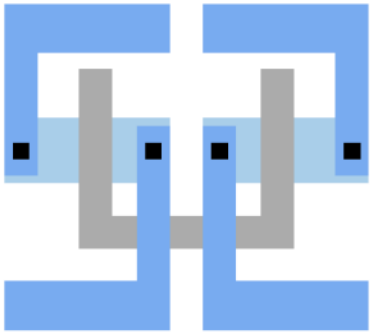
UNIVERSITÄT
HEIDELBERG

Slow-Control Status



Michael Ritzert

michael.ritzert@ziti.uni-heidelberg.de

PXD Meeting

DESY, Hamburg

23.01.2018

Outline

- LMU PS Firmware Updates
- Status at KEK
  - Phase II
  - Commissioning
- Discussion Items
  - Archiver
  - Access Controls
  - GUI

New LMU PS Firmware

- What?
  - Partial rewrite of the PS *firmware* to remove use of the XME protocol.
  - Also according changes to the IOC.
- Why?
  - There are problems where the XME protocol or its implementation could be contributing today.
  - There are foreseeable problems with the implementation when increasing the number of units.
  - All knowledge about XME was lost when Thorsten left the project.
  - Rebuilding the firmware requires specific, very old, versions of various software packages.
- Why now?
  - Analysis of tcpdump output taken during PS problems revealed some not-so-nice, deep-rooted, facts about XME.
- Note:
  - There is no guarantee this solves all (or even a single one) of the problems we see.

