

Grid Oversight, Status and Issues

Ron Trompert
COD



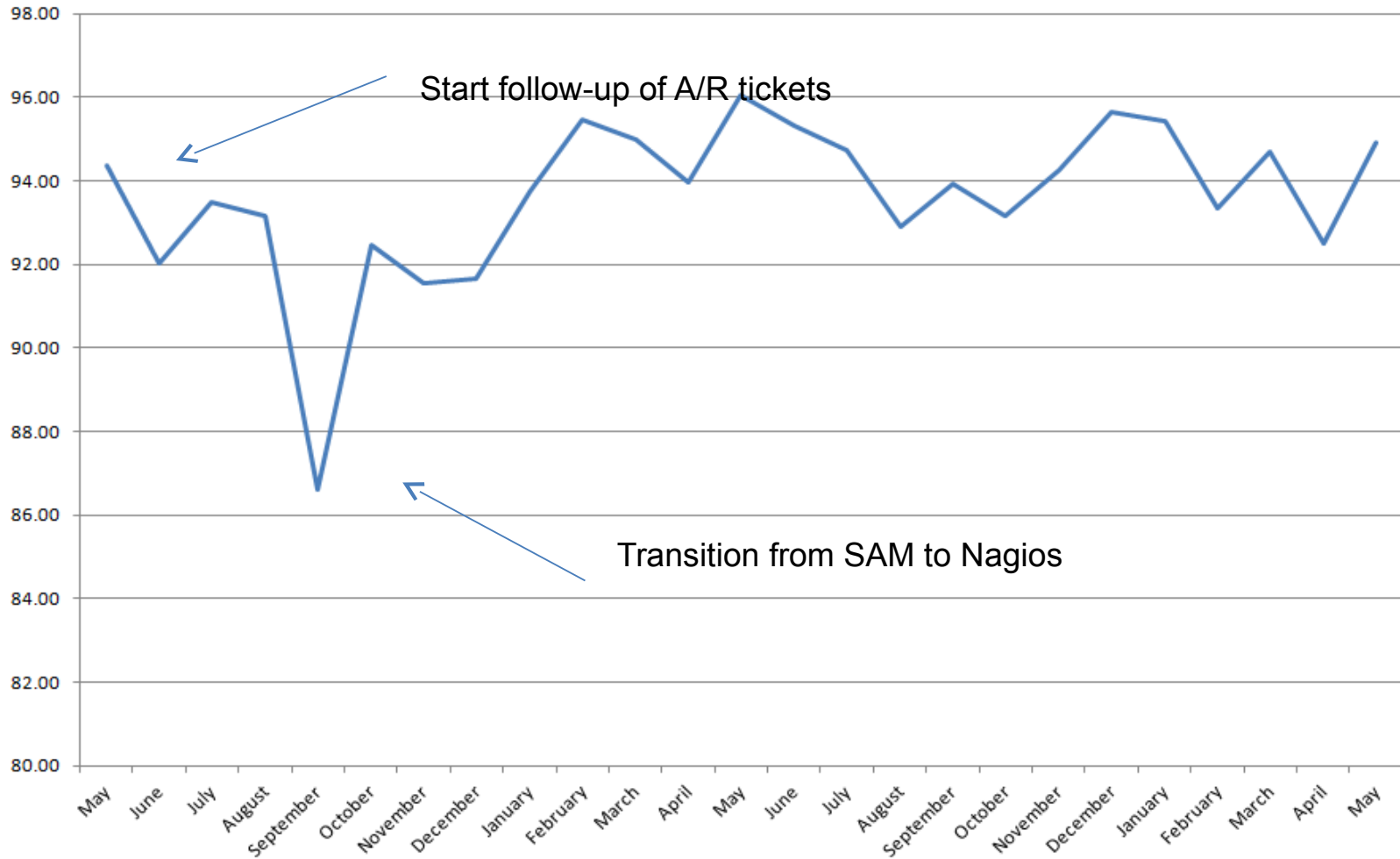
- CERN
- Central Europe (Austria, Czech Republic, Hungary, Poland, Slovakia, Slovenia)
- France
- Germany and Switzerland
- Ireland and UK
- Italy
- Northern Europe (Belgium, Denmark, Estonia, Finland, The Netherlands, Norway, Sweden)
- Russia
- South-East Europe (Bulgaria, Cyprus, Greece, Israel, Romania)
- South-West Europe (Portugal, Spain)

