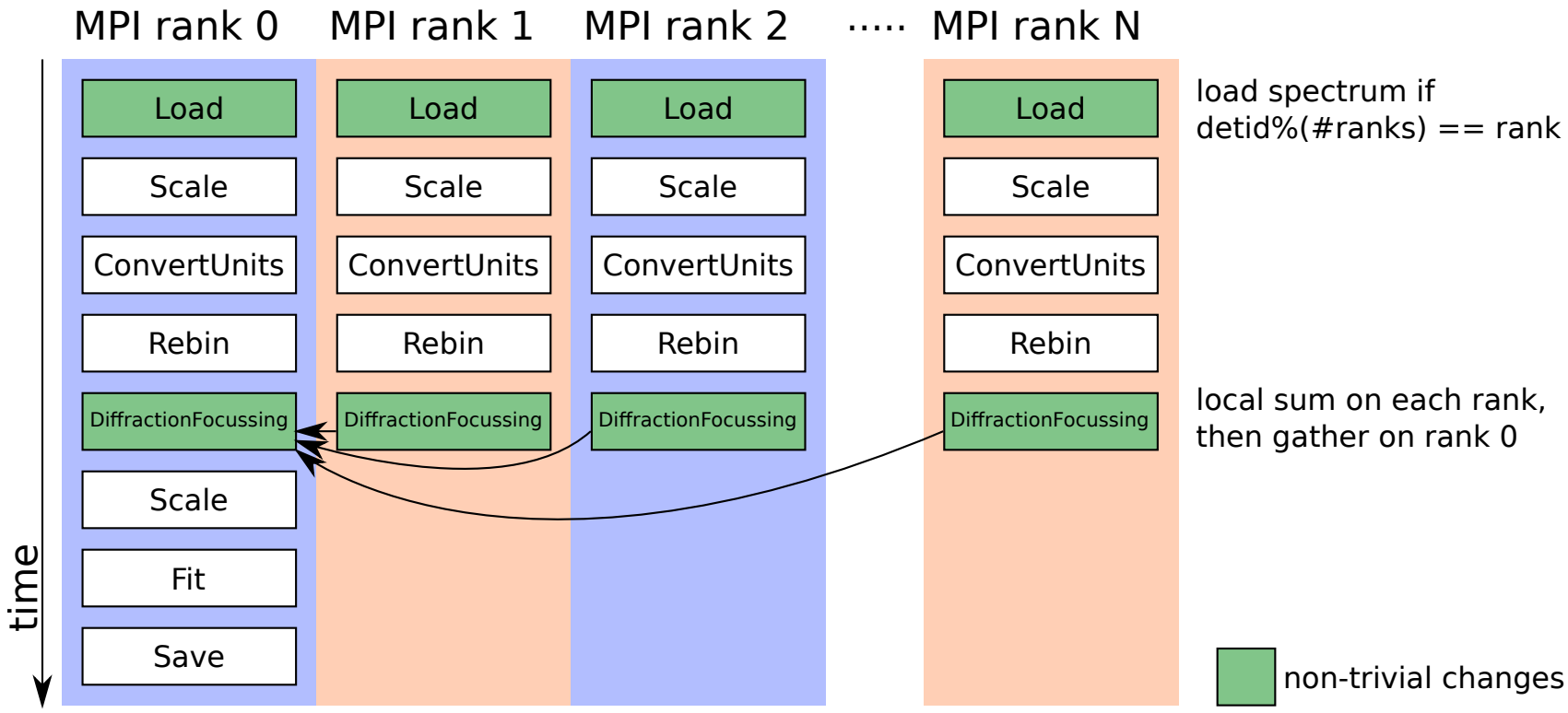
MPI-based data reduction in Mantid

Simon Heybrock

European Spallation Source

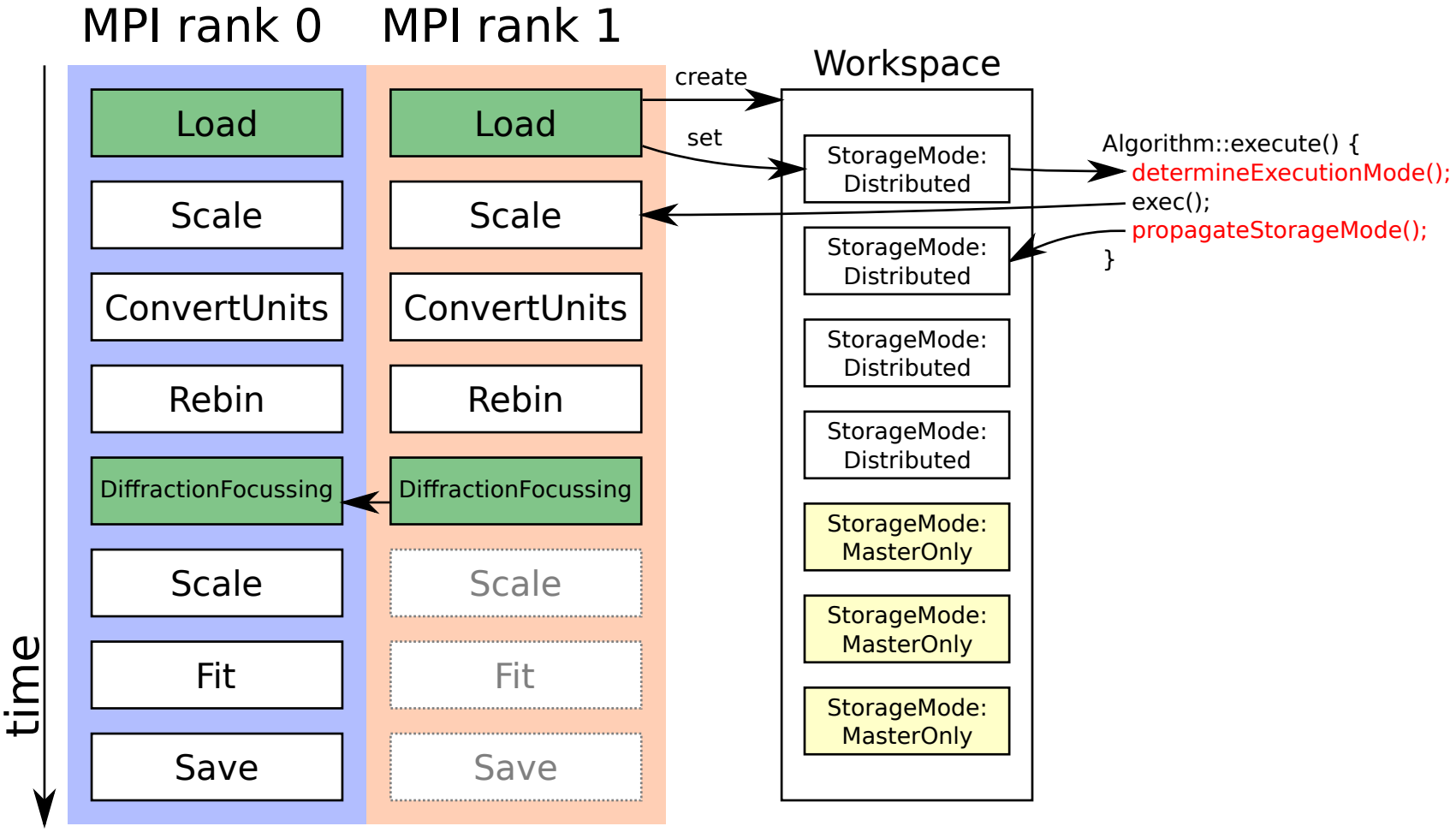
2015-09-24

MPI model based on by-spectrum work distribution



