

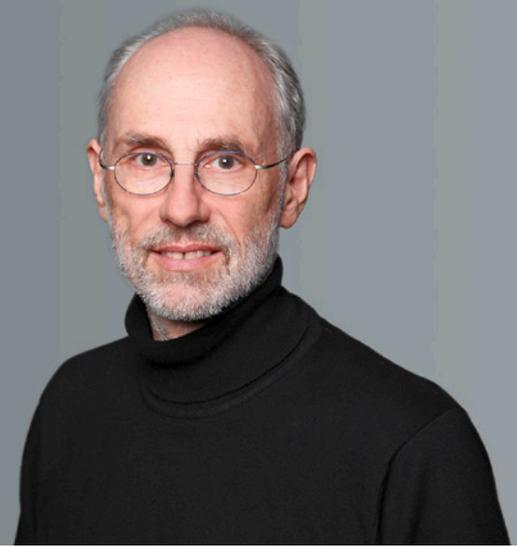


Tina Hill, Frauke Santacruz, Dr.-Ing. Janis Keuper, Dr. Abel Amirbekyan, Clemens Koch, Sven Breuner, Delger Lhamsuren, Dr. Mirko Rahn, Matthias Deller, Dr. Alexander Klauer, Christian Mohrbacher, Dr. Rui Mário da Silva Machado, Dr. Franz-Josef Pfreundt, Bernd Lietzow, Frank Kautz, Matthias Klein, Dr. Matthias Balzer, Bernd Lörwald, Dr. Daniel Grünwald, Ely Wagner Aguiar de Oliveira, Valentin Fütterling, Lukas Ristau, Dr. Martin Kühn, Dr. Dimitar Stoyanov, Dr. Tiberiu Rotaru

COMPETENCE CENTER HIGH PERFORMANCE COMPUTING

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

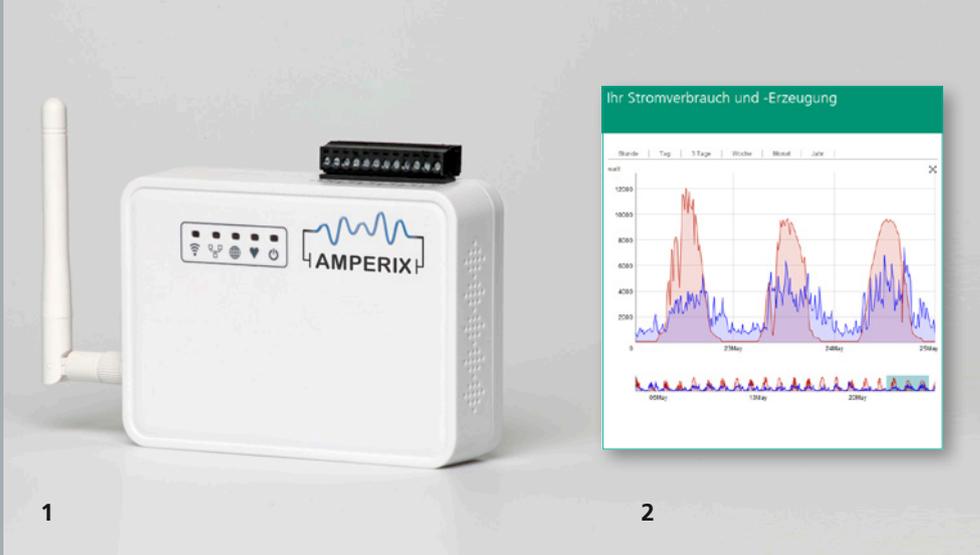




Hoch- und Höchstleistungsrechnen – High Performance Computing (HPC) – ist für die Wettbewerbsfähigkeit von Wissenschaft und Wirtschaft unerlässlich. Ohne detaillierte Simulationen sind moderne Grundlagenforschung in der Energieforschung, den Material- und Lebenswissenschaften oder auch der Klimaforschung 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 Anwenderinnen und 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) zu nennen, das die Programmierung skalierbarer eng gekoppelter Software erlaubt, also Software, die einerseits mit hoher Frequenz kleine Datenpakete austauscht und andererseits ein Problem bei Bereitstellung von mehr Ressourcen tatsächlich schneller löst. Setzt GPI explizite Formulierung der Kommunikationsmuster durch die Anwendung voraus, geht das Werkzeug GPI-Space weiter und automatisiert sowohl Parallelisierung und Fehlertoleranz als auch Speichermanagement. 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. 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.



