

Relazione di progetto

Corso di High Performance Computing

A.A. 2024-2025

Alessandro Valmori
February 14, 2025

1. Introduzione

Si sceglie di implementare la parallelizzazione dell'algoritmo Skyline in due diverse versioni, utilizzando prima OpenMP e poi MPI, due tecniche di parallelizzazione rispettivamente a memoria condivisa e a memoria distribuita. L'approccio alla parallelizzazione utilizzando i due paradigmi condivide la fase di partizionamento del dominio iniziale fra i processi, utilizzando in particolare un partizionamento statico a grana grossa.

2. Versione OpenMP

Per la versione a memoria condivisa utilizzando OpenMP, si sceglie di parallelizzare il confronto fra un punto e tutti gli altri. In particolare, il ciclo esterno, che itera attraverso i punti da esaminare, è stato parallelizzato utilizzando la direttiva `pragma omp parallel for`. Per calcolare il numero di punti nello skyline set (l'insieme di punti non dominati), è stato applicato il pattern di riduzione. Nella riduzione, il valore di r (il contatore che tiene traccia dei punti nello skyline) viene modificato da più thread senza conflitti grazie alla direttiva `reduction(-:r)`. Ogni volta che un punto è trovato dominante su un altro, il punto dominante viene segnato come non appartenente allo skyline (modificando l'array s), e il contatore r viene decrementato. Per evitare il controllo ripetuto di punti già rimossi dallo skyline, ogni punto viene esaminato una sola volta, e non appena si trova un punto che domina un altro, quest'ultimo viene rimosso.

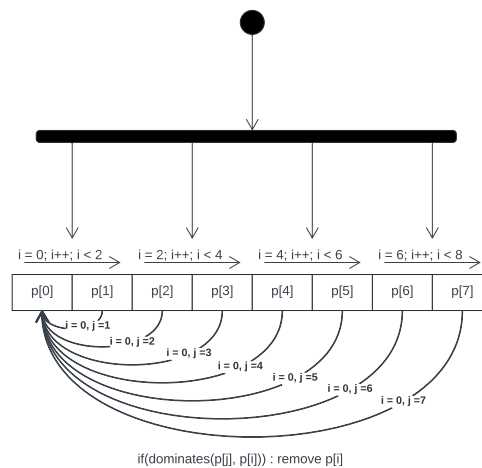


FIGURE 1. Schema concettuale del processo svolto dall'algoritmo parallelo con OpenMP

3. Versione MPI

Per la versione a memoria distribuita utilizzando MPI è stata implementata una soluzione che presenta un processo più scandito e verboso, che comincia con la suddivisione del dominio iniziale in parti il quanto più possibile eguali fra i processi. Questo partizionamento avviene staticamente e grossolanamente dopo che il processo master ha letto l'insieme di punti da standard input.

Una volta che ciascun processo detiene la propria partizione dell'input set originale, ognuno di essi applica la versione seriale dell'algoritmo Skyline a quest'ultima, ottenendo quindi un sottoinsieme di tale partizione costituito dai punti che non sono dominati da alcun altro punto all'interno della partizione.

In seguito, il processo master esegue una MPI_Gatherv di ciascuno skyline set locale in un unico array, di cui ne viene poi distribuita una copia a ciascun processo, in modo da consentire il confronto fra lo skyline set locale e tale array. Infine, dopo che ciascun processo ha rimosso dalla propria partizione tutti i punti dominati, avviene una chiamata alla funzione MPI_Reduce per ottenere il numero totale di punti presenti nello skyline finale, seguita da un'ulteriore chiamata a MPI_Gatherv per concatenare tutti i singoli array locali nello skyline finale.

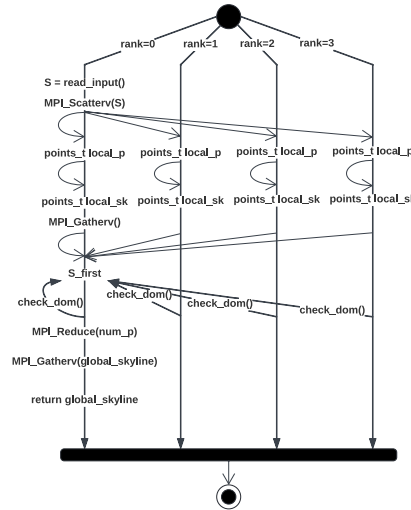


FIGURE 2. Schema concettuale del processo svolto dall'algoritmo parallelo con MPI

4. Prestazioni

Per valutare l'efficacia dell'implementazione parallela dell'algoritmo Skyline, sono stati condotti test mirati all'analisi della scalabilità forte e della scalabilità debole. L'obiettivo è verificare come il tempo di esecuzione varia al variare del numero di processi MPI, sia mantenendo fisso il carico di lavoro totale (strong scaling) sia mantenendo costante il carico per processo (weak scaling).

