SATURN 2020

Quality Attribute Concerns for Microservices at the Edge

Marc Novakouski

Tactical and AI-Enabled Systems Initiative

Software Engineering Institute

Document Markings

This material is based upon work funded and supported by the Department of Defense under Contract No. FA8702-15-D-0002 with Carnegie Mellon University for the operation of the Software Engineering Institute, a federally funded research and development center.

DM20-0013

Agenda

Introduction

Challenges at the Edge

Architectures for the Edge

Microservices in Edge Systems

Microservices: Scaling Down at the Edge

Introduction

Who is TAS? Who am I?

Tactical and AI-Enabled Systems Initiative

- SEI Team of 10+ researchers, engineers, and domain experts working on edge technologies
- Focus on edge technologies since 2010
- Research & Customer Projects includes ML/AI, Networking, HMI, Data Analysis, Mobile Platforms, Context Awareness, Security, Resource Management

Marc Novakouski

- SEI Senior Engineer, 10+ years; Raytheon Senior Software Engineer, 8+ years
- Research & Prototyping work includes Microservices, Mobile Devices, Networking
- Customers include DHS, Army, Navy, USAF, USMC, SOCOM

What is this talk about?

We are good at building **complex** systems in **safe** locations doing **relatively predictable** things that change **over time** and may need to scale **up**.

- E-Commerce
- AI/ML applications
- Searching, Mapping, Social Media

We are less adept at building systems in **unsafe** locations in **dynamic** situations that change **at a moment's notice**; i.e. at the "Edge." In these situations, typically we fall back to **simple, stand-alone** solutions.

- Disaster recovery & First Responders
- Soldiers under fire

Q: How do we realize complex, dynamic solutions at the Edge?

A: By building to scale **down**, instead of **up**

Microservices: Scaling Down at the Edge

Challenges at the Edge

What is “the Edge”? - 1

The **Humanitarian Edge**: First Responders & Humanitarian Aid

Cellular & Analog Infrastructure unavailable or limited

Power unavailable or limited to generators or batteries

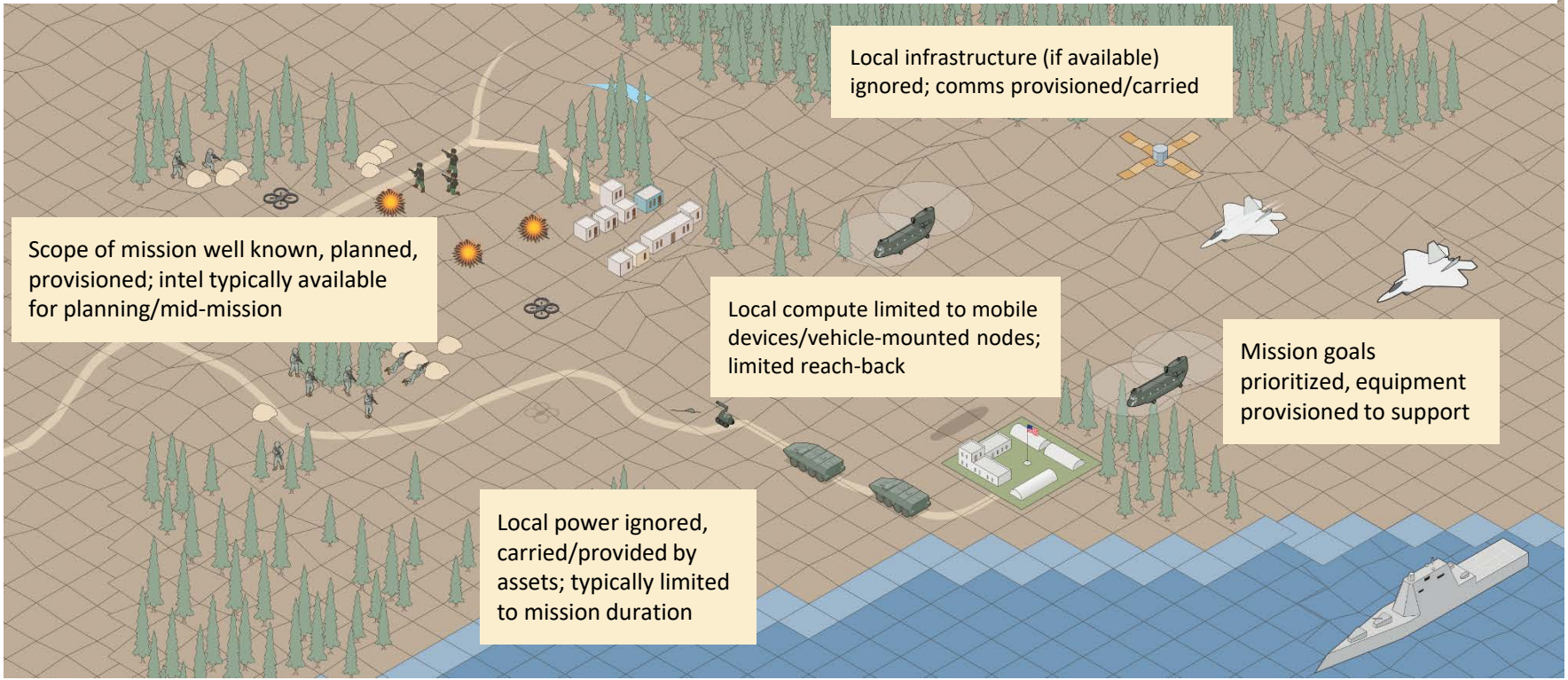
Scope of mission unknown, includes survey, discovery, triage of issues

