Computation of Service Availability Metrics in ACE

Rajesh Kalmady, Phool Chand, Kumar Vaibhav, Vibhuti Duggal, Kislay Bhatt, Digamber Sonvane

Computer Division, BARC, India

Version	Date	Comments
0.1	28.06.2011	First Draft

ACE's service availability module computes status, availability and reliability metrics of grid entities such as sites and services. These computed metrics are displayed in the Gridview/ACE frontend and other tools such as MyWLCG/MyEGI in the form of graphs and charts, with the ability for a user to drill-down from the availability of a site to individual test results that contributed to the computed figure.

1. Terms and Definitions:

This section describes certain terms which occur frequently in this document.

a) Service: A service is a single resource in the grid, such as a particular compute element or a storage element in some site. It should be noted that the service refers to the service endpoint rather than to a physical machine - it is quite possible that the same physical server run multiple services.

b) Service Flavour: A service flavour is a set of similar services. For example, multiple CREAM CEs in a site together make up the CREAM CE service flavour for the site. A service flavour can be made up of one or more services.

c) Site: A site in the grid is a collection of several service flavours such as LCG-CE, CREAM-CE, SRM, sBDII and so on. A site can be made up of one or more service flavours. A service flavour in a site can be made up of many similar services.

d) Status: Status of a service, service flavour or a site is the status of that entity at a given point in time. Possible status values are
   - 'OK' : The service, service flavour or site is working
   - 'WARNING' : The service, service flavour or site is working, but with warnings
   - 'CRITICAL' : The service, service flavour or site is not working
   - 'UNKNOWN' : The status cannot be computed because tests return 'UNKNOWN' status
   - 'MISSING' : The status cannot be computed because tests could not run or the validity period of the earlier test expired.
   - 'REMOVED': The service no longer exists

   These status values are mutually exclusive. The status of an entity can have only one value at a given point in time.

e) Scheduled Downtime: A service or a site could be declared to be in scheduled downtime during a specific period. The period of scheduled downtime is considered during computation of availability and reliability.

f) Service Metric: A Service Metric is a functional test for a given service flavour. Each service flavour has a set of service metrics that verify its functionality and

performance. Service Metrics are generated when Nagios tests are run on a particular service.

g) Profile: A profile defines which service metrics are to be considered to compute the status of a service of a particular flavour. Profiles are defined by individual VOs. A VO can define multiple profiles. Profiles also contain algorithms for computing statuses of service flavours and sites.

h) Algorithm: An Algorithm defines how to aggregate service statuses into service flavour status and how to aggregate service flavour statuses into site status. A VO can define several algorithms. Every profile has an associated algorithm.

i) Metric Result Store (MRS): It is the repository of service metrics that are generated in different sites by Nagios tests. These service metrics are transmitted to the central MRS repository using the messaging system. MRS also computes the status of individual services using these service metrics.

2. Availability and Reliability

Total Period (Total Time) : Total period is the entire period over which the metrics are being computed.

UP period (Uptime) : UP period is the period over which the status of the entity was either OK or WARNING and the entity was not in scheduled downtime. For availability computation status WARNING is considered as good as OK.

Up fraction : Up fraction is the fraction of the total time the service was in the UP period.

Up fraction = Up period / Total Period

Down Period (Downtime) : Down period is the period over which the status of the entity was CRITICAL and the entity was not in scheduled downtime.

Down fraction : Down fraction is the fraction of the total time the service was in the Down Period.

Down fraction = Down period / Total Period

UNKNOWN period : The time interval over which the status of the entity was either 'UNKNOWN' or 'MISSING' and the entity was not in scheduled downtime.

UNKNOWN fraction : UNKNOWN fraction is the fraction of the total time the status of the service was in the UNKNOWN period.

UNKNOWN fraction = UNKNOWN period / Total Period

Scheduled Down Period (Scheduled Downtime) : Scheduled Down period is the period over which the entity was declared to be in Scheduled Downtime.

Scheduled Down fraction : Scheduled Down fraction is the fraction of the total time the service was in the Scheduled Down Period.

Scheduled Down fraction = Scheduled Down period / Total Period

KNOWN period : The time interval over which the status of the entity was known during the given period. It indicates the accuracy of the computed availability and reliability metrics.

