

# 大数据开发技术

东北林业大学

卢洋



# 网络拓扑—节点距离计算



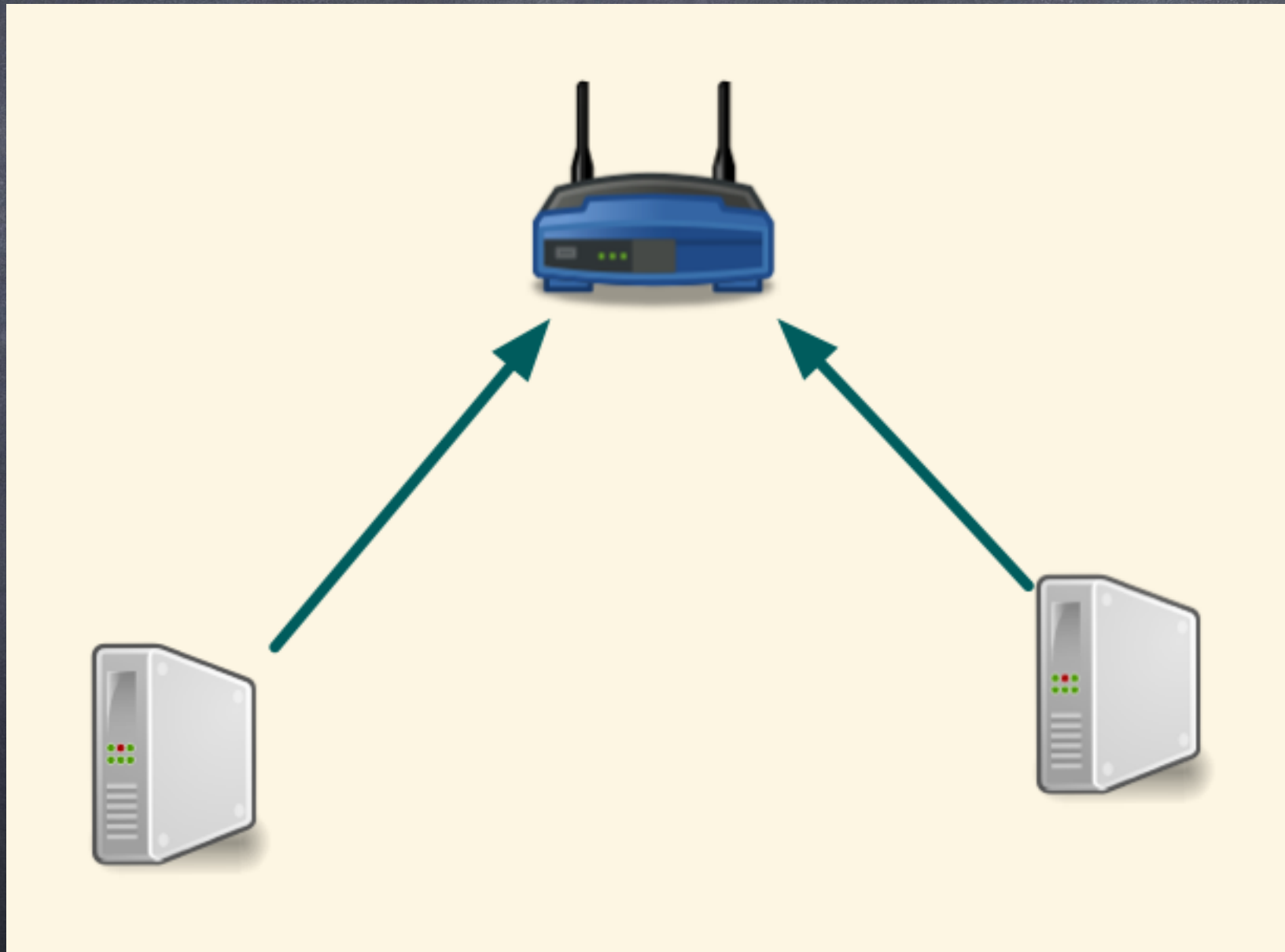
## 网络拓扑与节点距离的计算

- HDFS写数据过程中，NameNode会选择距离待上传数据距离最近的DataNode接收数据；
- 节点距离：两个节点到达最近共同祖先的距离总和。

最近路由器



共同祖先?





