Using Clustering to Detect and Mitigate Spam Distribution

Andrey Bakhmutov

Kaspersky Lab

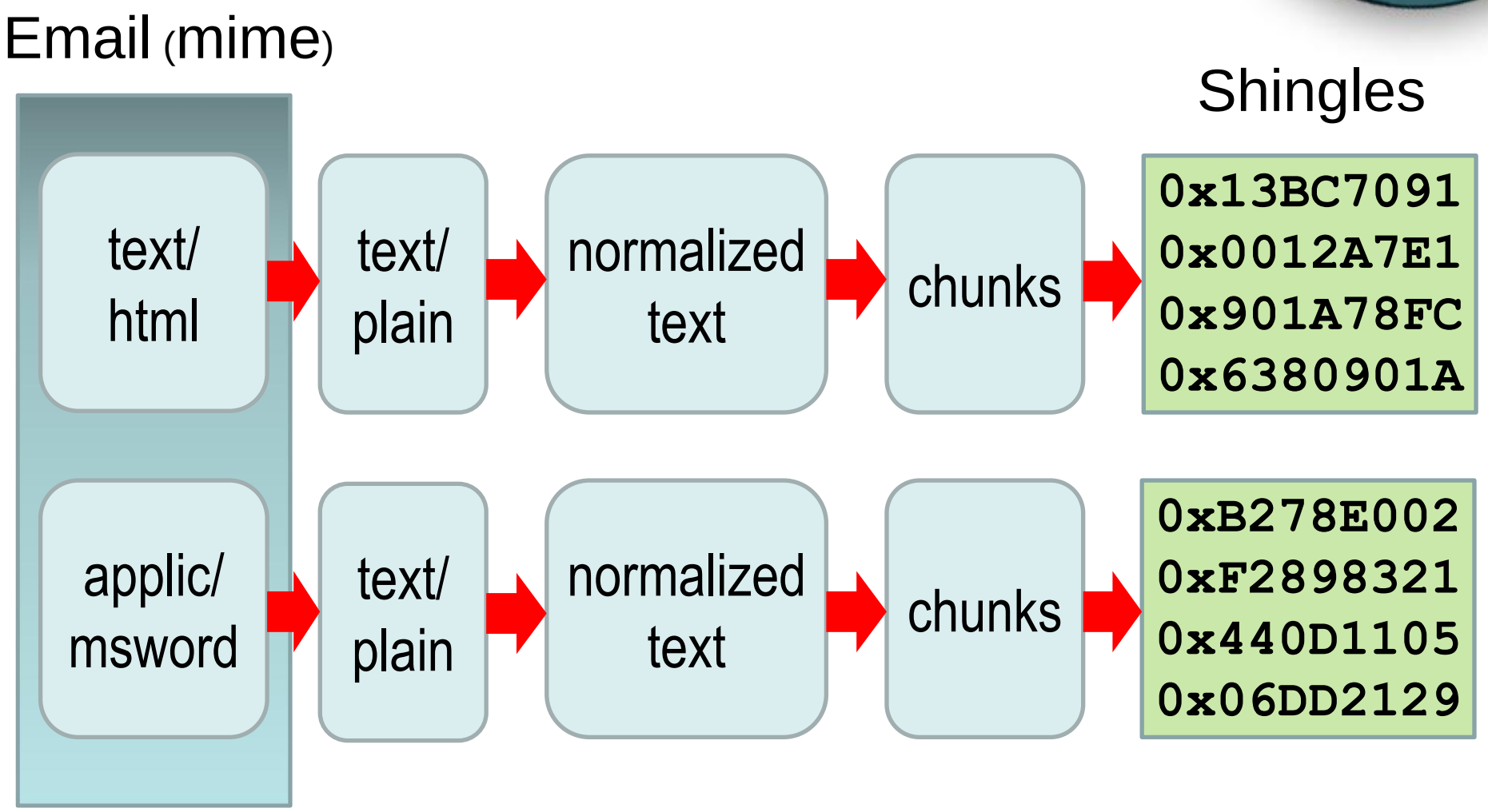
Andrey.Bakhmutov@kaspersky.com

Email clustering system requirements

An anti-spam clustering system should:

- be large-scale and distributed
- use an efficient, secure data model
  - small data size
  - inability to restore original content
- adopt an efficient algorithm for real-time clustering at high speed
- employ extra tools to identify spam clusters

Email processing stages

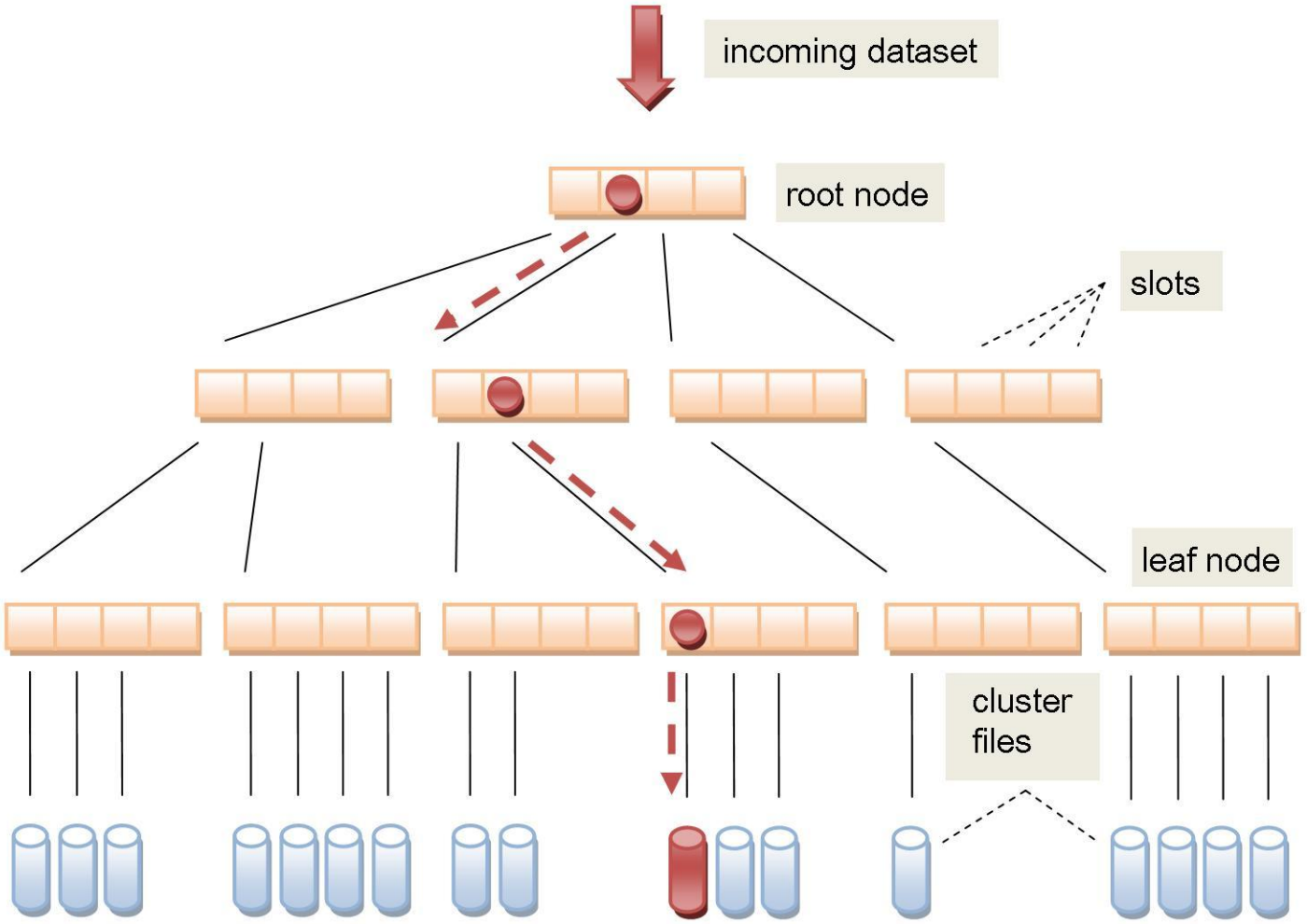


Shingles

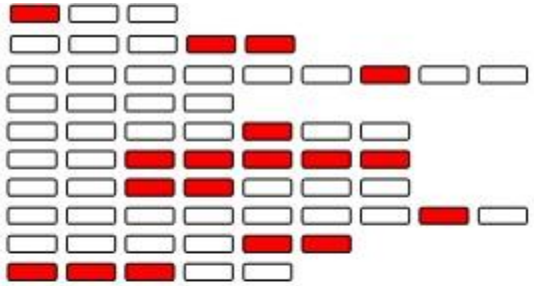
find out why hcg works extremely effective rapid weight loss supplement and where you can order hcg online from verified and trusted online http hcgkiloshot

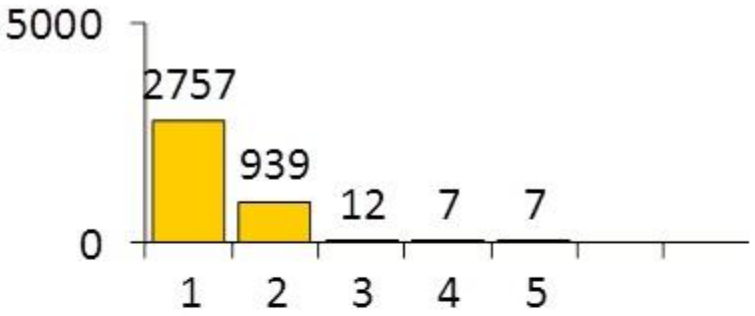
```
0xC05F5857
0x2EA6814D
0xB60257F0
0xA18E4FCD
0x20167ADD
0x31E713B8
```

