June 24, 2015

Guide for Trimming AWIPS-2 Cases for WES-2 Bridge

Introduction

Archived case sizes in AWIPS-2 are significantly larger than in AWIPS-1, for several reasons. First, even with internal compression enabled, the HDF file storage in AWIPS-2 consumes more disk space for many data types than the corresponding netCDF in AWIPS-1. We also have to store the archived metadata database records for each corresponding HDF file; this requirement did not exist in AWIPS-1. Accordingly, besides disk space, the number of database records is also a key statistic in determining the "size" of an archived case and affects the efficiency of some of the WES-2 Bridge data manipulation routines. Finally, higher-resolution datasets (for example, the HRRR model) exist in the AWIPS-2 era than did when many of the AWIPS-1 and WES-1 archiving routines were established; these also consume more disk space.

In addition to actual disk space consumption being a concern for archived cases, we also have a requirement to preserve both raw and processed versions of archived data. This requirement comes about because software changes in AWIPS-2 can make some of the processed versions of data unusable; consequently, the raw data needs to be preserved to enable re-creation of the processed data.

This guide assumes case data have already been archived and reside in separate Raw and Processed directories inside a case.

Strategies

… archive compreshenively then trim …. Rather than missing some critical data. Important for completeness, operational represenattiveness. and for situation awareness exercises. The exact data used in a simulation should be guided by the objectives of a simulation. For scientific case review, having more data to compare/contrast is helpful.

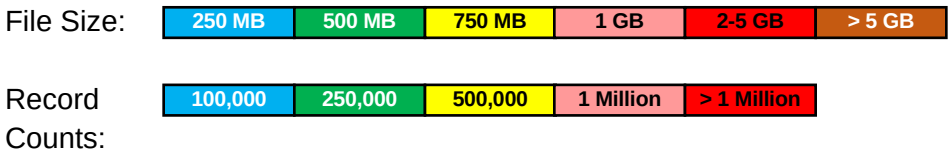
1. **Removing entire datatypes.**

Data in AWIPS-2 are organized by individual datatypes, for both raw and processed data. Some data types are more significant in how much disk space they typically consume. Another factor to consider in choosing data types to delete is the number of database records for that datatype. More database records can require a greater amount of time to load and unload cases using WES-2 Bridge. Finally, care must be exercised in deleting some types of raw data because there is not always a one to one relationship between raw and processed data owing to the fact that some processed data is constructed

using multiple raw data types and also that some types of raw data are split up into multiple processed datatypes.

Not all datasets are created equal. Some datasets consume orders of magnitudes of more space than others; hence, there is more “bang for the buck” in terms of both disk space as well as human and computer processing time in concentrating on datatypes that use the largest amounts of space. To aid in the determining which datatypes to trim, the following table gives daily estimates of disk space (both raw and processed) and the number of database records for each datatype for a whole day. Some of these estimates are dependent upon the weather; others are affected by local WFO AWIPS configurations. These are noted in the table.

The left set of cells of the tables show the file sizes of raw data, the middle set is for the file sizes of processed data, and the right triplet of cells is for a count of database records for processed data. For the file-size groups of cells, the left cell is for kilobytes (KB), the middle cell for megabytes (MB) and the right cell for (GB). This layout provides the ability to see the orders of magnitudes of file sizes during a quick glance. For the database record counts, a similar organization is used: left is for record counts less than 1,000; the middle for counts greater than 1,000 and less than 1,000,000 and the right cell for counts greater than 1 million. However, the casual user may have difficulty in discerning the significance of these numbers. Therefore, some table cells are color-filled with a simple scheme. No color-fill means the size of the data is not significant at all. The following chart shows the color scheme and the relative amount of significance:



The significance comes from the fact that file sizes become meaningful in the range of gigabytes, so this table attempts to call one’s attention to file types that begin to contribute to gigabytes beginning at 250 MB. Similarly, the database tables begin to take time to process when they have numbers of records in hundreds of thousands to millions. This information is presented to you in order to save your time. **There is no need to trim datasets when they are not significant in terms of size or database records; concentrate your efforts on datasets that do consume space.**

In the table, references are made to Raw data and Processed data. On the real-time AWIPS, raw data is located in /data_store and is situated in a Raw directory in an archived case with the same file structure as /data_store. Similarly, processed data on the real-time AWIPS is stored in /awips2/edex/data/hdf5 (with 7 days of storage in /archive) and is located in a Processed directory in a case.

The single largest datatype is gridded data, and the directory names can be confusing. Gridded raw data is in a grib or grib2 directory when ingested from the SBN, and is in a manual/grib directory when the raw data arrives via an LDAD system. In any case, all processed gridded data is in a grid directory. In the gridded data section, there are a number of notes that refer to WMO headers which can be used to refer to some of the gridded data by geographic sector. An explanation of how to do this follows the table. These WMO headers are documented in several external references by either NCEP, MDL, or various NWS Technical Information Notices. Because of the small font size used in the notes, these references are notated using circled numbers (e.g., ①) rather than superscripts and the reference URL is given below the table.

A note about warnings. The processed warning data type is the actual set of warnings that were issued. WES-2 Bridge has a way to filter warnings by CWA during simulation. Keeping this data type with a case is important for at least two reasons: (1) WFO staff can easily review their own performance after an event, and (2) eventually WES-2 Bridge will have some objective warning performance metrics that will use the actual office warnings as a comparison point.

DataType	Raw Data Size (/data_store)	Processed Data Size (/archive/*.h5 + /archive/*bin)	Processed Record Counts (/archive/*.bin)
acars	100 MB	500-750 MB	450,000 – 600,000
acarssounding		150-200 MB	5,000-9,000
Note: acarssounding data are processed data only and are derived from acars.			
airep	35 MB	50 MB	35,000
Note: Processed versions of airep text bulletins are in the text (fxatext) database and are included in processed data (.bin files) for text below.			
airmet	400 KB	400-800 KB	100
Note: The processed versions of airmet text bulletins are also in the text (fxatext) database and are included in processed data (.bin files) for text below.			
binlightning	6 MB	100 MB	1,400
Note: These results are for binlighting (from NLDN alone) without the addition of entlightning (total lightning)			
binlightning + entlightning		250 MB	2,700
Note: Raw data for entlightning is stored in an entlightning directory and presented separately below in the entlighting row. When entlighting data are processed by EDEX, both NLDN and ENTLN are stored in the binlightning processed data files (.bin and .h5)			
bufrascat	30 MB	1.8 GB	1,800,000

bufhdw	75 MB	1.5 GB	1,200,000
bufrmos	100 MB		

Note: Processed data for bufrmos are stored individually by the type of bufrmos: AVN, ETA, GFS, HPC, LAMP, and MRF and are shown individually below.

bufrmosAVN		280 MB	350,000
bufrmosETA		90 MB	105,000
bufrmosGFS		200 MB	250,000
bufrmosHPC		3.5 MB	4,300
bufrmosLAMP		800 MB	1,000,000
bufrmosMRF		40 MB	47,000
bufrmthdw	7 MB	180 MB	150,000
bufrncwf	2 MB	7 MB	4,000
bufrsigwx	400 KB	3 MB	600
bufrssmi	10 MB	40 MB	40,000
bufrua	3.5 MB	4 MB	2,500
ccfp		4 MB	2,500

Note: raw data for ccfp is stored in /data_store/forecasts with WMO headers FAUS28 KKCI, FAUS29 KKCI, and FAUS30 KKCI. Text bulletins are also in the text (fxatext) database and are included in processed data (.bin files) for text below.

climate	30 MB		

Note: The processed data for climate text bulletins are in the text (fxatext) database and are included in processed data (.bin files) for text below.

convsigmet	400 KB	700 KB	75

Note: The processed versions of convsigmet text bulletins are also in the text (fxatext) database and are included in processed data (.bin files) for text below.

cwa 1.5 MB 50

Note: The raw data for cwa (Center Weather Advisories, issued by CWSUs), are stored in forecasts, with WMO headers `FAUSii CCCC` where ii=20 through 26 or `FAAK20` through `FAAK29`, where CCCC = station ID of the CWSU, such as KZTL = Atlanta CWSU. The corresponding text bulletins are also in the text (fxatext) database, and are included in the processed data (.bin files) for text below.

cwat 25 MB 1,000

Note: CWAT is a decision assistance tool (DAT) which is part of the SCAN system. It is produced from a number of raw data inputs, including radar data, lightning, and RUC130 fields.

entlightning 200 MB

Note: The processed data for total lightning is included with binlightning data above.

ffg 3.5 MB 400

Note: The raw data for ffg is stored in /data_store/forecasts with WMO headers `FOUS61` through `FOUS64`. These are text-based versions of flash flood guidance (county-based). Text bulletins are also in the text (fxatext) database and are included in the processed data (.bin files) for text below.

ffmp 500 MB 7,500

Note: FFMP is a decision assistance tool (DAT). The size of its data is highly dependent on the local office configuration, especially the number of local radars and the geographic definitions of stream basins. It uses local radar data, high-resolution precipitation estimates (see grid/BHPE and grid/HPE), and gridded flash flood guidance (see grib/NWS_151 for more information on gridded FFG).

fire_wx_spot_fcst_reports 600 MB

Note: The processed data for fire_wx_spot_fsct_reports is stored in the text (fxatext) database and are included in the processed data (.bin files) for text below.

forecast 450 MB

Note: The processed data for forecast is stored in the text (fxatext) database. Particular text bulletins are also processed by other plugins, including ccfp, cwa, idft, and vaa, and processed data for those plugins are listed separately.

fssobs 10 MB – 200 MB 1,000 – 95,000

Note: FSSOBS is a decision assistance tool (DAT). The size of its data is highly dependent on the local office configuration, especially the inclusion of local stations that are ingested via LDAD. This tool uses metar, mesonet, maritime, buoys, and MAROB data as input.

gfe 2.5 GB 85,000

goessounding 2.5 GB 200-500 MB 175,000 - 400,000

Note: At the time of this compilation (Summer 2015), most WFOs were storing raw goessounding data, but not processing it or archiving processed goessounding data due

to local office configurations not having an appropriate goesBufr.spi file in /awips2/edex/data/utility/common_static/site/<SITE>/basemaps. This information comes from WDTD's system where the size of the processed data was highly dependent on this .spi file as well as the weather conditions.

The following section of the table consolidates and lists gridded datasets from /data_store/grib/, /data_store/grib2/, /archive/grid/, and /data_store/manual/grib/. Raw data is labeled as "grib", "grib2", or "manual/grib". Processed is labeled as "grid". The individual datatypes resume below the gridded section with the "idft" datatype.

