

# Comprehensive Benchmark Report: File Search Algorithms

Generated: 2025-11-02 16:44:19

Tested file sizes: 10000, 50000, 100000, 250000, 500000, 1000000

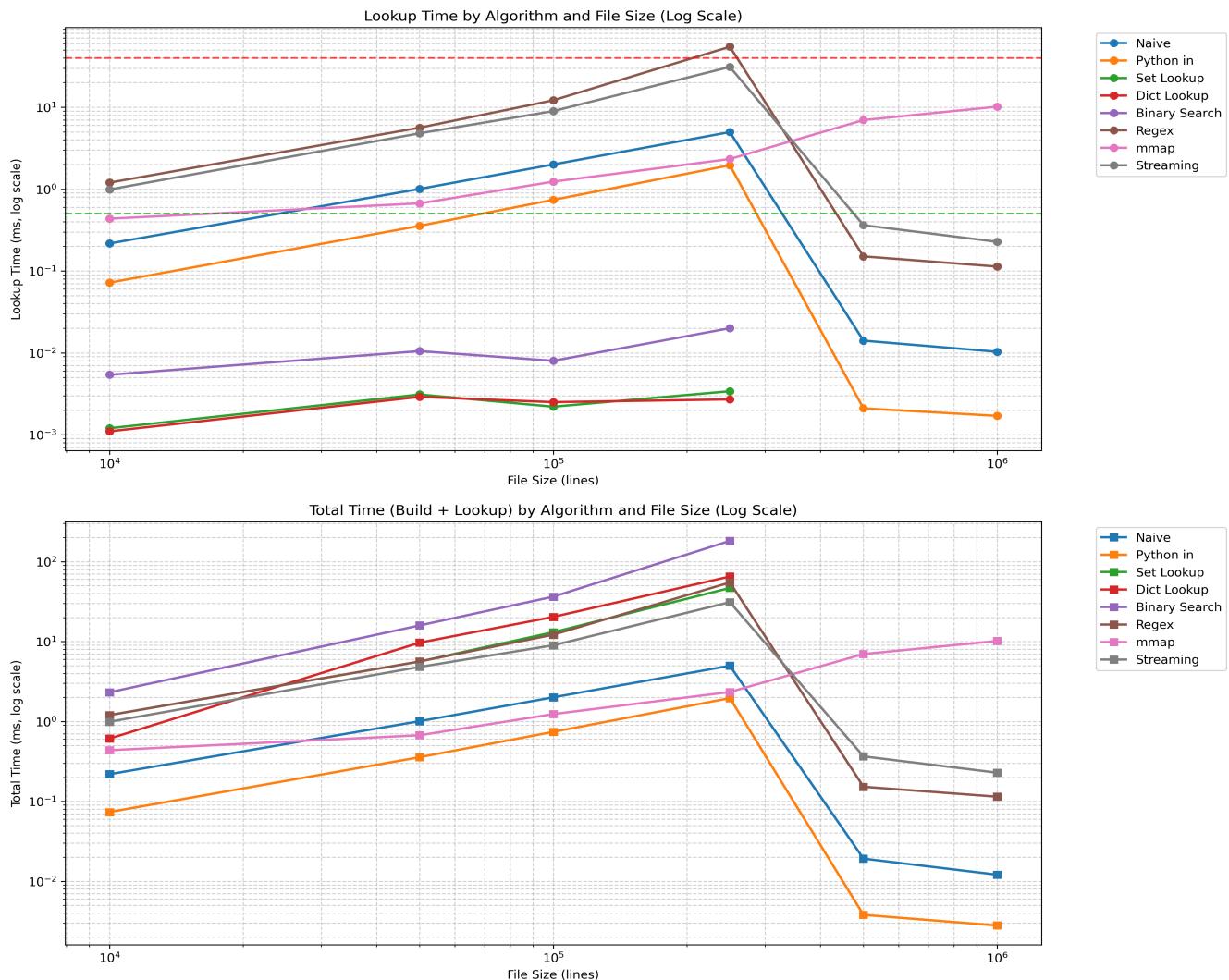
## Executive Summary

**Recommended algorithms (meet 0.5ms requirement):** Dict Lookup, Set Lookup, Binary Search

## Performance Requirement Validation (250k lines)

Algorithm	Scenario	Avg Time (ms)	Requirement	Status
Naive	REREAD_ON_QUERY=True	6.431	40ms	PASS
Python in	REREAD_ON_QUERY=True	1.983	40ms	PASS
Set Lookup	REREAD_ON_QUERY=False	0.006	0.5ms	PASS
Binary Search	REREAD_ON_QUERY=False	0.018	0.5ms	PASS
mmap	REREAD_ON_QUERY=True	1.679	40ms	PASS
Streaming	REREAD_ON_QUERY=True	21.404	40ms	PASS

## Performance Analysis Charts



## Algorithm Performance Summary

Algorithm	Avg Lookup (ms)	Max File Size	Meets 0.5ms	Scalability
Naive	1.370	1,000,000	NO	Good
Python in	0.521	1,000,000	NO	Good
Set Lookup	0.002	250,000	YES	N/A
Dict Lookup	0.002	250,000	YES	N/A
Binary Search	0.011	250,000	YES	N/A
Regex	12.346	1,000,000	NO	Good
mmap	3.634	1,000,000	NO	Moderate
Streaming	7.735	1,000,000	NO	Good

## Detailed Results (Sample)

Algorithm	File Size	Found	Build Time (ms)	Lookup Time (ms)	Total Time (ms)
Naive	10000	1	0.0021	0.2168	0.2189
Python in	10000	1	0.0015	0.072	0.0735
Set Lookup	10000	1	1.2012	0.0012	1.2024
Dict Lookup	10000	1	0.6092	0.0011	0.6103
Binary Search	10000	1	2.3118	0.0054	2.3172
Regex	10000	1	0.0013	1.1995	1.2008
mmap	10000	1	0.0012	0.4351	0.4363
Streaming	10000	1	0.0017	0.9895	0.9912
Naive	50000	1	0.002	1.0055	1.0075
Python in	50000	1	0.0012	0.3558	0.357
Set Lookup	50000	1	5.6186	0.0031	5.6217
Dict Lookup	50000	1	9.6767	0.0029	9.6796
Binary Search	50000	1	15.9	0.0105	15.9105
Regex	50000	1	0.003	5.6312	5.6342
mmap	50000	1	0.0014	0.67	0.6714