

Tintri IntelliFlash

Tintri IntelliFlash is a storage product, formerly owned by Western Digital, that evolved from the Tegile systems. In September 2019, Western Digital's IntelliFlash business was purchased by DDN and organized under its Tintri brand of enterprise storage solutions.

The system design leverages common hardware platforms with systems utilizing standard bridge-bay controllers and storage in a single 2U enclosure. Storage controller features are delivered through a combination of open source ZFS, combined with Western Digital developed features, CLI and GUI.

IntelliFlash systems are designed as general purpose, high performance storage systems competing with other midrange storage systems. The IntelliFlash systems compete against HPE Primera, Dell EMC Unity XT, NetApp EF Systems, IBM FlashSystem, Dell EMC XtremIO, Hitachi VSP F-Series and other leading all-flash and hybrid storage systems.

Highlights

- Multi-protocol, file and block access
 - FC
 - iSCSI
 - NFS
 - SMB 3.0
- All software included in license price
 - Deduplication
 - Compression
 - Snapshots (nearly unlimited)
 - Replication (async)
 - Live Dataset Migration
 - Hybrid Cloud Connector
- Three product lines
 - N-Series NVMe All Flash
 - H-Series Hybrid with NVMe Flash
- SAS disks + SAS SSD in either 3.5" LFF or 2.5" SFF
- PCIe attached SSDs with NVMe protocol in 2.5" SFF
- IntelliFlash Architecture -
 - Meta-data from ZFS filesystem is cached, accelerating some operations
 - Active / Active dual controllers (NFS is active / passive)
 - RAID 1, RAID 5, RAID 6, 3 parity RAID, 3 way mirror
 - Data may be pinned in flash, must be brought onto SSD, then remains until reboot
- SSDs as Cache
 - Separate read and write cache built from pool of SSDs
 - Default is 10 GB / SSD for write cache

- Remainder of SSD for level 2 read cache
- FlashVols
 - Volumes can be pinned in SSD for high performance access
- Controller with disks, along with expansion chassis
 - Controllers are 2U, including disk. InfiniFlash module is 3U
- VMware Integration
 - iSCSI VAAI - Thin provisioning, Block Zero, Full Copy, HW Assisted Locking
 - FCP VAAI - Thin provisioning, Block Zero, Full Copy, HW Assisted Locking
 - NAS VAAI - None
 - VMware vCenter - Web Management Console GUI plugin
- Encryption with internally managed keys

Overview

IntelliFlash storage systems provide multi-protocol storage with both spinning media and NAND flash solid state disks as cache in the hybrid models. In some configurations, an all-flash system is possible with all data residing on flash.

Target applications are those requiring high performance, particularly those with large capacity requirements. IntelliFlash systems are designed as general purpose, high performance storage systems competing with other midrange storage systems. Target applications are those requiring high performance, particularly those with large capacity requirement. Effective capacity with compression and deduplication is listed as 4X the raw capacity.

IntelliFlash systems are highly available (HA), providing all hot-swap components with active / active controllers with symmetric LUN access. This enables non-disruptive access during component failures and upgrades including host access paths or storage controllers. The controller firmware supports FCP and iSCSI block access along with CIFS and NFS file access via the HA controller nodes.

IntelliFlash software offers support for VMware Web client and vCenter plugins. The VMware support enables users to perform storage allocation along with performance and capacity management from within the VMware console. Additionally, support for VMware VAAI provides enhanced performance with VMware based workloads including thin reclamation, block zeroing and virtual machine cloning operations.

The ability to pin data into flash SSDs in their hybrid models may further enhance performance. Specific volumes or filesystem mount points for NFS may be designated as residing in flash. When pinned, data must be copied onto the SSD cache but will remain in the cache until the system is restarted.

IntelliFlash does support asynchronous storage replication, an included feature at no additional cost, which is a differentiating factor versus some next generation flash systems.

IntelliFlash is a flash-optimized storage software architecture that integrates multiple grades of storage media. This includes features for management, storage protocols choice, data services, metadata management, and media optimization.

