



Gepaart mit einer parallelen Architektur, schnellen Kommunikationsverfahren sowie skalierbaren Renderkernels eignet sich die XtreamView Render Engine für die interaktive Visualisierung sehr großer Datensätze, wie sie beispielsweise in der Seismik, der Medizin oder auch in der Film- und Computerspieleindustrie anfallen. Das Besondere des HPC-Visualisierungsansatzes ist die Verwendung von CPUs ohne die Einbindung von Grafikkarten.

COMPETENCE CENTER HIGH PERFORMANCE COMPUTING

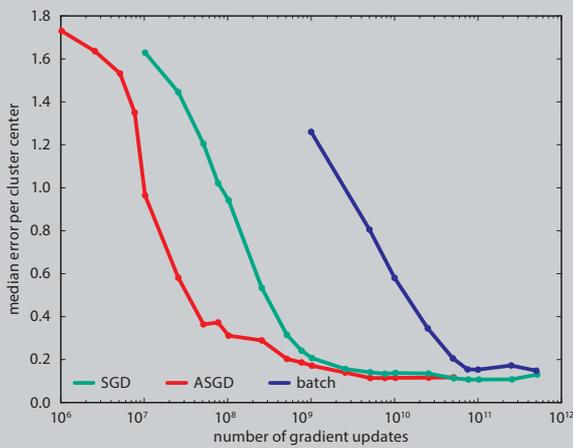
- **BIG DATA – BeeGFS, GPI, GPI-Space**
- **HPC TOOLS**
- **SEISMIC IMAGING**
- **VISUALISIERUNG GROSSER DATENMENGEN**
- **PERFORMANCE-OPTIMIERUNG**
- **E-ENERGIE, SMART GRIDS**



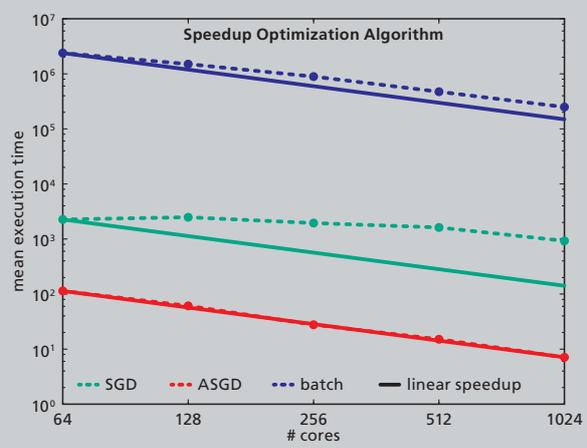


Hoch- und Höchstleistungsrechnen – High Performance Computing (HPC) – ist für die Wettbewerbsfähigkeit von Wissenschaft und Wirtschaft unerlässlich. Ohne detaillierte Simulationen ist moderne Grundlagenforschung in der Energieforschung, den Material- und Lebenswissenschaften oder auch der Klimaforschung undenkbar. Das gilt auch für Schlüsselbereiche der deutschen Wirtschaft. Ob elektronische Geräte, Autos, Flugzeuge, moderne Medikamente oder neuartige Operationsverfahren – sie alle basieren auf Erkenntnissen aus Simulationen. High Performance Computing erschließt neue Anwendungsfelder im Bereich der Simulation komplexer sozialer Phänomene oder anspruchsvoller Aufgaben in der Logistik. Gleichzeitig steigt die Komplexität der zugrundeliegenden HPC-Architekturen und ihre effiziente Nutzung stellt industrielle Anwender vor große Herausforderungen: Programme müssen heute massiv parallel arbeiten und dabei sowohl lokale Speicher berücksichtigen als auch Speicher, auf deren Inhalte nur über Netzwerke zugegriffen werden kann. Die sich zusätzlich eröffnenden Möglichkeiten durch die Verwendung spezialisierter Hardware, wie Grafikprozessoren, erhöhen die Komplexität der Programmierung weiter.