---

grib2/**118**

	2-8 MB	

Note: This data is related to the implementation of the Unrestricted Mesoscale Analysis (URMA). Most sites store their raw URMA data in grib2/<date><hour>/URMA, except for a precipitation analysis which is stored here in the 118 folder under grib2. The corresponding processed data is in grid/URMA25 with a forecast hour of F006 (e.g., grid/URMA25/SFC/URMA25-2015-06-10-00-FH-006).

grib/**3hr**
grid/**PROB3HR**

	2.5 MB			**6 MB**			**300**	

grid/**AK-NamDNG5**

		1.3 GB		**3,000**	

Note: The raw data for AK-NamDNG5 is stored in grib2/date>/<hour>/NMM_89/GRID255 and is included below as grib2/NMM_89. Within the NMM_89/GRID255 directory, the WMO headers for the raw data are like `[LM].A... KWBE` ①

grid/**AK-RTMA**

	335 MB		**300**		

Note: The raw data for AK-RTMA is stored in grib2/<date>/<hour>/RTMA/GRID255 and is included below as grib2/RTMA. Within the RTMA/GRID255 directory, AK-RTMA has WMO headers like: `L.AA98 KWBR` ②

grid/**AK-RTMA3**

		1.1 GB	**400**		

Note: The raw data for AK-RTMA3 is stored in grib2/<date>/<hour>/RTMA/GRID255, and is included below as grib2/RTMA. Within the RTMA/GRID255 directory, AK-RTMA3 has WMO headers like: `L.KA98 KWBR` ③

grid/**AKwave4**

	350 MB			**3,200**	

Note: The raw data for AKwave4 is stored in grib2/<date>/<hour>/GMGWM/GRID255, and is included below as grib2/GMGWM. Within the GMGWM/GRID255 directory, AKwave4 has WMO headers like: `E.F.88_KWBJ`. ④ ⑤

grid/**AKwave10**

	215 MB			**3,200**	

Note: The raw data for AKwave10 is stored in grib2/<date>/<hour>/GMGWM/GRID255, and is included below as grib2/GMGWM. Within the GMGWM/GRID255 directory, AK-wave10 has WMO headers like: `E.E.88_KWBJ`. ④ ⑤

grid/**AKWAVE239** – see grib/**NWW_122** below.

grid/**AUTOSPE** – see grib/**NWS_190** below.

grid/**Aviation** 2.5 GB 6,100

Note: The raw data for Aviation is stored in grib/RUC2/<date>/<hour>/GRID130. It comes from the Aviation Weather Center and has a CCCC in the WMO header of KKCI. In addition, the Aviation Weather Center produces a Current Icing Potential (CIP) product in grib format, and is identified as AWC_CIP in raw data with its processed data being stored in grid/Aviation. As of the time of this compilation (Summer 2015), the baseline pqact.conf file on WFO systems was not configured to correctly receive this product; however, earlier versions of AWIPS-2 did.

grid/**AVN211** 850 MB 32,000

Note: The raw data for AVN211 is stored in grib/<date>/<hour>/GFS/GRID211 (forecasts) and grib/<date>/<hour>/SSIGFS/GRID211 (initial hour analysis). This is an 80-km CONUS grid and is sometimes known as GFS80.

grid/**AVN225** 770 MB 10,000

Note: The raw data for AVN225 is stored in grib/</date>/<hour>/GFS/GRID225 (forecasts) and grib/<date>/<hour>/SSIGFS/GRID225 (initial hour analysis). This is a Hawaii/Pacific Grid.

grib/**AWC_NCWD** 65 MB 50 MB 300
grid/**NCWF**

grid/**BHPE** 5 – 50 MB 225

Note: The raw data for BHPE (Bias HPE) are not currently archived (grib files are stored in /awips2/edex/data/processing and XMRG files are in /awips2/edex/data/share/hydroapps). A future hydro archiver configuration will include the raw data. The file sizes are highly dependent upon local configuration (e.g., number of radars and the size of domain)

manual/***CMC2_nh*** 114 MB 340 MB 2,000
grid/**Canadian-NH**

manual/***CMC2_reg*** 125 MB 460 MB 1,300
grid/**Canadian-Reg**

grib2/**DGEX_115** 315 MB

Note: Both DGEX185 and DGEX186 are the processed versions of the DGEX_185 raw data (stored in GRID185 and GRID186 directories, respectively)

grid/**DGEX185** 405 MB 2,100

grid/**DGEX186** 235 MB 2,000

Note: DGEX186 is produced for an Alaska grid.

grib2/**ECMWF_144** | | 4.2 MB | |

Note: The processed versions of ECMWF_144 are stored as ECMF1 through ECMF8 and as the ECMF-NorthernHemisphere and ECMF-Tropical composites.

grid/**ECMF1** through grid/**ECMF4** | | 3.5 MB | | | 285 | |

Note: The results shown here are for each grid individually, so each grid separately consumes 3.5 MB of space for a total of 14 MB per day for all four grids. ECMF1 and ECMF2 cover the CONUS area. ECMF1 through ECMF4 also contribute to the ECMF-NorthernHemisphere composite.

grid/**ECMF5** through grid/**ECMF8** | | 1 MB | | | 90 | |

Note: The results shown here are for each grid individually, so each grid separately consumes 1 MB of space for a total of 4 MB per day for all four grids. ECMF5 and ECMF6 cover the CONUS area south of 35°N. ECMF5 through ECMF8 also contribute to the ECMF-Tropical composite.

grid/**ECMF-Northern Hemisphere** | | 13 MB | | | 285 | |

grid/**ECMF-Tropical** | | 2 MB | | | 40 | |

grib2/**ECMWF_HiRes** grid/**ECMWF_HiRes** | | 64 MB | | | 190 MB | | | 4,400 |

grid/**ENPWAVE253** – see grib/**NWW_124** below.

grid/**ENSEMBLE** | | | 1.8 GB | | 21,000 |

Note: The raw data for ENSEMBLE is stored in grib2/GFS and grib2/SSIGFS (in GRID038, GRID039, and GRID040 directories). ENSEMBLE is a composite of ENSEMBLE38, ENSEMBLE39 and ENSEMBLE40.

grid/**ENSEMBLE38** through **ENSEMBLE40** | | 450 MB | | | 21,000 |

Note: The sizes given here are for each grid individually, so all three grids total **450 MB * 3 = 1,350 MB (about 1.2 GB)** and a total of 63,000 database records. The raw data for ENSEMBLE38 through ENSEMBLE40 are stored in grib2/<date>/<hour>/GFS and grib2/<date>/<hour>/SSIGFS in GRID038 through GRID040 directories, respectively.

grid/**EPwave10** | | 125 MB | | | 3,300 |

Note: The raw data for EPwave10 is stored in grib2/<date>/<hour>/GMGWM/GRID255, and is included below as grib2/GMGM. Within the GMGWM/GRID255 directory, EPwave10 has WMO headers like `E.D.88 KWBJ.` ④ ⑤

grib2/**ESTOFS** | | | 2.1 GB |

Note: This data is from the **E**xtra-tropical **S**urge and **T**ide **O**perational **F**orecast **S**ystem. Both estofsPR and estofsUS are the processed versions of the ESTOFS raw data. They are stored in /grib2/<date>/<hour>/ESTOFS/GRID255 with WMO headers like `E[ECH]I.88_KWBM` and `E[ECH]P.88_KWBM` for estofsUS (US East Coast) and estofsPR (Puerto Rico), respectively. ⑥

grid/**estofsPR** | | 425 MB | | | 2,200 |

grid/**estofsUS** | | | 4 GB | | 2,200 |

grid/**ETA** | | 80 MB | | | 3,700 |

Note: Raw data for ETA is stored in grib</date>/<hour>/NAM_84/GRID211. This is an 80-km CONUS grid, and is sometimes known as NAM80.

grid/**ETA** (supplemental) | | 4 MB | | | 225 MB | | | 11,000 |

manual/grib/*__gribF##*__

Note: Raw data for ETA which is delivered via the SBN is stored in /data_store/grib/NAM_84/GRID211. Many WFOs also ingest supplemental ETA data for forecast hours beyond 60 through their LDAD/LDM. This extra data is stored in manual/grib with filenames that typically include a “gribF##” pattern where ## is the forecast hour. The file size and record count shown here include both the baseline and supplemental data for a single day.

grid/**ETA207** | | 50 MB | | | 3,500 |

Note: Raw data for ETA207 is stored in grib/<date>/<hour>/NAM_84/GRID207. This is a 95-km grid for Alaska.

grid/**ETA212** – see grib/**NMM_89** below.

grid/**ETA218** | | | 1.5 GB | | 30,000 |

Note: Raw data for ETA218 (also known as NAM12 for CONUS) is stored in grib2/<date>/<hour>/NAM_84/GRID218. This model is clipped or subgridded, so the actual size can vary as the size of the domain changes.

grid/**ETA242** | | | 8 GB | | 30,000 |

Note: Raw data for ETA242 (also known as AK-NAM12) is stored in grib2/<date>/<hour>/NAM_84/GRID242.

grid/**FFG-*** | | 250 KB – 4 MB | | | 10 - 80 | |

Note: This is gridded Flash Flood Guidance (FFG) from each RFC (sites can be ALR, FWR, KRF, MSR, ORN, PTR, RHA, RSA, STR, TAR, TIR, and TUA). The sizes presented here are an average of daily estimates for individual sites. A very liberal estimate for all RFCs could total about 48 MB (if all RFCs had the maximum size), but the actual total is typically much smaller. The raw data for gridded FFG is stored in grib/NWS_151.

grib2/**FORECASTER** 20 - 45 MB

Note: The processed version of FORECASTER data are stored as grid/TPCWindProb (shown below).

grib2/**GEFS** 120 MB 375 MB 2,500

grid/**gefs**

grib/**GFS** 810 MB

Note: The processed versions of the grib versions of GFS are stored in AVN211 (GFS/GRID211), AVN225 (GFS/GRID225), ENSEMBLE38 (GFS/GRID038), ENSEMBLE39 (GFS/GRID039), and ENSEMBLE40 (GFS/GRID040), all under grid.

grib2/**GFS** 8.4 GB

Note: The processed versions of the grib2 versions of GFS are stored in GFS160 (GFS/GRID160), GFS161 (GFS/GRID161), GFS201 (GFS/GRID201), GFS212 (GFS/GRID212), GFS213 (GFS/GRID213), GFS254 (GFS/GRID254). In addition, MOSGuide for CONUS is stored under grib2/<date>/<hour>/GFS/GRID197 and GFSGuide, MOSGuide-AK and GFS199 are all stored under grib2/<date>/<hour>/GFS/GRID255.

grid/**GFS160** 3.5 GB 42,000

Note: GFS160 is produced for a 47.5 km Alaska grid. Its raw data is stored in grib2/<date>/<hour>/GFS/GRID160 and grib2/<date>/<hour>/SSIGFS/GRID160, which are included in grib2/GFS above and grib2/SSIGFS below.

grid/**GFS161** 2 GB 42,000

Note: GFS161 is produced on a lat/lon grid for Puerto Rico, from raw data in grib2/<date>/<hour>/GFS/GRID161 and grib2/SSIGFS/GRID161, and included in grib2/GFS above and grib2/SSIGFS below.

grid/**GFS201** 280 MB 12,000

Note: GFS161 is produced on a 381-km Northern Hemisphere grid from raw data in grib2/<date>/<hour>/GFS/GRID201 and SSIGFS/<date>/<hour>/GRID201, and is included in grib2/GFS above and grib2/SSIGFS below.

grid/**GFS212** 3 GB 42,000

Note: GFS212 (the highest-resolution version) is produced for a 40-km CONUS grid from raw data in grib2/<date>/<hour>/GFS/GRID212 and grib2/<date>/<hour>/SSIGFS/GRID212 and is included in grib2/GFS above and grib2/SSIGFS below.

grid/**GFS213** 550 MB 14,000

Note: GFS213 is produced on a 90-km CONUS Polar-Stereographic grid from raw data in grib2/<date>/<hour>/GFS/GRID213 and SSIGFS/<date>/<hour>/GRID213 and is included in grib2/GFS above and grib2/SSIGFS below.

grid/**GFS254** 12 GB 42,000

Note: GFS254 is produced on a 40-km Pacific Region Mercator grid from raw data in grib2/<date>/<hour>/GFS/GRID254 and SSIGFS/<date>/<hour>/GRID254 and is included in grib2/GFS above and grib2/SSIGFS below.

grid/**GFSGuide** | 10 MB | 600

Note: GFSGuide is produced for Guam from raw data in grib2/<date>/<hour>/GFS/GRID255 with WMO headers like `[LM].G... KWBT`. ⑦

grid/**GFSLAMPTstorm** – see grib2/**LAMP** below.

grib2/**GLAMP** or **GLAMP25**
grid/**GLAMP25** | 1.4 GB | 1.5 – 4 GB | 4,400

Note: This model is clipped or subgridded, so the actual size can vary as the size of the domain changes.

