5 Ways Companies Are Using SUSE High Performance Computing in AI, Machine Learning and Analytics

BOV1043

Jeff Reser
Product & Solution Marketing
SUSE / jeff.reser@suse.com

Kay Tate
ISV Programs Manager
SUSE / ktate@suse.com
Prologue – Setting the Stage
HPC market factoids

HPC ROI is very high - $463 (on average) revenue per dollar; $44 average profit (or cost savings) per dollar invested in HPC\(^1\)

Worldwide HPC revenue expected to reach over $19.5 billion by 2022\(^1\)

Big data combined with HPC creating new solutions, adding many new users/buyers to the HPC space (AI/ML/DL and HPDA are hot new areas)

SUSE runs on nearly half (24) of the top 50 supercomputers (5 RH, 7 CentOS)\(^2\)

SUSE dominates top 100, CentOS gains share in “smaller” supercomputers\(^2\)

Commercial OS Share in Top 500 (represents 111 supercomputers in the list): SUSE 59%, RH 22%, bullx 15%, Ubuntu 4%\(^2\)

---

\(^1\) Hyperion Research, November 2018
\(^2\) Top500 Supercomputer Report, November 2018
Where does SUSE HPC fit?

SUSE does not have an AI/ML solution, but delivers an HPC platform for parallel computing that makes those solutions better (easier to develop, faster and more manageable)

Businesses around the world today are recognizing that a Linux-based HPC infrastructure is vital to supporting the analytics applications of tomorrow

High Performance AI, Machine Learning, Internet of Things and High Performance Data Analytics are powered by High Performance Computing and fueled by big data

HPC is not just for scientific research any longer, and is being adopted across banking, healthcare, utilities and manufacturing
SLE HPC product factoids

**HPC Module** with supported HPC packages – beyond an OS

Supports **Aarch64 (Arm)** and **x86-64** architectures

Many **IHV/ISV/CSP** partnerships

**Actively investing** in HPC

**Competitively priced**
HPC-related projects

- SUSE Package Hub
- Kubernetes, micro-services and containers
- OpenACC and accelerators
- Software-defined storage for HPC
- Arm 64 support gaining popularity
- HPC in the Cloud
- HPC module
SUSE HPC Reference Architecture

Clustering

Interface

Storage

Tools

SUSE Linux Enterprise for High Performance Computing

Compute Cluster

Compute Cluster

Compute Cluster

Compute Cluster

Compute Node

Compute Node

Compute Node

Compute Node

Compute Node

Login Node

Login Node

Login Node

Login Node

High Bandwidth Fabric

Timers

Storage

Lustre

pNFS

XFS

Spectrum Scale (GPFS)

HPC Module

Mockup: User environment management

Mockup: Terminal management

Mockup: Slurm workload management

Mockup: Memory allocation

Mockup: Parallel task executor

Mockup: Parallel remote commands

Mockup: Power management for clusters

Mockup: Monitor kernel RAS

Mockup: File system library

Mockup: Sample scripts
The Spectrum of AI Solutions

Artificial Intelligence
Examples are Google Maps and game play.

Machine Learning
Examples are cyber security, autonomous vehicles and F1 racing.

Neural Networks
Examples are facial and voice recognition.

Deep Learning
Examples are disease identification and energy demand optimization.

Convolutional Neural Networks
Examples are image/video recognition and medical image analysis.

