>BUSINESS MADE **SIMPLE**

Ethernet, MPLS, Where and How Much?

Tim Pearson

Leader, Asia Business Development

timp@nortel,com

MPLS JAPAN 2007

NORTEL

Introduction

> How far you push MPLS from the WAN into the MAN is a key decision facing service providers

> Robust new developments in ethernet technologies change the question from *"**how far**"* you push MPLS into the metro, to **"how much and why?"**

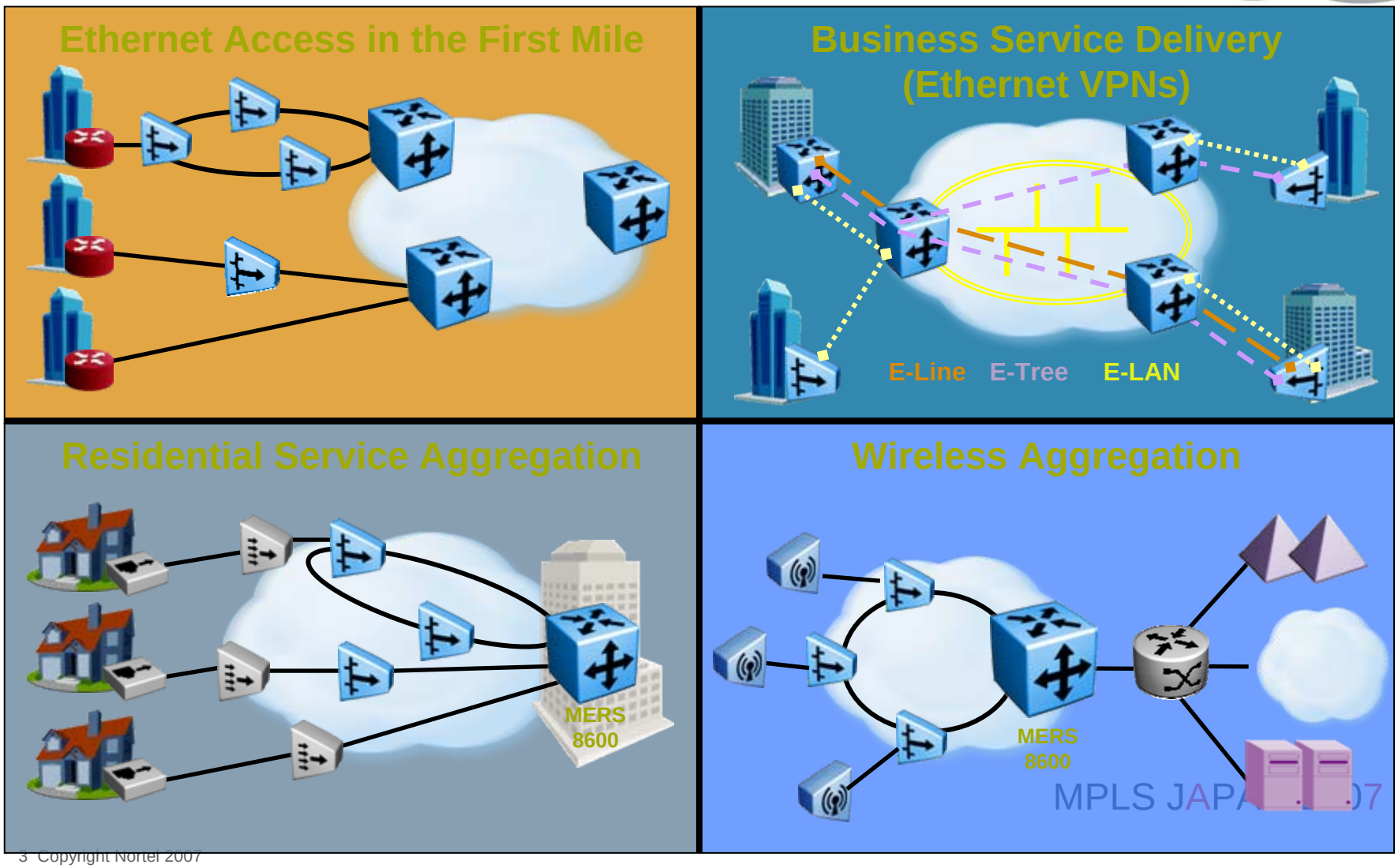
> The relationship of MPLS and Ethernet will dominate this decision due to technology and architecture fundamentals

> An Ethernet metro solution that offers transit for non-Ethernet services based on MPLS adaptations is *the* scalable next generation solution.

