Chasing a 25 year old Thunderbird security bug

Garming Sam, Douglas Bagnall

```
Patch to pass a Windows RODC test - Mozilla Thunderbird
File Edit View Go Message Enigmail Tools Help
                   Write Chat Address Book
Tag ▼
                                                                          Reply Reply All
                                                                                                   ⇒ Forward
                                                                                                                More ▼
 From Me <garming@catalyst.net.nz>
Subject Patch to pass a Windows RODC test
                                                                                                          20/09/16 17:33
    To Catalyst Samba Developers <samba-dev@catalyst.net.nz>😭
Things to do:
- Automate credentials pre-loading in RODC environments (add
Administrator to RODC replication group + RODC preload run)
- Give a proper server for the referrals (PDC or replication partner
instead of hardcoding)
- Add the same check in the deleted case
-0001-Let-me-cheat-here.patch-
From eeba21f2bf1fb4a28d498ba032edcb8625eb68c7 Mon Sep 17 00:00:00 2001
From: Ubuntu <ubuntu@gar-mas-0-2.samdom.example.com>
Date: Tue, 20 Sep 2016 04:25:34 +0000
Subject: [PATCH] Let me cheat here...
Signed-off-by: Ubuntu <ubuntu@qar-mas-0-2.samdom.example.com>
 source4/dsdb/samdb/ldb modules/samldb.c
                                               10 +++++++
 2 files changed, 35 insertions(+)
diff --qit a/source4/dsdb/samdb/ldb modules/repl meta data.c b/source4/dsdb/samdb/ldb modules/repl meta data.c
index 7a5906e..d870479 100644
--- a/source4/dsdb/samdb/ldb modules/repl meta data.c
+++ b/source4/dsdb/samdb/ldb modules/repl meta data.c
@@ -976,6 +976,30 @@ static int replmd add(struct ldb module *module, struct ldb request *req)
               quid = GUID random();
        if (!ldb request get control(req, DSDB CONTROL REPLICATED UPDATE OID)
                       && !ldb request get control(reg, DSDB CONTROL DBCHECK MODIFY RO REPLICA)) {
               bool rodc = false;
               struct loadparm context *lp ctx;
               char *referral:
               ret = samdb rodc(ldb, &rodc);
               if (ret != LDB SUCCESS) {
                       DEBUG(4, ( location ": unable to tell if we are an RODC\n"));
               } else if (rodc) {
                       ldb set errstring(ldb, "RODC modify is forbidden!");
                       lp ctx = talloc get type(ldb get opaque(ldb, "loadparm"),
                                      struct loadparm context);
                       referral = talloc asprintf(req,
                                      "ldap://qar-mas-0-0.%s/%s",
                                      lnofa dnodomain(ln ctv)
```

```
To: Catalyst Samba Developers <samba-dev@catalyst.net.nz>
From: Garming Sam <garming@catalyst.net.nz>
Subject: Patch to pass a Windows RODC test
Message-ID: <3ae65b63-914e-27ac-f39a-3e071156f597@catalyst.net.nz>
Date: Tue, 20 Sep 2016 17:33:01 +1200
User-Agent: Mozilla/5.0 (X11; Linux x86 64; rv:45.0) Gecko/20100101
Thunderbird/45.2.0
MIME-Version: 1.0
Content-Type: multipart/mixed;
boundary="-----C2AFA65155D3EB764BFA8504"
This is a multi-part message in MIME format.
-----C2AFA65155D3EB764BFA8504
Content-Type: text/plain: charset=utf-8
Content-Transfer-Encoding: 7bit
Things to do:
- Automate credentials pre-loading in RODC environments (add
Administrator to RODC replication group + RODC preload run)
- Give a proper server for the referrals (PDC or replication partner
instead of hardcoding)
- Add the same check in the deleted case
-----C2AFA65155D3EB764BFA8504
Content-Type: text/x-diff;
name="0001-Let-me-cheat-here.patch"
Content-Transfer-Encoding: 7bit
Content-Disposition: attachment;
filename="0001-Let-me-cheat-here.patch"
From eeba21f2bf1fb4a28d498ba032edcb8625eb68c7 Mon Sep 17 00:00:00 2001
From: Ubuntu <ubuntu@gar-mas-0-2.samdom.example.com>
Date: Tue, 20 Sep 2016 04:25:34 +0000
Subject: [PATCH] Let me cheat here...
Signed-off-by: Ubuntu <ubuntu@gar-mas-0-2.samdom.example.com>
 source4/dsdb/samdb/ldb modules/samldb.c
                                             10 +++++++
 2 files changed, 35 insertions(+)
```

Options?

Strategy 0: Ignore it

Strategy 1: GDB, standard debugger

Strategy 2: Valgrind

Strategy 3: Address Sanitizer

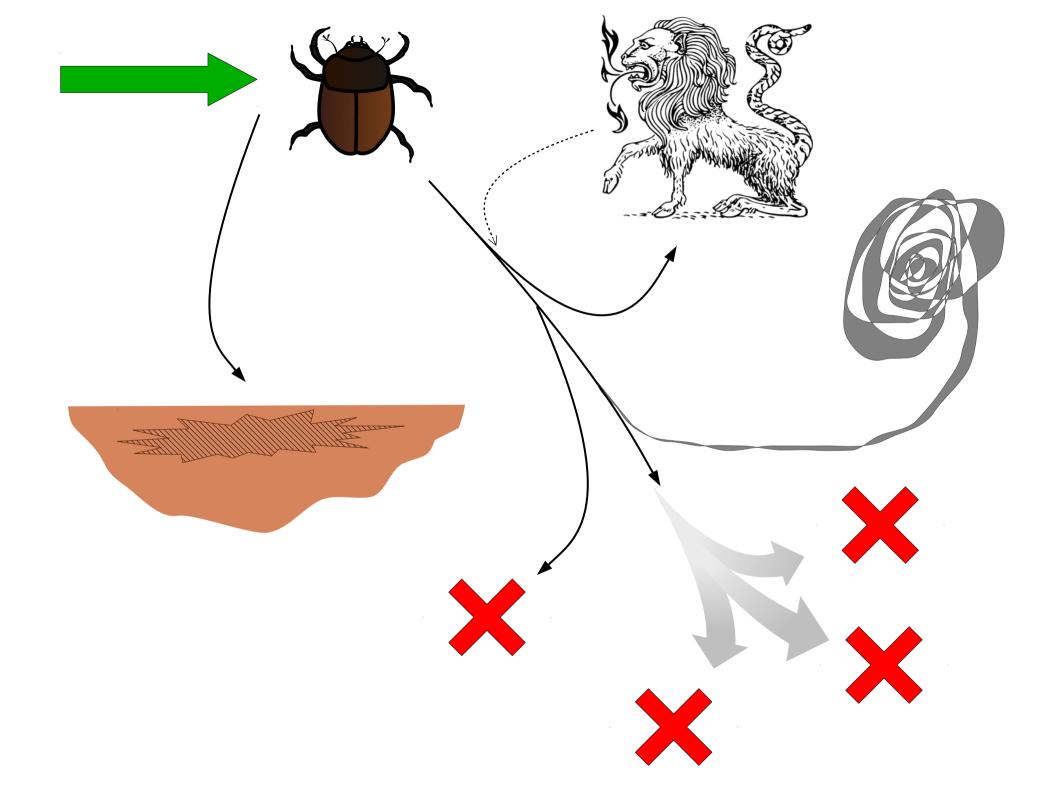
Strategy 4: Bisection of input

Strategy 5: ???

