

Universidade de Pernambuco

Programa de Pós-Graduação em Engenharia da Computação (PPGEC)

Proposta de Tese de Doutorado

Área: Computação Inteligente / IoT

Título: Otimização das redes de nevoeiro por técnicas de IA e arquitetura RISC-V para IoT

Orientador – Sérgio Murilo Maciel Fernandes (smurilo@ecomp.poli.br)

Coorientador – Edison de Queiroz Albuquerque (edison@ecomp.poli.br)

Descrição

Após o surgimento e evolução do paradigma IoT (*Internet of Things*), tem-se observado uma enorme expansão no número de dispositivos sensores implantados no dia-a-dia das pessoas. Nos últimos anos, muitos aplicativos de IoT têm surgido em vários domínios, como por exemplo, nos esportes (por meio dos aparelhos esportivos Fitbits e Apple Watches), na saúde (dispositivos de marcapassos e biochips), nos transportes (carros autônomos), nas residências e nas cidades inteligentes, bem como na manutenção preditiva em sistemas de missão crítica (usinas nucleares ou aviação) [1]. O termo IoT é a generalização de uma variedade de coisas (dispositivos) heterogêneas conectadas (por exemplo: sensores, atuadores, smartphones, computadores, automóveis, geladeiras, fogões, eletrodomésticos, entre outros) que podem interagir uns com os outros e cooperar com outros dispositivos vizinhos para alcançar objetivos comuns [2]. A troca de informação entre estes dispositivos tem proporcionado um crescimento da capacidade de geração de dados, resultando no acúmulo de uma grande variedade de informações que vão desde medições de fenômenos naturais, até comportamentos relacionados aos seres humanos. O avanço da Internet das Coisas possibilita a interconexão de todos os dispositivos entre si e com a Internet [3]. Isto demanda uma necessidade de controle, transferência, armazenamento e processamento seguro de uma grande quantidade de dados (da ordem de *terabytes*) e a utilização de técnicas de inteligência artificial (*Artificial Intelligence* ou AI) e de aprendizado de máquina (*machine learning*), como evidenciado em aplicações de inteligência cognitiva e de máquina em classificação de imagem, reconhecimento facial ou tradução de linguagem [1]. A combinação de processadores e FPGAs (*Field Programmable Gate Array*) em um único pacote resulta no conceito FPSoC (*Field-Programmable System-on-Chip*). A arquitetura do FPSoC é apresentada em [1], junto a uma análise da utilidade de seus recursos de hardware, para implementação dos algoritmos de aprendizado de máquina em dispositivos de IoT. Outros trabalhos como [4], [5] e [6] implantaram a tecnologia dos FPGAs para IoT com o objetivo de aumentar o desempenho dos sistemas, conectividade e a interoperabilidade dos dispositivos de borda (ou IoT), sendo capaz também de fornecer soluções mais rápidas, de maior eficiência energética para algoritmos de computação intensivos de dados com níveis de paralelismo. Os nós finais da Internet das Coisas exigem plataformas de processamento de baixa potência, caracterizadas por unidades dedicadas heterogêneas, controladas por um núcleo de processador que executa threads de controle simultâneo. Esse esquema de arquitetura se encaixa em um dos principais domínios de aplicativo de destino do conjunto de instruções RISC-V. O trabalho [7] apresenta um núcleo de processamento de código aberto compatível com RISC-V no lado do software e com a popular plataforma de processador Pulpino no lado do hardware, enquanto suporta *multithreading* intercalado para aplicativos IoT. À medida que o mercado de IoT cresce, o RISC-V é um excelente candidato para adoção devido à sua arquitetura personalizável. O RISC-V surgiu como uma tecnologia de processador que oferece novas oportunidades de design e o nível de desempenho e extensibilidade necessários no mundo hiperconectado atualmente [8]. Em [9] é proposto o GAP-8, um mecanismo de computação RISC-V IoT totalmente programável, de alto desempenho e baixo consumo. Segundo [10], o GAP-8 tornou-se um novo processador de aplicativos IoT/IA. Em particular, o GAP-8 é projetado para processamento de inferência de IoT, permitindo que aplicativos analisem dados diretamente dos sensores de imagem, áudio e movimento enquanto se comunicam em redes de taxas de dados relativamente baixas, mantendo baixos os requisitos de *throughput* de dados *upstream*. Este chip é particularmente adequado para todo um conjunto de outras aplicações, como amostragem de sinal de áudio e reconhecimento de fala, monitoramento de integridade da máquina.

