

https://dspace.org/

Chunked extendible arrays and its integration with the global array toolkit for parallel image processing

Nimako, Gideon

2016

http://hdl.handle.net/10539/22332

Nimako, Gideon (2016) Chunked extendible arrays and its integration with the global array toolkit for parallel image processing, University of the Witwatersrand, <http://hdl.handle.net/10539/22332>

Downloaded from a DSpace repository. https://dspace.org/

PhD Thesis Corrections

Chunked Extendible Arrays and its Integration with the Global Array Toolkit for Parallel Image Processing

Gideon Nimako (Student Number : 557286)

Introduction

In this document, we illustrate how the examiners comments are addressed for this PhD thesis submitted to the Faculty Of Engineering and the Built Environment on May 6, 2016. All minor corrections such punctuations, sentence constructs etc., have been corrected in the thesis. We present how we addressed the comments of each examiner.

Examiner 1

In this section we addressed the comments made by the examiner in the file `Externalexaminer1report.pdf`

List of Publications		
No.	Page(s)	Corrections & Comments
1	v	Additional spaces in references 1, 2, 3 have been removed.
2	v	The quotations around the title of the first publication have been removed for consistent citation style.
3	v-vi	DOI and links to publications have been added.
4	v	"In press" have been removed from publication 3.
Chapter 1		
No.	Page(s)	Corrections & Comments
5	4	Line 18. Missing reference is corrected
6	6	Line 15. The sentence "The organisations of ... two components;" was rephrased to "A PEXTA file has two components:"
7	8	The sentence with the phrase "compared with **the** conventional ..." has been corrected.
8	$1-9$	All occurrences of "organize" have been change to organise in the chapter.
9	9	Line 3 has been rephrased to how our parallel I/O scheme differs from parallel HDF5 file accesses.
10	9	On line 5, the acronyms PPI and AMEE has been expanded.
11	9	Line 8 has been corrected to reflect "Multidimensional Arrays and their application in....".
Chapter 2		
No.	Page(s)	Corrections & Comments
12	11	Line 13: We have given a brief explanation of shells.
13	12	Line 1: The last sentence in the previous paragraph has been rewritten to reflect the change required.
14	12	Line 6: specific class**es** of arrays. Corrected
15	12	Line 11: Required comma inserted.
16	13	Line 16: The sentence, "A number of research works use the mapping ..." has been slightly modified.
17	13	Line 17: "of" has been deleted in "One of such work is ..."
18	14	Line 2: The sentence with the phrase "and even allows" has been corrected.
19	14	Line 3: Extra space after (D, V) and before the comma has been removed.

20	14	Line 6: "Known limitations of each work presented". In each array scheme presented in the literature, we have indicated the limitation and challenges. For instance on page 12, line 2, we started outlining some of the challenges of the earlier computed access scheme in reference 5.
21	14	Line 9: Citation has been added.
22	15	Line 15: The sentence has been modified to indicate that the integration of extendible arrays into Global Array is actually work done in this thesis.
Chapter 3		
No.	Page(s)	Corrections & Comments
23	17	Line 1: "The" has been inserted.
24	17	Line 3: Citations have been added.
25	17	Line 4: "s" has been added to "advantage".
26	25	Line 3: The sentence with the phrase "in Appendix A.1 illustrates" has been orrected.
27	25	Line 3: The sentence with the phrase "for a 3-dimensional" has been orrected.
28	25	Line 12: The phrase "equation 3.3.1..." has been changed to "Equation 3.3.1..".
29	N/A	The words Section, Equation, Figure, etc., has been used consistently.
30	26	Line 5: The sentence with the phrase "The axial-vector" has been corrected.
31	27	Section 3.5: The number of random accesses has been added. We have also added an example to illustrate Algorithm 2.
Chapter 4		
No.	Page(s)	Corrections & Comments
32	31	Line 3: "Forward citation of the Lemma". We believe that citing the Lemma is important here if the reader wants an explanation to cubical extension.
33	33	Line 7: The sentence with the phrase "gray region (ghost region)" has been corrected.
34	33	Line 8: A comma has been added to the phrase "hand, if the".
35	36	Line 3: The sentence with the phrase "maintaining these" has been corrected.
36	36	Line 7: The sentence with the phrase "corresponds to" has been corrected.
36	38	Algorithm 3 has been formatted.
37	40	i_i is the ith index value of the k-dimensional index. = has been replaced with $\leftarrow$ in Algorithm 5.
38	42	Line 5: The sentence with the phrase "in a cyclic interruptible" has been corrected.
39	44	Line 9: The sentence with the phrase "two levels..." has been corrected.
Chapter 5		
No.	Page(s)	Corrections & Comments
40	47	Line 1: The sentence with the phrase "arrays used in" has been corrected.
41	47	Line 2: The sentence with the phrase "storage systems" has been corrected.
42	54	Line 17: The sentence with the phrase "a Hilbert curve" has been corrected.
43	56	Line 7: The sentence with the phrase "using Algorithm 7" has been corrected.
44	56	Algorithm 7: C_l has been changed to K for clarity.
45	58	We have directed the reader to reference 41 for further details of the Butz algorithm.
46	59	Figure 5.3.2 has been discussed in more detail.

