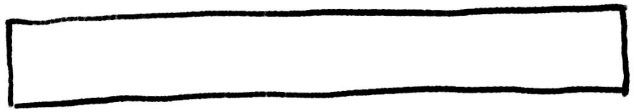
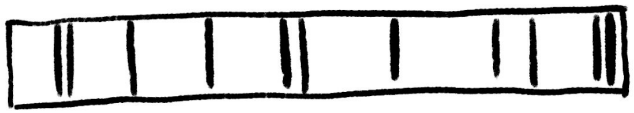
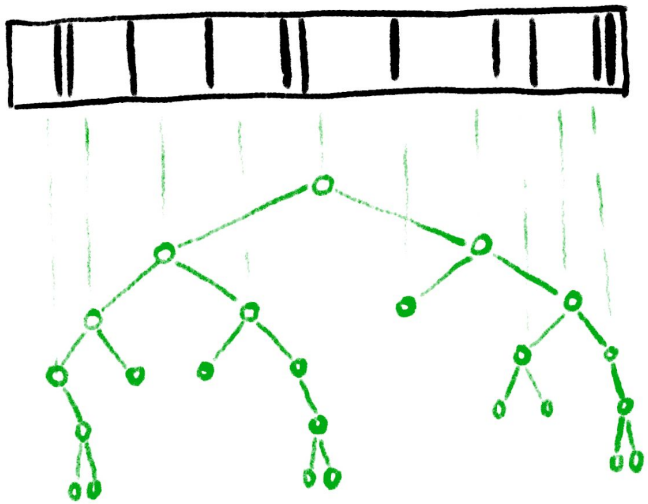
The power of two choices:
Why the Kademlia binary tree isn’t balanced and what we can do about it

