



Enabling Grids for E-scienceE

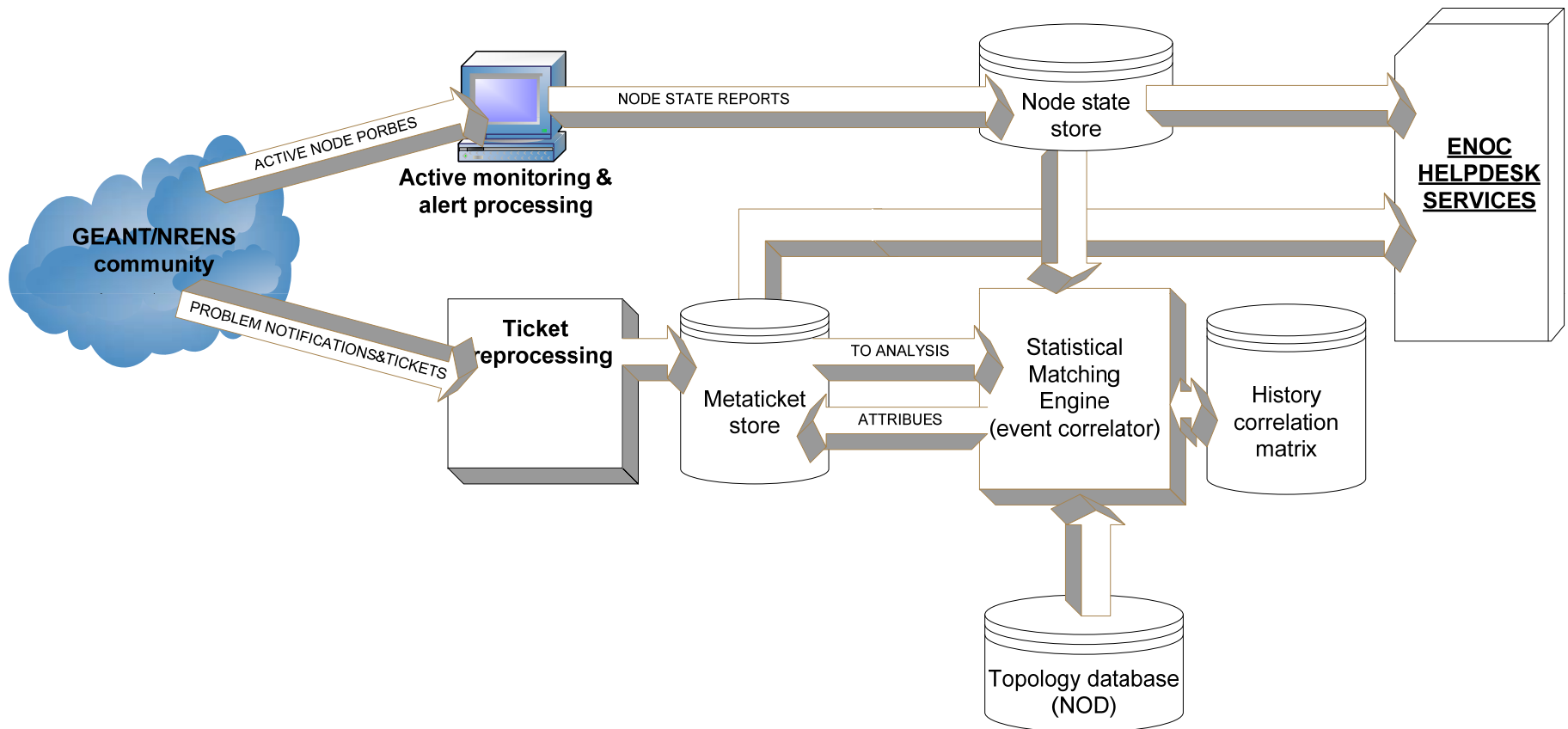
# Statistical matching approach for trouble ticket caring in multi- provider GEANT/NREN environment

