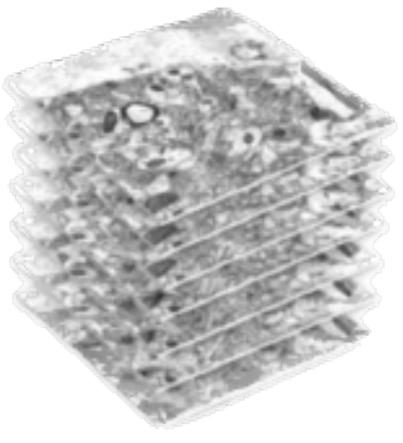
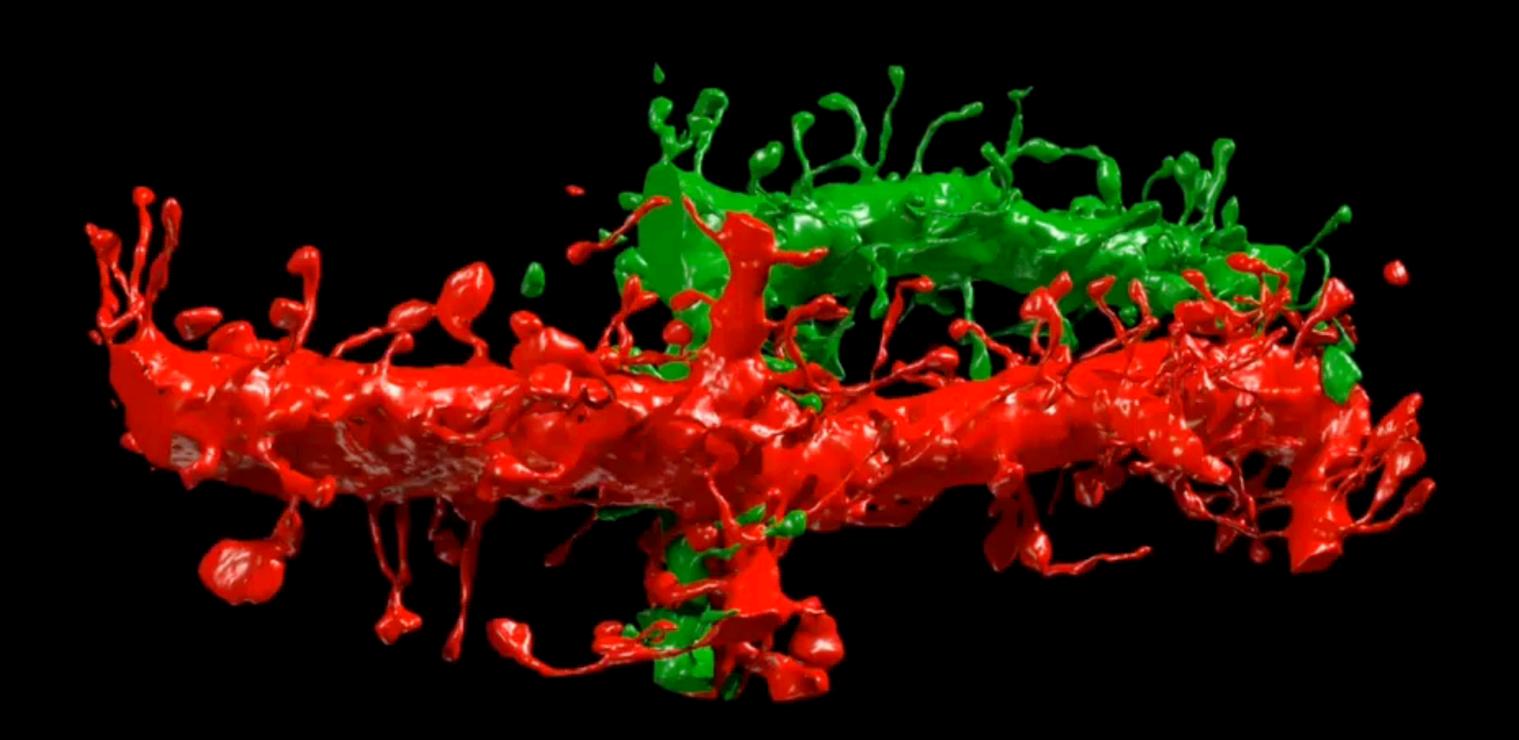


