

上一篇文章通过配置静态路由表初步学习了如何进行路由表的编写，从而加深了对路由表的认识，不过我们也说出了它的特点：配置方式简单、开销小，一般适用于小型网络，缺点是不能及时适应网络状态的变化。

与之对应的是动态路由选择，路由器可通过路由选择协议自动获取路由信息，能较好地适应网络状态的变化，适合大型网络，缺点是比较复杂、开销比较大。

由于篇幅问题，将动态路由选择协议分为了三个章节：RIP、OSPF、BGP，本文着重来讨论下RIP协议的基本工作原理和工作过程。

一、因特网使用的是动态路由选择

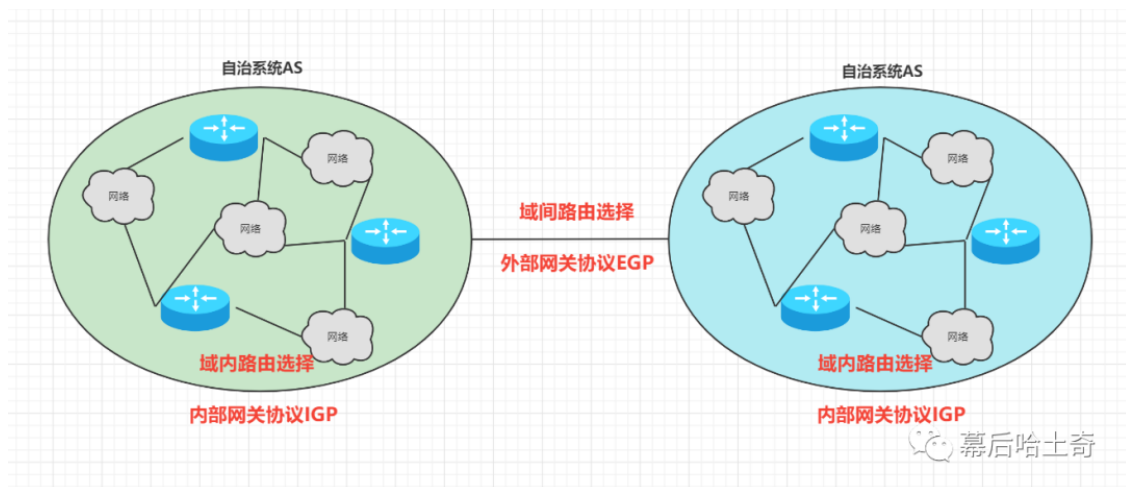
因特网是全球最大的互连网络，具有以下三个主要特点：

- 自适应：动态路由选择，能较好地适应网络状态的变化。
- 分布式：因特网中的各路由器通过相互间的信息交互，共同完成路由信息的获取和更新。
- 分层次：将整个因特网划分为许多较小的自治系统AS（Autonomous System）。例如我们可将一个较大的因特网服务提供商作为一个自治系统。在自治系统内部和自治系统外部，分别使用不同类别的路由选择协议，分别进行路由选择。

自适应和分布式很容易理解。

因特网太巨大了，使用动态路由选择实现自适应是最好的选择，此外，一个路由器肯定无法完成整个网络的互连，需要很多路由器互连才行，还记得我们把路由器比作一个个交棒的选手吗？众人拾柴火焰高，而要形成良好的团队合作氛围，必然需要进行充分沟通，路由器之间也会经常交互信息，更新自己的路由表，达到自适应和分布式的目的。

下面来着重说说分层次的特点。



自治系统就是几个路由器组成了一个小团体 🧑‍🤝‍🧑，小团体内部使用专用的协议进行通信，而小团体和小团体之间也使用专用的协议进行通信。

很容易理解，从AS划分的角度来说，AS之间的路由选择协议我们统称为外部网关协议EGP，各AS内部使用的路由选择协议统称为内部网关协议，EGP和IGP只是大的分类名称，并不是具体的路由选择协议。

另外，这里使用的是网关这个名词，这是因为在因特网早期的RFC文档中，没有使用路由器，而是使用网关这个名词，现在新的RFC中又改为使用路由器了，外部网关协议EGP可称为外部路由协议ERP，内部网关协议IGP可称为内部路由协议IRP。

