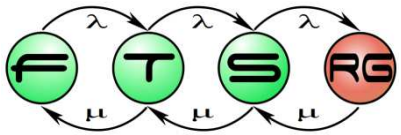
Complex event processing and the CoMiFin project

Gönczy László

gonczy@mit.bme.hu



CEP basics

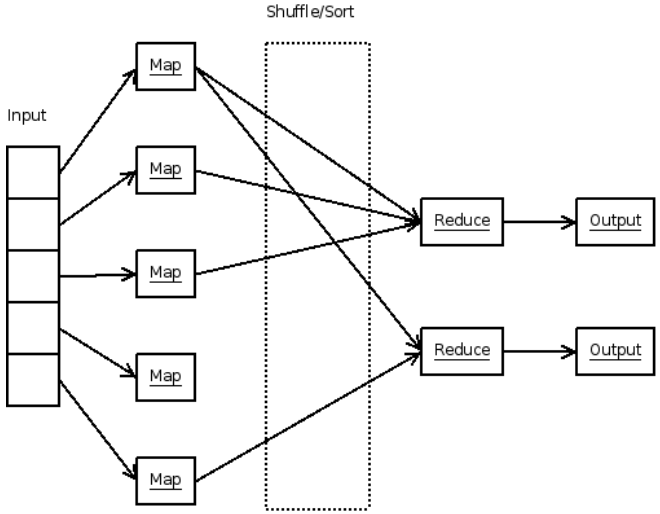
- Complex event
  - A logical composition of atomic events
- Main characteristics
  - Timing (e.g., sliding window)
  - Asynchron operation
  - Causility, hierarchy of events
  - Correlation
- SQL-like languages
  - E,g.. EPL: Event Processing Language, JAQL, …
  - Queries as basic steps of event processing
- Distributed data sources
  - Databases, online transaction handling systems, monitoring systems, etc.
- Scalability
  - Distributed, cloud-based deplyoment
  - ~100k/sec transcations

CEP application areas

- Business&finance
  - Investments, stock exchange
  - „Treasury”
  - Risk assessment
  - Transport tracking
- „Business Activity Monitoring”
- Online fraud detection/prevention
  - Transaction scanning
  - Online betting (pl. UEFA)
- Operation of large IT systems
  - Detection complex patterns in operation
  - Metrics evaluation
- Security
  - E.g.,. Early detection of dDOS
- http://www.complexevents.com/

Map/Reduce algorithm

- Map
  - Data split
- Reduce
  - Data processing
- Example
  - „Split a text to words, count occurences”
  - Implementation in multiple languages
- Apache implementation
  - Distributed architecture
  - Hadoop (+ Hadoop Distributed File System)
  - Schedule : Job Tracker, Task Tracker



CEP tools

- Esper
- Drools Fusion
- IBM InfoSphereStreams (System S)
- OpenESB - Intelligent Event Processor
- Apache Hadoop + extending projects
- Main features
  - Event processing logic
  - Throughput
  - Requested response time („low latency”)

Case study: CoMiFin project

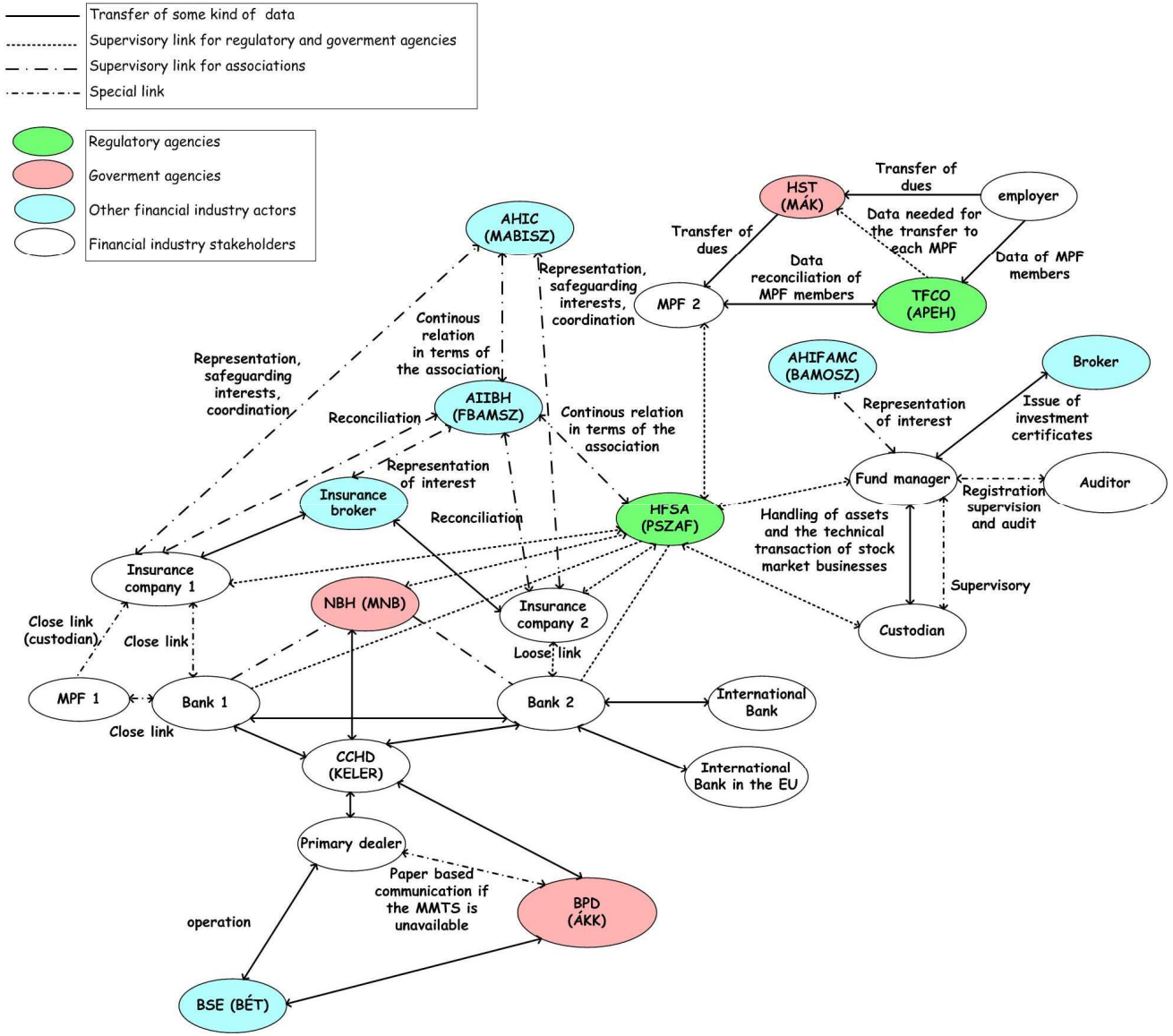
Case study: CoMiFin

- „Communication Middleware for Financial Infrastructures”
- Motivation
  - Banking systems are more and more dependable on IT services
  - Attacks are becoming more sophisticated
  - Critical infrastructures (network, telco, power supply)… interconnected
  - Traditional offline communication is slow (e.g., 8 days to close a bug)
- Aim
  - Scheme to set up and manage a secure environment (software, hardware, monitoring tools, etc.) for information exchange and analysis
- Demonstrator lead by a BME spin-off (OptXware)

Information Federation

- Purpose: detect events (attacks) in a collaborative way
  - „Security: collaboration, not competition”
  - Collaborative network of certified participants
  - Security measures implemented
  - Resilient by its distributed nature
  - Systematic model-based management and monitoring
- Privacy assurance
  - Working on data about transactions
  - No transaction details
  - Anonymization
  - … see ***comifin.eu***

Example: logical interconnections



Attack types (IT)

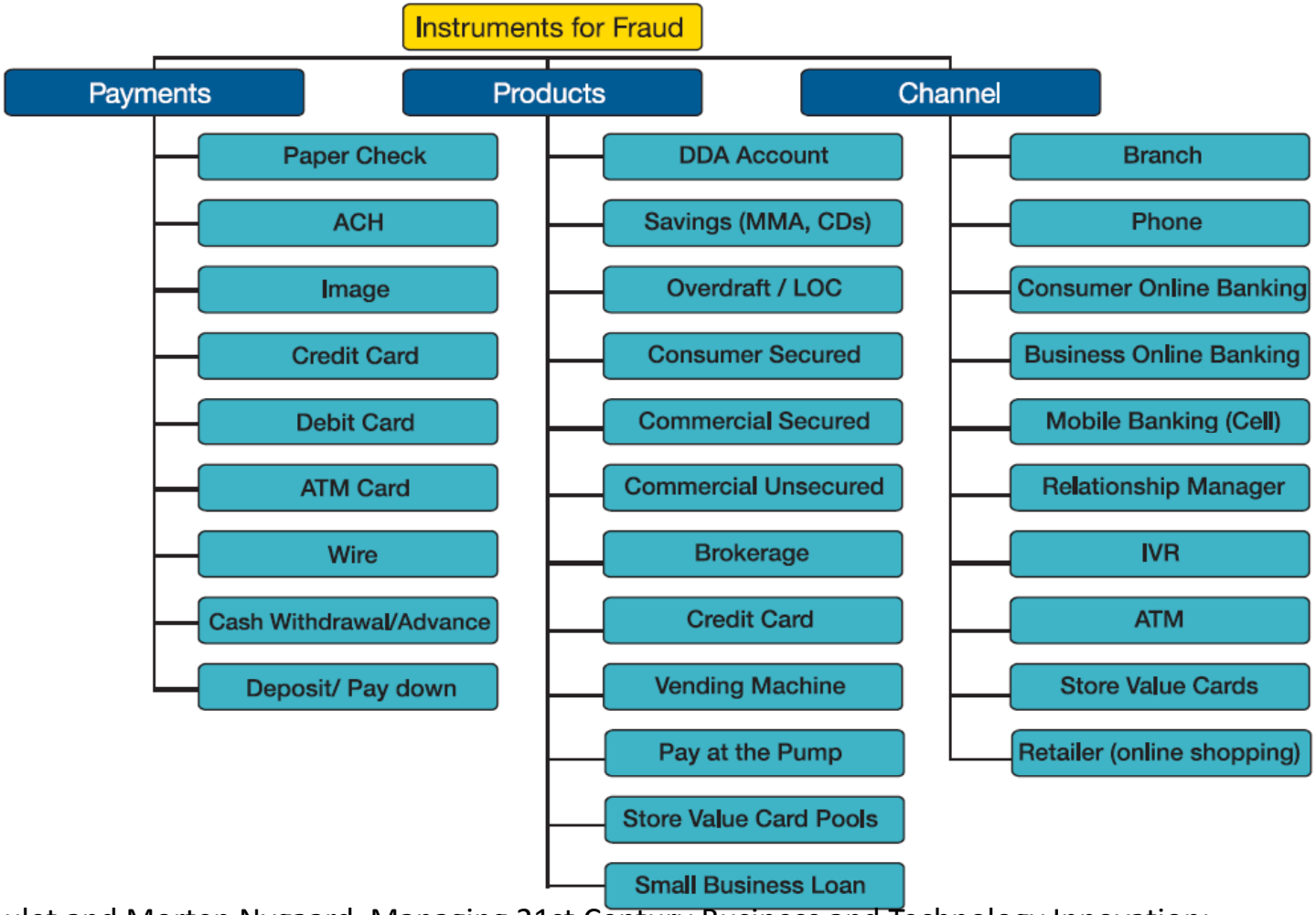
- Distributed Denial of Service (dDOS)
  - E.g., botnets
- Man in the Middle
  - E.g., phishing
- „Identity theft”