# 节点距离计算



集群-d1



集群-d2



机架-r1



机架-r2



机架-r3



机架-r4



机架-r5



机架-r6



n-0



n-1



n-2



n-0



n-1



n-2



n-0



n-1



n-2



n-0



n-1



n-2



n-0



n-1



n-2



n-0



n-1



n-2

# 1 同一节点: Distance( $d1/r1/n0$ , $d1/r1/n0$ )



集群-d1



集群-d2



机架-r1



机架-r2



机架-r3



机架-r4



机架-r5



机架-r6



n-0



n-1



n-2



n-0



n-1



n-2



n-0



n-1



n-2



n-0



n-1



n-2



n-0



n-1



n-2



n-0



n-1



n-2



自己跟自己  
0距离

1 同一节点:  $\text{Distance}(d1/r1/n0, d1/r1/n0)=0$



集群-d1



集群-d2



机架-r1



机架-r2



机架-r3



机架-r4



机架-r5



机架-r6



n-0



n-1



n-2



n-0



n-1



n-2



n-0



n-1



n-2



n-0



n-1



n-2



n-0



n-1



n-2



n-0



n-1



n-2

## 2 同一机架上的不同节点: Distance( $d1/r1/n1$ , $d1/r1/n2$ )



集群-d1



集群-d2



机架-r1



机架-r2



机架-r3



机架-r4



机架-r5



机架-r6



n-0



n-1



n-2



n-0



n-1



n-2



n-0



n-1



n-2



n-0



n-1



n-2



n-0



n-1



n-2



n-0



n-1

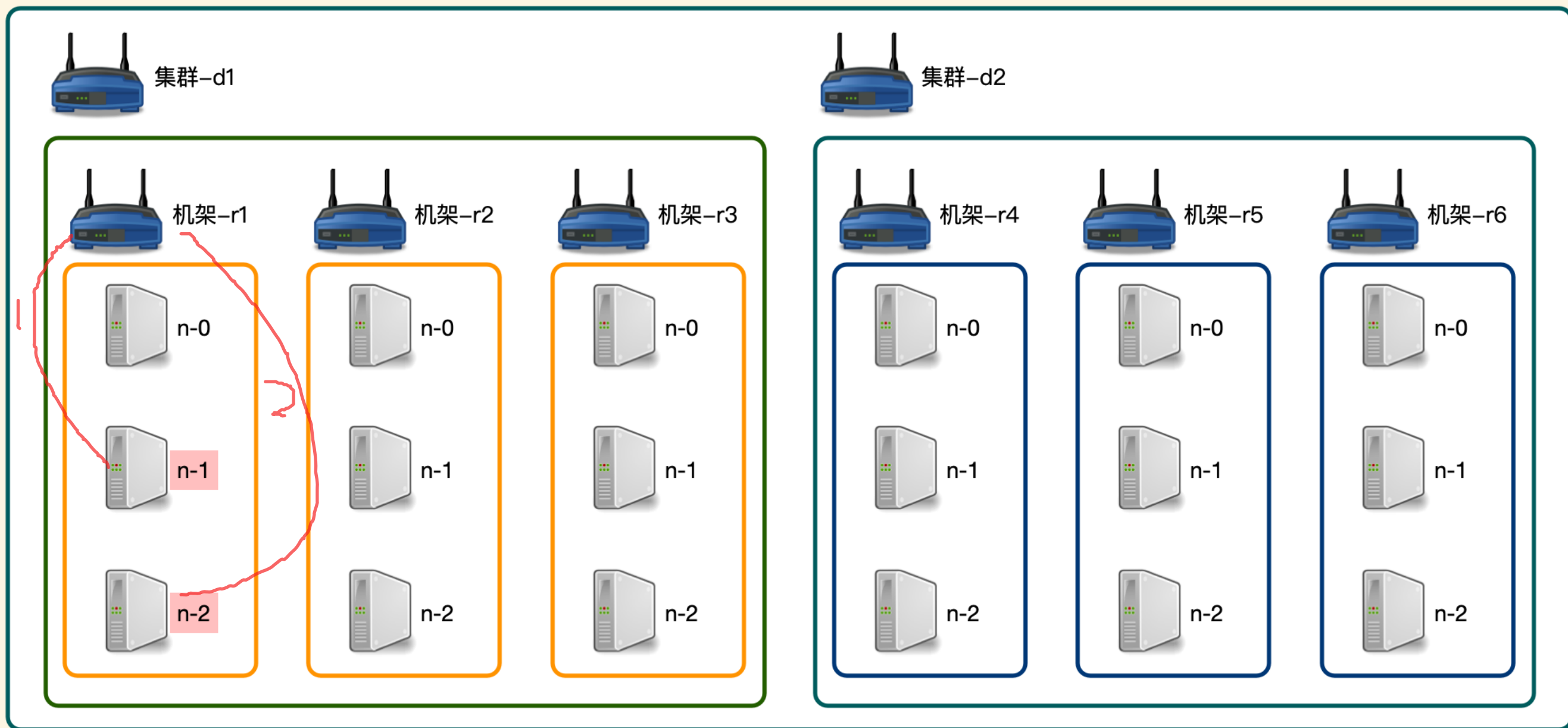


n-2



## 2 同一机架上的不同节点:

$\text{Distance}(d1/r1/n1, d1/r1/n2)=2$





### 3 同一数据中心不同机架上的节点: Distance( $d1/r2/n0$ , $d1/r3/n2$ )



集群-d1



集群-d2



机架-r1



机架-r2



机架-r3



机架-r4



机架-r5



机架-r6



