سال چهاردهم، شماره‌های ۵۱ و ۵۲ ، بهار و تابستان ۱۴۰۱
صص: ۲۰۹_۲۳۰

دو فصلنامه علمی
فناوری اطلاعات و ارتباطات ایران

رایانش با کارایی بالا: الزامات، نیازمندی‌های نسل‌های آتی و محورهای تحقیقاتی

احسان آریانیان* محمد مهدی اثنی‌عشری فاطمه احسانی بشلی** شقایق سادات حسینی بیان*** مسعود ده‌یادگاری** بهنام صمدی****

*گروه سکوهای فناوری اطلاعات، پژوهشکده فناوری اطلاعات، پژوهشگاه ارتباطات و فناوری اطلاعات

**دانشکده مهندسی کامپیوتر، دانشگاه صنعتی خواجه نصیرالدین طوسی، تهران

*** دانشکده مهندسی کامپیوتر، دانشگاه خوارزمی

تاریخ دریافت: ۱۴۰۰/۰۳/۲۹ تاریخ پذیرش: ۱۴۰۰/۰۹/۱۰

نوع مقاله: مروری

چکیده

حرکت فعلی جهان در جهت هرچه توانمندتر کردن سامانه‌های رایانش با کارایی بالا، نشان‌دهنده نیاز روزافزون به این فناوری است. بدیهی است که هرچه این نیاز افزایش یابد، این سامانه‌ها نیز لازم است که توانمندتر شوند تا بتوانند فعالیت‌های بیشتر و سنگین‌تری را اجرا نمایند. در یک نگاه کلان‌نگر، نسل‌های آتی رایانش با کارایی بالا در دو دسته کلی قرار می‌گیرند؛ نسل‌های رایانشی مبتنی بر فناوری‌های نوظهور نظیر نورومورفیک و کوآنتوم و نسل‌های رایانشی مرسوم که به سمت Exascale در حال حرکت هستند. با توجه به اینکه در آینده نزدیک، احتمال عملیاتی شدن کامل و استفاده در مقیاس وسیع از فناوری‌های نوظهور پایین است، در این مقاله، تمرکز بر نسل‌های رایانشی مرسوم قرار گرفته و سعی شده است الزامات و نیازمندی‌های آنها از جنبه‌های سخت‌افزاری و نرم‌افزاری مورد بررسی قرار گیرند. همچنین، فناوری‌های هوش مصنوعی و رایانش ابری به عنوان موتورهای محرکه رایانش با کارایی بالا در جهان مورد بررسی قرار گرفته‌اند تا تاثیر متقابل آنها بر رایانش با کارایی بالا مشخص گردد. در نهایت، موضوعات و محورهای تحقیقاتی در سطح جهان که مورد توجه هستند بررسی و تدوین شده است.

واژگان کلیدی: رایانش با کارایی بالا، محاسبات Exascale، آینده پژوهی، معماری‌های سخت‌افزاری، معماری‌های نرم‌افزاری

۱. مقدمه

رایانش با کارایی بالا، به سرعت در حال گسترش در سطح جهان است و بسیاری از کشورهای پیشرفته و نیز در حال پیشرفت، سرمایه‌گذاری‌های کلانی را در این راستا آغاز نموده‌اند تا بتوانند در جمع دارندگان ابررایانه‌های قدرتمند باشند. کشور ما نیز از حدود اوایل دهه ۱۳۸۰ به صورت کاملاً جدی درگیر ساخت ابررایانه‌ها بوده است و موفقیت‌های چشم‌گیری نیز تا کنون در این زمینه کسب کرده است . آخرین نمونه از این موفقیت‌ها و دستاوردها، راه‌اندازی ابررایانه سیمرغ بوده است که با سرمایه‌گذاری مشترک و

نویسنده مسئول: احسان آریانیان ehsan_arianyan@itrc.ac.ir

همکاری وزارت ارتباطات و فناوری اطلاعات، وزارت علوم، تحقیقات و فناوری، پژوهشگاه فناوری ارتباطات و اطلاعات و دانشگاه صنعتی امیرکبیر راه‌اندازی شده و توان پردازشی ۵۰۰TFLOPS دارد.

رقابت بسیار شدیدی در این حوزه بین کشورهای مختلف وجود دارد که جایگاه مناسبی را در صدر فهرست Top500[1] ابررایانه‌های جهان به خود اختصاص دهند. در این راستا، برنامه‌های آینده‌پژوهانه متعددی در آمریکا[2]، اتحادیه اروپا[3]، ترکیه[4]، عربستان[5]، چین[6] و بسیاری از کشورهای دیگر اجرا شده است، چرا که فناوری در این حوزه به شدت در حال پیشرفت است و عدم همسویی با این پیشرفت‌ها، می‌تواند به سرعت جایگاه یک کشور را در سطح جهان تغییر دهد.

بر این مبنا، در این مقاله سعی کرده‌ایم با نگاهی به آینده حوزه رایانش با کارایی بالا در سطح جهان، فناوری‌های سخت‌افزاری و نرم‌افزاری مورد نیاز برای رسیدن به توان پردازشی Exascale را مرور نماییم و الزامات و نیازمندی‌های آنها را تبیین نماییم. همچنین، موضوعات تحقیقاتی مطرح در این زمینه را بررسی کرده و مشخص نموده‌ایم که کدام موضوعات بیشتر مورد توجه هستند. این اطلاعات می‌تواند به سیاست‌گذاران، فعالین و محققین این حوزه در سطح کشور کمک کند که بتوانند مسیر درستی را جهت اقدامات آتی خود انتخاب نموده و گام‌های معقولی را در این زمینه بردارند، به نحوی که در نهایت منجر به نقش‌آفرینی هرچه موثرتر ایران در سطح جهان در این حوزه گردد.

ادامه این مقاله به این شکل سازماندهی شده است. در بخش دوم، در خصوص فناوری‌ها و نیازمندی‌های سخت‌افزاری نسل‌های آتی رایانش با کارایی بالا پرداخته شده است. بخش سوم، معماری‌ها و سامانه‌های نرم‌افزاری مرتبط را مورد توجه قرار داده است. بخش‌های چهارم و پنجم به معرفی تأثیرات متقابل یادگیری ماشین و رایانش ابری بر آینده رایانش با کارایی بالا اختصاص یافته است. این دو فناوری در شکل‌دهی به آینده رایانش با کارایی بالا اهمیت بسزایی دارند و از این رو مورد توجه ویژه قرار گرفته‌اند. بخش ششم به معرفی محورها و موضوعات تحقیقاتی مطرح در این حوزه می‌پردازد. در نهایت، بخش هفتم جمع‌بندی مقاله خواهد بود.

۲. معماری‌های سخت‌افزاری

۲_۱ مقدمه

با بزرگ شدن ابعاد مسئله‌های پیش رو به لحاظ پیچیدگی محاسباتی و حجم داده‌های مورد نیاز برای پردازش و نیاز به شبیه‌سازی‌های با دقت بسیار بالا در حوزه‌های مختلفی چون علوم، مهندسی و تجارت، گستره‌ای از تلاش‌ها در جهت تجمیع توان محاسباتی در اختیار و ایجاد سیستم‌های با توان محاسباتی بالاتر از کامپیوترهای شخصی برای حل مسائل بزرگ مطرح شده و در حال پیگیری است که مجموعه‌ی فعالیت‌های در این حوزه تحت عنوان رایانش با کارایی بالا شناخته می‌شود[7].

یکی از شناخته‌شده‌ترین انواع راه حل‌های رایانش با کارایی بالا، ابررایانه است. ابررایانه‌ها، آخرین فناوری مبتنی بر سیستم‌های محاسباتی بزرگ هستند که قادرند پروژه‌ها یا برنامه‌های عظیمی را اجرا کنند که بر روی سیستم‌های رایج قابل‌اجرا نیستند. یک ابررایانه شامل هزاران گره محاسباتی است که با هم کار می‌کنند تا یک یا چند کار را انجام دهند و با ترکیب قدرت محاسباتی، سرعت پردازش را برای انجام رایانش با کارایی بالا افزایش می‌دهند. عملکرد یک ابررایانه معمولاً برحسب تعداد عملیات نقطه شناور در ثانیه (FLOPS) به جای تعداد میلیون دستور در ثانیه (MIPS) اندازه‌گیری می‌شود. راه حل‌های رایانش با کارایی بالا دارای سه جز اصلی هستند: محاسبات، شبکه و ذخیره‌سازی [1].

۲_۲ راه حل‌های رایانش با کارایی بالا در محاسبات

در چند سال گذشته، کلاس جدیدی از سیستم‌های رایانش با کارایی بالا ظهور کرده است. این سیستم‌ها برای محاسبات سنگین از معماری‌های نامتعارف پردازنده استفاده می‌کنند و بیشتر از واحدهای پردازش مرکزی (CPU) برای کارهای غیرمحاسباتی مانند ورودی و خروجی و ارتباطات استفاده می‌کنند. نمونه‌های برجسته این سیستم‌ها شامل Cell-based Roadrunner محصول آزمایشگاه ملی لوس آلاموس و خوشه ATI GPU-based Tianhe-1 محصول دانشگاه ملی فناوری دفاعی چین است [2].

۲_۲_۱ واحد پردازش مرکزی

چندین سال است که انتخاب پردازنده‌های موجود برای رایانش با کارایی بالا از اهمیت زیادی برخوردار است. برای این نوع پردازش کمتر از سه نوع عمده وجود دارد که شامل معماری‌های X86، Arm و Power می‌باشد. امروزه اکثر قریب به اتفاق سیستم‌های رایانش با کارایی بالا، از پردازنده‌های اینتل و گاهی اوقات از پردازنده‌های گرافیکی Nvidia تغذیه می‌شوند.

در عمل، نسل دوم پردازنده‌های Epyc، معروف به "Rome"، بخش قابل توجهی از سهم بازار اینتل را در فضای سرور، هم در رایانش با کارایی بالا و هم در سایر نقاط مصرف خواهند کرد. انتظار می‌رود، پردازنده‌های جدید Xeon (Xeon SP) های ۱۴ نانومتری

[1] https://www.top500.org

[2] https://www.exascaleproject.org

[3] https://www.etp4hpc.eu

[4] https://www.truba.gov.tr

[5] https://www.hpc.kaust.edu.sa

[6] http://www.eguangzhou.gov.cn

[7] https://insidehpc.com/hpc-basic-training/what-is-hpc/

"Cooper Lake") و به ویژه Xeon SP های ۱۰ نانومتری " Ice Lake") بتوانند تا حدی این آسیب به بازار اینتل را کاهش دهند. استفاده از معماری Arm نیز به آرامی به رایانش با کارایی بالا رسیده است که نقطه قوت آن این است که IP آن قابل صدور است و بنابراین، این معماری می‌تواند پایه و اساس تعداد زیادی پردازنده‌های سفارشی باشد. Hyperion Research که مدتی است فروش Arm را در HPC ردیابی می‌کند، ۶۴,۷ درصد نرخ رشد مرکب سالانه[1] درآمد پردازنده Arm را در این فضا طی ۵ سال آینده پیش‌بینی می‌کند و انتظار دارد تا سال ۲۰۲۴، فروش تراشه‌های Arm مخصوص دستگاه‌های HPC به بیش از ۶۱۰،۰۰۰ برسد [3].

۲_۲_۲ شتاب دهنده‌های همه‌منظوره

شتاب دهنده‌ها اجزای محاسباتی حاوی واحدهای عملیاتی همراه با سیستم‌های حافظه و کنترل هستند که می‌توانند به راحتی به رایانه‌ها اضافه شوند تا بخش‌هایی از برنامه‌ها را تسریع کنند. دلیل اصلی استفاده از شتاب دهنده‌ها به دلیل نیاز به افزایش کارایی برنامه در جهت کاهش زمان محاسبه و یا افزایش اندازه مسئله‌ای است [2].

اصطلاح GPU computing به معنی استفاده از GPU (واحد پردازش گرافیک) به‌عنوان پردازنده مشترک برای تسریع پردازنده‌های مرکزی برای محاسبات علمی و مهندسی است. از آنجا که هر روز نیاز به قدرت محاسبات توسط GPU بیشتر می‌شود، محققان بیشتری سعی می‌کنند از پردازنده‌های مختلف رایانش با کارایی بالا و پردازنده‌های گرافیکی عمومی (GPGPU) برای بهبود عملکرد کد خود استفاده کنند. GPU با بارگیری برخی از قسمت‌های محاسباتی و وقت‌گیر کد، برنامه‌های در حال اجرا روی CPU را تسریع می‌کند. بقیه برنامه‌ها هنوز روی CPU اجرا می‌شوند. از دید کاربر، برنامه سریع‌تر اجرا می‌شود زیرا از قدرت پردازش کاملاً موازی GPU برای افزایش عملکرد استفاده می‌کند. این به‌عنوان محاسبات "ناهمگن " یا "ترکیبی" شناخته می‌شود.

پردازنده مرکزی از چهار تا هشت هسته پردازنده تشکیل شده است، در حالی که GPU از صدها هسته کوچک‌تر تشکیل شده است. آنها با هم کار می‌کنند تا داده‌های موجود در برنامه را خرد کنند. این معماری کاملاً موازی همان چیزی است که به پردازنده گرافیکی عملکرد بالایی را ارائه می‌دهد. تعدادی از برنامه‌های شتاب‌دهنده GPU وجود دارد که راهی آسان برای دستیابی به رایانش با کارایی بالا ارائه می‌دهند. NVIDIA پیشگام محاسبات توسط شتاب‌دهنده با GPGPU های خود بوده است. در گذشته، عملکرد GPU با اتصال PCIe محدود شده بود. امروزه، معماری Pascal شرکت NVIDIA از SMX2 استفاده می‌کند که اتصال NVLink را برای پهنای باند دوطرفه پرسرعت که پنج برابر سریعتر از PCIe است، امکان‌پذیر می‌کند [4].

توسعه دهندگان نرم‌افزار با استفاده از یک پلتفرم برنامه‌نویسی موازی طراحی شده توسط NVIDIA به نام "CUDA"، عملکرد معماری GPU موازی را کنترل می‌کنند. این پلتفرم با استفاده از OpenCL (آخرین نسخه‌ی آن 3.0 OpenCL است که در سپتامبر ۲۰۲۰ منتشر شد) به‌عنوان یک استاندارد باز، گسترده‌ترین مدل برنامه‌نویسی برای محاسبات GPU است. پردازنده‌های گرافیکی Tesla به‌عنوان شتاب‌دهنده‌های محاسباتی یا کمک پردازنده‌های بهینه‌شده برای کاربردهای محاسبات علمی و فنی طراحی شده‌اند. آخرین پردازنده‌های گرافیکی سری ۲۰ تسلا بر اساس آخرین پیاده‌سازی‌های پلتفرم CUDA موسوم به " Fermi architecture" ساخته شده‌اند [5,6].

امروز AMD با رونمایی از AMD Instinct MI100 سریع‌ترین پردازنده گرافیکی رایانش با کارایی بالا، گام بزرگی به سمت محاسبات پردازش سریع برمی دارد که به‌طور ویژه برای بارهای مورد نیاز در محاسبات علمی، هدف گذاری شده است. شتاب‌دهنده جدید AMD Instinct MI100، سریع‌ترین HPC GPU جهان و اولین GPU سرور x86 است که از سد 10 TFLOPS (FP64)پیشی گرفته است. MI100، همراه با پردازنده‌های AMD EPYC و بستر نرم‌افزاری باز 4.0 ROCm، با پشتیبانی از پلتفرم‌های جدید محاسباتی تسریع‌شده از شرکت‌های Dell ، GIGABYTE، HPE و Supermicro، برای پیشبرد اکتشافات جدید قبل از عصر Exascale طراحی شده است.

پردازنده گرافیکی AMD Instinct MI100 که بر اساس معماری جدید AMD CDNA ساخته شده است، در صورت جفت شدن با نسل دوم پردازنده‌های AMD EPYC، کلاس جدیدی از سیستم‌های شتاب‌دهنده را برای رایانش با کارایی بالا و AI امکان‌پذیر می‌کند. MI100 حداکثر عملکرد 11.5 TFLOPS با دقت FP64 و حداکثر عملکرد 46.1 TFLOPS با دقت FP32 برای بارهای محاسباتی در هوش مصنوعی و یادگیری ماشین ارائه می‌دهد. MI100 با استفاده از فناوری جدید AMD Matrix Core، در مقایسه با شتاب دهنده‌های نسل قبلی AMD، تقریباً ۷ برابر افزایش در عملکرد حداکثری محاسبات نقطه شناور FP16 برای بارهای آموزشی AI فراهم می‌کند [7].

