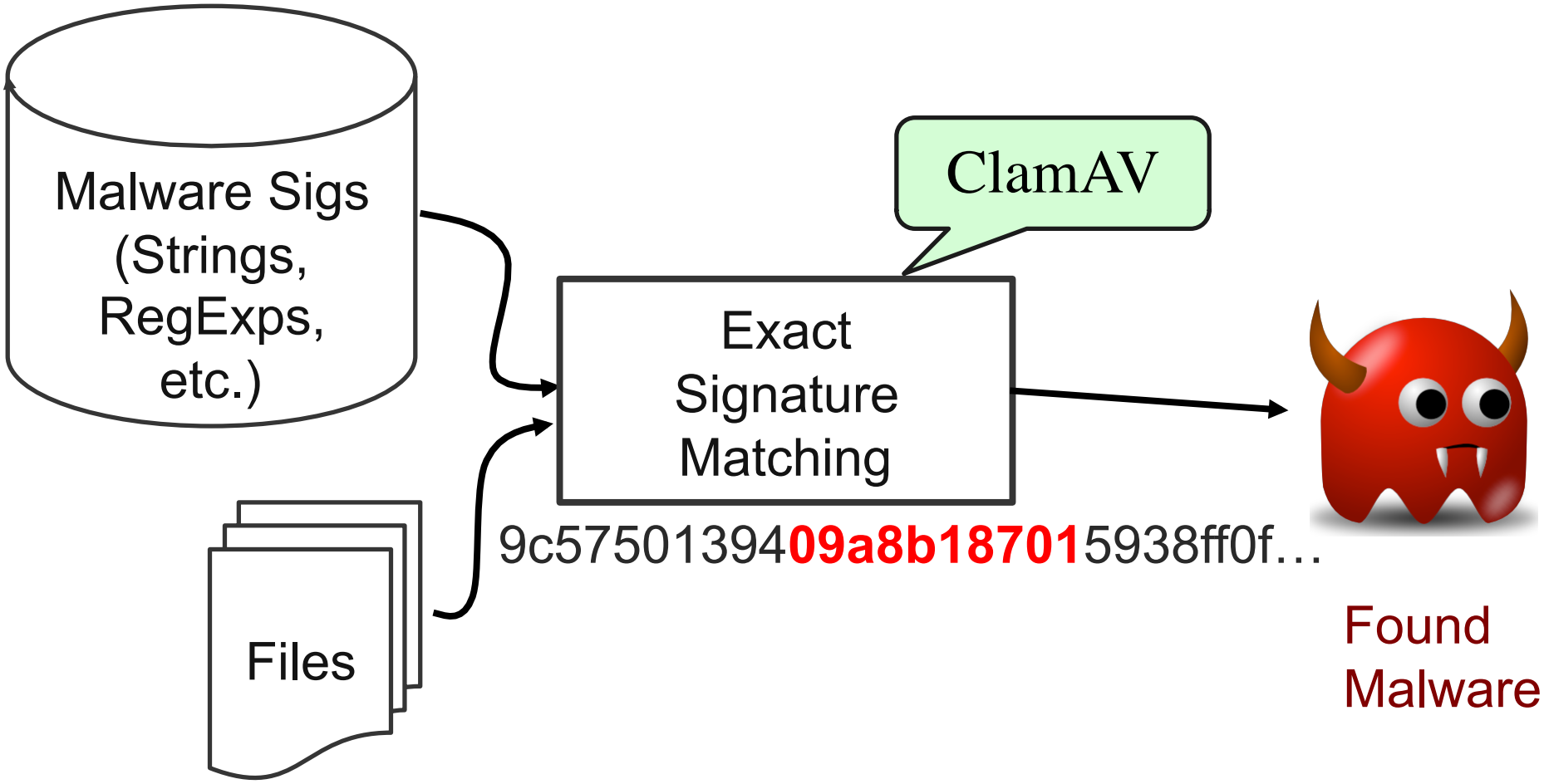
SplitScreen: Enabling Efficient, Distributed Malware Detection

Sang Kil Cha, Iulian Moraru, Jiyong Jang,
John Truelove, David Brumley, David G. Andersen

