

CLAIMS

WHAT IS CLAIMED IS:

- 1 1. ~~— A computer implemented method comprising:~~
2 under the control of one or more computer systems configured with executable-
3 instructions;
4 receiving a data retrieval request to retrieve a data object, the data retrieval-
5 request specifying a data object identifier;
6 creating a data retrieval job corresponding to the data retrieval request;
7 providing a job identifier associated with the data retrieval job that is usable for-
8 obtaining information about the data retrieval job; and
9 after providing the job identifier, processing the data retrieval job using at least in-
10 part the data object identifier to provide access to the data object.
- 1 2. ~~— The computer implemented method of claim 1, wherein the data object~~
2 identifier is provided in response to a previous storage request to store the data object.
- 1 3. ~~— The computer implemented method of claim 1, wherein processing the~~
2 data retrieval job comprises:
3 selecting the data retrieval job for execution;
4 determining, based at least in part on the data object identifier, one or more-
5 storage entities for storing one or more encoded data components, the one or more encoded data-
6 components are generated based at least in part on the data object;
7 causing retrieval of at least some of the one or more encoded data components-
8 from the determined storage entities; and
9 decoding the retrieved encoded data components to obtain the retrieved data-
10 object.
- 1 4. ~~— The computer implemented method of claim 3, wherein selecting the data~~
2 retrieval job is based at least in part on a batch processing schedule.
- 1 5. ~~— The computer implemented method of claim 1, further comprising~~
2 providing a status of the data retrieval job in response to a job status request that specifies the job
3 identifier.

1 6. — A system for providing archival data storage services, comprising:
2 one or more archival data storage devices;
3 a transient data store;
4 one or more processors;
5 memory, including executable instructions that, when executed by the one or
6 ~~more processors, cause the one or more processors to collectively at least:~~
7 receive a data storage request to store a data object;
8 cause storage of the data object in the transient data store;
9 provide a data object identifier associated with the data, the data object identifier
10 ~~encoding at least storage location information usable to locate the data object; and~~
11 after providing the data object identifier, cause storage of the data object in
12 ~~location specified by the storage location information.~~

1 7. — The system of claim 6, wherein the executable instructions, when executed
2 ~~by the one or more processors, further cause the one or more processors to collectively create a~~
3 ~~data storage job corresponding to the data storage request and wherein causing storage of the~~
4 ~~data object in location specified by the storage location information comprises processing the~~
5 ~~data storage job based at least in part on the storage location information.~~

1 8. — The system of claim 6, wherein causing storage of the data object in
2 ~~location specified by the storage location information comprises:~~
3 obtaining the data object from the transient data store;
4 obtaining one or more data encoding schemes;
5 encoding the data object with the one or more data encoding schemes to obtain a
6 ~~plurality of encoded data components; and~~
7 causing storage of the plurality of encoded data components in at least some of
8 ~~the one or more archival data storage devices.~~

1 9. — The system of claim 6, wherein the data object identifier encodes at least
2 ~~data validation information usable to validate integrity of the data object.~~

1 10.—The system of claim 9, wherein the executable instructions, when executed
2 by the one or more processors, further cause the one or more processors to collectively validate
3 integrity of the data object based at least in part on the data validation information

1 11.—The system of claim 6, wherein the executable instructions, when executed
2 by the one or more processors, further cause the one or more processors to:
3 receive a data retrieval request for the data object cause aggregation of the data
4 retrieval request with other data retrieval requests to form a request batch; and
5 cause validation of the request batch based at least in part on data object
6 identifiers corresponding to the request batch.

1 12.—The system of claim 6, wherein the executable instructions, when executed
2 by the one or more processors, further cause the one or more processors to receive a data
3 retrieval request for the data object; and
4 cause processing of the data retrieval request based at least in part on a power-
5 management schedule.

1 13.—The system of claim 6, wherein the executable instructions, when executed
2 by the one or more processors, further cause the one or more processors to cause partitioning of
3 pending data retrieval requests to facilitate parallel processing.

1 14.—The system of claim 6, wherein the executable instructions, when executed
2 by the one or more processors, further cause the one or more processors to provide status of a
3 data retrieval request in response to a status request that specifies the data object identifier.

1 15.—The system of claim 6, wherein the executable instructions, when executed
2 by the one or more processors, further cause the one or more processors to cause validation of
3 data integrity for the data object.

