Curtis Villamizar, Drew Perkins

Forward looking statements made in this presentation are the speculation and opinions of the authors and are provided here soley for informational purposes. No statement in this presentation should be interpreted as a statements of commercial intent.

Author affiliation is provided solely for identification purposes.

NANOG 42 Lightning Talks - Total Network Costs

We've been doing some simple total network cost analysis.
This is a glimpse at some key findings.

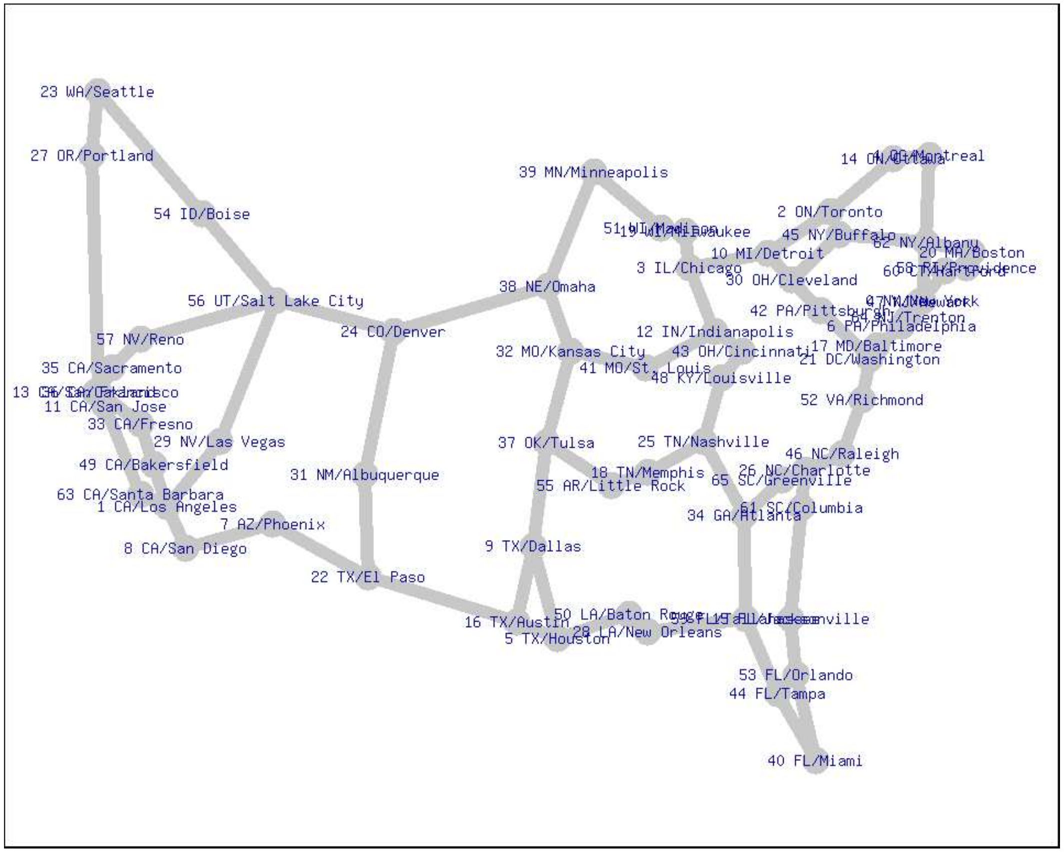
Contents :

Maps – a sample network

We need to rush through this. Look a the slides later.

Contents (previous contents page: 2) :

Fiber map



We had to start with a fiber topology.

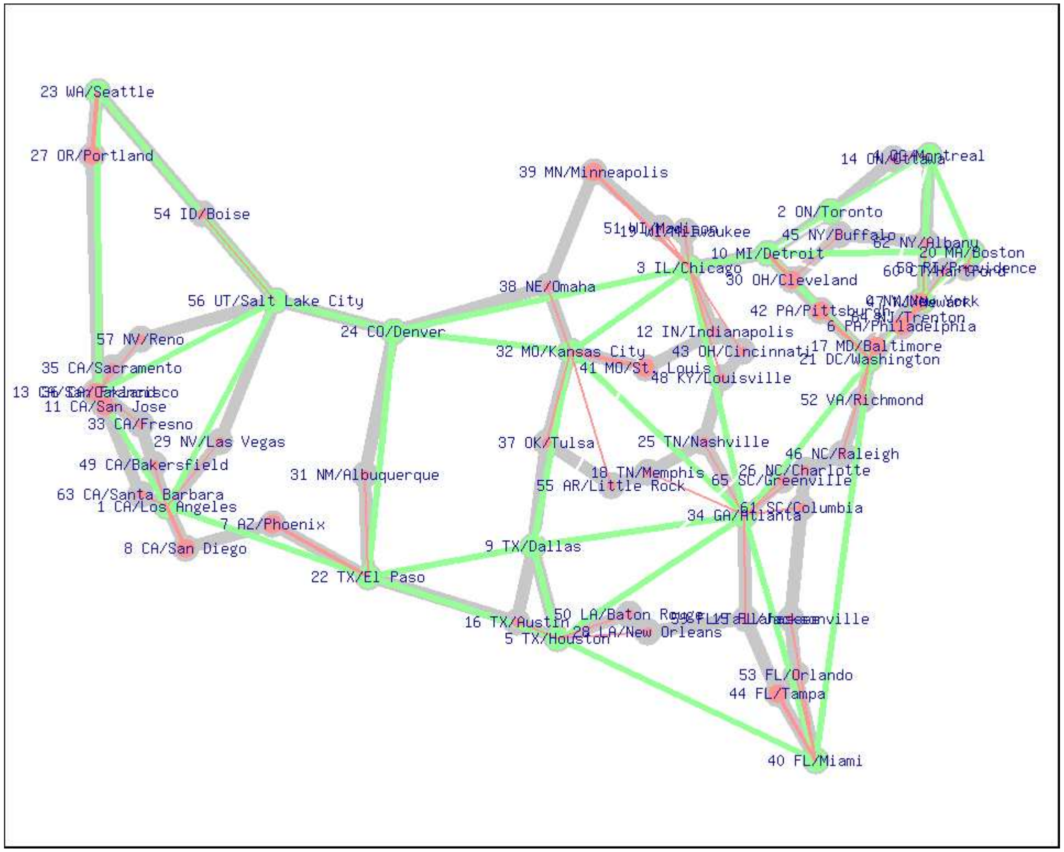
We made one up. We call this our "reference fiber network topology".