Contact Method	CY - 2006		CY - 2007		CY - 2008	
	Complaints	Percentages[1]	Complaints	Percentages[1]	Complaints	Percentages[1]
Internet - E-mail	138,195	45%	152,131	50%	193,817	52%
Mail	50,317	16%	42,330	14%	51,837	14%
Internet - Web Site/Others	46,687	15%	45,447	15%	40,596	11%
Phone	39,365	13%	33,733	11%	26,067	7%
Other	31,722	10%	33,481	11%	57,695	16%
Total Reporting Contact Method	306,286		307,122		370,012	

- … and many combinations

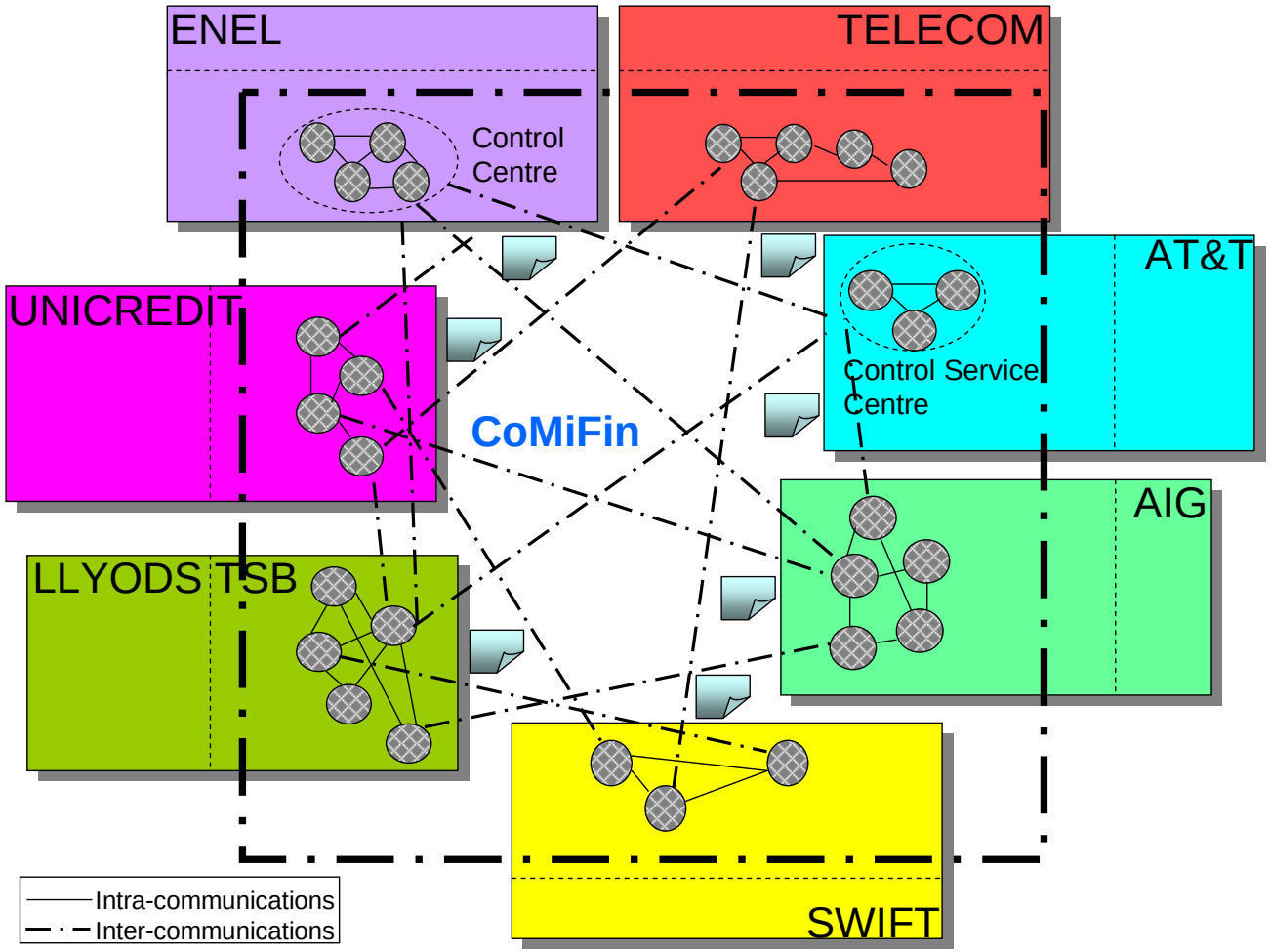
Attack surface

- Banking systems with lot of entry points…



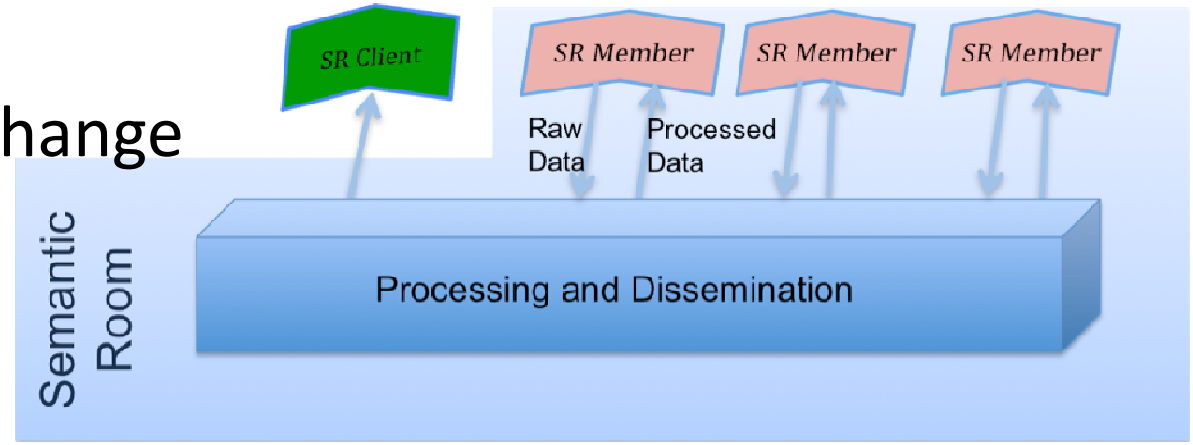
Martin Goulet and Morten Nygaard, Managing 21st Century Business and Technology Innovation: Reduce Operational Risk, Enable Compliance, Protect Privacy with IBM System z, IBM's Global Banking community, ibm.com/banking.

