

Tunneling Route Reduction Protocol (TRRP)

Conceptual Summary and Analysis for RRG 3/08

The problem:

The BGP distribution system for routes in the DFZ is insufficiently scalable. Small operators, individuals and mobile users are prevented from using provider-independent IP addresses because issues associated with memory consumption, processing and convergence time in the DFZ routers drive the cost above \$6200 per year per announced prefix.

General summary of TRRP's solution:

Each Autonomous System (AS) deploys one or more TRRP Ingress Transit Routers (ITRs). An AS-interior default route leads from the BGP routers to the ITR. The ITR finds an Egress Tunnel Router (ETR) for the destination IP address via a DNS lookup. The ITR then tunnels the packet via GRE to the "best" ETR. The ETR, which must have an IP address within space announced via BGP, knows a local-scope route to the destination IP address and delivers the packet. Longer route prefixes are then withdrawn from BGP until a comfortable BGP table size is attained.

TRRP is intended to supplement BGP. BGP is expected to continue doing what it does best: distributing short, large prefixes throughout the "default-free zone." TRRP takes over where BGP fails, routing provider independent addresses with a granularity up to and including individual IP addresses with almost no scalability penalty beyond it's initial implementation.

Advantages

1. **Unlimited scalability.** Because TRRP ITRs need only keep a small working set of destinations currently in use, scalability is limited only by the DNS itself, a protocol that has proven to have essentially unlimited scalability.
2. **Protocol agnostic.** Any network which can carry DNS traffic can also carry any routed protocol for which a TRRP mapping system has been defined.
3. **Mostly already deployed.** TRRP builds on the existing BGP system and reuses existing, widely deployed and well tested technologies like GRE and DNS. TRRP is explicitly designed to work with existing GRE routers as ETR endpoints and with existing DNS servers for the map lookups.
4. **Dirt Cheap.** Adding new maps in TRRP has the same infinitesimal cost as adding new names to DNS.

Disadvantages:

1. **Slow initial round trip.** Because TRRP uses "cached pull with TTL" to guarantee that changes in the routing maps are propagated to all interested parties, it suffers from a slow initial round trip when two hosts haven't recently communicated with each other. It has not yet been proven that this protocol-level delay won't negatively impact the user's application-level experience of the Internet.
2. **Cache-refresh overhead.** ITRs have to regularly refresh their map working set. While this has no user-visible impact and does not impact scalability, it does consume traffic overhead on the network.

Details about TRRP are available at: <http://bill.herrin.us/network/trrp.html>

Analysis

RRG design goals.

3.1 Improved Routing Scalability. TRRP's scalability is bounded only by the scalability of the DNS system. DNS has proven to have effectively unlimited scalability with almost no correlation between the number of records in the whole system and the system's cost.

3.2 Scalable support for Traffic Engineering. Supported to the individual IP address. Negligible cost for additional map records.

3.3. Scalable support for multi-homing. Supported to the individual IP address. Negligible cost for additional map records.

3.4. Scalable support for mobility. Provides a strong foundation for future work on mobility. Mobility includes numerous additional requirements not addressed by TRRP such as detecting and acquiring new towers, releasing old ones, disconnection and reconnection, automated billing, etc. An roaming system which acquires new tower ETRs at least one variably-settable-time-to-live prior to releasing the old ones should have seamless routing within the TRRP framework.

3.5. Simplified renumbering. TRRP discourages renumbering for all but DNS servers, ETRs and those few protocols (notably NTP) which suffer a significant negative reaction to the startup delay. Address space used by TRRP will generally be provider independent. Where renumbering is necessary, it is neither harder nor easier than today.

3.6. Decoupling location and identification. TRRP completely decouples the destination IP address from the nearby ETRs which currently serve it.

3.7. First-class elements.

I don't have a clear understanding of what "first class element" means in the RRG design goals context. I think it means: how adaptable is it without changing the protocol and redeploying the components.