- detectors from all ranks are interleaved ⇒ good load balance
- gathering a workspace in general not feasible in event mode, but can probably be avoided in most scenarios

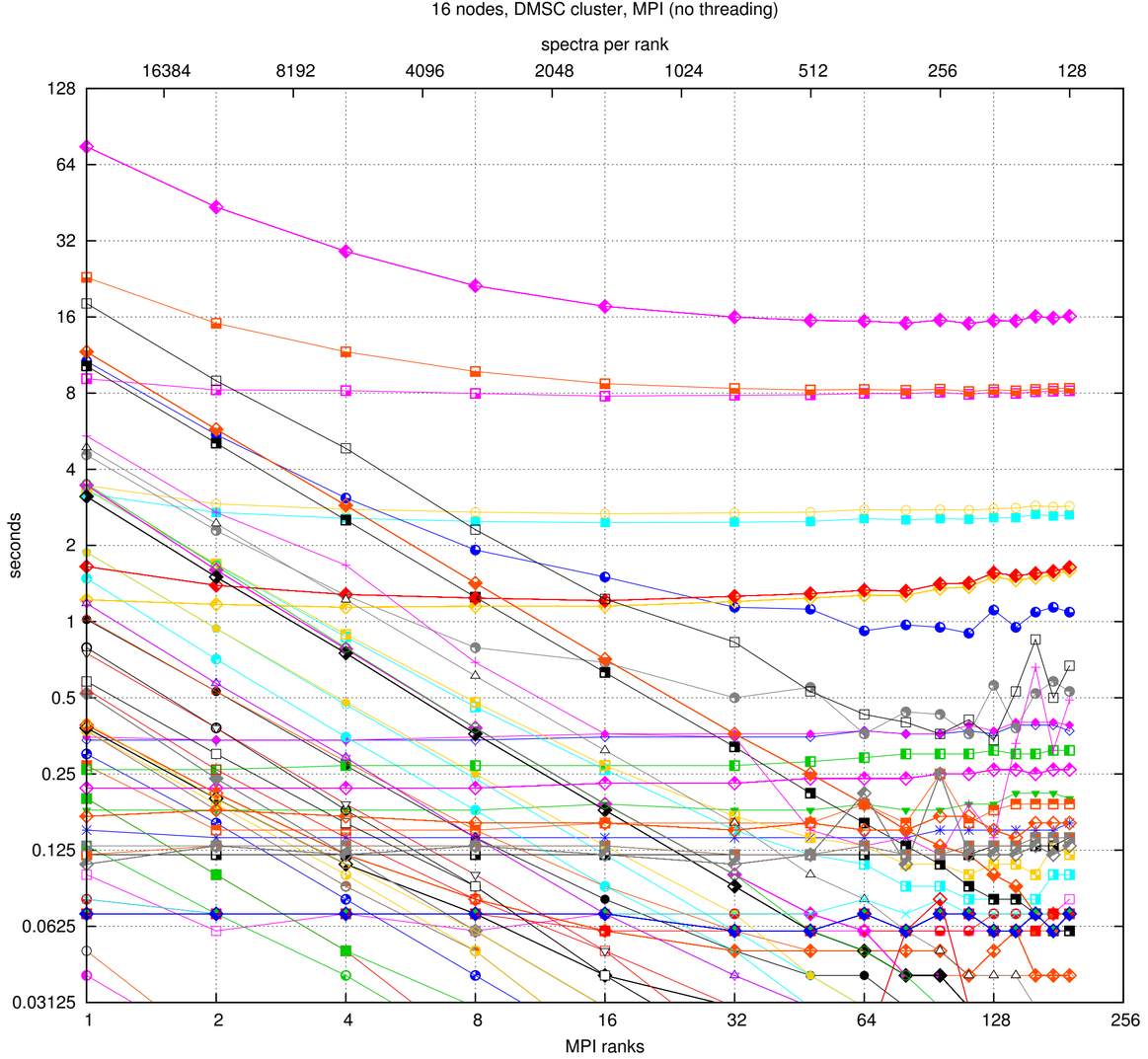
MPI model (implementation)