Compute unavailable or limited to mobile devices; limited reach-back

Medical & Humanitarian needs prioritized; infrastructure repair typically later

What is “the Edge”? - 2

The **Tactical Edge**: Warfighters & Military Assets



What are the Challenges?

	Humanitarian Edge	Tactical Edge
Limited Network	• Broken Infrastructure • Different networks in use (Bandwidth) • Coordination Required	• Military Hardware (rugged but limited) • Variety of Connectivity/Bandwidth limitations • Mission-specific data prioritization • Interference (terrain/opposing force)
Limited Power	• Unknown Mission Duration • Limited battery/plug in resources	• All power carried/vehicle borne • Mission-specific power prioritization
Limited Compute	• Limited/no reachback to cloud • Mobile devices only	• All compute carried/vehicle borne • Mission-specific compute prioritization
Unsafe Conditions	• Roads/Buildings/Infrastructure • Unknown Mission Scope • Limited attention • Prioritization of humanitarian equipment (medial, construction) over other resources (compute, power, network)	• Lives in danger • Focus on environment, adversaries; extremely limited attention • Loss of equipment (damage, jamming) • Adaptation to role changes (injury, casualty) • Security of equipment (zero-ize)

Microservices: Scaling Down at the Edge

Architectures for the Edge

Scenario Assumptions

Equipment

- A person operating at the edge will have a mobile device of some kind
  - Cell phone, raspberry pi, etc
- The mobile device will have limited connectivity & power
  - Spotty, low bandwidth, restricted access to cloud resources
  - Power should be rationed due to rare or unavailable charging opportunities

Mission

- A person operating at the edge will not have a full understanding of the situation
  - Situational awareness information is incomplete, old, or unavailable
- The mission will be dangerous
  - Teams will be small & distributed, if not individual
  - Majority of attention required to attend to situation
  - Possible damage to/loss of equipment, possible injury or loss of life

In general, a dangerous mission with limited compute, networking, power, manpower, and attention

