

Cognitive Powers — AI and Supercomputing Are Driving a New Era of Discovery and Innovation

Can machines think?

To someone using a pocket calculator, it certainly seems that way. The first one came on the market in the 1970s. That's not so long ago. The Beatles had only recently broken up, most cars drove with rear wheels, and Neil Armstrong had already taken his famous “small step” on the lunar surface. Anyone reading this who remembers a time before the pocket calculator knows that math problems or even anything beyond simple arithmetic required a sharp brain, and usually a sharp pencil. In that seemingly neolithic era, school children actually had to learn how to solve math problems—and even memorize multiplication tables.

The advent of the pocket calculator seemed like a small miracle. Suddenly we had palm-sized devices that even kids could afford yet seemed to have more brain power than humans—at least when it came to crunching numbers. However, anyone with a basic understanding of computer science knows machines like a pocket calculator, the first personal computers, or today's remarkable digital devices like Amazon's new domestic robot only appear to have cognitive powers. Their apparent intelligence is a mirage.

What actually happens at the level of software and chips in these smart machines is mind-numbingly simple. It just seems like intelligence because it happens at lightning speed. Let's have a look.

Unlike the Hindu-Arabic number system we learned as children, which uses ten numerals—the digits 0, 1, 2, 3, 4, 5, 6, 7, 8, 9—calculators and computers are based on the Binary number system. It uses only two digits, zero and one. While the inner workings of semiconductor chips and software that make digital machines work is incredibly complicated, they rely on a simple invention—the transistor. A transistor is an electrical machine that controls the flow of electrons through gates that are either open or closed. Think of a transistor gate like a light switch that can have only two states, On or Off.



Print enough microscopic transistors on a small chip of silicon and you have an integrated circuit (IC). Before an IC can become useful, however, the flow of electrons through the transistors' gates needs to be programmed in a pattern. Coders do this at chip hardware level with "machine language," which uses the Binary number system—using only 0 and 1—to turn the transistors' gates On or Off. A 1 in the code flips the switch on; a 0 flips it off. This programming can enable the chip to do things such as solving a simple problem like two plus two equals four.

Except in the case of very simple devices like pocket calculators, you need a vast number of transistors on a chip for much more than solving rudimentary arithmetic and math problems. Fortunately, it wasn't long before the chipmaking art had shrunk transistors to a point where you could have hundreds of thousands or even millions of transistors on a single piece of silicon the size of a postage stamp. An example of these massively integrated circuits is the central processing unit (CPU) that powered the first personal computers. And by creating programming languages that worked above the chip level, people could program these machines to do amazing things like running a spreadsheet application.

And today, as scientists have miniaturized transistors down to sizes measured in nanometers—billionths of a meter—we have massively integrated circuits that comprise whole computers on a chip. These system-on-a-chip designs (SoCs) power seemingly miraculous Lilliputian compute devices like your smartphone.

