# Techniques in Accelerating Query Processing on GPU

# File Decoders

Encoding	Throughput
Dictionary/Bitpacking	300 - 580 GB/s
Bitmask to Indices	570 GB/s
Sparse Bool	44.85 GB/s
Varint	490 GB/s
Mainly Constant	1000 GB/s
RLE Total Length	393 GB/s
RLE	522 GB/s

# Hash Probe

- Table size: 256 \* 1024 \* 1024
- Key: uint64
- Value: uint64

#### Hash Probe Performance No Tags No Partitioning With Tags Partitioned Partitioned with Tags 10 8 Billion Elements per Second 6 4 2 0 LF=0.5 MR=0.1 LF=0.5 MR=0.5 LF=0.8 MR=0.1 LF=0.8 MR=0.5

LF = Loading Factor, MR = Matching Rate

# Exchange

- OpenUCX provides a convenient interface for interacting with different hardware connections
- Numbers we got on Nvidia A100

Connection	Bandwidth
SYS	9.02 GiB/s
NV2	42.6 GiB/s
NV4	83 GiB/s

# Conclusion

GPU acceleration is useful in almost all aspects of query processing, not just project, select, hash join like studied in previous work, but also table scan and exchange. Even with hash join, there are some more techniques to be tried to further boost performance.