

# DataRouter: A Network Layer Service for Application-Layer Forwarding

Venkata K. Pingali, Runfang Zhou, Joseph D. Touch  
 USC Information Sciences Institute

{pingali, rzhou, touch}@isi.edu  
<http://www.isi.edu/datarouter>

Application overlays in the Internet such as Chord and CAN enable interesting and useful content-directed forwarding at a significant performance and complexity cost. Hop-by-hop connections used in such systems necessitate application-layer mechanisms to ensure end-to-end reliability and resequencing. Custom re-implementations of these transport-level functions within the application make it difficult to share code across overlay systems and hinder interoperability. Existing network-level overlays lack separate application endpoint identifiers and forwarding based on them. DataRouter is an open, generic packet forwarding facility that uses string rewriting. It augments the traditional, numeric address in an IPv4 or IPv6 header with application-provided identifier strings, embedded as tags in an enhanced IP Loose Source Route (LSR) option. Applications explicitly provide the forwarding information in the tags. A DataRouter indexes a forwarding table to identify the rewriting rule and rewrites the tag.

DataRouter's application-directed forwarding at the network layer has several advantages. Forwarding in kernel is much more efficient than forwarding at the user level (see Figure 2). Existing protocols for transport (TCP, UDP etc.), routing (RIP, BGP etc.) and security (IPsec) can be reused, which eliminates the need for reinventing the corresponding functionality in the application layer. The DataRouting facility and the code base can be shared by multiple overlays at the same time. By concatenating tags belonging to different overlay systems, it is possible to integrate different overlay systems without needing a gateway.

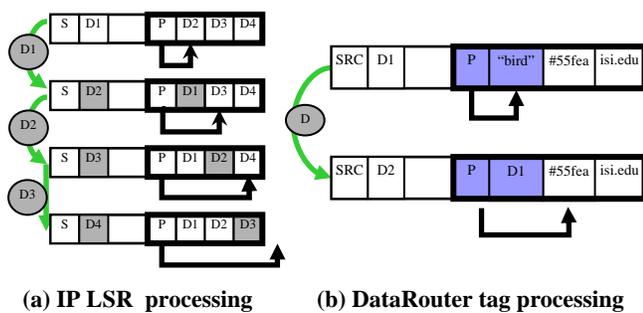


Figure 1 Loose Source Route processing with DataRouter tags

DataRouter tags are embedded in an enhanced IP Source Route option (see Figure 1). The handling of the source route of tags entries is similar to IP Loose Source Route except that the DataRouter tags go through a lookup and rewrite procedure to determine the DataRouter next hop and the final content of the option. DataRouter tag properties include (1) Routing class identifying the overlay (e.g., song hashes, www URLs), and (2) Rewriting rule matching policy: exact, longest, first, range etc.

and (3) Value: the tag string itself. Applications can set or read the tags. The tag rewriting rules are stored in a kernel forwarding table and can be manipulated by the application.

A preliminary implementation of the DataRouter has been completed in FreeBSD 5.0, using a new IPv4 option. It includes a preliminary API for inserting and reading tags and configuring rewriting rules. Applications can set one or more tags through `setsockopt()` system call or through a user-level library. The source route option containing the tag(s) is inserted into all packets originating in that socket. Rewriting rules are stored in a kernel forwarding table using the `droute` command.

On a dual-processor 2.4 GHz Xeon PC with 64-bit/66 MHz PCI gigabit Ethernet cards and running the FreeBSD 5.0, IPv4 packets are forwarded at 332K packets/sec, indicated as "IP/reg" in Figure 2. Forwarding the same packets on a kernel with DataRouter extension support, i.e., where the DataRouting capability is present but not used, does not affect performance measurably. DataRouted packets forwarded based on an exact match of a 32-bit hash decreases performance to around 271K packets/sec (Hash/DR), and forwarding based on a regular expression (in this case, `*.isi.jusc.edu`) results in 149K packets/sec. (RE/DR). These are simple baseline experiments, in which all packets for a test have the same header, and the forwarding tables have only one entry of each type (regular longest-prefix, hash, and regular expression).

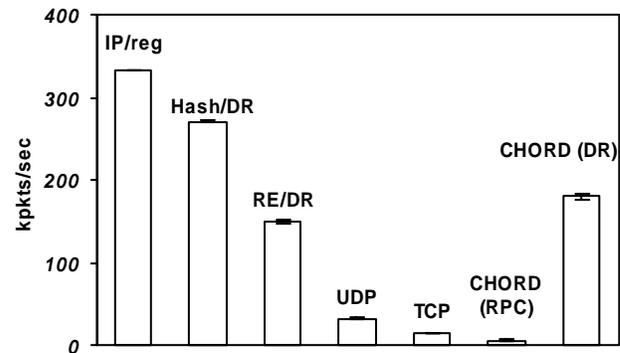


Figure 2 Forwarding rates for IP packets and Chord messages

A preliminary integration of Chord and DataRouter has been completed. Chord is able to forward 5,500 messages/sec using UDP as the transport protocol on the above mentioned platform, indicated as "CHORD (RPC)" in Figure 2. Chord integrated with DataRouter (CHORD (DR)) is 30x faster, at 180K messages/sec.