3.0 METHODOLOGY

3.1 Introduction

This research paper aims at studying and comparing the benefits of Co-location and proximity hosting by developing a model of Market microstructure and order matching process that clearly identifies the benefits of co-location based on the degree of proximity. We shall further develop a sample order matching system which can simulate the benefits of co-location using a sample HFT strategy. We shall also review relative expense of co-locating HFT servers for different global markets.

3.2 Data collection and Location of Study

Historical data is required as the training data for simulating the benefits of Co-location. For this project, BSE data will be downloaded from Yahoo Finance from Jan 2019 to Dec 2019. The instruments here have sufficient liquidity to support high frequency trading. We shall then narrow down to demonstrate the benefits of Colocation using futures trading as a sample. We will later generalize the results to the entire population after backtesting and testing the robustness of the model.

This paper seeks to study the benefits of Colocation and Proximity hosting at BSE in India.According to a report on Algo Trading, India boasts of the fastest colocation in the world. BSE is the fastest Exchange in the world with a speed of 6 microseconds. Market access across Equity, Equity Derivatives and Currency Derivatives segments (Table 3.1). BSE provides the fastest Co-location service in India with round trip network latency of less than10 microseconds. At BSE Colocation response for an order has round trip latency of about 16 microseconds (including 10 microseconds of Co-location network latency).

Table 3.1: The segment wise percentage of Algo orders coming from co-location

SEGMENT	**% ALGO ORDERS OUT OF TOTAL ORDERS IN BSE**	**% COLO ORDERS OUT OF TOTAL ALGO ORDERS IN BSE**
EQUITY	98.56	80.00
EQUITY DERIVATIVES	99.89	80.10
CURRENCY DERIVATIVES	94.93	91.88
INTEREST RATE DERIVTIVES	93.22	96.16

3.3 Method of data collection

Python is used as the programming language for collecting the data,developing the model and backtesting, due to its vast number of libraries and simplicity. Using the **yfinance** package (https://aroussi.com/post/python-yahoo-finance), historical BSE stock prices were downloaded from Yahoo Finance. Since the datatype of downloaded data will be a **pandas** DataFrame, it is simple to compute the stock returns by using the **pandas** package's **pct_change** method on the *Adjusted Close* prices

3.4 Hypotheses

Based on analysed academic literature from different scholars, we came to a conclusion that proximity Hosting and Co-Location to matching engines reduces latency, increases market liquidity, reduces bid ask spread, increases chances of making a high profit and increases market depth.We thereby state the central hypotheses of this paper as follows:

1) Proximity Hosting and Co-Location to matching engines in HFT reduces latency.
2) Proximity Hosting and Co-Location to matching engines in HFT has significant impact on market liquidity, efficiency and depth.

We wish to this through a regression model using future contracts at BSE,India.

3.5 Method of Analysis

In this research paper, we will use the regression model as a machine learning method to clearly show the impact and benefits of Colocation and proximity hosting. The reason why regression model is chosen instead of the other machine learning methods is that; first, this is the method preferred by most researchers and scholars as supervised learning method. Secondly, this technique is used for forecasting, time series modelling and finding the causal effect relationship between variables. This model best suits this project since we are simulating the benefits of co-location using a sample HFT strategy. We want to see the results of introducing Colocation at the Trading Exchanges to different global markets. Increase in profits, ,increase in market liquidity, reduction in bid ask spread, reduction of latency (expanded automatic execution and reduced the execution time) are all dependent on Proximity hosting which we shall carry as out independent variable.

3.6 Model design

3.6.1Market microstructure and order matching process

Several scholars document that algorithmic trading, in general, and high-frequency trading (HFT), specifically, are positively related to liquidity (e.g., **Brogaard**, 2010;**Hasbrouck** & **Saar**, 2012; **Hendershott**, **Jones**, & **Menkveld**, 2011; **Hendershott** &**Riordan**, 2009).Currently, exchanges have modified their market structure to attract more algorithmic trading. As **Hendershott** (2011, p. 31) point out, one way of doing so is by permitting "algorithmic traders to co-locate their servers in the market's data center."As stated before,Co-location reduces latency (the time it takes to make and execute trading decisions) and therefore enables co-located HFT market makers to more rapidly adjust their quotes as market conditions change. It also levels the playing field among competing HFT market makers who are also co-located by ensuring that none has a latency advantage over the other from an exchange perspective. Consequently, the introduction of co-location facilities is expected to increase liquidity by encouraging HFT and increase competition among HFT market makers.

In this study, we test the proposition that HFT increased and latency reduced following the introduction of co - location on the BSE by examining various proxies of the amount of HFT activity. We also directly test whether the introduction of co - location is associated with an improvement in Market liquidity, efficiency and depth.