However, most computers and smartphones are not necessarily smart. It is the astonishingly fast rate of transistor gates opening and closing and the speed of electrons flowing over the silicon and along tiny wire interconnects that makes the remarkable abilities of digital devices seem like intelligence. In fact, without the help of humans they are dumb as a brick. You still need smart coders to program their problem-solving, picture-taking, or other useful abilities.

```
1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051 1052 1053 1054 1055 1056 1057 1058 1059 1060 1061 1062 1063 1064 1065 1066 1067 1068 1069 1070 1071 1072 1073 1074 1075 1076 1077 1078 1079 1080 1081 1082 1083 1084 1085 1086 1087 1088 1089 1090 1091 1092 1093 1094 1095 1096 1097 1098 1099 1100 1101 1102 1103 1104 1105 1106 1107 1108 1109 1110 1111 1112 1113 1114 1115 1116 1117 1118 1119 1120 1121 1122 1123 1124 1125 1126 1127 1128 1129 1130 1131 1132 1133 1134 1135 1136 1137 1138 1139 1140 1141 1142 1143 1144 1145 1146 1147 1148 1149 1150 1151 1152 1153 1154 1155 1156 1157 1158 1159 1160 1161 1162 1163 1164 1165 1166 1167 1168 1169 1170 1171 1172 1173 1174 1175 1176 1177 1178 1179 1180 1181 1182 1183 1184 1185 1186 1187 1188 1189 1190 1191 1192 1193 1194 1195 1196 1197 1198 1199 1200 1201 1202 1203 1204 1205 1206 1207 1208 1209 1210 1211 1212 1213 1214 1215 1216 1217 1218 1219 1220 1221 1222 1223 1224 1225 1226 1227 1228 1229 1230 1231 1232 1233 1234 1235 1236 1237 1238 1239 1240 1241 1242 1243 1244 1245 1246 1247 1248 1249 1250 1251 1252 1253 1254 1255 1256 1257 1258 1259 1260 1261 1262 1263 1264 1265 1266 1267 1268 1269 1270 1271 1272 1273 1274 1275 1276 1277 1278 1279 1280 1281 1282 1283 1284 1285 1286 1287 1288 1289 1290 1291 1292 1293 1294 1295 1296 1297 1298 1299 1300 1301 1302 1303 1304 1305 1306 1307 1308 1309 1310 1311 1312 1313 1314 1315 1316 1317 1318 1319 1320 1321 1322 1323 1324 1325 1326 1327 1328 1329 1330 1331 1332 1333 1334 1335 1336 1337 1338 1339 1340 1341 1342 1343 1344 1345 1346 1347 1348 1349 1350 1351 1352 1353 1354 1355 1356 1357 1358 1359 1360 1361 1362 1363 1364 1365 1366 1367 1368 1369 1370 1371 1372 1373 1374 1375 1376 1377 1378 1379 1380 1381 1382 1383 1384 1385 1386 1387 1388 1389 1390 1391 1392 1393 1394 1395 1396 1397 1398 1399 1400 1401 1402 1403 1404 1405 1406 1407 1408 1409 1410 1411 1412 1413 1414 1415 1416 1417 1418 1419 1420 1421 1422 1423 1424 1425 1426 1427 1428 1429 1430 1431 1432 1433 1434 1435 1436 1437 1438 1439 1440 1441 1442 1443 1444 1445 1446 1447 1448 1449 1450 1451 1452 1453 1454 1455 1456 1457 1458 1459 1460 1461 1462 1463 1464 1465 1466 1467 1468 1469 1470 1471 1472 1473 1474 1475 1476 1477 1478 1479 1480 1481 1482 1483 1484 1485 1486 1487 1488 1489 1490 1491 1492 1493 1494 1495 1496 1497 1498 1499 1500 1501 1502 1503 1504 1505 1506 1507 1508 1509 1510 1511 1512 1513 1514 1515 1516 1517 1518 1519 1520 1521 1522 1523 1524 1525 1526 1527 1528 1529 1530 1531 1532 1533 1534 1535 1536 1537 1538 1539 1540 1541 1542 1543 1544 1545 1546 1547 1548 1549 1550 1551 1552 1553 1554 1555 1556 1557 1558 1559 1560 1561 1562 1563 1564 1565 1566 1567 1568 1569 1570 1571 1572 1573 1574 1575 1576 1577 1578 1579 1580 1581 1582 1583 1584 1585 1586 1587 1588 1589 1590 1591 1592 1593 1594 1595 1596 1597 1598 1599 1600 1601 1602 1603 1604 1605 1606 1607 1608 1609 1610 1611 1612 1613 1614 1615 1616 1617 1618 1619 1620 1621 1622 1623 1624 1625 1626 1627 1628 1629 1630 1631 1632 1633 1634 1635 1636 1637 1638 1639 1640 1641 1642 1643 1644 1645 1646 1647 1648 1649 1650 1651 1652 1653 1654 1655 1656 1657 1658 1659 1660 1661 1662 1663 1664 1665 1666 1667 1668 1669 1670 1671 1672 1673 1674 1675 1676 1677 1678 1679 1680 1681 1682 1683 1684 1685 1686 1687 1688 1689 1690 1691 1692 1693 1694 1695 1696 1697 1698 1699 1700 1701 1702 1703 1704 1705 1706 1707 1708 1709 1710 1711 1712 1713 1714 1715 1716 1717 1718 1719 1720 1721 1722 1723 1724 1725 1726 1727 1728 1729 1730 1731 1732 1733 1734 1735 1736 1737 1738 1739 1740 1741 1742 1743 1744 1745 1746 1747 1748 1749 1750 1751 1752 1753 1754 1755 1756 1757 1758 1759 1760 1761 1762 1763 1764 1765 1766 1767 1768 1769 1770 1771 1772 1773 1774 1775 1776 1777 1778 1779 1780 1781 1782 1783 1784 1785 1786 1787 1788 1789 1790 1791 1792 1793 1794 1795 1796 1797 1798 1799 1800 1801 1802 1803 1804 1805 1806 1807 1808 1809 1810 1811 1812 1813 1814 1815 1816 1817 1818 1819 1820 1821 1822 1823 1824 1825 1826 1827 1828 1829 1830 1831 1832 1833 1834 1835 1836 1837 1838 1839 1840 1841 1842 1843 1844 1845 1846 1847 1848 1849 1850 1851 1852 1853 1854 1855 1856 1857 1858 1859 1860 1861 1862 1863 1864 1865 1866 1867 1868 1869 1870 1871 1872 1873 1874 1875 1876 1877 1878 1879 1880 1881 1882 1883 1884 1885 1886 1887 1888 1889 1890 1891 1892 1893 1894 1895 1896 1897 1898 1899 1900 1901 1902 1903 1904 1905 1906 1907 1908 1909 1910 1911 1912 1913 1914 1915 1916 1917 1918 1919 1920 1921 1922 1923 1924 1925 1926 1927 1928 1929 1930 1931 1932 1933 1934 1935 1936 1937 1938 1939 1940 1941 1942 1943 1944 1945 1946 1947 1948 1949 1950 1951 1952 1953 1954 1955 1956 1957 1958 1959 1960 1961 1962 1963 1964 1965 1966 1967 1968 1969 1970 1971 1972 1973 1974 1975 1976 1977 1978 1979 1980 1981 1982 1983 1984 1985 1986 1987 1988 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023 2024 2025 2026 2027 2028 2029 2030 2031 2032 2033 2034 2035 2036 2037 2038 2039 2040 2041 2042 2043 2044 2045 2046 2047 2048 2049 2050 2051 2052 2053 2054 2055 2056 2057 2058 2059 2060 2061 2062 2063 2064 2065 2066 2067 2068 2069 2070 2071 2072 2073 2074 2075 2076 2077 2078 2079 2080 2081 2082 2083 2084 2085 2086 2087 2088 2089 2090 2091 2092 2093 2094 2095 2096 2097 2098 2099 2100 2101 2102 2103 2104 2105 2106 2107 2108 2109 2110 2111 2112 2113 2114 2115 2116 2117 2118 2119 2120 2121 2122 2123 2124 2125 2126 2127 2128 2129 2130 2131 2132 2133 2134 2135 2136 2137 2138 2139 2140 2141 2142 2143 2144 2145 2146 2147 2148 2149 2150 2151 2152 2153 2154 2155 2156 2157 2158 2159 2160 2161 2162 2163 2164 2165 2166 2167 2168 2169 2170 2171 2172 2173 2174 2175 2176 2177 2178 2179 2180 2181 2182 2183 2184 2185 2186 2187 2188 2189 2190 2191 2192 2193 2194 2195 2196 2197 2198 2199 2200 2201 2202 2203 2204 2205 2206 2207 2208 2209 2210 2211 2212 2213 2214 2215 2216 2217 2218 2219 2220 2221 2222 2223 2224 2225 2226 2227 2228 2229 2230 2231 2232 2233 2234 2235 2236 2237 2238 2239 2240 2241 2242 2243 2244 2245 2246 2247 2248 2249 2250 2251 2252 2253 2254 2255 2256 2257 2258 2259 2260 2261 2262 2263 2264 2265 2266 2267 2268 2269 2270 2271 2272 2273 2274 2275 2276 2277 2278 2279 2280 2281 2282 2283 2284 2285 2286 2287 2288 2289 2290 2291 2292 2293 2294 2295 2296 2297 2298 2299 2300 2301 2302 2303 2304 2305 2306 2307 2308 2309 2310 2311 2312 2313 2314 2315 2316 2317 2318 2319 2320 2321 2322 2323 2324 2325 2326 2327 2328 2329 2330 2331 2332 2333 2334 2335 2336 2337 2338 2339 2340 2341 2342 2343 2344 2345 2346 2347 2348 2349 2350 2351 2352 2353 2354 2355 2356 2357 2358 2359 2360 2361 2362 2363 2364 2365 2366 2367 2368 2369 2370 2371 2372 2373 2374 2375 2376 2377 2378 2379 2380 2381 2382 2383 2384 2385 2386 2387 2388 2389 2390 2391 2392 2393 2394 2395 2396 2397 2398 2399 2400 2401 2402 2403 2404 2405 2406 2407 2408 2409 2410 2411 2412 2413 2414 2415 2416 2417 2418 2419 2420 2421 2422 2423 2424 2425 2426 2427 2428 2429 2430 2431 2432 2433 2434 2435 2436 2437 2438 2439 2440 2441 2442 2443 2444 2445 2446 2447 2448 2449 2450 2451 2452 2453 2454 2455 2456 2457 2458 2459 2460 2461 2462 2463 2464 2465 2466 2467 2468 2469 2470 2471 2472 2473 2474 2475 2476 2477 2478 2479 2480 2481 2482 2483 2484 2485 2486 2487 2488 2489 2490 2491 2492 2493 2494 2495 2496 2497 2498 2499 2500 2501 2502 2503 2504 2505 2506 2507 2508 2509 2510 2511 2512 2513 2514 2515 2516 2517 2518 2519 2520 2521 2522 2523 2524 2525 2526 2527 2528 2529 2530 2531 2532 2533 2534 2535 2536 2537 2538 2539 2540 2541 2542 2543 2544 2545 2546 2547 2548 2549 2550 2551 2552 2553 2554 2555 2556 2557 2558 2559 2560 2561 2562 2563 2564 2565 2566 2567 2568 2569 2570 2571 2572 2573 2574 2575 2576 2577 2578 2579 2580 2581 2582 2583 2584 2585 2586 2587 2588 2589 2590 2591 2592 2593 2594 2595 2596 2597 2598 2599 2600 2601 2602 2603 2604 2605 2606 2607 2608 2609 2610 2611 2612 2613 2614 2615 2616 2617 2618 2619 2620 2621 2622 2623 2624 2625 2626 2627 2628 2629 2630 2631 2632 2633 2634 2635 2636 2637 2638 2639 2640 2641 2642 2643 2644 2645 2646 2647 2648 2649 2650 2651 2652 2653 2654 2655 2656 2657 2658 2659 2660 2661 2662 2663 2664 2665 2666 2667 2668 2669 2670 2671 2672 2673 2674 2675 2676 2677 2678 2679 2680 2681 2682 2683 2684 2685 2686 2687 2688 2689 2690 2691 2692 2693 2694 2695 2696 2697 2698 2699 2700 2701 2702 2703 2704 2705 2706 2707 2708 2709 2710 2711 2712 2713 2714 2715 2716 2717 2718 2719 2720 2721 2722 2723 2724 2725 2726 2727 2728 2729 2730 2731 2732 2733 2734 2735 2736 2737 2738 2739 2740 2741 2742 2743 2744 2745 2746 2747 2748 2749 2750 2751 2752 2753 2754 2755 2756 2757 2758 2759 2760 2761 2762 2763 2764 2765 2766 2767 2768 2769 2770 2771 2772 2773 2774 2775 2776 2777 2778 2779 2780 2781 2782 2783 2784 2785 2786 2787 2788 2789 2790 2791 2792 2793 2794 2795 2796 2797 2798 2799 2800 2801 2802 2803 2804 2805 2806 2807 2808 2809 2810 2811 2812 2813 2814 2815 2816 2817 2818 2819 2820 2821 2822 2823 2824 2825 2826 2827 2828 2829 2830 2831 2832 2833 2834 2835 2836 2837 2838 2839 2840 2841 2842 2843 2844 2845 2846 2847 2848 2849 2850 2851 2852 2853 2854 2855 2856 2857 2858 2859 2860 2861 2862 2863 2864 2865 2866 2867 2868 2869 2870 2871 2872 2873 2874 2875 2876 2877 2878 2879 2880 2881 2882 2883 2884 2885 2886 2887 2888 2889 2890 2891 2892 2893 2894 2895 2896 2897 2898 2899 2900 2901 2902 2903 2904 2905 2906 2907 2908 2909 2910 2911 2912 2913 2914 2915 2916 2917 2918 2919 2920 2921 2922 2923 2924 2925 2926 2927 2928 2929 2930 2931 2932 2933 2934 2935 2936 2937 2938 2939 2940 2941 2942 2943 2944 2945 2946 2947 2948 2949 2950 2951 2952 2953 2954 2955 2956 2957 2958 2959 2960 2961 2962 2963 2964 2965 2966 2967 2968 2969 2970 2971 2972 2973 2974 2975 2976 2977 2978 2979 2980 2981 2982 2983 2984 2985 2986 2987 2988 2989 2990 2991 2992 2993 2994 2995 2996 2997 2998 2999 3000 3001 3002 3003 3004 3005 3006 3007 3008 3009 3010 3011 3012 3013 3014 3015 3016 3017 3018 3019 3020 3021 3022 3023 3024 3025 3026 3027 3028 3029 3030 3031 3032 3033 3034 3035 3036 3037 3038 3039 3040 3041 3042 3043 3044 3045 3046 3047 3048 3049 3050 3051 3052 3053 3054 3055 3056 3057 3058 3059 3060 3061 3062 3063 3064 3065 3066 3067 3068 3069 3070 3071 3072 3073 3074 3075 3076 3077 3078 3079 3080 3081 3082 3083 3084 3085 3086 3087 3088 3089 3090 3091 3092 3093 3094 3095 3096 3097 3098 3099 3100 3101 3102 3103 3104 3105 3106 3107 3108 3109 3110 3111 3112 3113 3114 3115 3116 3117 3118 3119 3120 3121 3122 3123 3124 3125 3126 3127 3128 3129 3130 3131 3132 3133 3134 3135 3136 3137 3138 3139 3140 3141 3142 3143 3144 3145 3146 3147 3148 3149 3150 3151 3152 3153 3154 3155 3156 3157 3158 3159 3160 3161 3162 3163 3164 3165 3166 3167 3168 3169 3170 3171 3172 3173 3174 3175 3176 3177 3178 3179 3180 3181 3182 3183 3184 3185 3186 3187 3188 3189 3190 3191 3192 3193 3194 3195 3196 3197 3198 3199 3200 3201 3202 3203 3204 3205 3206 3207 3208 3209 3210 3211 3212 3213 3214 3215 3216 3217 3218 3219 3220 3221 3222 3223 3224 3225 3226 3227 3228 3229 3230 3231 3232 3233 3234 3235 3236 3237 3238 3239 3240 3241 3242 3243 3244 3245 3246 3247 3248 3249 3250 3251 3252 3253 3254 3255 3256 3257 3258 3259 3260 3261 3262 3263 3264 3265 3266 3267 3268 3269 3270 3271 3272 3273 3274 3275 3276 3277 3278 3279 3280 3281 3282 3283 3284 3285 3286 3287 3288 3289 3290 3291 3292 3293 3294 3295 3296 3297 3298 3299 3300 3301 3302 3303 3304 3305 3306 3307 3308 3309 3310 3311 3312 3313 3314 3315 3316 3317 3318 3319 3320 3321 3322 3323 3324 3325 3326 3327 3328 3329 3330 3331 3332 3333 3334 3335 3336 3337 3338 3339 3340 3341 3342 3343 3344 3345 3346 3347 3348 3349 3350 3351 3352 3353 3354 3355 3356 3357 3358 3359 3360 3361 
```

