REDUNDANCY

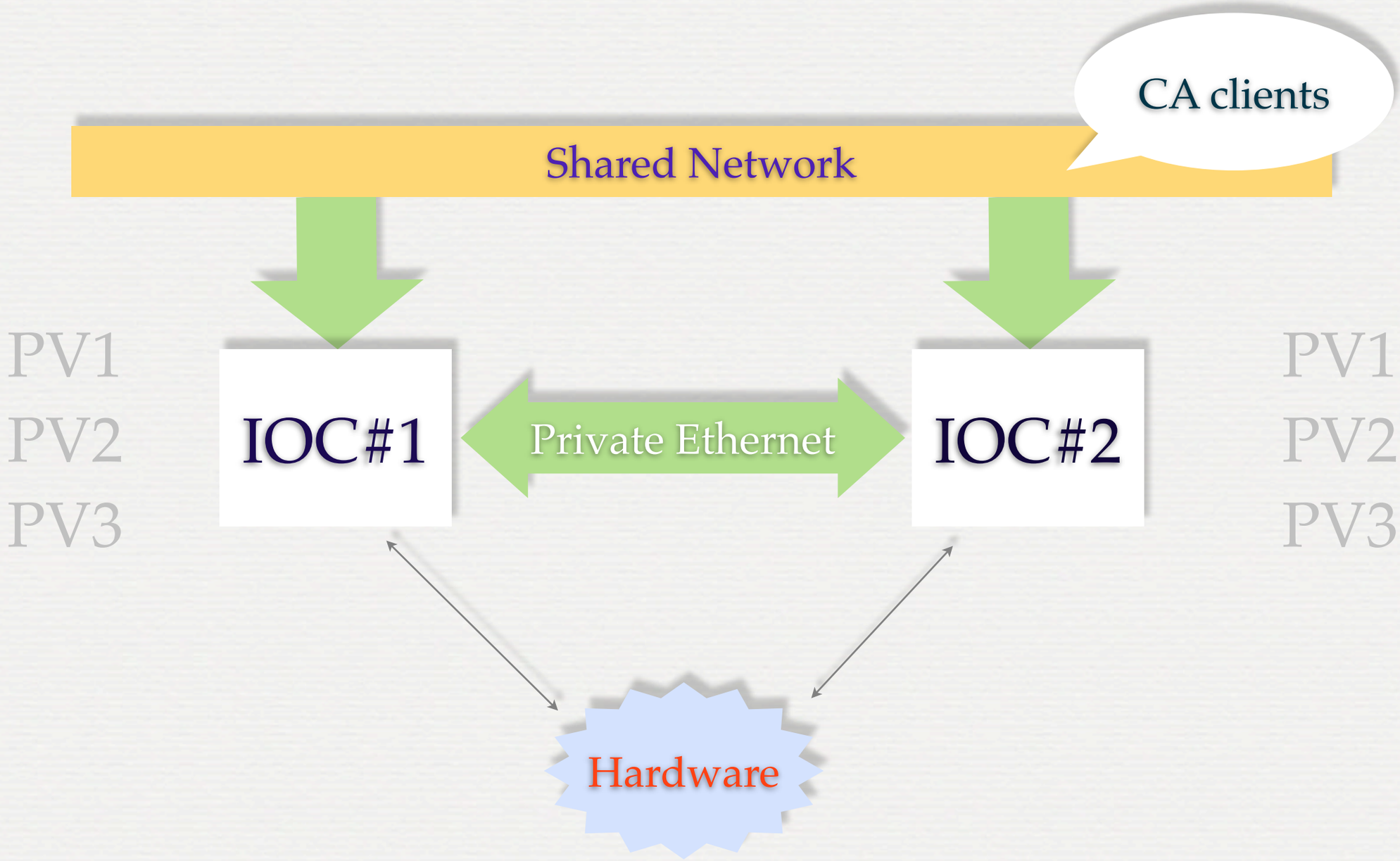
THE NEED FOR REDUNDANCY

- EPICS is a great software, but lacks redundancy support
- which is essential for some highly critical applications such as cryogenic plants

ORIGINAL EPICS REDUNDANCY

- Was developed by DESY in collaboration with SLAC
- support for vxWorks operating system only

WHAT IS REDUNDANT IOC?

