# A Simple Scheme for Processing Global VIIRS RDRs Via CSPP
## (A Poor-Man's DIY Guide)

Steve Finley

Cooperative Institute for Research in the Atmosphere
Colorado State University, Fort Collins, Colorado


CIRA
Connecting Models and Observations


- ISU
- CSU
- FAA
- DoD
- PrivacyNetworks
- CIRA

# Outline

- About CIRA

# Outline

- About CIRA
- CSPP Processing of VIIRS L1/L2 products

# Outline

- About CIRA
- CSPP Processing of VIIRS L1/L2 products
  - Questions

# Outline

- About CIRA
- CSPP Processing of VIIRS L1/L2 products
  - Questions
- CIRA Websites (and other shameless plugs)

# Outline

- About CIRA
- CSPP Processing of VIIRS L1/L2
  - Questions
- CIRA Websites (and other shameless plugs)
- Ceph data storage intro. (if time allows)

CIRA
Connecting Models and Observations

# About CIRA (not the whole list)

# About CIRA (not the whole list)

- CloudSat DPC (no DB)
  - Data Storage
  - L2 (L3?) Product Development
  - Product Dissemination

# About CIRA (not the whole list)

- CloudSat DPC (no DB)
- Mesoscale/Regional Scale Modeling
  - Linux parallel processing clusters
  - RAMS/WRF

# About CIRA (not the whole list)

- CloudSat DPC (no DB)
- Mesoscale/Regional Scale Modeling
- Data Assimilation – Data Fusion

# About CIRA (not the whole list)

- CloudSat DPC (no DB)
- Mesoscale/Regional Scale Modeling
- Data Assimilation – Data Fusion

# About CIRA (not the whole list)

- CloudSat DPC (no DB)
- Mesoscale/Regional Scale Modeling
- Data Assimilation – Data Fusion
- A few DoD projects (Andy Jones)
  - DPEAS

# About CIRA (not the whole list)

- CloudSat DPC (no DB)
- Mesoscale/Regional Scale Modeling
- Data Assimilation – Data Fusion
- A few DoD projects (Andy Jones)
- Proving Ground (Steve Miller)

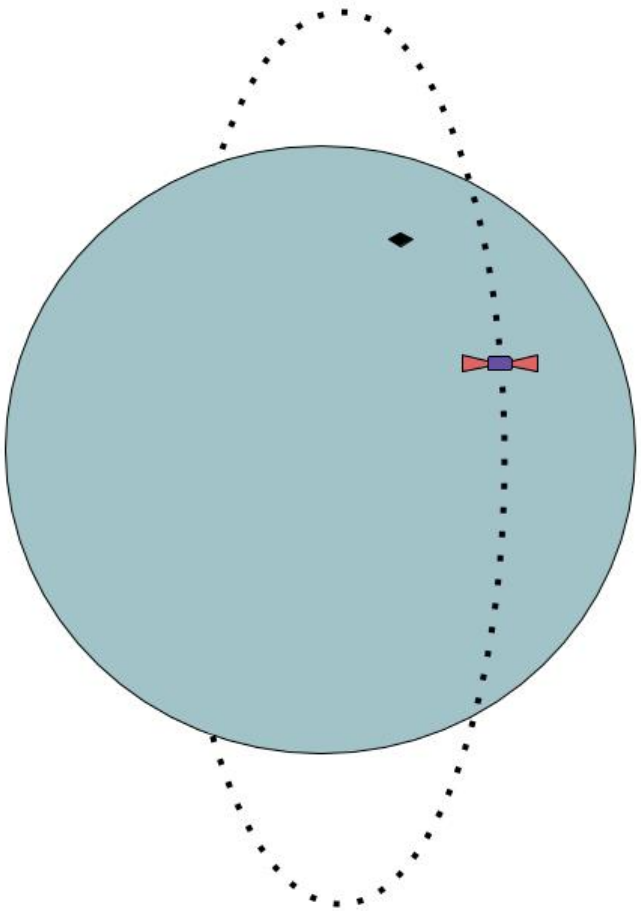
# About CIRA (not the whole list)