- Transition from 10 ROCs to now 37 NGIs
- Handover of first-line support and grid oversight

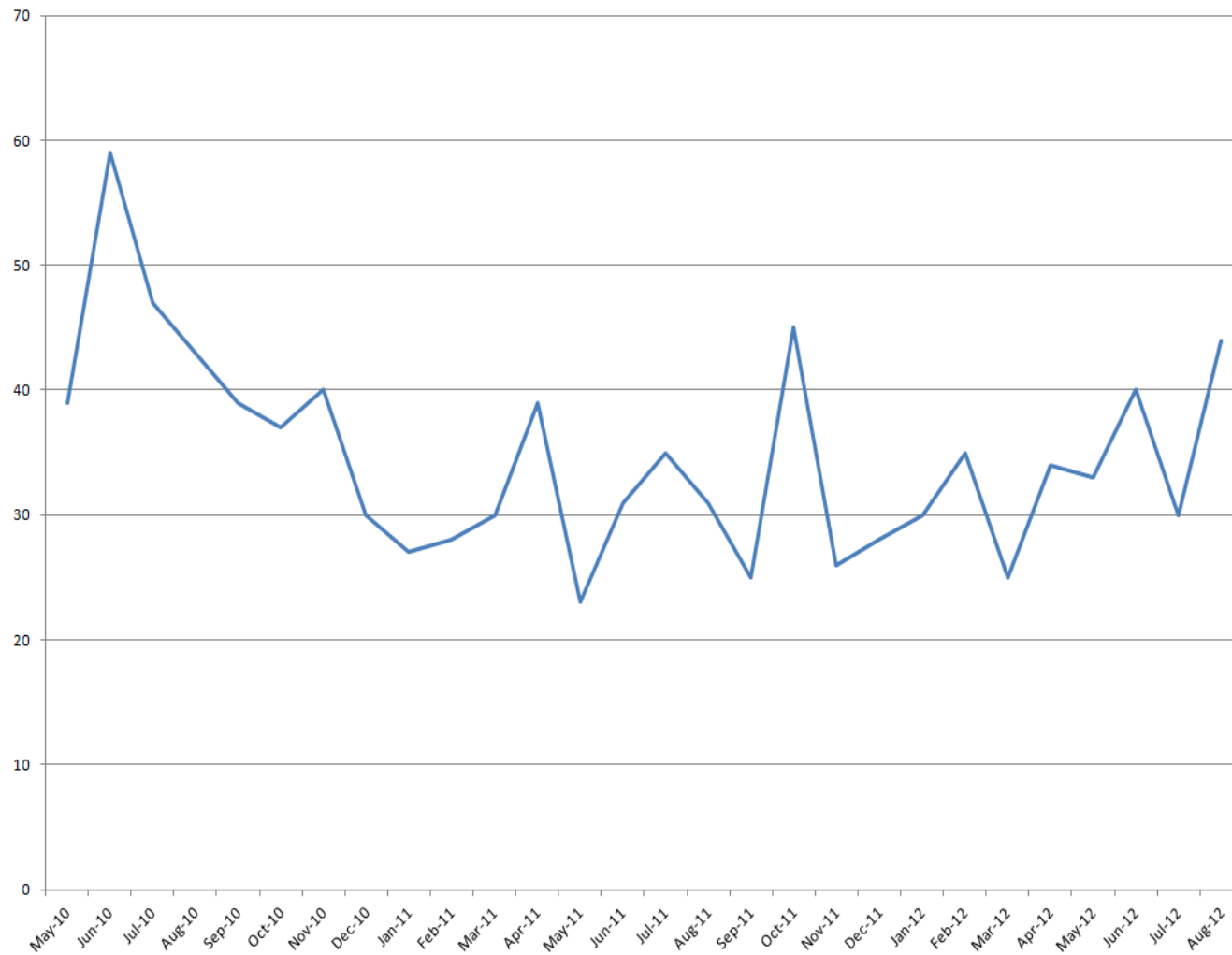


- Monthly follow-up of A/R by COD
 - GGUS tickets if site's A/R $< 70\%/75\%$. Site needs to give an explanation
 - GGUS ticket if sites availability $< 70\%$ for three consecutive months, the site qualifies for suspension.

