Services reliability and availability
SAM3

Rocío Rama

IT/SDC

Grid Deployment Board

11 March 2015

Table of contents

Service Availability Monitoring (SAM)

- System to verify **status of services and sites**
  - By submitting tests at regular intervals
- Multiple **sets of tests**:
  - Job submission, Storage, worker node, cvmfs
  - Generic/Experiment specific
- Two infrastructures:
  - Production (https://wlcg-sam.cern.ch)
  - Preproduction (http://wlcg-sam-dev.cern.ch)
- ATLAS/CMS use SAM results for blacklisting sites
- Monthly WLCG Availability and Reliability reports:
  - http://cern.ch/go/m76m
- **SAM3** in production since **November 2014**

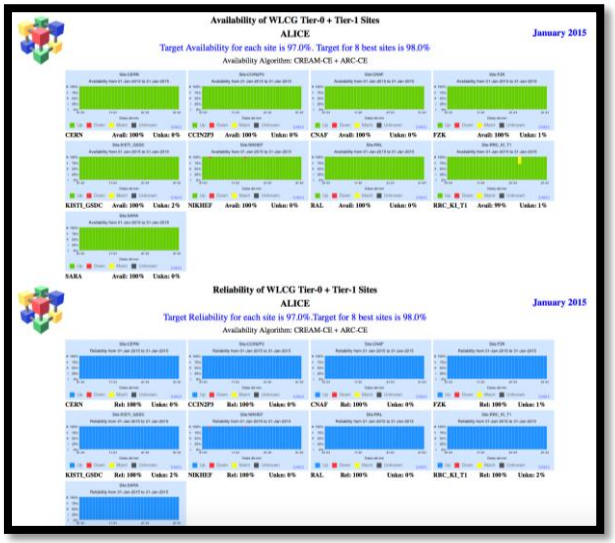
Availability/Reliability reports

- Created shortly after the end of the month
  - Draft, 10 calendar days for correction requests, final version

- **T0-T1 Summary** (1 page):
  - Availability plots for each site - Tier-0/1 sites
- **T0-T1 6 month History** (1 page)
  - Mean reliability table for 8 best sites - Tier-0/1 sites
  - Mean reliability table for all sites - Tier-0/1 sites
  - Mean reliability table each month/each site - Tier-0/1 sites
- **VOs 6 month details** (25 pages):
  - Availability/reliability plots for each site/month - Tier-0/1 sites
  - Mean availability/reliability tables for each site/day - Tier-0/1 sites
- **All Sites (T0-T1-T2)** (9 pages):
  - Availability/reliability plots for each site Tier-0/1 sites
  - 2 Tables with federation availabilities and reliabilities: sorted by name or availability)
  - Table with T2 capacities and 3 month historical availability and reliability

Report examples (January 2015)

T1

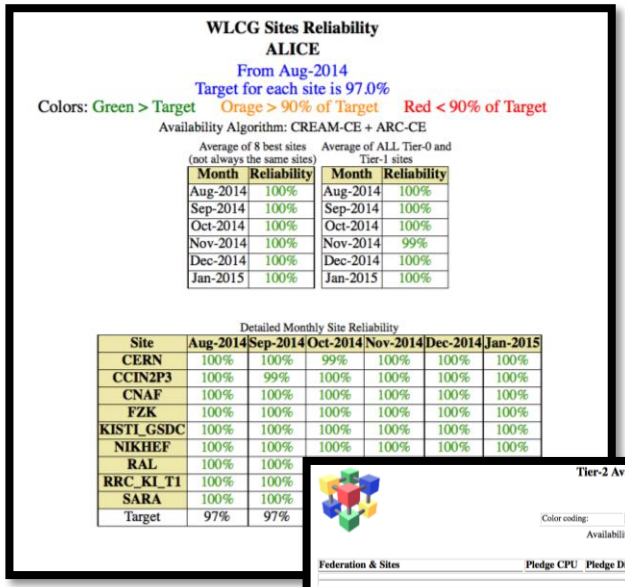


Availability of WLCG Tier-0 + Tier-1 Sites
ALICE
January 2015
Target Availability for each site is 97.0%. Target for 8 best sites is 98.0%
Availability Algorithm: CREAM-CE + ARC-CE

