

# 電子メールにおける送信元認証技術 の正しい利用と誤用の分析

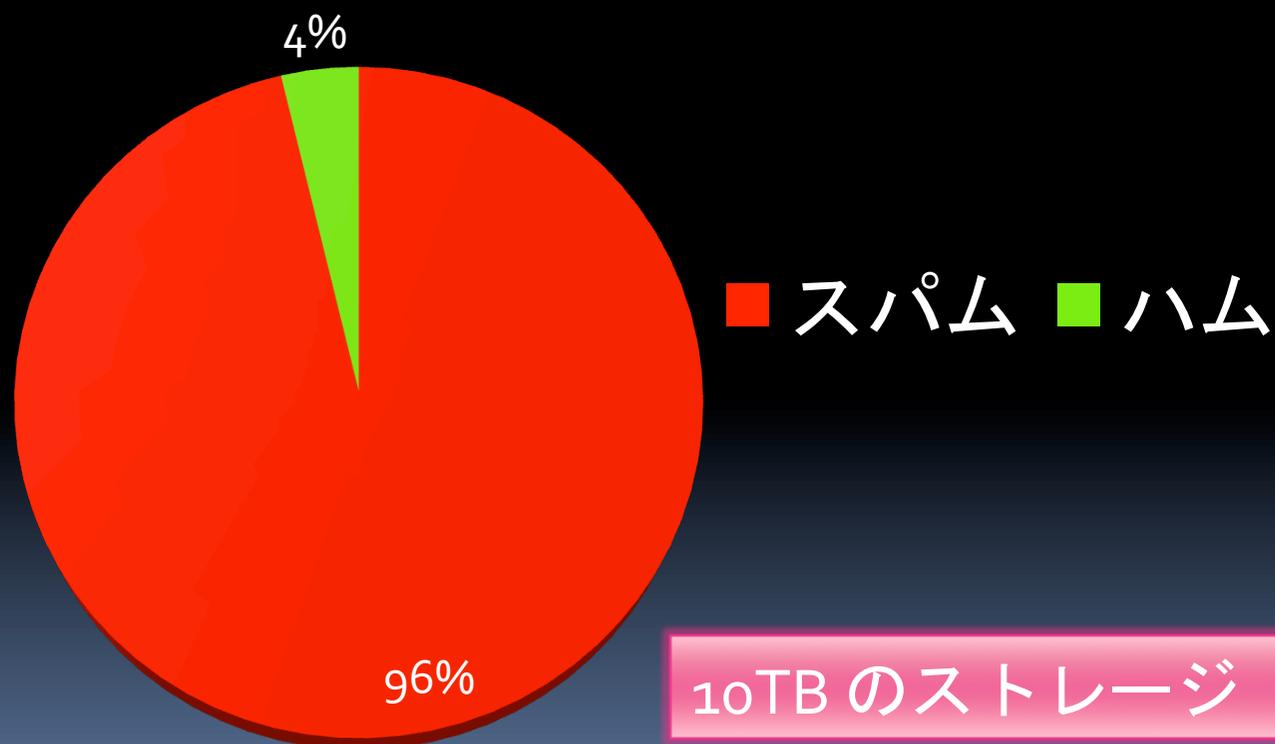
ICSS 2010年7月研究会

NTTサービスインテグレーション基盤研究所

森 達哉

# ある企業メールサーバの統計

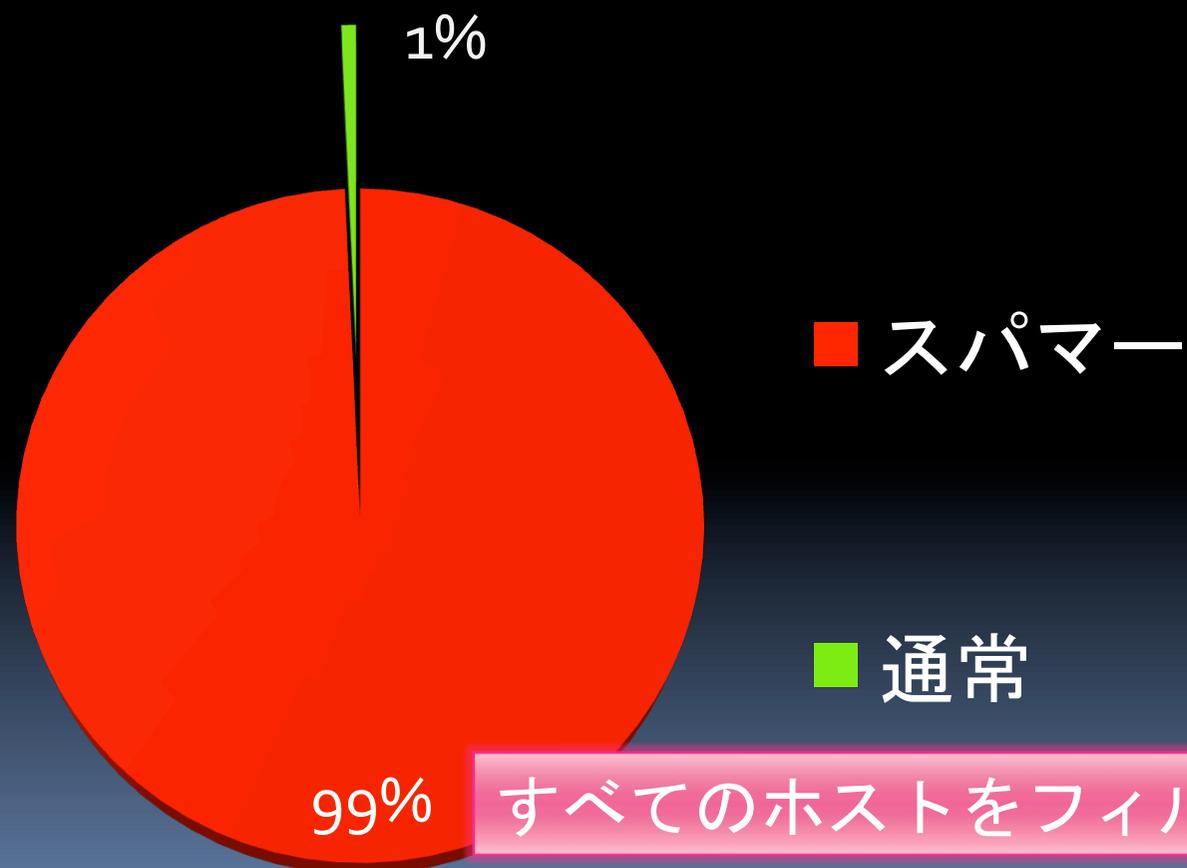
受信メッセージ: 約1800万通/1ヶ月の内訳



10TB のストレージ → 9.6TB が無駄!

# ある企業メールサーバの統計

メール送信ホスト: 約200万ホストの内訳



すべてのホストをフィルタ → 99%正解!!

