PWE3 Congestion Considerations

draft-stein-pwe3-congcons-01.pdf

Yaakov (J) Stein

David Black

Bob Briscoe

PW congestion as seen by PWE3

PWE3 was originally in the transport area
because handling the congestion issue was considered critical

The TDM PW drafts were accepted by the IESG
only after considerable work on their congestion considerations sections

The only draft that devoted entirely to a congestion issue
draft-stein-pwe3-ethpwcong
was extremely limited in scope
and was abandoned due to lack of interest in the WG

PWE3 as a WG has a long-standing commitment
to deal with the congestion problem

PW congestion as seen outside PWE3

The problem is often phrased as follows:

- PW traffic may be carried over IP networks
  - L2TPv3 PWs
  - TDM PWs have native UDP/IP mode
  - MPLS PWs can be carried over IP using RFC 4023 (with or without GRE)
- **Theorem:**
  If something is allowed by RFCs, then someone is going to do it
- **Corollary:**
  Someone is going to place PW traffic
  alongside and competing with TCP traffic
- **Conclusion:**
  In ***those*** cases, PWs MUST behave in a fashion
  that does not cause damage to congestion-responsive flows (RFC2914)
- **Felony:**
  PW traffic may not be inherently congestion-responsive
  and PWE3 has not defined any congestion mechanisms

What has been suggested

Several solutions have been offered:

- PWs should never be carried over IP
- All PW traffic must be carried over TCP
- All PW traffic must be carried over DCCP
- PWE3 must design its own
  TCP-friendly congestion response mechanism

Note, we adopt TCP friendliness (RFC 5348)
as a *safe operational envelope*
for the purposes of numerical analysis

In future work we may treat other conditions

What this draft says …

Careful analysis shows that this problem
may be much less serious than commonly imagined

We note that there are two distinct cases:

1) *elastic* PWs carrying congestion responsive traffic
   e.g., Ethernet PWs carrying mostly TCP traffic
2) *inelastic* PWs that can not respond to congestion
   e.g., TDM PWs (structure-agnostic or structure-aware)

We discover that

1) elastic PWs are automatically TCP-friendly
   and do not require any additional mechanisms
2) inelastic PWs are often TCP-friendly
   and usually do not require any additional mechanisms

Elastic PWs

Analyzed case:

Ethernet PWs carrying TCP traffic

in parallel with

TCP/IP packets

It has been proposed to encapsulate PW packets in TCP/IP to ensure that the PW does not endanger the TCP flows

However :

- there is 1 PW packet per 1 TCP/IP packet
- a single dropped packet causes the same back off to the TCP
- TCP flow is not rewarded or penalized for being inside PW

PW (as an aggregate of N flow) backs off much less (in percentage) than a single TCP flow

Inelastic PWs

Analyzed cases:

- E1, T1, E3, T3 TDM services
- SAToP or structure-aware encapsulations

Main idea

- TDM should have relatively low delay (N ms)
- SAToP service is valid for very low packet loss (0.5%?)
- structure-aware transport valid for higher packet loss (2%)

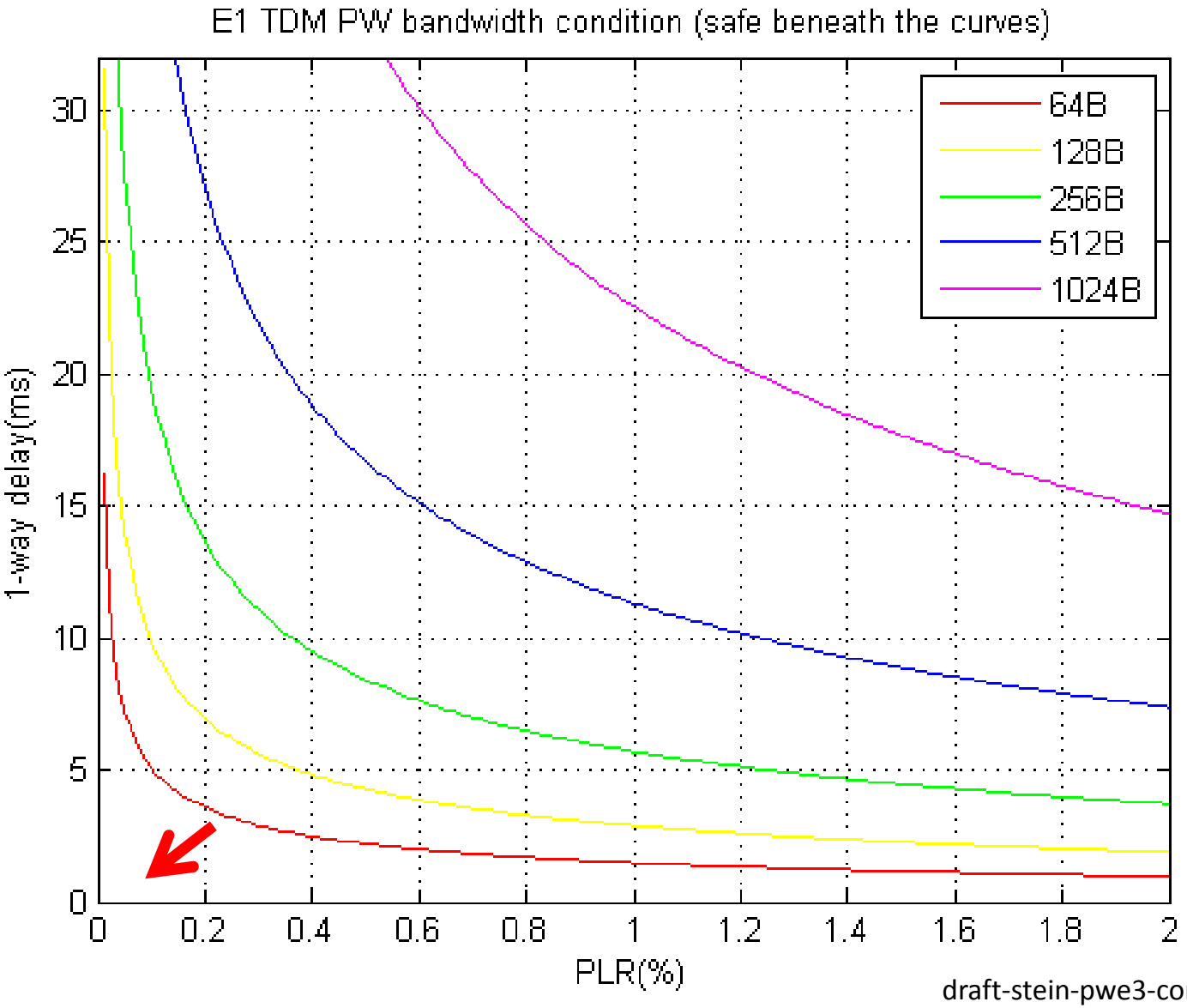
We can compare constant BW of TDM PW
with TCP's BW under the same delay and packet loss conditions

