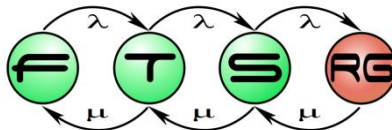


Komplex eseményfeldolgozás (CEP)

Gönczy László

gonczy@mit.bme.hu

Bergmann Gábor, Dávid István és az OptXware Kft. anyagainak felhasználásával



Tartalom

- Szabályalapú megközelítés felhasználása: komplex eseményfeldolgozó rendszerek
- CEP felhasználása
- Esettanulmány: CoMiFin
- Eseményfeldolgozás modell alapú tervezése

Kihívások

- Sok információforrás
 - „Szenzorok”
 - Felhasználói lépések szekvenciái
 - Logok
 - Külső szolgáltatások
- Sok esemény
 - Pl. ~százás nagyságrendű servermetrika, százás nagyságrendű server
- Sok „érdektelen” esemény közt néhány minta
- Párhuzamos, online adatfeldolgozás szükséges
 - Hagyományos adatbázis alapú módszerek lassúak lehetnek
 - Egyszerre nem fér el minden esemény egy feldolgozóegység memóriájában
- Feladat: események feldolgozása és korrelációja
 - Kis késleltetéssel
 - Aszinkron módon
- Kérdés: mit figyeljünk?

CEP alapelvek

- „Komplex esemény”
 - Több elemi esemény összekapcsolása
- Tulajdonságok
 - Időzítések figyelembevétele (pl. csúszóablak)
 - Aszinkron működés
 - Oksági kapcsolatok, hierarchikus események
 - Korreláció
 - „Forward chaining”
- SQL-szerű query nyelvek
 - Pl. EPL: Event Processing Language
 - Feldolgozási folyamatba láncolható lépések
 - Event-Condition-Action
- Elosztott adatforrások
 - Adatbázisok, beérkező kérések, megfigyelt események, stb.
- Skálázhatóság
 - Cloud környezet

Események szemantikája

■ Drools:

	Point-Point	Point-Interval	Interval-Interval
A before B			
A meets B			
A overlaps B			
A finishes B			
A includes B			
A starts B			
A coincides B			

Alapok:

○ Allen-féle intervallum logika, 1983...

```

rule "reasoning on events over time"
when
  $a : A( )
  $b : B( this after[-2,2] $a )
  $c : C( this after[-3,4] $a )
  $d : D( this after[1,2] $b, this after[2,3] $c )
  not E( this after[1,10] $d )
then
  // do something
end
    
```

<http://www.jboss.org/drools/drools-fusion.html>

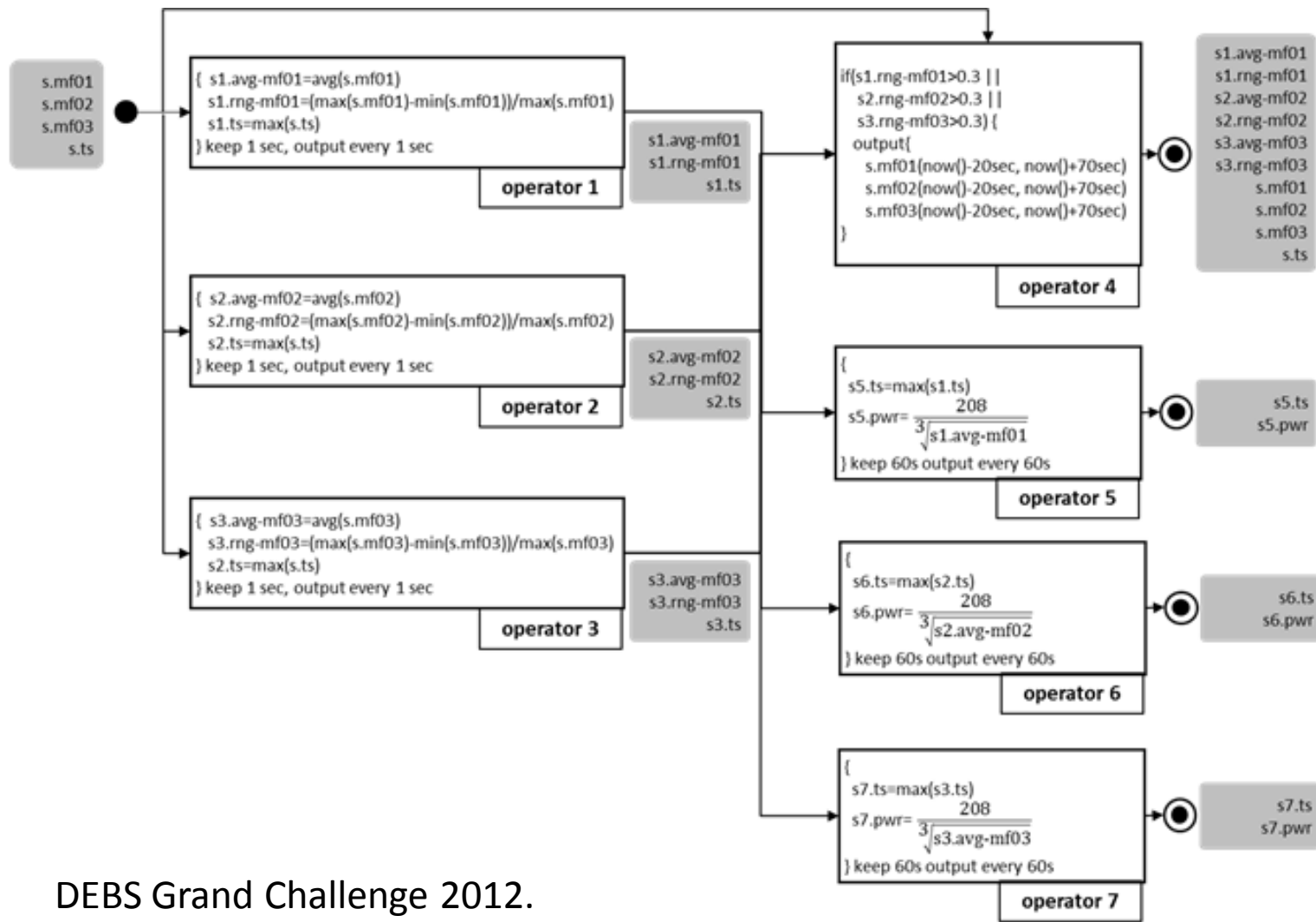
Tipikus CEP nyelvi elemek

- Esemény (típus definíció)
 - időbélyeg
- Eseményforrás
- Stream
- Időablak
- Időzítési operátorok
- Tolerancia (delta) az időbélyegekre
- Folyamatosan aktív lekérdezések

CEP alkalmazási területek

- Üzleti alkalmazások
 - Tőzsde, befektetések
 - „Treasury”
 - Kockázatkieértékelés
 - Hitelek árazása
 - Szállítmánykövetés
- „Business Activity Monitoring”
- Online visszaélések felderítése/megelőzése
 - Gyanús tranzakciók ellenőrzése
 - Fogadási adatok elemzése (pl. UEFA)
- Nagy IT rendszerek üzemeltetése
 - Komplex támadások felderítése
 - Metrika kiértékelés
- Biztonságtechnika
 - Pl. dDOS ellen
- <http://www.complexevents.com/>

Példa



DEBS Grand Challenge 2012.

CEP vs SzolgtInt

- Hogyan kapcsolódik a szolgáltatásorientált rendszerekhez?
- Döntéstámogatás
 - ~szabálykiértékelés
- Monitorozó logika
 - Működés helyessége
 - KPI kiértékelés
- CEP lehet maga is egy szolgáltatás
 - Eseményeket küld → folyamatokat indíthat
- CEP lehet az ESB része
 - Pl. tartalom alapú továbbítás
- Lehet az egész szolgáltatás CEP alapú...
 - Dinamikus folyamatok
 - Inkább a jövő...