另外，各自治系统内部可根据各自需求使用不同类别的内部网关协议，互不影响，比如一个AS内部网关协议使用【路由信息协议RIP】，另一个AS内部网关协议使用【开放式最短路径优先OSPF】，两个AS之间使用的外部网关协议可以为【边界网关协议BGP】。

其中，可供内部网关协议选择的协议还有很多，除了上面提及的RIP、OSPF，还有内部网关路由协议IGRP、增强型内部网关路由协议EIGRP、中间系统到中间系统IS-IS等等，而外部网关协议一般就是BGP了。

值得强调的是，路由选择协议是在路由器上运行的，路由器根据所使用的路由选择协议，周期性地与其他路由器进行路由信息的交互，互相告知各自知道的路由信息，来更新路由表。

补充1：实际上，路由器中还有一张转发表，转发表是根据路由表的信息得出的，使用特殊的数据结构使得可以快速查找，不过我们这里不严格区分路由表和转发表，还是以路由表来表述问题，便于问题的说明和分析。

补充2：路由器的各端口还应具有输入缓冲区和输出缓冲区，输入缓冲区用来暂存新进入路由器但还来不及处理的分组，输出缓冲区用来暂存已处理完毕但还来不及发送的分组。

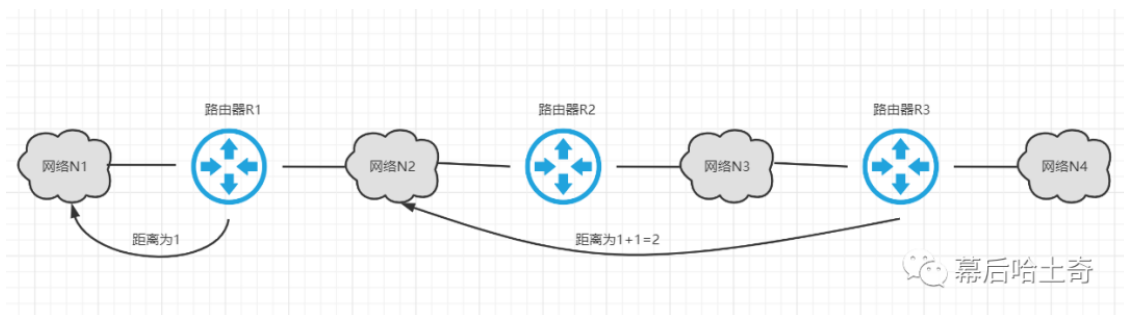
好了，下面我们着重来研究下RIP基本工作原理。

二、RIP基本工作原理

路由选择协议RIP，英文全称为Routing Information Protocol，是内部网关协议IGP中最先得到广泛使用的协议之一。

RIP要求自治系统AS内的每一个路由器都要维护从它自己到AS内其他每一个网络的距离记录，RIP使用“跳数”作为度量，来衡量到达目的网络的距离。

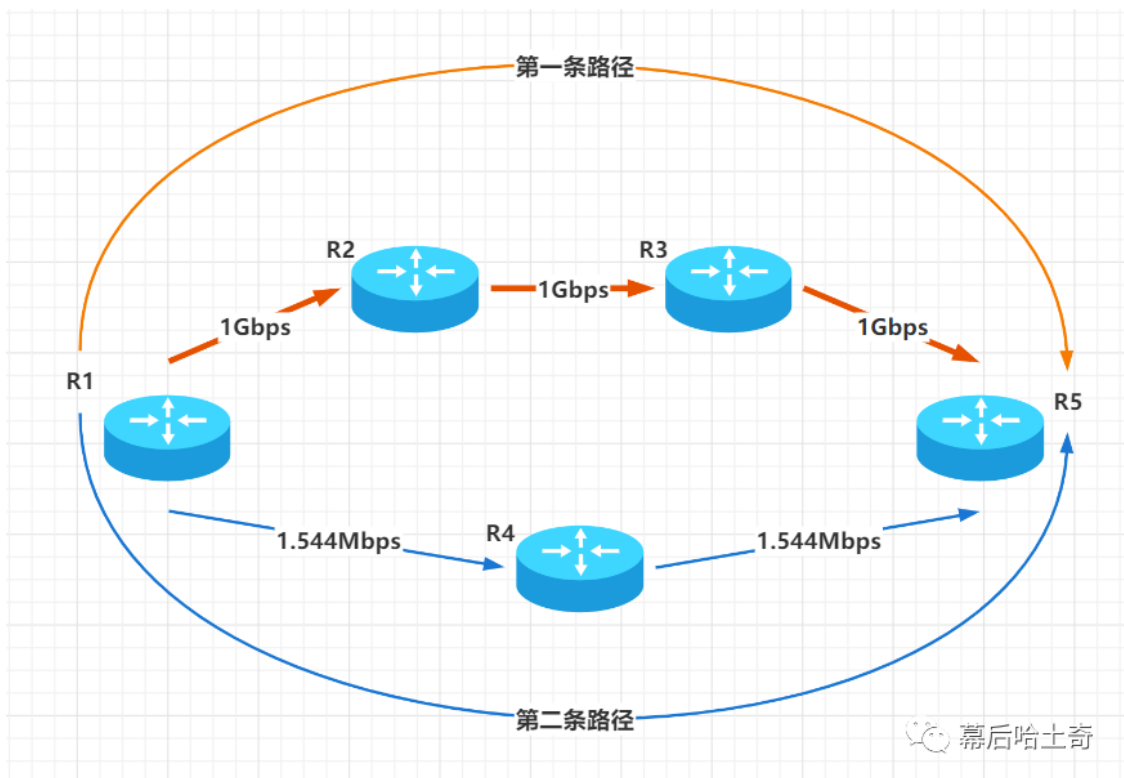
我们将路由器到其直连网络的距离定义为1，路由器到非直连网络的距离定义为所经过的路由器数加1。