Das Competence Center High Performance Computing beschäftigt sich in enger Zusammenarbeit mit industriellen und akademischen Partnern mit der Frage, wie die immer komplexer werdenden Prozessoren und Parallelrechner effizient genutzt werden können und stellt neben Werkzeugen zum Umgang mit Supercomputern auch komplette Softwarelösungen her. Bei den Werkzeugen ist das Global Address Space Programming Interface (GPI) mittlerweile ein europäischer Standard, zu nennen. Es erlaubt die Programmierung skalierbarer eng gekoppelter Software, also Software, die einerseits mit hoher Frequenz kleine Datenpakete austauscht und andererseits mit zunehmender Anzahl der Rechner proportional schneller wird. Setzt GPI explizite Formulierung der Kommunikationsmuster durch die Anwendung voraus, geht das Software-System **GPI-Space** weiter und automatisiert sowohl Parallelisierung und Fehlertoleranz als auch Speichermanagement und wird deshalb vorwiegend für Big Data-Aufgaben genutzt. Diese Werkzeuge und auch das am CC HPC entwickelte weltweit führende parallele Dateisystem **BeeGFS** werden erfolgreich in komplexen kundenspezifischen Anwendungen eingesetzt. Last but not least arbeitet das CC HPC am Management der Energiewende mit ihrer fluktuierenden Produktion erneuerbarer Energien. Die zeitliche Entkopplung von Energieerzeugung und -verbrauch, Verbrauchsprognose und -verlagerung, Optimierung des Eigenverbrauches sowie netzdienliches Management verteilter Batteriesysteme werden in allen Facetten behandelt. Umfangreiches Wissen über die Konstruktion und die Steuerung komplexer IT-Systeme fließt ein in die Verfolgung des Ziels einer sicheren, ökologischen und wirtschaftlichen Energieversorgung und Green by IT ist zu einem wichtigen Geschäftsfeld geworden.



1



2

LARGE SCALE MACHINE LEARNING

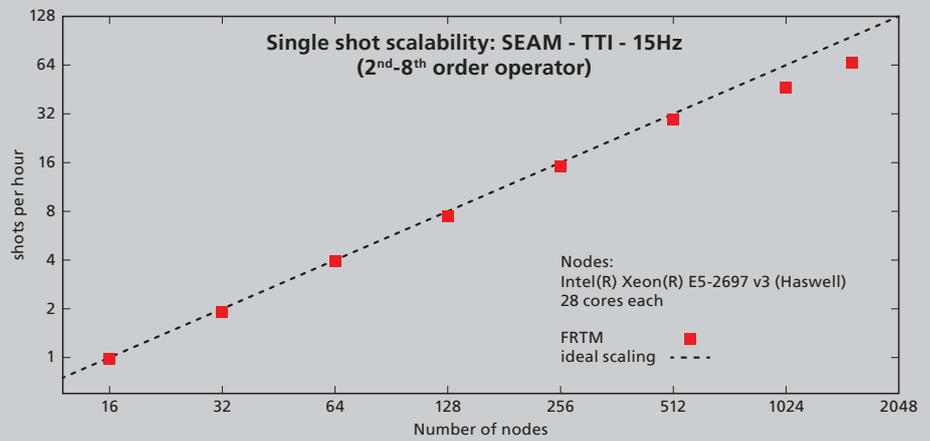
Im Bereich des Maschinellen Lernens (Machine Learning) wurden in den letzten Jahren beeindruckende Fortschritte erzielt. Von der Spracherkennung über die automatische Bildauswertung bis hin zum Erlernen komplexer Tätigkeiten wie z. B. dem Beherrschen von Videospielen: Algorithmen erlangen zunehmend Fähigkeiten, die bis vor kurzem nur von Menschen bewältigt werden konnten. Diese rasante Entwicklung geht einher mit einer stetig steigenden Komplexität der zugrunde liegenden Modelle, welche sich direkt in einem großen Bedarf an Rechenleistung und Speicher niederschlägt. So entwickelt sich Machine Learning immer mehr zu einer neuen Anwendung für das High Performance Computing. Der Focus der Arbeiten am CC HPC liegt dabei auf der Entwicklung skalierbarer Optimierungsalgorithmen für die verteilte Parallelisierung von sehr großen Machine-Learning-Problemen. Grundlage für diese Arbeiten sind dabei die am CC HPC entwickelten HPC-Komponenten, wie z. B. das parallele Filesystem **BeeGSFS** oder das Programmierframework **GPI2.0**, welche die effiziente Implementierung neuer Algorithmen wie **ASGD** (Asynchronous Stochastic Gradient Descent) erst ermöglichen.