SMART METER: BESSERES RAUMKLIMA, NIEDRIGERER STROMVERBRAUCH

Das EEG 2014 sieht vor, bis 2050 den Anteil des aus erneuerbaren Energien erzeugten Stroms am Bruttostromverbrauch kosteneffizient auf mindestens 80 Prozent auszubauen. Gleichzeitig steigen Energiekosten stetig. Durch die Projekte mySmartGrid und EMOS möchten wir diesem Trend entgegensteuern und Möglichkeiten zur individuellen Kosteneinsparung aufzeigen.

In mySmartGrid wurden intelligente Mess-, Regel- und Steuersysteme entwickelt, um den Energieverbrauch der Energieerzeugung anzupassen. Etwa 300 Testhaushalte wurden mit entsprechender Messhardware ausgestattet, die den Stromverbrauch erfasst. Über einen Internetanschluss werden Messwerte an die zentrale mySmartGrid-Website übertragen. Teilnehmer haben über einen Account Einblick in Verbrauchs- bzw. Erzeugungsdaten ihrer PV-Anlage, welche auf der Website in Form von Graphiken dargestellt werden. Die Teilnehmer können ihren Stromverbrauch analysieren und herausfinden, wo sich »Stromfresser« befinden. Ein weiterer Bestandteil der Website ist das Forum, eine Austauschplattform für die Teilnehmer untereinander sowie mit dem Projektteam. Ideen werden diskutiert, Probleme erörtert und gemeinsam gelöst. Teilnehmer, die über eine PV-Anlage verfügen, können durch die Installation einer wechselrichterunabhängigen PV-Anlagenüberwachung den Eigenverbrauch über individuelle Ertragsprognosen erhöhen. Eine Kernkomponente des Projektes ist das HexaBus-System: Um Haushaltsgeräte automatisch zu steuern, muss eine Kommunikation mit den Geräten umgesetzt werden. Der HexaBus als IPv6-basiertes Funksystem kann noch mehr: Neue Geräte können das System nutzen, um beliebige Zusatzfunktionen umzusetzen. Eine Waschmaschine könnte also warten, bis der Teilnehmer von unterwegs das Startzeichen gibt.

Aufbauend auf mySmartGrid untersucht das Projekt Energiemanagement für Mietwohnungen mit Open-Source Smart Metern (EMOS) seit 2013, wie ein optimales Raumklima erreicht werden kann. Ein optimales Raumklima spart Energie und damit klimaschädliches CO₂. Dies dient nicht nur der Umwelt, sondern trägt auch zum eigenen Wohlbefinden bei. Folglich kann durch effizientes Heizen und Lüften das Wohlfühl-Wohnklima energiesparend gesteuert werden. Basierend auf unserem Heimautomatisierungssystem HexaBus wurden Geräte entwickelt, die Temperatur und Luftfeuchte messen und Handlungsempfehlungen für Bewohner aufzeigen. Die aufgezeichneten Werte sind für Teilnehmer auf der mySmartGrid-Website durch visualisierte Darstellungen vergleichbar. Die Rückmeldung über Stromverbrauch und Raumklima der eigenen Wohnung hilft dabei, den Eigenverbrauch und damit Heiz- und Stromkosten zu senken.

Alle Ergebnisse dieser Projekte sind frei zugänglich und basieren auf Open-Source Komponenten.

1 Der Amperix – ein flexibles internetfähiges Energiemessgerät

2 mySmartGrid-Website: Grafische Darstellungen unterstützen das Verständnis des Haushaltsstromverbrauchs; Benutzer können in die Grafik hineinzoomen.

