

**2006-1: Residential Customer Privacy**

**2003-5: Distributed Information Server Use  
Requirements**

**2002-8: Privatizing POC Information**

**2002-4: Bulk Copies of ARIN's WHOIS**

**2001-7: Bulk ARIN WHOIS Data**

**2000-6: Bulk WHOIS Data Aggregation and Exchange**  
**WHOIS**

**2005-2: Directory Services Overhaul**

**2004-7: Residential Customer Privacy**

**By the Numbers**  
**2004-6: Privacy of Reassignment Information**

**2004-4: Purpose and scope of ARIN WHOIS  
directory**

**2003-16: POC Verification**

**2003-11: Purpose and scope of WHOIS**

.. .

# Purpose and Scope

- When discussing policy that affects WHOIS speakers will often base their opinion on what they *believe* to be in WHOIS.
  - ▶ Actual numbers are almost never used.
  - ▶ Most speakers seem to believe that other organizations put similar data into WHOIS as their own organization.

# Agenda

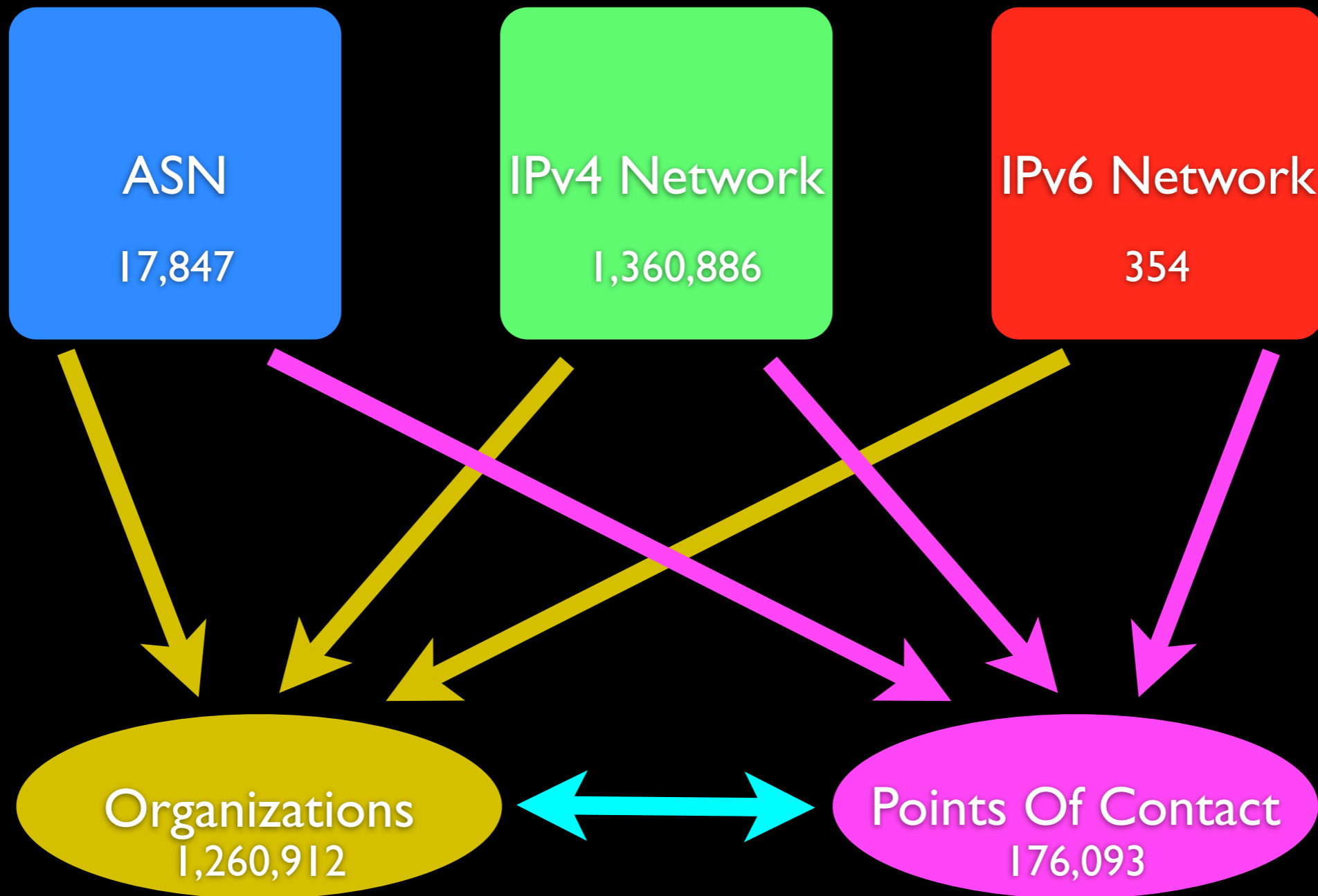
1. Structure of WHOIS, basic statistics.
2. Analysis of missing and incomplete data.
3. Analysis of stale data.
4. Analysis of the effect of 2003-3, Residential Privacy.
5. Analysis of Postal Codes in WHOIS.

# Disclaimer

- All analysis done on a data set retrieved on May 2nd, 2006.
- ARIN came into existence on December 22, 1997. Data with dates prior to that came from other sources.
- No data is included from RWHOIS servers. (More later.)