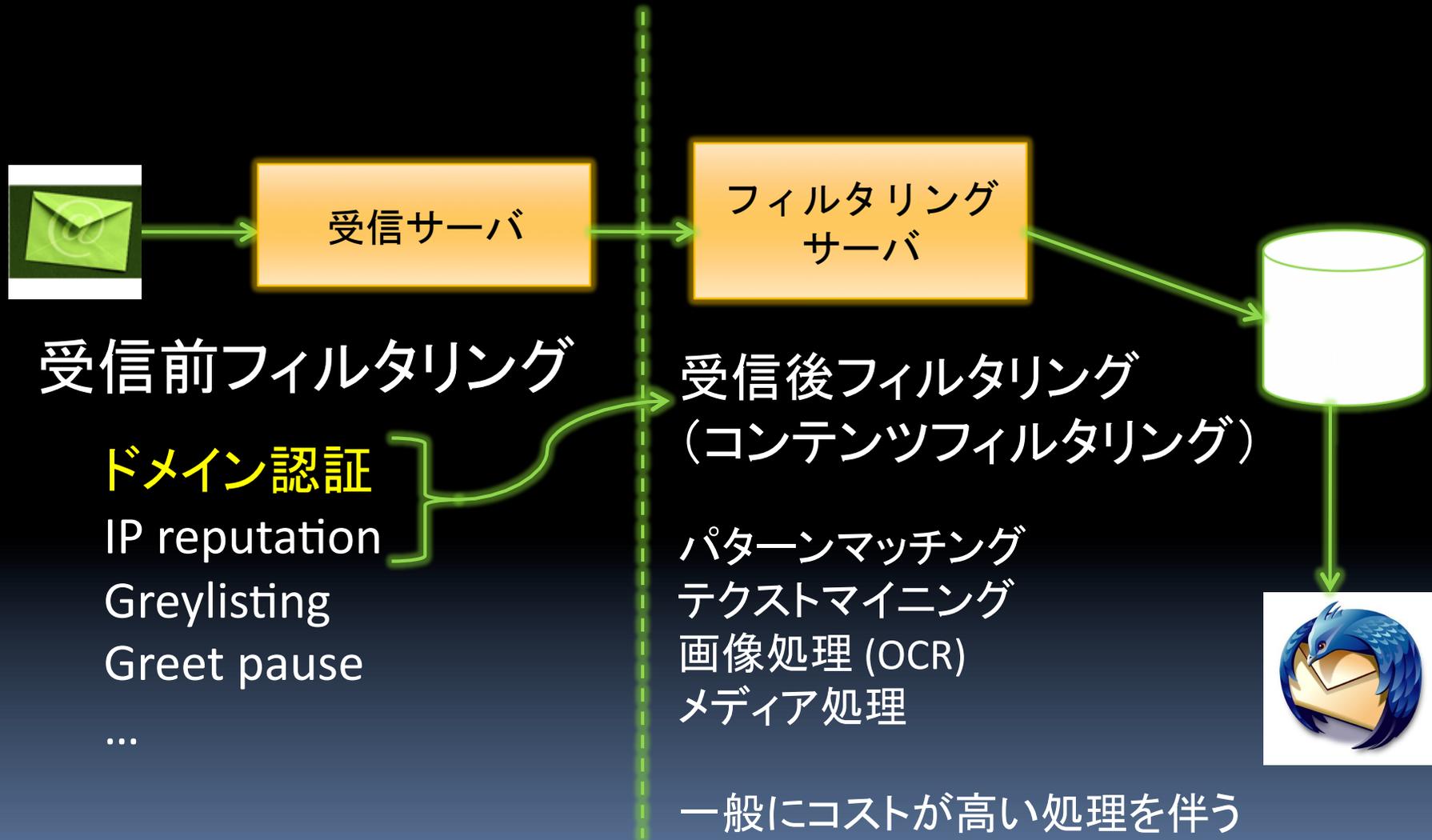
# 大多数のスパムは送信元を詐称

- ドメインを詐称
  - 「良い」ドメイン発のメールに見せかける
  - ホワइटリストにひっかける
  - 送信先のドメインを詐称
- ユーザアカウントを詐称
  - 送り先のドメイン上の適当なアカウント
  - スпам送信先リストからランダムに抽出

