

---

## **Synfinity II – A High-Speed Interconnect with 2GBytes/sec Self-Configurable Physical Link**

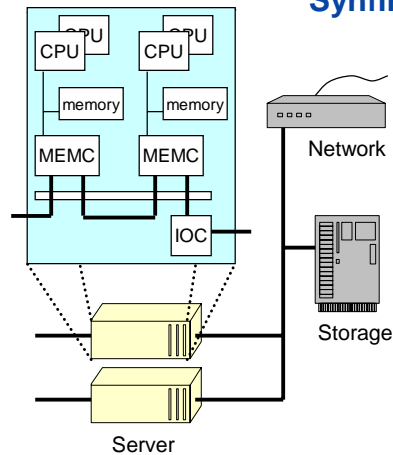
Yoichi Koyanagi, Takeshi Horie, Takashi Miyoshi  
And Mitsuo Ishii

## **Outline**

---

- 1. Design goals**
- 2. Architecture**
- 3. Synfinity II chip set – MIC2 and Router2**
- 4. Evaluation**
- 5. Related technologies**
- 6. Summary**

## Design Goals



### Synfinity II – Interconnect for cluster and back-plane connection

1. High bandwidth and low latency data transfer
2. Scalable and flexible architecture
3. RAS (Reliability, Availability and Serviceability) features
4. Cost-effectiveness

## Issues and Solutions

### 1. High bandwidth and low latency

Serialization overhead

→ Parallel-bit physical link

Routing table lookup overhead

→ Source routing

### 2. Scalability and flexibility

Dead-lock loop limits the network topology

→ Virtual lanes in a virtual channel

Application independence

→ Layered Architecture

Support both inter-chip and inter-cabinet connection

→ Self-configurable physical link

Centralized single clock degrades system scalability

→ Plesiochronous clocking

## Issues and Solutions (cont'd)

### 3. RAS features

Packet errors caused by noisy transmission lines

→ Packet retransmission

Unexpected unplugging or failure

→ Automatic reinitialization (hot-plug) physical link

Network management

→ In-band access (initialization, error/status collection)

Link failure on the path

→ Automatic rerouting

### 4. Cost effectiveness

External physical link transceiver

→ Integrated CMOS physical link

Expensive low-skew cable and PCB design

→ Automatic deskew mechanism

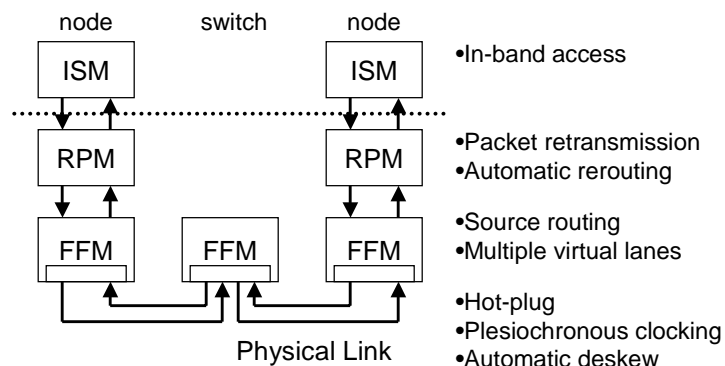
## Synfinity II Architecture

### Layered Architecture

ISM: Interconnect Services Manager

RPM: Reliable Packet Mover

FFM: Fast Frame Mover



## Physical Link

### 2GBytes/sec/direction parallel link

- ✓ 20bits data, 1bit clock (symbol rate: 1Gbps)

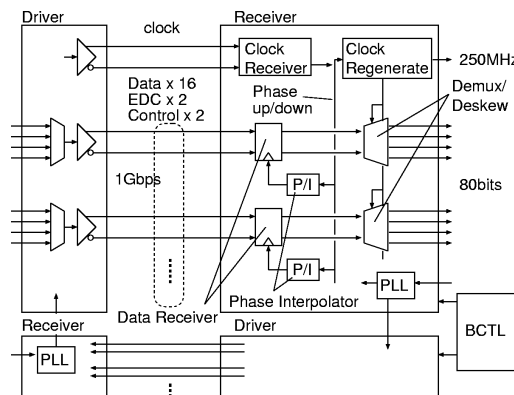
### Plesiochronous clocking

- ✓ Tolerate +/- 200ppm frequency difference

### Basic Control Logic (BCTL)

- ✓ Automatic deskew
- ✓ Automatic self-configuration (reinitialization)
  - Data rate: 2G, 1G, 500M, 250M Bytes/sec
  - Driver current adjustment to minimize power consumption
- ✓ Link exerciser for data transmission test