Carnegie Mellon University

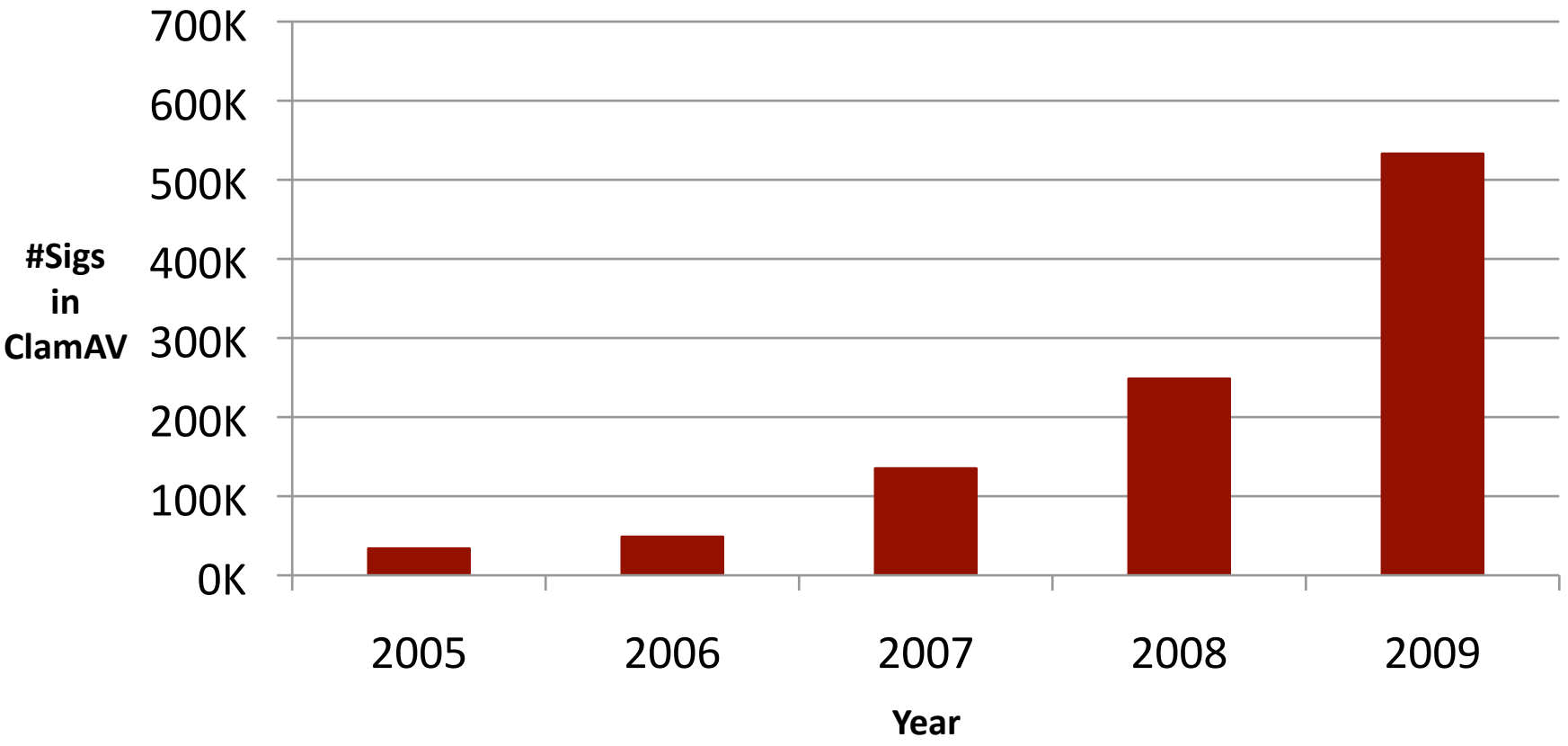
Malware Scanning



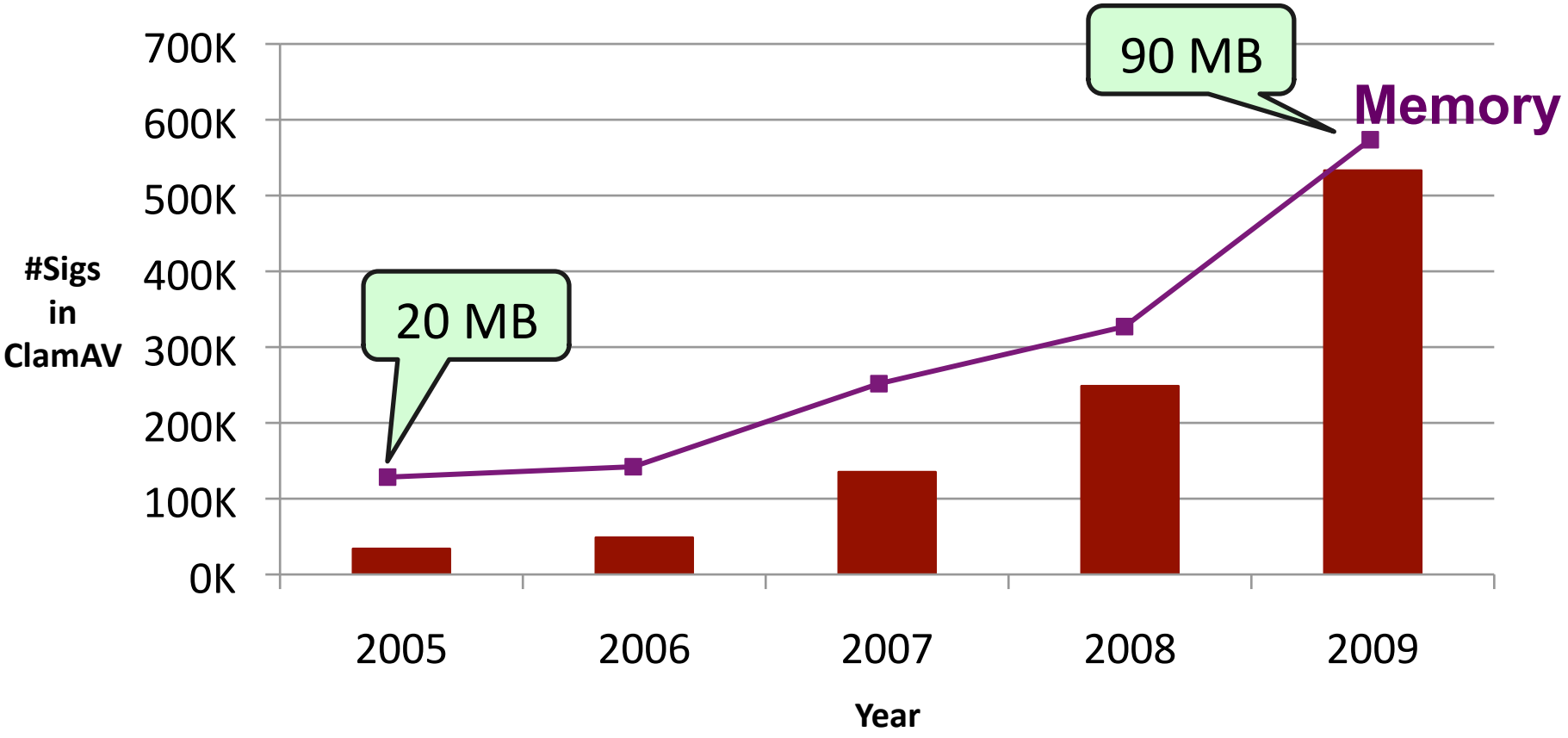
Signature-based Scanning

- **Currently fastest method**
  - Emerging alternatives slower (e.g., behavior- based)
  - Signature scanning likely part of practical solutions
- **Widely Deployed**
  - $2 billion industry
  - Symantec, Trend Micro, ClamAV, etc. all use signature-based scanning
- **Millions of existing signatures**