Transfer Learning
For example, knowledge gained while learning to recognize cars could apply when trying to recognize trucks.
Rise of High Performance AI

• HPAI is containerized, many small files in a hyperscale environment with data transfer within a single node (PCI bus)

• Traditional HPC distributes compute over many nodes, smaller number of large files with data transfer between multiple nodes
### AI/HPC use cases are popping up across industries

<table>
<thead>
<tr>
<th>Manufacturing</th>
<th>Banking &amp; Finance</th>
<th>Healthcare</th>
<th>Automotive &amp; Transportation</th>
<th>Security &amp; Safety</th>
<th>Retail</th>
<th>Energy / Oil &amp; Gas</th>
</tr>
</thead>
</table>
| • Computational Fluid Dynamics  
• Heat dissipation systems  
• AI advanced robotics  
• Automated systems  
• High performance designs | • Risk modeling in financial portfolios  
• Fraud detection in real time as transactions are processed between disparate systems  
• Cybercrime  
• High-freq trading applications  
• Pricing / regulatory compliance | • AI for precision and personalized medicine  
• On-demand diagnoses, diagnostic assistance and treatment plans  
• Drug discovery and research  
• Cancer research and cancer cell detection  
• Remote surgeries | • Automotive modeling and aerodynamic design  
• Engine performance and timing, fuel consumption  
• Safety systems  
• Self driving, AI driverless operations  
• Logistics and mapping | • Smart cities and power grids  
• Surveillance  
• Image analysis  
• Facial recognition and detection | • Affinity marketing  
• Smart inventory & supply chain management  
• Cyber fraud  
• Image tagging  
• Speech recognition  
• Sentiment analysis | • Air flow  
• Heating and cooling efficiencies  
• Wind turbines  
• Ocean exploration  
• Seismic imaging  
• Reservoir characterization  
• Sub-surface fault detection |
Casting the Players
Key Communities

- ceph
- OpenACC
- Linaro
- Lustre
- openHPC
- openSUSE
- openstack
Key Capabilities

- Slurm
- MUNGE
- TensorFlow
- OpenBLAS
- HDF
- MVAPICH
- ScaLAPACK
- OpenMP
- PAPI
- Ganglia
- netCDF
- Open MPI
- FFTW
Key ISVs

Sylabs.io

ANSYS

Microsoft Azure

Dassault Systemes

Bright Computing

RogueWave

MathWorks

UNIVA

Altair

Synopsys

nag

aws

IBM Watson

Big Data

AI / ML

HPC
Scene I – Consumer Goods
Consumer Goods – Appliance Design

Goals

• Reduce time in the development of innovative ovens, washing machines and refrigerators
• Simulate/test more products/variations in a virtual design space before committing to a physical prototype

Challenges

• Find the design “sweet spot” within complex customer, safety, regulatory and cost requirements
• Reduce customer returns and associated revenue loss incurred from products and packaging damaged in shipment
• Decrease costs and time-to-market delays associated with physical prototyping
Consumer Goods – HPC + AI/ML

- Saves time and money because prototypes are used for validation, not development
- Improves customer satisfaction with higher quality products that include more customer design features
- Saves money because packaging and product design innovations substantially reduce shipping damage (fewer returns)
- AI/ML applied across the entire product and service cycles – developing leaner, more optimized operations and personalized customer experiences
- Companies are benefiting by achieving higher sales, reducing manufacturing costs and improving the retail brand.
And there’s more …

- **Affinity marketing** – customer buying behaviors, analysis of buying patterns, precise ad targeting and what products to suggest to repeat customers

- **Smart inventory and supply chain management** – product trends, determining what to stock, determining what to stock and how much to buy

- **Cyber fraud** – credit card fraud, maintaining data privacy and combating cybercrime
Consumer Goods – ISV Solutions

Solutions

• **Create virtual design spaces**, where engineers can model advanced design solutions (including multiple requirements) and run these more complex simulations faster

• **Simulate full products** and corresponding packaging as an integrated “system” that withstands shipping rigors and arrives undamaged

Vendors

• **ANSYS** structural mechanics software

• Engineers may do multiple simulations in a day instead of spending weeks building a prototype – the company can now evaluate many more design alternatives and get better product to market faster

• Engineers can put multiple physical systems together in a virtual test to make better product tradeoffs – products can be optimized for energy efficiency or noise reduction or extra consumer features.
Many simulation types from electronics to fluid flow to structures and systems plus multi-physics