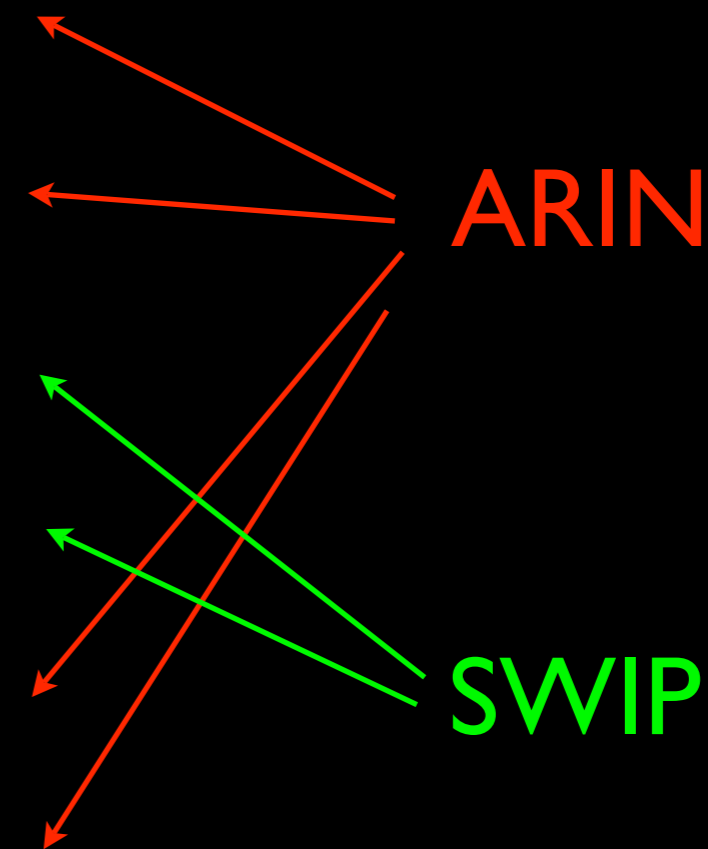
# WHOIS Structure



2,816,092 Records

# IPv4 Network Types

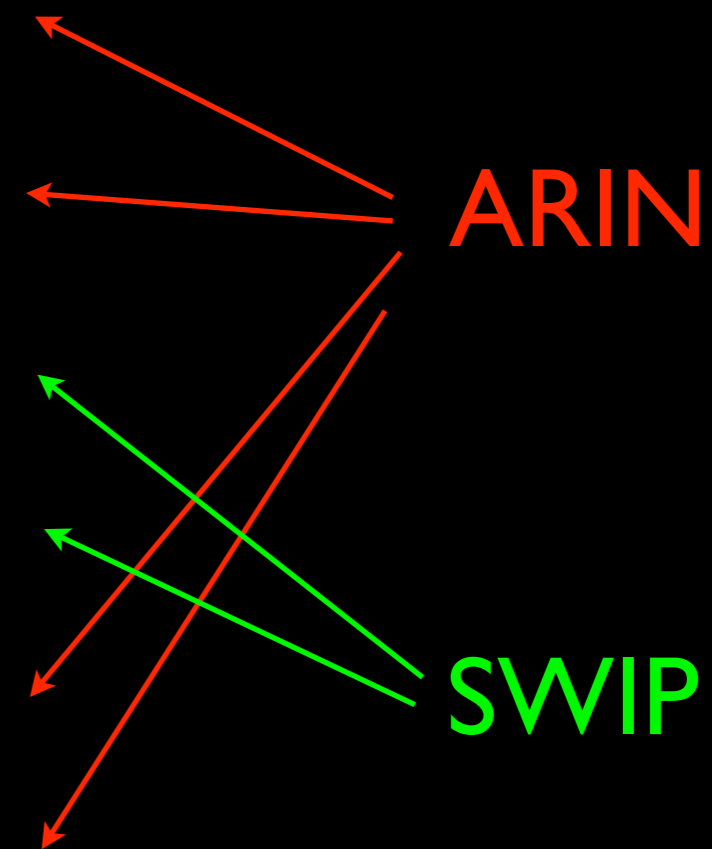
Assignment	28,555
Allocation	8,186
Reassignment	1,302,532
Reallocation	18,656
RIR	2,840
Reserved	33



97% of the records are from SWIP.

# IPv6 Network Types

Assignment	0
Allocation	247
Reassignment	49
Reallocation	50
RIR	8
Reserved	0





# Unique Organizations

	IPv4	IPv6	ASN
Assignment	20,502	0	14918
Allocation	3,469	207	
Reassignment	1,227,539	28	
Reallocation	6,143	40	

95.5% of the organizations are from SWIP.