CoMiFin topology

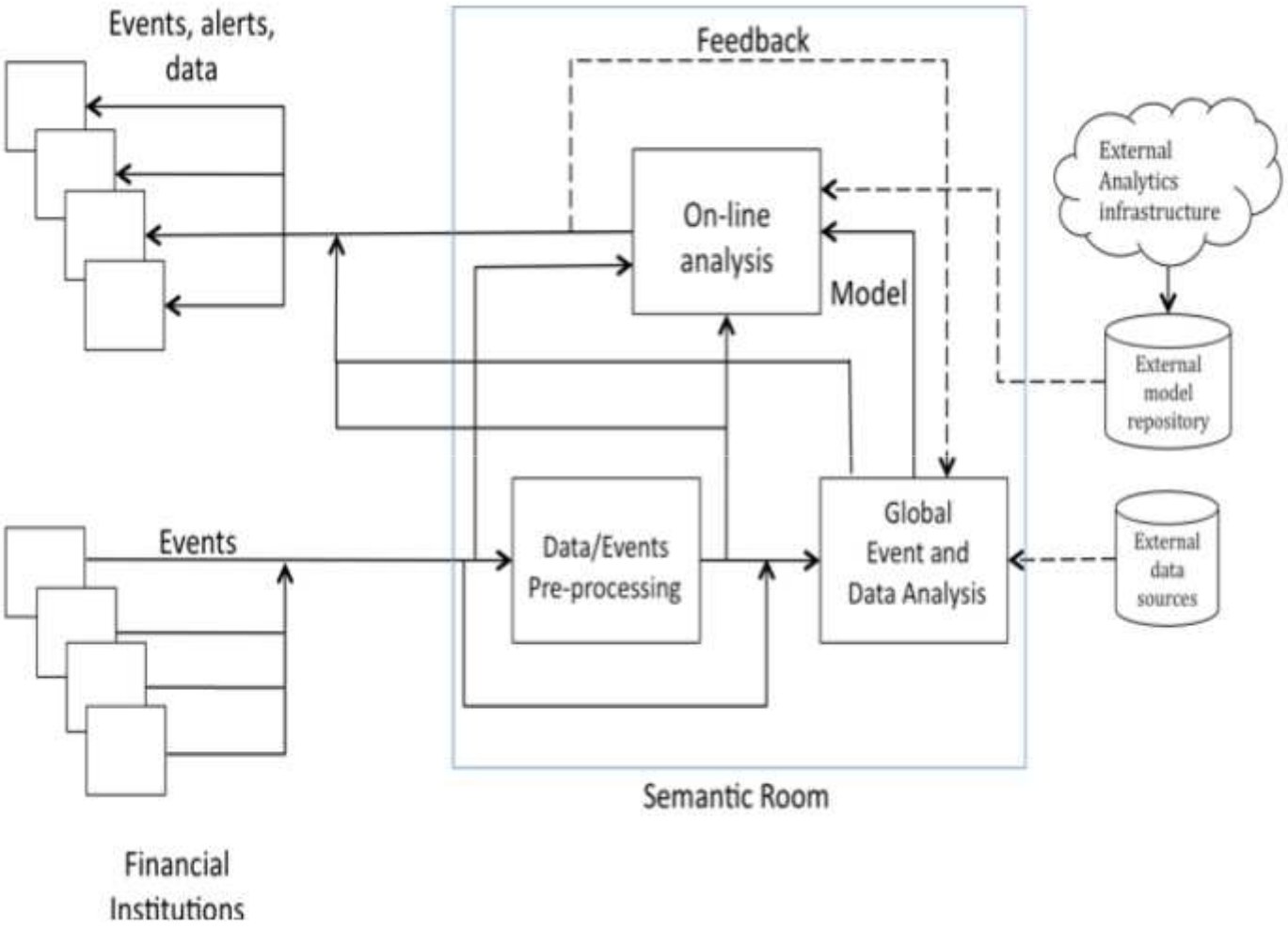


Semantic Room concept

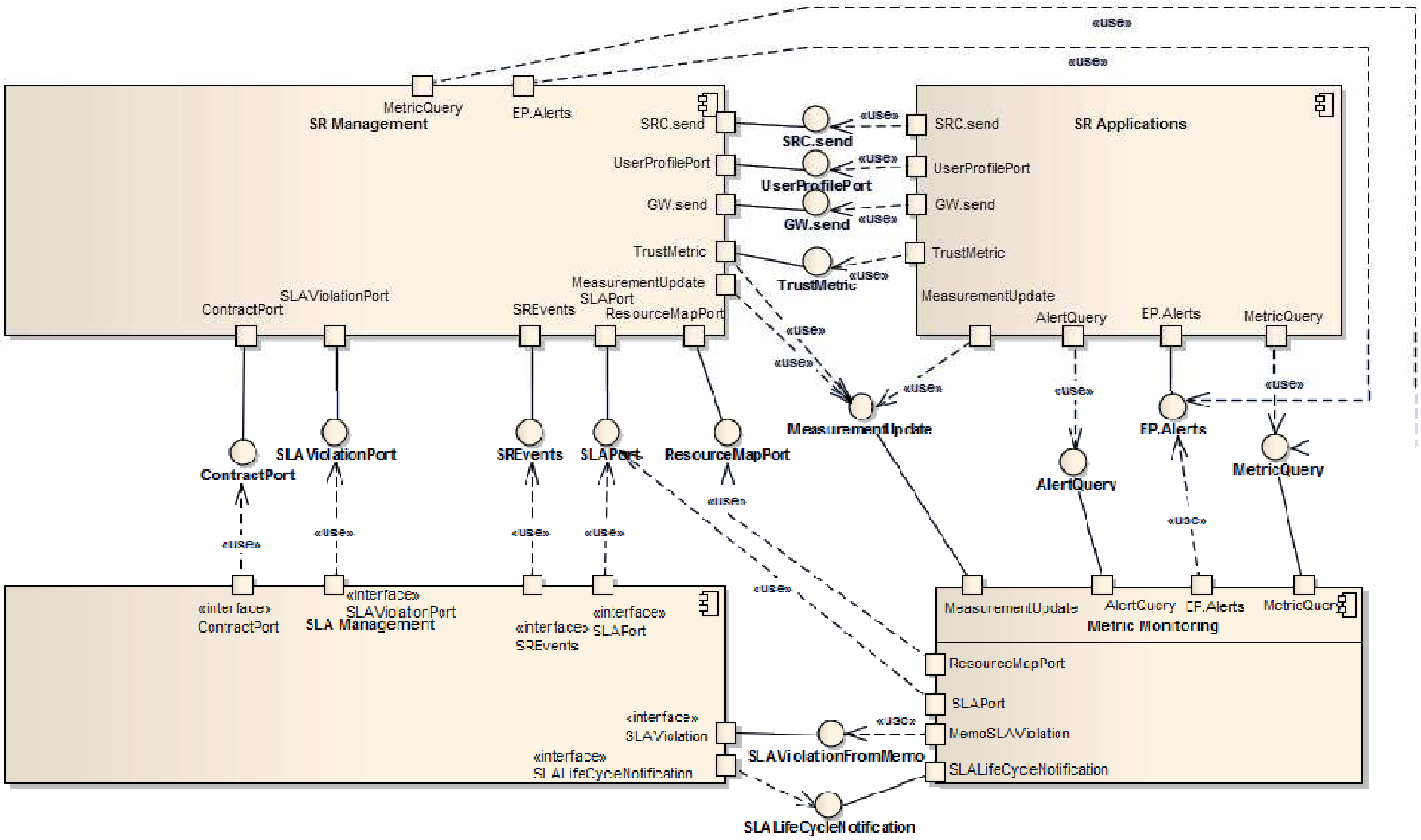
- Basic unit of information federation
- Logical separation of entities
- SR features : a „SaaS” offer of the underlying CoMiFin infrastructure
  - Membership management
  - Resource allocation
  - Monitoring
  - Information exchange



Online data analysis (CEP)



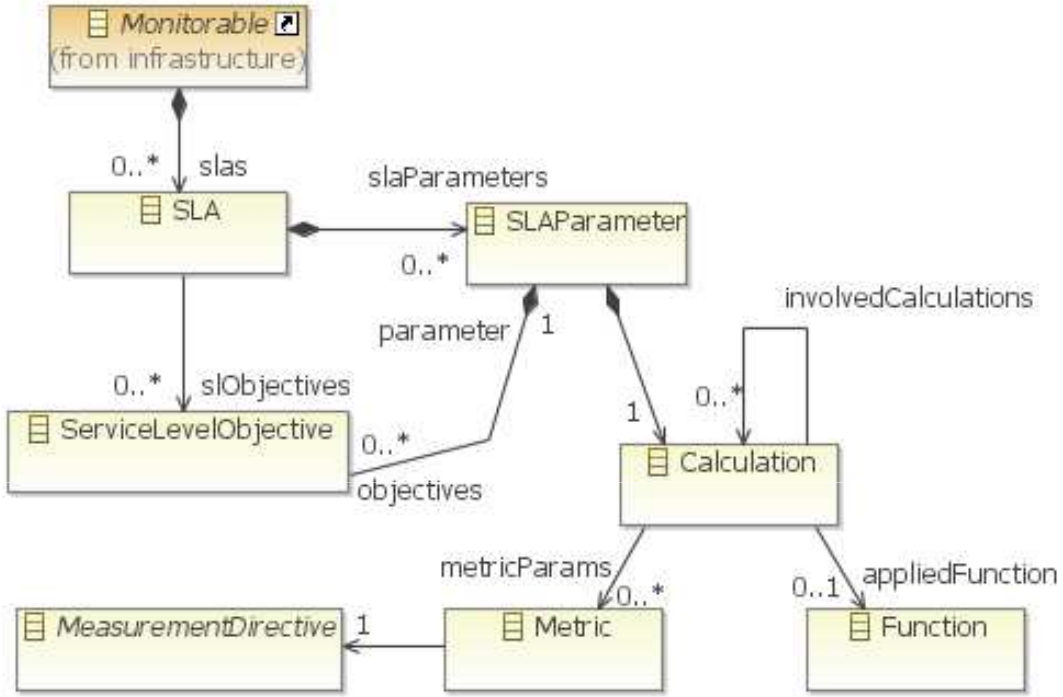
Logical components



Model-driven system monitoring

Levels for metrics

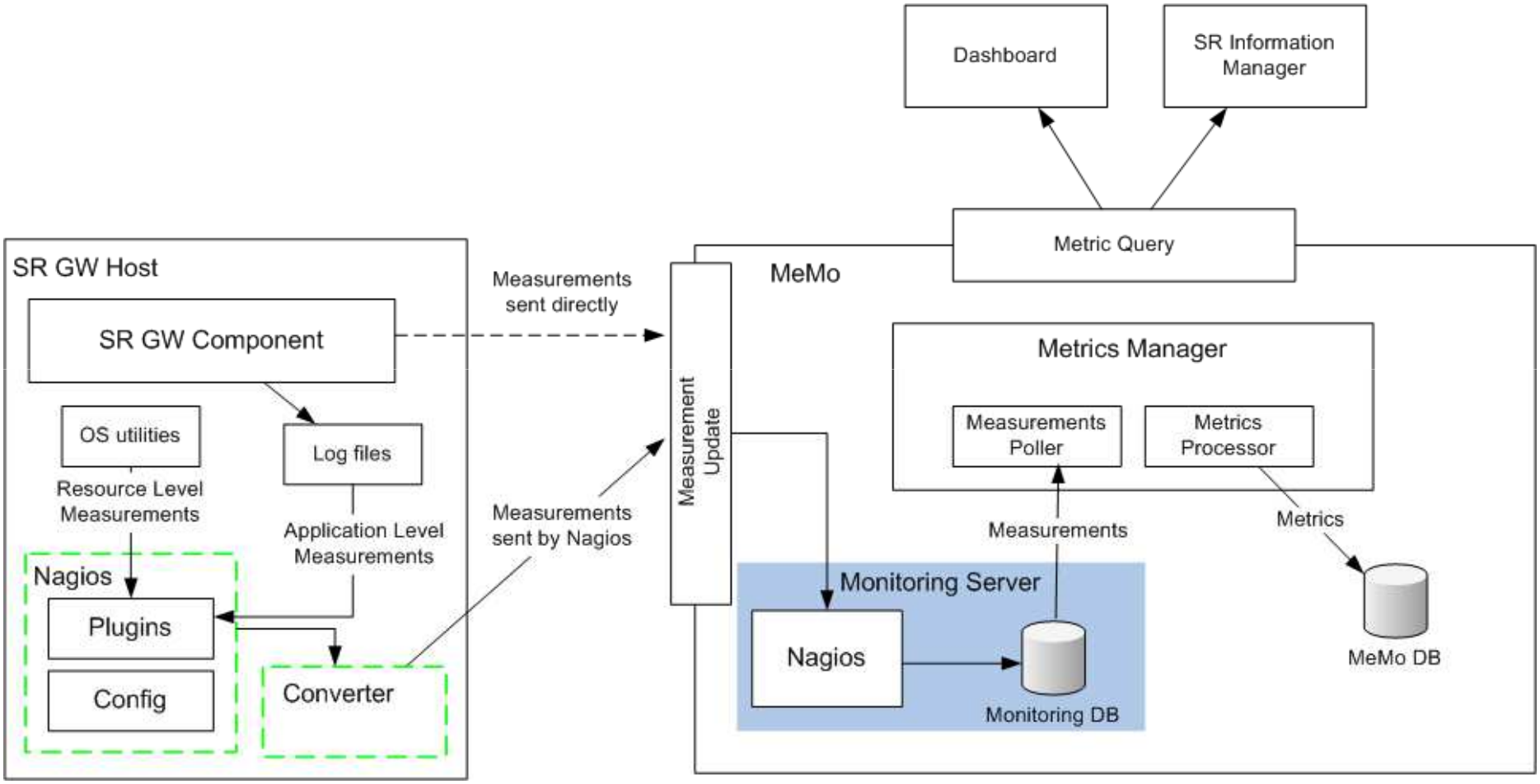
- Resource level
  - Network, Disk
  - Memory, CPU…
  - SR Administrator View
- Application level
  - Event processing
  - Evaluation of alerts
  - SR Administrator,
  - SR Manager View
  - FI specific parts. SR Member
- “Business level”
  - What CoMiFin produces?
  - Alerts are also displayed
  - SR Manager, SR Member



