The ATLANTA™ Chipset:

A Low-Cost Scalable VLSI Solution for High-Performance ATM Switching Systems

Fabio M. Chiussi, Joseph G. Kneuer and Vijay P. Kumar

High-Speed Networks Research Department
Bell Laboratories, Lucent Technologies
Holmdel, NJ 07733, USA

The ATLANTA™ Architecture

- Nonblocking
- Modularly scalable from 622 Mbps to 25 Mbps using commonly available implementation technology
- High buffer utilization while using distributed buffers
- Low complexity
- Provides path for future growth in features

The ATLANTA™ Architecture: Basics

Chiussi, Kneuer, Kumar

- Ingress and egress buffers
- Selective backpressure
  - No Head-of-Line (HOL) blocking
  - Sharing of distributed buffers
    - Buffer utilization comparable with centralized shared-memory switch
  - Majority of buffers in ingress and egress port cards
    - low-cost off-the-shelf RAM's, regardless of switching capacity
- Fabric
  - Up to 5 Gbps: standalone switching module
  - From 5 Gbps to 25 Gbps: *Memory/Space/Memory* Configuration

Lucent Technologies
Bell Labs Innovations

The ATLANTA™ Chipset: Essential Features

Chiussi, Kneuer, Kumar

- Complete set of building blocks for implementing ATLANTA™ switches from 622 Mbps to 25 Gbps
- 4 chips (2 for fabric, 2 for port cards)
- 622-Mbps input/output port rates (plus local header overhead)
- Up to 64K cell buffers per port
- Up to 64K VC's per port
- 4 Quality-of-Service (QoS) service classes per port
- Optimal support of multicast traffic
- Physical interface supports UTOPIA II multiplexing standard. Maximum 30 subports per port
- Support of redundant fabric configurations
- Port devices: full-duplex ingress and egress functionality

Lucent Technologies
Bell Labs Innovations

The ATLANTA™ Chipset: Fabric Devices

Chiussi, Kneuer, Kumar

- ATM Switch Module (ASX)
  - 5 Gps switching module
  - Operates either as 8x8 standalone switch, or as first stage, or as third stage in MSM configuration
  - Small integrated buffer (512 cells)
- ATM Crossbar Element (ACE)
  - 8x8 bufferless crossbar for second stage in MSM configuration
  - Implements one, two, or four crossbars, depending on fabric size

Lucent Technologies
Bell Labs Innovations

The ATLANTA™ Chipset: Port Devices

Chiussi, Kneuer, Kumar

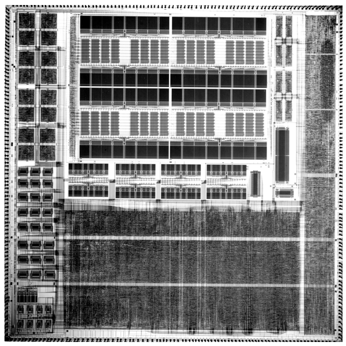
- ATM Layer Manager(ALM)
  - Input and output header processing, ATM-Forum-compliant policing, statistics collection, local header generation
- ATM BufferManager
  - Manages ingress and egress port buffers (up to 64K cells in shared buffer pool in external RAM)
  - Multicasting to subports within a port
  - Rate matching to subports
    - rate matching to subports with variable rate (*e.g.,* SAR's) also supported
  - Dual switch interfaces and monitoring logic for fault-tolerant redundant fabric configurations
  - Complete set of hooks for Available Bit Rate (ABR) and Operation, Administration & Maintenance (OA&M) support via external engine.
  - Can operate without the fabric as a low-cost 622-Mbps switching module

