



Non-Intrusive Deep Tracing of SCI Interconnect Traffic

Michael Manzke and Brian Coghlan



Non-Intrusive Deep Tracing of SCI Interconnect Traffic

- Why SCI Deep Trace Acquisition?
- Design Objectives
- Trace System Overview
- Hardware
- Software
- Future Work
- Conclusions



Why SCI Deep Trace Acquisition?

- Non-Intrusive spatial and temporal analysis
 - Analysis of:
 - Interconnect hardware
 - Interconnect topology
 - Driver, protocol, API and DSM software
 - Parallel executed algorithms

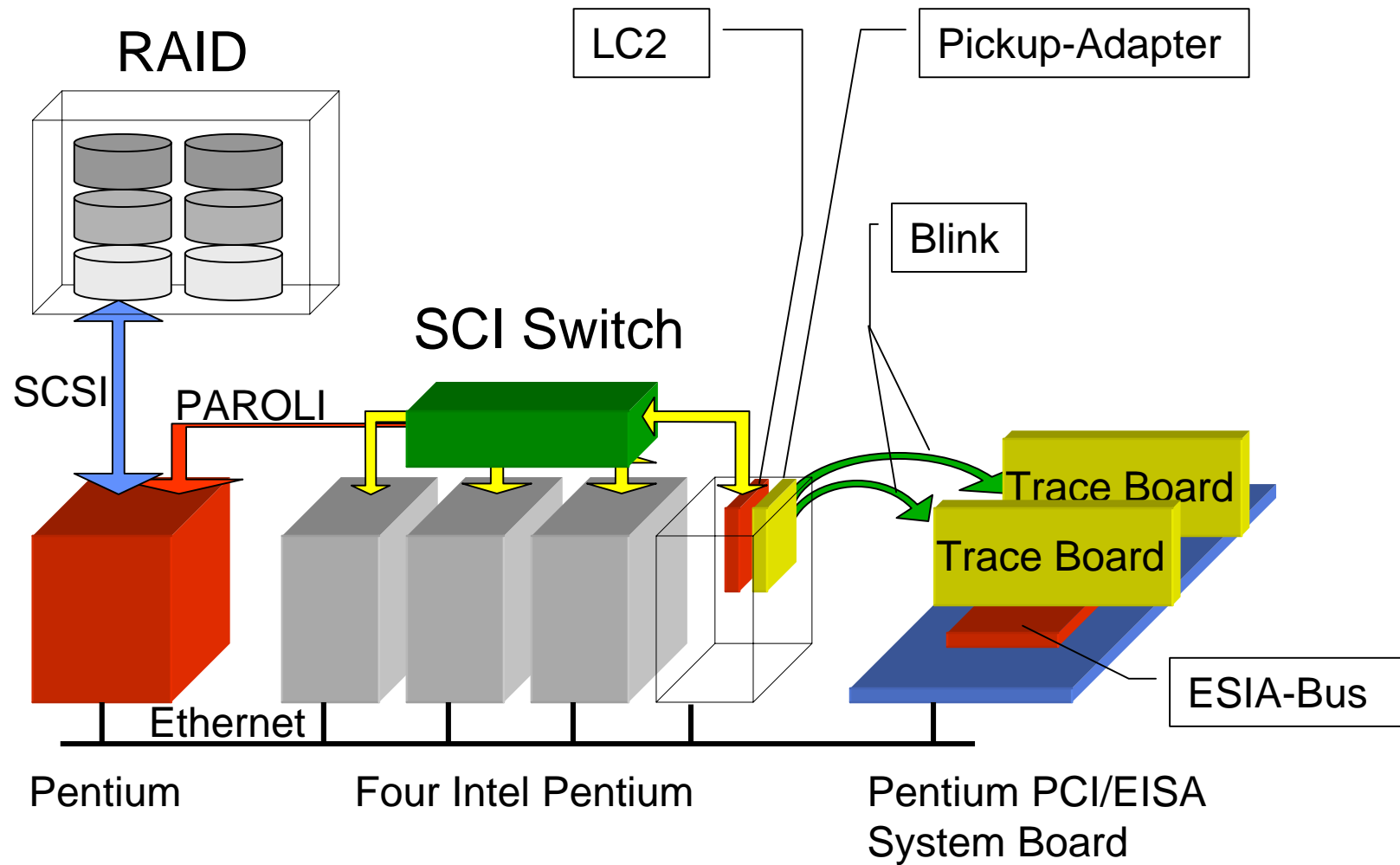


Design Objectives

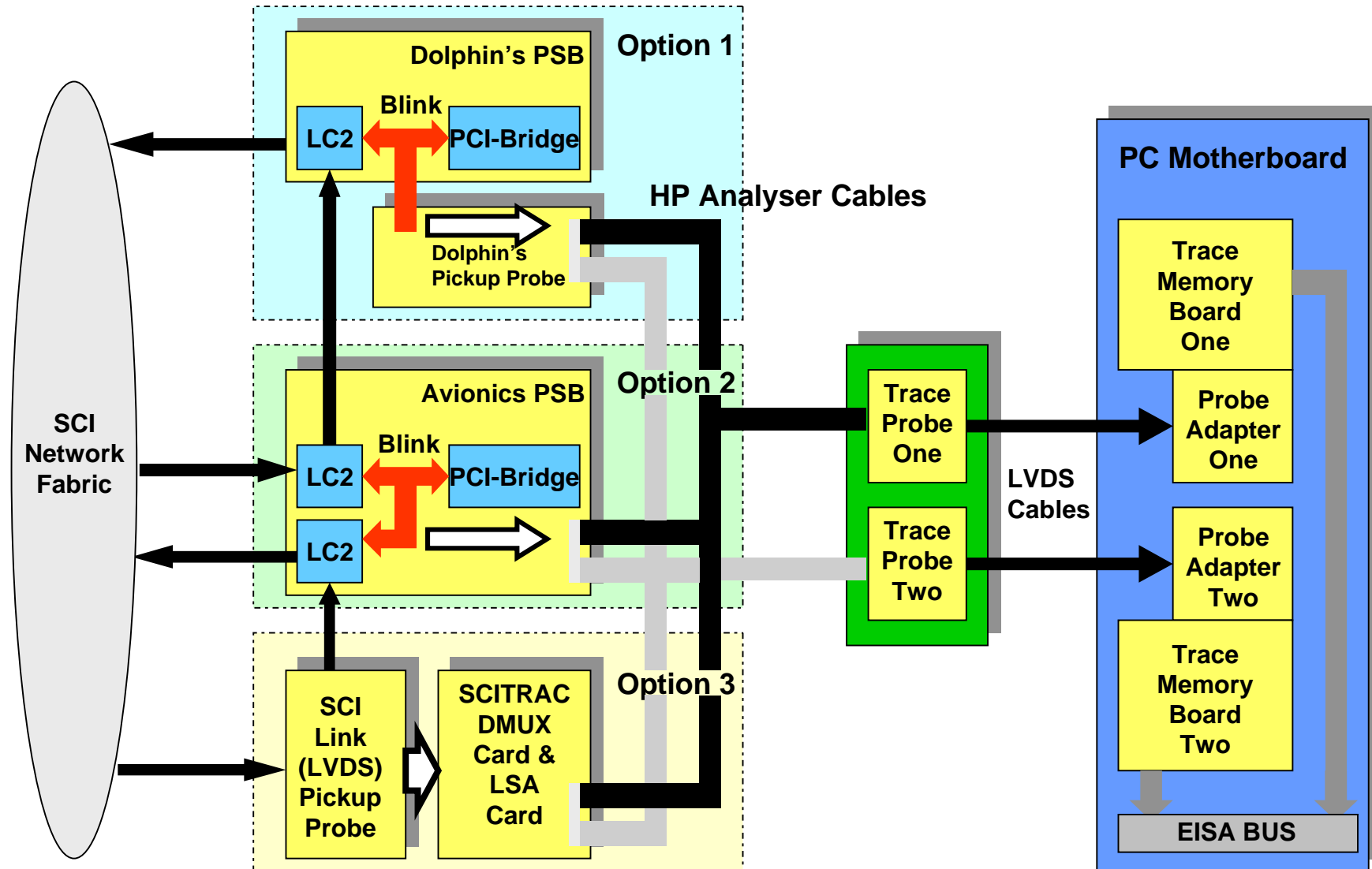
- Non-intrusive monitoring
- Very deep ($\gg 10$ Mbyte) traces
- Acquisition of all packets
- Synchronous trace acquisition on multiple nodes
- Allowance for various probes
- Adaptability to various SCI interface implementation
- Trace data storage in commercial relational database
- Ability to analyse causal relationships