We survey data for the four most actively traded futures contracts on the BSE before and after the introduction of co-location, to determine the effects on Market liquidity, efficiency and depth of the decision to allow co-location near the exchange server.We wish to survey :Share Price Index (SPI), 90-day Bank Accepted Bills (BABs), 3-year Government Bond, and 10-year Government Bond futures contracts.

Trading volume, the order - to - trade ratio, message traffic, and depth increased for the three interest rate futures contracts following co - location. On the other hand,trading volume, the order - to - trade ratio and message traffic fell while depth increased for the SPI contract following co - location. We tend to assume that this is caused by the tax on cash market equity message traffic that raised the cost and lowered the benefits of stock index arbitrage. In line with Chaboud, Hjalmarsson, Vega, and Chiquoine (2011), we do not observe any change in volatility after the introduction of co - location. This is contrary to Boehmer, Fong,and Wu (2012), who report evidence of increased short run volatility after the introduction of co - location and Hasbrouck and Saar (2012) and Hagströmer and Nordén (2013) who report a decrease in volatility when algorithmic trading increases. We also record economic significance of the increase in market liquidity after co - location.

For the purpose of this paper and for us to be in a position to demonstrate the benefits of introducing Colocation/proximity hosting, we will consider three related measures: (1) **Message Traffic**, (2) **Order-to-Trade Ratio**, and (3) **Algo Trade**. Boehmer (2012) and Hendershott (2011) highlight message traffic can include new order submissions, modifications, or order cancellations.

Order-to-Trade Ratio for contract ***i*** on day ***t*** is calculated as:

$$Order\text{-}to\text{-}Trade\ Ratio_{it} = \frac{Message\ Traffic_{it}}{Total\ Transactions_{it}}. \quad (1)$$

Hendershott (2011) suggest normalizing message traffic by trading volume, consistent with their approach we examine, Algo Trade calculated as follows for contract ***i*** on day ***t***:

$$Algo\ Trade_{it} = \frac{-Volume_{it}/100}{Message\ Traffic_{it}}. \quad (2)$$

To test the impact of the introduction of co-location on BSE we estimate for each futures contract (which can also be extrapolated to other Stocks)

$$Y_{it} = \alpha_i + \beta_i Colo_{it} + \varepsilon_{it}, \quad (3)$$

Where;

***Colo**$_{it}$* is a dummy variable set to zero prior to the introduction of co-location facilities on and set to one following its introduction.

***Y**$_{it}$* includes our proxies of the degree of algorithmic trading or a series of market quality and liquidity metrics, such as: bid–ask spreads, depth, open interest, trading activity, and volatility

The variable ***Depth***, aggregates over the trading day the average available total depth at the best bid and ask quotes. Trade Size, measures the average daily trade parcel, while ***Transactions*** counts the number of trades. Volume measures the total *size* of trade for trading day ***t***, Open Interest measures the ***open interest*** at the start of the trading day, while Volatility indicates contract ***i***'s ***price volatility*** on day ***t***, proxied by the daily price movement.

$$\ln\left(\frac{High_{it}}{Low_{it}}\right). \tag{4}$$

The main objective of our investigation is to identify the relation between increase in market liquidity, efficiency and depth, reduction in latency etc(algorithmic trading (AT)-Algo Trade) and Co-location. All factors kept constant, it is expected the introduction of co-location should increase the quantity of algorithmic trading.The securities Regulatory Commission might introduce trading fee to recover expanded market supervision costs due to introduction of Co-location. To test this proposition we estimate the following regression for each futures contract ***i***:

$$AT_{it} = \alpha_i + \beta_i Colo_{it} + \varphi_i CRC_{it} + \delta_i Open\ Interest_{it} + \gamma_i Volatility_{it} + \varepsilon_{it}, \tag{5}$$

Where;
AT_{it}, is a proxy for algorithmic trading: Message Traffic, Order-to-Trade Ratio, or AlgoTrade.
$Colo_{it}$ is a dummy variable set to zero prior to the introduction of co-location facilities and set to one following its introduction.
CRC_{it} is a dummy variable set to zero prior to the introduction of the cost recovery charge and set to one following its introduction.