Lucent Technologies
Bell Labs Innovations

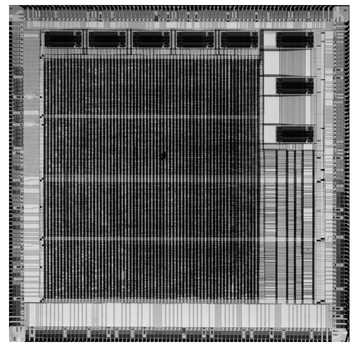
The ATLANTA™ Chipset: Pictures

Chiussi, Kneuer, Kumar

Fabric Devices



ASX



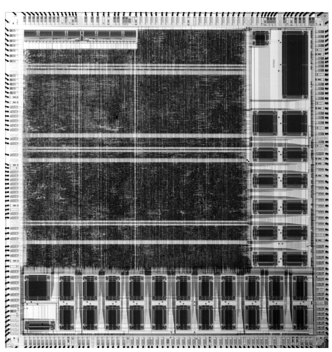
ACE

Lucent Technologies
Bell Labs Innovations

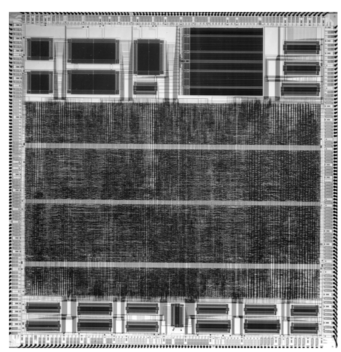
The ATLANTA™ Chipset: Pictures (cont.)

Chiussi, Kneuer, Kumar

Port Devices



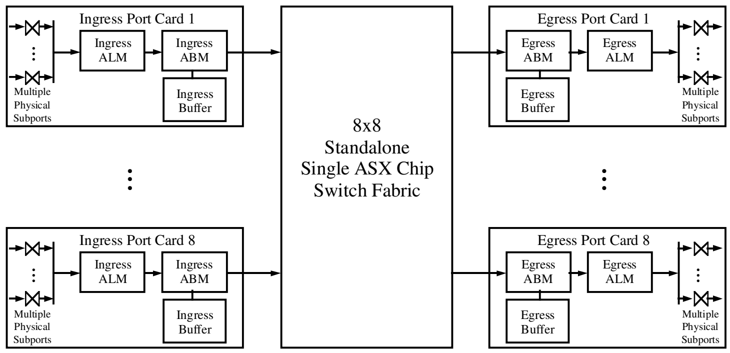
ALM



ABM

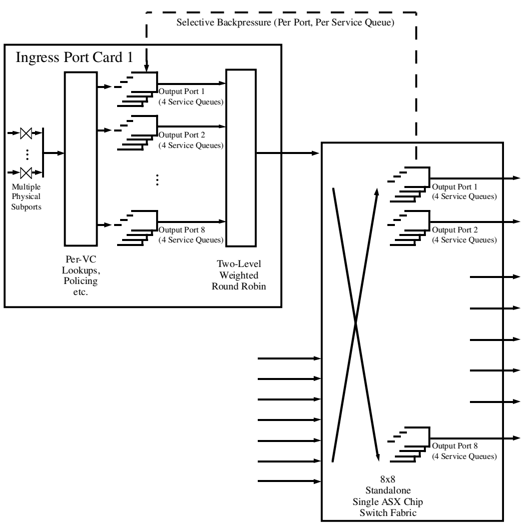
Lucent Technologies
Bell Labs Innovations

ATLANTA™ Architecture: Single-Stage Switch Fabric



(In reality, ingress and egress port devices are full duplex)

ATLANTA™ Architecture: Cell Flows



- One queueing group per output port
- Each queueing group: 4 service priorities, separate FIFO's for each priority
  - Round robin among groups
  - Weighted round robin (work-conserving) among priorities within group
- ABM: exact replica of queueing structure in ASX
- **Selective backpressure** between ASX and ABM