Note: Many WFOs use an incorrect WMO header in their local section of pqact.conf for LDM. The correct regular expression for the WMO headers for GLAMP25 should be `L.U.[0-9][0-9] KMDL`, but many sites actually use a pattern of `L...[0-9][0-9] KMDL`, which also stores EKDMOS generated from the NAEFS and stores a total of 5 GB/day, rather than an average of 1.4 GB for GLAMP25 only. If the files in edex_static/base/grib/models/ (or any corresponding site override files) are not configured to recognize EKDMOS, the EKDMOS data is stored as GribModel:7:14:114. ⑧

grib2/**GLWM**
grid/**GRLKWave** | 225 MB | 500 MB | 10,000

grib2/**GLWM25**
grid/**WW2-2km** | 550 MB | 1.1 GB | 10,000

grid/**GlobalWave** | 1.2 GB | 4,400

Note: The raw data for GlobalWave is stored in grib2/<date>/<hour>/GMGWM/GRID255, and is included below as grib2/GMGM. Within the GMGWM/GRID255 directory, EPwave10 has WMO headers like `E.A.88 KWBJ`. ④ ⑤

grib2/**GMGWM** | 1.3 GB

Note: The processed data for GMGWM is stored in grid/AKwave4, grid/AKwave10, grid/EPwave10, grid/GlobalWave, grid/WCWave4, grid/WCwave10, grid/WNAwave4, and grid/WNAwave10. The raw data is stored in grib2/<date>/<hour>/GRID255 and with these WMO headers for their respective domains: `E.A.88 KWBJ` (Globalwave), `E.B.88 KWBJ` (WNAwave10), `E.C.88 KWBJ` (WCwave10), `E.D.88 KWBJ` (WCwave10), `E.E.88 KWBJ` (AKwave10), `E.F.88 KWBJ` (AKwave4), `E.G.88 KWBJ` (WCwave4) and `E.H.88 KWBJ` (WNAwave4). ④ ⑤

grib2/**GMOS** or **GMOS25**
grid/**gmos25** | 850 MB | 500 MB – 1.8 GB | 2,000

Note: This model is clipped or subgridded, so the actual size can vary as the size of the domain changes.

grid/**GribModel:59:0:83** | | | 6 GB | | 7,800 |

Note: This is the processed version of the GSD version of HRRR (also sometimes known as HRRR.CR). GSD changed the Process ID of this model from 125 to 83 (the last number in the GribModel:*:*:*) in May 2015. The raw data for this version is in manual/grib/*HRRR*.

grid/**GribModel:7:14:114** | | | 2.3 GB | | 1,190 |

Note: This is EKDMOS (Ensemble Kernel Density MOS) from NAEFS (North American Ensemble Forecast System) from MDL. The WMO headers for EKDMOS for CONUS are L.[A-L][A-R].. KMDL. Some sites, intending to ingest GLAMP, used a WMO header pattern of `L...[0-9][0-9] KMDL` rather than `L.U.[0-9][0-9] KMDL`, and thus ingest both GLAMP and EKDMOS. If edex_static/…/grib/models/ have a configuration for Center 7, Subcenter 14, and Process ID 114, the system will not store the data in grid/GribModel:7:14:114, but whatever name is in the configuration file. Some sites have configured this data to store as grid/EKDMOS. EKDMOS for domains outside of CONUS have WMO headers that start with M, Y, and Z, respectively, for Alaska, Puerto Rico, and Hawaii. ⑧

grid/**GribModel:9:105:160** | | 100 KB | | 65 | |

Note: This is MPE-Local from SJU. Raw data is in grib/NWS_160.

grid/**GribModel:9:105:161** | | 100 KB | | 72 | |

Note: This is MPE-Mosaic from SJU. Raw data is in grib/NWS_161.

grid/**GribModel:9:105:171** | | 100 KB | | 36 | |

Note: This is QPE-XNAV from SJU. Raw data is from grib/NWS_171.

grib2/**GRID001** | | 2 MB | | | 24 MB | | 50 | |

grid/**SPCGuide**

grid/**Guam-RTMA** | | 7 MB | | 160 | |

Note: The raw data for Guam-RTMA is stored in grib2/<date>/<hour>/RTMA/GRID255 and is included below as grib2/RTMA. Within the RTMA/GRID255 directory, Guam-RTMA has WMO headers like `L.GA98 KWBR.` ②

grid/**GWW233** – see grib/**NOW** below.

grib2/**HFR** | | 45 MB | |

Note: The processed data for HFR (High-Frequency Radar) is stored by geographical domains (e.g., grid/HFR-US-6KM). The raw data are stored together in one HFR directory per hour.

grid/**HFR-EAST_6KM** | | 3 MB | | 48 | |

grid/**HFR-EAST_PR_6KM** | | 1 MB | | 48 | |

grid/**HFR-US_EAST_DELAWARE_1KM** | | 3.6 MB | | 48 | |

grid/**HFR-US_EAST_FLORIDA_2KM**		**2.4 MB**		**48**		
grid/**HFR-US_EAST_NORTH_2KM**		**2.6 MB**		**48**		
grid/**HFR-US_EAST_SOUTH_2KM**		**3.1 MB**		**48**		
grid/**HFR-US_EAST_VIRGINIA_1KM**		**3.5 MB**		**48**		
grid/**HFR-US_HAWAII_1KM**		**5.3 MB**		**48**		
grid/**HFR-US_HAWAII_2KM**		**2.4 MB**		**48**		
grid/**HFR-US_HAWAII_6KM**		**1.1 MB**		**48**		
grid/**HFR-US_WEST_500M**		**1.1 MB**		**48**		
grid/**HFR-US_WEST_CENCAL_2KM**		**2 MB**		**48**		
grid/**HFR-US_WEST_LOSANGELES_1KM**		**3.5 MB**		**48**		
grid/**HFR-US_WEST_LOSOSOS_1KM**		**2.1 MB**		**48**		
grid/**HFR-US_WEST_NORTH_2KM**		**2.8 MB**		**48**		
grid/**HFR-US_WEST_SANFRAN_2KM**						
grid/**HFR-US_WEST_SOCAL_2KM**		**2.0 MB**		**48**		
grid/**HFR-US_WEST_WASHINGTON_1KM**						
grid/**HFR-WEST_6KM**		**2.5 MB**		**48**		
grid/**HI-NamDNG5**		**220 MB**			**2,800**	

Note: The raw data for HI-NamDNG5 is stored in grib2/<date>/<hour>/NMM_89/GRID255 and is included below as grib2/NMM_89. Within the NMM_89/GRID255 directory, HI-NamDNG5 was WMO headers like `[LM].H... KWBE` ①

grid/**HI-RTMA**

Note: The raw data for HI-RTMA is stored in grib2/<date>/<hour>/RTMA/GRID255 and is included below as grib2/RTMA. Within the RTMA/GRID255 directory, HI-RTMA has WMO headers like `L.HA98 KWBR`. ②

grid/**HiResW-ARW-AK**

Note: The raw data for HiResW-ARW-AK is stored in grib2/<date>/<hour>/WRF_EM/GRID255 and is included below as grib2/WRF_EM. Within the WRF_EM/GRID255 directory, HiResW-ARW-AK has WMO headers like `[LM].C... KWBS.` ⑨

grid/**HiResW-ARW-East**

Note: The raw data for HiResW-ARW-East is stored in grib2/<date>/<hour>/WRF_EM/GRID255 and is included below as grib2/WRF_EM. Within the WRF_EM/GRID255 directory, HiResW-ARW-East has WMO headers like `[LM].A... KWBS.` ⑨ This model is clipped or subgridded, so the actual size can vary as the size of the domain changes.

grid/**HiResW-ARW-GU**

Note: The raw data for HiResW-ARW-GU is stored in grib2/<date>/<hour>/WRF_EM/GRID255 and is included below as grib2/WRF_EM. Within the WRF_EM/GRID255 directory, HiResW-ARW-GU has WMO headers like `[LM].F... KWBS.` ⑨

grid/**HiResW-ARW-HI**

Note: The raw data for HiResW-ARW-HI is stored in grib2/<date>/<hour>/WRF_EM/GRID255 and is included below as grib2/WRF_EM. Within the WRF_EM/GRID255 directory, HiResW-ARW-HI has WMO headers like `[LM].D... KWBS.` ⑨

grid/**HiResW-ARW-SJU**

Note: The raw data for HiResW-ARW-SJU is stored in grib2/<date>/<hour>/WRF_EM/GRID255 and is included below as grib2/WRF_EM. Within the WRF_EM/GRID255 directory, HiResW-ARW-SJU has WMO headers like `[LM].E... KWBS.` ⑨

grid/**HiResW-ARW-West**

Note: The raw data for HiResW-ARW-West is stored in grib2/<date>/<hour>/WRF_EM/GRID255 and is included below as grib2/WRF_EM. Within the WRF_EM/GRID255 directory, HiResW-ARW-West has WMO headers like `[LM].B... KWBS.` ⑨ This model is clipped or subgridded, so the actual size can vary as the size of the domain changes.

grid/**HiResW-NMM-AK**

Note: The raw data for HiResW-NMM-AK is stored in grib2/<date>/<hour>/WRF_NMM/GRID255 and is included below as grib2/WRF_NMM. Within the WRF_NMM/GRID255 directory, HiResW-NMM-AK has WMO headers like `[LM].C... KWBS.` ⑨

grid/**HiResW-NMM-East**

Note: The raw data for HiResW-NMM-East is stored in grib2/<date>/<hour>/WRF_NMM/GRID255 and is included below as grib2/WRF_NMM. Within the WRF_NMM/GRID255 directory, HiResW-NMM-East has WMO headers like `[LM].A... KWBS.` ⑨ This model is clipped or subgridded, so the actual size can vary as the size of the domain changes.

grid/**HiResW-NMM-GU** | | 170 MB | | | 3,800 |

Note: The raw data for HiResW-NMM-GU is stored in grib2/<date>/<hour>/WRF_NMM/GRID255 and is included below as grib2/WRF_NMM. Within the WRF_NMM/GRID255 directory, HiResW-NMM-GU has WMO headers like `[LM].F... KWBS`. ⑨

grid/**HiResW-NMM-HI** | | 200 MB | | | 3,800 |

Note: The raw data for HiResW-NMM-HI is stored in grib2/<date>/<hour>/WRF_NMM/GRID255 and is included below as grib2/WRF_NMM. Within the WRF_NMM/GRID255 directory, HiResW-NMM-HI has WMO headers like `[LM].D... KWBS`. ⑨

grid/**HiResW-NMM-SJU** | | 300 MB | | | 3,800 |

Note: The raw data for HiResW-NMM-SJU is stored in grib2/<date>/<hour>/WRF_NMM/GRID255 and is included below as grib2/WRF_NMM. Within the WRF_NMM/GRID255 directory, HiResW-NMM-SJU has WMO headers like `[LM].E... KWBS`. ⑨

grid/**HiResW-NMM-West** | | 660 MB | | | 3,800 |

Note: The raw data for HiResW-NMM-West is stored in grib2/<date>/<hour>/WRF_NMM/GRID255 and is included below as grib2/WRF_NMM. Within the WRF_NMM/GRID255 directory, HiResW-NMM-West has WMO headers like `[LM].B... KWBS`. ⑨ This model is clipped or subgridded, so the actual size can vary as the size of the domain changes.

grid/**HPCqpf** – see grib/**NCEP_QPF** below.

grid/**HPCqpfNDFD** – see grib2/**NCEP_QPF** below.

manual/***HPCWWD*** | | 41 MB | | | 70 MB | | | 120 |

grid/**HPCWWD**

grid/**HPE** | | 5 – 50 MB | | | 300-500 |

Note: The raw data HPE (High-Resolution Precipitation Estimator) is not currently archived (grib files are stored in /awips2/edex/data/processing and XMRG files are in /awips2/edex/data/share/hydroapps). A future hydro archiver configuration will include the raw data. The file sizes are highly dependent upon local configuration (e.g., number of radars and the size of domain)

grib2/**HRRR** | | | 45 GB | | | 9-13 GB | | | 32,000 |

grid/**HRRR**

Note: This model is clipped or subgridded, so the actual estimates can vary as the size of the domain changes.

grib2/**HRRR.CR** – see grid/**GribModel:59:0:83** above.

grib/**ICE_120** | 200 KB | | | | 2 MB | | | 5 |

grid/**SeaIce**

grib2/**LAMP** | 50 MB | 180 MB | 1,700

grid/**GFSLAMPTstorm**

Note: This model is clipped or subgridded, so the actual size can vary as the size of the domain changes.

manual/***LAPS*** | 290 MB | 750 MB | 4,100

grid/**LAPS**

Note: The sizes of LAPS data are highly dependent upon the local configuration (especially the size of the local domain). The sizes shown here are an average from several different AWIPS sites.

