



**Gridfore**

reactive  
intelligent  
platform



# Fast and robust complex event processing using



*The future is coming*



**Gridfore**  
reactive  
intelligent  
platform

# The Image of Digital Business



## DIGITAL PRODUCTS

Personal offers based on client behavior patterns determined by the data which was consolidated from different sources such as payments, operational logs, web and mobile analytics, open data of social networks and external analytical services



## THE IMPROVEMENT OF BUSINESS PROCESSES

Monitoring and forecasting of OLA and SLA process violations in real time: increase of service quality at the cash desk / operator line, prediction of service failures, fraud cases based on improper use of commercial equipment, etc.



## PROACTIVE INTERACTION WITH THE CLIENT

Flashing hot offers for cross-selling, churn prevention or risk management activities in response to subsequent updates to the customer profile or segmentation data marts, instead of traditional offline nightly recalculations based on outdated facts.



Gridfore  
reactive  
intelligent  
platform



# Classic BI solutions become ineffective

The path of data from source to business is **long,**  
**winding and expensive**

lengthy

Business can't use the data it needs,  
and when it really needed

Investments in development and infrastructure return **less**  
**value**

Customer profile quickly **loses relevance**

The cost of data is increased by  
**degraded performance**

expensive

Online forecasting of process failures **is not fast**  
**enough** to make proper reaction

**High cost** of real-time solutions

Just offline marketing campaigns

ineffective

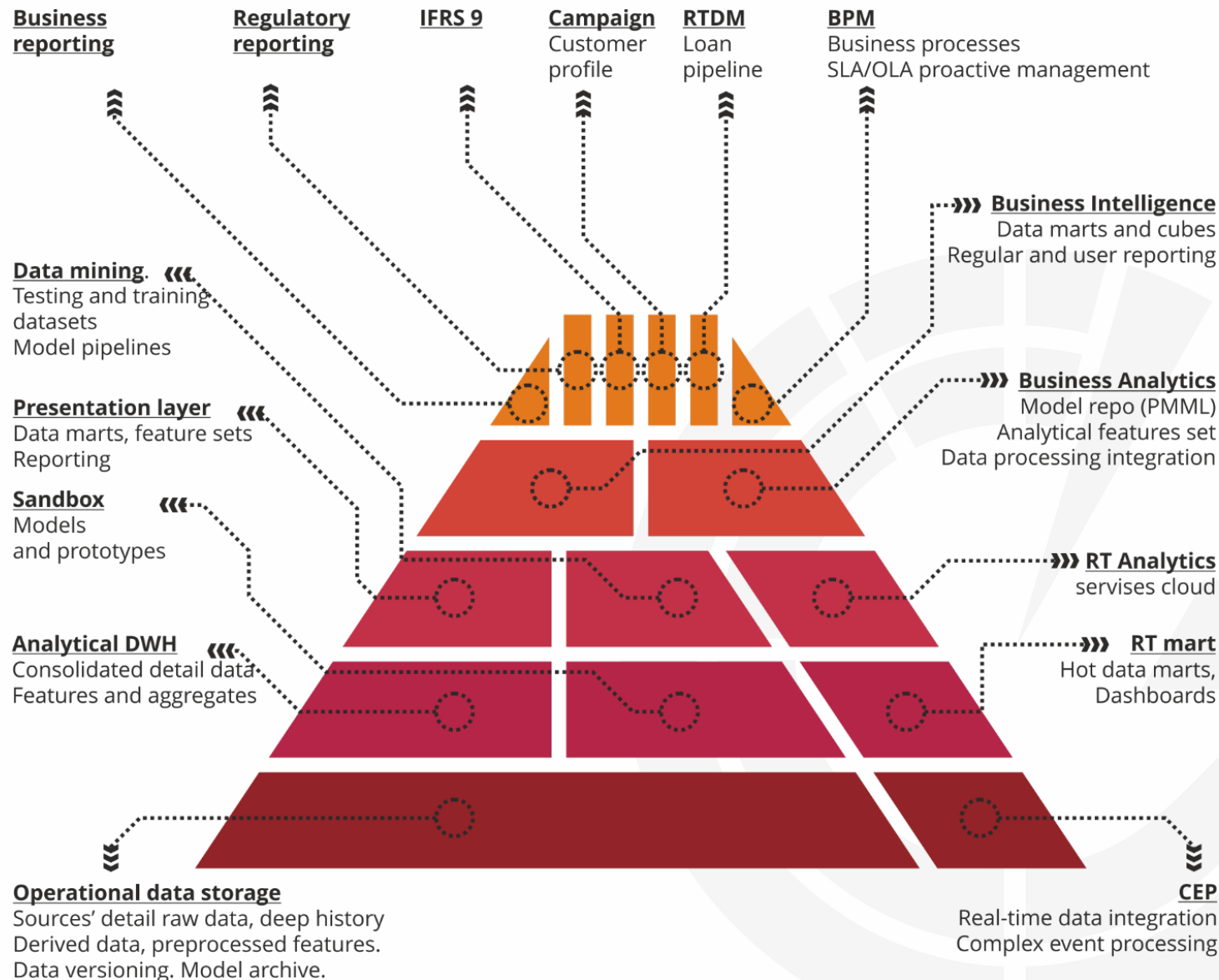
The data preparation time  
**overlaps the night** gap

# LOOKING OVER BIGDATA TOMB



**Gridfore**

reactive  
intelligent  
platform



# 5

# THE MAIN PRINCIPLES FOR MODERN ANALYTICAL PLATFORM



**Gridfore**  
reactive  
intelligent  
platform



Convergence of offline and real-time analytics



Factory of high-load analytical services with short Time2Market and good scaling



Support for modern data exchange standards, including IoT



Deep integration into business processes, corporate services and customer interaction tools



Support of data mining tools popular with analysts and data scientists



# RIGHT TOOL THAT FITS

We have redesigned the classic architecture of analytical solution to ensure maximum compliance with business objectives:

- Real-time sources and ultra-fast event triggering
- Low Time2Market to implement and test new ideas
- Simplification and cheapening of the operation and scaling process
- Compatibility with popular analysts toolkits
- Wide possibilities for integration
- Safe “sandbox” for experiments

Gridfore is the platform that can make value from both BigData and FastData

- ✓ Simple script language for application developing
- ✓ Rich ETL toolkit
- ✓ Smart data flow management
- ✓ Fast complex event processing
- ✓ Data Quality **beta**
- ✓ Machine Learning **beta**



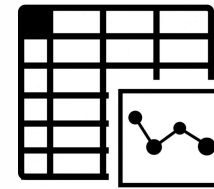
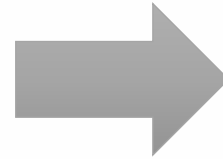
# EXTRACT THE RELEVANT ANALYTICS ON THE FLY

*Apache Ignite in-memory data Grid - fast and powerful data processing*



**1000+ GB**

*(raw source data)*



**< 1 GB**

*(useful features)*



In classical ETL solutions, the data must be loaded first, converted to a specific format, and then it is available to use.



Gridfore immediately identifies the relevant analytics directly **at load time**, combining various sources on the fly. We work with **near zero data latency**, allowing use of powerful analytics at the exact time we got the data from the source. We separate the useful features from bulk raw unstructured data, **reducing the resulting volumes** of useful analytics by thousand times, making subsequent processing much **faster and efficient**.



**Gridfore**

reactive  
intelligent  
platform



*ETL that works with any data source*



**Gridfore**  
reactive  
intelligent  
platform

# USE **ALL** YOUR DATA, REGARDLESS OF ITS FORM



You can use any data source - structured and unstructured - regardless of where and how it holds data: databases, network drives, SFTP, etc.