What in the world is Artificial Intelligence?

At the risk of boring computer scientist readers to tears, this elementary explanation of how computers work was necessary to help mere mortals understand one of today's hottest technology buzz phrases—Artificial Intelligence, or AI. Simply put, it involves humans programming computers and digital devices to make decisions and do things on their own. A key subset of AI, Machine Learning, involves machines using data to become “intelligent”—which means making better choices.

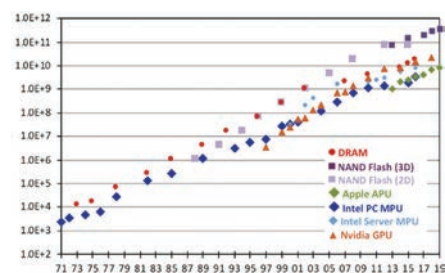
“Data is the source of a machine learning to be artificially intelligent,” According Dr. Eng Lim Goh, Senior Vice President & Chief Technology Officer, AI, Hewlett Packard Enterprise. “In this one sentence, I’ve linked data, machine learning, and AI. The key is the phrase in the middle, machine learning. A machine is learning from data, and ultimately through that learning, it becomes smart enough to make an informed choice. Additionally, those choices sometimes reflect human choices, because the digital device or computer is trained on the data humans choose to give the machine. This explains why we say the machine is artificially intelligent.”

The concept of artificial intelligence and machines thinking for us is not new. Almost three-hundred years ago, Jonathan Swift described “The Engine” in his satirical novel Gulliver’s Travels: “...whereas, by using this contrivance, the most ignorant person, at a reasonable charge, and with a little bodily labour, might write books in philosophy, poetry, politics, laws, mathematics, and theology, without the least assistance from genius or study.” And as if Swift could see the future, the book’s original sketch of The Engine looks remarkably like an IC.

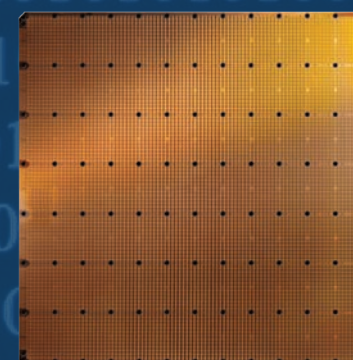
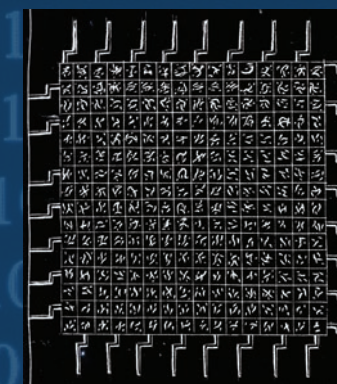
In the modern computing world, the conversation about artificial intelligence began in 1950 with the publication of Alan Turing’s “Computing Machinery and Intelligence.” In that paper, Turing—who is often referred to as the “father of computer science”—posed the question: “Can machines think?” To answer the question he offered the now famous “Turing Test.” It is remarkably simple: If a machine can engage in a conversation with a human expert without the human detecting it as a machine, the machine has demonstrated artificial intelligence.

How HPC gave AI a leg up

Trouble is, back in the 1950s computers did not really have enough horsepower to run AI applications in any meaningful sense. But that changed thanks to Moore’s Law, which most of us know is the prediction by Gordon Moore, co-founder and chairman emeritus of Intel Corporation, that the number of transistors in a dense IC would double approximately every two years. For decades, Gordon Moore’s prediction held true (Fig. 1). And the resulting growth of computing power is astonishing. While the first handheld calculator had a chip with approximately 2,000 transistors, we now have ICs packed with billions of transistors or more. In late 2020, Cerebras announced a Processing Unit (GPU) that has 1.2 trillion transistors.



(Fig. 1) Moore’s law predicted the doubling of transistors in a dense IC approximately every two years.



Johnathan Swift's
The Engine
circa 1726

Cerebras
GPU
circa 2020

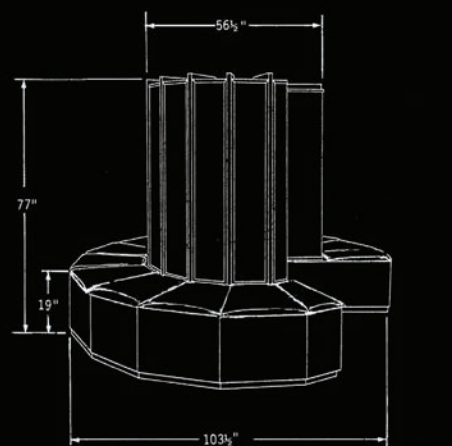
Wisconsin becomes a new superpower

The potential of integrating many of these powerful chips was not lost on some of the world's top scientists. One was Seymour Cray who had a passion for the art of designing computers and loved the challenge of coaxing more and more speed out of circuits and wires. In 1972, he formed Cray Research in Chippewa Falls, Wisconsin with the vision of building the world's fastest computing systems. Cray succeeded with the Cray-1. A masterpiece of engineering, it measured 8½ feet (2.60 meters) wide by 6½ feet (2 meters) high, tipped the scales at 10,500 pounds (4.76 metric tons), and incorporated 60 miles (96 kilometers) of wires.

Powered by newly designed ICs, the Cray-1 boasted more memory (one megabyte) and more speed (80 million computations per second) than any other computer in the world. And it retained the rank as the world's fastest supercomputer from 1976 to 1982. More importantly, it firmly established High Performance Computing (HPC) as the way to solve some of the world's toughest scientific and engineering challenges.

