The Truth about Data Integrity for MSP's
5 Questions to ask your Online Backup Software Provider

Introduction
Competition is fierce in the exploding online backup industry. With so many providers of Backup Software, whom can you trust with your customers' data? As a managed service provider, your customers are trusting you to employ solutions that will get them back their data when they come asking for it. Fewer issues are more sensitive than lost or corrupt data.

Finding a place to backup data is easy these days, but discerning which software provider can get back the *verifiably correct* data *all* the time *every* time is much harder. Slick websites and smooth-talking sales-people are no help here. This questionnaire will help you discover the empirical facts you need to determine whether or not to entrust your customers' data with an online backup provider.

Q1) Which established standards does your software follow for your cryptography?
In the complex world of cryptography, following well-established standards is the only sure path to safety. An excellent example is the proprietary GSM A5/1 cell phone encryption algorithm, which was subsequently broken. Another risk is that even if the encryption algorithm itself is standardized (such as AES), if the use of that algorithm (called *cipher mode*) does not follow standards, it is subject to serious flaws. For example, one software provider used AES in CTR mode, but chose to deviate from the NIST 800-38A standard and re-used the IV, causing their solution to become vulnerable to known-plaintext attacks. Ask about standards with respect to the following: encryption, hashing, and MAC algorithms, cipher modes, and pass phrase key generation.

Q2) Is your cryptography implementation well-known and open-source?
Cryptography is hard to implement correctly and securely, especially if it needs to be fast. Improper implementations are vulnerable to timing attacks. Bugs can also cause data corruption. Another danger is the presence of "back-doors" in an implementation that would allow access to the data without the encryption key – if a provider is using an established open-source cryptography library, the community has scrutinized the source code to make sure that it is correct, secure, and fast.

Q3) Which cryptographic primitives are used to protect the integrity of the data?
Some software providers focus so much on using cryptography to protect the confidentiality of your data that they do not consider another important aspect – data integrity. Encryption only provides secrecy but not data integrity. This is why cryptographic message authentication codes (MACs) must be used in addition to encryption. A MAC provides a cryptographic fingerprint that detects malicious tampering and accidental or silent corruption. Ask which MAC algorithm is used, whether MAC fingerprints are stored on disk, and whether the fingerprints are verified upon restore.

Q4) Which mechanisms and processes are used to protect against silent-data corruption?
Silent data corruption is caused by physical failures, corrupted or buggy firmware, misdirected writes, driver bugs, file system bugs, and human error. A recent study by CERN found that in a sample size of 8.7TB with 33700 files, 1 in 1500 files had some corruption, with an overall bit error ratio (BER) of 1×10^{-7}. Any hardware-only solution, including RAID, will not provide end-to-end coverage of all issues. Ask what technology is used to detect and repair silent data corruption, length of block checksums, where those checksums are verified and repaired, how much data redundancy is employed for repair, and whether the cipher mode is sensitive to single-bit errors. Be wary of providers that say integrity is provided through mirrored data centers without mechanisms specifically for silent data corruption – without detection mechanisms any data corruption will be silently mirrored to the other data center as well.

Q5) How often is the integrity of actively changing and archived data actively verified?
Frequently verifying data integrity mitigates risk through early detection and repair. Also, while systems should have a defense-in-depth solution so that silent data corruption never occurs, responsible solutions will have an

open reporting policy, which is especially important for regulatory compliance. Ask how often the integrity of actively changing is verified, how often the integrity of archived data (historical and non-changing data) is verified, and who is notified if corruption ever occurs.

Conclusion
No matter who you choose for your online backup software provider, make sure it's one that provides the highest level of data integrity protection. Finding a solid technical solution now will give you the confidence to sell to your customers without reservations or doubts.

For questions or feedback, contact our team at Support@BluPointe.net or by phone at **(800) 580-9402 - Ext 231**

The Truth about Data Integrity
Online Backup Provider Questionnaire Worksheet

	BluPointe	Provider 2	Provider 3
Q1) Which established standards do you follow for your cryptography?			
Standards for encryption algorithms?	AES-256 bit FIPS-197		
Standards for cipher modes?	CTR NIST 800-38A RFC 3686		
Standards for MAC algorithms?	HMAC-SHA-256 RFC 2404		
Standards for hashing algorithms?	SHA-256 RFC 2104		
Standards for pass phrase key generation?	PBKDF2 RFC 2898		
Standards for asymmetric cryptography?	RSA-3072 bit		
Q2) Is your cryptography implementation well-known and open source?			
Name of cryptography library?	OpenSSL		
Is cryptography library open source?	Yes		
Q3) Which cryptographic primitives are used to protect the integrity of the data?			
MAC sent/verified during transmission?	Yes		
Network MAC algorithm?	HMAC-SHA-1		
MAC stored on-disk with data?	Yes		
On-disk MAC verified during restore?	Yes		
MAC mismatches reported during restore?	Yes		
On-disk MAC algorithm?	HMAC-SHA-256		
On-disk MAC based on strong cryptography?	Yes		
Q4) Which mechanisms and processes are used to protect against silent-data corruption?			
Technology to detect silent-data corruption?	256-bit error-correcting software checksums, End2End hardware ECC		
Use of software-based checksums?	Yes		
Length of checksum?	256-bit		
Level of data redundancy for repair?	Close to Triple Mirror		
Estimated BER of detect/repair technology?	10^{-45}		
Cipher mode sensitive to single-bit errors?	No		
Q5) How often is the integrity of actively changing and archived data actively verified?			
How often is actively changing data verified?	Every Backup		
How often is archived data verified?	1-2 Times Monthly		
Is all redundant data verified as well?	Yes		
Corruption notification policy?	Reseller immediately contacted with file name and block #		

Refusal of a provider to disclose high-level information because of "security concerns" is in opposition to a well-known principle in cryptography called Kerckhoff's Law, and may be a sign of insecure design choices or lack of confidence in the security of the cryptographic primitives employed by their solution.