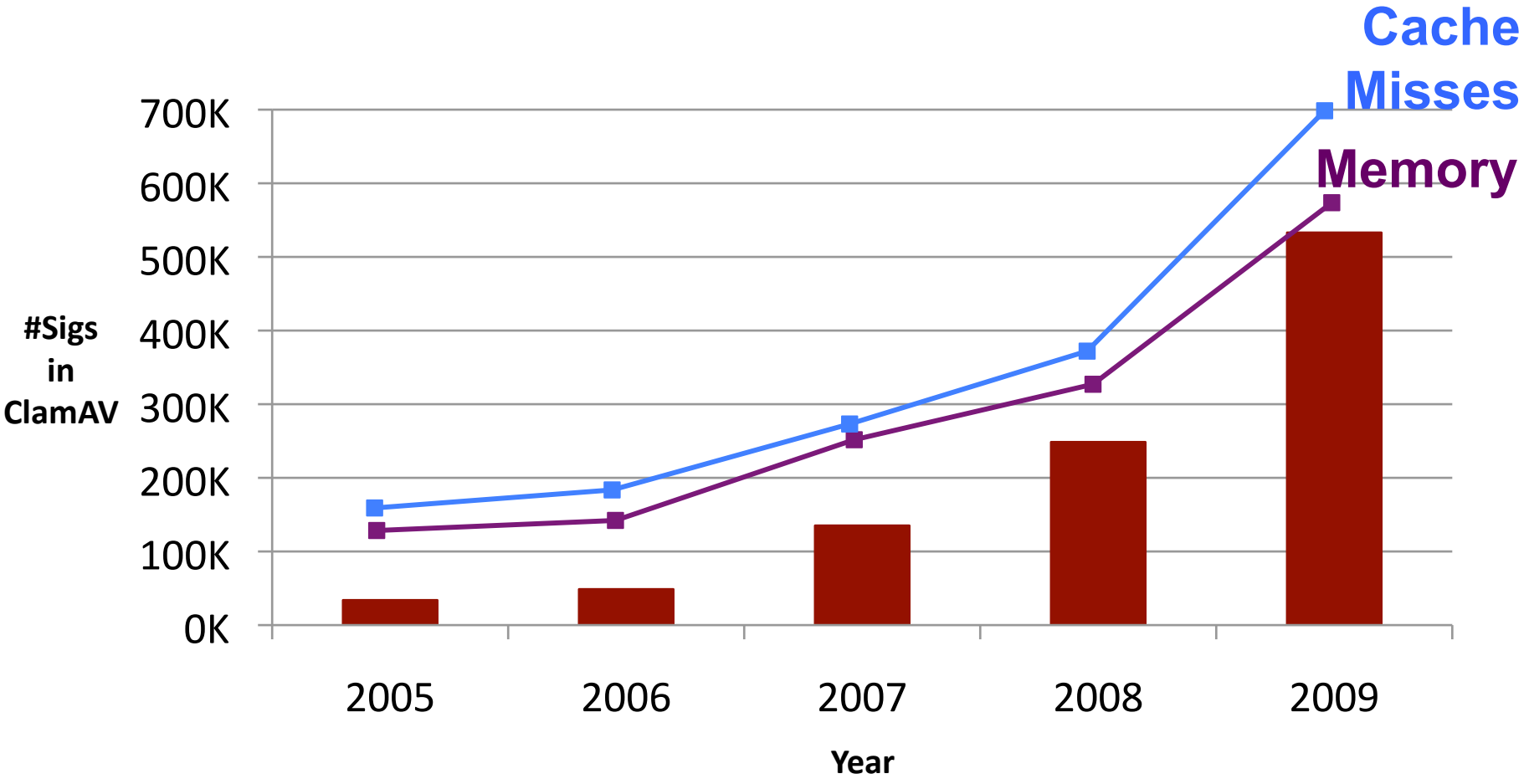
Signature Count is exploding



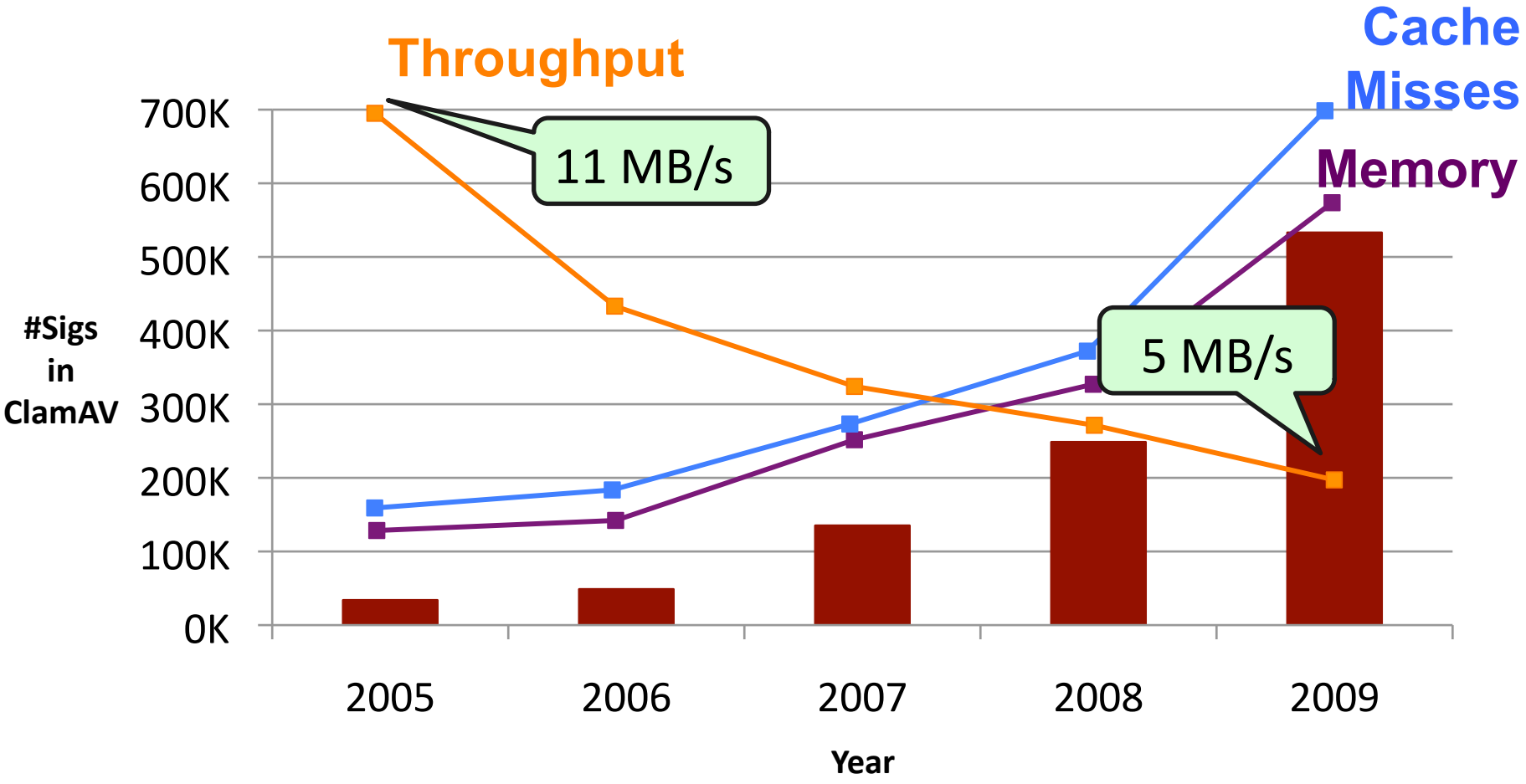
More Signatures = More Memory!



More Signatures