- `Workspace StorageMode` ⇒ `Algorithm ExecutionMode`
- default implementation for “trivially parallel algorithms” that can run in parallel out of the box

Test

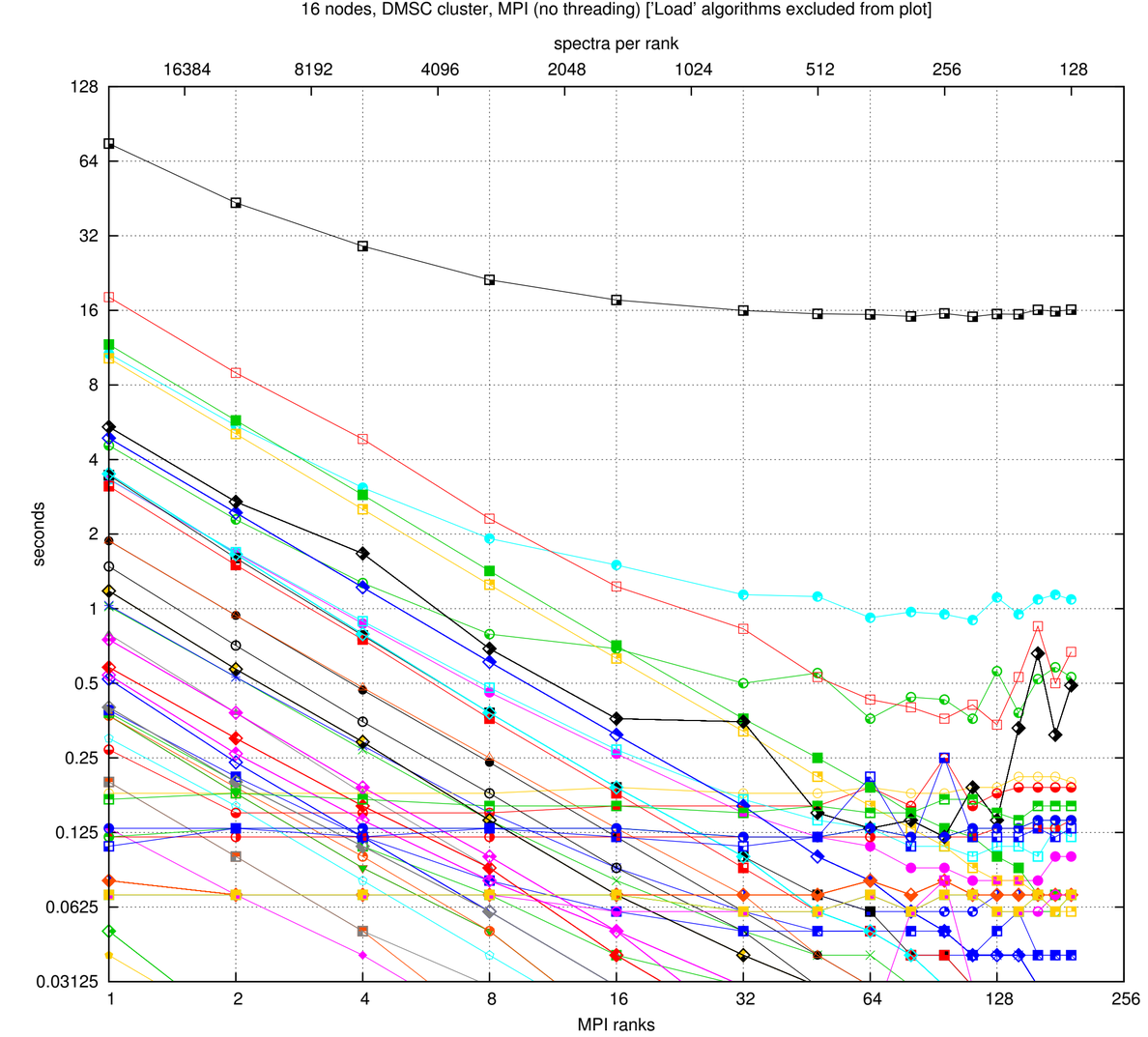
- `SNSPowderReduction`
  - disabled the existing MPI implementation
  - disabled `preserveEvents` in `DiffractionFocussing`
- test with POWGEN (PG3 system test)
  - data file with 17.9 M events
  - vanadium file with 65.8 M events
  - background file with 1.7 M events
- test runs on DMCS workstation and DMSC cluster (12 cores/node, Infiniband)
- MPI runs use only a single thread per rank



