روش های ذخیره سازی داده های رمز شده غیر تکراری در ابر

هدی جنّتی

پژوهشکده علوم کامپیوتر– پژوهشگاه دانش های بنیادی

hodajannati@ipm.ir

چشم‌انداز

- روش های ذخیره سازی داده های غیر تکراری در ابر
  - حذف داده تکراری مبتنی بر هدف
  - حذف داده تکراری مبتنی بر منبع
  - حذف داده تکراری مبتنی بر کاربر متقابل
- مراحل ذخیره سازی داده های رمز نشده در ابر
- مراحل ذخیره سازی داده های رمز شده در ابر
- نقاط ضعف روش های موجود
- کارهای آتی

روش های ذخیره سازی داده های غیر تکراری در ابر

جهان دیجیتال در سال ۲۰۲۰

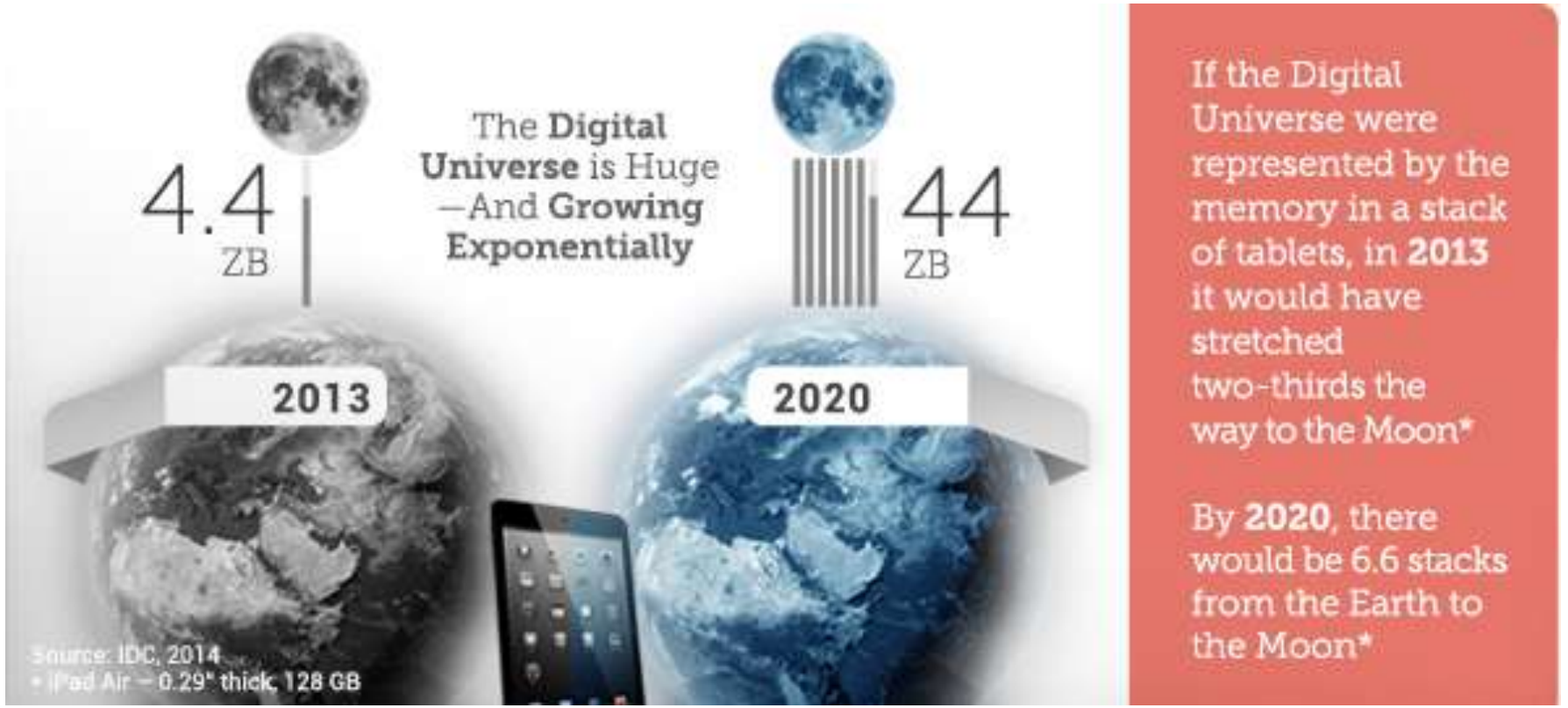


Image source: http://www.emc.com/leadership/digital-universe

ذخیره سازی داده ها در ابر

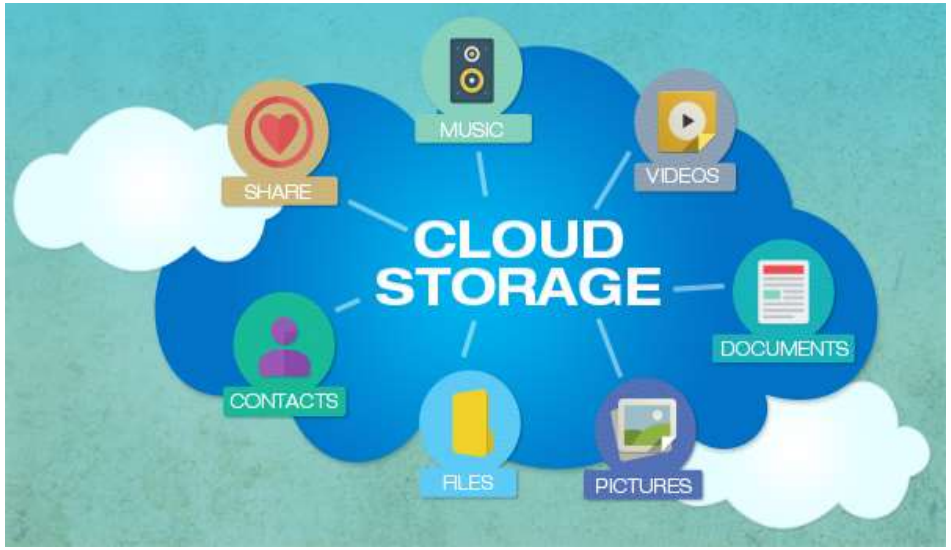


Image source: https://www.emaze.com/@AWRRZQI/CLOUD-COMPUTING-NEW.pptx
http://cloudnewsdaily.com/cloud-storage/
http://www.justcloud.org/2016