CLAIMS

WHAT IS CLAIMED IS:

1. A computer-implemented method for providing cost- effective and durable archival data storage service, comprising:
 - under the control of one or more computer systems of an archival data storage system that are configured with executable instructions,
 - receiving, over a communications network from a requesting computer system, an electronic retrieval request to retrieve a data object from persistent storage, the electronic retrieval request encoding an identifier of the data object;
 - in response to receipt of the electronic retrieval request:
 - initiating a retrieval job to obtain the data object from persistent storage;
 - providing, to the requesting computer system, a retrieval job identifier for the retrieval job that is usable to obtain a status of the retrieval job and that, after the retrieval job is completed, is usable to retrieve the data object;
 - receiving, over the communications network, an electronic download request to download at least a portion of the data object, the electronic download request encoding the retrieval job identifier; and
 - in response to receipt of the electronic download request and as a result of the retrieval job having been completed, providing the at least a portion of the data object for download.
2. The computer-implemented method of claim 1, further comprising:
 - receiving an electronic request for a status of the retrieval job, the electronic request for the status encoding the retrieval job identifier; and
 - in response to receipt of the electronic request for the status of the retrieval job, providing a current status of the retrieval job.
3. The computer-implemented method of claim 1, wherein:
 - the download request includes information specifying a portion of the data object; and
 - providing the at least a portion of the data object includes providing the at least the portion.

4. The computer-implemented method of claim 3, wherein providing the at least the portion includes providing a data verification value for the at least the portion, the data verification value usable to determine whether the provided at least the portion is corrupted.

5. The computer-implemented method of claim 1, further comprising, as a result of the electronic retrieval request including a notification parameter, providing a notification upon completion of the retrieval job prior to receipt of the electronic download request.

6. A computer-implemented method comprising:
under the control of one or more computer systems configured with
executable instructions,

receiving a first electronic message comprising a first identifier
corresponding to data stored by a storage system;

in response to receipt of the first electronic message, initiating a job in
connection with the data;

at a time after receiving the first electronic message, receiving a second
electronic message comprising a second identifier corresponding to the job, where the
second identifier is associated with the job and useable to retrieve the data corresponding
to the first identifier; and

in response to receipt of the second electronic message, providing
information associated with the job;

receiving a third electronic message comprising the second identifier; and
in response to receipt of the third electronic message, providing at least a
portion of the data for download.

7. The computer-implemented method of claim 6, wherein the first
identifier and the second identifier are different.

8. The computer-implemented method of claim 6, wherein:
the data comprises one or more data objects contained in a logical data
container; and

the first identifier is an identifier for the logical data container.

9. The computer-implemented method of claim 8, wherein the information associated with the job includes a collection of one or more identifiers that each identifies one of the one or more data objects.

10. The computer-implemented method of claim 8, wherein the information associated with the job includes a log of one or more actions performed in connection with the one or more data objects contained in the logical data container.

11. The computer-implemented method of claim 6, wherein:
the first electronic message originated from a customer computer system of a customer of the storage system;
the first electronic message includes metadata provided by the customer computer system; and
the information associated with the job includes at least a portion of the metadata.

12. A system, comprising:
one or more processors; and
memory storing executable instructions that, when executed by the one or more processors, cause the system to:
receive a first electronic message comprising a first identifier
corresponding to data stored by a storage system;
in response to receipt of the first electronic message, initiate a job in connection with the data;
at a time after receiving the first electronic message, receive a second electronic message comprising a second identifier corresponding to the job, where the second identifier is associated with the job and useable to retrieve the data corresponding to the first identifier; and
in response to receipt of the second electronic message, provide information associated with the job;
receive a third electronic message comprising the second identifier; and

in response to receipt of the third electronic message, provide at least a portion of the data for download.

13. The system of claim 12, wherein the first identifier and the second identifier are different.

14. The system of claim 12, wherein:
the data comprises one or more data objects contained in a logical data container; and
the first identifier identifies the logical data container.

15. The system of claim 14, wherein the information associated with the job includes a collection of one or more identifiers that each identifies one of the one or more data objects.

16. The system of claim 14, wherein the information associated with the job includes a log of one or more actions performed in connection with the one or more data objects contained in the logical data container.

17. The system of claim 12, wherein:
the first electronic message originated from a customer computer system of a customer of the storage system;
the first electronic message includes metadata provided by the customer computer system; and
the information associated with the job includes at least a portion of the metadata.

18. The system of claim 12, wherein the system provides the at least the portion of the data for download by moving the at least the portion of the data to another storage device of the storage system.

19. The system of claim 12, wherein the instructions, when executed by the one or more processors, cause the system to provide a notification of completion of the job.