3 Mobile mySmartGrid-Website: Ansicht mit intuitiven Icons beim Zugriff von Mobiltelefonen oder Tablets

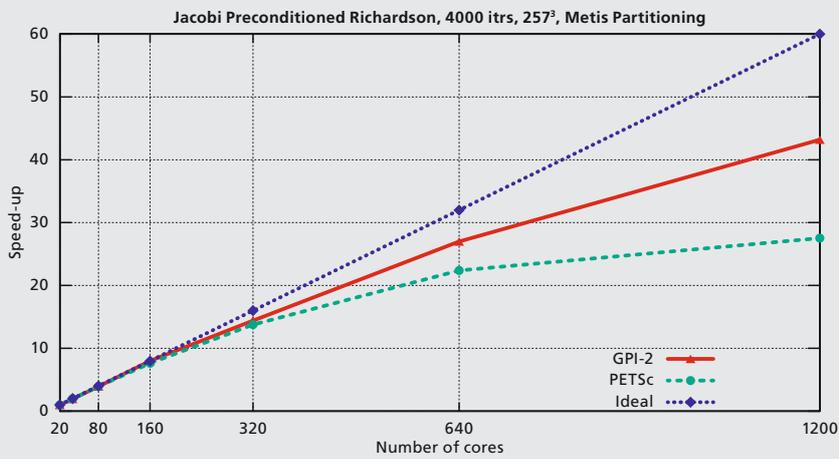


1 *Das Petascale-System
SuperMUC am Leibniz-
Rechenzentrum in München*

DEEP, EPIGRAMM UND ECA2CT: EU-PROJEKTE

Die Europäische Technologieplattform ETP4HPC hat die strategische Bedeutung des Hochleistungsrechnens als entscheidenden Vorteil für das europäische Innovationsvermögen erkannt und eine Agenda erarbeitet, um die europäische High Performance Computing-Technologie an die Weltspitze zu bringen. Das Fraunhofer ITWM ist Mitglied der ETP4HPC Plattform und die Abteilung HPC hat wichtige Beiträge in Bezug auf die Kommunikation von massiv parallelen Programmen und hochskalierbaren I/O-Systemen geleistet. Zusammen mit der Europäischen Kommission wird die von ETP4HPC vorgeschlagene Strategie im derzeitigen Forschungsrahmenprogramm Horizon2020 umgesetzt. Das primäre Ziel ist es, bis zum Jahr 2020 Computersysteme zu bauen, die im Exa-Bereich rechnen, d. h. 10^{18} Rechenoperationen pro Sekunde ausführen. Diese neuen Systeme stellen uns vor zahlreiche Herausforderungen. Zum Beispiel muss der Energieverbrauch der Rechnerkomponenten erheblich gesenkt werden, um einen Computer mit einer Leistungsfähigkeit im Exa-Bereich wirtschaftlich zu betreiben. Das erfordert radikale Innovationen bei der Architektur zukünftiger Supercomputer und den Einsatz von Computerelementen mit geringem Stromverbrauch. Auch die Entwicklung skalierbarer Software für die zukünftigen Systeme mit Millionen von Rechenkernen erfordert einen Paradigmenwechsel bei Systemsoftware und Programmiermodellen.

Im dem europäischen Projekt DEEP als auch dem Nachfolgeprojekt DEEP-ER wird eine solche Architektur entworfen und prototypisch umgesetzt. Es handelt sich um eine sogenannte Cluster-Booster-Architektur, bei der sich das Gesamtsystem aus einem Cluster von Rechenknoten mit vielen Rechenkernen und einem Booster zusammensetzt. Diese Kombination bietet maximalen Datenfluss und Skalierbarkeit des Boosters. Der Booster-Teil des Systems kann für den Teil der Programme genutzt werden, die hochskalierbar sind, während der traditionelle Cluster für die Programme genutzt werden kann, die den Datenfluss kontrollieren und Abhängigkeiten zu anderen Daten haben. Im Projekt DEEP hat sich herausgestellt, dass die wachsende Kluft zwischen I/O-Bandbreite und Rechnergeschwindigkeit eine relevante Herausforderung für die DEEP-Architektur ist, die in DEEP-ER angegangen wird. DEEP-ER erweitert die Computerarchitektur basierend auf vielfältigen Rechnerkomponenten durch ein hochskalierbares I/O-System. DEEP-ER wird einen Prototypen mit dem Intel Xeon Phi Prozessor, einer einheitlichen Hochgeschwindigkeitsvernetzung der Computerknoten, Permanentenspeicher auf den Computerknoten (NVM) und an das Netzwerk angeschlossenen Speicher (NAM) bereitstellen, um Hochgeschwindigkeitszugriffe auf den Speicher zu unterstützen. Im Projekt passt die Abteilung HPC das hochskalierbare, effiziente parallele I/O-System basierend auf dem parallelen Dateisystem BeeGFS an die DEEP-ER-Architektur an. Erweiterungen des Posix I/O-Standards werden es den Anwendungen erlauben, die verschiedenen Ebenen des Speichersystems effizient zu nutzen.



