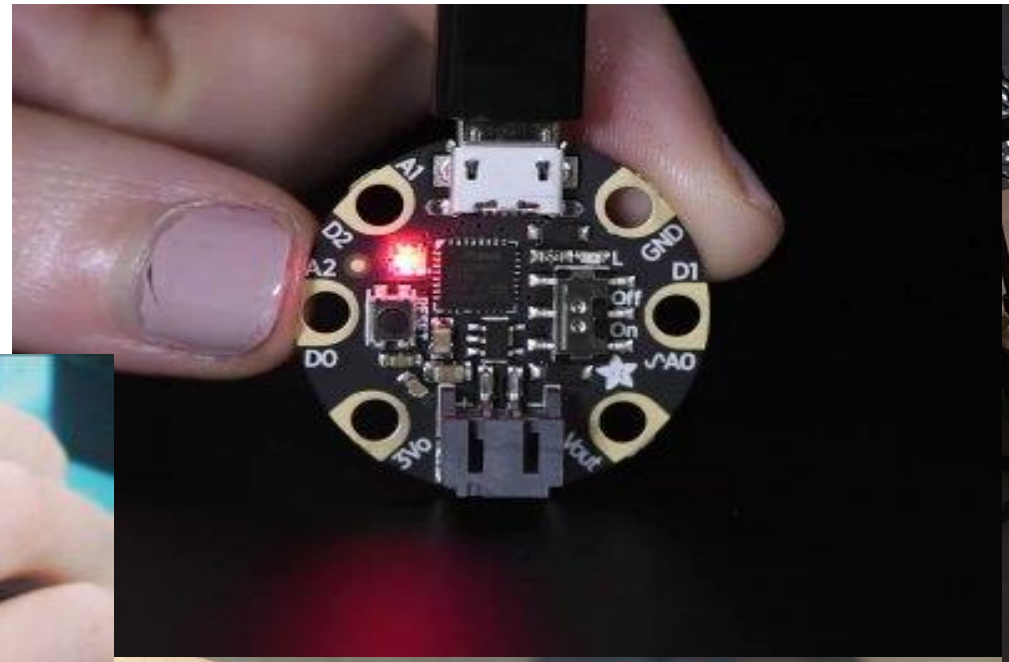


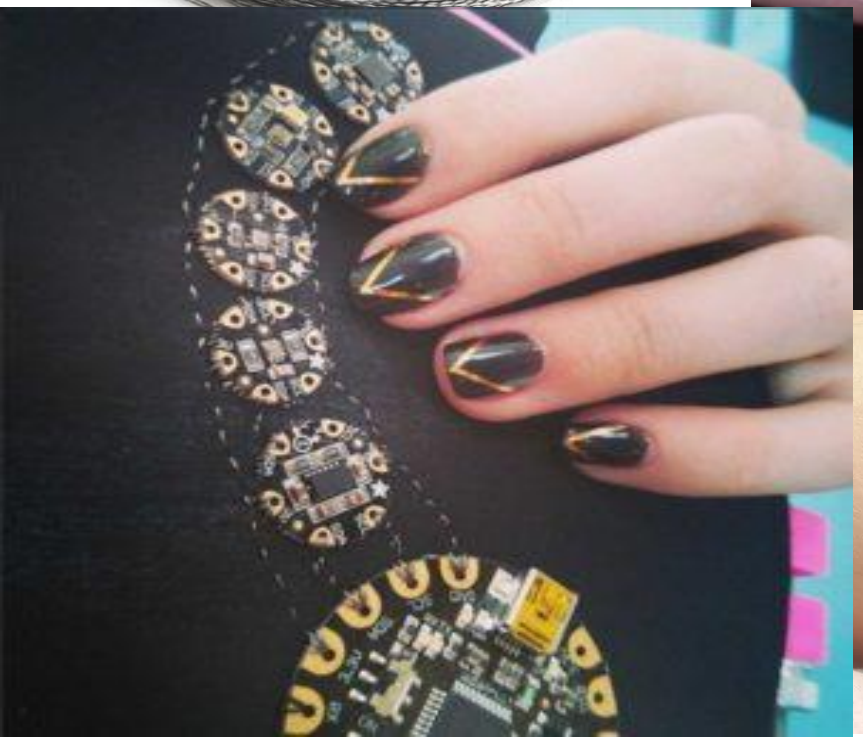
Example: Internet of Things (IoT)



CC BY 2.0, Source: https://www.flickr.com/photos/wilgengebroed/8249565455/ ,
Radar, GPS, lidar for navigation and obstacle avoidance (2007 DARPA Urban Challenge)

Wearable Electronics :)





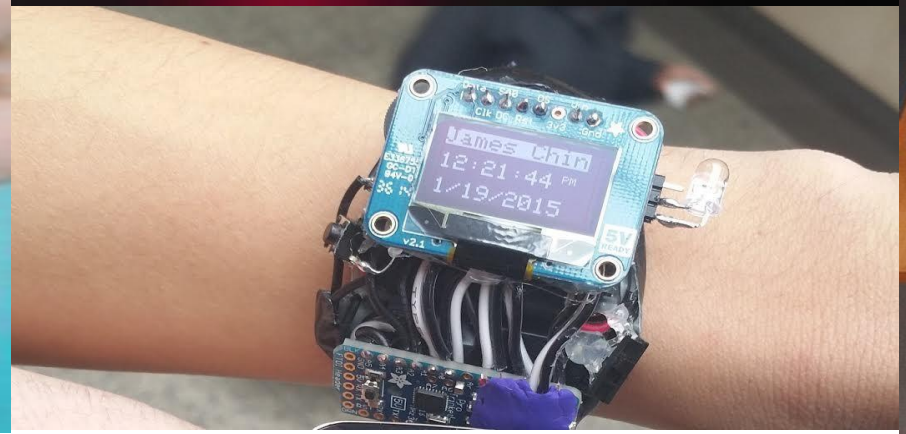
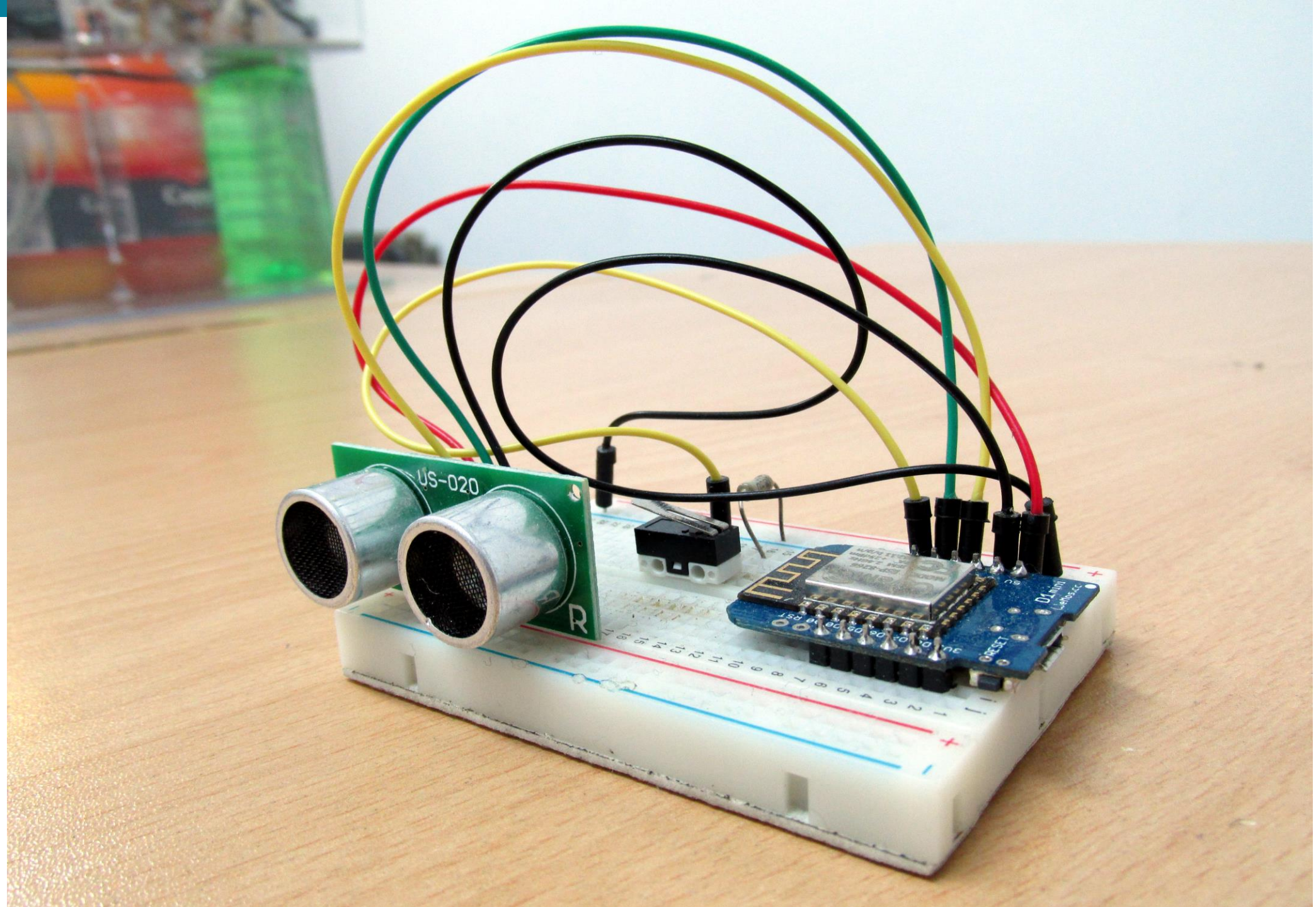
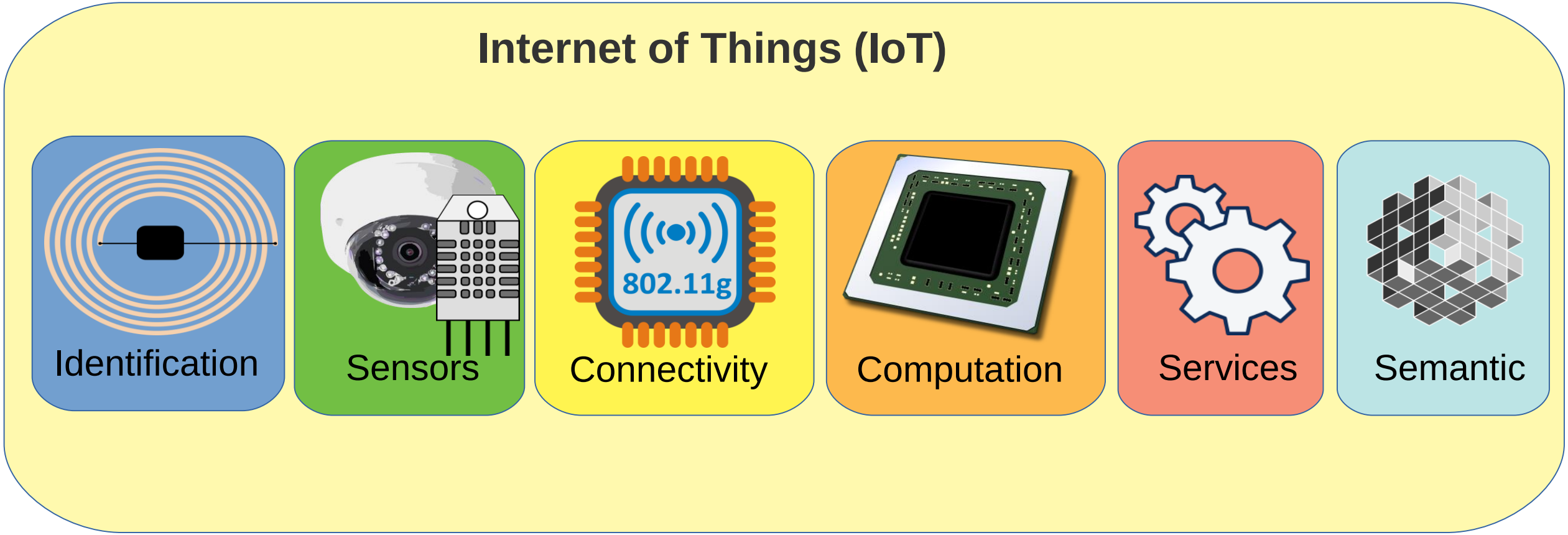