Trace System



Trace Hardware

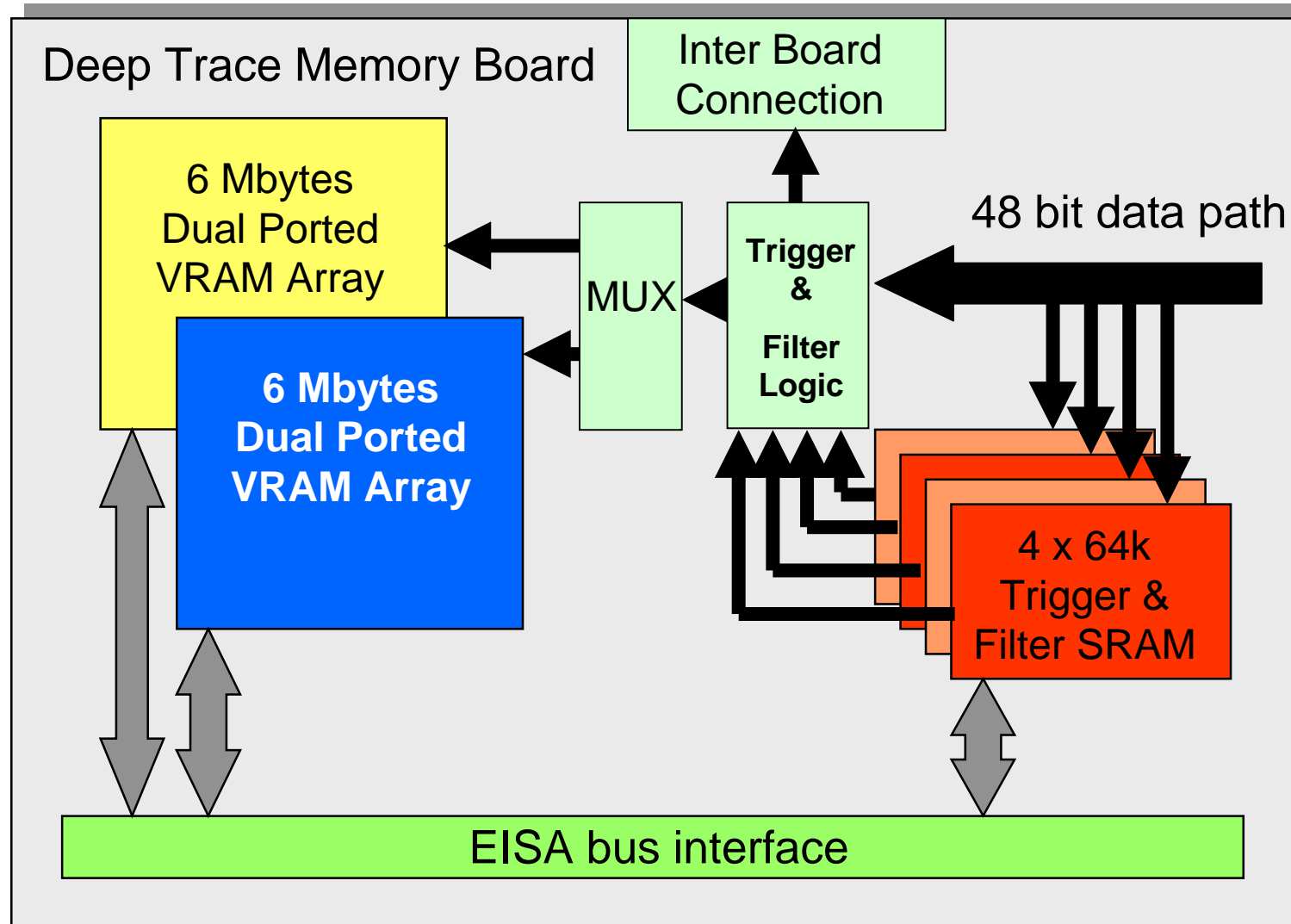


Hardware

- Trace probes, trace adapters and trace memory
- 96 bit-sample data path
- Maximum sample rate of 66 MHz (LC2 suitable)
- Test pattern generator in trace probes
- 24 Mbytes dual ported VRAM (on two boards)
 - > 2,000,000 samples
 - > 150,000 Request Send Packet with 64 bytes data
- Four 64k SRAMs for trigger and filter pattern



