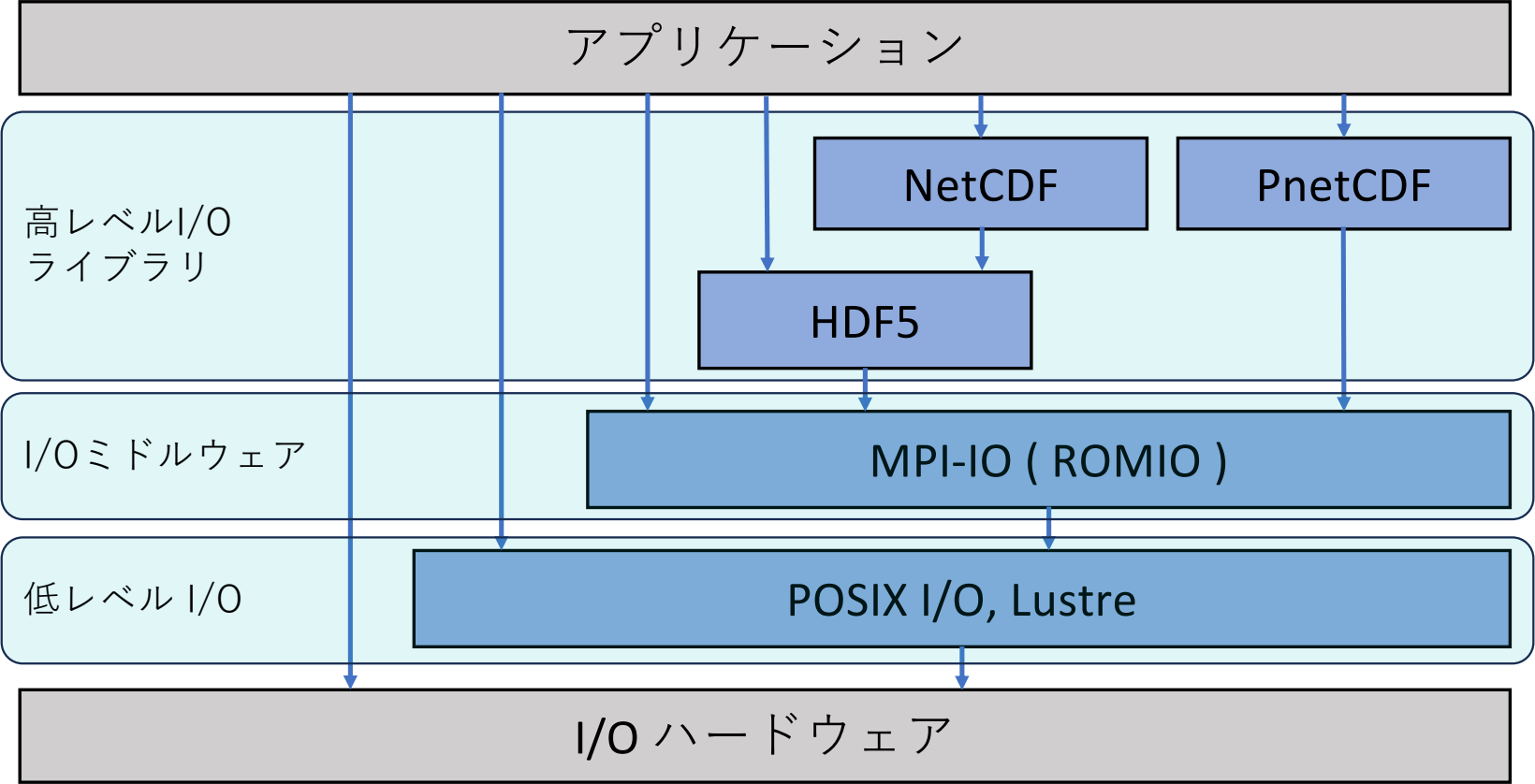
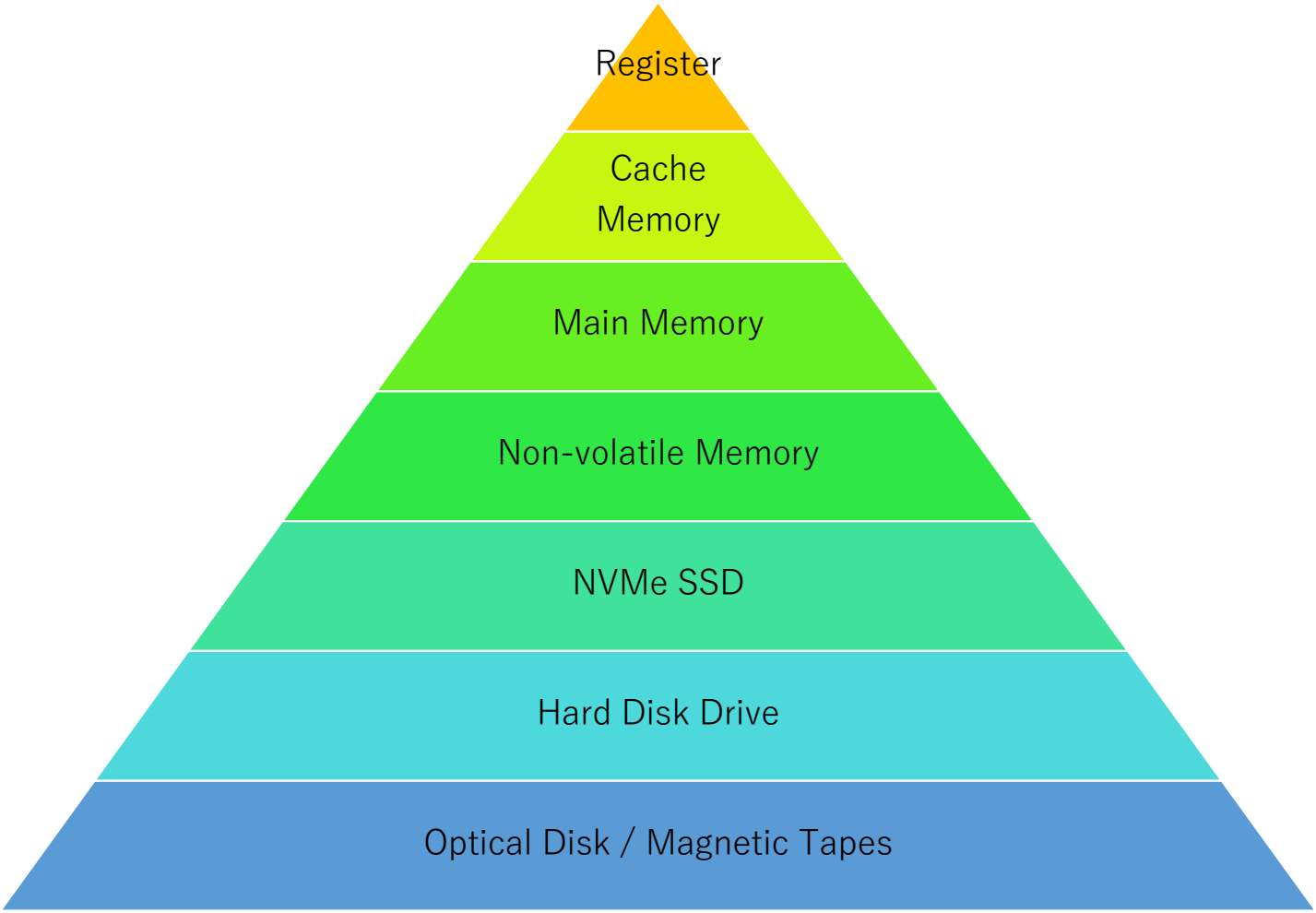