ذخیره سازی داده های غیر تکراری

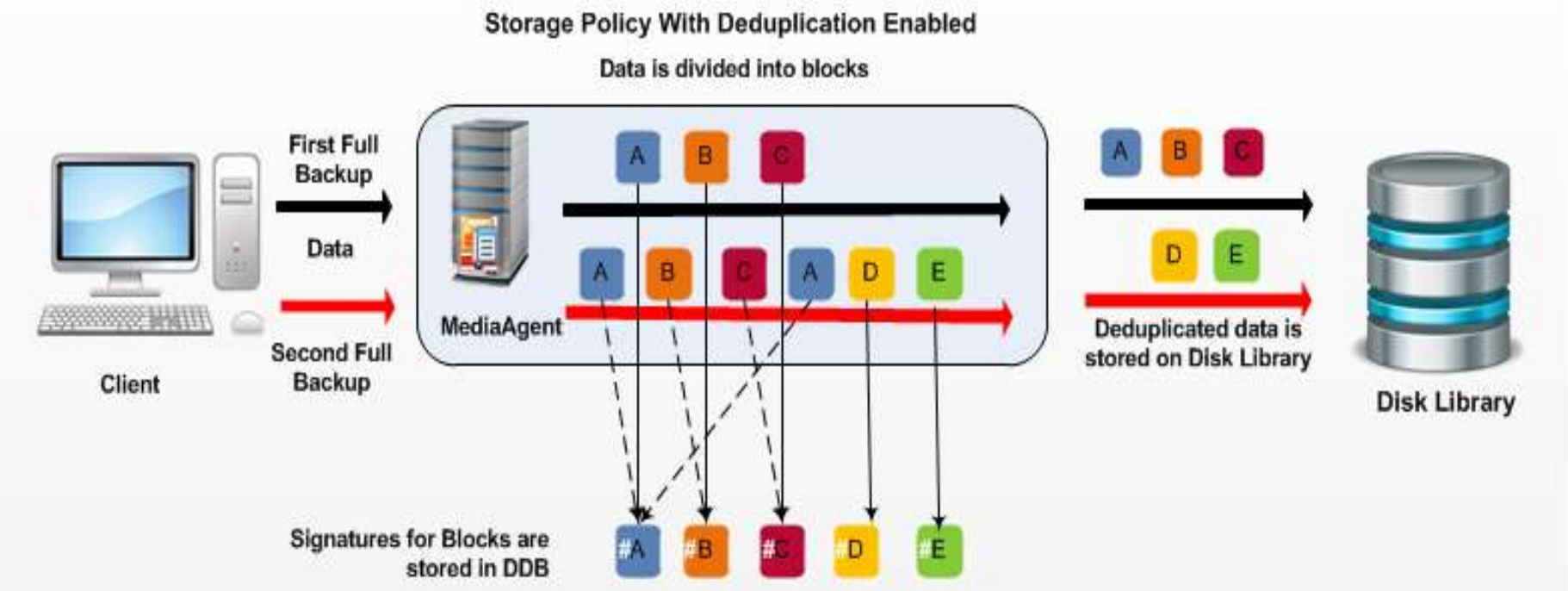


Image source:http://download.netapp.com/knowledge/docs/snapprotect/relsnap_protect100sp2b/215-07763_A0_books_online_100sp2b/books_online_1/english_us/prod_info/dedup_disk.htm

ذخیره سازی داده های غیر تکراری

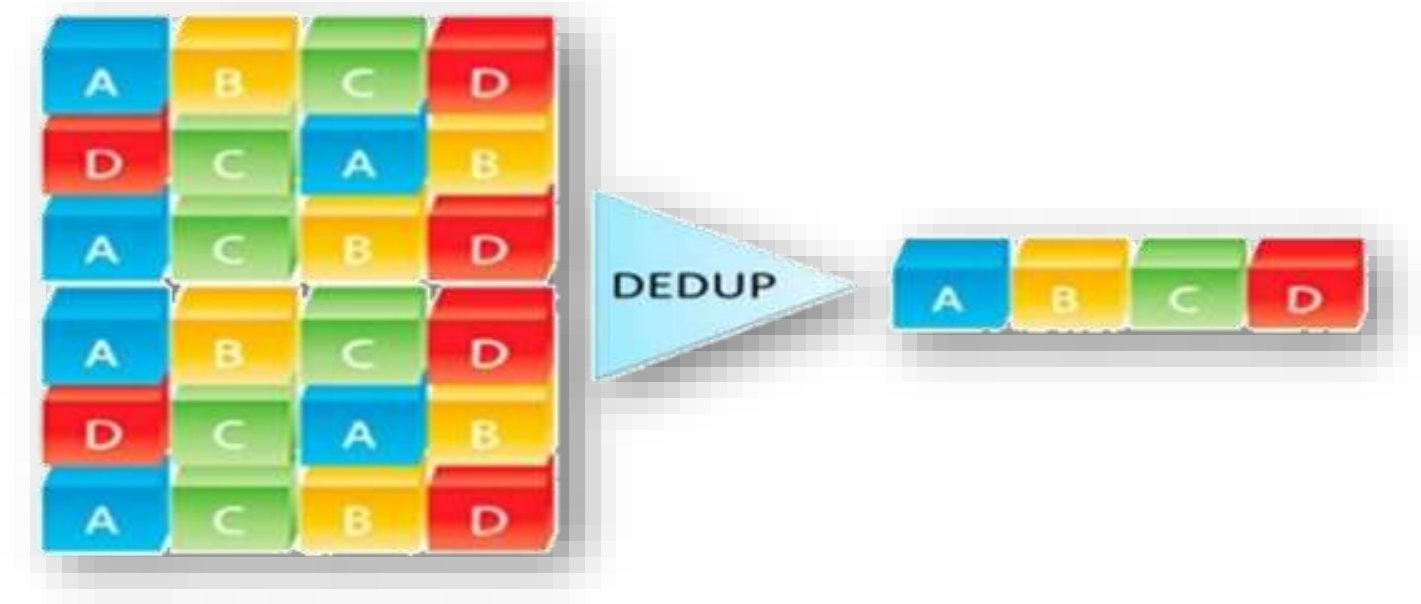


Image source: http://www.slideshare.net/naseem23/liquid-1-38097887

ذخیره سازی داده های غیر تکراری

- دسته بندی روش های ذخیره سازی داده های غیر تکراری در ابر
  - حذف داده تکراری مبتنی بر هدف (ابر)
  - حذف داده تکراری مبتنی بر منبع (کاربر)
  - حذف داده تکراری مبتنی بر کاربر متقابل

ذخیره سازی داده های غیر تکراری

- حذف داده تکراری مبتنی بر هدف (ابر)