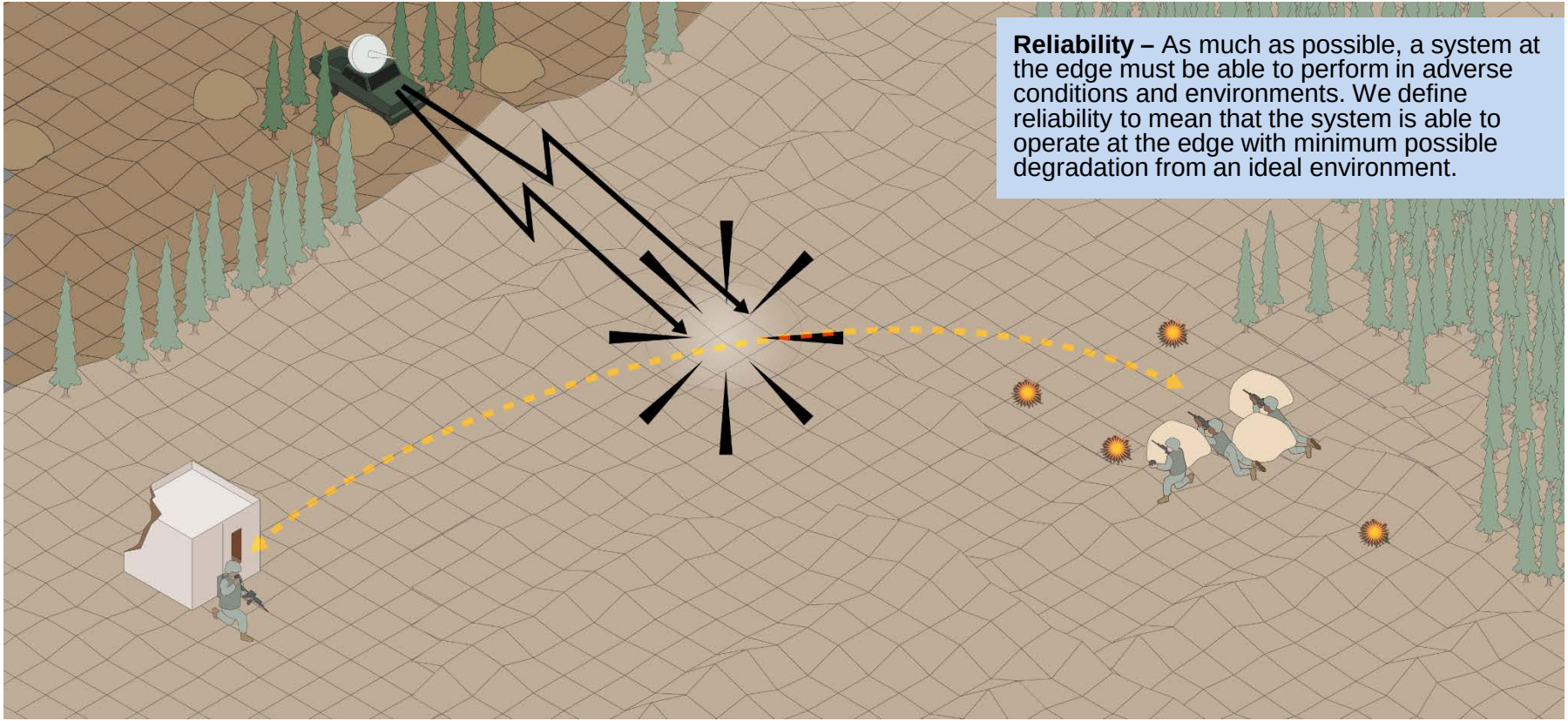
Quality Attributes for the Edge

Compensating for Edge challenges requires specific Architectural Features

TAS experience at the Humanitarian & Tactical Edge suggest:

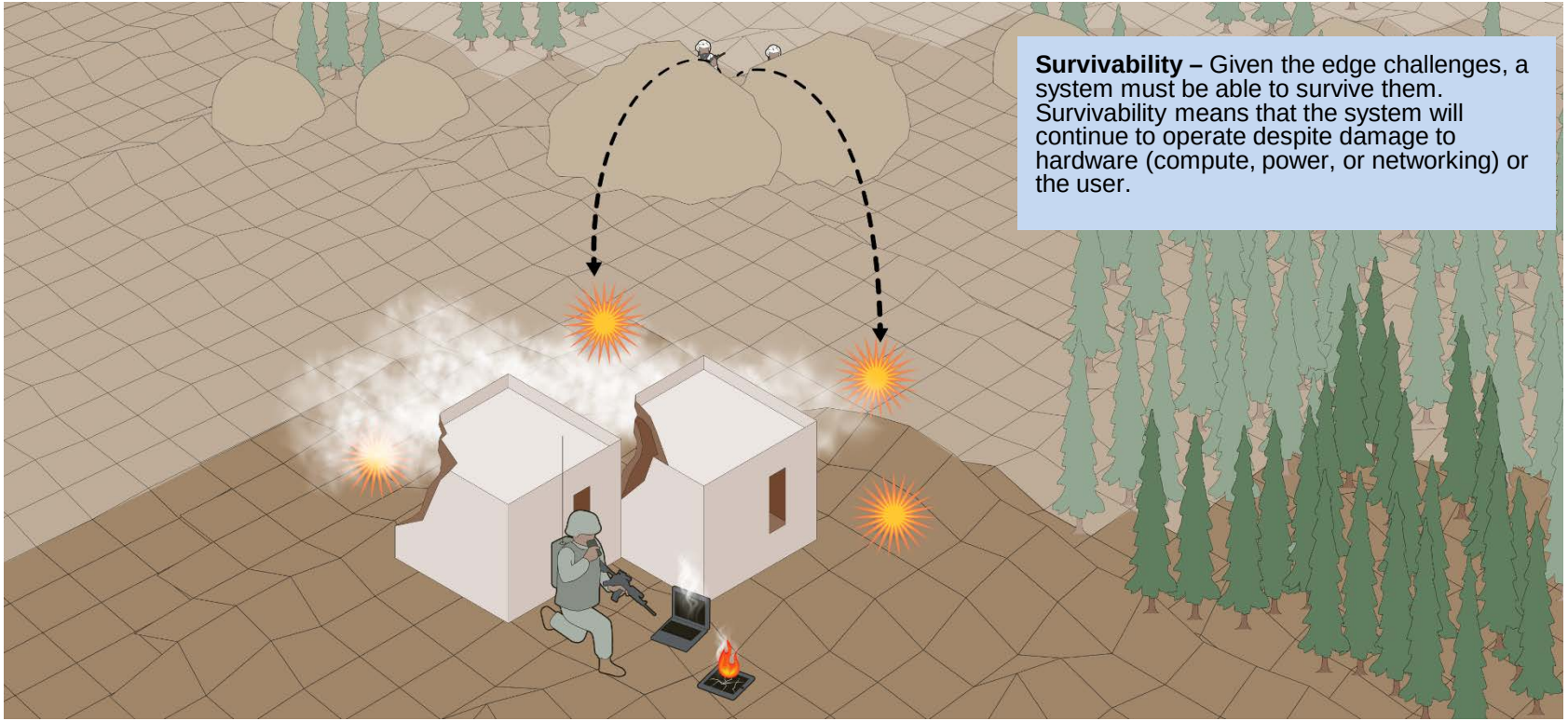
- Reliability
- Survivability
- Autonomy
- Adaptability
- Flexibility
- Distributability
- Openness

