

## \$README495

Fix: ZP15F495  
Status: Released  
Component: STACK  
Date: 2013/11/01  
Import: Medium  
Risk: High

### Symptoms:

IPN489I Duplicate RECV Discarded

IPN634E Duplicate ECB

Socket applications fail when running on multiple CPUs with a heavy load of external socket applications.

### Description:

ZP15F495 corrects a number of interrelated problems detected during heavy stress testing at CSI, IBM, and customer sites. These integrated changes have separate zap numbers for each affected module but for installation and discussion purposes are packaged and documented in the single ZP15F495.ZIP file. These changes have been extensively tested as a single group of updates and must be installed together.

THE LIB.SUBLIB USED TO INSTALL THIS ZAP MUST BE IN THE PHASE LIBDEF SEARCH CHAIN OF ALL EXTERNAL SOCKET APPLICATION PARTITIONS!!! OTHERWISE EXTERNAL SOCKET APPLICATIONS WILL FAIL WITH A "OS05I PHASE \$B SOCKET NOT FOUND" ERROR MESSAGE AND THE PARTITION WILL TERMINATE.

Many VSE sites have an outdated version of TCP/IP in the PRD1.BASE LIB.SUBLIB and will run into this problem unless the permanent phase LIBDEF search chain is updated to have the TCP/IP installation LIB.SUBLIB in front of PRD1.BASE in the permanent phase LIBDEF search chain for all partitions.

The following modules are contained in this zap:

ZP15F032 ASOCKET.PHASE This zap changes the method by which TCP socket requests are queued and processed from external partitions. All external partition socket requests are handled by the SVA resident ASOCKET program which runs as part of the external partition. ASOCKET by default anchors external socket requests on a single queue in the SOCKET $nn$  SDL entry where  $nn$  is the corresponding



# CSI INTERNATIONAL

TCP/IP stack id. ASOCKET will now CDLOAD the new \$B SOCKET.PHASE into private partition GETVIS, and by default external socket request will be allocated in 31-bit system GETVIS and anchored in the SOCKET $nn$  SDL entry. The new EXTPGVS ON command can be issued in the TCP/IP partition to cause external socket requests to be independently queued into a new partition control block anchored in the COMREG-IJBTCPP2 field. The TCP/IP partition CSOCKET program will then use access register mode to pull the TCP socket requests allocated in private 31-bit partition GETVIS from the external partition into the TCP/IP partition. Non-TCP requests are not affected by the EXTPGVS ON setting.

ZP15F032 was superseded with ZP15F092 for the ASOCKET.PHASE.  
ZP15F092 must be applied after ZP15F032.

ZP15F033 CSOCKET.PHASE  
ZP15F033 was superseded with ZP15F096 for the CSOCKET.PHASE.  
ZP15F096 must be applied after ZP15F033.

ZP15F041 SOCKPASS.PHASE

ZP15F043 IPNTYTCP.PHASE

CSOCKET, SOCKPASS, and IPNTYTCP are updated to support the new partition GETVIS socket request queues in conjunction with the changes described above for ASOCKET.

ZP15F044 \$B SOCKET.PHASE This is a new phase that is CDLOADED by ASOCKET during the first socket request from an external partition. \$B SOCKET allocates and maintains the new external partition control block (PXBLOK) and anchors it in the COMREG-IJBTCPP2 field. The PXBLOK contains the job start time, job step start time, flags, anchors for socket requests, and statistical counters for a single partition. It also provides services that run under the control of the external partition for writing messages and for optional locking of socket processing.

ZP15F030 SOCKOPT.A This zap provides the SOCKOPT.A macro that is used for generating a custom \$SOCKOPT.PHASE. The EXTPGVS ON command globally sets the use of private partition GETVIS for all external partition TCP socket requests. If EXTPGVS OFF is active (the default setting) then a custom \$SOCKOPT.PHASE can be generated with SOCFLG1=\$OPTPGSB to allow a partition to use private partition GETVIS for TCP socket requests.

The new SOCKOPT MAXSOCO= setting can be used to control the maximum number of outstanding TCP socket requests that a single external partition is allowed to have before socket requests are rejected. The default is MAXSOCO=00 which allows an unlimited number of outstanding socket requests from a single partition. This can be a problem if an application accidentally goes into a loop issuing socket requests without waiting for the ECB to be posted and overruns the TCP/IP partition



CSI INTERNATIONAL

with unnecessary socket requests. While it is normal to have a number of outstanding listens, receives, or even send requests, it is a good idea to limit the maximum number of outstanding socket requests that a single partition is allowed to have.

If/when the maximum number of outstanding (un-posted) sockets is reached a:

IPN960C SOCKET REQUEST REJECTED RS=SOCOEXCD message will be issued and the socket request will be rejected.

The new QUERY EXTERNAL command can be issued to display statistics with the number of outstanding socket requests for every active job step using socket requests. It is recommended that MAXSOCO=999 be specified to protect from an errant socket application. To use this option, EXTPGVS ON must be active.

The custom generated \$SOCKOPT.PHASE phase must be in the partition phase LIBDEF search chain for this option to be used.

ZP15F045 IPNIGARB.PHASE Updated to provide support for DIAG GETVIS command. The DIAG CLEANUP previously provided this support and can still be used but the DIAG GETVIS makes it clearer for monitoring GETVIS usage. The detection of dead partitions with outstanding socket requests now also uses the job step start time instead of the job start time which should improve clean up on multiple step jobs.

ZP15F046 IPNETAS.PHASE Updated for allocating a larger system control block (SCBLOK) that also anchors the external partition control blocks (PXBLOK). It will also cause a message similar to:

IPN694I SCBLOK at: 00366000 SCALLOCN to occur during the startup of TCP/IP.