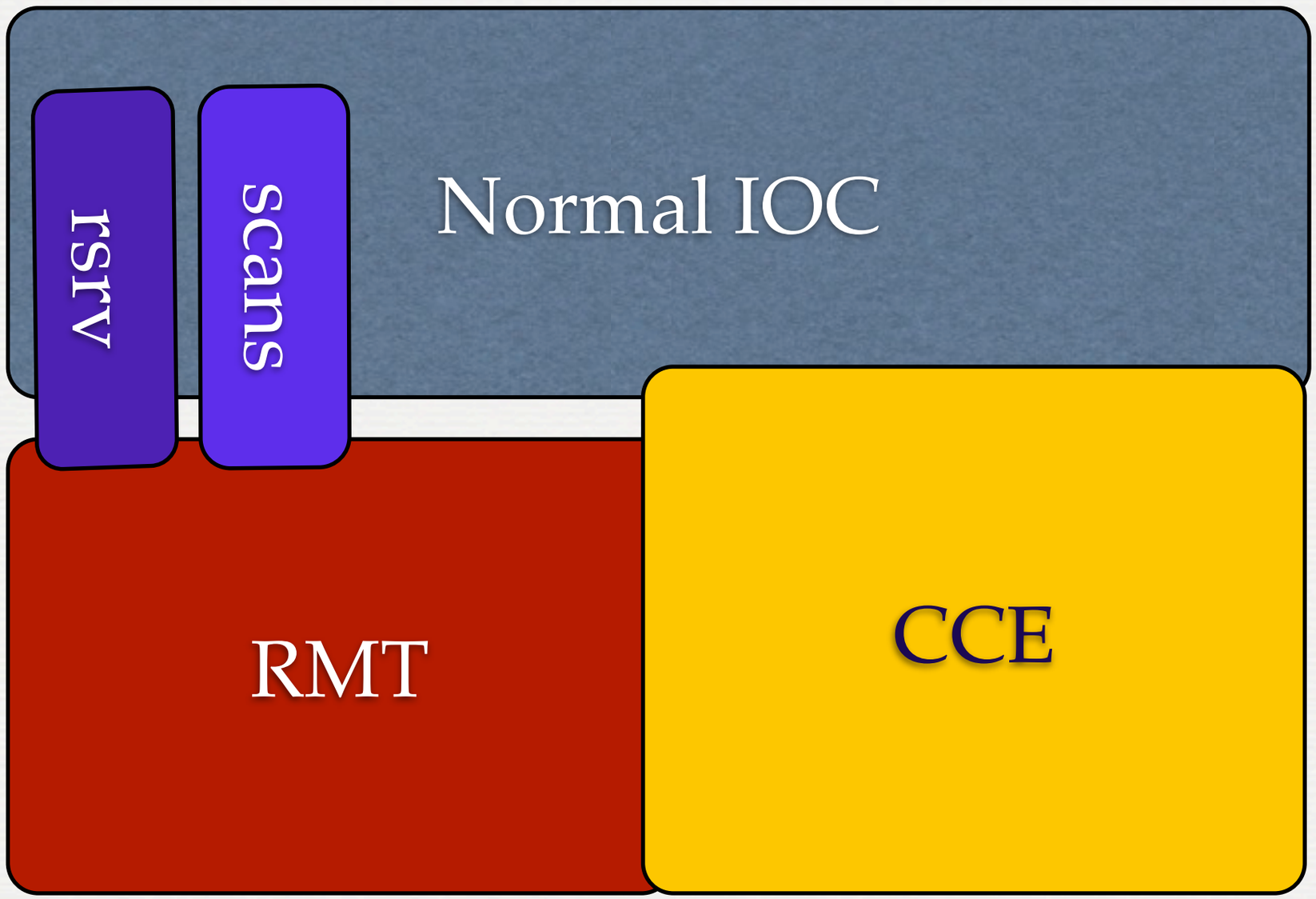


2. Redundancy

EPICS REDUNDANCY TERMINOLOGY

- RMT: Redundancy Monitoring Task - key component of EPICS redundancy implementation
- CCE: Continuos Control Executive - data “exchanger” for EPICS IOC
- RMT Driver: a piece of software which conforms to RMT API

REDUNDANT EPICS IOC INTERNALS



RMT FUNCTIONS

- Check “health” of the drivers
- And control drivers (start, stop, sync, etc...)
- Check connectivity with the network
- Communicate with the “partner”
- And decide when to switch to the partner

GENERALIZATION OF EPICS REDUNDANCY

- Other laboratories showed some interest in redundancy for EPICS, including KEK
- **Need for redundancy on platforms other than vxWorks**
- Could use RMT to **make other software redundant** on Linux and other systems
  - even EPICS unrelated software

GENERALIZATION OF EPICS REDUNDANCY

- all vxWorks specific code was replaced with EPICS/OSI (Operating System Independent) library calls
- additional libOSI functions were implemented

GENERALIZED VERSION

- works on vxWorks
- Linux
- Darwin (Mac OS X)
- and virtually on any EPICS supported OS
- can be used to add redundancy to other software

GENERALIZED VERSION

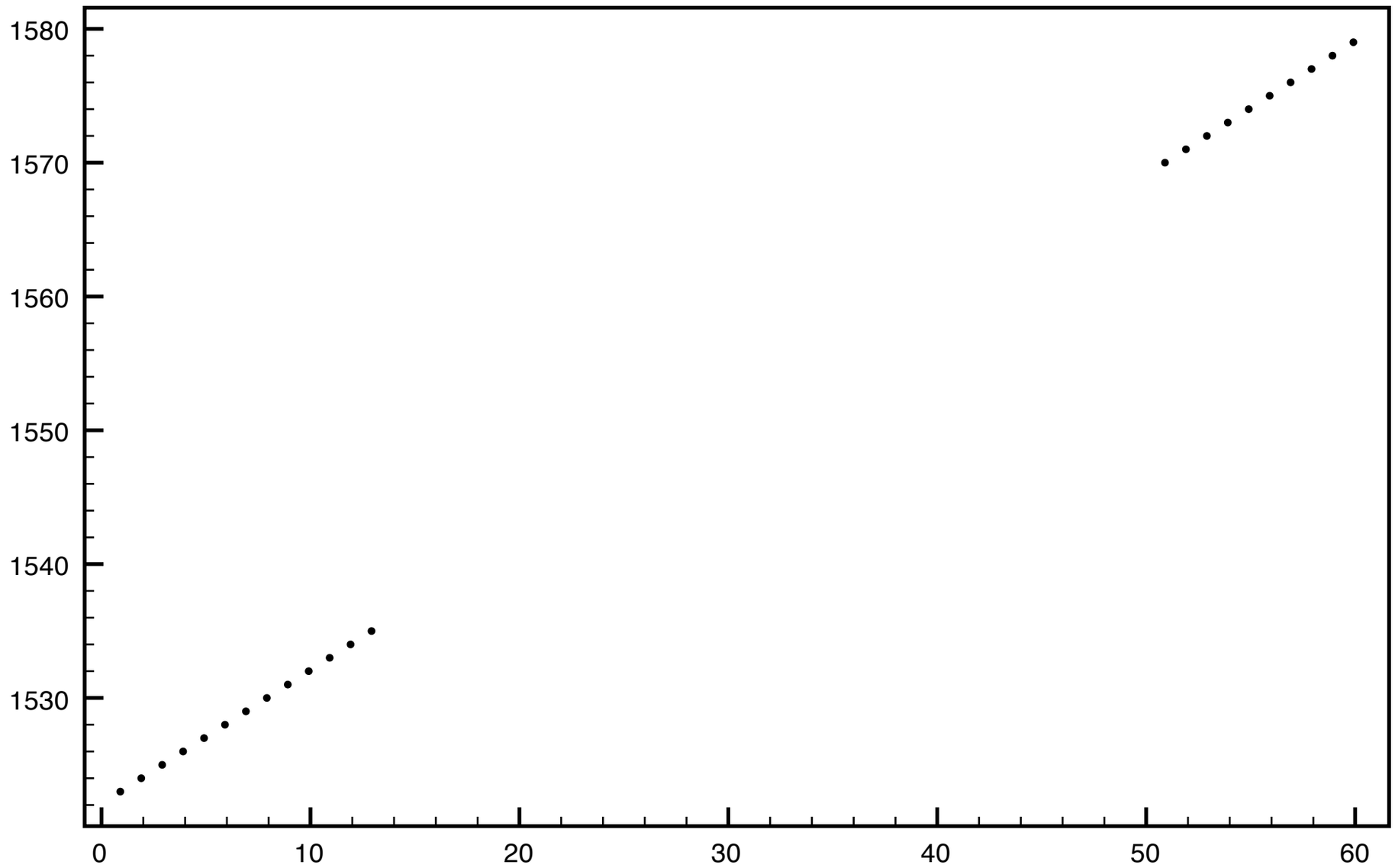
Allowed to include EPICS redundancy support into EPICS BASE distribution

