

# **ATM ABR Traffic Management: Our Recent Research Results**

**Raj Jain**

Department of CIS

The Ohio State University

**Contact:** [jain@cis.ohio-state.edu](mailto:jain@cis.ohio-state.edu)

<http://www.cis.ohio-state.edu/~jain/>

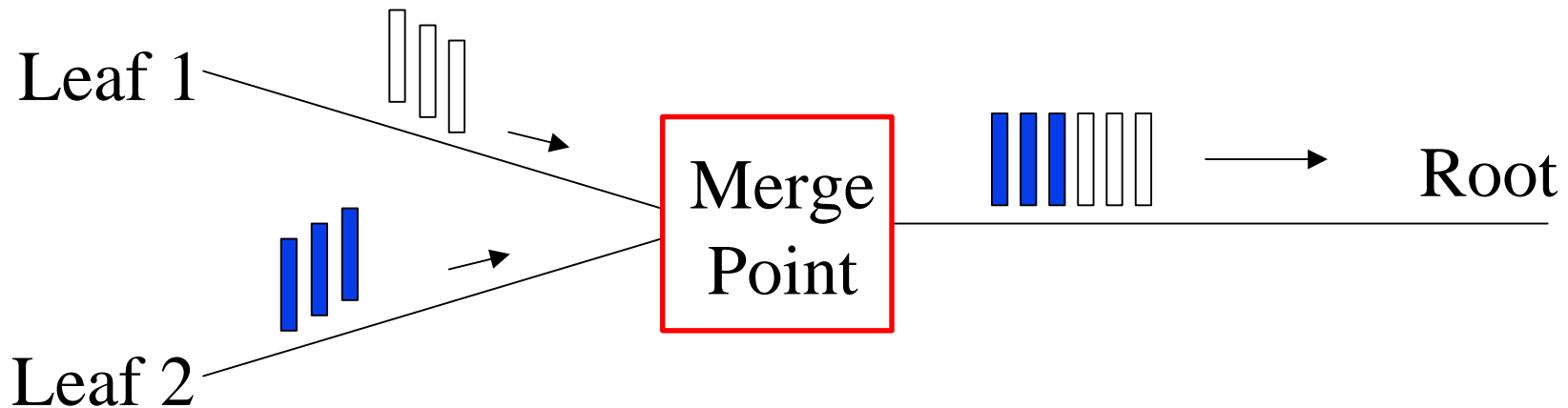


□ Recent Achievements:

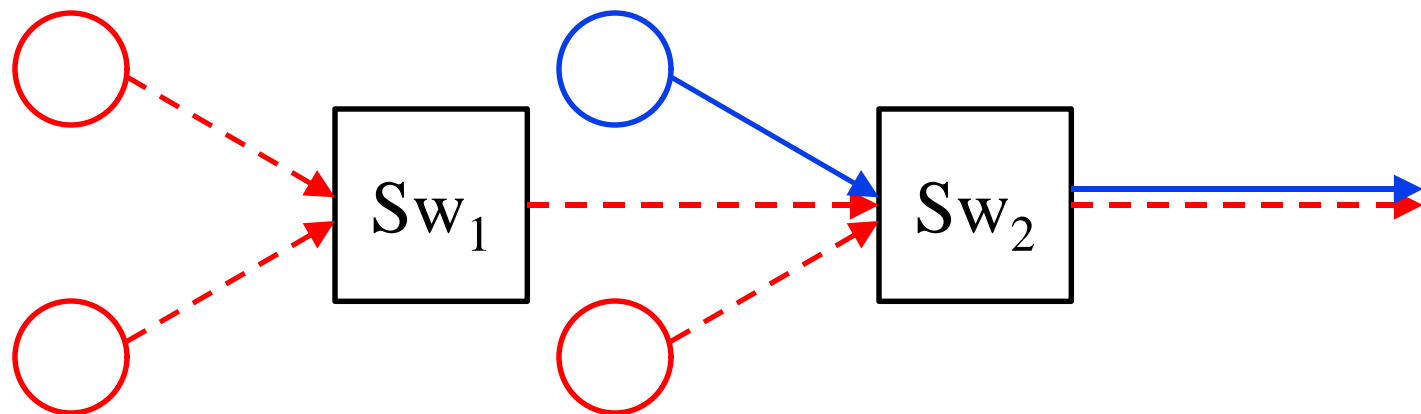
1. Multipoint Communication
2. Virtual Source/Virtual Destination
3. Real-Time ABR: Generalized Fairness