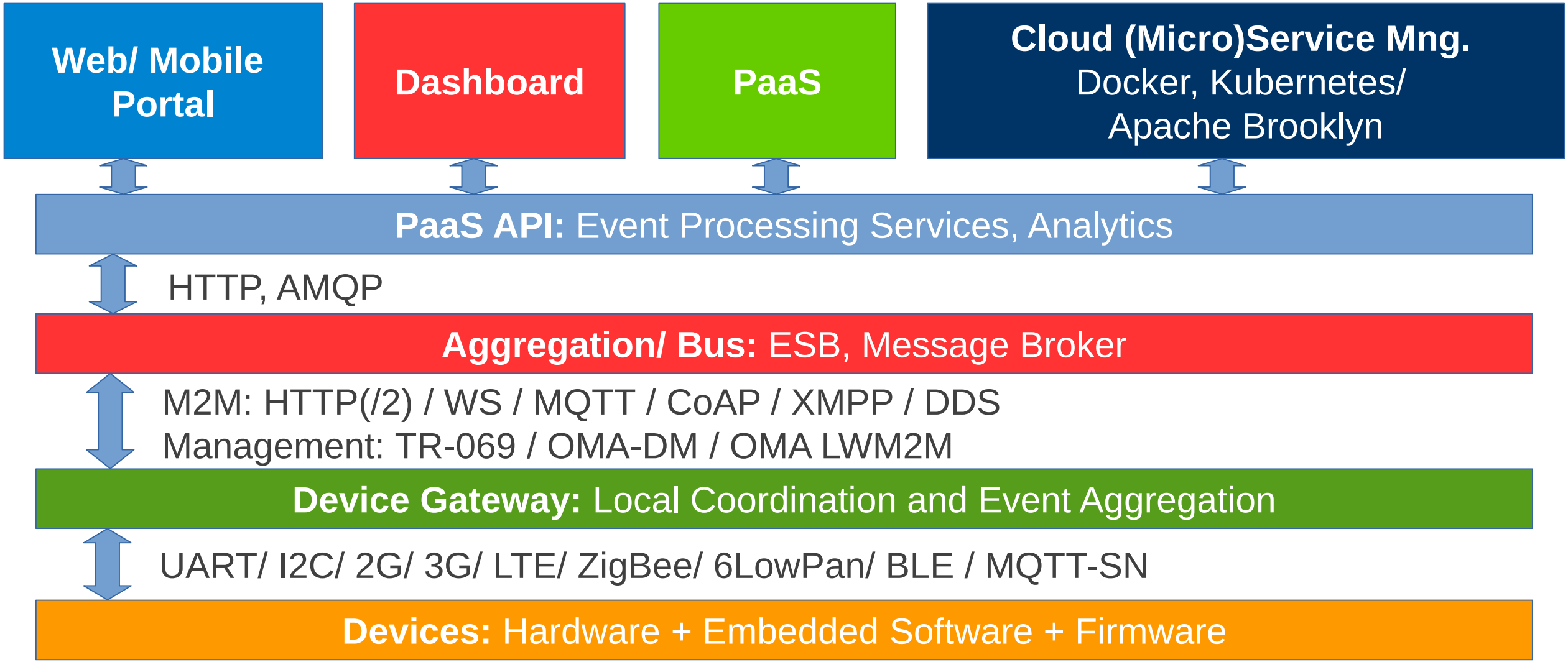
Photo credit: Adafruit, Becky Stern and others


US-020
R
D1 mini
RESET

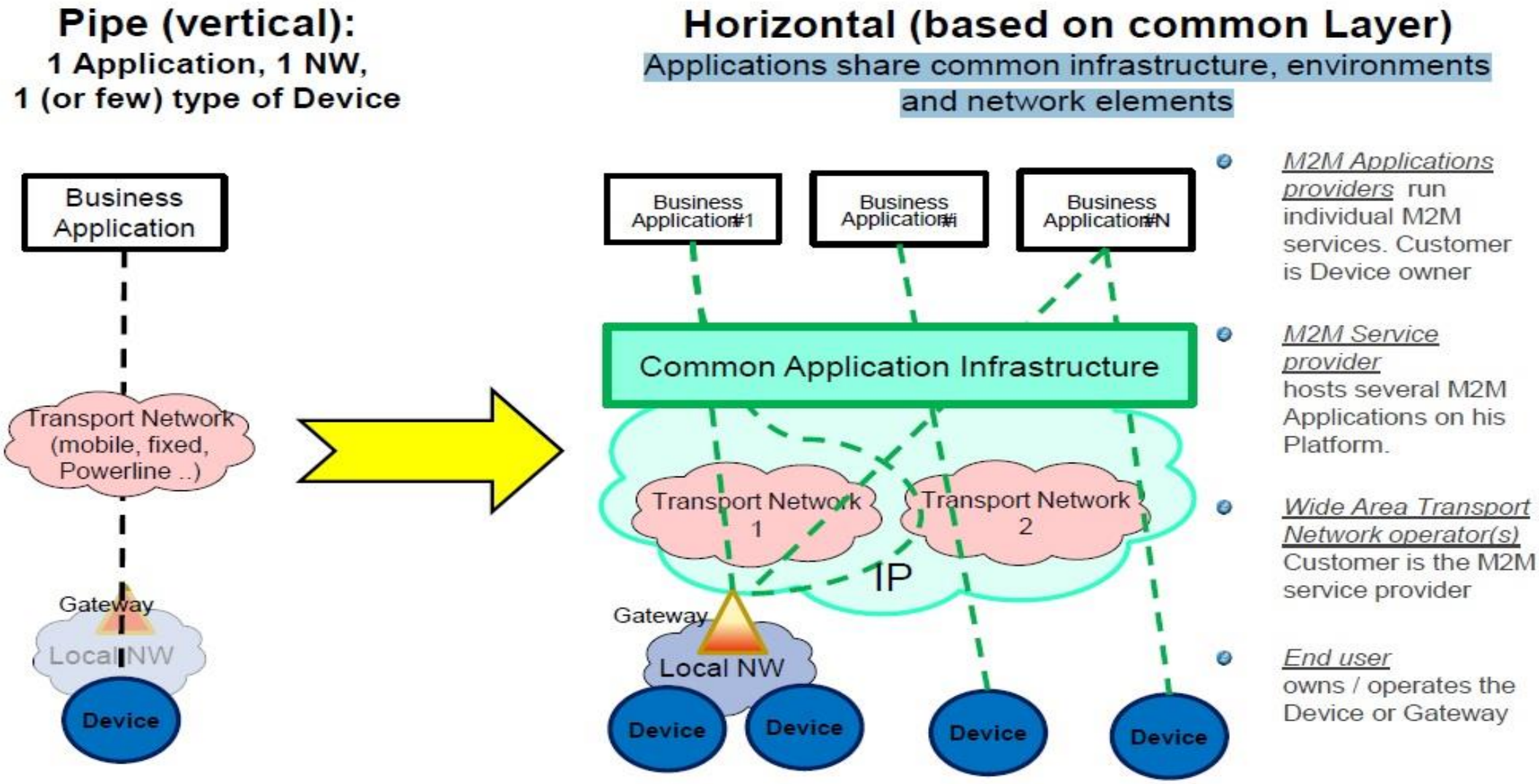
Key Elements of IoT



IoT Services Architecture



Vertical vs. Horizontal IoT



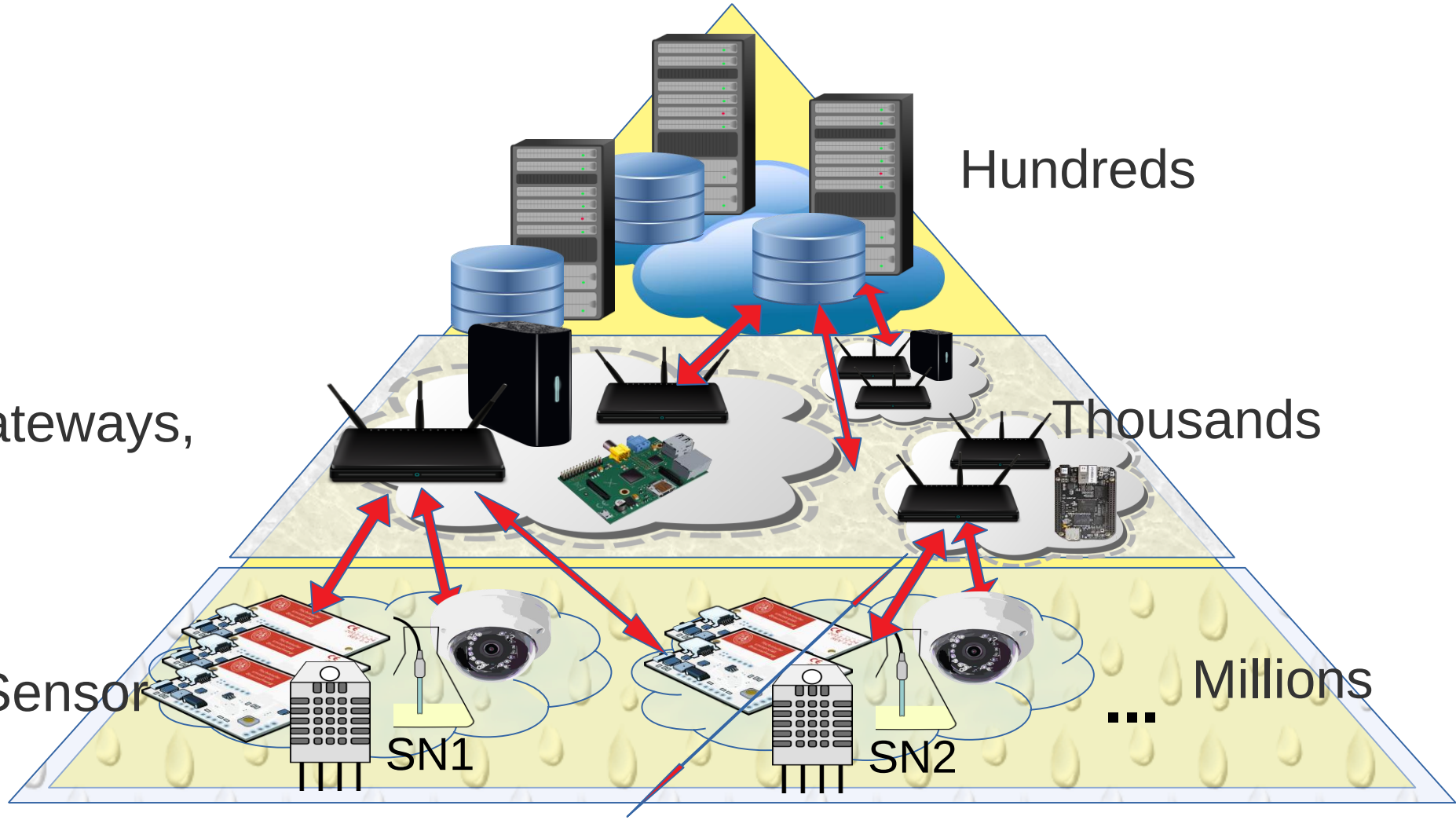
http://www.iot-week.eu/overview_2013/iot-week-2012/programme-1/tuesday-1/presentations/semantic-interoperability/K1_ETSI-M2M-oneM2M-and-the-need-for-semantics-IoTWeek2012.pdf, By Joerg Swetina (NEC)- ETSI M2M / oneM2M and the need for semantics

Cloud, Fog and Mist Computing

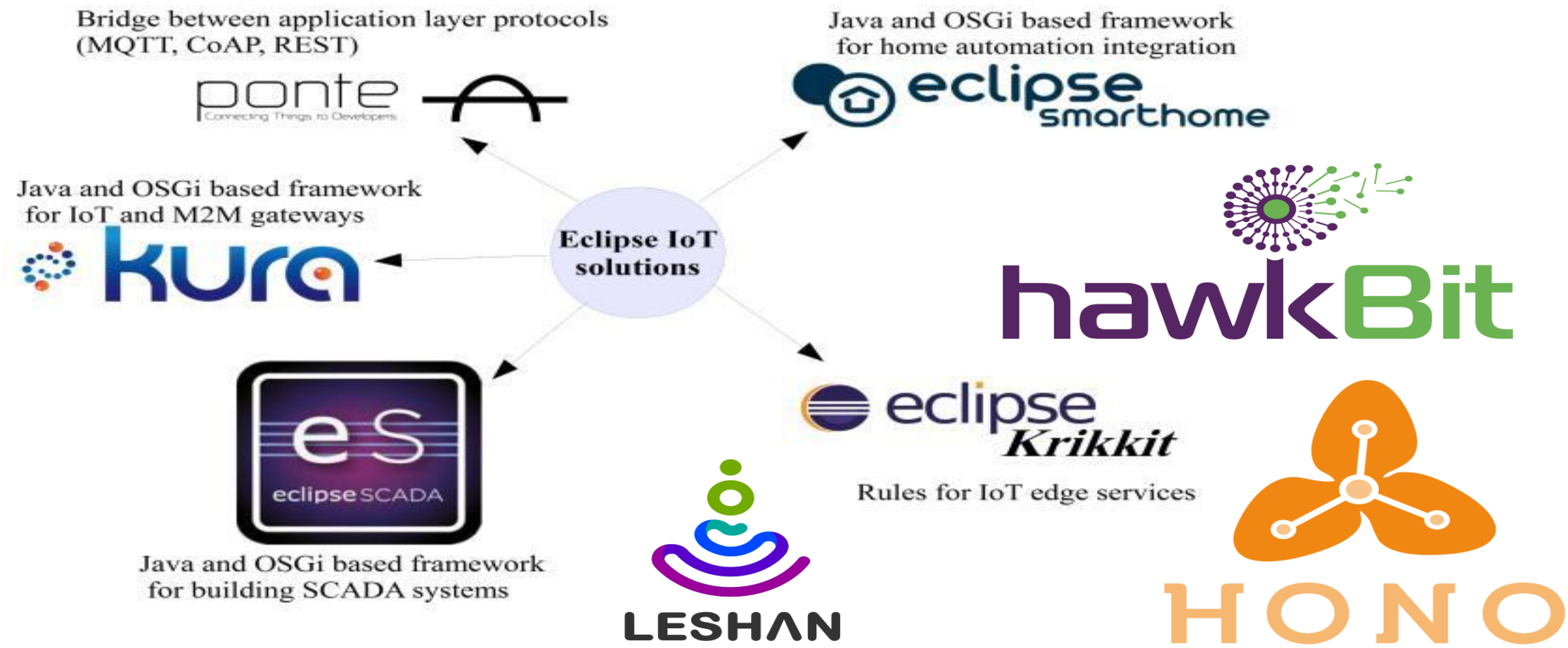
Cloud Computing
(Data-centers)

