

Compte rendu d'activité pour l'année écoulée

1 Les objectifs fixés

Présentation du sujet de thèse: L'univers du datacenter se révèle être beaucoup plus complexe que l'on y pense. il regroupe plusieurs métiers dont l'interaction devient difficile au vue des tâches à accomplir complètement différentes. Toutefois ces secteurs d'activités partagent un objectif commun qui est l'accessibilité et la disponibilité des services IT hébergés dans cette infrastructure. En effet au début de leur apparition, l'intérêt était dirigé vers les données contenues dans les serveurs afin qu'ils soient accessibles 24h/24 délaissant au passage l'impact que ces bâtiments avaient sur l'environnement. On peut citer la production de chaleur issues des équipements IT, la surconsommation d'électricité pour éviter toute interruption d'intensité et une utilisation d'eau assez importante afin de refroidir les serveurs.

Notre coeur de métier est l'optimisation des énergies (électrique, thermique) produites dans un souci de réduire les couts d'exploitations et les impacts écologiques de ces bâtiments. Un des soucis rencontrés lors de la réalisation de notre projet a été l'absence de référentielles correctes sur les équipements en fonctionnement. Par exemple certains appareils, remplacés par d'autres avec des caractéristiques différentes ne figurent nulle part dans la documentation, l'historique des maintenances est décousu avec beaucoup d'incompréhension et l'organisation physique des appareils est modifiée et pas mis à jour.

Face à ces contraintes, notre préoccupation est de savoir comment trouver l'architecture physique des équipements correspondant au fonctionnement du système à un instant donnée. L'idée nous est alors venu d'utiliser la variable constamment remontée dans le temps qui reflète l'état du système. Il s'agit des mesures physiques.

Travaux réalisés: le 1er trimestre a consisté à définir la fonction de corrélation entre 2 arcs. Cette fonction détermine le pourcentage de similarité entre l'ensemble de mesures prises sur les 2 arcs. En exécutant cette fonction sur tous les paires possibles d'arcs on obtient la matrice de corrélation des arcs. La matrice de corrélation est la matrice d'adjacence du graphe non orienté sous-jacent du réseau de flots à prédire.

Par la suite, une étude bibliographique sur la décomposition en cliques à partir d'un linegraph en a déduit que des travaux avaient été effectués seulement dans le cas où la matrice d'adjacence du graphe formait un *linegraph*. Durant le deuxième trimestre, on a supposé que cette matrice d'adjacence formait un *linegraph* et nous avons implémenté cet algorithme de décomposition en cliques. Nous avons considéré que des arcs couverts par une clique partagent un sommet en commun. Cependant nous ignorions lesquels de ces arcs appartenaient à l'ensemble d'arrivé ou à l'ensemble de départ de ce sommet.

Pendant le troisième trimestre, nous avons conduit les travaux sur l'identification des arcs entrants et sortants du noeud déduit de la clique. Nous avons défini une fonction d'orientation pour jouer ce rôle. Cette fonction est une fonction de décision qui depend d'une variable epsilon ϵ . Si cette variable est petite, la

fonction d'orientation génère beaucoup de bipartitions fausses positives. Par contre, si ϵ est grande, la fonction d'orientation génère beaucoup de bipartitions fausses négatives.

Une étude sur le bon intervalle de décision a été faite au début du quatrième trimestre et la prise en compte des erreurs de corrélation dans la matrice d'adjacence a été introduite dans l'algorithme de découpage en cliques. Cette dernière étape est toujours en cours.

Travaux prévisionnels:

- achever la correction de la matrice d'adjacence. Cette correction consiste à trouver le *linegraph* le plus proche de cette matrice d'adjacence. En d'autres termes c'est changer le minimum possible de valeurs de corrélation dans la matrice.
- tester cette algorithme sur des données réelles.