CAFS-ISP is a search engine, or Information Search Processor, exploiting highly parallel hardware techniques to enable information stored on rotating discs to be searched and evaluated at full disc transfer speed. It is intimately associated with conventional disc controllers on ICL's mainframe systems, being an option on 2900 series and fitted as standard equipment on Series 39.

The main elements of the hardware are a Logical Format Unit, which locates significant points in the data stream as it arrives from disc; 16 Key Channels which compare reference values, with optional masks, against actual values in the data; a Search Evaluation Unit which implements boolean logic, or quorum logic, or both types in combination, to determine whether a record qualifies as a 'hit'; a Retrieval Unit which extracts either all fields or a specified subset of fields from hit records; and a Retrieval Processor, capable of executing minor processing functions on hit records, and responsible for controlling their orderly transmission to the host processor.

A File Correlation Unit, which is a published extension to the design and which currently exists in prototype form, facilitates join operations by enabling a representation of the result of any CAFS scan to be created on the fly in a bit-map. The contents of such a bit-map can be associated, by suitable "and, or, not" logic, with any subsequent CAFS scan.

How CAFS works

A process running in the mainframe generates a 'task specification', which instructs the CAFS hardware:

- which part of the disc store is to be searched;
- what conditions identify a 'hit record';
- and which fields are to be retrieved.

The completed task specification is then exported to the CAFS hardware in a single outward transfer, and the elements of the
The processes of aligning the read heads with the target cylinder are carried out in the conventional manner. Indeed, it is worth stressing that the disc mechanism itself is completely standard. Data is then streamed off the disc at the full disc transfer speed and onto the byte-wide CAFS highway. Internally the current CAFS engine is capable of searching at up to 3.6 megabytes per second; the effective speed is chiefly determined by the transfer rate of the discs being used.

The Logical Format Unit interprets the data stream as it passes, locating such significant points as the start of a database page, the start of a record, the start of a field to be searched, and the start of a field to be retrieved. When such points are identified, control signals are sent to activate appropriate units in the search and retrieval mechanisms.

The Key Channels compare the reference values provided in the task with the actual values found in the record under consideration. During this process masks may be applied to the actual values, so that insignificant bits are excluded from the comparison. All Key Channels operate independently, so that, depending on the form of the task, all may be looking at different fields or several may be searching for different values in a common field. If a field is repeated, a Key Channel may make and register several comparisons within a single record. Each Key Channel is capable of comparing up to 256 bytes of data. As the data is not buffered, and the comparisons of different fields may take place at different times, a means is needed of registering the result of a comparison until evaluation can take place. For this purpose each Key Channel has an associated comparator latch, in the form of a 3-bit result store.

At the end of each record, the Search Evaluation Unit determines whether or not the record was a 'hit', by an appropriate combination of the results from the Key Channels. This Unit consists of a number of co-operating processors operating in a highly parallel manner, each capable in one step of accepting data from its associated comparator latch, from its own store, and from its neighbour, and capable also both of placing a result in its own store and of passing a result to a neighbour. The most complex evaluation which the unit can be called on to execute can be completed in some 4.5 microseconds, or the time taken for a mere 16 bytes to pass along the highway.

The Retrieval Unit performs the more simple process of extracting required fields from hit records. The machine is able to retrieve either complete records or selected fields. In order to save time, data is received into the store of the Retrieval Unit while the record is still being scanned and evaluated. If the record is not a hit, the data can be