International Journal of Research in Engineering and Innovation (IJREI)
journal home page: http://www.ijrei.com
ISSN (Online): 2456-6934

Performance analysis of FME based servers and cloud for data loading in big data and machine learning models for future data mining process, knowledge discovery in geo-spatial data

Aamir khan[1], Arif Ali Wani[2], Umesh Chandra[3], Parag Jain[4]

[1]*Uttarakhand Technical University, Dehradun*
[2]*Glocal University, Saharanpur (U.P)*
[3]*Banda university, Banda (U.P)*
[4]*RIT Roorkee Uttarakhand*

Abstract

With respect to absurdity of data during the last 10 years, various unwanted data is captured and it became important factor in decision making for future technologies. Servers and data storage cloud is not much more flexible to handle or manage all types of data and the performance of data analysis is decreasing day by day. Feature manipulation engine (FME) is used for data transformation with regards to various data formats that can handle and easily integrated with all servers and cloud data, so we can directly access data all over the world. Starting from creation of data to storage of data, furthermore, the proposed hybrid prediction model has better fault prediction for analyzing the server performance ahead of improving decision making in GIS data storing in cloud.

Keywords: FME, Data modeling, performance analysis, NADCON, GIS, NGS (National geodetic survey).

1. Introduction

The feature manipulation engine (FME) is an open source platform/ interface that smooth-running the translation of GIS data between geometric and digital formats. It is deliberating and substantially used with geographically information system and raster, vector graphic software, computer based design[1]. Safe software solution was developed this concept at Inc. of survey, British Colombia Canada [2]. FME basically work for location based application which may include useful application/ platform such as interactive geological, geographical and topographical mapping.

Example: Google earth and map quest.

Data exchange format is easily done by FME, it is very evaluating facilitated that transform geospatial data into a variety of formats, various data models and repositories for transfer, transmission to end developer and users[3]. This is continuous process is called special data extraction or mining, transformation and loading. FME include various modulated features[3]:

1. Interdisciplinary sites are migrating with FME.
2. Consolidation of data from known source.
3. Data distribution for multiple end users.
4. Developer tool kit.

2. FME servers

FME servers are real time operating system based servers that works hand in hand with all FME desktop applications. FME desktop gives a drag and drop interface and then publish it into FME servers at the push off button[1]. Using servers you can easily manage and scheduled the work space which is created to autonomously run at the specific time of interval, it may be converted into self-server process for remote users to run whenever they want/ act. All your work space is running in a real time operating system. Enterprise works with power of data flow across all over the cloud from simple cloud system to complex work flow cloud system with legacy systems[4].

Corresponding author: Aamir khan
Email Address: amirkhan@theglocaluniversity.in

Everywhere in the world FME servers includes RESTAPI and supports common protocols such as SMS, E-Mail and sockets, JMS and many more. It is basic support of 100 of files formats and web applications [3].

Profiling: Demo

Timer Name	Timer Context	User CPU Time	System CPU Time	Elapsed Real Time
FACTOR_TYPE	PointCloudSplitterFactory	18.140625	3.4375	30.965623
FACTOR_INSTANCE	RasterExtentsCoercer_PolygonGenerator	23.140625	2.46875	30.725689
FACTOR_TYPE	RasterSegmentationFactory	23.125	2.453215	30.723947
FACTOR_INSTANCE	ColorintensityDistribution_PointCloudSplitter_2	4.5	0.859375	8.077807
FACTOR_INSTANCE	ColorintensityDistribution_PointCloudSplitter	4.515625	0.796875	7.453812
FACTOR_INSTANCE	ColorintensityDistribution_PointCloudSplitter_4	4.34375	0.796875	7.162886
FACTOR_INSTANCE	ColorintensityDistribution_PointCloudSplitter_3	3.6875	0.71875	6.286207
FACTOR_INSTANCE	Clipper_Raster	4.5	0.078125	5.41969
FACTOR_TYPE	RasterClippingFactory	4.5	0.078125	5.419679
FACTOR_INSTANCE	FeatureReader	0.078125	0.765625	2.353204
FACTOR_TYPE	QueryFactory	0.078125	0.765625	2.353196
FACTOR_INSTANCE	PointCloudSplitter	1.09375	0.265625	1.985536
FACTOR_TYPE	TextStrokerFactory	0.125	0.515625	1.6135565
FACTOR_INSTANCE	ColorIntensityDistribution_TextStroker_2	0.125	0.515625	1.604116
FACTOR_TYPE	TeeFactory	0.859375	0.53125	1.552781

