

Median of a Running Stream

Given a running stream of integers, design an algorithm to print the **median** of the numbers in the stream each time a new integer is added to the stream.

Example

| | | | | | | | | | | |
|--------|---|---|---|-----|---|-----|---|-----|---|-----|
| Input | ↓ | | | | | | | | | |
| Stream | 4 | 6 | 2 | 5 | 3 | 1 | 7 | 9 | 8 | ... |
| Median | 4 | 5 | 4 | 4.5 | 4 | 3.5 | 4 | 4.5 | 5 | ... |

[4]
[4, 6]
[2, 4, 6]
[2, 4, 5, 6]
[2, 3, 4, 5, 6]
[1, 2, 3, 4, 5, 6]

assume
size of stream
is 'n'

Brute force
→ read a val. from the stream → C
→ insert it into sorted seq. → n
→ find me median → C
} n/2 =

x > median

5 4 1 2 3 [x]

median = 3

x =
eg: 10
[1, 2, 3, 4, 5, 10]
3.5

[1, 2, 3, 4, 5]
viz

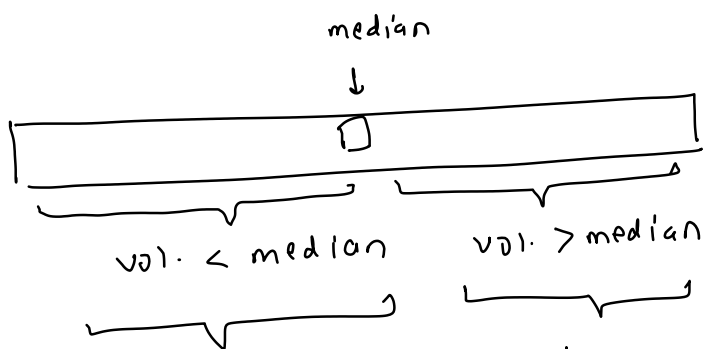
if x is > median then median ↑ es

5 4 1 2 3 [x] median = 3

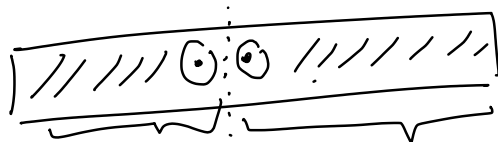
x < median

!!
median ↓ es

eg: x = 0
[0, 1, 2, 3, 4, 5]
2.5 < 3

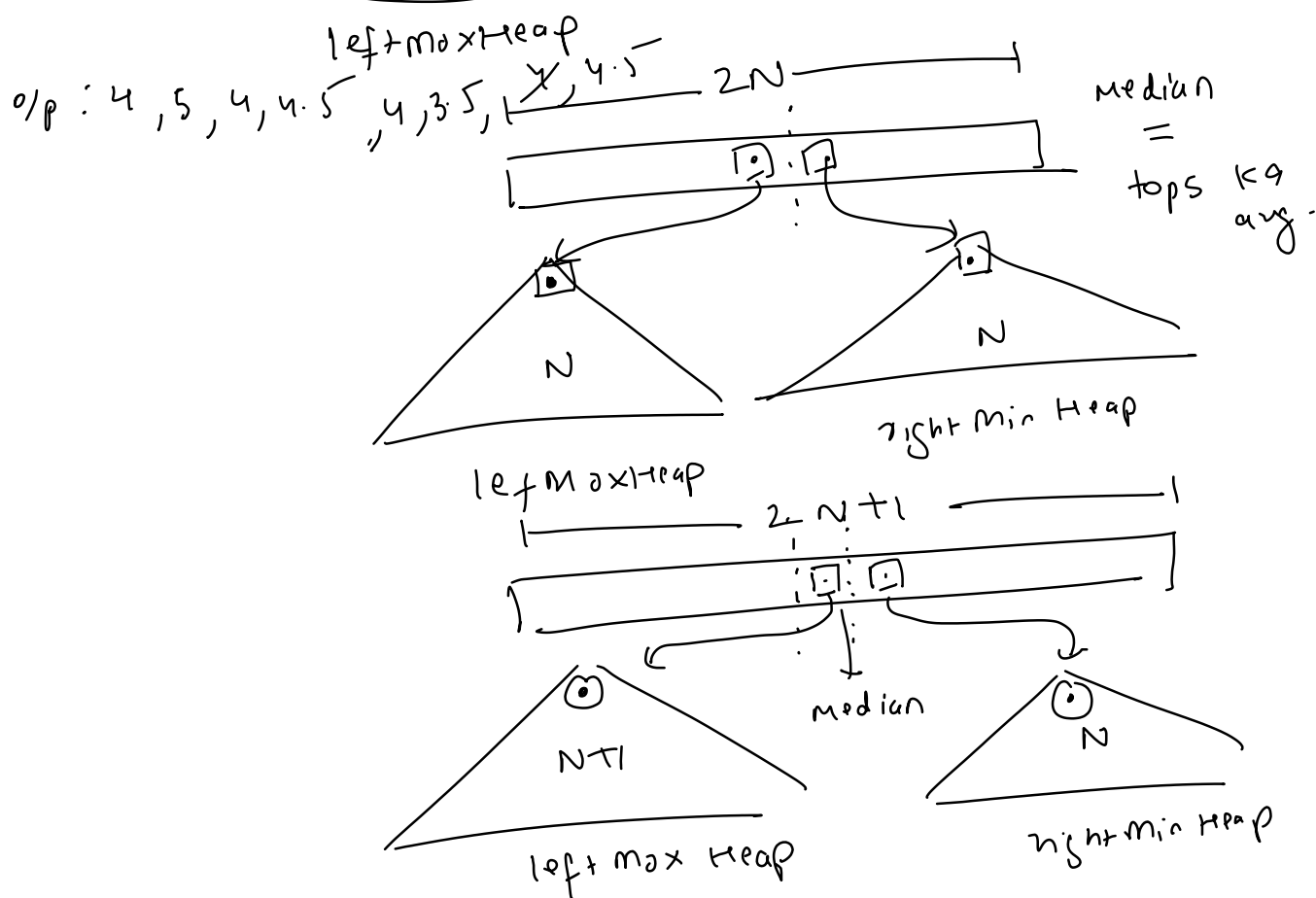


build a maxheap to track these values which are < median
 build a minheap to track these values > median
 leftMaxHeap rightMinHeap



| | | | | | | | | | |
|--------|---|---|---|---|---|---|---|---|---|
| Stream | 4 | 6 | 2 | 5 | 3 | 1 | 7 | 9 | 8 |
|--------|---|---|---|---|---|---|---|---|---|

median = ~~4~~ 5 ~~4~~ ~~4~~ 5 ~~4~~ 3 5 ~~4~~ ~~4.5~~



| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| ↑ | ↑ | ↑ | ↑ | — | — | — | — | — | — |
| x | ✓ | x | ✓ | | | | | | |

$\frac{n}{2}$ rebalance + $\frac{n}{2}$ no rebalance
 push + pop push

$$\begin{array}{ccc}
 \downarrow & & \downarrow \\
 n/2 \cdot \log n/2 + n/2 \cdot \log n/2 & & n/2 \cdot \log n/2 \\
 \underbrace{\hspace{10em}} & & \\
 3n/2 \log n/2 & \sim & \underline{n \log n}
 \end{array}$$