FMADIO20G Gen3

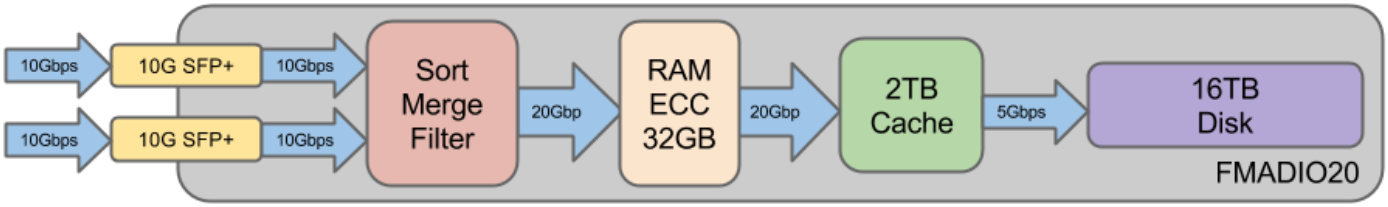
Introduction

Hardware

FMADIO 20G Gen3

The FMADIO 20G packet capture device is our entry level full sustained line rate 10Gbit capture to cache packet capture / packet sniffer devices. It is a compact 1U 650mm deep chassis featuring 3.2 nanosecond resolution hardware packet time stamps and sub 100ns world time accuracy via PTPv2 and PPS.

In addition there is 1-4TB of high bandwidth SSD flash storage which is written back into 16-64TB of raw magnetic disk drives. The system is unique by combining a hybrid SSD / HDD storage architecture to gain maximum cost savings with maximum disk storage and still be capable of sustained up to 4TB worth of line rate capture without any packet drops.



FMADIO20G Gen3 Block Architecture

System SKU

FMADIO20G Gen3 Capture SKU

Capture only SKUS have 32 CPUs and 32GB of DDR4 RAM

SKU	Description
FMAD-20Gv3-1U-1T-16T	FMAD20v3 1U 1TB Cache / 16TB Storage Packe Capture System
FMAD-20Gv3-1U-4T-40T	FMAD20v3 1U 1TB Cache / 40TB Storage Packe Capture System
FMAD-20Gv3-1U-4T-64T	FMAD20v3 1U 4TB Cache / 16TB Storage Packe Capture System

FMADIO20G Gen3 Analytics SKU

Analytics SKU have 96 CPUs and 386GB of DDR4 RAM, designed for on-system PCAP network processing

SKU	Description
FMAD-20Gv3-1U-4T-40T-ANALYTICS	FMAD20v3 1U 4TB Cache / 40TB Storage Packe Capture Analytics System 2.1Ghz 96CPU + 384G RAM

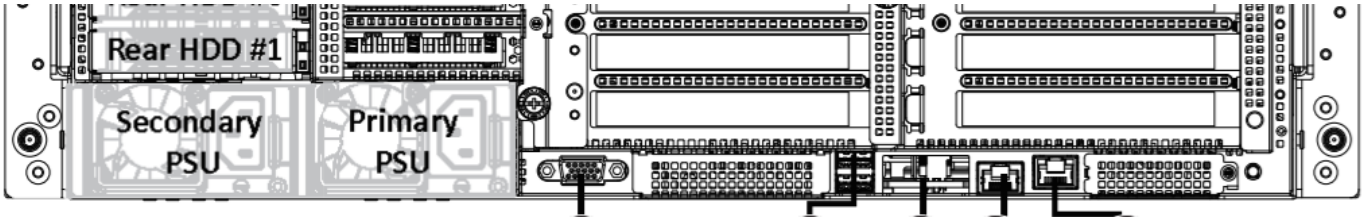
Hardware

FMADIO 20G Gen3 1U System

FMADIO 20G Gen3 1U Packet Capture System

FMADIO 20G Gen3 2U System

FMADIO 20G Gen3 2U Packet Capture System



FMADIO 20G 2U Rear view

Power Consumption

FMADIO20G Gen2 1U System

SKU	Description	Average	Max
Capture 1U	System Idle	150W	300W
Capture 1U	Full 20Gbps Sustained Capture	200W	300W
Capture 2U	System Idle		
Capture 2U	Full 20Gbps Sustained Capture		
Analytics 1U	System Idle		
Analytics 1U	Full 20GBps Sustained Capture		
Analytics 1U	96 CPUs maximum processing		

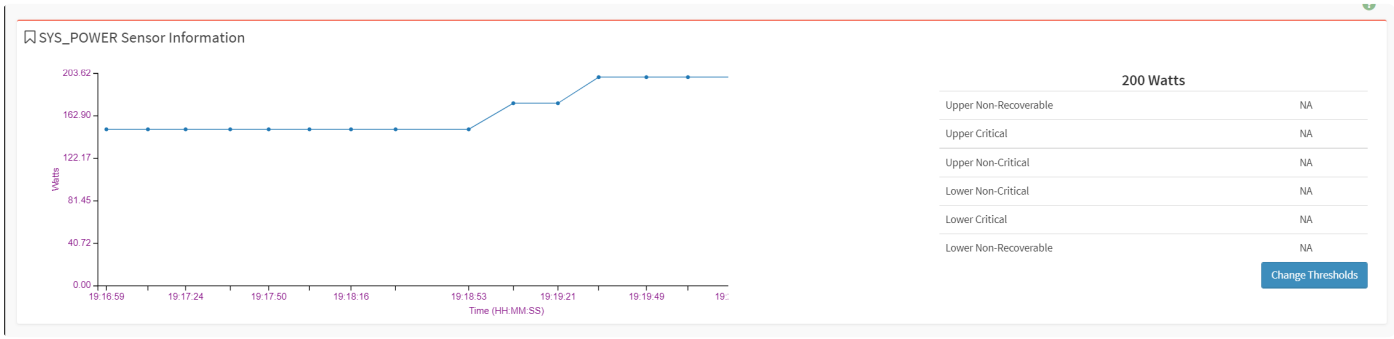
1U Capture Idle Power Consumption

FW: 7256 (2021/7/10)

FMADIO20G Gen3 1U Power Consumption Idle

1U Capture Full Capture Power consumption

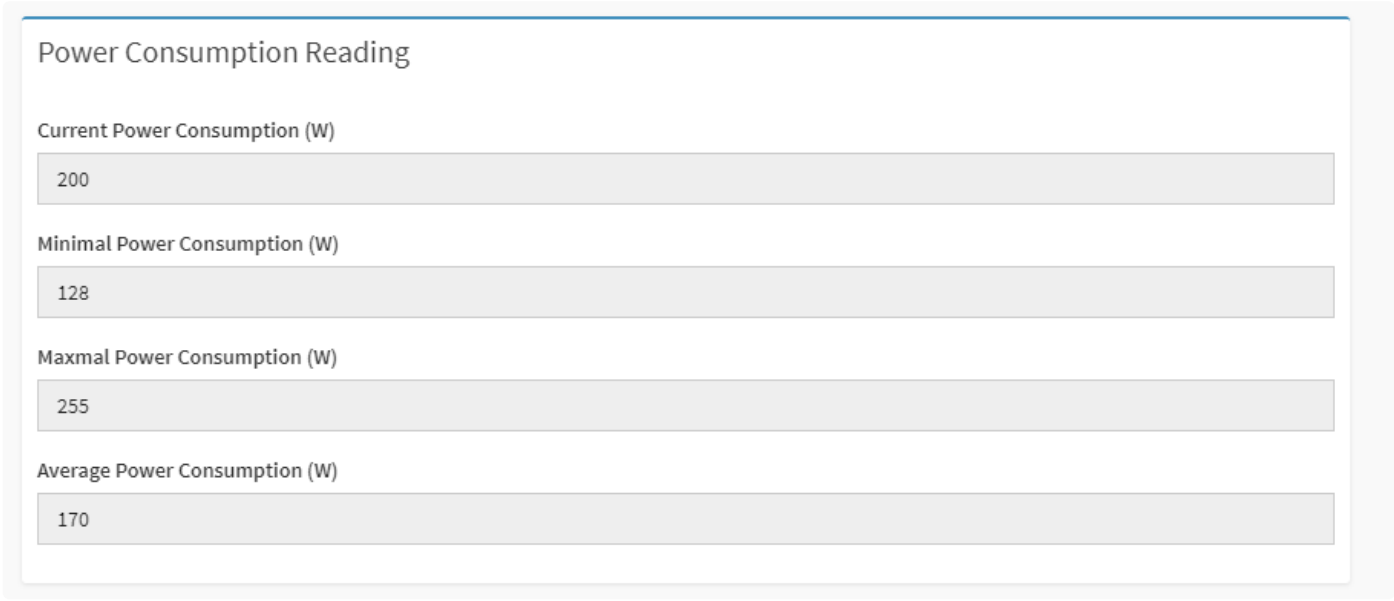
FW: 7256 (2021/7/10)



FMADIO20G Gen3 1U Power Consumption Full Capture

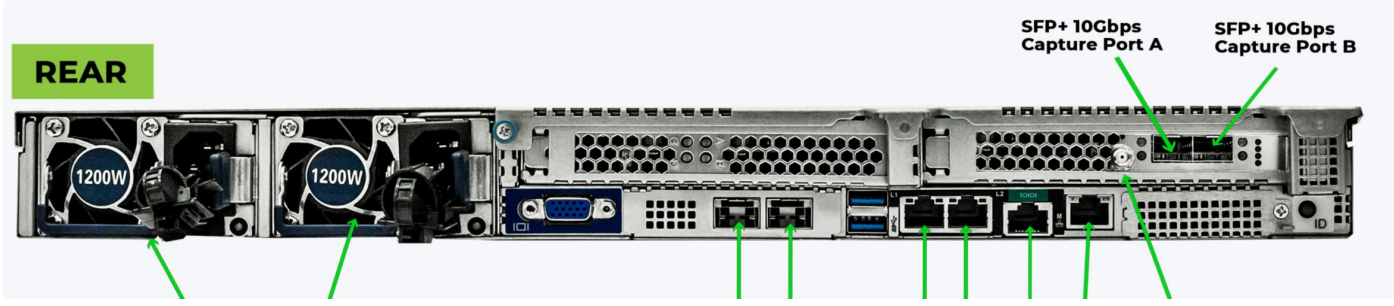
1U Maximum Power Consumption

FW: 7256 (2021/710)



Network Port Layout

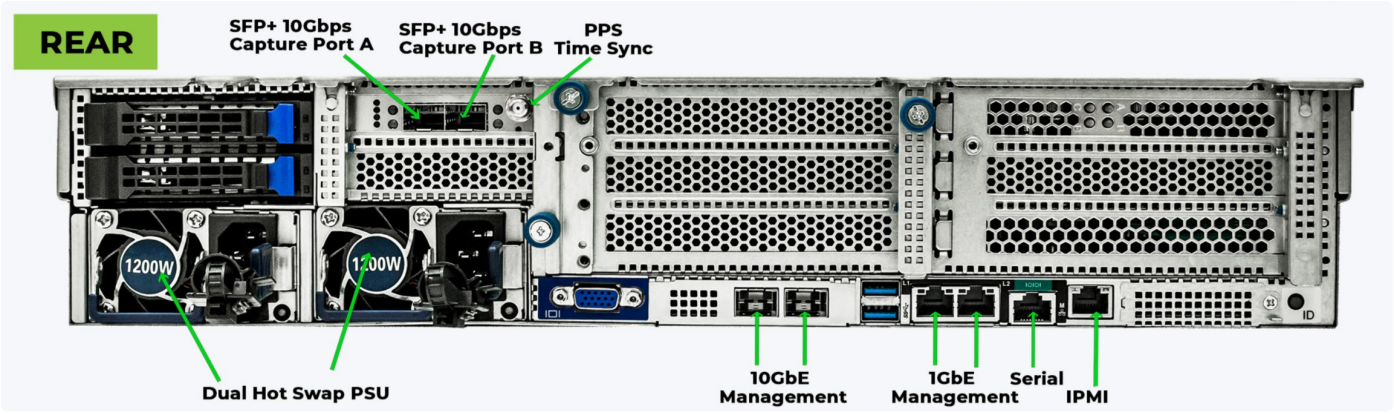
FMADIO 20G Gen3 1U Capture Port layout



FMADIO 20G Gen3 1U Analytics Port Layout

tbd*

FMADIO 20G Gen3 2U Capture Port layout



Performance

The performance of a capture system can be characterized in a number of different ways. We provide the following performance dimensions

---

Microburst Capture Speed (< 500msec)

This is the short time burst capture rate of the system. For the 100G Gen2 system this is burst capture rate that fills up the DDR buffer on the FMAD FPGA Capture card.

Burst Capture Speed (< 10min)

For the 100G Gen2 FMADIO Packet Capture system, all storage is on high speed NVMe SSDs, so the Burst Capture Speed is the same as the Sustained Capture Speed.

FMADIO 20G and 40G systems use a mixture of SSD and magnetic disk storage, so for those systems the Burst Capture Speed is higher than the Sustained Capture Speed.

Sustained Capture Speed (24/7)

We indicate this as the sustained capture rate, i.e. the capture rate that a system can sustain 24/7 without any packet loss. As mixing capture with downloads effects the capture speed, this performance metric is *Capture Only* with no simultaneous/concurrent downloads.

Capture and Download Speed

Performance metric is assuming no bottlenecks on the egress (download client) what is the capture performance while simultaneously downloading.

Download Only speed

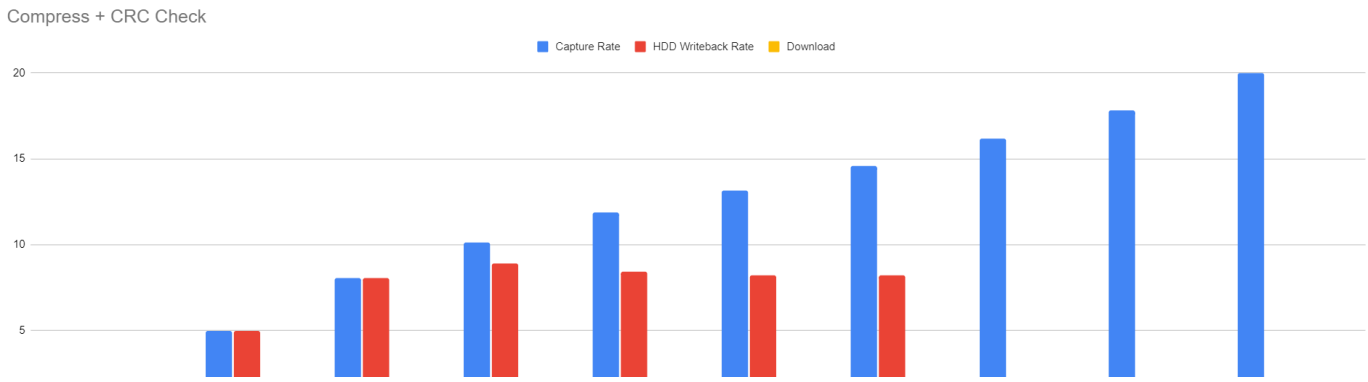
The other metric is Download only speed. This metric is used to calculate the maximum rate data can be moved off the device over 10G or 40G ethernet.