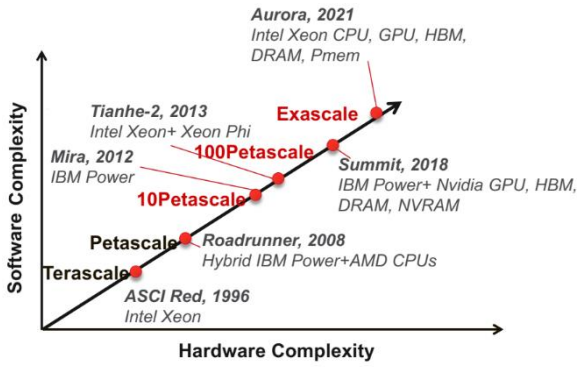
۲_۲_۳ شتاب‌دهنده‌های خاص‌منظوره

ASIC[2] (مدار مجتمع خاص‌منظوره) یک مدار مجتمع است که برای استفاده خاص تنظیم‌شده‌است، و نه برای استفاده عمومی

[1] Compound annual growth rate (CAGR)

[2] application-specific integrated circuit

مانند CPU. مدارات مجتمع خاص‌منظوره از GPU ها اختصاصی‌تر هستند، زیرا GPU هنوز یک پردازنده کاملاً موازی با هزاران واحد محاسباتی است که قادر به اجرای الگوریتم‌های مختلف است، در حالی که ASIC پردازنده‌ای است که برای انجام یک سری محاسبات بسیار کوچک (مثلاً فقط ضرب ماتریسی) طراحی شده است اما بسیار خوب انجام می‌شود. در حال حاضر حرکت‌های زیادی به سمت ASIC انجام شده است و نام‌های مختلفی برای این پردازنده‌ها وجود دارد: NPU[1]، XPU، TPU، IPU که به نظر می‌رسد IPU[2] و TPU مشهورترین نمونه‌های ASIC در حوزه‌ی رایانش با کارایی بالا باشند. به‌طور کلی، به نظر می‌رسد که کاهش تدریجی قانون مور[3] منجر به دامنه وسیع‌تری از کاربرد شتاب دهنده‌ها یا فناوری‌های نسبت به آنچه در سه دهه گذشته دیده‌ایم، می‌شود [8].



شکل ۱. نمودار پیچیدگی نرم‌افزاری و سخت افزاری ابررایانه‌ها [8]

۲_۲_۳_۱ شتاب دهنده‌های Graphcore IPU

مدل‌های دانشی که می‌خواهیم در سیستم‌های هوش ماشین ایجاد و دستکاری کنیم، به‌طور طبیعی به‌صورت گراف بیان می‌شوند که در آنها رأس‌ها ویژگی‌های داده را نشان می‌دهند و لبه‌ها همبستگی یا علل بین ویژگی‌های متصل را نشان می‌دهند. همه چارچوب‌های اصلی یادگیری ماشین مانند Tensorflow، MXnet، Caffe و غیره از گراف‌ها به‌عنوان ساختار اصلی داده‌ها استقبال کرده‌اند. این گراف‌ها موازی سازی زیادی را در داده‌ها و محاسبات نشان می‌دهد که می‌تواند توسط یک پردازنده کاملاً موازی و مخصوص پردازش گراف مورد بهره برداری قرار گیرد.

آخرین فناوری شرکت نیمه‌هادیGraphcores ، واحد پردازش هوشمند یا IPU است که یک پردازنده‌ی کاملاً موازی برای تسریع در عملیات هوش مصنوعی است. یک هسته اصلی و جدید برای زیرساخت‌های هوش مصنوعی، IPU-M2000 است که توسط چهار پردازنده IPU نسل دوم به نام Colossus Mk2 GC200، ساخته شده‌است. این هسته، محاسبات هوش مصنوعی با سرعت 1PetaFlops با دقت مختلط، حداکثر ۴۵۰ گیگابایت حافظه تبادل[4] (DRAM بیرون تراشه) و ۲,۸ ترابایت در ثانیه IPU-Fabric برای ارتباط با تأخیر فوق العاده کم را در یک تیغه باریک 1U[5] پوشش می‌دهد تا بیشترین میزان بارهای هوش ماشین را برطرف کند. همچنین IPU-M2000 دارای طراحی پیمانه‌ای و انعطاف پذیر است و به کاربر این امکان را می‌دهد تا با یک واحد شروع کند و مقیاس آن را به هزاران برساند. ظرفیت مقیاس سازی با فناوری اتصال بی نظیر IPU-Fabric تکمیل می‌شود. IPU-Fabric یک فناوری شبکه AI ویژه برای سیستم‌های IPU-M2000 و IPU-POD است که در نهایت سرعت پردازش حجم کار هوش مصنوعی را تسریع می‌کند. IPU-Fabric با پوشش ارتباط درون و بین رک‌ها، ضمن مقیاس دهی هزار IPU، تاخیر ارتباطی را با استفاده از توپولوژی ۳D ring نزدیک به ثابت حفظ می‌کند [9].

سیستم‌های IPU برای مراکز داده مجازی طراحی شده‌اند. آنها تخصیص و تأمین منابع سخت افزاری مجازی شده را با استفاده از ابزارهای استاندارد صنعتی مانند Docker و Kubernetes ارائه می‌دهند [10].

۲_۲_۳_۲ واحد پردازش تنسور Google TPU

واحد پردازش تنسور[6] (TPU) یک مدار مجتمع با کاربرد خاص(ASIC) است که توسط گوگل به‌طور خاص با هدف تسریع در محاسبات شبکه‌های عصبی، توسعه یافته است. این تراشه به‌طور خاص برای چارچوب TensorFlow Google طراحی شده است.

در مقایسه با واحد پردازش گرافیک، TPU برای حجم بالایی از محاسبات با دقت کم (مثلاً با دقت ۸ بیت) [11] و با عملیات ورودی/خروجی بیشتر در هر ژول[7] طراحی شده‌است.

انواع مختلف پردازنده‌ها برای انواع مختلف مدل‌های یادگیری ماشین مناسب هستند، TPU ها برای CNN مناسب هستند در حالی که GPU ها برای برخی شبکه‌های عصبی کاملاً متصل مزیت‌هایی دارند و CPU ها می‌توانند مزایایی برای شبکه‌های RNN داشته باشند [12].

Google از طریق سرویس Cloud TPU خود که بخشی از بستر Google Cloud Platform است و همچنین از طریق سرویس مبتنی بر نوت بوک Kaggle، امکان دسترسی اشخاص ثالث به TPU را میسر کرده است.

---

[1] Neural network Processing Unit

[2] Intelligence Processing Unit

[3] قانون مور بیان می‌کند که تعداد ترانزیستورها در یک مدار مجتمع متراکم (IC) تقریباً هر دو سال دو برابر می‌شود

[4] exchange memory

[5] rack unit

[6] Tensor Processing Unit

[7] joule

در مرکز داده Google، دستگاه‌های TPU با پیکربندی‌های زیر برای TPU v2 و TPU v3 در دسترس هستند. تک دستگاه‌های TPU که از طریق یک شبکه اختصاصی با سرعت بالا به یکدیگر متصل هستند و TPU Podها که ابررایانه‌هایی از دستگاه‌های TPU هستند که از طریق شبکه‌های اختصاصی پرسرعت به یکدیگر متصل می‌شوند. در جدول ۱ نسل‌های ۲ و ۳ TPU با یکدیگر مقایسه شده‌اند.

۲_۳ راه حل‌های رایانش با کارایی بالا در شبکه‌سازی

برای ساخت یک معماری رایانش با کارایی بالا، سرورهای محاسبات به‌صورت خوشه‌ای با هم شبکه می‌شوند. برنامه‌ها و الگوریتم‌های نرم‌افزاری به‌طور همزمان روی سرورهای موجود در خوشه اجرا می‌شوند. این خوشه برای ذخیره خروجی به شبکه ذخیره‌سازی داده متصل است. این اجزا با هم یکپارچه عمل می‌کنند تا مجموعه متنوعی از وظایف را انجام دهند.

جدول ۱. مقایسه نسل‌های مختلف TPU [13,14]

ویژگی	TPUv2	TPUv3
تعداد هسته در هر تراشه	۲	۲
تعداد تراشه در هر کارت	۴	۴
تعداد کارت در هر قفسه	۴	۲
تعداد قفسه در هر رک	۸	۱۶
تعداد رک در هر پاد	۴	۸
تعداد تراشه در هر پاد	۲۵۶	۱۰۲۴
نرخ کلاک (MHz)	۷۰۰	۹۴۰
حافظه HBM در هر تراشه (GB)	۱۶	۳۲
حافظه HBM در هر کارت (GB)	۶۴	۱۲۸
توان مصرفی در هر تراشه (W)	۲۸۰	۴۵۰
فناوری ساخت	کمتر از۱۲nm	کمتر از۱۲nm
حداکثر کارایی در هر تراشه (PFLOPS)	(۱۶بیت) ۴۶ (۳۲بیت) ۳	(۱۶ بیت) ۱۲۳ (۳۲ بیت) ۴
حداکثر کارایی در هر پاد (PFLOPS)	۱۱،۸	۱۲۶

برای کار با حداکثر عملکرد، هر یک از اجزا باید با سایر قطعات همگام باشد. به عنوان مثال، مؤلفه ذخیره‌سازی باید بتواند همگام با سرعت پردازش داده‌ها، آن‌ها را از سرورهای محاسباتی خوانده و یا ذخیره کند. به همین ترتیب، اجزای شبکه باید بتوانند از انتقال سریع داده‌ها بین سرورهای محاسباتی و ذخیره اطلاعات پشتیبانی کنند. اگر یک جز نتواند با مابقی هماهنگ شود، عملکرد کل زیرساخت رایانش با کاریی بالا آسیب می‌بیند [1].

اتصالات پرسرعت در رایانش با کارایی بالا را می‌توان تقریبا به دو دسته‌ی تقسیم کرد: اتصالات میان‌گره‌ای (اتصالاتی که بین گره‌های محاسباتی استفاده می‌شود) و اتصالات درون‌گره‌ای (اتصالاتی که در داخل هر گره محاسبه استفاده می‌شود) [15].

۲_۳_۱ اتصالات میان‌گره‌ای

شبکه‌های میان‌گره‌ای یک نیاز اساسی برای اتصال گره‌ها به منظور تشکیل یک سیستم مستقل و بزرگ هستند. تاکنون، برای سیستم‌های مقیاس متوسط و بزرگ، اتصال میان‌گره‌ای معمولاً به دو بخش اتصالات میان‌گره‌ای برای انتقال ترافیک کاربر و اتصالات میان‌گره‌ای برای انتقال اطلاعات مدیریتی تقسیم شده است. این تمایز عموماً به دلایل امنیتی انجام می‌شود. بسیاری از شبکه‌هایی که در حال حاضر مورد استفاده قرار می‌گیرند، شبکه‌های مبتنی بر سوئیچ هستند که در توپولوژی‌های مختلف مستقر شده‌اند تا پهنای باند کلی را بهینه کنند و تأخیر را پایین نگه دارند. برخی از اتصالات میان‌گره‌ای پرسرعت در رایانش با کارایی بالا عبارتند از: Ethernet ،ROCE ،InfiniBand ،Omni-Path ،Bull و... eXascale ،Slingshot

از میان موارد فوق، اترنت (Ethernet) به عنوان یکی از مهم‌ترین اتصالات میان‌گره‌ای است که به صورت گسترده‌ای در ابررایانه‌ها مورد استفاده قرار می‌گیرد. اترنت ۲۰۰ و ۴۰۰ گیگابیتی در پایان سال ۲۰۱۷ به‌عنوان یک استاندارد تصویب شده است و محصولات آن در حال حاضر در دسترس هستند. به عنوان مثال Mellanox هم کارتهای شبکه ۲۰۰ گیگابیت بر ثانیه و هم سوئیچ‌های ۴۰۰ گیگابیت بر ثانیه را در دسترس قرار داده است. دو نرخی که به احتمال زیاد به‌عنوان نرخ‌های بعدی تعیین می‌شوند، ۸۰۰ گیگابیت بر ثانیه و ۱،۶ ترابیت بر ثانیه است. در حال حاضر، هنوز هیچ استانداردی وجود ندارد که بتواند یک ترابیت بر ثانیه را از طریق اترنت فقط با یک کابل نوری (یا طول موج) فعال کند. اما وعده رسیدن به ترابیت اترنت نزدیک‌تر از آن است که به نظر می‌رسد. ائتلاف اترنت ۲۰۲۰ پیش بینی می‌کند که سرعت‌های ۸۰۰ گیگابیت بر ثانیه و ۱،۶ ترابیت در ثانیه احتمالاً بین ۲۰۲۳ تا ۲۰۲۵، به استاندارد IEEE تبدیل شوند [16].

ماژول‌های نوری احتمالاً یکی از مهم‌ترین اجزای فناوری ترابیت اترنت هستند، به ویژه فرستنده و گیرنده‌های مهمی مانند QSFP28، QSFP-DD، OSFP و CFP8. این ماژول‌های کوچک، قابل اتصال، کم مصرف و سازگار با فروشندگان مختلف هستند. امروزه، ماژول فرستنده/ گیرنده‌ی نوری QSFP28 به‌طور گسترده‌ای مورد استفاده قرار می‌گیرد. تجهیزات شبکه که از ۱۰۰ گیگابایت پشتیبانی می‌کنند اغلب با QSFP28 همراه هستند. به ویژه QSFP-DD دارای یک ماژول نوری جدید است که با مدولاسیون PAM4 بر روی فیبر نوری کار می‌کند. این ماژول قادر به پشتیبانی از ۲۰۰ گیگابیت بر ثانیه و حتی ۴۰۰ گیگابیت بر ثانیه از پهنای باند کل است [17].

۲_۳_۲ اتصالات درون‌گره‌ای

اتصال درون گره‌ای یکی از مهم‌ترین چالش‌های بهبود کارایی معماری‌های ناهمگن است. سرورهای مدرن شامل تعداد زیادی دستگاه مختلف است که نیاز به برقراری ارتباط با یکدیگر دارند. یک گره معمولی رایانش با کارایی بالا ممکن است شامل چندین پردازنده مرکزی (CPU)، پردازنده گرافیکی (GPU)، کارتهای شبکه (Network Adapter) و حافظه ذخیره‌سازی SSD (حافظه SSD) باشد که همه برای کارکرد آن باید به یکدیگر متصل شوند. یک سرور معمولی شامل چندین اتصال داخلی پر سرعت است. به عنوان مثال، گره‌ای با چندین CPU اینتل از UPI برای اتصال CPU ها به یکدیگر، از PCI-e برای اتصال حافظه، شبکه و GPU استفاده می‌کند، در حالی که از یک لینک پهنای باند بالا مانند NVIDIA NVLink، برای اتصال همه شتاب دهنده‌ها به یکدیگر استفاده می‌کند. برخی از اتصالات درون‌گره‌ای پرسرعت در رایانش با کارایی بالا عبارتند از: PCI-e، CXL[1]، CCIX[2]، UPI[3]/QPI[4]، Infinity Fabric، CAPI/OpenCAPI[5]، NVLin، Gen-Z و...

PCI express (PCI-e) اصلی‌ترین گذرگاه داخلی است که امروزه برای اتصال دستگاه‌ها به پردازنده‌ها در داخل یک گره استفاده می‌شود. استاندارد PCI-e به‌سرعت در حال پیشرفت است. در حال حاضر صنعت به‌سرعت در حال انتقال به PCI-e Gen 4 با نرخ انتقال 16GT/s[6] با توان نظری 31.5 GT/s برای درگاه x۱۶ با همان رمزگذاری PCI-e Gen 3 است. در حال حاضر، کمپانی‌های AMD و IBM و Intel سرورهایی با PCI-e Gen 4 ارائه می‌دهند. مرحله تکاملی بعدی PCI-e Gen 5 است که نرخ انتقال نسبت به Gen 4 به دو برابر می‌شود و تقریباً ۶۳ گیگابایت در ثانیه پهنای باند را از یک پورت x۱۶ می‌گیرید. CPU و GPU های پشتیبانی کننده از PCI-e Gen 5 انتظار می‌رود بین سال‌های ۲۰۲۱ و ۲۰۲۲ باشند [17].

