execution requests (i.e., execution latencies), in addition to (or as part of) communications latencies.

Summary

[0006a] Aspects of the invention are defined in the independent claims. Some optional features of the invention are defined in the dependent claims.

[0007] In various aspects the present disclosure provides systems, methods, and computer-executable instruction mechanisms (e.g., non-transient machine-readable programming structures) such as software-coded instruction sets and data, for the management of data processing by multiple networked computing resources. In particular, for example, the present disclosure provides systems, methods, and media useful in controlling the synchronization or coordination of related requests for processing of data using distributed network resources.

[0008] For example, in one aspect the present disclosure provides systems, methods, and media for coordinating processing of data by multiple networked computing resources. Such systems, for example, can include at least one processor configured to: receive from one or more data sources signals representing instructions for execution of at least one data process executable by a plurality of networked computing resources; divide the at least one data process into a plurality of data processing segments, each data processing segment to be routed to a different one of a plurality of networked execution processors; based at least partly on latencies in execution of prior data processing requests routed by the system to each of the plurality of networked execution processors, determine a plurality of timing parameters, each of the plurality of timing parameters to be associated with a corresponding one of the plurality of data processing segments, the plurality of timing parameters determined to cause synchronized execution of the plurality of data processing segments by the plurality of networked execution processors; route the plurality of data processing segments to the plurality of corresponding networked execution processors in a timing sequence based on the timing

parameters; determining a capture ratio for the data processing segments; and adjusting the timing parameters associated with each of the plurality of networked execution processors based on the capture ratio.

[0009] In some aspects the present disclosure provides systems, methods, and programming or other machine-interpretable instructions for causing synchronized/coordinated processing of data by multiple networked computing resources, such systems, for example, comprising at least one processor configured to execute machine-interpretable instructions and causing the system to: monitor execution of signal processing execution requests by each of the plurality of networked computing resources; determine at least one timing parameter associated with a latency in execution of signal processes between the system and each of the plurality of networked computing resources; and store the at least one timing parameter in machine-readable memory accessible by the at least one processor.

[0010] Monitoring of execution of signal processing execution requests according to such and other embodiments of the disclosure can be implemented on continual, periodic, and/or other suitable or desirable bases.

[0011] In various embodiments of the various aspects of the disclosure, the networked computing resources can include one or more exchange servers. The data sources can include one or more broker or trader systems or servers, the controlled signal processes can represent trades in financial interests, and the execution of signal processing execution requests represents the execution of transactions in financial interests, including for example stocks, bonds, options and contract interests, currencies and/or other intangible interests, and/or commodities. In such embodiments requests for execution of data processing procedures can be based wholly or partially on parameters including, for example, any one or more of current market data quotations, order routing rules, order characteristics, displayed liquidity of each networked computing resource, and a probable delay, or latency, in execution of an order quantity at each networked computing resource.

[0012] In the same and further aspects the disclosure provides systems for controlling or otherwise managing requests for processing of data by distributed computer resources, such systems including one or more processors configured to execute instructions for causing the system to: monitor execution of signal processing execution requests by each of the plurality of networked computing resources; determine at least one timing parameter associated with the latency in execution of signal processes between the system and each of the plurality of networked computing resources; and store the at least one timing parameter for each of the plurality of networked computing resources.

[0013] In accordance with one aspect, there is provided a system for coordinating processing of data by multiple networked computing resources. The system includes at least one processor configured to: monitor data associated with a plurality of networked computing resources, the monitored data including data associated with data processing segments previously routed to the plurality of networked computing resources; receive from one or more data sources signals representing instructions for execution of at least one data process executable by the plurality of networked computing resources; based on the monitored data: divide the at least one data process into at least one data processing segment, each data processing segment to be routed to one of the plurality of networked computing resources; determine a plurality of timing parameters, each of the plurality of timing parameters to be associated with a corresponding one of the plurality of networked computing resources, the plurality of timing parameters determined to cause synchronized execution of the at least one data processing segment by the plurality of networked computing processors; and route the at least one data processing segment to the plurality of corresponding networked computing processors in a timing sequence based on the timing parameters.