GDB (Standard debugger)

- + Works well for master divinators, super sleuths, and those-who-know-the-entire-project
- Needs symbols to be effective, building from scratch and lots of code navigation
- + Allows useful data extraction of the actual crash

Gives no clues as to what the cause is



Hints so far

Crash signatures:

```
arena_malloc | nsStringBuffer::Alloc | MimeInlineText_parse_eof

0x0 | CanonicalizeXPCOMParticipant | nsCycleCollector::ForgetSkippable

arena_dalloc | nsLDAPURL::Init | js::RunScript

arena_dalloc | GetMsgDBHdrFromURI | nsMailboxUrl::GetFolder
```

Memory corruption likely in email parsing (specifically in the attachment)

Scattered crashes in GC + memory alloc + JIT + XPCOM + parsing

Previous bugs had been found in parsing mime contents

```
==12752== by 0xF234569: mozJSComponentLoader::ObjectForLocation(ComponentLoaderInfo&, nsIFile*,
JS::MutableHandle<JSObject*>, JS::MutableHandle<JSScript*>, char**, bool, JS::MutableHandle<JS::Value>)
(mozJSComponentLoader.cpp:684)
==12752== by 0xF235710: mozJSComponentLoader::LoadModule(mozilla::FileLocation&)
(mozJSComponentLoader.cpp:390)
==12752== by 0xE9D17A8: nsComponentManagerImpl::KnownModule::Load()
(nsComponentManager.cpp:832)
==12752== by 0xE9D182A: nsFactoryEntry::GetFactory() (nsComponentManager.cpp:1865)
==12752== Address 0x1f277858 is 8 bytes after a block of size 32 free'd
==12752== at 0x923318F: g type free instance (in /usr/lib/x86 64-linux-gnu/libgobject-2.0.so.0.4002.0)
==12752==
            by 0x689A561: ??? (in /usr/lib/x86 64-linux-gnu/libgtk-3.so.0.1000.8)
            by 0x9218ACF: g object run dispose (in /usr/lib/x86 64-linux-gnu/libgobject-2.0.so.0.4002.0)
==12752==
==12752==
            by 0x686A0FF: ??? (in /usr/lib/x86 64-linux-gnu/libgtk-3.so.0.1000.8)
==12752==
            by 0x6861F07: ??? (in /usr/lib/x86 64-linux-gnu/libgtk-3.so.0.1000.8)
            by 0x9212331: g closure invoke (in /usr/lib/x86 64-linux-gnu/libgobject-2.0.so.0.4002.0)
==12752==
==12752==
            by 0x92240D3: ??? (in /usr/lib/x86 64-linux-gnu/libgobject-2.0.so.0.4002.0)
==12752==
            by 0x922BA28: g signal emit valist (in /usr/lib/x86 64-linux-gnu/libgobject-2.0.so.0.4002.0)
==12752==
            by 0x922BCE1: g signal emit (in /usr/lib/x86 64-linux-gnu/libgobject-2.0.so.0.4002.0)
            by 0x6A30A57: ??? (in /usr/lib/x86 64-linux-gnu/libgtk-3.so.0.1000.8)
==12752==
==12752==
            by 0x9218ACF: g object run dispose (in /usr/lib/x86 64-linux-gnu/libgobject-2.0.so.0.4002.0)
==12752==
            by 0x68BB73F: ??? (in /usr/lib/x86 64-linux-gnu/libgtk-3.so.0.1000.8)
==12752==
==12752== Invalid read of size 8
==12752== at 0xF284891: XPCWrappedNative::Init(XPCNativeScriptableCreateInfo const*)
(XPCWrappedNative.cpp:761)
==12752== by 0xF292BF5: XPCWrappedNative::GetNewOrUsed(xpcObjectHelper&,
XPCWrappedNativeScope*, XPCNativeInterface*, XPCWrappedNative**) (XPCWrappedNative.cpp:455)
==12752== bv 0xF268F0E: XPCConvert::NativeInterface2JSObject(JS::MutableHandle<JS::Value>.
nsIXPConnectJSObjectHolder**, xpcObjectHelper&, nsID const*, bool, nsresult*) (XPCConvert.cpp:829)
==12752== by 0xF283F07:
XPCWrappedNativeScope::GetComponentsJSObject(JS::MutableHandle<JSObject*>)
(XPCWrappedNativeScope.cpp:199)
==12752== by 0xF285CB0: XPCWrappedNativeScope::AttachComponentsObject(JSContext*)
(XPCWrappedNativeScope.cpp:226)
==12752== by 0xF285EAC: xpc::InitGlobalObject(JSContext*, JS::Handle<JSObject*>, unsigned int)
(nsXPConnect.cpp:466)
==12752== by 0xF2922BE: nsXPConnect::InitClassesWithNewWrappedGlobal(JSContext*, nsISupports*,
nsIPrincipal*, unsigned int, JS::CompartmentOptions&, nsIXPConnectJSObjectHolder**)
```

```
==12752== by 0x92240D3: ??? (in /usr/lib/x86 64-linux-gnu/libgobject-2.0.so.0.4002.0)
==12752==
            by 0x922BA28: g signal emit valist (in /usr/lib/x86 64-linux-gnu/libgobject-2.0.so.0.4002.0)
==12752==
            by 0x922BCE1: g signal emit (in /usr/lib/x86 64-linux-gnu/libgobject-2.0.so.0.4002.0)
            by 0x6A30A57: ??? (in /usr/lib/x86 64-linux-gnu/libgtk-3.so.0.1000.8)
==12752==
==12752==
            by 0x9218ACF: g object run dispose (in /usr/lib/x86 64-linux-gnu/libgobject-2.0.so.0.4002.0)
            by 0x68BB73F: ??? (in /usr/lib/x86 64-linux-gnu/libgtk-3.so.0.1000.8)
==12752==
==12752==
==12752== Invalid read of size 8
==12752== at 0xF2958DB: IsTaggedScope (xpcprivate.h:1756)
==12752==
            by 0xF2958DB: HasProto (xpcprivate.h:1772)
            by 0xF2958DB: XPCWrappedNative::GetProto() const (xpcprivate.h:1776)
==12752==
==12752==
            by 0xF284A40: XPCWrappedNative::Init(XPCNativeScriptableCreateInfo const*)
(XPCWrappedNative.cpp:781)
==12752== by 0xF292BF5: XPCWrappedNative::GetNewOrUsed(xpcObjectHelper&,
XPCWrappedNativeScope*, XPCNativeInterface*, XPCWrappedNative**) (XPCWrappedNative.cpp:455)
==12752== by 0xF268F0E: XPCConvert::NativeInterface2JSObject(JS::MutableHandle<JS::Value>,
nsIXPConnectJSObjectHolder**, xpcObjectHelper&, nsID const*, bool, nsresult*) (XPCConvert.cpp:829)
==12752== by 0xF283F07:
XPCWrappedNativeScope::GetComponentsJSObject(JS::MutableHandle<JSObject*>)
(XPCWrappedNativeScope.cpp:199)
==12752== by 0xF285CB0: XPCWrappedNativeScope::AttachComponentsObject(JSContext*)
(XPCWrappedNativeScope.cpp:226)
==12752== by 0xF285EAC: xpc::InitGlobalObject(JSContext*, JS::Handle<JSObject*>, unsigned int)
(nsXPConnect.cpp:466)
==12752== by 0xF2922BE: nsXPConnect::InitClassesWithNewWrappedGlobal(JSContext*, nsISupports*,
nsIPrincipal*, unsigned int, JS::CompartmentOptions&, nsIXPConnectJSObjectHolder**)
(nsXPConnect.cpp:513)
==12752== by 0xF233530: mozJSComponentLoader::PrepareObjectForLocation(JSContext*, nsIFile*,
nsIURI*, bool, bool*) (mozJSComponentLoader.cpp:581)
==12752== by 0xF234569: mozJSComponentLoader::ObjectForLocation(ComponentLoaderInfo&, nsIFile*,
JS::MutableHandle<JSObject*>, JS::MutableHandle<JSScript*>, char**, bool, JS::MutableHandle<JS::Value>)
(mozJSComponentLoader.cpp:684)
==12752== by 0xF235710: mozJSComponentLoader::LoadModule(mozilla::FileLocation&)
(mozJSComponentLoader.cpp:390)
==12752== by 0xE9D17A8: nsComponentManagerImpl::KnownModule::Load()
(nsComponentManager.cpp:832)
==12752== Address 0x1f277840 is 16 bytes inside a block of size 32 free'd
==12752== at 0x923318F; g type free instance (in /usr/lib/x86_64-linux-gnu/libgobiect-2.0.so.0.4002.0)
```

```
==12752== by 0xF292BF5: XPCWrappedNative::GetNewOrUsed(xpcObjectHelper&,
XPCWrappedNativeScope*, XPCNativeInterface*, XPCWrappedNative**) (XPCWrappedNative.cpp:455)
==12752== by 0xF268F0E: XPCConvert::NativeInterface2JSObject(JS::MutableHandle<JS::Value>.
nsIXPConnectJSObjectHolder**, xpcObjectHelper&, nsID const*, bool, nsresult*) (XPCConvert.cpp:829)
==12752== by 0xF283F07:
XPCWrappedNativeScope::GetComponentsJSObject(JS::MutableHandle<JSObject*>)
(XPCWrappedNativeScope.cpp:199)
==12752== by 0xF285CB0: XPCWrappedNativeScope::AttachComponentsObject(JSContext*)
(XPCWrappedNativeScope.cpp:226)
==12752== by 0xF285EAC: xpc::InitGlobalObject(JSContext*, JS::Handle<JSObject*>, unsigned int)
(nsXPConnect.cpp:466)
==12752== by 0xF2922BE: nsXPConnect::InitClassesWithNewWrappedGlobal(JSContext*, nsISupports*,
nsIPrincipal*, unsigned int, JS::CompartmentOptions&, nsIXPConnectJSObjectHolder**)
(nsXPConnect.cpp:513)
==12752== by 0xF233530: mozJSComponentLoader::PrepareObjectForLocation(JSContext*, nsIFile*,
nsIURI*, bool, bool*) (mozJSComponentLoader.cpp:581)
==12752== by 0xF234569: mozJSComponentLoader::ObjectForLocation(ComponentLoaderInfo&, nsIFile*,
JS::MutableHandle<JSObject*>, JS::MutableHandle<JSScript*>, char**, bool, JS::MutableHandle<JS::Value>)
(mozJSComponentLoader.cpp:684)
==12752== by 0xF235710: mozJSComponentLoader::LoadModule(mozilla::FileLocation&)
(mozJSComponentLoader.cpp:390)
==12752== by 0xE9D17A8: nsComponentManagerImpl::KnownModule::Load()
(nsComponentManager.cpp:832)
==12752== Address 0x1f277850 is 0 bytes after a block of size 32 free'd
==12752== at 0x923318F: g type free instance (in /usr/lib/x86 64-linux-gnu/libgobject-2.0.so.0.4002.0)
==12752== by 0x689A561: ??? (in /usr/lib/x86 64-linux-gnu/libgtk-3.so.0.1000.8)
==12752==
            by 0x9218ACF: g object run dispose (in /usr/lib/x86 64-linux-gnu/libgobject-2.0.so.0.4002.0)
==12752==
            by 0x686A0FF: ??? (in /usr/lib/x86 64-linux-gnu/libgtk-3.so.0.1000.8)
==12752==
            by 0x6861F07: ??? (in /usr/lib/x86 64-linux-gnu/libgtk-3.so.0.1000.8)
            by 0x9212331: g closure invoke (in /usr/lib/x86 64-linux-gnu/libgobject-2.0.so.0.4002.0)
==12752==
==12752==
            by 0x92240D3: ??? (in /usr/lib/x86 64-linux-gnu/libgobject-2.0.so.0.4002.0)
==12752==
            by 0x922BA28: g signal emit valist (in /usr/lib/x86 64-linux-gnu/libgobject-2.0.so.0.4002.0)
==12752==
             by 0x922BCE1: g signal emit (in /usr/lib/x86 64-linux-gnu/libgobject-2.0.so.0.4002.0)
             by 0x6A30A57: ??? (in /usr/lib/x86 64-linux-gnu/libgtk-3.so.0.1000.8)
==12752==
            by 0x9218ACF: g object run dispose (in /usr/lib/x86 64-linux-gnu/libgobject-2.0.so.0.4002.0)
==12752==
==12752==
             by 0x68BB73F: ??? (in /usr/lib/x86 64-linux-gnu/libgtk-3.so.0.1000.8)
==12752==
==12752== Invalid write of size 8
```

