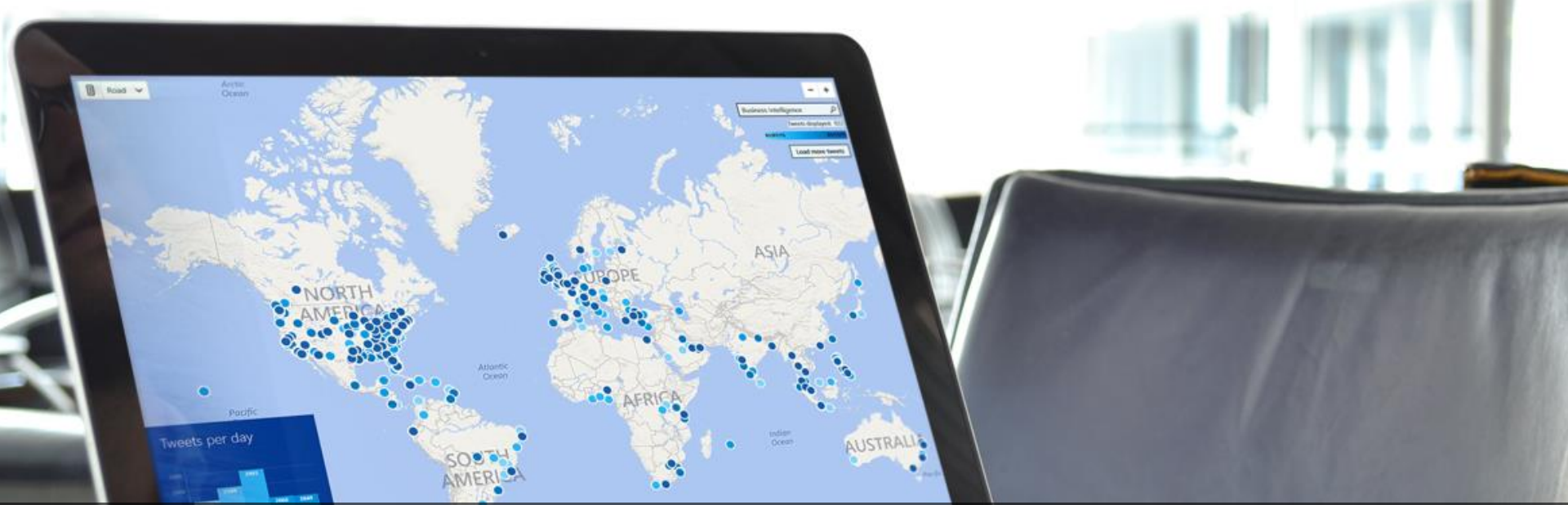



# Multi-Dimensional, In-GPU-Memory Databases: Streaming Conditional Calculations in Big Data Sets

Peter Strohm | GTC15 | San José | 03/17/2015

**2002**
Founded in Freiburg, Germany

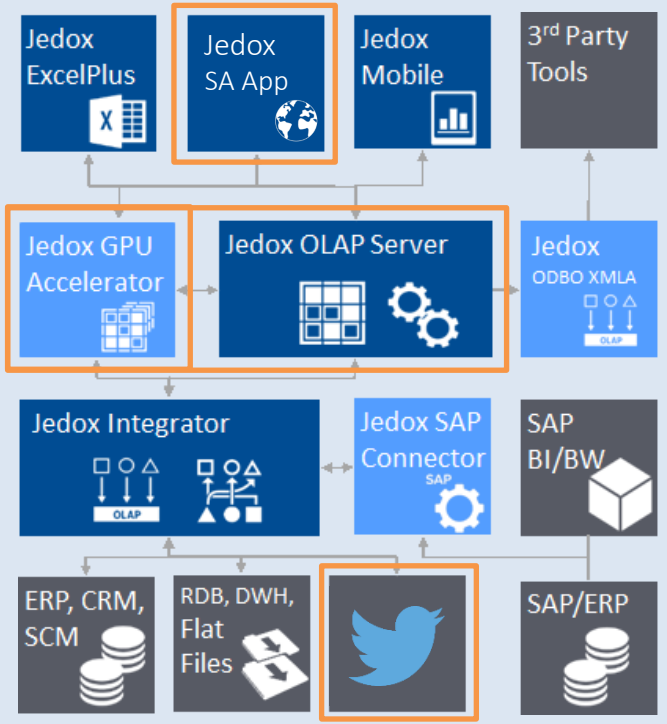
**Today**
Offices in Freiburg, Frankfurt, Düsseldorf, Paris, Boston

**120**
Global Business Partners

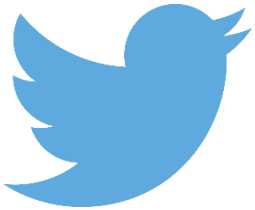
**125**
Countries with Jedox Users

**25**
Language Versions

Gartner 2013 Cool Vendor



# Big Data in real-time: Twitter



500.000.000

tweets per day

6.000.000

localized tweets per day

60.000.000

data points per day

> 4.000

tweets per minute

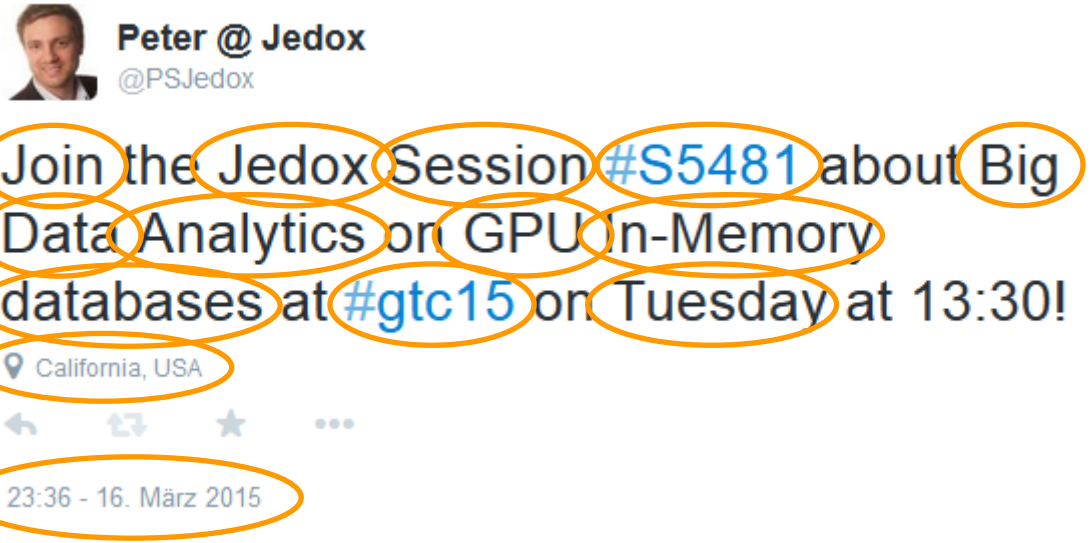
12 GB

data volume per GPU

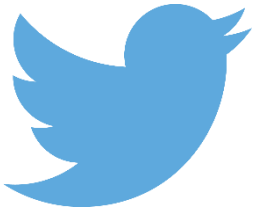
< 1 second

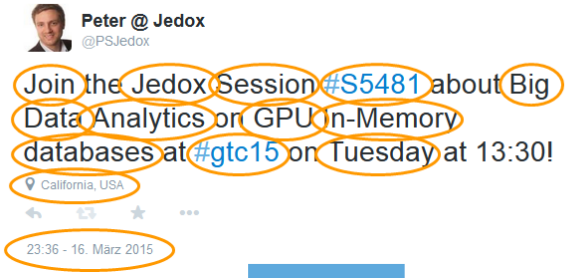
of computation time

# Big Data in real-time: Twitter



# Twitter: Big Data in Real-time



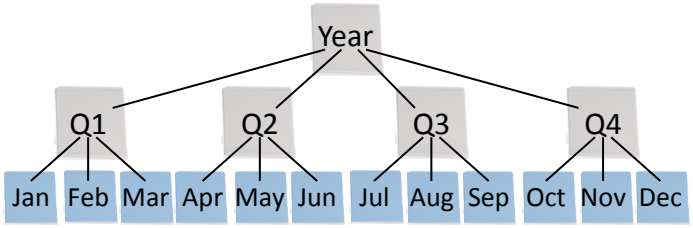


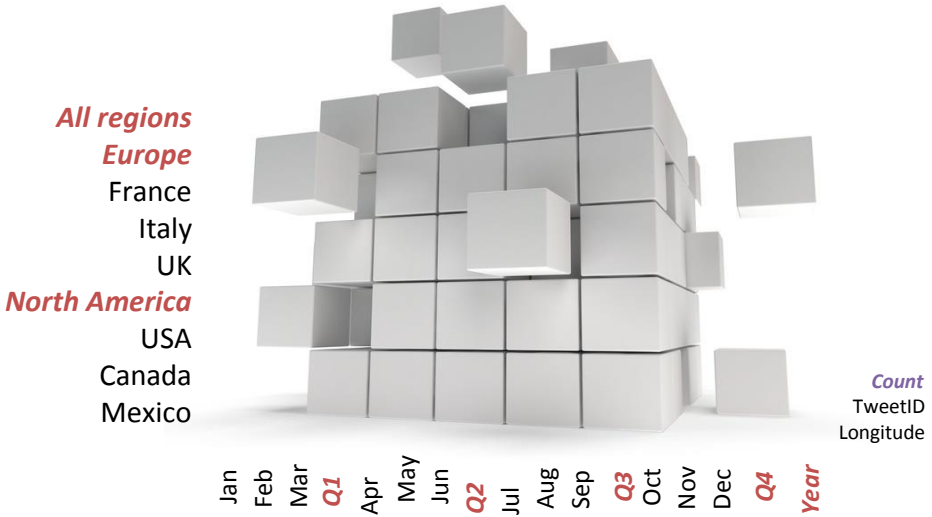

| Dictionary | Date | Hour | Minute | Second | Longitude | Latitude | Language | TweetID |
|---|---|---|---|---|---|---|---|---|
| Join | 2015-03-16 | 23 | 36 | 49 | -121 | 37 | en | 577599223119548416 |
| Jedox | 2015-03-16 | 23 | 36 | 49 | -121 | 37 | en | 577599223119548416 |
| Session | 2015-03-16 | 23 | 36 | 49 | -121 | 37 | en | 577599223119548416 |
| #S5481 | 2015-03-16 | 23 | 36 | 49 | -121 | 37 | en | 577599223119548416 |
| ... | | | | | | | | |

