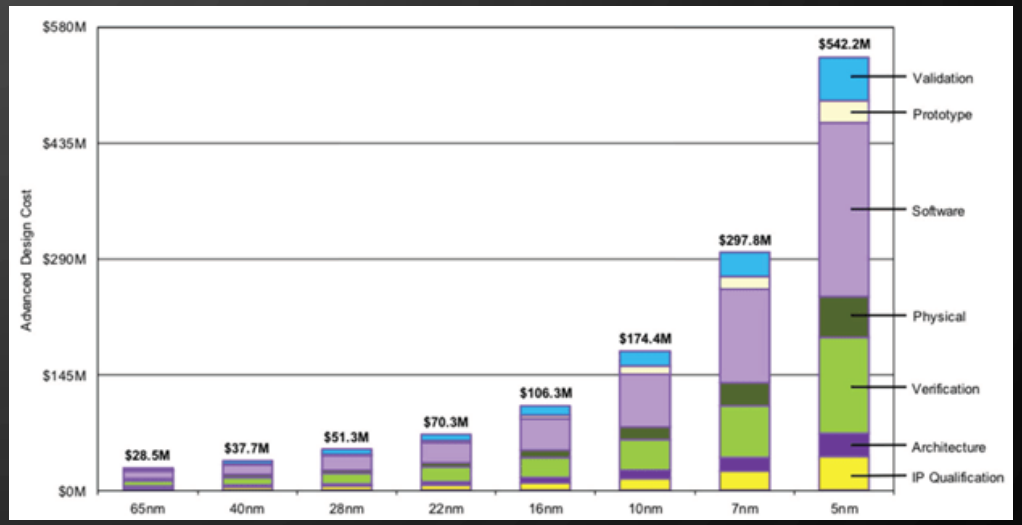
PROTOCOLLI DI INTERCONNESSIONE

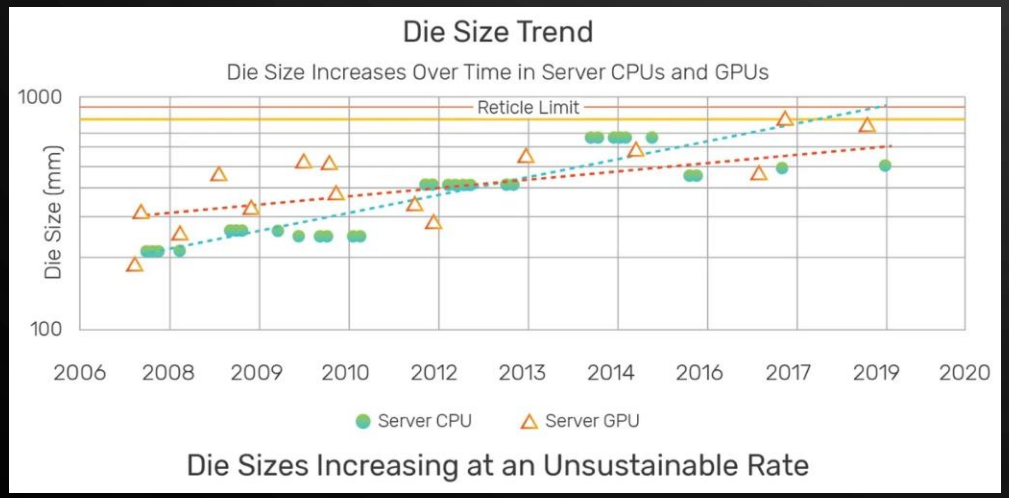
CORSO DI CALCOLATORI ELETTRONICI M

A CURA DI UMBERTO LAGHI - 1092993

SITUAZIONE ATTUALE – PROBLEMA

- Aumento funzionalità dei chip
- Aumento costo dei PP
- Problemi di scaling

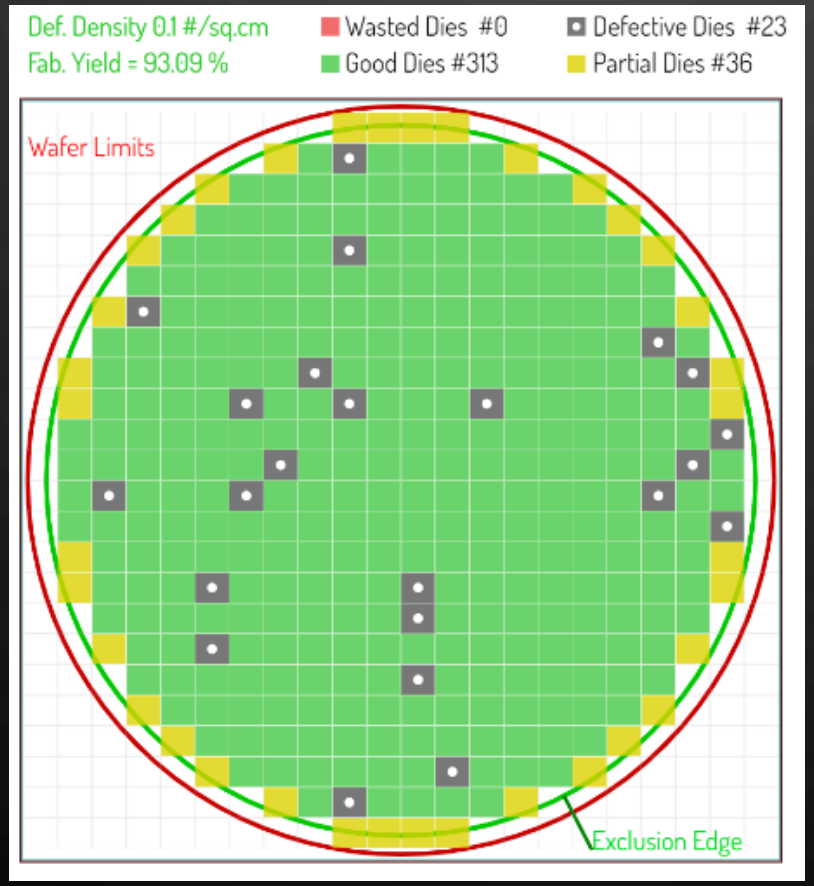




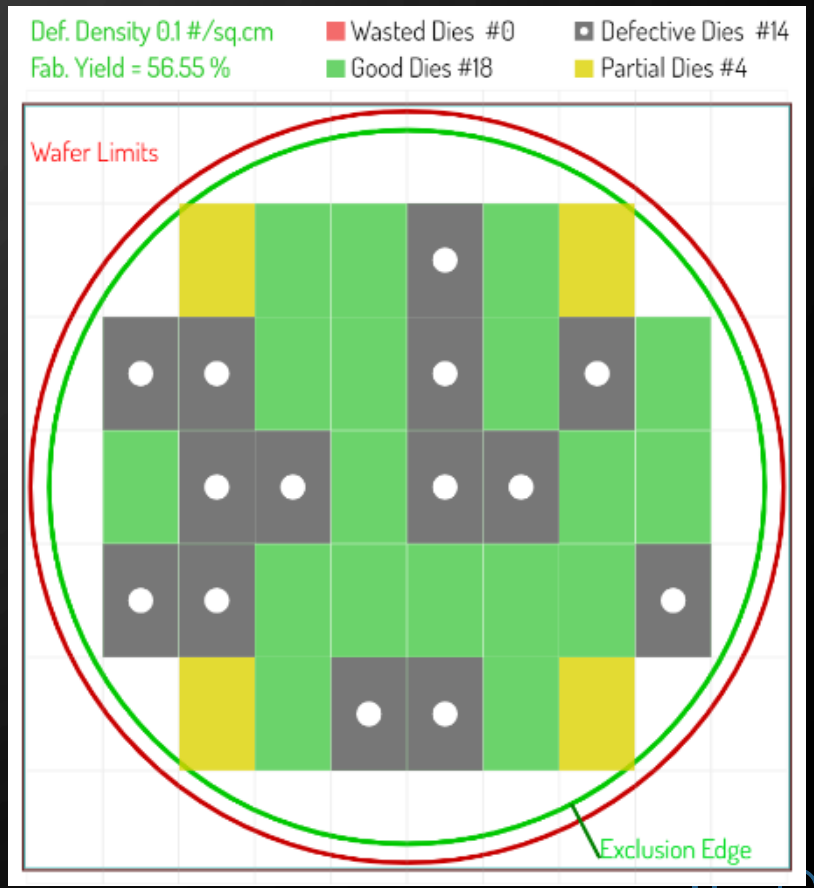
SITUAZIONE ATTUALE – CONSEGUENZE

- Resa minore
- Costi più alti

72 mm2

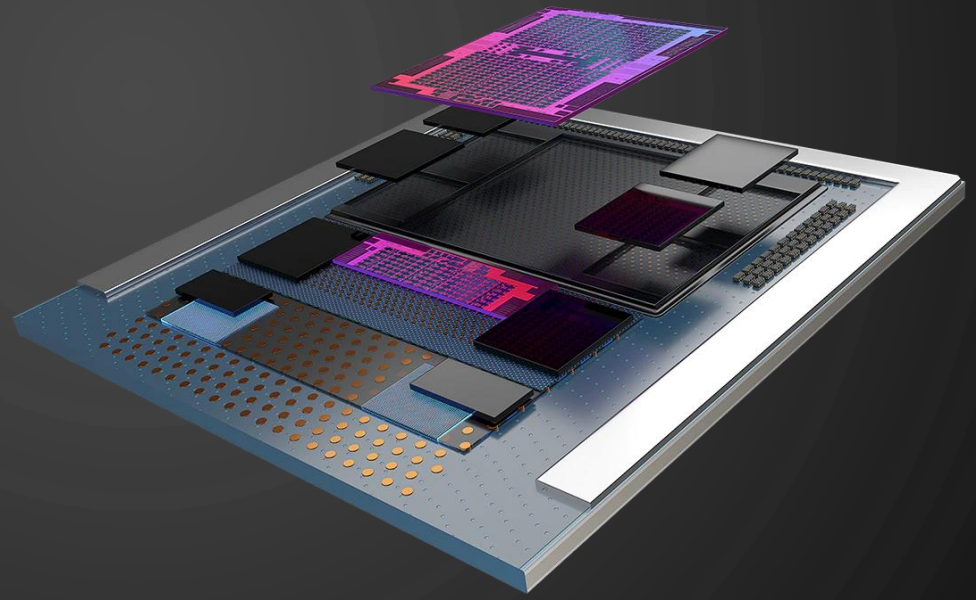


600 mm2

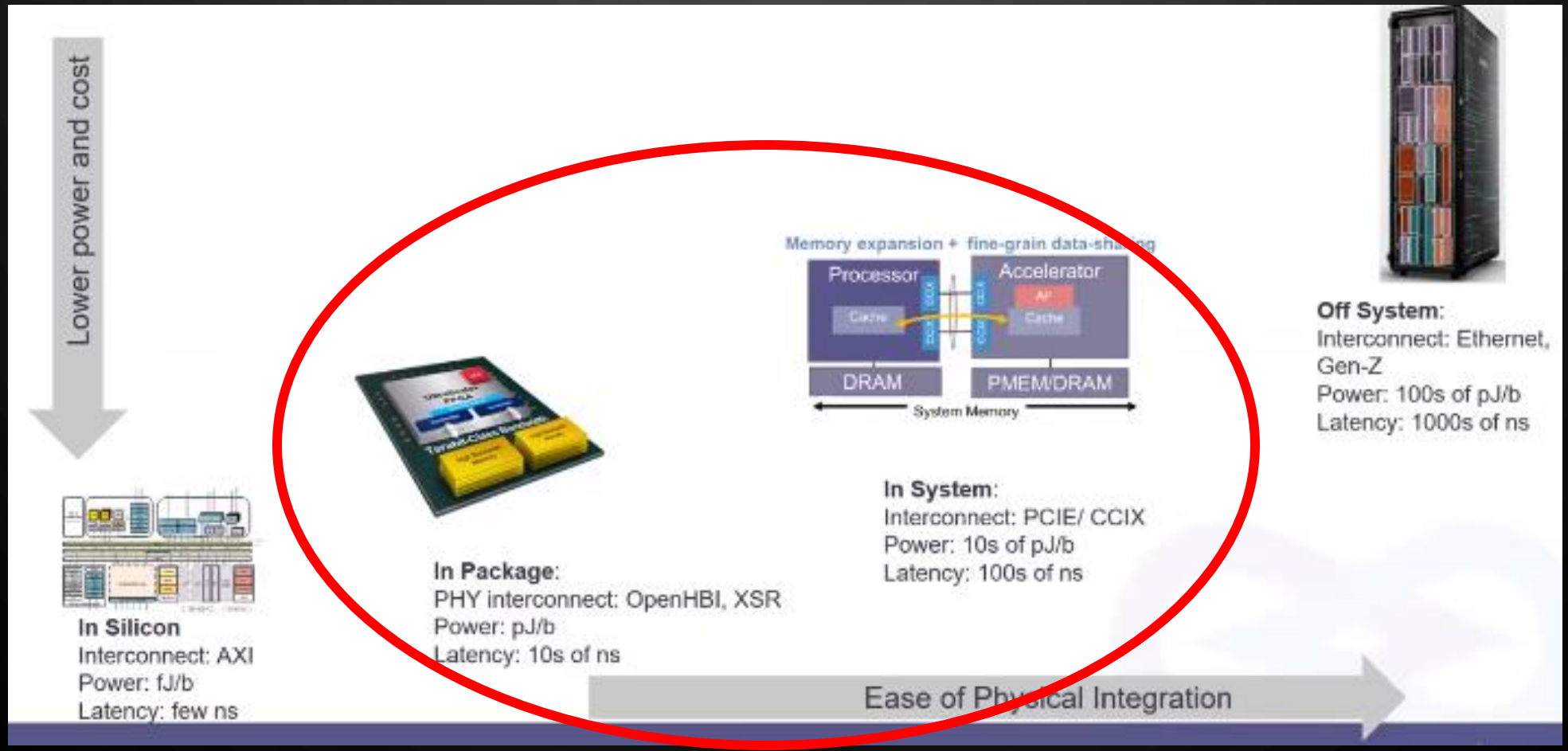


SOLUZIONE

- *Day of Reckoning* ipotizzato da Moore
- Collegare più chiplet o chip specifici



LIVELLI DI INTERCONNESSIONE



IDEE COMUNI

- Sfruttare infrastruttura già esistente
- Protocollo open
- Bassa latenza
- Alto bandwidth
- Power efficient
- Conveniente
- Astrazione a livello HW
- Cache coherent

