Leibniz-Rechenzentrum
der Bayerischen Akademie der Wissenschaften

E-Mailversand beim BVB

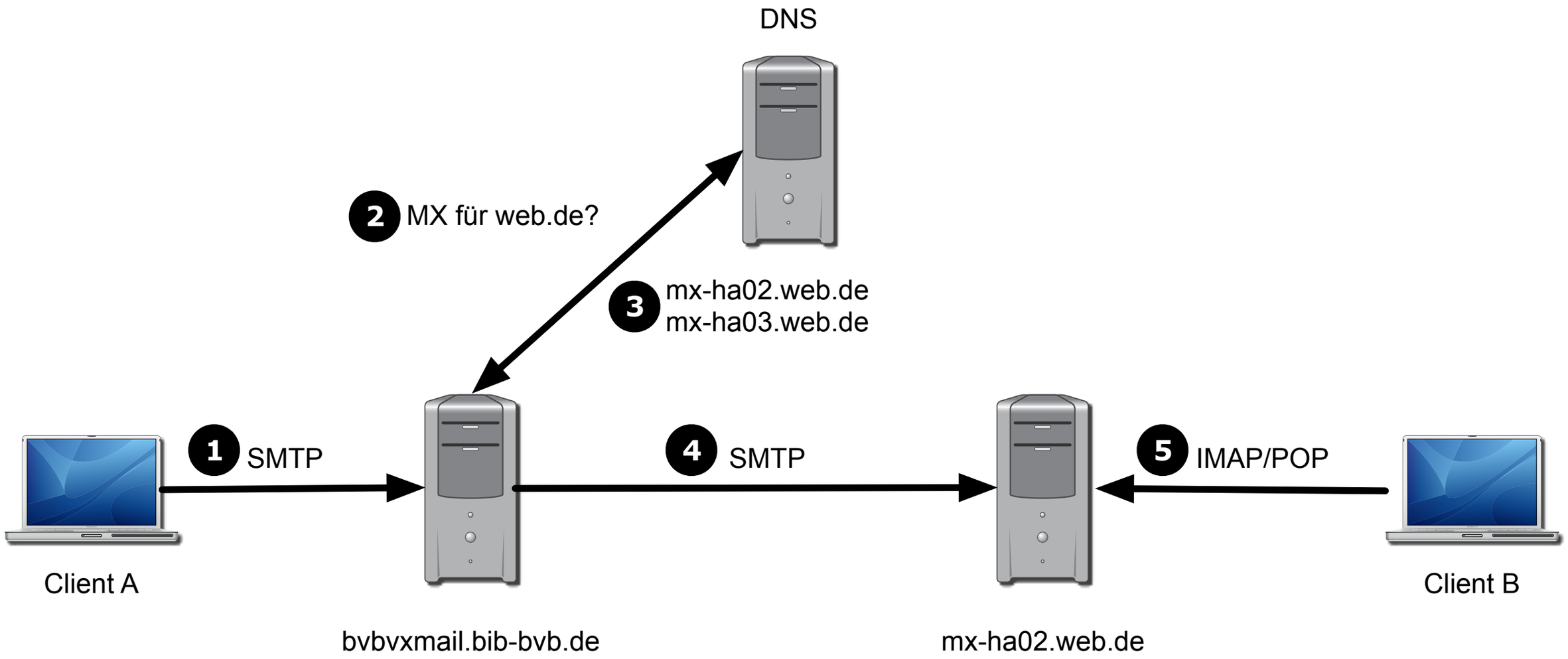
Bernhard Lichtinger

Vorbemerkungen

- E-Mail ist schon sehr alt, der RFC822 stammt von 1982
- Damals war das Internet noch freundlich und alles in ASCII
- Immer wieder erweitert durch Multipurpose Internet Mail Extensions (MIME):
  - Andere Zeichensätze als ASCII
  - Anhängen von Dateien
  - Übertragung von 8bit-Zeichen statt 7bit-Zeichen
  - HTML neben reinem Text für den Inhalt erlaubt
  - u.v.m.

Überblick E-Mail

- Vereinfachtes Beispiel: E-Mail von A@bib-bvb.de an B@web.de



0: E-Mail erstellen

- Analogie „Brief schreiben“
- Briefkopf = header
  - Absender = From:
  - Empfänger = To:
  - Betreff = Subject:
- Inhalt des Briefes = body
  - Typ = Content-Type
  - Evtl. mehrere Teile = multipart
- Anhänge = attachment

```
From: A@bib-bvb.de
To: B@web.de
Subject: Demomail
```

```
Content-Type: text/plain; charset="UTF-8"
Content-Type: multipart/mixed;
Content-Disposition: attachment;
  filename=„a.txt";
```

