



# QoS routing over the Internet: a BGP4-based approach

Christian JACQUENET

France Telecom R & D

[christian.jacquetnet@rd.francetelecom.com](mailto:christian.jacquetnet@rd.francetelecom.com)

Dresden Nov. 21, 2001



# QoS routing over the Internet

---

- Agenda:
  - Motivation and requirements
  - Proposal
  - Simulation work and preliminary results
  - Issues
  - Ongoing work
  - Conclusion



# Motivation and requirements

---

- QoS policy enforcement is currently restricted to the scope of an AS
- QoS information needs to be exchanged between domains
  - Existing BGP4 attributes can help in providing some kind of « QoS indication »
    - *E.g.* the « PREFER\_ME » and « AVOID\_ME » global values of the COMMUNITIES attribute
  - But a finer granularity would be useful
- Allow for a smooth migration
  - Gradual deployment of QoS route computation over the Internet
- Keep the approach scalable



- Use the BGP4 protocol for conveying QoS-related information between domains to:
  - Enable QoS-based route selection processes
  - Enhance peering agreements for the deployment of value-added IP services across domains
  - Contribute to the enforcement of end-to-end QoS policies
- Introduce a new optional transitive attribute:
  - The QOS\_NLRI attribute



# The QOS\_NLRI attribute

- Advertise « QoS routes », *i.e.* routes that can be depicted with specific QoS information
  - *E.g.* “route to network N1 experiences a 100 ms one-way transit delay”
- Provide QoS information associated to the destination prefixes
  - *E.g.* “EF-marked datagrams may use this route to network N2”

```
+-----+
| QoS Information Code (1 octet) |
+-----+
| QoS Information Sub-code (1 octet) |
+-----+
| QoS Information Value (2 octets) |
+-----+
| QoS Information Origin (1 octet) |
+-----+
| Address Family Identifier (2 octets) |
+-----+
| Subsequent Address Family Identifier (1 octet) |
+-----+
| Network Address of Next Hop (4 octets) |
+-----+
| Network Layer Reachability Information (variable) |
+-----+
```

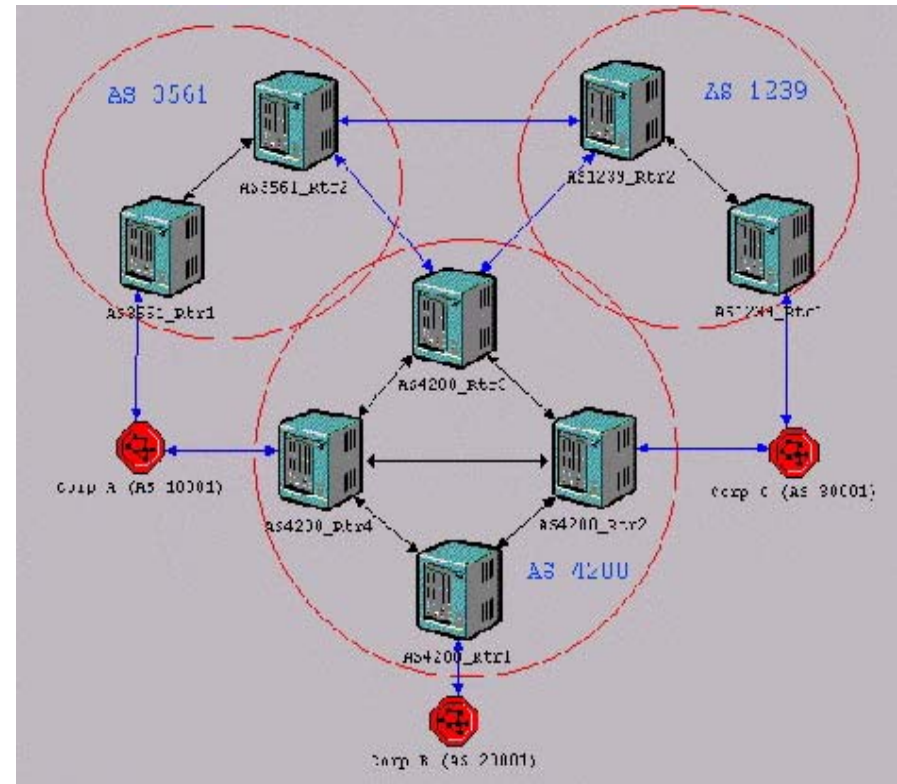
- The network model:
  - A mixed environment
    - QOS\_NLRI enabled and “classical” BGP peers
    - CPE/PE/P node taxonomy
  - Multiple domains



