Claims:

1. A forensic analysis method performed on a Distributed Computing System comprising a server and at least one client machine, the method comprising:

collecting data in a client machine of the Distributed Computing System, the collected data being a system call to a resource comprised in the Distributed Computing System, the system call being collected by a collection tool from between a user space and a kernel space of the client machine;

filtering the collected data to form a first data set, filtering comprising extracting file related system call data but not registry, process and memory related system call data;

applying a data reduction model to the first data set to form a second data set; and

processing the second data set in the server of the Distributed Computing System to provide for detection of suspect behaviour at the client machine, in which

the data reduction model is configured to extract a subset of data from the first data set to form the second data set, the subset of data comprising: user account identifier; and process and object identifier, the process and object identifier being one of: a process and physical resource identifier; and a process and logical object identifier.

2. The forensic analysis method according to claim 1 further comprising forming a data transfer object in dependence on the second data set and storing the data transfer object in data storage comprised in the server.

3. The forensic analysis method according to any one of the preceding claims in which a number of data elements of the second data set is no more than half a number of data elements of the first data set.

4. The forensic analysis method according to claim 3 in which the number of data elements of the second data set is no more than 10% of the number of data elements of the first data set.

5. The forensic analysis method according to any one of the preceding claims in which the second data set consists of no more than ten data elements.

6. The forensic analysis method according to any one of the preceding claims in which the data reduction model is applied to the first data set in the client machine to form the second data set and the second data set is conveyed from the client machine to the server by way of a communication channel.

7. The forensic analysis method according to any one of claims 1 to 5 in which the first data set is conveyed from the client machine to the server by way of a communication channel and the data reduction model is applied to the first data set in the server to form the second data set.

8. The forensic analysis method according to any one of the preceding claims in which the subset of data extracted by the data reduction model further comprises a time element reflecting when a process of the process and object identifier was invoked.

9. The forensic analysis method according to any one of the preceding claims in which the subset of data extracted by the data reduction model further comprises a location identifier which reflects a location of one of: a software agent acting for a user account that initiates the system call; and an object operated on by a process of the process and object identifier.

10. The forensic analysis method according to claim 9 in which the location identifier is in respect of one of a physical location in the Distributed Computing System and a logical location in the Distributed Computing System.

11. The forensic analysis method according to any one of the preceding claims in which the step of processing the second data set in the server comprises analysing the second data set to detect suspect behaviour, the step of analysing the second data set to detect suspect behaviour comprising comparing the second data set with plural reference data sets and detecting suspect behaviour in dependence on the

comparison, the plural reference data sets representing at least one of normal behaviour and suspect behaviour.

12. The forensic analysis method according to any one of the preceding claims in which the step of processing the second data set in the server comprises applying the second data set to an expert system comprising a model of normal behaviour, the model of normal behaviour being formed in dependence on normal reference data which reflects normal behaviour of the Distributed Computing System.

13. The forensic analysis method according to claim 12 in which an outlier detection algorithm is configured to apply the second data set to the model of normal behaviour and to determine whether or not the second data set is in accordance with the model of normal behaviour.

14. The forensic analysis method according to claim 12 or 13 further comprising modifying the model of normal behaviour in a supervised manner in dependence on user interaction with the DCS.

15. The forensic analysis method according to any one of claims 12 to 14 in which the model of normal behaviour comprises plural submodels, each submodel being separately formed from a different form of reference data.

16. The forensic analysis method according to claim 15 in which each submodel is formed from reference data relating to one of: a different system call; and a different user.

17. The forensic analysis method according to any one of the preceding claims in which the client machine is one of: a PC; and a mobile device.

18. The forensic analysis method according to any one of the preceding claims being an at least near real time forensic analysis method.

19. A computer program comprising program instructions for causing a computer to perform the method according to any one of the preceding claims.

20. The computer program according to claim 19 in which the program instructions are at least one of: embodied on a record medium; embodied in read only memory; stored in computer memory; and carried on an electrical carrier signal.

21. A Distributed Computing System comprising a server and at least one client machine, the Distributed Computing System being configured: to perform forensic analysis by collecting data in a client machine, the collected data being a system call to a resource comprised in the Distributed Computing System, the system call being collected by a collection tool from between a user space and a kernel space of the client machine; to filter the collected data to form a first data set, filtering comprising extracting file related system call data but not registry, process and memory related system call data; to apply a data reduction model to the first data set to form a second data set; and to process the second data set in the server of the Distributed Computing System to provide for detection of suspect behaviour at the client machine, in which the data reduction model is configured to extract a subset of data from the first data set to form the second data set, the subset of data comprising: user account identifier; and process and object identifier, the process and object identifier being one of: a process and physical resource identifier; and a process and logical object identifier.