[0014] In accordance with another aspect there is provided a method for coordinating processing of data by multiple networked computing resources.

What is claimed is:

1. A system for coordinating processing of data by multiple networked computing resources, the system comprising at least one processor configured to:

monitor data associated with a plurality of networked computing resources, the monitored data including data associated with data processing segments previously routed to the plurality of networked computing resources, wherein the data associated with the data processing segments previously routed to the plurality of networked computing resources includes data representing a multimodal distribution of historical execution latencies associated with the plurality of networked computing resources;

receive from one or more data sources signals representing instructions for execution of at least one data process executable by the plurality of networked computing resources;

based on the monitored data including the data representing the multimodal distribution of historical execution latencies:

divide the at least one data process into at least one data processing segment, each data processing segment to be routed to one of the plurality of networked computing resources, the respective data processing segments associated with at least one mode in the multi-modal distribution;

determine a plurality of timing parameters, each of the plurality of timing parameters to be associated with a corresponding one of the plurality of data processing segments, the plurality of timing parameters determined to cause synchronized execution of the at least one data processing segment by the plurality of networked computing resources; and

route the at least one data processing segment to the plurality of corresponding networked computing resources in a timing sequence based on the timing parameters.

2. The system of claim 1, wherein monitoring the data associated with the plurality of networked computing resources includes:

determining components of a latency associated with routing a data processing segment to one of the plurality of networked computing resources,

wherein determining the components of the latency is based on differences in latencies associated with routing different types of data processing segments to the one of the plurality of networked computing resources.

3. The system of claim 2, wherein the components of the latency include at least one range of random delay lengths introduced by at least one of the plurality of networked computing resources.

4. The system of claim 1, wherein dividing the at least one data process into at least one data processing segment includes:

selecting to which of the plurality of networked computing resources at least one of the plurality of data processing segments is to be routed; and

for each of the selected networked computing resources, determining a size of the corresponding at least one data processing segments.

5. The system of claim 4, wherein selecting to which of the plurality of networked computing resources and determining the size of the corresponding at least one data processing segments are based on a variance of the historical execution latencies for the plurality of networked computing resources.

6. The system of claim 5, wherein dividing the at least one data process into a plurality of data processing segments includes dividing the at least one data process into at least two data processing segments to be routed to the same networked computing resource, wherein the at least one processor is configured to determine the timing parameters for each of the at least two data processing segments.

7. The system of claim 5, wherein the plurality of data processing segments include at least two data processing segments to be routed to a particular networked computing resource based on the multimodal distribution, and wherein determining the plurality of timing parameters includes determining the timing parameters for the at least two data processing segments based on the multimodal distribution.

8. The system of claim 1, wherein the at least one processor is configured to:

determine whether liquidity in addition to posted liquidity is historically available at a particular networked computing resource,

and wherein dividing the at least one data process includes dividing the at least one data process into at least two data processing segments to be routed to the particular networked computing resource, the at least two data processing segments having timing parameters to target the posted available trade liquidity and the additional liquidity.

9. A method for coordinating processing of data by multiple networked computing resources, the method comprising:

monitoring data associated with a plurality of networked computing resources, the monitored data including data associated with data processing segments previously routed to the plurality of networked computing resources, wherein the data associated with the data processing segments previously routed to the plurality of networked computing resources includes data representing a multimodal distribution of historical execution latencies associated with the plurality of networked computing resources;

receiving from one or more data sources signals representing instructions for execution of at least one data process executable by the plurality of networked computing resources;

based on the monitored data including the data representing the multimodal distribution of historical execution latencies:

dividing the at least one data process into at least one data processing segment, each data processing segment to be routed to one of the plurality of networked computing resources, the respective data processing segments associated with at least one mode in the multi-modal distribution;

determining a plurality of timing parameters, each of the plurality of timing parameters to be associated with a corresponding one of the plurality of data processing segments, the plurality of timing parameters determined to cause synchronized execution of the at least one data processing segment by the plurality of networked computing resources; and

routing the at least one data processing segment to the plurality of corresponding networked computing resources in a timing sequence based on the timing parameters.

