Parallel patterns against an exabyte data lake using exascale heterogeneous computing

Mr Andreas Vermeulen
University of St Andrews
Saint Andrews, Fife KY16 9AJ
University of Dundee
Nethergate,Dundee DD1 4HN
a.f.vermeulen@dundee.ac.uk

Dr Vladimir Janjic
University of St Andrews
Saint Andrews, Fife KY16 9AJ
vj32@st-andrews.ac.uk

Mr Andy Cobley
University of Dundee
Nethergate, Dundee DD1 4HN
acobley@computing.dundee.ac.uk

ABSTRACT

An enhancement of a rapid information factory using exascale heterogeneous computing and parallel knowledge-extraction patterns to generate a deep learning source from a exabyte data lake.

Categories and Subject Descriptors

H.4 [**Information Systems Applications**]: Miscellaneous

General Terms

Theory, Framework, Application, Research, Hardware

Keywords

exabyte, exascale

1. RESEARCH QUESTION

Can a Rapid Information Factory using agile and lean six sigma manufactory principles to solve the issues generated by effective and efficient exascale heterogeneous computing of a quintillion bytes data lake into a value-add deep learning knowledge source?

2. RAPID INFORMATION FACTORY (RIF)

2.1 Rapid Information Factory Framework (RIFF)

The rapid information factory framework is a methodology, the result of research since 2008 , designed to guide a exascale [1] heterogeneous computing cluster to process a exabyte data lake. The framework will processes a quintillion calculations per second against quintillion bytes of disk storage. The framework generates a series of virtual factories that together process the data lake using enhanced custom designed parallel processes.

2.2 Rapid Information Factory Cluster (RIFC)

The rapid information factory cluster is a hetrogeniuos computing cluster.

2.3 Autonomous Node Transport (ANT)

????

2.4 Persistent Uniform Protocol Agreement (PUPA)

????

3. REFERENCES

[1] Bergman, K., Borkar, S., Campbell, D., Carlson, W., Dally, W., Denneau, M., Franzon, P., Harrod, W., Hill, K., Hiller, J., et al. Exascale computing study: Technology challenges in achieving exascale systems. *Defense Advanced Research Projects Agency Information Processing Techniques Office (DARPA IPTO), Tech. Rep 15* (2008).