



White Paper

The Case for Open Source-RDMA and the OpenFabrics Alliance

The OpenFabrics Alliance (www.openfabrics.org) develops and distributes open source software (OFED) that is the middleware foundation for billions of dollars in revenues for computer, network and storage hardware and software manufacturers as well as their customers in product development, business intelligence, financial services, national security and defense, media content distribution, scientific and academic research. OFED enabling RDMA and multi-gigabit (up to 56Gbps in 2011) networking boosts application performance by as much as 10X, ensures the lowest latency for data motion and computation and reduces the amount of hardware to achieve these benefits by up to 50 percent.

Date: August 2011

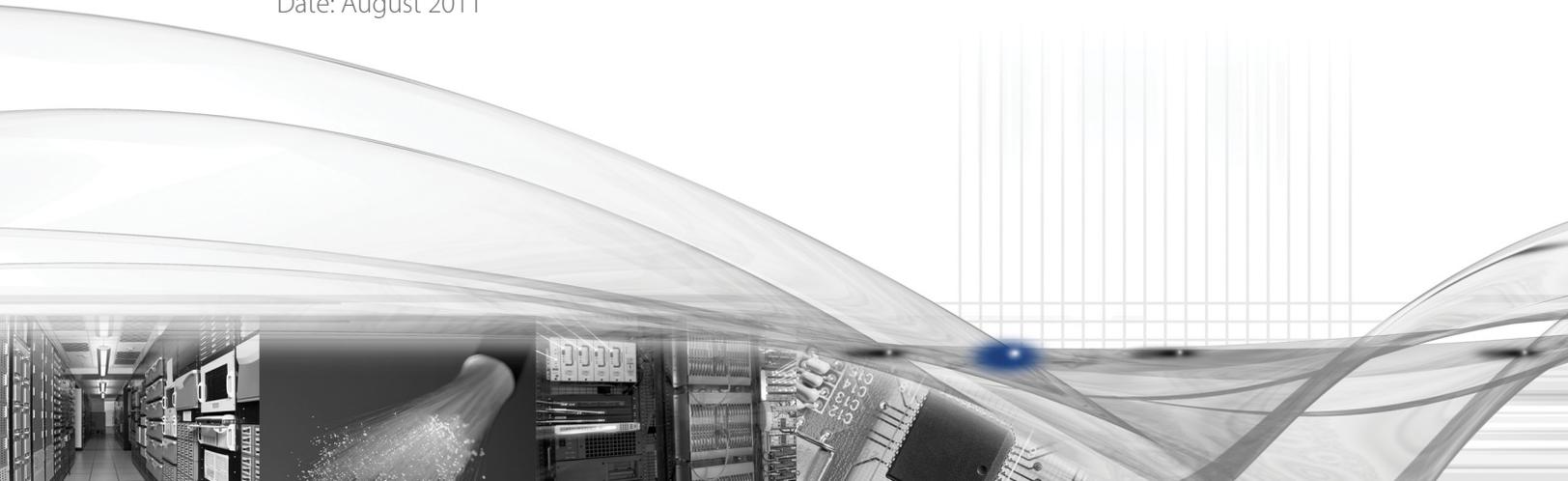


Table of contents

1. Introduction.....	1
2. What’s Unique about the OFA?.....	1
3. Benefits at a Glance.....	2
4. A Brief History of the Alliance.....	2
5. OFED Software Technology.....	3
5.1 Vendor-provided adapter device driver.....	4
5.2 Kernel bypass and RDMA API services.....	4
5.3 Upper Layer Protocols (ULPs) supported by OFED.....	4
5.4 OFED-Driven Vendor Interoperability.....	6
6. Typical OFED Use Cases.....	6
6.1 Server Clustering.....	6
6.2 Storage Clustering.....	6
6.3 Storage Area Networking (SAN)y.....	6
6.4 Network File Systems (NFS).....	6
6.5 Clustered File Systems.....	7
7. OFED Success Stories.....	7
7.1 In Academia and Research.....	7
7.2 In Weather Forecasting.....	7
7.3 In Automotive and Fluid Dynamics Design Industries.....	8
7.4 In Financial Services Industries).....	8
7.5 In Data Warehousing and OLTP Industries.....	8
7.6 In Storage Applications.....	8
7.7 In Managed Hosting or Cloud Services.....	8
8. Opportunities for OFED in Exascale, Web 2.0 & Cloud Deployments.....	9
9. Alliance Workgroups.....	9
10. OpenFabrics Alliance Membership.....	10
11. Membership Benefits and Call to Action.....	10

