Intellectual Property Office

BL O/440/16
20 September 2016

PATENTS ACT 1977

APPLICANT	Cloudera, Inc.
ISSUE	Whether patent application GB1417283.7 complies with section 1(2)
HEARING OFFICER	H Jones

DECISION

Introduction

1 The patent application relates to a system and method for performing data queries in a distributed computing cluster comprising a set of computers connected to each other across a network. The application was published as GB2520411 on 20 May 2015 and claims priority from an earlier US application filed on 1 October 2013.

2 The issue to be decided is whether the invention as claimed in the application consists solely of a program for a computer which the Act excludes from patentability. The examiner dealing with the application considers that the invention is excluded from patentability under section 1(2)(c) as a program for a computer as such. The applicant disagrees, and requested a hearing to decide the matter. The hearing was held on 2 August 2016 at which the applicant was represented by Mr James Leach of Mewburn Ellis.

The invention

3 The invention is concerned with improving the speed of querying data stored in separate computers (nodes) in a distributed computing cluster. The specification explains that this is achieved by taking advantage of the speed of the schema-on-write (SOW) model for formatting datasets while retaining the flexibility of the schema-on-read (SOR) model, the SOW model being an arrangement where data is made to conform (or formatted) to a plan or schema when it is stored in a database and the SOR model being an arrangement where data is formatted to a plan or schema as it is pulled out of a stored location. This is achieved by providing schemas specifying target formats in which data is stored at data nodes of the cluster and saving data at the nodes in an original format and a target format. Data is transformed from the original format to a target format according to a particular schema and stored at the node either when a query is made (SOR), when the data is stored (SOW) or at some point in between as determined by a format conversion schedule. The format conversion schedule can take into account factors such as the occurrence of an event, e.g. the creation, initial update or last update of the data, or when the data has been updated or searched for a certain number of times.

4 When a query is made, a query planner reviews relevant schema information to identify the available file formats in which data is stored. If only data in an original format is available, the query planner defines plan fragments for the original format. If data in a converted target format is also available, the query planner defines plan fragments for the target format.

5 An amended set of claims was filed on 11 February 2016, having two independent claims (1 and 5) are considered to relate to the same invention. For the purpose of this decision I need only set out the wording of claim 1:

> A system for performing queries on stored data in a distributed computing cluster, wherein the system comprises:
>
> a plurality of data nodes, wherein each data node comprises, respectively, a format conversion engine configured to convert data stored at the data node from an original format into one or more target formats optimized for relational database processing according to a predetermined schedule so that, following conversion, the data is stored at the data node in both the original format and the one or more target formats;
>
> a metastore that includes one or more schemas specifying one or more target formats in which data is stored at the data nodes;
>
> wherein each data node further comprises, respectively, a query engine having:
>
> a query planner configured to parse a query from a client to create query fragments based on the one or more schemas in the metastore such that when data in a target format specified by a schema in the metastore is available, the query fragments are created for that target format, and so that when data in a target format specified by a schema in the metastore is not available, the query fragments are created for the original format;
>
> a query coordinator configured to distribute the query fragments among the plurality of data nodes; and
>
> a query execution engine comprising:
>
> a transformation module configured to, if the query fragments have been created for a target format into which data stored at the data node has not been converted by the format conversion engine, transform the data stored at the data node into the target format for which the query fragments have been created based on the relevant schema in the metastore; and
>
> an execution module configured to execute the query fragments on the data in the format for which the query fragments have been created to obtain intermediate results that are aggregated and returned to the client.

