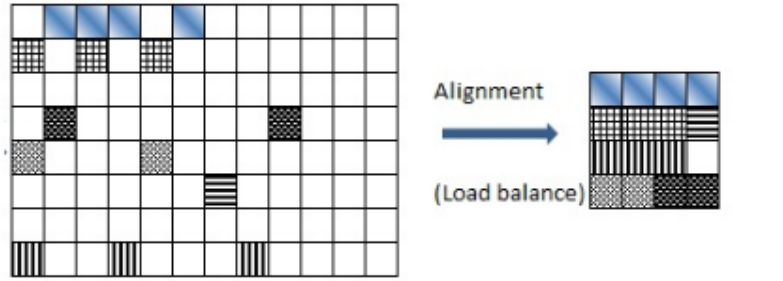
Question 5:

Find a published paper from an ACM or IEEE conference that discusses a novel sparse matrix format that was not covered in class. Discuss why the proposed format is superior to the CSR or CSC format. Make sure to cite your sources.

For this question, I read the paper "*An Efficient Sparse Matrix Multiplication for Skewed Matrix on GPU*", published in IEEE conference, which talked about a new sparse matrix format *Aligned COO* (extension to COO) format to optimize the performance of large sparse matrix having skewed distribution of non-zero elements. Coordinate list format (COO) format stores data as a list of tuples with three elements; row, column, value. The first element is the row index, the second element is column index, and the third element is the value to be stored in the row and column. Aligned COO format tries to set maximum alignment across the computing resources and helps to gain better performance without any extra memory overhead.

Aligned COO format stores data in column major order to support coalesced memory access. Below figure shows COO Alignment using data reordering.



Following are the reasons why Aligned COO format performs well in sparse matrix vector multiplication(SpMV) compared to CSR format.

- CSR provides higher concurrency through vectorization. Vectorization leads to splitting the row across the warp boundary causing an update loss problem and increases overhead of parallel reduction when insufficient data is to be processed by a warp.

- Load balancing among computational resources reduces thread scheduling overhead in Aligned COO Format. As a result, Aligned COO Format performs better compared to CSR.
- In Aligned COO format, each element of segments is executed in parallel to achieve synchronization-free concurrent execution. The process of segmentation ensures that the segment contains only one element of each row for synchronization free concurrent execution. Due to lack of synchronization in CSR format, update loss problem exists. As a result, Aligned COO Format performs better compared to CSR.
- CSR and ELL formats are only suitable for sparse matrices which have rows with similar row length. The SpMV implementation used here is a hybridization of SpMV ALIGNED COO seg kernel and SPMV ALIGNED COO flat kernels. Due to this implementation, Aligned COO format can work for a wide range of sparse patterns.
- Due to increased reuse of input vector and reduced fetch operation, Aligned COO format performs better than CSR.

References:

1. M. Shah and V. Patel, "An Efficient Sparse Matrix Multiplication for Skewed Matrix on GPU," 2012 IEEE 14th International Conference on High Performance Computing and Communication & 2012 IEEE 9th International Conference on Embedded Software and Systems, 2012, pp. 1301-1306, doi: 10.1109/HPCC.2012.192.
2. https://cmdlinetips.com/2019/01/3-ways-to-create-sparse-matrix-in-coo-format-with-scipy/