# CubeWerx advanced map tiling technology

Intelligent, on-demand web map tiles for big imagery

## Map tiling for big data

In the realm of geospatial data management, efficiently handling large-scale datasets is crucial. CubeWerx Stratos stands out as a robust platform designed for the rapid processing and serving of both raster and vector geospatial data. It excels in generating map tiles from extensive raster mosaics, capable of managing collections that include hundreds of thousands of scenes without the significant overhead typically associated with such tasks.

Stratos streamlines the update process within these collections, avoiding the lengthy and resource-intensive regeneration of image pyramids traditionally required for web map visualization. This efficiency is achieved with minimal storage requirements, utilizing less than 5% of the original data volume.
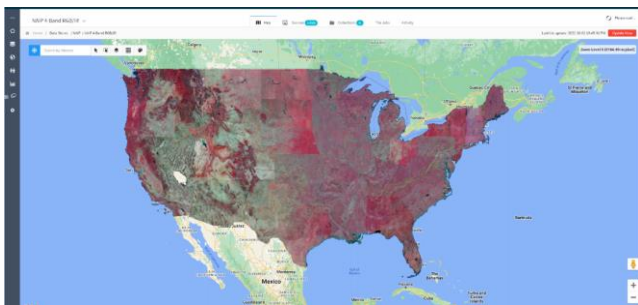


**Fig 1:** 247,923 individual satellite images, comprising over 100 TB of data – processing completed in 66 hours on 8 vCPU cores with 16GB of RAM (1.6 TB/hour)

## Leveraging Cloud Optimized GeoTIFFs

CubeWerx Stratos enhances its handling of large geospatial datasets through the adoption of Cloud-optimized GeoTIFF (COG) files. COGs, an advancement of the GeoTIFF format, are designed for efficient web access, supporting HTTP range requests and random-access from network storage. This format is inherently aligned with Stratos's operational efficiencies, particularly in web map tiling.

Stratos includes functionality to convert non-COG images to COG format, facilitating the integration of legacy data.

COGs' internal tiling and multi-resolution support negate the need for Stratos to generate additional map tiles, directly utilizing the COG's built-in structure for web service delivery. This approach significantly reduces storage requirements. However, for inputs not in COG format, Stratos automatically generates overview files, slightly increasing storage use but maintaining overall system efficiency.

## Processing data in situ

Stratos tackles the challenge of managing vast datasets through in situ data processing. This technique allows for the direct handling of large rasters within their storage locations, avoiding the need to duplicate, fully load into memory, or incorporate the entire dataset into a database.

This approach enables Stratos to efficiently offer value-added services over publicly hosted data, such as the USDA 1-meter 4-band dataset on AWS on the left, by leveraging existing cloud archives. In situ processing significantly enhances Stratos's capability to manage and serve large-scale geospatial data efficiently.

## Parallel image processing

The platform takes full advantage of every available thread on the data production hardware to split the job of tiling a large data set into as many parallel processes as possible. A smart controller process farms out the jobs to sub processes and monitors their execution. The number of threads to use is a configurable setting. The system can be instructed to limit itself to a given number so as not to starve other processes.

CubeWerx

## Efficient just-in-time tiling

The platform uses a streamlined approach to map tiling, focusing on efficiency without sacrificing user experience. By limiting the generation of map tiles to 8x the source data's resolution, Stratos significantly reduces the time needed to make datasets available online. This decision avoids the extensive processing typically associated with generating high-resolution tiles for the entire dataset, many of which may never be viewed.

For instances where finer detail is requested by users, Stratos employs a dynamic, on-demand tiling process. This method generates high resolution tiles only when needed, storing them in an ephemeral cache for quick retrieval on subsequent requests. This dynamic approach has been rigorously optimized to ensure that, even under the strain of serving hundreds of users simultaneously, the system remains responsive. Performance tests have shown that Stratos can maintain this level of service using just 8 vCPUs and 16GB of RAM.

## Minimal file management

More primitive tiling systems that need to generate complete tile sets typically produce millions of tiny files. This can cause problems with many file systems and makes maintenance extremely difficult. With the CubeWerx platform, only the ephemeral cache (high resolution) tiles are stored directly on the file system. All other tiles are kept in the database, which greatly simplifies management, backup and recovery.

## Dynamic reprojection and colormaps

Map tiles stored in the database are not converted into JPG or PNG images. They are stored as losslessly compressed GeoTIFF data. This means that they preserve the data types and multi-spectral characteristics of the original source data. This is important, because it allows the system to quickly apply different color maps to numeric data, perform multiband composition, or to accurately re-project data into different coordinate systems on the fly, without losing information.

## Smart scene management makes updates simple

The platform maintains a complete data catalog of each scene in a mosaic, including its geospatial footprint. This allows the tile update process to identify exactly which portions of the mosaic are affected by any changes and do the minimum amount of processing required to update the map. Even with very large mosaics, additions, updates and deletions happen very quickly.



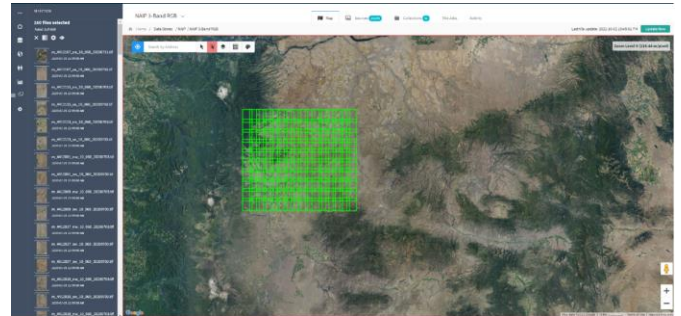**Fig 2:** Selection of multiple scenes for update using the spreadsheet functions



**Fig 3:** Selection of multiple scenes using the visual map tools

## Hundreds of spectral indices built-in

Stratos includes the Awesome Spectral Index library, a ready-to-use curated list of spectral indices for remote sensing applications. The server can apply any index on-the-fly for instantaneous streaming of informative web maps from multi-spectral data.

Visit https://cubewerx.com/ or email us at info@cubewerx.com today to learn more about our cutting-edge platform.

CubeWerx