BGP peer

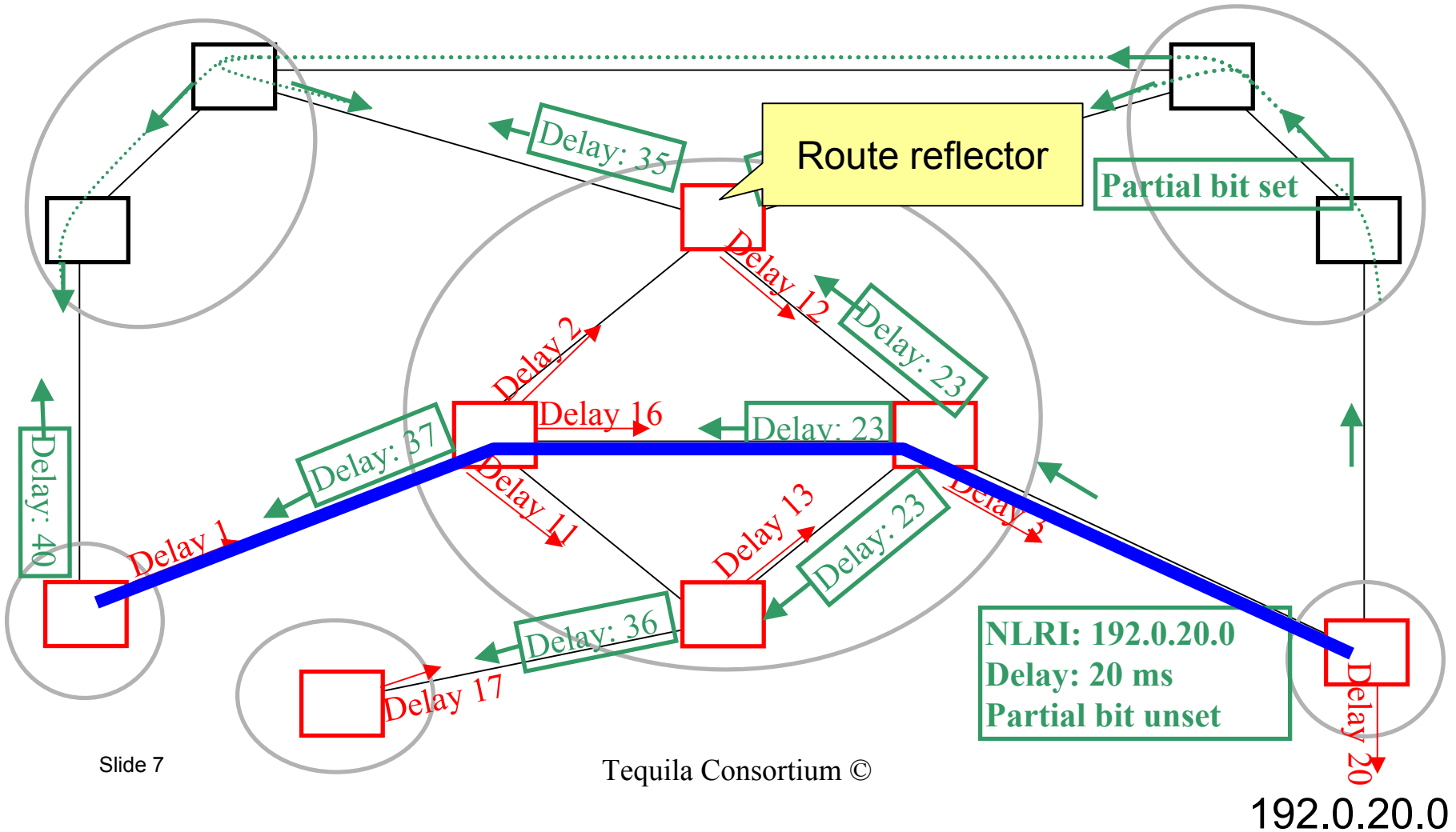


Access node



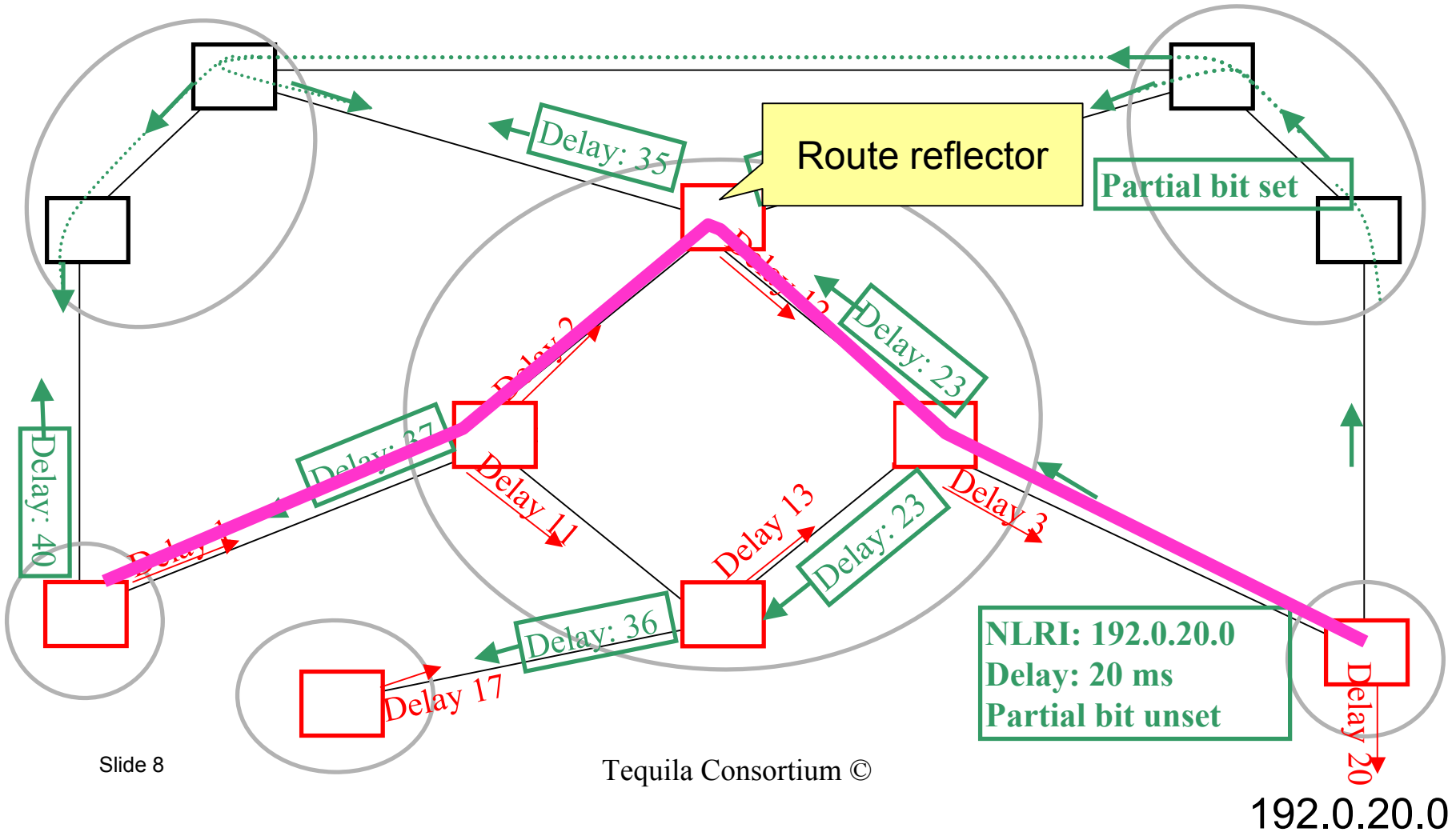


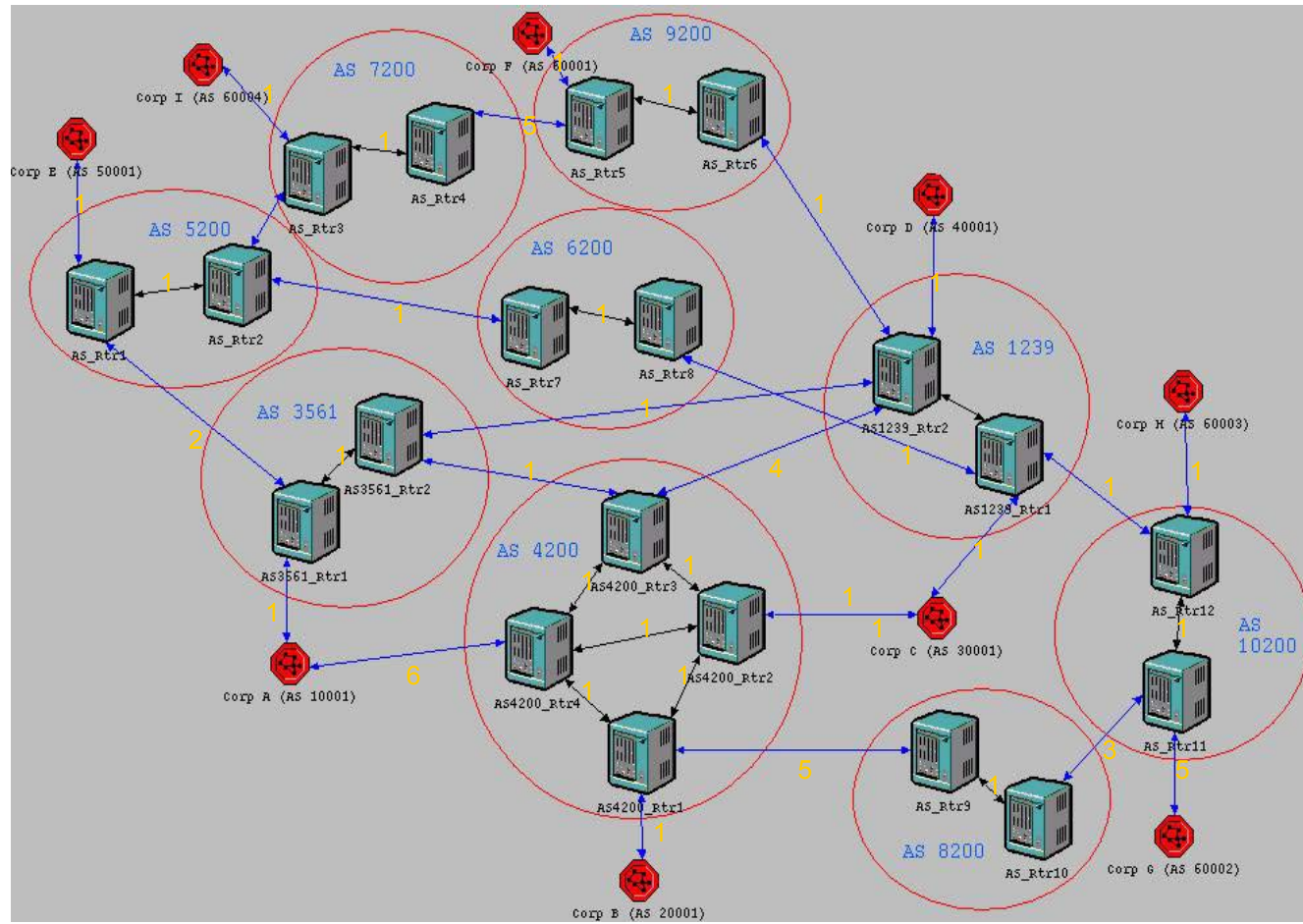
# Example of route selection





# Example of route selection







# Simulation parameters

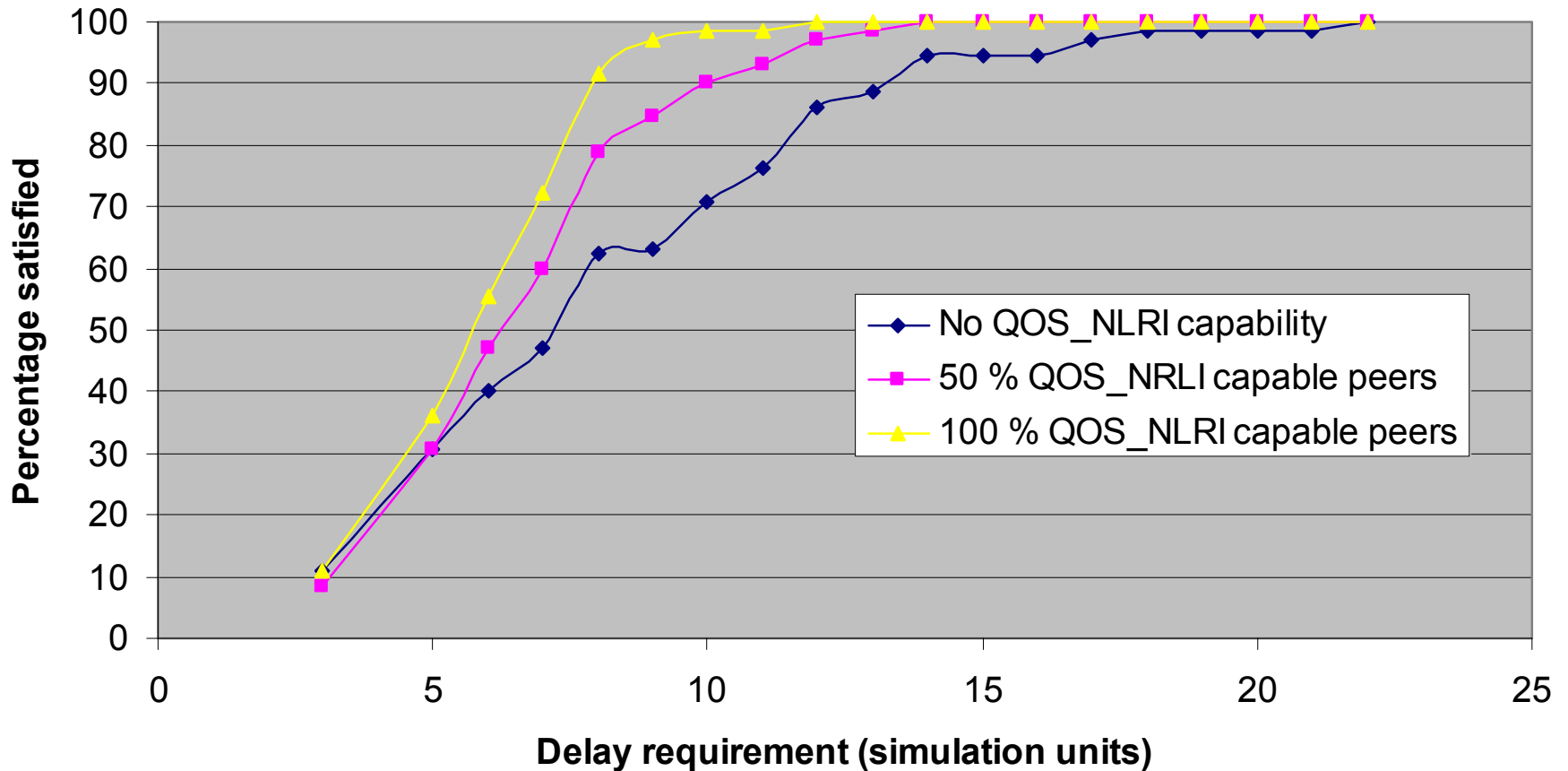
---

- Percentage of BGP speakers that are QOS\_NLRI capable
  - 0%: reference network (as a collection of autonomous systems)
  - $0% < x% < 100%$ :  $x%$  of the BGP peers are QOS\_NLRI-enabled
