

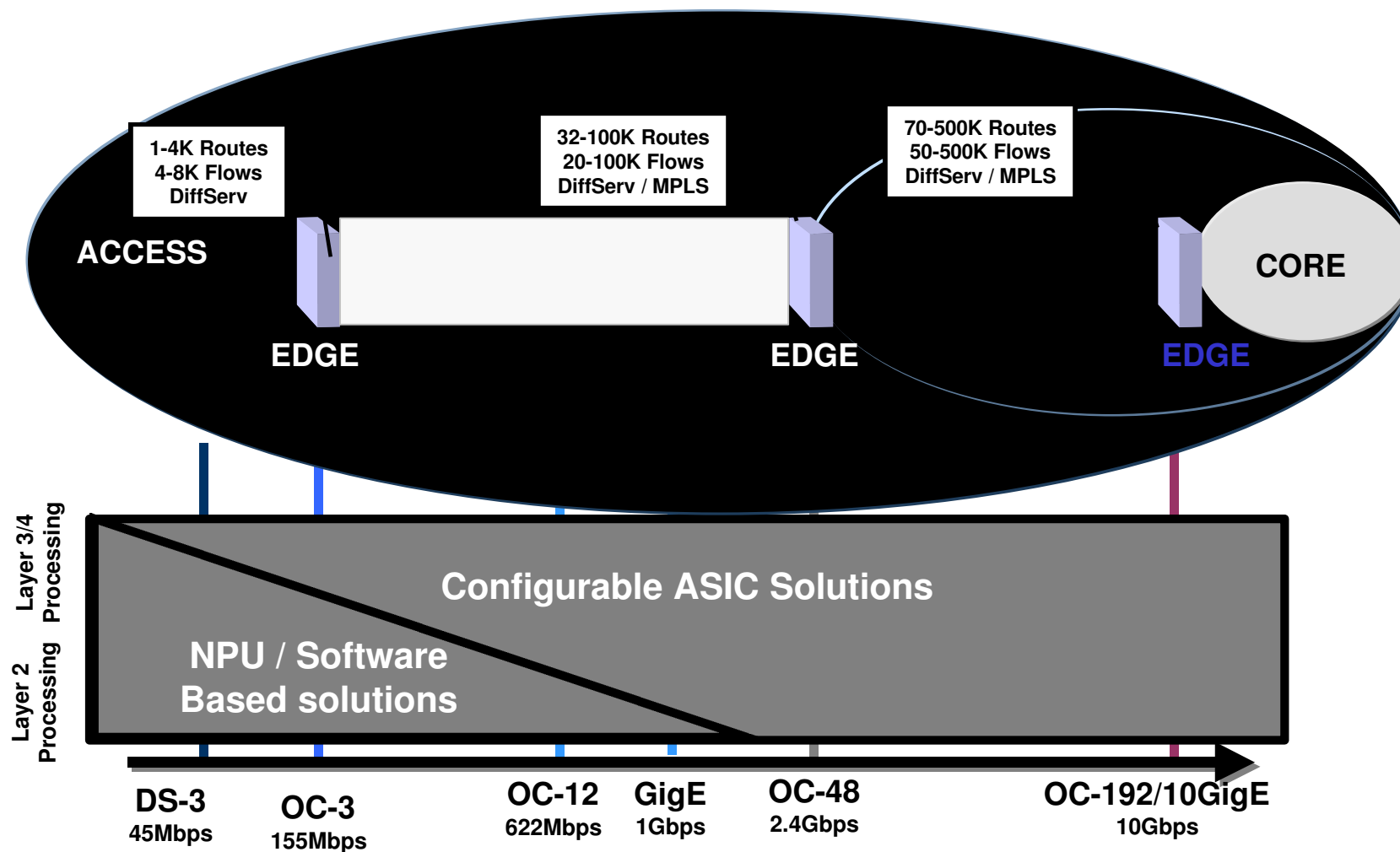


A Low Power Monolithic IP/MPLS Packet Forwarding Engine for Service Provider Networks.

