Peter Chochula

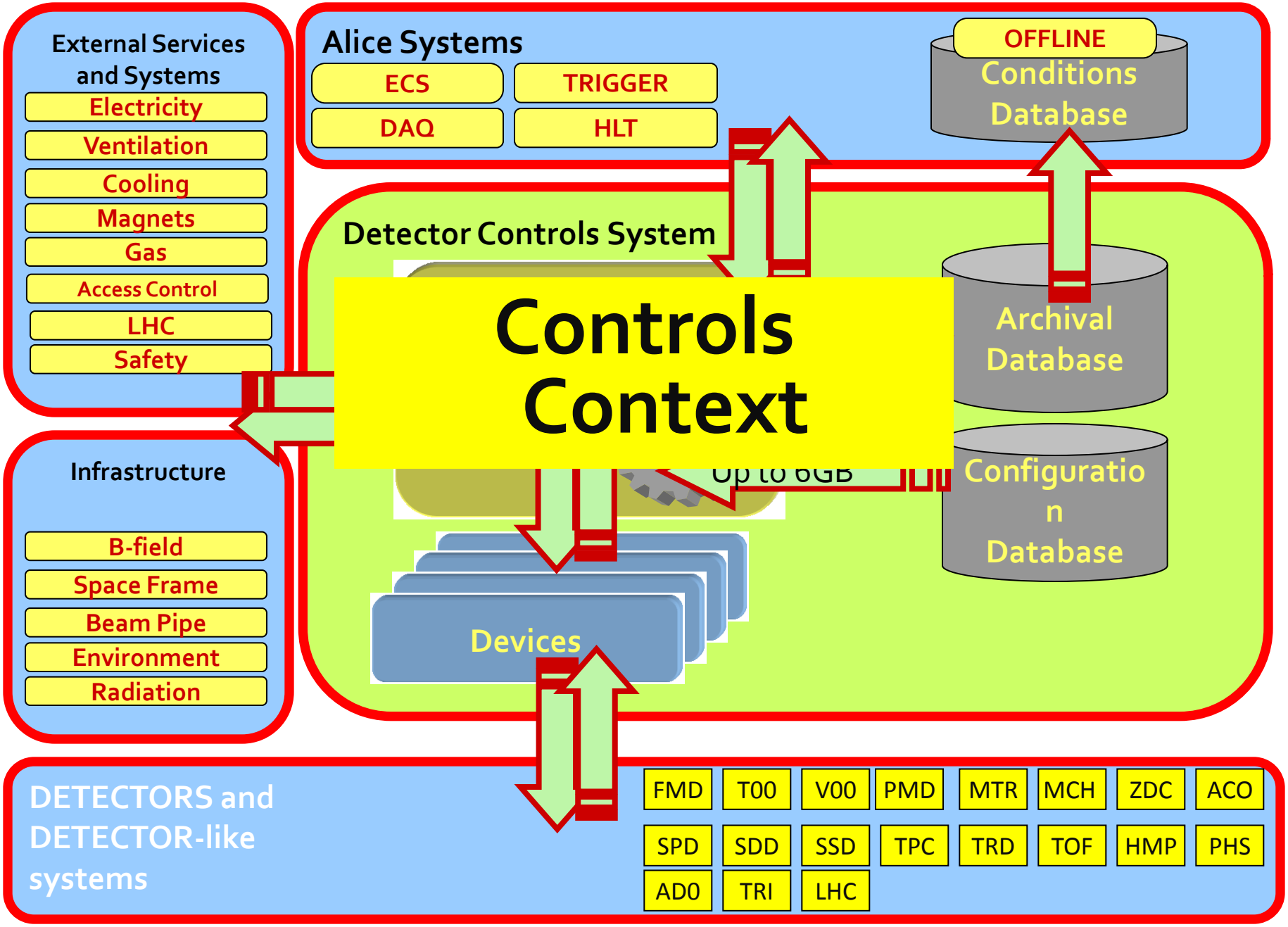
Databases in ALICE DCS

Outline

- DCS architecture in ALICE
- Databases in ALICE DCS
  - Layout
  - Interface to external systems
- Current status and experience
- Future plans

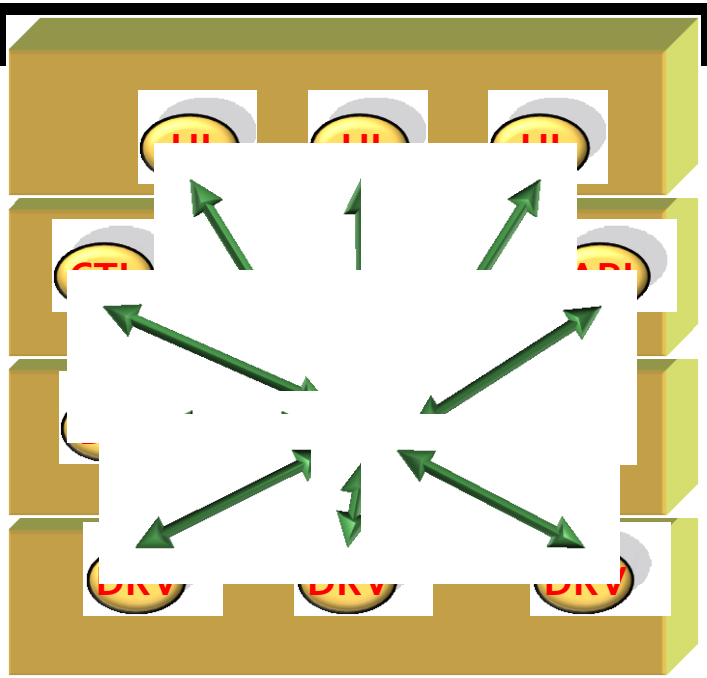
ALICE Detector Control System

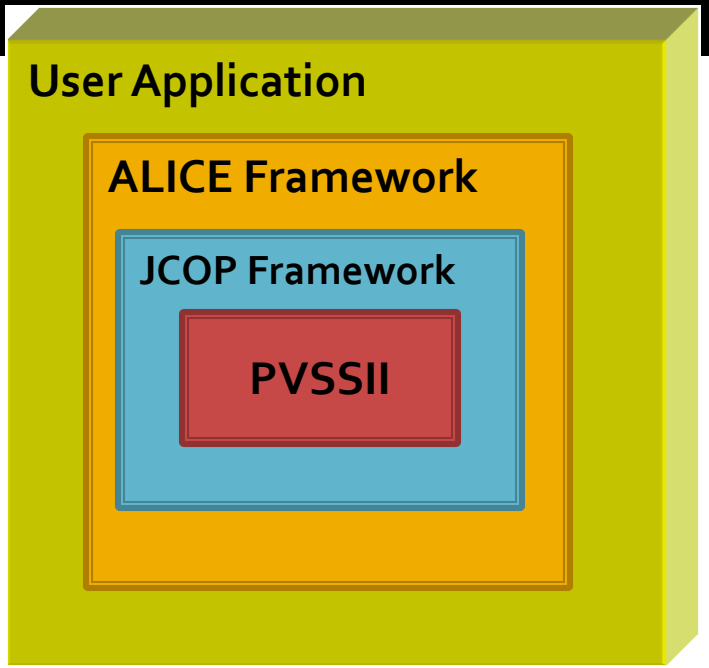
- ALICE DCS is responsible for safe and correct operation of the experiment
- DCS interacts with devices, configures them, monitors the operation and executes corrective actions
  - There are about 1200 network attached devices and ~300 directly connected devices controlled by the DCS
  - About 1 000 000 parameters are actively supervised by the system
- DCS interacts with many external systems
- Part of the acquired data is made available to offline for analysis (conditions data)


External Services and Systems
Electricity
Ventilation
Cooling
Magnets
Gas
Access Control
LHC
Safety
Alice Systems
ECS
TRIGGER
DAQ
HLT
OFFLINE
Conditions Database
Detector Controls System
Controls Context
Archival Database
Configuration Database
Up to 6GB
Devices
Infrastructure
B-field
Space Frame
Beam Pipe
Environment
Radiation
DETECTORS and DETECTOR-like systems
FMD
T00
V00
PMD
MTR
MCH
ZDC
ACO
SPD
SDD
SSD
TPC
TRD
TOF
HMP
PHS
AD0
TRI
LHC

