# “Local/Remote” Forwarding Decision in Switched Data Link Subnetworks

Yakov Rekhter

Cisco Systems

yakov@cisco.com

# draft-ietf-rolc-apr-03.txt

by

Y. Rekhter, D. Kandlur, C. Villamizar

*Disclaimer: opinions expressed are not necessarily shared by the rest of the co-authors.*

# Motivations

- SVC-based subnetworks allow dynamic dedication and redistribution of resources among the communicating entities
  - Use the ability to dynamically establish and tear down an SVC
  - Provided solely by the Link Layer mechanisms
- Need to be cognizant of certain overhead associated with SVC establishment and maintenance

# Motivations (cont.)

- Want to intelligently exploit Link Layer capabilities by TCP/IP applications
  - Move beyond PVCs

# SVC management

- Some of the issues:
  - SVC establishment and tear-down
  - Class of Service specification
  - SVC sharing

# SVC management in TCP/IP environment

- TCP/IP applications exhibit a wide range of traffic characteristics
  - short-lived, small volume of data (e.g., DNS query)
  - long-lived, small volume of data (e.g., telnet)
  - short-lived, large volume of data (e.g., FTP)
  - long-lived, large volume of data (e.g., video-conferencing)

# SVC management in TCP/IP environment (cont.)

- Coupling SVC management with parameters unrelated to QoS and/or traffic requirements would result in a suboptimal use of Link Layer resources
- Coupling TCP/IP QoS and/or traffic requirements with SVC management would enable better use of Link Layer resources

# IP subnet model

- IP subnet architecture distinguishes between:
  - "local" (reachable directly over a common subnet)
  - "remote" (can be reached only via routers)

# IP subnet model in the SVC-based environment

- “Local” requires an SVC (single IP hop)
- “Remote” requires to go through a router (multiple IP hops)
  - disallows direct SVC
- “Local” vs “Remote” is based solely on IP addressing information
- “Local” vs “Remote” is totally unrelated to the QoS and/or traffic characteristics of an application

*Observation:*

*Coupling "local/remote" with addressing information is unlikely to promote efficient exploitation of capabilities provided by SVC-based subnetworks.*

# QoS/Traffic driven “Local/ Remote”

- If we decouple host’s “local/remote” decision from the addressing information, then some of the options are:
  - Use an SVC to the first-hop router
  - Use an SVC to an egress router
  - Use an SVC to the destination (if the destination is on the same Data Link subnetwork as the source)
  - Use shared SVC
    - “Best-effort” vs dedicated to a particular type of service
  - Use dedicated SVC

# Possible options (summary)

| | Dedicated SVC | Shared SVC | |
|---|---|---|---|
| | | best-effort | ToS specific |
| SVC to the first-hop router | yes | yes | yes |
| SVC to an egress router | yes | yes | yes |
| SVC to the destination (provided that the dest is on the same subnetwork as the source) | yes | yes | yes |

# QoS/Traffic driven “Local/ Remote”

- Observations:
  - QoS/Traffic driven “local/remote” is more flexible, as compared to the IP subnet model
  - Enables to take into account QoS and/or traffic requirements of an application, and/or cost factors
    - Allows to improve SVC management
    - An important step towards better SVC management

# Extending the scope of “local”

- A source may have an SVC to any destination, provided that both the source and the destination are on a common Data Link subnetwork
  - Either dedicated or shared SVC
- Completely **decoupled** from the addressing information (in contrast with the IP subnet model)
- **Coupled** with the QoS/traffic requirements of the application

# Sufficient conditions for direct connectivity

- The ability to establish an SVC to a node requires node's Link Layer address
  - knowledge of the node's Link Layer address doesn't guarantee that Link Layer connectivity could be established

# Allowing the “remote” where applicable

- When the destination is not on the same Data Link subnetwork as the source, the options are:
  - Use first-hop (default) router, or
  - Establish a shortcut to an egress router
- When the destination is on the same Data Link subnetwork as the source, but QoS and/or traffic requirements do not justify a direct VC, the source should be able to go through its first-hop (default) router
- The source should be able to make its choice **irrespective** of the IP addresses (in contrast with the IP model)
-

# *Observations:*

*We formed a "router overlay" over an SVC-based Data Link subnetwork.*

*Use direct SVCs (single IP hop path across the subnetwork) when justified by the QoS/traffic requirements of an application.*

*Use the overlay for all other applications.*

# Address assignment

- Hosts and routers on a common Data Link subnetwork are partitioned into "*groups*"
- All hosts and routers within a "*group*" are assigned addresses out of a common prefix
  - the routers within a *group* advertise routes to the prefix
- Routers within a "*group*" act as "first-hop/last-hop" routers for all the hosts within the group
- Do we need a special term for the "*group*" ?
- Does address assignment need to be in the document ?

# Could the desired behavior be accommodated within the LIS model ?

- Yes, provided that:
  - each LIS contains just a host and a router, and
  - relaxing restriction to allow inter-LIS shortcuts
- Is that a good model ?

# Using point-to-point links

- Model the relation among hosts and router within each *group* as a collection of point-to-point links
- Numbered point-to-point links
  - for each router-host pair within a *group* use two IP addresses (one for the router and one for the host)
- Unnumbered point-to-point links
  - one address per host
  - one address per router
- Does this need to be in the document ?

# Conclusions

- To explore capabilities provided by SVC-based Data Link subnetworks, we need to allow SVC management be coupled with the QoS and/or traffic requirements
  - Enable direct connectivity for applications that could benefit from it
  - Rely on strict hop-by-hop paths for other applications
- Combined switched-based infrastructure and router-based overlay, and uses each for that which is best suited