```
(env311) PS F:\test> & F:/test/env311/Scripts/python.exe f:/test/net/zero copy breakthrough.py
INFO:__main__: \ 初始化高級零拷貝環境...
INFO:__main__: ☑ RetryIX SVM Core 已載入: OpenCL.dll
INFO:__main__: SVM粗粒度: ☑
INFO:__main__: SVM細粒度: ✓
INFO:__main__: SVM原子操作: 🗙
INFO:__main__:
             統一內存: 🗙
INFO: main :
              內存映射: 🔽
INFO: main : ☑ 環境初始化完成
INFO: main : 設備: gfx1010:xnack-
INFO:__main__:
               支持特性: ['svm_coarse', 'svm_fine', 'svm_atomics', 'unified_memory', 'map_buffer']
INFO:__main__: √ 初始化高級內存管理...
INFO: main : 初始化HOST PTR池...
INFO: main :
             初始化SVM池...
INFO:__main__:
             RetryIX SVM池初始化成功!分配 3 個SVM buffer
INFO:__main__:
             初始化映射內存池...
INFO: main : 6 開始全面零拷貝技術對比測試
INFO: __main__:
📊 測試策略: ['traditional', 'use_host_ptr', 'ultra_fast_host_ptr', 'map_buffer', 'svm_coarse']
INFO: main : 📊 測試大小: [1024, 10240, 102400, 1024000]
INFO:__main__:⋞ RetryIX SVM 測試已啟用
INFO: __main__:
▲ 測試策略: traditional
INFO: main :
               測試大小: 1024 元素 (4.0 KB)
F:\test\env311\Lib\site-packages\pyopencl\cache.py:496: CompilerWarning: Non-empty compiler output
encountered. Set the environment variable PYOPENCL COMPILER OUTPUT=1 to see more.
 _create_built_program_from_source_cached(
                總時間: 520.5 μs 🔸 亞毫秒級!
INFO: main :
INFO:__main__:
                 其中內核: 171.8 μs (33.0%)
INFO: main :
                 數據準備: 72.4 μs 🔸 超高效!
             測試大小: 10240 元素 (40.0 KB)
INFO: main :
INFO: main :
                總時間: 341.4 us ∳ 亞毫秒級!
INFO: main :
                 其中內核: 142.3 μs (41.7%)
INFO:__main__:
                 數據準備: 57.7 μs 🔸 超高效!
             測試大小: 102400 元素 (400.0 KB)
INFO: __main__:
INFO:__main__:
                 總時間: 696.5 μs 🔸 亞毫秒級!
INFO: main :
                 其中內核: 229.3 μs (32.9%)
INFO:__main__:
               測試大小: 1024000 元素 (4000.0 KB)
INFO: main :
                總時間: 5727.7 μs (5.73 ms)
INFO:__main__:
                 其中內核: 865.7 μs (15.1%)
INFO:__main__:
▲ 測試策略: use_host_ptr
INFO:__main__:
              測試大小: 1024 元素 (4.0 KB)
INFO: main :
                 總時間: 202.4 µs ∳ 亞毫秒級!
INFO:__main__:
                其中內核: 172.8 μs (85.4%)
INFO:__main_ :
                 數據準備: 7.8 µs 		 超高效!
             測試大小: 10240 元素 (40.0 KB)
INFO:__main__:
INFO: main :
                總時間: 119.7 μs 🔸 亞毫秒級!
INFO:__main__:
                 其中內核: 105.8 μs (88.5%)
INFO:__main__:
                數據準備: 5.5 μs 🔸 超高效!
             測試大小: 102400 元素 (400.0 KB)
INFO: main :
INFO: main :
                 總時間: 255.5 μs 🔸 亞毫秒級!
INFO: main :
                 其中內核: 224.9 μs (88.0%)
INFO:__main__:
                 數據準備: 15.5 μs 🔸 超高效!
INFO: main :
               測試大小: 1024000 元素 (4000.0 KB)
INFO: main :
                總時間: 1621.2 μs (1.62 ms)
INFO: __main_
INFO: __main_
                 其中內核: 1216.9 μs (75.1%)
▲ 測試策略: ultra_fast_host_ptr
INFO: main :
               測試大小: 1024 元素 (4.0 KB)
INFO:__main__:
                 總時間: 100.8 μs 🔸 亞毫秒級!
INFO: main :
                 其中內核: 94.9 μs (94.2%)
```

```
INFO: main :
               數據準備: 2.7 μs 🔸 超高效!
INFO:__main__:
             測試大小: 10240 元素 (40.0 KB)
INFO: main :
               總時間: 103.0 μs 🔸 亞毫秒級!
INFO:__main__:
               其中內核: 97.2 μs (94.3%)
               數據準備: 3.5 µs 		 超高效!
INFO: main :
INFO: main :
             測試大小: 102400 元素 (400.0 KB)
INFO:__main__:
               總時間: 220.2 μs 🔸 亞毫秒級!
               其中內核: 203.2 μs (92.3%)
INFO:__main__:
INFO: main :
               數據準備: 14.7 μs 🔸 超高效!
INFO: main :
             測試大小: 1024000 元素 (4000.0 KB)
INFO:__main__:
               總時間: 1538.4 μs (1.54 ms)
INFO:__main__:
               其中內核: 1212.3 μs (78.8%)
INFO: main :
▲ 測試策略: map buffer
INFO: main :
             測試大小: 1024 元素 (4.0 KB)
INFO: main :
               總時間: 846.3 μs 🔸 亞毫秒級!
INFO: main :
               其中內核: 187.7 μs (22.2%)
INFO:__main__:
             測試大小: 10240 元素 (40.0 KB)
INFO: main :
               總時間: 252.9 μs 🔸 亞毫秒級!