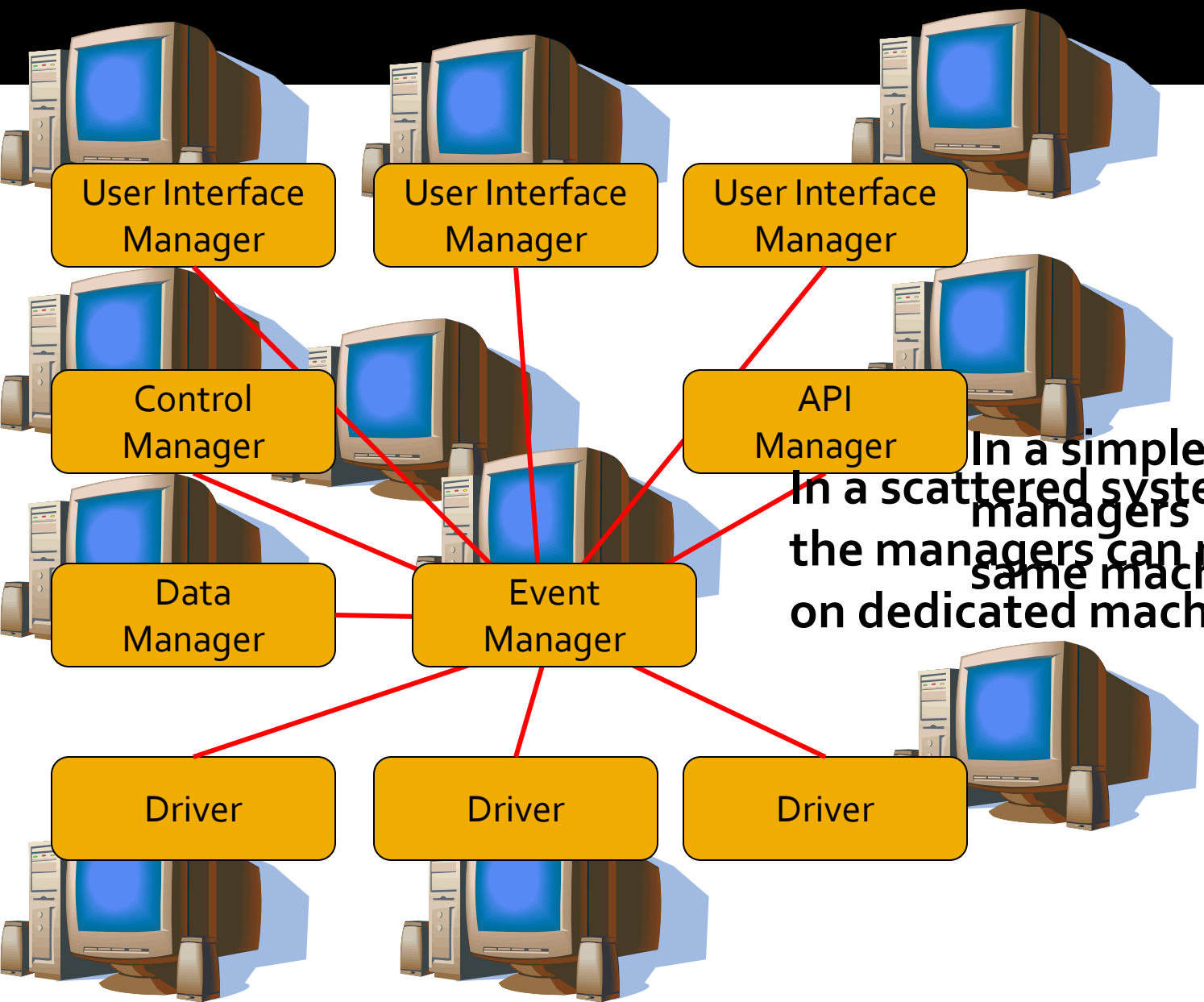
- Dataflow in ALICE DCS
  - 6GB of data is needed to fully configure ALICE-DCS for operation
  - Several stages of filtering applied to acquired data

DCS Software Architecture



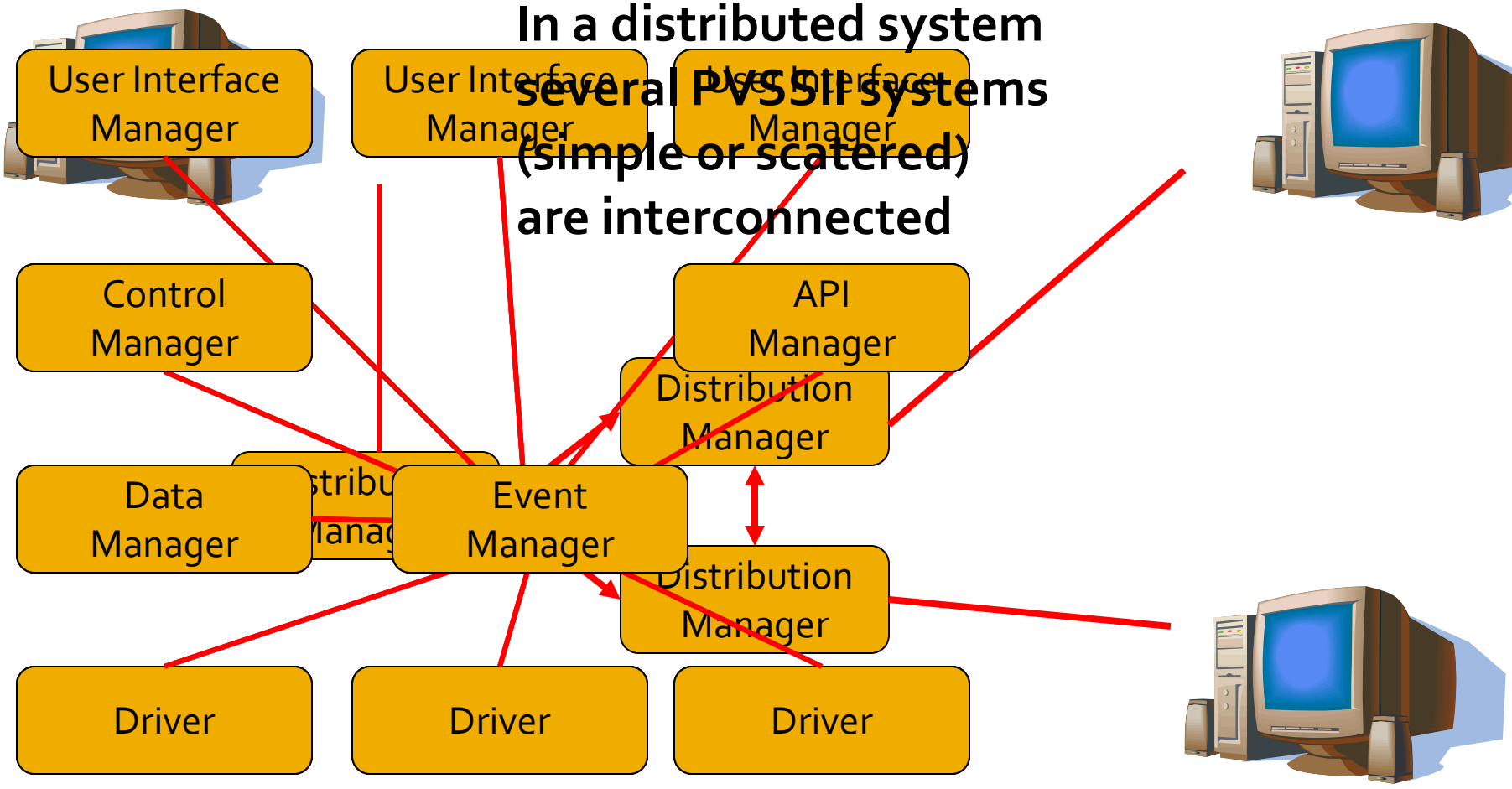


- PVSSII system is composed of specialized program modules (managers)
- Managers communicate via TCP/IP
- ALICE DCS is built from 100 PVSS systems composed of 1900 managers
- PVSSII is extended by JCOP and ALICE frameworks on top of which User applications are built