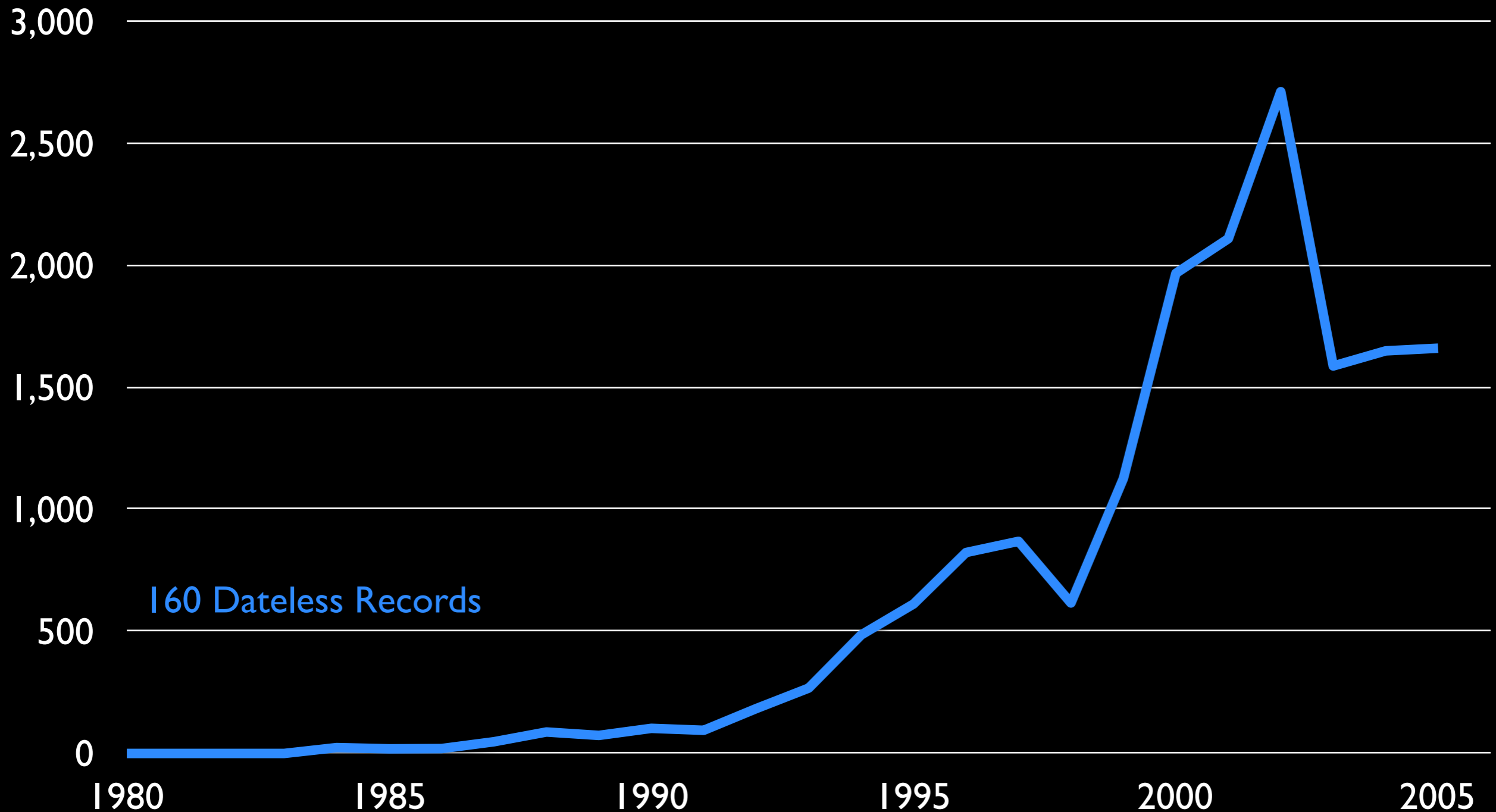
# Unique Organizations

	IPv4	IPv6	ASN
Assignment	1.39	N/A	1.19
Allocation	2.35	1.19	
Reassignment	1.06	1.75	
Reallocation	3.03	1.25	

Average records per organization.