Figure 1: Profiling performance demo for server[5]

Performance analysis of FME servers: As the technological world continuously evaluate the organization and faced with integrated various number of data source and applications with their operations rely upon. All applications have their own strength, every server have some performance to evaluate data inter exchange capability [1]which is given the best data transformation rate in the form of performance. Data rate for migration one format to[6] another format, how much time is taken for transmission measured in performance analysis. Performance analysis and testing is a phase where we ensure that software/ servers will perform well under their terms and conditions with expected work load. In this evaluation of features and functions are not concerned. Some features are calculated:-

1. Response time
2. Reliability
3. Resource used
4. Scalability

Performance testing is basically non-function testing technique which is performed in background to determine the system perimeters which is helpful for checking the responsiveness of servers, performance analysis of servers life cycle, behavior under various extreme conditions [7].

3. Big Data and machine learning models for data loading

FME servers 2016.1 has very much capability for data loading from web/hyper data and other third party authentic data. It allows user to take basic raw data from any type of resource and automatically loaded into FME[8] servers databases. But data after loading may be stored and supply into predefined database structure[7]. In the background data loading is working itself and auto configuration into two step process.

3.1 First step process

Access/ approach to the data as it is provided by the user to a workspace hosted on FME servers, then the work practice is executed taking data same as provided by the resource person then writing out to its final submission output on the screen. This complete process is thoroughly monitoring and handled by web services and notification services.

3.2 Second step process

For achieving more strength of data format conversation we are moving it between different format and server applications that will be often required at a specific level of manipulation. Simple data format can be completely transformed with the help of FME complete data transformation. FME server support more than 400 data format, FME tool allow to end user for most complex data component for reconfiguration and allow carry them into a practically any remote system.

4. Data loading with web services

For publishing data set, there may be dataset with parameters on the reader to allow the data must be read into workspace/ portal. Once your loading data is completely published into FME servers, FME provide an interface to drag and drop all browsing data files. Data set can located locally as well as remotely [3, 6], after that they will be provided as a URL or web services to access the data inside the workspace of FME servers.

5. Data loading with the notification services

Notification services can be push data into FME servers with message format. This message can be execute or passes or sent to the FME cloud servers in a variety of formats including HTTP, POST to the server through E-mails [9].
FME servers performance and profiling: Performance is all about time consumption as well as data processing and data loading time. Data set takes predefine time for processing, loading, storing and after that manipulation again for performance analysis. We will analyze where the time is going on. After this we can check profiling, the meaning of profiling is "What is taking so long?"

Profiling with FME

Timer Name	Timer Context	User CPU Time	System CPU time	Elapsed Real Time
FACTORY_TYPE	PIPComponentsFactory	0	0.015625	0.007439
FACTORY_TYPE	SortFactory	0.015625	0.015625	0.014168
FACTORY_TYPE	LabelFactory	0	0	0.007133
FACTORY_TYPE	CreationFactory	0	0	0.00258
FACTORY_TYPE	TransformFact	0.03125	0.171875	0.42796
FACTORY_TYPE	QueryFactory	0.046875	0.6875	2.201396
FACTORY_TYPE	RasterSegmentationFactory	21.859375	2.78125	29.551645
FACTORY_TYPE	SolidIntersectionFactory	0	0	0.000152
FACTORY_TYPE	NeighborColorSetterFactory	0	0	0.011963
FACTORY_TYPE	RasterClippingFactory	4.09375	0.0625	4.641151
FACTORY_TYPE	TextStrokerFactory	0.140625	0.5625	1.675614
FACTORY_TYPE	ConnectionFactory	0.046875	0	0.041158
FACTORY_TYPE	ClippingFactory	0.015625	0	0.032626
FACTORY_TYPE	RoutingFactory	0.015625	0.265625	0.329758
FACTORY_TYPE	AggregateFactory	0.34375	0	0.30847
FACTORY_TYPE	ChoppingFactory	0	0.03125	0.005066
FACTORY_TYPE	PointCloudSplitterFactory	16.71875	3.796875	25.46215
FACTORY_TYPE	VectorToRasterFactory	0	0	0.012541
FACTORY_TYPE	TeeFactory	0.96875	1.046875	2.416896
FACTORY_TYPE	AttrSetFactory	0	0	0.023371