**T. Asami, E. Chao, N. Dagli, J. Dickinson, J. Fiorenza, S. Fallow,
P. Gopi, O. Hassen, J. Hamada, N. Hudson, R. Krishnan, H. Luu,
C. Nabangxang, K. Peng, D. Peeters, P. Wang, A. Yoshida, J. Zoll**

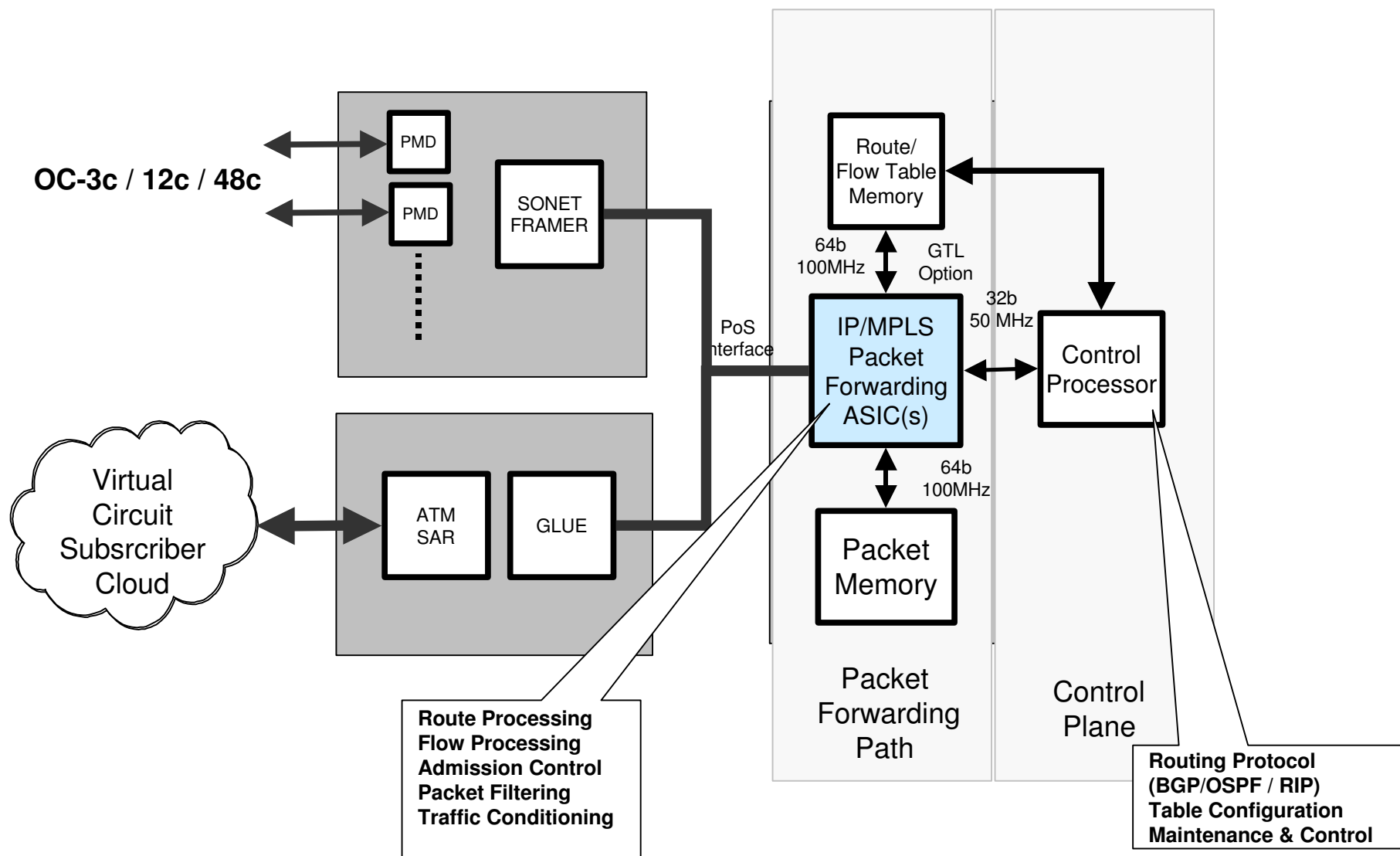
Presenter: Paramesh Gopi

Network Performance View



System View

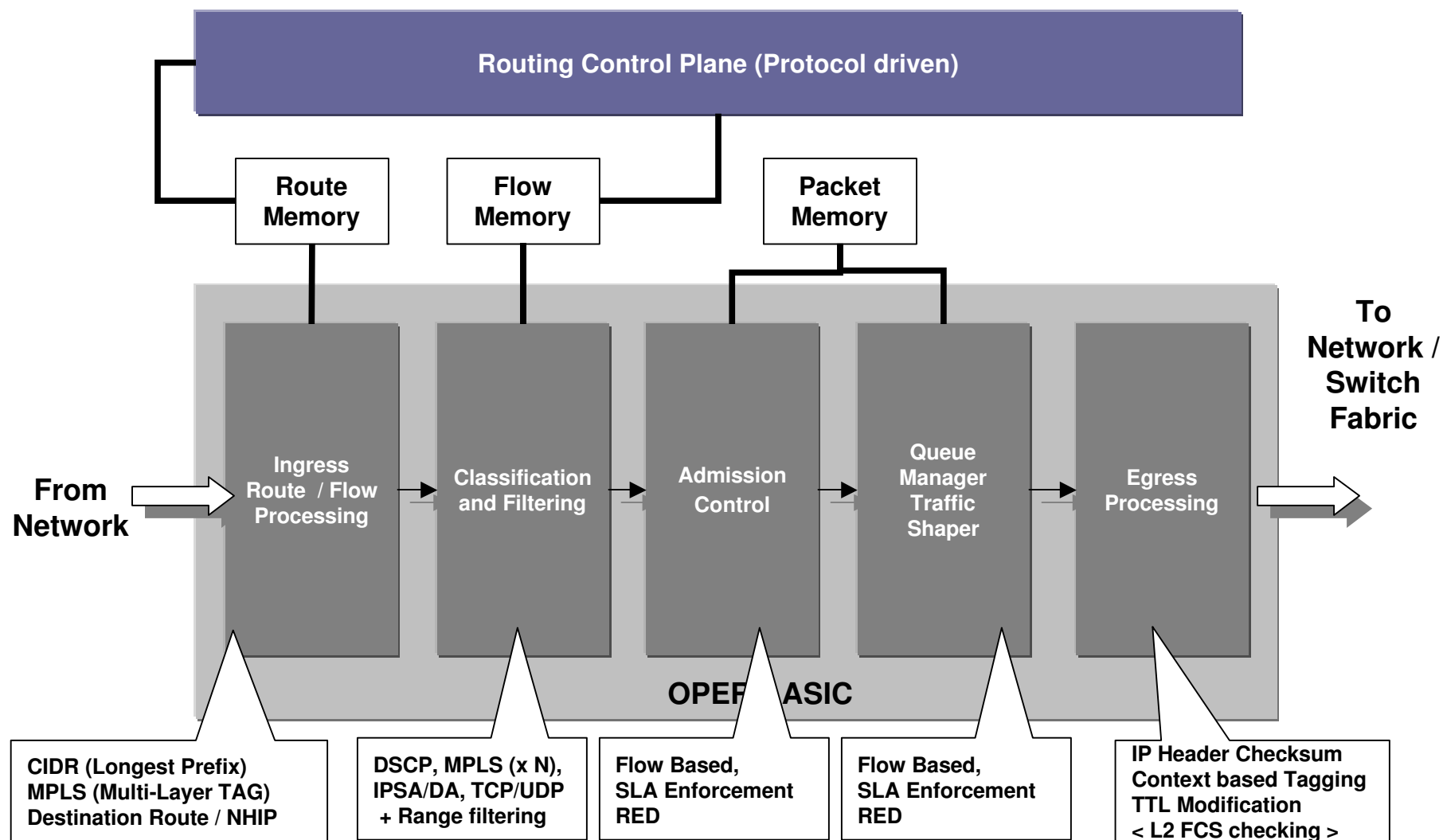
Wire-Speed Line Aggregation