## Kasthuri et al, Cell 2015

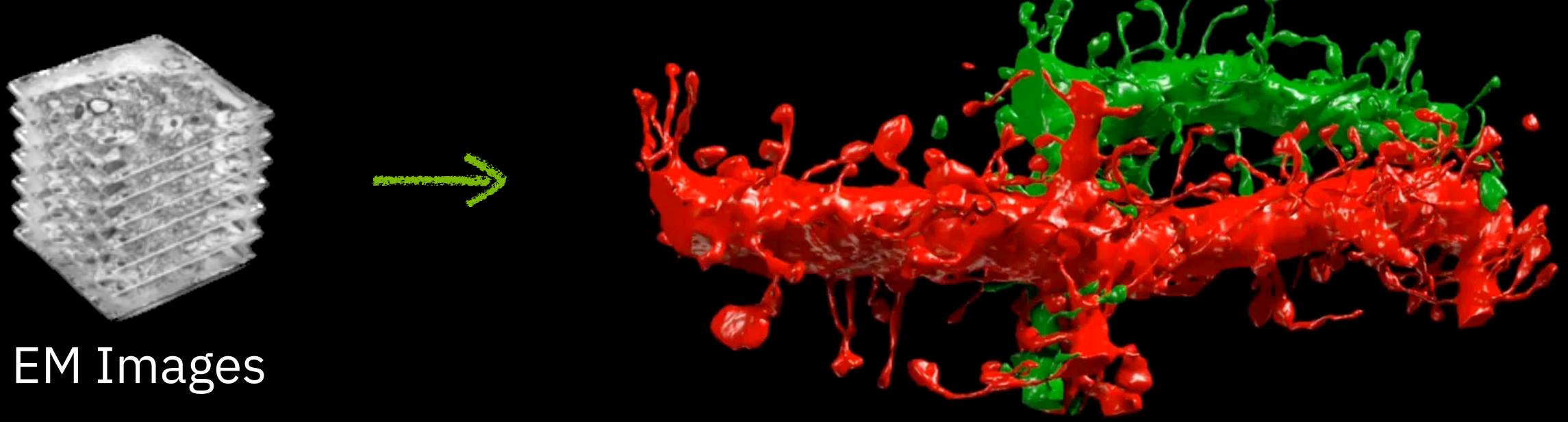


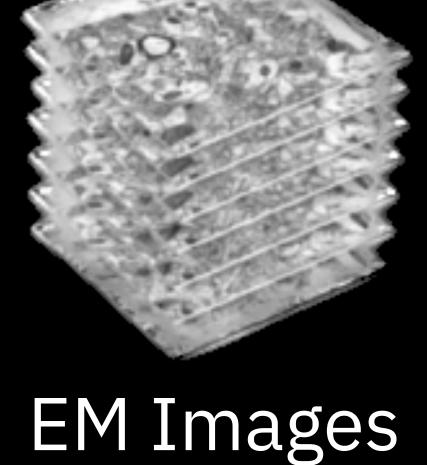






## Reconstructing the Brain Connectivity





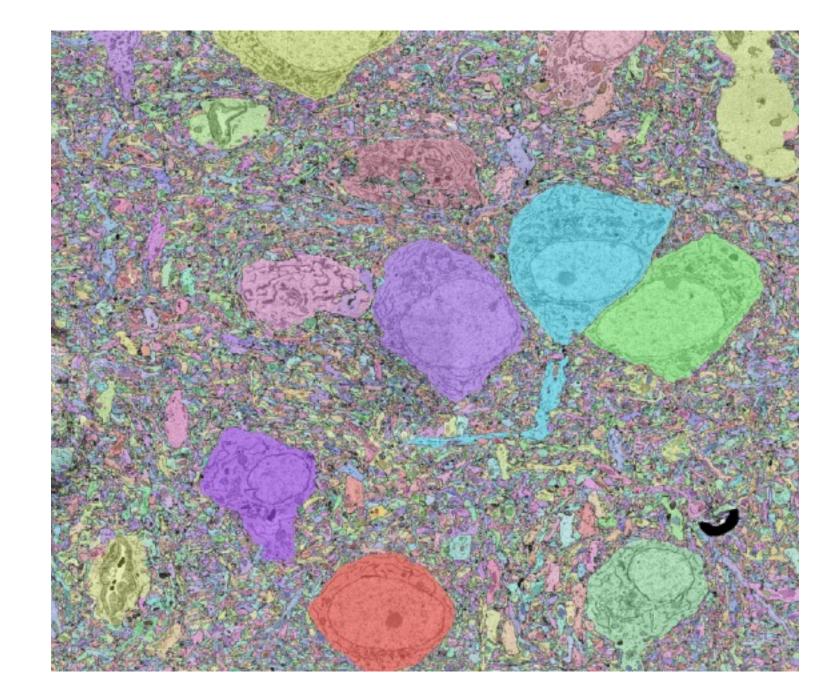


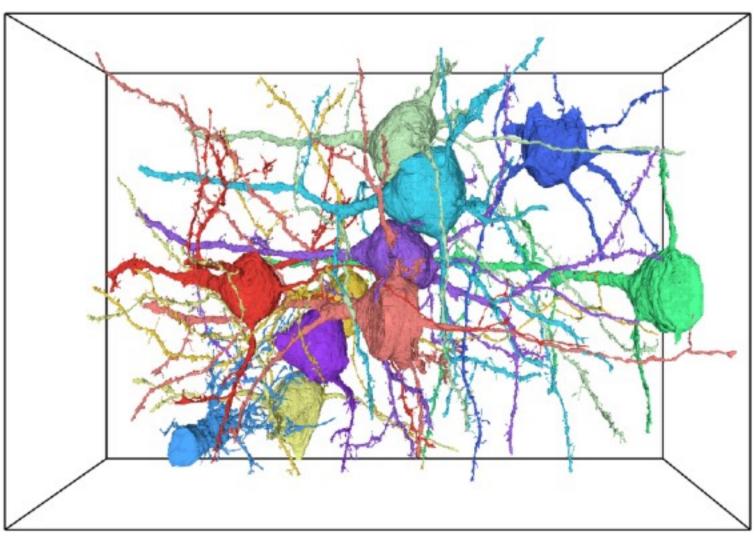
## Large-Scale Reconstruction

- Inference (and training) has scaled on CPU-based and GPU-based supercomputers (parallel granularity: overlapping subvolumes)
  - -Achieved million-way concurrency on Theta supercomputer
- Image stitching and alignment components are being scaled as well to ensure a scalable end-to-end pipeline

## **Exascale Inference Problem:**

- On a single GPU (A100), we achieve ~80 MegaVoxels/hour using 32-bit (There is still room for improvement here)
- In reduced precision (8-16 bits), we expect ~1 GigaVoxel/hour per GPU
- ■1 PetaVoxel (1mm³) will take ~1M GPU node hours
- Approximately, 24 hours on a system with 50K GPUs (considering overlapping sub-volumes)
- For a mouse brain (1cm³), 1 ExaVoxel, we would need ~3 years on an exascale system





Dong, et al, "Scaling Distributed Training of Flood-Filling Networks on HPC Infrastructure for Brain Mapping", 2019 IEEE/ACM Third Workshop on Deep Learning on Supercomputers (DLS) at SC19

