FIG. 1 — System Architecture

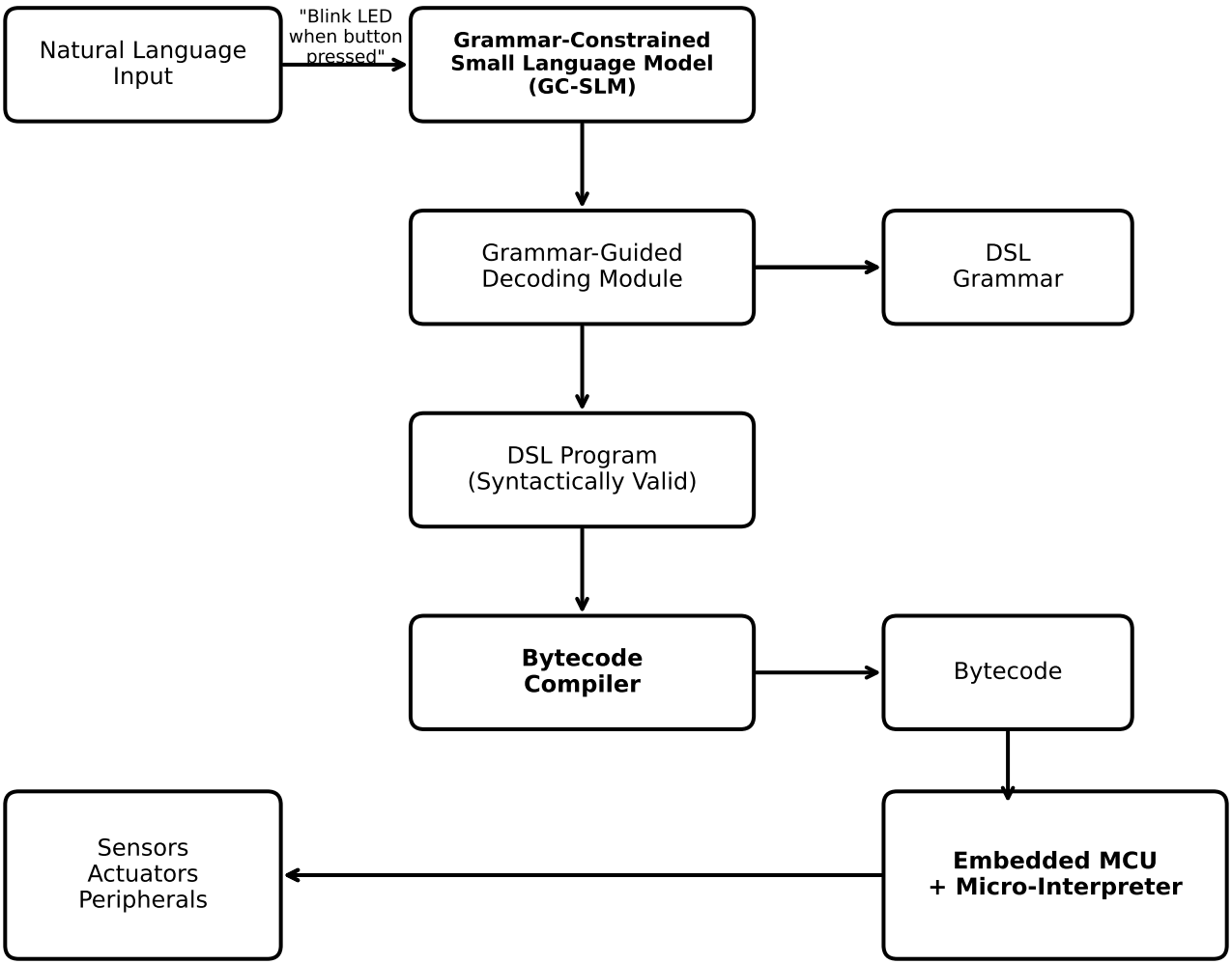

FIG. 2 — Grammar-Guided Decoding Flow

