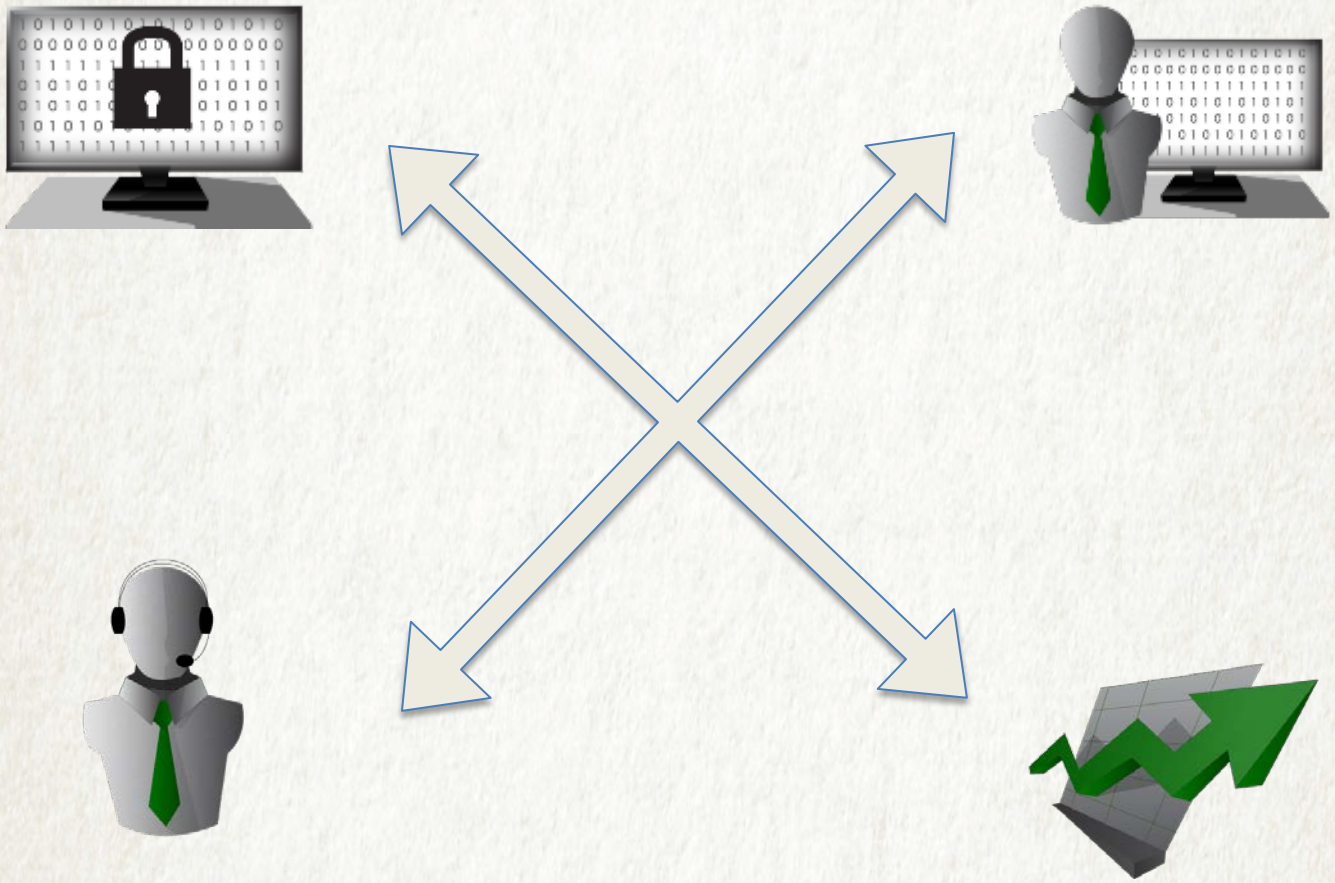
druva

Designing Flash-Intelligent Enterprise Applications

Somesh Jain

Enterprise Apps: Key Metrics



Scalable Performance

- No silver bullet
- Cost of performance
- Right-sizing hardware
- Fine-Tuning Software
- How Druva inSync does it?