since 3.14.10 base has all the “hooks” needed for redundant IOC

SOME NUMBERS

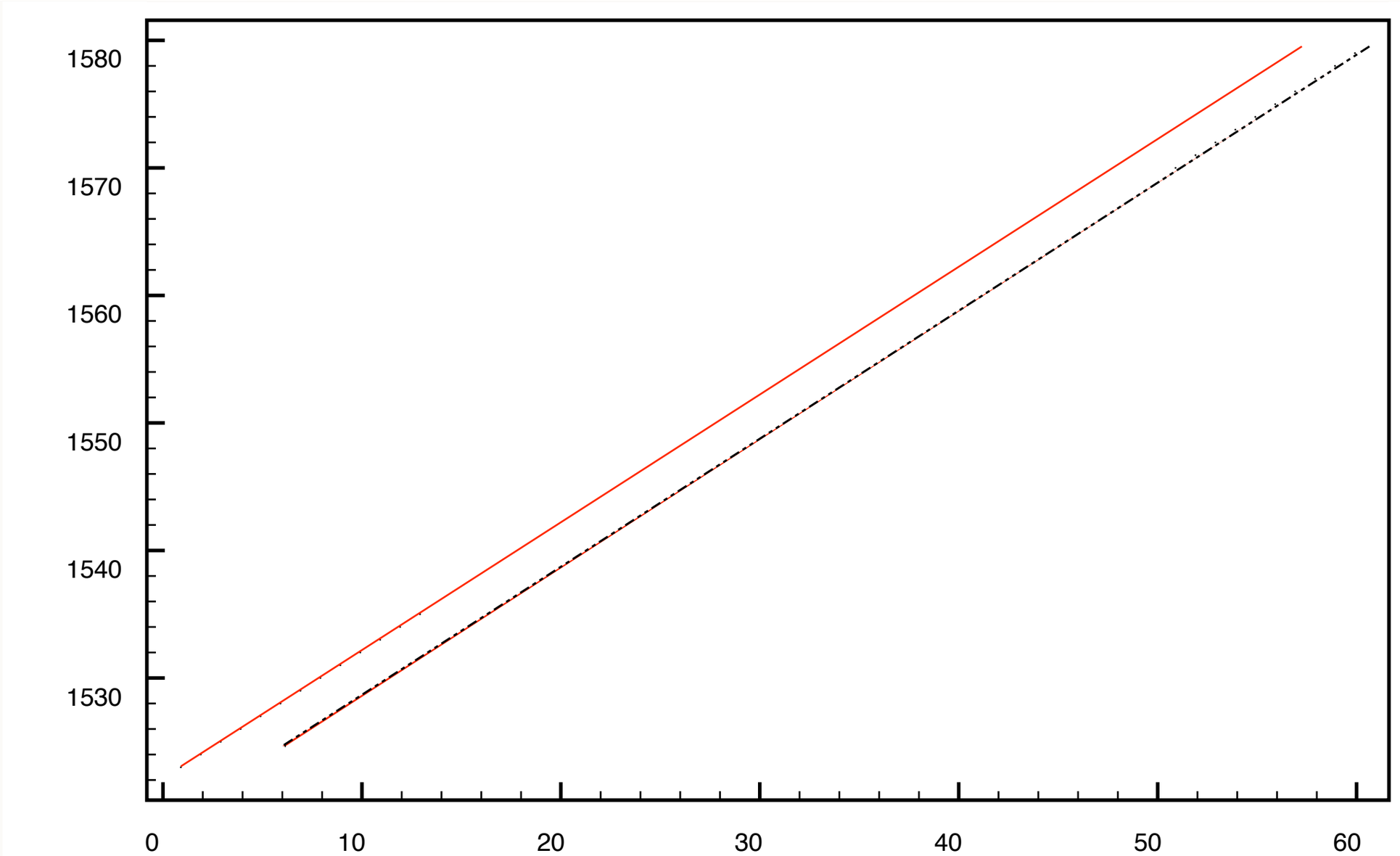
- switchover time < 3sec
  - in case of normal IOC it could be from several minutes to hours
- CCE can handle synchronization of ~5000/sec records

SWITCH OVER "TIME-LOSS"



2. Redundancy

SWITCH OVER "TIME-LOSS"



2. Redundancy

REDUNDANT CHANNEL ACCESS GATEWAY

CA GATEWAY

- very common program widely used in many laboratories

- used to make two or more subnets CA visible to each other

- and to provide access control, i.e. read ability for everyone outside control network

