

Référence	DSP-CSPIT-OSI-SCIENTIFIQUE-OCRI
<div>Description des indicateurs d'OCRI</div> <div>Date d'application : 01 septembre 2013</div>	

PROJET
DÉFINITIF

<input checked="" type="checkbox"/>
<input type="checkbox"/>

Résumé Le présent document contient une description succincte du fonctionnement des indicateurs utilisés dans OCRI.

Rédacteur	Approbateurs
Arnaud Le Gué (OBS)	Mohamed OUADAH (EDF Pexsi R & D)

Accessibilité

Libre
EDF - Gaz de France
EDF
Gaz de France
Diffusion restreinte
Confidentiel

<input type="checkbox"/>
<input type="checkbox"/>
<input checked="" type="checkbox"/>
<input type="checkbox"/>
<input type="checkbox"/>
<input type="checkbox"/>

PAGE DE SERVICE

Documents associés nécessaires à l'utilisation du présent document

Référence	Titre

Documents annulés par le présent document

Référence	Titre

Liste des modifications du document par rapport à la version précédente

Destinataires du document

Entités	Noms
EDF OSI R&D	Mohamed OUADAH

Lieu de stockage du document

Dernière mise à jour :

--

Sommaire

GENERALITES	5
TAUX D'OCCUPATION	5
SUIVI JOURNALIER REEL OU THEORIQUE DU TAUX D'OCCUPATION DES CPUS	5
Paramètres	5
Détails des requêtes vers la base de données	5
NŒUDS LES PLUS UTILISES	6
Paramètres	6
Détails des requêtes vers la base de données	6
SUIVI JOURNALIER DE LA CONSOMMATION CPU PAR PARTITION	6
Paramètres	6
Détails des requêtes vers la base de données	6
TAUX D'OCCUPATION DES SYSTEMES DE FICHIERS	7
Paramètres	7
Détails des requêtes vers la base de données	7
SUIVI DES JOBS	7
SUIVI HORAIRE DU NOMBRE DE JOBS EN ATTENTE ET EN PRODUCTION	7
Paramètres	8
Détails des requêtes vers la base de données	8
REPARTITION MENSUELLE DES JOBS EN ECHEC PAR STATUT	8
Paramètres	8
Détails des requêtes vers la base de données	8
CONSOMMATION HEURES CPU	9
REPARTITION MENSUELLE DES HEURES CONSOMMEES PAR DEPARTEMENT	9
Paramètres	9
Détails des requêtes vers la base de données	9
PLUS GRANDS CONSOMMATEURS	9
Paramètres	9
Détails des requêtes vers la base de données	9
NOMBRE D'HEURES CONSOMMEES MENSUELLEMENT PAR QUEUE	10
Paramètres	10
Détails des requêtes vers la base de données	10
DISPONIBILITE	10
TAUX DE DISPONIBILITE DE BOUT EN BOUT	10
Paramètres	10
Détails des requêtes vers la base de données	10
TAUX DE DISPONIBILITE DES NŒUDS DE CALCUL	11
Paramètres	11
Détails des requêtes vers la base de données	11
TEMPS D'ATTENTE	11

NOMBRE D'HEURES D'ATTENTE MENSUEL PAR QUEUE	11
Paramètres	11
Détails des requêtes vers la base de données	12
TEMPS D'ATTENTE PAR NOMBRE DE CPU UTILISES	12
Paramètres	12
Détails des requêtes vers la base de données	12

Généralités

- Tous les indicateurs ont comme paramètre le calculateur visé. Cela se traduit au niveau des requêtes par le fait que les jobs/nœuds/événements/utilisateurs sont toujours filtrés par cluster.
- Les jobs n'ayant pas de dates de lancement ne sont pas pris en compte.
- Les durées s'établissent depuis le début du premier jour (00:00:00) jusqu'à la dernière seconde de la dernière journée (23:59:59) ;
- Les heures de calculs d'un job sont déterminées par la différence entre sa fin et son début multiplié par son nombre de CPU ;
- Le temps d'attente d'un job est déterminé par la différence entre son début et la date de soumission.
- Il n'est pas possible d'obtenir des résultats sur un intervalle supérieur à 12 mois.