۲_۴ راه حل‌های رایانش با کارایی بالا در ذخیره‌سازی

حجم داده‌های در دسترس در مجموعه‌داده‌های در اختیار برای پردازش و به موازات آن کاربردهای دارای نیازمندی به پردازش سریع داده‌ها شدیداً در حال رشد هستند بنابراین عملیات ذخیره‌سازی در رایانش با کارایی بالا باید به گونه‌ای باشد که بتواند داده ها با سرعت نور پردازش شود تا بازده محاسبات در بالاترین سطح حفظ شود. داده هایی که باید از سیستم ذخیره‌سازی خوانده شوند و یا روی آن بنویسند، باید در حد اگزابایت (EB) باشد. سیستم ورودی و خروجی برای ذخیره‌سازی باید تا حد امکان از پردازنده مستقل باشند تا به محاسبات ارزشمندی که در حال انجام است وقفه‌ای وارد نشود. سیستم‌های ذخیره‌سازی رایانش با کارایی بالا این امکان را به پردازنده‌ها می‌دهند تا در حالی که سیستم در حال نوشتن و یاخواندن داده‌ها از روی دیسک است، به پردازش خود ادامه دهند. این مساله باید از امنیت داخلی بالایی برخوردار باشد، تحمل‌پذیری اشکال داشته باشد، به‌صورت پیمانه‌ای طراحی شده باشد و از همه مهمتر از نظر مقیاس به‌صورت یکپارچه باشد.

بیش از ۵۰ درصد از معماری ذخیره‌سازی جهانی، Luster را که یک فایل‌سیستم موازی متن باز برای پشتیبانی از خوشه های رایانش با کارایی بالا است، ترجیح می دهند. این فایل سیستم علاوه بر این، قابلیت های ذخیره‌سازی گسترده اطلاعات همراه با پیکربندی یکپارچه، مدیریت متمرکز، نصب ساده و مقیاس پذیری قدرتمندی را فراهم می کند [18].

Luster یک فایل‌سیستم موازی توزیع‌شده متن‌باز، تک فضای نام جهانی[7] و سازگار با POSIX[8] است که به منظور مقیاس پذیری، عملکرد بالا و در دسترس بودن بالا طراحی شده است. Luster بر روی سیستم عامل های مبتنی بر Linux اجرا می‌شود و از معماری شبکه کلاینت-سرور استفاده می‌کند. فضای ذخیره‌سازی توسط مجموعه ای از سرورها فراهم می شود که می‌تواند به جمعیت‌هایی که تا چند صد میزبان دارند برسند. سرورهای لوستر فقط یک فایل‌سیستم، می‌توانند ده‌ها پتابایت فضای ذخیره‌سازی با بیش از یک ترابایت در ثانیه توان ترکیبی را در اختیار هزاران کلاینت محاسباتی قرار دهند [19].

Luster فایل‌سیستمی است که نیازهای برنامه‌هایی که بر روی طیف وسیعی از سیستم‌ها (از رایانش با کارایی بالا در مقیاس کوچک تا بزرگترین ابررایانه‌ها) اجرا می‌شوند را برآورده می‌کند و با استفاده از بلوک های سازنده‌ی ذخیره‌سازی مبتنی بر شی برای ایجاد مقیاس پذیری بیشتر ایجاد شده است.

یک راه حل بهینه برای ذخیره‌سازی در رایانش با کارایی بالا آن است که براساس نیازهای رایانش با کارایی بالا، مقیاس بندی شود. در حالت ایده‌آل، باید ترکیبی مناسب از این دو مورد را داشته باشد: فضای ذخیره‌سازی سنتی (دیسک درایوها) و cloud (SSD ها و HDD ها). عملیات ورودی/خروجی پیچیده و پرمصرف داده را می‌توان به SSD هدایت کرد، در حالی که جریان داده‌های معمول را می‌توان توسط دیسک درایوها اداره کرد. ترکیبی کارآمد

[1] Compute Express Link

[2] Cache coherent interconnect for accelerators

[3] Ultra-Path Interconnect

[4] Quick Path Interconnect

[5] Coherent Accelerator Processor Interface

[6] Giga Transfers/s

[7] Global single-namespace

[8] Portable Operating System Interface

از ابر هیبریدی، فضای ذخیره‌سازی تعریف شده توسط نرم‌افزار و پیکربندی سخت‌افزار، در نهایت به کارایی مقیاس کمک می کند و نیاز به داشتن یک لایه ذخیره‌سازی جداگانه را از بین می‌برد. همچنین، حافظه رایانش با کارایی بالا باید از طریق سیستم ورودی و خروجی پیچیده از فایل سیستم‌های موازی هم پشتیبانی کند.

۳. معماری‌های نرم‌افزاری

۳_۱ مقدمه

گسترش و توسعه‌ی نرم افزارهای مختلف همگام با فناوری‌های سخت‌افزاری در روند توسعه‌ی رایانش با کارایی بالا از اهمیت بالایی برخوردار است. با توجه به پیشرفت و گسترش معماری‌ها و فناوری‌های سخت افزاری در راستای توسعه‌ی رایانش با کارایی بالا نیاز به گسترش و توسعه‌ی ابزارهای نرم‌افزاری متناسب با سخت‌افزار آن بیش از پیش وجود دارد. در این راستا بسیاری از سازمان‌ها و شرکت‌های توسعه دهنده‌ی مدل‌های برنامه نویسی و نرم‌افزار و ابزارهای توسعه‌ی رایانش با کارایی بالا، اقدام به نوآوری و تولید و یا گسترش فناوری‌های موجود پرداخته‌اند. در بخش‌های بعدی بررسی بر روی مدل‌های برنامه نویسی موازی رو به رشد و نرم افزارها و ابزارهای توسعه یافته در حوزه‌های مختلف و فعالیت‌های نرم‌افزاری برخی مراکز فعال در این زمینه بررسی می‌شود.

۳_۲ فناوری‌های نرم‌افزاری

تعداد بسیاری از استراتژی‌ها، پلت فرم‌ها و کتابخانه ها در سال‌های اخیر ارائه شده‌اند که هیچ کدام از آن‌ها را نمی‌توان به عنوان یک راه حل کلی برای معماری‌های موازی در حال حاضر در نظر گرفت. در میان مدل‌های نرم افزاری مختلف، مدل‌هایی که در دو سال اخیر روند رو به رشدی داشته‌اند و سرعت بیشتری در گرایش به سمت رشد و به کارگیری در راستای رایانش با کارایی بالا داشته‌اند بر طبق گزارش [20] مورد بررسی قرار گرفته‌اند.

۳_۲_۱ مدل برنامه نویسی مبتنی بر پیغام/حافظه‌ی توزیع شده

مدل‌های برنامه نویسی مبتنی بر پیغام مدل‌هایی هستند که پردازش‌ها از حافظه‌ی محلی خود در طول محاسبات استفاده می‌کنند. پردازش‌ها می‌توانند در یک ماشین و یا در میان تعداد دلخواهی از ماشین‌ها اجرا شوند. پردازش‌های موازی داده را از طریق انتقال پیام بین یکدیگر انجام می‌دهند. این ارتباطات می‌توانند به طور غیرهمزمان (زمانی که پیام می‌تواند قبل از اینکه دریافت کننده آماده باشد ارسال شود.) و یا همزمان (زمانی که دریافت کننده باید آماده باشد.) باشند. در ادامه مهم‌ترین و شناخته شده ترین مدل برنامه نویسی بر این مبنا معرفی می‌شود [21].

۳_۲_۱_۱ رابط انتقال پیام[1]

سیستم رابط انتقال پیام MPI یک استاندارد مشخص و قابل حمل است که به منظور به کار بردن در معماری‌های محاسباتی موازی مختلف استفاده می‌شود. این استاندارد مفاهیم و سینتکس‌های مجموعه‌ای از کتابخانه‌های معمول را برای استفاده در محدوده‌ی گسترده‌ای از کاربردها برای برنامه-های قابل حمل ارسال پیام در C, C++ و Fortran تعریف می-کند. سیستم‌های رابط انتقال پیام برای ماشین‌های پردازش انبوه استفاده می‌شود. چندین پیاده‌سازی کارآمد و آزمایش‌شده برای MPI وجود دارد که بسیاری از آن‌ها منبع باز و در دامنه‌های عمومی در اختیار گرفته‌اند. MPI یک پروتکل ارتباطی برای برنامه‌نویسی کامپیوترهای موازی است که هم ارتباط نقطه به نقطه و هم ارتباط جمعی را پشتیبانی می‌کند. استاندارد MPI تا کنون در سه نسخه‌ی MPI-1 و MPI-2 و MPI-3 معرفی شده است که MP-1محتوایی برای حافظه‌ی اشتراکی نداشت و MPI-2 تنها محتوای محدودی برای حافظه‌ی اشتراکی توزیع-شده داشت. برنامه نویسی قطعی حافظه اشتراکی در MPI-3 به طور کامل معرفی شده‌است. اکثر پیاده‌سازی‌های MPI شامل مجموعه‌ی مشخصی از روال‌ها هستند که به طور مستقیم در C, C++ و Fortran و هر زبان دیگری که بتواند رابط با این کتابخانه‌ها داشته باشد مانند Java, C# و Python قابل اجرا هستند.

MVAPICH یک پیاده‌سازی متن باز از MPI می‌باشد که برای استفاده در رایانش با کارایی بالا از آن استفاده می‌شود. تا کنون حدود ۱,۲ میلیون بارگیری داشته است و تعداد سازمان-هایی که از آن استفاده کرده‌اند حدود ۳۱۲۵ سازمان می‌باشد. این پروژه توسط آزمایشگاه محاسبات مبتنی بر شبکه[2]ی دانشگاه دولتی اوهایو هدایت می‌شود. این پیاده سازی که براساس استاندارد MPI 3.1 می‌باشد .بهترین عملکرد و قابلیت را برای سیستم‌های فوق سریع و سرورهایی که از فناوری-های شبکه‌ای InfinitBand,OmniPath,Ethernet/WARP,RoCE استفاده می‌کنند، دارد. همچنین آزمایشگاه محاسبات مبتنی بر شبکه یک نسخه‌ی جدید به نام MPI4cuML ارائه کرده است که برای پایتون ۳,۷ و یا کودا ۱۰,۱ یا ۱۰,۲ یا ۱۱ ساخته شده است که یک پشتیبانی بهینه از وظایف یادگیری ماشین در سطح MPI ارائه کرده است که سیستم پیام جمعی مانند Allreduce را نیز پشتیبانی می‌کند [22].

۳_۲_۲ مدل برنامه نویسی مبتنی بر حافظه‌ی اشتراکی

[1] Message Passing Interface

[2] Network-based computing laboratory

در مدل‌های برنامه نویسی مبتنی بر حافظه‌ی اشتراکی پردازش‌های موازی یک فضای آدرس عمومی که می‌توانند به طور غیرهمزمان به آن دسترسی داشته باشند به اشتراک می-گذارند. سازوکارهای بسیاری از جمله قفل‌ها[1] و سمافورها[2] برای کنترل دسترسی به حافظه‌ی اشتراکی و جلوگیری از ایجاد شرایط رقابتی و وقوع بن بست به کار می‌روند [21].

۳_۲_۲_۱ پردازش چندگانه‌ی باز[3] OpenMP

OpenMP یک رابط ساده و انعطاف پذیر برای توسعه‌ی برنامه نویسی موازی در C, C++ و Fortran است. ویژگی‌های OpenMP برای برنامه‌های کابردهای پیچیده‌ای که نیاز به الگوریتم‌های موازی سازی بی قاعده دارند استفاده می‌شوند. ویژگی‌های OpenMP راه به روزی برای ارائه‌ی همزمانی و موازی سازی است. آخرین نسخه‌ی منتشر شده‌ی آن در تاریخ ۱۳ نوامبر ۲۰۲۰ انتشار یافته است. ویژگی بارز OpenMP که آن را در حوزه‌ی رایانش با کارایی بالا کارآمد می‌کند این است که از نسخه‌ی ۴ که در جولای ۲۰۱۳ منتشر شد کار با شتاب دهنده ها را پشتیبانی می‌کند، و پس از آن در نسخه‌های ۴,۵ و ۵ پشتیبانی از شتاب دهنده ی را به میزان قابل توجهی بهینه کرده است. در نسخه‌ی ۴ و بعد از آن GPU و FPGA را پشتیبانی می‌کند [23].

۳_۲_۲_۲ Kokkos

یک مدل برنامه نویسی سطح بالا برای نوشتن برنامه‌های قابل حمل بر روی پلت فرم‌های رایانش با کارایی بالا است. کتابخانه‌ای است که از پیاده سازی ساده‌ی یک برنامه بر روی یک هسته تا اجرای کارآمد برروی انواع سخت افزارها مانند GPU و Intel Xeon Phis و CPUهای با چندین هسته را امکان پذیر می‌کند. Kokkos هسته‌ی C++ را به فناوری‌های برنامه نویسی مختلف از جمله CUDA، OpenMP و یا PThreads نگاشت می‌کند. Kokkos همچنین شکل‌هایی از ساختار داده را برای تطبیق با لایه‌های حافظه مانند آرایه‌های دوبعدی و سه بعدی فراهم می‌کند که عملکرد را برروی سخت افزارهای مختلف بهینه می‌کند. در حال حاضر Kokkos از چهار حالت اجرایی پشتیبانی می‌کند که شامل حالت MPI-Only برای CPU و Intel Phi، OpenMP برای CPUهای با تعداد هسته‌ی بالا و Intel Phi، CUDA برای GPUهای NVIDIA و HIP برای GPUهای AMD است. حالت مورد نظر را می‌توان در زمان ساختن برای اجرای منطبق با سخت افزار مورد نظر اعمال کرد [24].

3_2_3 مدل برنامه نویسی مبتنی بر وظیفه

در مدل برنامه نویسی مبتنی بر وظیفه تمرکز بر روی پردازش‌ها است. این پردازش‌ها یا وظایف به صورت متمایز در نظر گرفته می‌شوند وظایف می‌توانند با یکدیگر ازتباط برقرار کنند، این ارتباط می‌تواند در طول اجرای وظایف باشد و یا به عنوان یک ورودی پیش نیاز در آغاز وظیفه و درنهایت ارسال نتیجه به وظیفه‌ی دیگر در پایان انجام آن باشد [25].

۳_۳_۲_۱ OmpSs

OmpSsیک مدل برنامه نویسی جهت موازی سازی غیرهمزمان بر اساس ساختار OpenMP و StarSs است. موازی سازی غیرهمزمان در OmpSs با استفاده از وابستگی داده بین وظایف مختلف برنامه امکان پذیر شده است. OmpSs2 یک مدل برنامه نویسی ترکیب شده از روال های کتابخانه ای است که می-تواند به منظور توسعه‌ی برنامه های کاربردی همزمان استفاده شود. OmpSs2 مدل وظیفه‌ی OmpSs/OpenMP است که پشتیبانی وظایف تو در تو و وابستگی های ریز در سطوح تودرتویی مختلف را حفظ می‌کند که موازی سازی موثر برنامه ها را با استفاده از روش بالا به پایین امکان پذیر می‌کند. هدف فراهم کردن یک محیط تولیدی برای توسعه‌ی برنامه-های کاربردی برای سیستم‌های مدرن رایانش با کارایی بالا است [26].

۳_۲_۴ مدل برنامه نویسی ویژه‌ی GPU

در سال‌های اخیر GP-GPU[4] به عنوان یک فناوری رو به رشد برای سرعت بخشیدن توسعه‌ی الگویتم‌های پردازش موازی می‌باشد. بسیاری از معماری‌ها و مدل‌های برنامه نویسی بر مبنای GPU جایگاه مهمی در رایانش با کارایی بالا یافته‌اند. معرفی CUDA در سال ۲۰۰۶ توسط NVIDIA تحول بزرگی در محاسبات گرافیکی به وجود آورد و منجر به شبیه سازی بسیاری از الگوریتم‌های پردازش موازی شد. در پاسخ به توسعه‌ی مدل‌های برنامه نویسی ویژه‌ی GPU تولیدکنندگان GPU اقدام به افزایش هسته‌ها و قابلیت‌های محاسباتی شتاب دهنده‌های خود نمودند. اگرچه CUDA تنها به GPUهای NVIDIA اختصاص دارد، اما OpenCL مجموعه‌ی متنوعی از معماری‌های GPU تولیدکنندگان مختلف را پشتیبانی می‌کند [27].

