Elements of a Scalable Infrastructure for Weather Forecaster Access to Joint Polar Satellite System (JPSS) Data

John D. Evans, Ph.D.
Global Science & Technology, Inc. (GST)
NOAA Joint Polar Satellite System
Algorithm Management Project
Lanham / Greenbelt, MD

99th American Meteorological Society Annual Meeting
15th Annual Symposium on New Generation Operational Environmental Satellite Systems
5B Special Session on JPSS Series Satellite System — Part II — Jan. 9, 2019
Joint Polar Satellite System

- Polar-orbiting satellites, S-NPP and NOAA-20 (a.k.a. JPSS-1)

- Instruments:
  - Visible/Infrared Imaging Radiometer Suite (VIIRS)
  - Advanced Technology Microwave Sounder (ATMS)
  - Cross-track Infrared Sounder (CrIS)
  - Ozone Mapping and Profiler Suite (OMPS)

- Each satellite orbits 14x/day
- Each images the globe 2x/day
- Each produces 2TB/day globally

- Challenge: providing forecasters the data they need in a timely fashion
Most weather forecasters don’t need all of that data

- Forecasters in the field generally need regional data, rather than global.
- They also don’t need every single format / aggregation / variant of data
- No one forecasting office will need every data product.

<table>
<thead>
<tr>
<th>Sector</th>
<th>Daily VIIRS granules, per satellite</th>
<th>GB/day, per satellite</th>
<th>NWS sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alaska</td>
<td>128</td>
<td>74</td>
<td>5</td>
</tr>
<tr>
<td>Pacific</td>
<td>216</td>
<td>124</td>
<td>3</td>
</tr>
<tr>
<td>CONUS</td>
<td>78</td>
<td>45</td>
<td>128</td>
</tr>
<tr>
<td>Puerto Rico</td>
<td>23</td>
<td>13</td>
<td>1</td>
</tr>
<tr>
<td>All sectors</td>
<td>368</td>
<td>212</td>
<td></td>
</tr>
</tbody>
</table>

* For all products tagged by NWS as Key Performance Parameter (KPP), Critical, Supplemental High, or Suppl. Low
Each JPSS satellite sees (some of) CONUS about 10 times per day (78-80 VIIRS granules / day)
Near-real-time users: one overpass at a time

How to get data products from one overpass (~7 minutes of observations) to users quickly enough?
From each JPSS overpass, how many of these products can we afford to send to forecast offices?
Which products can we afford to get from each overpass?

Product volumes (MB) for a 7-minute JPSS overpass; and Cumulative bandwidth needs for timely (10-minute) delivery (==> “bandwidth budget”).

(Cryosphere products [for Alaska], not shown here, would require an add’l 1.6 to 3.4 Mbps.)
Getting JPSS data products to AWIPS

McMurdo (Antarctica)

Svalbard (Norway)

JPSS satellites (S-NPP, NOAA-20, etc.)

Direct Broadcast

NOAA Direct Broadcast Real Time Network

NWS Satellite Broadcast Network (SBN)

LDM / FTP

SBN uplink

LDM (Alaska)

SBN Broadcast

AWIPS sites

AWIPS-Data Delivery (data)

AWIPS-Data Delivery (metadata)

PDA standing subscriptions

Svalbard (Norway)

Science Mission Data downlink (half/full orbit)

NESDIS data processing & distribution systems (IDPS - NDE - PDA)

NWS AWIPS Network Control Facility (NCF)

McMurdo (Antarctica)
• “Mainstream” path into AWIPS forecaster workstations
• 69+ Mbps total; 6 Mbps available for polar satellite data
• Everyone receives the same thing – e.g., for JPSS:
  • VIIRS Near-Constant Contrast (NCC) Day-Night Band imagery for Alaska, Pacific, CONUS, Puerto Rico
  • NUCAPS soundings for Americas + Pacific & East Asia
  • VIIRS bands I1, I4, I5 imagery for Alaska region

The SBN disseminates satellite, model, radar and other products to NWS AWIPS field nodes
• SBN expanded from 30 Mbps to 60+ Mbps
• Only SBN-related product flows shown
SBN is nearly full already

Mb/s needed for 10-minute delivery of model and satellite data over a 24-hour period
(Not shown: NEXRAD radar data)
AWIPS Data Delivery (AWIPS-DD)

On-demand services, connecting AWIPS to ESPDS
Product Distribution and Access (PDA) and others
• Deliver only what users request
• Less need to pinpoint end-user needs
  • Less predictable usage patterns