We certainly can't do a NANOG presentation with a customer's topology.

This topology should look at least vaguely familiar.

Note that amplification and regen sites are not shown.

IP core nodes and links

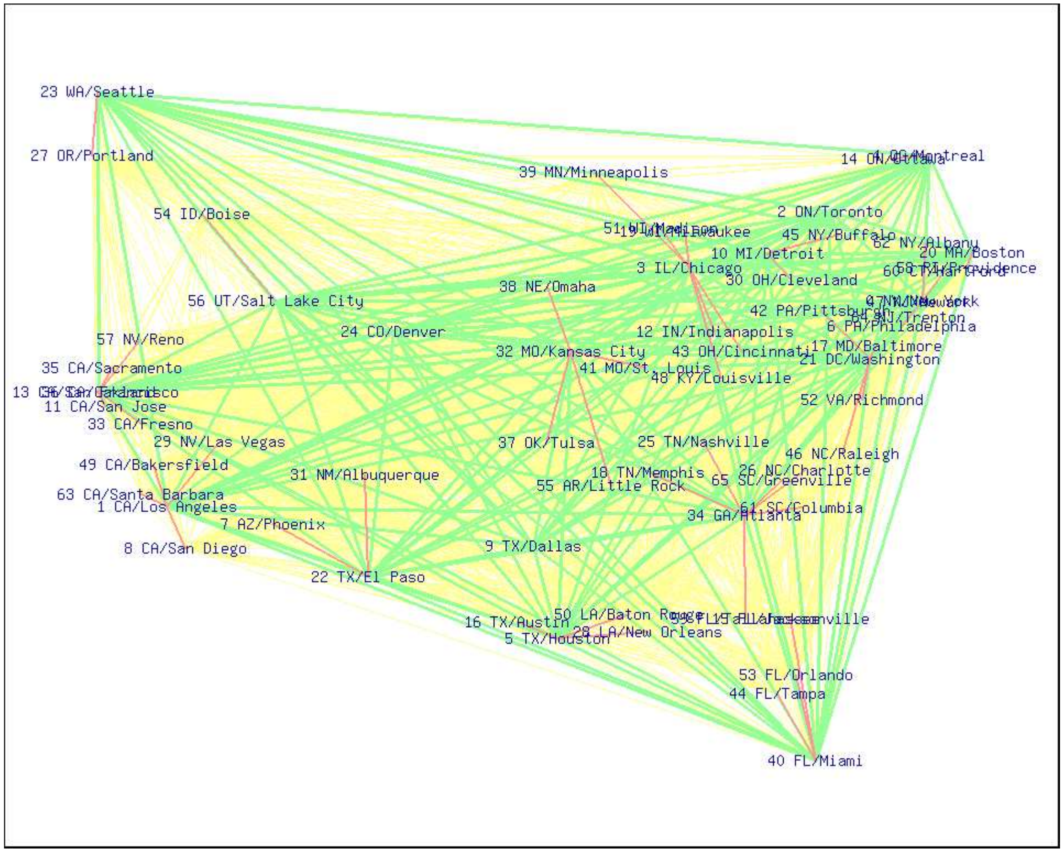


Core nodes were also hand picked.

Some simple software (a bit too simple) picked the core links.

The result is not too bad but maybe typical. In some cases junction nodes might be better off being core. A few nodes are higher degree than they need to be for smaller demands.

IP traffic demands



There are 66 nodes therefore 66 x 65 (unidirectional) city to city demands. Most demands are tiny.

A gravity model was used to obtain city to city demands. We plan to check the distribution against provider data but haven't done that.