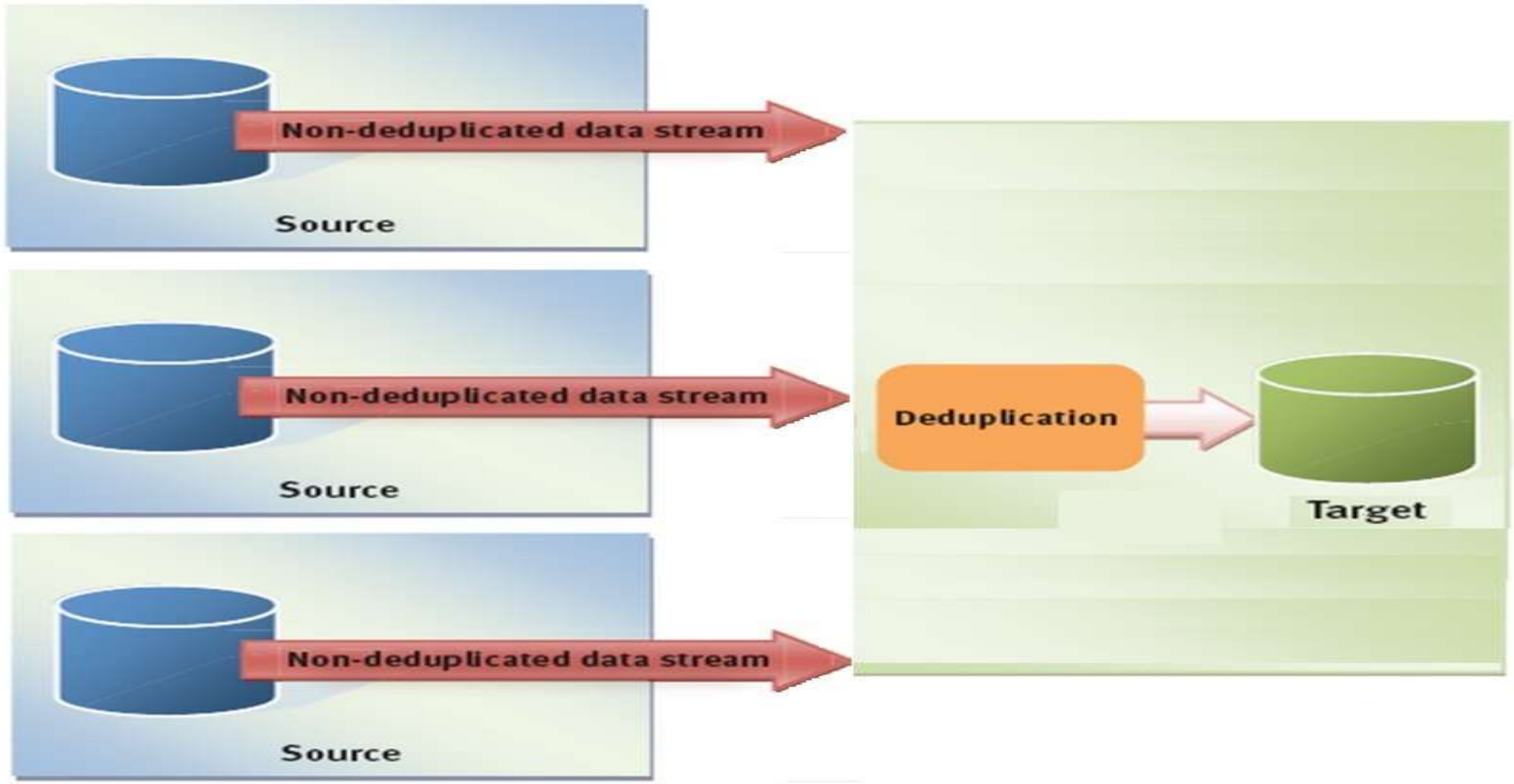


Image source: https://www.uk.insight.com/en-gb/shop/acronis/deduplication

ذخیره سازی داده های غیر تکراری

- حذف داده تکراری مبتنی بر منبع (هر کاربر)

