BGRP Plus: Quiet Grafting
Mechanisms for providing a scalable end-to-end QoS Solution

E. Nikolouzou, P. Sampatakos, L. Dimopoulou,
S. Salsano, I. S. Venieris

Art-QoS 2003 Workshop, Warsaw 24-25/03
Outline

- Inter-Domain Issues
- BGRP Plus Concepts
- Scalability Issues
  - Problems
  - Proposed Solution: Quiet Grafting
- Performance Studies
  - Simulation results
- Discussion on results
Inter-Domain Issues

- Inter-domain dynamic IP QoS is even more challenging
- Need for inter-domain resource reservation protocols
  - NSIS WG stresses this need
- Current Architectures
  - RSVP
  - DiffServ
- Our proposal
  - BGRP Plus (BGRP+): a solution to the scaling problem
BGRP+ Concepts

Inter-domain resource control mechanism

- Sink tree based aggregation for resource reservations over a network of DiffServ domains
- The aggregated reservations are negotiated between the BGRP+ agents
- Deployed at each BGP-capable border router of each DiffServ domain
  - Takes advantage of BGP running on the BRs aggregating the reservations according to the sink trees created by the BGP routing protocol
- Simple soft-state messages
BGRP+ Messages

- Downstream: PROBE message identifies sink tree and checks policy
- Upstream: GRAFT message makes reservation
- REFRESH: preserve the reservation state established
Scalability Issues

BGRP+ addresses scalability in terms of:

- State information ("Sink tree")
- Signalling messages (Quiet Grafting Mechanism)
Sink Trees

Objective: state information scalability

The sink tree is the logical representation of BGP routing

The root of the tree is the destination domain

Traffic is aggregated along a sink tree per destination domain

Reservations can be aggregated as well!

Autonomous Systems

Direction of data flows
Quiet Grafting

- Reduce the signalling load to improve scalability
  - Reduce the number of **PROBE** and **GRAFT** messages
  - Answer **PROBE** messages as early as possible before they reach the destination domain

- Requirements for quiet grafting:
  - identify the sink tree, to which the reservation belongs
  - have pre-reserved resources for this sink tree
  - provide means to contact the destination domain
Identification of the Sink Tree

**NLRI Labeling for Sink Tree Identification**

- Sink trees are identified by their root: the destination AS number and an identification of the border router in the destination AS (entry point)
- Bases its route selection on the shortest path
- NLRI information is propagated back from the root of the sink tree in GRAFT messages
- Is stored in each BGRP+ agent, which processes the message
Signalling in the Last Domain

- No resource reservation is performed in the last domain
  - resource reservations are carried out or pre-reserved resources are used up to the ingress border router of the destination domain

- Solution: direct communication between the initiating and the last domain
  - each domain should provide a standardized interface to its intra-domain resource control entity to reserve resources within the last domain
  - enhancing the information carried by the GRAFT message to back-propagate a reference to this interface as well as the IP address of the ingress border router of the last domain
Resource Cushions

Need for resource cushions
- PROBE message can be answered with a GRAFT message only, when pre-reserved resources are available on the sink tree of the reservation

Resource cushions
- Resources, which are reserved downstream but not upstream of a BGRPP agent are called resource cushions
- BGRPP agent uses two parameters to control the size of its resource cushions
  - RBS: Release block size
  - RP: Retain period
Return RBS of resources, if it was unused during the complete RP

- RBS and RP are configurable and determines performance of algorithm
Simulations

**Earliest point for quiet grafting**
- How far has a PROBE message to travel, until we can identify the corresponding sink tree?

**Effectiveness of resource cushions**
- Are resource cushions available where needed?
- What is the overhead of unnecessarily reserved resources?
Earliest Point for Quiet Grafting

Some definitions
- Sink tree: exists independently of actual reservations
- Populated node: node with an actual resource reservation for that sink tree

Basic considerations
- In a sparsely populated sink tree, the average distance from a node to the first populated node is rather high
- In a densely populated sink tree, the average distance from a node to the first populated node is rather low
- As the population of the tree increases, the average distance to the first possible point for quiet grafting decreases
Path Length Reduction

- Mean Path Length vs. Populated Nodes (%)
- Path length reduction (%) vs. Depth of Sink Tree

Poland, Warsaw 24-25/03  Art-QoS 2003
Effectiveness of Resource Cushions

- **Build resource cushions through delayed resource release**
  - Released resources are not immediately forwarded towards the sink tree root, but used to build a resource cushion

- **Release unused resources**
  - Resource cushions are released step-wise (Retain Period) and block-wise (Return Block Size)
  - When more than RBS resources are unused during a RP, then this block is released
Performance Study

- Sink tree with depth 4
- Resource Requests injected from the leaves of the sink tree
  - Traffic generators with exponential distributed inter-arrival time/holding time - request has fixed size: 1 unit
- Two different scenarios:
  - Variation of the Release Block Size
  - Variation of Retain Period
- Two load conditions:
  - Mean offered load = 20 unit
  - Mean offered load = 100 unit
- Results
  - Utilization of the overall resources
  - Signaling load
- Trade-off
  - Higher RP yields lower Utilization but lower Signaling load
- Load dependent
  - Higher load gives better results

Two load conditions:
- Mean offered load = 20 unit
- Mean offered load = 100 unit

Results
- Utilization of the overall resources
- Signaling load
- Trade-off
- Higher RP yields lower Utilization but lower Signaling load
- Load dependent
  - Higher load gives better results
Results for different RP (RBS=1 unit)

- **Signaling Overhead (%):**
  - Line 1: 100u
  - Line 2: 20u

- **Utilization (%):**
  - Line 1: 100u
  - Line 2: 20u

---

Poland, Warsaw 24-25/03
Art-QoS 2003
Results for different RPS (RP=10sec)

![Graph showing signaling overhead and utilization for different block sizes and signaling times]

- Signaling Overhead (%)
  - Block Size vs. Signaling Overhead for 20u and 100u signaling times.

- Utilization (%)
  - Block Size vs. Utilization for 20u and 100u signaling times.
Discussion on Performance Results

- **Proof of concept**
  - Delayed resource release can dramatically reduce the signalling load
  - Reasonable mean size of resource cushions (10% - 30%)
- **Especially effective in lively branches of the sink tree**
  - Best results when number of incoming requests is high
- **Optimal parameter sets will vary at different places in the sink tree**
  - Short RP and small RBS near the leafs
  - Longer RP and larger RBS near the sink