

VIRTUALISATION WITHIN THE CONTROL SYSTEM ENVIRONMENT AT THE AUSTRALIAN SYNCHROTRON

Andrew Starritt, Nicholas Hobbs, Ulrich Felzmann

Synchrotron Light Source Australia, 800 Blackburn Road, Clayton 3168, Victoria, Australia
e: uli.felzmann@synchrotron.org.au

INTRODUCTION

The Australian Synchrotron is a 3 GeV third-generation synchrotron light source built in Clayton near the Monash University Campus in Melbourne, Victoria and opened on 31 July 2007. Since 2012, the Australian Synchrotron is operated by the Australian Nuclear Science and Technology Organisation (ANSTO) under the name Synchrotron Light Source Australia (SLSA).

The core of the control system for the 100 MeV linac, the 100 MeV to 3 GeV booster, the 3 GeV storage ring and the 9 beamlines consists of more than 500 Input Output Controllers (IOCs) running EPICS. Most of those IOCs are Linux hosts (mostly CentOS), but also ~ 100 Libera systems, some Windows hosts, VME-bus computers and PLC-based controllers with EPICS embedded.

Each beamline reside in a separate subnet and VLAN and uses its dedicated set of isolated services with one service per host. Due to that philosophy, many of the IOCs do not have any special hardware attached. Those so-called "soft IOCs" and IOCs with only serial devices attached, that can be easily connected via a Ethernet to serial converter, can be easily converted into virtual machines (VM).