Valgrind (Runtime validator)

- + Will probably catch the bug
- Will also catch every other 'bug'
- + No need for recompilation
- Needs to be integrated into your project
- Supression files help, but not entirely
- Severe memory corruptions can crash valgrind

Not considering giving up my will to live

Address Sanitizer (Runtime validator)

- Unlike Valgrind, needs to be compiled in
- Needs project and build system knowledge
- + Will catch lots of bugs
- Won't necessarily help you catch your bug

No setup available for Thunderbird

Record-and-replay debugger (RR)

- Written by Robert O'Callahan, designed for Firefox
- Allows nearly real-time (1.0–1.5x) record and replay
- Uses CPU performance counters



Usage

\$> rr record thunderbird 'Patch to pass a Windows RODC test.eml'

\$> rr replay

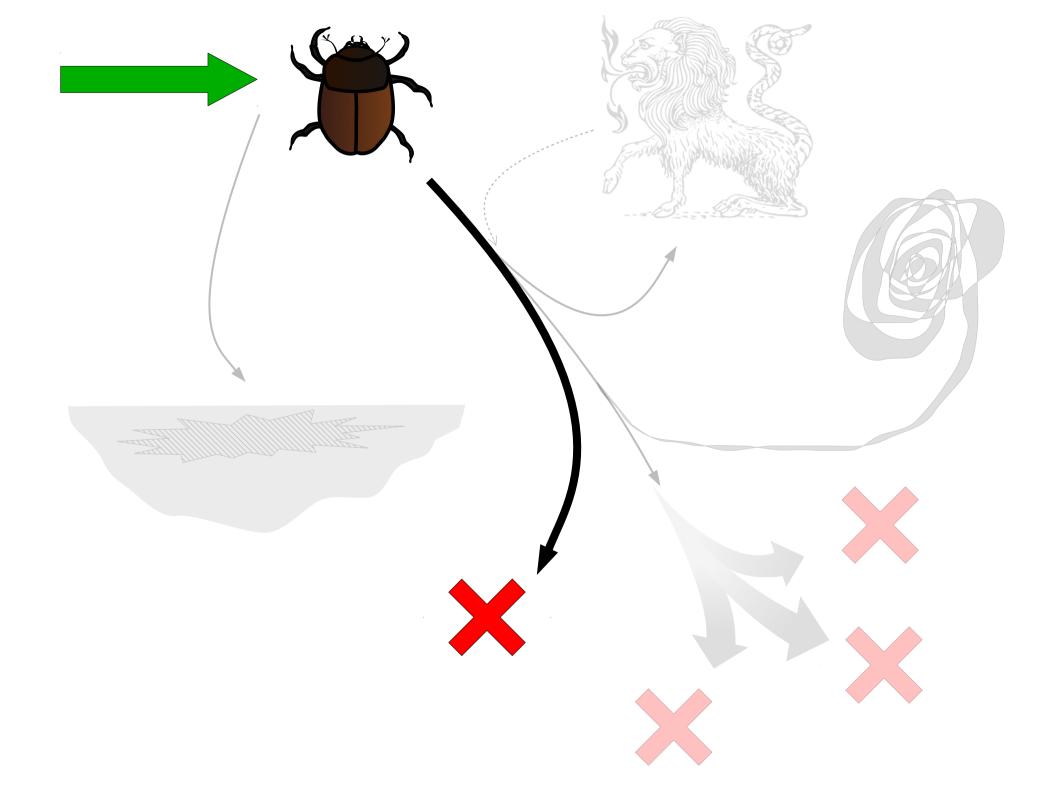
-M displays event numbers

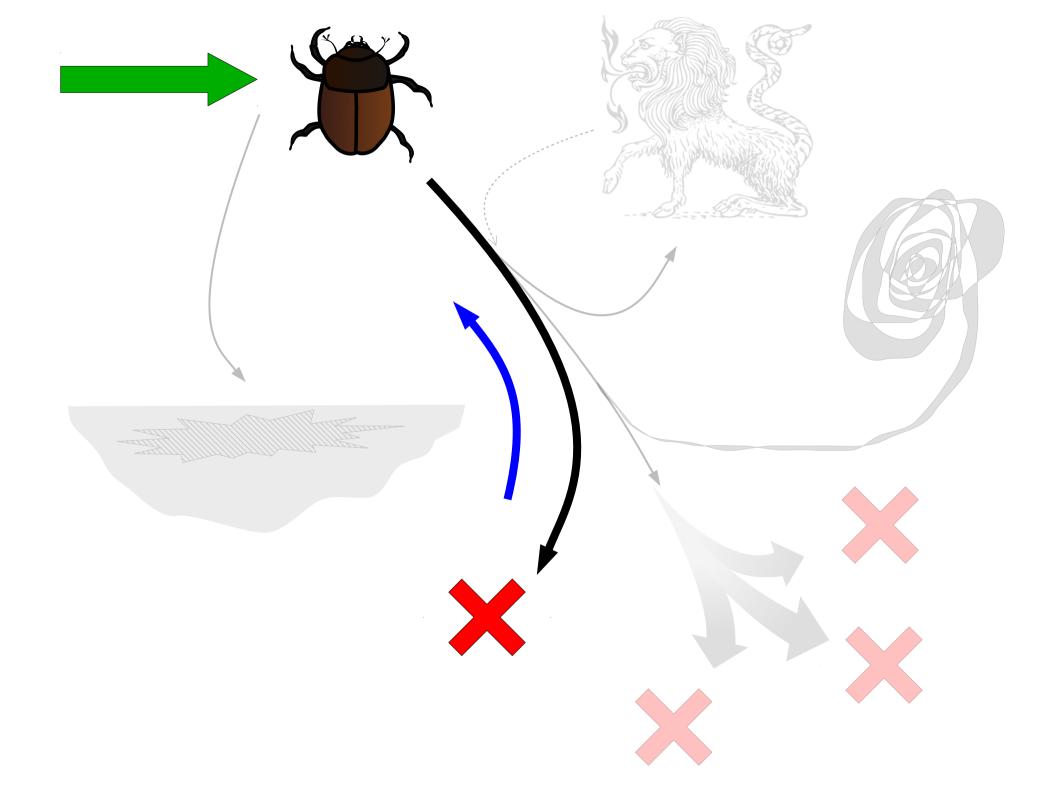
--goto <event number> replay normally until an event

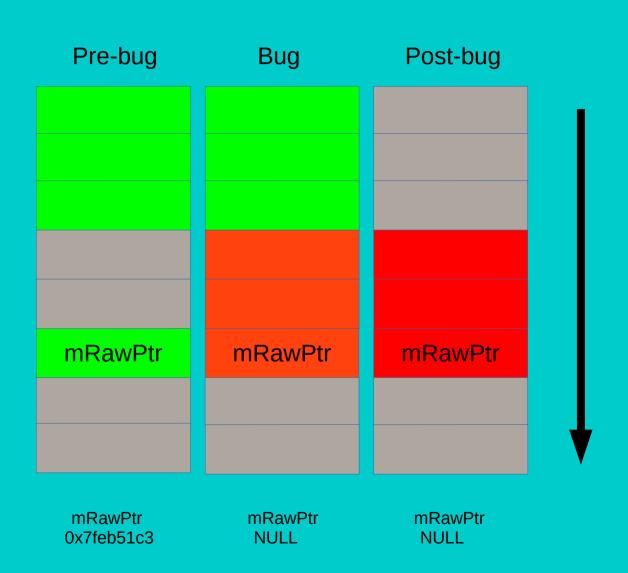
-p <PID> attach on new process

New commands: reverse-next, reverse-step, reverse-continue, reverse-finish

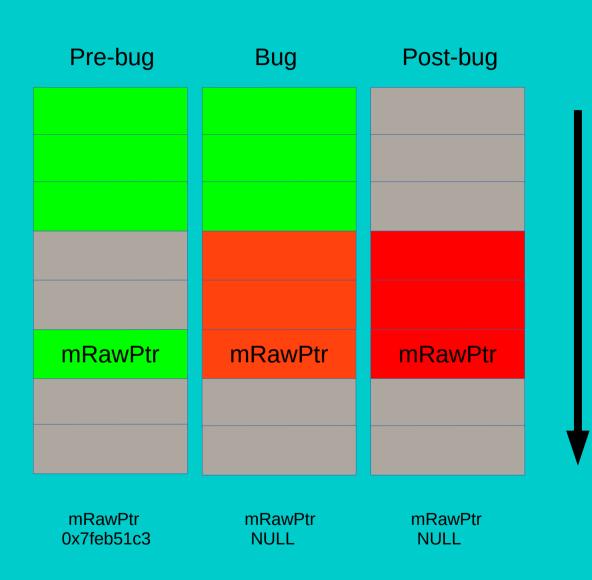
- + Completely deterministic (pids, addresses, registers)
- + No domain knowledge required
- + Replay repeatedly, can isolate cause from symptoms
- +/- Only native support for C, C++ and Rust







- 1. Thunderbird crash (SIGSEGV)
- 2. Set hardware watchpoint (mRawPtr)
- 3. Reverse-continue
- 4. Encounter exact location of bug



- 1. Thunderbird crash (SIGSEGV)
- 2. Set hardware watchpoint (mRawPtr)
- 3. Reverse-continue
- 4. Encounter exact location of bug

nsldapi_hex_unescape

```
Patch to pass a Windows RODC test - Mozilla Thunderbird
File Edit View Go Message Enigmail Tools Help
                   Write Chat Address Book
Tag ▼
                                                                          Reply Reply All
                                                                                                   ⇒ Forward
                                                                                                                More ▼
 From Me <garming@catalyst.net.nz>
Subject Patch to pass a Windows RODC test
                                                                                                          20/09/16 17:33
    To Catalyst Samba Developers <samba-dev@catalyst.net.nz>😭
Things to do:
- Automate credentials pre-loading in RODC environments (add
Administrator to RODC replication group + RODC preload run)
- Give a proper server for the referrals (PDC or replication partner
instead of hardcoding)
- Add the same check in the deleted case
-0001-Let-me-cheat-here.patch-
From eeba21f2bf1fb4a28d498ba032edcb8625eb68c7 Mon Sep 17 00:00:00 2001
From: Ubuntu <ubuntu@gar-mas-0-2.samdom.example.com>
Date: Tue, 20 Sep 2016 04:25:34 +0000
Subject: [PATCH] Let me cheat here...
Signed-off-by: Ubuntu <ubuntu@qar-mas-0-2.samdom.example.com>
 source4/dsdb/samdb/ldb modules/samldb.c
                                               10 +++++++
 2 files changed, 35 insertions(+)
diff --qit a/source4/dsdb/samdb/ldb modules/repl meta data.c b/source4/dsdb/samdb/ldb modules/repl meta data.c
index 7a5906e..d870479 100644
--- a/source4/dsdb/samdb/ldb modules/repl meta data.c
+++ b/source4/dsdb/samdb/ldb modules/repl meta data.c
@@ -976,6 +976,30 @@ static int replmd add(struct ldb module *module, struct ldb request *req)
               quid = GUID random();
        if (!ldb request get control(req, DSDB CONTROL REPLICATED UPDATE OID)
                       && !ldb request get control(reg, DSDB CONTROL DBCHECK MODIFY RO REPLICA)) {
               bool rodc = false;
               struct loadparm context *lp ctx;
               char *referral:
               ret = samdb rodc(ldb, &rodc);
               if (ret != LDB SUCCESS) {
                       DEBUG(4, ( location ": unable to tell if we are an RODC\n"));