20. The system of claim 19, wherein the notification is provided as a

result of the first electronic message specifying a notification parameter.

21. A non-transitory computer-readable storage medium having stored thereon instructions that, when executed by one or more processors of a computer system, cause the computer system to:

in response to receipt of a first electronic message comprising a first identifier corresponding to data stored by a storage system, initiate a job in connection with the data;

at a time after receiving the first electronic message, in response to receipt of a second electronic message comprising a second identifier corresponding to the job, where the second identifier is associated with the job and useable to retrieve the data corresponding to the first identifier, provide information associated with the job; and

in response to receipt of a third electronic message comprising the second identifier, provide the at least a portion of the data for download.

22. The non-transitory computer-readable storage medium of claim 21, wherein the first identifier and the second identifier are different.

23. The non-transitory computer-readable storage medium of claim 21, wherein:

the data comprises one or more data objects contained in a logical data container; and

the first identifier identifies the logical data container.

24. The non-transitory computer-readable storage medium of claim 23, wherein the information associated with the job includes a collection of one or more identifiers that each identifies one of the one or more data objects.

25. The non-transitory computer-readable storage medium of claim 23, wherein:

the first electronic message originated from a customer computer system of a customer of the storage system;

the first electronic message includes metadata provided by the customer

computer system; and

the information associated with the job includes at least a portion of the metadata.

26. The non-transitory computer-readable storage medium of claim 21, wherein the instructions that cause the computer system to provide the at least the portion of the data for download, when executed, cause the computer system to provide the at least the portion of the data for download by moving the at least the portion of the data to another storage device of the storage system.

27. The non-transitory computer-readable storage medium of claim 21, wherein the system provides the at least the portion at least in part by providing a data verification value for the at least the portion, the data verification value usable to determine whether the provided specified portion is corrupted.

28. The non-transitory computer-readable storage medium of claim 21, wherein the information associated with the job indicates whether data has been stored in a component of the storage system from which the data is downloadable.

29. The non-transitory computer-readable storage medium of claim 21, wherein the instructions, when executed by the one or more processors, cause the system to provide a notification of completion of the job.

30. The non-transitory computer-readable storage medium of claim 29, wherein the notification is provided as a result of the first electronic message specifying a notification parameter.

CLAIMS

WHAT IS CLAIMED IS:

1. A computer-implemented method for providing cost-effective and durable archival data storage service, comprising:
 - under the control of one or more computer systems of an archival data storage system that are configured with executable instructions,
 - receiving, over a communications network from a requesting computer system, an electronic retrieval request to retrieve a data object from persistent storage, the electronic retrieval request encoding an identifier of the data object;
 - in response to receipt of the electronic retrieval request:
 - initiating a retrieval job to obtain the data object from persistent storage;
 - providing, to the requesting computer system, a retrieval job identifier for the retrieval job that is usable to obtain a status of the retrieval job and that, after the retrieval job is completed, is usable to retrieve the data object;
 - receiving, over the communications network, an electronic download request to download at least a portion of the data object, the electronic download request encoding the retrieval job identifier; and
 - in response to receipt of the electronic download request and as a result of the retrieval job having been completed, providing the at least a portion of the data object for download.
2. The computer-implemented method of claim 1, further comprising:
 - receiving an electronic request for a status of the retrieval job, the electronic request for the status encoding the retrieval job identifier; and
 - in response to receipt of the electronic request for the status of the retrieval job, providing a current status of the retrieval job.
3. The computer-implemented method of claim 1, wherein:
 - the download request includes information specifying a portion of the data object; and
 - providing the at least a portion of the data object includes providing the at least the portion.

4. The computer-implemented method of claim 3, wherein providing the at least the portion includes providing a data verification value for the at least the portion, the data verification value usable to determine whether the provided at least the portion is corrupted.

5. The computer-implemented method of claim 1, further comprising, as a result of the electronic retrieval request including a notification parameter, providing a notification upon completion of the retrieval job prior to receipt of the electronic download request.

6. A computer-implemented method comprising:
under the control of one or more computer systems configured with executable instructions,
receiving a first electronic message comprising a first identifier corresponding to data stored by a storage system;
in response to receipt of the first electronic message, initiating a job in connection with the data;
at a time after receiving the first electronic message, receiving a second electronic message comprising a second identifier corresponding to the job, where the second identifier is associated with the job and useable to retrieve the data corresponding to the first identifier; and
in response to receipt of the second electronic message, providing information associated with the job;
receiving a third electronic message comprising the second identifier; and
in response to receipt of the third electronic message, providing at least a portion of the data for download.