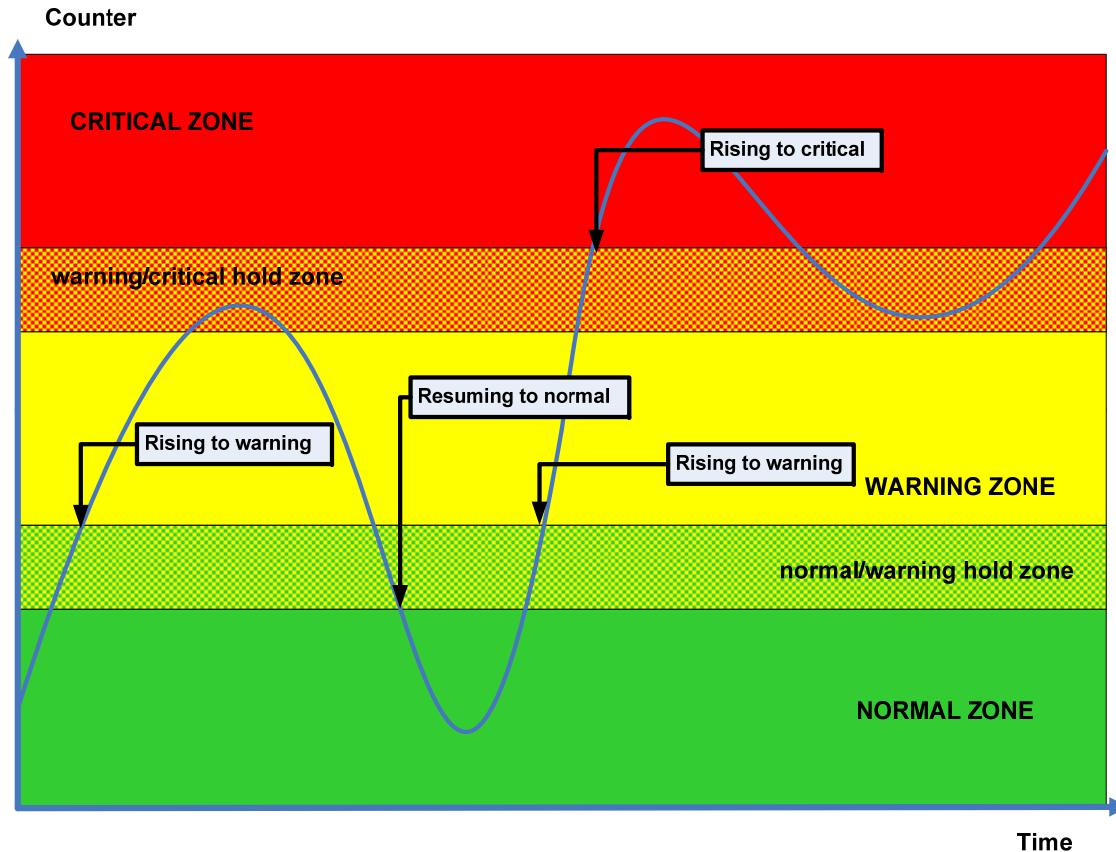
*Veniamin Konoplev (RRC-KI)  
All Hands meeting GARR/Rome  
25 March 2009*

[www.eu-egee.org](http://www.eu-egee.org)



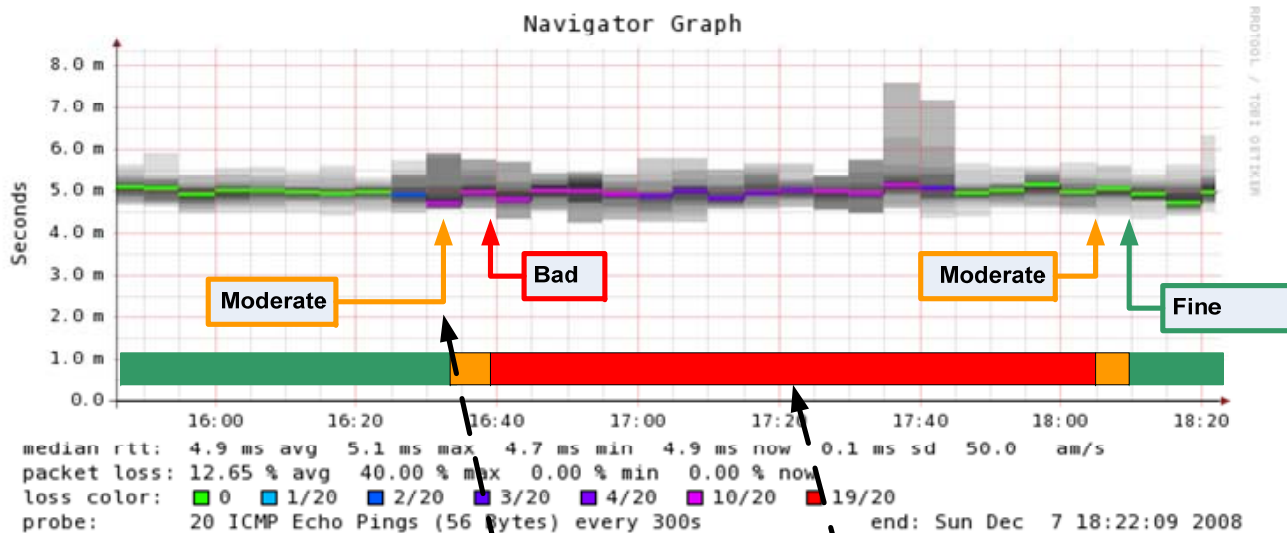
- Since 2006 ENOC deals with GEANT/NREN trouble ticket (TT) processing. It is subscribed to the main provider's notification streams and tries to take advantage of this additional information in daily ENOC operational tasks. In order for this process to be successful, the means for network impact estimation by TT content are needed. Thus ENOC staff could be aware of potential EGEE problems caused by GEANT/NREN notifications.
- The main problem – each NREN uses its own custom scheme of problem identification. Thus trouble ticket content is a very hard stuff to estimate “what” and “when” should we expect for EGEE infrastructure.
- Basic approach used by ENOC was to find a lexicographical similarity of TT location names with object names in the ENOC topological database (NOD). Due to NREN topological complexity and evolution this approach has serious restrictions.
- Further trouble ticket tool development focused on alternative statistical approach to associate TT content with real observed network connectivity degradations. This approach needs a minimal human support, can automatically react to topological evolutions and, with certain conditions, could supply the applicable results.





