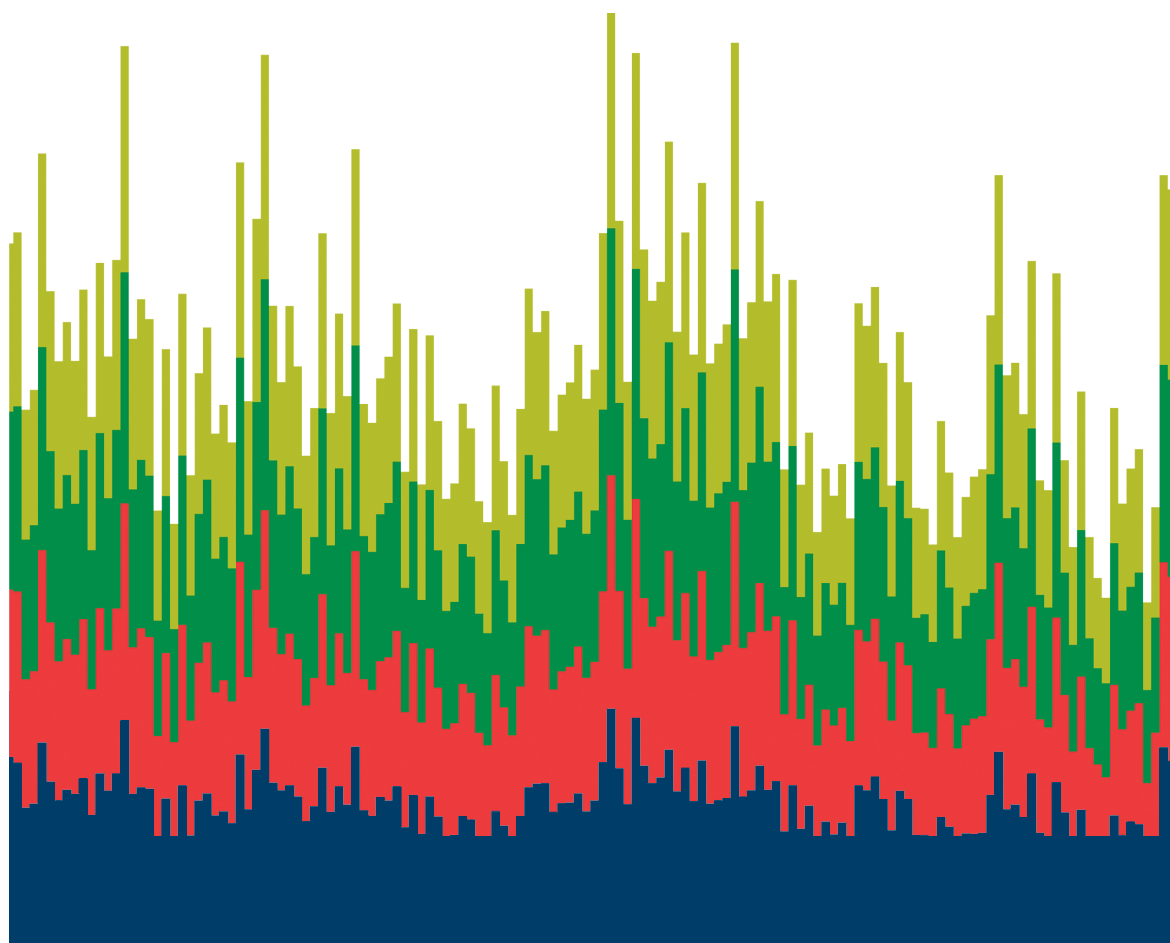




Analysis of I/O Requirements of Scientific Applications

Salem El Sayed Mohamed

IAS Series

Band / Volume 37

ISBN 978-3-95806-344-0

Mitglied der Helmholtz-Gemeinschaft

Forschungszentrum Jülich GmbH
Institute for Advanced Simulation (IAS)
Jülich Supercomputing Centre (JSC)

Analysis of I/O Requirements of Scientific Applications

Salem El Sayed Mohamed

Schriften des Forschungszentrums Jülich
Reihe IAS

Band / Volume 37

ISSN 1868-8489

ISBN 978-3-95806-344-0

Contents

Abstract	iii
Acknowledgements	v
List of Figures	ix
List of Tables	xiii
Abbreviations	xv
1 Introduction	1
1.1 Motivation	1
1.2 Research Goals	3
1.3 HPC I/O Architecture	4
1.3.1 Emerging I/O Architectures	7
2 I/O System Architecture	13
2.1 JUGENE I/O Stack	13
2.2 Storage Infrastructure	15
2.2.1 GPFS I/O Counters	16
3 Methodology: I/O Criteria	19
3.1 Related Work	21
3.2 Basic Quantities	23
3.2.1 Application Quantities	23
3.2.2 I/O Request Quantities	23
3.2.3 Filesystem Metadata Operation Quantities	25
3.3 Category 1: Aggregate Performance Numbers	26
3.4 Category 2: I/O Pattern Analysis	32
3.4.1 Request Size	34
3.4.2 Type Of I/O Operation	37
3.4.3 Spatiality Of I/O Requests	39
3.4.4 Temporal Intervals	44
3.4.5 Repetitive Behaviour	46
3.5 Category 3: Parallel I/O	48
3.6 Summarizing I/O Criteria	51

4	Performance Characterization: Analysing GPFS I/O Counters	55
4.1	Related Work And I/O Profiling Tools	55
4.1.1	I/O Measuring Tools	56
4.1.2	Analysis Process	59
4.1.3	Using Analysis Information	60
4.2	Reformatting GPFS I/O Counters	61
4.2.1	GPFS I/O Log Database	63
4.3	Job Database	64
4.4	Verifying Analysis Process	66
4.4.1	Verification Of GPFS I/O Counters Using I/O Benchmark	72
4.5	Evaluating JUGENE Job I/O	78
4.5.1	Filtering The Job List	79
4.5.2	Revisiting I/O Criteria	81
4.5.3	Category 1: Aggregate Performance Numbers	83
4.5.4	Category 2: I/O Pattern Analysis	99
4.5.5	Category 3: Parallel I/O	111
4.5.6	Further Analysing A Subset Of Jobs	116
4.5.7	Analysing Jobs Using Category 1: Aggregate Performance Numbers	119
4.5.8	Analysing Jobs Using Category 2: I/O Pattern Analysis	146
4.5.9	Analysing Jobs Using Category 3: Parallel I/O	150
4.6	General Notes on Analysing the GPFS I/O Counters	151
5	Performance Modeling: Modeling JUGENE I/O	153
5.1	Related Work	154
5.2	Modeling Framework (OMNET++)	155
5.3	Modelling JUGENE I/O	156
5.3.1	I/O Model Components	157
5.4	I/O Model Verification	162
5.4.1	Parameter Fitting Using GPFS I/O Logs	163
5.5	Future I/O Architectures	167
5.5.1	I/O Model Changes	168
5.5.2	Burst Buffers	171
5.6	Conclusions On Modelling System I/O	176
6	Conclusion	179
6.1	Future Work	182
A	I/O Criteria - Category 4: Application Details	185
B	I/O Model Parameter Fitting Using An I/O Benchmark	189
	Bibliography	193

IAS Series
Band / Volume 37
ISBN 978-3-95806-344-0