320 MB | 13,000

grid/**mesoEta212**

Note: Raw data for mesoEta212 (also known as NAM40 for CONUS) is stored in grib/<date>/<hour>/NAM_84/GRID212. This model is clipped or subgridded, so the actual size can vary as the size of the domain changes.

grid/**mesoEta215** | 250 MB | 4,200

Note: Raw data for mesoEta215 (also known as NAM20 for CONUS) is stored under grib/<date>/<hour>/NAM_84/GRID215. This model is clipped or subgridded, so actual size can vary as the size of the domain changes.

grid/**mesoEta216** | 250 MB | 6,800

Note: Raw data for mesoEta216 for Alaska (45 km) is stored under grib/<date>/<hour>/NAM_84/GRID216.

grid/**mesoEta217** | 85 MB | 1,100

Note: Raw data for mesoEta217 for Alaska (22.5 km) is stored under grib/<date>/<hour>/NAM_84/GRID217.

grid/**mesoEta237** | 50 MB | 3,000

Note: Raw data for mesoEta237 for Puerto Rico (32 km) is stored under grib/<date>/<hour>/NAM_84/GRID237.

grid/**MOSGuide** | 180 MB | 1,500

Note: Raw data for MOSGuide is stored under grib2/<date>/<hour>/GFS/GRID197. This model is clipped or subgridded, so the actual size can vary as the size of the domain changes.

grid/**MOSGuide-AK** | 80 MB | 1,700

Note: Raw data for MOSGuide is stored under grib2/<date>/<hour>/GFS/GRID255 and is included in the size estimates for grib2/GFS above. Within the GFS/GRID255 directory, the WMO headers for MOSGuide-AK are like `[LM].R... KWBQ`. ⑩ This model is clipped or subgridded, so the actual size can vary as the size of the domain changes.

manual/grib/**mpe/*** | 400 KB – 2 MB | 500 KB | 24

grid/**MPE-Local**

grid/**MPE-Mosaic** (where come from?)

grid/**MPE-Local-***

100 KB – 1 MB			**2 - 24**		

Note: This is Multi-sensor Precipitation Estimate for each RFC (sites can be ALR, MSR, ORN, RHA, RSA, STR, TAR, TIR, and TUA). The sizes presented here are an average of daily estimates for individual sites. A very liberal estimate for all RFCs could total about 10 MB (if all RFCs had the maximum size), but the actual total is typically much smaller. The raw data for MPE-Local-* is stored in grib/NWS_160.

grid/**MPE-Mosaic-***

50 KB – 700 KB			**10-20**		

Note: This is Multi-sensor Precipitation Estimate for each RFC (sites can be ALR, FWR, MSR, ORN, RHA, TAR, and TIR). The sizes presented here are an average of daily estimates for individual sites. A very liberal estimate for the total amount of space for all RFCs could total about 5 MB (if all RFCs had the maximum size), but the actual total is typically much smaller. The raw data for MPE-Mosaic-* is stored in grib/NWS_161.

grib2/**MRMS/***

grid/**MRMS**

		55 GB			**6-13 MB**		**10,000**	

Note: The size estimates come from WDTD's feed from NSSL's full MRMS feed, and is subgridded or clipped. A WFO's domain is likely a different size than WDTD's domain.

manual/grib/***MSAS***

grid/**MSAS**

	7 MB			**15 MB**		**300**		

Note: The sizes of MSAS data are highly dependent upon the local configuration (especially the size of the local domain). The sizes shown here are an average from several different AWIPS sites.

grib/**NAM_84**

		1.1 GB						

Note: The processed versions of the grib versions of NAM_84 are stored as ETA207 (NAM_84/GRID207 for Alaska 95 km), ETA (NAM_84/GRID211 commonly known as NAM80 for CONUS), mesoEta212 (NAM_84/GRID212 commonly known as NAM40 for CONUS), mesoEta215 (NAM_84/GRID215 commonly known as NAM20 for CONUS), mesoEta216 (NAM_84/GRID216 for Alaska 45 km), mesoEta217 (NAM_84/GRID217 for Alaska 22.5 km), and mesoEta237 (NAM_84/GRID237 for Puerto Rico 32 km), all under grid.

grib2/**NAM_84**

		4 GB						

Note: The processed versions of the grib2 versions of NAM_84 are stored in ETA218 (NAM_84/GRID218 commonly known as NAM12 for CONUS) and ETA12 (NAM_84/GRID242 for Alaska 12 km also known as AK-NAM12).

grib2/**NamDNG25**

grid/**namdng25**

		3 GB			**1 – 5.8 GB**		**2,000**	

Note: This model is clipped or subgridded, so the actual size can vary as the size of the domain changes.

grid/**NamDNG5**

					1-2 GB		**3,000**	

Note: The raw data for NamDNG5 is stored in grib2/<date>/<hour>/NMM_89/GRID255 and is included in the size estimates for grib2/NMM_89 below. Within the

NMM_89/GRID255 directory, NamDNG5 has WMO headers like `[LM].[MO]... KWBE`. ① This model is clipped or subgridded, so the actual size can vary as the size of the domain changes.

manual/grib/***NARRE***
grid/**NARRE-TL-CR**

175 MB	575 MB	17,500
14 MB	1 MB	42

grib/**NCEP_QPF**
grid/**HPCqpf**
This model is clipped or subgridded, so the actual size can vary as the size of the domain changes.

5 MB	45 MB	318

grib2/**NCEP_QPF**
grid/**HPCqpfNDFD**

grid/**NCWF** – see grib/**AWC_NCWD** above.

270 MB	420 MB	318

grib2/**NDFD** (KWBN)
grid/**HPCGuide**

40 MB	60 MB	190

grib2/**NDFD** (KWNH)
grid/**HPCGuide-AK**

15 MB	8 MB	500

grib/**NMM_89**
grid/**ETA212**
Note: This is also known as NAMWX. The raw data is stored in grib/<data>/<hour>/NMM_89/GRID212 and the processed data is stored in grid/ETA212. This model is clipped or subgridded, so the actual size can vary as the size of the domain changes.

1.6 GB

grib2/**NMM_89**
Note: The processed data is stored as grid/NamDNG5, grid/AK-NamDNG5, grid/HI-NamDNG5, and grid.PR-NamDNG5. The four domains of NamDNG5 are all stored together in a grib2/<date>/<hour>/NMM_89/GRID255 directory with their respective WMO headers: `[LM].[MO]... KWBE` (CONUS), `[LM].A... KWBE` (Alaska), `[LM].H... KWBE` (Hawaii), and `[LM].C... KWBE` (Puerto Rico). ①

44 MB	85 MB	1,100

grib/**NOW**
grid/**GWW233**

grid/**NOHRSC-SNOW** – see grib/**NWS_185** below.

grib/**NWS_151** 11 MB

Note: The processed data is stored in the following grid directories for each RFC: FFG-ALR, FFG-FWR, FFG-KRF, FFG-MSR, FFG-ORN, FFG-PTR, FFG-RHA, FFG-RSA, FFG-STR, FFG-TAR, FFG-TIR, and FFG-TUA.

grib/**NWS_152** 260 MB

Note: The processed data is stored in the following grid directories for each RFC: QPE-ALR, QPE-FWR, QPE-KRF, QPE-MSR, QPE-ORN, QPE-RHA, QPE-STR, QPE-TAR, QPE-TIR, and QPE-TUA.

grib/**NWS_159** 4 MB 400 KB 12

grid/**QPE-AUTO-TUA**

grib/**NWS_160** 25 MB

Note: The processed data is stored as MPE-Local-ALR, MPE-Local-FWR, MPE-Local-MSR, MPE-Local-ORN,MPE-Local-RHA, MPE-Local-RSA, MPE-Local-STR, MPE-Local-TAR MPE-Local-TIR MPE-Local-TUA and GribModel:9:105:160 (for SJU, the CCCC for these is KALR).

grib/**NWS_161** 55 MB

output is stored as MPE-Mosaic-ALR --- GribModel:9:105:160 is SJU – grid # wrong?), MPE-Mosaic-FWR, MPE-Mosaic-MSR, MPE-Mosaic-ORN,MPE-Mosaic-RHA, MPE-Mosaic-RSA, MPE-Mosaic-TAR MPE-Mosaic-TIR

grib/**NWS_171** 30 MB

output is stored as QPE-XNAV-ALR --- GribModel:9:105:160 is SJU – grid # wrong?), QPE-XNAV-FWR, QPE-XNAV-KRF, QPE-XNAV-MSR, QPE-XNAV-ORN,QPE-XNAV-RHA, QPE-XNAV-TAR QPE-XNAV-TIR, QPE-XNAV-TUA

grib/**NWS_172** 7 MB

Note: The processed data is stored in the following grid directories for each RFC: QPE-RFC-PTR, QPE-RFC-RSA, and QPE-RFC-STR.

grib/**NWS_180** 25 MB 18 MB 250

grid/**RFCqpf**

grib/**NWS_185** 2 MB 600 KB 10

grid/**NOHRSC-SNOW**

grib/**NWS_190** 28 MB 4 MB 25

grid/**AUTOSPE**

grib/**NWW_121**
grid/**WNAWAVE238**

56 MB	**90 MB**	**1,100**

grib/**NWW_122**
grid/**AKWAVE239**

14 MB	**35 MB**	**1,000**

grib/**NWW_124**
grid/**ENPWAVE253**

72 MB	**105 MB**	**1,100**

grid/**PROB3HR** – see grib/**3hr** above.

grid/**PR-NamDNG5**

90 MB	**3,000**

Note: The raw data for PR-NamDNG5 is stored in grib2/<date>/<hour>/NMM_89/GRID255 and is included in the size estimates for grib2/NMM_89 above. Within the NMM_89/GRID255 directory, PR-NamDNG5 has WMO headers like `[LM].[C]... KWBE`. ①

grid/**PR-RTMA**

18 MB	**400**

Note: The raw data for PR-RTMA is stored in grib2/<date>/<hour>/RTMA/GRID255 and is included in the size estimates for grib2/RTMA below. Within the RTMA/GRID255 directory, PR-RTMA has WMO headers like `L.CA98 KWBR`. ②

grid/**QPE-***

500 KB – 2 MB	**24**

Note: This is Quantitative Precipitation Estimates (QPE) for each RFC (sites can be ALR, FWR, KRF, ORN, RHA, RSA, STR, TAR, TIR, and TUA). The sizes presented here are an average of the daily estimates for individual sites. A very liberal estimate for all RFCs could be up to 20 MB (if all RFCs had the maximum size), but the actual total is typically much smaller. The raw data for QPE is stored in grib/NWS_152.

