

Example: LBL Forensics

i dont think this looks good:

```
Sep 20 00:30:37 <local-addr> /USR/SBIN/CRON[24948]:  
(root) CMD (/usr/share/hCtQEFtTsNlb.p2/.p-2.4a i &>  
/dev/null)
```

the ".p-2.4a" is one of the Phalanx backdoor signatures.

> checking logs, looks like the problems started after a reboot around 2:30 PM on the sixteenth. So, maybe have been something "dormant" waiting for a reboot well in advance of the *<elided>* account.

Can you pull the disks? I'll pick them up from you for imaging.

>> Its fairly strange that multiple computers, when ssh'd by *ATTACKER.uk* respond back with a connection back to an unspecified high port (if it was ident, that would be understandable) - note that <*VICTIM1*> is doing that, but also <*VICTIM2*>, and <*VICTIM3*> - other hosts that *ATTACKER.uk* probed [4 hostnames elided] are former compromised hosts...

- These are all new accounts, as of Wednesday 9/16. So it seems very likely that the NERSC <elided> account was sniffed on <*victim1*> because Robert indicates that he only logged into <elided> from <*host*>, and only to test it.

... Sounds like we want to reconnoiter with NERSC.

- Another clue: a previous NERSC incident involved user "<elided>" on <*victim1*> a few months ago.

- The login came from a host *<IP-address>* located in *<country>* which is quite unusual based on your history.
- Given the ~500 msec delay between the two and the consistently short data volume on the SSH connection, this very likely is the attacker issuing a single command via SSH to back-connect to their machine. The telltale is that the second connection lasts a number of seconds and transfers a good amount of data. It might be the transfer part of an scp, say.

That then suggests that any machine responding in this fashion is compromised, because the attacker was able to run a command on it.

- *<IP-address-2>* is exhibiting the same behavior as *<IP-address-1>* - a backchannel return response to an inbound ssh - suggest looking for connections to/from that IP as well.
- For both *<victim-1>* and *<victim-2>*, the `/usr/share/LecPuokMdTSR.p2` directory was used for the rootkit - might be a good idea to check for the existence of that directory - if it exists, please, please, PLEASE, don't access it, as that could affect timestamps, but just report.

Exploiting Underlying Structure for Detailed Reconstruction of an Internet-scale Event

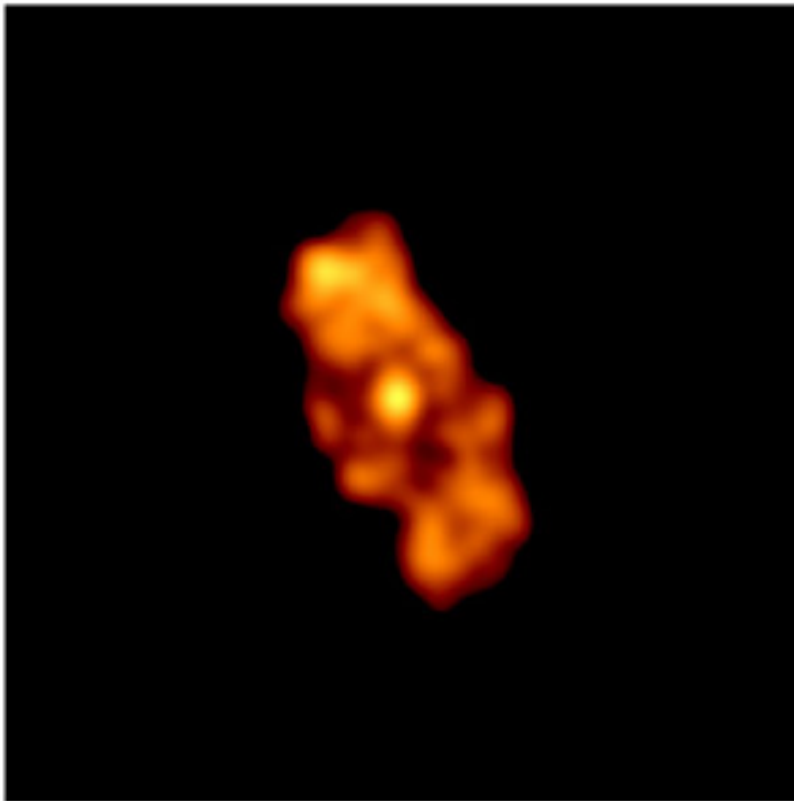
Abhishek Kumar (Georgia Tech / Google)

Vern Paxson (ICSI)

Nicholas Weaver (ICSI)

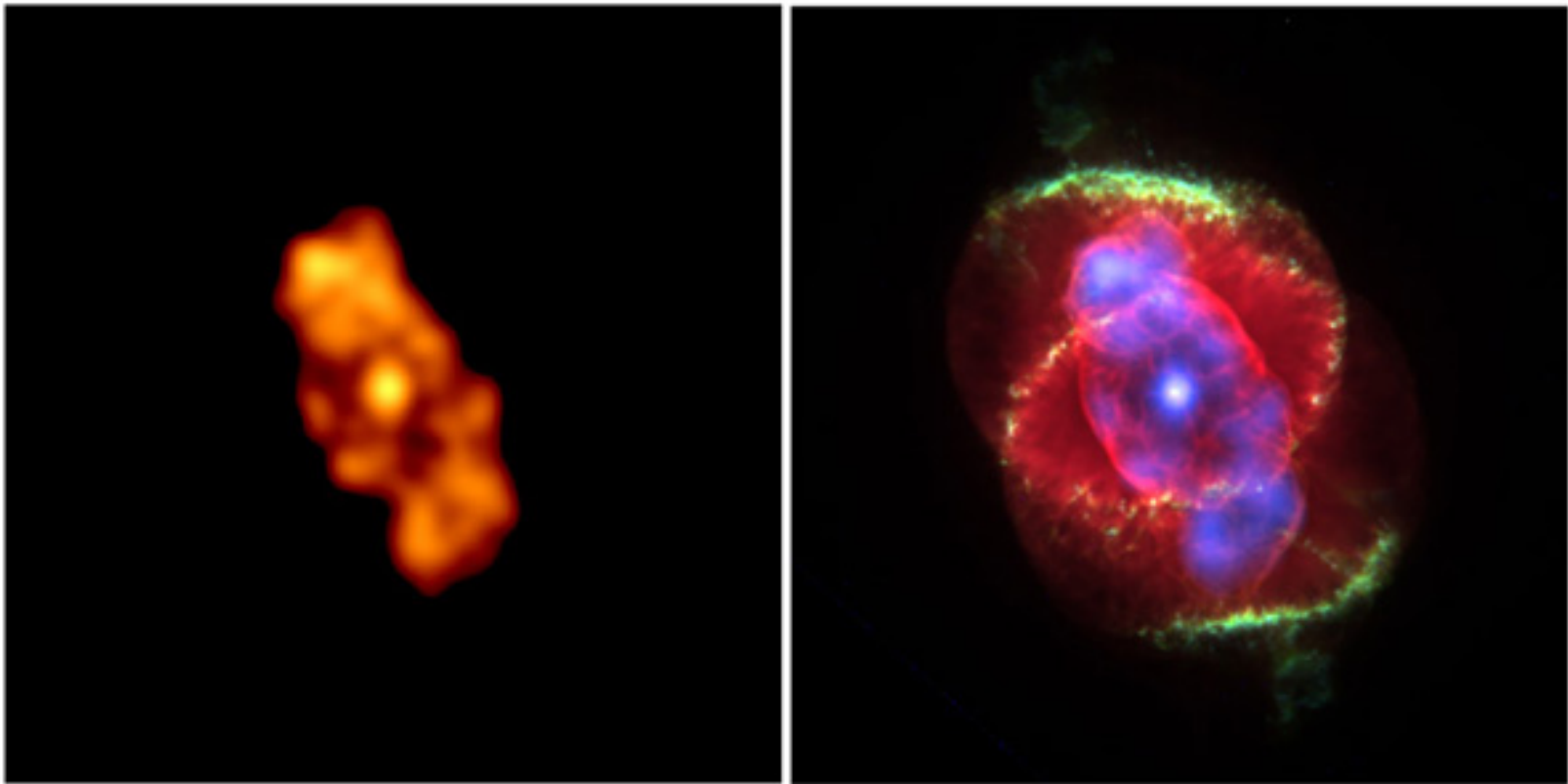
Proc. ACM Internet Measurement Conference 2005

Enhancing Telescope Imagery



NGC6543: Chandra X-ray Observatory Center (<http://chandra.harvard.edu>)

Enhancing Telescope Imagery



NGC6543: Chandra X-ray Observatory Center (<http://chandra.harvard.edu>)

The “Witty” Worm

- Released March 19, 2004.
- Exploited flaw in the *passive analysis* of Internet Security Systems products
- Worm fit in a *single* Internet packet
 - **Stateless**: When scanning, worm could “fire and forget”
- Vulnerable pop. (12K) attained in 75 minutes.
- Payload: *slowly corrupt random disk blocks*.
- Flaw had been announced the *previous day*.
- Written by a Pro.

Witty Telescope Data

- UCSD telescope recorded every Witty packet seen on /8 (2^{24} addresses).
 - *But with unknown losses*
- In the **best case**, we see ≈ 4 of every 1,000 packets sent by each Witty infectee.

? What can we figure out about the worm?

What Exactly Does Witty Do?

1. Seed the PRNG using system uptime.
2. Send 20,000 copies of self to randomly selected destinations.
3. Open physical disk chosen randomly between 0 .. 7.
4. If success:
5. Overwrite a randomly chosen block on this disk.
6. Goto line 1.
7. Else:
8. Goto line 2.

Generating (Pseudo-)Random Numbers

- *Linear Congruential Generator* (LCG)
proposed by Lehmer, 1948:

$$X_{i+1} = X_i * A + B \mod M$$

- Picking A, B takes care, e.g.:

$$A = 214,013$$

$$B = 2,531,011$$

$$M = 2^{32}$$

- Theorem: the *orbit* generated by these is a complete permutation of $0 \dots 2^{32}-1$
- Another theorem: we can *invert* this generator

srand(seed) { $X \leftarrow \text{seed}$ }

rand() { $X \leftarrow X * 214013 + 2531011$; return X }

main()

1. **srand**(get_tick_count());
2. for($i=0; i < 20,000; i++$)
3. $\text{dest_ip} \leftarrow \mathbf{rand()}_{[0..15]} \parallel \mathbf{rand()}_{[0..15]}$
4. $\text{dest_port} \leftarrow \mathbf{rand()}_{[0..15]}$
5. $\text{packetsize} \leftarrow 768 + \mathbf{rand()}_{[0..8]}$
6. $\text{packetcontents} \leftarrow \text{top-of-stack}$
7. sendto()
8. if(open_physical_disk($\mathbf{rand()}_{[13..15]}$))
9. write($\mathbf{rand()}_{[0..14]} \parallel 0x4e20$)
10. goto 1
11. else goto 2

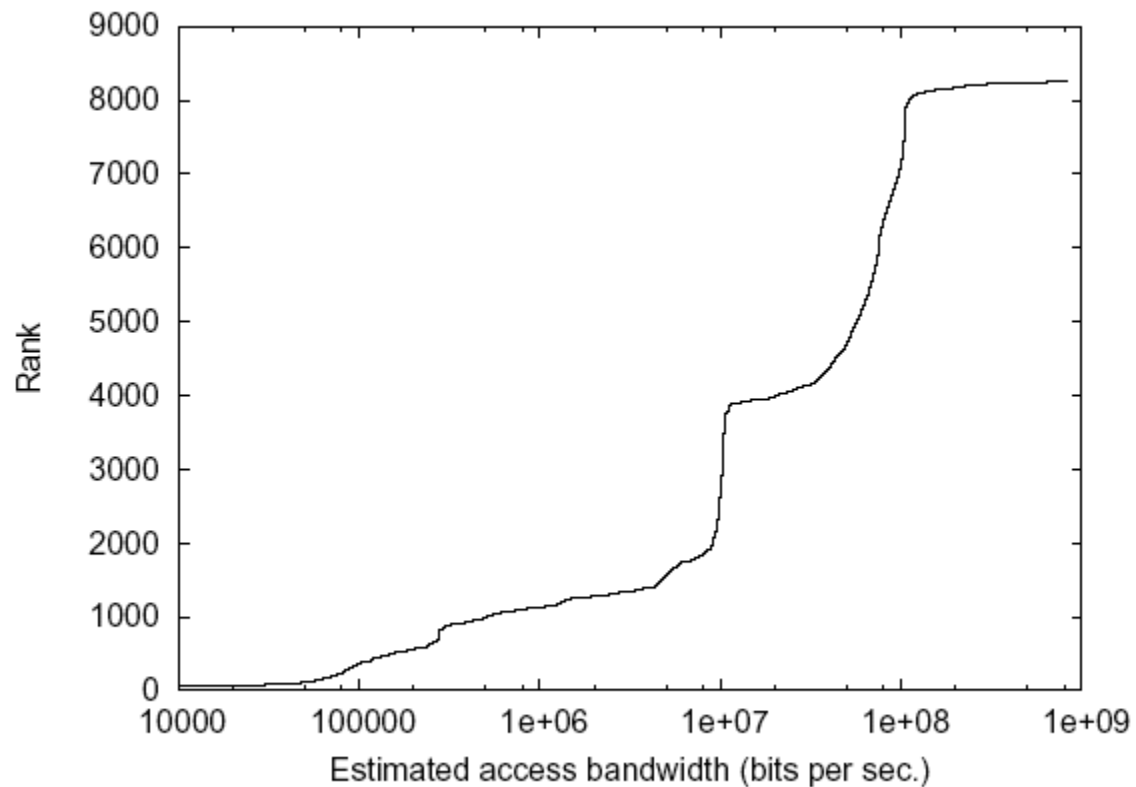
What Can We Do Seeing Just 4 Packets Per Thousand?

- Each packet contains bits from 4 consecutive PRNGs:
3. $dest_ip \leftarrow \text{rand()}_{[0..15]} \parallel \text{rand()}_{[0..15]}$
4. $dest_port \leftarrow \text{rand()}_{[0..15]}$
5. $packet_size \leftarrow 768 + \text{rand()}_{[0..8]}$
 - If first call to **rand()** returns X_i :
3. $dest_ip \leftarrow (X_i)_{[0..15]} \parallel (X_{i+1})_{[0..15]}$
4. $dest_port \leftarrow (X_{i+2})_{[0..15]}$
 - Given top 16 bits of X_i , now *brute force* all possible lower 16 bits to find which yield consistent top 16 bits for X_{i+1} & X_{i+2}
- ⇒ **Single** Witty packet suffices to extract infectee's *complete* PRNG state! Think of this as a **sequence number**.

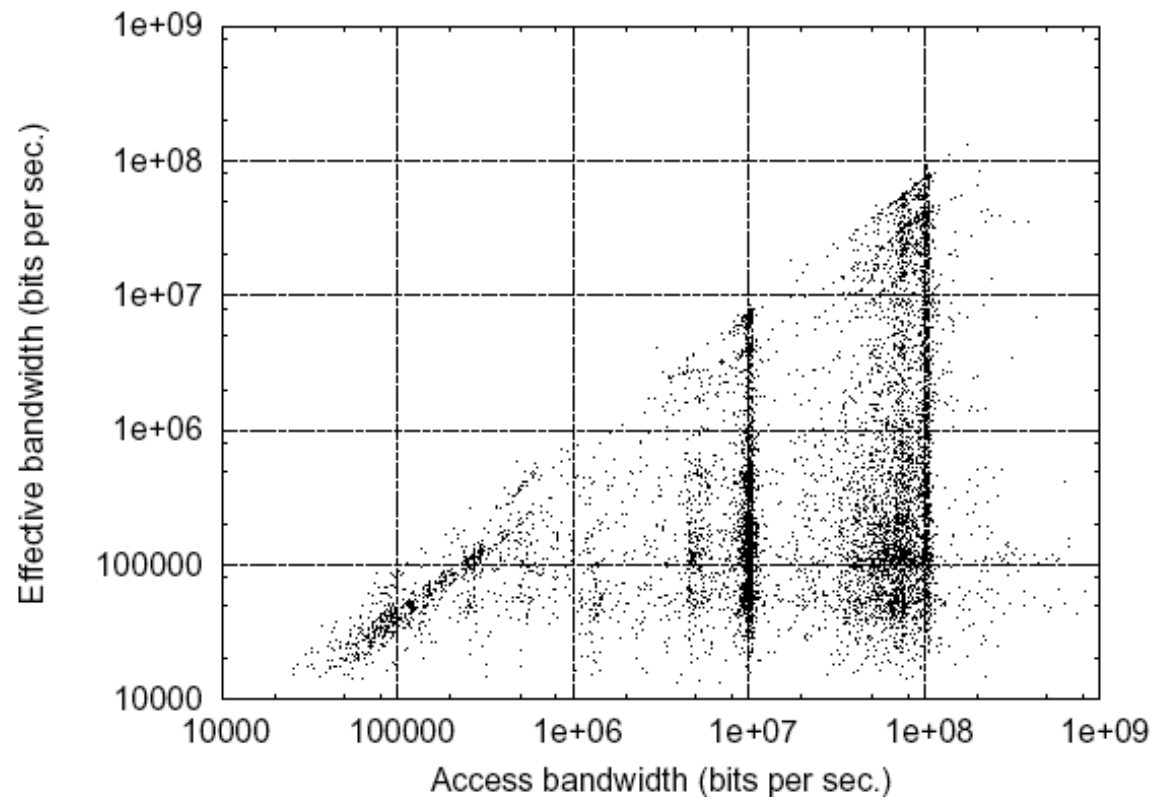
Cool, But So What?

