Workload Data Importer for LDX-Enterprise

Analyze Your Production Workload with Data from Arrays or Mgmt Tools

Introduction

The ground-breaking Workload Data Importer software module of Load DynamiX Enterprise helps you understand your production workload I/O profile.

Performance Planning Challenges

Applications, networking, and storage teams often don't know their own workload profiles, and their vendors really aren't in a position to help. Consequently, performance planning is guesswork, sometimes resulting in under provisioning, but more often, over provisioning. If application workload performance optimization, or storage cost optimization are attempted at all, it's nearly always done by trial and error.

The typical infrastructure improvement scenario is to deploy, listen for user complaints, and reactively adjust. Even if the deployment goes well initially, any number of changes can affect performance of the production storage infrastructure. Changes might be characterized as:

- Business driven
    - More users – sometimes from mergers and acquisitions
    - Fewer users – often from downsizing or consolidation
    - New customers – a typical Service Provider scenario
- Application driven
    - Software/DBMS upgrades
    - New features

- Infrastructure driven
    - Configuration – such as HBA settings
    - Firmware/OS updates
    - Server upgrades – physical or virtual

Current performance monitoring solutions are too siloed to provide comprehensive workload visibility. Changes may affect latency, throughput or IOPS, but Application Performance Monitors (APMs) or Infrastructure Performance Monitors (IPMs) can't correlate changes to their effects because they only see a part of the data path and because they often can't associate a workload with an infrastructure resource.

Workload Modeling Challenges

Until now, there have been two ways of doing workload modeling, quickly with poor realism, or slowly and painfully with better realism.

When engineers get tasked to run a workload simulation, they often seek a freeware tool, and cobble together a small lab based on borrowed PCs that no one else is using, because that's all they know. And they don't have to ask for a budget. They can use these command-line programs to build simple scripts that describe simple workload behavior … read/writes, block sizes, IOPS, and perhaps dedupe ratios. And for small tests of simple applications, this often makes sense. But if you are trying to build a realistic simulation to intelligently make million dollar decisions, you need to build realistic tests.

Gold Standard of Storage Workload Modeling

Workload Realism is defined by the combination of these three vectors and how they vary over time.

Access Pattern
R/W ratio
Seq / Rnd ratio
Command Mix
Load Levels

Data Content
Compression Ratio
Deduplication Ratio
Specific Byte Stream

Spatial Locality
Load across LUNs
R/W Hot Spots

VI is the only vendor that captures, models and simulates real workload behavior

---

Realistic simulations often contain a deep description of workload behavior, far beyond what scripting tools like Iometer can build. Using just block workloads as an example, here are the details than can be characterized:

The environment

- The number of physical initiators in the environment (HBA or NIC's)
- Number of storage ports
- Port bandwidth
- The average # of virtual ports per physical port
- Number of Logical Units (LU's) per storage port
- Size of each LUN
- IO hotspots of each LUN
- Zoning configuration
- The initiator to target fan-in ratio (the average number of initiators that access each target port)
- For Fibre Channel, the queue depth setting in the switch environment for the server HBAs

Access Patterns

- Read / Write ratio
- Read Random / Sequential ratio
- Write Random / Sequential ratio
- Distribution of I/O request sizes (block sizes)
- Distribution of I/O request locations [logical block address regions (LBA)]
  - o Data access
  - o Hot spots

- Read and Write cache hit percentage
- Read latencies
- Write latencies
- Metadata latencies
- Typical sizes of I/O active regions
- Data content compressibility (repeating patterns)
- Data content de-duplicability of the store data (duplicate files or file segments)

Load Patterns

- Read KB/s, Write KB/s
- IOPs per storage port or LUN
- IOPs per application type
- Read IOPs
- Write IOPs
- Metadata IOPs
- Total Bandwidth (MBs)
- Read Bandwidth (MBs)
- Write Bandwidth (MBs)
- Number of Connections
- Number of Pending/Queued exchanges

Temporality

- Access Patterns over time
- Load Patterns over time

If you would like to more fully explore the subject of workload characteristics and modeling for block, file or object workloads, ask your Virtual Instruments

representative or reseller for a free copy of our Workload Characterization whitepaper.

The more details you can include in your tests, the more realistic and accurate your simulation will be. Before the availability of Load DynamiX Enterprise and the Workload Data Importer, only engineers who work with these characteristics on a daily basis possess the experience and knowledge to incorporate a significant portion of these data points into models. So you were stuck with doing simple, unrealistic tests, or spending weeks or even months building better tests. Virtual Instruments and has now dramatically simplified this process and enables unsurpassed realism.

