Resilient Overlay Networks

David Andersen, Hari Balakrishnan, Frans Kaashoek, and Robert Morris
MIT Laboratory for Computer Science
ron@nms.lcs.mit.edu
http://nms.lcs.mit.edu/ron/

A Resilient Overlay Network (RON) is an architecture that allows distributed Internet applications to detect and recover from path outages and periods of degraded performance within several seconds, improving over today's wide-area routing protocols that take at least several minutes to recover.

RON is designed as an application-layer overlay on top of the existing Internet routing substrate. The RON nodes use a combination of active and passive measurements to monitor the functioning and quality of the Internet paths among themselves, and use this information to decide whether to route packets directly over the Internet or by way of other RON nodes, optimizing application-specific routing metrics.

RON is able to potentially provide faster failover than BGP for a variety of reasons. First, BGP's failover is relatively coarse; in 2000, Labovitz noted that, "Internet ... routing convergence is an order of magnitude slower than previously thought," taking an average of 3 minutes to recover from a simple BGP failure, and up to 15 minutes in the worst cases. Second, BGP cannot route around denial-of-service attacks, or even bad links congestion. Finally, BGP's routing policies are relatively coarse, and as a result, a *user* who is authorized to use a backup link may not see that link because of site-wide policies.

Results from two sets of measurements of a working RON deployed at sites scattered across the Internet demonstrate the benefits of our architecture; a summary of these results is shown below. For instance, over a 64-hour sampling period in March 2001 across a twelve-node RON, there were 32 significant outages lasting over thirty minutes each over the 132 measured paths. RON's routing mechanism was able to detect, recover, and route around *all* of them, in less than twenty seconds on average, showing that its methods for fault detection and recovery work well at discovering alternate paths in the Internet. Furthermore, RON was able to improve loss rate, latency, or throughput perceived by data transfers; for example, about 5% of the transfers doubled their TCP throughput and 5% of our transfers saw their loss probability reduced by 0.05. Table 1 shows a subset of these results.

Interestingly, we found that forwarding packets via at most one intermediate RON node is sufficient to overcome faults and improve performance in most cases. These improvements, particularly in the area of fault detection and recovery, demonstrate the benefits of moving some of the control over routing into the hands of end-systems.

Loss Rate	RON Win	No Change	Ron Loss
10%	557	165	113
20%	168	112	33
30%	131	84	18
40%	110	75	7
50%	106	69	7
60%	100	62	5
70%	93	57	1
80%	87	54	0
90%	85	48	2
100%	67	45	1

Table 1: Outage data for RON_2. A "RON win" at $p\%$ means that the loss rate of the direct Internet path was $\geq p\%$ and the RON loss rate was $< p\%$. Numbers in brackets show the contribution to the total outage number after eliminating all the (typically more reliable) Internet2 paths, which reflects the public Internet better.

RON raises as many questions as it answers, about scalability, security, and improved Internet routing. The wider use of RON provides a platform on which many of these future issues can be explored and measured. We encourage interested readers to visit our website (`http://nms.lcs.mit.edu/ron/`, or read our upcoming SOSP paper that describes and evaluates RON in more detail[1].

1. References

[1] ANDERSEN, D. G., BALAKRISHNAN, H., KAASHOEK, M. F., AND MORRIS, R. Resilient overlay networks. In *Proceedings of the 18th ACM Symposium on Operating Systems Principles (SOSP) (to appear)* (Oct. 2001).