- E.g., *Individual Access Bandwidth Estimation*
 - Suppose two consecutively-observed packets from source **S** arrive with states X_i and X_j
 - Compute $j-i$ by counting # of cranks forward from X_i to reach X_j
 - # packets sent between the two observed = $(j-i)/4$
 - **sendto** call in Windows is *blocking*
 - Ergo, access bandwidth of that infectee should be $(j-i)/4 * \text{size-of-those-packets} / \Delta T$
 - Note: works even in the presence of very heavy packet loss

Inferred Access Bandwidth of Individual Witty Infectees

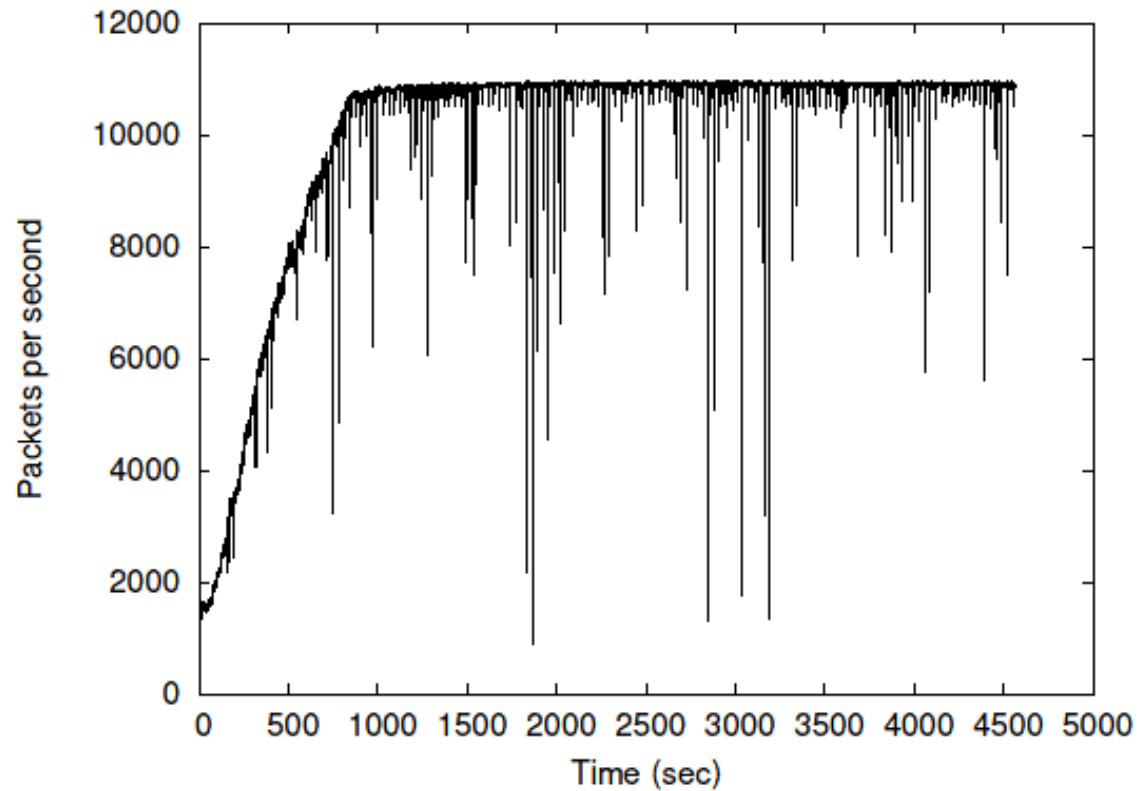


Precise Bandwidth Estimation vs. Rates Measured by Telescope



Systematic Telescope Loss

Measuring Effective Bandwidth



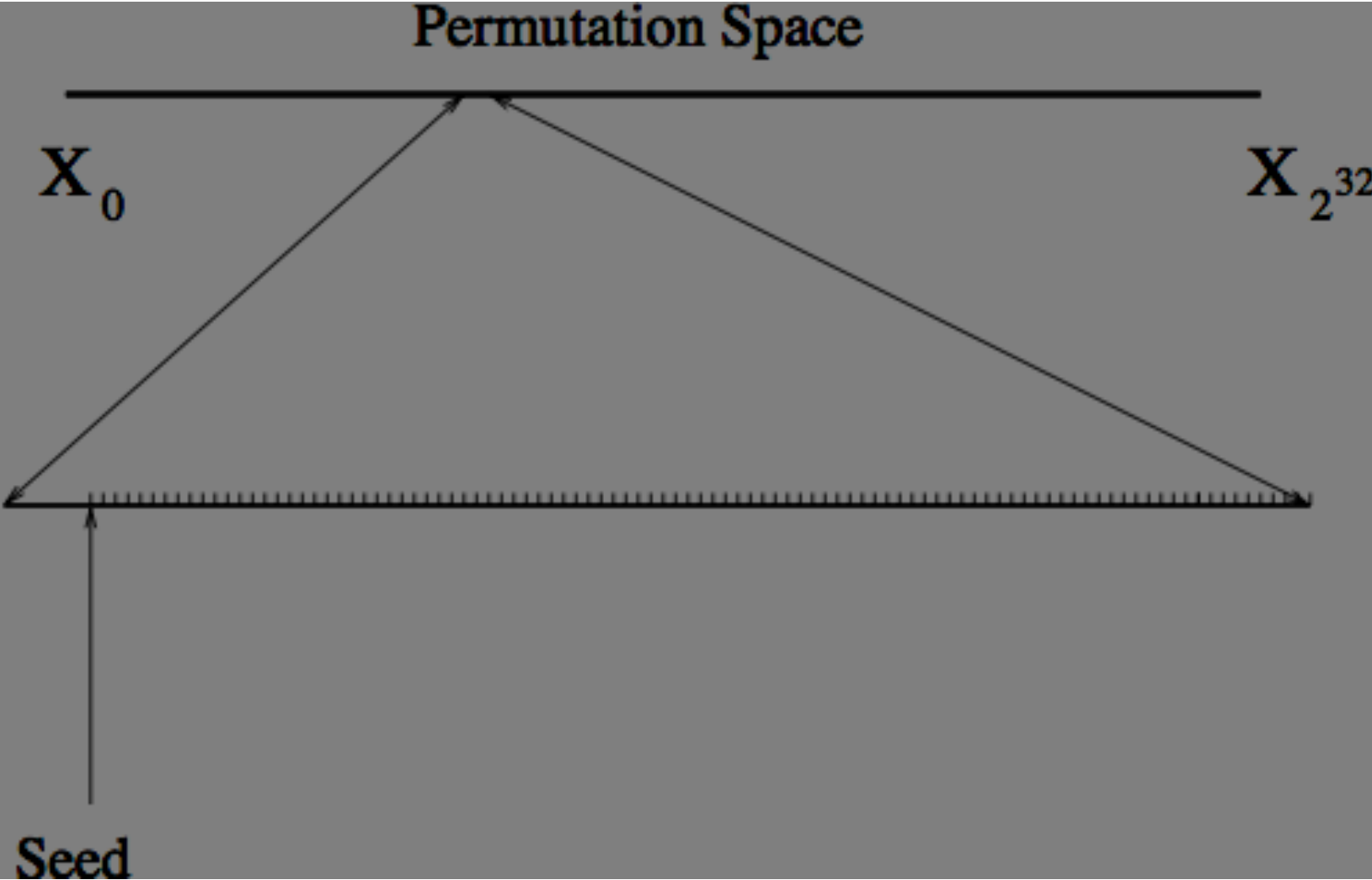
```
srand(seed) {  $X \leftarrow \text{seed}$  }  
rand() {  $X \leftarrow X * 214013 + 2531011$ ; return  $X$  }
```

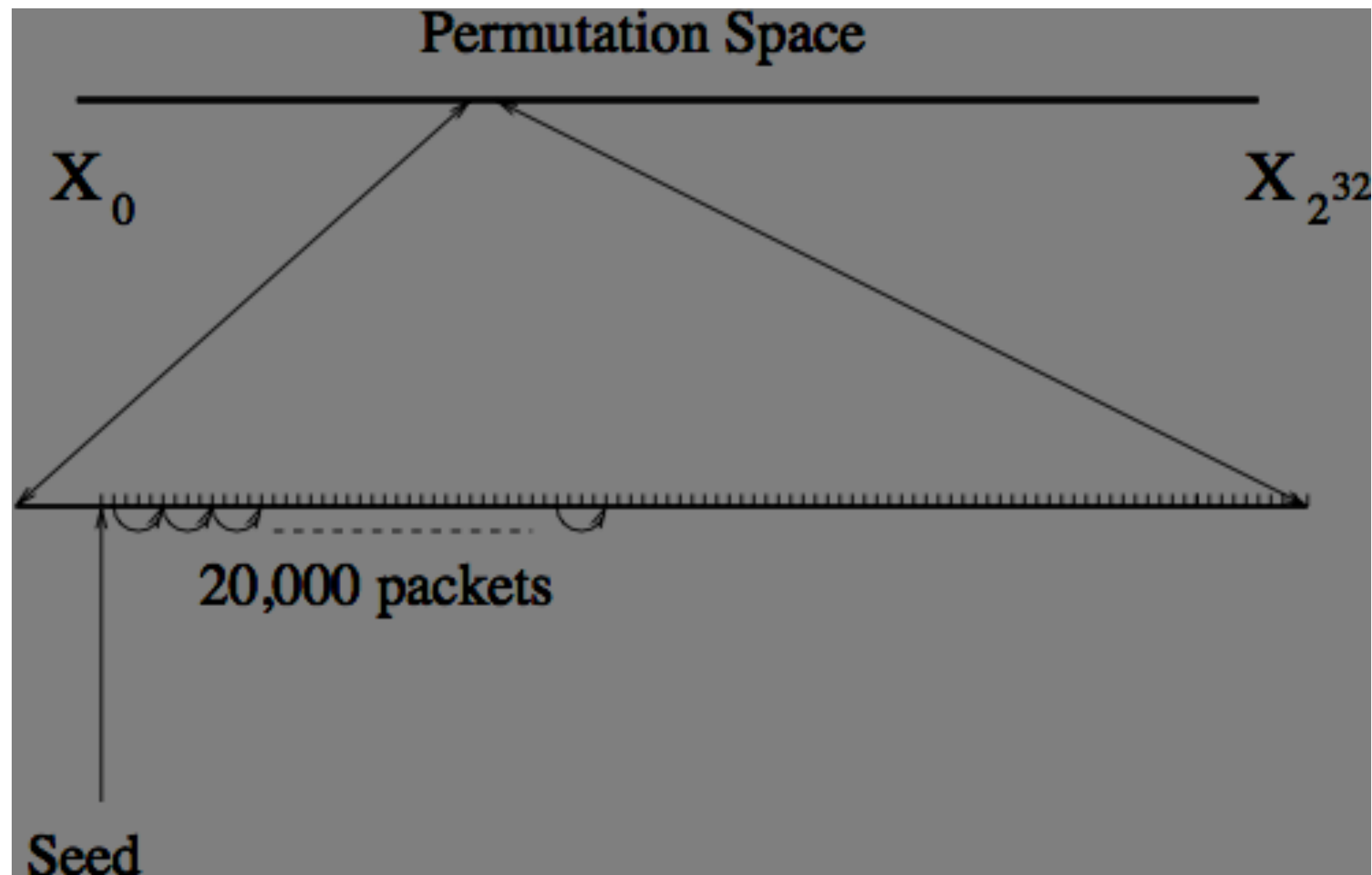
```
main()
```

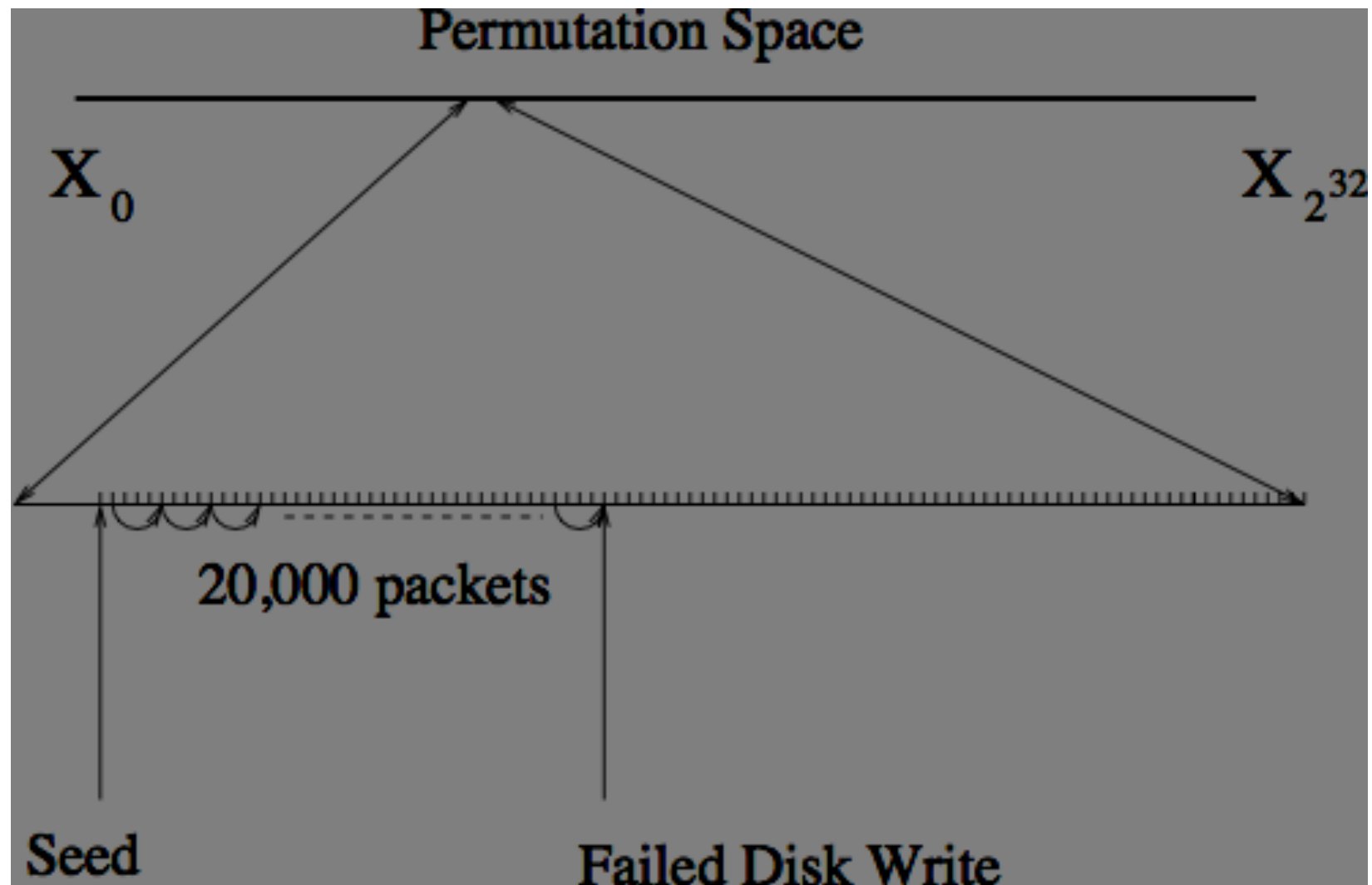
```
1.  srand(get_tick_count());  
2.  for(i=0;i<20,000;i++)  
3.      dest_ip  $\leftarrow$  rand()[0..15] || rand()[0..15] }  
4.      dest_port  $\leftarrow$  rand()[0..15] } 4 calls to rand()  
5.      packetsize  $\leftarrow$  768 + rand()[0..8] } per loop  
6.      packetcontents  $\leftarrow$  top-of-stack  
7.      sendto()  
8.      if(open_physical_disk(rand()[13..15] )) } Plus one more every 20,000  
9.          write(rand()[0..14] || 0x4e20) } packets, if disk open fails ...  
10.      goto 1 } ... Or complete reseeding if not  
11.  else goto 2
```