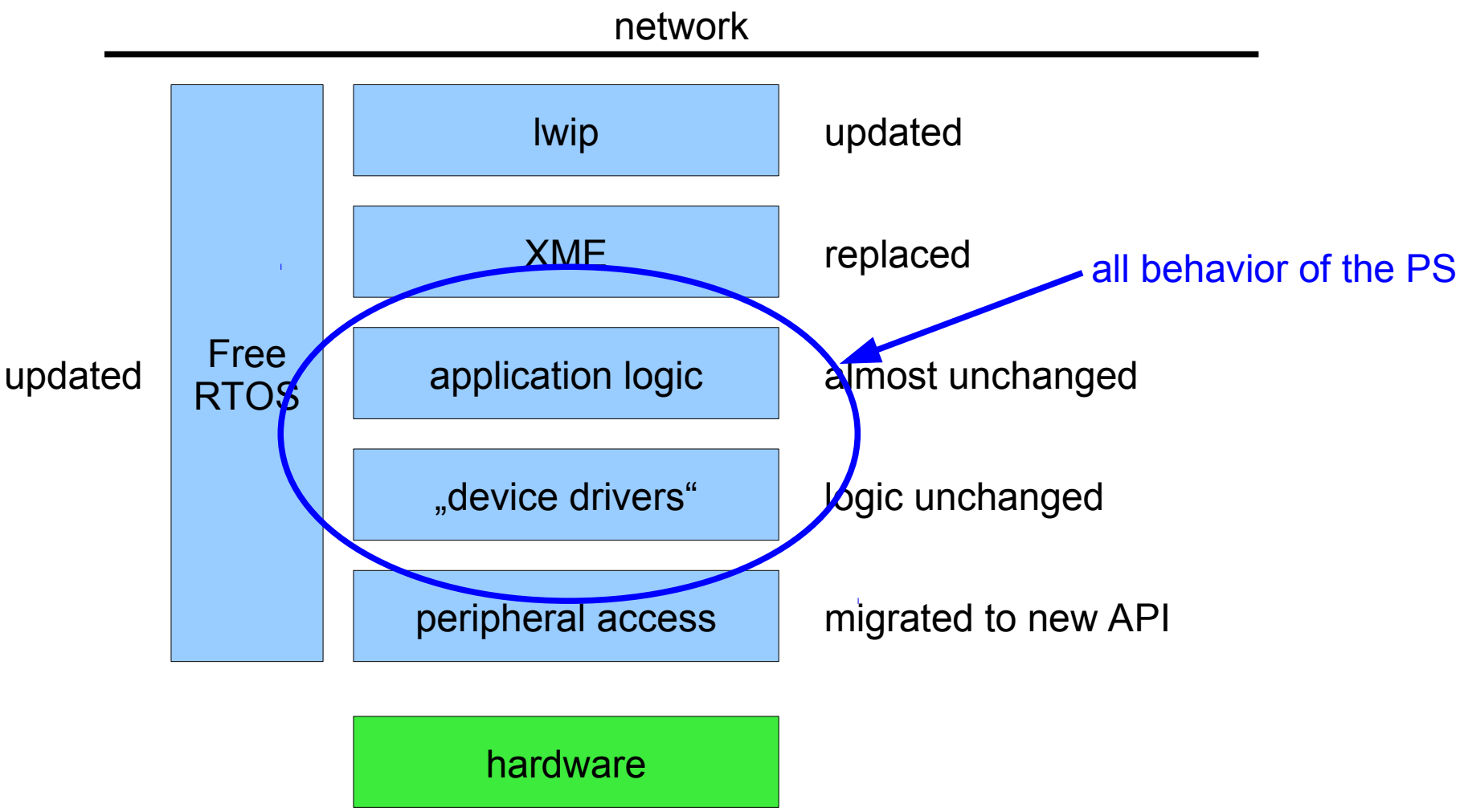
The Problems We See

- Most striking observation:
  It is not possible to reliably power up several units *in parallel* with the power-up sequence.
  Powering them one after the other is fine.
- ⇒ What's different is the number of packets on the network.
  Let's look at them.
- tcpdump shows that the TCP receiver in the PS goes down at some point.
  - Digging in would mean debugging XME code we don't know.
- It also shows that *every* command is sent to *all* units, not just the one that has to act on it!
  - This alone is a problem that has to be solved. Traffic goes with the number of units *squared*. 1600 times for 40 units…
    It's unlikely the µC could handle that.
  - We do not have the knowledge to solve it within the context of XME.

  ⇒ even without conclusive proof that it will solve any present problems, dumping XME will likely prevent problems ahead.

Is It Risky?

- **Yes**. It's the PS. We can break modules.
- But… The code consists of several blocks.
  - We need not touch the most critical ones.



What Do We Lose?

One should assume XME was chosen for a reason.

I really can't find a good one, though…

IMO: It's the wrong protocol for the job at hand.

- We do not need the flexibility it offers.
  We know we have one IOC, several PS.
- The PS µC is small. XME is targeted to µCs, but it's still a significant amount of code and overhead.
  It was always close to get the executable small enough for the µC.
  Now: Plenty of space. Even a debug build can be run.

⇒ I'd say we

- make the system scalable
- fix some bugs
- introduce a few new bugs…
- with manageable risk.

Changes on the IOC Side

- The XME layer has been replaced with just sockets.
- Further planned changes:
  - Voltages and currents are integers in the PS hardware, but presented to EPICS as floats.
    ⇒ ai/ao will become longin/longout.
  - Command PVs do not automatically reset to 0.
    ⇒ field(HIGH, “1”) on all bo records.
  - No really big changes, but actions by the users WILL be required.
    (as the minimum: The archiver complains on PV type changes.)

Status

- The firmware project has been migrated to the latest STM (µC manufacturer) development kit.
  - FreeRTOS and lwIP have manually been upgraded (to 10.0 / 2.0.3).
- A lot of code could be
  - dropped because it is now autogenerated by the STM tools.
    E.g. all µC peripheral initialization.
  - greatly simplified because the STM API offers suitable functionality.
- XME has been completely removed. Now:
  - Commands are sent to the PS directly via TCP to one unit.
    Before: Via TCP + XME to *all* units.
  - Monitoring data are sent by the PS via UDP multicast.
    Before: UDP multicast + XME.
- Interlock: The TCP connection must be active,
  i.e. send at least a „ping“ packet once per second.

What Next?

- Solve the last few (known) problems in the new firmware.
  - Uncommanded µC reboot every 300s.
  - A few more cleanups.
- Test it.
  - Voltage verification. Without and with module.
  - Sequence verification. One and multiple PS.

  Who volunteers?

- If it **doesn't** solve all problems…
  - Well… We'll have to keep on looking.
  - But now we better understand what's happening in the PS.
    Debugging should be a lot easier.

- If it **does** solve the power-up problems.
  - PSC integration can continue. Everything is in place.

