SR 3-5661194891 : Solaris 11 sshd version Sun_SSH_2.0 issue

Severity	3-Standard	Status	Development Working
Escalation Status	Never Escalated	Opened	May 2, 2012 9:40 PM (1+ month ago)
Last Updated	Jun 25, 2012 9:13 AM (Yesterday)		
Bug Reference	No Related Bugs	Attachments	View Attachments (22)
Related Articles	View Related Articles	Related SRs	No Related SRs
Field Service Tasks	No Field Service Tasks		
Support Identifier	17071810		
Serial Number	0710NNN1DC	Account Name	Medtronic Inc
Primary Contact	Chang-An Hsiao	Alternate Contact	
Service Address	4102 SOUTHPOINT BLVD JACKSONVILLE FL United States		
System/Configuration		Host	
Product	Solaris SPARC Operating System	Asset Name	
		Product Version	10 10/08 U6
Operating System	SPARC	OS Version	
Project		Project Milestone	
Problem Description	sshd version Sun_SSH_2.0 sends out packets being blocked by network device Strange thing is, not all SSH clients fail to connect to server. For example, OpenSSH_3.6.1p2 can connect fine while OpenSSH_4.3p2 cannot connect at all. When Solaris 10 was installed, all SSH clients could connect to the server without issues.		

History

ODM Action Plan — **Oracle Support**- Jun 25, 2012 9:13 AM (Yesterday)

Who: customer Chang-An Hsiao
What: continue to work with the firewall vendor to get the final fix from there. Use the provided IDR and the tunable as workaround until the firewall fix is available
When: by August, 3rd

Who: Solaris sustaining
What: continue to produce Solaris 11 (instead of IDR) to allow the tunable to be used
When: by August, 3rd

Update from Customer — **CHANG_AN@YAHOO.COM**- Jun 24, 2012 3:21 PM (Sunday)

Thanks for the clarification.

I will try to uninstall IDR instead of rolling back to older BE.

Regards,

Chang-An

ODM Action Plan — **Oracle Support**- Jun 24, 2012 9:54 AM (Sunday)

Who: customer Chang-An Hsiao
What: continue to work with the firewall vendor to get the final fix from there. Use the provided IDR and the tunable as workaround until the firewall fix is available
When: by August, 3rd

Who: Solaris sustaining
What: continue to produce Solaris 11 (instead of IDR) to allow the tunable to be used
When: by August, 3rd

Notes — **Oracle Support**- Jun 24, 2012 9:53 AM (Sunday)

Hi,

sorry to hear that we got no response from the gateway vendor yet.

Regarding the question about the boot environments (BE): The boot environment is a snapshot of the ZFS root pool (rpool) and rolling back to an older snapshot (older BE) will cause any cahnges on the root pool to be lost.
It is recommended that applications are not using the root pool but are using their own ZFS dataset - in which case they will be independent of any BE.
If your MySQL database is uusing information from the root pool (either the DB itself or the DB configuration) then rolling back to an older BE may indeed be prproblematic and in this case I would NOT recommend to go back to the old BE but rather deinstall the IDR.
See the output of "pkg info idr232" which will show the pkg command to be used to uninstall this IDR without the requirement to go back to an older BE. A reboot is still required though as this IDR replaces the kernel. The new reboot after IDR removal will still be from the current BE though.

Best regards,
Wolfgang Ley

Update from Customer — **CHANG_AN@YAHOO.COM**- Jun 23, 2012 6:07 PM (Saturday)

Thanks a lot for your help with the tunable workaround. I am still pushing the gateway router vendor to come up with an update but haven't got anything from them yet. I ping them every week and the response is he same --- they are looking into the case. I am very frustrated with them (auctioned). Not much more I can push.....

When I finally get their update, I can roll back to the original BE.

One question. I made some changes to MySQL before I roll back to original BE yesterday. After rolling back, the MySQL lost the changes. You would think it has been written to the database and BE should affect it but it's gone. Would I lose all the changes in the future when I roll back to the original BE?

ODM Action Plan — **Oracle Support**- Jun 23, 2012 11:50 AM (Saturday)

Who: customer Chang-An Hsiao
What: continue to work with the firewall vendor to get the final fix from there. Use the provided IDR and the tunable as workaround until the firewall fix is available
When: by August, 3rd

Who: Solaris sustaining
What: continue to produce Solaris 11 (instead of IDR) to allow the tunable to be used
When: by August, 3rd

Notes **Oracle Support**- Jun 23, 2012 11:48 AM (Saturday)
Thanks for the feedback and the information that the IDR installation works now (and the tunable works, too).
Please note that this tunable is just an insecure workaround and not a final solution. The final solution is to fix the firewall system and to avoid initial windows sizes which are smaller than then MSS (maximum segment size of the local ethernet packet).
Do you have any news from the firewwall vendor regarding the fix?

We will continue to include the fix (to allow the tunable) in an upciming Solaris 11 patch (instead of just the IDR) but this will take some time - and as said: this is only a workaround and not the solution (which needs to be provided by the firewall vendor as we cannot fix that).

Best regards,
Wolfgang Ley

Update from Customer **CHANG_AN@YAHOO.COM**- Jun 22, 2012 11:20 PM (Friday)
I rolled back to original BE and then destroyed the new BE and rebooted first to get back to original state.

Then, I renamed /default to /default_explorer and install the IDR package with truss capture again just in case.

This time, it didn't report any errors.

After reboot, the system tunable (set ip:tcp_init_wnd_chk = 512 in /etc/system) seems to help and SSH starts to work! Windows size attack protection seems to be disabled as anticipated!!!

Update from Customer **CHANG_AN@YAHOO.COM**- Jun 22, 2012 9:59 PM (Friday)
All right.

I will try renaming it and install the IDR again.

Have a nice weekend!

ODM Action Plan **Oracle Support**- Jun 22, 2012 4:32 PM (Friday)
Who: customer Chang-An Hsiao
What: rename the directory "/default" and tretry the pkg install command
When: by June, 29th

Notes **Oracle Support**- Jun 22, 2012 4:31 PM (Friday)
Hi,

no - please do not rename all of them as most of them are correct system defaults here.
Please only rename the /default directory and retry the pkg install afterwards. Thanks.

Best regards,
Wolfgang Ley.

Update from Customer **CHANG_AN@YAHOO.COM**- Jun 22, 2012 4:29 PM (Friday)
I did a thorough search from / and the following are all the default directories I got:

%find / -name default
/export/home/cph/.mozilla/default
/export/home/cph/.cpan/build/Template-Toolkit-2.20-0xicG0/t/test/lib/default
/export/home/cah/.gconf/apps/panel/profiles/default
/export/home/cah/.gnome2/panel2.d/default
/usr/lib/ocm/ccr/sysman/plugins/default
/usr/lib/ocm/ccr/config/default
/usr/lib/help/locale/C/default
/usr/lib/thunderbird/chrome/icons/default
/usr/lib/thunderbird/extensions/{e2fda1a4-762b-4020-b5ad-a41df1933103}/chrome/icons/default
/usr/lib/firefox/chrome/icons/default
/usr/share/pixmaps/pidgin/emotes/default
/usr/share/X11/xkb/compat/default
/usr/share/X11/xkb/types/default
/usr/share/X11/xkb/semantics/default
/usr/share/lib/pkg/web/_themes/default
/usr/share/espeak-data/voices/default
/usr/share/sounds/gnome/default
/etc/opt/SUNWexplo/default
/etc/explorer/default
/etc/default
/var/sadm/install/admin/default
/var/ocm/ccr/config/default
/default
/boot/grub/default
/opt/sfw/mkspecs/default

/opt/sfw/lib/xemacs/xemacs-packages/etc/ecb/ecb-images/default
/opt/sfw/kde/share/apps/mediacontrol/default
/opt/sfw/kde/share/apps/kmessedwords/themes/default
/opt/sfw/kde/share/apps/kwin4/grafix/default
/opt/SUNWexplo/output/explorer.000403a9.cahtoh02-2012.05.03.16.51/etc/default
/opt/SUNWexplo/default

I don't need to rename each single one, do I?

Notes **Oracle Support**- Jun 22, 2012 4:22 PM (Friday)
I would recommend to simply rename it and move it back after a successful pkg install operation.
Please also check for a "default" directory in your homedir and/or the directory where you placed the p5p file (the IDR0 for installation.

Best regards,
Wolfgang Ley.

Update from Customer **CHANG_AN@YAHOO.COM**- Jun 22, 2012 4:14 PM (Friday)
/default directory was created when I was collecting explorer.

I guess I can safely remove it for it only has a small explorer text file in it.

Created by SUNWsneep so that explorer can obtain the Chassis Serial Number.
Note that in order for explorer to execute successfully, several settings
are necessary, which SUNWsneep does not provide.
Be sure to run "explorer -g" to configure this file properly.

# Serial number for hostid 000403a9
EXP_SERIAL_000403a9="3CR8201F63"

ODM Action Plan **Oracle Support**- Jun 22, 2012 3:54 PM (Friday)
Who: customer Chang-An Hsiao
What: rename the directory "default" which causes the pkg install failure due to mkisofs problems and retry the pkg install (see previous notes for details)
When: by June, 29th

Notes **Oracle Support**- Jun 22, 2012 3:54 PM (Friday)
Hi,

the pkg install error was due to a directory called "default" which should be renamed before the pkg install (see previous notes).
If your testing has caused multiple boot environments to be created then I would suggest to cleanup first.
Boot from the "Solaris 11 11/11" grub menu entry and then verify with "beadm list" that the current boot environment is "solaris" (the default).
Use "beadm activate solaris" to make this older/unpatched boot environment the default again. A check with "beadm list" should show NR for the "solaris" environment.
Then use "beadm destory solaris-1" (and maybe solars-2 et.c) to remove the environments from the failed idr installation.

Renamed the "default" directory and run the pkg install again. This should no longer produce the bootarchive error and the tunable via /etc/system should work (after a reboot from the new environment solaris-1)).

Best regards,
Wolfgang Ley

ODM Action Plan **Oracle Support**- Jun 22, 2012 1:44 PM (Friday)
Who: customer Chang-An Hsiao
What: rename the directory "default" which causes the pkg install failure due to mkisofs problems and retry the pkg install (see previous notes for details)
When: by June, 29th

Notes **Oracle Support**- Jun 22, 2012 1:43 PM (Friday)
Hi,

the truss showed that the underlying problem is a failure of the "mkisofs" command which stumbles over a directory called "default" on your system.
Can you please check whether your have durectory called "default"? Unfortunately the full pathname is not available in the truss but my guess is that this is either /export/home/cah/default, /default or /root/default (or a directory called "default" if you have placed the idr232.2.p5p file elsewhere).

If you do have a directory /export/home/cah/default then please rename this directory (e.g. to "default.old"). Boot from the original (unpatched) boot environment and rnu the pkg install command again.
Please let me know if the pkg install works after renaming the "default" directory on your system. Thanks.

Bets regards,
Wolfgang Ley

Notes **Oracle Support**- Jun 22, 2012 1:39 PM (Friday)
I managed to reproduce the customers failure after creating a directory "default" (which then causes mkisofs to fail)

Notes **Oracle Support**- Jun 22, 2012 1:29 PM (Friday)
The problem is related to a directory called "default" which does exists on the customers system and causes the mkisofs to abort.
The open() is called without full pathname so it is unclear whether this is /defauklt or /export/home/cah/default

From the failing truss of the custoemr:
[....]
1722/1: 174.6601 openat(0xFFD19553, 0x080A4498, 0, 0666) = 5
1722/1: 0xFFD19553: AT_FDCWD
1722/1: 0x080A4498: "default"

[...]