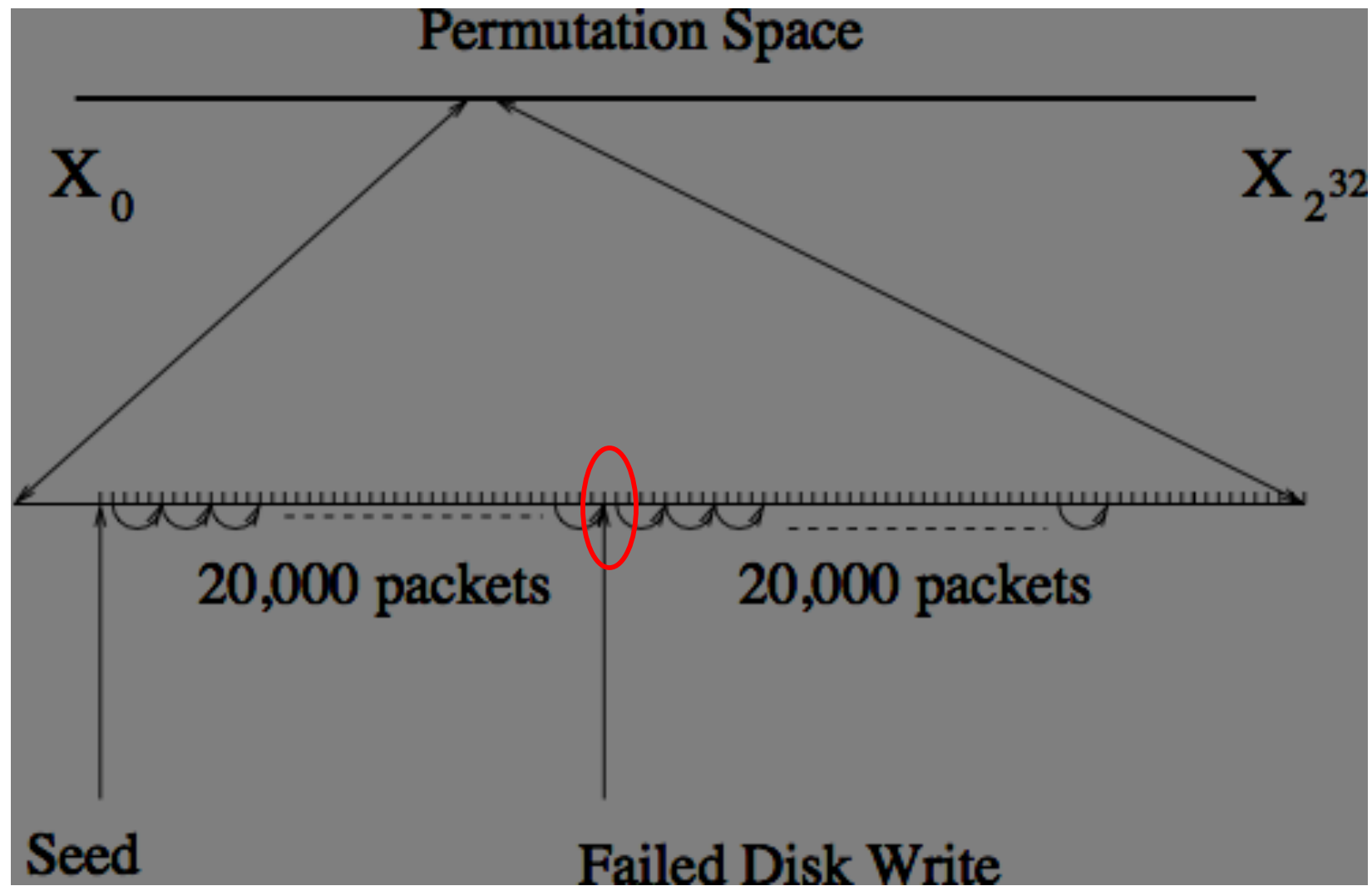
Witty Infectee Reseeding Events

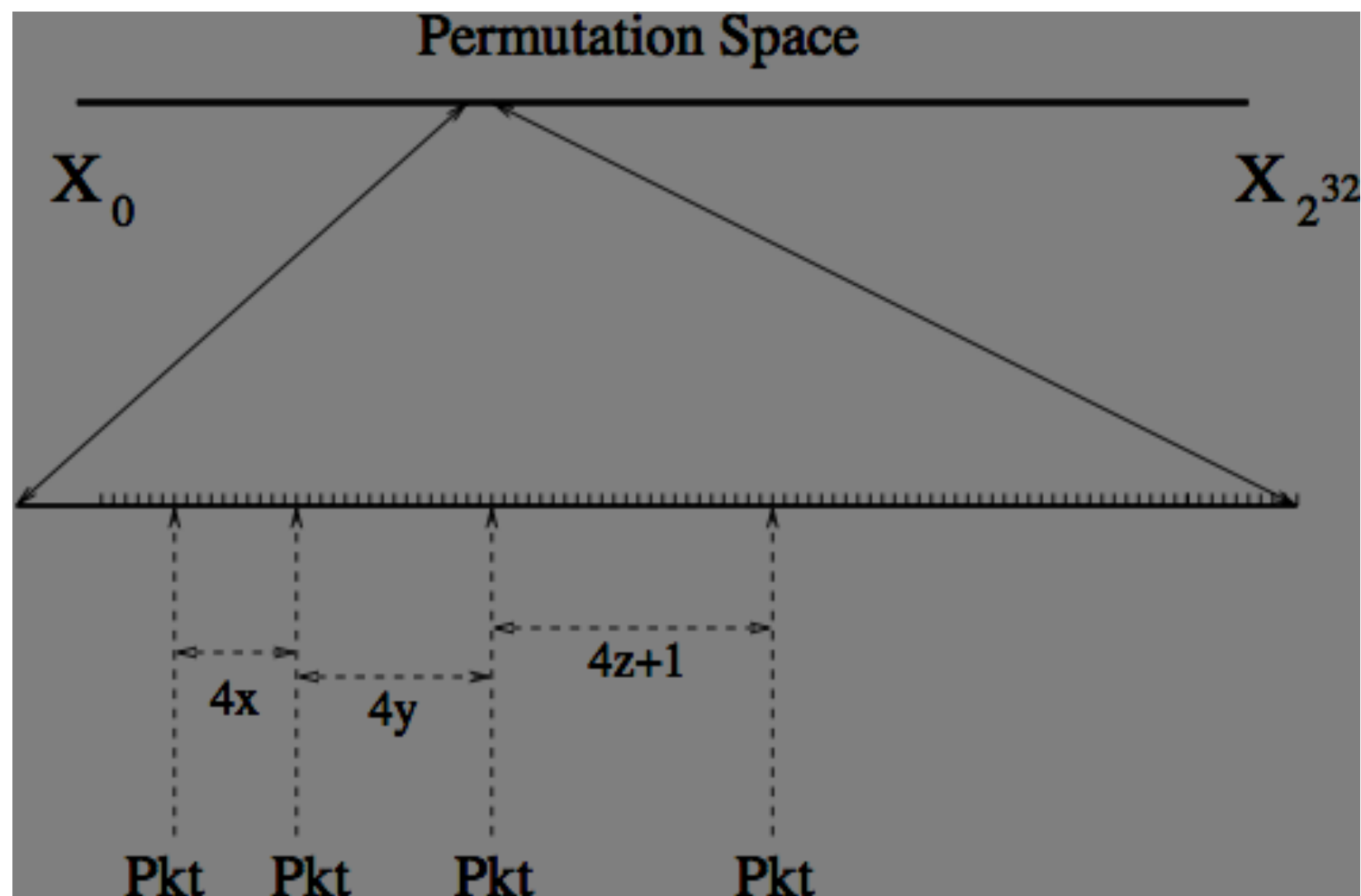
- For packets with state X_i and X_j :
 - If from the **same** batch of 20,000 then
 - $j - i = 0 \mod 4$
 - If from separate but adjacent batches, for which Witty **did not** reseed, then
 - $j - i = 1 \mod 4$
(but which of the 100s/1000s of intervening packets marked the phase shift?)
 - If from batches across which Witty reseeded, then **no apparent relationship**.