ATLANTA™ Architecture: Buffers

Chiussi, Kneuer, Kumar

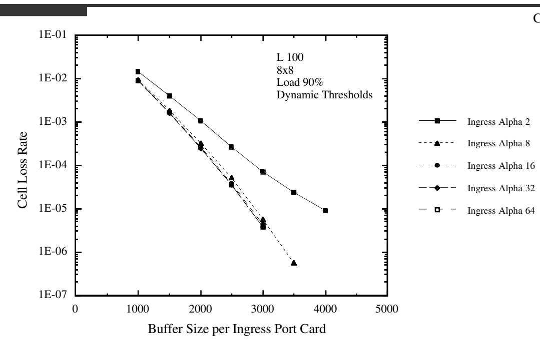
- ASX: Common 512-cell on-chip buffer pool
- ABM cell buffer (up to 64K cells) in external low-speed static RAM
- Buffer management:
  - each queue controlled by dynamic thresholds

    $T_j(t) = \max(\alpha_j (B_j - Q(t)), 0)$
  - ABM: dynamic threshold triggers discarding
  - ASX: dynamic threshold triggers selective backpressure
- ABM on egress: backpressure to the fabric (optional)

Lucent Technologies
Bell Labs Innovations

ATLANTA™ Stand-Alone Fabric: Performance

Chiussi, Kneuer, Kumar

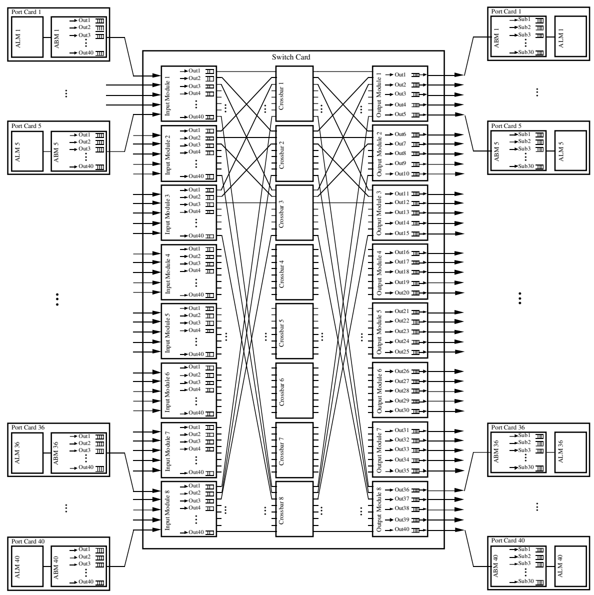


- *Example*: bursty traffic, geometrically distributed bursts, average burst length $L = 100$ cell, 90% load, single service class
- Selective backpressure
  - Backpressure applied *selectively* only to the traffic flows destined to congested port(s)
  - No HOL blocking
  - Buffers in ingress cards act as an extension of output buffers in the fabric
  - Buffer requirements less than a factor of two larger than centralized shared-memory switch
- Buffer requirements depend on α

Lucent Technologies
Bell Labs Innovations

ATLANTA™ Architecture: Multi-Stage Switch Fabric

Chiussi, Kneuer, Kumar



- **Memory/Space/ Memory** Configuration
- Nonblocking multi-stage routing network
- Expansion; multiple paths from each input to each output module
- All interconnection lines between stages have same rate as input and output ports

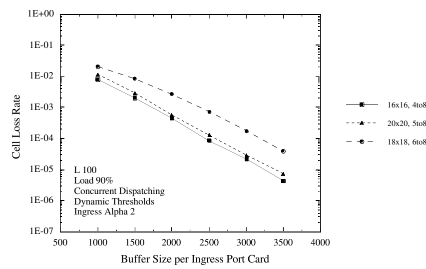
Memory/Space/Memory Configuration

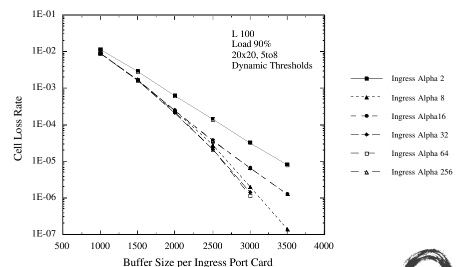
Chiussi, Kneuer, Kumar

- Achieves non-blocking behavior with low complexity
  - Buffers in the input and output stages
  - Bufferless second stage
  - Simple, efficient, fully distributed self-routing algorithm
    - *Concurrent Dispatching Algorithm*
    - independently run by each input module in the fabric
    - Very simple to implement.
- Small buffers in the fabric
  - Selective backpressure from the output to the input buffers in the fabric
- Cell sequence is preserved
  - Bufferless center stage, cells can be routed individually without affecting cell sequence

