

Spark bevezeto

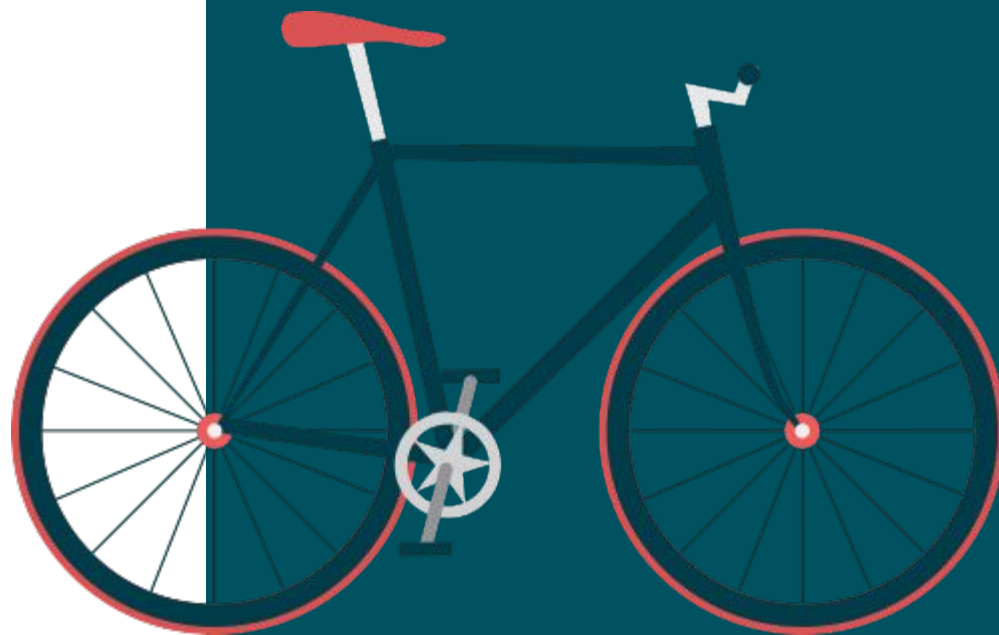
Mate Gulyas



GULYÁS MÁTÉ

CTO & Co-Founder

 [@gulyasm](https://twitter.com/gulyasm)



APACHE SPARK

MI AZ A SPARK?

1. Adat feldolgozó keretrendszer
2. Elosztott, több gépen folyik
adatfeldolgozás
3. In memory

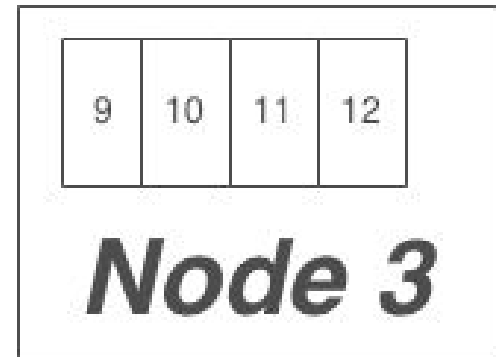
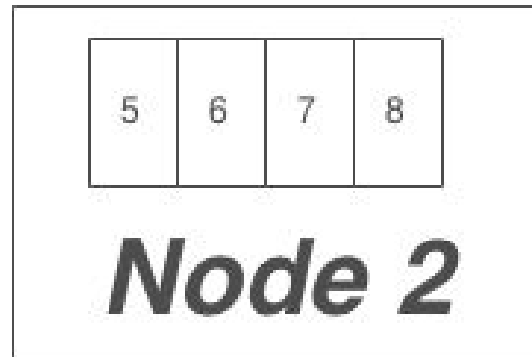
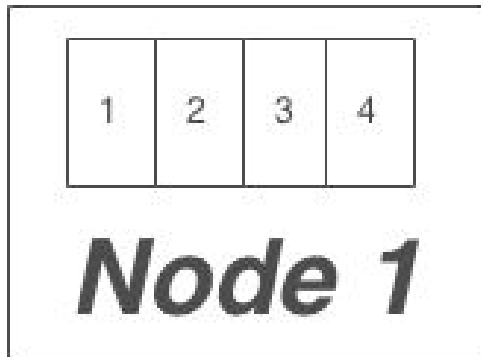
MIÉRT JÖTT LÉTRE?