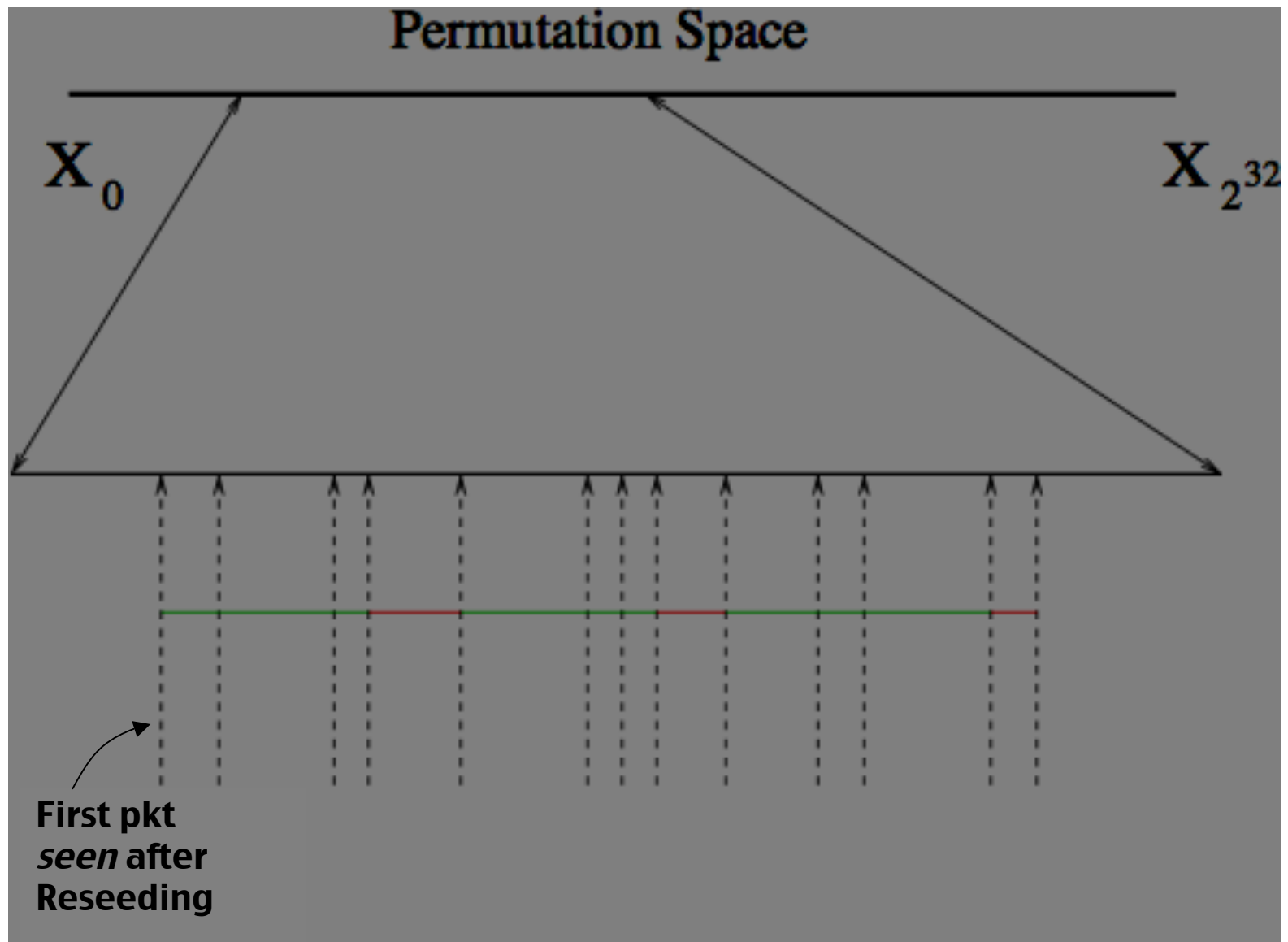




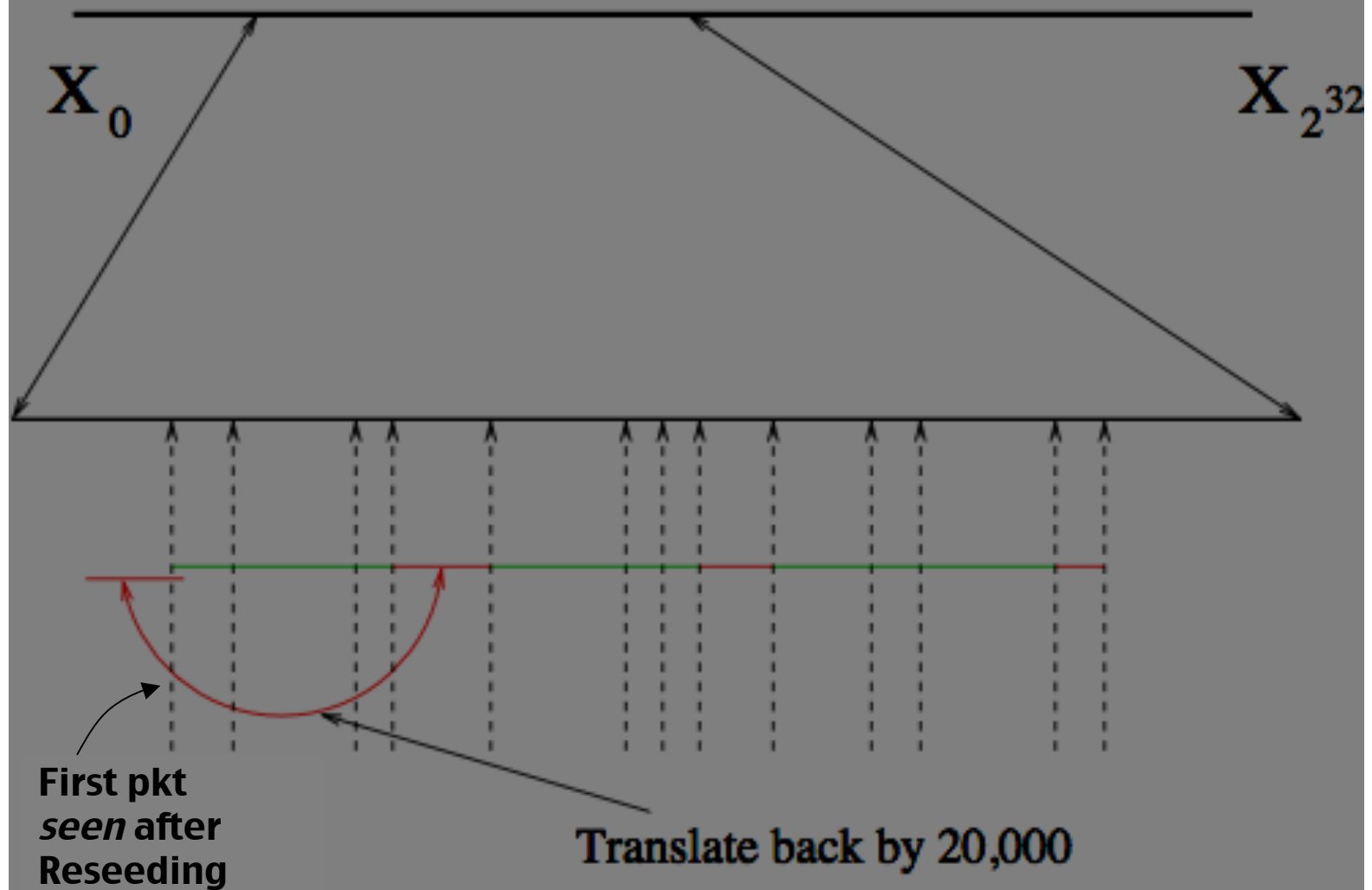




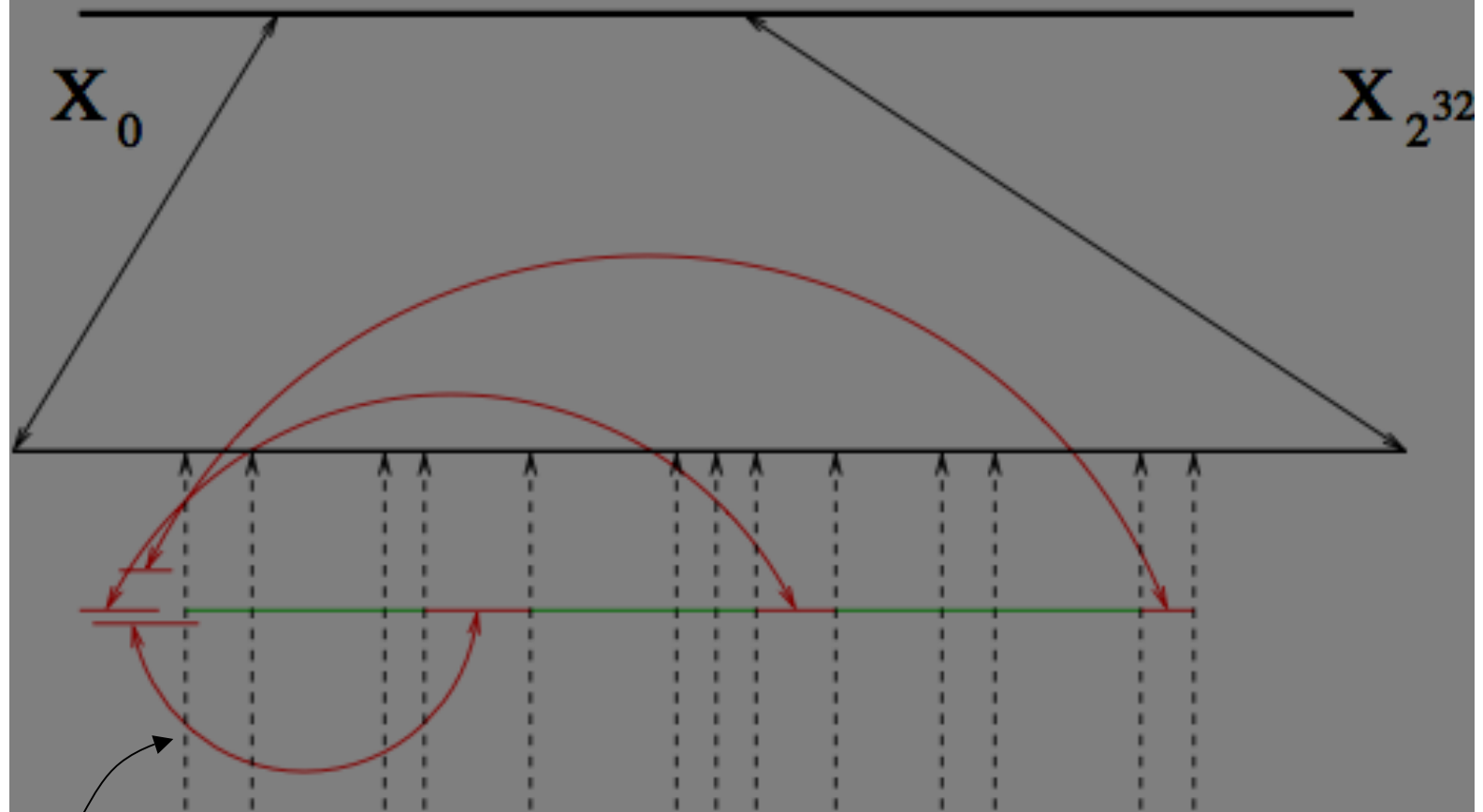




Permutation Space

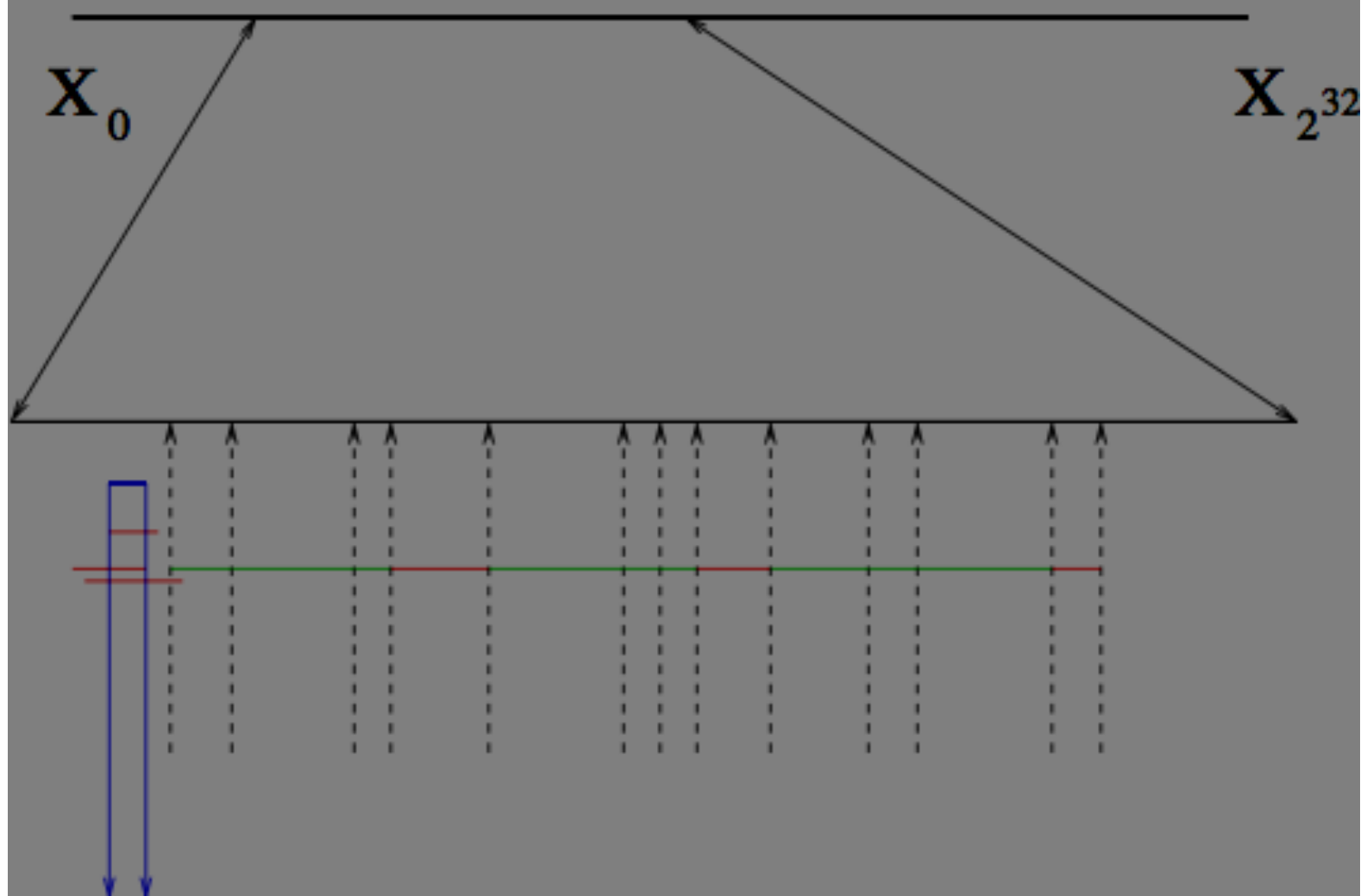


Permutation Space

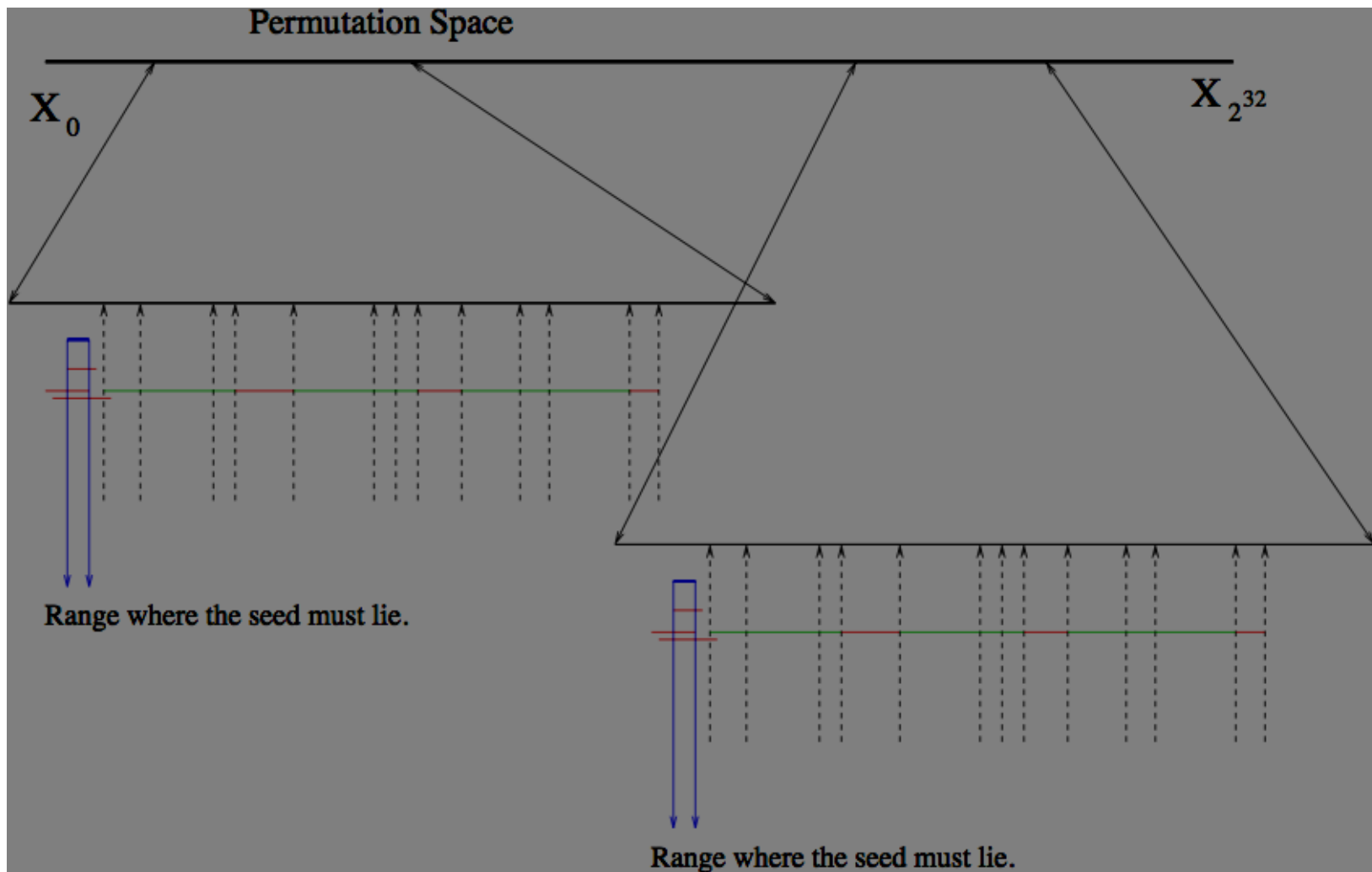


First pkt
seen after
Reseeding

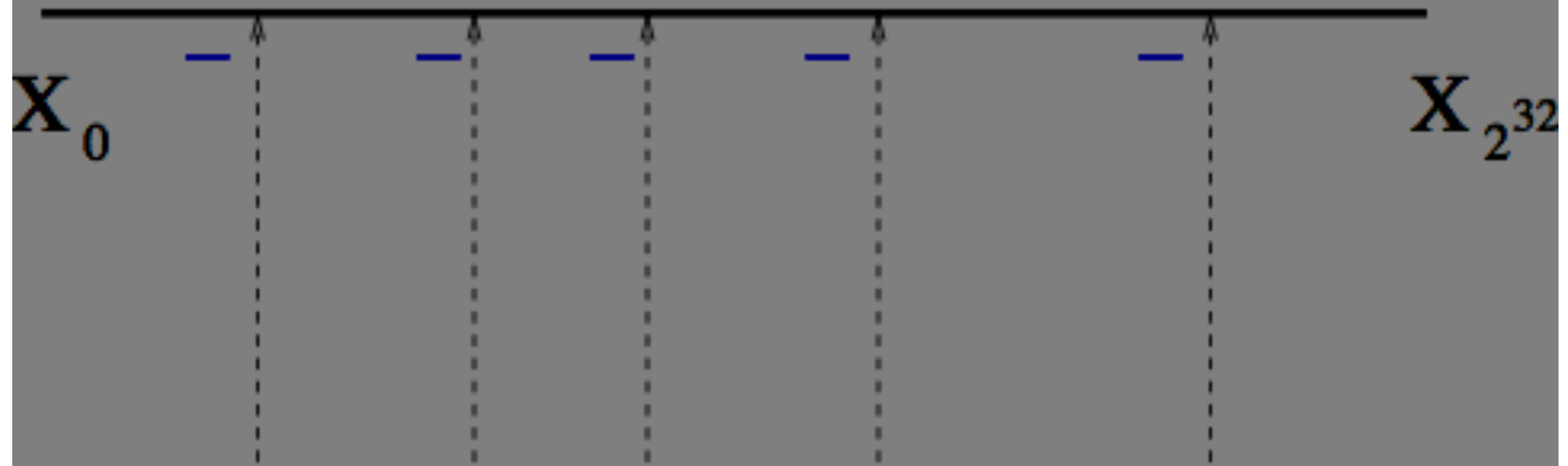
Permutation Space



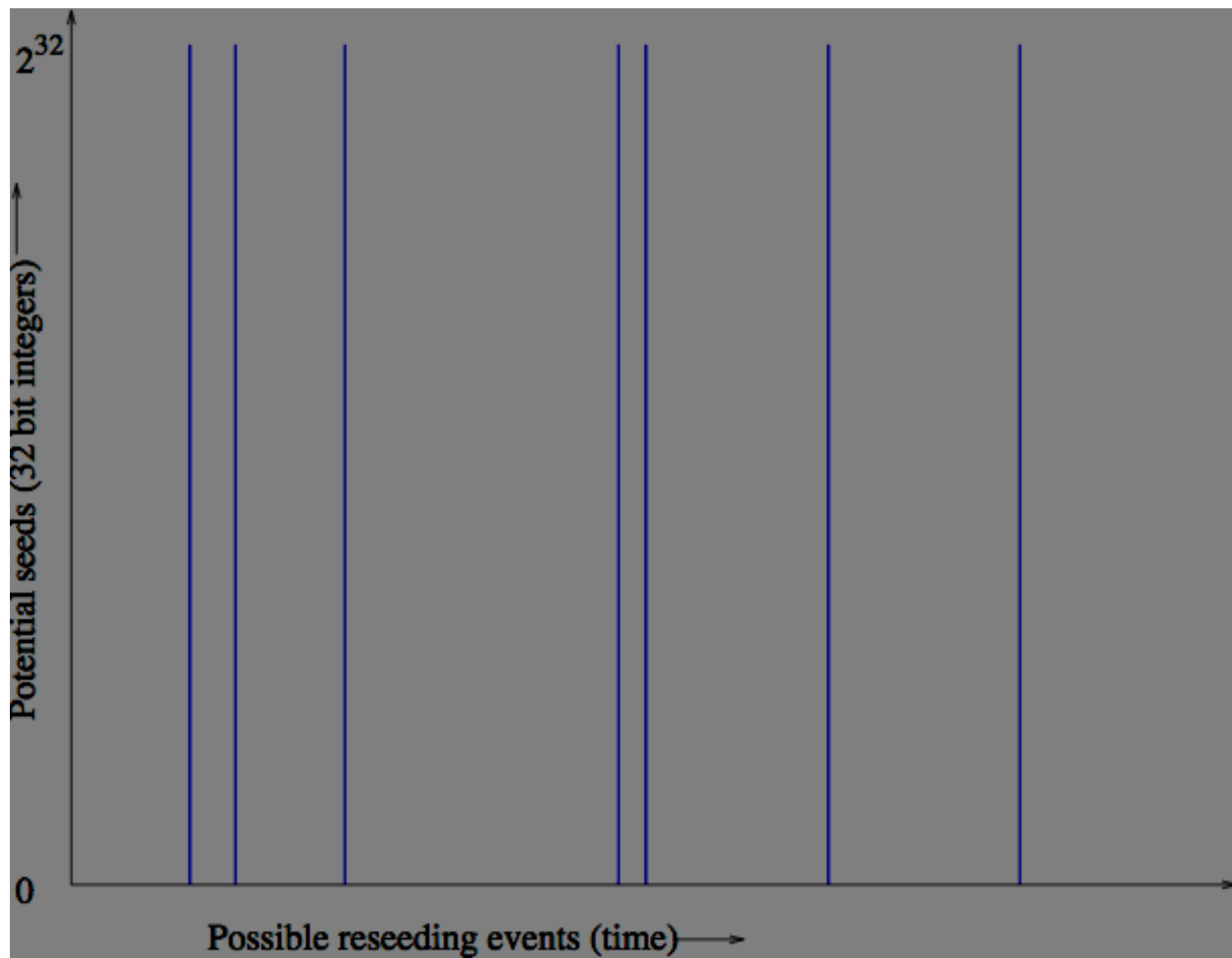
Range where the seed must lie.