# In-GPU-Memory OLAP-Database

1. All data in GPU memory

2. Hierarchical structure of consolidated elements



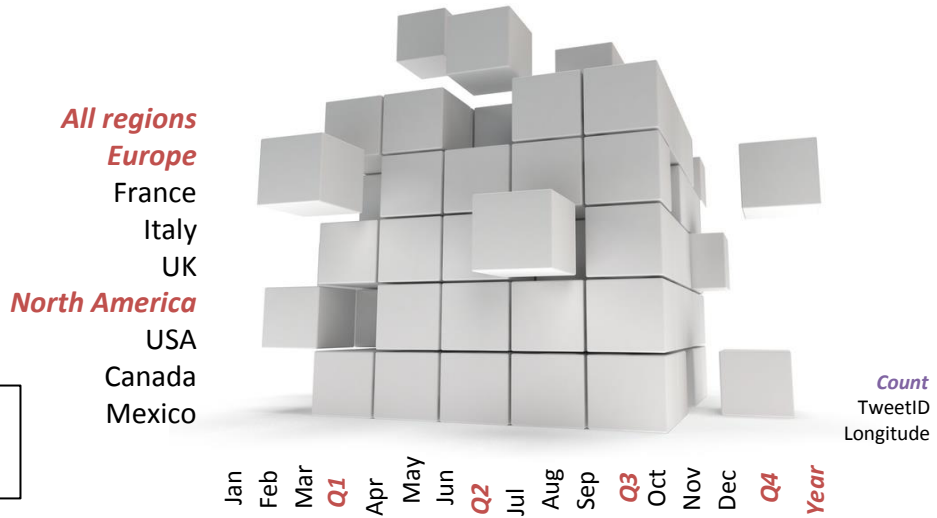



# In-GPU-Memory OLAP-Database

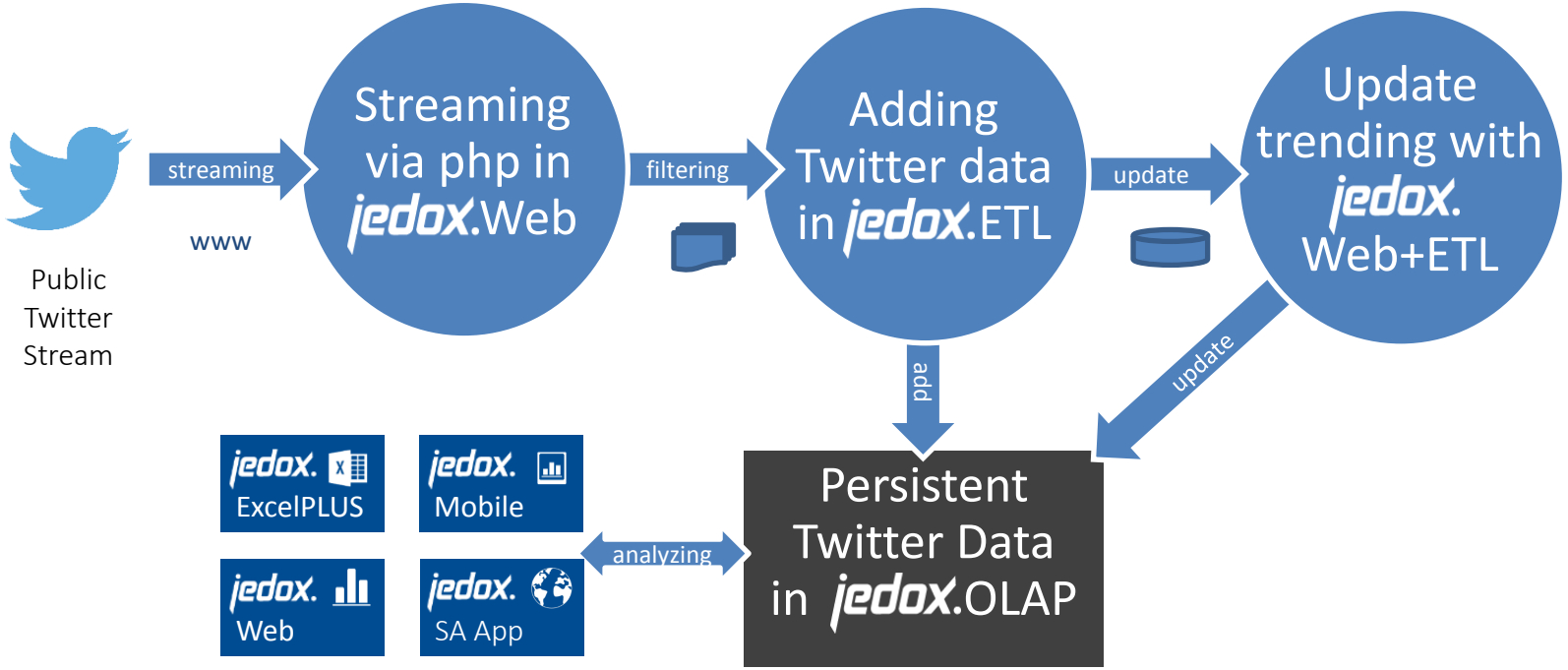
3 Calculation “on-the-fly”

4 Store only non-zero values

i In-GPU-Memory & “on-the-fly”



# Jedox Social Analytics Workflow



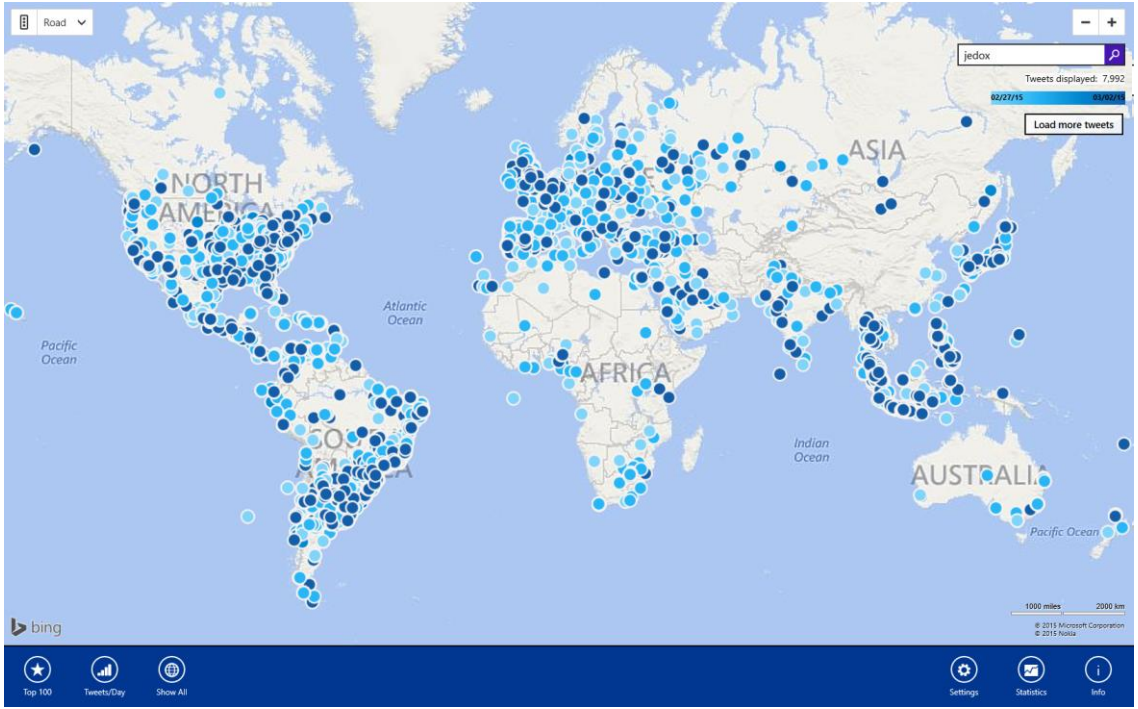
# Jedox Social Analytics with GPU

1. System: 2 x K40 with 12 GB GPURAM
   CPU with 128GB RAM


2. Data sets with > 500 million entries

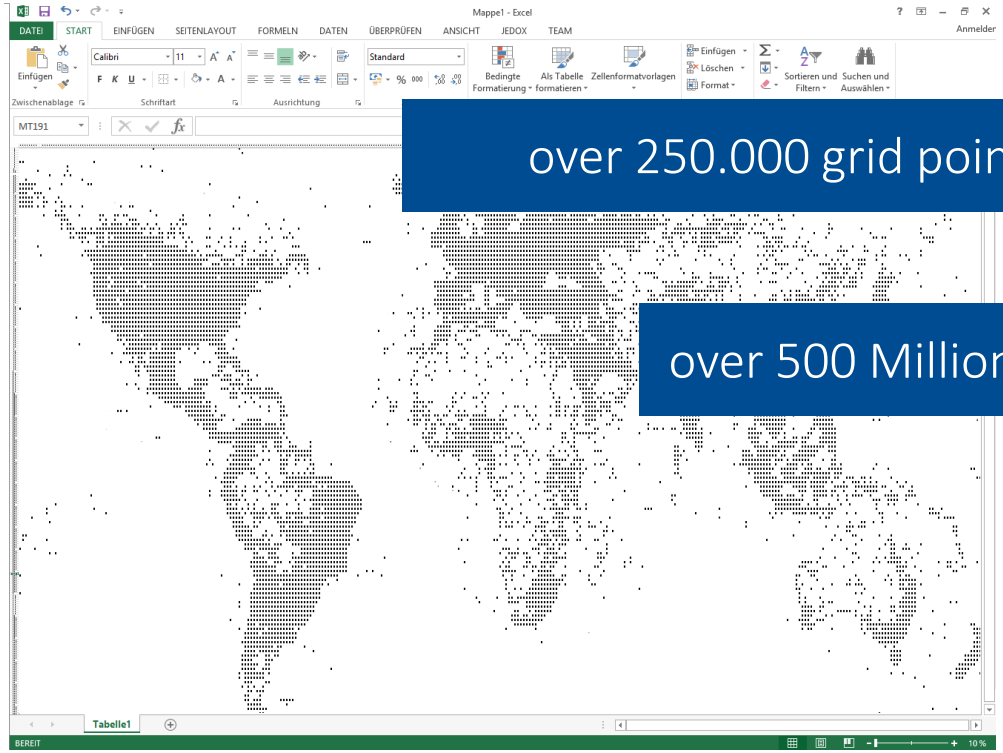
3. Calculating “on-the-fly” on ***all*** data entries