               } else if (rodc) {
                       ldb set errstring(ldb, "RODC modify is forbidden!");
                       lp ctx = talloc get type(ldb get opaque(ldb, "loadparm"),
                                      struct loadparm context);
                       referral = talloc asprintf(req,
                                      "ldap://qar-mas-0-0.%s/%s",
                                      lnofa dnodomain(ln ctv)
```

```
Patch to pass a Windows RODC test - Mozilla Thunderbird
File Edit View Go Message Enigmail Tools Help
                   Write Chat Address Book
Tag ▼
                                                                          Reply Reply All
                                                                                                   ➡ Forward
                                                                                                                More ▼
 From Me <garming@catalyst.net.nz>
Subject Patch to pass a Windows RODC test
                                                                                                          20/09/16 17:33
    To Catalyst Samba Developers <samba-dev@catalyst.net.nz>😭
Things to do:
- Automate credentials pre-loading in RODC environments (add
Administrator to RODC replication group + RODC preload run)
- Give a proper server for the referrals (PDC or replication partner
instead of hardcoding)
- Add the same check in the deleted case
-0001-Let-me-cheat-here.patch-
From eeba21f2bf1fb4a28d498ba032edcb8625eb68c7 Mon Sep 17 00:00:00 2001
From: Ubuntu <ubuntu@gar-mas-0-2.samdom.example.com>
Date: Tue, 20 Sep 2016 04:25:34 +0000
Subject: [PATCH] Let me cheat here...
Signed-off-by: Ubuntu <ubuntu@qar-mas-0-2.samdom.example.com>
 source4/dsdb/samdb/ldb modules/samldb.c
                                               10 +++++++
 2 files changed, 35 insertions(+)
diff --git a/source4/dsdb/samdb/ldb modules/repl meta data.c b/source4/dsdb/samdb/ldb modules/repl meta data.c
index 7a5906e..d870479 100644
--- a/source4/dsdb/samdb/ldb modules/repl meta data.c
+++ b/source4/dsdb/samdb/ldb modules/repl meta data.c
@@ -976,6 +976,30 @@ static int replmd add(struct ldb module *module, struct ldb request *req)
               quid = GUID random();
        if (!ldb request get control(req, DSDB CONTROL REPLICATED UPDATE OID)
                       && !ldb request get control(reg, DSDB CONTROL DBCHECK MODIFY RO REPLICA)) {
               bool rodc = false;
               struct loadparm context *lp ctx;
               char *referral:
               ret = samdb rodc(ldb, &rodc);
               if (ret != LDB SUCCESS) {
                       DEBUG(4, ( location ": unable to tell if we are an RODC\n"));
               } else if (rodc) {
                       ldb set errstring(ldb, "RODC modify is forbidden!");
                       lp ctx = talloc get type(ldb get opaque(ldb, "loadparm"),
                                      struct loadparm context);
                       referral = tattoc_asprintr(req,
                                      "ldap://qar-mas-0-0.%s/%s",
                                       locfo docdomain/lo ctyl
```

```
&& !ldb request get control(reg, DSDB CONTROL DBCHECK MODIFY RO REPLICA)) {
bool rodc = false;
struct loadparm context *lp ctx;
char *referral;
ret = samdb rodc(ldb, &rodc);
if (ret != LDB SUCCESS) {
       DEBUG(4, ( location ": unable to tell if we are an RODC\n"));
} else if (rodc) {
       ldb set errstring(ldb, "RODC modify is forbidden!");
        lp ctx = talloc get type(ldb get opaque(ldb, "loadparm"),
                        struct loadparm context);
        referral = talloc asprintf(rea
                        lpcfg dnsdomain(lp ctx),
                        ldb dn get linearized(reg->op.add.message->dn));
        ret = ldb module send referral(reg, referral);
        return ret;
```