如图所示，路由器R1到网络N1的距离认为是1跳，而路由器R3到网络N2则需要再经历一个路由器才行，需要2跳。

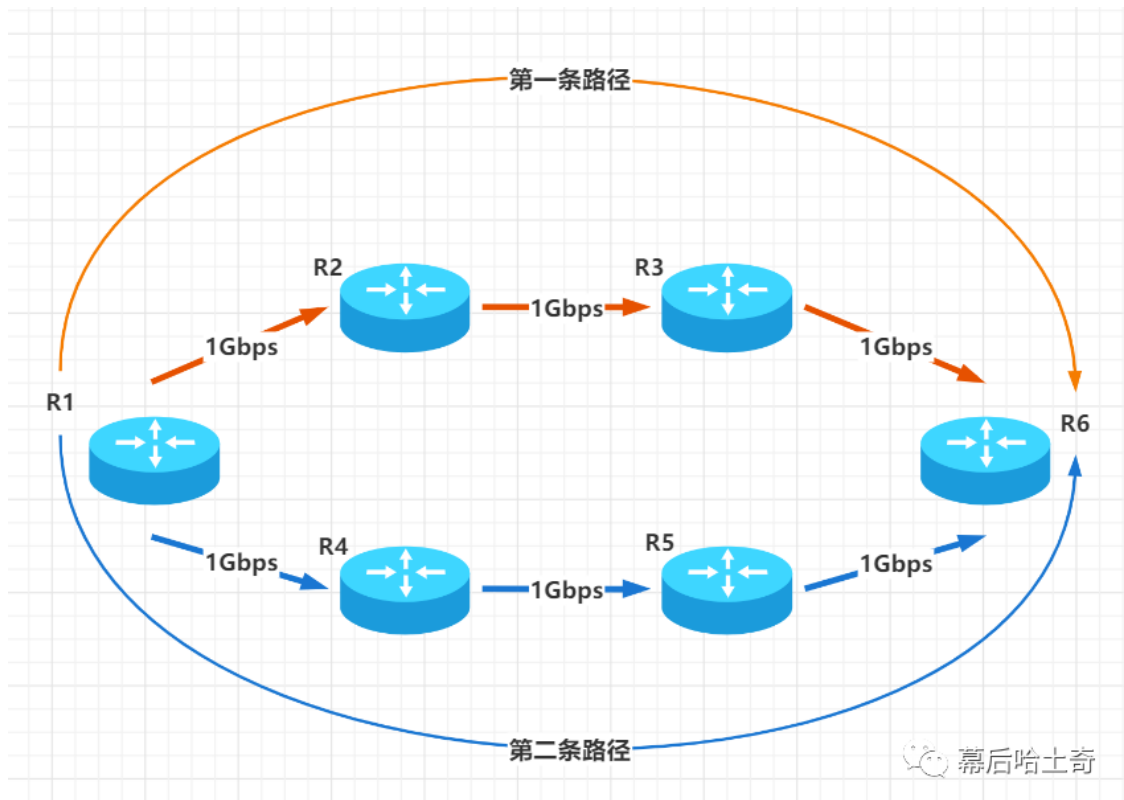
在RIP协议中，允许一条路径最多只能包含15个路由器，当距离等于16时相当于不可达，因此，RIP只适用于小型互联网。（需要说明的是，有些厂家的路由器并没有严格按照RIP标准文档的规定来实现RIP，例如思科路由器中的RIP将路由器到直连网络的距离定义为0）