If you want to more completely understand the limitations of simple testing, you can download this very useful whitepaper on the subject.

Workload Data Importer Summary

For the first time, you can create incredibly realistic models quickly and programmatically. What used to take weeks or months can now be done in minutes to hours. The Workload Data Importer (WDI) ingests files directly from your production array, performance monitoring tools, or from

VirtualWisdom, and builds the workload model for you – automatically. For the first time, you can now easily create hyper-realistic block and file (NFSv3) workload models. And run them at huge scale on Load DynamiX Enterprise load generation appliances. The WDI provides out-of-the-box policies for most major storage vendors and allows users to easily add profiles from newer vendors.

Workload Data Importer Details

Everything starts with the Load DynamiX Enterprise 5.x dashboard / homepage. The basic workflow described on the home page starts with the Library, which is the default content provided by Load DynamiX for acquiring, generating, and analysis tasks. The Acquire functionalities includes the Workload Data Importer, highlighted in the red box, which enables you to acquire and profile workloads from your live production storage environment.

To import and model a workload, simply select Workload Data Importer. (We cover other Load DynamiX Enterprise functions you see below in the dashboard, elsewhere.)
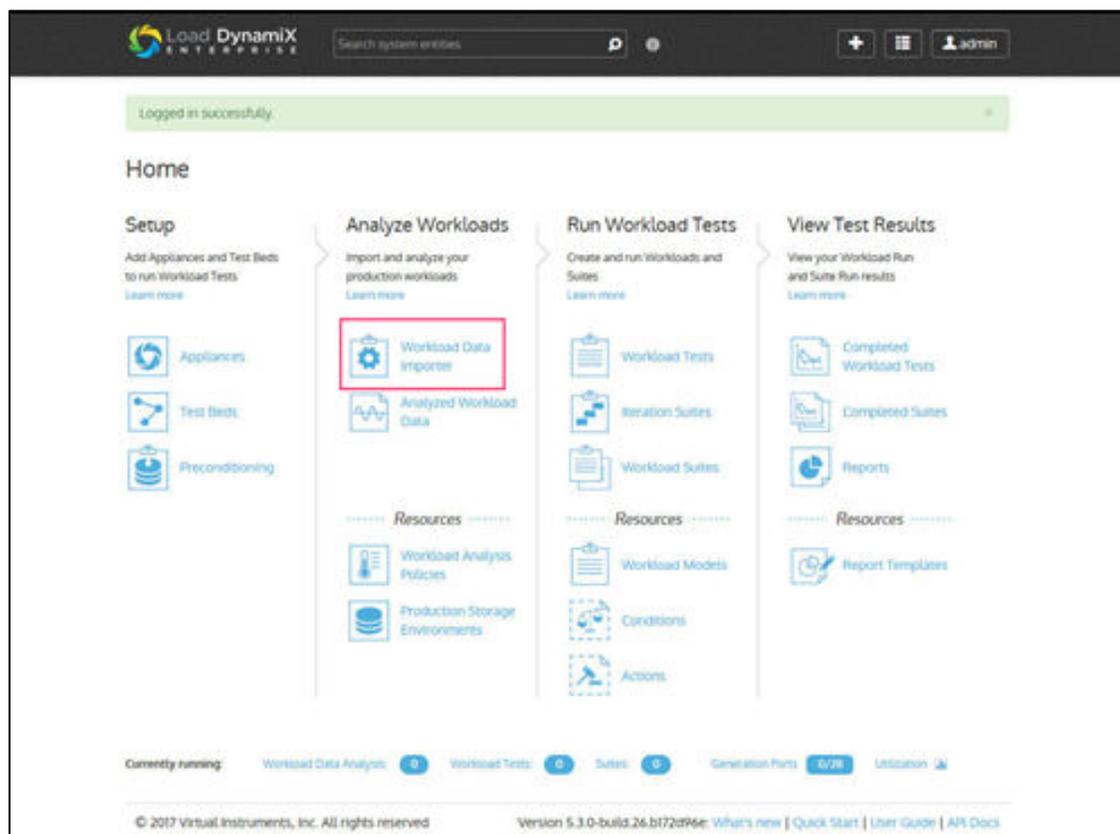