Evaluator Group comment: IntelliFlash is targeted as a primary, tier 1 storage where a customer would use IntelliFlash for a variety of enterprise applications including all their critical applications with their all-flash models and not limited to a few special workloads. Currently the support for data replication technologies expected in tier one storage systems that enables this usage is not yet available on IntelliFlash products.



Figure 1: Tintri IntelliFlash storage system (Source: Tintri)

Product Architecture

The IntelliFlash is a commodity controller design that supports multiple interfaces with pluggable modules for advanced interface connections including 16Gb/s Fibre Channel, 40GigE iSCSI and file protocols (NFS and CIFS/SMB 3). Each controller contains DRAM in varying capacity depending on the model, up to 1.5TB. SSD ports use SAS interfaces at 12Gb/s with four lanes (wide SAS) for an aggregated 24 Gb/s or PCIe attached with the NVMe protocol. The controllers have dual RJ45 Ethernet 10/100 BaseT connectors and USB serial management and configuration ports. All interface ports are small form factor pluggable (SFPs) connector interfaces. The battery for data and cache protection is a long life (lithium Ion) battery. The IntelliFlash systems are enclosures with two RAID active/active controllers with mirrored write back cache, host interface ports, and redundant power and cooling.

Flash is used to enhance the performance of the log-structured filesystem at the core of IntelliFlash systems. Use of Flash accelerates write performance by allowing the system to acknowledge a write and then coalesce data before storing it across solid-state drive modules. Recently accessed data is stored in a separate Flash read cache, enhancing read performance.

All current IntelliFlash models are dual controller systems for high availability and include data protection features. The IntelliFlash architecture includes multiple RAID levels, including dual (RAID 6) and triple parity RAID, along with two-way mirroring (RAID 1) and 3 way mirrors. Space efficient snapshots are included with the system, which supports a nearly unlimited number of both read only and read-write snapshots. Additionally, remote asynchronous replication over iSCSI is available.

Incoming data is written to flash acting as NVRAM and mirrored to the partner controller in HA configurations over a PCIe interconnect using Non-Transparent Bridging. Next, a write acknowledgement is sent to the client. In-line data deduplication and compression are available, using unique hashes as a first check and then a byte-wise comparison for final validation of data duplication. The tracking of duplicate blocks of data is handled by the MASS storage controller, which also performs

compression. Data may be pinned to SSD for high performance volumes dedicated to specific applications.