在RIP眼中，好的路由就是距离短的路由，也就是所通过路由器数量最少的路由。



可以看到，有两条路径，RIP根据距离会选择R1-》R4-》R5这条路由，尽管这条路由上各段链路的带宽都非常小。

再来看一种情况，有两条路径，分别是R1-》R2-》R3-》R6和R1-》R4-》R5-》R6两条路由，可以看到这两条路由的距离是相等的，该怎么选择呢？

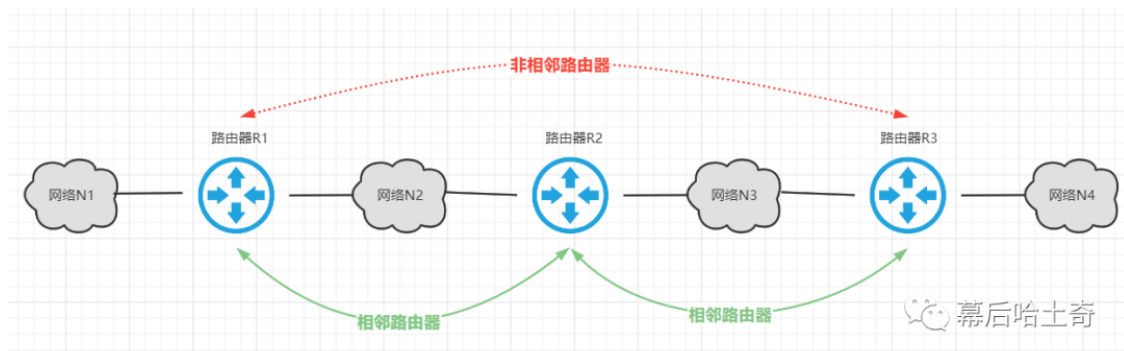


当到达同一目的网络有多条距离相等的路由时，则可进行等价负载均衡，即将通量均衡地分布到各条等价的路由上。

RIP包含以下三个要点：

- 和谁交换信息：仅和自己相邻路由器交换信息。
- 交换什么信息：自己的路由表。
- 何时交换信息：周期性交换，比如每隔30秒发送一次RIP更新报文。（每个RIP路由器的定时器都独立于网络中其他路由器，因此它们都同时广播的可能性很小。）

何为相邻路由器呢？看下图所示即可明白。



三、RIP基本工作过程

下面来说说RIP的基本工作过程，可以归纳为如下三个过程：

- 路由器刚开始工作时，只知道自己到直连网络的距离为1；
- 每个路由器仅和相邻路由器周期性地交换并更新路由信息；
- 若干次交换和更新后，每个路由器都知道到达本AS内各网络的最短距离和下一跳地址，称之为收敛。

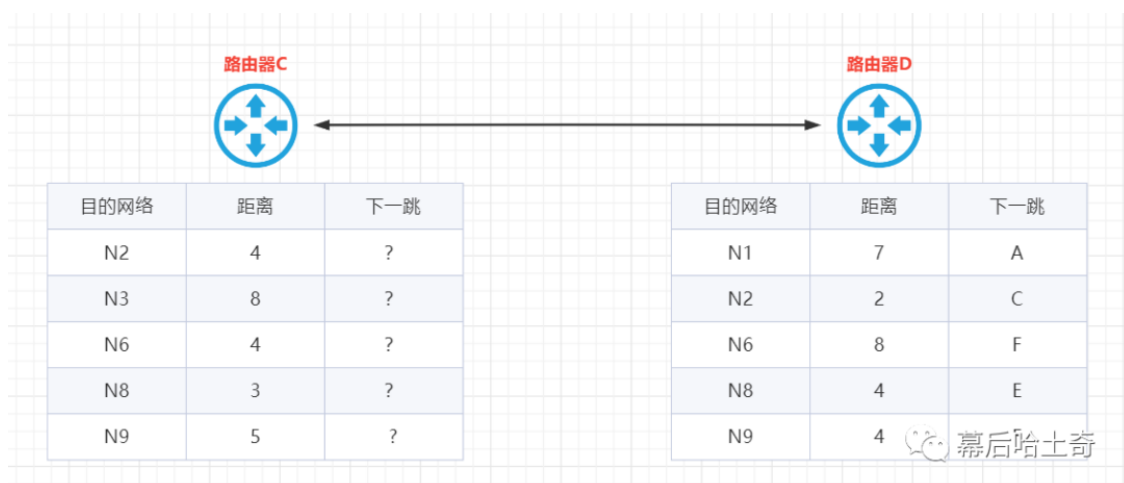
收敛：所有路由器对于所有链路所有网段的路径信息达成一致的过程，路由选择协议越好收敛速度越快。