Hewlett Packard Enterprise (HPE) acquired Cray Inc. in 2019, and since then has made major investments in Chippewa Falls, sometimes called the supercomputing capital of the world. Today, HPE Cray EX supercomputers continue to push the boundaries of HPC. The NERSC Perlmutter system at the Lawrence Berkeley National Laboratory is among the 10 ten fastest computers in the world. While not the first supercomputer, the Cray-1 helped to spawn a whole new field of high performance computing and a race for faster and faster speeds. As of mid 2021, the fastest supercomputer on the planet was Fugaku at the RIKEN Center for Computational Science in Kobe, Japan. With 7,630,848 cores, the Arm-based supercomputer achieves an impressive speed of 442 petaflops.

When you combine the enormous processing power of HPC with the potential of AI you get a new field of science. Today, this combination is ushering in a new age of insights that can benefit everything from engineering innovation to new medicines to helping meet basic human needs to breakthroughs like clean fusion energy.



- Dimensions
 - Base - 103½ inches diameter by 19 inches high
 - Columns - 56½ inches diameter by 77 inches high including height of base
- 24 chassis
- 1662 modules (16 banks); 113 module types
- Each module contains up to 288 IC packages per module
- Power consumption approximately 115 kw input for maximum memory
- Freon cooled with Freon/water heat exchange
- Three memory options
- Weight 10,500 lbs (maximum memory size)
- Three basic chip types
 - 5/4 NAND gates
 - Memory chips
 - Register chips

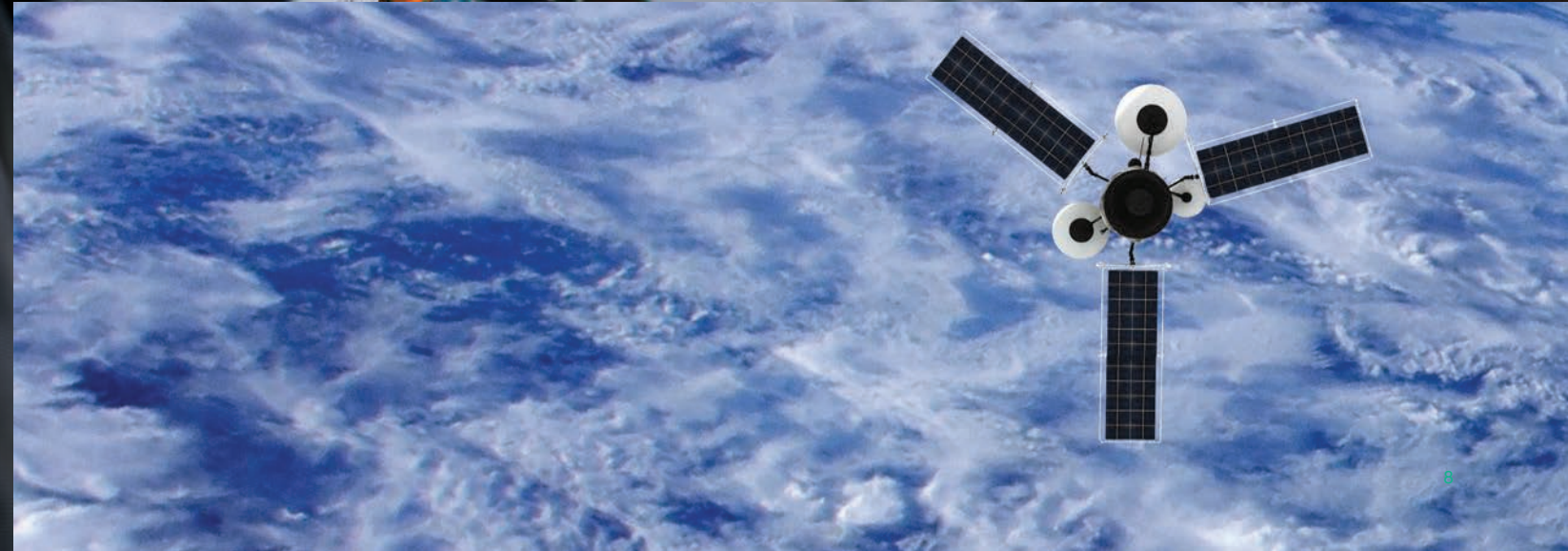
Figure 2-1. Physical organization of mainframe

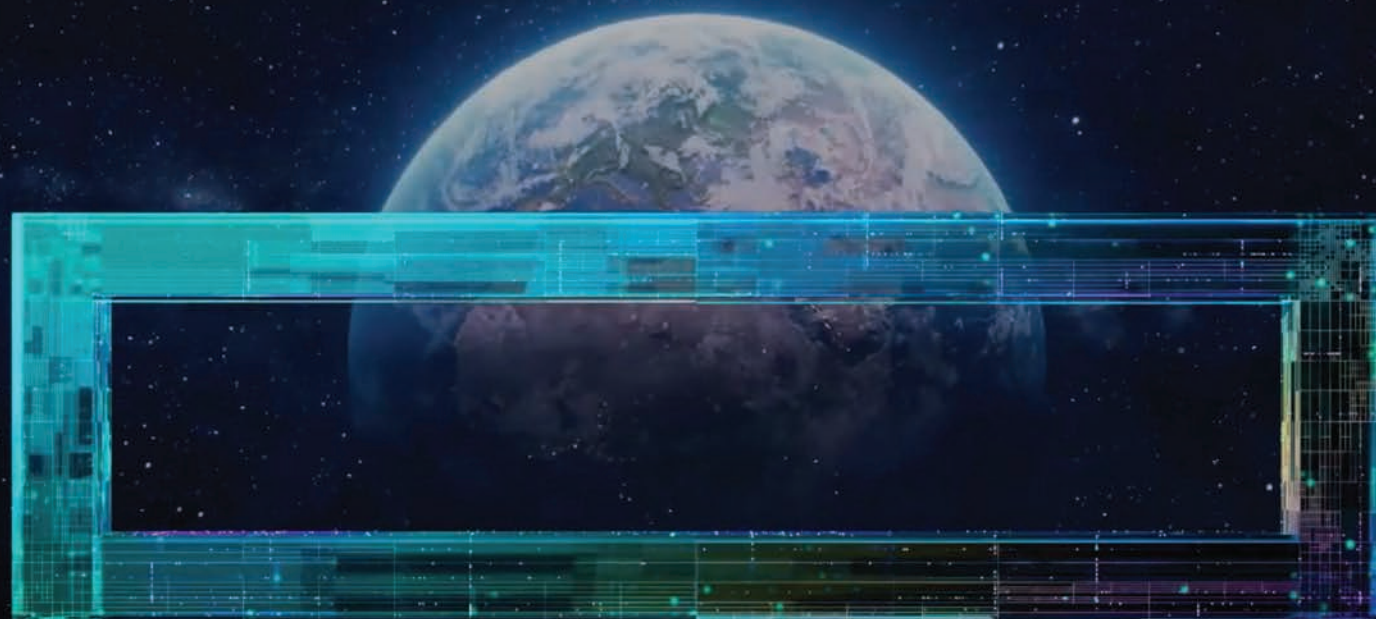
The Cray-1

HPC and AI are really taking off

Enabling this new era does not require gargantuan hardware. For example, Hewlett Packard Enterprise (HPE) tested the power of HPC in space by sending a mini supercomputer to the International Space Station (ISS) in 2017. Since the first crew arrived on ISS more than 20 years ago, scientist-astronauts have conducted countless hours of research for government agencies, academia, and commercial enterprises. And much of that requires computing power, which inspired a mission to send the first supercomputer beyond Earth – The Spaceborne Computer.

According to Hewlett Packard Enterprise and NASA, “The HPE Spaceborne Computer mission objective is to execute a long-duration operation of a Commercial Off-The-Shelf (COTS) high-performance computing (HPC) and Artificial Intelligence/Machine Learning (AI/ML) systems in the low Earth orbit internal ISS environment to determine conditions and parameters for continuance of correct computer operations.”





“We had many naysayers who claimed the supercomputing technology could not withstand the hazards of space travel,” explained Dr. Mark Fernandez, Principal Investigator on the Spaceborne Computer-2 project. “They fretted that everything from G-forces during the rocket launch to cosmic radiation would destroy the hardware.” In fact, prior to launch, Fortune magazine called it the computer that must survive ‘the shake, rattle, and roll’ of a space launch.