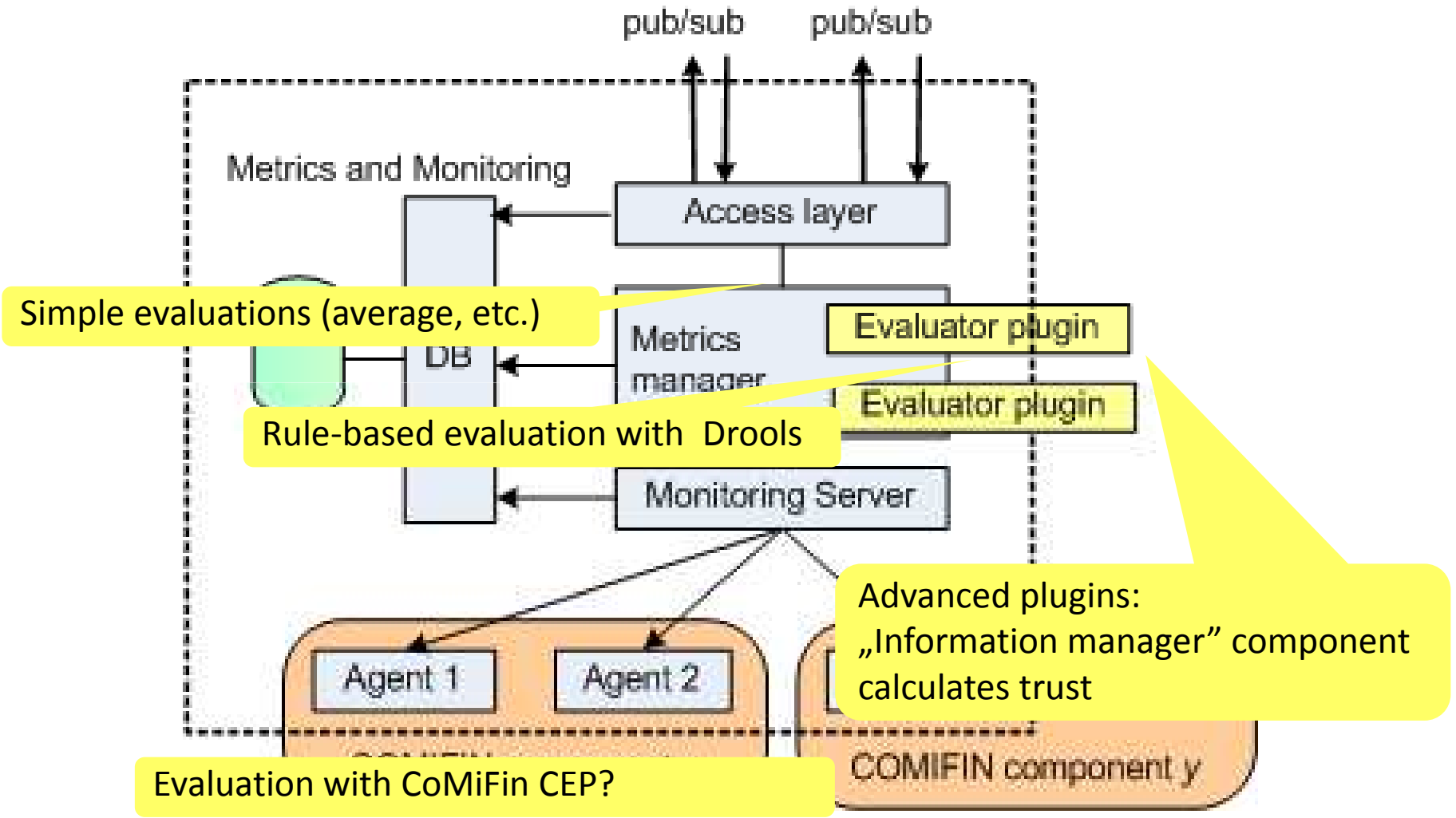
Example: Metric definition

Name	Performance capability of web server	
Definition	The number of threads that can receive and process HTTP requests.	
	States	Normal Operational State: The number of threads matches the estimated capacity of the underlying resources.
		Overloaded Operational State: The number of threads are higher than the maximal estimated for the underlying resources.
		Unusable Operational State: The number of processing threads cannot be determined.
	Measurement Unit	piece
	Range	non-negative integer
	Type	Low level
Measuring	The web server should be instrumented with an SNMP agent exposing this attribute.	
	Frequency	periodical, every 3 minutes
Evaluation	Method	classification (see states above)
	Time range	evaluation is instantaneous

Architectural Example: Monitoring of the Gateway



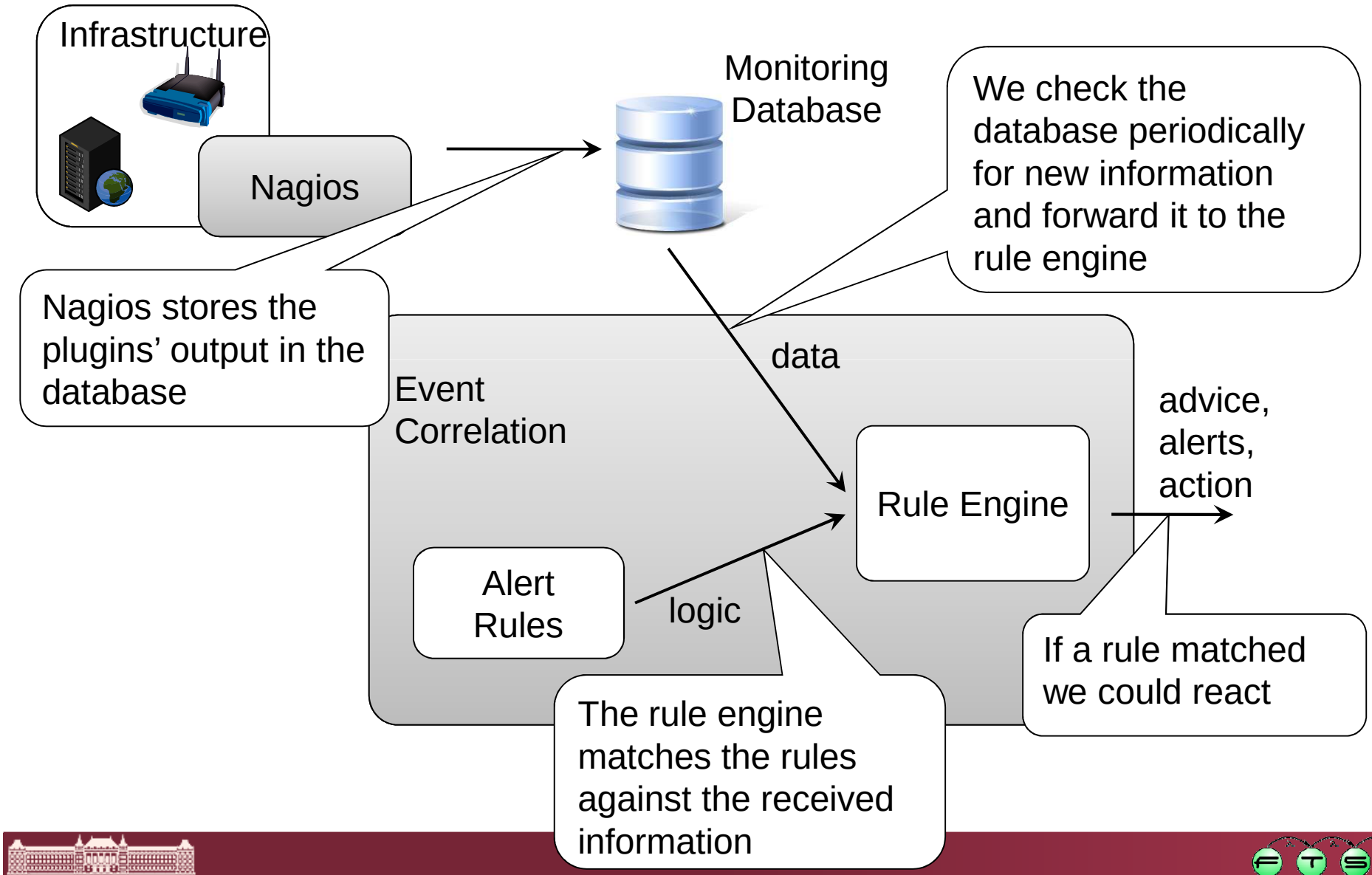
Advanced Evaluation



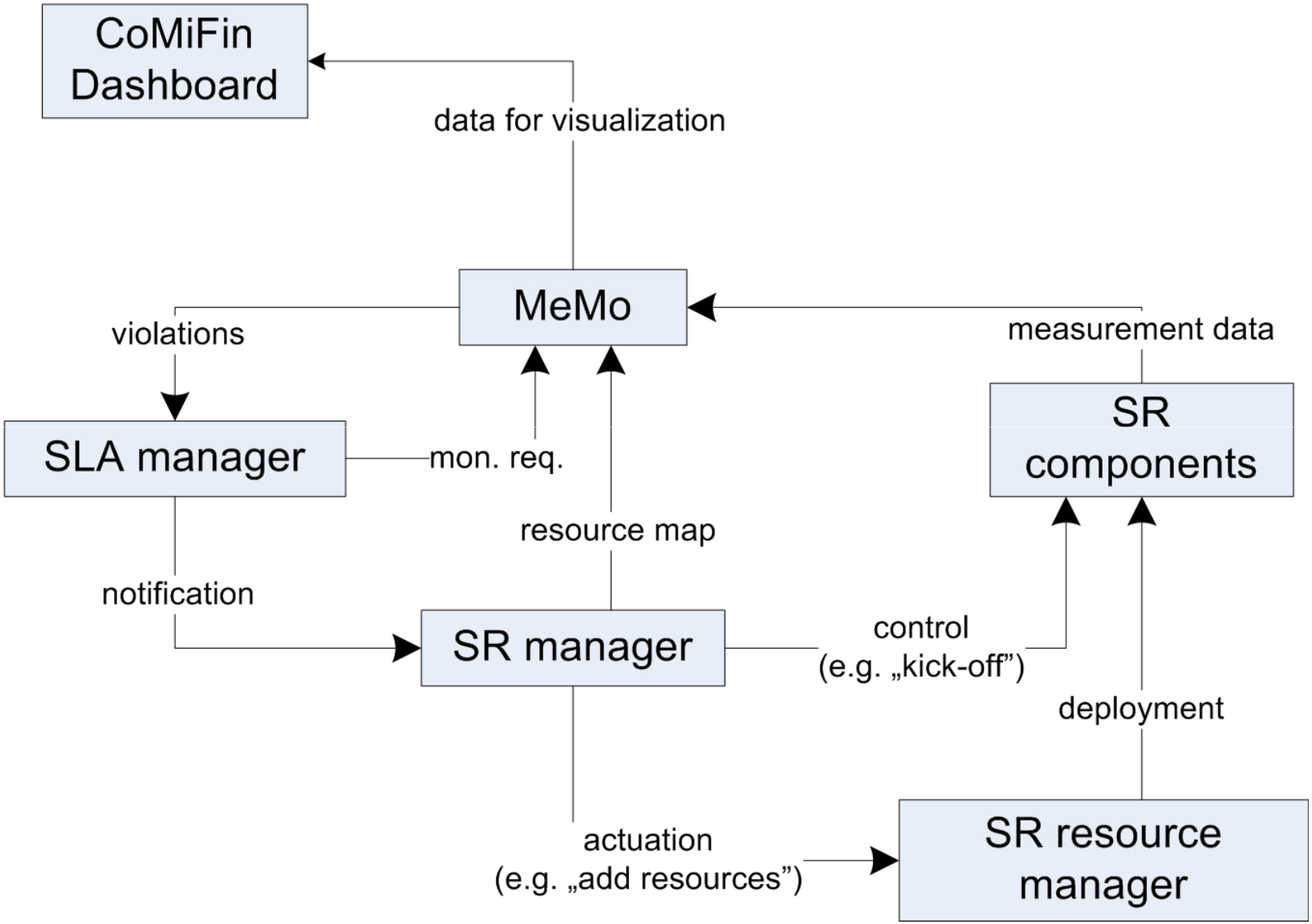
Evaluation of measurements

- Advanced „Evaluator plugin” for monitoring
- Nagios integrated with Drools
  - Nagios: monitoring
  - Drools: correlation of events sent by Nagios
- Motivation
  - **CoMiFin is itself a critical infrastructure**
    - advanced monitoring of resource should be ensured
    - e.g. COBIT/SOX/… directives
  - detecting meaningful patterns within CoMiFin could be beneficial