ldap/c-sdk/libraries/libldap/unescape.c

unescape.c deals with percent encoding (a.k.a. url encoding or RFC 1738 encoding)

```
2 short functions
```

```
/* **** BEGIN LICENSE BLOCK ****
* Version: MPL 1.1/GPL 2.0/LGPL 2.1
* The contents of this file are subject to the Mozilla Public License Version
 * 1.1 (the "License"); you may not use this file except in compliance with
 * the License. You may obtain a copy of the License at
 * http://www.mozilla.org/MPL/
 * Software distributed under the License is distributed on an "AS IS" basis,
 * WITHOUT WARRANTY OF ANY KIND, either express or implied. See the License
 * for the specific language governing rights and limitations under the
 * The Original Code is Mozilla Communicator client code, released
 * March 31, 1998.
 * The Initial Developer of the Original Code is
 * Netscape Communications Corporation.
 * Portions created by the Initial Developer are Copyright (C) 1998-1999
 * the Initial Developer. All Rights Reserved.
 * Contributor(s):
 * Alternatively, the contents of this file may be used under the terms of
 * either of the GNU General Public License Version 2 or later (the "GPL").
 * or the GNU Lesser General Public License Version 2.1 or later (the "LGPL"),
 * in which case the provisions of the GPL or the LGPL are applicable instead
 * of those above. If you wish to allow use of your version of this file only
 * under the terms of either the GPL or the LGPL, and not to allow others to
 * use your version of this file under the terms of the MPL, indicate your
 * decision by deleting the provisions above and replace them with the notice
 * and other provisions required by the GPL or the LGPL. If you do not delete
 * the provisions above, a recipient may use your version of this file under
 * the terms of any one of the MPL, the GPL or the LGPL.
 * ***** END LICENSE BLOCK ***** */
* LIBLDAP unescape.c -- LDAP URL un-escape routines
 * We also tolerate URLs that look like: <ldapurl> and <URL:ldapurl>
#include "ldap-int.h"
static int unhex( char c ):
nsldapi_hex_unescape( char *s )
* Remove URL hex escapes from s... done in place. The basic concept for
 * this routine is borrowed from the WWW library HTUnEscape() routine.
            for ( p = s; *s != '\0'; ++s ) {
                       if ( *s == '%' ) {
                                   if ( *++s != '\0' ) {
                                               *p = unhex( *s ) << 4;
                                   if ( *++s != '\0' ) {
                                               *p++ += unhex( *s );
                       } else {
            *p = '\0';
static int
unhex( char c )
            return( c >= '0' && c <= '9' ? c - '0'
               : c >= 'A' && c <= 'F' ? c - 'A' + 10
               : c - 'a' + 10 ):
```

}

url encoding

A "%" character followed by 2 hexadecimal digits encodes the byte those digits refer to.

decoding a single hexadecimal character

```
10
  B 11
  C 12
d D 13
e E 14
f F
     15
```

```
C means 12
c also means 12
3 means 3
```

decoding a single hexadecimal character with unhex()

```
... # $ % ... / 0 1 2 3 4 5 6 7 8 9 : ... ? @ A B C D E F G ... _ ` a b c d e f g h ... r s t u ...
0 1 2 3 4 5 6 7 8 9
```

Decoding a single hexadecimal character with unhex()

Decoding a single hexadecimal character with unhex()

Decoding a single hexadecimal character with unhex()

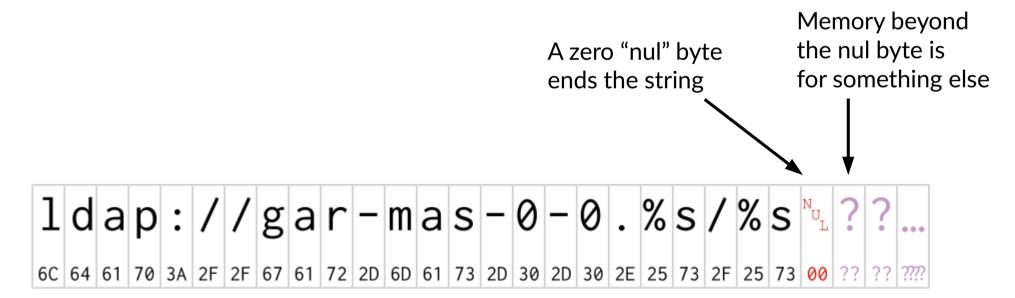
Reading a single hex byte

```
... # $ % ... / 0 1 2 3 4 5 6 7 8 9 : ... ? @ A B C D E F G ... _ ` a b c d e f g h ... r s t u ... -52 -51 -50 ... -40 0 1 2 3 4 5 6 7 8 9 -29 ... -24 -23 10 11 12 13 14 15 -16 ... 8 9 10 11 12 13 14 15 16 17 ... 27 28 29 39 ...
```

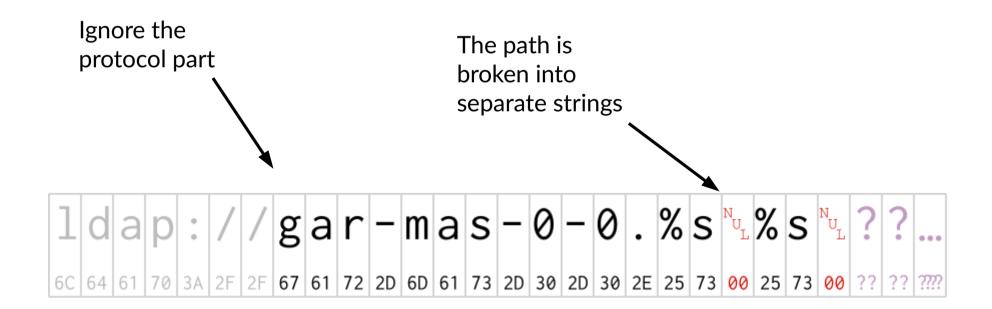
```
void
nsldapi_hex_unescape( char *s )
{
/*
 * Remove URL hex escapes from s... done in place. The basic concept for
 * this routine is borrowed from the WWW library HTUnEscape() routine.
 */
    char
            *p;
   for ( p = s; *s != '\0'; ++s ) {
            if ( *s == '%' ) {
                if ( *++s != '\0' ) {
                    *p = unhex( *s ) << 4;
                if ( *++s != '\0' ) {
                    *p++ += unhex( *s );
            } else {
                *p++ = *s;
    *p = ' \ 0';
```

