

iSCSI

iS not an Apple appliance

Daniel Braniss
danny@cs.huji.ac.il

BSDCan 2008, Ottawa



why

- learn a new OS/kernel
- check feasibility of writing a driver knowing little about:
 - SCSI
 - network
- see what iSCSI is about



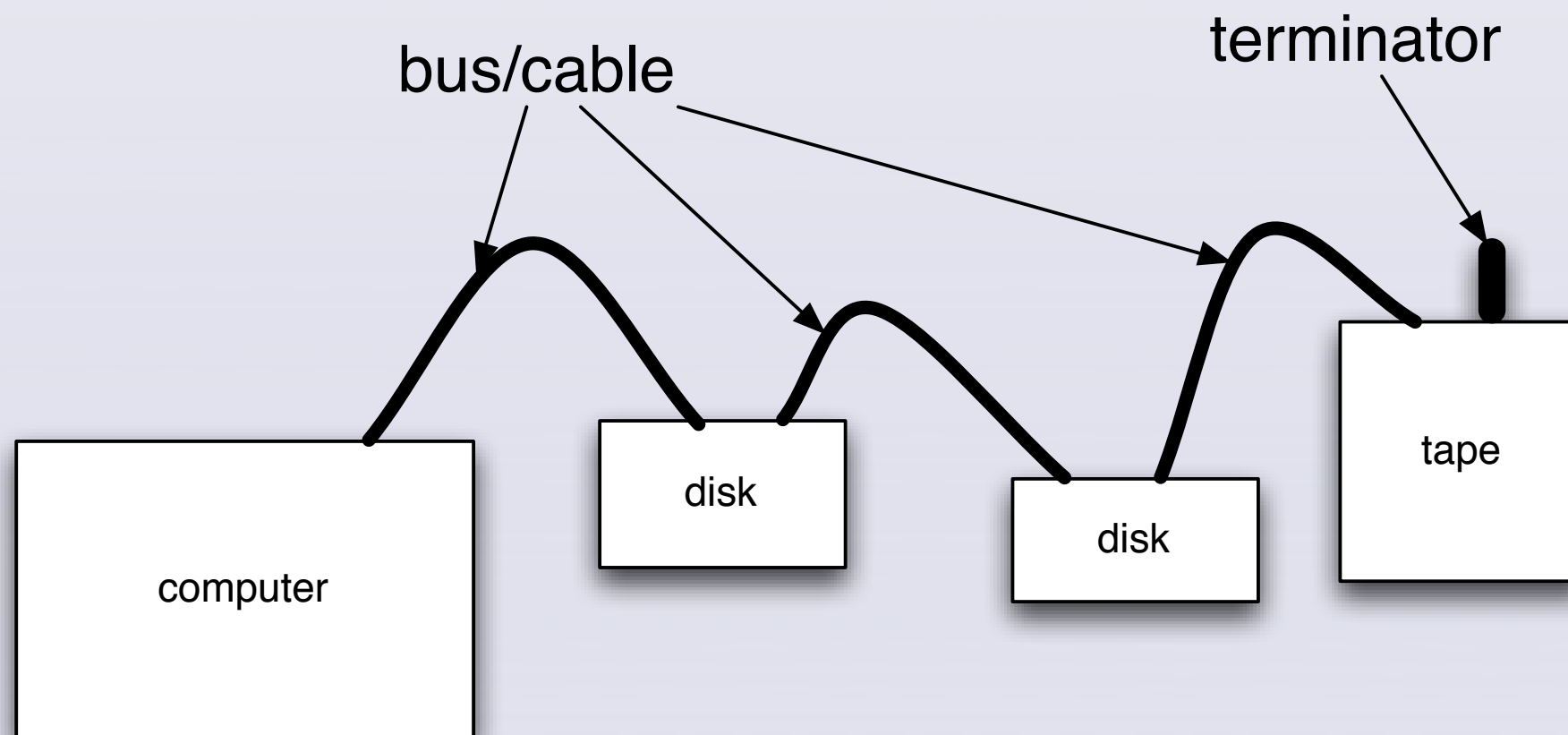
what's iSCSI

- iNTERNET Small Computer System Interface
- iNSECURE SCSI
- iNTERESTING

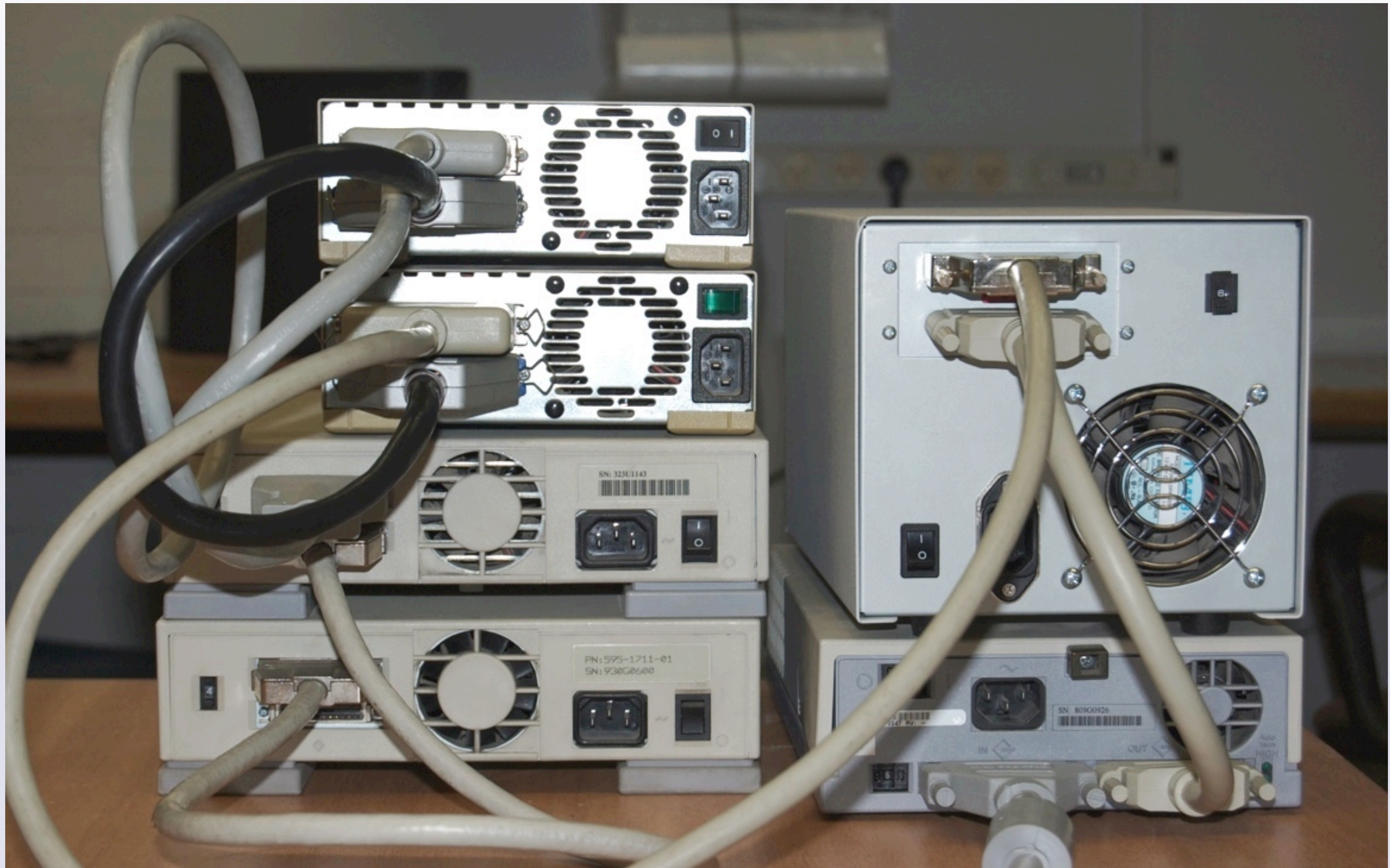


SCSI

Small Computer Storage Interface







SCSI ...