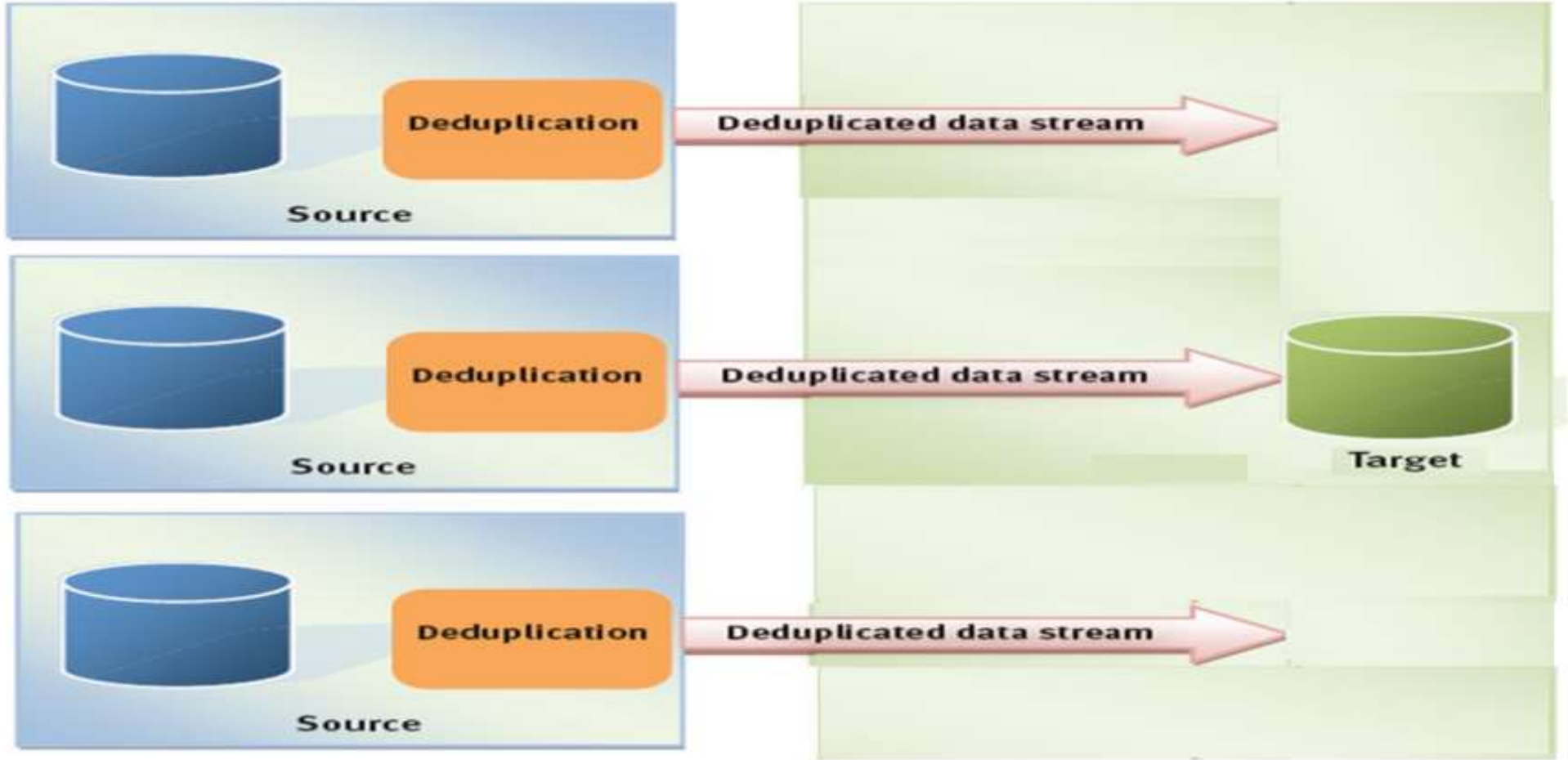


Image source: https://www.uk.insight.com/en-gb/shop/acronis/deduplication

ذخیره سازی داده های غیر تکراری

حذف داده تکراری مبتنی بر کاربر متقابل

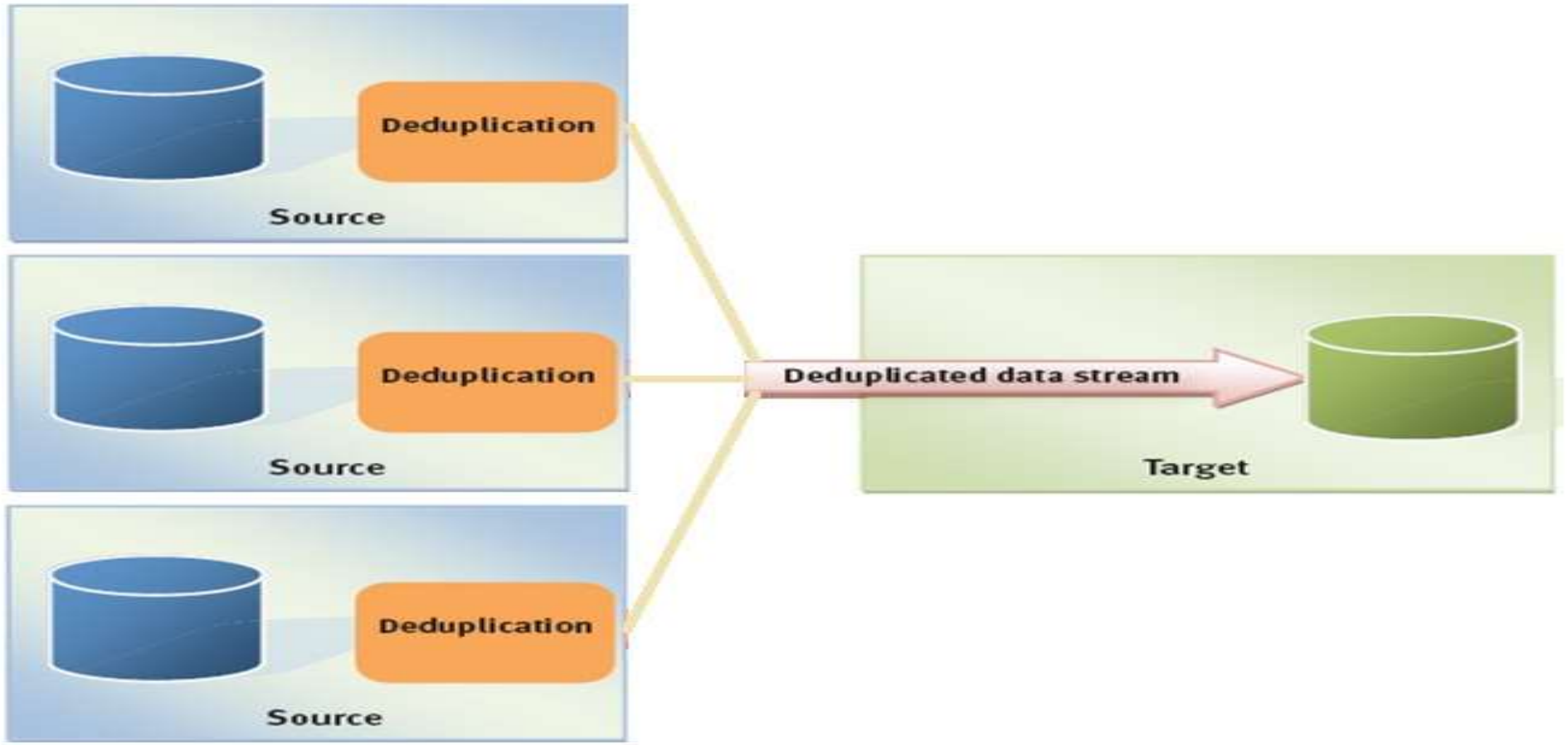


Image source: https://www.uk.insight.com/en-gb/shop/acronis/deduplication

مراحل ذخیره سازی داده های رمز نشده در ابر

حذف داده تکراری مبتنی بر کاربر متقابل (داده رمز نشده)

D1 = Hello → T1 = Hash(D1)

T1 →

← No duplication

D1 →

← Pointer1

D2 = Bye → T2 = Hash(D2)

T2 →

← No duplication

D2 →

← Pointer2

D1 = Hello → T1 = Hash(D1)

T1 →

← Duplication

← Pointer1

Tag	Cipher	
T1	D1	Pointer1
T2	D2	Pointer2
		Pointer3

حذف داده تکراری مبتنی بر کاربر متقابل (داده رمز نشده)

مراحل ذخیره سازی داده های رمز شده در ابر

حذف داده تکراری مبتنی بر کاربر متقابل (داده رمز شده)

حذف داده تکراری مبتنی بر کاربر متقابل (داده رمز شده)

- **مسأله:** حذف داده تکراری مبتنی بر کاربر متقابل، روی داده های رمز شده قابل اجرا نمی باشد.

