

Tvheadend - Bug #1407

Odd space(s) in channel name

2012-11-16 22:33 - Rene Herbrich

Status:	Fixed	Start date:	2012-11-16
Priority:	Normal	Due date:	
Assignee:		% Done:	100%
Category:	DVB	Estimated time:	0.00 hour
Target version:	3.2	Affected Versions:	3.2, 3.3
Found in version:	3.3.153~ged5d7e4		
Description			
Some channels do have some odd space(s) in its name. If I scan for channels e.g. on my tv, there are none. So I think that's no problem of my cable provider? I'm experiencing the problem since the first time I compiled tvheadend (6 months ago), so it's nothing new.			

Associated revisions

Revision 8a48fe82 - 2012-11-29 09:57 - Adam Sutton

Issue #1407 - skip unused ISO 8859 control codes.

This had the effect of placing unwanted spaces into various channel names where the broadcaster was sending these control chars.

Thanks to Rene Herbrich for diagnosing this.

Revision 44628138 - 2012-11-29 10:51 - Adam Sutton

Issue #1407 - forgot to strip unused escapes from ISO 6937.

Revision e444d80a - 2012-12-19 20:31 - Adam Sutton

Fix #1407 - skip unused ISO 8859 control codes.

This had the effect of placing unwanted spaces into various channel names where the broadcaster was sending these control chars.

Thanks to Rene Herbrich for diagnosing this.
(cherry picked from commit 8a48fe8214df98fc7cd7f23fb210971921d6282c)

History

#1 - 2012-11-20 18:32 - Adam Sutton

- Category changed from 11 to DVB
- Status changed from New to Need feedback

The only thing I'm aware that TVH does is to clean spaces from either end of the names, since some providers do indeed stick spaces at either end. It certainly should not insert any additional chars.

The only thing I can think is to try putting some printf()'s around the point the data is read from the tables. Look in src/dvb/dvb_tables.c in dvb_sdt_callback() and follow the code into dvb_desc_service() as required.

Adam

#2 - 2012-11-20 19:07 - Rene Herbrich

Will take some time, my HDHomeRun died shortly after I submitted this ticket. :(
Have to wait for the replacement...

#3 - 2012-11-26 15:57 - Adam Sutton

This has been reported by at least one other user. So we need to understand where these spaces are coming from.

#4 - 2012-11-26 16:02 - Christian Brauers

Yeah, think that was me ...

Having the same trouble. German DVB-C with Sundtek Stick!
On my TVs and EyeTV on iMac everything is fine!

#5 - 2012-11-26 16:06 - Adam Sutton

It would be good to know which services these are etc...?

But since the original report is HDHR I'm guessing these might both be cable providers? So basically it wouldn't surprise me at all if they are sending garbage on purpose!

Adam

#6 - 2012-11-26 22:19 - Rene Herbrich

I'll get my replacement hdhomerun this week, I hope...

If it helps: my cable provider is kabelbw.

But it would be strange, if it's the fault of my provider? I mean, why is there no problem with my TV?

#7 - 2012-11-27 02:10 - Bernd Schneider

Have the same problem on my PC-System.
I'm testing/using Openelec with my old Skystar2 to receive Astra 19.2.
I installed openelec many times, older revisions too.
It seems that always the same channels have this wrong additional space(s) in the channelname. The channels are on different transponders and I think I can exclude that this causes about a weak receiving signal.

#8 - 2012-11-27 21:16 - B. J.

Yes, same here. But this affects only service-names from the german provider "Sky". As a workaround I've edit the names of the affected channels after mapping services to these. I don't know if the odd spaces are really spaces (20H) or possibly some kind of whitespaces or other control chars (tab, null etc).

#9 - 2012-11-27 22:31 - Rene Herbrich

For me not only on channels broadcasted by "Sky" are affected. Additionally some other pay and free-tv channels have the spaces in its name. But what's striking: most of the channels with this problem belong to "Sky".