Aggregating city to city demands onto the regional links (in red) and core links (in green) makes for large demands, IP mux gain, and far less underfilled pipes.

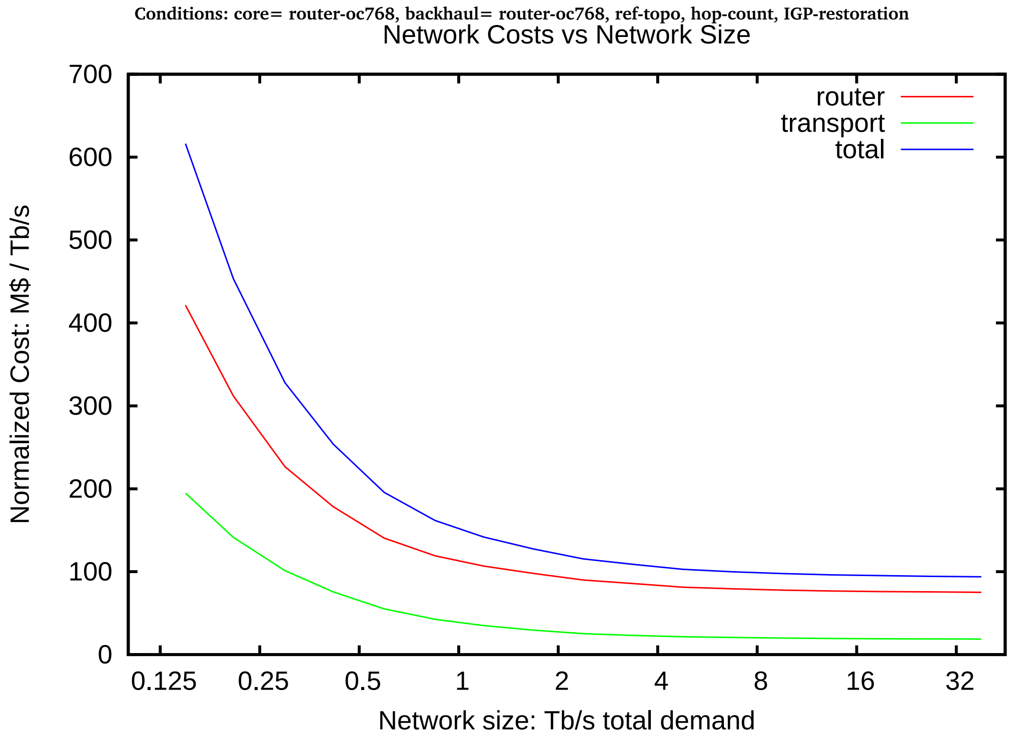
In graphs that follow the demands were scaled linearly over a range of 1 to 256, with the lower end of the range being somewhat smaller that the largest IP providers today.

Costs with various interface types

Vendor names are omitted to protect the guilty. Costs are based on major vendors and the steep discounts typically given to very large deployments. Please ignore the cost per bit values as YMMV.

Contents (previous contents page: 2) :

Costs for router OC-768



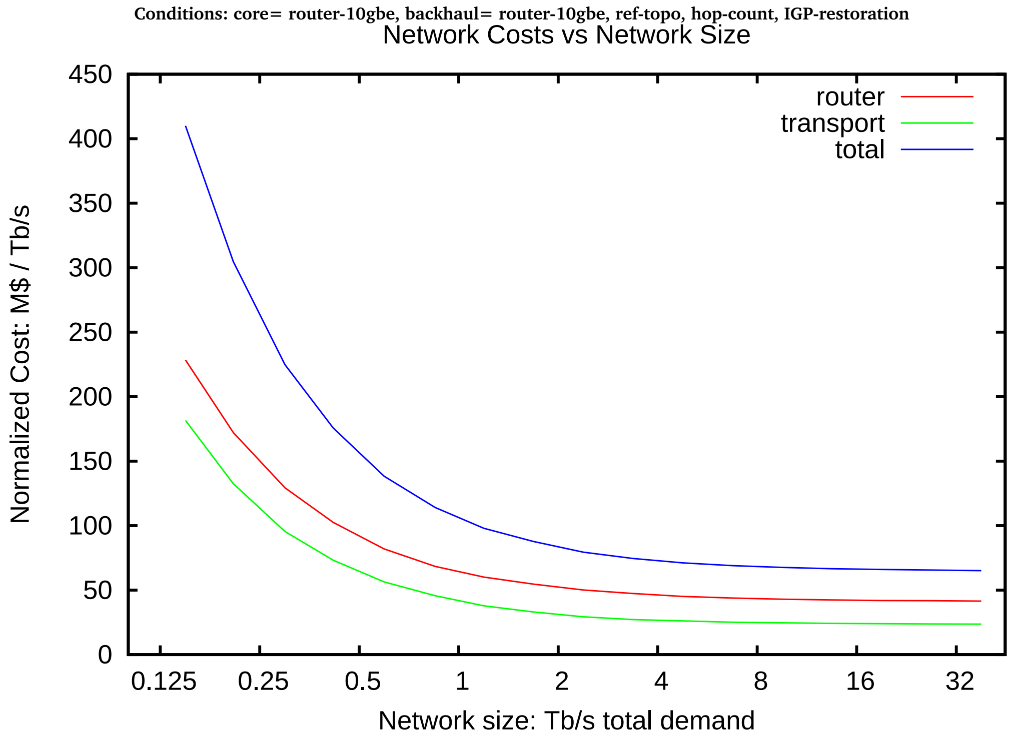
This buildout uses 1 x OC-768/SR on large routers and uses digital transport.