- CloudSat DPC (no DB)
- Mesoscale/Regional Scale Modeling
- Data Assimilation – Data Fusion
- A few DoD projects (Andy Jones)
- Proving Ground (Steve Miller)
- RAMMB (NESDIS/STAR)
  - NOAA/NWS product development
  - GVAR and GRB
  - Tropical Cyclones (NHC/JTWC)
  - S-NPP VIIRS EDR Imagery Evaluation

# Global CSPP VIIRS Processing

- Science RDR (Lo) data from remote servers (rather than DB)
- The Platform
- The Scheme
- The Metrics
- The Issues

# Global CSPP VIIRS Processing



# Global CSPP VIIRS Processing

- Science RDR (Lo) data from remote servers (rather than DB)
  - GRAVITE

# Global CSPP VIIRS Processing

- Science RDR (Lo) Data from remote servers (rather than DB)
  - GRAVITE
  - PDA

# Global CSPP VIIRS Processing

- Science RDR (Lo) Data from remote servers (rather than DB)
  - GRAVITE
  - PDA
  - DR sites (AK)

# Why not just grab all the VIIRS L1/L2 files from GRAVITE (or PDA)?

- Network bandwidth does not allow it in a timely manner via the Commodity Internet
- Local generation via CSPP is faster
- Responsible use of your local network

# The Platform

- 20-core Intel(R) Xeon(R) 2.20Ghz

# The Platform

- 20-core Intel(R) Xeon(R) 2.20Ghz
- 128GiB Memory

# The Platform

- 20-core Intel(R) Xeon(R) 2.20Ghz
- 128GB Memory
- CentOS 7 Linux

# The Platform

- 20-core Intel(R) Xeon(R) 2.20Ghz
- 128GB Memory
- CentOS 7 Linux
- 10Gbps network

# The Platform

- 20-core Intel(R) Xeon(R) 2.20Ghz
- 128GB Memory
- CentOS 7 Linux
- 10Gbps network
- ~$6500 (US)