CERN Avail: 100% Unkn: 0% | CCIN2P3 Avail: 100% Unkn: 0% | CNAF Avail: 100% Unkn: 0% | FZK Avail: 100% Unkn: 1%
KISTI_GSDC Avail: 100% Unkn: 2% | NIKHEF Avail: 100% Unkn: 0% | RAL Avail: 100% Unkn: 0% | RRC_KI_T1 Avail: 99% Unkn: 1%
SARA Avail: 100% Unkn: 0%

Reliability of WLCG Tier-0 + Tier-1 Sites
ALICE
January 2015
Target Reliability for each site is 97.0%.Target for 8 best sites is 98.0%
Availability Algorithm: CREAM-CE + ARC-CE

CERN Rel: 100% Unkn: 0% | CCIN2P3 Rel: 100% Unkn: 0% | CNAF Rel: 100% Unkn: 0% | FZK Rel: 100% Unkn: 1%
KISTI_GSDC Rel: 100% Unkn: 2% | NIKHEF Rel: 100% Unkn: 0% | RAL Rel: 100% Unkn: 0% | RRC_KI_T1 Rel: 100% Unkn: 2%



WLCG Sites Reliability
ALICE
From Aug-2014
Target for each site is 97.0%
Colors: Green > Target Orage > 90% of Target Red < 90% of Target
Availability Algorithm: CREAM-CE + ARC-CE

Average of 8 best sites (not always the same sites)

Month	Reliability
Aug-2014	100%
Sep-2014	100%
Oct-2014	100%
Nov-2014	100%
Dec-2014	100%
Jan-2015	100%

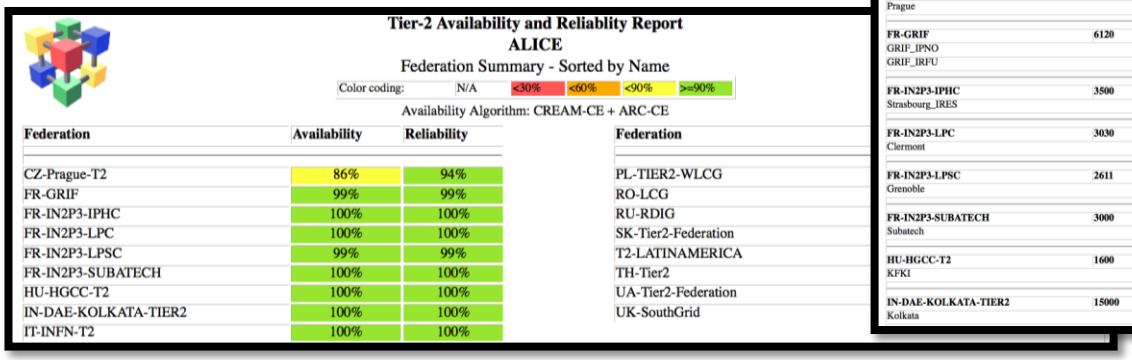
Average of ALL Tier-0 and Tier-1 sites

Month	Reliability
Aug-2014	100%
Sep-2014	100%
Oct-2014	100%
Nov-2014	99%
Dec-2014	100%
Jan-2015	100%

Detailed Monthly Site Reliability

Site	Aug-2014	Sep-2014	Oct-2014	Nov-2014	Dec-2014	Jan-2015
CERN	100%	100%	99%	100%	100%	100%
CCIN2P3	100%	99%	100%	100%	100%	100%
CNAF	100%	100%	100%	100%	100%	100%
FZK	100%	100%	100%	100%	100%	100%
KISTI_GSDC	100%	100%	100%	100%	100%	100%
NIKHEF	100%	100%	100%	100%	100%	100%
RAL	100%	100%				
RRC_KI_T1	100%	100%				
SARA	100%	100%				
Target	97%	97%				