Fog Computing
(Fog Nodes, Edge Gateways, Cloudlets)

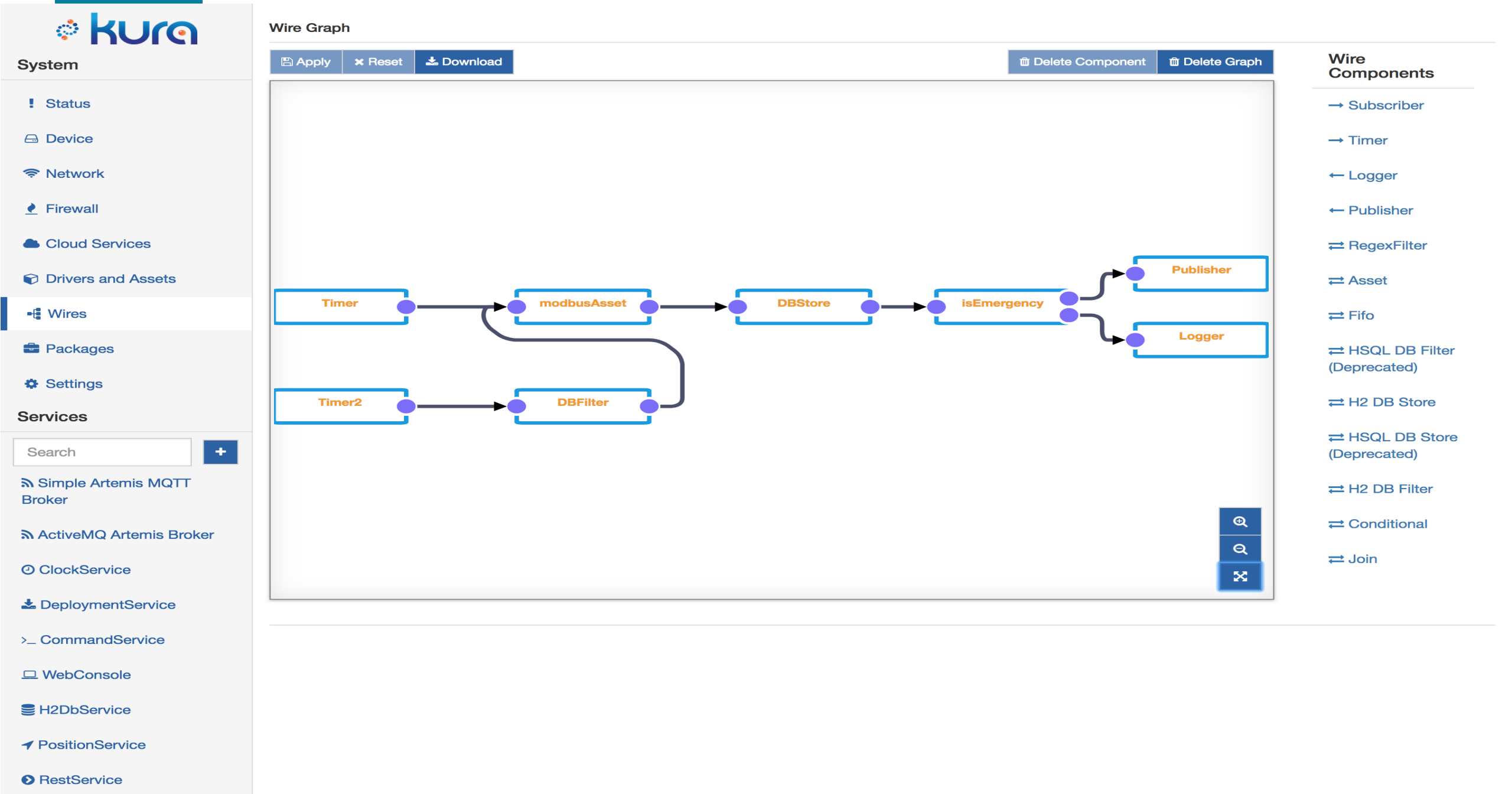
Mist Computing
(Smart IoT Devices, Sensor Networks)



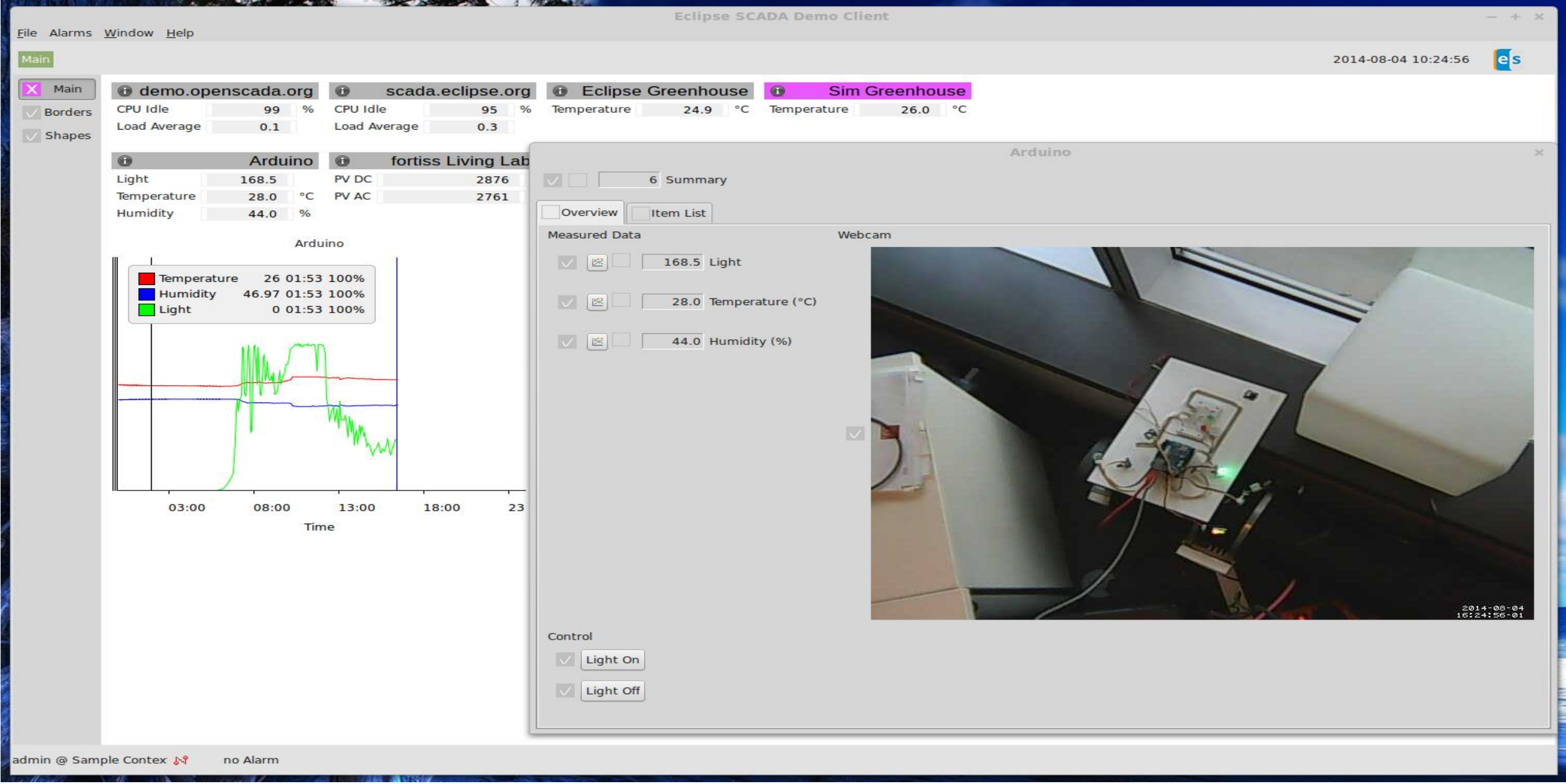
Eclipse IoT Platform



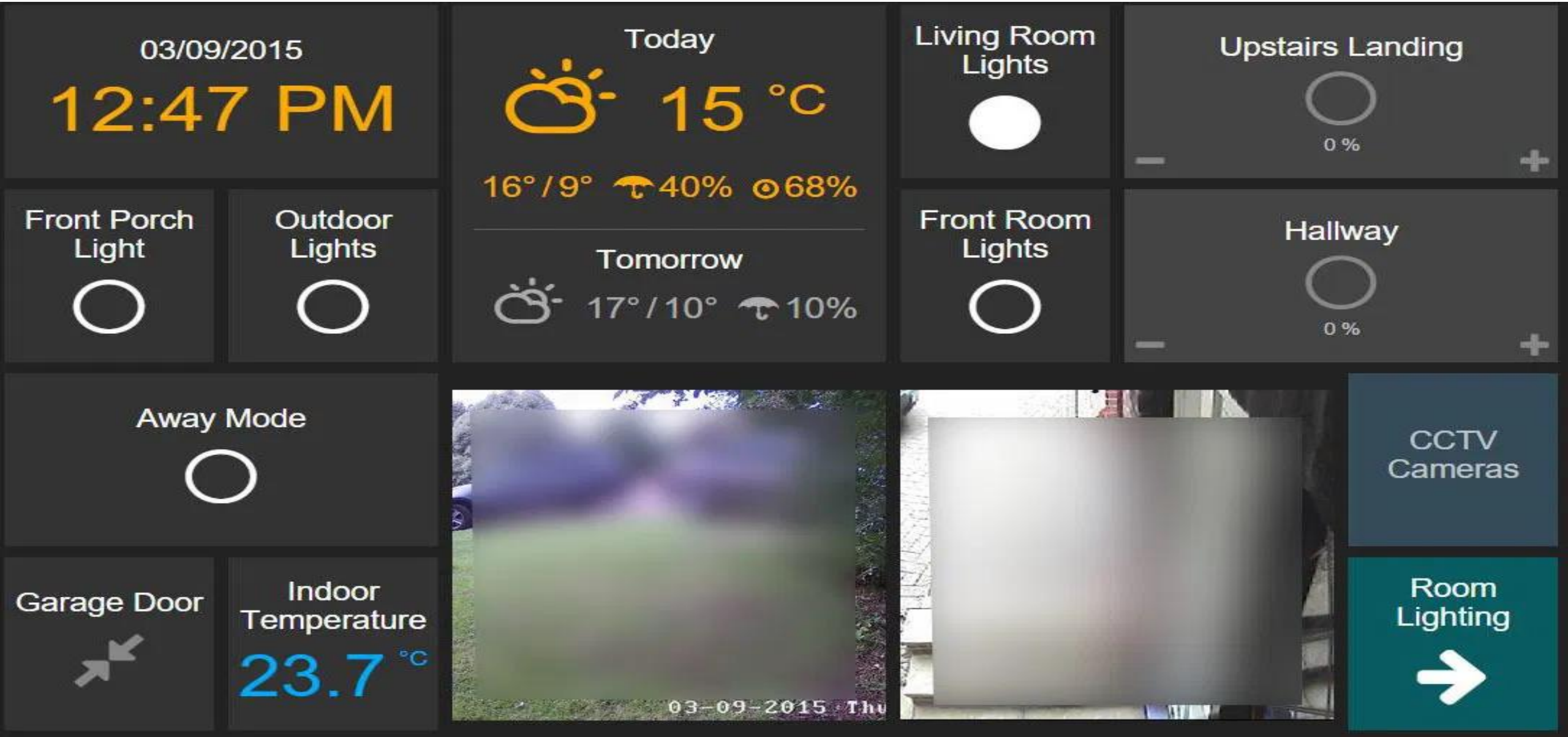
Based on: https://www.researchgate.net/publication/279177017_Internet_of_Things_A_Survey_on_Enabling_Technologies_Protocols_and_Applications, By Ala Al-Fuqaha et al. - Internet of Things: A Survey on Enabling Technologies, Protocols and Applications