FMADIO 20G 2U System

FMADIO 20G 2U Packet Capture system has 4TB of SSD Cache and 48TB-216TB worth of HDD Magnetic storage.

Compression Enabled, CRC Check, No Download

The default setting has Compression and CRC checks enabled. Its designed to get maximum total storage capacity via the use of compression and CRC Checks for data integrity. This specific dataset is incompressible, thus the writeback performance is the raw hardware performance.

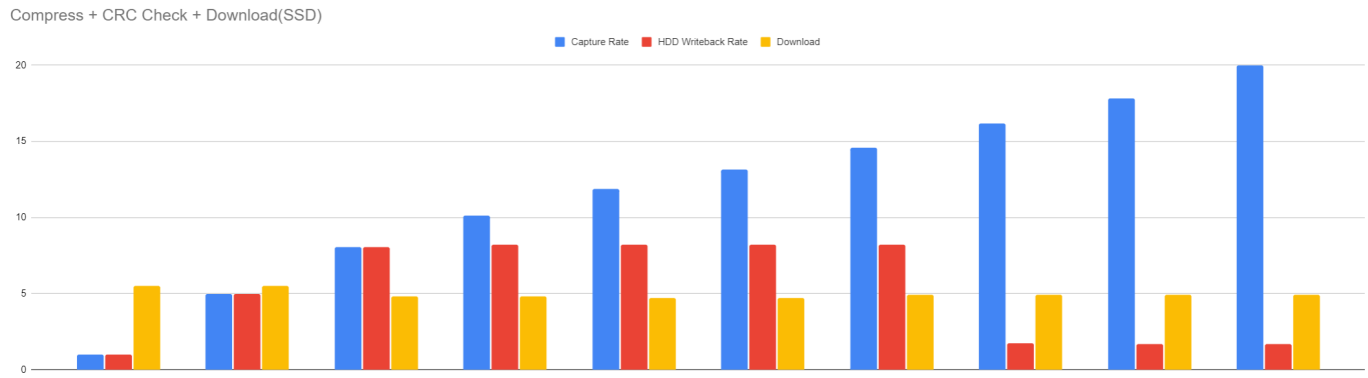


FMADIO20G 2U Default Performance

Testing: FMADIO20Gv3-2U-48TB System

Compression + CRC Check + Download (SSD)

Compression and CRC checks enabled and downloads that hit the SSD cache. e.g Download data is on SSDs and does not require access to the HDD. Download is using localhost to remove network performance from the test.



FMADIO20G 2U Default Performance Capture + Download(SSD)

Testing: FMADIO20Gv3-2U-48TB System

Compression + CRC Check + Download (HDD)

There is a difference when downloading from SSD cache vs HDD storage, as seen below. When a download has to fetch data from HDD magnetic storage, it dramatically effects the throughput of the HDD writeback. This is physical limitation of magnetic storage, as its a physical spinning disk which has poor random IO access performance. This is clearly seen in the significantly lower writeback and download speed, as shown below.

FMADIO20G 2U Default Performance Capture + Download(HDD)

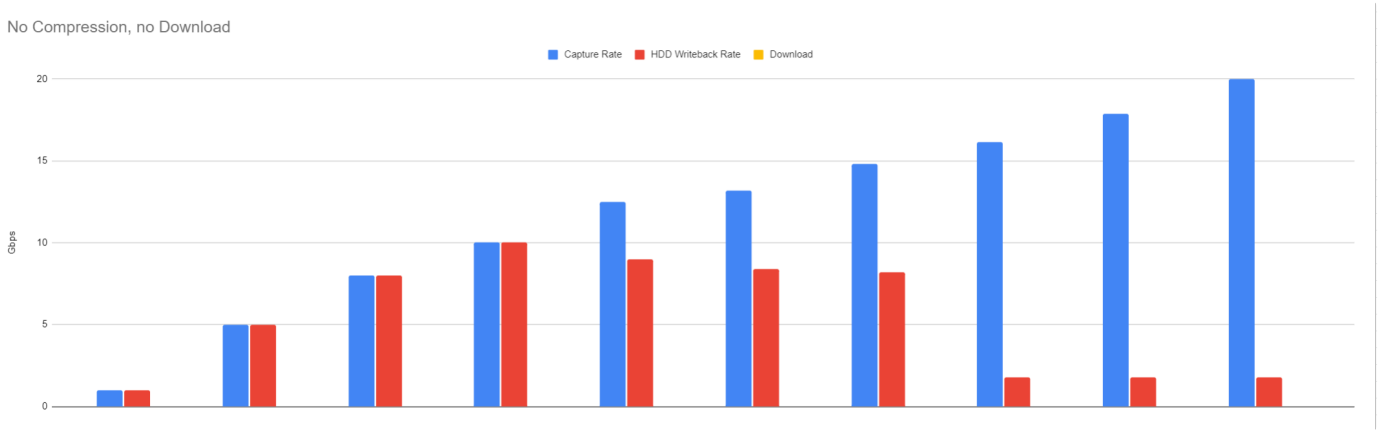
Testing: FMADIO20Gv3-2U-48TB System

No Compression, No CRC Check, No Download (Writeback Performance Mode)

This setting shows Writeback performance optimized mode per

https://docs.fmad.io/fmadio-documentation/v/fmadio20v3/settings/writeback#maximum-sustained-capture-rate).

This setup both Burst Capture and Sustained Capture rates @ 10Gbps are possible across the entire storage. However above 10Gbps Burst Capture is limited to SSD size (4TB) as beyond that the magnetic storage performance becomes a bottleneck.



FMADIO20G 2U Writeback Performance Mode

Testing: FMADIO20Gv3-2U-48TB System

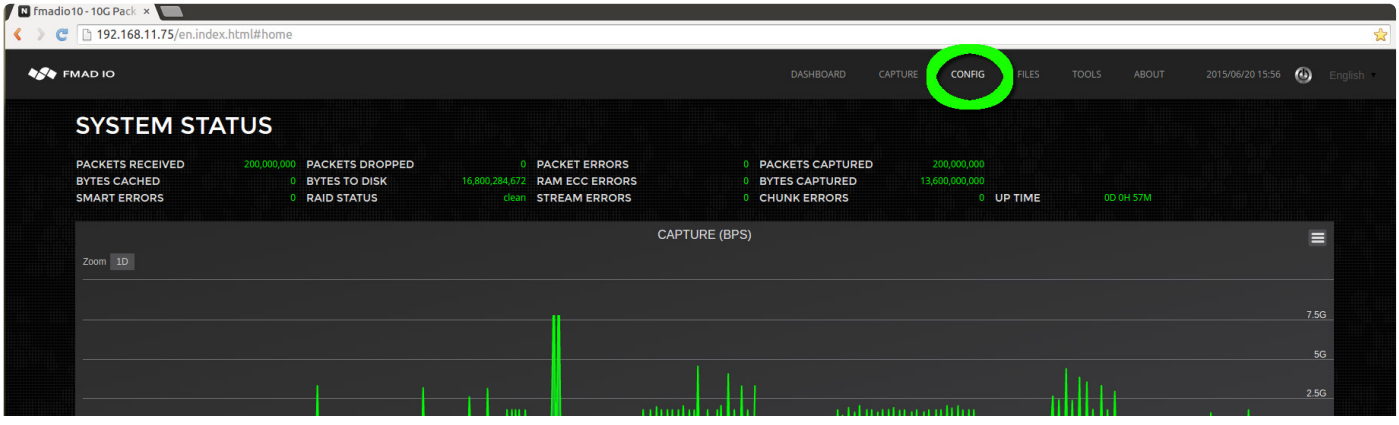
Configuration

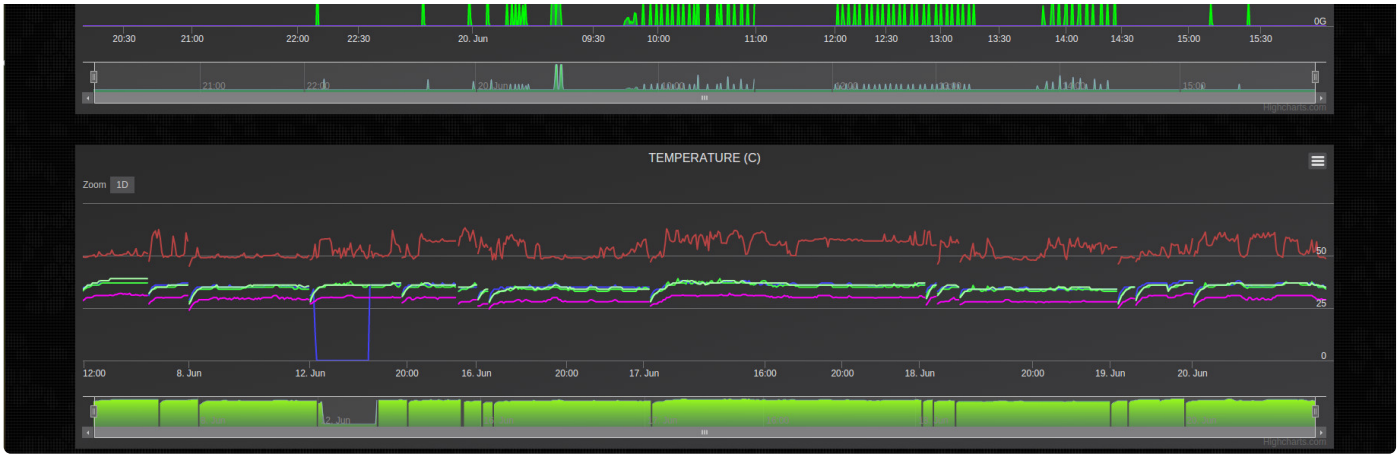
Network Configuration (GUI)

Network port configuration can be achieved using a) the web interface, b) SSH command line interface(CLI). Using the Web interface is the easiest route, however in highly constrained network environments a pure CLI based configuration can be easier

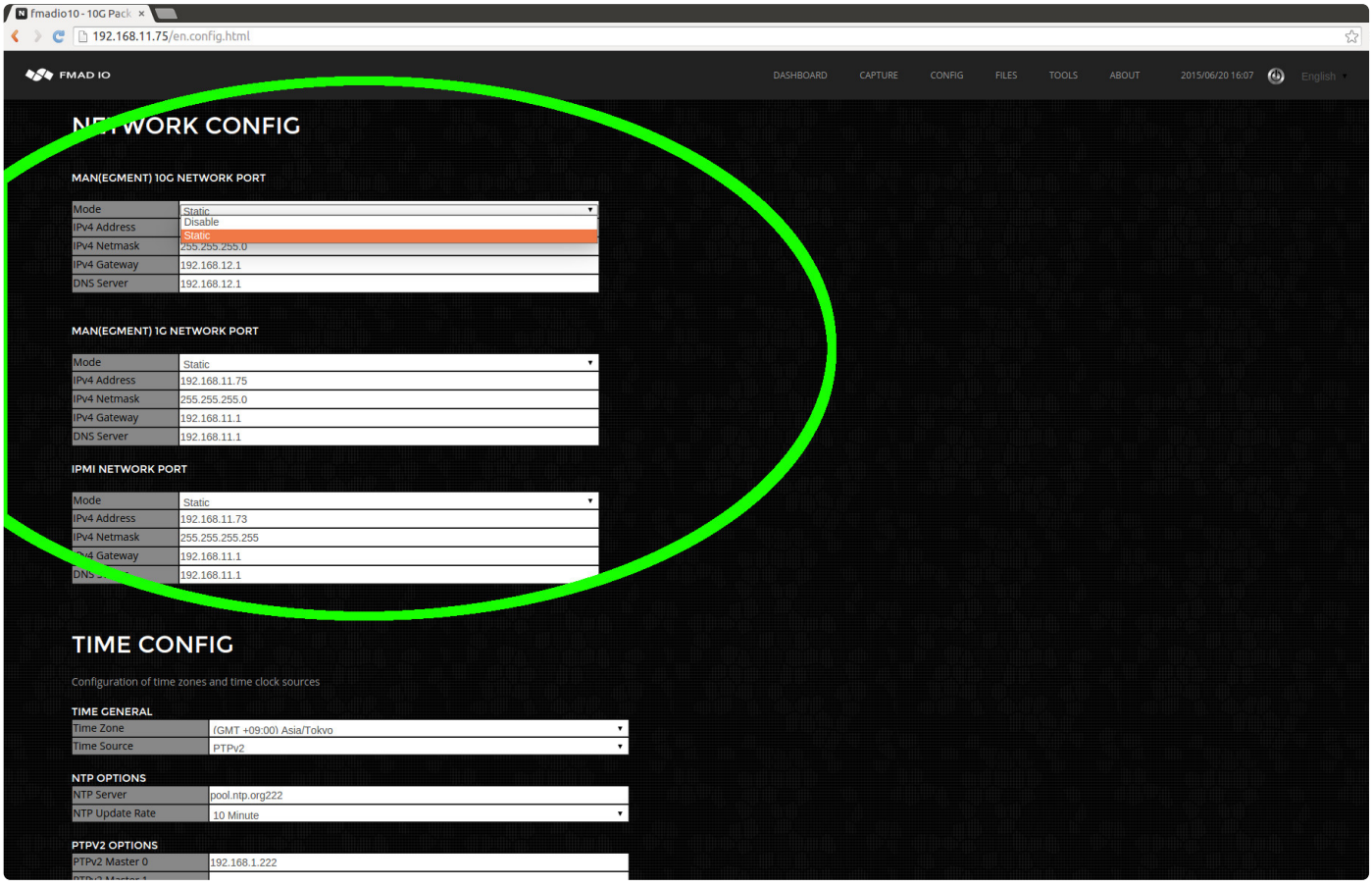
WEB INTERFACE: NETWORK CONFIG

From the dashboard page, Start by selecting the configuration menu option from as shown below (highlighted in green).