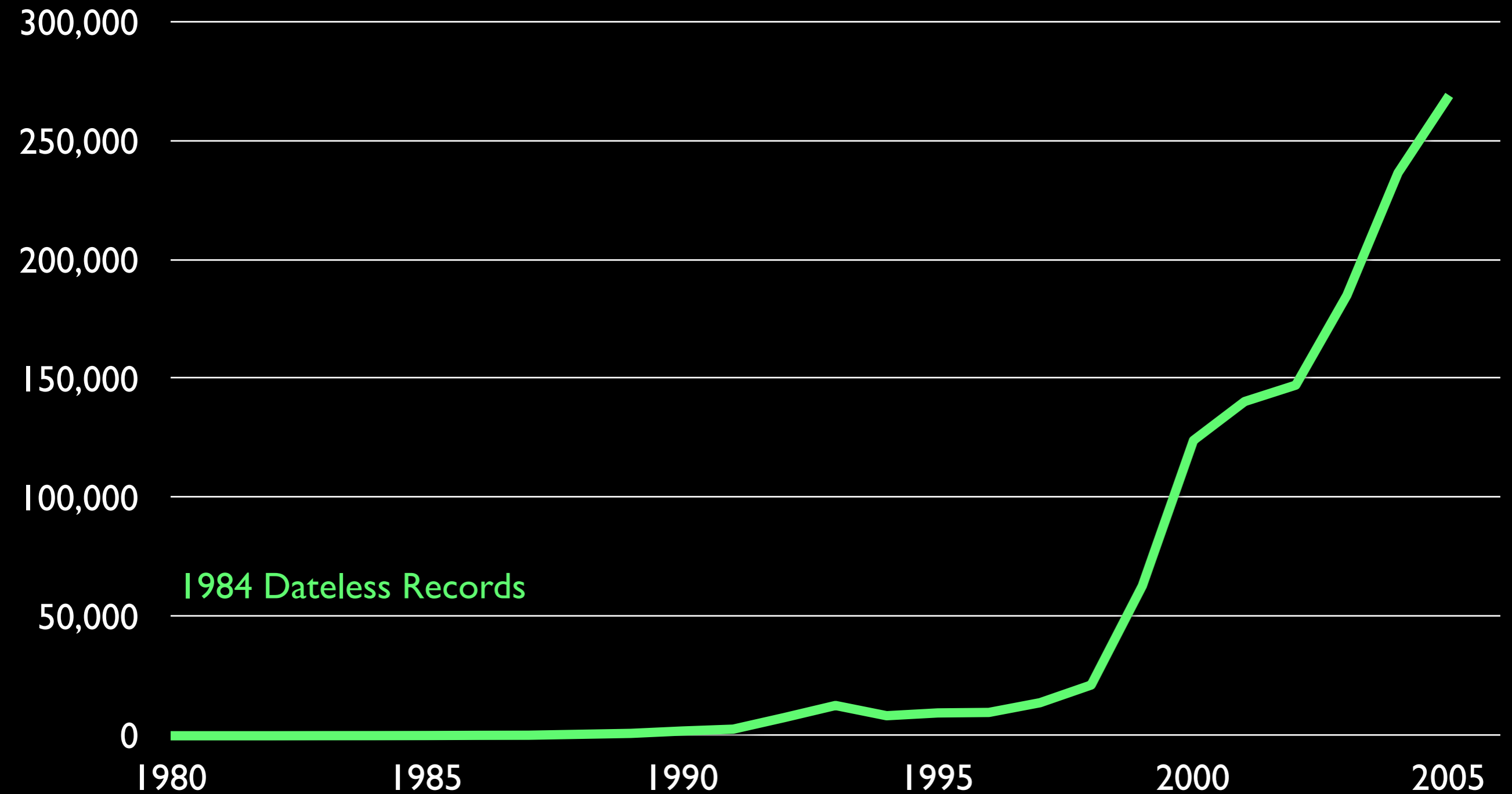
# Record Growth

## Autonomous System Numbers



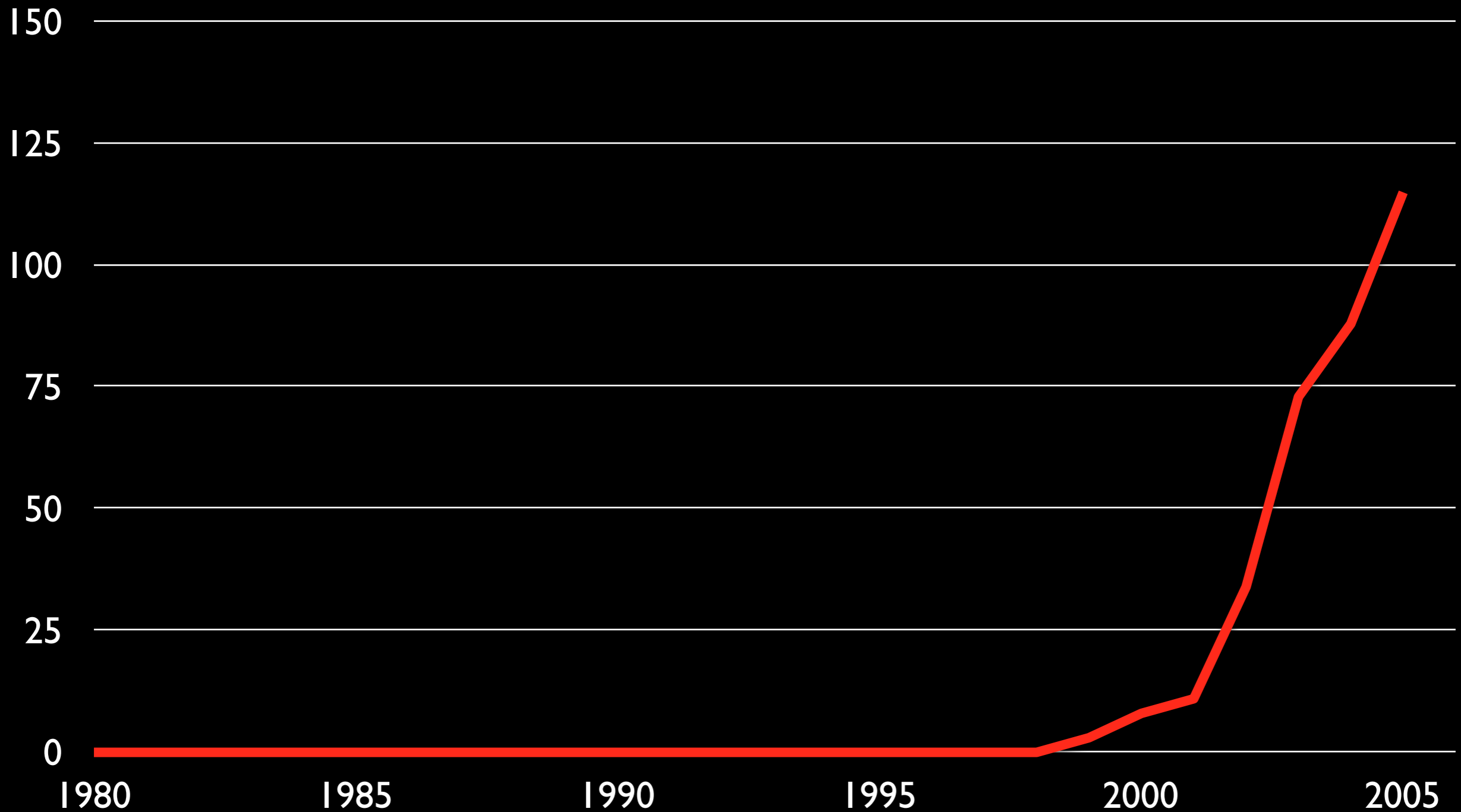
# Record Growth

## IPv4 Networks



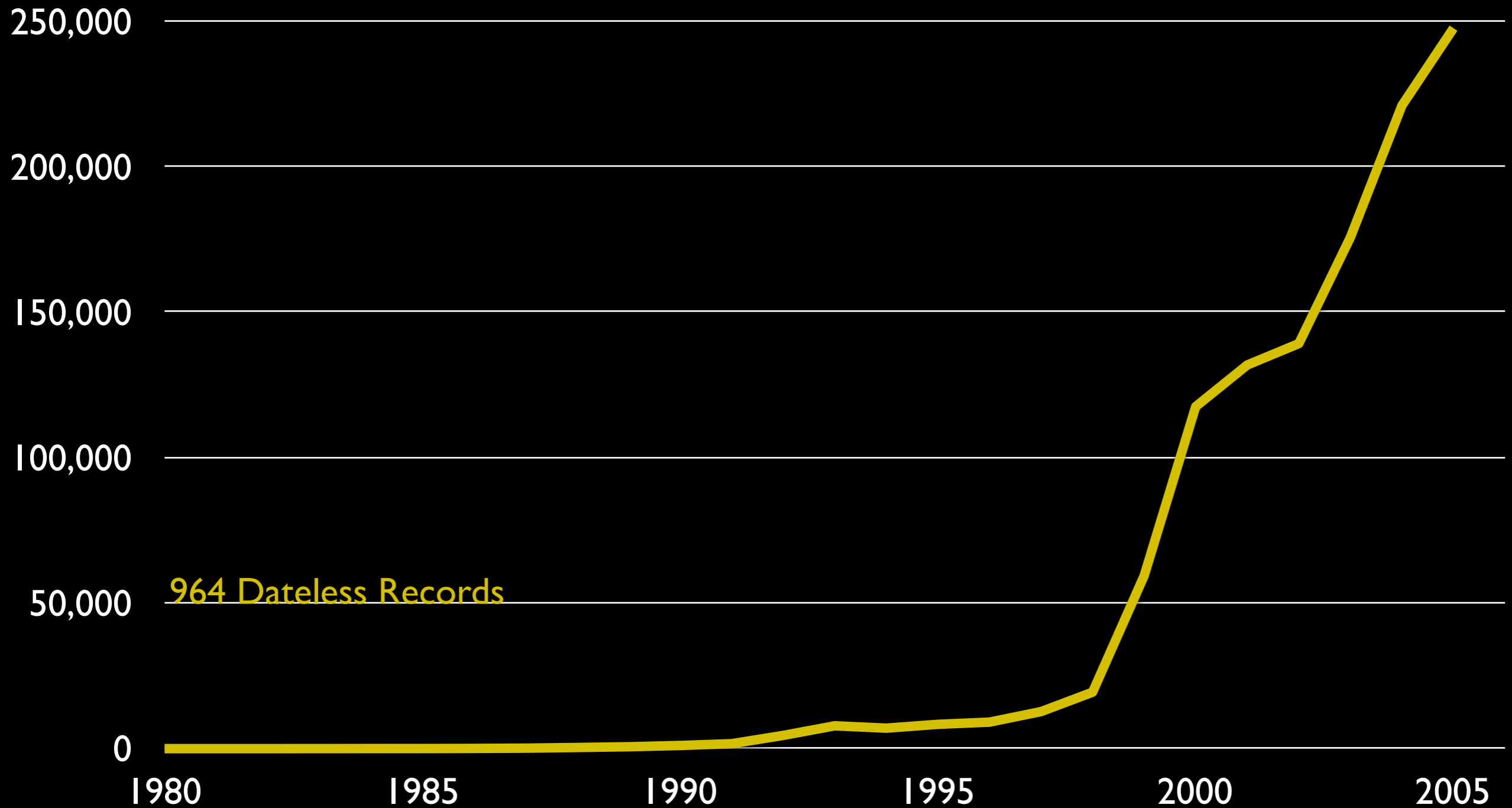
# Record Growth

## IPv6 Networks



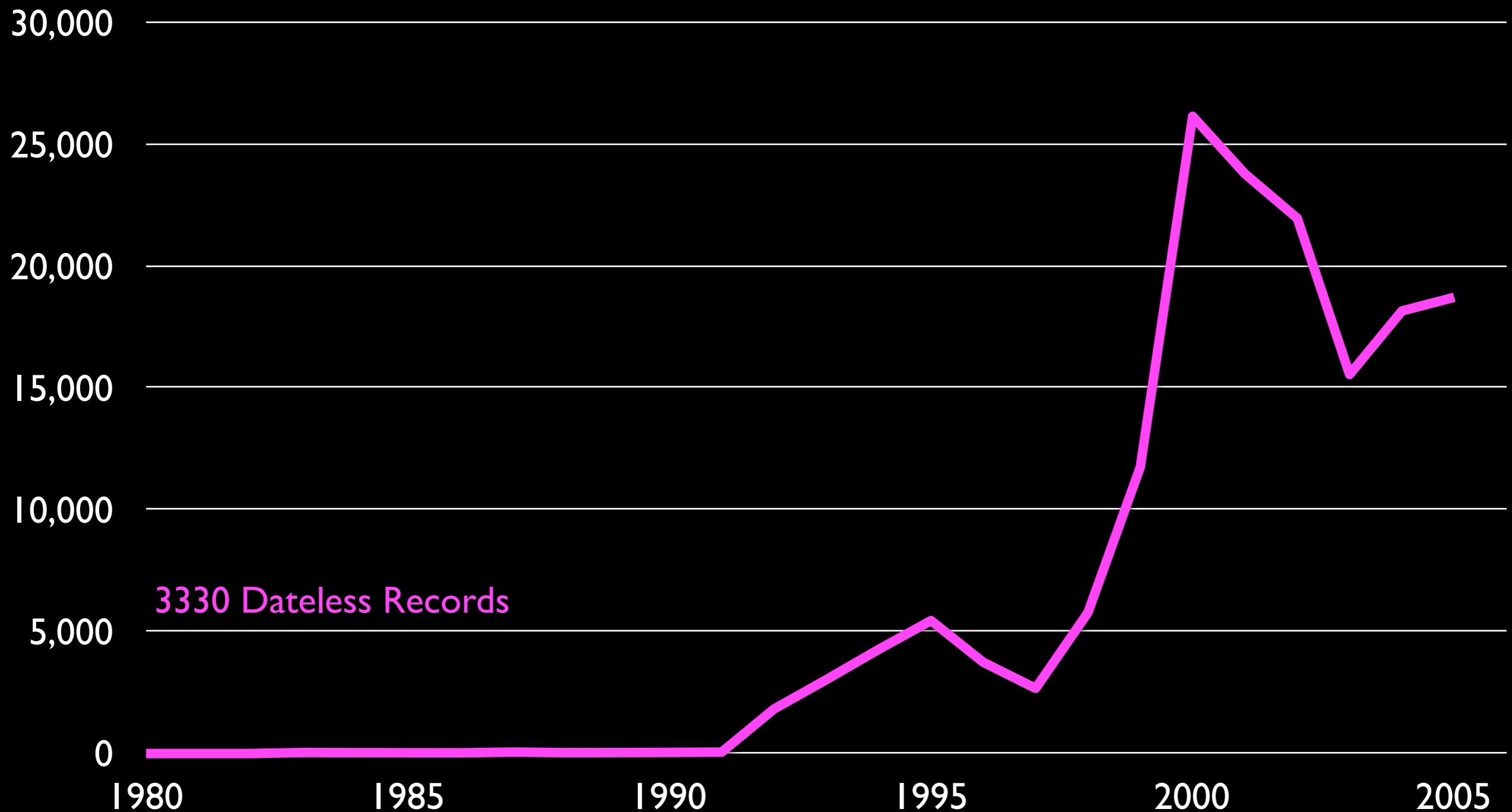
# Record Growth

## Organizations



# Record Growth

## Points Of Contact



# Pre-ARIN Data

Registration Date

Updated Date

ASN	21%	14%
Networks	5%	3%
IPv6	0%	0%
Orgs	4%	2%
POC	14%	10%



2

# Missing and Incomplete Data

- Some records do not contain all fields. Fields may be optional, or may have been added making them absent from earlier records.
- Historical data may have never been recorded, or may have been lost.