CA GATEWAY OPERATION

2. Redundancy

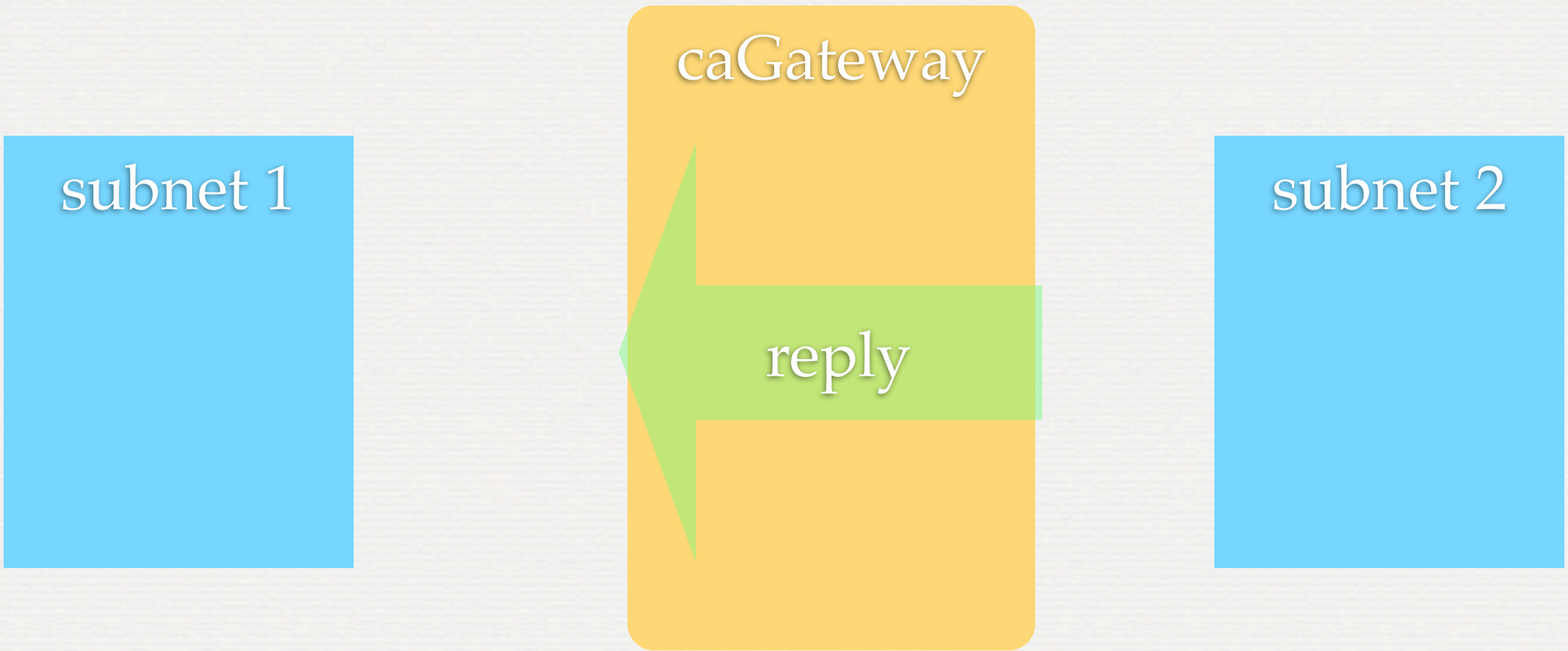
CA GATEWAY OPERATION

2. Redundancy

CA GATEWAY OPERATION

2. Redundancy

CA GATEWAY OPERATION



2. Redundancy

CA GATEWAY OPERATION

2. Redundancy

CA GATEWAY OPERATION

2. Redundancy

CA GATEWAY NEEDS REDUNDANCY

It is **single point of failure**: if it is not working whole subnet becomes unreachable for other subnet

REDUNDANT CA GATEWAY

- Has no critical internal state data to be synchronized between peers
- Can be redundant “out-of-the-box”, but client would see multiple replies
- would be very nice to have “load-balancing”, which would improve response time and improve throughput

CONFUSING REDUNDANCY

Client:

GW #1:

GW #2:

CONFUSING REDUNDANCY

Client:

-Who has PV?

GW #1:

GW #2:

CONFUSING REDUNDANCY

Client:

-Who has PV?

- I'm Confused !!!

GW #1:

-I have!

GW #2:

-I have!

LET'S ADD RMT

Client:

GW #1:

GW #2:

2. Redundancy

LET'S ADD RMT

Client:

GW #1:

GW #2:

2. Redundancy

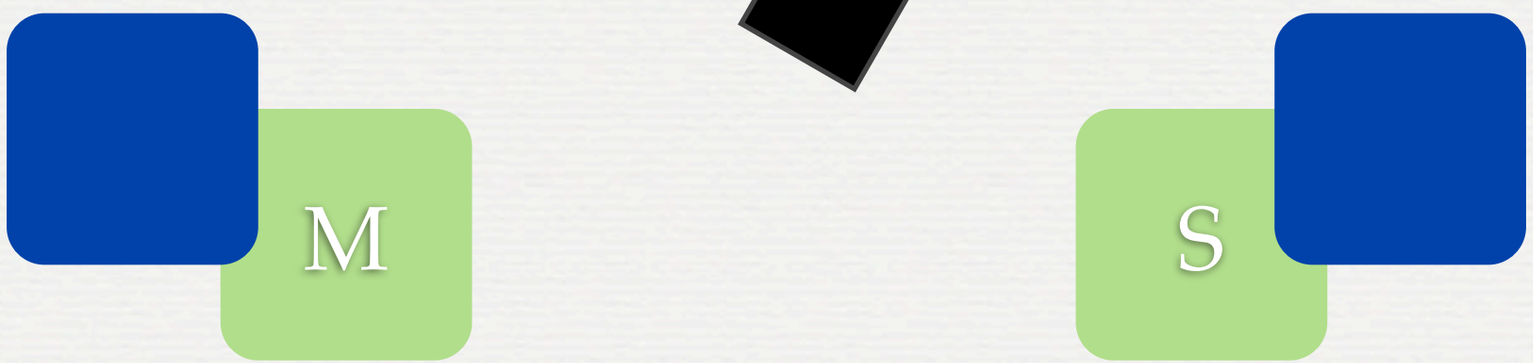
LET'S ADD RMT

Client:

-Who has PV?