monthly availability



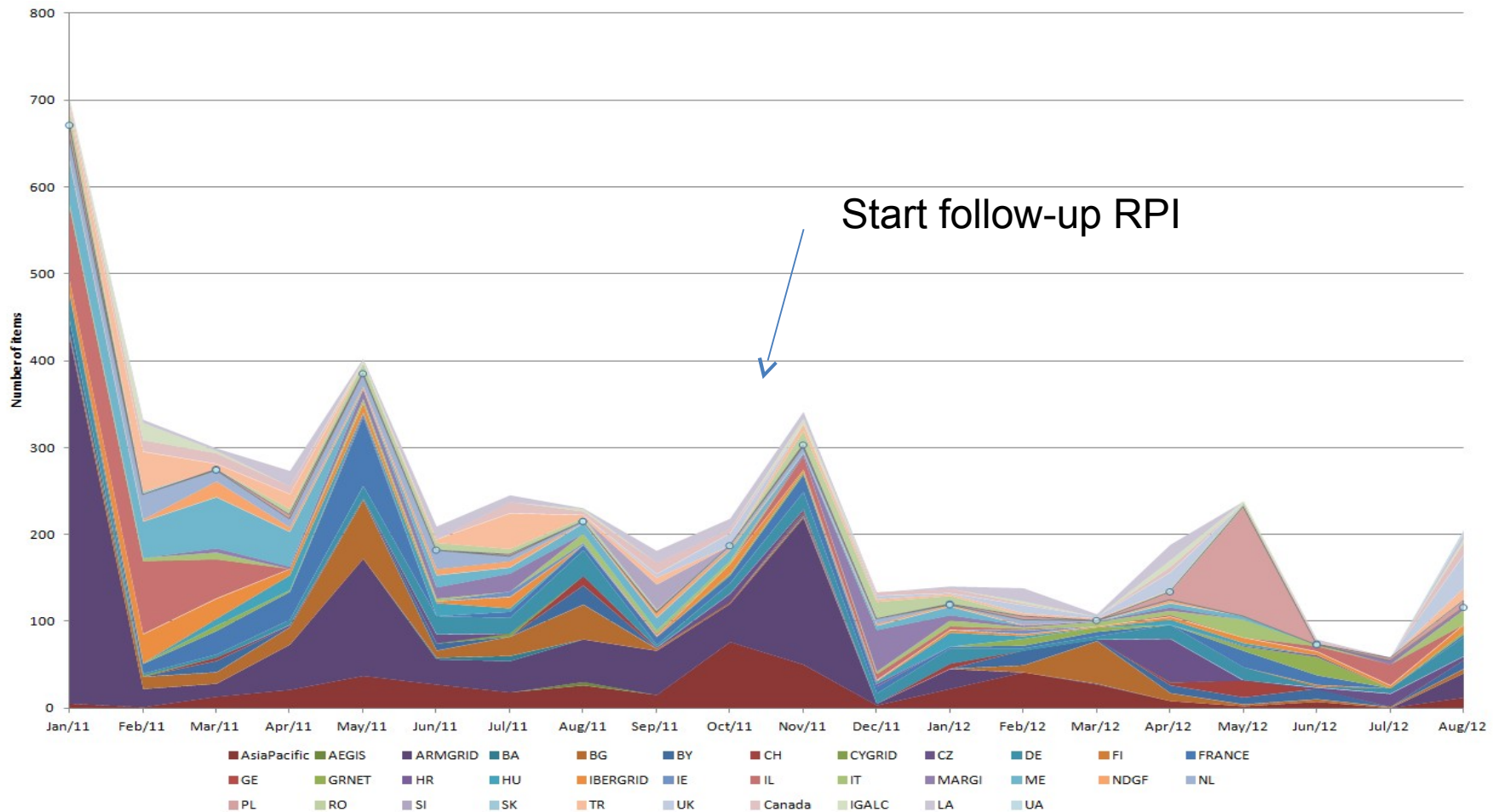
GGUS A/R tickets



- On average the availability is about 94% and the reliability is somewhat higher
 - Means that the grid is down for about 2 days every month
 - But the grid is **not** down for 2 days every month. 94% is the average availability of sites but it is **not** the availability of the Grid as a whole.
 - If the availability of the Grid is defined as the probability that the ops VO can store a file and run a job on the grid, the availability of the grid is much much higher