Die Implementierungen der meisten Algorithmen im Bereich des Machine Learning lösen im Kern ein numerisches Optimierungsproblem. In diesem Zusammenhang haben sich stochastische Gradientenabstiegs-Methoden (SGD) seit langem bewährt. Sie liefern gute Ergebnisse, sowohl im Hinblick auf die Konvergenz als auch bezüglich der resultierenden Genauigkeit. In jüngster Zeit wurden mehrere Ansätze zur Parallelisierung publiziert, die versuchen, SGD für den Einsatz bei der Lösung sehr großer Machine-Learning-Probleme zu skalieren. Dabei basieren die meisten dieser Ansätze auf dem MapReduce-Schema. Mit ASGD haben wir einen neuen Algorithmus vorgestellt, der in der Lage ist, SGD effizient auf verteilten Systemen zu parallelisieren. Dabei nutzten wir die Vorteile asynchroner Kommunikation. Im Vergleich mit bestehenden Verfahren bietet ASGD schnellere Konvergenz, bei linearer Skalierbarkeit und stabiler Genauigkeit. Mittels ASGD konnten bereits mehrere Machine-Learning-Algorithmen so implementiert werden, dass sie bis auf weit über tausend Rechenknoten parallel ausgeführt werden konnten.

Verwendung von ASGD zur Berechnung des K-Means-Algorithmus auf einem verteilten Speichersystem (HPC-Cluster). Vergleich zu MapReduce Batch-Optimierung und dem bisher besten Verfahren zur SGD-Parallelisierung. Testdaten: $\sim 1\text{ TB}$, $k = 10$, $d = 10$

1 Konvergenzeigenschaften

2 Skalierung vom ASGD im K-Means-Experiment (starke Skalierung) auf bis zu 1024 CPUs



1

HIGH PERFORMANCE COMPUTING MADE IN KAISERSLAUTERN FÜR EUROPA

1 Skalierbarkeit auf dem SuperMUC-Cluster

Das CC HPC engagiert sich seit einigen Jahren in Projekten der europäischen Förderprogramme. Während die ersten Projekte 2016 enden, fing die zweite Charge an Projekten im Spätherbst 2015 an. Also ist es eine gute Zeit Zwischenbilanz zu ziehen und einen Ausblick auf spannende neue Forschungsperspektiven zu geben.

