

# IDIA(2011)

## CASA Imaging I/O Performance

James Robnett  
Computing Infrastructure  
NRAO/NM

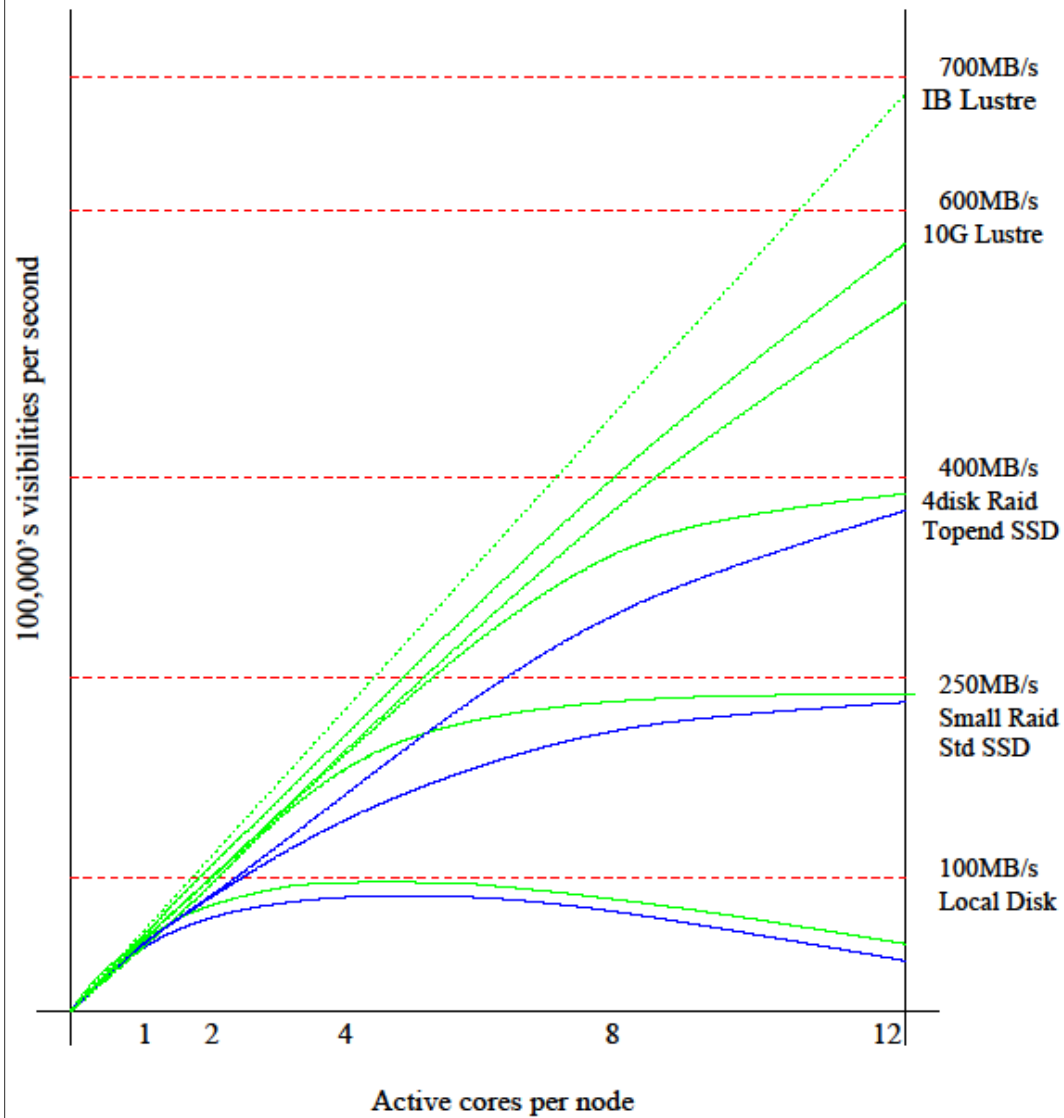
# IDIA(2011)

## CASA Imaging I/O Performance

- Measurement sets distributed to independent iPython processes
  - Separate processing and async I/O threads
  - I/O optimized for large (>1MB) sequential reads
- 100GB high frequency continuum EVLA data set
- Lustre filesystem feeding processing nodes over 40Gbit QDR Infiniband
- Async I/O from disk to CPU; memory stores images not raw visibilities
- 40 - 50MB/s I/O consumed per processing thread at 95% user CPU load
- 200 operations / unit I/O
- Lustre operating at 80% aggregate spindle speed (spindle speed limited)

Visibilities gridded/second VS CPU model and I/O subsystem

- 2.4 GHz 6 core Nehalem (E5645)
- 2.9 GHz 6 core Nehalem (E5670)



# IDIA(2011)

## CASA Imaging I/O Performance

- CPU and I/O financially balanced for simple imaging
- Industry increases over next decade result in 50 to 1 disparity in CPU to disk I/O rates
- $10^3$  operations per unit I/O for MS-MFS effectively increases life span of I/O subsystem
- $10^4$  or  $10^5$  ops per unit I/O cost for higher sensitivity further improves lifespan