MPLS JAPAN 2007

Metro Ethernet Applications

Overview of Deployment Scenarios



MPLS in 2007

> Defined in such a way that it touches most other networking technologies
  - E.g. GMPLS into L1 and optical, ubiquitous shim for packet processing

> A fundamental value is the combination of automation and autonomous resiliency offered by the set of MPLS control planes.

> MPLS dominates the WAN, the question has become will the virtues of MPLS automation stand up to the scaling requirements of the MAN.

> A key value of MPLS is the control plane, so question becomes how does control plane fare when pushed to the edge.

MPLS JAPAN 2007

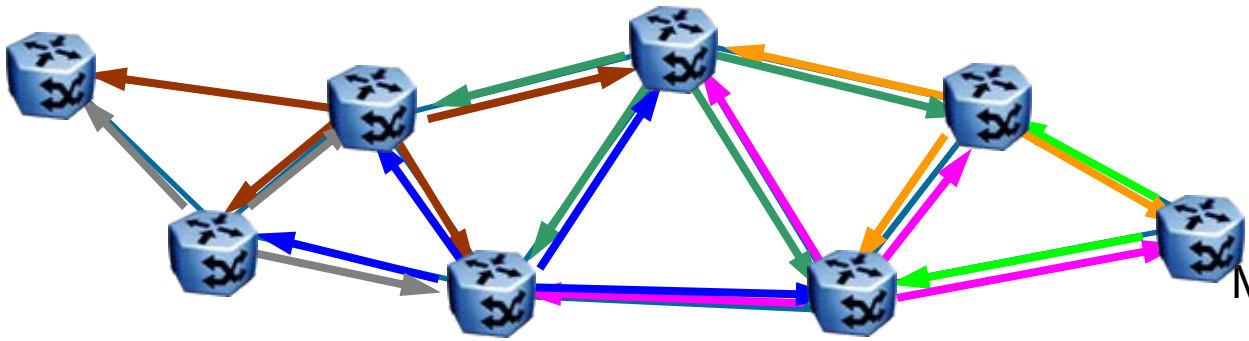
Nature of LDP

> LDP is a label flooding mechanism

> Each device when it learns of a FEC offers a local label for the FEC to it's peers
  - either via routing/Independent mode, or via LDP peers/ordered mode

> Per platform labels/liberal label retention is most common usage
  - Minimizes per-peer state, and simplifies FIB refresh as it produces a common per platform FIB

> Amount of state per LSR per interface goes up linearly with the number of PE /32 FECs in the network



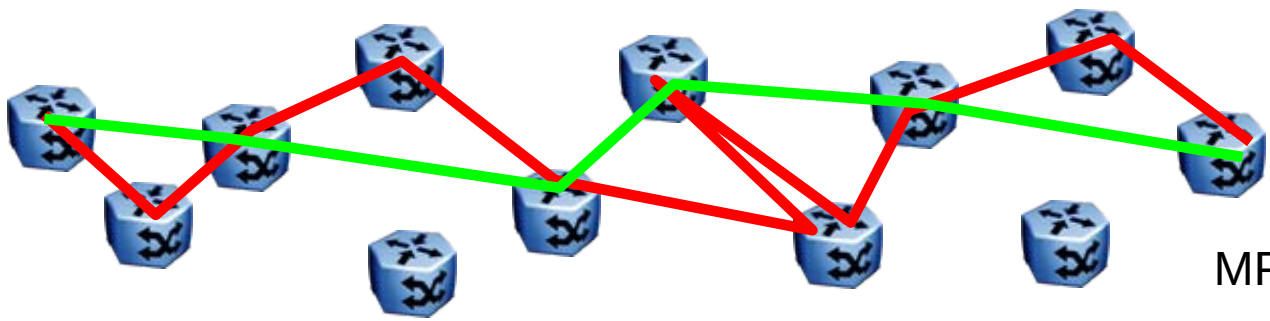
Liberal/Independent

Nature of RSVP-TE

> RSVP-TE is focused on P2P and more recently P2MP

> RSVP-TE hellos giving way to FRR or e2e BFD

> RSVP-TE + FRR is N-squared LSPs plus backup per potential point of failure
  - Per link for facility, per LSP per hop for detours



