What is claimed is:

1. A method for combining specifications of dataflow graphs, the method including:

receiving over an input device or port a plurality of dataflow graph specifications, including:

a first dataflow graph specification that specifies two or more components connected by links representing flows of data between components, including a first component and a second component, with at least one component representing a computation applied to data flowing into a port of the component, and at least one component representing a computation providing data flowing out of a port of the component, and

a second dataflow graph specification that specifies at least a third component, and at least one sub-graph interface, where the sub-graph interface includes at least one flow junction representing a connection between: (1) a flow of data outside the sub-graph interface to or from a port of the third component, and (2) a flow of data inside the sub-graph interface to or from a port of a component of the first dataflow graph specification; and

processing, using at least one processor, information including the first dataflow graph specification and the second dataflow graph specification, to generate a combined dataflow graph specification, the processing including:

identifying an association between the sub-graph interface and the first dataflow graph specification,

for at least a first flow junction on the sub-graph interface, determining a direction associated with transferring a value of a descriptor of a data or computational characteristic, and

transferring a value of a descriptor of a data or computational characteristic from the first flow junction to a component specified by the first dataflow graph specification or a component specified by the second dataflow graph specification, according to the determined direction.

2. The method of claim 1, wherein the first dataflow graph specification includes at least one indicator that indicates that a descriptor associated with the first component is identical to a descriptor associated with the second component.

3. The method of claim 2, wherein the determined direction corresponds to an inward transfer of a value of a first descriptor of a data or computational characteristic from the first flow junction on the sub-graph interface to the second component.

4. The method of claim 3, wherein the first descriptor is provided to the first flow junction from the third component.

5. The method of claim 3 or 4, wherein the first descriptor is provided to the first component from the second component.

6. The method of claim 1, wherein the second dataflow graph specification includes at least one indicator that indicates that a descriptor associated with the third component is identical to a descriptor associated with the sub-graph interface.

7. The method of claim 6, wherein the second dataflow graph specification includes at least one indicator that indicates that a descriptor associated with a fourth component is identical to a descriptor associated with the third component.

8. The method of claim 7, wherein the determined direction corresponds to an outward transfer of a value of a first descriptor of a data or computational characteristic from the first flow junction on the sub-graph interface to the third component.

9. The method of claim 8, wherein the first descriptor is provided to the first flow junction from the first component.

10. The method of claim 8 or 9, wherein the first descriptor is provided to the fourth component from the third component.

11. The method of any of claims 1 to 10, wherein the descriptor describes a data characteristic of data transferred between a port of the first component and a port of the second component.

12. The method of any of claims 1 to 11, wherein the data characteristic includes a format of fields of records within the transferred data.

13. The method of any of claims 1 to 10, wherein the descriptor describes a computational characteristic of a component of at least one of the first dataflow graph specification or the second dataflow graph specification.

14. The method of any of claims 1 to 10 and 13, wherein the computational characteristic includes a degree of parallelism of execution of a computation represented by the component.

15. The method of any of claims 1 to 14, wherein the first dataflow graph is encrypted.

16. The method of any of claims 1 to 15, wherein the second dataflow graph is encrypted.

17. The method of any of claims 1 to 16, wherein the two or more components of the first dataflow graph are distributed among a first plurality of sequentially executing phases in the first dataflow graph.

18. The method of claim 17, wherein the second dataflow graph includes a plurality of components, the plurality of components and the sub-graph interface distributed among a second plurality of sequentially executing phases in the second dataflow graph.

19. The method of claim 17, wherein preparing the dataflow graph for execution includes determining a number of phases for inclusion in the dataflow graph based on the first plurality of sequentially executing phases and the second plurality of sequentially executing phases.

20. Software stored in a non-transitory form on a computer-readable medium, for combining specifications of dataflow graphs, the software including instructions for causing a computing system to perform the method of any of claims 1 to 19 or 22-24.

21. A computing system for combining specifications of dataflow graphs, the computing system including:

an input device or port configured to receive a plurality of dataflow graph specifications, including the first dataflow graph specification and the second dataflow graph specification

; and

at least one processor configured to perform the method of any of claims 1 to 19 or 22-24.

22. The method of claim 1, wherein generating the combined dataflow graph specification includes replacing the sub-graph interface with the first dataflow graph specification in the second dataflow graph specification.

23. The method of claim 1, wherein the sub-graph interface is a placeholder configured for association with a set of multiple dataflow graph specifications, the set including the first dataflow graph specification, each dataflow graph specification in the set of multiple dataflow graph specifications conforming to the sub-graph interface.

24. The method of claim 23, wherein each dataflow graph in the set of multiple dataflow graph specifications is an implementation of the sub-graph interface.