# RAPPORT LABORATOIRE n°4 ASD2

# « TABLES DE HACHAGE »

Le laboratoire « Tables de hachage » est composé de deux parties.

Dans la première partie du labo nous avons pour but de comparer les 6 différentes fonctions de hachage fournies dans les données.

Nous devons répondre quelle fonction est la plus adapté pour le type de donnée fourni.

Pour passer d’un type de donnée à l’autre nous devons modifier le « max\_load\_factor ».

La valeur par défaut du « max\_load\_factor » fournie par STL est 1.0.

Les fonctions hash :

1. DirectoryInt - utilisation de la fonction hash STL
2. DirectoryLong - utilisation de la fonction hash STL
3. DirectoryPol<z> - h = (z\*h) + avs[i].

(z = « max\_load\_factor » et avs[i] correspond aux numéros AVS)

1. DirectoryStl - utilisation de la fonction hash STL
2. DirectorySha256 - opérations binaire ???
3. DirectoryCity - utilisation de la fonction hash implementé par google

Lors des tests de performances, un MAX\_LOAD\_FACTOR définit à 3 nous donne un pourcentage de collisions de 65%. A 20, nous donne 90%. A partir d’un facteur de 300, on tend vers 100%.

Le facteur MAX\_LOAD\_FACTOR va influencer le nombre de bucket. Il s’agit du taux de remplissage.

**Phase 2**

L’implémentation fournie, n’est pas optimale car la clé (le nom de la personne) ne permet pas d’identifier de manière unique cette personne. (Taux de collision pour un facteur de 0.5 de ~62% )

Une première solution serait de rendre simplement plus unique la clé utilisée. Par exemple, une composition du nom, prénom, genre, date de naissance (On tombe à un taux de collision de 16%).

Une deuxième solution serait d’implémenter une compression polynomiale sur la même clé composée (nom, prénom, genre, date de naissance)