Figure 1: Load DynamiX Enterprise dashboard, with WDI highlighted

Workload data can be exported from storage vendor monitoring tools or server utilities as Comma Separated Variable (CSV) text files. Details on how to do this is contained in the Workload Characterization Whitepaper

mentioned earlier. Below is the WDI input screen. Select your environment, name it, describe it, add tags, then import your CSV file. You can drag and drop, or select your file.
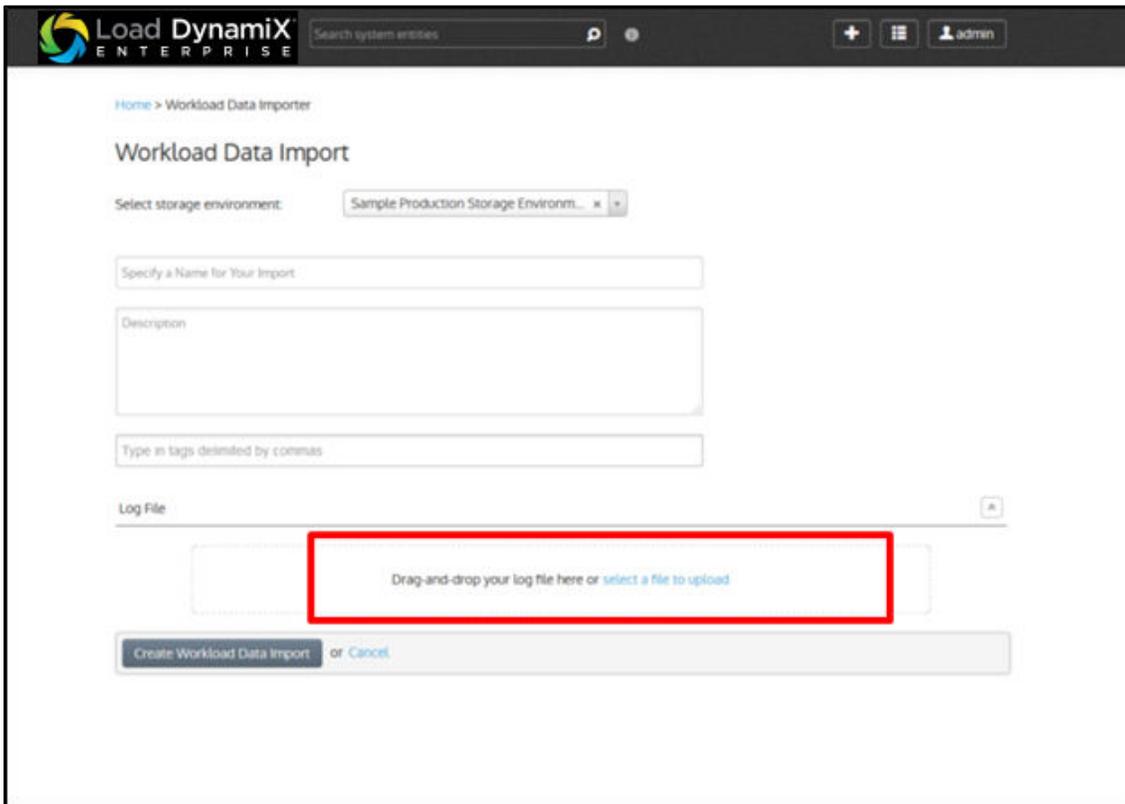

Figure 2: WDI input screen

Here's a small snippet of what a simple CSV file might contain

Date and Time	LDE V Number	Read I/O /sec	Write I/O /sec	Read Response Rate	Write Response Rate	Total Response Rate	Read Hit %	Read Xfer /sec	Write Xfer /sec	Sequential Total I/O /sec	Sequential Total Xfer /sec	Random Total I/O /sec	Random Total Xfer /sec
1 04 2016 00:10:00	0:24:00	195.26334	52.17667	340.00034	94.24151	299.9492	9.808588	1.2933333	0.2933333	2.27	0.02	195.17	1.566667
1 04 2016 00:10:00	0:24:01	196.26666	48.13333	339.39294	92.0433	303.5968	9.693264	1.3	0.2466667	0.9	0.006666667	193.5	1.536667
1 04 2016 00:10:00	0:24:02	194.86333	50.96667	339.4463	93.91167	300.6198	10.77256	1.2866666	0.2733333	1.74	0.016666668	194.09	1.543333
1 04 2016 00:10:00	0:24:03	195.87334	49.46	336.2149	95.1647	299.8599	11.65146	1.2933333	0.2733333	2.1066666	0.02	193.22667	1.546667
1 04 2016 00:10:00	0:24:04	195.35333	57.28667	337.47516	94.12151	303.0051	10.93618	1.29	0.24	0.7833334	0.006666667	191.85667	1.523333
1 04 2016 00:10:00	0:24:05	198.74	49.86333	334.5485	92.57019	298.163	11.83675	1.32	0.2733333	2.7533333	0.023333333	195.85	1.57
1 04 2016 00:10:00	0:24:06	195.81	48.32333	333.61197	94.04577	298.6601	12.257	1.3233334	0.25	1.04	0.01	193.09334	1.56
1 04 2016 00:10:00	0:24:07	197.13333	49.93	335.00027	94.17612	298.4239	11.90866	1.3066666	0.27	1.4533334	0.013333334	195.61	1.563333

