Business Continuity at DESY

… a collection of themes and thoughts

… covering among others measures, procedures and dependencies

Peter van der Reest, Yves Kemp, DESY IT

Hepix Spring 2014, 21.05.2014

General DESY risk assessment

- DESY performs a general, yearly risk assessment
  - This is a formal process
  - Risks from all possible fields, including financial and other external ones
  - Also covers IT

- Risk assessment performed by separate DESY entities
  - E.g. administration, machine control, …
  - Not always formal process
  - Written/Oral reports from units to directorate after incidents

- “DESY is an experiment-oriented laboratory”
  translates into “IT is second in priority for e.g. power and cooling after accelerators and experiments”
  - Does not mean that IT is neglected!

ISO 27001 certification

- Background: DESY project management office is asked by funding agencies to certify that its procedures and infrastructures conforms to ISO 27001
  - Includes IT … which is most of central IT
- External consultant first evaluating status and estimating work and costs of such a certification
- So far interviews with all relevant groups within IT
- First impression is that many requirements concerning setup and workflows are met, but formal documentation of processes should be enforced

Network and IDS

- Scanning networks and testing ports
  - Get to learn who does what -> “Who is running https server? HeartBleed”
  - See differences, e.g. when malware listens on ports
- Efforts to separate different networks
  - Or define relations between networks
  - Incident containment
- Investigations into flow monitoring
  - Checking for unusual patterns in network traffic
- Network interventions and glitches have huge impact

- Linux: Dedicated intrusion detection software on (most) systems
- Windows: No dedicated IDS, anti-virus also catches some intrusions

CC operation and Communication to users

> Operational aspects

- Control room, workdays 8:00-20:00 with operator-on-duty
- On-call operator all other times

> User Consulting Office (UCO)

- Generates user documentation
- Handles first level requests and trouble shooting
- Organizes communication with users in disaster situations

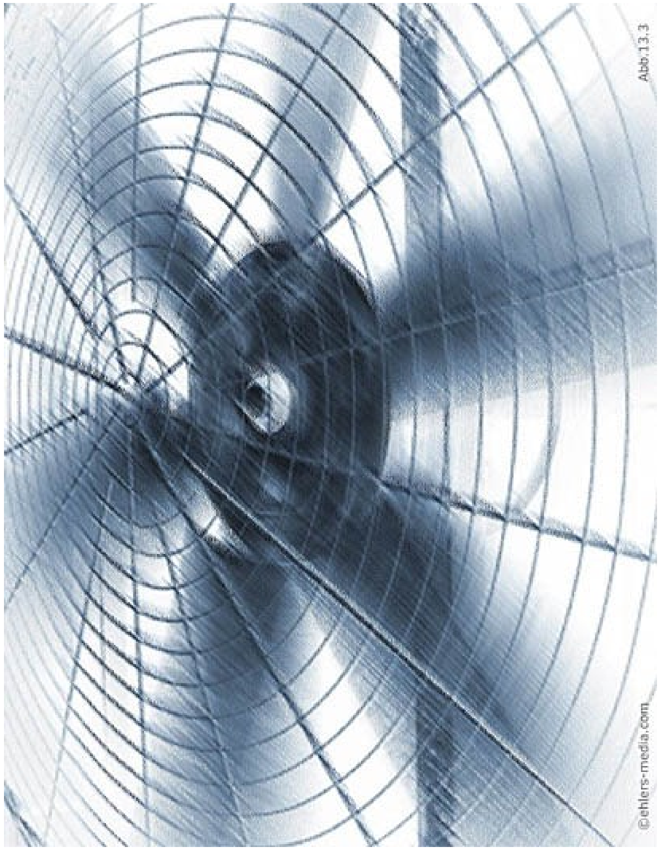
  e.g. also by pinning paper information about network outages to entry doors of buildings…

Computing Center and Power

- Three independent power lines to HH campus two used by IT in room 1 & 2 (same building)
- These two lines shared with other groups on campus
- Two independent lines with generally good and stable quality
- Have battery powered UPS – but mainly to flatten out voltage fluctuations or very short interruptions (~20 minutes)
- ~2 years ago, we had disturbances in internal power distribution system – complete black-out … other independent power feeds would not have helped

Cooling

- Climate (also in CC) not under IT control
  - ... The same for power distribution
  - More communication with infrastructure groups needed to make them understand our needs for separation and decoupling (which is more expensive)
- Cooling redundancy: Cold water ring
  - On HH campus, 8.4 MW total, 2 MW for IT
  - Two inputs: overhauled HERA cooling and new highly efficient PETRA III cooling
  - Currently ring not closed – more like a bus
- Cooling redundancy: Distribution in the CC
  - Recent incident: Work on increasing redundancy of in air cooling for room 1 resulted in cascade of short-circuits that stopped cooling of water-cooled racks im room 2
  - (Some) water-cooled racks react very fast to cooling disturbance because of small amount of air