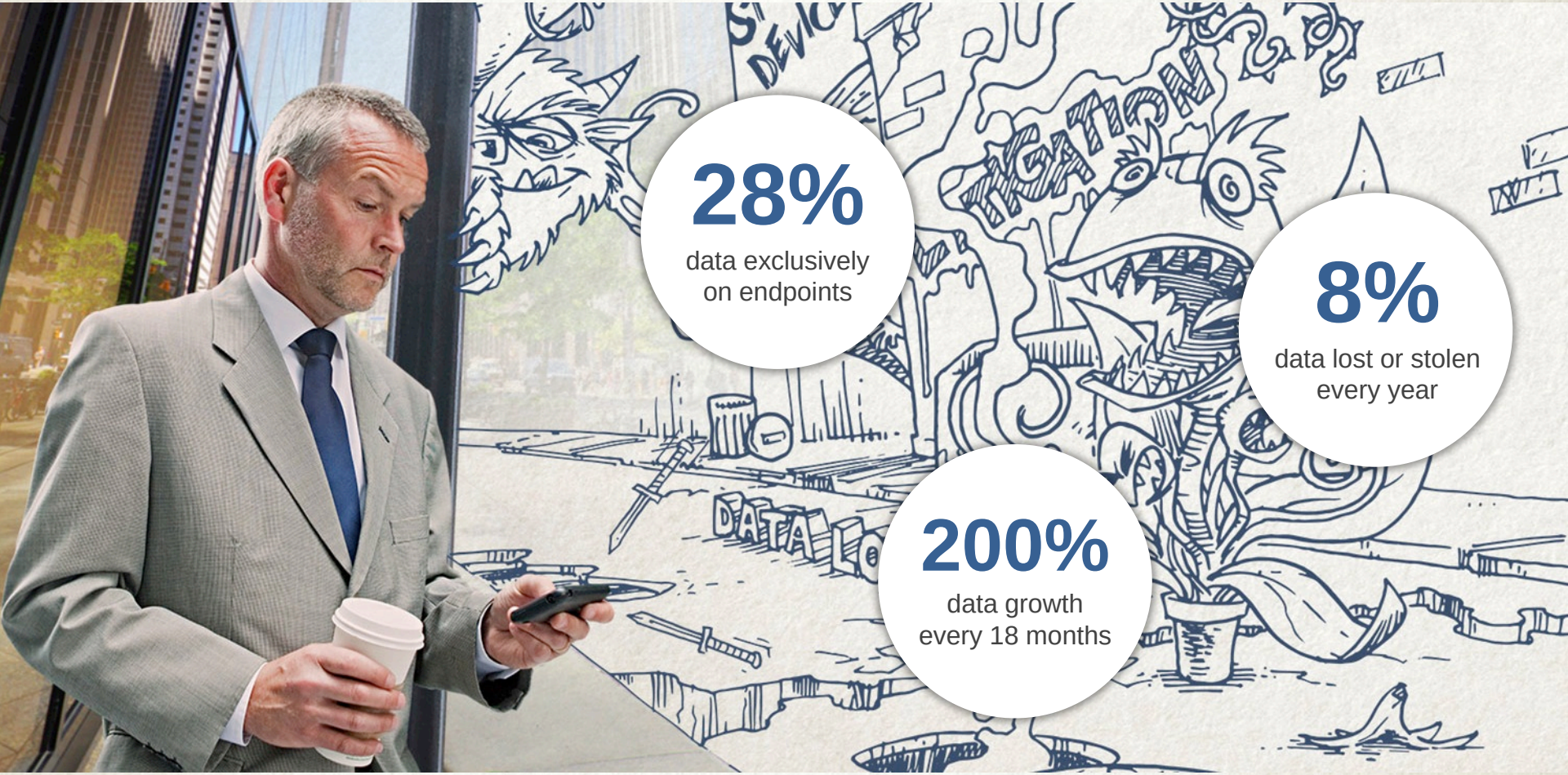
Err… But Who's Druva?
What does inSync do?