Quality Attributes for the Edge - Reliability



Reliability – As much as possible, a system at the edge must be able to perform in adverse conditions and environments. We define reliability to mean that the system is able to operate at the edge with minimum possible degradation from an ideal environment.

Quality Attributes for the Edge - Survivability

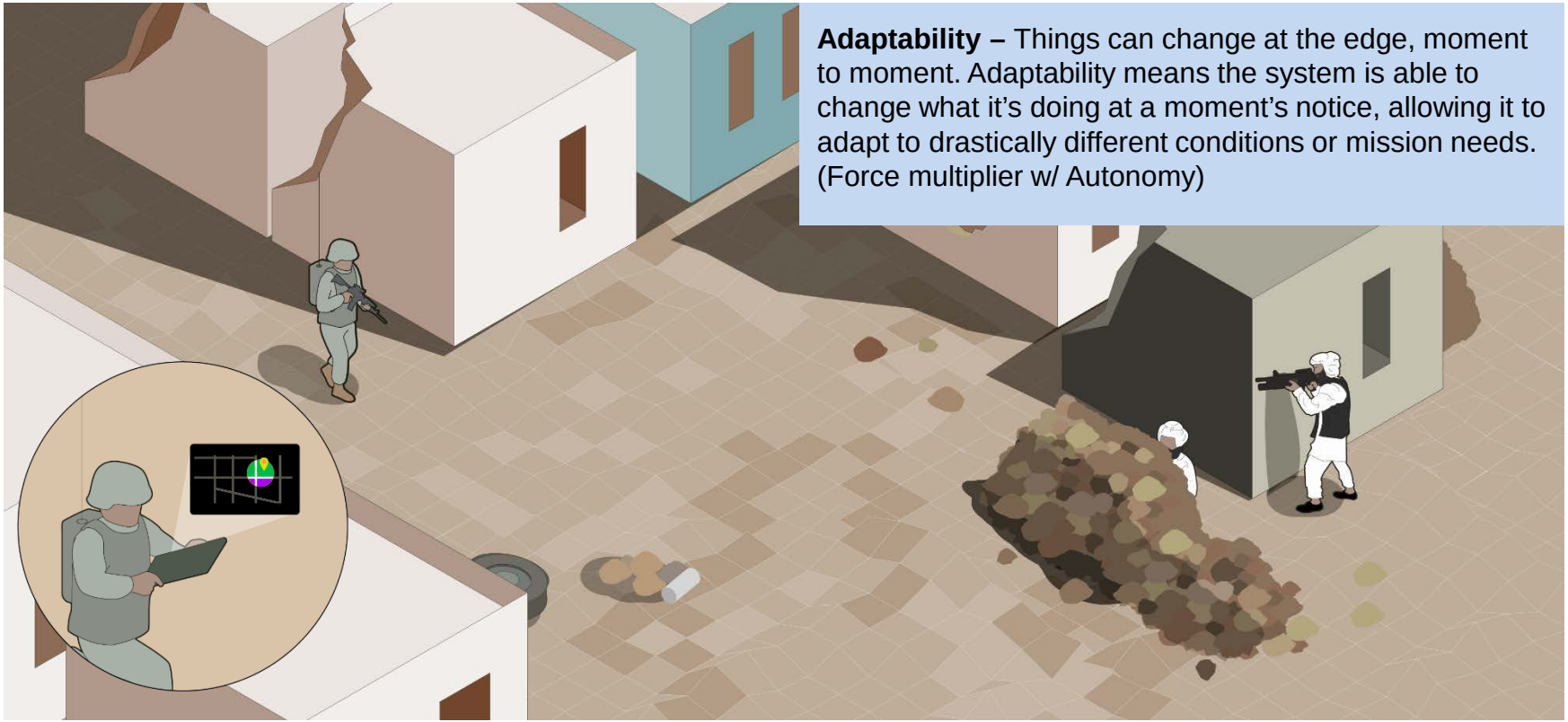


Survivability – Given the edge challenges, a system must be able to survive