Source: https://www.eclipse.org/kura/



Eclipse Smart Home – Open HAB



01/09/2017

1:40 AM

57.4 °F

Humidity: 52 %

Pressure: 30.12 in

Today	Tuesday	Wednesday
		
Mostly Cloudy	Clear	Clear
75 °F / 46 °F	70 °F / 45 °F	66 °F / 45 °F

Utility

Hallway

Family Room

Dining

Kitchen

Stereo Volume

17%

Stereo

34%

73.0 °F

Office: 77°

70

Front Yard

Porch

Driveway

Garage

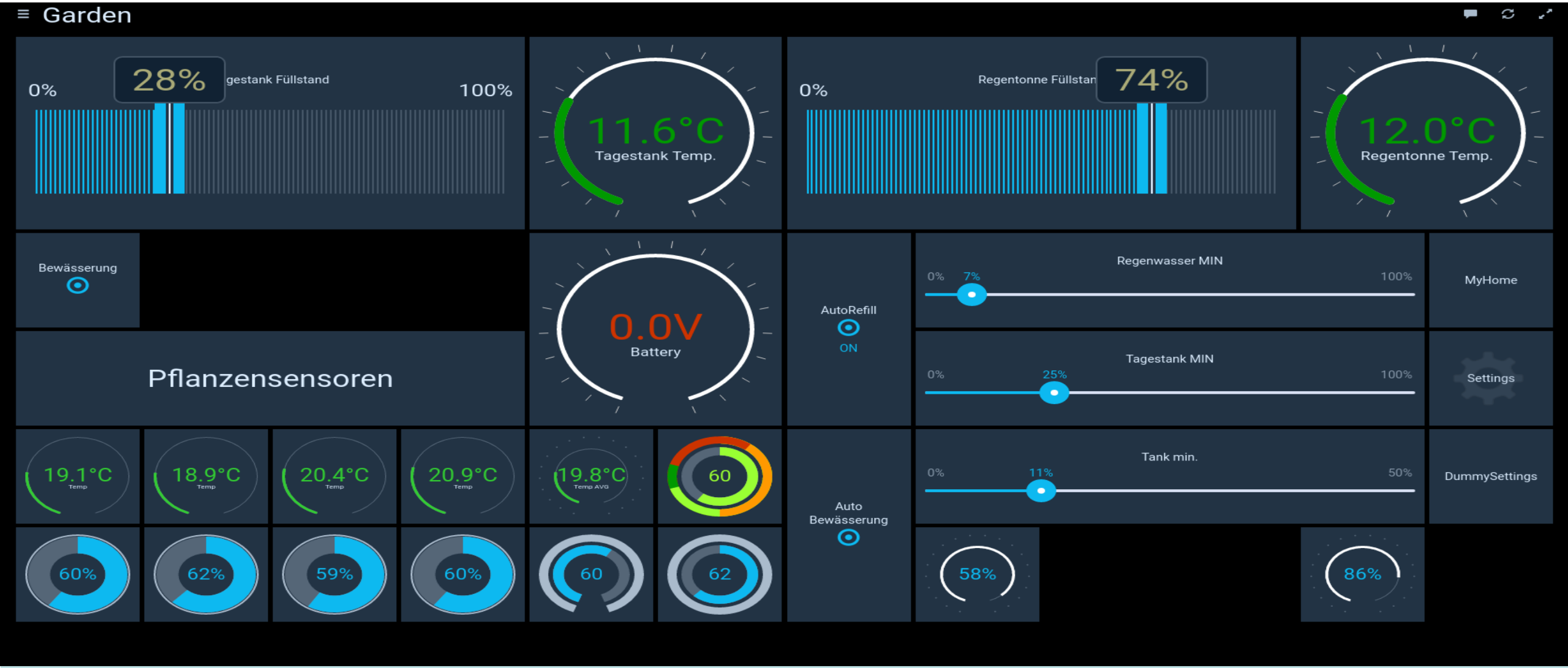
Living Room

Living Room Lights

Back Yard

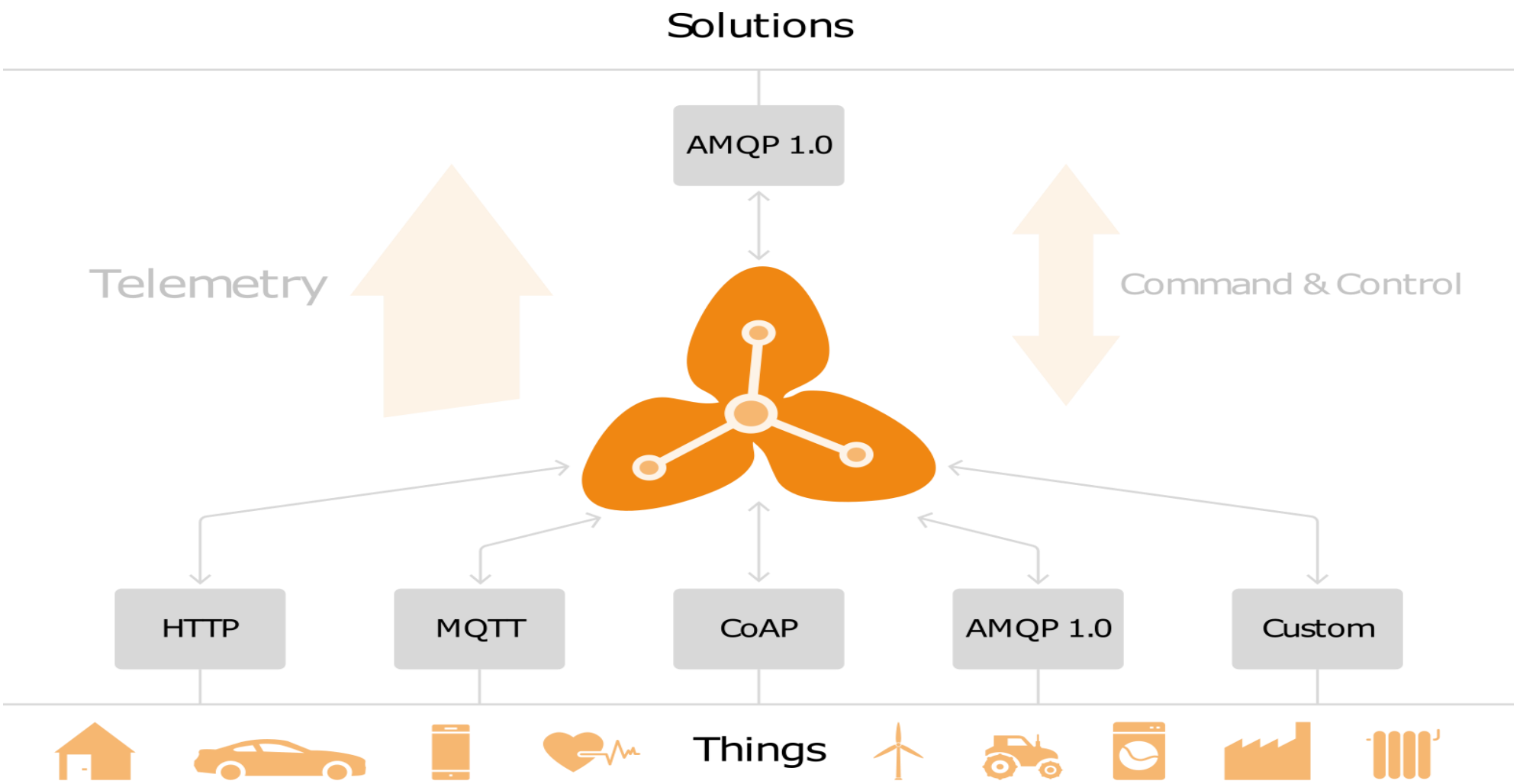
Wake Up

Open HAB: Hydroponics Dashboard



Source: https://community.openhab.org/t/aquaponics-goes-openhab/26456

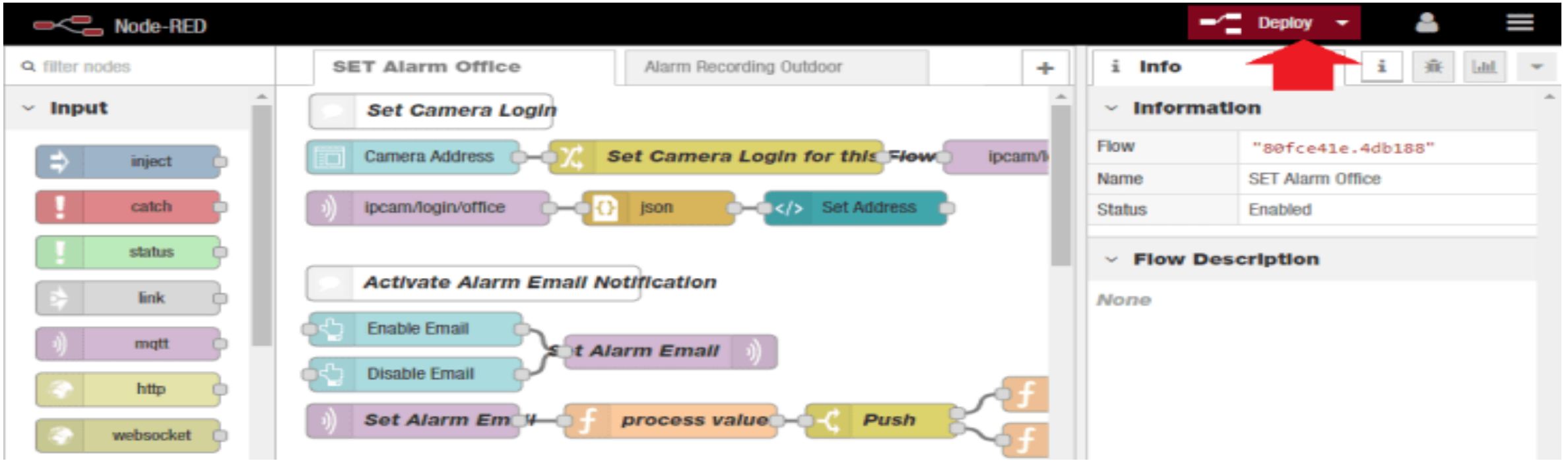
Eclipse Hono: Connect. Command. Control.



