Changes to HBase SLB to take into account heterogeneous region server configurations.

At a high level a new cost function can be added to SLB which looks at the proportion of regions allocated to each node with respect to the resources the node has and if a node has more than what it ideally should have increase the cost of the candidate cluster for the balancer to not favor it.

- Each server when checking in with master advertises its resource capacity for e.g. total memory (M) available
- Master stores the resource details from all the servers in the cluster and identifies the max capacity say (Mc) which is ideal
- Proportion of each node (Nx) with respect to the ideal can be calculated (Nx/Mc) to come-up with the number of nodes with max capacity (Mc)
- Summing up the ratios for all the nodes provides the ideal number of servers (W) in the cluster with all nodes having the max capacity (Mc)
- Total number of regions in the cluster (T) divided by the number of nodes in the ideal cluster (W) will determine the max number of regions (Mr) for any node with max capacity (Mc)
- (Nx/Mr)*W will be the number of regions all the nodes in the cluster should ideally have and that can be used to identify the skew.
- If there are more than the expected number of regions (from the calculation from the previous step) in a region server then the difference should be considered as a skew.
- The total of the skews from all the servers multiplied by a standard cost multiplier can be added to the total cost for the balancer to determine whether the candidate cluster is better one

To illustrate using an example

- In a 5 server cluster with 50 regions, each nodes advertises the capacity [100, 150, 200, 250, 300]. Note that the values are by server index starting at 1. The max capacity in this cluster is server 5 with value 300
- The proportion of each server capacity to the max is [0.33333334, 0.5, 0.6666667, 0.8333333, 1.0] and the ideal cluster has a size of 3.3333333 servers
- The max number of regions in a server in the ideal cluster will be 15. So in terms of number of regions per server can ideally serve is approx [5 8 10 12 15]
- If a candidate cluster has a region distribution of [2 12 12 12 12 12] then the skew will be [0 4 2 0 0] and the sum will be 6. By having a cost multiplier of say 500, the additional computed cost will be 3000 which can be used by the balancer to determine the best candidate in a SLB run

Note that this cost function can replace the *RegionCountSkewCostFunction* .

Having the data about the skew can then be used to create a candidate generator that can try to move regions from a highly skewed server to a non-skewed server and generate a candidate cluster that can be lower in cost.

If multiple resource capacities say memory and CPU need to be taken into account, it can be approached by having

a) Multiple cost functions or
b) A single cost function that amplifies the lack of resources by multiplying the proportion to the max of each resource of a server.

When a server is removed from the cluster (gracefully or not), master need to remove the data it stores about the capacity of the server so that it is not taken into account in computing the cost further.