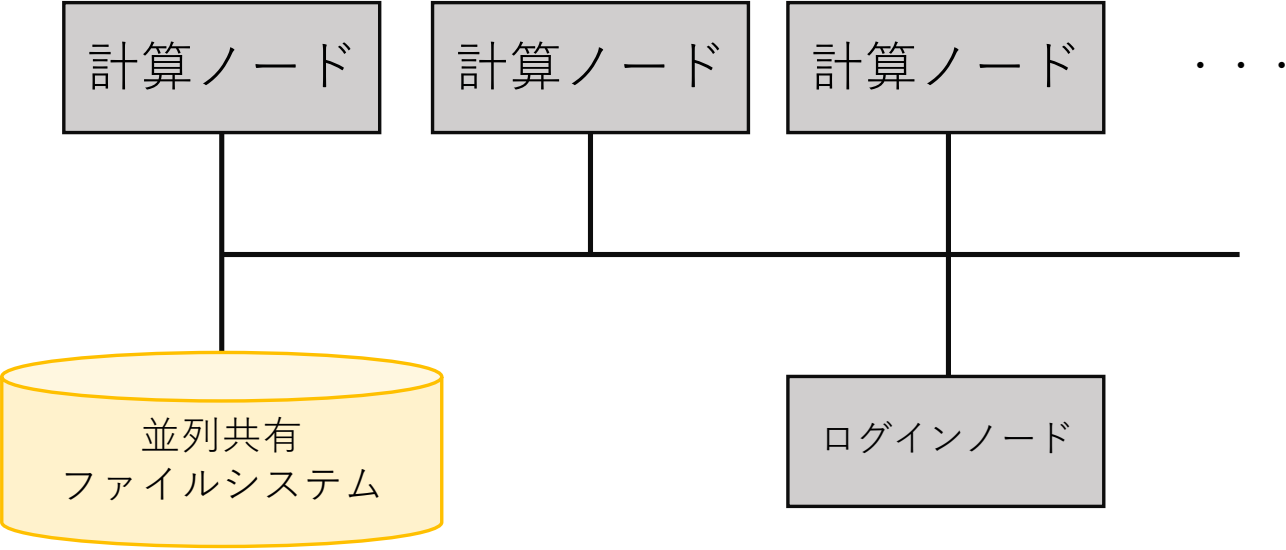
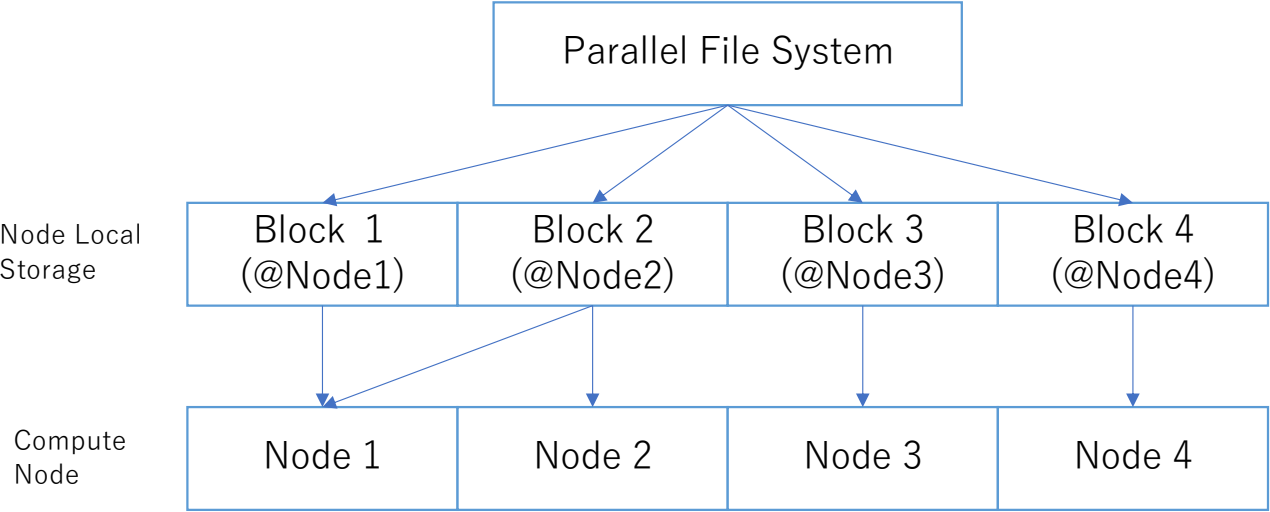