# 1. Multipoint-to-Point VCs

- ❑ A multipoint-to-point VC can have more than one concurrent sender
- ❑ Traffic at root =  $\Sigma$  Traffic originating from leaves



# Sources, VCs, and Flows



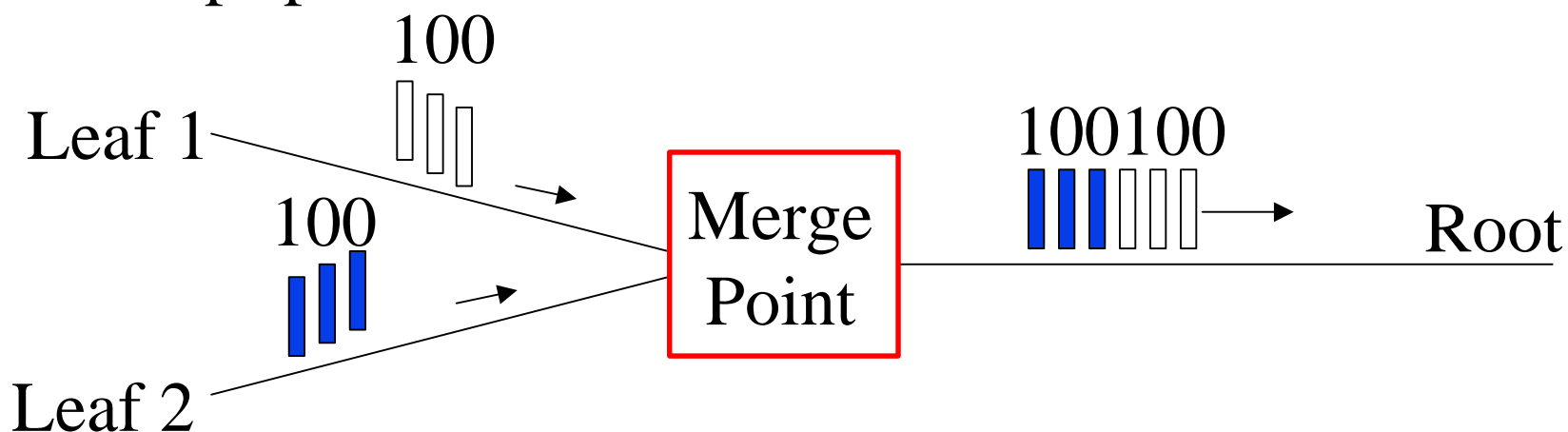
- Sw<sub>2</sub> has to deal with
  - Two VCs: Red and Blue
  - Four sources: Three red sources and one blue source
  - Three flows: Two red flows and one blue

# Fairness Definitions

- ❑ Source-based: N-to-one connection = N one-to-one connections  $\Rightarrow$  Use max-min fairness among sources
- ❑ VC/Source-based:
  1. Allocate bandwidth fairly among VCs
  2. For each VC, allocate fairly among its sources
- ❑ Flow-based: Flow = VC coming on an input link. Switch can easily distinguish flows.
- ❑ VC/Flow-based:
  1. Allocate bandwidth fairly among VCs
  2. For each VC, allocate fairly among its flows

# Mpt-pt Issues

- ❑ Cells of senders in the same multipoint-to-point VC cannot be distinguished
- ❑ Question: Can we achieve source-based fairness?  
Answer: Yes!
- ❑ We extended ERICA to achieve source based fairness for mpt-pt VCs