At low demands both router and transport capacity is underfilled. Underutiled interfaces, switching fabrics, chassis, and fibers yield higher cost per bit.

The transport cost is less than the router costs even at low demands.

Transport costs are very low compared to the router cost as demands increase.

Costs for router 10GbE



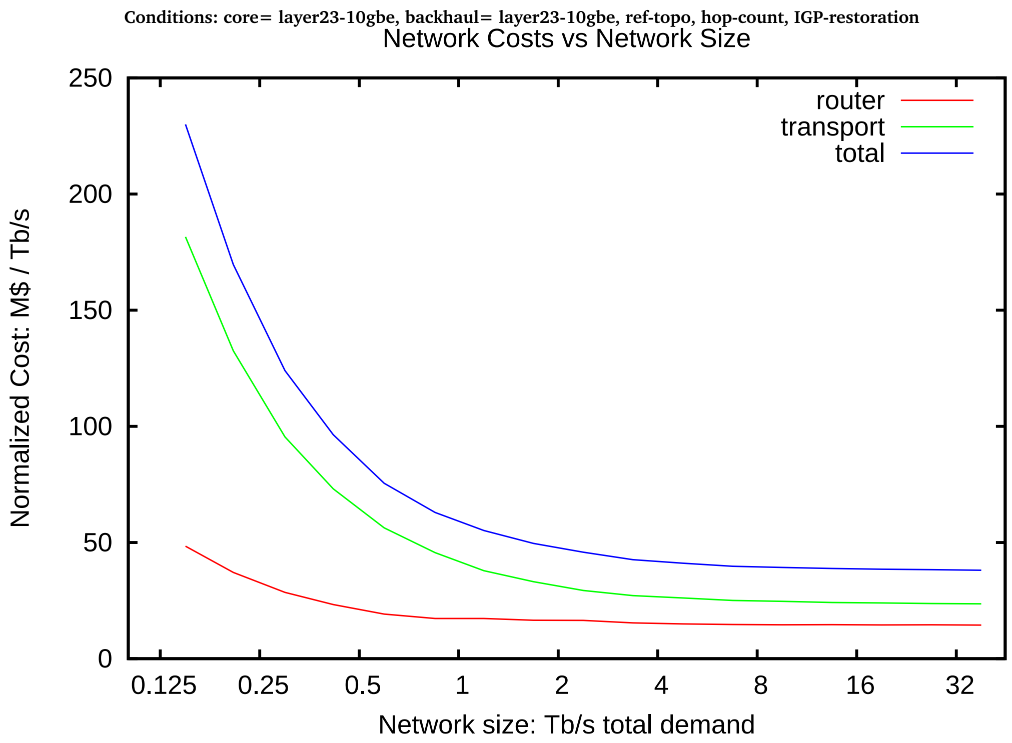
This buildout uses 10GbE on large routers and uses digital transport.

Router costs are much lower for 10GbE than for OC-768.

At low demands router and transport capacity is under-filled but router capacity is better utilized.

As demands increase, transport costs become lower than router costs by a very significant margin.

Costs for layer-2/3 10GbE



If the largest layer-2/3 switches are used in place of larger router and arranged in a CLOS, the router cost is reduced substantially.

As a CLOS becomes necessary at more sites, these small switches become less cost effective but still far more cost effective than large routers.

At high demands transport and router costs are close, with transport costs being somewhat higher.

IP router bypass

Contents (previous contents page: 2) :

Analysis using IP Router Bypass

An IP bypass link bypasses an IP core router on the fiber path.