Status at KEK

- PXD SC system fully up after the year-end shutdown.
  - pxdgw1/pxdioc1/pxdgw1 for the Phase II system.
    (pxdioc1 and pxddb1 are virtual machines on pxdgw1).
  - pxdctrl3/pxdctrl4 in the B3 control room.
  - Contact Botho or me for an account on PXD systems,
    Nakao-san for an account on bdaq (required!).
- Backup to KEK CC is established, but needs improvement.
  - Currently logs on to KEK CC using my account. Need to investigate how to get a group/role account @ KEK CC.
  - PostgreSQL (ConfigDB) write frequency is too low to create frequent database log dumps (after 16MB written).
    Now that's a problem you won't hear too often…
    We can switch to running periodic full dumps.
- The server for commissioning at KEK is up and running but not fully set up, yet.
  - To happen on my next trip to KEK.
- Final configuration to be prepared after the end of Phase II.
  - Also the commissioning system will be down for one or two days!

Commissioning / Phase III Preparations

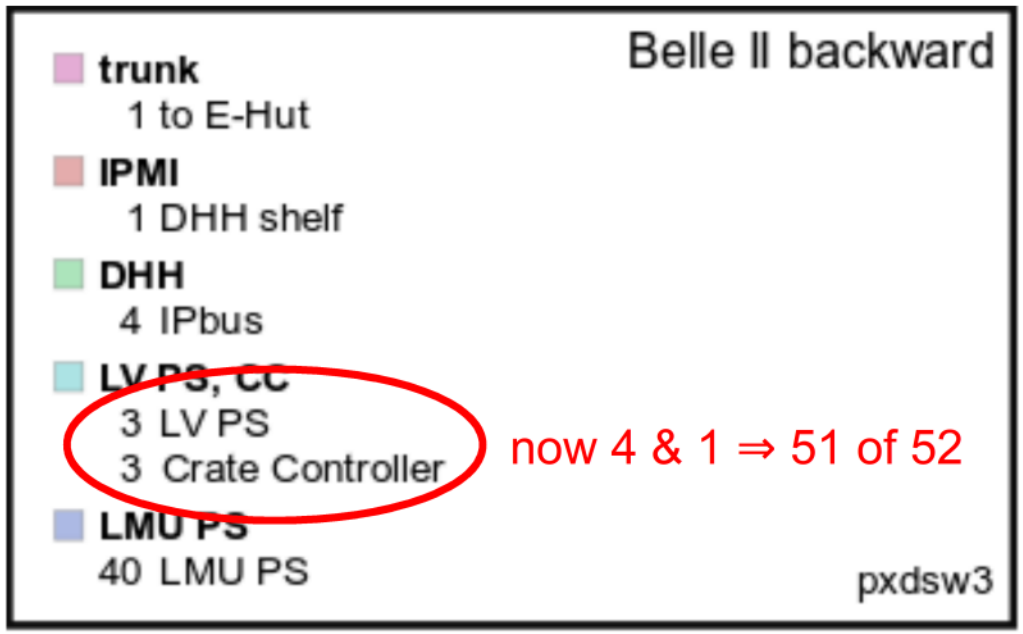
- SC-wise, we made first preparations for parallel Phase II and commissioning.
  - The phase II modules have invalid ids in the SC system to prevent confusion and PV name collisions.
  - There will be two versions of CSS for the two systems.
    If you insist, you can get the “wrong” data.
- There's a big difference between phase II and commissioning:
  - The phase II hardware is stable,
    the system is considered a production system.
  - The commissioning system will change/grow.
    ⇒ More user-serviceable parts, esp. (re)starting IOCs.

Other Systems: Commissioning / Phase III

- All SC hardware, but the Archiver server, is already at KEK.
  This is good — and bad:
- For the commissioning, we have to use alternative hardware.
  - The good old belle-iocpxd and belle-control servers.
  - The trusty „big switch“ + two smaller switches for all the PS.
- Order for the archiver server to be prepared.
  - Specs (HD, RAM) depend on logging decisions.
- Switch allocation: **Exceeded** on BWD side for phase III!
  - Some new hardware has been added since the planning was made.[1]
  - It's all just „small“ stuff (Raspberry Pis mostly), but they need their network connections.
  - I hear rumors that the racks are to be put in one place only. This would help.
  - Otherwise we need to connect devices to the FWD switch.

