Optimizing Lock and Cache Structures in the Coupling Facility

As a user, you must allocate and define sizes that are appropriate to your application needs for the lock structure and a cache structure in the coupling facility (CF).

This section provides guidelines for determining optimal sizes for these structures based on current experience.

Note:
There may be sites for which these guidelines are not appropriate.

This chapter covers the following topics:

- Cache Structure Size in the Coupling Facility

- Lock Structure Size in the Coupling Facility

Cache Structure Size in the Coupling Facility

The coupling facility cache structure must be large enough to retain

- "directory elements" for all blocks that reside in all the buffer pools; and

- enough "data elements" to keep changed blocks between buffer flushes (cast-outs).

Directory elements are used to keep track of the cluster members that have a particular block in their buffer pools so that the block can be invalidated should any member modify it.

If the number of directory elements is insufficient, the coupling facility reuses existing directory elements and invalidates the blocks associated with those directory elements, because they can no longer be tracked. These blocks must then be reread from the database and registered again the next time they are referenced and validated, even though they did not change.

It is generally better to reassign storage for data elements to keep more ASSO and DATA blocks in the coupling facility than to define too many directory elements in the cache structure. More data elements than necessary can be used to keep additional blocks to improve the local buffer efficiency.

The number of directory elements need not be greater than the sum of the sizes of all buffer pools divided by the smallest block size in use for ASSO and DATA.

When connecting to the cache structure during startup, the ADAX57 message reports the number of directory elements and data elements. The ADARUN parameters `DIRRATIO` and `ELEMENTRATIO` determine the ratio between the number of directory and data elements.

Lock Structure Size in the Coupling Facility

All nuclei in a database cluster share the lock structure.

The coupling facility uses a lock table (organized as a hash table) to allocate and find a specific lock entry. It uses lock record entries to maintain data associated with lock instances.

When the coupling facility receives a lock request (for example, to put an ISN of a file into hold status), it allocates specific lock table and lock record entries unless another member of the cluster has already made a conflicting allocation.

- another member holds the same lock (real contention); or

- another lock name hashes to the same lock table entry (false contention).

False contention is eventually detected and resolved by the lock manager. However, since contention resolution is much more expensive than a lock request (there is a difference of about two orders of magnitude), false contention should be avoided.

False contention depends on the number of lock table entries compared to the number of concurrent lock requests. The likelihood (and therefore the frequency) of false contention decreases if the number of lock table entries allocated in the lock structure is increased.

Locks are held for a variety of entities, for example unique descriptor values. These lock types tend to occur with very different frequencies. The amount of lock activity during a session for each lock type is displayed in a shutdown statistic.

It is often the case that ISN locks show the greatest activity. The sum of high-water marks for NH yields an upper limit for the number of ISN locks that were held concurrently during the session.

Since lock contention is significantly more expensive than lock requests without contention, the lock table should be made large enough so that only a very small percentage of all lock requests cause false contention. As a rule of thumb, the number of lock table entries should be at least 1,000 times higher than the maximum number of ISN locks held concurrently.

RMF-I and RMF-III have reports that indicate how many instances of false contention occurred within a monitoring interval.

The minimum lock structure size can be roughly estimated as:

```
(NU*2 + NH + LDEUQP/16 + MAXFILES*4 + 50) * 240 + 500,000 bytes
```

where `MAXFILES` is the maximum number of files in the database (set in ADADEF or ADAORD) and `NU`, `NH`, and `LDEUQP` are the ADARUN parameters of the cluster nuclei. The formula in parentheses (NU*2 + NH + LDEUQP/16 + MAXFILES*4 + 50) is used to calculate the minimum number of lock record entries that the cluster nuclei expect to have available.