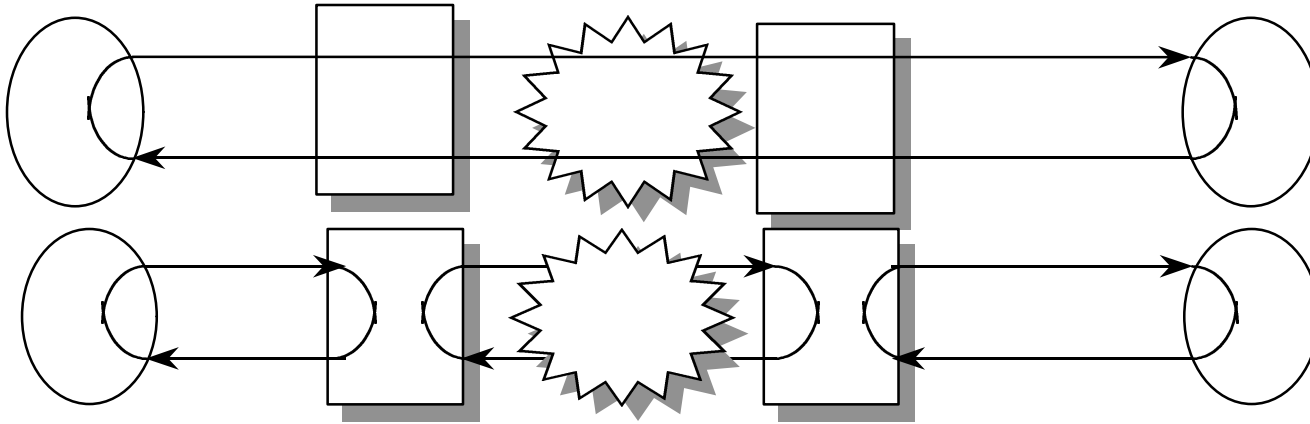
# Lessons Learnt

- ❑ Avoid determining the effective number of active sources
- ❑ Avoid estimation of rates of sources, or determining if a source is bottlenecked at this link
- ❑ Use only aggregate measurements
- ❑ Do not use CCR values from BRM cells
- ❑ CCR from FRM cells can be used
- ❑ Using the maximum CCR in an interval, and exponentially averaging the maximum ER in the previous interval can improve performance
- ❑ Do not turn around RM cells at merging point

# Impact I

- ❑ A new item was added to the TM living list describing the issues in Mpt-to-pt ABR
- ❑ A sample merge point algorithm, which applies to mpt-to-mpt also, was added to the living list
- ❑ Our fairness definitions were moved to this same living list item

## 2. Virtual Source / Virtual Destination (VS / VD)

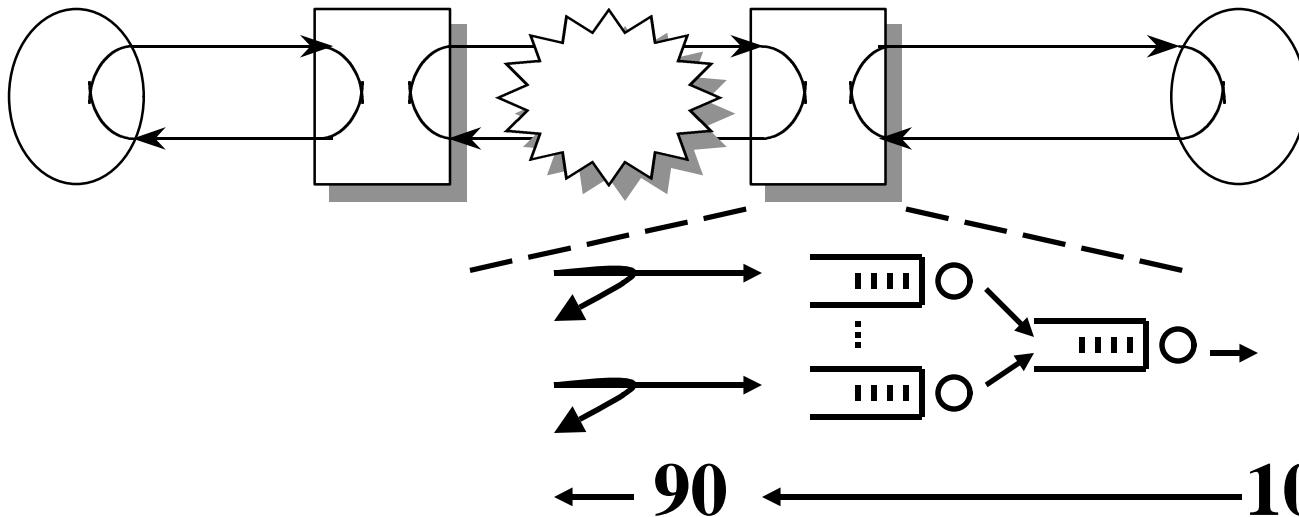


- ❑ Segments the end-to-end ABR control loop.
- ❑ Coupling between loops is implementation specific.
- ❑ ABR switches separated by non-ATM network could also implement VS/VD.