druva

Award winning Data Protection and Governance at the Edge

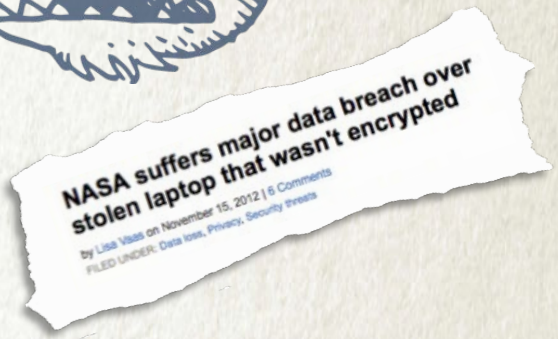
Ranked #1 by Gartner in 2012 and 2013

The Wild Frontier



The Risks Abound

Loss of data & productivity



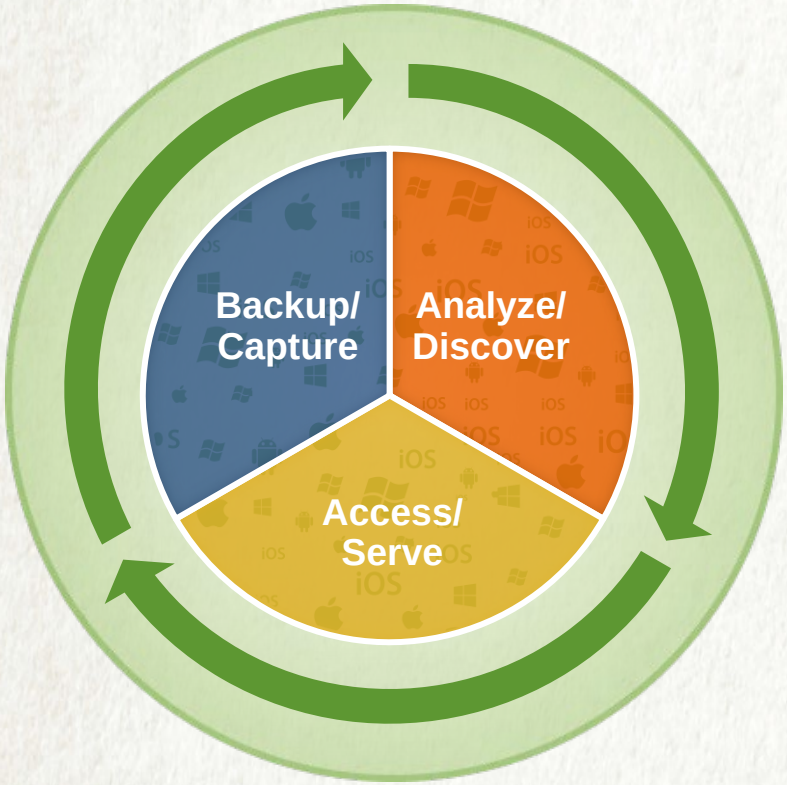
Data breaches

70%

edge devices unencrypted

Litigation & compliance exposure

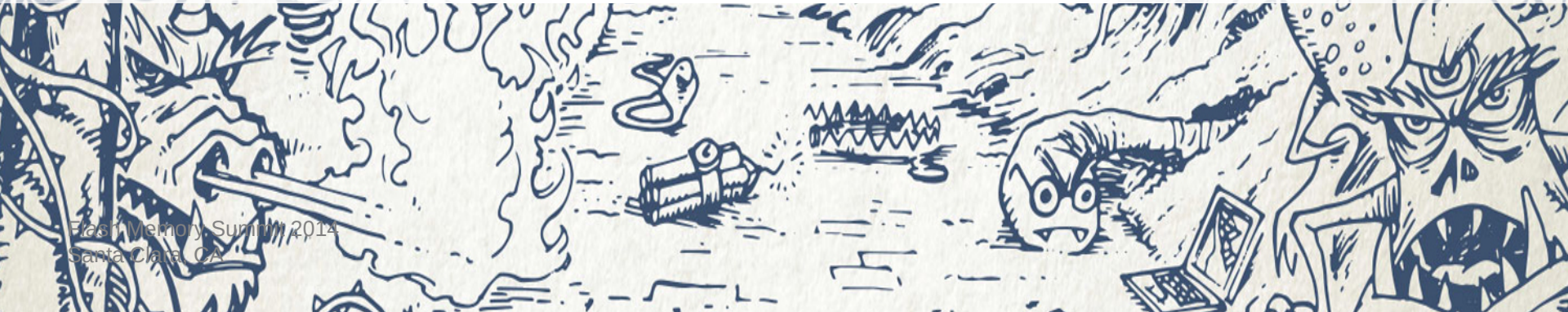
Defend Your Data in the Wild



- Protect
- Govern
- Access

druva

Performance & scalability challenges



inSync Backup USP

- Deduplication
  - Global: Duplicate backed up data from multiple laptops
  - App-aware: Variable block-size deduplication
  - Source-side : Don’t send if it exists on server
- Incremental backups
- WAN optimization

inSync Data Deduplication

- Choose file ***F1*** to backup on laptop ***L1***
- Generate finger print ***S1*** for the file's contents ***D1***

Backup F1 with S1

- Check if ***S1*** exists in dedupe index ***DD***

S1 Found

- In ***L1***'s backup index ***B1***, mark ***F1*** as backed up
- ***F1*** points to ***S1***

Send data for F1 and S1

S1 does not exist in ***DD***

- Store data ***D1***
- create new entry in ***DD*** for ***S1***

Druva inSync client

Druva inSync server

Deployment Challenges

Enterprise laptop backup

- Average backup size: *30GB*
- Number of laptops: *5000*
- Data storage capacity : 64TB + 8TB
  - Storage for data (like ***D1***)
  - Storage for indexes (***B1*** and ***DD***)