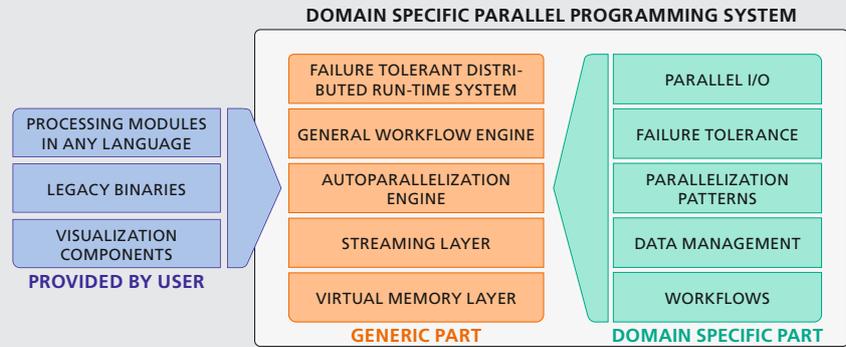
2

Generell gilt: Nutzt man mehr Kerne für dieselbe Problemgröße (starke Skalierung), sinkt die Rechenzeit für jeden Knoten, während die Kommunikationskosten konstant bleiben oder ansteigen. Das bedeutet, dass die Zeit bis zur Lösung des Problems von den Kommunikationskosten dominiert wird. Diese Entwicklung begünstigt asynchrone, nicht-blockierende Kommunikationsmodelle, die die Kommunikation an dem frühestmöglichen Zeitpunkt während der Ausführung des Programmes starten, so dass die Daten rechtzeitig empfangen werden, wenn sie benötigt werden. Außerdem erfordert die steigende Zahl von Computerkomponenten mehr Nebenläufigkeit und die Fähigkeit, granular zu kommunizieren. Das CC HPC stellt solch ein System bereit: GPI. Das Global Address Space Programming Interface (GPI) ist mit Blick auf die Skalierbarkeit paralleler Anwendungen entwickelt worden. GPI erlaubt eine effiziente, multi-threaded Kommunikation mit geringer Wartezeit und hoher Bandbreite, ohne Belastung der Prozessoren und unter Vermeidung von Zwischenkopien der Daten. GPI wird als Programmiermodell in zwei EU Projekten, und zwar EPIGRAM und EXA2CT, verwendet.

2 Skalierung einer Matrix-Vektor-Basisoperation im Vergleich

Das Ziel des Projektes EPIGRAM ist es, Programmiermodelle für Computer im Exa-Bereich vorzubereiten, indem die derzeitigen Beschränkungen grundlegend ausgewertet werden. EPIGRAM führt neue fortschrittliche Konzepte ein, um die technologische Lücke zu den Programmiermodellen im Exa-Bereich zu füllen. EPIGRAM analysiert die Philosophien verschiedener Programmiermodelle und versucht die besten Eigenschaften der Modelle miteinander zu kombinieren. Ein wichtiger Punkt ist die Interoperabilität zwischen verschiedenen Kommunikationsmodellen. GPI wird in den zwei Anwendungen des EPIGRAM Projektes verwendet: NEK5000, ein Strömungsdynamikprogramm (CFD), und IPIC3D, ein Programm zur Simulation von Weltraum und Fusionsplasmen.

Das Projekt EXA2CT bringt Experten aus der Entwicklung von Lösern numerischer Gleichungen und HPC-Softwarearchitekten von Programmiermodellen an einen Tisch. Die Weiterentwicklung von Simulationen als Wissenszweig beruht zunehmend auf rechenintensiven Modellen, die mehr Computerressourcen benötigen. Da die Leistungsfähigkeit eines einzelnen Prozessorkerns nur bis zu einer gewissen Grenze erhöht werden kann, werden Systeme im Exa-Bereich auf stark parallelisierten Programmen beruhen. Im Rahmen des Projektes unterstützt das CC HPC die Partner bei der Nutzung des GPI-Programmiermodells und entwickelt es weiter, um insbesondere die Entwicklung mathematischer Softwarebibliotheken zu unterstützen. Das wichtigste Ziel ist, dass die Programme nutzerfreundlich für die Anwendungsentwickler werden. Während des Projektes werden prototypische Anwendungen entwickelt, die als Blaupause für große realistische Anwendungsprogramme dienen.