# VS/VD Issues

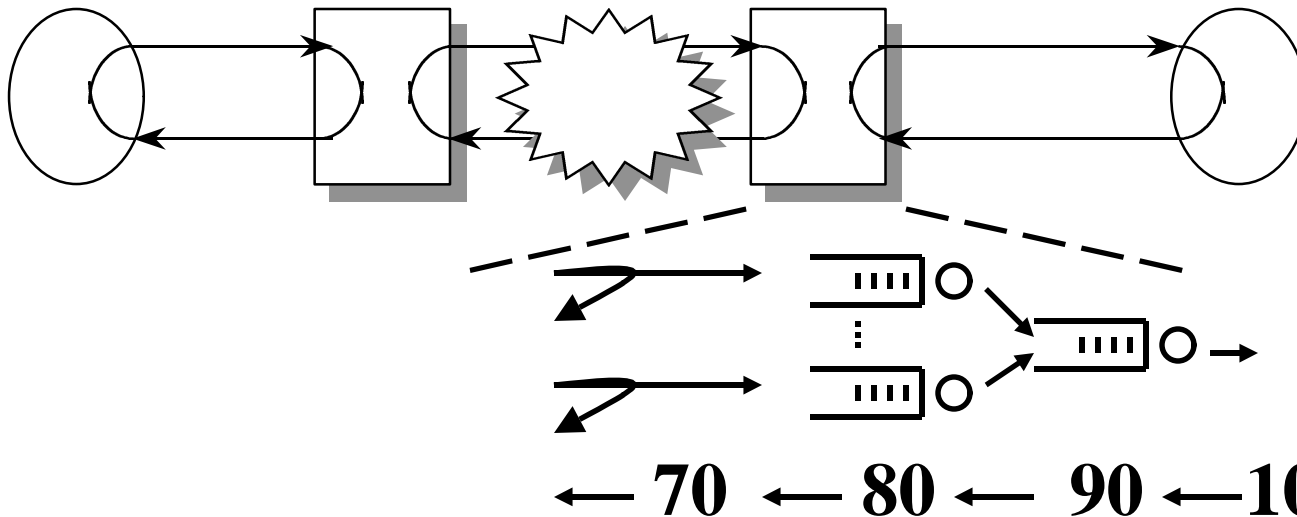
- ❑ Although TM4.0 allows VS/VD, it does not describe how the feedback must be passed from VS to previous VD.
- ❑ It is not clear if and when VS/VD help.
- ❑ Our Accomplishments:
  - Analyzed issues in designing rate allocation schemes for VS/VD switches.
  - Developed a per-VC rate allocation scheme for VS/VD.
  - Showed how VS/VD can help in buffer management across the network.

# Simple VS/VD Model



- ❑ Internal Service Rate =  $f(\text{Downstream Feedback, Switch algorithm using class } Q + \text{per-VC } Q)$
- ❑ Upstream Feedback = Internal Service Rate
- ❑ Example: Downstream = 100, Service = 90, Upstream feedback = 90 Mbps