= Poor Cache Performance!



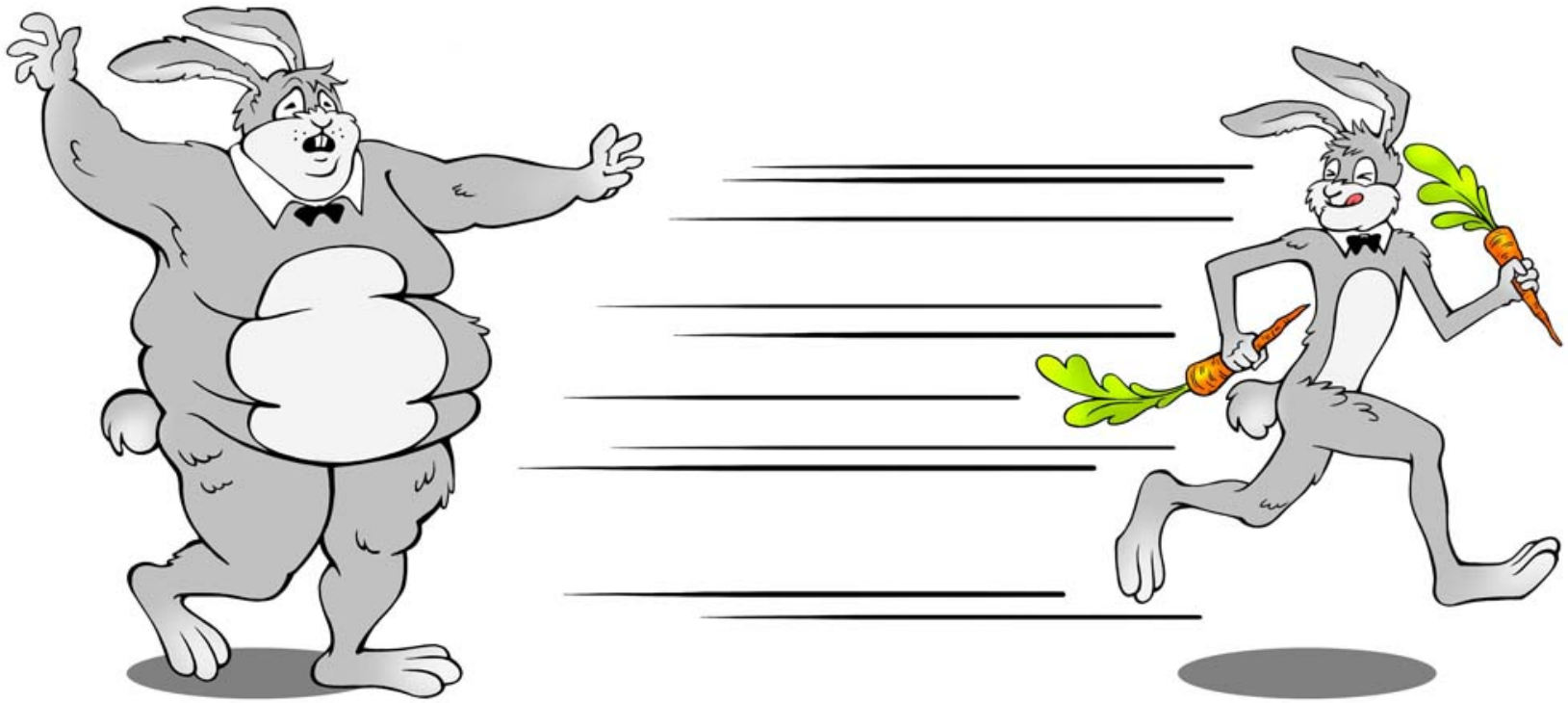
More Signatures = Slower!



SplitScreen:

≥ 2x the speed,

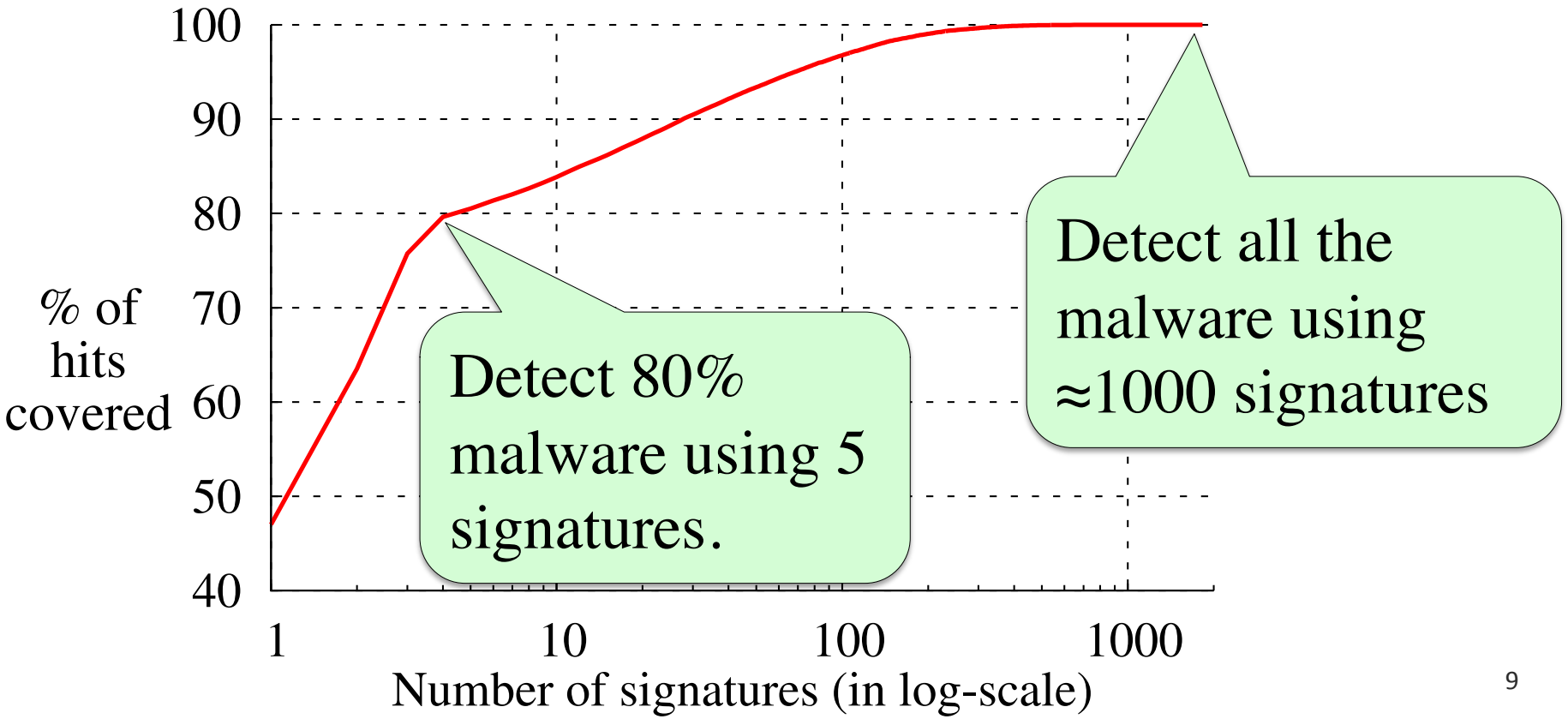
≤ ½ the memory.