If TDM PW consumes same or less BW then it is "friendly"

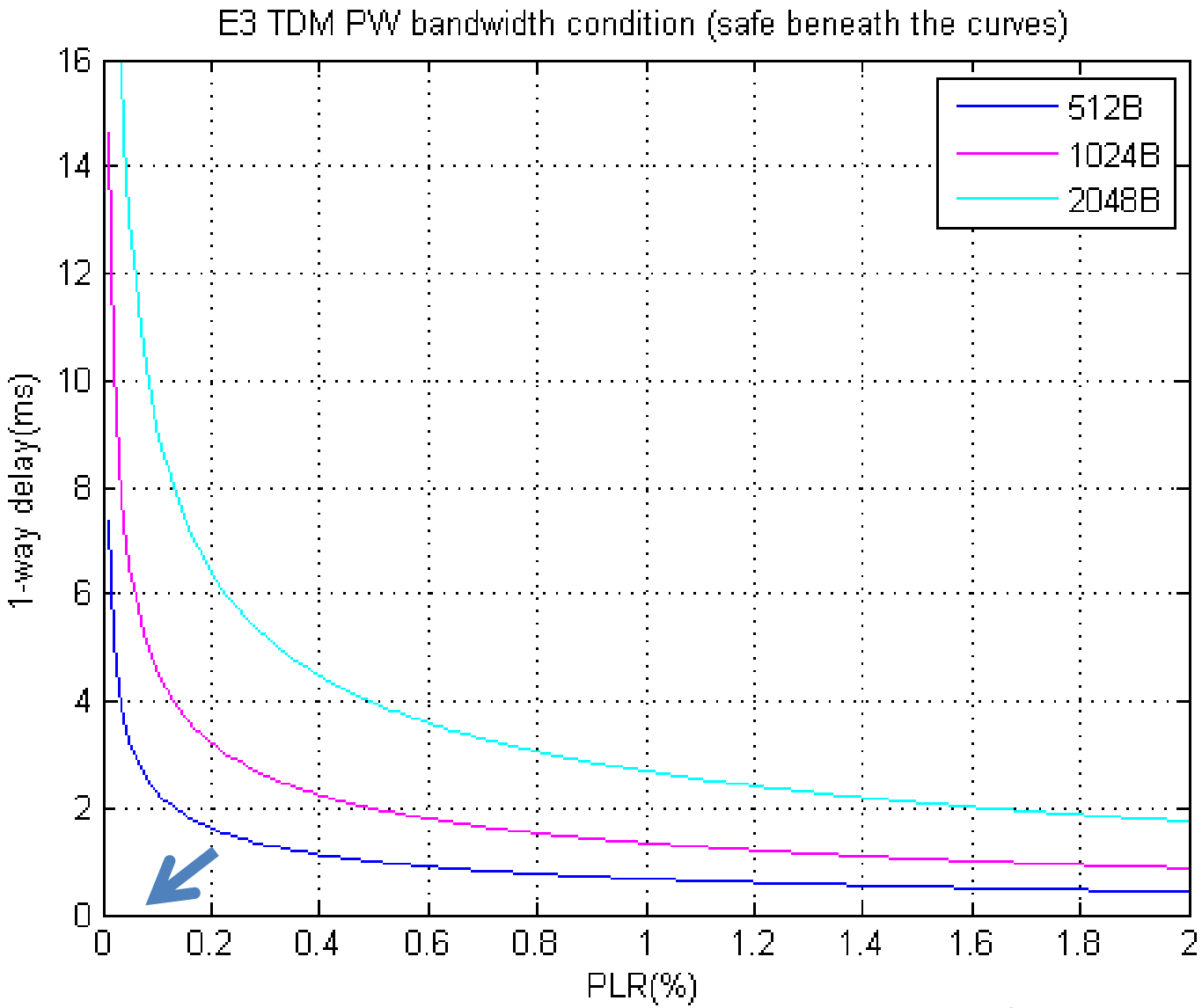
See figures (from pdf version of draft) for when this condition is obeyed

When condition is not obeyed, PW may cause congestion

E1 / TCP compatibility regions



E3 / TCP compatibility regions



Next steps

Explore more TDM cases

Tighten up the argument for inelastic PWs

- what happens when compete with short-lived TCP flows ?
- treat dynamic cases
- how much time to wait until shut-down ?
- give specific recommendations

Get more feedback from congestion-control community

Request that PWE3 accept this as a work item towards an *informational* RFC