# The Scheme

Requirements

# The Scheme

Requirements

- Need three RDRs *in sequence* to generate one good SDR

# The Scheme

Requirements

- Need three RDRs *in sequence* to generate one good SDR
  - h5check

# The Scheme

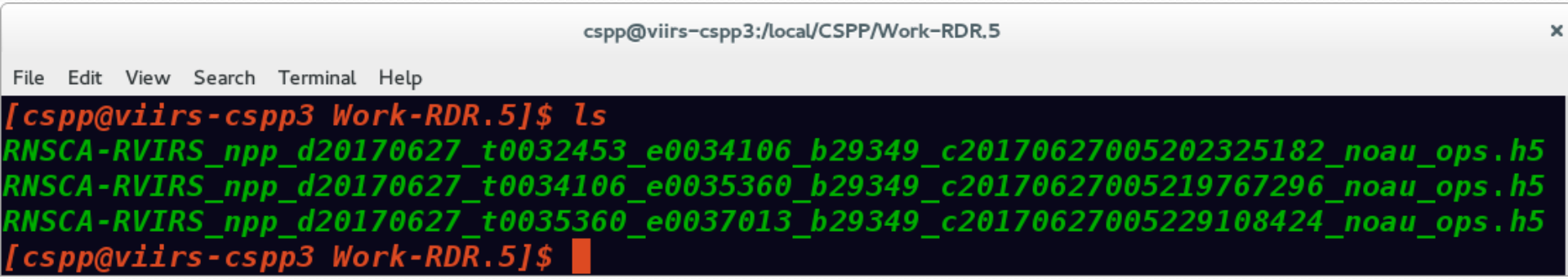
Requirements

- Need three RDRs *in sequence* to generate one good set of SDR

- Need three SDRs *in sequence* to generate EDRs

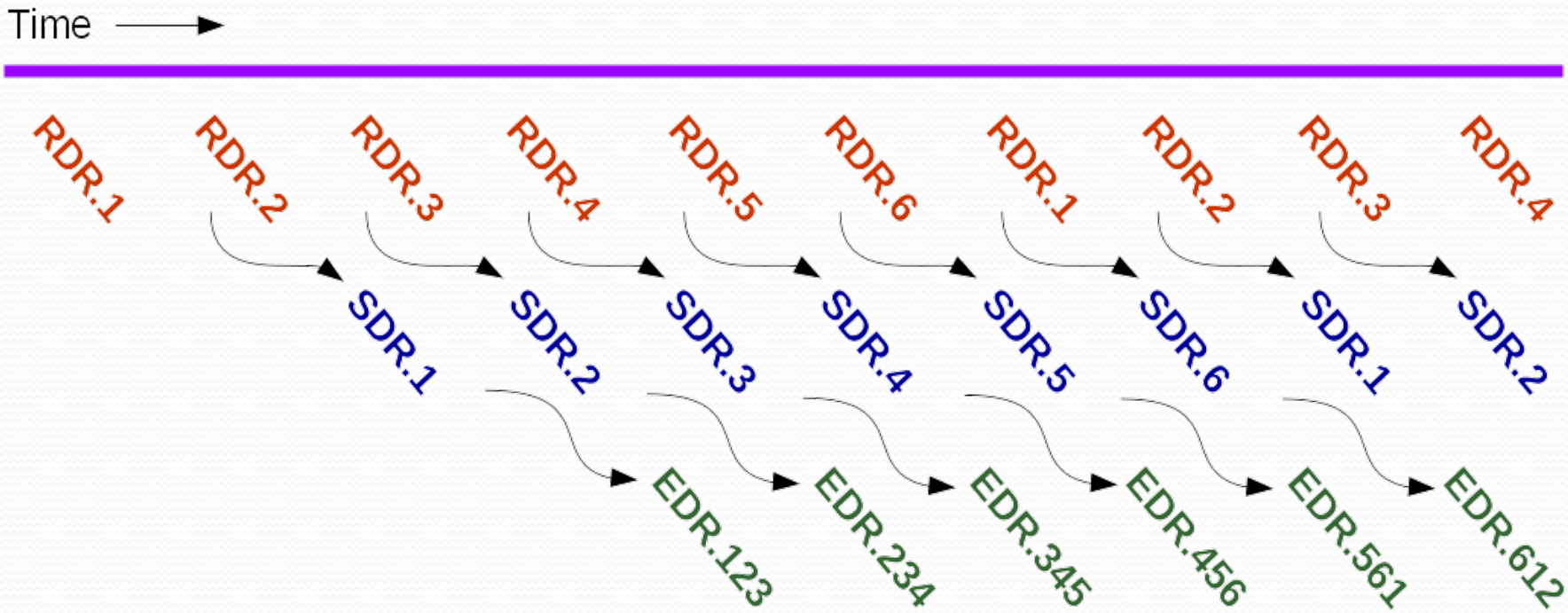
# The Scheme

## RDR sequence



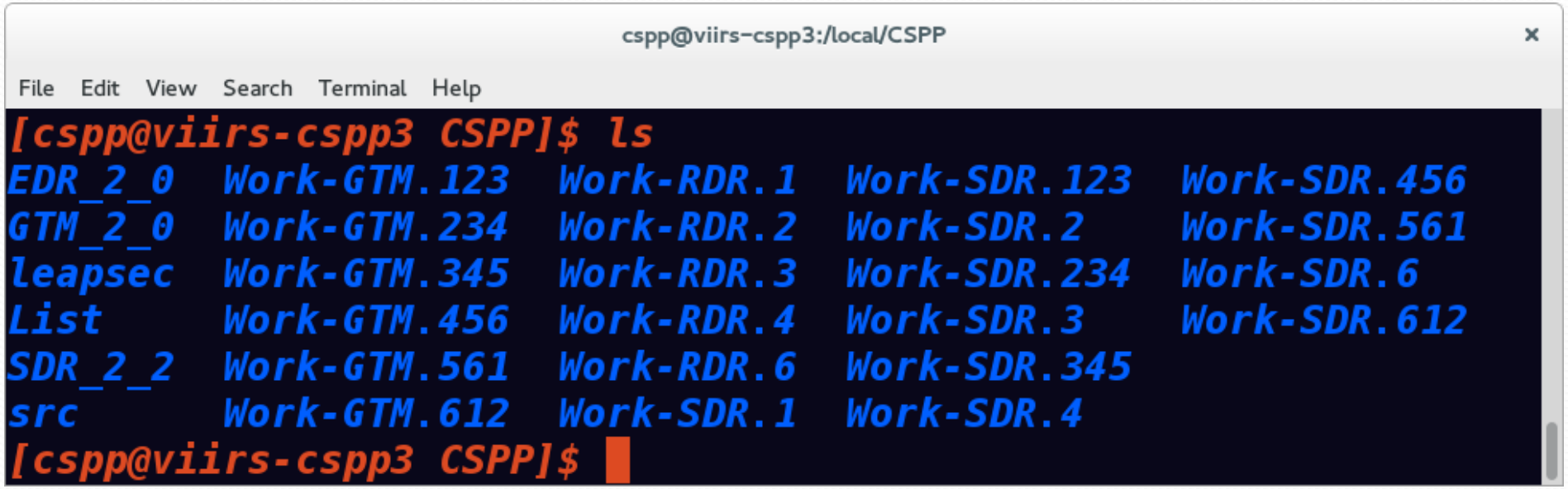
# The Scheme

## Timeline



# The Scheme

## Work directories



# The Scheme

## Crontab entries

cspp@viirs-cspp3:~

File Edit View Search Terminal Help

```
##################################################################################
# Process SDRs
* * * * * /local/home/cspp/bin/GenerateSDRs.sh 1 > /dev/null 2>&1
* * * * * /local/home/cspp/bin/GenerateSDRs.sh 2 > /dev/null 2>&1
* * * * * /local/home/cspp/bin/GenerateSDRs.sh 3 > /dev/null 2>&1
* * * * * /local/home/cspp/bin/GenerateSDRs.sh 4 > /dev/null 2>&1
* * * * * /local/home/cspp/bin/GenerateSDRs.sh 5 > /dev/null 2>&1
* * * * * /local/home/cspp/bin/GenerateSDRs.sh 6 > /dev/null 2>&1
# Process EDRs (GTM)
* * * * * /local/home/cspp/bin/GenerateGTMs.sh 123 > /dev/null 2>&1
* * * * * /local/home/cspp/bin/GenerateGTMs.sh 234 > /dev/null 2>&1
* * * * * /local/home/cspp/bin/GenerateGTMs.sh 345 > /dev/null 2>&1
* * * * * /local/home/cspp/bin/GenerateGTMs.sh 456 > /dev/null 2>&1
* * * * * /local/home/cspp/bin/GenerateGTMs.sh 561 > /dev/null 2>&1
* * * * * /local/home/cspp/bin/GenerateGTMs.sh 612 > /dev/null 2>&1
##################################################################################
```