1. Introduction

The impact of the software produced and licensed by the OpenFabrics Alliance (OFA) in the success of world class companies is arguably one of the best kept secrets in the computing server and storage networking industry. OFA software is delivering high performance networking in a wide range of applications and environments: in the billion dollar product portfolios of giant data warehousing and financial services companies; in the leading edge scientific research and surveillance/data recording in the government, weather forecasting, natural resource exploration and academia; in the products from the automotive, defense, entertainment and media industries; and more. OpenFabrics Software is the cornerstone and common thread of such success stories that improve our data analytics processes, and the productivity, innovation and time-to-market acceleration of solving real world problems. And it does so by delivering computer, server and storage networking software that takes application performance, clustering efficiency and overall IT ROI to unprecedented levels. The message is simple, as stated by Oracle in reference to its data warehousing and online transaction processing appliances that rely on OpenFabrics Software – “Get 10X the performance at half the hardware cost!”

This paper presents the reasons for joining the OFA, describes what OFA does, how it matters to the broader computing, networking and storage communities, the benefits of membership and how to join the Alliance as well as the benefits of OFED software and how it is available.

The Case for OFA at a Glance:

- OFA members include many tier 1 and influential companies in the industry
- OFED software enables RDMA and low latency communication across servers, networks and storage systems
- OFED software delivers increased application productivity and reduced capital expenditures and operating expenses
- OFED open source and community-oriented software repositories are used for inclusion of RDMA software in kernel, OS, networking and storage vendor distributions
- OFA interoperability testing programs ensure interoperability of different vendor products preventing vendor lock-in
- OFED software is used in nearly 50 percent of high performance computing clustering and their unified I/O storage, data warehousing and content delivery applications
- Multiple publicly available case studies in high performance computing and data center applications exemplify the value that OFED brings to its users
- The features and benefits of OFED are critical in future Exascale, Business Analytics, Web 2.0, Virtualization and Cloud Computing
- Membership in the Alliance brings many benefits, including: influencing the future of OFED software, product interoperability and certifications; access to online and onsite training curriculums; and marketing promotions at industry events

2. What’s Unique about the OFA?

The OFA is an open source software Alliance that focuses on developing, testing and licensing high performance networking software for servers and storage systems. It is unique in that its open source software – OpenFabrics Enterprise Distribution or OFED™ – powers some of the most powerful server, networks’ fabrics and storage systems in the world, in the high performance computing, financial services, data warehousing, online transaction processing and managed hosting services applications. OFED has been proven to be the most efficient and high performance server and storage clustering software, utilizing technologies such as kernel bypass and RDMA (Remote Direct Memory Access). The tremendous growth of clustering and standards-based high performance interconnects in the TOP500 (www.TOP500.org) is testament to this. A host of who’s who in the data center and high performance computing industries – such as AMD, Cisco, HP, IBM, Intel, Oracle, Microsoft, NetApp, the Tri-labs (Lawrence Livermore National Laboratory,

Los Alamos National Laboratory, Sandia National Laboratories), Mellanox, Qlogic, Chelsio and others – are its members. An active community of open source developers and commercial companies collaborate to create and release OFED for all popular Linux and Microsoft Windows server operating system platforms. OFED forms the basis of server clustering, low latency and RDMA software in operating system from Canonical, Debian, Microsoft, Novell, Oracle and Red Hat. OFED has also been ported to other operating systems and virtualization platforms such as VMware, Xen, KVM, etc.

3. Benefits at a Glance

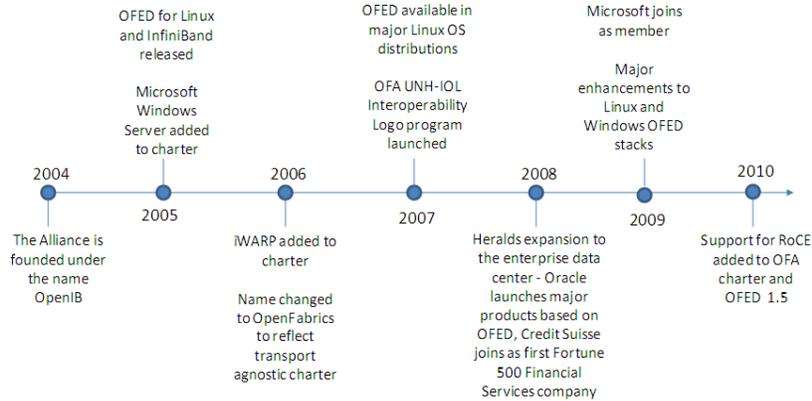
Before we dive into OFED as a technology, it is important to highlight its key benefits:

- 1. Increased productivity:** OFED helps maximize the number of job operations or transactions per minute. OFED helps researchers and developers complete simulations faster. OFED helps process more capital market data per second and make split-second and accurate trading decisions and transactions.
- 2. Increased ROI:** When OFED is used to cluster servers or storage systems, it boosts efficiency. What this means is that performance of the cluster can scale effectively as more servers and storage systems are added to the cluster. So, more transactions, more users and applications, better decisions, and more accurate results are possible with less server, storage and networking hardware. And through industry-wide collaborations driven by OFA, and transport-neutral software architecture, OFED fosters vendor interoperability, enabling multiple hardware solutions.
- 3. Energy Reduction:** Server and storage systems are the primary consumers of power in the data center. OFED helps reduce the number of server and storage systems needed to meet business and research goals, thereby reducing power consumption in the data center by a significantly higher order of magnitude.