## Physical Link Architecture



### Clock receiver

- Phase tracking

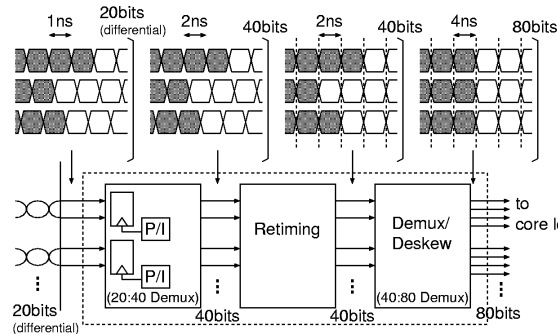
### Phase interpolator

- Timing generator

### PRD receiver [Gotoh et al. ISSCC99]

- Equalization for long cable

## Link Training Step (Initialization)



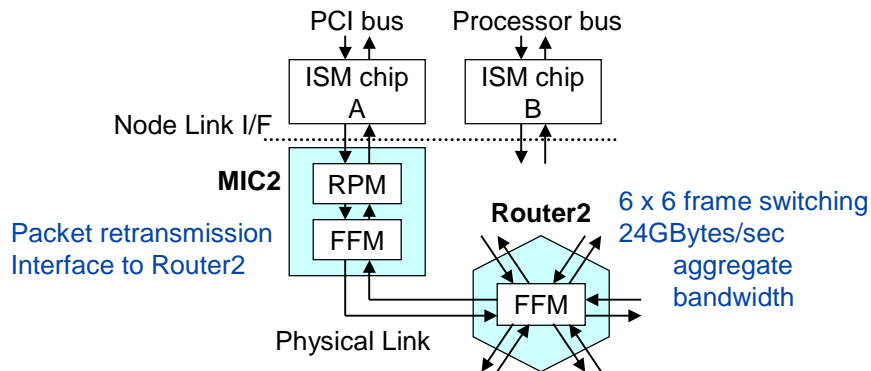
1. Phase adjust
2. Retiming
3. Deskew
4. Link exercise
5. Repeat until optimal speed/current

Maximum 4 bit-time skew can be compensated

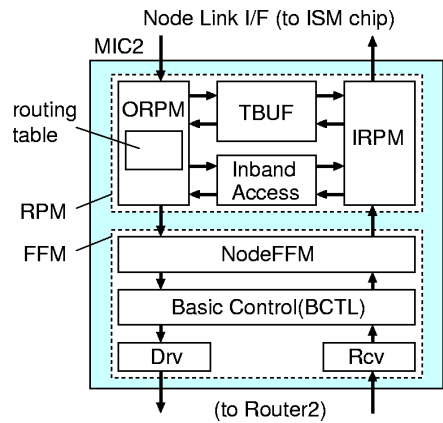
## Synfinity II Chip Set

**MIC2: Node Interface chip (RPM, FFM)**

**Router2: Switch chip (FFM)**



## MIC2 Microarchitecture



### Outgoing RPM (ORPM)

→ Retransmission

### Transmission buffer (TBUF)

→ Save packets

### NodeFFM

→ Interface to Router2

### Incoming RPM (IRPM)

→ Check packet, return acknowledge

## Router2 Microarchitecture

### Output Arbiter

→ Assign grant

### Cross-bar

→ Data switch

### Frame Buffer

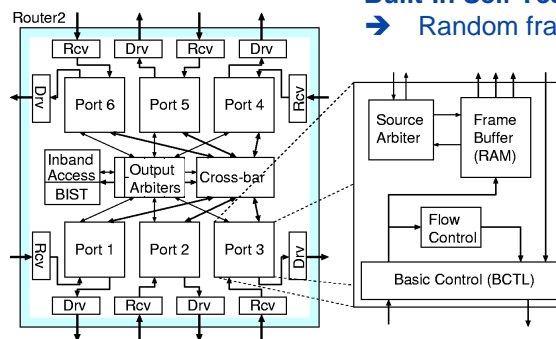
→ 20-entry frame pool (RAM)

### Source Arbiter

→ Arbitration request control

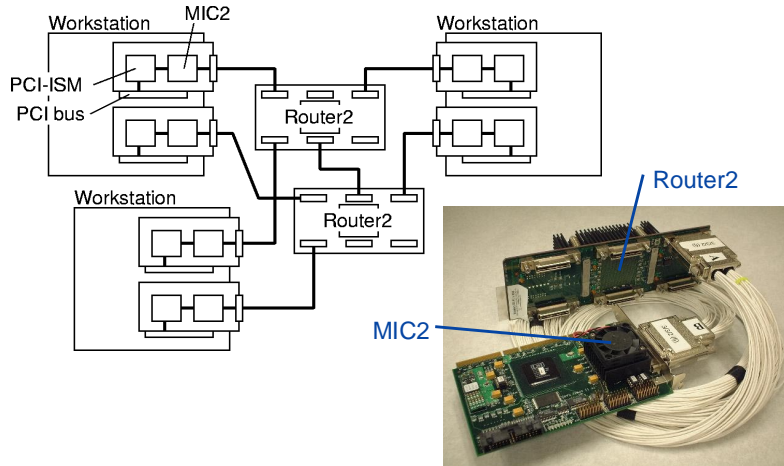
### Built-In Self Test (BIST)