**Structured (DBMS)**



**Semi-structured** (XML, JSON, HTML etc.)



**Full-text telemetry (logs)**



**Reports and documentation**



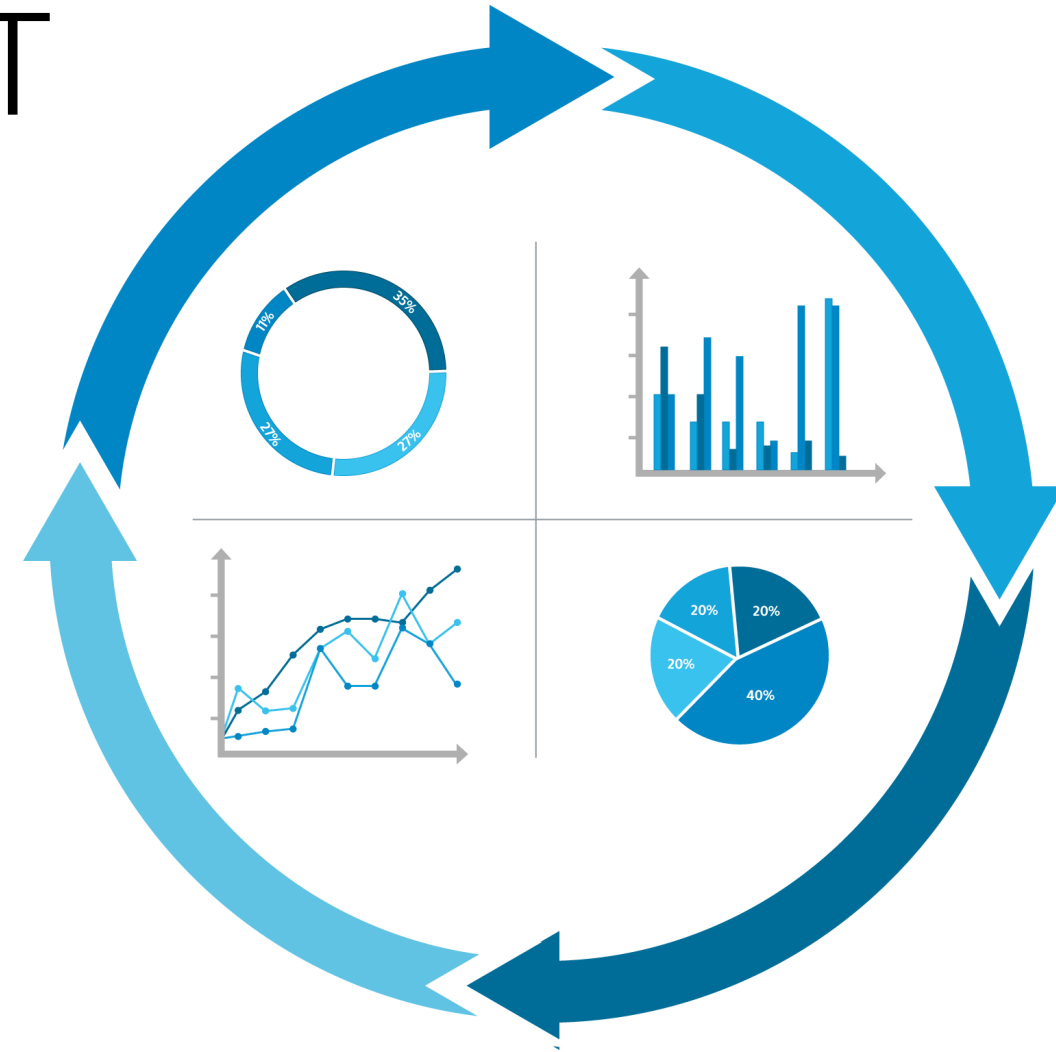
**External services (REST, SOAP)**





# INSTANT REACTION ON BUSINESS EVENT

The complex event triggering mechanics allows you to make intraday reporting, create real-time analytical services, conduct precise targeted marketing campaigns, proactively respond to emergency situations and make predictions based on the most relevant data you can have.





# DATA HYGIENE IN REAL TIME

Data quality checks – both standard and based on business rules - are performed on time you need it, just before or immediately after further processing

- ⚙ Simple rules (nulls, reference lists, ranges)
- ⚙ Complex cross-checks
- ⚙ Business rules
- ⚙ Trend control
- ⚙ Historical checks
- ⚙ Managing data flows based on the check results



*Gridfore Orchestrator – data flow  
management module*



**Gridfore**  
reactive  
intelligent  
platform

# FILIGREE CONDUCTING THE DATA FLOW



A special component of our platform monitors the dependencies between the data flows and starts updates of data marts in corresponding order, ensuring consistency, as soon as something important has changed.

Thanks to GO, you can also manage your own processes, reacting to new data from the sources, changes in data or CEP events triggering.



# MONITOR THE HEALTH OF YOUR BUSINESS

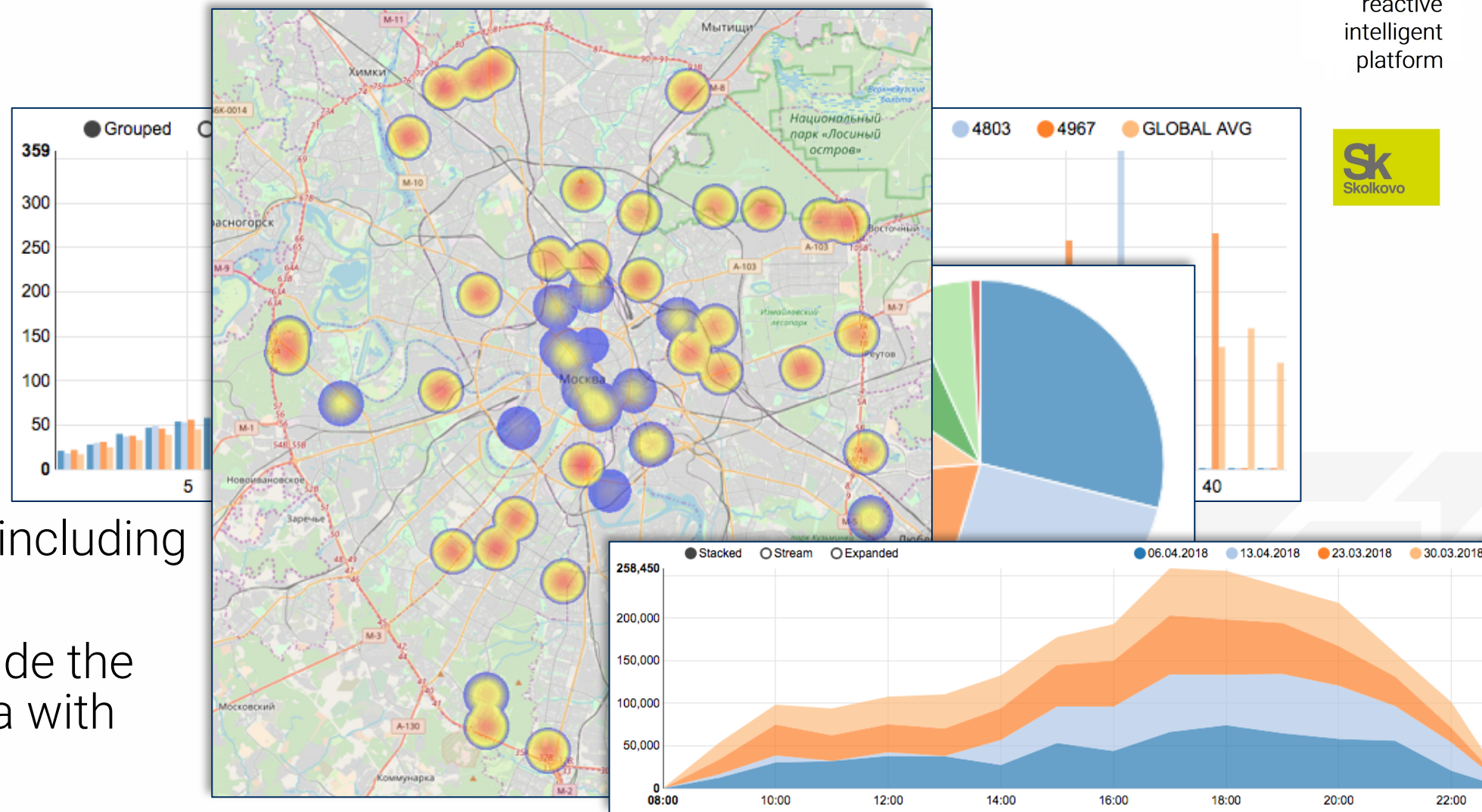
reactive intelligent platform



Real-time reporting on the state of things in your business here and now.

Use any BI tools, including open source.

Gridfore will provide the most current data with minimal latency.



# COMMERCIAL EQUIPMENT INTEGRATION

## Real-time cash register / ATM monitoring!

- Does not affect Infrastructure and SW performance
- No CR software modification needed
- 100+K devices at real-time
- The whole picture of the running business – at one place