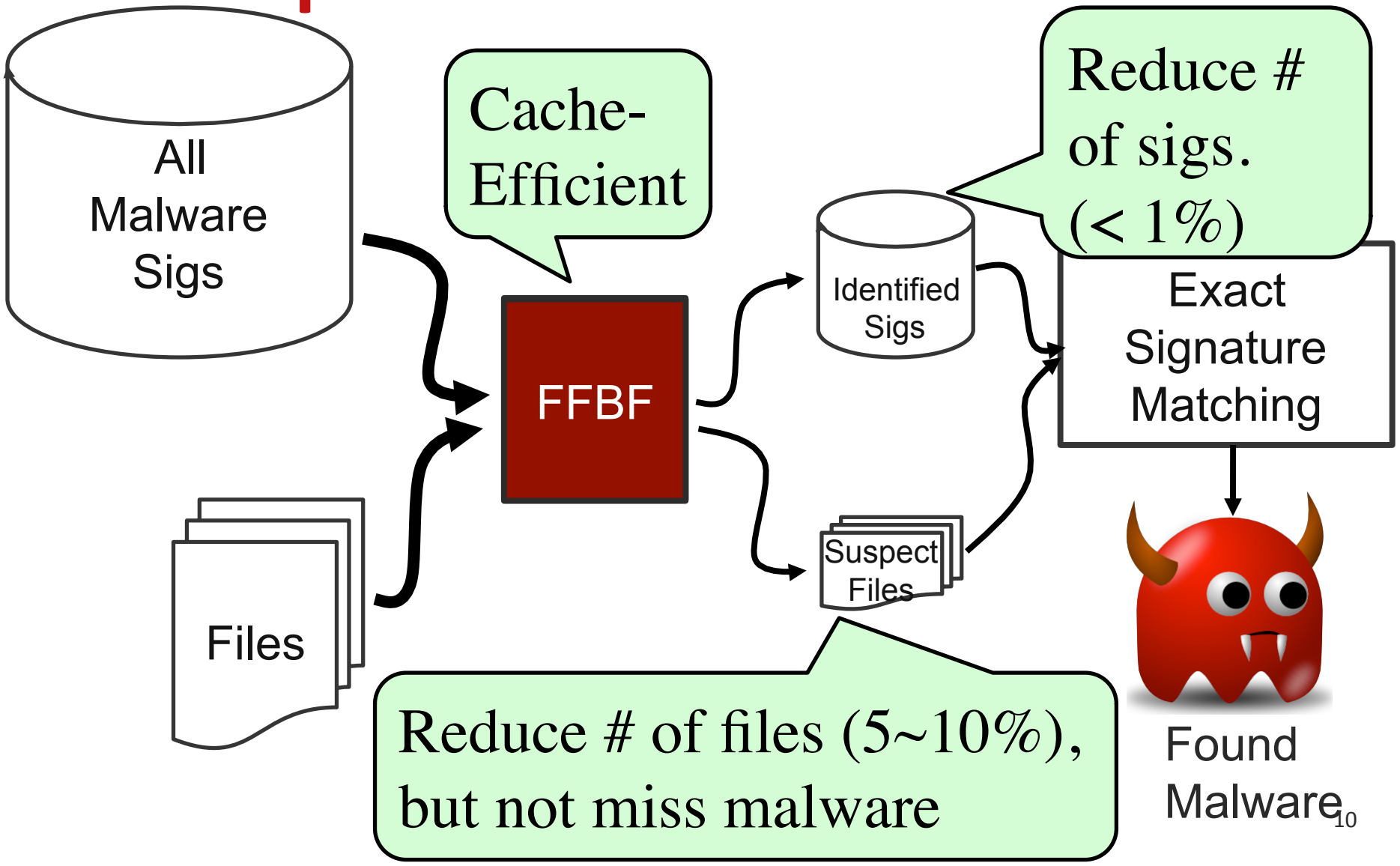
Opportunity: Few Signatures Matched

4 month study of CMU email malware

< 1% of signatures used by ClamAV for all malware



SplitScreen Architecture

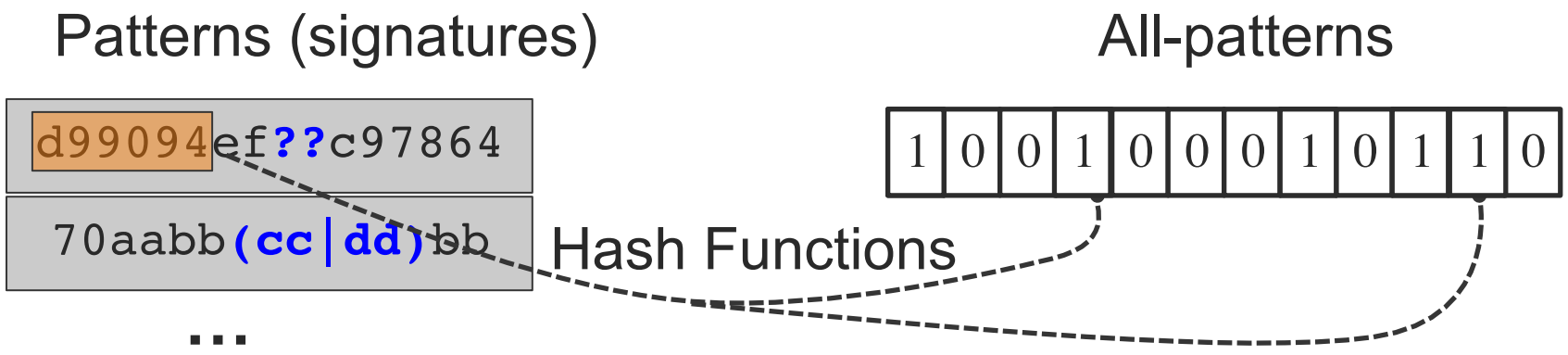


SplitScreen Design

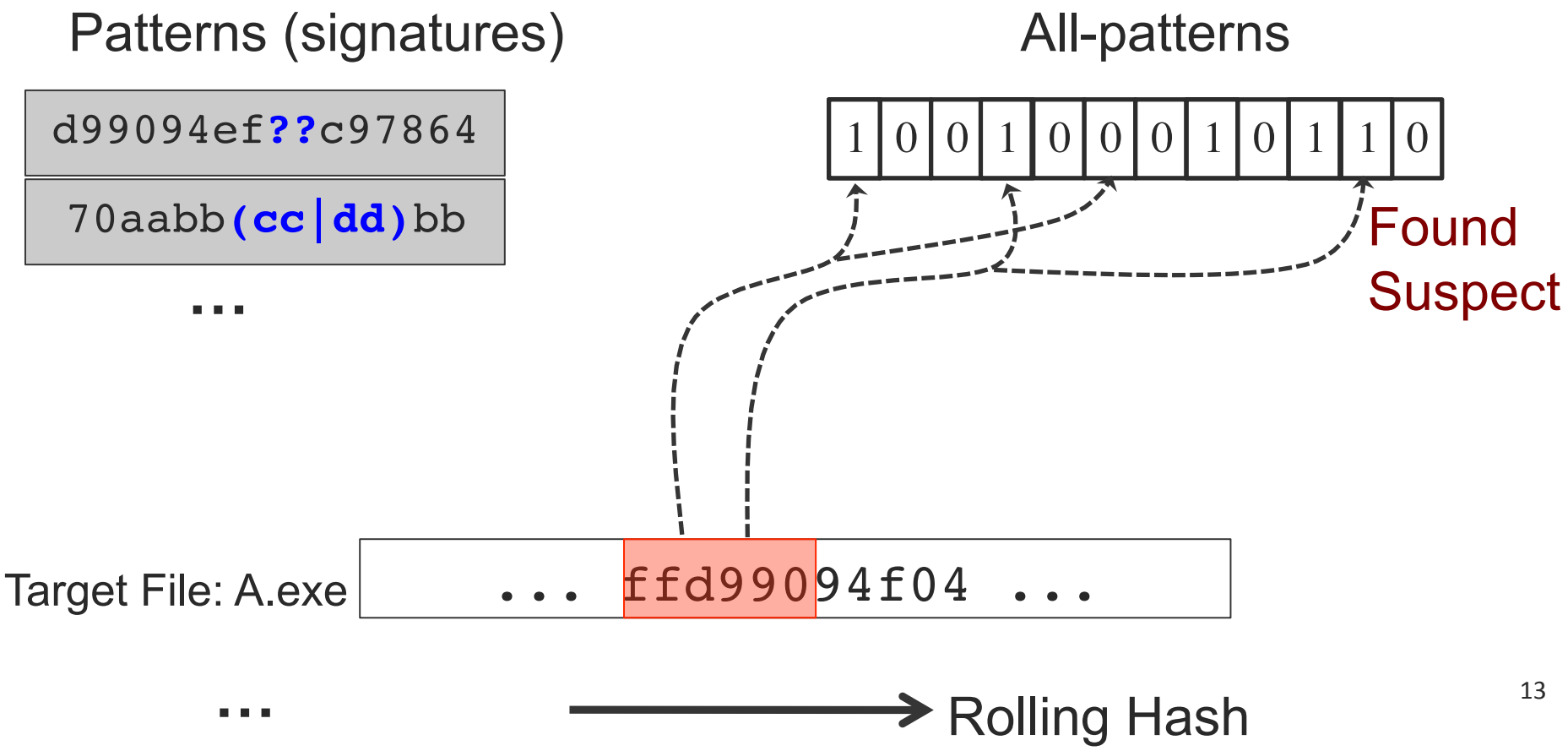
- Feed-Forward Bloom Filter (FFBF)
  - FFBF Initialization
  - FFBF Scanning: File filtering
  - FFBF Scanning: Pattern filtering

- Cache-efficient Bloom filter design

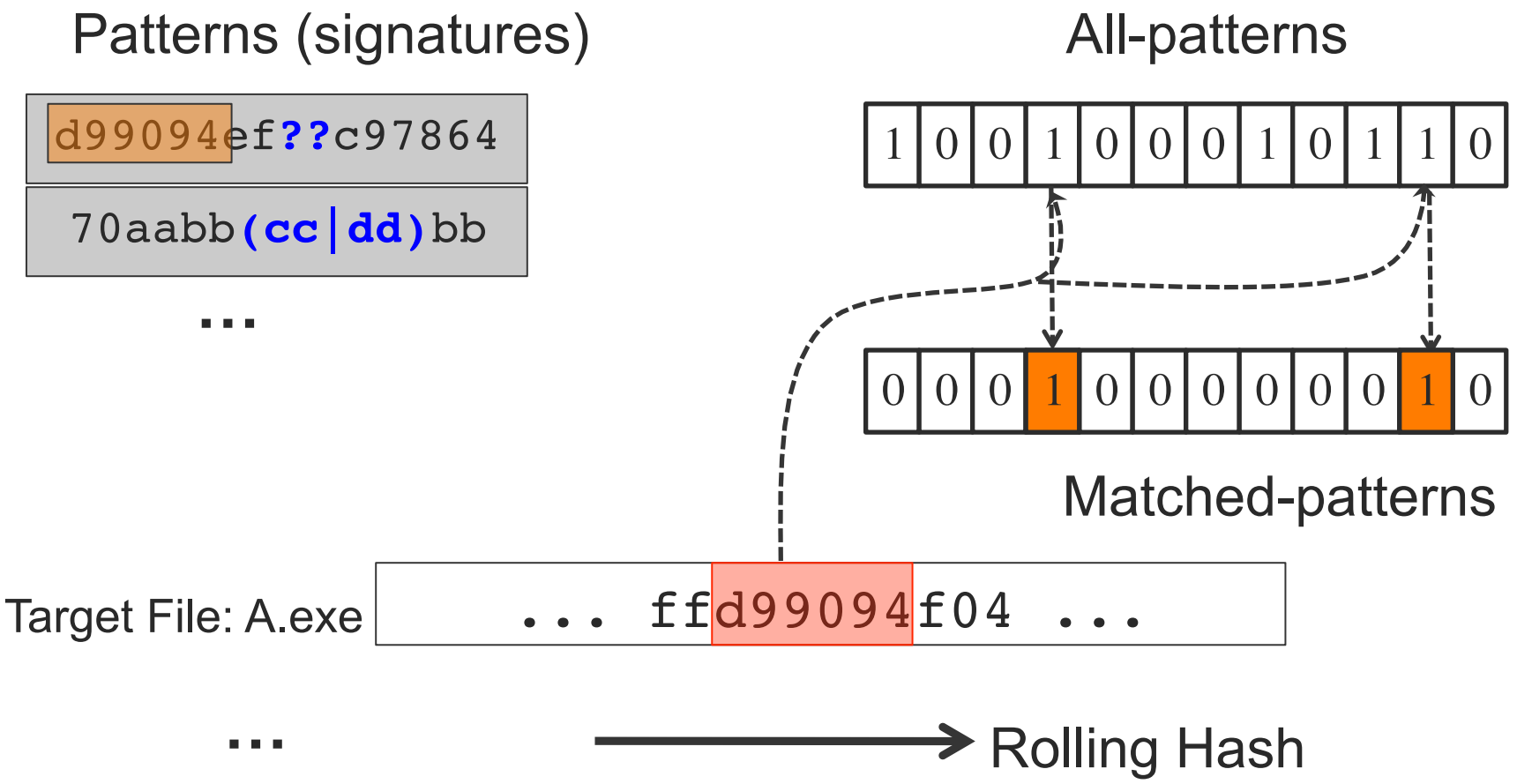
Feed-Forward Bloom Filter (FFBF): Initialization



