WINS-Replication

Stefan Metzmacher
SerNet – Service Network GmbH
Samba Team
metze@samba.org

http://samba.org/~metze/presentations/2006/metze_sambaxp2006_winsrepl.pdf

Who Am I?

- Student at the University of Applied Sciences Cologne
- First Samba-Patch in 2001
- Samba3-Work: VFS- and PASSDB-Modules, Quotas
- Samba Team Member since 2004
- Samba4-Work: Core-Infrastructure, ADS-Replication, ...
- SerNet Employee since 2005
- Samba4-Work: WINS-Replication, SMB2-Support, ...

Agenda

- Problem Space
- How to learn the replication protocol?
- Some important data structures
- What is Push- or Pull-Replication?
- How does the Replication protocol works?
- Replication conflict handling
- Server Implementation
- samba4wins

Problem Space

- All Windows versions use Netbios-Names for name resolution (dynamic registering/releasing of names)
- Win9x, NT4 and ME as default
- Win2k, WinXP, Win2k3 and Vista as fallback from DNS
- WINS – Windows Internet Name Service is used in routed networks with multiple subnets
- Large networks need multiple servers which hold the same Information to grant fault tolerance
- The replication uses a custom protocol on top of TCP/42
- Samba3 is a fast and stable WINS-Server, but doesn't support replication

How to learn the replication protocol?

- The basic packet format was already decoded in samba3
- A basic client library and torture test was already in samba4
- There's a good description of the WINS-Server architecture available under:

  http://msdn.microsoft.com/archive/default.asp? → url=/archive/en-us/dnarnetbios/html/msdn_winswp.asp
- To get a better feeling for the protocol, I added a dissector to the “ethereal” network sniffer.

NBT Name-Types

- UNIQUE
  - these are registered by workstation and servers
  - e.g.: MYSERVER<00>, MYSERVER<20>
  - only one ip-address
- NORMAL GROUP
  - these names are registered by all members of a WORKGROUP/DOMAIN
  - the WINS-Server always returns 255.255.255.255
- SPECIAL GROUP (Internet Groups)
  - registered by a group of computers
  - e.g.: MYDOMAIN<1C> for all domain controllers
  - 25 ip-addresses at maximum, more are just skipped
- MULTIHOMED
  - like UNIQUE names, but with up to 25 ip-addresses

Name-Record States

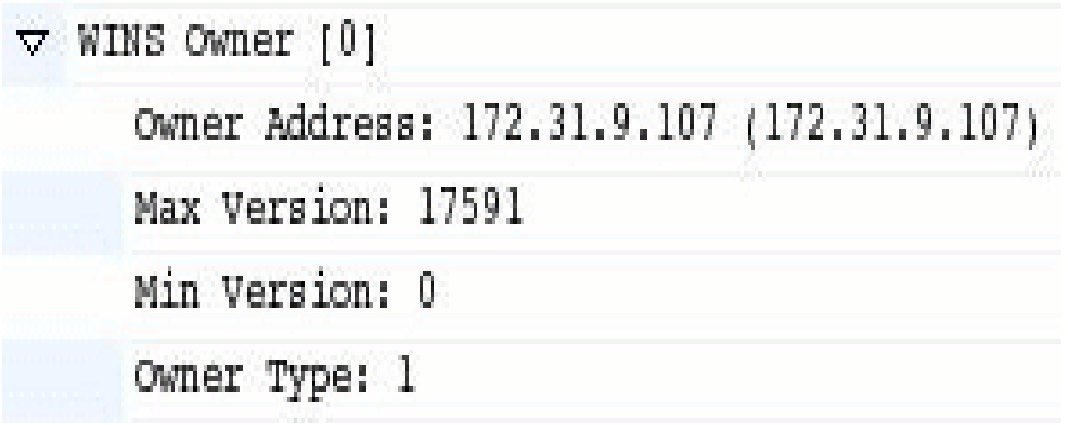
- ACTIVE
  - only active records are returned in WINS Name Queries
  - after 6 days without a WINS Name Release the record becomes RELEASED
- RELEASED
  - released records are not replicated
  - after 6 days they become TOMBSTONED
- TOMBSTONED
  - tombstoned records are replicated
  - after 1 day they will be removed from the database