Alle Forschungsdisziplinen benötigen heutzutage immer höhere Rechenleistung und die Fähigkeit, große Datenmengen zu analysieren. Daher ist ein übergeordnetes Ziel des Forschungsprogramms Horizon2020, bis zum Jahr 2020 Computersysteme zu bauen, die im Exa-Bereich funktionieren, d. h. 10^{18} Rechenoperationen pro Sekunde ausführen. Die ungeheure Rechenleistung wird durch die Kombination vieler einzelner Komponenten ermöglicht. So kann man bei heutiger Rechenleistung der Einzelkomponenten ausrechnen, dass etwa zwei Millionen Rechenkerne benötigt werden, um in den Exa-Bereich vorzudringen. Die Hardwarearchitektur hat enorme Auswirkungen auf die benötigte Software. Es ist essentiell, dass die Programme ihre Aufgaben nebenläufig abarbeiten. Zwischenergebnisse müssen zwischen den einzelnen Komponenten ausgetauscht werden, um den Fortschritt des gesamten Programmes zu gewährleisten. Generell gilt: Nutzt man mehr Rechenkerne für dieselbe Problemgröße (starke Skalierung), sinkt die Rechenzeit für jeden Knoten, während die Kommunikationskosten im besten Fall konstant bleiben. Bei sehr großen Systemen wird also die Zeit, die der Computer mit der Kommunikation verbringt, entscheidend. Es ist notwendig, durchdachte Konzepte zur Kommunikation der Daten zu verwenden. Das am Fraunhofer ITWM entwickelte Kommunikationsmodell GPI (Global Address Space Programming Interface) erlaubt eine asynchrone, multi-threaded Kommunikation unter Vermeidung von Zwischenkopien der Daten. GPI, das sich an der Speerspitze moderner Kommunikationsmodelle der HPC befindet, wird in zwei ausgereiften EU-Projekten (EXA2CT und EPIGRAM) verwendet.

In EPIGRAM werden die Grenzen des Kommunikationsmodells von GPI getestet, Erweiterungen und Veränderungen vorgeschlagen und die Anwendbarkeit auf verschiedenartige, heterogene Speichermodelle untersucht. Die Kompatibilität von GPI mit anderen auf dem HPC-Markt verfügbaren Kommunikationsmodellen ist erfolgreich getestet worden. Die Idee der Kompatibilität und Vereinbarkeit verschiedener HPC-Programmiermodelle wird in einem gerade gestarteten EU-Projekt mit Namen INTERTWINE noch ausführlicher beleuchtet werden. Die Skalierbarkeit von GPI ist im Rahmen des EPIGRAM-Projektes auf der neuesten Erweiterung des größten deutschen Supercomputers, dem SuperMUC am Leibnizrechenzentrum in München, getestet worden. Im Jahr 2015 verfügte der SuperMUC über eine Rechenleistung von etwa $2,5 \cdot 10^{15}$ Rechenoperationen pro Sekunde, ist also noch den Faktor 250 kleiner als ein Exa-Computersystem.

Es konnte gezeigt werden, dass GPI über drei Größenordnungen hinweg stark skaliert und Anwendungen über 90 % paralleler Effizienz erreichen. Bei den Tests wurden bis zu $4 \cdot 10^{14}$ Rechenoperationen pro Sekunde auf 84 000 Rechenkernen erreicht.

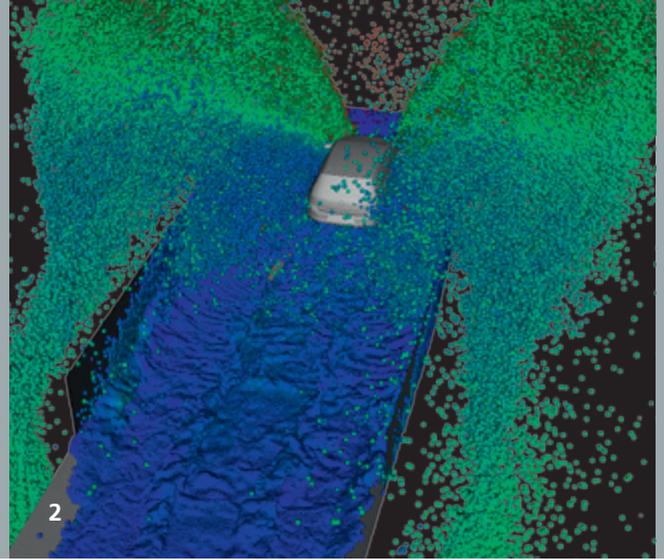
Schon beim Computercluster SuperMUC zeigt sich, dass es häufige Ausfälle bei den Hardwarekomponenten des Computersystems gibt. Die Anzahl der Hardwareausfälle wird mit der Anzahl der Hardwarekomponenten steigen. Damit ein Ausfall einer Komponente des Systems nicht sofort Auswirkungen auf die massiv parallelen Programme hat, muss die Hardware und Software fehlertolerant sein. Das muss auf allen Ebenen gewährleistet sein, also auch auf der Ebene der Kommunikationssoftware. Diese Herausforderung ist in dem europäischen Projekt EXA2CT angegangen worden. EXA2CT bringt Experten aus der Entwicklung von Lösern numerischer Gleichungen und HPC-Softwareentwickler an einen Tisch mit dem Ziel, nutzerfreundliche Kommunikationssoftware zu entwickeln. Wir haben im Projekt eine Bibliothek entwickelt, die die Fehlertoleranz von GPI noch weiter ausbaut und es Anwendungen erlaubt, Daten während der Rechnung in dem Speicher eines festgelegten Nachbarrechners abzulegen. Sollte es zu einem Ausfall eines Rechners kommen, können die Daten, die so verteilt worden sind, zurückkopiert werden und das gesamte Programm von einem kontrollierten Punkt wieder gestartet werden. Diese Methode ist deutlich schneller als das Kopieren der Daten von Festplattenspeichern.