Petar Maymounkov
LabWeek 2022 libp2p

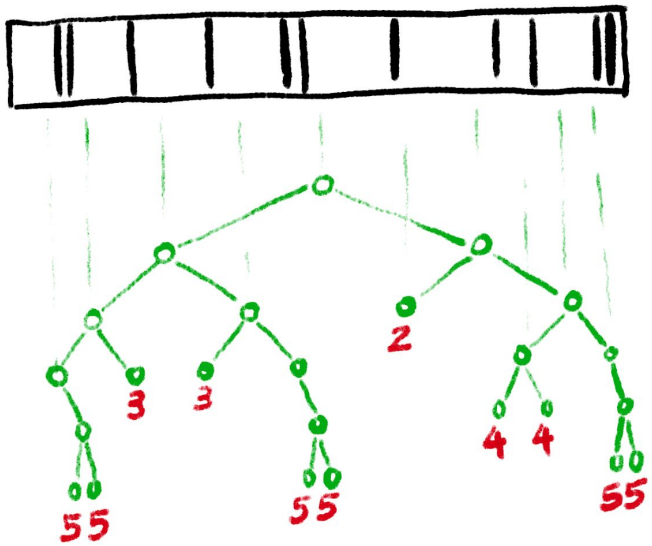


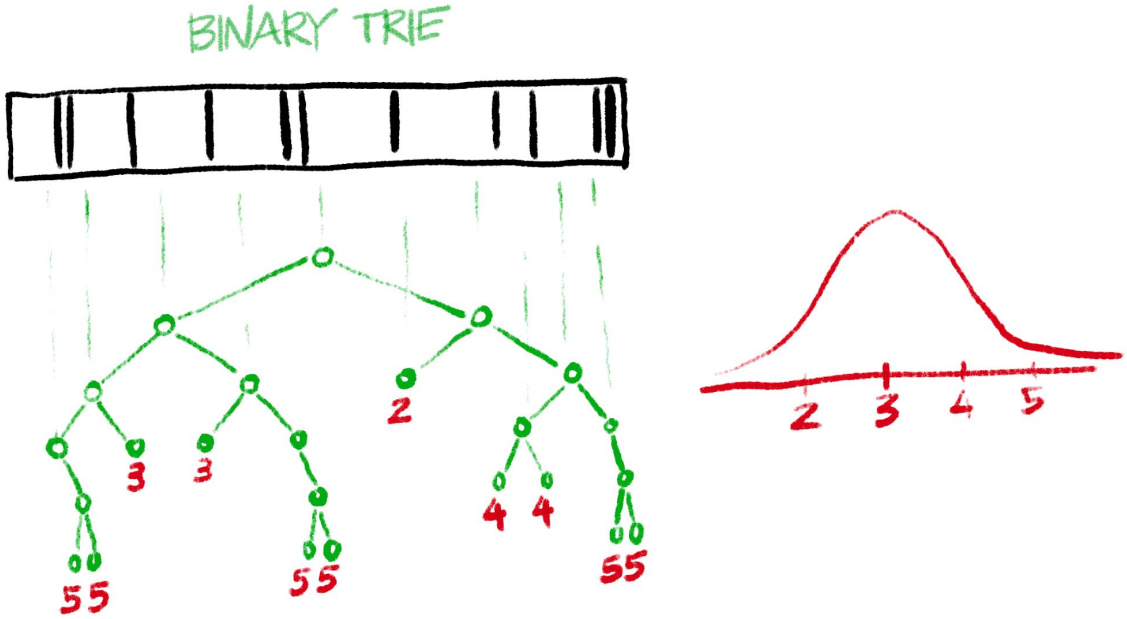


BINARY TRIE

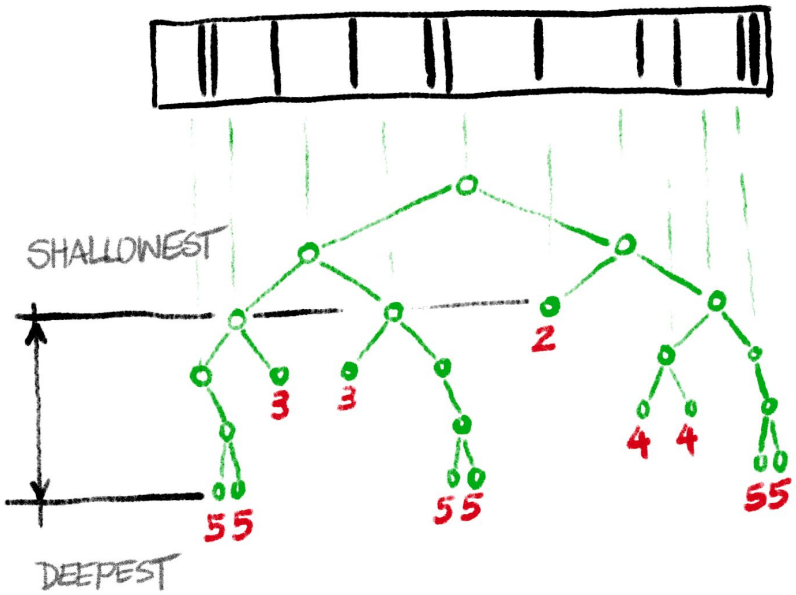


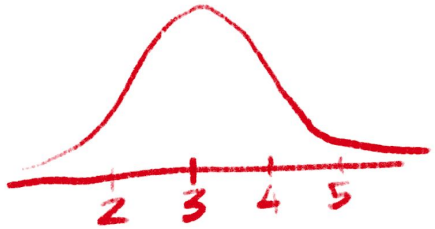
BINARY TRIE




BINARY TRIE
2
3
3
4
4
55
55
55
2
3
4
5

BINARY TRIE





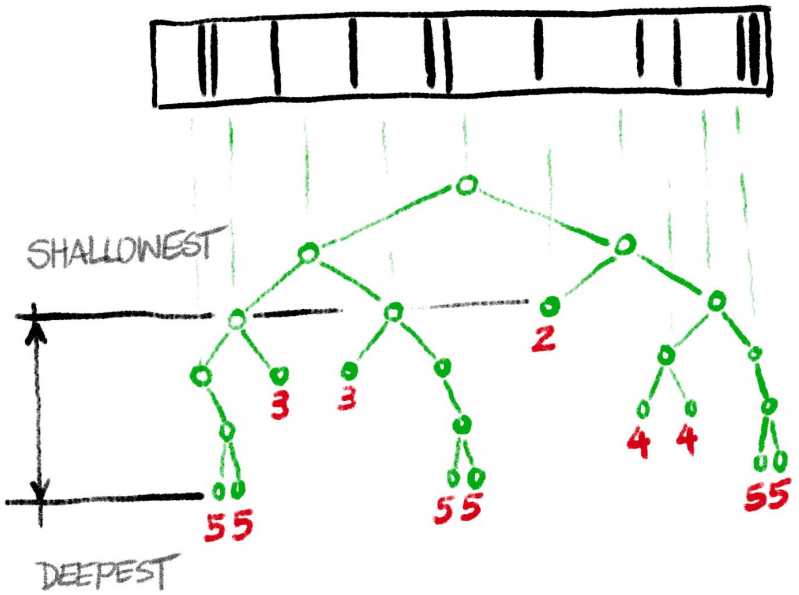
THEOREM:

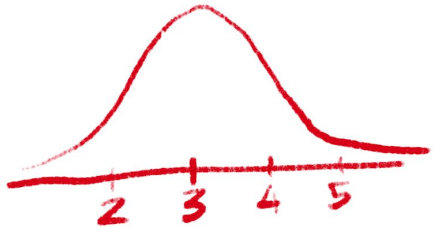
MEAN ~ LogN

DEVIATION ~ LogN

DEEPEST − SHALLOWEST ~ LogN

BINARY TRIE



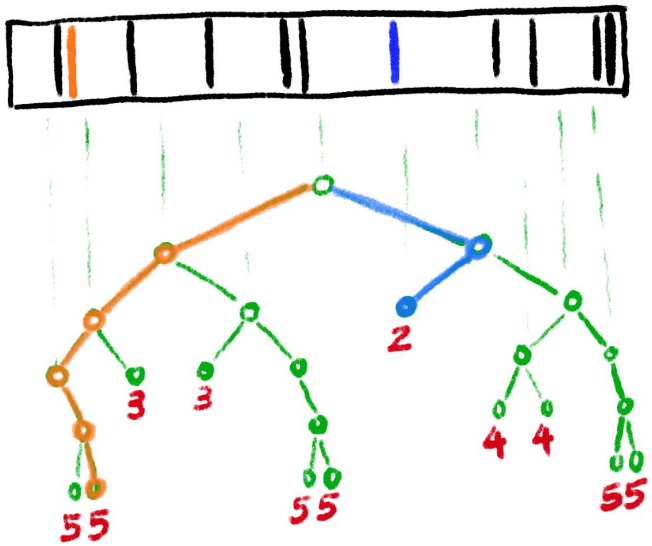


THEOREM:

MEAN ~ Log N
DEVIATION ~ Log N
DEEPEST − SHALLOWEST ~ Log N

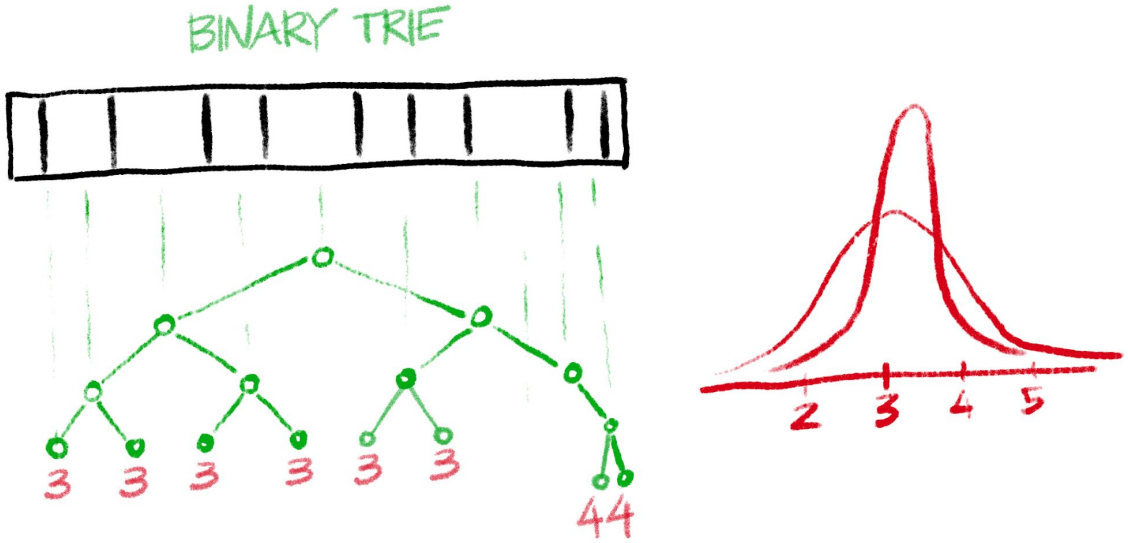
GITHUB.COM/LIBP2P/GO-LIBP2P-XOR
/PY-LIBP2P-XOR