- The state change is done during a database scavenging every 3 days
- Records can be static or dynamic
- Only dynamic records can be overwritten

Replication Partners

- Replication is only done between configured partners
- Not all WINS-Owners (WINS-Servers) needs to be partners.
- e.g. ServerA can get the records of ServerC via ServerB without directly replicating with server ServerC
- the meaning of Pull/Push-Partner is a bit confusing when you first hear it
- My Push-Partners will may get notifications from me
- My Push-Partners will be able to pull information from me
- I will may get notifications from my Pull-Partners
- I will try to pull information from my Pull-Partners in a configured interval
- typically servers are Pull- and Push-Partners

WINS-Owner structure



- This is one of the core structure in the replication protocol
- It represents one WINS-Server and a given range of version_ids
- The owner type is completely ignored

WINS-Name structure (single address)

```
▽ WINS Name [0]: _SAME_OWNER_A<00>: 127.0.65.2
    Name Len: 17
  ▷ Name: _SAME_OWNER_A<00> (Workstation/Redirector)
  ▽ Name Flags: 0x00000000
      .... .... .... .... .... .... .... ..00 = Record Type: Unique (0x00000000)
      .... .... .... .... .... .... .... 00.. = Record State: Active (0x00000000)
      .... .... .... .... .... .... ...0 .... = Local: False
      .... .... .... .... .... .... .00. .... = Host Type: B-node (0x00000000)
      .... .... .... .... .... .... 0... .... = Static: False
    Name Group Flag: 0x00000000
    Name Version Id: 31087
    IP Address: 127.0.65.2 (127.0.65.2)
    Unknown IP: 255.255.255.255 (255.255.255.255)
```

WINS-Name structure (address list)

```
WINS Name [0]: _SAME_OWNER_A<00>: 127.0.65.3, 127.0.65.4
    Name Len: 17
    Name: _SAME_OWNER_A<00> (Workstation/Redirector)
    Name Flags: 0x00000002
        .... .... .... .... .... .... .... ..10 = Record Type: Speical group (0x00000002)
        .... .... .... .... .... .... .... 00.. = Record State: Active (0x00000000)
        .... .... .... .... .... .... ...0 .... = Local: False
        .... .... .... .... .... .... .00. .... = Host Type: B-node (0x00000000)
        .... .... .... .... .... .... 0... .... = Static: False
    Name Group Flag: 0x00000001
    Name Version Id: 31092
    WINS Address List: 127.0.65.3, 127.0.65.4
        Num IPs: 2
        WINS IP [0]: 127.0.65.3
            IP Owner: 127.65.65.1 (127.65.65.1)
            IP Address: 127.0.65.3 (127.0.65.3)
        WINS IP [1]: 127.0.65.4
            IP Owner: 127.65.65.1 (127.65.65.1)
            IP Address: 127.0.65.4 (127.0.65.4)
    Unknown IP: 255.255.255.255 (255.255.255.255)
```

WINS-Replication Message-Types

- WREPL_START_ASSOCIATION
  - a request to create a replication association (context)
  - like the first half of an “open()” call
- WREPL_START_ASSOCIATION_REPLY
  - the reply that finishes the “open()” call
- WREPL_STOP_ASSOCIATION
  - this message destroy a replication context
  - like a “close()” call
- WREPL_REPLICATION
  - this messages require a valid association
  - there are some subcommands which do the replication

