

XMCP - Runtime Optimization

Closed-loop self-optimization as a hardware behavior

XMCP turns observation into action. The Optimization Layer adjusts the pipeline while the workload runs, then scores the result - grounded in systems and control theory, not static heuristics.

What it tunes

- Routing - execution paths chosen from measured congestion.
- Precision - per-operation 8/16/32-bit selection.
- Cache / prefetch - driven by predicted reuse.
- Sparse thresholds - how aggressively zero work is skipped.

Explainability and evidence

The Decision Center records the input, decision, reason and confidence for each stage, with a confidence heatmap and root-cause view. The Evidence Center attaches proof and a confidence score to each capability claim - all from runtime outputs.

Note: figures shown in XMCP Studio are produced at runtime on representative workloads; they illustrate behavior and are not a performance guarantee for a specific deployment.