BINARY TRIE

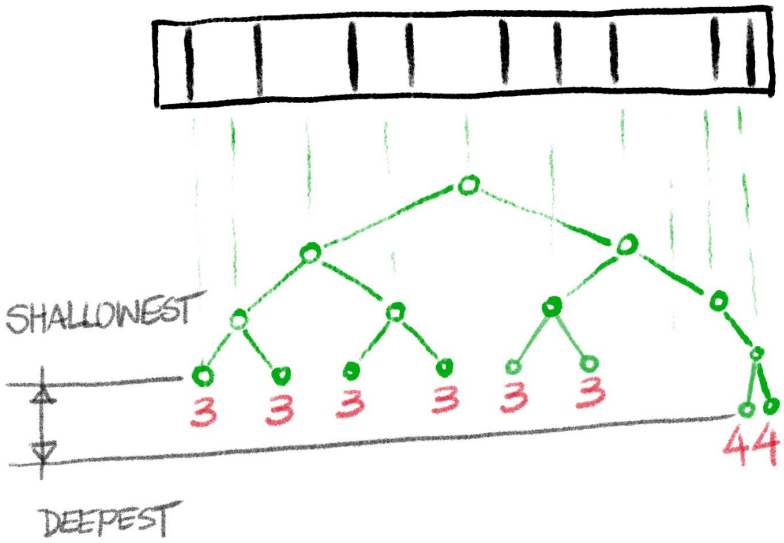


DRAW TWO
RANDOM NUMBERS

CHOOSE THE ONE
WHICH LANDS
SHALLOWER
IN
THE TRIE


BINARY TRIE
3
3
3
3
3
3
44
2
3
4
5

BINARY TRIE



THEOREM:

MEAN ~ Log N
DEVIATION ~ Log Log N
DEEPEST − SHALLOWEST ~ Log Log N

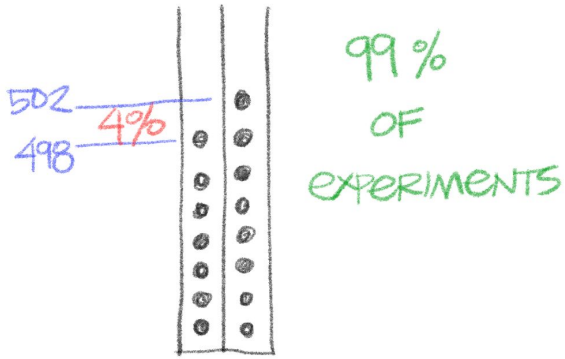
"THE POWER OF TWO CHOICES"

HAZARDOUS
MATERIAL
NEXT

PURE
MATH

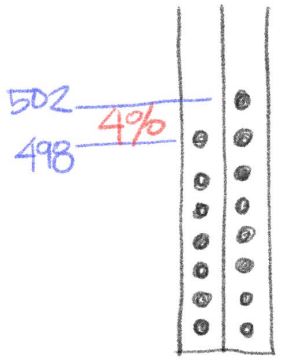
$N = 1000$

N = 1000



99%
OF
EXPERIMENTS

$N = 1000$



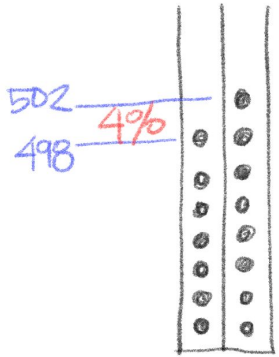
99%
OF
EXPERIMENTS

THEOREM:

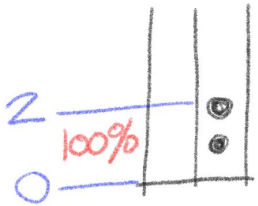
DIFFERENCE $\sim \sqrt{N}$

$$= 100\frac{\sqrt{N}}{N}\% = \frac{100}{\sqrt{N}}\% \to 0\%$$

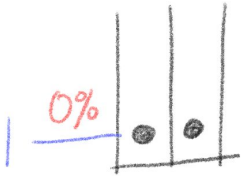
N=1000 . . . . N=2



99% OF EXPERIMENTS



50%



50%

THEOREM:

DIFFERENCE $\sim \sqrt{N}$

$$= 100\frac{\sqrt{N}}{N}\% = \frac{100}{\sqrt{N}}\% \to 0\%$$