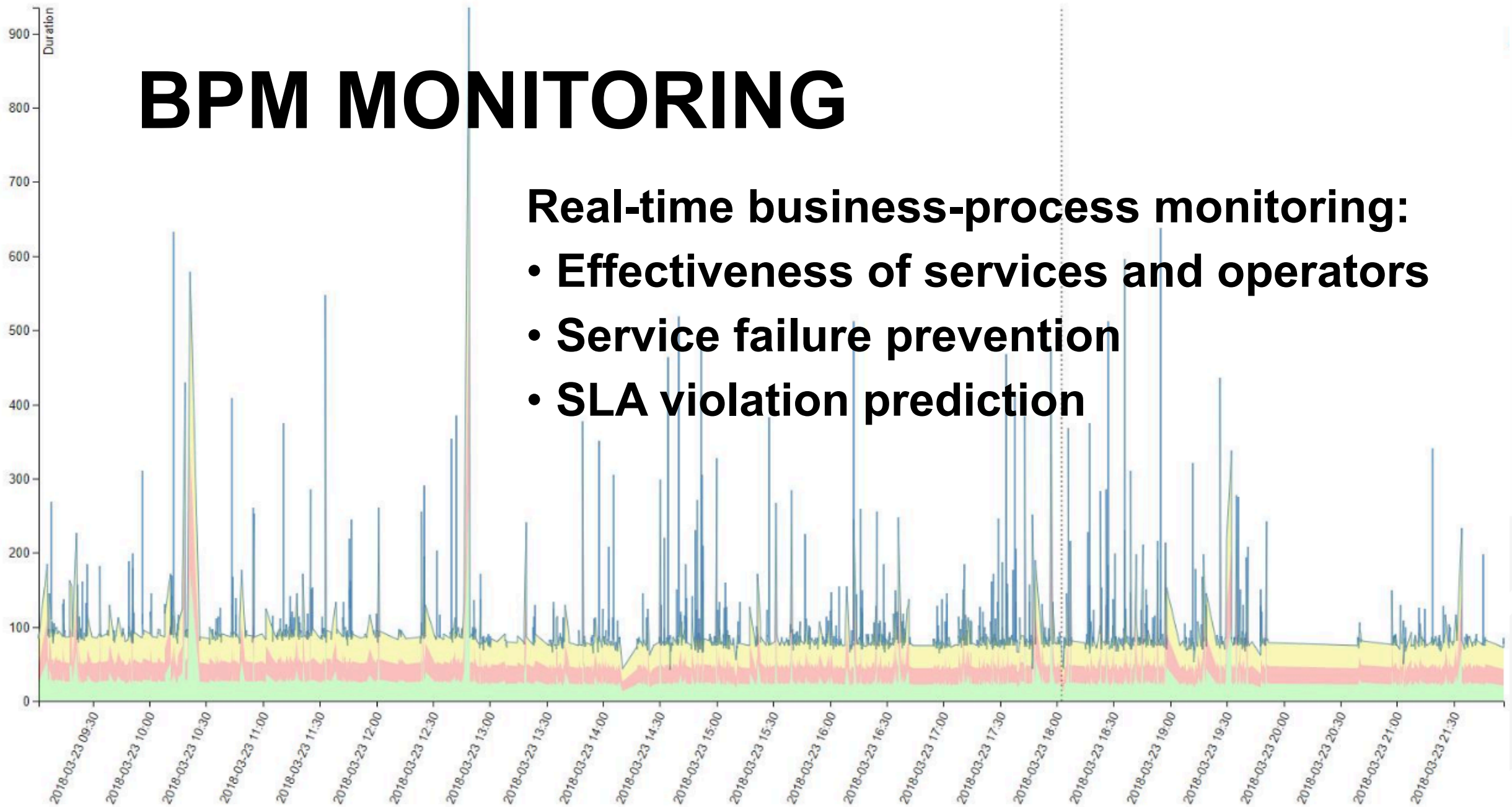
23	ХЛЕБЫ ФРИН КРИСП	2 * 107.99	=215.98..Б
24	ШОК РИТ АЛЬП МОЛ1	1 * 69.99	=69.99..А
25	ШОК РИТ АЛЬП МОЛ1	1 * 69.99	=69.99..А
26	ШОК РИТ АЛЬП МОЛ1	1 * 69.99	=69.99..А
27	ШОК РИТ АЛЬП МОЛ1	1 * 69.99	=69.99..А
28	БАЗИЛИК КР.БОГОС.	1 * 49.99	=49.99..Б
29	ШОК РИТ АЛЬП МОЛ1	1 * 69.99	=69.99..А
30	КРИМА ПИЦЦА ТОС1	1 * 69.99	=69.99..Б
31	ПЕЧЕРЬЕ ГУЛЛОН 17	2 * 92.99	=185.98..А
32	ПЕЧЕРЬЕ ГУЛЛОН 12	2 * 102.99	=205.98..А
33	ПЕЧЕРЬЕ ГУЛЛОН 12	2 * 102.99	=205.98..А
34	ТВОРОГ ПРОСТОК 2	2 * 96.99	=193.98..Б
35	СЫР АЛЬМЕТТЕ БЕЛ.	1 * 92.99	=92.99..Б
36	СЫР АЛЬМЕТТЕ БЕЛ.	1 * 92.99	=92.99..Б
37	СЫР АЛЬМЕТТЕ СЛ.1	3 * 92.99	=278.97..Б
38	СЛИВКИ ПАРМ 11 0.	2 * 39.49	=78.98..Б
39	СУПЕРЭКСТРА 8 11	6 * 18.99	=113.99..Б
40	СМЕТАН.ИВ 10 180	2 * 31.99	=63.98..Б
41	НАСЛОС Т.03.62,6 1	1 * 129.99	=129.99..Б
42	АЕРИКОСЫ	0.430 * 189.99	=81.70..А
43	СЫРОК СВИТ.ВАНИЛЬ	12 * 28.29	=339.48..Б
<b>ИТОГ</b>			<b>≐6719.17</b>
ЭЛЕКТРОННЫМИ ПОЛУЧЕНС:			≐6719.17
ЭЛЕКТРОННЫМИ			≐6719.17
А:СУММА НАС 182			≐568.31
Б:СУММА НАС 102			≐272.14
РН КИТ:С000628092036122			29.05.18 14:49
ЗН КИТ:С548420001019051			СМЕНА:348 ЧЕК:15
КАССОВЫЙ ЧЕК/ПРИКОД			
ИНН:7743543761			ИН:8710000100802956
Кассир:АТКШОВА А.			#5387
Сайт:ИНС:			www.nalor.ru
СНО:ОСН			ОД:С0584 ОП:0203868659

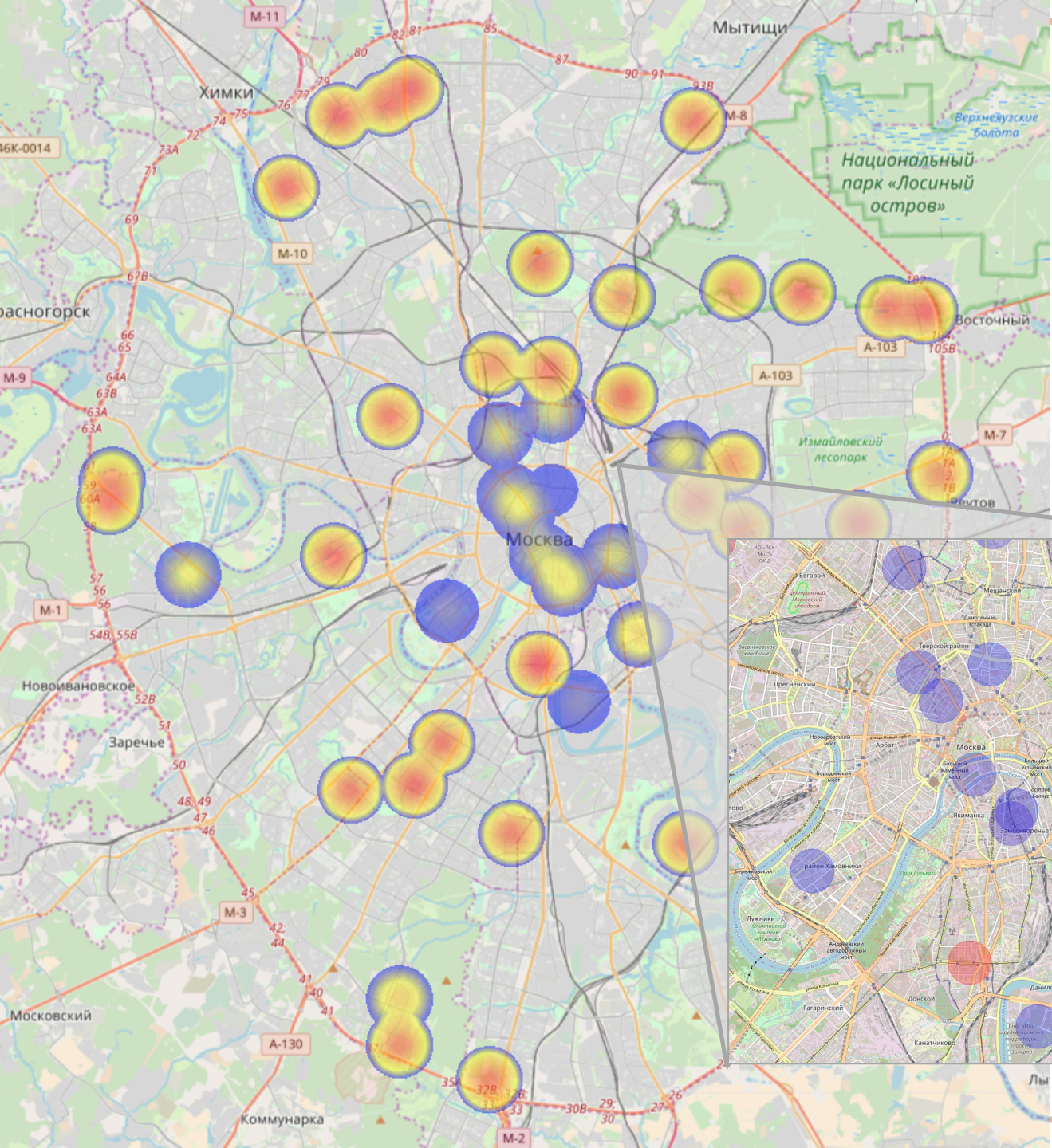


```
^-- Node [id=9945f9a6, name=datagrid, uptime=08:27:16
^-- H/N/C [hosts=5, nodes=6, CPUs=80]
^-- CPU [cur=1,87%, avg=12,55%, GC=0%]
^-- PageMemory [pages=3167]
^-- Heap [total=10480MB, free=15,36%, comm=12288MB]
^-- Non heap [used=221MB, free=-1%, comm=230MB]
^-- Public thread pool [active=0, idle=15, qSize=0]
^-- System thread pool [active=0, idle=16, qSize=0]
^-- Outbound message queue [size=0]
00:01:31.591 [grid-timeout-worker-#39%datagrid%] INFO o.
00:01:31.591 [grid-timeout-worker-#39%datagrid%] INFO o.
11:05:58.415 [main] INFO o.s.b.a.e.mvc.EndpointHandlerMa
11:05:58.421 [main] INFO o.s.b.a.e.mvc.EndpointHandlerMa
11:06:00.402 [main] INFO o.s.j.e.a.AnnotationMBeanExport
11:06:00.408 [main] INFO o.s.j.e.a.AnnotationMBeanExport
11:06:00.411 [main] INFO o.s.j.e.a.AnnotationMBeanExport
11:06:00.425 [main] INFO o.s.j.e.a.AnnotationMBeanExport
11:06:00.431 [main] INFO o.s.j.e.a.AnnotationMBeanExport
11:06:00.433 [main] INFO o.s.b.a.e.jmx.EndpointMBeanExpo
11:06:00.442 [main] INFO o.s.c.s.DefaultLifecycleProcess
11:06:00.444 [main] INFO o.s.b.a.e.jmx.EndpointMBeanExpo
19:20:57.711 [http-nio-0.0.0.0-8080-exec-10] INFO c.g.gr
19:20:57.715 [http-nio-0.0.0.0-8080-exec-10] INFO c.g.gr
```