From the reference truss of the lab system:
[....]
2052/1: 220.8747 openat(0xFFD19553, 0x08044E70, 0, 0666) Err#2 ENOENT
2052/1: 0xFFD19553: AT_FDCWD
2052/1: 0x08044E70: "/usr/lib/siconv/default"
2052/1: 220.8750 openat(0xFFD19553, 0x080A4490, 0, 0666) Err#2 ENOENT
2052/1: 0xFFD19553: AT_FDCWD
2052/1: 0x080A4490: "cp10000"
2052/1: 220.8750 openat(0xFFD19553, 0x08044E70, 0, 0666) Err#2 ENOENT
2052/1: 0xFFD19553: AT_FDCWD
2052/1: 0x08044E70: "/usr/lib/siconv/cp10000"
2052/1: 220.9005 faccessat(0xFFD19553, 0x08100E78, 4, 0) Err#2 ENOENT
2052/1: 0xFFD19553: AT_FDCWD
2052/1: 0x08100E78: "/usr/lib/iconv/geniconvtbl/binarytables/cp10000%UCS-2BE.bt"
2052/1: 220.9166 faccessat(0xFFD19553, 0x08100E78, 4, 0) Err#2 ENOENT
2052/1: 0xFFD19553: AT_FDCWD
2052/1: 0x08100E78: "/usr/lib/iconv/cp10000%UCS-2BE.so"
2052/1: 220.9169 openat(0xFFD19553, 0xFEF579E8, 0, 01001052760) = 5
2052/1: 0xFFD19553: AT_FDCWD
2052/1: 0xFEF579E8: "/usr/lib/iconv/alias"
2052/1: 220.9172 fstatat64(5, 0, 0x08045570, 0) = 0
[...]

So we need to know why the open for "default" worked on the customers system. Note that /usr/lib/siconv/default is not available in any package and that the open was done without this full pathname (on the customers system).

Notes **Oracle Support**- Jun 22, 2012 12:02 PM (Friday)

The creation of the boot archive fails due to a failing mkisofs command.
From the truss we can see:

[...]
1722/1: 174.6442 execve(0xFEB0A1A4, 0xFEB0A0CC, 0xFEB0A0F4, 0) argc = 8
1722/1: 0xFEB0A1A4: "/usr/bin/mkisofs"
1722/1: argv: /usr/bin/mkisofs -quiet -graft-points -dlrDJN
1722/: -relaxed-filenames -o
1722/: /tmp/tmpC6bMr1//platform/i86pc/amd64/archive-new-1706
1722/: /tmp/tmpC6bMr1//platform/i86pc/amd64/archive_cache
1722/: envp: _=*1706*/usr/bin/mkisofs AVAHI_COMPAT_NOWARN=1
1722/: EDITOR=/bin/vi HISTFILE=/root/.sh_history HISTSIZE=1024
1722/: HOME=/ IFS=
LANG=C LC_ALL= LC_COLLATE= LC_CTYPE=
1722/: LC_MESSAGES= LC_MONETARY= LC_NUMERIC= LC_TIME= LOGNAME=root
1722/: MAIL=/var/mail/root
1722/: PATH=/usr/bin:/bin:/etc:/usr/bin:./:/usr/openwin/bin:/opt/bin:/usr/sfw/bin:/sbin:/usr/sbin:/export/home/www/bin:/usr/local/bin:/usr/local/ssh/bin:
/export/home/cah/bin/script:/usr/ccs/bin:/usr/local/ssl/bin:/opt/sfw/bin:/opt/sfw/sbin:/etc:/usr/bin:./:/usr/openwin/bin:/opt/bin:/usr/sfw/bin:/sbin:/usr/sbin:/export
/home/www/bin:/usr/local/bin:/usr/local/ssh/bin:/root/bin/script:/usr/ccs/bin:/usr/local/ssl/bin:/opt/sfw/bin:/opt/sfw/sbin
1722/: PS1=cahtoh02:${PWD}% PWD=/ SHELL=/bin/ksh SHLVL=2
1722/: SUDO_COMMAND=/usr/bin/ksh -o vi SUDO_GID=4 SUDO_UID=1001
1722/: SUDO_USER=cah TERM=vt100 TZ=localtime USER=root USERNAME=root
1722/: A__z="*SHLVL
[...]
1722/: 174.6595 write(2, 0x080452E8, 64) = 64
1722/: W a r n i n g : c r e a t i n g f i l e s y s t e m t h a
1722/: t d o e s n o t c o n f o r m t o I S O - 9 6 6 0 .\n
1706/: 174.6595 read(5, 0x080C0744, 5120) = 64
1706/: W a r n i n g : c r e a t i n g f i l e s y s t e m t h a
1706/: t d o e s n o t c o n f o r m t o I S O - 9 6 6 0 .\n
1722/: 174.6596 brk(0x080FF948) = 0
1722/: 174.6597 brk(0x08101948) = 0
1722/: 174.6601 openat(0xFFD19553, 0x080A4498, 0, 0666) = 5
1722/: 0xFFD19553: AT_FDCWD
1722/: 0x080A4498: "default"
1722/: 174.6601 fstatat64(5, 0, 0x08044ED0, 0) = 0
1722/: d=0x030D0002 i=405610 m=0040555 l=2 u=0 g=2 sz=3
1722/: at = May 21 18:35:31 EDT 2012 [1337639731.501296776]
1722/: mt = May 7 12:14:16 EDT 2012 [1336407256.309535286]
1722/: ct = May 7 12:14:16 EDT 2012 [1336407256.316169370]
1722/: bsz=512 blks=3 fs=zfs
1722/: 174.6602 fstatat64(5, 0, 0x08044DE0, 0) = 0
1722/: d=0x030D0002 i=405610 m=0040555 l=2 u=0 g=2 sz=3
1722/: at = May 21 18:35:31 EDT 2012 [1337639731.501296776]
1722/: mt = May 7 12:14:16 EDT 2012 [1336407256.309535286]
1722/: ct = May 7 12:14:16 EDT 2012 [1336407256.316169370]
1722/: bsz=512 blks=3 fs=zfs
1722/: 174.6603 ioctl(5, 0x00005401, 0x08044E80) Err#25 ENOTTY
1722/: 174.6603 read(5, 0x080FFF64, 512) Err#21 EISDIR

```
1722/1:     174.6604     schedctl()                          = 0xFEDD1000
1722/1:     174.6604     llseek(5, 0, 1)                      = 0
1722/1:     174.6604     close(5)                            = 0
1722/1:     174.6605     write(2, 0x08045308, 47)                 = 47
1722/1:       U n k n o w n   c h a r s e t   ' d e f a u l t ' . \n K n o w n
1722/1:         c h a r s e t s   a r e : \n
1706/1:     174.6605     read(5, 0x080C0744, 5120)                 = 47
1706/1:       U n k n o w n   c h a r s e t   ' d e f a u l t ' . \n K n o w n
1706/1:         c h a r s e t s   a r e : \n
1722/1:     174.6605     openat(0xFFD19553, 0x08044FF0, 0140020004, 0)   Err#2 ENOENT
1722/1:       0xFFD19553: AT_FDCWD
1722/1:       0x08044FF0: "/usr/lib/siconv/"
1722/1:     174.6607     write(2, 0x08045268, 68)                 = 68
1722/1:       / u s r / b i n / m k i s o f s :   I n s t a l l a t i o n   p
1722/1:       r o b l e m :   ' / u s r / l i b / s i c o n v / '   m i s s i
1722/1:       n g . \n
1706/1:     174.6607     read(5, 0x080C0744, 5120)                 = 68
1706/1:       / u s r / b i n / m k i s o f s :   I n s t a l l a t i o n   p
1706/1:       r o b l e m :   ' / u s r / l i b / s i c o n v / '   m i s s i
1706/1:       n g . \n
1722/1:     174.6608     _exit(-1)
1706/1:     174.6609     read(5, 0x080C0744, 5120)                 = 0
1706/1:     174.6610     llseek(5, 0, 1)                      Err#29 ESPIPE
1706/1:     174.6611     close(5)                            = 0
1706/1:     174.6612     waitid(0, 1722, 0x08045830, 03)            = 0
1706/1:         siginfo: SIGCLD CLD_EXITED pid=1722 status=0xFFFFFFFF
[...]
1706/1:     174.6616     schedctl()                          = 0xFEDC4000
1706/1:     174.6617     write(2, 0x08072D50, 40)                 Err#9 EBADF
1706/1:       b o o t - a r c h i v e   c r e a t i o n   F A I L E D ,   c o
1706/1:       m m a n d :   '
1706/1:     174.6617     write(2, 0x0808CEA0, 63)                 Err#9 EBADF
1706/1:       W a r n i n g :   c r e a t i n g   f i l e s y s t e m   t h a
1706/1:       t   d o e s   n o t   c o n f o r m   t o   I S O - 9 6 6 0 .
1706/1:     174.6618     write(2, 0x0808CFA8, 26)                 Err#9 EBADF
1706/1:       U n k n o w n   c h a r s e t   ' d e f a u l t ' .
1706/1:     174.6618     write(2, 0x0808A310, 19)                 Err#9 EBADF
1706/1:       K n o w n   c h a r s e t s   a r e :
1706/1:     174.6619     write(2, 0x0808D998, 67)                 Err#9 EBADF
1706/1:       / u s r / b i n / m k i s o f s :   I n s t a l l a t i o n   p
1706/1:       r o b l e m :   ' / u s r / l i b / s i c o n v / '   m i s s i
1706/1:       n g .
1706/1:     174.6620     unlinkat(0xFFD19553, 0x080469D0, 0)          Err#2 ENOENT
1706/1:       0xFFD19553: AT_FDCWD
1706/1:       0x080469D0: "/tmp/tmpC6bMr1//platform/i86pc/amd64/archive-new-1706"
1706/1:     174.6630     fcntl(4, 34, 0x08047AC0)                 = 0
1706/1:         typ=F_UNLCK  whence=SEEK_SET start=0     len=0     sys=134511368 pid=134570274
1706/1:     174.6630     close(4)                            = 0
1706/1:     174.6631     write(1, 0x080BE734, 59)                 Err#9 EBADF
1706/1:       u p d a t i n g   / t m p / t m p C 6 b M r 1 / / p l a t f o r
1706/1:       m / i 8 6 p c / a m d 6 4 / b o o t _ a r c h i v e \n
[...]
1667/1:     174.6658     write(2, 0x0A4C9634, 88)                 = 88
1667/1:       p k g :   ' / s b i n / b o o t a d m   u p d a t e - a r c h i
1667/1:       v e   - R   / t m p / t m p C 6 b M r 1 '   f a i l e d . \n w
1667/1:       i t h   a   r e t u r n   c o d e   o f   1 . \n
[...]
```

From the error message above we may conclude that /usr/lib/siconv/ is missing on this system but this cannot be the real problem as this directory does not exists on my lab system either (where the bootadm/mkisofs command works without problems).

So this reference to /usr/lib/siconv/ is misleading and we need to continue to check why the mkisofs is failing here (but working in the lab).

I will create a reference truss from the working pkg install to compare against this failing truss. This will take some time...

Notes **Oracle Support**- Jun 22, 2012 9:24 AM (Friday)

Checking truss file for this reported error:

%pkg install -g ./idr232.2.p5p idr232
Packages to install: 1
Packages to update: 1
Create boot environment: Yes
Create backup boot environment: No

DOWNLOAD PKGS FILES XFER (MB)
Completed 2/2 8/8 0.8/0.8$<3>

PHASE ACTIONS
Removal Phase 1/1
Install Phase 7/7
Update Phase 4/4

PHASE ITEMS
Package State Update Phase 3/3
Package Cache Update Phase 1/1
Image State Update Phase 2/2

PHASE ITEMS
Reading Existing Index 8/8
Indexing Packages 2/2
pkg: '/sbin/bootadm update-archive -R /tmp/tmpsImdFP' failed.
with a return code of 1.