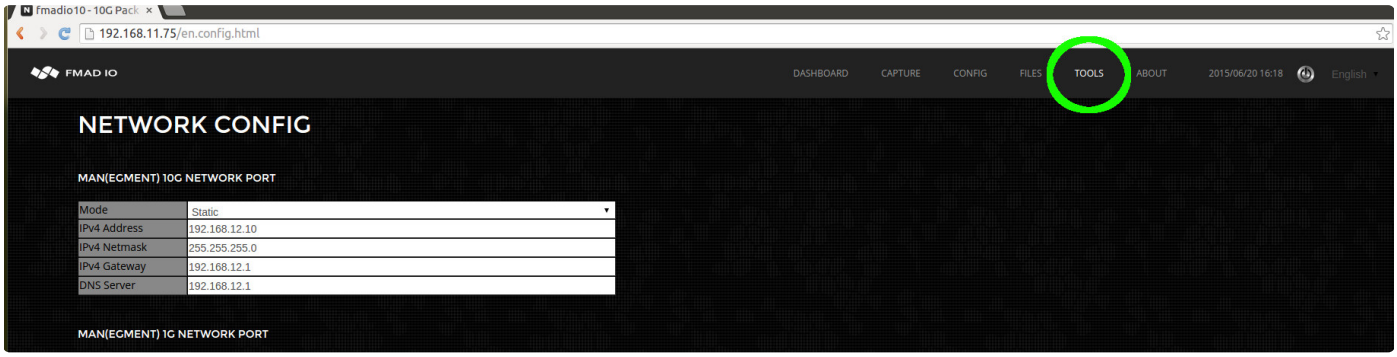




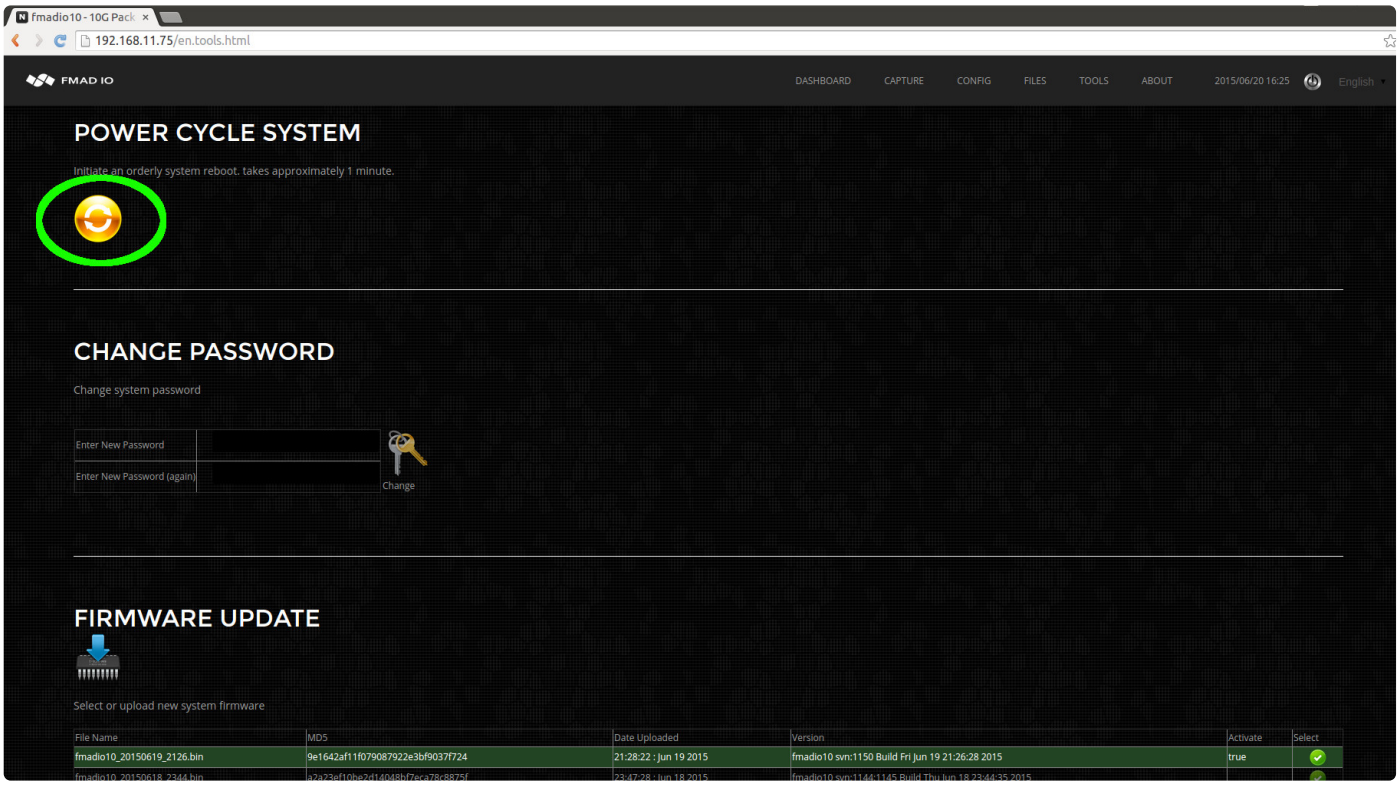
Then edit the network configuration`s IP/Netmask/Gateway/DNS setting as shown in the image below. After each field has been edited the system automatically saves and updates the system setting (save button is not required). After completing the update, refresh the web page to confirm the new settings.



Select the tools menu from the top toolbar, as shown in the image below.



And finally select the Power Cycle / Reboot button to restart the system



added some random comments

Network Configuration (CLI)

Modifying the network configuration setting in a restricted Colocation environment can be far easier to achieve via the command line. The first step is SSH into the system, change to the specified directory and view the current network settings, as shown below

```
aaron@display0:/tmp$ ssh fmadio@192.168.11.75
fmadio@192.168.11.75's password:
  _____                    .___.__       10G
_/ ____\_____  _____     __| _/|__|  ____
\   __\/     \ \__  \   / __ | |  | /  _ \
 |  | |  Y Y  \ / __ \_/ /_/ | |  |(  <_> )
 |__| |__|_|  /(____  /\____ | |__| \____/
            \/      \/      \/
=============================================
  -+ no user serviceable parts inside +-
fmadio@fmadio10-049:~$ cd /mnt/store0/etc
fmadio@fmadio10-049:/mnt/store0/etc$ cat network.lua
-- auto generated on Tue Apr 14 10:38:13 2015
local Config =
{
["sf0"] =

{
```

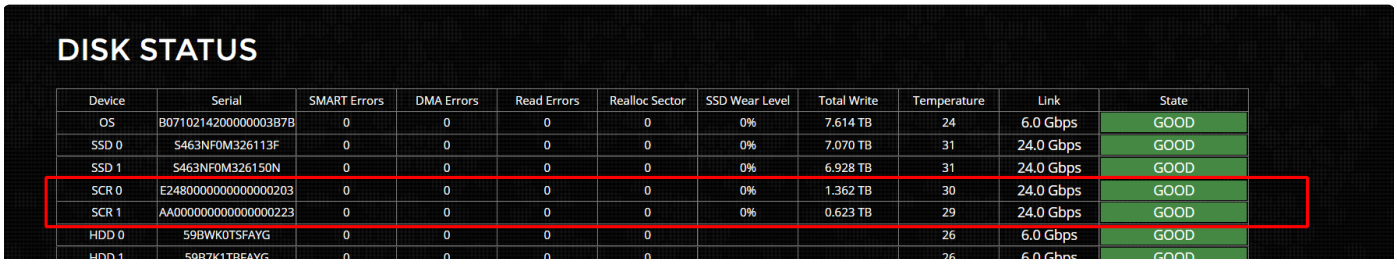
```
    ["Mode"]    = "disabled",
    ["Address"] = "192.168.1.2",
    ["Netmask"] = "255.255.255.0",
    ["Gateway"] = "192.168.1.1",
    ["DNS"]     = "192.168.1.1",
},
["sf1"] =
{
    ["Mode"]    = "static",
    ["Address"] = "192.168.12.10",
    ["Netmask"] = "255.255.255.0",
    ["Gateway"] = "192.168.12.1",
    ["DNS"]     = "192.168.12.1",
},
["eth0"] =
{
    ["Mode"]    = "static",
    ["Address"] = "192.168.11.75",
    ["Netmask"] = "255.255.255.0",
    ["Gateway"] = "192.168.11.1",
    ["DNS"]     = "192.168.11.1",
},
["bmc"] =
{
    ["Mode"]    = "static",
    ["Address"] = "192.168.11.73",
    ["Netmask"] = "255.255.255.255",
    ["Gateway"] = "192.168.11.1",
    ["DNS"]     = "192.168.11.1",
},
}
return Config
```

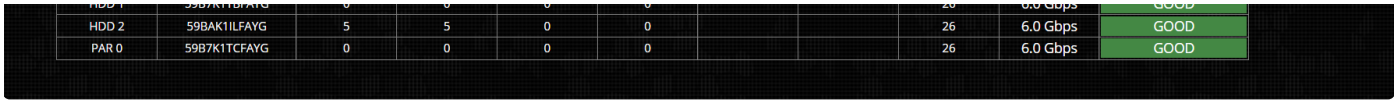
In the example configuration file above, the network ports are mapped as follows

Scratch Disk (EXT4)

When using FMADIO Packet capture system for analytics processing SSD resources can be split into Capture devices and Scratch disk space. In scratch disk space 1-16TB of SSD can be mounted as a general purpose file system used to store temporarily/intermediate network packet processing results.

The system should have scratch disks setup and visible on the GUI as follows, if this has not been configured contact support@fmad.io on how to configure





FMADIO Scratch Disk Network Analytics processing space

In the above example there are 2 disks SCR0 and SCR1 enabled for scratch disk these are seen on the file system as

```
fmadio@fmadio20v3-287:~$ ls -al /opt/fmadio/disk/scr*
lrwxrwxrwx    1 root     root            12 Jun  6 17:02 /opt/fmadio/disk/scr0 -> /dev/nvme
lrwxrwxrwx    1 root     root            12 Jun  6 17:02 /opt/fmadio/disk/scr1 -> /dev/nvme
fmadio@fmadio20v3-287:~$
```

NOTE: the /dev/* mount point may change from time to time, please use the /opt/fmadio/disk/scr* path name for all operations.

NOTE: Any change to the scratch disk configuration it is recommend to run a Quick Format to ensure SSD and Capture disks configuration are consistent

---

Specifcying Sratch Disks

By default all SSD are specified are dedicated to capture. This is specified in the configuration file

```
/opt/fmadio/etc/disk.lua
```

Capture disks are specified here

```
CacheDisk =
{
    ["S462NF0MA04379A"] = "ssd0",
    ["S462NF0MA05134B"] = "ssd1",
    ["S5JXNG0N108745Y"] = "ssd2",
    ["S5JXNG0N108746D"] = "ssd3",
}
```

In the above example we have 4 x SSD for capture. To convert half to capture and half to scratch disk modify as follows

```
CacheDisk =
{
    ["S462NF0MA04379A"] = "ssd0",
    ["S462NF0MA05134B"] = "ssd1",
}
,
ScratchDisk =
{
```

```
10      ["S5JXNG0N108745D"] = "scr0",
11  }
```

This is assigning the SSD Serial numbers to mount point /opt/fmadio/disk/scr0 and /opt/fmadio/disk/scr1. The actual Serial numbers for each system will be different, the mount point (scr0/scr1) is the same.

After updating confirm there are no syntax errors in the config file by running fmadiolua /opt/fmadio/etc/disk.lua as follows

```
1  fmadio@fmadio20n40v3-363:~$ fmadiolua /opt/fmadio/etc/disk.lua
2  fmad fmadlua Jun 13 2021
3  calibrating...
4  0 : 2992970378           2.9930 cycles/nsec offset:7.030 Mhz
5  Cycles/Sec 2992970378.0000 Std:       0 cycle std(  0.00000000) Target:3.00 Ghz
6  argv fmadiolua
7  failed to open self? [fmadiolua]
8  loading filename [/opt/fmadio/etc/disk.lua]
9  done 0.000038Sec 0.000001Min
10 fmadio@fmadio20n40v3-363:~$
```

Output as above shows correctly formatted file. Output per below shows configuration file with a syntax error (line 30 has some incorrect formatting)

```
1  fmadio@fmadio20n40v3-363:~$ fmadiolua /opt/fmadio/etc/disk.lua
2  fmad fmadlua Jun 13 2021
3  calibrating...
4  0 : 2992970368           2.9930 cycles/nsec offset:7.030 Mhz
5  Cycles/Sec 2992970368.0000 Std:       0 cycle std(  0.00000000) Target:3.00 Ghz
6  argv fmadiolua
7  failed to open self? [fmadiolua]
8  loading filename [/opt/fmadio/etc/disk.lua]
9  load status: 3 0
10 /opt/fmadio/etc/disk.lua:30: '}' expected (to close '{' at line 2) near 'ScratchDisk'
11 fmadio@fmadio20n40v3-363:~$
```

After confirming the configuration file syntax is correct, reboot the system. The mount points scr0 and scr1 should be visible as shown below.

```
1  fmadio@fmadio20n40v3-363:~$ ls -al /opt/fmadio/disk/scr*
2  lrwxrwxrwx    1 root     root            12 Jun  6 17:02 /opt/fmadio/disk/scr0 -> /dev/nvme
3  lrwxrwxrwx    1 root     root            12 Jun  6 17:02 /opt/fmadio/disk/scr1 -> /dev/nvme
4  fmadio@fmadio20n40v3-363:~$
```

Creating RAID0 partition

After /opt/fmadio/disk/scr[0-1] have been created. Next is creating a /dev/md1 RAID0 partition as follows

```
fmadio@fmadio20v3-287:~$ sudo mdadm --create /dev/md1 --force --level=raid0 --raid-devices=
mdadm: Defaulting to version 1.2 metadata
mdadm: array /dev/md1 started.
fmadio@fmadio20v3-287:~$
```

This creates a /dev/md1 partition as shown with lsblk command. Can see the /dev/md1 device

```
fmadio@fmadio20v3-287:~$ lsblk
NAME    MAJ:MIN RM   SIZE RO TYPE  MOUNTPOINT
sdd       8:48   0   3.7T  0 disk
sdb       8:16   0   3.7T  0 disk
nvme2n1 259:0    0 447.1G  0 disk
`-md1     9:1    0   894G  0 raid0
nvme1n1 259:3    0   477G  0 disk
sde       8:64   0   3.7T  0 disk
sdc       8:32   0   3.7T  0 disk
sda       8:0    0 238.5G  0 disk
|-sda2    8:2    0 223.6G  0 part  /mnt/store0
`-sda1    8:1    0  14.9G  0 part  /mnt/sda1
nvme0n1 259:1    0 447.1G  0 disk
`-md1     9:1    0   894G  0 raid0
nvme3n1 259:2    0   477G  0 disk
fmadio@fmadio20v3-287:~$
```

More detail via the mdadm --detail command

```
fmadio@fmadio20v3-287:~$ sudo mdadm --detail /dev/md1
/dev/md1:
        Version : 1.2
  Creation Time : Sun Jun  6 17:43:12 2021
     Raid Level : raid0
     Array Size : 937440256 (894.01 GiB 959.94 GB)
   Raid Devices : 2
  Total Devices : 2
    Persistence : Superblock is persistent

    Update Time : Sun Jun  6 17:43:12 2021
          State : clean
 Active Devices : 2
Working Devices : 2
 Failed Devices : 0
  Spare Devices : 0

     Chunk Size : 512K

           Name : fmadio20v3-287:1  (local to host fmadio20v3-287)
           UUID : 06234ac8:694ae295:3659e4fc:b59a7d55
         Events : 0