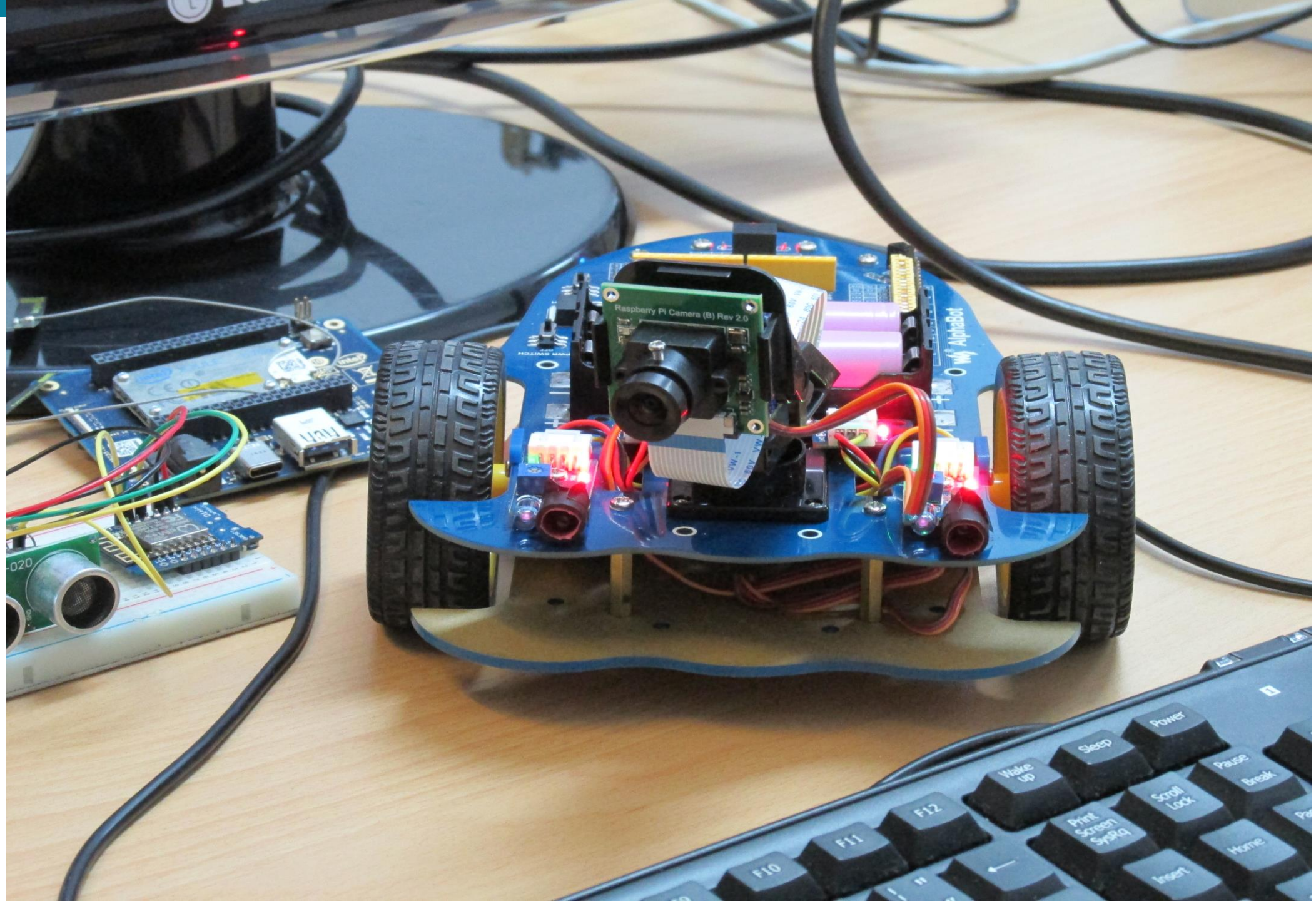
Source: https://www.eclipse.org/hono/

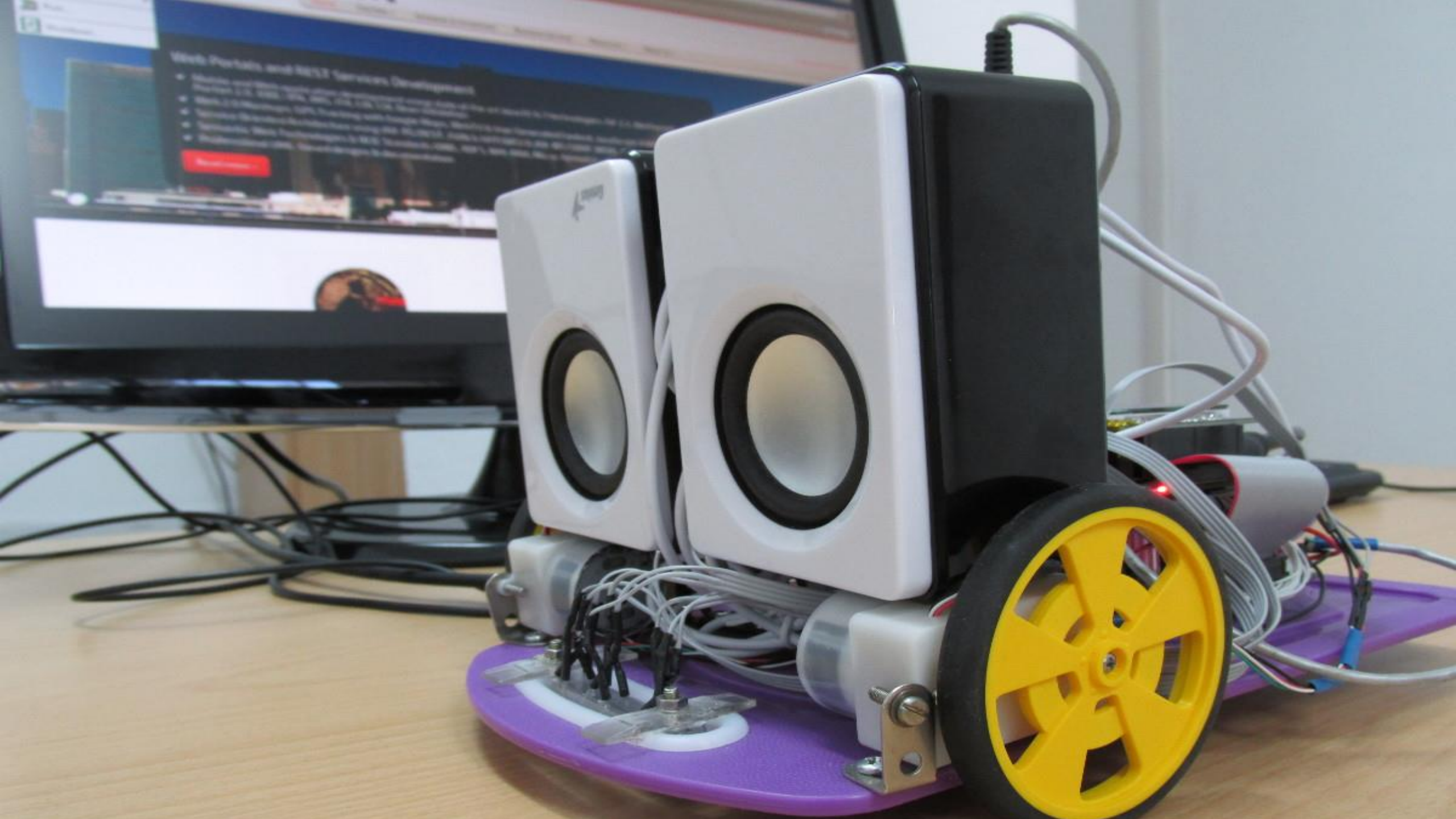
Node Red

Node-RED is a flow-based development tool for visual programming developed originally by IBM for wiring together hardware devices, APIs and online services as part of the Internet of Things.

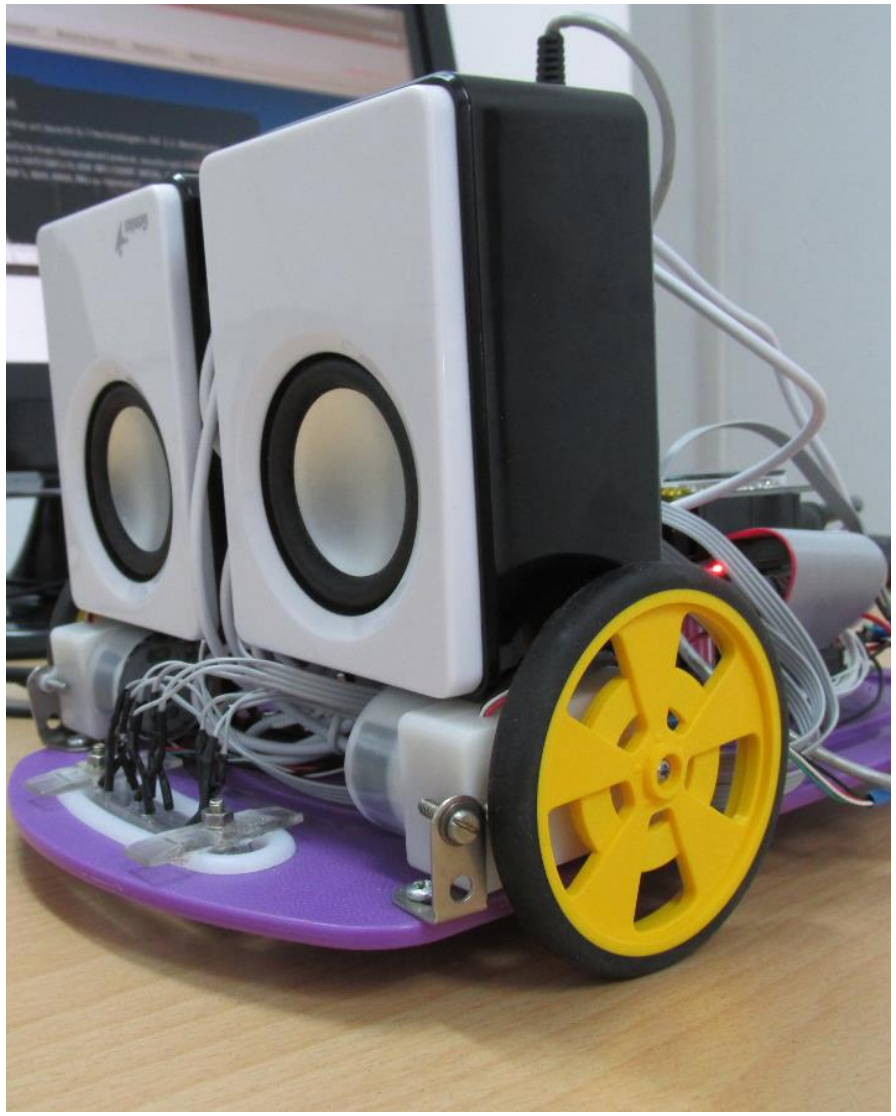


Source: https://nodered.org/


Raspberry Pi Camera (B) Rev 2.0
AlphaBot

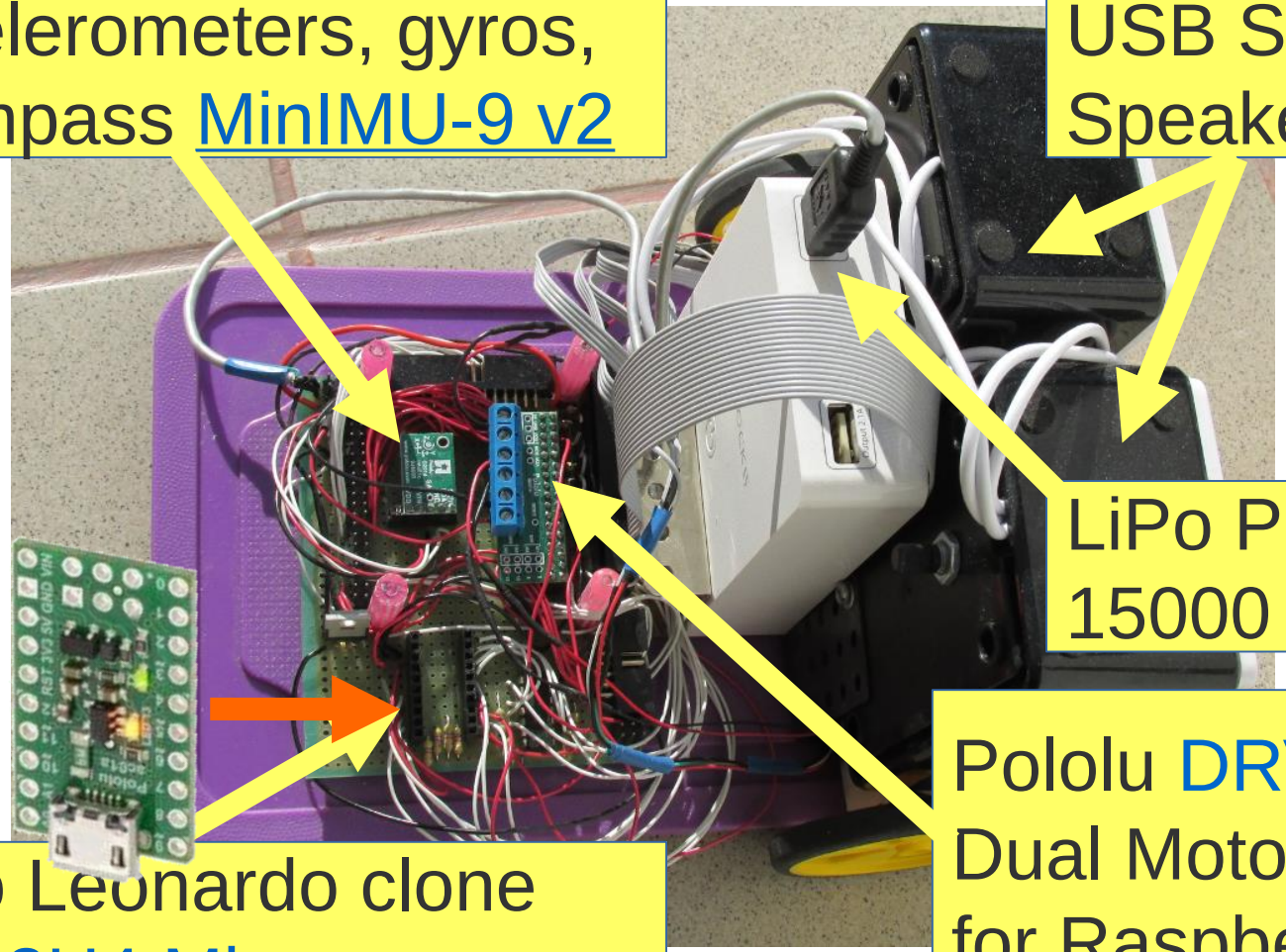


IPTPI: Raspberry Pi + Ardunio Robot



- Raspberry Pi 2 (quad-core ARMv7 @ 900MHz) + Arduino Leonardo clone A-Star 32U4 Micro
- *Optical encoders* (custom), IR optical array, 3D accelerometers, gyros, and compass MinIMU-9 v2
- **IPTPI** is programmed in Python, Java and Go using: Wiring Pi, Numpy, Pandas, Scikit-learn, Pi4J, Reactor, RxJava, GPIO(Go)

