[OpenVMS] Discussion Of Processes In The RWxxx Scheduling States
Last Technical Review: 10-FEB-2000

PRODUCT: COMPAQ OpenVMS Alpha, All Versions
             COMPAQ OpenVMS VAX, All Versions

COMPONENT: Scheduler

SOURCE: Compaq Computer Corporation

OVERVIEW:

This article describes the various RWxxx resource wait states used in
OpenVMS, as well as some basic information on what to look for, when
trying to determine a cause for the state.

The first portion of the article contains a list summarizing each
resource wait. Followed by a more detailed description of the state.

BACKGROUND:

A process enters a RWxxx state when a requested resource is currently
unavailable, e.g., nonpaged pool. Seeing a process in an RWxxx state
is not typically indicative of a system problem because the process
will become computable when the resource becomes available. However,
processes hung in a particular state, or repeated occurrences of a
given state, may be cause for concern. See the RELATED ARTICLES
section for a list of articles addressing troubleshooting methods
for some of the RWxxx states.

DESCRIPTION:

    Resource Wait State   Reason for Wait
    -------------------   ---------------
       *RWAST   AST wait
       *RWMBX   Mailbox full
       *RWNPG   Nonpaged dynamic memory
       *RWPAG   Paged dynamic memory
        RWPGF   Page file full
       *RWMPE   Modified page list empty
       *RWMPB   Modified page writer busy
        RWSCS   System Communication Services wait
        RWCLU   Cluster state transition
        RWCAP   CPU capability
       *RWCSV   OpenVMS cluster server process
        RWINS   Waiting for an innermode semaphore

       * = A detailed troubleshooting article exists for processes
           in one of these states.

RWAST - Is used when a process is waiting for something to happen.
        When that event has occurred, an AST will be delivered
        to the process. The most common occurrences happen while
        waiting for a quota to be returned, typically used up by I/O
        requests, or for a channel to become no longer busy.

RWMBX - Is used when a process attempts to write to a mailbox which
        is too full to accept the message being written. Normally
        this may be disabled with the SET PROCESS/NORESOURCE_WAIT
        command. If this is done, the process attempting to write
        to a full mailbox will get a return status of SS$_MBFULL.
        A process may also be put into RWMBX state when attempting
        to write to the OPCOM mailbox. This is also used by the
        security auditing code. Some of these waits can not be
        disabled with the SET PROCESS/NORESOURCE_WAIT command.

RWNPG - Is used when wanting to force a process to wait for nonpaged
        dynamic memory. This is controlled by the SYSGEN parameters
        IRPCOUNT, SRPCOUNT, LRPCOUNT, and NPAGEDYN. Using the SHOW
        MEMORY/POOL/FULL command can help in determining which one.

RWPAG - Is occasionally used by the XQP (extended QIO Processor) to
        cause a process to wait when paged dynamic memory is not
        available. Using SDA it is possible to determine how much
        memory is being requested with the command EXAMINE @AP+4.
        It is suggested that the system have about 40% of paged
        dynamic memory free. This may be seen with the DCL command
        SHOW MEMORY/POOL/FULL. This resource is controlled by the
        SYSGEN parameter PAGEDYN.

        The PAGEDYN resource will typically be consumed by logical
        name tables.

RWMPE - Is used when a process wants to verify that the entire modified
        page list is written before continuing execution. Currently,
        this is only used by the OPCCRASH program when crashing the
        system. Doing this also causes all other processes attempting
        to add pages to the modified page list to go into RWMPB state,
        until the modified page list is flushed.

RWMPB - Is used any time a process attempts to page fault a page onto
        the modified page list; and the modified page writer is writing
        the modified page list to their backing store address.
        The modified page writer is considered busy from the time the
        modified page list gets MPW_WAITLIMIT pages, until the modified
        page writer can bring it down to MPW_LOWAITLIMIT pages.

RWSCS - Is used by the distributed lock manager in a cluster environment
        when the requested lock information, or lock request, requires
        using the System Communication Services to communicate with the
        lock manager on another node of the cluster. Once the necessary
        communication is done, the process will continue executing.

RWCLU - Is used by the lock manager when it needs to stall a lock
        request. This is normally done when the cluster is in a
        cluster state transition.

        Note:
          With the new dynamic lock re-mastering in OpenVMS V5.5,
          processes might be suspended in RWCLU state for the
          duration of the lock tree transfer to another node.

          OpenVMS V5.5-2 introduced a special SYSGEN parameter
          PE1 to prohibit the transfer of larger trees to avoid
          unnecessary long RWCLU waits.

RWCAP - May be used by the VMS scheduling code when quorum is lost.
        However, this is not the normal case since once quorum is lost,
        no processes are allowed to continue executing.

            Note:
              In the case of a CLUEXIT bugcheck, it is very typical to see
              many processes in the RWCAP wait state.

RWCSV - Is the state a process is put in when waiting for communication
        with the CLUSTER_SERVER process. This should only be seen in a
        cluster environment. This process is normally in an HIB state.
        If it is in any other state, that is likely the cause of the
        RWCSV state, and should be looked into. If that process does
        not exist, restart it with the command:

                @SYS$SYSTEM:STARTUP CSP

        In OpenVMS, V5.2, processes enter RWCSV while waiting for the
        synchronous cluster-wide system services to return completion
        status from remote nodes.

RWINS - (Alpha only) Indicates that a process thread is trying to acquire
        the inner-mode semaphore. This mutex serializes thread execution
        in Executive, and/or Kernel mode while another thread of the same
        process is currently in Executive or Kernel mode.

        Multiple threads from a single process may run concurrently on
        different processors in user and supervisor mode. For threads to
        be allowed concurrent execution in an "inner mode", they must be
        considered "thread safe" by OpenVMS Alpha. If they're not
        considered "thread safe" they must acquire the "per process"
        inner-mode semaphore to serialize execution.