I'll try to insert some printf() asap and will report back then. My new device will arrive tomorrow, at least if the current tracking status is reliable.

#10 - 2012-11-28 23:08 - Rene Herbrich

- File TXT.rtf added

Here's a part of my debug output (the complete output is attached):

```

Nov 28 22:38:02 Channel Name Test: -----
Nov 28 22:38:02 Channel Name Test: ptr content0: 00000015 ( )
Nov 28 22:38:02 Channel Name Test: ptr content1: 0000000B ( )
Nov 28 22:38:02 Channel Name Test: ptr content2: 00000086 ( )
Nov 28 22:38:02 Channel Name Test: ptr content3: 00000041 (A)
Nov 28 22:38:02 Channel Name Test: ptr content4: 0000006C (l)
Nov 28 22:38:02 Channel Name Test: ptr content5: 00000062 (b)
Nov 28 22:38:02 Channel Name Test: ptr content6: 00000061 (a)
Nov 28 22:38:02 Channel Name Test: ptr content7: 0000006E (n)
Nov 28 22:38:02 Channel Name Test: ptr content8: 00000069 (i)
Nov 28 22:38:02 Channel Name Test: ptr content9: 00000061 (a)
Nov 28 22:38:02 Channel Name Test: ptr content10: 00000020 ( )
Nov 28 22:38:02 Channel Name Test: ptr content11: 00000073 (s)
Nov 28 22:38:02 Channel Name Test: ptr content12: 00000063 (c)
Nov 28 22:38:02 Channel Name Test: ptr content13: 00000072 (r)
Nov 28 22:38:02 Channel Name Test: ptr content14: 00000087 ( )
Nov 28 22:38:02 Channel Name Test: ptr content15: 00000065 (e)
Nov 28 22:38:02 Channel Name Test: name: Albania scr een RTV
[...]
Nov 28 22:38:02 Channel Name Test: -----
Nov 28 22:38:02 Channel Name Test: ptr content0: 0000000B ( )
Nov 28 22:38:02 Channel Name Test: ptr content1: 0000000B ( )
Nov 28 22:38:02 Channel Name Test: ptr content2: 00000086 ( )
Nov 28 22:38:02 Channel Name Test: ptr content3: 00000043 (C)