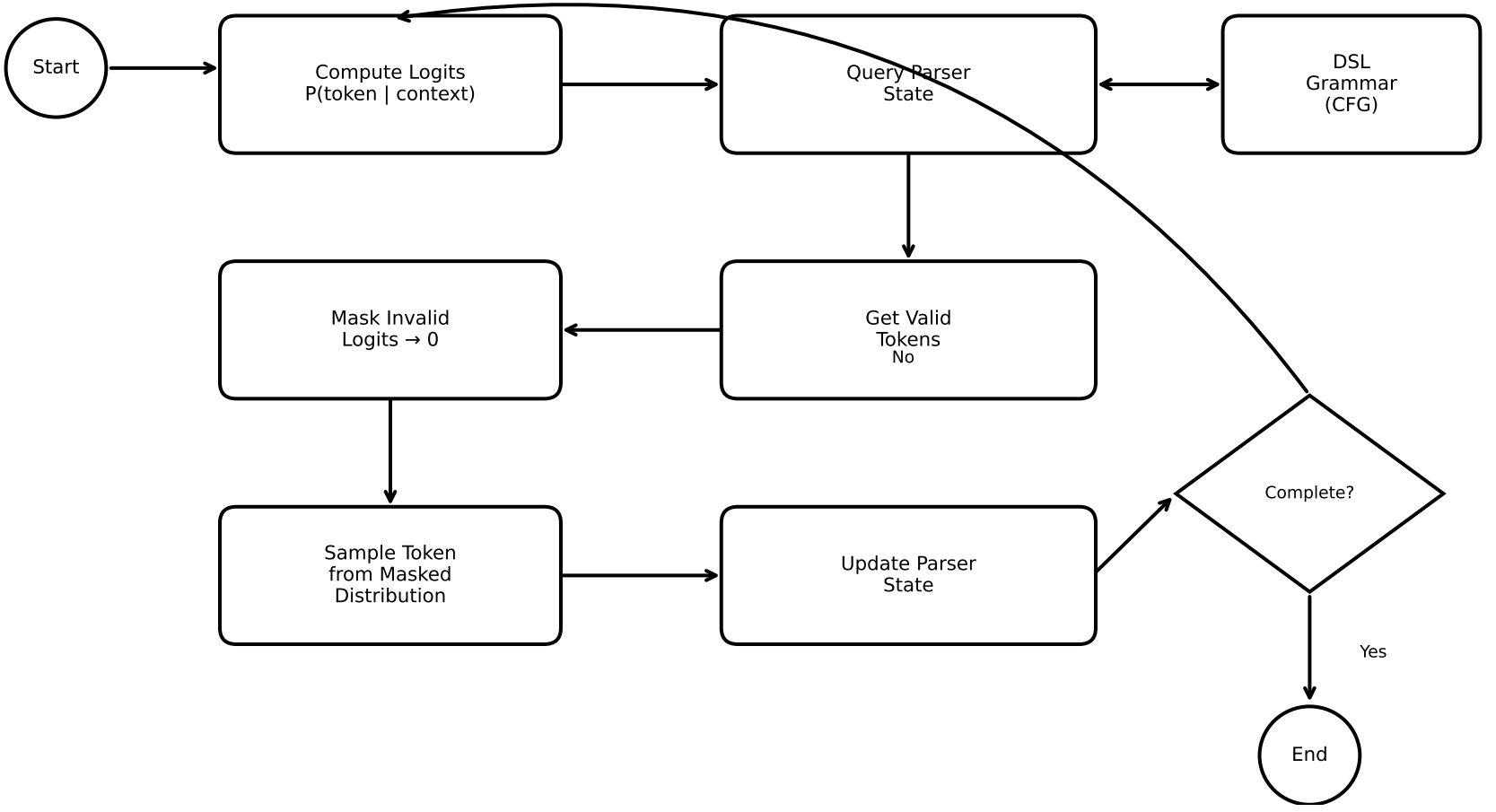

FIG. 3 — Sparse Low-Rank Adaptation (S-LoRA)