The string "Idap://gar-mas-0-0.%s/%s" in memory

C string handling:



"Idap://gar-mas-0-0.%s/%s" after preprocessing



```
for (p = s; *s != '\0'; ++s) {
        if ( *s == '%' ) {
            if ( *++s != '\0' ) {
                 *p = unhex( *s ) << 4;
                                                     does s point to a percent?
            }
            if ( *++s != '\0' ) {
                                                      no.
                *p++ += unhex( *s );
                                                      then copy it to p.
             }
                                                       (and move p along).
        } else {
            *p++ = *s;
}
*p = ' \setminus 0';
```



```
for (p = s; *s != '\0'; ++s)
        if ( *s == '%' ) {
            if ( *++s != '\0' ) {
                 *p = unhex( *s ) << 4;
                                                     then move s along
            }
            if ( *++s != '\0' ) {
                                                      does s point to NUL?
                 *p++ += unhex( *s );
                                                      no...
             }
                                                      keep going
        } else {
            *p++ = *s:
}
*p = ' \setminus 0';
```

```
for (p = s; *s != '\0'; ++s)
        if ( *s == '%' ) {
            if ( *++s != '\0' ) {
                *p = unhex( *s ) << 4;
            }
            if ( *++s != '\0' ) {
                *p++ += unhex( *s );
            }
        } else {
            *p++ = *s;
}
*p = ' \setminus 0';
```

g|a|r|-|m|a|s|-|0|-|0|. |%|s №| s|

70 3A 2F 2F 67 61 72 2D 6D 61 73 2D 30 2D 30 2E 25 73 00 25 73 00 ?? ?? ???

```
for (p = s; *s != '\0'; ++s)
                                                                  if ( *s == '%' ) {
                                                                                                                                                                                                                                                                                                                                                                                                                                 does s point to a percent?
                                                                                                   if ( *++s != '\0' ) {
                                                                                                                                    *p = unhex( *s ) << 4;
                                                                                                   }
                                                                                                                                                                                                                                                                                                                                                                                                                                                   YES.
                                                                                                   if ( *++s != '\0' ) {
                                                                                                                                *p++ += unhex( *s );
                                                                                                     }
                                                                   } else {
                                                                                                   *p++ = *s;
  }
*p = ' \setminus 0';
                                           d|a|p|: |/|/|g|a|r|-|m|a|s|-|0|-|0|. |%|s|_{v_{L}} | |s|_{v_{L}} | |s|
```

64 61 70 3A 2F 2F 67 61 72 2D 6D 61 73 2D 30 2D 30 2E 25 73 00 25 73 00 ?? ?? ?? ????

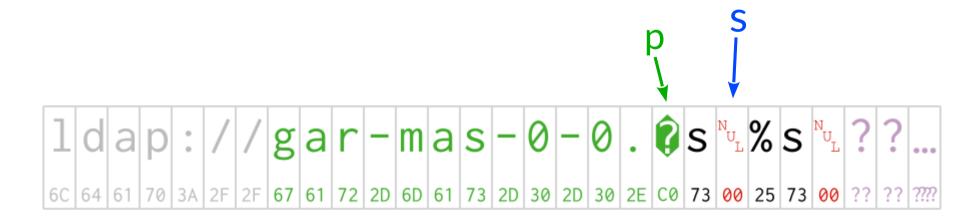
```
for (p = s; *s != '\0'; ++s)
                                                      move s along
        if ( *s == '%' ) {
                                                      does it point to NUL?
             if ( *++s != '\0' ) {
                 *p = unhex( *s ) << 4;
             }
             if ( *++s != '\0' ) {
                                                       no...
                 *p++ += unhex( *s );
                                                       then unhex it and write a
        } else {
                                                       temporary value to p
             *p++ = *s;
}
*p = ' \setminus 0';
```

|g|a|r|-|m|a|s|-|0|

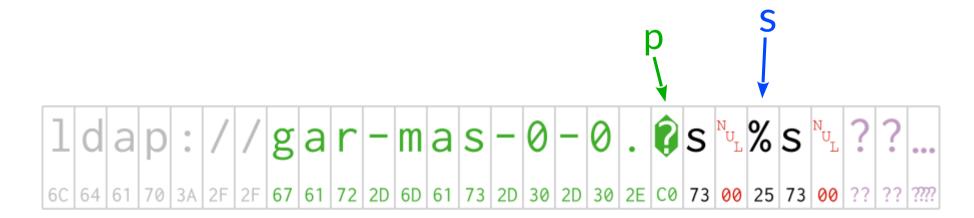
6C | 64 | 61 | 70 | 3A | 2F | 2F | 67 | 61 | 72 | 2D | 6D | 61 | 73 | 2D | 30 | 2D | 30 | 2E | C0 | 73 | 00 | 25 | 73 | 00 | ?? | ?? | ????

|-|0|. $||\mathbf{\hat{g}}||\mathbf{s}||^{\mathsf{N}_{\mathsf{U}_{\mathsf{L}}}}|\mathbf{\hat{s}}||^{\mathsf{N}_{\mathsf{U}_{\mathsf{L}}}}|?|$

```
for (p = s; *s != '\0'; ++s) {
                                                         move s again
        if ( *s == '%' ) {
                                                         does it point to NUL now?
            if ( *++s != '\0' ) {
                 *p = unhex( *s ) << 4;
            }
            if ( *++s != '\0' ) {
                                                      YES!
                 *p++ += unhex( *s );
                                                      don't unhex it
        } else {
                                                      don't move p on
            *p++ = *s;
}
*p = ' \setminus 0';
```

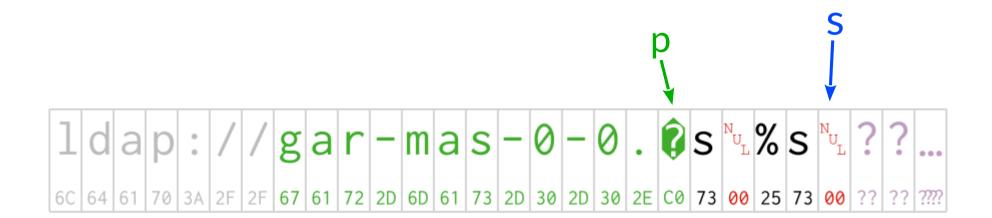


```
for (p = s; *s != '\0'; ++s)
                                                        move s again
        if ( *s == '%' ) {
                                                        does it point to NUL now?
            if ( *++s != '\0' ) {
                *p = unhex( *s ) << 4;
            }
                                                         no
            if ( *++s != '\0' ) {
                *p++ += unhex( *s );
                                                         but it is another '%'
        } else {
            *p++ = *s;
}
*p = ' \setminus 0';
```



```
for (p = s; *s != '\0'; ++s)
        if ( *s == '%' ) {
            if ( *++s != '\0' ) {
                *p = unhex( *s ) << 4;
            }
            if ( *++s != '\0' ) {
                *p++ += unhex( *s );
        } else {
            *p++ = *s;
}
*p = ' \setminus 0';
```

repeat the "%s $^{N_{U_L}}$ " pattern not moving p

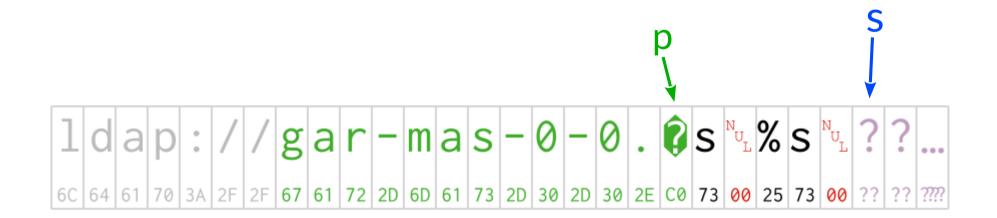


```
for (p = s; *s != '\0'; ++s) {
        if ( *s == '%' ) {
            if ( *++s != '\0' ) {
                 *p = unhex( *s ) << 4;
            }
            if ( *++s != '\0' ) {
                *p++ += unhex( *s );
        } else {
            *p++ = *s;
}
*p = ' \setminus 0';
```

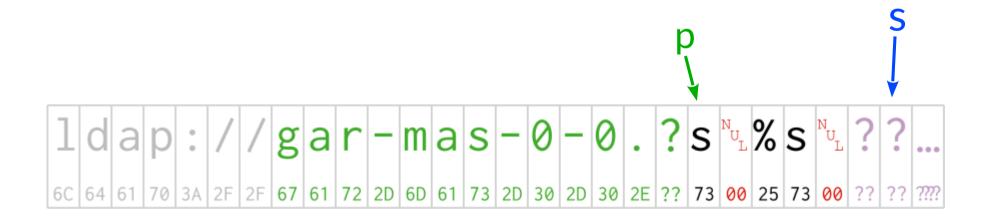
NOW does *s* point to NUL?