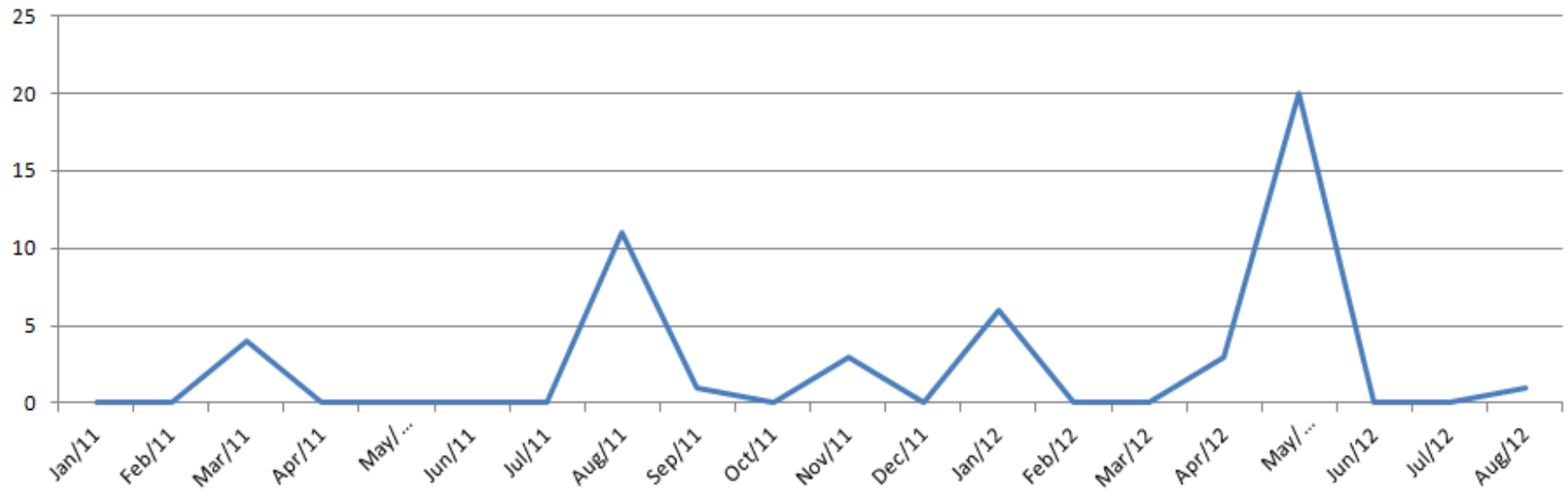
- The average availability seems to be fairly constant and the number of A/R GGUS tickets is fairly constant
- Hoped to increase the 70%/75% threshold but this is not an option.
- Questions:
 - Is the monthly follow-up of the A/R metrics beneficial?
 - If this activity is stopped, will the A/R drop?
 - Is it possible with the means that our resource centres have to increase the a/r further and if so, how?

