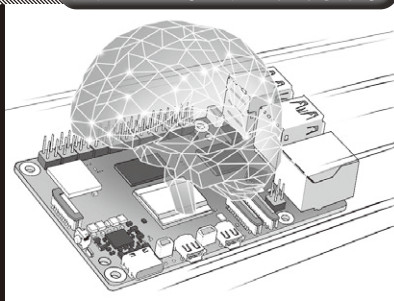


リソースの限られた環境での実行性能を徹底解析!

ラズパイの限界に挑戦!

ローカルLLM
動作検証レポート

新連載

第1回 スワップ発生時のSDカード/SSDの性能差を調べる

ご購入はこちら

澁谷 慎太郎

表1 ハードウェアの主な仕様と構成

	ラズベリー・パイ5と SSD	ラズベリー・パイ5 とSDカード
メイン・メモリ	8Gバイト	
プロセッサ	BCM2712	
アーキテクチャ	Arm Cortex-A76 クアッド・コア	
クロック周波数	2.4GHz	
ストレージ	Raspberry Pi SSD Kit 512Gバイト	Extreme Pro SDXC 128Gバイト A2 DSQXCQ-128G- GN6MA (サンディスク)
冷却	ラズベリー・パイ5用 公式アクティブ・クーラ	

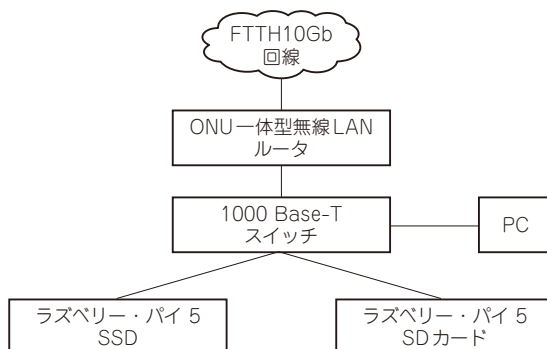


図1 実験環境のネットワーク構成

● 連載概要

現在、ChatGPT、Copilotなど数多くの生成AIがリリースされています。ChatGPTなどブラウザやAPI経由で使えるものだけでなく、ローカル環境で使えるLLM (Large Language Models, 大規模言語モデル) も、潤沢なハードウェア環境が必要なものや、ラズベリー・パイなどでも動作する軽量なもの、もっと軽量なもの、コード生成や数学、多言語に対応したものなど、さまざまなものがあります¹⁾。そこで、本連載では、ラズベリー・パイにローカルLLM実行環境を構築して、ローカルLLMがどこまで動くのか実験します。

● 今回の目的

2025年12月号 特集第4部では、ラズベリー・パイ5、4B、3B+で、5種のさまざまなサイズのLLM (Gemma, Qwen, Llama, Granite, Phi4-mini) がどの程度の速度で動くのか実験しました。このときは意図的にスワップを起こさせることはしませんでした。スワップが発生するとパフォーマンスが落ちることが予想されます。そこで今回は、わざとスワップが発生する条件を設定し、SSDとSDカードの差が果たして出るのか見てみたいと思います。加えて、実用パ

フォーマンスを度外視して大きなモデルも動かしてみたいと思います。そこで今回は、次の2つの実験をラズベリー・パイ5で行います。

実験①：メモリ制限を加え、意図的にスワップを発生させてSDカードとSSDのパフォーマンスを比較する

実験②：各種の設定やスワップ・サイズなどを調整して、20Bなど大きいサイズのモデルがどこまで動くか試してみる

実験環境

● ハードウェア環境

今回はSDカードを使ったときとSSDを使ったときの違いを検証するため、ラズベリー・パイ5の8Gバイトのモデルで、A2クラスで128GバイトのSDカードを搭載した機種と、NVMe SSD 512Gバイトを搭載した2つの機種を使います(表1)。SDカードには、Extreme Pro SDXC 128Gバイト A2 DSQXCQ-128G-GN6MA (サンディスク) を使用しています。このSDカードは、ラズベリー・パイとの相性も良く、コスト・パフォーマンスに優れています^{注1}。NVMe

注1：本誌2025年6月号p.136のコラムで、筆者がSDカードのパフォーマンス比較を書いています。