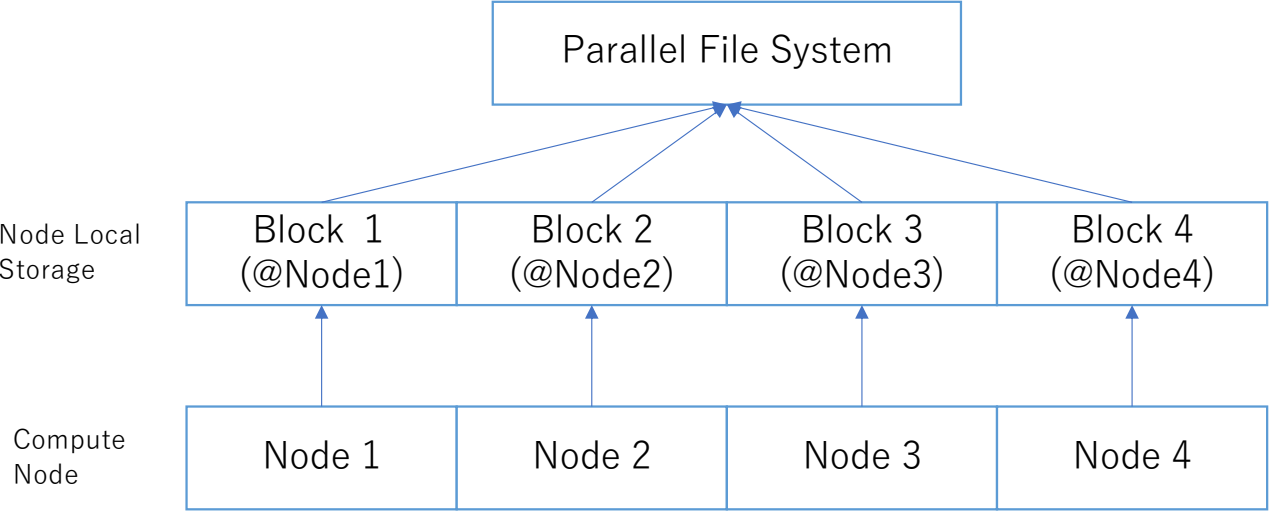
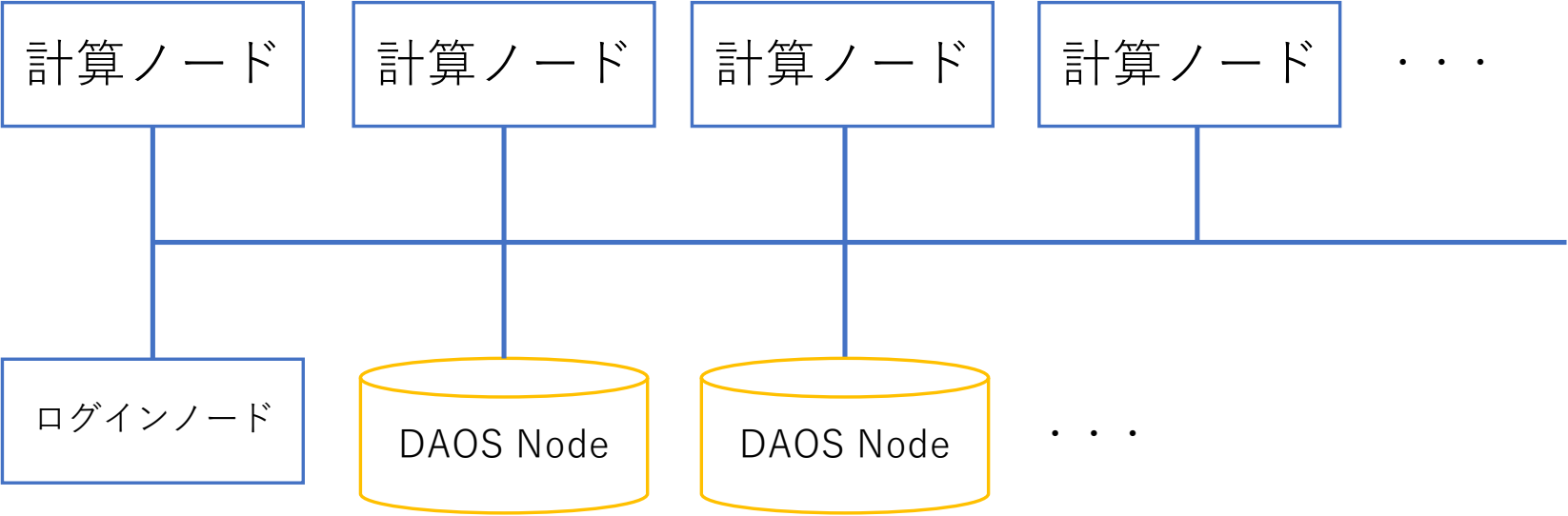