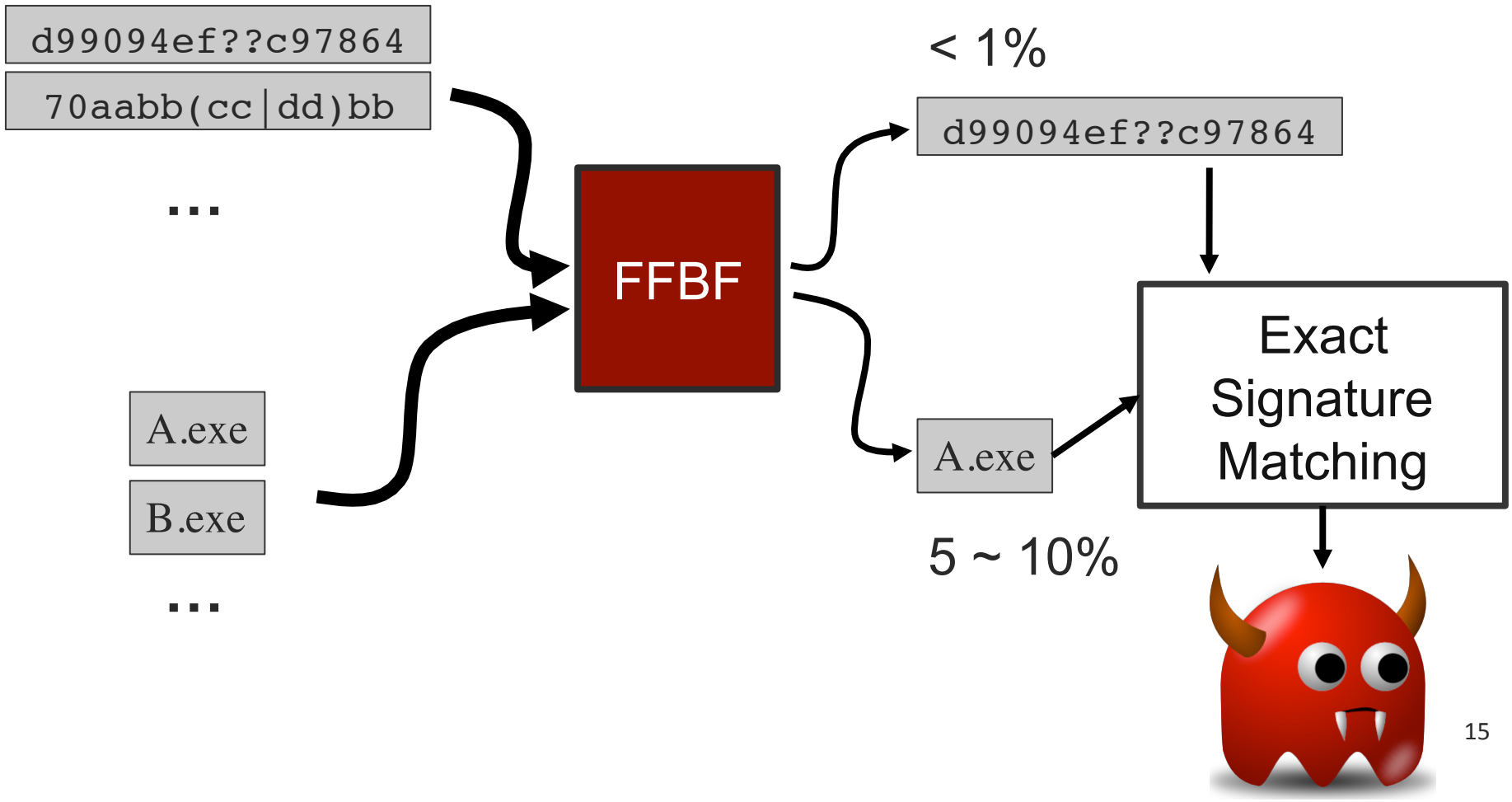
FFBF Scanning: File Filtering



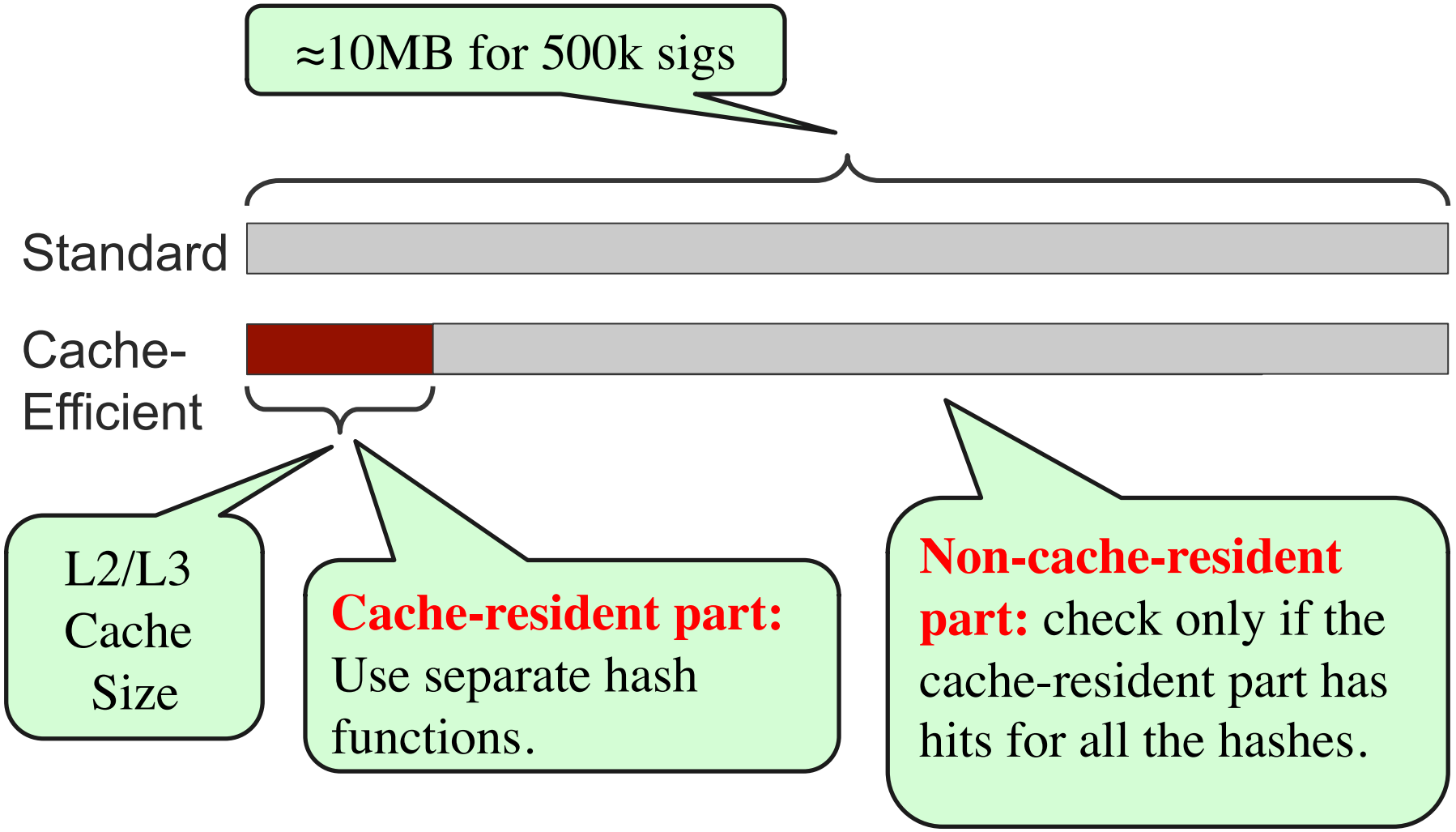
FFBF Scanning: Pattern Filtering



FFBF Scanning Recap



Cache-Efficient Bloom Filter

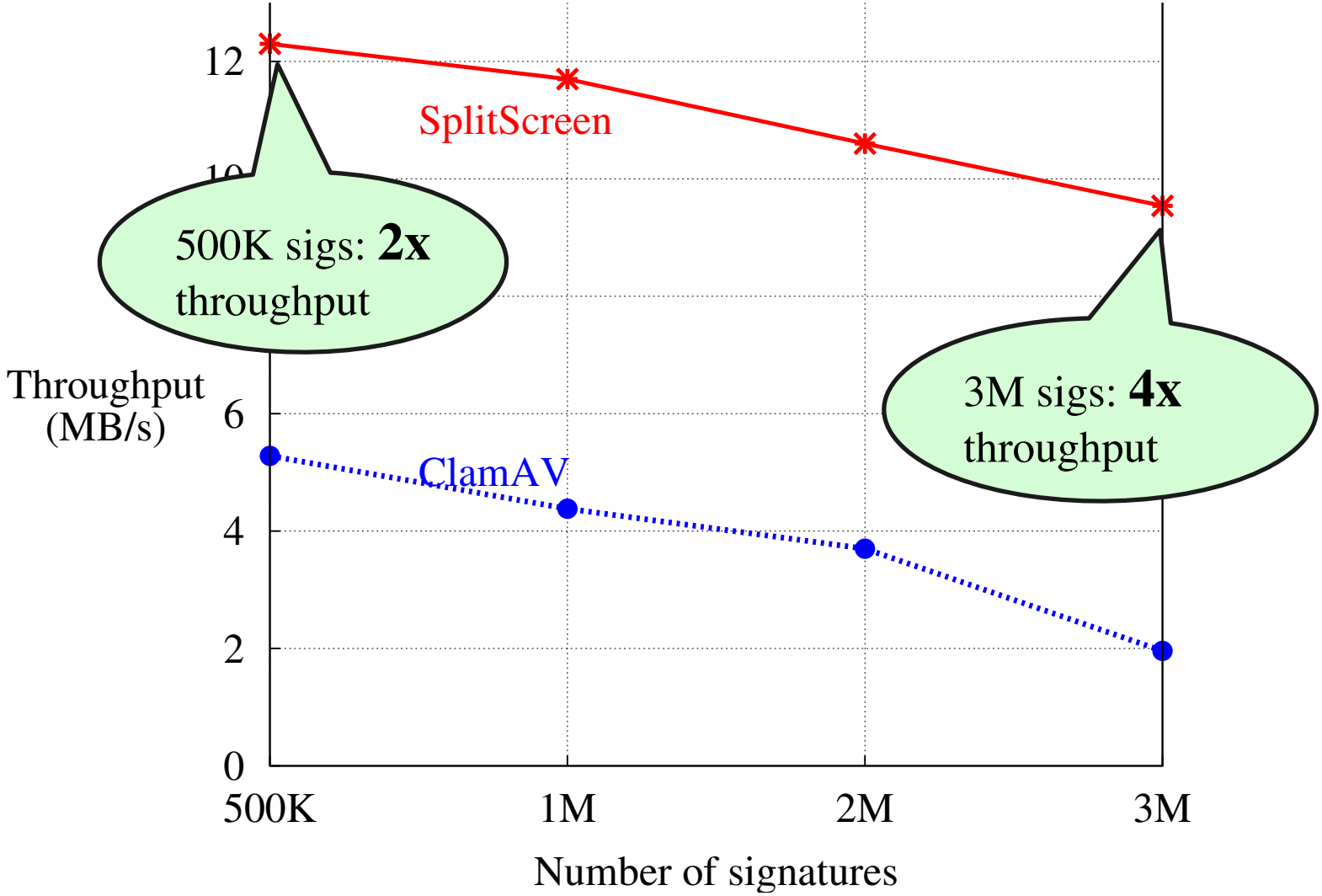


The rest of the talk