them. Survivability means that the system will continue to operate despite damage to hardware (compute, power, or networking) or the user.

Quality Attributes for the Edge - Autonomy

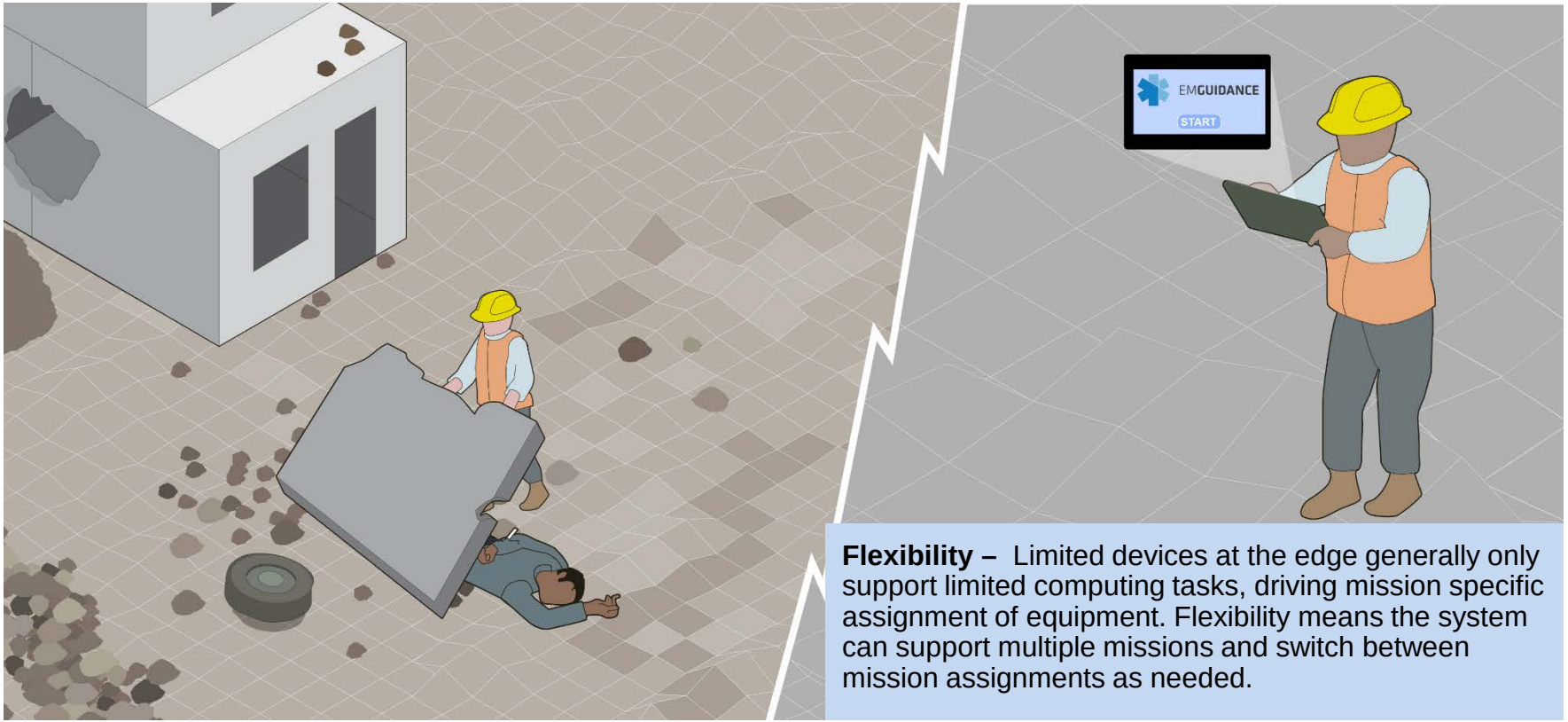
Autonomy – A system at the edge must have some level of intelligence in order to operate while the user is distracted by mission or environmental concerns. Autonomy means that the system can perform (some variable set of) actions on its own without user intervention.

