Technische Universität München
Lehrstuhl für Kommunikationsnetze
Prof. Dr.-Ing. Wolfgang Kellerer

Master‘s Thesis

VM Selection Heuristic for Financial Exchanges on the Cloud

Author:	Duclos-Cavalcanti, Daniel
Address:	230 W 55th Street
	10019 New York, NY
	U.S.A.
Matriculation Number:	03692475
Supervisor:	Muhamadd Haseeb & Navidreza Asadi
Begin:	03. April 2024
End:	03. October 2024

With my signature below, I assert that the work in this thesis has been composed by myself independently and no source materials or aids other than those mentioned in the thesis have been used.

München, 11.06.2014	
Place, Date	Signature

München, 11.06.2014	
Place, Date	Signature

Kurzfassung

A short abstract of the thesis in German.

Abstract

A short abstract of the thesis in English.

Contents

Chapter 1

Introduction

This chapter should give a short overview over the whole thesis. It should provide background information on the thesis topic, introduce the task definition and give a short outlook on the rest of the thesis.

Chapter 2

Background

In this chapter we discuss the required background on which this work builds upon. We start by discussing a quick overview of financial exchanges, the challenges to their migration to the cloud and finally the current state-of-the-art in the matter.

2.1 Financial Exchanges

Financial exchanges are modern redesign of a well known concept – the marketplace. By establishing a common platform to exchange goods, market participants are allowed to undergo the continuous process of bidding, an offer to buy an asset at a specified price, and asking, the intention to sell an asset at a specified price, for a given product. This procedure is called price discovery and serves to establish any asset's eventual true price. A true and fair price benefits both buyers and sellers, since a seller wishes to obtain just compensation, as well as a buyer wishes to pay nothing beyond what is needed.

Modern financial exchanges accomplish that through highly-engineered infrastructures composed of on-premise data centers, designed to provide a fair digital marketplace among participants co-located in the facility. A financial exchange's main components are summarized in the following segments and can also be seen as a diagramm on Figure 2.1

Orders. Trading orders can be either bids or asks. Bid orders are intents to buy an asset for a specified price, while ask orders indicate an offer to sell an asset at a specified price. Additionally, another relevant form of data is what is called a limit order book (**LOB**). A LOB is the amalgamation of all bids and asks concerning a specific asset and therefore summarize said asset's current market evaluation. Figure 2.2 illustrates an example of this data structure as presented to MPs.

Central Exchange Server. At the core of an exchange lies a matching engine (**ME**) that receives and processes trading orders from market participants. Orders are sent to the CES, which in return may or may not match an incoming order with a suitable counterpart.

Arriving orders, matched orders as well as periodic snapshots of the market (LOBs) are regular events that are broadcasted from the CES to the existing MPs.

Market Participants. Market participants are the co-located traders that are directly connected to an exchange's system, issuing orders and receiving data from the exchange.

Gateways. The gateways are structures placed between traders and the CES. Their responsabilities consist of routing orders and market data, as well as protecting the CES from abuse, e.g., unauthenticated or invalid orders. Within the on-premise clusters gateways are made to be equidistant to the CES as to ensure fairness regarding communication delays between the gateways and the central exchange.

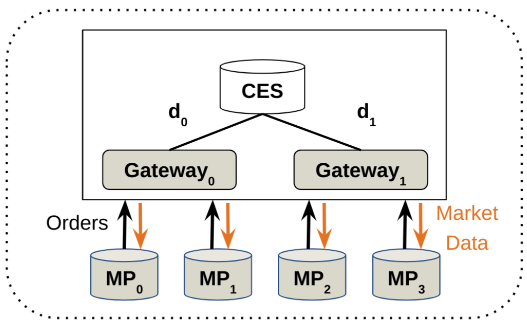

Figure 2.1: On-premise Infrastructure

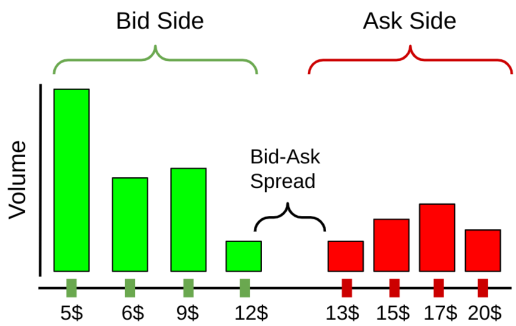

Figure 2.2: Limit Order Book

2.2 Migrating Financial Exchanges to the Cloud

Given the demanding network performance of financial exchanges, the use of on-premise data centers is the prevailing standard. However, the cloud displays well-known advantages as a medium for communication. These include flexibility, scalability, robustness and potential cost savings. On the other hand, the public cloud also exhibits high latency variance [HCW+24]. Additionally, it does not provide low-level engineering control to cloud-tenants such as switch-based multicast. As a result, cloud-based systems implement multicasting by directly unicasting a copy of a message to each recipient [GSPR22, GGS+21, GLR+24, GGM+23].

