

# Apache Crunch

SCOTT MCCLARY<sup>1,\*</sup>

<sup>1</sup>*School of Informatics and Computing, Bloomington, IN 47408, U.S.A.*

<sup>\*</sup>*Corresponding authors: scmccclar@indiana.edu*

*paper-002, March 25, 2017*

insert abstract here

© 2017 <https://creativecommons.org/licenses/>. The authors verify that the text is not plagiarized.

**Keywords:** Big-Data, Cloud, Hadoop, I524, MapReduce

<https://github.com/scottmcclary1/sp17/i524/blob/master/paper2/S17-IO-3011/report.pdf>

## 1. INTRODUCTION

insert introduction here

According to [1], Apache Crunch is...

## 2. ARCHITECTURE

insert architecture here

### 2.1. API

insert api here

### 2.2. Shell Access

insert shell access here

### 2.3. Graphical Interface

insert graphical interface here

#### 2.3.1. Science Gateways

insert science gateways here

## 3. LICENSING

insert licensing info here

## 4. ECOSYSTEM

insert ecosystem here

## 5. USE CASES

insert use cases here

### 5.1. Use Cases for Big Data

insert use cases for big data here

## 6. EDUCATIONAL MATERIAL

insert educational material here

## 7. CONCLUSION

insert conclusion here

## ACKNOWLEDGEMENTS

The authors would like to thank the School of Informatics and Computing for providing the Big Data Software and Projects (INFO-I524) course [2]. This paper would not have been possible without the technical support & edification from Gregor von Laszewski and his distinguished colleagues.

## AUTHOR BIOGRAPHIES



**Scott McClary** received his BSc (Computer Science) and Minor (Mathematics) in May 2016 from Indiana University and will receive his MSc (Computer Science) in May 2017 from Indiana University. His research interests are within scientific application performance analysis on large-scale HPC systems. He will begin working as a

Software Engineer with General Electric Digital in San Ramon, CA in July 2017.

## WORK BREAKDOWN

The work on this project was distributed as follows between the authors:

**Scott McClary.** He completed all of the work for this paper including researching and testing Apache Airavata as well as composing this technology paper.

## REFERENCES

- [1] Curoverse Inc, "Arvados | Introduction to Crunch," Web Page, accessed: 2017-2-22. [Online]. Available: <http://doc.arvados.org/user/tutorials/intro-crunch.html>
- [2] Gregor von Laszewski and Badi Abdul-Wahid, "Big Data Classes," Web Page, Indiana University, Jan. 2017. [Online]. Available: <https://cloudmesh.github.io/classes/>