```

```
25     Number   Major   Minor   RaidDevice State
26        0     259        0        0      active sync   /dev/nvme2n1
27        1     259        1        1      active sync   /dev/nvme0n1
28 fmadio@fmadio20v3-287:~$
```

Create EXT4 Filesystem

The block device /dev/md1 is block level only, it contains no mountable file system. Next create btrfs filesystem on the device as follows

```
1 fmadio@fmadio20v3-287:~$ sudo mkfs.ext4 /dev/md1
2
3 < update me >
4
5 fmadio@fmadio20v3-287:~$
```

Mount the Scratch Filesystem

By default FMADIO Packet Capture systems at boot time will start the RAID0 partition and mount /dev/md1 Scratch disk to /mnt/store1

If it fails to mount, please issue the following command

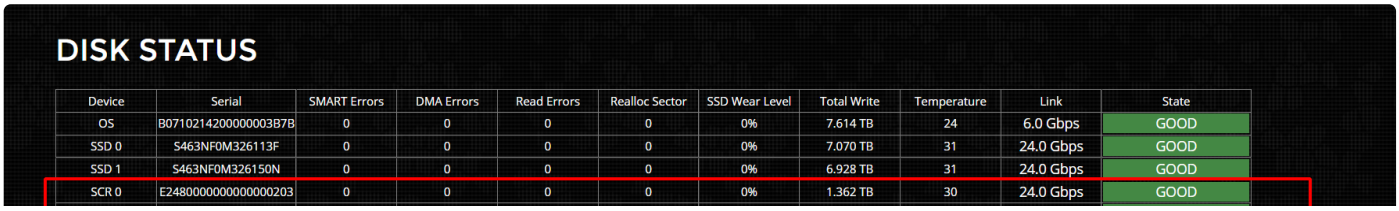
```
1 fmadio@fmadio20v3-287:~$ sudo mount -t ext4 /dev/md1 /mnt/store1
2 fmadio@fmadio20v3-287:~$
```

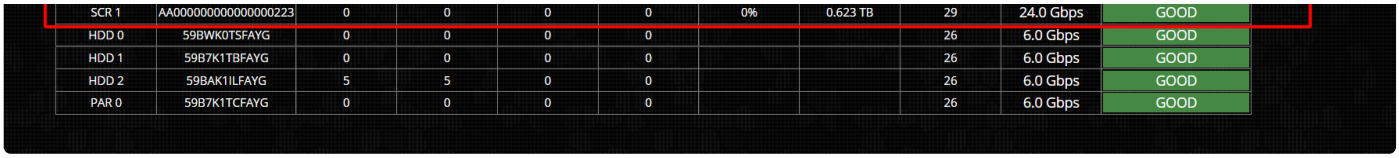
Scratch Disk (BTRFS)

FW: 7167+

When using FMADIO Packet capture system for analytics processing SSD resources can be split into Capture devices and Scratch disk space. In scratch disk space 1-16TB of SSD can be mounted as a general purpose file system used to store temporarily/intermediate network packet processing results.

The system should have scratch disks setup and visible on the GUI as follows, if this has not been configured contact support@fmad.io on how to configure





SCR 1	AA000000000000000223	0	0	0	0	0%	0.623 TB	29	24.0 Gbps	GOOD
HDD 0	59BWK0TSFAYG	0	0	0	0			26	6.0 Gbps	GOOD
HDD 1	59B7K1TBFAYG	0	0	0	0			26	6.0 Gbps	GOOD
HDD 2	59BAK1ILFAYG	5	5	0	0			26	6.0 Gbps	GOOD
PAR 0	59B7K1TCFAYG	0	0	0	0			26	6.0 Gbps	GOOD

FMADIO Scratch Disk Network Analytics processing space

In the above example there are 2 disks SCR0 and SCR1 enabled for scratch disk these are seen on the file system as

```
fmadio@fmadio20v3-287:~$ ls -al /opt/fmadio/disk/scr*
lrwxrwxrwx    1 root     root            12 Jun  6 17:02 /opt/fmadio/disk/scr0 -> /dev/nvme
lrwxrwxrwx    1 root     root            12 Jun  6 17:02 /opt/fmadio/disk/scr1 -> /dev/nvme
fmadio@fmadio20v3-287:~$
```

NOTE: the /dev/* mount point may change from time to time, please use the /opt/fmadio/disk/scr* path name for all operations.

Creating RAID0 partition

Start by creating a /dev/md1 RAID0 partition as follows

```
fmadio@fmadio20v3-287:~$ sudo mdadm --create /dev/md1 --force --level=raid0 --raid-devices=
mdadm: Defaulting to version 1.2 metadata
mdadm: array /dev/md1 started.
fmadio@fmadio20v3-287:~$
```

This creates a /dev/md1 partition as shown with lsblk command. Can see the /dev/md1 device

```
fmadio@fmadio20v3-287:~$ lsblk
NAME    MAJ:MIN RM   SIZE RO TYPE  MOUNTPOINT
sdd       8:48   0   3.7T  0 disk
sdb       8:16   0   3.7T  0 disk
nvme2n1 259:0    0 447.1G  0 disk
`-md1     9:1    0   894G  0 raid0
nvme1n1 259:3    0   477G  0 disk
sde       8:64   0   3.7T  0 disk

sdc       8:32   0   3.7T  0 disk
sda       8:0    0 238.5G  0 disk
|-sda2    8:2    0 223.6G  0 part  /mnt/store0
`-sda1    8:1    0  14.9G  0 part  /mnt/sda1
nvme0n1 259:1    0 447.1G  0 disk
`-md1     9:1    0   894G  0 raid0
nvme3n1 259:2    0   477G  0 disk
fmadio@fmadio20v3-287:~$
```

More detail via the mdadm --detail command

```
fmadio@fmadio20v3-287:~$ sudo mdadm --detail /dev/md1
```

```
/dev/md1:
        Version : 1.2
  Creation Time : Sun Jun  6 17:43:12 2021
     Raid Level : raid0
     Array Size : 937440256 (894.01 GiB 959.94 GB)
   Raid Devices : 2
  Total Devices : 2
    Persistence : Superblock is persistent

    Update Time : Sun Jun  6 17:43:12 2021
          State : clean
 Active Devices : 2
Working Devices : 2
 Failed Devices : 0
  Spare Devices : 0

     Chunk Size : 512K

           Name : fmadio20v3-287:1  (local to host fmadio20v3-287)
           UUID : 06234ac8:694ae295:3659e4fc:b59a7d55
         Events : 0

    Number   Major   Minor   RaidDevice State
       0     259        0        0      active sync   /dev/nvme2n1
       1     259        1        1      active sync   /dev/nvme0n1
fmadio@fmadio20v3-287:~$
```

Create BTRFS Filesystem

The block device /dev/md1 is block level only, it contains no mountable file system. Next create btrfs filesystem on the device as follows

```
fmadio@fmadio20v3-287:~$ sudo mkfs.btrfs /dev/md1
btrfs-progs v5.12.1
See http://btrfs.wiki.kernel.org for more information.

Detected a SSD, turning off metadata duplication.  Mkfs with -m dup if you want to force me
Performing full device TRIM /dev/md1 (894.01GiB) ...
Label:              (null)
UUID:               5432ceac-b3d1-4b1c-8b69-d622542a9184
Node size:          16384
Sector size:        4096
Filesystem size:    894.01GiB
Block group profiles:
  Data:             single            8.00MiB
  Metadata:         single            8.00MiB
  System:           single            4.00MiB
SSD detected:       yes
Zoned device:       no
Incompat features:  extref, skinny-metadata
Runtime features:
```

```
Checksum:           crc32c
Number of devices:  1
Devices:
   ID        SIZE  PATH
    1   894.01GiB  /dev/md1

fmadio@fmadio20v3-287:~$
```

Mount the BTRFS Filesystem

By default FMADIO Packet Capture systems at boot time mount BTRFS with lzo disk compression. Compression can be enabled or disabled with BTRFS on-the-fly. In this case we will mount it the same as capture system does at boot time.

```
fmadio@fmadio20v3-287:~$ sudo mount -t btrfs -o compress=lzo /dev/md1 /mnt/store1
fmadio@fmadio20v3-287:~$
```

Then check the mount point with lsblk. Below we can see /dev/md1 is mounted on /mnt/store1

```
fmadio@fmadio20v3-287:~$ lsblk
NAME    MAJ:MIN RM   SIZE RO TYPE  MOUNTPOINT
sdd       8:48   0   3.7T  0 disk
sdb       8:16   0   3.7T  0 disk
nvme2n1 259:0    0 447.1G  0 disk
`-md1     9:1    0   894G  0 raid0 /mnt/store1
nvme1n1 259:3    0   477G  0 disk
sde       8:64   0   3.7T  0 disk
sdc       8:32   0   3.7T  0 disk
sda       8:0    0 238.5G  0 disk
|-sda2    8:2    0 223.6G  0 part  /mnt/store0
`-sda1    8:1    0  14.9G  0 part  /mnt/sda1
nvme0n1 259:1    0 447.1G  0 disk
`-md1     9:1    0   894G  0 raid0 /mnt/store1
nvme3n1 259:2    0   477G  0 disk
fmadio@fmadio20v3-287:~$

```

Check BTRFS compression level

Checking the compression level with BTRFS requires calculating the raw storage and the compressed storage.

```
fmadio@fmadio20v3-287:/mnt/store1$ du -h -d 1
3.0G    ./cache
3.0G    .
fmadio@fmadio20v3-287:/mnt/store1$


fmadio@fmadio20v3-287:/mnt/store1$ sudo btrfs fi show
```

```
Label: none  uuid: 5432ceac-b3d1-4b1c-8b69-d622542a9184
        Total devices 1 FS bytes used 751.07MiB
        devid    1 size 894.01GiB used 2.02GiB path /dev/md1

fmadio@fmadio20v3-287:/mnt/store1$

```

In the above example we see /mnt/store1 has 3.0 GB worth of data (using du)

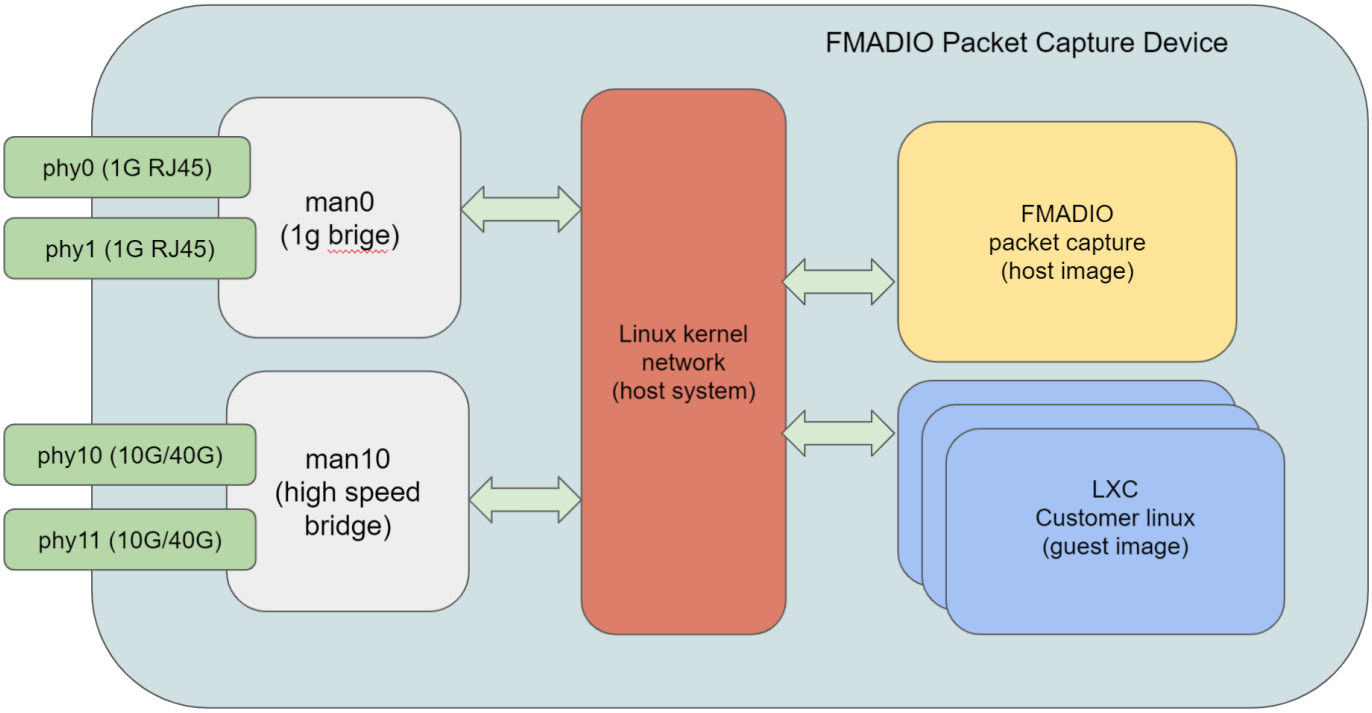
In the above example we see /mnt/store1 has used 751.MiB of actual storage capacity (using btrfs fi show)

Based on the above math (3112MB / 751MB) , the **compression rate is ~ x4.04**

Management Interface

FMADIO Systems have multiple 1G, 10G and 40G management interfaces, depending on the ordered SKU.