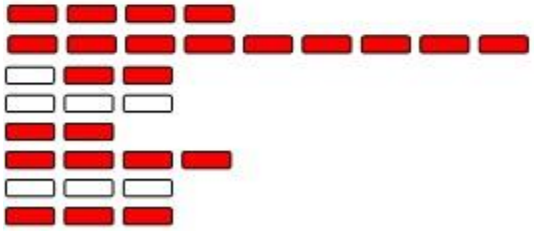
Clustering database

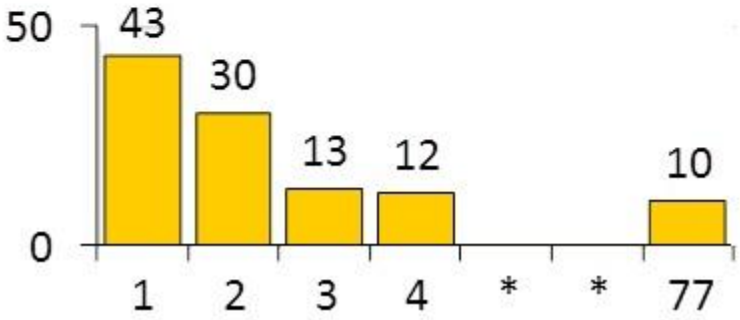


Grouping capabilities

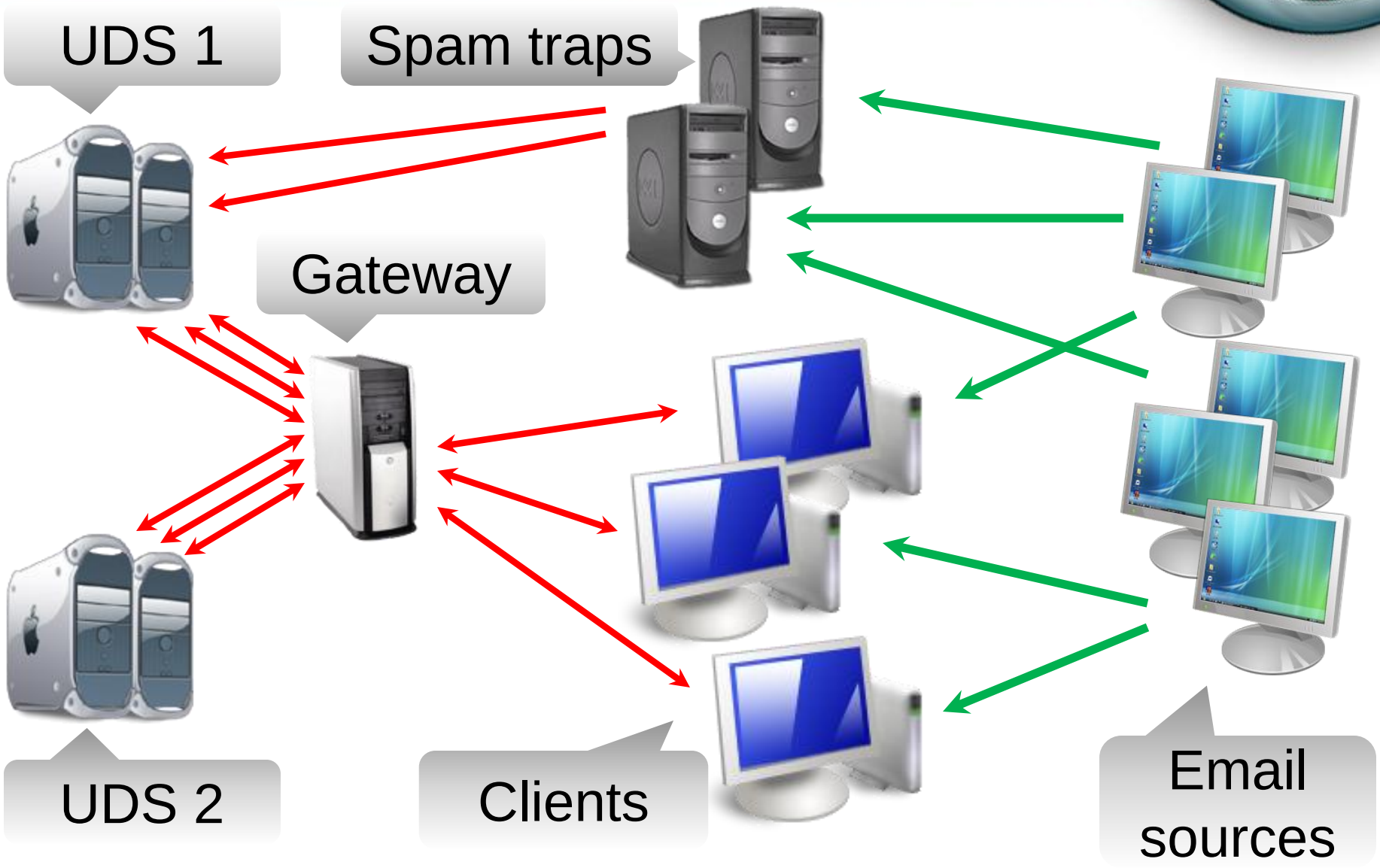








System architecture



Spam vs. mailing list sources

ip-109-38.hyper.net.id

ABTS-mum-dynamic-251.90.170.122.airtelbroadband.in

Static-177.143.195.14.tataidc.co.in

95x153x190x63.kubangsm.ru

112-79-36-62.live.vodafone.in

smtp293.usndr.com

smtp289.usndr.com

smtp131.usndr.com

smtp91.usndr.com

smtp171.usndr.com

Messages with auto-signatures

Here's to an all Manchester Europa League Final!!!!!

Chelsea mailing list

Chelsea@jesternix.net

http://jesternix.net/mailman/listinfo/chelsea_jesternix.net

Message refinement

Here's to an all Manchester Europa League Final!!!!!