PDLoadCharacterizations
DetermineChunking
LoadNexusLogs
LoadInstrument
LoadEventNexus
Load
FilterByLogValue
FilterBadPulses
PDDetermineCharacterizations
SortTableWorkspace
ConvertDiffCal
LoadCalFile
LoadDiffCal
MaskBins
RemovePromptPulse
CompressEvents
ExtractSpectra
CropWorkspace
MaskDetectors
AlignDetectors
ConvertUnits
MaskDetectors
RemoveLowResTOF
ConvertUnits
Rebin
SortEvents
DiffractionFocussing
ConvertUnits
Rebin
AlignAndFocusPowder
RenameWorkspace
Multiply
NormaliseByCurrent
SetUncertainties
SaveGSS
DeleteWorkspace
DetermineChunking
LoadNexusLogs
LoadInstrument
LoadEventNexus
FilterByLogValue
FilterBadPulses
CompressEvents
LoadEventAndCompress

PDDetermineCharacterizations
Multiply
NormaliseByCurrent
DetermineChunking
LoadNexusLogs
LoadInstrument
LoadEventNexus
FilterByLogValue
FilterBadPulses
CompressEvents
LoadEventAndCompress
PDDetermineCharacterizations
DeleteWorkspace
CompressEvents
ConvertUnits
SetSampleMaterial
CompressEvents
MultipleScatteringCylinderAbsorption
MaskBins
RemovePromptPulse
CompressEvents
ExtractSpectra
CropWorkspace
MaskDetectors
AlignDetectors
ConvertUnits
MaskDetectors
RemoveLowResTOF
ConvertUnits
Rebin
SortEvents
DiffractionFocussing
ConvertUnits
Rebin
AlignAndFocusPowder
ConvertUnits
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit

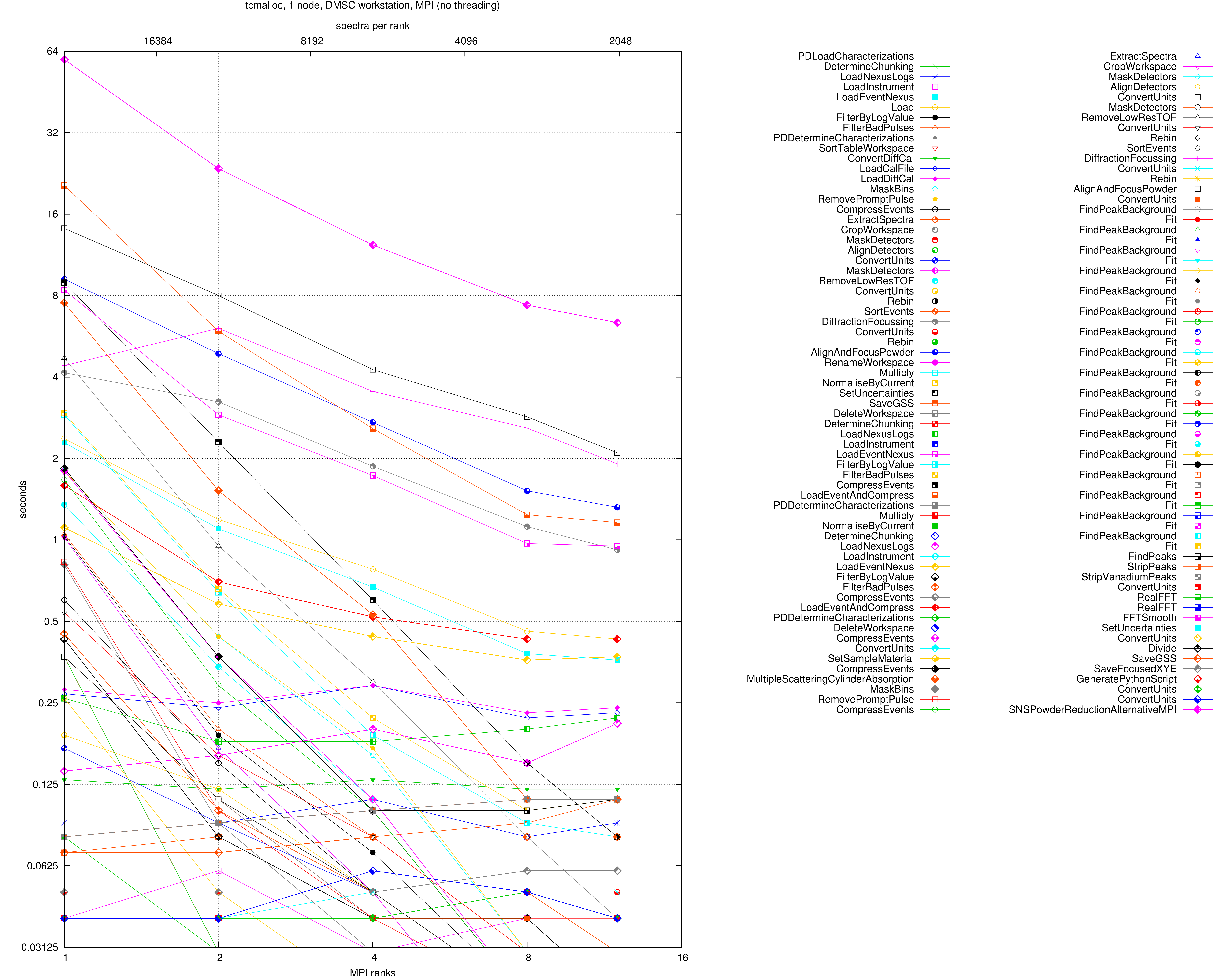
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeaks
StripPeaks
StripVanadiumPeaks
ConvertUnits
RealFFT
RealFFT
FFTSmooth
SetUncertainties
ConvertUnits
Divide
SaveGSS
SaveFocusedXYE
GeneratePythonScript
ConvertUnits
ConvertUnits
SNSPowderReductionAlternativeMPI