# BPM MONITORING

- Real-time business-process monitoring:**
- Effectiveness of services and operators
  - Service failure prevention
  - SLA violation prediction





# SUPPORT SERVICES



**Gridfore**  
reactive  
intelligent  
platform



**Identify problems in the operation, before the real incident occurs**

**Predict real-time failures**

# DATA THAT DRIVES MONEY



**Gridfore**

reactive  
intelligent  
platform



Real-time monitoring and forecasting of OLA and SLA violations in business processes: equipment failures, degradation in BPM processes

Marketing campaigns that react on the feedback or behavior of the customer

Interactive methods of churn counteraction

Targeted offers based on the most actual and complete customer profile

Reacting on fraud schemas

Real-time scoring for various RTDM cases

A graphic with the text 'BIG DATA' in blue and 'BIG PROFITS' in yellow. To the right of 'BIG PROFITS' is a small icon of a banknote with a person's face. The entire graphic is set against a background of a large, faint gear.

**BIG DATA**  
**BIG PROFITS**

~~PROBLEMS~~



*Simple as that*



**Gridfore**

reactive  
intelligent  
platform

# FITS TO YOUR INFRASTRUCTURE

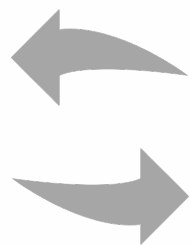
Gridfore easily integrates both classic and future data sources, and can be used best to prepare and upload data to DWH and BI solutions. Gridfore can function as unified operational data layer, to deliver both prepared historical data and low-latency raw operational data.



Without being a competitor to existing BI solutions, Gridfore can improve their efficiency and actuality, expand capabilities, significantly reduce the peak workload pressure on the infrastructure.



sources



Enterprise DWH & BI

*Easy to do what business needs*



**Gridfore**

reactive  
intelligent  
platform

# QUICKLY TEST YOUR NEW IDEAS

Application development using simple and expressive scripting language (DSL). Hot delivery into production provides high Time2Market and allows you to do it without the big help from IT professionals.

Using DSL requires no programming skills and is aimed at solving specific application tasks by devops engineer or analyst.

The complexity of DSL is similar to SQL, but it is more powerful in features.

Gridfore DSL also provides a quick code delivery process for mathematical model training algorithms (using PMML), data loading and processing workflows, event triggering, maintaining real-time data marts and so on.





**Gridfore**

reactive  
intelligent  
platform



Create services monetizing your data and provide them to partners to get additional income.

You can provide:

- High data quality
- As relevant as possible
- No personal data transfer

CREATE YOUR OWN OR USE  
PARTNERS'  
REAL-TIME ANALYTICAL  
SERVICES



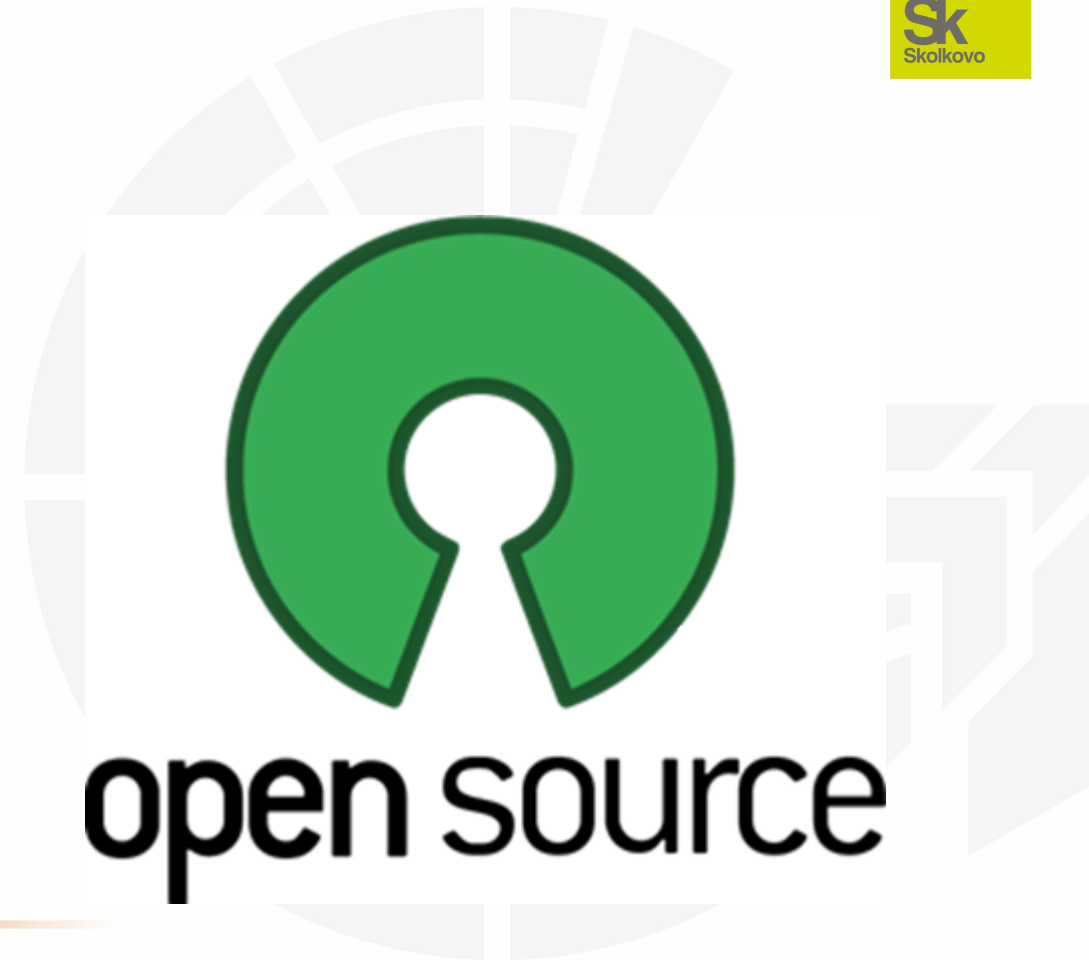
*Built on open technologies*

# MORE OPTIONS less capital investment

- Product includes only open source components
- Proven on popular Big Data solutions:
  - Hadoop (Hortonworks, Cloudera, Huawei FusionInsight)
  - Cassandra
- Many DBMS can be used as data storage
  - Postgres
  - Oracle
  - Teradata
  - any JDBC source



**Gridfore**  
reactive  
intelligent  
platform



Linear scale-out

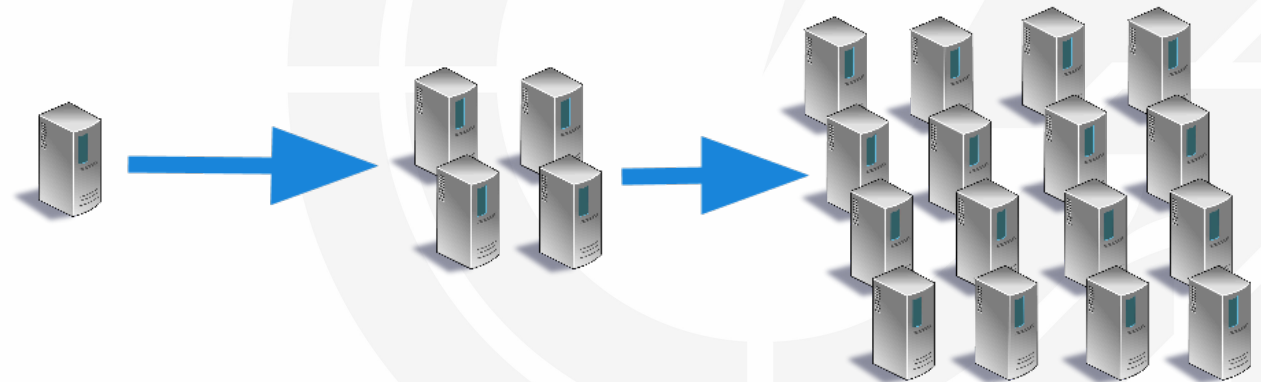


**Gridfore**  
reactive  
intelligent  
platform



# SCALE SOLUTION ALONG THE GROWTH OF YOUR BUSINESS

To scale the platform, you just need to connect additional hardware.  
Without stopping the processes



# TECHNOLOGY TESTED IN A REAL HIGHLOAD SOLUTION

The solutions developed by our team are successfully used by one of the World TOP-10 telecom for the implementation of complex event processing and real-time analytics



VEON

*We worked with Gridfore Company during 2017, aimed to build and deliver real-time data ingestion and complex event processing automation SW, in order to enable Realtime commercial analytical use cases. As a result of cooperation, we delivered Data Management Platform, satisfying our requirements and working in commercial operations. We found Gridfore as reliable, qualified partner in complex Java SW development.*

**Andrey Zheliezniak**

Head of Data management platform HUB,  
VEON Eurasia



**Gridfore**

reactive  
intelligent  
platform



# TESTED IN THE MOST SEVERE CONDITIONS

Together with Huawei, Gridfore solution was tested on Huawei FusionInsight and received the highest ratings.  
Gridfore is Huawei ISV partner.