IPTPI: Raspberry Pi + Ardunio Robot



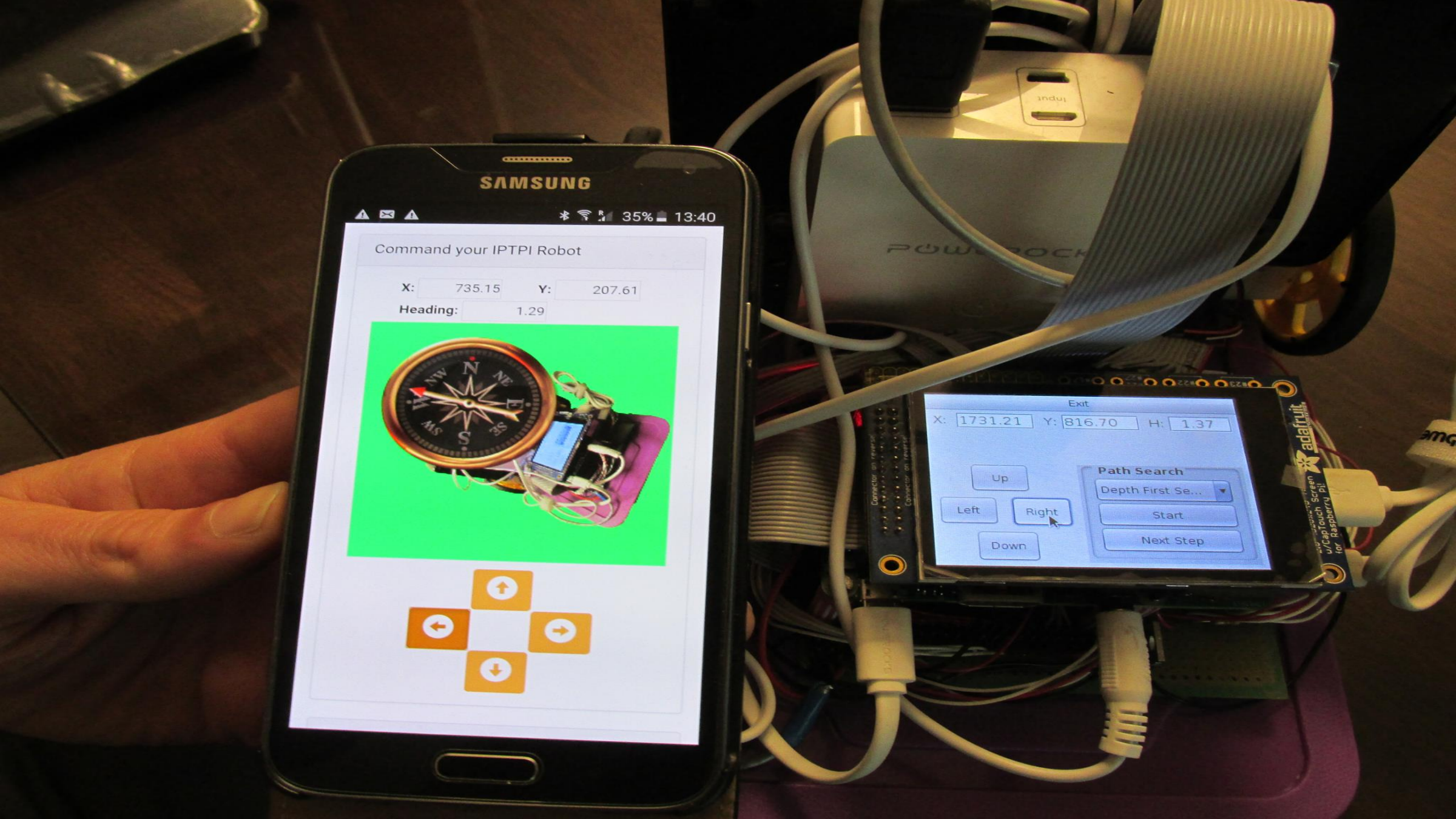
3D accelerometers, gyros, and compass MinIMU-9 v2

USB Stereo Speakers - 5V

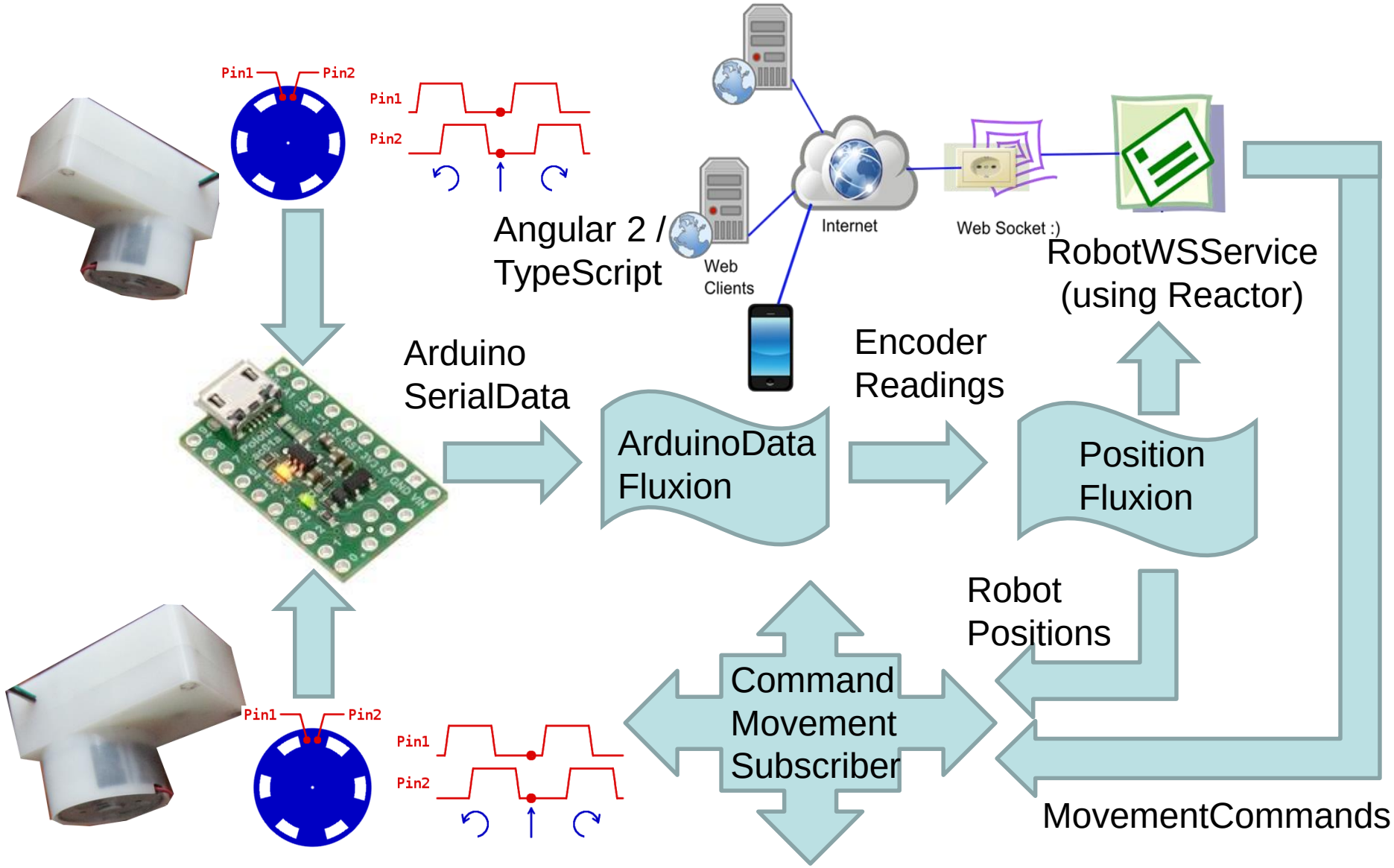
LiPo Powebank 15000 mAh

Pololu DRV8835 Dual Motor Driver for Raspberry Pi

Arduino Leonardo clone A-Star 32U4 Micro


SAMSUNG
35% 13:40
Command your IPTPI Robot
X: 735.15
Y: 207.61
Heading: 1.29
Exit
X: 1731.21
Y: 816.70
H: 1.37
Up
Left
Right
Down
Path Search
Depth First Se...
Start
Next Step

IPTPI Reactive Streams



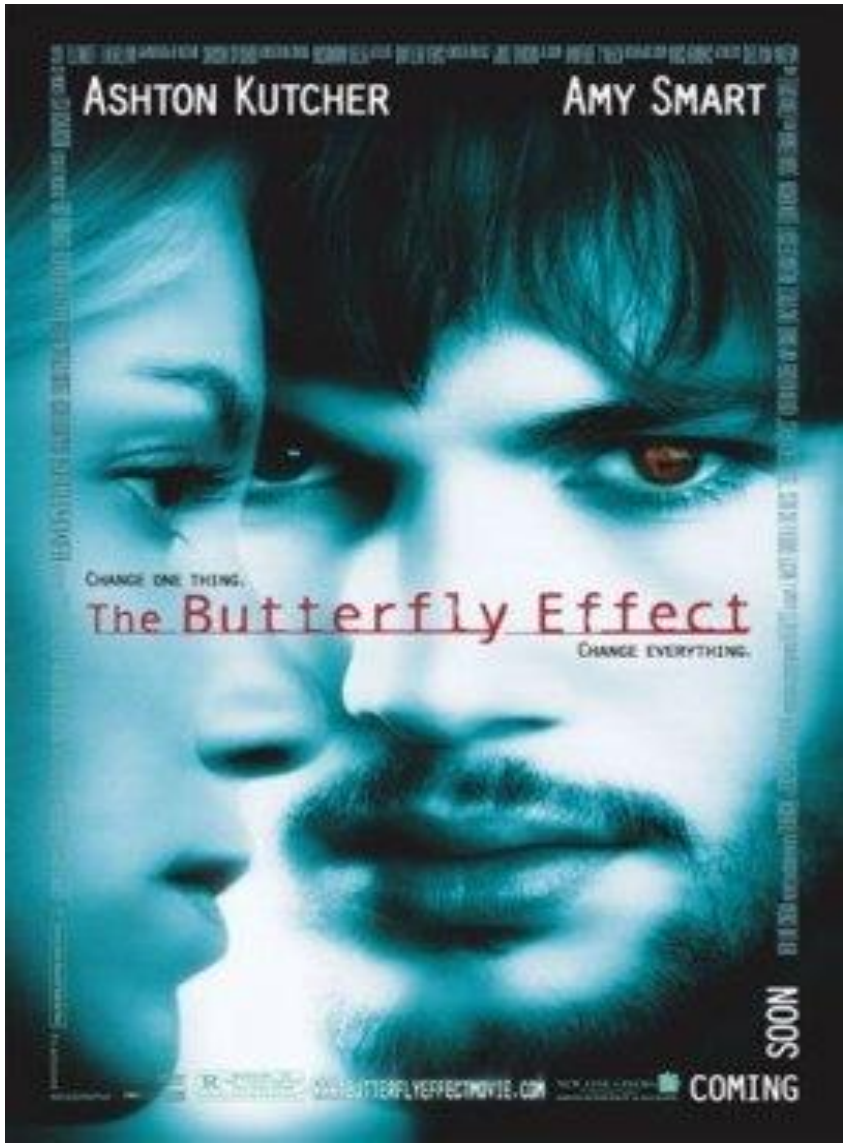
Functional Reactive Programming

Imperative and Reactive

We live in a Connected Universe

... there is hypothesis that all the things in the Universe are intimately connected, and you can not change a bit without changing all.

Action – Reaction principle is the essence of how Universe behaves.



Source: https://en.wikipedia.org/wiki/The_Butterfly_Effect#/media/File:Butterflyeffect_poster.jpg
Fair use, File:Butterflyeffect poster.jpg, Uploaded by Yaminator, Uploaded: 8 July 2008

Imperative vs. Reactive

- Reactive Programming: using static or dynamic data flows and propagation of change
- Example: a := b + c
- Functional Programming: evaluation of mathematical functions, avoids changing-state and mutable data, declarative programming
- Side effects free => much easier to understand and predict the program behavior.

Ex. (RxGo):

```
observable := rxgo.Just("Hello", "Reactive", "World", "from", "RxGo")().
        Map(ToUpper). // map to upper case
        Filter(LengthGreaterThan4) // greaterThan4 func filters values > 4
for item := range observable.Observe() {
        fmt.Println(item.V)
}
```

Functional Reactive Programming (FRP)

- According to Connal Elliot's (ground-breaking paper at Conference on Functional Programming, 1997), FRP is:

  (a) Denotative

  (b) Temporally continuous

- FRP is asynchronous data-flow programming using the building blocks of functional programming (map, reduce, filter, etc.) and explicitly modeling time

Reactive Programming

- ReactiveX (Reactive Extensions) - open source polyglot (http://reactivex.io):

  Rx = Observables + Flow transformations + Schedulers

- Go: RxGo, Kotlin: RxKotlin, Java: RxJava, JavaScript: RxJS, Python: RxPY, C#: Rx.NET, Scala: RxScala, Clojure: RxClojure, C++: RxCpp, Ruby: Rx.rb, Python: RxPY, Groovy: RxGroovy, JRuby: RxJRuby, ...
- Reactive Streams Specification (http://www.reactive-streams.org/):
  - Publisher – provider of potentially unbounded number of sequenced elements, according to Subscriber(s) demand (backpressure): onNext* (onError | onComplete)
  - Subscriber, Subscription
  - Processor = Subscriber + Publisher

ReactiveX: Observable vs. Iterable

Example code showing how similar high-order functions can be applied to an Iterable and an Observable