The values provided by OFED are well aligned with evolving trends in the industry – the scale and productivity of cloud, web 2.0 infrastructures and Exascale systems. The case studies and future trends presented in this article exemplify these value propositions.

4. A Brief History of the Alliance

Research communities in the Tri-labs and the U.S. Department of Energy envisioned the need for a networking software stack for very high performance I/O for servers and storage and the need for their standardization to help proliferation and enable economies of scale. The Alliance was founded in June 2004 with these goals in mind and was initially funded by the U.S. Department of Energy. It started with the name OpenIB and the charter of unifying multiple proprietary software stacks available from InfiniBand commercial vendors into a single Linux-based open source software stack suitable for upstream Linux kernel adoption and subsequent inbox distribution in Linux-based server operating systems. In 2005, the alliance committed itself to supporting Microsoft Windows, a move that would make the alliance's software stack truly cross-platform. In 2006, the organization again expanded its charter to include support for iWARP, an RDMA transport technology over Ethernet. The alliance modified its name to OpenFabrics and its charter to be transport agnostic, focusing on delivering the most comprehensive application programming interface (API) and upper layer protocols for high performance computing and enterprise data center applications. The alliance released the first version of the OpenFabrics Software, known as OFED in 2005. It supported InfiniBand and ran on Linux. Subsequent releases have added support for iWARP and Windows. In 2010, support for RoCE (RDMA over Converged Ethernet) was added to further expand its transport agnostic charter and to support all the choices for standards-based high performance I/O. OFED-1.5 is the latest release of the OpenFabrics Software available for download from its site. OFED versions have been available with OS distributions from Red Hat, Novell, Microsoft and Oracle for many years now. The following figure shows the OFA timeline pictorially.

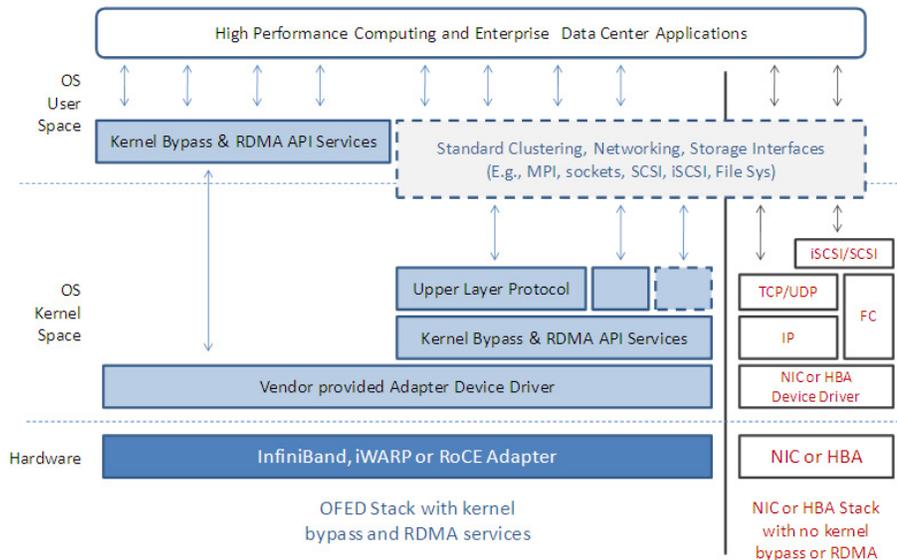


5. OFED Software Technology

This section provides a high level overview of the OFED software stack, focusing on compatibility to and use by applications to leverage the benefits mentioned above. OFED supports multiple application interfaces for computing, networking, clustering and storage applications. Comparisons with the more familiar TCP/UDP/IP based networking stack and iSCSI/SCSI/Fibre Channel storage networking stacks are provided to facilitate easier understanding.

The benefits delivered by OFED stem from its ability to complement standards-based high performance I/O adapter hardware to deliver extremely low node to node latency (nodes can be servers or storage systems), low CPU utilization (keeping CPUs free for use by applications), and a high bandwidth data path (making large data movements across nodes fast and efficient).