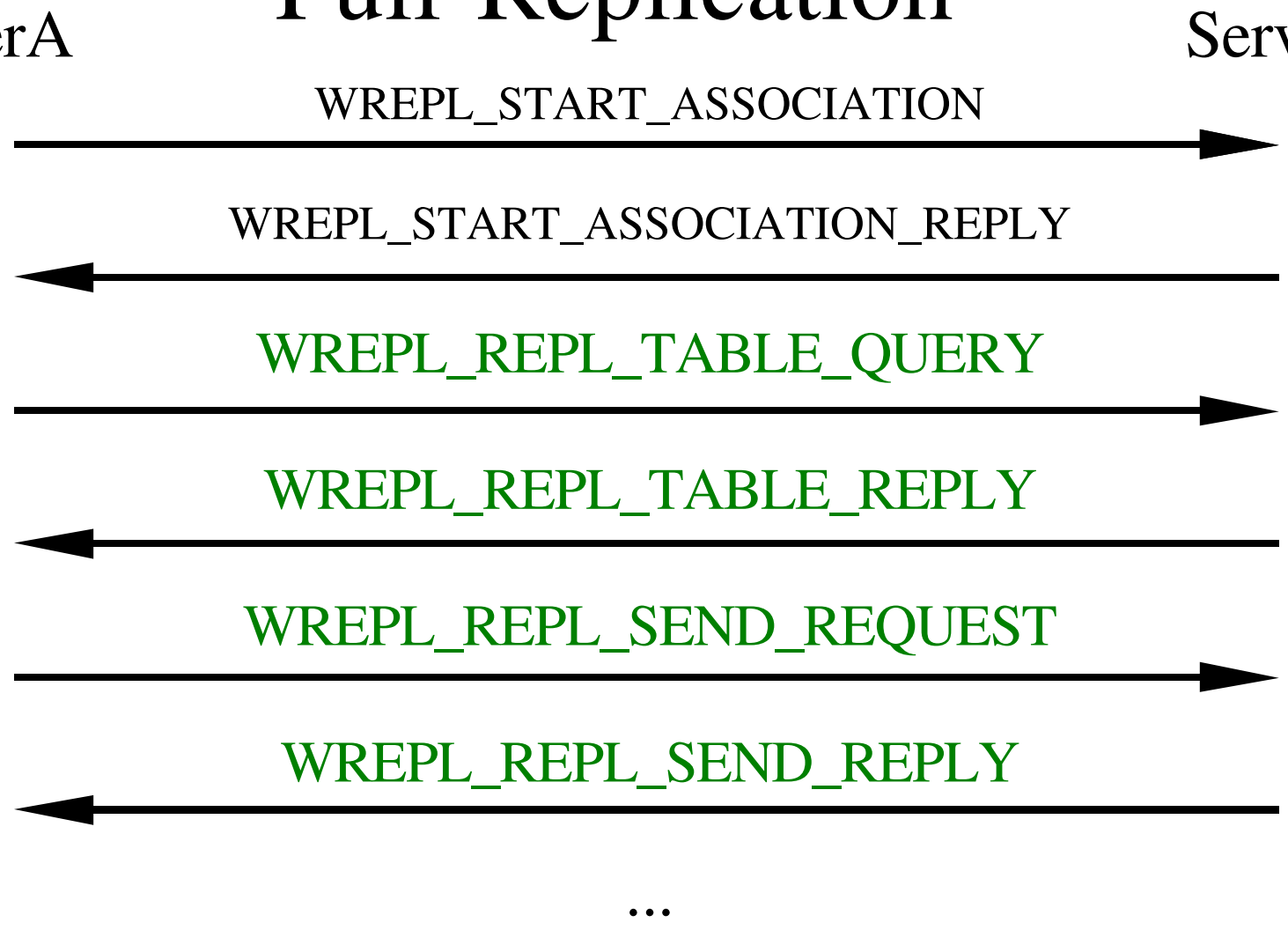
Replication Commands (Pull)