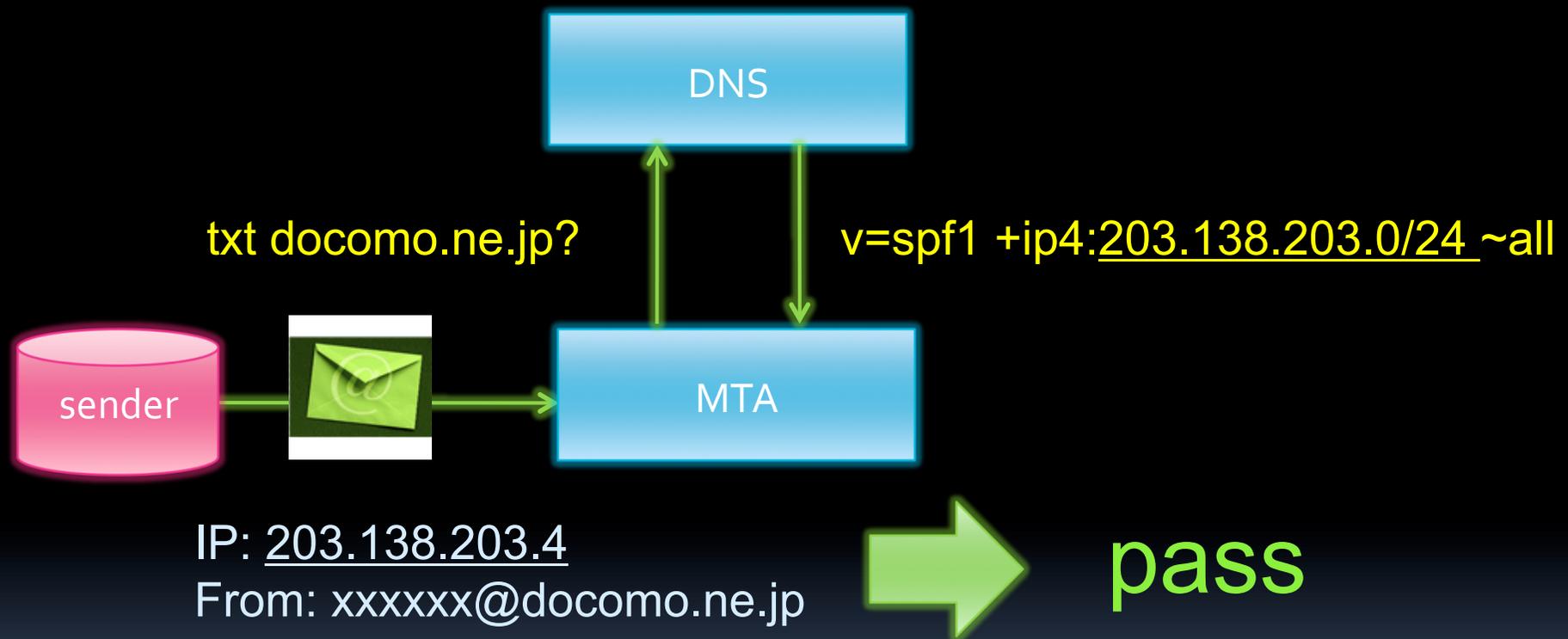
# 電子メール送信元認証技術

- 送信元認証技術
  - 1997年～ proposed by Paul Vixel
  - 2006年頃より本格的な普及開始
  - いずれも DNS をインタフェースとして使う
    - 送信元ドメインを使う為
- 代表的な電子メール送信元認証技術
  - SPF (Sender Policy Framework), Sender ID
  - DKIM (DomainKeys Identified Mail), DomainKeys