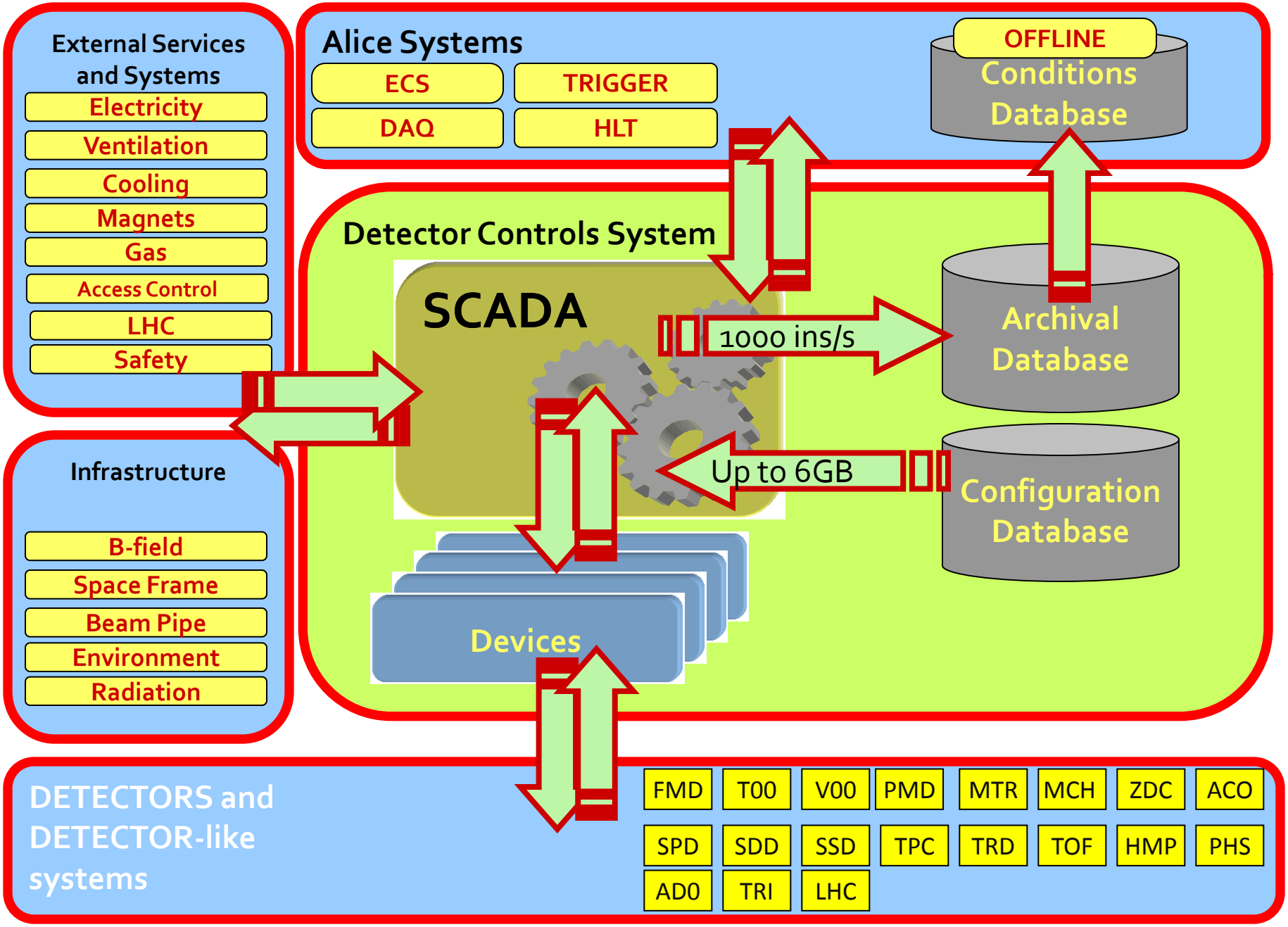


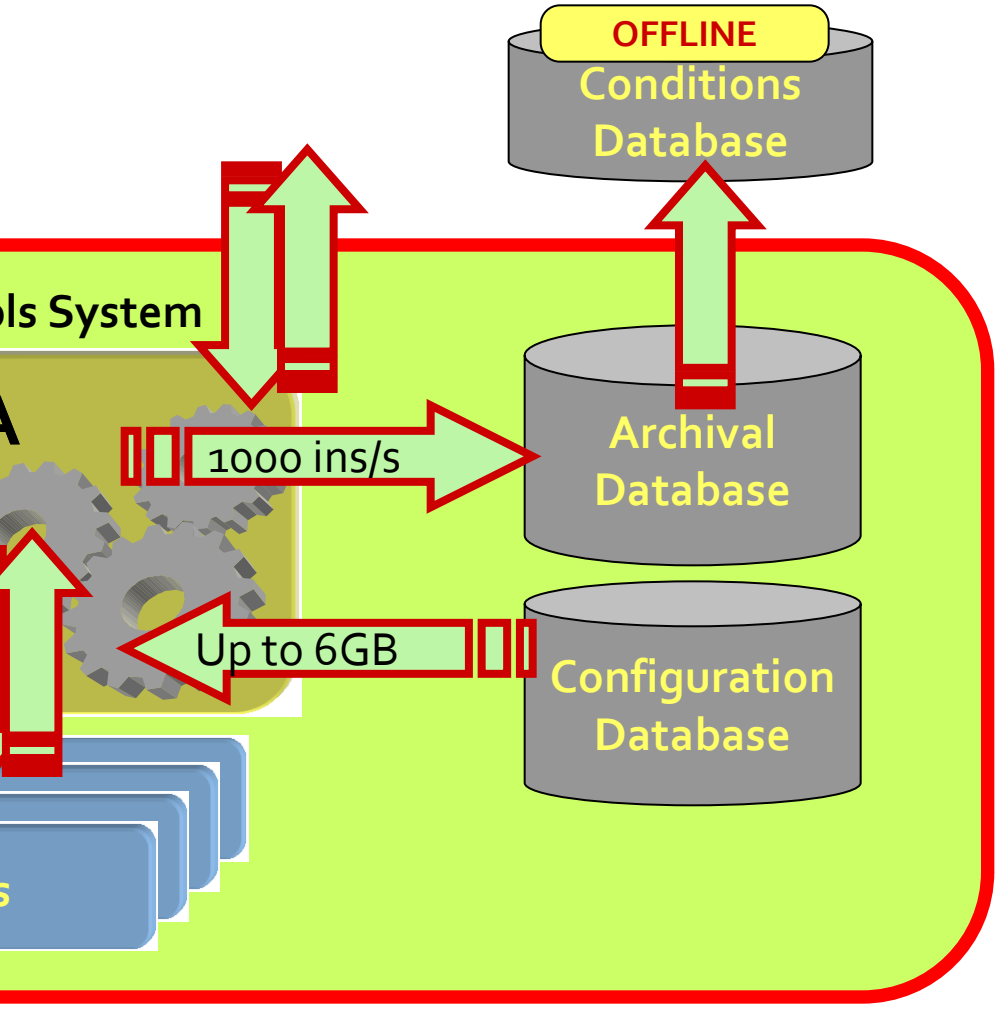
In a simple system all managers run on the same machine

In a scattered system, the managers can run on dedicated machines


In a distributed system
several PVSSII systems
(simple or scatered)
are interconnected
User Interface Manager
User Interface Manager
User Interface Manager
Control Manager
API Manager
Distribution Manager
Data Manager
Event Manager
Distribution Manager
Driver
Driver
Driver