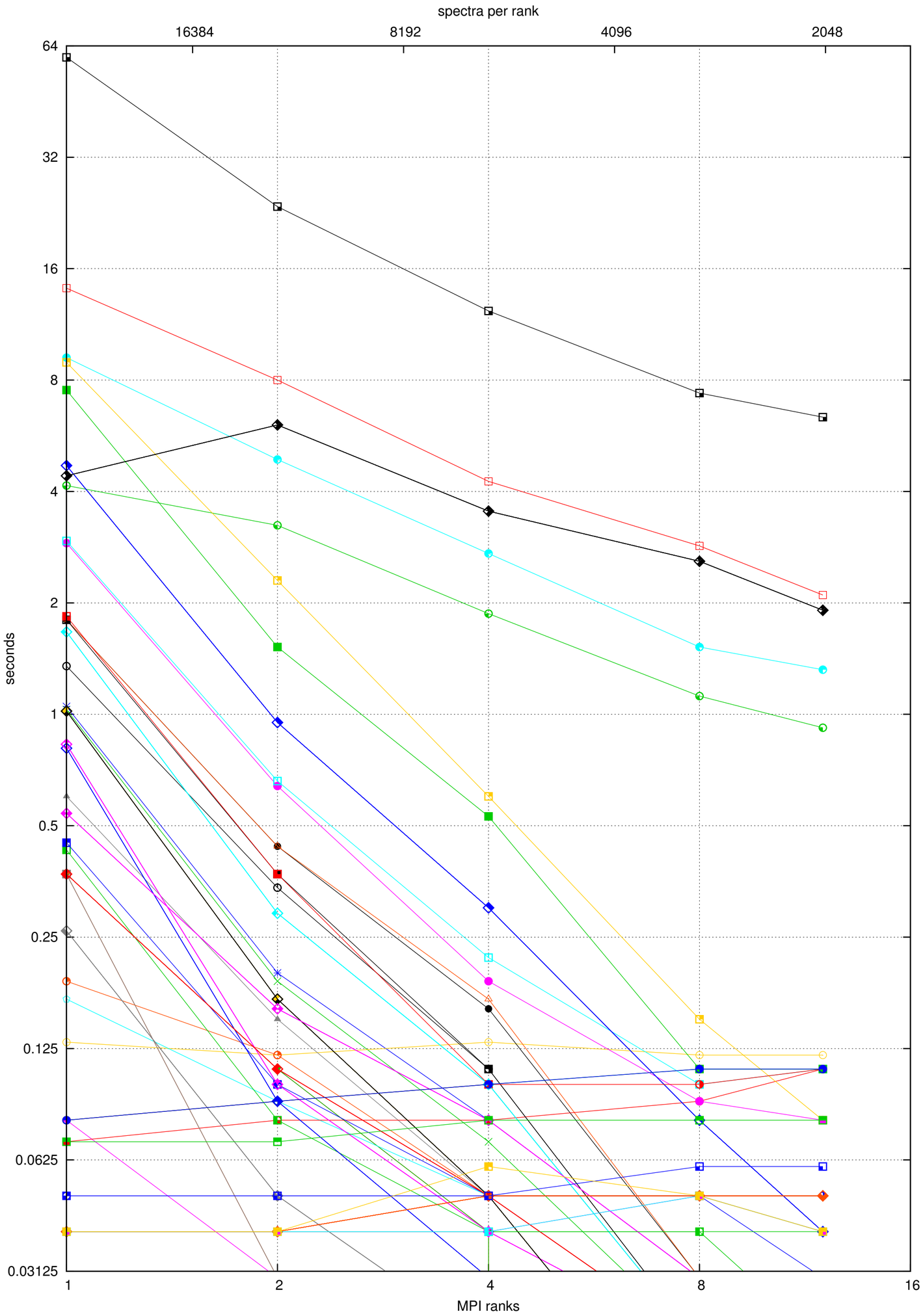
DetermineChunking
FilterByLogValue
FilterBadPulses
PDDetermineCharacterizations
SortTableWorkspace
ConvertDiffCal
MaskBins
RemovePromptPulse
CompressEvents
ExtractSpectra
CropWorkspace
MaskDetectors
AlignDetectors
ConvertUnits
MaskDetectors
RemoveLowResTOF
ConvertUnits
Rebin
SortEvents
DiffractionFocussing
ConvertUnits
Rebin
AlignAndFocusPowder
RenameWorkspace
Multiply
NormaliseByCurrent
SetUncertainties
SaveGSS
DeleteWorkspace
DetermineChunking
FilterByLogValue
FilterBadPulses
CompressEvents
PDDetermineCharacterizations
Multiply
NormaliseByCurrent
DetermineChunking
FilterByLogValue
FilterBadPulses

CompressEvents
PDDetermineCharacterizations
DeleteWorkspace
CompressEvents
ConvertUnits
SetSampleMaterial
CompressEvents
MultipleScatteringCylinderAbsorption
MaskBins
RemovePromptPulse
CompressEvents
ExtractSpectra
CropWorkspace
MaskDetectors
AlignDetectors
ConvertUnits
MaskDetectors
RemoveLowResTOF
ConvertUnits
Rebin
SortEvents
DiffractionFocussing
ConvertUnits
Rebin
AlignAndFocusPowder
ConvertUnits
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit

Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeaks
StripPeaks
StripVanadiumPeaks
ConvertUnits
RealFFT
RealFFT
FFTSmooth
SetUncertainties
ConvertUnits
Divide
SaveGSS
SaveFocusedXYE
GeneratePythonScript
ConvertUnits
ConvertUnits
SNSPowderReductionAlternativeMPI

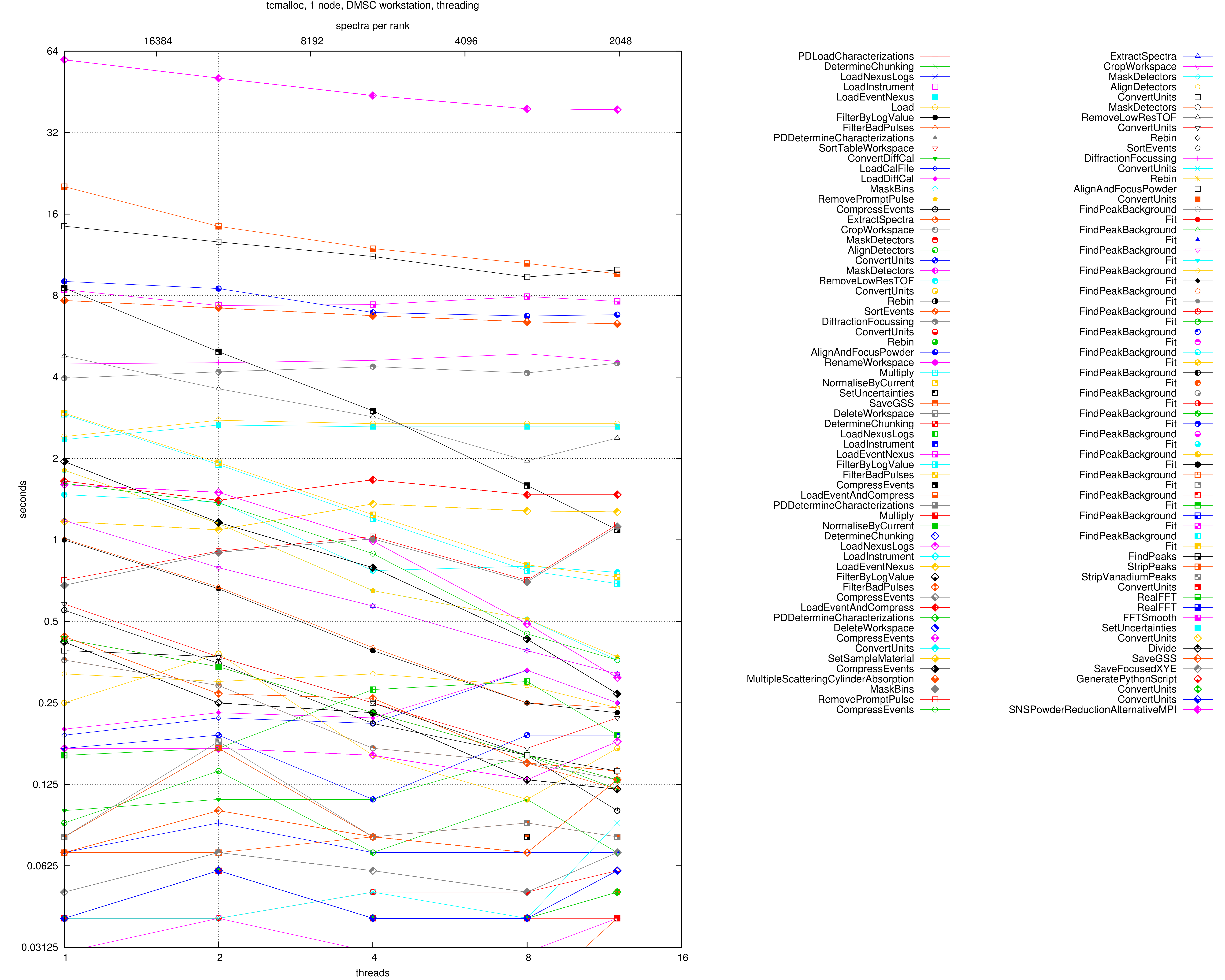

tcmalloc, 1 node, DMSC workstation, MPI (no threading)
spectra per rank
16384
8192
4096
2048
64
32
16
8
4
2
1
0.5
0.25
0.125
0.0625
0.03125
seconds
1
2
4
8
16
MPI ranks
PDLoadCharacterizations
DetermineChunking
LoadNexusLogs
LoadInstrument
LoadEventNexus
Load
FilterByLogValue
FilterBadPulses
PDDetermineCharacterizations
SortTableWorkspace
ConvertDiffCal
LoadCalFile
LoadDiffCal
MaskBins
RemovePromptPulse
CompressEvents
ExtractSpectra
CropWorkspace
MaskDetectors
AlignDetectors
ConvertUnits
MaskDetectors
RemoveLowResTOF
ConvertUnits
Rebin
SortEvents
DiffractionFocussing
ConvertUnits
Rebin
AlignAndFocusPowder
RenameWorkspace
Multiply
NormaliseByCurrent
SetUncertainties
SaveGSS
DeleteWorkspace
DetermineChunking
LoadNexusLogs
LoadInstrument
LoadEventNexus
FilterByLogValue
FilterBadPulses
CompressEvents
LoadEventAndCompress
PDDetermineCharacterizations
Multiply
NormaliseByCurrent
DetermineChunking
LoadNexusLogs
LoadInstrument
LoadEventNexus
FilterByLogValue
FilterBadPulses
CompressEvents
LoadEventAndCompress
PDDetermineCharacterizations
DeleteWorkspace
CompressEvents
ConvertUnits
SetSampleMaterial
CompressEvents
MultipleScatteringCylinderAbsorption
MaskBins
RemovePromptPulse
CompressEvents
ExtractSpectra
CropWorkspace
MaskDetectors
AlignDetectors
ConvertUnits
MaskDetectors
RemoveLowResTOF
ConvertUnits
Rebin
SortEvents
DiffractionFocussing
ConvertUnits
Rebin
AlignAndFocusPowder
ConvertUnits
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeaks
StripPeaks
StripVanadiumPeaks
ConvertUnits
RealFFT
RealFFT
FFTSmooth
SetUncertainties
ConvertUnits
Divide
SaveGSS
SaveFocusedXYE
GeneratePythonScript
ConvertUnits
ConvertUnits
SNSPowderReductionAlternativeMPI