# Missing and Incomplete Data

- Some records do not contain all fields. Fields may be optional, or may have been added making them absent from earlier records.
- Historical data may have never been recorded, or may have been lost.

# Incomplete

## Percentage of Records Missing the Listed Field

### ASN

AbuseHandle:	96.90%
NOCHandle:	95.48%
Comment:	91.57%
TechHandle:	23.94%

### Networks

NOCHandle:	99.65%
AbuseHandle:	98.06%
TechHandle:	82.71%
Comment:	79.61%
NetType:	0.01%
Parent:	0.01%

### IPv6 Networks

AbuseHandle:	94.92%
NOCHandle:	94.07%
Comment:	87.85%
TechHandle:	23.73%
Parent:	2.26%

### Orgs

ReferralServer:	99.97%
Comment:	99.87%
OrgNOCHandle:	99.87%
OrgAbuseHandle:	99.81%
OrgAdminHandle:	85.22%
OrgTechHandle:	85.22%
PostalCode:	2.52%
State/Prov:	2.02%
City:	1.94%
Street:	0.03%

### POCs

MobilePhone:	99.25%
FaxPhone:	94.48%
FirstName:	7.94%
LastName:	7.86%
State/Prov:	2.34%
PostalCode:	2.27%
Mailbox:	2.11%
City:	1.98%
OfficePhone:	0.34%
Street:	0.05%

# Incomplete

“assigned” or “allocated”

## Networks

NOCHandle:	97.59%
AbuseHandle:	97.04%
Comment:	89.22%
TechHandle:	14.61%
RegDate:	4.70%

36,741

## IPv6 Networks

Comment:	94.74%
AbuseHandle:	92.71%
NOCHandle:	91.90%
TechHandle:	16.91%

247

# RWHOIS & Data Analysis

- There are 346 unique RWHOIS servers listed in the database on 377 **Organization** records.
  - ▶ No attempt was made to pull down the data from those servers, and include it in this analysis.
  - ▶ Servers were tested for accessibility, 205 (59%) accepted a connection and an RWHOIS command.
    - \* No attempt was made to query actual data.

3

# Return to Sender

- 70 **POC** records have UUCP e-mail addresses. 9 have been updated since 1/1/1998.
- 69 **POC** records have BitNet e-mail addresses. 6 have been updated since 1/1/1998.



# Address Verification

- A test that could not be run, all of the services that could handle the volume of transactions in an automated fashion were commercial services.
- From looking at the data it's clear many companies have moved or gone out of business and not updated WHOIS records.

# Not Responsible, Anymore

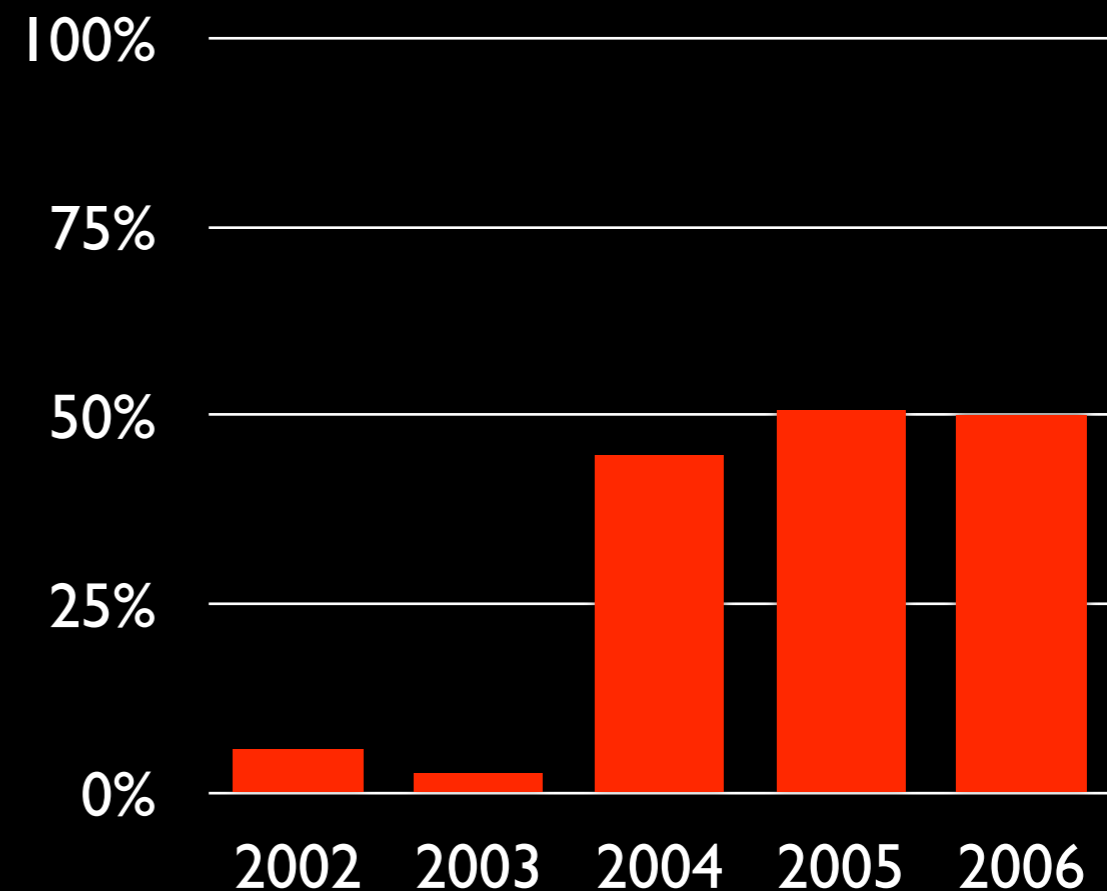
- 7725 POC contacts appear to be orphaned; that is they are not referenced from any ASN, Network, IPv6 Network, or Organization.