Nobody knows.

but if it doesn't, the character is copied to *p*



```
for ( p = s; *s != '\0'; ++s ) {
        if ( *s == '%' ) {
            if ( *++s != '\0' ) {
                 *p = unhex( *s ) << 4;
            }
            if ( *++s != '\0' ) {
                 *p++ += unhex( *s );
        } else {
            *p++ = *s;
}
*p = ' \setminus 0';
```

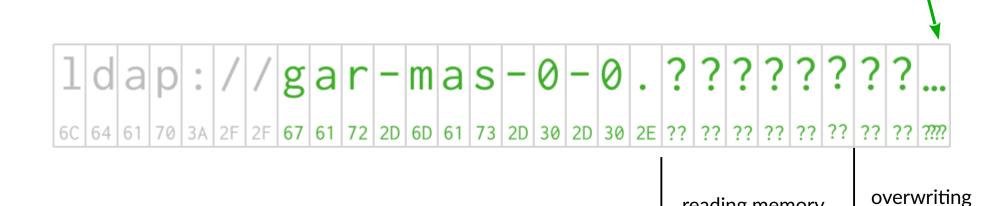


```
for (p = s; *s != '\0'; ++s)
        if ( *s == '%' ) {
            if ( *++s != '\0' ) {
                *p = unhex( *s ) << 4;
            }
            if ( *++s != '\0' ) {
                *p++ += unhex( *s );
        } else {
            *p++ = *s:
}
*p = ' \setminus 0';
```

Eventually s will hit a NUL and stop.

reading memory

memory



```
for (p = s; *s != '\0'; ++s) {
                                                               Thunderbird thinks it
         if ( *s == '%' ) {
                                                               stopped at the first NUL
              if ( *++s != '\0' ) {
                   *p = unhex( *s ) << 4;
                                                               and starts again
              }
              if ( *++s != '\0' ) {
                  *p++ += unhex( *s );
         } else {
              *p++ = *s;
                            a|r|-|m|a|s|-|0|-|0|.|%|s|^{v_{L}}%|s|^{v_{L}}?
  6C | 64 | 61 | 70 | 3A | 2F | 2F | 67 | 61 | 72 | 2D | 6D | 61 | 73 | 2D | 30 | 2D | 30 | 2E | 25 | 73 | <mark>00 | 25 | 73 | 00</mark> | ?? | ?? | ?????
                         gar-mas-0
```

Who else is using this?

```
/*
* Remove URL hex escapes from s... done in place. The basic concept for
* this routine is borrowed from the WWW library HTUnEscape() routine.
*/
    char
            *p;
   for (p = s; *s != '\0'; ++s) {
                                                        Google for this string
            if ( *s == '%' ) {
                if ( *++s != '\0' ) {
                    *p = unhex( *s ) << 4;
                if ( *++s != '\0' ) {
                    *p++ += unhex( *s );
            } else {
                *p++ = *s;
    *p = ' \setminus 0';
```



Videos

Images

News

Shopping

More

Settings

Tools

About 173 results (0.89 seconds)

apr_ldap_url.c - [svn.apache.org]svn.us.apache.org; svn.eu.apache.org
https://svn.apache.org/repos/asf/apr/apr-util/tags/1.2.10/ldap/apr_ldap_url.c ▼
... static void apr_ldap_pvt_hex_unescape(char *s) { /* * Remove URL hex escapes from s... done in place. The basic concept for * this routine is borrowed from ...

Copyright 2000-2004 The Apache Software Foundation * * Licensed ... https://svn.apache.org/repos/asf/apr/apr-util/tags/1.0.0/ldap/apr_ldap_url.c ... static void apr_ldap_pvt_hex_unescape(char *s) { /* * Remove URL hex escapes from s... done in place. The basic concept for * this routine is borrowed from ...

apr_ldap_url.c - The Apache Software Foundation!

https://svn.apache.org/repos/asf/apr/apr-util/branches/1.2.x/ldap/apr_ldap_url.c ▼ ... static void apr_ldap_pvt_hex_unescape(char *s) { /* * Remove URL hex escapes from s... done in place. The basic concept for * this routine is borrowed from ...

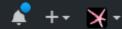
apr_ldap_url.c - The Apache Software Foundation!

https://svn.apache.org/repos/asf/apr/apr-util/branches/0.9.x/ldap/apr_ldap_url.c NULL) { LDAP_VFREE(ludp->lud_exts); } LDAP_FREE(ludp); } static void ldap pvt hex unescape(char*s) { /* * Remove URL hex escapes from s... done in ...

Apache/apr_Idap_url.c at master · omnigroup/Apache · GitHub
https://github.com/omnigroup/Apache/blob/master/apr-util/Idap/apr_Idap_url.c ▼
Remove URL hex escapes from s... done in place. The basic concept for. * this routine is borrowed from the WWW library HTUnEscape() routine. */. char *p;.

illumos-nexenta/unescape.c at master · Nexenta/illumos-nexenta ...
https://github.com/Nexenta/illumos-nexenta/blob/master/usr/src/lib/.../unescape.c ▼
Remove URL hex escapes from s... done in place. The basic concept for. * this routine is borrowed from the WWW library HTUnEscape() routine. */. char *p;.



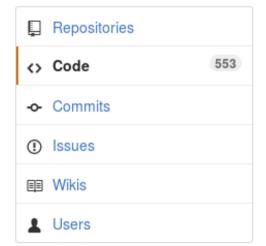


Search

"Remove URL hex escapes from s..."



Search



Languages

С	443
C++	2

Advanced search Cheat sheet

Showing 448 available code results 3

Sort: Best match ▼



С

Showing the top three matches. Last indexed on 16 Sep 2016.

```
#include "ldap-int.h"

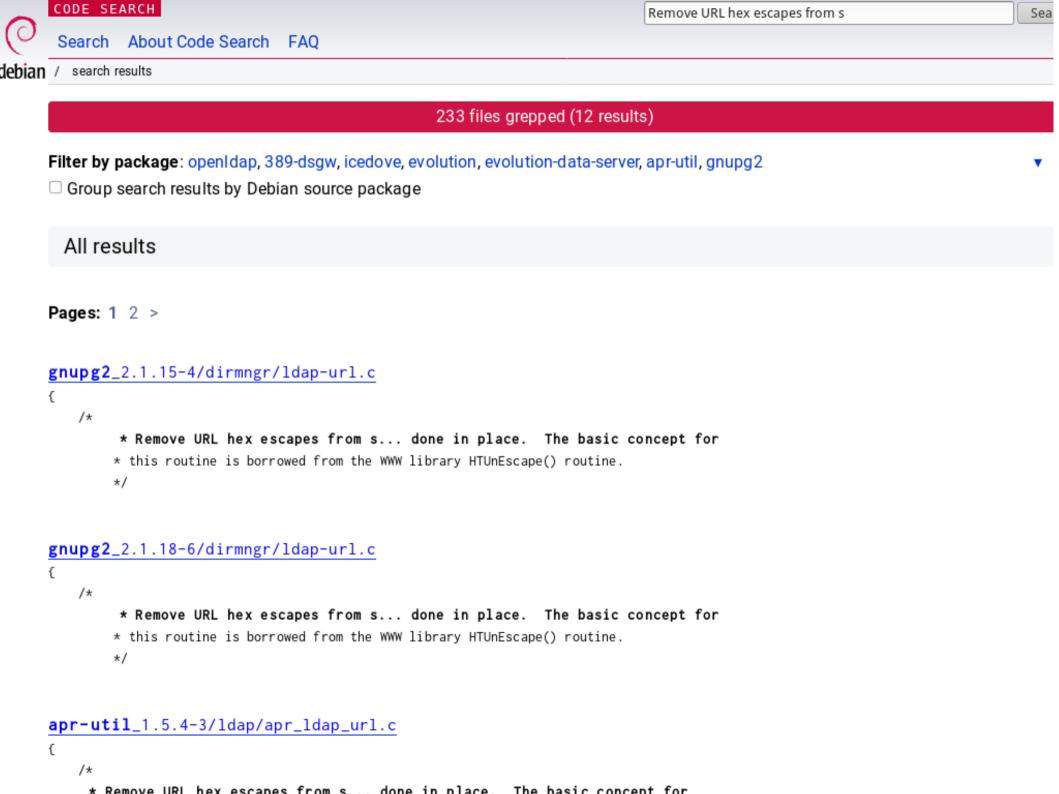
#include "late of the color of the late of the lat
```