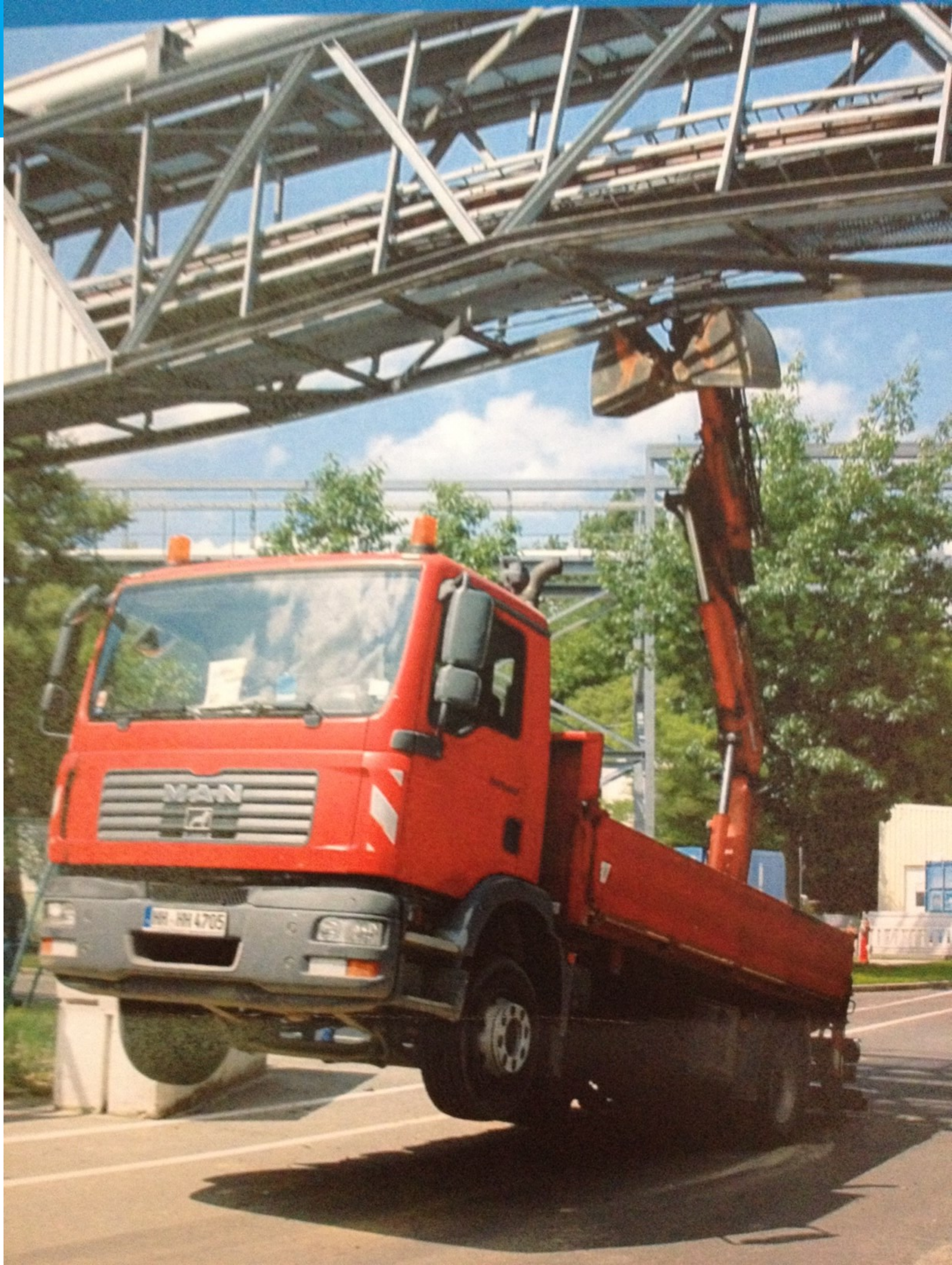


General comments on cooling and power

- IT depends on other DESY departments for climate and power
  - … recall “DESY is an experiment-oriented laboratory”
  - Generally good service and fast reaction
- Climate and power: Historically grown infrastructure
- Chasing single-point of failures?
  - We will discover unknown single-point-of-failures
  - Probably better to accept this fact and concentrate on optimizing reaction handling

One event we failed to prepare against (7/2013)