۳_۲_۴_۱ زبان محاسبه‌ی باز[5] OpenCL

OpenCL یک استاندارد متن باز و بدون حق امتیاز برای برنامه نویسی موازی بر روی شتاب دهنده‌های متنوع موجود در ابرکامپیوترها، سرورهای ابری است. OpenCL رابط برنامه نویسی

[1] Locks

[2] Semaphores

[3] Open Multi-Processing

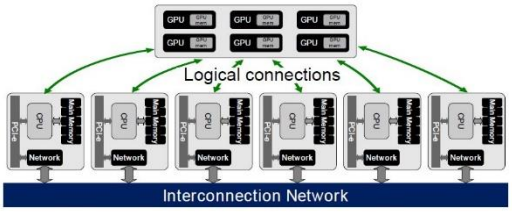
[4] General Purpose Graphical Processing Unit

[5] Open Computing Language

کاربردی سطح پایین برای اجرای محاسبات ناهمگون برروی GPU های قدرت گرفته از CUDA است. OpenCL به طور قابل توجهی سرعت و پاسخ دهی طیف گسترده ای از برنامه‌های کاربردی را در طبقه بندی های متعدد شامل نرم افزارهای علمی و پزشکی، پردازش بینایی و آموزش و استنتاج شبکه‌های عصبی را بهبود می‌بخشد. تاریخ آخرین نسخه‌ی انتشار یافته‌ی آن OpenCL3.0.6 در تاریخ ۳۰ سپتامبر ۲۰۲۰ است [28].

rCUDA ۲_۴_۲_۳

rCUDA و یا remoteCUDA یک فریم ورک نرم افزاری است که استفاده از سیستم‌های راه دور سازگار با CUDA را امکان پذیر می‌کند. به منظور امکان‌پذیر کردن کار با GPU از راه دور، این فریم ورک یک ابزار سازگار با CUDA بر روی ماشین‌های بدون GPU محلی ایجاد می‌کند. این ابزارهای مجازی، GPU فیزیکی که در یک میزبان از راه دور قرار دارند ارائه می‌کند. با استفاده از این تکنیک مجازی سازی GPU از راه دور rCUDA اجازه می‌دهد از هر گره در یک خوشه قابل دسترسی باشد. شکل ۲ نمای منطقی از عملکرد rCUDA را نشان می‌دهد [29].



شکل ۲. نمای منطقی از نحوه‌ی به کارگیری rCUDA در خوشه‌ها [29]

SYCL ۳_۴_۲_۳

یک مدل برنامه نویسی سطح بالا جهت پیاده سازی بر روی شتاب دهنده های سخت افزاری مختلف است. آخرین نسخه‌ی آن ۹ فوریه‌ی ۲۰۲۱ منتشر شده‌است. SYCL با استفاده از C++ استاندارد، قابلیت کد نویسی برای پردازشگرهای ناهمگون را فراهم می‌کند. SYCL بر اساس استاندارد OpenCL نوشته شده است. SYCL می‌تواند در پردازشگرهای متنوع و بسیاری از مسائل در بسیاری از دامنه‌ها شامل رایانش با کارایی بالا، خودکارسازی و یادگیری ماشین به کار رود. از مهم‌ترین ویژگی‌های SYCL می‌توان به اجرای موازی برروی سیستم‌های ناهمگون به صورت بسیار انعطاف پذیر اشاره کرد که امکان اولویت بندی الگوهای موازی را نیز فراهم می‌کند [30].

مدل برنامه نویسی مبتنی بر برنامه نویسی اسکلتی[1] ۵_۲_۳

برنامه نویسی اسکلتی دارای رویکرد طراحی بالا به پایین است. ساختار عمومی برنامه نویسی اسکلتی برمبنای map، reduce، farm scan، pipeline است که می‌تواند پیاده سازی-های موثری برای سیستم‌های کامپیوتری ناهمگون توزیع شده و موازی داشته باشد. برنامه نویسی اسکلتی دارای پیاده سازی‌هایی است که جزئیات سطح پایین مانند موازی سازی، همزمانی، برقراری ارتباط، مدیریت حافظه، استفاده از شتاب دهنده‌ها و دیگر بهینه سازی‌ها را در قسمتی جدا از دید کاربر قرار می‌دهد.

SkePU ۱_۵_۲_۳

یک رویکرد برنامه نویسی سطح بالا براساس الگوریتم‌های اسکلتی است. برنامه نویسی اسکلتی دارای پیاده سازی‌هایی است که جزئیات سطح پایین مانند موازی سازی، همزمانی، برقراری ارتباط، مدیریت حافظه، استفاده از شتاب دهنده ها و دیگر بهینه سازی‌ها را در قسمتی جدا از دید کاربر قرار می-دهد. SkePUیک فریم ورک منبع باز برمبنای برنامه‌نویسی اسکلتی است برای CPUهای چند هسته‌ای و سیستم‌های دارای چند GPU است که اجرا CUDA با OpenCL و اجرا بر روی خوشه ها را پشتیبانی می‌کند [31].

مدل برنامه نویسی مبتنی container ۶_۲_۳

containerها واحدهای استانداردی برای بسته بندی کردن برنامه‌های کاربردی و وابستگی‌های آن‌ها است به طوری که یک برنامه به طور قابل اطمینانی بتواند به سرعت در محیط های محاسباتی اجرا شود. container نرم افزار را از محیط خود جدا می‌کند و به طور یکنواخت اجرا می‌شود. به دلیل کاربردی بودن و اهمیت بالای آن‌ها مدل‌های برنامه نویسی مبتنی بر containerها طراحی و مطرح شده‌اند.

Boost.Compute ۱_۶_۲_۳

یک مدل برنامه نویسی براساس container است. containerها واحدهای استانداردی برای بسته بندی کردن برنامه‌های کاربردی و وابستگی‌های آن‌ها برای که یک برنامه هستند که به طور قابل اطمینانی بتواند به سرعت در محیط های محاسباتی اجرا شود. Boost.Compute یک کتابخانه‌ی C++ برای پردازش موازی براساس OpenCL است. دارای معماری لایه ای است و بر روی برنامه‌های کاربردی بر مبنای OpenCL عمل می‌کند. Boost.Compute امکان دسترسی به ابزارهای محاسباتی، صف‌های فرمان و بافرهای حافظه را می‌دهد [32].

مدیریت داده‌ها ۳_۳

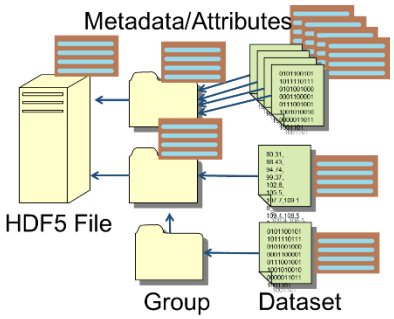
امروزه کشفیات علمی چه به صورت شهودی باشد و چه به صورت آزمایشگاهی باشد نیازمند تجزیه و تحلیل و غربال کردن داده‌های با حجم بالا و پیچیده است؛ به طور مثال شبیه سازی پلاسما بیلیون‌ها ذره را در یک اجرا شبیه سازی می‌کند، اما تجزیه و

[1] Skeleton programming

تحلیل نتایج آن نیاز به یک آرایه‌ی چند بعدی دارد که در موارد بیشتر از ۵۰ ترابایت بزرگی دارد و این تنها برای یک مرحله‌ی زمانی از شبیه سازی است. علاوه بر آن ابزارهای شبیه‌سازی داده‌های با حجم بالا تولید می‌کنند. یک تصویربرداری دو فوتون از مغز موش حدود بیشتر از ۱۰۰ گیگابایت داده در هر ساعت تولید می‌کند. [20] یکی از راهکارهای مدیریت این داده‌ها در نرم افزارهای حجم بالا استفاده از فایل‌های با فرمت مناسب است.

۳_۳_۱ فرمت فایل HDF5

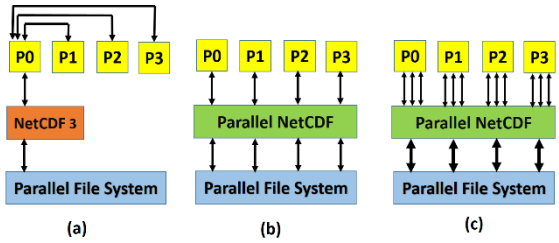
فرمت فایل HDF5 برای ذخیره و سازمان دهی داده‌های در حجم بالا استفاده می‌شود. HDF5 می‌تواند دو نوع از داده‌ها را ذخیره کند، مجموعه داده‌ها و گروه‌ها، که مجموعه داده یک آرایه‌ی چندبعدی از عناصر داده است و گروه ها ساختاری برای سامان دهی اشیا در فایل HDF5 است. با استفاده از این دو نوع می‌توان هر نوع داده‌ی علمی مانند تصاویر، آرایه‌ها، بردارها را ذخیره کرد که می‌توانند براساس نیازهای کاربر با یکدیگر ترکیب شوند. HDF5 تعداد فایل‌های موجود و تعداد اشیای موجود دریک فایل را محدود نمی‌کند. داده‌های ذخیره شده در آرایه های چندبعدی هرکدام می‌توانند داده‌های پیچیده باشند. HDF5 نیازی به بارگزاری تمام داده به طور یکجا ندارد و می‌توان هر قسمتی از فایل را که مورد نیاز باشد به آن دسترسی پیدا کرد [33]. در بسیاری از پروژه‌های ECP فرمت فایل HDF5 به کار می‌رود. از رویکردهای آینده‌ی فرمت فایل امکان دسترسی به داده‌های HDF5 بر روی سیستم‌های از راه دور برای خواندن است در حالیکه داده بروری فایل محلی قرار دارد؛ همچنین مدیریت داده‌ی خلوت در HDF5 بدون تغییر در برنامه‌های کاربردی موجود که در حال حاضر در مرحله‌ی طراحی است.



شکل ۳. ساختار فرمت فایل HDF5 [34]

۳_۳_۲ فرمت فایل Parallel NetCDF

یک کتابخانه‌ی ورودی، خروجی موازی برای دسترسی به فایل‌های Net CDF به شکل CDF-1,2 و ۵ است. فرمت فایل NetCDF انواع داده‌های بسیاری را پشتیبانی می‌کند و به کاربر اجازه می‌دهد اعداد صحیح ۶۴ بیتی را برای تعریف داده های با ابعاد بالا استفاده کند. هدف فراهم کردن موازی سازی برای ورودی و خروجی برنامه‌ها با عملکرد بالا است به طوری که تمام پردازش‌ها بتوانند به طور موازی به فایل‌های به اشتراک گذاشته شده دسترسی پیدا کنند. [35] شکل ۴ دسترسی به فایل در حالت NetCDF متوالی و ParallelNetCDF را نشان می‌دهد.



شکل ۴. مقایسه‌ی دسترسی به فایل در حالت NetCDF متوالی و Parallel NetCDF [35]

۳_۴ ارزیابی کارایی معماری‌های سخت افزاری و نرم افزاری

به منظور درک نحوه‌ی عملکرد و کارایی و میزان مصرف انرژی وابسته به منابع تخصیص یافته به یک کار، درک کارایی مدل‌ها ضروری است. منابع شامل تعداد گره‌ها، سیستم حافظه و شتاب دهنده‌های کامپیوتر هستند. مدل‌های ارزیابی ممکن به طور مستقیم توسط کاربر و یا به صورت اتوماتیک استفاده شوند. چالش‌های بسیاری از جمله اندازه‌ی سیستم، پیچیدگی سیستم (از نظر نوع حافظه و شتاب دهنده‌ها و ناهمگنی) و پیچیدگی نرم افزاری(زمان اجرا و قدرت کامپایلر)، لایه‌های برنامه، قابلیت انعطاف، فضای ذخیره سازی و جریان کاری ورودی و خروجی پیش روی مدل‌های ارزیابی هستند. مدل-های تحلیلی به طور معمول از سرعت اجرا در ارزیابی استفاده می‌کنند و ار آنجاییکه اندازه گیری و ارزیابی عملکرد آن‌ها بر روی سیستم‌های واقعی است در جایی که رفتار دقیق سیستم ناشناخته است عملکرد بهتری دارند. در ادامه مدل‌های ارزیابی مورد کاربرد رو به رشد بر اساس [20] بررسی شده-اند.

۳_۴_۱ مدل سلسله مراتبی MUSA

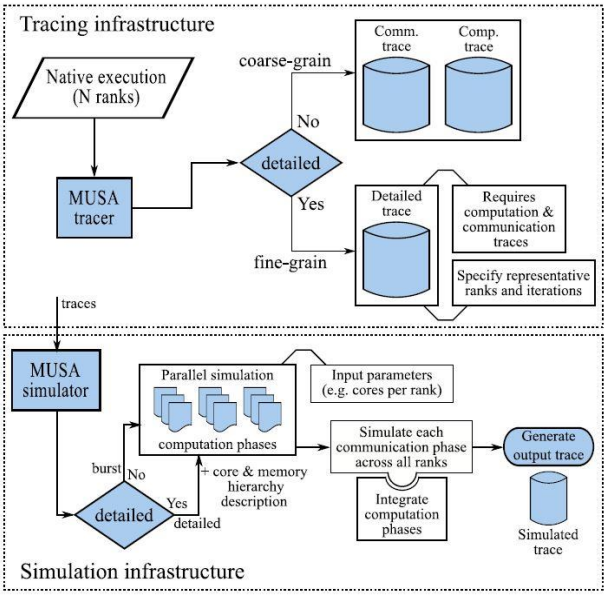
این مدل به صورت سلسله مراتبی بر اساس نمونه سازی از سطوح مختلف به صورت جزئی است که برای آینده‌ی سیستم‌های مقیاس اگزا مناسب است. MUSA یک رویکرد شبیه‌سازی چند سطحی است که تخمین وسیع و صحیحی از عملکرد سیستم‌های نسل آینده‌ی رایانش با کارایی بالا می‌زند. MUSA شبیه‌سازی در مقیاس بالا را امکان پذیر می‌کند. به این منظور MUSA از دو عنصر استفاد می‌کند: ردیابی زیرساخت-های ارتباطات، محاسبات و زمان اجراهای مختلف سیستم و یک زیرساخت شبیه سازی که این نتایج را برای شبیه سازی چندین سطح به کار می‌برد [36].

۳_۴_۱_۱ ردیابی

گام اول ردیابی اجرای برنامه‌های کاربردی در چند سطح است. ابتدا ردیابی هر فرآیند MPI با یک rank انجام می‌شود. درون یک rank چندین نخ در حال اجرای موازی ممکن است وجود داشته باشند که با زمان اجرای سیستم مدیریت می‌شوند. سپس دنبال کننده یک فایل شامل ارتباطات و محاسبات هر rank ایجاد می‌کند. این فایل شامل اطلاعاتی از ارتباطات MPI فازها است که شامل زمان شروع و پایان هر فاز ارتباطی برای همه‌ی rankها، نوع ارتباط (جمعی یا نقطه به نقطه) و اندازه-ی داده‌ی ارسال شده است. در همان زمانی که اطلاعات هر rank ذخیره می‌شود، ذخیره‌ی بازه‌های زمانی هر فاز محاسباتی، و زمان اجرای وقایع مانند قسمت‌های ایجاد، همگام سازی و قسمت‌های موازی نیز انجام می‌شود. می‌توان یک تحلیل معماری از فضای طراحی بزرگ با استفاده از همان مجموعه از فایل‌های ردیابی انجام داد [36].

۳_۴_۱_۲ شبیه‌سازی

در مرحله‌ی شبیه سازی MUSA از اطلاعات ارتباطات و محاسباتی که در مرحله‌ی ردیابی جمع آوری شده است استفاده می‌شود. هر سطح از فازهای محاسباتی با تعداد مشخصی از هسته‌ها و متغیرهای معماری و سلسله مراتب حافظه شبیه سازی می‌شود. به این منظور زمان فراخوانی برنامه‌های کاربردی سیستم را که با استفاده از زمان اجرای رویدادهای سیستم که در فایل ردیابی ذخیره شده‌است به کار می‌برد و از این طریق زمان اجرای سیستم را که شامل برنامه‌ریزی و همگام سازی تعداد هسته‌های شبیه سازی شده است، شبیه سازی می‌کند. شبیه سازی معماری در دو حالت انفجاری و یا دقیق انجام می‌شود که امکان شبیه سازی سریع و یا شبیه سازی با جزئیات بیشتر را فراهم می‌کند. بنابراین MUSA به کاربر اجازه می‌دهد تا محدودیت‌های مورد نظر برای متغیرهای ورودی را تعریف کند، انعطاف پذیری بالایی در راستای انتخاب فاز محاسباتی‌ای که برای پیاده‌سازی دقیق مورد نظر است دارد و در کنار آن عملکرد باقی فازها را پیش‌بینی می‌کند [36].