- Data storage IOPs
  - 55K RW IOPs random

Using SSDs or Flash storage

- High IOPs density
- High $/GB costs

Storage Sizing

- IOPs offered by RAID5 at 72TB capacity
  - 7000
- HDDs needed for 55K random IOPs on
  - 400
- Cost of 72TB SSDs
  - HUGE

Proposed Solutions

#1 - Split Data and Index

	Critical Accesses	Capacity	Recommended Storage
Data ***D1***	Large sequential writes	64 TB	HDD - RAID 5
Index ***B1+DD***	Random lookups & insertion	8 TB	SSD

#2 - Split Data, Index B1 & DD

	Critical Accesses	Capacity	Recommended Storage
Data ***D1***	Large sequential writes	64 TB	HDD - RAID 5
Index ***B1***	Random lookups, sequential index insertion	4.5 TB	HDD – RAID 10
Index ***DD***	Random lookups & insertion	3.5 TB	SSD

#3 – inSync HyperCache

90% dedupe happens on < 10% of data blocks and remaining 90% blocks are *unique*

- Frequently accessed finger-prints cache for faster lookups
- Updates queued for off-line commits, lower write IOPs online
- Smaller foot-print (350GB in our case)
- Use SSDs to store this

SSD as LUN Cache

- Block granularity Cache – ~100+ DD entries per DB page
- Noise introduced by recent updates - 90% non-dedupe entries

- *Poor Cache Efficiency*
- *Higher IOPs requirement*

Summary : Right-size your storage

- Know thy application access pattern
- Identify critical bottlenecks & Optimize them
- Use SSDs

Lower Enterprise Application TCO

Questions?

Thank You!

www.druva.com

somesh.jain@druva.com