Über die vorgestellten stark Software-orientierten Lösungen hinaus engagieren wir uns in drei europäischen HPC-Projekten, die spezielle Computerarchitekturen für den Exa-Bereich entwickeln: DEEP-ER, EXANODE und EXANEST. Für das Projekt DEEP-ER, das eine Cluster-Booster-Architektur vorschlägt, haben wir unser paralleles Dateisystem BeeGFS erweitert, um die verschiedenen Ebenen des Speichersystems effizient zu nutzen. Die gerade neu gestarteten Projekte EXANODE und EXANEST werden energiesparende Prozessoren und Nanotechnologie verwenden und ein systemweites vereinheitlichtes Speicherkonzept nutzen.

Viele der vorgestellten Ideen werden zum ersten Mal genutzt und es ist interessant zu sehen, wie sie die Forschungsrichtungen des High Performance Computing beeinflussen. Um unser umfangreiches Wissen in die europäische strategische Forschungsagenda einfließen zu lassen, arbeiten wir im Projekt EXDCI mit führenden Experten aus dem HPC-Bereich zusammen und tragen essentiell zu Entscheidungen für das wissenschaftliche Programm der EU bei. Es stehen spannende Zeiten bei dem Wettrennen zum ersten Exa-Computercluster bevor.



1



2

WELTKLASSE-3D AUS KAISERSLAUTERN

1 Diffuses Path Tracing auf dem Modell einer Boeing 777 mit ca. 300 Mio. Dreiecken (Detailansicht Fahrwerk); der Aufbau der BVH dauerte nur 15 Sekunden.

2 Die schnelle Aktualisierung der BVHs ermöglicht auch eine interaktive Visualisierung von vorberechneten Partikelsimulationen mit über 1 Mio. Partikel.