# Jedox Social Analytics Showcase



# Jedox Social Analytics Showcase

# Social Media Analytics: HeatMap




over 250.000 grid points

over 500 Million database entries

# Social Media Analytics: Trending

1

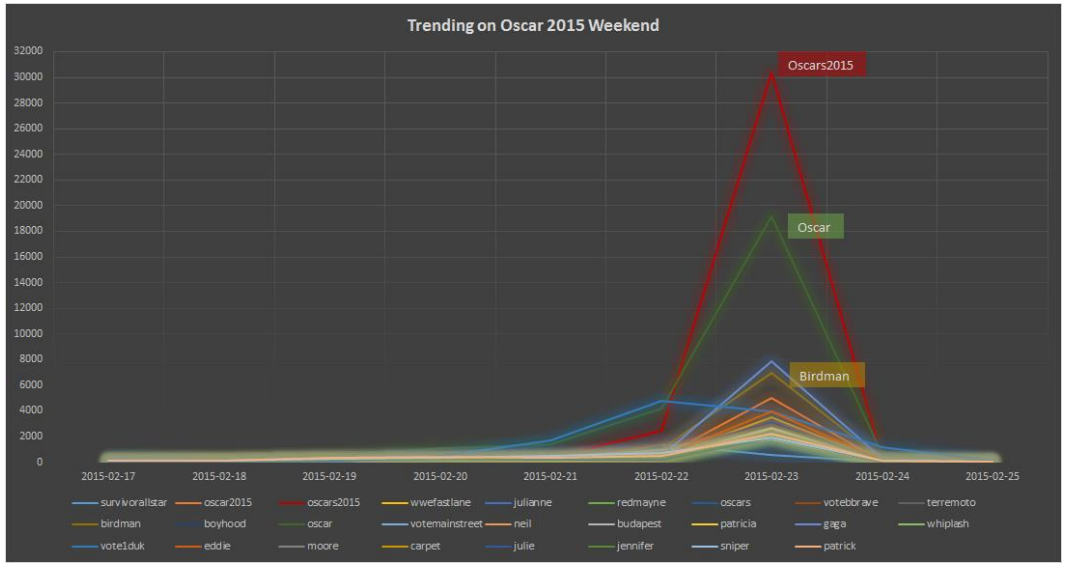
Top 100 *Trending Factor*

$$\frac{\text{\# current day}}{\text{\# all previous days}}$$

2

~ 28.000 dictionary words

> 80 Million DB entries



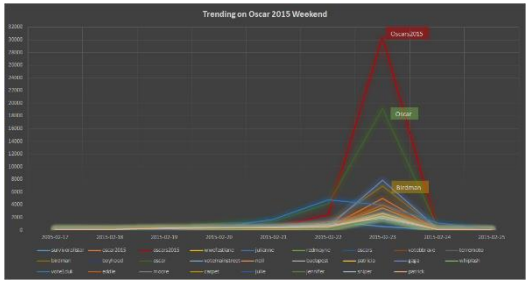
# Social Media Analytics: Trending

1. Assigning *Trending Factor* to **virtual cells** (e.g. [TF] = [B] / [C])

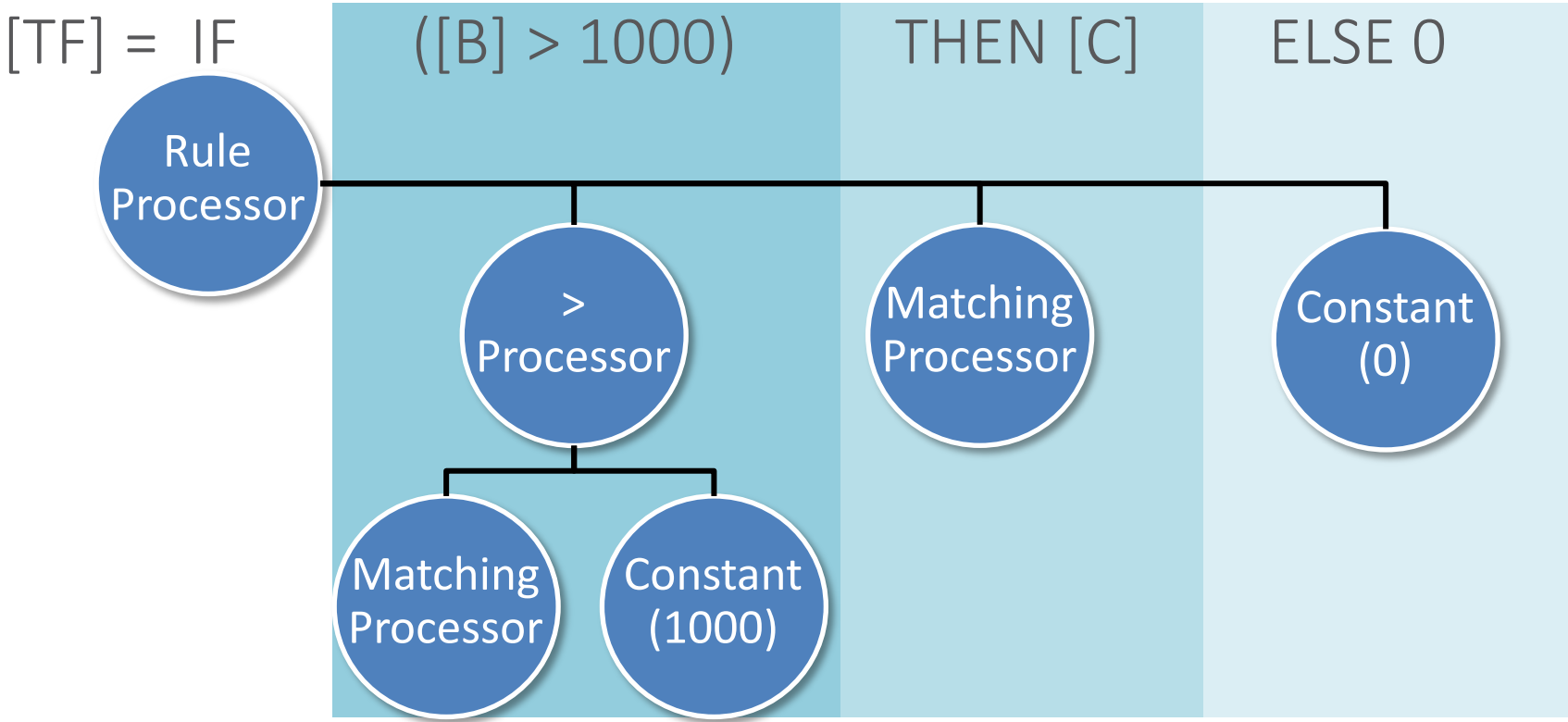
2. *Trending Factor* is **calculated on the fly**

3. Conditional Calculation: **IF-Rule**

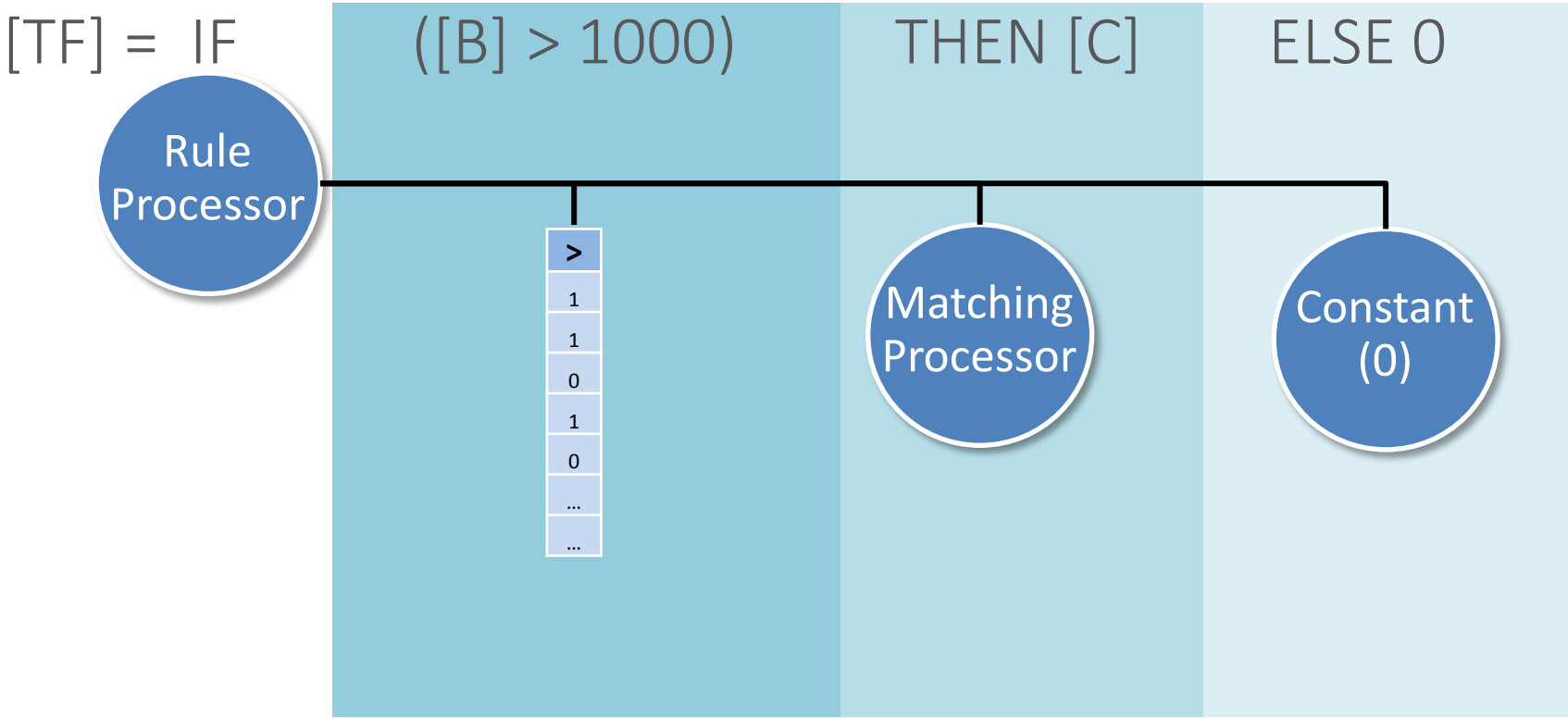
```
IF      ([count] >1000)
THEN [count,currentDay] / [count,previousWeek]
ELSE   0
```