```

```

Nov 28 22:38:02 Channel Name Test: ptr content4: 0000004C (L)
Nov 28 22:38:02 Channel Name Test: ptr content5: 00000041 (A)
Nov 28 22:38:02 Channel Name Test: ptr content6: 00000053 (S)
Nov 28 22:38:02 Channel Name Test: ptr content7: 00000053 (S)
Nov 28 22:38:02 Channel Name Test: ptr content8: 00000049 (I)
Nov 28 22:38:02 Channel Name Test: ptr content9: 00000043 (C)
Nov 28 22:38:02 Channel Name Test: ptr content10: 00000041 (A)
Nov 28 22:38:02 Channel Name Test: ptr content11: 00000087 ( )
Nov 28 22:38:02 Channel Name Test: ptr content12: 0000005F ( )
Nov 28 22:38:02 Channel Name Test: ptr content13: 00000004 ( )
Nov 28 22:38:02 Channel Name Test: ptr content14: 00000000 ( )
Nov 28 22:38:02 Channel Name Test: ptr content15: 00000000 ( )
Nov 28 22:38:02 Channel Name Test: name: CLASSICA
[...]
Nov 28 22:39:19 Channel Name Test: -----
Nov 28 22:39:19 Channel Name Test: ptr content0: 00000014 ( )
Nov 28 22:39:19 Channel Name Test: ptr content1: 0000000B ( )
Nov 28 22:39:19 Channel Name Test: ptr content2: 00000086 ( )
Nov 28 22:39:19 Channel Name Test: ptr content3: 00000045 (E)
Nov 28 22:39:19 Channel Name Test: ptr content4: 00000053 (S)
Nov 28 22:39:19 Channel Name Test: ptr content5: 00000050 (P)
Nov 28 22:39:19 Channel Name Test: ptr content6: 0000004E (N)
Nov 28 22:39:19 Channel Name Test: ptr content7: 00000087 ( )
Nov 28 22:39:19 Channel Name Test: ptr content8: 00000020 ( )
Nov 28 22:39:19 Channel Name Test: ptr content9: 00000041 (A)
Nov 28 22:39:19 Channel Name Test: ptr content10: 0000006D (m)
Nov 28 22:39:19 Channel Name Test: ptr content11: 00000065 (e)
Nov 28 22:39:19 Channel Name Test: ptr content12: 00000072 (r)
Nov 28 22:39:19 Channel Name Test: ptr content13: 00000069 (i)
Nov 28 22:39:19 Channel Name Test: ptr content14: 00000063 (c)
Nov 28 22:39:19 Channel Name Test: ptr content15: 00000061 (a)
Nov 28 22:39:19 Channel Name Test: name: ESPN America GB
[...]
Nov 28 22:39:19 Channel Name Test: -----
Nov 28 22:39:19 Channel Name Test: ptr content0: 00000013 ( )
Nov 28 22:39:19 Channel Name Test: ptr content1: 0000000B ( )
Nov 28 22:39:19 Channel Name Test: ptr content2: 00000086 ( )
Nov 28 22:39:19 Channel Name Test: ptr content3: 00000056 (V)
Nov 28 22:39:19 Channel Name Test: ptr content4: 00000048 (H)
Nov 28 22:39:19 Channel Name Test: ptr content5: 00000031 (l)
Nov 28 22:39:19 Channel Name Test: ptr content6: 00000020 ( )
Nov 28 22:39:19 Channel Name Test: ptr content7: 00000043 (C)
Nov 28 22:39:19 Channel Name Test: ptr content8: 0000006C (l)
Nov 28 22:39:19 Channel Name Test: ptr content9: 00000061 (a)
Nov 28 22:39:19 Channel Name Test: ptr content10: 00000073 (s)
Nov 28 22:39:19 Channel Name Test: ptr content11: 00000087 ( )
Nov 28 22:39:19 Channel Name Test: ptr content12: 00000073 (s)
Nov 28 22:39:19 Channel Name Test: ptr content13: 00000069 (i)
Nov 28 22:39:19 Channel Name Test: ptr content14: 00000063 (c)
Nov 28 22:39:19 Channel Name Test: ptr content15: 00000086 ( )
Nov 28 22:39:19 Channel Name Test: name: VH1 Clas sic GB

```

That's what I put into the function "dwb_desc_service" within "dwb_tables.c":

```

tvhlog(LOG_NOTICE, "Channel Name Test", "-----");
tvhlog(LOG_NOTICE, "Channel Name Test", "ptr content0: %08X (%c)", ptr[0], ptr[0]);
tvhlog(LOG_NOTICE, "Channel Name Test", "ptr content1: %08X (%c)", ptr[1], ptr[1]);
tvhlog(LOG_NOTICE, "Channel Name Test", "ptr content2: %08X (%c)", ptr[2], ptr[2]);
tvhlog(LOG_NOTICE, "Channel Name Test", "ptr content3: %08X (%c)", ptr[3], ptr[3]);
tvhlog(LOG_NOTICE, "Channel Name Test", "ptr content4: %08X (%c)", ptr[4], ptr[4]);
tvhlog(LOG_NOTICE, "Channel Name Test", "ptr content5: %08X (%c)", ptr[5], ptr[5]);
tvhlog(LOG_NOTICE, "Channel Name Test", "ptr content6: %08X (%c)", ptr[6], ptr[6]);
tvhlog(LOG_NOTICE, "Channel Name Test", "ptr content7: %08X (%c)", ptr[7], ptr[7]);
tvhlog(LOG_NOTICE, "Channel Name Test", "ptr content8: %08X (%c)", ptr[8], ptr[8]);
tvhlog(LOG_NOTICE, "Channel Name Test", "ptr content9: %08X (%c)", ptr[9], ptr[9]);
tvhlog(LOG_NOTICE, "Channel Name Test", "ptr content10: %08X (%c)", ptr[10], ptr[10]);
tvhlog(LOG_NOTICE, "Channel Name Test", "ptr content11: %08X (%c)", ptr[11], ptr[11]);
tvhlog(LOG_NOTICE, "Channel Name Test", "ptr content12: %08X (%c)", ptr[12], ptr[12]);
tvhlog(LOG_NOTICE, "Channel Name Test", "ptr content13: %08X (%c)", ptr[13], ptr[13]);
tvhlog(LOG_NOTICE, "Channel Name Test", "ptr content14: %08X (%c)", ptr[14], ptr[14]);
tvhlog(LOG_NOTICE, "Channel Name Test", "ptr content15: %08X (%c)", ptr[15], ptr[15]);
if((r = dwb_get_string_with_len(name, namelen, ptr, len, NULL, NULL)) < 0)
    return -1;

tvhlog(LOG_NOTICE, "Channel Name Test", "name: %s", name);