1. Hadoop MapReduce mindig diszkre perzisztál
2. Alacsony szintű API

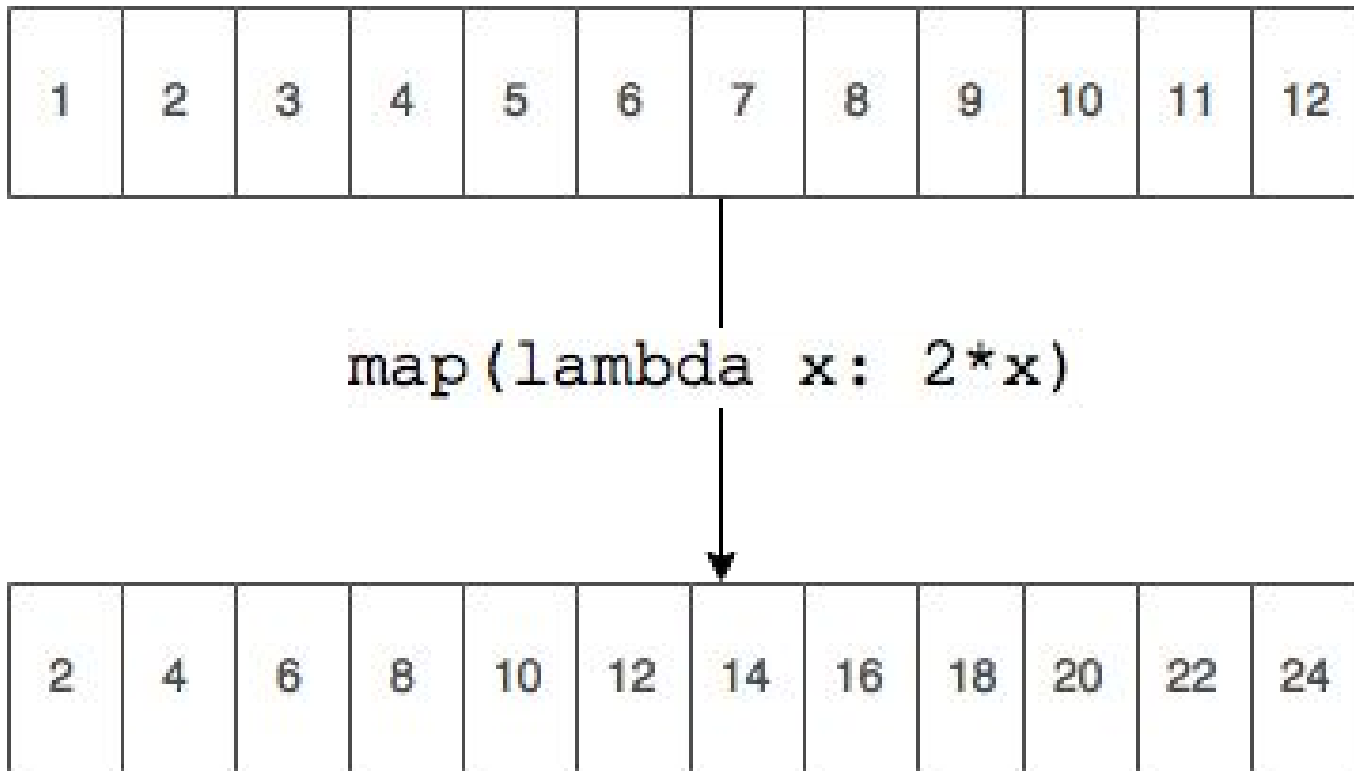
HOGYAN KELL ELKÉPZELNI?