More than 50 ANSYS workflows certified on SLE for engineering simulation across the entire product lifecycle, from digital exploration to operations and maintenance

• Racing aerodynamics testing and optimizations
• Integration testing for complex, distributed, smart systems in modern jet aircraft
• Development of fully-electric super sports car
Scene II – Energy & Utilities
Energy & Utilities – Sustainable Energy

Goals
- Diversify and increase pace of innovation
- Improve organizational and infrastructure efficiencies
- Increase responsiveness by aligning product supply
- Achieve safe and compliant operations

Challenges
- Increasing health/safety/environmental regulations and public scrutiny
- Manage complexities and risk of aging infrastructure and new technologies
- Changing priorities and demands of customers and society
Energy & Utilities – HPC + AI/ML

• Simulations and modeling help reduce cost and risk of future operations

• AI/ML/DL revolutionizing the way we produce, transmit and consume energy

• AI/ML/DL helps limit environmental impacts and optimizes supply/demand

• Energy smart grids powered by AI/ML, continuously collecting overwhelming amounts of data from millions of smart sensors nationwide to make timely decisions on how to best allocate energy resources
And there’s more …

Hyperloop

• Tube systems for passengers and vehicles rushing at speeds up to 700 miles/hour
• Magnetic accelerators and compressed air bearings remove frictional forces
• Design anticipates operational maintenance and variables such as earthquakes, power outages and passenger fluctuations
• Design must be safe, reliable, affordable and self-powered.
• HPC simulations enable iterative design changes and show results
Tidal energy

- Tidal turbines use an undulating membrane that “swims” in the moving current
- Tidal waters move the membrane and specialized electronics convert the motion into usable energy
- Tidal energy is more efficient than solar or wind energy, produces no pollution or waste, no visual impact, no noise, no harm to wildlife and does not impede navigation
- HPC simulations tests virtual models to ensure performance, practicality and cost-effectiveness before manufacturing physical prototype that meets sustainability targets
Energy & Utilities – ISV Solutions

Solutions

- **Create consolidated View** of Energy Plants -- the impact of potential changes can be validated to meet the original requirements
- **Expedite renewable energy development** -- such as biomass, water, wind, solar, fuel cells, tidal and wave, and hydropower plays an ever-increasing role in the global share of electricity generation

Vendors

- **Dassault Systemes** provides virtual representations of physical plants and systems, enabling validation and verification of plant design, construction safety procedures, operations and even personnel training.
- **ANSYS** enables engineering and research organizations to better develop new renewable energy equipment. From testing for the electrochemical performance of a fuel cell stack to optimizing the design of biomass reactors and photovoltaic collectors, ANSYS solutions help speed the development process and bring renewable technologies rapidly to market.
Simulation to explore real-world behavior, delivered on the 3DEXPERIENCE platform

- Collaborate, model and optimize operations
- Consumer experience and marketing effectiveness
- Data management for smarter decision-making
- Planet modeling and simulation
Scene III – Automotive
Automotive – Design

Goals

• Design, model, test and refine control strategies in an integrated design environment
• Use processes that reduce design times, enabling engineers to create and test their ideas quickly and with few hardware prototypes
• Provide a seamless environment to take designers’ ideas from concept, through verification to real product code

Challenges

• Speed up design, increase quality and reduce R&D costs by finding an alternative to traditional design methods
• Bridge the gaps in traditional automotive electronics development
Automotive – HPC + AI/ML

• HPC provides the best platform for running design and simulations along with the tools used in the design processes
• AI/ML assists with advanced safety features for connected vehicles
• AI cloud platforms ensure data is available when needed
• AI speeds up the process of filing claims when accidents occur
• Car manufacturing optimized in innovative new ways
• Driving monitoring enhanced by AI, detecting and monitoring driver behavior, recognition and infotainment control
Automotive – ISV Solutions

Solutions

• **Coordinate within the electronic control units (ECUs)** – the under-dash controllers that run the software that runs the vehicles.