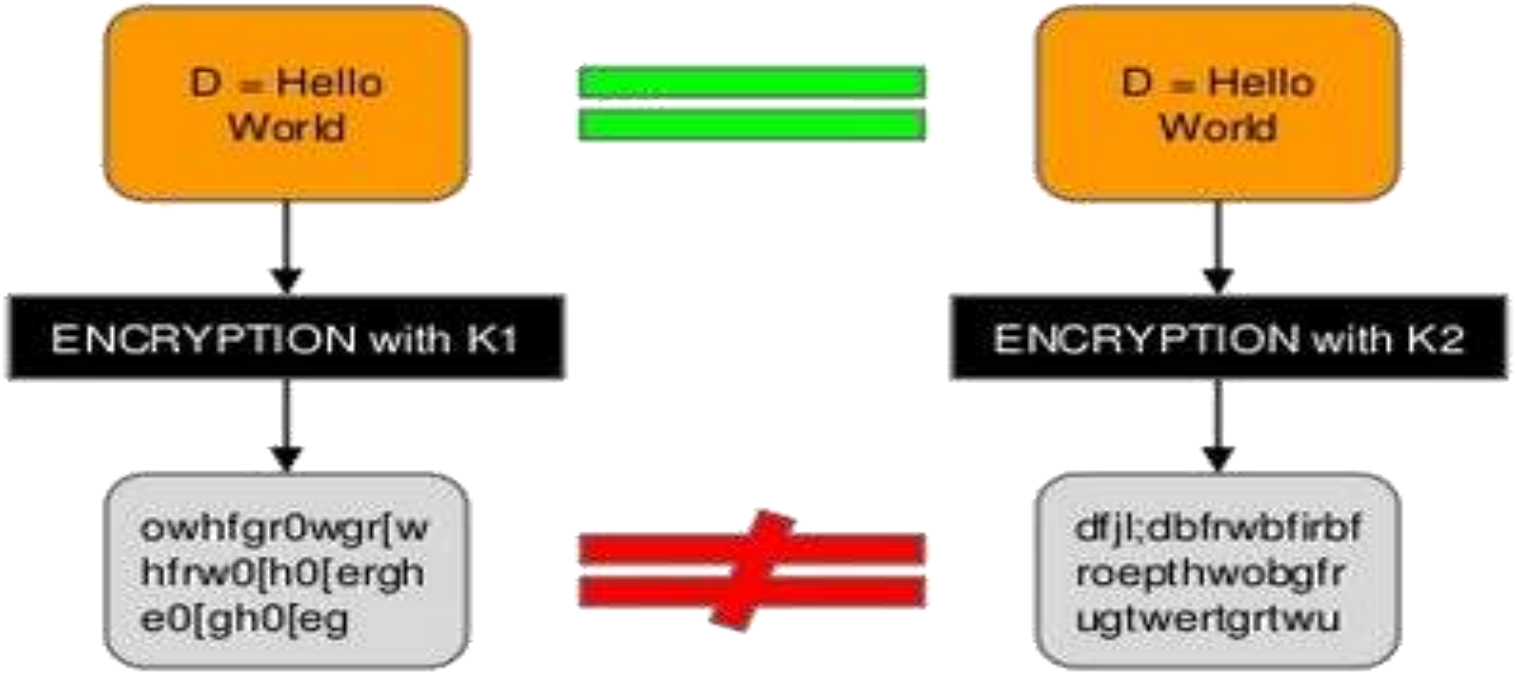


Image source: http://www.emc.com/leadership/digital-universe/2014iview/executive-summary.htm

حذف داده تکراری مبتنی بر کاربر متقابل (داده رمز شده)

D1 = Hello

T1 = Hash(D1)

C1 = Enc(D1,K1)

Tag	Cipher	
T1	C1	Pointer1
		Pointer2
		Pointer3

T1 →

← No duplication

C1 →

← Pointer1

D۱ = Hello

T1 = Hash(D1)

C2 = Enc(D1,K2)

T1 →

← Deduplication

← Pointer1

حذف داده تکراری مبتنی بر کاربر متقابل (داده رمز شده)

- **راه حل:** استفاده از کلید همگرا (رمز گذاری همگرا)

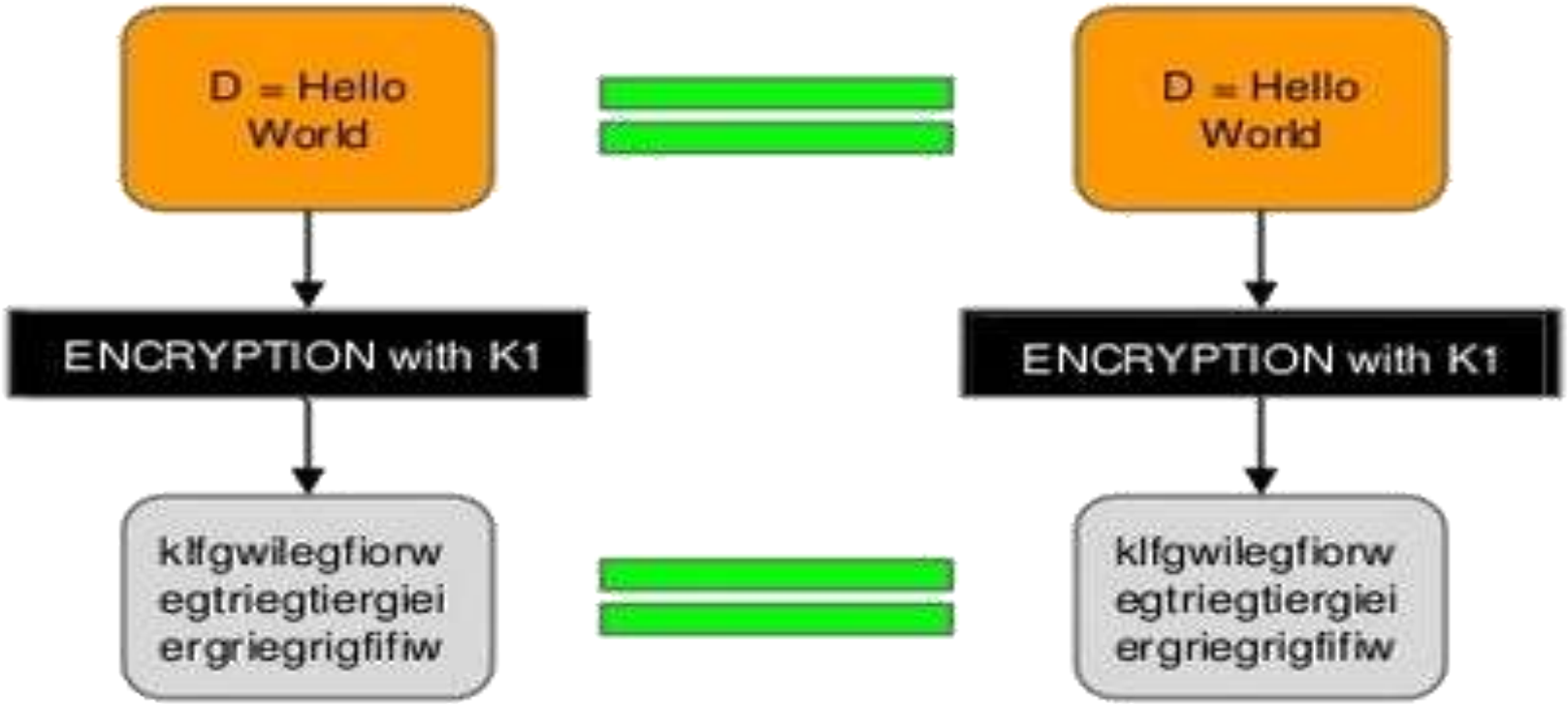


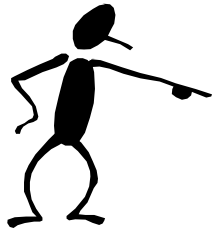
Image source: http://www.emc.com/leadership/digital-universe/2014iview/executive-summary.htm

حذف داده تکراری مبتنی بر کاربر متقابل (تولید کلید)