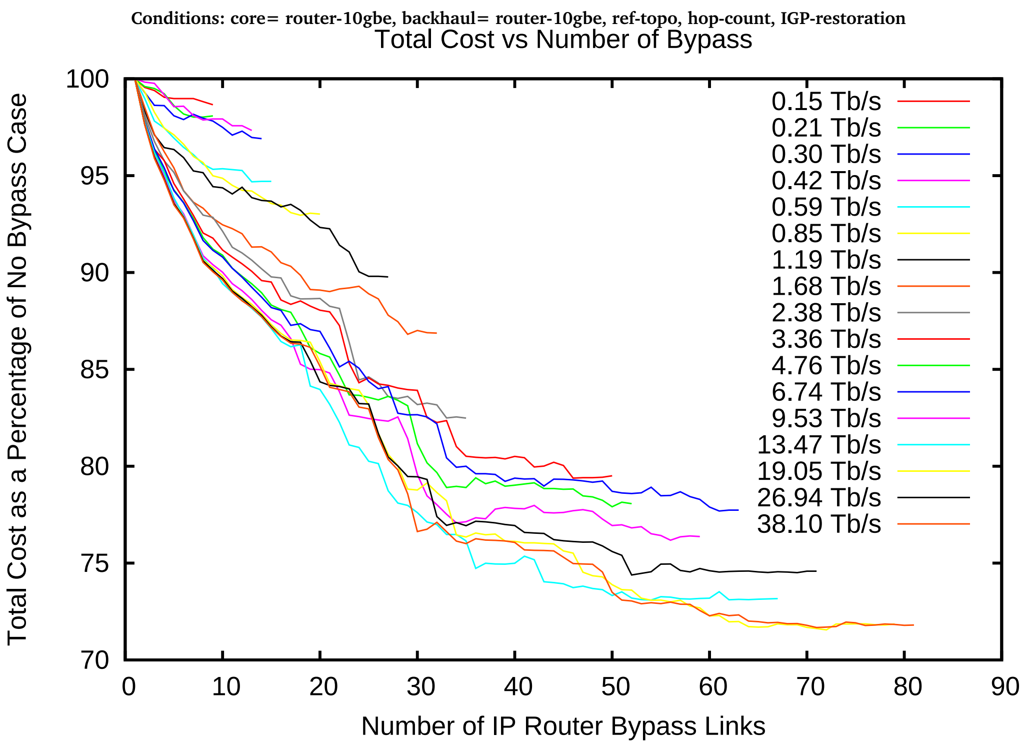
The goal is to reduce the number of router interfaces and transport tributary interfaces and therefore reduce cost.

A practical limit is imposed by traffic variations which can make the loss of multiplexing by the core router an issue.

Today only a small number of bypass links are cost effective.

Applying bypass eventually yields a full mesh of IP links, though the largest scale studied does not result in a full mesh.

IP Router Bypass Cost Savings

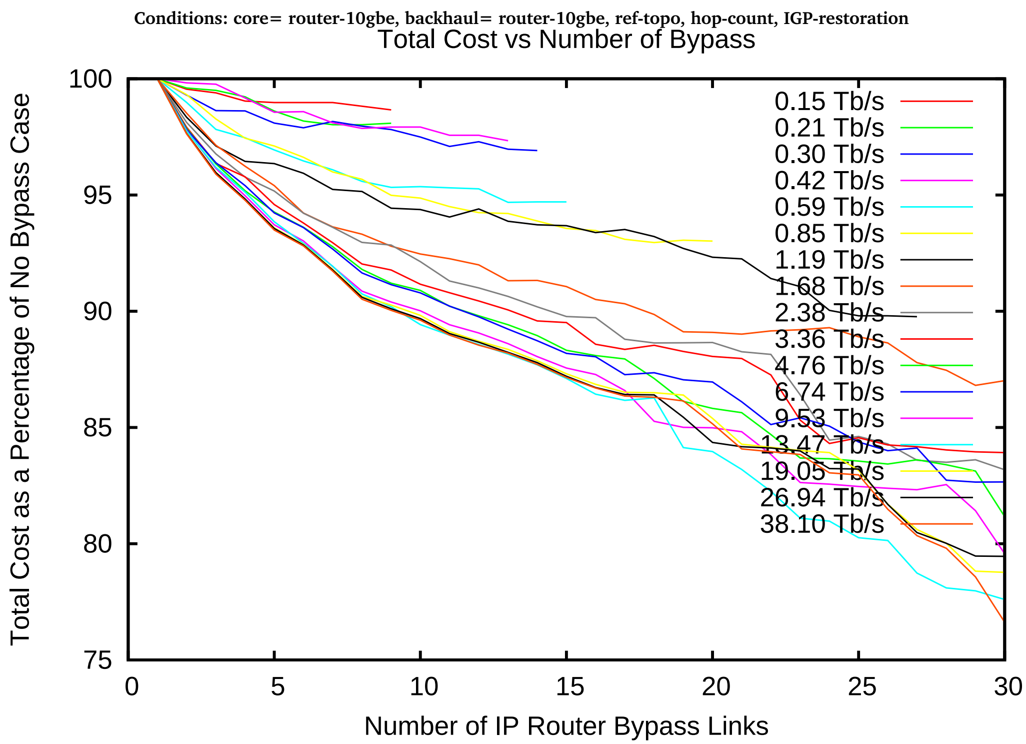


Software applies the bypass link which most reduces the number of router interface and iterates until no further reduction is possible.

Each line represents data at one scale (total demand) where the number of bypass links is varied.

Router cost savings are up to 35%. Transport cost savings are less dramatic. Total cost is reduced by 20-28% (shown) at scales from 4.8 to 38 Tb/s.