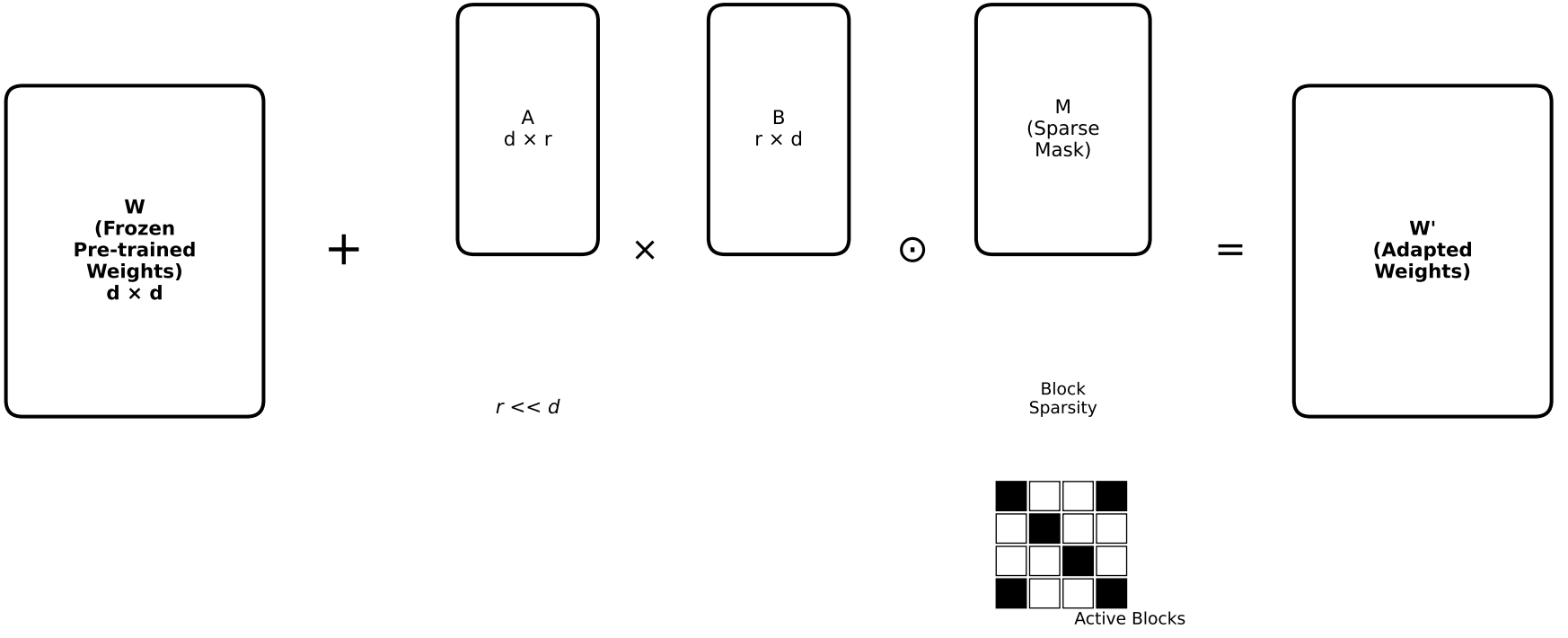

FIG. 4 — Micro-Interpreter Architecture

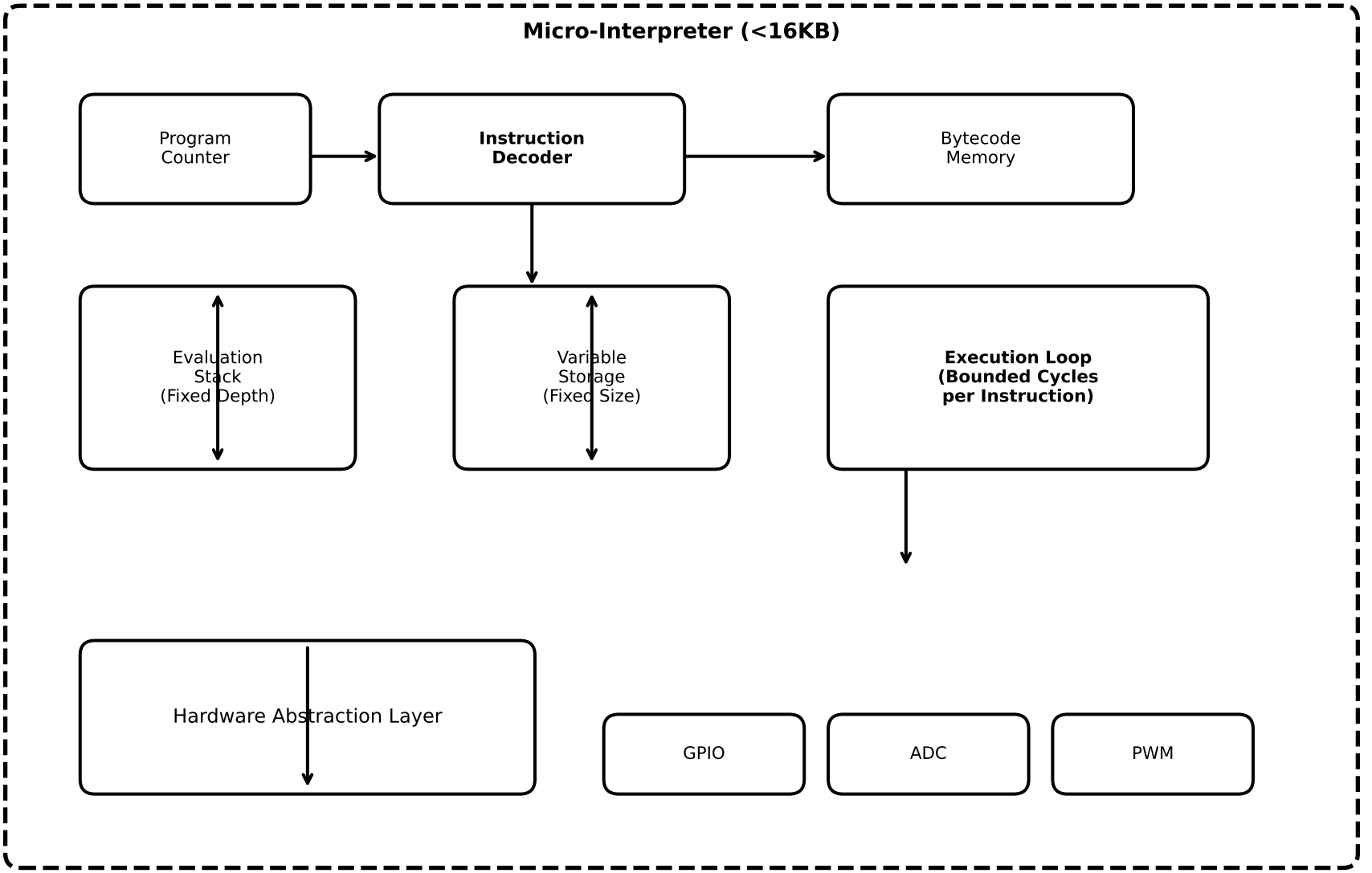

FIG. 5 — MCU Deployment Environment