grid/**QPE-AUTO-TUA** – see grib/**NWS_159** above.

grid/**QPE-RFC-***

200 – 600 KB	**10-24**

Note: This is Quantitative Precipitation Estimates (QPE) for each RFC (sites can be PTR, RSA, STR; sites not in grid/QPE*). The sizes presented here are an average of the daily estimates for individual sites. A very liberal estimate for all three RFCs could be up to 2 MB (if they all had the maximum size), but the actual total is typically much smaller. The raw data for QPE-RFC-* is stored in grib/NWS_ 172.

grid/**QPE-XNAV-***

100 – 600 KB	**10**

Note: This is XNAV version of Quantitative Precipitation Estimate for each RFC (sites can be ALR, FWR, KRF, MSR, ORN, RHA, TAR, TIR, TUA). The sizes presented here are an average of the daily estimates for individual sites. A very liberal estimate for all these RFCs could be up to 5 MB (if they all had the maximum size), but the actual total is typically much smaller. The raw data for QPE-XAV* is stored in grib/NWS_171.

grid/**RTGSSTHR** – see grib/**SST** below.

grib2/**RTMA** 3.2 GB

Note: The processed data stored as grid/RTMA (CONUS), grid/AK-RTMA, grid/HI-RTMA, grid/PR-RTMA, and grid/Guam-RTMA. The raw data is stored in grib2/<date>/<hour>/RTMA/GRID197 (CONUS) and in grib2/<date>/<hour>/RTMA/GRID255 with these respective WMO headers: `L.AA98 KWBR` (AK-RTMA), `L.HA98 KWBR` (HI-RTMA) `L.CA98 KWBR` (PR-RTMA), `L.GA98 KWBR` (Guam-RTMA). ②

grid/**RTMA** 515 MB 500

Note: The raw data for RTMA is stored in grib2/<date>/<hour>/RTMA/GRID197 and is included in the size estimates for grib2/RTMA above. This model is clipped or subgridded, so the actual size can vary as the size of the domain changes.

.

grid/**RTMA25** or
grid/**RTMA-Mosaic** 600 MB 550

Note: This model is clipped or subgridded, so the actual size can vary as the size of the domain changes. Sizes as high as 2.3 GB have been observed.

grib2/**RTOFS** 2 GB

Note: This is data from the **R**eal-**T**ime **O**cean **F**orecast **S**ystem. The processed data is stored as grid/RTOFS for several domains, with respective WMO headers: `E.A.01 KWBW` (Alaska), `E.B.01 KWBW` (Bering), `E.C.01 KWBW` (West CONUS), `E.D.01 KWBW` (Arctic), `E.G.01 KWBW` (Guam), `E.I.01 KWBW` (Gulf of Alaska), `E.K.01 KWBW` (West Atlantic), `E.H.01 KWBW` (Honolulu), `E.S.01 KWBW` (Samoa), `E.T.01 KWBW` (Tropical Pacific), `E.J.01 KWBW` (Hudson Baffin). As of the time of this compilation, all raw data for the RTOFS are stored together including the nowcasts (processed data is grid/RTOFS-Now-*) and forecasts (processed data is grid/RTOFS-*). ⑪

grid/**RTOFS-Alaska** 68 MB 700

grid/**RTOFS-Arctic** 150 MB 700

grid/**RTOFS-Bering** 260 MB 700

grid/**RTOFS-Guam** 320 MB 700

grid/**RTOFS-GulfAlaska** 14 MB 700

grid/**RTOFS-Honolulu** 410 MB 700

grid/**RTOFS-HudsonBaffin** 55 MB 700

grid/**RTOFS-Now-Alaska** 35 MB 350

grid/**RTOFS-Now-Arctic** 75 MB 350

grid/**RTOFS-Now-Bering** 130 MB 350

grid/**RTOFS-Now-Guam** 160 MB 350

grid/**RTOFS-Now-GulfAlaska** 7 MB 350

grid/**RTOFS-Now-Honolulu** 210 MB 350

grid/**RTOFS-Now-HudsonBaffin** 30 MB 350

grid/**RTOFS-Now-Samoa** 145 MB 350

grid/**RTOFS-Now-TropPaciLowres** 13 MB 350

grid/**RTOFS-Now-WestAtl** 140 MB 350

grid/**RTOFS-Now-WestConus** 180 MB 350

grid/**RTOFS-Samoa** 285 MB 700

grid/**RTOFS-TropPaciLowres** 25 MB 700

grid/**RTOFS-WestAtl** 275 MB 700

grid/**RTOFS-WestConus** 355 MB 700

grib/**RUC2** 1.6 GB

Note: The corresponding processed data is stored as grid/RUC236 (for grib/<date>/<hour>/RUC2/GRID236) and grid/Aviation (for grib/<date>/<hour>/RUC2/GRID130).

grib2/**RUC2** 5.9 GB

Note: The processed data is stored as grid/RUC130.

grid/**RUC130** 7 GB 135,000

Note: This model is clipped or subgridded, so the actual size can vary as the size of the domain changes.

grid/**RUC236** | 830 MB | 21,000

grid/**SeaIce**– see grib/**ICE_120** above.

grid/**SPCGuide** – see grib/**GRID001** above.

grib/**SPEC62MRF** | 380 MB

Note: The corresponding processed data is stored as grid/ENSEMBLE38 (for SPEC62MRF/GRID038), ENSEMBLE39 (for SPEC62MRF/GRID039), and ENSEMBLE40 (for SPEC62MRF/GRID040).

grib2/**SREF_113** | 285 MB

Note: The corresponding processed data is stored as grid/SREF212 (for SREF_113/GRID212), grid/SREF216 (for SREF_113/GRID216), and grid/SREF243 (for SREF_113/GRID255).

grid/**SREF212** | 475 MB | 12,000

Note: The raw data for SREF212 is stored in grib2/<date>/<hour>/SREF_113/GRID212 and is included in the size estimates for grib2/SREF_113 above. This is a 40-km CONUS grid.

grid/**SREF216** | 350 MB | 13,000

Note: The raw data for SREF216 is stored in grib2/<date>/<hour>/SREF_113/GRID216 and is included in the size estimates for grib2/SREF_113 above. This is a 45 km Alaska grid.

grid/**SREF243** | 280 MB | 12,000

Note: The raw data for SREF243 is stored in grib2/<date>/<hour>/SREF_113/GRID255 and is included in the size estimates for grib2/SREF243 above. This is a Pacific grid.

grib/**SSIGFS** | 32 MB

Note: The corresponding processed data of the grib versions of SSIGFS are stored in grid/AVN211 (for SSIGFS/GRID211), grid/AVN225 (for SSIGFS/GRID225), grid/ENSEMBLE38 (for SSIGFS/GRID038), grid/ENSEMBLE39 (for SSIGFS/GRID039), and grid/ENSEMBLE40 (for SSIGFS/GRID040). These are the 00-hour forecast (i.e., initialization field) grids of the GFS.

grib2/**SSIGFS** | 190 MB

Note: The corresponding processed data of the grib2 versions of SSIGFS are stored in grid/GFS160 (for SSIGFS/GRID160), grid/GFS161 (for SSIGFS/GRID161), grid/GFS201 (for SSIGFS/GRID201), grid/GFS212 (for SSIGFS/GRID212), grid/GFS213 (for SSIGFS/GRID213), and grid/GFS254 (for SSIGFS/GRID254). These are all the 00-hour forecast (i.e., initialization field) grids of the GFS.

grib/**SST** | 6 MB | 13 MB | 1

grid/**RTGSSTHR**

grib2/**TPC** 32 MB 300 MB 2,000

grid/**TPCSurgeProb**

Note: Other products may be stored in grib2/TPC because this is not a baseline pqact.conf configuration. These products are only generated during active tropical cyclone systems. This model is clipped or subgridded, so the actual size can vary as the size of the domain changes.

20 MB 800

grid/**TPCWindProb**

Note: The raw data for TPCWindProb is stored in grib2/<date>/<hour>/FORECASTER.

grib/**UKM_45** 35 MB

Note: The processed versions of UKM_45 are stored as grid/UKMET37 through grid/UKMET40 (for GRID037 through GRID040) and as the UKMET-NorthernHemisphere composite.

grid/**UKMET37** through grid/**UKMET40** 30 MB 1,200

Note: The results shown here are for each grid individually, so each grid separately consumes 30 MB of space for a total of 120 MB per day for all four grids.

grid/**UKMET-Northern Hemisphere** 160 MB 1,000

grib2/**URMA** 2 GB 2.5 GB 500

grid/**URMA25**

Note: This model is clipped or subgridded, so the actual size can vary as the size of the domain changes.

grid/**WCwave4** 280 MB 3,300

Note: The raw data for WCwave4 is stored in grib2/<date>/<hour>/GMGWM/GRID255, and is included above in the size estimates for grib2/GMGWM. Within the GMGWM/GRID255 directory, WCwave4 has WMO headers like: `E.G.88 KWBJ`. ④ ⑤

grid/**WCwave10** 120 MB 3,300

Note: The raw data for WCwave10 is stored in grib2/<date>/<hour>/GMGWM/GRID255, and is included above in the size estimates for grib2/GMGWM. Within the GMGWM/GRID255 directory, WCwave10 has WMO headers like: `E.C.88 KWBJ`. ④ ⑤

grid/**WNAwave4** 300 MB 3,300

Note: The raw data for WNAwave4 is stored in grib2/<date>/<hour>/GMGWM/GRID255, and is included above in the size estimates for grib2/GMGWM. Within the GMGWM/GRID255 directory, WNAwave4 has WMO headers like: `E.H.88 KWBJ`. ④ ⑤

grid/**WNAwave10** | | 300 MB | | | 3,300 |

Note: The raw data for WNAwave10 is stored in grib2/<date>/<hour>/GMGWM/GRID255, and is included above in the size estimates for grib2/GMGWM. Within the GMGWM/GRID255 directory, WNAwave10 has WMO headers like: `E.B.88 KWBJ`. ④ ⑤

grid/**WNAWAVE238** – see grib/**NWW_121** above.

grib2/**WRF_EM** | | | 2.5 GB |

Note: The corresponding processed versions of WRF_EM are stored as six versions of grid/HiResW-ARW for the different domains (AK, East, GU, HI, SJU, and West). The raw data is stored in grib2/<date>/<hour>/WRF_EM/GRID255 with respective WMO headers: `[LM]... KWBS` (AK), `[LM].A... KWBS` (East), `[LM].F... KWBS` (GU), `[LM].D... KWBS` (HI), `[LM].E... KWBS` (SJU), and `[LM].B... KWBS` (West). ⑨

grib2/**WRF_NMM** | | | 2.5 GB |

Note: The processed versions of WRF_NMM are stored as six versions of HiResW-NMM for the different domains (AK, East, GU, HI, SJU, and West). The raw data is stored in grib2/<date>/<hour>/WRF_NMM/GRID255 with respective WMO headers: `[LM]... KWBS` (AK), `[LM].A... KWBS` (East), `[LM].F... KWBS` (GU), `[LM].D... KWBS` (HI), `[LM].E... KWBS` (SJU), and `[LM].B... KWBS` (West). ⑨

Local or Regional Models via LDAD:

manual/grib
grid/OUNWRF | | | 25 GB | | | 62 GB | | | 160,000 |

manual/grib
grid/wrf-arw-rnk | | | 3.5 GB | | | 10 GB | | | 58,000 |

manual/grib
grid/ww3-2km (LOT) | | | 3.5 GB | | | 1.1 GB | | | 11,000 |

This is the end of the gridded data section. Regular datasets resume below this point.

