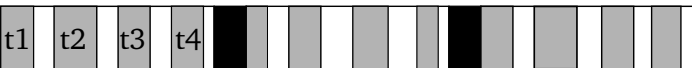


CPU: optimized for low latency



GPU: optimized for high throughput