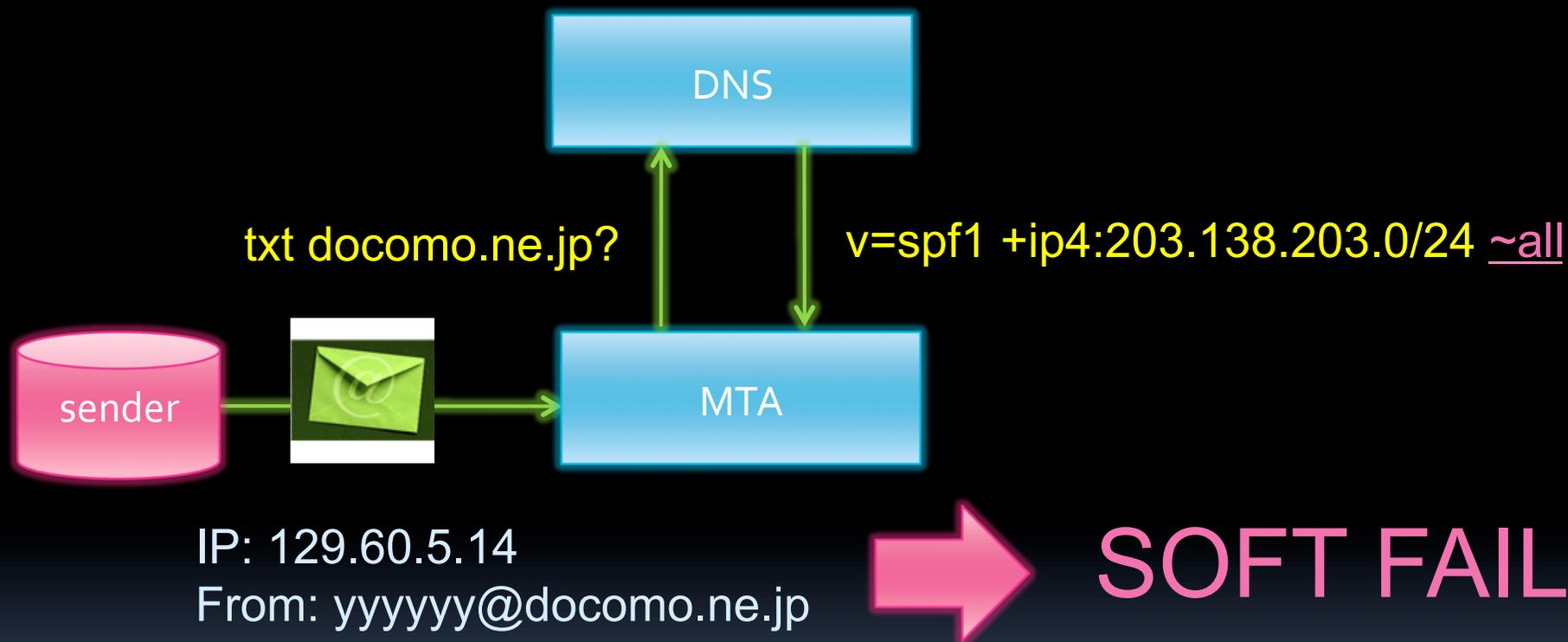
# 他の anti-spam 技術との関係



# SPF のメカニズム



# SPF のメカニズム



# ドメイン管理者が設定するポリシー (qualifiers)

- **+: PASS**
- **?: NEUTRAL (no policy).**
- **~: SOFTFAIL (typically accepted but tagged)**
- **-: FAIL (should be rejected)**
  
