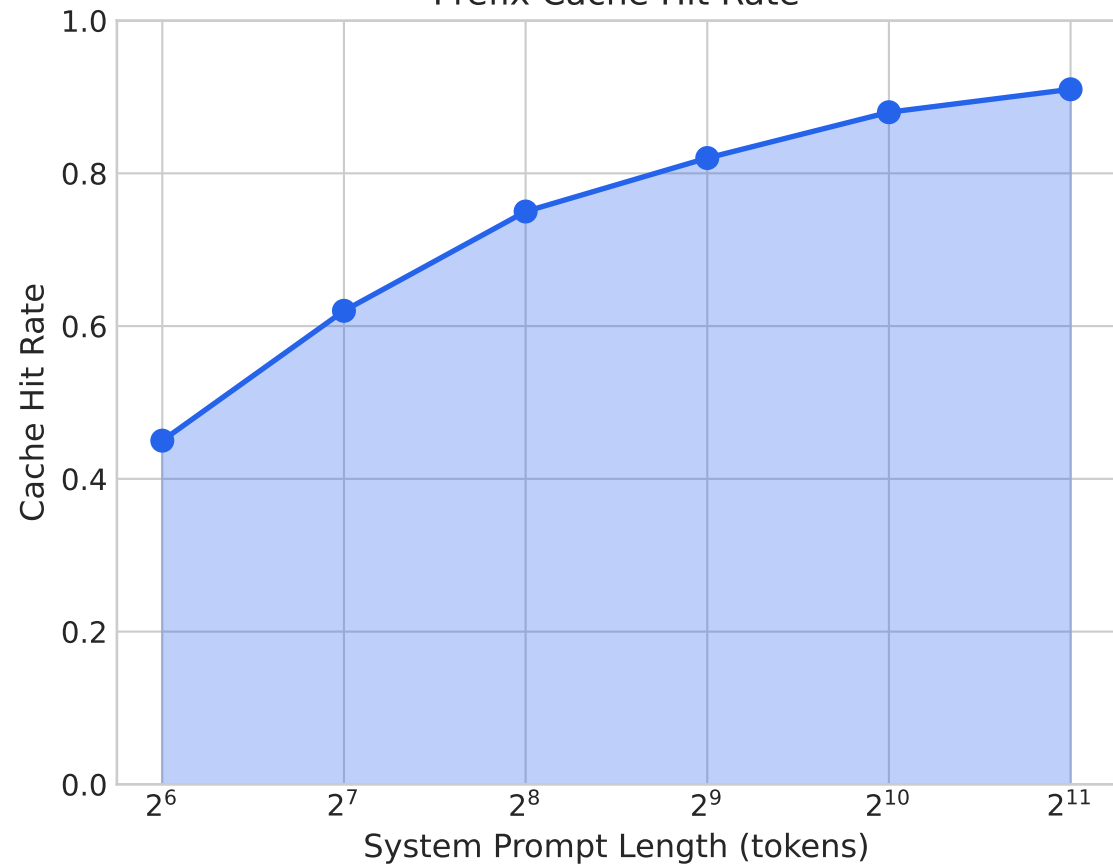


Prefix Cache Hit Rate



Prefill Latency with Prefix Caching