شکل ۵. نمای کلی ردیابی و شبیه سازی MUSA [36]

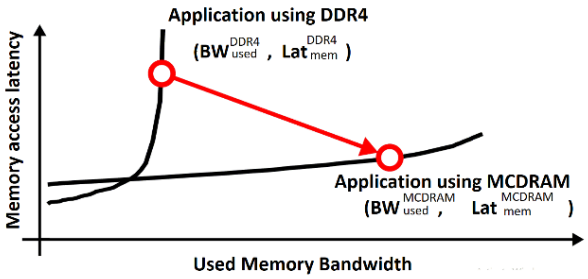
۳_۴_۱_۳ خروجی نهایی

پس از اینکه فازهای محاسباتی شبیه سازی شدند، MUSA اجرای ردیابی ارتباطات را به منظور شبیه سازی ارتباطات شبکه و تولید خروجی نهایی شبیه سازی تکرار می‌کند. در طول این فرآیند مدت زمان فازهای محاسباتی با نتایج به دست آمده از شبیه سازی‌ها (چه در حالت شبیه سازی انفجاری و چه در حالت شبیه سازی دقیق) جایگزین می‌شود و فازهای ارتباطی با استفاده از شبیه سازهای شبکه شبیه-سازی می‌شود. در پایان این فرآیند شبیه سازی به طور کامل انجام می‌شود و خروجی به منظور بصری سازی تولید می-شود. شکل ۵ نمای کلی از ردیابی شبیه سازی MUSA را نشان می‌دهد [36].

۳_۴_۲ مدل PROFET

یک مدل تحلیلی است که پیش‌بینی می‌کند چگونه عملکرد یک برنامه‌ی کاربردی و مصرف نیرو و انرژی آن با اجرا بر روی سیستم حافظه‌ی جدید تغییر می‌کند. این مدل از پارامترهای گرفته شده از هر قسمت اجرای برنامه‌ی کاربردی که در بازه های زمانی منظم تقسیم شده است، استفاده می‌کند. برای هر قسمت از شمارنده‌های اندازه گیری عملکرد یعنی تعداد چرخه‌ها، تعداد دستورالعمل‌ها و خواندن خطاهای حافظه‌ی پنهان سطح آخر، پهنای باند استفاده شده‌ی حافظه استفاده می‌شود. خروجی PROFET میزان عملکرد، قدرت و مصرف انرژی بر روی سیستم حافظه‌ی هدف است. مدل‌های قدرت و حافظه‌ی برای هر قسمت (هر بازه‌ی زمانی) اجرا می‌شوند، که برای هر قسمت یک زمان اجرا و قدرت و میزان مصرف انرژی را پیش‌بینی می‌کند. با جمع بستن با گذشت زمان، زمان کلی اجرا و مصرف انرژی برنامه به دست می‌آید. میانگین قدرت تقاضای برنامه برابر با حاصل تقسیم میزان کلی انرژی بر زمان کلی اجرا است. مدل عملکرد PROFET اطلاعات اجرای

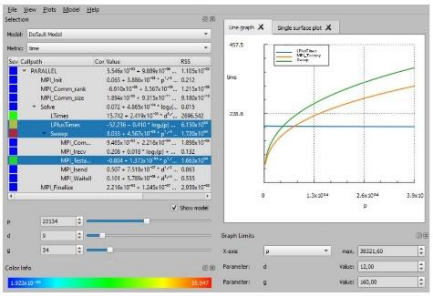
برنامه را از روی فایل ردیابی می‌خواند و موقعیت برنامه را برروی منحنی پهنای باند-تاخیر سیستم حافظه‌ی پایه مشخص می‌کند. سپس موقعیت برنامه‌ی کاربردی را بر روی منحنی پهنای باند حافظه-تاخیر برای سیستم حافظه‌ی هدف تخمین می‌زند و از آن برای پیش‌بینی عملکرد برنامه برروی سیستم حافظه‌ی هدف استفاده می‌کند [37]. شکل ۶ نحوه‌ی پیش‌بینی مدل PROFET برای سیستم حافظه‌ی هدف را نمایش می‌دهد.



شکل ۶. تخمین PROFET برای سیستم حافظه‌ی هدف [37]

۳_۴_۳ مدل Extra-P

Extar-P یک مدل تحلیلی برای شناسایی Scalability bugs است. Scalability bugs خطاهایی هستند که در محسابات در مقیاس بالا به وجود می‌آیند. Extra-P از اندازه گیری معیارهای عملکرد متنوع در قالب‌های اجرایی متفاوت به عنوان ورودی برای تولید مدل‌های عملکرد از قسمت‌های مختلف کد استفاده می‌کند. Extar-P با جستجوی Scalability bugs در کدهای گسترده و استفاده از اجرای تعدادی از آزمایش‌های عملکرد در مقیاس کوچک، و مقایسه‌ی تقریبی یا دقیق عملکرد بدترین نمونه با انتظارات عمل می‌کند. Extra-P علاوه بر لیست Scalability bugs ، مدل‌های قابل خواندن برای تمام متغیرهای عملکرد مانند عملیات ممیز شناور و یا بایت‌های ارسال شده با فراخوانی‌های MPI ایجاد می‌کند که می‌تواند برای تحلیل‌های بعدی و یا مقایسه برای شناسایی دلایل ریشه‌ای Scalability bugs استفاده شود [38]. شکل ۷ نمای خروجی مدل Extra-P براساس چند نمونه از متغیرهای عملکرد را نمایش می‌دهد.



شکل ۷. نمای خروجی مدل Extra-P [38]

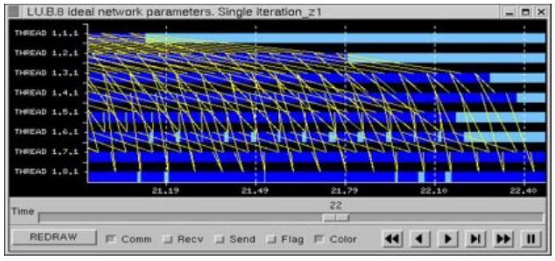
۳_۴_۴ مدل Dimemas

این مدل کاربر را قادر می‌کند تا برنامه‌های کاربردی موازی را بر روی یک ایستگاه کاری توسعه دهد و همزمان یک پیش‌بینی دقیق از عملکرد آن‌ها برروی ماشین موازی هدف داشته باشد. شبیه ساز Dimemas رفتار زمانی برنامه‌های کاربردی مدل شده برروی یک ماشین را با مجموعه‌ای از متغیرهای عملکرد بازسازی می‌کند.

مدل Dimemas از گره‌های SMP ساخته شده است. هر گره مجموعه‌ای از پردازشگرها و حافظه‌ی محلی دارد که برای ارتباطات درون گره‌ها استفاده می‌شود. ارتباطات درون شبکه-ای با دو متغیر نمایش داده می‌شود. تعداد پیوندها از گره به شبکه که با L نمایش داده شده است و تعداد باس‌های موجود شبکه که با B نمایش داده می‌شود. این متغیرها ظرفیت شبکه را محدود می‌کنند به طوری که تا تعداد B پیام به طور همزمان می‌توانند از شبکه استفاده کنند که اجازه‌ی تجزیه و تحلیل بحث شبکه را می‌دهد. متغیر L تعداد پیام‌هایی که به یک گره مشخص وارد و خارج می‌شوند را مشخص می‌کند، بنابراین تجزیه و تحلیل اتصال نیز می‌تواند مشخص شود.

رکوردهای فایل ردیابی شامل سه قسمت اصلی هستند: نقطه‌ی پایان ارتباط که اطلاعات مرتبط با وظایفی که در اتباط دخیل هستند را (مانند اندازه و شناسه‌های پیام) ذخیره می‌کند. اطلاعات رویدادها که این رکورد می‌تواند شامل هر نوعی از اطلاعات به طور مثال شروع و آغاز توابع، مقدار متغیرها، مقدار ثبات‌های درونی پردازشگرها باشد. مصرف CPU که شامل زمان گذرانده‌ی پردازشگر برای ارتباطات متوالی و رویدادها می‌باشد.

نتایج Dimemas شامل اطلاعات سراسری برنامه یعنی زمان اجرا و سرعت می‌باشد. برای هر وظیفه زمان محاسبه، زمان مسدود کردن، و اطلاعات مربوط به پیام‌ها مانند تعداد ارسال‌ها، تعداد دریافت‌ها را نیز شامل می‌شود [39]. شکل ۸ نمای خروجی مدل Dimemas از تحلیل عملکرد برنامه را نشان می‌دهد.

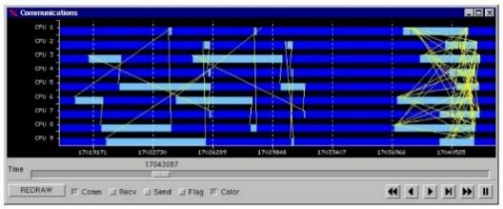


شکل ۸. نمای خروجی مدل Dimemas از تحلیل عملکرد برنامه [39]

۳_۴_۵ مدل Paraver

یک ابزار تحلیل عملکرد انعطاف پذیر است.Paraver در نیاز به درک کیفی از رفتار برنامه با تحلیل بصری و توانایی تحلیل کمّی دقیق از مسائل توسعه یافته است. قدرت بیان، انعطاف پذیری و قابلیت کارآمد کنترل ردیابی‌های گسترده از ویژگی‌های کلیدی در طراحی Paraver هستند. ویژگی‌های اساسی Paraver شامل تحلیل کمّی دقیق از عملکرد برنامه، تحلیل مقایسه‌ای همزمان از چندین ردیابی و مفاهیم قابل تنظیم اطلاعات بصری می‌باشد.

اطلاعات نمایش داده شده در خروجی Paraver شامل سه عنصر اصلی هستند که متغیرهای وابسته به زمان برای هر شئ ارائه شده. flagهایی که مرتبط با وقایع زمانی شئ ارائه شده هستند، و خطوط ارتباطی که اشیای نمایش داده شده را به یکدیگر مرتبط می‌کند. ماژول-های بصری مشخص می‌کنند که هرکدام از این عناصر چگونه نمایش داده شوند. ماژول بصری مقادیر و رویدادهایی که به آن داده می‌شود بدون مشخص کردن هیچ معنای پیش تعریف شده ای ارائه می‌شوند که این امر به منظور انعطاف پذیری ابزار است که بسته به ردیابی‌های تولید شده می‌توان آن‌ها را در یک مقیاس بزرگ بصری با یکدیگر ترکیب کرد [40]. شکل ۹ نمای خروجی مدل Paraver و ارتباطات بین پردازش‌ها را نمایش می‌دهد.

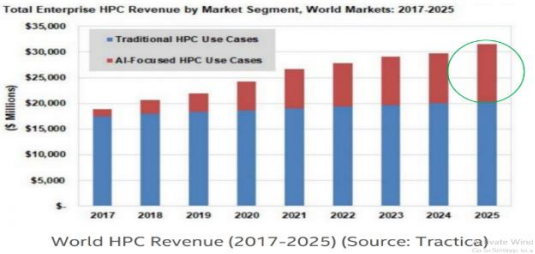


شکل ۹. نمای خروجی مدل Paraver [40]

۴. نقش هوش مصنوعی در رایانش با کارایی بالا

۴-۱. مقدمه

اغلب مراجع بررسی کننده‌ی روند تحولات رایانش با کارایی بالا (مانند [41] و [42])، موضوعی تحت عنوان همگرایی رایانش با کارایی بالا با هوش مصنوعی و پردازش کلان‌داده را به عنوان گرایشی مهم در جهت‌دهی به آینده‌ی رایانش با کارایی بالا معرفی می‌کنند. بر این اساس، امکانات مبتنی بر سیستم‌های رایانش با کارایی بالا بستر بسیار کارآمدی برای تسریع محاسبات مورد نیاز در هوش مصنوعی است و در سمت مقابل به‌کارگیری الگوریتم‌های مبتنی بر هوش مصنوعی می‌تواند به عملکرد بهینه‌تر و هوشمندانه‌تر سیستم‌های رایانش با کارایی بالا کمک کند. حوزه‌ی تحلیل کلان‌داده نیز در هر دو اثرگذاری یادشده، حضور دارد. شکل ۱۰ نقش موارد کاربرد رایانش با کارایی بالا با حضور AI را در رشد اقتصاد رایانش با کارایی بالا در سال‌های پیش رو مهم و اثرگذار پیش‌بینی می‌کند و بر این اساس در کنار تداوم اهمیت کاربردهای بدون حضور AI مانند شبیه‌سازی‌های فیزیکی، کاربردهای مبتنی بر AI را می‌توان بازیگر اصلی در رشد اقتصاد رایانش با کارایی بالا در سال‌های پیش رو دانست. بر اساس مفهوم همگرایی یاد شده، در ادامه موضوع جهت‌دهی هوش مصنوعی به آینده‌ی رایانش با کارایی بالا نخست از منظر امکانات مبتنی بر رایانش با کارایی بالا برای توسعه‌ی هوش مصنوعی و سپس از دیدگاه امکانات مبتنی بر هوش مصنوعی برای توسعه‌ی رایانش با کارایی بالا مطالعه خواهد شد و نهایتاً برخی اثرگذاری‌ها و جهت‌دهی‌های هوش مصنوعی به توسعه‌ی سیستم‌های رایانش با کارایی بالا به صورت موردی بررسی خواهند شد.



شکل ۱۰. پیش‌بینی موسسه‌ی Tractica از تاثیر AI بر آینده‌ی بازار رایانش با کارایی بالا [43]

۴-۲. امکانات مبتنی بر رایانش با کارایی بالا برای توسعه‌ی هوش مصنوعی

محاسبات مورد نیاز در بسیاری از الگوریتم‌های یادگیری ماشین به خصوص روش‌های مبتنی بر شبکه‌های عصبی به عنوان ابزاری بسیار پرکاربرد در این حوزه، عمدتا دارای ماهیت یک دستور چند داده (SIMD[1]) هستند و این موضوع، زمینه‌ی اصلی حضور مؤثر و کارآمد رایانش با کارایی بالا به عنوان بستر اجرای این محاسبات است. از دیگر سو، خروجی شبیه‌سازی‌های پدیده‌های طبیعی مبتنی بر رایانش با کارایی بالا می‌تواند کلان‌داده‌ی آموزشی موردنیاز برای الگوریتم‌های یادگیری ماشین در برخی کاربردها را فراهم کند. به طور خاص کاربرد رایانش با کارایی بالا در هوش مصنوعی در مرحله‌ی آموزش مدل‌ها اهمیت بسیار بالایی دارد و بسیار متداول است که آموزش یک مدل مورد استفاده در سیستم‌های درون‌کار یا پردازنده‌ی نه چندان قوی مانند پردازنده‌ی تلفن همراه بر بستر یک ابر رایانه انجام شود. مقاله‌ی مروری [۴۵]با بررسی مقالات متعددی، حوزه‌های کاربردی برای رایانش با کارایی بالا در هوش مصنوعی را یافته و مطالعه کرده‌است. در ادامه، ۴ مورد از کاربردها اجمالاً معرفی خواهند شد.

۴-۲-۱. حمل و نقل هوشمند

هوشمندسازی عملکرد وسایل حمل و نقل عمومی و خصوصی و نیز پردازش هوشمند اطلاعات دریافتی از این وسایل در مراکز داده‌ی شهر هوشمند مفهوم مهمی است که به دلیل کاربردهای وسیعی چون افزایش امنیت تردد و کاهش ترافیک مورد توجه ویژه‌ای قرار دارد. نظر به حجم بالای داده‌های ارسالی از خودروها در این سیستم‌ها پردازش اطلاعات در حجم مورد نظر قطعاً نیازمند امکانات رایانش با کارایی بالا است. در پروژه‌ی MCity در دانشگاه میشیگان به عنوان نمونه‌ی جامعی از پیاده‌سازی شهر هوشمند، زیر پروژه‌های متعددی در حوزه‌ی حمل و نقل هوشمند ایجاد شده‌است که برای پیاده‌سازی آن‌ها با همکاری شرکت DELL ابرارایانه‌ی

[1] Single Instruction Multiple Data

Great Lakes با به‌کارگیری ۴۴ واحد پردازنده‌ی گرافیکی NVIDIA Tensrocore V100) اولین GPU دارای توان پردازشی ۱۰۰ TFLOPS) مورد استفاده قرار گرفته است.

۴-۲-۲. مراقبت از سلامت

در سال ۲۰۲۰ یک مدل مبتنی بر هوش مصنوعی برای تشخیص سرطان سینه با کمک تصاویر X-Ray به دقت برابر با نیروی انسانی متخصص رادیولوژی دست یافت. پیشرفت‌هایی از این دست نشان دهنده‌ی گسترش کاربرد هوش مصنوعی در پزشکی است که به لحاظ نیازمندی‌های محاسباتی، لزوم به‌کارگیری رایانش با کارایی بالا را ایجاب می‌کند. به عنوان مثال در حوزه‌ی مراقبت اختصاصی از سلامت، تجویز راه‌کارهای درمانی و دارویی متناسب با ویژگی‌های ژنتیکی بیمار و سوابق وی موضوع مهمی است که راه‌کارهای مبتنی بر ترکیب پردازش درون‌کار و رایانش ابری برای آن در حال معرفی هستند [44].