Update from Customer **CHANG_AN@YAHOO.COM**- Jun 21, 2012 7:05 PM (Thursday)
Upload to gtcr successful for the file truss.out.gz.

Update from Customer **CHANG_AN@YAHOO.COM**- Jun 21, 2012 7:04 PM (Thursday)
After rebooting from the new BE (with IDR232), it works now!!

Update from Customer **CHANG_AN@YAHOO.COM**- Jun 21, 2012 6:45 PM (Thursday)
The truss.out is about 2 GB.
Can I upload here?
Or, do I need to upload to an FTP server?

-rw-r--r-- 1 root root 2028612961 Jun 21 14:32 truss.out

Let me try to compress it first to see how much I can shrink.

It has shrunk to around 240 MB, still big.

-rw-r--r-- 1 root root 245984814 Jun 21 14:32 truss.out.gz

Notes **Oracle Support**- Jun 21, 2012 6:14 PM (Thursday)
Who: customer Chang-An Hsiao
What: roll back to the original (unpatched) environment and collect truss data from the failing pkg install command
When: by June, 28th

Update from Customer **CHANG_AN@YAHOO.COM**- Jun 21, 2012 6:12 PM (Thursday)
I was able to boot from the new BE but it did not help.

I will generate the truss output for you shortly.
We an reconvene tomorrow.

Enjoy your evening!

CHang-An

ODM Action Plan **Oracle Support**- Jun 21, 2012 5:51 PM (Thursday)
Who: customer Chang-An Hsiao
What: boot from the new environment solaris-1. If this fails (and only then) boot again from the original boot environment and collect truss data as per previous note
When: by June, 29th

Notes **Oracle Support**- Jun 21, 2012 5:50 PM (Thursday)
So does the boot from the environment fail? If not then there is no need to rollback here.
The rollback as well as the truss data collection is ONLY to be used if the boot from the new environment fials.

Please note that this is the end of my shift here and any further updates will continue tomorrow starting at 09:00 UTC.

Bye,
Wolfgang Ley.

Update from Customer **CHANG_AN@YAHOO.COM**- Jun 21, 2012 5:45 PM (Thursday)
OK, I will try to rol back and try from there.

Thanks!

ODM Action Plan **Oracle Support**- Jun 21, 2012 5:30 PM (Thursday)
Who: customer Chang-An Hsiao
What: boot from the new environment solaris-1. If this fails (and only then) boot again from the original boot environment and collect truss data as per previous note
When: by June, 29th

Notes **Oracle Support**- Jun 21, 2012 5:29 PM (Thursday)

Installing the same pklg multiple times is NOT recommended. There is no additional logfile available here.

As per previous update: please try to boot from your new created boot environments (if you have created multiple due to the various tests then you may want to roll back to the very first and retry).
To see the list of created boot environments use "beadm list"
Use the beadm to activate the original solaris boot environement, reoboot the system to boot from the original (unpatches environment), delete the other boot environemnts and then retry the IDR installation ONCE (not multiple times).
Then try a boot from the new created boot environment. If this fails then boot back from the original enbvironment and collect the truss data as explained in my previous note. Please do NOT try to install the same IDR multiple times! Thanks.

Update from Customer **CHANG_AN@YAHOO.COM**- Jun 21, 2012 5:25 PM (Thursday)
I tried to install the package again and check the /tmp to see what's there during the installation.

I could see the tmpQRXs6W in /tmp but it still reported failed after the process.

pkg: '/sbin/bootadm update-archive -R /tmp/tmpQRXs6W' failed.
with a return code of 1.

A clone of solaris exists and has been updated and activated.
On the next boot the Boot Environment solaris-3 will be
mounted on '/'. Reboot when ready to switch to this updated BE.

Is there a log file for me to refer to this error message?

I couldn't find the information about this idr though.

%pkg info idr232
pkg: info: no packages matching the following patterns you specified are
installed on the system. Try specifying -r to query remotely:

idr232

Where do I find the solaris-1/2/3?

Notes **Oracle Support**- Jun 21, 2012 5:23 PM (Thursday)
You can try to boot from the solaris-1 boot environment. If this works then everything is ok.
If this fails then you will need to boot from your original boot environment and we would need to check why your pkg command fials here. In that case (after booting from the original boot environment and only if the boot of solaris-1 fails) please runn the following command to get the truss data of the failing pkg install command:

truss -aefdl -rall -wall -vall -xall -o /var/tmp/truss.out pkg install -f ./idr232.2.p5p idr232

If this fails again then please run bzip2 on the created outputfile /var/tmp/truss.out and sent that dilfe for analysis to me. Thanks.

Bye,
Wolfgang Ley.

Notes **Oracle Support**- Jun 21, 2012 5:18 PM (Thursday)
The directory is created during the pkg install and will be removed afterwards. So it is clear that you cannot see it right now anymore.
The temporary directory is used to hold the new boot archive which maybe several MB large and the required minimum is therefore to ahve 1GB RAM - but this should not be your problem either.

Without a truss from the failure there is really nothing I can do here. As said: the IDR installs cleanly here so this is not a general issue with the IDR but rather a system specific issue with your setup/system.

Update from Customer **CHANG_AN@YAHOO.COM**- Jun 21, 2012 5:11 PM (Thursday)
I was a root user and root owns / while /tmp is wide open for all users.

As you know too, I am not using non-global zone. I am not sure why it is showing the error message. I couldn't find "/tmp/tmpsImdFP". Could that be the reason if failed because it couldn't find it too?

ODM Action Plan **Oracle Support**- Jun 21, 2012 4:46 PM (Thursday)
Who: customer Chang-An Hsiao
What: verify that the provided IDR232.2 allows tuning via /etc/system and setting of ip:tcp_init_wnd_chk (as a workaround until the patches from the firewall vendor are available)
When: by June, 29th

Notes **Oracle Support**- Jun 21, 2012 4:44 PM (Thursday)
Hi,

an return code of 1 indicates a broken permission (or a missing privilege during pkg install). I have verified that the IDR installation works without this error message here in the lab so I do not know if/which permission (maybe / or /tmp is/was broken here.
Are you using non-global zones?This was not the case when I got the last explorer from you. If you are using zones then please see the document 1452392.1 on how to install the IDR (as the pkg install may indeed fial here).

A broken boot archive may likely cause the system to be not bootable. In this case you can still go back to the original boot environment (not the new created solaris-1) and retry the idr installation.

Bye,
Wolfgang Ley.

Update from Customer **CHANG_AN@YAHOO.COM**- Jun 21, 2012 4:39 PM (Thursday)
I got one failed message:

%pkg install -g ./idr232.2.p5p idr232
Packages to install: 1
Packages to update: 1
Create boot environment: Yes
Create backup boot environment: No

DOWNLOAD PKGS FILES XFER (MB)
Completed 2/2 8/8 0.8/0.8$<3>

PHASE ACTIONS
Removal Phase 1/1
Install Phase 7/7
Update Phase 4/4

PHASE ITEMS
Package State Update Phase 3/3
Package Cache Update Phase 1/1
Image State Update Phase 2/2

PHASE ITEMS
Reading Existing Index 8/8
Indexing Packages 2/2
pkg: '/sbin/bootadm update-archive -R /tmp/tmpsImdFP' failed.
with a return code of 1.

A clone of solaris exists and has been updated and activated.
On the next boot the Boot Environment solaris-1 will be
mounted on '/'. Reboot when ready to switch to this updated BE.

Should I go ahead and reboot it after having the following line in /etc/system?

set ip:tcp_init_wnd_chk = 512

Update from Customer **CHANG_AN@YAHOO.COM**- Jun 21, 2012 4:33 PM (Thursday)
Thanks very much!

I am in the process of installing the IDR and will reboot it to see how it works.

I will keep you posted shortly.

Chang-An

ODM Action Plan **Oracle Support**- Jun 21, 2012 3:28 PM (Thursday)
Who: customer Chang-An Hsiao
What: verify that the provided IDR232.2 allows tuning via /etc/system and setting of ip:tcp_init_wnd_chk (as a workaround until the patches from the firewall vendor are available)
When: by June, 29th

Notes **Oracle Support**- Jun 21, 2012 3:26 PM (Thursday)
Hi,

a new version of the IDR (Interim Diagnostic/Relief) 232 has been uploaded to this service request.
Please check the attachement section of this service request where you should be able to find a file called idr232.2.p5p
download the file idr232.2.p5p and install it by using the command "pkg install -g ./idr232.2.p5p idr232". See document 1452392.1 for more information about Solaris 11 IDRs and how to apply or remove them.

A reboot is requied after the IDR installation to activate the fix.
Once the fix is active you should be able to tune the lowest allowed initial window size to be even lower than an ethernet packet.
This can be done by using /etc/system and adding a line such as

set ip:tcp_init_wnd_chk = 512

(or use an even lower value such as 100).

Please ntoe that after any change in /etc/system another reboot is required to activate the new setting.

Please let me know if this IDR works for you and allows you to tune the window size checks until you get the final fix from your firewall vendor. Thanks.

Best regards,
Wolfgang Ley.

Notes **Oracle Support**- Jun 21, 2012 3:11 PM (Thursday)
Upload to gtcr successful for the file idr232.2.p5p.

Notes **Oracle Support**- Jun 21, 2012 3:06 PM (Thursday)
idr232.1.p5p successfully removed by support

Notes **Oracle Support**- Jun 21, 2012 3:03 PM (Thursday)
A new IDR 232.2 was created and this should work now. I will delete the currenrt IDR 232.1 from this SR and get the newer IDR version later today

Notes **Oracle Support**- Jun 21, 2012 2:15 PM (Thursday)
Update: Please do NOT use the provided IDR. The development group came back to me (as I had some concerns when testing the IDR in my lab) and they found a build issue with this IDR.
They will rebuild an IDR an dI will provide an update as soon as I do have news here.

ODM Action Plan **Oracle Support**- Jun 21, 2012 1:56 PM (Thursday)
Who: Solaris sustaining / development
What: provide binary fix for Solaris 11 tunable tcp_init_wnd_chk (bug escalation 7071362) --- Customer to check provided IDR on his own risk but in my lab we have indications that this IDR will not help yet
When: by June, 27th

Who: customer Chang-An Hsiao
What: contact vendor of the used firewall/NAT to get a fix for bogus window size advertised from there (point them to CVE-2008-4609)
When: by June, 27th

Notes **Oracle Support**- Jun 21, 2012 1:01 PM (Thursday)
I have received an IDR (Interim Diiagnostic/Relief) which is now attached to this customer.
The support document 1452392.1 explains how to install and remvoe the IDR the IDR.

However: A quick check in my lab here showed that the disassembling of the affected function did not changed. So I am not sure whether the tunable tcp_init_wnd_chk (in /etc/system) does now really work (after applying the IDR).
I do not have the test environment (your special firewall with the broken window sizes) so I cannot check the IDR to see whether the unable is working now.
I have asked our engineering for clarification regarding my concerns but if you want to test the IDR by yourself then please feel free to download and install the idr232.1.p5p (see attachment section of this SR)

Best regards,
Wolfgang Ley

Notes **Oracle Support**- Jun 21, 2012 12:21 PM (Thursday)
Upload to gtcr successful for the file idr232.1.p5p.

ODM Action Plan **Oracle Support**- Jun 13, 2012 8:57 AM (13 days ago)
Who: Solaris sustaining / development
What: provide binary fix for Solaris 11 tunable tcp_init_wnd_chk (bug escalation 7071362)
When: by June, 27th

Who: customer Chang-An Hsiao
What: contact vendor of the used firewall/NAT to get a fix for bogus window size advertised from there (point them to CVE-2008-4609)
When: by June, 27th

Update from Customer **CHANG_AN@YAHOO.COM**- Jun 12, 2012 4:04 PM (14 days ago)
Thanks for the update!!