*Glad to confirm the 1st phase – functional tests – completed with success today. The Gridfore Reactive Intelligent Platform (GRIP) solution functionality have been tested around 2 use cases. They mimic DW/ETL data flows with fast data transformations in memory on-the-go. I believe the FI-HD is now a “persistent” candidate to the Gridfore solution stack.*

*Thanks to Gridfore Team for the comprehensive demonstration and for such a vigorous inception!*

**Iurii Kuksa / 尤里**  
Solutions Architect, Finance



**Gridfore**  
reactive  
intelligent  
platform



**HUAWEI**

**4 weeks**

from the 1st day of work till the first MVP

**3 months**

to deliver your valuable business cases in production

**We use Agile approaches to obtain MPV at the end of each iteration which lasts no more than 2-3 weeks**

**As a result of the implementation, you obtain working solution, the application code delivery processes, CI/CD based on best DevOps practices**

**As part of the support, you automatically get all the new features of the platform**





**Gridfore**  
reactive  
intelligent  
platform

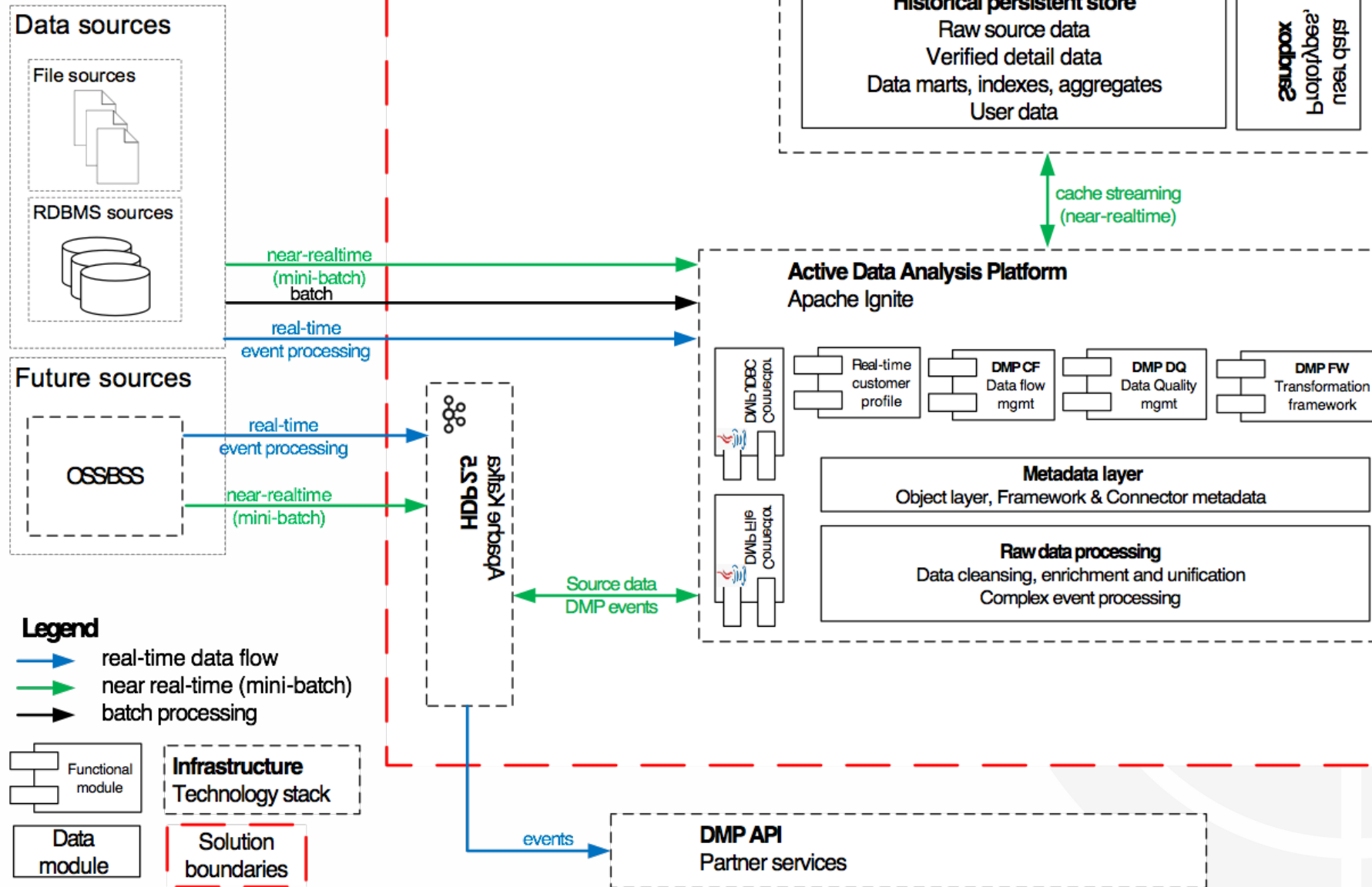


# Complex Event Processing Architecture



# Gridfore

reactive  
intelligent  
platform

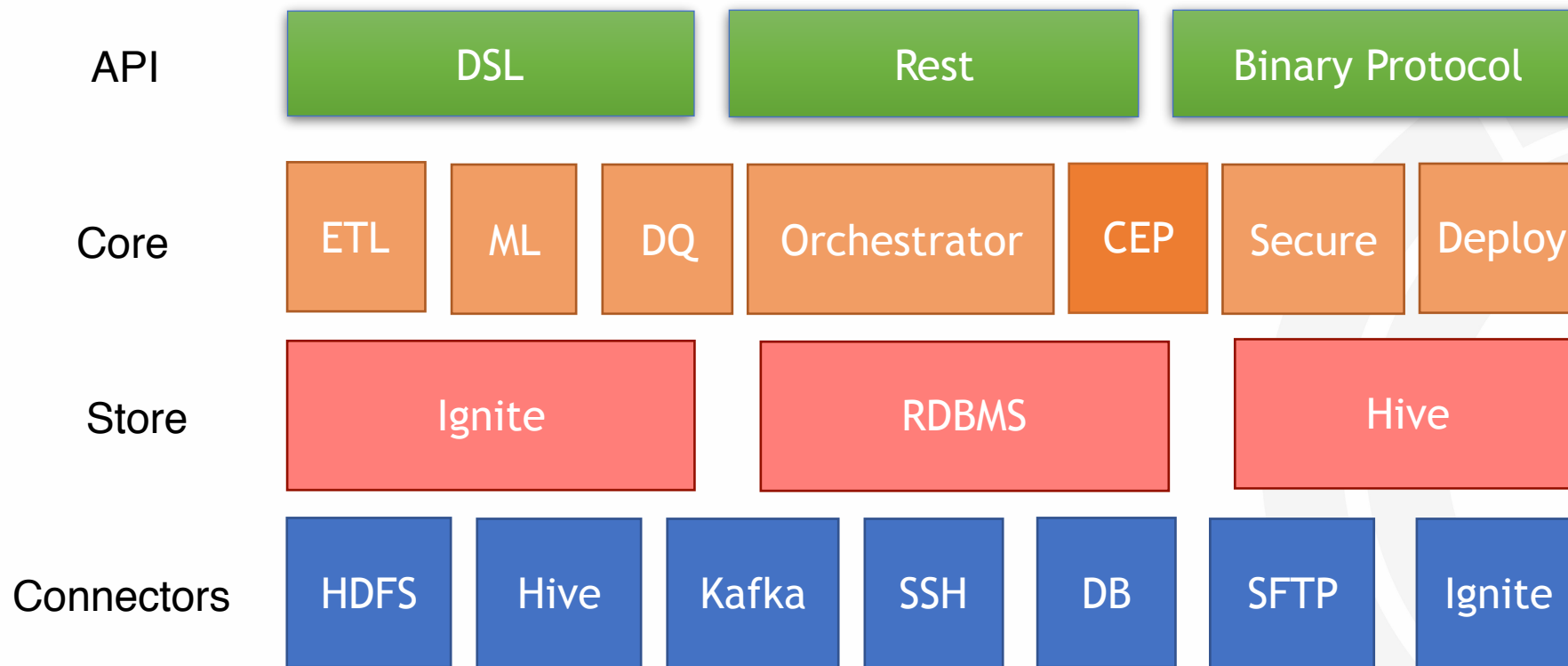




**Gridfore**  
reactive  
intelligent  
platform



# Gridfore Solution





**Gridfore**

reactive  
intelligent  
platform



# CEP Problems

Problem \ System type	Classic systems	Modern system
Hight Load (scaling)	-	+
Easy to develop (in-house team cost, Time2Market delay)	+	-
Less code	-	-
Fast code delivery	-	-
Free from IO bottlenecks (synchronisation using RDBMS)	-	+
Free from locks (trigger mechanism)	-	-
Free from network bottlenecks	-	-
Data & code collocation	-	+
Row based	+	-
Stream	-	+
CEP in most cases is scan by index	+	+