- Delay requirements for traffic between a source and a destination
  - Strongest (lowest-delay) requirements have less chance to be satisfied
- Transit delays on the links
  - The higher the delay on the links, the lower the percentage of serviced SLs



# Preliminary simulation results

- Satisfying delay requirements:





- Scalability:
  - How frequently should UPDATE messages be sent, according to changes of the « bandwidth conditions »?
  - Aggregation capabilities:
    - How to provide QoS indication with aggregated routes, and what would be the aggregation criteria?
- Stability:
  - PHB Id. and delay-related information should not yield flapping conditions
    - Dynamics of bandwidth information remains an issue
- Confidentiality of QoS information:
  - Already made publicly available by most ISPs
    - By means of looking glasses, for example



## Ongoing work

---

- Convey additional QoS information
- Update draft for the next IETF meeting
  - See current `draft-jacquet-qos-nlri-03.txt/pdf` on TEQUILA web site
- Ongoing prototype development
  - Based upon Zebra's code ([www.zebra.org](http://www.zebra.org))
- Additional simulation results by Q1 2002
  - Submit an applicability draft to the IETF



- On the approach:
  - NO modification of the BGP4 protocol
- On the simulation:
  - Preliminary results are encouraging
- On the remaining issues:
  - Technical feasibility has been demonstrated
  - Further simulation planned to investigate the scalability aspects