- The number of items that will appear on the COD dashboard
 - Alarms not handled within 72 hours
 - Expired tickets
 - Tickets open for more than one month
- GGUS tickets for ROD that are above 10 in one month

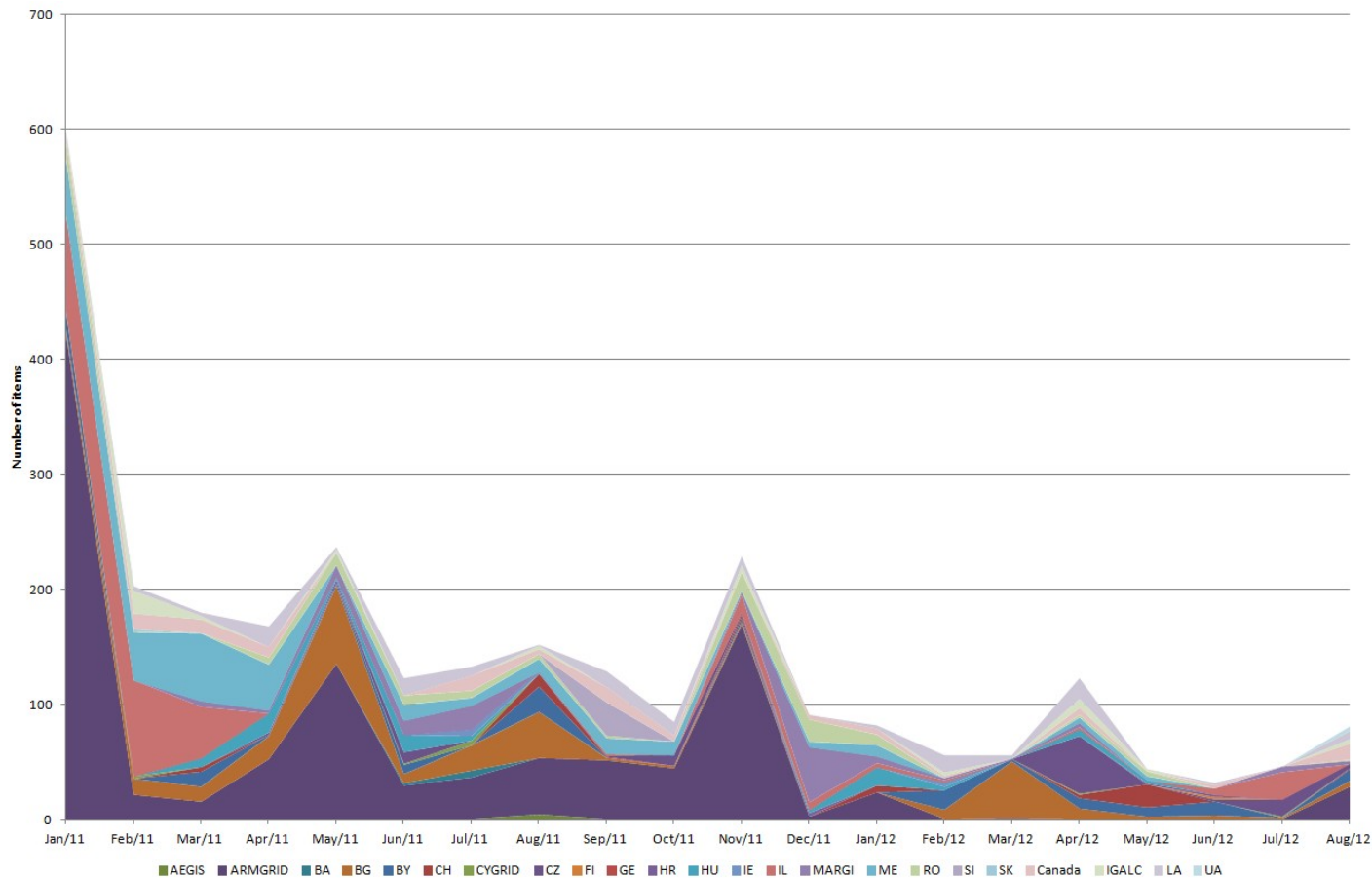


- Causes of “bad” performance
 - Holidays and in the past weekends
 - Ignored alarms
 - Problems with monitoring system