```

To make a long post short: The names contain the chars 0x86 at the beginning and especially 0x87.
At first I wanted to suggest to simply ignore them, but then I looked into the ASCII table.
These are legit chars used in foreign languages.

So, are you able to retrieve the country where tvheadend is running by the epg data?
If yes, you could leave them out. I know, not a nice solution as a channel name might contain such a char on purpose.

#11 - 2012-11-28 23:57 - Rene Herbrich

Well, forget the last part of my conclusion.

I've looked into the dvb specs: http://www.etsi.org/deliver/etsi_en/300400_300499/300468/01.11.01_60/en_300468v011101p.pdf
None of the code tables uses chars with a value of the 0x80 or 0x90 range (from page 104 to 114). So we could leave them out.

But what's really funny, I think: some of the channels propose to use a chinese code table [biggrin.png](#)

#12 - 2012-11-29 10:56 - Adam Sutton

- *Status changed from Need feedback to Accepted*

- *Target version set to 3.2*

- *Affected Versions 3.2, 3.3 added*

Rene,

I completely agree, I think those control codes should be completely skipped over. As you say the ETSI spec doc (and also the wiki entries for those code spaces) do not define any usage of those control chars.

Why they're even being used I don't know, I guess they must have some special meaning somewhere (or to someone). But I think for our purposes they should simply be removed from the strings.

I'll post a fix. Need to decide whether this warrants a backport to 3.2.

Adam

#13 - 2012-11-29 11:01 - Adam Sutton

- *Status changed from Accepted to Resolved*

#14 - 2012-11-29 11:43 - Rene Herbrich

Why only for iso8859? According to what I found out yesterday, none of the code tables use these chars, not even utf8.

#15 - 2012-11-29 11:44 - Adam Sutton

Because I'm an idiot and forgot to check the others [wink.png](#)

#16 - 2012-11-29 11:50 - Adam Sutton

I'm not sure what needs to be done for the UTF8 stuff, since we're not actually converting that. So I could try and strip the unwanted escape sequences, but to be honest most end clients would probably do that anyway.

The mistake here is simply that we replaced unused escape sequences (during translation) with spaces rather than completely ignoring them as we should.

I'll update the 6937 routine.

Adam

#17 - 2012-11-29 12:34 - Rene Herbrich

Mh, yesterday I found a utf8 code table telling me these chars are not valid. But they are? I don't know...

Well, you could leave them out and just wait if someone complains about missing chars?

#18 - 2012-12-19 21:36 - Adam Sutton

- *Status changed from Resolved to Fixed*

- *% Done changed from 0 to 100*

Applied in changeset [e444d80a23e3d356004c45516e2df8a8df593c37](#).

Files

HTS Tvheadend 3.3.153~ged5d7e4.jpg	196 KB	2012-11-16	Rene Herbrich
TXT.rtf	754 KB	2012-11-28	Rene Herbrich