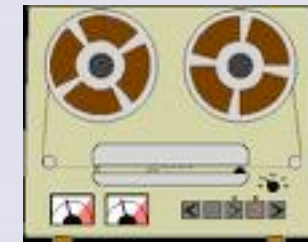
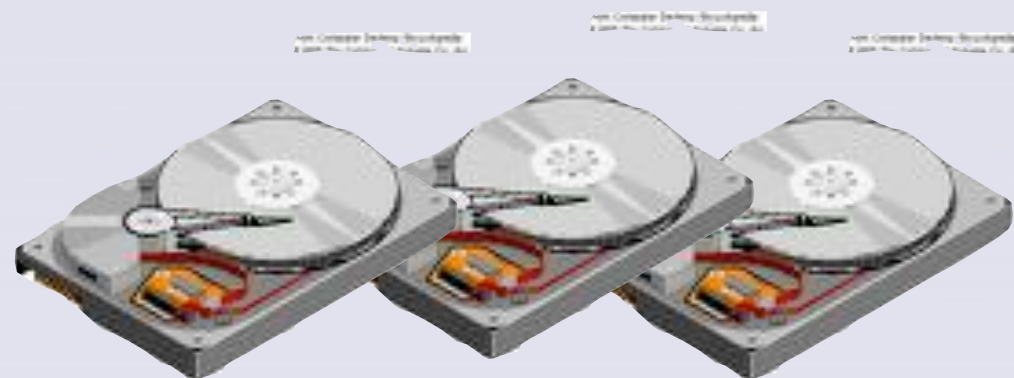
/dev/da0
/dev/sa0
/dev/ch0
...

tape driver
(sa)

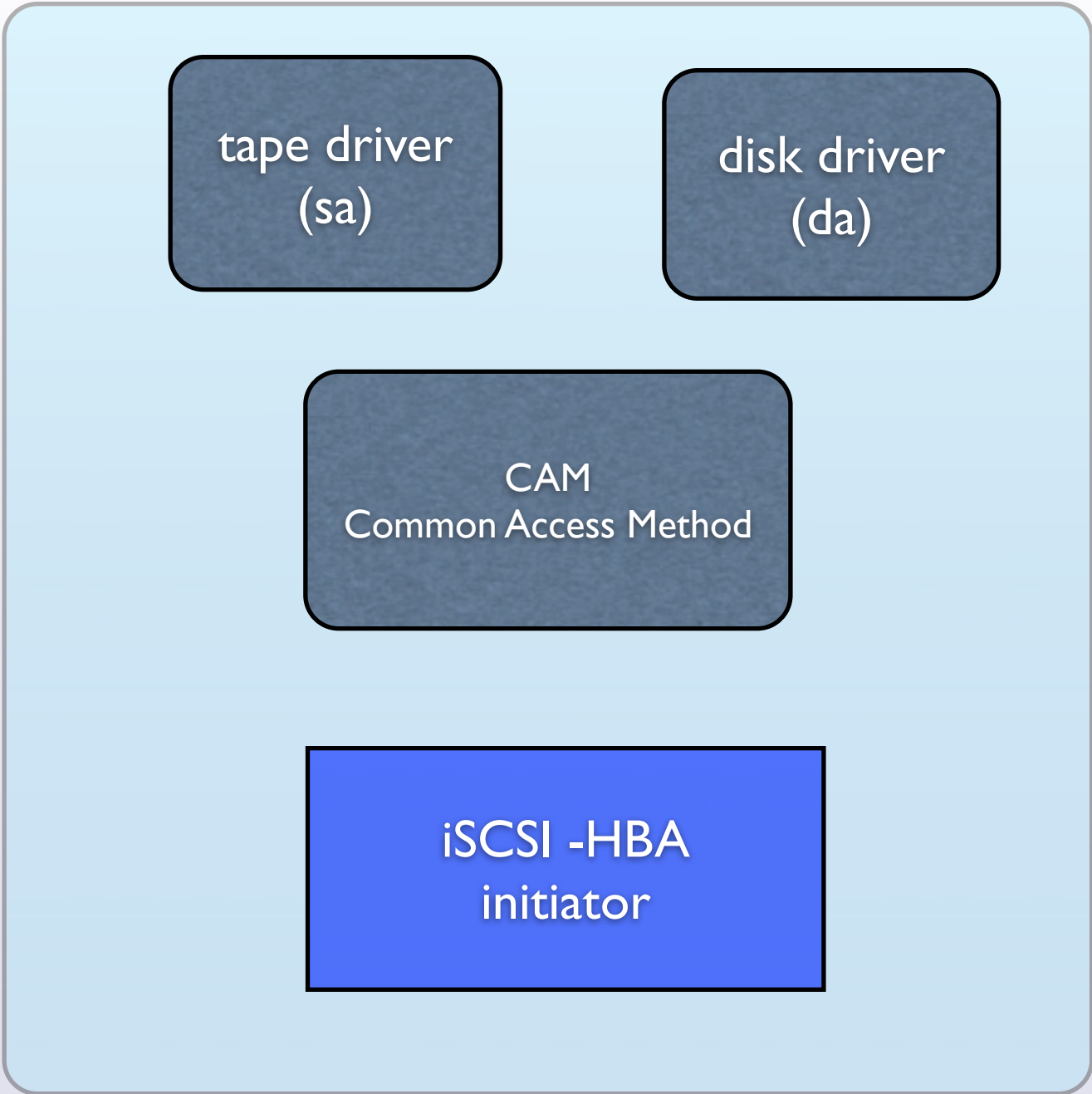
disk driver
(da)