Total period = (Up period + Down period + Scheduled Down period + Unknown period).

Known period = Total period – Unknown Period
= Up period + Down period + Scheduled Down period

For example, consider the status of a service was in the UP period for 15 minutes, DOWN period for 15 minutes, Scheduled DOWN for 15 minutes and was in the unknown period for 15 minutes in a particular hour.

Here, Total Period = 60 minutes
Known period = 45 minutes

Up period = 15 minutes
Up fraction = 15/60 = 0.25

Down period = 15 minutes
Down fraction = 15/60 = 0.25

Scheduled Down period = 15 minutes
Scheduled Down fraction = 15/60 = 0.25

Unknown period = 15 minutes
Unknown fraction = 15/60 = 0.25

Availability : Availability of a service instance, service or a site over a given period is defined as the fraction of time the same was in the UP Period during the known interval in the given period.

Availability = UP period / KNOWN period
= UP period / (Total period – UNKNOWN period)

Here divide the numerator as well as denominator by Total period, to derive the Availability formula for fractions,

= (UP period/Total period) / (1 – (UNKNOWN period/Total period))

Availability = Up fraction / (1 – UNKNOWN fraction)

Equivalently

Availability = Uptime / (Uptime + Downtime + Scheduled Downtime)
Or
Availability= Up fraction / (Up fraction + Down fraction + Scheduled Down fraction)

Availability is undefined (Not Applicable) if the known period is zero (that is if the status over the entire period is UNKNOWN).

Reliability : Reliability of a service instance, service or a site over a given period is defined as the ratio of the time interval it was in the UP Period over the time interval it was supposed (scheduled) to be UP during the known interval in the given period.

Reliability = UP period / (KNOWN period – Scheduled Downtime)
= UP period / (Total period – UNKNOWN period – Scheduled Downtime)

Here divide the numerator as well as denominator by Total period, to derive the Reliability formula for fractions,

= (UP period/Total period) / (1 – UNKNOWN period/Total period – Scheduled Downtime/Total period)

Reliability = Up fraction / (1- Scheduled Down fraction – UNKNOWN fraction)

Equivalently

Reliability = Uptime / (Uptime + Downtime)
Or
Reliability = Up fraction / (Up fraction + Down fraction)

Reliability is undefined (Not Applicable) if the status of the entity is either Scheduled Down or Unknown over the entire period.

The concept of availability and reliability can be illustrated in the following example:

In a particular hour, a service instance was under maintenance (Scheduled Down) for 15 minutes, with 'unknown' status for 15 minutes, 'UP' for 15 minutes and 'DOWN' for 15 minutes. Then for the above service:

Availability = (UP period)/(Total period – unknown period)
= 15/(60 – 15) = 33 %
Reliability = (UP period)/(total-scheduled_down-unknown)
= 15/(60-15-15) = 50 %

In other words, reliability is the ratio of the Uptime of a service to the time the service was scheduled to be UP. For computing the amount of time the service was scheduled to be UP, the time for scheduled downtime and unknown period are deducted from the total time.

It follows from this that if a service, service flavour or a site is in the Scheduled Down and/or Unknown periods for the entire period (Scheduled Down fraction + Unknown fraction = 1), then Reliability is undefined (Not Applicable) for that period. For example, in a particular hour, if the service is scheduled down for 30 minutes and of unknown status for 30 minutes,

Availability = 0/30 = 0 %
Reliability = 0/(60-30-30) = 0/0 (undefined)

In ACE, we first compute and store the 4 basic metrics which are UP fraction, Down Fraction, Scheduled Down fraction and Unknown fraction and then derive Availability and Reliability figures using them. Availability and Reliability figures are computed for different periodicities such as hourly, daily, weekly and monthly. The KNOWN period indicates the accuracy of the computed Availability and Reliability figures.

3. Metrics being computed

The following metrics are being computed:

1. Individual Service Status, Availability and Reliability
2. Individual Service Flavour Status, Availability and Reliability
3. Individual Site Status, Availability and Reliability

4. Computation of Status, Availability and Reliability Metrics

The status of services, service flavours and sites is computed on a continuous time scale for each profile using service metrics. Availability metrics like Uptime, Downtime, Scheduled Downtime, Unknown period, Availability and Reliability for services, service flavours and sites are computed on an hourly basis for each profile. These hourly metrics are later used to generate metrics for different periodicities such as daily and monthly.