I test sono stati eseguiti su una macchina con processore Intel Core i7-1165G7 di undicesima generazione (4 core fisici, 8 thread) e 16GB di RAM. L'algoritmo è stato eseguito 10 volte per ciascuna configurazione sperimentale, e il valore finale riportato corrisponde alla media dei tempi di esecuzione ottenuti, riducendo così l'effetto di eventuali variazioni dovute a fattori esterni, come il sistema operativo o altri processi in esecuzione.

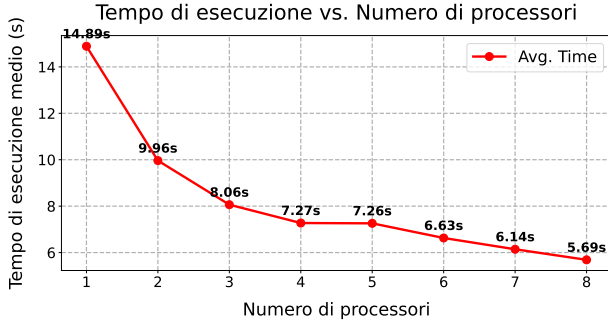
Per raccogliere i dati, è stato realizzato uno script bash che automatizza l'esecuzione delle prove. Lo script si occupa di lanciare il programma con diverse configurazioni di numero di processi MPI e dataset di input, redirezionando l'output standard e gli errori in un file di log per facilitare l'analisi successiva. Inoltre, un insieme di script Python è stato utilizzato per elaborare i risultati e generare grafici rappresentativi delle performance ottenute.

4.1. Strong scaling

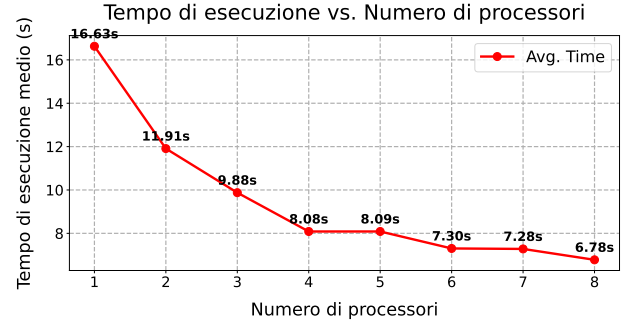
Lo strong scaling analizza il comportamento dell'algoritmo mantenendo costante la dimensione del problema e aumentando progressivamente il numero di processi MPI utilizzati per l'esecuzione. L'obiettivo è osservare come varia il tempo di esecuzione e verificare se l'algoritmo è in grado di sfruttare in modo efficiente un numero crescente di processi.

Per questi test, è stato scelto un dataset di dimensione fissa, rappresentativo di un caso d'uso realistico, e il numero di processi è stato variato da 1 a 8. Questo range è stato selezionato per valutare il comportamento dell'algoritmo sia in condizioni di esecuzione seriale (1 processo) sia quando si sfruttano tutte le risorse disponibili sulla macchina di test (8 processi, pari al numero massimo di thread del processore).

Idealmente, in uno scenario di perfetto strong scaling, il tempo di esecuzione dovrebbe ridursi proporzionalmente all'aumento del numero di processi. Tuttavia, in una situazione reale, il comportamento può essere influenzato da diversi fattori, tra cui l'overhead della comunicazione tra processi, il bilanciamento del carico e le latenze introdotte dalla gestione della memoria.



A. Tempo di esecuzione in funzione del numero di processi (Strong Scaling) di OpenMP



B. Tempo di esecuzione in funzione del numero di processi (Strong Scaling) di MPI

FIGURE 3. Analisi dello strong scaling dell'algoritmo

4.2. Scalabilità Debole

La scalabilità debole valuta come varia il tempo di esecuzione dell'algoritmo quando sia il numero di processi (o thread) che la dimensione del problema aumentano proporzionalmente. L'obiettivo è determinare se l'efficienza parallela rimane stabile con l'aggiunta di risorse computazionali.