CAM
Common Access Method

HBA
Host Bus Adaptor



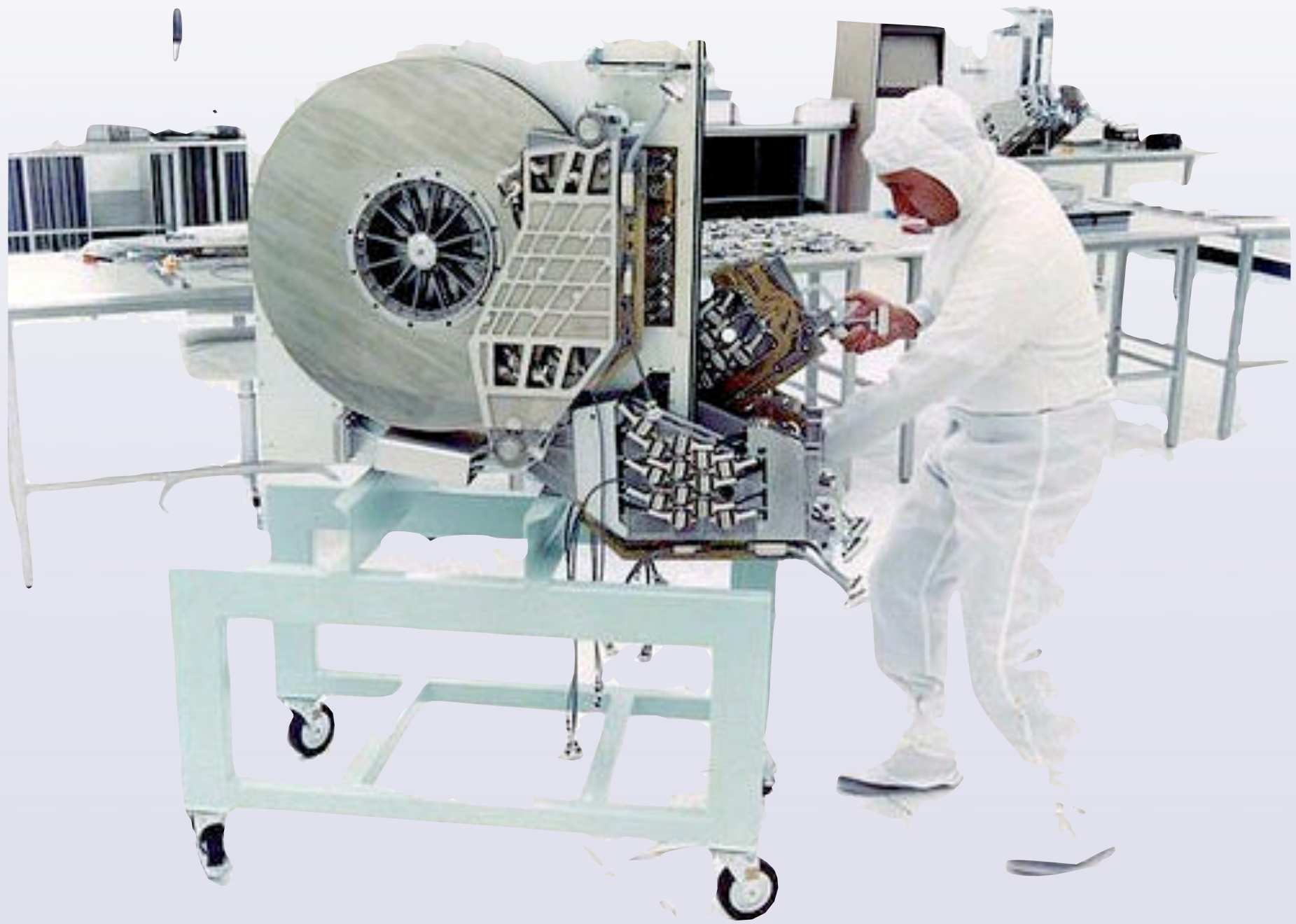
/dev/da0
/dev/sa0
/dev/ch0
...