PROTOCOLLI ANALIZZATI

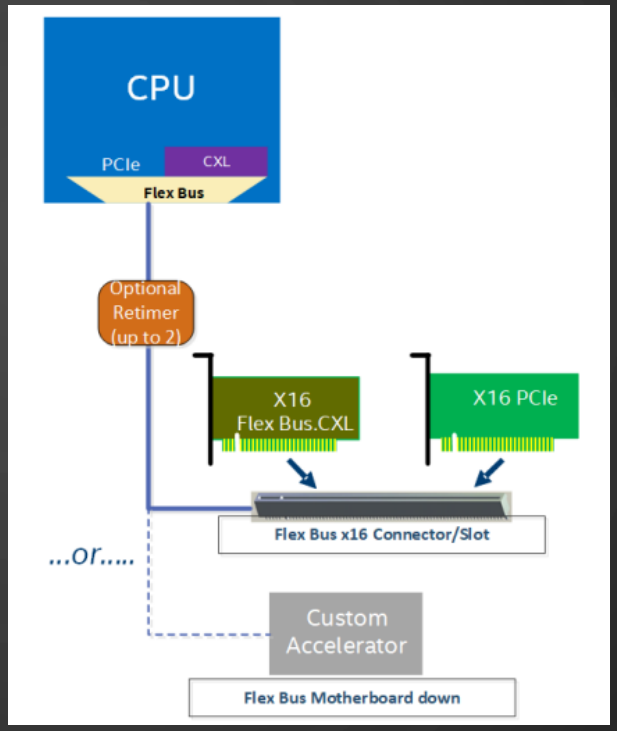
In package:

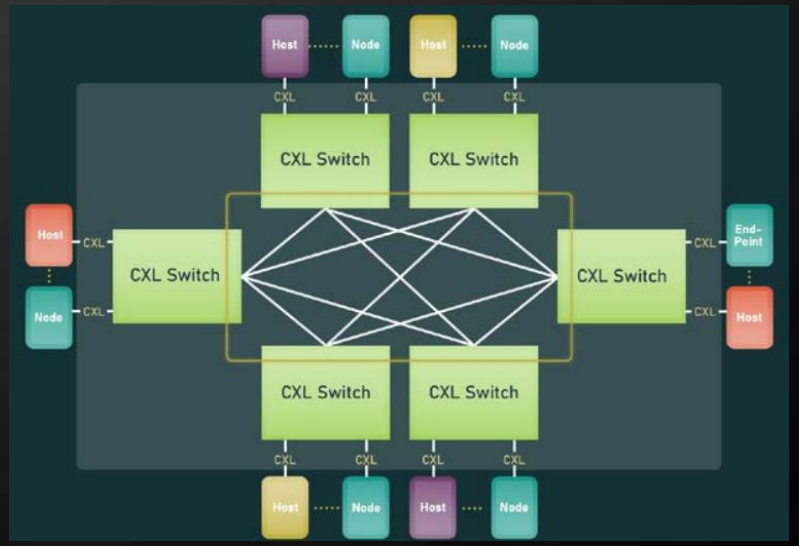
dalla versione 2.0

In system:

CXL 3.0

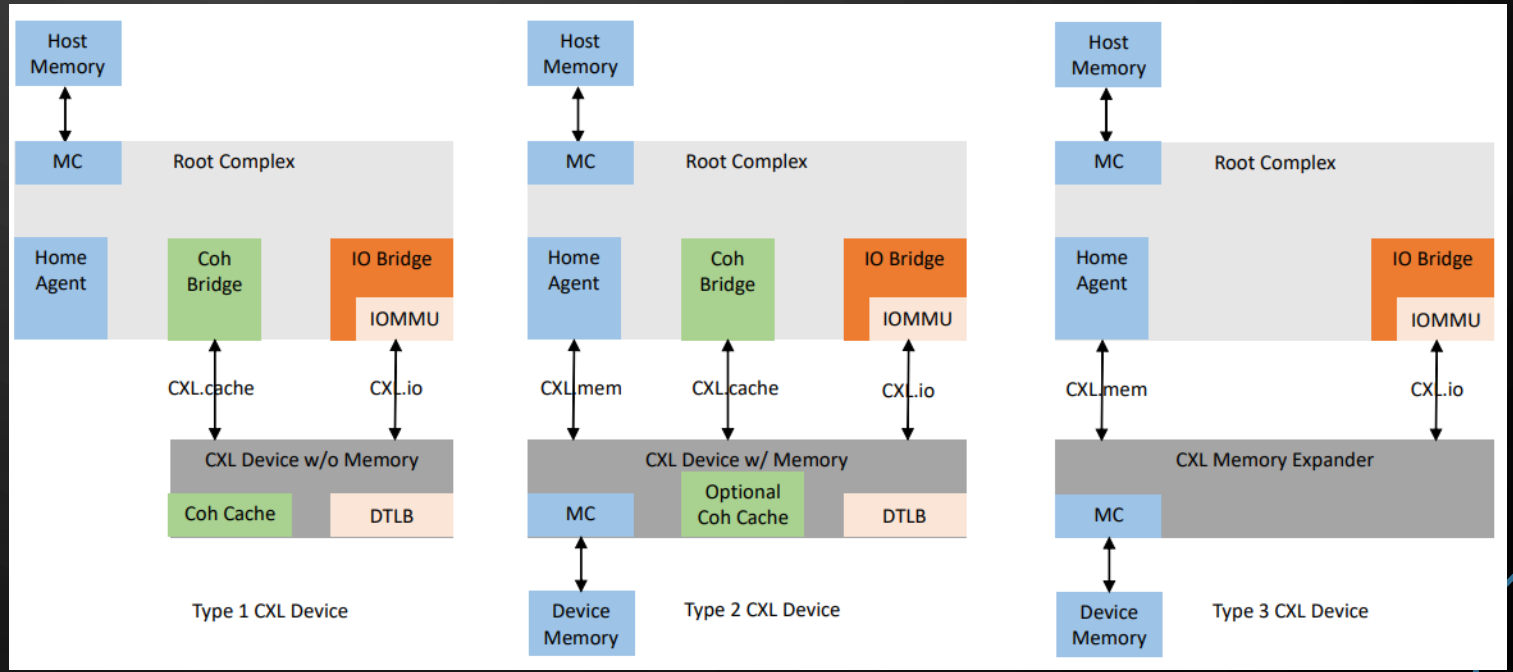
- Connessione tra CPU x86 e acceleratori
- Coerenza tra la memoria della CPU e quella dell'acceleratore
- Basato su PCIe 6.0
- Bandwidth teorico massimo di 256GB/s
- Topologia *fabrics*, non gerarchica





SOTTO-PROTOCOLLI E DISPOSITIVI

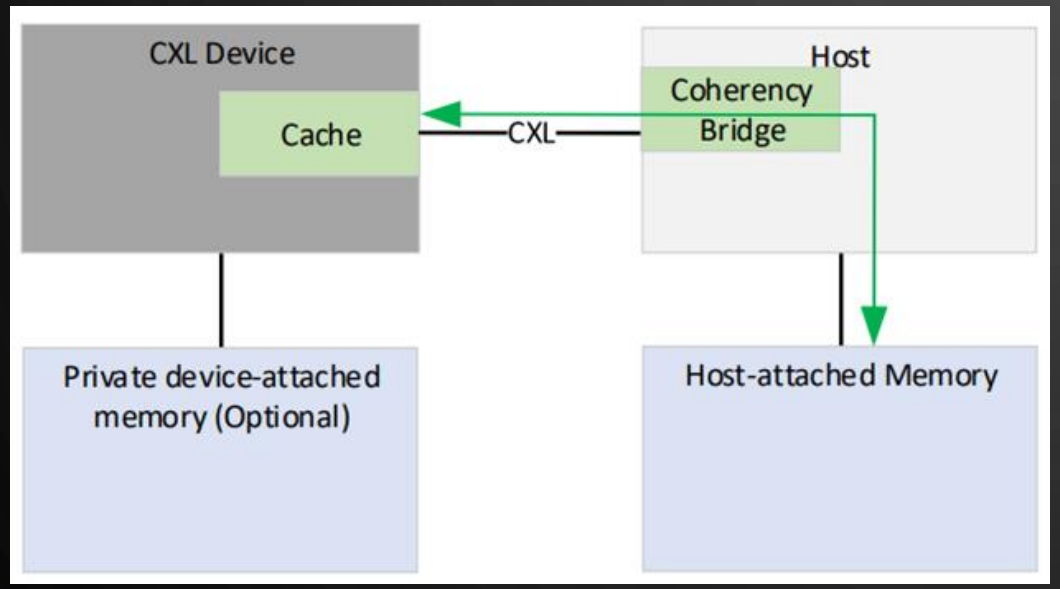
- Sotto-protocolli:
  - CXL.io
  - CXL.cache
  - CXL.memory (CXL.mem)
- Dispositivi:
  - Tipo 1
  - Tipo 2
  - Tipo 3



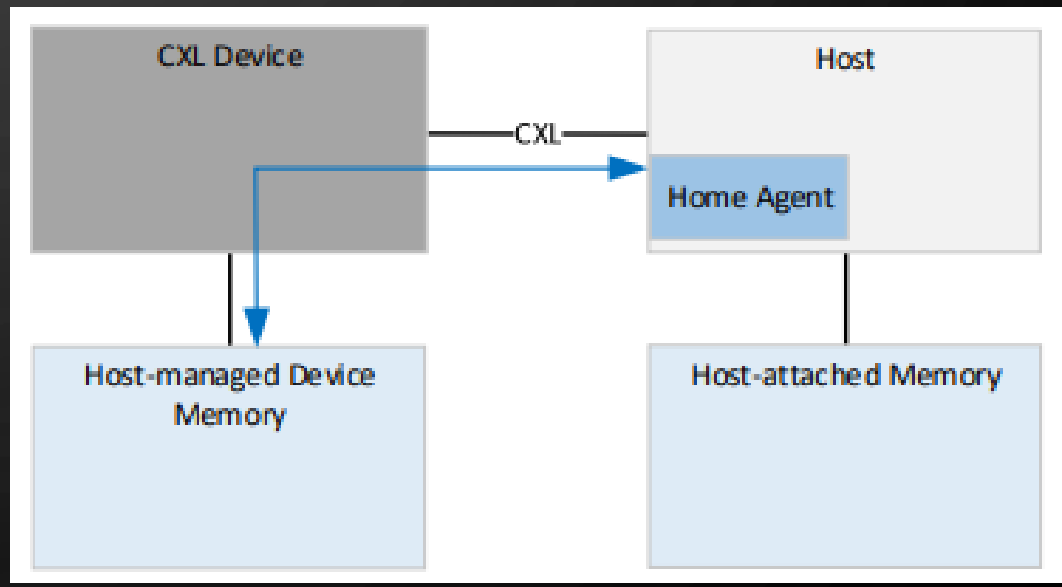
GESTIONE DELLA MEMORIA DEGLI ACCELERATORI

- Modello Master-Subordinate
- Gestita dall'host (CPU) → Host-Managed Device Memory (HDM)
- Modelli di coerenza memoria dell'acceleratore:
  - Host-only coherent (HDM-H)
  - Device coherent (HDM-D)
  - Host coherent using Back-Invalidation Snoop (HDM-DB)

TYPE 1 AND 3 DEVICES



Tipo 1 – Coerenza garantita con meccanismo di snooping

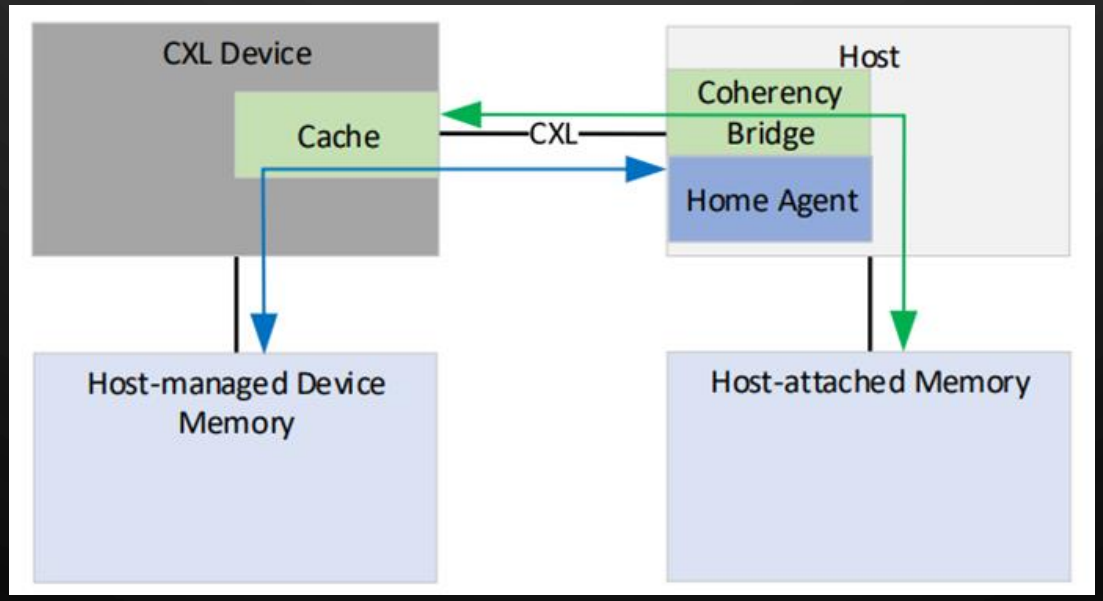


Tipo 3 – Espansione passiva di memoria gestita con HDM-H

TYPE 2 DEVICES

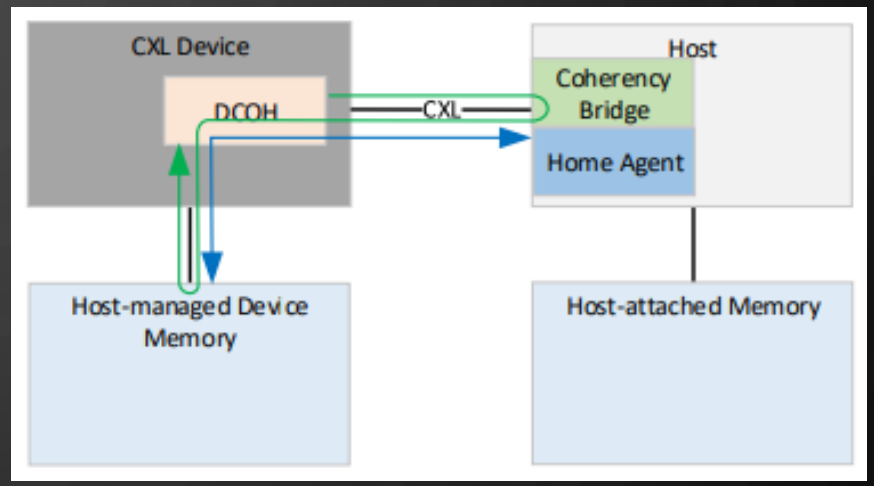
Coerenza garantita in due modi:

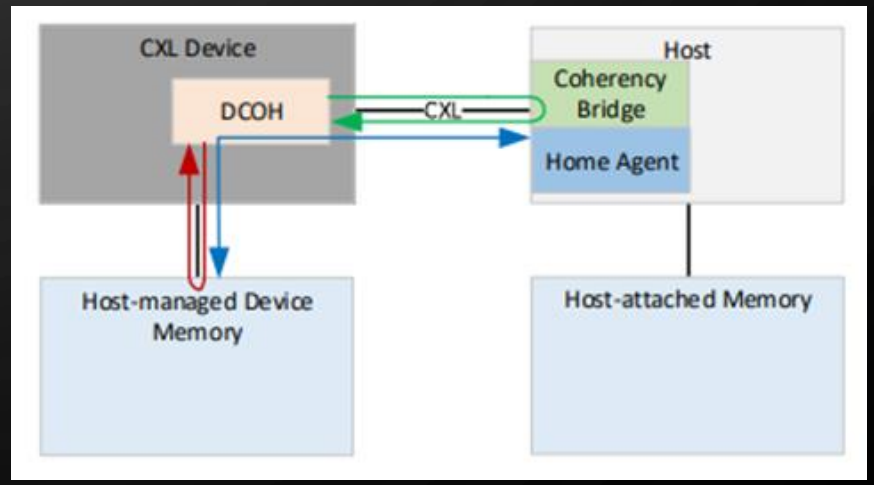
- HDM-D → Bias-Based coherency model
- HDM-DB → Back-Invalidate Snoop coherency model



BIAS-BASED COHERENCY

- Usa CXL.cache
- Due stati per la memoria del dispositivo:
  - Host Bias
  - Device Bias
- Minore Latenza

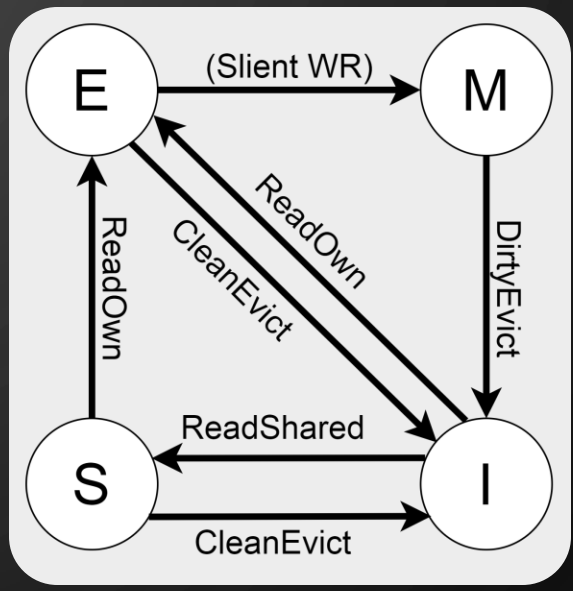


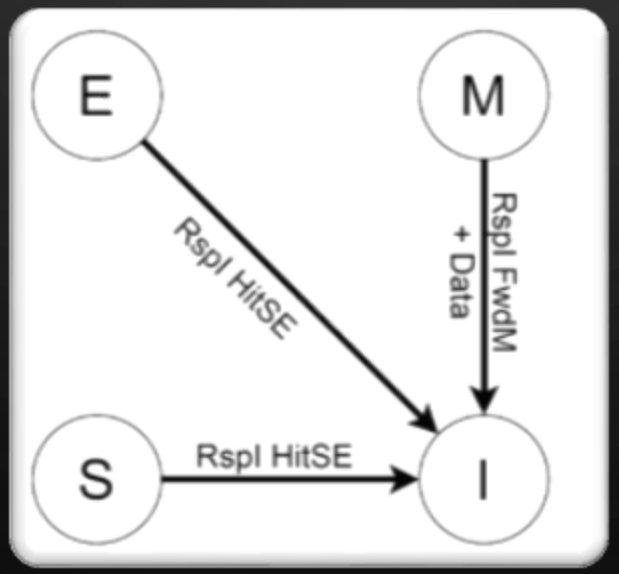


BACK-INVALIDATION SNOOP – 1

- Protocollo MESI
- Concetto di snoop
- Usa CXL.memory

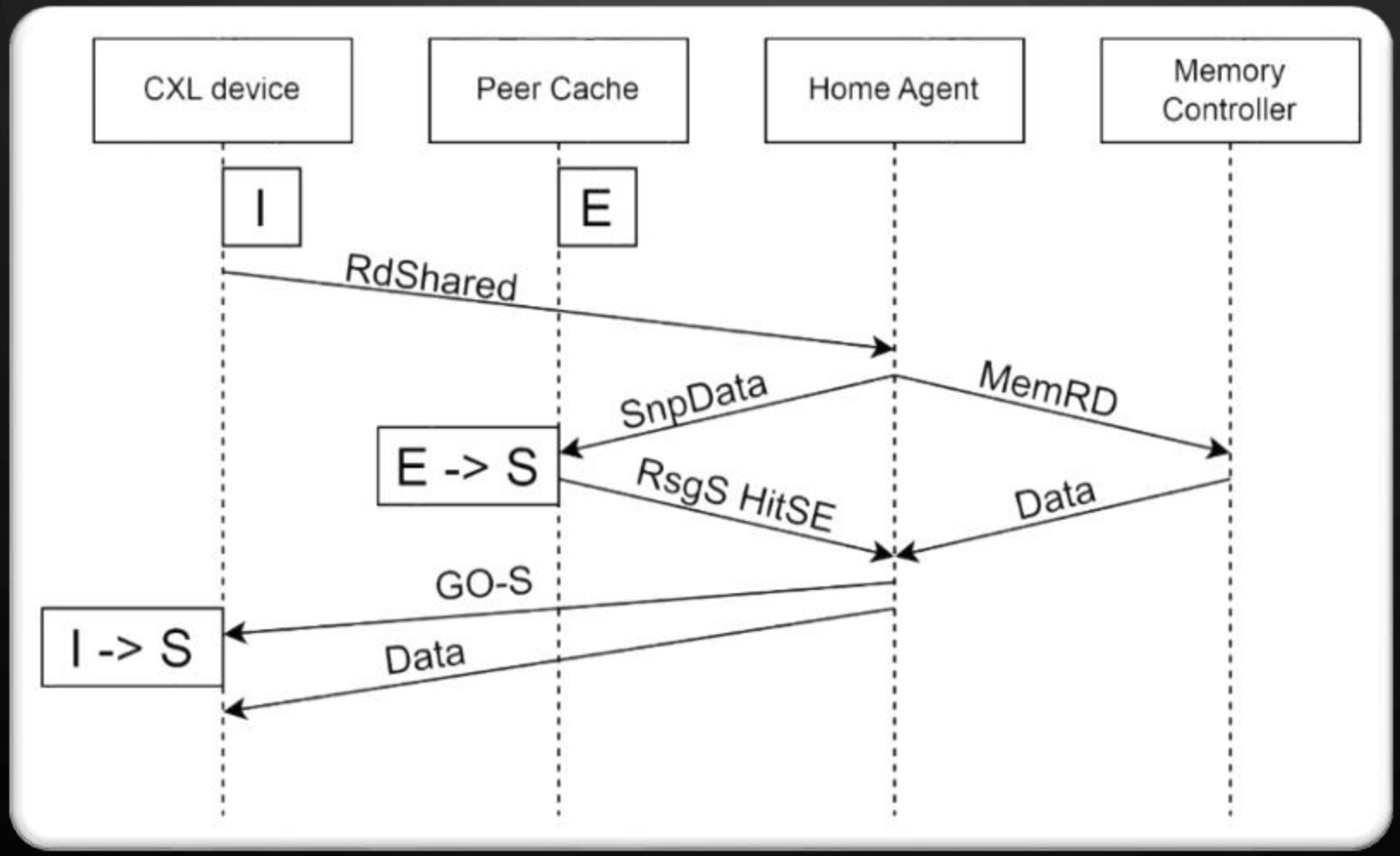
Richieste del dispositivo





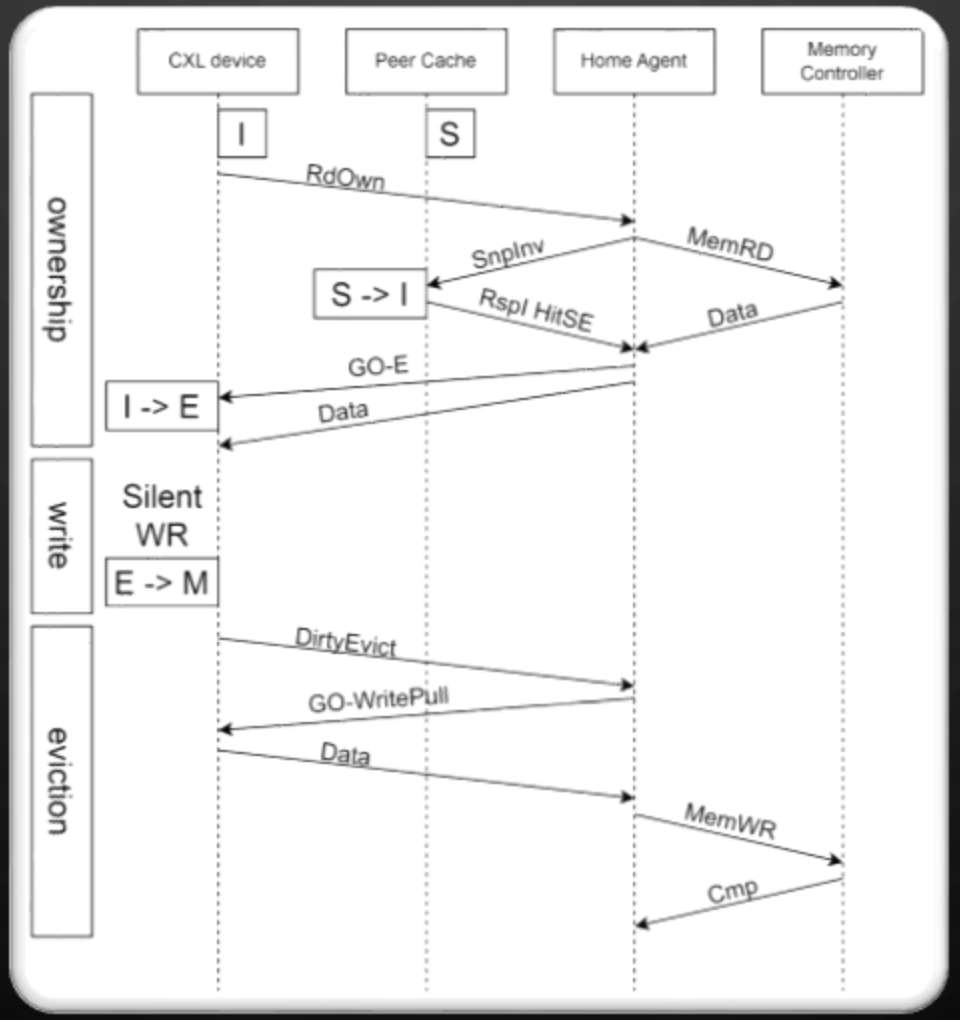
Cambio di stato in seguito a ricezione di SnpInv. Le risposte sono inviate all'Home Agent

BACK-INVALIDATION SNOOP – 2



Flusso di RD del dispositivo

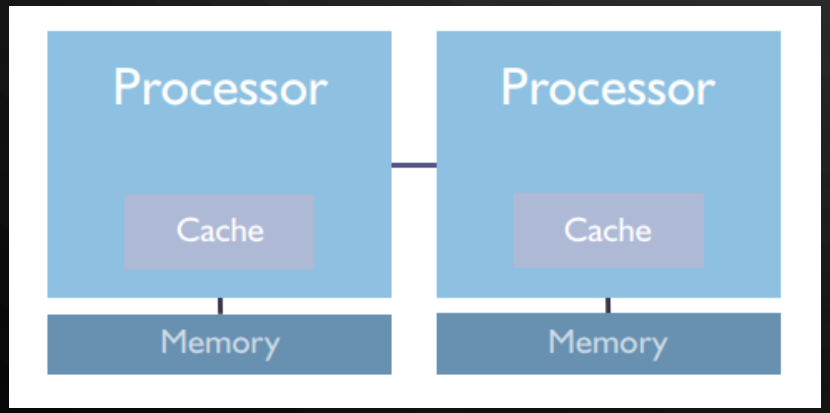
BACK-INVALIDATION SNOOP – 3

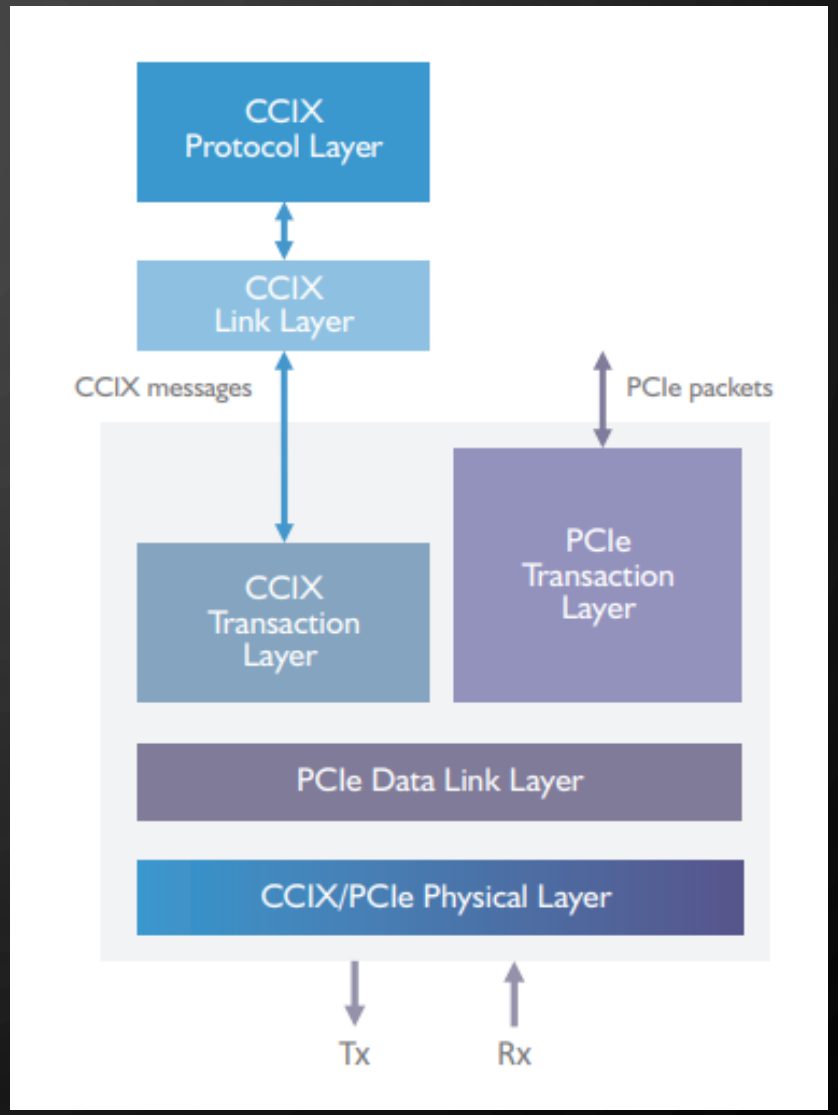


Flusso di WR del dispositivo

CCIX «see 6»