Elosztott számolás

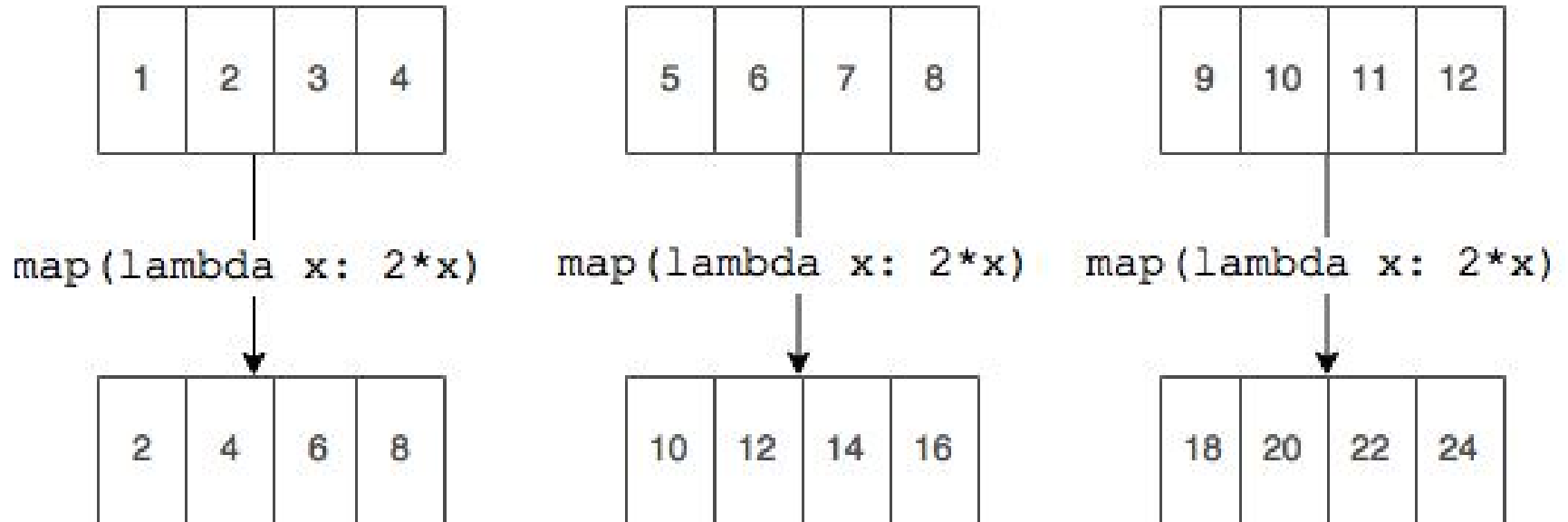
1	2	3	4	5	6	7	8	9	10	11	12
---	---	---	---	---	---	---	---	---	----	----	----