TRRP uses DNS to find maps. This map lookup can work over any protocol over which DNS can operate. TRRP ITRs are required to support GRE as a tunneling protocol but may support others. The DNS map entries can specify any tunneling protocol including protocols not yet written.

TRRP describes ETRs but could reasonably support a non-tunneling protocol such as "loose source routing" instead where the middleman was a loose source router instead of an egress tunnel router. The map design is flexible enough to change all this incrementally with zero negative impact on ongoing operation. TRRP anticipates that someone *will* create and deploy a better-tailored tunneling protocol than GRE without breaking or requiring a redeployment of TRRP.

Of course, the bindings between the destination IP address and its set of ETRs and tunneling protocols can also be changed at any time without any sort of software restart.

If that's what it means to be a first class element then yeah, got that.

3.8. Routing quality. As a map pull and cache protocol, TRRP suffers from an initial delay when two hosts start communicating with a sufficiently large lapse between communications. This affects only those routes for which TRRP supplements BGP; the short, large prefixes which remain in BGP are unaffected. Routes which follow TRRP usually (but are not required to) jump one additional hop to the side from a BGP router to an adjacent ITR for encoding. TRRP offers no other degradation to routing quality versus BGP today and may in some cases speed convergence and stability. Additional experimentation is necessary to determine whether the initial delay has a meaningfully negative user-visible impact.

3.9. Routing security. TRRP relies on DNS's proven and brutally well tested security model to guarantee routing security. It also relies on the remaining BGP system to properly route packets to the DNS servers and ETRs but offers no additional functionality to make BGP more secure. TRRP does reduce the total number of participants who can inject routes into the BGP system, reducing the scope of BGP's security problem.

3.10. Deployability. TRRP's technology is mostly already deployed. Most any router can be a TRRP GRE ETR. Any DNS server which supports TXT records can provide a basic TRRP map server. TRRP ITRs would need to be written and a trrp.arpa DNS hierarchy would need to be established, probably managed by the regional Internet registries (RIRs). TRRP is designed to run on generic COTS server hardware and is designed to have a [cashflow-positive impact on it's deployer from the very first ITR installation](#).

RADIR problem statement

4. Pressures on Routing Table Size

4.1. Traffic Engineering. See [RRG design goals 3.2](#).

4.2. Multihoming See [RRG design goals 3.3](#).

4.3. End Site Renumbering

Avoided as much as possible by using provider-independent address space. See [RRG design goals 3.5](#)

4.4. Acquisitions and Mergers. Renumbering avoided as above. If the slices are too small to economically stay in BGP, they move in to TRRP instead.

4.5. RIR Address Allocation Policies. TRRP frees the RIRs to assign provider independent address space down to the individual IP address level if desired. While it's still desirable for ISPs to have at least one large allocation in which to locate ETRs and DNS servers, the RIR's process is almost entirely decoupled from the routing table.

4.6. Dual Stack Pressure on the Routing Table. TRRP allows new protocols to be deployed at the edge without deploying new protocols in the core and vice versa.

4.7. Internal Customer Routes. TRRP does not directly impact internal customer routing.

4.8. IPv4 Address Exhaustion. TRRP does not directly impact IPv4 address exhaustion, however it could relieve the associated BGP stress from exhaustion by shunting the new, tiny prefixes into TRRP instead of BGP.

5. Pressures on Control Plane Load

TRRP's impact on the router control and forwarding planes is similar on an internet-wide scale to how MPLS impacts forwarding and control planes within a particular autonomous system (AS). The key differences are:

1. There are more "labels," at least one for each **Transit AS** which remains in the BGP system.
2. The "labelling" process happens in ITRs, much closer to the system edge where it can easily be handled on cheap COTS hardware.
3. Instead of a whole new protocol to propagate label routes, TRRP relies on existing BGP.
4. ITRs need only maintain a working set that includes destinations to which it has recently relayed packets. No ITR ever need know a significant fraction the entire set of address to

"label" maps.

[William Herrin](#) 3/3/2008