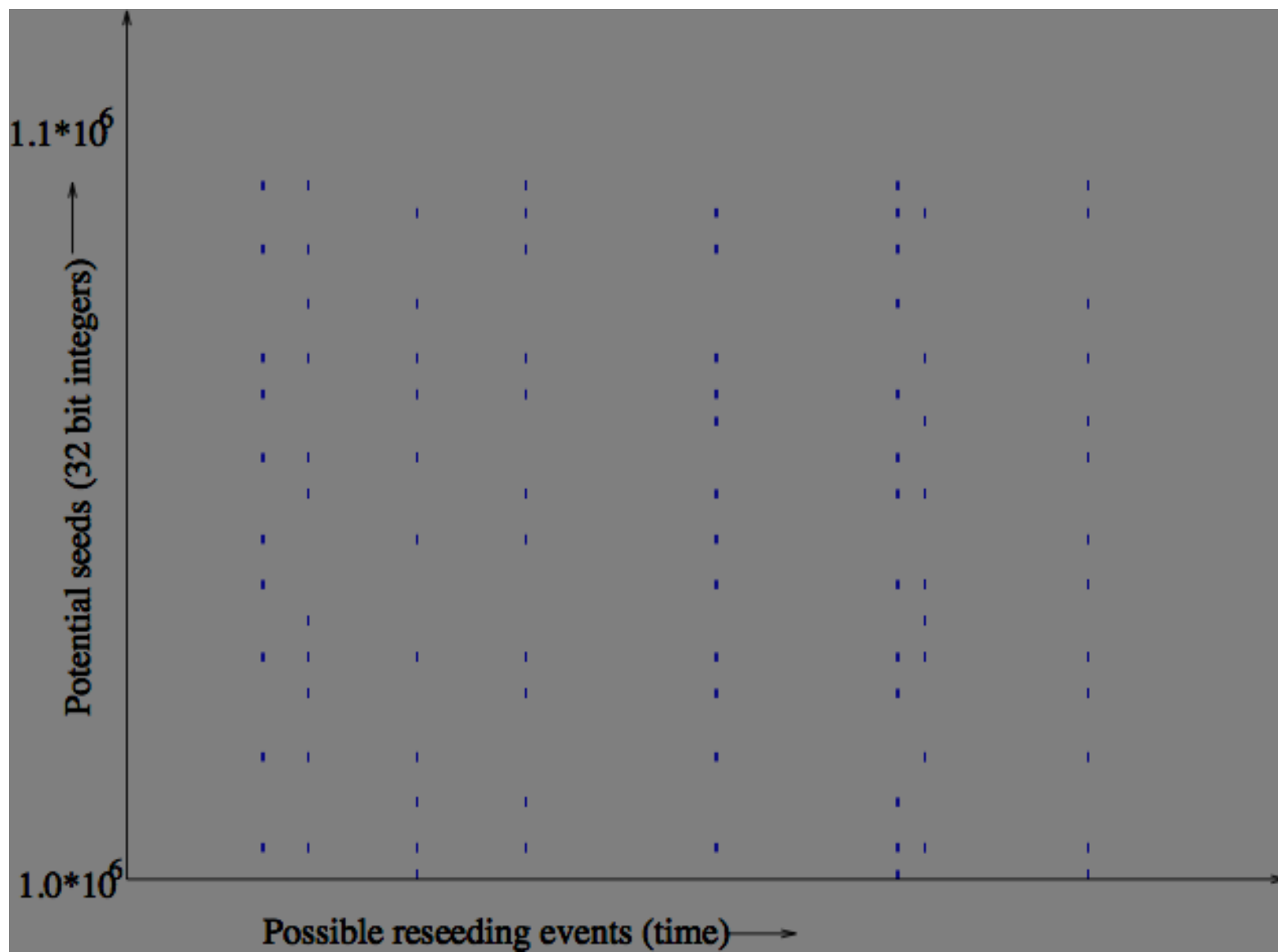


Permutation Space



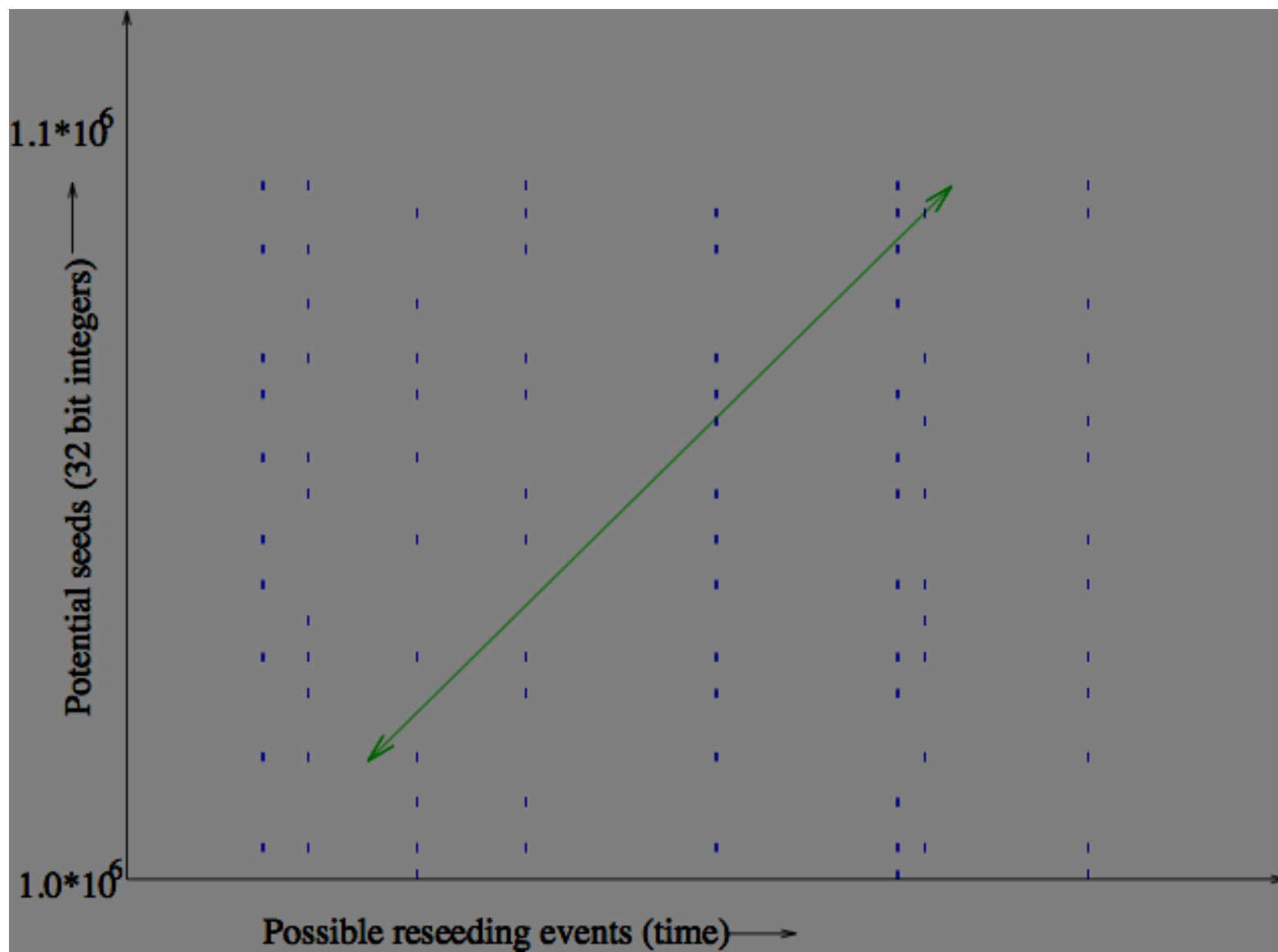
Packets unrelated to predecessors

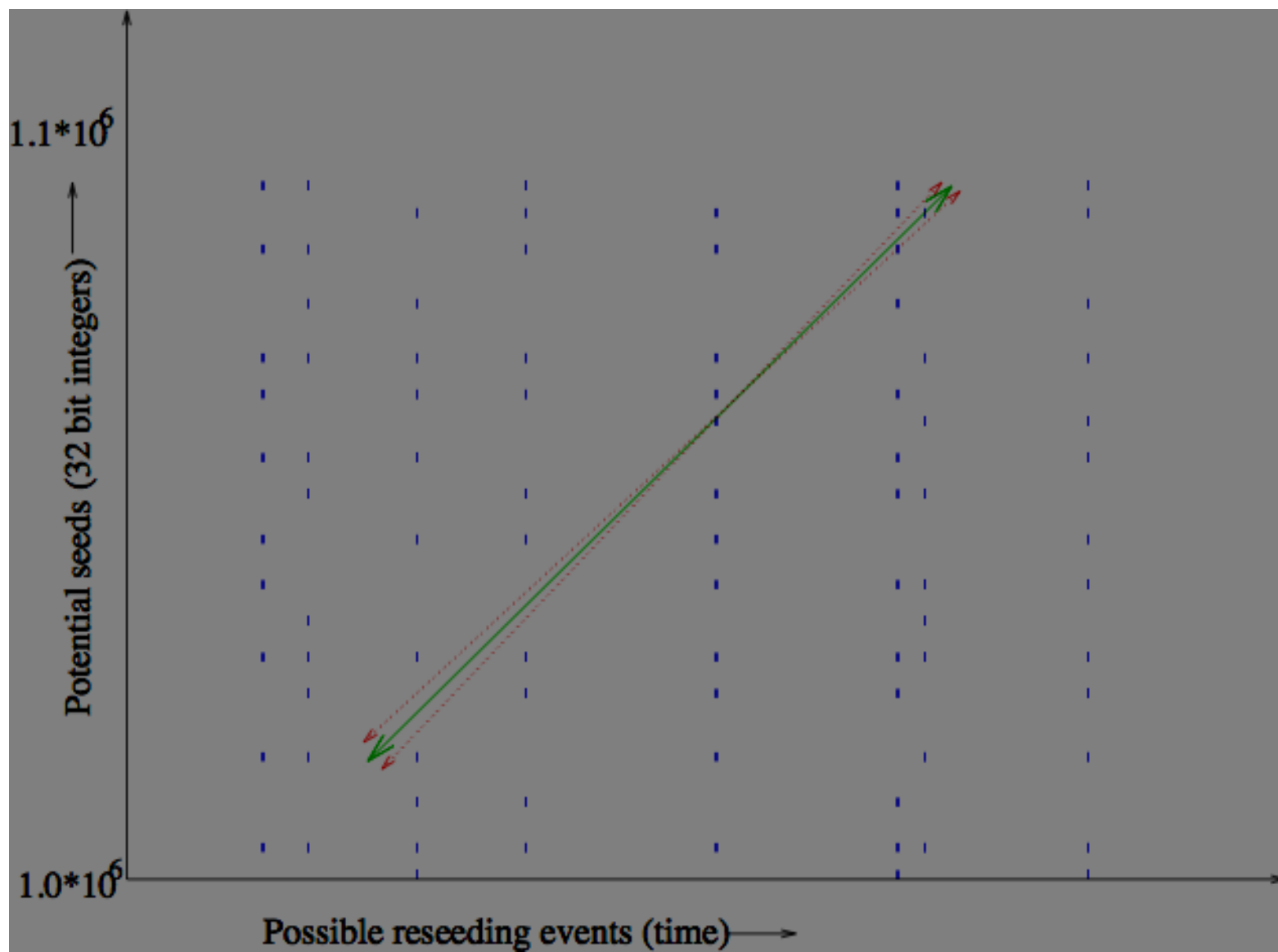


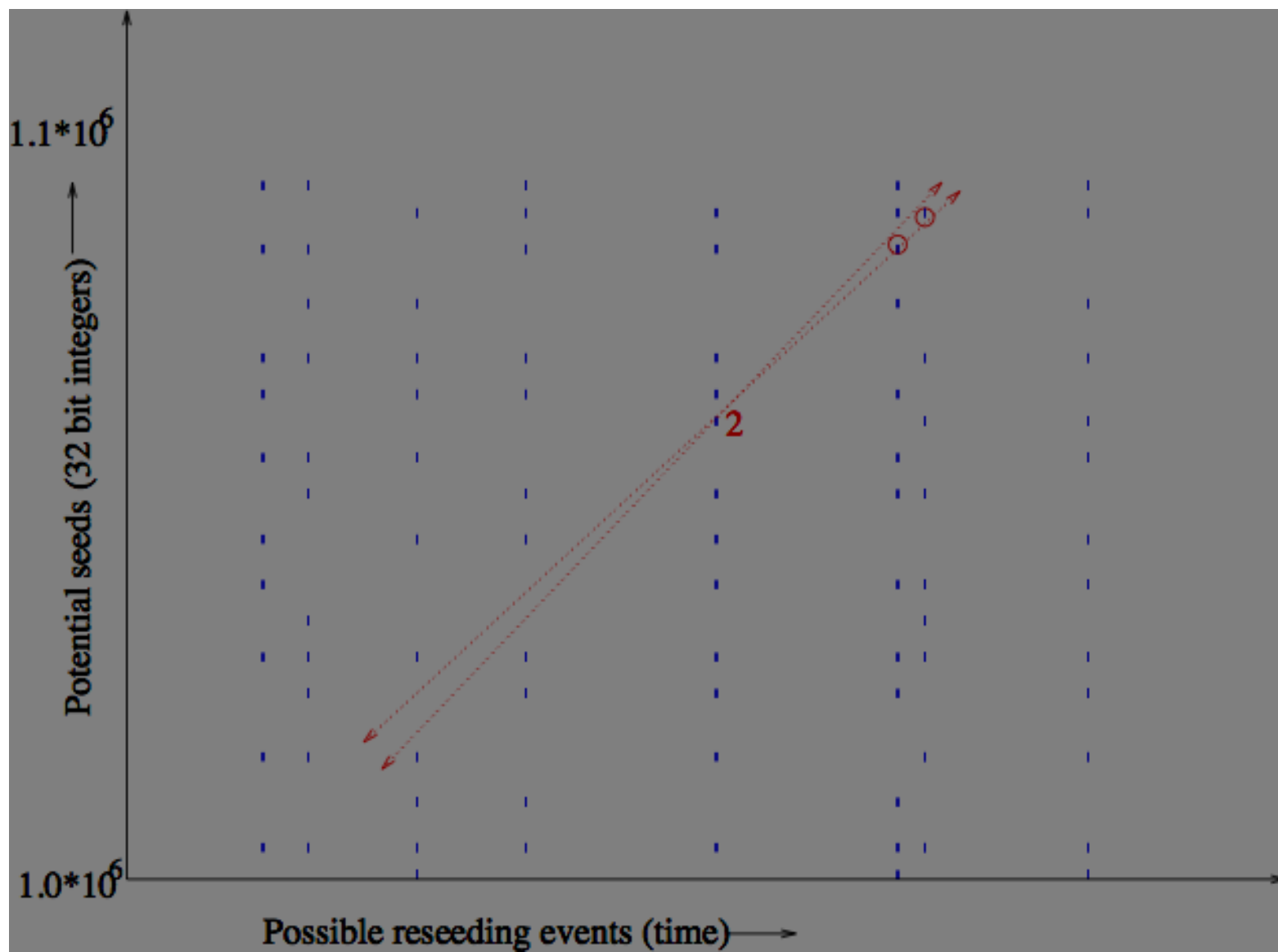


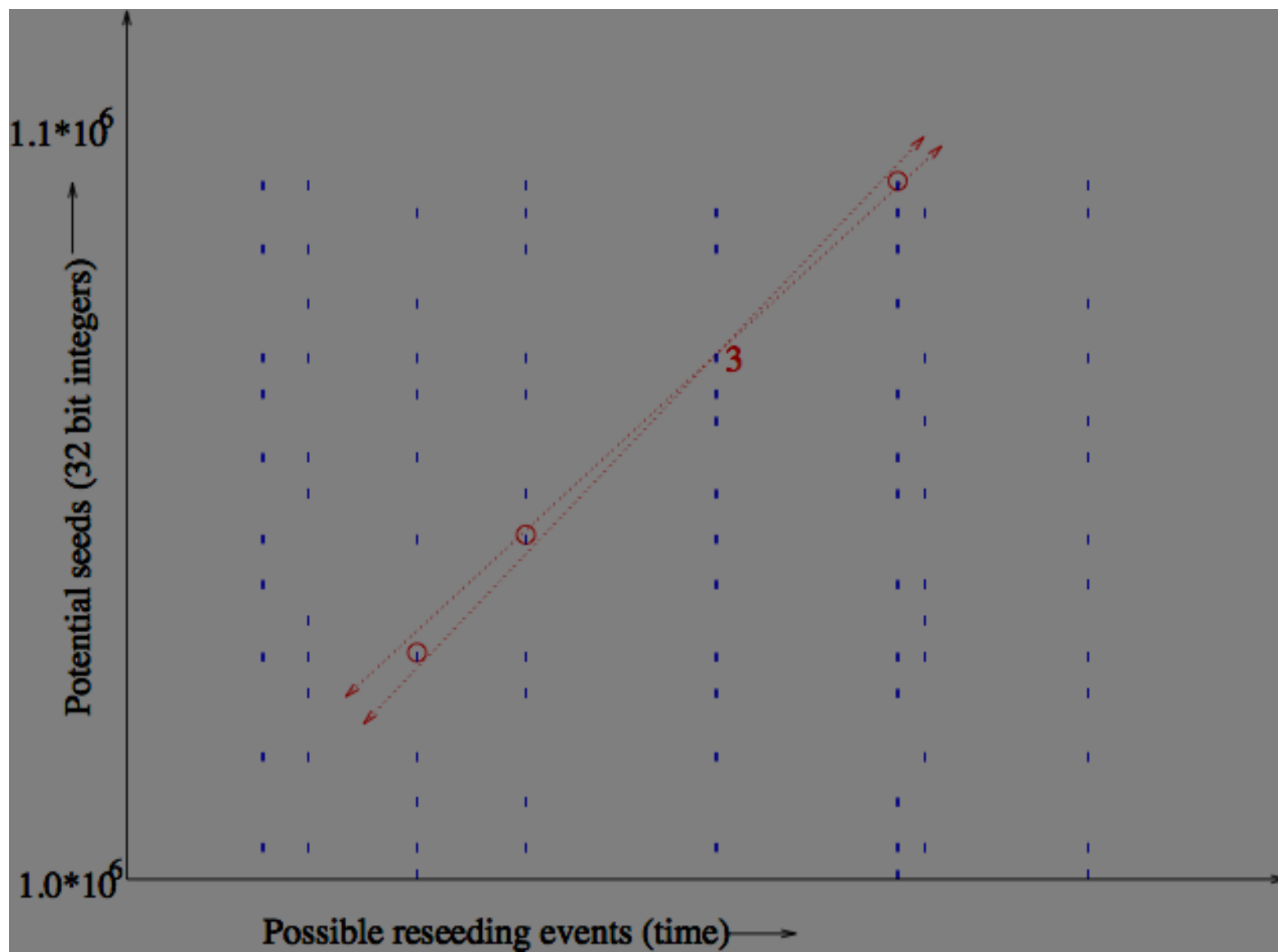
We Know Intervals in Which Each *First-Seed* Packet Occurs

