**Inference Latency vs. Network Bandwidth (ALBERT-large)**

Legend:
- OCC
- DINA
- MEIDA
- Ours

X-axis: Network Bandwidth (Mbps)
Y-axis: Inference Latency (ms)

*Lower is better*