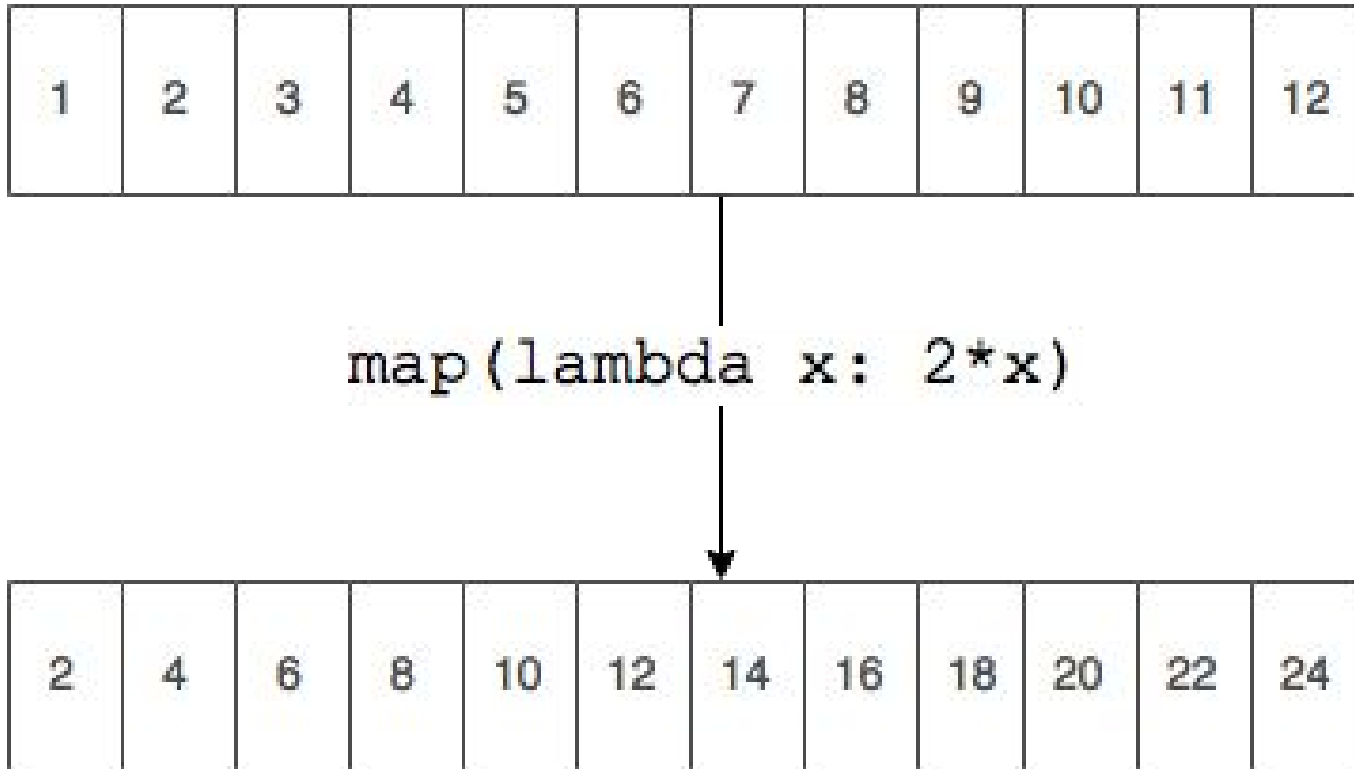
TRANSFORMATION



TRANSFORMATION



TRANSFORMATION



MIÉRT A HYPE?