HPE further felt the success of the mission hinged on sending standard hardware to the ISS. “Computing in space has many challenges, such as cosmic radiation, which can cause computing bits in the computer’s memory to flip from zero to one, or vice versa,” according to Dr. Goh. “Many worried that we needed to harden the Spaceborne computer to withstand the much higher levels of radiation in space to prevent bit flip and other potential problems.

However, hardening or modifying such a complex system in any way could take years or even a decade when you include NASA’s extensive qualification process. “Much of the compute hardware on the ISS is twenty years old and we wanted the Spaceborne Computer technology to be on the bleeding edge,” remarked Dr. Goh. “That explains why the Spaceborne team recognized the mission’s success hinged on using pure COTS hardware. Even modifying the computer to run on DC current, which is used on the space station, would risk not putting the state-of-the-art technology on the ISS.”

In 2017, HPE successfully put a standard, off-the-shelf teraflop system on the ISS. “We made no modifications whatsoever to the COTS hardware,” continued Dr. Goh. “For example, since engineers generally use software to minimize bit flip due to background radiation on Earth, we thought why not in space. The Spaceborne Computer not only uses similar software methods but also slows its compute speeds during episodes of high radiation to ensure proper operation. The DC power problem was solved by making a special DC/AC converter, which allows the standard AC hardware to run on DC power generated by the ISS’s solar panels.”

It all worked, and now the HPE supercomputer running on the ISS today proves it’s possible to do a trillion calculations a second in space. “The Spaceborne Computer was not only the smartest and fastest computer in space, but thanks to its ability to use energy from solar power and cooling courtesy of the freezing temperatures in space it was also the most energy-efficient computer in existence,” added Dr. Goh.

The Spaceborne Computer represented a huge leap forward, and its state-of-the-art speed continues to shorten time to discoveries that can benefit all of humanity. Nonetheless, proponents of deeper space exploration understood that computers in space, and now artificial intelligence, would be critical for these endeavors. For example, you need only travel 254 miles from the Earth’s surface to reach the ISS, and you would need to travel approximately a thousand times that distance 238,900 miles (383,000 km) to reach the moon. But at its greatest distance in orbit, Mars lies 238 million miles (378,000,000 km) from the moon. That’s deep space.

“We had many naysayers who claimed the supercomputing technology could not withstand the hazards of space travel. They fretted that everything from G-forces during the rocket launch to cosmic radiation would destroy the hardware.”

Dr. Mark Fernandez
Principal Investigator,
Spaceborne Computer-2

Accelerating space exploration with state-of-the-art edge computing and AI capabilities

It can take as long as 40 minutes for radio waves to make the round trip from Mars to Earth. This results in extreme latency of data signals, which means that computers on human Mars missions or colonies would need to think for themselves—without direction from the home planet. Enabling that autonomy represents a key goal of the next space-bound supercomputer, which builds on the success of the Spaceborne Computer proof-of-concept. In February 2021, the Spaceborne Computer-2 roared toward the ISS, aboard the 15th Northrop Grumman Resupply Mission to the ISS. The NG-15 spaceship was christened “SS. Katherine Johnson” in honor of the NASA mathematician who proved critical to the early success of the space program and was made famous in the recent film *Hidden Figures*.

Today, Spaceborne Computer-2 is accelerating space exploration while increasing self-sufficiency for astronauts. It delivers twice the compute speed with purpose-built edge computing capabilities powered by the HPE Edgeline EL4000 system and an HPE ProLiant DL360 Gen10 server. This combination can ingest and process data from a range of devices, including satellites and cameras, and process it in real time. Spaceborne Computer-2 also incorporates GPUs because running AI requires a different type of supercomputing system. The GPU capabilities support specific use cases and projects where using AI and machine learning techniques can augment conventional compute processes.

Deep space exploration requires computing at the ultimate edge

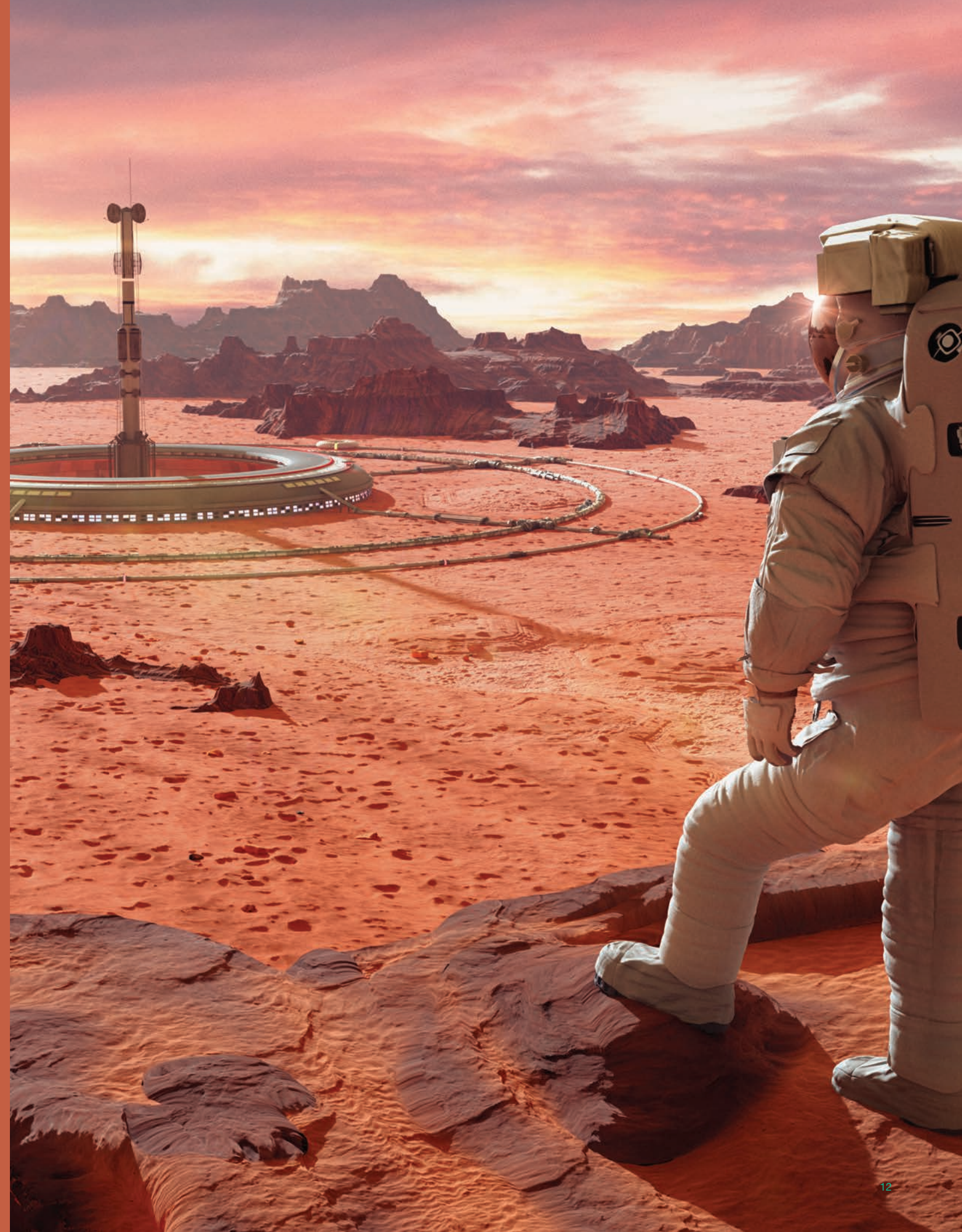
The combination of edge computing and AI will further enable astronauts to eliminate problems with longer latency and wait times associated with sending data over great distances. Even at the extreme edge, such as on another planet, autonomous supercomputing will help explorers tackle research and gain insights immediately for a range of projects, such as real-time monitoring of astronauts’ physiological conditions. By processing X-Rays, sonograms and other medical data supercomputers running AI and machine learning applications can speed time to diagnosis in space.