• **Focus on performance** for fuel injection and transmission controls to meet emissions standards while delivering improved performance,

Vendors

• **MathWorks** tools have found a real home in ECU development

• **Synopsys** provides tools and IP to help manufacturers create the special chips needed to address special needs in automotive systems

• Specifications for powertrain controllers, for example, now begin in the intuitive and self-documenting environment of Simulink and Stateflow, both powered by the industrial-strength computational, analytical and visualization capabilities of MATLAB.
MathWorks is the leading developer of mathematical computing software for engineers and scientists.

Its universal ECU and embedded coder facilitates rapid prototyping for many car types.

Synopsis delivers the world’s most advanced tools for silicon chip design, verification, IP integration and application security testing, certified and supported on SUSE.

- Automotive design and test
- Delivery of next generation IoT devices
- Cutting edge silicon chip design
Scene IV – Manufacturing
Manufacturing – Materials

Goals
• Build very strong, yet lightweight materials that can withstand extreme temperatures
• Exceed safety and compliance regulations for aerospace, deep space, automotive, oceanic and more

Challenges
• Developing materials requires deep understanding of their underlying properties, and that requires time-consuming computational processes on a massive scale
• Need an intuitive, easy to use and effective way of managing clusters and not take valuable time away from research
Manufacturing – HPC + AI/ML

• HPC modeling helps discover materials faster
• AI/ML predicts which compounds can create materials before setting foot in a lab
• Algorithms can predict which compounds combine to form interesting new materials
• AI can mine databases for “recipes” to make the materials
• Modeling helps quickly map out exactly what makes a material so much stronger or lighter
Bright Cluster Manager manages, provisions and monitors the HPC clusters, allowing researchers to focus on their work instead of managing IT infrastructure.

**Bright Computing** provides full support for HPC via GPUs, from drivers that render GPUs programmable via CUDA, to the CUDA-enabled tools, to CUDA-optimized libraries critical to materials science research.

**Univa** Grid Engine is a leading workload management solution that provides scheduling capabilities to AI/ML and GPU applications.

System admins have complete visibility and control over every aspect of their clusters and gain productivity advantages through the tight integration of the two products.

Clusters are automatically cloud-ready, enabling system administrators to extend their clusters into the cloud, and manage these nodes, with just a few easy steps.
Bright Cluster Manager delivers outstanding cluster management with quick deployments for data science readily done.

Bright Cluster Manager lets users monitor and build clusters of any size that are easy to provision, operate, monitor, manage and scale, all visible on one pane of glass.

Univa Grid Engine is the leading distributed resource management system that optimizes resources by transparently selecting the resources that are best suited for each segment of work.

Grid Engine software manages workload placement automatically, maximizes shared resources, supports extreme scale, provides enterprise-grade dependability and accelerates deployment of any container, application or service in any technology environment, on-premise or in the cloud.

Bright and Univa work together closely to help HPC users set up and manage their systems including their cloud images while smoothly provisioning and managing their workloads across the HPC space.
Scene V – Healthcare
Healthcare – Pharmaceuticals & Drug Research

Goals
• Faster drug discovery, design and validation
• More predictable drug treatment results among patients
• Consolidated view of servers, clusters and instances in the public cloud

Challenges
• Analyze and integrate different types of biomedical and healthcare data
• Integrate public cloud into the existing IT infrastructure with minimal administrative overhead, and no disruption to user workflow
• Find new and naturally occurring compounds from food sources rather than synthetic compounds
Healthcare – HPC + AI/ML

- ML predicts treatment results, matching drug interactions with patients
- ML/DL in drug design, discovery and origins virtually combines atoms to generate new molecular structures
- Combine AI and DNA analyses to discover ingredients in different food sources that have therapeutic qualities, including the management of chronic metabolic diseases.
- AI initiatives in precision medicine and health data applications
Healthcare – ISV Solutions