Figure 2: Profiling of FME Servers[10]

Profiling with FME: From last two years we are analyzing that FME has a great ability to perform profiling power that will give you result how plenty time is acquired or spend in fundamental functions of FME[8]. To turn it on do the following:

1. Select "Tools/edit header".
2. Directive "FME_PROFILE_RESULT_CSV<file path>"

6. FME usage for mobile/cloud mapping based GIS data

MEIT (Micro Engineering Tech Inc.) is a group of reading professional research and Development Company which is provided constructing and product services for geomatic and GIS data. This organization uses GIS data creation like points, line and polygon for basic insertion of data into data loading in FME servers. After this GIS data can be exported using FME desktop. FME data can be using mobile mapping and site/web based data for data modeling, after that complete this process an automatic servers [11] workspace analyses the background FME workspace and integrated with FME servers and clouds. Now it will be ready for publishing, in this entire process an auto config system is monitoring and analyzing the server functionality and analyzing the performance of the system/server. VISAT is a technology which will provide GIS data as a resource data which will be helpful for inform of some device captured data [12]. These data can be captured with the following technologies [11]:

1. Digital cameras
2. GPS (Global Positioning System)
3. Inertial navigation
4. Odometer
5. Laser Scanner.

VISAT can be defined as: 1. 6 cameras, 2 Laser scanners, integrated GPS, inertial navigation and odometer and Image capture at 1 to 10 meter interval.
Data from VISAT also export with the help of FME into CAD or other data formats.
GIS data and modeling with server performance: GIS data is very large and format can be vector and raster, these data first can be captured by VISAT technology[3]. After this data is to be inserted into FME predefined structure then process into performance with respect to time consuming for data creating, loading, storing to complete published data[13].