收敛时间：从网络拓扑发生变化到网络中所有路由器都知道这个变化的时间。

下面具体来看看在RIP协议下路由器是如何进行更新的。假设有路由器C和路由器D，他们互为相邻路由器，他们之间周期性地交换并更新路由信息。

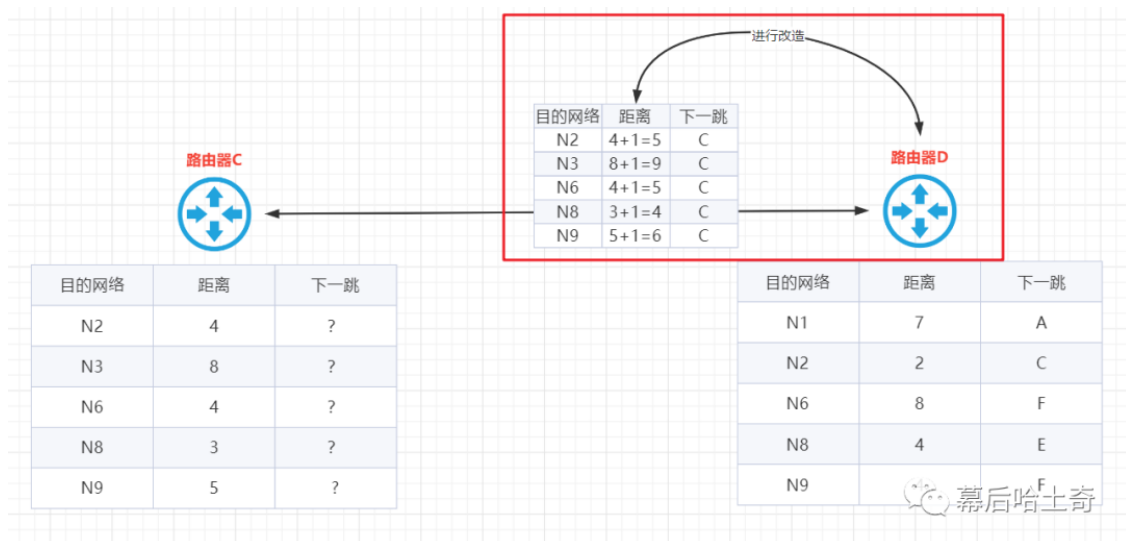


其中路由器C和路由器D的路由表假设为如下所示，其中路由器C的路由表中的到达各目的网络的下一跳都记为了问号，因为路由器D不关心路由器C的这些内容。



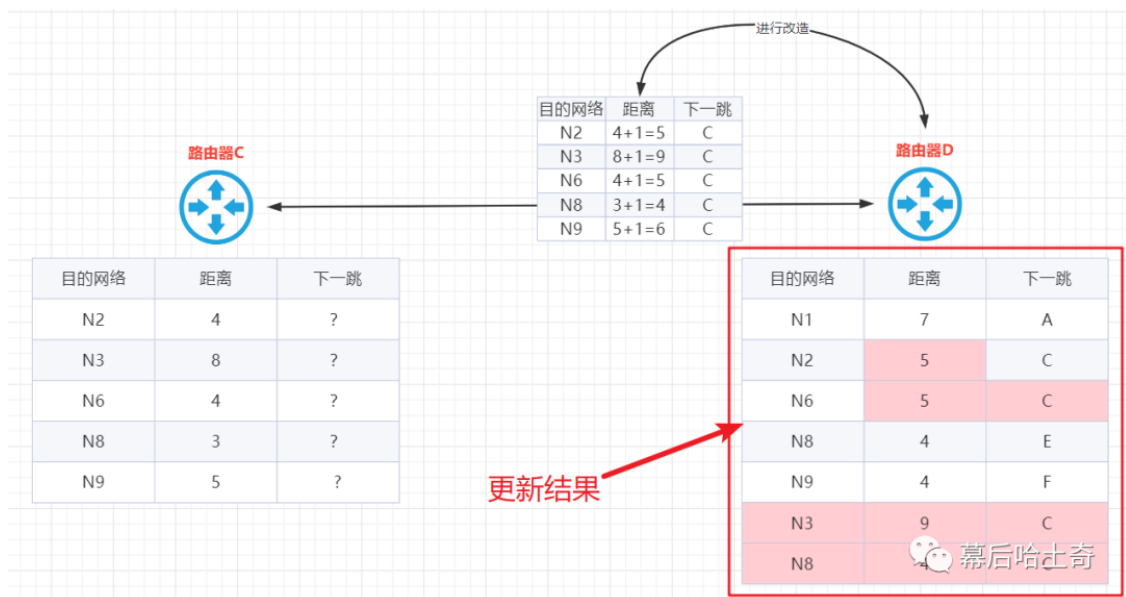
假设此时路由器C的RIP更新报文发送周期到了，则路由器C将自己路由表中的相关路由信息封装到RIP更新报文中发送给路由器D，我们可以简单理解为路由器C将自己的路由表发送给了路由器D。

路由器D收到后，对路由器C发来的路由表进行**改造**，将下一跳全部改为C，并且将距离全部加一。



这样的操作比较容易理解，因为路由器C告诉路由器D，它可以到达这些目的网络，那么路由器D作为路由器C的邻居路由器，当然也就可以通过路由器C这个跳板到达这些目的网络，只是跳数增加了一个。

好了，下面路由器D可以根据改造好的路由表，来**更新**自己先前的路由表。



可以看到，路由器D经过这一轮更新后发生了一些变化：

路由器D路由表更新前			路由器D路由表更新后		
目的网络	距离	下一跳	目的网络	距离	下一跳