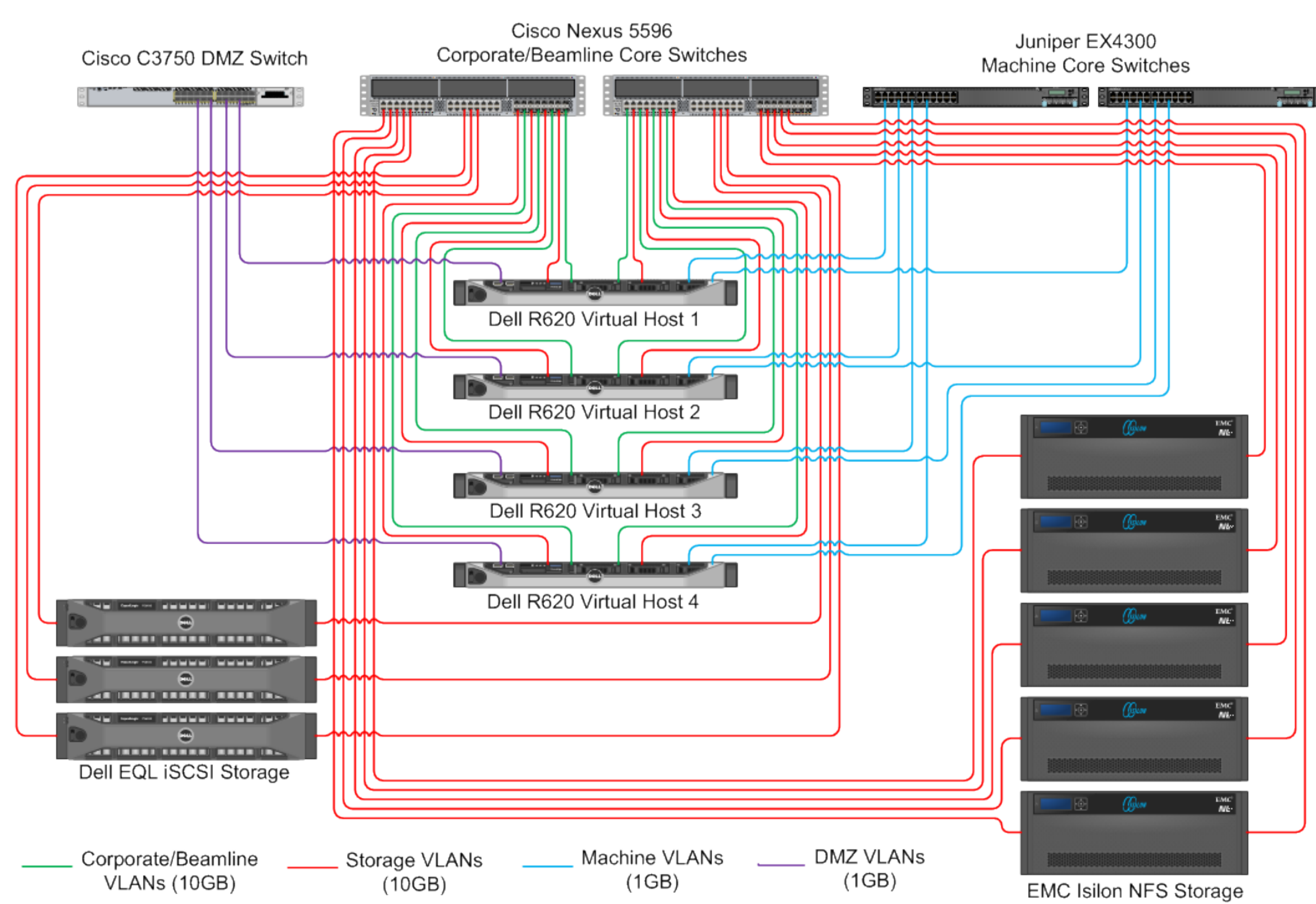
VIRTUAL ENVIRONMENT

Virtualisation, which refers to the creation of an environment that acts like a real computer with an operating system, offers a great way to increase the efficiency and availability of computing services while reducing risk and the total cost of ownership.

In June 2012, the Australian Synchrotron purchased its original hardware for the virtualisation platform. Since then, the initial hardware configuration has been extended substantially and consists currently of the following individual components:

- 4 x Dell R620 hypervisors, providing in total 8 Intel Xeon CPUs (E5-2650 @ 2 GHz) and 768 GB RAM
- 3 x clustered Dell EqualLogic iSCSI devices (104 TB storage group in total), hosting the virtual disk images and frequently changing data for corporate applications that require daily replication (backup)
- 5 x EMC Isilon NL400 nodes (674 TB in total), hosting the virtual disk images for IOCs via NFS as well all the user data from the beamlines
- 2 x CISCO Nexus 5596 switches, connecting the beamlines, corporate and storage networks
- 2 x Juniper EX4300 switches, connecting the machine
- 1 x CISCO Catalyst 3750 switch, connecting the DMZ
- VMware ESXi 6 with resource pooling enabled so the management of VMs can be delegated to the individual beamlines as a "self service".

Apart from services for the control system environment, the virtualisation platform also hosts database and application servers for the corporate business units (email, file server, www, etc) as well as network management (DHCP, DNS, NTP, PXE, etc) and monitoring (Observium, Zabbix) services for the facility.



Virtualisation Platform of the Australian Synchrotron. Not shown are the separate backup systems for the iSCSI and the NFS storage systems.