FRR
Primary and
Facility backup

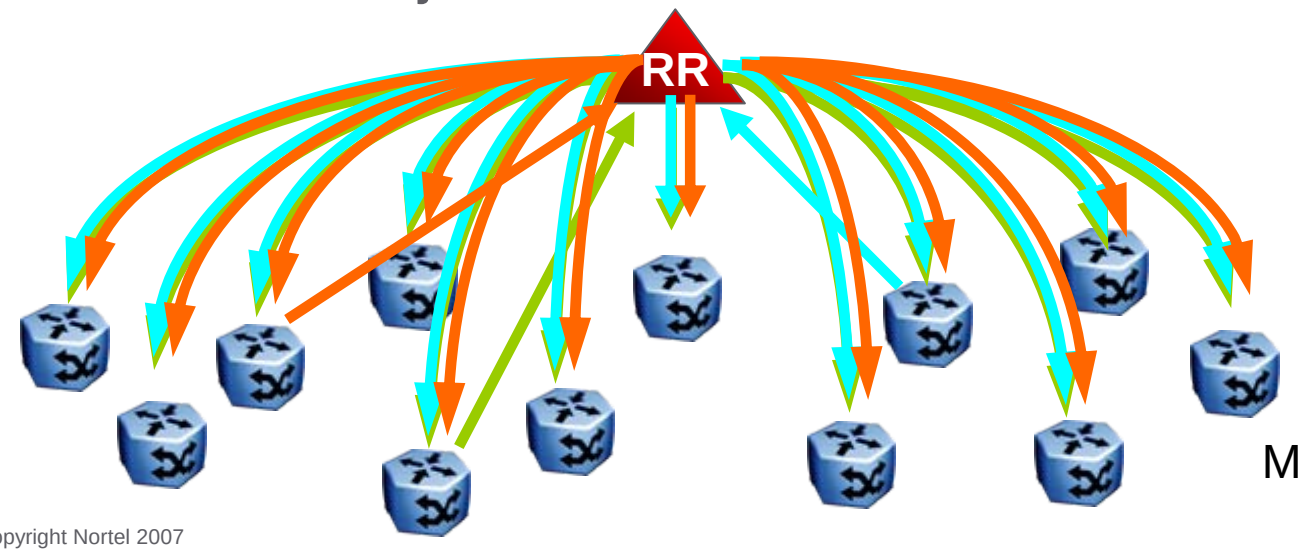
MPLS JAPAN 2007

Nature of BGP

> BGP in an MPLS context is basically a flooding mechanism for service information between PEs

> Primary uses are auto-discovery (L2VPN) and to relay customer reachability information (L3VPN)

> Basic “route target” receiver filtering paradigm and nominal full mesh of adjacencies needs route reflectors to scale



MPLS JAPAN 2007

Network build

> What does the network look like the closer to the edge you get?

- Order of magnitude more devices
- Progressively smaller community of interest for any given device
- Increased price pressure
- Increased difficulty of craftsperson access
  - Unmanned sites
- Desire for reduced complexity, footprint and powering requirements
  - Real estate and enclosure start to dominate overall cost

So what sort of MPLS options exist for the MAN?