Taux d'occupation

SUIVI JOURNALIER REEL OU THEORIQUE DU TAUX D'OCCUPATION DES CPUS

Fournit un graphique du taux d'occupation CPU sur une période choisie, selon la granularité. Les données sont produites de manière à être au prorata de la consommation des jobs sur chaque étape (cf granularité). Ainsi, si la granularité est « Jour », un job s'étendant sur plusieurs jours voit sa consommation répartie entre ces différents jours. Les jobs pris en compte sont ceux qui débutent ou se terminent sur la période étudiée et ayant un début clairement identifié. Que ce soit dans le mode réel ou théorique, les événements pris en compte suivent les mêmes règles.

Paramètres

- Mois début: Numéro du mois du début de la période étudiée;
- Année début: Année de début de la période étudiée;
- Mois fin : Numéro du mois de fin de la période étudiée;
- Granularité : Jour, semaine ou mois, les valeurs seront au prorata selon cette granularité ;
- Avec indisponibilité: Si cochée, en plus de l'occupation due aux jobs, une seconde courbe décrivant la somme de l'occupation et de l'indisponibilité (avec les valeurs associées dans le tableau) sera affichée.

Détails des requêtes vers la base de données

Les requêtes permettant de créer cet indicateur sont complexes. La première requête pour les jobs comporte 2 « étapes » :

1. Une sous requête générant une liste d'étapes (selon la granularité à minuit du dernier jour) avec la somme des CPU théorique disponibles pour chaque étape ;
2. Jointure avec la table jobs, pour obtenir par étape (en regroupant les jobs par étape):
 - On sélectionne les jobs dont :
 - soit la date de fin, soit la date de début, est comprise entre l'étape et l'étape +1 jour ;
 - si le job n'a pas de fin, il doit être en état RUNNING pour être pris en compte ;
 - Si le job a une fin, on fait par étape la différence entre :
 - Le plus tôt entre l'étape + 1 étape OU la date de fin (ainsi, si le job ne se termine pas après l'étape, on prend juste sa consommation sur l'étape) ;

- Le plus tard entre l'étape OU la date de début (de la même manière, si le job a commencé avant l'étape, on ne prendra sa part que sur l'étape) ;
- Si le job n'a pas de fin, on fait par jour la différence entre l'étape + 1 étape (pour arriver uniquement à la fin de journée) et le plus tard entre l'étape OU la date de début. Il est nécessaire de faire cette distinction sur la fin du fait que les jobs n'ayant pas de fin sont considérés comme ayant une fin en 1970, donc très tôt ;
- On multiplie cette différence par le nombre de cpu utilisés sur le total cpu du calculateur (cette consommation est précise à la seconde).

La seconde requête pour les événements est presque identique. Les quelques différences viennent uniquement de différences mineures entre les 2 tables (absence d'état et filtrage préalable des résultats avant la jointure non-nécessaire).

Nœuds les plus utilisés

Cet indicateur produit un top par mois des nœuds les plus utilisés, par heures CPU. Les jobs pris en compte sont ceux qui se terminent sur chacun des mois étudiés.

Paramètres

- Mois début : numéro du premier mois de la période étudiée ;
- Année début : première année de la période étudiée ;
- Mois fin : numéro du dernier mois de la période étudiée ;
- Nombre de nœuds : Taille du top.

Détails des requêtes vers la base de données

Cet indicateur génère une requête par mois compris dans la période.

Les nœuds sont :

- Sélectionnés par calculateur ;
- Joint avec les jobs via les noms des nœuds et l'id des jobs, dont la date de fin est comprise dans la période étudiée (à noter qu'ici, la période est découpée par mois) ;
- Regroupés par mois et nœuds ;
- Ordonnés par le total d'heures CPU descendant et limités aux 25 plus élevés.