アプリケーション
高レベルI/O
ライブラリ
NetCDF
PnetCDF
HDF5
I/Oミドルウェア
MPI-IO (ROMIO)
低レベル I/O
POSIX I/O, Lustre
I/O ハードウェア

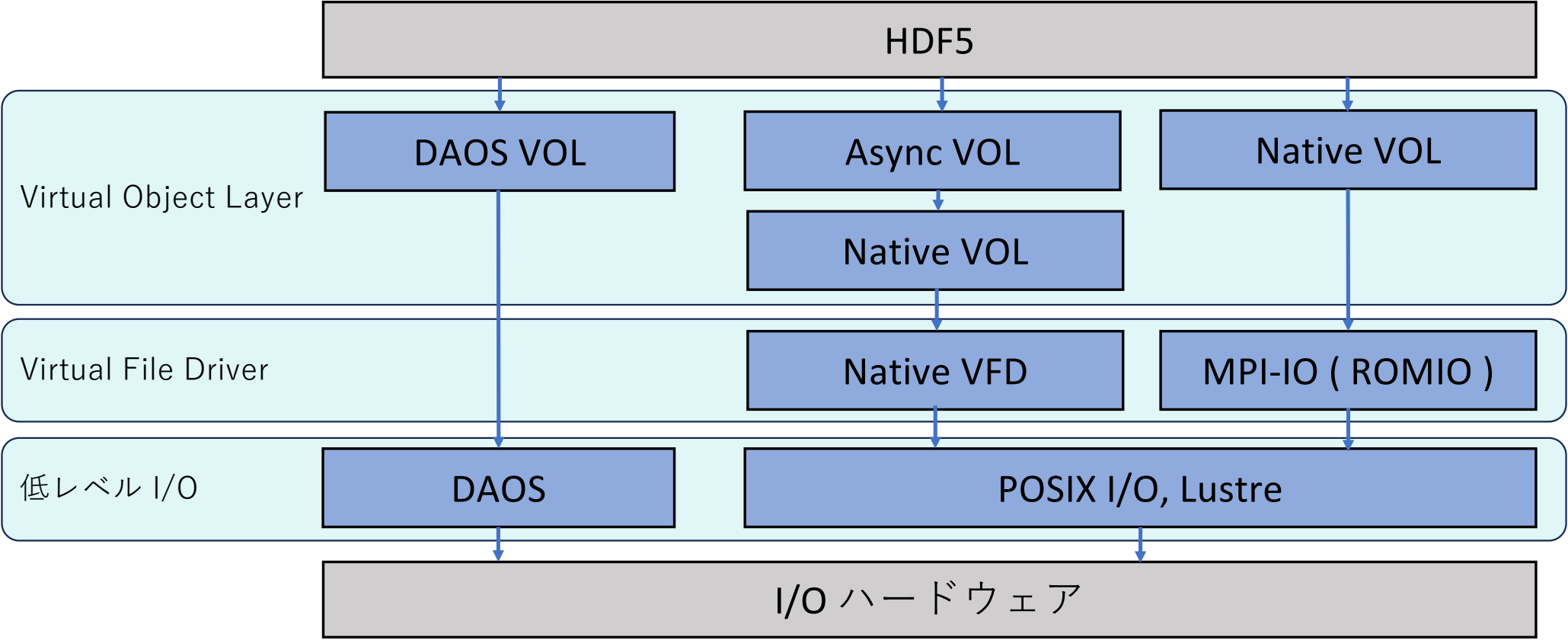

Register
Cache Memory
Main Memory
Non-volatile Memory
NVMe SSD
Hard Disk Drive
Optical Disk / Magnetic Tapes

