# Network Stack Scaling

Vijay Singh, NetApp

# Issues

- Lock contention.
- Leveraging stateless offload NIC features.
- Performance profiling tools.
- New RFC support.

# Lock Contention

- pcbinfo lock primary issue. PCBGROUP helps.
- Some locks still coarse grained. ACCEPT_LOCK() still since instance across vnets.
- Per-connection locks become an issue if flow is not serialized.
- Lock profiling at a high pps benchmark such as SFS is problematic.

# Use of NIC features

- RSS – we use it to provide the hash. Original intent of PCBGROUP was CPU affinity, but that has scaling issues.
- How to use the multiple Rx/Tx queues? Additional buffer space or notion of flow?
- For kernel consumers we have an issue where pcbinfo and accept locks are held across socket upcalls.

# Tools

- Lock profiling at high IO rates.
- Data on false sharing of cache-lines, where sw prefetching would help or mis-predicted branches.
- Where adaptive mutex isn’t helping.
- Some new RFC support such as TCP fast open (experimental) can help transactional workloads.