A fundamental characteristic of a financial exchange is the ability to disseminate data to market participants in a fair manner. That is, all MPs receive updates to the market almost simultaneously, as to not allow unwanted arbitrage opportunities. Without native solutions to enable multicasting and the significant serial delay of direct unicasting, a clear problem is presented to achieve the possible migration of financial exchanges to the public cloud.

2.3 State of the Art

2.3.1 Jasper: Fair Multicast for Fin. Exchanges on the Cloud

A siginificant project in this domain is Jasper [HGB+24], which offers a modern implementation of a financial exchange on the public cloud. This work realizes a multicast service that achieves low latency of less than 250 µs and a latency difference of less than 1 µs across 1,000 multicast receivers. To accomplish this, a series of techniques are employed.

Overlay Proxy Tree. This is the starting point for Jasper's design and borrow's the concept of trees to scale communication to a larger number of receivers, while improving latency in comparison to direct unicasting [CEM07, SST09]. The structure of the tree with Depth (**D**) and fan-out (**F**) has been empirically established as seen in Equations [2.1, 2.2].

$$F = 10 \tag{2.1}$$

$$D = \log_{10}(N) \tag{2.2}$$

Clock Synchronization. [GLY+18]

Chapter 3

Implementation/Results

3.1 Implementation

At the core of the tool.

3.2 Results

Results of the performed investigations are presented here. Interpretations for the observed effects are given and the impact of investigations is discussed.

Chapter 4

Conclusions and Outlook

The thesis is concluded here. The considered problem is repeated. The contribution of this work is highlighted and the results are recapitulated. Remaining questions are stated and ideas for future work are expressed.

Appendix A

Artifacts

The appendix may contain some listings of source code that has been used for simulations, extensive proofs or any other things that are strongly related to the thesis but not of immediate interest to the reader.

Appendix B

Abbreviations

CES	Central Exchange Server
MP	Market Participant

Bibliography

[CEM07] Tatsuhiro Chiba, Toshio Endo, and Satoshi Matsuoka. High-performance mpi broadcast algorithm for grid environments utilizing multi-lane nics. In *Seventh IEEE International Symposium on Cluster Computing and the Grid (CCGrid '07)*, pages 487–494, 2007.

[GGM+23] Prateesh Goyal, Eashan Gupta, Ilias Marinos, Chenxingyu Zhao, Radhika Mittal, and Ranveer Chandra. Dbo: Response time fairness for cloud-hosted financial exchanges, 2023.

[GGS+21] Ahmad Ghalayini, Jinkun Geng, Vighnesh Sachidananda, Vinay Sriram, Yilong Geng, Balaji Prabhakar, Mendel Rosenblum, and Anirudh Sivaraman. Cloudex: a fair-access financial exchange in the cloud. In *Proceedings of the Workshop on Hot Topics in Operating Systems*, HotOS '21, page 96–103, New York, NY, USA, 2021. Association for Computing Machinery.

[GLR+24] Junzhi Gong, Yuliang Li, Devdeep Ray, KK Yap, and Nandita Dukkipati. Octopus: A fair packet delivery service, 2024.

[GLY+18] Yilong Geng, Shiyu Liu, Zi Yin, Ashish Naik, Balaji Prabhakar, Mendel Rosunblum, and Amin Vahdat. Exploiting a natural network effect for scalable, fine-grained clock synchronization. In *Proceedings of the 15th USENIX Conference on Networked Systems Design and Implementation*, NSDI'18, page 81–94, USA, 2018. USENIX Association.

[GSPR22] Jinkun Geng, Anirudh Sivaraman, Balaji Prabhakar, and Mendel Rosenblum. Nezha: Deployable and high-performance consensus using synchronized clocks. *Proceedings of the VLDB Endowment*, 16(4):629–642, December 2022.

[HCW+24] Owen Hilyard, Bocheng Cui, Marielle Webster, Abishek Bangalore Muralikrishna, and Aleksey Charapko. Cloudy forecast: How predictable is communication latency in the cloud? In *2024 33rd International Conference on Computer Communications and Networks (ICCCN)*, pages 1–9. IEEE, 2024.

[HGB+24] Muhammad Haseeb, Jinkun Geng, Ulysses Butler, Xiyu Hao, Daniel Duclos-Cavalcanti, and Anirudh Sivaraman. Jasper: Scalable and fair multicast for financial exchanges in the cloud, 2024.

[SST09] Peter Sanders, Jochen Speck, and Jesper Larsson Träff. Two-tree algorithms for full bandwidth broadcast, reduction and scan. *Parallel Comput.*, 35(12):581–594, dec 2009.