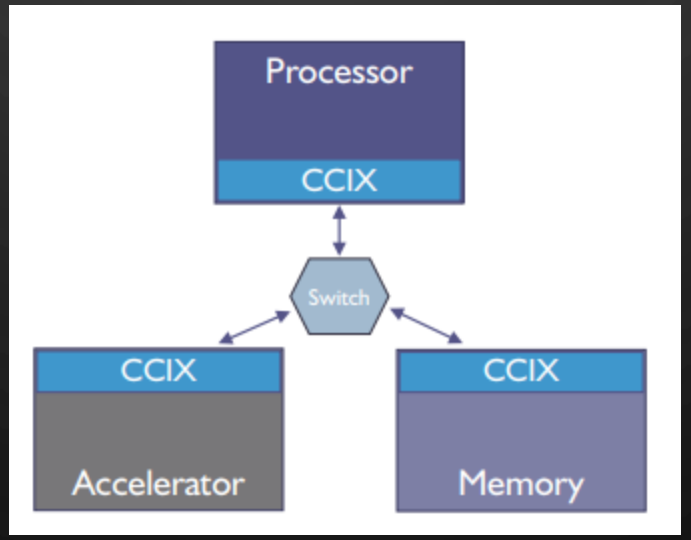
- Coerenza tra memorie dei chip
- Chip come peer → simmetrico
- Più livelli
- Agnostico al protocollo

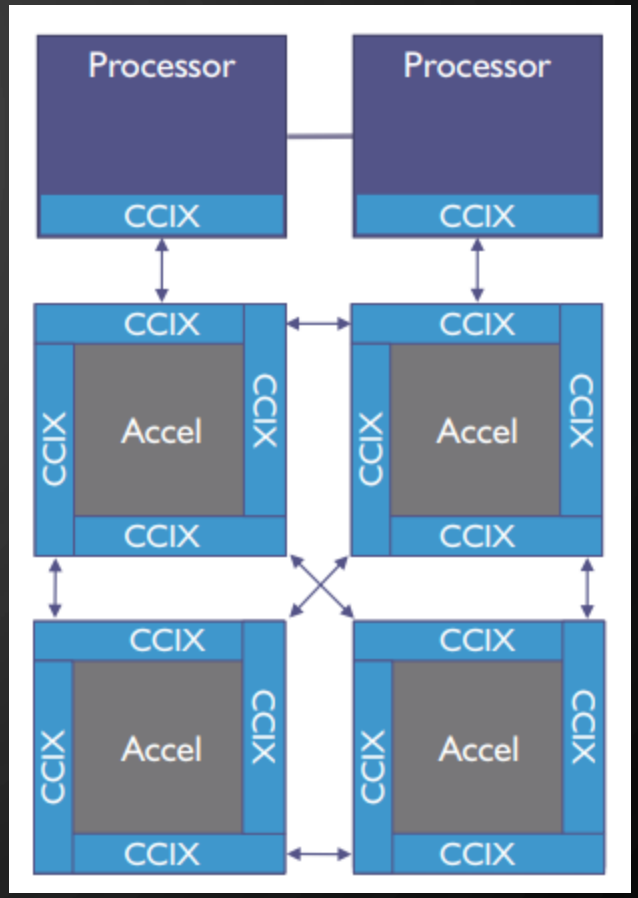




AGENTI E TOPOLOGIE

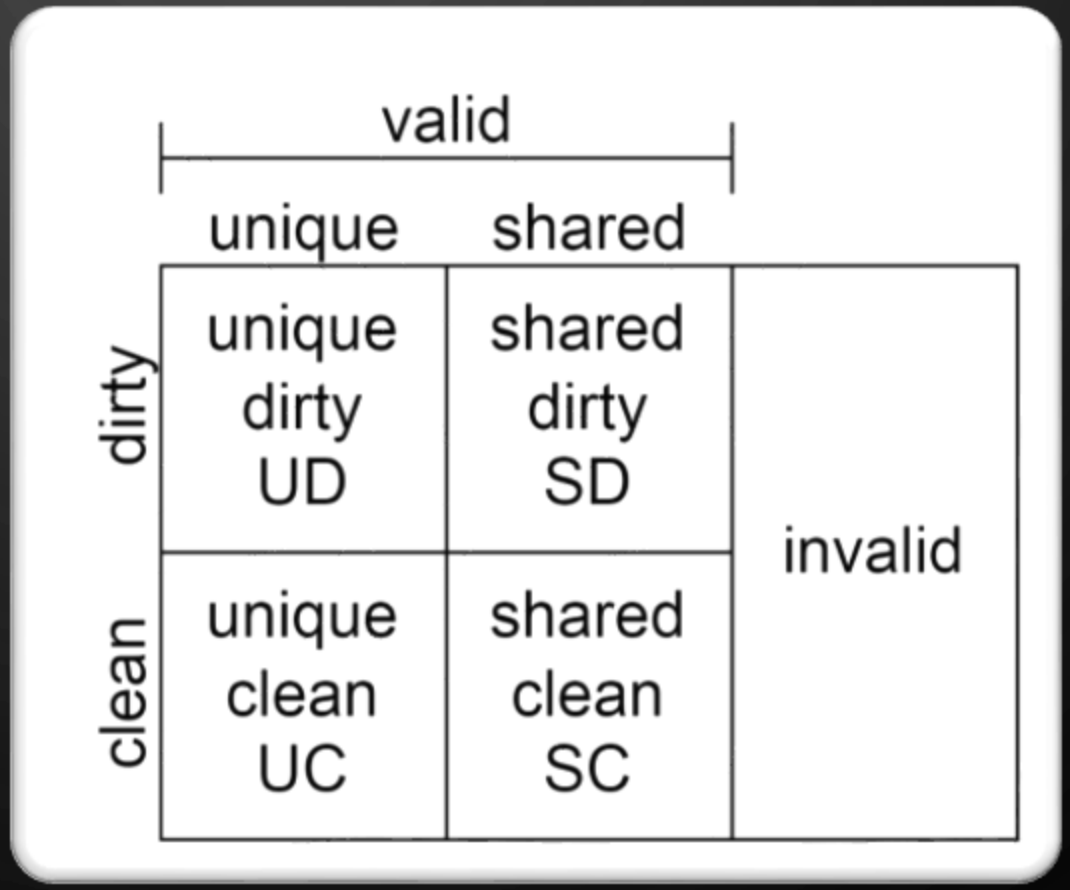
- Si occupano di orchestrare la coerenza
- Request Agent
- Home Agent
- Slave Agent
- Error Agent
- Più topologie possibili





PROTOCOLLO DI COERENZA

- Snoop
- Non-MESI

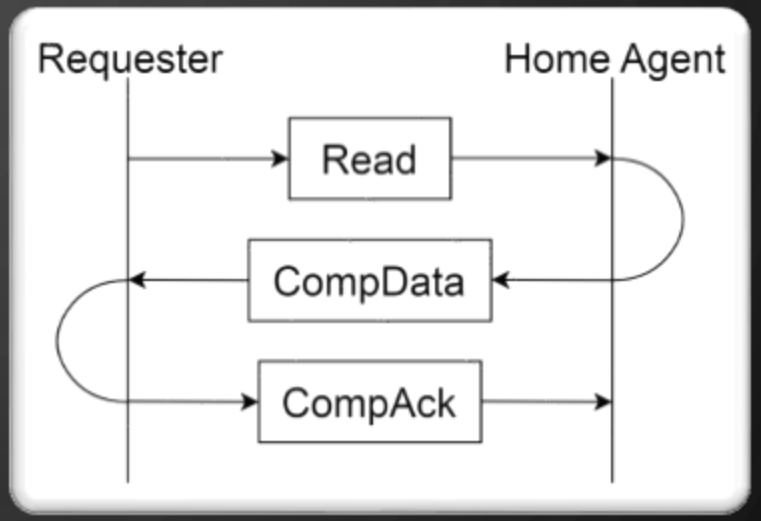


REQUEST AGENT

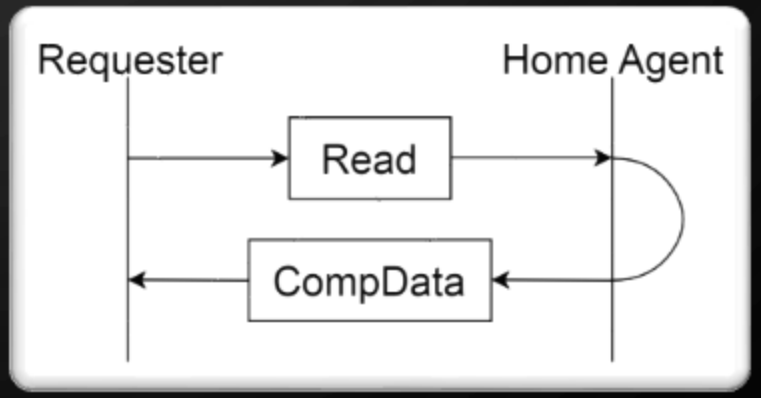
Tipi di richieste:

- Read: coerenti o non coerenti
- Write: coerenti o non coerenti
- Atomic
- Dataless

Read coerente

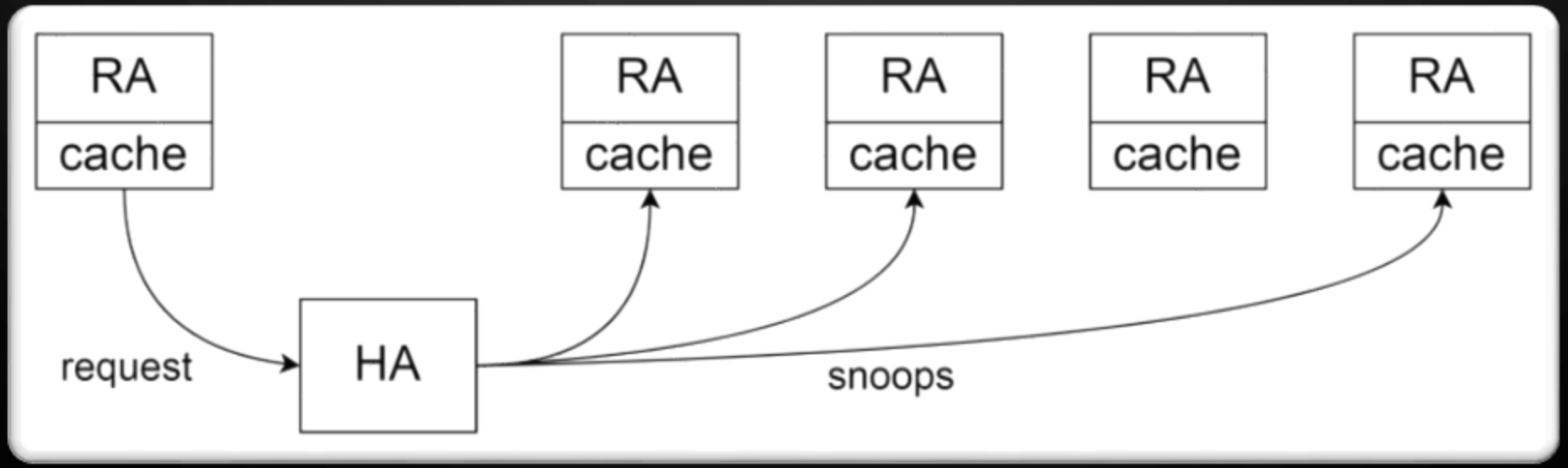


Read non coerente



HOME AGENT

- Controlla tutte le cache del sistema quando c'è una lettura dalla memoria
- Ha una struttura dati per capire quale delle RA ha la linea interessata
- Invia gli snoop