Hasonló technológiák

■ Szabálymotorok

- CEP esetén tipikusan nem tartjuk meg az adatot, nem módosítjuk (hierarchia, riasztások)
- Nincs: explicit időzítés, eseményforrás, operátorok

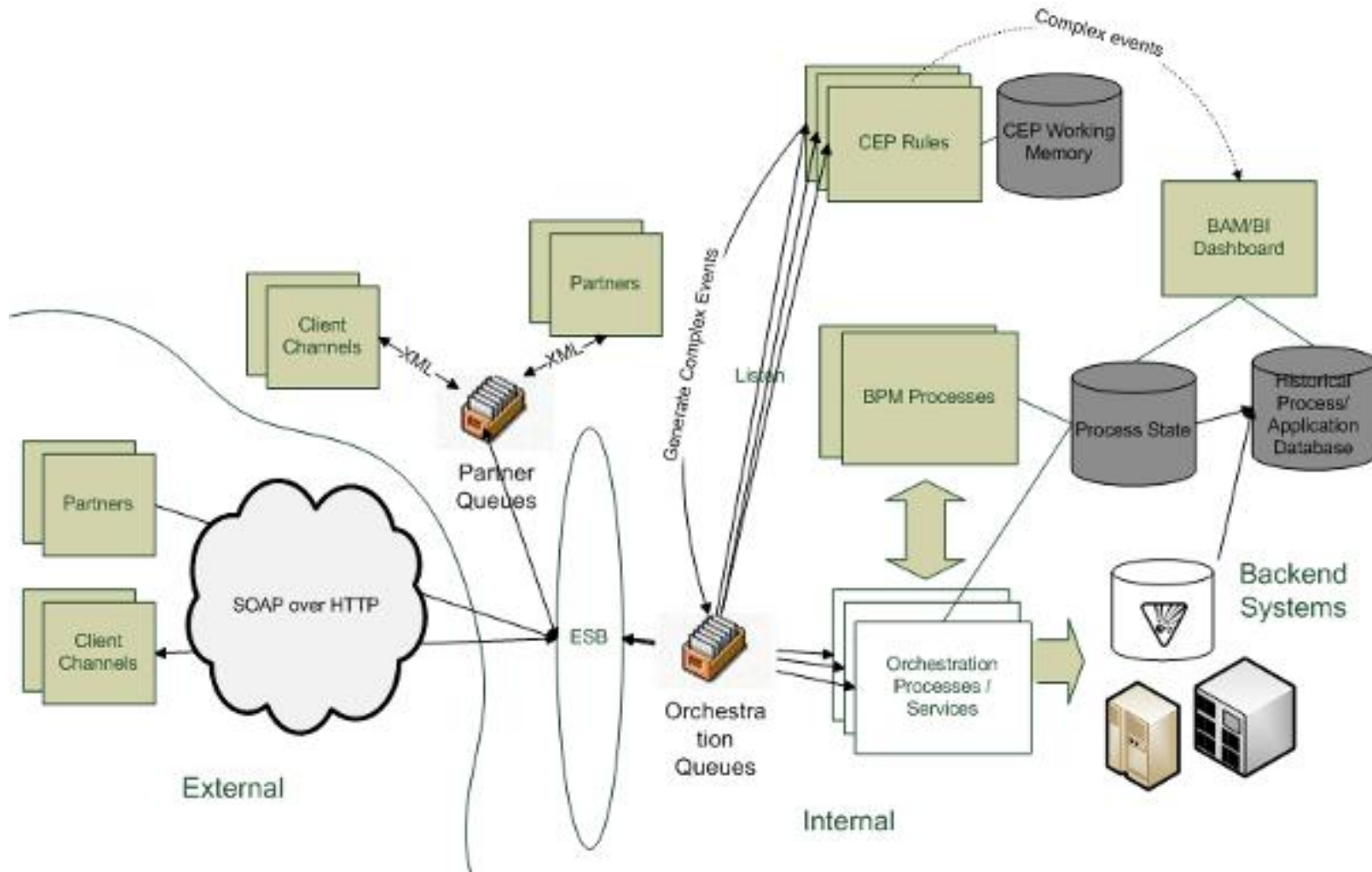
■ RDBM

- Lekérdezőnyelvek hasonlóak
- Nincs: időablak karbantartás, triggerek nem jól skálázódnak, válaszidő nehezen garantálható

■ „Döntéstámogató” rendszerek

- Nem realtime

Példa „architektúra”



<http://www.packtpub.com/article/cep-complex-event-processing-soa-service-oriented-architecture>

CEP eszközök

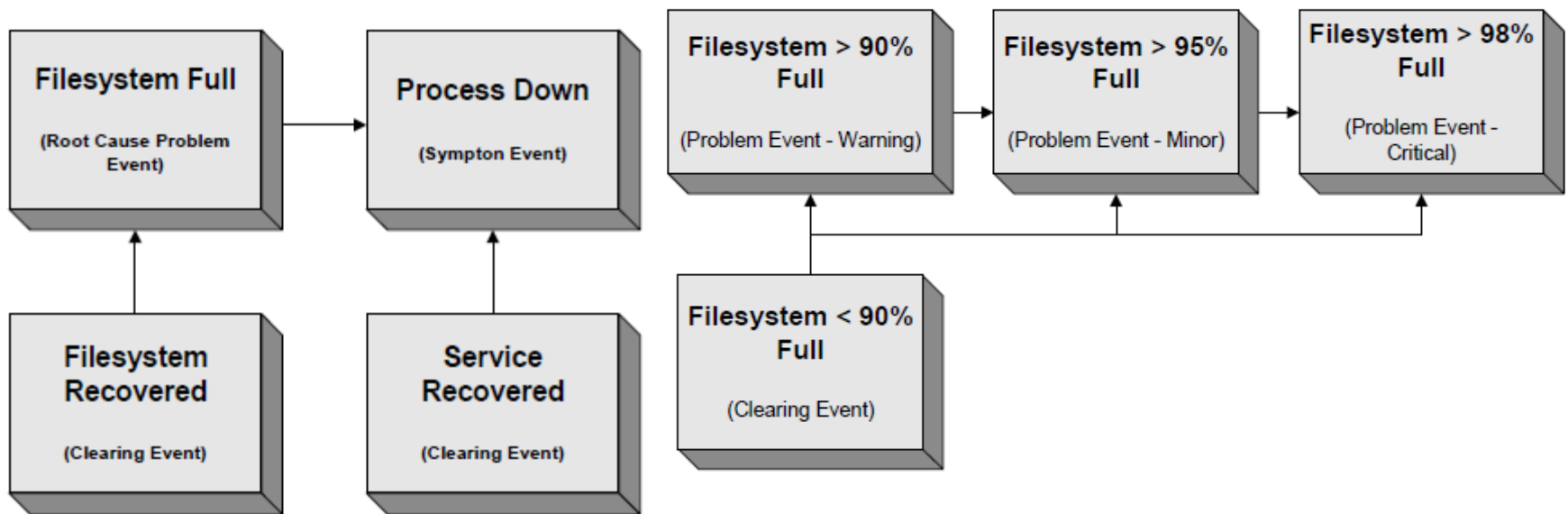
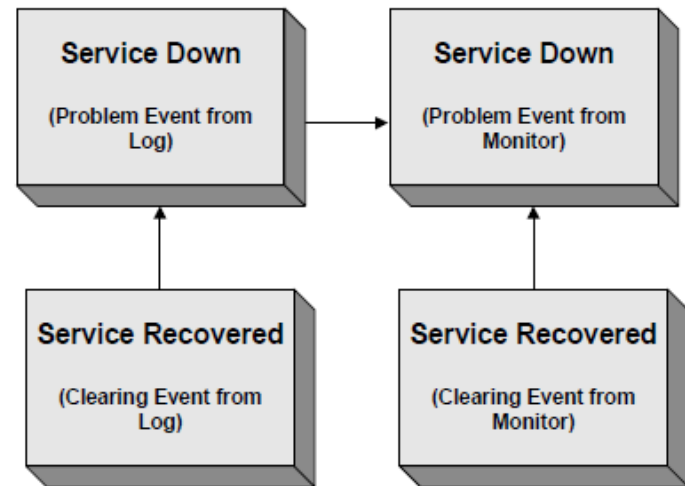
- Számptalan megoldás
 - Esper
 - Drools Fusion
 - IBM InfoSphereStreams (System S), WebSphere Decision Server
 - OpenESB - Intelligent Event Processor
 - Apache Hadoop + ráépülő projektek
 - TIBCO CEP
 - Microsoft StreamInsight
- Döntési szempontok
 - Eseményfeldolgozási logika
 - Áteresztőképeség
 - Elvárt válaszidő („low latency”)