Figure 3: Very small snippet of workload data (file)

Next, you would select the analysis policy you'd want to use. An Analysis Policy defines how a production workload will be analyzed and characterized. The output of the characterization can be used to create a test. The ideal workload is characteristically equivalent to the production workload. The Analysis Policy maps data from

different sources to a common set of parameters so that the data can be visualized. In addition, it defines what data will be considered when determining how to cluster workloads. And finally, it determines which workload will be created and how the data will be mapped to the workload. A set of default Analysis Policies are provided.
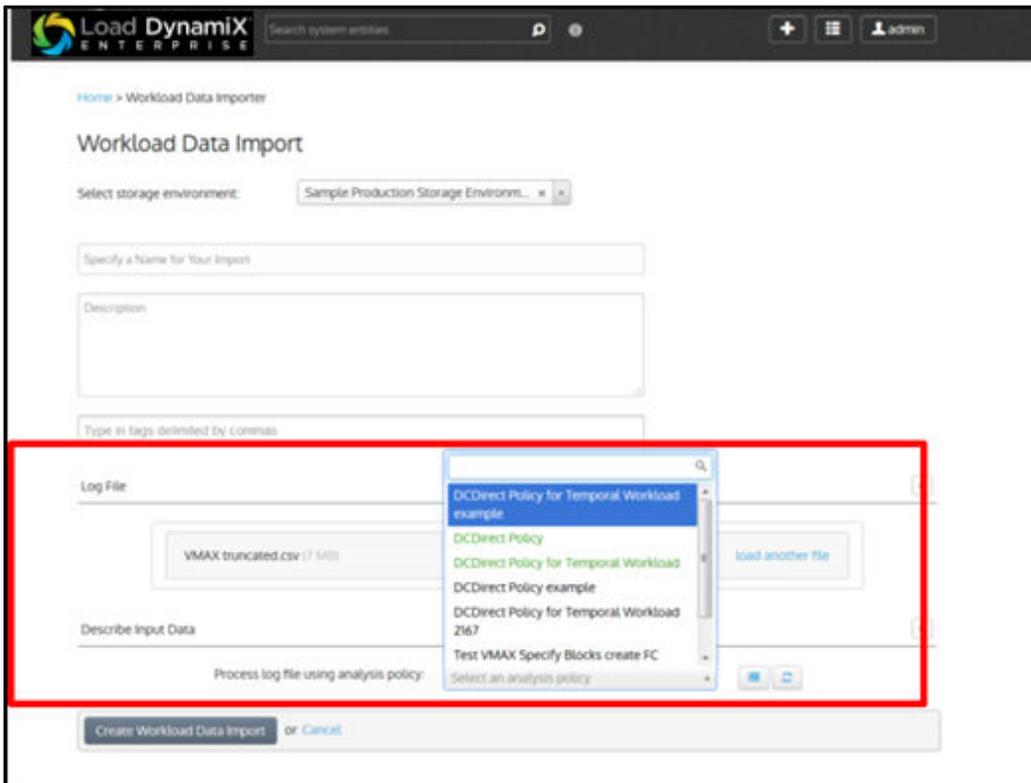