Chapter 6		
No.	Page(s)	Corrections & Comments
47	62	We indicated on line 5 that array elements used for these experiments were double floating point numbers.
48	62	Why is the dimension randomly selected? The dimension for extension was randomly selected because typically the request for extension from an external application is arbitrary.
49	63	Line 7: The sentence with the phrase "final set of" has been corrected.
50	N/A	The line symbols distinguish the different methods in the graphs in the chapter.
Chapter 7		
No.	Page(s)	Corrections & Comments
51	71	Line 6: NetCDF has been used consistently in this chapter.
52	71	Line 3: The sentence with the word "scientists" has been corrected.
53	71	Line 5: The sentence with the phrase "building blocks of" corrected.
54	71	Line 6: The acronym "e.g.," has been corrected.
55	71	Line 9: The acronym "NetCDF" has been corrected.
56	72	Line 3: The sentence with the phrase "multiple locations" has been corrected.
57	74	Line 10: The word "datasets" has been corrected.
58	74	Line 12: The sentence with the phrase "must be set to unlimited" has been corrected.
59	75	Line 11: The sentences with the phrases "HDF5. Extensions are...". have been corrected appropriately.
60	80	Line 2: The word "extendibility" has been corrected.
61	80	Line 3: The word "scientific" has been corrected.
Chapter 8		
No.	Page(s)	Corrections & Comments
62	84	Line 13: The sentence with the phrase "or functions" has been corrected.
63	85	Line 13: The sentence with the phrase "open the file" has been corrected.
64	86	Line 1: The sentence with the phrase "uses the HDF5" has been corrected.
65	88	Line 8: The sentence with the phrase "buffers by processes" has been corrected.
66	89 – 91	There are 18 routines in the PEXTA library. We only presented 4 routines which are peculiar to extending GA persistent arrays.
67	96	Line 3: The sentences with the phrases "... bounds. PEXTA writes however" have been corrected.
Chapter 9		
No.	Page(s)	Corrections & Comments
68	100	Lines 3-6: Reference has been added.
69	104	Lines 17: The sentence with the phrase "... have shown that a large number ..." has been corrected.
70	104	Line 19: The sentence with the phrase "... improved by focusing ..." has been corrected.
71	104	Line 21: The sentence with the phrase "... techniques for reducing ..." has been corrected.

72	105	Line 1: The sentence with the phrase "feature selection techniques" has been corrected.
73	105	Line 2: The sentence with the phrase "... or rearrange existing ..." has been corrected.
74	105	Line 11: The sentence with the phrase "use the" has been corrected.
75	105	Line 15: The sentences with the phrases "our implementation. The...". have been corrected.
76	106	m has been used consistently in Algorithm 8.
77	107	Line 3: The sentence with the phrase "unmixing techniques: linear and" has been corrected.
78	108	Line 5: Citation for ENVI software package has been added.
Chapter 10		
No.	Page(s)	Corrections & Comments
79	128	Line 3: The sentence with the phrase "update of the ..." has been corrected.
80	128	Line 12: The sentence with the phrase "usually parameterized ..." has been corrected.
81	129	Line 1: The sentence has been rephrased.
82	129	Line 8: The sentence with the phrase "... in some applications that require ..." has been corrected.
83	132	Line 5: The sentence with the phrase "will be discarded." has been corrected.