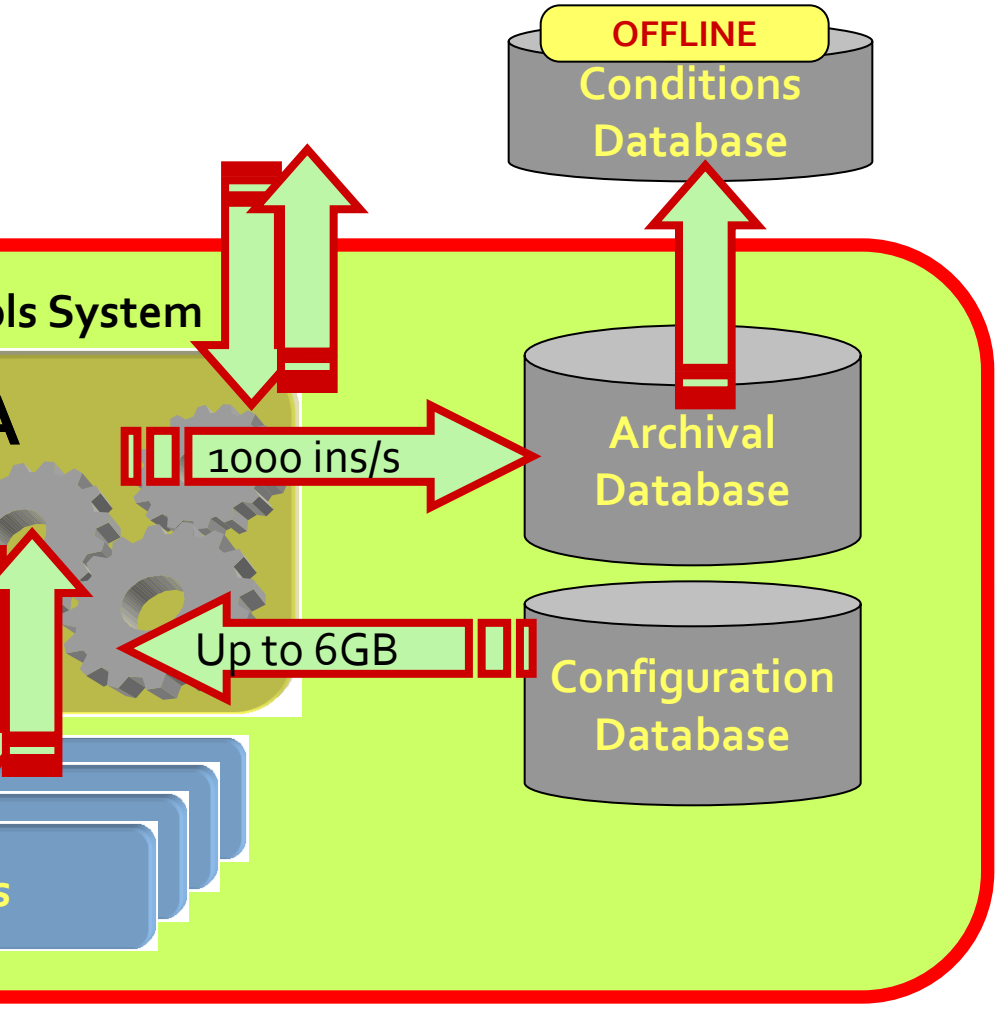
- ALICE DCS currently uses ~100 simple or scattered PVSS systems
  - Each system is unique and executes dedicated tasks
- All systems are interconnected to form one big distributed system
- Using DIP protocol, DCS directly interacts with additional highly specialized PVSS systems in ALICE (GAS, Detector Safety System..)


External Services and Systems
Electricity
Ventilation
Cooling
Magnets
Gas
Access Control
LHC
Safety
Alice Systems
ECS
TRIGGER
DAQ
HLT
OFFLINE
Conditions Database
Detector Controls System
SCADA
1000 ins/s
Archival Database
Up to 6GB
Configuration Database
Devices
Infrastructure
B-field
Space Frame
Beam Pipe
Environment
Radiation
DETECTORS and DETECTOR-like systems
FMD
T00
V00
PMD
MTR
MCH
ZDC
ACO
SPD
SDD
SSD
TPC
TRD
TOF
HMP
PHS
AD0
TRI
LHC



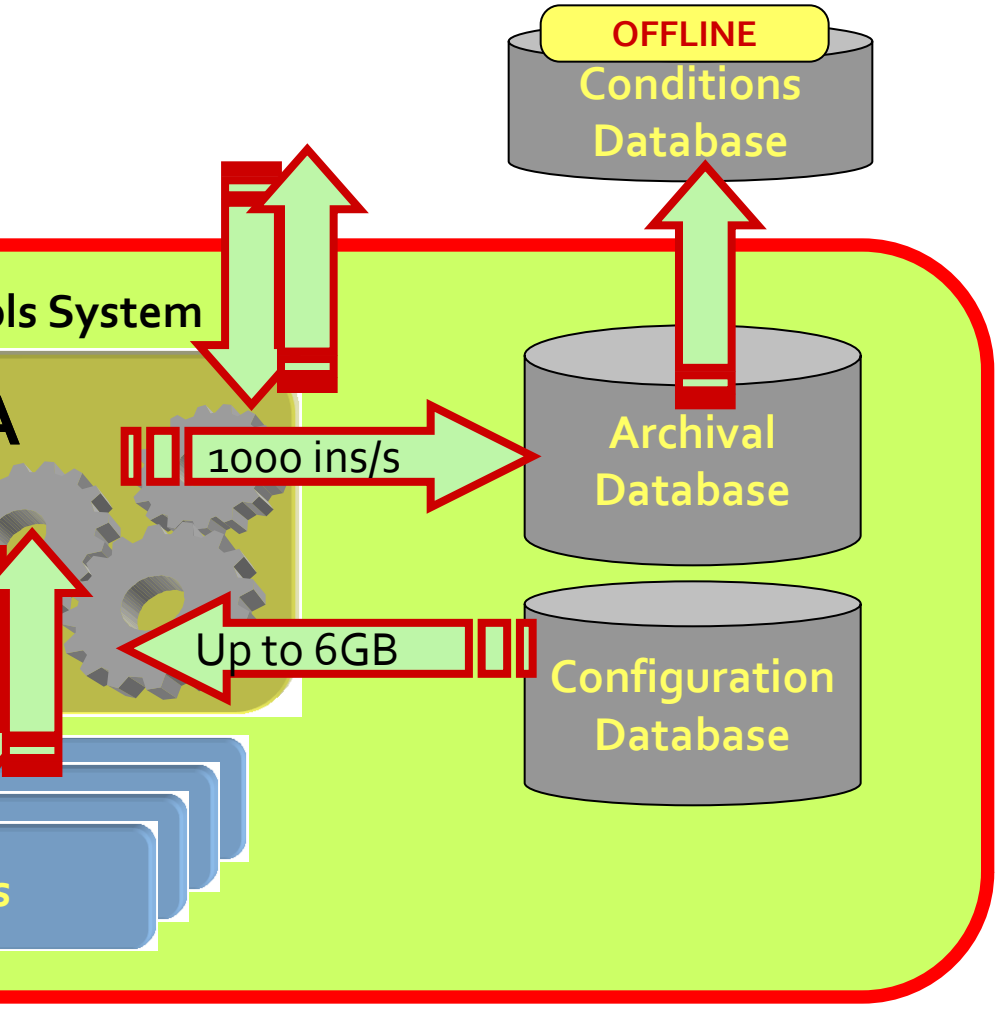
CONFIGURATION DATABASE:

- configuration of PVSS systems
- device settings
- front-end configuration
  - Stored mostly as code which is compiled online and sent to devices at the start of a run



ARCHIVAL DATABASE:

- Parameters acquired from devices



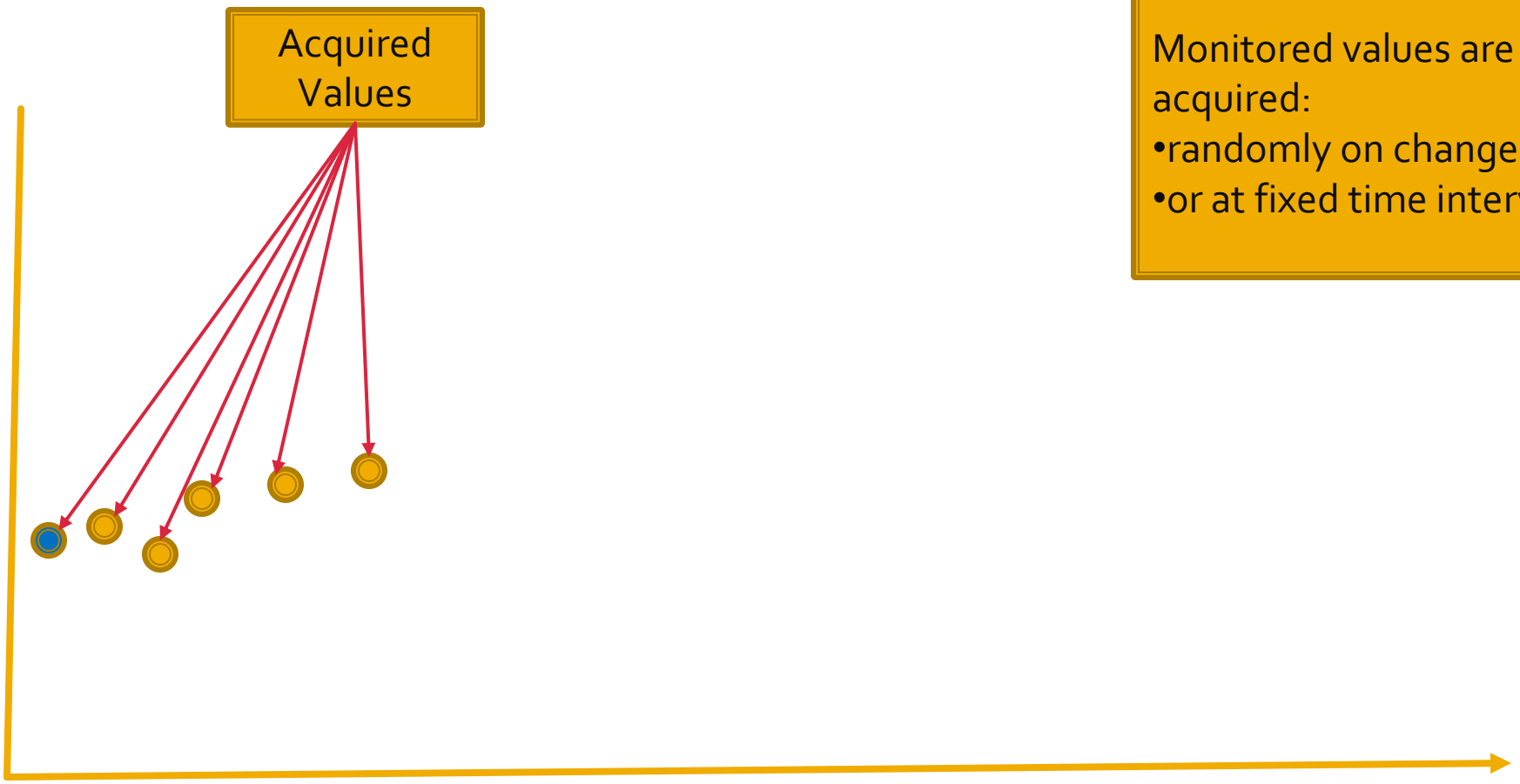
CONDITIONS DATABASE

- Stores a subset of archived data
- Implemented at OFFLINE side
- Populated after each run with data acquired during the run

Compression of acquired data amounts

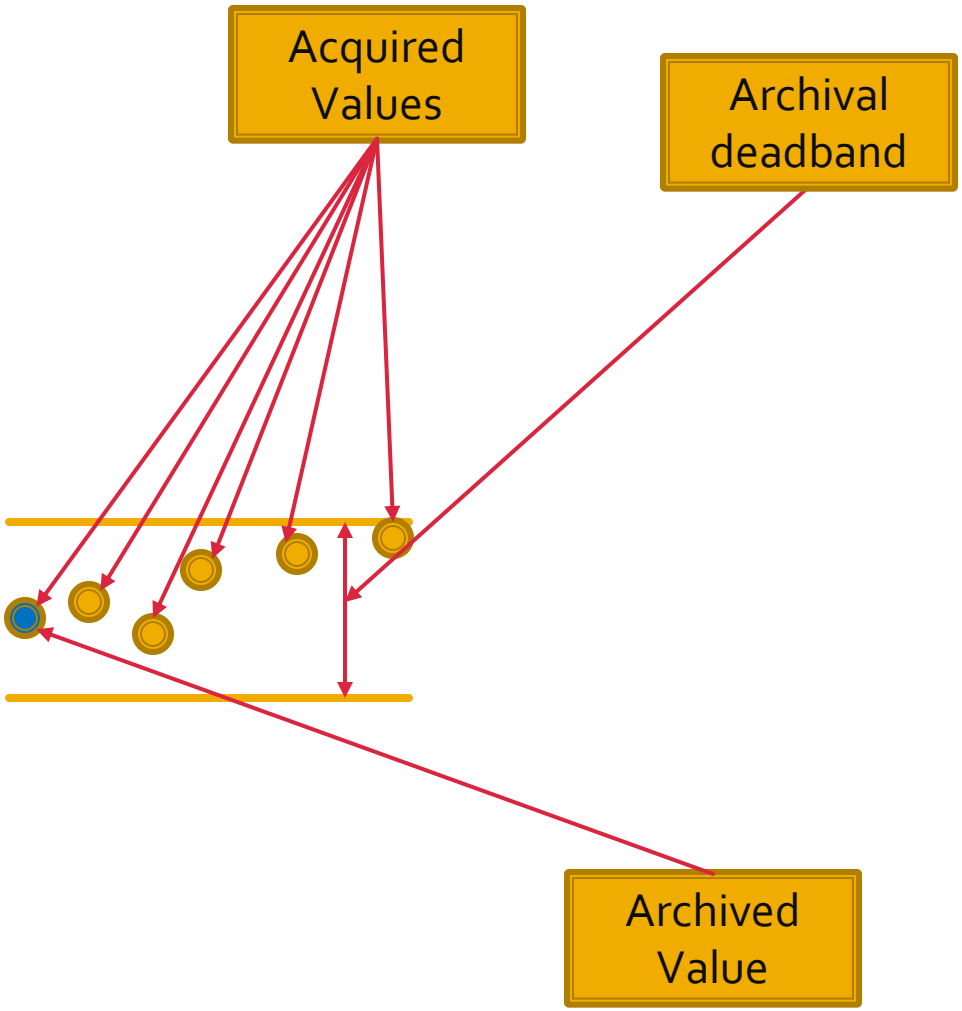
- Several stages of data compression and filtration are implemented at DCS
- Data sent to archive is filtered:
  - Only values significantly different from previously archived one are sent to archive
    - The set difference is large enough to suppress LSB and similar fluctuation
    - The set difference is small enough to preserve physics information
      - (Ex.: 0.1C change in the field cage affects track resolution in TPC but has no effect if detected in cavern environment)