10. The method of claim 9, wherein monitoring the data associated with the plurality of networked computing resources includes:

determining components of a latency associated with routing a data processing segment to one of the plurality of networked computing resources,

wherein determining the components of the latency is based on differences in latencies associated with routing different types of data processing segments to the one of the plurality of networked computing resources.

11. The method of claim 10, wherein the components of the latency include at least one range of random delay lengths introduced by at least one of the plurality of networked computing resources.

12. The method of claim 9, wherein dividing the at least one data process into at least one data processing segment includes:

selecting to which of the plurality of networked computing resources at least one of the plurality of data processing segments is to be routed; and

for each of the selected networked computing resources, determining a size of the corresponding at least one data processing segments.

13. The method of claim 12, wherein selecting to which of the plurality of networked computing resources and determining the size of the corresponding at least one data processing segments are

based on a variance of the historical execution latencies for the plurality of networked computing resources.

14. The method of claim 13, wherein dividing the at least one data process into a plurality of data processing segments includes dividing the at least one data process into at least two data processing segments to be routed to the same networked computing resource, wherein the at least one processor is configured to determine the timing parameters for each of the at least two data processing segments.

15. The method of claim 13, wherein the plurality of data processing segments include at least two data processing segments to be routed to a particular networked computing resource based on the multimodal distribution, and wherein determining the plurality of timing parameters includes determining the timing parameters for the at least two data processing segments based on the multimodal distribution.

16. The method of claim 9, wherein the method includes:

determining whether liquidity in addition to posted liquidity is historically available at a particular networked computing resource,

and wherein dividing the at least one data process includes dividing the at least one data process into at least two data processing segments to be routed to the particular networked computing resource, the at least two data processing segments having timing parameters to target the posted available trade liquidity and the additional liquidity.

17. A computer program or computer program product comprising instructions, or a non-transitory computer-readable medium having stored thereon instructions, which, when executed by at least one processor, configure the at least one processor to:

monitor data associated with a plurality of networked computing resources, the monitored data including data associated with data processing segments previously routed to the plurality of networked computing resources, wherein the data associated with the data processing segments previously routed to the plurality of networked computing resources includes data representing a multimodal distribution of historical execution latencies associated with the plurality of networked computing resources;

receive from one or more data sources signals representing instructions for execution of at least one data process executable by the plurality of networked computing resources;

based on the monitored data including the data representing the multimodal distribution of historical execution latencies:

divide the at least one data process into at least one data processing segment, each data processing segment to be routed to one of the plurality of networked computing resources, the respective data processing segments associated with at least one mode in the multi-modal distribution;

determine a plurality of timing parameters, each of the plurality of timing parameters to be associated with a corresponding one of the plurality of data processing segments, the plurality of timing parameters determined to cause synchronized execution of the at least one data processing segment by the plurality of networked computing resources; and

route the at least one data processing segment to the plurality of corresponding networked computing resources in a timing sequence based on the timing parameters.

18. The computer program, computer program product or non-transitory computer-readable medium of claim 17, wherein monitoring the data associated with the plurality of networked computing resources includes:

determining components of a latency associated with routing a data processing segment to one of the plurality of networked computing resources,

wherein determining the components of the latency is based on differences in latencies associated with routing different types of data processing segments to the one of the plurality of networked computing resources.

19. The computer program, computer program product or non-transitory computer-readable medium of claim 18, wherein the components of the latency include at least one range of random delay lengths introduced by at least one of the plurality of networked computing resources.

20. The computer program, computer program product or non-transitory computer-readable medium of claim 17, wherein the instructions configure the processor to:

determine whether liquidity in addition to posted liquidity is historically available at a particular networked computing resource,

and wherein dividing the at least one data process includes dividing the at least one data process into at least two data processing segments to be routed to the particular networked computing resource, the at least two data processing segments having timing parameters to target the posted available trade liquidity and the additional liquidity.