tcmalloc, 1 node, DMSC workstation, MPI (no threading) ['Load' algorithms excluded from plot]



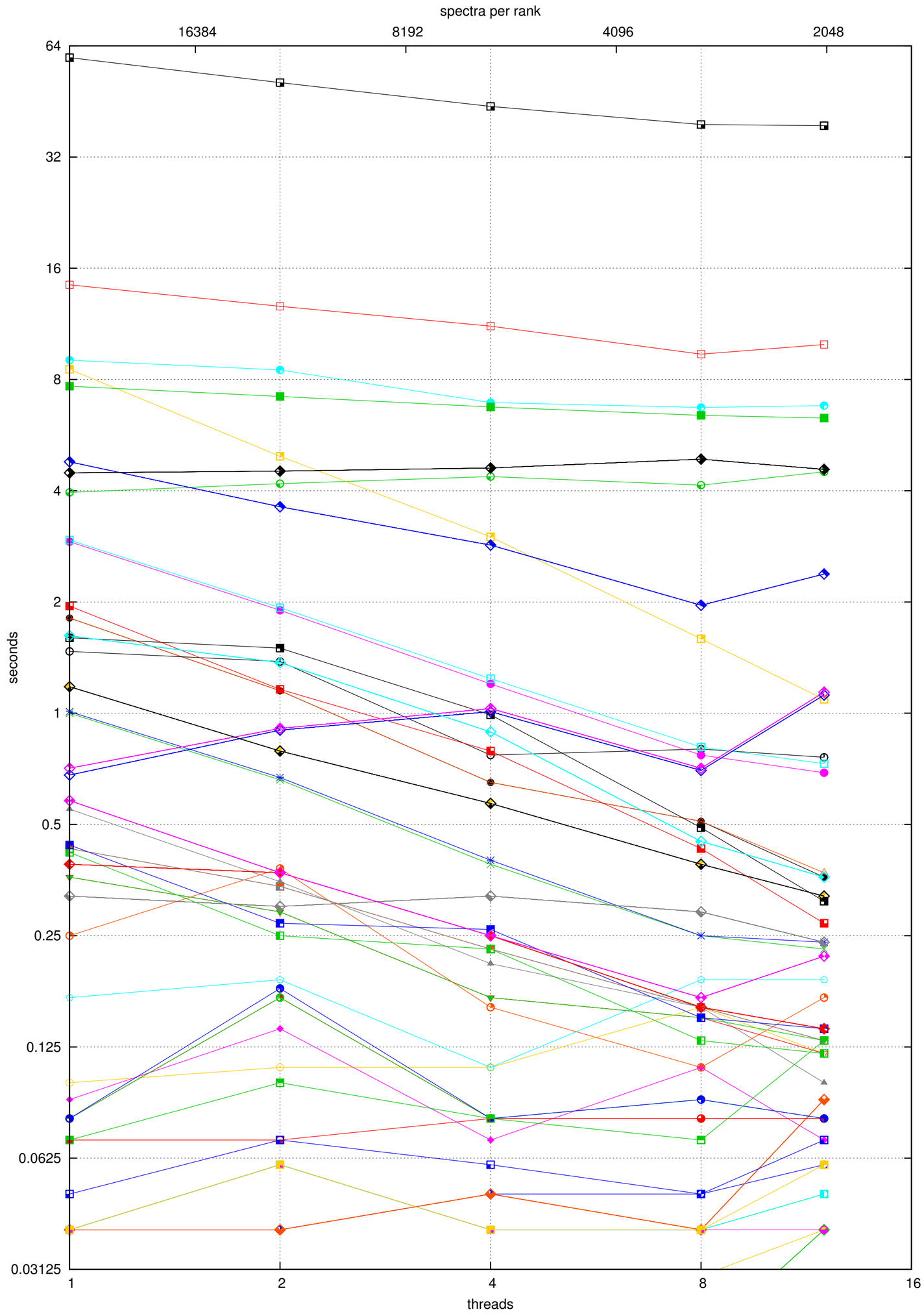
DetermineChunking
FilterByLogValue
FilterBadPulses
PDDetermineCharacterizations
SortTableWorkspace
ConvertDiffCal
MaskBins
RemovePromptPulse
CompressEvents
ExtractSpectra
CropWorkspace
MaskDetectors
AlignDetectors
ConvertUnits
MaskDetectors
RemoveLowResTOF
ConvertUnits
Rebin
SortEvents
DiffractionFocussing
ConvertUnits
Rebin
AlignAndFocusPowder
RenameWorkspace
Multiply
NormaliseByCurrent
SetUncertainties
SaveGSS
DeleteWorkspace
DetermineChunking
FilterByLogValue
FilterBadPulses
CompressEvents
PDDetermineCharacterizations
Multiply
NormaliseByCurrent
DetermineChunking
FilterByLogValue
FilterBadPulses
CompressEvents
PDDetermineCharacterizations
DeleteWorkspace
CompressEvents
ConvertUnits
SetSampleMaterial
CompressEvents
MultipleScatteringCylinderAbsorption
MaskBins
RemovePromptPulse
CompressEvents
ExtractSpectra
CropWorkspace
MaskDetectors
AlignDetectors
ConvertUnits
MaskDetectors
RemoveLowResTOF
ConvertUnits