N1	7	A	N1	7	A
N2	2	C	N2	5	C
N6	8	F	N6	5	C
N8	4	E	N8	4	E
N9	4	F	N9	4	F
			N3	9	C
			N8	4	C

下面一个一个来说更新的具体规则。

- 对于去N1目的网络，由于路由器C没有此路由信息，路由器D保留即可；
- 对于去N2目的网络，由于下一跳都是路由器C，那么则需要根据最新的信息更新距离，因为很有可能此时网络拓扑发生了变化。
- 对于去N3目的网络，路由器D原先没有，说明发现了新的网络，直接添加；
- 对于去N6目的网络，路由器D原先是通过F走8跳，而得知可以通过C走5跳即可，那么更新为更具优势的5跳。
- 对于去N8目的网络，路由器D原先是通过E走4跳，此时发来一条新的信息是通过C也是4跳可以到达，此时需要将C这个路由加上，作为负载均衡路由。
- 对于去N9目的网络，路由器D原先是通过F走4跳，路由器C告诉D需要通过C走6跳，显然后者不具备优势，被舍弃，还是保留原始通过F走4跳的方案。

那么将其上述各种情况总结为规则的话如下：

- 到达相同目的网络，相同下一跳，则需要按照最新的消息进行更新，因为很有可能网络拓扑发生了变化；
- 如果发现了新的网络路由，则直接添加；
- 到达相同目的网络，不同下一跳，若新路由有优势，则更新；
- 到达相同目的网络，不同下一跳，若新路由劣势，不更新；
- 到达相同目的网络，不同下一跳，若新路由距离一样长，则增加做等价负载均衡；

