Matrix Engines In Conclusion

- Increased throughput over FMA units
- Our ME-enabled GEMM microkernels get close to peak
- Our ME-enabled GEMM kernels get 80-90% of peak
- DLA Libraries need to adapt to support ME (New packing format not supported by BLIS)