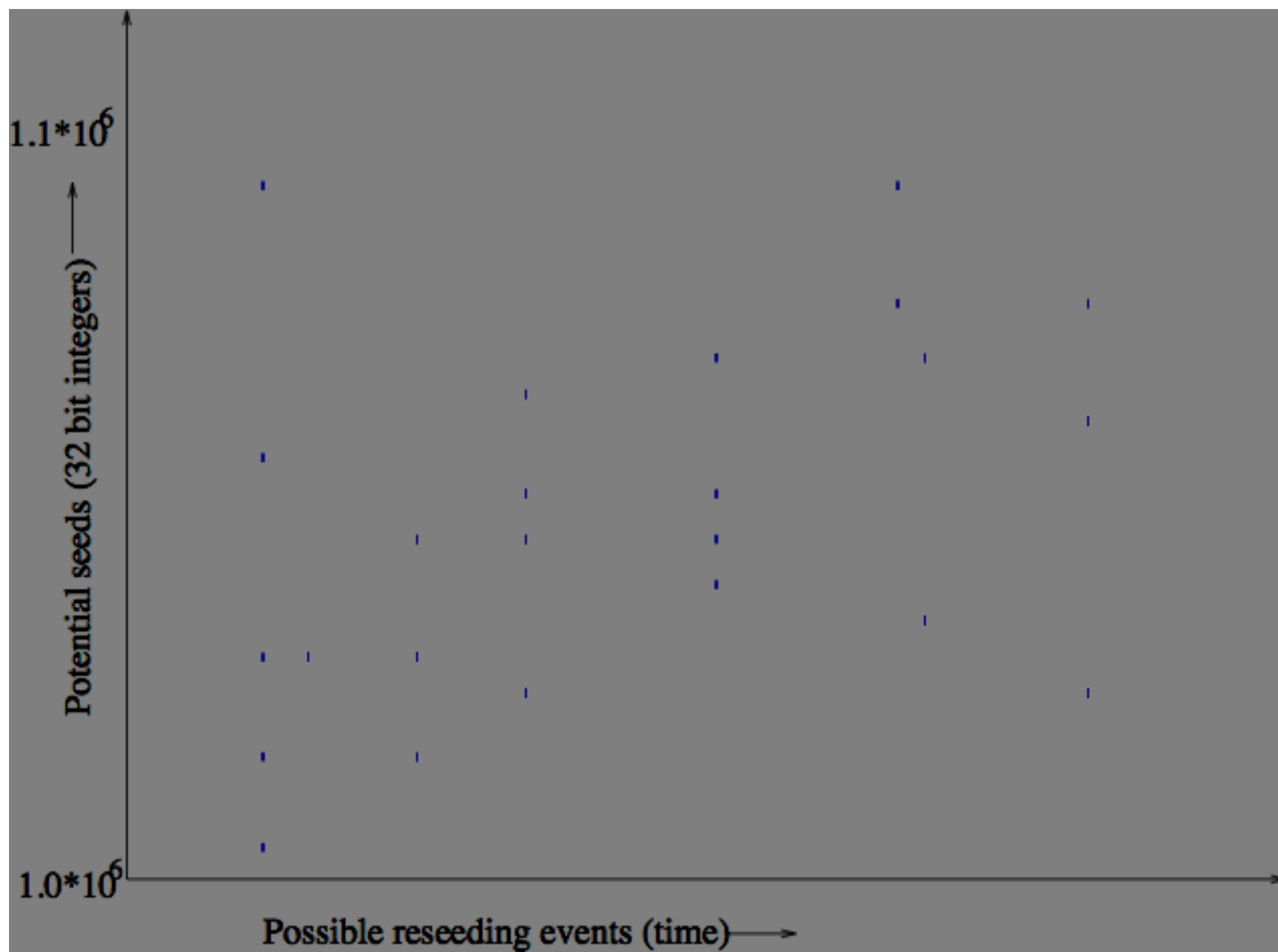
- ... but which among the 1,000s of candidates are the **actual** seeds?
- Entropy isn't all that easy to come by ...
- Consider
srand(get_tick_count())
i.e., uptime in msec
- The values used in repeated calls increase **linearly with time**