Architecturally, OFED implements a set of kernel bypass and remote DMA mechanisms for delivering the latency, utilization and bandwidth benefits for these data center traffic types – inter-processor, networking, clustering and storage – making it the highest performing converged I/O stack. OFED enables several industry standard RDMA, kernel bypass and transport specifications – InfiniBand, Rapid I/O, RoCE (hard and soft) and iWARP. The first uses InfiniBand Host Controller Adapters (HCAs) and switches, while the latter two operate over Ethernet Network Interface Cards (NICs) and switches. The goal of OFED is to deliver a set of APIs – both at kernel (called verbs) for maximum performance, and protocol level for excellent performance and application compatibility – in a way that is independent to the underlying interconnect and transport technology. In other words, the same OFED APIs and applications that use those APIs run seamlessly over InfiniBand-, iWARP- or RoCE-based hardware. This is shown pictorially in the following high level OFED stack diagram. It also shows on the right hand side a more familiar NIC or HBA software stack as a comparison.



Next, we look at the vendor-provided adapter device driver, kernel bypass and RDMA API services, and the Upper Layer Protocols supported by OFED in some more detail.

5.1 vendor-provided adapter device driver

This is the only OFED software layer that contains hardware dependent components. These device drivers support PCI-based adapter cards used in servers and storage systems for high performance I/O connectivity using the supported InfiniBand, iWARP, Rapid I/O and RoCE standards. Interconnect vendor device drivers currently included in OFED includes the following vendors: Chelsio, Intel, IBM, Mellanox, Mercury and QLogic.

5.2 Kernel bypass and RDMA API services

These set of API services are also referred to as verbs and they are available both as kernel space and user space components. The verbs supported by OFED are derived from the messages or programming semantics defined by the transports relevant to the standards that OFED supports, namely InfiniBand, iWARP, Rapid I/O and RoCE. Typically, the following messages or programming semantics are supported at the verbs API level:

- Direct Memory Access read from or, write to, a remote node (RDMA)
- Channel send or receive (that can also implement kernel bypass)
- Transaction-based operations
- Multicast transmissions
- Atomic operations

Note that while the verbs API set delivers the highest performance (bandwidth, latency, and CPU utilization), applications need to be ported to avail of the benefits. Many latency sensitive applications in the high performance computing and financial services sectors have been ported to run directly over the verbs interface.

5.3 Upper Layer Protocols (ULPs) supported by OFED

OFED supports a host of kernel and user level ULPs, geared toward providing a seamless interface to applications that can enjoy the benefits of OFED without requiring porting. The following are popular kernel level ULPs and relevant applications:

- **SRP (SCSI RDMA Protocol):** This storage protocol is available as both initiator and target components for OFED and is suitable for high performance block storage communications between servers and storage systems akin to Fiber Channel. The protocol provides a standard SCSI interface to applications while delivering high throughput and IOPs using underlying RDMA mechanisms in the OFED stack and vendor hardware. SRP is used as a RDMA storage protocol over InfiniBand.
- **iSER (iSCSI Extensions over RDMA):** This storage protocol is available as an initiator component within the OFED package and the target component is available from several members of the OFA. iSER is suitable for high performance block storage communications between servers and storage systems. The protocol provides a standard iSCSI interface to applications, including compatibility to iSCSI management software while delivering high throughput and IOPs using underlying RDMA mechanisms in the OFED stack and vendor hardware. iSER can be used as a RDMA storage protocol over iWARP and RoCE.
- **RDS (Reliable Datagram Service):** This low latency and high performance IPC (inter processor communication) and storage protocol is used, for example, in Oracle clustered database applications for both server to server and server to storage system communications. It implements both send, receive and RDMA semantics, with the former used for IPC and the latter used for storage applications. RDS is currently supported over InfiniBand and RoCE.

- **IPoIB (Internet Protocol over IB):** This protocol enables any IP application to operate over InfiniBand without requiring any change. It enables use of the TCP/UDP/IP sockets and protocols over InfiniBand. No RDMA or send receive semantics are supported. IPoIB can operate as both an unreliable datagram service (IPoIB-UD) or in connected mode (IPoIB-CM). IPoIB-CM supports larger MTUs to support higher effective bandwidth for IP applications running over InfiniBand.
- **SDP (Sockets Direct Protocol):** The purpose of SDP is to provide an RDMA accelerated alternative to the TCP protocol on IP. The goal is to do this in a manner which is transparent to the application. Originally designed for InfiniBand, SDP has been redefined as a transport agnostic protocol for RDMA network fabrics. SDP defines a standard wire protocol over an RDMA fabric to support stream sockets (SOCK_STREAM) network. SDP uses various RDMA network features for high-performance zero-copy data transfers. SDP is a pure wire-protocol level specification and does not go into any socket API or implementation specifics. SDP in OFED can be implemented over InfiniBand, iWARP or RoCE.
- **Lustre Parallel File System:** Lustre is the world's #1 parallel file system and is designed to enable I/O performance and scaling beyond the limits of traditional storage technology. Often used in High Performance Computing environments with InfiniBand and OFED, Lustre is also applicable to any enterprise storage environment where very high I/O bandwidth is required.
- **PureScale and GPFS:** The GPFS InfiniBand Remote Direct Memory Access (RDMA) code uses RDMA for NSD client file I/O requests. RDMA transfers data directly between the NSD client memory and the NSD server memory instead of sending and receiving the data over the TCP socket. Using RDMA improves performance, enhances bandwidth and decreases CPU utilization.