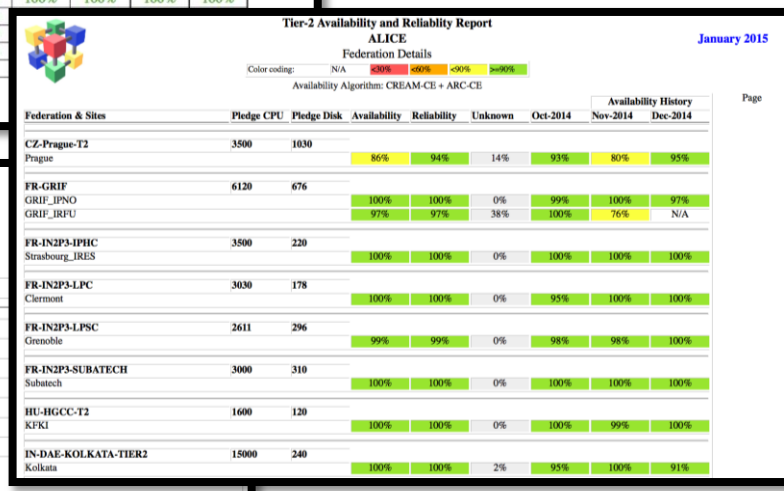
T2



Tier-2 Availability and Reliablity Report
ALICE
Federation Summary - Sorted by Name
Color coding: N/A <30% <60% <90% >=90%
Availability Algorithm: CREAM-CE + ARC-CE

Federation	Availability	Reliability
CZ-Prague-T2	86%	94%
FR-GRIF	99%	99%
FR-IN2P3-IPHC	100%	100%
FR-IN2P3-LPC	100%	100%
FR-IN2P3-LPSC	99%	99%
FR-IN2P3-SUBATECH	100%	100%
HU-HGCC-T2	100%	100%
IN-DAE-KOLKATA-TIER2	100%	100%
IT-INFN-T2	100%	100%

Federation
PL-TIER2-WLCG
RO-LCG
RU-RDIG
SK-Tier2-Federation
T2-LATINAMERICA
TH-Tier2
UA-Tier2-Federation
UK-SouthGrid



Tier-2 Availability and Reliablity Report
ALICE
Federation Details
January 2015
Color coding: N/A <30% <60% <90% >=90%
Availability Algorithm: CREAM-CE + ARC-CE

Federation & Sites	Pledge CPU	Pledge Disk	Availability	Reliability	Unknown	Oct-2014	Availability History Nov-2014	Availability History Dec-2014
CZ-Prague-T2	3500	1030						
Prague			86%	94%	14%	93%	80%	95%
FR-GRIF	6120	676						
GRIF_IPNO			100%	100%	0%	99%	100%	97%
GRIF_IRFU			97%	97%	38%	100%	76%	N/A
FR-IN2P3-IPHC	3500	220						
Strasbourg_IRES			100%	100%	0%	100%	100%	100%
FR-IN2P3-LPC	3030	178						
Clermont			100%	100%	0%	95%	100%	100%
FR-IN2P3-LPSC	2611	296						
Grenoble			99%	99%	0%	98%	98%	100%
FR-IN2P3-SUBATECH	3000	310						
Subatech			100%	100%	0%	100%	100%	100%
HU-HGCC-T2	1600	120						
KFKI			100%	100%	0%	100%	99%	100%
IN-DAE-KOLKATA-TIER2	15000	240						
Kolkata			100%	100%	2%	95%	100%	91%