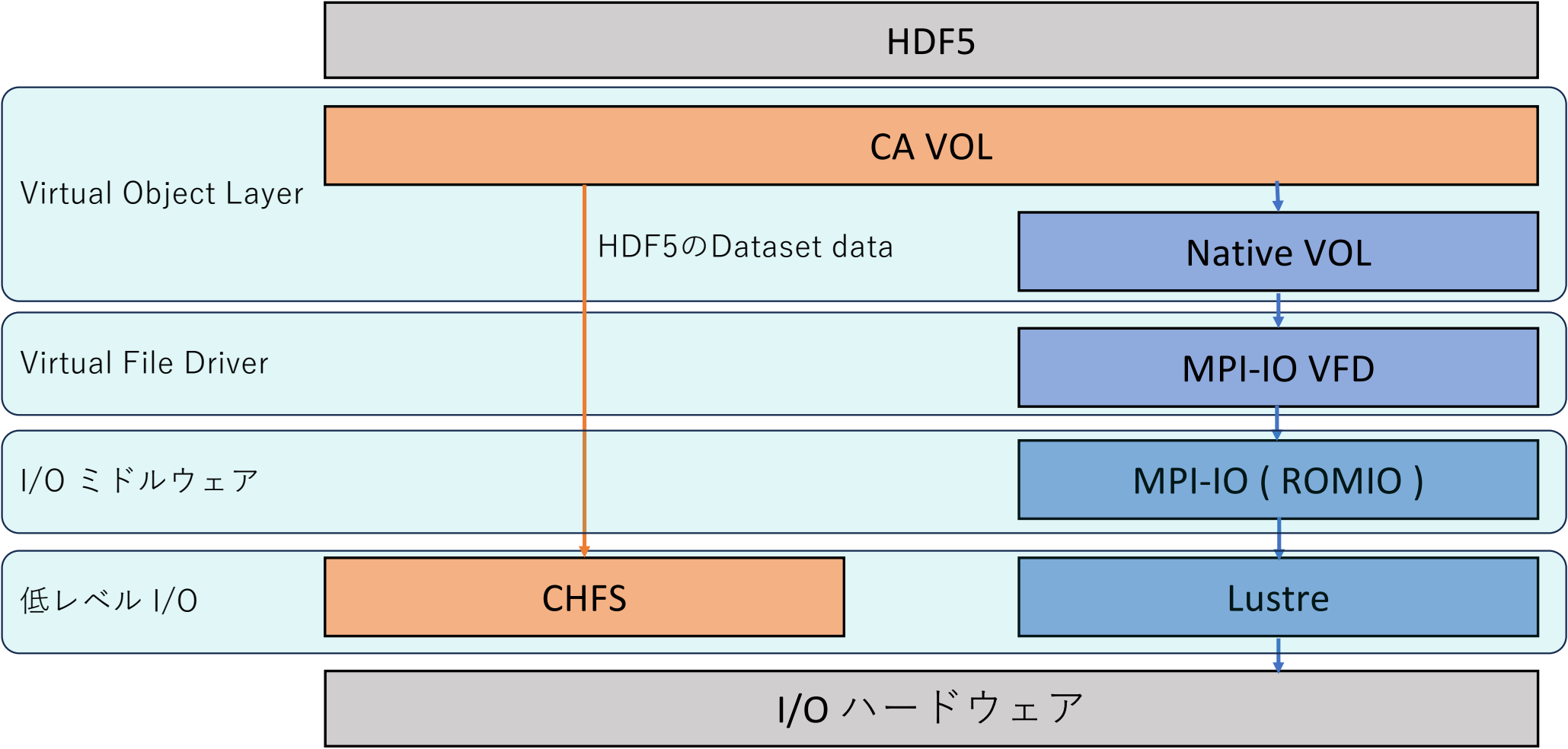

計算ノード
計算ノード
計算ノード
・・・
並列共有
ファイルシステム
ログインノード

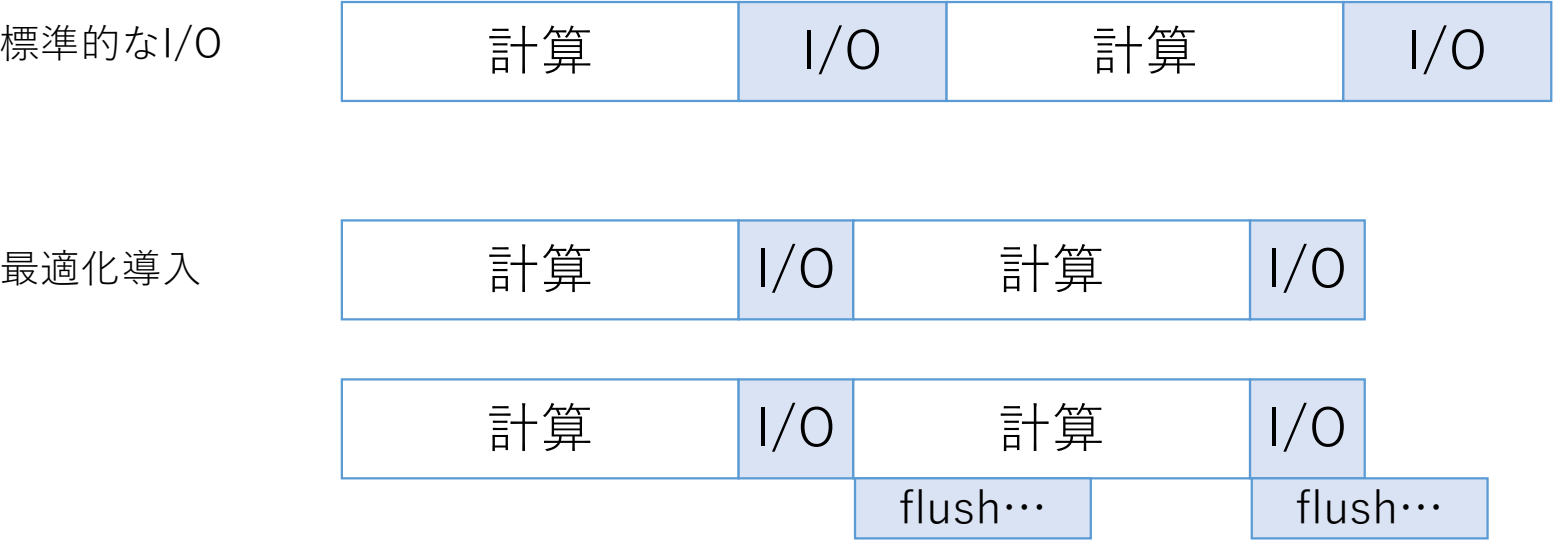

Parallel File System
Node Local Storage
Block 1 (@Node1)
Block 2 (@Node2)
Block 3 (@Node3)
Block 4 (@Node4)
Compute Node
Node 1
Node 2
Node 3
Node 4

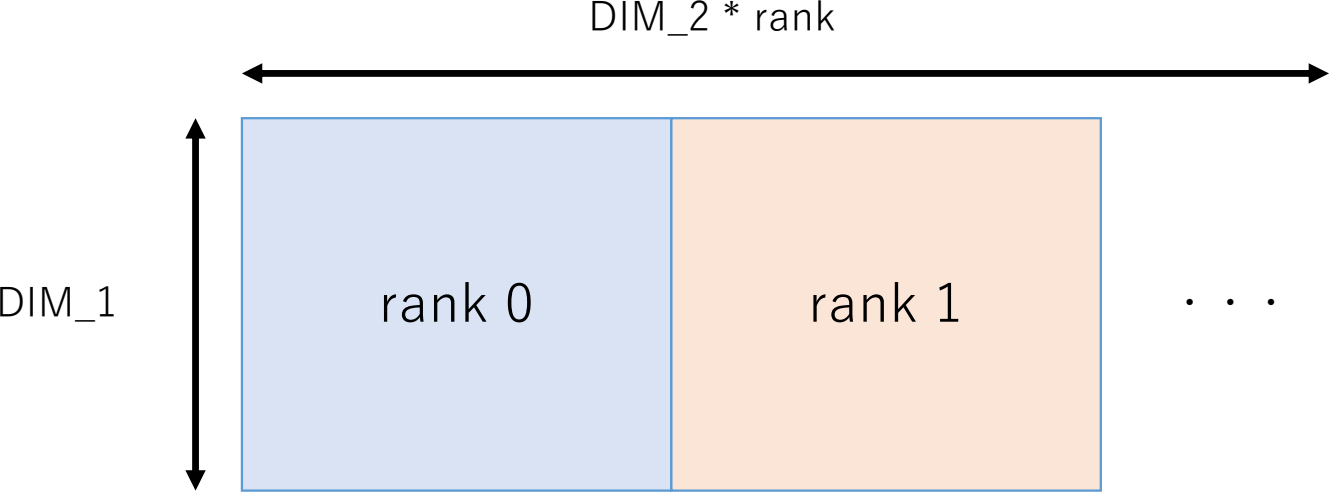

Parallel File System
Node Local Storage
Block 1 (@Node1)
Block 2 (@Node2)
Block 3 (@Node3)
Block 4 (@Node4)
Compute Node
Node 1
Node 2
Node 3
Node 4