Limited Use of Bypass

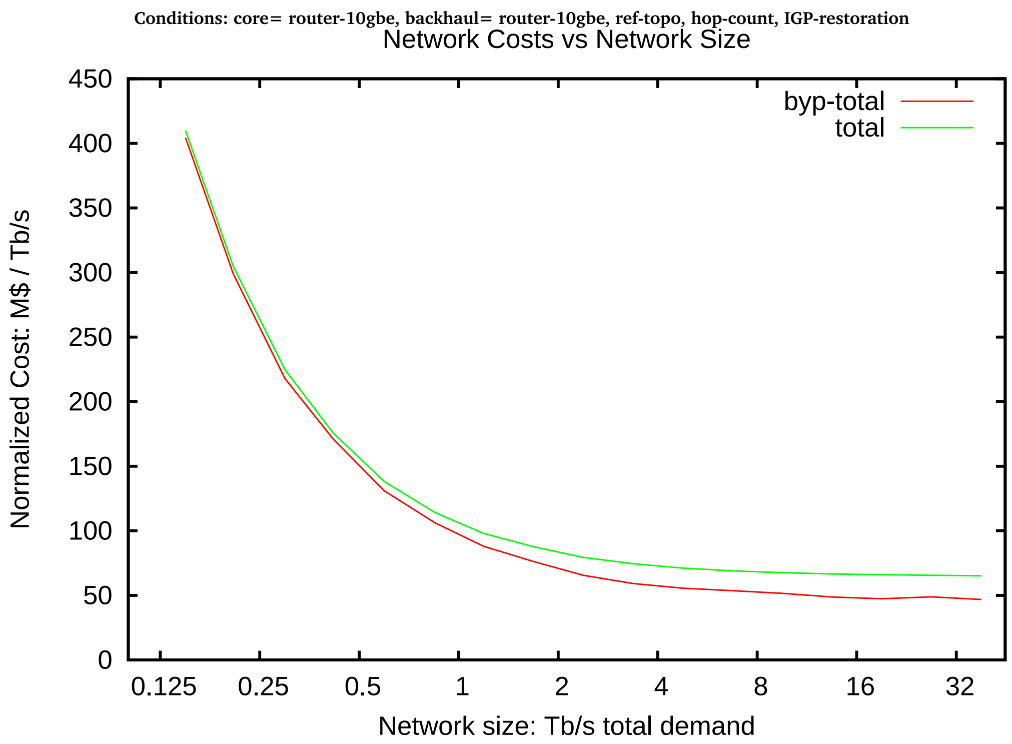


In this plot the X-axis is constrained to 30 bypass links.

Router cost savings of up to 25% are possible with 30 bypass links. Total cost savings of 20% are achieved at high scale (shown).

In all but the smallest scales, router cost savings of 10% to 15% can be achieved when adding only 10 bypass links. Total cost savings with 10 bypass links are 5% to 10% in all but small scales.

Maximum Use of IP Router Bypass

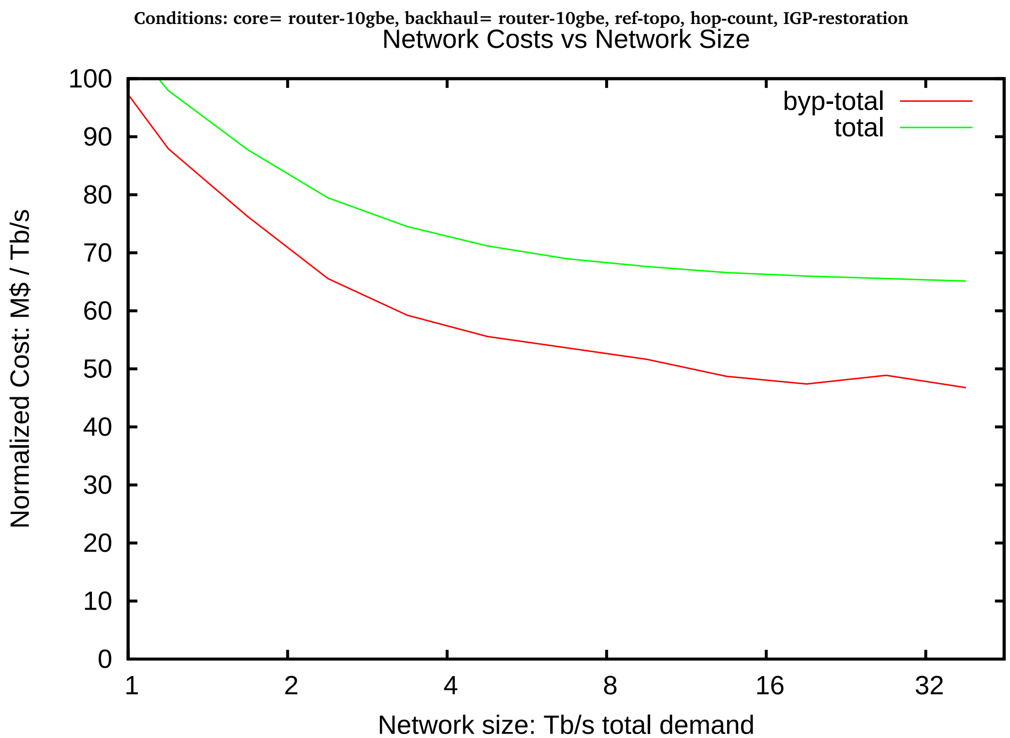


The cost per bit using router 10GbE with no bypass is compared to the cost per bit of applying bypass links until no further gain is possible.