Iterable

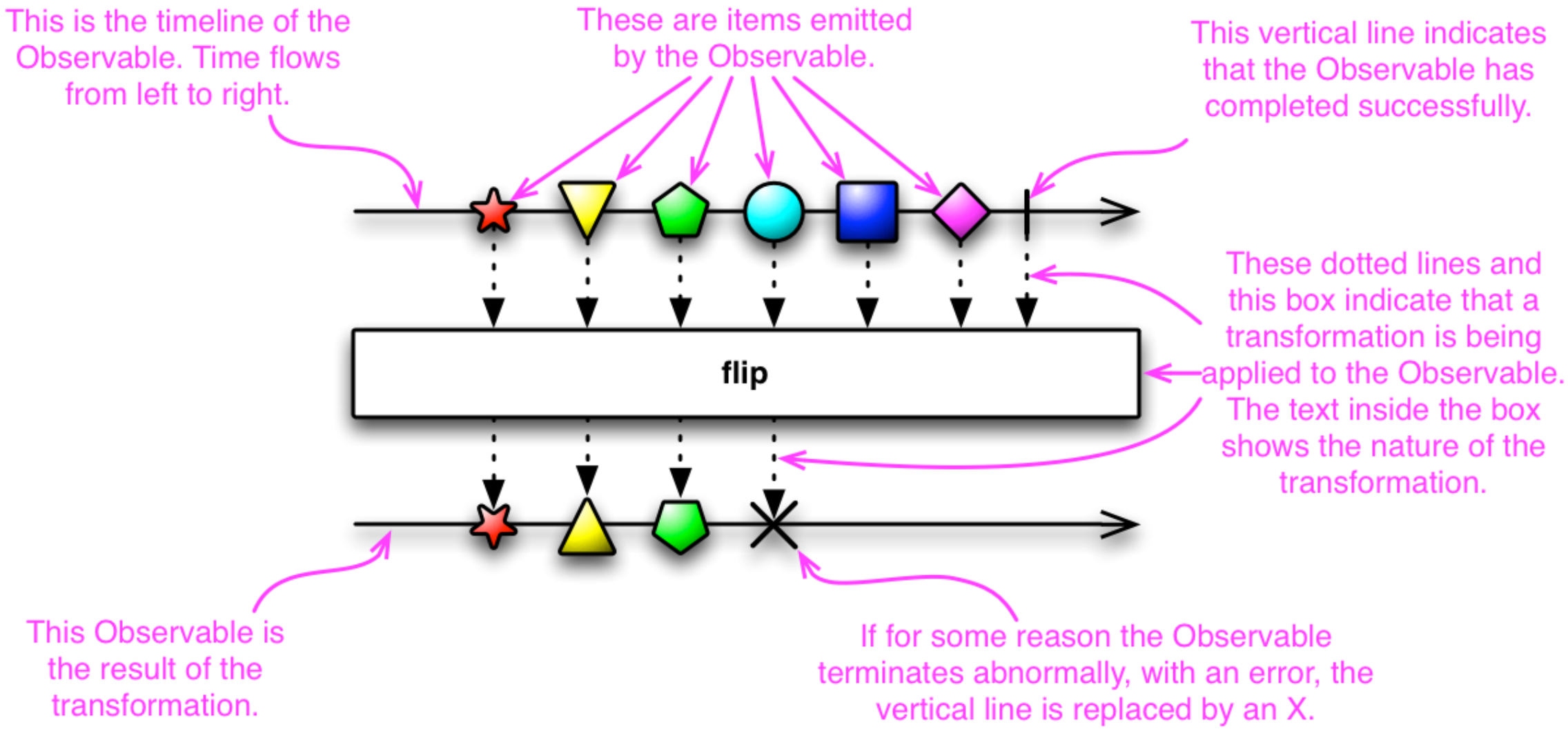
```
getDataFromLocalMemory()
  .skip(10)
  .take(5)
  .map({ s -> return s + " transforme
d" })
  .forEach({ println "next => " + it
})
```

Observable

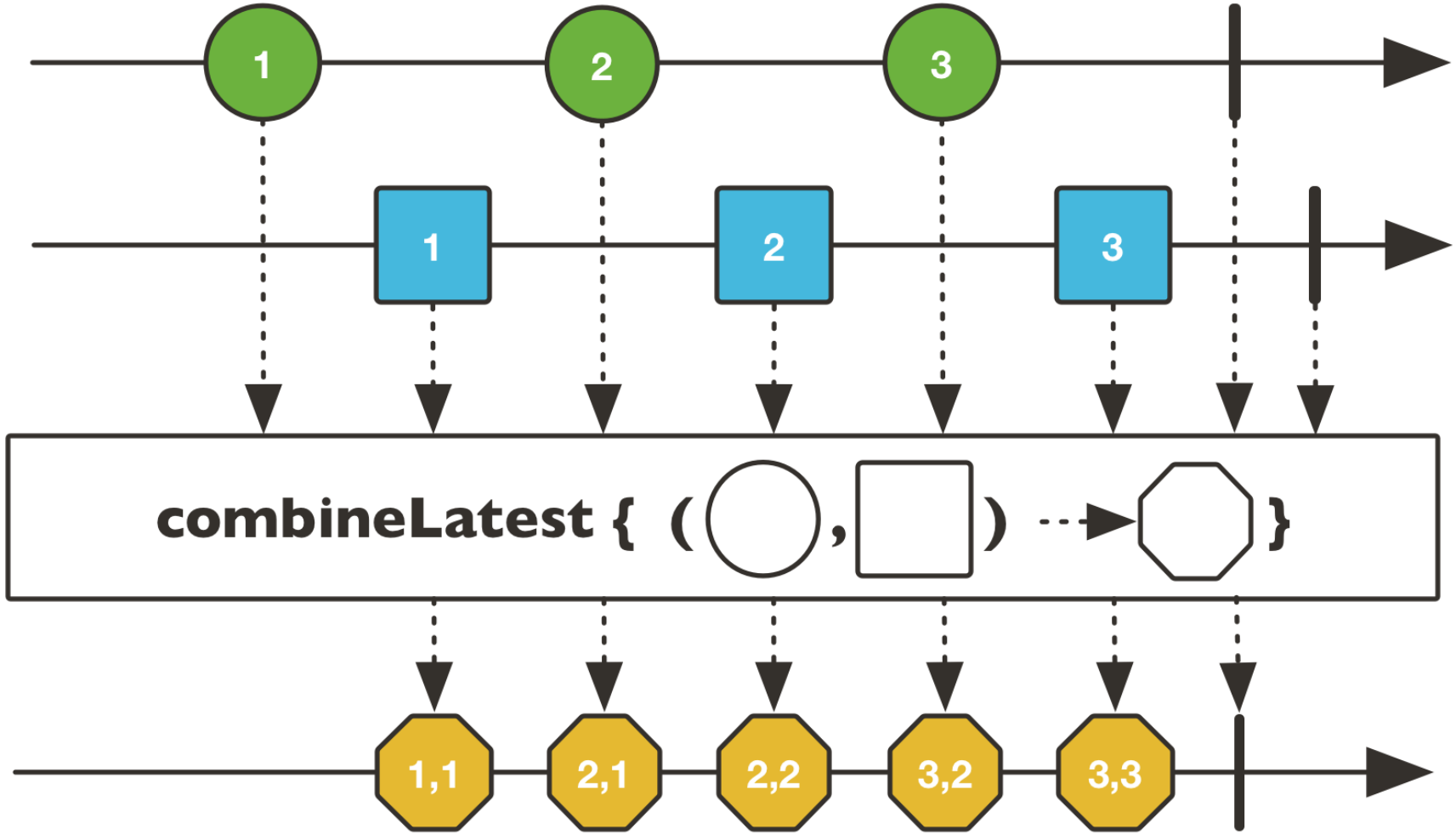
```
getDataFromNetwork()
  .skip(10)
  .take(5)
  .map({ s -> return s + " transformed"
})
  .subscribe({ println "onNext => " + it
})
```

You can think of the Observable class as a **"push"** equivalent to Iterable, which is a **"pull."** With an Iterable, the consumer pulls values from the producer and the thread blocks until those values arrive. By contrast, with an Observable the producer pushes values to the consumer whenever values are available. This approach is more flexible, because values can arrive synchronously or asynchronously.

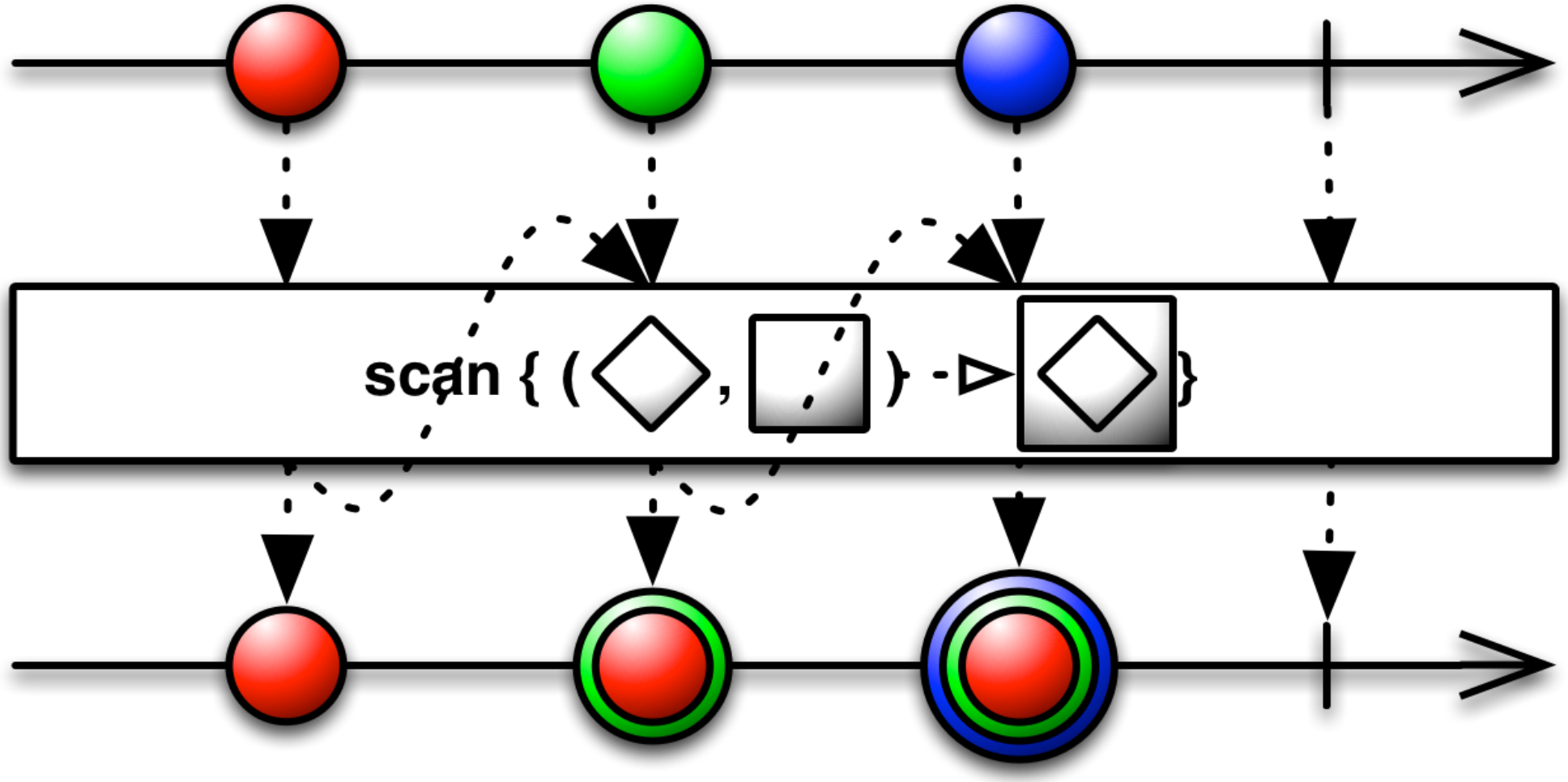
ReactiveX Observable – Marble Diagrams



Example: CombineLatest



Redux == Rx Scan Opearator



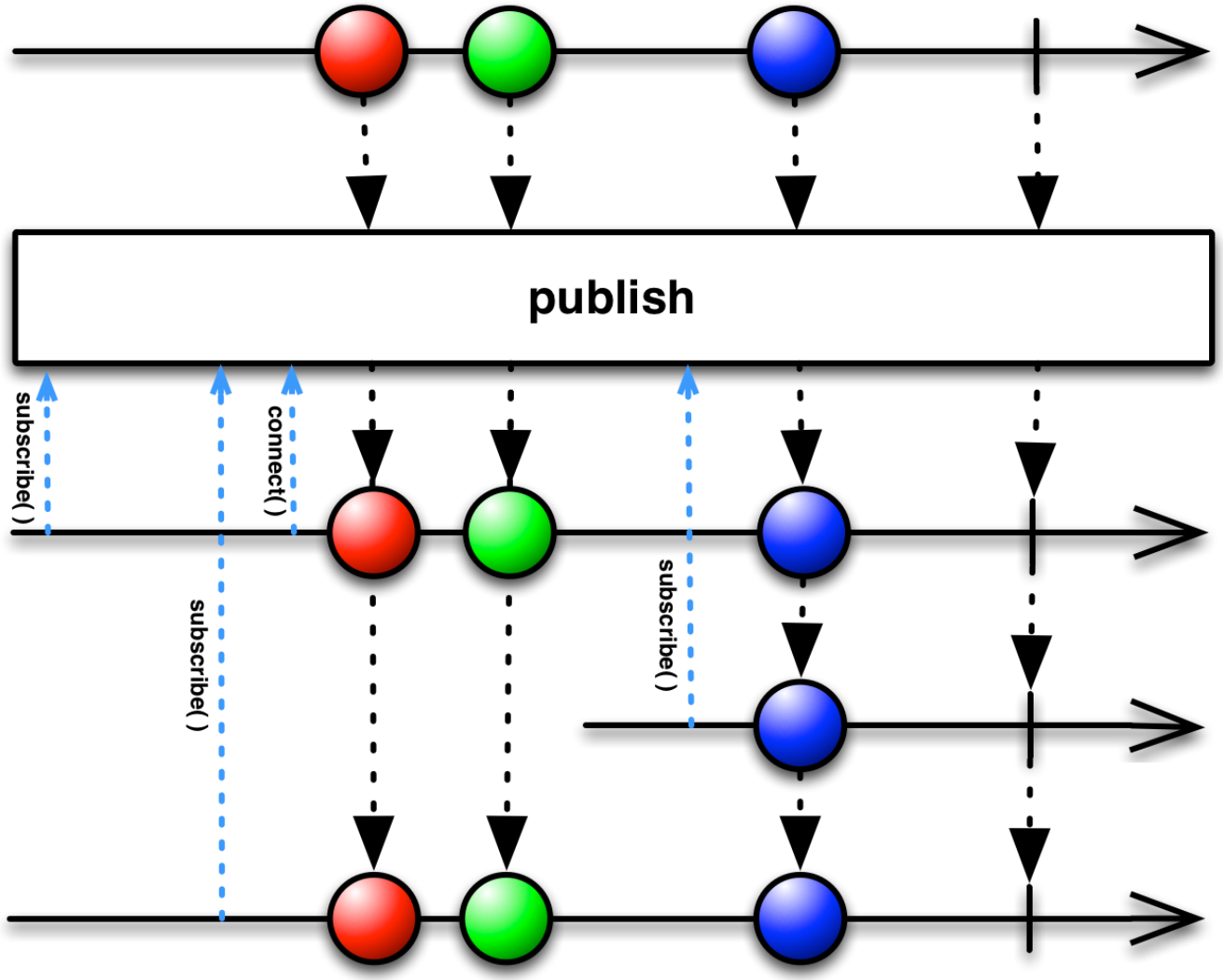
Source: RxJava 2 API documentation, http://reactivex.io/RxJava/2.x/javadoc/