GW #1:

GW #2:



2. Redundancy

LET'S ADD RMT

Client:

-Who has PV?

- OK!!!

GW #1:

-I have!

GW #2:

-I have!

2. Redundancy

LET'S ADD RMT

Client:

-Who has PV?

- OK!!!

GW #1:

-I have!

GW #2:

-I have!

2. Redundancy

REDUNDANCY ONLY

- RMT as separate process, which does all monitoring, health-checking and decision making
- Gateway is running as usual
- On “SLAVE” we block replies from the Gateway by firewall rule
- no modification to the source code of GW (!!!)
  - which means no new bugs whatsoever (!)

ADD LOAD BALANCING

- Inform GW about its partner status, whether it is alive
- Load-balance using “directory service”-feature of CA protocol

First query

Client:

GW #1:

GW #2:

2. Redundancy

First query

Client:

GW #1:

GW #2:

2. Redundancy

First query

Client:

-Who has PV?

GW #1:

GW #2:

2. Redundancy

First query

Client:

-Who has PV?

- OK!!!

GW #1:

-I have!

GW #2:

-I have!

First query

Client:

-Who has PV?

- OK!!!

GW #1:

-I have!

GW #2:

-I have!

SECOND QUERY

Client:

GW #1:

GW #2:

2. Redundancy

SECOND QUERY

Client:

-Who has PV2?

GW #1:

GW #2:

2. Redundancy

SECOND QUERY

Client:

-Who has PV2?

- OK!!!

GW #1:

-GW2 has!

GW #2:

-GW1 has!

2. Redundancy

SECOND QUERY

Client:

-Who has PV2?

- OK!!!

GW #1:

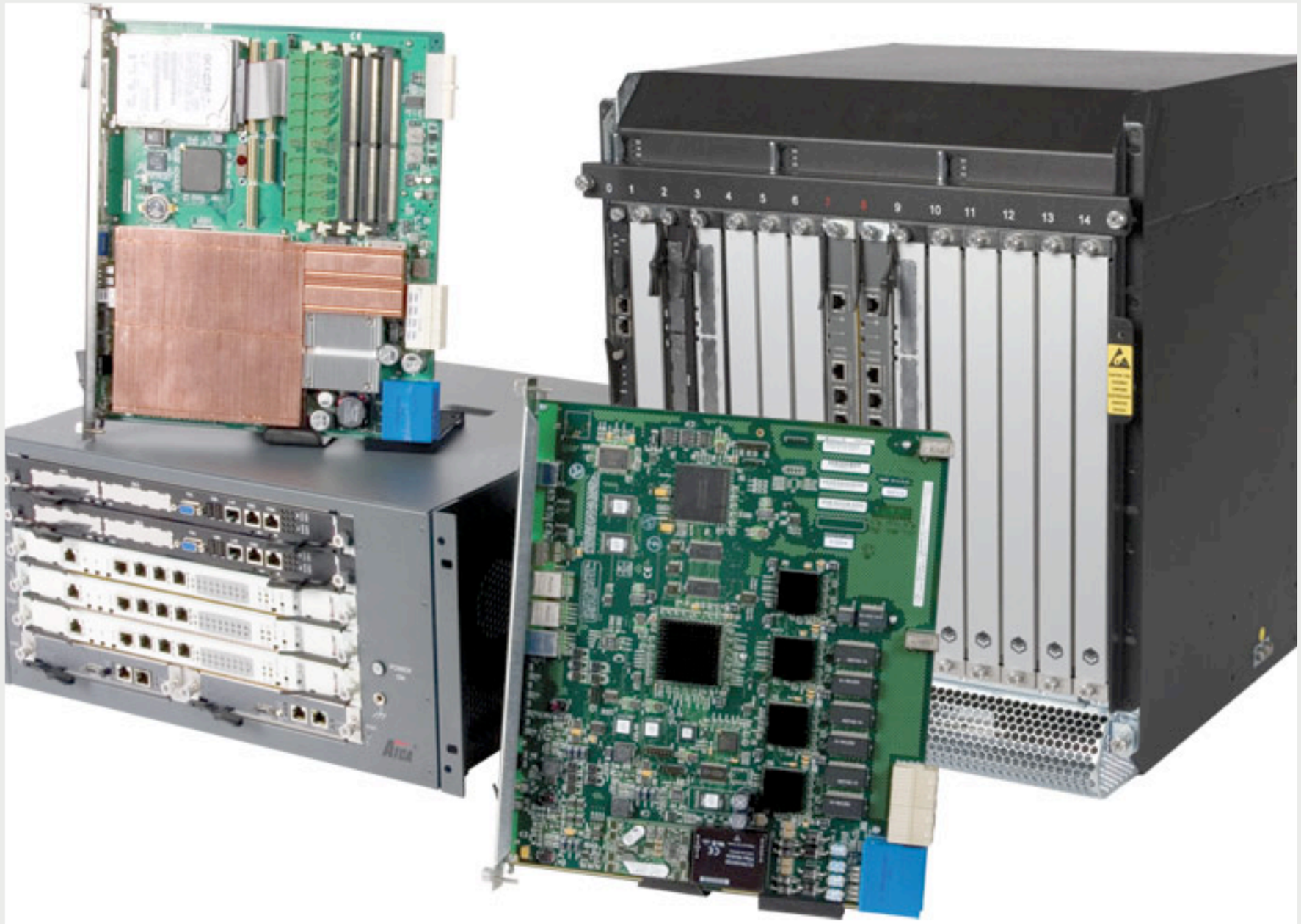
-GW2 has!