Management interfaces are all bridged by default per the following block diagram



FMADIO Management Interface Architecture

Using the above configuration allows

- LXC containers full pass-thru IP address (no NAT)
- Bonded management mode for Redundancy (Hot-Standby)
- Bonded management mode for Throughput (LAG)

Example ifconfig of the system is as follows

```

```

```
fmadio@fmadio20n40v3-363:~$ ifconfig -a
man0      Link encap:Ethernet  HWaddr 18:C0:4D:B4:0E:6C
          inet addr:192.168.2.225  Bcast:192.168.2.255  Mask:255.255.255.0
          inet6 addr: fe80::1ac0:4dff:feb4:e6c/64 Scope:Link
          UP BROADCAST RUNNING MULTICAST  MTU:1500  Metric:1
          RX packets:2371 errors:0 dropped:0 overruns:0 frame:0
          TX packets:136 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:1000
          RX bytes:911945 (890.5 KiB)  TX bytes:24130 (23.5 KiB)


man10     Link encap:Ethernet  HWaddr 18:C0:4D:17:7A:4E
          inet addr:192.168.15.225  Bcast:192.168.15.255  Mask:255.255.255.0
          inet6 addr: fe80::1ac0:4dff:fe17:7a4e/64 Scope:Link
          UP BROADCAST RUNNING MULTICAST  MTU:9200  Metric:1
          RX packets:5227 errors:0 dropped:0 overruns:0 frame:0
          TX packets:14 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:1000
          RX bytes:253670 (247.7 KiB)  TX bytes:1068 (1.0 KiB)

phy0      Link encap:Ethernet  HWaddr 18:C0:4D:B4:0E:6C
          inet6 addr: fe80::1ac0:4dff:feb4:e6c/64 Scope:Link
          UP BROADCAST RUNNING MULTICAST  MTU:1500  Metric:1
          RX packets:2372 errors:0 dropped:0 overruns:0 frame:0
          TX packets:151 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:1000
          RX bytes:945205 (923.0 KiB)  TX bytes:25260 (24.6 KiB)

phy1      Link encap:Ethernet  HWaddr 18:C0:4D:B4:0E:6D
          UP BROADCAST MULTICAST  MTU:1500  Metric:1
          RX packets:0 errors:0 dropped:0 overruns:0 frame:0
          TX packets:0 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:1000
          RX bytes:0 (0.0 B)  TX bytes:0 (0.0 B)


phy10     Link encap:Ethernet  HWaddr 18:C0:4D:17:7A:4E
          inet6 addr: fe80::1ac0:4dff:fe17:7a4e/64 Scope:Link
          UP BROADCAST RUNNING MULTICAST  MTU:9200  Metric:1
          RX packets:5262 errors:0 dropped:35 overruns:0 frame:0
          TX packets:27 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:1000
          RX bytes:331398 (323.6 KiB)  TX bytes:2066 (2.0 KiB)

phy11     Link encap:Ethernet  HWaddr 18:C0:4D:17:7A:4F
          BROADCAST MULTICAST  MTU:1500  Metric:1
          RX packets:0 errors:0 dropped:0 overruns:0 frame:0
          TX packets:0 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:1000
          RX bytes:0 (0.0 B)  TX bytes:0 (0.0 B)

fmadio@fmadio20n40v3-363:~$

```

And bridge settings

```
fmadio@fmadio20n40v3-363:~$ brctl show man0
bridge name     bridge id               STP enabled     interfaces
man0            8000.18c04db40e6c       no              phy0

fmadio@fmadio20n40v3-363:~$ brctl show man10
bridge name     bridge id               STP enabled     interfaces
man10           8000.18c04d177a4e       no              phy10

```

Management MTU Setting

By default MTU size is set to 1500B for maximum compatibility. This can be configure for 9200 Jumbo frame support to maximize download throughput. This is done by setting

```
["MTU"] = 9200,
```

For both man10 and phy10 network interfaces in the network configuration script below.

```
/opt/fmadio/etc/network.lua
```

This has to be set on both the man10 and phy10 (optionally phy11 if used) interfaces to be fully effective as per below example.

```
["man10"] =
{
    ["Mode"]    = "static",
    ["Address"] = "192.168.15.225",
    ["Netmask"] = "255.255.255.0",
    ["Gateway"] = "",
    ["DNS0"]    = "",
    ["DNS1"]    = "",
    ["MTU"]     = 9200,
}
,
["phy10"] =
{
    ["MTU"]  = 9200,
}

```

LACP Link Bonding

Requires FW:6508+

LACP or Link Bonding is critical for fail over / redundancy planning. FMADIO Packet Capture devices run on Linux thus we support LCAP/Bonding on the management interfaces.

```
/opt/fmadio/etc/network.lua
```

Add a bonded interface "bond0" as follows

```
["bond0"] =
{
    ["Mode"]    = "bond",
    ["Address"] = "192.168.1.2",
    ["Netmask"] = "255.255.255.0",
    ["Gateway"] = "192.168.1.1",
    ["DNS0"]    = "",
    ["DNS1"]    = "",
    ["Slave"]  = { "phy0", "phy1" }
},
```

In the above example the "Slave" field contains the list of physical interfaces the bonding runs on. This example is bonding the two 1G RJ45 interfaces on the system. To bond the 10G interfaces on a separate LCAP link (bond1), use the following:

```
["bond1"] =
{
    ["Mode"]    = "bond",
    ["Address"] = "192.168.1.2",
    ["Netmask"] = "255.255.255.0",
    ["Gateway"] = "192.168.1.1",
    ["DNS0"]    = "",
    ["DNS1"]    = "",
    ["Slave"]  = { "phy10", "phy11" }
},
```

LACP Bonding Mode

Requires FW: 6633+

By default 802.3ad bonding mode is used, full list of Linux bonding modes can be seen on kernel.org. Note "BondMode" specifies the Linux bonding mode to be used.

```
["bond1"] =
{
    ["Mode"]    = "bond",
    ["BondMode"] = "active-backup",
```

```
    ["Address"] = "192.168.1.2",
    ["Netmask"] = "255.255.255.0",
    ["Gateway"] = "192.168.1.1",
    ["DNS0"]    = "",
    ["DNS1"]    = "",
    ["Slave"]  = { "phy10", "phy11" }
},
```

Line Bonding mode options (details ripped from kernel.org)

Round-robin (balance-rr)

Transmit network packets in sequential order from the first available network interface (NIC) slave through the last. This mode provides load balancing and fault tolerance.

Active-backup (active-backup)

Only one NIC slave in the bond is active. A different slave becomes active if, and only if, the active slave fails. The single logical bonded interface's MAC address is externally visible on only one NIC (port) to avoid distortion in the network switch. This mode provides fault tolerance.

XOR (balance-xor)

Transmit network packets based on a hash of the packet's source and destination. The default algorithm only considers MAC addresses (layer2). Newer versions allow selection of additional policies based on IP addresses (layer2+3) and TCP/UDP port numbers (layer3+4). This selects the same NIC slave for each destination MAC address, IP address, or IP address and port combination, respectively. This mode provides load balancing and fault tolerance.

Broadcast (broadcast)

Transmit network packets on all slave network interfaces. This mode provides fault tolerance.

Default mode
IEEE 802.3ad Dynamic link aggregation (802.3ad, LACP)

Creates aggregation groups that share the same speed and duplex settings. Utilizes all slave network interfaces in the active aggregator group according to the 802.3ad specification. This mode is similar to the XOR mode above and supports the same balancing policies. The link is set up dynamically between two LACP-supporting peers.

Adaptive transmit load balancing (balance-tlb)

Linux bonding driver mode that does not require any special network-switch support. The outgoing network packet traffic is distributed according to the current load (computed relative to the speed) on each network

interface slave. Incoming traffic is received by one currently designated slave network interface. If this receiving slave fails, another slave takes over the MAC address of the failed receiving slave.

Adaptive load balancing (balance-alb)

includes balance-tlb plus receive load balancing (rlb) for IPV4 traffic, and does not require any special network switch support. The receive load balancing is achieved by ARP negotiation. The bonding driver intercepts the ARP Replies sent by the local system on their way out and overwrites the source hardware address with the unique hardware address of one of the NIC slaves in the single logical bonded interface such that different network-peers use different MAC addresses for their network packet traffic.

NOTE: PTPv2 and LCAP on the 10G Management interfaces are mutually exclusive.

FMAD Shark Config

Work in progress, contact support

LXC Configuration

Start by untaring the LXC container into the directory

```
/mnt/store0/lxc/lib/lxc/
```

As follows (Contact support for the tarball)

```
fmadio@fmadio20v3-287:/mnt/store0/lxc/lib/lxc$ sudo tar xf cloudshark5-1.tar.gz
fmadio@fmadio20v3-287:/mnt/store0/lxc/lib/lxc$ sudo ls -al cloudshark5
total 16
drwxrwx---    3 root     root          4096 Mar  4 12:20 .
drwxr-xr-x   12 fmadio   staff         4096 Jul 22 16:59 ..
-rw-r--r--    1 root     root           744 Jul 17 09:37 config
drwxr-xr-x   18 root     root          4096 Mar  4 12:30 rootfs
fmadio@fmadio20v3-287:/mnt/store0/lxc/lib/lxc$

```

Next Edit the LXC network config to allocate a specific IP address for CloudShark to run on. In the example below change the IPv4 address from 192.168.1.00/24 to the assigned IP/Prefix + update the IPv4 Gateway to be assigned correctly

```
root@fmadio20v3-287:/mnt/store0/lxc/lib/lxc# cat cloudshark5/config

# Distribution configuration
lxc.include = /usr/share/lxc/config/centos.common.conf
```

```
lxc.arch = x86_64
# Container specific configuration
lxc.rootfs.path = dir:/opt/fmadio/lxc/lib/lxc/cloudshark5/rootfs
lxc.uts.name = cloudshark5

# Network configuration
lxc.net.0.type = veth
lxc.net.0.link = man0
lxc.net.0.flags = up
lxc.net.0.ipv4.address = 192.168.1.100/24
#lxc.net.0.ipv4.gateway = 192.168.1.1
root@fmadio20v3-287:/mnt/store0/lxc/lib/lxc#

```

Next update the Same IPv4 address to RHEL7 network configuration file

```
root@fmadio20v3-287:/mnt/store0/lxc/lib/lxc# cat cloudshark5/rootfs/etc/sysconfig/network-s
DEVICE=eth0
BOOTPROTO=static
ONBOOT=yes
HOSTNAME=cloudshark5
NM_CONTROLLED=no
TYPE=Ethernet
IPADDR=192.168.1.100
NETMASK=255.255.255.0
GATEWAY=192.168.1.1
PREFIX=24
DNS1=192.168.1.1
root@fmadio20v3-287:/mnt/store0/lxc/lib/lxc#

```

FMADIO Configuration

edit config file

```
/mnt/store0/etc/time.lua
```

Specifically enable API v1 and **Cloudshark** as follows, specifically focusing on the changed items.

Please set

CloudShark = true

DownloadAPI = "v1"

as Shown below, if the fields do not exist, please create them

```
fmadio@fmadio20v3-287:/mnt/store0/etc$ cat time.lua

```

```
local Time=
{
.
.
.

["PCAP"] =
{
        .
        .

        ["CloudShark"]     = true,
        ["DownloadAPI"]     = "v1",
        .
        .

},
.
.
.
return Time
fmadio@fmadio20v3-287:/mnt/store0/etc$

```

Reboot the system or

sudo killall stream_http

sudo killall stream_command

TODO, deal with the hardcoded cloudshark ip. Not sure where to put that, probably in time.lua "CloudSharkIP"

Settings

HDD Writeback

FW: 7219+

Configuration options are in the specified config file. Please note all options requires capture to be stopped and started before settings are applied.

```
/opt/fmadio/etc/time.lua
```

Specifically the Writeback block, example as follows

```
["Writeback"] =
{
        ["IOPriority"]       = 11,
        ["EnableCompress"]   = true,
        ["CheckBlockCRC"]    = true,
        ["EnableECC"]        = true,
},
```

IOPriority

This setting changes the default writeback IO Priority, allows changing preference for faster Downloads(default) or faster sustained Writeback to magnetic storage.

Setting	Description
11	Writeback to Magnetic storage is lowest priority (Default Value)
30	Writeback to disk has higher priority than Download or Push speed

EnableCompress

Setting enables or disables disk compression. For faster sustained writeback to disk speeds, disable compression.

For 1U systems disabling compression makes little difference due to lower of HDD write bandwidth.

For 2U systems disabling compression improves sustained writeback to HDD performance. As the system has 12 spinning disks with an aggregate 10Gbps-20Gbps (depending on spindle position) write throughput.

Setting	Description
true	Enable Compression (Default setting)
false	Disable disk compression. Faster sustained writeback performance on 2U systems

CheckBlockCRC

This function checks the CRC when reading data from the SSDs. It calculates the CRC and checks for a match against the orignial captured CRC value, before writing the block to magnetic storage. This adds additional CPU overhead.

Disabling this improves the sustained write performance on 2U systems. On 1U systems there is little performance advantage

Setting	Description
true	Enable SSD CRC Checks before writing to Magnetic Storage (Default setting)
false	Do not check SSD CRC data check. This improve sustained writeback performance

EnableECC

This is mostly a debug setting. by disabling ECC it removes the RAID5 calculation and additional IO writeback. Turning the system into a RAID0 configuration. This is mostly for debug testing and not recommend for production systems

Setting	Description
true	Calculate ECC RAID5 Parity (Default setting)
false	No ECC calculation (RAID0 mode)

Recommended Settings

The default settings are recommended unless there is specific use cases.

Maximum Sustained Capture Rate

For Maximum sustained capture rate we recommended the following settings. This disables the compression and priorities Magnetic Storage writeback over download performance.

```
["Capture"] =
{
    ["Inline"]             = false,
    ["PortMode"]           = "2x10G",
    ["FlushPktCnt"]        = 2000,
    ["FlushPeriod"]        = 0,
    ["FlushIdle"]          = 1000000000,

    ["IOPriorityLevel"]    = 40,
    ["IOPriorityHighByte"] = 128*1024*1024,

    ["CaptureSizeMax"]     = 0,
```

```
    ["CaptureTimeMax"]     = 0,
    ["DisableCPUPriority"] = true,
},
["Writeback"] =
{
    ["IOPriority"]         = 50,
    ["EnableCompress"]     = false,
    ["EnableECC"]          = true,
    ["CheckBlockCRC"]      = false,
},
```

Capture

There are various tunable available only via configuration file editing the fie

```
/opt/fmadio/etc/time.lua
```

Then in the ["capture"] section editing each field, an example is shown below

```
["Capture"] =
{
        ["Inline"]      = false,
        ["PortMode"]    = "2x10G",
        ["FlushPktCnt"] = 2000,
        ["FlushPeriod"] = 0,
        ["FlushIdle"]   = 1000000000,
        ["IOPriorityLevel"]   = 90,
        ["IOPriorityHighByte"]   = 1048576,
        ["CaptureSizeMax"]   = 0,
        ["CaptureTimeMax"]   = 0,
},
```

CaptureSizeMax

When the total daily capture sizes start exceeding 10TB / day file sizes can get bit too large and difficult to work with. This setting sets the maximum size of a single capture, then rolling (losslessly) to a new Capture when the limit is reached.

Example below rolls the capture file every 1TB

```
 ["CaptureSizeMax"]   = 1e12,
```

CaptureTimeMax

In addition to maximum size, for large ((10TB+ daily) capture rates a more simpler approach is to roll the

capture every 1H. This reduces the size of each capture to something more manageable.