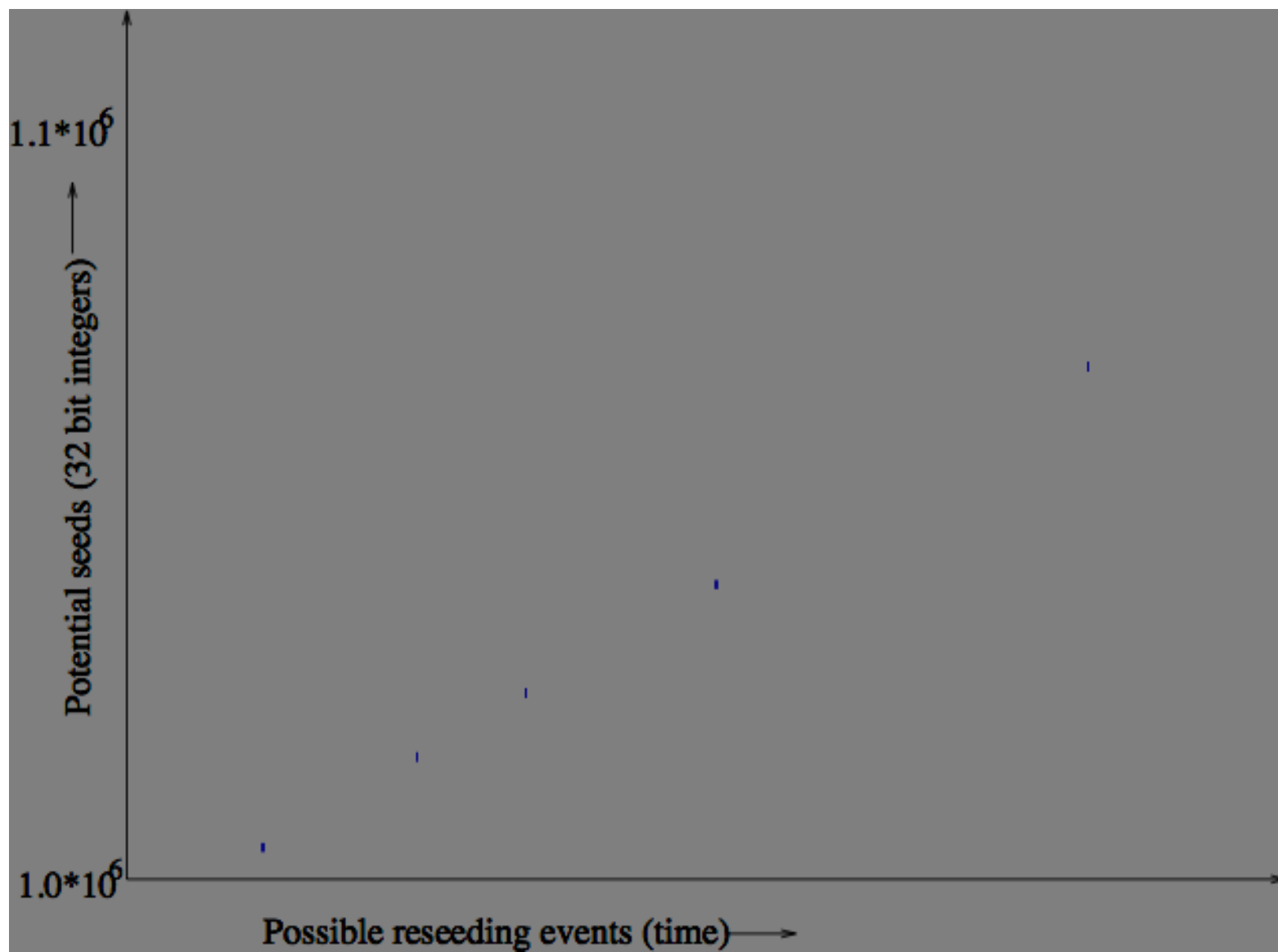


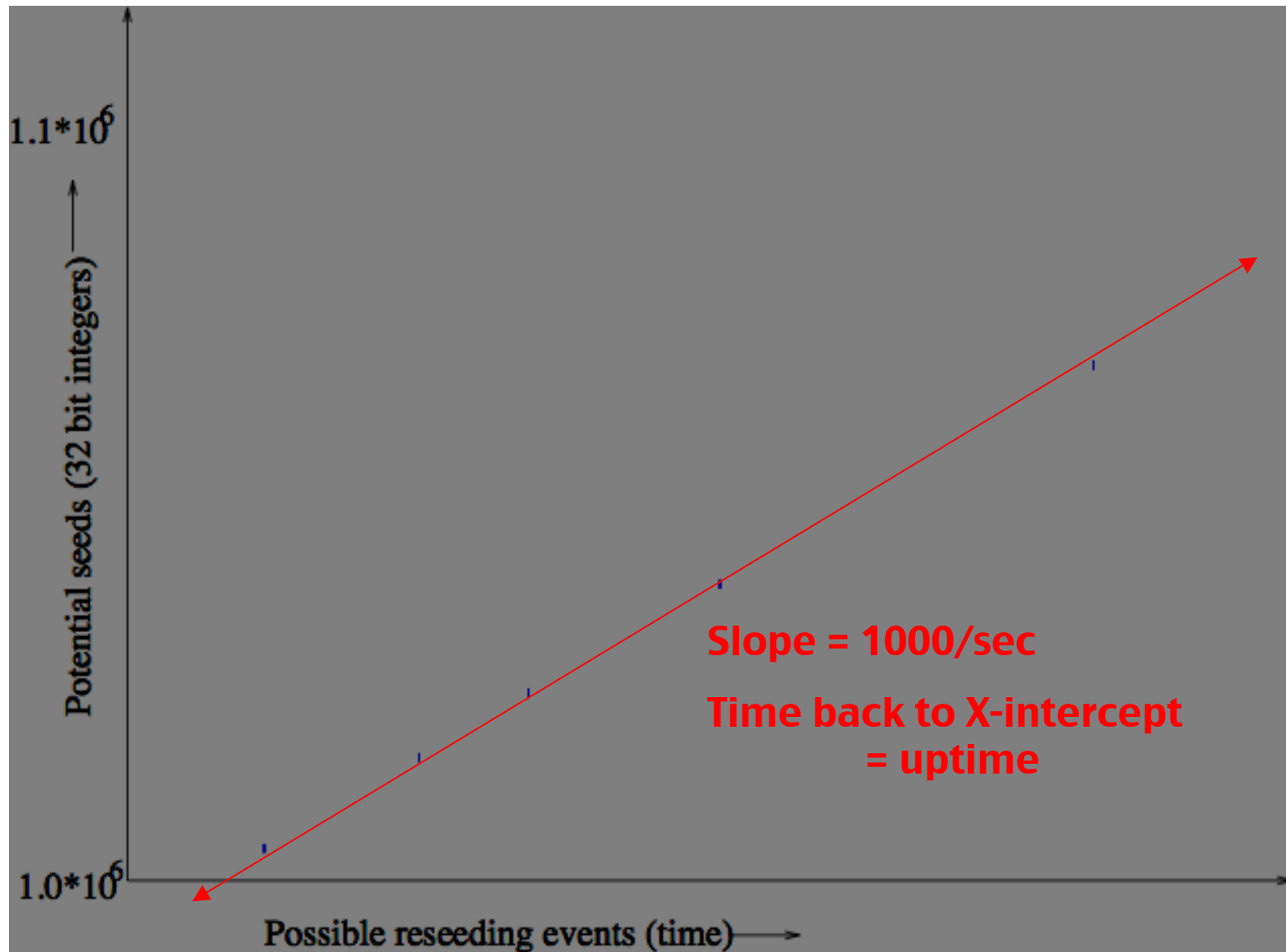




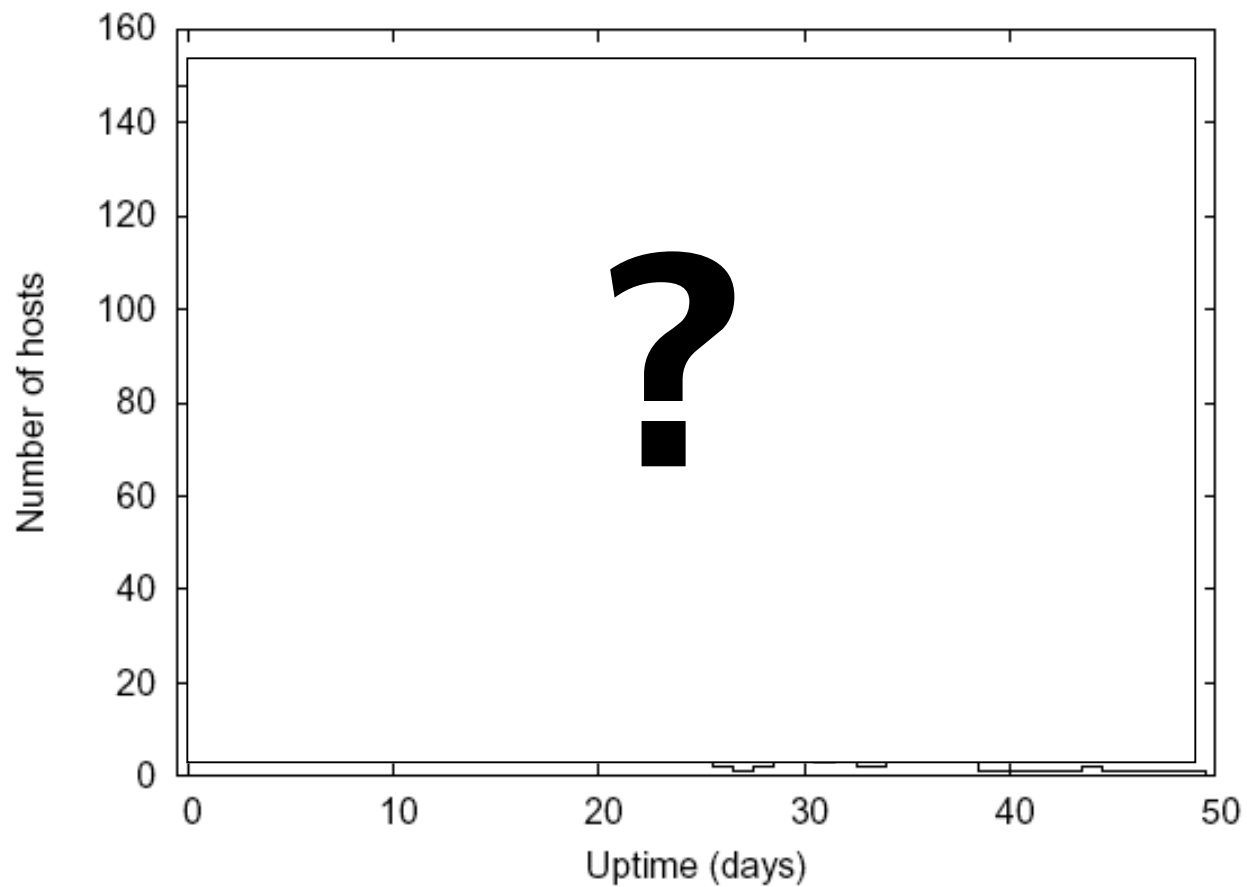




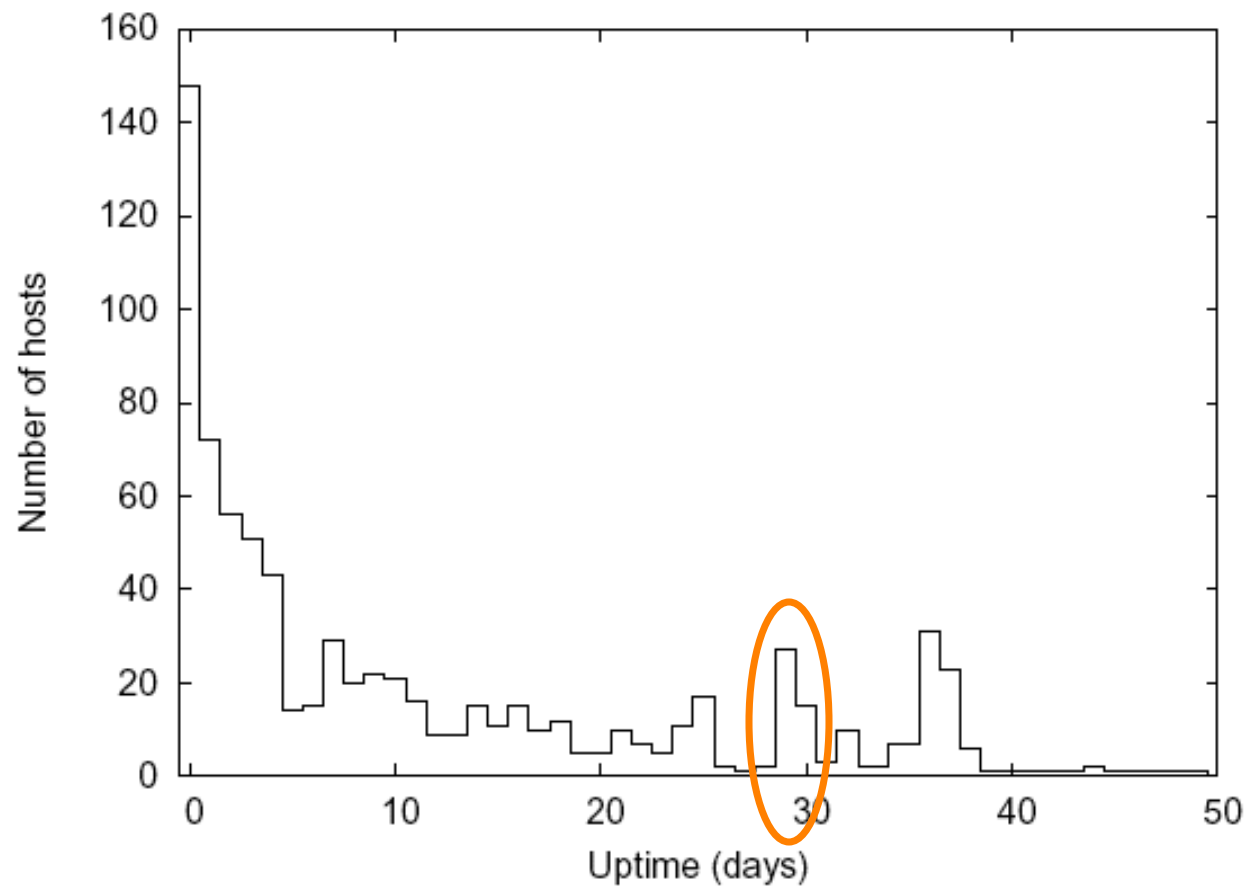




Uptime of 750 Witty Infectees



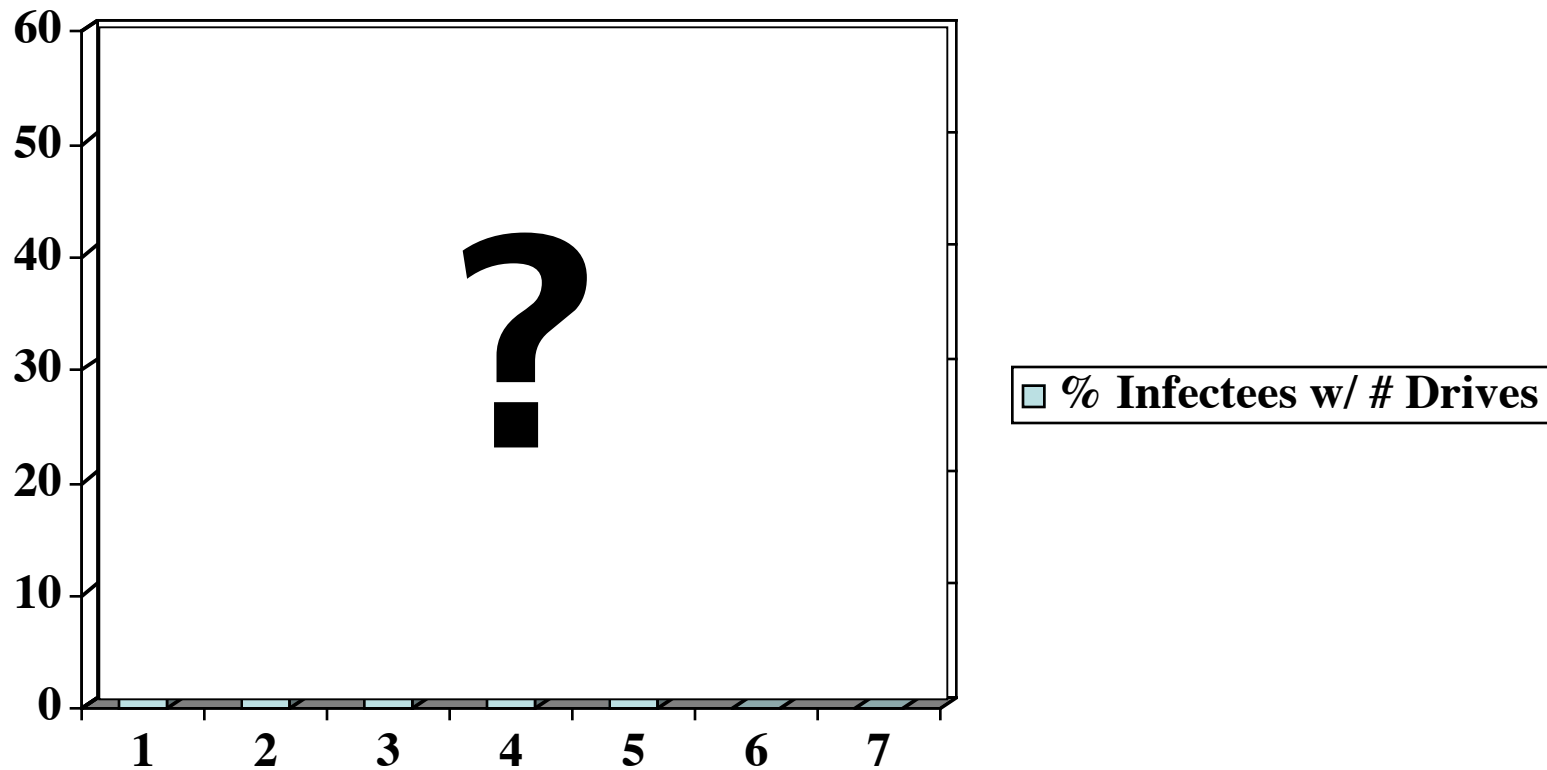
Uptime of 750 Witty Infectees



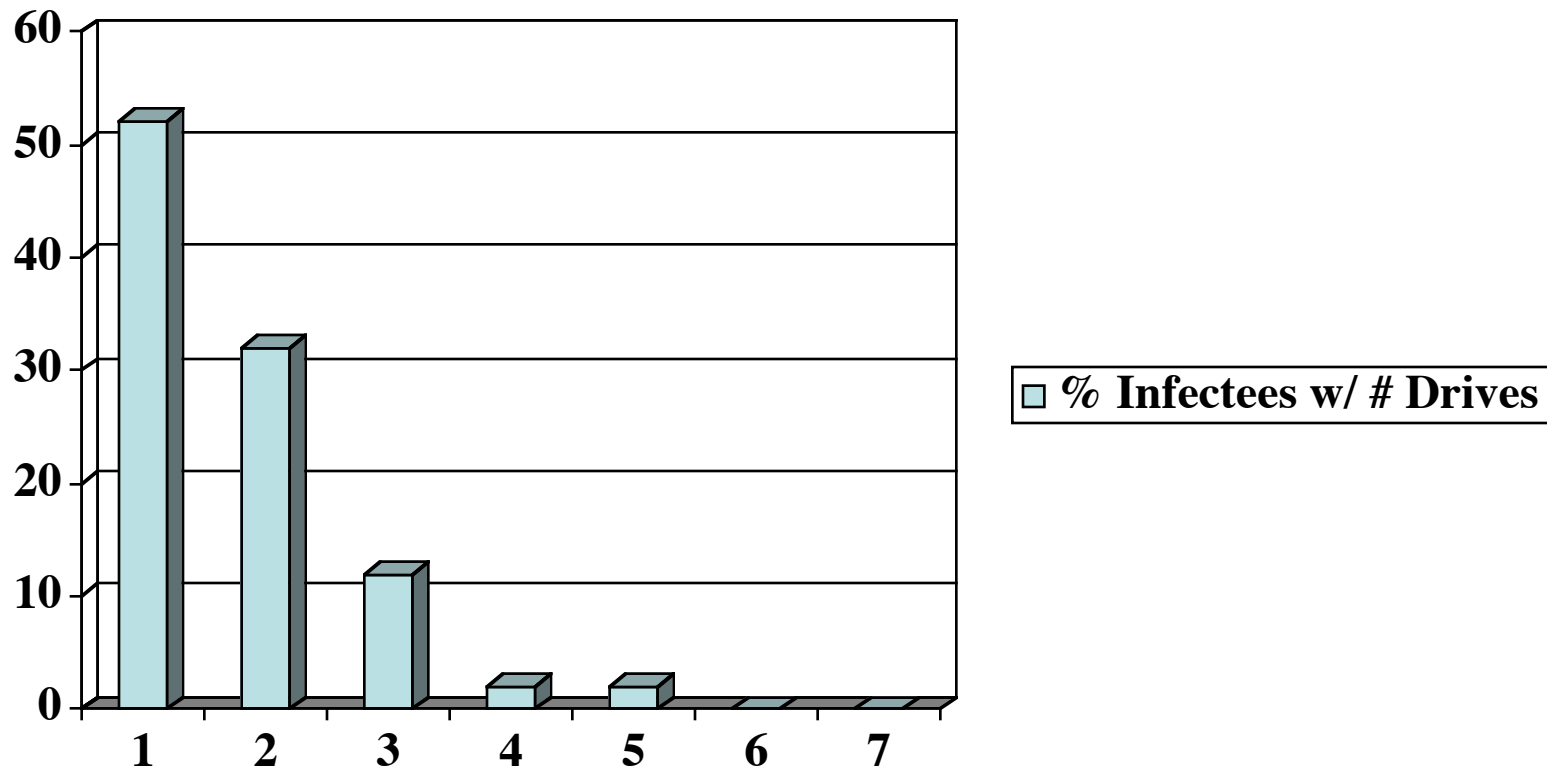
Given Exact Values of Seeds Used for Reseeding ...

- ... we know **exact** random # used at each subsequent disk-wipe test:
`if(open_physical_disk(rand()[13..15])`
- ... and its success, or failure, i.e., **number of drives attached** to each infectee ...

Disk Drives Per Witty Infectee



Disk Drives Per Witty Infectee



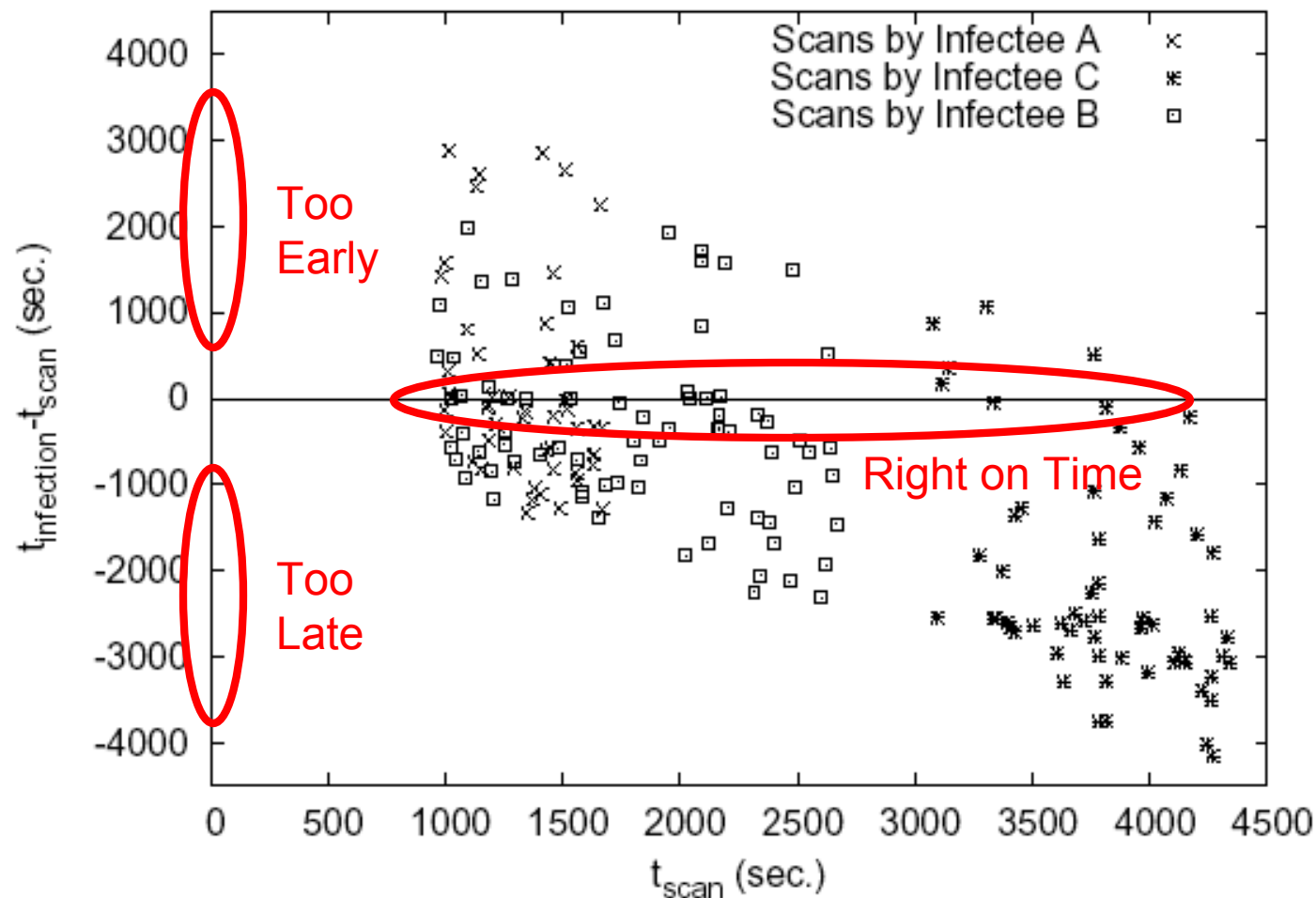
Given Exact Values of Seeds Used for Reseeding ...