INFO: __main__:
               其中內核: 155.2 μs (61.4%)
               數據準備: 29.4 μs 🔸 超高效!
INFO: main :
INFO: main :
             測試大小: 102400 元素 (400.0 KB)
INFO:__main__:
               總時間: 443.3 us ∮ 亞毫秒級!
               其中內核: 266.0 μs (60.0%)
INFO: __main__:
INFO: main :
               數據準備: 46.3 μs 🔸 超高效!
INFO:__main__:
             測試大小: 1024000 元素 (4000.0 KB)
INFO:__main__:
               總時間: 3364.9 μs (3.36 ms)
INFO: main :
               其中內核: 1518.7 μs (45.1%)
INFO: __main__:
INFO: main : 測試大小: 1024 元素 (4.0 KB)
INFO:__main__:
               總時間: 140.0 μs 🔸 亞毫秒級!
INFO: __main__:
               其中內核: 101.7 μs (72.7%)
INFO:__main__:
               數據準備: 16.0 μs 🔸 超高效!
INFO:__main__:
             測試大小: 10240 元素 (40.0 KB)
INFO:__main__:
               總時間: 134.2 μs 🔸 亞毫秒級!
               其中內核: 96.9 μs (72.2%)
INFO: main :
INFO:__main__:
               數據準備: 17.2 μs 🔸 超高效!
             測試大小: 102400 元素 (400.0 KB)
INFO:__main__:
INFO:__main__:
               總時間: 298.1 μs 🔸 亞毫秒級!
INFO: main :
               其中內核: 216.7 μs (72.7%)
INFO: main :
               數據準備: 51.3 µs ∳ 超高效!
INFO:__main__:
             測試大小: 1024000 元素 (4000.0 KB)
INFO:__main__:
               總時間: 1493.8 μs (1.49 ms)
INFO: __main__:
               其中內核: 1133.3 μs (75.9%)
INFO: main :
______
INFO:__main__: 

高級零拷貝技術效果分析
INFO: __main__:
📊 性能對比表 (時間單位: 微秒)
INFO:__main__:策略\大小
                                  1024(4KB) 10240(40KB) 102400(400KB)1024000(4000KB)
INFO: __main__:-----
INFO: __main__:traditional
                              521(1.00x)
                                           341(1.00x)
                                                        696(1.00x)
                                                                   5728(1.00x)
INFO: __main__:use_host_ptr
                              202(2.57x)
                                           120(2.85x)
                                                                   1621(3.53x)
                                                        256(2.73x)
INFO: main__:ultra_fast_host_ptr
                              101(5.16x)
                                           103(3.31x)
                                                        220(3.16x)
                                                                   1538(3.72x)
INFO:__main__:map_buffer
                              846(0.62x)
                                           253(1.35x)
                                                        443(1.57x)
                                                                   3365(1.70x)
INFO:__main__:svm_coarse
                               140(3.72x)
                                           134(2.54x)
                                                        298(2.34x)
                                                                   1494(3.83x)
INFO:__main__:
🚼 最佳策略分析:
INFO: __main__:
  數據大小 1024 (4.0 KB):
INFO: main :
               最佳策略: ultra fast host ptr
INFO: main :
               總時間: 100.8 μs
```

```
LLM Inference Optimization Techniques - Claude
2025/8/7 晚上10:58
                   計算占比: 94.2%
 INFO:__main__:
 INFO:__main__:
                   性能提升: 5.16倍
 INFO: main :
    數據大小 10240 (40.0 KB):
 INFO:__main__:
                   最佳策略: ultra_fast_host_ptr
                   總時間: 103.0 μs
 INFO: __main__:
 INFO:__main__:
                   計算占比: 94.3%
 INFO:__main__:
                   性能提升: 3.31倍
 INFO: __main__:
    數據大小 102400 (400.0 KB):
 INFO:__main__:
                   最佳策略: ultra_fast_host_ptr
 INFO:__main :
                   總時間: 220.2 μs
 INFO: main :
                   計算占比: 92.3%
 INFO:__main__:
                   性能提升: 3.16倍
 INFO:__main__:
    數據大小 1024000 (4000.0 KB):
 INFO: main :
                   最佳策略: svm_coarse
 INFO:__main__:
                   總時間: 1493.8 μs
 INFO: main :
                   計算占比: 75.9%
                   性能提升: 3.83倍
 INFO:__main__:
 INFO:__main__:
  亞毫秒級技術突破總結:
 INFO:__main__: 型 亞毫秒級測試占比: 15/20 (75.0%)
 INFO:__main__: ∳ 最快記錄: ultra_fast_host_ptr @ 1024元素 = 100.8 μs
 INFO:__main__: ☑ 最高計算占比策略: ultra_fast_host_ptr (89.9%)
 INFO: main : 亞毫秒級突破成功!數據傳輸延遲基本消除
 INFO: main :
  ♀ 性能推薦方案:
 INFO: __main__:
                 小數據(< 10KB): ultra_fast_host_ptr (101μs ∳ , 計算占比94.2%)
                 中數據(10KB-100KB): ultra_fast_host_ptr (103μs 🔸 ,計算占比94.3%)
 INFO: __main__:
                 大數據(100KB-1MB): ultra_fast_host_ptr (220μs ∳, 計算占比92.3%)
 INFO: __main__:
 INFO:__main__:
                 超大數據(> 1MB): svm_coarse (1.49ms, 計算占比75.9%)
 INFO: __main__:
```

👺 高級零拷貝技術測試完成! (env311) PS F:\test>