Message Oriented Middleware (MOM) plays a key role in enterprise data distribution. The strength of MOM is that it allows for communication between applications situated on heterogeneous operating systems and networks. MOM allows developers to by-pass the costly process of building explicit connections between these varied systems and networks. Advanced Message Queue Protocol (AMQP) has emerged as an open standard for MOM communication and utilizes OFED in a product called MRG from RedHat. Other messaging systems utilizing OFED are available from IBM, Tibco, Microsoft and other suppliers.

Besides the above kernel level ULPs, OFED also includes user level components such as the following. These components are transport neutral implementations (i.e., agnostic to the use of InfiniBand, iWARP or RoCE) that provide RDMA capabilities in user space. These are used for server to server messaging or IPC for both high performance computing and database applications:

- **UDAPL (User Direct Access Programming Library):** UDAPL is a specification defined by the DAT (Direct Access Transport) Collaborative (www.datcollaborative.org). It defines a single set of user APIs for all RDMA capable transports. The UDAPL mission is to define a Transport-independent and Platform-standard set of APIs that exploits RDMA capabilities in various RDMA capable interconnects. UDAPL is included with OFED and is tested with OFA supported RDMA transports and interconnects, namely InfiniBand, iWARP and RoCE.
- **MPI (Message Passing Interface):** MPI is a language-independent communications protocol used to program parallel computers. Both point-to-point and collective communication are supported. MPI is a message-passing application programmer interface, together with protocol and semantic specifications for how its features must behave in any implementation. MPI's goals are high performance, scalability, and portability. MPI remains the dominant model used in high-performance computing today. Various implementations of MPI are available in the industry. OFED includes the Ohio State University implementation of MVAPICH/MVAPICH2 (<http://mvapich.cse.ohio-state.edu/>) and the Open MPI implementation (available from www.Open-MPI.org).

OFED also includes certain management and diagnostics software components, primarily designed to help simplify or automate management and diagnostics of server clusters.

5.4 OFED-Driven Vendor Interoperability

OFA, in collaboration with the University of New Hampshire Interoperability Labs (UNH-IOL), has developed a comprehensive interoperability test program and manages events multiple times a year. Such events are used to test vendor interoperability using the latest OFED releases and test cases specific to the OFED components described above.

Next we look at how some of these components are used in end user application specific use cases.

6. Typical OFED Use Cases

OFED is used in server clustering, storage clustering and storage area networking, and networked and clustered file system applications. This section explains such specific use cases. The next section applies these use cases to specific market verticals and end user case studies and exemplifies how the OFED benefits highlighted earlier in this article are realized with real life applications.

6.1 Server Clustering

MPI and underlying OFED components such as vendor-provided adapter device driver, kernel bypass and RDMA API services are used extensively for server to server messaging in high performance applications such as academic and government research applications, weather forecasting applications, automotive and fluid mechanics design applications, computer aided design, electronic design automation, reservoir modeling, and oil and gas exploration. UDAPL, RDS, vendor-provided adapter device driver, kernel bypass and RDMA API services are used for IPC communications in clustered database applications for data warehousing and online transaction processing applications. vendor-provided adapter device driver, kernel bypass and RDMA API services are used for also used by many latency-sensitive financial services applications for scaling processing and analytics of large volumes of capital market data. SDP with vendor-provided adapter device driver, kernel bypass and RDMA API services can be used for sockets oriented application accelerations such as in application servers, distributed caching servers and communication between application servers and database servers.

6.2 Storage Clustering

Vendor-provided adapter device driver, kernel bypass and RDMA API services are used extensively for high availability and traffic mirroring, multi-pathing and other storage specific IPC traffic between clustered storage systems.

6.3 Storage Area Networking (SAN)

SRP or RDS with vendor-provided adapter device driver, kernel bypass and RDMA API services are used to replace separate Fibre Channel-based SAN services and provide consolidated IPC, networking and SAN services over InfiniBand, and delivering significantly higher storage throughput. iSER with vendor-provided adapter device driver, kernel bypass and RDMA API services are used to replace separate Fibre Channel based SAN services and provide consolidated IPC, networking and SAN services over iWARP or RoCE based 10GigE or 40GigE infrastructures, and delivering significantly higher storage throughput.