Archival smoothing principle



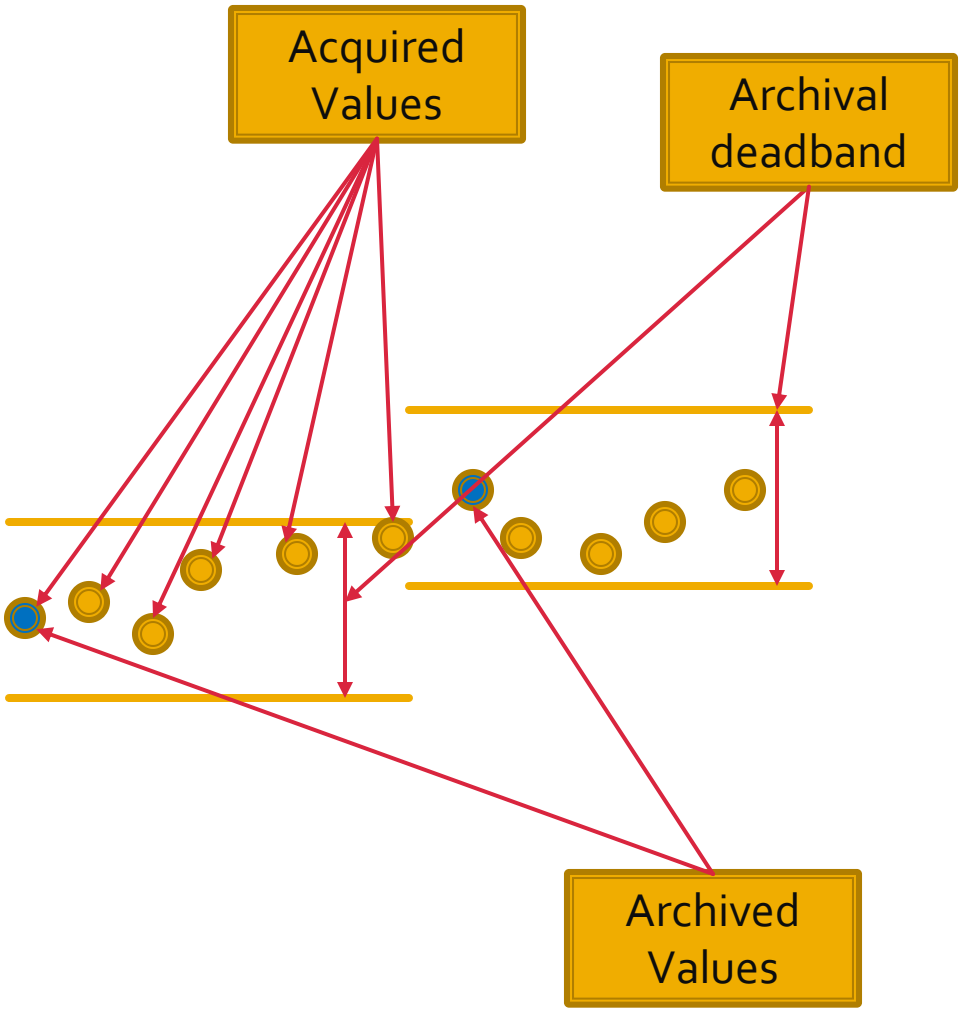
Monitored values are acquired:
- randomly on change
- or at fixed time intervals

Archival smoothing principle



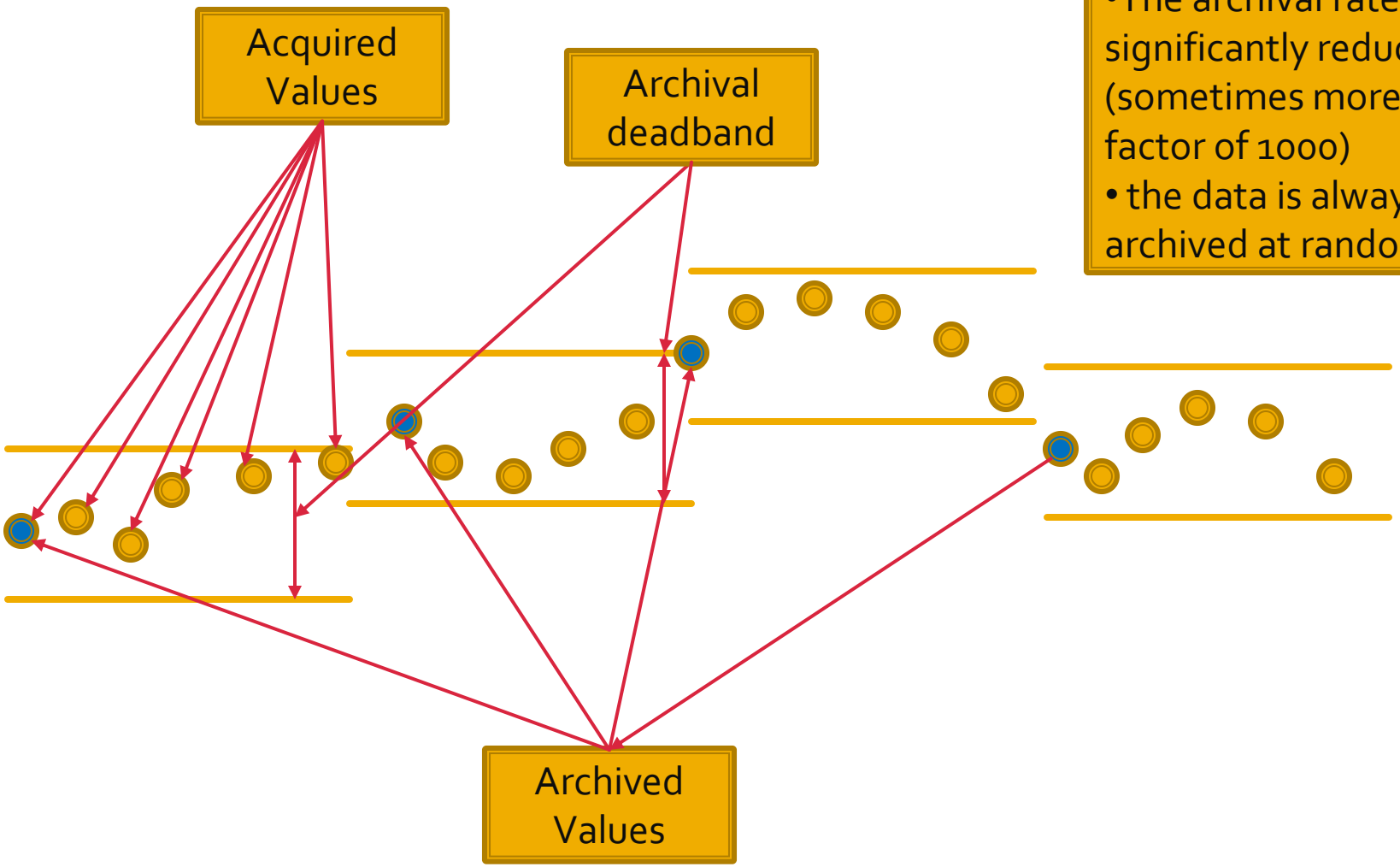
- A deadband is defined around each archived value
- Only values exceeding the deadband are sent to archive

Archival smoothing principle



If a value exceeds the deadband:
- it is archived
- new deadband is defined around it

Archival smoothing principle

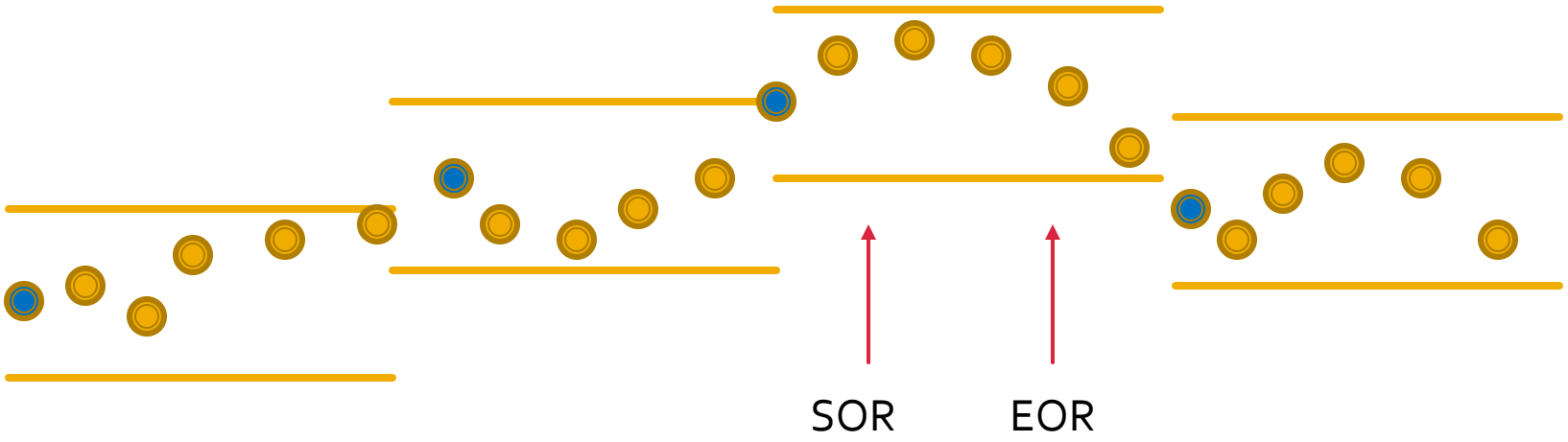


- The archival rate is significantly reduced (sometimes more than factor of 1000)
- the data is always archived at randomly