idft | | 15 MB | | | 21,000 |

Note: The raw data is stored under forecasts with WMO header FZXX41 KWNO. The corresponding text bulletins are also in the text (fxatext) database, and are included in the processed data (.bin files) for text below.

intlsigmet | 300 KB | | | 100 KB | | | 30 | | |

Note: The processed versions of intlsigmet text bulletins are also in the text (fxatext) database and are included in processed data (.bin files) for text below.

ldadhydro 3 MB | 10 – 25 MB | 2,000 – 15,000

Note: The raw data for ldadhydro comes from /data_store/ldad with files that are named like ldad.hydro.*.xml. /data_store/ldad is not part of the default archiver configuration so the local configuration should include /data_store/ldad to ensure the raw data is properly saved. The sizes of both raw and processed data are highly dependent on the local LDAD configuration.

ldadmesonet 30 - 175 MB | 140 – 900 MB | 50,000 – 350,000

Note: The raw data for ldadmesonet comes from /data_store/ldad with files that are named like ldad.mesonet.*.xml. /data_store/ldad is not part of the default archiver configuration so the local configuration should include /data_store/ldad to ensure the raw data is properly saved. The sizes of both raw and processed data are highly dependent on the local LDAD configuration.

ldad (SHEF) 44 MB

Note: The raw data included in this size estimate come from /data_store/ldad with files that are named like SXUS44 KWOHshef.* The corresponding processed data are stored in the WHFS (hd_ob92xxx) database.

lsr 500 KB | 3 MB | 500

maritime 65 MB

Note: The corresponding processed versions of maritime data are included in sfcobs. The corresponding text bulletins are also in the text (fxatext) database, and are included in the processed data (.bin files) for text below.

MAROB 250-400 KB

Note: The corresponding processed versions of MAROB data are part of sfcobs.

metar **obs** 260 MB | 350 MB | 170,000

Note: The raw data are stored in a metar directory and the processed data are stored in an obs directory. The corresponding text bulletins are also in the text (fxatext) database, and are included in the processed data (.bin files) for text below.

misc_adm_ **messages** 40 MB

Note: The processed versions of misc_adm_messages data are in the text (fxatext) database and are included in the processed data (.bin files) for text below.

misc_sfc_obs 2.5 MB

Note: The processed versions of misc_sfc_obs data is in the text (fxatext) database and are included in the processed data (.bin files) for text below.

modelsounding 1.5 GB | 1.5 GB | 720,000

nonconvsigmet | 500 KB | | | 100 KB | | | 30 | | |

Note: The processed versions of nonconvsigmet text bulletins are also in the text (fxatext) database and are included in processed data (.bin files) for text below.

nucaps | | 180 MB | | | 175 MB | | | 75,000 | |

obs – see **metar** above.

pirep | | 4-8 MB | | | 5 MB | | 300-1,200 | | |

Note: Processed versions of airep text bulletins are in the text (fxatext) database and are included in processed data (.bin files) for text below.

poessounding | | 70 MB | | | 5 MB | | | 3,500 | |

Note: At the time of this compilation (Summer 2015), most WFOs were storing raw poessounding data, but not processing it or archiving processed poessounding data due to local office configurations not having an appropriate poesBufr.spi file in /awips2/edex/data/utility/common_static/site/<SITE>/basemaps. This information comes from WDTD's system where the size of the processed data was highly dependent on this .spi file as well as the weather conditions.

preciprate | | | | | 9 - 30 MB | | | 1,000 | |

Note: preciprate is a decision assistance tool (DAT) which is part of the SCAN system. It is produced from radar data. The actual size is highly dependent on the local configuration.

profiler | 300 KB | | | | 3 MB | | 100 | | |

qc | | | | | 100-400 MB | | | 115,000 – 475,000 | |

Note: This data is quality control information that is related to the hydro and mesonet data ingested via LDAD. The size estimates therefore are highly dependent upon the local LDAD configuration.

qpf | | | | | 150 MB | | | 7,500 | |

Note: qpf is a decision assistance tool (DAT) which is part of the SCAN system. It is produced from local radar data and a grid from the RUC130 model.

radar (dedicated WSR-88D) | | | 3 -12 GB | | | 600 MB – 3.5 GB | | 12,000-40,000 | |

Note: These size estimates are per radar per day for a dedicated WSR-88D (data are processed through an RPG and a RadarServer). These estimates are also highly weather dependent, which is why size ranges are presented.

radar (SBN WSR-88D) | | 100 MB - 2 GB | | | 200 MB – 2 GB | | | 8,000-25,000 | |

Note: These size estimates are per radar per day for data ingested via the AWIPS Satellite Broadcast Network (SBN), or NOAAPORT. These estimates are also highly weather dependent, which is why size ranges are presented.

radar (local TDWR)

500 KB – 1 GB	250 – 700 MB	5,000-10,000

Note: These size estimates are per radar per day for a local TDWR (data are processed through an SPG and a RadarServer). These estimates are also highly weather dependent, which is why size ranges are presented.

radar (SBN TDWR)

30 MB – 500 MB	100 MB – 500 MB	4,000-7,000

Note: These size estimates are per radar per day for data ingested via the AWIPS Satellite Broadcast Network (SBN), or NOAAPORT. These estimates are also highly weather dependent, which is why size ranges are presented.

raobs

5 MB

Note: The processed versions of raobs data is in the text (fxatext) database and are included in the size estimates for text given below. These are the raw/coded TTAA/TTBB sounding observations. AWIPS-2 soundings that are plotted in NSHARP are derived from the bufrua versions.

redbook

8 MB	10 MB	700

regionalsat – see one of the **satellite** rows below.

sat
satellite

10 GB	8 GB	5,000

Note: These estimates are for baseline satellite data. Raw data are stored in the sat directory and processed data are in the satellite directory. Experimental and non-baseline data are ingested via LDAD and are given below under satellite.mcidas and regionalsat.

satellite.mcidas
satellite

24 - 50 GB	4.5 GB	550 - 1000

Note: The satellite.mcidas data are ingested via LDAD and the raw data appear in manual/satellite.mcidas. These size estimates are for one WFO's configuration. The processed files appear in the satellite directory alongside the baseline satellite data. Satellite.mcidas processed files can be identified through their AREA* directory names.

regionalsat
satellite

5 - 8 GB	1 - 3 GB	2,500

Note: The regionalsat data are ingested via LDAD and the raw data appear in the manual/regionalsat. These size estimates are based on a small number of WFO configurations which differed significantly from one another. The processed files appear in the satellite directory alongside the baseline satellite data. Some of the directory names for regionalsat data include eastConus, westConus, and conusONE.

scan

50 MB	4,000

Note: scan is a decision assistance tool (DAT) which is part of the SCAN system. It is produced from radar data and other data sources including RUC and lightning data.

sfcobs

330 MB | 120,000

Note: sfcobs is produced from raw data stored in maritime, MAROB, and synoptic directories.

shef

3 GB

Note: The processed versions of shef data are in the text (fxatext) database and hydro WHFS (hd_ob92xxx) databases. Local SHEF data ingested through LDAD are also included in the hydro database.

summaries

32 MB

Note: Processed versions of summaries text bulletins are in the text (fxatext) database and are included in processed data (.bin files) for text below. This raw data are WMO headers that begin with A.

svrwx

500 KB – 1 MB | 200 KB – 1.5 MB | 50 - 300

Note: The raw data for svrwx are text products that have WMO headers like `NWUS20`. Processed versions of summaries text bulletins are also in the text (fxatext) database and are included in processed data (.bin files) for text below

synoptic

4 MB

Note: The processed data are in the sfcobs directory. Some synoptic data is also processed by the shef plugin and is included in the WHFS (hd_ob92xxx) database. Processed versions of synoptic text bulletins are also in the text (fxatext) database and are included in processed data (.bin files) for text below

taf

35 MB | 65,000

Note: raw data is included in forecasts with WMO headers like `FT*`. The processed data are also in the text (fxatext) database and are included in the processed data (.bin files) for text below.

text

35 MB | 800 MB – 1.1 GB | 275,000-325,000

Note: The raw data contains text products that are not included in other directories. The processed data (.bin files) are extracts from the text (fxatext) database and include the processed versions of raw data in the following directories: airep, airmet, climate, convsigmet, fire_wx_spot_forecast_reports, forecast, intlsigmet, lsr, maritime, metar, misc_adm_messages, misc_sc_obs, nonconvsigmet, pirep, radar, raobs, shef, summaries, svrwx, synoptic, taf, tcg, tcm, tcs, upperair, wwa, and xml.

upperair

30 MB

Note: The processed versions are in the text (fxatext) database and are included in the size estimates for processed data (.bin) for text above. This data can contain radiosonde data outside US.

vaa

50 KB | 20

Note: The raw data for vaa are in the forecast directory with WMO headers like `FVXX20` through `FVXX27` and `FVAK20` through `FVAK24`. The processed data are also in the text (fxatext) database and are included in the size estimates for processed data (.bin) for text above.

viirs

8.9 GB | 700

Note: The raw data is located in satellite/VIIRS. As of the time of this compilation, baseline VIIRS imagery is only valid for Alaska, though some sites were experimenting with VIIRS data for the CONUS that was ingested via their LDAD.

vil | 500 MB | 2,500

Note: vil is a decision assistance tool (DAT) which is part of the SCAN system. It is produced from radar data.

wwa | 10 MB | 30-45 MB | 3,000-5,000

warning

Note: Raw data are stored in the wwa directory and the processed data are stored in the warning directory, which is the source of the warning polygon displays. These are also text products and the processed versions are also stored in the text (fxatext) database and are included in the size estimates for text above. In addition, the raw data for a watch product from SPC (WMO header `WWUS60 KWNS`) is stored in wwa and the processed version is stored in the wcp directory.

wcp | 100 KB | 25

Note: Raw data for wcp is stored in the wwa directory, using the WMO header `WWUS60 KWNS`.

xml | 12 MB

Note: The processed versions are in the text (fxatext) database and are included in the size estimates for processed data (.bin) for text above.

References for WMO Headers:

①http://www.nco.ncep.noaa.gov/pmb/changes/nam_dng_wmoheaders.shtml
②http://www.nws.noaa.gov/infoservicechanges/tin11-42rtma_oper.txt
③http://www.nws.noaa.gov/os/notification/tin13-39rtma_q4.htm
④ http://www.nws.noaa.gov/os/notification/tin08-23nww3.txt
⑤ http://www.nco.ncep.noaa.gov/pmb/docs/headers/awpwave/
⑥ http://www.nws.noaa.gov/os/notification/tin13-43estofs_noaaport_aaa.htm
⑦ http://www.nco.ncep.noaa.gov/pmb/changes/gfs_dng_guam.shtml
⑧ http://www.mdl.nws.noaa.gov/~naefs_ekdmos/EKDMOS_WMO_Headers.pdf
⑨ http://www.nco.ncep.noaa.gov/pmb/changes/hiresw_wmo_headers.shtml
⑩ http://www.nws.noaa.gov/mdl/synop/gmos/gmosAKheaders.pdf
⑪ http://www.nws.noaa.gov/os/notification/tin14-25rtofs_grib.htm

2. “Thinning” datasets: raw and processed data.

It is important to trim both raw and processed data. There is a “rawPlay” script that is used with WES-2 Bridge to re-process raw data. However, this reprocessing can be very time consuming, so removing unneeded raw data can save significant amounts of time.

“Thinning” datasets refers to the process of removing selected portions of data from within a datatype, rather than the whole datatype. It is therefore helpful to be familiar with the organization of the various archived datasets.