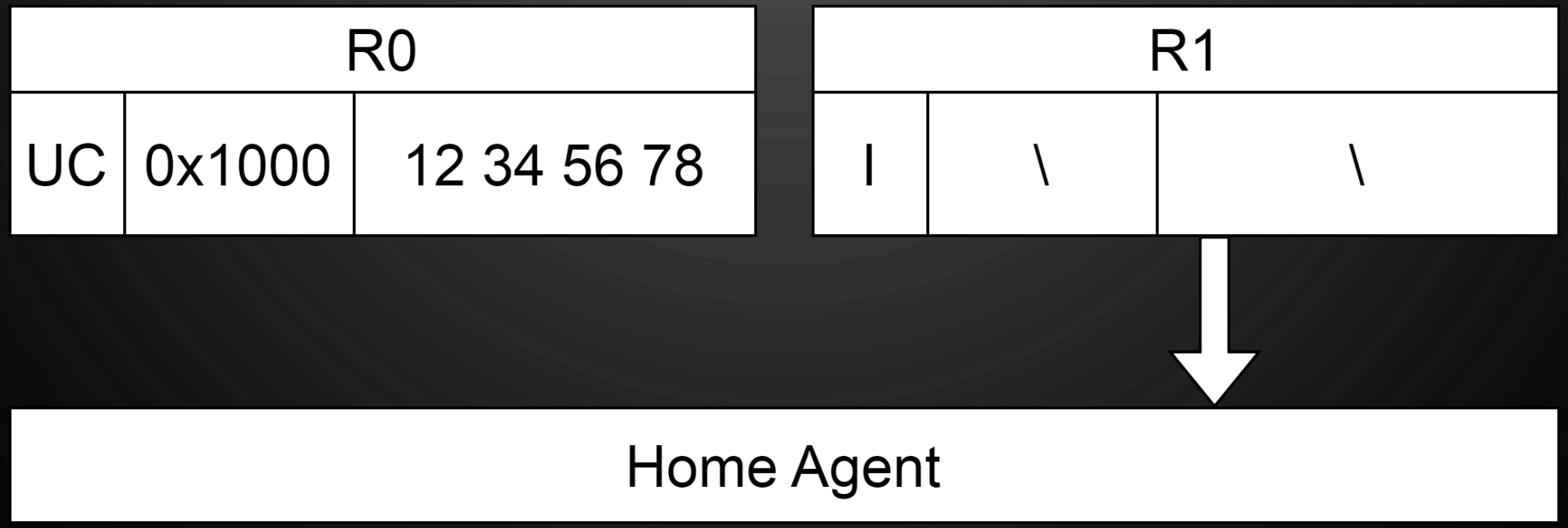
ESEMPIO

R0		
UC	0x1000	12 34 56 78

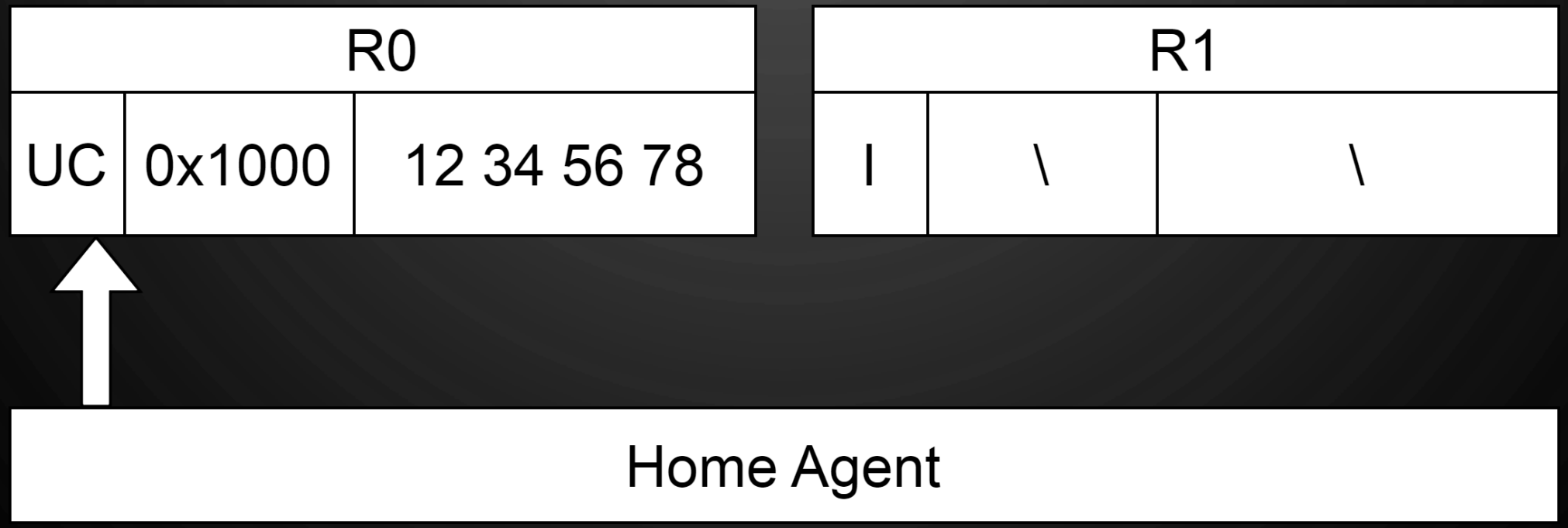
R1		
I	\	\

Home Agent

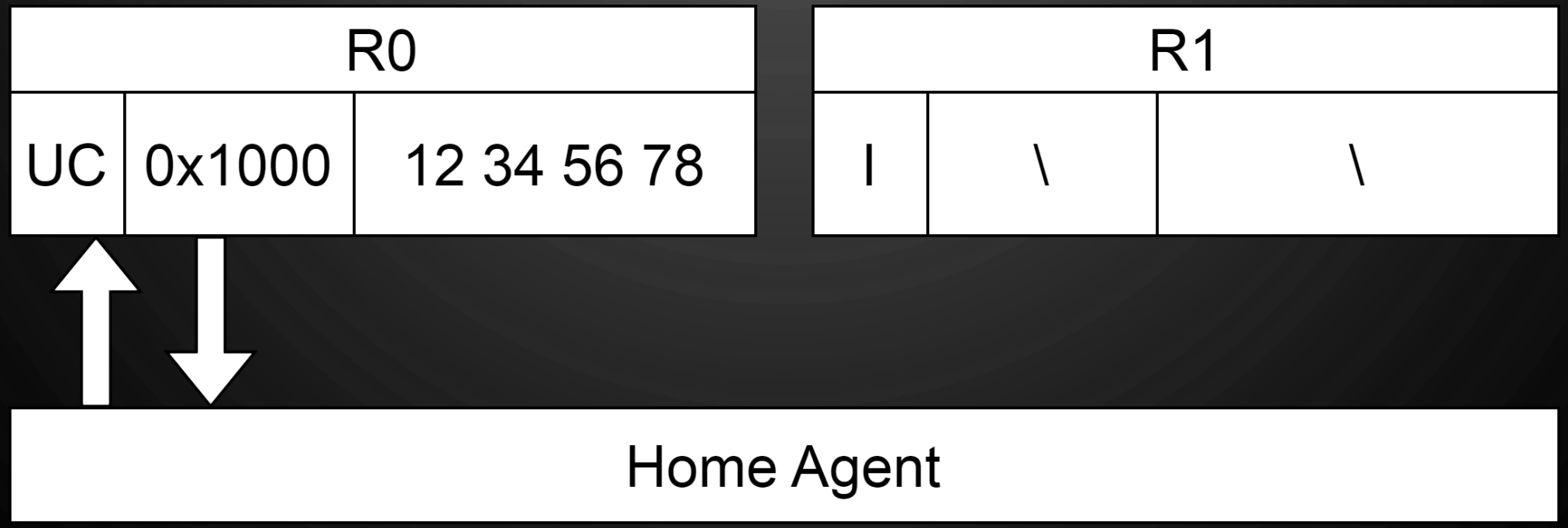
ESEMPIO



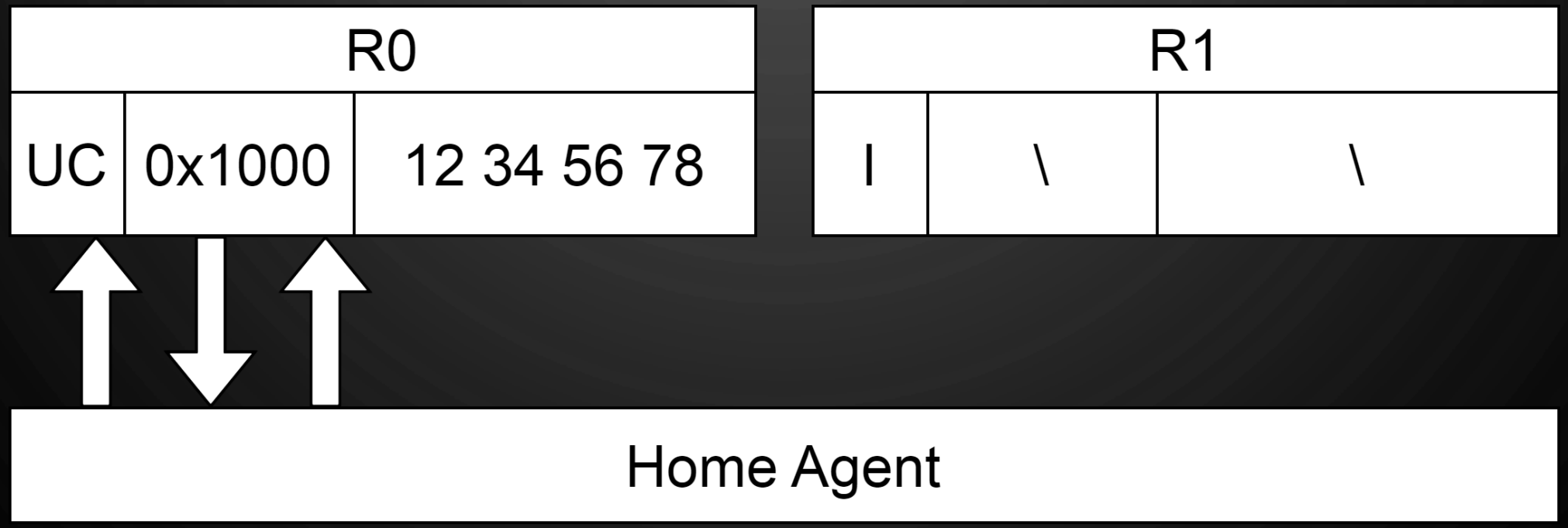
ESEMPIO



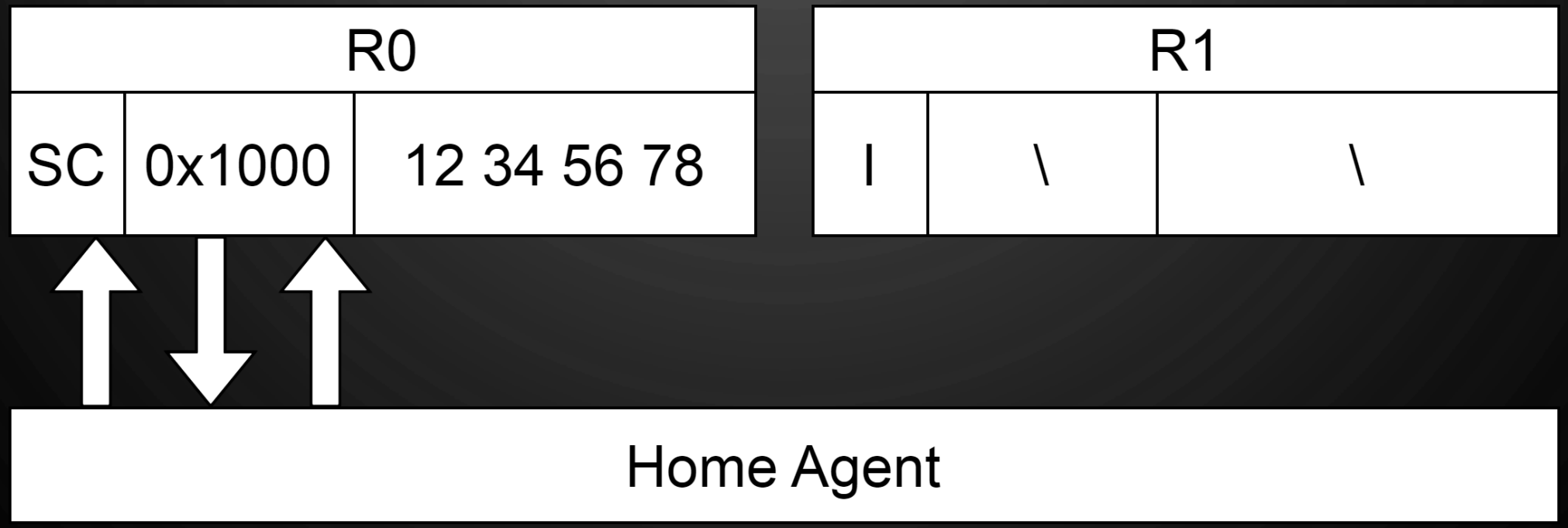
ESEMPIO



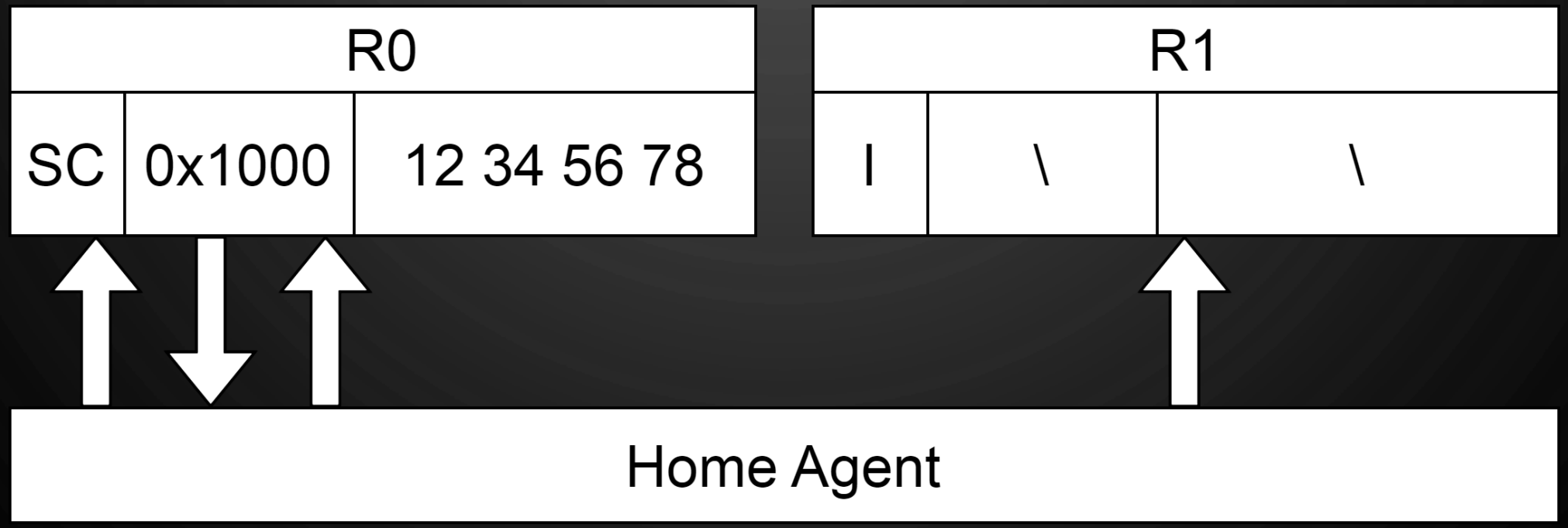
ESEMPIO



ESEMPIO



ESEMPIO



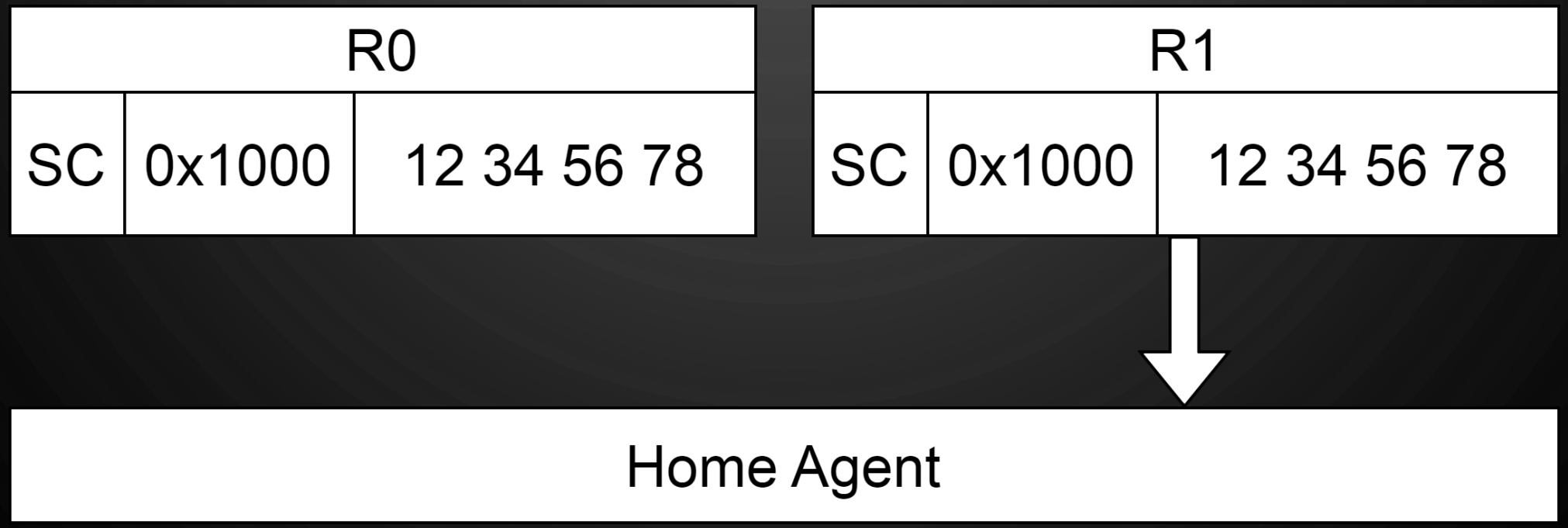
ESEMPIO

R0		
SC	0x1000	12 34 56 78

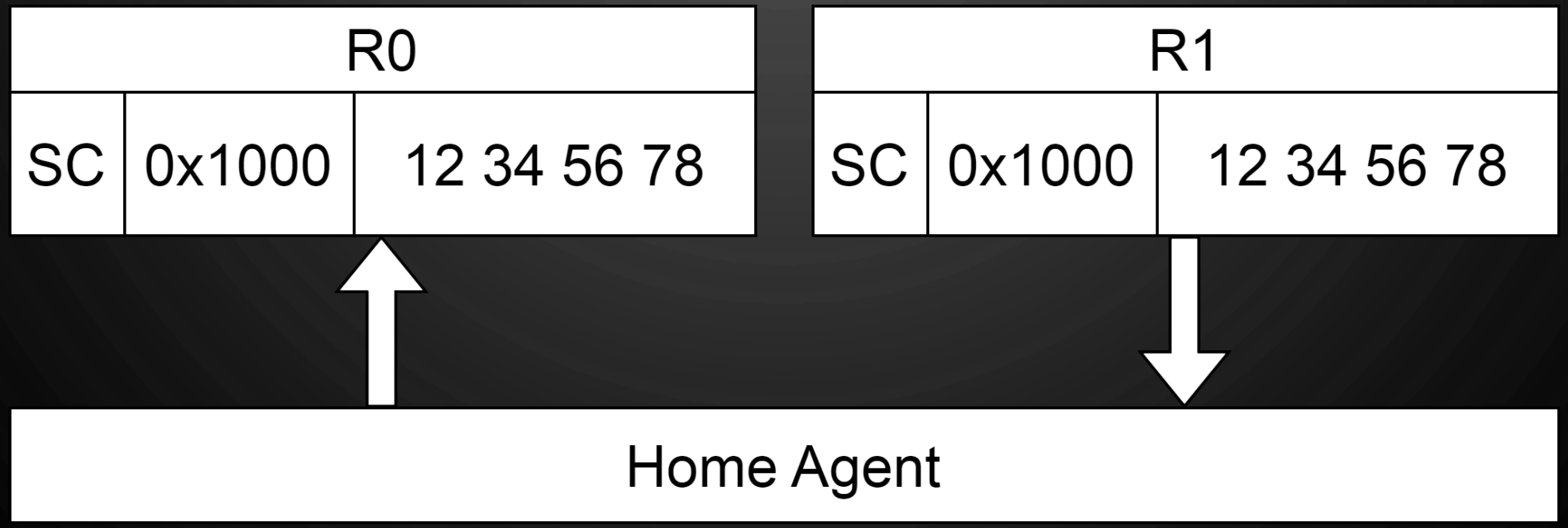
R1		
SC	0x1000	12 34 56 78

Home Agent

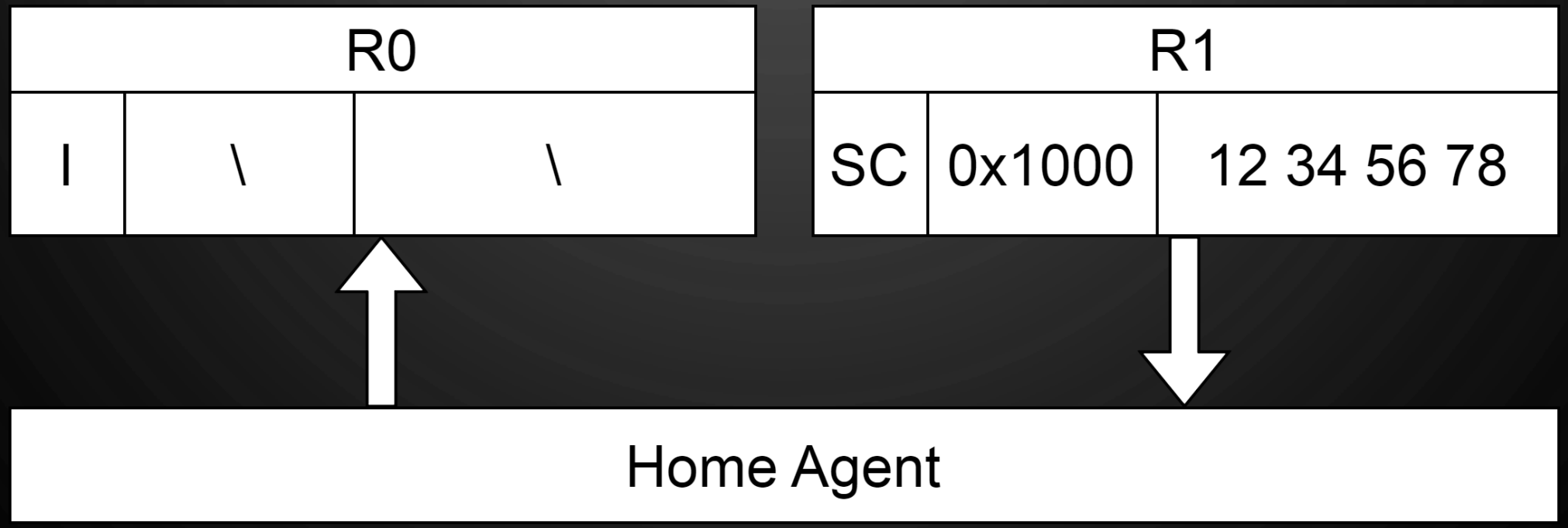
ESEMPIO



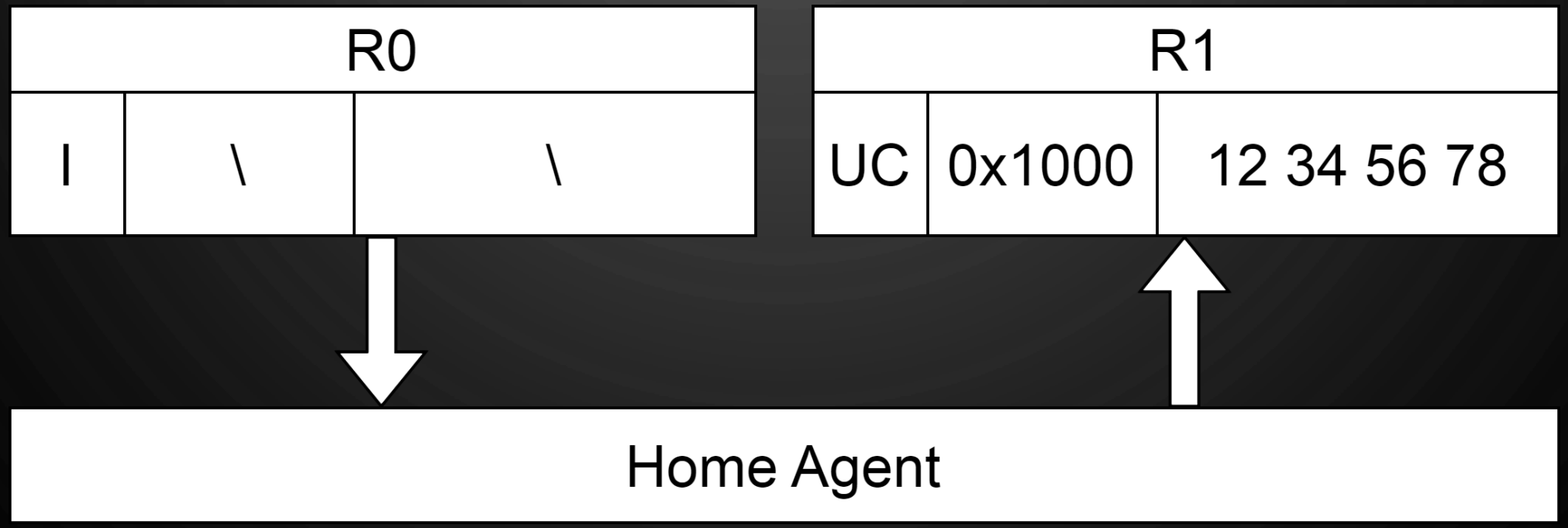
ESEMPIO



ESEMPIO



ESEMPIO



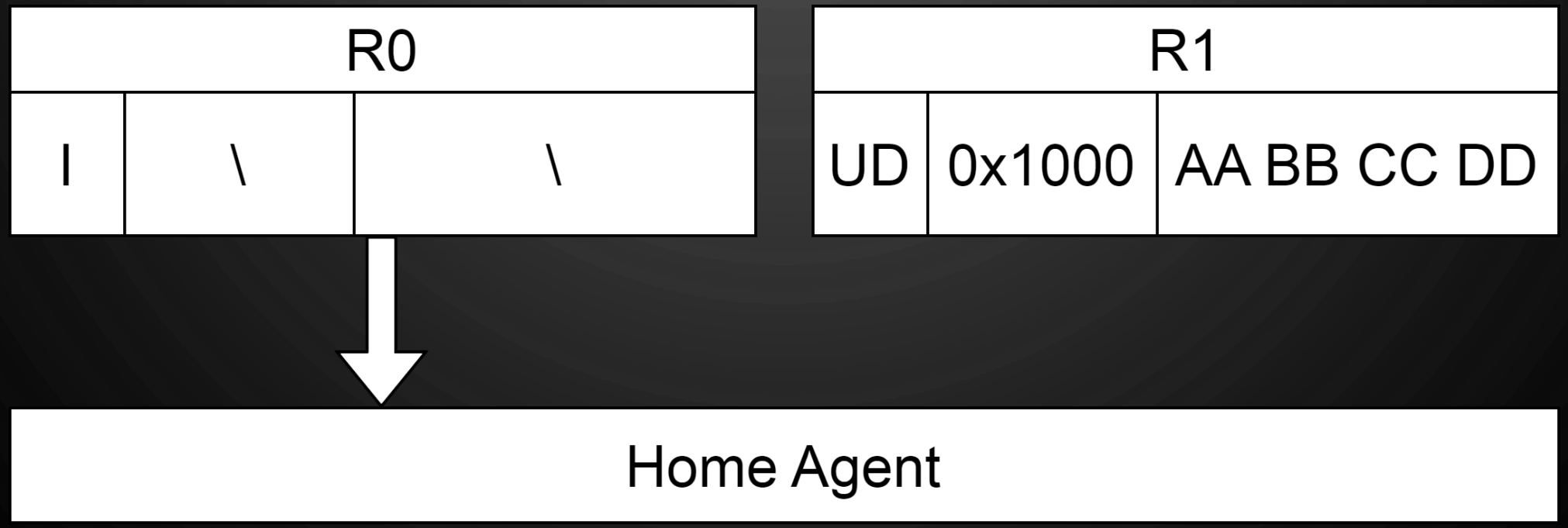
ESEMPIO

R0		
I	\	\

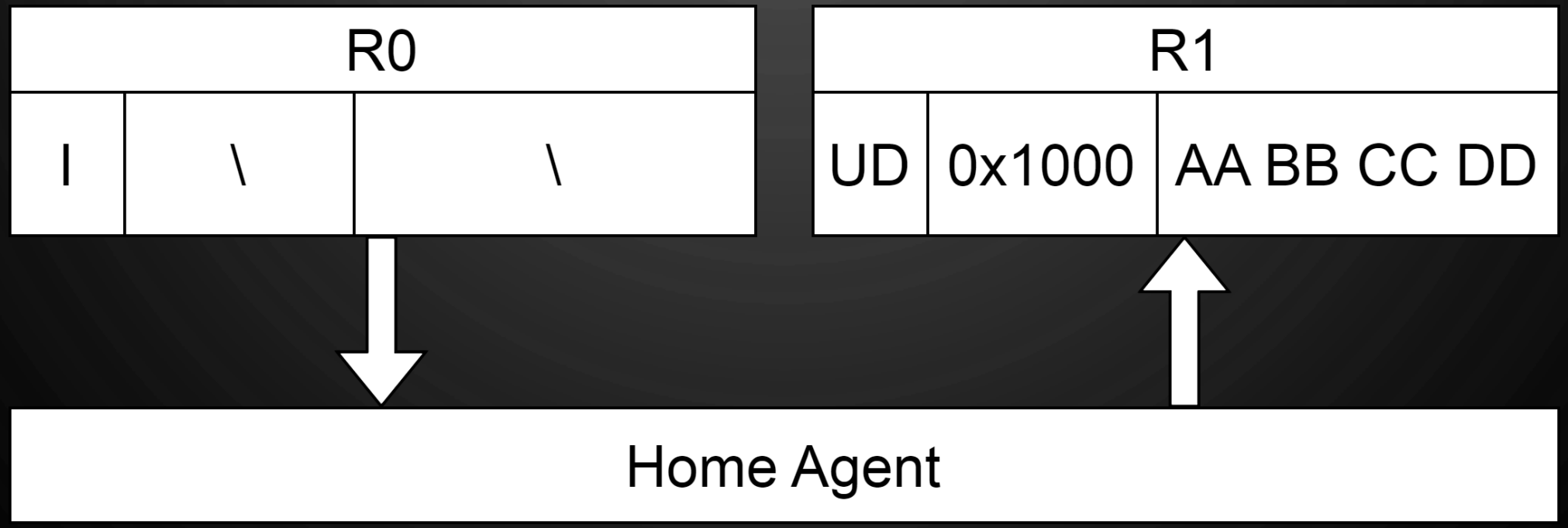