 - Regional SE down
 - Nagios problems
 - Top-BDII problems
 - Non-production service
 - These alarms should have been handled.
 - Close in nonOK status
 - Bad coordination
 - People go on holidays and forget to pass on their shift to a colleague
 - People that forgot that they were on shift

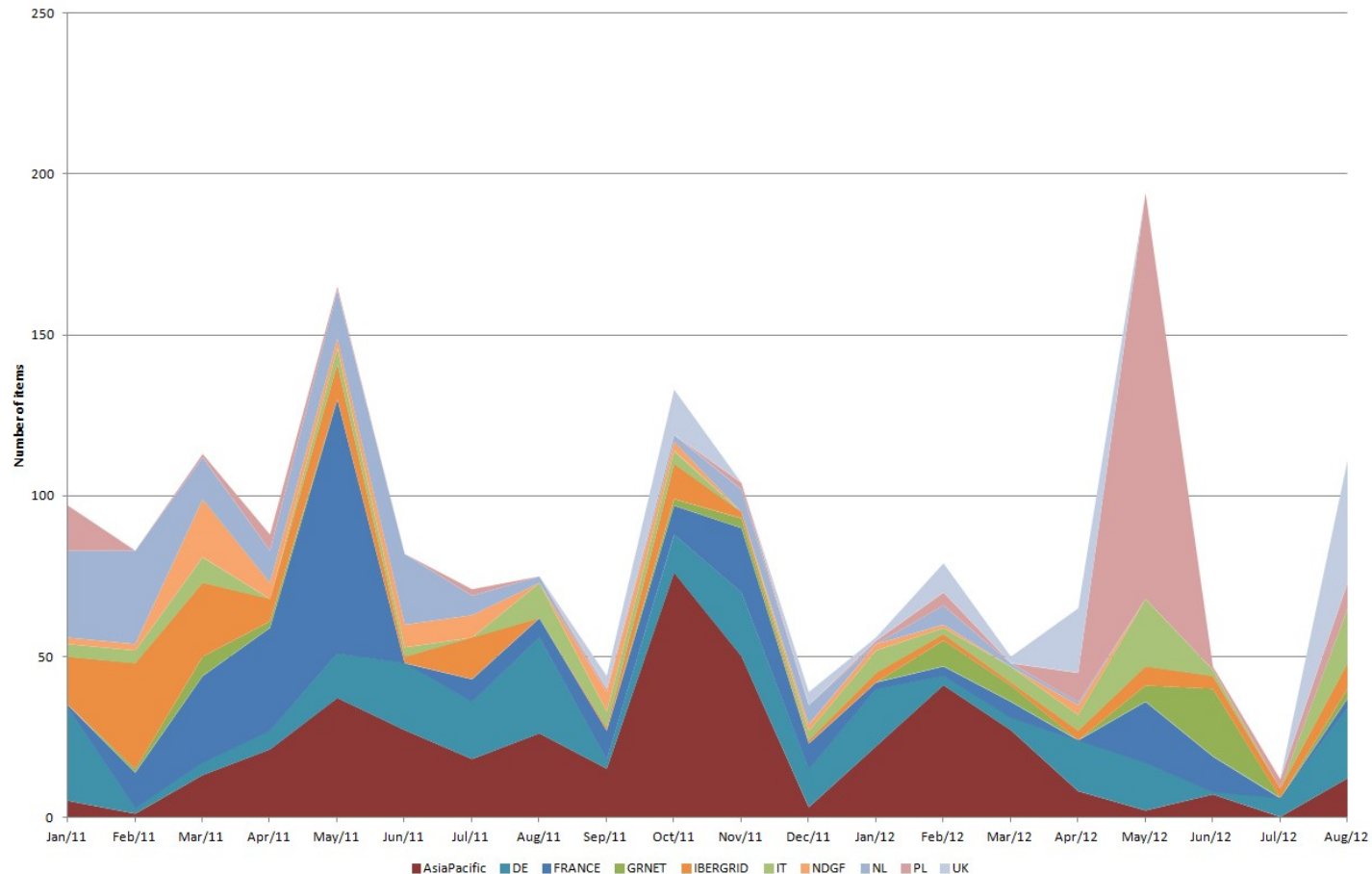
- ROD performance index of a typical ROD