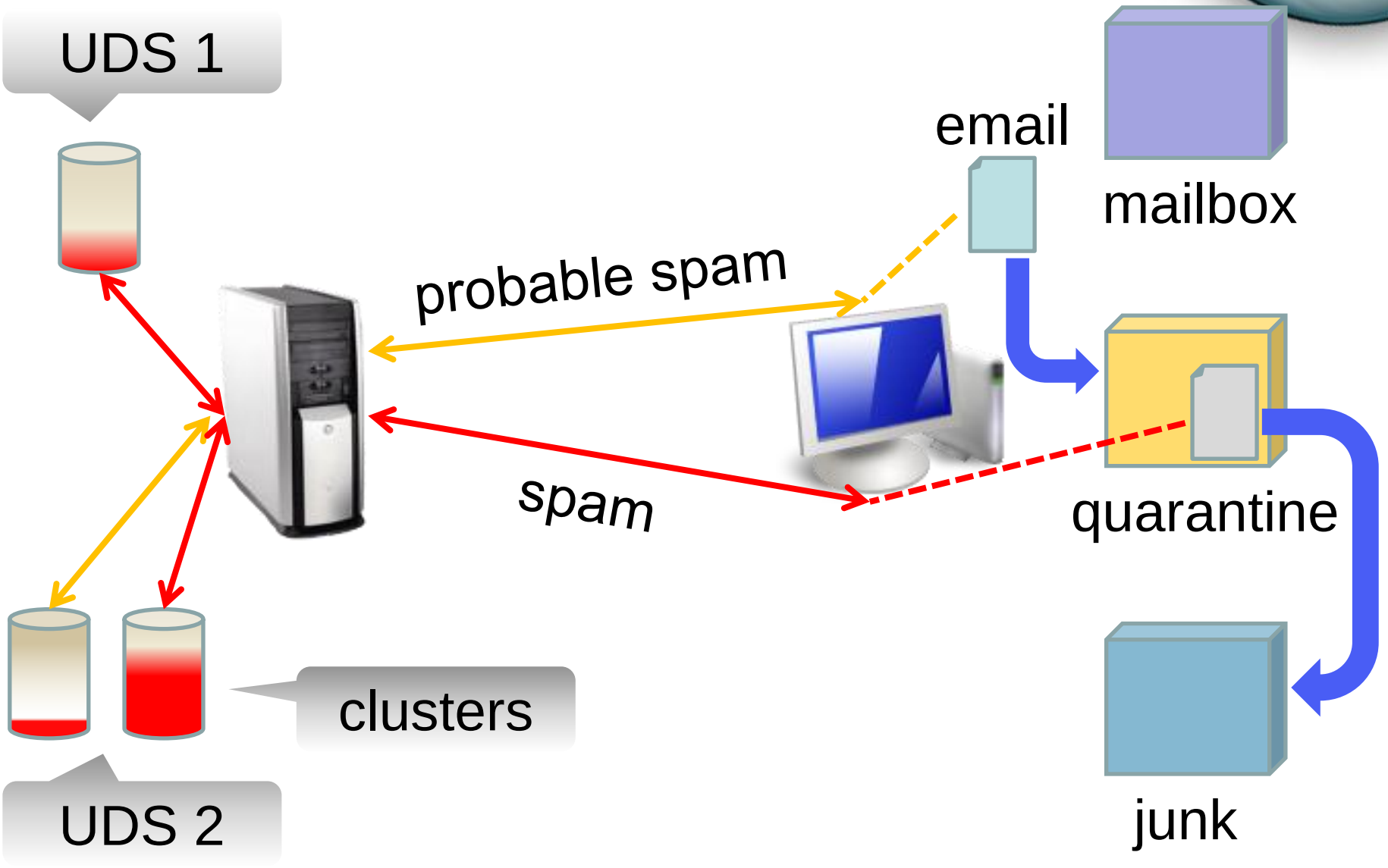
CONFIDENTIALITY NOTICE: This e-mail **message, including any attachments,** is **for the sole use** of **the intended recipient(s), and may contain confidential and privileged information. Any unauthorized review, use, disclosure** or **distribution** is **prohibited.** If **you are not the intended recipient, please contact the sender** by **reply** e-mail **and destroy all copies** of the original message**.**