- **例:**  
ibm.com → v=spf1 **-all**  
すべてのあらゆる送信元を REJECT してもOK
  - xxxx@ibm.com はすべて invalid なアドレス
    - xxxx@us.ibm.com などは valid

# SPFの特徴

- 完全な分散アーキテクチャ
  - CA やThird Trusted Party は存在しない
  - Spammer も SPF を使うことが可能
- ポリシーの柔軟性・自由度の高さ
  - ポリシーは sender (domain) が決める。  
ただし receiver は設定されたポリシーをどのように解釈するかという受信側の立場としてのポリシーを決める必要がある
  - 受信者として妥当なポリシーを定めるには、他の送信元がどのようなことをしているかを把握することが必要不可欠

# 本研究の狙いと課題

- 送信者（ドメイン管理者）あるいは受信者の立場でSPFを正しく運用するために必要な統計を得る。 Best current practice
- SPFの普及と利用状況の理解（サンプル）
- SPFの正しい利用と誤用の把握
- SPFの問題点に関する定量的理解

# 分析に用いるデータ

- ドメインリスト
  - 良性ドメイン・悪性ドメイン
  - 既存DB (Alexa top500, DBL), マニュアル収集
- 電子メール配送ログ
  - SPF 認証結果とメッセージスコアの相関 (スパム or ハム)
- IPレピュテーションリスト
  - 良性IPアドレスリスト・悪性IPアドレスリスト
  - DNSBL (spamhaus), DNSWL

# Alexa Top 500 sites (良性ドメイン)

The screenshot shows the Alexa website interface. At the top, the Alexa logo is followed by the text 'The Web Information Company'. Navigation links include 'Home', 'Top Sites' (which is highlighted), 'Site Info', 'What's Hot', 'Toolbar', and 'For Site Owners'. Below the navigation is a search bar with a 'Search' button. Underneath are tabs for 'Global', 'By Country', and 'By Category'. The main content area is titled 'Top Sites' with a globe icon and the subtitle 'The top 500 sites on the web. ?'. A list of the top 4 sites is shown:

- 1 Google**  
google.com  
Enables users to search the Web, Usenet, and images. Features include PageRank, caching and tra... More  
★★★★★ Search Analytics ▶ Audience ▶
- 2 Facebook**  
facebook.com  
A social utility that connects people, to keep up with friends, upload photos, share links and ... More  
★★★★★ Search Analytics ▶ Audience ▶
- 3 YouTube - Broadcast yourself**  
youtube.com  
YouTube is a way to get your videos to the people who matter to you. Upload, tag and share your... More  
★★★★★ Search Analytics ▶ Audience ▶
- 4 Yahoo!**  
yahoo.com  
Personalized content and search options. Chatrooms, free e-mail, clubs, and pager.  
★★★★★ Search Analytics ▶ Audience ▶

# SPAMHAUS DBL (悪性ドメイン)



The screenshot shows the Spamhaus website interface. At the top, the Spamhaus logo is on the left and 'THE' logo is on the right. A navigation bar contains links for Home, SBL, XBL, PBL, DBL (highlighted), DROP, and ROKSO. Below the navigation bar is a yellow banner with 'Blocklist Removal Center' on the left and 'About Sp' on the right. The main content area features a large 'DBL Advisory' header. Below this is the title 'The Domain Block List'. The text explains that the Spamhaus DBL is a realtime database of domains found in spam messages, used by mail server software to identify, classify, or reject spam. It also mentions that the DBL is queriable in realtime by mail systems throughout the Internet. A sidebar on the left contains sections for 'Blocklist Help' and 'Associated Documents', each with a list of links.

**SPAMHAUS** 

Home SBL XBL PBL **DBL** DROP ROKSO

Blocklist Removal Center About Sp

## DBL Advisory

### The Domain Block List

The Spamhaus DBL is a realtime database of domains (typically web site domains) found in spam messages. Mail server software capable of scanning email message body contents for URIs can use the DBL to identify, classify or reject spam containing DBL-listed domains.

The DBL is queriable in realtime by mail systems throughout the Internet, allowing mail server administrators to identify, tag or block incoming email containing domains which Spamhaus deems to be involved in the sending, hosting or origination of Unsolicited Bulk Email (aka "Spam"). The DBL database is maintained by a dedicated team of specialists working with an automated system that constantly analyses a large portion of the world's email flow and the domains used in spam emails.