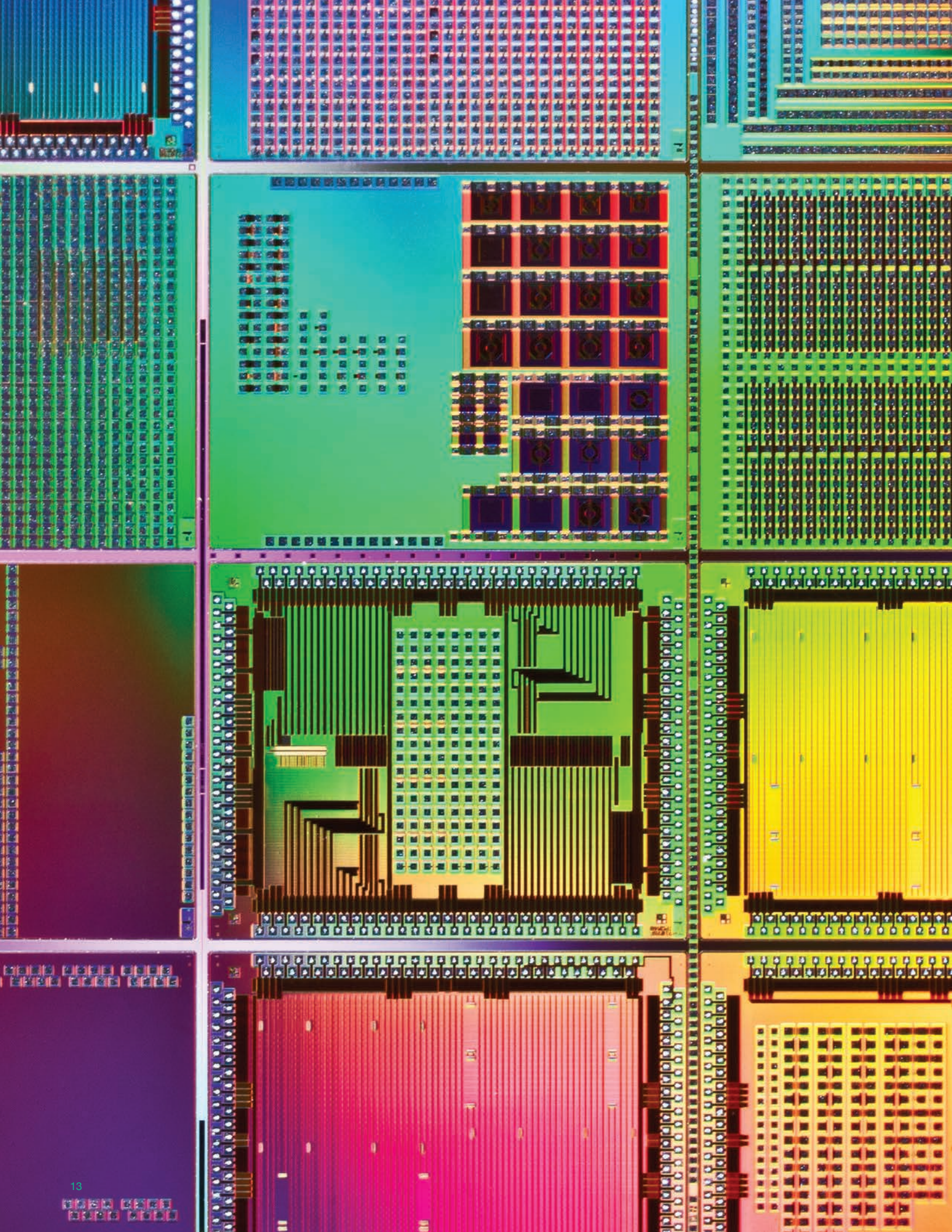
“Starting with the first Apollo mission to today’s exploration of Mars, onboard computing has always been essential to exploration,” says Professor Derek H. Warner of Cornell University’s Engineering School who has participated in the HPE Spaceborne Computer-2 project. His group at Cornell works to improve the science of industrial materials and the integrity of structures, key areas where astronauts on other planets might need to create new parts or repair spacecraft, rovers, or other critical equipment and infrastructure.

“From yesterday’s calculations of location from star positions, through tomorrow’s calculation of tool paths for manufacturing replacement parts in deep space, the HPE Spaceborne Computers represent a big step toward the ultimate in edge computing—deep space,” observed Professor Warner, whose team relies on HPC for their engineering research, which aims to increase the standards of living across the globe by enabling better performing and more sustainable civil engineering and environmental technologies.

“Starting with the first Apollo mission to today’s exploration of Mars, onboard computing has always been essential to exploration.”

Professor Derek H. Warner
Cornell University Engineering





“As part of its standing goal of deep space exploration, NASA has emphasized the necessity for earth-independent advanced manufacturing to maintain life-support and establish habitation abroad,” explains Professor Warner’s colleague, Cornell University researcher Terrence Moran, Principal Investigator on a NASA and HPE approved experiment running on Spaceborne Computer-2 aboard the ISS. “Due to the scarcity of resources in space, the ability to simulate these manufacturing processes before fabrication is critical. For reasons like this, onboard high-fidelity scientific computing becomes mission-critical for the advancement of human exploration.”

What is the difference between HPC and supercomputing?

The small size and weight of the Spaceborne Computers contrast dramatically against room-size computers like Fugaku. Today, high performance computing comes in many forms and sizes. An entry-level HPC system typically has multiple interconnected nodes, which often share the same operating system. This architecture enables users to combine the power of many microprocessors to make fast work of simulations, modeling and other technical computing challenges. Supercomputers tend to incorporate a very large number of CPUs or GPUs in a single machine or group of computers that are linked via extremely fast interconnects. According to Hyperion research, systems costing more than \$500,000 represent the border between “HPC” and “supercomputers.” These include scale-out systems with as many as 2,000 nodes and millions of cores, such as HPE solutions that ComnetCo has provided for major customers like the U.S. Department of Energy (DOE).

In any form, High Performance Computing helps solve many of the most complex problems in engineering, life sciences, medicine, weather and climate, cosmology, and many other fields. What differentiates supercomputing is the massive speeds at which these systems can solve problems. This holds true regardless of size, as even small-form machines like HPE’s first Spaceborne Computer, a proof-of-concept launched and installed on the International Space Station in 2017 for a one-year mission to operate in the harsh conditions of space. The mini supercomputer like the first Spaceborne Computer can achieve speeds at teraflop scale. That’s one trillion floating-point operations per second, or one million million operations, represented by 1 followed by 12 zeros.

Advances in supercomputing in recent years have led to huge scale-out systems capable of petaflop speeds. Imagine a room perhaps 10s of thousands of square feet in size filled to the brim with computers. But a key difference here between a petascale supercomputer and a typical IT data center is that all these computers are bound together by an extremely high throughput and low latency network. The supercomputing architecture makes it possible to solve one problem across all those computers. As a result, machines such as these can take an extremely complex problem that might otherwise require weeks to solve and bring that time down to minutes, or even seconds.



ComnetCo delivered Sawtooth, the latest supercomputer at the Idaho National Laboratory. Funded by the DOE, the system ranked #37 on the November TOP500 list of the fastest supercomputers in the world when it came online in late 2019.

This lightning speed explains why supercomputers and AI are being applied to some of the most pertinent and pressing problems facing the planet Earth and the people living on it. Take the COVID-19 pandemic as an example. Here scientists are applying supercomputers in several areas, the most notable being vaccine research. Supercomputers help in this field by, for example, running the gene expression data of volunteers participating in clinical trials. When the vaccine is injected, genes are expressed, and then researchers can use AI to run through the data and try to predict the efficacy of the vaccines, or sometimes even side effects.

Exascale, like the next moonshot

Perhaps not since the ambitious mission to put astronauts on the moon has humankind taken on a more ambitious technological challenge as achieving Exascale. That means building computers capable of calculating at least one quintillion (10^{18}) floating point operations per second; that's a billion billion, or 1 followed by 18 zeros. Many even doubted the possibility of achieving Exascale due to challenges such as power consumption costs.

However, HPE recognized Exascale was not only achievable but also held enormous possibilities holds for ushering in a new era of insights. Towards this ambitious goal, HPE acquired two of the world's pioneering companies on the path to Exascale computing: Silicon Graphics (SGI) and Cray.