**Gridfore**

reactive  
intelligent  
platform



# CEP Requirements

Problem \ System type	Solution
Hight Load (scaling)	Easy scaling
Easy to develop (in-house team cost, Time2Market delay)	DSL
Less code	DSL
Fast code delivery	Custom deploy mechanism
Free from IO bottlenecks (synchronisation using RDBMS)	Data in IMDG
Free from locks (trigger mechanism)	No Locks (Additional Thread Pool)
Free from network bottlenecks	Affinity
Data & code collocation	Affinity
Row based	
Stream	IgniteDataStreamer
CEP in most cases is scan by index	Index

# Why apache Ignite™



**Gridfore**

reactive  
intelligent  
platform

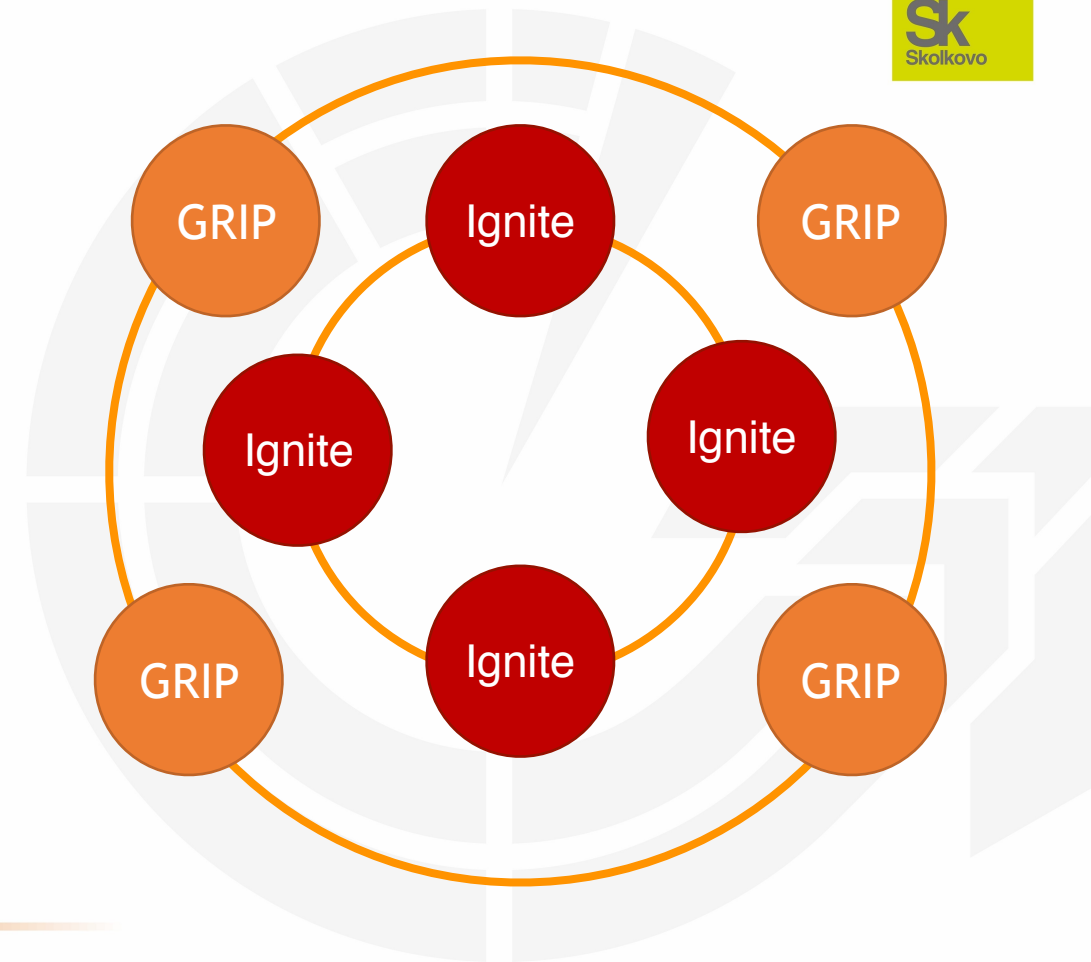


Problem \ System type	Solution
<b>Hight Load (scaling)</b>	<b>Clustered</b>
Platform maturity and risks	Active community. Fast-paced product
Less code	Flexible API
<b>Fast code delivery</b>	<b>ClassLoader</b>
Free from IO bottlenecks (synchronisation using RDBMS)	Data in IMDG
Free from locks (trigger mechanism)	EventListener (use cache thread)
<b>Free from network bottlenecks</b>	<b>Affinity</b>
<b>Data &amp; code collocation</b>	<b>Affinity</b>
Row based	IgniteCache#put IgniteCache#get
<b>Stream</b>	<b>IgniteDataStreamer</b>
CEP in most cases is scan by index	Index



# Scaling

- GRIP work over Ignite cluster
- Using Ignite engine to build cluster
- Cluster Health depends on TX health (<https://issues.apache.org/jira/browse/IGNITE-6380>)





**Gridfore**

reactive  
intelligent  
platform



# DSL

- Less code
- Abstraction layer
- Easy to use (less requirements to employee)
- Generate high optimised code based on data model

```
//trigger for building vip customers
trigger 'VIP Customers' of {
  // check VIP status
  if (CUSTOMER_STATUS in ['VIP', 'VIP+', 'VIP2018'])
  {
    // ACID update by key CUSTOMER_ID
    transaction CUSTOMER_VIP[CUSTOMER_ID] {
      CUSTOMER_NAME // add field
      TRANSITIONS_COUNT++ // update counter
      UPDATE_DATE = now

      // send message
      message 'Change VIP Customer' of {
        // get data from CUSTOMER_VIP
        data CUSTOMER_VIP

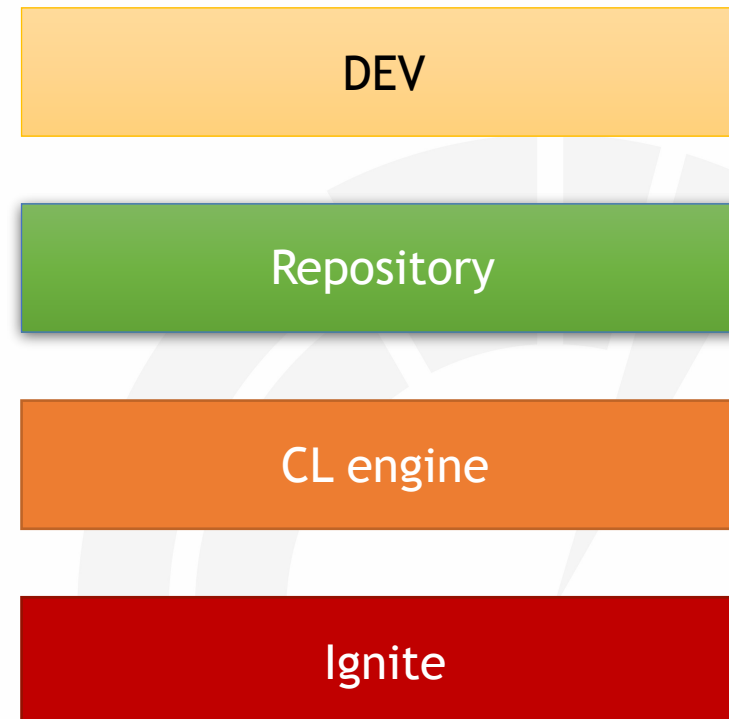
        // publish to Kafaka
        publish it to KAFKA topic 'VIP' format
        JSON
      }
    }
  }
}
```





# Custom deploy mechanism

- Use Ignite CL engine with custom implementation
- DEV publish code from IDE
- GRIP (re)load class on fly with
- GRIP store metadata - (re)load metadata