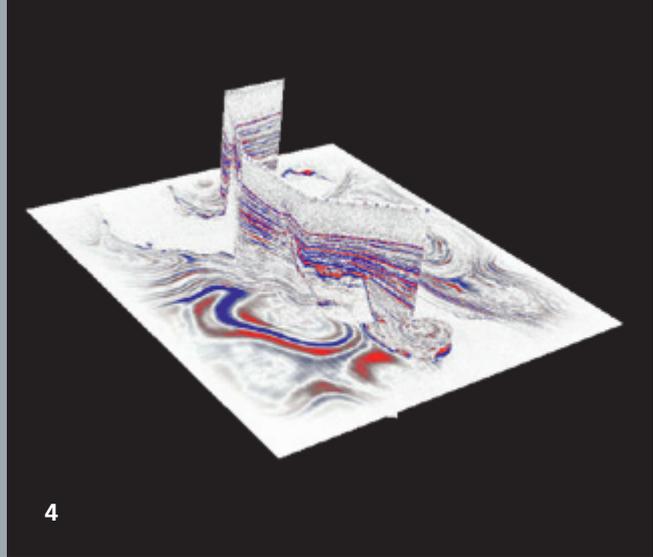
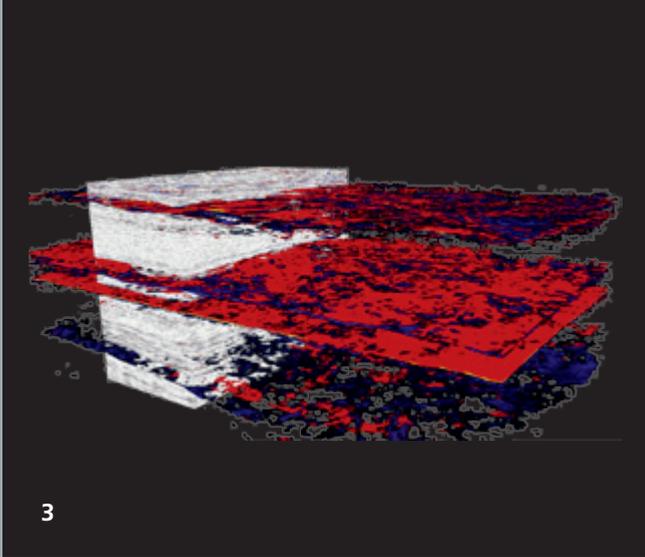
Bereits seit vielen Jahren beschäftigt sich das Competence Center High Performance Computing mit dem Thema 3D-Visualisierung. Dabei bestritt das CC HPC von Anfang an einen neuen Weg, durch die konsequente Verwendung von CPUs ohne die Einbindung von Grafikkarten. Gepaart mit einer parallelen Architektur, schnellen Kommunikationsverfahren sowie skalierbaren Renderkernels eignet sich die PV-4D Render Engine damit insbesondere für die interaktive Visualisierung sehr großer Datensätze, wie sie beispielsweise in der Seismik, der Medizin oder auch in der Film- und Computerspieleindustrie anfallen. Die Einsatzmöglichkeiten und -fähigkeiten der Engine sind dabei sehr umfassend. Angefangen bei der Visualisierung von Volumendaten mit oder ohne Volume Rendering über die effiziente Darstellung triangulierter Objekte und Szenen bis hin zur fotorealistischen Szenen, PV-4D ermöglicht gänzlich neue Einsatzszenarien.

Um ihren Spitzenplatz zu behaupten, entwickeln die Wissenschaftler in der Visualisierungsgruppe des CC HPC weitere Optimierungen und neue Verfahren. Dabei macht es die stetig wachsende Rechenleistung – auch dank neuer Hardwarearchitekturen wie Intels KNL – möglich, immer aufwändigere Visualisierungsmethoden für sehr große Datensätze in Echtzeit zur Verfügung zu stellen.

Fotorealismus mit Path Tracing

Nachdem PV-4D bereits fotorealistisches Ray Tracing in Echtzeit anbietet, folgt mit der Einbindung von Path Tracing nun der nächste Schritt. Im Gegensatz zu konventionellem Ray Tracing werden beim Path Tracing auch diffuse Oberflächen korrekt beleuchtet und so wesentlich realistischere Beleuchtungseffekte erzeugt. Um ein möglichst rauschfreies Bild zu bekommen, müssen dabei verhältnismäßig viele Strahlen pro Pixel verfolgt und gegen die Objekte in der Szene getestet werden. Entscheidend dabei sind zum einen schnelle Verfahren für den Aufbau sogenannter Bounding-Volume-Hierarchien (BVH) und zum anderen für den Test von Strahlen gegen diese BVH. Für beides werden am CC HPC gerade neue und effiziente Verfahren entwickelt. Dabei sind die Verfahren für die Strahltests – und zwar sowohl für kohärente als auch für inkohärente Strahlen – gemessen an anderen Lösungen wie z. B. Intel Embree bereits um einen Faktor 2 bis 4 schneller. Zielsetzung ist das interaktive Rendering ganzer Filme, vorausgesetzt natürlich, genügend Rechenleistung steht zur Verfügung.