Overview of advanced evaluation



CoMiFin Control Loop



Example metric

Name	Performance capability of web server	
Definition	The number of threads that can receive and process HTTP requests.	
	States	Normal Operational State: The number of threads matches the estimated capacity of the underlying resources.
		Overloaded Operational State: The number of threads are higher than the maximal estimated for the underlying resources.
		Unusable Operational State: The number of processing threads cannot be determined.
	Measurement Unit	piece
	Range	non-negative integer
	Type	Low level
Measuring	The web server should be instrumented with an SNMP agent exposing this attribute.	
	Frequency	periodical, every 3 minutes
Evaluation	Method	classification (see states above)
	Time range	evaluation is instantaneous

Rule-based event detection

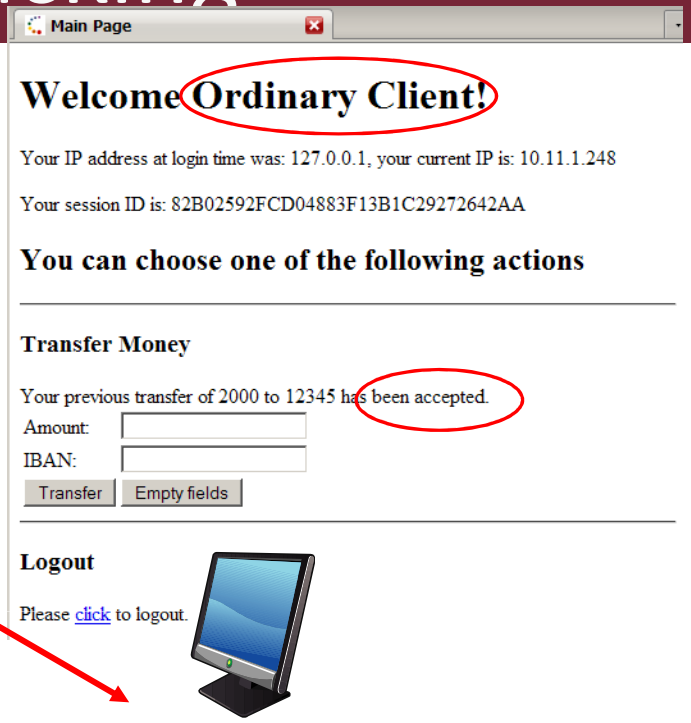
Session Hijack detection

- Session hijack: modify the sesison by observing/modifying user communication
- Configuration: sample banking app
  - User management
  - Transaction management
- Monitoring of application level information
  - Session ID
  - Client data
- „Fault injection”
  - Simulated attacks

Session Hijacking

Application Server

Ordinary Client



Bad Guy

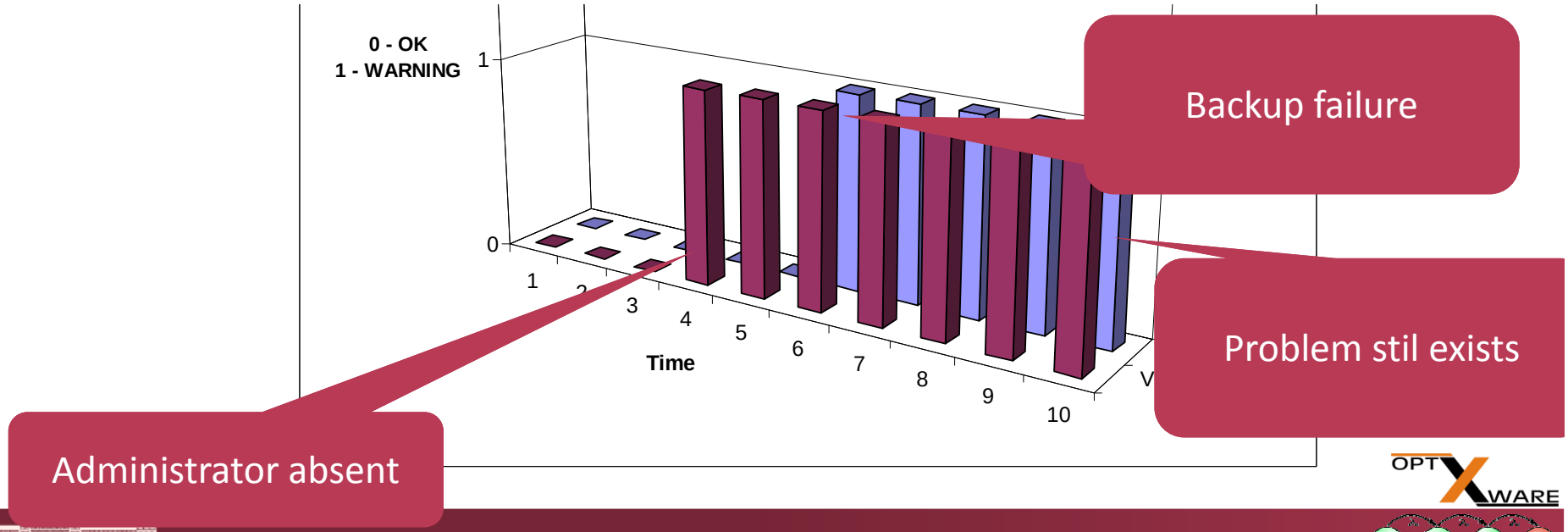
Session hijack detection

- Checking IP and session ID with Drools

```
WARNING: possible session hijack:
{
 currentAddress=127.0.0.1,
 remoteAddress=127.0.0.1,
 sessionId=21...B6
}
and the possible attacker with the same session:
{
 currentAddress=10.11.1.154,
 remoteAddress=127.0.0.1,
 sessionId=21...B6
}
```

Demo 2: Missing backup problem

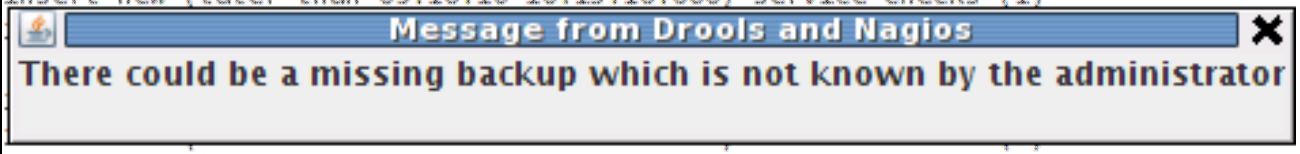
- Information from multiple aplications / event sources
- Related standards
  - e.g. COBIT PO-4 “Define the IT Processes, Organization and Relationships”
  - COBIT PO-9 “Manage IT Human Resources”



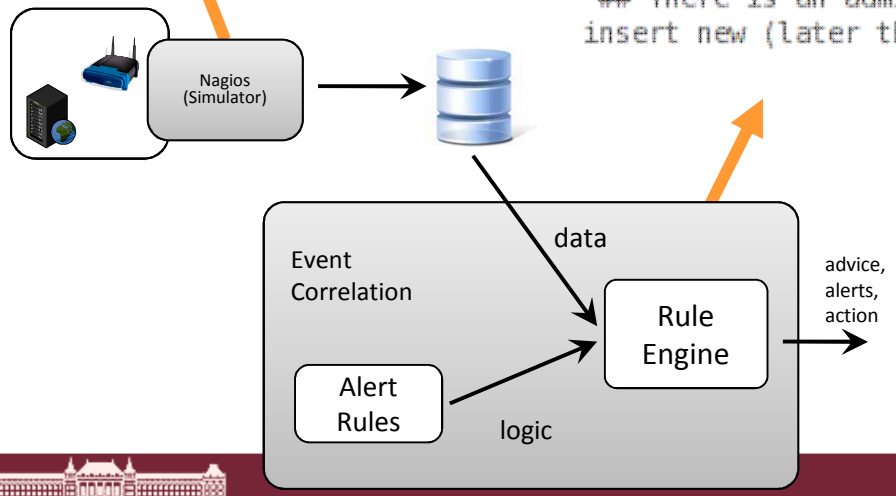
Drools based detection

```
[root@p2 NagiosDrools5]# sh runSimulator.sh "2009-10-10 10:10:10" 0.01
Press enter to start simulation

Press enter to stop simulation
Next sleep time is 1800
Next sleep time is 1200
Next sleep time is 1800
Next sleep time is 1200
Next sleep time is 1800
Next sleep time is 1200
Next sleep time is 1800
Next sleep time is 1200
Next sleep time is 1800
Next sleep time is 1200
Next sleep time is 1800
Next sleep time is 3000
Insertion of events finished

[root@p2 NagiosDrools5]#
```

```
INFO: Feeder has been started
Press enter to exit
insert new (later than 09.10.10 10:10:10.000) service checks (1)
insert new (later than 09.10.10 10:13:10.000) service checks (1)
 ## There is an administrator missing ServiceCheck(3): 10:15:10.000 (1)
insert new (later than 09.10.10 10:15:10.000) service checks (1)
```




```
 ## There is an administrator missing ServiceCheck(7): 10:25:10.000 (1)
insert new (later than 09.10.10 10:25:10.000) service checks (1)
 ## There could be a missing backup which is not known by the administrator
 ## Missing backup: ServiceCheck(8): 10:28:10.000 (1)
insert new (later than 09.10.10 10:28:10.000) service checks (1)
 ## There is an administrator missing ServiceCheck(9): 10:30:10.000 (1)
insert new (later than 09.10.10 10:30:10.000) service checks (1)
```

