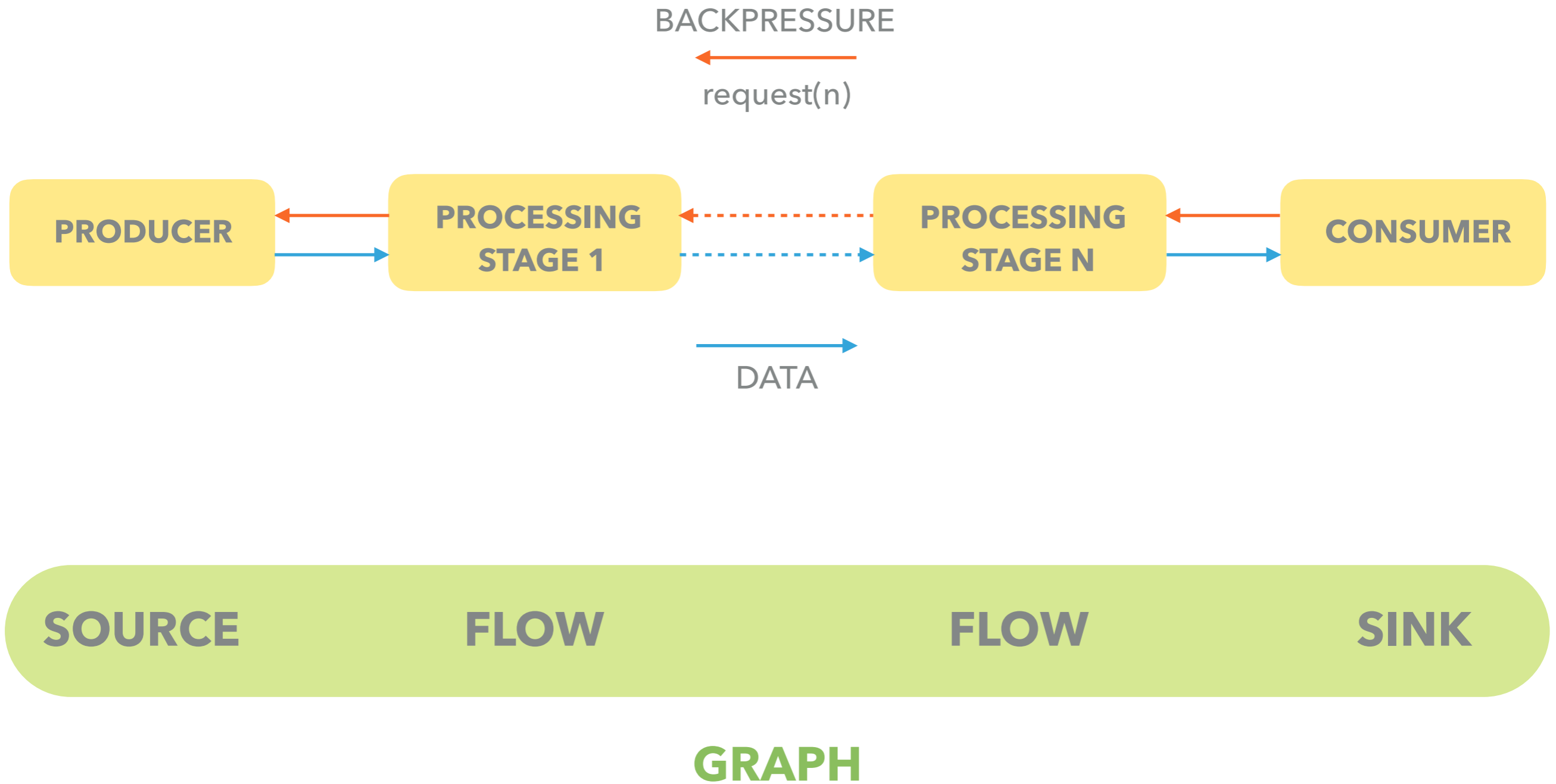


JACEK KUNICKI





AKKA STREAMS IN PRACTICE

STREAM PROCESSING



Stream processing

- Apache Storm 
- Apache Storm Trident
- Alibaba JStorm
- Twitter Heron 
- **Apache Spark Streaming**
- **Apache Flink**
- **Apache Beam**



- Apache Kafka Streams
- Apache Samza 
- Apache Gearpump
- Apache Apex
- Apache Ignite Streaming
- Apache S4 
- ...



AKKA STREAMS

- ▶ not distributed (unlike Spark Streaming etc.)



<http://www.slideshare.net/kamkasravi/gearpump-akka-streams>

AKKA STREAMS

- ▶ not distributed (unlike Spark Streaming etc.)
- ▶ reusable building blocks
- ▶ a graph is lazy until actually run

USE CASE

```
id;value  
380427;0.5683437990833604  
380427;0.31402364105656866  
956159;invalid_value  
956159;0.7374388244952665  
121101;0.23269941059069932  
121101;0.8915470447933084
```

.GZ

AVERAGE BY ID

CASSANDRA





AND MORE...

- ▶ graph DSL for non-trivial graphs
- ▶ build your own flow - see my blog post

THANK YOU!



jacek.kunicki.org/Id-2017



@rucek