- WREPL_REPL_TABLE_QUERY
  - this call doesn't have any specific data
  - it asks for the server owner/version_id table
- WREPL_REPL_TABLE_REPLY
  - this reply gives back the owner/version_id table of the server
- WREPL_REPL_SEND_REQUEST
  - this request asks for the name records in a given version_id range of a specific owner
- WREPL_REPL_SEND_REPLY
  - sends all name records in the requested version_id range

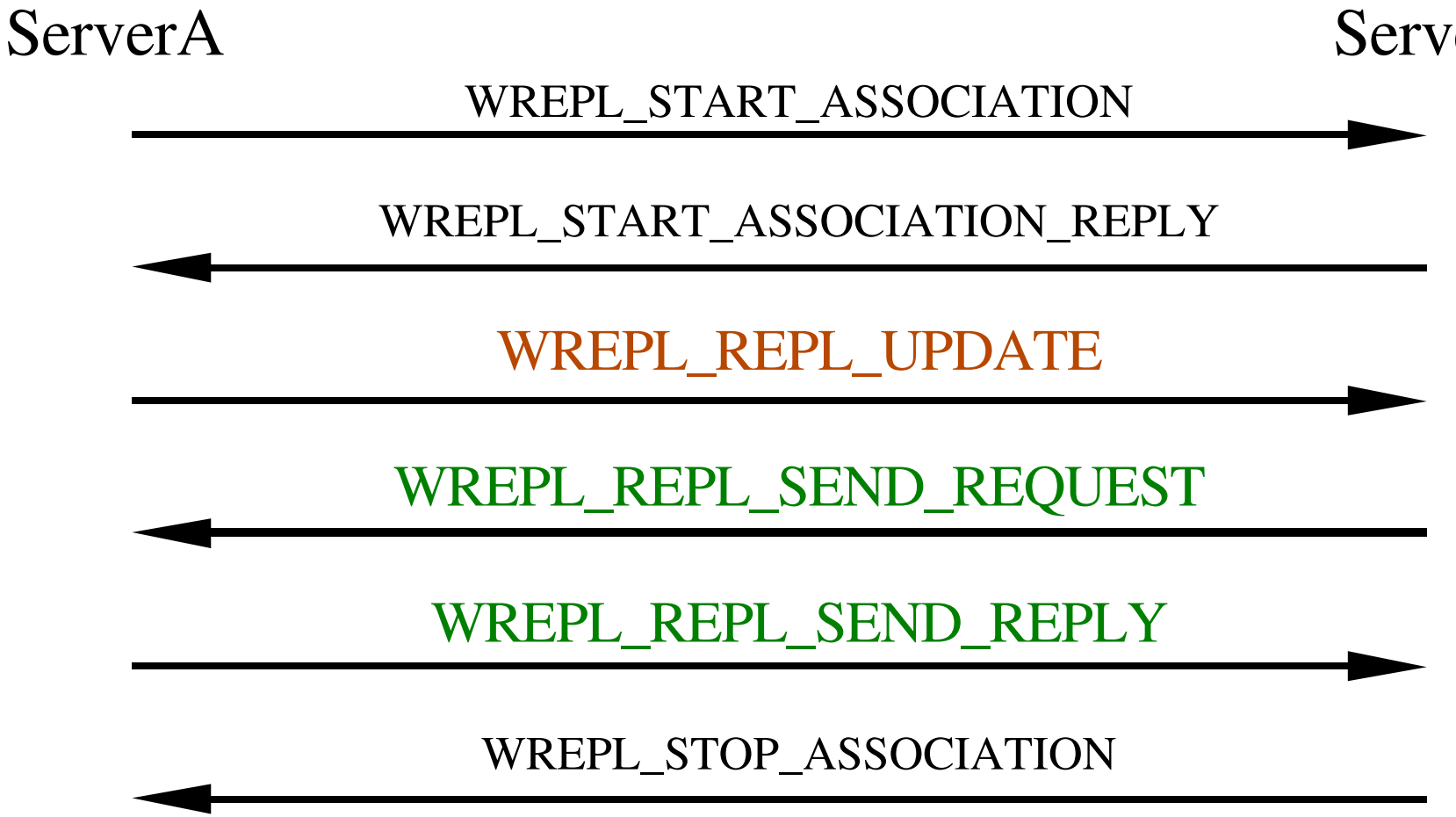
Replication Commands (Push)