۴-۲-۳. GeoAI

به‌کارگیری ترکیبی رایانش با کارایی بالا و هوش مصنوعی در علم فضا به منظور درک، تحلیل و بصری‌سازی پدیده‌های طبیعی بر مبنای اطلاعات مکانی، راه‌حل‌های متنوعی پدید آورده‌است که تحت عنوان علم نوظهور GeoAI مطالعه می‌شوند. یک مسئله‌ی بسیار مهم در این حوزه تحت عنوان Exposure Modeling به مدل‌سازی ریاضی و در جهت تخمین یا پیش‌بینی اثر یک عامل آلوده‌کننده‌ی محیط زیست می‌پردازد که چالش‌ها و اهداف پیش روی این حوزه در [45] معرفی شده‌است.

۴-۲-۴. شبیه‌سازی عملکرد مغز در علوم اعصاب محاسباتی

شبیه‌سازی عملکرد میلیاردها نورون و تریلیاردها سیناپس در مغز موجودات زنده مسئله‌ی مهمی در علوم اعصاب محاسباتی و هوش مصنوعی است که به وضوح نیازمند به‌کارگیری امکانات رایانش با کارایی بالا است. به عنوان مثال در پروژه‌ی BlueBrain[1] با تمرکز شبیه‌سازی مغز جوندگان و تعمیم نتایج به شبیه‌سازی مغز انسان ابررایانه‌ی شرکت IBM به نام BluGene با توان پردازشی حدود ۱۰۰TFLOPS و فضای ذخیره‌سازی ۱۰۰ترابایتی با ۸۰۰۰ واحد پردازنده‌ی مرکزی به کار گرفته شده‌است.

۴-۳. امکانات مبتنی بر هوش مصنوعی برای توسعه‌ی رایانش با کارایی بالا

به‌کارگیری الگوریتم‌های هوش مصنوعی برای عملکرد بهینه و هوشمند سیستم‌های رایانش با کارایی بالا با توجه به حجم بالای وظایف محاسباتی و تعدد منابع در دسترس در سیستم‌های رایانش با کارایی بالا موضوعی مورد نیاز است که به لحاظ ایجاد توازن بین توان مصرف انرژی و کارایی محاسباتی می‌تواند نقش کلیدی‌ای در تحقق عصر Exascale داشته باشد. در ادامه برخی صورت‌مسئله‌های مطرح در این حوزه معرفی خواهند شد.

برنامه‌ریزی وظایف[2]: انتخاب وظیفه‌ی محاسباتی مناسب در زمان مناسب برای تخصیص یک منبع در سیستم‌های رایانش با کارایی بالا با اهدافی چون بیشینه‌سازی utilization منابع، کمینه‌سازی زمان انتظار کاربران یا اولویت‌بندی وظایف انجام می‌شود. در سال‌های پیشین روش‌های Heuristic چون FCFS[3] برای این منظور مورد به‌کار گرفته شده‌اند و در راه‌حل‌های جدید تر الگوریتم‌های بهینه‌سازی نیز به مسئله وارد شده‌اند. در پژوهش [46] در سال ۲۰۲۱ به‌کارگیری یادگیری تقویتی برای این موضوع پیشنهاد شده‌است که موفق شده‌است علاوه بر افزایش کارایی محاسباتی، پاسخ مناسبی به ماهیت پویا و تغییرات ناگهانی در اهداف برنامه‌ریزی وظایف ایجاد کند.

تنظیم بهینه‌ی پارامترها در سیستم: انتخاب هوشمندانه‌ی مقادیر برای پارامترهای سیستم رایانش با کارایی بالا با توجه به ویژگی‌های برنامه‌ی محاسباتی مورد نظر می‌تواند افزایش قابل توجه کارایی محاسباتی را در مقایسه با استفاده از مقادیر پیش فرض برای پارامترها نتیجه دهد. به عنوان مثال در [47] برای این منظور با استخراج ۷۸ ویژگی مربوط به برنامه محاسباتی و طراحی الگوریتم انتخاب ویژگی و اعمال الگوریتم کلاس‌بندی، مقادیر پیشنهادی برای پارامترهای تعداد گره‌های محاسباتی و اندازه‌ی نوار و تعداد نوارها (پاارمترهای مربطو به فایل سیستم Lustre)به سیستم ارائه می‌شوند.

تنظیم بهینه پارامترهای برنامه محاسباتی اگزا مقیاس مسئله‌ی مطرح دیگری است که به عنوان مثال در پژوهش [48] که منجر به معرفی نرم‌افزار GPTune شده‌است الگوریتم‌های مبتنی بر بهینه‌سازی بیزی برای انتخاب پارامترهای برنامه در راستای بهینه‌سازی کمیت‌هایی چون زمان اجرا و حافظه‌ی مصرفی به‌کار گرفته شده‌اند. این ابزار تا کنون برای مسائلی در جبر خطی و برخی مسائل علمی چون plasma fusion به‌کارگرفته شده‌است و به طور میانگین منجر به ۲/۸ برابر کاهش زمان اجرای برنامه روی سیستم‌های دارای تعداد گره‌های محاسباتی بالا در حدود ۲۰۴۸ گره شده‌است.

۴-۴. جهت‌دهی توسعه‌ی هوش مصنوعی به آینده‌ی طراحی‌های رایانش با کارایی بالا

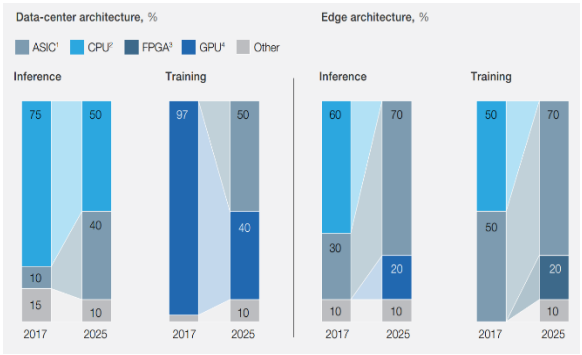
چنان‌که در بخش مربوط به معماری‌های سخت‌افزاری مطرح شد تعامل هوش مصنوعی و رایانش با کارایی بالا ، تاثیرات مشهودی بر طراحی‌های سیستم‌های رایانش با کارایی بالا گذاشته است که

[1] https://www.epfl.ch/research/domains/bluebrain/

[2] Task Scheduling

[3] First Come First Serve

معرفی معماری‌هایی چون TPU و NPU را می‌توان نتیجه‌ی این تأثیرات دانست. علاوه بر ایجاد معماری‌های جدید، تأثیر ملاحظات مربوط به توسعه‌ی هوش مصنوعی را به صورت موردی در برخی محصولات و فناوری‌های ایجاد شده در سال‌های اخیر و همچنین در آزمون‌های مقایسه‌ای ارزیابی کارایی محاسباتی سیستم‌های رایانش با کارایی بالا می‌توان دید. طبق شکل ۱۱ مطالعه‌ی آینده‌پژوهانه‌ی موسسه‌ی Mckinsey&company [49] پیش‌بینی می‌کند که طراحی‌های سخت‌افزاری مورد استفاده در HPC در ۴ حوزه‌ی ASIC, CPU, GPU, FPGA که در سال ۲۰۱۷ هر کدام با سهمی مختص به خود در بازار حضور دارند همگی در سال ۲۰۲۵ نیز حضور خواهند داشت اما سهم‌های فناوری‌ها تغییر خواهد کرد. بر این اساس افزایش چشمگیر نقش طراحی‌های اختصاصی ASIC در مرحله استنتاج در پردازش درون‌کار بسیار پررنگ خواهد بود و معماری CPU نیز همچنان در مراکز داده در مرحله‌ی استنتاج به‌کار گرفته خواهد شد.



شکل ۱۱. روند تغییر سهم فناوری های سخت افزاری در تعداد موارد استفاده در کاربردهای یادگیری عمیق [49]

جهت بررسی تاثیر هوش مصنوعی بر طراحی‌های رایانش با کارایی بالا می‌توان روند تحول برخی آزمون‌های مقایسه‌ای مشهور را به عنوان شاهدی برای این موضوع مطالعه کرد. محک Linpack در سال ۱۹۷۹ توسط جک دانگارا مطرح شد و تمرکز اصلی آن ارزیابی کارایی سیستم رایانشی در انجام محاسبات ماتریسی ممیز شناور بود. با معرفی این محک معیار FLOPS مطرح شد. در سال ۲۰۱۰ محک Graph500، از نظر تمرکز اصلی به سمت وظایف Data Intensive متمایل شده‌است و با تمرکز بر ارزیابی توانمندی یک سیستم در پیمایش یال‌ها در اجرای یک جستجوی BFS[1]، معیار GTEPS[2] معرفی شده‌است. در سال ۲۰۱۸ معیار MLPerf مشخصاً روی ارزیابی کارایی سیستم در اجرای الگوریتم‌های شبکه‌های عصبی عمیق بنا شده و در نسخه‌ی ۷ دو مجموعه داده به اندازه‌های ۸/۸ترابایت و ۵/۱ ترابایت برای انجام پیش‌بینی در پدیده‌های آب و هوایی در آن ایجاد شده‌است.

ردپای ملاحظات مربوط به توسعه‌ی هوش مصنوعی را می‌توان در طراحی سریع‌ترین ابررایانه‌ی جهان در زمان نگارش مقاله‌ی حاضر مشاهده کرد. در طراحی واحد محاسبات و منطق ابر رایانه‌ی Fugaku به این مفهوم توجه شده‌است که برای محاسبات مربوط به شبکه‌های عصبی دقت محاسبات ممیز شناور بیش از ۱۶ بیت مورد نیاز نیست (مقاله‌ی [50] از شرکت Intel کفایت دقت ۱۶ بیتی برای آموزش و ۸ بیتی برای استنتاج را نشان می‌دهد). لذا در طراحی Fugaku دقت ساده یا ۱۶ بیتی مورد استفاده قرار گرفته‌است. همچنین طراحی انحصاری شبکه‌ی Tofu نیز برای بهبود عملکرد سیستم در پیاده‌سازی شبکه‌های عصبی کانولوشنی صورت گرفته‌است. در فوریه‌ی سال ۲۰۲۱ ابر رایانه‌ی مذکور برای تولید کلان‌داده‌ی آموزشی از طریق شبیه‌سازی و نیز تولید و آموزش مدل پیش‌بینی کننده‌ی تسونامی مورد استفاده قرار گرفت.

به طور خلاصه، می‌توان نتیجه گرفت که تعامل دو سویه‌ی حوزه‌های رایانش فوق سریع و هوش مصنوعی، طراحان الگوریتم‌های هوش مصنوعی را به سمت مدل‌های با تعداد پارامترهای بیشتر و طراحان سیستم‌های رایانشی فوق سریع را به سمت طراحی سیستم‌های دارای تعداد گره‌های محاسباتی بیشتر هدایت کرده‌است.

۵. نقش رایانش ابری در رایانش با کارایی بالا

۵-۱. مقدمه

نظر به گسترش بسیار وسیع معماری‌های سخت‌افزاری و نرم‌افزاری مطرح در رایانش با کارایی بالا ، فناوری رایانش ابری به لحاظ امکان ایجاد استقلال تیم توسعه از پیچیدگی‌های این معماری‌ها نقش کلیدی و مهمی در این حوزه دارند. رایانش ابری در دو سناریوی All Cloud و Cloud Bursting می‌تواند بستر اجرای محاسبات رایانش با کارایی بالا باشد که در سناریوی All Cloud تمامی پیکربندی سیستم رایانش با کارایی بالا شامل ایجاد گره‌های محاسباتی و ذخیره‌سازی در سمت بستر ابری و توسط سازمان فراهم‌کننده‌ی رایانش ابری (CSP[3]) ایجاد شده و کاربر از طریق یک اتصال معمولی همچون SSH به ابر متصل شده و درخواست شروع محاسبه را ارسال می‌کند. در سناریوی Cloud Bursting، پیکربندی‌های یادشده تماماً در سمت کاربر ایجاد می‌شوند اما در زمان رسیدن حجم محاسبات به بیشینه‌ی توان سیستم، محاسبات مازاد به بستر ابری انتقال می‌یابند و گره‌های محاسباتی در ماشین‌های مجازی روی سرور ابری ایجاد می‌شوند. در این شرایط وجود یک ارتباط با تأخیر کم میان کاربر و بستر ابری و نیز فناوری‌های VM Orchestration اهمیت اساسی دارند.

[1] Breadth First Search

[2] Giga Traversed Edges per Second

[3] Cloud Service Provider

۵-۲. آینده‌ی رایانش با کارایی بالا بر بستر رایانش ابری

بر مبنای مطالعه‌ی آینده‌پژوهانه‌ی موسسه‌ی Hyperion [44] سهم بار محاسباتی مربوط به رایانش با کارایی بالا بر بستر رایانش ابری از سال ۲۰۲۰ تا ۲۰۲۳ از ۹/۸ به ۱۵ درصد خواهد رسید که نقطه‌ی عطفی در منحنی پیشرفت آن خواهد بود. در گزارش مذکور محرک‌های اصلی گرایش به رایانش ابری در رایانش با کارایی بالا معرفی شده‌اند که در ادامه اجمالاً مرور می‌شوند.

۱- خاصیت کشسانی[1]: در اغلب موارد توان پردازشی مورد نیاز در یک محاسبه در سازمان‌های کوچک بیش از توان محاسباتی در اختیار آن‌ها هست و از این رو خاصیت کشسانی در رایانش ابری را می‌توان اصلی‌ترین مزیت این فناوری در تعامل با رایانش با کارایی بالا دانست.

۲- تنوع نیازمندی‌های سخت‌افزاری و نرم‌افزاری: تنوع بسیار بالای معماری‌های شتاب‌دهنده در رایانش با کارایی بالا و نیز معرفی تعداد بسیار زیاد بسته‌های نرم‌افزاری مختص رایانش با کارایی بالا نیاز به استفاده از بسترهای رایانش ابری برای قرار گیری این معماری‌ها در دسترس طیف وسیع کاربران را تقویت می‌کند.

۳- کاهش هزینه از طریق کاهش زمان انتظار: در اغلب موارد به ویژه در سازمان‌های کوچک، کاهش زمان انتظار اجرای وظایف محاسباتی از چند روز به چند ساعت منجر به کاهش هزینه‌ها در سازمان می‌شود.

۴- اتصال بسیارقوی به داده‌های موجود در اینترنت: سرعت بسیار زیاد اتصال بسترهای رایانش ابری به اینترنت امکان بارگیری مجموعه‌داده‌های موجود در وب در زمان کوتاه را درکاربردهای مربوط به هوش مصنوعی فراهم می‌کند.

۵- Closely Coupled Hybrid Clouds: مفهومی نوظهور است که با گسترش فناوری G5 اهمیت روزافزونی می‌یابد. در فناوری‌های مرتبط با این موضوع پردازش‌های مربوط به رایانش با کارایی بالا در بستر ابری و پردازش درون‌کار به صورت همروند اجرا می‌شوند که مباحث مهمی چون حمل و نقل هوشمند و ارائه‌ی راه‌کارهای درمانی اختصاصی در پزشکی را تحت تأثیر جدی قرار می‌دهد.

در گزارش شرکت Hyperion پیش‌بینی شده‌است که فناوری Container در راستای تحقق HPC بر بستر رایانش ابری نیازمند تحولات مهمی خواهد بود که برای بررسی تحولات مورد انتظار این فناوری، گزارش شرکت Alibaba Clouds درباره‌ی گرایش‌های اصلی در تحولات مورد انتظار Kubernetes Container در سال ۲۰۲۱ [51] بررسی شد. بر اساس این گزارش، حائز اهمیت است که فناوری Kubernetes به امکانات خاص‌منظوره‌ی برنامه‌ریزی منابع مربوط به رایانش با کارایی بالا مجهز شود. در حال حاضر قابلیت‌های Capacity Scheduler و Batch Scheduler برای این منظور به‌کار گرفته می‌شوند. از دیگر سو در حال حاضر موضوع پشتیبانی Kubernetes از معماری GPU در GPUهای سری Ampere شرکت NVIDIA محقق شده‌است و قابلیت MIG[2] جهت استفاده‌ی همزمان از تعداد بالای GPU ایجاده شده‌است اما همچنان معماری‌های متنوع سخت‌افزاری دیگری چون NPU و RDMA[3] نیازمند پشتیبانی توسط این فناوری هستند. همچنین نظر به گسترش کاربردهای مربوط به پردازش داده‌های کلان، پشتیبانی از قالب‌های مختص Elastic Data Tasks مانند HDF, OSS, Ceph برای اعمال مختلفی چون جابه‌جایی، تکثیر، اخراج و تبدیل داده‌ها توسط Kebernetes نیاز حیاتی دیگری است که به گزارش مذکور از گرایش‌های اصلی مورد انتظار در سال ۲۰۲۱ است.