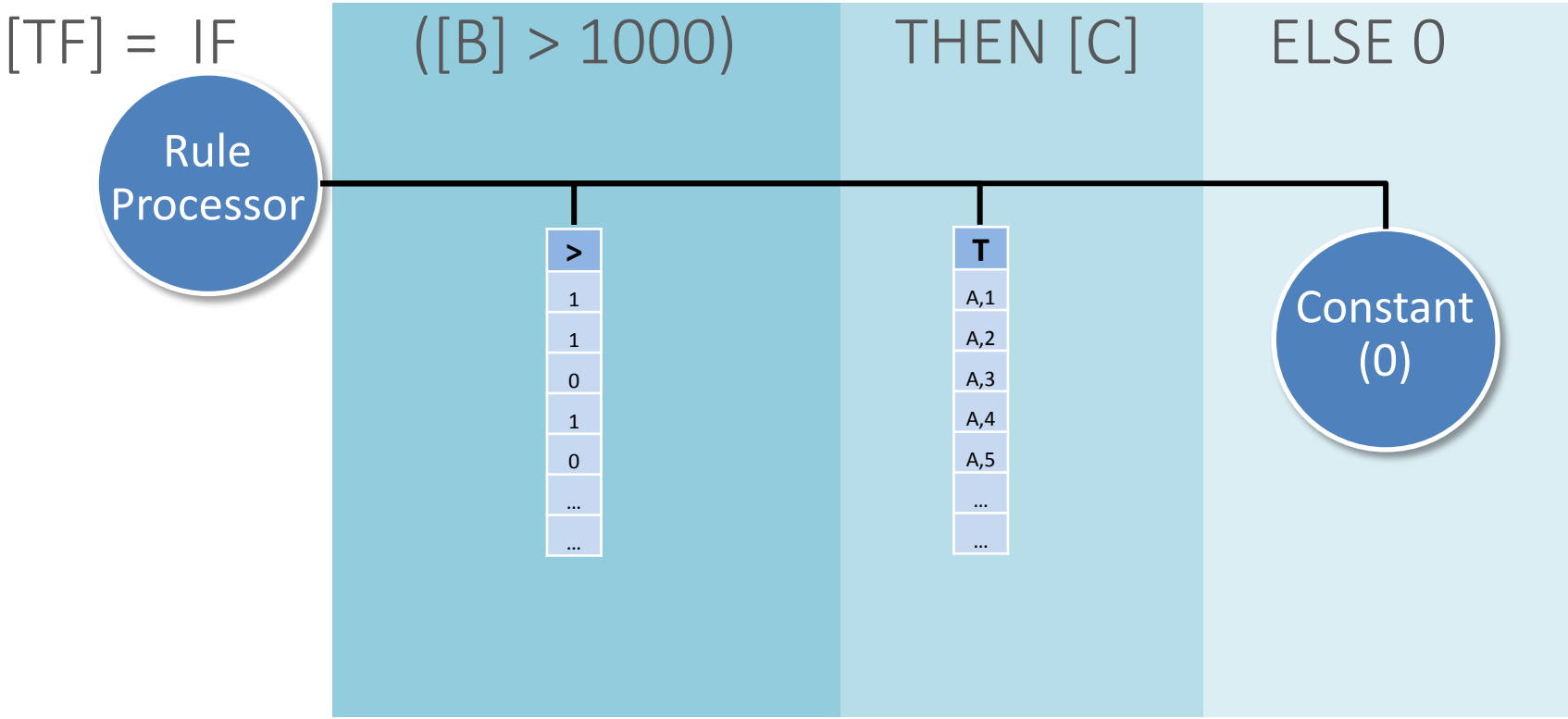
# Conditional Rule with constants



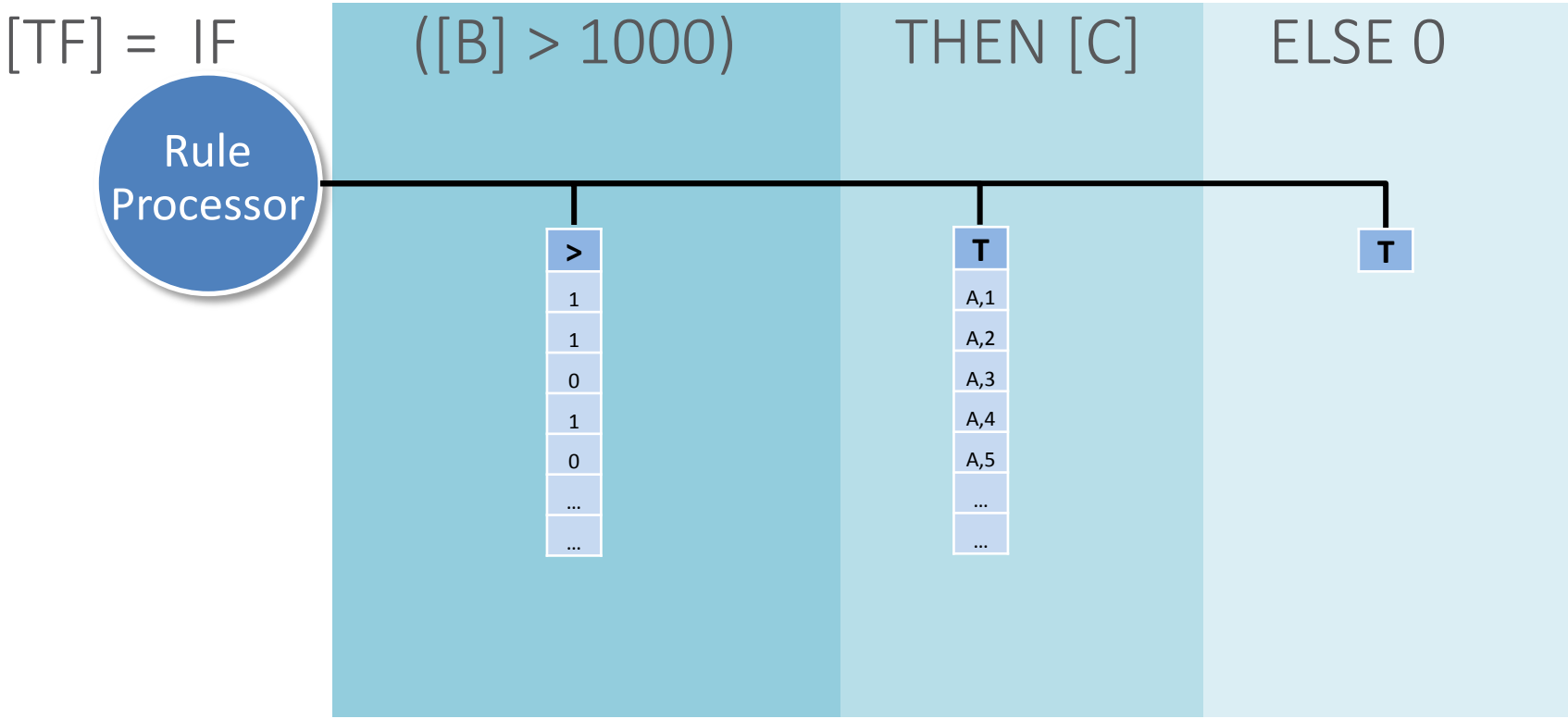
# Conditional Rule with constants



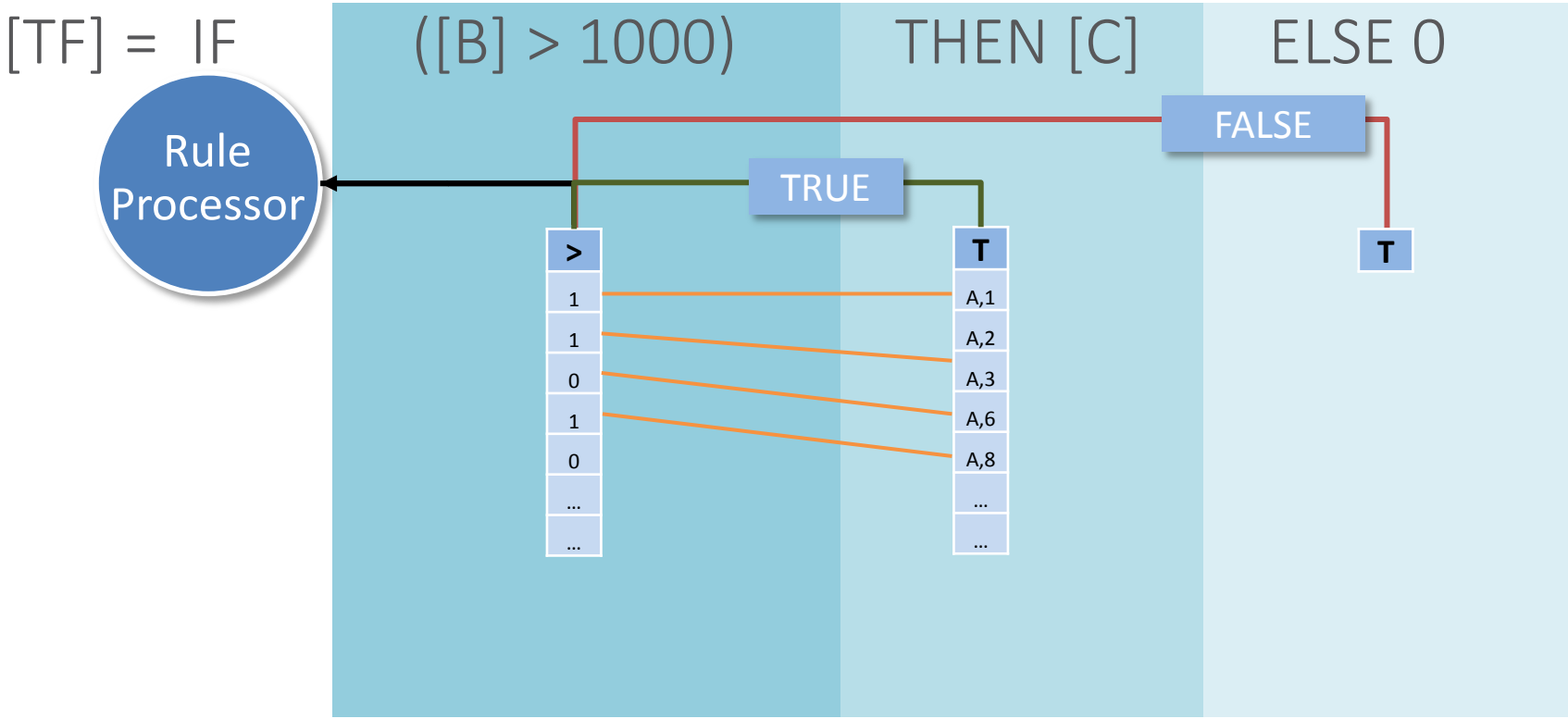
# Conditional Rule with constants



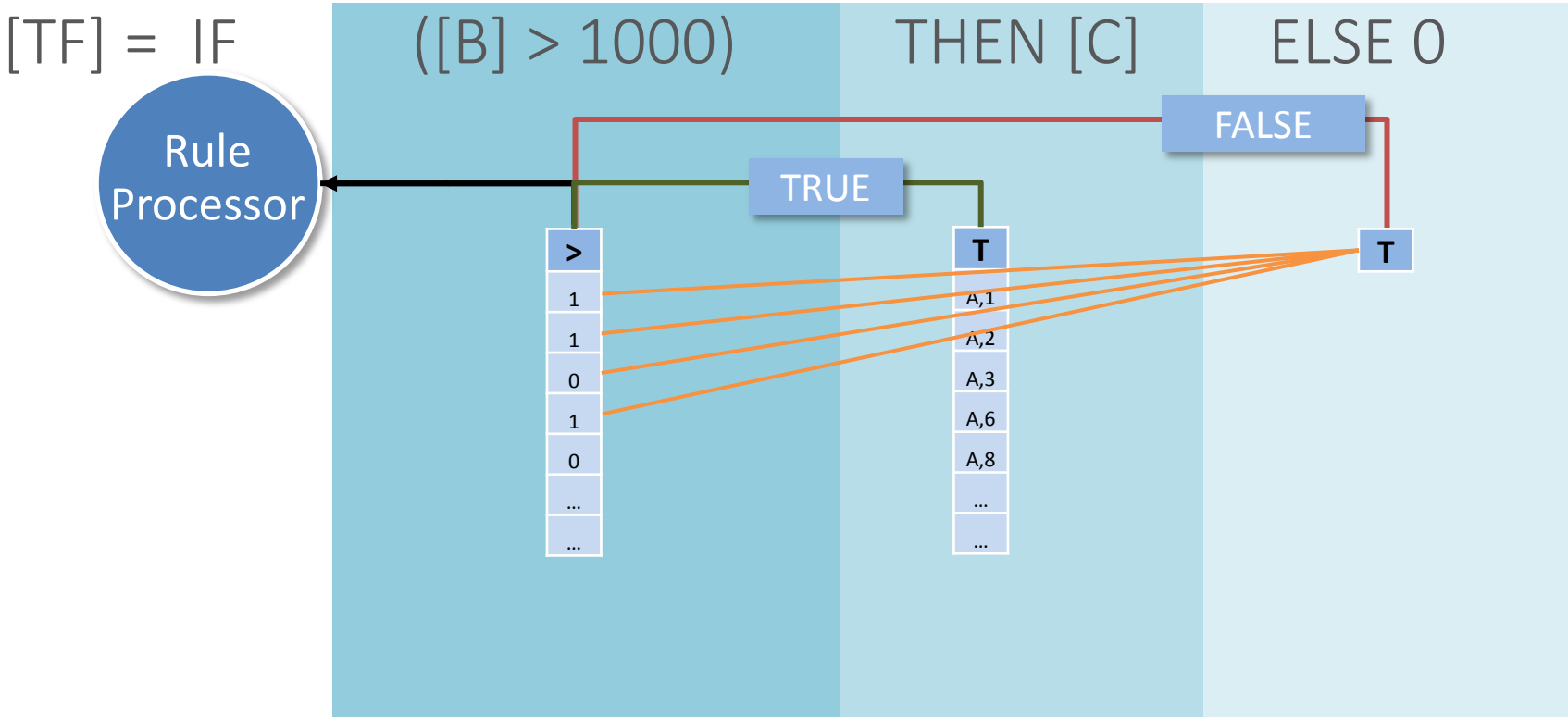
# Conditional Rules with constants



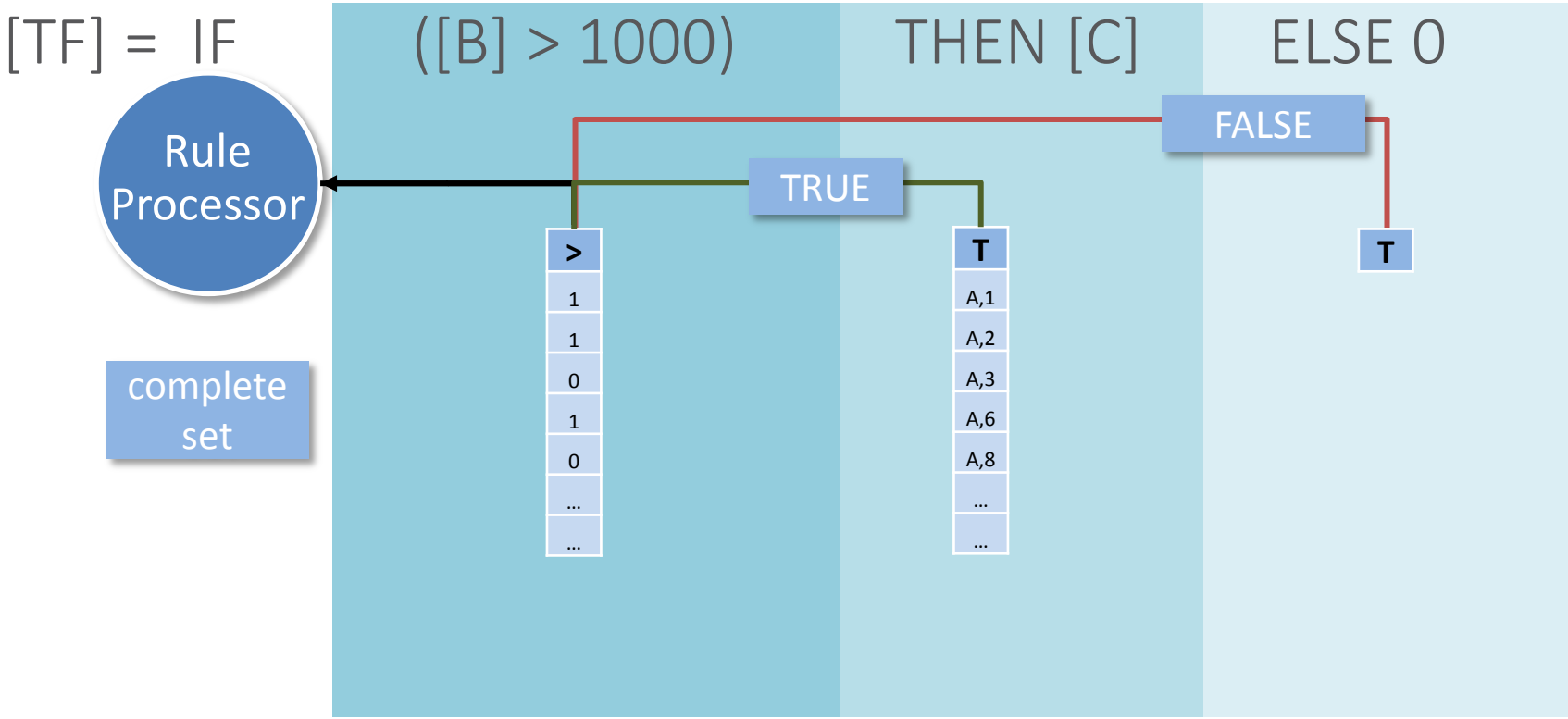