A. Processed Data

Most processed data, with the exceptions of bufrsigwx, gfe, grid, modelsounding, radar, redbook, satellite, sfcobs, and viirs, are organized simply by hour with an HDF file and a similarly-named hourly directory that contains the corresponding database extracts (.bin) files for that same hour. For example, a Processed/obs directory within an archived case might contain metar-2015-06-12-05.h5 and a corresponding metar-2015-06-12-05 directory which itself would contain metar-2015-06-12-05.bin.1. The bin files are numbered because each file is limited to 10,000 records to aid in performance; when data for a given hour contains more than 10,000 records, additional numbered bin files are created by the AWIPS-2 archiver.

Some processed data exist only in a PostgreSQL database and have no .h5 files, but they still have .bin files. (Remember, every piece of data in AWIPS-2 has a database component.) Text is a good example, where an hourly directory (such as text-2015-06-23-15) might contain, say, ten .bin files, named text-2015-06-23-15.bin.01 through text-2015-06-23-15.bin.10. At any rate, this hourly file and directory structure permits easy thinning of datasets by time to within an hour’s resolution. This structure is similar to AWIPS-1 where netCDF files were organized by hour.

The exceptions for the processed data storage are given below with an example to indicate the storage structure.

- bufrsigwx:

```
bufrsigwx/SWH/sigwxCAT-2015-06-20-18.h5
bufrsigwx/SHW/sigwxCAT-2015-06-20-18/sigwxCAT-2015-06-20-18.bin.1
```

⬆

Type (can be SWH, SWM, or SWBOTH)

- gfe: (XXX refers to a WFO site ID)

```
gfe/XXX/GFS20/2015_06_18_1800/XXX_GRID__GFS20_20150618_1800_MAXRH_SFC.h5
gfe/XXX/GFS20/2015_06_18_1800/XXX_GRID__GFS20_20150618_1800_MAXRH_SFC/XXX_GRID__GFS20_20150618_1800_MAXRH_SFC.bin.1
gfe/XXX/SAT/XXX_GRID__SAT_20150618_VisibleE_SFC.h5
gfe/XXX/SAT/XXX_GRID__SAT_20150618_VisibleE_SFC/XXX_GRID__SAT_20150618_VisibleE_SFC.bin.1
```

- grid:.

```
grid/HRRR/MB/HRRR-2015-06-19-20-FH-005.h5
grid/HRRR/MB/HRRR-2015-06-19-20-FH-005/HRRR-2015-06-19-20-FH-005.bin.1
```

model (coord type) — model run cycle — forecast hour

coord type is a coordinate (e.g, MB, FHAG, SFC, EA, meaning millibars [constant pressure], fixed height above ground, surface, or entire atmosphere – there are several others).

In AWIPS-1 grid data were only organized by hour, but in AWIPS-2, it is feasible to trim irrelevant forecasts that could be irrelevant to a simulation. For example, a typical severe weather case, there is usually no need to have forecasts for time periods longer than 36 or 48 hours.

In these sample filenames, the grids referenced in these files are the various constant pressure fields for the HRRR valid for June 20, 2015 at 01Z (five hours after the model is initialized at 20Z on June 19).

- modelsounding: Processed/modelsounding/model/cycle/hourly_directory. model can be either GFS or ETA, cycle is the time of the model run, and the hourly_directory is similar to the ones used in grid.

```
modelsounding/ETA/2015-06-19_12:00:00.0/modelsounding-ETA-2015-06-19-12-FH-11.h5
modelsounding/ETA/2015-06-19_12:00:00.0/modelsounding-ETA-2015-06-19-12-FH-11/modelsounding-ETA-2015-06-19-12-FH-11.bin.1
```

model — model run cycle — forecast hour

For modelsounding, the filename structure is similar to grid. The model can be either ETA or GFS.

- radar:

```
radar/kama/0.5/radar-kama-2015-06-16-17.h5
radar/kama/0.5/radar-kama-2015-06-16-17/radar-kama-2015-06-16-17.bin.1
```

radar site — elevation angle — product hour

elevation angle can be 0.0, for volumetric products that do not have a single elevation angle.
The radar HDF file contains all the radar products for a single radar and elevation that were observed in that hour.

This particular file contains reflectivity, velocity, spectrum width, and all the dual-pol products for the 0.5 elevation angle observed between 17:00Z and 17:59Z.

Radar data in AWIPS-1 were stored individual Level-3 product files (very similar to the raw radar data in AWIPS-2 that comes from an RPG) and had time resolution for each volume scan of the radar. If you are concerned about a case playback or simulation on WES-2 Bridge stopping at a particular volume-scan time, additional data filtering can be accomplished with WES-2 Bridge.

- redbook:

```
redbook/PEIB40/redbook-2015-06-17-06.h5
redbook/PEIB40/redbook-2015-06-17-06.bin.1
          ⬆
```

WMO Header ID

- satellite:

```
satellite/East CONUS/Imager Visible/satellite-2015-06-18-14.h5
satellite/East CONUS/Imager Visible/satellite-2015-06-18-14/satellite-2015-06-18-14.bin.1
              ⬆              ⬆
```

Sector Sensor

- sfcobs:

```
sfcobs/1001/sfc-2015-06-17-11.h5
sfcobs/1001/sfc-2015-06-17-11/sfc-2015-06-17-11.bin.1
          ⬆
```

Type (1001 through 1007)

1001	Synoptic Fixed Land	1005	Synoptic Moored Buoy
1002	Synoptic Mobile Land	1006	Drifting Buoy
1003	Synoptic Ship	1007	Synoptic MAROB
1004	Synoptic CMAN		

- viirs:
  `viirs/Alaska/Imagery/3.74/BT/viirs-2015-06-23-00.h5`
  `viirs/Alaska/Imagery/3.74/BT/viirs-2015-06-23-00/viirs-2015-06-23-00.bin.1`

B. Raw Data

The organization of raw data is controlled by the LDM pqact.conf file for data ingested via the SBN and additionally by the LDAD system for local data. For most datatypes (exceptions are acars, grib, grib2, ldad, manual, radar, redbook, satellite) raw data are organized simply by date and hour. For example, metars are stored in metar/YYYYMMDD/HH, such as metar/20150618/23). Each file begins with its WMO header.

The exceptions are given below:

- acars

  `acars/acars_decrypted/20150619/11/IUAX02_KARP_191104_275128503.acars.2015061911.9lOw5H`

- grib/grib2

  `grib2/20150619/15/RUC2/GRID130/1500Z_F009_TMPK-LTDJ10_KWBG_191500_65495296.grib2.2015061915`

  File Type | Model ID | Grid # | Cycle Time | Fcst Hour | Para-meter | WMO Header | Sequence Number

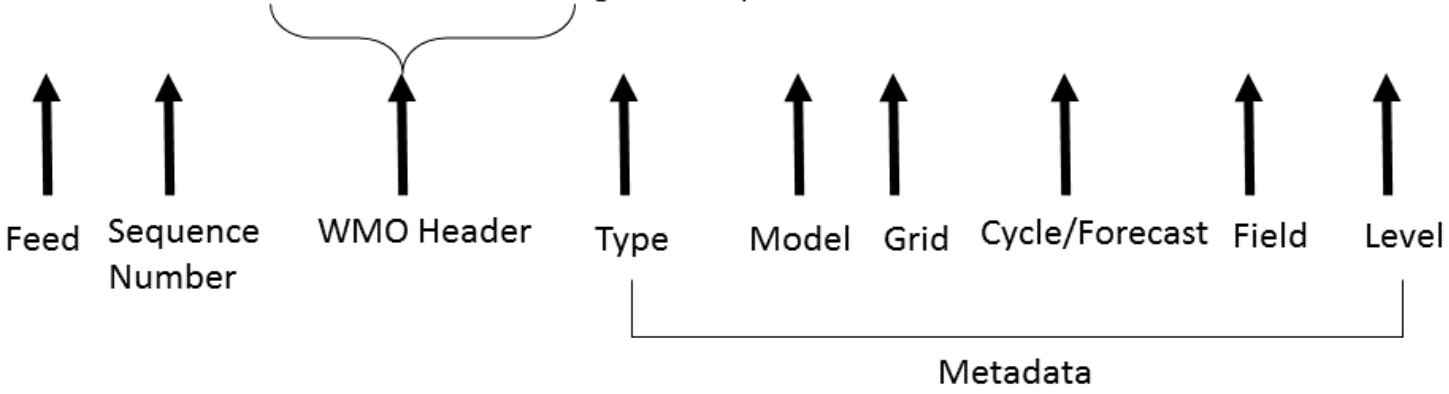
  Model ID is the name of a model or model process, such as RUC2 or NWS_161. The filename has a number of product-dependent fields (model cycle time, forecast hour, model parameter) followed by a dash, and then the WMO header, the date/time, LDM sequence number, etc. as in the regular filenames.

  Note how these product fields are determined. The LDM software extracts metadata from individual files in its product stream. This metadata is then used to name the files in /data_store. Because LDM is written and

maintained by Unidata at UCAR and not by the NWS or its contractor, some of the metadata determination can be incomplete, especially as new models are added. In addition, sometimes the metadata information for model grids uses a grid number for 255 to denote a wildcard or an undefined grid. In these cases, multiple domains may be stored commingled together in the same directory. Fortunately, in most cases, domains are also signified by the WMO headers and can be separated. In situations where this happens, the table above presents the WMO headers that represent various geographical domains in the notes below specific gridded datatypes. Examples are given below on how to use this information.

Several symptoms occur when new models are stored by the LDM prior to the metadata routines in the LDM being updated. The metadata can be incorrect, or non-existent. So the forecast hours and the parameter name may be missing from the filename. In addition, when a model is completely new, the LDM may not recognize the model's process ID. This ID is the primary information that distinguishes one model from another. When it cannot translate the process ID into a name, the LDM may use the process ID as the name; the model is therefore stored in a directlry that begins with anumber. For example, when the HRRR model was first transmitted over the SBN, the LDM at many sites stored the HRRR data in grib2/<date>/<hour>/83 rather than grib2/<date>/<hour>/HRRR.

```
NGRID 68942098  LTSA25 KWBE 231800 !grib2/ncep/NAM_84/#242/201506231800F000/TMPK/250 hPa PRES
NGRID 68942107  LWSA98 KWBE 231800 !grib2/ncep/NAM_84/#242/201506231800F000/CAPE/0 - NONE
NGRID 68942110  LHSA50 KWBE 231800 !grib2/ncep/NAM_84/#242/201506231800F000/HGHT/500 hPa PRES
NGRID 68942111  LTSA57 KWBE 231800 !grib2/ncep/NAM_84/#242/201506231800F000/TMPK/575 hPa PRES
NGRID 68942117  LHSA62 KWBE 231800 !grib2/ncep/NAM_84/#242/201506231800F000/HGHT/625 hPa PRES
NGRID 68942127  LRSA70 KWBE 231800 !grib2/ncep/NAM_84/#242/201506231800F000/RELH/700 hPa PRES
NGRID 68942136  LTSA97 KWBE 231800 !grib2/ncep/NAM_84/#242/201506231800F000/TMPK/0 - TROP
NGRID 68942154  LRSA94 KWBE 231800 !grib2/ncep/NAM_84/#242/201506231800F000/RELH/0 - FRZL
NGRID 68942156  LHSA94 KWBE 231800 !grib2/ncep/NAM_84/#242/201506231800F000/HGHT/0 - FRZL
NGRID 68941574  LTCA98 KWBR 231900 !grib2/ncep/RTMA/#255/201506231900F000/TMPK/2 m HGHT
NGRID 68941590  LRHA98 KWBR 231900 !grib2/ncep/RTMA/#255/201506231900F000/DWPK/2 m HGHT
NGRID 68941591  LTHA98 KWBR 231900 !grib2/ncep/RTMA/#255/201506231900F000/TMPK/2 m HGHT
```



The NAM_84 is given for GRID # 242, which turns out to be the 12km NAM for Alaska. Notice the RTMA has GRID # 255 and has two parameters that are identical (the bottom line and third from the bottom line both have Temperature at 2 m). The only difference between the two is the WMO header (LTCA98 vs LTHA98, which are domains for Puerto Rico and Hawaii, respectively).