# The Scheme

## Bash script

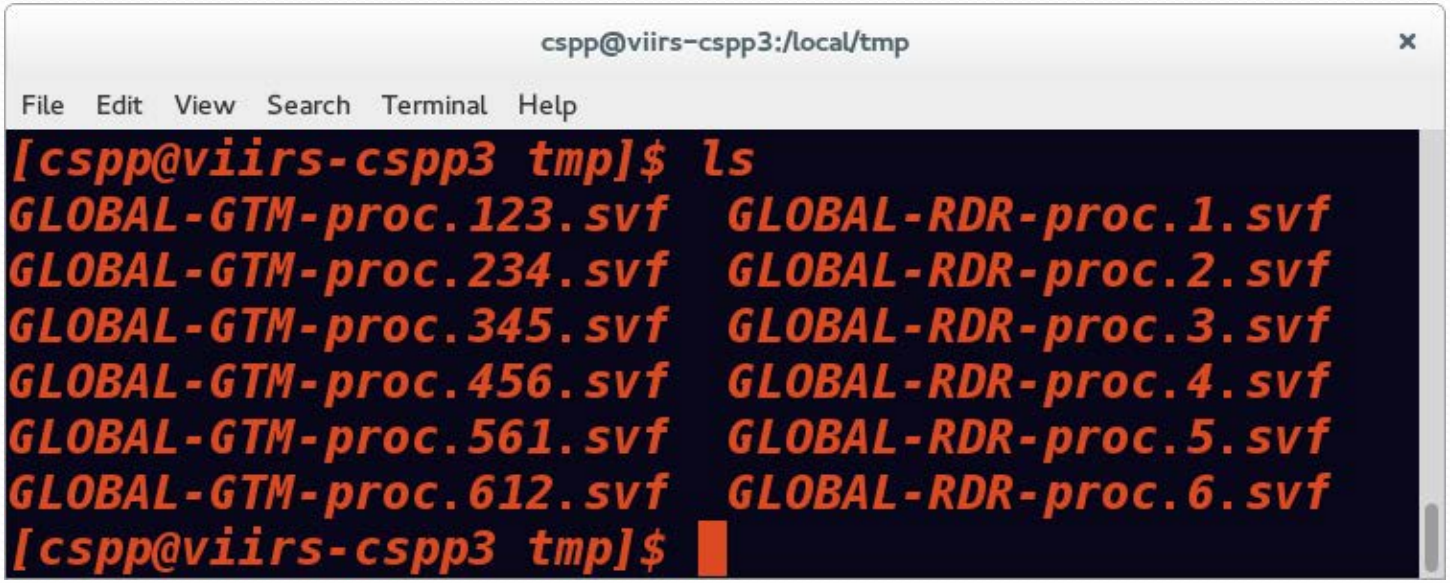
cspp@viirs-cspp3:~/bin

File Edit View Search Terminal Help

```
[cspp@viirs-cspp3 bin]$ cat GenerateSDRs.sh
#!/bin/bash
#
# shell script to try a GenerateSDR.pl run
# at a finer granularity than once per min
#
n=$1
k=4
while [ $k -gt 0 ]
    do
    /local/home/cspp/bin/GenerateSDRs.pl -n $n >> /home/cspp/SDR-$n.log 2>&1 &
    sleep 14
    k=`/usr/bin/expr $k - 1`
done
[cspp@viirs-cspp3 bin]$
```