The law

6 The relevant provision of the Act in relation to excluded inventions is section 1(2), which reads:

> *1(2) It is hereby declared that the following (among other things) are not inventions for the purposes of this Act, that is to say, anything which consists of –*
>
> *(a) a discovery, scientific theory or mathematical method;*
> *(b) a literary, dramatic, musical or artistic work or any other aesthetic creation whatsoever;*
> *(c) a scheme, rule or method for performing a mental act, playing a game or doing business, or a program for a computer;*
> *(d) the presentation of information;*
>
> *but the foregoing provision shall prevent anything from being treated as an invention for the purposes of this Act only to the extent that a patent or application for a patent relates to that thing as such.*

7 The examiner considers that the invention relates to the field of computer programming and that it is potentially caught by the exclusion to patentability set out in section 1(2)(c). In order to decide whether an invention relates to subject matter excluded from patentability under section 1(2), the Court of Appeal has said that the issue must be decided by answering the question of whether the invention reveals a technical contribution to the state of the art (cf *Symbian*[1], *Aerotel*[2]). The Court of Appeal in Aerotel set out the following four-step test to help decide the issue:

> 1) construe the claim;
> 2) identify the actual (or alleged) contribution;
> 3) ask whether it falls solely within the excluded subject matter;
> 4) check whether the actual or alleged contribution is actually technical in nature.

Arguments and analysis

8 With regard to the first of the four steps, the examiner has stated that the claims are perfectly clear in the context of the description and so this step causes no difficulty. I agree.

9 With regard to the second step of identifying the actual (or alleged) contribution, agreement has been reached between the examiner and the applicant that the contribution is as defined in the examiner's letter dated 17 May 2016 (paragraph 3) and lies in:

> A method for performing queries on stored data in a distributed computing environment having a plurality of data nodes (the system of claim 1), comprises a format conversion engine at each data node configured to convert data from an original format into one or more target formats (optimised for relational database processing), following conversion the data is stored at the data node in both the original format and one or more target formats, query fragments can be executed more quickly at the query execution engines (located at the data nodes), since the transformation of data stored at the data node into a target format for which the query fragments have been created has already been done,

[1] Symbian Ltd. v Comptroller-General of Patents [2008] EWCA Civ 1066
[2] Aerotel Ltd v Telco Holdings Ltd and Macrossan's Application [2006] EWCA Civ 1371

yet queries can still be executed even when data in the target format is not available through a query planner creating query fragments in the original format (in such circumstances), this provides a user with the benefits of both the "schema-on-read" model and the "schema-on-write" models.

10 I agree with this assessment of the contribution.

11 The main area of disagreement between the examiner and applicant is in relation to the third and fourth steps of the test regarding whether the invention falls within excluded subject matter and whether the contribution is technical in nature.

12 Where a claim involves the use of a computer program, it does not naturally follow that the claim must be excluded. Instead, the contribution of a claim to a computer program must be assessed by reference to the process the program will cause a computer to perform, because, as stated in *Astron Clinica*[3], the assessment is based on the substance of the invention. In the case of *Halliburton Energy Services' Applications*[4], HHJ Birss QC, as he then was, emphasised that "[a] computer programmed to perform a task which makes a contribution to the art which is technical in nature is a patentable invention and may be claimed as such." Therefore, a computer program that provides a technical contribution will not fall under the exclusion because it is more than a computer program as such. The crux of the matter therefore lies in determining whether the claimed invention makes a technical contribution.

13 One way of identifying whether a computer-implemented invention makes a "technical contribution" is to use the signposts set out by Lewison J in *AT&T/CVON*[5] and subsequently used by Mann J in *Gemstar v Virgin*[6]. In considering the signposts in *AT&T/CVON*, it goes without saying that these do not provide a definitive account of what is and what isn't technical, but they do provide useful guidance of where the Courts have determined a technical contribution can be made. The signposts are as follows:

- whether the claimed technical effect has a technical effect on a process which is carried on outside the computer;
- whether the claimed technical effect operates at the level of the architecture of the computer; that is to say whether the effect is produced irrespective of the data being processed or the applications being run;
- whether the claimed technical effect results in the computer being made to operate in a new way;
- whether there is an increase in the speed or reliability of the computer; and
- whether the perceived problem is overcome by the claimed invention as opposed to merely being circumvented.