Solutions

• Create treatments for gene-specific diseases based on better understanding of the genome
• Track and analyze patient outcomes across whole healthcare systems to identify best practices and share them.
• Ensure consistent and to-spec manufacturing of medical devices

Vendors

• Altair PBS Pro, they configured each of the four geographically distinct locations as peers from the workload-management perspective. PBS Pro maintains a real-time inventory of node availability on a per-location basis. Jobs submitted to PBS Pro are separated into six types based upon resource requirements that detail job size, processing and memory requirements, urgency, software image and other needs.
• Bright Cluster Manager, seamless integration with the public cloud can be achieved. The cloud can be easily used as an additional resource on the system with virtually no administrative overhead.
• SAS analytics helps hospital groups gather and analyze very large amounts of patient data to identify and promulgate best practices, improving healthcare on a wide scope.
Altair makes HPC faster, smarter & easy to manage with PBS Works™

Altair provides services for applications that streamline the workflow management of compute-intensive tasks including solvers, optimization, modeling, visualization and analytics

- Simulation technology to synthesize and optimize designs, processes and decisions
- Accurate, reliable vehicle component modeling
- Management and coordination of HPC development
- Innovative drone propulsion design
SAS is a market leader for multi-variate analyses, business intelligence, data management and advanced predictive analytics.

SAS just announced a $1B investment in AI capabilities, R&D and education.

- Business intelligence
- Advanced analytics
- Fraud and security intelligence
Epilogue
We examined 5 industry scenarios ...

<table>
<thead>
<tr>
<th>Industry</th>
<th>Scenario</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consumer Goods &amp; Retail</td>
<td>Appliance design</td>
</tr>
<tr>
<td>Energy &amp; Utilities</td>
<td>Sustainable energy</td>
</tr>
<tr>
<td>Automotive</td>
<td>Design</td>
</tr>
<tr>
<td>Manufacturing</td>
<td>Materials</td>
</tr>
<tr>
<td>Healthcare</td>
<td>Pharmaceuticals &amp; drug research</td>
</tr>
</tbody>
</table>
... and validated some conclusions

• SUSE does not have an AI/ML or GPU solution, but delivers an HPC platform that makes those solutions better (easier to develop, faster and more manageable)

• Businesses around the world today are recognizing that a Linux-based HPC infrastructure is vital to supporting the analytics applications of tomorrow

• HPC is not just for scientific research any longer, and is being adopted across banking, healthcare, utilities and manufacturing

• Key ISVs for SUSE HPC environments:
Questions?

Just for fun, if you figured out my puzzle just send the solution to jeff.reser@suse.com

The first 20 solution winners during SUSECON get a prize!
Unpublished Work of SUSE LLC. All Rights Reserved.
This work is an unpublished work and contains confidential, proprietary and trade secret information of SUSE LLC. Access to this work is restricted to SUSE employees who have a need to know to perform tasks within the scope of their assignments. No part of this work may be practiced, performed, copied, distributed, revised, modified, translated, abridged, condensed, expanded, collected, or adapted without the prior written consent of SUSE. Any use or exploitation of this work without authorization could subject the perpetrator to criminal and civil liability.

General Disclaimer
This document is not to be construed as a promise by any participating company to develop, deliver, or market a product. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. SUSE makes no representations or warranties with respect to the contents of this document, and specifically disclaims any express or implied warranties of merchantability or fitness for any particular purpose. The development, release, and timing of features or functionality described for SUSE products remains at the sole discretion of SUSE. Further, SUSE reserves the right to revise this document and to make changes to its content, at any time, without obligation to notify any person or entity of such revisions or changes. All SUSE marks referenced in this presentation are trademarks or registered trademarks of Novell, Inc. in the United States and other countries. All third-party trademarks are the property of their respective owners.