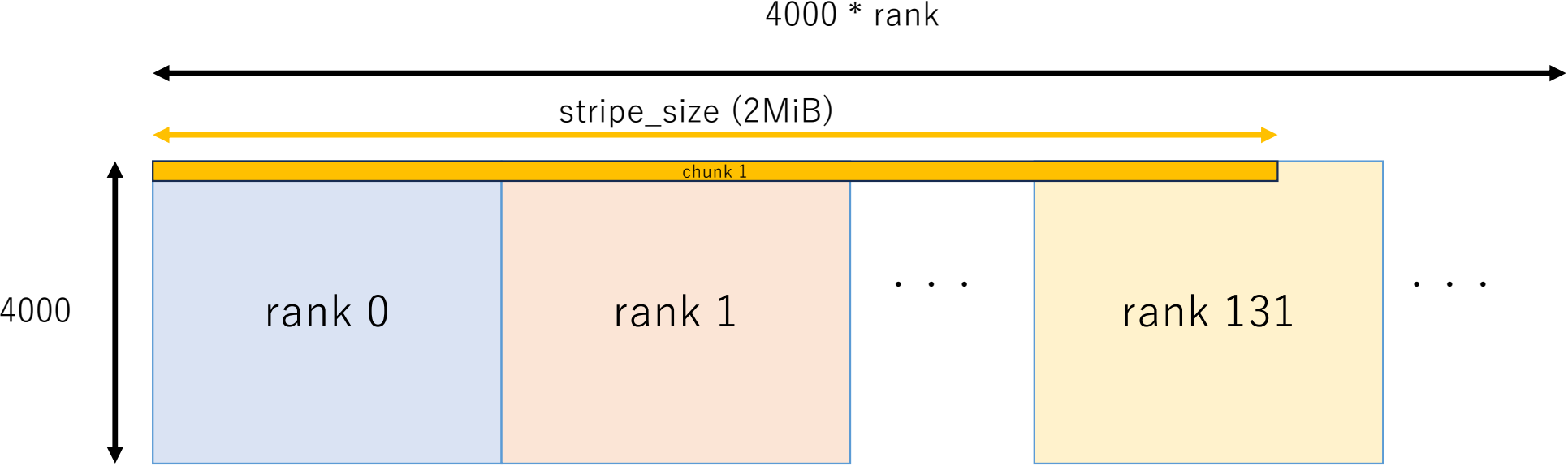

計算ノード
計算ノード
計算ノード
計算ノード
・・・
ログインノード
DAOS Node
DAOS Node
・・・

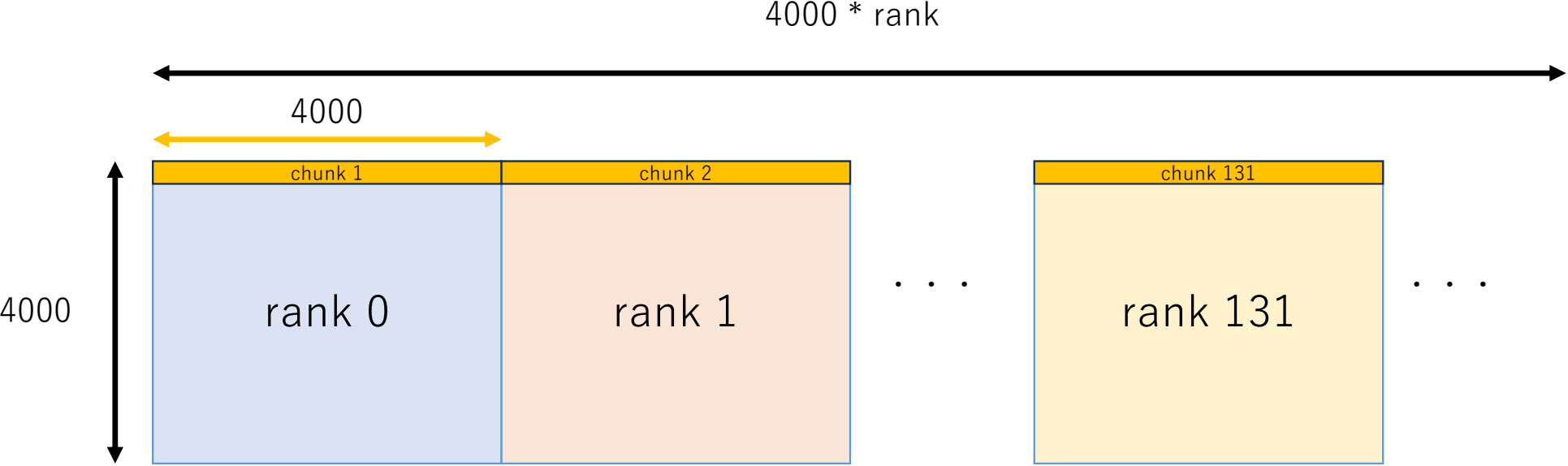

HDF5
Virtual Object Layer
DAOS VOL
Async VOL
Native VOL
Native VOL
Virtual File Driver
Native VFD
MPI-IO (ROMIO)
低レベル I/O
DAOS
POSIX I/O, Lustre
I/O ハードウェア


HDF5
Virtual Object Layer
CA VOL
HDF5のDataset data
Native VOL
Virtual File Driver
MPI-IO VFD
I/O ミドルウェア
MPI-IO (ROMIO)
低レベル I/O
CHFS
Lustre
I/O ハードウェア