I have been checking with Actiontec every week. The latest update I got was they are looking into the case. No further information is available yet.

Thanks!!

ODM Action Plan **Oracle Support**- Jun 12, 2012 1:26 PM (14 days ago)
Who: Solaris sustaining / development
What: provide binary fix for Solaris 11 tunable tcp_init_wnd_chk (bug escalation 7071362)
When: by June, 27th

Who: customer Chang-An Hsiao
What: contact vendor of the used firewall/NAT to get a fix for bogus window size advertised from there (point them to CVE-2008-4609)
When: by June, 27th

Notes **Oracle Support**- Jun 12, 2012 1:24 PM (14 days ago)
I haven;t received any fix (well - actually workaround) from our engineering yet. I have pinged them for the status and will provide an update as soon as I do have more news.
Have you heard anything back from the firewall/nat vendor regardingthe final fix?

ODM Action Plan **Oracle Support**- May 29, 2012 4:16 PM (28 days ago)
Who: Solaris sustaining / development
What: provide binary fix for Solaris 11 tunable tcp_init_wnd_chk (bug escalation 7071362)
When: by June, 12th

Who: customer Chang-An Hsiao
What: contact vendor of the used firewall/NAT to get a fix for bogus window size advertised from there (point them to CVE-2008-4609)
When: by June, 12th

Notes **Oracle Support**- May 29, 2012 4:16 PM (28 days ago)
Hi,

thanks for the update. I will let you know if I have news from our engineering regarding the option to disable the security protection (as a workaround).

Bye,
Wolfgang Ley.

Update from Customer **CHANG_AN@YAHOO.COM**- May 29, 2012 4:13 PM (28 days ago)

Hi, Wolfgang:

Thanks for the information. I know it may take some time for the engineering to come up with a fix.

I have checked with the vendor on a regular basis. They are back logged and just started to review my service request. I don't have any further information from the vendor. I will ping them again today to see if they really picked up the ticket. I also mentioned you may be able to join the conference if they need the input from you and they are aware of the potential help.

Thanks!

Chang-An

ODM Action Plan **Oracle Support**- May 29, 2012 2:18 PM (28 days ago)
Who: Solaris sustaining / development
What: provide binary fix for Solaris 11 tunable tcp_init_wnd_chk (bug escalation 7071362)
When: by June, 12th

Who: customer Chang-An Hsiao
What: contact vendor of the used firewall/NAT to get a fix for bogus window size advertised from there (point them to CVE-2008-4609)
When: by June, 12th

Notes **Oracle Support**- May 29, 2012 2:17 PM (28 days ago)
Hi,

the IDR (interim Diagnostic/Relief) fix to allow you to turn off the Solaris security prtection against bogus MSS advertisements (which are causing the connection drops here) are not yet available.
I have pinged our development regarding the status and will provide an update as soon as I do have news from there.

Have you received an update/solution from the firewall vendor yo actually fix the root cause there? Disabling the Solaris security prot4ections (once the IDR is available) is only a workaround and would open your system to certain denial of service attacks.

Best regards,
Wolfgang Ley

Update from Customer **CHANG_AN@YAHOO.COM**- May 14, 2012 4:42 PM (1+ month ago)
Thanks!!

ODM Action Plan **Oracle Support**- May 13, 2012 12:22 PM (1+ month ago)
Who: Solaris sustaining / development
What: provide binary fix for Solaris 11 tunable tcp_init_wnd_chk (bug escalation 7071362)
When: by May, 31st

Who: customer Chang-An Hsiao
What: contact vendor of the used firewall/NAT to get a fix for bogus window size advertised from there (point them to CVE-2008-4609)
When: by May, 31st

Notes **Oracle Support**- May 13, 2012 12:20 PM (1+ month ago)
Hi,

I have forwarded the request to our development and will let you know whether they will provide a binary fix for the unpatched kernel (which you are using right now) or whether some upgrade is required first.
I will provide an update as soon as I do have news from our development but this may take some time. Please continue to work with the firewall vendor for the final fix here. (while we continue to work here to provide a workaround to allow you to disable the security protections in Solaris).

Best regards,
Wolfgang Ley.

Update from Customer **CHANG_AN@YAHOO.COM**- May 13, 2012 5:55 AM (1+ month ago)
I see.

I can try the interim solution (binary fix).
Please let me know if I need to upgrade to certain kernel level before I can apply this binary fix.

Indeed, this will be applied on cahtoh02, the host I created the explorer on.

Thanks!

ODM Action Plan **Oracle Support**- May 12, 2012 2:15 PM (1+ month ago)
Who: customer Chang-An Hsiao
What: provide feedback on whether a binary fix for host cahtoh02 is requested. Continue to work with the firewall vendor to get the final fix from there.
When: by May, 25th

Notes **Oracle Support**- May 12, 2012 2:14 PM (1+ month ago)
Hi,

the real problem is with the firewall and the broken window size advertisement from there but we still have one issue with Solaris 11 where the tunable tcp_init_wnd_chk (to disable the security protection against the window size attacks) is not working.
We are in the progress of fixing this tunable. Do you want to test a binary fix for this (until the patch will become available)?
If you are interested in such a fix: The provided explorer from host cahtoh02 shows that you are not using any Solaris 11 patches yet. An interim solution (binary fix) will have to match a particular kernel build so we need confirmation that a requested fix should really be for the currently unpatched host cahtoh02.

Can you please let me know whether you want to test a fix (to renable the tunable which alows you to turn off the security protections against the window size attacks) and whether this is requested for the unpatched host cahtoh02? Thanks.

Best regards,
Wolfgang Ley.

ODM Action Plan **Oracle Support**- May 9, 2012 6:32 PM (1+ month ago)
who: service request owner Wolfgang Ley
What: provide an update as soon as we do have news from engaged engineering on a correction of the non-working tunable
When: by May, 26th

ODM Action Plan **Oracle Support**- May 9, 2012 4:43 PM (1+ month ago)
who: service request owner Wolfgang Ley
What: provide an update as soon as we do have news from engaged engineering on a correction of the non-working tunable
When: by May, 26th

Notes **Oracle Support**- May 9, 2012 4:41 PM (1+ month ago)
Hi,

well - the root cause as been found as the brogus incoming window size and not on the Solaris 11 host.
The bug identified on the Solaris 11 x86 host is the fact that th tunable to disable the security protection does not work here (bug creates is 7167477). This is not the bug causing your issue but a problem that we cannot workaround the issue.
I will check whether we can get a fix to provide the workaround variable on the Solaris 11 host but this is not a fix for your real problem (the bogus window size). Other hosts with similar protections may/will suffer similar issues when acting as a server here.

regarding your positive feedback to my supervisor/manager: You may get an invitation to provide feedback after the service request has been closed. The case closure information will also include a contact email address to opt-in for such a feedback if you are not automatically selected. Within this service request feedback you will find a question on whether you want to nominate me for an excellence service award.. The feedback there will be directly passed no to my manager.

However: the case closure may still take some time as I will first try to see whether we can provide the tunable here (I will let you know if I have news here). I will also be available for questions from Actiontec (if needed).
I have engaged our engineering regarding the non-working tunable and will provide an update as soon as I have news from there (which however may take some time).

Best regards,
Wolfgang Ley.

Update from Customer **CHANG_AN@YAHOO.COM**- May 9, 2012 4:14 PM (1+ month ago)
Hi, Wolfgang:

Thanks for helping find the cause and a bug in Solaris 11 x86.

I have forwarded your findings and suggestion to the vendor.

I know it takes time to fix bugs (either from Oracle/SUN or Actiontec) but I trust SUN more.

I guess I will have to wait and see what outcome comes from Actiontec.

Your efforts are highly appreciated!

Are there ways I can tell your supervisors about your hard work ? I want to give you the credits you deserve to receive.

Let me know!

Chang-An

ODM Action Plan **Oracle Support**- May 9, 2012 12:24 PM (1+ month ago)
Who: custo0mer Chang-An Hsiao
What: contact the vendor of the used firewall/NAT system to get a fix for the broken window size advertisement
When: by May, 17th

Notes **Oracle Support**- May 9, 2012 12:23 PM (1+ month ago)
Hi Chang-An,

I am afraid that the tunable on Solaris 11 x86 does not work (due to a compiler optimization). A bug has been filed to document this but right now we do not have tunable to turn off the TCP/IP security protection on Solaris (against CVE-2008-4609).
The only solution to this would be to contact the firewall/NAT vendor and ask them to correct the issue and advertise initial window sizes which are at least the MSS (MTU - TCP/IP header size). The TCP/IP standard recommends to advertise at least 4*MSS and usually a much higher value is used to get sufficient performance.

Best regards,
Wolfgang ley.

ODM Solution/Action Plan **Oracle Support**- May 9, 2012 12:21 PM (1+ month ago)
Contact the vendor to get a fix for the broken window size advertisements which are discarded by clients with the protection against the security vulnerability CVE-2008-4609 (see http://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2008-4609 for more details)

ODM Proposed Solution Justif **Oracle Support**- May 9, 2012 12:19 PM (1+ month ago)
Solaris side cannot be tuned (new bug has been created to document that tcp_init_wnd_chk tunable does not work on Solaris 11 x86(

ODM Proposed Solution(s) **Oracle Support**- May 9, 2012 12:18 PM (1+ month ago)
Contact the firewall vendor for a fix on the broken window size

ODM Cause Justification **Oracle Support**- May 9, 2012 12:17 PM (1+ month ago)

The provided tcpdump data proofs that the client is advertising a larger window size and is setting the window scale to 7 (multiply by 128).
The firewall is using this window scale option and then advertises a window size of 11. So the initial window size seen by the Solaris host is 11 * 128 = 1408 which is less than the required minimum of 1460 (maximum segment size) and the packet will therefore be dropped.
The kernel stattcp:0:tcpstat:tcp_zwin_ack_syn confirms that this check causes the connection abort.

ODM Cause Determination **Oracle Support**- May 9, 2012 12:15 PM (1+ month ago)

The issue is caused by a bogus incoming final ACK on the 3-way handhake when setting up a TCP connection. The sender (firewall/NAT in this case) is advertising an initial window size which is smaller than the segment size.
The issue is triggered by a remote client using window scale option (usualy Linux client). The used firewall is using a fixed window size (1460 which is the local MTU - TCP header) and when trying to use window scaling together with this fixed window size this fails.
The firewall limit is divded by the scale factor and the rounding error will then case a final /real window size which is smaller than the MSS.

Notes **Oracle Support**- May 9, 2012 12:12 PM (1+ month ago)

I have created a new bug to document that the tunable tcp_init_wnd_chk does not work on Solaris 11 x86 (due to a compiler optimization issue). The reference is 7167477 but I do not know if/when this bug will be visible in the mY Oracle Support portal.
So right now the security protection from Solaris against bogus window sizes (see security vulnerability CVE-2008-4609) is not tunable and the only place where this can be fixed is the firewall/NAT appliance which is advertising the bogus window sizes.

Best regards,
Wolfgang Ley.

ODM Action Plan **Oracle Support**- May 9, 2012 11:40 AM (1+ month ago)

Who: service request owner Wolfgang Ley
What: file a new bug to document that the tunable tcp_init_wnd_chk does not work on Solaris 11 x86 (due to compiler optimization).
Wehn: by May, 15th

Who: customer Chang-An Hsiao
What: contact vendor of the used firewall/NAT to get a fix for bogus window size advertised from there (point them to CVE-2008-4609)
When: by May, 17th

Notes **Oracle Support**- May 9, 2012 11:38 AM (1+ month ago)

Due to compiler optimization the tunable does not work on Solaris 11 x86. I will file a bug for this but right now it looks like we cannot cahnge much on the behaviour of the incoming packet handling with broken window sizes.

