CLAIMS

What is claimed is:

[c1] 1. A computer-implemented method comprising:

receiving, by one or more configured computing systems of a configurable workflow service, configuration information for a defined workflow of a client of the configurable workflow service, the configuration information specifying multiple workflow components of the defined workflow and interconnections between at least some of the multiple workflow components;

provisioning, by the one or more configured computing systems, multiple computing nodes to each implement at least one of the multiple workflow components, at least one of the multiple computing nodes being provided by the configurable workflow service for use by clients of the configurable workflow service; and

implementing the defined workflow for the client by initiating performance of defined activities on the multiple computing nodes that correspond to the implemented workflow components, the defined activities including:

one or more first activities involving obtaining input data from one or more data sources external to the configurable workflow service;

one or more second activities involving producing intermediate results data by performing data manipulation operations that are based at least in part on the obtained input data; and

one or more third activities involving producing output data based at least in part on the intermediate results data of the second activities.

[c2] 2. The method of claim 1 further comprising:

providing a graphical user interface, wherein the client specifies by the configuration information via the graphical user interface by editing a graph that logically represents the defined workflow, each of the multiple workflow

components being indicated within the graph by at least one of multiple nodes of the graph; and

wherein the implementing of the defined workflow includes physically instantiating a copy of the graph.

[c3] 3. The method of claim 1 wherein, for one or more of the multiple computing nodes, the computing node is implemented as a cluster of multiple computer systems that perform at least one of the defined activities in a distributed manner.

[c4] 4. The method of claim 1 wherein one or more of the multiple computing nodes are provided by an online execution service external to the configurable workflow service, and wherein the provisioning of the one or more computing nodes includes performing, by the one or more configured computing systems, one or more interactions with the online execution service corresponding to the one or more computing nodes.

[c5] 5. The method of claim 4 wherein the provided one or more computing nodes are associated by the online execution service with an account of the client with the online execution service, and wherein the one or more interactions with the online execution service include providing information on behalf of the client to the online execution service regarding the account of the client.

[c6] 6. The method of claim 4 wherein the provided one or more computing nodes are associated by the online execution service with the configurable workflow service in a manner that prevents the client from obtaining access to the one or more computing nodes.

[c7] 7. The method of claim 1 wherein the multiple workflow components include one or more source components and one or more data manipulation components and one or more destination components, each of the one or more

source components having an associated data source and associated instructions for accessing at least some of the input data from the associated data source, each of the one or more data manipulation components having associated instructions for performing at least one of the data manipulation operations, and each of the one or more destination components having associated instructions of handling at least some of the output data that is produced.

[c8] 8. The method of claim 7 wherein the multiple workflow components further include multiple of the source components, and wherein at least one of the source components is supplied by the client to the configurable workflow service to enable access to a data source that is specific to the client.

[c9] 9. The method of claim 7 wherein the multiple workflow components further include multiple of the data manipulation components, and wherein at least one of the data manipulation components has an associated data manipulation operation that is supplied by the client to the configurable workflow service to enable performance of a defined activity that is specific to the client.

[c10] 10. A system, comprising:

one or more processors; and

one or more modules of a workflow service that, when executed by at least one of the one or more processors, cause the system to:

receive configuration information from the client specifying multiple workflow components and interconnections for the defined workflow;

provision multiple computing nodes based at least in part on the received configuration information to implement the multiple workflow components; and

implement the defined workflow for the client by initiating performance of defined activities on the multiple computing nodes that correspond to the implemented workflow components, the performance of the defined activities including obtaining input data from one or more data sources, performing data manipulation operations that are based at least in part on the

obtained input data, and producing output data based at least in part on the performed data manipulation operations.

[c11] 11. The system of claim 10, wherein the one or more modules further cause the system to:

receive indications of one or more criteria for initiating implementation of the defined workflow; and

before the implementing of the defined workflow, determine that the one or more criteria are satisfied.

[c12] 12. The system of claim 11, wherein the one or more criteria correspond to periodic implementation of the defined workflow at multiple times based on changing input data that is available at those multiple times, and wherein the one or more modules further cause the system to perform the implementing of the defined workflow at each of those multiple times.

[c13] 13. The system of claim 11, wherein at least one of one or more criteria is defined as a pre-condition associated with at least one of the workflow components.

[c14] 14. The system of claim 10, wherein the one or more modules further cause the system to:

receive information about a post-condition criterion associated with one of the workflow components; and

determine that the post-condition criterion is satisfied after performing one or more of the defined activities that correspond to the one workflow component.