SUIVI JOURNALIER DE LA CONSOMMATION CPU PAR PARTITION

Cet indicateur n'est pas encore activé, les informations en base sont insuffisantes pour le mettre en œuvre.

Cet indicateur produit 2 graphiques. Le premier représente l'occupation théorique par partition, le second l'occupation réelle (prise en compte des nœuds inactifs sur le cluster). Les données sont produites de manière à être au prorata de la consommation des jobs sur chaque jour. Ainsi, un job s'étendant sur plusieurs jours voit sa consommation répartie entre ces différents jours. Les jobs pris en compte sont ceux qui débutent ou se terminent sur la période étudiée et ayant un début clairement identifié.

Paramètres

- Mois: Numéro du mois étudié;
- Année: Année du mois étudié.

Détails des requêtes vers la base de données

Les requêtes permettant de créer cet indicateur sont complexes. La première requête pour les jobs comporte 3 « étapes » :

3. Première sous-requête créant une table des jobs liés à des partitions (on fait ainsi disparaître l'intermédiaire des nœuds, ce qui réduit beaucoup la combinatoire pour la suite) :

- Dans la table job_nodes (décrivant les relations entre jobs et nœuds), on récupère les couples id de jobs et noms de nœuds dont les partitions sont distinctes (réduisant donc la relation à un seul nœud significatif) ;
 - On réduit la jointure avec les jobs qui se terminent ou débutent dans la période ;
 - On réduit la jointure avec les nœuds qui sont sur le calculateur choisi ;
4. Seconde sous requête générant une liste d'étapes journalières (minuit de chaque jour) par partition (ainsi, si 3 partitions, sur un mois de 30 jours, cela fera 90 étapes), avec la somme des CPU théoriques disponibles pour chaque étape ;
 5. Jointure des 2 sous-requêtes, pour obtenir par jour (en regroupant les jobs par jour et partition):
 - On sélectionne les jobs dont :
 - soit la date de fin, soit la date de début, est comprise entre l'étape et l'étape +1 jour ;
 - si le job n'a pas de fin, il doit être en état RUNNING pour être pris en compte ;
 - Si le job a une fin, on fait par jour la différence entre :
 - Le plus tôt entre l'étape + 1 jour OU la date de fin (ainsi, si le job ne se termine après l'étape, on prend juste sa consommation sur la journée) ;
 - Le plus tard entre l'étape OU la date de début (de la même manière, si le job a commencé avant l'étape, on ne prendra sa part que sur la journée) ;
 - Si le job n'a pas de fin, on fait par jour la différence entre l'étape + 1 jour (pour arriver uniquement à la fin de journée) et le plus tard entre l'étape OU la date de début. Il est nécessaire de faire cette distinction sur la fin du fait que les jobs n'ayant pas de fin sont considérés comme ayant une fin en 1970, donc très tôt ;
 - On multiplie cette différence par le nombre de cpu utilisés sur le total cpu de la partition (cette consommation est précise à la seconde).

La seconde requête pour les événements d'indisponibilité des nœuds est plus simple.

1. On reprend la même sélection d'étapes que la précédente requête ;
2. Les événements d'indisponibilité visant directement des nœuds précis, on peut faire une jointure directe entre un événement et un nœud, et les regrouper par partition, avec les mêmes mécanismes que la requête précédente pour les dates ;

TAUX D'OCCUPATION DES SYSTEMES DE FICHIERS

Produit un graphique des taux d'occupation de chacun des systèmes de fichiers du calculateur en pourcentage sur un mois. La fréquence d'échantillonnage est au jour.

Paramètres

- Mois: Numéro du mois étudié;
- Année: Année du mois étudié.

Détails des requêtes vers la base de données

On détermine via une jointure les systèmes de fichier appartenant au clusteur étudié, puis on récupère une ligne par jour et système de fichier indiquant le taux d'occupation.