6.4 Network File Systems (NFS)

NFS over RDMA (NFS-RDMA) implementations over vendor-provided adapter device driver, kernel bypass and RDMA API service can deliver high NFS throughput performance for applications such as backup required with data warehousing and online transaction

processing applications, especially when they already use OFED-based send receive and RDMA services for IPC and block storage access.

6.5 Clustered File Systems

Use of clustered file systems is popular in high performance computing as seen by the proliferation of the use of products like Lustre and GPFS (Global Parallel File System). These products use vendor-provided adapter device driver, kernel bypass and RDMA API service to deliver high performance and efficient clustered file systems. Clustered file systems are being increasingly used in cloud storage systems and web 2.0 applications where massive scaling with commodity storage server components is critical.

Next, we discuss real end user success stories using OFED components and use cases we just discussed.

7. OFED Success Stories

Since the release of OFED in 2005 by the OpenFabrics Alliance, within a short span of five years, its use has proliferated into many industry sectors, providing distinct and tangible benefits that translate directly to the productivity and efficiency benefits in large government, defense and academia institutions, and profitability and competitiveness benefits in Fortune 500 commercial companies and institutions. The following are just a few examples of such success stories.

7.1 In Academia and Research

- **Julich:** The Julich Petascale project uses OFED-based technologies to boast the most efficiency in the Top 10 supercomputers listing by TOP500.org. It achieved 274 Tera Flops using 3000 server nodes and 26,000 CPU cores at 92% efficiency. (It is useful to know that other supercomputers that use lower performing I/O technologies that do not use OFED reach only 50-60% efficiency). Benefits: Efficiency, ROI, Greener.
- **Texas Advanced Computing Center (TACC):** The mission of TACC is to enable scientific discovery and enhance society through the application of advanced computing technologies. The TACC Ranger Cluster is one of the largest in the world with 82 server racks with about 4K server nodes, more than 62K CPU cores, 123 Terabytes of memory, 72 storage servers serving 1.7 Petabytes of disk capacity, and delivers 579 Tera Flops performance using OFED-based technologies. Benefits: Performance, Productivity, Greener, ROI.

7.2 In Weather Forecasting

- **NASA Center for Climate Simulation (NCCS):** Its 1,200-node cluster located at the Goddard Space Flight Center in Greenbelt, Maryland comprises 14,400 processors and OFED-based technologies to double NCCS computational capabilities to more than 300 trillion floating-point operations per second. At these new performance levels, NCCS users will have the ability to fine-tune global model resolutions, capturing smaller-scale features in the atmosphere and oceans so that they can better understand and predict climate change. Benefits: Productivity, ROI, Betterment of Human Lives.
- **University of Colorado at Boulder:** OFED-based technologies used with the new University of Colorado at Boulder 1,368-node cluster supports research into climate change and quick response to severe weather occurrences. The new cluster ranks #31 on the June 2010 TOP500 list of supercomputers with peak performance of 152 Tflops and an efficiency rating of 86 percent. Benefits: Productivity, ROI, Greener, Betterment of Human Lives.

7.3 In Automotive and Fluid Dynamics Design Industries

- **General Motors:** GM uses OFED-based technologies in vehicle design and simulations where vehicle development process time reduced from 42 months to 18 months. Benefits: Productivity, Profitability
- **Sikorsky:** Sikorsky uses OFED-based technologies in its flagship CH-53K program (CH-53K Super Stallion is a large, heavy-lift cargo helicopter) reducing simulations duration from 4 days to several hours. Benefits: Productivity, Profitability.

7.4 In Financial Services Industries

- **NYSE:** NYSE Technologies' Data Fabric, using industry standard hardware, offers 10x times the throughput with 1/10th the latency of traditional, low-latency IP-based middleware. This breakthrough performance is achieved by leveraging the hardware acceleration afforded by the latest generation of servers and OFED-based technologies that bypass the OS and IP stack – two bottlenecks that traditionally plague IP-based middleware offerings. Benefits: Performance, Profitability.
- **NASDAQ and Singapore Exchange (SGX):** A new trading engine – SGX Reach – delivered through NASDAQ OMX's technology and OFED-based technologies achieved a record-breaking, average order response time providing financial customers with the fastest execution capability in the world. Benefits: Performance, Profitability.