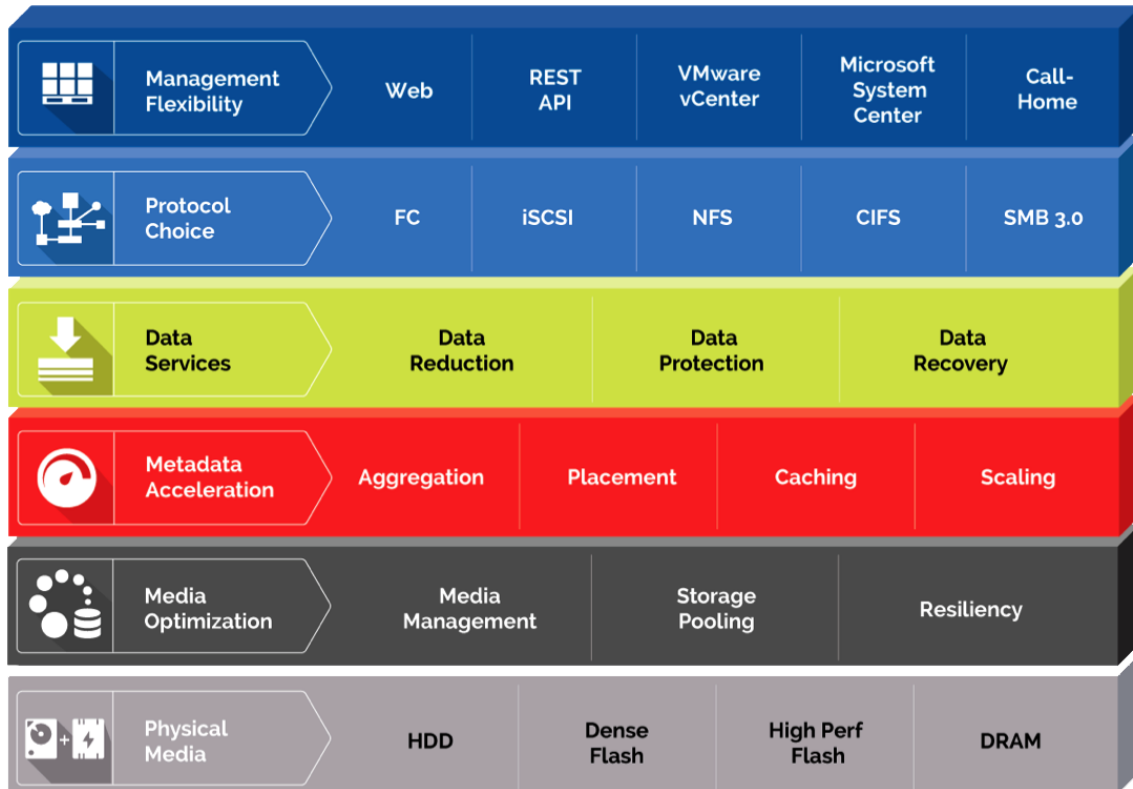


Figure 2 - IntelliFlash Software Architecture

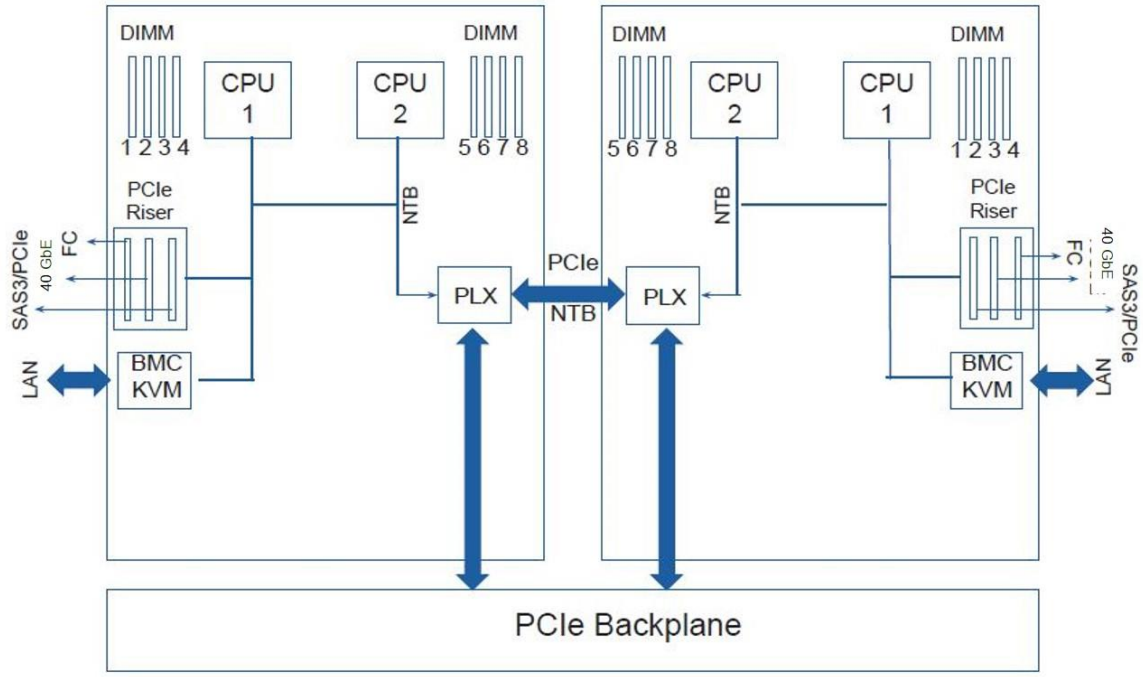


Figure 3: IntelliFlash N-Series Hardware Architecture

Tintri IntelliFlash Summary

N Series

Model	Cache	Storage Devices	Connection Ports	Base Raw Capacity	Base + Max Expansion Raw
N6100	448 GB RAM	24 dual ported PCIe SSDs	8 @ (16 Gb FC or 40 GbE)	46 – 184 TB	920 TB
N6200	448 GB RAM	24 dual ported PCIe SSDs	8 @ (16 Gb FC or 40 GbE)	46 - 384 TB	1.8 PB