Limitations of SAM2

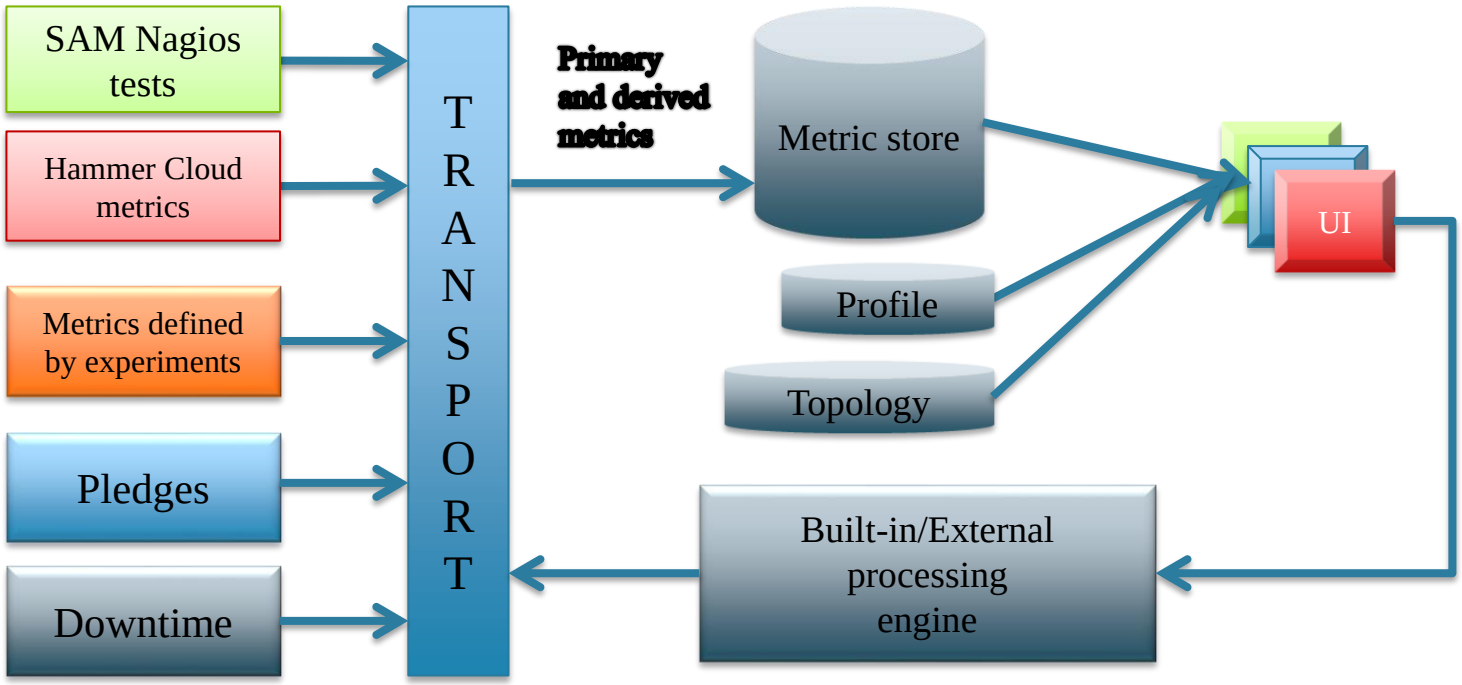
- **Functionality**:
  - Restriction on services and hosts that could be monitored
  - Limited set of options for the profile algorithm
  - Inconsistencies in site naming conventions for the reports
  - Multiple interfaces presenting the same data (SUM/mywlcg)
    - Some cases, even replicating it (importing data in SSB)
- **Design**
  - SAM2 was designed for a fully distributed operational model (EGI)
    - No clear split between different components
    - ATP/MRS/POEM/ACE implemented in a single DB schema sharing certain tables
- **Operations/Maintenance**
  - Heavy operational support.
    - 1 dedicated FTE on rota for operations, reports corrections
    - Load decreased after moving from OPS tests to the Experiment-specific tests
  - Difficult to evolve the system due to the internal dependencies

SAM3 benefits in functionality

- Experiments have **more power** (and more responsibility!)
  - Definition of sites and services
  - Definition of profiles
  - Injecting their own metrics
  - Possibility to overwrite (recalculation)
- **More flexibility** in the algorithm for profiles
  - Metrics defined at the site level
  - Combine ‘any’/’all’ services
  - Multiple roles
- Creation of reports from the UI
- For site administrators
  - Possibility to integrate in **local nagios**

SAM3 benefits in design

- Layered design
- Centralized deployment



SAM3 benefits in operations

- **Common schema** with the Site Status Board applications
- Combining different UI: MyWLCG and SUM
  - Less services to maintain
- Recalculation can be done by:
  - Experiments
  - WLCG monitoring team

SAM3 concepts (I)

- **Metrics**:
  - Measurement taken over a period of time
    - (Entity, Value, timestamp, validity)
  - Several types:
    - Status (e.g. JobSumit, status of CE at a site). Possible values: OK, WARN, CRITICAL, DOWNTIME, UNKNOWN
    - Numerical (e.g. # finished jobs per hour, pledged resources)
    - Mapping (e.g. name of the site for a given service, tier of a site)

- **Vofeed:**
  - XML provided by the experiments with the site names and services
  - Used to aggregate services into sites and naming conventions

- **Profile:**
  - Combination of metrics following an algorithm
  - One critical per experiment (used for the reports)

SAM3 concepts (II)

- **Downtime:**
  - Declaration of a site intervention
  - Entries collected from GOCDB and OIM
  - **ONLY SCHEDULED OUTAGES** are considered for Avl/Rel
- **Availability:**
  - Percentage of time that an instance is working over all known states
  - Avl = (OK +WARN) / (OK + WARN + CRITICAL + DOWNTIME)
- **Reliability:**
  - Percentage of time in working state over all known states except downtime
  - Rel= (OK +WARN) / (OK + WARN + CRITICAL)