- WREPL_REPL_UPDATE / WREPL_REPL_UPDATE2
- WREPL_REPL_INFORM / WREPL_REPL_INFORM2
  - with this messages a server informs the partner that it has new information, which can be pulled down
  - the content is the same as in a TABLE_REPLY
- NT4 uses UPDATE / UPDATE2
  - as reaction to this message the Client/Server Roles are turned on the TCP connection
  - then the “new” client then pull information like in the PULL-Replication case and close the connection afterwards
- Win2k/Win2k3 use INFORM / INFORM2
  - this messages doesn't have a reply and the connection is persistent
  - this messages just trigger a normal PULL-Replication on the peer

Pull-Replication



Push-Replication UPDATE



Push-Replication INFORM

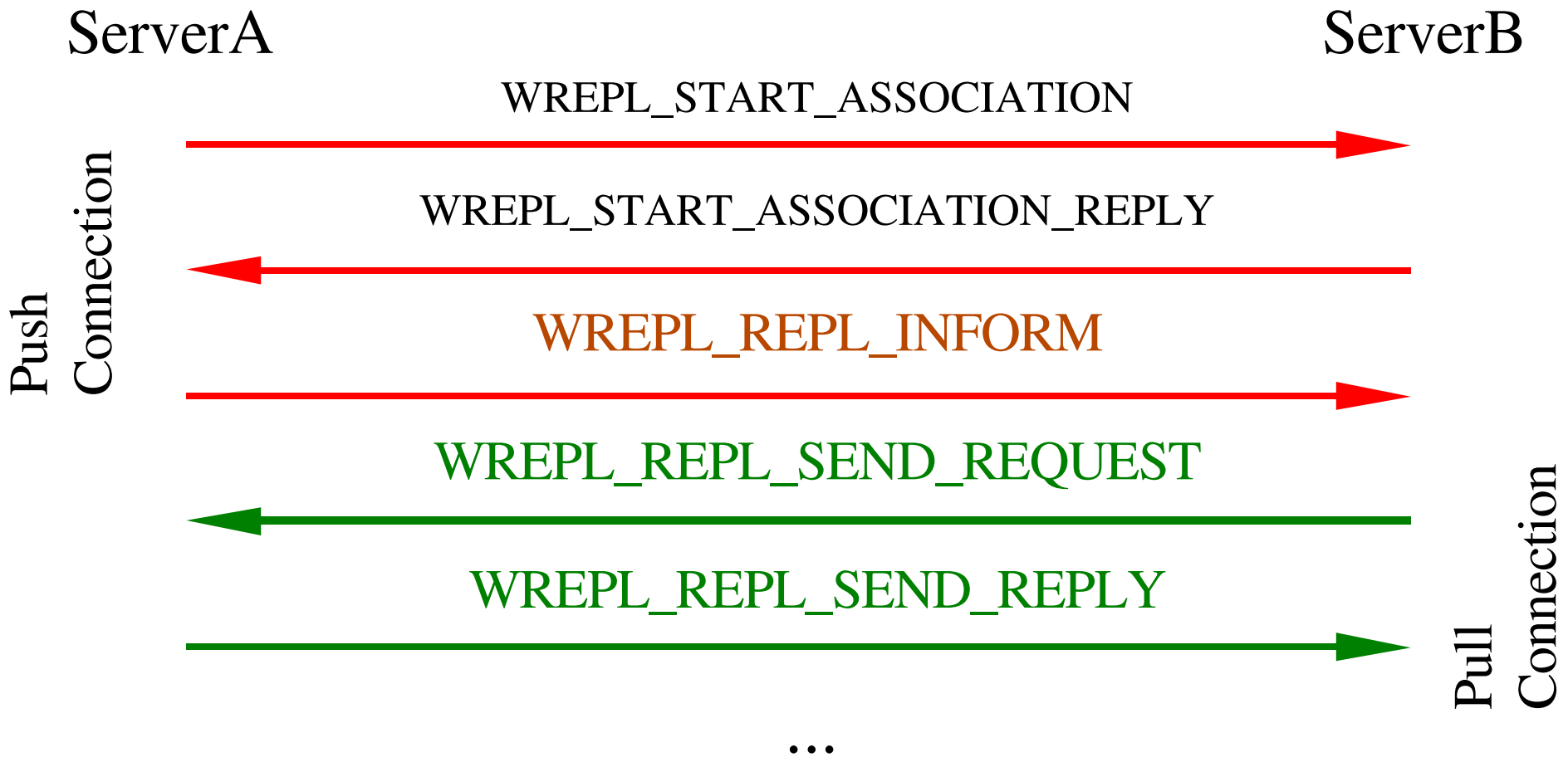
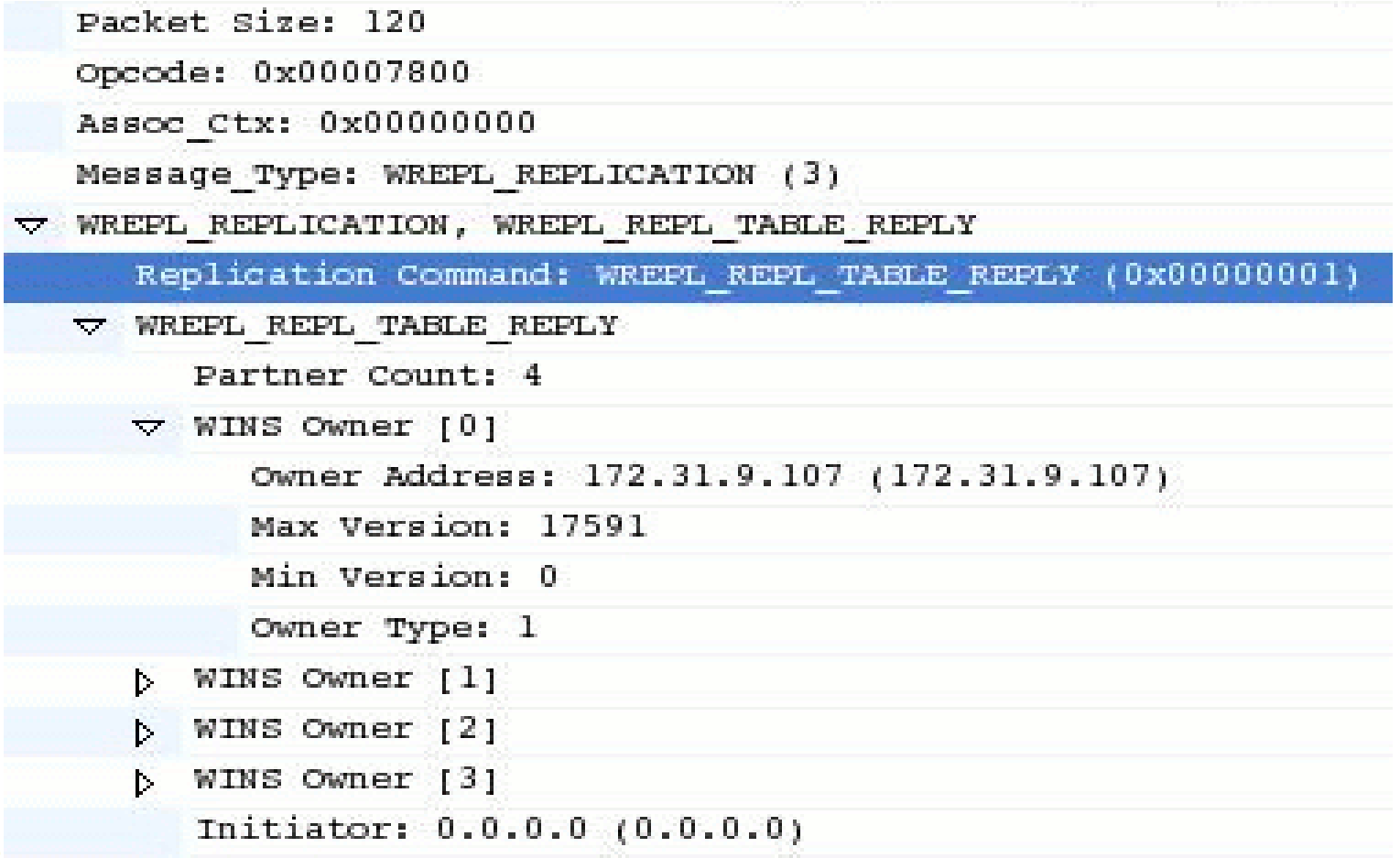
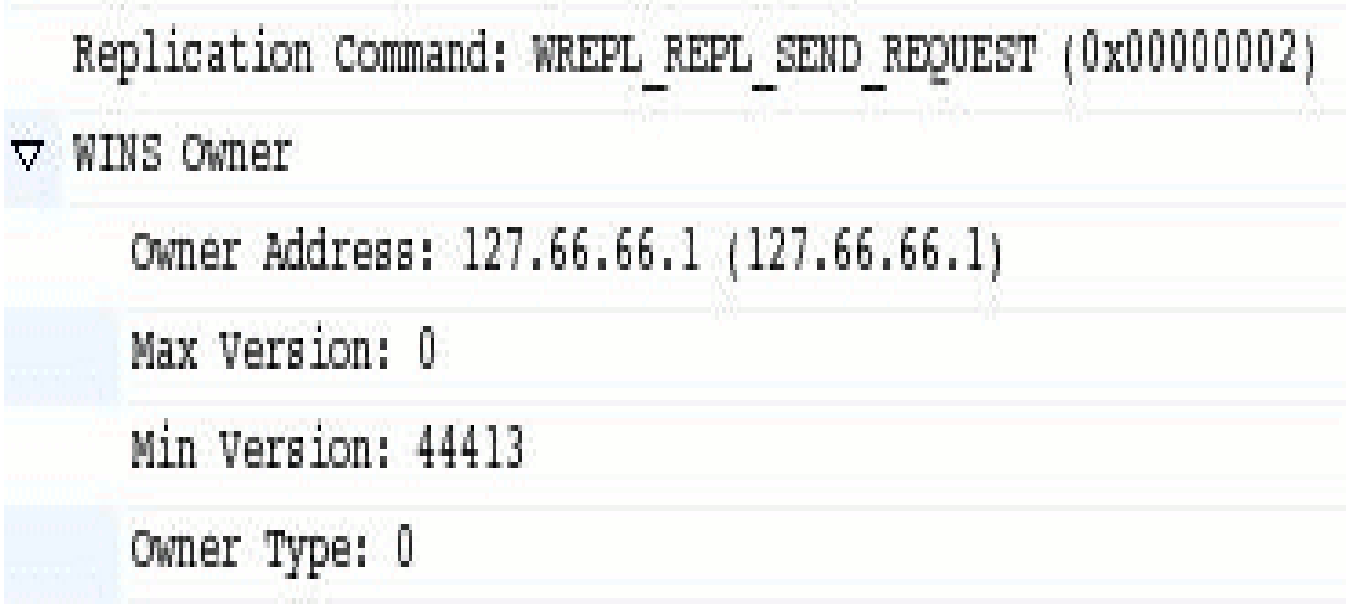