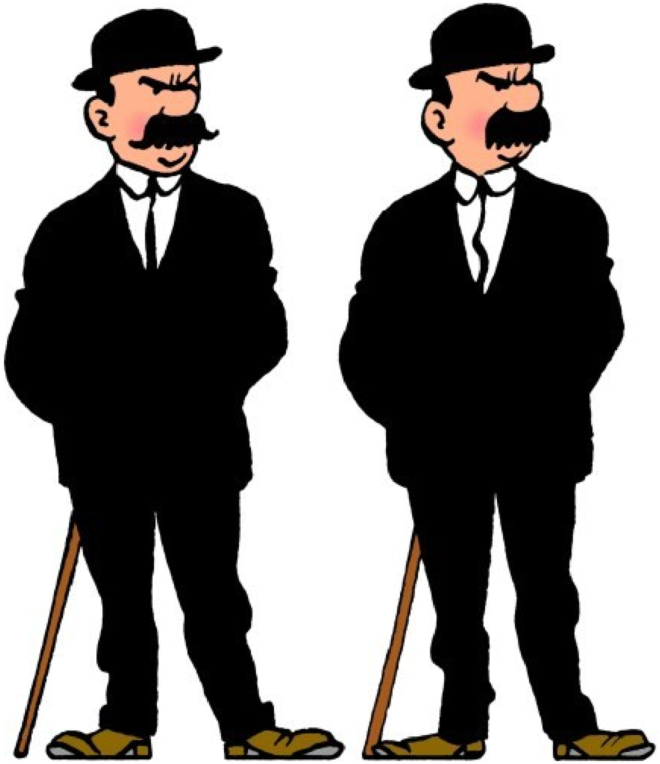
- One of our two lines was cut
- Transformer on second line overheated
- On batteries for ~20 minutes … power came back in last second
- No set procedures, but the whole crew reacted well – we survived!
- … and we were lucky: The helium line above was not in use …



High-Availability, Server and service redundancy

High-Availability & Redundancy :

- Whenever possible, set up systems in high-availability mode
- Using VMware + Cisco UCS to build infrastructure for mission critical applications
  - … spread over Computer Rooms 1&3 (~500m apart)
  - … e.g. for EDMS, Person management systems, Mail, …
- Classic Cold/Warm/Hot standby
- Load-Balancer with fail-over: F5 & Poise (own development, advanced metrics)
- Fail-Over cluster etc. whenever necessary and possible



Configuration management

- General tendency towards common and widespread tools
  - WDS/WSUS for Windows well established
  - Migration to Puppet for Linux (actually consolidation of Quattor/Salad+WBOOM/FAI)
- Introducing version control management in configurations puppet
  - Enables roll-back, auditing, …
- Automate configuration as much as possible
  - Fast reinstall with guaranteed results
- Make secret handling processes (pw, keys, certs,…) audi
  - See Sven’s talk
- Using vanilla distributions with only minimal changes
  - E.g. discontinue HEP ENV / HEP X11

Backup & Archive & Tapes

- Backup & Archive & Tapes:
  - For TSM backups data is saved redundantly in two locations (HH and ZN)
  - For selected archive data sets two copies are held:
    one online in silo, other offline in former atomic shelter
  - Other methods of redundant data keeping are considered, e.g. cloud storage syncing: although this is not backup it might help users with broken notebooks
- Desaster recovery
  - of notebooks&desktops: TSM backup methods are sufficient (or not needed: $HOME on network FS)
  - of RAID-Arrays without copy/backup: Very rare, rapid escalation to external data rescue experts … costly but usually successful

Human Continuity _ 1

- as workload is high, for some services we do not have n+1 (n=1) redundancy
  - even when desirable, budgets won’t allow for it
- absence or exit of colleagues can leave holes
  - illness
  - leaving DESY usually before new recruitment has finishe
  - spreading tasks over remaining staff will only work for limited time
- standardization, use of widespread tools and products
  - Allows for hiring external fire-fighters

Human Continuity _ 2

- past cases have raised awareness of importance of up-to-date documentation
  - In disaster situations
  - Knowledge transfer after changes in personnel
- and even more of the independent check that this documentation is understandable and complete
  - many minor details are taken as common knowledge (by the author…)
- unfortunately, this also increases workload
  - but can well be built into operating procedures

… being a Scientific Computing Center

- In the end, our mission is to serve Scientists and enable Science

Need to find a balance between

- **Stable, well documented infrastructures and workflows**
- **Flexible environment to ad-hoc**
  - Deploy non-standard hardware and software
  - Bypass procedures in case of needs from scientists
  - … and later include in standardization and documentation

This is what distinguishes us from commercial hosters