[1] knowing earlier probably wouldn't have made much of a difference: we planned with the biggest switch…

Phase III Switch Allocation

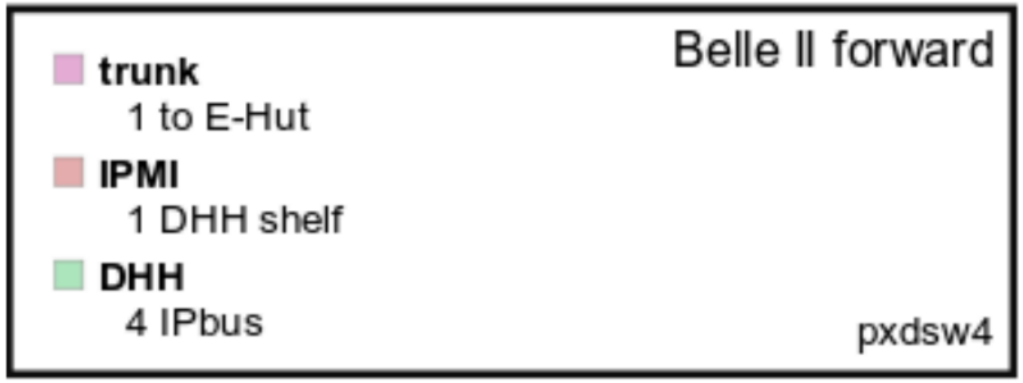


now 4 & 1 ⇒ 51 of 52

52 out of 52 available planned.
Some env. sensors to be added.

Options

- Move one rack.
- Buy another 28-port switch.
- Buy another 52-port switch.
  Part number is Cisco SG300-52.
- Run cables BWD→FWD.



6 out of 28 available planned.

Time For Some Discussions

- Archiver
  - set of PVs to archive
- Access Controls
  - type of restrictions
  - size of user groups with full access
- Shifter GUI
  - PXD top
  - subsystem top
  - + one or two levels of not-complete-expert views.

Archiver

Some statistics from the live system @ KEK:

- 1 „active“ PV can produce about 0.36 GB per year.
- For 40 units, that's 14 GB a year. For 7 years, that's 100 GB!
- So „Let's add these 10 PVs per module.“ is „Let's add 1 TB of storage.“!
- This is not exact math, but it demonstrates the problems we face.

- The appliance developers say 100k PVs per archiver is OK.
  - Extrapolating today's use, we'd have 450k for modules+DHH alone.