- RPI new NGIs



- RPI old NGIs (former EGEE ROCs)



- Causes of “bad” performance
 - Holidays
 - Ignored alarms
 - Problems with monitoring system
 - Regional SE down
 - Nagios problems
 - Top-BDII problems
 - Non-production service
 - These alarms should have been handled.
 - Close in nonOK status
 - Bad coordination
 - People go on holidays and forget to pass on their shift to a colleague
 - People that forgot that they were on shift

- There are no real persistent issues, only transient ones
- Trend is decaying which is good
- New NGIs are doing fine

- Site certification
 - Some NGIs “certify” sites to get them to make the tests run. This is bad practice. Exposes users to sites that have problems. Bad for your NGIs A/R.
 - This is how it should go down:
 - Set the site to “uncertified”
 - Add site to your NGIs nagios and separate toplevel BDII where your Nagios looks at.
 - Site should configure this BDII in yaim
 - When OK for three days the site is certified.

- Non OK alarms
 - Should not be closed in principle and a ticket should be generated, but....
 - There are cases when it is OK to close them
 - Site in downtime
 - Some times an alarm is closed with the explanation that the BDII is broken.
 - This is not a valid reason to close an alarm

- Escalation procedure
 - Sometimes tickets opened to sites are dragging along for too long.
 - It is good to follow the escalation procedure (<https://wiki.egi.eu/wiki/PROC01>) and take care of the timing. This helps you to resolve a site issue quickly.

- The unknowns
 - Please have a look at the “Performance records/Resource centres” section of:
https://wiki.egi.eu/wiki/Availability_and_reliability_monthly_statistics





























[Performance reports](#)

[\[edit\]](#)

[Resource Centres](#)

[\[edit\]](#)

January 2008 - April 2010  (EGEE)

Avallability/Rellability	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2010	-	-	-	-	05/10 	06/10 	07/10 	08/10 	09/10 	10/10 	11/10 	12/10 
2011	01/11 	02/11 	03/11 	04/11 	05/11 	06/11 	07/11 	08/11 	09/11 	10/11 	11/11 	12/11 
2012	01/12 	02/12 	03/12 	04/12 	05/12 	06/12 	07/12 	08/12 	[09/12]	[10/12]	[11/12]	[12/12]

- The unknowns
 - Broken monitoring

95 %	95 %	23 %
79 %	79 %	12 %
100 %	100 %	22 %
87 %	87 %	22 %
98 %	98 %	22 %
76 %	76 %	26 %
93 %	93 %	22 %

- Broken site

92 %	100 %	0 %
88 %	88 %	1 %
86 %	86 %	22 %
96 %	96 %	0 %



- GGUS, COD support unit
- Email: central-operator-on-duty@mailman.egi.eu