Planning for KAPB move

8 May 2017
Danny Jacobs

Summary and Goals

1) On site visitors: Aaron Parsons +2 UCB students. May 15- May 25
   Off site help:
   Danny Jacobs (US Pacific time: Early morning, late evening SA time)
   Peter Williams (after May 14, US Eastern time 3 hours closer to SA)
   Jack Hickish (US Pacific time)
2) Goal: Move all HERA equipment from CMC to KAPB
3) Goal: Move SKA network attachment point to single port in KAPB.
4) Goal: Prepare HERA container for its move.

Plan points:

1) HERA currently uses 16 IPs. The eventual goal is to put the private HERA subnet behind an externally managed login node. Once we have this we will put our mission critical systems on the HERA subnet. For now they must remain accessible from the padloper login node.
2) Currently HERA uses two fibers one dedicated to internal data traffic (often maxed out to 1Gbps) and one public KAT. The KAPB has only one fiber available to HERA. We will therefore combine traffic onto a single fiber. The HP switch in the KAPB will form a HERA pvt and SKA public vlans and extend those to its mate in the HERA container, both vlans will be split evenly across the available ports.
3) Our strategy for managing the SKA-HERA interface in the KAPB and in the HERA container, is to provide switches at each end which will be configured to provide access to the HERA and SKA vlans. This removes the necessity of altering the configuration internal HERA switches which are heavily customized.

Sources

http://herawiki.berkeley.edu/doku.php/network_v1
github.com/ProjectFiles/network_power_config/
This document:
https://docs.google.com/document/d/1zxxAMYZ-B0MZXDAQixQ8gBf8-FulslYb3mZqtvzy_V8/edit?usp=sharing

Schedule for May 2017 Trip

Monday May 15:

- Aaron P., +2 students arrive on site

Tuesday May 16:

- Label, document, move as much CMC as possible.
- NB: Check for flashing disks
- End of day goals:
    - Primary

        - KAPB: HP switch, netgear switch, qmaster, and 1 still machine
        - HERA container: HP switch installed and fiber connected
    - Secondary
        - More of the CMC as time permits

Wednesday May 17:

- Peter Macfarlane day trip.
- Configure two HP switches for HERA Pvt and SKA public Vlans. Verify correct connections between machines on each vlans, and outside for the “public machines”.
- Aaron and team to assist Peter.

Thursday May 18:

- Setup remaining RTP/Librarian machines in KAPB.
- End of day goal: return to fully operational correlator, RTP, Librarian operations.

Friday May 19:

- Margin

Saturday May 20:

- Margin

Sunday May 21 - 26

- Deconstruct HERA container in prep for move. Store correlator machines etc in KAPB?

Supplied by SKA

Fiber connection to the SKA Switch (The Cisco switch)
What kind of connector on our end?
Peter Macfarlane for switch setup! Site trip 17 May 2017?

Rack #11?

Power PDUs

New things

HERA Gateway switch (HP 2810-24G, shipped from UCB March 2017)

HERA Container switch (HP 2810-24G, ordered to UCB on 4 May 2017)

Things moved from the CMC

The plan is to move everything without any config changes. Add the HP switch as an interface layer which provides both HERA.pvt and SKA public vlans.

RTP:

Qmaster (SKA “public”, HERA private, and IPMI)
Still1-4 (RTP Private)
Cask1,2 (RTP Private)

Librarian:

Pot 6 (SKA “public”, HERA private, and IPMI)
Pot 1(SKA “public”, HERA private,and IPMI)

Power switch (APC)

Dell power connect switch

HERA container
Paper1:

Eth0 gigabit (public KAT 192.168.216.100 - port)
Eth1 Gigabit (10.0.1.1 -- roaches/X-engs, netboot, “HERA pvt”)
Eth1:1 Gigabit (10.0.100.1-- X-box IPMI)
eth2 mellanox 10G (10.0.2.1 -- roach/xeng debug?)

Two racks of correlator hardware.
Etc other things.

Notes:
The X engines are unreachable from outside the HERA container switch. This might be on purpose?

We require 1000BASE-LX -- i.e., 1GbE over pairs of fiber at 1310nm. Will be using HP J4859C transceivers --https://h10057.www1.hp.com/ecomcat/hpcatalog/specs/provisioner/99/J4859C.htm

Questions for meeting with Peter Macfarlane

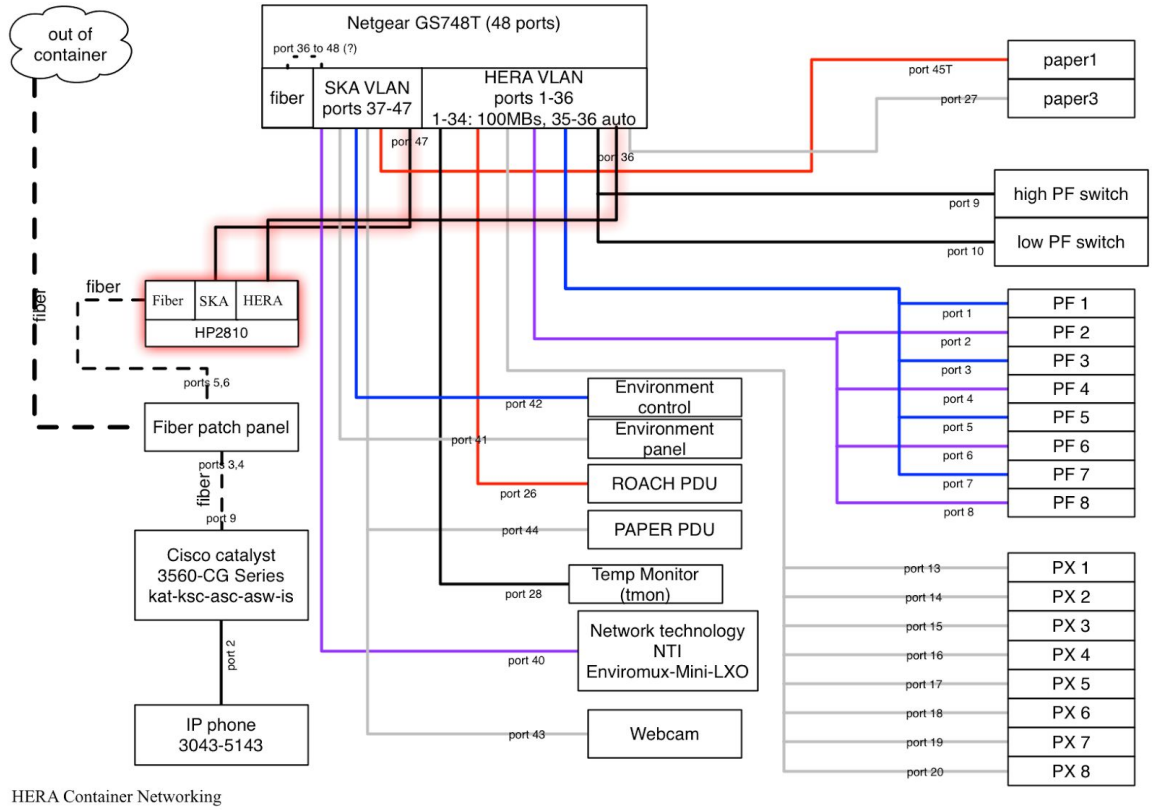
What is the new SKA subnet and what are our IPs?
Verify staged plan for ~16 current public ips, then transition to login node +NAT+DNS to have fewer publics.