0: Beispiel

root <root@bvbxma2.bib-bvb.de> Gestern 11:46

An: mon-lokalsys@bib-bvb.de

SEC: unbekannte Meldung von ueiov2

```
Aug  5 11:36:06 ueiov2 ifup:     eth0      device: VMware VMXNET3 Ethernet Controller (rev 01)
Aug  5 11:36:07 ueiov2 SuSEfirewall2: SuSEfirewall2 not active
```

```
Date: Wed, 5 Aug 2015 11:46:07 +0200
To: <mon-lokalsys@bib-bvb.de>
Subject: SEC: unbekannte Meldung von ueiov2
User-Agent: Heirloom mailx 12.5 7/5/10
Content-Type: text/plain; charset="us-ascii"
Content-Transfer-Encoding: 7bit
Message-ID: <20150805094607.3A05C94283@bvbxma2.bib-bvb.de>
From: root <root@bvbxma2.bib-bvb.de>
MIME-Version: 1.0

Aug  5 11:36:06 ueiov2 ifup:     eth0      device: VMware VMXNET3 Ethernet
Controller (rev 01)
Aug  5 11:36:07 ueiov2 SuSEfirewall2: SuSEfirewall2 not active
```

1: E-Mail abschicken

- „In Umschlag stecken, aber nicht zukleben“
- Client überträgt per Simple Mail Transfer Protocol (SMTP) die E-Mail an den Mailserver:

```
220 bvbvxmail.bib-bvb.de ESMTP Postfix
HELO bvbxma2.bib-bvb.de
250 bvbvxmail.bib-bvb.de
MAIL FROM: <test@bvbxma2.bib-bvb.de>
250 2.1.0 Ok
RCPT TO: <lichtinger@lrz.de>
250 2.1.5 Ok
DATA
354 End data with <CR><LF>.<CR><LF>
Subject: Testemail
From: <test@bvbxma2.bib-bvb.de>
To: <lichtinger@lrz.de>
Emailtext

.
250 2.0.0 Ok: queued as 9305154F54B
quit
221 2.0.0 Bye
```

envelope from

envelope to

email header

email body

2+3: Zielserver finden

- Zuerst wird im DNS nach mail exchange (MX) Einträgen angefragt, welche Mailserver für die Zieldomain zuständig sind
- Ein MX-Eintrag besteht aus einer Priorität und einem Hostnamen
- Es sind mehrere Einträge erlaubt
- Z.B. hat web.de 2 Einträge mit gleicher Priorität 100:

```
$ host -t mx web.de
web.de mail is handled by 100 mx-ha02.web.de.
web.de mail is handled by 100 mx-ha03.web.de.
```

4: E-Mail an Zielserver zustellen

- Genauso wie ein E-Mailclient sprechen die Mailserver per SMTP miteinander, einer als Client und der andere als Server
- Dabei werden die selben envelope from und to Adressen benutzt, wie zuvor schon der E-Mailclient benutzt hat (die from und to header werden ignoriert)
- WENN nicht der Mailserver aufgrund seiner Konfiguration Veränderungen vorgenommen hat:
  - Umschreiben der Domain: z.B. aus „lrz-muenchen.de“ wird „lrz.de“
  - Auswerten von Aliases: z.B. aus „lichtinger@lrz.de“ wird „bernhard.lichtinger@lrz.de“
  - Weiterleitungen: z.B. „root@bib-bvb.de“ an „bvb@lrz.de“
  - u.v.m.
- Zielserver speichert die E-Mail dann ab und stellt sie dem Client zur Abholung bereit

5: Client holt E-Mail ab

- Für den Abruf von E-Mails werden im wesentlichen 2 Protokolle benutzt:
  - Post Office Protocol v3 (POP3)
  - Internet Message Access Protocol (IMAP)
- IMAP ist die modernere Variante mit viel mehr Funktionen als POP3
- Groupwaresysteme wie MS Exchange nutzen primär proprietäre Protokolle, können aber (meist) auch IMAP und POP3

Bounces

- Kann ein Server eine E-Mail nicht zustellen, weil z.B. die Empfängeradresse falsch ist, dann schickt er eine „Bounce E-Mail“ an den **envelope from** (sender) zurück
- Diese E-Mail enthält dann (leider oft recht technisch formuliert) den Grund, warum die E-Mail nicht zugestellt werden konnte und als Anhang die ursprüngliche E-Mail
- Die häufigsten Gründe sind:
  - Empfängeradresse existiert nicht
  - Empfängermailbox ist voll
  - E-Mail bzw. Anhang ist zu groß für den Zielserver
  - Zielserver hält die E-Mail für Spam

Spam

- Ist ein großes Problem, da es bei SMTP möglich ist, den Absender zu fälschen und jeder Rechner im Internet E-Mail versenden kann
- Bei der Entstehung des Protokolls dachte keiner daran, dass es missbraucht werden könnte, der Fokus war auf „simple“
- Über die Jahre wurden diverse Gegenmassnahmen entwickelt
- Es gibt kein universelles Gegenmittel, dazu müsste man von heute auf morgen SMTP abschalten und auf ein neues Protokoll wechseln: utopisch, es wird immer eine große Menge Altsysteme geben, die nicht umgestellt werden können
- Wenn es je einen Nachfolger geben wird, wird es sicher ein so träger Umstieg wie von IPv4 auf IPv6

Anti-Spam

- 2 Klassen von Gegenmassnahmen:
- Während dem SMTP-Dialog:
  - Benötigt wenig Rechenleistung
  - Verhindert einen Großteil des Spams
  - Hat nur die envelope-Daten als Kriterien
  - Zielen meist darauf ab zu prüfen, ob der Absender legitim ist
- Nach der Annahme der E-Mail
  - Benötigt viel Rechenleistung
  - Hat die komplette E-Mail zur Untersuchung
  - In Deutschland darf die E-Mail nicht gelöscht werden und muss zugestellt werden
  - Untersuchen hauptsächlich den Inhalt auf Werbung bzw. Viren