4.1 Computation of Service Status: 'ServiceStatus'

This metric indicates the status of a particular service, for example, a specific CE, over a period of time. The responsibility of computation of this status lies with the Metric Result Store (MRS). This status is determined by using results of all the service metrics that were run on the particular service. The list of service metrics to be used for the service is defined by each VO in the profile. Using all these input parameters, MRS derives the status of the service as below:

If (service is no longer present)
Then
Service Status = REMOVED

Else If (at least one service metric is CRITICAL)
Then
    Service Status = CRITICAL
Else If (at least one service metric is missing (expired but not updated))
Then
    Service Status = MISSING
Else If (at least one service metric is in 'UNKNOWN' status)
Then
    Service Status = UNKNOWN
Else If (at least one service metric is in 'WARNING' status)
Then
    Service Status = WARNING
Else
    /* It means all service metrics are OK */
    Service Status = OK

4.2 Computation of Service Flavour Status: 'ServiceFlavourStatus'

This metric is computed by combining status of all production services of a particular service flavour (computed in 6.1), for example, all individual Cream CEs in a site, at a given point in time. The operation that is to be used in combining the statuses is defined by the VO in the profile. There are 2 operations that are supported : OR and AND.

Service Flavour status using OR operation is derived as below :

If (All services are OK)
Then
    Service Flavour Status = OK
Else If (at least one service status is WARNING)
Then
    Service Flavour Status = WARNING
    /* One or more (but not all) services are OK */
Else If (at least one service status is CRITICAL)
Then
    Service Flavour Status = CRITICAL
Else If (at least one service status is UNKNOWN)
Then
    Service Flavour Status = UNKNOWN
Else
    Service Flavour Status = MISSING

Service Flavour status using AND operation is derived as below :

If (at least one Service Status is CRITICAL)

Then
Service Flavour Status = CRITICAL
Else If (at least one Service Status is MISSING)
Then
Service Flavour Status = MISSING
Else If (at least one Service Status is 'UNKNOWN')
Then
Service Flavour Status = UNKNOWN
Else If (at least one Service Status is 'WARNING')
Then
Service Flavour Status = WARNING
Else
/* It means all Service statuses are OK */
Service Flavour Status = OK

Services in 'REMOVED' status are not considered for computation of service flavour status.

4.3 Computation of Site Status: 'SiteStatus'

This metric is computed by combining the status of all service flavours (computed in 4.2) provided by the site at a given point in time. There are several possible ways by which service flavours could be combined in order to derive the site status. The actual algorithm for performing the computation is defined by the VO in the profile. Examples of such algorithms are given below:

- CE AND sBDII AND SRM
  - Perform AND operation of statuses of CE, sBDII and SRM
- (CE OR CREAMCE OR ArcCE) AND SRM AND sBDII
  - Perform OR of all CE flavours and then AND it with SRM and sBDII
- CE OR sBDII OR SRM
  Perform OR of statuses of CE, sBDII, SRM

Such complex expressions using AND and OR operators are supported by ACE and VOs can define the type of computation they want for calculating site status.

The AND operation of 2 statuses is described below:

If (at least one Status is CRITICAL)
Then
Result Status = CRITICAL
Else If (at least one Status is MISSING)
Then
Result Status = MISSING
Else If (at least one Status is 'UNKNOWN')

```
Then
        Result Status = UNKNOWN
Else If (at least one Status is 'WARNING' )
Then
        Result Status = WARNING
Else
        /* It means all statuses are OK */
        Result Status = OK
```

Similarly, the OR operation is described below:

```
If ( All  statuses are OK )
Then
        Result Status = OK
Else If ( at least one status is WARNING )
Then
        Result Status = WARNING
        /* One or more (but not all) statuses are OK */
Else If ( at least one status is CRITICAL )
Then
        Result Status = CRITICAL
Else If ( at least one status is UNKNOWN )
Then
        Result Status = UNKNOWN
Else
        Result Status = MISSING
```

In order to compute the site status, the expression defined in the algorithm is evaluated using the AND and OR operations described above.