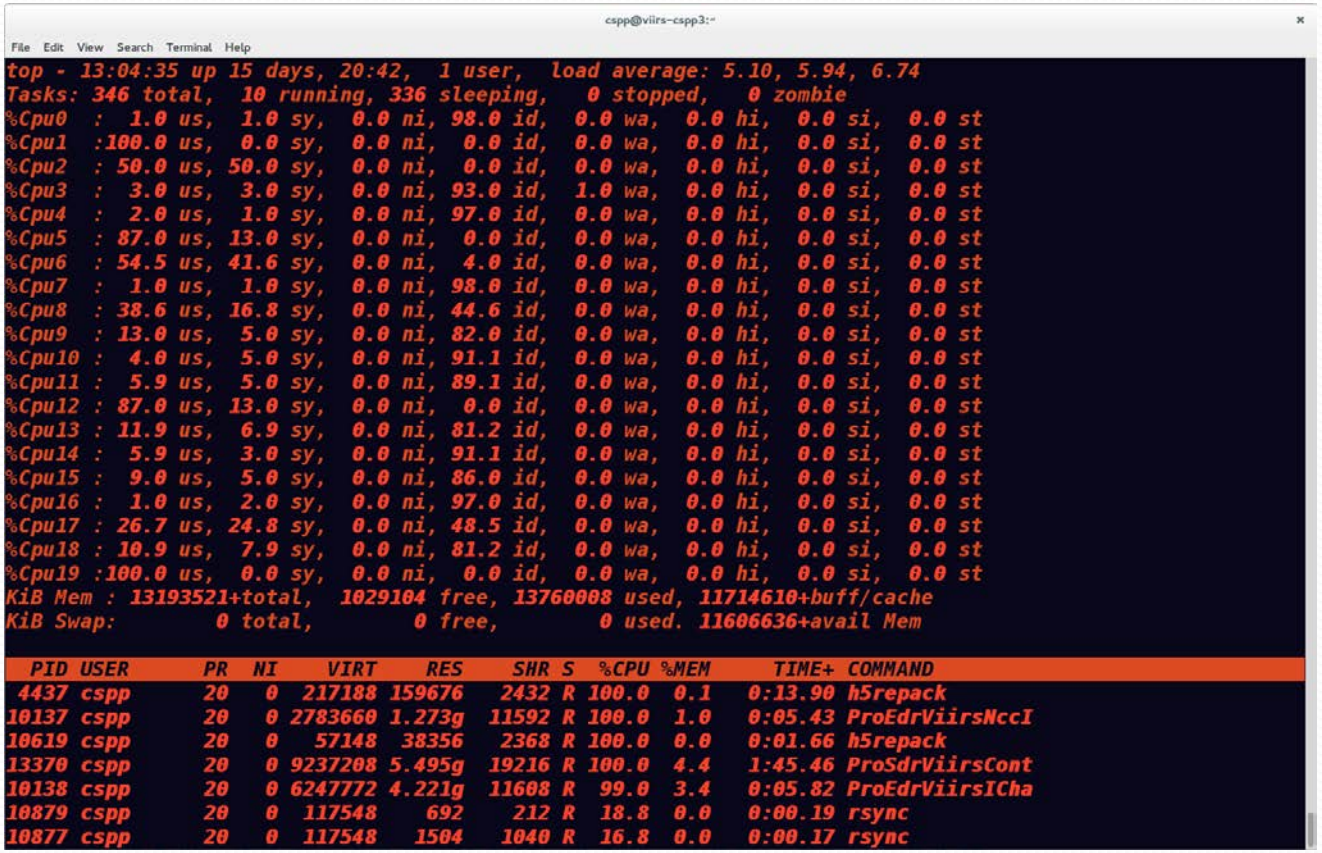
# The Scheme

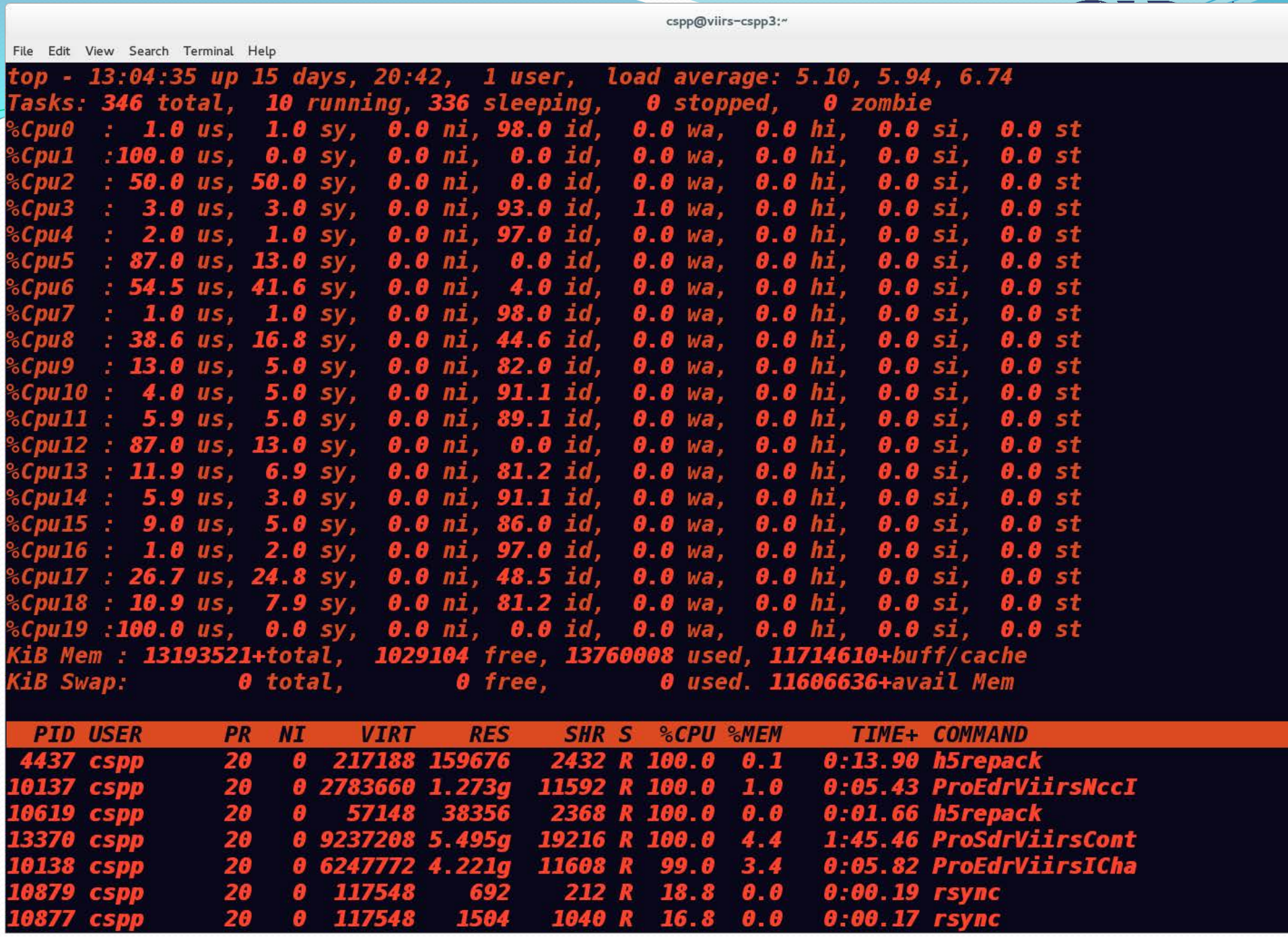
Lock files



# The Scheme

Top




cspp@viirs-cspp3:~
File Edit View Search Terminal Help
top - 13:04:35 up 15 days, 20:42, 1 user, load average: 5.10, 5.94, 6.74
Tasks: 346 total, 10 running, 336 sleeping, 0 stopped, 0 zombie
%Cpu0 : 1.0 us, 1.0 sy, 0.0 ni, 98.0 id, 0.0 wa, 0.0 hi, 0.0 si, 0.0 st
%Cpu1 :100.0 us, 0.0 sy, 0.0 ni, 0.0 id, 0.0 wa, 0.0 hi, 0.0 si, 0.0 st
%Cpu2 : 50.0 us, 50.0 sy, 0.0 ni, 0.0 id, 0.0 wa, 0.0 hi, 0.0 si, 0.0 st
%Cpu3 : 3.0 us, 3.0 sy, 0.0 ni, 93.0 id, 1.0 wa, 0.0 hi, 0.0 si, 0.0 st
%Cpu4 : 2.0 us, 1.0 sy, 0.0 ni, 97.0 id, 0.0 wa, 0.0 hi, 0.0 si, 0.0 st
%Cpu5 : 87.0 us, 13.0 sy, 0.0 ni, 0.0 id, 0.0 wa, 0.0 hi, 0.0 si, 0.0 st
%Cpu6 : 54.5 us, 41.6 sy, 0.0 ni, 4.0 id, 0.0 wa, 0.0 hi, 0.0 si, 0.0 st
%Cpu7 : 1.0 us, 1.0 sy, 0.0 ni, 98.0 id, 0.0 wa, 0.0 hi, 0.0 si, 0.0 st
%Cpu8 : 38.6 us, 16.8 sy, 0.0 ni, 44.6 id, 0.0 wa, 0.0 hi, 0.0 si, 0.0 st
%Cpu9 : 13.0 us, 5.0 sy, 0.0 ni, 82.0 id, 0.0 wa, 0.0 hi, 0.0 si, 0.0 st
%Cpu10 : 4.0 us, 5.0 sy, 0.0 ni, 91.1 id, 0.0 wa, 0.0 hi, 0.0 si, 0.0 st
%Cpu11 : 5.9 us, 5.0 sy, 0.0 ni, 89.1 id, 0.0 wa, 0.0 hi, 0.0 si, 0.0 st