SKA IPs

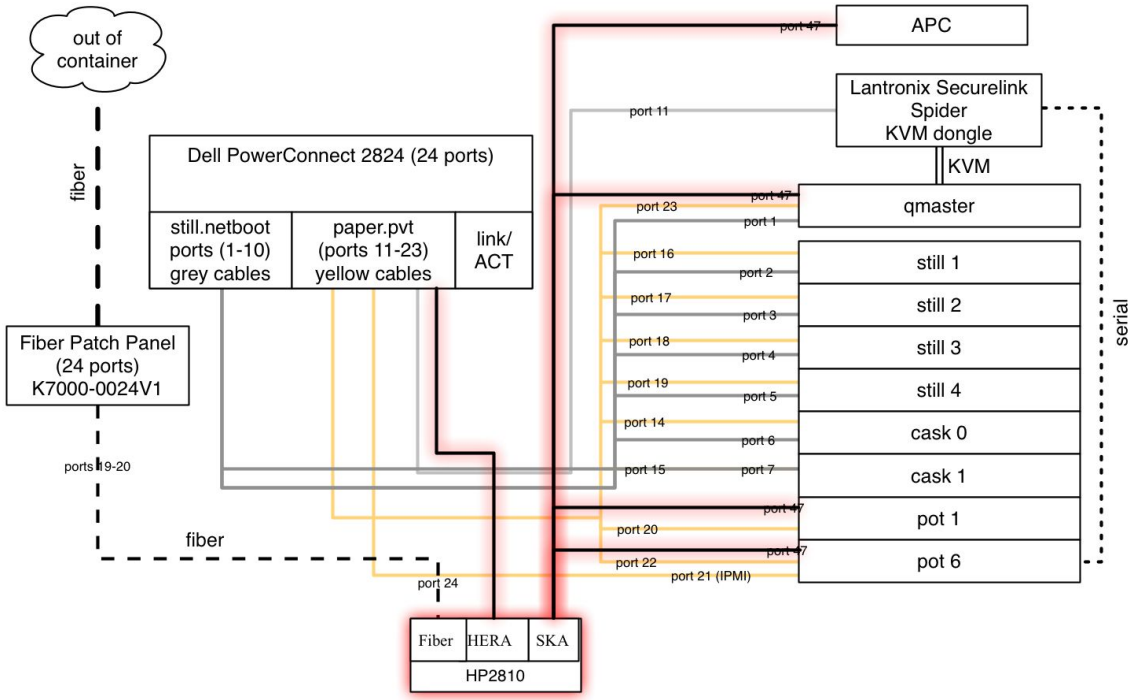
We currently have ~15 “public” machines.
Needs for public IP are: needing to see out, can be fixed with NAT and DNS on a HERA login node. Other times to have a clear path for data movement (ie avoiding ssh tunnels). Mission critical things (power strips, web cams) are put on public ips to avoid having to tunnel through the HERA login nodes which are currently both in containers (which can overheat etc).

If we can setup reliable login nodes off site which provide NAT, DNS etc, then most of them can move to private. The only things which will need true SKA access will be things accessing SKA services. Data routing for maximum bandwidth is an open question and depends on where the HERA login nodes get hosted.

Network diagram after change

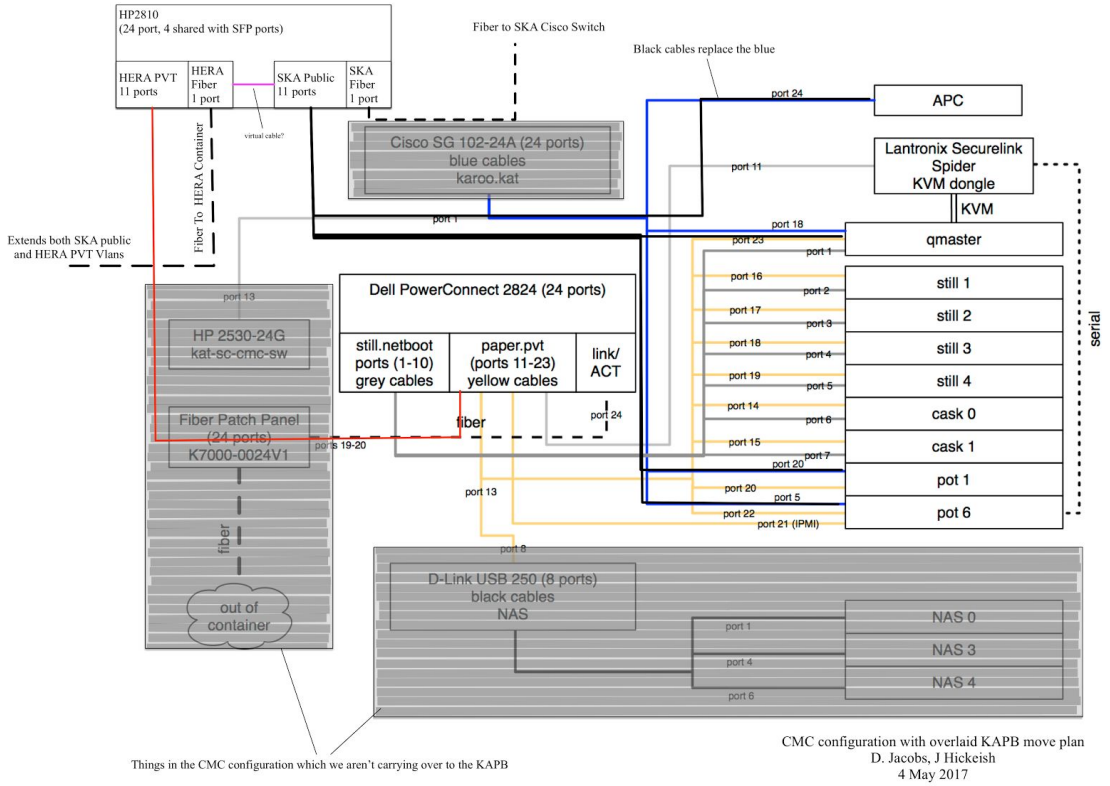


HERA Container Networking
Initial Drawing: B. Hazelton October 2016
Updated to show planned SKA interface: D. Jacobs May 2017 Note: Additions for May 2017 highlighted in red

