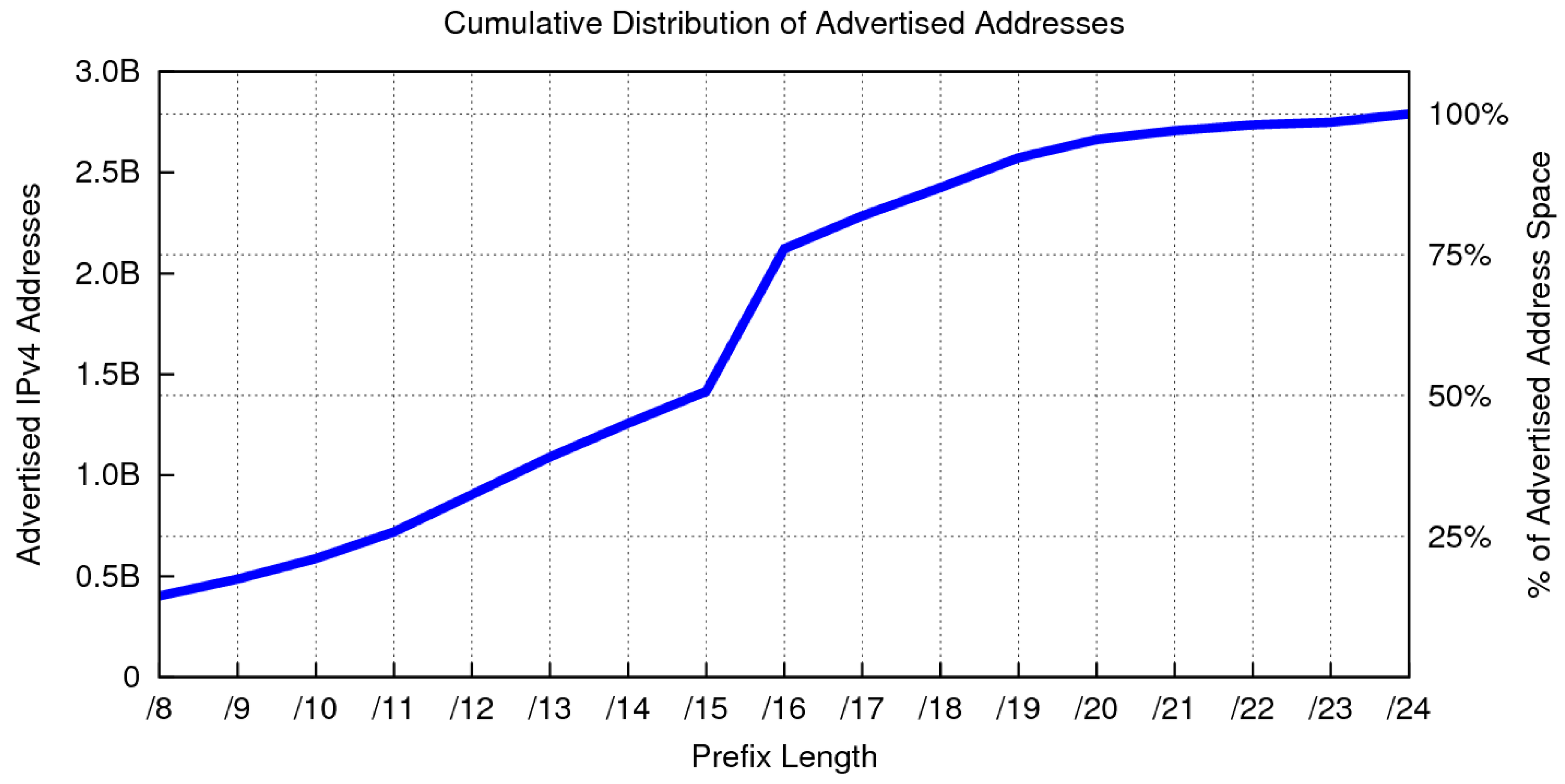
Deterministic, Reduced-Visibility Inter-Domain Forwarding