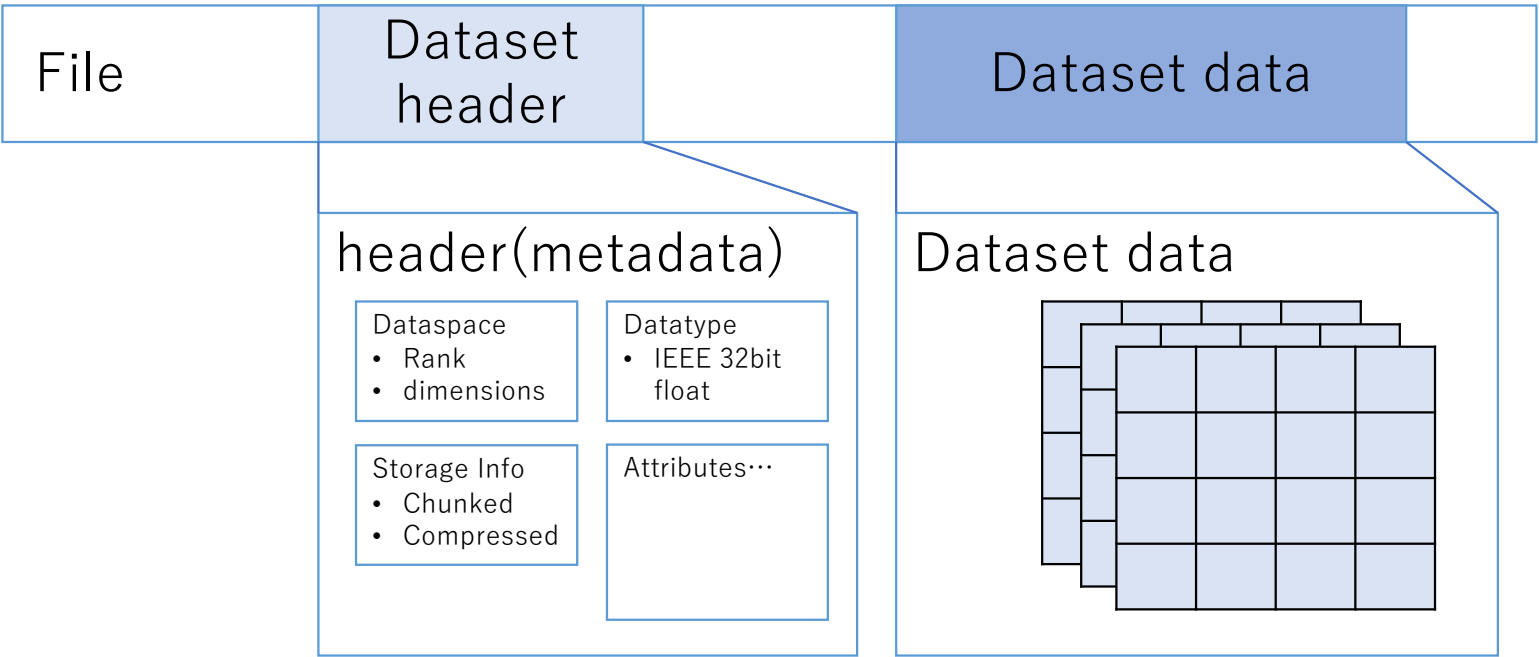

標準的なI/O
計算
I/O
計算
I/O
最適化導入
計算
I/O
計算
I/O
計算
I/O
計算
I/O
flush…
flush…

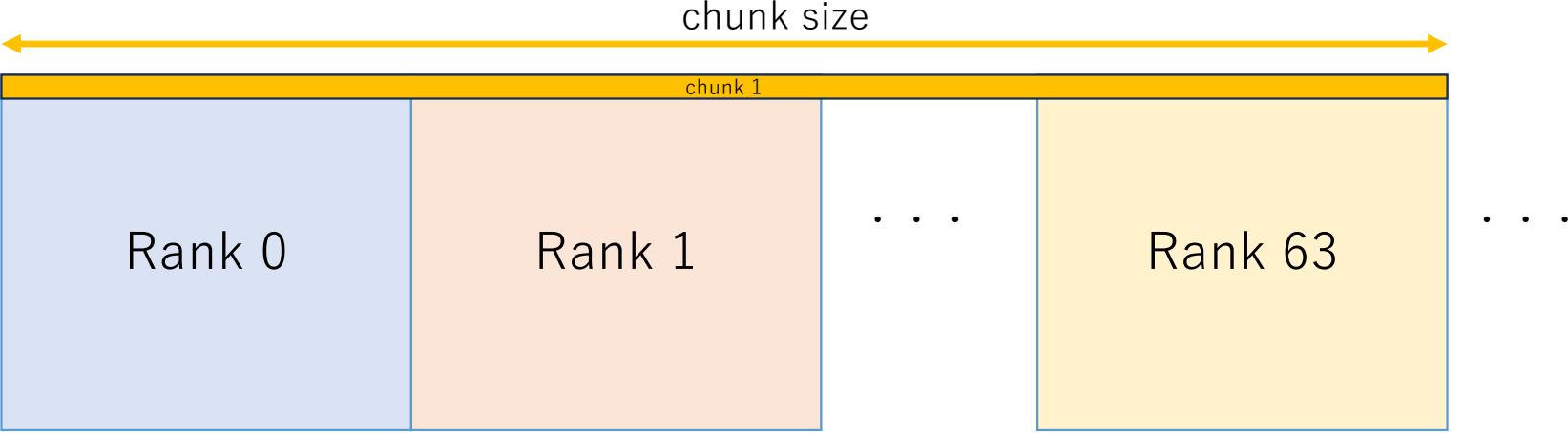

DIM_2 * rank
DIM_1
rank 0
rank 1
. . .

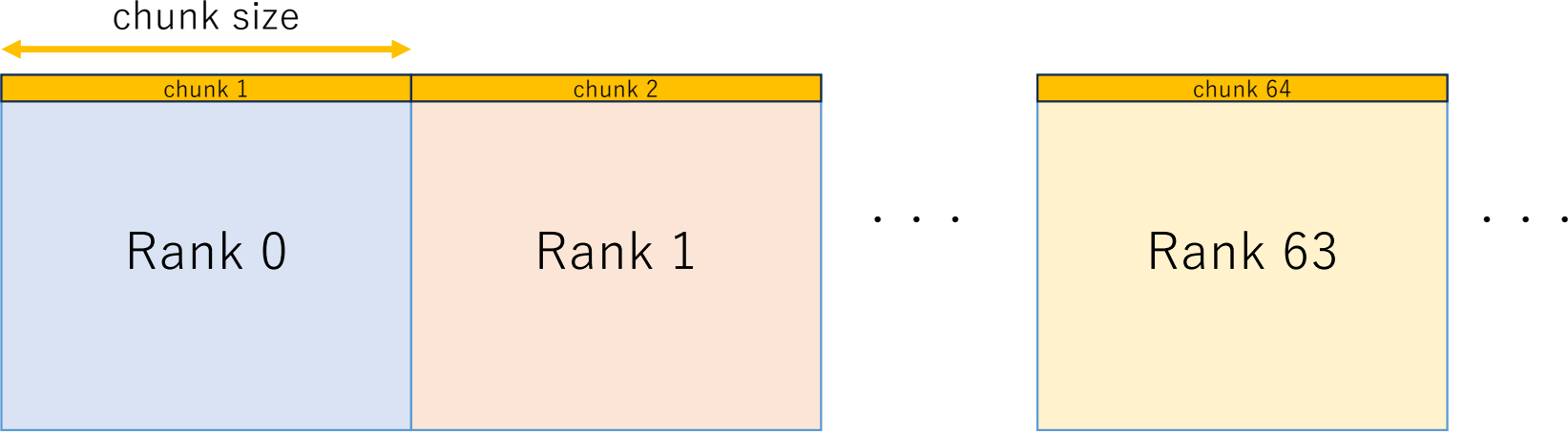

4000 * rank
stripe_size (2MiB)
chunk 1
4000
rank 0
rank 1
. . .
rank 131
. . .