Example below rolls the capture every 1H (units are in nano seconds)

```
["CaptureTimeMax"]   = 60*60*1e9,
```

DisableCPUPriority

This setting a debug only as it (potentially) reduces Capture performance, specifically on 100G and higher capture systems. Only enable this if directed by FMADIO Support

```
 ["DisableCPUPriority"]   = true,
```

Confirm the setting by checking log file /mnt/store0/log/stream_capture f20.cur where the following log entries will be seen.

```
[20210617_175614] fNIC100_RxPollLoop  : 2319 | Disable RxPoll Loop SCHED_RR
[20210617_175614] fNIC100_RxIndexLoop : 2575 | Disable RxIndex 0 Loop SCHED_RR
[20210617_175614] fNIC100_RxIndexLoop : 2575 | Disable RxIndex 0 Loop SCHED_RR
[20210617_175614] fNIC100_RxIndexLoop : 2575 | Disable RxIndex 1 Loop SCHED_RR
```

FlushPktCnt

This is the number of packets to send (per pipeline) when the Capture pipline has to be flushed. Default is 2000

```
    ["FlushPktCnt"] = 2000,
```

FlushPeriod

When in continuous output flush mode this is the period (in nanoseconds) between flushes. Disabling constant period flushing set this to 0. Default is 0 (disable)

```
    ["FlushPeriod"] = 0,
```

FlushIdle

This is the idle packet activity timeout. If no ***new*** packets are received within this period, the pipeline gets flushed. Default value is 1e9 (1 second)

```
    ["FlushIdle"]   = 1e9,
```

ManualOffset

Disables captures midnight roll and instead applies a manual offset to the capture roll time.

Value	Description
0	Midnight roll enabled
nil	Midnight roll enabled
1	Midnight roll disabled
<nanosecond amount>	Manual time offset (e.g. GMT offset -9*60*1e9)

Value in nanoseconds, example is offset of -9 hours backwards (GMT - 9) from the local timezone. This may helpful when the local timeones midnight does not align with an appropriate time to roll the capture file (e.g. a Pool of FMADIO global probes should roll the capture at midnight GMT irrespective of the local timezone)

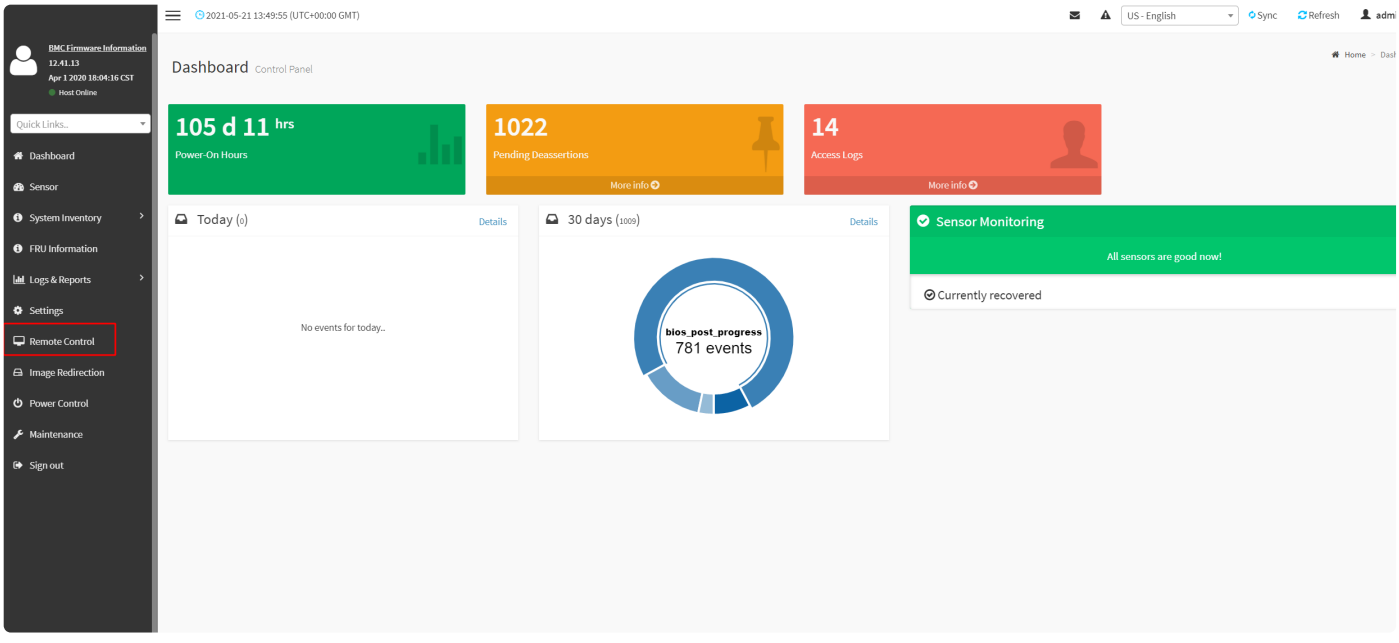
```
1     ["MidnightRollDisable"]   = -9*60*1e9,
```

Troubleshooting

System Recovery

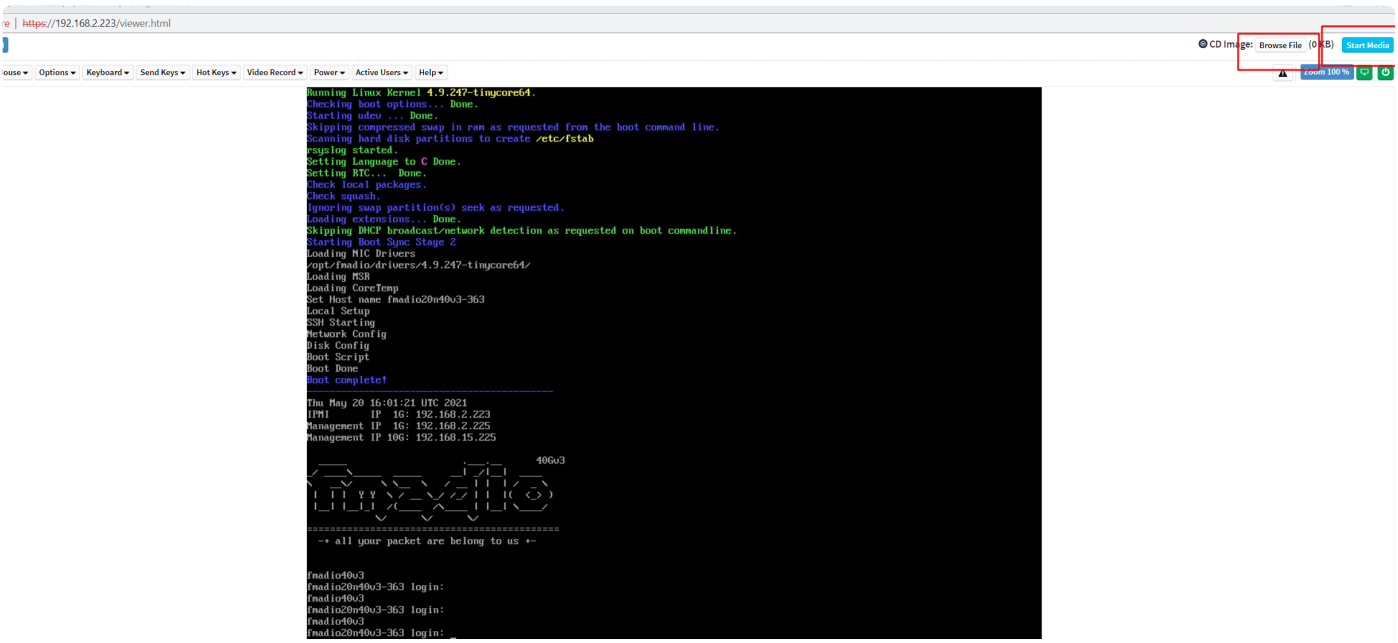
In the unlikely event of a complete boot failure, system can be recovered by booting via the Virtual CDROM interface over a HTML BMC connection

Start by going to the BMC interface (default IP is 192.168.0.93) contact us for default login/password



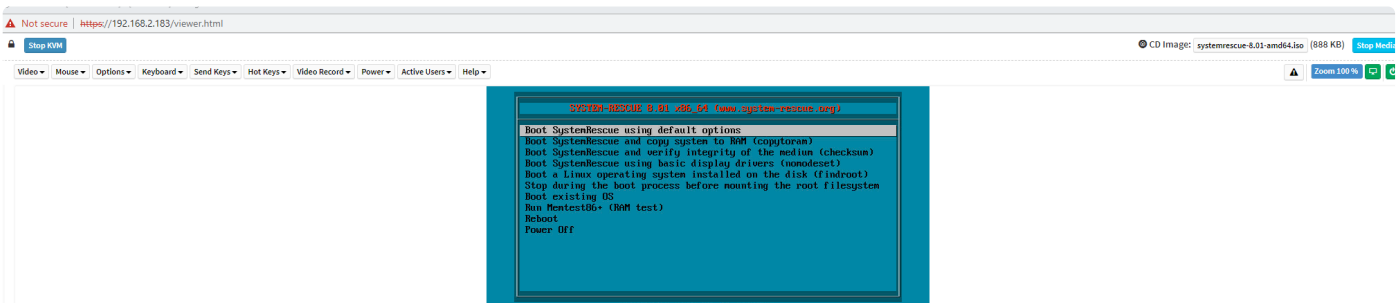
Start the Remote HTML KVM

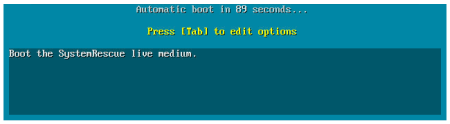
Will look like this. Select Brose Files, selecting an ISO image + Start the Media



System will boot Ubuntu (for example), we are using (systemrescue 8.01 amd64)

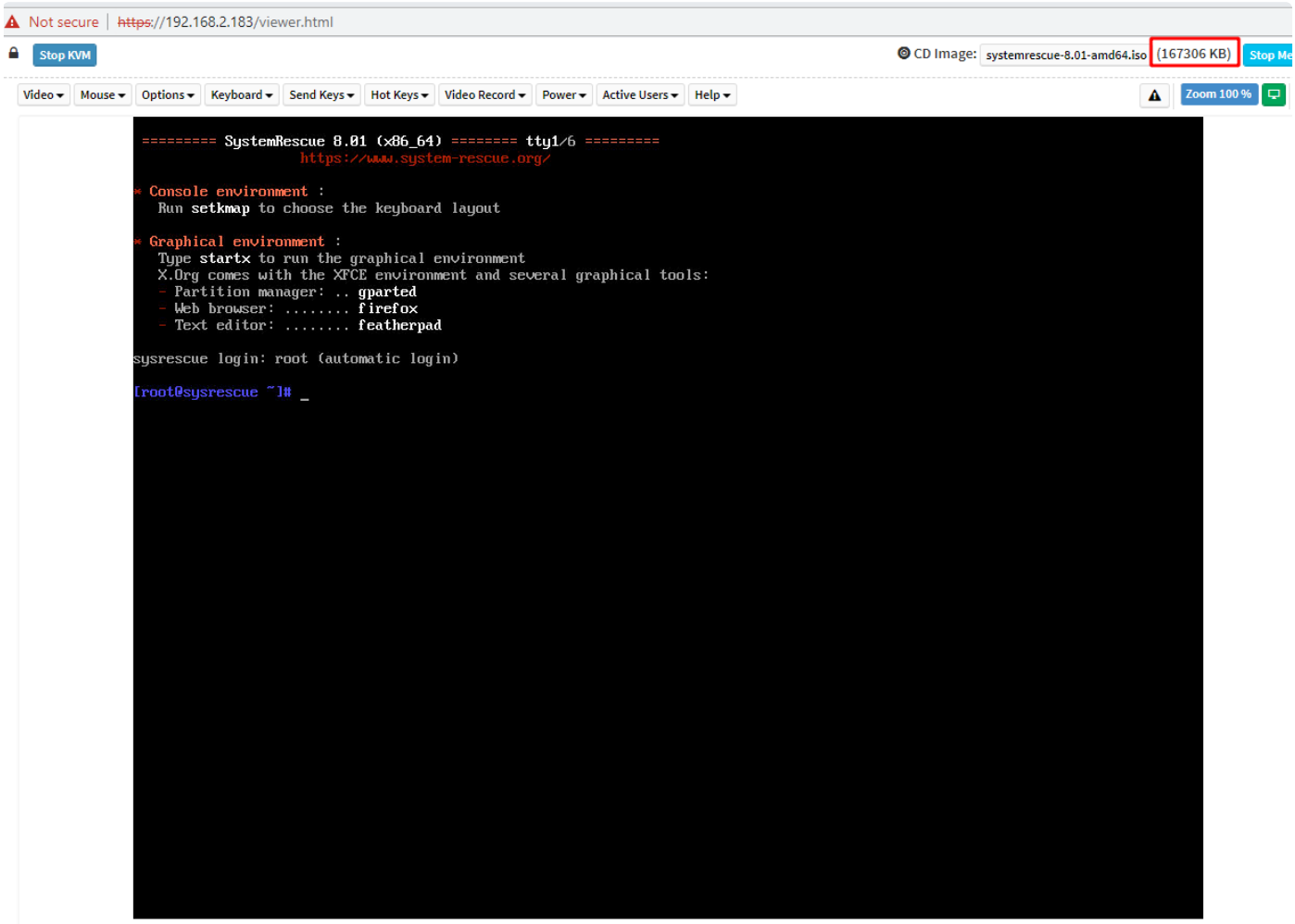
https://sourceforge.net/projects/systemrescuecd/files/sysresccd-x86/8.01/systemrescue-8.01-amd64.iso/download





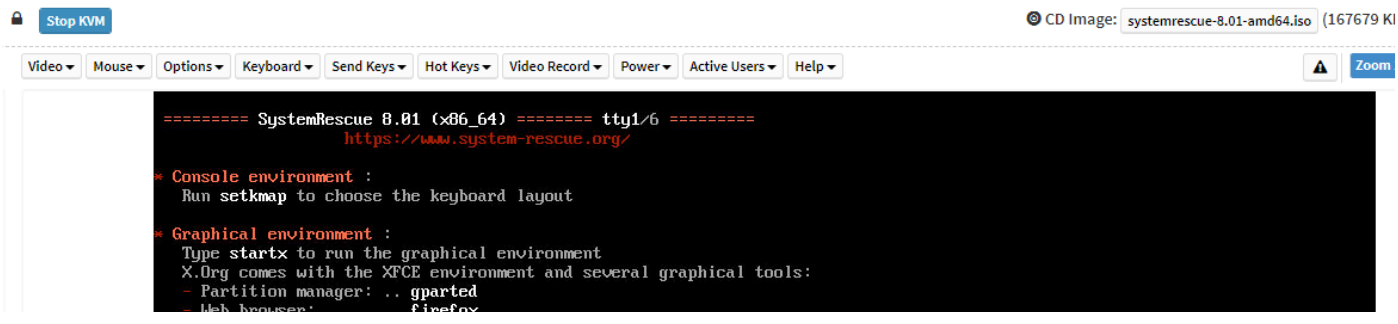
System will boot as follows, it may take several minutes depending on the speed of the HTML <-> FMADIO System connection. Recommend the closer the HTML instance is to the FMAD device the better.