Quality Attributes for the Edge - Adaptability



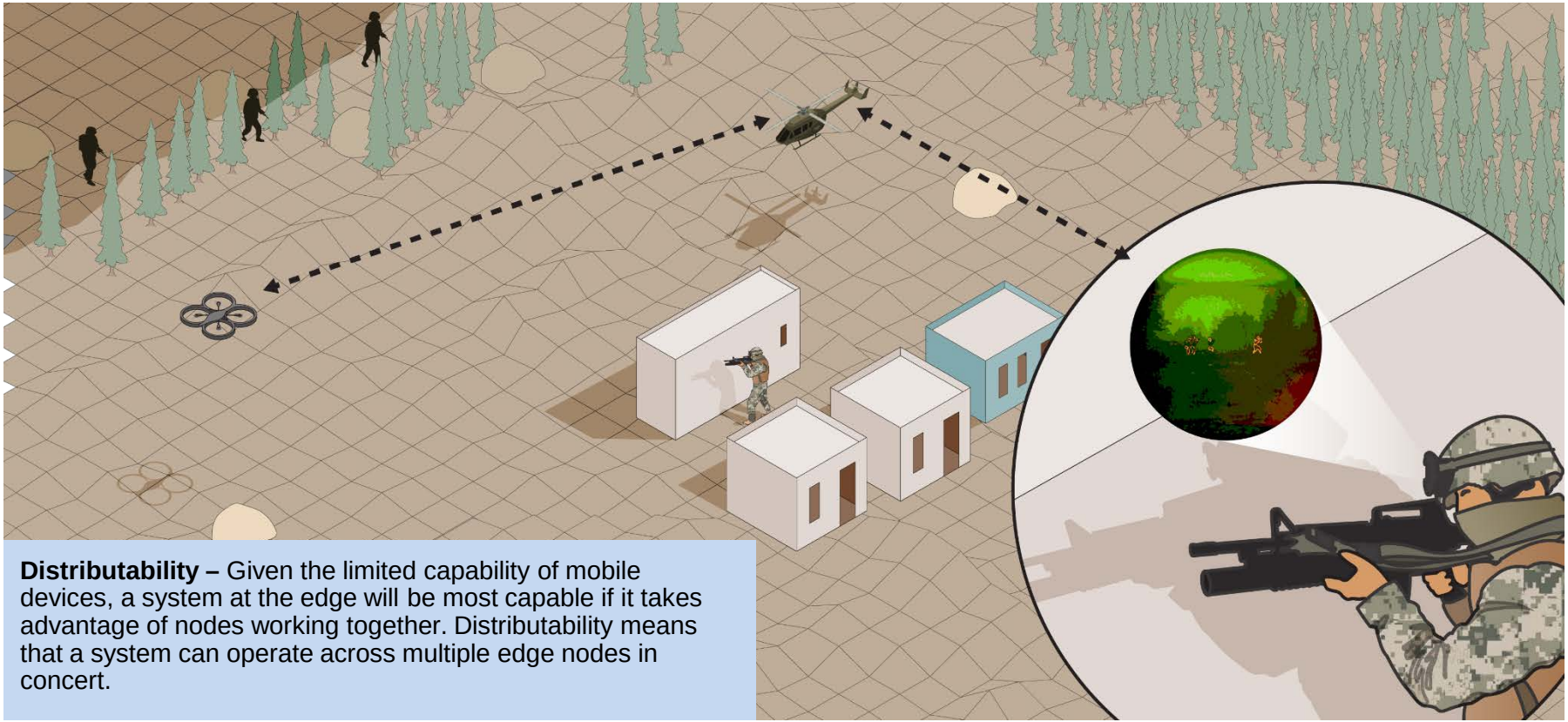
Adaptability – Things can change at the edge, moment to moment. Adaptability means the system is able to change what it's doing at a moment's notice, allowing it to adapt to drastically different conditions or mission needs. (Force multiplier w/ Autonomy)

Quality Attributes for the Edge - Flexibility



Flexibility – Limited devices at the edge generally only support limited computing tasks, driving mission specific assignment of equipment. Flexibility means the system can support multiple missions and switch between mission assignments as needed.