Trace Memory Board



Trace Instrument GUI

The image displays two overlapping windows from the SCI Trace Instrument GUI.

SCI Trace Instrument - SRAM Window:

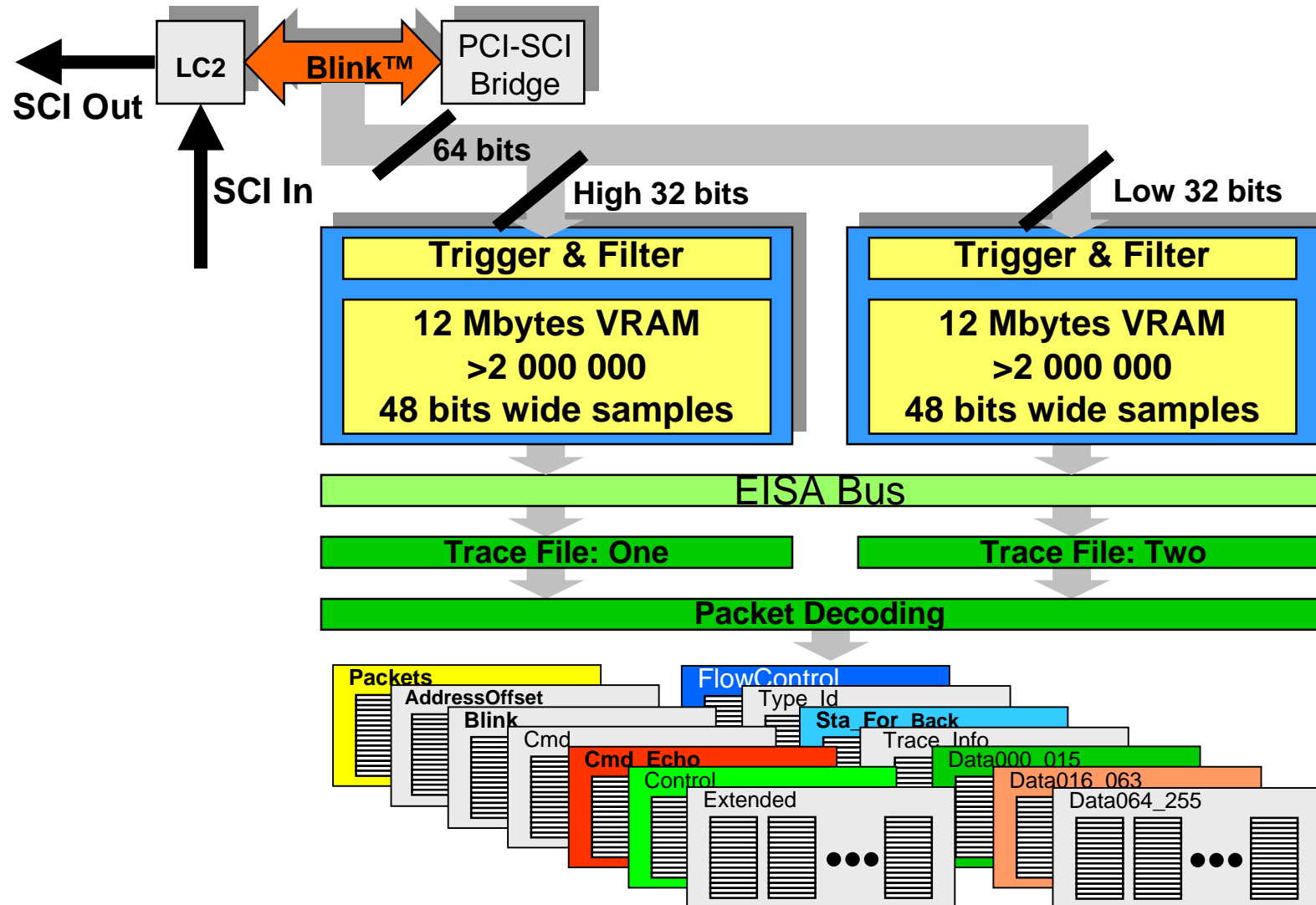
- Write into SRAM:** Includes a 'Write' button and a 'Trigger/Trace' section with checkboxes for Channel #3, Channel #2 (checked), Channel #1, and Channel #0.
- Broken into three blocks of 16 bits:** Six symbol input fields (Symbol 1 to Symbol 6) each containing '0000', followed by a 'Hex' label.
- Read from SRAM:** Six checkboxes for SRAM5 D[3..0], SRAM4 D[3..0], SRAM3 D[3..0], SRAM2 D[3..0], SRAM1 D[3..0], and SRAM0 D[3..0], all of which are checked.
- 96 bit Triggers:** A grid of dropdown menus for SRAM 5, SRAM 4, SRAM 3, SRAM 2, SRAM 1, and SRAM 0, each with four channels (Cha. 0 to Cha. 3).
- Buttons:** 'SCAN', 'Wipe', and 'Exit'.