- ... we know exact random # used at each subsequent disk-wipe test:

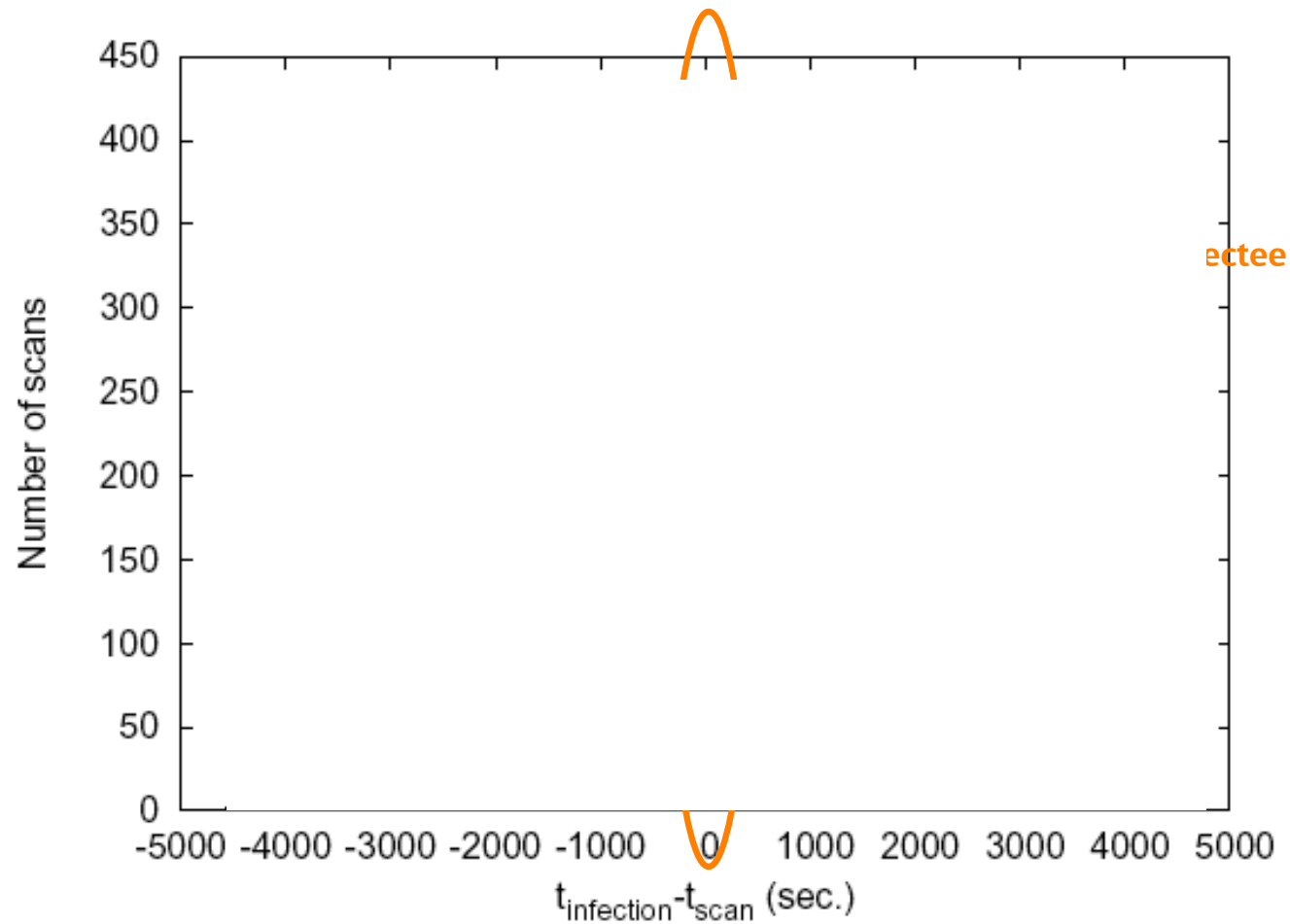
if(open_physical_disk(**rand()**_[13..15])

- ... and its success, or failure, i.e., number of drives attached to each infectee ...
- ... and, more, generally, *every packet each infectee sent*
 - Can compare this to when new infectees show up
 - i.e., *Who-Infected-Whom*

Time Between Scan by Known Infectee and New Source Arrival At Telescope



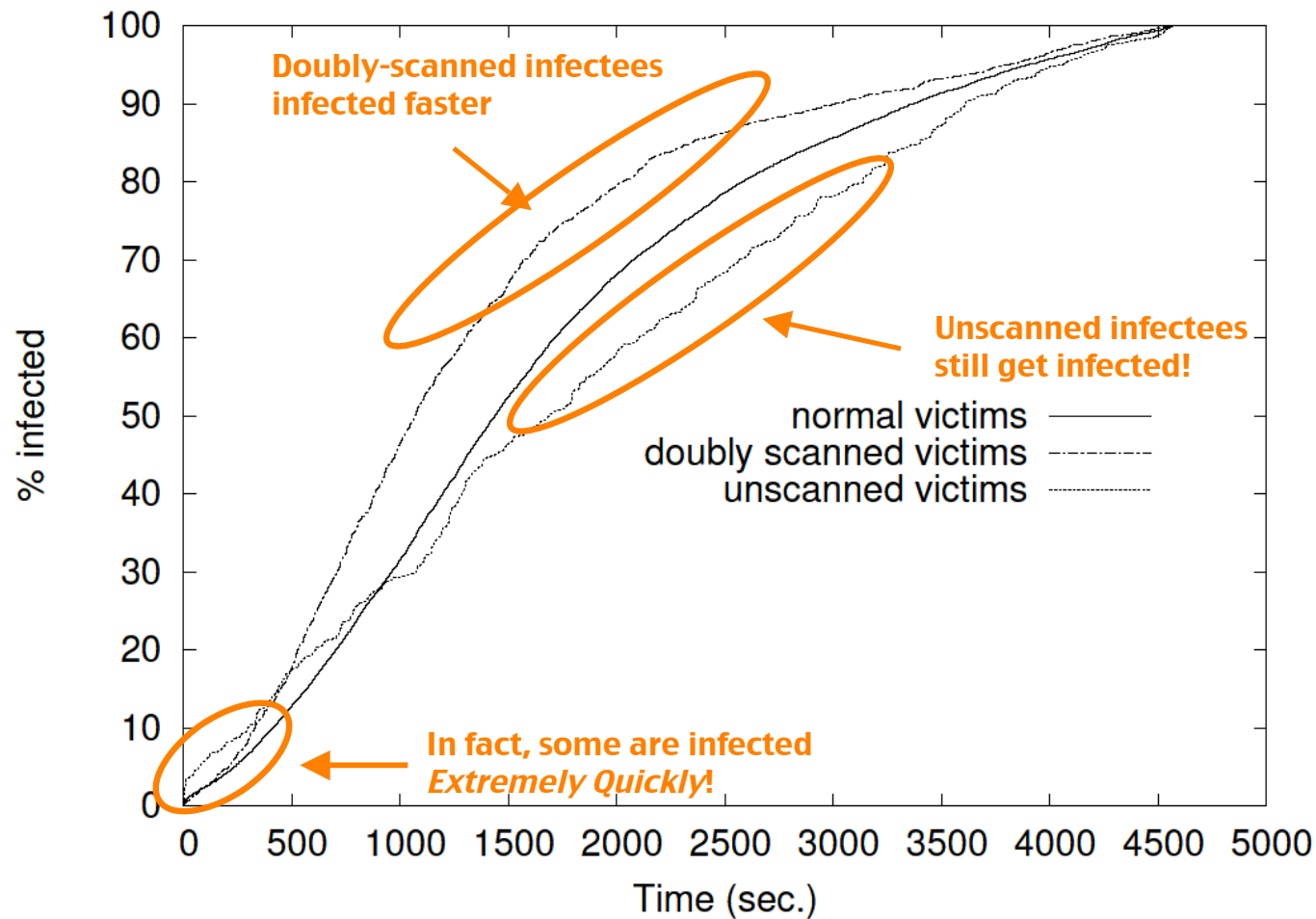
Infection Attempts That Were Too Early, Too Late, or *Just Right*



Witty is Incomplete

- Recall that LCD PRNG generates a **complete** orbit over a permutation of $0..2^{32}-1$.
- **But:** Witty author **didn't use all 32 bits** of single PRNG value
 - $dest_ip \leftarrow (X_i)_{[0..15]} \parallel (X_{i+1})_{[0..15]}$
 - Knuth recommends top bits as having better pseudo-random properties
- **But²:** This does **not** generate a complete orbit!
 - Misses 10% of the address space
 - Visits 10% of the addresses (exactly) twice
- So: were 10% of the potential infectees **protected**?

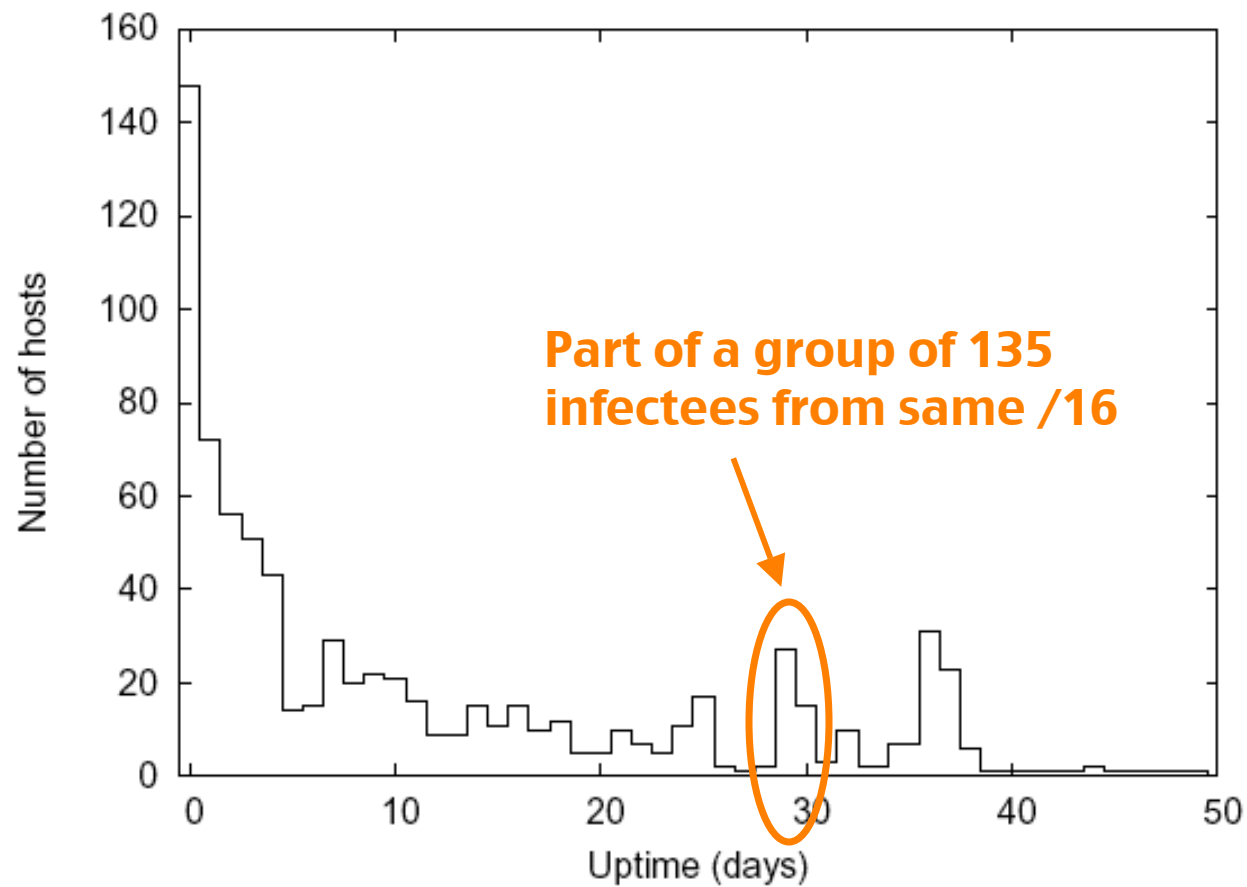
Time When Infectees Seen At Telescope



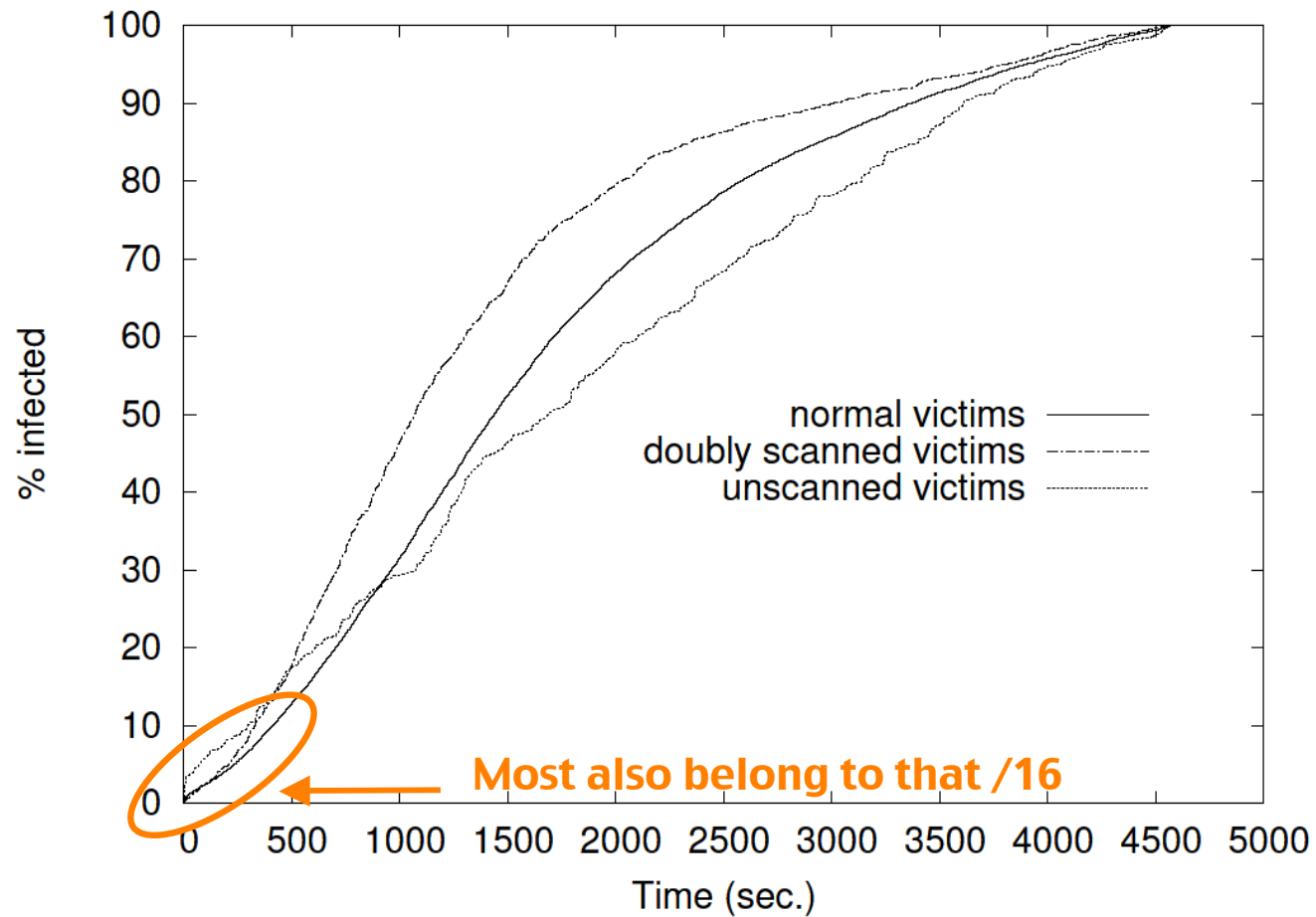
How Can an Unscanned Infectee Become Infected?

- Multihomed host infected via another address
 - Might show up with normal speed, but not *early*
- DHCP or NAT aliasing
 - Would show up *late*, certainly not *early*
- Could they have been *passively infected* extra quickly because they had *large cross-sections*?
- Just what are those hosts, anyway?

Uptime of 750 Witty Infectees



Time When Infectees Seen At Telescope



Analysis of the Extra-Quick Hosts

- Initial infectees exhibit super-exponential growth ⇒ they **weren't found by random scanning**
 - Hosts in prevalent /16 numbered x.y.z.**4** in consecutive /24 subnets
 - “Lineage” analysis reveals that these subnets *not* sufficiently visited at onset to account for infection
 - One possibility: they monitored networks *separate* from their own subnet
 - But: if so, strange to number each **.4** in adjacent subnets ...
- ⇒ Unlikely infection was due to passive monitoring ...

Alternative: Witty Started With A “Hit List”

- ...Unlikely infection was due to passive monitoring ...
- Prevalent /16 = **U.S. military base**
- Attacker knew of ISS security software installation at military site \Rightarrow ***ISS insider***
(or *ex-insider*)
- Fits with **very rapid development** of worm after public vulnerability disclosure

Are All The Worms In Fact Executing Witty?

- Answer: **No**
 - There is *one* “infectee” that probes addresses **not on the orbit**.
 - Each probe contains Witty contagion, but lacks randomized payload size.
 - Shows up very near beginning of trace.
- ⇒ *Patient Zero* - machine attacker used to launch Witty. (Really, *Patient Negative One*.)
- European retail ISP
 - Information passed along to Law Enforcement

Summary of Witty Telescope Forensics

- Understanding a measurement's **underlying structure** adds enormous analytic power
- Cuts both ways: makes **anonymization** much harder than one would think
- With enough effort, worm “**attribution**” can be possible
 - But: a **lot** of work
 - And: **no guarantee of success**