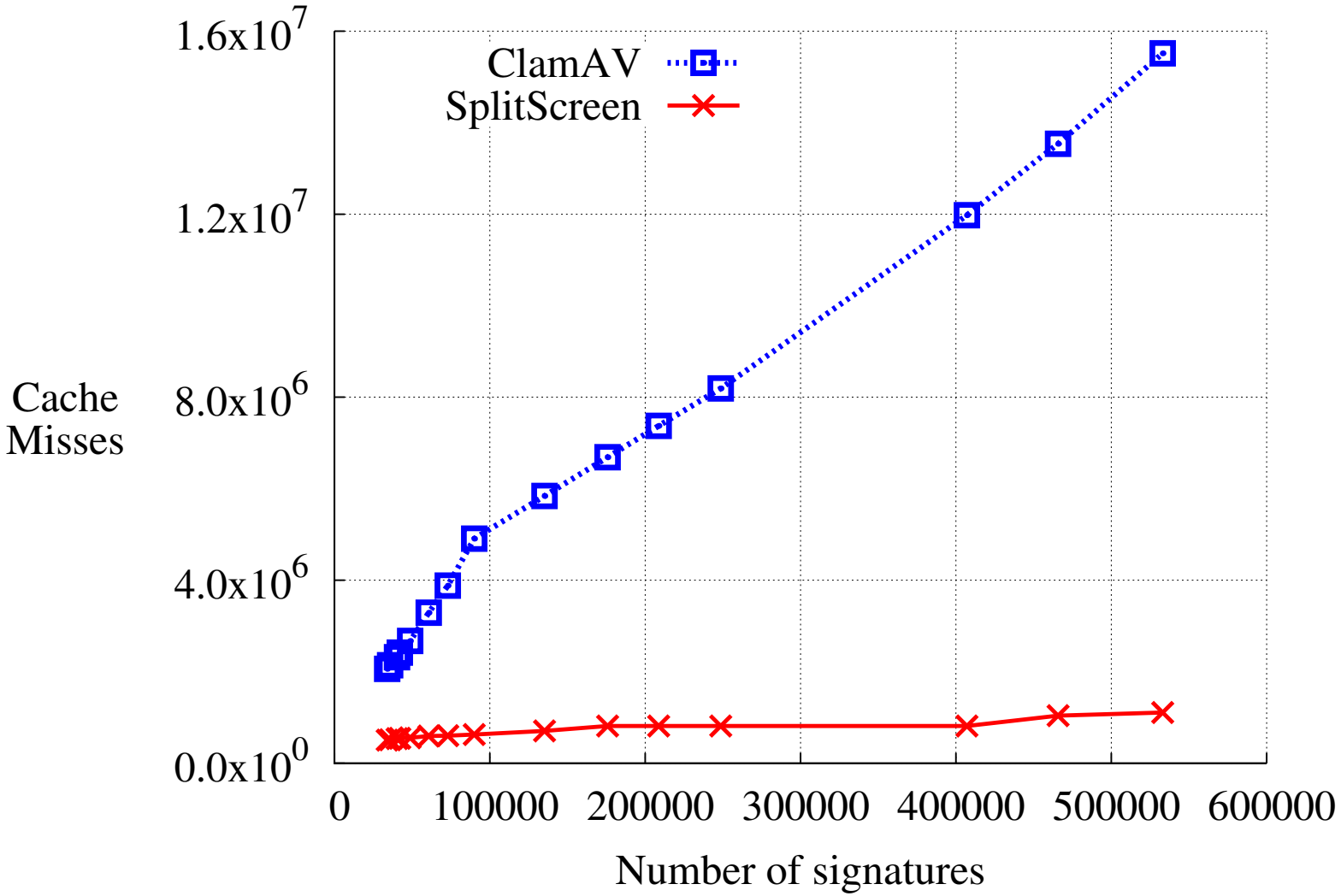
- Evaluation of SplitScreen
  on Intel 2.4 GHz Core 2 Quad with 4 GB of RAM
  - Throughput
  - Cache performance
  - Memory use

- On-demand signature distribution

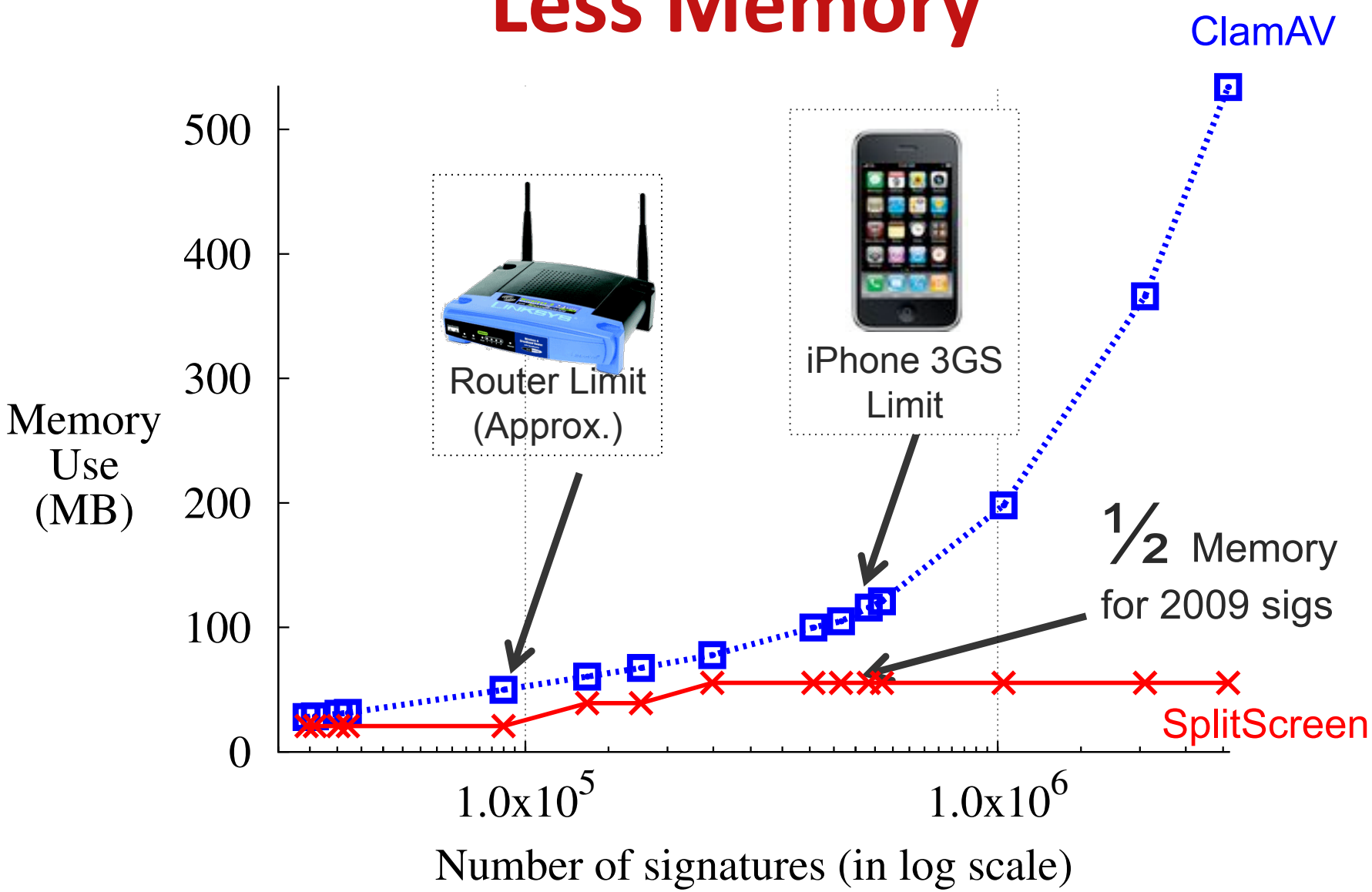
Throughput (1.6 GB Clean Files)