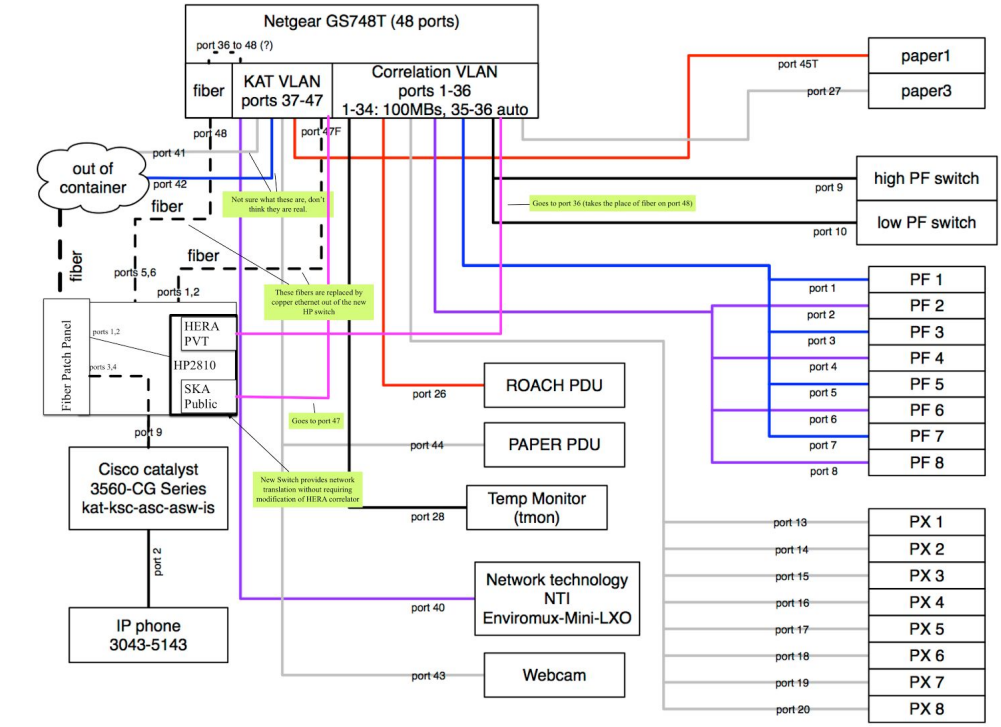


HERA RTP/Librarian Networking
Initial Drawing (in CMC): B. Hazelton October 2016
Updated to show planned SKA interface (and move to KAPB): D. Jacobs May 2017

Redlined original networking diagrams to show changes



CMC configuration with overlaid KAPB move plan
D. Jacobs, J Hickeish
4 May 2017



HERA container/correlator configuration with overlaid network move plan
D. Jacobs, J Hickeish
4 May 2017

Roadmap for HERA Network Configuration and Bandwidth

HERA Team

May 1, 2017

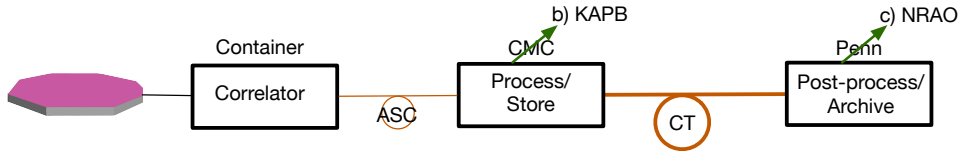
1 Introduction

HERA is an international experiment to detect and characterize the Epoch of Reionization (EOR). The telescope is located at the South African SKA site in the Karoo Astronomy Reserve. This note summarizes the overall network configuration and bandwidth for HERA as relevant to the interfaces with SKA-SA infrastructure. HERA construction and observing are proceeding in parallel.

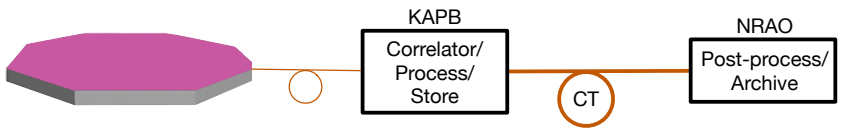
The HERA correlator is currently located next to the telescope array. In late 2017 or early 2018, the correlator will be moved to the KAPB. At that time the correlator will ingest raw voltage streams at terabit rates. These will be fed over a bundle of ∼96 fibers from the array to the KAPB. This future upgrade is outside the scope of this summary, other than to note the future bandwidth requirements to the US.

2 HERA network requirements prior to new correlator in 2018

Figure 1 shows the very high-level network configuration for 2017 and 2018. The correlator currently lives in the 'HERA container' in the middle of the HERA telescope array. The correlator data files are transferred for processing and storage in the CMC container using a dedicated fiber. The processed files are then transferred off-continent to the USA, currently to a cluster at the University of Pennsylvania in Philadelphia, PA. Over the next month or two, the CMC-based equipment will be moved to the KAPB and the USA node will be moved to the National Radio Astronomy Observatory (NRAO) site in Socorro, NM. The single pair of 1Gbe fiber between the HERA container and the processing cluster is adequate for operations through fall of 2017.



Phase 1: Current implementation; upgrade to 1b) move to KAPB and 1c) move to NRAO.



Phase 2: Full node/KAPB implementation.

Figure 1: High-level organization of network. See text for explanation.

Prior to the move of the equipment from the CMC to the KAPB, there is a desire to update the network configuration to accommodate the evolving needs of the KAT/SKA network and HERA. There is no strict

requirement that this must happen before the physical move, but doing so allows the network change to be deployed separately from the major hardware relocation. When the equipment is relocated to the KAPB additional storage will be added to the system and the switches will be upgraded.