In this paper, our aim is to test the impact of co-location (reduced latency) on the HFT market (market liquidity, market volatility, market depth, stock prices bid ask spread).To evaluate and measure the effect of co-location we estimate the following regression for our four measures of market liquidity as a proxy;

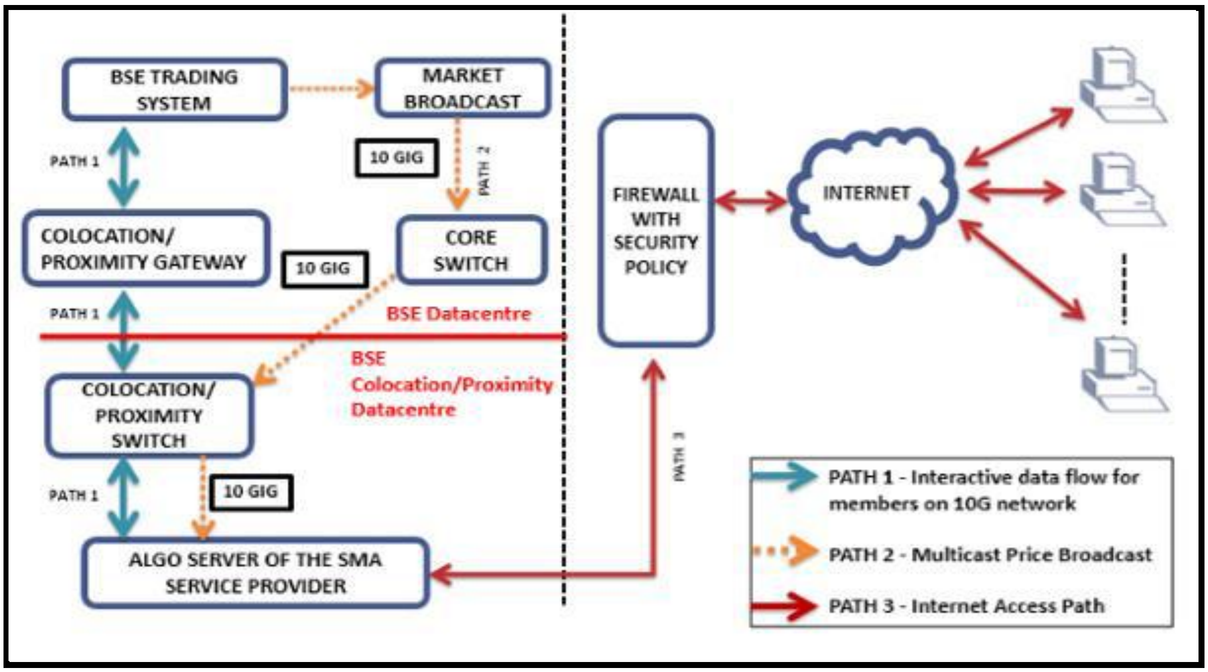
$$Liquidity_{it} = \alpha_i + \delta_i Open\ Interest_{it} + \gamma_i Volatility_{it} + \beta_i Colo_{it} + \varepsilon_{it}, \tag{6}$$

Where;
$Liquidity_{it}$ is proxied via: dollar tick spread, percentage tick spread, percentage of trades at the minimum tick, and total depth at the best prevailing quotes. In-line with the scholars that seeks to explain the determinants of liquidity, we control for price volatility and the level of trading activity at HFT.
Price volatility measures the risk to liquidity providers per unit of time. Hence, the higher the price volatility, the greater the compensation sought by liquidity providers. To control for the level of trading open interest is introduced as a control for the level of liquidity.

3.6.2 Co-location Design and architecture at BSE

This research paper seeks to study the benefits of Colocation and Proximity hosting at BSE in India. According to a report on Algo Trading, India boasts of the fastest colocation in the world. BSE is the fastest Exchange in the world with a speed of 6 microseconds. The following architectural design show how colocation is done at BSE.

Figure 3.1



Exchange provides open access to all members to setup in the Exchange colocation. Trading members do not have to incur any cost for setting up servers in BSE co-location. The Exchange provides both rack space and servers to the members.

In order to facilitate usage of newer technology/development amongst the market participants, BSE has launched a Technology Programme. The details are given below:

Under the Technology Programme, BSE bears the cost on behalf of the member for:

- The software license cost for Algo trading and market access. Algo software is provided to

members for trading from their Co-location and Non Co-location sites registered with Exchange.

- The infrastructure cost for Co-location rack space
- The Algo trading servers at Co-location
- The internet connectivity between BSE Co-location and member's office

With this technological assistance in place, BSE Levels the playing field among competing HFT market makers who are also co-located by ensuring that none has a latency advantage over the other from an exchange perspective.
This formed a good case study to simulate the benefits of co-Location to HFT market makers since all of them have a level playing field.

3.6.2 Sample order Matching system Design

According to capital.com,order matching systems were first introduced in the United States in the early 1980s. In 1982, the Chicago Stock Exchange (then the Mid West Stock Exchange) launched a MAX system to provide automated stock order execution.

An order matching system is an electronic system that matches buy and sell orders for a security on a stock market, a commodity on a commodity market and any other electronically traded financial instruments

This research paper aims at based on the model above, develop a sample order matching system which can simulate the benefits of co-location using a sample HFT strategy.

Figure 3.2 Sample Order Matching System