GW #2:

-GW1 has!

2. Redundancy

REDUNDANT IOC ON ATCA



ADVANCED TELECOM. COMPUTING ARCHITECTURE

Example boards and crates

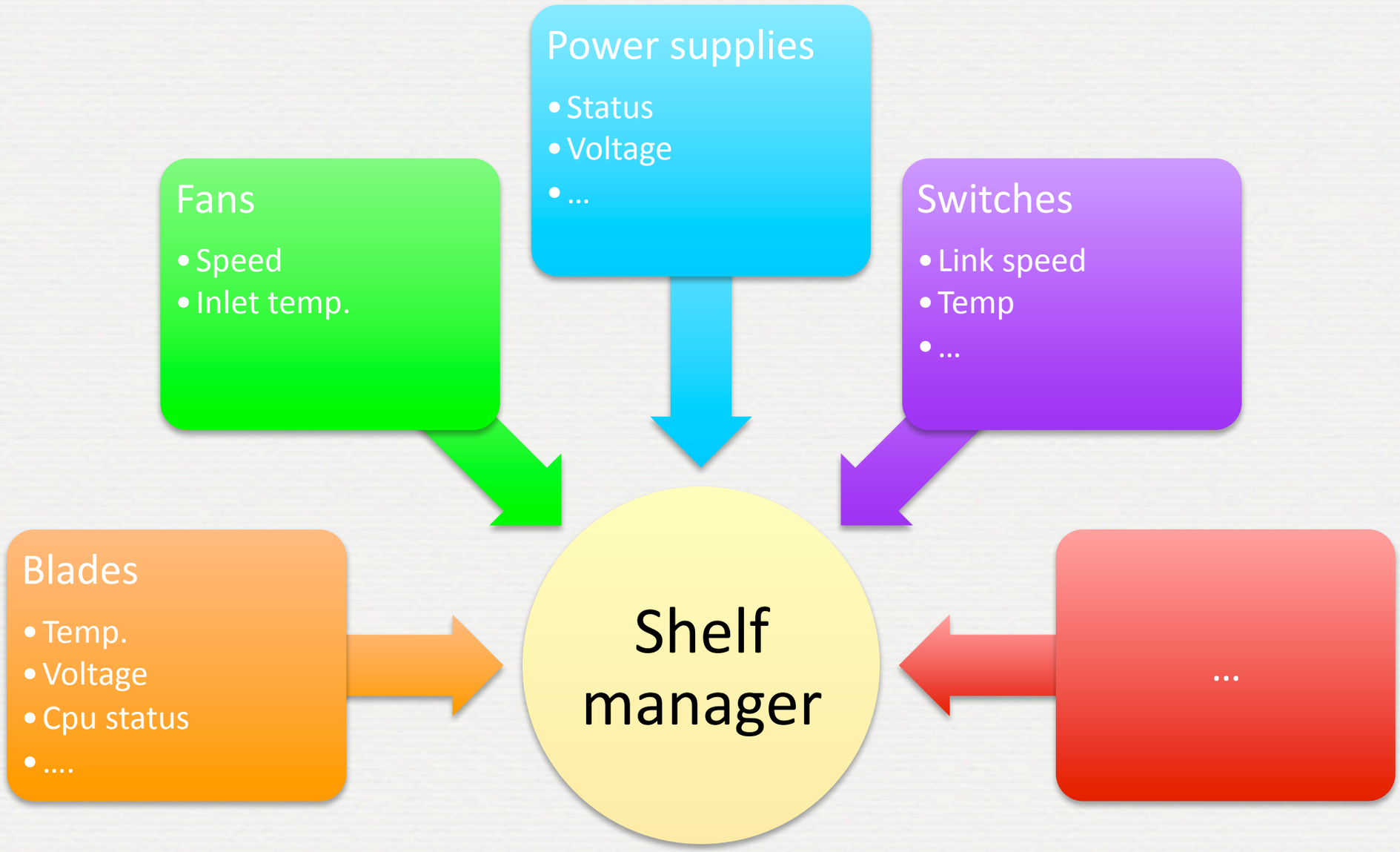
ADVANCED TELECOM COMPUTING ARCHITECTURE

- ATCA is a relatively new standard targeted as a platform for Highly Available applications