Figure 2 shows a schematic of the proposed network arrangement. In order to ensure smooth HERA operations through mid-2018, we provide the following **requirements**:

1. There must be a login portal allowing secure access to HERA nodes from the Internet.
2. There must be 32 IP addresses reserved for HERA nodes on the SKA/KAT network so that they may be accessed directly from our login portal.
3. Some HERA nodes must be able to access to the SKA/KAT network for integration with CAM (as both data sinks and data sources).
4. There must be a dedicated fiber pair between the HERA container and the HERA processing cluster (currently located in the CMC; soon KAPB).
5. The network must support a data rate to the US of 200 Mbps by mid-2017, growing to 400 Mbps by mid-2018.

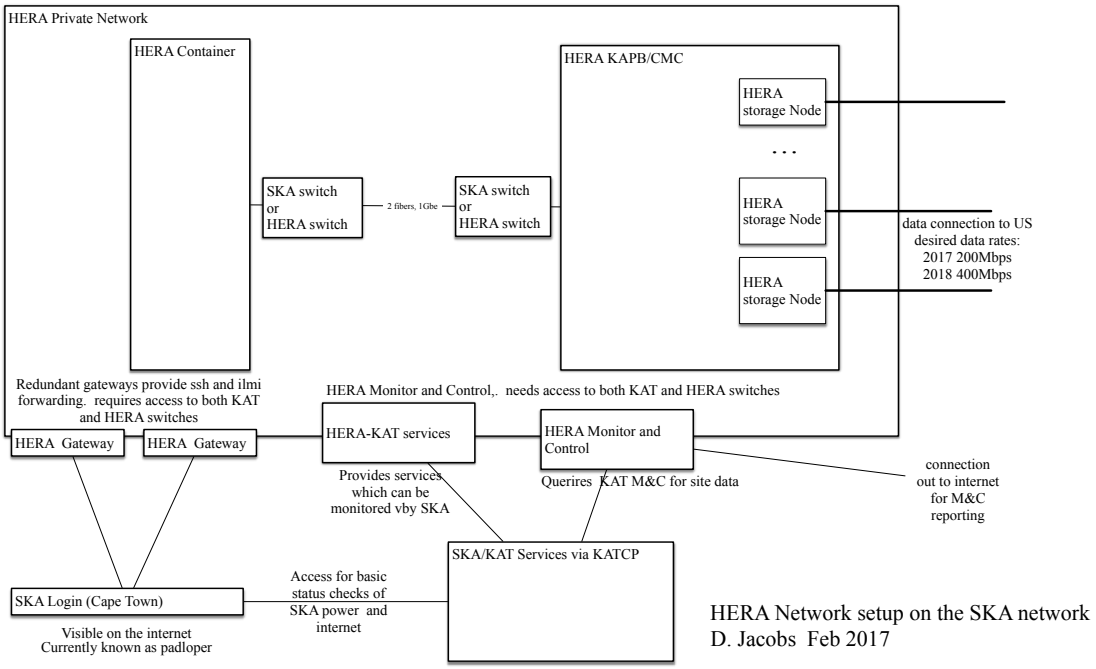


Figure 2: Schematic of notional network arrangement.

3 Data Rate

HERA currently records at a data rate of 200 Mbps. In 2018 this will increase to 1.5 Gbps. Currently the average bandwidth to the US is 60 Mbps, so we can only transfer a small fraction of our data. This limits analysis capabilities and means that mission-critical data are not replicated off-site. Moving data at the currently-attainable rate of 60 Mbps causes conflict with SKA SA staff Internet use, so copies to the US are limited to night-time hours, further reducing the effective bandwidth available to the project.

The desired data rate to support current operations is 200 Mbps. The target for mid-2018 is 400 Mbps.

4 Action Plan

Options are limited by only having a single fiber into the KAPB available for HERA use and by a desire to have a single connection point for all HERA machines onto the SKA network. This will limit the possibility of loopbacks. On the HERA side there is a desire to let several machines be able to access the outside world and some smaller subset of machines to be visible on the SKA net. In the current setup, HERA is connected to the legacy KAT network, which is kept wholly separate from the SKA net, and because most machines are behind a NAT it is difficult to monitor traffic and bandwidth usage from individual machines. The KAT network is being phased out.

In the new setup HERA will be connected to a single level 3 port on the KAPB Cisco network and given a subnet of some range of IPs. Using virtual networks in the HERA switches a private network and a public network will be created at the ska connection point and then both networks will be extended out to the HERA container. See diagrams in Figure 3. The change will proceed in two steps. First we will make all the necessary networking changes, tapping into the KAPB via the available fiber and creating the networks in the CMC. Second we will move the HERA processing machines and the main connection point into the KAPB, replacing the CMC switch with a fiber pass through.

