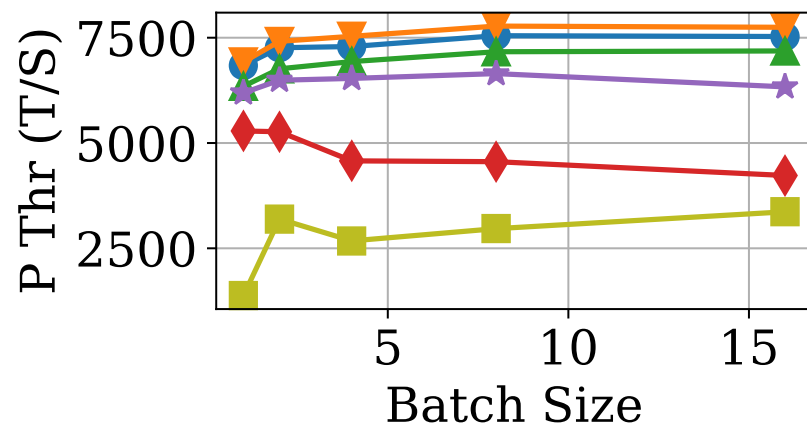
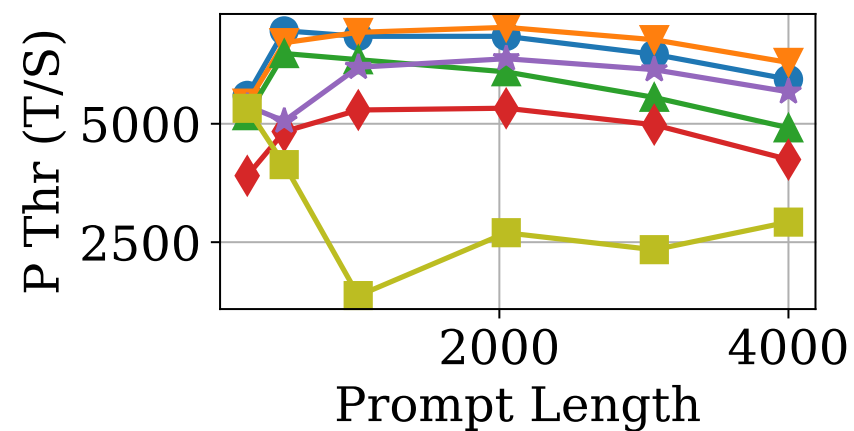


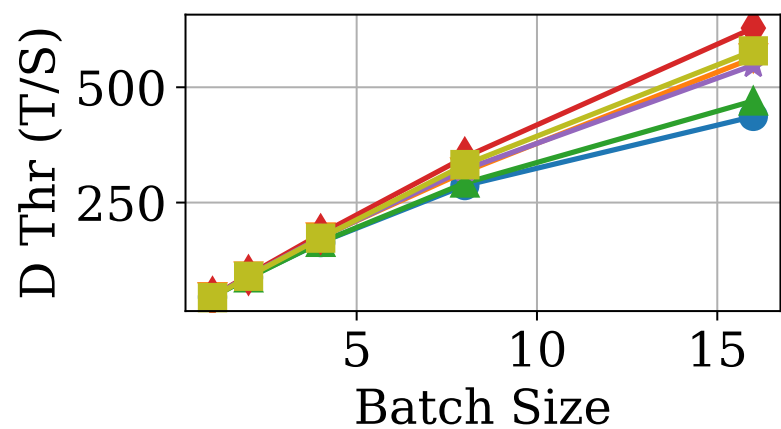
● FP16   
 ▼ K-4   
 ▲ G-4   
 ◆ H2O   
 ★ Stream   
 ■ SnapKV



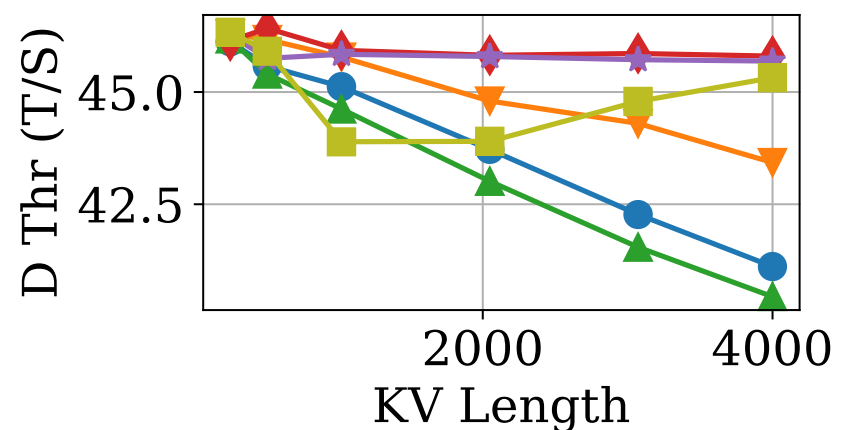
**Prefill, Prompt Length 1024**



**Prefill: Batch Size 1**



**Decode: Prompt Length 1024**



**Decode: Batch Size 1**