What is deduplication?

Data deduplication removes multiple copies of the same data stored in a system, thereby reducing storage consumption.

Are all deduplication technologies the same?

No, there are many varieties of deduplication that share the same name. Key variations among deduplication solutions include:

 - Some require more memory per usable capacity
 - Some can scale more than others
 - Some run post-process versus inline

The ability to deliver scalability over 256TB is a requisite for today's data stores. With IOPS speeds driven by flash technology, a minimum of 200K IOPS is required today. With today's RAM costs, efficient RAM/scalability ratios are required.

What is compression?

Compression reduces data redundancies by simplifying patterns.

What makes compression different than deduplication?

Deduplication, as the name implies, reduces the size required for lots of copies of the same or highly similar data – for example, operating system images in virtual desktop infrastructure (VDI) deployments. Compression reduces data by detecting and simplifying patterns, but does not require exact duplicate data, so it tends to benefit more general workloads like databases. Deduplication, compression, space-efficient snapshots, and thin provisioning should be leveraged together to provide maximum data reduction for mixed enterprise environments.

What is the best "use case" for deduplication?

The best use case for deduplication is in unstructured data such as VM/ VDI files or backup data. Data types that have high redundancy across many files typically yield the highest storage savings with deduplication.

What is the best "use case" for compression?

The best use case for compression is for structured data such as databases, where reduction by simplifying patterns will yield optimal storage savings.

IBM FlashSystem and Permabit FAQ

Do deduplication and compression complement each other?

Yes, because businesses typically have BOTH structured and unstructured data in their data stores. In fact, the greatest data reduction and strongest storage savings occurs when both dedupe and compression are run together, alongside thin provisioning and space-efficient snapshots.

How do I deploy Permabit SANBlox with FlashSystem?

Permabit SANBlox appliances are connected through a Fibre Channel SAN switch to either FlashSystem 900 or FlashSystem V9000 products. Data should be directed to SANblox first as a "front-end" for FlashSystem. A SANblox appliance can be installed in under 30 minutes with full configuration completed in a few hours. SANblox appliances are installed in pairs, each of which is cabled to Fibre Channel SAN and Ethernet switches (for management). After cables are connected, FlashSystem volumes are provisioned for SANblox, and the SANblox appliances are configured to offer storage to application hosts.

If I have a FlashSystem V9000, can I use a Permabit SANblox appliance for dedupe only, to complement IBM Real-time Compression?

Yes. Permabit SANBlox is available in a "deduplication only" version for IBM FlashSystem V9000, allowing clients to leverage the enterprise features from the FlashSystem V9000 platform, including IBM Real-time Compression, migration, thin-provisioning, and IBM Easy Tier®.