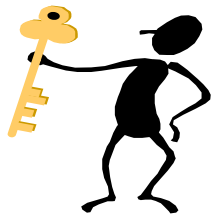
- تولید کلید همگرا برای داده های یکسان
  - ***Douceur et al.*** [1]: تولید کلید همگرا با بکارگیری داده (K=Hash(data))
    - آسیب پذیر به حمله واژه نامه ای توسط کاربران (***Bellare et al*** [2])
  - ***Bellare et al.*** [2]: تولید کلید همگرا توسط ابر
  - ***Bellare et al.*** [2]: تولید کلید همگرا هم با بکارگیری داده و هم با بکارگیری کلید خصوصی ابر
    - آسیب پذیر به حمله واژه نامه ای توسط ابر

حذف داده تکراری مبتنی بر کاربر متقابل (تولید کلید)

- ***Bellare et al.*** [2]: تولید کلید همگرا هم با بکارگیری داده و هم با بکارگیری کلید خصوصی ابر



User



Key Server

$K = Hash(D)r^e$

K →

$K' = K^d = Hash(D)^d r$

K' ←

$K_D = r^{-1}(K') = Hash(D)^d$

حذف داده تکراری مبتنی بر کاربر متقابل (تولید کلید)

- ذخیره سازی کلید همگرا (مدیریت کلید)
  - ***Bellare et al.*** [2]: هر کاربر، کلید همگرای رمز شده با کلید خصوصی خودش را برای ابر ارسال کند و ابر کلید همگرای رمز شده برای هر کاربر را به صورت جداگانه ذخیره کند.
  - ***Li et al.*** [3]: هر کاربر، کلید همگرا را با به کارگیری طرح تسهیم راز به *l* سهم تقسیم کند. سپس هر سهم را در اختیار یکی از سرورهای ابر قرار دهد.
  - ***Chen et al.*** [4]: هر کلید همگرایی توسط هر کاربر به صورت جداگانه ذخیره شود.

حذف داده تکراری مبتنی بر کاربر متقابل (داده رمز شده)

حذف داده تکراری مبتنی بر کاربر متقابل (تولید برچسب)

- **مسأله:** سازگاری برچسب با داده رمز شده
- تولید برچسب با استفاده از داده رمز نشده
- تولید برچسب با استفاده از داده رمز شده (Bellare et al [5])

حذف داده تکراری مبتنی بر کاربر متقابل (تولید برچسب)

تولید برچسب با داده رمز نشده

D1 = Hello

Tag1 = Hash(D1)

C1 = Enc(D1,K1)

Tag1

No duplication

C1

Pointer1

Tag	Cipher
Tag1	C1

Pointer1

Pointer2

Pointer3

D1 = Hello

Tag1 = Hash(D1)

Tag1

Duplication

Pointer1

حذف داده تکراری مبتنی بر کاربر متقابل (تولید برچسب)

- نقطه ضعف تولید برچسب با داده رمز نشده(Bellare et al [5])

Tag	Cipher	
Tag1	C1	Pointer1
Tag2	C2	Pointer2
		Pointer3

D2 = Bye

Tag2 = Hash(D2)

C2 = Enc(D4,K2)

Tag2

No duplication

C2

Pointer2

D2 = Bye

Tag2 = Hash(D2)

Tag2

Duplication

Pointer2

حذف داده تکراری مبتنی بر کاربر متقابل (تولید برچسب)

- تولید برچسب با داده رمز شده (Bellare et al [5])

D1 = Hello → C1 = Enc(D1,K1) → Tag1 = Hash(C1)

Tag1 →

← No duplication

C1 →

← Pointer1

Tag	Cipher	
Tag1	C1	Pointer1
		Pointer2
		Pointer3

حذف داده تکراری مبتنی بر کاربر متقابل (تولید برچسب)

- تولید برچسب با داده رمز شده (Bellare et al [5])

D1 = Hello → C1 = Enc(D4,K1) → Tag1 = Hash(C1)

Tag1 →

← No duplication

C1 →

← Pointer1

Tag	Cipher	
Tag1	C1	Pointer1
Tag'1	C'1	Pointer2
		Pointer3

D1 = Hello → C'1 = Enc(D1,K1) → Tag'1 = Hash(C'1)

Tag'1 →

← No duplication

C'1 →

← Pointer2

حذف داده تکراری مبتنی بر کاربر متقابل (داده رمز شده)

نقطه ضعف در حذف داده تکراری مبتنی بر کاربر متقابل (داده رمز شده)

- **مسأله:** بازیابی داده توسط کاربر نامعتبر
  - کاربر نامعتبر: کاربری که به کل داده اولیه دسترسی نداشته است.

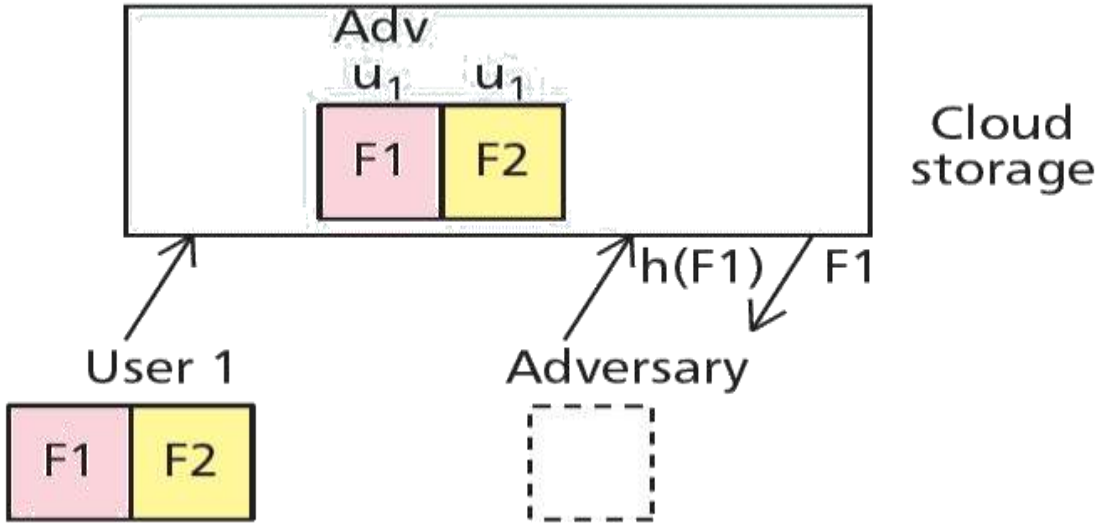


Image source: Ref [6]

حذف داده تکراری مبتنی بر کاربر متقابل (داده رمز شده)

- **مسأله:** بازیابی داده توسط کاربر نامعتبر
  - کاربر نامعتبر: کاربری که به کل داده اولیه دسترسی نداشته است.
- **راه حل:** برای دسترسی به محل ذخیره سازی هر داده در ابر، کاربران باید به ابر اثبات کنند که به کل داده دسترسی دارند نه اینکه فقط به برچسب آن دسترسی داشته باشند (Halevi et al. [6]). بدین منظور از پروتکل های اثبات مالکیت Proof of Ownership (POW) باید استفاده شود.