Anti-Spam-Massnahmen 1

- Reverse record lookup:
  - Prüft, ob IP-Adresse, zugehöriger Hostname und Hostname im HELO zusammenpassen:

```
HELO bvbxma2.bib-bvb.de
bvbxma2.bib-bvb.de has address 193.174.96.25
25.96.174.193.in-addr.arpa domain name pointer bvbxma2.bib-bvb.de.
```

- Blacklists:
  - Listen von IP-Adressen, die als Spamversender bekannt sind
  - Diverse Anbieter mit diversen Kriterien, wann eine IP gelistet wird, mit teils sehr unterschiedlicher Qualität
- SMTP-Callout:
  - Der annehmende Mailserver tut so, als ob er eine E-Mail an die Absenderadresse zustellen wollte
  - Nur wenn der Mailserver die Absenderadresse als Empfänger akzeptiert, wird die E-Mail angenommen

Anti-Spam-Massnahmen 2

- Greylisting:
  - Pflegt eine Datenbank mit Tripeln: Absender- und Empfängeradresse und Client-IP
  - Versucht eine Client-IP zum ersten Mal eine E-Mail abzuliefern, wird dies zunächst mit einem temporären Fehler abgelehnt
  - Im Gegensatz zu den meisten Spam-Versendern versucht ein echter Mailserver mehrmals die E-Mail zuzustellen
  - Nach einer definierten Zeitspanne wird dann die E-Mail angenommen und in der Datenbank vermerkt
  - Bei nachfolgenden E-Mails von der gleichen IP wird dann sofort die E-Mail angenommen

Anti-Spam-Massnahmen 3

- DomainKeys (DKIM):
  - Der Mailserver signiert die ausgehenden E-Mails (header UND body)
  - Der öffentliche Schlüssel wird im DNS hinterlegt
  - Bei signierten E-Mails kann man sicher sein, dass sie authentisch sind
  - (fast) JEDE Veränderung an einer E-Mail führt zu einer invaliden Signatur
- Sender Policy Framework (SPF) lookup:
  - Beim SPF wird bei einer Domain im DNS hinterlegt, von welchen Servern E-Mails von dieser Domain versandt werden dürfen
  - Es wird nur der „envelope from“ geprüft, „from:“-header ist egal
  - Wenn zu streng eingestellt, können auch „gute“ E-Mails abgelehnt werden
  - Beispiel uni-bw:

```
$ host -t txt unibw.de
unibw.de descriptive text "v=spf1 mx a:bvbvxmail.bib-bvb.de -all"
```