SAM3 Recomputations

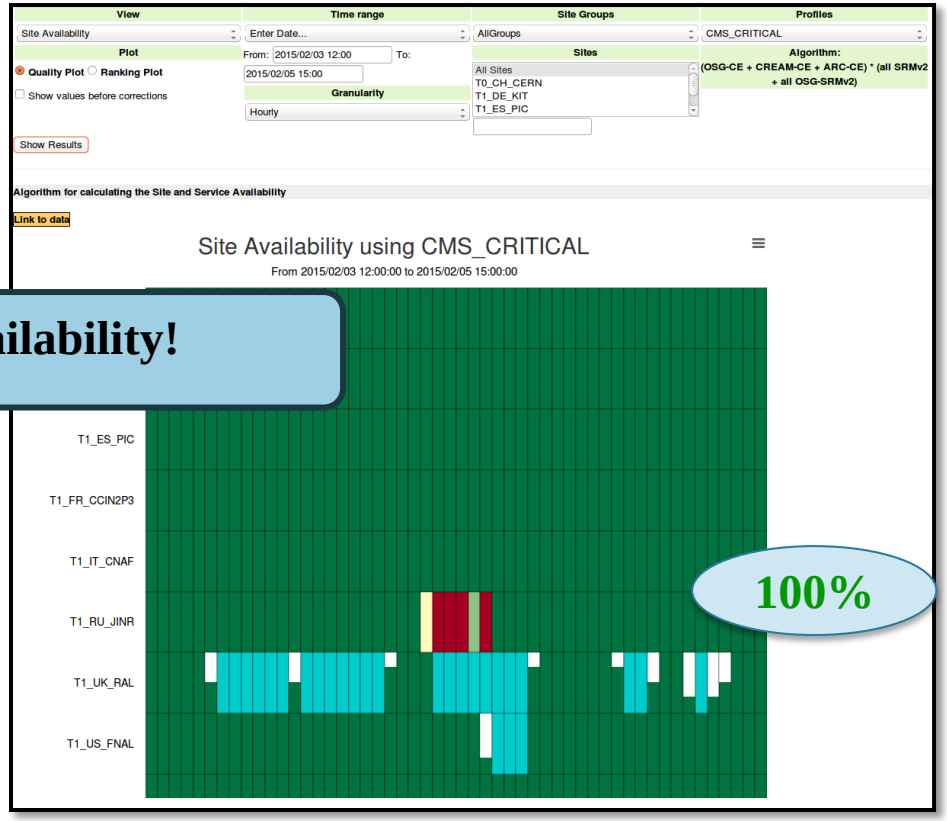
- Profiles can be corrected by the privileged users in the experiments or in the WLCG monitoring team.
  - https://twiki.cern.ch/twiki/bin/view/ArdaGrid/ProfileCorrections

- Advantages:
  - Adjust manually the availability of sites
    - For non site-related problems
  - Record of before/after recomputations
  - Can be visualized at any moment

SAM3 Recomputations

Before

After



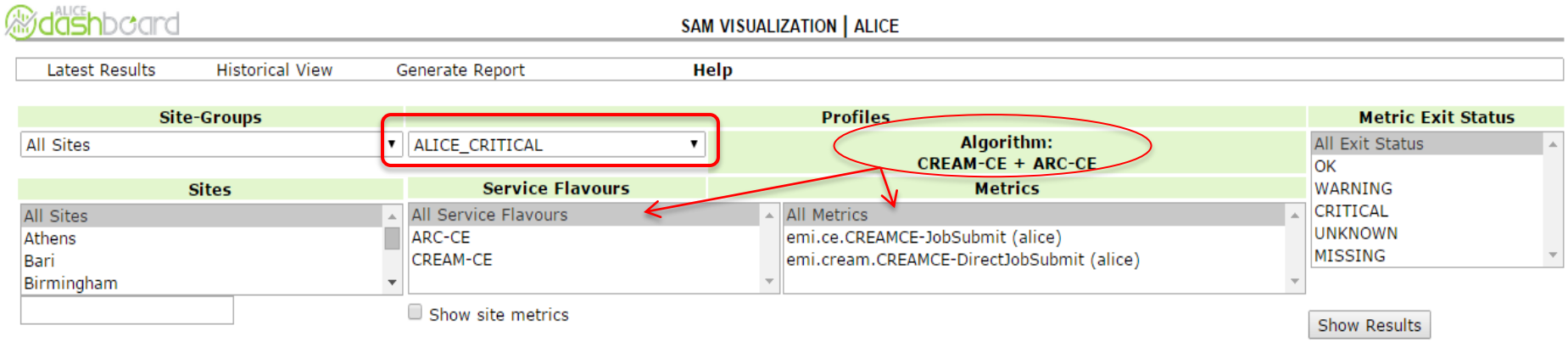
Adjust Availability!