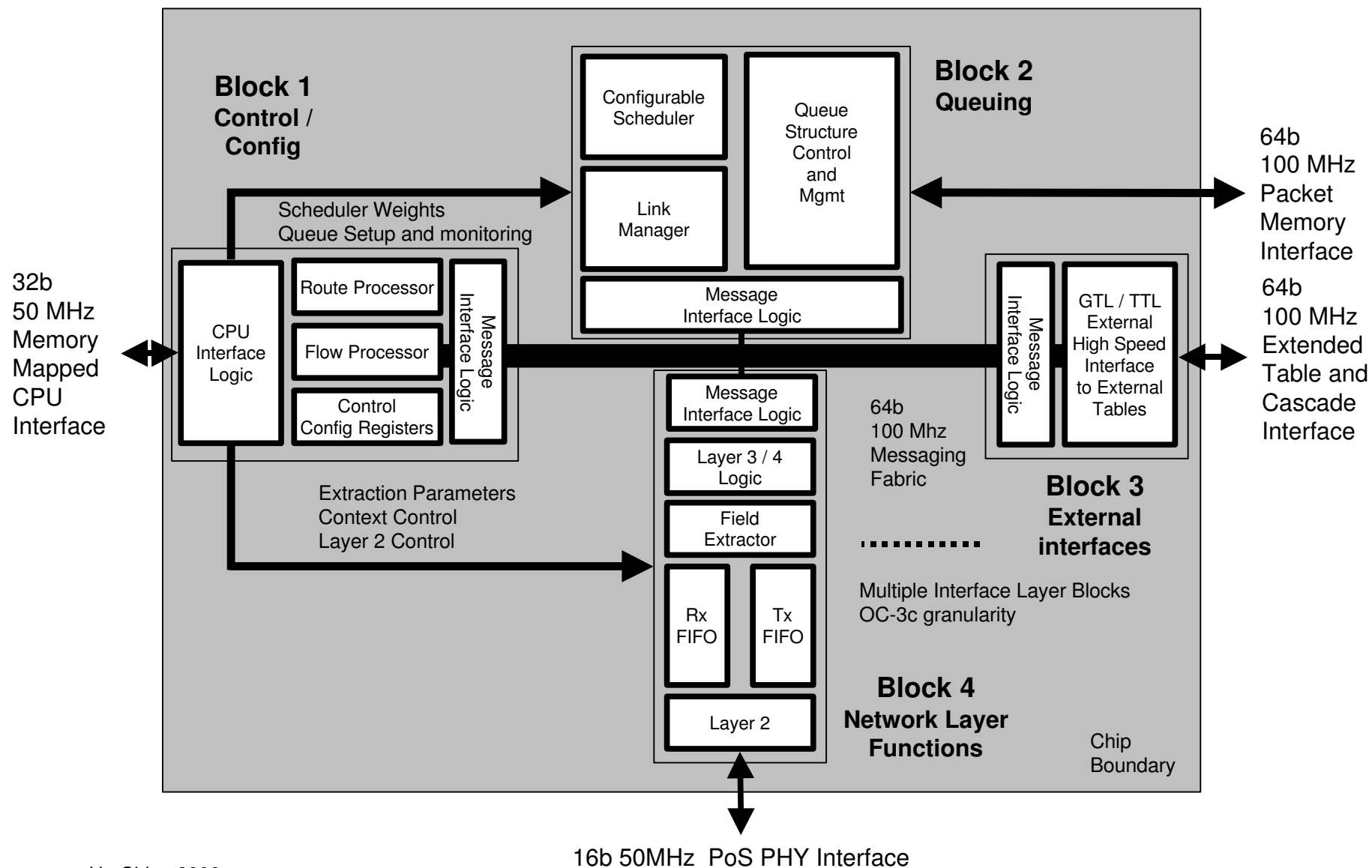
System Requirements

- **Route Processing**
 - Full Internet Feed (100K prefixes)
 - Full Multi-Layer MPLS Lookup (4 LSR prefixes)
- **Flow Classification**
 - IP Src / Dest Address, Protocol Field
 - DiffServ Code Points
 - Layer 4 TCP/UDP Port numbers
 - Layer 4 TCP/UDP Port ranges
 - 64K Flows
- **Admission Control**
 - Flow Aggregates based on Flow Classification events
- **Policy driven congestion recovery**
 - RED, WRED
- **Traffic Conditioning**
 - Fair Queueing
 - Weighted Round Robin
 - 4 Delivery Priorities / 4 drop precedences per interface
- **Maximize Connection Density**
 - CO POP space is a premium commodity
- **Minimize the power figure per rack foot**
 - Carriers are limited by their infrastructural power / heat management requirements