Maximum IP router bypass != IP full mesh.

Continuing to add bypass links until a full mesh is reached increases cost. At very high scale this increase is small but only if 10 Gb/s interfaces are used (later slides).

Maximum Bypass - Large Scale

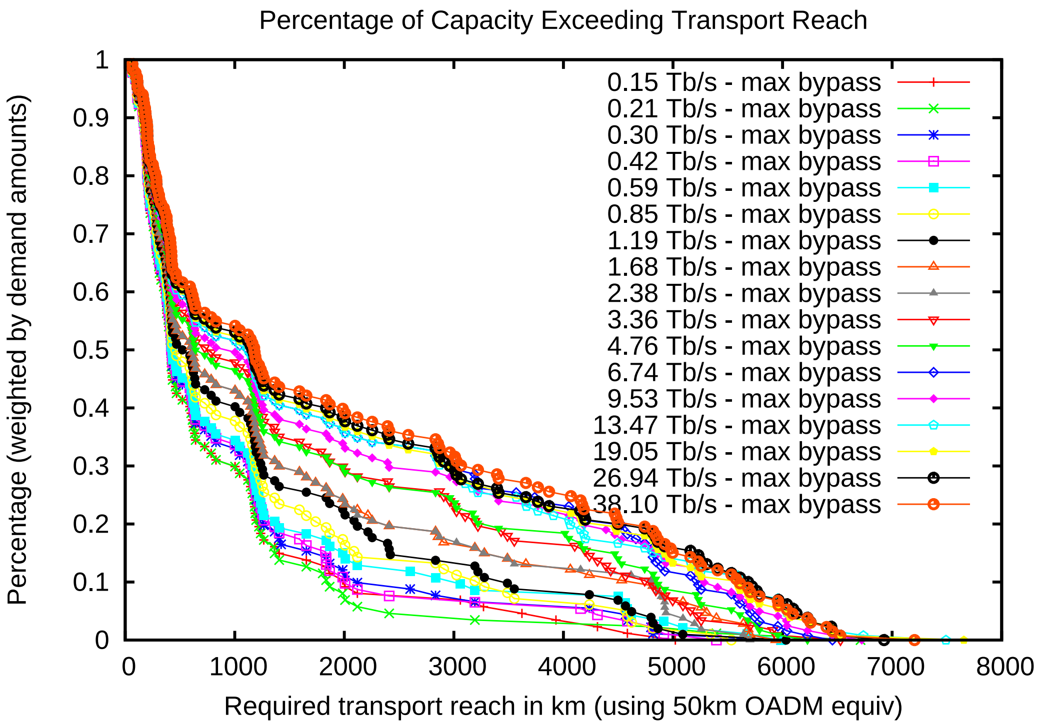


Total cost per bit of using no bypass and using maximum bypass are shown for large scale only.

Total cost savings (router plus transport) reaches about 20% and levels out.

Due to the scale of the plot on the prior slide the cost difference is more difficult to see on that slide.

Impact of Bypass on Reach Requirements



At small scale little or no bypass is used, therefore there is no impact on reach requirements.

As bypass links are added, the fraction of total capacity that must span long distances and pass a large number of OADM increases dranatically.

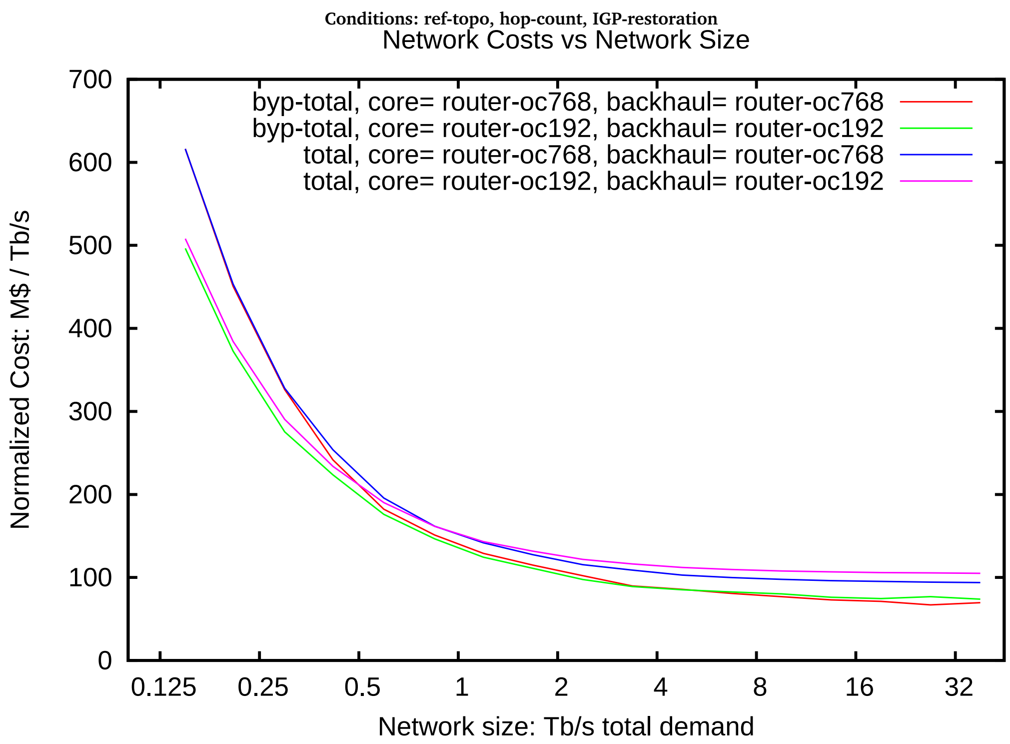
30% of capacity exceeds a 3,000 km reach when a 50 km penalty per OADM is added (25% with no OADM penalty) in this topology.

