

Contents

1	Input/Output and Postprocessing	1
2	Parallel I/O Basics	3
2.1	The I/O Software Stack	3
2.2	Parallel File Systems	5
2.3	MPI-IO	7
2.4	Data Format Libraries	8
2.5	Applying I/O Lessons to Applications	10
2.6	Parallel I/O Today and Tomorrow	11
	References	12
3	ESM I/O Layers	13
3.1	ESM I/O Layers: Design Considerations	13
3.2	An ESM Parallel I/O Class Library	14
3.2.1	mpp_io	15
3.2.2	Other ESM Parallel I/O Libraries	17
3.3	ESM Post-Processing Tools and Libraries	18
3.4	Discussion	19
	References	20
4	Data Storage	21
4.1	Hierarchical Storage Management: An Introduction	21
4.2	An Abstraction for HSMs: A Three-Level Storage Model	22
4.3	Discussion: HSM Strategies	24
	References	24
5	Data Representation	25
5.1	Scientific Data File Formats	25
5.1.1	NetCDF	26
5.1.2	HDF	26
5.1.3	GRIB	28

5.2	Remote Data Access	28
5.2.1	Bulk Access	29
5.2.2	Remote Subsetting in Index Space	29
5.2.3	Remote Subsetting in Coordinate Space	30
5.3	Georeferencing Coordinate Systems and Metadata	31
5.3.1	CF Conventions	31
5.3.2	Other Georeferencing Conventions	32
5.3.3	Discovery Versus Use Metadata	32
5.4	Data Models	33
5.4.1	ISO/OGC Models	34
5.5	Data Aggregation	34
5.6	Case Study: Attribution and Detection	36
	References	37
6	Data Analysis and Visualization	39
6.1	Background and History	39
6.2	The Analysis Process	41
6.3	Current Platforms Used for Data Analysis and Visualization	43
6.4	Community Tools and Environments	44
6.5	Use Case Example	46
6.6	Future Challenges and Directions	47
7	Future Perspectives	49
	References	51
	Glossary	53
	Index	57