MPLS JAPAN 2007

Option 1: MPLS to the Edge

> This means

- relatively flat LDP PSN (/32 labels)
- Multi area solutions to achieve some degree of isolation
- Large number of labels on any given interface
- VPLS N-squared becomes a problem->H-VPLS->resiliency
- RSVP-TE N-squared becomes a problem
- Amount of BGP configuration and route reflectors is a concern

The wall you will hit is the amount of telemetry collected that is of little value

MPLS JAPAN 2007

Option 2: MPLS L2-L3 Islands

> We're seeing a trend to L3 WANs and Emulated L2 MANs
  - 2547 in the core, VPLS in the metro

> We assume VPLS is the current choice in the perceived absence of other solutions

> Produces a decoupled solution
  - L2 MAN, L3 WAN

This produces a complex metro simply because L2 is emulated instead of natively recreated

MPLS JAPAN 2007

Option 3: Ethernet PSN

> Third option is to produce decoupled islands but using Ethernet instead of MPLS

> Much of the world already doing Ethernet

- Swiss Army Knives have Ethernet RJ45 on the back
- Broadband is already going Ethernet
  - DOCSIS, DSLF “TR-101 migrating to Ethernet aggregation”

> Issue is other legacy services, and how far do you invest in continuing to support/transition declining markets

- FR/ATM, CES etc.

MPLS and Ethernet

- MPLS initially defined as IP helper, so universal carriage over any link layer is defined
  - If it carries IP, it carries MPLS
- Ethernet is ubiquitous, so universal carriage over any link layer is defined
  - 802.3, 802.17, GFP/SONET/SDH, MPLS etc.
  - Ethernet has it's own link layers as well
- MPLS & Ethernet each can encapsulate and carry the other
- PEs have Ethernet UNIs and NNIs

However…

… not all semantics of the relationship translate evenly

- Ethernet carrying MPLS starts as a broadcast medium and filters to produce more selective behaviors
  - The atomic unit is a broadcast LAN segment
- When MPLS PEs are directly connected by Ethernet, MPLS declares itself redundant
  - PHP
- MPLS carrying Ethernet clones P2P tunnels to produce a split horizon “full” mesh
  - The atomic unit is a P2P PW

Emulating Broadcast with a PW Mesh

> Places a lot of burden on the PEs

> May keep some state out of the core
  - But at the expense of multicast inefficiency AND
  - N-squared PW control, OAM and telemetry

> H-VPLS improves the situation, but does not change the fundamentals
  - “many” MPLS PWs to emulate Ethernet “one” Ethernet segment

> Clearly carrying MPLS with Ethernet is different than carrying Ethernet with MPLS

Why emulate broadcast with connections built on a broadcast media?

So, when considering the network…

- It is the relationship between MPLS and Ethernet that will dominate choices
- If MPLS recreation of Ethernet LAN segments has clear scaling issues, lets take that off the table
- Lets do Ethernet WITH Ethernet and leave MPLS to do what it does well

Eliminate the scaling burden that Ethernet as the dominant service will place on MPLS

Ethernet Tools to Scale

- Defined solutions:
  - Link layer: LAG, RPR etc.
  - 802.1ad QinQ,
- Separation of “Ethernet as a service” from “Ethernet as infrastructure”..***802.1ah MACinMAC***
  - Secure the MAN/WAN
  - Push customer state to the edges
  - Easy to instrument known end points
  - Carrier in complete control of all aspects of Ethernet forwarding
- OAM… ***802.1ag CFM/Y.1731***
  - Fault and performance management OAM for Ethernet
- Configure “Ethernet infrastructure”…***802.1Qay PBB-TE***
  - Engineered p2p and p2mp