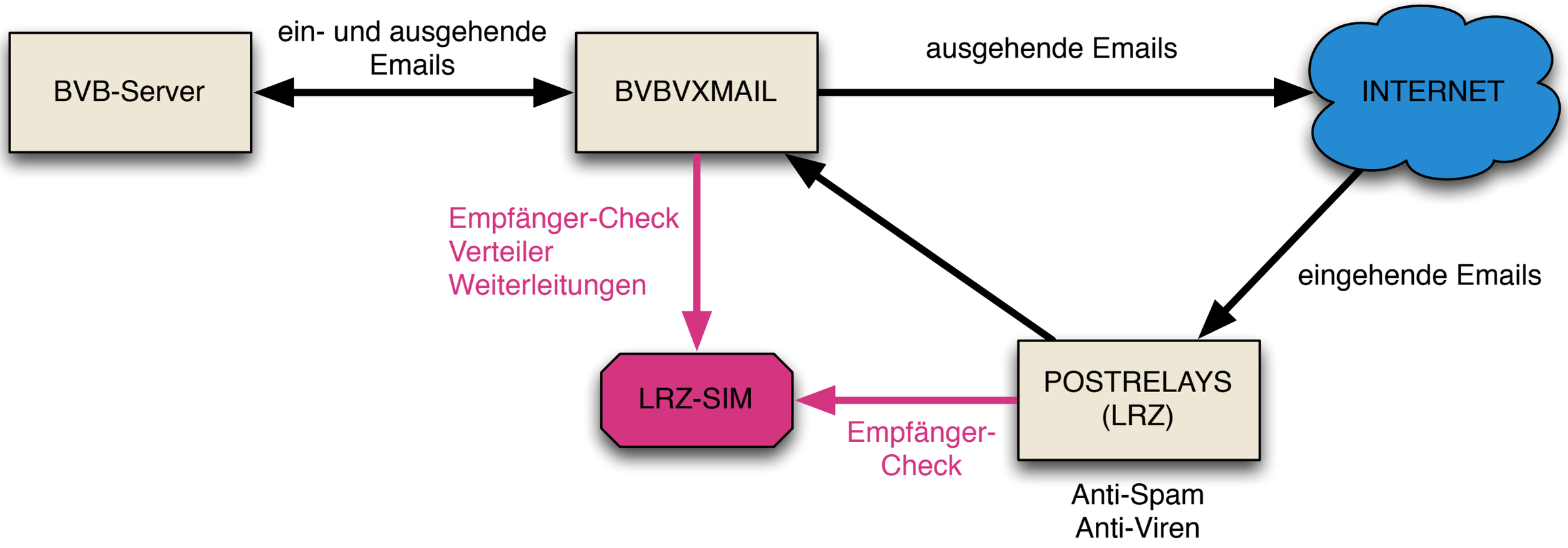
Anti-Spam-Massnahmen 4

- DMARC:
  - Setzt auf SPF und DKIM auf
  - Strengere Regeln: Auch die „from:“-header müssen passen
  - Admin einer Domain kann sich Reports schicken lassen, erleichtert die Fehlersuche und liefert Fremdserver, die die eigene Domain als Absender verwenden
  - Mailserver des Empfängers wertet aus, was Absender-Domain eingestellt hat
- Probleme mit DMARC:
  - Weiterleitungen und Mailinglisten funktionieren nicht mehr wie gewohnt, die E-Mails werden als „gefälscht“ eingestuft und abgelehnt
- Verbreitung von SPF und DMARC nimmt zu:
  - Seit Juni 2016: 1und1, gmx, web.de haben SPF-Prüfung eingehend aktiviert
  - Ab Juni 2016: gmail.com DMARC „scharf“ geschaltet für gmail.com, eingehend wird schon länger geprüft

Anti-Spam-Massnahmen 5

- Inhaltsanalyse nach Annahme der E-Mail
- Beispiel spamassassin:
  - Diverse Module vergeben Plus- und Minuspunkte
  - Beim Überschreiten eines definierten Punktestands wird die E-Mail als Spam eingestuft
  - Wortlisten
  - Reguläre Ausdrücke
  - Bayesfilter (muss mit guten und schlechten E-Mails trainiert werden: ham vs. spam)
  - u.v.m.
- Virenscanner

E-Mailversand im BVB



E-Mailversand im BVB

- LDAP-Server (LRZ-SIM):
  - Liste der gültigen E-Mailadressen
  - Weiterleitungen
  - Verteilerlisten mon-admin, mon-lokalsys, mon-nvs, …
  - VD17-Verteilerlisten
- postrelay.lrz.de:
  - MX für bib-bvb.de, vd17.de
  - Antispam und Virenschutz für eingehende E-Mails
- bvbvxmail.bib-bvb.de (aka mailhost.bib-bvb.de):
  - Zentraler Mailserver für ein- und ausgehende E-Mail
  - IMAP/POP3-Server
  - Mailinglisten mit mailman

Besonderheiten beim E-Mailversand im BVB

- bvbvxmail ändert den envelope from:
  - Aus `KENNUNG@HOST.bib-bvb.de` wird `bounce+KENNUNG_HOST@bib-bvb.de`
  - Der Teil zwischen „+“ und „@„ wird beim E-Mailtransport ignoriert, effektiv bleibt `bounce@bib-bvb.de` übrig
  - So erfüllen wir die Bedingung, dass die Absenderadresse zustellbar sein muss und wir nicht alle möglichen Adressen im LDAP pflegen müssen
  - Es bleibt trotzdem sichtbar, wer der eigentliche Absender war
  - Wenn Antwortemails an den absendenden Rechner zurückgehen sollen, können wir Ausnahmen definieren
  - Aus `root@HOST.bib-bvb.de` wird `root+HOST@bib-bvb.de`. Diese E-Mails landen in einer gemeinsamen Mailbox vom SV-Team

Besonderheiten beim E-Mailversand im BVB

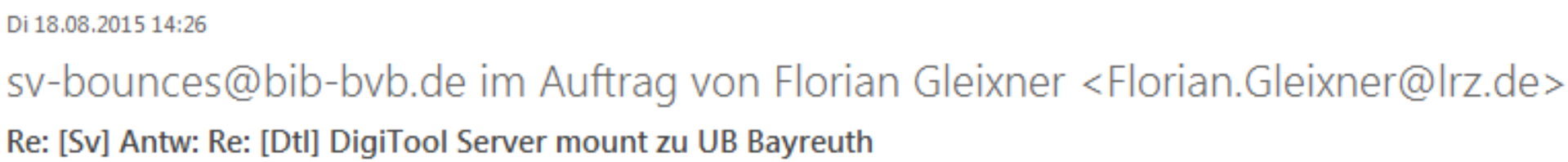
- Zum Unterdrücken von automatischen Antworten (Abwesenheitsbenachrichtigungen, etc.) fügt die bvbvxmail folgende Header hinzu:
  - precedence: bulk
  - Auto-Submitted: auto-generated
  - X-Auto-Response-Suppress: DR,RN,NRN,OOF,AutoReply
- Das funktioniert für (fast) alle Mailsysteme
- Wird nur gemacht, wenn die Absender-Domain bib-bvb.de ist