Notes **Oracle Support**- May 9, 2012 11:37 AM (1+ month ago)

From the source code we can see:

```
[...]
static uint32_t tcp_init_wnd_chk = 4096;
[...]
pinit_wnd = ntohs(tcpha->tha_win) << tcp->tcp_snd_ws;
if (pinit_wnd < tcp->tcp_mss &&
pinit_wnd < tcp_init_wnd_chk) {
freemsg(mp);
TCP_STAT(tcps, tcp_zwin_ack_syn);
tcp->tcp_second_ctimer_threshold =
tcp_early_abort * SECONDS;
return;
}
[...]
```

But checking the disassemler code of tcp_ionput_data() we can see
the following generated code:

```
[...]
tcp_input_data+0x3c8: movw 0xe(%rbx),%ax
tcp_input_data+0x3cc: bswap %eax
tcp_input_data+0x3ce: shrl $0x10,%eax
tcp_input_data+0x3d1: movl 0x30(%r15),%ecx
tcp_input_data+0x3d5: shll %cl,%eax
tcp_input_data+0x3d7: cmpl 0x5c(%r15),%eax
tcp_input_data+0x3db: jae +0x59 <tcp_input_data+0x436>
tcp_input_data+0x3dd: cmpl $0x1000,%eax
tcp_input_data+0x3e2: jb +0xb <tcp_input_data+0x3ef>
tcp_input_data+0x3e4: movl 0xffffffffffffed8(%rbp),%edx
tcp_input_data+0x3ea: jmp +0x6fb <tcp_input_data+0xaea>
tcp_input_data+0x3ef: movq %r12,%rdi
tcp_input_data+0x3f2: call +0x3e92b71 <freemsg>
tcp_input_data+0x3f7: movq %gs:0x10,%rax
tcp_input_data+0x400: movslq 0x4(%rax),%rax
tcp_input_data+0x404: movq 0xfffffffffffffe88(%rbp),%rcx
tcp_input_data+0x40b: movq 0x1b8(%rcx),%rcx
tcp_input_data+0x412: movq (%rcx,%rax,8),%rax
tcp_input_data+0x416: incq 0x1f8(%rax)
tcp_input_data+0x41d: movl -0x37c347db(%rip),%eax <tcp_early_abort
[...]
```

The problem is the fixed comparision here:
tcp_input_data+0x3dd: cmpl $0x1000,%eax

Note that this uses the fixed value 0x1000 instead of the variable tcp_init_wnd_chk

This is a problem of the C-Compiler optimization which has detected that the
variable tcp_init_wnd_chk is declared as 4096 and only used once for the comparison
while no other code changes or uses this values. As a result the compiler is using
directly the values 4096 (0x1000) instead of the variable --- which breaks the option
to change this value here :-(

Update from Customer **CHANG_AN@YAHOO.COM**- May 8, 2012 6:24 PM (1+ month ago)
Thanks a lot!

I don't need immediate response, for I enjoy working with you and I don't want to take the risk of getting next engineer who may not be as good as you are. :)

Enjoy your evening.
We can reconvene tomorrow.

Thanks!

ODM Action Plan **Oracle Support**- May 8, 2012 6:21 PM (1+ month ago)
Who: service request owner Wolfgang Ley
What: continue code inspection to see why the proposed workaround did not helped here
When: by May 17th,

Who: customer Chang-An Hsiao
What: contact vendor of the used firewall/NAT to get a fix for bogus window size advertised from there (point them to CVE-2008-4609)
When: by May, 17th

Notes **Oracle Support**- May 8, 2012 6:20 PM (1+ month ago)
Strange that even 1 will nto help . Will check the new dumps tomorrow (as I am already offline for quite some time). Please note my usual working hours of 09:00 - 17:00 UTC *GMT).
If you need help outside these working hours then please contact the support line and ask for the next available engineer (which then however would also cause myself to no longer be involved here).

Best regards,
Wolfgang ley

Update from Customer **CHANG_AN@YAHOO.COM**- May 8, 2012 6:18 PM (1+ month ago)
After tcpdump again, it increased to 3:

%kstat -p tcp:0:tcpstat:tcp_zwin_ack_syn
tcp:0:tcpstat:tcp_zwin_ack_syn 3

Update from Customer **CHANG_AN@YAHOO.COM**- May 8, 2012 6:16 PM (1+ month ago)
Upload to gtcr successful for the file tcpdump_server.2012050802.

Update from Customer **CHANG_AN@YAHOO.COM**- May 8, 2012 6:15 PM (1+ month ago)
Upload to gtcr successful for the file tcpdump_client.2012050802.

ODM Action Plan **Oracle Support**- May 8, 2012 6:12 PM (1+ month ago)
Who: customer Chang-An Hsiao
What: set the value of tcp_init_wnd_chk (in /etc/system) to 1 to disable the security check/protection against bogus window size advertisements. Provide feedback whether this workaround helps here.
When: by May, 15th

Update from Customer **CHANG_AN@YAHOO.COM**- May 8, 2012 6:12 PM (1+ month ago)
After changing kstat -p tcp:0:tcpstat:tcp_zwin_ack_syn to 1, it still failed.....

%kstat -p tcp:0:tcpstat:tcp_zwin_ack_syn
tcp:0:tcpstat:tcp_zwin_ack_syn 2

%mdb -k
Loading modules: [unix genunix specfs dtrace mac cpu.generic uppc pcplusmp scsi_vhci zfs ip hook neti arp usba kssl sd fctl s1394 sockfs lofs random idm crypto nfs sppp sata cpc fcip logindmux ptm ufs ipc]
> tcp_init_wnd_chk/D
tcp_init_wnd_chk:
tcp_init_wnd_chk: 1

Do you need 2 more tcpdump files?

Update from Customer **CHANG_AN@YAHOO.COM**- May 8, 2012 6:04 PM (1+ month ago)
Ah....You are in Germany! Got you.

Let me try to reset the value of tcp_init_wnd_chk to 1 and see how it takes it.

I will just send Actiontec the tcpdump files and ask them to load into wireshark.

Will update you shortly.

ODM Action Plan **Oracle Support**- May 8, 2012 5:52 PM (1+ month ago)

Who: customer Chang-An Hsiao
What: set the value of tcp_init_wnd_chk (in /etc/system) to 1 to disable the security check/protection against bogus window size advertisements. Provide feedback whether this workaround helps here.
When: by May, 15th

Notes **Oracle Support**- May 8, 2012 5:50 PM (1+ month ago)
Hi,

you can simply send him the tcpdumps that you have send me. These can be loaded into wireshark without problems (which is what I did here).
I can talk to Actiontec if required (if they do not get or see the window scale and window size problem here). Please note that my usual working hours are 09:00 - 17:00 UTC (GMT+0) as I am located in Germany, Europe.

I now already found the bug in the code why the workaround was not working here. It looks like we have applied our sending window scale (which is 1) instead of the receiving window scale (the 128 from the client) for this check.
As a result the window size of 11 will be used for the check which indeed still fails to pass the workaround.

Can you please retry the test and set tcp_init_wnd_chk to 1 in /etc/system./ This will turn off the security protection against the mentioned window size secruity issue but I have also checked that this tunable is only used in one particular part (the failing size chekc) and setting this value to 1 does not have any other/additional imapcts.

Best regards,
Wolfgang Ley.

Update from Customer **CHANG_AN@YAHOO.COM**- May 8, 2012 5:39 PM (1+ month ago)
Thank you so much! I know it will take some time and I do appreciate your taking extra miles to troubleshoot.

I have also opened a ticket with Actiontec. The assigned engineer will ask me for wireshark (formerly known ethereal) printout for analysis. Hopefully, he will provide me the instruction of what informaiton he wants to collect from me.

Then, we can compare the findings between Actiontec and Oracle/SUN.

Would you be willing to talk directly with Actiontec if that happens in the near future?

ODM Action Plan **Oracle Support**- May 8, 2012 5:25 PM (1+ month ago)
Who: service request owner Wolfgang Ley
What: continue code inspection to see why the proposed workaround did not helped here
When: by May 17th,

Who: customer Chang-An Hsiao
What: contact vendor of the used firewall/NAT to get a fix for bogus window size advertised from there (point them to CVE-2008-4609)
When: by May, 15th

Notes **Oracle Support**- May 8, 2012 5:23 PM (1+ month ago)
The setting should not make any difference and 128 should be ok, too. The kstat however shows that the workaround did not helped and the packet was still rejected due to the window size security checks.
The tcpdump dump data confirms this (the connection was done on the ssh standard port this time). On the server side the packet #14 shows the bogus advertised incoming window size of 11 (with scale factor 128 leads us to 1406 which is less than one segment size of 1460).

I will need to dig through the code to see why this tunable did not helped here.. This will take some time though.

Note that all we can try to do here is to find a workaround on the Solaris die. The real problem is still with the firewall/NAT vendor and the final solution needs to be provided from there. Setting an initial window size (withouth any previous data) which is less than a local segment size (ethernet MTU - tcp/ip header) is seriously broken and considered an attack as explained in the CVE-2008-4609 entry.

Update from Customer **CHANG_AN@YAHOO.COM**- May 8, 2012 5:14 PM (1+ month ago)
%mdb -k
Loading modules: [unix genunix specfs dtrace mac cpu.generic uppc pcplusmp scsi_vhci zfs ip hook neti arp usba kssl sd fctl s1394 sockfs lofs random idm crypto nfs sppp sata cpc fcip logindmux ptm ufs ipc]
> tcp_init_wnd_chk/D
tcp_init_wnd_chk:
tcp_init_wnd_chk: 128

Update from Customer **CHANG_AN@YAHOO.COM**- May 8, 2012 5:12 PM (1+ month ago)
I have uploaded 2 tcpdump files a few seconds ago.

Before tcpdump:

%kstat -p tcp:0:tcpstat:tcp_zwin_ack_syn
tcp:0:tcpstat:tcp_zwin_ack_syn 1

After tcpdump:

%kstat -p tcp:0:tcpstat:tcp_zwin_ack_syn
tcp:0:tcpstat:tcp_zwin_ack_syn 2

BTW, I set the "tcp_init_wnd_chk" to 128 for testing purpose and the tcpdump is done under 128, not 512. Does it make any differences?

Update from Customer **CHANG_AN@YAHOO.COM**- May 8, 2012 5:09 PM (1+ month ago)
Upload to gtcr successful for the file tcpdump_server.2012050801.

Update from Customer **CHANG_AN@YAHOO.COM**- May 8, 2012 5:09 PM (1+ month ago)

Upload to gtcr successful for the file tcpdump_client.2012050801.

Update from Customer **CHANG_AN@YAHOO.COM**- May 8, 2012 5:02 PM (1+ month ago)
All right.
I will send the 2 tcpdump files shortly.

%kstat -p tcp:0:tcpstat:tcp_zwin_ack_syn
tcp:0:tcpstat:tcp_zwin_ack_syn 1

This is before the tcpdump value.

ODM Action Plan **Oracle Support**- May 8, 2012 9:12 AM (1+ month ago)
Who: customer Chang-An Hsiao
What: check kstat parameter tcp:0:tcpstat:tcp_zwin_ack_syn before and after the failure. Provide two new tcpdumps files from the failure to a system with the workaround applied.
When: by May, 15th

Notes **Oracle Support**- May 8, 2012 9:11 AM (1+ month ago)
Hi,

sorry for the late response but I was not in the office on Monday (as explained in an earlier update from me).
I am surprised that the workaround did not helped here and if you have used mdb then this is indeed the ocrrect way to check the setting.
So we may have more than one issue with this firewall/NAT here (or the workaround does not help as expected)..