Alert	Priority	Rising	Falling	History	Fast
Unreachable	1	70%	50%	12	3
Bad	2	7%	5%	18	4
Moderate	3	2%	1%	24	5
Fine	4	Fake alert always active			

## JINR-LCG2



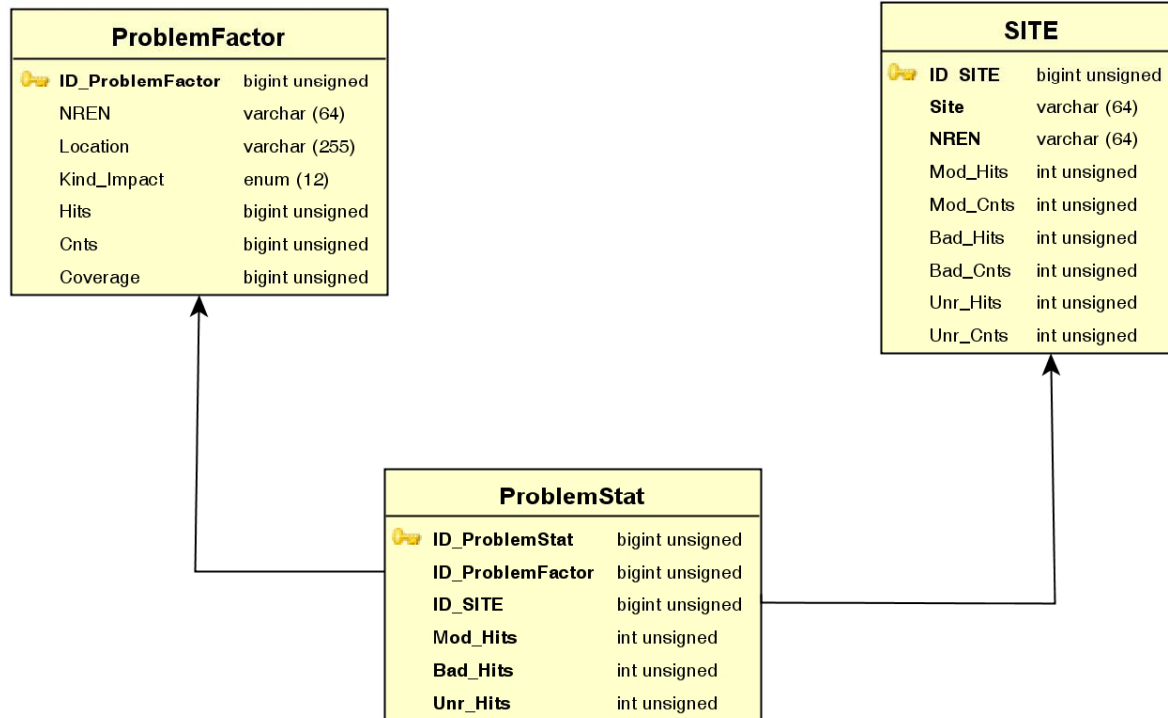
(1) Alert notifications (2) Resulting node status

### The matcher output (3)

```
Dec 7 16:34:06 loss-moderate (rising)
Dec 7 16:39:07 loss-bad (rising)
Dec 7 18:04:07 loss-moderate (resuming)
Dec 7 18:09:08 loss-fine (resuming)
```