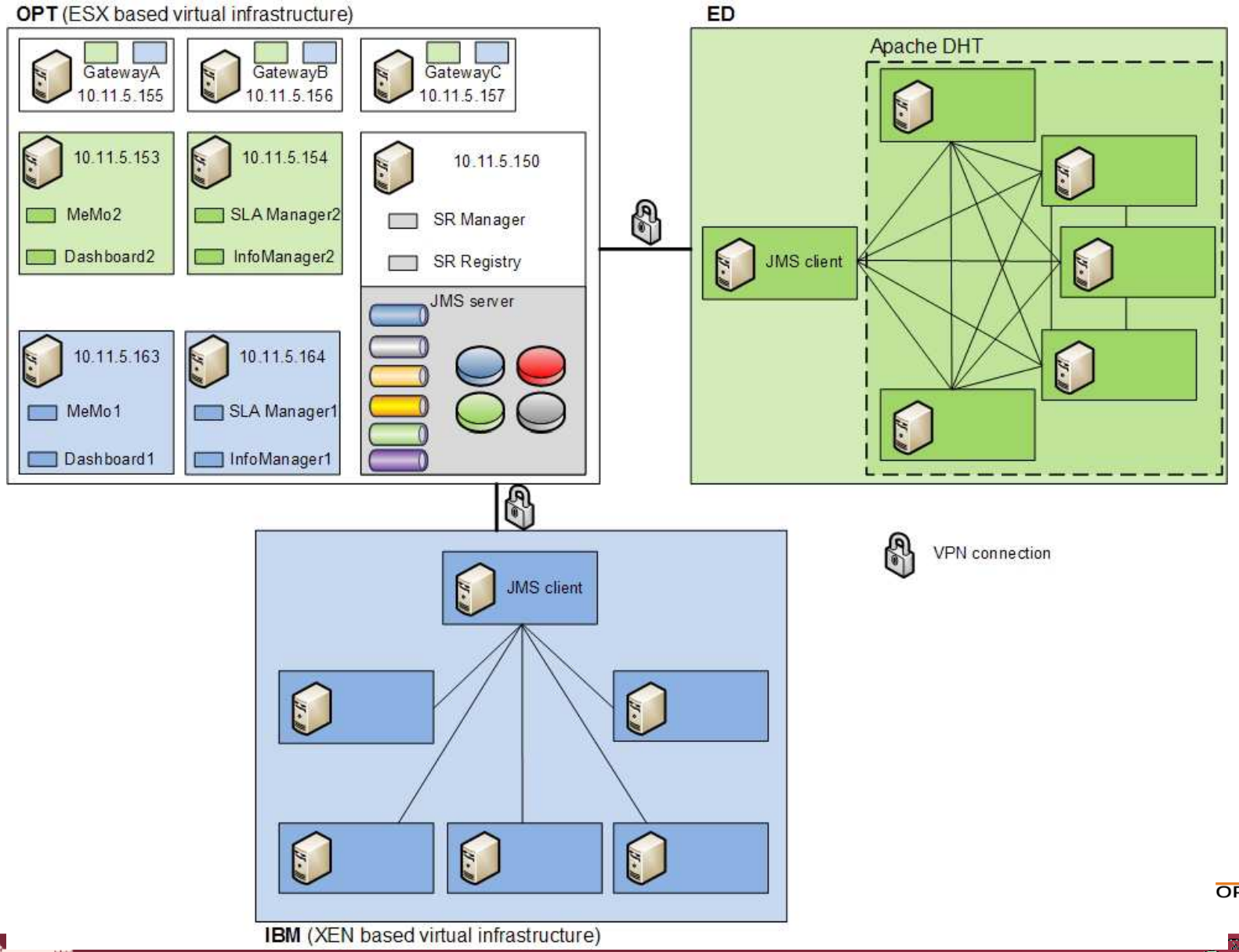


Connectivity of Semantic Rooms

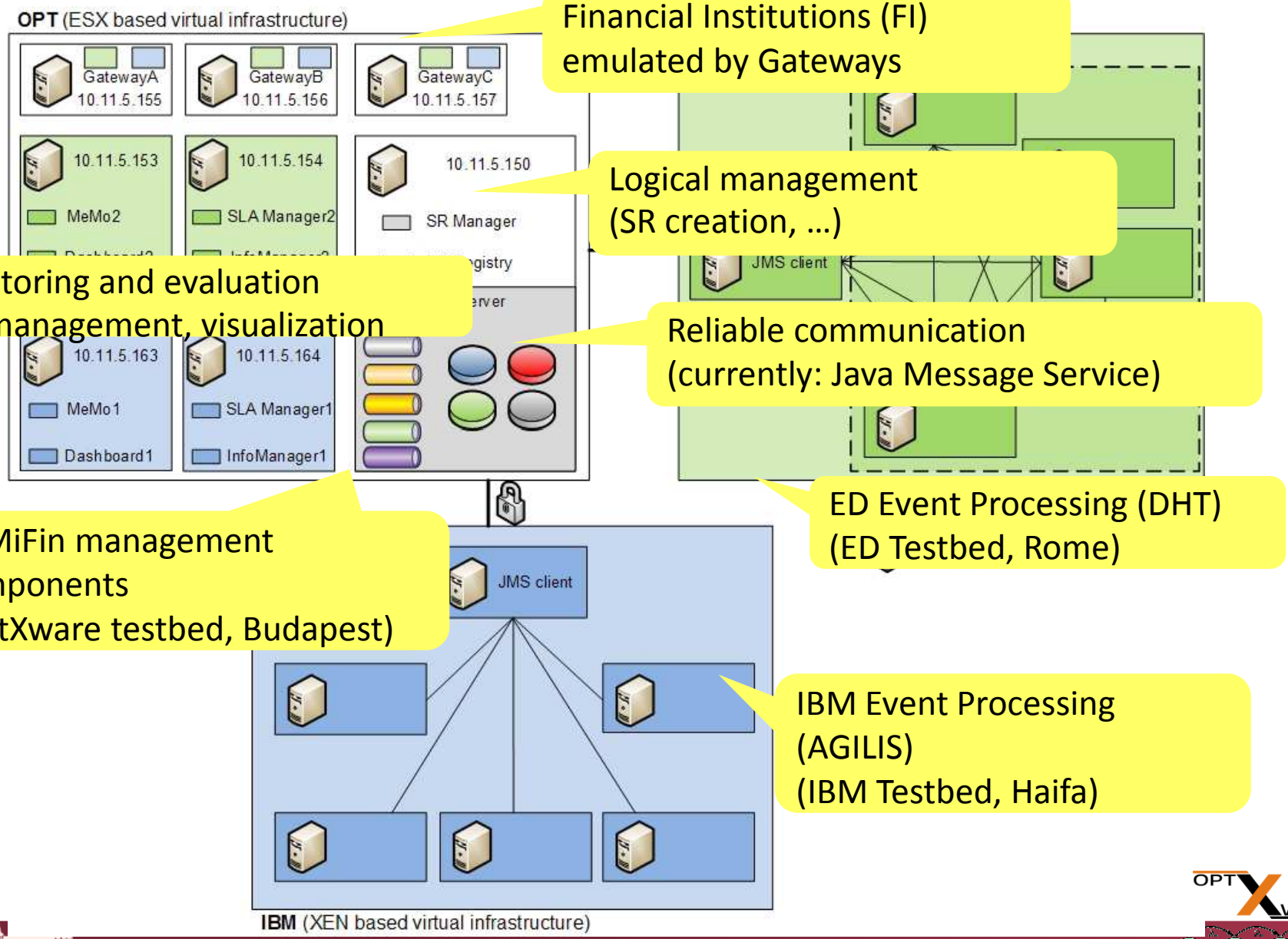
- Vision: CoMiFin as a platform (D6.5)
  - “Framework for information sharing among FIs”
  - “A trusted third party platform”
- Connectivity in CoMiFin
  - Allows for propagating alerts among SRs
  - Improves the quality of Event Processing with additional input
  - Improves the capability of CoMiFin to detect attacks
  - Facilitates the creation of new services built on a combination of SRs

  → improves the ability of FIs to prevent attacks

Architecture



Architecture



Financial Institutions (FI)
emulated by Gateways

Logical management
(SR creation, …)

Monitoring and evaluation
SLA management, visualization

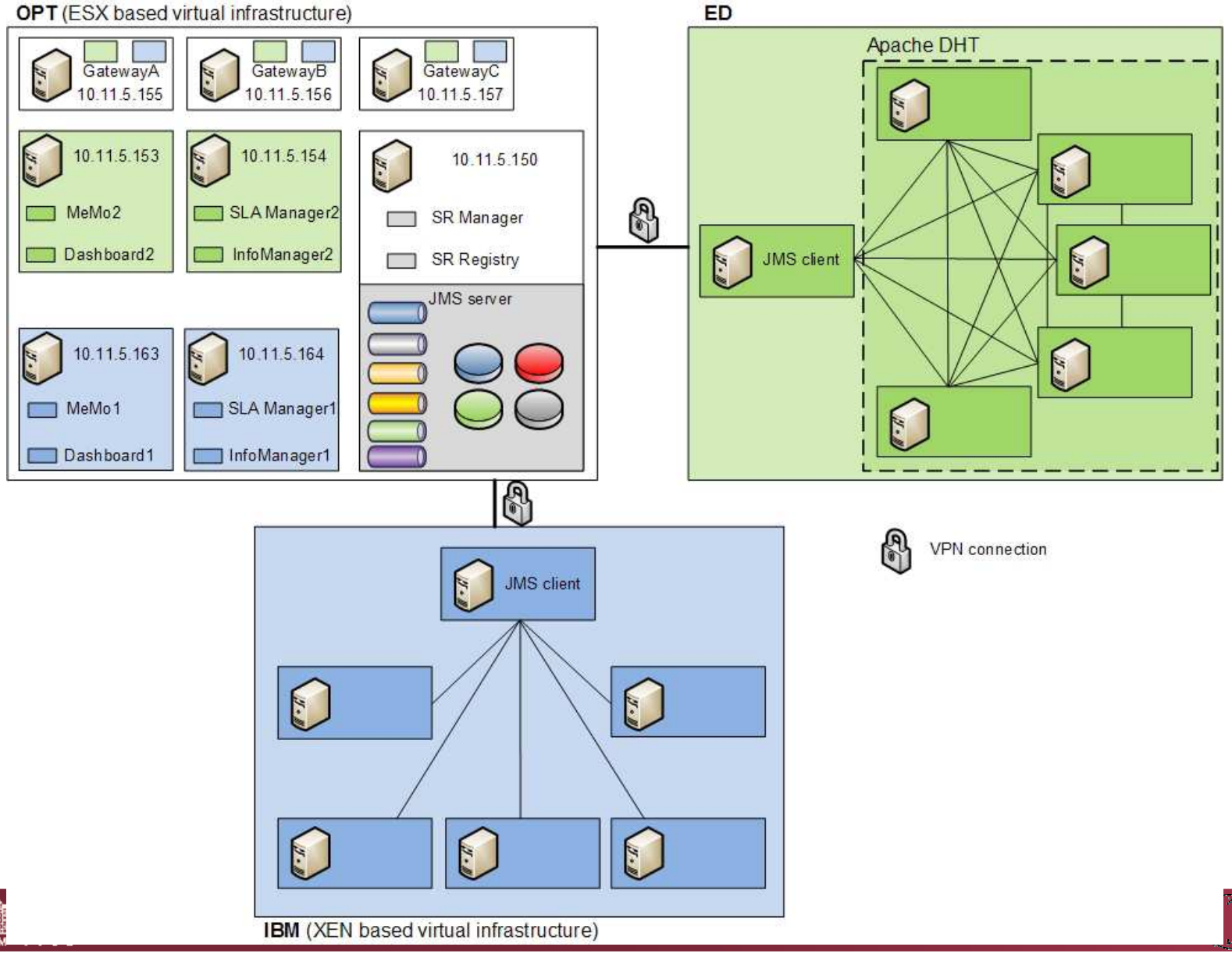
Reliable communication
(currently: Java Message Service)

ED Event Processing (DHT)
(ED Testbed, Rome)