Besonderheiten bei SLES-VMs

- Im Auslieferungszustand werden alle Adressen (to und from!) umgeschrieben:
  - Alles von `@hostname` nach `root+hostname@srv.mwn.de`
  - `root+hostname@srv.mwn.de` wird an das SV-Team weitergeleitet
  - Das sorgt auch wieder für zustellbare Absenderadressen
  - Auch hier können wir Ausnahmen definieren

mailman

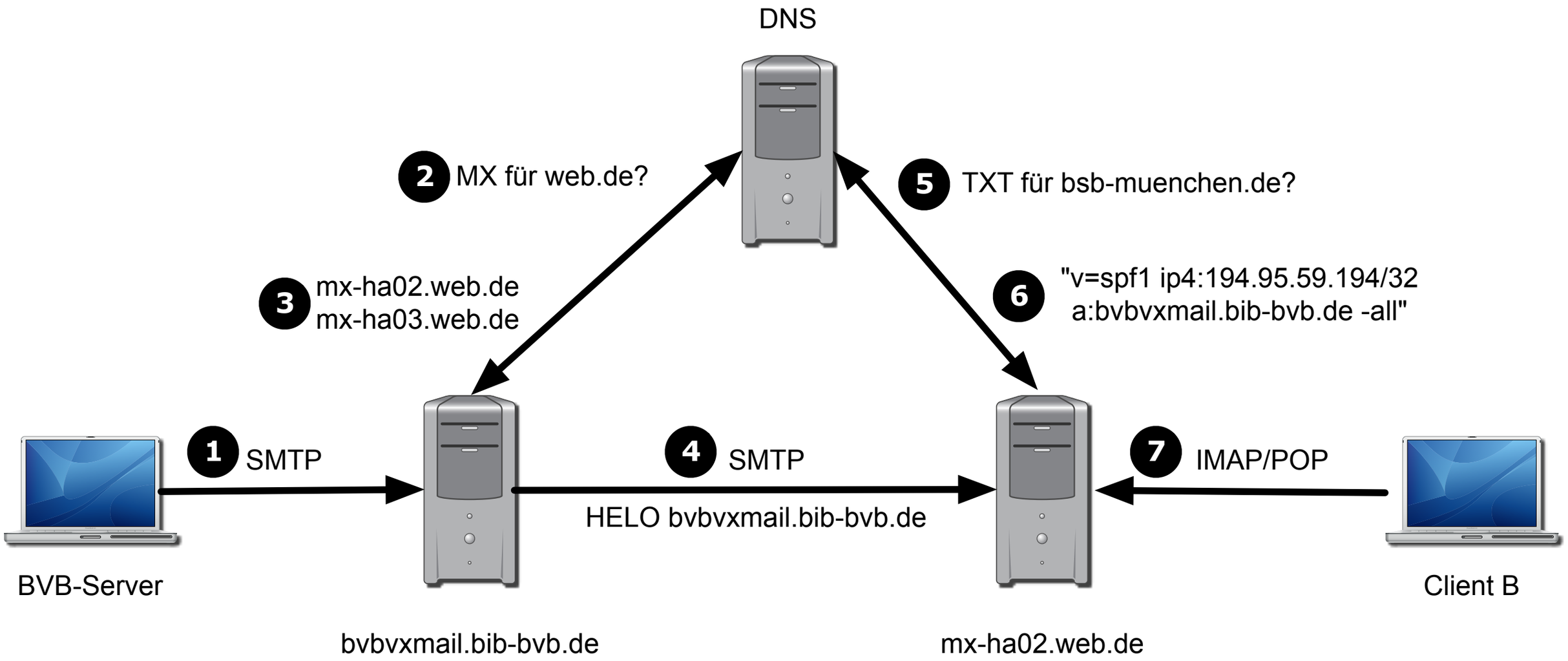
- Der mailman benutzt spezielle envelope Absenderadressen der Form LISTE-bounces@bib-bvb.de
- Somit landen Rückläufer (bounces) beim mailman und nicht beim eigentlichen Absender der E-Mail
- Erzeugt eine Empfängeradresse zu viele bounces, wird sie vom mailman automatisch deaktiviert
- mailman setzt auch den „sender:“ header, den manche E-Mailclients als „gesendet im Auftrag von“ anzeigen:



- Manche Benutzer klicken beim Antworten auf den falschen Knopf und schreiben dann z.B. an sv-bounces@bib-bvb.de anstatt an sv@bib-bvb.de. „Antworten“ bzw. „reply“ sollte immer das richtige tun

