

GIS PROCESSING ON COMPUTE CLUSTERS

by

Nathan Thomas Kerr

A Thesis Presented in Partial Fulfillment  
of the Requirements for the Degree  
Master of Science

ARIZONA STATE UNIVERSITY

August 2009

GIS PROCESSING ON COMPUTE CLUSTERS

by

Nathan Thomas Kerr

has been approved

August 2009

Graduate Supervisory Committee:

Dr. Daniel Stanzione, Chair  
Dr. Robert Pahle  
Dr. Yi Chen

ACCEPTED BY THE GRADUATE COLLEGE

## ABSTRACT

Abstract goes here

Dedication goes here

## ACKNOWLEDGEMENTS

Acknowledgments go here.

## TABLE OF CONTENTS

	Page
CHAPTER 1 Introduction . . . . .	1
1.1 Current methods . . . . .	1
1.2 Significance . . . . .	1
1.3 Attempted solutions . . . . .	1
1.4 Needed solution, cluster programming model . . . . .	2
CHAPTER 2 Related Work . . . . .	3
2.1 Desktop GIS . . . . .	3
2.2 Database GIS . . . . .	3
2.3 GIS libraries . . . . .	3
2.4 Parallel DB . . . . .	3
2.5 Parallel GRASS . . . . .	3
2.6 MRGIS MapReduce GIS . . . . .	4
CHAPTER 3 Requirements . . . . .	5
3.1 easy to use to form new operations . . . . .	5
CHAPTER 4 Design/Implementation . . . . .	6
CHAPTER 5 Experimental Setup/Design . . . . .	7
5.1 Operation set . . . . .	7
5.2 Dataset Description . . . . .	7
5.3 Hardware setup . . . . .	7
5.4 implementations . . . . .	8
5.4.1 PostGIS . . . . .	8
5.4.2 Hadoop Prototype . . . . .	8
5.4.3 ClusterGIS . . . . .	8
CHAPTER 6 Results . . . . .	9
6.1 Performance Analysis . . . . .	9
6.2 Comparisons . . . . .	9
CHAPTER 7 Recommendations . . . . .	10
7.1 Applicable Problem Spaces . . . . .	10
CHAPTER 8 Conclusion . . . . .	11
8.1 Future Work . . . . .	11

# **Chapter 1**

## **Introduction**

GIS processing is important

1.1 Current methods

Desktop GIS

Database GIS

1.2 Significance

Limitations of single processor implementations

1.3 Attempted solutions

Parallel DB

Parallel GRASS

Problems with desktop programs on clusters

#### 1.4 Needed solution, cluster programming model

batch

mpi

scalable architecture

spmd

Thesis statement goes here

How this document goes about supporting the thesis statement.

# **Chapter 2**

## **Related Work**

2.1 Desktop GIS

ArcInfo, QuantumGIS, GRASS GIS

2.2 Database GIS

PostGIS, Oracle Spatial, Client-Server Paradise

2.3 GIS libraries

JTS, GEOS

2.4 Parallel DB

Paradise

2.5 Parallel GRASS

Other converted programs?

## 2.6 MRGIS MapReduce GIS

# **Chapter 3**

## **Requirements**

batch mode processing

data distribution options

standard geospatial operations

scalable, high performance

dataset centric

### 3.1 easy to use to form new operations

Sample program:

initClusterGIS

loadData (distributed/replicated)

process (user code here)

saveData (distributed/replicated)

finalizeClusterGIS

# **Chapter 4**

## **Design/Implementation**

How to fulfill requirements

Architecture of solution

Split dataset across tasks (MPI I/O)

Each task works on their own part

MPI to communicate if needed

# **Chapter 5**

## **Experimental Setup/Design**

How to show how well requirements were filled

### **5.1 Operation set**

These operations are representative of problem space.

Data access based: parallelization is based off data decomp, so operations are too.

OGC standard requirements - survey of methods requiring 1, 2, or more data points.

### **5.2 Dataset Description**

Full datasets; 34k employers, 1.2m parcels

sub datasets, how to generate from full

listing of datasets used

### **5.3 Hardware setup**

Saguaro

## 5.4 implementations

### 5.4.1 PostGIS

### 5.4.2 Hadoop Prototype

### 5.4.3 ClusterGIS

# **Chapter 6**

## **Results**

### 6.1 Performance Analysis

vary procs, maintain data

vary data, maintain procs

vary data, vary procs

### 6.2 Comparisons

PostGIS as baseline

PostGIS to Hadoop

Hadoop to ClusterGIS

PostGIS to ClusterGIS

# **Chapter 7**

## **Recommendations**

Synthesis

### 7.1 Applicable Problem Spaces

Analysis of problem types that are good/bad for this approach

# Chapter 8

## Conclusion

### 8.1 Future Work

additional decomposition methods, combined with alternative MPI communicators

addition of preprocessing methods (indexing, etc)

chunking of replicated dataset

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Quisque a augue nec sapien varius vulputate et nec purus. Vivamus venenatis massa non dolor blandit sed venenatis urna lobortis. Fusce condimentum euismod condimentum. Suspendisse ligula neque, scelerisque eget viverra vel, ornare eget nibh. Sed nec mi metus. Aliquam erat volutpat. Pellentesque sit amet libero risus, id viverra est. Suspendisse sit amet velit et quam varius laoreet. Morbi quis laoreet ante. In metus sapien, venenatis eu sodales nec, sollicitudin eu quam. Nullam consequat scelerisque placerat. Phasellus dapibus mauris vel lacus dictum venenatis. Pellentesque vestibulum bibendum purus, at vestibulum felis ullamcorper ut. Integer mattis rhoncus urna, at pellentesque magna accumsan consequat. Sed dictum ultricies viverra. In nec pharetra elit. Mauris faucibus pharetra odio faucibus pharetra. Fusce eget dolor est, sed adipiscing quam. Lorem ipsum dolor sit amet, consectetur adipiscing