I will nbeed to have new data from the system with the applied workaround. Can you please first check the current value of the kernel statistics counter for the dropped connections due to bogus window size advertizement in the first ACK from the client?:
kstat -p tcp:0:tcpstat:tcp_zwin_ack_syn
Then please start tcpdump on the server and the failing client and retry the failing connection.
After the connection failure please stop the tcpdumps and then recheck the kernel statistic above to see whether this has increased here.
Send the two tpcudmps so I can check how the connection looks now. Thanks.

Best regards,
Wolfgang Ley.

Update from Customer **CHANG_AN@YAHOO.COM**- May 7, 2012 6:36 PM (1+ month ago)
I used mdb to check the value and the 512 is set as planned. Did not help SSH connection though.

%mdb -k
Loading modules: [unix genunix specfs dtrace mac cpu.generic uppc pcplusmp scsi_vhci zfs ip hook neti arp usba kssl sd fctl s1394 sockfs lofs random idm crypto nfs sppp sata fcp cpc fcip nsmb logindmux ptm ufs ipc]
> tcp_init_wnd_chk/D
tcp_init_wnd_chk:
tcp_init_wnd_chk: 512

Update from Customer **CHANG_AN@YAHOO.COM**- May 7, 2012 6:00 PM (1+ month ago)
Sorry. The model of the Actiontec is MI424WR-GEN2.

Update from Customer **CHANG_AN@YAHOO.COM**- May 7, 2012 5:14 PM (1+ month ago)
Thanks a lot for the extensive information! Very informative!

After adding "set ip:tcp_init_wnd_chk = 512" to /etc/system and reboot the server, the connection still failed. How do I check if this setting is really in there after the reboot? Any command to check?

The network device is Actiontec's WI424WR.

ODM Action Plan **Oracle Support**- May 6, 2012 9:37 AM (1+ month ago)
Who: customer Chang-An Hsiao
What: contact the vendor of the firewall/NAT box for a fix. Apply workaround on affected Solaris systems. See previous two notes for details.
report back status of this service request (i.e. if workaround has been activated and whether there are any remaining questions for Oracle or whether we can close this service request)
When: by May, 16th

Notes **Oracle Support**- May 6, 2012 9:36 AM (1+ month ago)
Just out of interest: What is the type/model/vendor of the used firewall/NAT box with this window size problem? Would be good to know which setups are affected by this. Thanks.

Notes **Oracle Support**- May 6, 2012 9:32 AM (1+ month ago)
Please see my previous note for the final analysis from our Oracle Solaris side here.
The problem is not with Sun SSH but the 3-way handhake is not completed because the firewall/NAT box uses a fixed/small window size (smaller than the window size sent from the OpenSSH client) but does use the larger window scale option value 7 (multiply by 128) which is used by the OpenSSH 4.x client side.
The window size problem needs to be fixed by the firewall/NAT vendor but you can workaround this problem on the Solaris 11 side by accepting the broken ACK anyway. This can be done by adding the following line to the file /etc/system and reboot the system with "init 6" to activate the new value:

set ip:tcp_init_wnd_chk = 512

Please let me know if this helps for your failing OpenSSH testcase here.

Note that patches Solaris 10 clients (see previous note for bug and patch details) will have the same security check applied to incoming ACK packets and will also ignore the packet. Please keep this in mind if you have other Solaris systems behind a similar firewall/NAT box.

I was working over the weekend on this case and analysis. I will therefore not be in the office on Monday but return on Tuesday. If you have any questions or concerns then I can address them on Tuesday.

Notes **Oracle Support**- May 6, 2012 9:25 AM (1+ month ago)

From the analysis so far we can see that the difference between the working and the failing tcp connection was the incoming last ACK of the initial 3-way TCP handshake.

In the working case the remote side sends an ACK with a window size of 1460 (scaling 1) while the failing connections sends an ACK with a window size of 11 and scale 128 (which would give us 1408).

Checking onnv-clone/usr/src/uts/common/inet/tcp/tcp_input.c we can see the following checks for this incoming ACK:

```
2766 /*
2767 * No sane TCP stack will send such a small window
2768 * without receiving any data. Just drop this invalid
2769 * ACK. We also shorten the abort timeout in case
2770 * this is an attack.
2771 */
2772 pinit_wnd = ntohs(tcpha->tha_win) << tcp->tcp_snd_ws;
2773 if (pinit_wnd < tcp->tcp_mss &&
2774 pinit_wnd < tcp_init_wnd_chk) {
2775 freemsg(mp);
2776 TCP_STAT(tcps, tcp_zwin_ack_syn);
2777 tcp->tcp_second_ctimer_threshold =
2778 tcp_early_abort * SECONDS;
2779 return;
2780 }
```

The check is against mss (536) and tcp_init_wnd_chk (4096) and the used value 11 << 7 (1408) should be ok here. It seems that we still drop the ACK due to this check here. From the explroer we can see that the associated kstat value tcp_zwin_ack_syn is 129 so this check seems to have kicked in here. This would happen if the default mss value has been tuned (not done here) or that TCP setup has adjusted mss to a higher value.

Checking the code we can see that the mss will indeed be set to MTU - 40 (space for the tcp header) and this would be 1460 here. And that is our problem as the 11 << 128 is lower than 1460 and the ACK therefore ignored.

The sanity check was implemented to fix the bug
6759500 [CVE-2008-4609] FICORA #193744 TCP vulnerabilities
and this bugfix is integrated in Solaris 11. Note that this fix is also in Solaris 10 starting with patch 144488-06 (Sparc) and 144489-06 (x86)

The implemented check assures that the used window size is not smaller than the the default MSS (derived from the MTU - header size) but if the window size is larger than tcp_init_wnd_chk (4096) then we allow it anyway.

```
*
* To protect TCP against attacker using a small window and requesting
* large amount of data (DoS attack by conuming memory), TCP checks the
* window advertised in the last ACK of the 3-way handshake. TCP uses
* the tcp_mss (the size of one packet) value for comparion. The window
* should be larger than tcp_mss. But while a sane TCP should advertise
* a receive window larger than or equal to 4*MSS to avoid stop and go
* tarrfic, not all TCP stacks do that. This is especially true when
* tcp_mss is a big value.
*
* To work around this issue, an additional fixed value for comparison
* is also used. If the advertised window is smaller than both tcp_mss
* and tcp_init_wnd_chk, the ACK is considered as invalid. So for large
* tcp_mss value (say, 8K), a window larger than tcp_init_wnd_chk but
* smaller than 8K is considered to be OK.
*/
tatic uint32_t tcp_init_wnd_chk = 4096;
```

We can misuse this second limit and tune this is a value < 1408 to allow the broken ACK to be accepted anyway.

So a workaround for this broken sender (the firewall/NAT box) would be to allow ACKs with a window size smaller than the MSS, e.g. 512. This can be done on the Solaris side by adding this line to /etc/system and rebooting the system:

```
set ip:tcp_init_wnd_chk = 512
```

The problem is triggered byt the newer OpenSSH client setting a larger

buffer size which causes the window scaling option to be used (set to 128) and the firewall/NAT box with a fixed window size of 1460 now computes 1460/128 which is 11.4 so the window size of 11 will be used (which in turn results to 11 << 7 = 1408 which is smaller than the MSS).

Note that the OpenSSH client is only the trigger and that client is not doing anything wrong. The problem is the firewall/NAT box with the fixed smaller window size and the additional problem that is uses the clients window scale option even for the smaller window sizes (which causes the rounding problem and the total window size smaller than MSS).

Update from Customer **CHANG_AN@YAHOO.COM**- May 4, 2012 8:16 PM (1+ month ago)

Good finding!
I did see the different window sizes/scales before when I was viewing the snoop information.

I actually reached out to the network device manufacturer and asked for their back line engineers to look at this. I haven't got any responses back yet.

Thank yoiu very much for your assitance.
Have a nice weekend!

ODM Action Plan **Oracle Support**- May 4, 2012 6:47 PM (1+ month ago)

Who: customer Chang-An Hsiao
What: check whether the used window size on the failing OpenSSH 4.x client can be reduced and investigate why the firewall/NAT system is blocking/ignoring the resend packets from the Solaris 11 with the SYn/ACK
When: by May, 25th

Who: service request owner Wolfgang Ley
What: check why the Solaris 11 system is ignoring the ACK pacjet if the client uses window size of 11 and scale 128
When: by May, 25th

Notes **Oracle Support**- May 4, 2012 6:44 PM (1+ month ago)

Comparing the two tcpdump sets (working vs. failing) we can see that the only difference is the used TCP window size advertised by the client.

On the working connection we can see the initial SYN packet from the client:
[...]
Window size value: 5840
[...]
Options: (20 bytes)
Maximum segment size: 1460 bytes
TCP SACK Permitted Option: True
Timestamps: TSval 3859729267, TSecr 0
Kind: Timestamp (8)
Length: 10
Timestamp value: 3859729267
Timestamp echo reply: 0
No-Operation (NOP)
Window scale: 0 (multiply by 1)
Kind: Window Scale (3)
Length: 3
Shift count: 0
[Multiplier: 1]
[...]

On the failing connection we can see a different window scale from the client:
[...]
Window size value: 5840
[...]
Options: (20 bytes)
Maximum segment size: 1460 bytes
TCP SACK Permitted Option: True
Timestamps: TSval 4165269939, TSecr 0
Kind: Timestamp (8)
Length: 10
Timestamp value: 4165269939
Timestamp echo reply: 0
No-Operation (NOP)
Window scale: 7 (multiply by 128)
Kind: Window Scale (3)
Length: 3
Shift count: 7
[Multiplier: 128]
[...]

This is indeed a change in OpenSSH (on the client side) which is now using a larger window size to increase performance. So the difference here is OpenSSH 3.x with the default TCP Window Scale (which works with your firewall/NAT) and OpenSSH 4.x with the increased TCP Window Scale (which is blocked by your firewall/NAT).

Note that the Sun system is always answering with the same Window Scale
(for both clients - working and failing):
[...]
Window size value: 64436
[...]
Options: (20 bytes)
TCP SACK Permitted Option: True
Timestamps: TSval 249505423, TSecr 3859729267
Kind: Timestamp (8)
Length: 10
Timestamp value: 249505423
Timestamp echo reply: 3859729267
Maximum segment size: 1460 bytes
No-Operation (NOP)
Window scale: 1 (multiply by 2)
Kind: Window Scale (3)
Length: 3
Shift count: 1
[Multiplier: 2]
[...]

The big difference is in the way the firewall/NAT system is handling
these different window sizes.
If we look at the tcpdump data from the Sun server side: In the good
example the final ACK (third packet of the TCP connection) from the
firewall to the Sun has this window size:
[...]
Window size value: 1460
[Calculated window size: 1460]
[Window size scaling factor: 1]
[...]
while in the failing connection we can see the following window size:
[...]
Window size value: 11
[Calculated window size: 1408]
[Window size scaling factor: 128]
[...]

Note that the used window sizes between client and firewall are
different. So the firewall/NAT is even changing those values here.
If we look at the tcpdump on the client then the good (working)
connection is using window size 5840 with scale 1 while the bad
(failing) client is using window sieze 46 with scale 128 (which
will give us 5888).

So the trigger is different window scale advertized by the client and
I do not know the used clients to determine whether this can be changed
in your failing OpenSSH 4.x client.
This however still leaves us with two open questions:
1) Why does the Solaris 11 system ignore the incoming final ACK
if the sender is using window size 11 and scale 128 (which gives
us a total of 1408) but accepting the ACK if the sender is using
window size 1460 and scale 1 (which gives us a total of 1460)
2) Why does the firewall/NAT ignore the resends of the SYn/ACK from
the Solaris 11 system back to the client (which then finally
causes the connection abort)?

