

## Log Structured Merge Trees aplicada no jogo The Last Mage

O Log Structured Merge Tree é uma estrutura de dados que provê acesso indexado para arquivos com um grande volume de conteúdo. Mantém os dados em duas ou mais estruturas separadas que são sincronizadas de forma eficiente. É ideal para aplicações que a escrita tem um custo alto, pois isso é otimizado por essas estruturas. Os dados são escritos de forma ordenada e sequencial em String Sorted Tables (SSTables) quando atingem um limite de memória volátil. Um grande detalhe a ser tratado nessa técnica é que a SSTable criada pela LSM é imutável, o que faz com que as atualizações nas mesmas impliquem em duas SSTables com dados parecidos. Para este problema, juntamos duas SSTables na intenção de remover as duplicatas e deixar o arquivo o mais atualizado de acordo com o estado atual do Sistema. As árvores LSM causam alguma amplificação de gravação: os dados devem ser gravados no log de gravação antecipada e depois liberados no disco, onde serão eventualmente relidos e gravados novamente durante o processo de compactação.

Pode ser utilizado em nosso projeto com a finalidade de acelerar o tempo de execução de um log que seja muito grande, já que suas operações são somente executadas em memória principal. Essa estrutura possibilita uma escrita otimizada e uma busca indexada, assim torna fácil a busca por dados. Para sua implementação, primeiramente seria necessário criar snapshots da memtable de forma sincronizada com a escrita no log.

Os snapshots poderiam ser criados determinando um valor  $n$  de operações a serem armazenadas no log e, após essas  $n$  operações, realizando uma descarga desses dados para um snapshot e então resetar a memtable. Após o snapshot ser concluído, temos agora SSTABLES, já que os arquivos em disco não correspondem mais a snapshots do banco de dados. A partir disso, é necessário compactar essas SSTABLES, para isso, o programa leria os registros de maneira sequencial, eliminando dados desnecessários e diminuindo o número de arquivos.

Em seguida a compactação pode ser feita, unindo dois segmentos de dados no merge sort (comparar dois arquivos e escolhendo a menor chave da vez). Por último, para diminuir o problema de encontrar as sstables pertinentes, implementaríamos um filtro Bloom associado a cada SSTABLE.