1

BIG DATA: KOMPETENZEN UND PROJEKTE

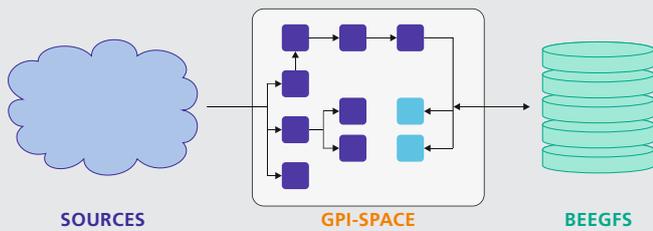
1 Die Kernkomponenten von GPI-Space in der Übersicht, zusammen mit Benutzer- und HPC-spezifischen Komponenten

Neben klassischen HPC-Anwendungen wie Simulationen hat sich Big Data in den letzten Jahren als ein wichtiges Anwendungsfeld für Großrechner etabliert. Absolute Performance und Durchsatz bei der Verarbeitung von Daten spielen für Unternehmen aus den unterschiedlichsten Bereichen eine immer größere Rolle. Die Analyse extrem großer Datenmengen erlaubt nicht nur neue Geschäftsmodelle für die Industrie, auch Forschung und Entwicklung können basierend auf fortschrittlicher Technologie neue Antworten erhalten. Personalisierte Industrie, Identifikation von Krankheitsmarkern, Genomanalysen für alle, Gebäudesteuerung oder Prozessüberwachung für einzelne Bauteile sind hier Stichworte.

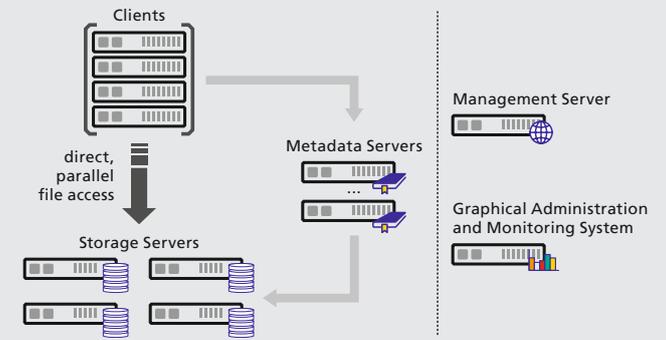
Das CC HPC hat in den letzten Jahren verstärkt an dezidiert anderen und neuen Ansätzen zu den scheinbar alternativlosen Standards im Bereich Big Data geforscht und gearbeitet. Aufbauend auf GPI und unter Berücksichtigung und Weiterentwicklung von Paradigmen aus dem Cloud-Umfeld erstellt das CC HPC die Entwicklungsplattform und Laufzeitumgebung GPI-Space. Es handelt sich dabei um ein Werkzeug, das die Entwicklung und fehlertolerante Ausführung paralleler Software erheblich vereinfacht und für Big Data-Anwendungen prädestiniert ist. GPI-Space löst dazu zwei der wichtigsten und immer noch offenen Probleme bei der Verarbeitung extrem großer Daten: das Programmiermodell für entsprechende Anwendungen und die Ausführungsumgebung.

Kennzeichnend für GPI-Space ist die Abkopplung des Anwendungsentwicklers von den Anforderungen, die an einen Entwickler für ein HPC-System gestellt werden; es gibt ihm also die Freiheit, sich auf die Lösung seines anwendungsspezifischen Problems zu konzentrieren. HPC-Experten können sich unabhängig davon und zeitgleich um die für die jeweilige Anwendungsdomäne relevanten Parallelisierungsprobleme wie I/O, Zerlegbarkeitsmuster, Fehlertoleranz, Datenmanagement und Workflows kümmern. GPI-Space führt diese Domänen- und HPC-Komponenten zusammen und ermöglicht die fehlertolerante, dynamische, parallele Ausführung selbst bereits existierender sequentieller Programme. Neben der Möglichkeit, existierende Daten auf einem schnellen lokalen Dateisystem zu analysieren, wurde auch ein Streaming-Layer integriert, der die Datenverarbeitung direkt von der Quelle und ohne Zwischenspeicherung erlaubt.