Register Window [EISA Slot-2]:

- Registers:** Includes tabs for Registers, VRAM, SRAM, Tests, Stats, and Verify.
- Table:** A table with columns: Address[Hex], Sample[Dec], Contents[Hex], Sub-Bank[Dec], Bank[Dec], and Row[Dec]. It contains 12 rows of data.
- Buttons:** 'Log', 'Dump', 'Wipe', 'Update', and 'Exit'.
- Trace Board Slot (3):** Includes fields for Sample Stop Count (00FF8000), Address Map (FF000000), and Serial Buffer Byte Enables (5.0, all checked).
- Mode:** A dropdown menu set to 'SRAMLD'.
- Status:** A checkbox labeled 'Done'.
- Head Transfer Pointer:** A field containing '00FF8000' with a 'Write' button.
- EISA ID:** A field containing 'FFFFFFF'.
- Trigger/Trace Config:** Checkboxes for Chan. 3.0 (checked), Trace Continuously (checked), and Trigger Immediately (checked).
- Buttons:** 'Quick Test', 'Update', 'BoardReset', 'TimeReset', and 'Re-map'.

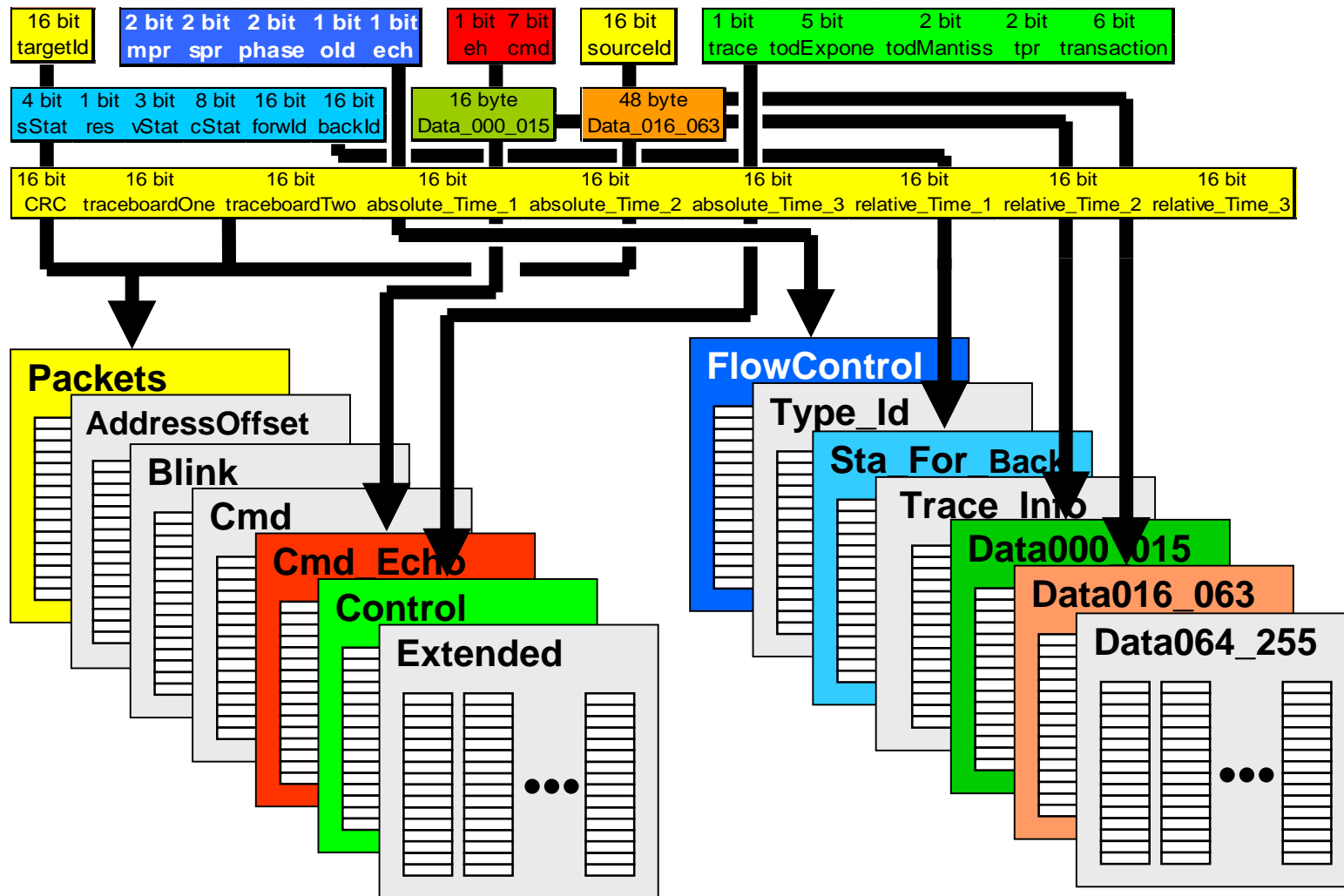


Trace Data Flow

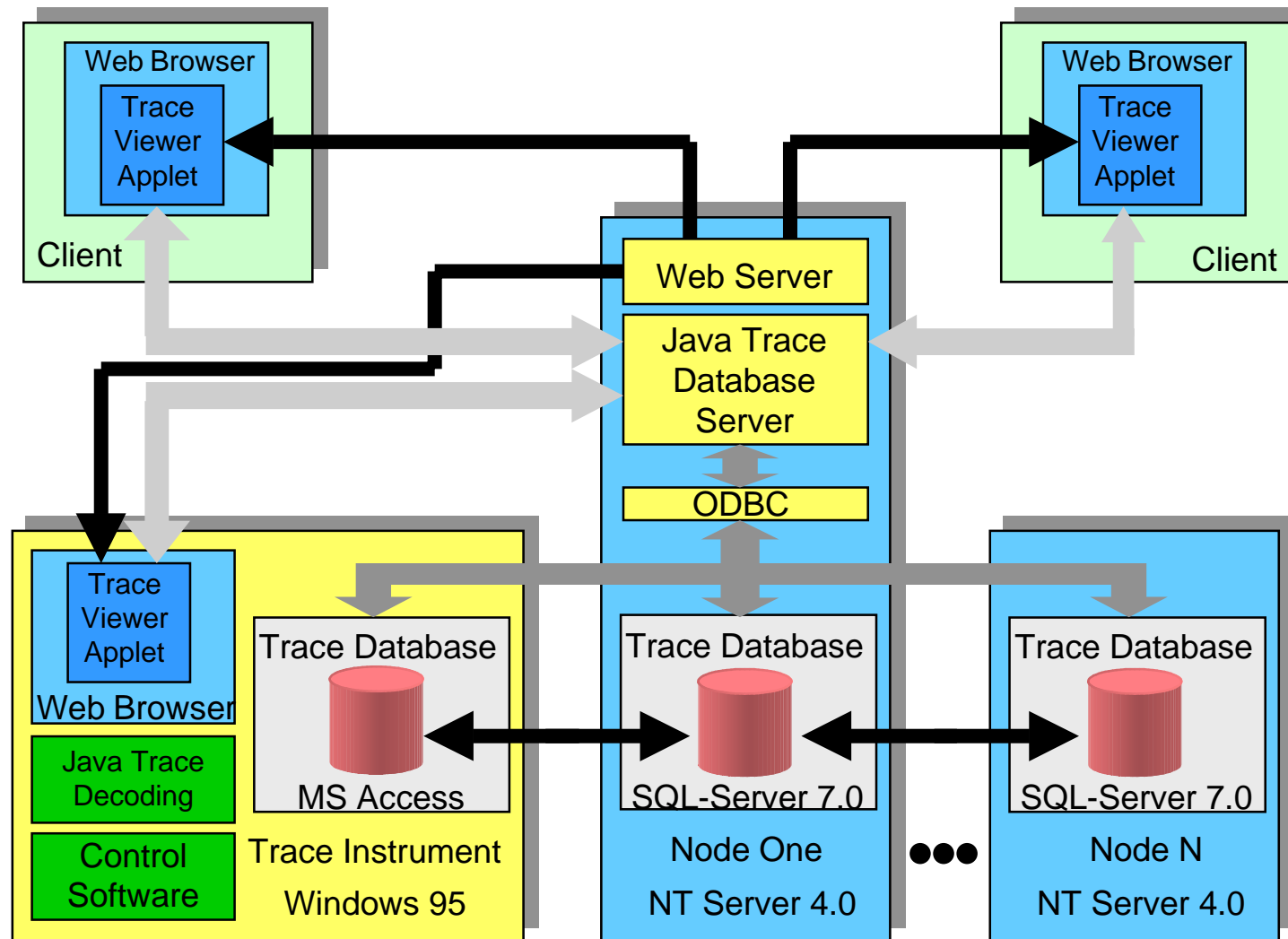


SCI Packet Decoding

Response-send-packet with 64 bytes data



Trace System Software



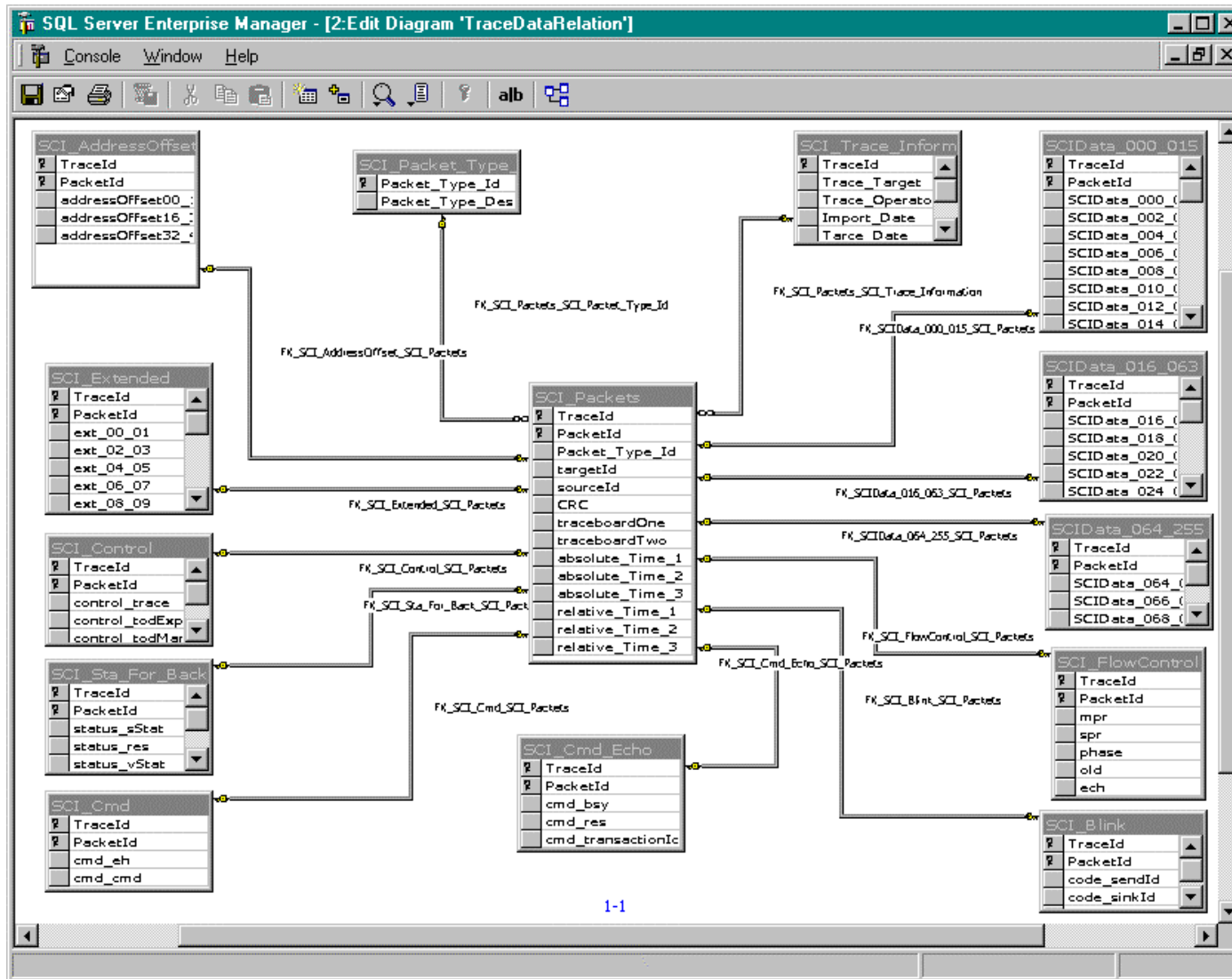
Trace Database

- Trace_ID and Packet_ID uniquely identify every packet
- DB accommodates all SCI packet types (Blink & cable)
- A main table is shared by all packets
- Table fields exhaustively enumerate packet information
- Table fields preserve the maximum level of detail

Query: All request-send packets with targetId = X
source Id = Y
addressOffset > A
addressOffset < B



Trace Database