Better Cache Performance



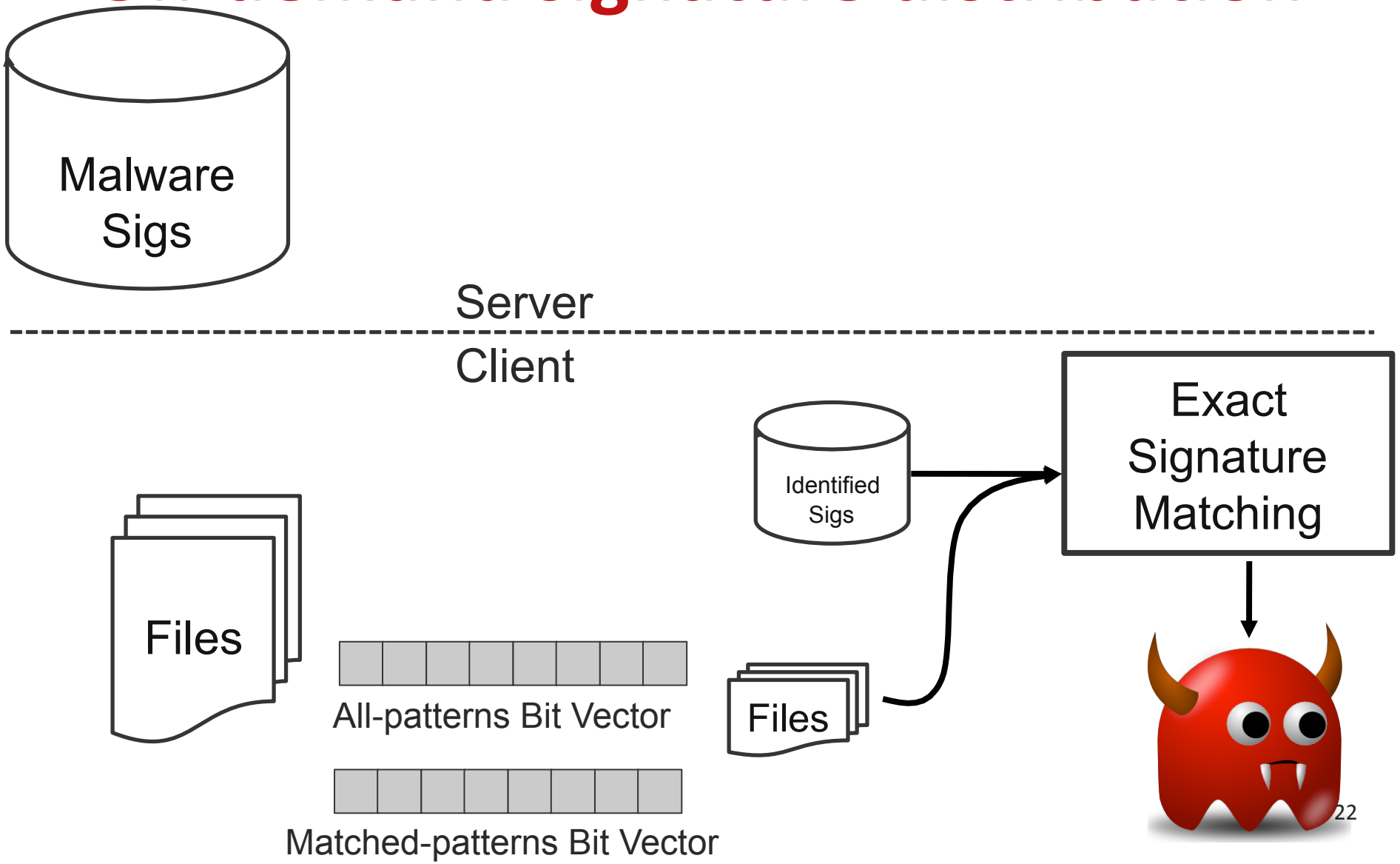
Less Memory


ClamAV
500
400
300
200
100
0
Memory Use (MB)
Router Limit (Approx.)
iPhone 3GS Limit
½ Memory for 2009 sigs
SplitScreen
1.0x10^5
1.0x10^6
Number of signatures (in log scale)

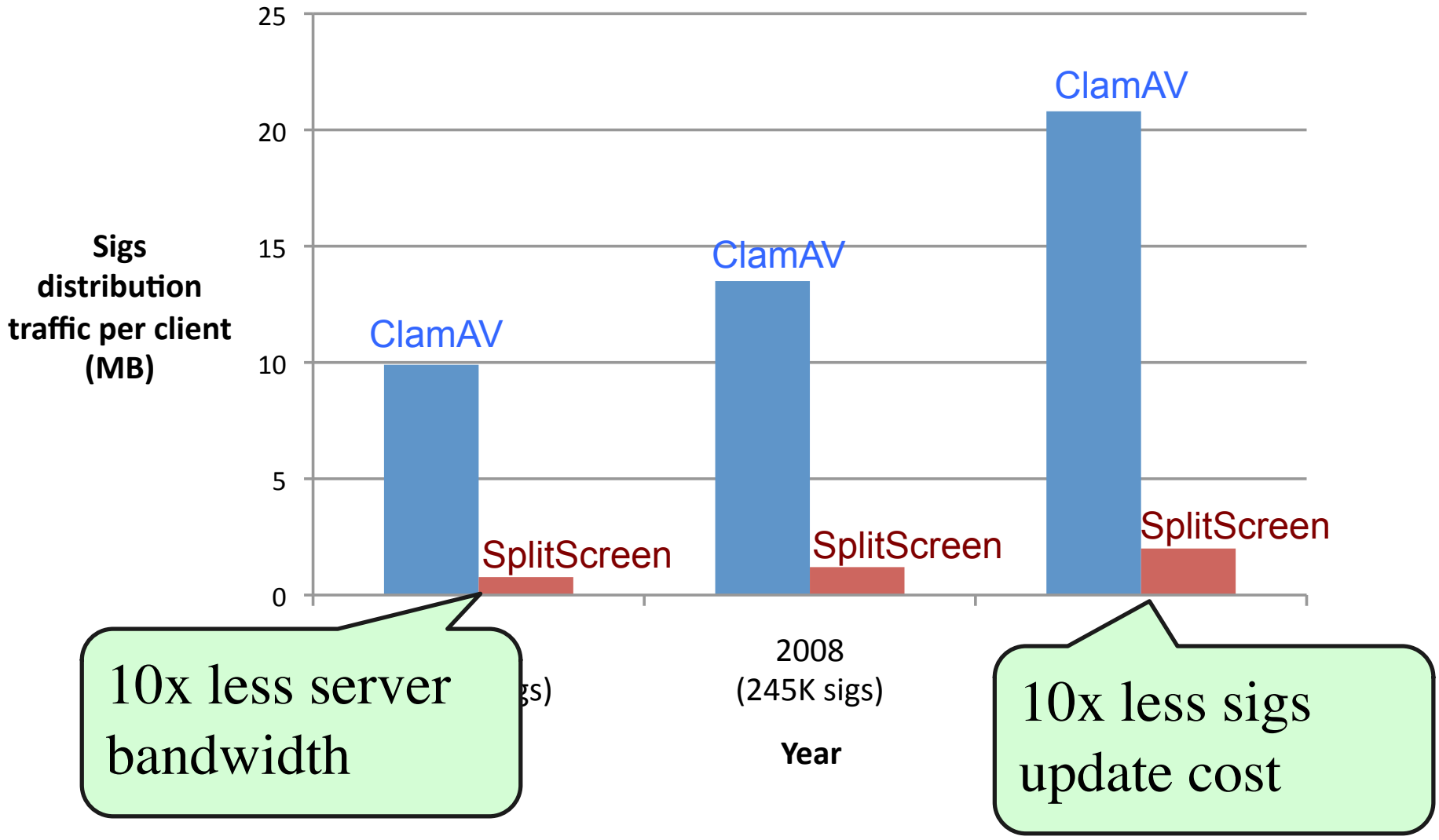
Reducing Signature Distribution Cost

- Option 1: Private data in cloud
  - Cloud-based virus scanning, e.g., CloudAV
  - SplitScreen accelerates scanning
- Option 2: Private data not in cloud
  - On-demand signature distribution

On-demand signature distribution



Lower Initial Signature Distribution Cost



More details in paper

- Handling short signatures
- Choosing which part of signature for FFBF
- FFBF tuning
  - Number of hash functions
  - Size of the sliding window
  - FFBF size

SplitScreen:

- **Malware scanning at 2x - 4x higher throughput, 1/2 - 1/10 memory**
- **Enables malware scanning on weak devices**
  - Embedded systems,…, maybe iPads
- **Enables on-demand signature distribution**
  - Minimize overhead given millions of sigs
- Source code at http://security.ece.cmu.edu/