Der in GPI-Space verwendete virtuelle Speicher basiert auf einem schnellen internen Hauptspeicher und nicht, wie in anderen Big Data-Lösungen, auf einem langsamen externen Plattenspeicher. Allein dadurch verkürzen sich Antwortzeiten erheblich, insbesondere wenn gleiche Daten unter mehreren Gesichtspunkten betrachtet werden sollen. Der virtuelle Speicher ist zudem unabhängig von konkreten Anwendungen und erlaubt deren einfache und direkte Kopplung. Die Laufzeitumgebung von GPI-Space ist nicht nur fehlertolerant, sondern auch voll dynamisch



2



3

und erlaubt es, die Größe und Struktur der verwendeten Maschine während der Laufzeit eines Algorithmus zu verändern bzw. in verschiedenen Phasen die jeweils optimale Topologie zu verwenden. Auch das ist ein großer Vorteil gegenüber existierenden Lösungen, die üblicherweise während der gesamten Laufzeit so viele Ressourcen bereithalten müssen, wie der hungrigste Teil der Anwendung benötigt.

BeeGFS – bisher unter dem Namen FhGFS bekannt – ist das am CC HPC entwickelte parallele Dateisystem. Von Anfang an waren Skalierbarkeit, maximale Performance, Flexibilität sowie einfache Bedienbarkeit die Eckpfeiler bei der Entwicklung und sind heute die Eigenschaften, die es für eine immer breiter und globaler werdende Nutzerbasis so attraktiv machen. Um das beizubehalten, wurden im vergangenen Jahr vor allem auch interne Optimierungen vorgenommen, aber auch neue sichtbare Features wie Quota-Unterstützung hinzugefügt. 2014 wurden Verkauf sowie Support an ThinkParQ – ein Spin-Off des Fraunhofer ITWM – ausgelagert und auch der neue Produktname vorgestellt. Das CC HPC bleibt aber weiter treibende Kraft hinter der Entwicklung und hat jüngst das »Scalability Lab« in Betrieb genommen, mit dem Dateisystem und Komponenten fit für die Herausforderungen in einer Exascale-Umgebung gemacht werden.

Ziel einer Initiative von Fraunhofer und einigen Partnern aus der Industrie ist die Entwicklung eines Echtzeit-Überwachungssystems für intelligente Energiezähler (Smartmeter) in großen Gebäudekomplexen wie Krankenhäusern, Universitäten, Hotels, Bürogebäuden etc. Diese Smartmeter sind in Verteilerkästen an Schlüsselpositionen innerhalb der Gebäude angebracht und liefern Echtzeitdaten direkt an eine auf GPI-Space basierende Anwendung. Dort werden die Daten mithilfe modernster Algorithmen auf bestimmte Muster hin untersucht. Dies ermöglicht eine nicht-invasive, gerätespezifische Lastüberwachung, ohne jedes Gerät selbst an ein Smartmeter anschließen zu müssen. Die Analyse der Echtzeitdaten sowie Laufzeitanalysen mit gespeicherten Daten erlauben es, den Energieverbrauch des Gebäudes zu optimieren und zu planen. Mittels Data Mining kann dann u. a. eine Vorhersage über den zukünftigen Energieverbrauch auf Grundlage vorangegangener Jahre, Monate, Wochen und Tage erstellt werden, die auch zur optimalen Nutzung von vor Ort erzeugter Energie, z. B. aus Solarzellen, genutzt werden kann. Neben dieser Planung lassen sich mit dieser Methode auch einzelne Geräte mit hohem Energieverbrauch einfach überwachen, so dass diese bei Nichtnutzung automatisch ausgeschaltet oder Fehlfunktionen erkannt werden können.

2 Mithilfe des Streaming Layers kann GPI-Space Daten sowohl direkt von Quellen als auch von Dateisystemen verarbeiten und kombinieren.

3 Übersicht über die Architektur von BeeGFS: Sowohl Anzahl als auch physische Platzierung von Storage Server, Metadaten-Server und Client-Prozessen sind variabel und ermöglichen dem Nutzer größtmögliche Flexibilität.