The DBL is both a domain URI Blocklist and RHSBL. It is intended primarily for message body URI checks but it can additionally be used for connection checks at the SMTP level and header domain

**Blocklist Help**

Blocked? To check, get info and resolve listings go to  
▶ Blocklist Removal Center

**Associated Documents**

- ▶ DBL FAQs
- ▶ DNSBL Usage Terms
- ▶ How Blocklists Work
- ▶ Datafeed Service

# Alexa Top 500 sites に対する SPF普及率の分析

Lars » Eggert

[About](#) [Contact](#) [CV](#) [Papers](#) [Talks](#) [Students](#) [Software](#) [Other](#)

## Global SPF Deployment

**53.5%**

493 sites tested  
0 DNS errors  
264 with SPF

### China

**23.4%**

483 sites tested  
1 DNS error  
113 with SPF

### Germany

**41.9%**

493 sites tested  
1 DNS error  
206 with SPF

### Finland

**34.8%**

480 sites tested  
0 DNS errors  
167 with SPF

### India

**47.3%**

488 sites tested

### IETF\*

**28.1%**

2470 sites tested

### South Korea

**57.9%**

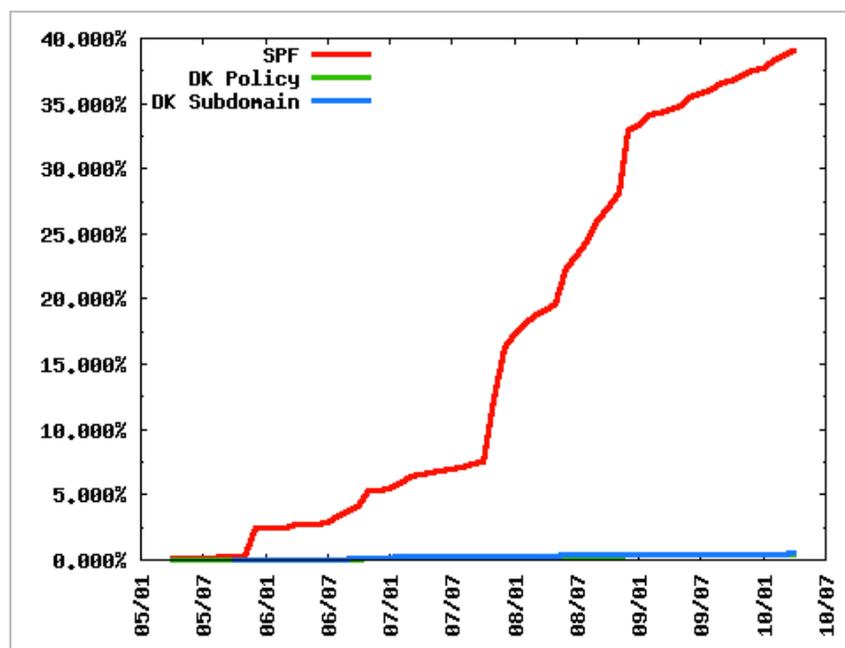
480 sites tested

# .jp における SPF 普及率の分析

## ドメイン認証の普及率に対する測定結果

WIDE antispam WG

WIDE プロジェクトは、JPRS と共同研究契約を結び、2005年4月からドメイン認証の普及率を毎月測定している。JPRS からは、jp 以下のドメインの一覧を提供して頂き、それらのドメインに対して 認証情報の有無を検査している。



2010年4月現在、JPドメインにおけるドメイン認証技術の おおよその普及率：

# 分析項目

- グローバルな視点
  - SPFの**詳細な**普及状況
    - 良性、悪性別のカウント, sender ID, IPv6, parser
  - SPFで登録されている prefix 長
    - /8 以下の大きなアドレス空間を登録しているのは怪しいケースが多い→定量的に分析
- ローカルな受信側の視点
  - SPFで認証されたメッセージの統計
  - SPFで認証された送信者の統計

# SPF の普及状況

Domain set	<i>N</i>	<i>S</i> (%)
Legitimate domains		
ALEXA	500	289 (58%)
LGD	1,675	754 (45%)
Spamming domains		
DBL	66,356	24,522 (37%)
LBD	45,123	9,590 (21%)