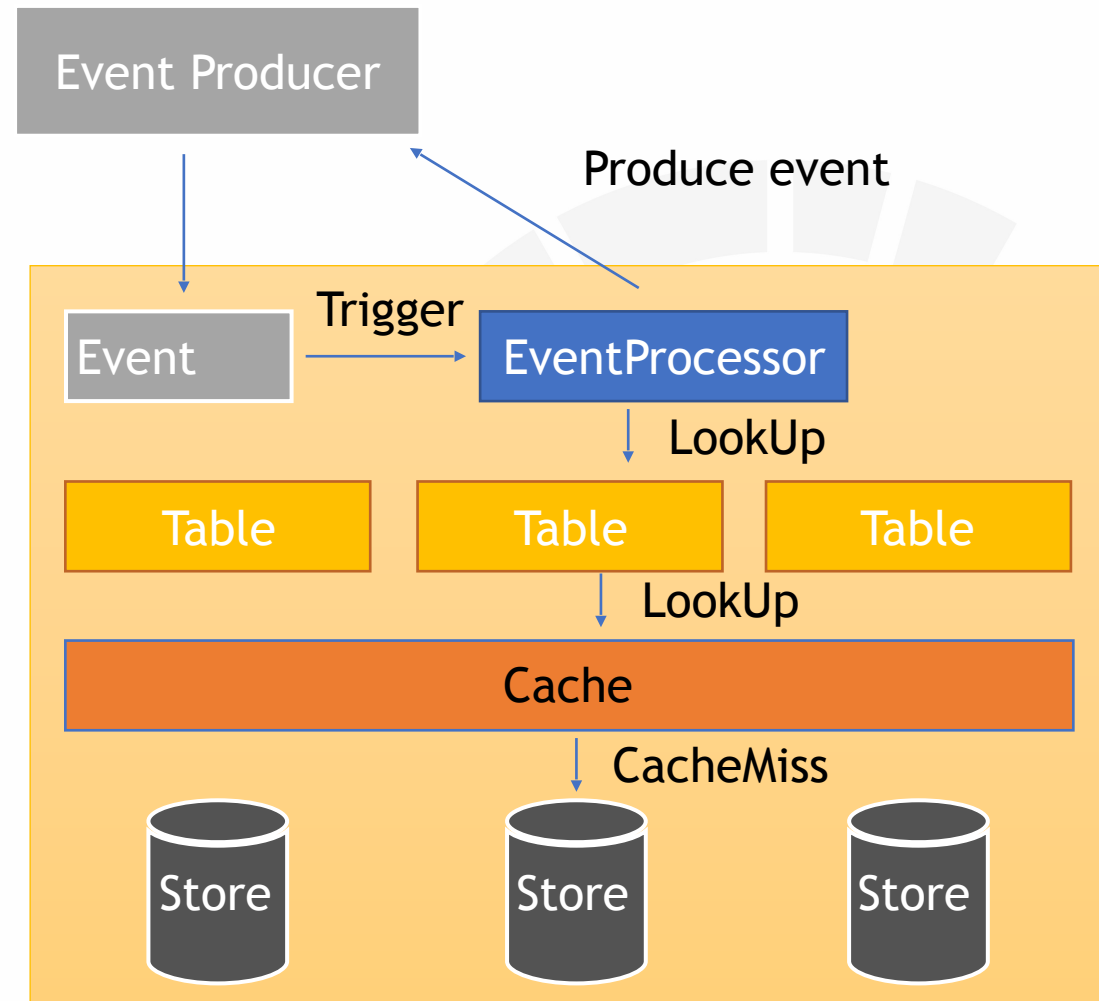
**Gridfore**

reactive  
intelligent  
platform



# Data in RDBMS

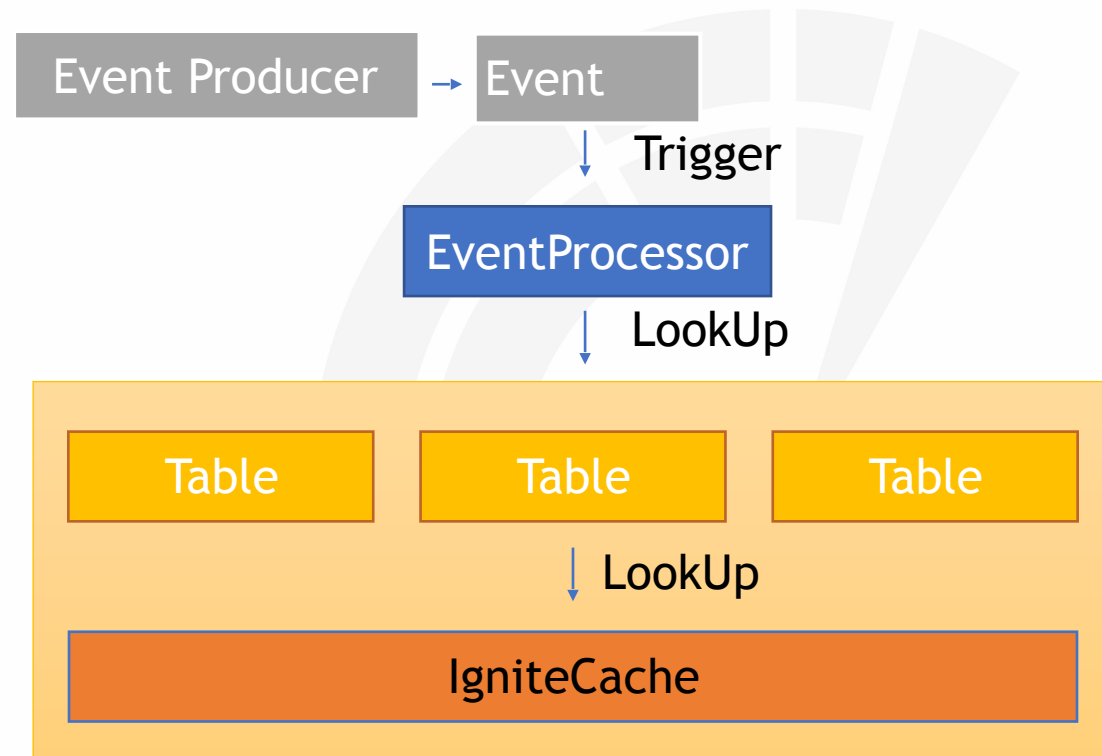
- EventProcessor could be loaded per session
- LookUp into DISK is slow
- MVCC
- Data organisation (columnar?)
- Low memory cost





# Data in IMDG

- EventProcessor is always available
- Could be row based or batch
- High memory cost