R1		
UD	0x1000	AA BB CC DD

Home Agent

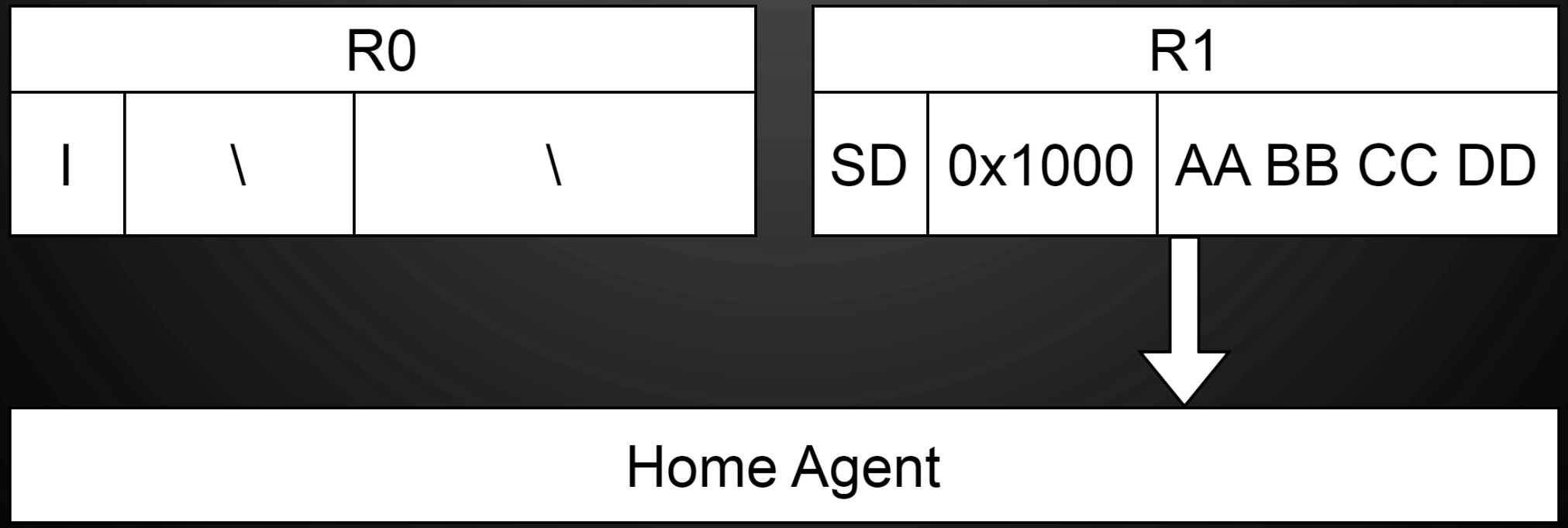
ESEMPIO



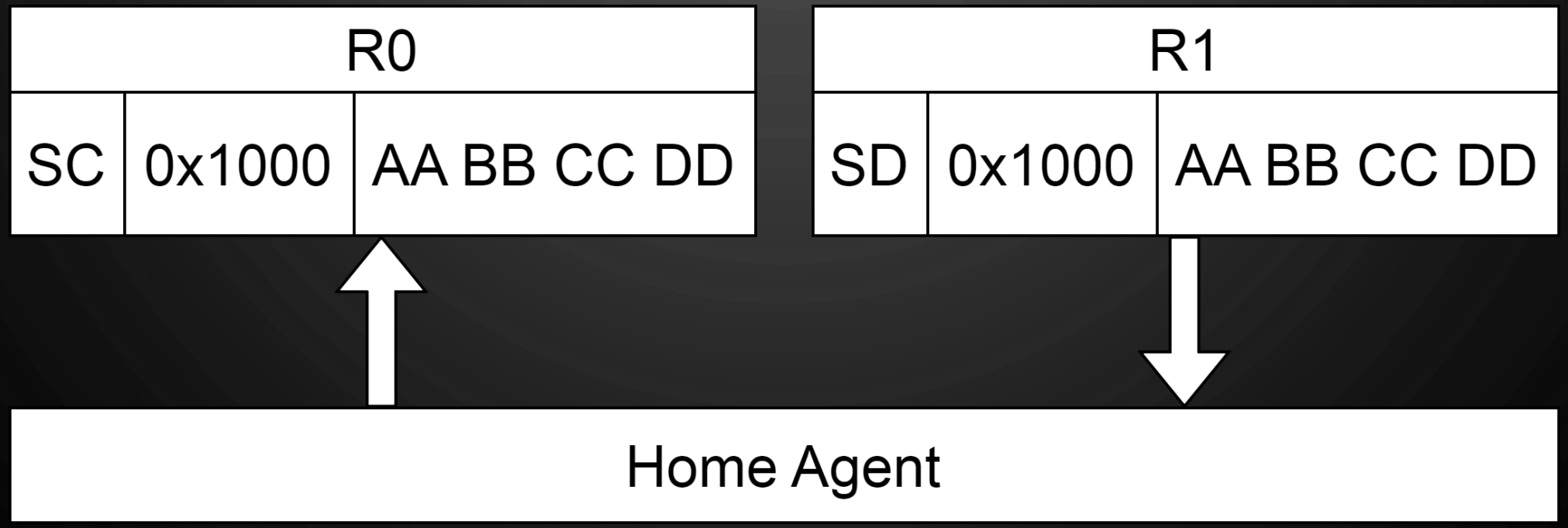
ESEMPIO



ESEMPIO

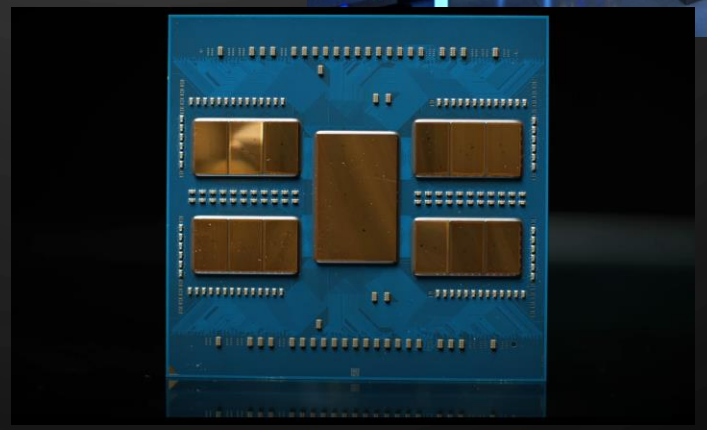


ESEMPIO



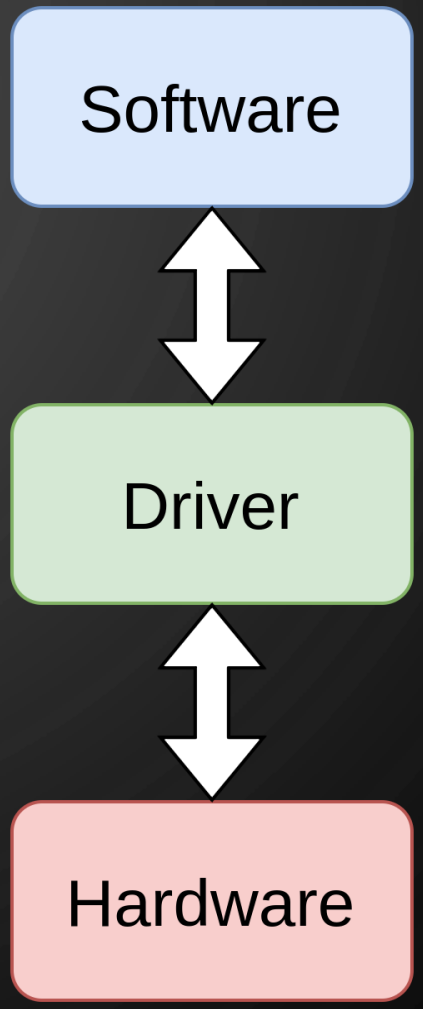
CXL VS CCIX

- Modello
- Dispositivi connessi
- Usecase
- Consumo energetico



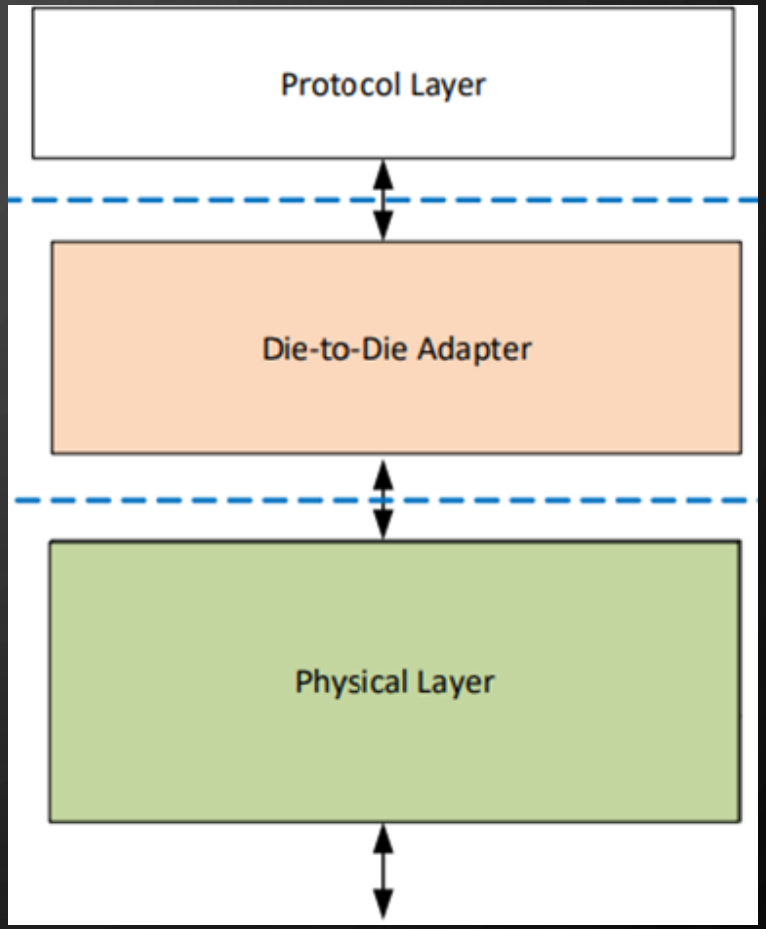
OBIETTIVO DI CXL & CCIX

- Gestione a livello HW della coerenza
- Memoria unificata
- Trasparenza per il SW



UCIe

- Connessione in package
- Mappa altri protocolli
- Protocollo stacked



PROTOCOL LAYER

- Può mappare:
  - CXL 2.0 e 3.0
  - PCIe Gen5 e Gen6
  - Streaming Protocol
- Trasmette i dati a D2D Adapter attraverso FDI
- Modalità di mappatura:
  - FLIT
  - Raw