Figure 4: Selecting from the default Analysis Policies

After selecting the policy, WDI parses the file and creates the workload model. This generally takes a few minutes. Afterwards, you are presented with a standard Load DynamiX Enterprise editing screen, where you can adjust any parameter by the usual drop downs or slider adjustments to conduct a series of "what-if" analysis. And it's ready to run through the Workload Analyzer, for results similar to the screenshots below.
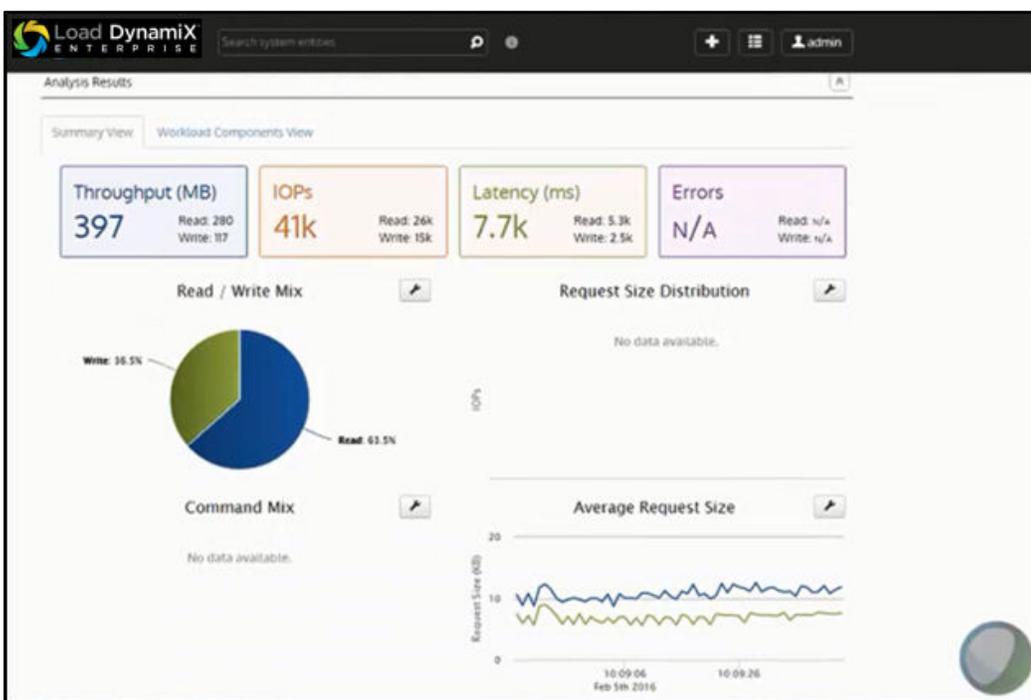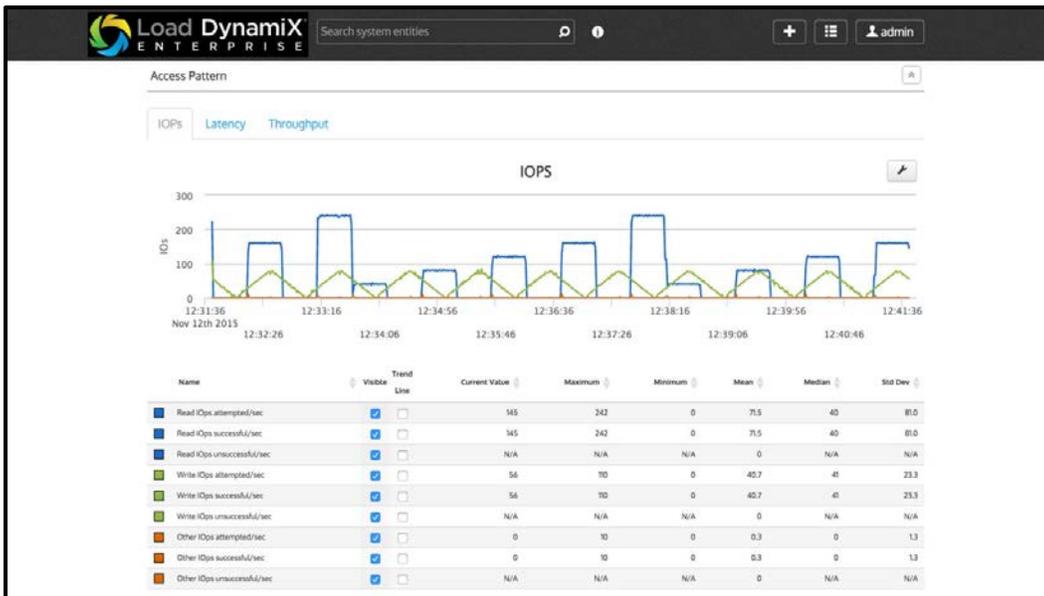
Figures 5 and 6: Snippets of Load DynamiX Enterprise Workload Analyzer output.

Summary

Load DynamiX Enterprise makes it easy to acquire and model workloads, taking what used to consume days or weeks, and compresses that to minutes. You can now take any workload, model it with WDI, manage it with Load DynamiX Enterprise, point the Load DynamiX appliance at any target array, and perform your "what if" analysis.

Integrated with VirtualWisdom, it provides a complete solution suite for Storage Performance Testing and Monitoring.

Sales
Sales@virtualinstruments.com
1.888-522.2557

Website
virtualinstruments.com