MPLS JAPAN 2007

Future Innovations – 2008/2009

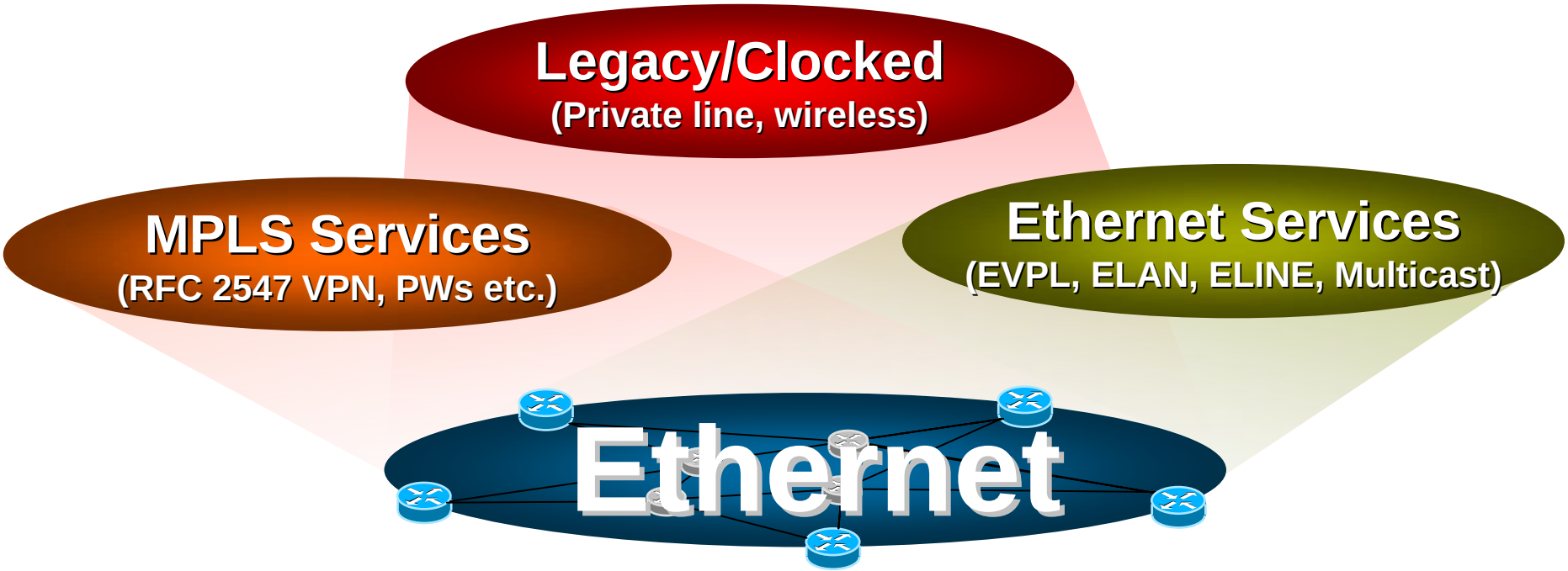
> Infrastructure

- L1/L2 clock distribution via Ethernet PHY/Ethernet frame
- eDCO – DSP dispersion compensation for optical
  - 1600km reach with no regeneration or engineering of fiber
  - Control plane
  - Routing for link state bridging plus connection automation

> Services

- PW adaptations of legacy L2 onto Ethernet directly
  - “Dry Martini”
- Simple and scalable L2VPN side by side with PBT
  - MACinMAC + PBT + Routing system + loop suppression
  - near zero-config ELAN +ETREE

The vision – Converging on Ethernet



Ethernet with LANs and connections as the infrastructure of choice

MPLS JAPAN 2007

In Conclusion…

- Ethernet is fundamentally the most flexible networking technology
  - It is a commodity with vast untapped capabilities
- We can re-purpose Ethernet switches with new control software to broaden their role in the network
  - TE, SPF in addition to bridging and multicast….
- This can be done with minimal changes to Ethernet standards
  - *And they are well underway*
- This lets us radically delayer and simplify the network

MPLS JAPAN 2007

For Further Reading

Fedyk, Don, David Allan, et. al. “GMPLS control of Ethernet” IETF Draft June 2006 <http://www.ietf.org/internet-drafts/draft-fedyk-gmpls-ethernet-pbt-01.txt>.

Allan, David, Nigel Bragg, et. al. "Pseudo Wires over Provider Backbone Transport. " IETF Draft July 2006 <http://www.ietf.org/internet-drafts/draft-allan-pw-o-pbt-01.txt>.

Allan, Bragg, McGuire, Reid. "Ethernet as Carrier Transport Infrastructure." IEEE Communications 44(2006): 134-139.

Allan, Bragg “Taking Control: The evolving role of control and data planes” Nortel Technical Journal, Issue 4 http://www2.nortel.com/go/technicaljournal_index.jsp?locale=en-US

Fedyk, Bottorff “Provider Link state Bridging”,

http://www.ieee802.org/1/files/public/docs2007/aq-fedyk-provider-link-state-bridging-0107-01.pdf

MPLS JAPAN 2007

Questions?