FLIT Mode

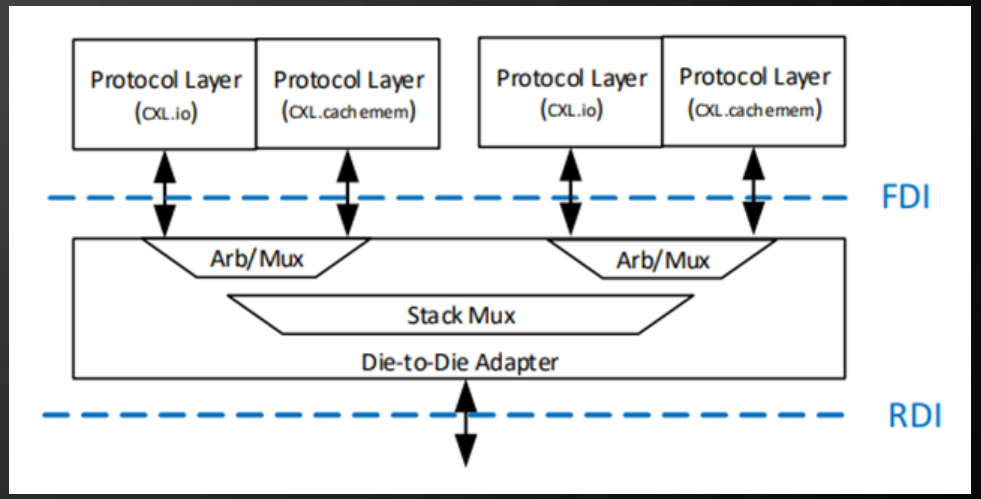
Byte								
0	Flit Hdr (Byte 0)	Flit Hdr (Byte 1)	Flit Chunk 0 62B (from Protocol Layer)					
64	Flit Chunk 1 58B (from Protocol Layer)					DLP Bytes 2:5	CRC0 (Byte 0)	CRC0 (Byte 1)
128	Flit Chunk 2 64B (from Protocol Layer)							
192	Flit Chunk 3 52B (from Protocol Layer)			TLP 4B (Optional)	2B Rsvd	Flit_Marker 4B	CRC1 (Byte 0)	CRC1 (Byte 1)

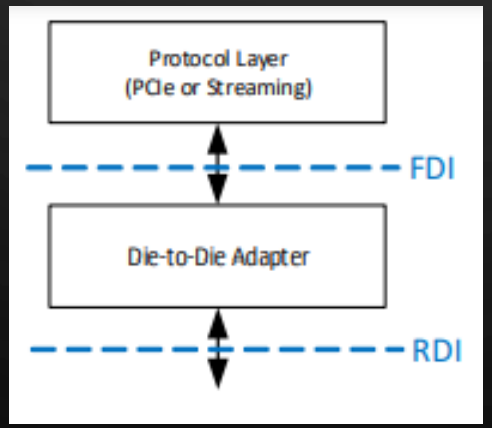
Raw Mode

Byte	
0	64B (from Protocol Layer)

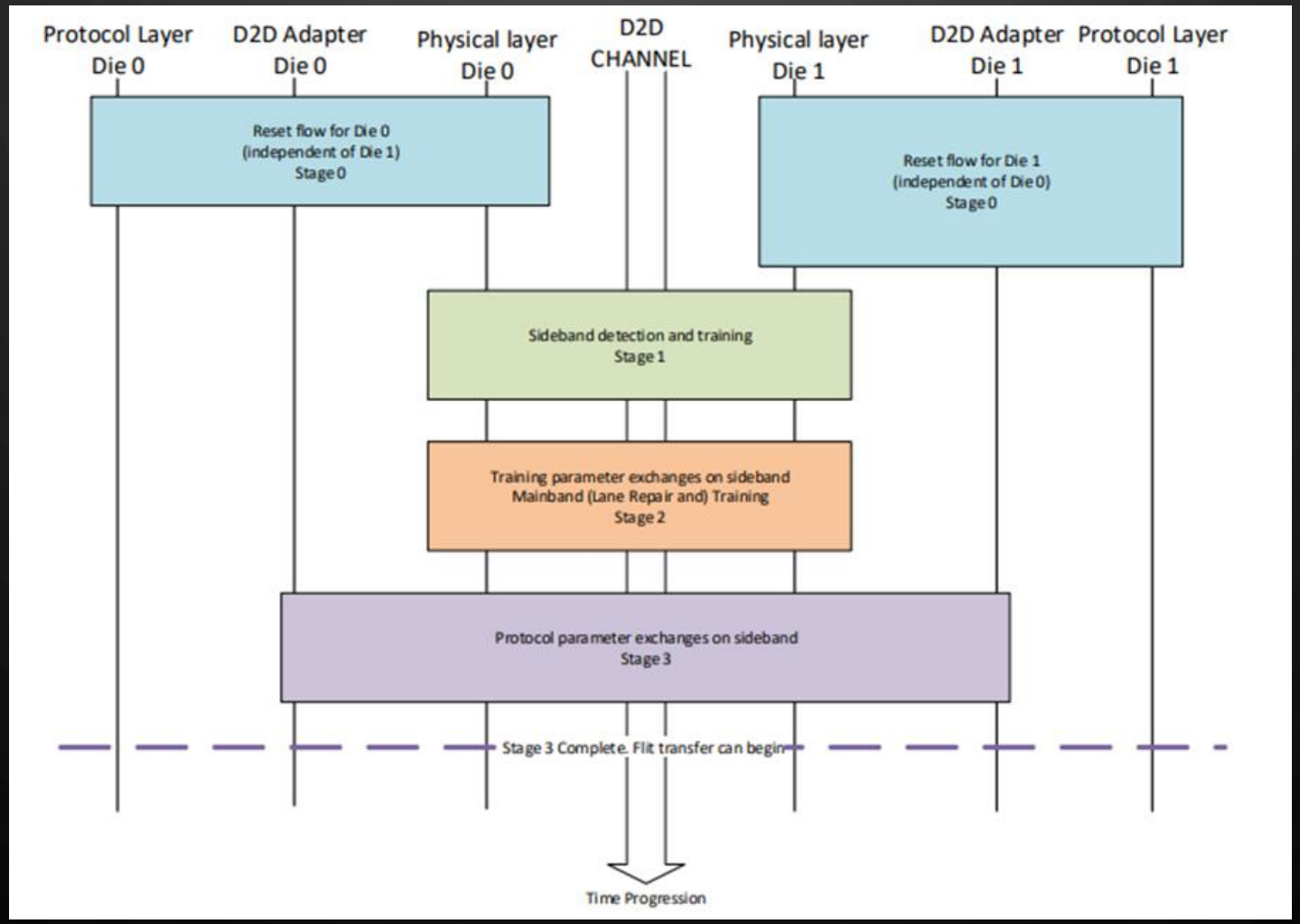
DIE TO DIE ADAPTER

- Trasferire i dati in maniera affidabile
- Gestire il MUX in caso di più protocol layer
- Gestire lo stato del collegamento
- Negoziare protocollo → Link Initialization
- Fare il CRC
- Gestire Retry