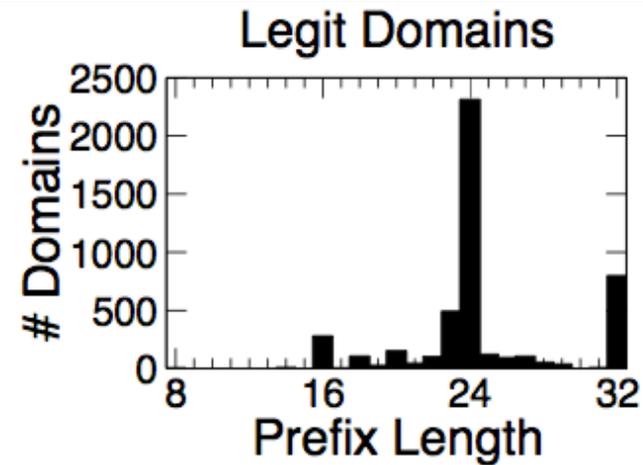
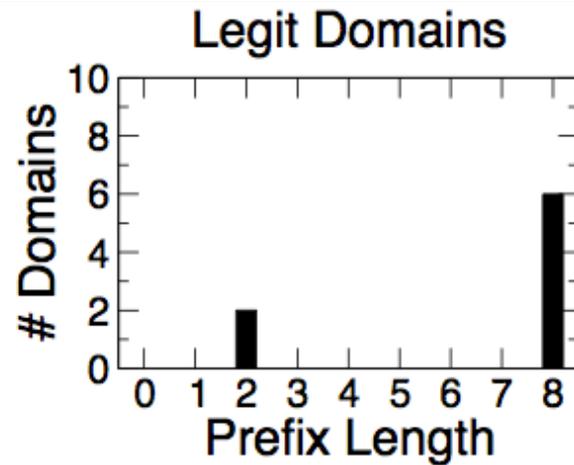
良性ドメイン 約50% の普及

悪性ドメイン 約30% の普及

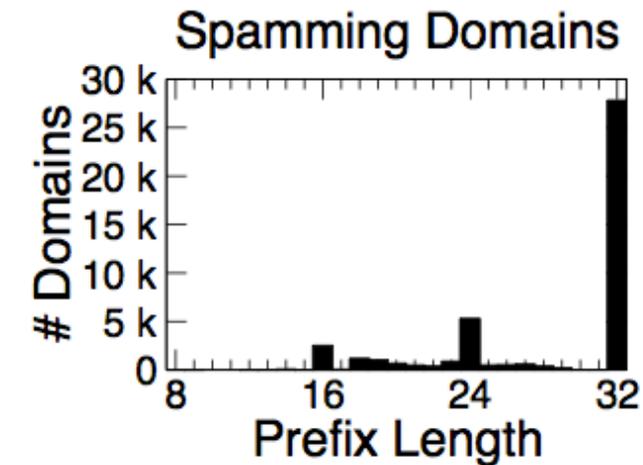
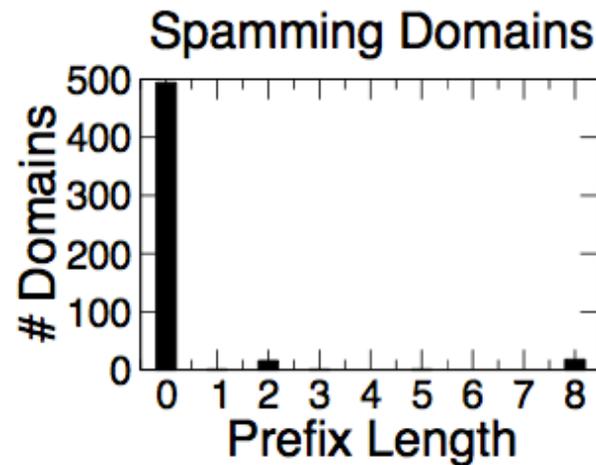
Domain set	SPF	Sender ID	both	IPv6
Legitimate domains				
ALEXA	289	31	31	1
LGD	754	33	33	9
Spamming domains				
DBL	24,439	801	798	0
LBD	9,560	190	170	2

# SPFに登録された prefix 長分布

良性ドメイン



悪性ドメイン



# SPFに登録された prefix 長分布

- Spammer は /0 を使う
  - 任意のIPアドレスから送ってもOK
    - dynamic IP を持つボットから送信できる
  - SPF 認証チェックは通過
- Spammer は /32 も使う
  - 専用ホスティングサーバ
  - 良いIPと悪いIPを混ぜる
    - Google の IP を数個混ぜるパターンなどもあり

# SPF認証と実際に配送されたメッセージ

Auth	Total	# of Spam	# of Ham	# of Other
PASS	1,273,994	174,380	1,073,666	25,948
NEUTRAL	48,956	22,750	24,933	1,273
SOFTFAIL	271,253	108,517	156,097	6,639
FAIL	110,867	85,152	24,574	1,141
NONE	2,269,749	1,175,843	1,065,464	28,442
TOTAL	3,974,819	1,566,642	2,344,734	63,443