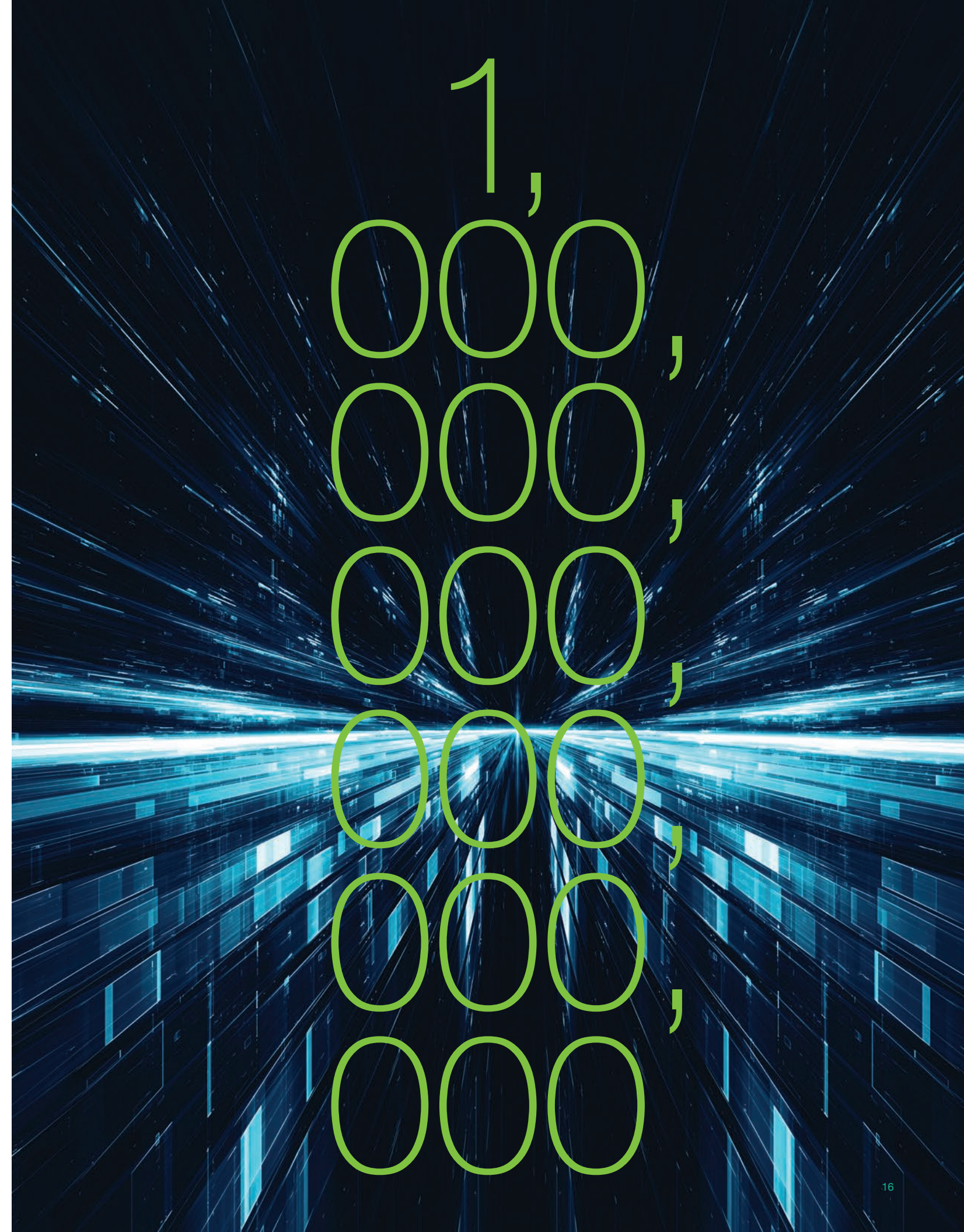
Today, many advanced technologies, intellectual property, and people gained in the SGI acquisition are key to this Exascale strategy. The next-generation MC990 X and HPE Superdome Flex servers produced by HPE are based on the SGI ccNUMA technology. These HPE Integrity systems handle technical computing workloads that are often too challenging to run in a cluster. These technologies have also helped HPE along the path to Exascale.

The culmination of a colossal undertaking, an HPE Cray EX system called 'Frontier' will become the first Exascale system in the world.



Since the two companies merged, HPE has committed to keeping the Cray brand and building on its technologies to advance HPC and AI technologies. Frontier, an upcoming exascale system based on the HPE Cray EX supercomputer commissioned by the U.S. DOE's Oak Ridge National Laboratory (ORNL), will be capable of 1.5 exaflops, or 1.5 quintillion calculations per second. How fast is that? Imagine that if all 7.7 billion people on Earth used a pocket calculator to do one calculation per second, it would take six years to do what Frontier could do in one second.

So, Exascale is here today and it's providing the incredible processing speed required for today's new massive AI, converged modeling, simulation, and analytics workloads.





Speed matters

Whether it's racing to find new medicines, predicting the landfall of a hurricane in time to get people out of harm's way, or finding financial fraud before the money is gone, speed can make a big difference. Take for example the critical field of cybersecurity. It seems not a day goes by without a headline about this or that data breach or crippling cyber attack. By 2025, the cost of cyberattacks worldwide will cost \$10 trillion. Increasingly sophisticated cybercriminals may steal your private data, turn off your power, or close your hospital.

In fall of 2020, for instance, the FBI and other federal agencies issued a rare warning that they had credible information of "an increased and imminent cybercrime threat" to U.S. hospitals and healthcare providers. They warned that groups of malicious cybercriminals could unleash a wave of data-scrambling extortion attempts called ransomware attacks against the U.S. healthcare system. If successful, these cyber attacks could paralyze hospital information systems just as nationwide cases of COVID-19 were spiking. Trouble is, even with a warning, ransomware attacks are very difficult to detect before the damage is done.

Ransomware, malware, DDoS attacks, all rely to some extent on the element of surprise.

Whether security teams must protect a hospital's databases or the controls on a nuclear power plant, the greatest fear is of the unknown. Conventional cybersecurity defenses such as anti-malware programs operate using information regarding known cyberattack methods. They are defenseless against never-before-seen strains of malware, which security experts call "zero day" threats. Predicting unknown threats is an exciting application where the speed and the power of HPC coupled with AI and machine learning can act like over-the-horizon radar for security teams.

So how can AI help the defenders become more proactive? A key involves predicting unknowns, which is a bit like looking into a crystal ball and predicting the future. One method of doing this is to sit down with someone who has been working in cybersecurity for 30 years, learn what they know, and then write a set of rules. People have done this in the financial services business by extracting knowhow from experienced and successful traders before they retire, and then codifying it in rules.

"The issue with that method is the scenarios constantly evolve, the data changes, and so on. This explains why these rules need to be augmented by an AI system that's learning from data all the time," explains Dr. Goh. "Essentially, you constantly replace old rules learned from history with new rules based on very recent history in the data. The machine learns from the new data that comes in and makes a prediction on that data. It can't actually see the threats in a crystal ball, but its prediction can come very close." This information gained from AI can give defenders time to fill the holes in their attack surface, or spot malware inside a network or system that has already evaded their defenses—before the damage is done.

So how can AI help the defenders become more proactive? A key here is predicting unknowns, which is a bit like looking into a crystal ball and predicting the future.