Trace Database

Applet Viewer: SCIDatabaseMain

Applet

Code (blink)					Traget	Cmd(blink)	Flow Control					Cmd	
sendId	sinkId	vd	re	transId	targetId	priority	mpr	spr	todMantiss	tpr	transaction	eh	cmd
ac61	ac61	ac61	ac61	ac61	ac61	ac61	acff	acff	phase	old	ech	6eff	6eff

Cmd Echo			Source	Control				
bsy	res	tnsactionId	sourceId	trace	todExpone	todMantiss	tpr	transaction
dc43	dc43	dc43	e423	ac61	6eff	fb56	4298	efff

Address Offset			Status - Forwld - BackId						Extented Header							
00 to 16	16 to 31	32 to 47	sStat	res	vStat	CStat	forwld	backId	00&01	02&03	04&05	06&07	08&10	10&11	12&13	14&15
ac61	ac61	ac61	ac61	6eff	fb56	4298	efff	6df3	ac61	ac61	ac61	ac61	acff	ac61	ac61	ac61

16 Byte Data							
00&01	02&03	04&05	06&07	08&10	10&11	12&13	14&15
ac61	6eff	fb56	4298	efff	6df3	78e4	3232

48 Byte Data												
byte	00&01	02&03	04&05	06&07	08&10	10&11	12&13	14&15	16&17	18&19	20&21	22&23
016 to 031	ad3d	b345	5ef1	76cd	dc43	ffab	52fe	4319	eebe	f451	5e1f	4561

Database Host: CAGraidserver.cs.tcd.ie

Exit Reset

Applet started.



Future Work

- Implementation of:
 - State (gantt) & causality (hasse) diagrams
 - Statistical analysis
- The instrument will be used to Validate:
 - Global cluster state estimation algorithm
 - SCI Simulation tuning



Conclusion

- Non-intrusive measurement of SCI traffic
- Analysis of the true temporal behavior
- Deep traces > 150,000 packets per node
- Synchronous trace acquisition
- Trace Database provides maximum detail
- Trace DB is well understood and easy to use