4

# Residential Privacy

- Implemented in May of 2004 under policy proposal 2003-3.
- Allows residential customers to have the street name suppressed.
- Graph shows the percentage of new **Org** records by year that have “private” in the street address

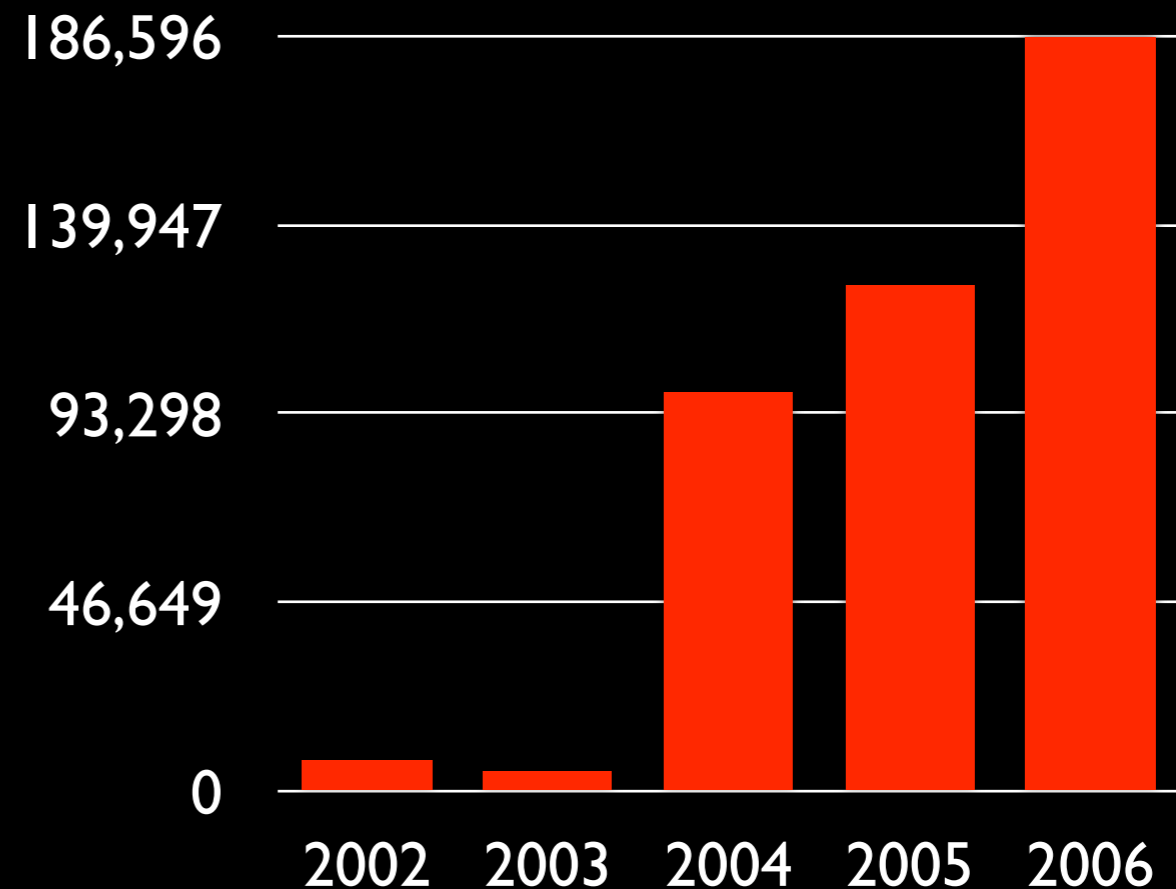
Percentage of new **Org** records marked as Private.



# Residential Privacy

- Out of 1,260,912 organizations in the database, 294,066 (23%) are labeled under the “Residential Privacy” policy.
- Unfortunately, this is only an approximation based on searching for “Private” or “Pvt” in the street address, there is no flag for the residential privacy policy.

Number of new **Org** records marked as Private.



2006 projected from January 1 to May 2 data.

# Outside the Policy?

- Anecdotal evidence suggests that some ISP's were providing privacy by listing their own address.
- We can check for this behavior by looking for street addresses that were used repeatedly.

# Most Popular Streets

Street	Count
268 Bush St (Suite 5000)	104511
2701 W 15th Street PMB 236	43166
1701 Alma St	28742
303 Second Street	22495
1 Mount Pleasant Road	12806
non pub	7518
2623 Camino Ramon	4077
310 Orange St	3581
300 Viger Est	3125
3075 Sanders RdG2E	2997
100 Carpenter Drive Suite 206	2845

# Most Popular Streets

Street	Count
Pacific Bell Internet Services	104511
SBC	43166
SBC	28742
SBC	22495
Rogers Cable	12806
SBC	7518
SBC	4077
South Western Bell	3581
VideoTron	3125
Allstate Insurance	2997
Network Access Solutions (DSLNet?)	2845



# Privacy Totals

- 1,260,912 total organizations.
- 294,066 (23%) under 2003-3, Residential Privacy.
- 232,866 (18%) are listed as the ISP, based on the top 10 streets only.
- 41% of all org records in the database do not point directly at the end user.