7.5 In Data Warehousing and OLTP Industries

- **Allegro Group:** Allegro is the No. 1 e-commerce company in Eastern Europe and the second-largest online auction business in the world. With its business growing 40 percent a year and its data volumes growing 60 percent, Allegro Group brought in Oracle Exadata to support its data warehouse. With Oracle Exadata (that uses OFED-based technologies), Allegro gets a single source of truth and vastly improved performance. Queries that were running in 24 hours are now running in less than 30 minutes. Benefits: Productivity, Profitability.
- **Pacific Gas and Electric Utility Services:** PG&E has been able to reduce hardware costs by five times, resulting in an annual savings of over \$5 million by re-platforming to Oracle Real Application Clusters (RAC) which utilizes OFED-based technologies to deliver very high query throughput, significantly reduced table scan times with exponentially increasing data set sizes. Benefits: ROI, Greener.

7.6 In Storage Applications

- The **Oracle Exadata II Storage System** utilizes OFED-based technologies with flash-based caching technologies to improve query throughput to 50 Gb/s which is more than four times higher than the closest competing solutions. Benefits: Productivity, ROI, Greener.
- The **IBM SONAS** (Scale Out Networked Attached Storage) is a cloud storage appliance that used OFED-based technologies to manage multiple Petabytes of storage and up to a billion files in a single file system and deliver Scale-out performance to satisfy bandwidth hungry applications. Benefits: Performance, Scaling, TCO.

7.7 In Managed Hosting or Cloud Services

- **Accenture:** Global online airline transaction processing application hosted by Accenture used OFED-based software components and OFED supported hardware to reduce provisioning time to 12 hours from 200 hours, achieved 30% I/O power savings and 70% reduced cabling costs. Benefits: Productivity, TCO, Greener.

² As demonstrated and referenced at VMworld 2009, where Intalio received a Best of VMworld award.

Oracle's Exalogic is designed to revolutionize data center standardization and consolidation, enabling enterprises to bring together tens, hundreds, or even thousands of disparate, performance-sensitive workloads with maximum reliability, availability, and security. Oracle Exalogic's unique high-bandwidth, low-latency interconnect fabric means that complex, distributed applications can run with a responsiveness simply not achievable with traditional servers used in data centers today.

- **Mitsubishi Manufacturing:** Intalio is a provider of private cloud appliance solutions and provides solutions to end users like Mitsubishi and others in the aerospace, defense, business services, communications and media, construction, energy, healthcare, financial services and other industries. Using commodity server and storage components and OFED-based technologies, the Intalio Cloud Appliance delivers more virtual machines (with required compute, I/O and memory resources) per server while reducing hardware acquisition costs by 75% and cloud management and administration costs by 50%. Benefits: Productivity, TCO, Greener.

8. Opportunities for OFED in Exascale, Web 2.0 & Cloud Deployments

Needless to say, OFA and its flagship OFED software has played a vital role in the proliferation of the use of high performance I/O in mainstream scientific engineering applications of today, delivering exemplary benefits and helping users and companies find solutions, understand the data they are dealing with and implement more efficient operations – whether they be product design, manufacturing efficiency, highly leveraged financial transactions, customer services or content generation for entertainment. This is the beginning of more widespread and significant adoption of OFED's key value-adds which are fast becoming the key architectural ingredient of current societal trends for faster data motion and shorter time to get key information, exemplified by the massive scale of web 2.0, cloud computing deployments, clusters for data analytics and scalable systems for scientific research. Clustering is the preferred technology for a growing set of web 2.0 deployments using open source Hadoop and Memcached applications. Cloud computing and virtualization requirements are changing the networking paradigm where traditional highly oversubscribed vertical traffic patterns (server to access layer to core layer to the enterprise edge) are giving way to the more dominant horizontal server-rack to server-rack traffic patterns to enable efficient scaling, accelerating the need for high bandwidth, low latency and fully subscribed networks. The result is a flattening of the network to more clustering-oriented architectures and a compelling synergy to the benefits that OFED brings. The success that OFED has seen to date is just the tip of the iceberg; the chances of its use proliferating in the largest data centers around the world is promising.

9. Alliance Workgroups

The OFA is comprised of a board of directors with a chair, co-chair, treasurer business manager and available counsel. At the heart of its operations are an executive working group and multiple other groups that perform its daily operations through a community of contributors. The workgroups are led by a chair and a co-chair. The Executive Workgroup (XWG) is delegated by the Board and the members the responsibility to run the OFA, make core decisions and bring matters for discussion to the board for approval. The Enterprise Work Group (EWG) and Windows Workgroup (WWG) form the core development teams for OFED for Linux and Windows respectively. OFA hosted servers function as the code repository and website for the code produced by these two workgroups. Code is contributed to OFA under both the GPL and BSD dual licenses. Code for Linux can be downloaded to a user or uploaded to kernel.org, freely under either GPL or BSD – for OFED for Windows only the BSD license applies distributions. The Marketing Workgroup (MWG) is responsible for member recruiting, OFA promotions, conferences and events activities, training and education, web site and other collateral development and conduct of OFA workshops, including the annual International Developers Workshop. The Interoperability Workgroup (IWG) develops interoperability test plans and manages UNH-IOL interoperability test events, test labs and results. Interoperability events at UNH-IOL are held twice a year. OFA is currently working on delivering comprehensive training curriculum to teach RDMA, send/receive, kernel bypass programming methods with hands-on examples and both web-based and onsite delivery at end user sites.