[c15] 15. The system of claim 10, wherein the one or more modules further cause the system to:

receive indications of one or more criteria for initiating implementation of the defined workflow;

determine that the one or more criteria are not satisfied; and

in response to determining that the one or more criteria are not satisfied, generate an alarm notification for the client to indicate that the one or more criteria are not satisfied.

CLAIMS

What is claimed is:

1. A system, comprising:

one or more processors; and

one or more modules of a configurable workflow service that are configured to, when executed by at least one of the one or more processors, provide functionality of a defined workflow for a client by:

receiving configuration information from the client specifying multiple workflow components and interconnections for the defined workflow, wherein one or more of the specified multiple workflow components are selected by the client from a plurality of predefined workflow components provided by the configurable workflow service;

provisioning multiple computing nodes based at least in part on the received configuration information to implement the multiple workflow components; and

implementing the defined workflow for the client by initiating performance of defined activities on the multiple computing nodes that correspond to the implemented workflow components, the performance of the defined activities including obtaining input data from one or more data sources, performing data manipulation operations that are based at least in part on the obtained input data, and producing output data based at least in part on the performed data manipulation operations.

2. The system of claim 1 further comprising a plurality of computing nodes for use in implementing workflows for multiple clients of the configurable workflow service, wherein at least some of the multiple

computing nodes are a subset of the plurality of computing nodes and are provided for use in the implementing of the defined workflow for the client in exchange for one or more fees paid by the client, and wherein the one or more modules include software instructions for execution by at least one of the one or more processors.

3. The system of claim 1 wherein the one or more modules consist of one or more means for performing the providing of the functionality of the defined workflow for the client.

4. A computer-implemented method comprising:

receiving, by one or more configured computing systems of a configurable workflow service, configuration information for a defined workflow of a client of the configurable workflow service, the configuration information specifying multiple workflow components of the defined workflow and interconnections between at least some of the multiple workflow components, wherein one or more of the specified multiple workflow components are selected by the client from a plurality of predefined workflow components provided by the configurable workflow service for use by clients of the configurable workflow service;

provisioning, by the one or more configured computing systems of the configurable workflow service, multiple computing nodes to each implement at least one of the multiple workflow components, at least one of the multiple computing nodes being provided by the configurable workflow service for use by clients of the configurable workflow service; and

implementing the defined workflow for the client by initiating, by the one or more configured computing systems of the configurable workflow service, performance of defined activities on the multiple computing nodes

that correspond to the implemented workflow components, the defined activities including:

one or more first activities involving obtaining input data from one or more data sources external to the configurable workflow service;

one or more second activities involving producing intermediate results data by performing data manipulation operations that are based at least in part on the obtained input data; and

one or more third activities involving producing output data based at least in part on the intermediate results data of the second activities.

5. The method of claim 4 wherein the implementing of the defined workflow for the client further includes at least one of storing the produced output data in one or more storage locations specified by the client, and generating one or more reports that are based at least in part of the produced output data and that are provided to the client.

6. The method of claim 4 further comprising:

receiving indications of one or more criteria for initiating implementation of the defined workflow; and

before the implementing of the defined workflow, determining, by the one or more configured computing systems, that the one or more criteria are satisfied.

7. The method of claim 6 wherein the one or more criteria correspond to periodic implementation of the defined workflow at multiple times based on changing input data that is available at those multiple times, and wherein the method further comprises performing the implementing of the defined workflow at each of those multiple times.

8. The method of claim 6 wherein at least one of one or more criteria is defined as a pre-condition associated with at least one of the multiple workflow components.

9. The method of claim 6 further comprising receiving information about a post-condition criterion associated with one of the multiple workflow components, and wherein the implementing of the defined workflow further includes determining that the post-condition criterion is satisfied after the performing of one or more of the defined activities that correspond to the one workflow component.

10. The method of claim 6 wherein the determining that the one or more criteria are satisfied occurs at a first time, and wherein the method further comprises:

determining, by the one or more configured computing systems, that the one or more criteria are not satisfied at a second time; and

in response, generating an alarm notification for the client to indicate that the one or more criteria are not satisfied.

11. The method of claim 4 further comprising:

providing, by the one or more configured computing systems, a graphical user interface of the configurable workflow service to the client, wherein the client specifies the configuration information via the graphical user interface by editing a graph that logically represents the defined workflow, each of the multiple workflow components being indicated within the graph by at least one of multiple nodes of the graph; and

wherein the implementing of the defined workflow includes physically instantiating a copy of the graph.

12. The method of claim 4 further comprising implementing the defined workflow for the client at multiple distinct times, and wherein one or more of the multiple computing nodes are maintained between the multiple distinct times to enable the provisioning of the one or more computing nodes to occur only once.