全スパムの大多数はSPFに非対応。10%はSPFにPASS  
全ハムの5%がSOFTFAIL or FAIL

# 送信元IPアドレスとSPF認証

良性IPアドレス

悪性IPアドレス

Auth	Total	DNSWL	Local WL	End-host BL	Spam Gang BL	Open Proxy BL	Local BL
PASS	7.53 K	1.87 K	3.30 K	94	26	32	150
NEUTRAL	2.08 K	198	184	1.22 K	2	452	785
SOFTFAIL	4.51 K	255	451	2.79 K	39	1.09 K	1.62 K
FAIL	3.18 K	134	142	2.31 K	18	827	1.44 K
NONE	131 K	1.65 K	3.56 K	112 K	258	41.0 K	20.2 K
TOTAL	142 K	3.43 K	6.75 K	115 K	336	41.9 K	20.6 K

良性IPアドレスで SOFTFAIL/FAIL のものが存在(少数)

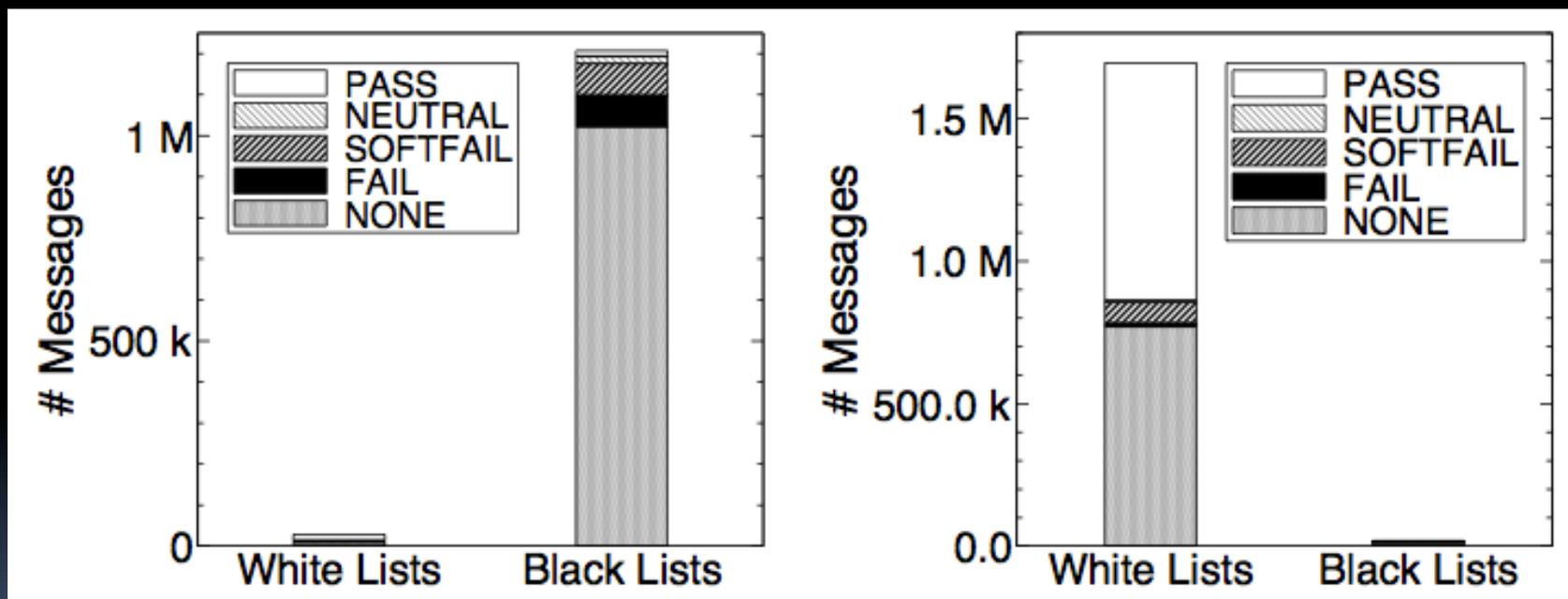
悪性IPアドレスでPASSするものが存在(少数)

アドレスの数としては少数であるが、これらは比較的多くのメッセージを送信している

# 良性・悪性IPアドレス発メッセージとSPF認証結果

#spam

#ham



# まとめ

- 半数の良性ドメインおよび約30%の悪性ドメインはSPFに対応
- SpammerがSPFを利用する目的
  - 認証をパスする(広いIPアドレス空間), レピュテーションを上げる(良いIPを混ぜる)
- スパムの10%はSPF認証をパスしている
- 5%のハムがSPFでSOFTFAILかREJECTとなる
  - 転送の問題、ポータブルなメールアドレスの問題等
- 問題となる送信者の数自体はそれほど多くないが、メッセージ数は無視できない
  - このような送信者が存在することを考慮した上で後段のフィルタリングを実施する必要あり
  - 運用で解決できるケースもあり