Critical **OK**

**Blue – Unknown

Example of new SAM3 functionalities

ALICE_CRITICAL



http://wlcg-sam-alice.cern.ch/

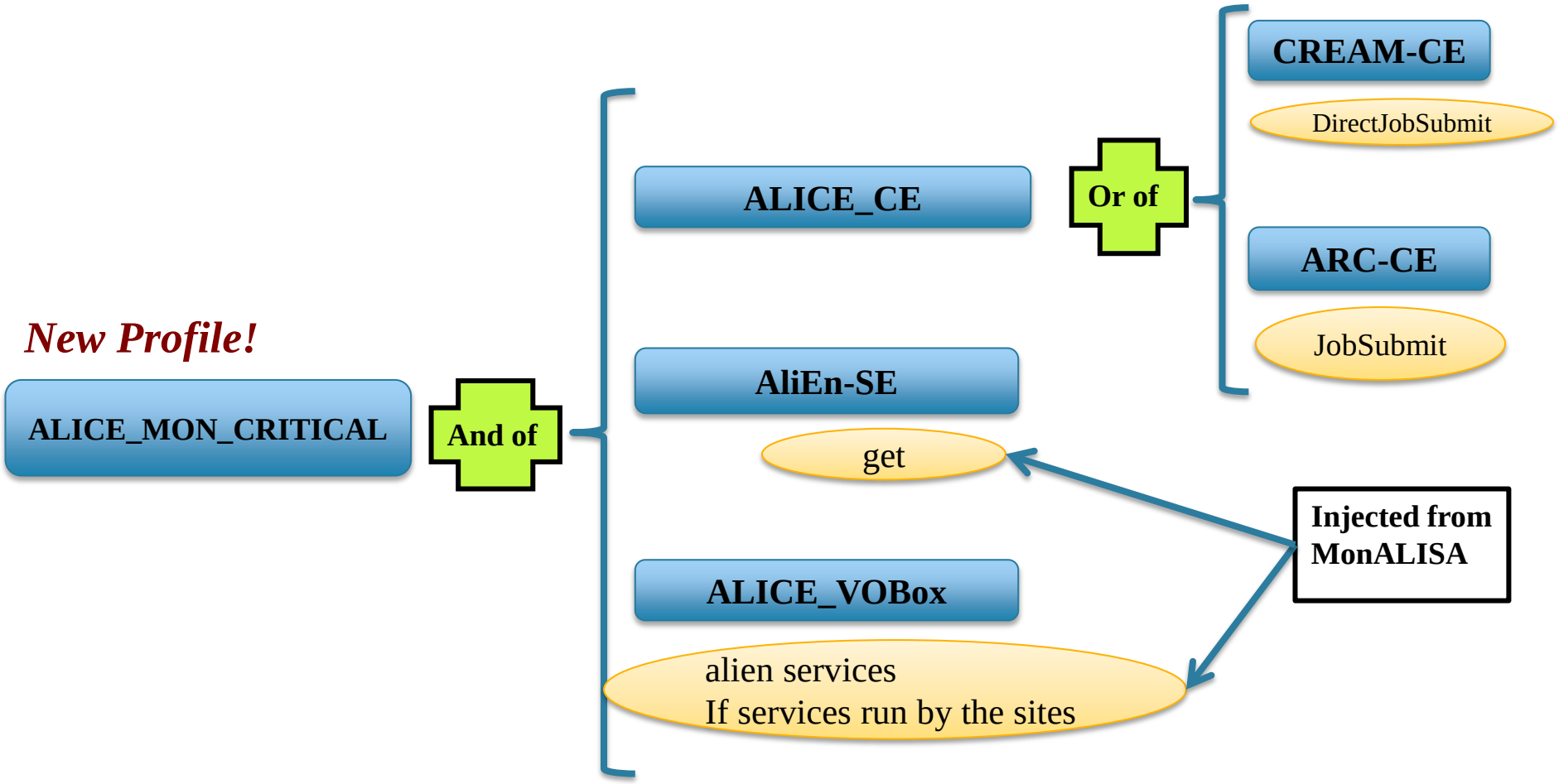
ALICE_CRITICAL issues

- Critical profile: ARC-CE OR CREAM-CE
  - Sites without those services do not appear (at the moment, 19 sites)
  - **Solution:** include AliEn-CE or VOBox in algorithm
- Storage ignored from availability
  - **Solution:** include AliEn-SE in algorithm
- NDGF does not appear as a single T1
  - The individual sites appear as 'NorduGrid Sites',
  - **Solution:** modify Vofeed to define NDGF-T1
- OSG sites do not appear in the reports
  - Since OSG sites do not run the services of the current 'critical_profile' they were not marked as T2 to avoid having them appear with no data
  - **Solution:** after the AliEn-CE or VOBox is included in the algorithm, mark them back as a T2