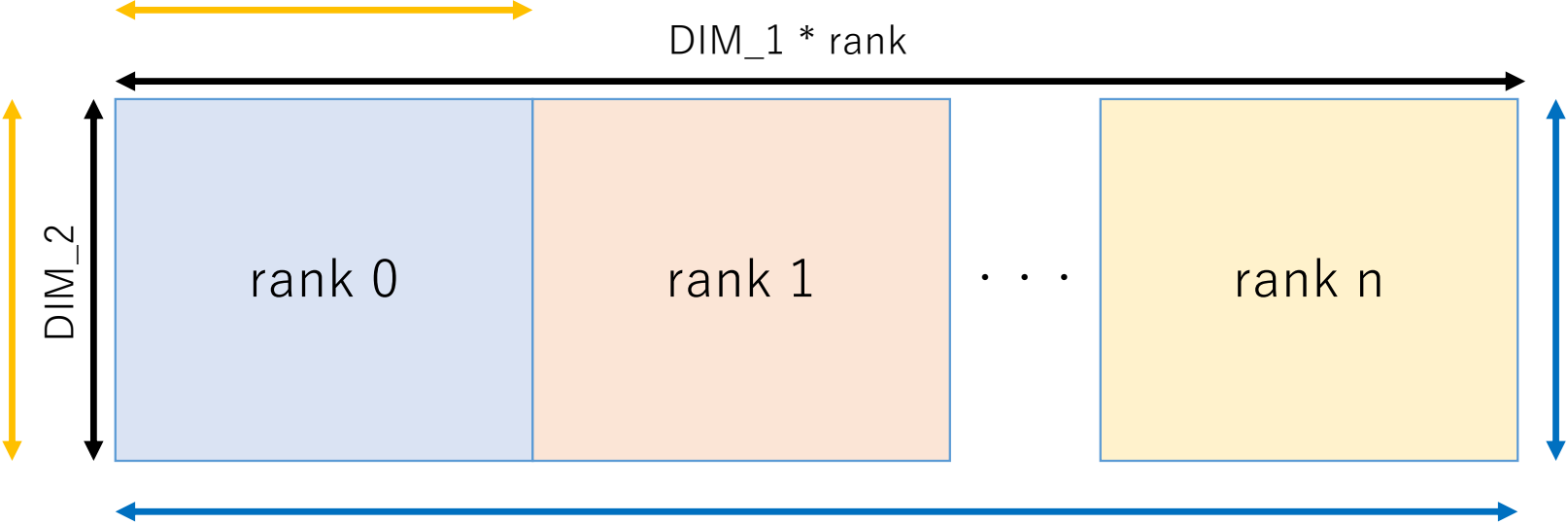

4000 * rank
4000
chunk 1
chunk 2
chunk 131
4000
rank 0
rank 1
rank 131

HDF5 File Layout

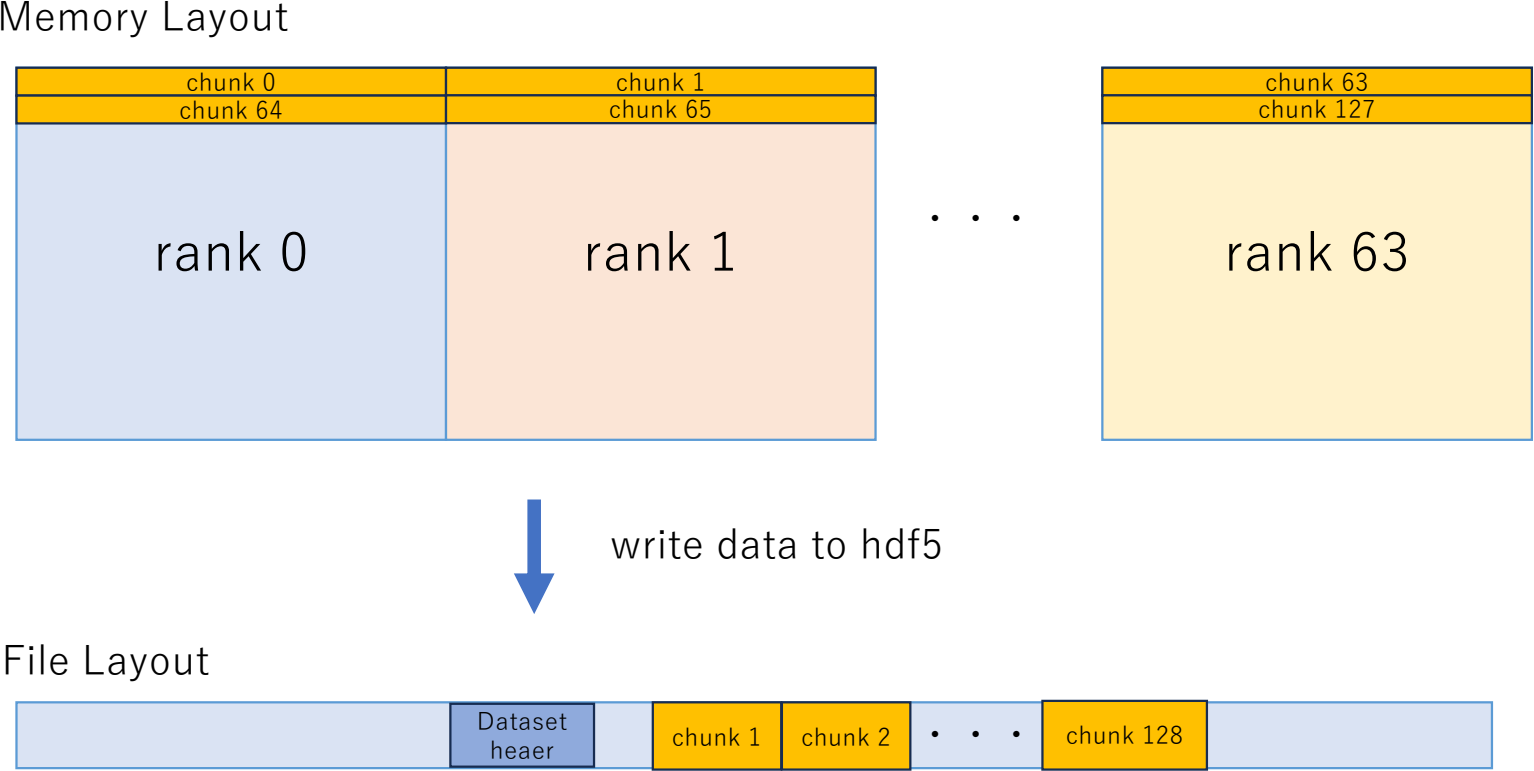



chunk size
chunk 1
Rank 0
Rank 1
Rank 63


chunk size
chunk 1
chunk 2
chunk 64
Rank 0
Rank 1
Rank 63



H5Sget_select_bounds() 全体のレイアウトのうちどこか
H5Sget_simple_extent_dims() 全体のレイアウトの大きさ
H5Sget_simple_extent_ndims() 次元数
H5Tget_size() 格納している値のサイズ


Memory Layout
chunk 0
chunk 64
chunk 1
chunk 65
chunk 63
chunk 127
rank 0
rank 1
rank 63
write data to hdf5
File Layout
Dataset heaer
chunk 1
chunk 2
chunk 128