Suivi des jobs

SUIVI HORAIRE DU NOMBRE DE JOBS EN ATTENTE ET EN PRODUCTION

Fournit un graphique en barre produit à partir d'un échantillonnage des jobs en production et en attente. Ces informations se basent sur l'état des jobs à chaque heure (00h00, 01h00, 02h00, etc.). Ce n'est donc pas une moyenne sur l'heure.

Un job est considéré comme actif à une heure donnée si son lancement et sa fin d'exécution chevauche cette heure. Il est en attente à une heure donnée si sa soumission et son lancement chevauche cette heure.

Du fait de la fréquence horaire, la production de cet indicateur sur un mois complet est particulièrement lourde.

Paramètres

- Jour début : jour de début de la période étudiée (celle-ci étant au maximum d'un mois) ;
- Jour fin : jour de fin de la période étudiée ;
- Mois : Numéro du mois sur lequel se déroule l'étude ;
- Mode :
 - Tous : l'ensemble des jobs sont concernées ;
 - Nouveau : seuls sont pris en compte les jobs soumis pendant la période étudiée ;

Détails des requêtes vers la base de données

Cet indicateur se fait en 2 étapes similaires, avec une variation sur la clause pour récupérer soit les jobs en production, soit les jobs en attente. Les jobs en production sont ceux où l'étape est comprise entre le début et la fin de l'exécution. Les jobs en attentes sont ceux où l'étape est comprise entre la soumission et le début de l'exécution.

La requête comporte 3 structures :

- la génération d'une série d'étapes pour chaque heure sur la période étudiée ;
- la génération d'une liste de jobs, sélectionnés par :
 - la date de fin de jobs, qui doit être postérieure au début de la période étudiée ;
 - la date de soumission, qui doit être antérieure à la fin de la période étudiée ;
- pour chacune des étapes, le total des jobs et CPU correspondant de la liste filtrée, regroupés par étape ;
- En cas de mode nouveau, seuls les jobs dont la date de soumission est comprise dans la période sont pris en compte.

REPARTITION MENSUELLE DES JOBS EN ECHEC PAR STATUT

Produit un graphique en barre faisant la sommes des jobs par état qui ne soit pas dans l'état « COMPLETE », regroupés par mois (en abscisse). Contrairement à la majorité des indicateurs, du fait que de nombreux jobs étudiés n'ont pas une date de fin valide, ce sont les jobs dont la date de soumission est comprise dans chacun des mois étudiés.

Paramètres

- Mois début : numéro du premier mois de la période étudiée ;
- Année début : première année de la période étudiée ;
- Mois fin : numéro du dernier mois de la période étudiée ;

Détails des requêtes vers la base de données

2 requêtes sont effectuées :

1. Une première requête récupérant la liste des différents états différents de l'état « COMPLETE » sur la période étudiée ;

2. La seconde récupère les jobs dont :
- La date de soumission est dans la période étudiée ;
 - L'état est différent de l'état « COMPLETE » ;
 - Regroupé par numéro de mois et état.

Consommation Heures CPU

REPARTITION MENSUELLE DES HEURES CONSOMMEES PAR DEPARTEMENT

Fournit un graphique en barre fournissant la part de consommation répartie sur plusieurs mois de chaque département. Les jobs pris en compte sont ceux qui se terminent dans le mois étudié, calcul n'est donc pas réalisé au prorata.

Paramètres

- Mois début : numéro du premier mois de la période étudiée ;
- Année début : première année de la période étudiée ;
- Mois fin : numéro du dernier mois de la période étudiée ;

Détails des requêtes vers la base de données

Récupère les jobs dont :

- La date de fin est comprise dans la période étudiée ;
- Sont reliés à un utilisateur via :
 - Correspondance de l'UID du job avec l'UID de l'utilisateur ;
 - Correspondance du clusteur de l'utilisateur au clusteur étudié ;
 - La date de soumission du job est comprise entre la date de création et la date de destruction du compte utilisateurs ;
- Groupés par numéro du mois et par département des utilisateurs.