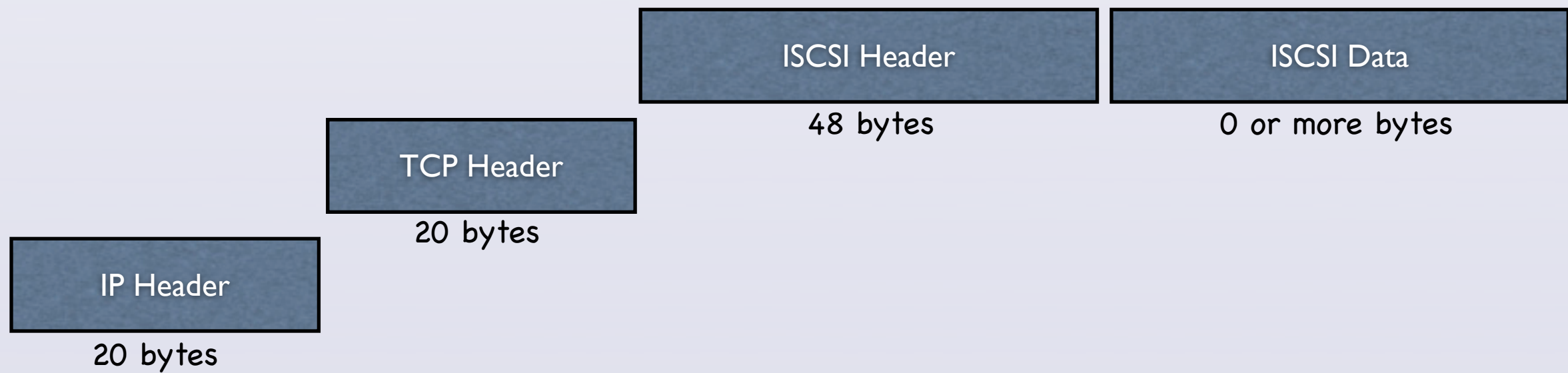
iscsi target

iscsi target

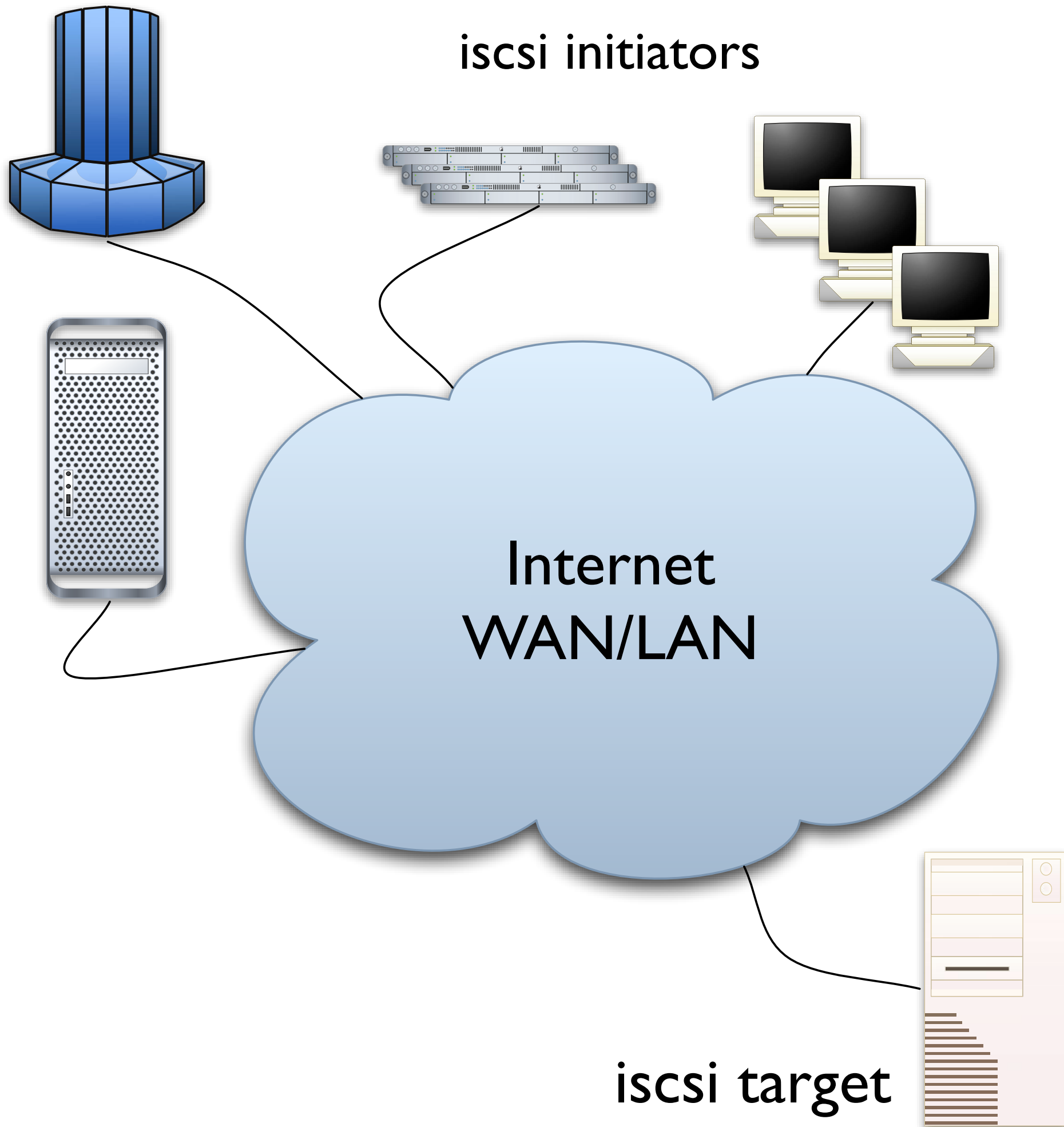




from bus to ether



iscsi initiators



iscsi target

iSCSI RFC 3720

- 257 pages – 14395 lines
- 7226 lines of code
- first products shipped in 2001 (IBM, Cisco)
- the RFC is from 2003/2004.



iSCSI login

there are 2 types of sessions:

- Discovery:
 - the response is a list of targets
- Normal:
 - goes into Full Feature mode/phase



first steps

1. read RFC for the nth. time
2. write some code
 - 2.1. translate designers language into programming language.
3. run
4. goto 1



iSCSI login ...

- needed to start a session
- establish authenticity of both/either/neither target and/or initiator based on
 - target name - `iqn.1992-08.com.netapp:sn.50405391`
 - initiator name - `iqn.2005-il.ac.huji.cs:xhost`
 - if CHAP then also some shared secret
- negotiate some options:
 - header/data CRC



iSCSI PDU

Protocol Data Unit

	byte 0	byte 1	byte 2	byte 3
0	Basic Header Segment (BHS)			
...				
48	Additional Header Segment 1(optional)			
...	Additional Header Segment 2(optional)			
...	...			
...	Additional Header Segment n(optional)			
k	Header-Digest (optional)			
l	Data-Segment (optional)			
m	Data-Digest (optional)			