10. OpenFabrics Alliance Membership

OFA members are a worldwide community including top-tier silicon, system, server, network equipment, storage, OS vendor (OSV), and Independent Software Vendor (ISV) companies. Current OFA members include: Allston Trading, AMD, Appro, Chelsio, Cisco, Cray, DataDirect Networks, Dell, Endace, Flextronics, HP, IBM, Intel, the Tri-Labs (LANL, LLNL, Sandia), Mellanox, Microsoft, NetApp, Oracle, QLogic, SGI, System Fabric Works, Voltaire and Xsigo. These members and a larger community of developers and end users help define the charter of the Alliance to benefit the overall industry, help foster its growth, and enable industry-ready solutions in response to the growing demand of today's high-performance applications and infrastructures.

OFA's membership levels are tailored for levels of interest and involvement:

- **Promoter** – organizations and enterprises that wish to become a candidate for a seat on the Alliance's Board of Directors; strongly influence the technology directions and affect the development process, testing, release preparation and problem resolution in the OpenFabrics Enterprise Distribution (OFED); and help direct the Alliance's promotional, marketing and educational activities. Dues are US\$3,000 initiation, US\$10,000 annual membership.
- **Adopter** – organizations and enterprises that wish to contribute to, and participate in, the development process and promotional activities but do NOT feel the need to strongly affect processes, have a Board seat, or influence technology directions and promotion. Dues are US\$3,000 initiation, US\$5,000 annual membership.
- **Supporter** – organizations and enterprises that wish to use the OpenFabrics Software, leverage the Alliance's promotional activities and stay informed about the work of the Alliance, but not actively contribute. Dues are US\$3,000 initiation, US\$1,500 annual membership.
- **Academic** – educational institutions that wish to contribute to the technology directions of OFED; participate in the development process, release testing, preparation and problem resolution; and/or institute OpenFabrics education and research into their institution's curricula. Dues are US\$2,000 annual membership.
- **Individual** – individual users or developers of OFED who wish to contribute to the technology directions of OFED and participate in the development process, release testing, preparation and problem resolution. Individual member applicants will verify their independence or their organization does not wish for privacy, security or government regulation to become a member at another level of membership. Dues are US\$200 annual membership.
- **Consulting Participant** – organizations and individuals that the Alliance selects based on their technical achievements and their standing in the OpenFabrics community. Honorary participation is by invitation only and there are no dues.

11. Membership Benefits and Call to Action

Previous sections established the tangible benefits that OFA software delivers to the industry and the promise it holds for future trends in large markets. This section is focused on OFA membership and how it enables collaboration with industry thought leaders who are influencing the future of how and what OFA delivers.

OFA membership offers benefits to many audiences.

- **End users:** can enforce product interoperability requirements to help ensure there is no I/O or OEM vendor lock-in. OFA membership also provides the ability to influence OEM suppliers while also receiving access to hands-on and onsite RDMA programming training curriculums.
- **OSVs:** may fine tune their platform offerings with the latest OFED releases and influence features and packaging requirements for easier integration into their distributions.

- **OEMs and I/O device vendors:** receive access to logo programs to ensure interoperability and functionality of their products that utilizes OFED software. OEMs and device vendors also benefit from cost-effective marketing promotions and participation at industry events, as well as access to the popular OFA Web site that serves as a repository of OFED source code and presentations by industry luminaries at OFA hosted events. Membership and participation help OEMs device vendors keep tabs on the latest features and use cases, as well as influence future OFED features, roadmaps and positioning.

Collaboration with industry thought leaders and defining the direction of future high performance I/O technologies is a key ingredient of OFA membership. There are many instances of member-driven initiatives in the past, for example:

- The Tri-Labs drove single stack and open source requirements to help inclusion of OFED in Red Hat distributions
- Financial services end users influenced the need to support real time OS distributions, fostered the race to zero-latency and the need for improved documentation and training curriculum
- Data warehousing software suppliers have influenced the addition new upper layer protocols such as RDS and feature additions in UDAPL
- Managed hosting service providers have promoted the use of channel interfaces for I/O virtualization benefits and reiterated need for Windows support and
- Multiple OEMs have influenced how OFED needs to be ready for commercial use, and prioritization of features that has resulted in the numerous benefits described above in this article.

We end this article with a call to the industry to join the OpenFabrics Alliance, enjoy the numerous benefits of membership and be part of the even greater success story to come!