PLUS GRANDS CONSOMMATEURS

Produit un top sous forme de liste des plus grand « consommateurs » répartis éventuellement sur plusieurs mois (cf paramètres), selon 3 variables (cf les modes). Quelque soit le mode, les jobs pris en compte sont ceux qui se terminent durant le mois étudié.

Paramètres

- Mois début : numéro du premier mois de la période étudiée ;
- Année début : première année de la période étudiée ;
- Mois fin : numéro du dernier mois de la période étudiée ;
- Nombre de résultats : Taille du top ;
- Mode :
 - Nombre de jobs : top des utilisateurs ayant soumis le plus de jobs ;
 - Heure de calculs : top des utilisateurs ayant consommés le plus d'heures CPU ;
 - Temps d'attente : top des utilisateurs dont les jobs ont accumulé les plus longues attentes.

Détails des requêtes vers la base de données

Quel que soit le mode, les jobs récupérés sont :

- Ceux dont la fin est comprise dans la période ;
- Sont reliés à un utilisateur via :
 - Correspondance de l'UID du job avec l'UID de l'utilisateur ;
 - Correspondance du clusteur de l'utilisateur au clusteur étudié ;

- La date de soumission du job est comprise entre la date de création et la date de destruction du compte utilisateur ;

Ensuite, ils sont regroupés et ordonnés différemment selon le mode :

- Nombre de jobs : regroupés par numéro de mois et noms d'utilisateur, ordonnés par nombre de jobs descendant ;
- Heures de calcul : regroupés par numéro de mois et noms d'utilisateur, ordonnés par la somme des heures de calculs ;
- Temps d'attente : regroupés par numéro de mois et noms d'utilisateur et département, ordonnés par la moyenne des temps d'attente ;

NOMBRE D'HEURES CONSOMMEES MENSUELLEMENT PAR QUEUE

Produit un graphique en barre représentant le nombre d'heures consommées par queue regroupées par mois. Les jobs pris en compte sont ceux qui se terminent dans chacun des mois étudiés.

Paramètres

- Mois début : numéro du premier mois de la période étudiée ;
- Année début : première année de la période étudiée ;
- Mois fin : numéro du dernier mois de la période étudiée ;

Détails des requêtes vers la base de données

2 requêtes sont effectuées :

1. Une première requête récupérant la liste des différentes queues comportant au moins un job qui s'est terminé sur la période ;
2. La seconde récupère les jobs dont :
 - La date de fin est dans la période étudiée ;
 - Regroupé par queue.

Disponibilité

TAUX DE DISPONIBILITE DE BOUT EN BOUT

Cet indicateur mesure la complétude de jobs « sonde », répondant à certains profils. Pour chaque heure, 12 jobs par profil sont lancés et doivent aboutir pour obtenir une disponibilité de 100% sur l'heure (soit 12*24 par jour, etc). Chaque courbe correspond à un scénario. Le tableau permet d'obtenir le détail des jobs « sondes » (tout scénario confondus). Les jobs sondes sont lancés sous l'utilisateur « **rd-ap-hpcstats** ».

Paramètres

- Mois début : numéro du premier mois de la période étudiée ;
- Année début : première année de la période étudiée ;
- Mois fin : numéro du dernier mois de la période étudiée ;
- Granularité : Heure, Jour, Mois, les valeurs seront au prorata selon cette granularité ;

Détails des requêtes vers la base de données

La requête comporte 4 phases :

1. Une sous-requête générant une liste d'étapes (selon la granularité à minuit du dernier jour) avec la somme des CPU théorique disponibles pour chaque étape ;
2. Une sous-requête générant une liste de profils ;
3. Une sous-requête générant une liste de jobs filtrés sur la base de l'utilisateur, regroupés par profils ;
4. Jointure des 3 sous-requêtes.

