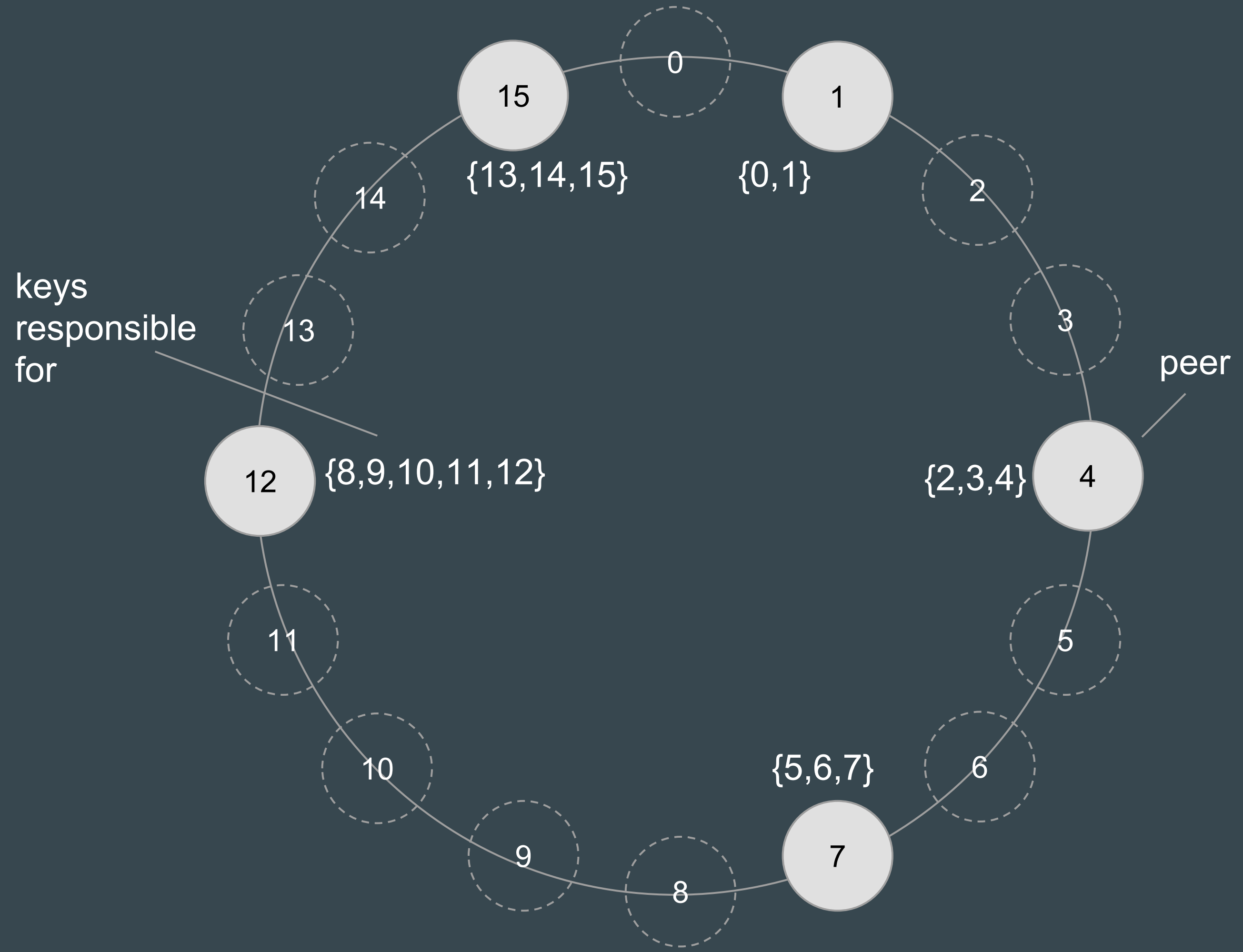
The Chord DHT

- *m*-bit *key* unique for every peer and piece of content.
- Define: *succ(k)* for any key *k* is peer that exists with smallest $id \geq k$.
- Content with key *k* hosted by *succ(k)*.