WHY RUN RIOC ON ATCA

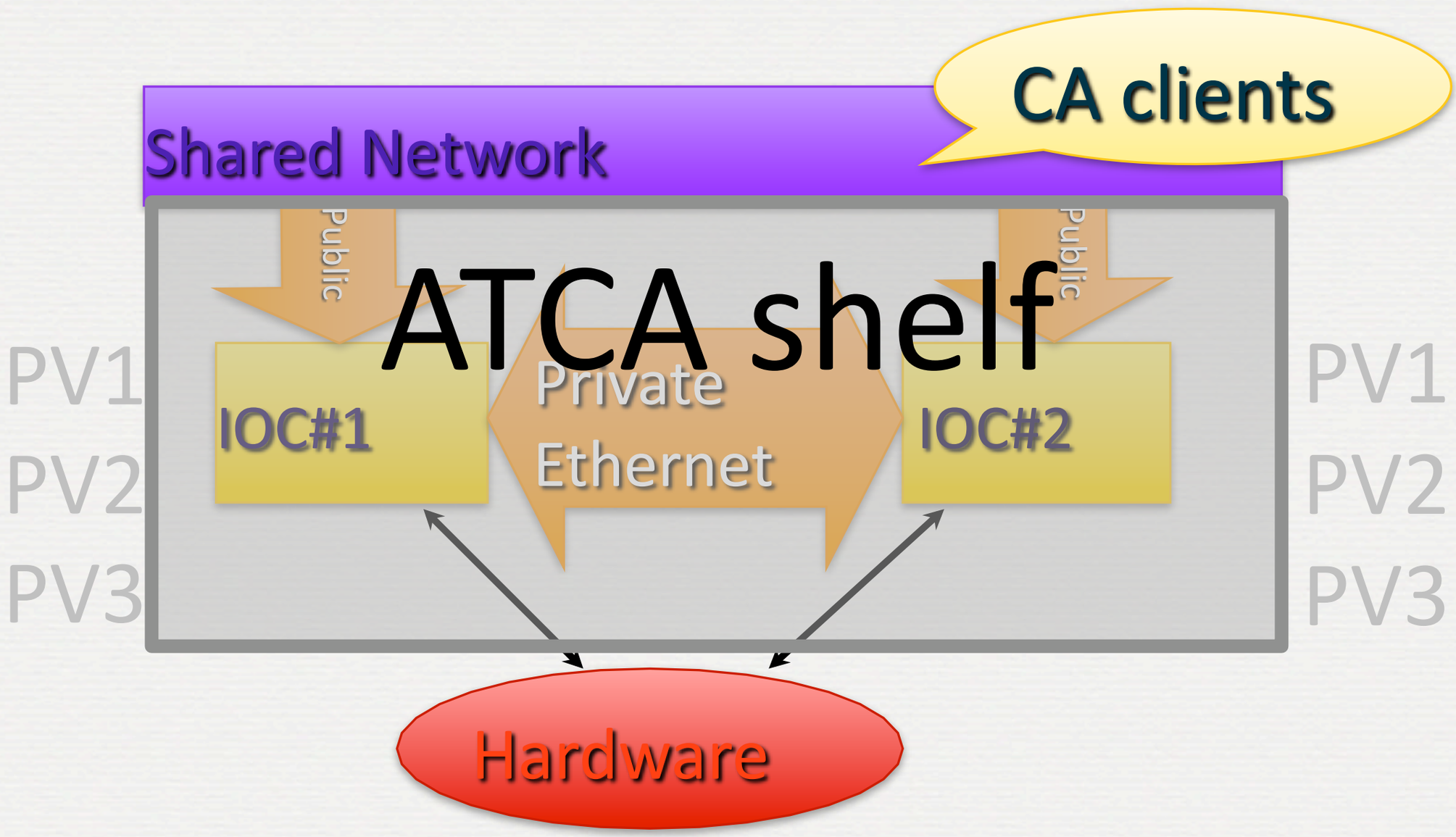
- ATCA is a modern industry standard for HA applications
- Very reliable (99.999% design availability)
- ATCA is suggested as a platform for the ILC control system
- ATCA is a **hardware designed for critical applications** and RIOC is a **software designed for critical applications**