Por outro lado, é necessário o fornecimento de infra-estrutura de rede e uma computação eficiente para aplicativos IoT através de alta velocidade e baixo tempo de resposta [11]. A computação em nuvem apresenta grande capacidade de armazenamento e processamento para aplicações de IoT, além de armazenar uma grande quantidade de dados e ser acessada em qualquer lugar do mundo [12]. Devido ao alto armazenamento e alto poder de processamento, os dados podem ser acessados eficientemente. Apesar da plataforma da computação em nuvem (*cloud computing*) ter se tornado acessível ao processamento de dados de uma grande gama de aplicativos em geral e oferecer suporte aos dispositivos de IoT, por frequentemente utilizar solução balanceada para agregar flexibilidade e poder computacional eficiente para algoritmos de aprendizado de máquina, apresenta uma arquitetura em nuvem

centralizada e uma resposta inadequada para os requisitos de alta mobilidade, reconhecimento de localização e baixa latência, necessários a IoT. Deste modo o conceito de computação de nevoeiro emerge, onde dados e computação estão próximos ao usuário final, eliminando os problemas acima mencionados e fornecendo baixa latência, reconhecimento de localização e melhoria da qualidade dos serviços (QoS) para aplicações em tempo real [13], atuando como interface entre IoT e a computação em nuvem, além de otimizar o envio de dados para a nuvem. Esta proposta de doutorado, inspirada nestas tecnologias, propõe otimização na comunicação entre os dispositivos eletrônicos e entre estes e os seres humanos, e também na minimização/redução do impacto dos gargalos com relação ao acúmulo de informações destes dispositivos no sistema por meio de computação em nevoeiro e algoritmos de inteligência artificial. Desta forma, com a contribuição de técnicas de IA, dentre elas as redes neurais profundas e algoritmos de aprendizagem de máquina e com o suporte do RISC-V, ou mesmo de FPGAs, dispositivos interligados com IoT deverão ser capazes de suportar requisitos de alta mobilidade e reconhecimento de localização, com baixa latência e alta segurança no armazenamento, processamento e transmissão dos dados.

Referências Bibliográficas

- [1] Roberto Fernandez Molanes, Kasun Amarasinghe, Juan Rodriguez-Andina, and Milos Manic. Deep learning and reconfigurable platforms in the internet of things: Challenges and opportunities in algorithms and hardware. *IEEE Industrial Electronics Magazine*, 12(2):36–49, 2018.
- [2] Guto Leoni Santos, Patricia Takako Endo, Matheus Felipe Ferreira da Silva Lisboa, Leylane Grazielle Ferreira da Silva, Djamel Sadok, Judith Kelner, Theo Lynn, et al. Analyzing the availability and performance of an e-health system integrated with edge, fog and cloud infrastructures. *Journal of Cloud Computing*, 7(1):16, 2018.
- [3] Luiz Bittencourt, Roger Immich, Rizos Sakellariou, Nelson Fonseca, Edmundo Madeira, Marilia Curado, Leandro Villas, Luiz da Silva, Craig Lee, and Omer Rana. The internet of things, fog and cloud continuum: Integration and challenges. *Internet of Things*, 2018.
- [4] Tiago Gomes, Sandro Pinto, Adriano Tavares, and Jorge Cabral. Towards an fpga-based edge device for the internet of things. In *Emerging Technologies & Factory Automation (ETFA), 2015 IEEE 20th Conference on*, pages 1–4. IEEE, 2015.
- [5] Saman Biookaghazadeh, Ming Zhao, Fengbo Ren, George Pavlou, Alex Afanasyev, Lixia Zhang, Ying-Dar Lin, Lior Khernmsh, and Drew Roselli. Are fpgas suitable for edge computing? In *{USENIX} Workshop on Hot Topics in Edge Computing (HotEdge 18)*, 2018.
- [6] Alfonso Rodríguez, Juan Valverde, Jorge Portilla, Andrés Otero, Teresa Riesgo, and Eduardo de la Torre. Fpga-based high-performance embedded systems for adaptive edge computing in cyber-physical systems: The artico3 framework. *Sensors*, 18(6):1877, 2018.
- [7] Abdallah Cheikh, Gianmarco Cerutti, Antonio Mastrandrea, Francesco Menichelli, and Mauro Olivieri. The microarchitecture of a multi-threaded risc-v compliant processing core family for iot end-nodes. *arXiv preprint arXiv:1712.04902*, 2017.
- [8] Risc-v: All hype or real hope for the processor market? <https://www.allaboutcircuits.com/industry-articles/risc-v-all-hype-or-real-hope-for-the-processor-market/>. Accessed: 2019-01-17.
- [9] Eric Flamand, Davide Rossi, Francesco Conti, Igor Loi, Antonio Pullini, Florent Rotenberg, and Luca Benini. Gap-8: A risc-v soc for ai at the edge of the iot. In *2018 IEEE 29th International Conference on Application-specific Systems, Architectures and Processors (ASAP)*, pages 1–4. IEEE, 2018.
- [10] The risc-v momentum continues with the gap8, a new iot/ai application processor. <https://fuse.wikichip.org/news/990/the-risc-v-momentum-continues-with-the-gap8-a-new-iot-ai-application-processor/>. Accessed: 2019-01-17.
- [11] Ashkan Yousefpour, Genya Ishigaki, and Jason P Jue. Fog computing: Towards minimizing delay in the internet of things. In *Edge Computing (EDGE), 2017 IEEE International Conference on*, pages 17–24. IEEE, 2017.
- [12] Pengfei Hu, Sahraoui Dhelim, Huansheng Ning, and Tie Qiu. Survey on fog computing: architecture, key technologies, applications and open issues. *Journal of Network and Computer Applications*, 98:27–42, 2017.
- [13] Ricardo Silva, Jorge Sá Silva, and Fernando Boavida. Opportunistic fog computing: Feasibility assessment and architectural proposal. In *Integrated Network and Service Management (IM), 2017 IFIP/IEEE Symposium on*, pages 510–516. IEEE, 2017.