4.4 Hourly Availability and Reliability Computation

- The Hourly Up fraction, Scheduled Down fraction, and Unknown fraction for a service instance, a service and a site are computed on an hourly basis from the respective status information and scheduled downtime information for that hour.
- The Hourly Availability and Reliability of a service instance, a service and a site are computed on an hourly basis from the Up fraction, Scheduled Down fraction, and Unknown fraction for that hour.

4.5 Computation of availability and reliability for higher periodicities

- The Daily, Weekly and Monthly Up fraction, Scheduled Down fraction, and Unknown fraction figures are computed from the corresponding Hourly figures by averaging over the required time periods.
- The Daily, Weekly and Monthly Availability and Reliability figures are computed directly from the Up fraction, Scheduled Down fraction, and Unknown fraction figures over the corresponding periods and **not** by averaging the Hourly Availability or Reliability figures.

This may look strange but it can be easily seen from the example below that it is not possible to obtain availability or reliability figures for higher periodicities by averaging hourly availabilities or reliabilities.

Consider a service which has the following figures for a day of 24 hours:

- Hours 00-12 : 100 % UP, 0% down, 0% scheduled down, 0% unknown in each hour
    - For these 12 hours, hourly availability = 100 %, hourly reliability = 100%
- Hours 13-24 : 0% UP, 10% down, 40% scheduled down, 50% unknown in each hour
    - For these 12 hours, hourly availability = 0%, hourly reliability = 0%

It can be seen that the service status was up for 12 hours, down for 1.2 hours, scheduled down for 4.8 hours and unknown for 6 hours in the whole period of 24 hours.

Daily availability for the service for this day = 12/(24 - 6) = 67 %
Daily reliability for the service for this day = 12/ (24 - 4.8 - 6) = 12/13.2 = 91%

If we compute daily Availability by averaging the hourly availabilities for the 24 hours, we will obtain a daily Availability figure of (100*12 + 0*12) / 24 = 50 % !

Similarly, if we compute daily reliability by averaging the hourly reliabilities for the 24 hours, we will obtain a daily reliability figure of (100*12 + 0*12) / 24 = 50 % !

4.6 Computation of Overall Availability and Overall Reliability for Federations/Regions/NGIs

The overall (aggregate) Availability / Reliability for a Federation / Region / NGI is computed by calculating a weighted average of the Availabilities / Reliabilities of all the sites belonging to the Federation / Region / NGI. The weight factor is the installed capacity of a site as denoted by the HEPSPEC06 benchmark. This is done in order to have the sites' contribution to the Availability / Reliability number of the Federation / Region / NGI proportionate to its size that is the installed capacity it provides. The installed capacity of a site is obtained from sources such as GridMap and REBUS, which in turn compute it from the number of logical CPUs and the HEPSPEC06 rating of each cpu, as published by the site in the BDII.

If a site doesn't publish the installed capacity, it is taken as 1 for computing the averages. So if all the sites in a particular Federation / Region / NGI do not publish the numbers, the weighted averaging is same as simple average. On the other hand, if few sites in the Federation / Region / NGI publish the numbers and others don't then, the contribution of the sites not publishing the numbers to the Federation / Region / NGI overall numbers will be negligible.

5. Profile used for official Availability Reports

ACE reports are being used as the official Availability/Reliability reports for WLCG and EGI sites from May 2011 for OPS VO. The OPS Profile that is being used for the reports is called 'WLCG_CREAM_LCGCE_CRITICAL'. This profile is defined by OPS VO and has the following features:

- LCG CE, CREAM CE, ArcCE, OSGCE, SRM and sBDII are used to compute availabilities
- Status of multiple services of the same flavour in a site are ORed to derive the service flavour status. For example, if there are two CREAM CEs in a site, the status of the CREAM CE service flavour is the OR of the statuses of the individual CREAM CEs.
- Site status is derived by the following expression:
  (OSG-CE or CREAM-CE or CE or ARC-CE or gLite-CE) and (OSG-SRMv2 or OSG-BestmanXrootd) and OSG-GridFtp and (SRM or SRMv2) and sBDII

It can be seen from the above expression that the algorithm aggregates similar service flavours together with an OR operation and then performs AND across different service types.

It should also be noted that if a site does not provide a service that is listed in the expression, the service is ignored during the status computation. For example, if a site provides only ArcCE, CREAM CE and sBDII, the expression is evaluated with only these services. The other services such as OSGCE are removed from the expression for that site.