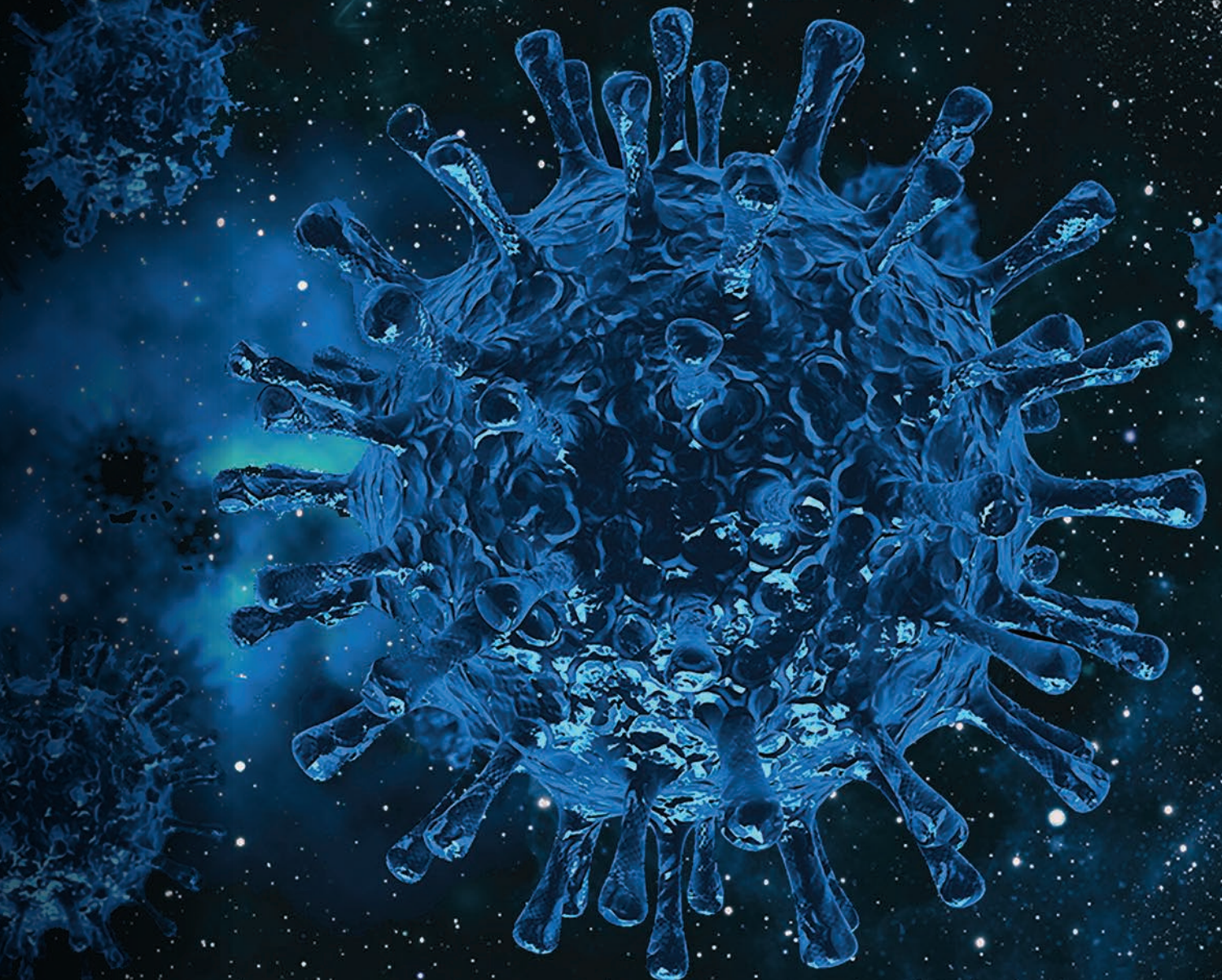
That's the Holy Grail, finding a drug that you may be able to put in the body, and then every time a COVID-19 virus comes, the drug attaches itself to the spike's tip before it even touches the human's ACE-2 receptor.

"Another area benefitting from the speed delivered by this new era of supercomputing coupled with AI is the critical field of 3D electron microscopes," explains Dr. Goh. "For example, people researching COVID-19, have focused on the virus' spike protein, because this is what attaches to our ACE-2 receptor and then starts the infection. However, the very small size of the virus makes this difficult. So we use cryo-electron microscopy (Cryo-EM). In other words, researchers freeze the live virus, and then put it under a Cryo-EM microscope. And then they take 4K by 4K video (and in some cases 8K by 8K video) of the virus while tilting it slightly each time, just very slightly. And they do this for two days. The issue with this method arises from the terabytes and terabytes of data exhaust flowing out of the process, and the resulting data backlog on the microscope."

One innovative research team solved this problem by developing a technique to extract the data from the microscope at a very high rate. So once data was extracted, a machine learning system developed together with HPE helped to compose all these images of the virus at different angles. Once an AI-based system created a 3D model of the virus, the researchers could clearly see the shape of the spike, which is important because the shape fits very nicely with our ACE-2 receptor. And apparently, in the Delta variant the shape fits the ACE-2 receptor even better than the Alpha and Beta variants. This tight fit explains why Delta is more infectious.

With that information, supercomputing also helps during drug development. For example, the intense processing speed of a supercomputer makes it possible to go through a 100,000-molecule library to see if it can find one that fits even better than the tip of the virus that attaches to the ACE-2 receptor. That's the Holy Grail, finding a drug that you could put in the body, and then every time a COVID-19 virus comes, the drug attaches itself to the spike's tip before it even touches the human's ACE-2 receptor.

"But in addition to the fact that it takes time to pull the data out from each microscope, the next problem has to do with HPC simulation of molecular docking," continues Dr. Goh. "To dock one molecule from 100,000 molecules in the libraries and try it with the tip requires a lot of HPC resources, which can get very expensive. So what HPE did was to build an AI tool to go through the library, train on past successes, and use AI-augmented molecular docking to choose drug candidates that were more likely to succeed. The researchers could then do more faithful simulation with those shortlisted drugs. Without AI augmenting supercomputing by pre-selecting a shortlist of candidates to do much more detailed work, the HPC would need to simplify its physics in order to go through the 100,000-molecule libraries in a timely manner."





The Intelligent Edge

As edge devices—whether on an oil pipeline in Texas, a patient’s wrist, or a rover on Mars—gain more memory and processing power they become more capable of solving problems autonomously. This makes AI and machine learning increasingly important.

“More and more, we are thinking at some point a smart edge device should not only be running a trained AI/ML learning model given to it by humans, but should also be doing learning on its own based on the data it’s collecting,” explained Dr. Goh. “This is the next forward-thinking concept. However, one catch here is that each sensor will be looking at its own compartmentalized data, and therefore will be highly biased towards the data it’s seeing. Eliminating this bias by sharing AI/ML outcomes from many edge devices is one reason why we came up with a concept called ‘Swarm Learning’, which appeared on the cover of the journal *Nature*.”

The beauty of Swarm Learning is that it allows for the insights generated from data to be shared without sharing the source data itself. Plus, there is no central node which aggregates the data. This decentralized, distributed model shares only the insights gleaned from the data, often derived using AI models. “We believe this approach will usher in a new era in research analytics,” continued Dr. Goh. “One area where we are looking at applying swarm learning is in enabling medical researchers around the world to use a private blockchain to share insights without compromising patients’ privacy.”

“Fast and reliable detection of patients with severe and heterogeneous illnesses is a major goal of precision medicine,” according to the *Nature* article. “Patients with leukemia can be identified using machine learning on the basis of their blood transcriptomes [the complete set RNA molecules expressed in a blood cell]. However, there is an increasing divide between what is technically possible and what is allowed because of privacy legislation. Here, to facilitate the integration of any medical data from any data owner worldwide without violating privacy laws, we introduce Swarm Learning—a decentralized machine-learning approach that unites edge computing, blockchain-based peer-to-peer networking and coordination while maintaining confidentiality without the need for a central coordinator...”¹

This new concept offers unprecedented opportunities to use the ever-increasing volumes of data produced and collected at the edge. It represents just one of many ways the way the intelligent edge, with machines learning from this data, is already benefitting humans on Earth—and soon perhaps on other planets.

¹ Warnat-Herresthal, S., Schuitze, H., Shastry, K.L. et al. *Swarm Learning for decentralized and confidential clinical machine learning*. *Nature* 594, 265–270 (2021).

To Exascale and beyond

This example represents just one of many benefits delivered by the tremendous speed of today's supercomputers coupled with AI. Plus, new systems coming in the Exascale era will enable never-before-possible insights and innovations like sequencing the wheat genome to find new strains that could offer resistance to climate change. AI and supercomputing can save lives or save the planet. They are helping to find vaccines and drugs to fight new variants of COVID-19, and modeling the earth's oceans and atmosphere to help prevent global warming. This combination could also find solutions to sustainable energy, such as solving the challenge of nuclear fusion. It's already enabling smart cities and making autonomous vehicles smarter. Who knows what will come next.

To learn more about how you can take advantage of HPC and AI, explore [HPE.com](https://www.hpe.com), and [comnetco.com](https://www.comnetco.com).



About ComnetCo

Two decades of experience in the evolution of HPC helps ComnetCo configure powerful compute and storage systems. This virtually unequalled track record includes delivering some of the world's fastest supercomputers. Together with its primary partner, HPE, ComnetCo helps optimize systems for the unique needs of researchers in Higher Education, Research Institutes, Global Enterprises, and Federal Government Agencies. These solutions—which include purpose-built platforms for AI—help scientists and engineers speed time to discovery in fields ranging from pharmaceutical research like new vaccines to industrial companies creating new materials to supporting deep space exploration. For more information, visit: www.comnetco.com



About Hewlett Packard Enterprise

Hewlett Packard Enterprise (NYSE: HPE) is the global edge-to-cloud company that helps organizations accelerate outcomes by unlocking value from all of their data, everywhere. Built on decades of reimagining the future and innovating to advance the way people live and work, HPE delivers unique, open and intelligent technology solutions delivered as a service – spanning Compute, Storage, Software, Intelligent Edge, High Performance Computing and Mission Critical Solutions – with a consistent experience across all clouds and edges, designed to help customers develop new business models, engage in new ways, and increase operational performance. For more information, visit: www.hpe.com