Table 1: Tintri IntelliFlash N Series Models

H Series Hybrid NVMe

Model	Flash Capacity (Raw)	HDD Capacity	Ethernet Ports	FC Ports
H6200	46 – 368 TB	360 – 5040 TB	16x 10GbE, or 8x 25GbE, or 8x 100Gb	8x 8G or 16x 16G FC

Table 2: Tintri IntelliFlash H Series Models

Advanced Features and Functions

The features offered in the IntelliFlash storage software architecture are supported by all hybrid and all flash arrays in the product family

System Management

There are multiple management options:

- Web UI
- VMware vCenter web client and desktop client plugins
- Microsoft System Center Virtual Machine Manager (SCVMM)

Provisioning – IntelliFlash virtualizes the physical media into a pool that can be allocated as LUNs or file shares. Pool capacity can be expanded online when needed.

Striping – IntelliFlash uses dynamic stripe widths to reduce overhead and read-modify-write operations that happen with traditional RAID. Unallocated data is not copied when a drive is rebuilt.

Data Reduction and Capacity Optimization

Inline deduplication and compression are used in IntelliFlash to reduce the number of writes that actually get written to SSDs. Deduplication can be enabled at the storage pool, per LUN, or file share level and supports different block sizes for each LUN from 4KB to 128KB. Compression can also be enabled at the storage pool, per LUN, or file share level for capacity optimization with a choice of compression algorithms to meet the needs of different workloads. Thin provisioning is also used in IntelliFlash for capacity optimization.

Live Dataset Migration

Live Dataset Migration allows IntelliFlash storage pools to be moved from one IntelliFlash array to another IntelliFlash array. iSCSI/FC LUNs can be migrated while the system is operating in order to maintain productivity and continuity across storage systems.

Hybrid-Cloud Data Mobility

IntelliFlash supports hybrid cloud capabilities by using the S3 Cloud Connector. IntelliFlash can migrate snapshots to AWS S3 as well as other S3 enabled clouds and object storage systems. This feature can be leveraged to provide additional data protection and cost savings.

Local/Remote Replication

Local

IntelliFlash software has reservationless thin snapshots and thin clones. Thin point-in-time copies can be made of a single or multiple volumes. Thin clones are writeable and mountable. Space allocations are only made for change blocks after the initial copy for both thin snapshots and thin clones.

Remote

Periodic asynchronous replication with data compression is used for less traffic on the WAN across datacenters and optimizes the amount of data transmitted for replication. This is what IntelliFlash uses to facilitate disaster recovery.

Caching and Acceleration

The adaptive caching algorithms are optimized for common I/O patterns to gain additional performance. Multiple levels of cache (DRAM and SSDs) are provided to handle multiple media latency levels. The devices used for caching are mirrored for data protection and will automatically expand space for metadata as needed for scaling. The read cache is persistent across controllers in the case of failures. This can be useful in retaining performance levels when there is a controller failure.

Media Management

Alignment – Misalignment can cause degradation in workload performance as much as 30% in some cases. Writes are aligned on sector boundaries and page sizes to both mitigate fragmentation and eliminate additional writes caused by misalignment.

Wear Balancing – IntelliFlash keeps track of wear on SSDs and moves data to uniformly balance wear across the array.

Garbage Collection – IntelliFlash works with the host operating system to identify blocks marked for deletion and continually frees up blocks to minimize write amplification.