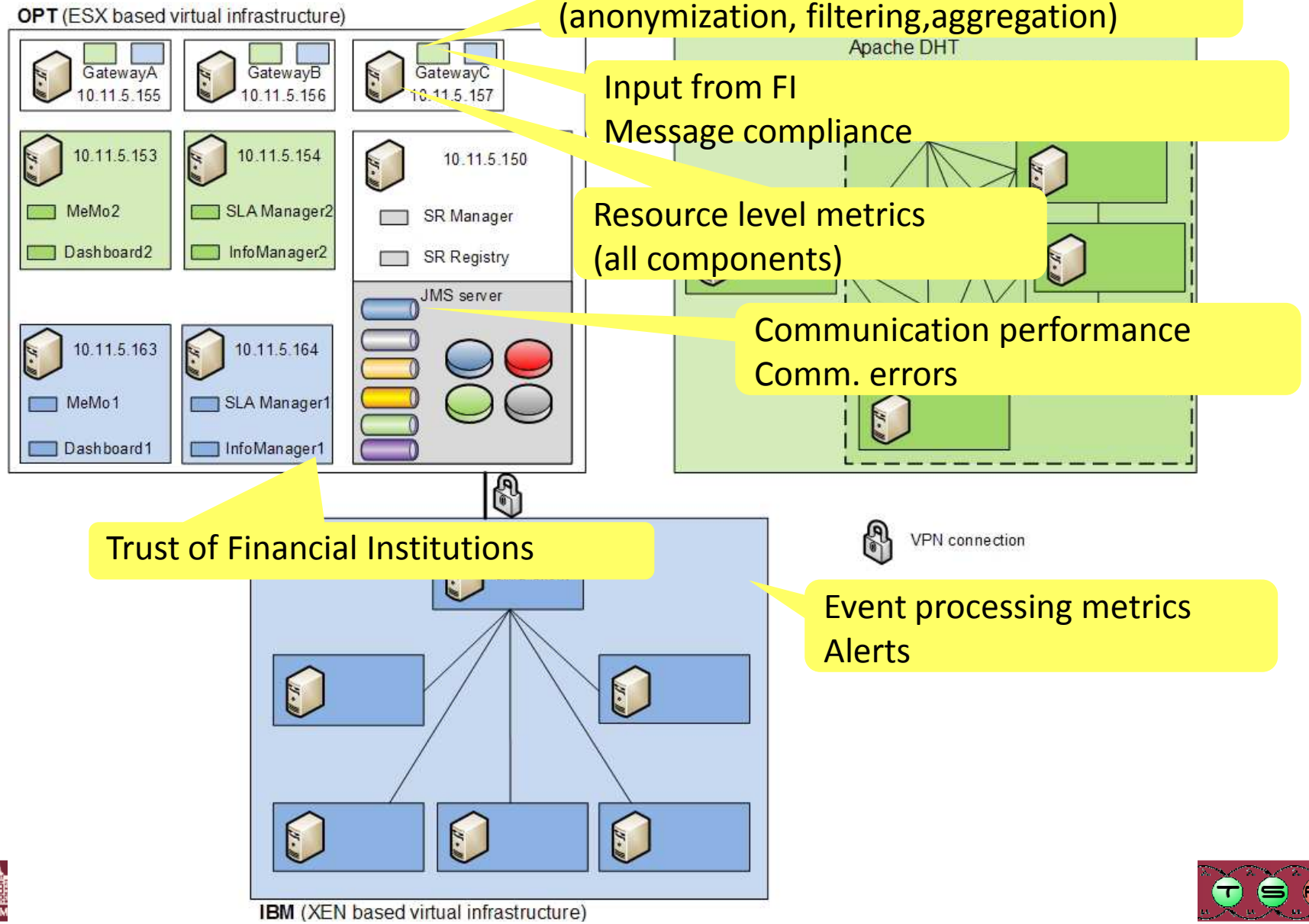
CoMiFin management
components
(OptXware testbed, Budapest)

IBM Event Processing
(AGILIS)
(IBM Testbed, Haifa)

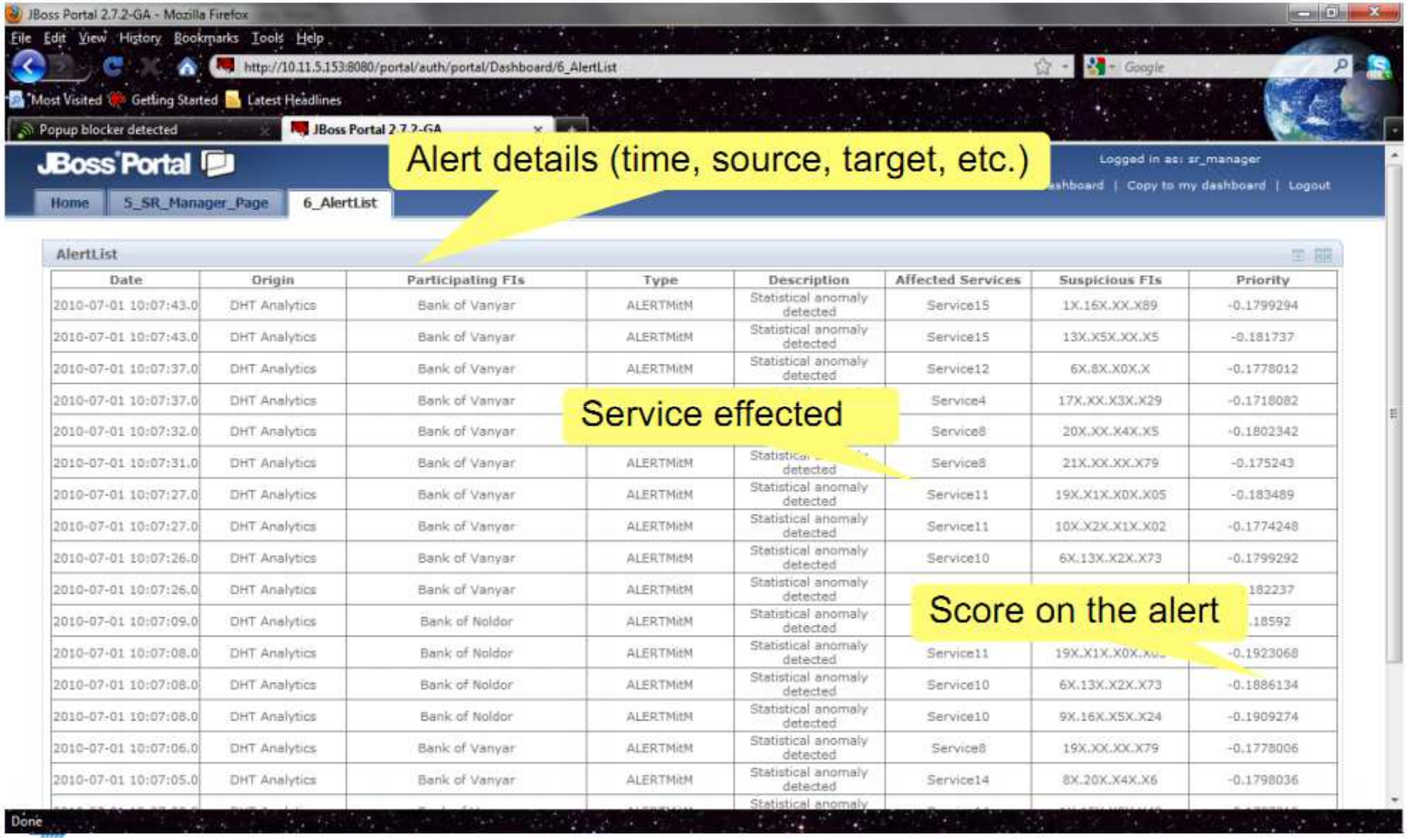
What to Measure in CoMiFin?



What to Measure in CoMiFin?



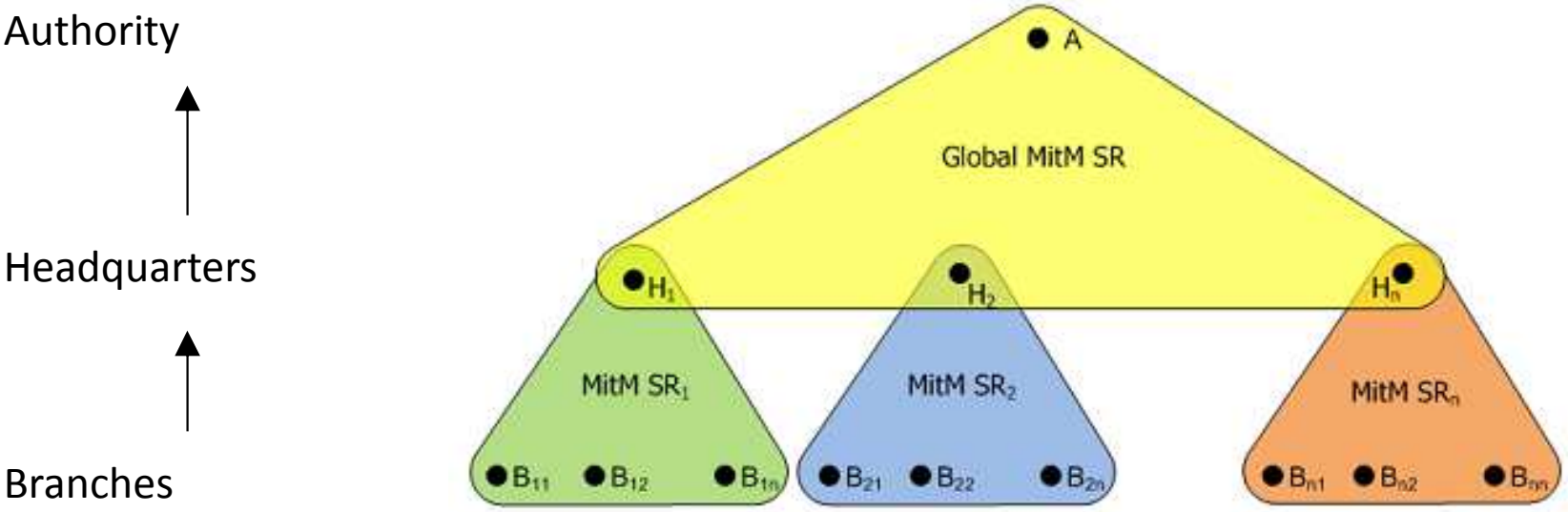
Sample results



Date	Origin	Participating FIs	Type	Description	Affected Services	Suspicious FIs	Priority
2010-07-01 10:07:43.0	DHT Analytics	Bank of Vanyar	ALERTMitM	Statistical anomaly detected	Service15	1X.16X.XX.X89	-0.1799294
2010-07-01 10:07:43.0	DHT Analytics	Bank of Vanyar	ALERTMitM	Statistical anomaly detected	Service15	13X.X5X.XX.X5	-0.181737
2010-07-01 10:07:37.0	DHT Analytics	Bank of Vanyar	ALERTMitM	Statistical anomaly detected	Service12	6X.8X.X0X.X	-0.1778012
2010-07-01 10:07:37.0	DHT Analytics	Bank of Vanyar	ALERTMitM	Statistical anomaly detected	Service4	17X.XX.X3X.X29	-0.1718082
2010-07-01 10:07:32.0	DHT Analytics	Bank of Vanyar	ALERTMitM	Statistical anomaly detected	Service8	20X.XX.X4X.X5	-0.1802342
2010-07-01 10:07:31.0	DHT Analytics	Bank of Vanyar	ALERTMitM	Statistical anomaly detected	Service8	21X.XX.XX.X79	-0.175243
2010-07-01 10:07:27.0	DHT Analytics	Bank of Vanyar	ALERTMitM	Statistical anomaly detected	Service11	19X.X1X.X0X.X05	-0.183489
2010-07-01 10:07:27.0	DHT Analytics	Bank of Vanyar	ALERTMitM	Statistical anomaly detected	Service11	10X.X2X.X1X.X02	-0.1774248
2010-07-01 10:07:26.0	DHT Analytics	Bank of Vanyar	ALERTMitM	Statistical anomaly detected	Service10	6X.13X.X2X.X73	-0.1799292
2010-07-01 10:07:26.0	DHT Analytics	Bank of Vanyar	ALERTMitM	Statistical anomaly detected			182237
2010-07-01 10:07:09.0	DHT Analytics	Bank of Noldor	ALERTMitM	Statistical anomaly detected			.18592
2010-07-01 10:07:08.0	DHT Analytics	Bank of Noldor	ALERTMitM	Statistical anomaly detected	Service11	19X.X1X.X0X.X...	-0.1923068
2010-07-01 10:07:08.0	DHT Analytics	Bank of Noldor	ALERTMitM	Statistical anomaly detected	Service10	6X.13X.X2X.X73	-0.1886134
2010-07-01 10:07:08.0	DHT Analytics	Bank of Noldor	ALERTMitM	Statistical anomaly detected	Service10	9X.16X.X5X.X24	-0.1909274
2010-07-01 10:07:06.0	DHT Analytics	Bank of Vanyar	ALERTMitM	Statistical anomaly detected	Service8	19X.XX.XX.X79	-0.1778006
2010-07-01 10:07:05.0	DHT Analytics	Bank of Vanyar	ALERTMitM	Statistical anomaly detected	Service14	8X.20X.X4X.X6	-0.1798036

