ENABLING THE TACTICAL CLOUD

Prof. Ken Birman
Dept of Computer Science, Cornell University
http://www.cs.cornell.edu/ken; 607-255-9199

Cloud Computing platforms have the capacity to host staggering amounts of data and to perform massive computations, with impressive cost benefits relative to smaller-scale computing models. Mobile users with powerful handheld devices are eager consumers for a new class of data-rich information applications. This would seem to be an ideal marriage.

There are good reasons to be happy about the development, starting with economics. One recent Microsoft study cited 10x improvements on cost of management, power consumption and cooling and predicts a further 10x in all three dimensions. For example, simply by locating cloud computing systems close to power sources such as hydroelectric dams, the data center can benefit from power without that power needing to be transmitted over long distances, with the inevitable transmission losses. Conversely, transmitting data is relatively cheap. In view of these economies of scale, the cloud looks like the green computing platform of the future. Moreover, application development for the cloud is well supported both by powerful tools and a rich theory and engineering knowledge base (see (1)).

Our interest involves the needs that arise as our tactical military community and other mission-critical computing users shift towards cloud-computing models. Like other mobile users, deployed military units need fingertip access to a diversity of information while operating in the field, but unlike those users, are often operating at the end of networks that were designed as stovepipes to support dedicated applications, such as the plane-to-plane signaling systems used in squadrons of Apache helicopters (see (2),(3)). These networks work well for their intended purpose, but perhaps they could do much more. Can we find ways to leverage the huge deployed base to link the military end-user with cloud computing applications?

There are several issues here. One is the communication model: today's cloud is poorly designed for settings in which the last hop might traverse a tactical military communications link; these often have low bandwidth, can exhibit high latency or frequent disconnects, and may have a rigid channel structure (as in the case of Link 16, which treats each kind of data separately). A second reflects the properties that cloud systems can – and cannot – guarantee. Applications such as medical care, banking or control of the future "smart" electric power grid all bring security, reliability, responsiveness, fault-tolerance or other requirements. Not all need continuous availability, but there are issues associated with the way that the cloud handles disruptions, which are common in modern systems, and can results in inconsistencies such as incorrect product prices or missing images on a web page). Today's cloud is *inconsistent by default*; tomorrow's applications might need stronger (or at least different) properties.

Preliminary study of these questions suggests to us that some of these challenges might be low-hanging fruit, on which focused research could yield quick progress. These include:

- *Limited support for time-critical response.* Existing cloud computing technologies achieve astonishing speed by caching precomputed answers. They also assume the client has a *fast*, reliable connection. These assumptions break down in military and other nationally critical settings: there is a need for disconnected modes of operation, better

security, and ways to guarantee rapid response to fast-changing conditions. For example, we don't really have a way to securely share real-time video feeds in the field, short of shipping the data back to a secure home site and then back out again.

- *Poor availability.* Even if the tactical user could cache substantial content and communicate directly with his counterparts, the cloud expects data to be "bounced off" some sort of home data center. When the quality of the reachback link is poor, all services break down even if many protocols could (at least in principle) adapt themselves to favor secured peer-to-peer modalities or other information management paradigms that operate off potentially stale data – but data that is readily available.

- *Poor reliability and security.* Cloud platforms employ a heavily virtualized 3-tier architecture that leaves it to the client to cope with abrupt service crash/restart events. Front-tier servers are not merely stateless, but the cloud will often force crashes for purposes such as load balancing (migration) or system management. The security model secures the client-to-server path, but within the cloud platform itself protection is weak and the trend towards virtualization prevents us from using TPM security hardware. While we know how to secure groups that share data (like video feeds) the cloud security standards completely ignore such options. Thus the tactical user ends up in an artificially weak security enclave simply because of the inattention from service-computing vendors.

- *Difficulty customizing content, services and the network layer itself.* Modern clouds are stovepipes: one can opt for Google Wave or go with Microsoft Silverlight, but the options don't mix. Yet all modern clouds "embrace inconsistency." Services run on stored, often stale content. But in military, medical, banking and other sensitive applications demand that data be pulled from varied sources, hosted in many places, and information quality can determine mission outcomes.

Our Cornell-based research effort has begun to explore these kinds of questions. For the issues listed above, we see good hope for progress, and these are still early days. Over time, we are confident that it will be possible to create a range of new options: tools both for building cloud solutions with stronger properties, and for improving the connectivity choices when those applications need to support mobile users operating at the end of tactical network links.

FURTHER READING

1. **Reliable Distributed Systems Technologies, Web Services, and Applications**. Birman, Kenneth P. 2005, XXXVI, 668 p. 145 illus., Hardcover ISBN: 0-387-21509-3
2. **Link 16 (**http://en.wikipedia.org/wiki/Link_16) is a military tactical data exchange network used by North Atlantic Treaty Organization (NATO). Link 16 supports the exchange of tactical status reports, text messages, imagery data and offers two channels of digital voice.
3. **Warfighter Information Network-Tactical (WIN-T).** WIN-T is an Army GIG protocol originally intended for use in the Army's Future Combat System (cancelled in May 2009). http://www.globalsecurity.org/military/systems/ground/win-t.htm, http://www.gdc4s.com/content/detail.cfm?item=538ca0ca-4675-4291-84e8-bf047859281f