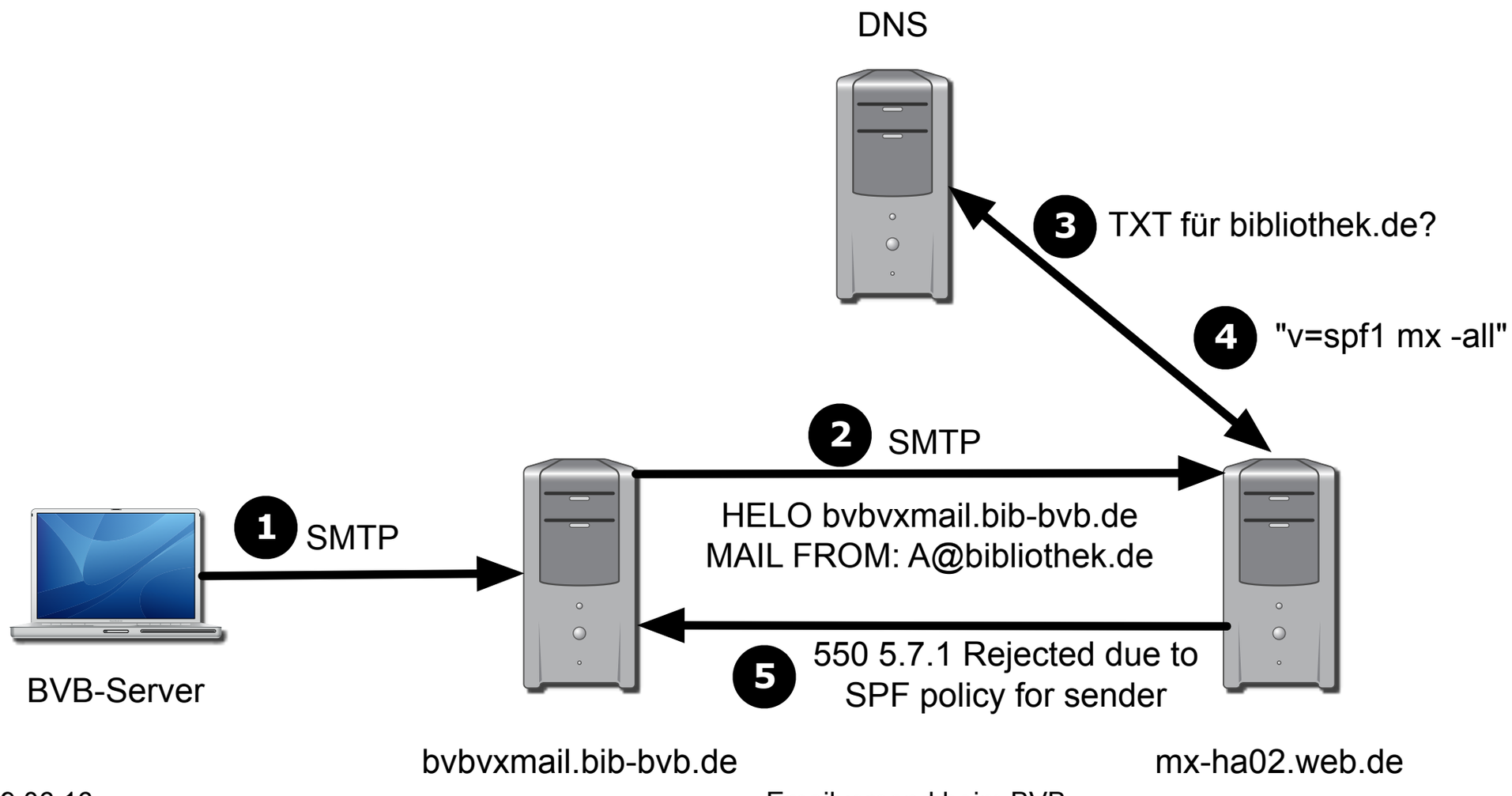
Versand E-Mail und SPF

- Beispiel: E-Mail von BVB-Server mit Absender A@bsb-muenchen.de an B@web.de



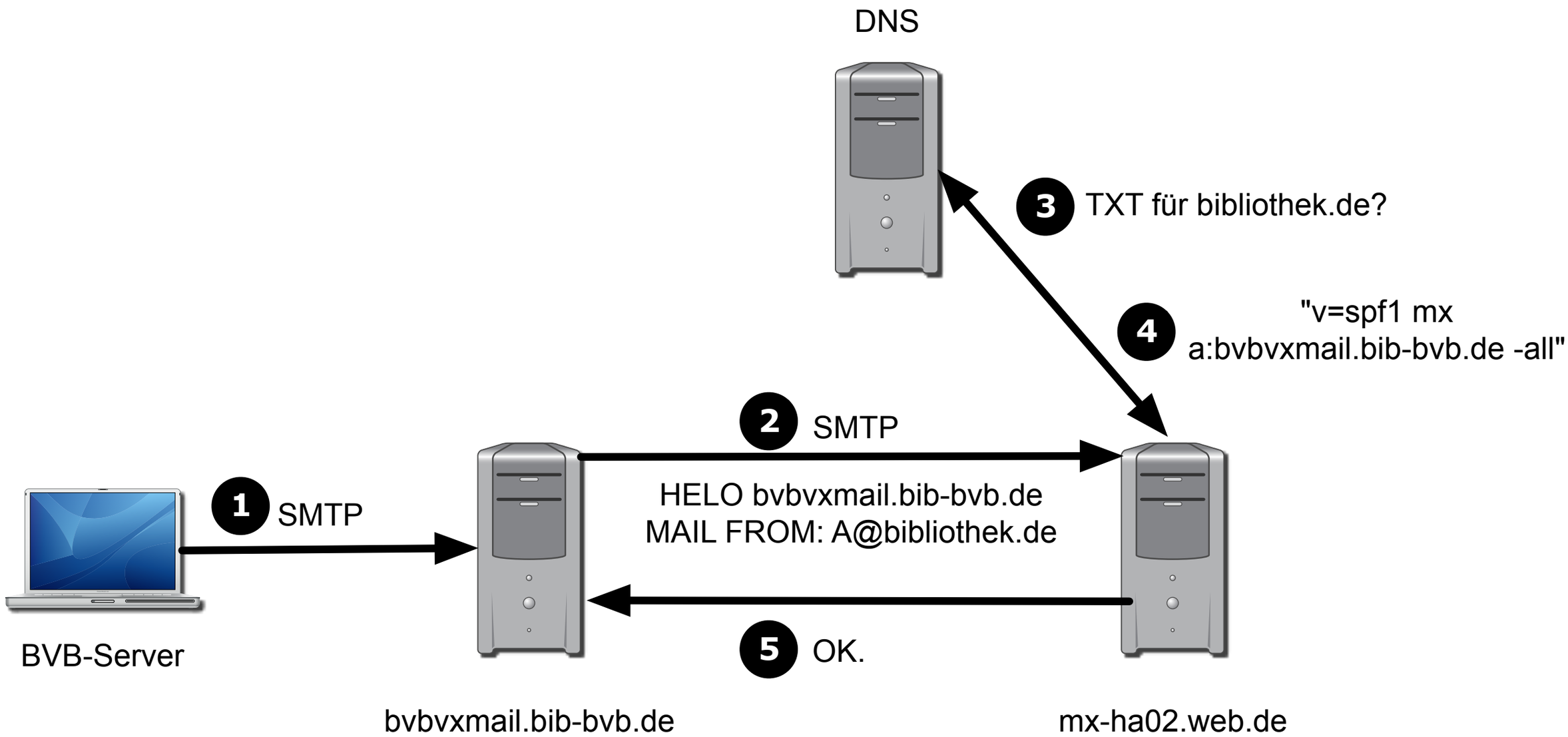
Probleme mit SPF, DMARC & Co

- Problem:
  BVB-Rechner versenden E-Mails mit der Absender-Domain der einzelnen Bibliotheken. Sobald die Bibliotheken SPF, DKIM und