Metadata Management

For performance and less fragmentation, IntelliFlash separates metadata and user data. Metadata is stored on a low latency media for higher performance of data services like deduplication, compression, snapshots, clones, and thin provisioning. Metadata is managed in a multi-layered data structure that is distributed across multiple SSDs for data protection.

Reliability, Availability, Serviceability Features

The IntelliFlash systems maintain data integrity and provide data availability via the following mechanisms:

- Dual controllers with InfiniBand interconnect and failover for high availability
- Redundant, hot-swappable components:
 - Controllers
 - Fans
 - Power supplies
- Redundant media fabrics – aggregating I/O ports within and across controllers
- Online hot sparing
- Non-disruptive firmware and component upgrades
- Checksums of every data block written - The checksum and the data are stored on separate nodes for further integrity.
- Auto-healing to maintain data integrity - If a block is considered corrupted, a mirror copy is used.
- Persistent cache backup in case of power outage
- Storage of configuration metadata on each disk drive
- Alert thresholding
- Error notification and logging
- Remote call home
- Data Recovery – Snapshots are used to recover data when data is lost or corrupted
- Multiple RAID levels, including dual parity (RAID 6) and triple parity

Evaluator Group Comments

IntelliFlash systems are multi-protocol storage designed for use in general purpose environments with high performance requirements or multiple workloads running on the same storage system. The multi-protocol file and block access provides both application and host connectivity options, which are particularly well suited to deployments where flexibility is an advantage. This product offers advanced enterprise storage features like embedded data reduction technology for capacity optimization and local/remote replication. It is unusual for a new-design all-flash array to be offered with a full suite of advanced functions.

Perceived Strengths:

IntelliFlash has several strengths, namely multi-protocol access on all of their systems, along with a broad range of products that include hybrid, all-flash and high-density flash systems. Additionally, since IntelliFlash is based on ZFS, many of their features are derived from a proven enterprise storage software that provides the foundation of their design.

The asynchronous local/remote replication and inline data reduction capabilities are also significant features that can differentiate IntelliFlash from other systems. For some workloads, such as VDI data deduplication can have a significant impact on storage capacity utilization, while data compression can be significant advantage for some data sets and database applications.

For virtualized environments, the number one performance problem reported is storage performance. Flash optimized storage is a response to customer need for storage performance. Solid-state storage is well suited for accelerating virtual infrastructure performance in transaction-intensive datacenters while saving OPEX of power and cooling. Even through the initial cost of solid-state storage can be higher than HDDs the savings in power and cooling will more than make up for the difference.

IntelliFlash continues the focus on advanced storage features and performance. The IntelliFlash provides economic value with the ability to customize the use of solid state technology and pin applications in flash to accelerate applications. In addition to the performance gains in traditional applications and server virtualization environments, other workloads in geo-science, video and imaging, telemetry, visualization and stereo graphics, simulations will benefit from IntelliFlash solid state performance. The customizable performance gains with IntelliFlash will allow a customer to consolidate more storage systems with greater efficiency rather than focus on over-provisioning capacity to meet performance demands.

Perceived Challenges:

IntelliFlash has had success in mid-market enterprise accounts. However, they have not established an entry into large enterprises on a consistent basis, where their competitors have the advantage of a direct sales force. This is in part due to Western Digital's sales approach. It remains to be seen how the IntelliFlash will capture the market under DDN.



More detailed information is available at <http://evaluatorgroup.com>

Copyright 2021 Evaluator Group, Inc. All rights reserved.

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying and recording, or stored in a database or retrieval system for any purpose without the express written consent of Evaluator Group Inc. The information contained in this document is subject to change without notice. Evaluator Group assumes no responsibility for errors or omissions. Evaluator Group makes no expressed or implied warranties in this document relating to the use or operation of the products described herein. In no event shall Evaluator Group be liable for any indirect, special, inconsequential or incidental damages arising out of or associated with any aspect of this publication, even if advised of the possibility of such damages. The Evaluator Series is a trademark of Evaluator Group, Inc. All other trademarks are the property of their respective companies.