حذف داده تکراری مبتنی بر کاربر متقابل (داده رمز شده)

حذف داده تکراری مبتنی بر کاربر متقابل (داده رمز شده)

- راه حل: پروتکل اثبات مالکیت (Halevi et al. [6])

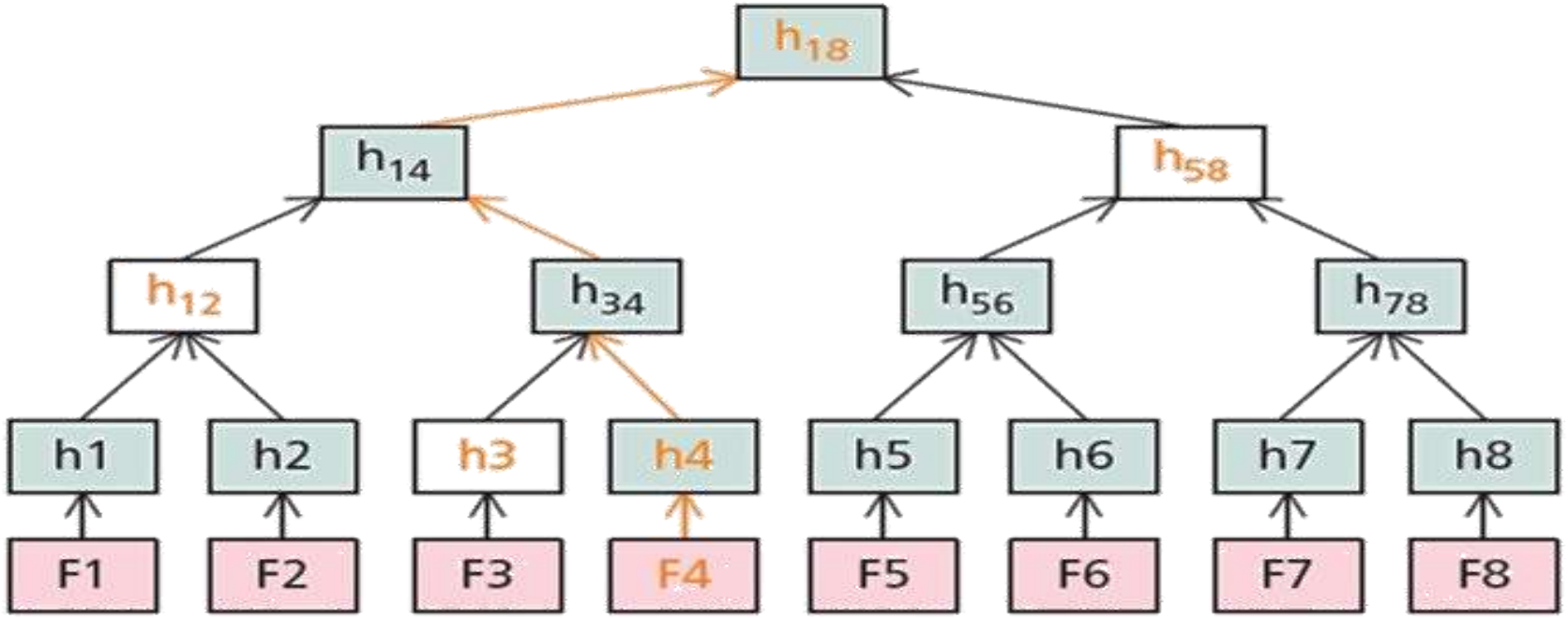


Image source: Ref [6]

نقاط ضعف روش های موجود

نقطه ضعف ها در حذف داده تکراری مبتنی بر کاربر متقابل (داده رمز شده)

- **مسأله:** تجسس وجود داده خاص در ابر (Yu et al. [7])

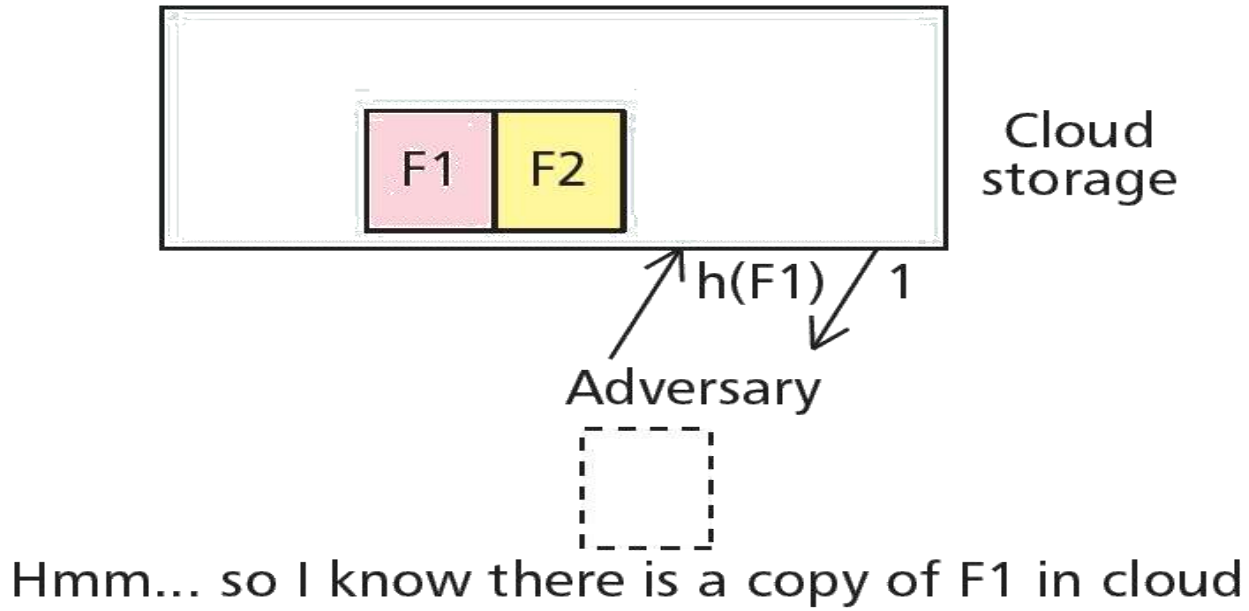


Image source: Ref [7]

نقطه ضعف ها در حذف داده تکراری مبتنی بر کاربر متقابل (داده رمز شده)

- **مسأله:** استفاده از کانال به عنوان کانال پنهان (Yu et al. [7])

