🎯 COLD START vs CACHE WARMING - Detailed Explanation
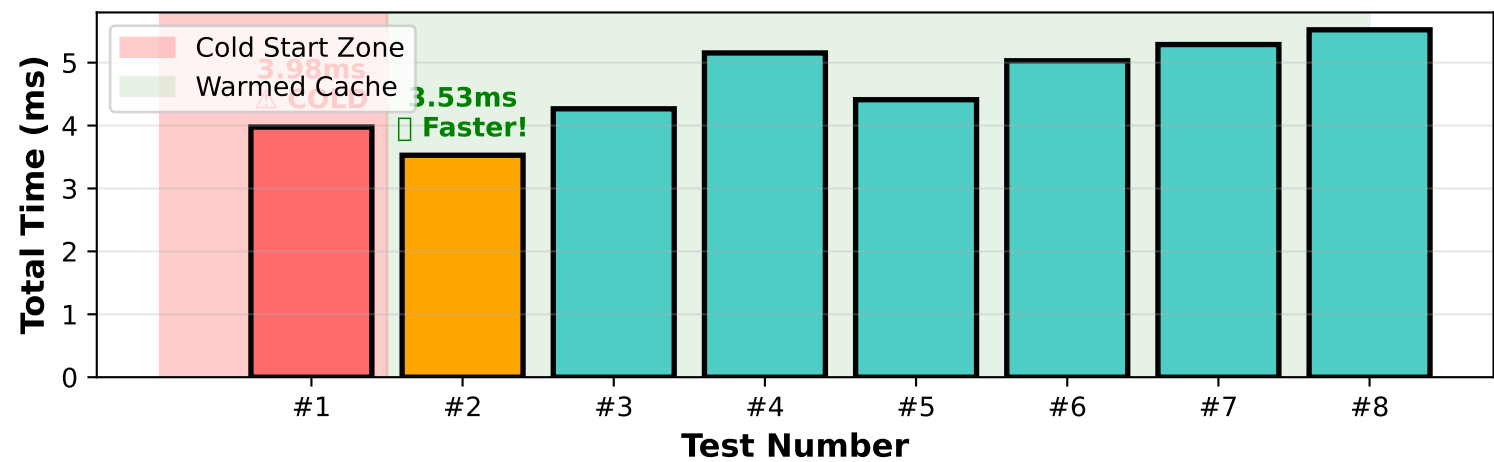## Why the first test is slower and how to fix it
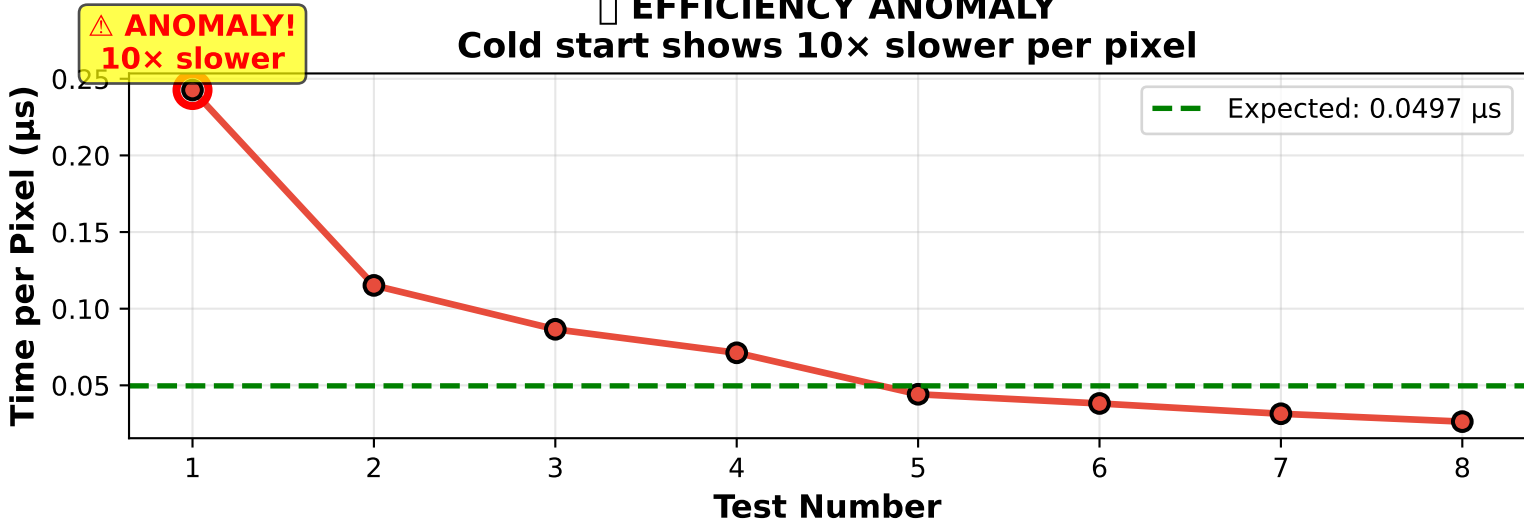
🐌 COLD START EFFECT
First test is SLOWER