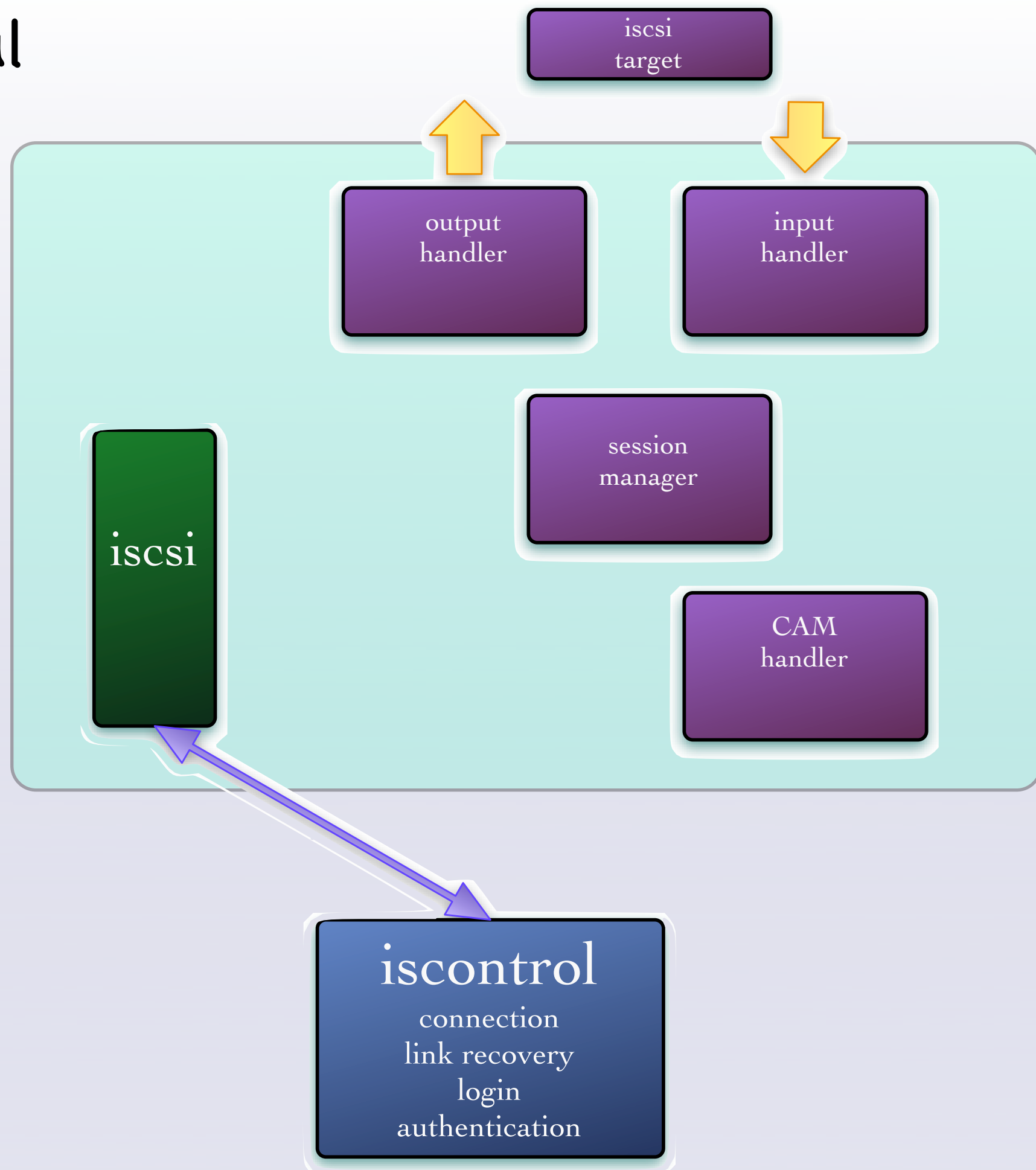
iSCSI PDU ...

Basic Header Segment - BHS

	byte 0	byte 1	byte 2	byte 3
0	. Opcode	F	Opcode-Specific fields	
4	TotalAHSLength	DataSegmentLength		
8	LUN or Opcode-Specific Fields			
12				
16	Initiator Task Tag			
20 ...	Opcode-Specific Fields			
48				



architectural view



iscontrol ...

Globals

#

port = 3260

myiscsi {

targetname = iqn.1900.com.com:sn.123456

initiatorname = iqn.2005-01.il.ac.huji.cs:somebody

targetaddress = iscsi1; tags = 8; port=3333;
}

chaptest {

targetaddress = iscsi2

targetname = iqn.1900.com.com:sn.123456

initiatorname = iqn.2005-01.il.ac.huji.cs:nobody

authmethod = CHAP

chapsecret = 0x3713c3336d9a224c2791c873d3d2b174

tags = 8

}



iscsi_initiator

- kernel side of the protocol.
 - consists of several threads
 - handles PDU queues
 - iSCSI <-> CAM
 - about 4500 line of code



iscsi_initiator ...