Chelsea mailing list

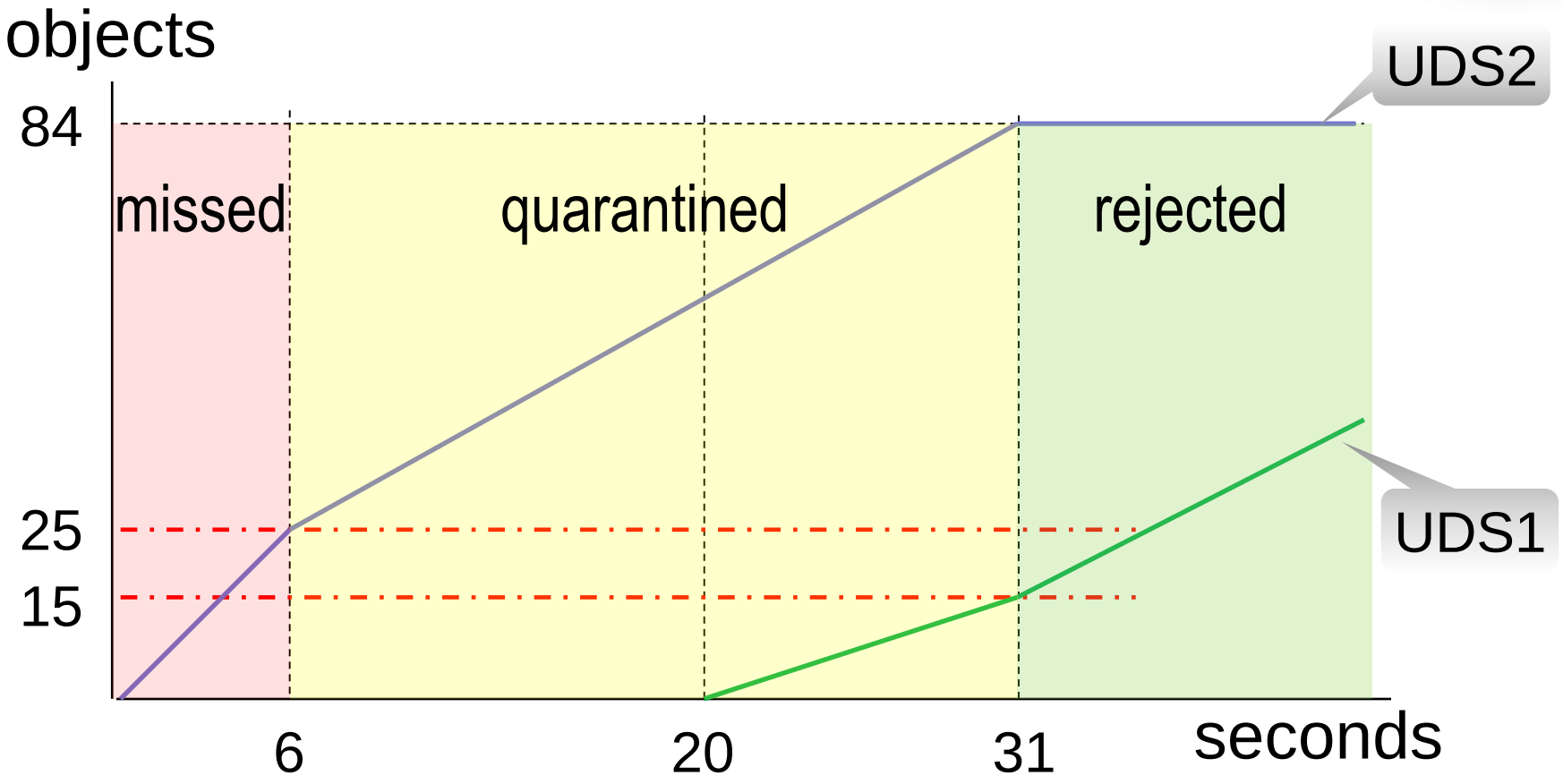
Chelsea@jesternix.net

http://jesternix.net/mailman/listinfo/chelsea_jesternix.net

Quarantine usage



Spam campaign mitigation



THANK YOU

Andrey Bakhmutov

Kaspersky Lab

Andrey.Bakhmutov@kaspersky.com