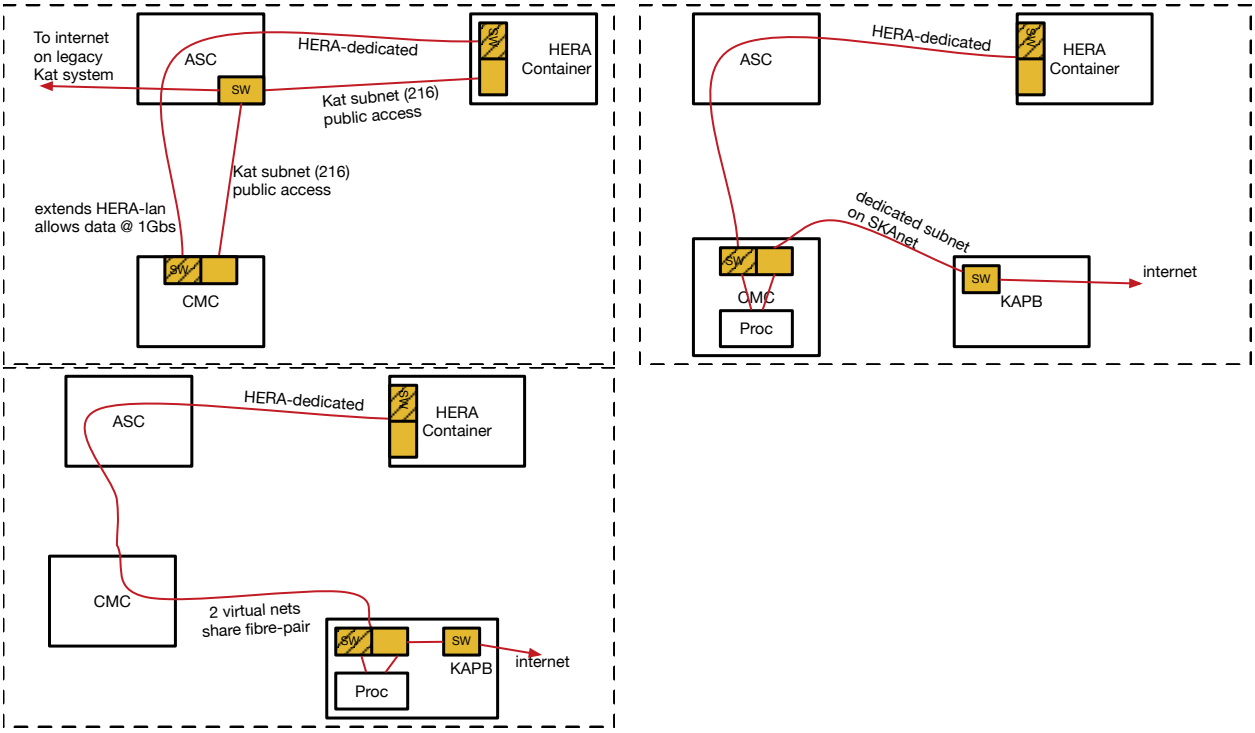


Figure 3: Notes from meeting Feb 13 2017. Clockwise from top left, the current networking setup with the HERA container and the HERA processing cluster in the CMC, the proposed intermediate setup moving HERA off the KAT network and onto the KAPB Cisco network and moving from two fibers (private and KAT) to a single fiber hosting two virtual nets, then bottom left moving the processing cluster into the KAPB, replacing the router in the CMC with a fiber pass-through.

SA Networking

External SSH access: `padloper.kat.ac.za` (IP: 196.24.41.252), **port 2222**.

Machines on the KAT network have IP addresses in the range 192.168.216.100 - 192.168.216.125.

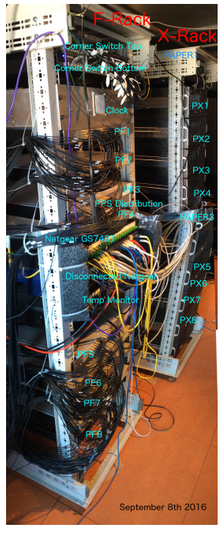
Host names ending in `karoo` are registered with the KAT/SKA network and are pingable from `padloper`, either as written or as (e.g.) `paper1.karoo.kat.ac.za`. Other hostnames are hardcoded in various `/etc/hosts` files and may not be defined consistently on all machines (yay).

Power Connections numbers run left-to-right for horizontally aligned power strips and top-to-bottom for vertically aligned power strips. Numbers include the strip's input. Machines with multiple power connections have them listed from top-to-bottom or left-to-right.

Name	KAT IP address	HERA IP address	Role	OS	Access	Network Connection	Power Connection
paper1.karoo	192.168.216.100	10.0.1.1	correlator control	Ubuntu 12.04		Netgear GS748T, 45T	PAPER PDU,3
paperqmaster.karoo	192.168.216.110	10.0.1.20		Ubuntu 12.04		DELL PowerConnect 2824,1(still),23(pvt)	CMC UPS, 2
Infrastructure							
paperwebcam.karoo	192.168.216.101		camera	N/A	admin:admin	Netgear GS748T, 43	
papersensor.karoo	192.168.216.103		networked temperature switch (Enviromux-mini-lxo?)	N/A	root:nti	Netgear GS748T,40	F-rack ML, 1
paper104.karoo	192.168.216.104		Labjack	N/A			
paperpdu.karoo	192.168.216.105		Networked power switch	N/A	apc:apc	Netgear GS748T, 44	
paper106.karoo	192.168.216.106		Netgear GS748T	N/A	P9ls4R*@		Frack, LL, 1
paper107.karoo	192.168.216.107		Liebert chiller HVAC first IP	N/A			
paper108.karoo	192.168.216.108		Liebert chiller HVAC second IP	N/A			
paperstillpdu.karoo	192.168.216.117		APC Power supply in CMC	N/A	apc:apc	Cisco SG 102-24, 24	
tmon.paper.pvt		10.0.1.39	HERA container temperature monitoring Raspberry Pi	?	?		
apc.paper.pvt, roachpdu		10.0.1.240	ROACH APC Power supply	N/A	apc:apc	Netgear GS748T, 26	
dellswitch.mgt.pvt		192.168.2.1	Switch for private network (Dell PowerConnect 2824)		telnet: admin, or HTTP: admin:(empty password)		CMC-rack, UL 6a
CISCO SG 102-24	?	?	Cisco CMC Switch				CMC UL 6b
Cisco catalyist 3560 CG Series	?	?	Cisco phone switch !Power over ethernet!				HERA UR 3
Data storage							
pot0, paper118.karoo	?	10.0.1.30	data storage pot (in US)	Centos 6.1			
pot1, paper109.karoo	192.168.216.118	10.0.1.31	data storage pot; specs: upenn062812_pot1.pdf	Centos 6.1		Dell Powerconnect 2824, 20	CMC APC, 3
pot2, paperpot2.karoo	192.168.216.119	10.0.1.32	data storage pot; specs: silicon_mechanics_pot2.pdf	Centos 6.1			
“pot3”: paperpot3.karoo	192.168.216.120	10.0.1.33	data storage pot specs: silicon_mechanics_quote_269207_pot3.pdf	Centos 6.1			
“pot6”: paper121.karoo	192.168.216.121	10.0.1.36	data storage pot specs: silicon_mechanics_quote_297449_pot6.pdf	CentOS 7.1.1503 (Core)		Dell Powerconnect 2824, 21 (IPMI) ; Dell Powerconnect 28248, 22 (PRIV) ; CISCO SG201-24, 5 (KAT)	CMC APC, 5
Processing							
still1.paper.pvt		10.0.{1,2}.21	processing host	Ubuntu 12.04		Dell Powerconnect 2824, 2(still),16(pvt)	CMC UL, 1
still2.paper.pvt		10.0.{1,2}.22	processing host	Ubuntu 12.04		Dell Powerconnect 2824, 3(still),17(pvt)	CMC UL, 2
still3.paper.pvt		10.0.{1,2}.23	processing host	Ubuntu 12.04		Dell Powerconnect 2824, 4(still),18(pvt)	CMC UL, 3
still4.paper.pvt		10.0.{1,2}.24	processing host	Ubuntu 12.04		Dell Powerconnect 2824, 5(still),19(pvt)	CMC UL, 4
cask0	?		?			Dell Powerconnect 2824, 6(still),14(pvt)	CMC APC, 6
cask1	?		?			Dell PowerConnect 2824, 7(still),15(pvt)	CMC APC, 7
NAS0	?		?			D-Link USB-250, 1	CMC NAS 4
NAS3	?		?			D-Link USB-250, 4	CMC NAS 5
NAS4	?		?			D-Link USB-250, 6	CMC NAS 1
Correlator							
PF1	?	?	F-Engine			Netgear GS748T, 1	X-rack UL, 2
PF2	?	?	F-Engine	?		Netgear GS748T, 2	X-rack UL, 3
PF3	?	?	F-Engine	?		Netgear GS748T, 3	Roach PDU, 5
PF4	?	?	F-Engine	?		Netgear GS748T, 4	Roach PDU, 4