Hot and Cold Event Streams

- PULL-based (Cold Event Streams) – Cold streams (e.g. RxJava Observable / Flowable or Reactor Flow / Mono) are streams that run their sequence when and if they are subscribed to. They present the sequence from the start to each subscriber.
- PUSH-based (Hot Event Streams) – emit values independent of individual subscriptions. They have their own timeline and events occur whether someone is listening or not. When subscription is made observer receives current events as they happen.
- *Example:* mouse events

56

Converting Cold to Hot Stream



Source: RxJava 2 API documentation, http://reactivex.io/RxJava/2.x/javadoc/

Reactive Programming

❖ Microsoft® opens source polyglot project ReactiveX (Reactive Extensions) [http://reactivex.io]:

Rx = Observables + LINQ + Schedulers :)

Java: RxJava, JavaScript: RxJS, C#: Rx.NET, Scala: RxScala, Clojure: RxClojure, C++: RxCpp, Ruby: Rx.rb, Python: RxPY, Groovy: RxGroovy, JRuby: RxJRuby, Kotlin: RxKotlin ...

❖ Reactive Streams Specification [http://www.reactive-streams.org/] used by:

❖ (Spring) Project Reactor [http://projectreactor.io/]

❖ Actor Model – Akka (Java, Scala) [http://akka.io/]

Reactive Streams Spec.

- Reactive Streams – provides standard for asynchronous stream processing with non-blocking back pressure.
- Minimal set of interfaces, methods and protocols for asynchronous data streams
- April 30, 2015: has been released version 1.0.0 of Reactive Streams for the JVM (Java API, Specification, TCK and implementation examples)
- Java 9: java.util.concurrent.Flow

Reactive Streams Spec.

- **Publisher** – provider of potentially unbounded number of sequenced elements, according to Subscriber(s) demand.

Publisher.subscribe(Subscriber) => **onSubscribe onNext* (onError | onComplete)?**

- **Subscriber** – calls **Subscription.request(long)** to receive notifications
- **Subscription** – one-to-one **Subscriber** ↔ **Publisher**, request data and cancel demand (allow cleanup).
- **Processor** = **Subscriber** + **Publisher**

FRP = Async Data Streams

- **FRP** is asynchronous data-flow programming using the building blocks of functional programming (e.g. map, reduce, filter) and explicitly modeling time
- Used for GUIs, robotics, and music. Example (RxJava):

```
Observable.from(
     new String[]{"Reactive", "Extensions", "Java"})
    .take(2).map(s -> s + " : on " + new Date())
    .subscribe(s -> System.out.println(s));
```

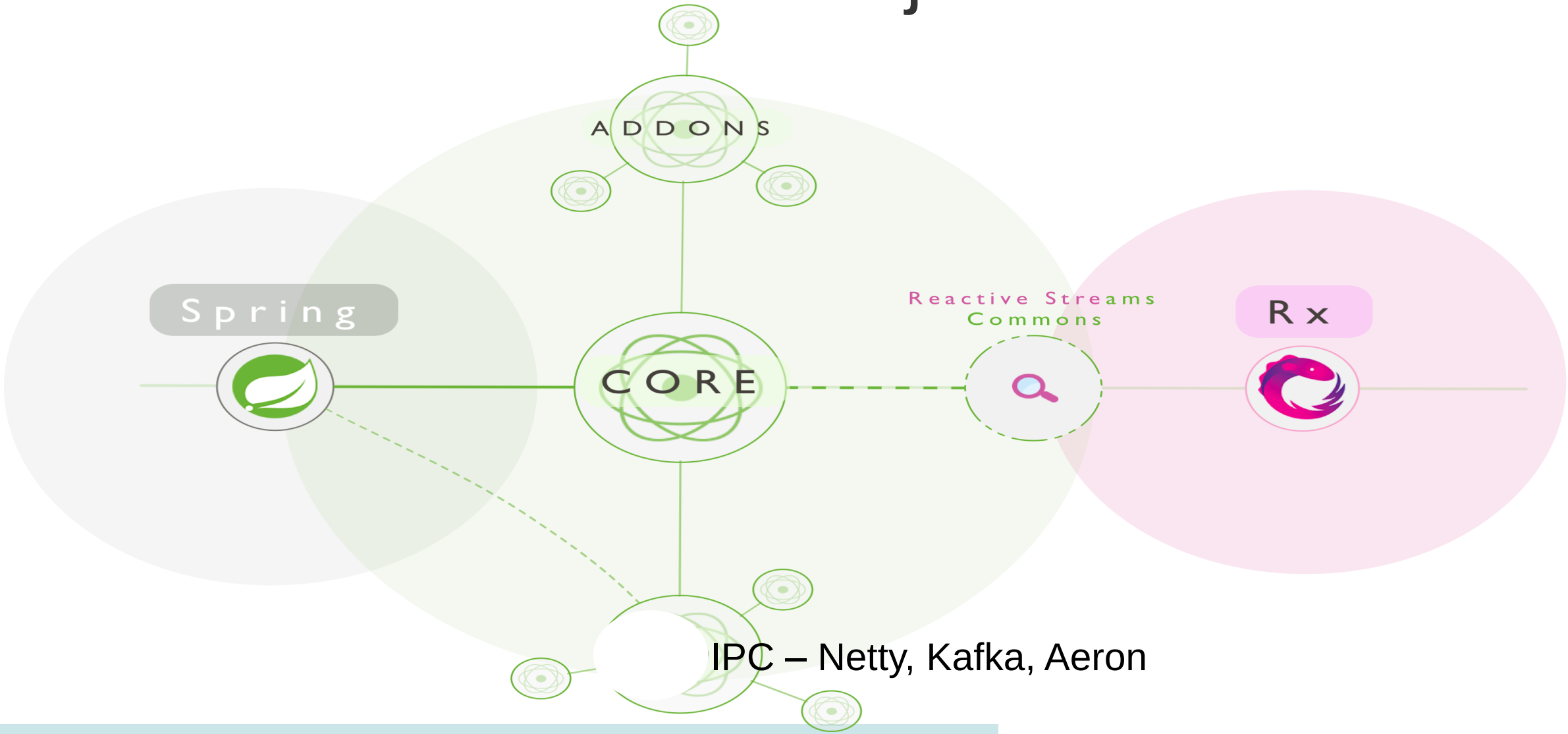
Result:

```
Reactive : on Wed Jun 17 21:54:02 GMT+02:00 2015
Extensions : on Wed Jun 17 21:54:02 GMT+02:00 2015
```

Project Reactor

- Reactor project allows building high-performance (low latency high throughput) non-blocking asynchronous applications on JVM.
- Reactor is designed to be extraordinarily fast and can sustain throughput rates on order of 10's of millions of operations per second.
- Reactor has powerful API for declaring data transformations and functional composition.
- Makes use of the concept of Mechanical Sympathy built on top of Disruptor / RingBuffer.

Reactor Projects



https://github.com/reactor/reactor, Apache Software License 2.0

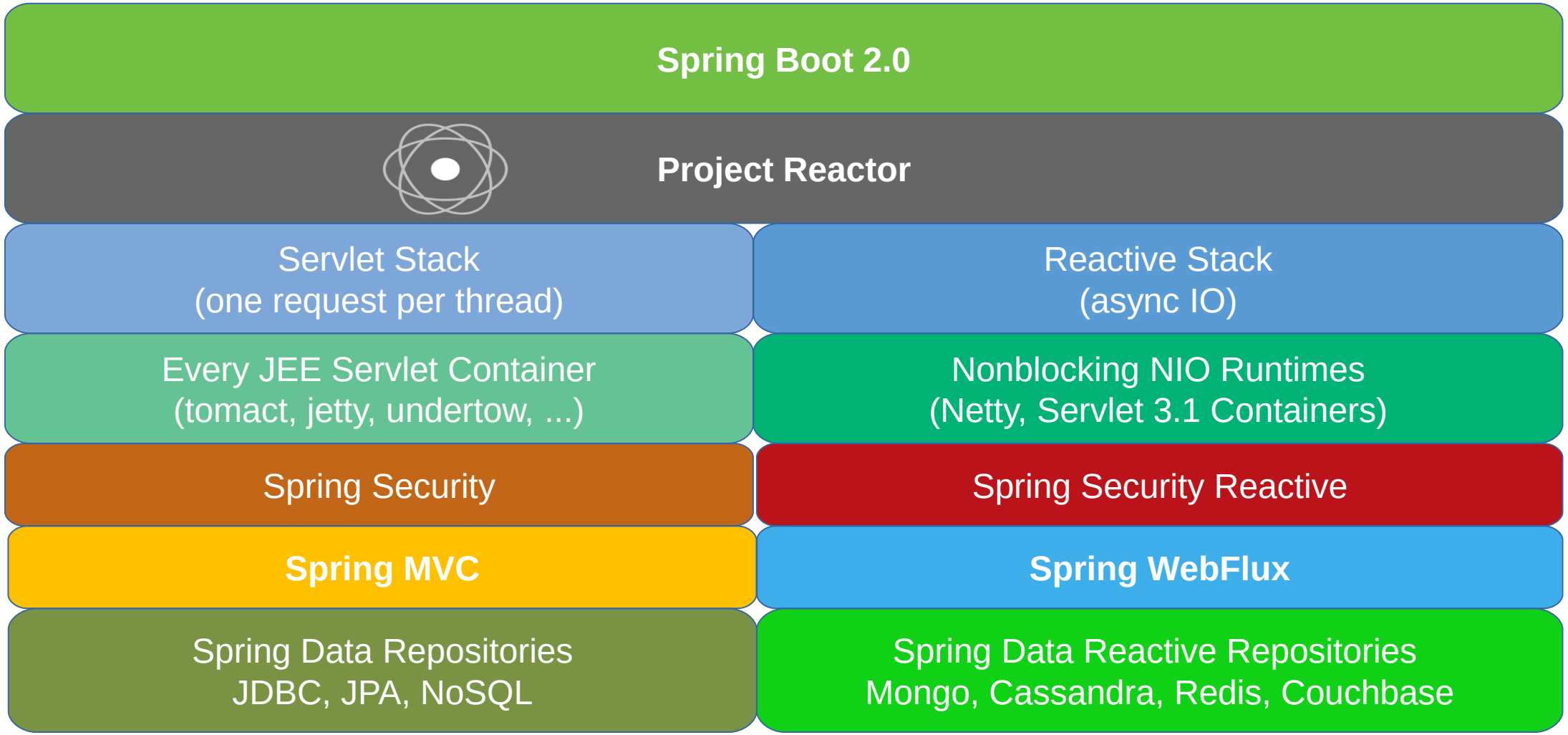
Hot Stream Example - Reactor

```
EmitterProcessor<String> emitter =
        EmitterProcessor.create();
FluxSink<String> sink = emitter.sink();
emitter.publishOn(Schedulers.single())
        .map(String::toUpperCase)
        .filter(s -> s.startsWith("HELLO"))
        .delayElements(Duration.of(1000, MILLIS))
.subscribe(System.out::println);
sink.next("Hello World!"); // emit - non blocking
sink.next("Goodbye World!");
sink.next("Hello Trayan!");
Thread.sleep(3000);
```

Top New Features in Spring 5

- Reactive Programming Model
- Spring Web Flux
- Reactive DB repositories & integrations + hot event streaming: MongoDB, CouchDB, Redis, Cassandra, Kafka
- JDK 8+ and Java EE 7+ baseline
- Testing improvements – WebTestClient (based on reactive WebFlux WebClient)
- Kotlin functional DSL

Spring 5 Main Building Blocks



Source: https://spring.io

Thank's for Your Attention!

Trayan Iliev

IPT – Intellectual Products & Technologies

http://iproduct.org/

https://github.com/iproduct

https://twitter.com/trayaniliev

https://www.facebook.com/IPT.EACAD