Slides presented in the "ALICE Tier-1/Tier-2 Workshop 2015" in Torino.

The sites aware about it

ALICE_MON_CRITICAL



ALICE_MON_CRITICAL

- Advantages:
  - Sites without the services ARC-CE or CREAM-CE appear
    - Algorithm VOBox included
  - Storage from availability is not ignored
    - Including AliEn-SE
  - NDGF and OSG sites appear
    - Algorithm VOBOx included
- Both profiles in parallel until end of March
- **Beginning of April** new profile will be used for the reports

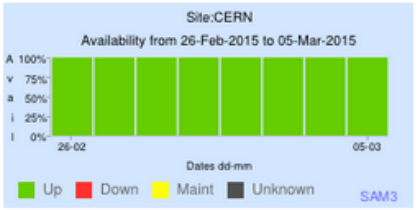
Reports ALICE_MON_CRITICAL

Availability of WLCG Tier-0 + Tier-1 Sites

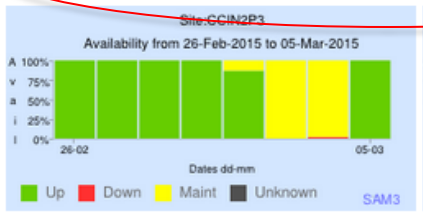
ALICE **March 2015**

Target Availability for each site is 97.0%. Target for 8 best sites is 98.0%

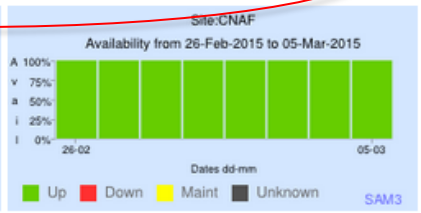
Availability Algorithm: @ALICE_CE * @ALICE_VOBOX * AliEn-SE