- Cold Start Zone
- Warmed Cache

3.98ms ❄ COLD
3.53ms ⚡ Faster!

Total Time (ms) / Test Number (#1 – #8)

📉 EFFICIENCY ANOMALY
Cold start shows 10× slower per pixel

⚠ ANOMALY! 10× slower

Expected: 0.0497 μs

Time per Pixel (μs) / Test Number (1 – 8)

❄ COLD START (Lần chạy đầu tiên)

1. 📀 Load thư viện từ disk:
   • numpy, PIL, matplotlib
   • Phải đọc từ SSD/HDD → CHẬM

2. 🧊 CPU Cache Miss:
   • L1/L2/L3 cache đều trống
   • Phải fetch từ RAM → CHẬM

3. 🐍 Python Initialization:
   • JIT compilation
   • Module imports
   • Function lookups → CHẬM

4. 💾 OS Overhead:
   • Memory allocation
   • Page faults
   • Disk I/O → CHẬM

📊 KẾT QUẢ: Test #1 = 2.55ms
(Chậm hơn 3× so với expected!)

🔥 CACHE WARMED (Sau lần chạy đầu)

1. 🚀 Thư viện đã trong RAM:
   • numpy, PIL đã loaded
   • Không cần đọc disk → NHANH

2. 🔥 CPU Cache Hot:
   • L1/L2/L3 đã có data
   • Cache hit rate cao → NHANH

3. ⚡ Python đã init:
   • Functions đã compiled
   • Imports đã cached → NHANH

4. ✅ OS Cache Ready:
   • Memory pages allocated
   • No page faults → NHANH

💡 GIẢI PHÁP: CACHE WARMING

🎯 Strategy:
   1. Chạy 1 test WARMUP trước
   2. KHÔNG tính kết quả warmup
   3. Chỉ tính tests sau warmup

🔧 Implementation:

```
# Warmup (không tính)
print("🔥 Cache warming...")
run_single_test(size=256)
results = []  # Clear!

# Real benchmark (có tính)
for size in [128, 256, ..., 1024]:
    result = run_single_test(size)
    results.append(result)
```

✅ KẾT QUẢ:
   • Test đầu tiên đã NHANH
   • Không có cold start anomaly
   • Dữ liệu chính xác hơn!
   • Benchmark đáng tin cậy hơn!

⚡ IMPACT OF CACHE WARMING
Warmup eliminates cold start

- 🔴 Without Warmup
- 🟢 With Warmup

2.55ms ⚠ Slow — Lần 1 (Cold)
2.46ms Better — Lần 2 (Warm)
4.95ms Stable — Lần 3-8 (Stable)

Average Time (ms)