Quality Attributes for the Edge - Distributability

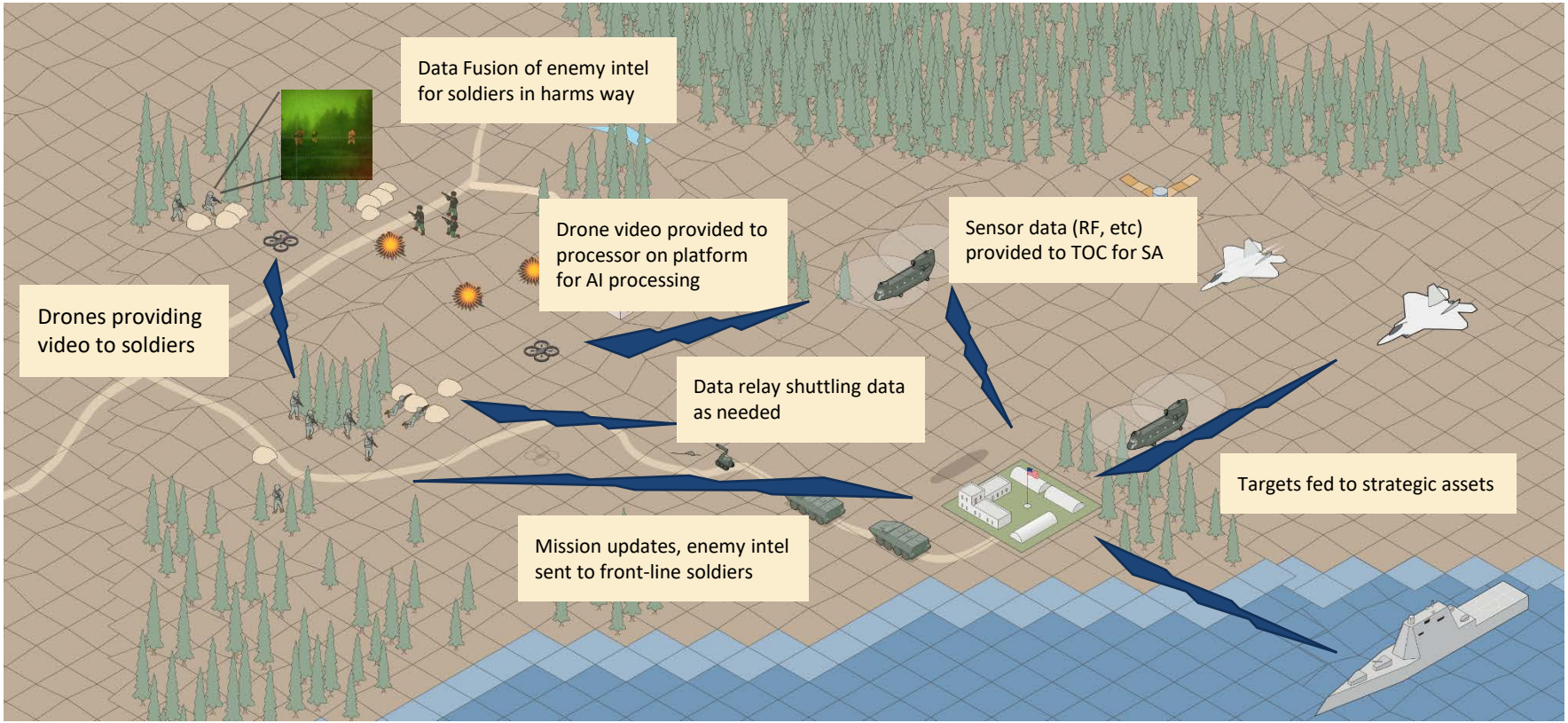


Distributability – Given the limited capability of mobile devices, a system at the edge will be most capable if it takes advantage of nodes working together. Distributability means that a system can operate across multiple edge nodes in concert.

Quality Attributes for the Edge - Openness

Openness – Given the challenges at the edge, and the historical mix of software and devices which are provided to soldiers and first responders at the edge, we cannot expect that one vendor or organization will solve all of these challenges. Openness means that the system must use well-defined, open, and well-supported interfaces and platforms so that any organization with an effective software or hardware solution can easily integrate their contribution with the system.

Nodes and Data Flow at the Tactical Edge



Microservices: Scaling Down at the Edge

Microservices in Edge Systems

Why Microservices?

Microservices allow us to address the Quality Attributes of interest at the edge

- **Reliability & Survivability –** Monitoring, Quick restart, Hot Swap, Statelessness
- **Adaptability & Flexibility –** Statelessness, Small Footprint, Service Orchestration
- **Distributability & Openness –** Most platforms support Docker/TCP/UDP; Small Footprint; Standard Interfaces
- **Autonomy** – Sensor monitoring & adaptation

