Segue como eu realizei os testes:

**Criação do banco:**  
 Criei o banco usando SimpleStrategy para replicação e fator de replicação 1.

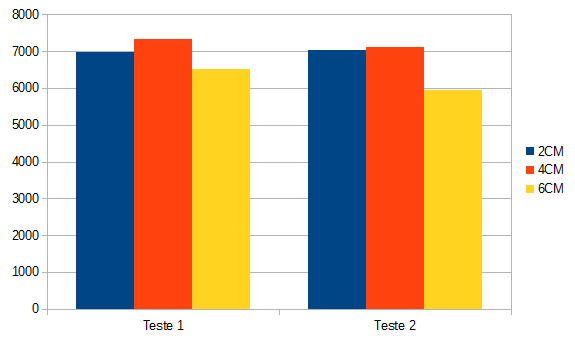
 Apenas uma tabela foi criando, utilizando um subconjunto das colunas do dos arquivos do bolsa família: *uf, periodo, valor, nis\_favorecido, nome\_favorecido, cod\_municipio, nome\_municipio, fonte*.

 A chave primária escolhida foi *uf, periodo, valor e nis\_favorecido*. Esses são os campos que identificam unicamente um registro.

 Além disso foi utilizado ordenamento nas colunas *periodo*(ascendente), *valor*(descendente), buscando uma melhora no tempo de busca.

**Configuração dos clusters:**  
Foram configurados dois clusters, utilizando no total 12 computadores do laboratório, separados em 6 e 6. Em cada um os testes foram realizados utilizando 2, 4 e 6 máquinas.

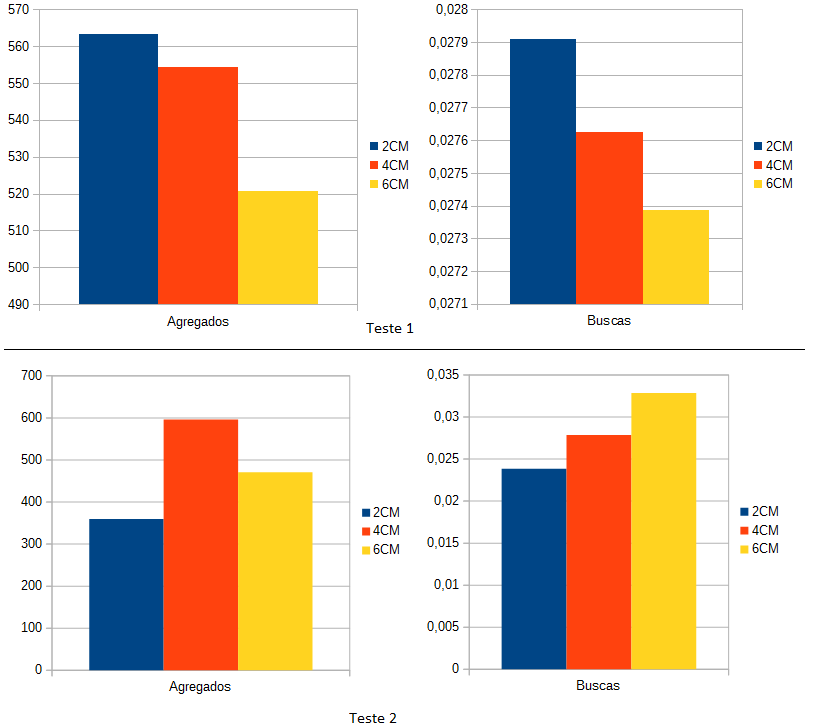
**Inserção:**  
 Baixei os 24 arquivos do bolsa família(de 07/2014 a 06/2016) e realizei a inserção por meio de um programa escrito em java utilizando o driver da Datastax. A inserção é realizada apenas em um nó do cluster.  
 Foram realizadas 10 inserções para cada tamanho de cluster(2, 4 e 6 máquinas). Os tempos foram anotados e o gráfico foi feito fazendo-se a média de 8 resultados, ignorando o maior e o menor tempo.



(Onde Teste 1 é o teste com o primeiro conjunto de 6 máquinas e Teste 2 com o segundo conjunto de outras 6 máquinas)

**Consultas:**Foram realizadas dois tipos de consultas: agregação e busca por registro.  
Na consulta por agregação eu pedi a soma dos valores recebidos em algumas ufs.  
Na consulta por busca eu busquei por alguns registros específicos utilizando toda a chave primária como parâmetro.

Assim como na inserção, foram realizadas 10 repetições em cada cluster, ignorando-se os extremos para criação do gráfico.



Como é possível ver, alguns testes estão como seria o esperado, mas como o outro não está não sei se o resultado é confiável o suficiente para utilizar.