# SBC & Residential Privacy

- There are 284,956 records entered for SBC customers that have “Private” in the street address, and “SBC” in the network name.
- That makes SBC responsible for 97% of the records that could be identified as having “private” in the street address, that is appear to be compliant with the 2003-3 policy.

# What else can we tell?

- Even though the street address is missing, we can still look at city, state, and zip.
  - ▶ Note that looking at ZIP includes the previous two.
- We can also try to pull information out of the names.

5

# Postal Codes

- A long running debate as centered around the inclusion (or exclusion) of postal codes, typically in conjunction with residential privacy.
- What can we tell from the post codes already in WHOIS?

# Amazing Facts

- There are organizations from 195 different countries in ARIN's WHOIS database.
  - ▶ US accounts for 94.39% of all records.
  - ▶ Canada accounts for 4.79% of all records.
  - ▶ The next 5 countries are: Great Britain, Mexico, Argentina, Australia, and Venezuela.
- Only 4% of the US ZIP codes have ZIP+4 information.
- There are 24,829 ZIP codes in the database.  
(There are approximately 29,470 ZIP codes total.)

# Top Ten US ZIP Codes

ZIP	Count	Location
75075	328,440	Allen, TX
94104	113,059	San Francisco, CA
94107	48,468	San Francisco, CA
60606	12,758	Chicago, IL
94583	5,822	San Ramon, CA
60062	4,757	Northbrook, IL
06510	3,769	New Haven, CT
20164	3,116	Sterling, VA
75053	3,114	Grand Praire, TX
75025	2,481	Plano, TX

# Top Ten US ZIP Codes

ZIP	Count	Providers
75075	328,440	SBC
94104	113,059	SBC/PBI
94107	48,468	SBC/PBI
60606	12,758	SBC
94583	5,822	SBC
60062	4,757	UUNet/ATT/TDS
06510	3,769	SBC/DSLNet/ATT
20164	3,116	DigiCorp/UUNet
75053	3,114	SBC
75025	2,481	SBC



# SBC

- The single largest user of the residential privacy policy.
- Has a large number of customer records pointed at SBC addresses, rather than directly to the customers.
  - ▶ Is that allowed under the policy?
- Where are all the other end user ISP's?

# Summary

# WHOIS, By the Numbers

1. Trends, pre/post ARIN data, breakdown by type.
2. Missing Data
3. Data inconsistency / integrity issues.
4. Residential Privacy
5. Postal Code Analysis

Is anyone now interested in doing more research?

One more thing...

# A sample record.

```
% whois -a NET-70-247-246-128-1
```

```
CustName:    VICTOR B WILLIAMS MD-050602021747
Address:     Private Address
City:        Plano
StateProv:   TX
PostalCode:  75075
Country:     US
RegDate:     2005-06-02
Updated:     2005-06-02
```

```
% whois -a 70.247.246.128
```

```
SBC Internet Services SBCIS-SIS80 (NET-70-240-0-0-1)
70.240.0.0 - 70.255.255.255
VICTOR B WILLIAMS MD-050602021747 SBC07024724612829050602021758
(NET-70-247-246-128-1)
70.247.246.128 - 70.247.246.135
```

# Is the Assignment Private?

Place the customer name, “Victor B Williams, MD” into [www.google.com](http://www.google.com), hit “I’m feeling lucky.”

## Victor B Williams, MD

Little Rock, AR 72204

### Contact Victor B Williams, MD

Phone: (501) 224-8206

### Business Profile for Victor B Williams, MD

Services: Doctors & Physicians: Family & General Practice

### In the Neighborhood

#### Nearby Businesses

- [Nearby Banks](#)
- [Nearby Book Stores](#)
- [Nearby Coffee Houses](#)
- [Nearby Florists](#)
- [Nearby Gas Stations](#)
- [Nearby Gyms](#)
- [Nearby Grocery Stores](#)
- [Nearby Ice Cream](#)
- [Nearby Nightclubs](#)
- [Nearby Restaurants](#)



Approximate Location

[Larger Map and Driving Directions](#)

# Are you sure?

```
% traceroute 70.247.246.129
traceroute to 70.247.246.129 (70.247.246.129), 64 hops max, 40 byte packets
 1  abc.123.uandme (xxx.xxx.xxx.xxx)  0.534 ms  0.879 ms  0.318 ms
 2  abc.123.uandme (xxx.xxx.xxx.xxx)  1.267 ms  0.895 ms  0.537 ms
 3  ex2-g8-0s1.eqabva.sbcglobal.net (206.223.115.79)  1.139 ms  36.207 ms  1.828 ms
 4  bb2-p2-1.hrndva.sbcglobal.net (151.164.40.54)  2.828 ms  0.907 ms  1.290 ms
 5  core2-p2-0.crhnva.sbcglobal.net (151.164.191.101)  1.559 ms  1.511 ms  1.021 ms
 6  core1-p1-0.crhnva.sbcglobal.net (151.164.188.17)  1.256 ms  2.743 ms  1.845 ms
 7  core1-p11-0.crdltx.sbcglobal.net (151.164.243.217)  49.716 ms  36.900 ms  35.803 ms
 8  bb1-p2-0.rcsntx.sbcglobal.net (151.164.40.38)  35.257 ms  35.976 ms  35.384 ms
 9  151.164.42.222 (151.164.42.222)  42.994 ms  43.297 ms  43.225 ms
10  dist1-vlan30.ltrkar.sbcglobal.net (151.164.64.225)  43.214 ms  43.201 ms  43.276 ms
11  rback6-g1-0.ltrkar.sbcglobal.net (151.164.64.138)  43.297 ms  43.253 ms  43.747 ms
12  adsl-70-247-246-134.dsl.ltrkar.swbell.net (70.247.246.134)  59.325 ms  57.810 ms  59.387 ms
```

Plus, it is not only the first result returned by google, but the *only* result.

Questions?  
Discussion?