%Cpu12 : 87.0 us, 13.0 sy, 0.0 ni, 0.0 id, 0.0 wa, 0.0 hi, 0.0 si, 0.0 st
%Cpu13 : 11.9 us, 6.9 sy, 0.0 ni, 81.2 id, 0.0 wa, 0.0 hi, 0.0 si, 0.0 st
%Cpu14 : 5.9 us, 3.0 sy, 0.0 ni, 91.1 id, 0.0 wa, 0.0 hi, 0.0 si, 0.0 st
%Cpu15 : 9.0 us, 5.0 sy, 0.0 ni, 86.0 id, 0.0 wa, 0.0 hi, 0.0 si, 0.0 st
%Cpu16 : 1.0 us, 2.0 sy, 0.0 ni, 97.0 id, 0.0 wa, 0.0 hi, 0.0 si, 0.0 st
%Cpu17 : 26.7 us, 24.8 sy, 0.0 ni, 48.5 id, 0.0 wa, 0.0 hi, 0.0 si, 0.0 st
%Cpu18 : 10.9 us, 7.9 sy, 0.0 ni, 81.2 id, 0.0 wa, 0.0 hi, 0.0 si, 0.0 st
%Cpu19 :100.0 us, 0.0 sy, 0.0 ni, 0.0 id, 0.0 wa, 0.0 hi, 0.0 si, 0.0 st
KiB Mem : 13193521+total, 1029104 free, 13760008 used, 11714610+buff/cache
KiB Swap: 0 total, 0 free, 0 used. 11606636+avail Mem
PID USER PR NI VIRT RES SHR S %CPU %MEM TIME+ COMMAND
4437 cspp 20 0 217188 159676 2432 R 100.0 0.1 0:13.90 h5repack
10137 cspp 20 0 2783660 1.273g 11592 R 100.0 1.0 0:05.43 ProEdrViirsNccI
10619 cspp 20 0 57148 38356 2368 R 100.0 0.0 0:01.66 h5repack
13370 cspp 20 0 9237208 5.495g 19216 R 100.0 4.4 1:45.46 ProSdrViirsCont
10138 cspp 20 0 6247772 4.221g 11608 R 99.0 3.4 0:05.82 ProEdrViirsICha
10879 cspp 20 0 117548 692 212 R 18.8 0.0 0:00.19 rsync
10877 cspp 20 0 117548 1504 1040 R 16.8 0.0 0:00.17 rsync