- Session Manager
 - for each 'session':
 - starts a session-manager
 - which starts a receiver on the given connection/socket
 - loops for something to do till:
 - terminated from user land
 - user land died



iscsi_initiator – receiver

- waits till it gets a 48 byte header
- calculates size of PDU
- process the received PDU
 - if it's a SCSI command pass it to the CAM

will run as long as there is a connection



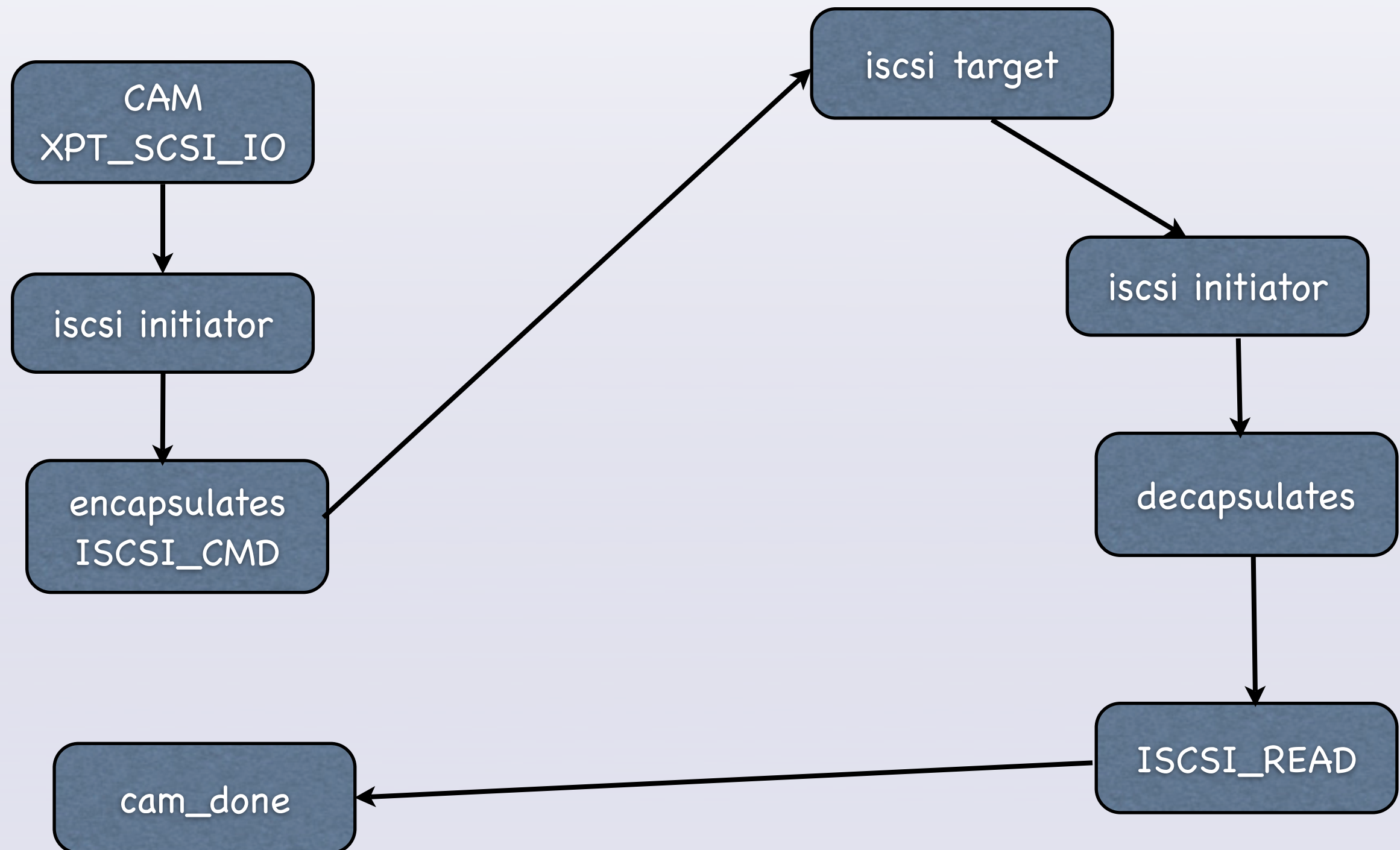
iscsi_initiator – dispatcher

- goes through the PDU queues for something to do
- PDU queues:
 - Immediate
 - R2T(Request to Transmit)
 - command (SCSI) – only if window size allows.
 - hold – waiting for ack

one per session



iscsi_initiator Read



iscsi initiator ...

- sync with the shutdown process
 - `iscontrol` gets killed.
 - we loose the socket/connection
- finally the buffers get flushed.



iscsi_initiator ...

- Recovery:



numbers ...



conclusions



to do list

- recovery
- header/data digest
- boot from an iSCSI
- make the CAM aware of virtual scsi devices?
- clean up
- target?