TAUX DE DISPONIBILITE DES NŒUDS DE CALCUL

Produit une courbe indiquant la disponibilité des nœuds de calculs, au prorata selon la granularité choisie. Seules les sorties de nœuds ayant un début identifié sont prises en compte.

Paramètres

- Mois début: Numéro du mois du début de la période étudiée;
- Année début: Année de début de la période étudiée;
- Mois fin : Numéro du mois de fin de la période étudiée;
- Granularité : Jour, semaine ou mois, les valeurs seront au prorata selon cette granularité ;

Détails des requêtes vers la base de données

La requête permettant de créer cet indicateur est complexe. Elle comporte 2 « étapes » :

1. Une sous requête générant une liste d'étapes (selon la granularité à minuit du dernier jour) avec la somme des CPU théorique disponibles pour chaque étape ;
2. Jointure avec la table des événements, pour obtenir par jour (en regroupant les événements selon la granularité):
 - On sélectionne les événements dont :
 - soit la date de fin, soit la date de début, est comprise entre l'étape et l'étape +1 étape ;
 - Si l'évènement a une fin, on fait par étape la différence entre :
 - Le plus tôt entre l'étape + 1 étape OU la date de fin (ainsi, si le job ne se termine pas après l'étape, on prend juste sa consommation sur la journée) ;
 - Le plus tard entre l'étape OU la date de début (de la même manière, si l'évènement a commencé avant l'étape, on ne prendra sa part que sur la journée) ;
 - Si l'évènement n'a pas de fin, on fait par étape la différence entre l'étape + 1 étape et le plus tard entre l'étape OU la date de début. Il est nécessaire de faire cette distinction sur la fin du fait que les événements n'ayant pas de fin sont considérés comme ayant une fin en 1970, donc très tôt ;
 - On multiplie cette différence par le nombre de cpu utilisés sur le total cpu du calculateur (cette consommation est précise à la seconde).

Temps d'attente

NOMBRE D'HEURES D'ATTENTE MENSUEL PAR QUEUE

Produit deux graphiques en barre reprenant les 5 queues ayant eu le plus d'attente (soit en moyenne, soit au total, voir les modes). Le premier graphique correspond au choix du mode et détermine l'ordre des barres. Le second graphique est basé sur le nombre de jobs, mais ordonné selon le choix du mode. Le document fournit en plus une liste de toutes les queues (non limitée aux 5 premières), avec les informations souhaitées.

Les jobs pris en compte sont pour chaque mois ceux qui se terminent durant le mois. Le calcul n'est donc pas effectué au prorata.

Paramètres

- Mois début : numéro du premier mois de la période étudiée ;
- Année début : première année de la période étudiée ;
- Mois fin : numéro du dernier mois de la période étudiée ;
- Mode :
 - Moyenne : fait la moyenne du nombre d'heures d'attentes moyennes des jobs par queue ;
 - Nombre : fait la somme du nombre d'heures d'attentes moyennes des jobs par queue.

Détails des requêtes vers la base de données

Récupère les jobs dont :

- La date de fin est comprise dans la période étudiée ;
- Groupés par numéro du mois et par queue.

TEMPS D'ATTENTE PAR NOMBRE DE CPU UTILISES

Produit des tops mensuels du temps d'attente selon le nombre de CPU utilisés par les jobs. Les jobs sont donc regroupés selon leur nombre de CPU et l'on vérifie les temps d'attente sur ce critère. Chacun des tops ne concerne qu'un mois sur la période étudiée, et les jobs pris en compte sont ceux qui se terminent sur ce mois.

Paramètres

- Mois début : numéro du premier mois de la période étudiée ;
- Année début : première année de la période étudiée ;
- Mois fin : numéro du dernier mois de la période étudiée ;
- Nombre de résultats: Taille du top.

Détails des requêtes vers la base de données

Les jobs sont :

- Ceux qui se terminent durant le mois étudié ;
- Regroupés par nombre de CPU pour sommer leur temps d'attente ;
- Ordonnés par la somme des temps d'attente.