7. The computer-implemented method of claim 6, wherein the first identifier and the second identifier are different.

8. The computer-implemented method of claim 6, wherein:
the data comprises one or more data objects contained in a logical data container; and

the first identifier is an identifier for the logical data container.

9. The computer-implemented method of claim 8, wherein the information associated with the job includes a collection of one or more identifiers that each identifies one of the one or more data objects.

10. The computer-implemented method of claim 8, wherein the information associated with the job includes a log of one or more actions performed in connection with the one or more data objects contained in the logical data container.

11. The computer-implemented method of claim 6, wherein:
the first electronic message originated from a customer computer system of a customer of the storage system;
the first electronic message includes metadata provided by the customer computer system; and
the information associated with the job includes at least a portion of the metadata.

12. A system, comprising:
one or more processors; and
memory storing executable instructions that, when executed by the one or more processors, cause the system to:
receive a first electronic message comprising a first identifier corresponding to data stored by a storage system;
in response to receipt of the first electronic message, initiate a job in connection with the data;
at a time after receiving the first electronic message, receive a second electronic message comprising a second identifier corresponding to the job, where the second identifier is associated with the job and useable to retrieve the data corresponding to the first identifier; and
in response to receipt of the second electronic message, provide information associated with the job;
receive a third electronic message comprising the second identifier; and

in response to receipt of the third electronic message, provide at least a portion of the data for download.

13. The system of claim 12, wherein the first identifier and the second identifier are different.

14. The system of claim 12, wherein:
the data comprises one or more data objects contained in a logical data container; and
the first identifier identifies the logical data container.

15. The system of claim 14, wherein the information associated with the job includes a collection of one or more identifiers that each identifies one of the one or more data objects.

16. The system of claim 14, wherein the information associated with the job includes a log of one or more actions performed in connection with the one or more data objects contained in the logical data container.

17. The system of claim 12, wherein:
the first electronic message originated from a customer computer system of a customer of the storage system;
the first electronic message includes metadata provided by the customer computer system; and
the information associated with the job includes at least a portion of the metadata.

18. The system of claim 12, wherein the system provides the at least the portion of the data for download by moving the at least the portion of the data to another storage device of the storage system.

19. The system of claim 12, wherein the instructions, when executed by the one or more processors, cause the system to provide a notification of completion of the job.

20. The system of claim 19, wherein the notification is provided as a

result of the first electronic message specifying a notification parameter.

21. A non-transitory computer-readable storage medium having stored thereon instructions that, when executed by one or more processors of a computer system, cause the computer system to:

in response to receipt of a first electronic message comprising a first identifier corresponding to data stored by a storage system, initiate a job in connection with the data;

at a time after receiving the first electronic message, in response to receipt of a second electronic message comprising a second identifier corresponding to the job, where the second identifier is associated with the job and useable to retrieve the data corresponding to the first identifier, provide information associated with the job; and

in response to receipt of a third electronic message comprising the second identifier, provide the at least a portion of the data for download.

22. The non-transitory computer-readable storage medium of claim 21, wherein the first identifier and the second identifier are different.

23. The non-transitory computer-readable storage medium of claim 21, wherein:

the data comprises one or more data objects contained in a logical data container; and

the first identifier identifies the logical data container.

24. The non-transitory computer-readable storage medium of claim 23, wherein the information associated with the job includes a collection of one or more identifiers that each identifies one of the one or more data objects.

25. The non-transitory computer-readable storage medium of claim 23, wherein:

the first electronic message originated from a customer computer system of a customer of the storage system;

the first electronic message includes metadata provided by the customer

computer system; and

the information associated with the job includes at least a portion of the metadata.

26. The non-transitory computer-readable storage medium of claim 21, wherein the instructions that cause the computer system to provide the at least the portion of the data for download, when executed, cause the computer system to provide the at least the portion of the data for download by moving the at least the portion of the data to another storage device of the storage system.

27. The non-transitory computer-readable storage medium of claim 21, wherein the system provides the at least the portion at least in part by providing a data verification value for the at least the portion, the data verification value usable to determine whether the provided specified portion is corrupted.

28. The non-transitory computer-readable storage medium of claim 21, wherein the information associated with the job indicates whether data has been stored in a component of the storage system from which the data is downloadable.

29. The non-transitory computer-readable storage medium of claim 21, wherein the instructions, when executed by the one or more processors, cause the system to provide a notification of completion of the job.

30. The non-transitory computer-readable storage medium of claim 29, wherein the notification is provided as a result of the first electronic message specifying a notification parameter.