ATCA SHELF MANAGER



Data is exchanged through redundant Intelligent Platform Management Bus IPMB

"PLAIN" RIOC ON ATCA



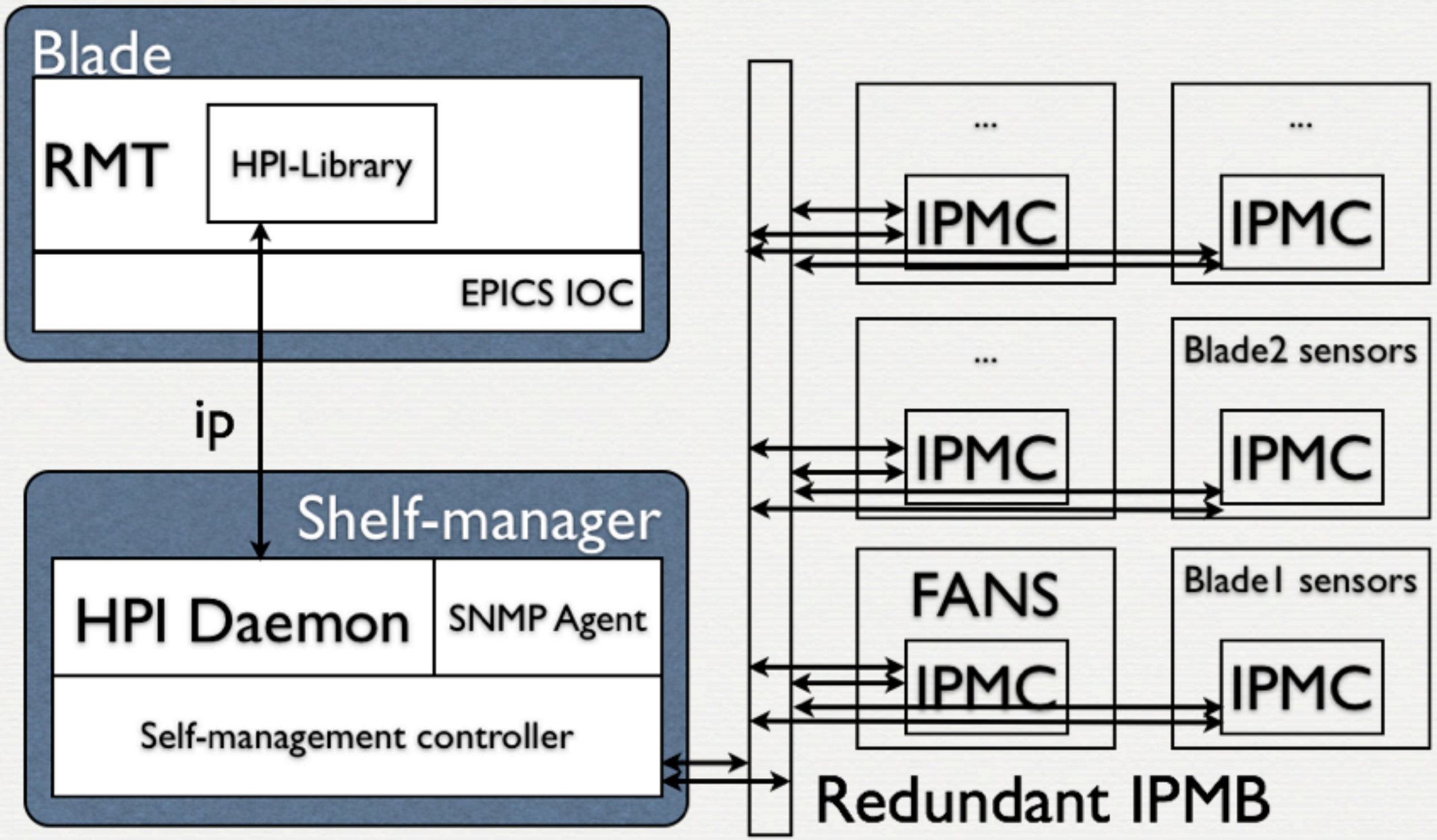
“PLAIN” RIOC ON ATCA

- can run RIOC on ATCA without modification
- But does not know anything about the “smart” hardware of ATCA
- Basically is same as running on two normal PCs

BENEFITS OF USING ATCA-"AWARE" RIOC

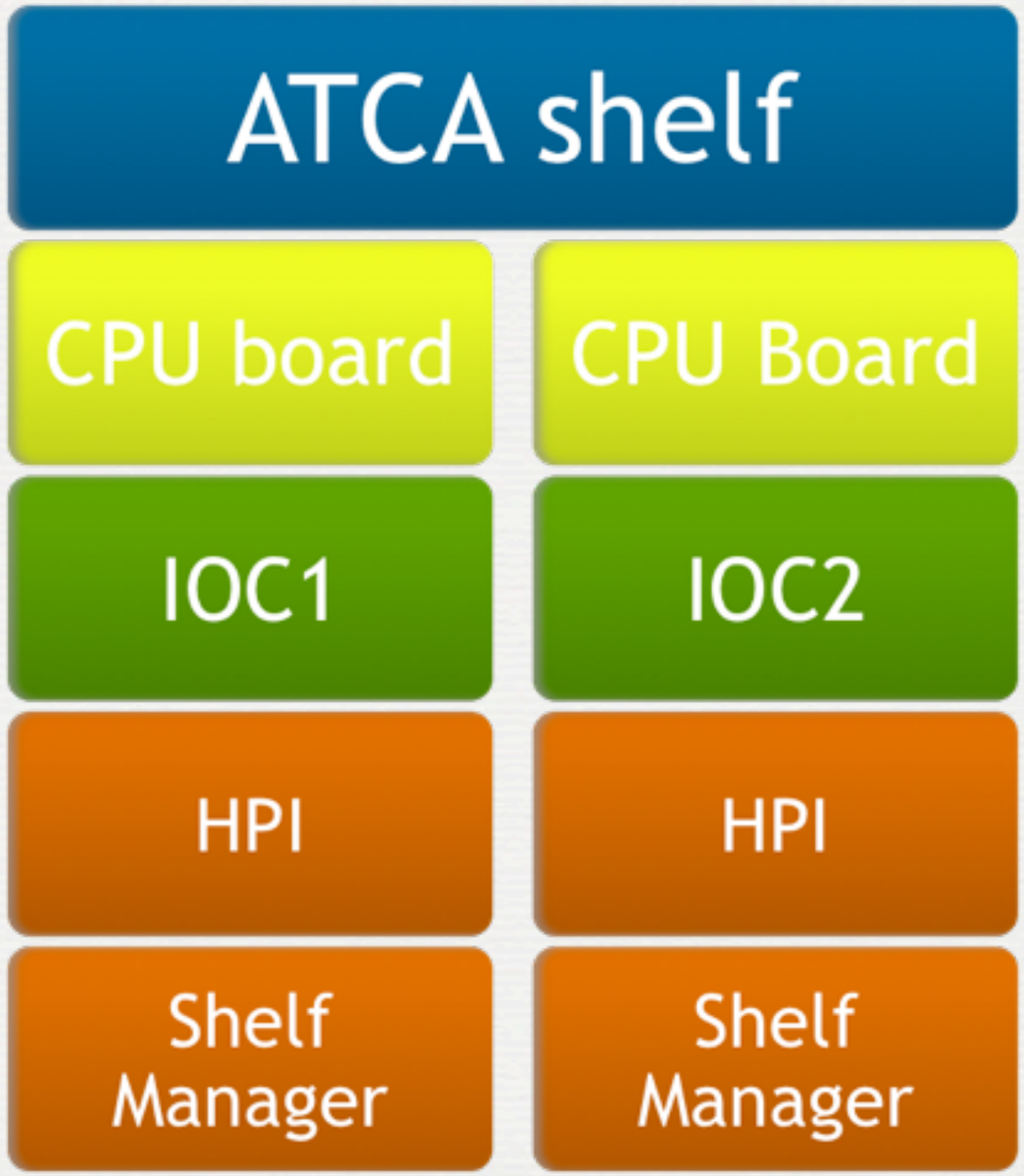
- Failures can be "predicted"
- i.e. temperature starts to rise and the CPU is still working -> we can initiate fail-over procedure before actual hardware fails -> fail-over occurs in more stable and controlled environment
- Client connections can be gracefully closed
- Allowing the client to reconnect to back-up IOC within 1 second
- In case of "real" hardware failure reconnect would occur only after 30 seconds

ATCA-"AWARE" RIOC



HPI USAGE EXAMPLE – REDUNDANT EPICS IOC

- HPI (Hardware Platform Interface) is used to monitor the health of each blade and the shelf
- This information is used to make decision on failover



HPI USAGE EXAMPLE – REDUNDANT EPICS IOC

- HPI is Platform independent
- Instead of ATCA we can use “conventional” server PC
- OpenHPI has /dev/ sysfs mappings on Linux