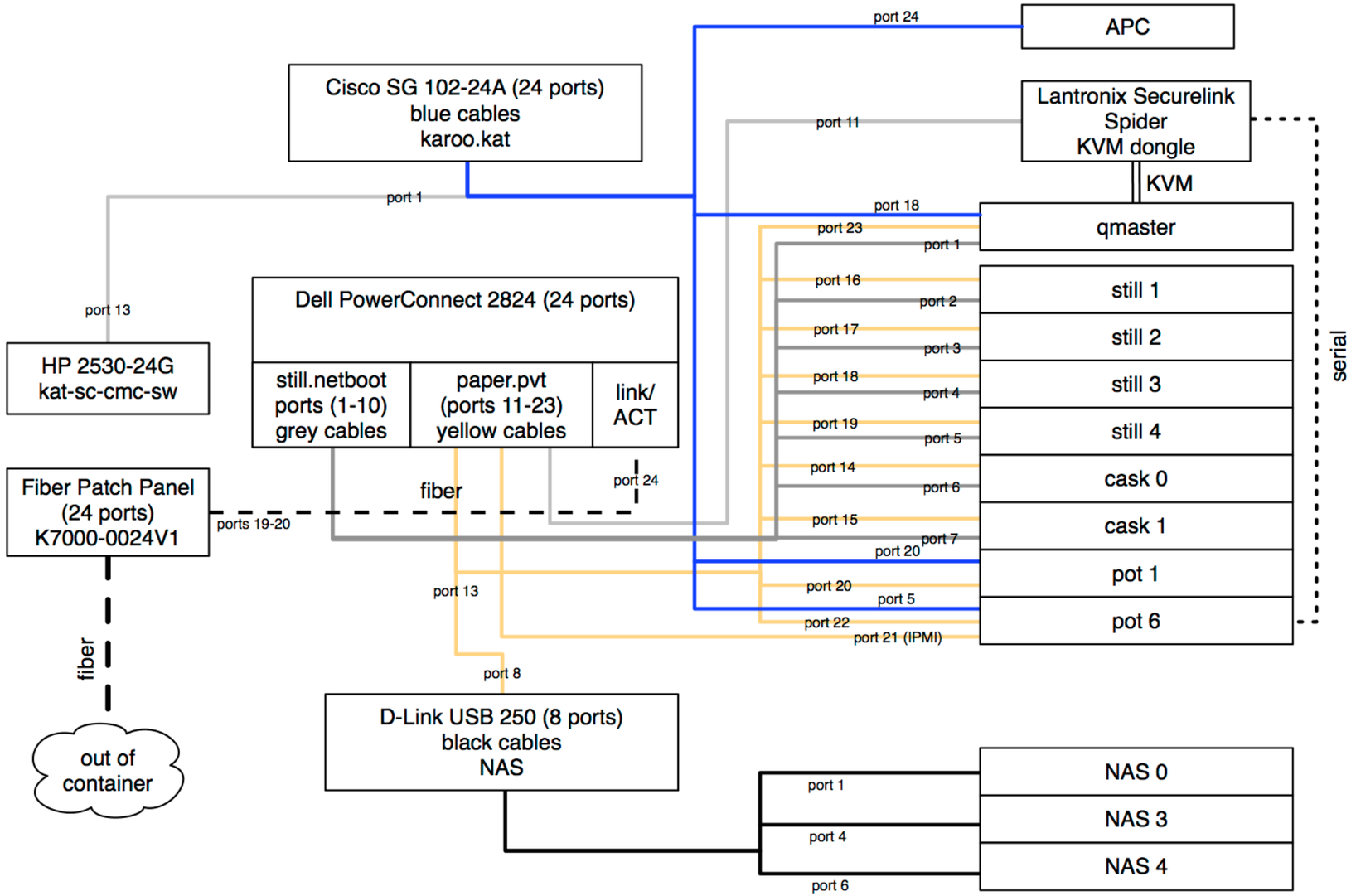
PF5	?	?	F-Engine	?		Netgear GS748T, 5	Roach PDU, 6
PF6	?	?	F-Engine	?		Netgear GS748T, 6	Roach PDU, 5
PF7	?	?	F-Engine	?		Netgear GS748T, 7	Roach PDU, 8
PF8	?	?	F-Engine	?		Netgear GS748T, 8	Roach PDU, 7
Corner Switch (top)	?	?	Corner turn?	?		Netgear GS748T, 9	F-rack ML, 2; F-rack UR, 1
Corner Switch (bottom)	?	?	Corner turn?	?		Netgear GS748T, 10	F-rack ML,3;F-rack UR, 2
Clock	?	?	?	?			PAPER PDU, 8
PPS Distribution	?	?	?	?			PAPER PDU, 4
Temp Monitor	?	?	?	?		Netgear GS748T, 28	PAPER PDU, 9
PX1	?	?	X-Engine	?		Netgear GS748T, 13	PAPER PDU 6,7
PX2	?	?	X-Engine	?		Netgear GS748T, 14	X-rack, UL 1,2
PX3	?	?	X-Engine	?		Netgear GS748T, 15	X-rack, UL 3,4
PX4	?	?	X-Engine	?		Netgear GS748T, 16	X-rack MR 1, X-rack LL 1
PX5	?	?	X-Engine	?		Netgear GS748T, 17	X-rack LL 2, X-rack MR 2
PX6	?	?	X-Engine	?		Netgear GS748T, 18	X-rack B 3, X-rack LR 5
PX7	?	?	X-Engine	?		Netgear GS748T, 19	X-rack LR 6, X-rack B 4
PX8	?	?	X-Engine	?		Netgear GS748T, 20	X-rack B 5, X-rack LR 7
Out of Band management (IPMI,KVM etc)							
pot0x.mgt.pvt		10.0.200.30	pot0 IPMI (in theory)				
pot1x.mgt.pvt		10.0.200.31	pot1 IPMI (in theory)				
pot2x.mgt.pvt		10.0.200.32	pot2 IPMI (in theory)				
pot3x.mgt.pvt		10.0.200.33	pot3 IPMI (in theory)				
pot6.ipmi		10.0.3.6	remote management host	supermicro IPMI and remote KVM	see below		
still1x.mgt.pvt		10.0.200.21	still1 IPMI				
still2x.mgt.pvt		10.0.200.22	still2 IPMI				
still3x.mgt.pvt		10.0.200.23	still3 IPMI				
still4x.mgt.pvt		10.0.200.24	still4 IPMI				
(none)		10.0.3.20	Lantronix Securelink Spider connected to `qmaster`		sysadmin : (standard P… password)		