Total costs

These plots put information from prior plots together for easier compasison.

Contents (previous contents page: 2) :

Total costs for POS and 10GbE

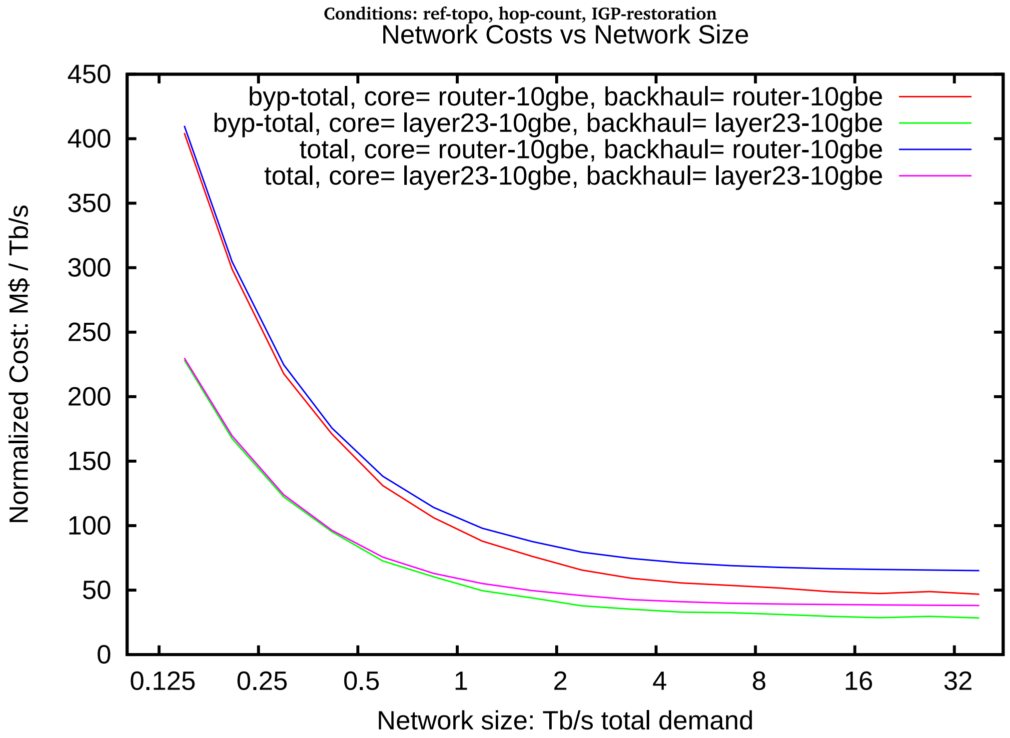


IP router bypass gives OC-192 a slight cost advantage over OC-768 within a range where 10Gb/s provides more opportunity for bypass than 40 Gb/s

At high scale OC-192 and OC-768 have nearly equal opportunity for bypass. At high scale OC-768 and OC-192 are about equally cost effective.

At high scales IP router bypass can save one quarter to one third of total network cost.

Total costs for 10GbE



With or without bypass, OC-192 and OC-768 are more costly than 10GbE.

Choice of router 10GbE vs Layer-2/3 switch 10GbE has very significant impact on costs.

Use of bypass has significant impact of costs.

The impact on total cost is greater for router 10GbE than layer-2/3 10GbE due to the cost contribution of transport which is only slightly reduced when bypass is used.

Other work in progress

Interesting topics not covered in this presentation include:

Impact of using MPLS/TE on network costs.

Restoration techniques.

Space and power requirements (easy enough to determine).

Examination of router DWM / alien wave solutions.

Future technology that exist as research today.

Summary

Total cost for OC-768 is higher than total cost for OC-192. OC-768 remains more costly for the foreseable future if router bypass is used.

The total cost of POS is much higher than the cost of 10GbE regardless of router vendor.

Layer-2/3 switches provide even lower cost even when used in a CLOS configuration.

IP router bypass can substantially reduce cost. Practical limitations on IP router bypass are imposed by the amount of traffic change. We need provider data to quantify how much bypass is expected to be practical.