DMARC für ihre Domain einschalten, kann es passieren, dass die E-Mails von den BVB-Rechnern nicht mehr ankommen
- Beispiel: Absender A@bibliothek.de an B@web.de



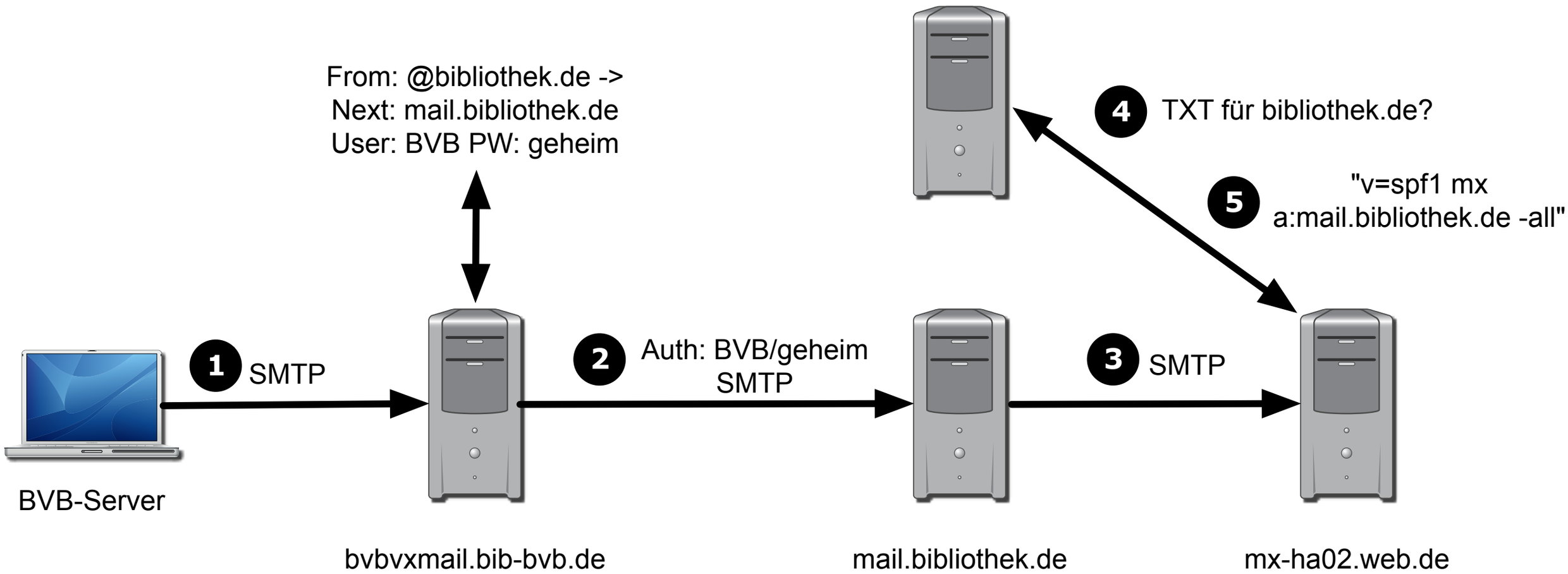
Lösung für SPF

- Die Bibliothek trägt auch die bvbvxmail.bib-bvb.de als erlaubten Absende-Server ein: `"v=spf1 mx a:bvbvxmail.bib-bvb.de -all"`