Physical Computer Locations



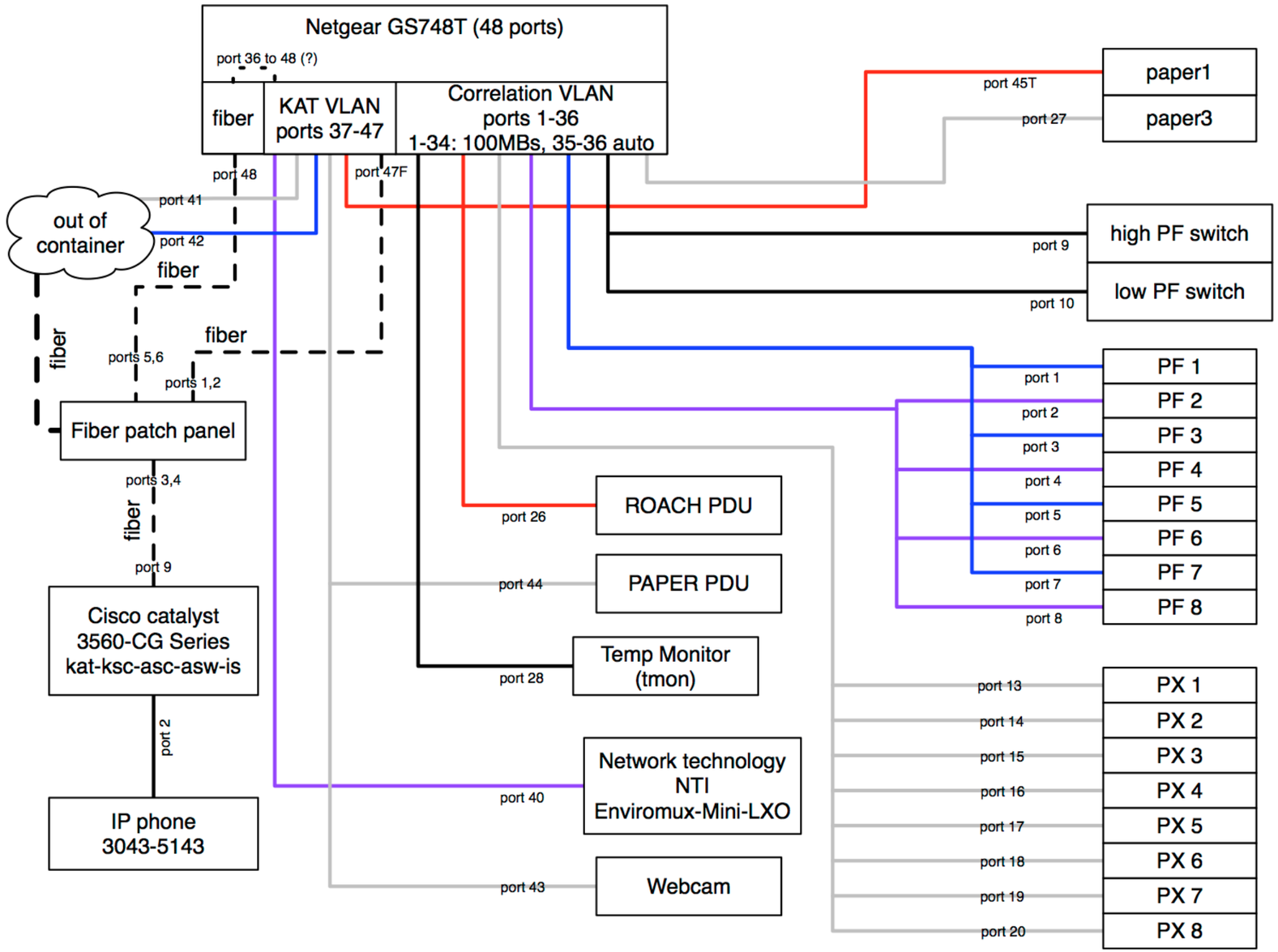
Power Supply Legend

Colored solid lines are ethernet cables, color gives cable color

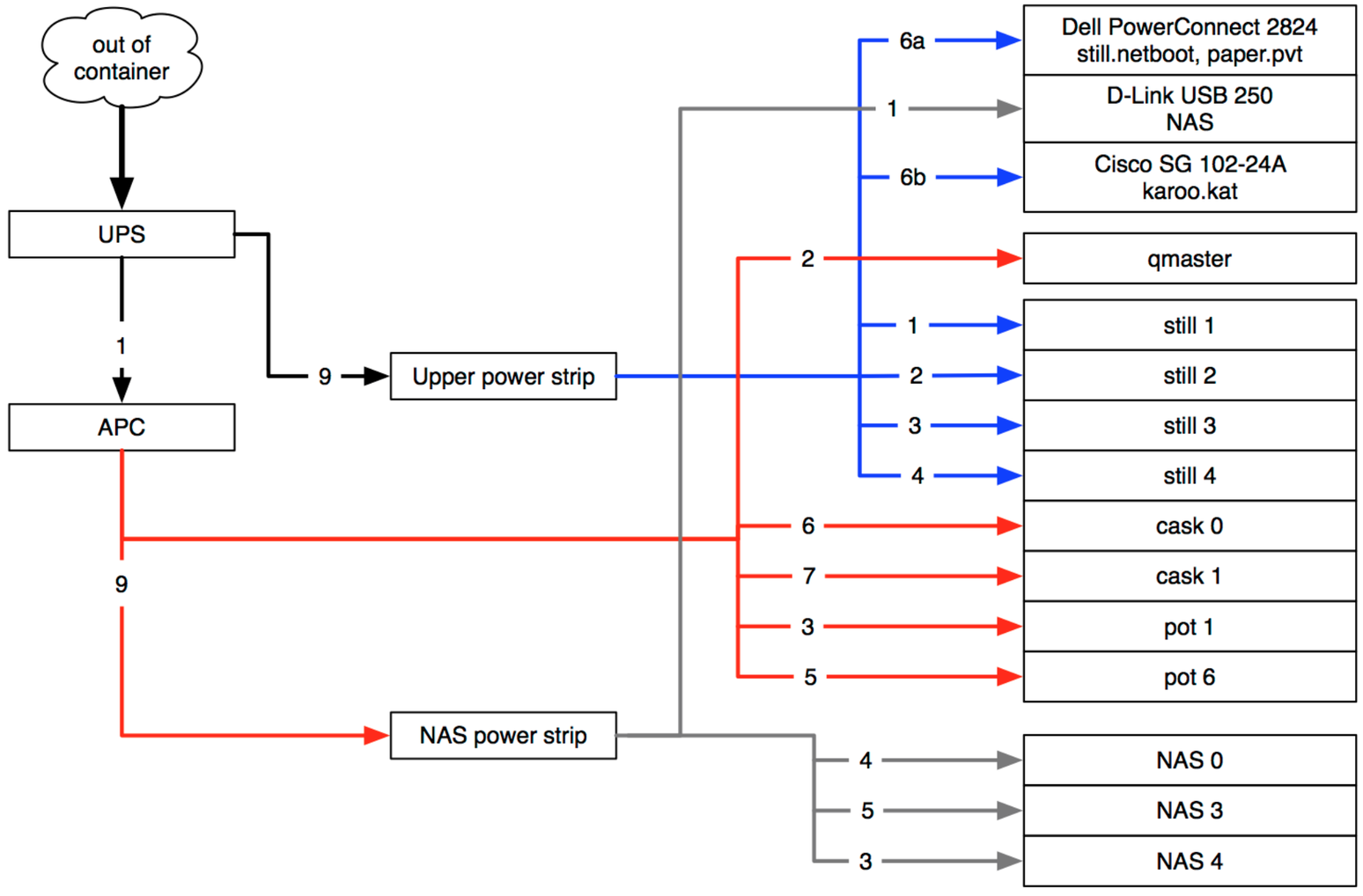


HERA Container network

Colored solid lines are ethernet cables, color gives cable color

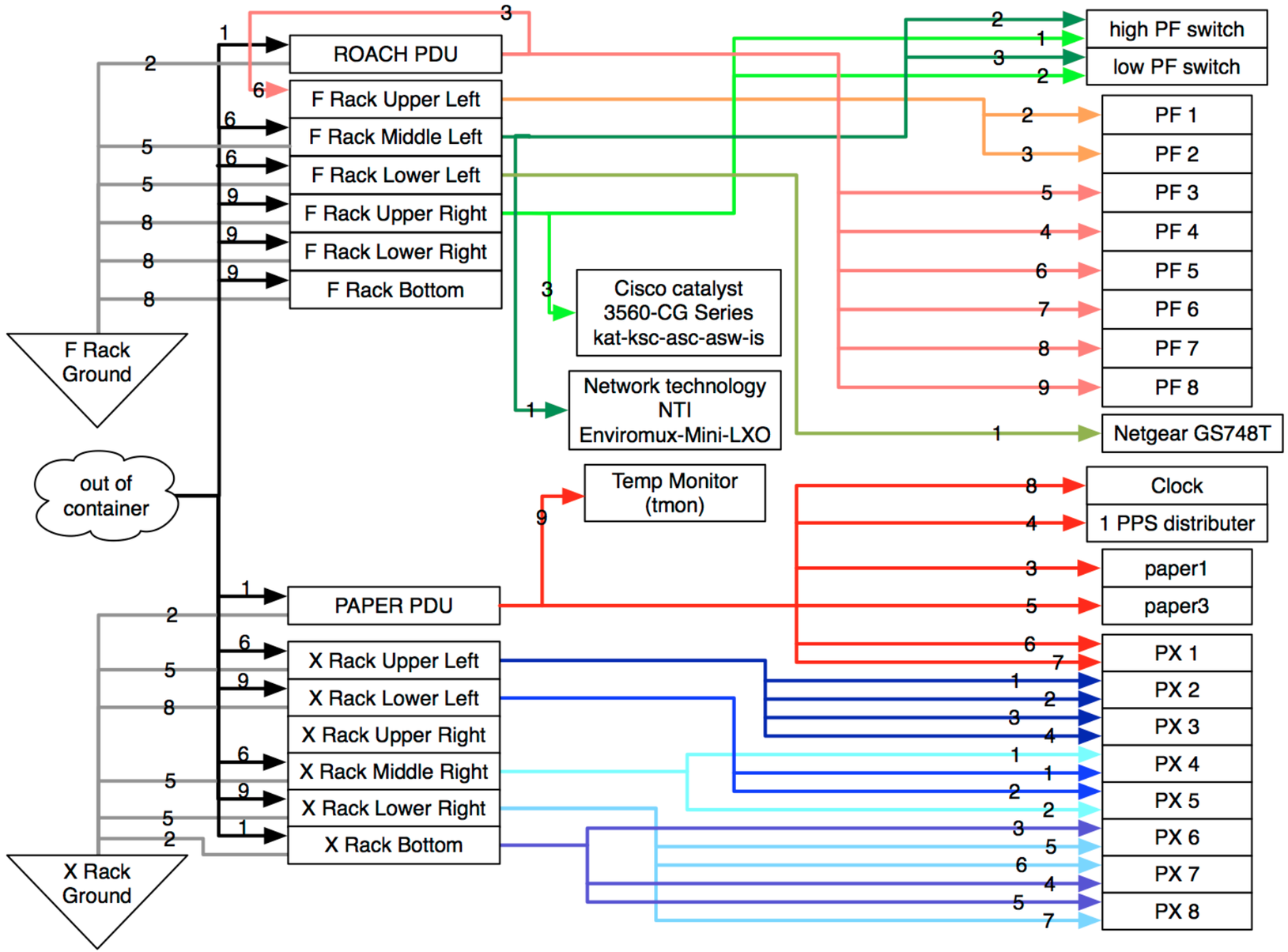


CMC HERA power



Colored solid lines are power cables, color used to distinguish sources.

HERA Container power



Colored solid lines are power cables, color used to distinguish sources.

pdf versions:

CMC HERA network diagram CMC HERA power diagram HERA Container network diagram HERA Container power diagram

Older, somewhat out-of-date diagrams