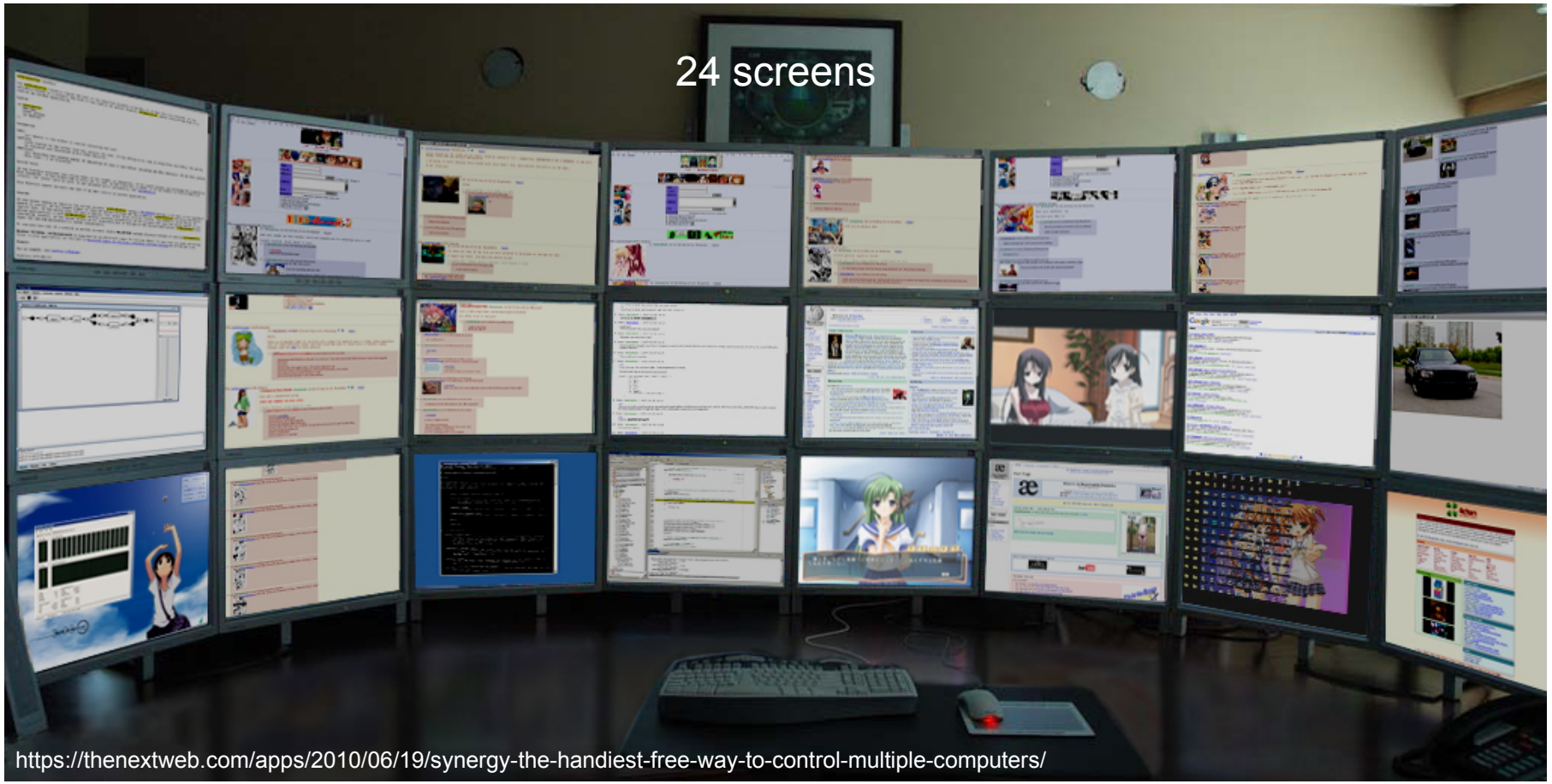
- Please, let's be honest about what's required, what's „nice to have“, and what is just noise.
- „Everything is required“ is the wrong answer…
- All voltages and currents are logged ⇒ O(70) PVs/module ⇒ 7 TB.
- Let's assume we have another O(10) TB available.

Access Controls

- In case you haven't noticed yet:
  Write access to PVs at KEK is restricted.
- Current implementation:
  - Several defined PV groups, roughly per subsystem.
  - Per-user access level defined by rules on the pxdgw1 CA gateway.
- This is not a complete lockout (it's easy to circumvent).
  - But it puts the bar higher to do thing you maybe shouldn't do.
- Goal: Protection of the system (e.g. bad voltage settings) and the operation (e.g. accidental misconfiguration of DAQ parts).
- We have to discuss how to handle this in the future.
  - At least all module-related settings should be constrained to a small group of experts.
    Of course, this does **not** include operating the system via RC/PSC!
  - ONSEN has PVs that are expert-only.
  - For IBBelle, almost all PVs are expert-only.
  - FOS data are read-only for all but the experts.
  - Others?

PXD GUI Over Time

- For one module, we had one screen for CSS.
- For two modules (TB), we had two screens for CSS.
- For four modules (Phase II), we have one BIG screen for CSS.
- This could be our control room for one side (FWD or BWD)…



https://thenextweb.com/apps/2010/06/19/synergy-the-handiest-free-way-to-control-multiple-computers/

GUI in the Future

- Overall design: As outlined in the GUI guidelines.
  - Small, easy to comprehend OPIs.
  - Navigation instead of more open OPIs.
  - Trustworthy alarm system instead of eying everything.
- Shifter views: Assume no knowledge about PXD at all.
  - Simple yes/no or value stable/moving displays.
  - No abbreviations nobody knows.