۵-۲-۱.فناوری‌ها و اهداف برخی کمپانی‌های پیش‌رو در رایانش با کارایی بالا بر بستر ابری

جهت مطالعه‌ی مسیر پیش رو در تکامل و توسعه‌ی رایانش با کارایی بالا بر بستر رایانش ابری، فعالیت‌های صورت گرفته و اهداف برخی کمپانی‌های مطرح در این حوزه بررسی می‌شود.

کمپانی Rescale به گزارش Yahoo Finance بزرگترین شرکت ارائه‌دهنده‌ی خدمات رایانش با کارایی بالا بر بستر Cloud است که با ایجاد یک API و رابط کاربری قوی، امکانات متعددی برای کاربران در سطوح شرکت‌های بزرگ تا محققین انفرادی جهت تعریف وظایف محاسباتی رایانش با کارایی بالا ایجاد کرده‌است. این کمپانی در گزارشی که در سال ۲۰۲۱ منتشر کرده‌است [52] رویکردهای اصلی این شرکت در این سال را اعلام کرده‌است. بر این اساس پیش‌بینی شده‌است که با توجه تخصصی شدن تمرکز شرکت‌های ارائه دهنده‌ی خدمات رایانش ابری در حوزه‌ی مختلف، تعداد نسبی کمپانی‌های استفاده کننده از رایانش با کارایی بالا که به صورت همزمان نیازمند دریافت خدمات از بیش از یک CSP هستند از ۳۵٪ در سال ۲۰۱۸ به ۵۰٪ در سال ۲۰۲۰ افزایش یافته‌است. لذا توسعه‌ی فناوری‌های Multi-Cloud از اصلی‌ترین اهداف کمپانی طبق گزارش مذکور است. شکل ۱۲ مقایسه‌ی یادشده را به طور دقیق‌تر نشان می‌دهد.

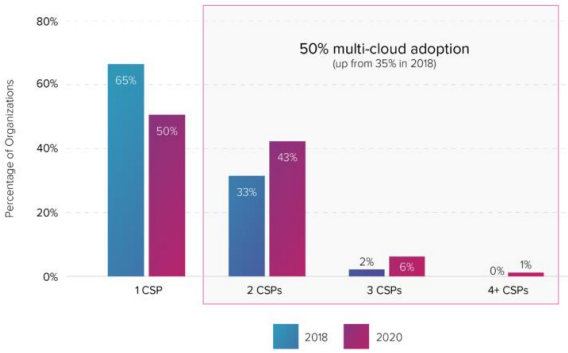
نکته‌ی شایان توجه دیگر در این گزارش، طرح موضوع تعدد بسیار بالای بسته‌های نرم‌افزاری و معماری‌های سخت‌افزاری معرفی شده برای رایانش با کارایی بالا است که شرکت‌های فعال در ارائه این

[1] Elasticity

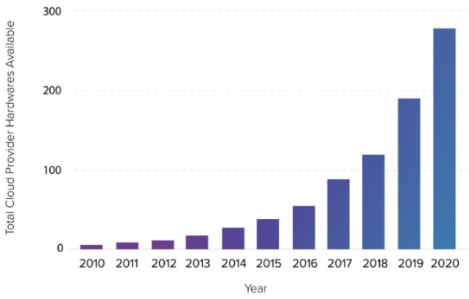
[2] Multi Instance GPU

[3] Remote Direct Memory Access

خدمات بر بستر رایانش ابری را نیازمند صرف انرژی قابل توجهی برای مدیریت و مهندسی دانش این گستره‌ی وسیع از امکانات می‌کند. تعداد بسته‌های نرم‌افزاری مورد پشتیبانی توسط کمپانی در زمان انتشار گزارش حدود ۶۰۰ بسته در حوزه‌های مختلف علوم و مهندسی گزارش شده‌است. همچنین شکل ۱۳ برگرفته از این گزارش افزایش چشمگیر تعداد معماری‌های اختصاصی رایانش با کارایی بالا در سال‌های اخیر را نشان می‌دهد.



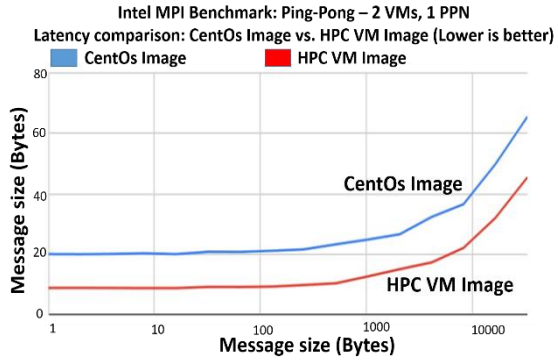
شکل ۱۲. تغییرات نسبت سازمان‌های استفاده کننده از خدمات بیش از یک CSP در رایانش با کارایی بالا [52]



شکل ۱۳. افزایش تعداد معماری‌های اختصاصی رایانش با کارایی بالا در دهه‌ی اخیر [52]

فناوری مطرح دیگر در Google Cloud Platform ایجاد شده‌است. کمپانی Google با ایجاد و افزودن بسترهای Google Genomics (برای پردازش کلان‌داده‌های زیست‌شناسی بر بستر ابری) و Google AI Platform (فراهم‌کننده‌ی طیف وسیعی از ابزارهای هوش مصنوعی) بستر ابری پرکاربردی مبتنی بر فناوری Multi-Cloud برای شرکت‌های بزرگ و کوچک ایجاد کرده‌است. در فوریه‌ی ۲۰۲۱ با معرفی قابلیت Pre-Tuned HPC VM images که یک ماشین مجازی بر پایه‌ی CentoOS7 است مدت‌زمان ارسال پیام در بارهای محاسباتی Tightly Coupled MPI کاهش مناسبی نسبت به دیگر فناوری‌های برپایه‌ی CentOS ایجاد کرده‌است [53].

شرکت Amazon با ایجاد قابلیت‌های Amazon SageMaker و Amazon Lake Formation و AWS Glue امکان استخراج مجموعه‌داده‌های آموزشی برای کاربردهای هوش مصنوعی از Amazon Simple Storage Service و نیز انجام تحلیل‌های مبتنی بر هوش مصنوعی در یک محیط Jupyter Notebook را ایجاد کرده‌است. کمپانی Microsoft با معرفی فناوری Azure Batch پشتیبانی رایانش با کارایی بالا بر بستر Cloud را در هر دو سناریوی All Cloud و Cloud Bursting را فراهم کرده و با توجه به نیازمندی‌هایی وظایف محاسباتی Tightly Coupled به فناوری‌های شبکه‌ای قوی مانند Infiniband یا دسترسی مستقیم از راه دو به حافظه (RDMA) حرکت به سمت پشتیبانی این فناوری‌ها را آغاز کرده‌است. تاکید اصلی در طراحی Azure Batch توجه به ماهیت غیر قابل پیش‌بینی بار محاسباتی در رایانش با کارایی بالا است که با توجه به مقادیر به دست‌آمده در روند محاسبات می‌تواند تغییرات چشمگیر داشته باشد لذا با ایجاد VM Pool سیستم تعداد VM ها را به صورت پویا و متناسب با بار محاسباتی کم و زیاد می‌کند.



شکل ۱۴. کاهش تاخیر MPI با استفاده از HPC VM Images [53]

به طور خلاصه، کمپانی‌های بزرگ به دلیل پشتیبانی طیف بسیار وسیع سخت‌افزارها و نرم‌افزارهای موجود در رایانش با کارایی بالا و نیز امکان دسترسی همزمان به قابلیت‌های چندین سرویس ابری و کمپانی‌های کوچک علاوه بر موارد یادشده به دلیل کاهش هزینه و زمان انتظار برای پیکربندی یک سیستم رایانش با کارایی بالا و همین‌طور کاهش چشمگیر زمان اجرای محاسبات بر بستر Cloud به استفاده از رایانش با کارایی بالا بر بستر رایانش ابری تمایل رو به رشدی دارند.

۶. موضوعات تحقیقاتی

۶-۱. مقدمه

به منظور مشخص کردن زمینه‌های تحقیقاتی رایانش با کارایی بالا که در حال حاضر در سطح جهان مورد توجه است، ابتدا لیستی از بازیگران فعال و مطرح در این حوزه شناسایی شده و سپس بررسی شده است که هریک از آنها در حال حاضر روی چه موضوعاتی در حال تحقیق هستند. در نهایت، با جمع‌بندی نتایج حاصل از این بررسی‌ها، به لیستی از زمینه‌های تحقیقاتی مطرح در حوزه رایانش با کارایی بالا دست یافته‌ایم. در ادامه این بخش، ابتدا لیست بازیگران مطرح در اکوسیستم رایانش با کارایی بالا را ارائه خواهیم

کرد و پس از آن، به تفکیک در خصوص زمینه‌های تحقیقاتی هریک اشاره خواهیم نمود. در نهایت، جمع‌بندی این اطلاعات در اختیار قرار خواهد گرفت.

۶-۲. لیست بازیگران مطرح در حوزه تحقیقات رایانش با کارایی بالا

پیدا کردن لیست کاملی از بازیگران فعال در حوزه رایانش با کارایی بالا چندان عملیاتی به نظر نمی‌رسد. اما بازیگران مطرح و شناخته شده‌تری در این حوزه وجود دارند که در ادامه به صورت مختصر معرفی می‌شوند:

• ISC-Group[1]: یکی از قدیمی‌ترین و معروف‌ترین گروه‌های فعال در حوزه رایانش با کارایی بالا است که در سال ۱۹۸۶ ایجاد شده است. انتشار لیست Top500 یکی از مهم‌ترین فعالیت‌های این گروه است.

• ECP: پروژه Exascale Computing Project[2] فعالیتی است که به صورت مشترک توسط وزارت انرژی[3] و اداره امنیت هسته‌ای ملی[4] ایالات متحده آمریکا اجرا می‌شود. هدف از این پروژه، حفظ برتری ایالات متحده در حوزه رایانش فوق سریع در جهان و حرکت به سمت توان پردازشی Exascale در این حوزه است.

• ETP4HPC[5]: این اتحادیه غیردولتی و غیرانتفاعی اروپایی عضوی از EuroHPC JU است که در سال ۲۰۱۲ تأسیس شده است. هدف‌گذاری آن افزایش توانمندی‌های فنی و تحقیقاتی اتحادیه اروپا در زمینه رایانش با کارایی بالا است به نحوی که بتواند بهره‌وری حداکثری اقتصادی و اجتماعی را از این فناوری برای کشورهای عضو این اتحادیه فراهم سازد.

• JLESC[6]: این آزمایشگاه سازمانی مجازی و بین‌المللی است که هدف‌گذاری آن گذار از توان پردازشی Petascale به Exascale برای اعضای آن است. مؤسسین این آزمایشگاه مؤسسه ملی تحقیقات در علم و فناوری دیجیتال (INRIA) فرانسه و دانشگاه ایلینویز هستند.

• HiPEAC[7]: شبکه‌ای از محققین دانشگاهی و صنعتی در سطح اروپا با بیش از ۲۰۰۰ عضو است که از سال ۲۰۰۴ فعالیت‌های خود را آغاز نموده است. هدف‌گذاری این شبکه پیشبرد معماری‌های سخت‌افزاری و سامانه‌های پردازشی با هم‌افزایی بین محققین از دانشگاه‌ها و مؤسسات تحقیقاتی مختلف در سطح اروپاست.

• BDEC[8]: با این ایده آغاز به کار کرده است که برنامه‌های جهانی در حوزه رایانش با کارایی بالا باید توجه ویژه‌ای به بحث Big Data داشته باشند. در حقیقت، اگرچه جهت‌گیری‌های جهانی به سمت رایانش Exascale است، اما با توجه به اهمیت موضوع Big Data، لازم است که در این جهت‌گیری‌ها، ساختارهای مورد نیاز داده‌های عظیم هم مورد توجه قرار گیرد.

• SPEC[9]: یک شرکت غیرانتفاعی است که تولید کننده ابزارها و Benchmarkهای استاندارد در حوزه ارزیابی کارایی و انرژی مصرفی سامانه‌های پردازش فوق سریع است.

• VI-HPS[10]: مؤسسه مجازی متشکل از ۱۴ مؤسسه تحقیقاتی در سراسر جهان است که هدف‌گذاری آن تولید نرم‌افزارهایی جهت کار با سوپرکامپیوترهاست.

• JARA-CSD[11]: شرکت Forschungszentrum Jülich و دانشگاه RWTH Aachen در آلمان با همکاری یکدیگر مرکز شبیه سازی و علم داده را پایه گذاری کرده‌اند که هدف‌گذاری آن در ۴ حوزه تحقیقاتی، آموزشی، توسعه زیرساخت‌ها و انجام شبیه‌سازی‌ها و فعالیت‌های مرتبط با علم داده برای متقاضیان تقسیم‌بندی شده است.

• OpenMP[12]: OpenMP ARB مالک استاندارد OpenMP در زمینه برنامه‌نویسی رایانش با کارایی بالا است که به صورت مداوم فعالیت توسعه و تکمیل این استاندارد را انجام می‌دهد. متشکل از بیش از ۳۰ دانشگاه، مؤسسه تحقیقاتی و شرکت‌های نرم‌افزاری و سخت‌افزاری است.

• Nvidia[13]: شرکت Nvidia مهم‌ترین و معروف‌ترین شرکت تولید کننده کارت‌های گرافیکی در سطح جهان است که با توجه به فراگیری استفاده از آنها در بحث رایانش با کارایی بالا ، اعتبار قابل توجهی را در این عرصه کسب نموده است. این شرکت در سال ۱۹۹۳ تأسیس گردیده است.

زمینه‌های تحقیقاتی بازیگران فوق‌الذکر مورد بررسی قرار گرفت و بر اساس آن، جدول ۲ تهیه گردید. اطلاعات در این جدول بر اساس تعداد بازیگرانی که روی یک موضوع تحقیقاتی توافق داشته‌اند مرتب شده است، به نحوی که ردیف‌های ابتدایی جدول نشان‌دهنده موضوعاتی است که بیشترین تعداد بازیگران روی آنها

[1] https://www.isc-group.com/home.html

[2] https://www.exascaleproject.org

[3] Department of Energy (DoE)

[4] National Nuclear Security Administration (NNSA)

[5] https://www.etp4hpc.eu

[6] https://jlesc.github.io

[7] https://www.hipeac.net

[8] https://www.spec.org/

[9] https://www.vi-hps.org/

[10] https://www.vi-hps.org/

[11] https://www.jara.org/en/research/center-for-simulation-and-data-sciences

[12] https://www.openmp.org/

[13] https://nvidia.com

در حال فعالیت هستند. بر این اساس، می‌توان بیان نمود که کار روی الگوریتم‌ها و روش‌های ریاضی مورد نیاز در برنامه‌هایی که باید روی بسترهای رایانش با کارایی بالا اجرا شوند، در بالاترین اولویت پژوهشی قرار دارد. همچنین، توسعه برنامه‌های کاربردی مورد نیاز سایر علوم هم، با توجه به آنکه عملاً نیاز بسیاری از علوم دیگر را در بر می‌گیرد، اولویت بالایی دارد. در زمینه پیشبرد اهداف رایانش با کارایی بالا ، دو موضوع معماری‌های سخت‌افزاری و مدل‌های برنامه‌نویسی موازی مهم‌ترین اولویت‌های تحقیقاتی و پژوهشی را شکل می‌دهند. مؤلفه‌های سخت‌افزاری و بهره‌وری انرژی هم موضوعات بعدی در این رتبه‌بندی هستند. بهره‌وری انرژی و رایانش سبز هم که با توجه به اهمیت بالای آن در دنیای امروزی قطعا اولویت بالایی دارد و لازم است که توجه ویژه‌ای نسبت به آن صورت پذیرد.