- **Purpose:** to find dependencies between TT metadata and potential impact to EGEE sites.
- **Input data:**
  - Slices [Interval, NREN, Location, KindImpact] from the ticket store.
  - Actual monitoring alert events.
- **Base analyzing statistic element is Hit.**
  - Hit associates a ticket to a site if site experienced a problem within the ticket time interval.
  - Hit has a severity corresponding to the most hardest alert for the ticket time interval
- **Hit statistic is calculated and stored for:**
  - Ticket problem identification triplets: [NREN, Location, KindImpact]
  - Sites: [NREN, Site]
  - Problem-to-site associations: [NREN, Location, KindImpact, Site]
- **Hit statistic is periodically recalculated over last several months. Thus, the system stays up-to-date with changes of NREN topology and identification scheme.**

# History correlation matrix



Powered by yFiles

# Matching summary for GEANT2 tickets

Location	Tickets	Hits	Ratio	Validated
AMSTERDAM-NEW YORK	3	2	0.6667	Yes
REDIRIS BACKUP	3	2	0.6667	Yes
TEIN2-DE	4	2	0.5000	Yes
CYNET BACKUP	3	1	0.3333	Yes
REDIRIS	3	1	0.3333	Yes
RT1.VIE.AT	3	1	0.3333	Yes
<i>ULAKBIM BACKUP</i>	<i>12</i>	<i>3</i>	<i>0.2500</i>	<i>No</i>



# Matching details for GEANT2 tickets

Ticket Location	Affected NRENs	Problem prediction probability		Prediction Significance
		Bad connectivity	Unreachable	
AMSTERDAM-NEW YORK	CAnet	66%	0%	9%
REDIRIS BACKUP	REDIRIS	33%	33%	20%
TEIN2-DE	JANET	0%	25%	83%
	FCCN	25%	0%	19%
	CAnet	25%	0%	9%
	CESNET	25%	0%	22%
CYNET BACKUP	SANET	33%	0%	6%
REDIRIS	REDIRIS	33%	0%	20%
RT1.VIE.AT	SWITCH	0%	33%	8%
	JANET	0%	33%	83%
	IUCC	0%	33%	25%
ULAKBIM BACKUP	SWITCH	8%	0%	8%
	PIONER	8%	0%	17%
	ULAKNET	8%	0%	7%
	SANET	8%	0%	6%
	RENATER	8%	0%	3%

# Matching summary for NREN tickets

NREN	Location	Tickets	Hits	Ratio	Validated
REDIRIS	ES / NODO REGIONAL AST	3	3	1.0000	Yes
<i>RENATER</i>	<i>FR / PARIS-RÉUNION</i>	<i>4</i>	<i>3</i>	<i>0.7500</i>	<i>No</i>
RENATER	FR / CACHAN	3	2	0.6667	???
RENATER	FR / CADARACHE	3	2	0.6667	???
RENATER	FR / STRASBOURG	3	2	0.6667	Yes
<i>RENATER</i>	<i>FR / CAYENNE-MEDIASERV</i>	<i>4</i>	<i>2</i>	<i>0.5000</i>	<i>No</i>
RENATER	FR / ORSAY	8	4	0.5000	Yes
<i>RENATER</i>	<i>FR / PARIS-MAYOTTE</i>	<i>4</i>	<i>2</i>	<i>0.5000</i>	<i>No</i>
RENATER	FR / GRENOBLE	5	2	0.4000	Yes
RENATER	FR / LE MANS – TOURS	3	1	0.3333	???
RENATER	FR / JUSSIEU	3	1	0.3333	Yes
RENATER	FR / DIJON	3	1	0.3333	???
REDIRIS	ES / STM-4 ARA-CAT	3	1	0.3333	???
<i>RENATER</i>	<i>FR / MAYOTTE</i>	<i>3</i>	<i>1</i>	<i>0.3333</i>	<i>No</i>
...	...	...	...	...	...

- The period analysis (3 months) is too short. The majority of matched locations have only 3-4 corresponding tickets resulting in casual errors.
- The prototype was located in RBENT while connection from RBNET to GEANT experiences casual connectivity degradation and these cases can be mistakenly treated as target sites connectivity degradation. Current system revision can detect such cases but not with 100% accuracy.
- Ticked are not categorized yet (KindImpact field is ignored). Thus all ticket are processed even those not significantly affecting the network.

- A new methodology which helps ticket interpretation in large-scale multi-provider environment is represented here.
- We succeed to obtain acceptable results for GEATN2 tickets matched against NREN connectivity.
- NREN tickets matching were too poor to be practically used yet. This problem needs a detail analysis and further elaborations.
- Future work plan includes:
  - production version setup
  - automatic ticket ranking based on matching results
  - further tuning of matching algorithm