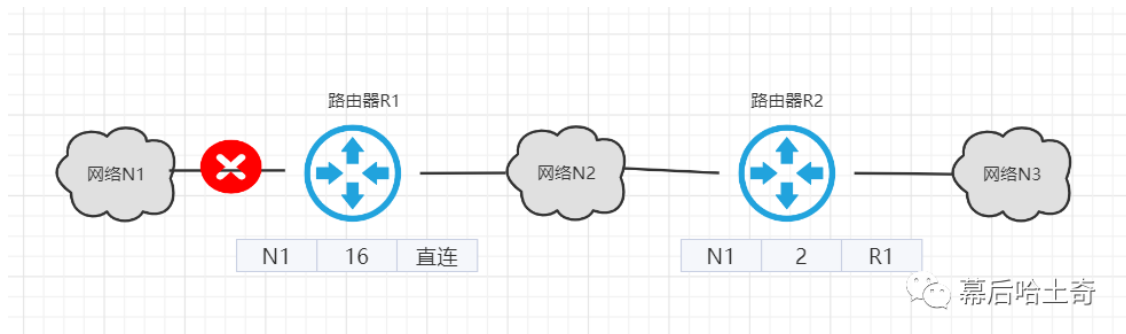
虽然说起来比较啰嗦，实际上规则还是比较简单的。

四、RIP坏消息传播慢问题

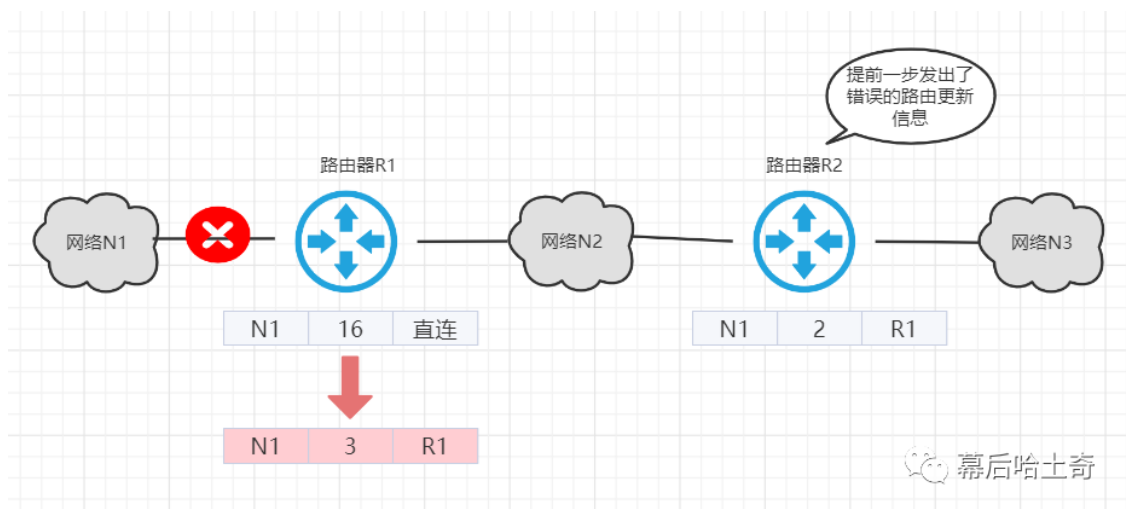
RIP存在坏消息传播的慢的问题。

看完下面的解释就会明白为什么称之为坏消息传播慢了。

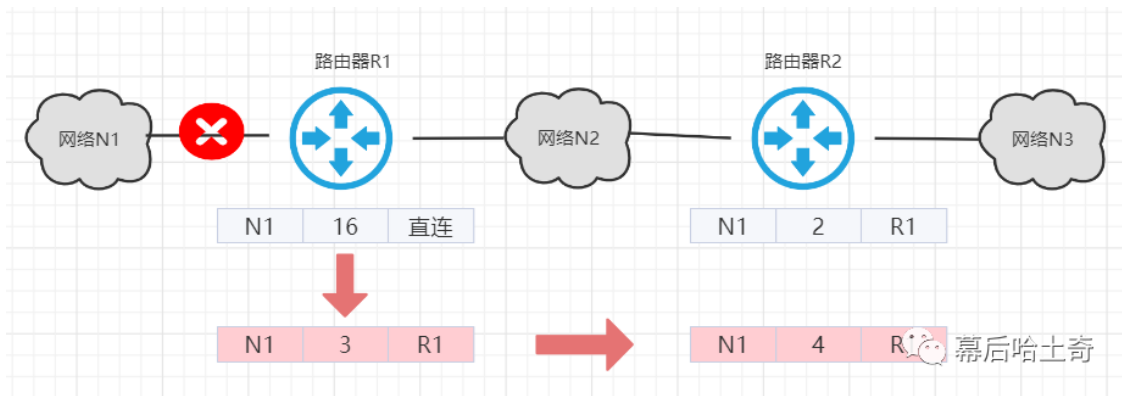
我们假设下面一种情况，路由器R1和直连网络N1之间出现故障，此时路由器R1将跳数置为16即表示不可达。此时路由器R2还保留着原来的路由信息即通过R1到N1的距离为2。