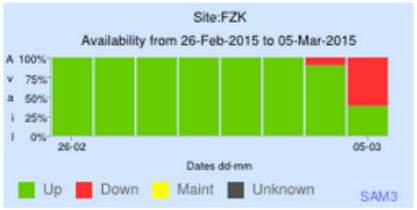
CERN **Avail: 100%** **Unkn: 6%**



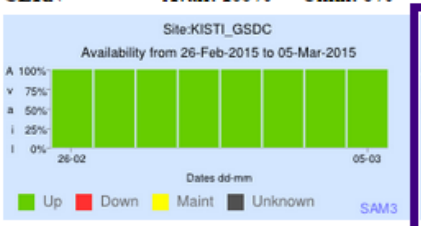
CCIN2P3 **Avail: 73%** **Unkn: 16%**



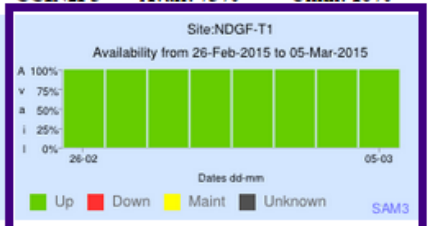
CNAF **Avail: 100%** **Unkn: 6%**



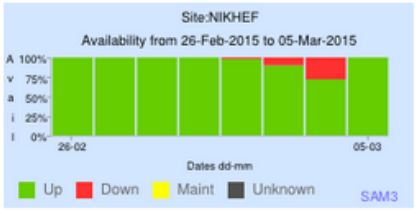
FZK **Avail: 91%** **Unkn: 6%**



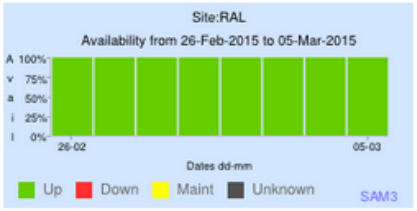
KISTI_GSDC **Avail: 100%** **Unkn: 6%**



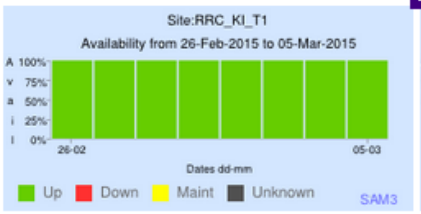
NDGF-T1 **Avail: 100%** **Unkn: 6%**



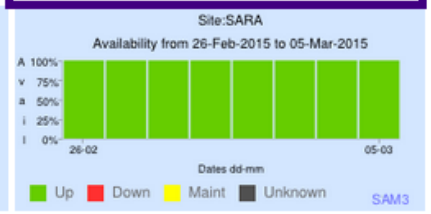
NIKHEF **Avail: 95%** **Unkn: 0%**



RAL **Avail: 100%** **Unkn: 6%**



RRC_KI_T1 **Avail: 100%** **Unkn: 6%**



SARA **Avail: 100%** **Unkn: 6%**

New profile ATLAS_AnalysisAvailability

- ASAP (ATLAS Site Availability and Performance):
  - New metric defined by ATLAS for Run II
  - Measure the availability for all Tiers
  - Simpler and easier to maintain
    - Analyzing the data during the **last two hours**
    - All measurements are averaged over certain period of time
- Metric injected into SAM3, and available for reports

http://cern.ch/go/H6CH

ATLAS_CRITICAL and ATLAS_AnalysisAvailability

Profile Name	Algorithm	Flavour Metrics
ATLAS_AnalysisAvailability	@SiteAvailability	SiteAvailability: ASAP ;
ATLAS_CRITICAL	(OSG-CE + CREAM-CE + ARC-CE) * (&SRMv2 + &OSG-SRMv2)	CREAM-CE: org.sam.CONDOR-JobSubmit (/atlas/Role_lcgadmin),org.atlas.WN-swspa SRMv2: org.atlas.SRM-VOGet (/atlas/Role_production),org.atlas.SRM-VOPut (/atlas/Ro OSG-SRMv2: org.atlas.SRM-VOPut (/atlas/Role_production),org.atlas.SRM-VOGet (/atl OSG-CE: org.atlas.WN-swspace (/atlas/Role_lcgadmin),org.atlas.WN-swspace (/atlas/R ARC-CE: org.atlas.WN-swspace (/atlas/Role_pilot),org.sam.CONDOR-JobSubmit (/atlas

- Discussed and agreed at the ATLAS International Computing Boards during 2014, reported in ATLAS Jamboree in December and ATLAS SW week in January
- Sites below 80% of ASAP over the past 3 months are not effective for ATLAS
  - These sites will be contacted by the ATLAS ICB representatives
- Start categorizing the sites as from the 1st of January 2015

Future Developments

- Investigate NoSQL storage
  - ElasticSearch, HADOOP
- Combine data from different SSB instances
- More flexibility in algorithms
  - “Not” operator
  - Numerical metrics
    - Number of completed jobs > threshold
  - Different weight for components:
    - 80% CE + 20% SRM
- Create reports from any profile

Conclusion

- SAM calculates the site and service availability and reliability
- SAM3 offers new functionality and flexibility
  - Definition of algorithm
  - Metrics and topology defined by experiments
  - More responsibility to the experiments!
- ATLAS and ALICE started to introduce new profiles

- We encourage sites and VOs to explore new functionalities and to provide your feedback!

Combination of metrics

OPERATION	PRIORITY
OR	
AND	
AND IF DATA	
<m> OVERWRITE <n>	IF <N> !=  THEN <N> ELSE <M>
ANY <m>, <t>	OR of all instances in <m> that have the same value in <t>
ALL <m>, <t>	AND IF DATA of all instance in <m> that have the same value in <t>
FILTER <m>, <t>='v'	Take only the instances of <m> that have a value of 'v' in metric <t>

OK WARNING CRITICAL DOWNTIME UNKNOWN D DATA

* Will be converted to if there is at least one more metric