4.2.1. Metodologia di Scalabilità

La complessità parallela dell'algoritmo skyline può essere espressa come:

$$O\left(\frac{N^2 D}{p}\right)$$

dove:

- N = numero di punti,
- D = numero di dimensioni (assunto costante),
- p = numero di processi (o thread).

Ogni thread è responsabile di un sottoinsieme dei confronti totali N^2 , dato da:

$$\frac{N^2}{p}$$

Per mantenere il carico di lavoro per processo costante, dobbiamo garantire che questo termine rimanga invariato all'aumentare di p . Ciò porta alla condizione:

$$\frac{(N')^2}{p'} = \frac{N^2}{p}$$

dove:

- N' è la nuova dimensione del problema con p' thread,
- N è la dimensione originale con p thread.

4.2.2. Derivazione del Fattore di Scalabilità

Riorganizzando la formula per N' , otteniamo:

$$N' = N \times \sqrt{\frac{p'}{p}}$$

Questo significa che, per mantenere il carico di lavoro per processo costante, la dimensione del problema N deve essere aumentata di un fattore:

$$\sqrt{p}$$

Ad esempio:

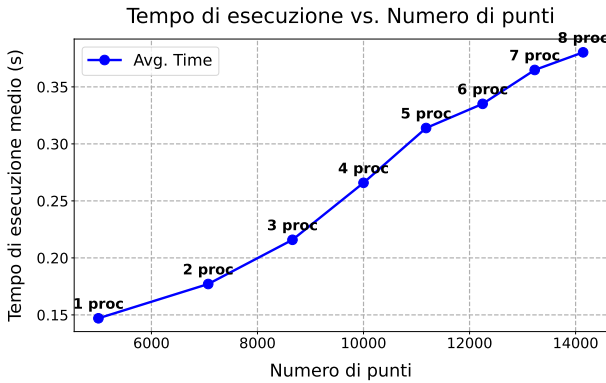
- Se il numero di thread p raddoppia ($p' = 2p$), la dimensione del problema deve aumentare di un fattore:

$$N' = N \times \sqrt{2}$$

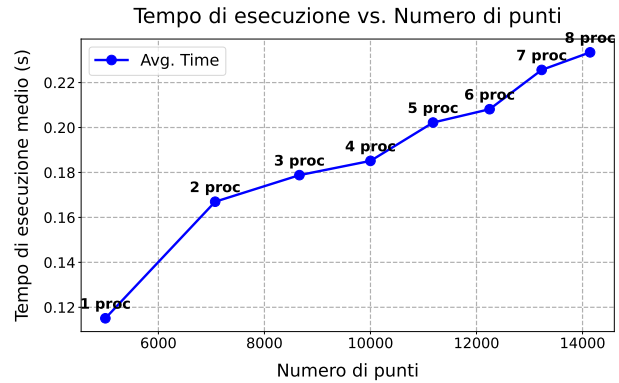
- Se p aumenta di un fattore 4, allora $N' = N \times 2$.
- Se p aumenta di un fattore 9, allora $N' = N \times 3$.

4.2.3. Risultati Sperimentali

Per analizzare la scalabilità debole, abbiamo eseguito test in cui sia il numero di thread p che la dimensione del dataset N sono stati aumentati secondo la regola di scalabilità derivata sopra.



A. Risultati della scalabilità debole con OpenMP.



B. Risultati della scalabilità debole con MPI.

FIGURE 4. Analisi della scalabilità debole per le implementazioni OpenMP e MPI.

I risultati in Figura 4 mostrano che:

- OpenMP: Il tempo di esecuzione rimane relativamente stabile per un numero ridotto di thread, ma aumenta con valori più elevati a causa dell'overhead di sincronizzazione e della contesa della memoria condivisa.
- MPI: La performance scala meglio inizialmente, ma l'overhead di comunicazione diventa significativo all'aumentare del numero di processi.

In uno scenario ideale di scalabilità debole, il tempo di esecuzione dovrebbe rimanere costante con l'aumento simultaneo della dimensione del problema e delle risorse computazionali. Tuttavia, limitazioni reali come contesa della memoria, overhead di comunicazione e squilibri nel bilanciamento del carico contribuiscono a deviazioni da questo comportamento ideale.

5. Conclusioni