Eseményfeldolgozás lépései

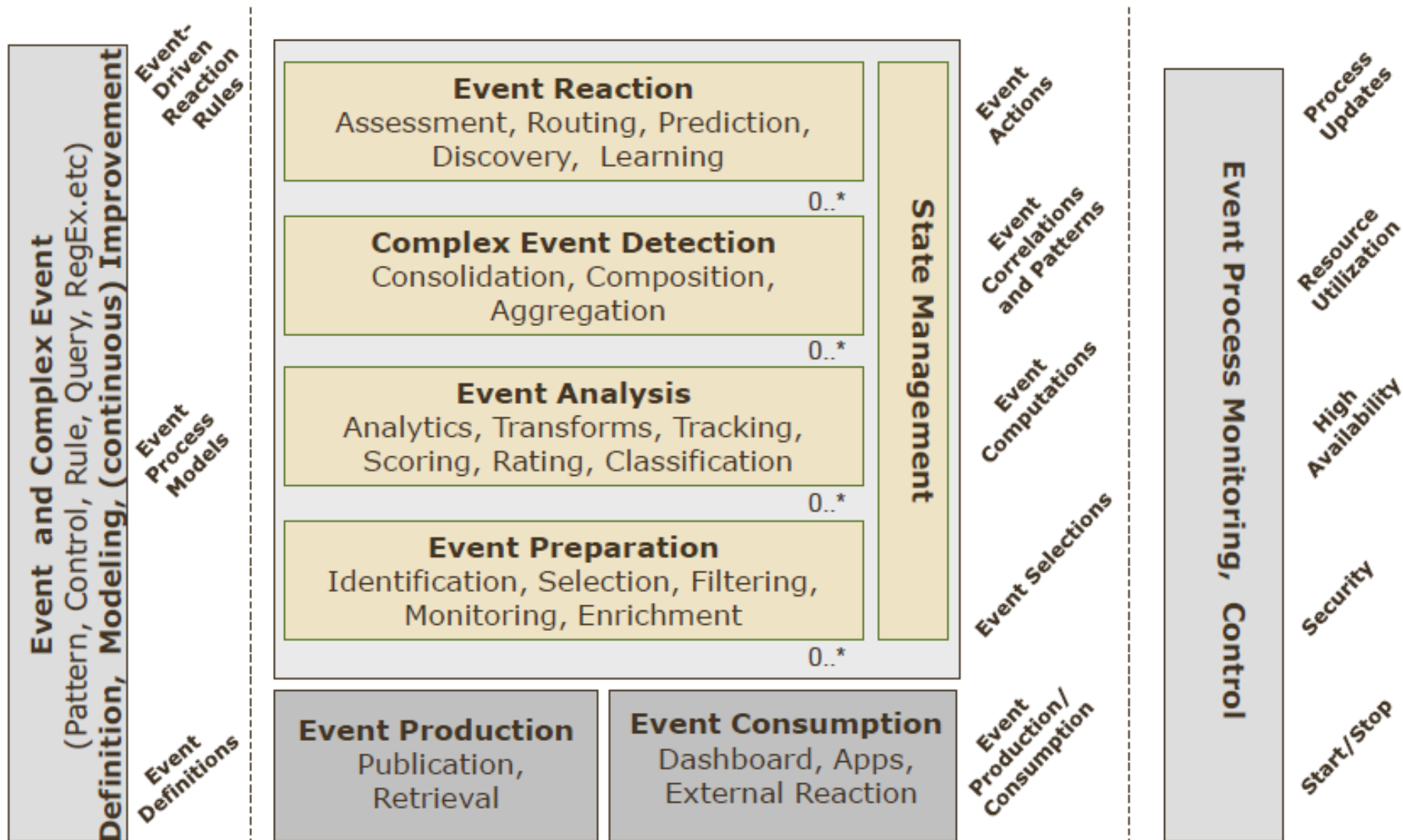
- Előkészítés
 - Események azonosítása („Mi honnan jön?”)
 - Események kiválasztása/szűrése
 - Események kiegészítése
 - Aggregálás
- Elemzés
 - Események „osztályozása” (rating, scoring, classification)
 - Elemzési minták (pl. elnyomás, topológia alapú függőségek figyelembevétele)
 - Események → komponens állapot
- Feldolgozás
 - Továbbítás
 - Előrejelzés
 - Esemény alapú tanulás

Mit kezdünk az eseményekkel?

- Korreláció
 - Szolgáltatás leáll- újraindul
- Eszkaláció
- Ok-hatás analízis



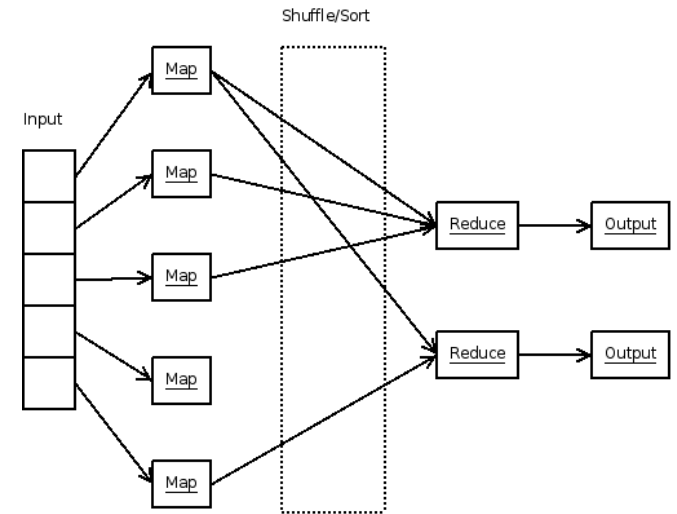
Referencia architektúra/feladatok



Event Processing Technical Group, 2011.

Példa: Map/Reduce algoritmus

- Map lépés
 - adat felosztása
- Reduce lépés
 - adat feldolgozása
- Példa
 - szöveg felosztása szavakra, szavak számának megállapítása
- Számos programnyelven
- Apache implementáció
 - Elosztott megoldás
 - Hadoop (+ Hadoop Distributed File System)
 - Ütemezés : Job Tracker, Task Tracker



CEP Benchmarking

■ Kihívások

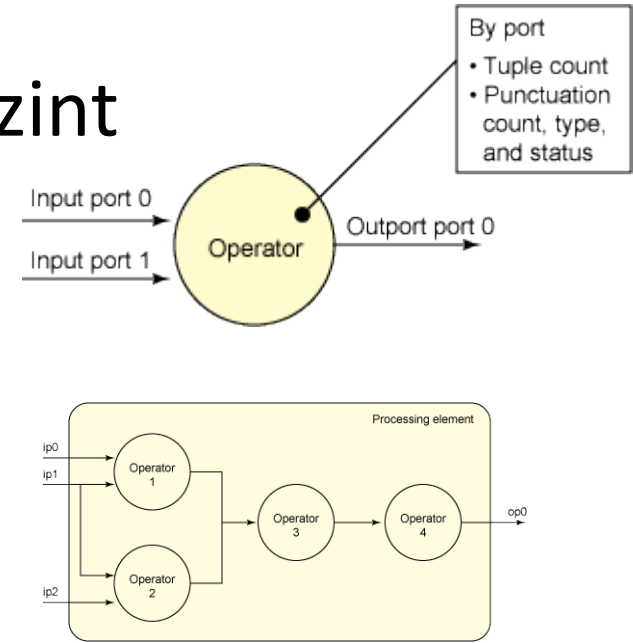
- Nagy tömegű eseménynél ritka illeszkedés
- Feldolgozás/adattovábbítás overhead
- Kapcsolat a feldolgozott események és az eredmény közt
- Terhelésgenerálás is lehet szűk keresztmetszet

■ Mit mérünk?