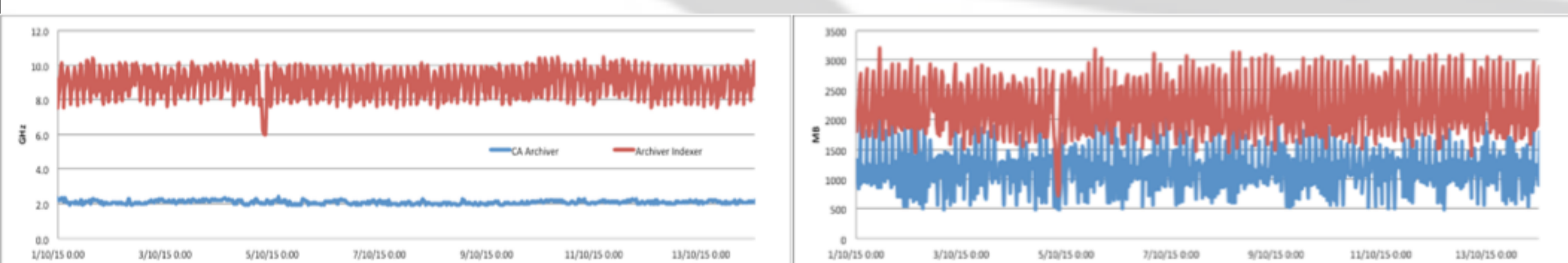
APPLICATIONS

Currently more than 200 virtual machines are active, of which approximately 50 are virtual IOCs (vIOC). In addition, there are 19 EPICS gateways, 3 EPICS CA Archivers, ~ 15 build boxes, as well as VMs hosting version control, bug tracking and continuous integration testing systems.

Many of the vIOCs have been created from physical machines using the vCenter Converter from VMware, which works successfully most of the time. Some physical hosts with an old Linux operating systems such as CentOS4 or Debian 3 required a manual conversion.

The median run-time configuration parameters for a vIOC are:

- 1 vCPU (9.5 MHz), 672 MB RAM allocated, 13 GB disk space used (20 GB thin-provisioned)



The most resource intense VMs are the EPICS Archiver Indexer (red) and the EPICS CA Archiver (blue).

Further to **hardware virtualisation**, some VMs provide further nested virtualisation in the form of **operating-system-level virtualisation**.

- 19 EPICS Gateways providing Channel Access between the different subnets are operated as LXC instances on a single VM to reduce resource consumption.
- About 10 former Cosylab μ IOCs (MCS8) driving various beamline optics and originally built on Debian 3 are now run within a chroot environment on individual CentOS 6 (32bit) VMs. Due to the unavailability of the source code of the EPICS applications, copying the entire file-system into a chroot jail proved to be the most suitable method to replace those critical μ IOCs, that started failing recently.

LESSONS LEARNT

With the introduction of a new technology, there are lessons to be learnt as issues and defects with the new setup and configuration are identified.

- VMs are abstract for many users and administrators, and problems with the VM itself are often wrongly attributed to the hypervisor. End users often find comfort in the fact that they can 'see and touch' physical hardware, and are resistant to migrate to a virtual platform, which is perceived as being beyond their control.
- A failure in one hypervisor can crash many services, and the recovery can take a considerable time. Also, expert knowledge and presence is often required after a disruptive event involving the hypervisors.
- Dell Equallogic iSCSI storage is very cost-effective and provides useful features like the ability to mount iSCSI volumes within a VM, and replication and snapshotting for disaster recovery planning. However, it does not provide good redundancy and the iSCSI protocol is not very tolerant of network disruptions. EMC Isilon storage provides a higher level of redundancy and the NFS protocol is more tolerant of network disruptions. The EMC Isilon cluster has proved to be a more robust storage platform for virtual control systems.
- Initially we deployed VMware vCenter Server on MS Server, with the database containing all the information about running VMs being hosted on a separate MS SQL server. Since switching to the latest OpenSUSE based vCenter Server Appliance with inbuilt vProgress database, the stability of the Virtualisation Environment has greatly improved.
- Virtual machine snapshots are very useful for rapid rollback and are the traditional method for backup processes to gain access to virtual disks. However VMware snapshotting has proved to be detrimental to EPICS processes. Instead, we utilise storage level snapshots as a means of gaining access to virtual disks for backup purposes without disrupting VM operation.

CONCLUSION & OUTLOOK

Virtualisation has significantly improved the resilience of the control systems at the Australian Synchrotron while removing effectively high risk IOCs and reducing the overall hardware footprint tremendously. New services can be deployed and tested very easily and in a very cost-effective matter. The live migration of control systems between different hypervisors works well and VMware's resource scheduler manages the allocated resources effectively so a self-service was offered without the risk of one beamline consuming all the resources. To further reduce the efforts required to backup all the control system VMs frequently, a configuration management solution is being investigated so VMs can be rebuild from scratch on a short time-scale.