elit. Quisque a augue nec sapien varius vulputate et nec purus. Vivamus venenatis massa non dolor blandit sed venenatis urna lobortis. Fusce condimentum euismod condimentum. Suspendisse ligula neque, scelerisque eget viverra vel, ornare eget nibh. Sed nec mi metus. Aliquam erat volutpat. Pellentesque sit amet libero risus, id viverra est. Suspendisse sit amet velit et quam varius laoreet. Morbi quis laoreet ante. In metus sapien, venenatis eu sodales nec, sollicitudin eu quam. Nullam consequat scelerisque placerat. Phasellus dapibus mauris vel lacus dictum venenatis. Pellentesque vestibulum bibendum purus, at vestibulum felis ullamcorper ut. Integer mattis rhoncus urna, at pellentesque magna accumsan consequat. Sed dictum ultricies viverra. In nec pharetra elit. Mauris faucibus pharetra odio faucibus pharetra. Fusce eget dolor est, sed adipiscing quam. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Quisque a augue nec sapien varius vulputate et nec purus. Vivamus venenatis massa non dolor blandit sed venenatis urna lobortis. Fusce condimentum euismod condimentum. Suspendisse ligula neque, scelerisque eget viverra vel, ornare eget nibh. Sed nec mi metus. Aliquam erat volutpat. Pellentesque sit amet libero risus, id viverra est. Suspendisse sit amet velit et quam varius laoreet. Morbi quis laoreet ante. In metus sapien, venenatis eu sodales nec, sollicitudin eu quam. Nullam consequat scelerisque placerat. Phasellus dapibus mauris vel lacus dictum venenatis. Pellentesque vestibulum bibendum purus, at vestibulum felis ullamcorper ut. Integer mattis rhoncus urna, at pellentesque magna accumsan consequat. Sed dictum ultricies viverra. In nec pharetra elit. Mauris faucibus pharetra odio faucibus pharetra. Fusce eget dolor est, sed adipiscing quam. Integer eu metus at nisi tincidunt ultricies quis ut odio. Vestibulum lacinia faucibus risus a tristique. Proin pretium, dolor ut vulputate iaculis, erat arcu semper quam, interdum egestas nibh erat id sem. Quisque tincidunt nisi in orci lobortis tempor porttitor ante aliquam. Proin varius, diam eu mattis cursus, risus lacus ultricies ipsum, nec iaculis quam lacus nec magna. Cras eget mollis massa. Mauris tincidunt urna

vitae sapien convallis tincidunt. Mauris luctus tincidunt risus sit amet elementum. Nullam ut eros et magna posuere feugiat non sed erat. Pellentesque blandit, risus id placerat molestie, felis ante ullamcorper mi, ac laoreet urna dolor vel dolor. Maecenas in urna quis libero aliquam faucibus eget eu est. Nunc orci tellus, dictum nec eleifend eu, elementum sed nibh. Morbi et nisi risus, at egestas ipsum. Nunc eget orci sit amet metus porttitor aliquet. Vestibulum facilisis semper nunc ut tincidunt. In lobortis orci in eros hendrerit eleifend. Nam commodo mollis odio ut consectetur. Duis et enim arcu. Aenean lacinia facilisis metus. In hac habitasse platea dictumst. Suspendisse consequat ornare est, vitae porttitor enim viverra bibendum. Nulla sapien urna, pulvinar nec vulputate in, vehicula id est. Donec elit metus, rhoncus at iaculis et, consectetur eu lectus. Aenean sed mauris massa, vel ullamcorper justo. Maecenas et odio quis ligula ultricies elementum eu eget orci. Sed facilisis, mauris vitae vestibulum lobortis, lorem mi aliquet odio, sit amet dignissim nulla tortor nec risus. Vestibulum sit amet urna sit amet eros adipiscing consectetur. Vestibulum sodales tristique cursus. Aellentesque fermentum purus ac urna convallis id aliquam nulla scelerisque. Etiam auctor iaculis pharetra. liquam aliquam auctor velit quis interdum. Curabitur lectus metus, convallis non rutrum in, dictum tincidunt velit. Nulla sollicitudin, orci sit amet venenatis laoreet, metus risus porta tellus, in fermentum erat diam eget diam. Nulla turpis ante, pulvinar non tincidunt id, luctus vel nunc. Duis bibendum tincidunt urna, quis aliquet ligula aliquam vitae. Nam imperdiet ipsum ac lacus facilisis placerat. Nunc placerat nunc felis, a condimentum turpis. Praesent diam libero, feugiat eget viverra eu, scelerisque eget orci. Donec dignissim venenatis dolor, in placerat erat rutrum non. Nulla ut viverra dui. Pellentesque fermentum purus ac urna convallis id aliquam nulla scelerisque. Etiam auctor iaculis pharetra. Vestibulum ornare sollicitudin fermentum. Maecenas non vestibulum tortor. In hac habitasse platea dictumst. Phasellus iaculis viverra feugiat. Pellentesque sem enim, vehicula eget lacinia vel, laoreet et do-

lor. Vestibulum viverra lobortis nisl et dapibus. Fusce vehicula quam quis libero mattis fringilla. Mauris ut lorem eu lorem commodo pellentesque. Quisque ullamcorper ligula at velit tempus vitae varius purus fringilla. Nunc aliquam velit nec sapien cursus adipiscing. Cras sit amet neque ante. Maecenas ipsum tellus, facilisis sit amet lobortis at, semper ac ligula. Etiam pellentesque velit nec nisl vehicula scelerisque. Vivamus iaculis congue varius.