ZP15F062 CMDPARS.PHASE

ZP15F063 CMDEXEC.PHASE CMDPARS and CMDEXEC were updated to add new commands.

EXTPGVS OFF is the default setting and will use the same old method of allocating external partition socket requests in 31-bit system GETVIS storage.

EXTPGVS ON will cause the allocation of external partition socket requests in 31-bit private partition storage and be anchored as documented above in the ASOCKET description.

QUERY EXTERNAL will display statistics from external socket application partitions. The statistics are reset at the start of a new job step. The following messages and statistics are displayed:



CSI INTERNATIONAL

IPN685I F9 last socket job step 11:27:36.545

IPN685I identifies the external job step start time.

IPN686I F9 csoc=3 + ntcp=6 = 9

IPN686I csoc= is the total number of TCP external socket requests copied from the external partition to the TCP/IP partition by the CSOCKET program.

IPN686I ntcp= is the total number of non-TCP requests from the external partition processed by the ASOCKET program.

IPN687I spas=6 is the total number of non-asynchronous external socket requests posted for this partition by the SOCKPASS program.

IPN687I async=0 is the total number of asynchronous external socket requests posted for this partition by the SOCKPASS program.

IPN687I ntlk=2 is the total number of notalk=yes and fast=yes socket requests received by the ASOCKET program. The ECB associated with these types of socket requests is immediately posted by ASOCKET before the request has even been processed by the TCP/IP partition.

IPN688I F9 Socket requests(9) do not match posts(8) asoc(9)

IPN688I is only issued when the number of socket requests does not match the number of posts. Although this can indicate a problem it can also be normal for an application to have some passive open listens or receives outstanding, but if the number is excessively high or is continually growing larger, then the socket application should be analyzed for possible processing and logic problems.

IPN689I F9 Total passive opens(1) successful connects(0)

IPN689I is only issued when passive (listen state) opens are detected, and contains the number of passive opens issued and the number of connections that have been completed by the IPNTYTCP program. Passive opens are normally used by server socket applications and these numbers indicate the number of connections that have occurred.

ZP15F064 IPCCSERV.PHASE Some previously DIAG HASHING messages have now been changed to be issued without DIAG HASHING active. This is to help clearly identify problems that may be occurring during problem source identification.



CSI INTERNATIONAL

ZP15F067 MSKELIP.PHASE was updated for some changes to messages.

IPI508 xxxxxxxx job step time changed px=addr cc=addr

This message informs that the job step time has changed and previous socket requests from the prior step are being discarded.

IPN684, IPN685, IPN686, IPN687, IPN688, and IPN689 are new messages produced by the QUERY EXTERNAL command and provide statistical counters for socket requests for each external partition.

ZP15F070 FTPDAEMN.PHASE was updated to handle a condition where a new connection was established during the brief window of time (1/30th of a second) that the ftp server was establishing a new connection. A status command was used to verify that the server was back in the listen state but a new client had already connected and the status was established. This condition would cause the FTPDAEMN to shut down and no further connections could be established with it. A recommended circumvention to avoid client connection requests from being rejected is by issuing a secondary DEFINE FTPD on the same port (default 21) to always maintain a listen state for new connections.

ZP15F079 IPMSSG.PHASE When running a large number of FTPBATCH jobs a VSE system could run out of available tasks to attach. When this condition occurs the FTPBATCH job will hang (not terminate) and cannot issue any error messages because the attach of the message writer subtask has failed. This zap changes the attach of the message writer to wait for a free subtask and issues a:

IPN866W NO SUBTASKS AVAILABLE IPMSSG IS WAITING FOR A SUBTASK

message. The VSE workload should either be reduced or increase the number of available subtasks in the system.

ZP15F080 FTPBATCH.PHASE used a control call to verify the version of CDLOADed phases and this control call could fail when running a large number of FTPBATCH jobs simultaneously. The FTPBATCH job would then fail with a CDLOAD error. This zap corrects this false CDLOAD failure. It also now checks the return code for the attach of the message writer subtask in conjunction with ZP15F079 IPMSSG as mentioned above to avoid a hung FTPBATCH job.

ZP15F081 IPNACONT.PHASE provides a "free port" service to select the next available free port available for use. The free port range is established by the PORTRANGE LO=4096 HI=65535 but can be set to any desired range. It is possible for a site to use a port in this free port range for a server application and this could cause a problem for FTPBATCH and other applications when a port was returned that was being used by the server application. In the failing system a telnet daemon was defined on port 5100 and the FTPBATCH connected to the telnet daemon and of course this did not



CSI INTERNATIONAL



have good results (the FTPBATCH job would hang and would have to be manually cancelled). The free port service code will now scan to verify that the free local port is not currently in use by any other connections in the TCP/IP stack. In addition it is recommended that server applications (TELNETD, FTPD, etc.) should not be assigned ports in the established free port range.

ZP15F495 IPNET.PHASE Updated to post socket requests during shutdown.

ZP15F034 IPNRBSDC.PHASE This phase replacement zap corrects the following issues in the BSD socket interface.

The abort() function could hang due to the changes to ASOCKET and is corrected in the zap.

The socket check routine could incorrectly detect sockets that had been processed by givesocket() to be a dead socket. This could then cause a takesocket() failure within an application such as the CICS EZAL to terminate. In addition a diagnostic message will now be issued to indicate the number of sockets in the 'given' state during the socket check processing to assist in application debugging. The message will be similar to this:

BSD116I 37 sockets flagged given

A selectex() function could ABEND with a segment translation exception due to an un-cleared high-order bit in the last parameter passed.

CICS web services could ABEND during an asynchronous accept() due to an ECONNRESET being returned during the accept processing.



CSI INTERNATIONAL