ATLANTA™ Multi-Stage Fabric: Performance

- Expansion factor as small as 6:8 sufficient to achieve nonblocking behavior
- ASX device supports 4:8, 5:8, and 6:8 expansion factors
  - buffer requirements in the ingress depend on expansion factor.
- *Example*: bursty traffic, geometrically distributed bursts, average burst length $L = 100$ cell, 90% load, single service class, dynamic thresholding parameter $\alpha = 2$
- Buffer requirements decrease as α increases



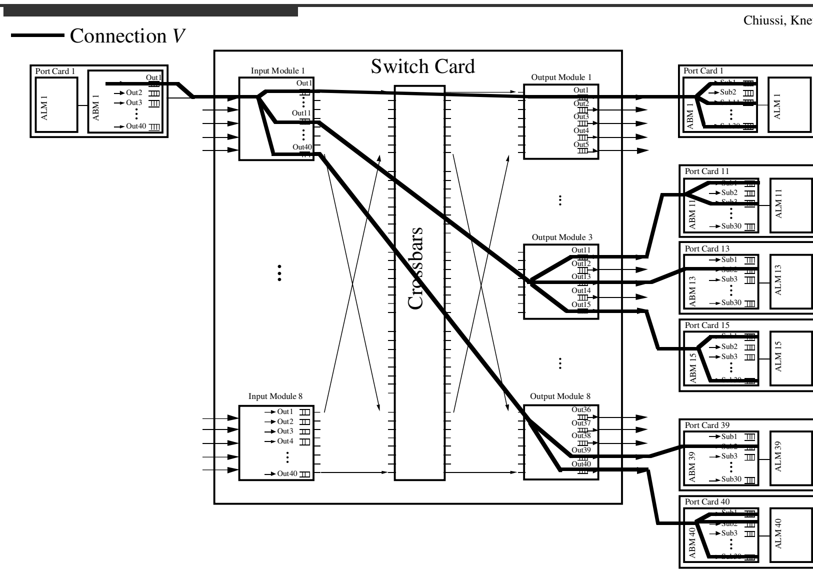


ATLANTA™ Architecture: Multicasting

- Minimum multicast tree
  - Cells always replicated as far downstream as possible in the switch
    - No replication in ingress ABM
    - In first stage of multi-stage fabric only one replication to each desired output module
  - Minimizes expansion in traffic volume internally to the switch due to multicasting
  - Excellent performance
- Replication does not require external control
  - embedded in self-routing structure
- Single copy of a multicast cell locally stored in each buffer
  - replicated (if necessary) only when the cell is sent to following stage in the switch
- Efficient and robust interaction between multicasting scheme and selective backpressure scheme

ATLANTA™ Multicasting: Example

Chiussi, Kneuer, Kumar



Lucent Technologies
Bell Labs Innovations

ATLANTA™ Chipset: Conclusions

Chiussi, Kneuer, Kumar

- Four chips, complete VLSI solution for implementing ATLANTA™ switches ranging in capacity from 622 Mbps to 25 Gbps
- Innovative architecture
  - Nonblocking
  - Scalable, modular
  - Highly-distributed
  - Low complexity
  - Excellent multicasting
- Extensive set of features and hooks
- Majority of buffers in low-cost off-the-shelf RAM's
- Due to its distributed nature, the ATLANTA™ architecture is a very desirable platform for future growth.
  - Even larger switching and buffering capacities
  - Higher link rates
  - More sophisticated buffer management for better provisioning of QoS guarantees (*e.g.*, Per-VC queueing)

Lucent Technologies
Bell Labs Innovations