# Conditional Rules on GPU



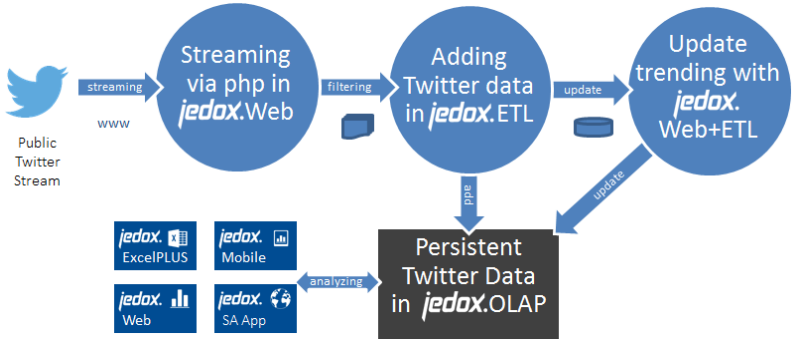
# Conditional Rules on GPU



# Conditional Rules on GPU



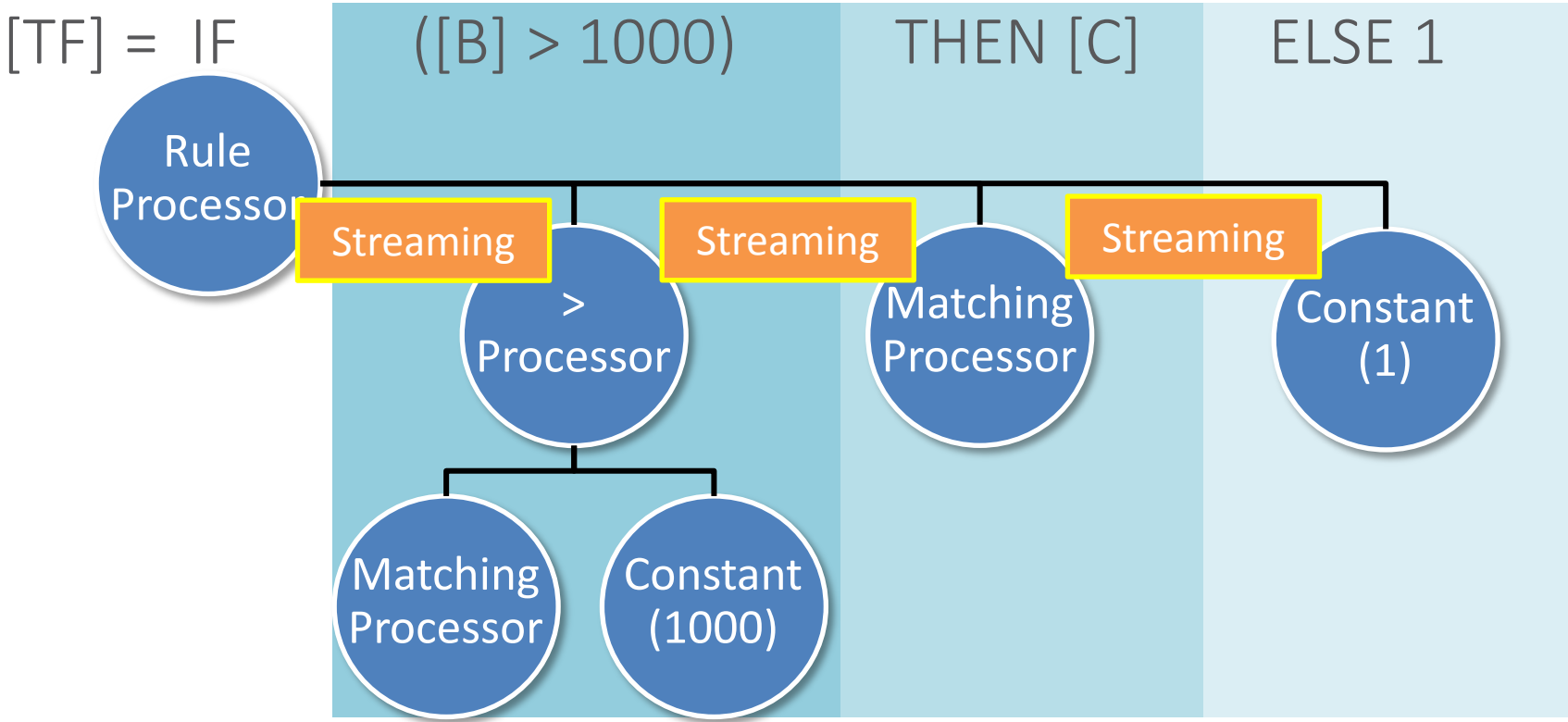
# Conditional Rules on GPU



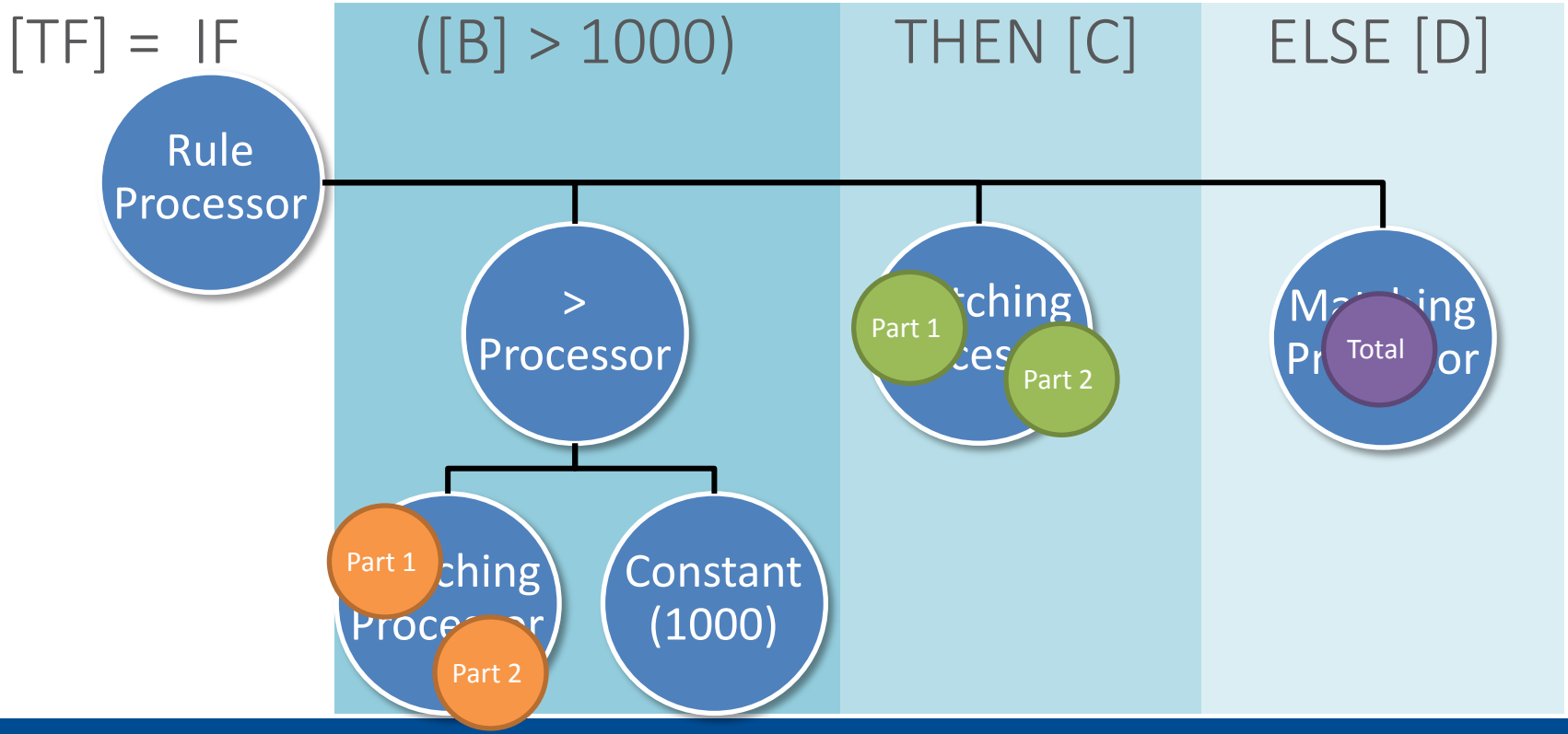

GPU MEMORY

SPACE FOR CALCULATIONS

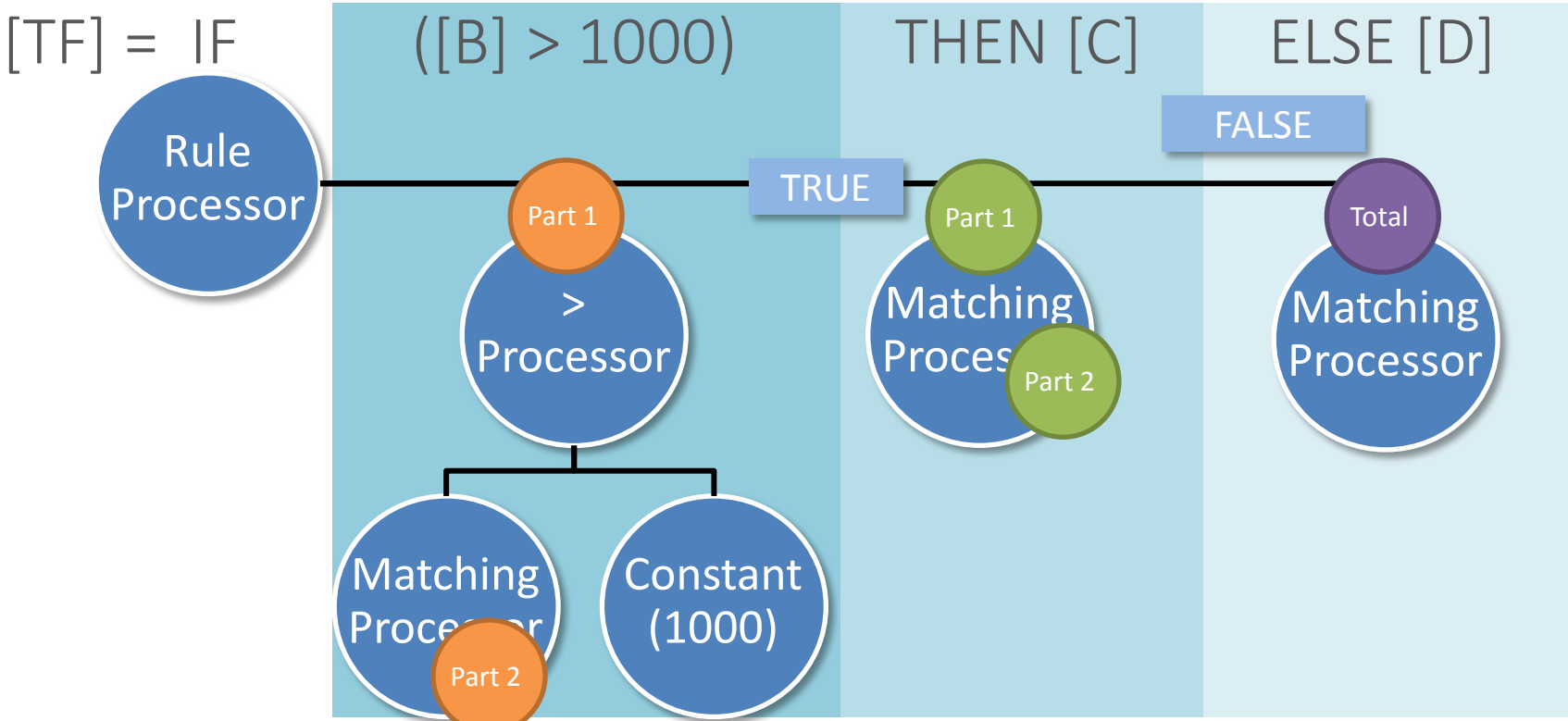