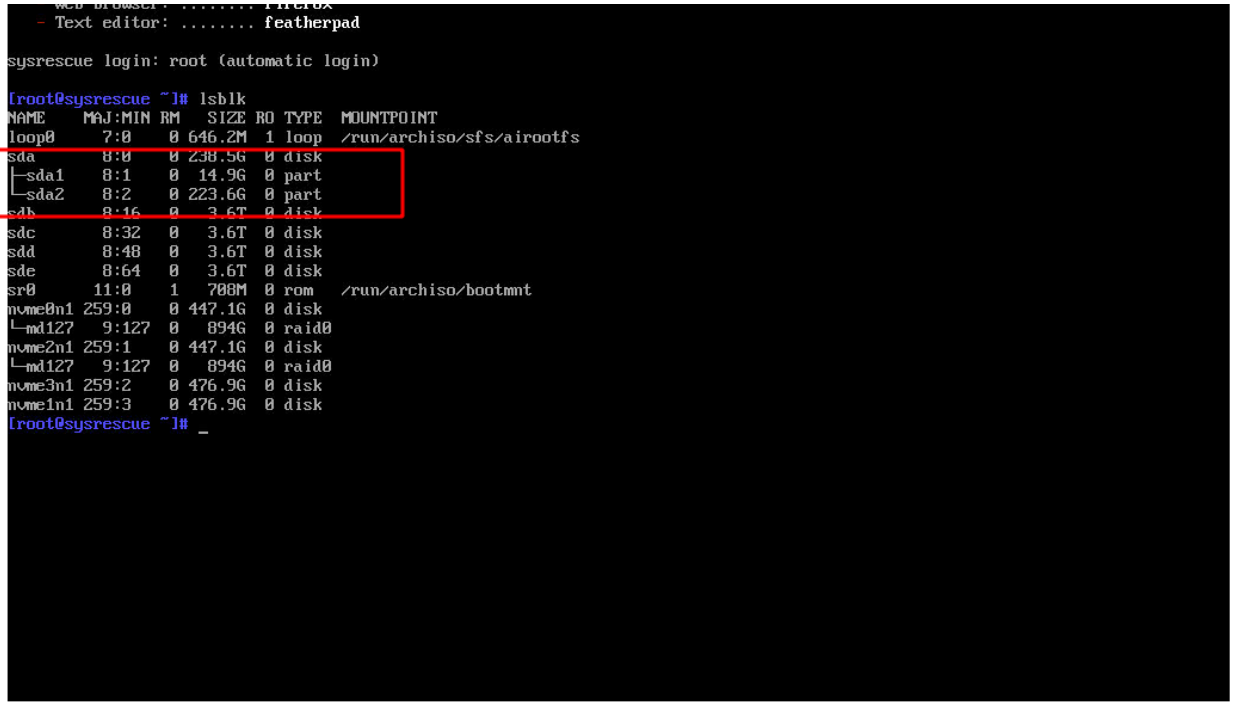
If a particular boot stage is taking too long Ctrl-C can skip it



After SystemRescue CD has booted, the above is seen. Note the total number of bytes transfered over the Virtual ISO.

First step is to find the FMADIO OS and Persistant storage devices, Use the "lsblk" tool





Looking foor a small (15GB) partition as the OS boot disk. In this case its sda1 and a large (224GB or larger) partition for the Persistent storage

Sometimes its easier to work over SSH. To access the system find or assign an IP address to the a reachable interface

SystemRescuelCD by default has iptables setup. Disable all iptables as follows

```
iptables -F
iptables -X
systemctl stop iptables
```

Then setup a password for the root account

```
passwd
```

Then ssh access to the system is possible

```
aaron@ingress:~$ ssh root@192.168.2.121
The authenticity of host '192.168.2.121 (192.168.2.121)' can't be established.
ECDSA key fingerprint is SHA256:v2CQjmUL70YpMJh39GWhcyqanKUU4eqLXxjTg/2i35Q.
Are you sure you want to continue connecting (yes/no)? yes
Warning: Permanently added '192.168.2.121' (ECDSA) to the list of known hosts.
root@192.168.2.121's password:
[root@sysrescue ~]#
```

Next mount the FMAD OS and Persistant storage disks. They may be sda* or nvme0n1p* in this example its mapped to sda

```
[root@sysrescue ~]# cd /mnt
[root@sysrescue /mnt]# mkdir system
[root@sysrescue /mnt]# mkdir store0
[root@sysrescue /mnt]# mount /dev/sda1 system/
[root@sysrescue /mnt]# mount /dev/sda2 store0/
[root@sysrescue /mnt]#
```

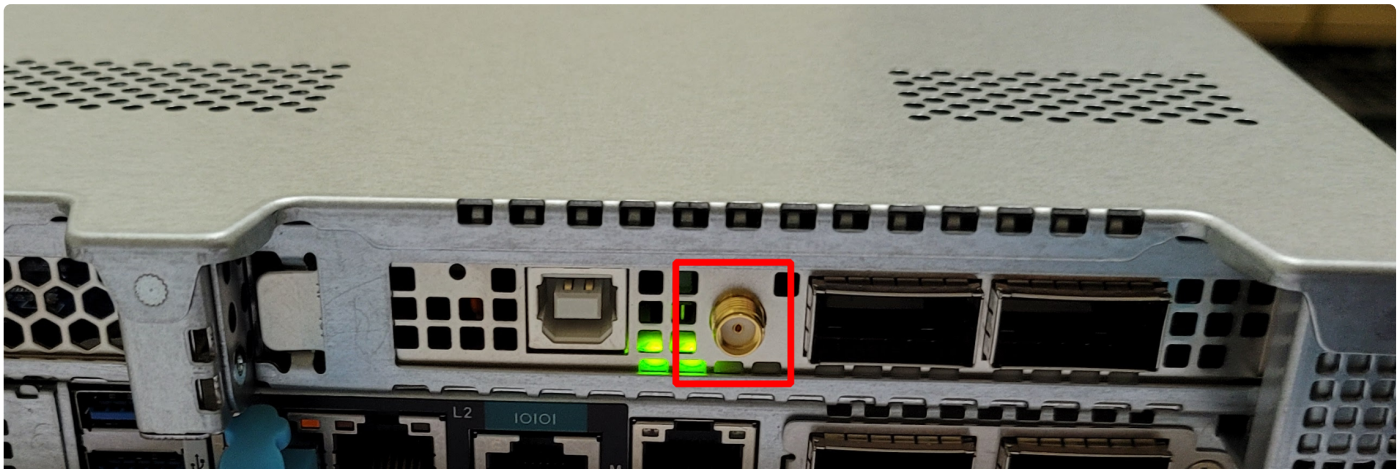
Next check the contents, it should look roughtly like this

```
[root@sysrescue /mnt]# ls -al /mnt/system/
total 64
drwxr-xr-x 5 root root  8192 Jan  1  1970 .
drwxr-xr-x 1 root root    80 May 22 08:53 ..
drwxr-xr-x 3 root root  8192 Apr 18 15:16 boot
drwxr-xr-x 2 root root 40960 Apr 18 15:16 firmware
drwxr-xr-x 5 root root  8192 May 11 10:13 tce
[root@sysrescue /mnt]# ls -al /mnt/store0
total 17244
drwxrwxrwx 32 root   root      4096 May 11 08:32 .
drwxr-xr-x  1 root   root        80 May 22 08:53 ..
drwxr-xr-x  3 root   root      4096 May 11 11:29 etc
drwxr-xr-x  2 root   root      4096 Dec 17  2019 filter
drwxrwxrwx  4 root   root  17477632 May 22 08:37 log
drwx------  2 root   root     16384 Dec 16  2019 lost+found

drwxr-xr-x  6   1002 games     4096 Oct 12  2019 lxc
drwxr-xr-x  3   1002 games     4096 Aug  5  2020 plugin_data
drwxr-xr-x  2 root   root      4096 Dec 17  2019 stream
drwx------  4 nobody root      4096 Dec 29 03:23 tmp
drwxrwxrwx 10 root   root      4096 Mar 23 10:25 tmp2
[root@sysrescue /mnt]#
```

Global Time Synchronization

FMADIO Packet captures systems have the ability to get < 100nsec global world time synchronization using PTPv2 + PPS signal.



Example SMA PPS Connector (FMAD100G)

Pulse Per Second (PPS) Clock

PPS time synchronization is a one pule per second signal, typically from a PTP Grand Master, or GPS based global time master. Its connected over a SMA connector directly to the FMADIO Capture FPGA.

The PPS signal disciplines the start of the global time second, it does not set the actual world time, e.g. its used to set when the start of a second begins.

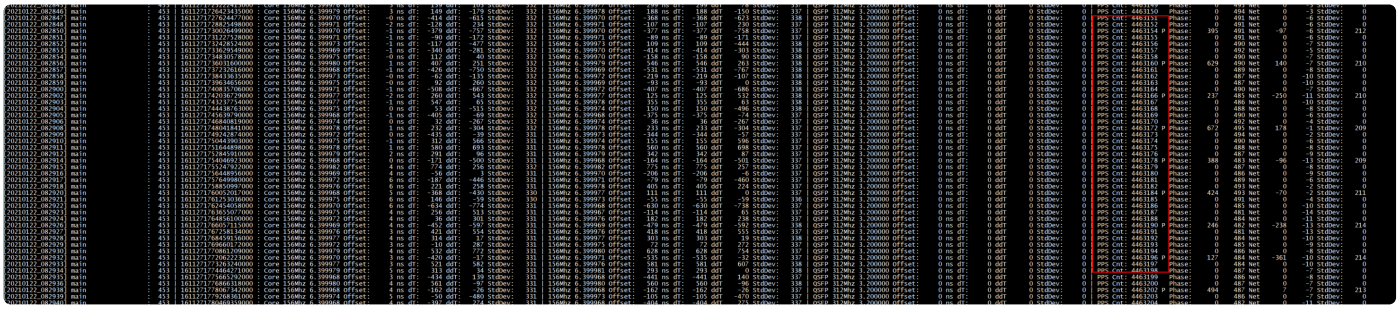
```
Pulese 0 | YYYYY-MM-DD HH:MM:DD:(SS + 0) 000.000.000  nanos
Pulese 1 | YYYYY-MM-DD HH:MM:DD:(SS + 1) 000.000.000  nanos
Pulese 2 | YYYYY-MM-DD HH:MM:DD:(SS + 2) 000.000.000  nanos
Pulese 3 | YYYYY-MM-DD HH:MM:DD:(SS + 3) 000.000.000  nanos
Pulese 4 | YYYYY-MM-DD HH:MM:DD:(SS + 4) 000.000.000  nanos
```

FMADIO system uses PTPv2 to set the YYYY-MM-DD HH:MM:SS part of the time.

To confirm PPS signal is active please check the log file

```
/mnt/store0/log/fnic_clock.cur
```

The following screenshot shows an example PPS signal incrementing every 1 second



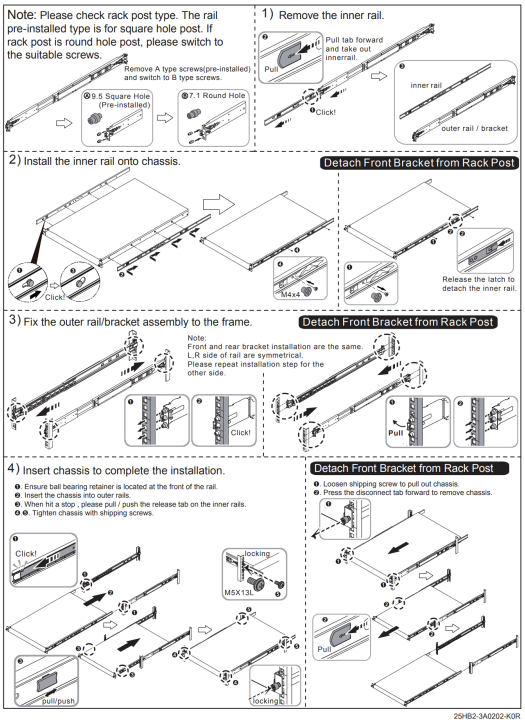
PPS Signal Incrementing

If the PPS Cnt is not incrementing, there may be a problem with the SMA connection or PPS voltage/pulse specification. Please contact support for further assistance.

Maintenance

Rack mount install

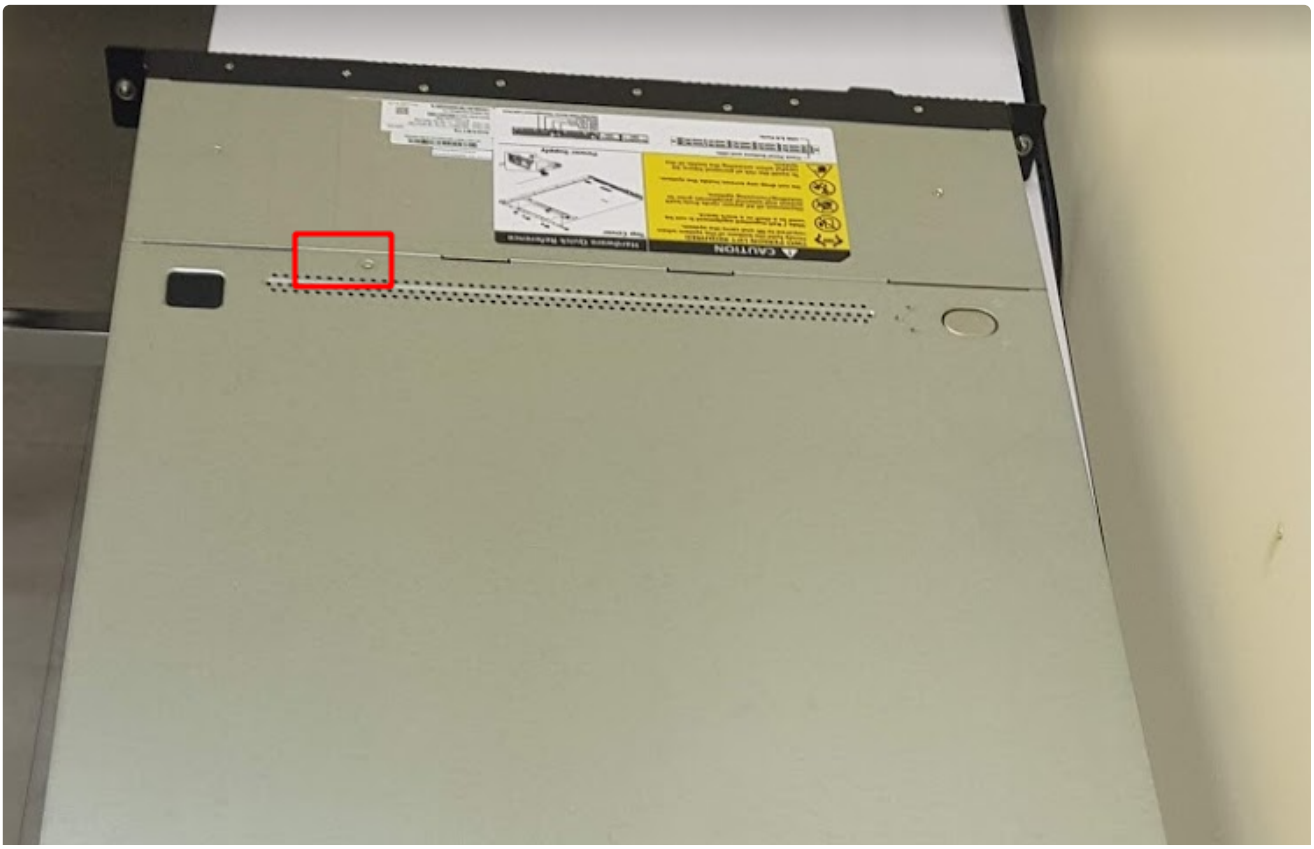
Instructions for rack mounting the system. Rails are the same for both 1U and 2U chassis