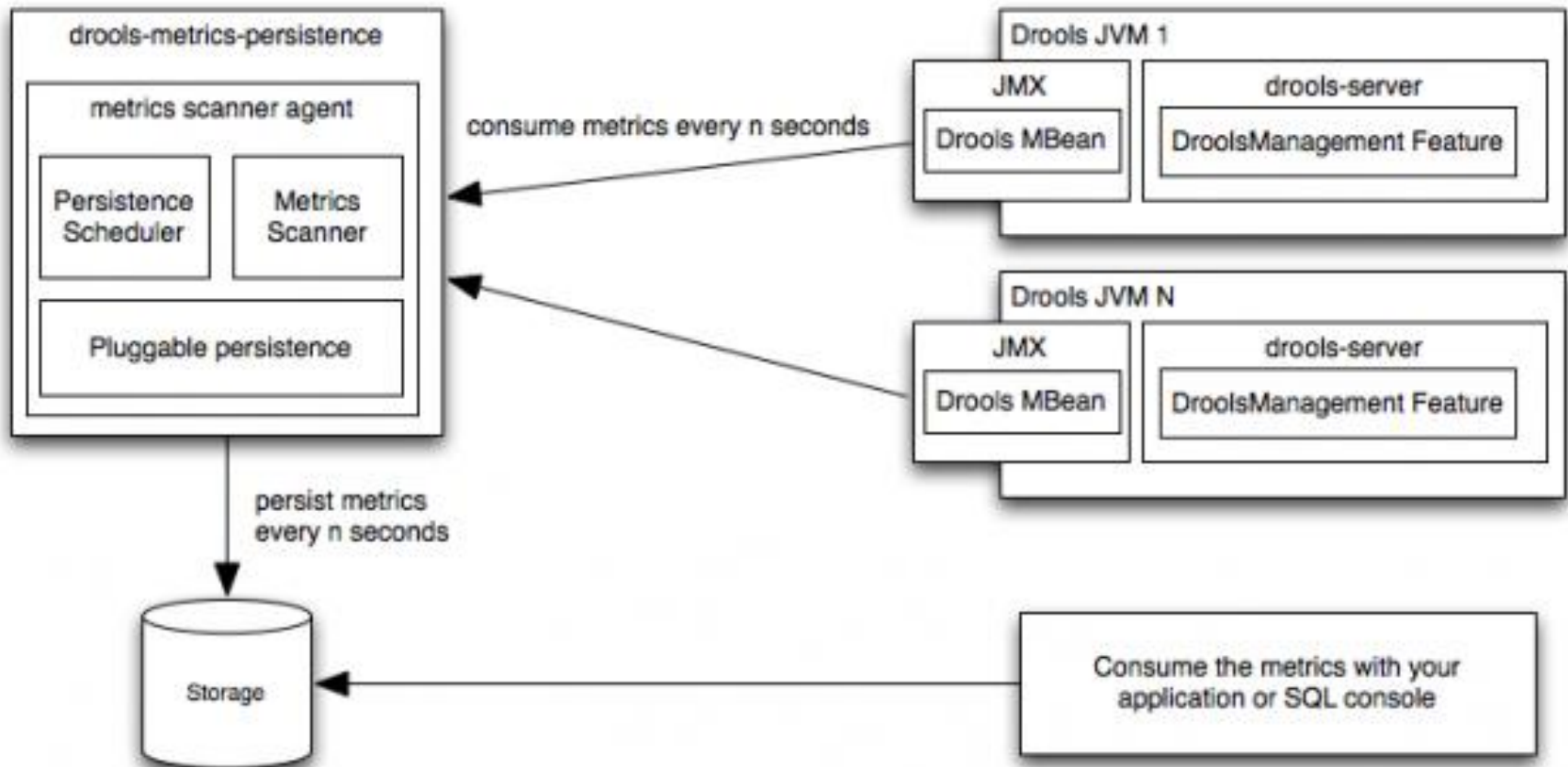
- Áteresztőképesség (bejövő/kimenő események)
- Válaszidő
- Skálázhatóság
- ... elvárt pontosság mellett
- Mögötte: query plan sharing, átfedő lekérdezések kiszűrése
- http://bicep.dei.uc.pt/images/3/34/BiCEP_wosp2010.pdf

Hogyan mérjük az eseményfeldolgozást?

- Példa: IBM InfoSphere Streams
- Operátor/feldolgozási egység szint
 - Feldolgozott/eldobott adatok
 - Továbbított adatok
 - Sorhossz
- Feldolgozási egység szintje