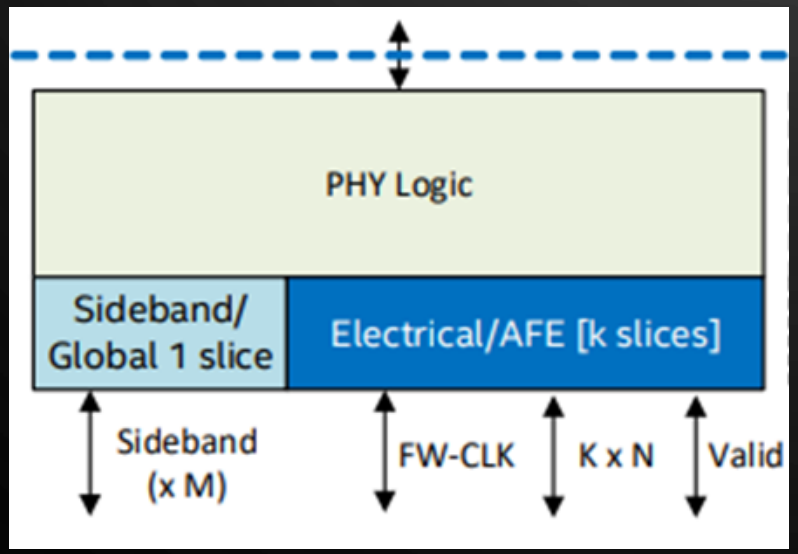


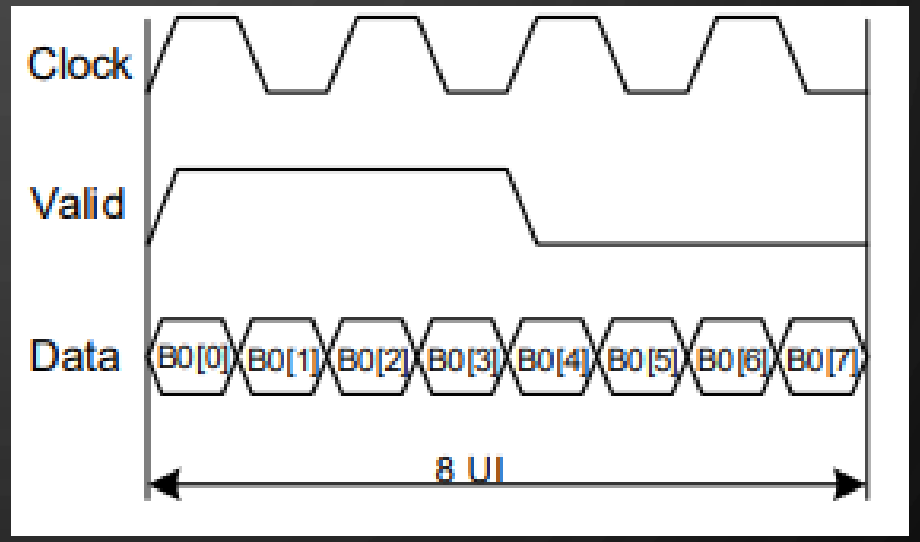
DIE TO DIE ADAPTER – Link Initialization



LOGICAL PHYSICAL LAYER

- Power Management
- Mapping byte-linea
- Data lane repair



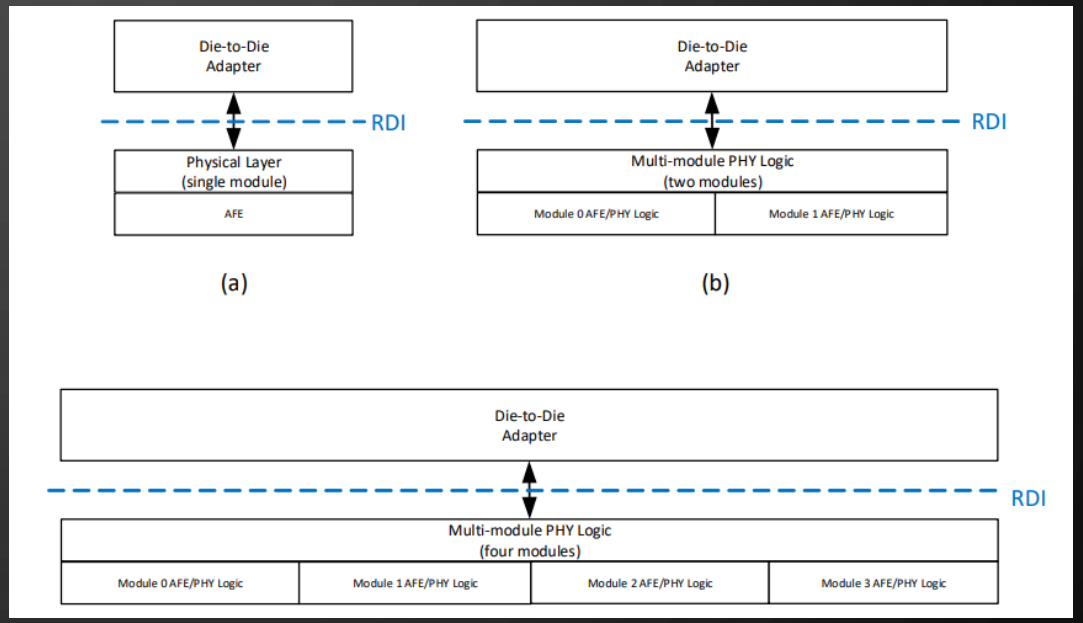


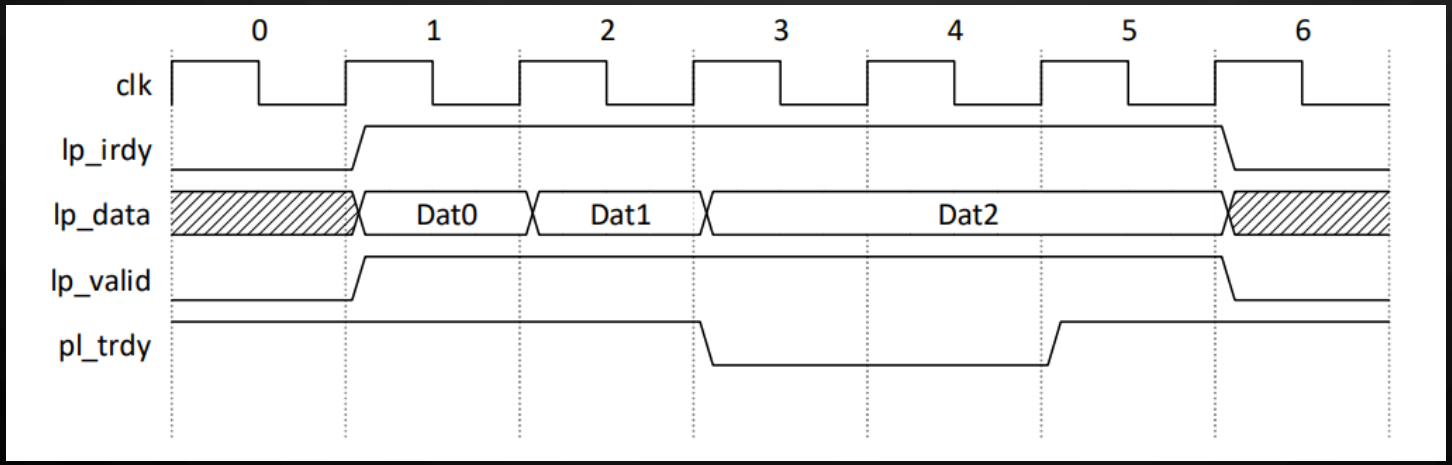
CARATTERISTICHE DEI PACKAGE

	Standard Package	Advanced Package
Transfer rate (GT/s)	4, 8, 12, 16, 24, 32	
Numero di linee	16	64
Bump pitch (µm)	100-130	25-55
Lunghezza canale (mm)	≤ 25	≤ 2

INTERFACCE - RDI

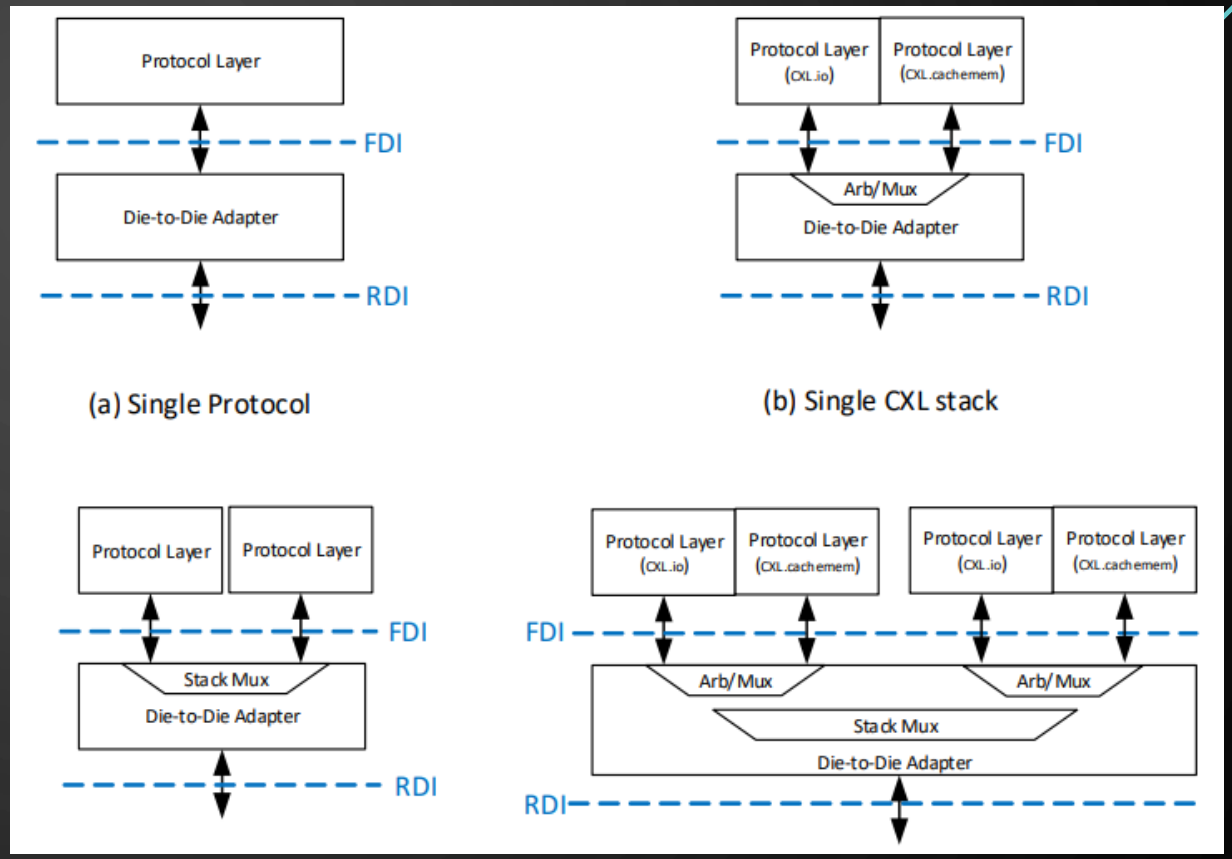
- Unico
- Bidirezionale
- Tra D2D Adapter e Logic PHY
- Stesso clock tra le due parti

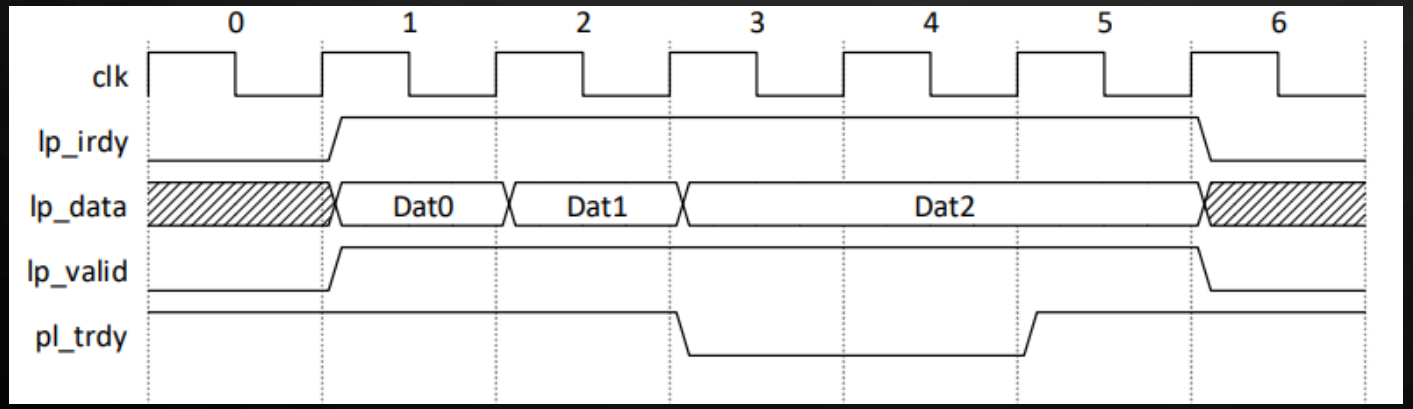




INTERFACCE – FDI

- Uno per ogni Protocol Layer
- Gestione di DLLP per PM, FC, Ack
- Stesso clock tra le due parti





UCIe AD ALTO LIVELLO

- Consistente con il protocollo mappato
- Retrocompatibile

TARGET PRESTAZIONI

Metrica	Transfer rate	Standard Package	Advanced Package
Bandwidth per lato del die (GB/s per mm)	4 GT/s	28	165
	8 GT/s	56	329
	12 GT/s	84	494
	16 GT/s	112	658
	24 GT/s	168	988
	32 GT/s	224	1317
Latenza (ns)		≤ 2	

DOMANDE

GRAZIE PER L'ATTENZIONE