14 The examiner and the applicant agree that the first signpost is clearly not relevant in this case: the invention relates to a software-implemented way of processing queries and storing data in a distributed computing cluster and therefore does not have a technical effect on a process outside the computer.

[3] Astron Clinica Ltd & Ors v The Comptroller General of Patents, Designs and Trade Marks [2008] RPC 14
[4] Halliburton Energy Services Inc's Applications [2012] RPC
[5] AT&T Knowledge Ventures/CVON Innovations v Comptroller of Patents [2009] EWHC 343 (Pat)
[6] Gemstar-TV Guide International Inc & Ors v Virgin Media Ltd & Anor [2009] EWHC 3068 (Ch) [2010] RPC 10

15 The examiner and the applicant disagree in relation to the second signpost. The applicant argues that the second clause of this signpost sets out the intended meaning behind the first clause, namely whether a technical effect is produced irrespective of the data being processed or the applications being run. At the hearing, Mr Leach identified two technical effects resulting from the invention: the improved flexibility of the SOR model and some speed benefits of the SOW model. He said that the invention provides a faster and better computer system for performing queries on stored data in a distributed computing cluster, this being irrespective of the data content of the queries being processed. He also noted that a similar technical effect, i.e. a faster computer, was found to be patentable in *Symbian*.

16 The examiner says that he does not consider the effect of the contribution to be at the level of the architecture of the computer as it relates to the application level for processing data. Mr Leach disagrees, arguing that if this were the case then most, if not all, computer programs would not be patentable, whereas case law suggests otherwise.

17 The second signpost set out in *AT&T/CVON* derives from a number of cases considered by the courts here in the UK and by the EPO Boards of Appeal, e.g. *IBM (T 0006/83)* and *Gale*[7]*,* where it is clear that the "relevant" technical effect is found at a much higher level of generality within the computer, for example in the internal workings of processors and the transmission equipment irrespective of the nature of the data and the way in which a particular application program operates on the data files. In my view, the examiner's assessment is based on the correct reading of Lewison J's second signpost; the technical effect in the present case is found in the way in which the computer program operates on the data files and is not concerned with the internal workings of processors or how various elements of the computer system interact with each other. The second signpost is not relevant in this case.

18 The examiner and applicant further disagree in relation to the third and fourth signposts, i.e. whether the claimed technical effect results in the computer being made to operate in a new way, and whether there is an increase in the speed or reliability of the computer. The examiner says that he does not see that the computer (/hardware) is operating as a computer in a new way except insofar as any computer running a new program operates in a new way, nor does he see that the computer is operating more efficiently or effectively as a computer, instead it is only operating more effectively in processing queries. Mr Leach says that this does not correctly reflect the case law in this area. He argues that a computer that is operating more effectively in processing queries on stored data in a distributed computing cluster (by performing those queries more quickly while retaining the benefits of a SOR model) is "making the computer a better computer in the sense of running more efficiently and effectively as a computer" as in *HTC v Apple*[8], paragraph 51. Mr Leach repeats the point made earlier that a similar technical effect, i.e. a faster computer, was found to be patentable in *Symbian* (paragraphs 56 and 59).

19 I shall deal with the question of whether the invention results in a more efficient/effective computer first, i.e. the fourth signpost. It seems to me that the overall system claimed cannot be faster than known SOW systems for all queries (data) as there will be a need to convert data from an original format to a target format for some queries at the time those queries are executed. Furthermore, there

[7] Re Gales's Application [1991] RPC 305
[8] HTC v Apple [2013] EWCA Civ 451

are scenarios in which the overall system will be no faster than the known SOR systems: according to the independent claims, when data in a target format is not available the "query fragments are created for the original format", and the applicant has acknowledged that when query fragments in the original format are executed, they are executed on data in the original format. In such a scenario the system seems to operate as a SOR system and so is no faster than a purely SOR system.