Die schnellen Verfahren zum Aufbau und zur Aktualisierung der BVH ermöglichen auch neue Wege in der Visualisierung von Partikel-Simulationen. Bisher wurden diese oft Bild für Bild berechnet und anschließend als Videoclip wiedergegeben. Mit Fraunhofer-Technologie ist nun eine interaktive Visualisierung solcher Szenen möglich; mit völlig neuen Einsichten.



Neben den Anstrengungen, den technologischen Spitzenplatz in der 3D-Visualisierung zu behaupten, schlug das CC HPC in diesem Jahr auch den Weg zur stärkeren Kommerzialisierung dieser Technologie ein. Ganz im Sinne des Fraunhofer-Konzepts »Intrapreneurship« wurde mit **XtreemView** ein Startup im Unternehmen angestoßen, um diese Kommerzialisierung voranzutreiben. Gefördert wird die Kommerzialisierung – als einziges gefördertes Softwareprojekt – im Programm »Fraunhofer-Innovator«, das Technologieprojekte systematisch am Übergang vom Forschungsergebnis zum marktreifen Produkt unterstützt mit dem Ziel, die Technologien in Spin-offs oder durch Lizenzverträge am Markt verfügbar zu machen. Nicht zuletzt wegen der langjährigen Präsenz auf diesem Markt richtet sich die erste Kommerzialisierungsphase auf den Öl- und Gasmarkt und bietet – neben der PV-4D Engine – auch einen einfach zu bedienenden, schnellen, parallelen Viewer für seismische Daten: XtreemView.

XtreemView bringt Spitzentechnologie zum Endnutzer

XtreemView nutzt die Stärken der PV-4D Engine und stellt sie direkt dem Endnutzer zur Verfügung – egal, ob dieser täglich und nur gelegentlich mit der Visualisierung seismischer Daten zu tun hat. Das Programm stellt die echten 32bit Float-Werte auf dem Bildschirm dar und erlaubt das einfache Überblenden zweier beliebiger Datensätze, z. B. von Seismik und einem Geschwindigkeitsfeld. Dieses Extra an Informationen liefert einen entscheidenden Mehrwert bei der Interpretation und Analyse der Daten. Unterschiedliche Darstellungsformen, gepaart mit Volume Rendering, die Einbeziehung sehr großer und zahlreicher seismischer Horizonte sowie die Möglichkeit, beliebige Schnittflächen zu definieren, runden die Möglichkeiten für den Nutzer ab.

Der entscheidende Vorteil gegenüber anderen Viewern ist aber die einfache Skalierbarkeit, die es ermöglicht, die Hardware an die Problemgröße anzupassen. Und nicht nur lassen sich mit zwei Rechenknoten doppelt so große Daten visualisieren, durch die parallele Implementierung sämtlicher I/O Operationen sind auch die Ladezeiten nicht länger eine große Hürde: Je mehr Knoten laden, desto schneller ist der Ladevorgang. Geladen wird dabei von Datenformaten, die in der Industrie bereits seit Jahrzehnten verbreitet sind: SEG-Y, JavaSeis, SU u. a.

Vorgestellt wurde XtreemView bereits auf den beiden großen jährlichen Messen der Öl- und Gasindustrie, der EAGE in Madrid sowie der SEG in New Orleans, wo der Viewer großen Anklang beim Publikum fand.

3 *Zwei seismische Datensätze mit zusammen ~ 120 GB, einer als solider Körper, der andere mit Volume Rendering dargestellt.*

4 *XtreemView ermöglicht es auch, beliebige Pfade in den Daten zu selektieren und die Volumendaten auf den so entstandenen Ebenen darzustellen.*