- Proposal for “home” display:
  - PSC and RC states. Global and PXD.
  - IBBelle (“cooling!”) state (setpoint, “setpoint reached”).
  - top-level PXD DQM plots.
    - E.g. some trends of values that should be stable over time in stable run conditions ⇒ fluctuations could indicate problems.
  - PXD global alarm state.
  - PXD log messages with high (≥ warning) level.

Second-Level GUI

- Per “subsystem” second level.
  - Definition of “subsystem” is not so obvious, here:
    - A shifter wouldn't call for “ONSEN problems”, but for “DAQ problems”.
    - ⇒ First step must be to identify which of the subsystems as we know them is involved.
- To be accessed by the shifter in cause of problems that cannot be immediately resolved (via guidelines in the alarm system).
- Content: The data you would ask for when woken from your dreams in your role as standby expert at 3am.
  - Imagine a call from somebody without any knowledge about PXD, i.e. a shifter.
  - To narrow down where the problem is, what are the first questions you would ask? The answers should be found on this page.
  - Of course, eventually you will have to guide the shifter further down towards the data relevant for this very error or log on yourself.

Requested SC Development Time Allocation

- Fine-tuning of the PS IOC's voltage/current monitoring.
  - Compares a channel's output (:cur) against the settings (:set).
  - Goal: No warnings generated during routine operation.
  - Get estimated time constants for voltage steps (required to silence mismatch alarms).
  - Identify well-defined special situations (e.g. current limit active) during normal startup/shutdown.
  - Time: Possible at any time. Schedule one day with one module. @DESY? @KEK?
  - May need to be redone after sequence changes.

- Migration from phase II + commissioning to final setup.
  - Plan with two days of SC service downtime @ KEK.
  - Time: After the end of phase II, before the start of phase III.

Thank you!

Archiver Stats @ KEK

PV Name	Storage Rate (KB/hour)	Storage Rate (MB/day)	Storage Rate (GB/year)	Details	Quick chart
PXD:MI:CSAA_USA15_LiveCounter:PosSt	64.83308323489817	1.5195253883179258	0.5416277018906669		
PXD:O03S1:MEMORY:OCC	43.20507135397601	1.0126188598588128	0.36094324594576827		
PXD:O03S1:FREEBUFS:cur	43.1255854202106	1.010755908286186	0.36027920559029086		
PXD:H1132:dhc_data:CNT:cur	43.11842779985433	1.010588151559086	0.3602194094912758		
PXD:O01M1:LUTR-NoDatcon:CNT	43.086523761331584	1.009840400656209	0.3599528771870276		
PXD:O01M1:LUTR-OPS:CNT	43.084858341766974	1.0098013673851636	0.35993896396053193		
PXD:H1131:dhc_data:CNT:cur	43.068438142178266	1.0094165189573032	0.35980178654239814		
PXD:O03S1:LUTR-OPS:CNT	43.03807519217932	1.0087048873167028	0.3595481287798794		
PXD:H1131:dhp_data:CNT:cur	42.999052787053685	1.0077902996965706	0.35922212830981276		
PXD:O03S1:LUTW-OPS:CNT	42.95931493880761	1.0068589438783033	0.3588901508941218		
PXD:H1132:dhp_data:CNT:cur	42.6563905818213	0.9997591542614367	0.3563594641654535		
PXD:O03S2:MEMORY:OCC	42.24756960903665	0.9901774127117966	0.35294409730449783		
PXD:H2132:dhc_data:CNT:cur	40.92254056998758	0.959122044609084	0.3418745569163239		
PXD:H2131:dhc_data:CNT:cur	40.91133618197005	0.958859441764923	0.3417809533634735		
PXD:O03S2:LUTR-OPS:CNT	40.79362752095408	0.9561006450223613	0.34079759319644715		
PXD:O03S2:FREEBUFS:cur	40.75378311445562	0.9551667917450537	0.3404647255731881		
PXD:O03S2:LUTW-OPS:CNT	40.610085701665554	0.9517988836327864	0.3392642505136397		
PXD:H2131:dhp_data:CNT:cur	40.51594225888324	0.949592396692576	0.3384777585867092		
PXD:H2132:dhp_data:CNT:cur	40.491576473958105	0.9490213236083931	0.3382742022627573		