20 The applicant argues that "an increase in processing speed is one of the technical effects that is repeatedly viewed as being adequate to put a computer program firmly outside the realm of excluded matter". On this I agree. The applicant uses this point to further argue that because the system of the independent claims processes queries quicker it falls outside the realm of excluded matter. On this point I find myself unable to agree. There is a distinction to be made between "processing speed of a computer" and the "speed at which a system comprising a computer processes queries". Take for example a SOW system. All data is processed (parsed) prior to any queries being executed on that parsed data. When a query is executed it appears to be executed quickly, but this is because the data has been pre-processed. Contrast this with the SOR system, the data still has to be processed (parsed) but this is done at the time the query is made and so to the user the system seems (and is) slower. However, the computer is no quicker at processing the data in the SOW system, it is just processing at a different time, i.e. the data is pre-processed, and so at the time of execution of a query the SOW system seems to compute quicker. The same is true of the system as defined in the independent claims in that some query fragments will be in a format which accords with data at a node and so will be able to be executed without there being a need to parse data at the time the query is to be executed. However that data has at some point in time been parsed and there is no evidence to suggest that the processing of the data (parsing) step is any quicker in the computer of the disclosed system of the invention than it is in any other system.

21 Even if I were to accept the applicant’s argument that the invention does result in the faster querying of data in a distributed computing cluster, it seems to me that the technical effect of increased speed is not as a result of the computer itself running any faster; what is “faster” is the way in which certain data queries are processed. I am therefore of the opinion that the requirement of the fourth signpost is not met. I also note that when he reviewed the fourth signpost at paragraph 51 of *HTC v Apple*, Kitchin LJ said that “this is, to my mind, another illustration of the still broader question whether the invention solves a technical problem within the computer.” In my view, the present invention does not solve a technical problem within the computer.

22 Returning to the third signpost of whether the claimed technical effect results in the computer being made to operate in a new way. In part the computer seems to perform as the known SOW model and in part according to the known SOR model. During operation the following operations are performed: i) a known schema is applied to data at each node and the data is saved in that format; ii) data is also saved at the node in the original format; iii) query fragments are created for known schemas for which data is known to exist; iv) query fragments may also be created in the original format; v) where the query fragments and data at a data node conform to the same schema the fragments are executed; vi) where the data at a node conforms to a different schema than the query fragments, the data are transformed using the relevant schema so that they conform to the format of the query fragments; vii) those

query fragments are then executed on the transformed data. Throughout all of these operations the computer seems to be operating conventionally and therefore is not considered to operate in a new way. I do not consider the requirement of the third signpost to be met.

23 The fifth signpost relates to whether a technical problem is solved or circumvented. In order for this signpost to have any bearing on assessing whether an invention falls outside the exclusion, the problem the invention solves has to be a technical one. I find myself in agreement with the examiner in that the problem is not a technical one. The invention is apparently not faster than a schema-on-write process and is no more flexible than a schema-on-read process, and seems to be a compromise based on personal choice rather than a process that provides a solution to a technical problem.

24 There are two final points to deal with. The first is whether the present invention provides the same technical effect, i.e. a faster computer, as was found patentable in *Symbian*, and I am satisfied that the two cases can be distinguished on their facts. The second concerns the corresponding EP patent application which the applicant says has not been subject to objection under the Article 52(2)(c) EPC at the European Patent Office despite the requirement of section 130(7) that the law regarding excluded subject matter should be interpreted to have the same effect in the UK as under the EPC. As the examiner quite rightly says in his pre-hearing report, the IPO is bound by UK law and precedent cases in reaching its decisions, which is what I have done in this decision.

Conclusion

25 I find that the claimed invention is excluded under section 1(2) because it relates to a computer program as such. I therefore refuse the application under section 18(3).

Appeal

26 Any appeal must be lodged within 28 days after the date of this decision.

H JONES
Deputy Director, acting for the Comptroller