# Conditional Rule on GPU



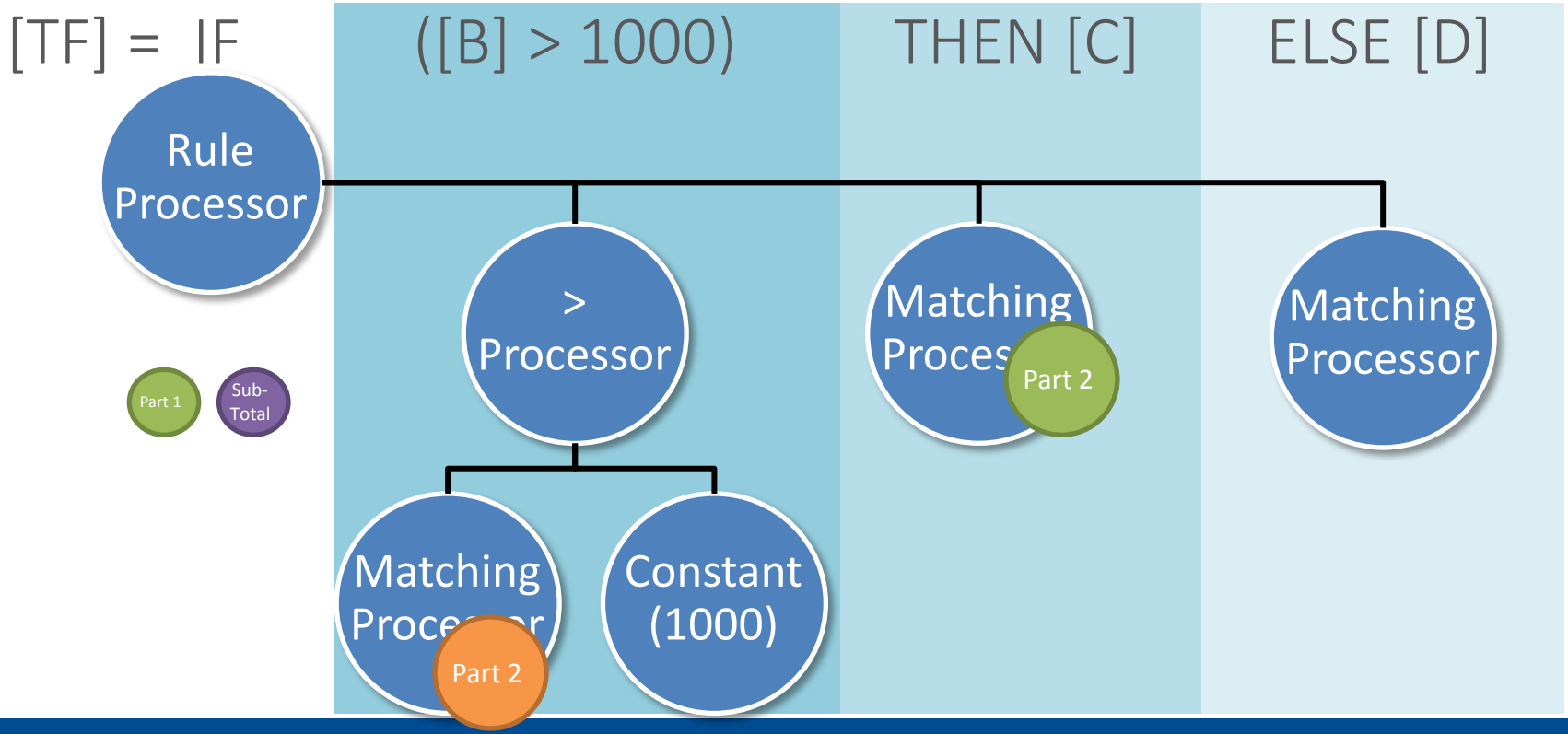
# Streaming in Conditional Rules



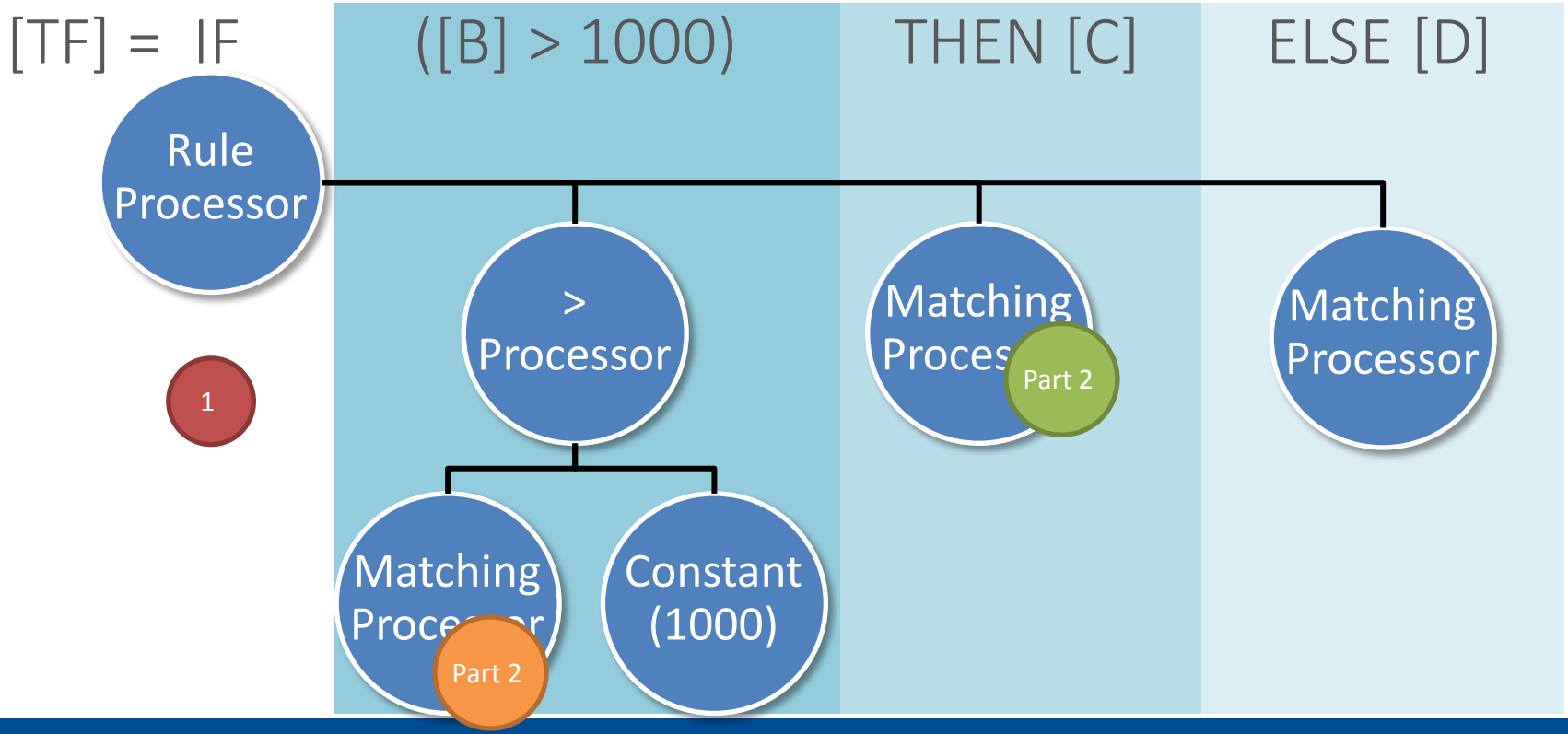
# Streaming in Conditional Rules



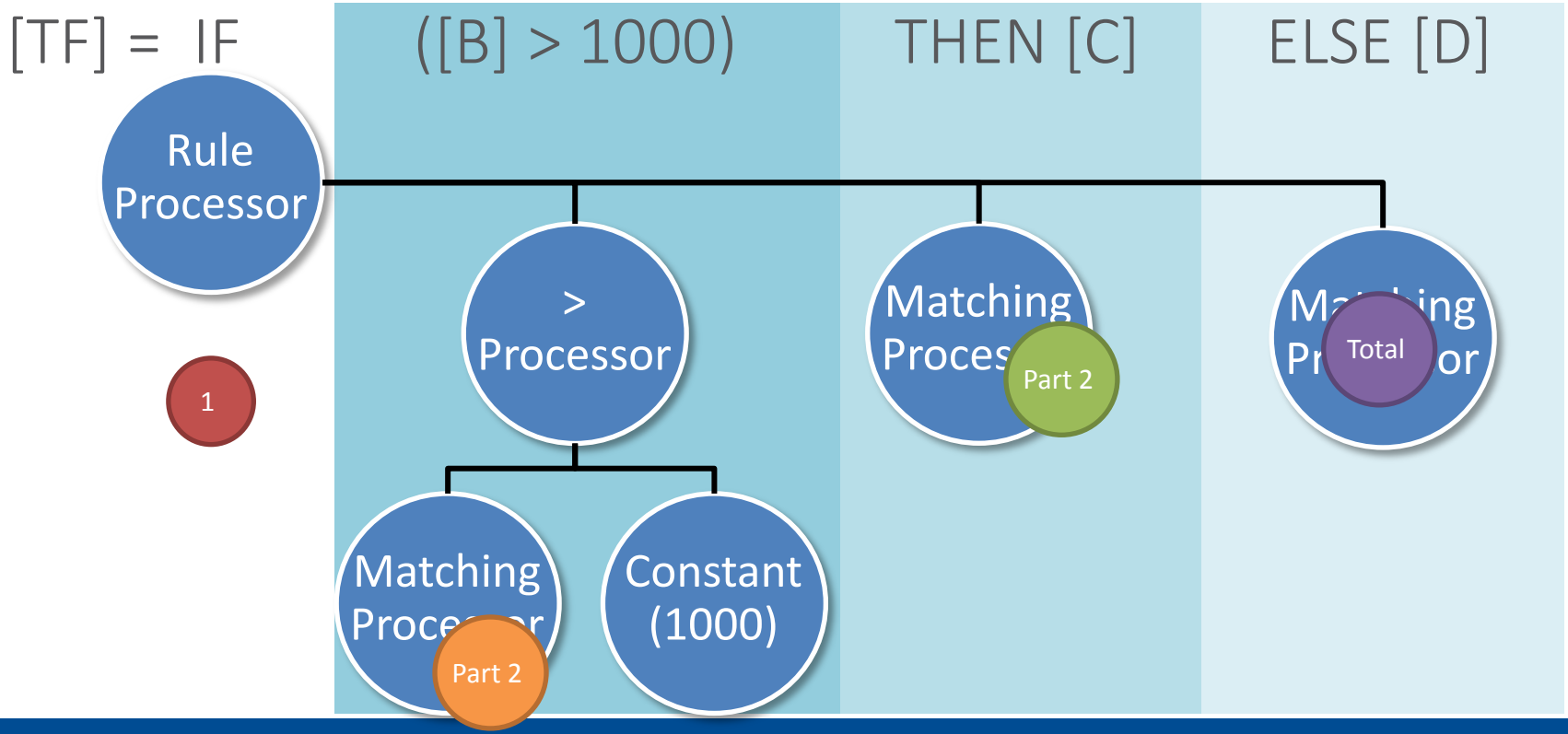
# Streaming in Conditional Rules



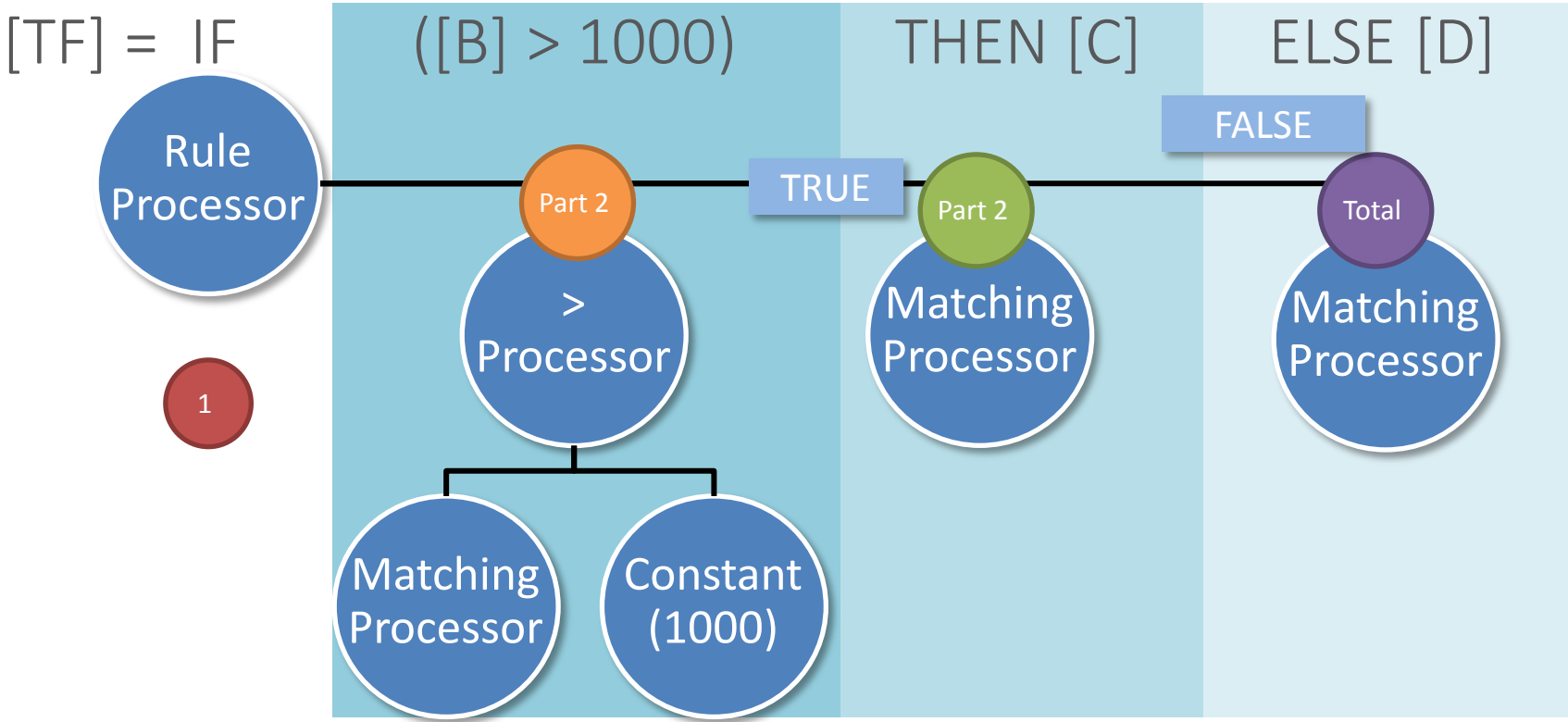
# Streaming in Conditional Rules



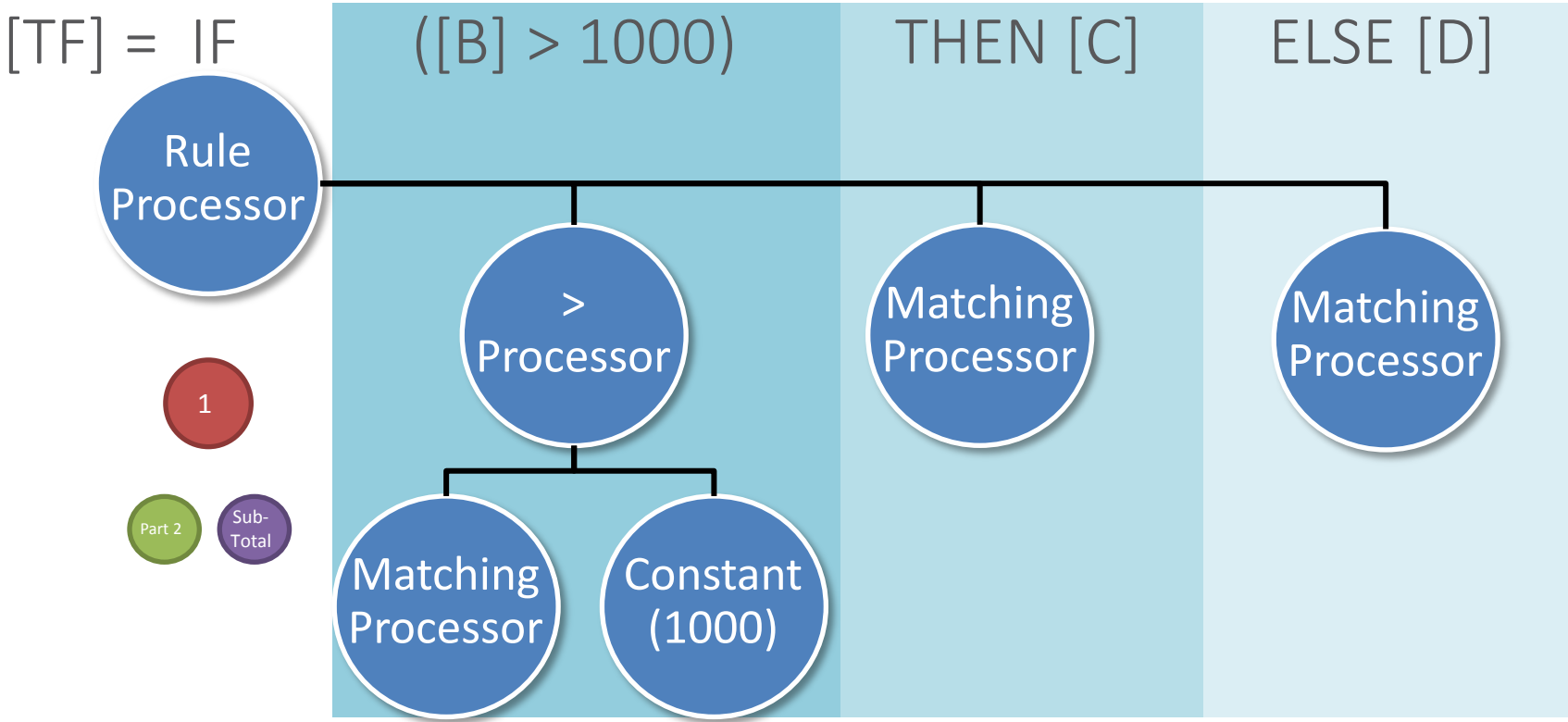