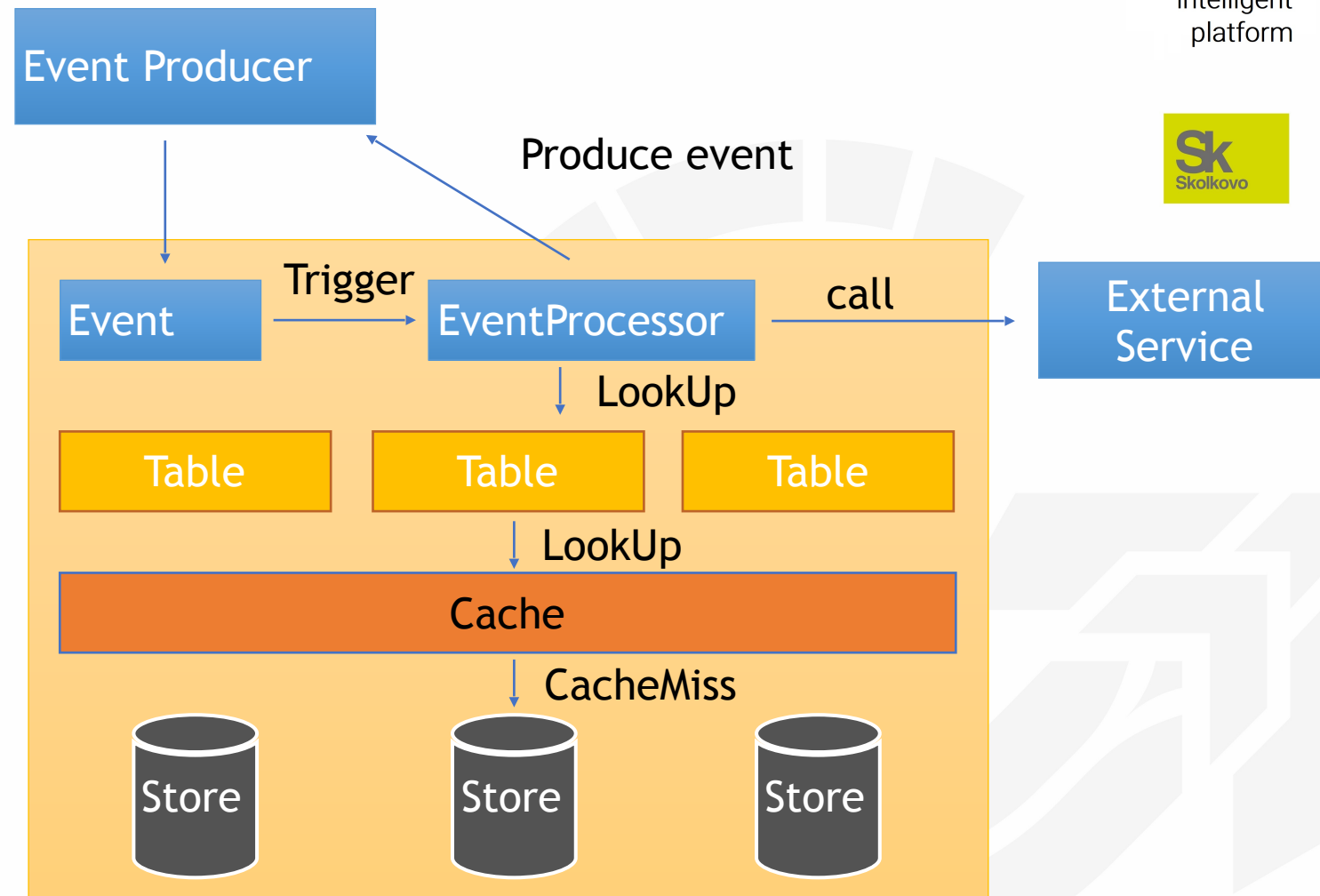
**Gridfore**

reactive  
intelligent  
platform



# Locks

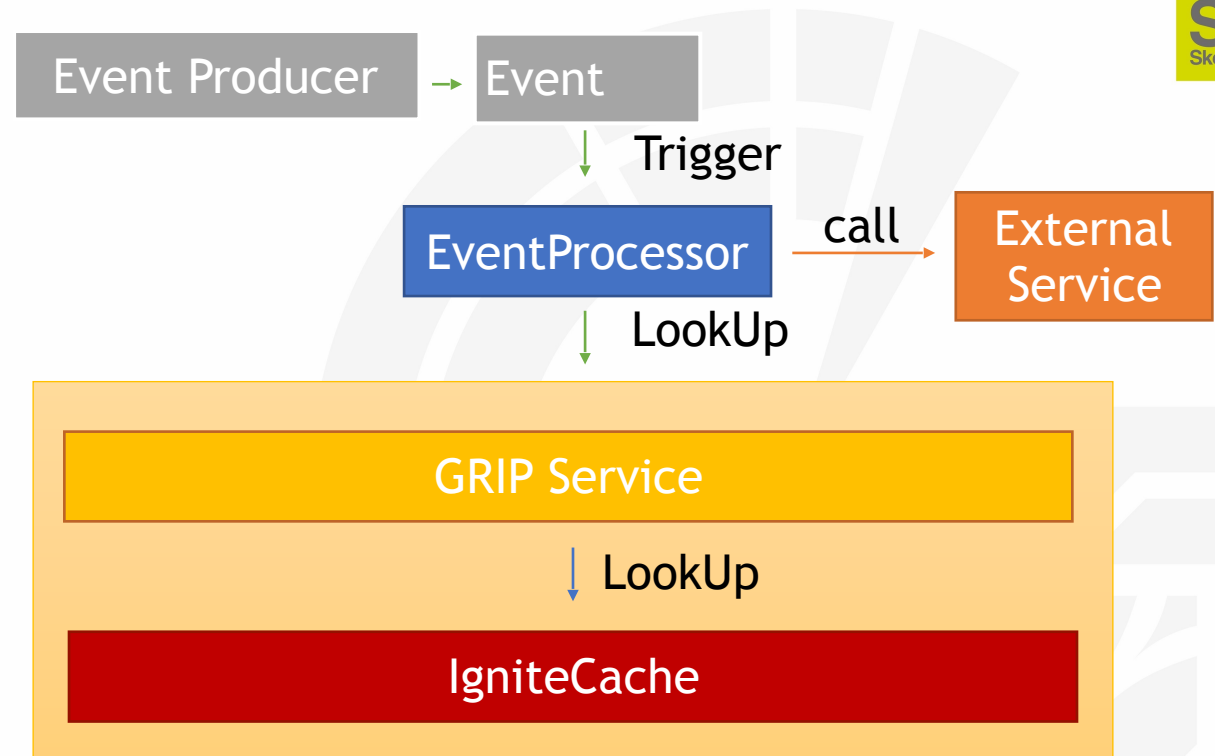
- EventProcessor block DB pool
- Heavy job could block DB pool - low DB performance





# Locks

- EventProcessor doesn't block DB pool
- Heavy job doesn't block DB pool - low DB performance





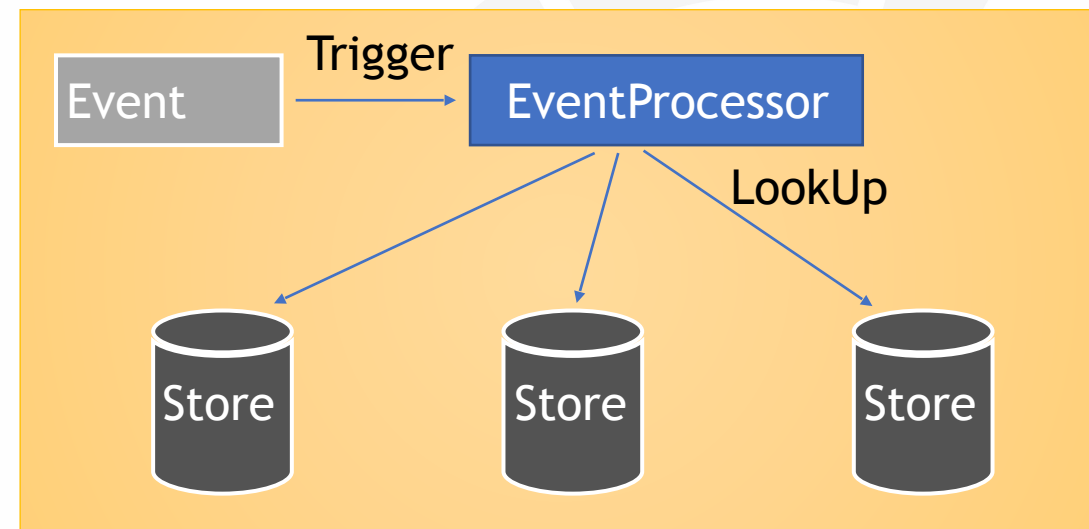
**Gridfore**

reactive  
intelligent  
platform



# Data Collocation

- Avoid network IO
- Compute jobs must be also collocated with data





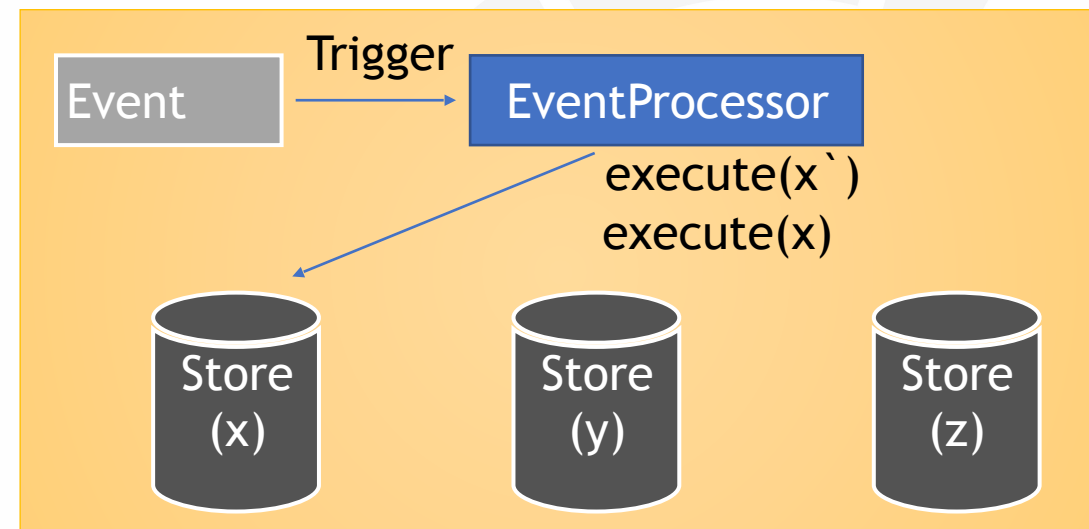
**Gridfore**

reactive  
intelligent  
platform



# Data Collocation

- Use affinity
- Execute code by affinity
- Partition strategy
- Logical partition





**Gridfore**  
reactive  
intelligent  
platform



# Compare with another systems

Problem \ System type	GRIP	Classic Systems	Modern Systems
Hight Load (scaling)	+	-	+
Easy to develop (in-house team cost, Time2Market delay)	+	+	-
Less code	+	-	-
Fast code delivery	+	-	-
Free from IO bottlenecks (synchronisation using RDBMS)	+	-	+
Free from locks (trigger mechanism)	+	-	-
Free from network bottlenecks	+	-	-
Data & code collocation	+	-	+
Row based	+	+	-
Stream	+	-	+





**Gridfore**

reactive  
intelligent  
platform



# Benefits

- Easy scaling with Ignite
- DSL to avoid tones of code
- Custom deploy mechanism with Ignite
- Data in IMDG
- No Locks for EventProcessor (Additional Thread Pool)
- Affinity could help to colocate data



**Gridfore**

reactive  
intelligent  
platform



# Problems

- Java related problems (GC pause)
- IMDG common problems
  - Loose data on critical fault (OOM etc.)
- Ignite have problems
  - No network data compression (resource cost)
  - Data compression (resource cost)
  - Cluster health (<https://issues.apache.org/jira/browse/IGNITE-6380>)
- Code and metadata versioning
- Strong consistency or high performance
- Network IO latency (nightly backups, admins reconfigure hardware)



**Gridfore**

reactive  
intelligent  
platform



# How we solve problems

Problem	Solution
Java related problems (GC pause)	<ul style="list-style-type: none"><li>• Monitoring</li><li>• async-profiler (<a href="https://github.com/jvm-profiling-tools/async-profiler/releases">https://github.com/jvm-profiling-tools/async-profiler/releases</a>)</li></ul>
Lose data on OOM or any fault	<ul style="list-style-type: none"><li>• Monitoring</li><li>• Async persist with low latency</li><li>• Restore state on cold start</li></ul>
Code and metadata versioning	Prototype
Strong consistency or high performance	Tradeoff
Network IO latency (nightly backups, admins reconfigure hardware)	Monitoring

# Performance

- ETL process that read from Kafka to update customer profile (3 EventProcessor)
- 3 nodes x (300GB Mem, Xeon E5 series 56 core)
- TX pessimistic repeatable read
- Custom Ignite 2.3 build



**Gridfore**

reactive  
intelligent  
platform



# Performance



**Gridfore**

reactive  
intelligent  
platform



	Per Thread	Per Node
Processed entries	2285951	68578530
Duration	120	120
RPS	19050	571488
CPU load	16 %	915 %



**Gridfore**

reactive  
intelligent  
platform



# Questions?

Viktor Khodiakov

+7 925 010 9263

[vkhodyakov@gridfore.com](mailto:vkhodyakov@gridfore.com)

© Gridfore, 2018

[sales@gridfore.com](mailto:sales@gridfore.com)

Bolshoy boulevard 42 bld. 1 office 1436, Skolkovo Innovation Center, Moscow, 121205, Russia