- *m = 4*



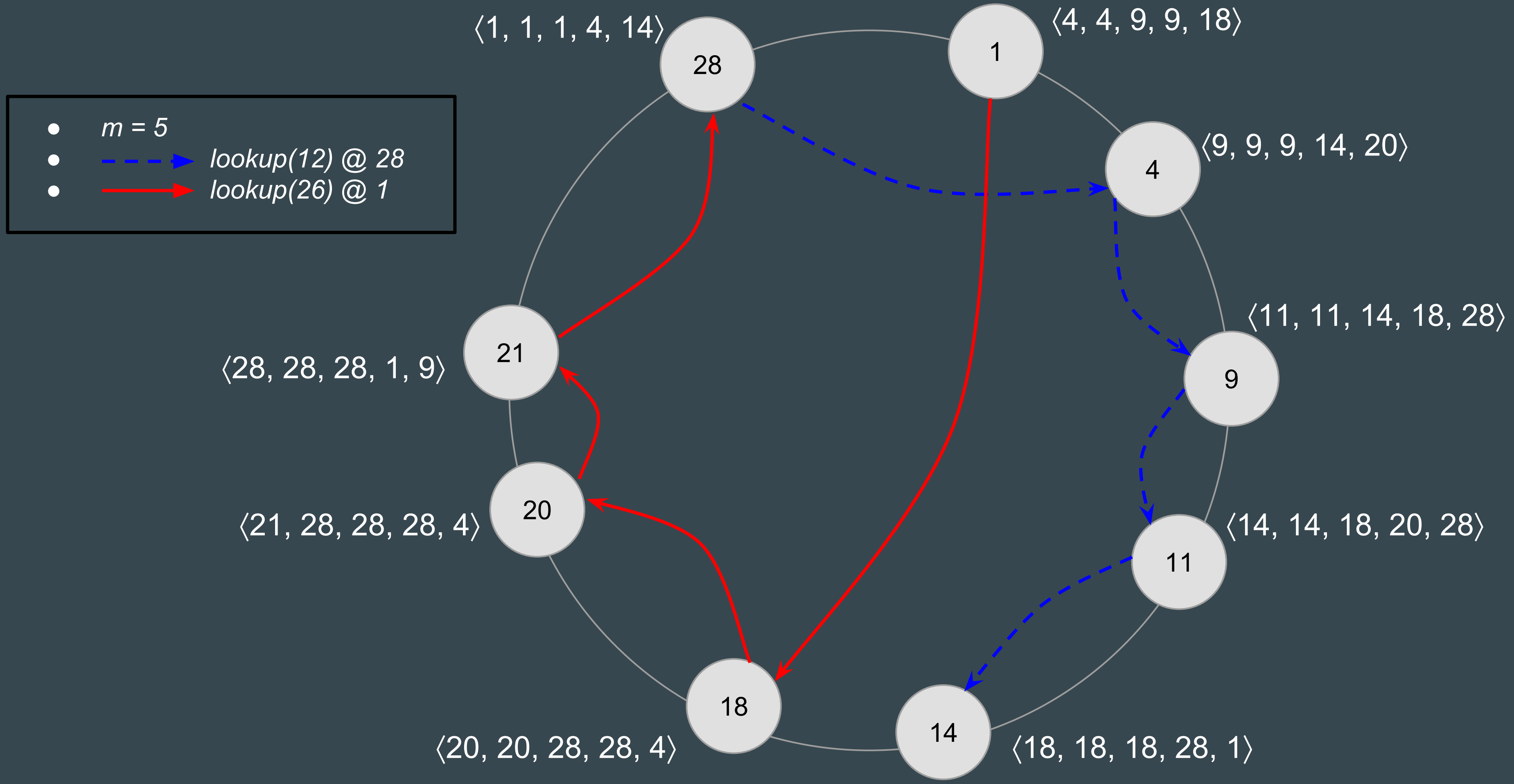
keys
responsible
for
peer
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
{0,1}
{2,3,4}
{5,6,7}
{8,9,10,11,12}
{13,14,15}

lookup(k)

- We may want to *lookup(k)* at any peer.
- Query is routed to *succ(k)*. How?

The Chord Approach

- Maintain a "finger table" (routing table) at Peer *p*, $FT_p[\,]$.
- *m* entries
- $FT_p[i] = succ(p + 2^{i-1})$, for all $i \in [1,m]$


m = 5
lookup(12) @ 28
lookup(26) @ 1
28
⟨1, 1, 1, 4, 14⟩
1
⟨4, 4, 9, 9, 18⟩
4
⟨9, 9, 9, 14, 20⟩
9
⟨11, 11, 14, 18, 28⟩
11
⟨14, 14, 18, 20, 28⟩
14
⟨18, 18, 18, 28, 1⟩
18
⟨20, 20, 28, 28, 4⟩
20
⟨21, 28, 28, 28, 4⟩
21
⟨28, 28, 28, 1, 9⟩

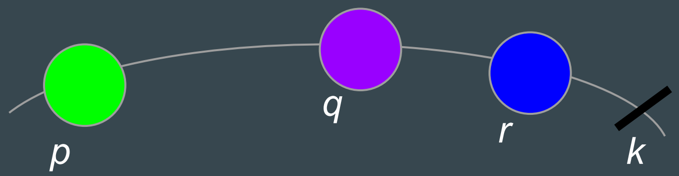
Claims

- We go around the circle at most once
  - Termination guaranteed
- Expected number of “hops” is $O(\log n)$, where n = # peers.
  - Each hop covers half the arc-length.
  - Hand-wave: peers are equidistant under randomness assumption.

Each hop covers > half the arc-length…

Claim:

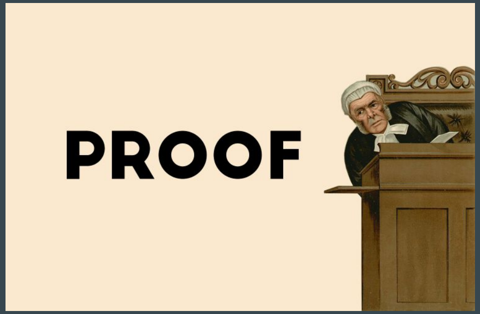
Suppose we do *lookup*(k) @ peer p. Assume that r is the peer immediately before *succ*(k). Suppose that the next hop at p is q. Then: q - p > (r - p)/2.



Proof

Suppose $q = FT_p[j]$.

Then, $q \geq p + 2^{j-1}$. And, $r < p + 2^j$.



Artwork credit

- … (to be completed)