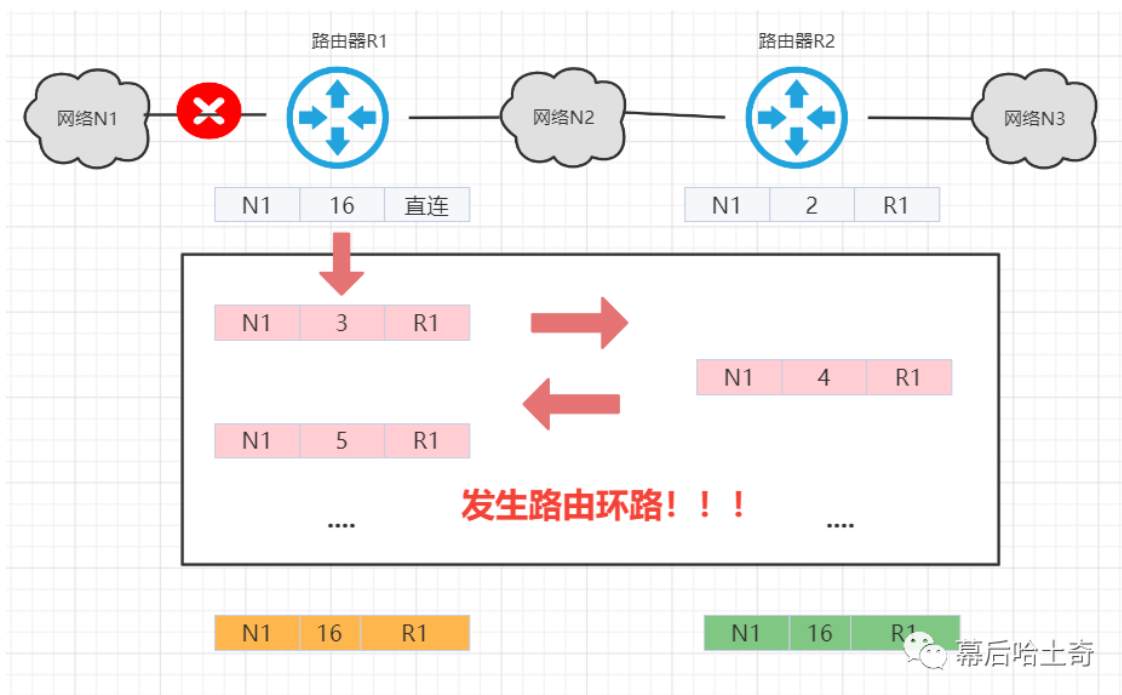
此时要发送RIP更新路由信息了，不过是路由器R2先发出来的，此时就会造成一个问题：**路由器R1收到相邻路由器R2的RIP更新路由后，将自己的路由表被误导更新为了如下：**



从上帝视角知道，此时发生了错误，但是路由器R1和R2都浑然不知，此时R1也发出路由更新信息，导致R2也进行错误的更新：



那么下面就会没完没了地错误地更新，直到变成16跳，R1和R2这才发现被耍了：



“坏消息传播得慢”指得就是路由器R1和N1之间的网络故障信息没有及时通知到相邻路由器，导致出现了以上路由环路问题，因此“坏消息传播得慢”问题又称为“路由环路”或“距离无穷计数”问题，这是距离向量算法的一个固有问題，可以采取多种措施减少出现这种问题的概率或减少该问题带来的危害。

- 16跳不可达可减少该问题带来的危害；
- 当路由表发生变化时就立即发送更新报文，即触发更新，而不是周期性发送；
- “水平切割”：由于路由器可能收到它自己发送的路由信息，而这种信息是无用的，解决办法就是在路由信息传送过程中，不再把路由信息发送到接收到此路由信息的接口上，从而在一定程度上避免了环路的产生。

RIP的度量值为跳数，跳数越小路径越优，而不考虑带宽，这一点实际上是不科学的，下面需要来了解下OSPF是如何工作的，看下它的工作原理，下篇文章见！