In short, they are ideal for scaling down and maintaining capability at the Edge

Example Technologies

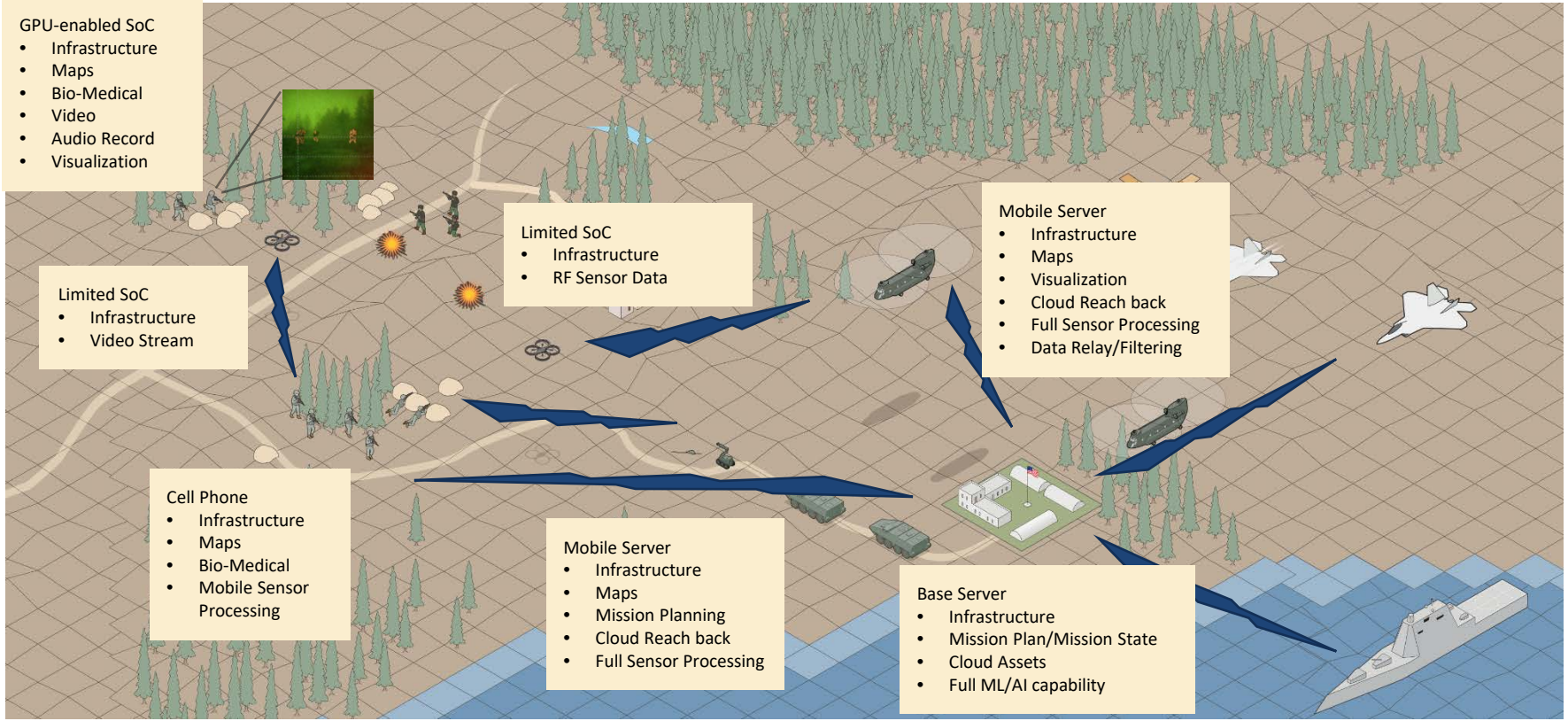
Hardware Devices

- **Comms –** Iridium/Inmarsat; BT/WiFi Android MANET (e.g. Zello); Harris 117G/152a/163; Trellisware TW-400; Silvus Streamcaster 4200
- **Compute –** Mobile Phones/Tablets; Rasberry Pi; Beaglebone; Intel Compute Cards; Portable Embedded Computing Modules (Nvidia Jetson *)
- **Combined –** Persistent Systems MPU5

Software Technologies

- **Hosting –** Hypervisor (VMware, KVM, Xen, etc); Virtual Machines; Containers (Docker)
- **Orchestration/Coordination –** Kubernetes; ROS; DARPA Collaborative Operations in Denied Environments (CODE)
- **Middleware –** ZeroMQ; DDS; ActiveMQ; RabbitMQ
- **Encoding –** XML; CORBA; JSON; Protocol Buffers

Nodes and Microservice Allocation - Tactical Edge



Contact Information

Presenter

- Marc Novakouski
- Senior Member of the Technical Staff
- Email: novakom@sei.cmu.edu

Software Solutions Division/Tactical and AI-Enabled Systems Initiative/Team

- Kevin Pitstick
- Keegan Williams
- Craig Mazzotta
- Jacob Ratzlaff