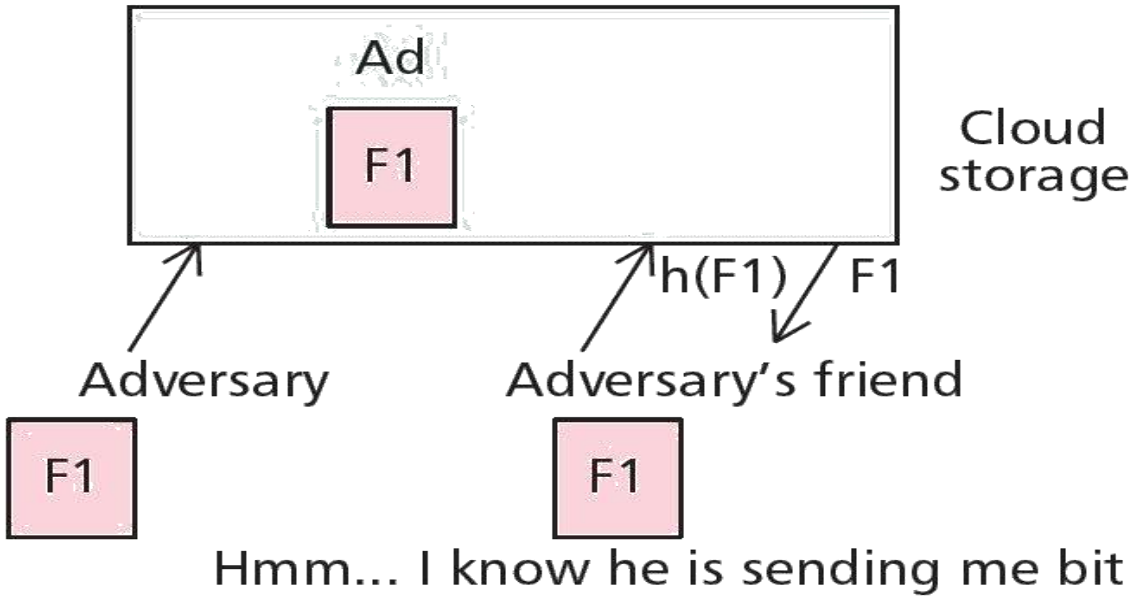


Image source: Ref [7]

نقطه ضعف ها در حذف داده تکراری مبتنی بر کاربر متقابل (داده رمز شده)

- **مسأله:** استفاده از شبکه به عنوان شبکه انتقال داده ([7] .Yu et al)

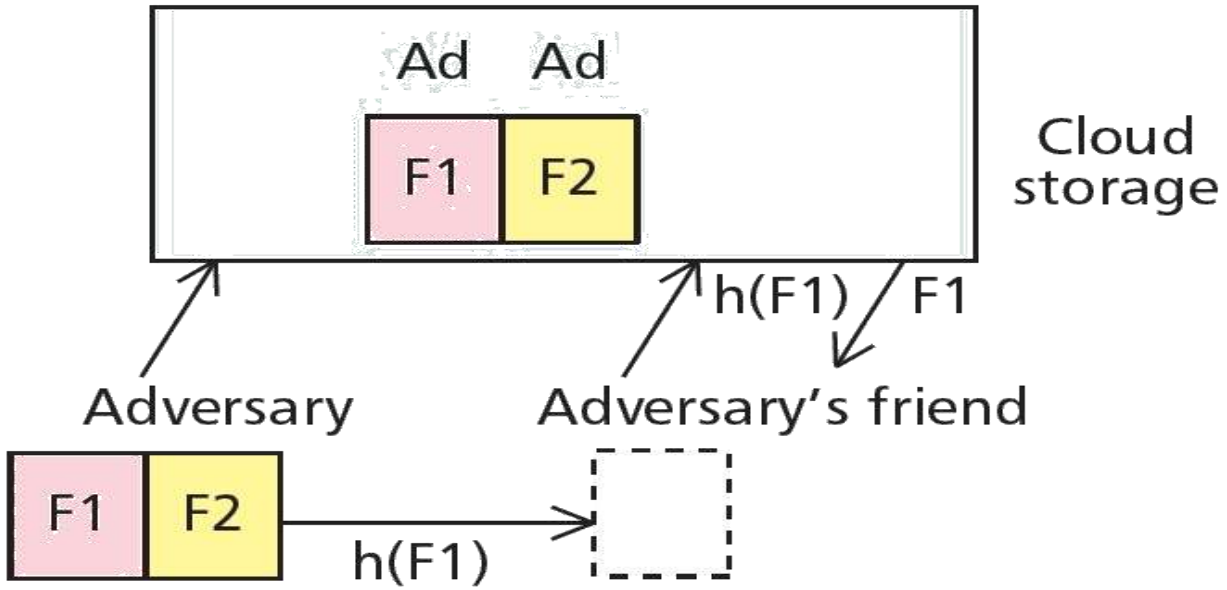


Image source: Ref [7]

کارهای آتی

- تولید کلید همگرایی
- نحوه ذخیره سازی کلید همگرایی
- جلوگیری از حملات موجود
  - تجسس وجود داده خاص در ابر
  - استفاده از کانال به عنوان کانال پنهان
  - استفاده از شبکه به عنوان شبکه انتقال داده
- تطابق با ویژگی های دیگر در ابر

مراجع

1. J. R. Douceur, A. Adya, W. J. Bolosky, D. Simon, and M. Theimer, “Reclaiming space from duplicate files in a serverless distributed file system,” in *Proc. of the 22nd International Conference on Distributed Computing Systems (ICDCS’02)*, pp. 617-624, 2002.
2. M. Bellare, S. Keelveedhi, and T. Ristenpart, “Dupless: server aided encryption for deduplicated storage,” in *Proc. of the 22nd USENIX Conference on Security (SEC’13)*, pp. 179–194, 2013.
3. J. Li, X. Chen, M. Li, J. Li, P. P. C. Lee, and W. Lou, “Secure Deduplication with Efficient and Reliable Convergent Key Management,” *IEEE Transactions on Parallel and Distributed Systems*, vol. 25(6), pp. 1615–1625, 2014.
4. R. Chen, Y. Mu, G. Yang, and F. Guo,“BL-MLE: block-level message locked encryption for secure large file deduplication,” *IEEE Transactions on Information Forensics and Security*, vol. 10(12), pp. 2643–2652, 2015.
5. M. Bellare, S. Keelveedhi, and T. Ristenpart, “Message-locked encryption and secure deduplication,” in *Proc. of the Advances in Cryptology (EUROCRYPT’13)*, vol 7881 of LNCS, pp. 296–312, 2013.
6. S. Halevi, D. Harnik, B. Pinkas, and A. Shulman-Peleg, “Proofs of ownership in remote storage systems,” in *Proc. of the 18th ACM Conference on Computer and Communications Security (CCS’11)*, pp. 491–500, 2011.
7. Chia-Mu Yu, Chi-Yuan Chen, and Han-Chieh Chao, “Proof of Ownership in Deduplicated Cloud Storage with Mobile Device Efficiency,” *IEEE Network*, vol. 29(2), pp. 51 - 55, 2015.

با تشکر از توجه شما