Examiner 2

In this section we addressed the comments made by the examiner in the file `Externalexaminer2report.pdf`

Comment 1**: Citation of ADIOS**		
No.	Page(s)	Corrections & Comments
1	86	In Section 8.2 we explained (as well as citation for ADIOS) why PEXTA was not compared to ADIOS.
Comment 2**: Implication PEXTA on Applications and Future Contributions**		
No.	Page(s)	Corrections & Comments
2	85, 98 – 123	In Section 10.1.5 we summarise the impact of PEXTA on scientific applications in which the data grows by continually tiling of swathes of datasets to existing dataset. One of such applications is hyperspectral image acquisition and processing. PEXTA I/O is also suitable for Multi-dimensional On-Line Analytical Processing (MOLAP) used in data warehousing.

Examiner 3

In this section we addressed the comments made by the examiner in the file `Internalexaminerreport.pdf`

Comment 1**: Clarity on candidates contributions and background**		
No.	Page(s)	Corrections & Comments

1	20	We started Section 3.3 with "In this section we present the mapping function of extendible arrays that forms basis of the main contribution of this thesis."
2	25, 26	We started Section 3.4 with "Although inverse mapping function of extendible dense arrays are not frequently used in scientific applications, in certain specific scenarios such as nearest neighbourhood search, it is useful and handy. As part of the contribution of this thesis we present a function $\mathcal{F}^{*-1}$, similar to the one discussed in Section 3.2."
3	31	We have added a leading paragraph to indicate the fact that Chapter 4 presents the the main contribution of the thesis.
4	78	We started Section 7.3 with "This section presents one of the contributions of this thesis. We provide routines that extend HDF5 and NetCDF to support extendibility across multiple dimensions."
Comment 2: Clarity experimental datasets in Chapter 6 and 8		
No.	Page(s)	Corrections & Comments
5	62, 63	The experiments in Chapter 6 used simulated datasets. We have added a section (Section 6.3) to the chapter describing how the array data cubes and elements were allocated.
6	91, 92	The experiments in Chapter 8 used simulated datasets. We have added a section (Section 8.5.1) to the chapter describing how the array data cubes and elements were generated and allocated.
7	51, 92	The performance of our allocation scheme is rather sensitive to the chunk size. The estimation of the optimal chunk size has been extensively explored in reference [33] of the thesis (See pages 50 and 51).
Comment 3: Code Availability		
No.	Page(s)	Corrections & Comments
8	131	Yes the code for the core chunked extendible array with O_2-Tree indexing structure is readily available and will be made available on suitable code distribution repository platform. PEXTA on the other cannot be used in isolation. It is to be embedded in Global Array Toolkit (GA). The version of the GA used for this thesis is 5.2. An ideal model to make this work have high impact is to collaborate with GA group at the Pacific Northwest National Laboratories (PNNL) to incorporate PEXTA as one of GA I/O library. We have added this as on page 131 as one of the activities outline for future direction.
Comment 4: Annotated Thesis and Other Corrections		
No.	Page(s)	Corrections & Comments
9	27, 61	Under Section 3.5 in Chapter 3, we have added the exact Intel Quad Core processor specifying L2 Cache size. We have also indicated the compiler used in these experiments and option flags set. Similarly, we modified section Section 6.1 in Chapter 6 to reflect these details.
10	57, 58	The Butz iterative algorithm is an alternative approach of computing the displacement within a chunk when Hilbert curve is utilized. The explanation of $\alpha, \eta, \tau, \tilde{\tau}, \rho$ and σ are explained on page 58. Further details of the algorithm are found in references 42 and 43.

11	See Correction & Comments	All punctuations, grammar, structure and the comments in the annotated thesis have been addressed. There following are the pages of those changes: title page, ii, 2, 3, 4, 5, 12, 13, 21, 23, 24, 28, 31, 45, 49, 50, 52 – 55, 58, 61, 63, 71 – 78, 80 – 83, 87, 92, 95, 104, 112 – 113, 116 – 118, 121 – 122 and 125 – 126.
12	136 – 146	All uncorrected references have been corrected.