Table Query/Reply

- WREPL_REPL_TABLE_QUERY has no data
- WREPL_REPL_TABLE_REPLY



Send Request



Send Reply

```
Replication Command: WREPL_REPL_SEND_REPLY (0x00000003)
▽ WREPL_REPL_SEND_REPLY
    Num Names: 1
  ▽ WINS Name [0]: _SAME_OWNER_A<00>: 127.0.65.2
      Name Len: 17
    ▷ Name: _SAME_OWNER_A<00> (Workstation/Redirector)
    ▽ Name Flags: 0x00000000
        .... .... .... .... .... .... .... ..00 = Record Type: Unique (0x00000000)
        .... .... .... .... .... .... .... 00.. = Record State: Active (0x00000000)
        .... .... .... .... .... .... ...0 .... = Local: False
        .... .... .... .... .... .... .00. .... = Host Type: B-node (0x00000000)
        .... .... .... .... .... .... 0... .... = Static: False
      Name Group Flag: 0x00000000
      Name Version Id: 31087
      IP Address: 127.0.65.2 (127.0.65.2)
      Unknown IP: 255.255.255.255 (255.255.255.255)
```

Update(2)/Inform(2)

```
Replication Command: WREPL_REPL_UPDATE (0x00000004)
▽ WREPL_REPL_TABLE_REPLY
    Partner Count: 1
  ▽ WINS Owner [0]
      Owner Address: 127.66.66.1 (127.66.66.1)
      Max Version: 44413
      Min Version: 0
      Owner Type: 1
    Initiator: 0.0.0.0 (0.0.0.0)
```

Exploring the conflict resolving

- The replication uses a Multi-Master-Topology
- Each Server is able to handle registration/releases
- Conflicts can happen when different machines register the same name, and this conflicts maybe detected during replication
- This conflicts need to be resolved to have consistent information on all servers
- There's some information about the resolving in the Microsoft documentation, but it's unclear on some important corner cases
- I created about 300 tests in samba4's smbtorture, to demonstrate the resolving algorithm

Server Implementation

- I was trying to finish the half working “wrepld” from samba3 a few year ago, but I fail...
- As I already was quite familiar with the samba4 architecture and a wrote a lot of that stuff myself
- This time I tried it using the samba4 framework where the hardest part (asynchronous packet handling) was already solved in a generic and relatively easy to use way
- The development directly went into the main SAMBA_4_0 subversion-tree
- I step by step tried to get all torture tests working against the server until everything was working fine

A WINS-Partner in wins_config.ldb

dn: CN=WINSSERVER-02,CN=PARTNERS
objectClass: wreplPartner
address: 192.168.9.9
type: 0x3
pullInterval: 1800
pullRetryInterval: 30
pushChangeCount: 0
pushUseInform: 0

A Name-Record in wins.ldb

dn: name=DOMAIN1,type=0x1C
objectClass: winsRecord
type: 0x1C
name: DOMAIN1
recordType: 2
recordState: 0
nodeType: 0
isStatic: 0
expireTime: 20060110201752.0Z
winsOwner: 192.168.9.9
versionID: 55654
address: 192.168.9.56;winsOwner:192.168.9.9;expireTime:20060110201752.0Z;
address: 192.168.4.66;winsOwner:192.168.4.9;expireTime:20060103135509.0Z;

More Details

- Fetch the Samba4 source code from svn
  - see http://devel.samba.org/
- And take a look at the related code in:
  - source/wrepl_server/
  - source/libcli/winsrepl/
  - source/nbt_server/wins/
  - source/torture/nbt/

samba4wins

- samba4wins is a SAMBA_4_0 snapshot with most features disabled
- only the NBT and WINSREPL services are active
- It can run in parallel with samba-3.0.21 or higher

- SerNet started this project because of many requests from corporate and public customers in late 2005
- Sponsors of the initial project are Computacenter, Fujitsu Siemens Computers (FSC) and the Linux Solutions Group (LiSoG e.V.)
- http://www.enterprisesamba.org/index.php?id=88

More info...

- Are there any Questions?
- Many thanks for your attention!