- ldad: LDAD data have no directory structure; it only has filenames.

```
LDAD.hydro.1434623540.CoCoRaHS.decoded.xml
LDAD.mesonet.1434914444.RAWS.decoded.xml
LDAD.mesonet.NOS-NWLON.1435075200.51.msas_qc.decoded.xml
SXUS44 KWOHshef.1434877170
```

- radar: Radar data have three structures, depending upon the source of the radar data.

  AWIPS Satellite Broadcast Network (SBN) or NOAAPORT:

  ```
  radar/20150623/19/KSRX/N0R/KSRX_N0R_231949-SDUS54_KTSA_290068923.rad.2015062319
  ```

  The product ID is the special three-character ID that is used for NOAAPORT dissemination. A table of these IDs is provided in the AWIPS-2 Data Type Reference.

  RPG: Radar Product Generator for WSR-88D

  ```
  /data_store/radar/kmkx/Z/elev0_5/res0_25/az0_5/level256/kmkx.153.20150620_0222
  ```

  The data storage for WSR-88D raw data from an RPG is very similar to the old AWIPS-1 radar storage; the difference is that the radar ID and product code number are prefixed to the volume scan time in the filename.

  SPG: Supplemental Product Generator for TDWR

  ```
  /data_store/radar/tmke/Z/elev4_3/res0_15/level256/tmke.180.20150620_2159
  /data_store/radar/tmke/Z/elev4_3/res0_15/level256/netcdf/Reflectivity/20150620_220037.gz
  ```

  Some of the base products (Reflectivity, Velocity, and Spectrum Width) for TDWR have extra netcdf versions created for use in FSI. The filename time for the netcdf file apparently is a file write time, and not the volume scan time (note the time mismatch between these two corresponding files).

- Redbook

  ```
  /data_store/redbook/20150622/16/PXS0024CN/1616Z_PXS0024CN_CONVOL_NMCGPHFW1-PMWE98_KWNS_286294876.rb.2015062216
  /data_store/redbook/20150622/18/PXSF001CN/1823Z_PXSF001CN_MCDSUM_NMCGPHMCD-PGNA00_KWNS_286585379.rb.2015062218
  /data_store/redbook/20150622/18/PXSF000CN/1800Z_PXSF000CN_WNDGEO_NMCGPH9AM-PYWA31_KWBC_286629048.rb.2015062218
  ```

  The first two products are from the Storm Prediction Center (convective outlook and a summary of current mesoscale convective discussions) and the last product is a geostrophic wind plot.

- Satellite

  ```
  /data_store/sat/20150622/20/GOES-13/2045Z_VIS_1km_EAST-CONUS-TIGE01_KNES_216202.satz.2015062220
  ```

Examples of Thinning Archived Cases

Example 1. Trimming processed model data for forecast periods greater than 36 hours.

Let's say you have placed a copy of a case of processed data to be trimmed in /data1/mycase. In this situation, the processed grid directory will be /data1/mycase/Processed/grid. (Remember, working with the original archived case is dangerous).

So from the grid directory, you have the following example directories and files:
ETA218/FHAG/ETA218-2015-06-18-12-FH-039.h5
ETA218/FHAG/ETA218-2015-06-18-12-FH-039/ETA218-2015-06-18-12-FH-039.bin.1

Given this directory structure, the following sequence of Linux shell commands will delete all processed model files with forecast hours > 36.

```
cd /data1/mycase/Processed/grid
rm –rf */*/*-FH-03[789]*          Deletes data with forecast hours -037, -038, and -039
rm –rf */*/*-FH-0[456789]*        Deletes data with forecast hours between 40 and 99
rm –rf */*/*-FH-[12]*             Deletes data with forecast hours greater than 100
                                  (the longest forecast period is 10 days or 240 hours)
```

Example 2. Trimming raw model data for forecast periods greater than 36 hours.

For this example, we'll use the same basic premise as the previous example, except we'll be working in a Raw data directory in /data1/mycase/Raw. Model data that follows the proper storage convention in pqact.conf (based on having correct metadata) will be in either a grib or grib2 directory. For example:

```
/data1/mycase/Raw/grib/20150623/12/GFS/GRID211/1200Z_F120_RH-YRQO30_KWBC_231200_289553275.grib.2015062316
```

Here are commands that will delete all files with F037 F038 or F039 in their filenames, representing forecast hours:

```
$ cd /data1/mycase/Raw
$ find grib/ grib2/ -name '????Z_F03[789]_*-*.grib*' –print –delete
```

deletes raw data with forecast hours F037, F038, and F039.

The ? matches any one character of a filename, so the ????Z_F ensures we are actually deleting files that match our expected structure. The forecast hour always has three digits. The square brackets denote a set of characters to be matched, so F03[789] actually means F037, F038, and F039. The remaining _*-*.grib* just provides insurance that we are actually going to delete files that still match our expected structure with an underscore after the forecast hour and the hyphen before the WMO header. The -print displays a log on the screen of the files that are actually deleted.

The two commands below, respectively, delete the files with forecast hours between 40 and 99 and those with a first digit of 1 or 2. There are no forecast hours greater than 240 (ten days).

```
$ find grib/ grib2/ -name '????Z_F0[456789]?_*-*.grib* -print -delete
$ find grib/ grib2/ –name '????Z_F[12]??_*-*.grib* -print -delete
```

It is important to be especially careful with the find command with the -delete flag. Always make sure you have it at the end of the command; having it at the beginning can delete all files in the given directories. You may even want to run a similar command first without deleting the files to ensure you will delete only the desired files. A sample command to do this is:

```
$ find grib/ grib2/ -depth –name '????Z_F[12]??_*-*.grib* -print
```

These find commands make take a while to run because they are searching every single file within your grib and grib2 directories. In a typical case, there can easily be millions of files that are being searched and compared.

Example 3. Trimming raw model data for models by geographic domain.

There are two methods that may need to be applied to thin raw model data by geographic domain. The first is to delete whole GRID directories that are known to be outside the area of interest. For example, if the case is desired to be valid for only the CONUS, then grid directories for Alaska, Hawaii, and Puerto Rico may be deleted. The second method is to delete files from within GRID255 directories that are for domains outside the area of interest by using known WMO headers that signify domains.

For either of these methods, you want to consider the estimated daily sizes presented in the table above. If the sizes are not significant, then it may not be worth investing your time to weed out extra data. For example, there may be a tendency to delete many of the NWS_* grids (for various hydro products like FFG, QPE, QPF, MPE, etc.) and their corresponding processed directories. However, the total sizes of these products are small compared to other larger datasets like radar, satellite, and high-resolution models. So the advice of this writer is to simply leave the hydro products alone in the case.

In judging significance of data sizes, you'll need to look at both the raw and the corresponding processed directories.

Method A: Whole Directories. Delete files outside of CONUS. This is a list of OCONUS grids from the table above:

Directory	GRID	Model (Domain)	Size
grib2**/DGEX_115**	GRID186	DGEX186 (Alaska)	235 MB
grib**/NAM_84**	GRID207	ETA207 (Alaska)	50 MB
grib2**/NAM_84**	GRID242	ETA242 (Alaska)	8 GB
grib2**/GFS / SSIGFS**	GRID160	GFS160 (Alaska)	3.5 GB
grib2**/GFS / SSIGFS**	GRID161	GFS161 (Puerto Rico)	2 GB
grib2**/GFS / SSIGFS**	GRID254	GFS254 (Pacific)	12 GB
grib**/NAM_84**	GRID216	mesoEta216 (Alaska)	250 MB
grib**/NAM_84**	GRID217	mesoEta217 (Alaska)	85 MB
grib**/NAM_84**	GRID237	mesoEta237 (Puerto Rico)	50 MB

The sizes comes from the processed data that correspond to the raw data directories. The estimated total space saved in processed data when this data is reprocessed is approximately 25 GB. Simply delete these directories one-by-one.

```
$ cd /data1/mycase/grib2
$ rm –rf */*/DGEX_115/GRID2196
```

(etc.)

Method B: From GRID255 directories using WMO headers:

The notation used in the table above for WMO headers is regular expression notation where a dot (.) character represents any single character and characters within square brackets represents a set of characters. For example, [LM] means both L and M. The regular expressions can be changed to filename patterns simply by replacing each dot with a question mark.

Again, by looking through the table, here are the various models that have GRID255 directories that contain data for multiple domains.

Model Type	Directory	Domain	WMO Header	Header Combo
NamDNG	/grib2/NMM_89	Alaska	[LM].A... KWBE	[LM].[ACH]... KWBE
		Hawaii	[LM].H... KWBE	
		Puerto Rico	[LM].C... KWBE	
RTMA	/grib2/RTMA	Alaska	L.AA98 KWBR	L.[ACGHK]A98 KWBR
		Alaska 3-km	L.KA98 KWBR	
		Guam	L.GA98 KWBR	
		Hawaii	L.HA98 KWBR	
		Puerto Rico	L.CA98 KWBR	
HiResW-ARW HiResW-NMM	/grib2/WRF_EM /grib2/WRF_NMM	Alaska	[LM].C… KWBS	[LM].[CDEF]…. KWBS
		Guam	[LM].F… KWBS	
		Hawaii	[LM].D… KWBS	
		Puerto Rico	[LM].E… KWBS	
MOSGudie-AK	/grib2/GFS	Alaska	[LM].R… KWBQ	

Example commands using ls and rm or find and given below for the NamDNG:

- ls and rm

```
$ cd /data1/mycase/Raw/grib2
$ ls */*/NMM_89/GRID255/*-[LM]?[ACH]???_KWBE*.grib*
$ rm */*/NMM_89/GRID255/*-[LM]?[ACH]???_KWBE*.grib*
```

- find

```
$ cd /data1/mycase/Raw
$ find grib2/ –depth –regex ‘.*NMM_89/GRID255/.*-[LM].[ACH]..._KWBE.*grib.*’ –print
$ find grib2/ -regex ‘.*NMM_89/GRID255/.*-[LM]?[ACH]???_KWBE*.grib*’ –print –delete
```

Example 4. For an interior CONUS WFO, you may want to remove whole models that are for marine applications (e.g., wind/wave models, ocean and extra-topical surge). Here is a partial list of models that you could consider removing (or not even archiving for an interior CONUS site in the first place.) Again, some of these take up very little space, but many of them are significant consumers of space.

Raw: ESTOFS (2.1 GB), GLWM (225 MB) or GLWM25 (550 MB), GMGWM (1.3 GB), NWW_121 (56 MB), NWW_122 (14 MB), NWW_124 (72 MB), RTOFS (2 GB)

Processed: AKwave4 (350 MB), AKwave10 (215 MB), AKWAVE239 (35 MB), ENPWAVE253 (105 MB), EPwave10 (125 MB), estofsPR (425 MB), estofsUS (4 GB), GlobalWave (12 GB), GRLKWave (500 MB) or WW2-2km (higher resolution of Great Lakes Model; 1.1 GB), RTOFS-* (3.3 GB total), WCwave10 (120 MB), WCwave4 (280 MB), WNAwave4 (300 MB), WNAwave10 (300 MB), WNAWAVE238 (90 MB).