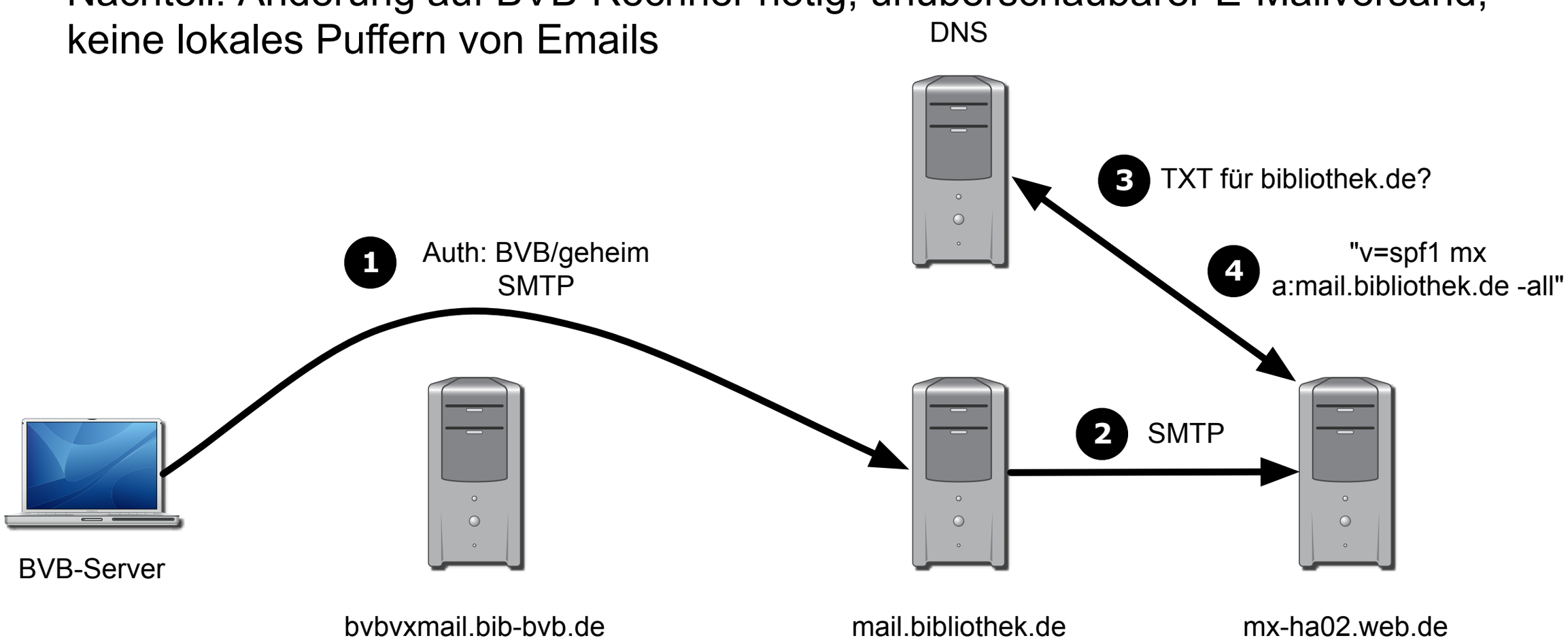
Saubere Lösung bei DKIM/DMARC

- bvbvxmail.bib-bvb.de versendet die E-Mails via Mailserver der jeweiligen Bibliothek
  - Benötigt „Freischaltung“ via IP-Adresse, Benutzer/Passwort oder Zertifikat
  - Vorteil: keine Änderung auf BVB-Rechnern nötig, zentrale Stelle für E-Mailversand



Unschöne Lösung

- BVB-Rechner selbst versendet via Mailserver der jeweiligen Bibliothek (Solaris: msmtp, Linux: postfix)
  - Benötigt „Freischaltung“ via IP-Adresse, Benutzer/Passwort oder Zertifikat
  - Benötigt Freischaltung in den Firewalls für SMTP-Verbindung (Port 25 bzw. 587)
  - Nachteil: Änderung auf BVB-Rechner nötig, unüberschaubarer E-Mailversand, keine lokales Puffern von Emails



Fazit

- Früher war E-Mail einfach
- Heute ist es eine eigene Wissenschaft
- Als E-Mailversender muss man sich an die Gepflogenheiten halten, damit man nicht als Spamversender gilt
- SPF und DMARC muss man vorsichtig einschalten, sonst kann man schnell E-Mails verlieren, wenn zu streng eingestellt bzw. absendende Server vergessen werden
- SPF: „`-all`“ nur dann einstellen, wenn man sicher ist, ALLE legitimen Absender-Server eingetragen zu haben. Besser nur „`~all`“, dann werden E-Mails schlimmstenfalls nur als Spam markiert

Nicht vergessen: Webformulare

- Oft gibt es Webseiten, über die man E-Mails versenden kann
  Bsp: Trefferliste aus dem OPAC verschicken
- Diese trifft diese ganze Problematik mit SPF/DMARC auch!

Zum Nachlesen

- Syntax der SPF-Einträge:
  http://www.openspf.org/SPF_Record_Syntax
- Überblick über DMARC:
  https://dmarc.org/overview/
- Mehr Informationen zu DKIM:
  http://www.dkim.org