I will continue to work on the first question but this may take some
time as I do not have such a questionable sender with window size 11
and scale 128 here.
You will need to work on the second question and should also check the
available options on your OpenSSH 4.x client to determine whether the
window size / window scale can be modified there to use something which
works here.

ODM Issue Verification **Oracle Support**- May 4, 2012 6:43 PM (1+ month ago)

When trying to open an ssh connection from an OpenSSH 4.x client over a firewall/NAT device to the Solaris 11 sun system then the TCP connection cannot be established.
Using OpenSSH 3.x on the client will not cause this problem.

Update from Customer **CHANG_AN@YAHOO.COM**- May 4, 2012 4:34 PM (1+ month ago)

Understood what you have in mind.

However, it doesn't make sense when all components (SSH clients, network device, IPs, ...) remain intact and the only change is form Solaris 10 to Solaris 11. Maybe it is not SUN_SSH per se but there must be something different that is causing the network device to block it. All SSH connection command is just "ssh <server>". No options were given.

That's why I am trying to get some help from Oracle/SUN. Your assistance is much appreciated.

ODM Action Plan **Oracle Support**- May 4, 2012 4:28 PM (1+ month ago)
Who: service request owner Wolfgang ley
What: review additional tcpdumps from the working connection to see if this gives any hont on why the firewall/NAT may block the other conenction but passes this one
When: by May, 8th

Notes **Oracle Support**- May 4, 2012 4:27 PM (1+ month ago)
sun SSh is not invovled in this connection at all. The TCP connection does not get established and therefore the server application is involved here at all.

It maybe a difference on the used OpenSSH client versions (e.g. setting different TCP options) but this is not supported by Oracle and you would have to contact whoever supports your OpenSSH client. The Sun SSH server side is _NOT_ involved here as this server is not participating in any of the blocked traffic.

I will take a look at the two new tcpdumps to see if we can still help you with your firewall issue (but as said: this has nothing to do with Sun SSH server here)

Best regards,
Wolfgang Ley

Update from Customer **CHANG_AN@YAHOO.COM**- May 4, 2012 4:23 PM (1+ month ago)
Indeed.

We can not tell what the philosophy behind the design of the network device. However, we should know why SUN SSHD is sending different packets back to the clients.

Do we know what differences are between SUN_SSH and OpenSSH?

Both un-filtered tcpdump files from client and server have been uploaded.

Thanks.

ODM Action Plan **Oracle Support**- May 4, 2012 4:20 PM (1+ month ago)
Who: customer Chang-An Hsiao
What: provid eunflitered tcpdump (client and server) from the working connection as reference. Maybe we can see something here (if not then there isn't much more we can do as we do not have influence on the firewall/nat device here)
When: by May, 15th

Update from Customer **CHANG_AN@YAHOO.COM**- May 4, 2012 4:19 PM (1+ month ago)
Upload to gtcr successful for the file tcpdump_server.2012050401.

Update from Customer **CHANG_AN@YAHOO.COM**- May 4, 2012 4:19 PM (1+ month ago)
Upload to gtcr successful for the file tcpdump_client.2012050401.

Notes **Oracle Support**- May 4, 2012 4:18 PM (1+ month ago)
The truss is not needed right now. The tcpdump is sufficient for a start.

Even if we find the difference in the TCP packets:"Solaris has no influence on the blocking firewall/NAT device and you will have to investigate that device to see why that device is blocking the packets. We do not know the logic used on that device, sorry.

Update from Customer **CHANG_AN@YAHOO.COM**- May 4, 2012 4:10 PM (1+ month ago)
You are right.
The traffic is indeed blocked by the firewall/NAT network device due to invalid packet it received from the Solaris 11 server. That's why I want to find out why the packets are different between successful and failed connections. We know SSHD is the same. The only difference is the client.

In order to do a fair comparison, I will generate the un-filtered tcpdump files for the successful connection from both server and client.

Do you still need truss file? Or, just tcpdumps? I will upload them shortly.

Thanks!

ODM Action Plan **Oracle Support**- May 4, 2012 12:42 PM (1+ month ago)
Who: customer Chang-An Hsiao
What: check the system doing the TCP reqrite to determine why the packets the server is (re)sending are not reaching the client. Also provide an unfiltered binary tcpdump from the working connection as reference. See previous note for analysis details.
When: by May, 16th

Notes **Oracle Support**- May 4, 2012 12:41 PM (1+ month ago)
Checking the client snoop we can see the new connection starting with packet #30 (SYN).
The connection was opened on May, 3rd 19:30:45 UTC
The TCP connection is from 10.0.32.22 (port 56418) to 71.189.165.164 (port 22222).
The TCP connections gets correctly established (SYN/ACK in #32 and ACK in #33).
There is no further tarffic until the RST packet in #120 (at 19:31:18 UTC)

Checking the server snoop we can see the incoming connection in packet #9 (SYN).
Note that the used IP addresses have changed here -- so this is not the direct
connection from the client but it was modified by some firewall/NAT/... system.
The server send back the SYN/ACK (packet #10) and gets the final ACK in #11.
This last ACK however is ignored and the server resends its SYN/ACK (retries
in packets 12, 13, 14, 15, 16, 19) but these resends are not getting answered
by the remote side. Note that a single incling ACK may indeed get lost and a
resend of a SYN/ACK has to be ACKed again by the remote side --- which however
is not done here.

The system between the client and server (doing the tcp rewrting) is not responding to the retransmitted SYN/ACK packets and this causes the connection abort. We can argue why we need a resent in the first place but such resends are part of the TCP standard.

In your working example: Are you really using the same client IP using the same rewriting system (firewall/nat) between the client and the server? Or is that connection using some other IP and/or gateway?
It would be interesting to see the working connection to check whether we can see the (ignored) retransmits there, too. Can you please provide an unfiltered tcpdump from the working connection as reference? It may also be worth to note that the client is using some TCP options and I would like to compare them with the working connection here, too.

The analysis here clearly shows that the issue is outside of the Sun system and with the system doing the TCP rewrite. The retransmits are not reaching the end client system and there is nothing the server side can do here (as we are sending these packets out).

ODM Action Plan **Oracle Support**- May 3, 2012 7:39 PM (1+ month ago)
Who: service request owner Wolfgang Ley
What downloa dand review new data (two binary unfiltered tcpdump and the explroer).
When: by May, 7th

Update from Customer **CHANG_AN@YAHOO.COM**- May 3, 2012 7:41 PM (1+ month ago)
Upload to gtcr successful for the file explorer.000403a9.cahtoh02_2012.05.03.16.51.tar.gz.

Update from Customer **CHANG_AN@YAHOO.COM**- May 3, 2012 7:36 PM (1+ month ago)
All right.

I have uploaded the explorer gz file and 2 no filtering tcpdump files.

Thanks!

Update from Customer **CHANG_AN@YAHOO.COM**- May 3, 2012 7:35 PM (1+ month ago)
Upload to gtcr successful for the file tcpdump_server.out.no_filtering.

Update from Customer **CHANG_AN@YAHOO.COM**- May 3, 2012 7:34 PM (1+ month ago)
Upload to gtcr successful for the file tcpdump_client.out.no_filtering.

ODM Action Plan **Oracle Support**- May 3, 2012 7:28 PM (1+ month ago)
Who: customer Chang-An Hsiao
What: provide the missing explorer data of the Solaris 11 host (see previous notes)
When:"by May, 15th

Notes **Oracle Support**- May 3, 2012 7:26 PM (1+ month ago)
Got two binary tcpdump files(which however only show traffic for one connection - so they seem to be filtered already) but still waiting for the explorer.

Update from Customer **CHANG_AN@YAHOO.COM**- May 3, 2012 7:25 PM (1+ month ago)
Just "tcpdump -i <NIC>"?

Update from Customer **CHANG_AN@YAHOO.COM**- May 3, 2012 7:24 PM (1+ month ago)
Upload to gtcr successful for the file tcpdump_client.out.

Update from Customer **CHANG_AN@YAHOO.COM**- May 3, 2012 7:24 PM (1+ month ago)
Upload to gtcr successful for the file tcpdump_server.out.

ODM Action Plan **Oracle Support**- May 3, 2012 7:24 PM (1+ month ago)
Who: cusotmer Chang-An Hsiao
what: provide unfiltered binary tcpdump data from both ends and the missing explorer. Also run a testcase by using IP addresses instead of hostnames
When: by May, 15th

Notes **Oracle Support**- May 3, 2012 7:22 PM (1+ month ago)
Hi,

I do not want any port filtering but a tcpdump without any filtering (to ensure that aall traffic, including arp etc.) is visible.

Best regards,
Wolfgang ley.

Update from Customer **CHANG_AN@YAHOO.COM**- May 3, 2012 7:19 PM (1+ month ago)
Like I mentioned, it is behind firewall that's why you see one real IP in the front and private IP in the back. There's no concern here. I am sure both ssh clients connect to the same Solaris 11 server through firewall.

I can provide the binary tcpdump files but why do you not want port filtering?

ODM Action Plan **Oracle Support**- May 3, 2012 7:14 PM (1+ month ago)
Who: cusotmer Chang-An Hsiao
what: provide unfiltered binary tcpdump data from both ends and the missing explorer. Also run a testcase by using IP addresses instead of hostnames
When: by May, 15th

Notes **Oracle Support**- May 3, 2012 7:13 PM (1+ month ago)

Additional note: the last tcpdump text files do not just have a different port number but also different IP addressed (e.g. not 192.168.x.x as the target but 71.189.165.164).
Maybe the two clients have different hostname resolving here? Can you please check the two clients and use the same Solaris 11 server IP address as the destination (and ensure that this IP address is indeed rachable from the client, e.g. testing with ping)?

ODM Action Plan **Oracle Support**- May 3, 2012 7:02 PM (1+ month ago)
Who: cusotmer Chang-An Hsiao
what: provide unfiltered binary tcpdump data from both ends and the missing explorer
When: by May, 15th

Notes **Oracle Support**- May 3, 2012 7:00 PM (1+ month ago)
Sorry but the text output files czannot be used with any analysis program. Can you please upload unfiltered binary data (i.e. not restricted to a single port and binary output file)? Thanks.

I am also still waiting for the explroer data.

Update from Customer **CHANG_AN@YAHOO.COM**- May 3, 2012 6:54 PM (1+ month ago)
Upload to gtcr successful for the file tcpdump_client.txt.

Update from Customer **CHANG_AN@YAHOO.COM**- May 3, 2012 6:54 PM (1+ month ago)
Upload to gtcr successful for the file tcpdump_server.txt.

Update from Customer **CHANG_AN@YAHOO.COM**- May 3, 2012 6:54 PM (1+ month ago)
Upload to gtcr successful for the file sshd.truss.failed_22222.

Update from Customer **CHANG_AN@YAHOO.COM**- May 3, 2012 6:53 PM (1+ month ago)
I am uploading 3 files for your review.

This time, I changed the port to 22222.

ODM Action Plan **Oracle Support**- May 3, 2012 6:32 PM (1+ month ago)
Who: customer Chang-An Hsiao
What: provide tcpdump output from the client and the server at the same itme during the failing connection. Provide explroer of the Solaris 11 syste,
When: by May, 15th

Notes **Oracle Support**- May 3, 2012 6:31 PM (1+ month ago)
Collecting tcpdump data on just one end (either client or server) will not help here, sorry.
Please ensure that you collect the network data at the same time on both endpoints. Thanks.