Traps of the smoothing

The data is sent to archive at random intervals

- For very short runs or for extremely stable values the archived data might be missing
  - A dedicated mechanism forces the archival of each value required by offline at the start and end of each run (SOR/EOR)

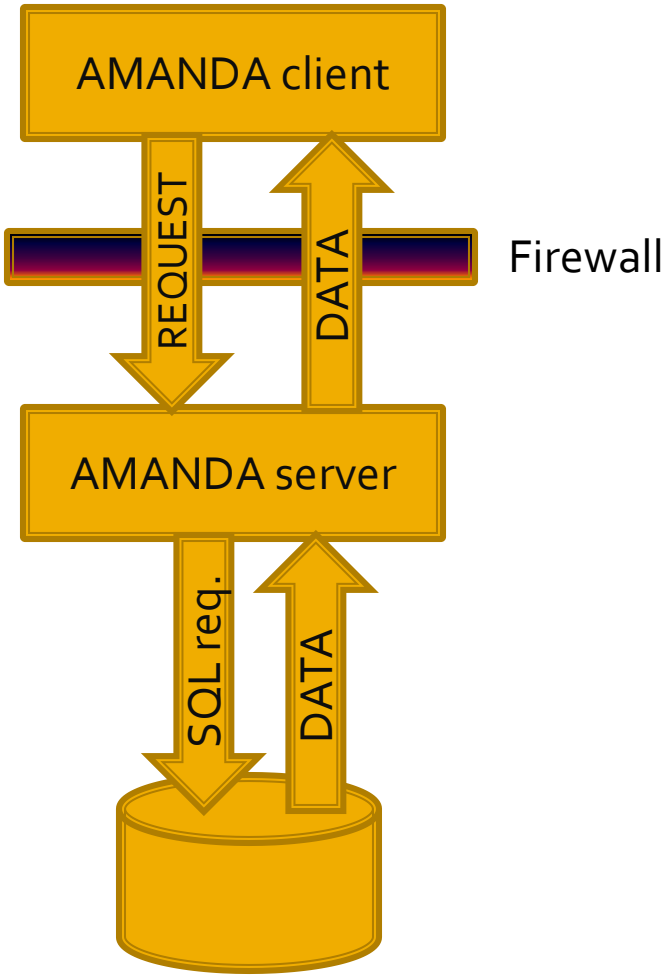


Accessing the archived data

- All PVSS system have direct access to all archived data using PVSS built-in interface
  - Online analysis purposes
  - Interactive analysis by shift crew and experts
- External system can access data only via dedicated server (AMANDA)
  - Protection of archive
  - Load balancing

AMANDA server

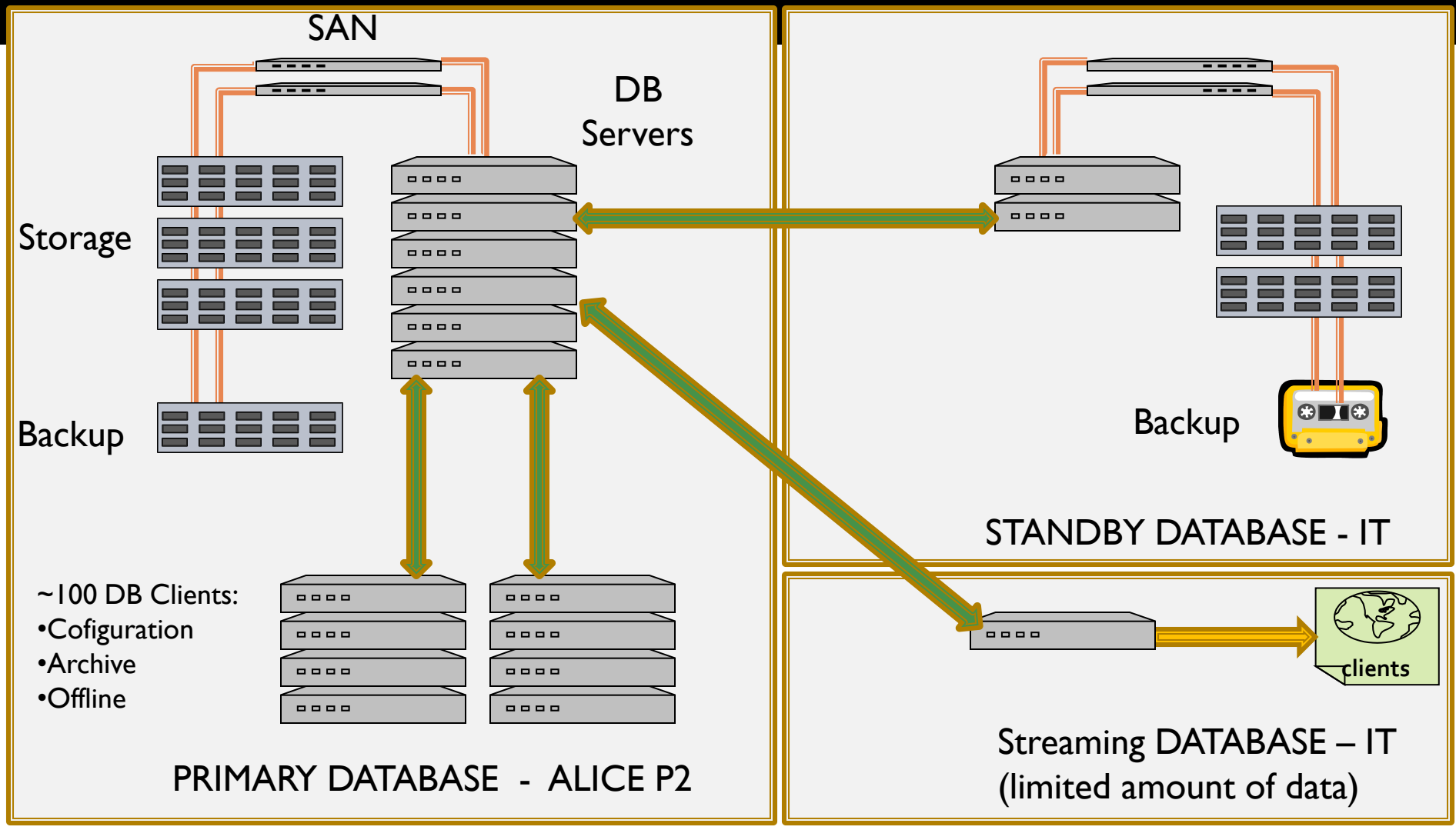
- AMANDA is a ALICE-grown client-server software used to access ALICE DCS data
  - Client sends request for data, indicating names of the required parameters and requested time interval (without knowing the archive structure)
  - Server retrieves the data and sends it back to the client
- Several AMANDA servers are deployed in ALICE
- Multiple requests are queued in the servers and processed sequentially
- In case of overload, it is enough to kill AMANDA server
- AMANDA servers operate across the secured network boundary