# The Metrics

- RDR filesize ~20MiB - ~100MiB  ~56GiB/day
- 20X expansion to SDRs/EDRs  ~1.2TiB/day
- ~118 sec to generate one set of SDRs/EDRs
- ~43min to process imagery SDRs/EDRs from a pass
- ~77GiB per orbit

# The Metrics

- RDR filesize ~20MiB - ~100MiB ~56GiB/day
- 20X expansion to SDRs/EDRs ~1.2TiB/day
- ~118 sec to generate one set of SDRs/EDRs
- ~43min to process imagery SDRs/EDRs from a pass
- ~77GiB per orbit

# The Metrics

- RDR filesize ~20MiB - ~100MiB ~56GiB/day
- 20X expansion to SDRs/EDRs ~1.2TiB/day
- ~118 sec to generate one set of SDRs/EDRs
- ~43min to process imagery SDRs/EDRs from a pass
- ~77GiB per orbit

# The Metrics

- RDR filesize ~20MiB - ~100MiB  ~56GiB/day
- 20X expansion to SDRs/EDRs  ~1.2TiB/day
- ~118 sec to generate one set of SDRs/EDRs
- ~43min to process imagery SDRs/EDRs from a pass
- ~77GiB per orbit

# The Metrics

- RDR filesize ~20MiB - ~100MiB ~56GiB/day
- 20X expansion to SDRs/EDRs ~1.2TiB/day
- ~118 sec to generate one set of SDRs/EDRs
- ~43min to process imagery SDRs/EDRs from a pass
- ~77GiB per orbit

# Issues

- Bad RDR files
- Late or missing RDR files
- "Duplicate" RDR file
- Power outages

# Questons

The End

# CIRA Websites

- rammb.cira.colostate.edu
  - RAMMSDIS Online
  - TC Real-Time
- rammb-slider.cira.colostate.edu
  - GOES-16
  - Himawari-8
- www.cloudsat.cira.colostate.edu

# Ceph Storage

- Inexpensive (or expensive)
- Redundant/reliable (failure domains)
- Scalable
- Cephfs is very fast