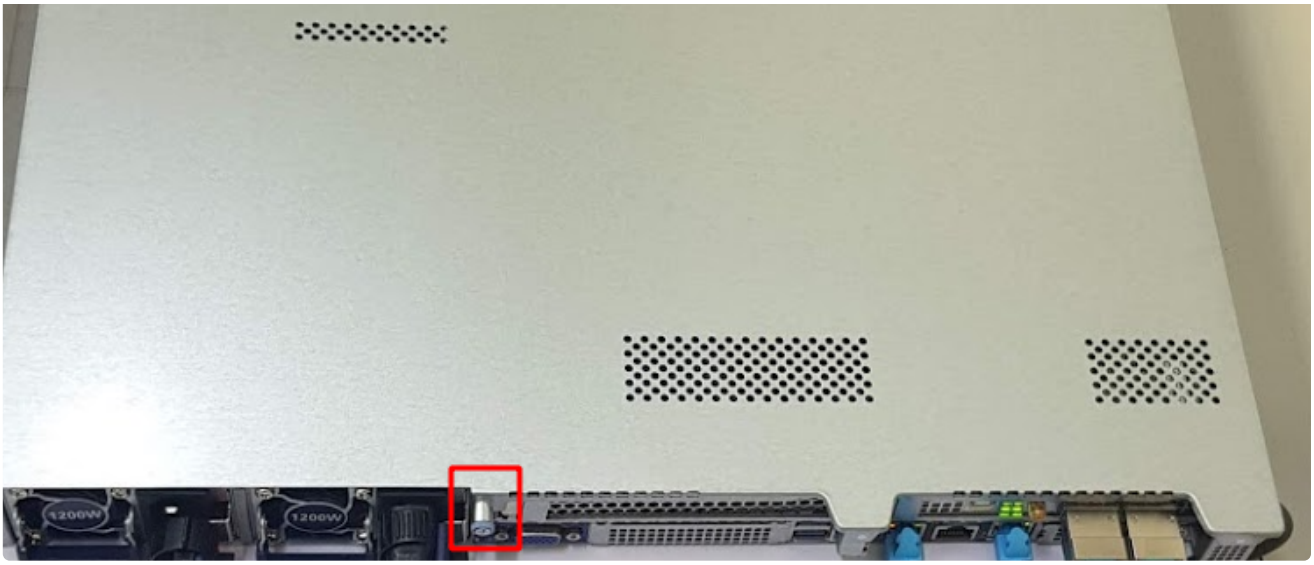


SSD Hardware replacement

Replacement of SSDs is straight forward but requires unracking and removing PCIe devices

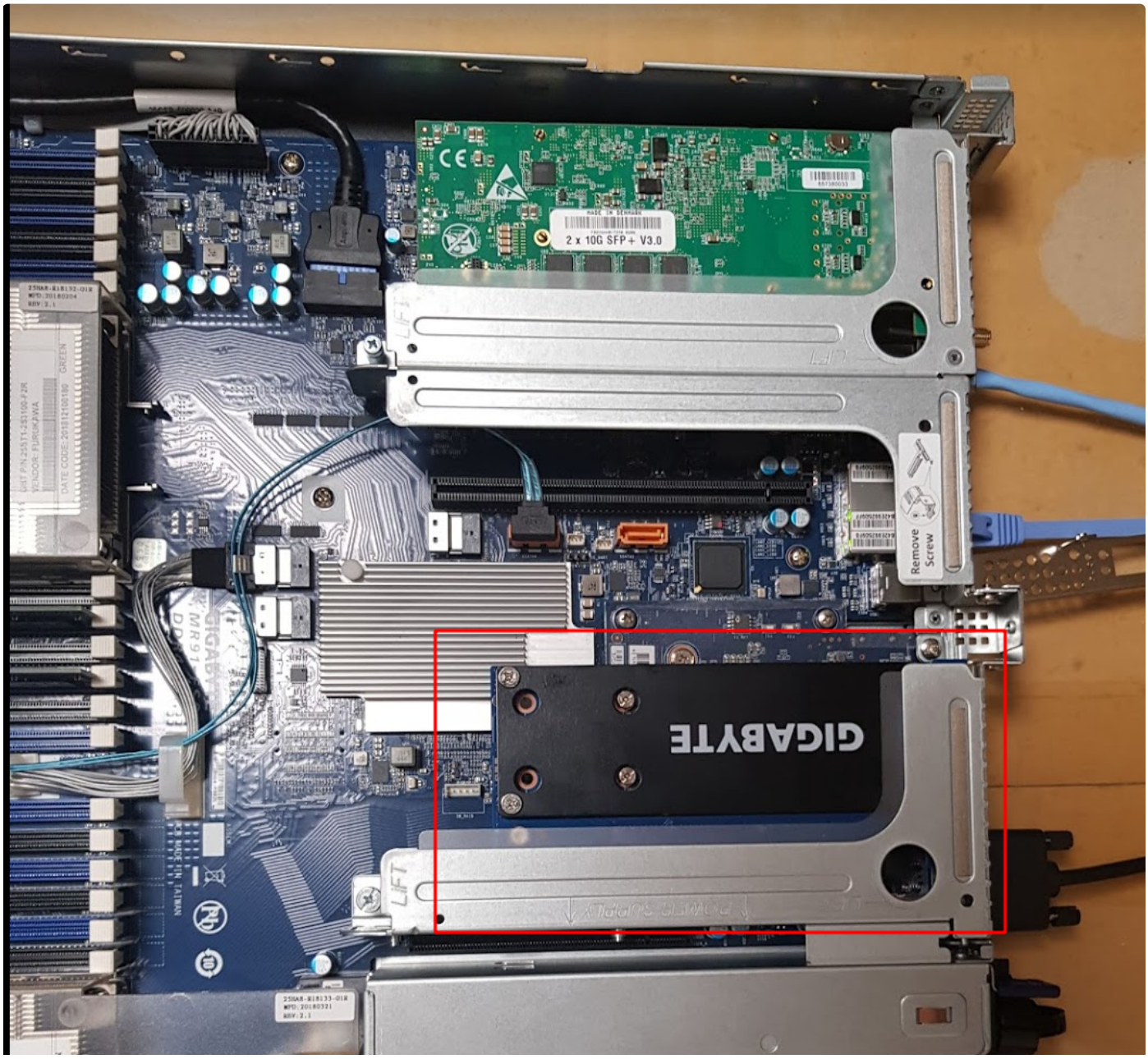
Once unracked remove the top cover by un-screwing the system as follows





FMADIO20Gv3 Remove Cover

Once removed the system looks like the following. The SSDs are located at the following location, shown in RED with the GIGABYTE thermal heatsink



FMADIO20G Gen3 PCIe slot layout

To remove the M.2 SSD drives remove the screwes shown as follows

The PCIe riser is detached from the motherboard

SSD PCIe card is removed from the Riser

SSD PCIe black heatsink is removed from the board

After the PCIe card has been removed, the SSDs are accessible as follows

Remove the screws highlighted in RED above to remove and replace the SSDs

After completing the SSD replacement, reverse the above steps to complete the installation

HDD Hardware replacaement

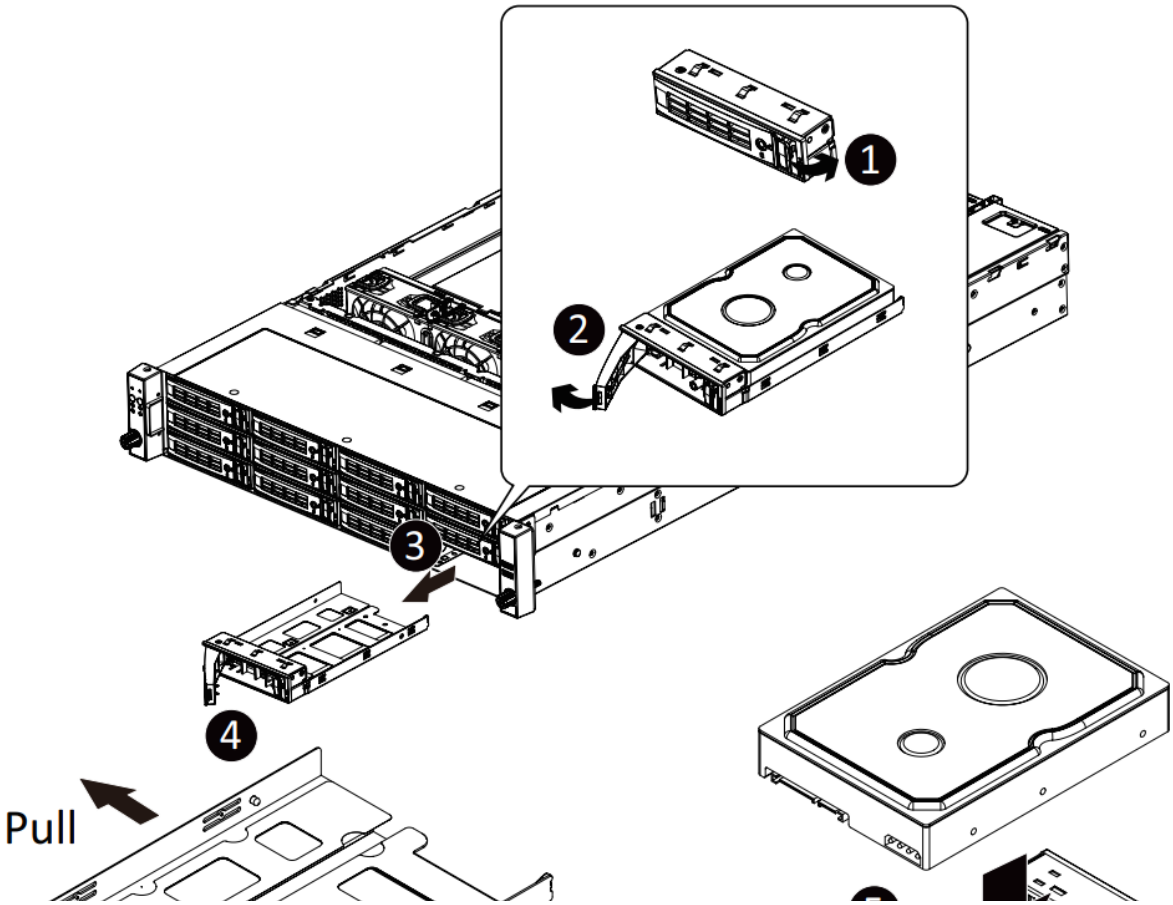
3-6 Installing the Hard Disk Drive

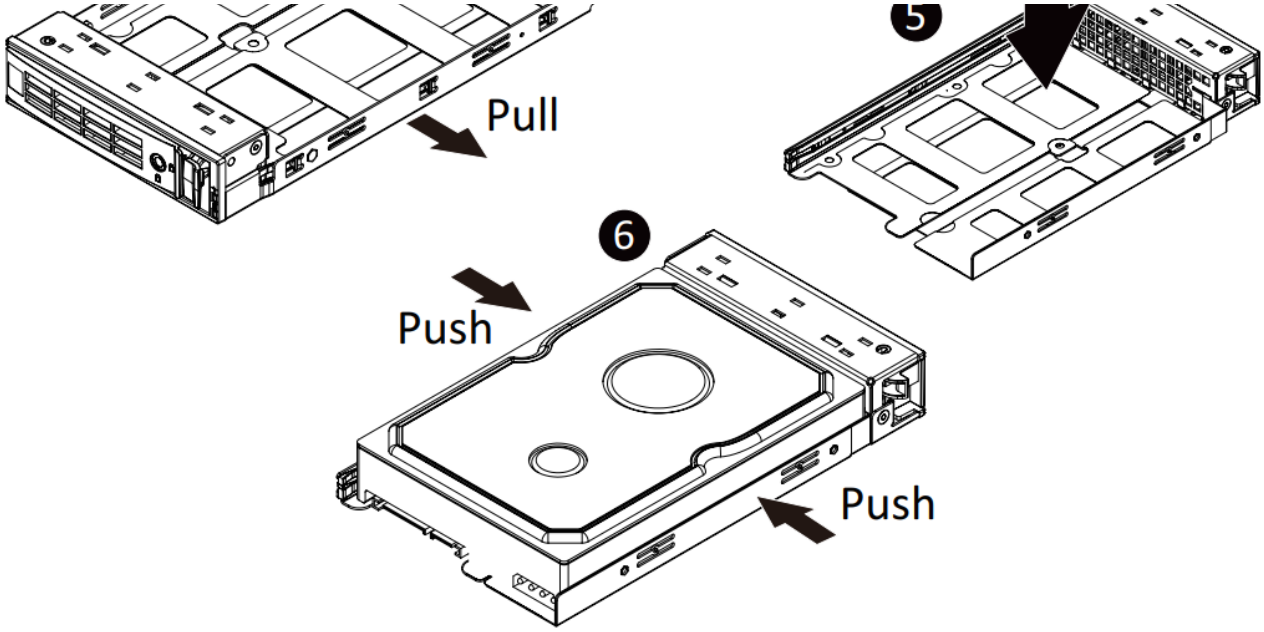
Read the following guidelines before you begin to install the Hard disk drive:

- Take note of the drive tray orientation before sliding it out.
- The tray will not fit back into the bay if inserted incorrectly.
- Make sure that the HDD is connected to the HDD connector on the backplane.

Follow these instructions to install the Hard disk drive:

1. Press the release button.
2. Pull the locking lever to remove the HDD tray.
3. Pull apart the HDD tray.
4. Slide hard disk into the tray.
5. Push together to secure the hard drive.




Pull
5
6
Push
Push