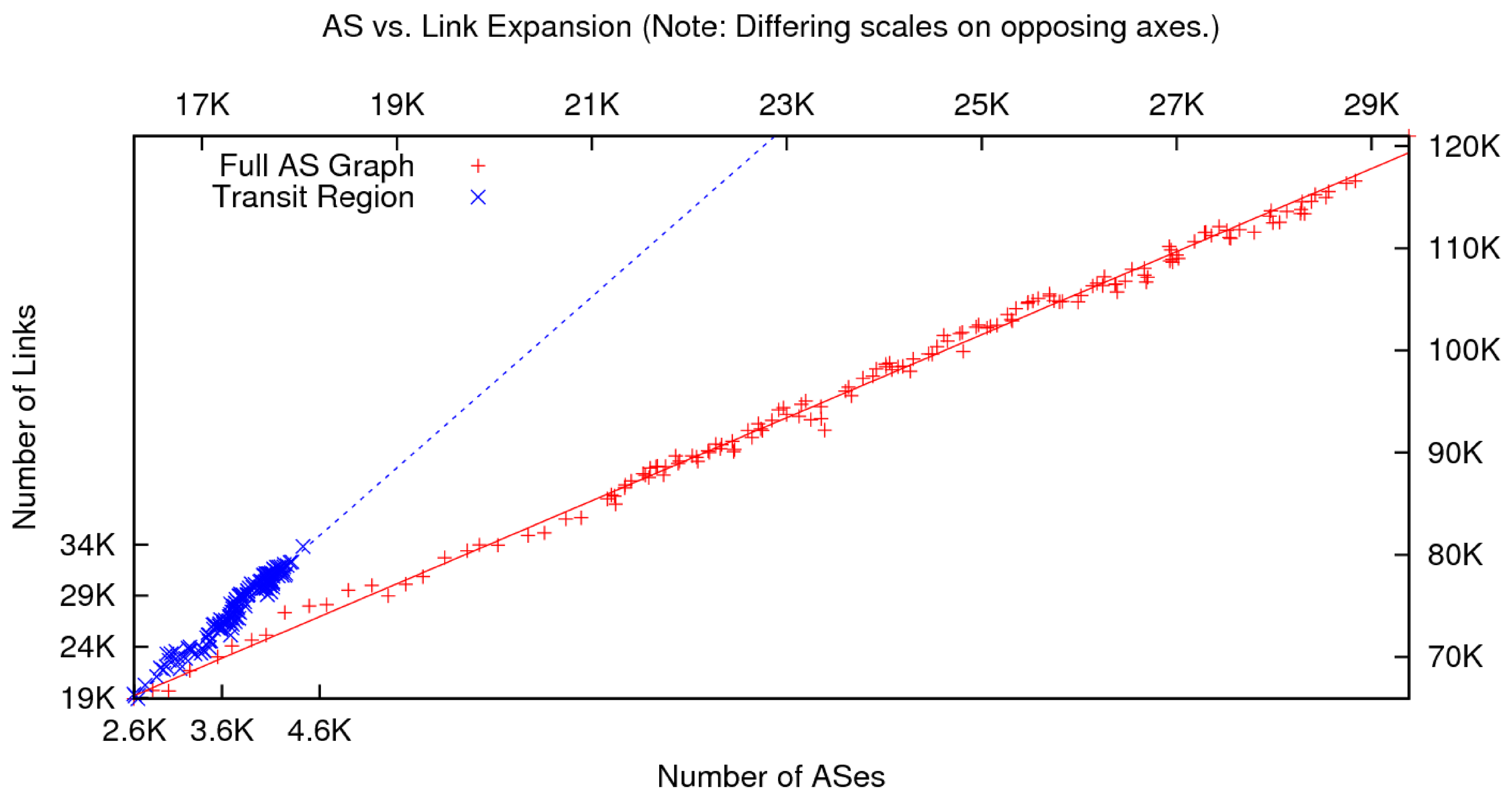
Stephen Strowes and Colin Perkins
sds@dcs.gla.ac.uk, csp@csperkins.org
Department of Computing Science
University of Glasgow

Problem: Increased Forwarding State

Inter-domain forwarding state at core Internet routers is growing at approximately 17% per annum. Table growth is exacerbated by adoption of edge-site multi-homing, over and above the natural growth of the number of advertised prefixes.

Larger forwarding tables incur more time-consuming lookups, threatening a router's ability to forward packets at line-rate as table sizes increase.





Advertised Address Space

Long prefixes dominate tables: The figure above shows 75% of the advertised address space is covered by /8's through /16's; these account for 4.3% of the advertised prefixes.

55% of the advertised prefixes are /24's, which cover only 1.5% of the advertised address space. A further 58% of these /24's are de-aggregates from other advertisements.

Network Structure

Autonomous Systems (ASes) which offer transit services are well-connected, offering many potential routes between pairs of ASes with a mean AS out-degree between 7 and 8. This region continues to grow in density (the figure above shows a higher rate of link creation).

The density of this region leads to a shallow graph, with a mean distance between pairs of ASes between 3 and 4 AS hops.

Reducing State

We define a prefix's *utility* based on the length of the prefix and the distance travelled from its origin. Shorter prefixes, and prefixes few hops from their origin, have the greatest utility. Prefixes of length /8 to /16, and those advertised by direct neighbours, have a utility of 1.

With this utility measure, our FIB generation algorithm randomly selects a subset of the advertised prefixes:

FIB Generation:

```
On BGP update:
    r = random()
    if r > utility(prefix):
        Install path in FIB
```

To recover from missing entries, we forward packets between transit ASes by hashing over IP headers to retain determinism. We assert that this graph is dense enough to recover from missing FIB entries by deterministically forwarding packets between other transit networks.

Deterministic Hash-Based Forwarding:

```
On data packet:
    perform longest prefix match
    if (match found):
        Forward packet
    else if (no default route):
        Forward on port selected by hashing IP header (not incoming port)
    else:
        Use default route
```

Ongoing Work

We have developed a new AS graph simulator to test these ideas, amongst others. Our analysis will focus on the relationship between the size of forwarding tables generated at a given node, vs the effect on path stretch across the network, when altering the utility measure of a prefix (*u*).

References

[1] T. Bu, L. Gao, D. Towsley. On Characterizing BGP Routing Table Growth. Computer and Telecomm Networking, 45(1), 2004.
[2] G. Huston. BGP in 2008. http://potaroo.net/ispcol/2009-03/bgp2008.html
[3] D. Meyer et al. Report from the IAB Workshop on Routing and Addressing (RFC 4984), Sept. 2007.