AWIPS-DD & PDA use an asynchronous protocol
• Loosely based on the Open Geospatial Consortium
  Web Coverage Service (WCS) standard
• A fairly complex protocol, but has been shown to work with
  GOES-R data; AWIPS team is now adapting it to polar data
• Fetches discrete JPSS product files (no on-demand “tailoring”)

Data, metadata, requests, and responses travel via
OneNWSnet TCP/IP fiber-optic network
• Now 100Mbps => a workable “bandwidth budget”
Direct Broadcast: Game-changing latency

For comparison: latencies via Svalbard / McMurdo

Distribution of intervals between granule timestamp and product delivery time
NOAA-20 and SNPP VIIRS KPPs (Imagery EDRs for Alaska)
via Svalbard/McMurdo -> PDA -> NCF (Feb. 11, 2018)

Total Img. EDRs delivered: 775 (NOAA-20); 878 (S-NPP)
Ensuring scalability

• Current technologies (esp. AWIPS-DD / PDA via OneNWSnet’s 100 Mbps) will give forecasters timely access to the data products they need... For now. Handling many more users and more data will require

• Reducing unnecessary data movement – *e.g.*,
  • Produce and disseminate smaller (“thinned”) versions of products
  • Subset data on demand by location, time, or parameter

• Limiting server loads – *e.g.*,
  • Tiered Content Distribution Network
  • Conditional (or on-demand) processing

• Emphasizing simplicity, fault tolerance, interoperability
Reducing data movement:

Thinned products for NWS forecasters

VIIRS Aerosol Optical Depth
(9-171 MB)

VIIRS Aerosol Optical Depth (thinned for NWS)
(6-14 MB)

VIIRS Volcanic Ash
(6-120 MB)

VIIRS Volcanic Ash (thinned for NWS)
(6-24 MB)

CrIS / ATMS NUCAPS
(3 MB)

CrIS / ATMS NUCAPS (thinned for NWS)
(0.2 MB)
Reducing data movement: Streamlined, interoperable Web services

Full OGC Web Coverage Service (WCS) capability would include on-demand selection and subsetting by location, time, and field/parameter;

- Perhaps also Resampling, Aggregation, and Reprojection

This may further reduce data transfer volumes (4x ~ 100x)

Use of actual OGC WCS protocol would also bring

- Simpler client-server interaction
- Possibility of COTS solutions
- Interoperable & reusable service

WCS client / server interactions

- Get content summary (GetCapabilities)
- Describe A, B, C (DescribeCoverage)
- Query A, B, C by time, location, etc. (GetCoverage)
Instead of having all end users fetch products from a single site, maybe distribute to one or more tiers of “edge” servers – likely via Cloud Computing.

After each CONUS overpass of a JPSS satellite, up to ~100 sites may opt to fetch ~5-10 products from each of ~5 granules, within a short time.

- 5,000 concurrent file transfers!? 
Limiting server loads:

Conditional / On-demand product generation

- Instead of running the full suite of algorithms on every data granule, maybe generate some less-frequently used products only under certain circumstances, or only when requested by users.

- Cloud Computing would allow rapid and temporary “scale-out” of processing resources when needed.

JPSS algorithm run times and interdependencies

Maybe run some specialized products only by request?
Summary

• Providing forecaster access to Joint Polar Satellite System data products is a significant challenge;

• However, given expected patterns of data supply and demand from polar orbiting satellites, and the new OneNWSnet bandwidth, PDA and AWIPS-Data Delivery will be able (for now) to provide forecasters with timely access to the JPSS data products they need.

• As data volumes and usage grow, we will need more scalable approaches to product generation, distribution, and access – for example:
  • Reduce data movement via thinned products, improved data access services, etc.;
  • Use industry-standard protocols for simplicity, versatility, and resilience; and
  • Limit server loads via tiered content distribution and conditional processing.
Near-real-time users: Mb/s, not GB/day

Method: Collected file sizes and file creation times for 44 SNPP and AMSR2 products of interest to AWIPS, intersecting the NWS CONUS region on July 18, 2018. Summed MB (received / sent) per 10-minute interval to infer Mbps needed to “clear the buffer” in each interval. Stacked-area chart above shows Mbps needed for the 20 largest products (~90% of daily data volume).

Findings: Total daily data volume for this region: 45GB / day; but bandwidth needs vary from 0 Mbps to over 58 Mbps (e.g., at 15:10 above).
Sidebar: WCS extensions under development

WCS extensions for swath data (EOX @ OGC Testbed 14)

WCS GetCorridor extension (UK Met Office)