# Streaming in Conditional Rules



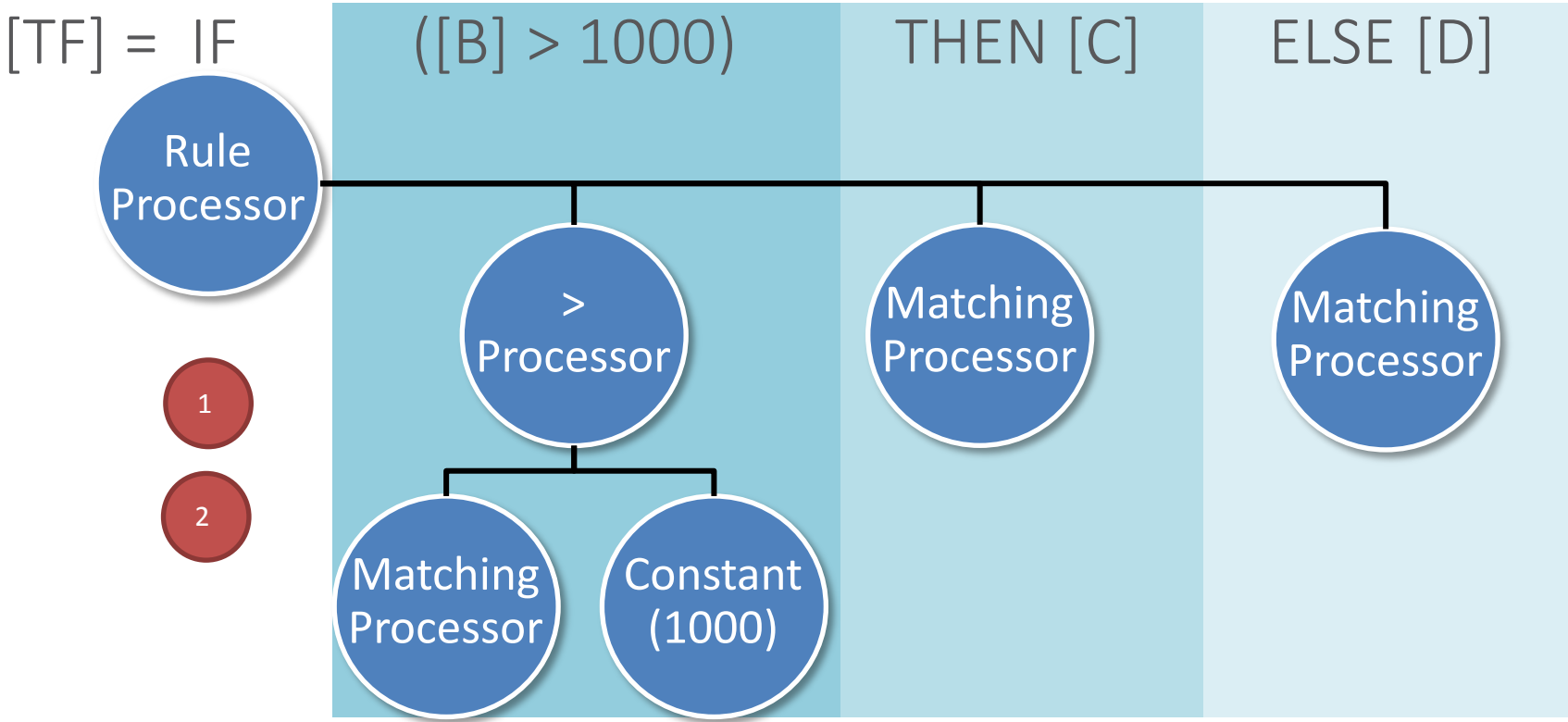
# Streaming in Conditional Rules



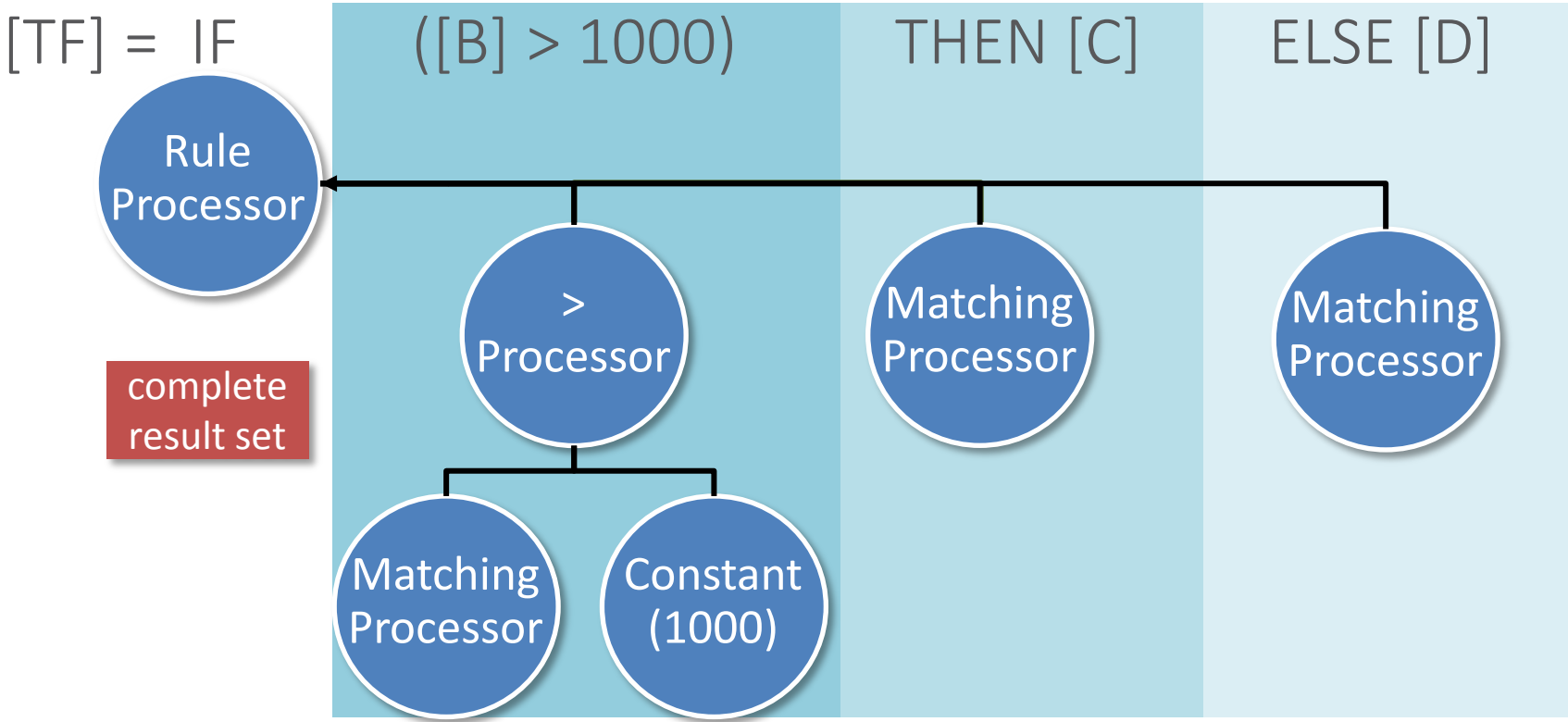
# Streaming in Conditional Rules



# Streaming in Conditional Rules



# Streaming in Conditional Rules

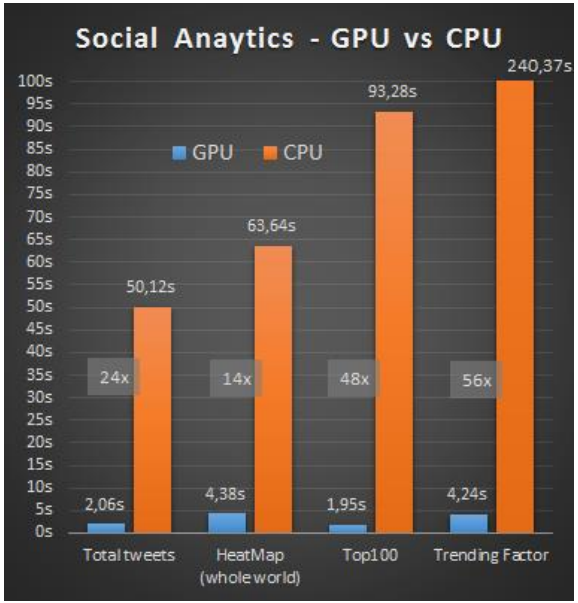


# Features: Performance with GPU

1. Top 100
2. Heat Map
3. Trending Factor

! Speed-up up to 50x and more!

! In-GPU-Memory database & calculations



# Jedox Social Analytics Showcase

Jedox Social Analytics App download at:
www.jedox.com/en/jedox-social-analytics

GPU- & Social Analytics App-Team:

Alexander Haberstroh

Steffen Wittmer

Leo Mehlig