ODM Action Plan **Oracle Support**- May 3, 2012 6:24 PM (1+ month ago)
Who: customer Chang-An Hsiao
What: continue debugging on the TCP network layer. Provide snoop/.tcpdump data from both endpoints as well as the explorer of the Solaris 11 if additional help is requested.
When: by May, 15th

Notes **Oracle Support**- May 3, 2012 6:22 PM (1+ month ago)
Hi,

this additional information confirms that this has nothing to do with the application and the used ssh version but is a problem on the TCP layer.
Wou will have to analyze the network to check why the incoming SYN is not being answered by an ACK. The good example shows that this is not due to a settning on the Sun Solaris system.
Please collect snoop/tcpdump data from both ends and let me know if you need help analyzing that data. If such additional network debugging help is requested then please provide the snoop/tcpdump data from both endpoints as well as the explorer from the Solaris 11 system.

Best regards,
Wolfgang Ley.

Update from Customer **CHANG_AN@YAHOO.COM**- May 3, 2012 6:19 PM (1+ month ago)
When attempting from OpenSSH 4.3p2, solaris 11 seemed to get the request:

192.168.1.225.2222 144.15.223.8.14161 0 0 128872 0 SYN_RCVD

After certain time (75 seconds?), it dropped.

Update from Customer **CHANG_AN@YAHOO.COM**- May 3, 2012 6:17 PM (1+ month ago)
the tcpsump output files were collected on solaris 11 server.

I am collecting tcpdump from the client now.

Update from Customer **CHANG_AN@YAHOO.COM**- May 3, 2012 6:10 PM (1+ month ago)
Upload to gtcr successful for the file tcpdump.out.failed.

Update from Customer **CHANG_AN@YAHOO.COM**- May 3, 2012 6:09 PM (1+ month ago)
Upload to gtcr successful for the file tcpdump.out.

Update from Customer **CHANG_AN@YAHOO.COM**- May 3, 2012 6:07 PM (1+ month ago)
I also have tcpdump files if you need.
I can also create snoop information if needed.

Update from Customer **CHANG_AN@YAHOO.COM**- May 3, 2012 6:06 PM (1+ month ago)
It is behind a firewall and that's why it has the private IP.

Please note that both SSH connection attempts were from the same 144.15.223.8. The one with OpenSSH 3.6.1p2 worked but OpenSSH4.3p2 failed. There's no routing issue. Both Linux servers are in the same subnet and the 144.15.223.8 is the outgoing router interface IP.

ODM Action Plan **Oracle Support**- May 3, 2012 5:52 PM (1+ month ago)
Who: customer Chang-An Hsiao
What: confirm that the truss was running at the time of the connection failure. if not, then redo the test. If yes, then continue analysis on the underlying network layer as the problem is there and not with ssh. ---- also run a test with ping from the failing client
When: by May, 15th

Notes **Oracle Support**- May 3, 2012 5:51 PM (1+ month ago)
My previous note already explains that the failing client connection is due to a problem on the TCP layer (client cannot even pen a TCP connection with the server).
The used IP addresses of the server are on a private subnet which is usually not routed. So this maybe a routing problem on the failing client side.
Can you please try (from the failing client) to check whether the server is reachable by ping, i.e.
ping 192.168.1.225

The truss from the working connection shows that the connection from client 144.15.223.8 was successful but the the connection attempt from the failing client never reached the server.

ODM Action Plan **Oracle Support**- May 3, 2012 5:43 PM (1+ month ago)
Who: customer Chang-An Hsiao
What: confirm that the truss was running at the time of the connection failure. if not, then redo the test. If yes, then continue analysis on the underlying network layer as the problem is there and not with ssh.
When: by May, 15th

Notes **Oracle Support**- May 3, 2012 5:42 PM (1+ month ago)
Hi,

checking the truss data from the failing connection we can see the following information:

```
[....]
13944/1: 0.0283 bind(3, 0x080D01A8, 32, SOV_SOCKBSD) = 0
13944/1: AF_INET6 name = :: port = 2222
13944/1: scope id = 0 source id = 0x0
13944/1: flow class = 0x00 flow label = 0x00000
13944/1: 0.0289 write(2, 0x08046DC0, 33) = 33
13944/1: S e r v e r l i s t e n i n g o n : : p o r t 2 2 2 2
13944/1: .
13944/1: 0.0290 write(2, 0xFEF646DC, 2) = 2
13944/1: 0xFEF646DC: "\r\n"
13944/1: 0.0291 listen(3, 5, SOV_DEFAULT) = 0
13944/1: 0.0291 sigaction(0x00000001, 0x08047130, 0x080471B0) = 0
13944/1: new: hand = 0xFEEB7479 mask = 0xFFBFFEFF 0xFFFFFFFF 0x000000FF 0 flags = 0x0002
13944/1: old: hand = 0x00000000 mask = 0 0 0 0 flags = 0x0000
13944/1: 0.0292 sigaction(0x0000000F, 0x08047130, 0x080471B0) = 0
13944/1: new: hand = 0xFEEB7479 mask = 0xFFBFFEFF 0xFFFFFFFF 0x000000FF 0 flags = 0x0002
13944/1: old: hand = 0x00000000 mask = 0 0 0 0 flags = 0x0000
13944/1: 0.0293 sigaction(0x00000003, 0x08047130, 0x080471B0) = 0
13944/1: new: hand = 0xFEEB7479 mask = 0xFFBFFEFF 0xFFFFFFFF 0x000000FF 0 flags = 0x0002
13944/1: old: hand = 0x00000000 mask = 0 0 0 0 flags = 0x0000
13944/1: 0.0293 sigaction(0x00000012, 0x08047130, 0x080471B0) = 0
13944/1: new: hand = 0xFEEB7479 mask = 0xFFBFFEFF 0xFFFFFFFF 0x000000FF 0 flags = 0x20002
13944/1: old: hand = 0x00000000 mask = 0 0 0 0 flags = 0x20000
13944/1: pollsys(0x08047130, 1, 0x00000000, 0x00000000) (sleeping...)
13944/1: fd=3 ev=POLLRDNORM rev=0xFFFFFEF6
13944/1: 75.6624 Received signal #2, SIGINT, in pollsys() [default]
13944/1: 75.6626 pollsys(0x08047130, 1, 0x00000000, 0x00000000) Err#4 EINTR
13944/1: fd=3 ev=POLLRDNORM rev=0xFFFFFEF6
[...]
```

So the server started to list on port 2222 but never got any incoming request from the client. This has nothing to do with the sshd or the usefd ssh version on the server but the underlying TCP connection is not reaching the server at all.
The truss was then stopped 75 seconds afterwards.

So either the truss was not running during the time the client did the failing connection (in which case the testrun needs to be redone) or the problem is outside of the Sun server as the incoming connection never reaches our system (e.g. blocked by a firewall between the client and the server).
If the truss was started before the failing client connect (and still running while the client connects) then please redo this test.
If the truss however was really from the time the client could not connect then the problem is outside of the Sun System as the problem is not with ssh application layer but with the failing TCP connection. In this case analysis needs to continue on the network layer to see where this TCP connection is blocked.

So the next possible actions are either:
1) Redo the test to ensure that truss of the server was running at the same time the failing client tries to connect (this action is only useful if the first set of data collection was not already done this way)
OR
2) Continue with network analysis to check why the incoming TCP conneciton on port 2222 does not reach the server. In this case please collect the snoop output (or tcpdump on non-Solaris hosts) of the client and the server side during the testrun.

Best regards,
Wolfgang Ley.

ODM Action Plan **Oracle Support**- May 3, 2012 5:29 PM (1+ month ago)

Who: service request owner Wolfgang ley
What: analyze provided data (truss from failing and working connectgion as wekll as associated client outputs)
When: by May, 7th

Notes **Oracle Support**- May 3, 2012 5:29 PM (1+ month ago)
Hi,

thanks for the data. I will download and analyze them (and of course provide the results afterwards).

Best regards,
Wolfgang ley

Update from Customer **CHANG_AN@YAHOO.COM**- May 3, 2012 5:26 PM (1+ month ago)
I have uploaded 4 files for your review. They are in pairs:

Successful ssh connection from OpenSSH_3.6.1p2
---

ssh_connection
sshd.truss

Failed ssh connection from OpenSSH_4.3p2
---

ssh_failed_connection
sshd.truss.failed

Update from Customer **CHANG_AN@YAHOO.COM**- May 3, 2012 5:21 PM (1+ month ago)
Upload to gtcr successful for the file sshd.truss.

Update from Customer **CHANG_AN@YAHOO.COM**- May 3, 2012 5:20 PM (1+ month ago)
Upload to gtcr successful for the file sshd.truss.failed.

Update from Customer **CHANG_AN@YAHOO.COM**- May 3, 2012 5:19 PM (1+ month ago)
Upload to gtcr successful for the file ssh_failed_connection.

Update from Customer **CHANG_AN@YAHOO.COM**- May 3, 2012 5:19 PM (1+ month ago)
Upload to gtcr successful for the file ssh_connection.

ODM Action Plan **Oracle Support**- May 3, 2012 11:16 AM (1+ month ago)
Who: customer Chang-An Hsiao
What: provide explorer and truss output from the server and ssh client (either output or truss). See previous note for details.
When: by May, 15th

Notes **Oracle Support**- May 3, 2012 11:11 AM (1+ month ago)
Hi,

checking the service request information I can see that you are reporting that some outgoing ssh packets "are being blocked by network device". Where have you seen such blocked packets? Is this on the Sun system itself or is there some firewall (or other external network device) in volved here?
Can you please provide an explorer output from the Solaris 11 host (see document 1312847.1 for more information about the Explro3er data collector tool)?

I would also need a truss of a debug sshd to check what is going on here. On the Solaris 11 host please start this additional ssh server on port 2222:
truss -aefdl -rall -wall -vall -xall -o /var/tmp/sshd.truss /usr/lib/ssh/sshd -ddd -p 2222
Now use the failing OpenSSH_4.3p2 client to connect to this debug server:
a) If the client is not running Solaris: % ssh -vvv -p 2222 <servername>
b) If the client is running Solaris: % truss -aefdl -rall -wall -vall -xall -o /var/tmp/ssh.truss ssh -vvv -p 2222 <servername>
Note that <servername> has to be replaced by the hostname of the failing Solaris 11 server.

The truss should automatically exit after the connection attempt. Please sent the file /var/tmp/sshd.truss from the Solaris 11 system and the client output (either the screen output if not Solaris or the truss output if Solaris).
I will then check the sshd truss data and the explorer to determine the root cause of this issue.

Best regards,
Wolfgang Ley.

ODM Issue Clarification **Oracle Support**- May 3, 2012 11:01 AM (1+ month ago)
Som SSH clients cannot connect to a Solaris 11 SSH server while other clients work ok.
All clients can connect to Solaris 10 ssh servers without problems.

Notes **Oracle Support**- May 3, 2012 11:00 AM (1+ month ago)
Hi,

your service request has been forwarded to the network support group (for ssh support) and I am the new owner of this issue.
I will review the information available in this service request and provide an update afterwards (later today).

My business hours are 09:00 - 17:00 UTC. If you need help outside these hours then please contact the supportline to get the next available engineer.

Best regards,
Wolfgang Ley

ODM Action Plan **Oracle Support**- May 2, 2012 10:07 PM (1+ month ago)

Who: Next Available Engineer
What: Take ownership of the SR and contact the customer
When: Within the SLA

Customer Problem Description **CHANG_AN@YAHOO.COM**- May 2, 2012 9:40 PM (1+ month ago)

1) ### Impact on Business ###
SSH client can not connect to server

2) What is the OS version and the kernel patch level of the system?
5.11 11.0

3) What is the Firmware level of the system?
N/A