→ Random frame generator



## Evaluation

### Workstation cluster configuration



## Evaluation Results

### Physical link initialization

- Automatically initialized successfully

### Self-Configuration

- 2GBytes/sec in 5m cable, 1GBytes/sec in 15m cable

### Message passing test program

- Ran stably with 2GBytes/sec link

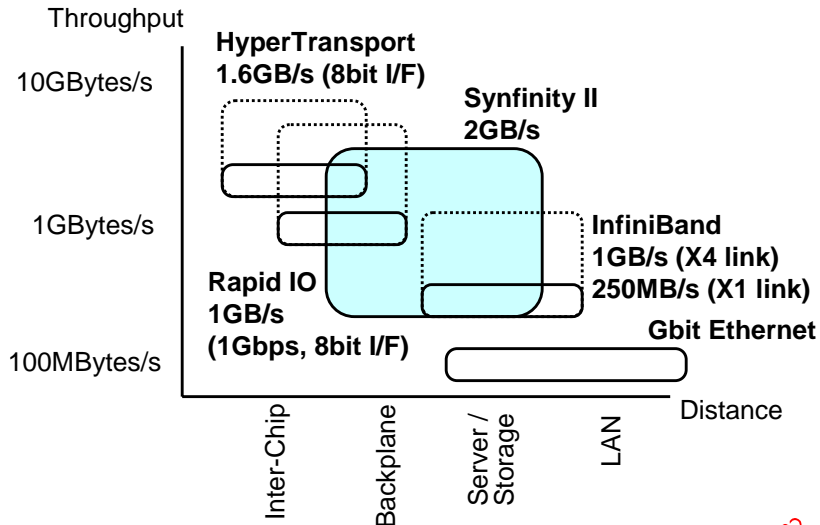
### Unplug and re-plug the cable while sending data

- Data transfer was resumed automatically

### Measure BER using link exerciser and BIST

- Under  $10^{-12}$  with 2GBytes/sec data transfer

## Related Technologies



## Summary

### Summary

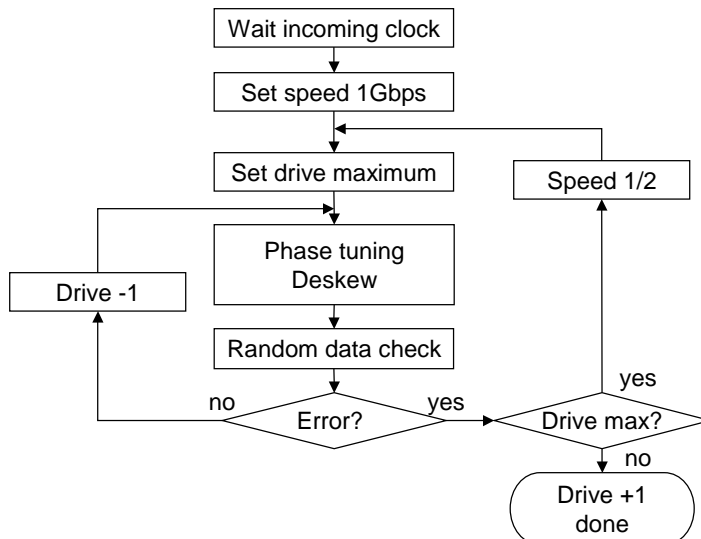
- ✓ Synfinity II realizes High-speed low latency interconnect infrastructure with RAS functions

### Future goals

- ✓ Enhance throughput for high-performance server systems
- ✓ Leverage other interconnects with Synfinity II physical link technology

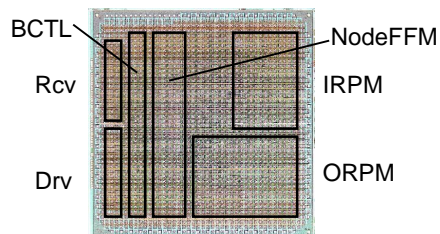
## Appendix

## Automatic Link Configuration



## MIC2 – A Node Interface Chip

<b>Functions</b>	RPM – Packet re-transmission FFM – Interface to Router2
<b>Frequency</b>	250MHz (FFM), 66/200MHz (RPM)
<b>Die size</b>	10mm x 10.2mm 0.25um CMOS
<b>Gates</b>	380K gates



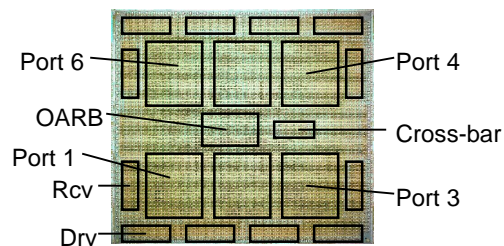
Fujitsu Laboratories of America

19

FUJITSU

## Router2 – A Switch Chip

<b>Functions</b>	FFM – 6 x 6 frame switching
<b>Aggregate bandwidth</b>	24GBytes/sec
<b>Frequency</b>	250MHz
<b>Die size</b>	15.6mm x 15.6mm 0.25um CMOS
<b>Gates</b>	1.2M gates



Fujitsu Laboratories of America

20

FUJITSU