13. The method of claim 4 wherein, for one or more of the multiple computing nodes, the computing node is implemented as a cluster of multiple computer systems that perform at least one of the defined activities in a distributed manner.

14. The method of claim 4 wherein the provisioning of the multiple computing nodes includes, for each of the multiple computing nodes, loading at least one of a software program or configuration data on the computing node.

15. The method of claim 4 wherein one or more of the multiple computing nodes are provided by an online execution service external to the configurable workflow service, and wherein the provisioning of the one or more computing nodes includes performing, by the one or more configured computing systems, one or more interactions with the online execution service corresponding to the one or more computing nodes.

16. The method of claim 15 wherein the provided one or more computing nodes are associated by the online execution service with an account of the client with the online execution service, and wherein the one or more interactions with the online execution service include providing information

on behalf of the client to the online execution service regarding the account of the client.

17. The method of claim 15 wherein the provided one or more computing nodes are associated by the online execution service with the configurable workflow service in a manner that prevents the client from obtaining access to the one or more computing nodes.

18. The method of claim 4 wherein the multiple workflow components include one or more source components and one or more data manipulation components and one or more destination components, each of the one or more source components having an associated data source and associated instructions for accessing at least some of the input data from the associated data source, each of the one or more data manipulation components having associated instructions for performing at least one of the data manipulation operations, and each of the one or more destination components having associated instructions of handling at least some of the output data that is produced.

19. The method of claim 18 wherein the multiple workflow components further include multiple of the source components, and wherein at least one of the source components is supplied by the client to the configurable workflow service to enable access to a data source that is specific to the client.

20. The method of claim 18 wherein the multiple workflow components further include multiple of the data manipulation components, and wherein at least one of the data manipulation components has an associated data

manipulation operation that is supplied by the client to the configurable workflow service to enable performance of a defined activity that is specific to the client.

21. The method of claim 4 wherein the client is a customer of the configurable workflow service that provides one or more fees to the configurable workflow service in exchange for at least one of the provisioning of the multiple computing nodes or the implementing of the defined workflow for the client.

22. The method of claim 4 further comprising, for each of multiple additional clients of the configurable workflow service:

receiving configuration information for an additional defined workflow of the additional client;

storing the configuration information for an additional defined workflow of the additional client on one or more storage mechanisms of the configurable workflow service; and

implementing the additional defined workflow on behalf of the additional client on one or more computing nodes at one or more times.

23. A non-transitory computer-readable medium having stored contents that configure a computing system of a configurable workflow service to:

provision, by the configured computing system of the configurable workflow service, multiple computing nodes based on configuration information supplied for a defined workflow by a customer of the configurable workflow service, wherein the provisioning includes configuring the multiple computing nodes to implement multiple workflow components for the defined

workflow, the defined workflow including interconnections between at least some of the multiple workflow components, and wherein one or more of the multiple workflow components are selected by the customer from a plurality of predefined workflow components provided by the configurable workflow service; and

implement the defined workflow for the customer by initiating performance of defined activities on the multiple computing nodes that correspond to the implemented workflow components in exchange for one or more fees paid by the customer, the performance of the defined activities including obtaining input data from one or more data sources, performing data manipulation operations that are based at least in part on the obtained input data, and producing output data based at least in part on the performed data manipulation operations.

24. The non-transitory computer-readable medium of claim 23 wherein the configurable workflow service provides a plurality of computing nodes for use in implementing workflows for multiple clients of the configurable workflow service, wherein at least some of the multiple computing nodes are a subset of the plurality of computing nodes, and wherein the stored contents are software instructions that, when executed, program the configured computing system.

25. The non-transitory computer-readable medium of claim 23 wherein the computing system is further configured to, before the provisioning of the multiple computing nodes:

provide one or more interfaces for use by clients of the configurable network service in specifying information; and

receive configuration information from the customer for the defined workflow via at least one of the provided interfaces, wherein the received configuration information specifies the one or more workflow components selected by the customer,

and wherein the defined activities are performed in accordance with the received configuration information.

26. The non-transitory computer-readable medium of claim 23 wherein at least some of the multiple computing nodes are provided by one or more program execution services external to the configurable workflow service, and wherein the computing system is further configured to interact with the one or more program execution services on behalf of the customer as part of the provisioning of the multiple computing nodes.

27. The non-transitory computer-readable medium of claim 23 wherein at least some of the multiple computing nodes are provided by the customer via one or more computer systems under control of the customer at a location remote from the configured computing system, and wherein the computing system is further configured to interact with the one or more computer systems as part of the provisioning of the multiple computing nodes.