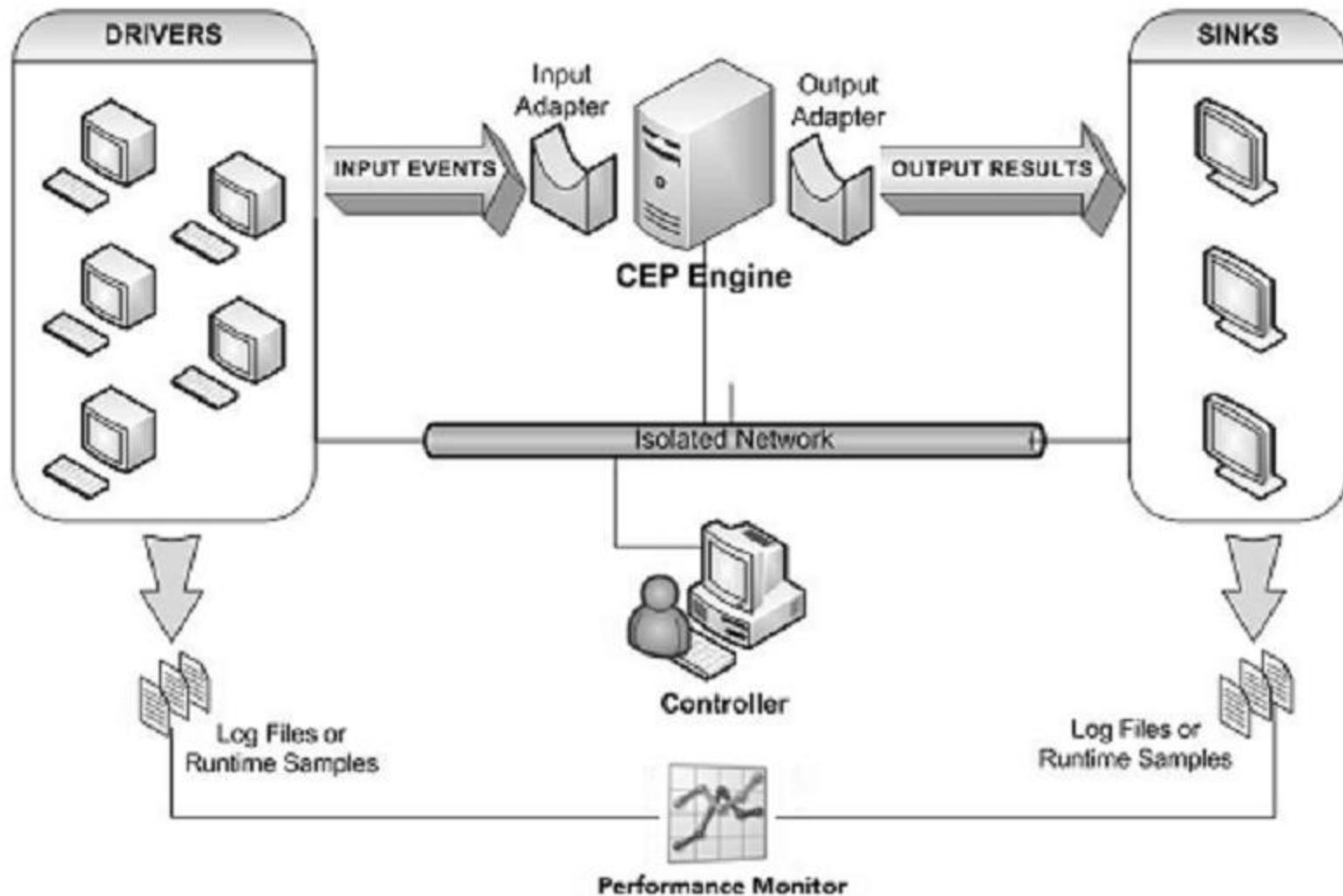


Példa: Drools metrika gyűjtés



<http://lucazamor.wordpress.com/2011/01/07/drools-metrics-persistence/>

Példa: FinCOS

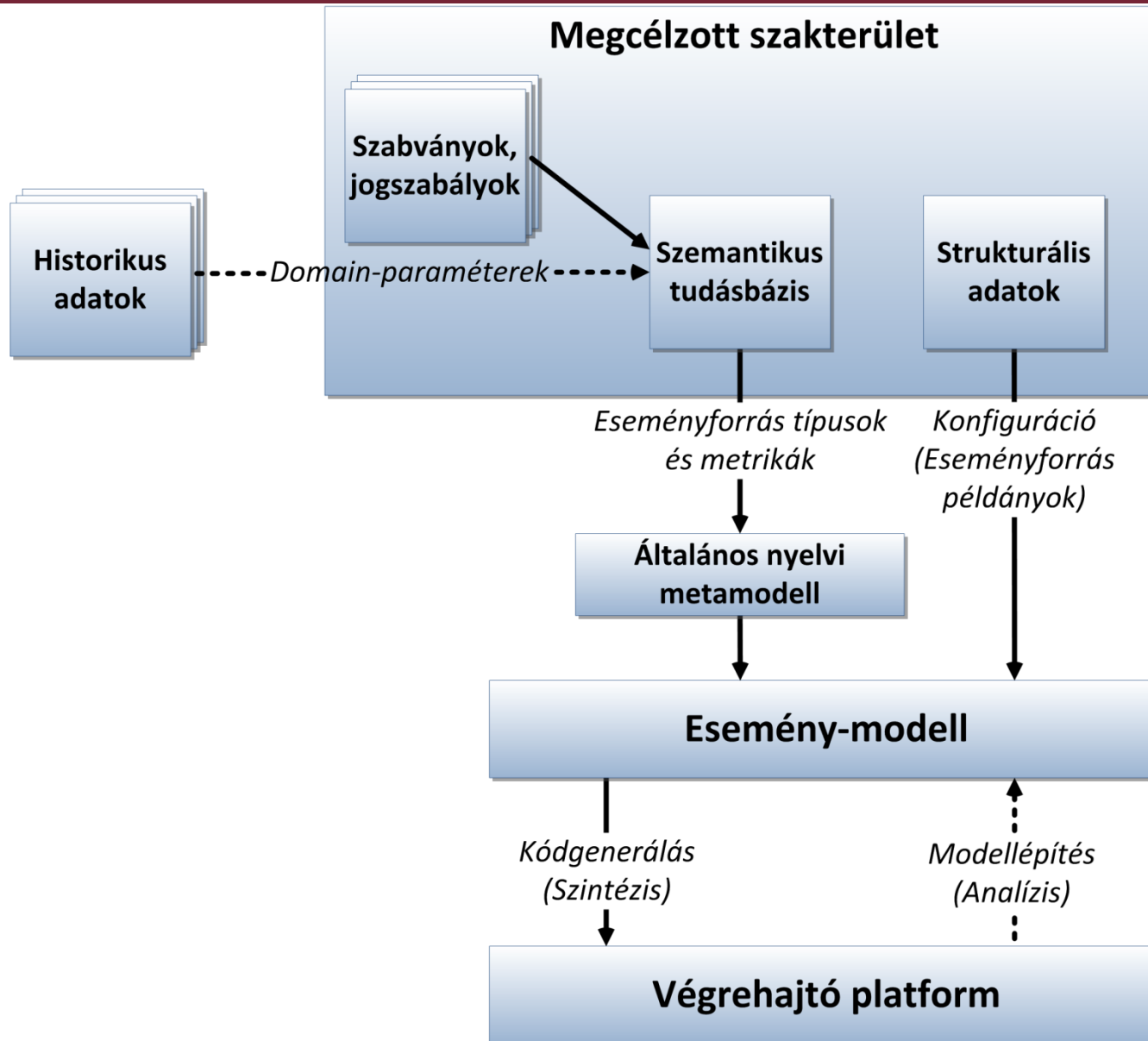


<https://code.google.com/p/fincos/>

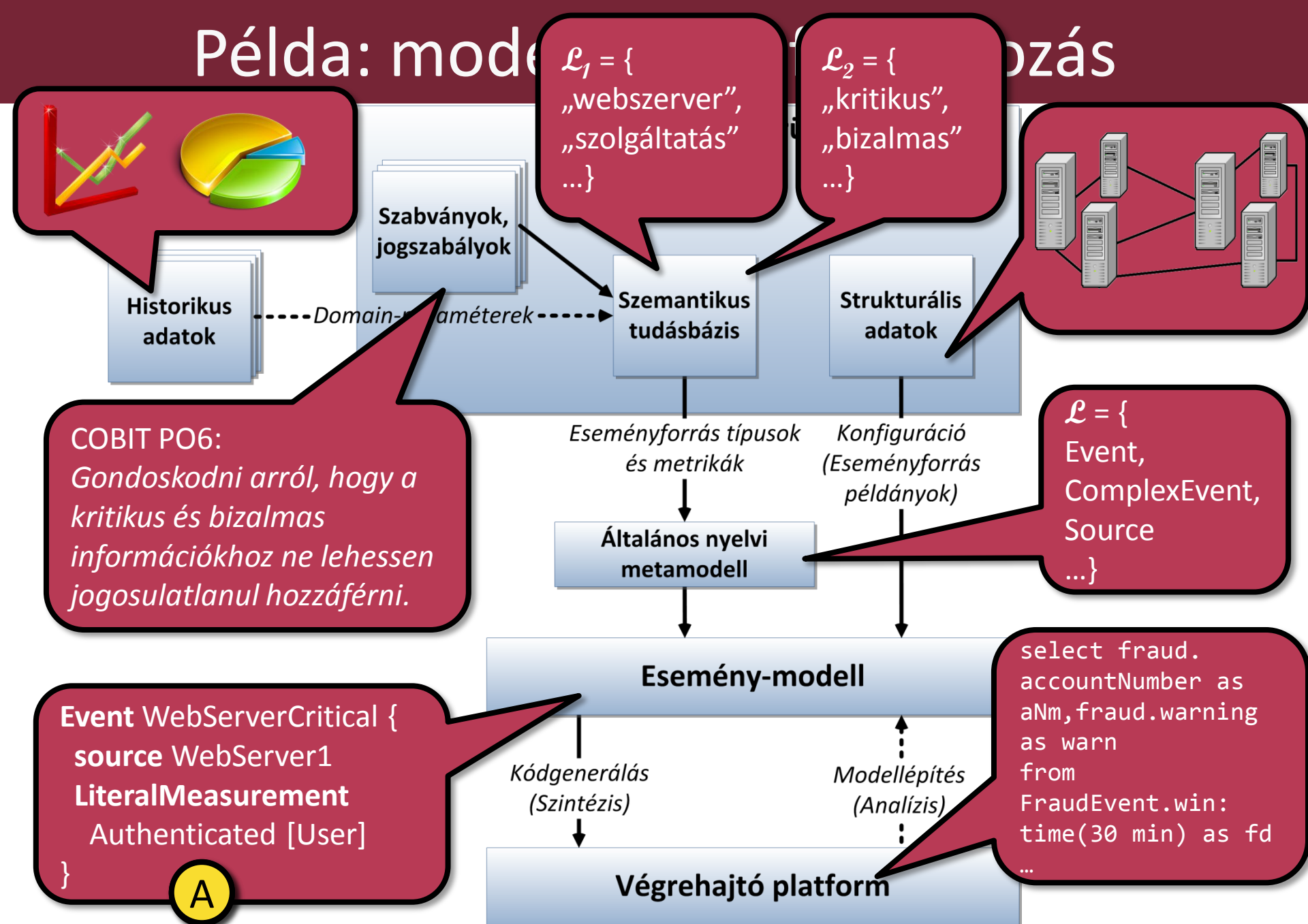
Kihívások

- Események szemantikája
 - Mit jelent? Melyiket figyeljük? (~100 eseményforrás)
 - Milyen kapcsolata van a rendszer dinamikus működésével? (folyamatok)
- Minta alapú tanulás
 - Pl. küszöbértékek hangolása
- Ritka események hatékony azonosítása
- Teljesség? Helyesség?
- Eseményleírás definíciója
 - Nincs egységes szabvány (BEMN, RuleML,)