Rebin
SortEvents
DiffractionFocussing
ConvertUnits
Rebin
AlignAndFocusPowder
ConvertUnits
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeaks
StripPeaks
StripVanadiumPeaks
ConvertUnits
RealFFT
RealFFT
FFTSmooth
SetUncertainties
ConvertUnits
Divide
SaveGSS
SaveFocusedXYE
GeneratePythonScript
ConvertUnits
ConvertUnits
SNSPowderReductionAlternativeMPI



tcmalloc, 1 node, DMSC workstation, threading
spectra per rank
16384
8192
4096
2048
64
32
16
8
4
2
1
0.5
0.25
0.125
0.0625
0.03125
seconds
1
2
4
8
16
threads
PDLoadCharacterizations
DetermineChunking
LoadNexusLogs
LoadInstrument
LoadEventNexus
Load
FilterByLogValue
FilterBadPulses
PDDetermineCharacterizations
SortTableWorkspace
ConvertDiffCal
LoadCalFile
LoadDiffCal
MaskBins
RemovePromptPulse
CompressEvents
ExtractSpectra
CropWorkspace
MaskDetectors
AlignDetectors
ConvertUnits
MaskDetectors
RemoveLowResTOF
ConvertUnits
Rebin
SortEvents
DiffractionFocussing
ConvertUnits
Rebin
AlignAndFocusPowder
RenameWorkspace
Multiply
NormaliseByCurrent
SetUncertainties
SaveGSS
DeleteWorkspace
DetermineChunking
LoadNexusLogs
LoadInstrument
LoadEventNexus
FilterByLogValue
FilterBadPulses
CompressEvents
LoadEventAndCompress
PDDetermineCharacterizations
Multiply
NormaliseByCurrent
DetermineChunking
LoadNexusLogs
LoadInstrument
LoadEventNexus
FilterByLogValue
FilterBadPulses
CompressEvents
LoadEventAndCompress
PDDetermineCharacterizations
DeleteWorkspace
CompressEvents
ConvertUnits
SetSampleMaterial
CompressEvents
MultipleScatteringCylinderAbsorption
MaskBins
RemovePromptPulse
CompressEvents
ExtractSpectra
CropWorkspace
MaskDetectors
AlignDetectors
ConvertUnits
MaskDetectors
RemoveLowResTOF
ConvertUnits
Rebin
SortEvents
DiffractionFocussing
ConvertUnits
Rebin
AlignAndFocusPowder
ConvertUnits
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeaks
StripPeaks
StripVanadiumPeaks
ConvertUnits
RealFFT
RealFFT
FFTSmooth
SetUncertainties
ConvertUnits
Divide
SaveGSS
SaveFocusedXYE
GeneratePythonScript
ConvertUnits
ConvertUnits
SNSPowderReductionAlternativeMPI

tcmalloc, 1 node, DMSC workstation, threading ['Load' algorithms excluded from plot]



DetermineChunking
FilterByLogValue
FilterBadPulses
PDDetermineCharacterizations
SortTableWorkspace
ConvertDiffCal
MaskBins
RemovePromptPulse
CompressEvents
ExtractSpectra
CropWorkspace
MaskDetectors
AlignDetectors
ConvertUnits
MaskDetectors
RemoveLowResTOF
ConvertUnits
Rebin
SortEvents
DiffractionFocussing
ConvertUnits
Rebin
AlignAndFocusPowder
RenameWorkspace
Multiply
NormaliseByCurrent
SetUncertainties
SaveGSS
DeleteWorkspace
DetermineChunking
FilterByLogValue
FilterBadPulses
CompressEvents
PDDetermineCharacterizations
Multiply
NormaliseByCurrent
DetermineChunking
FilterByLogValue
FilterBadPulses
CompressEvents
PDDetermineCharacterizations
DeleteWorkspace
CompressEvents
ConvertUnits
SetSampleMaterial
CompressEvents
MultipleScatteringCylinderAbsorption
MaskBins
RemovePromptPulse
CompressEvents
ExtractSpectra
CropWorkspace
MaskDetectors
AlignDetectors
ConvertUnits
MaskDetectors
RemoveLowResTOF
ConvertUnits

Rebin
SortEvents
DiffractionFocussing
ConvertUnits
Rebin
AlignAndFocusPowder
ConvertUnits
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeakBackground
Fit
FindPeaks
StripPeaks
StripVanadiumPeaks
ConvertUnits
RealFFT
RealFFT
FFTSmooth
SetUncertainties
ConvertUnits
Divide
SaveGSS
SaveFocusedXYE
GeneratePythonScript
ConvertUnits
ConvertUnits
SNSPowderReductionAlternativeMPI