# Correct VS/VD Model

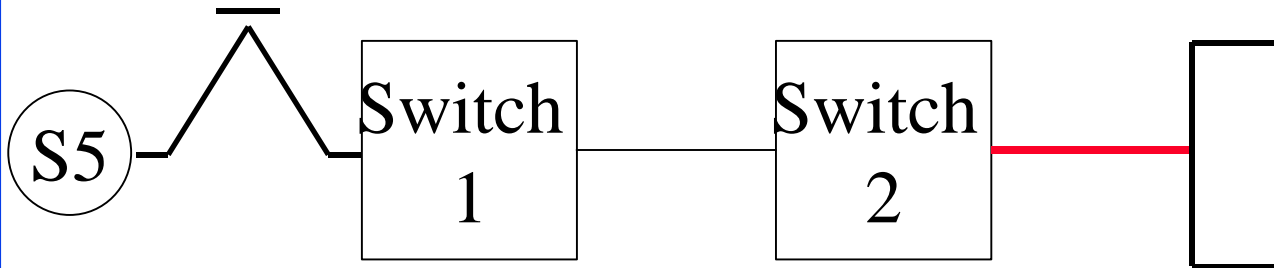


- ❑ Internal Service Rate =  $f(\text{Downstream Feedback, Switch algorithm using class } Q)$
- ❑  $ACR_{ij} = f(\text{Internal service rate, end system rules})$
- ❑ Upstream feedback =  $f(VC \ Q)ACR_{ij}$
- ❑ Example: Downstream = 100, Service = 90,  $ACR=80$ , Upstream feedback=70 Mbps

# Simulation Results

- Without VS/VD:
  - Single control loop for the entire connection.
  - All queues are in the bottleneck switch.
  - Buffer requirements for terrestrial switch are proportional to satellite propagation delay.
- With VS/VD:
  - Control loop broken at each switch.
  - Queues remain at the switch between the satellite and the terrestrial loop (satellite switch).
  - Terrestrial switch only requires small buffers.

## Results (Cont)



- ❑ With correct implementation of VS/VD:  
Maximum queue at each switch  
 $\leq$  A small multiple of bandwidth delay product of the previous loop  
 $\Rightarrow$  Can help isolate long-delay hops from short-delay hops.
- ❑ Workgroup switches on satellite paths will not need buffering proportional to round-trip even if they are the bottleneck.

# Impact II

- Our extension of the ERICA switch algorithm including VS/VD was accepted for inclusion as a sample algorithm in the TM5.0 baseline text.

# 3. Real-Time ABR: Generalized Fairness

- ❑ Real-time applications need non-zero Minimum Cell Rate (MCR)
- ❑ In TM4, Distribution of excess bandwidth (fairness) is implementation specific.
- ❑ TM4.0 has five examples of fair distribution
- ❑ We have shown that two of the examples are not meaningful and have proposed a sixth example that is a superset of the remaining three definitions
- ❑ We developed a switch algorithm that implements the proposed definition

# TM4.0 Definitions

1.  $B(i) = B/n$

2.  $B(i) = \text{MCR}(i) + (B-M)/n$

3.  $B(i) = \text{Max}\{\text{MCR}(i), \text{Max-Min Share}\}$

4.  $B(i) = B * (\text{MCR}(i)/M)$

5.  $B(i) = w(i) * B / \text{Sum}(w(j))$

❑ Definition 5 does not always guarantee MCR

❑ Definition 3 may result in total of fair share being more than the capacity

❑ Notation:  $n = \#$  of active VCs bottlenecked here

$B =$  Bandwidth available for the bottlenecked VCs

$$M = \sum \text{MCR}(I)$$

# General Definition

- Fair Share

$$B(i) = MCR(i) + \frac{w(i) (B - M)}{\sum_{j=1,n} w(j)}$$

- This definition is a superset of 1, 2, 4 in TM4.0
- Always ensures MCR

# Mapping to TM 4.0

- $w(i) = w, \text{MCR}(i)=0: B(i) = B/n$   
This is Definition 1 (Max-min Fair).
- $w(i) = w: B(i) = \text{MCR}(i) + (B - M)/n$   
This is Definition 2 (MCR plus equal share)
- $w(i) = \text{MCR}(i):$   
$$B(i) = \text{MCR}(i) + (B-M) \text{MCR}(i) / M$$
$$= B^* (\text{MCR}(i)/M)$$
  
This is Definition 4 (Proportional to MCR)

# Pricing Function

- $T$  = Small time interval,  $W$  = Number of bits  
 $R$  = Average rate  $W/T$
- Cost  $C = f(W, R)$ . If  $C$  is restricted to continuous differentiable functions:  $C = \sum_{ij} a_{ij} W^i R^j$
- For all values of  $W$  and  $R$ :
  - $C \geq 0$     $\partial C / \partial W \geq 0$     $\partial C / \partial R \geq 0$
  - $\partial(C/W) / \partial W \leq 0$  [Economy of Scale]
  - $\partial(C/R) / \partial R \leq 0$  [Economy of Scale]
- The only function that satisfies all 5 conditions is:

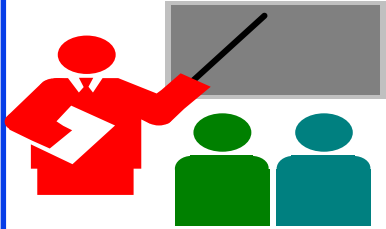
$$C = a_{00} + a_{10}W + a_{01}R + a_{11}WR$$

# A Simple Pricing Fn

- $f()$  is non-decreasing w.r.t to  $W$   
 $f()$  is non-increasing w.r.t to  $T \Rightarrow$  non-decreasing w  $R$
- A simple function satisfying these requirements is:  
$$C = c + w W + r R$$
  
Here,       $c =$  Fixed cost per connection  
               $w =$  Cost per bit (How much)  
               $r =$  Cost per Mbps (How fast)
- This cost function implies that the excess bandwidth should be allocated using the proposed generalized fairness function

# Impact III

The proposed definition was added to TM5.0 baseline text being developed now.



# Summary

1. Multipoint Communication: Source-based fairness can be achieved even though sources can not be distinguished in an mpt-pt VC
2. Virtual Source/Virtual Destination: The buffering required at each VS/VD switch is proportional to the bandwidth-delay product of the next loop
3. Real-Time ABR: Generalized Fairness based on charging policies is a superset of current TM4 policies.
4. Extensions of ERICA to cover the above three cases have been developed

# Our Recent ATM Forum Contributions

- All our contributions and papers are available on-line at <http://www.cis.ohio-state.edu/~jain/>  
See **Recent Hot Papers** for tutorials.
- S. Fahmy et al, "Proposed appendix on sample ABR point-to-multipoint algorithms," ATM Forum/98-0293, April 1998, <http://www.cis.ohio-state.edu/~jain/atmf/a98-0293.htm>

# Contributions (Cont)

- ❑ S. Fahmy et al, "Proposed modifications to the baseline text and living list on multipoint ABR behavior," ATM Forum/98-0150, February 1998, <http://www.cis.ohio-state.edu/~jain/atmf/a98-0150.htm>
- ❑ B. Vandalore et al, "Generalized Fairness support in Switch Algorithms" ATM Forum/98-0151, February 1998, <http://www.cis.ohio-state.edu/~jain/atmf/a98-0151.htm>

## Contributions (Cont)

- ❑ S. Fahmy et al, "Determining the number of active ABR sources in switch algorithms," ATM Forum/98-0154, February 1998, <http://www.cis.ohio-state.edu/~jain/atmf/a98-0154.htm>
- ❑ R. Goyal et al, "Per-VC Rate Allocation Techniques for ABR Feedback in VS/VD Networks" ATM Forum/97-1086r1, February 1998, <http://www.cis.ohio-state.edu/~jain/atmf/a97-1086r1.htm>

# Contributions (Cont)

- ❑ B. Vandalore et al, "Design and Analysis of Queue Control Function for Switch Schemes," ATM Forum/97-1087, February 1998, <http://www.cis.ohio-state.edu/~jain/atmf/a97-1087r1.htm>
- ❑ S. Fahmy et al, "A switch algorithm for ABR multipoint-to-point connections," ATM Forum/97-1085, December 1997, <http://www.cis.ohio-state.edu/~jain/atmf/a97-1085.htm>

# Our Recent Papers

- ❑ R. Goyal et al, "Per-VC Rate Allocation Techniques for ATM-ABR Virtual Source Virtual Destination Networks," Submitted to Globecom'98, February 1998, <http://www.cis.ohio-state.edu/~jain/papers/globecom98.htm>
- ❑ R. Goyal et al, "Design Issues for providing Minimum Rate Guarantees to the ATM Unspecified Bit Rate Service", Proceedings of ATM'98, May 1998, <http://www.cis.ohio-state.edu/~jain/papers/atm98.htm>

# Papers (Cont)

- S. Fahmy et al, "On Determining the Fair Bandwidth Share for ABR Connections in ATM Networks," Proceedings of the IEEE International Conference on Communications (ICC) 1998, June 1998, <http://www.cis.ohio-state.edu/~jain/papers/neff.htm>

# Thank You!



This work was sponsored in part by  
Rome Laboratory/C3BC under  
Contract #F30602-96-C-0156