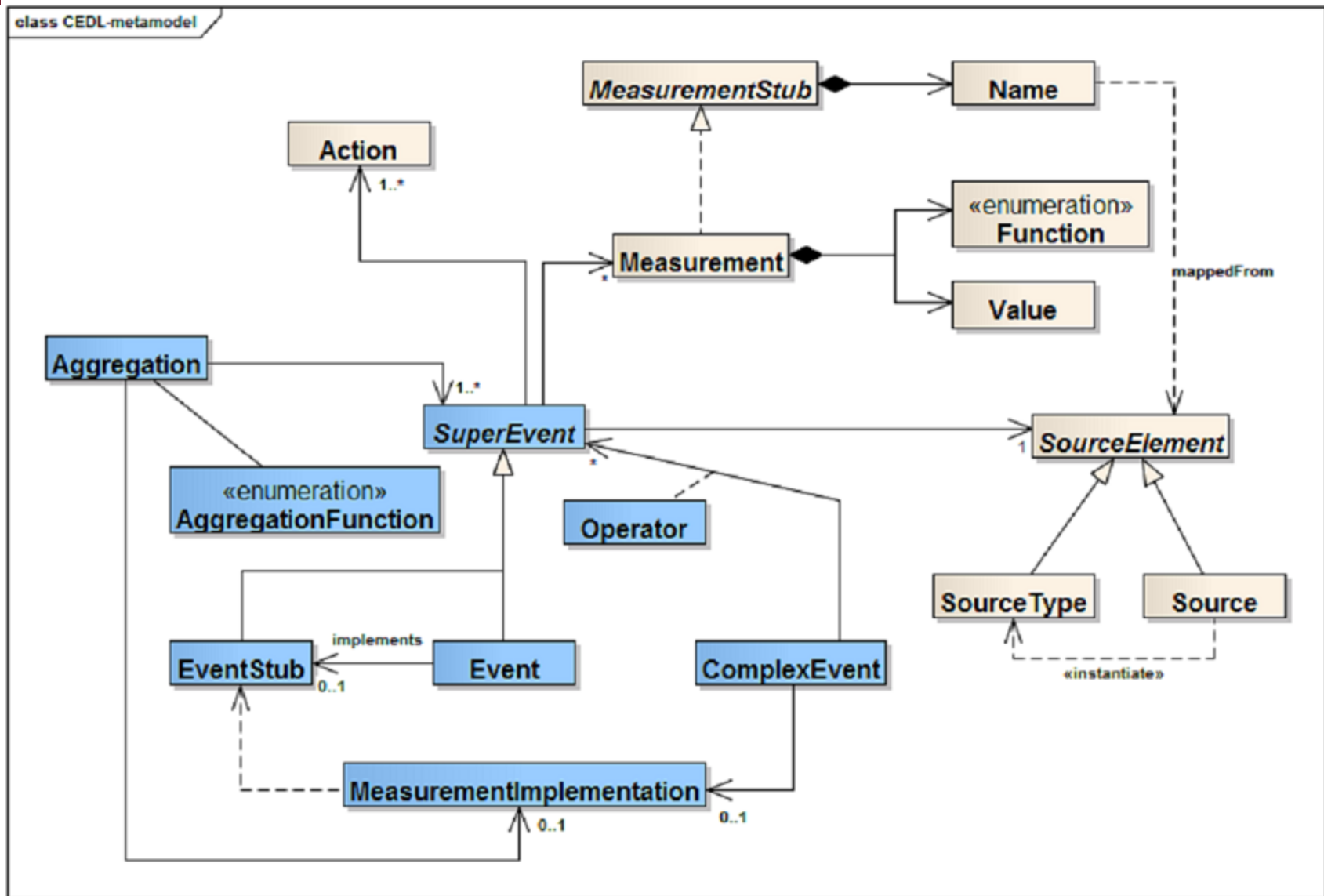
Példa: modell alapú feldolgozás



Példa: modellalkotás



Eseményleíró nyelv metamodellje



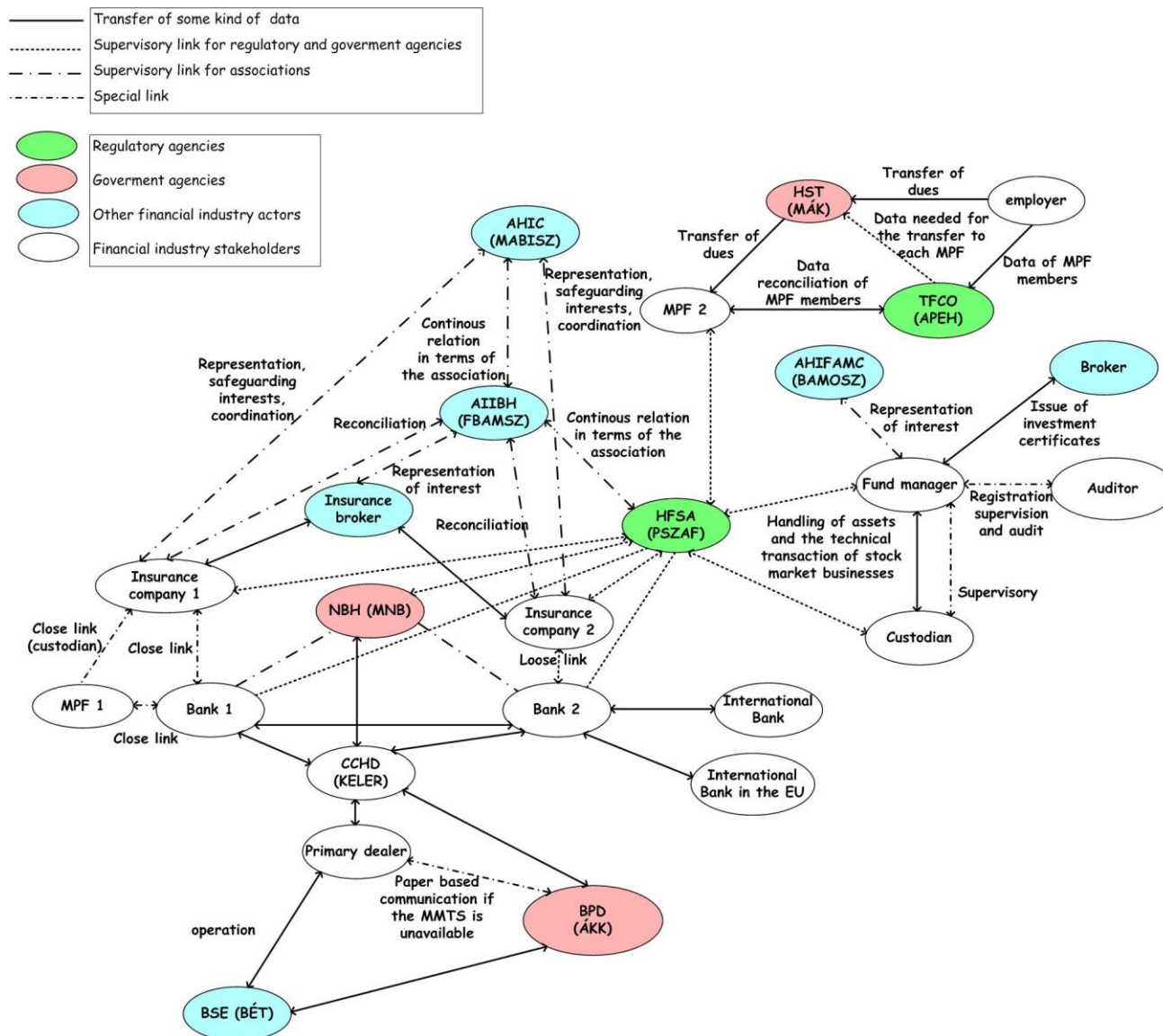
Esettanulmány: CoMiFin

Szolgáltatásalapú rendszerek, modellvezérelt fejlesztés,
komplex eseményfeldolgozás,...