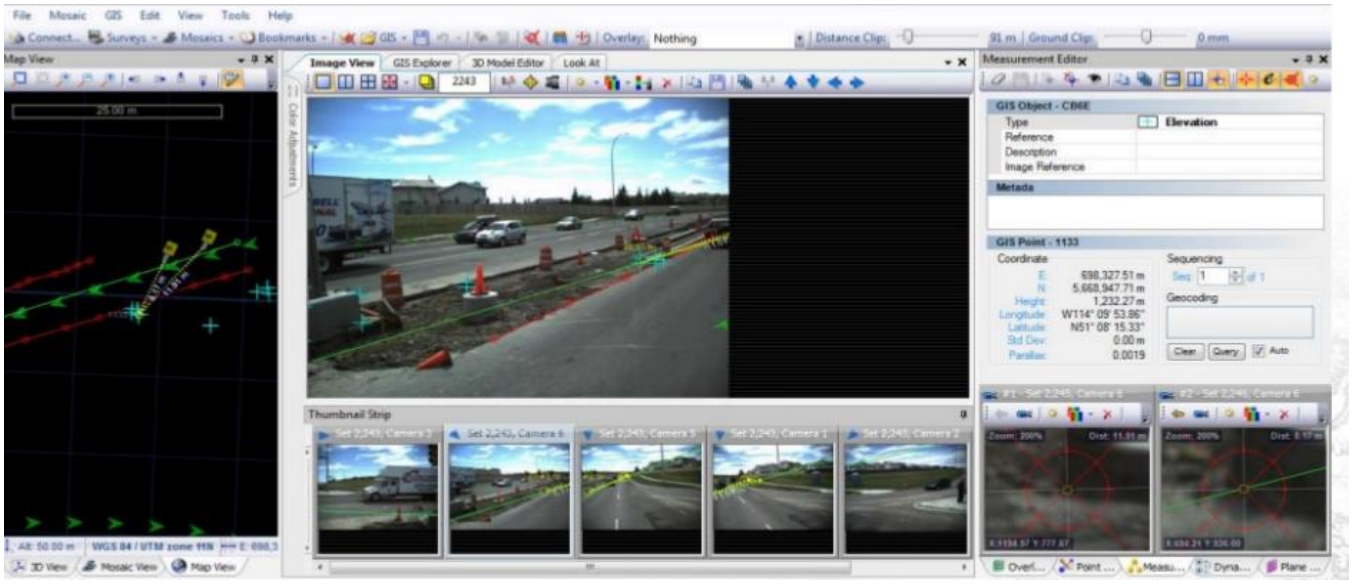

Figure 3: VISAT construction example

6. Conclusion

FME is a new feature technology which is used to construct all data format into a common or particular data format for common platform. In this paper we can only analyze the server capabilities and performance on the behalf of data loading. Various other internal technologies and functions are work together for a common objective. The main focus to conclude performance testing[5], so that we can give further research on this technology so that server can load data with high accuracy and efficiency with VISAT system.

References

[1] A. Khan, S. Ahmad, M. Hussain, and C. Science, "Case Study of 802 . 11a with open-air-interface transmission with simulation methodology," no. 07, pp. 140–142, 2018.

[2] N. Tyagi, S. Ahmad, A. Khan, and M. M. Afzal, "Sentiment Analysis Evaluating the Brand Popularity of Mobile Phone by Using Revised Data Dictionary," vol. 7, no. 3, pp. 53–61, 2018.

[3] I. Kureshi, C. Pulley, J. Brennan, V. Holmes, S. Bonner, and Y. James, "Advancing Research Infrastructure Using OpenStack," pp. 63–69, 2013.

[4] "Implementation of a Geoserver application for GIS data distribution and manipulation," 2013.

[5] R. Cheng, R. Knopp, C. Ho, K. Hsu, and T. Liu, "Demo Abstract - Design and Implementation of an Open Source NB-IoT eNB," no. April, 2018.

[6] S. Paudel, "Investigation , Analysis and Implementation of Open Source Mobile Communication Software," no. September, 2016.

[7] U. Chandra and P. Jain, "Advanced Data Mining Approaches of Knowledge Data Discovery for Geospatial Data," vol. 8, no. 12, pp. 24–29, 2017.

[8] C. Nahum, P. Batista, and A. Klautau, "Emulation of 4G / 5G Network Using OpenAirInterface," pp. 989–990, 2017.

[9] "FME SUCCESS - perspectives on safe and effective storage of CO2," no. June, 2017.

[10] T. X. Tran, A. Younis, and D. Pompili, "Understanding the Computational Requirements of Virtualized Baseband Units using a Programmable Cloud Radio Access Network Testbed."

[11] U. Nations, E. Commission, F. O. R. Europe, C. Of, and E. Statisticians, "united nations economic commission for europe Meeting of the 2016 / 2017 Bureau For discussion and recommendations Item II (a) of the Provisional Agenda In-depth review of data integration Prepared by the secretariat and the participants and project lea," 2017.

[12] O. Sefraoui and M. Aissaoui, "OpenStack : Toward an Open-Source Solution for Cloud Computing," vol. 55, no. 03, pp. 38–42, 2012.

[13] A. Biotech *et al.*, "Appendix - 1 : List of ICI Journal Titles Pending Evaluation."

Cite this article as: Aamir khan, Arif Ali wani, Umesh Chandra, Parag Jain, Performance analysis of FME based servers and cloud for data loading in big data and machine learning models for future data mining process, knowledge discovery in geo-spatial data, International journal of research in engineering and innovation (IJREI), vol 3, issue 1 (2019), 68-71.