n-0



n-1



n-2



n-0



n-1



n-2



n-0



n-1



n-2



n-0



n-1



n-2



n-0



n-1



n-2



n-0



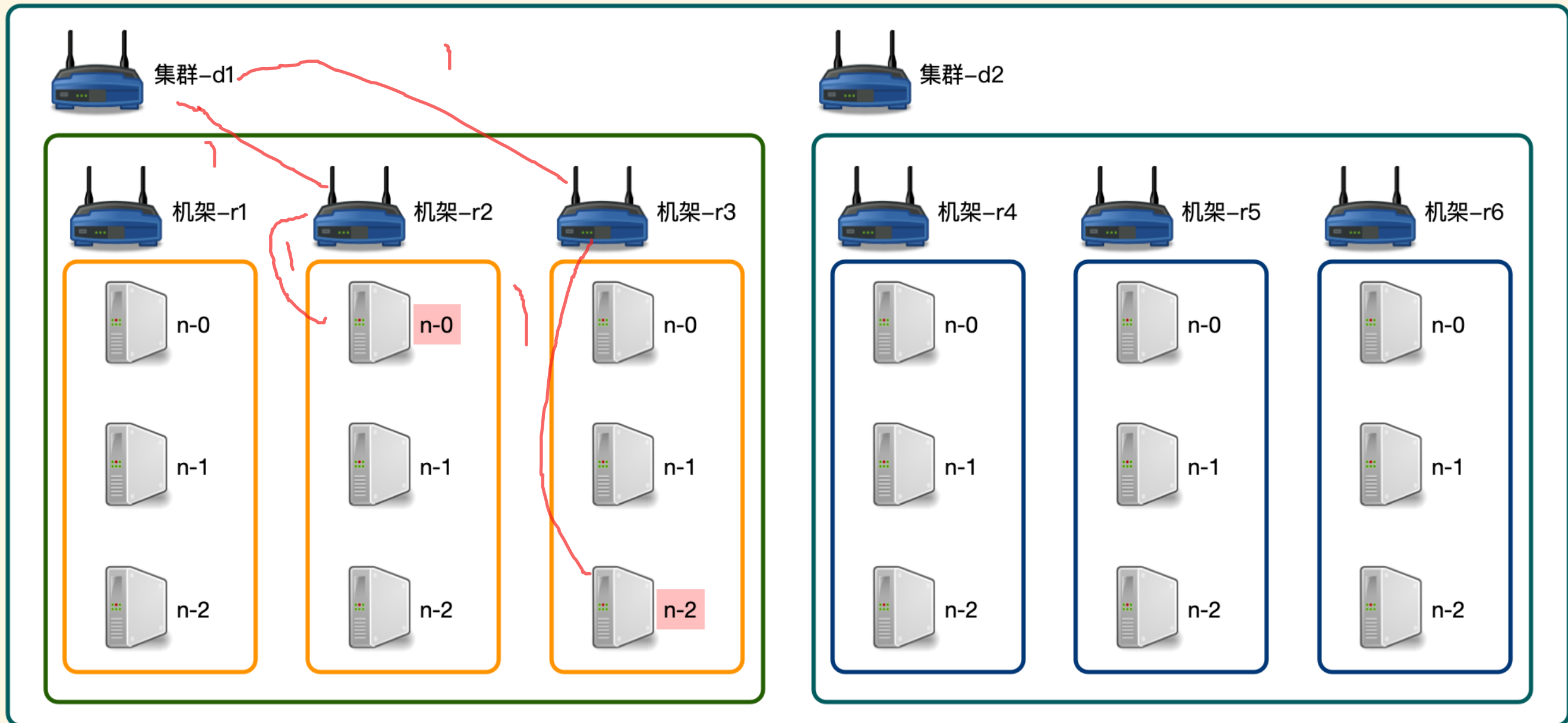
n-1



n-2



### 3 同一数据中心不同机架上的节点: $\text{Distance}(d1/r2/n0, d1/r3/n2)=4$





## 4 不同数据中心的节点: Distance( $d1/r2/n1$ , $d2/r4/n1$ )



集群-d1



集群-d2



机架-r1



机架-r2



机架-r3



机架-r4



机架-r5



机架-r6



n-0



n-1



n-2



n-0



n-1



n-2



n-0



n-1



n-2



n-0



n-1



n-2



n-0



n-1



n-2



n-0



n-1

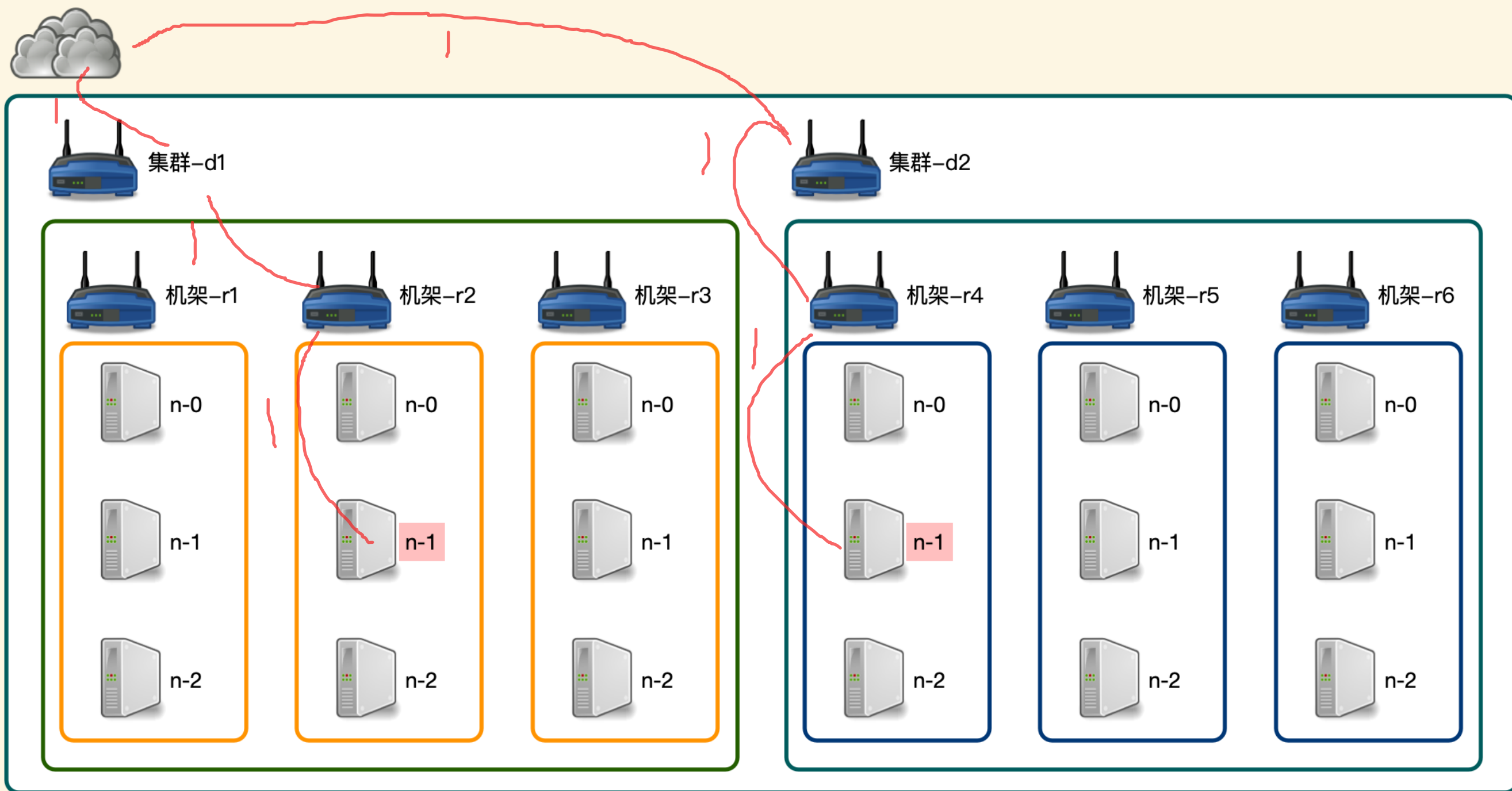


n-2



#### 4 不同数据中心的节点:

$\text{Distance}(d1/r2/n1, d2/r4/n1)=6$





# 网络拓扑