nxmirrors/onny - unescape.c

С

Showing the top three matches. Last indexed on 14 Sep 2016.

```
30  #include "ldap-int.h"
31
32
33  static int unhex( char c );
34
35
36  void
37  nsldapi_hex_unescape( char *s )
38  {
```



Many of these have used openIdap's version (fixed in 2002).

Most Github projects are dead forks.

Of the projects in Debian, only icedove and 389-dsgw have the bug.

Illumos (formerly Solaris) is also affected.

For a little while we thought we found a similar problem in Firefox.

The fix

```
for (p = s; *s != '\0'; ++s) {
        if ( *s == '%' ) {
             if ( *++s != '\0' ) {
                 *p = unhex( *s ) << 4;
             if ( *++s == '\0' ) {
                 break;
            }
if ( *++s != '\0' ) {
                 *p++ += unhex( *s );
             *p = unhex( *s ) << 4;
             if (*++s == '\0')
                 break;
             *p++ += unhex( *s );
+
        } else {
             *p++ = *s;
```

```
void
nsldapi_hex_unescape( char *s )
/*
 * Remove URL hex escapes from s... done in place. The basic concept for
 * this routine is borrowed from the WWW library HTUnEscape() routine.
 */
    char *p;
    for (p = s; *s != '\0'; ++s) {
         if ( *s == '%' ) {
              if (*++s == '\0')
                  break;
              *p = unhex( *s ) << 4;
                                                           if you hit a NUL byte, stop
              if ( *++s == ' \setminus 0'
                  break;
              *p++ += unhex( *s );
         } else {
              *p++ = *s;
    *p = ' \setminus 0';
```

How can we call it a 25 year old bug when Thunderbird was first released in 2003?

Answer (part 1): the file is older

```
* The Original Code is Mozilla Communicator client code, released
* March 31, 1998.
*
* The Initial Developer of the Original Code is
* Netscape Communications Corporation.
* Portions created by the Initial Developer are Copyright (C) 1998-1999
```

and contains conflicting copyright claims:

* the Initial Developer. All Rights Reserved.

- * Copyright (c) 1996 Regents of the University of Michigan.
- * All rights reserved.

Dates in other versions:

- /* Copyright 2000-2005 The Apache Software Foundation or its licensors, as
 * applicable.
- /* Portions Copyright 1998-2002 The OpenLDAP Foundation
- * All rights reserved.
- * Portions Copyright (c) 1992-1996 Regents of the University of Michigan.
- * All rights reserved.
- * Copyright (c) 1999 Apple Computer, Inc. All rights reserved.

It probably dates to around 1996

back to the source:

```
* Remove URL hex escapes from s... done in place. The basic concept for
* this routine is borrowed from the WWW library HTUnEscape() routine.
*/
```

HTUnEscape may skip over NUL byte...

- To: www-lib-bugs@w3.org
- Subject: HTUnEscape may skip over NUL byte...
- From: Markku Savela <msa@msa.tte.vtt.fi>
- Date: Thu, 14 Nov 1996 14 17.16 +0200 (EET)
- From www-lib-bugs-request@www10.w3.org Thu Nov 14 07: 17:19 1996
- Message-Id: <199611141217.OAA16267@msa.tte.vtt.fi>
- Reply-to: msa@hemuli.tte.vtt.fi (Markku Savela)

If HTUnEscape is given a string like "....%2", it will step over the

• X-List-URL: http://www.w3.org/pub/WWW/Archives/Public/www-lib-bugs/

```
terminating NUL byte, and with bad luck, if the next bytes are
non-NUL, it may trash memory too. I think following minimal fix will
repair the problem:
*** HTEscape.c.ORIG
                     Wed May 1 02:31:29 1996
--- HTEscape.c Thu Nov 14 13:02:49 1996
*** 112,119 ****
          if (*p == HEX ESCAPE) {
            if (*p) *q = from hex(*p++) * 16;
            if (*p) *q = FROMASCII(*q + from hex(*p));
            p++, q++;
        } else {
            *q++ = *p++;
--- 112,119 ----
          if (*p == HEX ESCAPE) {
            p++;
            if (*p) *q = from hex(*p++) * 16;
           if (*p) *q = FROMASCII(*q + from hex(*p)), ++p;
            q++;
        } else {
            *q++ = *p++;
```

Disclaimer! This is of course my view of the matter. Someone in W3C should probably verify the patch (or invent a neater one).

three weeks later – fix as a compile-time option

```
Author: Eric Prud'hommeaux <eric@w3.org>
       Thu Dec 5 23:20:33 1996 +0000
Date:
    goin' back to Cali
diff --git a/Library/src/HTEscape.c b/Library/src/HTEscape.c
index f83abea...6385928 100644
--- a/Library/src/HTEscape.c
+++ b/Library/src/HTEscape.c
@@ -112,8 +112,13 @@ PUBLIC char * HTUnEscape (char * str)
         if (*p == HEX_ESCAPE) {
            p++:
            if (*p) *q = from_hex(*p++) * 16;
+#ifdef UNTESTED /* suggestion from Markku Savela - I just copied it in - EGP */
            if (*p) *q = FROMASCII(*q + from_hex(*p)), ++p;
+
            a++:
+#else /* UNTESTED */
            if (*p) *q = FROMASCII(*q + from_hex(*p));
            p++, q++;
+#endif /* !UNTESTED */
        } else {
            *q++ = *p++;
```

18 months later

```
commit 0ccbc2c9fac13f7687e5cbe05ae11431b2ec89b2
Author: Henrik Frystyk Nielsen <frytstik@w3.org>
Date: Fri Mar 20 17:53:18 1998 +0000
   version 5.1k
diff --git a/Library/src/HTEscape.c b/Library/src/HTEscape.c
@@ -111,14 +110,15 @@ PUBLIC char * HTUnEscape (char * str)
     while(*p) {
         if (*p == HEX_ESCAPE) {
            p++;
            if (*p) *q = from_hex(*p++) * 16;
-#ifdef UNTESTED /* suggestion from Markku Savela - I just copied it in - EGP */
            if (*p) *q = FROMASCII(*q + from_hex(*p)), ++p;
            if (*p) *q = HTAsciiHexToChar(*p++) * 16;
+#if 1
           /* Suggestion from Markku Savela */
            if (*p) *q = FROMASCII(*q + HTAsciiHexToChar(*p)), ++p;
            a++:
-#else /* UNTESTED */
           if (*p) *q = FROMASCII(*q + from_hex(*p));
+#else
            if (*p) *q = FROMASCII(*q + HTAsciiHexToChar(*p));
            p++, q++;
-#endif /* !UNTESTED */
+#endif
```

Actually only a 22 year old bug

Author: Henrik Frystyk Nielsen <frytstik@w3.org>

Date: Sun Nov 13 11:31:16 1994 +0000

Add HTEscape module

Trying to coordinate releases with the least possible effort

- get Illumos and Red Hat (389-dsgw) to sign up to Thunderbird bugzilla
- tell them to sort it out
- [weeks pass; everyone waits]
- Mozilla almost releases, others say no
- [weeks pass]
- Mozilla releases anyway

Garming was awarded a bug bounty

but is ineligible because we are contractors.