OPEN SOURCE

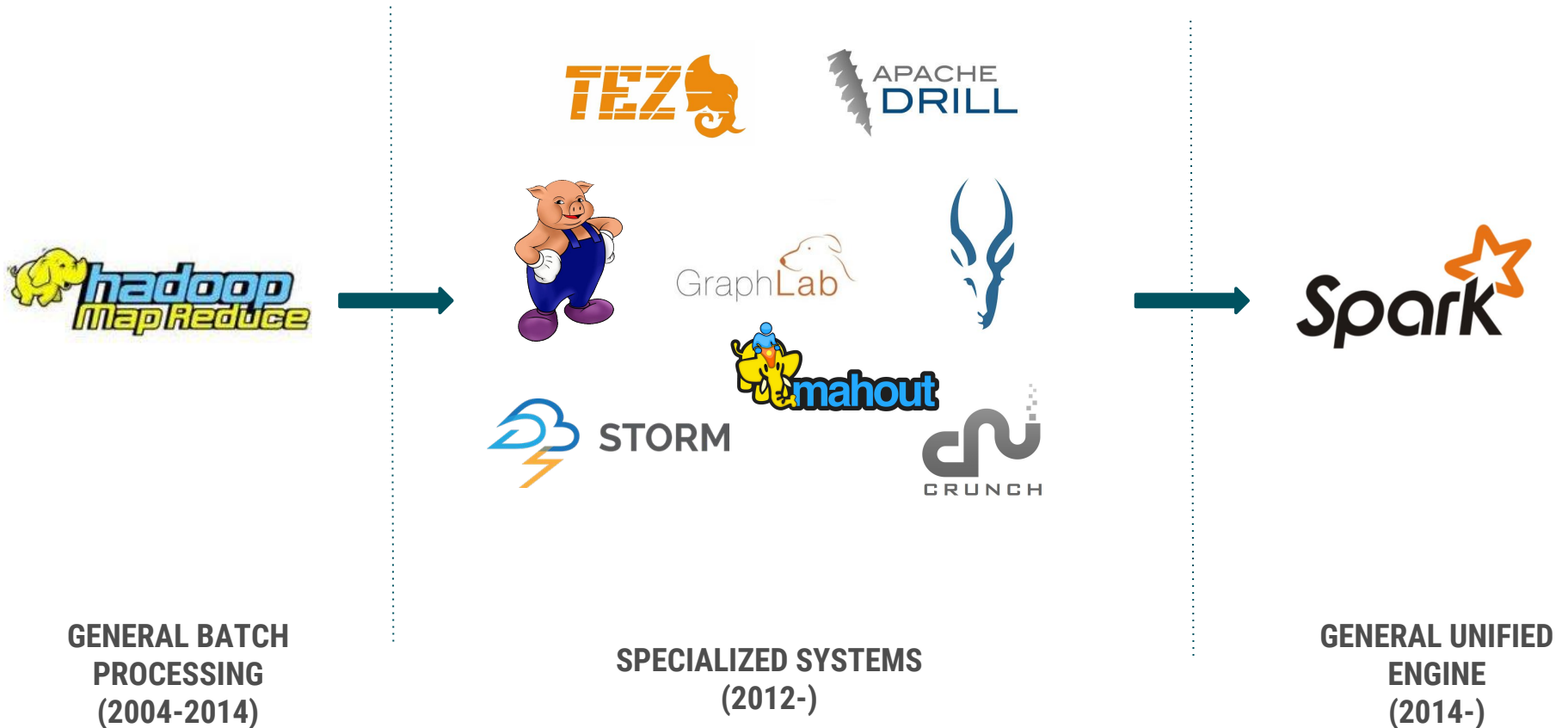
1. **Community**
2. **Ecosystem**
3. **Open source**

SPARK

as a

PLATFORM

Unified Data Processing Engine



UNIFIED STACK

Spark Core

Spark SQL

Spark Streaming

MLlib

GraphX

Cluster Managers



Egységes platform

1. Code újrafelhasználás
2. Üzemeltetési költség
3. Know how

HIGH LEVEL

API



LONGEST LINE

```
JavaRDD<String> lines = context.textFile(  
    "/home/gulyasm/text"  
);
```

```
JavaRDD<Integer> lengths = lines.map(  
    (Function<String, Integer>) String::length  
);
```

```
Integer max = length.reduce(  
    (Function2<Integer, Integer, Integer>  
Integer::max  
);
```

```
JavaRDD<String> lines = context.textFile(  
    "/home/gulyasm/text"  
);
```

```
JavaRDD<Integer> lengths = lines.map(  
    (Function<String, Integer>) String::length  
);
```

```
Integer max = length.reduce(  
    (Function2<Integer, Integer, Integer>  
Integer::max  
);
```

```
JavaRDD<String> lines = context.textFile(
    "/home/gulyasm/text"
);

JavaRDD<Integer> lengths = lines.map(
    (Function<String, Integer>) String::length
);

Integer max = length.reduce(
    (Function2<Integer, Integer, Integer>)
Integer::max
);
```

```
JavaRDD<String> lines = context.textFile(
    "/home/gulyasm/text"
);

JavaRDD<Integer> lengths = lines.map(
    (Function<String, Integer>) String::length
);

Integer max = length.reduce(
    (Function2<Integer, Integer, Integer>)
Integer::max
);
```

THANK YOU!

MATE GULYAS

gulyasm@enbrite.ly

[@gulyasm](https://twitter.com/gulyasm)

[@enbritely](https://twitter.com/enbritely)