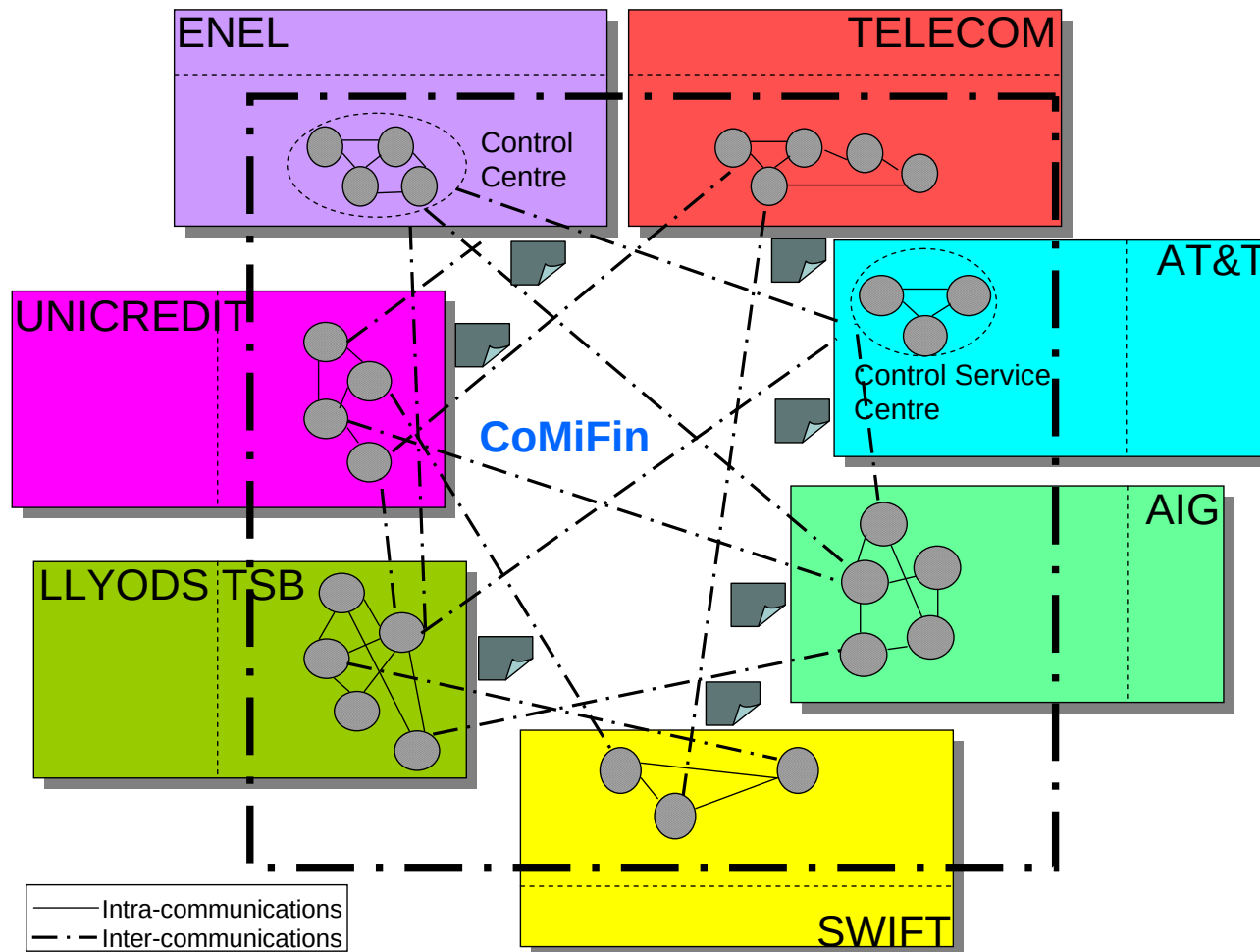
Esettanulmány: CoMiFin

- „Communication Middleware for Financial Infrastructures”
- Motiváció
 - Banki rendszerek egyre erősebben függenek külső szolgáltatóktól
 - Támadások egyre kifinomultabbak
 - Kritikus infrastruktúrák (pl. mobilhálózat, áramellátás, Internet) elleni komplex támadások kivédése
 - Hagyományos kommunikáció lassú (példa: 8 nap egy eset lezárása)
- Cél
 - Scheme to set up and manage a secure environment (software, hardware, monitoring tools, etc.) for information exchange and analysis
- Tanszéki spin-off (OptXware) vezette a demonstrátor fejlesztését

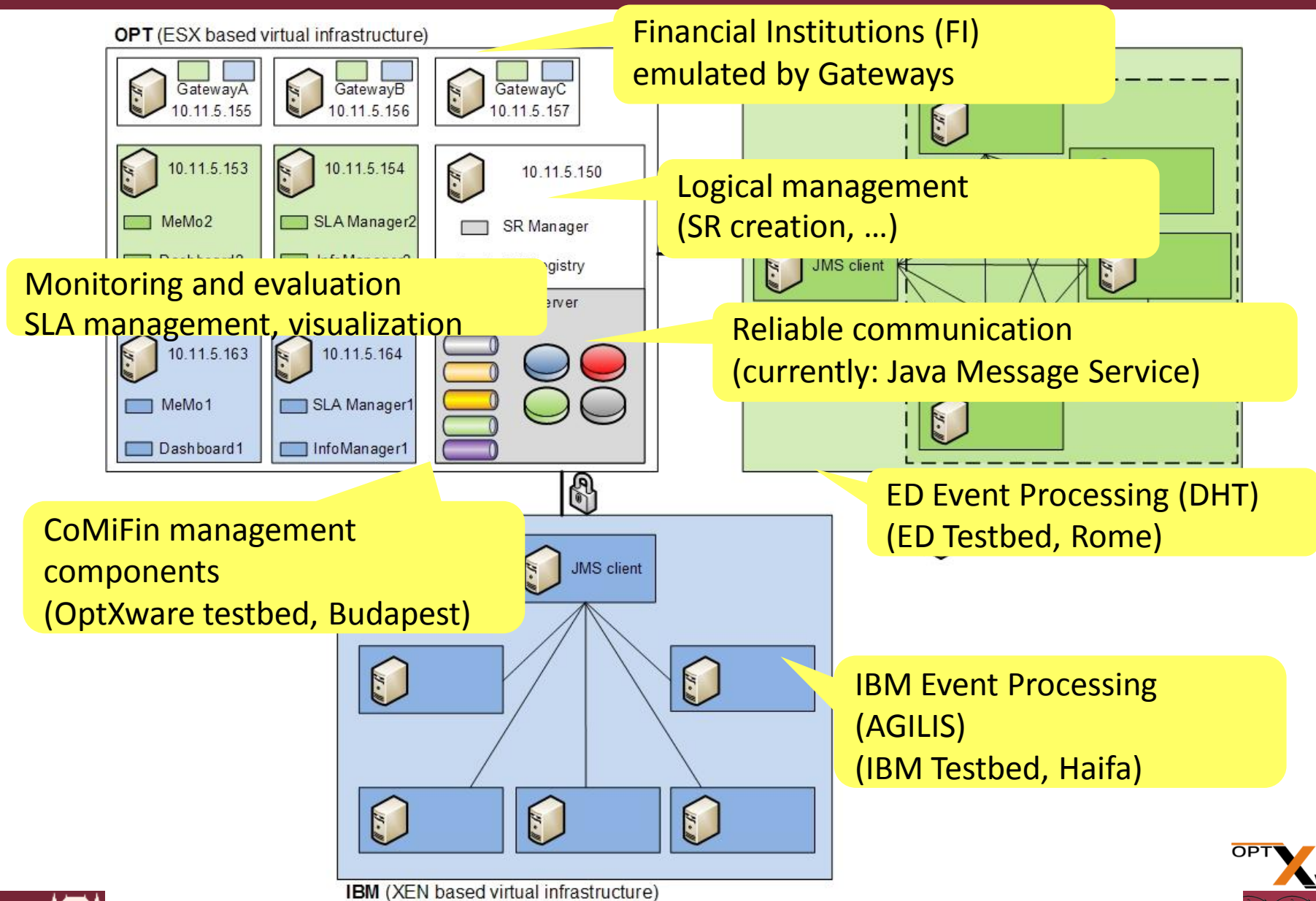
Példa: magyar infrastruktúra



Logikai architektúra



Architektúra



Eredmények megjelenítése

JBoss Portal 2.7.2-GA - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://10.11.5.153:8080/portal/auth/portal/Dashboard/6_AlertList

Most Visited Getting Started Latest Headlines

Popup blocker detected JBoss Portal 2.7.2-GA

JBoss Portal

Home 5_SR_Manager_Page 6_AlertList

Alert details (time, source, target, etc.)

Logged in as: sr_manager

Dashboard | Copy to my dashboard | Logout

AlertList

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			Service4	17X.XX.X3X.X29	-0.1718082
2010-07-01 10:07:32.0	DHT Analytics	Bank of Vanyar			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			-0.182237
2010-07-01 10:07:09.0	DHT Analytics	Bank of Noldor	ALERTMitM	Statistical anomaly detected			-0.18592
2010-07-01 10:07:08.0	DHT Analytics	Bank of Noldor	ALERTMitM	Statistical anomaly detected	Service11	19X.X1X.X0X.X05	-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
				Statistical anomaly			

Service effected

Score on the alert

Done

Esettanulmány: IT infrastruktúra események monitorozása

Dávid István, Gönczy László

Modellalapú fejlesztési módszer komplex események
feldolgozásához (European Dependable Computing
Conference 2012, Mesterpróba 2012)