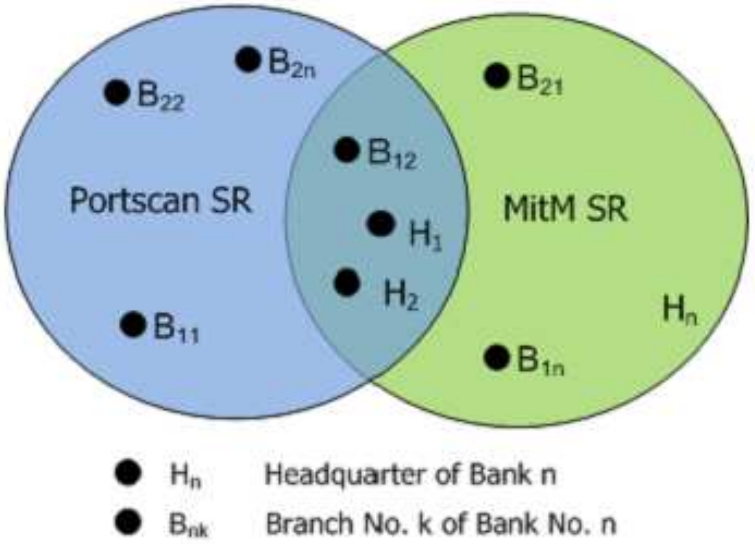
SR Composition (1)

- **Vertical, hierarchical composition**
  - The goal of the collaborating SRs is identical
  - The “topmost” SR has an aggregated global view
  - The underlying SRs have detailed local view
  - Example: Banks and the Financial Supervisory Authority



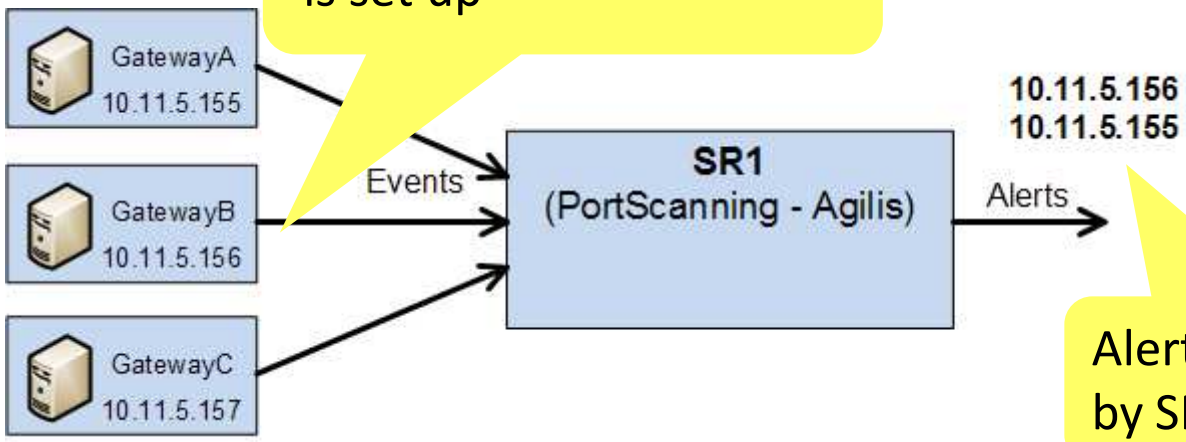
SR Composition (2)

- **Horizontal, sequential composition**
  - Communicating SRs have different goals
  - Complex/combined attacks can be detected by information fusion
  - E.g. , alerts generated in Portscan SR contributes to the Blacklist managed in MitM SR

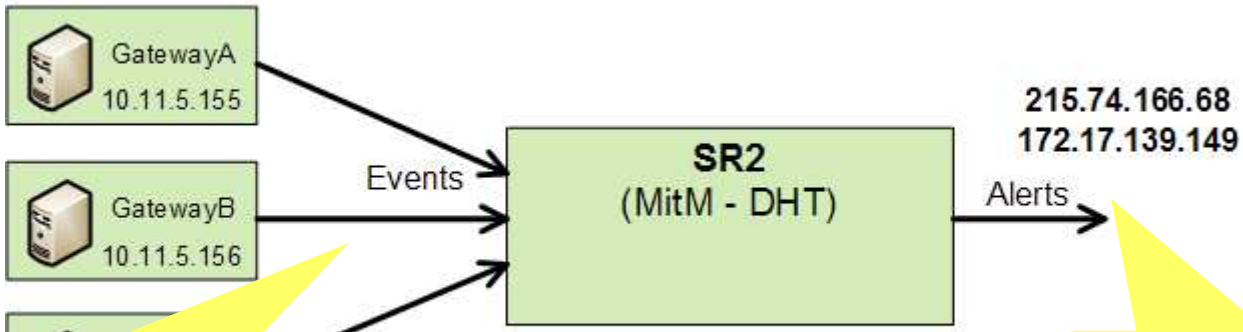


Demo – Scenario

SR1 (Port Scanning)
is set up



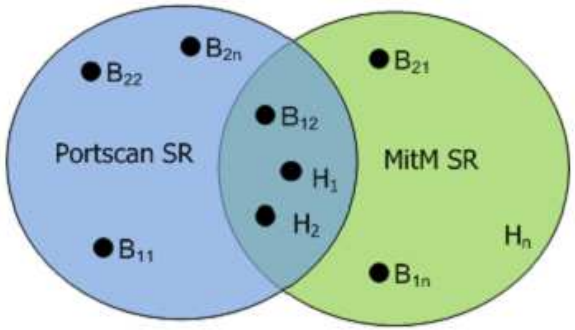
Alerts generated
by SR1



SR2 (MitM)
is set up

Alerts generated
by SR2 on its own

Demo - Scenario



SR1 (Port Scanning) is set up

GatewayA 10.11.5.155

GatewayB 10.11.5.156

GatewayC

Events

SR1 (PortScanning - Agilis)

SRC1

Alerts

10.11.5.155
10.11.5.156
10.11.5.157

SR Connectivity between the two SRs is activated

SRC notifies SR2 about suspicious IPs

10.11.5.155
10.11.5.156
10.11.5.157

GatewayA 10.11.5.155

GatewayB

Events

SRC2

SR2 (MitM - DHT)

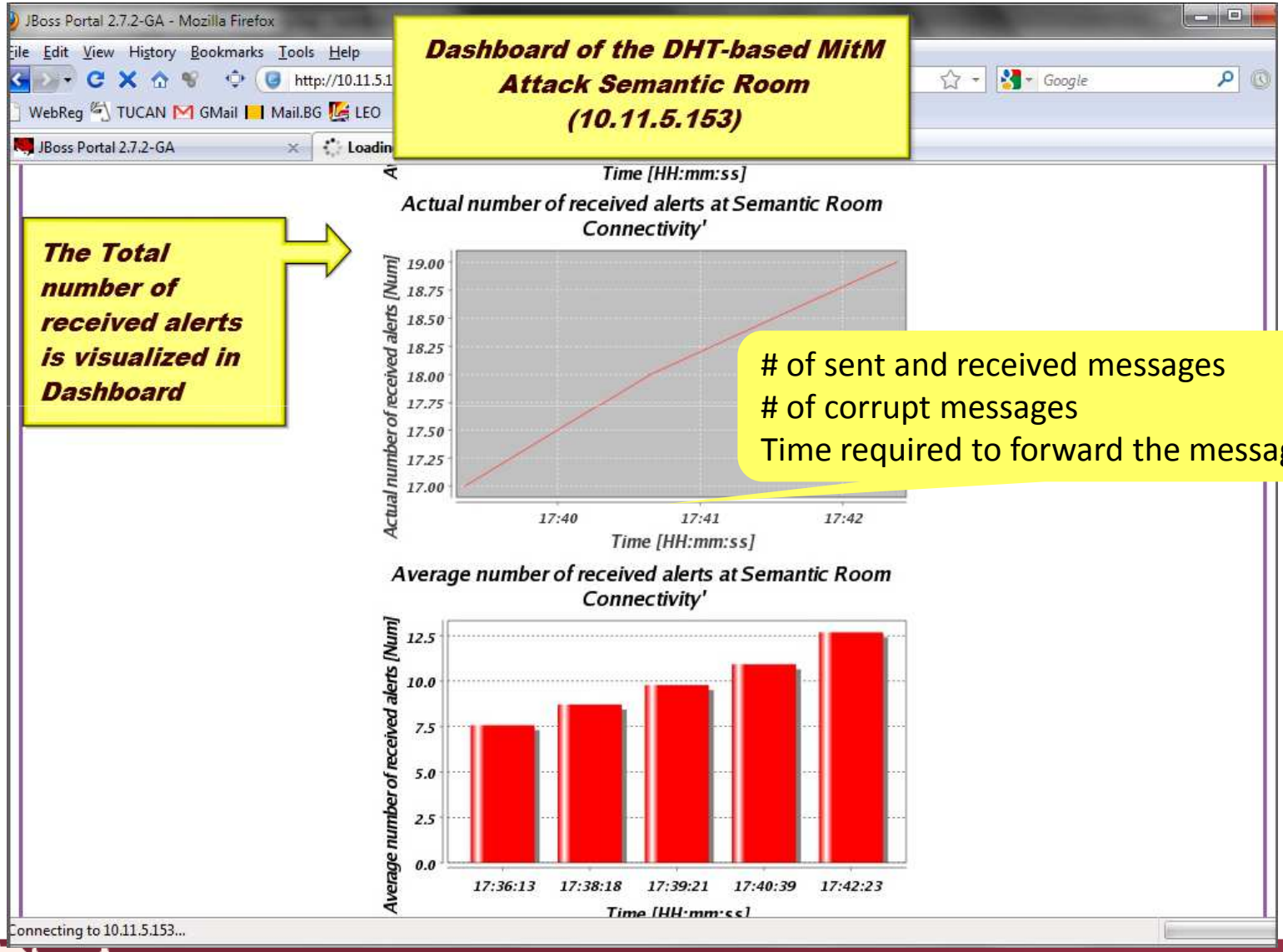
Alerts

215.74.166.68
172.17.139.149
10.11.5.155
10.11.5.156

SR2 (MitM) is set up

Attacks from suspicious IPs also considered

SR Connectivity Metrics



# of sent and received messages
# of corrupt messages
Time required to forward the messages

Conclusions

- Prototype applications
  - Italy, Norway
- Still a lot to do to be „autonomous”
  - Self-* properties
  - Online modifications (pattern, thresholds…)