جدول ۲. موضوعات تحقیقاتی مورد توجه بازیگران مهم رایانش فوق سریع

موضوعات تحقیقاتی / بازیگران	ISC-Group	ECP	ETP4HPC	JLESC	HiPEAC	BDEC	SPEC	VI-HPS	JARA-CSD	OpenMP	Nvidia	مجموع
الگوریتم ها و روشهای ریاضی	✓		✓	✓	✓	✓				✓	✓	۷
برنامه های کاربردی	✓	✓	✓	✓	✓				✓			۶
معماری های سخت افزاری	✓		✓	✓	✓						✓	۵
مدل های برنامه نویسی موازی	✓	✓			✓	✓				✓		۵
مؤلفه های سخت افزاری	✓		✓	✓							✓	۴
بهره وری انرژی					✓	✓		✓	✓			۴
ابزارهای سنجش کارایی				✓			✓	✓				۳

به منظور ریزدانه‌تر کردن موضوعات تحقیقاتی فوق‌الذکر، مجموعه‌ای از مهم‌ترین کنفرانس‌های بین‌المللی در حوزه رایانش فوق‌سریع یا حوزه‌های نزدیک به آن، با جستجو در وب‌سایت http://www.wikicfp.com، انتخاب شده و موضوعات تحقیقاتی پیشنهادی آنها ذیل دسته‌بندی فوق تدوین گردید. نتیجه در جدول ۳ قابل مشاهده است.

جدول ۳. تدوین ریزدانه‌تر موضوعات تحقیقاتی با توجه به پیشنهادات کنفرانس‌های مرتبط که در سال ۲۰۲۱ برگزار شده‌اند.

دسته‌بندی موضوعات تحقیقاتی	موضوعات تحقیقاتی
الگوریتم‌ها و روش‌های ریاضی	• High-performance scientific and engineering computing • Numerical Solvers • Mixed Arithmetics • Mathematical, probabilistic and statistical models and theories • Large scale optimization
برنامه‌های کاربردی	• Computational science and scalable methods • Using HPC for scalable multi-scale, multi-physics, and high-fidelity computational science • Structured and unstructured meshes using extreme scale computing • Computational biology, earth sciences, cosmology, fluid dynamics, plasma modeling • High Performance Systems for Medical Applications • Machine Learning for Weather and Climate • Machine Learning in Life Sciences • Education, health, cost/energy-efficient design, smart cities, emerging markets, and interdisciplinary applications
معماری‌های سخت‌افزاری	• Parallel and distributed system architectures • Exascale Systems • Multicore, Many-core and Multithreaded Architectural Approaches • HW Heterogeneity • System-on-chip • Reconfigurable architectures • Node and system architecture for HPC and Big Data clusters • Energy-efficient cluster architectures • 3D-stacking
مدل‌های برنامه‌نویسی موازی	• Parallel and distributed software technologies • Parallel and distributed algorithms • Peer-to-peer computing • High performance Adaptive and Evolvable Computing • Grid and cluster computing • Web services and Internet computing • Cloud computing • Synchronization and concurrency control • High-level parallel programming and performance models (e.g. BSP, CGM, LogP, MPM, etc.) and too • Declarative parallel and distributed programming methodologies based on functional, logical, data-flow, actor, and other paradigms • Efficient code generation, auto-tuning and optimization for parallel and distributed programs • High-level parallel methods for large structured and semi-structured datasets • Novel task-based runtime environments • Hybrid programming techniques in applications and libraries (e.g., MPI+X) • Programming models for Big Data processing • Stream Processing • Networking, network function virtualization, software-defined networking • Memory policies and management • Approximate, analog, inexact, probabilistic computing • Large-scale databases • Simulation of Exascale Systems
مؤلفه‌های سخت‌افزاری	• RISC-V as Next-Generation HPC Components • Support and integration of non-volatile memory • GPUs, FPGAs and accelerator architectures • ASIC and FPGA Advances • ISA & processors • System and network interconnects • New Memory and Storage Technologies • Off-Chip/On-Chip Interconnection Networks • Interconnect/memory architectures • Chiplets • Accelerator-based architectures • Architectures and compiler techniques to accelerate deep neural networks, neuromorphic accelerators
بهره‌وری انرژی	• Programming language and compilation techniques for reducing energy and data movement • HPC for the Energy Transition • High performance Energy Aware Information Systems • Energy-efficient cluster architectures • Energy-efficient middleware • Power- and energy-management for clouds, datacenters, and exascale systems
ابزارهای سنجش کارایی	• Performance evaluation and measurement • Early hardware evaluations • Performance modeling for AI/ML/DL applications • Scalable tools and instrumentation infrastructure for measurement, monitoring, and/or visualization of performance • Performance Modeling and Tuning for Exascale Systems

- Profiling and performance monitoring of task-based environments
- Stochastic and non-deterministic models and metrics for performance, reliability, safety, correctness, and security
- New algorithms for the analysis or simulation of stochastic, probabilistic and non-deterministic models

رسیدن به لیستی از موضوعات تحقیقاتی مطرح، ۲۰ کنفرانس معتبری که در نیمه دوم سال ۲۰۲۱ در حوزه رایانش با کارایی بالا برگزار شده‌اند، انتخاب شده و فراخوان‌های مقالات آنها مورد بررسی قرار گرفت. بر این اساس، لیستی از موضوعات تحقیقاتی به دست آمد که ذیل محورهای فوق‌الذکر دسته‌بندی شدند.

۷. جمع‌بندی

در این مقاله، هدف بررسی الزامات و نیازمندی‌های سخت‌افزاری و نرم‌افزاری برای آینده رایانش با کارایی بالا ، با تمرکز بر فناوری‌های دست‌یافتنی در بازه زمانی کوتاه مدت بود. در حقیقت، اگرچه فناوری‌هایی نظیر پردازش کوآنتومی احتمالاً یکه تاز آینده رایانش با کارایی بالا در ۱۰ تا ۲۰ سال آینده خواهند بود، اما در ۳ تا ۵ سال آینده، به صورت جدی وارد عرصه سرویس‌دهی عمومی نمی‌شوند. از این رو، تمرکز بر توان پردازشی مقیاس اگزا بوده است. فناوری‌های سخت‌افزاری در این زمینه مورد بررسی قرار گرفتند و نشان داده شد که شتاب‌دهنده‌ها، شامل GPUها، TPUها و IPUها، سهم بسزایی در این زمینه دارند. همچنین، معماری‌های پردازنده‌ای خاص منظوره (ASIC) و FPGAها نیز نقش آفرینی بسیار جدی در بازار آینده رایانش فوق سریع خواهند داشت. فناوری ارتباطی ترابیت اترنت و ذخیره‌سازی در مقیاس اگزابایت با تکیه بر فایل سیستم Luster دیگر جنبه‌های سخت‌افزاری را تشکیل می‌دهند.

در حوزه نرم‌افزار، واسط برنامه‌نویسی MPI همچنان مهم‌ترین و کارآمدترین فناوری ارتباطی مابین گره‌های پردازشی خواهد بود. اما به منظور ساده‌تر شدن استفاده از آن، تنوع بالایی از ابزارها ارائه شده‌اند که روی MPI سوار شده و می‌توانند به برنامه‌نویسان این حوزه کمک کنند که با سهولت بیشتری به توسعه نیازمندی‌های خود بپردازند. OpenMP برای ارتباطات درون گره‌ای و OpenCL برای ارتباطات میان گره‌های پردازشی ناهمسان ارائه شده‌اند. فرمت فایل‌های HDF5 و Parallel NetCDF هم طبق آنچه که در پروژه‌های آینده‌پژوهشی کشورهای توسعه یافته مشاهده شده است، مهم‌ترین فرمت‌هایی هستند که ابزارها و نرم‌افزارهای حوزه رایانش با کارایی بالا از آنها استفاده خواهند نمود.

در نهایت، با بررسی فعالیت‌های بازیگران مطرح در حوزه تحقیقات رایانش با کارایی بالا ، لیستی از مهم‌ترین محورهای پژوهشی در این حوزه استخراج گردید. در این لیست، به ترتیب «ارائه الگوریتم‌ها و روش‌های ریاضی»، «توسعه برنامه‌های کاربردی مورد نیاز سایر علوم» ، «معماری‌های سخت‌افزاری»، «مدل‌های برنامه‌نویسی موازی»، «مؤلفه‌های سخت‌افزاری» و «بهره‌وری انرژی» جای گرفته‌اند. به منظور ریزدانه‌تر کردن این محورها و

مراجع

[1] "What is High-Performance Computing?," ed. https://www.netapp.com/data-storage/high-performance-computing/what-is-hpc/, 2021.

[2] "High Performance Computing with Accelerators," ed. https://www.seminarsonly.com/computer%20science/High-Performance-Computing-with-Accelerators.php, 2021.

[3] "HPC IN 2020: COMPUTE ENGINE DIVERSITY GETS REAL," ed. https://www.nextplatform.com/2020/01/13/hpc-in-2020-compute-engine-diversity-gets-real/, 2020.

[4] "HPC PROCESSORS – GREATER PERFORMANCE OVER TRADITIONAL CPUS," ed. https://www.aspsys.com/solutions/hpc-processors/, 2020.

[5] "What is GPU Computing?," ed. https://www.e4company.com/en/2021/02/what-is-gpu-computing/, 2021.

[6] "What Is a GPU?," ed. https://www.intel.com/content/www/us/en/products/docs/processors/what-is-a-gpu.html, 2020.

[7] "AMD Announces World's Fastest HPC Accelerator for Scientific Research¹," ed. https://www.amd.com/en/press-releases/2020-11-16-amd-announces-world-s-fastest-hpc-accelerator-for-scientific-research, 2020.

[8] "Re-Imagining Codesign," ed. https://www.orau.gov/ASCR-CoDesign/, 2021.

[9] "GRAPHCORE IPU-M-2000," ed. https://www.boston.co.uk/products/ipu-m2000.aspx, 2021.

[10] "An AI acceleration solution for production environments," ed. https://www.hpcwire.com/2021/02/22/massively-scalable-ai-that-seamlessly-works-with-your-existing-infrastructures-power-availability-and-budget/, 2021.

[11] L. Armasu, "Google's Big Chip Unveil For Machine Learning: Tensor Processing Unit With 10x Better Efficiency (Updated)," *Tom's Hardware*, 2016.

[12] Y. E. Wang, G.-Y. Wei, and D. Brooks, "Benchmarking tpu, gpu, and cpu platforms for deep learning," *arXiv preprint arXiv:1907.10701*, 2019.

[13] P. Teich, "Tearing apart Google's TPU 3.0 AI coprocessor," *Retrieved June*, vol. 12, p. 2018, 2018.

[14] N. P. Jouppi *et al.*, "A domain-specific supercomputer for training deep neural networks," *Communications of the ACM*, vol. 63, no. 7, pp. 67-78, 2020.

[15] P. G. A. Tekin.Tuncer Durak. Piechurski. Kaliszan. Aylin Sungur. Robertsén, "State-of-the-Art and Trends for Computing and Interconnect Network Solutions for HPC and AI," 2020.

[16] "2020 Roadmap," ed. https://ethernetalliance.org/technology/2020-roadmap/, 2020.

[17] "Terabit Ethernet: The New Hot Trend in Data Centers," ed. https://www.lanner-america.com/blog/terabit-ethernet-new-hot-trend-data-centers/, 2019.

simulating the CPU," *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, vol. 3, no. 2, pp. 1-33, 2019.

[38] "Extra-P." https://www.scalasca.org/scalasca/software/extra-p/download.html (accessed 2021).

[39] "Dimemas: predict parallel performance using a single cpu machine." BSC. https://tools.bsc.es/dimemas (accessed 2021).

[40] "Paraver: a flexible performance analysis tool." BSC. https://tools.bsc.es/paraver (accessed 2021).

[41] R. Farber, "AI-HPC is Happening Now," 5082598570, 2017.

[42] M. Savonin. "How to Leverage High Performance Computing (HPC) for AI." (accessed.

[43] V. Lee, "Progress in HPC + AI Convergence High performance Computing (HPC)."

[44] S. Conway, A. Norton, B. Sorensen, and E. Joseph, "Technology Spotlight The Future of HPC Cloud Computing Sponsored by Google," 2019.

[45] T. Vopham, J. E. Hart, F. Laden, and Y. Y. Chiang, "Emerging trends in geospatial artificial intelligence (geoAI): Potential applications for environmental epidemiology," *Environmental Health: A Global Access Science Source*, vol. 17, pp. 1-6, 2018, doi: 10.1186/s12940-018-0386-x.

[46] Y. Fan, Z. Lan, T. Childers, P. Rich, W. Allcock, and M. E. Papka, "Deep Reinforcement Agent for Scheduling in HPC," *arXiv preprint arXiv:2102.06243*, 2021.

[47] J. Bang *et al.*, "HPC Workload Characterization Using Feature Selection and Clustering," presented at the Proceedings of the 3rd International Workshop on Systems and Network Telemetry and Analytics, 2020.

[48] Y. Liu *et al.*, "GPTune: Multitask learning for autotuning exascale applications," presented at the Proceedings of the ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming, PPOPP, 2021.

[49] G. Batra, Z. Jacobson, S. Madhav, A. Queirolo, and N. Santhanam, "Artificial-intelligence hardware: New opportunities for semiconductor companies," *McKinsey & Company, New York, NY, USA, Tech. Rep*, 2018.

[50] A. Rodriguez. "Lowering Numerical Precision to Increase Deep Learning Performance." https://www.intel.com/content/www/us/en/artificial-intelligence/posts/lowering-numerical-precision-increase-deep-learning-performance.html (accessed.

[51] C. Y. A. C. Zhang Kai. "The Development Trends of Six Major Container Technologies in 2021." https://www.alibabacloud.com/blog/the-development-trends-of-six-major-containertechnologies-in-2021_597227 (accessed.

[52] E. Hsu, "Cloud Adoption is Driving HPC Toward Digital R&D," 2021. [Online]. Available: http://www.bigcompute.org/blog/state-of-cloud-hpc

[53] "Google Cloud Announces 'Pre-Tuned' HPC VM Images." https://www.hpcwire.com/off-the-wire/google-cloud-announces-pre-tuned-hpc-vmimages (accessed.

[18] "High-Performance Computing Storage," ed: https://dzone.com/articles/high-performance-computing-storage-hybrid-cloud-pa, 2020.

[19] "Introduction to Lustre," ed. https://wiki.lustre.org/Introduction_to_Lustre, 2017.

[20] D. P. C. Prof. Dr. Theo Ungerer, "Eurolab4HPC Long-Term Vision on High-Performance Computing (2nd Edition)," 2020.

[21] "Introduction to Parallel Computing Tutorial " Livermore Computing Center. https://hpc.llnl.gov/training/tutorials/introduction-parallel-computing-tutorial (accessed 2021).

[22] "MVAPICH: MPI over InfiniBand, Omni-Path, Ethernet/iWARP, and RoCE." Network-Based Computing Laboratory. https://mvapich.cse.ohio-state.edu/ (accessed.

[23] "OpenMP 5.0 API Reference Guide," 2019.

[24] "KOKKOS package." Sandia Corporation. https://lammps.sandia.gov/doc/Speed_kokkos.html (accessed 2021).

[25] C. Kessler and J. Keller, "Models for parallel computing: Review and perspectives," *Mitteilungen-Gesellschaft für Informatik eV, Parallel-Algorithmen und Rechnerstrukturen*, vol. 24, pp. 13-29, 2007.

[26] "The OmpSs Programming Model." 2021 Programming Models Group BSC. https://pm.bsc.es/ompss (accessed April, 2021).

[27] V. K. Pallipuram, M. Bhuiyan, and M. C. Smith, "A comparative study of GPU programming models and architectures using neural networks," *The Journal of Supercomputing*, vol. 61, no. 3, pp. 673-718, 2012.

[28] "The OpenCL Specification," 2020. [Online]. Available: https://www.khronos.org/registry/OpenCL/specs/3.0-unified/pdf/OpenCL_API.pdf

[29] "rCUDA v20.07 alpha User's Guide," 2020.

[30] T. K. S. W. Group, "SYCL 2020 Specification (revision 3)," 2020.

[31] "SkePU Autotunable Multi-Backend Skeleton Programming Framework for Multicore CPU and Multi-GPU Systems." https://skepu.github.io/ (accessed 2021).

[32] J. Szuppe, "Boost. Compute: A parallel computing library for C++ based on OpenCL," in *Proceedings of the 4th International Workshop on OpenCL*, 2016, pp. 1-39.

[33] "WHAT IS HDF5?" https://support.hdfgroup.org/HDF5/whatishdf5.html (accessed 2021).

[34] "HDF5, the new ICM+ data format." University of Cambridge. https://icmplus.neurosurg.cam.ac.uk/home/icm-features/hdf5-new-icm-data-format/ (accessed 2021).

[35] "PnetCDF." Center Of Ultra-Scale Computing and Information Security. http://cucis.ece.northwestern.edu/projects/PnetCDF/ (accessed 2021).

[36] T. Grass *et al.*, "MUSA: a multi-level simulation approach for next-generation HPC machines," in *SC'16: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*, 2016: IEEE, pp. 526-537.

[37] M. Radulovic, R. Sánchez Verdejo, P. Carpenter, P. Radojković, B. Jacob, and E. Ayguadé, "PROFET: modeling system performance and energy without