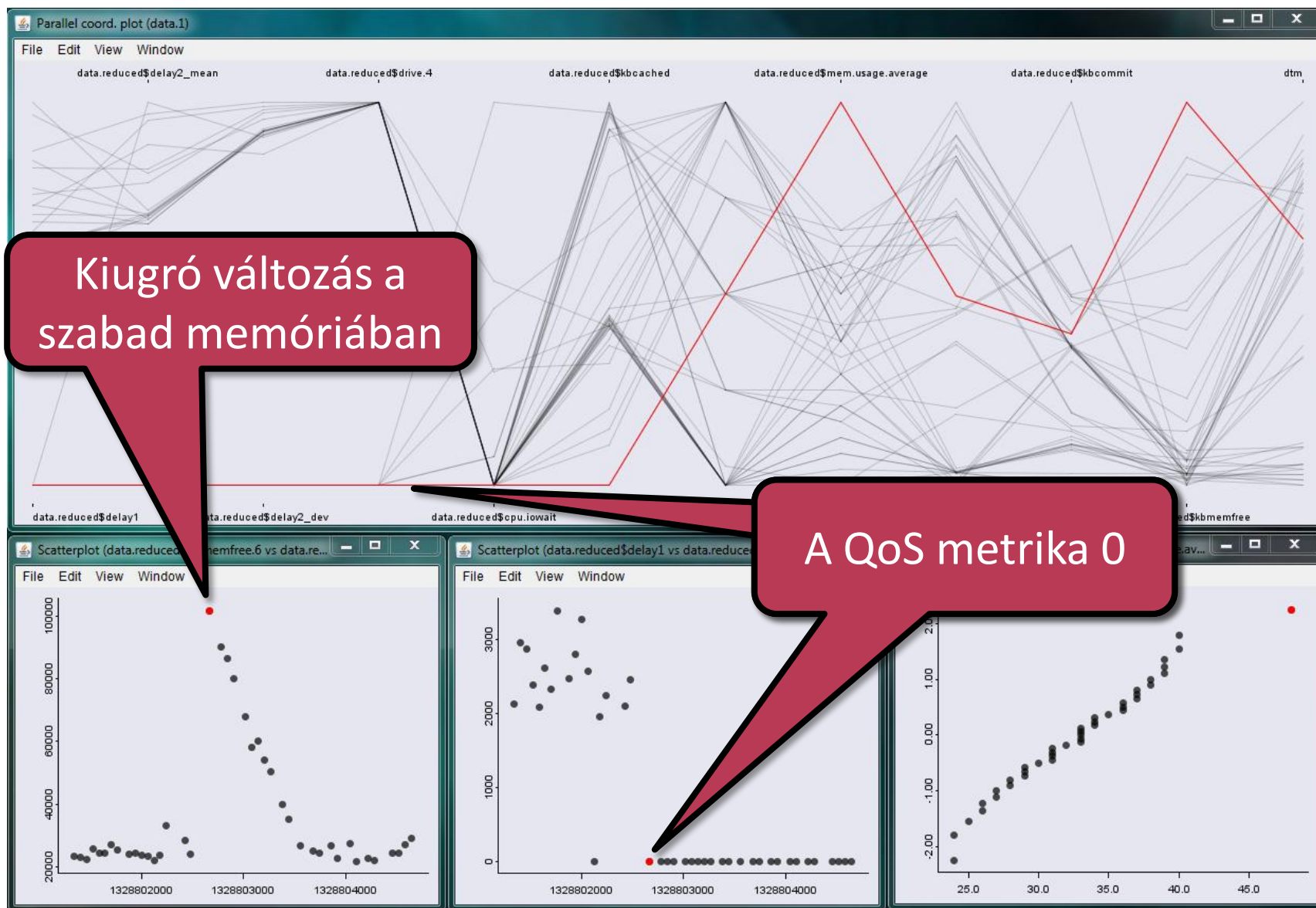
Esettanulmány



- BigBlueButton
- Többféle mért attribútum
 - CPU terheltség, rendelkezésre álló memória, cache...
 - QoS metrikák: az audiovizuális adat késleltetése
- A rendszer túlterheltségének karakterisztikája:
 - A mért késleltetés 0 (definíció szerint)
 - Hirtelen növekedés a szabad memóriában
 - Fail-silent működés
- A cél:
 - Detektáljuk a túlterhelést és konfiguráljuk újra a rendszert
- Az **eseményfolyam** ebben az esetben: Az infrastruktúra elemek metrikáinak folyamatosan mért értékei

*Imre Kocsis, András Pataricza, Zoltán Micskei, István Szombath, András Kövi & Zolt Kocsis
Cloud Based Analytics for Cloud Based Applications, ICA CON 2012.*

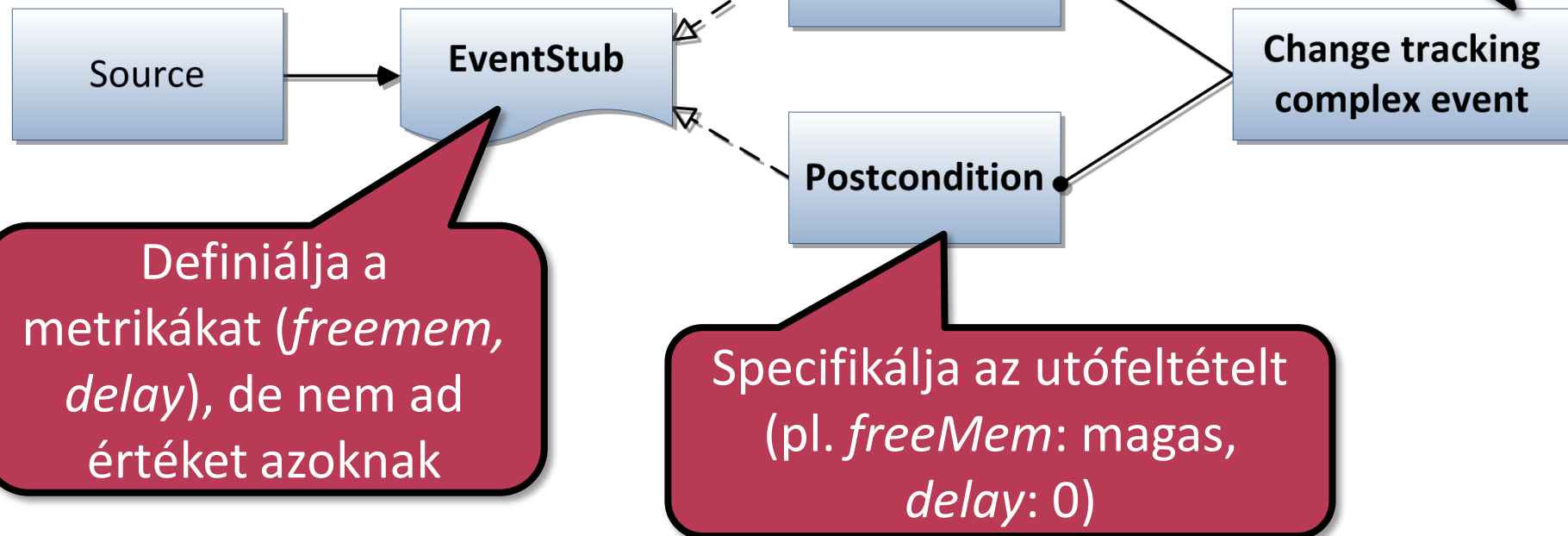
Túlterhelés vizuálisan ábrázolva



A modellezés áttekintése

- *Change tracking* tervezési minta

Specifikálja az előfeltételt
(pl. *freeMem*: nem magas,
delay: nem 0)



Források

- <http://www.complexevents.com/>
- <http://www.slideshare.net/isvana/epts-debs2011-event-processing-reference-architecture-and-patterns-tutorial-v1-2>
- <http://www.ibm.com/developerworks/data/library/techarticle/dm-1203infostreamsfeatures1/index.html>
- <http://www.jboss.org/drools/drools-fusion.html>
- <http://www.thetibcoblog.com/2010/03/04/how-does-cep-fit-into-bpm-and-soa-environments/>
- <http://tdk.aut.bme.hu/Conf/TDK2011/szoftver/Modellalapu-fejlesztési>
- <http://www.packtpub.com/article/cep-complex-event-processing-soa-service-oriented-architecture>
- <http://books.google.hu/books?id=tx2NXQEo47EC&printsec=frontcover&hl=hu#v=onepage&q&f=false>
- <http://bpt.hpi.uni-potsdam.de/pub/Public/GeroDecker/edoc2007-eventlanguage.pdf>
- [http://www.soa.si/wp-content/documents/clanki/WSDL and BPEL extensions for Event Driven Architecture.pdf](http://www.soa.si/wp-content/documents/clanki/WSDL_and_BPEL_extensions_for_Event_Driven_Architecture.pdf)
- <http://public.dhe.ibm.com/software/data/sw-library/infosphere/casestudy/Assessing-Transport-Systems-casestudy-in-Dublin.pdf>
- <http://www.slideshare.net/opher.etzion/debs2009-event-processing-languages-tutorial>