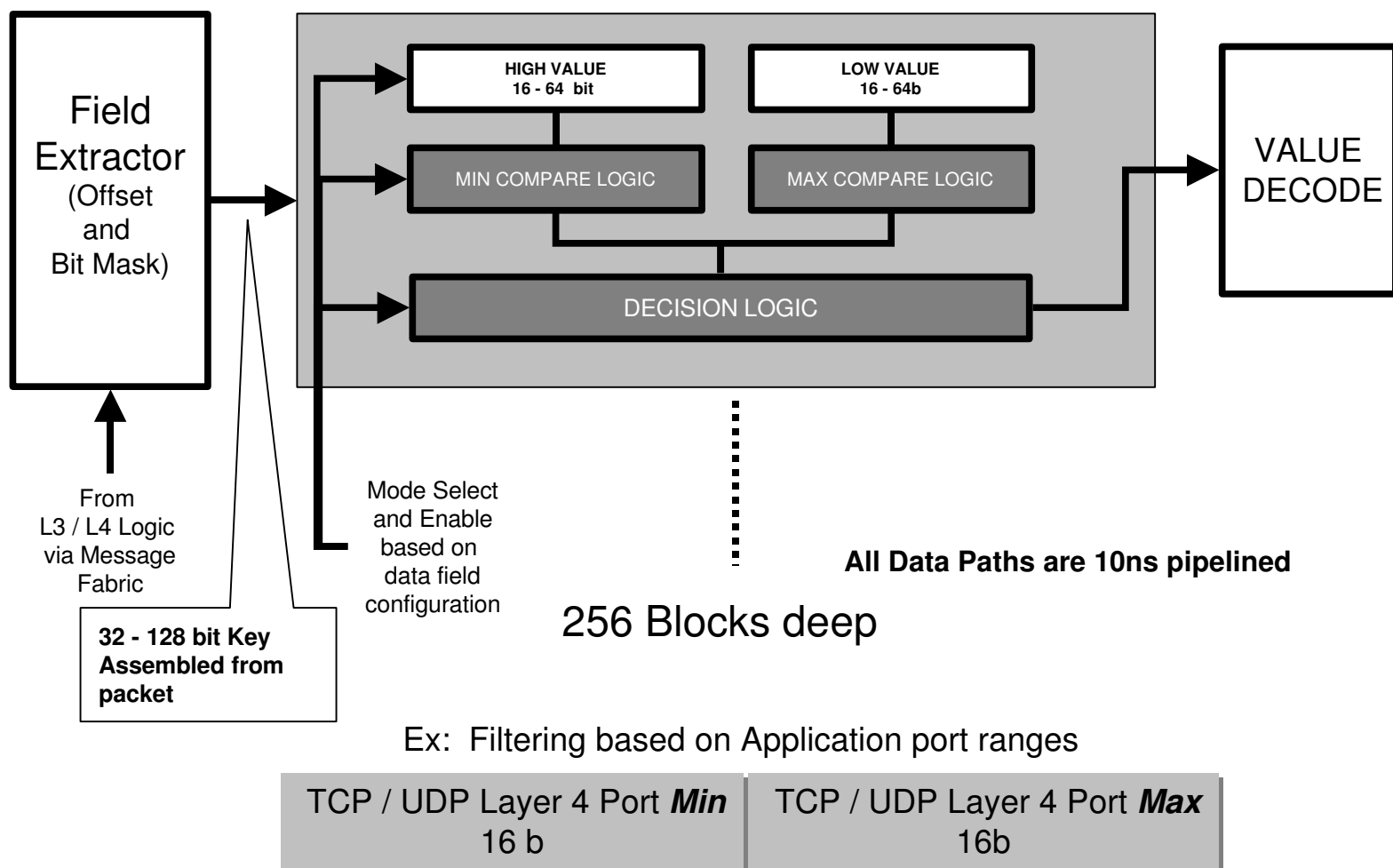
Micro-Architecture

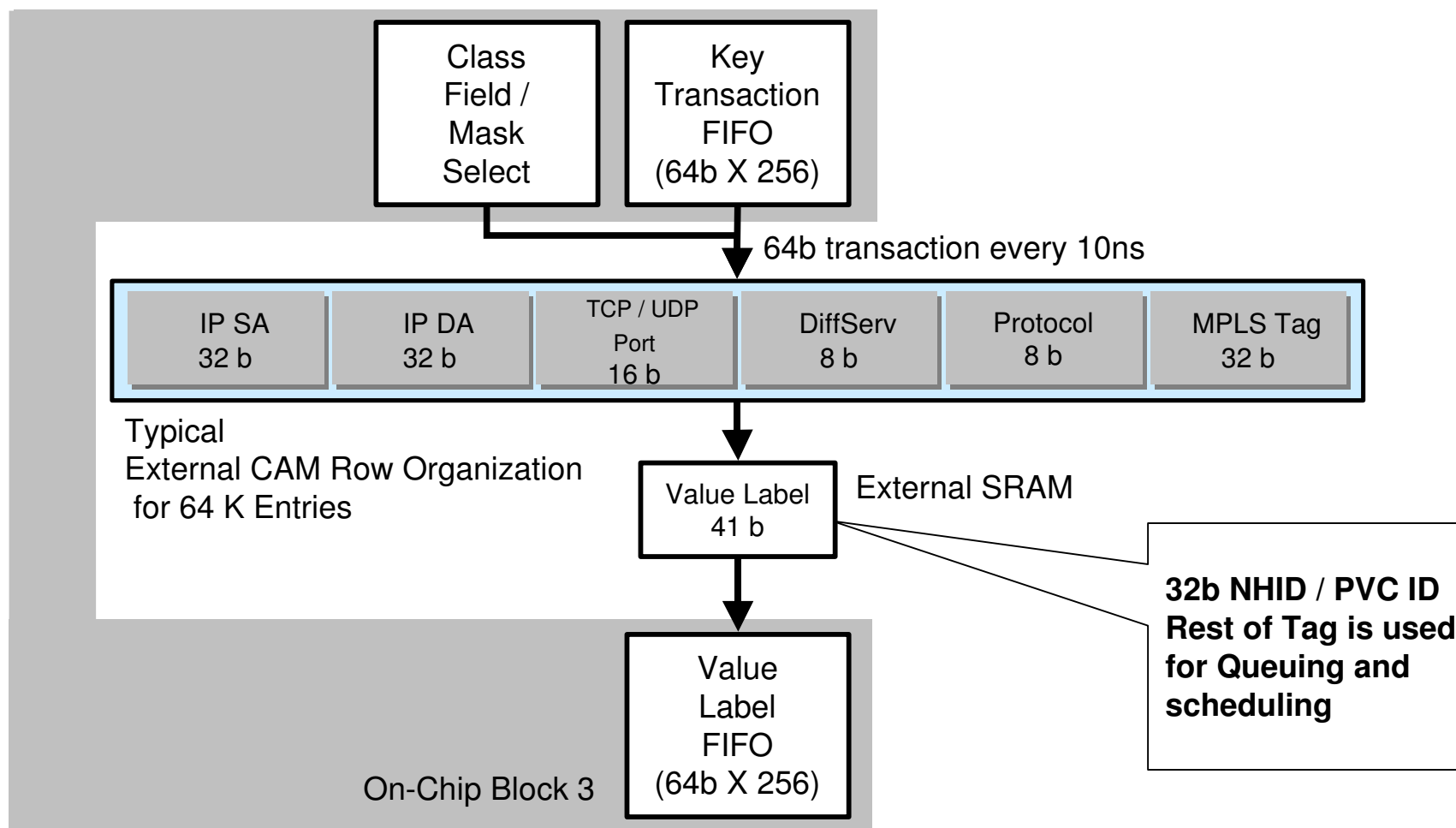
Functional Partitioning



Micro-Architecture

Pre-Emptive Multi-Field Range based classifier

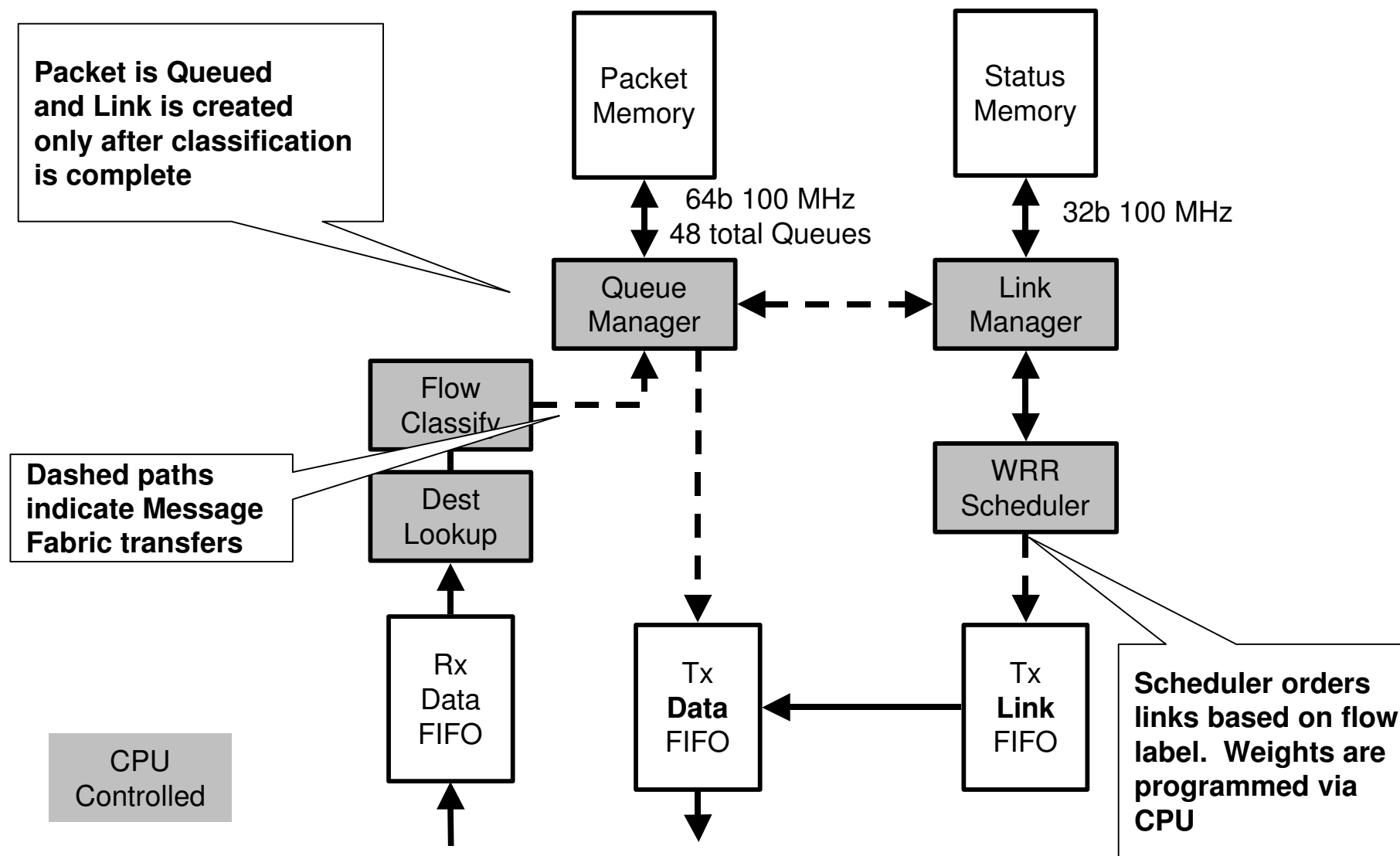






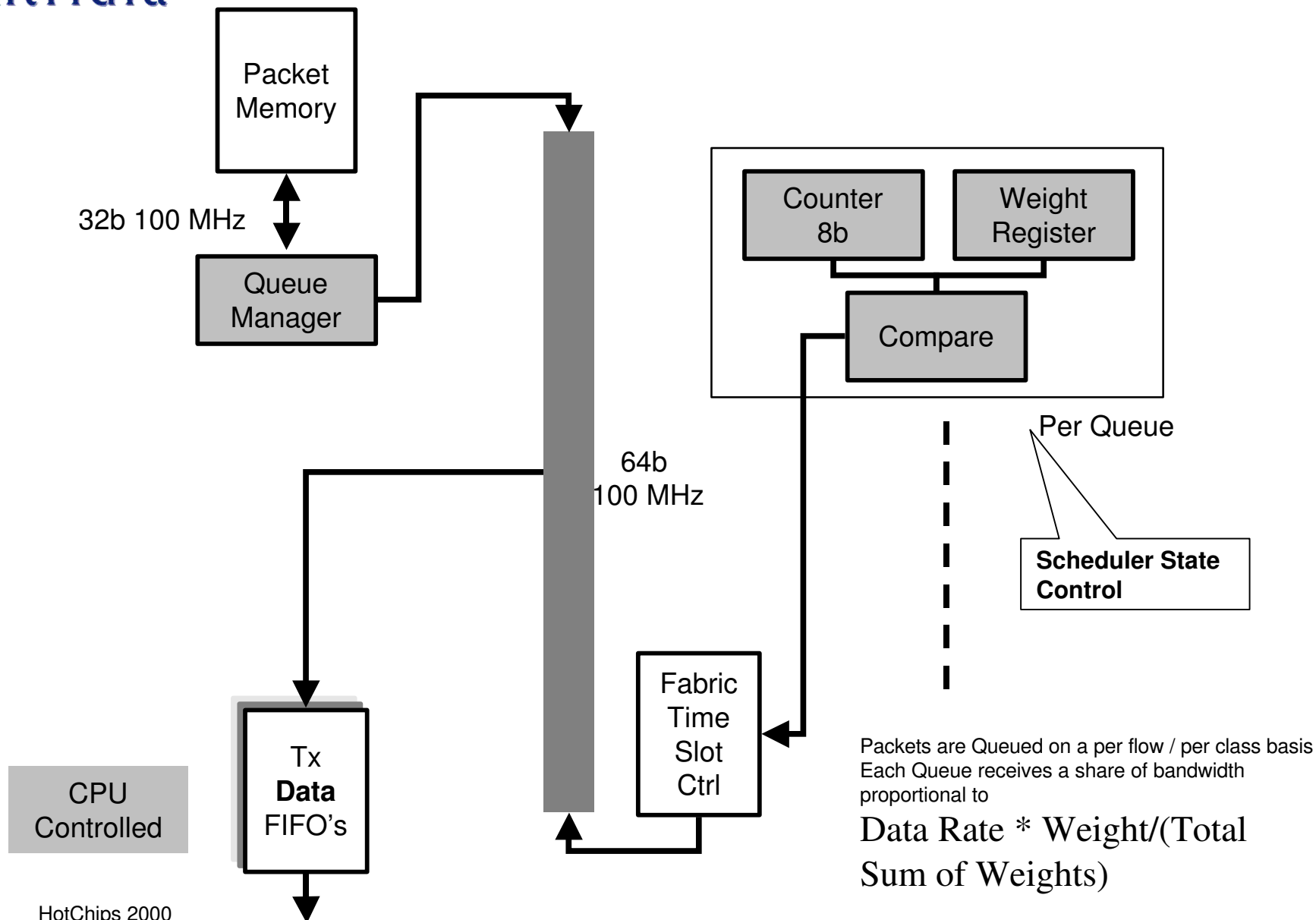
Micro-Architecture

Queuing Processes



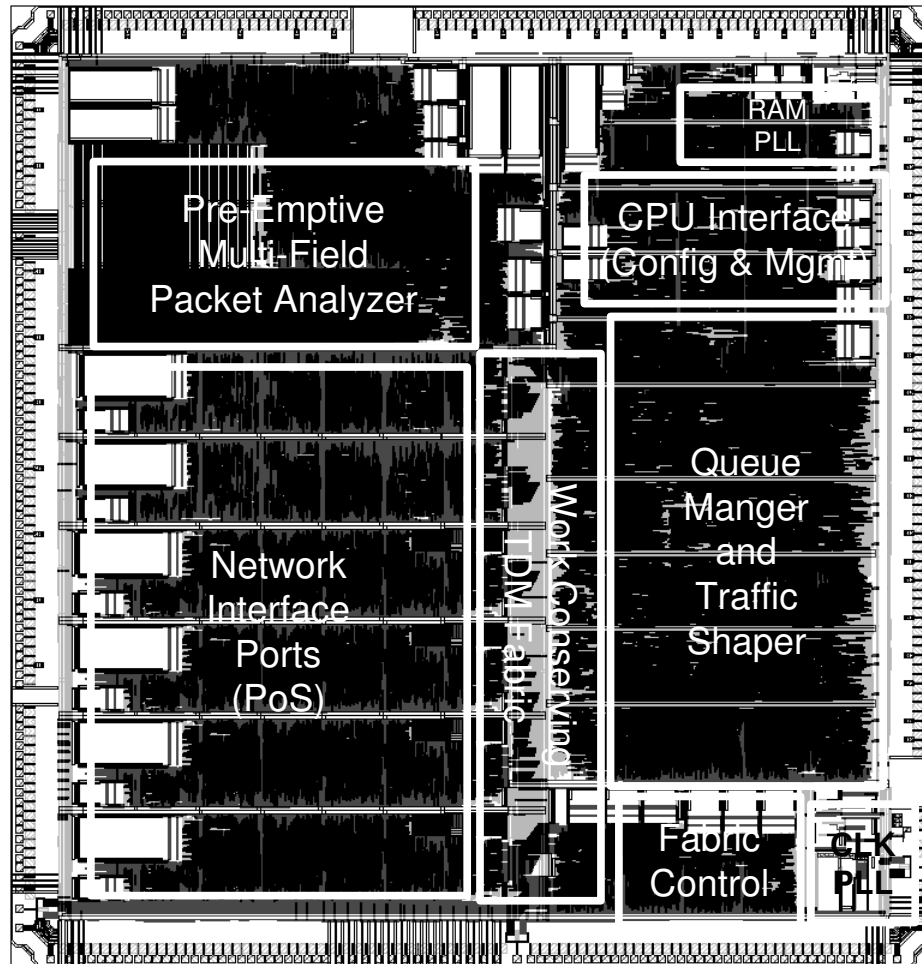
Micro-Architecture

Scheduler (Egress Detail)



Micro-Architecture

Floorplan





Silicon Validation

Key Metrics

Process	0.25um TSMC CMOS SP5M
Gate Count	2.5 Million
Package	520 HPBGA
Power Dissipation (nominal)	6.5W
Memory	ZBT SRAM (Status / Packet)
Maximum Wire-speed Throughput (40 byte packets)	4.3 Million packets per second
Maximum forwarding Latency	7 μ s
Maximum Lookup table density	100,000 IPv4 Routes 64,000 128b Arbitrary Flows



Silicon Architecture

Key Innovations

- **Packet Classification and Filtering**
 - Multi-value denominative packet classification methodology
 - 10ns pipelined non-algorithmic circuit structure
- **Dual Stage Memory Architecture**
 - Small Packet cache to maintain sustained 40byte packet transfers indefinitely
 - Main Packet memory to store large packets
- **Scalable Work Conserving TDM Messaging Fabric**
 - 6.4 Gbps
 - Single cycle (10ns) agent access
 - Single cycle arbitration
 - Deterministic transaction resolution results in extremely small latency variance
 - 10% signalling overhead