DCS database implementation

- The main database service for the DCS is based on ORACLE
- The DBA tasks are provided by DSA section of the IT-DB, based on a SLA between ALICE and IT
- PRIMARY database servers and storage are located in ALICE pit
- STANDBY database and tape backups are located in IT

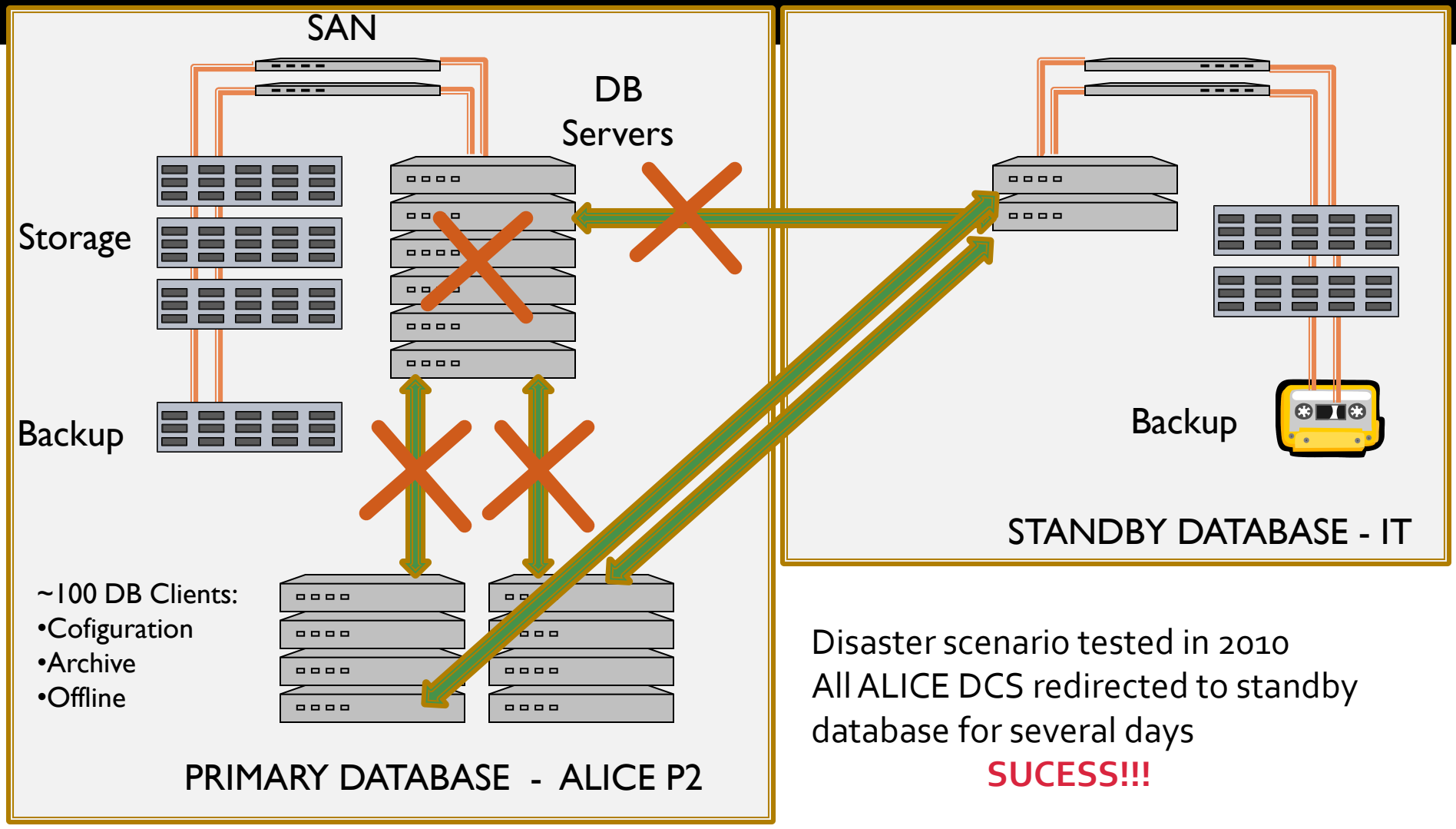
DCS Database Service



Making the service bulletproof

- The DB is backed-up directly in ALICE site to a dedicated array
  - Fast recovery
  - Full backup
- The whole DB is mirrored on STANDBY database in IT
- The STANDBY database is backed up on taped
- In case of DB connectivity problems, the clients can accumulate data in local buffers and dump them to DB once the connection is restored.
  - Lifetime of local buffers is ~days

DCS Database Service



Disaster scenario tested in 2010
All ALICE DCS redirected to standby database for several days

SUCESS!!!

Some numbers on ALICE DCS

- Number of clients: ~100
- The ALICE DCS DB is tuned and tested for:
  - steady insertion rate of ~1000 inserts/s
  - peak rate of 150 000 inserts/s
- Current DB size:
  - ~3TB
  - 2 schemas/detector

Operational experience

- ALICE DB service is in uninterrupted and stable operation since more than 3 years
  - Initial problems caused by instabilities of RAID arrays solved by firmware upgrades
  - Operational procedures fine-tuned, to match IT and ALICE requirements
    - Updates only during LHC technical stops, etc..
- The typical operational issues are caused by clients:
  - Misconfigured smoothing (client overload)
  - Missing data (stuck client, lost SOR/EOR signals)
    - However, big improvements on stability during the last year (credits to EN-ICE guys)!

Operational experience

- The smooth and stable operation of the ORACLE database for ALICE DCS is a big achievement
- **Hereby we wish to express out thanks to the members of the IT-DB DS team for their highly professional help and approach!**

Additional databases in ALICE DCS

- There are some other databases deployed in ALICE DCS, but their use is very light:
  - MySQL – for bookkeeping on file exchange servers
    - Zero maintenance, local backup solution
  - SQL Server – as a storage for system monitoring tools (MS SCOM)
    - Used as a out-of-the box solution, but growing quickly (will need to move to a new server)

Future plans

- Currently the service fulfils ALICE needs
- No major architectural changes planned in the near future (before the long LHC shutdown)
- HW and SW upgrades still foreseen

Future plans (Hardware)

- A replacement of the DB service in ALICE counting room is prepared for this year
  - Hardware (blade servers, SAN infrastructure and arrays) being currently installed and commissioned
- We wish to keep the retired DB service in operation:
  - Testing purposes
  - SW validation

Future plans (Data amounts)

- No significant increase of data volume from detectors planned before the long LHC shutdown
  - During the shutdown new detector modules will be added to ALICE. This might double the amount of data
- New project – the ALICE-LHC interface currently store data to files (luminosities, trigger counters, etc.)
  - Aim to move to ORACLE
  - Currently estimating the load – comparable with present DCS archival
  - We will probably use the retired DB for validation and tests before we make decision on the final setup:
    - Extension of the newly installed service
    - Or stay with files

Future plans (CLIENTS)

- We will focus mainly on clients
  - New AMANDA servers to be developed
  - Fine tuning of PVSS clients
- According to operational experience, we need to improve the monitoring
  - Quality of the produced data (online checks)
  - Amounts of produced data (regular scans?)
- We are getting more request for accessing the data from local analysis code
  - Currently we are able to satisfy the needs with AMANDA
  - We plan to explore and possibly extend the streaming setup

Additional "concerns"

- There are 3 big players in the business
  - DATABASE
  - ALICE experiment (clients, procedures)
  - PVSS
- Matching of various requirements might create clashes
  - Deployment of patches on PVSS or DB side might collide with experiment status (critical runs…)
  - Compatibility of patches needs to be validated in a formal way (IT-EN/ICE-Experiments)
    - ALICE can for example offer the retired DB service and the test cluster in ALICE lab for validation and participate in the tests

Conclusion

- The main DB service in ALICE DCS is based on ORACLE
- The operational experience is very positive (stability, reliability) on server side
  - Small issues on clients side, being constantly improved
- No major modifications expected before the LHC long shutdown
- Several upgrades ongoing

- Again, thank to IT-DB experts for smooth operation of this critical service